

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

Sample: b'Electoral_division_of_Apsley'

The Electoral division of Apsley is one of the 15 electorates or seats in the [1 1 Tasmanian Legislative Council 1] . It is the second-largest 1 upper house electorate in the state by area after 2 Murchison . The total area of the Apsley division is 19,204 km ² . The electorate created in 1999 is named after the [1 Apsley River 1] and the . These were named by William Lyne in honour of Lord Apsley , earl of [2 Bathurst 2] . As of October 2010 , there were 22,593 enrolled voters in Apsley . The current member in the is who was elected in 2004 and re-elected unopposed in 2010 . Members serve six-year terms , before facing re-election : the next election in Apsley is due in 2016 . Towns within the division include : 3 Pipers River , 4 Scottsdale , 5 Evandale , [3 6 Swansea 3] , 7 Derby , 8 Lilydale , 9 Bridport , [2 2 Campbell Town 2] , , [4 10 St Helens 4] , 11 Branhholm , 12 Avoca , 13 Fingal , 14 Bicheno , 15 Bagdad , 16 Bellingham , 17 Tomahawk , 18 Ross , 19 St Marys , 20 Rossarden

1: 579457 Tasmanian_Legislative_Council

2: 10003933 Campbell_Town,_Tasmania

1: 579457 Tasmanian_Legislative_Council

[winner: 540100 Upper_gastrointestinal_bleeding,
pem_pos=1, final_scores=-1.1705443,
final_local_scores=-1.0947464, pem=-1.637799,
similarity=-19.673096]

RECALL MISS: 579457 Tasmanian_Legislative_Council]
[winner: 16752663 House,_North_Carolina, pem_pos=16,
final_scores=-1.3283705, final_local_scores=-1.2727723,
pem=-4.8077354, similarity=-7.747141]

RECALL MISS: 579457 Tasmanian_Legislative_Council]

2: 4475584 Electoral_division_of_Murchison

[winner: 674227 Murchison,_New_Zealand, pem_pos=1,
final_scores=-0.111209795, final_local_scores=-0.06323868,
pem=-1.7426276, similarity=11.679225]

gt: 4475584 Electoral_division_of_Murchison, pem_pos=7,
final_scores=-0.39710537, final_local_scores=-0.40774083,
pem=-2.7027931, similarity=6.6201224]

3: 16349526 Pipers_River,_Tasmania

[winner: 16349526 Pipers_River,_Tasmania, pem_pos=1,
final_scores=-0.066337615, final_local_scores=0.020881861,
pem=-0.523729, similarity=7.5609894]

gt: 16349526 Pipers_River,_Tasmania, pem_pos=1,
final_scores=-0.066337615, final_local_scores=0.020881861,
pem=-0.523729, similarity=7.5609894]

4: 782484 Scottsdale,_Tasmania

[winner: 106659 Scottsdale,_Arizona, pem_pos=0,

final_scores=-0.049045265, final_local_scores=0.00091195107,
 pem=-0.2987771, similarity=5.7417235|
 gt: 782484 Scottsdale, Tasmania, pem_pos=2,
 final_scores=-0.2972033, final_local_scores=-0.33085364,
 pem=-1.925832, similarity=4.6893544]
 5: 3469529 Evandale, Tasmania
 [winner: 3469529 Evandale, Tasmania, pem_pos=0,
 final_scores=-0.07984151, final_local_scores=-0.102392696,
 pem=-0.32480422, similarity=2.8004465|
 gt: 3469529 Evandale, Tasmania, pem_pos=0,
 final_scores=-0.07984151, final_local_scores=-0.102392696,
 pem=-0.32480422, similarity=2.8004465]
 6: 508153 Swansea, Tasmania
 [winner: 46394 Swansea, pem_pos=0,
 final_scores=0.12265919, final_local_scores=0.10359898,
 pem=-0.40837145, similarity=9.402101|
 gt: 508153 Swansea, Tasmania, pem_pos=14,
 final_scores=-0.5981176, final_local_scores=-0.634683,
 pem=-3.154351, similarity=2.3033576]
 7: 14188982 Derby, Tasmania
 [winner: 18727355 Derby, pem_pos=0,
 final_scores=0.0299332, final_local_scores=-0.013843305,
 pem=-0.40742427, similarity=5.8922377|
 RECALL MISS: 14188982 Derby, Tasmania]
 8: 3300543 Lilydale, Tasmania
 [winner: 609313 Lilydale, Victoria, pem_pos=1,
 final_scores=-0.20746607, final_local_scores=-0.22764759,
 pem=-0.84514964, similarity=1.8923098|
 gt: 3300543 Lilydale, Tasmania, pem_pos=4,
 final_scores=-0.46234083, final_local_scores=-0.49111137,
 pem=-1.6663253, similarity=-1.5042934]
 9: 2443252 Bridport, Tasmania
 [winner: 3884688 Bridport F.C., pem_pos=3,
 final_scores=-0.22360714, final_local_scores=-0.17431924,
 pem=-1.6752547, similarity=7.9979753|
 gt: 2443252 Bridport, Tasmania, pem_pos=4,
 final_scores=-0.49141073, final_local_scores=-0.4902284,
 pem=-1.7629721, similarity=-0.95236444]
 10: 782443 St_Helens, Tasmania
 [winner: 1095032 St_Helens RLFC, pem_pos=0,
 final_scores=0.04077135, final_local_scores=0.09785506,
 pem=-1.1285291, similarity=13.147019|
 gt: 782443 St_Helens, Tasmania, pem_pos=5,
 final_scores=-0.57173306, final_local_scores=-0.5790283,
 pem=-2.8974087, similarity=2.5669124]
 11: 29254978 Branxholm, Tasmania
 [winner: 29254978 Branxholm, Tasmania, pem_pos=0,
 final_scores=-0.18436198, final_local_scores=-0.16326004,
 pem=-0.12516314, similarity=-0.10163599|
 gt: 29254978 Branxholm, Tasmania, pem_pos=0,
 final_scores=-0.18436198, final_local_scores=-0.16326004,
 pem=-0.12516314, similarity=-0.10163599]

12: 3301380 Avoca, Tasmania
[winner: 1289754 Avoca, County Wicklow, pem_pos=6,
final_scores=-0.4134306, final_local_scores=-0.4161743,
pem=-2.3492887, similarity=4.4460387]
gt: 3301380 Avoca, Tasmania, pem_pos=18,
final_scores=-0.665864, final_local_scores=-0.7076347,
pem=-2.9074965, similarity=-1.2161002]
13: 16495493 Fingal, Tasmania
[winner: 15417965 Sporting Fingal F.C., pem_pos=10,
final_scores=-0.07762349, final_local_scores=-0.004885316,
pem=-2.6322248, similarity=18.258373]
gt: 16495493 Fingal, Tasmania, pem_pos=4,
final_scores=-0.3908563, final_local_scores=-0.39007407,
pem=-2.2657728, similarity=4.7707496]
14: 778053 Bicheno, Tasmania
[winner: 778053 Bicheno, Tasmania, pem_pos=0,
final_scores=-0.34395486, final_local_scores=-0.32948118,
pem=0.0, similarity=-5.7426624]
gt: 778053 Bicheno, Tasmania, pem_pos=0,
final_scores=-0.34395486, final_local_scores=-0.32948118,
pem=0.0, similarity=-5.7426624]
15: 1812992 Bagdad, Tasmania
[winner: 4492 Baghdad, pem_pos=1,
final_scores=-0.24494277, final_local_scores=-0.20361064,
pem=-1.572928, similarity=6.5673923]
gt: 1812992 Bagdad, Tasmania, pem_pos=11,
final_scores=-0.5225417, final_local_scores=-0.51294565,
pem=-2.584407, similarity=2.8367991]
16: 3310407 Bellingham, Tasmania
[winner: 39300 Bellingham, Washington, pem_pos=0,
final_scores=-0.03125246, final_local_scores=-0.042024463,
pem=-0.46201012, similarity=5.348102]
gt: 3310407 Bellingham, Tasmania, pem_pos=6,
final_scores=-0.7031906, final_local_scores=-0.6804484,
pem=-2.9254289, similarity=-0.30728918]
17: 4763872 Tomahawk, Tasmania
[winner: 31238 Tomahawk (missile), pem_pos=0,
final_scores=-0.36062238, final_local_scores=-0.23994622,
pem=-0.9611242, similarity=2.155982]
gt: 4763872 Tomahawk, Tasmania, pem_pos=16,
final_scores=-0.8811913, final_local_scores=-0.8407799,
pem=-3.0813878, similarity=-4.2437696]
18: 9381228 Ross, Tasmania
[winner: 417624 Ross, pem_pos=0,
final_scores=-0.54052776, final_local_scores=-0.56143963,
pem=-1.9870358, similarity=-1.8589559]
RECALL MISS: 9381228 Ross, Tasmania]
19: 8396834 St Marys, Tasmania
[winner: 2579292 St Marys, New South Wales, pem_pos=0,
final_scores=-0.16942394, final_local_scores=-0.19783466,
pem=-0.9705591, similarity=3.4639835]
gt: 8396834 St Marys, Tasmania, pem_pos=1,

final_scores=-0.2510092, final_local_scores=-0.2380027,
 pem=-1.4234394, similarity=4.728122]
 20: 2510597 Rossarden,_Tasmania
 [winner: 2510597 Rossarden,_Tasmania, pem_pos=0,
 final_scores=-0.42362553, final_local_scores=-0.3434185,
 pem=0.0, similarity=-6.1585803]
 gt: 2510597 Rossarden,_Tasmania, pem_pos=0,
 final_scores=-0.42362553, final_local_scores=-0.3434185,
 pem=0.0, similarity=-6.1585803]
 1: 579457 Tasmanian_Legislative_Council
 [pem_pos=0, final_scores=0.24172474,
 final_local_scores=0.26682144, pem=0.0, similarity=12.052197]
 2: 10003933 Campbell_Town,_Tasmania
 [pem_pos=0, final_scores=0.11818378,
 final_local_scores=0.118225634, pem=0.0, similarity=7.6178017]
 1: 19381647 Apsley_River_(New_South_Wales)
 [pem_pos=0, final_scores=0.20385803,
 final_local_scores=0.2886502, pem=0.0, similarity=12.703611]
 2: 323764 Bathurst,_New_South_Wales
 [pem_pos=0, final_scores=0.21090882,
 final_local_scores=0.18765518, pem=-0.7060237,
 similarity=13.529184]
 3: 46394 Swansea
 [pem_pos=0, final_scores=0.12265919,
 final_local_scores=0.10359898, pem=-0.40837145,
 similarity=9.402101]
 4: 1095032 St_Helens_RLFC
 [pem_pos=0, final_scores=0.04077135,
 final_local_scores=0.09785506, pem=-1.1285291,
 similarity=13.147019]

Sample: b'Big_Blue_River'

_ -LRB- [1 Indiana 1] -RRB- The [2 Big Blue River 2] is an 1
 tributary of the 2 Driftwood River in east-central [1 1 India
 na 1] in the [2 2 United States 2] . Via the Driftwood , 3 Wh
 ite , 4 Wabash and [3 5 Ohio 3] rivers , it is part of the 6 w
 atershed of the [3 3 Mississippi River 3] . The Big Blue
 rises in northeastern [4 4 Henry County 4] and flows
 generally southwestwardly through 7 Rush , 8 Hancock , 9 Shel
 by and 10 Johnson counties , past the towns of [4 11 New Cast
 le 4] , [5 5 Knightstown 5] , 12 Carthage , [5 13 Morristown 5
] , [6 14 Shelbyville 6] and [7 15 Edinburgh 7] . It joins 16
 Sugar Creek to form the Driftwood River west of [8 Edinburgh 8
] . At [9 Shelbyville 9] , it collects the

1: 21883857 Indiana
 2: 3434750 United_States
 3: 19579 Mississippi_River
 4: 96114 Henry_County,_Indiana
 5: 112336 Knightstown,_Indiana
 1: 72465 Tributary

[winner: 72465 Tributary, pem_pos=1,
 final_scores=-1.1491978, final_local_scores=-1.1569805,

pem=-0.012074053, similarity=-30.371233|
 gt: 72465 Tributary, pem_pos=1, final_scores=-1.1491978,
 final_local_scores=-1.1569805, pem=-0.012074053,
 similarity=-30.371233]

2: 4089879 Driftwood_River
 [winner: 4089879 Driftwood_River, pem_pos=0,
 final_scores=-0.031500474, final_local_scores=0.04450488,
 pem=0.0, similarity=5.417827|
 gt: 4089879 Driftwood_River, pem_pos=0,
 final_scores=-0.031500474, final_local_scores=0.04450488,
 pem=0.0, similarity=5.417827]

3: 427269 White_River_(Indiana)
 [winner: 273285
 Race_and_ethnicity_in_the_United_States_Census, pem_pos=0,
 final_scores=-0.24935229, final_local_scores=-0.18806764,
 pem=-0.2843238, similarity=0.023596399|
 RECALL MISS: 427269 White_River_(Indiana)]

4: 427204 Wabash_River
 [winner: 112661 Wabash,_Indiana, pem_pos=1,
 final_scores=-0.033164963, final_local_scores=-0.07950798,
 pem=-1.5962621, similarity=10.397758|
 gt: 427204 Wabash_River, pem_pos=0,
 final_scores=-0.058676675, final_local_scores=-0.10298222,
 pem=-1.492543, similarity=9.133199]

5: 22388 Ohio_River
 [winner: 22199 Ohio, pem_pos=0, final_scores=0.40550464,
 final_local_scores=0.35704744, pem=-0.19147073,
 similarity=15.785969|
 gt: 22388 Ohio_River, pem_pos=1,
 final_scores=-0.30191267, final_local_scores=-0.38723797,
 pem=-3.2031271, similarity=9.952863]

6: 191162 Drainage_basin
 [winner: 2722676 European_Watershed, pem_pos=2,
 final_scores=-1.4463984, final_local_scores=-1.4543307,
 pem=-1.38281, similarity=-31.79048|
 gt: 191162 Drainage_basin, pem_pos=0,
 final_scores=-1.4501384, final_local_scores=-1.5444269,
 pem=0.0, similarity=-41.99907]

7: 96073 Rush_County,_Indiana
 [winner: 25432 Rush_(band), pem_pos=0,
 final_scores=-0.1818763, final_local_scores=-0.1181147,
 pem=-0.49376196, similarity=3.2500887|
 gt: 96073 Rush_County,_Indiana, pem_pos=9,
 final_scores=-0.40739602, final_local_scores=-0.49332193,
 pem=-3.9413664, similarity=10.801765]

8: 96118 Hancock_County,_Indiana
 [winner: 65693 Hancock,_Michigan, pem_pos=14,
 final_scores=-0.26055586, final_local_scores=-0.30173835,
 pem=-3.4741447, similarity=13.978178|
 gt: 96118 Hancock_County,_Indiana, pem_pos=5,
 final_scores=-0.29950368, final_local_scores=-0.37336594,
 pem=-3.128637, similarity=9.961744]

9: 96071 Shelby County, Indiana
[winner: 127818 Shelby, North Carolina, pem_pos=0,
final_scores=-0.056971952, final_local_scores=-0.10087523,
pem=-1.9416804, similarity=11.638556]
gt: 96071 Shelby County, Indiana, pem_pos=4,
final_scores=-0.31073964, final_local_scores=-0.3718049,
pem=-2.6451175, similarity=7.3788733]

10: 96106 Johnson County, Indiana
[winner: 54533 Lyndon B. Johnson, pem_pos=0,
final_scores=-0.20688462, final_local_scores=-0.21394762,
pem=-2.2601202, similarity=9.995974]
gt: 96106 Johnson County, Indiana, pem_pos=4,
final_scores=-0.49830705, final_local_scores=-0.5889494,
pem=-3.40783, similarity=5.046598]

11: 112341 New Castle, Indiana
[winner: 108923 New Castle, Delaware, pem_pos=0,
final_scores=0.20221587, final_local_scores=0.1914981,
pem=-1.385504, similarity=17.338984]
gt: 112341 New Castle, Indiana, pem_pos=3,
final_scores=0.101254925, final_local_scores=0.07257444,
pem=-1.7905742, similarity=15.9929]

12: 112584 Carthage, Indiana
[winner: 6555 Carthage, pem_pos=0,
final_scores=0.029901966, final_local_scores=0.10429013,
pem=-0.2933236, similarity=8.797078]
gt: 112584 Carthage, Indiana, pem_pos=14,
final_scores=-0.65400696, final_local_scores=-0.67225975,
pem=-3.0879302, similarity=0.82078487]

13: 112589 Morristown, Indiana
[winner: 125569 Morristown, New Jersey, pem_pos=0,
final_scores=0.27478713, final_local_scores=0.25551704,
pem=-0.6796117, similarity=15.410685]
gt: 112589 Morristown, Indiana, pem_pos=3,
final_scores=-0.36128187, final_local_scores=-0.34857172,
pem=-2.014812, similarity=4.644499]

14: 112590 Shelbyville, Indiana
[winner: 115426 Shelbyville, Kentucky, pem_pos=0,
final_scores=0.1701823, final_local_scores=0.13398993,
pem=-0.96834135, similarity=13.354231]
gt: 112590 Shelbyville, Indiana, pem_pos=1,
final_scores=0.12986849, final_local_scores=0.09056698,
pem=-1.0989844, similarity=12.76886]

15: 112120 Edinburg, Indiana
[winner: 9602 Edinburg, pem_pos=0,
final_scores=0.22536997, final_local_scores=0.22887069,
pem=-0.11916062, similarity=11.567685]
gt: 112120 Edinburg, Indiana, pem_pos=3,
final_scores=-0.64540046, final_local_scores=-0.6843824,
pem=-3.540854, similarity=2.922093]

16: 4090378 Sugar Creek (Driftwood River)
[winner: 150737 Sugar Creek, Missouri, pem_pos=0,
final_scores=-0.085232124, final_local_scores=-0.10099223,

pem=-1.6490352, similarity=10.043613|
gt: 4090378 Sugar_Creek_(Driftwood_River), pem_pos=15,
final_scores=-0.5615634, final_local_scores=-0.5454602,
pem=-2.660092, similarity=2.278086]
1: 21883857 Indiana
[pem_pos=0, final_scores=0.5028475,
final_local_scores=0.4691217, pem=-0.27230176,
similarity=19.570059]
2: 3434750 United_States
[pem_pos=0, final_scores=0.48562002,
final_local_scores=0.4924919, pem=-0.073457584,
similarity=19.186129]
3: 19579 Mississippi_River
[pem_pos=0, final_scores=0.42003685,
final_local_scores=0.4234764, pem=0.0, similarity=16.727093]
4: 96114 Henry_County,_Indiana
[pem_pos=1, final_scores=0.13328458,
final_local_scores=0.097486466, pem=-1.5978819,
similarity=15.688434]
5: 112336 Knightstown,_Indiana
[pem_pos=0, final_scores=0.17777391,
final_local_scores=0.16819975, pem=0.0, similarity=9.109129]
1: 21883857 Indiana
[pem_pos=0, final_scores=0.27217895,
final_local_scores=0.21715134, pem=-0.27230176,
similarity=12.050761]
2: 1461489 Big_Blue_River_(Kansas)
[pem_pos=0, final_scores=0.07688622,
final_local_scores=0.13290575, pem=-0.061755188,
similarity=8.39172]
3: 22199 Ohio
[pem_pos=0, final_scores=0.40550464,
final_local_scores=0.35704744, pem=-0.19147073,
similarity=15.785969]
4: 108923 New_Castle,_Delaware
[pem_pos=0, final_scores=0.20221587,
final_local_scores=0.1914981, pem=-1.385504,
similarity=17.338984]
5: 125569 Morristown,_New_Jersey
[pem_pos=0, final_scores=0.27478713,
final_local_scores=0.25551704, pem=-0.6796117,
similarity=15.410685]
6: 115426 Shelbyville,_Kentucky
[pem_pos=0, final_scores=0.1701823,
final_local_scores=0.13398993, pem=-0.96834135,
similarity=13.354231]
7: 9602 Edinburgh
[pem_pos=0, final_scores=0.22536997,
final_local_scores=0.22887069, pem=-0.11916062,
similarity=11.567685]
8: 9602 Edinburgh
[pem_pos=0, final_scores=0.3963024,

final_local_scores=0.41558838, pem=-0.11916062,
similarity=17.139711]
9: 115426 Shelbyville, Kentucky
[pem_pos=0, final_scores=0.119899616,
final_local_scores=0.07906376, pem=-0.96834135,
similarity=11.715124]

Sample: b'Markus_Prock'

[1 Markus Prock 1] -LRB- born 22 June 1964 -RRB- is an [1 1 A
ustria 1] n 1 luge r who competed between 1983 and 2002 .
Born in [2 2 Innsbruck 2] , [2 Prock 2] competed in six
winning three medals in the men 's singles event with two
silvers -LRB- 2 1992 , 3 1994 -RRB- and one bronze -LRB- 4 200
2 -RRB- . At the [3 3 FIL World Luge Championships 3] , [3
Prock 3] won 13 medals , including five gold -LRB- Men 's
singles : 5 1987 , 6 1996 ; Mixed team : 1996 , 7 1997 , 8 199
9 -RRB- , four silvers -LRB- Men 's singles : 9 1990 , 1997 ;
Mixed team : 10 1991 , 11 1993 -RRB- , and four bronzes -LRB-
Men 's singles : 12 1995 , 13 2001 ; Mixed team : 1995 , 14 20
01 -RRB- . [4 Prock 4] won six medals at the [4 4 FIL European
Luge Championships 4] with five in men 's singles -LRB-
gold : 15 1994 , 16 1998 , 17 2002 ; silver : 18 1988 , 19 199
0 -RRB- and one medal in the mixed team event -LRB- bronze :
2002 -RRB- . He also won the overall ten times in men 's
singles -LRB- 1987-8 , 1990-1 , 1991-2 , 1992-3 , 1993-4 ,
1994-5 , 1995-6 , 1996-7 , 1998-9 , 2001-2 -RRB- . As of
2007 , he was the sports director of the [5 Tyrolean 5] luge
association . He also manages his nephew , [5 5 Gregor Schlie
renzauer 5] , who was second in the 2006-07 entering the [6 20
FIS Nordic World Ski Championships 6] 2007 in [6 6 Sapporo 6]
 , [7 7 Japan 7]

1: 26964606 Austria

2: 55857 Innsbruck

3: 5127885 FIL_World_Luge_Championships

4: 6569633 FIL_European_Luge_Championships

5: 8234102 Gregor_Schlierenzauer

6: 6792879 Sapporo

7: 15573 Japan

1: 37779 Luge

[winner: 37779 Luge, pem_pos=0, final_scores=-0.8989632,
final_local_scores=-1.0268521, pem=-0.43050212,
similarity=-24.212465]

gt: 37779 Luge, pem_pos=0, final_scores=-0.8989632,
final_local_scores=-1.0268521, pem=-0.43050212,
similarity=-24.212465]

2: 113357 1992_Winter_Olympics

[winner: 34669 1992, pem_pos=0, final_scores=-0.6203924,
final_local_scores=-0.47061217, pem=-0.76294684,
similarity=-5.805269]

gt: 113357 1992_Winter_Olympics, pem_pos=14,
final_scores=-0.958559, final_local_scores=-1.0775259,
pem=-4.6746583, similarity=-2.6442888]

3: 113198 1994_Winter_Olympics
[winner: 18948337 1994, pem_pos=0,
final_scores=-0.8293264, final_local_scores=-0.74862003,
pem=-0.7539553, similarity=-14.1504755]
gt: 113198 1994_Winter_Olympics, pem_pos=19,
final_scores=-1.5140008, final_local_scores=-1.7043984,
pem=-4.842105, similarity=-20.440815]
4: 38678 2002_Winter_Olympics
[winner: 35502 2002, pem_pos=0, final_scores=-0.82790375,
final_local_scores=-0.7135766, pem=-0.5897454,
similarity=-13.997709]
gt: 38678 2002_Winter_Olympics, pem_pos=16,
final_scores=-1.4843979, final_local_scores=-1.6589355,
pem=-4.980688, similarity=-18.330475]
5: 11171723 FIL_World_Luge_Championships_1987
[winner: 34760 1987, pem_pos=0, final_scores=-0.5372608,
final_local_scores=-0.3607607, pem=-0.72282183,
similarity=-2.7452881]
RECALL MISS: 11171723 FIL_World_Luge_Championships_1987]
6: 11209322 FIL_World_Luge_Championships_1996
[winner: 34636 1996, pem_pos=0, final_scores=-0.6591366,
final_local_scores=-0.4732498, pem=-0.78063476,
similarity=-5.787792]
RECALL MISS: 11209322 FIL_World_Luge_Championships_1996]
7: 11209429 FIL_World_Luge_Championships_1997
[winner: 34601 1997, pem_pos=0, final_scores=-0.8242394,
final_local_scores=-0.7464737, pem=-0.69383013,
similarity=-14.413397]
RECALL MISS: 11209429 FIL_World_Luge_Championships_1997]
8: 11209548 FIL_World_Luge_Championships_1999
[winner: 34553 1999, pem_pos=0, final_scores=-1.4222561,
final_local_scores=-1.4485366, pem=-0.6240096,
similarity=-35.74405]
RECALL MISS: 11209548 FIL_World_Luge_Championships_1999]
9: 11171979 FIL_World_Luge_Championships_1990
[winner: 34635 1990, pem_pos=0, final_scores=-0.7624348,
final_local_scores=-0.6615462, pem=-0.79115707,
similarity=-11.349709]
RECALL MISS: 11171979 FIL_World_Luge_Championships_1990]
10: 11172117 FIL_World_Luge_Championships_1991
[winner: 34703 1991, pem_pos=0, final_scores=-0.9398814,
final_local_scores=-0.81783813, pem=-0.6661314,
similarity=-16.693682]
RECALL MISS: 11172117 FIL_World_Luge_Championships_1991]
11: 11194033 FIL_World_Luge_Championships_1993
[winner: 34598 1993, pem_pos=0, final_scores=-0.71352786,
final_local_scores=-0.5193826, pem=-0.7063994,
similarity=-7.56819]
RECALL MISS: 11194033 FIL_World_Luge_Championships_1993]
12: 11209174 FIL_World_Luge_Championships_1995
[winner: 34658 1995, pem_pos=0, final_scores=-1.1709796,
final_local_scores=-1.1349043, pem=-0.68201965,

similarity=-26.069164|
RECALL MISS: 11209174 FIL_World_Luge_Championships_1995]
13: 11209911 FIL_World_Luge_Championships_2001
[winner: 34551 2001, pem_pos=0, final_scores=-0.61097807,
final_local_scores=-0.3999558, pem=-0.5613713,
similarity=-4.792942|
RECALL MISS: 11209911 FIL_World_Luge_Championships_2001]
14: 11209911 FIL_World_Luge_Championships_2001
[winner: 34551 2001, pem_pos=0, final_scores=-0.58131915,
final_local_scores=-0.367558, pem=-0.5613713,
similarity=-3.8261259|
RECALL MISS: 11209911 FIL_World_Luge_Championships_2001]
15: 11255597 FIL_European_Luge_Championships_1994
[winner: 18948337 1994, pem_pos=0,
final_scores=-0.7547229, final_local_scores=-0.6671271,
pem=-0.7539553, similarity=-11.718564|
RECALL MISS: 11255597
FIL_European_Luge_Championships_1994]
16: 11255860 FIL_European_Luge_Championships_1998
[winner: 34647 1998, pem_pos=0, final_scores=-1.0400767,
final_local_scores=-1.0216889, pem=-0.6850515,
similarity=-22.674107|
RECALL MISS: 11255860
FIL_European_Luge_Championships_1998]
17: 11272218 FIL_European_Luge_Championships_2002
[winner: 35502 2002, pem_pos=0, final_scores=-0.64015585,
final_local_scores=-0.50849056, pem=-0.5897454,
similarity=-7.8775334|
RECALL MISS: 11272218
FIL_European_Luge_Championships_2002]
18: 11255296 FIL_European_Luge_Championships_1988
[winner: 34670 1988, pem_pos=0, final_scores=-0.6985559,
final_local_scores=-0.58156097, pem=-0.7886237,
similarity=-8.976568|
RECALL MISS: 11255296
FIL_European_Luge_Championships_1988]
19: 11255372 FIL_European_Luge_Championships_1990
[winner: 34635 1990, pem_pos=0, final_scores=-0.8868613,
final_local_scores=-0.7974633, pem=-0.79115707,
similarity=-15.405747|
RECALL MISS: 11255372
FIL_European_Luge_Championships_1990]
20: 6438947 FIS_Nordic_World_Ski_Championships_2007
[winner: 1280856 FIS_Nordic_World_Ski_Championships,
pem_pos=0, final_scores=0.20225985,
final_local_scores=0.21282834, pem=0.0, similarity=10.440935|
gt: 6438947 FIS_Nordic_World_Ski_Championships_2007,
pem_pos=5, final_scores=-0.77057207,
final_local_scores=-0.84907454, pem=-5.371155,
similarity=7.960809]
[winner: 36165 2007, pem_pos=0, final_scores=-0.60281104,
final_local_scores=-0.48836097, pem=-0.24805899,

```

similarity=-9.13497|
  RECALL MISS: 6438947
FIS_Nordic_World_Ski_Championships_2007]
1: 26964606 Austria
  [pem_pos=0, final_scores=0.6297452,
final_local_scores=0.6675045, pem=-0.25262418,
similarity=25.383188]
2: 55857 Innsbruck
  [pem_pos=0, final_scores=0.5436453,
final_local_scores=0.5374139, pem=0.0, similarity=20.127216]
3: 5127885 FIL_World_Luge_Championships
  [pem_pos=0, final_scores=0.24393772,
final_local_scores=0.24401543, pem=0.0, similarity=11.37162]
4: 6569633 FIL_European_Luge_Championships
  [pem_pos=0, final_scores=0.20725863,
final_local_scores=0.22332045, pem=0.0, similarity=10.754041]
5: 8234102 Gregor_Schlierenzauer
  [pem_pos=0, final_scores=0.50391334,
final_local_scores=0.54292315, pem=0.0, similarity=20.291622]
6: 6792879 Sapporo
  [pem_pos=0, final_scores=0.29171926,
final_local_scores=0.2861371, pem=-0.028839061,
similarity=12.785446]
7: 15573 Japan
  [pem_pos=0, final_scores=0.7049734,
final_local_scores=0.8186067, pem=-0.13549899,
similarity=29.255432]
1: 9382308 Markus_Prock
  [pem_pos=0, final_scores=0.22392617,
final_local_scores=0.19778451, pem=0.0, similarity=9.991998]
2: 9382308 Markus_Prock
  [pem_pos=0, final_scores=0.5242255,
final_local_scores=0.5258159, pem=0.0, similarity=19.781109]
3: 9382308 Markus_Prock
  [pem_pos=0, final_scores=0.47881776,
final_local_scores=0.47621483, pem=0.0, similarity=18.300913]
4: 9382308 Markus_Prock
  [pem_pos=0, final_scores=0.582249,
final_local_scores=0.58919775, pem=0.0, similarity=21.672548]
5: 148994 Tyrol_(state)
  [pem_pos=1, final_scores=0.30094463,
final_local_scores=0.35360563, pem=-1.0005876,
similarity=20.08336]
6: 1280856 FIS_Nordic_World_Ski_Championships
  [pem_pos=0, final_scores=0.20225985,
final_local_scores=0.21282834, pem=0.0, similarity=10.440935]

```

Sample: b'Noel_Purcell'

_ -LRB- water_polo -RRB- [1 Noel Mary Purcell 1] -LRB- 25
December 1899 or 14 November 1891 -- 31 January 1962 -RRB-
was an [2 Irish 2] rugby union and water polo player . He
represented both [3 1 Great Britain 3] and [4 2 Ireland 4] as

a 3 water polo player at the 4 1920 and 5 1924 [5 Summer Olympics 5] respectively . He also played 6 rugby union for [1 1 Belvedere College 1] , [2 2 Clongowes Wood College 2] and [6 7 Ireland 6] . He is possibly the first athlete to have represented two different nations at the [7 Olympics 7] At the 8 1920 Summer Olympics Purcell , together with [3 3 Charles Sydney Smith 3] , , [4 4 Charles Bugbee 4] , , and [5 5 William Henry Dean 5] , was a member of the [8 Great Britain 8] water polo team that won the gold medal . Then in 9 1924 he captained [9 10 Ireland 9] . In the 11 1921 [10 Five Nations Championship 10] , [11 Purcell 11] played in all four games for [6 6 Ireland 6] . He made his debut against [12 12 England 12] on 12 February . He then played against [7 7 Scotland 7] on 26 February and 13 Wales on 12 March . He made his final appearance for Ireland against [13 14 France 13]

1: 977628 Belvedere_College

2: 969176 Clongowes_Wood_College

3: 10772467 Charles_Sydney_Smith

4: 10772644 Charles_Bugbee

5: 10772794 William_Henry_Dean

6: 955500 Ireland_national_rugby_union_team

7: 758049 Scotland_national_rugby_union_team

1: 6084493 Great_Britain_at_the_1920_Summer_Olympics

[winner: 13530298 Great_Britain, pem_pos=0,
final_scores=0.16351293, final_local_scores=0.12495175,
pem=-0.50213265, similarity=10.549198]

RECALL MISS: 6084493

Great_Britain_at_the_1920_Summer_Olympics]

2: 7098311 Ireland_at_the_1924_Summer_Olympics

[winner: 14560 Republic_of_Ireland, pem_pos=1,
final_scores=0.15973125, final_local_scores=0.1308653,
pem=-1.5040284, similarity=16.174135]

RECALL MISS: 7098311 Ireland_at_the_1924_Summer_Olympics]

3: 47426 Water_polo

[winner: 47426 Water_polo, pem_pos=0,
final_scores=-0.69153786, final_local_scores=-0.76076347,
pem=-0.031322464, similarity=-18.44265]

gt: 47426 Water_polo, pem_pos=0,

final_scores=-0.69153786, final_local_scores=-0.76076347,
pem=-0.031322464, similarity=-18.44265]

4: 8473140 Water_polo_at_the_1920_Summer_Olympics

[winner: 34676 1920, pem_pos=0, final_scores=-0.69331306,
final_local_scores=-0.5626169, pem=-0.8917648,
similarity=-7.8503413]

RECALL MISS: 8473140

Water_polo_at_the_1920_Summer_Olympics]

5: 8472920 Water_polo_at_the_1924_Summer_Olympics

[winner: 34675 1924, pem_pos=0, final_scores=-0.80541784,
final_local_scores=-0.7149822, pem=-0.80803895,
similarity=-12.852541]

RECALL MISS: 8472920

Water_polo_at_the_1924_Summer_Olympics]

6: 25405 Rugby_union
 [winner: 19774201 Rugby_Virginia, pem_pos=7,
 final_scores=-0.8386906, final_local_scores=-0.71619165,
 pem=-2.956099, similarity=-1.2071489]
 gt: 25405 Rugby_union, pem_pos=3,
 final_scores=-1.7924142, final_local_scores=-1.9811466,
 pem=-2.0110612, similarity=-44.09519]
 [winner: 481130 Union_(American_Civil_War), pem_pos=0,
 final_scores=-1.3410168, final_local_scores=-1.3958554,
 pem=-1.132797, similarity=-31.40507]
 RECALL MISS: 25405 Rugby_union]

7: 955500 Ireland_national_rugby_union_team
 [winner: 14560 Republic_of_Ireland, pem_pos=1,
 final_scores=0.22496231, final_local_scores=0.19680879,
 pem=-1.5040284, similarity=18.14202]
 gt: 955500 Ireland_national_rugby_union_team, pem_pos=2,
 final_scores=-0.28640145, final_local_scores=-0.36032152,
 pem=-3.3784852, similarity=11.709727]

8: 8473140 Water_polo_at_the_1920_Summer_Olympics
 [winner: 34676 1920, pem_pos=0, final_scores=-1.0365286,
 final_local_scores=-0.9375277, pem=-0.8917648,
 similarity=-19.038427]
 RECALL MISS: 8473140
 Water_polo_at_the_1920_Summer_Olympics]
 [winner: 27566 Summer_Olympic_Games, pem_pos=0,
 final_scores=-1.0970486, final_local_scores=-1.2261238,
 pem=-0.4112264, similarity=-30.263958]
 RECALL MISS: 8473140
 Water_polo_at_the_1920_Summer_Olympics]

9: 8472920 Water_polo_at_the_1924_Summer_Olympics
 [winner: 34675 1924, pem_pos=0, final_scores=-0.5425455,
 final_local_scores=-0.42783406, pem=-0.80803895,
 similarity=-4.2834673]
 RECALL MISS: 8472920
 Water_polo_at_the_1924_Summer_Olympics]

10: 22421332 Ireland_men's_national_water_polo_team
 [winner: 14560 Republic_of_Ireland, pem_pos=1,
 final_scores=0.04663749, final_local_scores=0.002016008,
 pem=-1.5040284, similarity=12.329015]
 RECALL MISS: 22421332
 Ireland_men's_national_water_polo_team]

11: 16275605 1921_Five_Nations_Championship
 [winner: 26547 Six_Nations_Championship, pem_pos=0,
 final_scores=0.15547977, final_local_scores=0.14367428,
 pem=0.0, similarity=8.37724]
 RECALL MISS: 16275605 1921_Five_Nations_Championship]
 [winner: 34724 1921, pem_pos=0, final_scores=-0.84224695,
 final_local_scores=-0.7319295, pem=-0.7163948,
 similarity=-13.856656]
 RECALL MISS: 16275605 1921_Five_Nations_Championship]

12: 691024 England_national_rugby_union_team
 [winner: 9904 England_national_football_team, pem_pos=1,

final_scores=0.03632067, final_local_scores=0.01875937,
pem=-2.776464, similarity=19.748373]
gt: 691024 England_national_rugby_union_team, pem_pos=3,
final_scores=-0.0992506, final_local_scores=-0.1464557,
pem=-3.7874734, similarity=20.316051]
13: 770930 Wales_national_rugby_union_team
[winner: 580291 Wales_national_football_team, pem_pos=1,
final_scores=-0.053981796, final_local_scores=-0.08746129,
pem=-2.4526134, similarity=14.817386]
gt: 770930 Wales_national_rugby_union_team, pem_pos=2,
final_scores=-0.086380765, final_local_scores=-0.12157509,
pem=-2.723462, similarity=15.272281]
14: 955656 France_national_rugby_union_team
[winner: 5843419 France, pem_pos=0,
final_scores=0.10934396, final_local_scores=0.09267387,
pem=-0.16240834, similarity=7.738488]
gt: 955656 France_national_rugby_union_team, pem_pos=2,
final_scores=-0.33068097, final_local_scores=-0.36829287,
pem=-4.2920833, similarity=16.440136]
1: 977628 Belvedere_College
[pem_pos=0, final_scores=0.037804767,
final_local_scores=0.046901003, pem=0.0, similarity=5.489332]
2: 969176 Clongowes_Wood_College
[pem_pos=0, final_scores=0.20974086,
final_local_scores=0.23841175, pem=0.0, similarity=11.204395]
3: 10772467 Charles_Sydney_Smith
[pem_pos=0, final_scores=0.034316465,
final_local_scores=0.11182223, pem=0.0, similarity=7.4267116]
4: 10772644 Charles_Bugbee
[pem_pos=0, final_scores=0.36854136,
final_local_scores=0.49098462, pem=0.0, similarity=18.741673]
5: 10772794 William_Henry_Dean
[pem_pos=0, final_scores=0.13175277,
final_local_scores=0.22841245, pem=0.0, similarity=10.905996]
6: 955500 Ireland_national_rugby_union_team
[pem_pos=2, final_scores=0.048198357,
final_local_scores=0.017189085, pem=-3.3784852,
similarity=22.975397]
7: 758049 Scotland_national_rugby_union_team
[pem_pos=2, final_scores=0.03670208,
final_local_scores=0.01891601, pem=-3.3931751,
similarity=23.106817]
1: 9738060 Noel_Purcell_(water_polo)
[pem_pos=0, final_scores=0.19256006,
final_local_scores=0.22371635, pem=0.0, similarity=10.765855]
2: 14560 Republic_of_Ireland
[pem_pos=3, final_scores=0.49463385,
final_local_scores=0.49138397, pem=-2.2576945,
similarity=31.031284]
3: 13530298 Great_Britain
[pem_pos=0, final_scores=0.16351293,
final_local_scores=0.12495175, pem=-0.50213265,

```

similarity=10.549198]
4: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.15973125,
    final_local_scores=0.1308653, pem=-1.5040284,
    similarity=16.174135]
5: 27566 Summer_Olympic_Games
    [pem_pos=0, final_scores=0.041913554,
    final_local_scores=0.030169562, pem=-0.4112264,
    similarity=7.226346]
6: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.22496231,
    final_local_scores=0.19680879, pem=-1.5040284,
    similarity=18.14202]
7: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.08425404,
    final_local_scores=0.07642013, pem=-2.7333555,
    similarity=21.234655]
8: 13530298 Great_Britain
    [pem_pos=0, final_scores=0.16093926,
    final_local_scores=0.12214038, pem=-0.50213265,
    similarity=10.465301]
9: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.04663749,
    final_local_scores=0.002016008, pem=-1.5040284,
    similarity=12.329015]
10: 26547 Six_Nations_Championship
    [pem_pos=0, final_scores=0.15547977,
    final_local_scores=0.14367428, pem=0.0, similarity=8.37724]
11: 9738060 Noel_Purcell_(water_polo)
    [pem_pos=0, final_scores=0.20062838,
    final_local_scores=0.23252976, pem=0.0, similarity=11.028865]
12: 9904 England_national_football_team
    [pem_pos=1, final_scores=0.03632067,
    final_local_scores=0.01875937, pem=-2.776464,
    similarity=19.748373]
13: 5843419 France
    [pem_pos=0, final_scores=0.10934396,
    final_local_scores=0.09267387, pem=-0.16240834,
    similarity=7.738488]

```

Sample: b'Alabama_State_Route_13'

State Route 13 -LRB- SR-13 -RRB- is a 1 state highway in the western part of the [1 2 U.S. 1] state of [1 1 Alabama 1] . Except for a portion roughly between and , SR-13 is the unsigned designation for . Thus , while the total distance of the route is over , as an independently signed route , SR-13 is only long . The southern terminus of SR-13 and US-43 is at their junction with and unsigned in [2 Mobile 2] . The northern terminus of the route is on US-43 at the [2 2 Tennes see 2] state line north of 3 Killen in [3 4 Lauderdale County 3] . As a signed route , the southern terminus of the route is at the junction of US-43 and in southern [4 5 Fayette Coun

ty 4] , and the northern terminus as at US-43 and south of [5
6 Russellville 5] in [6 7 Franklin County 6] . East of ,
SR-13 diverts from US-43 and becomes a signed route that
parallels the to the east . SR-13 leads through rural areas
of [7 Fayette 7] , 8 Walker , 9 Winston , 10 Marion and 11 Fr
anklin

1: 303 Alabama

2: 30395 Tennessee

1: 332996 State_highway

[winner: 332996 State_highway, pem_pos=0,
final_scores=-1.1061984, final_local_scores=-1.1770797,
pem=-0.024095139, similarity=-30.90566|

gt: 332996 State_highway, pem_pos=0,
final_scores=-1.1061984, final_local_scores=-1.1770797,
pem=-0.024095139, similarity=-30.90566]

2: 18618239 U.S._state

[winner: 3434750 United_States, pem_pos=0,
final_scores=0.47905046, final_local_scores=0.49457842,
pem=-0.14075686, similarity=19.614376|

gt: 18618239 U.S._state, pem_pos=6,
final_scores=-0.29907054, final_local_scores=-0.3760306,
pem=-3.9799929, similarity=14.51203]

[winner: 229379 States_of_Brazil, pem_pos=0,
final_scores=-1.6816367, final_local_scores=-1.7032278,
pem=0.0, similarity=-46.738003|

RECALL MISS: 18618239 U.S._state]

3: 104831 Killen,_Alabama

[winner: 104831 Killen,_Alabama, pem_pos=0,
final_scores=-0.04625699, final_local_scores=-0.051804632,
pem=-0.8343759, similarity=7.0812254|

gt: 104831 Killen,_Alabama, pem_pos=0,
final_scores=-0.04625699, final_local_scores=-0.051804632,
pem=-0.8343759, similarity=7.0812254]

4: 80771 Lauderdale_County,_Alabama

[winner: 94809 Lauderdale_County,_Mississippi, pem_pos=0,
final_scores=0.16234754, final_local_scores=0.15527019,
pem=-0.39773273, similarity=10.886217|

gt: 80771 Lauderdale_County,_Alabama, pem_pos=1,
final_scores=0.12624948, final_local_scores=0.11101949,
pem=-0.47271967, similarity=9.973479]

5: 77837 Fayette_County,_Alabama

[winner: 91959 Fayette_County,_Pennsylvania, pem_pos=0,
final_scores=0.08409928, final_local_scores=0.049865916,
pem=-1.2421032, similarity=12.3325615|

gt: 77837 Fayette_County,_Alabama, pem_pos=7,
final_scores=-0.121317804, final_local_scores=-0.15612122,
pem=-2.15599, similarity=11.155353]

6: 104710 Russellville,_Alabama

[winner: 115325 Russellville,_Kentucky, pem_pos=1,
final_scores=0.07219984, final_local_scores=0.033040658,
pem=-1.1734573, similarity=11.457155|

gt: 104710 Russellville,_Alabama, pem_pos=4,

final_scores=-0.09260675, final_local_scores=-0.13551503,
pem=-1.8308225, similarity=10.001971]
7: 77838 Franklin_County,_Alabama
[winner: 93120 Franklin_County,_Ohio, pem_pos=0,
final_scores=0.03917311, final_local_scores=0.009019285,
pem=-1.836298, similarity=14.34494|
gt: 77838 Franklin_County,_Alabama, pem_pos=10,
final_scores=-0.18286629, final_local_scores=-0.23774979,
pem=-2.7638488, similarity=12.025024]
8: 81321 Walker_County,_Alabama
[winner: 81321 Walker_County,_Alabama, pem_pos=3,
final_scores=-0.34646088, final_local_scores=-0.4249161,
pem=-3.253684, similarity=9.10341|
gt: 81321 Walker_County,_Alabama, pem_pos=3,
final_scores=-0.34646088, final_local_scores=-0.4249161,
pem=-3.253684, similarity=9.10341]
9: 77366 Winston_County,_Alabama
[winner: 77366 Winston_County,_Alabama, pem_pos=2,
final_scores=-0.095192164, final_local_scores=-0.13777533,
pem=-2.2205565, similarity=12.053953|
gt: 77366 Winston_County,_Alabama, pem_pos=2,
final_scores=-0.095192164, final_local_scores=-0.13777533,
pem=-2.2205565, similarity=12.053953]
10: 80788 Marion_County,_Alabama
[winner: 112080 Marion,_Illinois, pem_pos=0,
final_scores=-0.12465066, final_local_scores=-0.17707083,
pem=-2.8057163, similarity=14.063487|
gt: 80788 Marion_County,_Alabama, pem_pos=18,
final_scores=-0.5377819, final_local_scores=-0.6096806,
pem=-3.9809852, similarity=7.544844]
11: 77838 Franklin_County,_Alabama
[winner: 93120 Franklin_County,_Ohio, pem_pos=7,
final_scores=-0.2996285, final_local_scores=-0.3784392,
pem=-3.773688, similarity=13.318235|
RECALL MISS: 77838 Franklin_County,_Alabama]
1: 303 Alabama
[pem_pos=0, final_scores=0.39680195,
final_local_scores=0.3371772, pem=-0.32435498,
similarity=15.915647]
2: 30395 Tennessee
[pem_pos=0, final_scores=0.5036296,
final_local_scores=0.46547985, pem=-0.27977526,
similarity=19.502022]
1: 3434750 United_States
[pem_pos=0, final_scores=0.47905046,
final_local_scores=0.49457842, pem=-0.14075686,
similarity=19.614376]
2: 20952 Mobile,_Alabama
[pem_pos=0, final_scores=0.33515292,
final_local_scores=0.31801957, pem=-0.6921803,
similarity=17.344234]
3: 94809 Lauderdale_County,_Mississippi

[pem_pos=0, final_scores=0.16234754,
final_local_scores=0.15527019, pem=-0.39773273,
similarity=10.886217]
4: 91959 Fayette_County,_Pennsylvania
[pem_pos=0, final_scores=0.08409928,
final_local_scores=0.049865916, pem=-1.2421032,
similarity=12.3325615]
5: 115325 Russellville,_Kentucky
[pem_pos=1, final_scores=0.07219984,
final_local_scores=0.033040658, pem=-1.1734573,
similarity=11.457155]
6: 93120 Franklin_County,_Ohio
[pem_pos=0, final_scores=0.03917311,
final_local_scores=0.009019285, pem=-1.836298,
similarity=14.34494]
7: 91959 Fayette_County,_Pennsylvania
[pem_pos=1, final_scores=0.049715266,
final_local_scores=-0.0037131011, pem=-2.3120737,
similarity=16.552322]

Sample: b'Christine'

_ -LRB- name -RRB- Christine is a feminine name of [1 1 Greek 1] or origin . It is derived from the word [2 Christ 2] , which is the [3 Greek 3] translation of the [4 Hebrew 4] word '' [1 1 Messiah 1] '' . As a [5 2 French 5] variant of the [6 3 Greek 6] name Christina , it is a name in regular usage in [7 4 French 7] , [8 5 English 8] , [9 6 German 9] , [2 2 Scand inavia 2] n , [10 7 Dutch 10] and cultures , and it is often associated with the meaning `` follower of [11 Christ 11] . '' Other ways to say Christine are [12 Christina 12] , , , , Kirsten , , -LRB- most used in [3 3 Spain 3] and [4 4 Latin Am erican 4] countries -RRB- , etc. . It is a popular name ; for example , in 1968 , it was the 14th most frequently-given name for girls in the [5 5 United States 5]

1: 19616 Messiah
2: 26740 Scandinavia
3: 26667 Spain
4: 18524 Latin_America
5: 3434750 United_States

1: 66540 Ancient_Greece
[winner: 12108 Greece, pem_pos=1, final_scores=0.4833687,
final_local_scores=0.4722942, pem=-2.3247876,
similarity=30.826471]

gt: 66540 Ancient_Greece, pem_pos=4,
final_scores=-0.32772, final_local_scores=-0.38411176,
pem=-2.8341312, similarity=8.039497]

2: 10597 French_language

[winner: 5843419 France, pem_pos=0,
final_scores=0.50546795, final_local_scores=0.4865604,
pem=-0.80223274, similarity=22.972313]

gt: 10597 French_language, pem_pos=1,
final_scores=-0.23064257, final_local_scores=-0.29334655,

pem=-0.9498388, similarity=0.5010377]
3: 66540 Ancient_Greece
[winner: 12108 Greece, pem_pos=1,
final_scores=0.44721526, final_local_scores=0.43280208,
pem=-2.3247876, similarity=29.647947]
gt: 66540 Ancient_Greece, pem_pos=4,
final_scores=-0.3127095, final_local_scores=-0.36771515,
pem=-2.8341312, similarity=8.528805]
4: 962731 French_people
[winner: 5843419 France, pem_pos=0,
final_scores=0.4442951, final_local_scores=0.42244005,
pem=-0.80223274, similarity=21.058834]
gt: 962731 French_people, pem_pos=2,
final_scores=-0.13882759, final_local_scores=-0.18755051,
pem=-2.5292845, similarity=12.247474]
5: 18803164 English_people
[winner: 8569916 English_language, pem_pos=0,
final_scores=0.2339351, final_local_scores=0.21016347,
pem=-0.640266, similarity=13.843276]
gt: 18803164 English_people, pem_pos=2,
final_scores=-0.28960764, final_local_scores=-0.28868774,
pem=-2.2084396, similarity=7.4845324]
6: 152735 Germans
[winner: 11867 Germany, pem_pos=1,
final_scores=0.37797636, final_local_scores=0.35426262,
pem=-1.0465823, similarity=20.353094]
gt: 152735 Germans, pem_pos=2, final_scores=-0.19531798,
final_local_scores=-0.25438118, pem=-2.1953664,
similarity=8.437215]
7: 1656551 Dutch_people
[winner: 21148 Netherlands, pem_pos=0,
final_scores=0.38555866, final_local_scores=0.37350726,
pem=-0.7943452, similarity=19.555687]
gt: 1656551 Dutch_people, pem_pos=2,
final_scores=-0.21122353, final_local_scores=-0.26034382,
pem=-2.0349581, similarity=7.386953]
1: 19616 Messiah
[pem_pos=0, final_scores=0.1686111,
final_local_scores=0.19682911, pem=-0.56475556,
similarity=13.034715]
2: 26740 Scandinavia
[pem_pos=0, final_scores=0.2657023,
final_local_scores=0.23402154, pem=-0.08773311,
similarity=11.550489]
3: 26667 Spain
[pem_pos=0, final_scores=0.59545606,
final_local_scores=0.6019739, pem=-0.18857051,
similarity=23.07929]
4: 18524 Latin_America
[pem_pos=0, final_scores=0.39751804,
final_local_scores=0.39422518, pem=-0.13736542,
similarity=16.601194]

5: 3434750 United_States
[pem_pos=0, final_scores=0.50103205,
final_local_scores=0.5000532, pem=-0.073457584,
similarity=19.411774]
1: 12108 Greece
[pem_pos=1, final_scores=0.4833687,
final_local_scores=0.4722942, pem=-2.3247876,
similarity=30.826471]
2: 5185 Christ
[pem_pos=0, final_scores=0.4056517,
final_local_scores=0.43358052, pem=-0.13607849,
similarity=17.768637]
3: 12108 Greece
[pem_pos=1, final_scores=0.26038653,
final_local_scores=0.22872007, pem=-2.3247876,
similarity=23.557732]
4: 13450 Hebrew_language
[pem_pos=0, final_scores=0.22256036,
final_local_scores=0.20641041, pem=-0.08056865,
similarity=10.687555]
5: 5843419 France
[pem_pos=0, final_scores=0.50546795,
final_local_scores=0.4865604, pem=-0.80223274,
similarity=22.972313]
6: 12108 Greece
[pem_pos=1, final_scores=0.44721526,
final_local_scores=0.43280208, pem=-2.3247876,
similarity=29.647947]
7: 5843419 France
[pem_pos=0, final_scores=0.4442951,
final_local_scores=0.42244005, pem=-0.80223274,
similarity=21.058834]
8: 8569916 English_language
[pem_pos=0, final_scores=0.2339351,
final_local_scores=0.21016347, pem=-0.640266,
similarity=13.843276]
9: 11867 Germany
[pem_pos=1, final_scores=0.37797636,
final_local_scores=0.35426262, pem=-1.0465823,
similarity=20.353094]
10: 21148 Netherlands
[pem_pos=0, final_scores=0.38555866,
final_local_scores=0.37350726, pem=-0.7943452,
similarity=19.555687]
11: 5185 Christ
[pem_pos=0, final_scores=0.42207938,
final_local_scores=0.45152527, pem=-0.13607849,
similarity=18.304144]
12: 49613 Christina,_Queen_of_Sweden
[pem_pos=0, final_scores=0.032038733,
final_local_scores=0.054940447, pem=-0.8929986,
similarity=10.58551]

Sample: b'D-block_contraction'

d-block contraction -LRB- sometimes called scandide contraction -RRB- is a term used in chemistry to describe the effect of having full d orbitals on the 1 period 4 elements . The elements in question are the 2 Ga , 3 Ge , 4 As , 5 Se and 6 Br . Their electronic configurations include completely filled d orbitals -LRB- d10 -RRB- . d-block contraction is best illustrated by comparing some properties of the 7 group 1 3 elements to highlight the effect on gallium . Gallium can be seen to be anomalous . The most obvious effect is that the sum of the first three s of gallium is higher than that of aluminium , whereas the trend in the group would be for it to be lower . The second table below shows the trend in the sum of the first three ionization potentials for the elements 8 B , [1 1 Al 1] , 9 Sc , 10 Y , 11 La . 12 Sc , 13 Y , 14 La are 15 group 3 element s and have three valence electrons above a 16 noble gas electron core . In contrast to the 17 group 13 elements this sequence shows a smooth reduction . Other effects of the d-block contraction are that the Ga³⁺ ion is smaller than expected , being closer in size to [1 Al³⁺ 1] + . Care must be taken in interpreting the ionization potentials for indium and thallium , other effects e.g. the 18 inert pair effect become increasingly important for the heavier members of the group . The cause of the d-block contraction is the poor shielding of the nuclear charge by the electrons in the d orbitals . The outer valence electrons are more strongly attracted by the nucleus causing the observed increase in ionization potentials . The d-block contraction can be compared to the 19 lanthanide contraction

1: 904 Aluminium

1: 181554 Period_4_element

[winner: 181554 Period_4_element, pem_pos=0, final_scores=-0.6112713, final_local_scores=-0.6362541, pem=-0.16705407, similarity=-13.988913]

gt: 181554 Period_4_element, pem_pos=0, final_scores=-0.6112713, final_local_scores=-0.6362541, pem=-0.16705407, similarity=-13.988913]

2: 12241 Gallium

[winner: 2534573 Ga-Adangbe_people, pem_pos=2, final_scores=-0.32205048, final_local_scores=-0.18875375, pem=-1.7439415, similarity=7.9407496]

gt: 12241 Gallium, pem_pos=0, final_scores=-0.39467973, final_local_scores=-0.44050032, pem=-0.814508, similarity=-4.6262736]

3: 12242 Germanium

[winner: 12242 Germanium, pem_pos=0, final_scores=-0.27748168, final_local_scores=-0.3172616, pem=-0.20409533, similarity=-4.268104]

gt: 12242 Germanium, pem_pos=0, final_scores=-0.27748168, final_local_scores=-0.3172616, pem=-0.20409533, similarity=-4.268104]

4: 897 Arsenic no overlap with any filtered span
5: 27117 Selenium
[winner: 27117 Selenium, pem_pos=0,
final_scores=-0.35139614, final_local_scores=-0.39675766,
pem=-0.37795234, similarity=-5.6949635|
gt: 27117 Selenium, pem_pos=0, final_scores=-0.35139614,
final_local_scores=-0.39675766, pem=-0.37795234,
similarity=-5.6949635]
6: 3756 Bromine
[winner: 3756 Bromine, pem_pos=0,
final_scores=-0.24710055, final_local_scores=-0.2747061,
pem=-0.11717643, similarity=-3.4708414|
gt: 3756 Bromine, pem_pos=0, final_scores=-0.24710055,
final_local_scores=-0.2747061, pem=-0.11717643,
similarity=-3.4708414]
7: 144929 Boron_group
[winner: 15484199 Group_13, pem_pos=0,
final_scores=-0.7499654, final_local_scores=-0.62241507,
pem=0.0, similarity=-14.484393|
gt: 144929 Boron_group, pem_pos=2,
final_scores=-1.2981819, final_local_scores=-1.4081702,
pem=-3.281238, similarity=-20.089012]
[winner: 5659 Chemical_element, pem_pos=0,
final_scores=-0.99628586, final_local_scores=-1.0769894,
pem=-0.7723687, similarity=-23.849537|
RECALL MISS: 144929 Boron_group]
8: 3755 Boron
[winner: 3755 Boron, pem_pos=8, final_scores=-0.6193715,
final_local_scores=-0.70996803, pem=-2.7747905,
similarity=-2.0074067|
gt: 3755 Boron, pem_pos=8, final_scores=-0.6193715,
final_local_scores=-0.70996803, pem=-2.7747905,
similarity=-2.0074067]
9: 27116 Scandium
[winner: 27116 Scandium, pem_pos=0,
final_scores=-0.15275784, final_local_scores=-0.18547684,
pem=-0.34960154, similarity=0.45590144|
gt: 27116 Scandium, pem_pos=0, final_scores=-0.15275784,
final_local_scores=-0.18547684, pem=-0.34960154,
similarity=0.45590144]
10: 20373503 Yttrium no overlap with any filtered span
11: 17744 Lanthanum
[winner: 17744 Lanthanum, pem_pos=0,
final_scores=-0.3852964, final_local_scores=-0.42289126,
pem=-0.8045514, similarity=-4.15493|
gt: 17744 Lanthanum, pem_pos=0, final_scores=-0.3852964,
final_local_scores=-0.42289126, pem=-0.8045514,
similarity=-4.15493]
12: 27116 Scandium
[winner: 27116 Scandium, pem_pos=0,
final_scores=-0.3994075, final_local_scores=-0.4549041,
pem=-0.34960154, similarity=-7.584346|

gt: 27116 Scandium, pem_pos=0, final_scores=-0.3994075,
final_local_scores=-0.4549041, pem=-0.34960154,
similarity=-7.584346]
13: 20373503 Yttrium no overlap with any filtered span
14: 17744 Lanthanum
[winner: 17744 Lanthanum, pem_pos=0,
final_scores=-0.38797516, final_local_scores=-0.42581743,
pem=-0.8045514, similarity=-4.242253]
gt: 17744 Lanthanum, pem_pos=0,
final_scores=-0.38797516, final_local_scores=-0.42581743,
pem=-0.8045514, similarity=-4.242253]
15: 306609 Group_3_element
[winner: 5659 Chemical_element, pem_pos=0,
final_scores=-0.986991, final_local_scores=-1.0668362,
pem=-0.63089556, similarity=-24.3159]
RECALL MISS: 306609 Group_3_element]
[winner: 8355 December_3, pem_pos=13,
final_scores=-1.0141913, final_local_scores=-0.8425192,
pem=-3.995643, similarity=0.6761865]
gt: 306609 Group_3_element, pem_pos=17,
final_scores=-1.5058988, final_local_scores=-1.6055486,
pem=-4.233116, similarity=-20.802717]
[winner: 19447 Group_(mathematics), pem_pos=1,
final_scores=-1.3034513, final_local_scores=-1.3285847,
pem=-1.2354953, similarity=-28.83909]
RECALL MISS: 306609 Group_3_element]
16: 21140 Noble_gas
[winner: 21140 Noble_gas, pem_pos=0,
final_scores=-0.80249214, final_local_scores=-0.8543718,
pem=0.0, similarity=-21.406445]
gt: 21140 Noble_gas, pem_pos=0,
final_scores=-0.80249214, final_local_scores=-0.8543718,
pem=0.0, similarity=-21.406445]
17: 144929 Boron_group
[winner: 15484199 Group_13, pem_pos=0,
final_scores=-0.75749385, final_local_scores=-0.6306387,
pem=0.0, similarity=-14.729804]
gt: 144929 Boron_group, pem_pos=2,
final_scores=-1.3645673, final_local_scores=-1.4806862,
pem=-3.281238, similarity=-22.253033]
[winner: 5659 Chemical_element, pem_pos=0,
final_scores=-1.0922278, final_local_scores=-1.1817913,
pem=-0.7723687, similarity=-26.977036]
RECALL MISS: 144929 Boron_group]
18: 827232 Inert_pair_effect
[winner: 248159 Inert_gas, pem_pos=0,
final_scores=-0.445782, final_local_scores=-0.47290984,
pem=0.0, similarity=-10.02286]
RECALL MISS: 827232 Inert_pair_effect]
[winner: 80506 Twisted_pair, pem_pos=3,
final_scores=-1.3555694, final_local_scores=-1.40727,
pem=-2.2308412, similarity=-25.774364]

RECALL MISS: 827232 Inert_pair_effect]
 [winner: 10165 Effects_unit, pem_pos=1,
 final_scores=-1.4655893, final_local_scores=-1.5071452,
 pem=-1.0228662, similarity=-35.324]
 RECALL MISS: 827232 Inert_pair_effect]
 19: 2168393 Lanthanide_contraction
 [winner: 2168393 Lanthanide_contraction, pem_pos=0,
 final_scores=-0.706058, final_local_scores=-0.722447,
 pem=0.0, similarity=-17.469545]
 gt: 2168393 Lanthanide_contraction, pem_pos=0,
 final_scores=-0.706058, final_local_scores=-0.722447,
 pem=0.0, similarity=-17.469545]
 1: 904 Aluminium
 [pem_pos=0, final_scores=0.12662734,
 final_local_scores=0.010229722, pem=-0.52665496,
 similarity=7.25902]
 1: 904 Aluminium
 [pem_pos=0, final_scores=0.27227384,
 final_local_scores=0.16932642, pem=0.0, similarity=9.142751]

Sample: b'Kenneth_Willis_Clark_Collection'

The Kenneth Willis Clark Collection of Greek Manuscripts in [
 1 1 Duke University Library 1] contains ninety-eight 2 manusc
 ript s -- in both roll and codex form -- dating from the 9th
 to the 17th century . Professor H. Branscomb of the [2 3 Duke
 Divinity School 2] bought a manuscript of the [3 Greek 4 New T
 estament 3] in the [1 1 Munich 1] bookshop . The manuscript
 after his arriving to the Library became Duke Greek Ms. 1 .
 It was happen in 19 February 1931 , and it was the beginning
 of the collection . Although according the primary intention
 was to collect only manuscripts of the [4 New Testament 4] ,
 today the collection contains a variety of materials . In the
 collection are 27 manuscripts which contain texts of the [2 2
 New Testament 2] . Among this number -- Mss . 5 4 , 6 5 , 7 6
 , 8 15 , 9 22 , 10 25 , 11 31 , 12 38 , , and 13 64 . Ms.
 60 , also known as 14 Codex Daltonianus , is most notable
 among this group . Written in the latter half of the 11th
 century , it contains commentaries to the four . There are
 also some 15 Lectionaries in collection , represented by
 MSS . 16 1 , 17 2 , 18 10 , 19 12 , 20 , 20 24 , 21 27 , 22 28
 , 23 39 , 24 43 , 25 65 , 26 82 , 27 83 , 85 . Two
 lectionaries are more notable : Ms 65 and 85 . 28 Ms 65 -LRB-
 ø 1839 -RRB- was written in 11th century . Ms. 85 -LRB- ø 451
 -RRB-

1: 19058 Munich

2: 21433 New_Testament

1: 53273 Duke_University

[winner: 23804859 Duke_University_Libraries, pem_pos=0,
 final_scores=0.22237127, final_local_scores=0.30087504,
 pem=0.0, similarity=13.068424]

RECALL MISS: 53273 Duke_University]

2: 19167 Manuscript

[winner: 19167 Manuscript, pem_pos=0,
final_scores=-0.9288668, final_local_scores=-1.0634081,
pem=-0.0290608, similarity=-27.486471|
gt: 19167 Manuscript, pem_pos=0,
final_scores=-0.9288668, final_local_scores=-1.0634081,
pem=-0.0290608, similarity=-27.486471]
3: 3067115 Duke_Divinity_School
[winner: 53273 Duke_University, pem_pos=1,
final_scores=0.1840835, final_local_scores=0.24230605,
pem=-0.6853952, similarity=15.0478945|
gt: 3067115 Duke_Divinity_School, pem_pos=0,
final_scores=0.15632586, final_local_scores=0.20457941,
pem=0.0, similarity=10.194771]
4: 21433 New_Testament
[winner: 2056409 Novum_Testamentum_Graece, pem_pos=0,
final_scores=0.10951446, final_local_scores=0.117259294,
pem=-0.055702526, similarity=7.891883|
gt: 21433 New_Testament, pem_pos=3,
final_scores=-0.62439626, final_local_scores=-0.7371274,
pem=-4.67904, similarity=7.5377073]
5: 20491388 Minuscule_2268
[winner: 27433944 4,_5,_6, pem_pos=0,
final_scores=-0.5952538, final_local_scores=-0.43497333,
pem=0.0, similarity=-8.8907585|
RECALL MISS: 20491388 Minuscule_2268]
6: 20483851 Minuscule_2612
[winner: 27433944 4,_5,_6, pem_pos=0,
final_scores=-0.5952538, final_local_scores=-0.43497333,
pem=0.0, similarity=-8.8907585|
RECALL MISS: 20483851 Minuscule_2612]
7: 20484167 Minuscule_2613
[winner: 27433944 4,_5,_6, pem_pos=0,
final_scores=-0.5952538, final_local_scores=-0.43497333,
pem=0.0, similarity=-8.8907585|
RECALL MISS: 20484167 Minuscule_2613]
8: 20488406 Minuscule_2615
[winner: 19674 May_15, pem_pos=1, final_scores=-1.006759,
final_local_scores=-0.8282372, pem=-3.1076534,
similarity=-3.7266352|
RECALL MISS: 20488406 Minuscule_2615]
9: 20760372 Minuscule_2491
[winner: 8728 December_22, pem_pos=8,
final_scores=-0.90528965, final_local_scores=-0.7815787,
pem=-3.401866, similarity=-0.7342795|
RECALL MISS: 20760372 Minuscule_2491]
10: 20978273 Minuscule_1813
[winner: 2733 April_25, pem_pos=0,
final_scores=-0.95987225, final_local_scores=-0.8663816,
pem=-2.8852773, similarity=-6.074255|
RECALL MISS: 20978273 Minuscule_1813]
11: 20680284 Minuscule_2766
[winner: 19653 May_31, pem_pos=1,

final_scores=-0.9188203, final_local_scores=-0.6963668,
 pem=-2.705463, similarity=-1.9785334|
 RECALL MISS: 20680284 Minuscule_2766]

12: 20673375 Minuscule_2757
 [winner: 21882494 House_at_38_Salem_Street, pem_pos=16,
 final_scores=-1.2693396, final_local_scores=-1.2189572,
 pem=-3.9271958, similarity=-10.929698|
 RECALL MISS: 20673375 Minuscule_2757]

13: 20673375 Minuscule_2757
 [winner: 20314 March_5, pem_pos=15,
 final_scores=-1.0038557, final_local_scores=-0.8120435,
 pem=-3.9054012, similarity=1.0948946|
 RECALL MISS: 20673375 Minuscule_2757]

14: 20847120 Codex_Daltonianus
 [winner: 20847120 Codex_Daltonianus, pem_pos=0,
 final_scores=0.010296389, final_local_scores=0.07468371,
 pem=0.0, similarity=6.3184237|
 gt: 20847120 Codex_Daltonianus, pem_pos=0,
 final_scores=0.010296389, final_local_scores=0.07468371,
 pem=0.0, similarity=6.3184237]

15: 1520065 Lectionary
 [winner: 1520065 Lectionary, pem_pos=0,
 final_scores=-0.11021927, final_local_scores=-0.110422686,
 pem=0.0, similarity=0.7944796|
 gt: 1520065 Lectionary, pem_pos=0,
 final_scores=-0.11021927, final_local_scores=-0.110422686,
 pem=0.0, similarity=0.7944796]

16: 20679472 Minuscule_1780
 [winner: 19348 May_1, pem_pos=3, final_scores=-1.045038,
 final_local_scores=-0.87987554, pem=-3.7402232,
 similarity=-1.8276129|
 RECALL MISS: 20679472 Minuscule_1780]

17: 21057079 Lectionary_1619
 [winner: 19349 May_2, pem_pos=4, final_scores=-1.0180092,
 final_local_scores=-0.8413478, pem=-3.709198,
 similarity=-0.8465887|
 RECALL MISS: 21057079 Lectionary_1619]

18: 21054737 Lectionary_1965
 [winner: 22406918 HMS_St_Lawrence_(1813), pem_pos=0,
 final_scores=-0.8217399, final_local_scores=-0.73242205,
 pem=0.0, similarity=-17.767221|
 RECALL MISS: 21054737 Lectionary_1965]

19: 20944761 Lectionary_1966
 [winner: 19497 May_12, pem_pos=3,
 final_scores=-1.0246266, final_local_scores=-0.8421059,
 pem=-3.2948976, similarity=-3.1222436|
 RECALL MISS: 20944761 Lectionary_1966]

20: 20962686 Lectionary_1967
 [winner: 19459 May_24, pem_pos=3,
 final_scores=-0.9095174, final_local_scores=-0.70011836,
 pem=-3.5294495, similarity=2.3904781|
 RECALL MISS: 20962686 Lectionary_1967]

21: 20986847 Lectionary_2144
 [winner: 19624 May_27, pem_pos=2,
 final_scores=-0.92441314, final_local_scores=-0.7552667,
 pem=-3.0017982, similarity=-2.1247084]
 RECALL MISS: 20986847 Lectionary_2144]

22: 20986953 Lectionary_648
 [winner: 19672 May_28, pem_pos=2,
 final_scores=-0.95923805, final_local_scores=-0.79469156,
 pem=-3.0133107, similarity=-3.2386165]
 RECALL MISS: 20986953 Lectionary_648]

23: 21043766 Lectionary_2138
 [winner: 1345303 Az-Zumar, pem_pos=6,
 final_scores=-1.0096203, final_local_scores=-0.8489648,
 pem=-3.4948552, similarity=-2.2395246]
 RECALL MISS: 21043766 Lectionary_2138]

24: 21057480 Lectionary_2145
 [winner: 11535228 National_Highway_43_(India), pem_pos=3,
 final_scores=-1.1858056, final_local_scores=-1.0793153,
 pem=-3.187552, similarity=-10.784804]
 RECALL MISS: 21057480 Lectionary_2145]

25: 20884101 Lectionary_1839
 [winner: 14350881 No._65_Squadron_RAF, pem_pos=0,
 final_scores=-1.0125197, final_local_scores=-0.9411036,
 pem=-2.422572, similarity=-10.820372]
 RECALL MISS: 20884101 Lectionary_1839]

26: 21057115 Lectionary_1623
 [winner: 34756 1982, pem_pos=2, final_scores=-1.0079955,
 final_local_scores=-0.8561584, pem=-2.9408827,
 similarity=-5.4667835]
 RECALL MISS: 21057115 Lectionary_1623]

27: 21043631 Lectionary_302
 [winner: 12046944 Malaysia_Federal_Route_83, pem_pos=6,
 final_scores=-1.080836, final_local_scores=-0.96484935,
 pem=-3.114813, similarity=-7.764478]
 RECALL MISS: 21043631 Lectionary_302]

28: 20884101 Lectionary_1839
 [winner: 20884101 Lectionary_1839, pem_pos=0,
 final_scores=-0.17966218, final_local_scores=-0.15856592,
 pem=0.0, similarity=-0.6422104]
 gt: 20884101 Lectionary_1839, pem_pos=0,
 final_scores=-0.17966218, final_local_scores=-0.15856592,
 pem=0.0, similarity=-0.6422104]

1: 19058 Munich
 [pem_pos=0, final_scores=0.4877528,
 final_local_scores=0.59195846, pem=-0.119245164,
 similarity=22.403408]

2: 21433 New_Testament
 [pem_pos=0, final_scores=0.17642964,
 final_local_scores=0.15355644, pem=-0.03720609,
 similarity=8.8744755]

1: 23804859 Duke_University_Libraries
 [pem_pos=0, final_scores=0.22237127,

final_local_scores=0.30087504, pem=0.0, similarity=13.068424]
 2: 53273 Duke_University
 [pem_pos=1, final_scores=0.1840835,
 final_local_scores=0.24230605, pem=-0.6853952,
 similarity=15.0478945]
 3: 2056409 Novum_Testamentum_Graece
 [pem_pos=0, final_scores=0.10951446,
 final_local_scores=0.117259294, pem=-0.055702526,
 similarity=7.891883]
 4: 21433 New_Testament
 [pem_pos=0, final_scores=0.14732347,
 final_local_scores=0.121762365, pem=-0.03720609,
 similarity=7.925677]

Sample: b'2009_European_Pairs_Speedway_Championship'

The 2009 [1 European Pairs Speedway Championship 1] will be the 6th 1 UEM 2 European Pairs Speedway Championship season . The Final was held on 26 September 3 2009 in [1 1 Miskolc 1] , [2 2 Hungary 2] ; it was second Final in [2 Hungary 2] , but first in [3 Miskolc 3] . The championship was won by [4 4 Czech Republic 4] pair and they beat [5 5 Russia 5] and the defending Champions [6 6 Poland 6] . In the Final will be the defending Champion [7 7 Poland 7] , [8 8 Czech Republic 8] -LRB- 2nd place in 2008 Final -RRB- , [9 9 Russia 9] -LRB- 3rd place -RRB- , host team [10 10 Hungary 10] -LRB- 4th place -RRB- and [11 11 Latvia 11] -LRB- 5th place -RRB- . A last finalist will be determined in one Semi-Final . In [3 3 Ljubljana 3] , [4 4 Slovenia 4] on May 13 will be [12 12 Austria 12] -LRB- 6th place -RRB- , [13 13 Germany 13] -LRB- 7th place -RRB- , [14 14 Ukraine 14] , [15 15 Finland 15] , host team [16 16 Slovenia 16] , [17 17 Italy 17] and [18 18 Croatia 18]

1: 235256 Miskolc
 2: 13275 Hungary
 3: 42374 Ljubljana
 4: 27338 Slovenia

1: 12407732 Union_Européenne_de_Motocyclisme
 [winner: 3526132 Universidade_Estadual_de_Maringá,
 pem_pos=2, final_scores=-0.71393913,
 final_local_scores=-0.63951737, pem=-1.0986123,
 similarity=-9.020336]
 gt: 12407732 Union_Européenne_de_Motocyclisme,
 pem_pos=0, final_scores=-0.7735017,
 final_local_scores=-0.7530339, pem=0.0, similarity=-18.38232]
 2: 12290511 European_Pairs_Speedway_Championship
 [winner: 12290511 European_Pairs_Speedway_Championship,
 pem_pos=0, final_scores=-0.3859684,
 final_local_scores=-0.31074482, pem=0.0,
 similarity=-5.1835327]
 gt: 12290511 European_Pairs_Speedway_Championship,
 pem_pos=0, final_scores=-0.3859684,
 final_local_scores=-0.31074482, pem=0.0,
 similarity=-5.1835327]

3: 4051722 2009_in_sports
[winner: 35983 2009, pem_pos=0, final_scores=-1.068618,
final_local_scores=-0.99466264, pem=-0.9296576,
similarity=-20.537382]
RECALL MISS: 4051722 2009_in_sports]

4: 13444681 Czech_Republic_national_speedway_team
[winner: 5321 Czech_Republic, pem_pos=0,
final_scores=0.23353504, final_local_scores=0.14975053,
pem=-0.27794555, similarity=10.070079]
RECALL MISS: 13444681
Czech_Republic_national_speedway_team]
[winner: 80506 Twisted_pair, pem_pos=3,
final_scores=-1.4692601, final_local_scores=-1.4291079,
pem=-2.2308412, similarity=-26.426052]
RECALL MISS: 13444681
Czech_Republic_national_speedway_team]

5: 12529786 Russia_national_speedway_team
[winner: 25127146 Russia_national_football_team,
pem_pos=2, final_scores=0.115149245,
final_local_scores=0.1382274, pem=-3.5373242,
similarity=27.451214]
RECALL MISS: 12529786 Russia_national_speedway_team]

6: 12546053 Poland_national_speedway_team
[winner: 22936 Poland, pem_pos=0,
final_scores=0.23260222, final_local_scores=0.15434644,
pem=-0.15333635, similarity=9.529586]
gt: 12546053 Poland_national_speedway_team, pem_pos=5,
final_scores=-0.5643017, final_local_scores=-0.56760335,
pem=-4.642386, similarity=12.397314]

7: 12546053 Poland_national_speedway_team
[winner: 22936 Poland, pem_pos=0,
final_scores=0.24737541, final_local_scores=0.17048389,
pem=-0.15333635, similarity=10.01116]
gt: 12546053 Poland_national_speedway_team, pem_pos=5,
final_scores=-0.5227117, final_local_scores=-0.5221726,
pem=-4.642386, similarity=13.753059]

8: 13444681 Czech_Republic_national_speedway_team
[winner: 5321 Czech_Republic, pem_pos=0,
final_scores=0.27685565, final_local_scores=0.19650644,
pem=-0.27794555, similarity=11.465368]
RECALL MISS: 13444681
Czech_Republic_national_speedway_team]

9: 12529786 Russia_national_speedway_team
[winner: 25391 Russia, pem_pos=0,
final_scores=0.31358832, final_local_scores=0.28457588,
pem=-0.20299491, similarity=13.685943]
RECALL MISS: 12529786 Russia_national_speedway_team]

10: 18408057 Hungary_national_speedway_team
[winner: 13275 Hungary, pem_pos=0,
final_scores=0.37609392, final_local_scores=0.30863282,
pem=-0.32577527, similarity=15.071548]
RECALL MISS: 18408057 Hungary_national_speedway_team]

11: 19830531 Latvia_national_speedway_team
[winner: 17514 Latvia, pem_pos=0, final_scores=0.4053151,
final_local_scores=0.34694666, pem=-0.35144013,
similarity=16.354479]
RECALL MISS: 19830531 Latvia_national_speedway_team]

12: 22288201 Austria_national_speedway_team
[winner: 26964606 Austria, pem_pos=0,
final_scores=0.12178485, final_local_scores=0.040544957,
pem=-0.25262418, similarity=6.6734653]
RECALL MISS: 22288201 Austria_national_speedway_team]

13: 19390936 Germany_national_speedway_team
[winner: 11867 Germany, pem_pos=0,
final_scores=0.3144186, final_local_scores=0.2617061,
pem=-0.15529114, similarity=12.744043]
RECALL MISS: 19390936 Germany_national_speedway_team]

14: 22287747 Ukraine_national_speedway_team
[winner: 31750 Ukraine, pem_pos=0,
final_scores=0.2733392, final_local_scores=0.21496272,
pem=-0.35407752, similarity=12.430158]
RECALL MISS: 22287747 Ukraine_national_speedway_team]

15: 21120863 Finland_national_speedway_team
[winner: 10577 Finland, pem_pos=0,
final_scores=0.3244447, final_local_scores=0.27523282,
pem=-0.2445894, similarity=13.633325]
RECALL MISS: 21120863 Finland_national_speedway_team]

16: 20169096 Slovenia_national_speedway_team
[winner: 27338 Slovenia, pem_pos=0,
final_scores=0.3536641, final_local_scores=0.30592868,
pem=-0.27845925, similarity=14.73354]
RECALL MISS: 20169096 Slovenia_national_speedway_team]

17: 19344084 Italy_national_speedway_team
[winner: 14532 Italy, pem_pos=0, final_scores=0.5461361,
final_local_scores=0.5444481, pem=-0.20612453,
similarity=21.458067]
RECALL MISS: 19344084 Italy_national_speedway_team]

18: 24130724 Croatia_national_speedway_team
[winner: 5573 Croatia, pem_pos=0,
final_scores=0.44115096, final_local_scores=0.3909747,
pem=-0.35021794, similarity=17.661716]
RECALL MISS: 24130724 Croatia_national_speedway_team]

1: 235256 Miskolc
[pem_pos=0, final_scores=0.49894744,
final_local_scores=0.49536514, pem=0.0, similarity=18.872396]

2: 13275 Hungary
[pem_pos=0, final_scores=0.5373705,
final_local_scores=0.48480296, pem=-0.32577527,
similarity=20.328817]

3: 42374 Ljubljana
[pem_pos=0, final_scores=0.5896382,
final_local_scores=0.5762139, pem=0.0, similarity=21.285086]

4: 27338 Slovenia
[pem_pos=0, final_scores=0.43358606,

final_local_scores=0.39323133, pem=-0.27845925,
similarity=17.338823]
1: 12290511 European_Pairs_Speedway_Championship
[pem_pos=0, final_scores=0.03214924,
final_local_scores=0.16293016, pem=0.0, similarity=8.951874]
2: 13275 Hungary
[pem_pos=0, final_scores=0.61333877,
final_local_scores=0.56778675, pem=-0.32577527,
similarity=22.80522]
3: 235256 Miskolc
[pem_pos=0, final_scores=0.45458418,
final_local_scores=0.44690502, pem=0.0, similarity=17.42625]
4: 5321 Czech_Republic
[pem_pos=0, final_scores=0.23353504,
final_local_scores=0.14975053, pem=-0.27794555,
similarity=10.070079]
5: 25127146 Russia_national_football_team
[pem_pos=2, final_scores=0.115149245,
final_local_scores=0.1382274, pem=-3.5373242,
similarity=27.451214]
6: 22936 Poland
[pem_pos=0, final_scores=0.23260222,
final_local_scores=0.15434644, pem=-0.15333635,
similarity=9.529586]
7: 22936 Poland
[pem_pos=0, final_scores=0.24737541,
final_local_scores=0.17048389, pem=-0.15333635,
similarity=10.01116]
8: 5321 Czech_Republic
[pem_pos=0, final_scores=0.27685565,
final_local_scores=0.19650644, pem=-0.27794555,
similarity=11.465368]
9: 25391 Russia
[pem_pos=0, final_scores=0.31358832,
final_local_scores=0.28457588, pem=-0.20299491,
similarity=13.685943]
10: 13275 Hungary
[pem_pos=0, final_scores=0.37609392,
final_local_scores=0.30863282, pem=-0.32577527,
similarity=15.071548]
11: 17514 Latvia
[pem_pos=0, final_scores=0.4053151,
final_local_scores=0.34694666, pem=-0.35144013,
similarity=16.354479]
12: 26964606 Austria
[pem_pos=0, final_scores=0.12178485,
final_local_scores=0.040544957, pem=-0.25262418,
similarity=6.6734653]
13: 11867 Germany
[pem_pos=0, final_scores=0.3144186,
final_local_scores=0.2617061, pem=-0.15529114,
similarity=12.744043]

14: 31750 Ukraine
 [pem_pos=0, final_scores=0.2733392,
 final_local_scores=0.21496272, pem=-0.35407752,
 similarity=12.430158]
 15: 10577 Finland
 [pem_pos=0, final_scores=0.3244447,
 final_local_scores=0.27523282, pem=-0.2445894,
 similarity=13.633325]
 16: 27338 Slovenia
 [pem_pos=0, final_scores=0.3536641,
 final_local_scores=0.30592868, pem=-0.27845925,
 similarity=14.73354]
 17: 14532 Italy
 [pem_pos=0, final_scores=0.5461361,
 final_local_scores=0.5444481, pem=-0.20612453,
 similarity=21.458067]
 18: 5573 Croatia
 [pem_pos=0, final_scores=0.44115096,
 final_local_scores=0.3909747, pem=-0.35021794,
 similarity=17.661716]

Sample: b'Iowa'

's _ 10th_congressional_district Iowa 's 10th congressional district existed from 1883 to 1933 , when [1 1 Iowa 1] sent eleven congressmen to the [2 2 United States House of Representatives 2] . The district , known as `` The Big Tenth , '' covered large areas of north-central [1 Iowa 1] . From 1883 to 1886 , the district included the north-central [2 Iowa 2] counties of 1 Worth , 2 Cerro Gordo , 3 Franklin , 4 Hardin , 5 Story , 6 Wright , 7 Hamilton , [3 8 Humboldt 3] , 9 Hancock , 10 Kossuth , 11 Winnebago , 12 Webster , and 13 Boone . Reapportionment in 1886 reflected the increasing population balance between eastern and western [4 Iowa 4] , resulting in a westward shift of the district 's boundaries . From 1886 until 1933 , the district was made up of Boone , 14 Calhoun , 15 Carroll , 16 Crawford , 17 Emmet , 18 Greene , Hamilton , Humboldt , Kossuth , [5 19 Palo Alto 5] , 20 Pocahontas , Winnebago , and Webster counties . After 1886 , the boundaries of the district never changed ; the [3 3 Iowa General Assembly 3] refused to reapportion its districts until the loss of two seats following the 1930 census left the State with no other choice . The district was predominantly rural , especially in its 1886 reconfiguration . During that period , it included only one of [6 Iowa 6] 's twenty largest cities -- [4 4 Fort Dodge 4] -- and included counties that had been relatively slow to settle . However , by 1895 the large area of the 10th District , coupled with increased migration to its small towns , caused it to have the largest population of any of [7 Iowa 7] 's congressional districts . By 1890 , the continued disproportionate increase in the population of the 10th and 21 11th Districts caused some to predict that the General Assembly would need to reduce the

area of each district , but no such change occurred . By 1921 , the 10th District had over 100,000 more residents than the 1st District , and nearly 60,000 more than the population of an ideally-sized [8 Iowa 8] congressional district . Every congressman elected from this district was a member of the [5 5 Republican Party 5] . As a general matter , the most influential event during each election year was not the November general election , but the [9 Republican Party 9] 's district nominating convention -LRB- or later , the [10 Republican 10] primary -RRB- . Two 10th district congressmen -LRB- [6 6 Jonathan P. Dolliver 6] and -RRB- became well-known members of the [11 U.S. 11] Senate and sought national office . Under the nine-district plan adopted by the [12 Iowa General Assembly 12] in 1931 , the boundaries of the old 10th district were preserved as the new 8th district . The last congressman elected by the old 10th district , [7 7 Fred C. Gilchrist 7] , was elected in the new 8th district in 1932 . Ten years later , when the 1940 census caused Iowa to lose another seat , the new 6th district included all of the old 10th district 's counties , plus [13 Wright County 13] . It was only after that district elected a [14 Democrat 14] -LRB- [8 8 Merwin Coad 8]

1: 26810748 Iowa

2: 19468510 United_States_House_of_Representatives

3: 1213502 Iowa_General_Assembly

4: 114334 Fort_Dodge,_Iowa

5: 32070 Republican_Party_(United_States)

6: 350053 Jonathan_P._Dolliver

7: 11585581 Fred_C._Gilchrist

8: 11585696 Merwin_Coad

1: 95903 Worth_County,_Iowa

[winner: 96668 Worth_County,_Georgia, pem_pos=0, final_scores=-0.15053079, final_local_scores=-0.134058, pem=-1.9321554, similarity=10.596517]

gt: 95903 Worth_County,_Iowa, pem_pos=15, final_scores=-0.47135878, final_local_scores=-0.49175674, pem=-3.2995057, similarity=7.3579364]

2: 96023 Cerro_Gordo_County,_Iowa

[winner: 1270798 Battle_of_Cerro_Gordo, pem_pos=0, final_scores=-0.11634724, final_local_scores=-0.03644286, pem=-1.0530221, similarity=8.728683]

gt: 96023 Cerro_Gordo_County,_Iowa, pem_pos=5, final_scores=-0.15692022, final_local_scores=-0.16173172, pem=-1.5748976, similarity=7.827854]

3: 96003 Franklin_County,_Iowa

[winner: 95721 Franklin_County,_Kansas, pem_pos=1, final_scores=-0.26041257, final_local_scores=-0.28015378, pem=-3.353311, similarity=13.965192]

RECALL MISS: 96003 Franklin_County,_Iowa]

4: 95995 Hardin_County,_Iowa

[winner: 93112 Hardin_County,_Ohio, pem_pos=4, final_scores=-0.081718996, final_local_scores=-0.088121116,

pem=-2.263782, similarity=13.770802|
gt: 95995 Hardin_County,_Iowa, pem_pos=6,
final_scores=-0.21200679, final_local_scores=-0.2286171,
pem=-2.77283, similarity=12.346404]
5: 95939 Story_County,_Iowa
[winner: 95939 Story_County,_Iowa, pem_pos=1,
final_scores=-0.17993969, final_local_scores=-0.22066751,
pem=-2.4111757, similarity=10.616903|
gt: 95939 Story_County,_Iowa, pem_pos=1,
final_scores=-0.17993969, final_local_scores=-0.22066751,
pem=-2.4111757, similarity=10.616903]
6: 95902 Wright_County,_Iowa
[winner: 94881 Wright_County,_Minnesota, pem_pos=2,
final_scores=-0.36363268, final_local_scores=-0.40020117,
pem=-3.3966992, similarity=10.6186905|
gt: 95902 Wright_County,_Iowa, pem_pos=16,
final_scores=-0.46325085, final_local_scores=-0.50120556,
pem=-4.2298465, similarity=12.135303]
7: 95998 Hamilton_County,_Iowa
[winner: 14288 Hamilton,_Ontario, pem_pos=0,
final_scores=-0.2216955, final_local_scores=-0.23708363,
pem=-1.8266565, similarity=6.9483066|
RECALL MISS: 95998 Hamilton_County,_Iowa]
8: 95991 Humboldt_County,_Iowa
[winner: 80544 Humboldt_County,_California, pem_pos=1,
final_scores=0.033354446, final_local_scores=0.027769804,
pem=-1.6302307, similarity=13.783867|
gt: 95991 Humboldt_County,_Iowa, pem_pos=9,
final_scores=-0.37681943, final_local_scores=-0.41052535,
pem=-3.2458613, similarity=9.490316]
9: 95996 Hancock_County,_Iowa
[winner: 95496 Hancock_County,_Maine, pem_pos=1,
final_scores=-0.22158362, final_local_scores=-0.23348492,
pem=-2.8017802, similarity=12.358573|
RECALL MISS: 95996 Hancock_County,_Iowa]
10: 95981 Kossuth_County,_Iowa
[winner: 95981 Kossuth_County,_Iowa, pem_pos=1,
final_scores=-0.01635547, final_local_scores=-0.026555404,
pem=-0.92249054, similarity=8.313893|
gt: 95981 Kossuth_County,_Iowa, pem_pos=1,
final_scores=-0.01635547, final_local_scores=-0.026555404,
pem=-0.92249054, similarity=8.313893]
11: 95906 Winnebago_County,_Iowa
[winner: 96151 Winnebago_County,_Illinois, pem_pos=0,
final_scores=-0.023670018, final_local_scores=-0.039677903,
pem=-1.3697716, similarity=10.354676|
gt: 95906 Winnebago_County,_Iowa, pem_pos=18,
final_scores=-0.40053248, final_local_scores=-0.41577652,
pem=-3.2365646, similarity=9.283054]
12: 95907 Webster_County,_Iowa
[winner: 97657 Webster_Parish,_Louisiana, pem_pos=7,
final_scores=-0.29250377, final_local_scores=-0.3629895,

pem=-3.345952, similarity=11.45319|
RECALL MISS: 95907 Webster_County,_Iowa]

13: 96032 Boone_County,_Iowa
[winner: 4512 Boone,_North_Carolina, pem_pos=0,
final_scores=-0.16706225, final_local_scores=-0.18077382,
pem=-1.8851963, similarity=8.947052|
gt: 96032 Boone_County,_Iowa, pem_pos=9,
final_scores=-0.45028988, final_local_scores=-0.48758715,
pem=-3.0839436, similarity=6.3101053]

14: 96027 Calhoun_County,_Iowa
[winner: 52110 John_C._Calhoun, pem_pos=0,
final_scores=-0.15356389, final_local_scores=-0.18848887,
pem=-1.9563721, similarity=9.103886|
gt: 96027 Calhoun_County,_Iowa, pem_pos=14,
final_scores=-0.43934554, final_local_scores=-0.46239844,
pem=-3.363293, similarity=8.580931]

15: 96026 Carroll_County,_Iowa
[winner: 96282 Carroll_County,_Illinois, pem_pos=2,
final_scores=-0.36669362, final_local_scores=-0.40314668,
pem=-2.7271717, similarity=6.889795|
gt: 96026 Carroll_County,_Iowa, pem_pos=13,
final_scores=-0.6196279, final_local_scores=-0.66138905,
pem=-3.7520165, similarity=4.7565937]

16: 96015 Crawford_County,_Iowa
[winner: 96266 Crawford_County,_Illinois, pem_pos=2,
final_scores=-0.3172018, final_local_scores=-0.35238692,
pem=-2.8027444, similarity=8.815544|
gt: 96015 Crawford_County,_Iowa, pem_pos=14,
final_scores=-0.6418071, final_local_scores=-0.6713304,
pem=-3.7151763, similarity=4.2595816]

17: 96006 Emmet_County,_Iowa
[winner: 95099 Emmet_County,_Michigan, pem_pos=0,
final_scores=-0.21486188, final_local_scores=-0.20823039,
pem=-1.6098409, similarity=6.6302676|
gt: 96006 Emmet_County,_Iowa, pem_pos=3,
final_scores=-0.33697614, final_local_scores=-0.3246438,
pem=-2.1135745, similarity=5.8956423]

18: 96001 Greene_County,_Iowa
[winner: 54338 Greene_County,_New_York, pem_pos=0,
final_scores=-0.3248551, final_local_scores=-0.3588684,
pem=-2.5238388, similarity=7.105392|
gt: 96001 Greene_County,_Iowa, pem_pos=16,
final_scores=-0.47273552, final_local_scores=-0.5006937,
pem=-3.920678, similarity=10.46927]

19: 95962 Palo_Alto_County,_Iowa
[winner: 53445 Palo_Alto,_California, pem_pos=0,
final_scores=0.30769664, final_local_scores=0.345766,
pem=-0.09296338, similarity=14.9136095|
RECALL MISS: 95962 Palo_Alto_County,_Iowa]

20: 95960 Pocahontas_County,_Iowa
[winner: 237523 Pocahontas, pem_pos=0,
final_scores=-0.037464783, final_local_scores=-0.012317084,

```

pem=-0.6703434, similarity=7.367578|
    gt: 95960 Pocahontas County, Iowa, pem_pos=19,
final_scores=-0.6530591, final_local_scores=-0.6834154,
pem=-4.0777345, similarity=5.870588 ]
21: 22384888 Iowa's 11th congressional district
    [winner: 3912581 11th Ward of New Orleans, pem_pos=7,
final_scores=-1.3289322, final_local_scores=-1.3098361,
pem=-3.283266, similarity=-17.143497]
    RECALL MISS: 22384888 Iowa's 11th congressional district]
1: 26810748 Iowa
    [pem_pos=0, final_scores=0.1348619,
final_local_scores=0.05667238, pem=-0.3546882,
similarity=7.7097797]
2: 19468510 United States House of Representatives
    [pem_pos=0, final_scores=0.13216855,
final_local_scores=0.07930829, pem=0.0, similarity=6.4564304]
3: 1213502 Iowa General Assembly
    [pem_pos=0, final_scores=0.1523795,
final_local_scores=0.13182095, pem=0.0, similarity=8.023513]
4: 114334 Fort Dodge, Iowa
    [pem_pos=0, final_scores=0.3172369,
final_local_scores=0.31725404, pem=0.0, similarity=13.557206]
5: 32070 Republican Party (United States)
    [pem_pos=0, final_scores=0.47936535,
final_local_scores=0.445419, pem=-0.16706309,
similarity=18.29042]
6: 350053 Jonathan P. Dolliver
    [pem_pos=0, final_scores=0.4183373,
final_local_scores=0.46274346, pem=0.0, similarity=17.898901]
7: 11585581 Fred C. Gilchrist
    [pem_pos=0, final_scores=0.23045205,
final_local_scores=0.2841375, pem=0.0, similarity=12.568942]
8: 11585696 Merwin Coad
    [pem_pos=0, final_scores=0.096450254,
final_local_scores=0.13308737, pem=0.0, similarity=8.061305]
1: 26810748 Iowa
    [pem_pos=0, final_scores=0.08253367,
final_local_scores=-0.004805878, pem=-0.3546882,
similarity=5.8751454]
2: 26810748 Iowa
    [pem_pos=0, final_scores=0.034430966,
final_local_scores=-0.0573508, pem=-0.3546882,
similarity=4.3071003]
3: 80544 Humboldt County, California
    [pem_pos=1, final_scores=0.033354446,
final_local_scores=0.027769804, pem=-1.6302307,
similarity=13.783867]
4: 26810748 Iowa
    [pem_pos=0, final_scores=0.11089985,
final_local_scores=0.030497491, pem=-0.3546882,
similarity=6.9286685]
5: 53445 Palo Alto, California

```

[pem_pos=0, final_scores=0.30769664,
 final_local_scores=0.345766, pem=-0.09296338,
 similarity=14.9136095]
 6: 26810748 Iowa
 [pem_pos=0, final_scores=0.09976421,
 final_local_scores=0.01833351, pem=-0.3546882,
 similarity=6.565671]
 7: 26810748 Iowa
 [pem_pos=0, final_scores=0.07872118,
 final_local_scores=-0.008970454, pem=-0.3546882,
 similarity=5.7508664]
 8: 26810748 Iowa
 [pem_pos=0, final_scores=0.0977665,
 final_local_scores=0.016151294, pem=-0.3546882,
 similarity=6.5005493]
 9: 32070 Republican_Party_(United_States)
 [pem_pos=0, final_scores=0.44867504,
 final_local_scores=0.4118945, pem=-0.16706309,
 similarity=17.289982]
 10: 32070 Republican_Party_(United_States)
 [pem_pos=0, final_scores=0.39910418,
 final_local_scores=0.3577459, pem=-0.1310191,
 similarity=15.4780655]
 11: 3434750 United_States
 [pem_pos=0, final_scores=0.2592427,
 final_local_scores=0.24932837, pem=-0.14075686,
 similarity=12.295627]
 12: 1213502 Iowa_General_Assembly
 [pem_pos=0, final_scores=0.1643476,
 final_local_scores=0.1448943, pem=0.0, similarity=8.413648]
 13: 94881 Wright_County,_Minnesota
 [pem_pos=0, final_scores=0.1779059,
 final_local_scores=0.21114281, pem=-0.23455217,
 similarity=11.666166]
 14: 5043544 Democratic_Party_(United_States)
 [pem_pos=0, final_scores=0.3966301,
 final_local_scores=0.360974, pem=-0.061652232,
 similarity=15.197172]

Sample: b'Derrick_Schofield'

[1 Derrick Schofield 1] -LRB- born 1928 in [1 1 Rochdale 1]
 -- died 25 August 1999 -LRB- aged 71 -RRB- -RRB- was an [2 1 E
 nglish 2] professional 2 rugby league footballer of the 1940s
 and '50s , playing at representative level for [3 3 Great Bri
 tain 3] , [4 4 England 4] , and [5 5 Lancashire 5] , and at
 club level for [2 2 Rochdale Hornets 2] -LRB- twice -RRB- , [3
 3 Halifax 3] , [4 4 Castleford 4] and [5 5 Dewsbury 5] , as
 a , or , i.e. number 2 or 5 , or , 11 or 12 . [6 Derrick
 Schofield 6] won caps for [7 6 England 7] while at [8
 Rochdale 8] in 1952 against [9 Wales 9] , in 1953 against [10
 France 10] , and won a cap for [6 6 Great Britain 6] while at
 [11 Halifax 11] in 1955 against [12 New Zealand 12] . [13

Derrick Schofield [13] also represented [7 7 Great Britain 7] while at [14 Halifax 14] between 1952 and 1956 against [15 France 15] -LRB- 1 non - 7 Test match -RRB- . [16 Derrick Schofield 16] played Right - , i.e. number 12 , in [8 8 Halifax 8] 's 4-4 draw with [9 9 Warrington 9] in the [10 10 Challenge Cup 10] final at [11 11 Wembley Stadium 11] on Saturday 1 May 1954 , and the 4-18 defeat in the [12 12 Challenge Cup 12] final replay at [13 13 Odsal Stadium 13] , [14 14 Bradford 14]

1: 635367 Rochdale
2: 2934756 Rochdale_Hornets
3: 1196374 Halifax_RLFC
4: 1216025 Castleford_Tigers
5: 2437935 Dewsbury_Rams
6: 951941 Great_Britain_national_rugby_league_team
7: 951941 Great_Britain_national_rugby_league_team
8: 1196374 Halifax_RLFC
9: 852392 Warrington_Wolves
10: 21050680 Challenge_Cup
11: 143673 Wembley_Stadium_(1923)
12: 21050680 Challenge_Cup
13: 3411441 Odsal_Stadium
14: 23861381 Bradford

1: 18803164 English_people
[winner: 9316 England, pem_pos=1, final_scores=0.663885, final_local_scores=0.6924115, pem=-1.3527626, similarity=32.109184]

gt: 18803164 English_people, pem_pos=2, final_scores=-0.44976336, final_local_scores=-0.48594794, pem=-2.2084396, similarity=1.5978951]

2: 25735 Rugby_league
[winner: 19774201 Rugby_Virginia, pem_pos=7, final_scores=-0.89117694, final_local_scores=-0.7682995, pem=-2.956099, similarity=-2.7621517]

gt: 25735 Rugby_league, pem_pos=14, final_scores=-1.8050313, final_local_scores=-2.0194643, pem=-3.3002021, similarity=-38.22812]

[winner: 1424274 League_(unit), pem_pos=2, final_scores=-1.8086343, final_local_scores=-1.8708549, pem=-1.8401146, similarity=-41.733498]

gt: 25735 Rugby_league, pem_pos=5, final_scores=-1.9384995, final_local_scores=-2.165258, pem=-3.5378149, similarity=-41.286716]

3: 951941 Great_Britain_national_rugby_league_team
[winner: 13530298 Great_Britain, pem_pos=0, final_scores=0.03156261, final_local_scores=0.04486066, pem=-0.50213265, similarity=8.15912]

gt: 951941 Great_Britain_national_rugby_league_team, pem_pos=2, final_scores=-0.01343213, final_local_scores=-0.08573794, pem=-2.6635387, similarity=16.01586]

4: 990422 England_national_rugby_league_team

[winner: 9316 England, pem_pos=0,
final_scores=0.27527595, final_local_scores=0.26791516,
pem=-0.19053856, similarity=13.121014]
gt: 990422 England_national_rugby_league_team,
pem_pos=5, final_scores=-0.45111865,
final_local_scores=-0.56605667, pem=-4.237946,
similarity=10.244062]
5: 493717 Rugby_League_War_of_the_Roses
[winner: 749919 Lancashire_County_Cricket_Club,
pem_pos=1, final_scores=0.13087718,
final_local_scores=0.12932831, pem=-1.7295884,
similarity=17.354898]
gt: 493717 Rugby_League_War_of_the_Roses, pem_pos=4,
final_scores=-0.25251424, final_local_scores=-0.3256777,
pem=-2.9341326, similarity=10.327109]
6: 990422 England_national_rugby_league_team
[winner: 9316 England, pem_pos=0,
final_scores=0.07047175, final_local_scores=0.04419768,
pem=-0.19053856, similarity=6.444838]
gt: 990422 England_national_rugby_league_team,
pem_pos=5, final_scores=-0.34414035,
final_local_scores=-0.4491991, pem=-4.237946,
similarity=13.731325]
7: 16926380 Test_match_(rugby_league)
[winner: 30863 Test_cricket, pem_pos=0,
final_scores=-0.0049604177, final_local_scores=-0.012878388,
pem=-0.23950954, similarity=5.0078864]
RECALL MISS: 16926380 Test_match_(rugby_league)]
1: 635367 Rochdale
[pem_pos=0, final_scores=0.15767737,
final_local_scores=0.15674251, pem=-0.7956193,
similarity=13.093923]
2: 2934756 Rochdale_Hornets
[pem_pos=0, final_scores=0.26107275,
final_local_scores=0.21793604, pem=0.0, similarity=10.593359]
3: 1196374 Halifax_RLFC
[pem_pos=3, final_scores=0.08373012,
final_local_scores=0.011287868, pem=-2.7129183,
similarity=19.179836]
4: 1216025 Castleford_Tigers
[pem_pos=1, final_scores=0.16044353,
final_local_scores=0.1049643, pem=-0.78848183,
similarity=11.509944]
5: 2437935 Dewsbury_Rams
[pem_pos=1, final_scores=0.11833142,
final_local_scores=0.060192823, pem=-1.8209132,
similarity=15.788397]
6: 951941 Great_Britain_national_rugby_league_team
[pem_pos=2, final_scores=0.1902392,
final_local_scores=0.14292651, pem=-2.6635387,
similarity=22.839663]
7: 951941 Great_Britain_national_rugby_league_team

[pem_pos=2, final_scores=0.12006037,
final_local_scores=0.066266775, pem=-2.6635387,
similarity=20.551983]
8: 1196374 Halifax_RLFC
[pem_pos=3, final_scores=0.14444576,
final_local_scores=0.07761049, pem=-2.7129183,
similarity=21.159037]
9: 852392 Warrington_Wolves
[pem_pos=1, final_scores=0.33639222,
final_local_scores=0.2818528, pem=-1.6982093,
similarity=21.73589]
10: 21050680 Challenge_Cup
[pem_pos=0, final_scores=0.2874788,
final_local_scores=0.24159709, pem=-0.38845974,
similarity=13.411957]
11: 143673 Wembley_Stadium_(1923)
[pem_pos=1, final_scores=0.31747156,
final_local_scores=0.28192568, pem=-0.27926707,
similarity=14.021636]
12: 21050680 Challenge_Cup
[pem_pos=0, final_scores=0.265041,
final_local_scores=0.21708718, pem=-0.38845974,
similarity=12.680532]
13: 3411441 Odsal_Stadium
[pem_pos=0, final_scores=0.30095035,
final_local_scores=0.2729198, pem=0.0, similarity=12.234184]
14: 23861381 Bradford
[pem_pos=0, final_scores=0.16871588,
final_local_scores=0.1428566, pem=-0.6938952,
similarity=12.126349]
1: 22532855 Derrick_Schofield
[pem_pos=0, final_scores=0.20990889,
final_local_scores=0.21562427, pem=0.0, similarity=10.524371]
2: 9316 England
[pem_pos=1, final_scores=0.663885,
final_local_scores=0.6924115, pem=-1.3527626,
similarity=32.109184]
3: 13530298 Great_Britain
[pem_pos=0, final_scores=0.03156261,
final_local_scores=0.04486066, pem=-0.50213265,
similarity=8.15912]
4: 9316 England
[pem_pos=0, final_scores=0.27527595,
final_local_scores=0.26791516, pem=-0.19053856,
similarity=13.121014]
5: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.13087718,
final_local_scores=0.12932831, pem=-1.7295884,
similarity=17.354898]
6: 22532855 Derrick_Schofield
[pem_pos=0, final_scores=0.5066168,
final_local_scores=0.5397325, pem=0.0, similarity=20.196407]

7: 9316 England
 [pem_pos=0, final_scores=0.07047175,
 final_local_scores=0.04419768, pem=-0.19053856,
 similarity=6.444838]
 8: 451166 Rochdale_A.F.C.
 [pem_pos=1, final_scores=0.22540252,
 final_local_scores=0.20094606, pem=-0.95097107,
 similarity=15.257874]
 9: 770930 Wales_national_rugby_union_team
 [pem_pos=2, final_scores=0.034652755,
 final_local_scores=-0.008760005, pem=-2.723462,
 similarity=18.638908]
 10: 5843419 France
 [pem_pos=0, final_scores=0.11238365,
 final_local_scores=0.15878749, pem=-0.16240834,
 similarity=9.711451]
 11: 1196374 Halifax_RLFC
 [pem_pos=3, final_scores=0.066538945,
 final_local_scores=-0.011297822, pem=-2.7129183,
 similarity=18.505835]
 12: 154957 New_Zealand_national_rugby_union_team
 [pem_pos=4, final_scores=0.39297092,
 final_local_scores=0.39736283, pem=-3.8192244,
 similarity=36.717346]
 13: 22532855 Derrick_Schofield
 [pem_pos=0, final_scores=0.4419474,
 final_local_scores=0.46909106, pem=0.0, similarity=18.088326]
 14: 1196374 Halifax_RLFC
 [pem_pos=3, final_scores=0.047901884,
 final_local_scores=-0.031655967, pem=-2.7129183,
 similarity=17.898306]
 15: 250197 France_national_football_team
 [pem_pos=1, final_scores=0.098895654,
 final_local_scores=0.096781135, pem=-3.3831942,
 similarity=25.37619]
 16: 22532855 Derrick_Schofield
 [pem_pos=0, final_scores=0.45038074,
 final_local_scores=0.47830325, pem=0.0, similarity=18.363235]

Sample: b'Southwest_Tennessee_Development_District'

[1 Southwest Tennessee Development District 1] , SWTDD is a
 planning organization , the [2 Southwest Tennessee
 Development District 2] promotes the renewal and
 revitalization of both rural and urban communities through
 betterment of an economic base -LRB- Economic Development -
 RRB- , physical infrastructure -LRB- Community Development -
 RRB- and quality of life -LRB- Area Agency on Aging &
 Disability -RRB- for an eight county area . SWTDD serves the
 following counties in [3 Tennessee 3] : [4 1 Chester County 4]
 , 2 Decatur County , [1 1 Hardeman County 1] , [5 3 Hardin Co
 unty 5] , [2 2 Haywood County 2] , [6 4 Henderson County 6] ,
 5 Madison County and [3 3 McNairy County 3] . [7 Southwest

Tennessee Development District 7] also serves twelve counties through a regional economic development initiative called REDI . [8 6 Chester County 8] , [4 4 Crockett County 4] , 7 Decatur County , , [9 8 Gibson County 9] , [5 5 Hardeman County 5] , 9 Hardin County , 10 Haywood County , [10 11 Henderson County 10] , [11 12 Lauderdale County 11] , [6 6 McNairy County 6] , and [7 7 Tipton County 7] . The SWTDD office is located in [12 13 Jackson 12] , [13 Tennessee 13]

1: 91734 Hardeman County, Tennessee

2: 91731 Haywood County, Tennessee

3: 91709 McNairy County, Tennessee

4: 91757 Crockett County, Tennessee

5: 91734 Hardeman County, Tennessee

6: 91709 McNairy County, Tennessee

7: 91687 Tipton County, Tennessee

1: 91762 Chester County, Tennessee

[winner: 91970 Chester County, Pennsylvania, pem_pos=0, final_scores=0.114855394, final_local_scores=0.109811425, pem=0.0, similarity=7.366705]

gt: 91762 Chester County, Tennessee, pem_pos=1, final_scores=-0.042026415, final_local_scores=-0.10065623, pem=-0.9702275, similarity=6.3621783]

2: 91753 Decatur County, Tennessee

[winner: 96792 Decatur County, Georgia, pem_pos=1, final_scores=-0.016367272, final_local_scores=-0.05458033, pem=-0.8800056, similarity=7.2465343]

gt: 91753 Decatur County, Tennessee, pem_pos=3, final_scores=-0.12550245, final_local_scores=-0.17166562, pem=-1.2037162, similarity=5.512864]

3: 91733 Hardin County, Tennessee

[winner: 95595 Hardin County, Kentucky, pem_pos=0, final_scores=0.057785258, final_local_scores=0.009281501, pem=-0.8375672, similarity=8.921512]

gt: 91733 Hardin County, Tennessee, pem_pos=3, final_scores=-0.051116496, final_local_scores=-0.107343435, pem=-1.0956672, similarity=6.8447795]

4: 91730 Henderson County, Tennessee

[winner: 91558 Henderson County, Texas, pem_pos=0, final_scores=0.13650198, final_local_scores=0.092487395, pem=-0.57645446, similarity=9.9845705]

gt: 91730 Henderson County, Tennessee, pem_pos=4, final_scores=-0.16163014, final_local_scores=-0.21836595, pem=-1.2486126, similarity=4.363387]

5: 91714 Madison County, Tennessee

[winner: 96208 Madison County, Illinois, pem_pos=2, final_scores=-0.056830987, final_local_scores=-0.10678053, pem=-1.8813272, similarity=11.134119]

gt: 91714 Madison County, Tennessee, pem_pos=10, final_scores=-0.26348084, final_local_scores=-0.34906197, pem=-2.5886054, similarity=7.750246]

6: 91762 Chester County, Tennessee

[winner: 91970 Chester County, Pennsylvania, pem_pos=0,

final_scores=0.044108525, final_local_scores=0.032531247,
pem=0.0, similarity=5.06051|
gt: 91762 Chester_County,_Tennessee, pem_pos=1,
final_scores=-0.15958965, final_local_scores=-0.2290762,
pem=-0.9702275, similarity=2.5298698]
7: 91753 Decatur_County,_Tennessee
[winner: 96792 Decatur_County,_Georgia, pem_pos=1,
final_scores=0.010621026, final_local_scores=-0.025099725,
pem=-0.8800056, similarity=8.126294|
gt: 91753 Decatur_County,_Tennessee, pem_pos=3,
final_scores=-0.10167748, final_local_scores=-0.14564046,
pem=-1.2037162, similarity=6.289507]
8: 91744 Gibson_County,_Tennessee
[winner: 96123 Gibson_County,_Indiana, pem_pos=0,
final_scores=0.12755154, final_local_scores=0.08535178,
pem=-0.39675155, similarity=8.794377|
gt: 91744 Gibson_County,_Tennessee, pem_pos=1,
final_scores=-0.013316482, final_local_scores=-0.052709997,
pem=-0.7913899, similarity=6.8204427]
9: 91733 Hardin_County,_Tennessee
[winner: 95595 Hardin_County,_Kentucky, pem_pos=0,
final_scores=0.0113980025, final_local_scores=-0.04759343,
pem=-0.8375672, similarity=7.224251|
gt: 91733 Hardin_County,_Tennessee, pem_pos=3,
final_scores=-0.096506104, final_local_scores=-0.15747929,
pem=-1.0956672, similarity=5.3486257]
10: 91731 Haywood_County,_Tennessee
[winner: 91731 Haywood_County,_Tennessee, pem_pos=1,
final_scores=0.02483429, final_local_scores=-0.029521026,
pem=-0.09158807, similarity=3.706817|
gt: 91731 Haywood_County,_Tennessee, pem_pos=1,
final_scores=0.02483429, final_local_scores=-0.029521026,
pem=-0.09158807, similarity=3.706817]
11: 91730 Henderson_County,_Tennessee
[winner: 91558 Henderson_County,_Texas, pem_pos=0,
final_scores=0.120695755, final_local_scores=0.07522151,
pem=-0.57645446, similarity=9.469322|
gt: 91730 Henderson_County,_Tennessee, pem_pos=4,
final_scores=-0.16372927, final_local_scores=-0.22065893,
pem=-1.2486126, similarity=4.29496]
12: 91720 Lauderdale_County,_Tennessee
[winner: 80771 Lauderdale_County,_Alabama, pem_pos=1,
final_scores=0.0730664, final_local_scores=0.033717662,
pem=-0.47271967, similarity=7.666638|
gt: 91720 Lauderdale_County,_Tennessee, pem_pos=2,
final_scores=-0.076650664, final_local_scores=-0.120558456,
pem=-0.8327129, similarity=5.02043]
13: 135276 Jackson,_Tennessee
[winner: 30395 Tennessee, pem_pos=0,
final_scores=0.4839236, final_local_scores=0.46159798,
pem=-0.27977526, similarity=19.386179|
RECALL MISS: 135276 Jackson,_Tennessee]

[winner: 16167 Jackson,_Mississippi, pem_pos=0,
final_scores=0.11175327, final_local_scores=0.09610471,
pem=-2.1336284, similarity=18.560673]
gt: 135276 Jackson,_Tennessee, pem_pos=3,
final_scores=-0.10369542, final_local_scores=-0.1537481,
pem=-3.3921297, similarity=17.948492]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7033957, final_local_scores=-1.659179,
pem=-4.6728287, similarity=-20.011927]

RECALL MISS: 135276 Jackson,_Tennessee]

1: 91734 Hardeman_County,_Tennessee
[pem_pos=0, final_scores=0.16202198,
final_local_scores=0.12000108, pem=-0.14857256,
similarity=8.478745]
2: 91731 Haywood_County,_Tennessee
[pem_pos=1, final_scores=0.04725091,
final_local_scores=0.0014448762, pem=-0.09158807,
similarity=4.630901]
3: 91709 McNairy_County,_Tennessee
[pem_pos=0, final_scores=0.13142876,
final_local_scores=0.10091215, pem=0.0, similarity=7.101133]
4: 91757 Crockett_County,_Tennessee
[pem_pos=0, final_scores=0.15100722,
final_local_scores=0.13548282, pem=0.0, similarity=8.132791]
5: 91734 Hardeman_County,_Tennessee
[pem_pos=0, final_scores=0.10747318,
final_local_scores=0.059121013, pem=-0.14857256,
similarity=6.6619625]
6: 91709 McNairy_County,_Tennessee
[pem_pos=0, final_scores=0.105260566,
final_local_scores=0.07232736, pem=0.0, similarity=6.2481055]
7: 91687 Tipton_County,_Tennessee
[pem_pos=0, final_scores=0.12818618,
final_local_scores=0.08719279, pem=0.0, similarity=6.6917195]
1: 25362068 Southwest_Tennessee_Development_District
[pem_pos=0, final_scores=0.23466985,
final_local_scores=0.25611666, pem=0.0, similarity=11.732745]
2: 25362068 Southwest_Tennessee_Development_District
[pem_pos=0, final_scores=0.15521483,
final_local_scores=0.1693241, pem=0.0, similarity=9.142682]
3: 30395 Tennessee
[pem_pos=0, final_scores=0.48146814,
final_local_scores=0.4589157, pem=-0.27977526,
similarity=19.306135]
4: 91970 Chester_County,_Pennsylvania
[pem_pos=0, final_scores=0.114855394,
final_local_scores=0.109811425, pem=0.0, similarity=7.366705]
5: 95595 Hardin_County,_Kentucky
[pem_pos=0, final_scores=0.057785258,
final_local_scores=0.009281501, pem=-0.8375672,
similarity=8.921512]
6: 91558 Henderson_County,_Texas

[pem_pos=0, final_scores=0.13650198,
final_local_scores=0.092487395, pem=-0.57645446,
similarity=9.9845705]
7: 25362068 Southwest_Tennessee_Development_District
[pem_pos=0, final_scores=0.20831488,
final_local_scores=0.22732785, pem=0.0, similarity=10.87363]
8: 91970 Chester_County,_Pennsylvania
[pem_pos=0, final_scores=0.044108525,
final_local_scores=0.032531247, pem=0.0, similarity=5.06051]
9: 96123 Gibson_County,_Indiana
[pem_pos=0, final_scores=0.12755154,
final_local_scores=0.08535178, pem=-0.39675155,
similarity=8.794377]
10: 91558 Henderson_County,_Texas
[pem_pos=0, final_scores=0.120695755,
final_local_scores=0.07522151, pem=-0.57645446,
similarity=9.469322]
11: 80771 Lauderdale_County,_Alabama
[pem_pos=1, final_scores=0.0730664,
final_local_scores=0.033717662, pem=-0.47271967,
similarity=7.666638]
12: 16167 Jackson,_Mississippi
[pem_pos=0, final_scores=0.11175327,
final_local_scores=0.09610471, pem=-2.1336284,
similarity=18.560673]
13: 30395 Tennessee
[pem_pos=0, final_scores=0.4839236,
final_local_scores=0.46159798, pem=-0.27977526,
similarity=19.386179]

Sample: b'Marcel_Hirscher'

[1 Marcel Hirscher 1] -LRB- born March 2 , 1989 -RRB- is a 1
World Cup 2 alpine ski racer from [1 1 Austria 1] . Born in [2
2 Annaberg-Lungötz 2] , [2 3 Salzburg 2] , [3 Hirscher 3]
made his World Cup debut in March 4 2007 . At the 5 2010 Wint
er Olympics , [4 Hirscher 4] placed fourth in the and fifth
in the at . He placed fourth in the at the 6 2009 World Champ
ionships , but broke his ankle the weekend preceding the 7 201
1 World Championships , which ended his 8 2011 season .
Returning after injury , [5 Hirscher 5] had his best season
in 9 2012 with 9 victories and a total of 14 podiums , all in
the two technical events -LRB- except for one third place in
the season 's last -RRB- . He won the World Cup overall and
giant slalom titles , and placed third in the slalom . In
October 2012 , [6 Hirscher 6] was awarded the Skieur d'Or
Award by members of the International Association of Ski
Journalists for his performances during the previous season .
[7 Hirscher 7] won the overall World Cup title again in 10 201
3 with 6 victories ; he also won the slalom title and was
runner-up in giant slalom . [8 Hirscher 8] scored a total of
18 podium finishes out of 19 races in the two technical
events . The only race in either slalom or giant slalom where

he finished outside top 3 was the giant slalom in [9 Adelboden 9] , where after leading after the first run and with a second advantage over [3 3 Ted Ligety 3] in the second several gates before the final [10 Marcel Hirscher 10] made a crucial mistake and finished only 16th . He became the first male racer to retain the overall World Cup title since fellow [11 Austrian 11] [4 4 Stephan Eberharter 4] in 11 2002 and 12 2003

1: 26964606 Austria

2: 3001568 Annaberg-Lungötz

3: 4071698 Ted_Ligety

4: 601813 Stephan_Eberharter

1: 183628 FIS_Alpine_Ski_World_Cup

[winner: 183628 FIS_Alpine_Ski_World_Cup, pem_pos=2, final_scores=-0.111150935, final_local_scores=-0.15266353, pem=-2.7628422, similarity=14.558695]

gt: 183628 FIS_Alpine_Ski_World_Cup, pem_pos=2, final_scores=-0.111150935, final_local_scores=-0.15266353, pem=-2.7628422, similarity=14.558695]

2: 39576 Alpine_skiing

[winner: 2619936 Racer_(Kennywood), pem_pos=2, final_scores=-0.9470594, final_local_scores=-0.8580008, pem=-1.4941926, similarity=-13.389091]

RECALL MISS: 39576 Alpine_skiing]

[winner: 679433 Ski_warfare, pem_pos=2, final_scores=-0.96079427, final_local_scores=-0.93283445, pem=-0.69268847, similarity=-19.98098]

gt: 39576 Alpine_skiing, pem_pos=1, final_scores=-0.9821312, final_local_scores=-1.0693592, pem=-0.65621436, similarity=-24.253502]

[winner: 23160842 Alpine_Mendocino_County,_California, pem_pos=9, final_scores=-1.1588448, final_local_scores=-1.0792501, pem=-3.0050297, similarity=-11.775445]

gt: 39576 Alpine_skiing, pem_pos=13, final_scores=-1.9678247, final_local_scores=-2.1460795, pem=-3.601453, similarity=-40.36832]

3: 55859 Salzburg_(state)

[winner: 40335 Salzburg, pem_pos=0, final_scores=0.3554343, final_local_scores=0.44158256, pem=-0.11955855, similarity=17.917595]

gt: 55859 Salzburg_(state), pem_pos=1, final_scores=-0.12381153, final_local_scores=-0.09285042, pem=-1.7480965, similarity=10.825292]

4: 4243107 2007_Alpine_Skiing_World_Cup

[winner: 8793874 March_2007, pem_pos=0, final_scores=-0.44887036, final_local_scores=-0.32200912, pem=-0.09072522, similarity=-5.026304]

RECALL MISS: 4243107 2007_Alpine_Skiing_World_Cup]

5: 4054119 Alpine_skiing_at_the_Winter_Olympics

[winner: 27566 Summer_Olympic_Games, pem_pos=2, final_scores=-0.07870634, final_local_scores=-0.055689752,

pem=-2.7333555, similarity=17.292233|
RECALL MISS: 4054119
Alpine_skiing_at_the_Winter_Olympics]
[winner: 178749 2010_Winter_Olympics, pem_pos=0,
final_scores=-0.63484776, final_local_scores=-0.68222666,
pem=0.0, similarity=-16.26929|
RECALL MISS: 4054119
Alpine_skiing_at_the_Winter_Olympics]
6: 5403315 FIS_Alpine_World_Ski_Championships_2009
[winner: 218926 IAAF_World_Championships_in_Athletics,
pem_pos=0, final_scores=-0.2786131,
final_local_scores=-0.26579583, pem=-2.1703603,
similarity=7.9605923|
RECALL MISS: 5403315
FIS_Alpine_World_Ski_Championships_2009]
[winner: 35983 2009, pem_pos=0, final_scores=-1.0309379,
final_local_scores=-0.9375366, pem=-0.9296576,
similarity=-18.832626|
RECALL MISS: 5403315
FIS_Alpine_World_Ski_Championships_2009]
7: 5417172 FIS_Alpine_World_Ski_Championships_2011
[winner: 218926 IAAF_World_Championships_in_Athletics,
pem_pos=0, final_scores=-0.24308269,
final_local_scores=-0.22698423, pem=-2.1703603,
similarity=9.118808|
RECALL MISS: 5417172
FIS_Alpine_World_Ski_Championships_2011]
[winner: 36225 2011, pem_pos=0, final_scores=-1.2040131,
final_local_scores=-1.0990072, pem=-1.4232541,
similarity=-20.966978|
RECALL MISS: 5417172
FIS_Alpine_World_Ski_Championships_2011]
8: 28382662 2011_Alpine_Skiing_World_Cup
[winner: 36225 2011, pem_pos=0, final_scores=-1.0396045,
final_local_scores=-0.91941583, pem=-1.4232541,
similarity=-15.607612|
RECALL MISS: 28382662 2011_Alpine_Skiing_World_Cup]
9: 32863312 2012_Alpine_Skiing_World_Cup
[winner: 47374 2012, pem_pos=0, final_scores=-0.83148307,
final_local_scores=-0.63887495, pem=-1.4461981,
similarity=-7.1109405|
RECALL MISS: 32863312 2012_Alpine_Skiing_World_Cup]
10: 32863343 2013_Alpine_Skiing_World_Cup
[winner: 46945 2013, pem_pos=0, final_scores=-0.7904183,
final_local_scores=-0.61914134, pem=-1.490928,
similarity=-6.2788005|
RECALL MISS: 32863343 2013_Alpine_Skiing_World_Cup]
11: 225184 2002_Alpine_Skiing_World_Cup
[winner: 35502 2002, pem_pos=0, final_scores=-0.7271396,
final_local_scores=-0.590358, pem=-0.5897454,
similarity=-10.320621|
RECALL MISS: 225184 2002_Alpine_Skiing_World_Cup]

12: 4241148 2003_Alpine_Skiing_World_Cup
[winner: 36163 2003, pem_pos=0, final_scores=-0.6910986,
final_local_scores=-0.52357745, pem=-0.50449735,
similarity=-8.7913475]
RECALL MISS: 4241148 2003_Alpine_Skiing_World_Cup]

1: 26964606 Austria
[pem_pos=0, final_scores=0.45962054,
final_local_scores=0.5281544, pem=-0.25262418,
similarity=21.224703]

2: 3001568 Annaberg-Lungötz
[pem_pos=0, final_scores=0.17449598,
final_local_scores=0.30570287, pem=0.0, similarity=13.212496]

3: 4071698 Ted Ligety
[pem_pos=0, final_scores=0.6475063,
final_local_scores=0.7471384, pem=0.0, similarity=26.385813]

4: 601813 Stephan Eberharter
[pem_pos=0, final_scores=0.8910734,
final_local_scores=0.9834791, pem=0.0, similarity=33.43869]

1: 25418931 Marcel Hirscher
[pem_pos=0, final_scores=0.47293574,
final_local_scores=0.42460805, pem=0.0, similarity=16.760864]

2: 40335 Salzburg
[pem_pos=0, final_scores=0.3554343,
final_local_scores=0.44158256, pem=-0.11955855,
similarity=17.917595]

3: 25418931 Marcel Hirscher
[pem_pos=0, final_scores=0.38192552,
final_local_scores=0.32519323, pem=0.0, similarity=13.794127]

4: 25418931 Marcel Hirscher
[pem_pos=0, final_scores=0.42158204,
final_local_scores=0.36851192, pem=0.0, similarity=15.0868435]

5: 25418931 Marcel Hirscher
[pem_pos=0, final_scores=0.3320737,
final_local_scores=0.27073765, pem=0.0, similarity=12.1690645]

6: 25418931 Marcel Hirscher
[pem_pos=0, final_scores=0.29834646,
final_local_scores=0.23389578, pem=0.0, similarity=11.06963]

7: 25418931 Marcel Hirscher
[pem_pos=0, final_scores=0.44565326,
final_local_scores=0.3948061, pem=0.0, similarity=15.871514]

8: 25418931 Marcel Hirscher
[pem_pos=0, final_scores=0.41854858,
final_local_scores=0.3651983, pem=0.0, similarity=14.987959]

9: 630747 Adelboden
[pem_pos=0, final_scores=0.2266921,
final_local_scores=0.26835126, pem=0.0, similarity=12.09785]

10: 25418931 Marcel Hirscher
[pem_pos=0, final_scores=0.76087576,
final_local_scores=0.7391387, pem=0.0, similarity=26.147087]

11: 26964606 Austria
[pem_pos=0, final_scores=0.59440947,
final_local_scores=0.6753908, pem=-0.40674555,

similarity=26.456667]

Sample: b'Harvard_Crimson_men'

's _ lacrosse The [1 Harvard Crimson 1] men 's lacrosse team represents [1 1 Harvard University 1] in 1 National Collegiate Athletic Association -LRB- NCAA -RRB- 2 Division I men 's 3 lacrosse . Harvard competes as a member of the [2 2 Ivy League 2] and plays its home games at Cumnock Turf and 4 Harvard Stadium in [2 5 Cambridge 2] , [3 Massachusetts 3] . [4 Harvard 4] fielded its first lacrosse team in 1878 , and the following year , joined the United States National Amateur Lacrosse Association alongside [5 6 New York University 5] and nine club teams . In 1881 , [6 Harvard 6] defeated [7 7 Princeton 7] to win the first intercollegiate lacrosse tournament . In 1882 , the Intercollegiate Lacrosse Association was formed , and the following season also inducted the newly established [8 8 Yale 8] lacrosse team . [9 Harvard 9] and Princeton dominated the league throughout the 1880s , and the Crimson claimed the title in 1882 , 1885 , 1886 , and 1887 . The United States Intercollegiate Lacrosse League -LRB- USILL -RRB- was formed in 1905 , divided into a Northern Division and Southern Division . Championships were awarded in each division based on record and 9 strength of schedule . Harvard was named the Northern Division champions six consecutive seasons from 1908 to 1913 , and again in 1915 . In 1941 , Navy refused to play the integrated [10 Harvard 10] team , so its athletic director 10 ordered home its one black player rather than forfeit the game . The Crimson secured the [11 Ivy League 11] championship with the best league record in 1964 , 1980 , and 1990 . In 11 1971 , the [12 NCAA 12] established the 12 national championship tournament . [13 Harvard 13] made its first appearance in 13 1980 , when it lost in the first round to 14 Johns Hopkins , 16 -- 12 . The Crimson returned to the event in 15 1988 and were edge , 10 -- 9 , by 16 Navy in the opening round . In 17 1990 , [14 Harvard 14] won its first [15 NCAA 15] tournament game when it defeated 18 Notre Dame , 9 -- 3 . In the quarterfinals , the Crimson were beaten handily , 18 -- 3 , by 19 North Carolina . [16 Harvard 16] returned to the quarterfinals in 20 1996 , after beating [17 21 Hofstra 17] , 15 -- 12 , and then fell to eventual national runners-up 22 Virginia , 23 -- 12 . It was a decade before the Crimson again reached the tournament . They were beaten in the 23 2006 first round by 24 Syracuse

1: 18426501 Harvard_University

2: 14975 Ivy_League

1: 60706 National_Collegiate_Athletic_Association

[winner: 60706 National_Collegiate_Athletic_Association,

pem_pos=0, final_scores=-0.037866727,

final_local_scores=-0.040693447, pem=0.0,

similarity=2.8753393]

gt: 60706 National_Collegiate_Athletic_Association,

pem_pos=0, final_scores=-0.037866727,
final_local_scores=-0.040693447, pem=0.0,
similarity=2.8753393]
2: 669931 Division_I_(NCAA)
[winner: 669931 Division_I_(NCAA), pem_pos=1,
final_scores=-0.08607574, final_local_scores=-0.08958043,
pem=-0.7993879, similarity=5.7636495|
gt: 669931 Division_I_(NCAA), pem_pos=1,
final_scores=-0.08607574, final_local_scores=-0.08958043,
pem=-0.7993879, similarity=5.7636495]
3: 10448339 College_lacrosse
[winner: 22794284 Cornell_Big_Red_men's_lacrosse,
pem_pos=4, final_scores=-1.3831592,
final_local_scores=-1.4409854, pem=-1.6087462,
similarity=-30.163555|
gt: 10448339 College_lacrosse, pem_pos=7,
final_scores=-2.0274868, final_local_scores=-2.1875656,
pem=-3.987601, similarity=-39.506413]
4: 765891 Harvard_Stadium
[winner: 765891 Harvard_Stadium, pem_pos=0,
final_scores=-0.052112356, final_local_scores=-0.014474444,
pem=0.0, similarity=3.6577666|
gt: 765891 Harvard_Stadium, pem_pos=0,
final_scores=-0.052112356, final_local_scores=-0.014474444,
pem=0.0, similarity=3.6577666]
5: 5685 Cambridge,_Massachusetts
[winner: 1645518 Massachusetts, pem_pos=0,
final_scores=0.35943097, final_local_scores=0.3509206,
pem=-0.1423749, similarity=15.33614|
RECALL MISS: 5685 Cambridge,_Massachusetts]
[winner: 5685 Cambridge,_Massachusetts, pem_pos=2,
final_scores=0.25164872, final_local_scores=0.21609375,
pem=-1.848646, similarity=20.591606|
gt: 5685 Cambridge,_Massachusetts, pem_pos=2,
final_scores=0.25164872, final_local_scores=0.21609375,
pem=-1.848646, similarity=20.591606]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.6954796, final_local_scores=-1.6821685,
pem=-4.6728287, similarity=-20.69798|
RECALL MISS: 5685 Cambridge,_Massachusetts]
6: 11130928 NYU_Violets
[winner: 7954455 New_York_University, pem_pos=0,
final_scores=0.33247298, final_local_scores=0.28840986,
pem=0.0, similarity=12.696439|
gt: 11130928 NYU_Violets, pem_pos=1,
final_scores=-0.25579506, final_local_scores=-0.24873172,
pem=-1.5996877, similarity=5.366413]
7: 22860847 Princeton_Tigers_men's_lacrosse
[winner: 2204744 Michelle_Obama, pem_pos=0,
final_scores=0.13463996, final_local_scores=0.105921775,
pem=0.0, similarity=7.25063|
RECALL MISS: 22860847 Princeton_Tigers_men's_lacrosse]

8: 27511211 Yale_Bulldogs_men's_lacrosse
 [winner: 34273 Yale_University, pem_pos=0,
 final_scores=0.22784914, final_local_scores=0.16055393,
 pem=-0.3216871, similarity=10.630346]
 RECALL MISS: 27511211 Yale_Bulldogs_men's_lacrosse]
 [winner: 219560 Lacrosse_at_the_1904_Summer_Olympics,
 pem_pos=5, final_scores=-1.7405457,
 final_local_scores=-1.7519503, pem=-1.6087462,
 similarity=-39.443367]
 RECALL MISS: 27511211 Yale_Bulldogs_men's_lacrosse]

9: 15671039 Strength_of_schedule
 [winner: 277641 Strength_of_materials, pem_pos=3,
 final_scores=-1.3179563, final_local_scores=-1.2771249,
 pem=-1.1479203, similarity=-27.779669]
 RECALL MISS: 15671039 Strength_of_schedule]
 [winner: 16990186 UEFA_Euro_2008_schedule, pem_pos=15,
 final_scores=-1.5677845, final_local_scores=-1.4783802,
 pem=-4.296762, similarity=-16.661644]
 RECALL MISS: 15671039 Strength_of_schedule]

10: 23936857 1941_Harvard-Navy_lacrosse_game
 [winner: 18152723 Battle_of_Baxi, pem_pos=3,
 final_scores=-0.9593899, final_local_scores=-0.81798905,
 pem=-1.0986123, similarity=-14.346287]
 RECALL MISS: 23936857 1941_Harvard-Navy_lacrosse_game]
 [winner: 13254093 Irving_Langmuir_House, pem_pos=19,
 final_scores=-1.0189463, final_local_scores=-0.95438844,
 pem=-3.583519, similarity=-4.903408]
 RECALL MISS: 23936857 1941_Harvard-Navy_lacrosse_game]
 [winner: 273285
 Race_and_ethnicity_in_the_United_States_Census, pem_pos=0,
 final_scores=-1.15149, final_local_scores=-1.1156094,
 pem=-1.1573607, similarity=-22.908388]
 RECALL MISS: 23936857 1941_Harvard-Navy_lacrosse_game]
 [winner: 6112406 Clan_Home, pem_pos=10,
 final_scores=-1.376574, final_local_scores=-1.3665049,
 pem=-3.6764636, similarity=-16.696339]
 RECALL MISS: 23936857 1941_Harvard-Navy_lacrosse_game]
 [winner: 153914 Cassette_deck, pem_pos=8,
 final_scores=-1.6220677, final_local_scores=-1.6292841,
 pem=-2.561675, similarity=-30.600588]
 RECALL MISS: 23936857 1941_Harvard-Navy_lacrosse_game]

11: 10059979 1971_NCAA_Division_I_Men's_Lacrosse_Championship
 [winner: 34748 1971, pem_pos=0, final_scores=-0.9419217,
 final_local_scores=-0.8826219, pem=-0.6167389,
 similarity=-18.895563]
 RECALL MISS: 10059979
 1971_NCAA_Division_I_Men's_Lacrosse_Championship]

12: 714456 NCAA_Men's_Lacrosse_Championship
 [winner: 574966 W._P._McGee_Trophy, pem_pos=1,
 final_scores=-1.0495586, final_local_scores=-1.0136743,
 pem=-0.18232158, similarity=-25.168854]
 gt: 714456 NCAA_Men's_Lacrosse_Championship, pem_pos=0,

final_scores=-1.0546178, final_local_scores=-1.1138742,
 pem=0.0, similarity=-29.150517]
 13: 18160817 1980_NCAA_Division_I_Men's_Lacrosse_Championship
 [winner: 34640 1980, pem_pos=0, final_scores=-0.47251606,
 final_local_scores=-0.25803643, pem=-0.78379405,
 similarity=0.65178466]
 RECALL MISS: 18160817
 1980_NCAA_Division_I_Men's_Lacrosse_Championship]
 14: 16309951 Johns_Hopkins_Blue_Jays_men's_lacrosse
 [winner: 38420 Johns_Hopkins_University, pem_pos=1,
 final_scores=-0.12220353, final_local_scores=-0.20671722,
 pem=-0.82426006, similarity=2.4033127]
 gt: 16309951 Johns_Hopkins_Blue_Jays_men's_lacrosse,
 pem_pos=3, final_scores=-0.26492584,
 final_local_scores=-0.27074018, pem=-1.8312414,
 similarity=5.968861]
 15: 17045179 1988_NCAA_Division_I_Men's_Lacrosse_Championship
 [winner: 34670 1988, pem_pos=0, final_scores=-0.7764109,
 final_local_scores=-0.6700485, pem=-0.7886237,
 similarity=-11.617211]
 RECALL MISS: 17045179
 1988_NCAA_Division_I_Men's_Lacrosse_Championship]
 16: 22790903 Navy_Midshipmen_men's_lacrosse
 [winner: 20518076 United_States_Navy, pem_pos=0,
 final_scores=-0.28867096, final_local_scores=-0.31014276,
 pem=-1.0602508, similarity=0.6002432]
 RECALL MISS: 22790903 Navy_Midshipmen_men's_lacrosse]
 17: 11716277 1990_NCAA_Division_I_Men's_Lacrosse_Championship
 [winner: 27625098
 Pennsylvania_gubernatorial_election, 1990, pem_pos=2,
 final_scores=-0.75856185, final_local_scores=-0.6669851,
 pem=-1.0986123, similarity=-9.840028]
 RECALL MISS: 11716277
 1990_NCAA_Division_I_Men's_Lacrosse_Championship]
 18: 26454820 Notre_Dame_Fighting_Irish_men's_lacrosse
 [winner: 1245667 Notre_Dame_Fighting_Irish_football,
 pem_pos=1, final_scores=0.028561994,
 final_local_scores=0.048558593, pem=-1.3761536,
 similarity=13.022535]
 gt: 26454820 Notre_Dame_Fighting_Irish_men's_lacrosse,
 pem_pos=10, final_scores=-0.5420824,
 final_local_scores=-0.51660556, pem=-3.524415,
 similarity=7.8394895]
 19: 22807864 North_Carolina_Tar_Heels_men's_lacrosse
 [winner: 8252650
 North_Carolina_Tar_Heels_men's_basketball, pem_pos=3,
 final_scores=-0.30280453, final_local_scores=-0.3060665,
 pem=-3.8903022, similarity=16.112146]
 RECALL MISS: 22807864
 North_Carolina_Tar_Heels_men's_lacrosse]
 20: 18160656 1996_NCAA_Division_I_Men's_Lacrosse_Championship
 [winner: 34636 1996, pem_pos=0, final_scores=-0.7404882,

final_local_scores=-0.5711659, pem=-0.78063476,
similarity=-8.709804|
RECALL MISS: 18160656
1996_NCAA_Division_I_Men's_Lacrosse_Championship]
21: 27301066 Hofstra_Pride_men's_lacrosse
[winner: 418091 Hofstra_University, pem_pos=0,
final_scores=0.06224282, final_local_scores=0.016356781,
pem=0.0, similarity=4.5778317|
gt: 27301066 Hofstra_Pride_men's_lacrosse, pem_pos=1,
final_scores=-0.16440812, final_local_scores=-0.083662316,
pem=-1.247153, similarity=8.375275]
22: 22860850 Virginia_Cavaliers_men's_lacrosse
[winner: 32432 Virginia, pem_pos=0,
final_scores=-0.047424838, final_local_scores=-0.06628968,
pem=-0.2023721, similarity=3.2120278|
RECALL MISS: 22860850 Virginia_Cavaliers_men's_lacrosse]
23: 11123672 2006_NCAA_Division_I_Men's_Lacrosse_Championship
[winner: 36164 2006, pem_pos=0, final_scores=-0.32036084,
final_local_scores=-0.15954678, pem=-0.3060146,
similarity=0.99267375|
RECALL MISS: 11123672
2006_NCAA_Division_I_Men's_Lacrosse_Championship]
24: 18292722 Syracuse_Orange_men's_lacrosse
[winner: 10189860 Syracuse_Orange_men's_basketball,
pem_pos=4, final_scores=-0.24342191,
final_local_scores=-0.22062832, pem=-2.744889,
similarity=12.4328575|
gt: 18292722 Syracuse_Orange_men's_lacrosse, pem_pos=12,
final_scores=-0.52526027, final_local_scores=-0.5293158,
pem=-3.746009, similarity=8.665253]
1: 18426501 Harvard_University
[pem_pos=0, final_scores=0.23137508,
final_local_scores=0.16679764, pem=0.0, similarity=9.067287]
2: 14975 Ivy_League
[pem_pos=0, final_scores=0.118957266,
final_local_scores=0.093854666, pem=-0.013410744,
similarity=6.9634533]
1: 18426501 Harvard_University
[pem_pos=0, final_scores=0.26853025,
final_local_scores=0.20738402, pem=-0.040016375,
similarity=10.49608]
2: 5685 Cambridge_Massachusetts
[pem_pos=2, final_scores=0.25164872,
final_local_scores=0.21609375, pem=-1.848646,
similarity=20.591606]
3: 1645518 Massachusetts
[pem_pos=0, final_scores=0.35943097,
final_local_scores=0.3509206, pem=-0.1423749,
similarity=15.33614]
4: 18426501 Harvard_University
[pem_pos=1, final_scores=0.1422645,
final_local_scores=0.065963015, pem=-0.16400388,

similarity=6.9500594]
 5: 7954455 New_York_University
 [pem_pos=0, final_scores=0.33247298,
 final_local_scores=0.28840986, pem=0.0, similarity=12.696439]
 6: 18426501 Harvard_University
 [pem_pos=1, final_scores=0.20901941,
 final_local_scores=0.1388826, pem=-0.16400388,
 similarity=9.126125]
 7: 2204744 Michelle_Obama
 [pem_pos=0, final_scores=0.13463996,
 final_local_scores=0.105921775, pem=0.0, similarity=7.25063]
 8: 34273 Yale_University
 [pem_pos=0, final_scores=0.22784914,
 final_local_scores=0.16055393, pem=-0.3216871,
 similarity=10.630346]
 9: 18426501 Harvard_University
 [pem_pos=0, final_scores=0.07590769,
 final_local_scores=-0.006990671, pem=-0.16400388,
 similarity=4.772976]
 10: 18426501 Harvard_University
 [pem_pos=1, final_scores=0.3016219,
 final_local_scores=0.24003676, pem=-0.16400388,
 similarity=12.144767]
 11: 14975 Ivy_League
 [pem_pos=0, final_scores=0.14004712,
 final_local_scores=0.1168921, pem=-0.013410744,
 similarity=7.650936]
 12: 60706 National_Collegiate_Athletic_Association
 [pem_pos=0, final_scores=0.03511603,
 final_local_scores=0.05059287, pem=-0.19246244,
 similarity=6.646146]
 13: 2204744 Michelle_Obama
 [pem_pos=0, final_scores=0.14671786,
 final_local_scores=0.119115025, pem=0.0, similarity=7.6443424]
 14: 2204744 Michelle_Obama
 [pem_pos=0, final_scores=0.12002833,
 final_local_scores=0.08659473, pem=0.0, similarity=6.6738725]
 15: 60706 National_Collegiate_Athletic_Association
 [pem_pos=0, final_scores=0.05356209,
 final_local_scores=0.07074238, pem=-0.19246244,
 similarity=7.2474475]
 16: 2204744 Michelle_Obama
 [pem_pos=0, final_scores=0.16714434,
 final_local_scores=0.14142784, pem=0.0, similarity=8.310202]
 17: 418091 Hofstra_University
 [pem_pos=0, final_scores=0.06224282,
 final_local_scores=0.016356781, pem=0.0, similarity=4.5778317]

Sample: b'Inspectorates-General'

_ -LRB- Turkey -RRB- Inspectorates-General or General
 Inspectorates -LRB- 1 Turkish : Umumi Mufettişlikler -RRB-
 was a regional governorship whose authorities prevailed over

civilian , military and judicial institutions under their domain of the direct command of [1 1 Mustafa Kemal 1] in order to establishing 2 authoritarian rule and to consolidate the authority in the process of 3 Turkification of religious and ethnic minorities . The got the law numbered 1164 and dated June 25 , 1927 passed . On January 1 , 1928 , the First Inspectorate-General -LRB- Birinci Umumi Müfettişlik -RRB- including [1 4 Diyarbakır 1] , [2 2 Elâzığ 2] , [2 5 Urfa 2] , [3 6 Bitlis 3] , [4 7 Van 4] , , [5 8 Siirt 5] ve [6 9 Mardin 6] provinces , was established with the center in [3 3 Diyarbakır 3] . Before , on February 19 , 1934 , the Second Inspectorate-General -LRB- İkinci Umumi Müfettişlik -RRB- including [4 4 Kırklareli 4] , [7 10 Edirne 7] , [5 5 Tekirdağ 5] and [8 11 Çanakkale 8] provinces , was established with the center in [6 6 Edirne 6] On August 25 , 1935 , the Third Inspectorate-General -LRB- Üçüncü Umumi Müfettişlik -RRB- including [7 7 Ağrı 7] , [9 12 Kars 9] , [10 13 Artvin 10] , [11 14 Rize 11] , [12 15 Trabzon 12] , [8 8 Gümüşhane 8] , [13 1 6 Erzincan 13] and [14 17 Erzurum 14] provinces , was established with the center in [9 9 Erzurum 9] . Before [10 10 Dersim Rebellion 10] , on June 6 . 1936 , the Fourth Inspectorate-General -LRB- Dördüncü Umumi Müfettişlik -RRB- was established in historical [15 Dersim 15] region that includes [11 11 Tunceli 11] , [16 18 Elazığ 16] and [12 12 Bingöl 12] provinces , with the center in

1: 1329090 Mustafa_Kemal_Atatürk

2: 1095357 Elazığ_Province

3: 642286 Diyarbakır

4: 1704047 Kırklareli_Province

5: 1949365 Tekirdağ_Province

6: 216887 Edirne

7: 1364160 Ağrı_Province

8: 2150286 Gümüşhane_Province

9: 966195 Erzurum

10: 9727104 Dersim_Rebellion

11: 291275 Tunceli_Province

12: 1704063 Bingöl_Province

1: 29992 Turkish_language

[winner: 29992 Turkish_language, pem_pos=0, final_scores=-0.24538182, final_local_scores=-0.30427733, pem=-0.7274845, similarity=-1.0343554]

gt: 29992 Turkish_language, pem_pos=0, final_scores=-0.24538182, final_local_scores=-0.30427733, pem=-0.7274845, similarity=-1.0343554]

2: 21347657 Authoritarianism

[winner: 2275708 Authoritarian_personality, pem_pos=1, final_scores=-1.109407, final_local_scores=-1.0339104, pem=-0.6798454, similarity=-23.067131]

gt: 21347657 Authoritarianism, pem_pos=0, final_scores=-1.3016953, final_local_scores=-1.3383539, pem=-0.041395877, similarity=-35.624317]

[winner: 252311 Rule_of_inference, pem_pos=3,

final_scores=-1.3332279, final_local_scores=-1.2771263,
pem=-2.4380946, similarity=-20.763542|
RECALL MISS: 21347657 Authoritarianism]

3: 4445638 Turkification
[winner: 4445638 Turkification, pem_pos=0,
final_scores=-0.18936253, final_local_scores=-0.16600986,
pem=0.0, similarity=-0.8643524|
gt: 4445638 Turkification, pem_pos=0,
final_scores=-0.18936253, final_local_scores=-0.16600986,
pem=0.0, similarity=-0.8643524]

4: 793511 Diyarbakır_Province
[winner: 642286 Diyarbakır, pem_pos=0,
final_scores=0.10244112, final_local_scores=0.039389014,
pem=-0.50045717, similarity=7.9867234|
gt: 793511 Diyarbakır_Province, pem_pos=1,
final_scores=-0.0040014237, final_local_scores=-0.049801767,
pem=-0.95192075, similarity=7.780222]

5: 1353138 Şanlıurfa_Province
[winner: 285533 Şanlıurfa, pem_pos=0,
final_scores=0.17359321, final_local_scores=0.15733999,
pem=0.0, similarity=8.785052|
gt: 1353138 Şanlıurfa_Province, pem_pos=1,
final_scores=-0.0062499493, final_local_scores=-0.040904813,
pem=-0.65073806, similarity=6.4078465]

6: 1704064 Bitlis_Province
[winner: 2072949 Bitlis, pem_pos=0,
final_scores=0.15821104, final_local_scores=0.12917459,
pem=-0.1863189, similarity=8.957771|
gt: 1704064 Bitlis_Province, pem_pos=1,
final_scores=0.04264514, final_local_scores=0.03740084,
pem=-0.3003207, similarity=6.8390193]

7: 757618 Van_Province
[winner: 63685 Van, Turkey, pem_pos=3,
final_scores=0.04724835, final_local_scores=-0.02249825,
pem=-3.4721067, similarity=22.300177|
gt: 757618 Van_Province, pem_pos=1,
final_scores=-0.0030261725, final_local_scores=-0.058238268,
pem=-1.6961808, similarity=11.575862]

8: 3055567 Siirt_Province
[winner: 3055450 Siirt, pem_pos=0,
final_scores=0.1488014, final_local_scores=0.12621251,
pem=0.0, similarity=7.856146|
gt: 3055567 Siirt_Province, pem_pos=1,
final_scores=-0.13618225, final_local_scores=-0.17277032,
pem=-0.66207296, similarity=2.5343575]

9: 1446282 Mardin_Province
[winner: 2308621 Mardin, pem_pos=0,
final_scores=0.15583389, final_local_scores=0.12365851,
pem=-0.46888694, similarity=10.329809|
gt: 1446282 Mardin_Province, pem_pos=1,
final_scores=-0.043165997, final_local_scores=-0.07628902,
pem=-0.5370033, similarity=4.7334037]

10: 890427 Edirne_Province
[winner: 216887 Edirne, pem_pos=0,
final_scores=0.32514852, final_local_scores=0.2841229,
pem=0.0, similarity=12.568507|
gt: 890427 Edirne_Province, pem_pos=3,
final_scores=-0.17586607, final_local_scores=-0.21005167,
pem=-1.4105282, similarity=5.4920235]
11: 886788 Çanakkale_Province
[winner: 1871641 Çanakkale, pem_pos=0,
final_scores=0.057140514, final_local_scores=0.047175214,
pem=-0.64046264, similarity=8.980451|
gt: 886788 Çanakkale_Province, pem_pos=1,
final_scores=-0.10043143, final_local_scores=-0.11013053,
pem=-0.7487625, similarity=4.8750854]
12: 886782 Kars_Province
[winner: 1120351 Kars, pem_pos=0,
final_scores=0.17618866, final_local_scores=0.1250161,
pem=-0.43661144, similarity=10.194803|
gt: 886782 Kars_Province, pem_pos=1,
final_scores=-0.118478626, final_local_scores=-0.17134564,
pem=-0.91911316, similarity=3.9746976]
13: 1315356 Artvin_Province
[winner: 1887344 Artvin, pem_pos=0,
final_scores=0.15891986, final_local_scores=0.13186854,
pem=0.0, similarity=8.024933|
gt: 1315356 Artvin_Province, pem_pos=1,
final_scores=-0.048752055, final_local_scores=-0.051708035,
pem=-0.54772675, similarity=5.525265]
14: 886778 Rize_Province
[winner: 1331862 Rize, pem_pos=0,
final_scores=0.15068446, final_local_scores=0.13819817,
pem=-0.586935, similarity=11.405665|
gt: 886778 Rize_Province, pem_pos=1,
final_scores=0.017603293, final_local_scores=-0.0077110827,
pem=-0.6304818, similarity=7.288257]
15: 1000368 Trabzon_Province
[winner: 62653 Trabzon, pem_pos=0,
final_scores=0.29488534, final_local_scores=0.2509829,
pem=-0.001391186, similarity=11.58711|
gt: 1000368 Trabzon_Province, pem_pos=4,
final_scores=-0.04370834, final_local_scores=-0.07962334,
pem=-1.47423, similarity=9.730686]
16: 1796598 Erzincan_Province
[winner: 1210336 Erzincan, pem_pos=0,
final_scores=0.28454977, final_local_scores=0.23585811,
pem=-0.25802472, similarity=12.531367|
gt: 1796598 Erzincan_Province, pem_pos=1,
final_scores=-0.0048106015, final_local_scores=-0.036850065,
pem=-0.84438527, similarity=7.5819316]
17: 902883 Erzurum_Province
[winner: 966195 Erzurum, pem_pos=0,
final_scores=0.3245036, final_local_scores=0.28674048,

pem=0.0, similarity=12.646621|
gt: 902883 Erzurum_Province, pem_pos=3,
final_scores=0.034767523, final_local_scores=-0.017124727,
pem=-1.2874299, similarity=10.579921]
18: 1095357 Elazığ_Province
[winner: 2341852 Elâzığ, pem_pos=0,
final_scores=0.19994922, final_local_scores=0.14560014,
pem=-0.25418612, similarity=9.817016|
gt: 1095357 Elazığ_Province, pem_pos=1,
final_scores=-0.112709895, final_local_scores=-0.15002352,
pem=-1.4940909, similarity=7.73781]
1: 1329090 Mustafa_Kemal_Atatürk
[pem_pos=0, final_scores=0.5257526,
final_local_scores=0.5394385, pem=0.0, similarity=20.187634]
2: 1095357 Elazığ_Province
[pem_pos=0, final_scores=0.07215391,
final_local_scores=0.058705166, pem=-0.20110513,
similarity=6.935233]
3: 642286 Diyarbakır
[pem_pos=0, final_scores=0.17619868,
final_local_scores=0.119957924, pem=-0.50045717,
similarity=10.391061]
4: 1704047 Kırklareli_Province
[pem_pos=0, final_scores=0.13989548,
final_local_scores=0.14979759, pem=-0.4182041,
similarity=10.83423]
5: 1949365 Tekirdağ_Province
[pem_pos=0, final_scores=0.056764945,
final_local_scores=0.04218757, pem=-0.47348717,
similarity=7.923571]
6: 216887 Edirne
[pem_pos=0, final_scores=0.34513003,
final_local_scores=0.30594972, pem=0.0, similarity=13.219863]
7: 1364160 Ağrı_Province
[pem_pos=0, final_scores=0.122477904,
final_local_scores=0.09101176, pem=-0.29364827,
similarity=8.402591]
8: 2150286 Gümüşhane_Province
[pem_pos=0, final_scores=0.16083859,
final_local_scores=0.1802974, pem=0.0, similarity=9.470147]
9: 966195 Erzurum
[pem_pos=0, final_scores=0.40403754,
final_local_scores=0.3730464, pem=0.0, similarity=15.222161]
10: 9727104 Dersim_Rebellion
[pem_pos=0, final_scores=0.09174828,
final_local_scores=0.107836396, pem=0.0, similarity=7.307766]
11: 291275 Tunceli_Province
[pem_pos=0, final_scores=0.31824994,
final_local_scores=0.29216254, pem=0.0, similarity=12.808426]
12: 1704063 Bingöl_Province
[pem_pos=0, final_scores=0.15175872,
final_local_scores=0.13939637, pem=-0.44590068,

similarity=10.674457]

1: 642286 Diyarbakır

[pem_pos=0, final_scores=0.10244112,
final_local_scores=0.039389014, pem=-0.50045717,
similarity=7.9867234]

2: 285533 Şanlıurfa

[pem_pos=0, final_scores=0.17359321,
final_local_scores=0.15733999, pem=0.0, similarity=8.785052]

3: 2072949 Bitlis

[pem_pos=0, final_scores=0.15821104,
final_local_scores=0.12917459, pem=-0.1863189,
similarity=8.957771]

4: 63685 Van, Turkey

[pem_pos=3, final_scores=0.04724835,
final_local_scores=-0.02249825, pem=-3.4721067,
similarity=22.300177]

5: 3055450 Siirt

[pem_pos=0, final_scores=0.1488014,
final_local_scores=0.12621251, pem=0.0, similarity=7.856146]

6: 2308621 Mardin

[pem_pos=0, final_scores=0.15583389,
final_local_scores=0.12365851, pem=-0.46888694,
similarity=10.329809]

7: 216887 Edirne

[pem_pos=0, final_scores=0.32514852,
final_local_scores=0.2841229, pem=0.0, similarity=12.568507]

8: 1871641 Çanakkale

[pem_pos=0, final_scores=0.057140514,
final_local_scores=0.047175214, pem=-0.64046264,
similarity=8.980451]

9: 1120351 Kars

[pem_pos=0, final_scores=0.17618866,
final_local_scores=0.1250161, pem=-0.43661144,
similarity=10.194803]

10: 1887344 Artvin

[pem_pos=0, final_scores=0.15891986,
final_local_scores=0.13186854, pem=0.0, similarity=8.024933]

11: 1331862 Rize

[pem_pos=0, final_scores=0.15068446,
final_local_scores=0.13819817, pem=-0.586935,
similarity=11.405665]

12: 62653 Trabzon

[pem_pos=0, final_scores=0.29488534,
final_local_scores=0.2509829, pem=-0.001391186,
similarity=11.58711]

13: 1210336 Erzincan

[pem_pos=0, final_scores=0.28454977,
final_local_scores=0.23585811, pem=-0.25802472,
similarity=12.531367]

14: 966195 Erzurum

[pem_pos=0, final_scores=0.3245036,
final_local_scores=0.28674048, pem=0.0, similarity=12.646621]

15: 291275 Tunceli_Province
[pem_pos=0, final_scores=0.30519134,
final_local_scores=0.27789798, pem=0.0, similarity=12.382743]
16: 2341852 Elâzığ
[pem_pos=0, final_scores=0.19994922,
final_local_scores=0.14560014, pem=-0.25418612,
similarity=9.817016]

Sample: b'Charles_Fitzgerald'

-LRB- rugby -RRB- Charles James Fitzgerald -LRB- born 6
June 1899 in 1 Fairhall -- died 8 May 1961 in -RRB- was a [1 2
New Zealand 1] 3 rugby football er who represented [2 New
Zealand 2] in 4 rugby union and 5 rugby league . Fitzgerald
was a utility back in both codes . [3 Fitzgerald 3]
represented 6 Marlborough between 1917 and 1921 , appearing
for the union in every backline position except fullback . He
played for a combined [4 7 Nelson 4] - 8 Marlborough - [1 1 Go
lden Bay-Motueka 1] side against the 1921 [2 2 Springboks 2]
kicking a penalty goal in the side 's 26 -- 3 defeat . In
that game he lined up alongside fellow future dual-
international 9 Jim O'Brien . Fitzgerald played for the in
1922 , scoring a try in the inter-island match . Following
this he was selected for [3 3 New Zealand 3] , playing in
five matches on a tour of [5 New Zealand 5] and [6 Australia 6
] . [7 Fitzgerald 7] represented [4 4 New Zealand 4] -LRB- RU
-RRB- in the 12 -- 11 victory over 10 Wairarapa at [5 5 Carte
rton 5] on 19 July 1922 , the 26 -- 19 victory over [8 11 New
South Wales 8] at [9 Sydney 9] on 29 July 1922 , the 24 -- 6
victory over Metropolitan Union at [10 Sydney 10] on 2 August
1922 , the 6 -- 8 defeat to [11 12 New South Wales 11] at [12
Sydney 12] on 7 August 1922 , and scored a 13 penalty in the
45 -- 11 victory over [6 6 Manawatu 6] - 14 Wellington XV at [7
7 Palmerston North 7] on 16 August 1922 but never played a
Test match for his country . In 1923 [13 Fitzgerald 13] moved
to [8 8 Christchurch 8] , joining the Marist Old Boys club .
However the club was expelled from the in April 1924 and [14
Fitzgerald 14] followed to club to 15 rugby league . [15
Fitzgerald 15] excelled in rugby league and in his debut year
he made the national side , playing in two [16 Test 16]
matches against . He played for the [17 16 South Island 17]

1: 18961210 Golden_Bay-Motueka_Rugby_Union
2: 376115 South_Africa_national_rugby_union_team
3: 154957 New_Zealand_national_rugby_union_team
4: 154957 New_Zealand_national_rugby_union_team
5: 594856 Carterton, New_Zealand
6: 6091534 Manawatu_Rugby_Union
7: 314133 Palmerston_North
8: 628782 Christchurch

1: 17709857 Fairhall
[winner: 17709857 Fairhall, pem_pos=0,
final_scores=-0.082776606, final_local_scores=0.005394131,
pem=0.0, similarity=4.2506847]

gt: 17709857 Fairhall, pem_pos=0,
final_scores=-0.082776606, final_local_scores=0.005394131,
pem=0.0, similarity=4.2506847]
2: 1286042 New_Zealanders
[winner: 154957 New_Zealand_national_rugby_union_team,
pem_pos=4, final_scores=0.31596237,
final_local_scores=0.26455736, pem=-3.8192244,
similarity=32.754166]
RECALL MISS: 1286042 New_Zealanders]
3: 25402 Rugby_football
[winner: 19774201 Rugby_Virginia, pem_pos=7,
final_scores=-0.8311649, final_local_scores=-0.7136388,
pem=-2.956099, similarity=-1.1309667]
gt: 25402 Rugby_football, pem_pos=1,
final_scores=-2.3732927, final_local_scores=-2.6307664,
pem=-1.8082088, similarity=-64.58428]
[winner: 10568 Association_football, pem_pos=0,
final_scores=-1.7479732, final_local_scores=-1.8593895,
pem=-0.33534196, similarity=-49.574535]
gt: 25402 Rugby_football, pem_pos=16,
final_scores=-2.4302135, final_local_scores=-2.6929438,
pem=-4.722696, similarity=-50.590343]
4: 154957 New_Zealand_national_rugby_union_team
[winner: 19774201 Rugby_Virginia, pem_pos=7,
final_scores=-0.7677432, final_local_scores=-0.6443602,
pem=-2.956099, similarity=0.9364461]
RECALL MISS: 154957
New_Zealand_national_rugby_union_team]
[winner: 481130 Union_(American_Civil_War), pem_pos=0,
final_scores=-1.3217024, final_local_scores=-1.3762803,
pem=-1.132797, similarity=-30.82091]
RECALL MISS: 154957
New_Zealand_national_rugby_union_team]
5: 952064 New_Zealand_national_rugby_league_team
[winner: 19774201 Rugby_Virginia, pem_pos=7,
final_scores=-0.7988952, final_local_scores=-0.6783891,
pem=-2.956099, similarity=-0.079044506]
RECALL MISS: 952064
New_Zealand_national_rugby_league_team]
[winner: 1424274 League_(unit), pem_pos=2,
final_scores=-1.7461228, final_local_scores=-1.8073611,
pem=-1.8401146, similarity=-39.838715]
RECALL MISS: 952064
New_Zealand_national_rugby_league_team]
6: 6875302 Marlborough_Rugby_Union
[winner: 6875302 Marlborough_Rugby_Union, pem_pos=13,
final_scores=-0.20493057, final_local_scores=-0.22337133,
pem=-3.0740142, similarity=14.140834]
gt: 6875302 Marlborough_Rugby_Union, pem_pos=13,
final_scores=-0.20493057, final_local_scores=-0.22337133,
pem=-3.0740142, similarity=14.140834]
7: 18961061 Nelson_Rugby_Union

[winner: 61978 Nelson, New Zealand, pem_pos=0,
final_scores=0.06919144, final_local_scores=0.029353023,
pem=-1.7804265, similarity=14.6479025]
RECALL MISS: 18961061 Nelson_Rugby_Union]

8: 6875302 Marlborough_Rugby_Union
[winner: 6875302 Marlborough_Rugby_Union, pem_pos=13,
final_scores=-0.21609189, final_local_scores=-0.23556337,
pem=-3.0740142, similarity=13.776999]
gt: 6875302 Marlborough_Rugby_Union, pem_pos=13,
final_scores=-0.21609189, final_local_scores=-0.23556337,
pem=-3.0740142, similarity=13.776999]

9: 31928300 Andrew_O'Brien_(rugby)
[winner: 2887800 Jim_O'Brien_(basketball, born_1952),
pem_pos=0, final_scores=-0.83819705,
final_local_scores=-0.8102898, pem=-0.7657296,
similarity=-15.926795]
gt: 31928300 Andrew_O'Brien_(rugby), pem_pos=7,
final_scores=-1.1508534, final_local_scores=-1.2367492,
pem=-3.3065815, similarity=-14.835646]

10: 25952690 Wairarapa_Rugby_Football_Union
[winner: 6585848 Wairarapa_Bush_Rugby_Football_Union,
pem_pos=3, final_scores=0.0112879425,
final_local_scores=0.031666785, pem=-1.5077807,
similarity=13.234259]
gt: 25952690 Wairarapa_Rugby_Football_Union, pem_pos=6,
final_scores=-0.73986286, final_local_scores=-0.78331596,
pem=-4.9894605, similarity=7.8474646]

11: 4044378 New_South_Wales_Rugby_Union
[winner: 781440 New_South_Wales_Waratahs, pem_pos=3,
final_scores=0.21005742, final_local_scores=0.17226511,
pem=-2.886913, similarity=24.92993]
RECALL MISS: 4044378 New_South_Wales_Rugby_Union]

12: 4044378 New_South_Wales_Rugby_Union
[winner: 781440 New_South_Wales_Waratahs, pem_pos=3,
final_scores=0.22055627, final_local_scores=0.18373352,
pem=-2.886913, similarity=25.27217]
RECALL MISS: 4044378 New_South_Wales_Rugby_Union]

13: 19992219 Penalty_(rugby_union)
[winner: 2112397 Sanctions_(law), pem_pos=4,
final_scores=-1.2078, final_local_scores=-1.1598709,
pem=-1.6643215, similarity=-21.472307]
gt: 19992219 Penalty_(rugby_union), pem_pos=12,
final_scores=-1.7417275, final_local_scores=-1.8365602,
pem=-5.2298737, similarity=-22.276041]

14: 5783293 Wellington_Rugby_Football_Union
[winner: 5783293 Wellington_Rugby_Football_Union,
pem_pos=3, final_scores=-0.025785148,
final_local_scores=-0.08594096, pem=-3.5096571,
similarity=20.611124]
gt: 5783293 Wellington_Rugby_Football_Union, pem_pos=3,
final_scores=-0.025785148, final_local_scores=-0.08594096,
pem=-3.5096571, similarity=20.611124]

15: 25735 Rugby_league
 [winner: 19774201 Rugby_Virginia, pem_pos=7,
 final_scores=-0.81262124, final_local_scores=-0.6933827,
 pem=-2.956099, similarity=-0.5264845]
 gt: 25735 Rugby_league, pem_pos=14,
 final_scores=-1.70133, final_local_scores=-1.9029279,
 pem=-3.3002021, similarity=-34.75044]
 [winner: 1424274 League_(unit), pem_pos=2,
 final_scores=-1.63618, final_local_scores=-1.6872654,
 pem=-1.8401146, similarity=-36.254818]
 gt: 25735 Rugby_league, pem_pos=5,
 final_scores=-1.6523713, final_local_scores=-1.849448,
 pem=-3.5378149, similarity=-31.862316]

16: 28705488 South_Island_rugby_league_team
 [winner: 62049 South_Island, pem_pos=0,
 final_scores=0.15590273, final_local_scores=0.15683869,
 pem=0.0, similarity=8.770092]
 gt: 28705488 South_Island_rugby_league_team, pem_pos=8,
 final_scores=-0.45853245, final_local_scores=-0.5238575,
 pem=-4.052085, similarity=10.492628]

1: 18961210 Golden_Bay-Motueka_Rugby_Union
 [pem_pos=0, final_scores=0.17988746,
 final_local_scores=0.26625112, pem=0.0, similarity=12.035177]

2: 376115 South_Africa_national_rugby_union_team
 [pem_pos=0, final_scores=0.42525822,
 final_local_scores=0.41760468, pem=0.0, similarity=16.55187]

3: 154957 New_Zealand_national_rugby_union_team
 [pem_pos=4, final_scores=0.42843753,
 final_local_scores=0.38819563, pem=-3.8192244,
 similarity=36.44378]

4: 154957 New_Zealand_national_rugby_union_team
 [pem_pos=4, final_scores=0.38601863,
 final_local_scores=0.34185946, pem=-3.8192244,
 similarity=35.061016]

5: 594856 Carterton_New_Zealand
 [pem_pos=0, final_scores=0.046052262,
 final_local_scores=0.062566385, pem=-0.5388591,
 similarity=8.887218]

6: 6091534 Manawatu_Rugby_Union
 [pem_pos=0, final_scores=0.3238443,
 final_local_scores=0.31638712, pem=-0.703368,
 similarity=17.356361]

7: 314133 Palmerston_North
 [pem_pos=0, final_scores=0.44545162,
 final_local_scores=0.42592198, pem=0.0, similarity=16.800074]

8: 628782 Christchurch
 [pem_pos=0, final_scores=0.48415244,
 final_local_scores=0.4633217, pem=-0.16020472,
 similarity=18.787374]

1: 154957 New_Zealand_national_rugby_union_team
 [pem_pos=4, final_scores=0.31596237,
 final_local_scores=0.26455736, pem=-3.8192244,

similarity=32.754166]
2: 154957 New_Zealand_national_rugby_union_team
[pem_pos=4, final_scores=0.4820587,
final_local_scores=0.44676864, pem=-3.8192244,
similarity=38.191715]
3: 1010867 James_FitzGerald
[pem_pos=0, final_scores=0.038304374,
final_local_scores=0.042710662, pem=-0.26905325,
similarity=6.8274374]
4: 61978 Nelson,_New_Zealand
[pem_pos=0, final_scores=0.06919144,
final_local_scores=0.029353023, pem=-1.7804265,
similarity=14.6479025]
5: 4913064 New_Zealand
[pem_pos=0, final_scores=0.30089432,
final_local_scores=0.28167453, pem=-0.23795462,
similarity=13.789477]
6: 4689264 Australia
[pem_pos=0, final_scores=0.3792612,
final_local_scores=0.3884784, pem=-0.17168477,
similarity=16.616333]
7: 1010867 James_FitzGerald
[pem_pos=0, final_scores=0.15492277,
final_local_scores=0.17009851, pem=-0.26905325,
similarity=10.628945]
8: 781440 New_South_Wales_Waratahs
[pem_pos=3, final_scores=0.21005742,
final_local_scores=0.17226511, pem=-2.886913,
similarity=24.92993]
9: 27862 Sydney
[pem_pos=0, final_scores=0.4749558,
final_local_scores=0.48305297, pem=-0.1443489,
similarity=19.28997]
10: 27862 Sydney
[pem_pos=0, final_scores=0.38673472,
final_local_scores=0.38668483, pem=-0.1443489,
similarity=16.414152]
11: 781440 New_South_Wales_Waratahs
[pem_pos=3, final_scores=0.22055627,
final_local_scores=0.18373352, pem=-2.886913,
similarity=25.27217]
12: 27862 Sydney
[pem_pos=0, final_scores=0.40355575,
final_local_scores=0.40505928, pem=-0.1443489,
similarity=16.962482]
13: 1010867 James_FitzGerald
[pem_pos=0, final_scores=0.08617948,
final_local_scores=0.09328455, pem=-0.26905325,
similarity=8.336663]
14: 1010867 James_FitzGerald
[pem_pos=0, final_scores=0.12629335,
final_local_scores=0.13710287, pem=-0.26905325,


```

similarity=9.644289]
15: 1010867 James_FitzGerald
    [pem_pos=0, final_scores=0.24732648,
final_local_scores=0.27103555, pem=-0.26905325,
similarity=13.641108]
16: 30863 Test_cricket
    [pem_pos=0, final_scores=0.031438425,
final_local_scores=0.028040782, pem=-0.23950954,
similarity=6.228996]
17: 62049 South_Island
    [pem_pos=0, final_scores=0.15590273,
final_local_scores=0.15683869, pem=0.0, similarity=8.770092]

```

Sample: b'Energy_in_Sudan'

Energy in [1 Sudan 1] describes and 1 electricity production , consumption and imports in [1 1 Sudan 1] . [2 Sudan 2] is a net energy exporter . Primary energy use in [3 Sudan 3] was 179 2 kWh and 4 kWh per million persons in 2008 . The world share of energy production in [4 3 Africa 4] was 12 percent of oil and 7 percent of gas in 2009 . In 2010 , major energy producers in [5 Africa 5] were [6 4 Algeria 6] , [7 5 Angola 7] , [8 6 Cameroon 8] , [9 7 Democratic Republic of the Congo 9] , [10 8 Equatorial Guinea 10] , [11 9 Gabon 11] , [12 10 Libya 12] , [13 11 Nigeria 13] and [14 Sudan 14] . According to the [2 2 OECD 2] and the [3 3 World Bank 3] , the 12 population growth of from 2004 to 2008 was 16.4 percent -LRB- in comparison to the world average of 5.3 percent , [15 13 India 15] at 5.6 percent and [16 14 Nigeria 16] at 17.6 percent -RRB- . According to [4 4 Transparency International 4] , companies operating in [17 Sudan 17] -LRB- based on reporting in 2009 -RRB- included four national companies : [18 China National Petroleum Corporation 18] -LRB- -RRB- , [19 Kuwait Petroleum Corporation 19] -LRB- -RRB- , the Oil and Natural Gas Corporation -LRB- , the [20 Indian 20] 15 NOC -RRB- and 16 Petronas

1: 27421 Sudan

2: 33853117 Organisation_for_Economic_Co-operation_and_Development

3: 33309 World_Bank

4: 66241 Transparency_International

1: 9540 Electricity_generation

[winner: 2144978 Nashville_Electric_Service, pem_pos=2, final_scores=-1.169374, final_local_scores=-1.1287754, pem=-1.0984836, similarity=-23.621471]

gt: 9540 Electricity_generation, pem_pos=6, final_scores=-1.7858993, final_local_scores=-1.9091661, pem=-4.272555, similarity=-29.648792]

2: 180234 Kilowatt_hour

[winner: 180234 Kilowatt_hour, pem_pos=0, final_scores=-0.6388603, final_local_scores=-0.60949814, pem=0.0, similarity=-14.098927]

gt: 180234 Kilowatt_hour, pem_pos=0,

final_scores=-0.6388603, final_local_scores=-0.60949814,
pem=0.0, similarity=-14.098927]
3: 33895300 Energy_in_Africa
[winner: 5334607 Africa, pem_pos=0,
final_scores=0.58649033, final_local_scores=0.55717087,
pem=-0.13197383, similarity=21.434498]
RECALL MISS: 33895300 Energy_in_Africa]
4: 31815852 Energy_in_Algeria
[winner: 358 Algeria, pem_pos=0, final_scores=0.46637338,
final_local_scores=0.45088297, pem=-0.33694586,
similarity=19.377325]
RECALL MISS: 31815852 Energy_in_Algeria]
5: 23295232 Energy_in_Angola
[winner: 701 Angola, pem_pos=0, final_scores=0.4073859,
final_local_scores=0.3626141, pem=-0.39249876,
similarity=17.045311]
RECALL MISS: 23295232 Energy_in_Angola]
6: 23391633 Energy_in_Cameroon
[winner: 5447 Cameroon, pem_pos=0, final_scores=0.443857,
final_local_scores=0.41652757, pem=-0.36813655,
similarity=18.52171]
RECALL MISS: 23391633 Energy_in_Cameroon]
7: 23401977 Energy_in_the_Democratic_Republic_of_the_Congo
[winner: 76762 Democratic_Republic_of_the_Congo,
pem_pos=0, final_scores=0.29039812,
final_local_scores=0.24872339, pem=-0.0055517,
similarity=11.542306]
RECALL MISS: 23401977
Energy_in_the_Democratic_Republic_of_the_Congo]
8: 23610440 Energy_in_Equatorial_Guinea
[winner: 9366 Equatorial_Guinea, pem_pos=0,
final_scores=0.24086563, final_local_scores=0.19648397,
pem=-0.26655748, similarity=11.402768]
RECALL MISS: 23610440 Energy_in_Equatorial_Guinea]
9: 23642688 Energy_in_Gabon
[winner: 12027 Gabon, pem_pos=0, final_scores=0.33235997,
final_local_scores=0.29830134, pem=-0.3021721,
similarity=14.634878]
RECALL MISS: 23642688 Energy_in_Gabon]
10: 31795980 Energy_in_Libya
[winner: 17633 Libya, pem_pos=0, final_scores=0.29673576,
final_local_scores=0.25065538, pem=-0.3054436,
similarity=13.23082]
RECALL MISS: 31795980 Energy_in_Libya]
11: 32115480 Energy_in_Nigeria
[winner: 21383 Nigeria, pem_pos=0,
final_scores=0.5084452, final_local_scores=0.4745258,
pem=-0.31455088, similarity=19.961086]
RECALL MISS: 32115480 Energy_in_Nigeria]
12: 940606 Population_growth
[winner: 12702333 Population_(album), pem_pos=1,
final_scores=-1.0536003, final_local_scores=-0.97300017,

```

pem=-0.6897683, similarity=-21.195486|
  RECALL MISS: 940606 Population_growth]
  [winner: 3304834 Growth_investing, pem_pos=6,
final_scores=-1.4272434, final_local_scores=-1.3841238,
pem=-2.9371083, similarity=-21.242847|
  gt: 940606 Population_growth, pem_pos=4,
final_scores=-1.7498139, final_local_scores=-1.8769214,
pem=-1.5838009, similarity=-43.30841 ]
13: 14598 Demographics_of_India
  [winner: 14533 India, pem_pos=0, final_scores=0.6864038,
final_local_scores=0.7032413, pem=-0.13622136,
similarity=25.81663|
  RECALL MISS: 14598 Demographics_of_India]
14: 21386 Demographics_of_Nigeria
  [winner: 21383 Nigeria, pem_pos=0,
final_scores=0.61917084, final_local_scores=0.59547675,
pem=-0.31455088, similarity=23.570503|
  RECALL MISS: 21386 Demographics_of_Nigeria]
15: 7075619 National_oil_company
  [winner: 883983 Network_operations_center, pem_pos=1,
final_scores=-0.71239275, final_local_scores=-0.6535844,
pem=-0.9708481, similarity=-10.134927|
  gt: 7075619 National_oil_company, pem_pos=17,
final_scores=-1.7278409, final_local_scores=-1.8170202,
pem=-7.969358, similarity=-6.7951875 ]
16: 1123978 Petronas
  [winner: 1123978 Petronas, pem_pos=0,
final_scores=-0.06354673, final_local_scores=-0.0025391728,
pem=0.0, similarity=4.013939|
  gt: 1123978 Petronas, pem_pos=0,
final_scores=-0.06354673, final_local_scores=-0.0025391728,
pem=0.0, similarity=4.013939 ]
1: 27421 Sudan
  [pem_pos=0, final_scores=0.48294133,
final_local_scores=0.42319113, pem=-0.2848764,
similarity=18.267782]
2: 33853117 Organisation_for_Economic_Co-
operation_and_Development
  [pem_pos=0, final_scores=0.2837385,
final_local_scores=0.31749672, pem=-0.0033873406,
similarity=13.582869]
3: 33309 World_Bank
  [pem_pos=0, final_scores=0.31050724,
final_local_scores=0.3314279, pem=0.0, similarity=13.980183]
4: 66241 Transparency_International
  [pem_pos=0, final_scores=0.15107752,
final_local_scores=0.18675682, pem=0.0, similarity=9.662909]
1: 27421 Sudan
  [pem_pos=0, final_scores=0.47002804,
final_local_scores=0.40908527, pem=-0.2848764,
similarity=17.846836]
2: 27421 Sudan

```

[pem_pos=0, final_scores=0.5530152,
final_local_scores=0.49973613, pem=-0.2848764,
similarity=20.552038]
3: 27421 Sudan
[pem_pos=0, final_scores=0.46673077,
final_local_scores=0.40548348, pem=-0.2848764,
similarity=17.739351]
4: 5334607 Africa
[pem_pos=0, final_scores=0.58649033,
final_local_scores=0.55717087, pem=-0.13197383,
similarity=21.434498]
5: 5334607 Africa
[pem_pos=0, final_scores=0.5827972,
final_local_scores=0.55313677, pem=-0.13197383,
similarity=21.314112]
6: 358 Algeria
[pem_pos=0, final_scores=0.46637338,
final_local_scores=0.45088297, pem=-0.33694586,
similarity=19.377325]
7: 701 Angola
[pem_pos=0, final_scores=0.4073859,
final_local_scores=0.3626141, pem=-0.39249876,
similarity=17.045311]
8: 5447 Cameroon
[pem_pos=0, final_scores=0.443857,
final_local_scores=0.41652757, pem=-0.36813655,
similarity=18.52171]
9: 76762 Democratic_Republic_of_the_Congo
[pem_pos=0, final_scores=0.29039812,
final_local_scores=0.24872339, pem=-0.0055517,
similarity=11.542306]
10: 9366 Equatorial_Guinea
[pem_pos=0, final_scores=0.24086563,
final_local_scores=0.19648397, pem=-0.26655748,
similarity=11.402768]
11: 12027 Gabon
[pem_pos=0, final_scores=0.33235997,
final_local_scores=0.29830134, pem=-0.3021721,
similarity=14.634878]
12: 17633 Libya
[pem_pos=0, final_scores=0.29673576,
final_local_scores=0.25065538, pem=-0.3054436,
similarity=13.23082]
13: 21383 Nigeria
[pem_pos=0, final_scores=0.5084452,
final_local_scores=0.4745258, pem=-0.31455088,
similarity=19.961086]
14: 27421 Sudan
[pem_pos=0, final_scores=0.48508042,
final_local_scores=0.4255277, pem=-0.2848764,
similarity=18.337511]
15: 14533 India

[pem_pos=0, final_scores=0.6864038,
 final_local_scores=0.7032413, pem=-0.13622136,
 similarity=25.81663]
 16: 21383 Nigeria
 [pem_pos=0, final_scores=0.61917084,
 final_local_scores=0.59547675, pem=-0.31455088,
 similarity=23.570503]
 17: 27421 Sudan
 [pem_pos=0, final_scores=0.42192394,
 final_local_scores=0.3565389, pem=-0.2848764,
 similarity=16.278748]
 18: 2646197 China_National_Petroleum_Corporation
 [pem_pos=0, final_scores=0.104476795,
 final_local_scores=0.15854225, pem=0.0, similarity=8.82093]
 19: 3781829 Kuwait_Petroleum_Corporation
 [pem_pos=0, final_scores=0.11135046,
 final_local_scores=0.22124454, pem=0.0, similarity=10.692091]
 20: 14533 India
 [pem_pos=0, final_scores=0.31792057,
 final_local_scores=0.30072936, pem=-0.618294,
 similarity=16.426456]

Sample: b'Jeffris_Hopkins'

Jeffris David Hopkins -LRB- born 23 August 1950 -RRB- is a
 former [1 1 Welsh 1] 2 cricket er . Hopkins was a right-
 handed who fielded as a 3 wicket-keeper . He was born at [1 1
 Bridgend 1] , [2 4 Glamorgan 2] . Having played for the [3 5 G
 lamorgan Second XI 3] since in 1967 and 1968 , Hopkins joined
 [2 2 Middlesex 2] for the 6 1969 season , making his 7 first-
 class debut in that same season against [3 3 Lancashire 3] in
 the . He also made his [4 4 List A 4] debut in that same
 season , against [5 5 Essex 5] in the 8 Player 's County Leag
 ue . With opportunities limited at [4 Middlesex 4] due to
 presence of regular wicket-keeper [6 6 John Murray 6] ,
 Hopkins made just three further first-class appearances , the
 last of which came against [7 7 Northamptonshire 7] in the ,
 as well as appearing in just one more [5 List A 5] match ,
 against [8 8 Derbyshire 8] in the 9 1971 [6 John Player League
 6] . Hopkins failed to make an impression at [7 Middlesex
 7] , scoring 8 runs at an 10 average of 1.60 in first-class
 cricket , while taking 9 . He later appeared for in their
 only [8 11 ICC Trophy 8] appearance in 1979 , playing in
 three matches against the 12 Netherlands , 13 Israel and the [
 9 14 United States 9] . He made scores of 92 against [10
 Israel 10] and 72 against the [11 United States 11] . Hopkins
 ' brother , 15 John

1: 353261 Bridgend
 2: 822103 Middlesex_County_Cricket_Club
 3: 749919 Lancashire_County_Cricket_Club
 4: 683459 List_A_cricket
 5: 1622252 Essex_County_Cricket_Club
 6: 2092404 John_Murray_(cricketer)

7: 530308 Northamptonshire_County_Cricket_Club
8: 1829984 Derbyshire_County_Cricket_Club
1: 2030193 Welsh_people
[winner: 69894 Wales, pem_pos=1, final_scores=0.36745566,
final_local_scores=0.39341658, pem=-1.0776843,
similarity=21.690664|
gt: 2030193 Welsh_people, pem_pos=2,
final_scores=-0.26991522, final_local_scores=-0.29640168,
pem=-1.6894228, similarity=4.431839]
2: 25675557 Cricket
[winner: 330091 Rhabdophoridae, pem_pos=10,
final_scores=-1.512907, final_local_scores=-1.4873891,
pem=-2.56456, similarity=-26.350471|
gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.8911823, final_local_scores=-2.11202,
pem=-0.07532314, similarity=-58.52756]
3: 463313 Wicket-keeper
[winner: 463313 Wicket-keeper, pem_pos=0,
final_scores=-1.4918221, final_local_scores=-1.6813035,
pem=-0.004312502, similarity=-46.060284|
gt: 463313 Wicket-keeper, pem_pos=0,
final_scores=-1.4918221, final_local_scores=-1.6813035,
pem=-0.004312502, similarity=-46.060284]
4: 52520 Glamorgan
[winner: 1622320 Glamorgan_County_Cricket_Club,
pem_pos=1, final_scores=0.521551,
final_local_scores=0.47116488, pem=-0.7903395,
similarity=22.448202|
gt: 52520 Glamorgan, pem_pos=0,
final_scores=-0.10534006, final_local_scores=-0.14145291,
pem=-0.5980491, similarity=3.1207597]
5: 1622320 Glamorgan_County_Cricket_Club
[winner: 1622320 Glamorgan_County_Cricket_Club,
pem_pos=0, final_scores=0.5053688,
final_local_scores=0.4534883, pem=0.0, similarity=17.622707|
gt: 1622320 Glamorgan_County_Cricket_Club, pem_pos=0,
final_scores=0.5053688, final_local_scores=0.4534883,
pem=0.0, similarity=17.622707]
6: 5672124 1969_English_cricket_season
[winner: 34610 1969, pem_pos=0, final_scores=-0.62720245,
final_local_scores=-0.43419114, pem=-0.6617498,
similarity=-5.2687182|
RECALL MISS: 5672124 1969_English_cricket_season]
[winner: 690345 Season(sports), pem_pos=0,
final_scores=-1.3992523, final_local_scores=-1.4732351,
pem=-1.219945, similarity=-33.24031|
RECALL MISS: 5672124 1969_English_cricket_season]
7: 11040 First-class_cricket
[winner: 662844 First-class_citizen, pem_pos=3,
final_scores=-1.2722157, final_local_scores=-1.2003198,
pem=-1.5935546, similarity=-23.064228|
gt: 11040 First-class_cricket, pem_pos=0,

final_scores=-1.3184547, final_local_scores=-1.511544,
 pem=-0.053819276, similarity=-40.725098]
 8: 28797906 1969_Player's_County_League
 [winner: 1177737 Pro40, pem_pos=0,
 final_scores=-0.13926996, final_local_scores=-0.19953619,
 pem=-1.0986123, similarity=4.1095796]
 RECALL MISS: 28797906 1969_Player's_County_League]
 [winner: 463313 Wicket-keeper, pem_pos=8,
 final_scores=-0.39841896, final_local_scores=-0.48692662,
 pem=-3.0503128, similarity=6.1469274]
 RECALL MISS: 28797906 1969_Player's_County_League]
 [winner: 1415027 David_Berman_(musician), pem_pos=5,
 final_scores=-2.661836, final_local_scores=-2.7817676,
 pem=-3.0012724, similarity=-62.602398]
 RECALL MISS: 28797906 1969_Player's_County_League]
 9: 30520920 1971_John_Player_League
 [winner: 1177737 Pro40, pem_pos=0,
 final_scores=0.33127058, final_local_scores=0.322701,
 pem=0.0, similarity=13.719754]
 gt: 30520920 1971_John_Player_League, pem_pos=7,
 final_scores=-0.74768275, final_local_scores=-0.7794038,
 pem=-5.0369525, similarity=8.22248]
 [winner: 34748 1971, pem_pos=0, final_scores=-0.87579507,
 final_local_scores=-0.8125549, pem=-0.6167389,
 similarity=-16.804625]
 RECALL MISS: 30520920 1971_John_Player_League]
 10: 18933273 Batting_average
 [winner: 875871 Bowling_average, pem_pos=0,
 final_scores=-0.9162057, final_local_scores=-1.0425833,
 pem=0.0, similarity=-27.023054]
 gt: 18933273 Batting_average, pem_pos=1,
 final_scores=-1.177981, final_local_scores=-1.3159251,
 pem=-0.6931472, similarity=-31.410675]
 11: 2675465 1979_ICC_Trophy
 [winner: 1952482 ICC_World_Cup_Qualifier, pem_pos=0,
 final_scores=0.28204054, final_local_scores=0.27915385,
 pem=0.0, similarity=12.42022]
 gt: 2675465 1979_ICC_Trophy, pem_pos=1,
 final_scores=-0.124408334, final_local_scores=-0.13291852,
 pem=-1.3137479, similarity=7.2675242]
 12: 2164343 Netherlands_national_cricket_team
 [winner: 9647657 Netherlands_national_football_team,
 pem_pos=1, final_scores=0.015734121,
 final_local_scores=0.055663288, pem=-2.660505,
 similarity=20.219057]
 gt: 2164343 Netherlands_national_cricket_team,
 pem_pos=2, final_scores=-0.13691124,
 final_local_scores=-0.20055175, pem=-3.8293223,
 similarity=18.929298]
 13: 4399971 Israel_national_cricket_team
 [winner: 9282173 Israel, pem_pos=0,
 final_scores=-0.10291532, final_local_scores=-0.05601327,

pem=-0.21279564, similarity=3.575381|
gt: 4399971 Israel_national_cricket_team, pem_pos=7,
final_scores=-0.4883852, final_local_scores=-0.5107697,
pem=-4.876911, similarity=15.368727]
14: 980450 United_States_national_cricket_team
[winner: 3434750 United_States, pem_pos=0,
final_scores=0.34585726, final_local_scores=0.4007507,
pem=-0.073457584, similarity=16.448387|
gt: 980450 United_States_national_cricket_team,
pem_pos=4, final_scores=-0.58108085,
final_local_scores=-0.6666165, pem=-5.176109,
similarity=12.345032]
15: 12848577 John_Hopkins_(cricketer)
[winner: 12493 Gospel_of_John, pem_pos=0,
final_scores=-0.6376158, final_local_scores=-0.5692801,
pem=-1.6259884, similarity=-4.056362|
RECALL MISS: 12848577 John_Hopkins_(cricketer)]
1: 353261 Bridgend
[pem_pos=0, final_scores=0.15148108,
final_local_scores=0.17339706, pem=-0.26986,
similarity=10.731768]
2: 822103 Middlesex_County_Cricket_Club
[pem_pos=1, final_scores=0.6025608,
final_local_scores=0.5560711, pem=-1.5695794,
similarity=29.2196]
3: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.572352,
final_local_scores=0.530295, pem=-1.7295884,
similarity=29.320545]
4: 683459 List_A_cricket
[pem_pos=0, final_scores=0.31167173,
final_local_scores=0.26365134, pem=-0.020267159,
similarity=12.067811]
5: 1622252 Essex_County_Cricket_Club
[pem_pos=1, final_scores=0.8598421,
final_local_scores=0.84258723, pem=-2.3459642,
similarity=41.991913]
6: 2092404 John_Murray_(cricketer)
[pem_pos=7, final_scores=0.13539608,
final_local_scores=0.17421299, pem=-2.8732252,
similarity=24.913622]
7: 530308 Northamptonshire_County_Cricket_Club
[pem_pos=1, final_scores=0.47221142,
final_local_scores=0.4241026, pem=-1.2069724,
similarity=23.309483]
8: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.70425326,
final_local_scores=0.6733696, pem=-1.4828352,
similarity=32.24829]
1: 69894 Wales
[pem_pos=1, final_scores=0.36745566,
final_local_scores=0.39341658, pem=-1.0776843,


```

similarity=21.690664]
2: 1622320 Glamorgan_County_Cricket_Club
   [pem_pos=1, final_scores=0.521551,
final_local_scores=0.47116488, pem=-0.7903395,
similarity=22.448202]
3: 1622320 Glamorgan_County_Cricket_Club
   [pem_pos=0, final_scores=0.5053688,
final_local_scores=0.4534883, pem=0.0, similarity=17.622707]
4: 822103 Middlesex_County_Cricket_Club
   [pem_pos=1, final_scores=0.7495508,
final_local_scores=0.71663535, pem=-1.5695794,
similarity=34.011158]
5: 683459 List_A_cricket
   [pem_pos=0, final_scores=0.17209731,
final_local_scores=0.11118749, pem=-0.020267159,
similarity=7.517985]
6: 1177737 Pro40
   [pem_pos=0, final_scores=0.33127058,
final_local_scores=0.322701, pem=0.0, similarity=13.719754]
7: 822103 Middlesex_County_Cricket_Club
   [pem_pos=1, final_scores=0.72723,
final_local_scores=0.69225323, pem=-1.5695794,
similarity=33.283546]
8: 1952482 ICC_World_Cup_Qualifier
   [pem_pos=0, final_scores=0.28204054,
final_local_scores=0.27915385, pem=0.0, similarity=12.42022]
9: 3434750 United_States
   [pem_pos=0, final_scores=0.34585726,
final_local_scores=0.4007507, pem=-0.073457584,
similarity=16.448387]
10: 9282173 Israel
    [pem_pos=0, final_scores=0.14695857,
final_local_scores=0.23271719, pem=-0.21279564,
similarity=12.191673]
11: 3434750 United_States
    [pem_pos=0, final_scores=0.30698144,
final_local_scores=0.35828477, pem=-0.073457584,
similarity=15.18112]

```

Sample: b'2009_Team_Speedway_Junior_European_Championship'

The 2009 Team Speedway Junior European Championship will be the 2nd 1 UEM 2 Team Speedway Junior European Championship season . The Final will be held on August 23 , 3 2009 in [1 1 Holsted 1] , [2 2 Denmark 2] . The championship was won by [1 4 Poland 1] -LRB- 40 points -RRB- , who they beat defending champion [2 5 Sweden 2] -LRB- 38 points -RRB- , host team [3 6 Denmark 3] -LRB- 26 points -RRB- and [4 7 Czech Republic 4] -LRB- 15 points -RRB- In the Final will be the defending Champions [5 8 Sweden 5] , host team [6 9 Denmark 6] -LRB- 3rd place in 2008 Final -RRB- and [7 10 Poland 7] -LRB- 4th place -RRB- . A last finalist will be determined in one Semi-Final . In 11 Bockhorn , [3 3 Germany 3] on 18 July will be ,

[8 12 Czech Republic 8] , [9 13 Russia 9] and host team 14 Germany -LRB- 2nd place -RRB- . The [10 15 Russian 10] team withdrew and was replaced by [11 16 Finland 11]

1: 146491 Holsted

2: 76972 Denmark

3: 11867 Germany

1: 12407732 Union_Européenne_de_Motocyclisme

[winner: 12407732 Union_Européenne_de_Motocyclisme,
pem_pos=0, final_scores=-0.28745747,
final_local_scores=-0.23151276, pem=0.0,
similarity=-2.8190897]

gt: 12407732 Union_Européenne_de_Motocyclisme,
pem_pos=0, final_scores=-0.28745747,
final_local_scores=-0.23151276, pem=0.0,
similarity=-2.8190897]

2: 15776048 Team_Speedway_Junior_European_Championship

[winner: 34379447
European_Men's_Junior_Handball_Championship, pem_pos=0,
final_scores=-0.59250367, final_local_scores=-0.5440804,
pem=0.0, similarity=-12.14673]

RECALL MISS: 15776048

Team_Speedway_Junior_European_Championship]

[winner: 26252745 Speedway, West_Virginia, pem_pos=4,
final_scores=-1.2462146, final_local_scores=-1.1995561,
pem=-1.7755163, similarity=-22.0519]

RECALL MISS: 15776048

Team_Speedway_Junior_European_Championship]

[winner: 8155197 Volleyball_at_the_1972_Summer_Olympics,
pem_pos=15, final_scores=-2.4624991,
final_local_scores=-2.59312, pem=-4.1031766,
similarity=-50.98045]

RECALL MISS: 15776048

Team_Speedway_Junior_European_Championship]

3: 4051722 2009_in_sports

[winner: 35983 2009, pem_pos=0, final_scores=-0.88985306,
final_local_scores=-0.7988814, pem=-0.9296576,
similarity=-14.69488]

RECALL MISS: 4051722 2009_in_sports]

4: 26761992 Poland_national_under-19_speedway_team

[winner: 771860 Poland_national_football_team, pem_pos=1,
final_scores=0.04930769, final_local_scores=0.05693227,
pem=-3.319454, similarity=23.840393]

RECALL MISS: 26761992

Poland_national_under-19_speedway_team]

5: 27831566 Sweden_national_under-19_speedway_team

[winner: 5058739 Sweden, pem_pos=0,
final_scores=0.094111755, final_local_scores=0.008694902,
pem=-0.2135508, similarity=5.5105085]

RECALL MISS: 27831566

Sweden_national_under-19_speedway_team]

6: 25008089 Denmark_national_under-19_speedway_team

[winner: 76972 Denmark, pem_pos=0,

final_scores=0.29369736, final_local_scores=0.2402991,
 pem=-0.26762357, similarity=12.716096|
 RECALL MISS: 25008089
 Denmark_national_under-19_speedway_team]
 7: 24999333 Czech_Republic_national_under-19_speedway_team
 [winner: 5321 Czech_Republic, pem_pos=0,
 final_scores=0.20431845, final_local_scores=0.12990889,
 pem=-0.27794555, similarity=9.477964|
 RECALL MISS: 24999333
 Czech_Republic_national_under-19_speedway_team]
 8: 27831566 Sweden_national_under-19_speedway_team
 [winner: 5058739 Sweden, pem_pos=0,
 final_scores=0.21680968, final_local_scores=0.14272371,
 pem=-0.2135508, similarity=9.510196|
 RECALL MISS: 27831566
 Sweden_national_under-19_speedway_team]
 9: 25008089 Denmark_national_under-19_speedway_team
 [winner: 76972 Denmark, pem_pos=0,
 final_scores=0.13463588, final_local_scores=0.066548616,
 pem=-0.26762357, similarity=7.5310345|
 RECALL MISS: 25008089
 Denmark_national_under-19_speedway_team]
 10: 26761992 Poland_national_under-19_speedway_team
 [winner: 22936 Poland, pem_pos=0,
 final_scores=0.12121059, final_local_scores=0.05015774,
 pem=-0.15333635, similarity=6.420387|
 RECALL MISS: 26761992
 Poland_national_under-19_speedway_team]
 11: 4501372 Bockhorn, Lower_Saxony
 [winner: 4501372 Bockhorn, Lower_Saxony, pem_pos=0,
 final_scores=-0.055266842, final_local_scores=0.07962364,
 pem=0.0, similarity=6.4658413|
 gt: 4501372 Bockhorn, Lower_Saxony, pem_pos=0,
 final_scores=-0.055266842, final_local_scores=0.07962364,
 pem=0.0, similarity=6.4658413]
 12: 24999333 Czech_Republic_national_under-19_speedway_team
 [winner: 5321 Czech_Republic, pem_pos=0,
 final_scores=0.1617379, final_local_scores=0.08339611,
 pem=-0.27794555, similarity=8.089931|
 RECALL MISS: 24999333
 Czech_Republic_national_under-19_speedway_team]
 13: 27829178 Russia_national_under-19_speedway_team
 [winner: 25127146 Russia_national_football_team,
 pem_pos=2, final_scores=0.12802921,
 final_local_scores=0.14267558, pem=-3.5373242,
 similarity=27.583956|
 RECALL MISS: 27829178
 Russia_national_under-19_speedway_team]
 14: 25124605 Germany_national_under-19_speedway_team
 [winner: 250204 Germany_national_football_team,
 pem_pos=1, final_scores=-0.060416594,
 final_local_scores=-0.08300698, pem=-3.3613694,

similarity=19.892267|
RECALL MISS: 25124605
Germany_national_under-19_speedway_team]
15: 27829178 Russia_national_under-19_speedway_team
[winner: 25391 Russia, pem_pos=1,
final_scores=0.46001536, final_local_scores=0.4483729,
pem=-1.0515022, similarity=23.188286|
RECALL MISS: 27829178
Russia_national_under-19_speedway_team]
[winner: 140801 Team, pem_pos=0, final_scores=-1.7375393,
final_local_scores=-1.8384824, pem=-0.5173818,
similarity=-47.960663|
RECALL MISS: 27829178
Russia_national_under-19_speedway_team]
16: 25008142 Finland_national_under-19_speedway_team
[winner: 10577 Finland, pem_pos=0,
final_scores=0.2937925, final_local_scores=0.22852236,
pem=-0.2445894, similarity=12.239391|
RECALL MISS: 25008142
Finland_national_under-19_speedway_team]
1: 146491 Holsted
[pem_pos=0, final_scores=0.14562495,
final_local_scores=0.23718241, pem=0.0, similarity=11.167709]
2: 76972 Denmark
[pem_pos=0, final_scores=0.48156875,
final_local_scores=0.44552004, pem=-0.26762357,
similarity=18.840298]
3: 11867 Germany
[pem_pos=0, final_scores=0.671771,
final_local_scores=0.65596354, pem=-0.15529114,
similarity=24.509472]
1: 771860 Poland_national_football_team
[pem_pos=1, final_scores=0.04930769,
final_local_scores=0.05693227, pem=-3.319454,
similarity=23.840393]
2: 5058739 Sweden
[pem_pos=0, final_scores=0.094111755,
final_local_scores=0.008694902, pem=-0.2135508,
similarity=5.5105085]
3: 76972 Denmark
[pem_pos=0, final_scores=0.29369736,
final_local_scores=0.2402991, pem=-0.26762357,
similarity=12.716096]
4: 5321 Czech_Republic
[pem_pos=0, final_scores=0.20431845,
final_local_scores=0.12990889, pem=-0.27794555,
similarity=9.477964]
5: 5058739 Sweden
[pem_pos=0, final_scores=0.21680968,
final_local_scores=0.14272371, pem=-0.2135508,
similarity=9.510196]
6: 76972 Denmark

[pem_pos=0, final_scores=0.13463588,
 final_local_scores=0.066548616, pem=-0.26762357,
 similarity=7.5310345]
 7: 22936 Poland
 [pem_pos=0, final_scores=0.12121059,
 final_local_scores=0.05015774, pem=-0.15333635,
 similarity=6.420387]
 8: 5321 Czech_Republic
 [pem_pos=0, final_scores=0.1617379,
 final_local_scores=0.08339611, pem=-0.27794555,
 similarity=8.089931]
 9: 25127146 Russia_national_football_team
 [pem_pos=2, final_scores=0.12802921,
 final_local_scores=0.14267558, pem=-3.5373242,
 similarity=27.583956]
 10: 25391 Russia
 [pem_pos=1, final_scores=0.46001536,
 final_local_scores=0.4483729, pem=-1.0515022,
 similarity=23.188286]
 11: 10577 Finland
 [pem_pos=0, final_scores=0.2937925,
 final_local_scores=0.22852236, pem=-0.2445894,
 similarity=12.239391]

Sample: b'Bradley_Dale_Peveto'

[1 Bradley Dale Peveto 1] -LRB- born December 28 , 1962 -
 RRB- is an [2 American 2] 1 football coach and former
 player . He grew up in the small town of [3 2 Orangefield 3] ,
 [4 Texas 4] and attended [1 1 Southern Methodist University 1]
 where he played 3 defensive back from 1982 to 1986 . He began
 his coaching career in 1987 as an assistant at [2 2 Trinity Va
 lley Community College 2] , and also served as an assistant
 at [5 4 Stephen F. Austin 5] , [6 5 Southern Miss 6] , [7 6 Ar
 kansas 7] , 7 Northwestern State , [8 8 Houston 8] , [9
 Middle Tennessee 9] and 9 LSU . [10 Peveto 10] served as the
 head football coach at [3 3 Northwestern State University 3]
 from 2008 to 2012 . During his four years as the Demons '
 head coach , he compiled an overall record of 14 wins and 30
 losses . In December 2012 [11 Peveto 11] was hired to serve
 as the secondary coach at 10 Kentucky . [12 Peveto 12]
 started his coaching career as the secondary coach at [4 4 Tr
 inity Valley Community College 4] in 1987 . From Trinity
 Valley , [13 Peveto 13] served as an assistant coach at [14 11
 Stephen F. Austin 14] , [15 12 Southern Miss 15] , [16 13 Ark
 ansas 16] , [17 14 Northwestern State 17] , [18 15 Houston 18]
 , [19 Middle Tennessee 19] before being hired at 16 LSU in
 2005 . As part of the LSU staff , he coached the special
 teams and linebackers from 2005 through 2007 , including the 1
 7 2007 national championship team
 1: 479109 Southern_Methodist_University
 2: 11192606 Trinity_Valley_Community_College
 3: 188638 Northwestern_State_University

4: 11192606 Trinity_Valley_Community_College

1: 18951490 American_football

[winner: 18951490 American_football, pem_pos=1,
final_scores=-2.5260422, final_local_scores=-2.7878206,
pem=-1.8467362, similarity=-69.06157]

gt: 18951490 American_football, pem_pos=1,
final_scores=-2.5260422, final_local_scores=-2.7878206,
pem=-1.8467362, similarity=-69.06157]

2: 6395100 Orangefield,_Texas

[winner: 29810 Texas, pem_pos=0, final_scores=0.68020576,
final_local_scores=0.70058817, pem=-0.16191302,
similarity=25.877169]

RECALL MISS: 6395100 Orangefield,_Texas]

[winner: 6395100 Orangefield,_Texas, pem_pos=0,
final_scores=0.09584899, final_local_scores=0.17646003,
pem=0.0, similarity=9.355632]

gt: 6395100 Orangefield,_Texas, pem_pos=0,
final_scores=0.09584899, final_local_scores=0.17646003,
pem=0.0, similarity=9.355632]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8692335, final_local_scores=-1.855299,
pem=-4.6728287, similarity=-25.864538]

RECALL MISS: 6395100 Orangefield,_Texas]

3: 359688 Defensive_back

[winner: 8654693 Telus_Cup_-_Defensive, pem_pos=3,
final_scores=-1.1309888, final_local_scores=-1.0709414,
pem=-1.3500465, similarity=-20.527552]

RECALL MISS: 359688 Defensive_back]

[winner: 19779519 Back_Peninsula, pem_pos=4,
final_scores=-1.2969329, final_local_scores=-1.2243805,
pem=-3.0864334, similarity=-15.663735]

RECALL MISS: 359688 Defensive_back]

4: 23208476 Stephen_F._Austin_Lumberjacks_football

[winner: 639824 Stephen_F._Austin_State_University,
pem_pos=1, final_scores=0.13790588,
final_local_scores=0.0905789, pem=-0.46631515,
similarity=9.328662]

gt: 23208476 Stephen_F._Austin_Lumberjacks_football,
pem_pos=4, final_scores=-0.98075145,
final_local_scores=-1.0428271, pem=-6.335814,
similarity=7.424814]

5: 7958395 Southern_Miss_Golden_Eagles_football

[winner: 4610689 University_of_Southern_Mississippi,
pem_pos=0, final_scores=0.054963663,
final_local_scores=-0.000729233, pem=-0.60437495,
similarity=7.3546357]

gt: 7958395 Southern_Miss_Golden_Eagles_football,
pem_pos=7, final_scores=-0.409872,
final_local_scores=-0.47091782, pem=-3.1873496,
similarity=7.3698893]

6: 7957059 Arkansas_Razorbacks_football

[winner: 1930 Arkansas, pem_pos=0,

final_scores=0.09709932, final_local_scores=0.070160866,
pem=-0.32007957, similarity=7.924096|
gt: 7957059 Arkansas_Razorbacks_football, pem_pos=2,
final_scores=-0.26630497, final_local_scores=-0.34154308,
pem=-3.2738736, similarity=11.70122]
7: 23208349 Northwestern_State_Demons_football
[winner: 188638 Northwestern_State_University, pem_pos=0,
final_scores=0.029160753, final_local_scores=-0.036920004,
pem=0.0, similarity=2.9879463|
gt: 23208349 Northwestern_State_Demons_football,
pem_pos=3, final_scores=-0.8167255,
final_local_scores=-0.8717114, pem=-3.9464245,
similarity=-0.46262342]
8: 9570822 Houston_Cougars_football
[winner: 13774 Houston, pem_pos=0,
final_scores=0.16769932, final_local_scores=0.12578681,
pem=-0.3232624, similarity=9.601394|
gt: 9570822 Houston_Cougars_football, pem_pos=4,
final_scores=-0.6036723, final_local_scores=-0.69000536,
pem=-4.0497565, similarity=5.521784]
9: 5978520 LSU_Tigers_football
[winner: 5978520 LSU_Tigers_football, pem_pos=1,
final_scores=-0.09303604, final_local_scores=-0.12554947,
pem=-1.4477273, similarity=8.216033|
gt: 5978520 LSU_Tigers_football, pem_pos=1,
final_scores=-0.09303604, final_local_scores=-0.12554947,
pem=-1.4477273, similarity=8.216033]
10: 7957688 Kentucky_Wildcats_football
[winner: 16846 Kentucky, pem_pos=0,
final_scores=-0.21593359, final_local_scores=-0.24699238,
pem=-0.28237545, similarity=-1.7454308|
gt: 7957688 Kentucky_Wildcats_football, pem_pos=3,
final_scores=-0.2445984, final_local_scores=-0.29071474,
pem=-3.635306, similarity=15.183563]
11: 23208476 Stephen_F._Austin_Lumberjacks_football
[winner: 639824 Stephen_F._Austin_State_University,
pem_pos=1, final_scores=0.11329572,
final_local_scores=0.06369601, pem=-0.46631515,
similarity=8.526423|
gt: 23208476 Stephen_F._Austin_Lumberjacks_football,
pem_pos=4, final_scores=-1.0336295,
final_local_scores=-1.1005883, pem=-6.335814,
similarity=5.701101]
12: 7958395 Southern_Miss_Golden_Eagles_football
[winner: 4610689 University_of_Southern_Mississippi,
pem_pos=0, final_scores=0.044185445,
final_local_scores=-0.010119438, pem=-0.60437495,
similarity=7.074414|
gt: 7958395 Southern_Miss_Golden_Eagles_football,
pem_pos=7, final_scores=-0.38566387,
final_local_scores=-0.43875524, pem=-3.1873496,
similarity=8.329684]

13: 7957059 Arkansas_Razorbacks_football
 [winner: 1930 Arkansas, pem_pos=0,
 final_scores=0.15155391, final_local_scores=0.12964424,
 pem=-0.32007957, similarity=9.699199]
 gt: 7957059 Arkansas_Razorbacks_football, pem_pos=2,
 final_scores=-0.28226256, final_local_scores=-0.3589743,
 pem=-3.2738736, similarity=11.181037]

14: 23208349 Northwestern_State_Demons_football
 [winner: 188638 Northwestern_State_University, pem_pos=0,
 final_scores=0.057639226, final_local_scores=0.00059872866,
 pem=0.0, similarity=4.10758]
 gt: 23208349 Northwestern_State_Demons_football,
 pem_pos=3, final_scores=-0.7836956,
 final_local_scores=-0.8345831, pem=-3.9464245,
 similarity=0.6453568]

15: 9570822 Houston_Cougars_football
 [winner: 13774 Houston, pem_pos=0,
 final_scores=0.3744756, final_local_scores=0.35165846,
 pem=-0.3232624, similarity=16.341854]
 gt: 9570822 Houston_Cougars_football, pem_pos=4,
 final_scores=-0.60585856, final_local_scores=-0.6923935,
 pem=-4.0497565, similarity=5.450516]

16: 5978520 LSU_Tigers_football
 [winner: 277905 Louisiana_State_University, pem_pos=0,
 final_scores=-0.0129978955, final_local_scores=-0.06165634,
 pem=-0.83875704, similarity=6.8110557]
 gt: 5978520 LSU_Tigers_football, pem_pos=1,
 final_scores=-0.09858763, final_local_scores=-0.13161373,
 pem=-1.4477273, similarity=8.035063]

17: 8986611 2007_LSU_Tigers_football_team
 [winner: 36165 2007, pem_pos=0, final_scores=-0.528099,
 final_local_scores=-0.40906194, pem=-0.24805899,
 similarity=-6.7685275]
 RECALL MISS: 8986611 2007_LSU_Tigers_football_team]
 [winner: 17278765
 College_football_national_championships_in_NCAA_Division_I_FB
 S, pem_pos=0, final_scores=-0.82416433,
 final_local_scores=-0.89238423, pem=-1.0235064,
 similarity=-16.974825]
 RECALL MISS: 8986611 2007_LSU_Tigers_football_team]
 [winner: 140801 Team, pem_pos=0, final_scores=-1.3823942,
 final_local_scores=-1.4725604, pem=-0.5173818,
 similarity=-37.040825]
 RECALL MISS: 8986611 2007_LSU_Tigers_football_team]

1: 479109 Southern_Methodist_University
 [pem_pos=0, final_scores=0.30008417,
 final_local_scores=0.26863825, pem=0.0, similarity=12.106414]

2: 11192606 Trinity_Valley_Community_College
 [pem_pos=0, final_scores=0.18046792,
 final_local_scores=0.19250384, pem=0.0, similarity=9.834412]

3: 188638 Northwestern_State_University
 [pem_pos=0, final_scores=0.2894091,

final_local_scores=0.2537721, pem=0.0, similarity=11.662779]
4: 11192606 Trinity_Valley_Community_College
[pem_pos=0, final_scores=0.16043796,
final_local_scores=0.17062414, pem=0.0, similarity=9.181478]
1: 34151888 Bradley_Dale_Peveto
[pem_pos=0, final_scores=0.27423233,
final_local_scores=0.2820238, pem=0.0, similarity=12.505865]
2: 3434750 United_States
[pem_pos=0, final_scores=0.43478245,
final_local_scores=0.47358936, pem=-0.12994486,
similarity=18.929226]
3: 6395100 Orangefield,_Texas
[pem_pos=0, final_scores=0.09584899,
final_local_scores=0.17646003, pem=0.0, similarity=9.355632]
4: 29810 Texas
[pem_pos=0, final_scores=0.68020576,
final_local_scores=0.70058817, pem=-0.16191302,
similarity=25.877169]
5: 639824 Stephen_F._Austin_State_University
[pem_pos=1, final_scores=0.13790588,
final_local_scores=0.0905789, pem=-0.46631515,
similarity=9.328662]
6: 4610689 University_of_Southern_Mississippi
[pem_pos=0, final_scores=0.054963663,
final_local_scores=-0.000729233, pem=-0.60437495,
similarity=7.3546357]
7: 1930 Arkansas
[pem_pos=0, final_scores=0.09709932,
final_local_scores=0.070160866, pem=-0.32007957,
similarity=7.924096]
8: 13774 Houston
[pem_pos=0, final_scores=0.16769932,
final_local_scores=0.12578681, pem=-0.3232624,
similarity=9.601394]
9: 637662 Middle_Tennessee_State_University
[pem_pos=1, final_scores=0.04541762,
final_local_scores=-0.0028606951, pem=-1.4743476,
similarity=12.022076]
10: 34151888 Bradley_Dale_Peveto
[pem_pos=0, final_scores=0.04956551,
final_local_scores=0.03660941, pem=0.0, similarity=5.182211]
11: 34151888 Bradley_Dale_Peveto
[pem_pos=0, final_scores=0.107714966,
final_local_scores=0.10012886, pem=0.0, similarity=7.077758]
12: 34151888 Bradley_Dale_Peveto
[pem_pos=0, final_scores=0.05114545,
final_local_scores=0.038335234, pem=0.0, similarity=5.2337127]
13: 34151888 Bradley_Dale_Peveto
[pem_pos=0, final_scores=0.13970272,
final_local_scores=0.13507062, pem=0.0, similarity=8.12049]
14: 639824 Stephen_F._Austin_State_University
[pem_pos=1, final_scores=0.11329572,

final_local_scores=0.06369601, pem=-0.46631515,
 similarity=8.526423]
 15: 4610689 University_of_Southern_Mississippi
 [pem_pos=0, final_scores=0.044185445,
 final_local_scores=-0.010119438, pem=-0.60437495,
 similarity=7.074414]
 16: 1930 Arkansas
 [pem_pos=0, final_scores=0.15155391,
 final_local_scores=0.12964424, pem=-0.32007957,
 similarity=9.699199]
 17: 188638 Northwestern_State_University
 [pem_pos=0, final_scores=0.057639226,
 final_local_scores=0.00059872866, pem=0.0, similarity=4.10758]
 18: 13774 Houston
 [pem_pos=0, final_scores=0.3744756,
 final_local_scores=0.35165846, pem=-0.3232624,
 similarity=16.341854]
 19: 637662 Middle_Tennessee_State_University
 [pem_pos=1, final_scores=0.054339573,
 final_local_scores=0.014848828, pem=-1.4743476,
 similarity=12.550563]

Sample: b'Goldie_Hexagon_Racing'

Goldie Hexagon Racing was a motor racing team from the [1 1 U
 nited Kingdom 1] , founded by Paul Michaels . It raced in [2 2
 Formula One 2] in , using customer [3 3 Brabham 3] cars . [1
 Goldie Hexagon Racing 1] first entered a [2 Formula One 2]
 race in 1972 , at the non-Championship 1 World Championship Vi
 ctory Race . They entered a 2 March 721 for [4 4 British 4]
 racing driver [5 5 John Watson 5] . [3 Watson 3] qualified
 the car in tenth position and finished the race in sixth
 position . In [4 Goldie Hexagon Racing 4] returned to [5
 Formula One 5] and raced a full Championship season , again
 with [6 John Watson 6] , entering [6 6 Brabham 6] cars . The
 season started in [7 3 Argentina 7] , where [8 Watson 8]
 qualified the car in 20th position and finished 12th . In 4 Br
 azil [9 Watson 9] qualified in 15th position , but retired on
 lap 27 . He also failed to finish in [10 5 South Africa 10] .
 In [11 6 Spain 11] [12 Watson 12] qualified in 15th position
 and finished 11th , and he finished in the same position in [1
 3 7 Belgium 13] . In 8 Monaco [14 Watson 14] scored [15
 Goldie Hexagon Racing 15] 's first World Championship
 points , when he finished in sixth position . In [16 9 Sweden
 16] [17 Watson 17] finished in 11th position . In 10 the Neth
 erlands [18 Watson 18] finished 7th . In [19 11 France 19]
 Goldie Hexagon Racing entered two cars , one for [20 Watson 20
] and one for [7 7 Carlos Pace 7] . While [21 Watson 21]
 finished in 16th position , [22 Pace 22] failed to qualify .
 Two races later , in [23 12 Austria 23] , [24 Watson 24]
 finished in fourth position , scoring three World
 Championship points . In [25 13 Italy 25] [26 Watson 26]
 qualified the car in fourth position , his highest of the

season , but only finished in seventh place . In the [27 14 US A 27]

1: 31717 United_Kingdom

2: 10854 Formula_One

3: 4996 Brabham

4: 31717 United_Kingdom

5: 30874971 John_Watson_(racing_driver)

6: 4996 Brabham

7: 1219709 Carlos_Pace

1: 25329710 1972_World_Championship_Victory_Race

[winner: 18414537 1971_World_Championship_Victory_Race,
pem_pos=1, final_scores=-0.06269969,
final_local_scores=-0.015079811, pem=-0.18232158,
similarity=4.6311946]

gt: 25329710 1972_World_Championship_Victory_Race,
pem_pos=0, final_scores=-0.11496802,
final_local_scores=-0.06251501, pem=0.0,
similarity=2.2241402]

2: 1347020 March_Engineering

[winner: 1347020 March_Engineering, pem_pos=2,
final_scores=-0.057666734, final_local_scores=-0.09604007,
pem=-1.530765, similarity=9.548224]

gt: 1347020 March_Engineering, pem_pos=2,
final_scores=-0.057666734, final_local_scores=-0.09604007,
pem=-1.530765, similarity=9.548224]

3: 1122599 1974_Argentine_Grand_Prix

[winner: 18951905 Argentina, pem_pos=0,
final_scores=0.35472822, final_local_scores=0.42541307,
pem=0.0, similarity=16.784887]

gt: 1122599 1974_Argentine_Grand_Prix, pem_pos=2,
final_scores=-0.8775434, final_local_scores=-0.92348486,
pem=-5.32301, similarity=5.4784384]

4: 1122600 1974_Brazilian_Grand_Prix

[winner: 3383 Brazil, pem_pos=0,
final_scores=0.0076596886, final_local_scores=0.05693133,
pem=-0.21313342, similarity=6.94771]

RECALL MISS: 1122600 1974_Brazilian_Grand_Prix]

5: 1122602 1974_South_African_Grand_Prix

[winner: 17416221 South_Africa, pem_pos=0,
final_scores=0.35396725, final_local_scores=0.44421232,
pem=-0.26251382, similarity=18.773485]

RECALL MISS: 1122602 1974_South_African_Grand_Prix]

6: 1122604 1974_Spanish_Grand_Prix

[winner: 26667 Spain, pem_pos=0, final_scores=0.08884616,
final_local_scores=0.13662681, pem=-0.18857051,
similarity=9.192405]

RECALL MISS: 1122604 1974_Spanish_Grand_Prix]

7: 1122615 1974_Belgian_Grand_Prix

[winner: 3343 Belgium, pem_pos=0,
final_scores=0.59269506, final_local_scores=0.66286165,
pem=-0.24180195, similarity=25.185783]

RECALL MISS: 1122615 1974_Belgian_Grand_Prix]

8: 1122618 1974_Monaco_Grand_Prix
[winner: 19261 Monaco, pem_pos=0,
final_scores=-0.48911136, final_local_scores=-0.52716136,
pem=-0.58243024, similarity=-8.474487]
RECALL MISS: 1122618 1974_Monaco_Grand_Prix]

9: 1122620 1974_Swedish_Grand_Prix
[winner: 5058739 Sweden, pem_pos=0,
final_scores=0.18511046, final_local_scores=0.24777249,
pem=-0.2135508, similarity=12.645061]
RECALL MISS: 1122620 1974_Swedish_Grand_Prix]

10: 1122622 1974_Dutch_Grand_Prix
[winner: 21148 Netherlands, pem_pos=0,
final_scores=-0.19796845, final_local_scores=-0.19813958,
pem=-0.24940795, similarity=-0.4668479]
RECALL MISS: 1122622 1974_Dutch_Grand_Prix]
[winner: 1978556 In_forma_pauperis, pem_pos=1,
final_scores=-1.3661504, final_local_scores=-1.3102856,
pem=-0.8070914, similarity=-30.622732]
RECALL MISS: 1122622 1974_Dutch_Grand_Prix]

11: 1122623 1974_French_Grand_Prix
[winner: 5843419 France, pem_pos=0,
final_scores=0.43424094, final_local_scores=0.49551028,
pem=-0.16240834, similarity=19.75993]
RECALL MISS: 1122623 1974_French_Grand_Prix]

12: 1122627 1974_Austrian_Grand_Prix
[winner: 26964606 Austria, pem_pos=0,
final_scores=0.636746, final_local_scores=0.72913,
pem=-0.25262418, similarity=27.222218]
RECALL MISS: 1122627 1974_Austrian_Grand_Prix]

13: 1122628 1974_Italian_Grand_Prix
[winner: 14532 Italy, pem_pos=0, final_scores=0.35493547,
final_local_scores=0.4183895, pem=-0.20612453,
similarity=17.696228]
RECALL MISS: 1122628 1974_Italian_Grand_Prix]

14: 1122631 1974_United_States_Grand_Prix
[winner: 3434750 United_States, pem_pos=0,
final_scores=0.3761242, final_local_scores=0.4352697,
pem=-0.13038418, similarity=17.78808]
gt: 1122631 1974_United_States_Grand_Prix, pem_pos=11,
final_scores=-0.8890104, final_local_scores=-0.9768381,
pem=-4.7775908, similarity=0.9201962]

1: 31717 United_Kingdom
[pem_pos=0, final_scores=0.4019745,
final_local_scores=0.4506216, pem=-0.09933331,
similarity=18.07735]

2: 10854 Formula_One
[pem_pos=0, final_scores=0.44848424,
final_local_scores=0.42076284, pem=-0.08942214,
similarity=17.132406]

3: 4996 Brabham
[pem_pos=0, final_scores=0.34354478,
final_local_scores=0.3316158, pem=0.0, similarity=13.98579]

4: 31717 United_Kingdom
[pem_pos=0, final_scores=0.49369282,
final_local_scores=0.5508099, pem=-0.48936236,
similarity=23.188208]
5: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.39676112,
final_local_scores=0.3695745, pem=-0.64885646,
similarity=18.647135]
6: 4996 Brabham
[pem_pos=0, final_scores=0.40938824,
final_local_scores=0.40955347, pem=0.0, similarity=16.311605]
7: 1219709 Carlos_Pace
[pem_pos=0, final_scores=0.2562716,
final_local_scores=0.24775967, pem=0.0, similarity=11.4833555]
1: 34495595 Goldie_Hexagon_Racing
[pem_pos=0, final_scores=0.071542755,
final_local_scores=0.08783102, pem=0.0, similarity=6.710766]
2: 10854 Formula_One
[pem_pos=0, final_scores=0.28193492,
final_local_scores=0.23883301, pem=-0.08942214,
similarity=11.7032585]
3: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.43187445,
final_local_scores=0.40793055, pem=-0.64885646,
similarity=19.791756]
4: 34495595 Goldie_Hexagon_Racing
[pem_pos=0, final_scores=0.07128094,
final_local_scores=0.08754501, pem=0.0, similarity=6.7022305]
5: 10854 Formula_One
[pem_pos=0, final_scores=0.46229893,
final_local_scores=0.4358533, pem=-0.08942214,
similarity=17.582737]
6: 30874971 John_Watson_(racing_driver)
[pem_pos=0, final_scores=0.41505206,
final_local_scores=0.39206535, pem=-0.64885646,
similarity=19.318306]
7: 18951905 Argentina
[pem_pos=0, final_scores=0.35472822,
final_local_scores=0.42541307, pem=0.0, similarity=16.784887]
8: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.42922735,
final_local_scores=0.4050148, pem=-0.64885646,
similarity=19.704744]
9: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.19065939,
final_local_scores=0.14695045, pem=-0.64885646,
similarity=12.00359]
10: 17416221 South_Africa
[pem_pos=0, final_scores=0.35396725,
final_local_scores=0.44421232, pem=-0.26251382,
similarity=18.773485]
11: 26667 Spain

[pem_pos=0, final_scores=0.08884616,
final_local_scores=0.13662681, pem=-0.18857051,
similarity=9.192405]
12: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.19999544,
final_local_scores=0.15463793, pem=-0.64885646,
similarity=12.233]
13: 3343 Belgium
[pem_pos=0, final_scores=0.59269506,
final_local_scores=0.66286165, pem=-0.24180195,
similarity=25.185783]
14: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.132708,
final_local_scores=0.081136614, pem=-0.64885646,
similarity=10.039574]
15: 34495595 Goldie_Hexagon_Racing
[pem_pos=0, final_scores=0.15164156,
final_local_scores=0.1753268, pem=0.0, similarity=9.321815]
16: 5058739 Sweden
[pem_pos=0, final_scores=0.18511046,
final_local_scores=0.24777249, pem=-0.2135508,
similarity=12.645061]
17: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.17785065,
final_local_scores=0.13295889, pem=-0.64885646,
similarity=11.586054]
18: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.21361731,
final_local_scores=0.16951776, pem=-0.64885646,
similarity=12.677043]
19: 5843419 France
[pem_pos=0, final_scores=0.43424094,
final_local_scores=0.49551028, pem=-0.16240834,
similarity=19.75993]
20: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.29986423,
final_local_scores=0.26372936, pem=-0.64885646,
similarity=15.488505]
21: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.21775822,
final_local_scores=0.17655182, pem=-0.64885646,
similarity=12.886953]
22: 1219709 Carlos_Pace
[pem_pos=0, final_scores=0.33477348,
final_local_scores=0.33351105, pem=0.0, similarity=14.042348]
23: 26964606 Austria
[pem_pos=0, final_scores=0.636746,
final_local_scores=0.72913, pem=-0.25262418,
similarity=27.222218]
24: 30874971 John_Watson_(racing_driver)
[pem_pos=1, final_scores=0.42937267,
final_local_scores=0.40519768, pem=-0.64885646,

```

similarity=19.710201]
25: 14532 Italy
    [pem_pos=0, final_scores=0.35493547,
final_local_scores=0.4183895, pem=-0.20612453,
similarity=17.696228]
26: 30874971 John_Watson_(racing_driver)
    [pem_pos=1, final_scores=0.21289419,
final_local_scores=0.16872784, pem=-0.64885646,
similarity=12.653471]
27: 3434750 United_States
    [pem_pos=0, final_scores=0.3761242,
final_local_scores=0.4352697, pem=-0.13038418,
similarity=17.78808]

```

Sample: b'Harry_Taylor'

_ -LRB- rugby_league -RRB- William Henry `` Harry '' Taylor was an [1 1 English 1] professional 2 rugby league footballer of the 1890s , 1900s and 1910s , playing at representative level for 3 Great Britain -LRB- 4 Captain -RRB- , [2 5 England 2] , and [3 6 Yorkshire 3] , and at club level for [4 7 Hull 4] , as a , i.e. number 1 , making his début in the 5-10 away defeat to 8 Bradford on 1 January 1898 , and was 9 captain of [5 Hull 5] during the 10 1902 -- 03 , 11 1903 -- 04 , 12 1907 -- 08 , and 13 1907 -- 09 seasons . and at club level coached for [6 Hull 6] . Harry Taylor won caps for [7 14 England 7] while at [8 Hull 8] in 1908 against [9 Wales 9] , and [10 New Zealand 10] , and won 3 caps for 15 Great Britain while at [11 Hull 11] in 1908 against [12 New Zealand 12] -LRB- 3 matches -RRB- in which he was captain in all 3 tests . At , [1 1 Leeds 1] on Saturday 25 January 1908 , he was the against [13 16 New Zealand 13] in the first ever 17 Test for Great Britain . Harry captained [14 18 Hull 14] in the [2 2 Challenge Cup 2] final in 1908 and 1909 , which was lost to [3 3 Hunslet 3] 0-14 at [15 Fartown Ground 15] , [16 Huddersfield 16] and [4 4 Wakefield Trinity 4] 0-17 at , [5 5 Leeds 5] , respectively . He played in the 7-7 draw in 1910 against [17 19 Leeds 17] at [18 Fartown Ground 18] , [19 Huddersfield 19] , but did not play in the Cup Final replay which was lost 12-26 , again at [20 Fartown Ground 20] , [21 Huddersfield 21] .

```

1: 8262427 Leeds
2: 21050680 Challenge_Cup
3: 3165161 Hunslet_Hawks
4: 1095973 Wakefield_Trinity_Wildcats
5: 8262427 Leeds

```

```

1: 18803164 English_people
    [winner: 9316 England, pem_pos=1,
final_scores=0.64623255, final_local_scores=0.6699385,
pem=-1.3527626, similarity=31.438547]
    gt: 18803164 English_people, pem_pos=2,
final_scores=-0.470195, final_local_scores=-0.5164208,
pem=-2.2084396, similarity=0.68852454 ]
2: 25735 Rugby_league

```

[winner: 19774201 Rugby_Virginia, pem_pos=7,
final_scores=-0.883521, final_local_scores=-0.7526033,
pem=-2.956099, similarity=-2.2937446|
gt: 25735 Rugby_league, pem_pos=14,
final_scores=-1.8916311, final_local_scores=-2.0963042,
pem=-3.3002021, similarity=-40.52117]

[winner: 21211 National_Football_League, pem_pos=1,
final_scores=-1.8989128, final_local_scores=-2.057844,
pem=-1.6728458, similarity=-48.22326|
gt: 25735 Rugby_league, pem_pos=5,
final_scores=-2.0654728, final_local_scores=-2.2861998,
pem=-3.5378149, similarity=-44.895863]

3: 951941 Great_Britain_national_rugby_league_team
[winner: 951941 Great_Britain_national_rugby_league_team,
pem_pos=2, final_scores=-0.17261454,
final_local_scores=-0.22168311, pem=-2.6635387,
similarity=11.958984|
gt: 951941 Great_Britain_national_rugby_league_team,
pem_pos=2, final_scores=-0.17261454,
final_local_scores=-0.22168311, pem=-2.6635387,
similarity=11.958984]

4: 1874452 Captain_(sports)
[winner: 6461120 Captain_(United_States), pem_pos=0,
final_scores=-0.93284625, final_local_scores=-0.95385396,
pem=-1.5108879, similarity=-16.158749|
gt: 1874452 Captain_(sports), pem_pos=10,
final_scores=-1.4369824, final_local_scores=-1.5941768,
pem=-3.53238, similarity=-24.27407]

5: 990422 England_national_rugby_league_team
[winner: 9316 England, pem_pos=0,
final_scores=0.21996938, final_local_scores=0.20431075,
pem=-0.19053856, similarity=11.222931|
gt: 990422 England_national_rugby_league_team,
pem_pos=5, final_scores=-0.52093107,
final_local_scores=-0.6097272, pem=-4.237946,
similarity=8.940846]

6: 493717 Rugby_League_War_of_the_Roses
[winner: 869945 Yorkshire_County_Cricket_Club, pem_pos=1,
final_scores=0.07814403, final_local_scores=0.045645386,
pem=-1.7779728, similarity=15.120754|
gt: 493717 Rugby_League_War_of_the_Roses, pem_pos=3,
final_scores=-0.29787272, final_local_scores=-0.34771973,
pem=-3.1132855, similarity=10.643594]

7: 1095938 Hull_F.C.
[winner: 451358 Hull_City_A.F.C., pem_pos=11,
final_scores=0.1137694, final_local_scores=0.040640235,
pem=-3.7682815, similarity=25.795|
gt: 1095938 Hull_F.C., pem_pos=1,
final_scores=-0.12998973, final_local_scores=-0.21215189,
pem=-2.238242, similarity=9.930585]

8: 23739721 Bradford_Park_Avenue_A.F.C.
[winner: 1095011 Bradford_Bulls, pem_pos=9,

final_scores=-0.15158309, final_local_scores=-0.23489785,
pem=-3.818427, similarity=17.845093|
RECALL MISS: 23739721 Bradford_Park_Avenue_A.F.C.]
9: 1874452 Captain_(sports)
[winner: 3319715 Australian_Army_officer_rank_insignia,
pem_pos=11, final_scores=-1.5290124,
final_local_scores=-1.508954, pem=-2.7026498,
similarity=-26.243052|
gt: 1874452 Captain_(sports), pem_pos=2,
final_scores=-1.6282052, final_local_scores=-1.8030586,
pem=-1.7624387, similarity=-40.132736]
10: 23823439 1902-03_Northern_Rugby_Football_Union_season
[winner: 3453736 FabricLive.03, pem_pos=2,
final_scores=-0.8160016, final_local_scores=-0.6530472,
pem=-2.3070445, similarity=-2.8524516|
RECALL MISS: 23823439 1902-
03_Northern_Rugby_Football_Union_season]
[winner: 34652 1902, pem_pos=0, final_scores=-0.8918889,
final_local_scores=-0.79242367, pem=-0.78011304,
similarity=-15.3154125|
RECALL MISS: 23823439 1902-
03_Northern_Rugby_Football_Union_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3534744, final_local_scores=-1.246913,
pem=0.0, similarity=-33.120655|
RECALL MISS: 23823439 1902-
03_Northern_Rugby_Football_Union_season]
11: 23825984 1903-04_Northern_Rugby_Football_Union_season
[winner: 12641051 Now_04_(Australian_series), pem_pos=0,
final_scores=-0.59336936, final_local_scores=-0.43990865,
pem=-1.6957035, similarity=0.18346114|
RECALL MISS: 23825984 1903-
04_Northern_Rugby_Football_Union_season]
[winner: 34639 1903, pem_pos=0, final_scores=-1.0461375,
final_local_scores=-0.97119594, pem=-0.6820626,
similarity=-21.183548|
RECALL MISS: 23825984 1903-
04_Northern_Rugby_Football_Union_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3905377, final_local_scores=-1.287399,
pem=0.0, similarity=-34.328846|
RECALL MISS: 23825984 1903-
04_Northern_Rugby_Football_Union_season]
12: 23877335 1907-08_Northern_Rugby_Football_Union_season
[winner: 579790 British_Rail_Class_08, pem_pos=2,
final_scores=-0.7818421, final_local_scores=-0.7571454,
pem=-2.272334, similarity=-6.1477094|
RECALL MISS: 23877335 1907-
08_Northern_Rugby_Football_Union_season]
[winner: 34649 1907, pem_pos=0, final_scores=-1.0061466,
final_local_scores=-0.9174123, pem=-0.65885776,
similarity=-19.70473|

RECALL MISS: 23877335 1907-
 08_Northern_Rugby_Football_Union_season]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.27811, final_local_scores=-1.1645889,
 pem=0.0, similarity=-30.663946]
 RECALL MISS: 23877335 1907-
 08_Northern_Rugby_Football_Union_season]
 13: 23890924 1908-09_Northern_Rugby_Football_Union_season
 [winner: 35983 2009, pem_pos=0, final_scores=-0.82913065,
 final_local_scores=-0.7280261, pem=-1.8222004,
 similarity=-7.726629]
 RECALL MISS: 23890924 1908-
 09_Northern_Rugby_Football_Union_season]
 [winner: 34649 1907, pem_pos=0, final_scores=-1.0336154,
 final_local_scores=-0.9474178, pem=-0.65885776,
 similarity=-20.600153]
 RECALL MISS: 23890924 1908-
 09_Northern_Rugby_Football_Union_season]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.2993599, final_local_scores=-1.1878011,
 pem=0.0, similarity=-31.356642]
 RECALL MISS: 23890924 1908-
 09_Northern_Rugby_Football_Union_season]
 14: 990422 England_national_rugby_league_team
 [winner: 9904 England_national_football_team, pem_pos=1,
 final_scores=0.17202316, final_local_scores=0.10598582,
 pem=-2.776464, similarity=22.351383]
 gt: 990422 England_national_rugby_league_team,
 pem_pos=5, final_scores=-0.40232813,
 final_local_scores=-0.47959775, pem=-4.237946,
 similarity=12.824169]
 15: 951941 Great_Britain_national_rugby_league_team
 [winner: 951941 Great_Britain_national_rugby_league_team,
 pem_pos=2, final_scores=0.013040885,
 final_local_scores=-0.018882751, pem=-2.6635387,
 similarity=18.010952]
 gt: 951941 Great_Britain_national_rugby_league_team,
 pem_pos=2, final_scores=0.013040885,
 final_local_scores=-0.018882751, pem=-2.6635387,
 similarity=18.010952]
 16: 952064 New_Zealand_national_rugby_league_team
 [winner: 154957 New_Zealand_national_rugby_union_team,
 pem_pos=4, final_scores=0.3664053,
 final_local_scores=0.38009977, pem=-3.8192244,
 similarity=36.202183]
 gt: 952064 New_Zealand_national_rugby_league_team,
 pem_pos=3, final_scores=-0.21615912,
 final_local_scores=-0.25756526, pem=-3.804206,
 similarity=17.091314]
 17: 16926380 Test_match_(rugby_league)
 [winner: 30863 Test_cricket, pem_pos=0,
 final_scores=-0.13945663, final_local_scores=-0.16127431,

pem=-0.23950954, similarity=0.57945585|
RECALL MISS: 16926380 Test_match_(rugby_league)]
18: 1095938 Hull_F.C.
[winner: 451358 Hull_City_A.F.C., pem_pos=11,
final_scores=0.11653747, final_local_scores=0.04366392,
pem=-3.7682815, similarity=25.885233|
gt: 1095938 Hull_F.C., pem_pos=1,
final_scores=-0.14625224, final_local_scores=-0.22991621,
pem=-2.238242, similarity=9.400462]
19: 1095024 Leeds_Rhinos
[winner: 7609513 Leeds_United_A.F.C., pem_pos=4,
final_scores=0.32816964, final_local_scores=0.27026588,
pem=-3.421221, similarity=30.760117|
gt: 1095024 Leeds_Rhinos, pem_pos=1,
final_scores=0.11103319, final_local_scores=0.05771771,
pem=-2.7286434, similarity=20.650911]
1: 8262427 Leeds
[pem_pos=0, final_scores=0.3136406,
final_local_scores=0.29020324, pem=-0.25452003,
similarity=14.134076]
2: 21050680 Challenge_Cup
[pem_pos=0, final_scores=0.23671322,
final_local_scores=0.19931445, pem=-0.38845974,
similarity=12.150158]
3: 3165161 Hunslet_Hawks
[pem_pos=1, final_scores=0.19211723,
final_local_scores=0.14452282, pem=-0.50400555,
similarity=11.143422]
4: 1095973 Wakefield_Trinity_Wildcats
[pem_pos=0, final_scores=0.15376444,
final_local_scores=0.107951134, pem=0.0, similarity=7.31119]
5: 8262427 Leeds
[pem_pos=0, final_scores=0.107421204,
final_local_scores=0.064939916, pem=-0.25452003,
similarity=7.411769]
1: 9316 England
[pem_pos=1, final_scores=0.64623255,
final_local_scores=0.6699385, pem=-1.3527626,
similarity=31.438547]
2: 9316 England
[pem_pos=0, final_scores=0.21996938,
final_local_scores=0.20431075, pem=-0.19053856,
similarity=11.222931]
3: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.07814403,
final_local_scores=0.045645386, pem=-1.7779728,
similarity=15.120754]
4: 451358 Hull_City_A.F.C.
[pem_pos=11, final_scores=0.1137694,
final_local_scores=0.040640235, pem=-3.7682815,
similarity=25.795]
5: 451358 Hull_City_A.F.C.

[pem_pos=11, final_scores=0.25959814,
final_local_scores=0.19993597, pem=-3.7682815,
similarity=30.5487]
6: 451358 Hull_City_A.F.C.
[pem_pos=11, final_scores=0.17437918,
final_local_scores=0.10684723, pem=-3.7682815,
similarity=27.770748]
7: 9904 England_national_football_team
[pem_pos=1, final_scores=0.17202316,
final_local_scores=0.10598582, pem=-2.776464,
similarity=22.351383]
8: 451358 Hull_City_A.F.C.
[pem_pos=11, final_scores=0.09735601,
final_local_scores=0.022711098, pem=-3.7682815,
similarity=25.259958]
9: 580291 Wales_national_football_team
[pem_pos=1, final_scores=0.21147095,
final_local_scores=0.15503058, pem=-2.4526134,
similarity=22.053827]
10: 154957 New_Zealand_national_rugby_union_team
[pem_pos=4, final_scores=0.3659919,
final_local_scores=0.3796482, pem=-3.8192244,
similarity=36.18871]
11: 451358 Hull_City_A.F.C.
[pem_pos=11, final_scores=0.1856231,
final_local_scores=0.11912948, pem=-3.7682815,
similarity=28.137276]
12: 154957 New_Zealand_national_rugby_union_team
[pem_pos=4, final_scores=0.36760348,
final_local_scores=0.38140857, pem=-3.8192244,
similarity=36.24124]
13: 154957 New_Zealand_national_rugby_union_team
[pem_pos=4, final_scores=0.3664053,
final_local_scores=0.38009977, pem=-3.8192244,
similarity=36.202183]
14: 451358 Hull_City_A.F.C.
[pem_pos=11, final_scores=0.11653747,
final_local_scores=0.04366392, pem=-3.7682815,
similarity=25.885233]
15: 463412 Fartown_Ground
[pem_pos=0, final_scores=0.21777667,
final_local_scores=0.21863654, pem=0.0, similarity=10.614264]
16: 236211 Huddersfield_Town_F.C.
[pem_pos=3, final_scores=0.3077867,
final_local_scores=0.24458805, pem=-2.2859883,
similarity=23.820269]
17: 7609513 Leeds_United_A.F.C.
[pem_pos=4, final_scores=0.32816964,
final_local_scores=0.27026588, pem=-3.421221,
similarity=30.760117]
18: 463412 Fartown_Ground
[pem_pos=0, final_scores=0.18946384,

final_local_scores=0.1877091, pem=0.0, similarity=9.691327]
 19: 236211 Huddersfield_Town_F.C.
 [pem_pos=3, final_scores=0.28540558,
 final_local_scores=0.22014001, pem=-2.2859883,
 similarity=23.09069]
 20: 463412 Fartown_Ground
 [pem_pos=0, final_scores=0.14977373,
 final_local_scores=0.14435366, pem=0.0, similarity=8.397514]
 21: 152475 Huddersfield
 [pem_pos=0, final_scores=0.17112778,
 final_local_scores=0.13907123, pem=-0.081461765,
 similarity=8.682878]

Sample: b'South_East_Lancashire'

-LRB- UK_Parliament_constituency -RRB- South East
 Lancashire was a of the of the 1 Parliament of the [1 United K
 ingdom 1] . It was represented by two 2 Members of Parliament
 . The constituency was created by the 3 Reform act of 1867 by
 the splitting of the [2 4 South Lancashire 2] constituency
 into South-West and South-East divisions . The constituency
 was abolished by the 5 Redistribution of Seats Act 1885 ,
 being divided into eight single member divisions of 6 Eccles ,
 7 Radcliffe-cum-Farnworth , 8 Gorton , 9 Heywood , 10 Middle
 ton , 11 Prestwich , [3 12 Stretford 3] and [4 13 Westhoughton
 4] . This constituency comprised the hundred of [5 Lancashire
 5] except for those parts of the hundred lying in the of [6 14
 Ashton-under-Lyne 6] , [7 15 Bolton 7] , [8 16 Bury 8] , [9 17
 Manchester 9] , [10 18 Rochdale 10] , [11 19 Salford 11] and [12
 20 Stalybridge 12]

1: 13964 Parliament_of_the_United_Kingdom
 [winner: 31717 United_Kingdom, pem_pos=0,
 final_scores=0.28211004, final_local_scores=0.29183844,
 pem=-0.09933331, similarity=13.3389435]
 gt: 13964 Parliament_of_the_United_Kingdom, pem_pos=16,
 final_scores=-0.57457584, final_local_scores=-0.63916075,
 pem=-4.3771396, similarity=8.8194475]
 [winner: 24406 Parliament, pem_pos=1,
 final_scores=-0.73577344, final_local_scores=-0.75333947,
 pem=-1.8263724, similarity=-8.459343]
 gt: 13964 Parliament_of_the_United_Kingdom, pem_pos=0,
 final_scores=-0.8505272, final_local_scores=-0.9453964,
 pem=-1.7932186, similarity=-14.371]
 [winner: 1961717 Supreme_Court_of_South_Australia,
 pem_pos=0, final_scores=-1.2068497,
 final_local_scores=-1.2641997, pem=0.0,
 similarity=-33.636528]
 RECALL MISS: 13964 Parliament_of_the_United_Kingdom]
 2: 98066 Member_of_Parliament
 [winner: 98066 Member_of_Parliament, pem_pos=0,
 final_scores=-0.37871498, final_local_scores=-0.43082502,
 pem=-0.014730592, similarity=-8.686858]
 gt: 98066 Member_of_Parliament, pem_pos=0,

final_scores=-0.37871498, final_local_scores=-0.43082502,
 pem=-0.014730592, similarity=-8.686858]
 3: 547757 Reform_Act_1867
 [winner: 26036 Reform_Judaism, pem_pos=0,
 final_scores=-0.18966734, final_local_scores=-0.14903376,
 pem=-1.0147479, similarity=5.160605|
 RECALL MISS: 547757 Reform_Act_1867]
 [winner: 34787 1867, pem_pos=0, final_scores=-0.7209299,
 final_local_scores=-0.54278624, pem=-0.23064093,
 similarity=-10.85385|
 gt: 547757 Reform_Act_1867, pem_pos=16,
 final_scores=-1.310131, final_local_scores=-1.4351954,
 pem=-3.1764655, similarity=-21.465265]
 [winner: 23572029 Suppression_of_Heresy_Act_1414,
 pem_pos=18, final_scores=-1.4719434,
 final_local_scores=-1.4436677, pem=-3.2915623,
 similarity=-21.092182|
 RECALL MISS: 547757 Reform_Act_1867]
 4: 7818360 South_Lancashire_(UK_Parliament_constituency)
 [winner: 50435 Lancashire, pem_pos=1,
 final_scores=0.194001, final_local_scores=0.114157796,
 pem=-0.65742916, similarity=11.071611|
 gt: 7818360
 South_Lancashire_(UK_Parliament_constituency), pem_pos=0,
 final_scores=0.070912674, final_local_scores=0.0955676,
 pem=0.0, similarity=6.941641]
 [winner: 12174399
 Nigerian_National_Assembly_delegation_from_Katsina,
 pem_pos=14, final_scores=-1.4708694,
 final_local_scores=-1.3861759, pem=-3.2927792,
 similarity=-19.369894|
 RECALL MISS: 7818360
 South_Lancashire_(UK_Parliament_constituency)]
 5: 2821917 Redistribution_of_Seats_Act_1885
 [winner: 2821917 Redistribution_of_Seats_Act_1885,
 pem_pos=0, final_scores=-0.19253094,
 final_local_scores=-0.21825892, pem=0.0,
 similarity=-2.4235685|
 gt: 2821917 Redistribution_of_Seats_Act_1885, pem_pos=0,
 final_scores=-0.19253094, final_local_scores=-0.21825892,
 pem=0.0, similarity=-2.4235685]
 [winner: 34715 1885, pem_pos=0, final_scores=-0.9563015,
 final_local_scores=-0.8699424, pem=-0.4892705,
 similarity=-19.210375|
 gt: 2821917 Redistribution_of_Seats_Act_1885, pem_pos=9,
 final_scores=-1.6752033, final_local_scores=-1.8378532,
 pem=-3.2905319, similarity=-32.861065]
 6: 1811196 Eccles_(UK_Parliament_constituency)
 [winner: 1742688 Eccles_Greater_Manchester, pem_pos=0,
 final_scores=-0.2616604, final_local_scores=-0.358118,
 pem=-1.037052, similarity=-0.95759237|
 gt: 1811196 Eccles_(UK_Parliament_constituency),

pem_pos=4, final_scores=-0.5941787,
final_local_scores=-0.6357702, pem=-2.3725758,
similarity=-1.9805033]
7: 12378279 Radcliffe-cum-
Farnworth_(UK_Parliament_constituency)
[winner: 12378279 Radcliffe-cum-
Farnworth_(UK_Parliament_constituency), pem_pos=0,
final_scores=-0.12132096, final_local_scores=-0.072735436,
pem=0.0, similarity=1.9191424|
gt: 12378279 Radcliffe-cum-
Farnworth_(UK_Parliament_constituency), pem_pos=0,
final_scores=-0.12132096, final_local_scores=-0.072735436,
pem=0.0, similarity=1.9191424]
8: 1811251 Manchester_Gorton_(UK_Parliament_constituency)
[winner: 1414171 Gorton, pem_pos=0,
final_scores=0.0005633086, final_local_scores=-0.049154863,
pem=-0.59120226, similarity=5.8378835|
gt: 1811251
Manchester_Gorton_(UK_Parliament_constituency), pem_pos=6,
final_scores=-0.29942814, final_local_scores=-0.3095775,
pem=-2.718763, similarity=9.63636]
9: 8369764 Heywood_(UK_Parliament_constituency)
[winner: 1219535 Heywood, Greater_Manchester, pem_pos=0,
final_scores=-0.0010742545, final_local_scores=-0.06854321,
pem=-1.0117869, similarity=7.5465016|
gt: 8369764 Heywood_(UK_Parliament_constituency),
pem_pos=6, final_scores=-0.4991989,
final_local_scores=-0.5169281, pem=-2.8233256,
similarity=4.017231]
10: 8369835 Middleton_(UK_Parliament_constituency)
[winner: 1089547 Middleton, Greater_Manchester,
pem_pos=0, final_scores=-0.2258073,
final_local_scores=-0.31305656, pem=-1.8317558,
similarity=4.708853|
RECALL MISS: 8369835
Middleton_(UK_Parliament_constituency)]
11: 8369910 Prestwich_(UK_Parliament_constituency)
[winner: 1711362 Prestwich, pem_pos=0,
final_scores=-0.08056469, final_local_scores=-0.15378484,
pem=-0.026270874, similarity=-0.35666823|
gt: 8369910 Prestwich_(UK_Parliament_constituency),
pem_pos=2, final_scores=-0.37067345,
final_local_scores=-0.36915913, pem=-1.7545353,
similarity=2.6147037]
12: 7052621 Stretford_(UK_Parliament_constituency)
[winner: 1074030 Stretford, pem_pos=0,
final_scores=0.044207975, final_local_scores=-0.016619816,
pem=0.0, similarity=3.5937445|
gt: 7052621 Stretford_(UK_Parliament_constituency),
pem_pos=1, final_scores=-0.20189665,
final_local_scores=-0.22395003, pem=-1.2037575,
similarity=3.952818]

13: 7047930 Westhoughton_(UK_Parliament_constituency)
 [winner: 1069951 Westhoughton, pem_pos=0,
 final_scores=0.072450176, final_local_scores=0.039361164,
 pem=0.0, similarity=5.2643285|
 gt: 7047930 Westhoughton_(UK_Parliament_constituency),
 pem_pos=2, final_scores=-0.1428477,
 final_local_scores=-0.1569942, pem=-0.9693272,
 similarity=4.676045]

14: 1811178 Ashton-under-Lyne_(UK_Parliament_constituency)
 [winner: 308036 Ashton-under-Lyne, pem_pos=0,
 final_scores=0.3112682, final_local_scores=0.26443103,
 pem=-0.014164876, similarity=12.057894|
 gt: 1811178 Ashton-under-Lyne_(UK_Parliament_constituency), pem_pos=1,
 final_scores=-0.05879368, final_local_scores=-0.0582826,
 pem=-0.9682261, similarity=7.6158075]

15: 8369513 Bolton_(UK_Parliament_constituency)
 [winner: 352418 Bolton, pem_pos=0,
 final_scores=0.1572753, final_local_scores=0.08200839,
 pem=-0.7678214, similarity=10.712539|
 gt: 8369513 Bolton_(UK_Parliament_constituency),
 pem_pos=7, final_scores=-0.76436436,
 final_local_scores=-0.8198626, pem=-3.8248525,
 similarity=0.42352125]

16: 7388094 Bury_(UK_Parliament_constituency)
 [winner: 451156 Bury_F.C., pem_pos=0,
 final_scores=0.08063473, final_local_scores=0.09178941,
 pem=-0.72067994, similarity=10.748061|
 gt: 7388094 Bury_(UK_Parliament_constituency),
 pem_pos=6, final_scores=-0.6552835,
 final_local_scores=-0.6999676, pem=-3.4291272,
 similarity=1.8494095]

17: 8369424 Manchester_(UK_Parliament_constituency)
 [winner: 20206 Manchester, pem_pos=0,
 final_scores=0.30117077, final_local_scores=0.24625555,
 pem=-0.3209542, similarity=13.183868|
 RECALL MISS: 8369424
 Manchester_(UK_Parliament_constituency)]

18: 1820424 Rochdale_(UK_Parliament_constituency)
 [winner: 635367 Rochdale, pem_pos=0,
 final_scores=0.18147437, final_local_scores=0.1089717,
 pem=-0.7956193, similarity=11.668347|
 gt: 1820424 Rochdale_(UK_Parliament_constituency),
 pem_pos=3, final_scores=-0.3469937,
 final_local_scores=-0.36653012, pem=-2.3874807,
 similarity=6.135214]

19: 1798095 Salford_(UK_Parliament_constituency)
 [winner: 182231 Salford_Greater_Manchester, pem_pos=0,
 final_scores=0.10366003, final_local_scores=0.030604824,
 pem=-1.0645502, similarity=10.792212|
 gt: 1798095 Salford_(UK_Parliament_constituency),
 pem_pos=5, final_scores=-0.51154387,


```
final_local_scores=-0.535236, pem=-2.5821002,  
similarity=2.1590645 ]  
20: 4675605 Stalybridge_(UK_Parliament_constituency)  
    [winner: 166173 Stalybridge, pem_pos=0,  
final_scores=0.037459776, final_local_scores=-0.033003077,  
pem=0.0, similarity=3.1048353]  
    gt: 4675605 Stalybridge_(UK_Parliament_constituency),  
pem_pos=2, final_scores=-0.2670191,  
final_local_scores=-0.2657784, pem=-1.3431205,  
similarity=3.4624534 ]  
1: 31717 United_Kingdom  
    [pem_pos=0, final_scores=0.28211004,  
final_local_scores=0.29183844, pem=-0.09933331,  
similarity=13.3389435]  
2: 50435 Lancashire  
    [pem_pos=1, final_scores=0.194001,  
final_local_scores=0.114157796, pem=-0.65742916,  
similarity=11.071611]  
3: 1074030 Stretford  
    [pem_pos=0, final_scores=0.044207975,  
final_local_scores=-0.016619816, pem=0.0,  
similarity=3.5937445]  
4: 1069951 Westhoughton  
    [pem_pos=0, final_scores=0.072450176,  
final_local_scores=0.039361164, pem=0.0, similarity=5.2643285]  
5: 50435 Lancashire  
    [pem_pos=0, final_scores=0.30238056,  
final_local_scores=0.23334503, pem=-0.20611861,  
similarity=12.174099]  
6: 308036 Ashton-under-Lyne  
    [pem_pos=0, final_scores=0.3112682,  
final_local_scores=0.26443103, pem=-0.014164876,  
similarity=12.057894]  
7: 352418 Bolton  
    [pem_pos=0, final_scores=0.1572753,  
final_local_scores=0.08200839, pem=-0.7678214,  
similarity=10.712539]  
8: 451156 Bury_F.C.  
    [pem_pos=0, final_scores=0.08063473,  
final_local_scores=0.09178941, pem=-0.72067994,  
similarity=10.748061]  
9: 20206 Manchester  
    [pem_pos=0, final_scores=0.30117077,  
final_local_scores=0.24625555, pem=-0.3209542,  
similarity=13.183868]  
10: 635367 Rochdale  
    [pem_pos=0, final_scores=0.18147437,  
final_local_scores=0.1089717, pem=-0.7956193,  
similarity=11.668347]  
11: 182231 Salford,_Greater_Manchester  
    [pem_pos=0, final_scores=0.10366003,  
final_local_scores=0.030604824, pem=-1.0645502,
```

similarity=10.792212]
12: 166173 Stalybridge
[pem_pos=0, final_scores=0.037459776,
final_local_scores=-0.033003077, pem=0.0,
similarity=3.1048353]

Sample: b'Thomas_Land'

_ -LRB- Drayton_Manor -RRB- [1 Thomas Land 1] is a themed area at [1 1 Drayton Manor Theme Park 1] in [2 2 Staffordshire 2] , [2 England 2] , based on the TV series . It stands on the former site of Robinsons Land , a themed area sponsored by the soft drinks brand of the 1 same name . The construction began while the park was still open in September 2007 , and many of the Robinsons Land rides were moved to other areas of the park for the remainder of the season . All but the Veteran Cars -LRB- now Sodor 's [3 Classic Cars 3] -RRB- were removed and sold to Funland at [3 3 Hayling Island 3] . [4 Thomas Land 4] officially opened to the public on 15 March 2008 . 2 Narrow gauge models of [5 3 Thomas 5] , 4 Percy and are used on a miniature railway that takes visitors through [6 Thomas Land 6] . The line features , where stationary models are kept , 5 Knapford Station , which serves as the main station and a refreshment stand , Dryaw , and Tidmouth Hault . It also has many sidings and sheds along the way , one of which carries a sleeping model of 6 James . In early 2009 , [4 4 HIT Entertainment 4] announced that the original models from the 7 television series would be put on display at [7 Thomas Land 7] in an exhibit called `` Discover Thomas & Friends '' , as production was to move to [8 Canada 8] , now that the series is to be animated with 8 CGI . The models were then put on display in 2010 . So far , the layout features Great Waterton , a town that first appeared in '' 9 The Great Discovery '' , the Fishing Village and [9 Brendam Docks 9] . Many other buildings have been seen from previous seasons from the show , such as `` Neptune 's Refreshments '' , `` Ben 's Books '' , Toby 's Shed and Bluff 's Cove signal box . Models have also appeared in the workshop , including : , 10 Terence , 11 Trevor , Elizabeth , Butch , 3 Horrid Lorries , a canal boat , , , 2 red and white express carriages , 2 green and cream express carriages , 12 Thomas , 13 Edward , 14 Toby

1: 1520959 Drayton_Manor_Theme_Park

2: 58733 Staffordshire

3: 13859 Hayling_Island

4: 369198 HIT_Entertainment

1: 5989312 Robinsons_(drink)

[winner: 26517843 Carlos_Trillo_Name, pem_pos=9,
final_scores=-0.87059784, final_local_scores=-0.76120645,
pem=-2.3025851, similarity=-6.104389]

RECALL MISS: 5989312 Robinsons_(drink)]

[winner: 1363640 300_(comics), pem_pos=0,
final_scores=-1.1152979, final_local_scores=-1.0949085,

pem=0.0, similarity=-28.584541|
RECALL MISS: 5989312 Robinsons_(drink)]
2: 21932 Narrow_gauge_railway
[winner: 476512 Narrow_Street, pem_pos=4,
final_scores=-0.54938793, final_local_scores=-0.5613067,
pem=-1.7917595, similarity=-2.9169328|
gt: 21932 Narrow_gauge_railway, pem_pos=0,
final_scores=-0.6182039, final_local_scores=-0.6058303,
pem=-0.47957304, similarity=-11.381479]
[winner: 30860417 Needle_gauge_comparison_chart,
pem_pos=7, final_scores=-1.1722193,
final_local_scores=-1.0923922, pem=-2.360724,
similarity=-15.671465|
gt: 21932 Narrow_gauge_railway, pem_pos=5,
final_scores=-1.5105937, final_local_scores=-1.5806305,
pem=-2.2368073, similarity=-30.915348]
3: 19357739 Thomas_the_Tank_Engine
[winner: 46758 Edward_Thomas_(poet), pem_pos=0,
final_scores=0.15527867, final_local_scores=0.2209566,
pem=0.0, similarity=10.683498|
RECALL MISS: 19357739 Thomas_the_Tank_Engine]
4: 1042770 Percy_the_Small_Engine
[winner: 1042770 Percy_the_Small_Engine, pem_pos=1,
final_scores=-0.24506108, final_local_scores=-0.25213605,
pem=-2.0153356, similarity=7.525178|
gt: 1042770 Percy_the_Small_Engine, pem_pos=1,
final_scores=-0.24506108, final_local_scores=-0.25213605,
pem=-2.0153356, similarity=7.525178]
5: 3809802 Knapford
[winner: 3809802 Knapford, pem_pos=0,
final_scores=-0.2413949, final_local_scores=-0.12322918,
pem=0.0, similarity=0.41230834|
gt: 3809802 Knapford, pem_pos=0,
final_scores=-0.2413949, final_local_scores=-0.12322918,
pem=0.0, similarity=0.41230834]
6: 1057312 James_the_Red_Engine
[winner: 528805 James_(band), pem_pos=0,
final_scores=-0.2155215, final_local_scores=-0.17345867,
pem=-1.6834015, similarity=8.06796|
gt: 1057312 James_the_Red_Engine, pem_pos=8,
final_scores=-0.57334745, final_local_scores=-0.5979072,
pem=-4.008076, similarity=8.043511]
7: 231475 Thomas_&_Friends
[winner: 19508643 Television_program, pem_pos=0,
final_scores=-1.1652569, final_local_scores=-1.2451274,
pem=-0.13869494, similarity=-32.31313|
RECALL MISS: 231475 Thomas_&_Friends]
8: 31626763 Computer-generated_imagery
[winner: 31626763 Computer-generated_imagery, pem_pos=0,
final_scores=-0.0038725436, final_local_scores=0.029184878,
pem=-0.20161216, similarity=6.057046|
gt: 31626763 Computer-generated_imagery, pem_pos=0,

```

final_scores=-0.0038725436, final_local_scores=0.029184878,
pem=-0.20161216, similarity=6.057046 ]
9: 12331162 The_Great_Discovery
  [winner: 12331162 The_Great_Discovery, pem_pos=0,
final_scores=-0.13466223, final_local_scores=-0.10877774,
pem=0.0, similarity=0.8435682]
  gt: 12331162 The_Great_Discovery, pem_pos=0,
final_scores=-0.13466223, final_local_scores=-0.10877774,
pem=0.0, similarity=0.8435682 ]
10: 1112013 Terence_the_Tractor
  [winner: 30756 Terence, pem_pos=0,
final_scores=-0.35530287, final_local_scores=-0.28096154,
pem=-0.13038038, similarity=-3.585711]
  gt: 1112013 Terence_the_Tractor, pem_pos=2,
final_scores=-0.76605254, final_local_scores=-0.77304506,
pem=-3.1749837, similarity=-1.7134387 ]
11: 1948836 Trevor_the_Traction_Engine
  [winner: 462912 Trevor_Chappell, pem_pos=7,
final_scores=-0.40434468, final_local_scores=-0.33537167,
pem=-3.6336124, similarity=13.841702]
  gt: 1948836 Trevor_the_Traction_Engine, pem_pos=2,
final_scores=-0.74988693, final_local_scores=-0.76253784,
pem=-2.2218184, similarity=-6.583342 ]
12: 19357739 Thomas_the_Tank_Engine
  [winner: 1111005 Oldfield_Thomas, pem_pos=0,
final_scores=-0.40208906, final_local_scores=-0.32479268,
pem=-1.9636645, similarity=5.0759645]
  RECALL MISS: 19357739 Thomas_the_Tank_Engine]
13: 1057308 Edward_the_Blue_Engine
  [winner: 18835362 Edward_VIII, pem_pos=0,
final_scores=-0.3885324, final_local_scores=-0.39258194,
pem=-1.8709836, similarity=2.548984]
  gt: 1057308 Edward_the_Blue_Engine, pem_pos=6,
final_scores=-0.8663638, final_local_scores=-0.9358636,
pem=-3.6069622, similarity=-4.2231026 ]
14: 1338265 Toby_the_Tram_Engine
  [winner: 8025663 Toby_Flenderon, pem_pos=0,
final_scores=-0.4754893, final_local_scores=-0.4133703,
pem=-1.6262579, similarity=0.5977616]
  gt: 1338265 Toby_the_Tram_Engine, pem_pos=3,
final_scores=-0.6145394, final_local_scores=-0.6521767,
pem=-2.0361786, similarity=-4.2994843 ]
1: 1520959 Drayton_Manor_Theme_Park
  [pem_pos=0, final_scores=0.32154733,
final_local_scores=0.33037642, pem=0.0, similarity=13.948804]
2: 58733 Staffordshire
  [pem_pos=0, final_scores=0.45527625,
final_local_scores=0.5016479, pem=-0.06319042,
similarity=19.403526]
3: 13859 Hayling_Island
  [pem_pos=0, final_scores=0.3227778,
final_local_scores=0.35867587, pem=0.0, similarity=14.793316]

```

4: 369198 HIT_Entertainment
 [pem_pos=0, final_scores=0.22688542,
 final_local_scores=0.26144397, pem=0.0, similarity=11.891723]
 1: 19804219 Thomas_Land_(Drayton_Manor)
 [pem_pos=0, final_scores=0.06629373,
 final_local_scores=0.03484504, pem=0.0, similarity=5.1295586]
 2: 9316 England
 [pem_pos=0, final_scores=0.5783771,
 final_local_scores=0.62539434, pem=-0.19053856,
 similarity=23.788906]
 3: 338360 Classic_car
 [pem_pos=0, final_scores=0.039706454,
 final_local_scores=0.081294745, pem=0.0, similarity=6.5157104]
 4: 19804219 Thomas_Land_(Drayton_Manor)
 [pem_pos=0, final_scores=0.12974499,
 final_local_scores=0.10415591, pem=0.0, similarity=7.197933]
 5: 46758 Edward_Thomas_(poet)
 [pem_pos=0, final_scores=0.15527867,
 final_local_scores=0.2209566, pem=0.0, similarity=10.683498]
 6: 19804219 Thomas_Land_(Drayton_Manor)
 [pem_pos=0, final_scores=0.17438464,
 final_local_scores=0.15291795, pem=0.0, similarity=8.65309]
 7: 19804219 Thomas_Land_(Drayton_Manor)
 [pem_pos=0, final_scores=0.13461734,
 final_local_scores=0.11553255, pem=0.0, similarity=7.5374346]
 8: 5042916 Canada
 [pem_pos=0, final_scores=0.54773694,
 final_local_scores=0.6427671, pem=-0.11662793,
 similarity=23.905405]
 9: 6197379 Fictional_locations_in_Thomas_&_Friends
 [pem_pos=0, final_scores=0.07144593,
 final_local_scores=0.09736997, pem=0.0, similarity=6.995427]

Sample: b'Zielona_Góra'

_ -LRB- parliamentary_constituency -RRB- [1 Zielona Góra 1]
 is a [1 1 Polish 1] that is coterminous with the [2 2 Lubusz V
 oivodeship 2] . It elects twelve members of the [2 Sejm 2]
 and three members of the Senate . The district has the number
 ' 8 ' and is named after the city of [3 3 Zielona Góra 3] .
 It includes the 1 counties of [3 2 Gorzów 3] , [4 3 Krosno Od
 rzańskie 4] , [5 4 Międzyrzecz 5] , [6 5 Nowa Sól 6] , 6 Słub
 ice , 7 Strzelce-Drezdenko , [7 8 Sulęcín 7] , [8 9 Świebodzin
 8] , 10 Wschowa , 11 Żagań , 12 Żary , and [9 13 Zielona Góra
 9] and the of [4 4 Gorzów Wielkopolski 4]
 1: 22936 Poland
 2: 75149 Lubusz_Voivodeship
 3: 74687 Zielona_Góra
 4: 296922 Gorzów_Wielkopolski
 1: 382207 Powiat
 [winner: 221378 Counties_of_Denmark, pem_pos=0,
 final_scores=-1.0775602, final_local_scores=-1.0734687,
 pem=0.0, similarity=-27.944733]

RECALL MISS: 382207 Powiat]

2: 396876 Gorzów County

[winner: 296922 Gorzów Wielkopolski, pem_pos=3,
final_scores=0.06675996, final_local_scores=-0.0013224185,
pem=-2.063693, similarity=15.272933]

gt: 396876 Gorzów County, pem_pos=2,
final_scores=-0.2707547, final_local_scores=-0.23951104,
pem=-1.9459101, similarity=7.5243874]

3: 396898 Krosno Odrzańskie County

[winner: 717388 Krosno Odrzańskie, pem_pos=0,
final_scores=0.39179415, final_local_scores=0.38935405,
pem=-0.20975207, similarity=16.849478]

gt: 396898 Krosno Odrzańskie County, pem_pos=1,
final_scores=-0.083433226, final_local_scores=-0.094077885,
pem=-1.8755937, similarity=11.482011]

4: 396881 Międzyrzecz County

[winner: 390509 Międzyrzecz, pem_pos=0,
final_scores=0.14134507, final_local_scores=0.096736476,
pem=-0.22557366, similarity=8.203227]

gt: 396881 Międzyrzecz County, pem_pos=1,
final_scores=-0.28295738, final_local_scores=-0.30369452,
pem=-1.9200342, similarity=5.4683075]

5: 396931 Nowa Sól County

[winner: 386901 Nowa Sól, pem_pos=0,
final_scores=0.15758125, final_local_scores=0.13101009,
pem=-0.2001606, similarity=9.08782]

gt: 396931 Nowa Sól County, pem_pos=1,
final_scores=-0.34446627, final_local_scores=-0.3474503,
pem=-1.8718021, similarity=3.9002542]

6: 396863 Słubice County

[winner: 656704 Słubice, pem_pos=0,
final_scores=-0.0709565, final_local_scores=-0.04863491,
pem=-0.14134355, similarity=3.4069984]

gt: 396863 Słubice County, pem_pos=1,
final_scores=-0.48991197, final_local_scores=-0.53801876,
pem=-2.7612212, similarity=3.0501106]

7: 396848 Strzelce-Drezdenko County

[winner: 396848 Strzelce-Drezdenko County, pem_pos=0,
final_scores=-0.075116426, final_local_scores=-0.07059167,
pem=0.0, similarity=1.9831165]

gt: 396848 Strzelce-Drezdenko County, pem_pos=0,
final_scores=-0.075116426, final_local_scores=-0.07059167,
pem=0.0, similarity=1.9831165]

8: 396893 Sulęcín County

[winner: 657859 Sulęcín, pem_pos=0,
final_scores=0.08498718, final_local_scores=0.07887527,
pem=-0.2739519, similarity=7.9333014]

gt: 396893 Sulęcín County, pem_pos=1,
final_scores=-0.35728297, final_local_scores=-0.38417163,
pem=-2.0657113, similarity=3.858924]

9: 396890 Świebodzin County

[winner: 386898 Świebodzin, pem_pos=0,

final_scores=0.1850494, final_local_scores=0.19141439,
pem=-0.15484726, similarity=10.643983|
gt: 396890 Świebodzin County, pem_pos=1,
final_scores=-0.44493735, final_local_scores=-0.47370243,
pem=-2.3944194, similarity=2.9747126]
10: 396925 Wschowa County
[winner: 2931642 Wschowa, pem_pos=0,
final_scores=0.0132458955, final_local_scores=0.02862002,
pem=0.0, similarity=4.9437914|
gt: 396925 Wschowa County, pem_pos=1,
final_scores=-0.39264342, final_local_scores=-0.41399994,
pem=-0.8911313, similarity=-3.4187613]
11: 396900 Żagań County
[winner: 386899 Żagań, pem_pos=0,
final_scores=0.005610183, final_local_scores=-0.025844604,
pem=-0.14355381, similarity=4.0991263|
gt: 396900 Żagań County, pem_pos=1,
final_scores=-0.49846923, final_local_scores=-0.53777754,
pem=-2.2366807, similarity=0.20477691]
12: 396834 Żary County
[winner: 386902 Żary, pem_pos=0,
final_scores=-0.010633245, final_local_scores=-0.052367486,
pem=-0.23571242, similarity=3.8088036|
gt: 396834 Żary County, pem_pos=1,
final_scores=-0.5177879, final_local_scores=-0.55295193,
pem=-1.9031228, similarity=-2.0619974]
13: 396861 Zielona Góra County
[winner: 74687 Zielona Góra, pem_pos=0,
final_scores=0.2956118, final_local_scores=0.23541081,
pem=-0.07172165, similarity=11.504874|
gt: 396861 Zielona Góra County, pem_pos=8,
final_scores=-0.89760023, final_local_scores=-0.97817165,
pem=-5.1805615, similarity=3.0718174]
1: 22936 Poland
[pem_pos=0, final_scores=0.4847663,
final_local_scores=0.5020717, pem=-0.81518245,
similarity=23.505623]
2: 75149 Lubusz Voivodeship
[pem_pos=0, final_scores=0.16400407,
final_local_scores=0.15829939, pem=0.0, similarity=8.813683]
3: 74687 Zielona Góra
[pem_pos=0, final_scores=0.3201881,
final_local_scores=0.26225662, pem=-0.07172165,
similarity=12.306007]
4: 296922 Gorzów Wielkopolski
[pem_pos=0, final_scores=0.07944618,
final_local_scores=0.023342729, pem=-0.012637078,
similarity=4.8550286]
1: 74687 Zielona Góra
[pem_pos=0, final_scores=0.3001901,
final_local_scores=0.24041185, pem=-0.07172165,
similarity=11.654116]

2: 28167 Sejm
 [pem_pos=0, final_scores=0.08307175,
 final_local_scores=0.10947879, pem=0.0, similarity=7.356778]
 3: 296922 Gorzów Wielkopolski
 [pem_pos=3, final_scores=0.06675996,
 final_local_scores=-0.0013224185, pem=-2.063693,
 similarity=15.272933]
 4: 717388 Krosno Odrzańskie
 [pem_pos=0, final_scores=0.39179415,
 final_local_scores=0.38935405, pem=-0.20975207,
 similarity=16.849478]
 5: 390509 Międzyrzecz
 [pem_pos=0, final_scores=0.14134507,
 final_local_scores=0.096736476, pem=-0.22557366,
 similarity=8.203227]
 6: 386901 Nowa Sól
 [pem_pos=0, final_scores=0.15758125,
 final_local_scores=0.13101009, pem=-0.2001606,
 similarity=9.08782]
 7: 657859 Sulęcín
 [pem_pos=0, final_scores=0.08498718,
 final_local_scores=0.07887527, pem=-0.2739519,
 similarity=7.9333014]
 8: 386898 Świebodzin
 [pem_pos=0, final_scores=0.1850494,
 final_local_scores=0.19141439, pem=-0.15484726,
 similarity=10.643983]
 9: 74687 Zielona Góra
 [pem_pos=0, final_scores=0.2956118,
 final_local_scores=0.23541081, pem=-0.07172165,
 similarity=11.504874]

Sample: b'Derek_Morgan'

-LRB- cricketer -RRB- [1 Derek Clifton Morgan 1] -LRB-
 born 26 February 1929 -RRB- is a former [1 1 English 1] 1 cri
 cket er who played for [2 2 Derbyshire 2] between 1950 and
 1969 , captaining the side between 1965 and 1969 . An all-
 rounder , he is the only [2 Derbyshire 2] cricketer besides [3
 Leslie Townsend 3] to have achieved the double of 10,000 runs
 and 1,000 wickets . [4 Morgan 4] was born in [3 3 Muswell Hill
 3] , [5 London 5] . He made his first-class debut for [6
 Derbyshire 6] in the 2 1950 season against [4 4 Northamptonsh
 ire 4] in May , when he took two wickets in the first innings
 and was not out in his only batting innings . He played
 occasionally for the second XI but was soon established as a
 first team regular . In the 3 1951 season he took 6 for 93
 against [5 5 Gloucestershire 5] He made his first century of
 109 not out against [6 6 Kent 6] in the 4 1955 season . In
 the 5 1960 season he took 7 wickets for 38 against [7 7 Yorks
 hire 7] . In the 6 1962 season he scored centuries against [8
 8 Sussex 8] , [9 9 Somerset 9] and [10 10 Hampshire 10] . He
 scored 113 not out against [7 Pakistan Eaglets 7] in the 7 196

3 season and was nominated man of the match in the club 's first [11 11 Gillette Cup 11] match against [8 Hampshire 8] . In this season he achieved his best first class score of 147 against [9 Hampshire 9] and made 115 against [10 Somerset 10] . He was also nominated man of the match in the first round [1 1 Gillette Cup 11] game in the 8 1964 season , although [12 Derbyshire 12] lost . [13 Morgan 13] became [14 Derbyshire 14] captain in the 9 1965 season and that year achieved his best bowling performance of 7 for 33 against [12 12 Glamorgan 12] . In the 1965 [13 13 County Championship 13] , [15 Derbyshire 15] was twelfth , rising to eighth in the 10 1966 season . [16 Morgan 16] scored 103 not out against [17 Warwickshire 17] , but the team sank to bottom of the championship in 11 1967 , 12 1968 and 13 1969 , his last three years as captain . However they were runners up in the [18 Gillette Cup 18] in 1969 . [19 Morgan 19] finished his first-class career with 17,842 runs for [14 14 Derbyshire 14]

1: 9316 England

2: 1829984 Derbyshire_County_Cricket_Club

3: 94184 Muswell_Hill

4: 530308 Northamptonshire_County_Cricket_Club

5: 1622322 Gloucestershire_County_Cricket_Club

6: 1177719 Kent_County_Cricket_Club

7: 869945 Yorkshire_County_Cricket_Club

8: 1763691 Sussex_County_Cricket_Club

9: 1622178 Somerset_County_Cricket_Club

10: 1919370 Hampshire_County_Cricket_Club

11: 1899702 Friends_Provident_Trophy

12: 1622320 Glamorgan_County_Cricket_Club

13: 683599 County_Championship

14: 1829984 Derbyshire_County_Cricket_Club

1: 25675557 Cricket

[winner: 330091 Rhabdophoridae, pem_pos=10, final_scores=-1.4264811, final_local_scores=-1.3757377, pem=-2.56456, similarity=-23.01857|

gt: 25675557 Cricket, pem_pos=0, final_scores=-1.6167725, final_local_scores=-1.8012996, pem=-0.07532314, similarity=-49.255043]

2: 31144749 Derbyshire_County_Cricket_Club_in_1950

[winner: 19283873 1950, pem_pos=0, final_scores=-0.80430096, final_local_scores=-0.6748046, pem=-0.90795875, similarity=-11.110182|

RECALL MISS: 31144749

Derbyshire_County_Cricket_Club_in_1950]

[winner: 690345 Season_(sports), pem_pos=0, final_scores=-1.691784, final_local_scores=-1.7805855, pem=-1.219945, similarity=-42.412262|

RECALL MISS: 31144749

Derbyshire_County_Cricket_Club_in_1950]

3: 24582274 Derbyshire_County_Cricket_Club_in_1951

[winner: 34602 1951, pem_pos=0, final_scores=-0.55616486, final_local_scores=-0.34088683, pem=-0.7317484,

similarity=-2.1036682|
RECALL MISS: 24582274
Derbyshire_County_Cricket_Club_in_1951]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.5932778, final_local_scores=-1.6729825,
pem=-1.219945, similarity=-39.20117|
RECALL MISS: 24582274
Derbyshire_County_Cricket_Club_in_1951]
4: 31523322 Derbyshire_County_Cricket_Club_in_1955
[winner: 34851 1955, pem_pos=0, final_scores=-0.34182248,
final_local_scores=-0.08898421, pem=-0.7263205,
similarity=5.3840904|
RECALL MISS: 31523322
Derbyshire_County_Cricket_Club_in_1955]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.4285558, final_local_scores=-1.4930487,
pem=-1.219945, similarity=-33.83159|
RECALL MISS: 31523322
Derbyshire_County_Cricket_Club_in_1955]
5: 31636994 Derbyshire_County_Cricket_Club_in_1960
[winner: 34664 1960, pem_pos=0, final_scores=-0.6232178,
final_local_scores=-0.42938134, pem=-0.86076504,
similarity=-4.0429087|
RECALL MISS: 31636994
Derbyshire_County_Cricket_Club_in_1960]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.5571471, final_local_scores=-1.6335151,
pem=-1.219945, similarity=-38.023384|
RECALL MISS: 31636994
Derbyshire_County_Cricket_Club_in_1960]
6: 31775806 Derbyshire_County_Cricket_Club_in_1962
[winner: 34717 1962, pem_pos=0, final_scores=-0.5615926,
final_local_scores=-0.3751806, pem=-0.7481338,
similarity=-3.0379565|
RECALL MISS: 31775806
Derbyshire_County_Cricket_Club_in_1962]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.530143, final_local_scores=-1.6040173,
pem=-1.219945, similarity=-37.14311|
RECALL MISS: 31775806
Derbyshire_County_Cricket_Club_in_1962]
7: 33898082 Derbyshire_County_Cricket_Club_in_1963
[winner: 34648 1963, pem_pos=0, final_scores=-0.6816637,
final_local_scores=-0.5024052, pem=-0.6371062,
similarity=-7.4383783|
RECALL MISS: 33898082
Derbyshire_County_Cricket_Club_in_1963]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.3993589, final_local_scores=-1.4611554,
pem=-1.219945, similarity=-32.879826|
RECALL MISS: 33898082
Derbyshire_County_Cricket_Club_in_1963]

8: 27770988 Derbyshire_County_Cricket_Club_in_1964
[winner: 34550 1964, pem_pos=0, final_scores=-0.51380235,
final_local_scores=-0.29839426, pem=-0.70862895,
similarity=-0.96133304]
RECALL MISS: 27770988
Derbyshire_County_Cricket_Club_in_1964]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.6087145, final_local_scores=-1.6898446,
pem=-1.219945, similarity=-39.704372]
RECALL MISS: 27770988
Derbyshire_County_Cricket_Club_in_1964]
9: 33986216 Derbyshire_County_Cricket_Club_in_1965
[winner: 34750 1965, pem_pos=0, final_scores=-0.62386984,
final_local_scores=-0.41376787, pem=-0.69137967,
similarity=-4.498116]
RECALL MISS: 33986216
Derbyshire_County_Cricket_Club_in_1965]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.4665756, final_local_scores=-1.5345795,
pem=-1.219945, similarity=-35.07095]
RECALL MISS: 33986216
Derbyshire_County_Cricket_Club_in_1965]
10: 25205089 Derbyshire_County_Cricket_Club_in_1966
[winner: 34691 1966, pem_pos=0, final_scores=-0.7546696,
final_local_scores=-0.6119532, pem=-0.7275667,
similarity=-10.21557]
RECALL MISS: 25205089
Derbyshire_County_Cricket_Club_in_1966]
[winner: 1340236 Eastertide, pem_pos=2,
final_scores=-1.5931017, final_local_scores=-1.5928206,
pem=-2.3723626, similarity=-30.541952]
RECALL MISS: 25205089
Derbyshire_County_Cricket_Club_in_1966]
11: 34319636 Derbyshire_County_Cricket_Club_in_1967
[winner: 34749 1967, pem_pos=0, final_scores=-0.60264534,
final_local_scores=-0.386899, pem=-0.63886935,
similarity=-3.9818544]
RECALL MISS: 34319636
Derbyshire_County_Cricket_Club_in_1967]
12: 27734235 Derbyshire_County_Cricket_Club_in_1968
[winner: 34655 1968, pem_pos=0, final_scores=-0.60541546,
final_local_scores=-0.3783511, pem=-0.7100701,
similarity=-3.3395665]
RECALL MISS: 27734235
Derbyshire_County_Cricket_Club_in_1968]
13: 26711773 Derbyshire_County_Cricket_Club_in_1969
[winner: 34610 1969, pem_pos=0, final_scores=-0.5763222,
final_local_scores=-0.37068492, pem=-0.6617498,
similarity=-3.3735664]
RECALL MISS: 26711773
Derbyshire_County_Cricket_Club_in_1969]
1: 9316 England

[pem_pos=1, final_scores=0.4871903,
final_local_scores=0.5315185, pem=-1.3527626,
similarity=27.30782]
2: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.68516964,
final_local_scores=0.63640887, pem=-1.4828352,
similarity=31.14531]
3: 94184 Muswell_Hill
[pem_pos=0, final_scores=0.2537008,
final_local_scores=0.32303032, pem=0.0, similarity=13.729582]
4: 530308 Northamptonshire_County_Cricket_Club
[pem_pos=1, final_scores=0.4386924,
final_local_scores=0.3787101, pem=-1.2069724,
similarity=21.95488]
5: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.59637994,
final_local_scores=0.545045, pem=-1.379793,
similarity=27.858473]
6: 1177719 Kent_County_Cricket_Club
[pem_pos=1, final_scores=0.7782529,
final_local_scores=0.7450834, pem=-2.5096242,
similarity=39.972214]
7: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.51529795,
final_local_scores=0.4532104, pem=-1.7779728,
similarity=27.283308]
8: 1763691 Sussex_County_Cricket_Club
[pem_pos=1, final_scores=0.6096005,
final_local_scores=0.577467, pem=-1.5566913,
similarity=29.78801]
9: 1622178 Somerset_County_Cricket_Club
[pem_pos=1, final_scores=0.7492563,
final_local_scores=0.7090763, pem=-2.4139037,
similarity=38.377148]
10: 1919370 Hampshire_County_Cricket_Club
[pem_pos=1, final_scores=0.7589672,
final_local_scores=0.7226103, pem=-1.9232318,
similarity=36.11268]
11: 1899702 Friends_Provident_Trophy
[pem_pos=0, final_scores=0.45179003,
final_local_scores=0.41502, pem=-0.08377493,
similarity=16.930317]
12: 1622320 Glamorgan_County_Cricket_Club
[pem_pos=1, final_scores=0.7367887,
final_local_scores=0.70136034, pem=-0.7903395,
similarity=29.317694]
13: 683599 County_Championship
[pem_pos=0, final_scores=0.28107297,
final_local_scores=0.21273011, pem=0.0, similarity=10.438004]
14: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.52230847,
final_local_scores=0.45850772, pem=-1.4828352,

similarity=25.836388]

1: 9396739 Derek_Morgan_(cricketer)

[pem_pos=0, final_scores=0.42385417,
final_local_scores=0.4233151, pem=0.0, similarity=16.72228]

2: 1829984 Derbyshire_County_Cricket_Club

[pem_pos=1, final_scores=0.7025866,
final_local_scores=0.6554342, pem=-1.4828352,
similarity=31.713062]

3: 3249168 Leslie_Townsend_(cricketer)

[pem_pos=1, final_scores=0.6999014,
final_local_scores=0.7643512, pem=0.0, similarity=26.899477]

4: 9396739 Derek_Morgan_(cricketer)

[pem_pos=0, final_scores=0.5286997,
final_local_scores=0.5378429, pem=0.0, similarity=20.140018]

5: 17867 London

[pem_pos=0, final_scores=0.34338176,
final_local_scores=0.40388155, pem=-0.09086737,
similarity=16.636494]

6: 1829984 Derbyshire_County_Cricket_Club

[pem_pos=1, final_scores=0.66187245,
final_local_scores=0.6109602, pem=-1.4828352,
similarity=30.385872]

7: 41343909 Pakistan_Eaglets_cricket_team

[pem_pos=0, final_scores=0.30940437,
final_local_scores=0.35573068, pem=0.0, similarity=14.705426]

8: 1919370 Hampshire_County_Cricket_Club

[pem_pos=1, final_scores=0.7454403,
final_local_scores=0.7078342, pem=-1.9232318,
similarity=35.67173]

9: 1919370 Hampshire_County_Cricket_Club

[pem_pos=1, final_scores=0.86872214,
final_local_scores=0.84250087, pem=-1.9232318,
similarity=39.690453]

10: 1622178 Somerset_County_Cricket_Club

[pem_pos=1, final_scores=0.62698144,
final_local_scores=0.57550955, pem=-2.4139037,
similarity=34.39125]

11: 1899702 Friends_Provident_Trophy

[pem_pos=0, final_scores=0.3872053,
final_local_scores=0.34447104, pem=-0.08377493,
similarity=14.824997]

12: 1829984 Derbyshire_County_Cricket_Club

[pem_pos=1, final_scores=0.750433,
final_local_scores=0.7076991, pem=-1.4828352,
similarity=33.272755]

13: 9396739 Derek_Morgan_(cricketer)

[pem_pos=0, final_scores=0.65019804,
final_local_scores=0.67056143, pem=0.0, similarity=24.100603]

14: 1829984 Derbyshire_County_Cricket_Club

[pem_pos=1, final_scores=0.7293009,
final_local_scores=0.68461555, pem=-1.4828352,
similarity=32.583893]

15: 1829984 Derbyshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.76479095,
 final_local_scores=0.723383, pem=-1.4828352,
 similarity=33.74079]
 16: 9396739 Derek_Morgan_(cricketer)
 [pem_pos=0, final_scores=0.7159994,
 final_local_scores=0.74243945, pem=0.0, similarity=26.245586]
 17: 667292 Warwickshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.70383406,
 final_local_scores=0.66214335, pem=-1.1093602,
 similarity=29.882265]
 18: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.30396754,
 final_local_scores=0.2535464, pem=-0.08377493,
 similarity=12.111624]
 19: 9396739 Derek_Morgan_(cricketer)
 [pem_pos=0, final_scores=0.6865655,
 final_local_scores=0.7102874, pem=0.0, similarity=25.286106]

Sample: b'Hugh_Waddell'

-LRB- rugby_union -RRB- Hugh Waddell -LRB- born -RRB- is a
 [1 1 Scottish 1] 2 rugby union and professional 3 rugby league
 footballer of the 1980s and '90s playing at representative
 has played level for [2 4 Great Britain 2] , [3 5 England
 3] , and [4 6 Scotland 4] , and at club level for 7 Blackpool
 Borough , [5 8 Oldham 5] , [6 9 Leeds 6] , [1 1 Manly-
 Warringah Sea Eagles 1] , [2 2 Sheffield Eagles 2] , [3 3 Wake-
 field Trinity 3] , [4 4 Rochdale Hornets 4] , [5 5 Swinton Li-
 ons 5] , [7 10 Carlisle 7] , [8 11 Barrow 8] , 12 Egremont Ra-
 ngers , and , as a , or . `` Hugh Waddell - [9 Prop 9] .
 Formerly a 13 soccer player in [6 6 Burton-on-Trent 6] ,
 Waddell first played 14 League for 15 Blackpool Borough and
 won 16 England Honours in 1984 . Transferred to [10 17 Oldham
 10] and débuted for 18 GB in 1988 , starring in the 19 Lions '
 third 20 Test victory over [11 21 Australia 11] . Signed for [12
 22 Leeds 12] , then 23 Sheffield . Winner in 1988 , 5 24 Te-
 st 25 cap s. '' Hugh Waddell won caps for [13 26 England 13]
 while at Blackpool Borough in 1984 against [14 Wales 14] ,
 won caps for 27 Scotland , and won caps for [7 7 Great Britain
 7] while at [15 28 Oldham 15] in 1988 against [16 France 16] -
 LRB- 2 matches -RRB- , Australia , and [17 New Zealand 17] ,
 and while at [18 29 Leeds 18]

1: 885301 Manly-Warringah_Sea_Eagles
 2: 3367769 Sheffield_Eagles
 3: 1095973 Wakefield_Trinity_Wildcats
 4: 2934756 Rochdale_Hornets
 5: 2000369 Swinton_Lions
 6: 223526 Burton_upon_Trent
 7: 951941 Great_Britain_national_rugby_league_team
 1: 34850722 Scottish_people

[winner: 26994 Scotland, pem_pos=0,
 final_scores=0.47949666, final_local_scores=0.518235,

pem=-0.47951394, similarity=22.162552|
gt: 34850722 Scottish_people, pem_pos=1,
final_scores=-0.34999952, final_local_scores=-0.38228852,
pem=-1.5054524, similarity=0.86834574]
2: 25405 Rugby_union
[winner: 19774201 Rugby_Virginia, pem_pos=7,
final_scores=-0.9025453, final_local_scores=-0.77074134,
pem=-2.956099, similarity=-2.8350203|
gt: 25405 Rugby_union, pem_pos=3,
final_scores=-2.315932, final_local_scores=-2.5630064,
pem=-2.0110612, similarity=-61.459045]
[winner: 481130 Union_(American_Civil_War), pem_pos=0,
final_scores=-1.2884321, final_local_scores=-1.3296266,
pem=-1.132797, similarity=-29.428665|
RECALL MISS: 25405 Rugby_union]
3: 25735 Rugby_league
[winner: 19774201 Rugby_Virginia, pem_pos=7,
final_scores=-0.91246945, final_local_scores=-0.78158206,
pem=-2.956099, similarity=-3.1585283|
gt: 25735 Rugby_league, pem_pos=14,
final_scores=-1.8128462, final_local_scores=-2.0231566,
pem=-3.3002021, similarity=-38.338306]
[winner: 21211 National_Football_League, pem_pos=1,
final_scores=-1.669908, final_local_scores=-1.8138919,
pem=-1.6728458, similarity=-40.943245|
gt: 25735 Rugby_league, pem_pos=5,
final_scores=-1.8006874, final_local_scores=-2.009875,
pem=-3.5378149, similarity=-36.649776]
4: 951941 Great_Britain_national_rugby_league_team
[winner: 13530298 Great_Britain, pem_pos=0,
final_scores=0.046056077, final_local_scores=0.06581539,
pem=-0.50213265, similarity=8.784451|
gt: 951941 Great_Britain_national_rugby_league_team,
pem_pos=2, final_scores=-0.117175356,
final_local_scores=-0.16810006, pem=-2.6635387,
similarity=13.558009]
5: 990422 England_national_rugby_league_team
[winner: 9316 England, pem_pos=0,
final_scores=0.27696788, final_local_scores=0.27453017,
pem=-0.19053856, similarity=13.318419|
gt: 990422 England_national_rugby_league_team,
pem_pos=5, final_scores=-0.5129961,
final_local_scores=-0.60984963, pem=-4.237946,
similarity=8.937193]
6: 990481 Scotland_national_rugby_league_team
[winner: 758049 Scotland_national_rugby_union_team,
pem_pos=2, final_scores=0.07144295,
final_local_scores=0.03720367, pem=-3.3931751,
similarity=23.652557|
gt: 990481 Scotland_national_rugby_league_team,
pem_pos=6, final_scores=-0.34869027,
final_local_scores=-0.43216136, pem=-4.2446737,

```

similarity=14.276352 ]
7: 2946790 Blackpool_Borough
  [winner: 2946790 Blackpool_Borough, pem_pos=0,
  final_scores=-0.025735125, final_local_scores=-0.06564048,
  pem=0.0, similarity=2.13087]
  gt: 2946790 Blackpool_Borough, pem_pos=0,
  final_scores=-0.025735125, final_local_scores=-0.06564048,
  pem=0.0, similarity=2.13087 ]
8: 1098719 Oldham_Roughyeds
  [winner: 451131 Oldham_Athletic_A.F.C., pem_pos=3,
  final_scores=0.31646645, final_local_scores=0.25841403,
  pem=-2.8632221, similarity=27.37195]
  gt: 1098719 Oldham_Roughyeds, pem_pos=1,
  final_scores=-0.17868735, final_local_scores=-0.25470492,
  pem=-2.0696447, similarity=7.7438593 ]
9: 1095024 Leeds_Rhinos
  [winner: 7609513 Leeds_United_A.F.C., pem_pos=4,
  final_scores=0.370637, final_local_scores=0.3146934,
  pem=-3.421221, similarity=32.085922]
  gt: 1095024 Leeds_Rhinos, pem_pos=1,
  final_scores=-0.16649568, final_local_scores=-0.2481224,
  pem=-2.7286434, similarity=11.524034 ]
10: 7564926 Carlisle_RLFC
  [winner: 451157 Carlisle_United_F.C., pem_pos=12,
  final_scores=0.5323921, final_local_scores=0.5006202,
  pem=-3.5999887, similarity=38.606518]
  RECALL MISS: 7564926 Carlisle_RLFC]
11: 1262740 Barrow_Raiders
  [winner: 712826 Barrow_A.F.C., pem_pos=0,
  final_scores=0.5857907, final_local_scores=0.5666875,
  pem=-1.0930984, similarity=26.945238]
  gt: 1262740 Barrow_Raiders, pem_pos=3,
  final_scores=-0.053459063, final_local_scores=-0.122482926,
  pem=-2.3147502, similarity=13.02255 ]
12: 1290378 National_Conference_League
  [winner: 35810826 Egremont_Rangers_ARLFC, pem_pos=0,
  final_scores=-0.03184037, final_local_scores=-0.03865783,
  pem=-0.6931472, similarity=6.7055283]
  gt: 1290378 National_Conference_League, pem_pos=1,
  final_scores=-0.058451578, final_local_scores=-0.08660333,
  pem=-0.6931472, similarity=5.274739 ]
13: 10568 Association_football
  [winner: 10568 Association_football, pem_pos=0,
  final_scores=-1.4124292, final_local_scores=-1.5564768,
  pem=0.0, similarity=-42.358658]
  gt: 10568 Association_football, pem_pos=0,
  final_scores=-1.4124292, final_local_scores=-1.5564768,
  pem=0.0, similarity=-42.358658 ]
14: 25735 Rugby_league
  [winner: 14419588 The_Football_League, pem_pos=0,
  final_scores=0.021952108, final_local_scores=-0.055971533,
  pem=-1.7772934, similarity=12.084609]

```


gt: 25735 Rugby_league, pem_pos=6,
 final_scores=-0.58118045, final_local_scores=-0.677749,
 pem=-3.0560107, similarity=0.48339215]
 15: 2946790 Blackpool_Borough
 [winner: 2946790 Blackpool_Borough, pem_pos=0,
 final_scores=-0.06605227, final_local_scores=-0.10968083,
 pem=0.0, similarity=0.8166182]
 gt: 2946790 Blackpool_Borough, pem_pos=0,
 final_scores=-0.06605227, final_local_scores=-0.10968083,
 pem=0.0, similarity=0.8166182]
 16: 990422 England_national_rugby_league_team
 [winner: 9904 England_national_football_team, pem_pos=1,
 final_scores=-0.44169414, final_local_scores=-0.57585704,
 pem=-2.776464, similarity=2.0038319]
 gt: 990422 England_national_rugby_league_team,
 pem_pos=5, final_scores=-1.0187658,
 final_local_scores=-1.1615994, pem=-4.237946,
 similarity=-7.528117]
 17: 1098719 Oldham_Roughyeds
 [winner: 451131 Oldham_Athletic_A.F.C., pem_pos=3,
 final_scores=0.32371825, final_local_scores=0.26633555,
 pem=-2.8632221, similarity=27.608343]
 gt: 1098719 Oldham_Roughyeds, pem_pos=1,
 final_scores=-0.14027148, final_local_scores=-0.21274143,
 pem=-2.0696447, similarity=8.996134]
 18: 951941 Great_Britain_national_rugby_league_team
 [winner: 12570 Gigabyte, pem_pos=0,
 final_scores=-0.4634424, final_local_scores=-0.35148543,
 pem=-0.41590655, similarity=-4.1375484]
 RECALL MISS: 951941
 Great_Britain_national_rugby_league_team]
 19: 951941 Great_Britain_national_rugby_league_team
 [winner: 8577 Detroit_Lions, pem_pos=1,
 final_scores=-0.15680245, final_local_scores=-0.16013682,
 pem=-1.79462, similarity=9.0703335]
 gt: 951941 Great_Britain_national_rugby_league_team,
 pem_pos=10, final_scores=-0.5389998,
 final_local_scores=-0.6286382, pem=-3.8479764,
 similarity=6.255789]
 [winner: 2026886 Āstika_and_nāstika, pem_pos=6,
 final_scores=-2.0125406, final_local_scores=-1.9964516,
 pem=-4.060443, similarity=-33.407063]
 RECALL MISS: 951941
 Great_Britain_national_rugby_league_team]
 20: 16926380 Test_match_(rugby_league)
 [winner: 30863 Test_cricket, pem_pos=0,
 final_scores=-0.280506, final_local_scores=-0.3110525,
 pem=-0.23950954, similarity=-3.8902242]
 RECALL MISS: 16926380 Test_match_(rugby_league)]
 21: 951978 Australia_national_rugby_league_team
 [winner: 4792249
 Australia_national_association_football_team, pem_pos=2,

final_scores=0.10667975, final_local_scores=0.0787586,
 pem=-3.687066, similarity=26.490866|
 gt: 951978 Australia_national_rugby_league_team,
 pem_pos=4, final_scores=-0.3341822,
 final_local_scores=-0.40505624, pem=-4.3727627,
 similarity=15.78179]
 22: 1095024 Leeds_Rhinos
 [winner: 7609513 Leeds_United_A.F.C., pem_pos=4,
 final_scores=0.42540842, final_local_scores=0.37452286,
 pem=-3.421221, similarity=33.871353|
 gt: 1095024 Leeds_Rhinos, pem_pos=1,
 final_scores=-0.20886533, final_local_scores=-0.2944048,
 pem=-2.7286434, similarity=10.142874]
 23: 3367769 Sheffield_Eagles
 [winner: 88885 Sheffield, pem_pos=0,
 final_scores=-0.3711695, final_local_scores=-0.44329113,
 pem=-0.25810194, similarity=-7.735379|
 RECALL MISS: 3367769 Sheffield_Eagles]
 24: 16926380 Test_match_(rugby_league)
 [winner: 30863 Test_cricket, pem_pos=0,
 final_scores=-0.23487082, final_local_scores=-0.261203,
 pem=-0.23950954, similarity=-2.4026155|
 RECALL MISS: 16926380 Test_match_(rugby_league)]
 25: 3191695 Cap_(sport)
 [winner: 1793788 Montenegrin_cap, pem_pos=7,
 final_scores=-1.3072429, final_local_scores=-1.2756355,
 pem=-1.7908002, similarity=-24.239141|
 gt: 3191695 Cap_(sport), pem_pos=0,
 final_scores=-2.1456223, final_local_scores=-2.4013877,
 pem=-0.45111725, similarity=-65.11923]
 26: 990422 England_national_rugby_league_team
 [winner: 9904 England_national_football_team, pem_pos=1,
 final_scores=0.20295964, final_local_scores=0.13116622,
 pem=-2.776464, similarity=23.102818|
 gt: 990422 England_national_rugby_league_team,
 pem_pos=5, final_scores=-0.45726407,
 final_local_scores=-0.54902065, pem=-4.237946,
 similarity=10.752452]
 27: 990481 Scotland_national_rugby_league_team
 [winner: 578101 Scotland_national_football_team,
 pem_pos=1, final_scores=0.011057779,
 final_local_scores=-0.07010245, pem=-2.6956348,
 similarity=16.656998|
 gt: 990481 Scotland_national_rugby_league_team,
 pem_pos=6, final_scores=-0.3331468,
 final_local_scores=-0.41770515, pem=-4.2446737,
 similarity=14.707754]
 28: 1098719 Oldham_Roughyeds
 [winner: 451131 Oldham_Athletic_A.F.C., pem_pos=3,
 final_scores=0.30247492, final_local_scores=0.24313039,
 pem=-2.8632221, similarity=26.915855|
 gt: 1098719 Oldham_Roughyeds, pem_pos=1,

```

final_scores=-0.075499, final_local_scores=-0.14198732,
pem=-2.0696447, similarity=11.107577 ]
29: 1095024 Leeds_Rhinos
    [winner: 7609513 Leeds_United_A.F.C., pem_pos=4,
final_scores=0.18666361, final_local_scores=0.11373043,
pem=-3.421221, similarity=26.088787]
    gt: 1095024 Leeds_Rhinos, pem_pos=1,
final_scores=-0.03730786, final_local_scores=-0.107004315,
pem=-2.7286434, similarity=15.735278 ]
1: 885301 Manly-Warringah_Sea_Eagles
    [pem_pos=0, final_scores=0.3397025,
final_local_scores=0.3357374, pem=0.0, similarity=14.108787]
2: 3367769 Sheffield_Eagles
    [pem_pos=0, final_scores=0.2370257,
final_local_scores=0.20433584, pem=0.0, similarity=10.187502]
3: 1095973 Wakefield_Trinity_Wildcats
    [pem_pos=0, final_scores=0.2501576,
final_local_scores=0.21801814, pem=0.0, similarity=10.595809]
4: 2934756 Rochdale_Hornets
    [pem_pos=0, final_scores=0.2458358,
final_local_scores=0.21802956, pem=0.0, similarity=10.596149]
5: 2000369 Swinton_Lions
    [pem_pos=0, final_scores=0.18238388,
final_local_scores=0.14213857, pem=0.0, similarity=8.331411]
6: 223526 Burton_upon_Trent
    [pem_pos=0, final_scores=0.30002338,
final_local_scores=0.3273528, pem=0.0, similarity=13.858573]
7: 951941 Great_Britain_national_rugby_league_team
    [pem_pos=2, final_scores=0.035452917,
final_local_scores=-0.0011358857, pem=-2.6635387,
similarity=18.540554]
1: 26994 Scotland
    [pem_pos=0, final_scores=0.47949666,
final_local_scores=0.518235, pem=-0.47951394,
similarity=22.162552]
2: 13530298 Great_Britain
    [pem_pos=0, final_scores=0.046056077,
final_local_scores=0.06581539, pem=-0.50213265,
similarity=8.784451]
3: 9316 England
    [pem_pos=0, final_scores=0.27696788,
final_local_scores=0.27453017, pem=-0.19053856,
similarity=13.318419]
4: 758049 Scotland_national_rugby_union_team
    [pem_pos=2, final_scores=0.07144295,
final_local_scores=0.03720367, pem=-3.3931751,
similarity=23.652557]
5: 451131 Oldham_Athletic_A.F.C.
    [pem_pos=3, final_scores=0.31646645,
final_local_scores=0.25841403, pem=-2.8632221,
similarity=27.37195]
6: 7609513 Leeds_United_A.F.C.

```

[pem_pos=4, final_scores=0.370637,
final_local_scores=0.3146934, pem=-3.421221,
similarity=32.085922]
7: 451157 Carlisle_United_F.C.
[pem_pos=12, final_scores=0.5323921,
final_local_scores=0.5006202, pem=-3.5999887,
similarity=38.606518]
8: 712826 Barrow_A.F.C.
[pem_pos=0, final_scores=0.5857907,
final_local_scores=0.5666875, pem=-1.0930984,
similarity=26.945238]
9: 1195281 Rugby_league_positions
[pem_pos=1, final_scores=0.032120958,
final_local_scores=0.008673802, pem=-0.08842231,
similarity=4.8294106]
10: 451131 Oldham_Athletic_A.F.C.
[pem_pos=3, final_scores=0.32371825,
final_local_scores=0.26633555, pem=-2.8632221,
similarity=27.608343]
11: 4792249 Australia_national_association_football_team
[pem_pos=2, final_scores=0.10667975,
final_local_scores=0.0787586, pem=-3.687066,
similarity=26.490866]
12: 7609513 Leeds_United_A.F.C.
[pem_pos=4, final_scores=0.42540842,
final_local_scores=0.37452286, pem=-3.421221,
similarity=33.871353]
13: 9904 England_national_football_team
[pem_pos=1, final_scores=0.20295964,
final_local_scores=0.13116622, pem=-2.776464,
similarity=23.102818]
14: 580291 Wales_national_football_team
[pem_pos=1, final_scores=0.14551528,
final_local_scores=0.07375988, pem=-2.4526134,
similarity=19.628548]
15: 451131 Oldham_Athletic_A.F.C.
[pem_pos=3, final_scores=0.30247492,
final_local_scores=0.24313039, pem=-2.8632221,
similarity=26.915855]
16: 250197 France_national_football_team
[pem_pos=1, final_scores=0.19825594,
final_local_scores=0.18540174, pem=-3.3831942,
similarity=28.020805]
17: 154957 New_Zealand_national_rugby_union_team
[pem_pos=4, final_scores=0.36623245,
final_local_scores=0.36806786, pem=-3.8192244,
similarity=35.84313]
18: 7609513 Leeds_United_A.F.C.
[pem_pos=4, final_scores=0.18666361,
final_local_scores=0.11373043, pem=-3.421221,
similarity=26.088787]

Sample: b'Arthur_Keegan'

Arthur `` Ollie `` [1 Keegan 1] -LRB- born 6 November 1938 in [1 1 Dewsbury 1] district -- died 3 November 2008 -RRB- was an [2 1 English 2] professional 2 rugby league footballer of the 1950s , '60s and '70s , and 3 coach of the 1970s , playing at representative level for 4 Great Britain , [3 5 England 3] , and [4 6 Yorkshire 4] , and at club level for 7 Hull , 8 Bramley , and 9 Batley , as a , i.e. number 1 , and was 10 captain of Hull during the 11 1965 -- 66 , 12 1966 -- 67 , 13 1967 -- 68 , 14 1968 -- 69 , 15 1969 -- 70 , and 16 1970 -- 71 seasons , and coaching at representative level for [5 17 Yorkshire 5] , and at club level for 18 Bramley , whilst serving in the [6 19 Duke of Wellington 6] 's Regiment . [7 Arthur Keegan 7] won caps for [8 20 England 8] while at [9 Hull 9] in 1969 against Wales , and [10 France 10] , and won caps for 21 Great Britain while at [11 Hull 11] in 1966 against Australia -LRB- 2 matches -RRB- , in 1967 against [12 France 12] -LRB- 2 matches -RRB- , and Australia -LRB- 3 matches -RRB- , in 1968 against [13 France 13] , and in 1969 against [14 France 14] . [15 Arthur Keegan 15] played , and scored five s in [16 22 Hull 16] 's 13-30 defeat to [2 2 Wigan 2] in the [3 3 Challenge Cup 3] final at [4 4 Wembley Stadium 4]

1: 149867 Dewsbury

2: 1095029 Wigan_Warriors

3: 21050680 Challenge_Cup

4: 143673 Wembley_Stadium_(1923)

1: 18803164 English_people

[winner: 9316 England, pem_pos=1, final_scores=0.6394939, final_local_scores=0.62719077, pem=-1.3527626, similarity=30.162868]

gt: 18803164 English_people, pem_pos=2, final_scores=-0.44448525, final_local_scores=-0.5030521, pem=-2.2084396, similarity=1.0874732]

2: 25735 Rugby_league

[winner: 19774201 Rugby_Virginia, pem_pos=7, final_scores=-0.9322802, final_local_scores=-0.8067922, pem=-2.956099, similarity=-3.9108508]

gt: 25735 Rugby_league, pem_pos=14, final_scores=-1.8308368, final_local_scores=-2.0263972, pem=-3.3002021, similarity=-38.43501]

[winner: 1424274 League_(unit), pem_pos=2, final_scores=-1.8807969, final_local_scores=-1.9873192, pem=-1.8401146, similarity=-45.209023]

gt: 25735 Rugby_league, pem_pos=5, final_scores=-1.9984587, final_local_scores=-2.209499, pem=-3.5378149, similarity=-42.60695]

3: 626967 Coach_(sport)

[winner: 7001560 Coach_(basketball), pem_pos=3, final_scores=-1.32797, final_local_scores=-1.3623538, pem=-1.991629, similarity=-25.734854]

gt: 626967 Coach_(sport), pem_pos=0,

final_scores=-1.5334122, final_local_scores=-1.6885898,
pem=-0.5765157, similarity=-43.165997]
4: 951941 Great_Britain_national_rugby_league_team
[winner: 13530298 Great_Britain, pem_pos=0,
final_scores=0.02918829, final_local_scores=-0.0027100146,
pem=-0.50213265, similarity=6.7395163]
gt: 951941 Great_Britain_national_rugby_league_team,
pem_pos=2, final_scores=-0.13868555,
final_local_scores=-0.17548683, pem=-2.6635387,
similarity=13.337573]
5: 990422 England_national_rugby_league_team
[winner: 9316 England, pem_pos=0,
final_scores=0.38327366, final_local_scores=0.3473091,
pem=-0.19053856, similarity=15.490288]
gt: 990422 England_national_rugby_league_team,
pem_pos=5, final_scores=-0.5909234,
final_local_scores=-0.68195087, pem=-4.237946,
similarity=6.7855477]
6: 493717 Rugby_League_War_of_the_Roses
[winner: 36637 Yorkshire, pem_pos=0,
final_scores=0.12962513, final_local_scores=0.083374366,
pem=-0.31620604, similarity=8.297348]
gt: 493717 Rugby_League_War_of_the_Roses, pem_pos=3,
final_scores=-0.37659585, final_local_scores=-0.41929293,
pem=-3.1132855, similarity=8.507707]
7: 1095938 Hull_F.C.
[winner: 451358 Hull_City_A.F.C., pem_pos=11,
final_scores=-0.016595244, final_local_scores=-0.069287896,
pem=-3.7682815, similarity=22.514524]
gt: 1095938 Hull_F.C., pem_pos=1,
final_scores=-0.09931818, final_local_scores=-0.15331131,
pem=-2.238242, similarity=11.686505]
8: 3753301 Bramley_Buffaloes
[winner: 3753301 Bramley_Buffaloes, pem_pos=3,
final_scores=0.025878564, final_local_scores=0.015980065,
pem=-1.6497065, similarity=13.5379505]
gt: 3753301 Bramley_Buffaloes, pem_pos=3,
final_scores=0.025878564, final_local_scores=0.015980065,
pem=-1.6497065, similarity=13.5379505]
9: 2817741 Batley_Bulldogs
[winner: 2817741 Batley_Bulldogs, pem_pos=1,
final_scores=-0.016303286, final_local_scores=-0.04305786,
pem=-1.6426295, similarity=11.737655]
gt: 2817741 Batley_Bulldogs, pem_pos=1,
final_scores=-0.016303286, final_local_scores=-0.04305786,
pem=-1.6426295, similarity=11.737655]
10: 1874452 Captain_(sports)
[winner: 3319715 Australian_Army_officer_rank_insignia,
pem_pos=11, final_scores=-1.5069345,
final_local_scores=-1.4917707, pem=-2.7026498,
similarity=-25.730274]
gt: 1874452 Captain_(sports), pem_pos=2,

final_scores=-1.582673, final_local_scores=-1.750149,
pem=-1.7624387, similarity=-38.55381]
11: 24328558 1965-66_Northern_Rugby_Football_League_season
[winner: 6954400 No._66_Squadron_RAF, pem_pos=1,
final_scores=-0.52975214, final_local_scores=-0.50297505,
pem=-3.0754452, similarity=5.804683]
RECALL MISS: 24328558 1965-
66_Northern_Rugby_Football_League_season]
[winner: 34750 1965, pem_pos=0, final_scores=-0.69329643,
final_local_scores=-0.5107978, pem=-0.69137967,
similarity=-7.3936825]
RECALL MISS: 24328558 1965-
66_Northern_Rugby_Football_League_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.4199436, final_local_scores=-1.3215768,
pem=0.0, similarity=-35.348778]
RECALL MISS: 24328558 1965-
66_Northern_Rugby_Football_League_season]
12: 24328799 1966-67_Northern_Rugby_Football_League_season
[winner: 34691 1966, pem_pos=0, final_scores=-0.7252018,
final_local_scores=-0.57832396, pem=-0.7275667,
similarity=-9.2120075]
RECALL MISS: 24328799 1966-
67_Northern_Rugby_Football_League_season]
[winner: 10679909 No._67_Squadron_RAF, pem_pos=0,
final_scores=-0.8643495, final_local_scores=-0.8239116,
pem=-2.4578676, similarity=-7.1311855]
RECALL MISS: 24328799 1966-
67_Northern_Rugby_Football_League_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3212786, final_local_scores=-1.2138002,
pem=0.0, similarity=-32.132504]
RECALL MISS: 24328799 1966-
67_Northern_Rugby_Football_League_season]
13: 24328953 1967-68_Northern_Rugby_Football_League_season
[winner: 34749 1967, pem_pos=0, final_scores=-0.56501377,
final_local_scores=-0.3377627, pem=-0.63886935,
similarity=-2.5155296]
RECALL MISS: 24328953 1967-
68_Northern_Rugby_Football_League_season]
[winner: 34812 68, pem_pos=0, final_scores=-0.952055,
final_local_scores=-0.86525685, pem=-2.1829414,
similarity=-9.860102]
RECALL MISS: 24328953 1967-
68_Northern_Rugby_Football_League_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.2906343, final_local_scores=-1.180326,
pem=0.0, similarity=-31.133572]
RECALL MISS: 24328953 1967-
68_Northern_Rugby_Football_League_season]
14: 24329050 1968-69_Northern_Rugby_Football_League_season
[winner: 34655 1968, pem_pos=0, final_scores=-0.56439525,

final_local_scores=-0.35232413, pem=-0.7100701,
 similarity=-2.56287|
 RECALL MISS: 24329050 1968-
 69_Northern_Rugby_Football_League_season]
 [winner: 34813 69, pem_pos=0, final_scores=-1.0521705,
 final_local_scores=-0.9966067, pem=-1.8609622,
 similarity=-15.530818|
 RECALL MISS: 24329050 1968-
 69_Northern_Rugby_Football_League_season]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.295297, final_local_scores=-1.1854193,
 pem=0.0, similarity=-31.285564|
 RECALL MISS: 24329050 1968-
 69_Northern_Rugby_Football_League_season]
 15: 24329822 1969-70_Northern_Rugby_Football_League_season
 [winner: 34610 1969, pem_pos=0, final_scores=-0.6665154,
 final_local_scores=-0.4998784, pem=-0.6617498,
 similarity=-7.228957|
 RECALL MISS: 24329822 1969-
 70_Northern_Rugby_Football_League_season]
 [winner: 34932 70, pem_pos=0, final_scores=-0.9626194,
 final_local_scores=-0.87677026, pem=-1.7523105,
 similarity=-12.545524|
 RECALL MISS: 24329822 1969-
 70_Northern_Rugby_Football_League_season]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.2866839, final_local_scores=-1.1760108,
 pem=0.0, similarity=-31.004799|
 RECALL MISS: 24329822 1969-
 70_Northern_Rugby_Football_League_season]
 16: 24330088 1970-71_Northern_Rugby_Football_League_season
 [winner: 34726 1970, pem_pos=0, final_scores=-0.7146688,
 final_local_scores=-0.55964273, pem=-0.80837256,
 similarity=-8.215088|
 RECALL MISS: 24330088 1970-
 71_Northern_Rugby_Football_League_season]
 [winner: 35283 71, pem_pos=3, final_scores=-1.1682278,
 final_local_scores=-1.0806184, pem=-3.066478,
 similarity=-11.48211|
 RECALL MISS: 24330088 1970-
 71_Northern_Rugby_Football_League_season]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.2709643, final_local_scores=-1.1588395,
 pem=0.0, similarity=-30.492367|
 RECALL MISS: 24330088 1970-
 71_Northern_Rugby_Football_League_season]
 17: 493717 Rugby_League_War_of_the_Roses
 [winner: 36637 Yorkshire, pem_pos=0,
 final_scores=0.23648737, final_local_scores=0.20010513,
 pem=-0.31620604, similarity=11.780827|
 gt: 493717 Rugby_League_War_of_the_Roses, pem_pos=3,
 final_scores=-0.50810784, final_local_scores=-0.5629498,

pem=-3.1132855, similarity=4.220699]
18: 3753301 Bramley_Buffaloes
[winner: 3753301 Bramley_Buffaloes, pem_pos=3,
final_scores=-0.020500198, final_local_scores=-0.04460153,
pem=-1.6497065, similarity=11.730075]
gt: 3753301 Bramley_Buffaloes, pem_pos=3,
final_scores=-0.020500198, final_local_scores=-0.04460153,
pem=-1.6497065, similarity=11.730075]
19: 473355 Duke_of_Wellington's_Regiment
[winner: 8474 Arthur_Wellesley,_1st_Duke_of_Wellington,
pem_pos=0, final_scores=0.049712226,
final_local_scores=0.075265676, pem=0.0,
similarity=6.3357906]
RECALL MISS: 473355 Duke_of_Wellington's_Regiment]
[winner: 166653 Regiment, pem_pos=0,
final_scores=-0.11393337, final_local_scores=-0.075951666,
pem=-0.011168689, similarity=1.8839008]
RECALL MISS: 473355 Duke_of_Wellington's_Regiment]
[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.3663871, final_local_scores=-2.4585044,
pem=-3.4067376, similarity=-50.75059]
RECALL MISS: 473355 Duke_of_Wellington's_Regiment]
20: 990422 England_national_rugby_league_team
[winner: 9316 England, pem_pos=0,
final_scores=0.11969258, final_local_scores=0.05938682,
pem=-0.19053856, similarity=6.8981123]
gt: 990422 England_national_rugby_league_team,
pem_pos=5, final_scores=-0.5512242,
final_local_scores=-0.63858545, pem=-4.237946,
similarity=8.079659]
21: 951941 Great_Britain_national_rugby_league_team
[winner: 13530298 Great_Britain, pem_pos=0,
final_scores=0.00037176907, final_local_scores=-0.034187675,
pem=-0.50213265, similarity=5.80016]
gt: 951941 Great_Britain_national_rugby_league_team,
pem_pos=2, final_scores=-0.16024716,
final_local_scores=-0.19903961, pem=-2.6635387,
similarity=12.634711]
22: 1095938 Hull_F.C.
[winner: 451358 Hull_City_A.F.C., pem_pos=11,
final_scores=0.21857844, final_local_scores=0.19515592,
pem=-3.7682815, similarity=30.406055]
gt: 1095938 Hull_F.C., pem_pos=1,
final_scores=0.0040108413, final_local_scores=-0.03827223,
pem=-2.238242, similarity=15.119501]
1: 149867 Dewsbury
[pem_pos=0, final_scores=0.16516404,
final_local_scores=0.14412668, pem=0.0, similarity=8.39074]
2: 1095029 Wigan_Warriors
[pem_pos=1, final_scores=0.2574461,
final_local_scores=0.24928156, pem=-1.8027086,
similarity=21.332182]

3: 21050680 Challenge_Cup
[pem_pos=0, final_scores=0.27530587,
final_local_scores=0.2694079, pem=-0.38845974,
similarity=14.241887]
4: 143673 Wembley_Stadium_(1923)
[pem_pos=1, final_scores=0.12795143,
final_local_scores=0.0909823, pem=-0.27926707,
similarity=8.323504]
1: 346979 Kevin_Keegan
[pem_pos=1, final_scores=0.111659035,
final_local_scores=0.116808355, pem=-0.4493757,
similarity=10.019282]
2: 9316 England
[pem_pos=1, final_scores=0.6394939,
final_local_scores=0.62719077, pem=-1.3527626,
similarity=30.162868]
3: 9316 England
[pem_pos=0, final_scores=0.38327366,
final_local_scores=0.3473091, pem=-0.19053856,
similarity=15.490288]
4: 36637 Yorkshire
[pem_pos=0, final_scores=0.12962513,
final_local_scores=0.083374366, pem=-0.31620604,
similarity=8.297348]
5: 36637 Yorkshire
[pem_pos=0, final_scores=0.23648737,
final_local_scores=0.20010513, pem=-0.31620604,
similarity=11.780827]
6: 8474 Arthur_Wellesley,_1st_Duke_of_Wellington
[pem_pos=0, final_scores=0.049712226,
final_local_scores=0.075265676, pem=0.0, similarity=6.3357906]
7: 22887149 Arthur_Keegan
[pem_pos=0, final_scores=0.39765328,
final_local_scores=0.45318228, pem=0.0, similarity=17.613577]
8: 9316 England
[pem_pos=0, final_scores=0.11969258,
final_local_scores=0.05938682, pem=-0.19053856,
similarity=6.8981123]
9: 56474 Kingston_upon_Hull
[pem_pos=0, final_scores=0.112628266,
final_local_scores=0.07634628, pem=-0.9178084,
similarity=11.359223]
10: 5843419 France
[pem_pos=0, final_scores=0.2198966,
final_local_scores=0.21881077, pem=-0.16240834,
similarity=11.502665]
11: 56474 Kingston_upon_Hull
[pem_pos=0, final_scores=0.059699222,
final_local_scores=0.018529356, pem=-0.9178084,
similarity=9.633851]
12: 250197 France_national_football_team
[pem_pos=1, final_scores=0.15649278,

final_local_scores=0.13916135, pem=-3.3831942,
 similarity=26.640902]
 13: 5843419 France
 [pem_pos=0, final_scores=0.16866796,
 final_local_scores=0.16285124, pem=-0.16240834,
 similarity=9.832721]
 14: 5843419 France
 [pem_pos=0, final_scores=0.17290245,
 final_local_scores=0.16552478, pem=-0.16240834,
 similarity=9.912504]
 15: 22887149 Arthur_Keegan
 [pem_pos=0, final_scores=0.53975827,
 final_local_scores=0.60841036, pem=0.0, similarity=22.245892]
 16: 451358 Hull_City_A.F.C.
 [pem_pos=11, final_scores=0.21857844,
 final_local_scores=0.19515592, pem=-3.7682815,
 similarity=30.406055]

Sample: b'Major_League_Baseball_on_TSN'

` [1 Major League Baseball 1] on TSN ' is a television
 presentation of [1 1 Major League Baseball 1] games on [2
 Canada 2] 's 1 TSN -LRB- The Sports Network -RRB- . [3 TSN 3]
 has broadcast [4 Major League Baseball 4] games since they
 went on the air in 1984 . Their current [5 MLB 5] schedule
 consists of simulcasting [6 2 ESPN 6] 's [2 2 Sunday Night Ba
 seball 2] broadcasts on [3 3 TSN2 3] . [7 TSN 7] had
 previously broadcast [8 Sunday Night Baseball 8] from
 1990-2000 . [9 TSN 9] has also broadcast [4 4 Toronto Blue Ja
 ys 4] -LRB- 3 1984 - 4 2009 -RRB- and [5 5 Montreal Expos 5] -
 LRB- 5 1985 - 6 1999 ; 7 2001 -RRB- games . [10 Toronto Blue
 Jays 10] baseball aired on [11 TSN 11] from September 1 , 8 19
 84 until September 23 , 9 2009 . When 10 The Sports Network
 launched in 1984 , the first program to air was a [12 11 Minn
 esota Twins 12] - [13 12 Toronto Blue Jays 13] game . In 13 19
 85 , [14 TSN 14] became the majority broadcaster of [6 6 Toro
 nto Blue Jays 6] games , a position which they would hold
 till the 14 2002 season . From 15 1993 - 16 2000 , [7 7 Baseb
 all Tonight 7] , a pregame show hosted by former Blue Jay [8 8
 Pat Tabler 8] preceded the game , with following it . As
 previously mentioned , the last game aired on September 23 ,
 2009 , as acquired the complete rights for the 17 2010 season
 . [15 Blue Jays 15] broadcasts were sponsored by 18 Labatt
 from 19 1984 - 20 2000 and by 21 Rogers from 22 2001 - 23 2009

1: 38776 Major_League_Baseball
 2: 4315483 Sunday_Night_Baseball
 3: 9215238 TSN2
 4: 30847 Toronto_Blue_Jays
 5: 20153 Montreal_Expos
 6: 30847 Toronto_Blue_Jays
 7: 1721061 Baseball_Tonight
 8: 2650908 Pat_Tabler
 1: 354723 The_Sports_Network

[winner: 354723 The_Sports_Network, pem_pos=0,
final_scores=-0.71926314, final_local_scores=-0.85133606,
pem=-0.17330989, similarity=-20.373367|
gt: 354723 The_Sports_Network, pem_pos=0,
final_scores=-0.71926314, final_local_scores=-0.85133606,
pem=-0.17330989, similarity=-20.373367]
2: 5876694 ESPN_Major_League_Baseball
[winner: 77795 ESPN, pem_pos=0, final_scores=0.21161895,
final_local_scores=0.17888927, pem=-0.09327439,
similarity=9.935367|
gt: 5876694 ESPN_Major_League_Baseball, pem_pos=1,
final_scores=-0.2558552, final_local_scores=-0.32882103,
pem=-2.9683454, similarity=10.419362]
3: 12207528 1984_Toronto_Blue_Jays_season
[winner: 34578 1984, pem_pos=0, final_scores=-0.99855906,
final_local_scores=-0.9502908, pem=-0.9448207,
similarity=-19.130777|
RECALL MISS: 12207528 1984_Toronto_Blue_Jays_season]
4: 19598205 2009_Toronto_Blue_Jays_season
[winner: 35983 2009, pem_pos=0, final_scores=-0.9168696,
final_local_scores=-0.8221318, pem=-0.9296576,
similarity=-15.388717|
RECALL MISS: 19598205 2009_Toronto_Blue_Jays_season]
5: 13991220 1985_Montreal_Expos_season
[winner: 34846 1985, pem_pos=0, final_scores=-0.8470697,
final_local_scores=-0.7448379, pem=-0.6656445,
similarity=-14.517857|
RECALL MISS: 13991220 1985_Montreal_Expos_season]
6: 13450695 1999_Montreal_Expos_season
[winner: 34553 1999, pem_pos=0, final_scores=-1.1340694,
final_local_scores=-1.1472082, pem=-0.6240096,
similarity=-26.751806|
RECALL MISS: 13450695 1999_Montreal_Expos_season]
7: 13955617 2001_Montreal_Expos_season
[winner: 34551 2001, pem_pos=0, final_scores=-0.57067895,
final_local_scores=-0.33476323, pem=-0.5613713,
similarity=-2.847465|
RECALL MISS: 13955617 2001_Montreal_Expos_season]
8: 12207528 1984_Toronto_Blue_Jays_season
[winner: 34578 1984, pem_pos=0, final_scores=-0.9723417,
final_local_scores=-0.92165244, pem=-0.9448207,
similarity=-18.276152|
RECALL MISS: 12207528 1984_Toronto_Blue_Jays_season]
9: 19598205 2009_Toronto_Blue_Jays_season
[winner: 35983 2009, pem_pos=0, final_scores=-0.90589017,
final_local_scores=-0.8101385, pem=-0.9296576,
similarity=-15.030814|
RECALL MISS: 19598205 2009_Toronto_Blue_Jays_season]
10: 354723 The_Sports_Network
[winner: 354723 The_Sports_Network, pem_pos=0,
final_scores=0.006176755, final_local_scores=-0.058903143,
pem=0.0, similarity=2.3319256|

gt: 354723 The_Sports_Network, pem_pos=0,
final_scores=0.006176755, final_local_scores=-0.058903143,
pem=0.0, similarity=2.3319256]
11: 9514043 1984_Minnesota_Twins_season
[winner: 20050 Minnesota_Twins, pem_pos=0,
final_scores=0.34807616, final_local_scores=0.3057657,
pem=-0.052076127, similarity=13.497569]
RECALL MISS: 9514043 1984_Minnesota_Twins_season]
12: 12207528 1984_Toronto_Blue_Jays_season
[winner: 30847 Toronto_Blue_Jays, pem_pos=0,
final_scores=0.36373186, final_local_scores=0.3050145,
pem=-0.018220792, similarity=13.291041]
gt: 12207528 1984_Toronto_Blue_Jays_season, pem_pos=15,
final_scores=-0.47042674, final_local_scores=-0.4676777,
pem=-3.0252838, similarity=6.5852413]
13: 12207474 1985_Toronto_Blue_Jays_season
[winner: 34846 1985, pem_pos=0, final_scores=-0.93018275,
final_local_scores=-0.8356263, pem=-0.6656445,
similarity=-17.227165]
RECALL MISS: 12207474 1985_Toronto_Blue_Jays_season]
14: 12206056 2002_Toronto_Blue_Jays_season
[winner: 35502 2002, pem_pos=0, final_scores=-0.7854227,
final_local_scores=-0.6715317, pem=-0.5897454,
similarity=-12.743006]
RECALL MISS: 12206056 2002_Toronto_Blue_Jays_season]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.6905342, final_local_scores=-1.8599467,
pem=-1.219945, similarity=-44.78056]
RECALL MISS: 12206056 2002_Toronto_Blue_Jays_season]
15: 12206431 1993_Toronto_Blue_Jays_season
[winner: 34598 1993, pem_pos=0, final_scores=-0.76156205,
final_local_scores=-0.57235765, pem=-0.7063994,
similarity=-9.149072]
RECALL MISS: 12206431 1993_Toronto_Blue_Jays_season]
16: 12206178 2000_Toronto_Blue_Jays_season
[winner: 34548 2000, pem_pos=0, final_scores=-0.64259183,
final_local_scores=-0.56148773, pem=-0.8862115,
similarity=-7.8468447]
RECALL MISS: 12206178 2000_Toronto_Blue_Jays_season]
17: 24370197 2010_Toronto_Blue_Jays_season
[winner: 43226 2010, pem_pos=0, final_scores=-0.95795846,
final_local_scores=-0.8469611, pem=-1.6731367,
similarity=-12.086516]
RECALL MISS: 24370197 2010_Toronto_Blue_Jays_season]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.6497445, final_local_scores=-1.8153901,
pem=-1.219945, similarity=-43.450897]
RECALL MISS: 24370197 2010_Toronto_Blue_Jays_season]
18: 17565 Labatt_Brewing_Company
[winner: 17565 Labatt_Brewing_Company, pem_pos=0,
final_scores=-0.0014439076, final_local_scores=0.008778468,
pem=0.0, similarity=4.35168]

gt: 17565 Labatt_Brewing_Company, pem_pos=0,
final_scores=-0.0014439076, final_local_scores=0.008778468,
pem=0.0, similarity=4.35168]
19: 12207528 1984_Toronto_Blue_Jays_season
[winner: 34578 1984, pem_pos=0, final_scores=-0.96409947,
final_local_scores=-0.91264904, pem=-0.9448207,
similarity=-18.007473]
RECALL MISS: 12207528 1984_Toronto_Blue_Jays_season]
20: 12206178 2000_Toronto_Blue_Jays_season
[winner: 34548 2000, pem_pos=0, final_scores=-0.50632817,
final_local_scores=-0.41264042, pem=-0.8862115,
similarity=-3.4049442]
RECALL MISS: 12206178 2000_Toronto_Blue_Jays_season]
21: 372874 Rogers_Communications
[winner: 9485388 Robbie_Rogers, pem_pos=18,
final_scores=-0.28743568, final_local_scores=-0.28582096,
pem=-4.2416224, similarity=18.626848]
gt: 372874 Rogers_Communications, pem_pos=1,
final_scores=-0.29074425, final_local_scores=-0.31496215,
pem=-2.1934192, similarity=6.618768]
22: 12206127 2001_Toronto_Blue_Jays_season
[winner: 34551 2001, pem_pos=0, final_scores=-0.6260736,
final_local_scores=-0.3952735, pem=-0.5613713,
similarity=-4.653212]
RECALL MISS: 12206127 2001_Toronto_Blue_Jays_season]
23: 19598205 2009_Toronto_Blue_Jays_season
[winner: 35983 2009, pem_pos=0, final_scores=-0.81195635,
final_local_scores=-0.70753, pem=-0.9296576,
similarity=-11.968772]
RECALL MISS: 19598205 2009_Toronto_Blue_Jays_season]
1: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.40522474,
final_local_scores=0.3604947, pem=0.0, similarity=14.847594]
2: 4315483 Sunday_Night_Baseball
[pem_pos=0, final_scores=0.14103968,
final_local_scores=0.10850559, pem=0.0, similarity=7.3277364]
3: 9215238 TSN2
[pem_pos=0, final_scores=0.04613398,
final_local_scores=0.049831286, pem=0.0, similarity=5.576778]
4: 30847 Toronto_Blue_Jays
[pem_pos=0, final_scores=0.325454,
final_local_scores=0.26320177, pem=-0.018220792,
similarity=12.043266]
5: 20153 Montreal_Expos
[pem_pos=0, final_scores=0.3524801,
final_local_scores=0.30371013, pem=-0.14979224,
similarity=13.967622]
6: 30847 Toronto_Blue_Jays
[pem_pos=0, final_scores=0.39461738,
final_local_scores=0.33875227, pem=-0.018220792,
similarity=14.297844]
7: 1721061 Baseball_Tonight

[pem_pos=0, final_scores=0.23071851,
final_local_scores=0.22376427, pem=0.0, similarity=10.767285]
8: 2650908 Pat_Tabler
[pem_pos=0, final_scores=0.39652675,
final_local_scores=0.41773176, pem=0.0, similarity=16.555662]
1: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.41887194,
final_local_scores=0.3754022, pem=0.0, similarity=15.292464]
2: 5042916 Canada
[pem_pos=0, final_scores=0.5191032,
final_local_scores=0.60028744, pem=-0.11662793,
similarity=22.637728]
3: 354723 The_Sports_Network
[pem_pos=0, final_scores=0.12134947,
final_local_scores=0.07299629, pem=-0.17330989,
similarity=7.210554]
4: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.43747324,
final_local_scores=0.39572132, pem=0.0, similarity=15.898826]
5: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.4622802,
final_local_scores=0.4228192, pem=0.0, similarity=16.707481]
6: 77795 ESPN
[pem_pos=0, final_scores=0.21161895,
final_local_scores=0.17888927, pem=-0.09327439,
similarity=9.935367]
7: 354723 The_Sports_Network
[pem_pos=0, final_scores=0.121205255,
final_local_scores=0.07283875, pem=-0.17330989,
similarity=7.205853]
8: 4315483 Sunday_Night_Baseball
[pem_pos=0, final_scores=0.19510226,
final_local_scores=0.16756073, pem=0.0, similarity=9.090059]
9: 354723 The_Sports_Network
[pem_pos=0, final_scores=0.16364197,
final_local_scores=0.11919442, pem=-0.17330989,
similarity=8.589198]
10: 30847 Toronto_Blue_Jays
[pem_pos=0, final_scores=0.36547768,
final_local_scores=0.30692154, pem=-0.018220792,
similarity=13.347951]
11: 354723 The_Sports_Network
[pem_pos=0, final_scores=0.085845426,
final_local_scores=0.034213528, pem=-0.17330989,
similarity=6.0531993]
12: 20050 Minnesota_Twins
[pem_pos=0, final_scores=0.34807616,
final_local_scores=0.3057657, pem=-0.052076127,
similarity=13.497569]
13: 30847 Toronto_Blue_Jays
[pem_pos=0, final_scores=0.36373186,
final_local_scores=0.3050145, pem=-0.018220792,

```
similarity=13.291041]
14: 354723 The_Sports_Network
    [pem_pos=0, final_scores=0.14777257,
final_local_scores=0.101859525, pem=-0.17330989,
similarity=8.071891]
15: 30847 Toronto_Blue_Jays
    [pem_pos=0, final_scores=0.31758827,
final_local_scores=0.25460967, pem=-0.47340602,
similarity=14.262229]
```

Sample: b'East_Mississippi_State_Hospital'

The East Mississippi State Hospital -LRB- EMSH -RRB- is a mental health facility of the 1 Mississippi Department of Mental Health located in [1 1 Meridian 1] , [2 2 Mississippi 2] . The facility is the third largest employer in the [1 Meridian 1] area . A regional center , EMSH serves 31 counties , including [3 3 Alcorn 3] , [2 2 Covington 2] , [3 3 Choctaw 3] , 4 Clarke , 5 Clay , 6 Forrest , 7 George , 8 Greene , 9 Jackson , 10 Jasper , [4 11 Jefferson Davis 4] , 12 Jones , [4 4 Kemper 4] , 13 Lamar , [5 5 Lauderdale County 5] , [6 6 Leake 6] , 14 Lowndes , 15 Marion , [7 7 Neshoba 7] , 16 Newton , 17 Noxubee , [8 8 Oktibbeha 8] , 18 Perry , 19 Prentiss , 20 Scott , 21 Smith , [9 9 Tippah 9] , [5 22 Tishomingo 5] , 23 Wayne , 24 Webster , and 25 Winston . Due to advocacy from [10 10 Dorothea Dix 10] , on March 8 , 1882 the passed legislation establishing the [6 East Mississippi State Insane Asylum 6] . The hospital opened in a location west of [11 11 Meridian 11] in 1885 . In 1898 the facility 's name was changed to the [7 East Mississippi Insane Hospital 7] . In the 1930s the facility received its current name . In 1984 the MDMH Board of Mental Health established a catchment area to EMSH so that admissions to EMSH and the [12 12 Mississippi State Hospital 12]

```
1: 106031 Meridian,_Mississippi
2: 16949861 Mississippi
3: 94856 Alcorn_County,_Mississippi
4: 94813 Kemper_County,_Mississippi
5: 94809 Lauderdale_County,_Mississippi
6: 94781 Leake_County,_Mississippi
7: 94772 Neshoba_County,_Mississippi
8: 94768 Oktibbeha_County,_Mississippi
9: 94742 Tippah_County,_Mississippi
10: 149355 Dorothea_Dix
11: 106031 Meridian,_Mississippi
12: 28326263 Mississippi_State_Hospital
```

```
1: 28327487 Mississippi_Department_of_Mental_Health
    [winner: 28327487
```

```
Mississippi_Department_of_Mental_Health, pem_pos=0,
final_scores=-0.020590946, final_local_scores=0.0595631,
pem=0.0, similarity=5.8671947]
```

```
gt: 28327487 Mississippi_Department_of_Mental_Health,
pem_pos=0, final_scores=-0.020590946,
```



```

final_local_scores=0.0595631, pem=0.0, similarity=5.8671947 ]
2: 94836 Covington_County,_Mississippi
   [winner: 115276 Covington,_Kentucky, pem_pos=0,
final_scores=0.13637085, final_local_scores=0.10836181,
pem=-1.4255807, similarity=15.075976]
   gt: 94836 Covington_County,_Mississippi, pem_pos=8,
final_scores=-0.37464935, final_local_scores=-0.44701126,
pem=-2.975717, similarity=6.932417 ]
3: 94844 Choctaw_County,_Mississippi
   [winner: 7222 Choctaw, pem_pos=0,
final_scores=0.040286496, final_local_scores=0.03900917,
pem=-0.2529774, similarity=6.6295547]
   gt: 94844 Choctaw_County,_Mississippi, pem_pos=4,
final_scores=-0.41366488, final_local_scores=-0.48382896,
pem=-2.234952, similarity=1.805308 ]
4: 94842 Clarke_County,_Mississippi
   [winner: 96806 Clarke_County,_Georgia, pem_pos=3,
final_scores=-0.3806644, final_local_scores=-0.45514748,
pem=-3.0728943, similarity=7.218081]
   gt: 94842 Clarke_County,_Mississippi, pem_pos=2,
final_scores=-0.4556055, final_local_scores=-0.50840724,
pem=-2.6096716, similarity=3.109627 ]
5: 57775 Clay_County,_Mississippi
   [winner: 36980 Clay, pem_pos=0, final_scores=-0.31198952,
final_local_scores=-0.21452782, pem=-1.8237997,
similarity=7.605881]
   gt: 57775 Clay_County,_Mississippi, pem_pos=12,
final_scores=-0.60100293, final_local_scores=-0.66911066,
pem=-3.8110967, similarity=4.8474517 ]
6: 94832 Forrest_County,_Mississippi
   [winner: 94832 Forrest_County,_Mississippi, pem_pos=1,
final_scores=-0.10305299, final_local_scores=-0.1549497,
pem=-1.6931186, similarity=8.673147]
   gt: 94832 Forrest_County,_Mississippi, pem_pos=1,
final_scores=-0.10305299, final_local_scores=-0.1549497,
pem=-1.6931186, similarity=8.673147 ]
7: 94830 George_County,_Mississippi
   [winner: 4682118 George_Air_Force_Base, pem_pos=7,
final_scores=-0.4824019, final_local_scores=-0.43780145,
pem=-3.3049498, similarity=8.997674]
   gt: 94830 George_County,_Mississippi, pem_pos=9,
final_scores=-0.75691086, final_local_scores=-0.83481604,
pem=-3.816703, similarity=-0.06703761 ]
8: 94829 Greene_County,_Mississippi
   [winner: 77842 Greene_County,_Alabama, pem_pos=4,
final_scores=-0.35263664, final_local_scores=-0.4234197,
pem=-3.177409, similarity=8.733268]
   gt: 94829 Greene_County,_Mississippi, pem_pos=11,
final_scores=-0.5292198, final_local_scores=-0.6102277,
pem=-3.4434304, similarity=4.605211 ]
9: 94820 Jackson_County,_Mississippi
   [winner: 16167 Jackson,_Mississippi, pem_pos=0,

```

final_scores=-0.06894001, final_local_scores=-0.14022923,
pem=-2.1336284, similarity=11.507997|
RECALL MISS: 94820 Jackson_County,_Mississippi]

10: 94819 Jasper_County,_Mississippi
[winner: 136192 Jasper,_Texas, pem_pos=6,
final_scores=-0.31755018, final_local_scores=-0.3838142,
pem=-3.0658038, similarity=9.308249|
gt: 94819 Jasper_County,_Mississippi, pem_pos=11,
final_scores=-0.5288122, final_local_scores=-0.60146433,
pem=-3.4044101, similarity=4.6545305]

11: 94816 Jefferson_Davis_County,_Mississippi
[winner: 16134 Jefferson_Davis, pem_pos=0,
final_scores=0.16155969, final_local_scores=0.17516652,
pem=0.0, similarity=9.317031|
gt: 94816 Jefferson_Davis_County,_Mississippi,
pem_pos=2, final_scores=-0.2547679,
final_local_scores=-0.28771347, pem=-1.878611,
similarity=5.7199483]

12: 94814 Jones_County,_Mississippi
[winner: 95983 Jones_County,_Iowa, pem_pos=3,
final_scores=-0.5229087, final_local_scores=-0.57983303,
pem=-3.2642715, similarity=4.5379553|
gt: 94814 Jones_County,_Mississippi, pem_pos=5,
final_scores=-0.60604495, final_local_scores=-0.6908157,
pem=-3.6630518, similarity=3.394641]

13: 94810 Lamar_County,_Mississippi
[winner: 108621 Lamar,_Colorado, pem_pos=0,
final_scores=-0.109679624, final_local_scores=-0.12415394,
pem=-1.5721211, similarity=8.934151|
gt: 94810 Lamar_County,_Mississippi, pem_pos=2,
final_scores=-0.3054061, final_local_scores=-0.36366352,
pem=-2.1473339, similarity=4.914803]

14: 94777 Lowndes_County,_Mississippi
[winner: 94777 Lowndes_County,_Mississippi, pem_pos=0,
final_scores=-0.08313103, final_local_scores=-0.13176997,
pem=-1.3272722, similarity=7.3753467|
gt: 94777 Lowndes_County,_Mississippi, pem_pos=0,
final_scores=-0.08313103, final_local_scores=-0.13176997,
pem=-1.3272722, similarity=7.3753467]

15: 94775 Marion_County,_Mississippi
[winner: 112080 Marion,_Illinois, pem_pos=0,
final_scores=-0.0368167, final_local_scores=-0.06633365,
pem=-2.8057163, similarity=17.368105|
RECALL MISS: 94775 Marion_County,_Mississippi]

16: 94771 Newton_County,_Mississippi
[winner: 105945 Newton,_Massachusetts, pem_pos=2,
final_scores=-0.030468598, final_local_scores=-0.051041722,
pem=-2.4490397, similarity=15.884786|
gt: 94771 Newton_County,_Mississippi, pem_pos=12,
final_scores=-0.62748647, final_local_scores=-0.7049982,
pem=-4.2920284, similarity=6.3918786]

17: 94769 Noxubee_County,_Mississippi

[winner: 94769 Noxubee County, Mississippi, pem_pos=0,
final_scores=0.015546158, final_local_scores=0.0063441247,
pem=0.0, similarity=4.279034|
gt: 94769 Noxubee County, Mississippi, pem_pos=0,
final_scores=0.015546158, final_local_scores=0.0063441247,
pem=0.0, similarity=4.279034]
18: 94763 Perry County, Mississippi
[winner: 91935 Perry County, Pennsylvania, pem_pos=2,
final_scores=-0.3651067, final_local_scores=-0.3934734,
pem=-3.3755941, similarity=10.704687|
gt: 94763 Perry County, Mississippi, pem_pos=3,
final_scores=-0.46222186, final_local_scores=-0.52745557,
pem=-3.5940645, similarity=7.894469]
19: 94758 Prentiss County, Mississippi
[winner: 94758 Prentiss County, Mississippi, pem_pos=0,
final_scores=-0.050707057, final_local_scores=-0.07096936,
pem=-0.8635264, similarity=6.6678367|
gt: 94758 Prentiss County, Mississippi, pem_pos=0,
final_scores=-0.050707057, final_local_scores=-0.07096936,
pem=-0.8635264, similarity=6.6678367]
20: 94753 Scott County, Mississippi
[winner: 95527 Scott County, Kentucky, pem_pos=3,
final_scores=-0.39712787, final_local_scores=-0.47868982,
pem=-3.457054, similarity=8.60465|
gt: 94753 Scott County, Mississippi, pem_pos=9,
final_scores=-0.47046417, final_local_scores=-0.5381808,
pem=-4.077295, similarity=10.202288]
21: 94749 Smith County, Mississippi
[winner: 91449 Smith County, Texas, pem_pos=19,
final_scores=-0.4213547, final_local_scores=-0.4909772,
pem=-4.25418, similarity=12.5728655|
RECALL MISS: 94749 Smith County, Mississippi]
22: 94741 Tishomingo County, Mississippi
[winner: 130282 Tishomingo, Oklahoma, pem_pos=0,
final_scores=0.13553558, final_local_scores=0.116927505,
pem=-0.22434364, similarity=8.799078|
gt: 94741 Tishomingo County, Mississippi, pem_pos=1,
final_scores=0.035019144, final_local_scores=-0.0037010014,
pem=-0.683598, similarity=7.6967793]
23: 94731 Wayne County, Mississippi
[winner: 125671 Wayne, New Jersey, pem_pos=4,
final_scores=-0.2663964, final_local_scores=-0.28739464,
pem=-3.1917994, similarity=12.870784|
gt: 94731 Wayne County, Mississippi, pem_pos=19,
final_scores=-0.7713247, final_local_scores=-0.8556965,
pem=-4.3030267, similarity=1.95455]
24: 94729 Webster County, Mississippi
[winner: 97657 Webster Parish, Louisiana, pem_pos=7,
final_scores=-0.2353784, final_local_scores=-0.3054368,
pem=-3.345952, similarity=13.170676|
gt: 94729 Webster County, Mississippi, pem_pos=14,
final_scores=-0.5352177, final_local_scores=-0.59985733,

pem=-3.8044386, similarity=6.877901]
25: 94726 Winston_County,_Mississippi
[winner: 77366 Winston_County,_Alabama, pem_pos=2,
final_scores=-0.16562739, final_local_scores=-0.20335606,
pem=-2.2205565, similarity=10.096894]
gt: 94726 Winston_County,_Mississippi, pem_pos=7,
final_scores=-0.48310524, final_local_scores=-0.53557867,
pem=-2.7474399, similarity=3.047981]
1: 106031 Meridian,_Mississippi
[pem_pos=0, final_scores=0.13128392,
final_local_scores=0.08321938, pem=-1.4366089,
similarity=14.385648]
2: 16949861 Mississippi
[pem_pos=1, final_scores=0.32412064,
final_local_scores=0.27959475, pem=-0.31944752,
similarity=14.170583]
3: 94856 Alcorn_County,_Mississippi
[pem_pos=0, final_scores=0.071426496,
final_local_scores=0.05037582, pem=-0.73905826,
similarity=9.612142]
4: 94813 Kemper_County,_Mississippi
[pem_pos=0, final_scores=0.042428657,
final_local_scores=0.012663126, pem=-0.33444196,
similarity=6.2863536]
5: 94809 Lauderdale_County,_Mississippi
[pem_pos=0, final_scores=0.17298149,
final_local_scores=0.1490323, pem=-0.39773273,
similarity=10.700067]
6: 94781 Leake_County,_Mississippi
[pem_pos=0, final_scores=0.054660782,
final_local_scores=0.03168574, pem=-0.6044618,
similarity=8.322436]
7: 94772 Neshoba_County,_Mississippi
[pem_pos=0, final_scores=0.1859713,
final_local_scores=0.17411837, pem=0.0, similarity=9.285752]
8: 94768 Oktibbeha_County,_Mississippi
[pem_pos=0, final_scores=0.06557764,
final_local_scores=0.049206585, pem=0.0, similarity=5.5581355]
9: 94742 Tippah_County,_Mississippi
[pem_pos=0, final_scores=0.06795569,
final_local_scores=0.045317516, pem=0.0, similarity=5.442078]
10: 149355 Dorothea_Dix
[pem_pos=0, final_scores=0.3489458,
final_local_scores=0.42390132, pem=0.0, similarity=16.739775]
11: 106031 Meridian,_Mississippi
[pem_pos=0, final_scores=0.07973753,
final_local_scores=0.026912779, pem=-1.4366089,
similarity=12.705346]
12: 28326263 Mississippi_State_Hospital
[pem_pos=0, final_scores=0.044630274,
final_local_scores=0.114450395, pem=0.0, similarity=7.5051413]
1: 106031 Meridian,_Mississippi

[pem_pos=0, final_scores=0.074294165,
final_local_scores=0.020966738, pem=-1.4366089,
similarity=12.5279045]
2: 115276 Covington_Kentucky
[pem_pos=0, final_scores=0.13637085,
final_local_scores=0.10836181, pem=-1.4255807,
similarity=15.075976]
3: 7222 Choctaw
[pem_pos=0, final_scores=0.040286496,
final_local_scores=0.03900917, pem=-0.2529774,
similarity=6.6295547]
4: 16134 Jefferson_Davis
[pem_pos=0, final_scores=0.16155969,
final_local_scores=0.17516652, pem=0.0, similarity=9.317031]
5: 130282 Tishomingo_Oklahoma
[pem_pos=0, final_scores=0.13553558,
final_local_scores=0.116927505, pem=-0.22434364,
similarity=8.799078]
6: 150238 Ed_Emswiller
[pem_pos=0, final_scores=0.04495193,
final_local_scores=0.11520639, pem=-0.057158444,
similarity=7.838538]
7: 150238 Ed_Emswiller
[pem_pos=0, final_scores=0.049532518,
final_local_scores=0.12020999, pem=-0.057158444,
similarity=7.9878554]

Sample: b'Leighton_Hodges'

Leighton Hodges -LRB- born 25 November 1975 -RRB- is a [1 1 W
elsh 1] rugby referee who mainly referees club rugby in such
tournaments as the [1 1 Pro 12 1] . He has refereed 2 rugby se
vens matches , making his debut in the at [2 3 London 2] and [3
4 Edinburgh 3] . Hodges referred 4 matches at the 5 2011 [4
IRB Junior World Championship 4] in [2 2 Italy 2] -- [5 6 Ita
ly 5] v [6 7 New Zealand 6] , [7 8 England 7] v 9 Scotland , [8
10 Argentina 8] v [9 11 Italy 9] , and the third place game
12 Australia v [10 13 France 10] He refereed his
international 14 rugby union debut in the 15 2012 end of year
tests . There he refereed three matches , including [11 16 Un
ited States 11] v 17 Russia on 9 November 2012 in a match
that was part of the newly formed 18 2012 [12 International Ru
gby Series 12] , [3 3 Ireland Wolfhounds 3] v [13 19 Fiji 13]
on 17 November 2012 , and [14 20 United States 14] v. [15 21 R
omania 15]

1: 846910 Pro_12
2: 14532 Italy
3: 11365198 Ireland_Wolfhounds

1: 862966 Welsh_Rugby_Union

[winner: 69894 Wales, pem_pos=1, final_scores=0.52951866,
final_local_scores=0.54198474, pem=-1.0776843,
similarity=26.124233]

RECALL MISS: 862966 Welsh_Rugby_Union]

2: 512323 Rugby_sevens
 [winner: 19774201 Rugby_Virginia, pem_pos=7,
 final_scores=-0.8950008, final_local_scores=-0.7772067,
 pem=-2.956099, similarity=-3.0279608]
 gt: 512323 Rugby_sevens, pem_pos=19,
 final_scores=-2.1956272, final_local_scores=-2.442049,
 pem=-3.412151, similarity=-50.23009]
 [winner: 3679017 Sevens_(card_game), pem_pos=3,
 final_scores=-1.379013, final_local_scores=-1.3623548,
 pem=-1.84092, similarity=-26.554462]
 gt: 512323 Rugby_sevens, pem_pos=0,
 final_scores=-1.8046594, final_local_scores=-2.014976,
 pem=-0.85355985, similarity=-51.399403]

3: 21551889 2009_London_Sevens
 [winner: 17867 London, pem_pos=0,
 final_scores=0.65251213, final_local_scores=0.66964424,
 pem=-0.09086737, similarity=24.567383]
 RECALL MISS: 21551889 2009_London_Sevens]

4: 21551983 2009_Edinburgh_Sevens
 [winner: 9602 Edinburgh, pem_pos=0,
 final_scores=0.3948943, final_local_scores=0.40834832,
 pem=-0.11916062, similarity=16.923655]
 RECALL MISS: 21551983 2009_Edinburgh_Sevens]

5: 31159252 2011_IRB_Junior_World_Championship
 [winner: 15573676 IRB_Junior_World_Championship,
 pem_pos=0, final_scores=0.10536961,
 final_local_scores=0.1387333, pem=-0.015915431,
 similarity=8.316342]
 gt: 31159252 2011_IRB_Junior_World_Championship,
 pem_pos=3, final_scores=-0.7339897,
 final_local_scores=-0.7708624, pem=-5.247024,
 similarity=9.619775]
 [winner: 36225 2011, pem_pos=0, final_scores=-1.1666504,
 final_local_scores=-1.1070125, pem=-1.4232541,
 similarity=-21.205872]
 RECALL MISS: 31159252 2011_IRB_Junior_World_Championship]

6: 36231436 Italy_national_under-20_rugby_union_team
 [winner: 362466 Italy_national_football_team, pem_pos=1,
 final_scores=0.18768851, final_local_scores=0.17165363,
 pem=-2.9573529, similarity=25.294744]
 RECALL MISS: 36231436
 Italy_national_under-20_rugby_union_team]

7: 15243060 New_Zealand_national_under-20_rugby_union_team
 [winner: 154957 New_Zealand_national_rugby_union_team,
 pem_pos=4, final_scores=0.36592096,
 final_local_scores=0.35986447, pem=-3.8192244,
 similarity=35.59832]
 RECALL MISS: 15243060
 New_Zealand_national_under-20_rugby_union_team]

8: 23258121 England_national_under-20_rugby_union_team
 [winner: 9904 England_national_football_team, pem_pos=1,
 final_scores=0.07519175, final_local_scores=0.03689927,

pem=-2.776464, similarity=20.289703|
 RECALL MISS: 23258121
 England_national_under-20_rugby_union_team]
 9: 36231242 Scotland_national_under-20_rugby_union_team
 [winner: 758049 Scotland_national_rugby_union_team,
 pem_pos=2, final_scores=0.026891008,
 final_local_scores=-0.01953721, pem=-3.3931751,
 similarity=21.959297|
 RECALL MISS: 36231242
 Scotland_national_under-20_rugby_union_team]
 10: 36230933 Argentina_national_under-20_rugby_union_team
 [winner: 454699 Argentina_national_football_team,
 pem_pos=1, final_scores=0.09256016,
 final_local_scores=0.068279624, pem=-2.3085454,
 similarity=18.681541|
 RECALL MISS: 36230933
 Argentina_national_under-20_rugby_union_team]
 11: 36231436 Italy_national_under-20_rugby_union_team
 [winner: 14532 Italy, pem_pos=0, final_scores=0.27340615,
 final_local_scores=0.2614565, pem=-0.20612453,
 similarity=13.013033|
 RECALL MISS: 36231436
 Italy_national_under-20_rugby_union_team]
 12: 32131416 Australia_national_under-20_rugby_union_team
 [winner: 620967 Australia_national_rugby_union_team,
 pem_pos=3, final_scores=-0.17162427,
 final_local_scores=-0.2361756, pem=-4.217898,
 similarity=19.979347|
 RECALL MISS: 32131416
 Australia_national_under-20_rugby_union_team]
 13: 31859980 France_national_under-20_rugby_union_team
 [winner: 5843419 France, pem_pos=0,
 final_scores=0.4102196, final_local_scores=0.40199453,
 pem=-0.16240834, similarity=16.969233|
 RECALL MISS: 31859980
 France_national_under-20_rugby_union_team]
 14: 25405 Rugby_union
 [winner: 19774201 Rugby_Virginia, pem_pos=7,
 final_scores=-0.85495967, final_local_scores=-0.7334678,
 pem=-2.956099, similarity=-1.7227039|
 gt: 25405 Rugby_union, pem_pos=3,
 final_scores=-1.8630456, final_local_scores=-2.0800378,
 pem=-2.0110612, similarity=-47.0463]
 [winner: 481130 Union_(American_Civil_War), pem_pos=0,
 final_scores=-1.3514466, final_local_scores=-1.4108343,
 pem=-1.132797, similarity=-31.852068|
 RECALL MISS: 25405 Rugby_union]
 15: 35139300 2012_end-of-year_rugby_union_tests
 [winner: 47374 2012, pem_pos=0, final_scores=-0.73866135,
 final_local_scores=-0.53653145, pem=-1.4461981,
 similarity=-4.056806|
 RECALL MISS: 35139300 2012_end-of-

year_rugby_union_tests]
 [winner: 24088322 Tests_(album), pem_pos=2,
 final_scores=-1.0289501, final_local_scores=-0.88600045,
 pem=-1.0854045, similarity=-16.44771|
 RECALL MISS: 35139300 2012_end-of-
 year_rugby_union_tests]
 [winner: 294288 1973_in_art, pem_pos=9,
 final_scores=-1.0484709, final_local_scores=-0.9662922,
 pem=-3.2149317, similarity=-7.2630754|
 RECALL MISS: 35139300 2012_end-of-
 year_rugby_union_tests]
 [winner: 1603769 End_(American_football), pem_pos=0,
 final_scores=-1.3407434, final_local_scores=-1.4185877,
 pem=-0.72865474, similarity=-34.28123|
 RECALL MISS: 35139300 2012_end-of-year_rugby_union_tests]
 16: 1145226 United_States_national_rugby_union_team
 [winner: 3434750 United_States, pem_pos=0,
 final_scores=0.24359341, final_local_scores=0.22643453,
 pem=-0.073457584, similarity=11.246445|
 gt: 1145226 United_States_national_rugby_union_team,
 pem_pos=6, final_scores=-0.4802736,
 final_local_scores=-0.56533754, pem=-5.19937,
 similarity=15.493895]
 17: 4921535 Russia_national_rugby_union_team
 [winner: 25127146 Russia_national_football_team,
 pem_pos=2, final_scores=0.022860304,
 final_local_scores=0.011995614, pem=-3.5373242,
 similarity=23.684204|
 RECALL MISS: 4921535 Russia_national_rugby_union_team]
 18: 37106837 2012_International_Rugby_Series
 [winner: 37106837 2012_International_Rugby_Series,
 pem_pos=0, final_scores=0.07425557,
 final_local_scores=0.14784, pem=0.0, similarity=8.501554|
 gt: 37106837 2012_International_Rugby_Series, pem_pos=0,
 final_scores=0.07425557, final_local_scores=0.14784, pem=0.0,
 similarity=8.501554]
 [winner: 47374 2012, pem_pos=0, final_scores=-0.76320934,
 final_local_scores=-0.5633464, pem=-1.4461981,
 similarity=-4.857018|
 RECALL MISS: 37106837 2012_International_Rugby_Series]
 19: 1074616 Fiji_national_rugby_union_team
 [winner: 10707 Fiji, pem_pos=0, final_scores=0.14814894,
 final_local_scores=0.16498995, pem=-0.39221013,
 similarity=11.146242|
 RECALL MISS: 1074616 Fiji_national_rugby_union_team]
 20: 1145226 United_States_national_rugby_union_team
 [winner: 3434750 United_States, pem_pos=0,
 final_scores=0.40216488, final_local_scores=0.39964974,
 pem=-0.073457584, similarity=16.415533|
 gt: 1145226 United_States_national_rugby_union_team,
 pem_pos=6, final_scores=-0.5208979,
 final_local_scores=-0.60971344, pem=-5.19937,


```
similarity=14.169629 ]
21: 979061 Romania_national_rugby_union_team
    [winner: 25445 Romania, pem_pos=0,
    final_scores=0.46601838, final_local_scores=0.48285985,
    pem=-0.21241479, similarity=19.654358]
    gt: 979061 Romania_national_rugby_union_team, pem_pos=3,
    final_scores=-0.41068834, final_local_scores=-0.4678069,
    pem=-3.8066964, similarity=10.830828 ]
1: 846910 Pro_12
    [pem_pos=0, final_scores=0.039299563,
    final_local_scores=0.03097403, pem=-0.011834457,
    similarity=5.0783973]
2: 14532 Italy
    [pem_pos=0, final_scores=0.43622768,
    final_local_scores=0.4393143, pem=-0.20612453,
    similarity=18.320665]
3: 11365198 Ireland_Wolfhounds
    [pem_pos=1, final_scores=0.1872745,
    final_local_scores=0.20528328, pem=-0.0076628514,
    similarity=10.257447]
1: 69894 Wales
    [pem_pos=1, final_scores=0.52951866,
    final_local_scores=0.54198474, pem=-1.0776843,
    similarity=26.124233]
2: 17867 London
    [pem_pos=0, final_scores=0.65251213,
    final_local_scores=0.66964424, pem=-0.09086737,
    similarity=24.567383]
3: 9602 Edinburgh
    [pem_pos=0, final_scores=0.3948943,
    final_local_scores=0.40834832, pem=-0.11916062,
    similarity=16.923655]
4: 15573676 IRB_Junior_World_Championship
    [pem_pos=0, final_scores=0.10536961,
    final_local_scores=0.1387333, pem=-0.015915431,
    similarity=8.316342]
5: 362466 Italy_national_football_team
    [pem_pos=1, final_scores=0.18768851,
    final_local_scores=0.17165363, pem=-2.9573529,
    similarity=25.294744]
6: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.36592096,
    final_local_scores=0.35986447, pem=-3.8192244,
    similarity=35.59832]
7: 9904 England_national_football_team
    [pem_pos=1, final_scores=0.07519175,
    final_local_scores=0.03689927, pem=-2.776464,
    similarity=20.289703]
8: 454699 Argentina_national_football_team
    [pem_pos=1, final_scores=0.09256016,
    final_local_scores=0.068279624, pem=-2.3085454,
    similarity=18.681541]
```

9: 14532 Italy
 [pem_pos=0, final_scores=0.27340615,
 final_local_scores=0.2614565, pem=-0.20612453,
 similarity=13.013033]
 10: 5843419 France
 [pem_pos=0, final_scores=0.4102196,
 final_local_scores=0.40199453, pem=-0.16240834,
 similarity=16.969233]
 11: 3434750 United_States
 [pem_pos=0, final_scores=0.24359341,
 final_local_scores=0.22643453, pem=-0.073457584,
 similarity=11.246445]
 12: 37106837 2012_International_Rugby_Series
 [pem_pos=0, final_scores=0.07425557,
 final_local_scores=0.14784, pem=0.0, similarity=8.501554]
 13: 10707 Fiji
 [pem_pos=0, final_scores=0.14814894,
 final_local_scores=0.16498995, pem=-0.39221013,
 similarity=11.146242]
 14: 3434750 United_States
 [pem_pos=0, final_scores=0.40216488,
 final_local_scores=0.39964974, pem=-0.073457584,
 similarity=16.415533]
 15: 25445 Romania
 [pem_pos=0, final_scores=0.46601838,
 final_local_scores=0.48285985, pem=-0.21241479,
 similarity=19.654358]

Sample: b'Davit_Kubriashvili'

Davit Kubriashvili -LRB- born 12 March 1986 -RRB- is a [1 1 G eorgia 1] n 1 rugby union player . His position is and he currently plays for [2 2 RC Toulon 2] in the 2 Top 14 and the [1 3 Georgia 1] national team . He played for [2 Georgia 2] in the 4 2011 [3 Rugby World Cup 3] . After arriving in [4 France 4] from [5 Georgia 5] in Fédérale 1 to play for 5 Massy , aged 21 Kubriashvili moved up the leagues to the 6 Top 14 with [6 7 Montpellier 6] , however after one season there where he did n't get much game time he moved to newly promoted 8 Toulon in 2008 where he has since established himself as a regular member of their squad and has played over 100 matches for the club . After arriving before the , Kubriashvili soon worked his way to becoming first choice in 2009 and remained first choice for the following season . In 2010 , [7 Toulon 7] signed former [8 All Black 8] [3 3 Carl Ha yman 3] who was seen as one of the best tighthead props in world rugby , but Kubriashvili again kept his place as first choice for the keeping [9 Hayman 9] on the bench , and Kubriashvili was considered one the best tighthead props in the league that season . However after playing at the 9 2011 [10 World Cup 10] , Kubriashvili lost a bit of form and [11 Hayman 11] became first choice again and Kubriashvili has been the second choice since . Kubriashvili made his debut

for [12 10 Georgia 12] in 2008 against [13 11 Portugal 13] , but has largely had to play as second choice to [4 4 Davit Zi rakashvili 4] through most of his international career . In 2011 after a successful season for [14 Toulon 14] , Kubriashvili became the first Georgian to play for the [15 Barbarians 15] and he featured in their victories over [16 12 England 16] and [17 13 Wales 17]

1: 48768 Georgia_(country)

2: 2370506 RC_Toulonnais

3: 6535201 Carl_Hayman

4: 8174661 Davit_Zirakashvili

1: 25405 Rugby_union

[winner: 19774201 Rugby_Virginia, pem_pos=7, final_scores=-0.93021935, final_local_scores=-0.7999022, pem=-2.956099, similarity=-3.7052383|

gt: 25405 Rugby_union, pem_pos=3, final_scores=-1.9077208, final_local_scores=-2.1327808, pem=-2.0110612, similarity=-48.620255]

[winner: 481130 Union_(American_Civil_War), pem_pos=0, final_scores=-1.3303902, final_local_scores=-1.3833134, pem=-1.132797, similarity=-31.030792|

RECALL MISS: 25405 Rugby_union]

2: 972609 Top_14

[winner: 972609 Top_14, pem_pos=0, final_scores=-0.002967313, final_local_scores=-0.02722387, pem=0.0, similarity=3.2772985|

gt: 972609 Top_14, pem_pos=0, final_scores=-0.002967313, final_local_scores=-0.02722387, pem=0.0, similarity=3.2772985]

3: 565815 Georgia_national_rugby_union_team

[winner: 48768 Georgia_(country), pem_pos=2, final_scores=0.12239276, final_local_scores=0.14257705, pem=-1.3466994, similarity=15.6680565|

gt: 565815 Georgia_national_rugby_union_team, pem_pos=14, final_scores=-0.40040046, final_local_scores=-0.47810864, pem=-4.915364, similarity=16.55251]

[winner: 140801 Team, pem_pos=0, final_scores=-1.1684914, final_local_scores=-1.2256429, pem=-0.5173818, similarity=-29.672314|

RECALL MISS: 565815 Georgia_national_rugby_union_team]

[winner: 302395 National_(brand), pem_pos=3, final_scores=-1.4862915, final_local_scores=-1.494194, pem=-2.9526155, similarity=-24.443233|

RECALL MISS: 565815 Georgia_national_rugby_union_team]

4: 2179022 2011_Rugby_World_Cup

[winner: 25406 Rugby_World_Cup, pem_pos=0, final_scores=0.20415239, final_local_scores=0.17112833, pem=-0.078096084, similarity=9.621222|

gt: 2179022 2011_Rugby_World_Cup, pem_pos=2, final_scores=-0.13640614, final_local_scores=-0.19258496, pem=-1.9034305, similarity=8.693744]

[winner: 36225 2011, pem_pos=1, final_scores=-1.0619677,
final_local_scores=-0.9807681, pem=-1.4232541,
similarity=-17.438486]
RECALL MISS: 2179022 2011_Rugby_World_Cup]

5: 37629820 RC_Massy
[winner: 2016341 Massy, _Essonne, pem_pos=0,
final_scores=-0.24562114, final_local_scores=-0.19106215,
pem=-0.087658875, similarity=-1.1352602]
gt: 37629820 RC_Massy, pem_pos=13,
final_scores=-0.5265361, final_local_scores=-0.54331285,
pem=-3.925774, similarity=9.225143]

6: 972609 Top_14
[winner: 972609 Top_14, pem_pos=0,
final_scores=-0.035797954, final_local_scores=-0.06308636,
pem=0.0, similarity=2.2070901]
gt: 972609 Top_14, pem_pos=0, final_scores=-0.035797954,
final_local_scores=-0.06308636, pem=0.0,
similarity=2.2070901]

7: 1723001 Montpellier_Hérault_Rugby
[winner: 1640865 Montpellier_HSC, pem_pos=1,
final_scores=0.26472932, final_local_scores=0.2752185,
pem=-1.476352, similarity=20.331413]
gt: 1723001 Montpellier_Hérault_Rugby, pem_pos=2,
final_scores=-0.35040647, final_local_scores=-0.38692576,
pem=-2.2784064, similarity=4.9334054]

8: 2370506 RC_Toulonnais
[winner: 50992 Toulon, pem_pos=0,
final_scores=0.02055566, final_local_scores=0.0470493,
pem=-0.12135109, similarity=6.153684]
gt: 2370506 RC_Toulonnais, pem_pos=1,
final_scores=-0.112923205, final_local_scores=-0.15924674,
pem=-1.7526085, similarity=8.86843]

9: 2179022 2011_Rugby_World_Cup
[winner: 11370 FIFA_World_Cup, pem_pos=0,
final_scores=0.08600889, final_local_scores=0.0537522,
pem=-1.001087, similarity=11.137853]
RECALL MISS: 2179022 2011_Rugby_World_Cup]
[winner: 36225 2011, pem_pos=1, final_scores=-1.021956,
final_local_scores=-0.9370613, pem=-1.4232541,
similarity=-16.134188]
RECALL MISS: 2179022 2011_Rugby_World_Cup]

10: 565815 Georgia_national_rugby_union_team
[winner: 1008603 Georgia_national_football_team,
pem_pos=5, final_scores=0.041587368,
final_local_scores=0.02702105, pem=-3.736231,
similarity=25.214281]
gt: 565815 Georgia_national_rugby_union_team,
pem_pos=14, final_scores=-0.40221387,
final_local_scores=-0.48055422, pem=-4.915364,
similarity=16.479528]

11: 4919609 Portugal_national_rugby_union_team
[winner: 576328 Portugal_national_football_team,

pem_pos=1, final_scores=0.31679744,
final_local_scores=0.3074221, pem=-2.453637,
similarity=26.607061|
gt: 4919609 Portugal_national_rugby_union_team,
pem_pos=3, final_scores=-0.2682609,
final_local_scores=-0.31983095, pem=-4.167849,
similarity=17.210732]
12: 691024 England_national_rugby_union_team
[winner: 9904 England_national_football_team, pem_pos=1,
final_scores=0.13218845, final_local_scores=0.100482404,
pem=-2.776464, similarity=22.187151|
gt: 691024 England_national_rugby_union_team, pem_pos=3,
final_scores=-0.081250325, final_local_scores=-0.14947599,
pem=-3.7874734, similarity=20.22592]
13: 770930 Wales_national_rugby_union_team
[winner: 580291 Wales_national_football_team, pem_pos=1,
final_scores=0.036781028, final_local_scores=-0.0047880113,
pem=-2.4526134, similarity=17.284521|
gt: 770930 Wales_national_rugby_union_team, pem_pos=2,
final_scores=-0.027601942, final_local_scores=-0.08531919,
pem=-2.723462, similarity=16.354229]
1: 48768 Georgia_(country)
[pem_pos=2, final_scores=0.27105868,
final_local_scores=0.31032562, pem=-1.3466994,
similarity=20.67401]
2: 2370506 RC_Toulonnais
[pem_pos=0, final_scores=0.35595405,
final_local_scores=0.36054286, pem=0.0, similarity=14.849031]
3: 6535201 Carl_Hayman
[pem_pos=0, final_scores=0.44137293,
final_local_scores=0.49705195, pem=0.0, similarity=18.922735]
4: 8174661 Davit_Zirakashvili
[pem_pos=0, final_scores=0.64525473,
final_local_scores=0.7599915, pem=0.0, similarity=26.769377]
1: 48768 Georgia_(country)
[pem_pos=2, final_scores=0.12239276,
final_local_scores=0.14257705, pem=-1.3466994,
similarity=15.6680565]
2: 1008603 Georgia_national_football_team
[pem_pos=5, final_scores=0.10980497,
final_local_scores=0.102369905, pem=-3.736231,
similarity=27.462841]
3: 25406 Rugby_World_Cup
[pem_pos=0, final_scores=0.20415239,
final_local_scores=0.17112833, pem=-0.078096084,
similarity=9.621222]
4: 5843419 France
[pem_pos=0, final_scores=0.6942664,
final_local_scores=0.73545086, pem=-0.16240834,
similarity=26.920235]
5: 48768 Georgia_(country)
[pem_pos=1, final_scores=0.5743419,

final_local_scores=0.64161646, pem=-1.3466994,
 similarity=30.560387]
 6: 1640865 Montpellier_HSC
 [pem_pos=1, final_scores=0.26472932,
 final_local_scores=0.2752185, pem=-1.476352,
 similarity=20.331413]
 7: 6618038 Sporting_Toulon_Var
 [pem_pos=2, final_scores=0.11175822,
 final_local_scores=0.131852, pem=-1.877353,
 similarity=18.233778]
 8: 154957 New_Zealand_national_rugby_union_team
 [pem_pos=0, final_scores=0.25172263,
 final_local_scores=0.2206747, pem=0.0, similarity=10.675086]
 9: 6535201 Carl_Hayman
 [pem_pos=0, final_scores=0.33452195,
 final_local_scores=0.39430666, pem=0.0, similarity=15.85661]
 10: 11370 FIFA_World_Cup
 [pem_pos=0, final_scores=0.08600889,
 final_local_scores=0.0537522, pem=-1.001087,
 similarity=11.137853]
 11: 6535201 Carl_Hayman
 [pem_pos=0, final_scores=0.42652994,
 final_local_scores=0.49481142, pem=0.0, similarity=18.855873]
 12: 1008603 Georgia_national_football_team
 [pem_pos=5, final_scores=0.041587368,
 final_local_scores=0.02702105, pem=-3.736231,
 similarity=25.214281]
 13: 576328 Portugal_national_football_team
 [pem_pos=1, final_scores=0.31679744,
 final_local_scores=0.3074221, pem=-2.453637,
 similarity=26.607061]
 14: 6618038 Sporting_Toulon_Var
 [pem_pos=2, final_scores=0.03550367,
 final_local_scores=0.048555464, pem=-1.877353,
 similarity=15.748042]
 15: 569077 Barbarian_F.C.
 [pem_pos=0, final_scores=0.26790202,
 final_local_scores=0.26183572, pem=-0.17108923,
 similarity=12.833823]
 16: 9904 England_national_football_team
 [pem_pos=1, final_scores=0.13218845,
 final_local_scores=0.100482404, pem=-2.776464,
 similarity=22.187151]
 17: 580291 Wales_national_football_team
 [pem_pos=1, final_scores=0.036781028,
 final_local_scores=-0.0047880113, pem=-2.4526134,
 similarity=17.284521]

Sample: b'Laurie_Johnson'

_ -LRB- cricketer -RRB- Hubert Laurence Johnson -LRB- born 8
 November 1927 -RRB- is a [1 West Indies 1] born 1 cricket er
 who played for [1 1 Derbyshire 1] between 1949 and 1966 . He

scored over 14,000 runs for the club in the first-class game . Johnson was born at Pine Hill , 2 St Michael , [2 2 Barbados 2] . He was educated at , well known as the cradle of cricket in [2 Barbados 2] , where he benefited from the coaching of 3 Leslie Arthur `` Bessie `` Walcott . Johnson was invited to the BCA trials in 1945 and in 1946 he was part of the [3 Barbados 3] team to [4 Trinidad 4] , but did not play . However , he scored centuries against Combermere , 4 Spartan and 5 Wanderers , and then migrated to [5 England 5] to train as a sugar industry engineer . He played for the Swarkestone Cricket Club , and appeared for [6 Derbyshire 6] in 2nd XI matches in 1947 and 1948 . He made his first class debut against [3 3 Gloucestershire 3] in June 1949 and made 23 and 6 in the match . He played regularly in the first and second [7 Derbyshire 7] teams in 1949 and 1950 , but then returned to [4 4 British Guiana 4] . He reappeared for [8 Derbyshire 8] in 1955 and was a first team regular for the next ten years . He built up his scoring rate to reach over 1400 runs in 1959 and then became a top scorer . In 1960 he hit his first centuries with 140 against [5 5 Glamorgan 5] , 130 against [6 6 Essex 6] , 113 against [7 7 Lancashire 7] and 109 against [8 8 Leicestershire 8] . In 1961 he scored 122 against [9 Gloucestershire 9] , 119 against [10 Essex 10] , 116 against [11 Leicestershire 11] and 112 against [9 9 Hampshire 9] . In 1962 he made his top score of 154 against [12 Leicestershire 12] , and scored 108 against [10 10 Somerset 10] and 114 against [11 11 Sussex 11] . He scored five centuries in 1964 with 137 against [12 12 Nottinghamshire 12] , 132 against [13 Somerset 13] , 101 against [13 13 Warwickshire 13] , 101 not out against the [14 Australians 14] and 100 against [14 14 Middlesex 14] . The run rate fell back in 1965 and 966 and he ended his career at the end of the 1966 season . For the last four years he played in the ListA 6 Gillette Cup competition . Johnson was a right-hand batsman and played 606 innings in 351 first class matches to make 14,286 runs . He scored 16 centuries with a top score of 154 and an average of 26.40 . He was a right-arm off-break bowler and took 21 first class wickets with an average of 39.14 and a best performance of 3 for 12 . He also stood in as wicket keeper for [15 15 Bob Taylor 15]

- 1: 1829984 Derbyshire_County_Cricket_Club
- 2: 3455 Barbados
- 3: 1622322 Gloucestershire_County_Cricket_Club
- 4: 199278 British_Guiana
- 5: 1622320 Glamorgan_County_Cricket_Club
- 6: 1622252 Essex_County_Cricket_Club
- 7: 749919 Lancashire_County_Cricket_Club
- 8: 1622318 Leicestershire_County_Cricket_Club
- 9: 1919370 Hampshire_County_Cricket_Club
- 10: 1622178 Somerset_County_Cricket_Club
- 11: 1763691 Sussex_County_Cricket_Club
- 12: 667341 Nottinghamshire_County_Cricket_Club

13: 667292 Warwickshire_County_Cricket_Club
 14: 822103 Middlesex_County_Cricket_Club
 15: 2513807 Bob_Taylor_(cricketer)
 1: 25675557 Cricket
 [winner: 330091 Rhabdophoridae, pem_pos=10,
 final_scores=-1.3304976, final_local_scores=-1.2781799,
 pem=-2.56456, similarity=-20.107248|
 gt: 25675557 Cricket, pem_pos=0,
 final_scores=-1.3517034, final_local_scores=-1.5237193,
 pem=-0.07532314, similarity=-40.97149]
 2: 551655 Saint_Michael,_Barbados
 [winner: 551655 Saint_Michael,_Barbados, pem_pos=1,
 final_scores=-0.038328692, final_local_scores=-0.037576914,
 pem=-1.7600404, similarity=12.539716|
 gt: 551655 Saint_Michael,_Barbados, pem_pos=1,
 final_scores=-0.038328692, final_local_scores=-0.037576914,
 pem=-1.7600404, similarity=12.539716]
 3: 3842929 Leslie_Walcott
 [winner: 3842929 Leslie_Walcott, pem_pos=0,
 final_scores=-0.121387675, final_local_scores=-0.07705967,
 pem=0.0, similarity=1.7900987|
 gt: 3842929 Leslie_Walcott, pem_pos=0,
 final_scores=-0.121387675, final_local_scores=-0.07705967,
 pem=0.0, similarity=1.7900987]
 4: 32260258 Spartan_Cricket_Club
 [winner: 5263232 Spartan_League, pem_pos=17,
 final_scores=-0.38854343, final_local_scores=-0.38996306,
 pem=-2.9723463, similarity=8.616519|
 RECALL MISS: 32260258 Spartan_Cricket_Club]
 5: 32271463 Wanderers_Cricket_Club
 [winner: 1598477 Wanderers_F.C., pem_pos=0,
 final_scores=-0.059146762, final_local_scores=-0.0601016,
 pem=-0.6287954, similarity=5.7156487|
 gt: 32271463 Wanderers_Cricket_Club, pem_pos=13,
 final_scores=-0.7284263, final_local_scores=-0.7624879,
 pem=-4.306225, similarity=4.753473]
 6: 1899702 Friends_Provident_Trophy
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=-0.61499596, final_local_scores=-0.73740196,
 pem=-0.08377493, similarity=-17.460253|
 gt: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=-0.61499596, final_local_scores=-0.73740196,
 pem=-0.08377493, similarity=-17.460253]
 1: 1829984 Derbyshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.62902814,
 final_local_scores=0.58067936, pem=-1.4828352,
 similarity=29.482233]
 2: 3455 Barbados
 [pem_pos=0, final_scores=0.12136333,
 final_local_scores=0.13060749, pem=-0.3637528,
 similarity=9.9654455]
 3: 1622322 Gloucestershire_County_Cricket_Club

[pem_pos=1, final_scores=0.54727954,
final_local_scores=0.4900458, pem=-1.379793,
similarity=26.217186]
4: 199278 British_Guiana
[pem_pos=0, final_scores=0.1309141,
final_local_scores=0.1584008, pem=0.0, similarity=8.816709]
5: 1622320 Glamorgan_County_Cricket_Club
[pem_pos=1, final_scores=0.7798124,
final_local_scores=0.75197226, pem=-0.7903395,
similarity=30.828053]
6: 1622252 Essex_County_Cricket_Club
[pem_pos=1, final_scores=0.9133242,
final_local_scores=0.89229804, pem=-2.3459642,
similarity=43.475384]
7: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.6875058,
final_local_scores=0.64801216, pem=-1.7295884,
similarity=32.833458]
8: 1622318 Leicestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.714117,
final_local_scores=0.6804194, pem=-1.0465667,
similarity=30.086176]
9: 1919370 Hampshire_County_Cricket_Club
[pem_pos=1, final_scores=0.64978576,
final_local_scores=0.60758406, pem=-1.9232318,
similarity=32.680065]
10: 1622178 Somerset_County_Cricket_Club
[pem_pos=1, final_scores=0.71313673,
final_local_scores=0.67135656, pem=-2.4139037,
similarity=37.251514]
11: 1763691 Sussex_County_Cricket_Club
[pem_pos=1, final_scores=0.49908572,
final_local_scores=0.45816922, pem=-1.5566913,
similarity=26.227926]
12: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.5828174,
final_local_scores=0.53428197, pem=-1.1067253,
similarity=26.052298]
13: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.8254128,
final_local_scores=0.79832447, pem=-1.1093602,
similarity=33.946182]
14: 822103 Middlesex_County_Cricket_Club
[pem_pos=1, final_scores=0.6031488,
final_local_scores=0.55114615, pem=-1.5695794,
similarity=29.07263]
15: 2513807 Bob_Taylor_(cricketer)
[pem_pos=0, final_scores=0.40701777,
final_local_scores=0.43124253, pem=-0.89295405,
similarity=21.814873]
1: 294567 British_West_Indies
[pem_pos=4, final_scores=0.07467358,

final_local_scores=0.13036537, pem=-3.0585637,
similarity=24.613022]
2: 3455 Barbados
[pem_pos=0, final_scores=0.2272604,
final_local_scores=0.24628395, pem=-0.3637528,
similarity=13.417461]
3: 3455 Barbados
[pem_pos=0, final_scores=0.2093186,
final_local_scores=0.22668526, pem=-0.3637528,
similarity=12.832598]
4: 452912 Trinidad
[pem_pos=0, final_scores=0.14296357,
final_local_scores=0.18899998, pem=-0.34123632,
similarity=11.585546]
5: 9316 England
[pem_pos=0, final_scores=0.505619,
final_local_scores=0.5319402, pem=-0.19053856,
similarity=21.000048]
6: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.75691444,
final_local_scores=0.7203757, pem=-1.4828352,
similarity=33.651047]
7: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.64249104,
final_local_scores=0.5953856, pem=-1.4828352,
similarity=29.921093]
8: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6461738,
final_local_scores=0.59940845, pem=-1.4828352,
similarity=30.041145]
9: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.563193,
final_local_scores=0.5074289, pem=-1.379793,
similarity=26.735933]
10: 1622252 Essex_County_Cricket_Club
[pem_pos=1, final_scores=0.85707754,
final_local_scores=0.83085716, pem=-2.3459642,
similarity=41.641865]
11: 1622318 Leicestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.8164306,
final_local_scores=0.79218143, pem=-1.0465667,
similarity=33.42138]
12: 1622318 Leicestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.775508,
final_local_scores=0.74747974, pem=-1.0465667,
similarity=32.08739]
13: 1622178 Somerset_County_Cricket_Club
[pem_pos=1, final_scores=0.69161195,
final_local_scores=0.64784396, pem=-2.4139037,
similarity=36.54985]
14: 374833 Australia_national_cricket_team
[pem_pos=2, final_scores=0.5663454,

final_local_scores=0.5430053, pem=-1.4598138,
similarity=28.232769]

Sample: b'Greater_London_Council_election'

, _ 1970 The third election to the [1 1 Greater London Council 1] was held on 9 April 1970 and saw a [1 Conservative 1] victory with a reduced majority . In addition to the 100 councillors , there were sixteen who divided 11 [2 Conservative 2] and 5 [3 Labour 3] , so that the [4 Conservatives 4] actually had 76 seats to 40 for Labour following the election . The poll in [5 1 Hammersmith 5] was delayed to 27 April due to the death of one of the [6 Labour 6] candidates . The original result for 2 Harrow turned out to have an enormous mistake and a recount was ordered on 20 April which did not alter the result in terms of those elected but did identify more than 25,000 votes which had been incorrectly recorded . With an electorate of 5,524,384 , there was a turnout of 35.2 % . [7 Labour 7] recovered from its mauling three years previously , but did so primarily in working-class areas . Consequently , relatively few seats changed hands : [8 Labour 8] won back 3 Camden , [9 4 Greenwich 9] , [10 5 Hammersmith 10] , 6 Wandsworth , and one seat in [11 7 Lambeth 11] . The results did enable [12 Labour 12] to take back control of the [2 2 Inner London Education Authority 2] and were one of the factors used by 8 Prime Minister [3 3 Harold Wilson 3] in deciding to call a 9 general election soon after . Among those who were first elected to the GLC in 1970 were 10 Tony Banks -LRB- [13 Labour 13] , [14 11 Hammersmith 14] , later Minister for Sport -RRB- and 12 Sir George Young -LRB- [15 Conservative 15] , [16 13 Ealing 16] , later a cabinet minister under [4 4 John Major 4] -RRB- . The election is also significant as it was at a meeting in support of the [17 Conservative 17] candidates in [18 14 Lambeth 18] that [19 John Major 19] met [5 5 Norma Johnson 5] , who became his wife . These parties were created by a group of students standing in [6 6 Haringey 6] , who declared that they intended to make a mockery of the election . No seats changed hands in byelections during this term . The [20 Conservatives 20] retained on 2 December 1971 after the death of Seton Forbes-Cockell , and 15 Barnet on 19 October 1972 after the death of Arthur Peacock . [21 Labour 21] retained 16 Wandsworth

1: 61268 Greater_London_Council

2: 163714 Inner_London_Education_Authority

3: 52271 Harold_Wilson

4: 15898 John_Major

5: 1783408 Norma_Major

6: 94306 London_Borough_of_Haringey

1: 95390 London_Borough_of_Hammersmith_and_Fulham

[winner: 192166 Hammersmith, pem_pos=0,
final_scores=0.3527524, final_local_scores=0.33510602,
pem=-0.014193171, similarity=14.1671295]

gt: 95390 London_Borough_of_Hammersmith_and_Fulham,
pem_pos=9, final_scores=-0.62879175,
final_local_scores=-0.7287019, pem=-5.433722,
similarity=11.893221]
2: 94427 London_Borough_of_Harrow
[winner: 188956 Harrow_School, pem_pos=0,
final_scores=-0.05831732, final_local_scores=-0.10453881,
pem=-0.7965306, similarity=5.3017235]
gt: 94427 London_Borough_of_Harrow, pem_pos=2,
final_scores=-0.51446235, final_local_scores=-0.5840428,
pem=-1.3742449, similarity=-5.8659306]
3: 26598633 London_Borough_of_Camden
[winner: 26598633 London_Borough_of_Camden, pem_pos=1,
final_scores=-0.18457709, final_local_scores=-0.2558608,
pem=-1.8784031, similarity=6.6693645]
gt: 26598633 London_Borough_of_Camden, pem_pos=1,
final_scores=-0.18457709, final_local_scores=-0.2558608,
pem=-1.8784031, similarity=6.6693645]
4: 93948 Royal_Borough_of_Greenwich
[winner: 95840 Greenwich, pem_pos=0,
final_scores=0.03131263, final_local_scores=0.0037761778,
pem=-0.5636761, similarity=7.26776]
gt: 93948 Royal_Borough_of_Greenwich, pem_pos=3,
final_scores=-0.29705578, final_local_scores=-0.34059355,
pem=-2.923222, similarity=9.822658]
5: 95390 London_Borough_of_Hammersmith_and_Fulham
[winner: 192166 Hammersmith, pem_pos=0,
final_scores=0.083617136, final_local_scores=0.04111661,
pem=-0.014193171, similarity=5.393899]
gt: 95390 London_Borough_of_Hammersmith_and_Fulham,
pem_pos=9, final_scores=-0.7040778,
final_local_scores=-0.81094044, pem=-5.433722,
similarity=9.439058]
6: 95361 London_Borough_of_Wandsworth
[winner: 94365 Wandsworth, pem_pos=0,
final_scores=0.019347414, final_local_scores=-0.024875745,
pem=-0.33546227, similarity=5.171667]
gt: 95361 London_Borough_of_Wandsworth, pem_pos=1,
final_scores=-0.13187605, final_local_scores=-0.18523218,
pem=-1.2775803, similarity=5.509695]
7: 95362 London_Borough_of_Lambeth
[winner: 298747 Lambeth, pem_pos=0,
final_scores=0.23634006, final_local_scores=0.20618868,
pem=-0.31364357, similarity=11.948437]
gt: 95362 London_Borough_of_Lambeth, pem_pos=1,
final_scores=-0.011258081, final_local_scores=-0.062666655,
pem=-1.6327958, similarity=11.099012]
8: 24150 Prime_Minister_of_the_United_Kingdom
[winner: 1248688 Prime_(New_Zealand), pem_pos=1,
final_scores=-0.6432801, final_local_scores=-0.6056334,
pem=-2.1396754, similarity=-2.3477085]
RECALL MISS: 24150 Prime_Minister_of_the_United_Kingdom]

[winner: 739237 Minister_(government), pem_pos=0,
final_scores=-1.210765, final_local_scores=-1.321888,
pem=-0.49105036, similarity=-32.687653|
RECALL MISS: 24150 Prime_Minister_of_the_United_Kingdom]
9: 159614 United_Kingdom_general_election,_1970
[winner: 12795 General_election, pem_pos=0,
final_scores=-0.6784188, final_local_scores=-0.7266624,
pem=-0.5947298, similarity=-14.361109|
RECALL MISS: 159614
United_Kingdom_general_election,_1970]
10: 414437 Tony_Banks,_Baron_Stratford
[winner: 270262 Tony_Banks_(musician), pem_pos=0,
final_scores=-0.11571066, final_local_scores=-0.032823585,
pem=0.0, similarity=3.1101916|
gt: 414437 Tony_Banks,_Baron_Stratford, pem_pos=2,
final_scores=-0.15758911, final_local_scores=-0.1871138,
pem=-0.8918017, similarity=3.35562]
11: 95390 London_Borough_of_Hammersmith_and_Fulham
[winner: 192166 Hammersmith, pem_pos=0,
final_scores=0.11240013, final_local_scores=0.07255764,
pem=-0.014193171, similarity=6.3321624|
gt: 95390 London_Borough_of_Hammersmith_and_Fulham,
pem_pos=9, final_scores=-0.7649912,
final_local_scores=-0.87747914, pem=-5.433722,
similarity=7.4534097]
12: 414480 Sir_George_Young,_6th_Baronet
[winner: 414480 Sir_George_Young,_6th_Baronet, pem_pos=0,
final_scores=-0.11501549, final_local_scores=-0.15944147,
pem=0.0, similarity=-0.6683386|
gt: 414480 Sir_George_Young,_6th_Baronet, pem_pos=0,
final_scores=-0.11501549, final_local_scores=-0.15944147,
pem=0.0, similarity=-0.6683386]
13: 94074 London_Borough_of_Ealing
[winner: 150301 Ealing, pem_pos=0,
final_scores=0.09984161, final_local_scores=0.058998078,
pem=-0.21419199, similarity=7.0151424|
gt: 94074 London_Borough_of_Ealing, pem_pos=1,
final_scores=-0.13491683, final_local_scores=-0.17109837,
pem=-1.4627016, similarity=6.938195]
14: 95362 London_Borough_of_Lambeth
[winner: 298747 Lambeth, pem_pos=0,
final_scores=0.3459571, final_local_scores=0.32770193,
pem=-0.31364357, similarity=15.5746355|
gt: 95362 London_Borough_of_Lambeth, pem_pos=1,
final_scores=0.05255498, final_local_scores=0.013686568,
pem=-1.6327958, similarity=13.377544]
15: 94333 London_Borough_of_Barnet
[winner: 451173 Barnet_F.C., pem_pos=0,
final_scores=0.021827027, final_local_scores=0.06515397,
pem=-0.6797195, similarity=9.730457|
gt: 94333 London_Borough_of_Barnet, pem_pos=2,
final_scores=-0.33451328, final_local_scores=-0.35439634,

pem=-2.1153727, similarity=5.0175447]
16: 95361 London_Borough_of_Wandsworth
[winner: 94365 Wandsworth, pem_pos=0,
final_scores=-0.12165347, final_local_scores=-0.1788978,
pem=-0.33546227, similarity=0.575342]
gt: 95361 London_Borough_of_Wandsworth, pem_pos=1,
final_scores=-0.2760387, final_local_scores=-0.342708,
pem=-1.2775803, similarity=0.8103034]
1: 61268 Greater_London_Council
[pem_pos=0, final_scores=0.24112044,
final_local_scores=0.21880886, pem=-0.0051736925,
similarity=10.647541]
2: 163714 Inner_London_Education_Authority
[pem_pos=0, final_scores=0.26948637,
final_local_scores=0.2649749, pem=0.0, similarity=11.997092]
3: 52271 Harold_Wilson
[pem_pos=0, final_scores=0.40270257,
final_local_scores=0.37522393, pem=0.0, similarity=15.287143]
4: 15898 John_Major
[pem_pos=0, final_scores=0.4096676,
final_local_scores=0.39390618, pem=0.0, similarity=15.844659]
5: 1783408 Norma_Major
[pem_pos=1, final_scores=0.067197725,
final_local_scores=0.07943472, pem=-0.5389965,
similarity=9.391351]
6: 94306 London_Borough_of_Haringey
[pem_pos=0, final_scores=0.25655442,
final_local_scores=0.2509135, pem=0.0, similarity=11.577473]
1: 32113 Conservative_Party_(UK)
[pem_pos=0, final_scores=0.09539835,
final_local_scores=0.0061776936, pem=-0.55914843,
similarity=7.3148036]
2: 32113 Conservative_Party_(UK)
[pem_pos=0, final_scores=0.10639028,
final_local_scores=0.018184721, pem=-0.55914843,
similarity=7.673117]
3: 19279158 Labour_Party_(UK)
[pem_pos=0, final_scores=0.050291672,
final_local_scores=-0.042522296, pem=-0.3416913,
similarity=4.678933]
4: 32113 Conservative_Party_(UK)
[pem_pos=0, final_scores=0.12143113,
final_local_scores=0.034614563, pem=-0.44721895,
similarity=7.5547266]
5: 192166 Hammersmith
[pem_pos=0, final_scores=0.3527524,
final_local_scores=0.33510602, pem=-0.014193171,
similarity=14.1671295]
6: 40312187 Israeli_Labor_Party
[pem_pos=7, final_scores=0.033809707,
final_local_scores=0.07278758, pem=-3.1950269,
similarity=23.636892]

7: 19279158 Labour_Party_(UK)
[pem_pos=0, final_scores=0.08787377,
final_local_scores=0.0012388229, pem=-0.3416913,
similarity=5.9848523]
8: 19279158 Labour_Party_(UK)
[pem_pos=0, final_scores=0.10740666,
final_local_scores=0.022575542, pem=-0.3416913,
similarity=6.6215825]
9: 95840 Greenwich
[pem_pos=0, final_scores=0.03131263,
final_local_scores=0.0037761778, pem=-0.5636761,
similarity=7.26776]
10: 192166 Hammersmith
[pem_pos=0, final_scores=0.083617136,
final_local_scores=0.04111661, pem=-0.014193171,
similarity=5.393899]
11: 298747 Lambeth
[pem_pos=0, final_scores=0.23634006,
final_local_scores=0.20618868, pem=-0.31364357,
similarity=11.948437]
12: 19279158 Labour_Party_(UK)
[pem_pos=0, final_scores=0.11196299,
final_local_scores=0.02755265, pem=-0.3416913,
similarity=6.770109]
13: 19279158 Labour_Party_(UK)
[pem_pos=0, final_scores=0.04034303,
final_local_scores=-0.0533897, pem=-0.3416913,
similarity=4.354628]
14: 192166 Hammersmith
[pem_pos=0, final_scores=0.11240013,
final_local_scores=0.07255764, pem=-0.014193171,
similarity=6.3321624]
15: 32113 Conservative_Party_(UK)
[pem_pos=0, final_scores=0.0928099,
final_local_scores=0.003350228, pem=-0.55914843,
similarity=7.230427]
16: 150301 Ealing
[pem_pos=0, final_scores=0.09984161,
final_local_scores=0.058998078, pem=-0.21419199,
similarity=7.0151424]
17: 32113 Conservative_Party_(UK)
[pem_pos=0, final_scores=0.097681716,
final_local_scores=0.008671939, pem=-0.55914843,
similarity=7.389237]
18: 298747 Lambeth
[pem_pos=0, final_scores=0.3459571,
final_local_scores=0.32770193, pem=-0.31364357,
similarity=15.5746355]
19: 15898 John_Major
[pem_pos=0, final_scores=0.40838242,
final_local_scores=0.39074975, pem=0.0, similarity=15.750464]
20: 32113 Conservative_Party_(UK)

[pem_pos=0, final_scores=0.102652535,
final_local_scores=0.014101803, pem=-0.44721895,
similarity=6.942585]
21: 19279158 Labour_Party_(UK)
[pem_pos=0, final_scores=0.20015164,
final_local_scores=0.12388533, pem=-0.3416913,
similarity=9.644869]

Sample: b'1996_Big_12_Championship_Game'

The 1996 Dr. Pepper Big 12 Championship Game was held on December 7 , 1996 at [1 1 Trans World Dome 1] -LRB- now [1 Edward Jones Dome 1] -RRB- in . The divisional winners from the [2 2 Big 12 Conference 2] squared off in the 1st edition of the game . The [2 1 Texas Longhorns 2] represented the [3 South Division 3] and the [4 2 Nebraska Cornhuskers 4] represented the [5 North Division 5] . [6 Texas 6] won the contest 37 -- 27 , keyed by a from their own 28-yard line with slightly more than 2 minutes left in the 4th quarter when the score was 30-27 in their favor . The matched up the winner of the North and South divisions of the [3 3 Big 12 Conference 3] . The game was first played in 3 1996 , when the conference was assembled to include all of the teams from the [4 4 Big Eight Conference 4] as well as four teams that had formerly been members of the [5 5 Southwest Conference 5] . The championship game was modeled on the 4 SEC format , which was the first conference in college football to have 5 a conference championship game . Today , six conferences in the top-level hold championship games -- the [7 6 ACC 7] , [8 7 Big Ten 8] , [9 8 C-USA 9] , 9 MAC , [10 SEC 10] , and [11 10 Pac-12 11] . However , the [12 Big 12 12] no longer holds such a game . After the 11 2010 season , two [13 Big 12 13] members , 12 Colorado and 13 Nebraska , left for other conferences as part of . As a result , the 14 2010 edition of this game was the last for the foreseeable future . The 15 2011 season

1: 452735 Edward_Jones_Dome

2: 78362 Big_12_Conference

3: 78362 Big_12_Conference

4: 415250 Big_Eight_Conference

5: 2005590 Southwest_Conference

1: 36981279 1996_Texas_Longhorns_football_team

[winner: 2834813 Texas_Longhorns, pem_pos=0,
final_scores=0.22787087, final_local_scores=0.18361175,
pem=-0.50125355, similarity=12.294949]

RECALL MISS: 36981279 1996_Texas_Longhorns_football_team]

2: 22894201 1996_Nebraska_Cornhuskers_football_team

[winner: 159450 Nebraska_Cornhuskers, pem_pos=0,
final_scores=0.22904961, final_local_scores=0.18397078,
pem=-0.55016685, similarity=12.571661]

gt: 22894201 1996_Nebraska_Cornhuskers_football_team,
pem_pos=8, final_scores=-0.4047746,
final_local_scores=-0.43849084, pem=-3.1058836,

similarity=7.89455]
 3: 4166606 1996_NCAA_Division_I-A_football_season
 [winner: 34636 1996, pem_pos=0, final_scores=-0.6974426,
 final_local_scores=-0.50211185, pem=-0.78063476,
 similarity=-6.649093]
 RECALL MISS: 4166606 1996_NCAA_Division_I-
 A_football_season]
 4: 79371 Southeastern_Conference
 [winner: 79371 Southeastern_Conference, pem_pos=0,
 final_scores=-0.01842089, final_local_scores=-0.10929851,
 pem=-0.43542784, similarity=3.1959517]
 gt: 79371 Southeastern_Conference, pem_pos=0,
 final_scores=-0.01842089, final_local_scores=-0.10929851,
 pem=-0.43542784, similarity=3.1959517]
 5: 3135359 SEC_Championship_Game
 [winner: 3135359 SEC_Championship_Game, pem_pos=0,
 final_scores=-1.0778738, final_local_scores=-1.2120645,
 pem=0.0, similarity=-32.080708]
 gt: 3135359 SEC_Championship_Game, pem_pos=0,
 final_scores=-1.0778738, final_local_scores=-1.2120645,
 pem=0.0, similarity=-32.080708]
 6: 8006361 ACC_Championship_Game
 [winner: 78023 Atlantic_Coast_Conference, pem_pos=0,
 final_scores=0.13620035, final_local_scores=0.06839983,
 pem=-0.32217512, similarity=7.8829384]
 RECALL MISS: 8006361 ACC_Championship_Game]
 7: 28272857 Big_Ten_Football_Championship_Game
 [winner: 78358 Big_Ten_Conference, pem_pos=0,
 final_scores=0.2638232, final_local_scores=0.20710224,
 pem=0.0, similarity=10.270057]
 gt: 28272857 Big_Ten_Football_Championship_Game,
 pem_pos=17, final_scores=-1.1269152,
 final_local_scores=-1.2500579, pem=-7.5430083,
 similarity=7.8055468]
 8: 7092507 Conference_USA_Football_Championship_Game
 [winner: 78428 Conference_USA, pem_pos=0,
 final_scores=0.11061503, final_local_scores=0.043349385,
 pem=0.0, similarity=5.383345]
 gt: 7092507 Conference_USA_Football_Championship_Game,
 pem_pos=2, final_scores=-0.76254845,
 final_local_scores=-0.85239875, pem=-4.5666604,
 similarity=3.4866433]
 9: 3597670 MAC_Championship_Game
 [winner: 19006979 Macintosh, pem_pos=0,
 final_scores=-0.14850013, final_local_scores=-0.04014186,
 pem=-0.48007193, similarity=5.502506]
 RECALL MISS: 3597670 MAC_Championship_Game]
 10: 29049711 Pacific-12_Football_Championship_Game
 [winner: 78455 Pacific-12_Conference, pem_pos=0,
 final_scores=0.116616115, final_local_scores=0.043234423,
 pem=-0.20016693, similarity=6.4684534]
 gt: 29049711 Pacific-12_Football_Championship_Game,

```

pem_pos=2, final_scores=-0.870297,
final_local_scores=-0.95212, pem=-4.9904327,
similarity=2.815302 ]
11: 26186877 2010_NCAA_Division_I_FBS_football_season
[winner: 43226 2010, pem_pos=0, final_scores=-1.0650116,
final_local_scores=-0.94035435, pem=-1.6731367,
similarity=-14.873559]
RECALL MISS: 26186877
2010_NCAA_Division_I_FBS_football_season]
[winner: 690345 Season(sports), pem_pos=0,
final_scores=-1.6123023, final_local_scores=-1.760351,
pem=-1.219945, similarity=-41.80842]
RECALL MISS: 26186877
2010_NCAA_Division_I_FBS_football_season]
12: 5740017 Colorado_Buffaloes_football
[winner: 5399 Colorado, pem_pos=0,
final_scores=-0.057775706, final_local_scores=-0.00059995055,
pem=-0.26926568, similarity=5.536118]
gt: 5740017 Colorado_Buffaloes_football, pem_pos=3,
final_scores=-0.30069536, final_local_scores=-0.38781327,
pem=-3.5544286, similarity=11.846127 ]
13: 4108481 Nebraska_Cornhuskers_football
[winner: 21647 Nebraska, pem_pos=0,
final_scores=-0.09615983, final_local_scores=-0.05141601,
pem=-0.4226844, similarity=4.85398]
gt: 4108481 Nebraska_Cornhuskers_football, pem_pos=1,
final_scores=-0.2236598, final_local_scores=-0.3206975,
pem=-3.0283203, similarity=10.987936 ]
14: 28911258 2010_Big_12_Championship_Game
[winner: 28302093 2010_Military_Bowl, pem_pos=0,
final_scores=-0.7759112, final_local_scores=-0.842337,
pem=0.0, similarity=-21.047304]
RECALL MISS: 28911258 2010_Big_12_Championship_Game]
15: 30246548 2011_NCAA_Division_I_FBS_football_season
[winner: 30863870
2011_Thai_Division_2_League_North_Eastern_Region, pem_pos=0,
final_scores=-0.8794391, final_local_scores=-0.86434627,
pem=0.0, similarity=-21.704103]
RECALL MISS: 30246548
2011_NCAA_Division_I_FBS_football_season]
1: 452735 Edward_Jones_Dome
[pem_pos=0, final_scores=0.2189927,
final_local_scores=0.22554207, pem=0.0, similarity=10.820338]
2: 78362 Big_12_Conference
[pem_pos=0, final_scores=0.30024868,
final_local_scores=0.22447446, pem=0.0, similarity=10.788479]
3: 78362 Big_12_Conference
[pem_pos=0, final_scores=0.19832842,
final_local_scores=0.11314207, pem=0.0, similarity=7.466098]
4: 415250 Big_Eight_Conference
[pem_pos=0, final_scores=0.24287818,
final_local_scores=0.1890131, pem=0.0, similarity=9.730241]

```

5: 2005590 Southwest_Conference
[pem_pos=0, final_scores=0.17484169,
final_local_scores=0.10274342, pem=0.0, similarity=7.1557813]
1: 452735 Edward_Jones_Dome
[pem_pos=0, final_scores=0.18284784,
final_local_scores=0.18605933, pem=0.0, similarity=9.642095]
2: 2834813 Texas_Longhorns
[pem_pos=0, final_scores=0.22787087,
final_local_scores=0.18361175, pem=-0.50125355,
similarity=12.294949]
3: 1952084 AFC_South
[pem_pos=0, final_scores=0.048450187,
final_local_scores=0.049992755, pem=-0.35126805,
similarity=7.4918466]
4: 159450 Nebraska_Cornhuskers
[pem_pos=0, final_scores=0.22904961,
final_local_scores=0.18397078, pem=-0.55016685,
similarity=12.571661]
5: 1952065 AFC_North
[pem_pos=1, final_scores=0.08081709,
final_local_scores=0.058424473, pem=-0.35354918,
similarity=7.7558713]
6: 30857 Texas_Rangers_(baseball)
[pem_pos=13, final_scores=0.056238458,
final_local_scores=0.06414598, pem=-4.928425,
similarity=32.805496]
7: 78023 Atlantic_Coast_Conference
[pem_pos=0, final_scores=0.13620035,
final_local_scores=0.06839983, pem=-0.32217512,
similarity=7.8829384]
8: 78358 Big_Ten_Conference
[pem_pos=0, final_scores=0.2638232,
final_local_scores=0.20710224, pem=0.0, similarity=10.270057]
9: 78428 Conference_USA
[pem_pos=0, final_scores=0.11061503,
final_local_scores=0.043349385, pem=0.0, similarity=5.383345]
10: 79371 Southeastern_Conference
[pem_pos=0, final_scores=0.066318795,
final_local_scores=-0.016733274, pem=-0.43542784,
similarity=5.958283]
11: 78455 Pacific-12_Conference
[pem_pos=0, final_scores=0.116616115,
final_local_scores=0.043234423, pem=-0.20016693,
similarity=6.4684534]
12: 78362 Big_12_Conference
[pem_pos=0, final_scores=0.07763888,
final_local_scores=-0.021197341, pem=0.0,
similarity=3.4571419]
13: 78362 Big_12_Conference
[pem_pos=0, final_scores=0.12986712,
final_local_scores=0.038358524, pem=0.0, similarity=5.234408]

Sample: b'Calumet_Region'

The [1 Calumet Region 1] is the name given to the geographic areas drained by the and the of northwestern [1 1 Indiana 1] in the [2 United States 2] . It is part of the [2 2 Great Lakes Basin 2] , which eventually reaches the Atlantic Ocean . This region includes the northern parts of 1 Lake and 2 Porter counties in [3 Indiana 3] , as well as the easternmost counties of northern [4 Illinois 4] . Since much of this region is on the south shore of [3 3 Lake Michigan 3] , it is sometimes referred to as the `` [5 South Shore 5] '' . Because it was initially cut off from the rest of the state due to natural geographic barriers like the [6 Kankakee Marsh 6] to the south , the [7 Calumet Region 7] was the last-settled portion of [8 Indiana 8] . It is a sub-region of the greater [4 4 Northwest Indiana 4] region and the even larger [9 3 Great Lakes 9] region . The [10 Calumet Region 10] includes the [11 Indiana 11] cities of [5 5 Gary 5] , [6 6 East Chicago 6] , [7 7 Hammond 7] , [12 4 Hobart 12] , 5 Whiting , [8 8 Crown Point 8] , and [13 6 Valparaiso 13] . The [14 Calumet Region 14] includes land from the following townships in [15 Porter County 15] : 7 Portage , 8 Union , [16 9 Westchester 16] , 10 Liberty , 11 Center , 12 Pine , 13 Jackson , [17 14 Washington 17] . It includes land located within the following townships of [18 Lake County 18] : 15 North , 16 St. John , 17 Calumet , [19 18 Hobart 19] , and 19 Ross . The name Calumet is said to come from [20 French 20] interpretations of either the [21 20 Delaware 21] or [9 9 Potawatomi 9] names for the rivers in question , or is a corruption of the Old French term Chalemel , which means `` reed '' . The word appears on early maps as Cal-La-Mick , Kil-La-Mick , Calumic , etc. . The first known print reference to refer to this area as a distinct geographical region is the 1755 map created by John Mitchell . In this map , however , he referred to the geographic region as `` Quadoche '' , a name that the [10 10 Iroquois 10] had given to the 21 Huron

1: 21883857 Indiana

2: 10200423 Great_Lakes_Basin

3: 17948 Lake_Michigan

4: 3652659 Northwest_Indiana

5: 12650 Gary,_Indiana

6: 112413 East_Chicago,_Indiana

7: 112415 Hammond,_Indiana

8: 112411 Crown_Point,_Indiana

9: 351662 Potawatomi

10: 19195965 Iroquois

1: 96102 Lake_County,_Indiana

[winner: 96102 Lake_County,_Indiana, pem_pos=5,
final_scores=-0.3002342, final_local_scores=-0.3861146,
pem=-3.117275, similarity=9.519509]

gt: 96102 Lake_County,_Indiana, pem_pos=5,
final_scores=-0.3002342, final_local_scores=-0.3861146,
pem=-3.117275, similarity=9.519509]

2: 96080 Porter_County,_Indiana
 [winner: 96080 Porter_County,_Indiana, pem_pos=2,
 final_scores=-0.22655676, final_local_scores=-0.30197877,
 pem=-3.0391955, similarity=11.605682|
 gt: 96080 Porter_County,_Indiana, pem_pos=2,
 final_scores=-0.22655676, final_local_scores=-0.30197877,
 pem=-3.0391955, similarity=11.605682]

3: 432536 Great_Lakes_region
 [winner: 12010 Great_Lakes, pem_pos=0,
 final_scores=0.2969007, final_local_scores=0.29199696,
 pem=-0.09949329, similarity=13.344544|
 gt: 432536 Great_Lakes_region, pem_pos=1,
 final_scores=-0.13095225, final_local_scores=-0.16984221,
 pem=-2.4588683, similarity=12.39299]
 [winner: 2125190 Regions_of_Togo, pem_pos=18,
 final_scores=-1.3942642, final_local_scores=-1.3203542,
 pem=-3.5198548, similarity=-16.170773|
 RECALL MISS: 432536 Great_Lakes_region]

4: 112417 Hobart,_Indiana
 [winner: 13699 Hobart, pem_pos=0,
 final_scores=0.22387283, final_local_scores=0.24126172,
 pem=-0.38929793, similarity=13.406507|
 gt: 112417 Hobart,_Indiana, pem_pos=1,
 final_scores=-0.25890046, final_local_scores=-0.31445163,
 pem=-2.60779, similarity=8.887417]

5: 112428 Whiting,_Indiana
 [winner: 112428 Whiting,_Indiana, pem_pos=1,
 final_scores=0.021290645, final_local_scores=0.009272367,
 pem=-1.7850239, similarity=14.073656|
 gt: 112428 Whiting,_Indiana, pem_pos=1,
 final_scores=0.021290645, final_local_scores=0.009272367,
 pem=-1.7850239, similarity=14.073656]

6: 112550 Valparaiso,_Indiana
 [winner: 70566 Valparaíso, pem_pos=0,
 final_scores=0.23938425, final_local_scores=0.27357626,
 pem=-0.48874232, similarity=14.911631|
 gt: 112550 Valparaiso,_Indiana, pem_pos=1,
 final_scores=0.017631963, final_local_scores=-0.036492586,
 pem=-1.494299, similarity=11.126932]

7: 13745232 Portage_Township,_Porter_County,_Indiana
 [winner: 138875 Portage,_Wisconsin, pem_pos=3,
 final_scores=-0.06858912, final_local_scores=-0.11305708,
 pem=-2.2836843, similarity=13.134894|
 gt: 13745232 Portage_Township,_Porter_County,_Indiana,
 pem_pos=14, final_scores=-0.5008289,
 final_local_scores=-0.50195783, pem=-3.2644558,
 similarity=6.8629093]

8: 13745249 Union_Township,_Porter_County,_Indiana
 [winner: 3434750 United_States, pem_pos=2,
 final_scores=-0.12060547, final_local_scores=-0.18243048,
 pem=-2.977907, similarity=14.839946|
 RECALL MISS: 13745249

Union_Township,_Porter_County,_Indiana]

9: 13745265 Westchester_Township,_Porter_County,_Indiana
 [winner: 48367 Westchester_County,_New_York, pem_pos=0,
 final_scores=0.23624693, final_local_scores=0.20619312,
 pem=-0.6583652, similarity=13.823219|
 gt: 13745265
 Westchester_Township,_Porter_County,_Indiana, pem_pos=6,
 final_scores=-0.20258181, final_local_scores=-0.21326923,
 pem=-2.102641, similarity=9.159822]

10: 13745213 Liberty_Township,_Porter_County,_Indiana
 [winner: 122628 Liberty,_Missouri, pem_pos=1,
 final_scores=-0.45082444, final_local_scores=-0.5067565,
 pem=-2.801813, similarity=4.2037826|
 RECALL MISS: 13745213
 Liberty_Township,_Porter_County,_Indiana]

11: 13745193 Center_Township,_Porter_County,_Indiana
 [winner: 652569 Center_(basketball), pem_pos=0,
 final_scores=-0.18421666, final_local_scores=-0.11719811,
 pem=-1.6317346, similarity=9.465914|
 RECALL MISS: 13745193
 Center_Township,_Porter_County,_Indiana]

12: 13745222 Pine_Township,_Porter_County,_Indiana
 [winner: 39389 Pine, pem_pos=0, final_scores=-0.29333633,
 final_local_scores=-0.25906944, pem=-1.1852114,
 similarity=2.8039277|
 gt: 13745222 Pine_Township,_Porter_County,_Indiana,
 pem_pos=11, final_scores=-0.6481157,
 final_local_scores=-0.6534628, pem=-3.5339787,
 similarity=3.807406]

13: 13745207 Jackson_Township,_Porter_County,_Indiana
 [winner: 16167 Jackson,_Mississippi, pem_pos=0,
 final_scores=-0.10088754, final_local_scores=-0.15835693,
 pem=-2.1336284, similarity=10.96703|
 RECALL MISS: 13745207
 Jackson_Township,_Porter_County,_Indiana]

14: 13745261 Washington_Township,_Porter_County,_Indiana
 [winner: 108956 Washington,_D.C., pem_pos=0,
 final_scores=0.06814389, final_local_scores=0.026783943,
 pem=-1.7496481, similarity=14.403857|
 RECALL MISS: 13745261
 Washington_Township,_Porter_County,_Indiana]

15: 13745137 North_Township,_Lake_County,_Indiana
 [winner: 21650 North_Carolina, pem_pos=5,
 final_scores=-0.27562767, final_local_scores=-0.3616569,
 pem=-3.4297683, similarity=11.948763|
 RECALL MISS: 13745137
 North_Township,_Lake_County,_Indiana]

16: 13745157 St._John_Township,_Lake_County,_Indiana
 [winner: 1039408 Saint_John,_U.S._Virgin_Islands,
 pem_pos=0, final_scores=0.0063105077,
 final_local_scores=0.009384632, pem=-1.6205289,
 similarity=13.182457|

gt: 13745157 St._John_Township,_Lake_County,_Indiana,
pem_pos=10, final_scores=-0.31338024,
final_local_scores=-0.28299809, pem=-2.8685744,
similarity=11.24424]
17: 13745061 Calumet_Township,_Lake_County,_Indiana
[winner: 99047 Calumet,_Michigan, pem_pos=2,
final_scores=-0.06325768, final_local_scores=-0.089131296,
pem=-1.6584138, similarity=10.448569]
gt: 13745061 Calumet_Township,_Lake_County,_Indiana,
pem_pos=9, final_scores=-0.4084527,
final_local_scores=-0.38257152, pem=-2.5098257,
similarity=6.321838]
18: 13745128 Hobart_Township,_Lake_County,_Indiana
[winner: 13699 Hobart, pem_pos=0, final_scores=0.1199566,
final_local_scores=0.12774903, pem=-0.38929793,
similarity=10.019061]
gt: 13745128 Hobart_Township,_Lake_County,_Indiana,
pem_pos=7, final_scores=-0.64001805,
final_local_scores=-0.60554653, pem=-3.515367,
similarity=5.1361084]
19: 13745143 Ross_Township,_Lake_County,_Indiana
[winner: 417624 Ross, pem_pos=0,
final_scores=-0.38180134, final_local_scores=-0.35391775,
pem=-1.9870358, similarity=4.333911]
RECALL MISS: 13745143
Ross_Township,_Lake_County,_Indiana]
20: 230270 Lenape
[winner: 7930 Delaware, pem_pos=0,
final_scores=0.21631928, final_local_scores=0.16964975,
pem=-0.5414915, similarity=12.0971155]
gt: 230270 Lenape, pem_pos=1, final_scores=-0.35443956,
final_local_scores=-0.39764515, pem=-3.4374223,
similarity=10.916426]
21: 140832 Wyandot_people
[winner: 51056 Lake_Huron, pem_pos=3,
final_scores=-0.295155, final_local_scores=-0.33640736,
pem=-2.3050523, similarity=6.585879]
gt: 140832 Wyandot_people, pem_pos=11,
final_scores=-0.32994682, final_local_scores=-0.36491358,
pem=-3.111064, similarity=10.118414]
1: 21883857 Indiana
[pem_pos=0, final_scores=0.5261436,
final_local_scores=0.4849437, pem=-0.27230176,
similarity=20.04222]
2: 10200423 Great_Lakes_Basin
[pem_pos=0, final_scores=0.22942464,
final_local_scores=0.28928417, pem=0.0, similarity=12.722529]
3: 17948 Lake_Michigan
[pem_pos=0, final_scores=0.43437117,
final_local_scores=0.4166605, pem=0.0, similarity=16.523693]
4: 3652659 Northwest_Indiana
[pem_pos=0, final_scores=0.35714668,

final_local_scores=0.34184593, pem=0.0, similarity=14.291078]
5: 12650 Gary,_Indiana
[pem_pos=0, final_scores=0.31092697,
final_local_scores=0.2686027, pem=-0.50188345,
similarity=14.834673]
6: 112413 East_Chicago,_Indiana
[pem_pos=0, final_scores=0.42676193,
final_local_scores=0.41485882, pem=0.0, similarity=16.469927]
7: 112415 Hammond,_Indiana
[pem_pos=0, final_scores=0.104226515,
final_local_scores=0.0583452, pem=-1.4899528,
similarity=13.933445]
8: 112411 Crown_Point,_Indiana
[pem_pos=0, final_scores=0.28715056,
final_local_scores=0.28408265, pem=-0.65300006,
similarity=16.118422]
9: 351662 Potawatomi
[pem_pos=0, final_scores=0.26023835,
final_local_scores=0.25590166, pem=0.0, similarity=11.726329]
10: 19195965 Iroquois
[pem_pos=0, final_scores=0.2665816,
final_local_scores=0.29030225, pem=-0.19568375,
similarity=13.817069]
1: 3652659 Northwest_Indiana
[pem_pos=0, final_scores=0.29920226,
final_local_scores=0.27855045, pem=0.0, similarity=12.402214]
2: 3434750 United_States
[pem_pos=0, final_scores=0.47220212,
final_local_scores=0.46989346, pem=-0.073457584,
similarity=18.511745]
3: 21883857 Indiana
[pem_pos=0, final_scores=0.5130655,
final_local_scores=0.47065783, pem=-0.27230176,
similarity=19.6159]
4: 14849 Illinois
[pem_pos=0, final_scores=0.5011683,
final_local_scores=0.4583513, pem=-0.16995399,
similarity=18.692066]
5: 953415 South_Shore,_Chicago
[pem_pos=0, final_scores=0.08158104,
final_local_scores=0.08090353, pem=-1.2455401,
similarity=13.277476]
6: 15341085 Kankakee_Outwash_Plain
[pem_pos=0, final_scores=0.31742626,
final_local_scores=0.37423676, pem=0.0, similarity=15.257684]
7: 3652659 Northwest_Indiana
[pem_pos=0, final_scores=0.3148291,
final_local_scores=0.2956204, pem=0.0, similarity=12.911615]
8: 21883857 Indiana
[pem_pos=0, final_scores=0.4801494,
final_local_scores=0.43470198, pem=-0.27230176,
similarity=18.542906]

9: 12010 Great_Lakes
 [pem_pos=0, final_scores=0.2969007,
 final_local_scores=0.29199696, pem=-0.09949329,
 similarity=13.344544]
 10: 3652659 Northwest_Indiana
 [pem_pos=0, final_scores=0.36288255,
 final_local_scores=0.34811148, pem=0.0, similarity=14.478054]
 11: 21883857 Indiana
 [pem_pos=0, final_scores=0.40969628,
 final_local_scores=0.35774267, pem=-0.27230176,
 similarity=16.246286]
 12: 13699 Hobart
 [pem_pos=0, final_scores=0.22387283,
 final_local_scores=0.24126172, pem=-0.38929793,
 similarity=13.406507]
 13: 70566 Valparaíso
 [pem_pos=0, final_scores=0.23938425,
 final_local_scores=0.27357626, pem=-0.48874232,
 similarity=14.911631]
 14: 3652659 Northwest_Indiana
 [pem_pos=0, final_scores=0.41293257,
 final_local_scores=0.4027835, pem=0.0, similarity=16.109577]
 15: 96080 Porter_County,_Indiana
 [pem_pos=0, final_scores=0.26418632,
 final_local_scores=0.23864621, pem=0.0, similarity=11.211391]
 16: 48367 Westchester_County,_New_York
 [pem_pos=0, final_scores=0.23624693,
 final_local_scores=0.20619312, pem=-0.6583652,
 similarity=13.823219]
 17: 108956 Washington,_D.C.
 [pem_pos=0, final_scores=0.06814389,
 final_local_scores=0.026783943, pem=-1.7496481,
 similarity=14.403857]
 18: 96102 Lake_County,_Indiana
 [pem_pos=3, final_scores=0.09412165,
 final_local_scores=0.051428378, pem=-1.6649153,
 similarity=14.678507]
 19: 13699 Hobart
 [pem_pos=0, final_scores=0.1199566,
 final_local_scores=0.12774903, pem=-0.38929793,
 similarity=10.019061]
 20: 5843419 France
 [pem_pos=0, final_scores=0.31088966,
 final_local_scores=0.34486985, pem=-0.80223274,
 similarity=18.743984]
 21: 7930 Delaware
 [pem_pos=0, final_scores=0.21631928,
 final_local_scores=0.16964975, pem=-0.5414915,
 similarity=12.0971155]

Sample: b'Darren_Shadford'

Darren James Shadford -LRB- born 4 March 1975 -RRB- is a

former [1 1 English 1] 2 cricket er . Shadford is a right-handed who bowls right-arm 3 medium-fast . He was born at [1 1 Oldham 1] , 4 Greater Manchester . Shadford made his debut for [2 2 Lancashire 2] in a [3 3 List A 3] match against [4 4 Gloucestershire 4] in the . He made two further appearances in that seasons competition , against [5 5 Nottinghamshire 5] and [6 6 Durham 6] . He made his 5 first-class debut the following season against [7 7 Essex 7] in the , with Shadford making a further first-class appearance in 1995 against [8 8 Surrey 8] . He did n't appear for [2 Lancashire 2] in 1996 , but did feature for the county in 1997 , making eight first-class and eight [3 List A 3] appearances . However , in 1998 he made just a single first-class appearance against [4 Essex 4] in the , as well as making a single [5 List A 5] appearance against [9 9 Sussex 9] in the [6 6 AXA League 6] . It was also in 1998 that he made a single appearance for the [10 10 Lancashire Cricket Board 10] in the [11 11 MCCA Knockout Trophy 11] against [12 12 Shropshire 12] . He left [7 Lancashire 7] at the end of the 7 1999 season . Since his debut for the county in 1994 , Shadford made eleven first-class appearances for the county , taking 23 wickets at an 8 average of 42.73 , with best figures of 5/80 . This was his only and came against [13 13 Warwickshire 13] in 1997 . With the bat , he scored 120 runs at a 9 batting average

1: 179445 Oldham

2: 749919 Lancashire_County_Cricket_Club

3: 683459 List_A_cricket

4: 1622322 Gloucestershire_County_Cricket_Club

5: 667341 Nottinghamshire_County_Cricket_Club

6: 1176371 Durham_County_Cricket_Club

7: 1622252 Essex_County_Cricket_Club

8: 766575 Surrey_County_Cricket_Club

9: 1763691 Sussex_County_Cricket_Club

10: 27793434 Lancashire_Cricket_Board

11: 7543416 MCCA_Knockout_Trophy

12: 3702447 Shropshire_County_Cricket_Club

13: 667292 Warwickshire_County_Cricket_Club

1: 18803164 English_people

[winner: 9316 England, pem_pos=1, final_scores=0.532094, final_local_scores=0.57786316, pem=-1.3527626, similarity=28.690836]

gt: 18803164 English_people, pem_pos=2, final_scores=-0.41312593, final_local_scores=-0.44894034, pem=-2.2084396, similarity=2.702275]

2: 25675557 Cricket

[winner: 330091 Rhaphidophoridae, pem_pos=10, final_scores=-1.4694947, final_local_scores=-1.4261112, pem=-2.56456, similarity=-24.521818]

gt: 25675557 Cricket, pem_pos=0, final_scores=-1.7673416, final_local_scores=-1.9719133, pem=-0.07532314, similarity=-54.346497]

3: 689836 Seam_bowling

[winner: 678695 Fast_bowling, pem_pos=0,
final_scores=-0.8577326, final_local_scores=-0.9927245,
pem=0.0, similarity=-25.535166]
RECALL MISS: 689836 Seam_bowling]

4: 79819 Greater_Manchester
[winner: 79819 Greater_Manchester, pem_pos=0,
final_scores=0.017768964, final_local_scores=0.034356907,
pem=0.0, similarity=5.1149917]
gt: 79819 Greater_Manchester, pem_pos=0,
final_scores=0.017768964, final_local_scores=0.034356907,
pem=0.0, similarity=5.1149917]

5: 11040 First-class_cricket
[winner: 662844 First-class_citizen, pem_pos=3,
final_scores=-1.277025, final_local_scores=-1.1981988,
pem=-1.5935546, similarity=-23.00093]
gt: 11040 First-class_cricket, pem_pos=0,
final_scores=-1.2772774, final_local_scores=-1.4738853,
pem=-0.053819276, similarity=-39.601288]

6: 30679337 1998_AXA_League
[winner: 1177737 Pro40, pem_pos=0,
final_scores=0.2931345, final_local_scores=0.27385774,
pem=0.0, similarity=12.262175]
RECALL MISS: 30679337 1998_AXA_League]

7: 7635102 1999_English_cricket_season
[winner: 1269575 1998-99_FA_Premier_League, pem_pos=0,
final_scores=-1.1626945, final_local_scores=-1.2149999,
pem=0.0, similarity=-32.16831]
RECALL MISS: 7635102 1999_English_cricket_season]

8: 875871 Bowling_average
[winner: 875871 Bowling_average, pem_pos=0,
final_scores=-0.9666564, final_local_scores=-1.101682,
pem=0.0, similarity=-28.786674]
gt: 875871 Bowling_average, pem_pos=0,
final_scores=-0.9666564, final_local_scores=-1.101682,
pem=0.0, similarity=-28.786674]

9: 18933273 Batting_average
[winner: 15652278 Batting_(material), pem_pos=2,
final_scores=-0.95831186, final_local_scores=-0.8561532,
pem=-1.1616125, similarity=-15.142576]
gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.2676781, final_local_scores=-1.4106045,
pem=-0.29142123, similarity=-36.420742]
[winner: 875871 Bowling_average, pem_pos=1,
final_scores=-1.1507887, final_local_scores=-1.3028185,
pem=-0.8657023, similarity=-30.081163]
gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.2126478, final_local_scores=-1.3504922,
pem=-0.5012068, similarity=-33.48603]

1: 179445 Oldham
[pem_pos=0, final_scores=0.08598383,
final_local_scores=0.10147461, pem=-0.46675745,
similarity=9.656218]

2: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.6075911,
final_local_scores=0.552065, pem=-1.7295884,
similarity=29.970205]
3: 683459 List_A_cricket
[pem_pos=0, final_scores=0.21437971,
final_local_scores=0.14908817, pem=-0.020267159,
similarity=8.649017]
4: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.594209,
final_local_scores=0.5406686, pem=-1.379793,
similarity=27.727873]
5: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.66067034,
final_local_scores=0.61341536, pem=-1.1067253,
similarity=28.413795]
6: 1176371 Durham_County_Cricket_Club
[pem_pos=2, final_scores=0.693634,
final_local_scores=0.65653056, pem=-2.3952909,
similarity=36.70786]
7: 1622252 Essex_County_Cricket_Club
[pem_pos=1, final_scores=0.9737219,
final_local_scores=0.95419484, pem=-2.3459642,
similarity=45.32251]
8: 766575 Surrey_County_Cricket_Club
[pem_pos=1, final_scores=0.7628823,
final_local_scores=0.7242096, pem=-1.7952642,
similarity=35.464497]
9: 1763691 Sussex_County_Cricket_Club
[pem_pos=1, final_scores=0.60568964,
final_local_scores=0.5684116, pem=-1.5566913,
similarity=29.517778]
10: 27793434 Lancashire_Cricket_Board
[pem_pos=0, final_scores=0.2227139,
final_local_scores=0.21318558, pem=0.0, similarity=10.451596]
11: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.33365303,
final_local_scores=0.29326582, pem=0.0, similarity=12.84135]
12: 3702447 Shropshire_County_Cricket_Club
[pem_pos=1, final_scores=0.19260667,
final_local_scores=0.1444155, pem=-2.279101,
similarity=20.793465]
13: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.78119963,
final_local_scores=0.7433219, pem=-1.1093602,
similarity=32.304794]
1: 9316 England
[pem_pos=1, final_scores=0.532094,
final_local_scores=0.57786316, pem=-1.3527626,
similarity=28.690836]
2: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.591657,

```

final_local_scores=0.5346594, pem=-1.7295884,
similarity=29.450787]
3: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.24060385,
final_local_scores=0.17773405, pem=-0.020267159,
similarity=9.503867]
4: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.9256601,
final_local_scores=0.9016946, pem=-2.3459642,
similarity=43.755795]
5: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.27862072,
final_local_scores=0.21926168, pem=-0.020267159,
similarity=10.7431345]
6: 1177737 Pro40
    [pem_pos=0, final_scores=0.2931345,
final_local_scores=0.27385774, pem=0.0, similarity=12.262175]
7: 749919 Lancashire_County_Cricket_Club
    [pem_pos=1, final_scores=0.471546,
final_local_scores=0.4034564, pem=-1.7295884,
similarity=25.535427]

```

Sample: b'Electoral_district_of_Mount_Hawthorn'

The [1 Electoral 1] district of [2 Mount Hawthorn 2] was a 1 Legislative Assembly 2 electorate in the 3 state of [1 1 West ern Australia 1] . The district was named for the inner northern suburb of [3 4 Mount Hawthorn 3] , which fell within its borders . [4 Mount Hawthorn 4] was a new seat created under the Redistribution of Seats Act 1929 , which took effect at the . The [2 2 Labor 2] member for 5 Leederville and 6 Collier Government minister , ran for and won the seat , holding it until his retirement from politics at the 7 1947 election , in which [5 Labor 5] lost government after 14 years in office . of the [3 3 Liberal Party 3] narrowly won the seat , but with a redistribution prior to the 8 1950 elec tion reducing the likelihood of a repeat , Nimmo opted to contest the new seat of [4 4 Wembley Beaches 4] . The redistribution had also merged the 9 North-West Labor-held seats of [6 10 Pilbara 6] and [7 11 Roebourne 7] , so Pilbara MLA contested and won Mount Hawthorn , which he held until his retirement in 1968 . then held the seat until its abolishment at the 12 1983 election . The redistribution which effected the end of the seat split its voters between the seat of [8 13 Subiaco 8] , 14 Balcatta -LRB- which Bertram contested and won -RRB- and [9 15 Perth 9]

```

1: 33613 Western_Australia
2: 1495 Australian_Labor_Party
3: 18453 Liberal_Party_of_Australia
4: 19246345 Electoral_district_of_Wembley_Beaches
1: 579448 Western_Australian_Legislative_Assembly
    [winner: 579433 New_South_Wales_Legislative_Assembly,
pem_pos=1, final_scores=-0.35075504,

```

```

final_local_scores=-0.4309219, pem=-1.9991245,
similarity=2.1016939|
    gt: 579448 Western_Australian_Legislative_Assembly,
pem_pos=3, final_scores=-0.43944544,
final_local_scores=-0.5324647, pem=-2.598477,
similarity=2.3308258 ]
2: 1726491 Electoral_districts_of_Western_Australia
    [winner: 3236919 New_Zealand_electorates, pem_pos=0,
final_scores=-0.8791917, final_local_scores=-0.95282483,
pem=-0.60937357, similarity=-21.030613|
    gt: 1726491 Electoral_districts_of_Western_Australia,
pem_pos=2, final_scores=-1.1603682,
final_local_scores=-1.2485936, pem=-1.646076,
similarity=-24.219193 ]
3: 4689096 States_and_territories_of_Australia
    [winner: 229379 States_of_Brazil, pem_pos=0,
final_scores=-1.4868279, final_local_scores=-1.4918029,
pem=0.0, similarity=-40.428665|
    RECALL MISS: 4689096 States_and_territories_of_Australia]
4: 2604866 Mount_Hawthorn,_Western_Australia
    [winner: 19248371 Electoral_district_of_Mount_Hawthorn,
pem_pos=1, final_scores=0.08115743,
final_local_scores=0.08330053, pem=-0.1711949,
similarity=7.5065517|
    gt: 2604866 Mount_Hawthorn,_Western_Australia,
pem_pos=0, final_scores=0.054108337,
final_local_scores=0.07514843, pem=0.0, similarity=6.332292 ]
5: 1732143 Electoral_district_of_Leederville
    [winner: 1732152 Leederville,_Western_Australia,
pem_pos=0, final_scores=-0.07316853,
final_local_scores=-0.019161575, pem=0.0,
similarity=3.5178933|
    gt: 1732143 Electoral_district_of_Leederville,
pem_pos=1, final_scores=-0.1578252,
final_local_scores=-0.18267897, pem=-0.3233385,
similarity=0.39657348 ]
6: 22686206 First_Collier_Ministry
    [winner: 22686206 First_Collier_Ministry, pem_pos=0,
final_scores=-0.0442079, final_local_scores=0.037488475,
pem=0.0, similarity=5.2084436|
    gt: 22686206 First_Collier_Ministry, pem_pos=0,
final_scores=-0.0442079, final_local_scores=0.037488475,
pem=0.0, similarity=5.2084436 ]
7: 19515919 Western_Australian_state_election,_1947
    [winner: 19515919
Western_Australian_state_election,_1947, pem_pos=0,
final_scores=-0.7795357, final_local_scores=-0.8003025,
pem=0.0, similarity=-19.79291|
    gt: 19515919 Western_Australian_state_election,_1947,
pem_pos=0, final_scores=-0.7795357,
final_local_scores=-0.8003025, pem=0.0, similarity=-19.79291 ]
8: 19504551 Western_Australian_state_election,_1950

```

[winner: 19283873 1950, pem_pos=0,
final_scores=-0.8440544, final_local_scores=-0.71772057,
pem=-0.90795875, similarity=-12.39088|
RECALL MISS: 19504551
Western_Australian_state_election, 1950]
[winner: 1442451 Elections_in_Rwanda, pem_pos=2,
final_scores=-1.3579601, final_local_scores=-1.3035219,
pem=-3.50929, similarity=-15.725916|
RECALL MISS: 19504551
Western_Australian_state_election, 1950]
9: 12544844 North_West_Australia
[winner: 2855278 North-
West_(European_Parliament_constituency), pem_pos=0,
final_scores=0.021581456, final_local_scores=0.07960528,
pem=-0.6758157, similarity=10.140484|
gt: 12544844 North_West_Australia, pem_pos=12,
final_scores=-1.071506, final_local_scores=-1.1338127,
pem=-6.349139, similarity=4.7820888]
10: 4284858 Electoral_district_of_Pilbara
[winner: 491918 Pilbara, pem_pos=0,
final_scores=0.13029192, final_local_scores=0.16972229,
pem=0.0, similarity=9.154565|
gt: 4284858 Electoral_district_of_Pilbara, pem_pos=1,
final_scores=-0.021136463, final_local_scores=-0.046403408,
pem=-0.5026536, similarity=5.4384513]
11: 1732099 Electoral_district_of_Roebourne
[winner: 1452070 Roebourne, Western_Australia, pem_pos=0,
final_scores=0.17006908, final_local_scores=0.2170991,
pem=-0.09918329, similarity=11.107758|
gt: 1732099 Electoral_district_of_Roebourne, pem_pos=2,
final_scores=-0.079745546, final_local_scores=-0.07820167,
pem=-0.77086896, similarity=5.9481244]
12: 18799336 Western_Australian_state_election, 1983
[winner: 34755 1983, pem_pos=0, final_scores=-0.8038364,
final_local_scores=-0.6705099, pem=-0.6866735,
similarity=-12.185403|
RECALL MISS: 18799336
Western_Australian_state_election, 1983]
[winner: 9457 Election, pem_pos=0,
final_scores=-1.168341, final_local_scores=-1.2506118,
pem=-0.60297334, similarity=-29.951973|
RECALL MISS: 18799336
Western_Australian_state_election, 1983]
13: 19248949 Electoral_district_of_Subiaco
[winner: 676712 Subiaco, Western_Australia, pem_pos=2,
final_scores=0.06913675, final_local_scores=0.07211836,
pem=-1.1465836, similarity=12.477168|
gt: 19248949 Electoral_district_of_Subiaco, pem_pos=6,
final_scores=-0.2845928, final_local_scores=-0.31372148,
pem=-1.9343126, similarity=5.2467318]
14: 4291138 Electoral_district_of_Balcatta
[winner: 671523 Balcatta, Western_Australia, pem_pos=0,

final_scores=-0.047449574, final_local_scores=0.0010140985,
pem=0.0, similarity=4.1199756]
gt: 4291138 Electoral_district_of_Balcatta, pem_pos=1,
final_scores=-0.13765065, final_local_scores=-0.15680619,
pem=-0.008759185, similarity=-0.5420624]
15: 4292591 Electoral_district_of_Perth
[winner: 24355 Perth, pem_pos=0, final_scores=0.06271155,
final_local_scores=0.04414025, pem=-0.32334754,
similarity=7.1653605]
gt: 4292591 Electoral_district_of_Perth, pem_pos=8,
final_scores=-0.4736129, final_local_scores=-0.56691515,
pem=-3.715271, similarity=7.3760557]
1: 33613 Western_Australia
[pem_pos=0, final_scores=0.5068224,
final_local_scores=0.52519816, pem=-0.03175048,
similarity=19.935337]
2: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.2675817,
final_local_scores=0.23345396, pem=-0.2258703,
similarity=12.284762]
3: 18453 Liberal_Party_of_Australia
[pem_pos=1, final_scores=0.14125131,
final_local_scores=0.105492026, pem=-1.6500671,
similarity=16.211126]
4: 19246345 Electoral_district_of_Wembley_Beaches
[pem_pos=0, final_scores=0.08703779,
final_local_scores=0.10609305, pem=0.0, similarity=7.255741]
1: 14056 Prince-elector
[pem_pos=0, final_scores=0.052120253,
final_local_scores=0.13539362, pem=0.0, similarity=8.130129]
2: 19248371 Electoral_district_of_Mount_Hawthorn
[pem_pos=1, final_scores=0.052781507,
final_local_scores=0.052304164, pem=-0.1711949,
similarity=6.581558]
3: 19248371 Electoral_district_of_Mount_Hawthorn
[pem_pos=1, final_scores=0.08115743,
final_local_scores=0.08330053, pem=-0.1711949,
similarity=7.5065517]
4: 19248371 Electoral_district_of_Mount_Hawthorn
[pem_pos=1, final_scores=0.07718457,
final_local_scores=0.078960806, pem=-0.1711949,
similarity=7.3770456]
5: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.36031252,
final_local_scores=0.33474827, pem=-0.2258703,
similarity=15.307586]
6: 491918 Pilbara
[pem_pos=0, final_scores=0.13029192,
final_local_scores=0.16972229, pem=0.0, similarity=9.154565]
7: 1452070 Roebourne,_Western_Australia
[pem_pos=0, final_scores=0.17006908,
final_local_scores=0.2170991, pem=-0.09918329,


```
similarity=11.107758]
8: 676712 Subiaco, Western_Australia
   [pem_pos=2, final_scores=0.06913675,
   final_local_scores=0.07211836, pem=-1.1465836,
   similarity=12.477168]
9: 24355 Perth
   [pem_pos=0, final_scores=0.06271155,
   final_local_scores=0.04414025, pem=-0.32334754,
   similarity=7.1653605]
```

Sample: b'Electoral_district_of_Wembley_Beaches'

The Electoral district of [1 Wembley Beaches 1] was a 1 Legislative Assembly 2 electorate in the 3 state of [1 1 Western Australia 1]. The district was named for the inner western suburb of [2 4 Wembley 2], which fell within its borders. Wembley Beaches came into existence at the 5 1950 election, and was largely created from the northern and coastal sections of the [3 6 Leederville 3], which had grown considerably in population in the years immediately following [2 2 World War II 2]. The seat included the inner suburbs of [3 3 Floreat Park 3], 7 Jolimont and 8 Wembley, and areas further north including 9 Doubleview, [4 4 Gwelup 4], 10 Innaloo, 11 North Beach, [4 12 Scarborough 4], 13 Trigg and [5 5 Watermans Bay 5]. While reasonably safe for the [6 6 Liberal Party 6], [5 Wembley Beaches 5] was won for a term by the [7 7 Labor 7], who were relatively strong in the [6 Scarborough 6] and Innaloo region throughout the seat's existence. The seat was abolished ahead of the 14 1962 election, with its area being split between the seats of [7 15 Wembley 7] and 16 Karrinyup.

```
1: 33613 Western_Australia
2: 32927 World_War_II
3: 2590878 Floreat, Western_Australia
4: 671579 Gwelup, Western_Australia
5: 2219469 Watermans_Bay, Western_Australia
6: 18453 Liberal_Party_of_Australia
7: 1495 Australian_Labor_Party
1: 579448 Western_Australian_Legislative_Assembly
   [winner: 579433 New_South_Wales_Legislative_Assembly,
   pem_pos=1, final_scores=-0.36947727,
   final_local_scores=-0.41678616, pem=-1.9991245,
   similarity=2.5235324]
   gt: 579448 Western_Australian_Legislative_Assembly,
   pem_pos=3, final_scores=-0.47411817,
   final_local_scores=-0.530041, pem=-2.598477,
   similarity=2.4031541 ]
2: 1726491 Electoral_districts_of_Western_Australia
   [winner: 3236919 New_Zealand_electorates, pem_pos=0,
   final_scores=-0.91467494, final_local_scores=-0.9536116,
   pem=-0.60937357, similarity=-21.05409]
   gt: 1726491 Electoral_districts_of_Western_Australia,
   pem_pos=2, final_scores=-1.2002015,
```

final_local_scores=-1.2574089, pem=-1.646076,
 similarity=-24.482258]
 3: 4689096 States_and_territories_of_Australia
 [winner: 445321 Political_divisions_of_the_United_States,
 pem_pos=0, final_scores=-1.4906304,
 final_local_scores=-1.5311127, pem=-0.5713872,
 similarity=-38.494453]
 RECALL MISS: 4689096 States_and_territories_of_Australia]
 4: 2590879 Wembley,_Western_Australia
 [winner: 94377 Wembley, pem_pos=1,
 final_scores=0.13595052, final_local_scores=0.12277472,
 pem=-1.0958498, similarity=13.712957]
 gt: 2590879 Wembley,_Western_Australia, pem_pos=4,
 final_scores=-0.23251207, final_local_scores=-0.22508991,
 pem=-2.460814, similarity=10.75487]
 5: 19504551 Western_Australian_state_election,_1950
 [winner: 19283873 1950, pem_pos=0,
 final_scores=-0.86828715, final_local_scores=-0.7623086,
 pem=-0.90795875, similarity=-13.721475]
 RECALL MISS: 19504551
 Western_Australian_state_election,_1950]
 [winner: 9457 Election, pem_pos=0,
 final_scores=-1.256715, final_local_scores=-1.3296411,
 pem=-0.60297334, similarity=-32.310368]
 RECALL MISS: 19504551
 Western_Australian_state_election,_1950]
 6: 1732143 Electoral_district_of_Leederville
 [winner: 1732152 Leederville,_Western_Australia,
 pem_pos=0, final_scores=0.1304725,
 final_local_scores=0.19688994, pem=0.0, similarity=9.9653015]
 gt: 1732143 Electoral_district_of_Leederville,
 pem_pos=1, final_scores=-0.13716024,
 final_local_scores=-0.11583833, pem=-0.3233385,
 similarity=2.3912313]
 7: 2646827 Jolimont,_Western_Australia
 [winner: 1467725 East_Melbourne,_Victoria, pem_pos=3,
 final_scores=-0.07727249, final_local_scores=-0.10716602,
 pem=-1.6341306, similarity=9.778321]
 gt: 2646827 Jolimont,_Western_Australia, pem_pos=1,
 final_scores=-0.18368886, final_local_scores=-0.1423102,
 pem=-0.5013852, similarity=2.5695038]
 8: 2590879 Wembley,_Western_Australia
 [winner: 94377 Wembley, pem_pos=1,
 final_scores=0.019616798, final_local_scores=-0.026871637,
 pem=-1.0958498, similarity=9.247211]
 gt: 2590879 Wembley,_Western_Australia, pem_pos=4,
 final_scores=-0.3617647, final_local_scores=-0.36071807,
 pem=-2.460814, similarity=6.7074556]
 9: 671518 Doubleview,_Western_Australia
 [winner: 671518 Doubleview,_Western_Australia, pem_pos=0,
 final_scores=-0.06956667, final_local_scores=-0.03006149,
 pem=0.0, similarity=3.1926181]

gt: 671518 Doubleview, Western_Australia, pem_pos=0,
final_scores=-0.06956667, final_local_scores=-0.03006149,
pem=0.0, similarity=3.1926181]
10: 671601 Innaloo, Western_Australia
[winner: 671601 Innaloo, Western_Australia, pem_pos=0,
final_scores=0.008831635, final_local_scores=0.023063645,
pem=0.0, similarity=4.777978]
gt: 671601 Innaloo, Western_Australia, pem_pos=0,
final_scores=0.008831635, final_local_scores=0.023063645,
pem=0.0, similarity=4.777978]
11: 672286 North_Beach, Western_Australia
[winner: 672286 North_Beach, Western_Australia,
pem_pos=1, final_scores=-0.041798383,
final_local_scores=-0.043881625, pem=-1.6962514,
similarity=12.004677]
gt: 672286 North_Beach, Western_Australia, pem_pos=1,
final_scores=-0.041798383, final_local_scores=-0.043881625,
pem=-1.6962514, similarity=12.004677]
12: 672298 Scarborough, Western_Australia
[winner: 381565 Scarborough, North_Yorkshire, pem_pos=0,
final_scores=0.096330866, final_local_scores=0.06886648,
pem=-0.9726719, similarity=11.434367]
gt: 672298 Scarborough, Western_Australia, pem_pos=7,
final_scores=-0.30083507, final_local_scores=-0.32597685,
pem=-3.4132688, similarity=12.9238]
13: 672310 Trigg, Western_Australia
[winner: 672310 Trigg, Western_Australia, pem_pos=1,
final_scores=0.002014324, final_local_scores=0.07231817,
pem=-1.0366617, similarity=11.885359]
gt: 672310 Trigg, Western_Australia, pem_pos=1,
final_scores=0.002014324, final_local_scores=0.07231817,
pem=-1.0366617, similarity=11.885359]
14: 19504023 Western_Australian_state_election, 1962
[winner: 34717 1962, pem_pos=0, final_scores=-0.5272548,
final_local_scores=-0.34551784, pem=-0.7481338,
similarity=-2.1527605]
RECALL MISS: 19504023
Western_Australian_state_election, 1962]
[winner: 9457 Election, pem_pos=0,
final_scores=-1.0542762, final_local_scores=-1.1085074,
pem=-0.60297334, similarity=-25.711294]
RECALL MISS: 19504023
Western_Australian_state_election, 1962]
15: 19244792 Electoral_district_of_Wembley
[winner: 143673 Wembley_Stadium_(1923), pem_pos=0,
final_scores=0.08735989, final_local_scores=0.10280597,
pem=-0.8375099, similarity=11.712157]
gt: 19244792 Electoral_district_of_Wembley, pem_pos=10,
final_scores=-0.62165207, final_local_scores=-0.6317835,
pem=-2.6790187, similarity=-0.19504617]
16: 19279300 Electoral_district_of_Karrinyup
[winner: 672245 Karrinyup, Western_Australia, pem_pos=0,

```

final_scores=0.019155547, final_local_scores=0.045212284,
pem=0.0, similarity=5.4389377|
    gt: 19279300 Electoral_district_of_Karrinyup, pem_pos=1,
final_scores=-0.25682664, final_local_scores=-0.19342107,
pem=-0.4691671, similarity=0.869046 ]
1: 33613 Western_Australia
    [pem_pos=0, final_scores=0.4895926,
final_local_scores=0.5109265, pem=-0.03175048,
similarity=19.509443]
2: 32927 World_War_II
    [pem_pos=0, final_scores=0.40951347,
final_local_scores=0.4601755, pem=0.0, similarity=17.822268]
3: 2590878 Floreat,_Western_Australia
    [pem_pos=0, final_scores=0.2161005,
final_local_scores=0.24066687, pem=0.0, similarity=11.271692]
4: 671579 Gwelup,_Western_Australia
    [pem_pos=0, final_scores=0.12090032,
final_local_scores=0.15583682, pem=0.0, similarity=8.740194]
5: 2219469 Watermans_Bay,_Western_Australia
    [pem_pos=0, final_scores=0.13171117,
final_local_scores=0.17025468, pem=0.0, similarity=9.170452]
6: 18453 Liberal_Party_of_Australia
    [pem_pos=1, final_scores=0.13405837,
final_local_scores=0.13991144, pem=-1.6500671,
similarity=17.238272]
7: 1495 Australian_Labor_Party
    [pem_pos=0, final_scores=0.27095526,
final_local_scores=0.28997073, pem=-0.2258703,
similarity=13.971335]
1: 19246345 Electoral_district_of_Wembley_Beaches
    [pem_pos=0, final_scores=0.092967734,
final_local_scores=0.12144089, pem=0.0, similarity=7.7137513]
2: 94377 Wembley
    [pem_pos=1, final_scores=0.13595052,
final_local_scores=0.12277472, pem=-1.0958498,
similarity=13.712957]
3: 1732152 Leederville,_Western_Australia
    [pem_pos=0, final_scores=0.1304725,
final_local_scores=0.19688994, pem=0.0, similarity=9.9653015]
4: 381565 Scarborough,_North_Yorkshire
    [pem_pos=0, final_scores=0.096330866,
final_local_scores=0.06886648, pem=-0.9726719,
similarity=11.434367]
5: 19246345 Electoral_district_of_Wembley_Beaches
    [pem_pos=0, final_scores=0.043935135,
final_local_scores=0.06788023, pem=0.0, similarity=6.115394]
6: 381565 Scarborough,_North_Yorkshire
    [pem_pos=0, final_scores=0.17645828,
final_local_scores=0.15639353, pem=-0.9726719,
similarity=14.04635]
7: 143673 Wembley_Stadium_(1923)
    [pem_pos=0, final_scores=0.08735989,

```

final_local_scores=0.10280597, pem=-0.8375099,
similarity=11.712157]

Sample: b'Larry_Worrell'

Lawrence ` Larry ' Roosevelt Worrell -LRB- born 28 August 1943 in [1 1 Saint Thomas 1] , [1 1 Barbados 1] -RRB- is a [2 2 Barbadian 2] born former [3 3 English 3] . [4 Worrell 4] was a right-handed who bowled right-arm 4 off break . [5 Worrell 5] made his 5 first-class debut for [2 2 Hampshire 2] , playing a single match in 1969 against the touring [3 3 New Zealanders 3] . During the 1969 season [6 Worrell 6] also represented [4 4 Dorset 4] in the , playing against [5 5 Berkshire 5] , [6 6 Cornwall 6] and [7 7 Somerset Second XI 7] . In 1971 [7 Worrell 7] returned to [8 Hampshire 8] for his second spell at the club . [9 Worrell 9] made his [8 8 County Championship 8] debut against [9 9 Nottinghamshire 9] . [10 Worrell 10] played 32 first-class matches for [11 Hampshire 11] , the last of which came against [10 10 Leicestershire 10] in 1972 . [12 Worrell 12] scored 289 runs at an of 11.56 . [13 Worrell 13] scored a single half-century in his career , making exactly 50 runs . This was his highest first-class score . An off break bowler , [14 Worrell 14] took 65 wickets at an 6 average of 32.55 , with best figures of 5-67 . In 1972 [15 Worrell 15] made his only career apearance against [11 11 Lancashire 11]

1: 3455 Barbados

2: 1919370 Hampshire_County_Cricket_Club

3: 478115 New_Zealand_national_cricket_team

4: 3712026 Dorset_County_Cricket_Club

5: 3702129 Berkshire_County_Cricket_Club

6: 3702042 Cornwall_County_Cricket_Club

7: 1622178 Somerset_County_Cricket_Club

8: 683599 County_Championship

9: 667341 Nottinghamshire_County_Cricket_Club

10: 1622318 Leicestershire_County_Cricket_Club

11: 749919 Lancashire_County_Cricket_Club

1: 1528036 Saint_Thomas,_Barbados

[winner: 441496 Saint_Thomas,_U.S._Virgin_Islands,

pem_pos=0, final_scores=0.16474216,

final_local_scores=0.26016185, pem=-0.419408,

similarity=14.134267]

gt: 1528036 Saint_Thomas,_Barbados, pem_pos=3,

final_scores=-0.2902698, final_local_scores=-0.22179833,

pem=-2.0187566, similarity=8.4491205]

2: 6960207 Barbadian_people

[winner: 3455 Barbados, pem_pos=0,

final_scores=0.1749276, final_local_scores=0.20559764,

pem=-0.4921499, similarity=12.901545]

gt: 6960207 Barbadian_people, pem_pos=1,

final_scores=-0.070479214, final_local_scores=-0.033942416,

pem=-0.90654206, similarity=8.00672]

3: 18803164 English_people

[winner: 9316 England, pem_pos=1,
final_scores=0.12913458, final_local_scores=0.14825925,
pem=-1.3527626, similarity=15.870598]
gt: 18803164 English_people, pem_pos=2,
final_scores=-0.34531116, final_local_scores=-0.36807844,
pem=-2.2084396, similarity=5.115356]
4: 711518 Off_break
[winner: 22499 Off_spin, pem_pos=0,
final_scores=-1.112432, final_local_scores=-1.2416017,
pem=0.0, similarity=-32.96216]
RECALL MISS: 711518 Off_break]
5: 11040 First-class_cricket
[winner: 662844 First-class_citizen, pem_pos=3,
final_scores=-1.2636801, final_local_scores=-1.1735387,
pem=-1.5935546, similarity=-22.265026]
gt: 11040 First-class_cricket, pem_pos=0,
final_scores=-1.4079711, final_local_scores=-1.5796523,
pem=-0.053819276, similarity=-42.757584]
6: 875871 Bowling_average
[winner: 875871 Bowling_average, pem_pos=0,
final_scores=-0.94364613, final_local_scores=-1.0514276,
pem=0.0, similarity=-27.286985]
gt: 875871 Bowling_average, pem_pos=0,
final_scores=-0.94364613, final_local_scores=-1.0514276,
pem=0.0, similarity=-27.286985]
1: 3455 Barbados
[pem_pos=0, final_scores=0.21822707,
final_local_scores=0.2528957, pem=-0.3637528,
similarity=13.61477]
2: 1919370 Hampshire_County_Cricket_Club
[pem_pos=1, final_scores=0.5781464,
final_local_scores=0.56505054, pem=-1.9232318,
similarity=31.41078]
3: 478115 New_Zealand_national_cricket_team
[pem_pos=2, final_scores=0.31482977,
final_local_scores=0.29957604, pem=-1.155384,
similarity=19.312817]
4: 3712026 Dorset_County_Cricket_Club
[pem_pos=1, final_scores=0.35530442,
final_local_scores=0.3580382, pem=-2.4537704,
similarity=28.11827]
5: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.55532795,
final_local_scores=0.572414, pem=-2.301417,
similarity=33.687153]
6: 3702042 Cornwall_County_Cricket_Club
[pem_pos=2, final_scores=0.27912968,
final_local_scores=0.28623927, pem=-3.34357,
similarity=30.814516]
7: 1622178 Somerset_County_Cricket_Club
[pem_pos=0, final_scores=0.5497446,
final_local_scores=0.5319583, pem=0.0, similarity=19.964409]

8: 683599 County_Championship
[pem_pos=0, final_scores=0.25089222,
final_local_scores=0.21770129, pem=0.0, similarity=10.586353]
9: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.56239045,
final_local_scores=0.54884154, pem=-1.1067253,
similarity=26.486784]
10: 1622318 Leicestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.6944286,
final_local_scores=0.7032519, pem=-1.0465667,
similarity=30.767544]
11: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.36448067,
final_local_scores=0.3332194, pem=-1.7295884,
similarity=23.439415]
1: 441496 Saint_Thomas,_U.S._Virgin_Islands
[pem_pos=0, final_scores=0.16474216,
final_local_scores=0.26016185, pem=-0.419408,
similarity=14.134267]
2: 3455 Barbados
[pem_pos=0, final_scores=0.1749276,
final_local_scores=0.20559764, pem=-0.4921499,
similarity=12.901545]
3: 9316 England
[pem_pos=1, final_scores=0.12913458,
final_local_scores=0.14825925, pem=-1.3527626,
similarity=15.870598]
4: 897236 Frank_Worrell
[pem_pos=8, final_scores=0.06079133,
final_local_scores=0.012411654, pem=-3.0445223,
similarity=21.016687]
5: 897236 Frank_Worrell
[pem_pos=8, final_scores=0.2623009,
final_local_scores=0.23253024, pem=-3.0445223,
similarity=27.585464]
6: 897236 Frank_Worrell
[pem_pos=8, final_scores=0.27447134,
final_local_scores=0.24452901, pem=-3.0445223,
similarity=27.943531]
7: 897236 Frank_Worrell
[pem_pos=8, final_scores=0.19951184,
final_local_scores=0.16394275, pem=-3.0445223,
similarity=25.538677]
8: 1919370 Hampshire_County_Cricket_Club
[pem_pos=1, final_scores=0.26007843,
final_local_scores=0.21854928, pem=-1.9232318,
similarity=21.070494]
9: 897236 Frank_Worrell
[pem_pos=8, final_scores=0.3692121,
final_local_scores=0.34739596, pem=-3.0445223,
similarity=31.013287]
10: 897236 Frank_Worrell

```

    [pem_pos=8, final_scores=0.2577142,
final_local_scores=0.22622436, pem=-3.0445223,
similarity=27.397284]
11: 1919370 Hampshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.66946095,
final_local_scores=0.66479784, pem=-1.9232318,
similarity=34.38744]
12: 897236 Frank_Worrell
    [pem_pos=8, final_scores=0.2974112,
final_local_scores=0.26958728, pem=-3.0445223,
similarity=28.69132]
13: 897236 Frank_Worrell
    [pem_pos=8, final_scores=0.3016935,
final_local_scores=0.27426505, pem=-3.0445223,
similarity=28.830914]
14: 897236 Frank_Worrell
    [pem_pos=8, final_scores=0.33411622,
final_local_scores=0.30968195, pem=-3.0445223,
similarity=29.887825]
15: 897236 Frank_Worrell
    [pem_pos=8, final_scores=0.05895005,
final_local_scores=0.010400355, pem=-3.0445223,
similarity=20.956665]

```

Sample: b'Interstate_691'

Interstate 691 -LRB- abbreviated I-691 -RRB- is a portion of the [1 1 Interstate Highway System 1] in [2 2 Connecticut 2] beginning at 1 Interstate 91 in [3 3 Meriden 3] and ending at 2 Interstate 84 near the [1 3 Cheshire 1] - [4 4 Southington 4] town line . It is in length , including of the exit ramp to the merge with westbound 4 I-84 . I-691 is also known as the Henry D. Altobello Highway for its entire length . I-691 is the main east -- west highway of the city of [5 5 Meriden 5] . The freeway actually begins in Middlefield as [2 5 Route 66 2] , technically becoming I-691 at the junction with 6 I-91 -LRB- Exit 11 -RRB- . However , westbound signage indicates I-691 begins at the start of the freeway -LRB- just west of Exit 13 -RRB- , while eastbound signage shows I-691 ending at the 7 Route 15 interchange -LRB- at eastbound Exit 10 about west of the interchange with 8 I-91 -RRB- . To go from 9 I-91 northbound to I-691 westbound -LRB- or from I-691 eastbound to 10 I-91 southbound -RRB- , one must actually use 11 Route 1 5 . In the 1940s , the I-691 routing was part of a planned US 6A Expressway from [3 Southington 3] to [4 Willimantic 4] . A section of the expressway -LRB- from its Middlefield terminus west to Exit 8 -RRB- first opened in 1966 . By 1968 , the US 6A designation was dropped in favor of [5 12 Route 66 5] . The highway was extended west to Exit 4 by 1971 . The connection to 13 I-84 was eventually completed in 1987 , with the renumbering to I-691 done at the same time . The portion east of 14 I-91 remained as 15 Route 66 . Environmental and community groups successfully blocked attempts to extend the

freeway east of its present terminus due to potential impacts on a reservoir that provides drinking water for the local area . A compromise was reached in the late 1990s allowing [6 6 CONNDOT 6] to convert 16 Route 66 from a 2-lane road to a 4-lane divided highway from the eastern end of the I-691 freeway to 17 Route 9 in [7 7 Middletown 7] . Construction on the 18 Route 66

1: 43950 Interstate_Highway_System

2: 6466 Connecticut

3: 256957 Meriden,_Connecticut

4: 108731 Southington,_Connecticut

5: 256957 Meriden,_Connecticut

6: 6049641 Connecticut_Department_of_Transportation

7: 275138 Middletown,_Connecticut

1: 26216916 Interstate_91_in_Connecticut

[winner: 92121 Interstate_91, pem_pos=0, final_scores=-0.06937465, final_local_scores=-0.11699155, pem=0.0, similarity=0.5984516|

gt: 26216916 Interstate_91_in_Connecticut, pem_pos=1, final_scores=-0.9204367, final_local_scores=-1.0153718, pem=-5.1752076, similarity=1.9325786]

2: 27555853 Interstate_84_in_Connecticut

[winner: 43950 Interstate_Highway_System, pem_pos=0, final_scores=-0.104521036, final_local_scores=-0.14647385, pem=-0.8229691, similarity=4.1940746|

RECALL MISS: 27555853 Interstate_84_in_Connecticut]

[winner: 13119357 Jalan_Kampung_Raja, pem_pos=4, final_scores=-1.0318645, final_local_scores=-0.9369792, pem=-2.986353, similarity=-7.6313624|

RECALL MISS: 27555853 Interstate_84_in_Connecticut]

3: 108793 Cheshire,_Connecticut

[winner: 7406 Cheshire, pem_pos=0, final_scores=0.24195157, final_local_scores=0.2578718, pem=-0.29490227, similarity=13.388846|

gt: 108793 Cheshire,_Connecticut, pem_pos=1, final_scores=-0.08353707, final_local_scores=-0.13906273, pem=-2.4974012, similarity=13.52106]

4: 27555853 Interstate_84_in_Connecticut

[winner: 1519632 Interstate_84_(Pennsylvania-Massachusetts), pem_pos=0, final_scores=-0.22239538, final_local_scores=-0.28108436, pem=-0.59677744, similarity=-1.0530366|

gt: 27555853 Interstate_84_in_Connecticut, pem_pos=6, final_scores=-0.459561, final_local_scores=-0.5233995, pem=-1.9962234, similarity=-0.673797]

5: 3353377 Connecticut_Route_66

[winner: 79717 U.S._Route_66, pem_pos=0, final_scores=0.059119806, final_local_scores=0.03163749, pem=-0.5774016, similarity=8.173838|

gt: 3353377 Connecticut_Route_66, pem_pos=3, final_scores=-0.4323163, final_local_scores=-0.45719385, pem=-2.633317, similarity=4.7665234]

6: 26216916 Interstate_91_in_Connecticut
[winner: 92121 Interstate_91, pem_pos=0,
final_scores=-0.5820185, final_local_scores=-0.6769771,
pem=0.0, similarity=-16.112635]
gt: 26216916 Interstate_91_in_Connecticut, pem_pos=1,
final_scores=-1.1038346, final_local_scores=-1.2157061,
pem=-3.302166, similarity=-14.231692]

7: 1339510 Connecticut_Route_15
[winner: 1339510 Connecticut_Route_15, pem_pos=0,
final_scores=-0.17140669, final_local_scores=-0.1921786,
pem=-0.72608405, similarity=2.3032777]
gt: 1339510 Connecticut_Route_15, pem_pos=0,
final_scores=-0.17140669, final_local_scores=-0.1921786,
pem=-0.72608405, similarity=2.3032777]

8: 26216916 Interstate_91_in_Connecticut
[winner: 92121 Interstate_91, pem_pos=0,
final_scores=-0.5217856, final_local_scores=-0.61118186,
pem=0.0, similarity=-14.149174]
gt: 26216916 Interstate_91_in_Connecticut, pem_pos=1,
final_scores=-1.060357, final_local_scores=-1.1682134,
pem=-3.302166, similarity=-12.81441]

9: 26216916 Interstate_91_in_Connecticut
[winner: 92121 Interstate_91, pem_pos=0,
final_scores=-0.18642421, final_local_scores=-0.24485041,
pem=0.0, similarity=-3.2171118]
gt: 26216916 Interstate_91_in_Connecticut, pem_pos=1,
final_scores=-0.748038, final_local_scores=-0.8270524,
pem=-3.302166, similarity=-2.6334887]

10: 26216916 Interstate_91_in_Connecticut
[winner: 92121 Interstate_91, pem_pos=0,
final_scores=-0.14440219, final_local_scores=-0.19894774,
pem=0.0, similarity=-1.8472844]
gt: 26216916 Interstate_91_in_Connecticut, pem_pos=1,
final_scores=-0.7000506, final_local_scores=-0.77463347,
pem=-3.302166, similarity=-1.0692022]

11: 1339510 Connecticut_Route_15
[winner: 1339510 Connecticut_Route_15, pem_pos=0,
final_scores=-0.41776633, final_local_scores=-0.4612891,
pem=-0.72608405, similarity=-5.7275167]
gt: 1339510 Connecticut_Route_15, pem_pos=0,
final_scores=-0.41776633, final_local_scores=-0.4612891,
pem=-0.72608405, similarity=-5.7275167]

12: 3353377 Connecticut_Route_66
[winner: 79717 U.S._Route_66, pem_pos=0,
final_scores=0.034651563, final_local_scores=0.0049096346,
pem=-0.5774016, similarity=7.3762255]
gt: 3353377 Connecticut_Route_66, pem_pos=3,
final_scores=-0.5375615, final_local_scores=-0.5721582,
pem=-2.633317, similarity=1.3357586]

13: 27555853 Interstate_84_in_Connecticut
[winner: 1519632 Interstate_84_(Pennsylvania-Massachusetts), pem_pos=0, final_scores=-0.2700228,

```

final_local_scores=-0.3331101, pem=-0.59677744,
similarity=-2.605588|
    gt: 27555853 Interstate_84_in_Connecticut, pem_pos=6,
final_scores=-0.53001213, final_local_scores=-0.6003567,
pem=-1.9962234, similarity=-2.970355 ]
14: 26216916 Interstate_91_in_Connecticut
    [winner: 92121 Interstate_91, pem_pos=0,
final_scores=-0.2534868, final_local_scores=-0.3181061,
pem=0.0, similarity=-5.403208|
    gt: 26216916 Interstate_91_in_Connecticut, pem_pos=1,
final_scores=-0.8469592, final_local_scores=-0.9351088,
pem=-3.302166, similarity=-5.858105 ]
15: 3353377 Connecticut_Route_66
    [winner: 79717 U.S._Route_66, pem_pos=0,
final_scores=-0.036104202, final_local_scores=-0.08790048,
pem=-0.5774016, similarity=4.606587|
    gt: 3353377 Connecticut_Route_66, pem_pos=3,
final_scores=-0.49067014, final_local_scores=-0.5205146,
pem=-2.633317, similarity=2.876906 ]
16: 3353377 Connecticut_Route_66
    [winner: 79717 U.S._Route_66, pem_pos=0,
final_scores=-0.026500687, final_local_scores=-0.07741011,
pem=-0.5774016, similarity=4.9196405|
    gt: 3353377 Connecticut_Route_66, pem_pos=3,
final_scores=-0.554559, final_local_scores=-0.59030354,
pem=-2.633317, similarity=0.79426503 ]
17: 3339405 Connecticut_Route_9
    [winner: 849593 Massachusetts_Route_9, pem_pos=0,
final_scores=-0.104369506, final_local_scores=-0.14522967,
pem=-0.69708306, similarity=3.5466156|
    gt: 3339405 Connecticut_Route_9, pem_pos=1,
final_scores=-0.23488493, final_local_scores=-0.27134305,
pem=-1.5314926, similarity=4.320791 ]
18: 3353377 Connecticut_Route_66
    [winner: 79717 U.S._Route_66, pem_pos=0,
final_scores=-0.0273128, final_local_scores=-0.07829722,
pem=-0.5774016, similarity=4.893167|
    gt: 3353377 Connecticut_Route_66, pem_pos=3,
final_scores=-0.49286413, final_local_scores=-0.52291125,
pem=-2.633317, similarity=2.805385 ]
1: 43950 Interstate_Highway_System
    [pem_pos=0, final_scores=0.19865547,
final_local_scores=0.20283079, pem=0.0, similarity=10.142589]
2: 6466 Connecticut
    [pem_pos=0, final_scores=0.07138328,
final_local_scores=0.014830828, pem=-0.24788867,
similarity=5.8803515]
3: 256957 Meriden, Connecticut
    [pem_pos=0, final_scores=0.3666075,
final_local_scores=0.32509923, pem=-0.6049158,
similarity=17.080948]
4: 108731 Southington, Connecticut

```

[pem_pos=0, final_scores=0.23989706,
 final_local_scores=0.19170696, pem=-0.10104706,
 similarity=10.36014]
 5: 256957 Meriden, Connecticut
 [pem_pos=0, final_scores=0.31769377,
 final_local_scores=0.2716685, pem=-0.6049158,
 similarity=15.486469]
 6: 6049641 Connecticut Department of Transportation
 [pem_pos=0, final_scores=0.171135,
 final_local_scores=0.18965918, pem=0.0, similarity=9.749521]
 7: 275138 Middletown, Connecticut
 [pem_pos=0, final_scores=0.03156914,
 final_local_scores=-0.025361806, pem=-1.6361735,
 similarity=12.230632]
 1: 7406 Cheshire
 [pem_pos=0, final_scores=0.24195157,
 final_local_scores=0.2578718, pem=-0.29490227,
 similarity=13.388846]
 2: 79717 U.S. Route 66
 [pem_pos=0, final_scores=0.059119806,
 final_local_scores=0.03163749, pem=-0.5774016,
 similarity=8.173838]
 3: 108731 Southington, Connecticut
 [pem_pos=0, final_scores=0.36482894,
 final_local_scores=0.32817602, pem=-0.10104706,
 similarity=14.43265]
 4: 108880 Willimantic, Connecticut
 [pem_pos=0, final_scores=0.37954193,
 final_local_scores=0.3706501, pem=0.0, similarity=15.150652]
 5: 79717 U.S. Route 66
 [pem_pos=0, final_scores=0.034651563,
 final_local_scores=0.0049096346, pem=-0.5774016,
 similarity=7.3762255]

Sample: b'Antonio_Rossi'

[1 Antonio Rossi 1] -LRB- born December 19 , 1968 in [1 1 Le
 cco 1] -RRB- is an [2 2 Italian 2] 1 sprint canoer who has
 competed since the early 1990s . Competing in five , he won
 five medals which included three golds -LRB- K-1 500 m : 2 199
 6 , K-2 1000 m : 1996 , 3 2000 -RRB- , one silver -LRB- K-2
 1000 m : 4 2004 -RRB- , and one bronze -LRB- K-2 500 m : 5 199
 2 -RRB- . [2 Rossi 2] has also been successful at the [3 3 ICF
 Canoe Sprint World Championships 3] with seven medals . This
 includes three golds -LRB- K-2 1000 m : 6 1995 , 7 1997 , 8 19
 98 -RRB- , three silvers -LRB- K-2 1000 m : 9 1993 , 10
 1994 ; K-4 200 m : 1998 -RRB- , and one bronze -LRB- K-1 500
 m : 1997 -RRB- . He was the [3 Italian 3] flag bearer at the
 opening ceremony of the 11 2008 [4 Summer Olympics 4] in [4 4
 Beijing 4] . [5 Rossi 5] 's wife , [5 5 Lucia Micheli 5] ,
 competed in the 12 K-4 500 m event at the 13 1992 [6 Summer Ol
 ympics 6] in [6 6 Barcelona 6] . His clus is 14 G.S. Fiamme Gi
 alle

1: 249828 Lecco
2: 14532 Italy
3: 5192170 ICF_Canoe_Sprint_World_Championships
4: 18603746 Beijing
5: 24663466 Lucia_Micheli
6: 4443 Barcelona
1: 1900663 Canoe_racing
[winner: 3485435 Sprint_(missile), pem_pos=1,
final_scores=-0.92919135, final_local_scores=-0.8435664,
pem=-1.6455486, similarity=-12.135242]
RECALL MISS: 1900663 Canoe_racing]
[winner: 1900663 Canoe_racing, pem_pos=0,
final_scores=-0.9927998, final_local_scores=-1.0579875,
pem=0.0, similarity=-27.482742]
gt: 1900663 Canoe_racing, pem_pos=0,
final_scores=-0.9927998, final_local_scores=-1.0579875,
pem=0.0, similarity=-27.482742]
2: 72311 1996_Summer_Olympics
[winner: 34636 1996, pem_pos=0, final_scores=-0.6754549,
final_local_scores=-0.4757415, pem=-0.78063476,
similarity=-5.8621492]
gt: 72311 1996_Summer_Olympics, pem_pos=9,
final_scores=-1.073277, final_local_scores=-1.2046692,
pem=-4.291002, similarity=-8.524886]
3: 71432 2000_Summer_Olympics
[winner: 34548 2000, pem_pos=0, final_scores=-0.78352475,
final_local_scores=-0.72476524, pem=-0.8862115,
similarity=-12.719372]
gt: 71432 2000_Summer_Olympics, pem_pos=12,
final_scores=-1.0495039, final_local_scores=-1.1687205,
pem=-4.788432, similarity=-4.746997]
4: 77741 2004_Summer_Olympics
[winner: 35524 2004, pem_pos=0, final_scores=-0.8217242,
final_local_scores=-0.68987995, pem=-0.4915334,
similarity=-13.824646]
gt: 77741 2004_Summer_Olympics, pem_pos=11,
final_scores=-1.7592685, final_local_scores=-1.9495835,
pem=-4.876049, similarity=-27.573034]
5: 82755 1992_Summer_Olympics
[winner: 34669 1992, pem_pos=0, final_scores=-0.9017809,
final_local_scores=-0.7855763, pem=-0.76294684,
similarity=-15.204426]
gt: 82755 1992_Summer_Olympics, pem_pos=7,
final_scores=-1.5974128, final_local_scores=-1.7761385,
pem=-4.0996614, similarity=-26.619205]
6: 18802185 1995_ICF_Canoe_Sprint_World_Championships
[winner: 34658 1995, pem_pos=0, final_scores=-1.010925,
final_local_scores=-0.96604675, pem=-0.68201965,
similarity=-21.030119]
RECALL MISS: 18802185
1995_ICF_Canoe_Sprint_World_Championships]
7: 18810455 1997_ICF_Canoe_Sprint_World_Championships

[winner: 34601 1997, pem_pos=0, final_scores=-0.9971357,
final_local_scores=-0.9315335, pem=-0.69383013,
similarity=-19.93595]

RECALL MISS: 18810455

1997_ICF_Canoe_Sprint_World_Championships]

8: 18826631 1998_ICF_Canoe_Sprint_World_Championships

[winner: 34647 1998, pem_pos=0, final_scores=-1.5016935,
final_local_scores=-1.5401425, pem=-0.6850515,
similarity=-38.145794]

RECALL MISS: 18826631

1998_ICF_Canoe_Sprint_World_Championships]

9: 18788871 1993_ICF_Canoe_Sprint_World_Championships

[winner: 34598 1993, pem_pos=0, final_scores=-0.6993738,
final_local_scores=-0.53384995, pem=-0.7063994,
similarity=-7.9999256]

RECALL MISS: 18788871

1993_ICF_Canoe_Sprint_World_Championships]

10: 18800905 1994_ICF_Canoe_Sprint_World_Championships

[winner: 18948337 1994, pem_pos=0,
final_scores=-0.87257135, final_local_scores=-0.8131324,
pem=-0.7539553, similarity=-16.075655]

RECALL MISS: 18800905

1994_ICF_Canoe_Sprint_World_Championships]

11: 77745 2008_Summer_Olympics

[winner: 27566 Summer_Olympic_Games, pem_pos=0,
final_scores=0.17073114, final_local_scores=0.12588656,
pem=-0.4112264, similarity=10.082731]

gt: 77745 2008_Summer_Olympics, pem_pos=2,
final_scores=-0.49783188, final_local_scores=-0.5618603,
pem=-3.4497817, similarity=6.0831304]

[winner: 35825 2008, pem_pos=0, final_scores=-1.0334609,
final_local_scores=-0.95612854, pem=-0.29331458,
similarity=-22.84798]

gt: 77745 2008_Summer_Olympics, pem_pos=12,
final_scores=-2.1839776, final_local_scores=-2.4121358,
pem=-5.467841, similarity=-38.15826]

12: 19834529 Canoeing_at_the_1992_Summer_Olympics_-
_Women's_K-4_500_metres

[winner: 6805315 K-4_(Kansas_highway), pem_pos=0,
final_scores=-0.20662948, final_local_scores=-0.046656087,
pem=0.0, similarity=2.6974022]

RECALL MISS: 19834529

Canoeing_at_the_1992_Summer_Olympics_-
_Women's_K-4_500_metres]

[winner: 89276 Sprint_(running), pem_pos=0,
final_scores=-0.6449482, final_local_scores=-0.71758693,
pem=-1.5988557, similarity=-8.629687]

RECALL MISS: 19834529

Canoeing_at_the_1992_Summer_Olympics_-Women's_K-4_500_metres]

13: 82755 1992_Summer_Olympics

[winner: 27566 Summer_Olympic_Games, pem_pos=0,
final_scores=0.11050181, final_local_scores=0.060095176,

```

pem=-0.4112264, similarity=8.119386|
    gt: 82755 1992_Summer_Olympics, pem_pos=17,
final_scores=-0.6523152, final_local_scores=-0.73467654,
pem=-4.174166, similarity=4.865265 ]
    [winner: 34669 1992, pem_pos=0, final_scores=-0.7963847,
final_local_scores=-0.67044693, pem=-0.76294684,
similarity=-11.768738|
    gt: 82755 1992_Summer_Olympics, pem_pos=7,
final_scores=-1.3779072, final_local_scores=-1.5363619,
pem=-4.0996614, similarity=-19.463793 ]
14: 35670529 Gruppi_Sportivi_Fiamme_Gialle
    [winner: 3495345 Fiamme, pem_pos=0,
final_scores=-0.93799984, final_local_scores=-0.841974,
pem=0.0, similarity=-21.03647|
    RECALL MISS: 35670529 Gruppi_Sportivi_Fiamme_Gialle]
    [winner: 4204370 Geprüfte_Sicherheit, pem_pos=1,
final_scores=-1.3670254, final_local_scores=-1.3048904,
pem=-1.8700523, similarity=-24.681185|
    RECALL MISS: 35670529 Gruppi_Sportivi_Fiamme_Gialle]
1: 249828 Lecco
    [pem_pos=0, final_scores=0.22003956,
final_local_scores=0.24619839, pem=-0.120320596,
similarity=12.091086]
2: 14532 Italy
    [pem_pos=0, final_scores=0.68489814,
final_local_scores=0.7215846, pem=-0.75948656,
similarity=29.753443]
3: 5192170 ICF_Canoe_Sprint_World_Championships
    [pem_pos=0, final_scores=0.10418309,
final_local_scores=0.13568825, pem=0.0, similarity=8.138921]
4: 18603746 Beijing
    [pem_pos=0, final_scores=0.41583115,
final_local_scores=0.4548527, pem=-0.042161036,
similarity=17.892704]
5: 24663466 Lucia_Micheli
    [pem_pos=0, final_scores=0.2373756,
final_local_scores=0.3355488, pem=0.0, similarity=14.103158]
6: 4443 Barcelona
    [pem_pos=0, final_scores=0.5238186,
final_local_scores=0.5399936, pem=-0.37077636,
similarity=22.220537]
1: 4683072 Antonio_Rossi
    [pem_pos=0, final_scores=0.36522204,
final_local_scores=0.3968563, pem=0.0, similarity=15.932696]
2: 4683072 Antonio_Rossi
    [pem_pos=0, final_scores=0.63722867,
final_local_scores=0.69005704, pem=0.0, similarity=24.682392]
3: 14532 Italy
    [pem_pos=1, final_scores=0.7677713,
final_local_scores=0.81211096, pem=-0.75948656,
similarity=32.45493]
4: 27566 Summer_Olympic_Games

```

```

    [pem_pos=0, final_scores=0.17073114,
final_local_scores=0.12588656, pem=-0.4112264,
similarity=10.082731]
5: 4683072 Antonio_Rossi
    [pem_pos=0, final_scores=0.51417714,
final_local_scores=0.55564195, pem=0.0, similarity=20.671177]
6: 27566 Summer_Olympic_Games
    [pem_pos=0, final_scores=0.11050181,
final_local_scores=0.060095176, pem=-0.4112264,
similarity=8.119386]

```

Sample: b'South_Carolina_Highway_200'

```

[1 South Carolina 1] Highway 200 -LRB- SC 200 -RRB- is a
primary 1 state highway in the [2 U.S. 2] state of [1 1 South
Carolina 1] . It connects the cities of [3 2 Winnsboro 3] , [4
3 Great Falls 4] , and [5 4 Lancaster 5] . SC 200 is mostly a
rural two-lane highway , traversing northwest from [2 2 Winns
boro 2] , through [6 5 Great Falls 6] and [7 6 Lancaster 7] ,
to the [3 3 North Carolina 3] state line . Established in
1937 as a renumbering part of 7 SC 93 to match 8 NC 200 ; it
originally traversed from to the [4 4 North Carolina 4] state
line . By 1952 , SC 200 was extended to its current southern
terminus in [8 9 Winnsboro 8] ; this replaced 10 SC 93 from [9
11 Lancaster 9] to [10 12 Great Falls 10] and 13 SC 22 from [1
1 14 Great Falls 11] to 15 Winnsboro

```

1: 27956 South_Carolina

2: 134477 Winnsboro,_South_Carolina

3: 21650 North_Carolina

4: 21650 North_Carolina

1: 332996 State_highway

```

    [winner: 332996 State_highway, pem_pos=0,
final_scores=-1.0693733, final_local_scores=-1.1295995,
pem=-0.024095139, similarity=-29.488754]

```

```

    gt: 332996 State_highway, pem_pos=0,
final_scores=-1.0693733, final_local_scores=-1.1295995,
pem=-0.024095139, similarity=-29.488754 ]

```

2: 134477 Winnsboro,_South_Carolina

```

    [winner: 115578 Winnsboro,_Louisiana, pem_pos=1,
final_scores=0.23531984, final_local_scores=0.22233835,
pem=-0.64974207, similarity=14.258131]

```

```

    gt: 134477 Winnsboro,_South_Carolina, pem_pos=0,
final_scores=0.1936727, final_local_scores=0.15557936,
pem=-0.4733174, similarity=11.306485 ]

```

3: 134438 Great_Falls,_South_Carolina

```

    [winner: 123354 Great_Falls,_Montana, pem_pos=0,
final_scores=0.50423855, final_local_scores=0.48646295,
pem=-0.63523185, similarity=22.061226]

```

```

    gt: 134438 Great_Falls,_South_Carolina, pem_pos=3,
final_scores=0.029542074, final_local_scores=0.028438836,
pem=-2.3000362, similarity=17.446339 ]

```

4: 134549 Lancaster,_South_Carolina

```

    [winner: 82388 Lancaster,_Pennsylvania, pem_pos=1,

```


final_scores=0.1297891, final_local_scores=0.075496525,
pem=-2.1130877, similarity=17.833979|
gt: 134549 Lancaster, South Carolina, pem_pos=13,
final_scores=-0.3438301, final_local_scores=-0.43046135,
pem=-3.7676427, similarity=11.732915]
5: 134438 Great_Falls, South Carolina
[winner: 123354 Great_Falls, Montana, pem_pos=0,
final_scores=0.41701704, final_local_scores=0.3917457,
pem=-0.63523185, similarity=19.234674|
gt: 134438 Great_Falls, South Carolina, pem_pos=3,
final_scores=-0.034651518, final_local_scores=-0.053474218,
pem=-2.3000362, similarity=15.00189]
6: 134549 Lancaster, South Carolina
[winner: 82388 Lancaster, Pennsylvania, pem_pos=1,
final_scores=0.1565562, final_local_scores=0.10522404,
pem=-2.1130877, similarity=18.721107|
gt: 134549 Lancaster, South Carolina, pem_pos=13,
final_scores=-0.29419905, final_local_scores=-0.37592807,
pem=-3.7676427, similarity=13.360296]
7: 10481800 South_Carolina_Highway_93
[winner: 27956 South_Carolina, pem_pos=0,
final_scores=-0.20687522, final_local_scores=-0.30784965,
pem=-1.67687, similarity=4.0219455|
RECALL MISS: 10481800 South_Carolina_Highway_93]
[winner: 22062456 National_Highway_93_(India), pem_pos=1,
final_scores=-1.036109, final_local_scores=-0.95013064,
pem=-2.7963648, similarity=-9.057013|
RECALL MISS: 10481800 South_Carolina_Highway_93]
8: 8093071 North_Carolina_Highway_200
[winner: 8093071 North_Carolina_Highway_200, pem_pos=0,
final_scores=-0.2959134, final_local_scores=-0.24887237,
pem=0.0, similarity=-3.337135|
gt: 8093071 North_Carolina_Highway_200, pem_pos=0,
final_scores=-0.2959134, final_local_scores=-0.24887237,
pem=0.0, similarity=-3.337135]
9: 134477 Winnsboro, South Carolina
[winner: 115578 Winnsboro, Louisiana, pem_pos=1,
final_scores=0.13736413, final_local_scores=0.11533663,
pem=-0.64974207, similarity=11.064986|
gt: 134477 Winnsboro, South Carolina, pem_pos=0,
final_scores=0.088096544, final_local_scores=0.04025346,
pem=-0.4733174, similarity=7.86493]
10: 10481800 South_Carolina_Highway_93
[winner: 27956 South_Carolina, pem_pos=0,
final_scores=-0.2799979, final_local_scores=-0.38772506,
pem=-1.67687, similarity=1.6383041|
RECALL MISS: 10481800 South_Carolina_Highway_93]
[winner: 10275067 Malaysia_Federal_Route_93, pem_pos=4,
final_scores=-1.0606797, final_local_scores=-0.9381755,
pem=-3.1110113, similarity=-6.989151|
RECALL MISS: 10481800 South_Carolina_Highway_93]
11: 134549 Lancaster, South Carolina

[winner: 82388 Lancaster, Pennsylvania, pem_pos=1,
final_scores=0.08648081, final_local_scores=0.028188795,
pem=-2.1130877, similarity=16.422222]
gt: 134549 Lancaster, South Carolina, pem_pos=13,
final_scores=-0.40670902, final_local_scores=-0.49914703,
pem=-3.7676427, similarity=9.683197]
12: 134438 Great_Falls, South Carolina
[winner: 123354 Great_Falls, Montana, pem_pos=0,
final_scores=0.27902073, final_local_scores=0.24100563,
pem=-0.63523185, similarity=14.736291]
gt: 134438 Great_Falls, South Carolina, pem_pos=3,
final_scores=-0.14660206, final_local_scores=-0.17576319,
pem=-2.3000362, similarity=11.352543]
13: 5994113 South_Carolina_Highway_22
[winner: 27956 South_Carolina, pem_pos=0,
final_scores=-0.21683873, final_local_scores=-0.31873327,
pem=-1.67687, similarity=3.6971567]
RECALL MISS: 5994113 South_Carolina_Highway_22]
[winner: 19660 May_22, pem_pos=1,
final_scores=-0.79314053, final_local_scores=-0.5833006,
pem=-3.022482, similarity=3.119586]
RECALL MISS: 5994113 South_Carolina_Highway_22]
14: 134438 Great_Falls, South Carolina
[winner: 123354 Great_Falls, Montana, pem_pos=0,
final_scores=0.15323885, final_local_scores=0.10424179,
pem=-0.63523185, similarity=10.654985]
gt: 134438 Great_Falls, South Carolina, pem_pos=3,
final_scores=-0.28780776, final_local_scores=-0.3303182,
pem=-2.3000362, similarity=6.740313]
15: 134477 Winnsboro, South Carolina
[winner: 134477 Winnsboro, South Carolina, pem_pos=0,
final_scores=-0.09900643, final_local_scores=-0.17079632,
pem=-0.4733174, similarity=1.5667838]
gt: 134477 Winnsboro, South Carolina, pem_pos=0,
final_scores=-0.09900643, final_local_scores=-0.17079632,
pem=-0.4733174, similarity=1.5667838]
1: 27956 South_Carolina
[pem_pos=0, final_scores=0.51122254,
final_local_scores=0.48145753, pem=-0.27085266,
similarity=19.930304]
2: 134477 Winnsboro, South_Carolina
[pem_pos=0, final_scores=0.20977114,
final_local_scores=0.17316449, pem=-0.4733174,
similarity=11.831261]
3: 21650 North_Carolina
[pem_pos=0, final_scores=0.42987728,
final_local_scores=0.39373112, pem=-0.21680297,
similarity=17.018442]
4: 21650 North_Carolina
[pem_pos=0, final_scores=0.47162622,
final_local_scores=0.43933553, pem=-0.21680297,
similarity=18.37937]

1: 27956 South_Carolina
 [pem_pos=0, final_scores=0.14480232,
 final_local_scores=0.08119908, pem=-0.27085266,
 similarity=7.985794]
 2: 3434750 United_States
 [pem_pos=0, final_scores=0.4855258,
 final_local_scores=0.48893434, pem=-0.14075686,
 similarity=19.445946]
 3: 115578 Winnsboro,_Louisiana
 [pem_pos=1, final_scores=0.23531984,
 final_local_scores=0.22233835, pem=-0.64974207,
 similarity=14.258131]
 4: 123354 Great_Falls,_Montana
 [pem_pos=0, final_scores=0.50423855,
 final_local_scores=0.48646295, pem=-0.63523185,
 similarity=22.061226]
 5: 82388 Lancaster,_Pennsylvania
 [pem_pos=1, final_scores=0.1297891,
 final_local_scores=0.075496525, pem=-2.1130877,
 similarity=17.833979]
 6: 123354 Great_Falls,_Montana
 [pem_pos=0, final_scores=0.41701704,
 final_local_scores=0.3917457, pem=-0.63523185,
 similarity=19.234674]
 7: 82388 Lancaster,_Pennsylvania
 [pem_pos=1, final_scores=0.1565562,
 final_local_scores=0.10522404, pem=-2.1130877,
 similarity=18.721107]
 8: 115578 Winnsboro,_Louisiana
 [pem_pos=1, final_scores=0.13736413,
 final_local_scores=0.11533663, pem=-0.64974207,
 similarity=11.064986]
 9: 82388 Lancaster,_Pennsylvania
 [pem_pos=1, final_scores=0.08648081,
 final_local_scores=0.028188795, pem=-2.1130877,
 similarity=16.422222]
 10: 123354 Great_Falls,_Montana
 [pem_pos=0, final_scores=0.27902073,
 final_local_scores=0.24100563, pem=-0.63523185,
 similarity=14.736291]
 11: 123354 Great_Falls,_Montana
 [pem_pos=0, final_scores=0.15323885,
 final_local_scores=0.10424179, pem=-0.63523185,
 similarity=10.654985]

Sample: b'Arnold_Townsend'

[1 Arnold Frederick Townsend 1] -LRB- 29 March 1912 -- 27
 February 1994 -RRB- was an [1 1 English 1] 1 cricket er who
 played for [2 2 Derbyshire 2] from 1934 to 1950 . [2 Townsend
 2] was born in [3 3 Long Eaton 3] , [3 2 Derbyshire 3] . He
 played his debut game for [4 Derbyshire 4] against [4 4 Leice
 stershire 4] in August 1934 , which was his only game in the 3

1934 season when he scored 7 and 0 . He played six games in the 4 1935 season without distinction , often being denied the chance to play a second innings and during [5 Derbyshire 5] 's [6 5 Championship-winning 6] 1936 season he played only twice . He appeared in three games in the 6 1937 season but in the 7 1938 season came to the fore playing more matches and hitting two half-centuries . In the 8 1939 season he played a fuller season and made his top score of 142 not out against [5 5 Somerset 5] . [7 Townsend 7] went to [8 South Africa 8] during the war and played one first-class match against a team of [9 South Africans 9] , playing for an Air Force XI . He played for a [10 Derbyshire 10] team in 1945 and when the championship restarted in the 9 1946 season he became a first-choice batsman . Against [6 6 Warwickshire 6] , he scored 105 against the economical , penetrating bowling of [11 Test 11] player [7 7 Eric Hollies 7] . During in the 10 1947 season , [12 Townsend 12] scored 137 against [8 8 Yorkshire 8] and 102 against [9 9 Surrey 9] and managed eight half-centuries , and was the most successful [13 Derbyshire 13] batsman during the season . In the 11 1948 season , he scored 102 not out against [10 10 Lancashire 10] . In the 12 1949 season he sustained an eye injury when fielding at short leg and this seriously affected his ability to focus . As a result he played fewer games in 1949 and in 13 1950 played only twice before ending his first class career . In his last three years he was also playing for [14 Derbyshire 14] second XI in the minor counties competition . [15 Townsend 15] was a right-handed batsman and played 200 innings in 117 first class matches with an average of 23.13 and a top score of 142 not out . He bowled 11 overs without taking a wicket . [16 Wisden 16] described him as `` the epitome of -LSB- [17 Derbyshire 17] 's -RSB- patient watchful batting '' . [18 Townsend 18] died in [11 11 Derby 11] at the age of 81 . His brother , [19 Leslie Townsend 19] , was a four-time

1: 9316 England

2: 1829984 Derbyshire_County_Cricket_Club

3: 412761 Long_Eaton

4: 1622318 Leicestershire_County_Cricket_Club

5: 1622178 Somerset_County_Cricket_Club

6: 667292 Warwickshire_County_Cricket_Club

7: 611473 Eric_Hollies

8: 869945 Yorkshire_County_Cricket_Club

9: 766575 Surrey_County_Cricket_Club

10: 749919 Lancashire_County_Cricket_Club

11: 18727355 Derby

1: 25675557 Cricket

[winner: 330091 Rhaphidophoridae, pem_pos=10,
final_scores=-1.5157044, final_local_scores=-1.4719307,
pem=-2.56456, similarity=-25.889162|

gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.7641548, final_local_scores=-1.9660239,

pem=-0.07532314, similarity=-54.170742]
2: 71070 Derbyshire
[winner: 1829984 Derbyshire_County_Cricket_Club,
pem_pos=1, final_scores=0.21026997,
final_local_scores=0.13062617, pem=-1.4828352,
similarity=16.051746|
gt: 71070 Derbyshire, pem_pos=0,
final_scores=0.11725317, final_local_scores=0.12291086,
pem=-0.34993473, similarity=9.660618]
3: 25708351 Derbyshire_County_Cricket_Club_in_1934
[winner: 7620901 1934_English_cricket_season, pem_pos=1,
final_scores=-0.778771, final_local_scores=-0.8168649,
pem=-0.50209194, similarity=-17.556711|
gt: 25708351 Derbyshire_County_Cricket_Club_in_1934,
pem_pos=4, final_scores=-1.011251,
final_local_scores=-1.0756097, pem=-2.1559815,
similarity=-16.284061]
4: 24764708 Derbyshire_County_Cricket_Club_in_1935
[winner: 24764708 Derbyshire_County_Cricket_Club_in_1935,
pem_pos=2, final_scores=-0.81603354,
final_local_scores=-0.8521286, pem=-1.0116009,
similarity=-15.838263|
gt: 24764708 Derbyshire_County_Cricket_Club_in_1935,
pem_pos=2, final_scores=-0.81603354,
final_local_scores=-0.8521286, pem=-1.0116009,
similarity=-15.838263]
5: 24368896 Derbyshire_County_Cricket_Club_in_1936
[winner: 683599 County_Championship, pem_pos=0,
final_scores=0.04166542, final_local_scores=-0.04109952,
pem=0.0, similarity=2.8632214|
RECALL MISS: 24368896
Derbyshire_County_Cricket_Club_in_1936]
[winner: 34673 1936, pem_pos=0, final_scores=-0.7700331,
final_local_scores=-0.6022886, pem=-0.8083407,
similarity=-9.487898|
RECALL MISS: 24368896
Derbyshire_County_Cricket_Club_in_1936]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.5094352, final_local_scores=-1.573431,
pem=-1.219945, similarity=-36.23036|
RECALL MISS: 24368896
Derbyshire_County_Cricket_Club_in_1936]
6: 27038709 Derbyshire_County_Cricket_Club_in_1937
[winner: 34684 1937, pem_pos=0, final_scores=-0.7753122,
final_local_scores=-0.6428251, pem=-0.6998789,
similarity=-11.287421|
RECALL MISS: 27038709
Derbyshire_County_Cricket_Club_in_1937]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.4655608, final_local_scores=-1.5255048,
pem=-1.219945, similarity=-34.800144|
RECALL MISS: 27038709

Derbyshire_County_Cricket_Club_in_1937]
7: 27912063 Derbyshire_County_Cricket_Club_in_1938
[winner: 19283852 1938, pem_pos=0,
final_scores=-0.68679804, final_local_scores=-0.5487125,
pem=-0.81177574, similarity=-7.870399]
RECALL MISS: 27912063
Derbyshire_County_Cricket_Club_in_1938]
[winner: 1340236 Eastertide, pem_pos=2,
final_scores=-1.5721462, final_local_scores=-1.5765209,
pem=-2.3723626, similarity=-30.055538]
RECALL MISS: 27912063
Derbyshire_County_Cricket_Club_in_1938]
8: 25538336 Derbyshire_County_Cricket_Club_in_1939
[winner: 34614 1939, pem_pos=0, final_scores=-1.1127572,
final_local_scores=-1.0379188, pem=-0.65361583,
similarity=-23.329393]
RECALL MISS: 25538336
Derbyshire_County_Cricket_Club_in_1939]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.5273744, final_local_scores=-1.5930269,
pem=-1.219945, similarity=-36.815136]
RECALL MISS: 25538336
Derbyshire_County_Cricket_Club_in_1939]
9: 24518361 Derbyshire_County_Cricket_Club_in_1946
[winner: 34631 1946, pem_pos=0, final_scores=-0.83290136,
final_local_scores=-0.71719635, pem=-0.5980445,
similarity=-14.060596]
RECALL MISS: 24518361
Derbyshire_County_Cricket_Club_in_1946]
[winner: 1340236 Eastertide, pem_pos=2,
final_scores=-1.4970171, final_local_scores=-1.4944539,
pem=-2.3723626, similarity=-27.606497]
RECALL MISS: 24518361
Derbyshire_County_Cricket_Club_in_1946]
10: 26854754 Derbyshire_County_Cricket_Club_in_1947
[winner: 34584 1947, pem_pos=0, final_scores=-0.93605596,
final_local_scores=-0.8447544, pem=-0.55926085,
similarity=-18.078096]
RECALL MISS: 26854754
Derbyshire_County_Cricket_Club_in_1947]
[winner: 1340236 Eastertide, pem_pos=2,
final_scores=-1.596456, final_local_scores=-1.6030757,
pem=-2.3723626, similarity=-30.84799]
RECALL MISS: 26854754
Derbyshire_County_Cricket_Club_in_1947]
11: 27886945 Derbyshire_County_Cricket_Club_in_1948
[winner: 34612 1948, pem_pos=0, final_scores=-0.6831848,
final_local_scores=-0.53703153, pem=-0.67699766,
similarity=-8.254761]
RECALL MISS: 27886945
Derbyshire_County_Cricket_Club_in_1948]
[winner: 690345 Season_(sports), pem_pos=0,

final_scores=-1.6116421, final_local_scores=-1.6850765,
pem=-1.219945, similarity=-39.56208|
RECALL MISS: 27886945
Derbyshire_County_Cricket_Club_in_1948]
12: 25081484 Derbyshire_County_Cricket_Club_in_1949
[winner: 21288296 1949_Baltimore_Colts_season, pem_pos=0,
final_scores=-0.7218572, final_local_scores=-0.61220753,
pem=0.0, similarity=-14.179782|
RECALL MISS: 25081484
Derbyshire_County_Cricket_Club_in_1949]
[winner: 1340236 Eastertide, pem_pos=2,
final_scores=-1.627273, final_local_scores=-1.6367385,
pem=-2.3723626, similarity=-31.85255|
RECALL MISS: 25081484
Derbyshire_County_Cricket_Club_in_1949]
13: 31144749 Derbyshire_County_Cricket_Club_in_1950
[winner: 3245375 James_Warburg, pem_pos=0,
final_scores=-0.67000186, final_local_scores=-0.63399017,
pem=-0.0840831, similarity=-14.372561|
RECALL MISS: 31144749
Derbyshire_County_Cricket_Club_in_1950]
1: 9316 England
[pem_pos=1, final_scores=0.50629824,
final_local_scores=0.54341036, pem=-1.3527626,
similarity=27.662695]
2: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6167481,
final_local_scores=0.57480806, pem=-1.4828352,
similarity=29.30702]
3: 412761 Long_Eaton
[pem_pos=0, final_scores=0.32040375,
final_local_scores=0.35982582, pem=0.0, similarity=14.827633]
4: 1622318 Leicestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.7071721,
final_local_scores=0.6815526, pem=-1.0465667,
similarity=30.119995]
5: 1622178 Somerset_County_Cricket_Club
[pem_pos=1, final_scores=0.51683456,
final_local_scores=0.47123027, pem=-2.4139037,
similarity=31.279346]
6: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6526392,
final_local_scores=0.62043387, pem=-1.1093602,
similarity=28.637571]
7: 611473 Eric_Hollies
[pem_pos=0, final_scores=0.6143159,
final_local_scores=0.6385103, pem=0.0, similarity=23.144133]
8: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6880723,
final_local_scores=0.65530115, pem=-1.7779728,
similarity=33.3141]
9: 766575 Surrey_County_Cricket_Club

[pem_pos=1, final_scores=0.7300553,
final_local_scores=0.7034188, pem=-1.7952642,
similarity=34.84406]
10: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.5273295,
final_local_scores=0.4865008, pem=-1.7295884,
similarity=28.013638]
11: 18727355 Derby
[pem_pos=0, final_scores=0.38098544,
final_local_scores=0.40573716, pem=-0.40742427,
similarity=18.413355]
1: 10767655 Arnold_Townsend
[pem_pos=0, final_scores=0.16817717,
final_local_scores=0.14989913, pem=0.0, similarity=8.563002]
2: 3249168 Leslie_Townsend_(cricketer)
[pem_pos=0, final_scores=0.46256953,
final_local_scores=0.4687034, pem=0.0, similarity=18.076757]
3: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.21026997,
final_local_scores=0.13062617, pem=-1.4828352,
similarity=16.051746]
4: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.70691437,
final_local_scores=0.67330104, pem=-1.4828352,
similarity=32.246246]
5: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.681753,
final_local_scores=0.645816, pem=-1.4828352,
similarity=31.426037]
6: 683599 County_Championship
[pem_pos=0, final_scores=0.04166542,
final_local_scores=-0.04109952, pem=0.0, similarity=2.8632214]
7: 3249168 Leslie_Townsend_(cricketer)
[pem_pos=0, final_scores=0.39737564,
final_local_scores=0.39655328, pem=0.0, similarity=15.923654]
8: 17416221 South_Africa
[pem_pos=0, final_scores=0.50536615,
final_local_scores=0.60817826, pem=-0.26251382,
similarity=23.666557]
9: 759263 South_Africa_national_cricket_team
[pem_pos=0, final_scores=0.23486711,
final_local_scores=0.1940088, pem=-0.71672493,
similarity=13.776984]
10: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.5221207,
final_local_scores=0.47144204, pem=-1.4828352,
similarity=26.222372]
11: 30863 Test_cricket
[pem_pos=0, final_scores=0.6249754,
final_local_scores=0.63192, pem=-0.23950954,
similarity=24.249956]
12: 3249168 Leslie_Townsend_(cricketer)

[pem_pos=0, final_scores=0.6450298,
 final_local_scores=0.6680135, pem=0.0, similarity=24.024569]
 13: 1829984 Derbyshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.6940992,
 final_local_scores=0.6593024, pem=-1.4828352,
 similarity=31.828499]
 14: 1829984 Derbyshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.7317889,
 final_local_scores=0.70047265, pem=-1.4828352,
 similarity=33.057102]
 15: 3249168 Leslie_Townsend_(cricketer)
 [pem_pos=0, final_scores=0.37353778,
 final_local_scores=0.37144977, pem=0.0, similarity=15.174515]
 16: 1582694 Wisden_Cricketers'_Almanack
 [pem_pos=0, final_scores=0.2050678,
 final_local_scores=0.17120805, pem=0.0, similarity=9.198903]
 17: 1829984 Derbyshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.46008158,
 final_local_scores=0.4036737, pem=-1.4828352,
 similarity=24.20003]
 18: 3249168 Leslie_Townsend_(cricketer)
 [pem_pos=0, final_scores=0.22318198,
 final_local_scores=0.2072089, pem=0.0, similarity=10.27324]
 19: 3249168 Leslie_Townsend_(cricketer)
 [pem_pos=1, final_scores=0.44450778,
 final_local_scores=0.44936985, pem=0.0, similarity=17.499805]

Sample: b'Cecelia_Joyce'

[1 Cecelia Nora Isobel Mary Joyce 1] -LRB- born 25 July 1983
 in [1 1 Wicklow 1] , [2 2 Republic of Ireland 2] -RRB- is an [2
 2 Irish 2] . A right-handed and 1 leg spin , she has played
 twenty for the [3 2 Ireland 3] women 's cricket team since
 July 2001 . Joyce made her [4 ODI 4] debut for [5 Ireland 5]
 against [6 3 Australia 6] on 14 July 2001 , in the second
 match of a series . She also played in the third match of the
 series , and against [7 4 Scotland 7] in the 5 European Champ
 ionship . She next played in the [3 3 IWCC Trophy 3] in the [4
 4 Netherlands 4] in July 2003 . The following year , she
 played three [8 ODIs 8] against [9 New Zealand 9] in [5 5 Dub
 lin 5] and in 2005 , played in the [10 6 World Cup 10] in [6 6
 South Africa 6] . She also played against [11 Australia 11]
 and in the European Championship in 2005 . She played four [12
 ODIs 12] in 2006 , two each against [13 7 India 13] and the 8
 Netherlands . In One-Day International matches , [14 Joyce 14]
 has scored 283 9 runs at an of 16.64 , with a top score of 77
 against the [15 Netherlands 15] . She has taken just one ,
 against Scotland . Joyce comes from a cricketing family . Her
 twin sister [16 10 Isobel 16] has played [17 11 Test 17] and [18
 ODI 18] 12 cricket for Ireland , whilst her three brothers
 13 Dominick , 14 Ed and 15 Gus have all played for the [19 16
 Ireland 19] men 's team . Ed has also played for [7 7 England
 7]

```

1: 500459 Wicklow
2: 14560 Republic_of_Ireland
3: 11393571 Women's_Cricket_World_Cup_Qualifier
4: 21148 Netherlands
5: 8504 Dublin
6: 17416221 South_Africa
7: 759125 England_cricket_team
1: 18121 Leg_spin
  [winner: 18121 Leg_spin, pem_pos=0,
  final_scores=-0.8535134, final_local_scores=-0.94980335,
  pem=-0.31609803, similarity=-22.535324]
  gt: 18121 Leg_spin, pem_pos=0, final_scores=-0.8535134,
  final_local_scores=-0.94980335, pem=-0.31609803,
  similarity=-22.535324 ]
2: 7045206 Ireland_women's_cricket_team
  [winner: 14560 Republic_of_Ireland, pem_pos=1,
  final_scores=0.21453689, final_local_scores=0.21749035,
  pem=-1.5040284, similarity=18.7592]
  RECALL MISS: 7045206 Ireland_women's_cricket_team]
  [winner: 330091 Rhaphidophoridae, pem_pos=10,
  final_scores=-1.302258, final_local_scores=-1.2641575,
  pem=-2.56456, similarity=-19.688793]
  RECALL MISS: 7045206 Ireland_women's_cricket_team]
  [winner: 14708158 TEAMS_(cable_system), pem_pos=3,
  final_scores=-1.4698904, final_local_scores=-1.455847,
  pem=-3.9775972, similarity=-17.724873]
  RECALL MISS: 7045206 Ireland_women's_cricket_team]
  [winner: 7277939 Diotima_of_Mantineia, pem_pos=12,
  final_scores=-1.5541263, final_local_scores=-1.5103753,
  pem=-3.218876, similarity=-23.478148]
  RECALL MISS: 7045206 Ireland_women's_cricket_team]
  [winner: 33880936 English_possessive, pem_pos=0,
  final_scores=-1.8891633, final_local_scores=-1.9510751,
  pem=-0.7098606, similarity=-50.273922]
  RECALL MISS: 7045206 Ireland_women's_cricket_team]
3: 2154946 Australia_national_women's_cricket_team
  [winner: 374833 Australia_national_cricket_team,
  pem_pos=1, final_scores=0.6780094,
  final_local_scores=0.6669749, pem=-3.4989398,
  similarity=43.021355]
  gt: 2154946 Australia_national_women's_cricket_team,
  pem_pos=8, final_scores=-0.32560956,
  final_local_scores=-0.40895176, pem=-5.0368776,
  similarity=19.2771 ]
4: 7076384 Scotland_national_women's_cricket_team
  [winner: 1859357 Scotland_national_cricket_team,
  pem_pos=3, final_scores=0.037993893,
  final_local_scores=-0.025484204, pem=-3.70665,
  similarity=23.486553]
  RECALL MISS: 7076384
  Scotland_national_women's_cricket_team]
5: 11382347 Women's_European_Cricket_Championship

```

[winner: 249510 UEFA_European_Championship, pem_pos=0,
final_scores=-0.10041313, final_local_scores=-0.114650905,
pem=-1.0168781, similarity=6.1982427|

RECALL MISS: 11382347

Women's_European_Cricket_Championship]

6: 3977980 Women's_Cricket_World_Cup

[winner: 7239 Cricket_World_Cup, pem_pos=4,
final_scores=0.39214844, final_local_scores=0.376472,
pem=-3.1731894, similarity=32.580685|

RECALL MISS: 3977980 Women's_Cricket_World_Cup]

7: 2281524 India_national_women's_cricket_team

[winner: 407754 India_national_cricket_team, pem_pos=1,
final_scores=0.7229236, final_local_scores=0.7215928,
pem=-3.5152702, similarity=44.740067|

gt: 2281524 India_national_women's_cricket_team,
pem_pos=6, final_scores=-0.19171782,
final_local_scores=-0.23863465, pem=-4.7643886,
similarity=22.877865]

8: 7046272 Netherlands_national_women's_cricket_team

[winner: 9647657 Netherlands_national_football_team,
pem_pos=1, final_scores=0.006498918,
final_local_scores=0.012834668, pem=-2.660505,
similarity=18.940966|

RECALL MISS: 7046272

Netherlands_national_women's_cricket_team]

9: 896043 Run_(cricket)

[winner: 300747 Run_(baseball), pem_pos=0,
final_scores=-1.0733824, final_local_scores=-1.1023681,
pem=0.0, similarity=-28.80715|

gt: 896043 Run_(cricket), pem_pos=2,
final_scores=-1.1768863, final_local_scores=-1.2900975,
pem=-1.2719624, similarity=-27.492239]

10: 11395781 Isobel_Joyce

[winner: 11344005 Cecelia_Joyce, pem_pos=0,
final_scores=0.13878451, final_local_scores=0.1570727,
pem=0.0, similarity=8.777076|

gt: 11395781 Isobel_Joyce, pem_pos=3,
final_scores=-0.38864765, final_local_scores=-0.38679218,
pem=-3.2053905, similarity=9.978475]

11: 1330193 Women's_Test_cricket

[winner: 30863 Test_cricket, pem_pos=0,
final_scores=0.11727555, final_local_scores=0.051941574,
pem=-0.23950954, similarity=6.942243|

gt: 1330193 Women's_Test_cricket, pem_pos=3,
final_scores=-0.41847324, final_local_scores=-0.49833643,
pem=-2.9393153, similarity=5.2028136]

12: 25675557 Cricket

[winner: 330091 Rhaphidophoridae, pem_pos=10,
final_scores=-1.4283559, final_local_scores=-1.4019003,
pem=-2.56456, similarity=-23.799313|

gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.8605348, final_local_scores=-2.0750444,

```

pem=-0.07532314, similarity=-57.424133 ]
13: 6488937 Dominick_Joyce
    [winner: 6488937 Dominick_Joyce, pem_pos=1,
final_scores=-0.17263357, final_local_scores=-0.1908723,
pem=-1.2868257, similarity=5.391661]
    gt: 6488937 Dominick_Joyce, pem_pos=1,
final_scores=-0.17263357, final_local_scores=-0.1908723,
pem=-1.2868257, similarity=5.391661 ]
14: 2059682 Ed_Joyce
    [winner: 375320 Ed_(TV_series), pem_pos=0,
final_scores=-0.44976026, final_local_scores=-0.3920285,
pem=-0.7337421, similarity=-3.6189961]
    RECALL MISS: 2059682 Ed_Joyce]
15: 11341066 Gus_Joyce
    [winner: 2962285 Gus_Smith, pem_pos=5,
final_scores=-0.47488594, final_local_scores=-0.38983303,
pem=-2.981962, similarity=8.67269]
    gt: 11341066 Gus_Joyce, pem_pos=10,
final_scores=-0.9161188, final_local_scores=-0.93672216,
pem=-3.5277367, similarity=-4.679564 ]
16: 2217991 Ireland_cricket_team
    [winner: 14560 Republic_of_Ireland, pem_pos=1,
final_scores=0.21401678, final_local_scores=0.21166298,
pem=-1.5040284, similarity=18.585299]
    gt: 2217991 Ireland_cricket_team, pem_pos=3,
final_scores=-0.33073765, final_local_scores=-0.43441823,
pem=-3.715907, similarity=11.3334875 ]
    [winner: 7809948 Men's_colleges_in_the_United_States,
pem_pos=9, final_scores=-1.2400868,
final_local_scores=-1.2142346, pem=-2.5649493,
similarity=-18.196875]
    RECALL MISS: 2217991 Ireland_cricket_team]
    [winner: 14708158 TEAMS_(cable_system), pem_pos=3,
final_scores=-1.4428288, final_local_scores=-1.4262865,
pem=-3.9775972, similarity=-16.842726]
    RECALL MISS: 2217991 Ireland_cricket_team]
    [winner: 33880936 English_possessive, pem_pos=0,
final_scores=-1.9110082, final_local_scores=-1.9749374,
pem=-0.7098606, similarity=-50.986027]
    RECALL MISS: 2217991 Ireland_cricket_team]
1: 500459 Wicklow
    [pem_pos=1, final_scores=0.20612602,
final_local_scores=0.25421172, pem=-0.84520566,
similarity=16.272259]
2: 14560 Republic_of_Ireland
    [pem_pos=0, final_scores=0.5413046,
final_local_scores=0.5691752, pem=-0.2122539,
similarity=22.229305]
3: 11393571 Women's_Cricket_World_Cup_Qualifier
    [pem_pos=0, final_scores=0.09995441,
final_local_scores=0.110965446, pem=0.0, similarity=7.401143]
4: 21148 Netherlands

```

[pem_pos=0, final_scores=0.48976642,
final_local_scores=0.5261734, pem=-0.24940795,
similarity=21.148096]
5: 8504 Dublin
[pem_pos=0, final_scores=0.48725933,
final_local_scores=0.50771344, pem=-0.23150389,
similarity=20.49985]
6: 17416221 South_Africa
[pem_pos=0, final_scores=0.51579744,
final_local_scores=0.5633032, pem=-0.26251382,
similarity=22.327394]
7: 759125 England_cricket_team
[pem_pos=2, final_scores=0.09784739,
final_local_scores=0.031783104, pem=-3.3054674,
similarity=23.01383]
1: 11344005 Cecelia_Joyce
[pem_pos=0, final_scores=0.22698008,
final_local_scores=0.25341293, pem=0.0, similarity=11.6520605]
2: 14560 Republic_of_Ireland
[pem_pos=3, final_scores=0.35421962,
final_local_scores=0.36481327, pem=-2.2576945,
similarity=27.254162]
3: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.21453689,
final_local_scores=0.21749035, pem=-1.5040284,
similarity=18.7592]
4: 381447 One_Day_International
[pem_pos=0, final_scores=0.47838074,
final_local_scores=0.44731402, pem=-0.096647784,
similarity=17.96404]
5: 2217991 Ireland_cricket_team
[pem_pos=3, final_scores=0.2272356,
final_local_scores=0.18083483, pem=-3.715907,
similarity=29.693867]
6: 374833 Australia_national_cricket_team
[pem_pos=1, final_scores=0.6780094,
final_local_scores=0.6669749, pem=-3.4989398,
similarity=43.021355]
7: 1859357 Scotland_national_cricket_team
[pem_pos=3, final_scores=0.037993893,
final_local_scores=-0.025484204, pem=-3.70665,
similarity=23.486553]
8: 381447 One_Day_International
[pem_pos=1, final_scores=0.47460955,
final_local_scores=0.44319457, pem=-0.010680528,
similarity=17.373606]
9: 478115 New_Zealand_national_cricket_team
[pem_pos=1, final_scores=0.3425743,
final_local_scores=0.30747116, pem=-3.142959,
similarity=30.357164]
10: 7239 Cricket_World_Cup
[pem_pos=4, final_scores=0.39214844,

```

final_local_scores=0.376472, pem=-3.1731894,
similarity=32.580685]
11: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.65646166,
final_local_scores=0.64343727, pem=-3.4989398,
similarity=42.318943]
12: 381447 One_Day_International
    [pem_pos=1, final_scores=0.2830881,
final_local_scores=0.23398644, pem=-0.010680528,
similarity=11.130417]
13: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.7229236,
final_local_scores=0.7215928, pem=-3.5152702,
similarity=44.740067]
14: 11344005 Cecelia_Joyce
    [pem_pos=1, final_scores=0.18869151,
final_local_scores=0.2115885, pem=0.0, similarity=10.403935]
15: 21148 Netherlands
    [pem_pos=0, final_scores=0.032975152,
final_local_scores=0.027198374, pem=-0.24940795,
similarity=6.2576857]
16: 11344005 Cecelia_Joyce
    [pem_pos=0, final_scores=0.13878451,
final_local_scores=0.1570727, pem=0.0, similarity=8.777076]
17: 30863 Test_cricket
    [pem_pos=0, final_scores=0.11727555,
final_local_scores=0.051941574, pem=-0.23950954,
similarity=6.942243]
18: 381447 One_Day_International
    [pem_pos=0, final_scores=0.23803063,
final_local_scores=0.18282124, pem=-0.096647784,
similarity=10.071049]
19: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.21401678,
final_local_scores=0.21166298, pem=-1.5040284,
similarity=18.585299]

```

Sample: b'Victor_Croome'

Victor Croome -LRB- 30 November 1899 -- 1 September 1973 -
RRB- was an [1 English 1] 1 cricket er . A right-handed and
occasional 2 wicket-keeper , he played 3 first-class cricket
for the [2 4 Royal Air Force 2] between 1928 and 1930 . Born
in [1 1 Westminster 1] , [2 2 London 2] in 1899 , the son of [3
3 Gloucestershire 3] player , Victor made his first-class
debut for the [3 RAF 3] against the [4 5 Army 4] in 6 1928 .
He also played against the [5 7 Royal Navy 5] the same year .
He played against the same opponents in 8 1929 . After a
final match for the [6 RAF 6] against the [7 Army 7] in 9 1930
, the remainder of his recorded cricket career took place in
the [4 4 Far East 4] . He played for the [8 10 Straits Settle
ments 8] against the [9 11 Federated Malay States 9] in 1932
and 1933 , and for 12 Malaya against and [10 13 Shanghai 10]

```

in [5 5 Hong Kong 5] in 1933 . He died in [6 6 Essex 6]
1: 37432 Westminster
2: 17867 London
3: 1622322 Gloucestershire_County_Cricket_Club
4: 11672 Far_East
5: 13404 Hong_Kong
6: 40100758 Essex
1: 25675557 Cricket
  [winner: 330091 Rhaphidophoridae, pem_pos=10,
  final_scores=-1.4535118, final_local_scores=-1.4032018,
  pem=-2.56456, similarity=-23.838152|
  gt: 25675557 Cricket, pem_pos=0,
  final_scores=-1.6726049, final_local_scores=-1.8267975,
  pem=-0.07532314, similarity=-50.01595 ]
2: 463313 Wicket-keeper
  [winner: 463313 Wicket-keeper, pem_pos=0,
  final_scores=-1.1150404, final_local_scores=-1.1911674,
  pem=0.0, similarity=-31.457094|
  gt: 463313 Wicket-keeper, pem_pos=0,
  final_scores=-1.1150404, final_local_scores=-1.1911674,
  pem=0.0, similarity=-31.457094 ]
3: 11040 First-class_cricket
  [winner: 662844 First-class_citizen, pem_pos=3,
  final_scores=-1.1982752, final_local_scores=-1.130214,
  pem=-1.5935546, similarity=-20.972128|
  gt: 11040 First-class_cricket, pem_pos=0,
  final_scores=-1.3457764, final_local_scores=-1.4754806,
  pem=-0.053819276, similarity=-39.64889 ]
  [winner: 330091 Rhaphidophoridae, pem_pos=10,
  final_scores=-1.3905873, final_local_scores=-1.3344665,
  pem=-2.56456, similarity=-21.786955|
  RECALL MISS: 11040 First-class_cricket]
4: 11458234 Royal_Air_Force_cricket_team
  [winner: 25679 Royal_Air_Force, pem_pos=0,
  final_scores=0.21891849, final_local_scores=0.16786125,
  pem=0.0, similarity=9.099028|
  gt: 11458234 Royal_Air_Force_cricket_team, pem_pos=1,
  final_scores=-0.4120268, final_local_scores=-0.4363978,
  pem=-2.9001539, similarity=6.83822 ]
5: 11458285 British_Army_cricket_team
  [winner: 32087 United_States_Army, pem_pos=0,
  final_scores=0.116301104, final_local_scores=0.08196536,
  pem=-1.138116, similarity=12.724975|
  gt: 11458285 British_Army_cricket_team, pem_pos=18,
  final_scores=-0.6477233, final_local_scores=-0.6959794,
  pem=-4.377645, similarity=7.1266136 ]
6: 7620874 1928_English_cricket_season
  [winner: 34674 1928, pem_pos=0, final_scores=-0.871837,
  final_local_scores=-0.7805572, pem=-0.8454088,
  similarity=-14.606206|
  RECALL MISS: 7620874 1928_English_cricket_season]
7: 3090550 Royal_Navy_Cricket_Club

```

[winner: 26061 Royal_Navy, pem_pos=0,
final_scores=0.14615004, final_local_scores=0.11085649,
pem=0.0, similarity=7.3978915|
gt: 3090550 Royal_Navy_Cricket_Club, pem_pos=1,
final_scores=-0.46044952, final_local_scores=-0.47658634,
pem=-3.026984, similarity=6.328636]
8: 7620878 1929_English_cricket_season
[winner: 6358567 1929_Palestine_riots, pem_pos=0,
final_scores=-0.9140144, final_local_scores=-0.923975,
pem=0.0, similarity=-23.483543|
RECALL MISS: 7620878 1929_English_cricket_season]
9: 5671900 1930_English_cricket_season
[winner: 34911 1930, pem_pos=0, final_scores=-0.89528835,
final_local_scores=-0.76168287, pem=-0.94197285,
similarity=-13.517828|
RECALL MISS: 5671900 1930_English_cricket_season]
10: 14528859 Straits_Settlements_cricket_team
[winner: 51158 Straits_Settlements, pem_pos=0,
final_scores=0.22400062, final_local_scores=0.21589404,
pem=0.0, similarity=10.532422|
gt: 14528859 Straits_Settlements_cricket_team,
pem_pos=1, final_scores=0.033922657,
final_local_scores=0.07371679, pem=-0.601217,
similarity=9.55908]
11: 14438658 Federated_Malay_States_cricket_team
[winner: 336685 Federated_Malay_States, pem_pos=0,
final_scores=0.23988287, final_local_scores=0.25579274,
pem=0.0, similarity=11.723079|
gt: 14438658 Federated_Malay_States_cricket_team,
pem_pos=1, final_scores=-0.055152193,
final_local_scores=-0.0029204637, pem=-0.56818223,
similarity=7.092424]
12: 14293374 Malaya_cricket_team
[winner: 3850267 British_Malaya, pem_pos=0,
final_scores=-0.051039413, final_local_scores=-0.10590574,
pem=-0.90195495, similarity=5.8342457|
RECALL MISS: 14293374 Malaya_cricket_team]
13: 14540512 Shanghai_cricket_team
[winner: 27643 Shanghai, pem_pos=0,
final_scores=0.22503586, final_local_scores=0.22395566,
pem=-0.10072064, similarity=11.320731|
gt: 14540512 Shanghai_cricket_team, pem_pos=7,
final_scores=-0.40007606, final_local_scores=-0.39513493,
pem=-3.874159, similarity=13.366379]
1: 37432 Westminster
[pem_pos=0, final_scores=0.27757508,
final_local_scores=0.25556323, pem=-1.0392612,
similarity=17.367893]
2: 17867 London
[pem_pos=0, final_scores=0.5804946,
final_local_scores=0.5750118, pem=-0.09086737,
similarity=21.743362]

3: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.23476018,
final_local_scores=0.25794357, pem=-1.379793,
similarity=19.290792]
4: 11672 Far_East
[pem_pos=0, final_scores=0.16572689,
final_local_scores=0.1462465, pem=-0.034746617,
similarity=8.642957]
5: 13404 Hong_Kong
[pem_pos=0, final_scores=0.47236627,
final_local_scores=0.4653334, pem=-0.22402541,
similarity=19.194475]
6: 40100758 Essex
[pem_pos=0, final_scores=0.32175237,
final_local_scores=0.29199627, pem=-0.576716,
similarity=15.939735]
1: 9316 England
[pem_pos=1, final_scores=0.6091266,
final_local_scores=0.5978722, pem=-1.3527626,
similarity=29.287947]
2: 25679 Royal_Air_Force
[pem_pos=0, final_scores=0.21891849,
final_local_scores=0.16786125, pem=0.0, similarity=9.099028]
3: 25679 Royal_Air_Force
[pem_pos=0, final_scores=0.27694142,
final_local_scores=0.23124248, pem=-0.044596832,
similarity=11.232974]
4: 32087 United_States_Army
[pem_pos=0, final_scores=0.116301104,
final_local_scores=0.08196536, pem=-1.138116,
similarity=12.724975]
5: 26061 Royal_Navy
[pem_pos=0, final_scores=0.14615004,
final_local_scores=0.11085649, pem=0.0, similarity=7.3978915]
6: 25679 Royal_Air_Force
[pem_pos=0, final_scores=0.29942286,
final_local_scores=0.25580004, pem=-0.044596832,
similarity=11.96582]
7: 32087 United_States_Army
[pem_pos=0, final_scores=0.12994702,
final_local_scores=0.096871436, pem=-1.138116,
similarity=13.169802]
8: 51158 Straits_Settlements
[pem_pos=0, final_scores=0.22400062,
final_local_scores=0.21589404, pem=0.0, similarity=10.532422]
9: 336685 Federated_Malay_States
[pem_pos=0, final_scores=0.23988287,
final_local_scores=0.25579274, pem=0.0, similarity=11.723079]
10: 27643 Shanghai
[pem_pos=0, final_scores=0.22503586,
final_local_scores=0.22395566, pem=-0.10072064,
similarity=11.320731]

Sample: b'Gmina_Przedecz'

Gmina Przedecz is an urban-rural 1 gmina -LRB- administrative district -RRB- in 2 Koło County , [1 1 Greater Poland Voivodeship 1] , in west-central [2 2 Poland 2] . Its seat is the town of 3 Przedecz , which lies approximately north-east of [3 3 Koło 3] and east of the regional capital [4 4 Poznań 4] . The gmina covers an area of , and as of 2006 its total population is 4,319 -LRB- out of which the population of Przedecz amounts to 1,771 , and the population of the rural part of the gmina is 2,548 -RRB- . Apart from the town of Przedecz , Gmina Przedecz contains the villages and settlements of 4 Arkuszewo , 5 Broniszewo , 6 Chrustowo , 7 Dziewczopólko , 8 Dziwie , 9 Holenderki , 10 Jasieniec , 11 Józefowo , 12 Katarzyna , 13 Kłokoczyn , 14 Łączewna , 15 Lipiny , 16 Mieczysławowo , 17 Nowa Wieś Wielka , 18 Rogóźno , 19 Rybno , 20 Zalesie , 21 Żarowo , 22 Zbijewo-Kolonia and 23 Zbijewo-Parcele A . Gmina Przedecz is bordered by the gminas of 24 Babiak , 25 Chodecz , 26 Chodów , 27 Dąbrowice , 28 Izbica Kujawska and 29 Kłodawa

1: 41498 Greater_Poland_Voivodeship

2: 22936 Poland

3: 390527 Koło

4: 23723 Poznań

1: 1234326 Gmina

[winner: 1234326 Gmina, pem_pos=0, final_scores=-0.49148732, final_local_scores=-0.5554253, pem=0.0, similarity=-12.485287]

gt: 1234326 Gmina, pem_pos=0, final_scores=-0.49148732, final_local_scores=-0.5554253, pem=0.0, similarity=-12.485287]

2: 5816537 Koło County

[winner: 5816537 Koło County, pem_pos=0, final_scores=0.013989881, final_local_scores=-0.0068752468, pem=0.0, similarity=3.884542]

gt: 5816537 Koło County, pem_pos=0, final_scores=0.013989881, final_local_scores=-0.0068752468, pem=0.0, similarity=3.884542]

3: 8417121 Przedecz

[winner: 8417121 Przedecz, pem_pos=0, final_scores=0.02869086, final_local_scores=0.12983257, pem=0.0, similarity=7.9641757]

gt: 8417121 Przedecz, pem_pos=0, final_scores=0.02869086, final_local_scores=0.12983257, pem=0.0, similarity=7.9641757]

4: 20994484 Arkuszewo

[winner: 20994484 Arkuszewo, pem_pos=0, final_scores=-0.18779281, final_local_scores=-0.14993432, pem=0.0, similarity=-0.3846264]

gt: 20994484 Arkuszewo, pem_pos=0, final_scores=-0.18779281, final_local_scores=-0.14993432, pem=0.0, similarity=-0.3846264]

5: 20994495 Broniszewo, Koło County
 [winner: 21159769 Broniszewo, Września County, pem_pos=1,
 final_scores=-0.21361892, final_local_scores=-0.19339392,
 pem=-0.83173335, similarity=2.8415477]
 gt: 20994495 Broniszewo, Koło County, pem_pos=4,
 final_scores=-0.37057084, final_local_scores=-0.4046895,
 pem=-1.229506, similarity=-1.3007842]

6: 20994497 Chrustowo, Koło County
 [winner: 21159919 Chrustowo, Września County, pem_pos=1,
 final_scores=-0.22867781, final_local_scores=-0.21613678,
 pem=-0.96639764, similarity=2.8951807]
 gt: 20994497 Chrustowo, Koło County, pem_pos=2,
 final_scores=-0.3677631, final_local_scores=-0.4033361,
 pem=-1.0067255, similarity=-2.4719112]

7: 20994507 Dziewczopólko
 [winner: 20994507 Dziewczopólko, pem_pos=0,
 final_scores=-0.08017455, final_local_scores=-0.005527124,
 pem=0.0, similarity=3.9247727]
 gt: 20994507 Dziewczopólko, pem_pos=0,
 final_scores=-0.08017455, final_local_scores=-0.005527124,
 pem=0.0, similarity=3.9247727]

8: 20994511 Dziwie
 [winner: 20994511 Dziwie, pem_pos=0,
 final_scores=-0.25160408, final_local_scores=-0.22343051,
 pem=0.0, similarity=-2.577899]
 gt: 20994511 Dziwie, pem_pos=0,
 final_scores=-0.25160408, final_local_scores=-0.22343051,
 pem=0.0, similarity=-2.577899]

9: 20994527 Holenderki
 [winner: 20994527 Holenderki, pem_pos=0,
 final_scores=-0.31346652, final_local_scores=-0.28514016,
 pem=0.0, similarity=-4.419438]
 gt: 20994527 Holenderki, pem_pos=0,
 final_scores=-0.31346652, final_local_scores=-0.28514016,
 pem=0.0, similarity=-4.419438]

10: 20994541 Jasieniec, Greater Poland Voivodeship
 [winner: 21669246 Jasieniec, Pomeranian Voivodeship,
 pem_pos=3, final_scores=-0.26863882,
 final_local_scores=-0.23186715, pem=-1.6572468,
 similarity=6.1827]
 gt: 20994541 Jasieniec, Greater Poland Voivodeship,
 pem_pos=9, final_scores=-0.54166025,
 final_local_scores=-0.591653, pem=-1.9595277,
 similarity=-2.9101756]

11: 20994545 Józefowo, Gmina Przedecz
 [winner: 21147090 Józefowo, Szamotuły County, pem_pos=7,
 final_scores=-0.4680471, final_local_scores=-0.44505438,
 pem=-3.27258, similarity=8.605201]
 RECALL MISS: 20994545 Józefowo, Gmina Przedecz]

12: 20994551 Katarzyna, Greater Poland Voivodeship
 [winner: 18325773 Katarzyna, Kuyavian-
 Pomeranian Voivodeship, pem_pos=1, final_scores=-0.42936814,

```

final_local_scores=-0.41083634, pem=-1.4764667,
similarity=-0.14120746|
    gt: 20994551 Katarzyna,_Greater_Poland_Voivodeship,
pem_pos=2, final_scores=-0.47250587,
final_local_scores=-0.47558552, pem=-1.9425563,
similarity=0.46121565 ]
13: 20994559 Kłokoczyn,_Greater_Poland_Voivodeship
    [winner: 19017125 Kłokoczyn,_Lesser_Poland_Voivodeship,
pem_pos=1, final_scores=-0.14179704,
final_local_scores=-0.13603173, pem=-0.71667767,
similarity=3.9276595|
    gt: 20994559 Kłokoczyn,_Greater_Poland_Voivodeship,
pem_pos=0, final_scores=-0.24469063,
final_local_scores=-0.2593531, pem=-0.6701577,
similarity=-0.005479414 ]
14: 20994567 Łączewna,_Greater_Poland_Voivodeship
    [winner: 18349477 Łączewna,_Kuyavian-
Pomeranian_Voivodeship, pem_pos=0, final_scores=-0.26335177,
final_local_scores=-0.2125659, pem=-0.47692412,
similarity=0.33991015|
    gt: 20994567 Łączewna,_Greater_Poland_Voivodeship,
pem_pos=1, final_scores=-0.2718066,
final_local_scores=-0.2499697, pem=-0.9694005,
similarity=1.9018689 ]
15: 20994572 Lipiny,_Gmina_Przedecz
    [winner: 20994308 Lipiny,_Gmina_Osiek_Mały, pem_pos=13,
final_scores=-0.49140936, final_local_scores=-0.52815646,
pem=-2.7264693, similarity=3.1554344|
    gt: 20994572 Lipiny,_Gmina_Przedecz, pem_pos=19,
final_scores=-0.60342586, final_local_scores=-0.5778127,
pem=-2.8143551, similarity=2.151531 ]
16: 21147708 Mieczysławowo,_Greater_Poland_Voivodeship
    [winner: 18378468 Mieczysławowo,_Kuyavian-
Pomeranian_Voivodeship, pem_pos=0, final_scores=-0.2929567,
final_local_scores=-0.25685686, pem=-0.542865,
similarity=-0.6232239|
    gt: 21147708 Mieczysławowo,_Greater_Poland_Voivodeship,
pem_pos=1, final_scores=-0.389508,
final_local_scores=-0.37024587, pem=-0.8700779,
similarity=-2.227543 ]
17: 20994583 Nowa_Wieś_Wielka,_Greater_Poland_Voivodeship
    [winner: 21822515 Nowa_Wieś_Wielka,_Lidzbark_County,
pem_pos=0, final_scores=-0.1301803,
final_local_scores=-0.078719676, pem=-1.0090001,
similarity=7.22766|
    gt: 20994583
Nowa_Wieś_Wielka,_Greater_Poland_Voivodeship, pem_pos=4,
final_scores=-0.4213845, final_local_scores=-0.45164913,
pem=-2.571918, similarity=4.5980916 ]
18: 20994592 Rogóżno,_Greater_Poland_Voivodeship
    [winner: 18772761 Rogóżno,_Łask_County, pem_pos=2,
final_scores=-0.4529416, final_local_scores=-0.45658526,

```

pem=-2.0406556, similarity=1.5617003|
gt: 20994592 Rogóźno, Greater_Poland_Voivodeship,
pem_pos=6, final_scores=-0.617651,
final_local_scores=-0.6765986, pem=-2.7782545,
similarity=-0.9927597]
19: 20994598 Rybno, Koło County
[winner: 18327386 Rybno, Kuyavian-Pomeranian_Voivodeship,
pem_pos=5, final_scores=-0.4840021,
final_local_scores=-0.48190373, pem=-1.89712,
similarity=0.025577605|
gt: 20994598 Rybno, Koło County, pem_pos=11,
final_scores=-0.53237253, final_local_scores=-0.58095,
pem=-2.2008023, similarity=-1.278688]
20: 20994607 Zalesie, Koło County
[winner: 18325532 Zalesie, Nakło County, pem_pos=8,
final_scores=-0.65768707, final_local_scores=-0.6647704,
pem=-3.702426, similarity=4.386006|
RECALL MISS: 20994607 Zalesie, Koło County]
21: 20994614 Żarowo, Greater_Poland_Voivodeship
[winner: 21999448 Żarowo, West_Pomeranian_Voivodeship,
pem_pos=0, final_scores=-0.28647888,
final_local_scores=-0.28657368, pem=-0.30796677,
similarity=-2.7874463|
gt: 20994614 Żarowo, Greater_Poland_Voivodeship,
pem_pos=1, final_scores=-0.47792256,
final_local_scores=-0.53218377, pem=-1.3277981,
similarity=-4.5709386]
22: 20994619 Zbijewo-Kolonia
[winner: 20994619 Zbijewo-Kolonia, pem_pos=0,
final_scores=-0.1406438, final_local_scores=-0.12914877,
pem=0.0, similarity=0.23565634|
gt: 20994619 Zbijewo-Kolonia, pem_pos=0,
final_scores=-0.1406438, final_local_scores=-0.12914877,
pem=0.0, similarity=0.23565634]
23: 20994621 Zbijewo-Parcele_A
[winner: 20994621 Zbijewo-Parcele_A, pem_pos=0,
final_scores=-0.2422672, final_local_scores=-0.15873149,
pem=0.0, similarity=-0.6471513|
gt: 20994621 Zbijewo-Parcele_A, pem_pos=0,
final_scores=-0.2422672, final_local_scores=-0.15873149,
pem=0.0, similarity=-0.6471513]
24: 15115154 Gmina_Babiak
[winner: 20989645 Babiak, Koło County, pem_pos=0,
final_scores=-0.19998454, final_local_scores=-0.20875466,
pem=-0.5682184, similarity=0.9501179|
gt: 15115154 Gmina_Babiak, pem_pos=1,
final_scores=-0.23475751, final_local_scores=-0.24910802,
pem=-0.7116439, similarity=0.52586305]
25: 14921091 Gmina_Chodecz
[winner: 14921091 Gmina_Chodecz, pem_pos=1,
final_scores=-0.19569562, final_local_scores=-0.17962678,
pem=-0.3678916, similarity=0.72994334|

gt: 14921091 Gmina_Chodecz, pem_pos=1,
 final_scores=-0.19569562, final_local_scores=-0.17962678,
 pem=-0.3678916, similarity=0.72994334]
 26: 15115177 Gmina_Chodów
 [winner: 15115177 Gmina_Chodów, pem_pos=1,
 final_scores=-0.50101995, final_local_scores=-0.5467084,
 pem=-1.4408985, similarity=-4.3893256]
 gt: 15115177 Gmina_Chodów, pem_pos=1,
 final_scores=-0.50101995, final_local_scores=-0.5467084,
 pem=-1.4408985, similarity=-4.3893256]
 27: 14930378 Gmina_Dąbrowice
 [winner: 18142846 Dąbrowice, Bydgoszcz County, pem_pos=0,
 final_scores=-0.41087696, final_local_scores=-0.4090067,
 pem=-1.6821973, similarity=1.0321879]
 gt: 14930378 Gmina_Dąbrowice, pem_pos=2,
 final_scores=-0.6016954, final_local_scores=-0.6039487,
 pem=-2.1559815, similarity=-2.208759]
 28: 14921160 Gmina_Izbica_Kujawska
 [winner: 14921160 Gmina_Izbica_Kujawska, pem_pos=1,
 final_scores=-0.05389814, final_local_scores=0.033207238,
 pem=-0.41459754, similarity=7.335329]
 gt: 14921160 Gmina_Izbica_Kujawska, pem_pos=1,
 final_scores=-0.05389814, final_local_scores=0.033207238,
 pem=-0.41459754, similarity=7.335329]
 29: 15115247 Gmina_Kłodawa, Greater_Poland_Voivodeship
 [winner: 7339057 Kłodawa, pem_pos=0,
 final_scores=-0.4557629, final_local_scores=-0.44220474,
 pem=-0.53251386, similarity=-6.2106657]
 gt: 15115247 Gmina_Kłodawa, Greater_Poland_Voivodeship,
 pem_pos=2, final_scores=-0.7096542,
 final_local_scores=-0.76019573, pem=-2.5447466,
 similarity=-4.757316]
 1: 41498 Greater_Poland_Voivodeship
 [pem_pos=0, final_scores=0.0648496,
 final_local_scores=0.07735142, pem=0.0, similarity=6.3980336]
 2: 22936 Poland
 [pem_pos=0, final_scores=0.34281474,
 final_local_scores=0.38497216, pem=-0.15333635,
 similarity=16.411917]
 3: 390527 Koło
 [pem_pos=0, final_scores=0.09151624,
 final_local_scores=0.06648275, pem=-0.314798,
 similarity=7.7856116]
 4: 23723 Poznań
 [pem_pos=0, final_scores=0.32362038,
 final_local_scores=0.3332282, pem=-0.056552384,
 similarity=14.341448]

Sample: b'Katrine_Lunde_Haraldsen'

[1 Katrine Lunde 1] -LRB- born 30 March 1980 in [1 1 Kristia
 nsand 1] -RRB- is a [2 2 Norwegian 2] team handball
 goalkeeper , playing for the [2 Hungarian 2] club [3 3 Győri A

udi ET0 KC 3] . She is the twin sister of fellow handball player [4 4 Kristine Lunde-Borgersen 4] . Her sporting achievements include gold medals with the [3 1 Norwegian 3] national team at the 2 Olympic Games , World Championships and 3 European Championships , as well as club victories in the [4 4 EHF Champions League 4] , and national championships in both [5 Denmark 5] and [6 Hungary 6] . She started her club career in , and later played for [7 5 Kristiansand 7] , 6 Våg and [5 5 Aalborg DH 5] . She won the [8 7 Champions League 8] with [9 Viborg 9] in and . In 2010 she signed with [6 6 Hungarian 6] club [7 7 Győri Audi ET0 KC 7] . [10 Lunde 10] made her debut on the [11 8 Norwegian 11] national team in 2002 , and has played 135 matches and scored 2 goals . She is a four-time [12 European 12] champion -LRB- 9 2004 , 10 2006 , 11 2008 , and 1 2 2010 -RRB- . She is World champion from 13 2011 , received a silver medal at the 14 2007 World Women 's Handball Championship , and a bronze medal at the 15 2009 World Championship . In August 2008 , together with her twin sister , she won the gold medal at the 16 2008 [13 Summer Olympics 13] in [8 8 Beijing 8] . She was selected into the tournament 's All-Star Team , and was also overall top goalkeeper with a 42 % save rate . [14 Lunde 14] was again named goalkeeper of the All-Star Team in the 17 2008 [15 European Championship 15]

1: 178143 Kristiansand

2: 21241 Norway

3: 20981736 Győri_Audi_ET0_KC

4: 16367027 Kristine_Lunde-Borgersen

5: 20763538 Aalborg_DH

6: 13275 Hungary

7: 20981736 Győri_Audi_ET0_KC

8: 18603746 Beijing

1: 8145203 Norway_women's_national_handball_team

[winner: 21241 Norway, pem_pos=0, final_scores=0.5760649, final_local_scores=0.5880959, pem=-0.605967, similarity=24.93501]

RECALL MISS: 8145203

Norway_women's_national_handball_team]

[winner: 35434210 Italy_national_athletics_team, pem_pos=6, final_scores=-1.1955812, final_local_scores=-1.2234697, pem=-3.4006476, similarity=-13.927814]

RECALL MISS: 8145203

Norway_women's_national_handball_team]

2: 1471137 Handball_at_the_Summer_Olympics

[winner: 22576 Olympic_Games, pem_pos=0, final_scores=-0.19019821, final_local_scores=-0.22489946, pem=-0.5587218, similarity=0.41668049]

RECALL MISS: 1471137 Handball_at_the_Summer_Olympics]

3: 159071 European_Women's_Handball_Championship

[winner: 1825838 European_Athletics_Championships, pem_pos=0, final_scores=-0.16027062,

final_local_scores=-0.17556357, pem=-2.1033094,
 similarity=10.28867|
 RECALL MISS: 159071
 European_Women's_Handball_Championship]
 4: 6951794 EHF_Women's_Champions_League
 [winner: 5140101 EHF_Champions_League, pem_pos=0,
 final_scores=0.21675824, final_local_scores=0.23223421,
 pem=0.0, similarity=11.020045|
 gt: 6951794 EHF_Women's_Champions_League, pem_pos=1,
 final_scores=0.0843467, final_local_scores=0.08690861,
 pem=-0.7240753, similarity=10.620873]
 5: 7915806 Kristiansands_IF
 [winner: 178143 Kristiansand, pem_pos=0,
 final_scores=0.0995671, final_local_scores=0.10482876,
 pem=0.0, similarity=7.2180123|
 gt: 7915806 Kristiansands_IF, pem_pos=3,
 final_scores=-0.30875754, final_local_scores=-0.2629883,
 pem=-1.3844901, similarity=3.7706904]
 6: 27906176 IK_Våg
 [winner: 27906176 IK_Våg, pem_pos=0,
 final_scores=-0.004916236, final_local_scores=0.04052484,
 pem=0.0, similarity=5.299055|
 gt: 27906176 IK_Våg, pem_pos=0,
 final_scores=-0.004916236, final_local_scores=0.04052484,
 pem=0.0, similarity=5.299055]
 7: 6951794 EHF_Women's_Champions_League
 [winner: 44220 UEFA_Champions_League, pem_pos=0,
 final_scores=0.08727674, final_local_scores=0.0821998,
 pem=-0.7442114, similarity=10.589856|
 gt: 6951794 EHF_Women's_Champions_League, pem_pos=17,
 final_scores=-0.42748445, final_local_scores=-0.48293006,
 pem=-3.4099705, similarity=8.222067]
 8: 8145203 Norway_women's_national_handball_team
 [winner: 21241 Norway, pem_pos=0,
 final_scores=0.53745806, final_local_scores=0.5459238,
 pem=-0.605967, similarity=23.676512|
 RECALL MISS: 8145203
 Norway_women's_national_handball_team]
 [winner: 35434210 Italy_national_athletics_team,
 pem_pos=6, final_scores=-1.240597,
 final_local_scores=-1.2726426, pem=-3.4006476,
 similarity=-15.395229|
 RECALL MISS: 8145203
 Norway_women's_national_handball_team]
 9: 2963558 2004_European_Women's_Handball_Championship
 [winner: 35524 2004, pem_pos=0, final_scores=-0.46927667,
 final_local_scores=-0.30362034, pem=-0.4915334,
 similarity=-2.2978885|
 RECALL MISS: 2963558
 2004_European_Women's_Handball_Championship]
 10: 8127818 2006_European_Women's_Handball_Championship
 [winner: 36164 2006, pem_pos=0, final_scores=-0.32235196,

final_local_scores=-0.12883446, pem=-0.3060146,
similarity=1.909191|
RECALL MISS: 8127818
2006_European_Women's_Handball_Championship]
11: 17958777 2008_European_Women's_Handball_Championship
[winner: 35825 2008, pem_pos=0, final_scores=-0.48290473,
final_local_scores=-0.36592886, pem=-0.29331458,
similarity=-5.235244|
RECALL MISS: 17958777
2008_European_Women's_Handball_Championship]
12: 22108228 2010_European_Women's_Handball_Championship
[winner: 43226 2010, pem_pos=0, final_scores=-1.0383527,
final_local_scores=-0.95318747, pem=-1.6731367,
similarity=-15.256524|
RECALL MISS: 22108228
2010_European_Women's_Handball_Championship]
13: 25559959 2011_World_Women's_Handball_Championship
[winner: 36225 2011, pem_pos=0, final_scores=-0.9123226,
final_local_scores=-0.8020156, pem=-1.4232541,
similarity=-12.1041565|
RECALL MISS: 25559959
2011_World_Women's_Handball_Championship]
14: 7369096 2007_World_Women's_Handball_Championship
[winner: 7331700 2009_World_Men's_Handball_Championship,
pem_pos=0, final_scores=-0.21609992,
final_local_scores=-0.23635048, pem=0.0,
similarity=-2.9634564|
RECALL MISS: 7369096
2007_World_Women's_Handball_Championship]
[winner: 36165 2007, pem_pos=0, final_scores=-0.93195605,
final_local_scores=-0.8722172, pem=-0.24805899,
similarity=-20.590004|
RECALL MISS: 7369096
2007_World_Women's_Handball_Championship]
[winner: 11242642 World_(Bee_Gees_song), pem_pos=6,
final_scores=-1.1448537, final_local_scores=-1.0814669,
pem=-3.2688804, similarity=-10.406737|
RECALL MISS: 7369096
2007_World_Women's_Handball_Championship]
[winner: 24359424 Women_(album), pem_pos=10,
final_scores=-1.3958168, final_local_scores=-1.3428984,
pem=-3.3705711, similarity=-17.655363|
RECALL MISS: 7369096
2007_World_Women's_Handball_Championship]
[winner: 1415027 David_Berman_(musician), pem_pos=5,
final_scores=-2.2037704, final_local_scores=-2.239802,
pem=-3.0012724, similarity=-46.429058|
RECALL MISS: 7369096
2007_World_Women's_Handball_Championship]
15: 7369130 2009_World_Women's_Handball_Championship
[winner: 1747960 FIVB_Volleyball_World_Championship,
pem_pos=2, final_scores=-0.032736674,

```

final_local_scores=-0.0044688284, pem=-2.334111,
similarity=16.649612|
    RECALL MISS: 7369130
2009_World_Women's_Handball_Championship]
    [winner: 35983 2009, pem_pos=0, final_scores=-0.9981439,
final_local_scores=-0.9148245, pem=-0.9296576,
similarity=-18.154852|
    RECALL MISS: 7369130
2009_World_Women's_Handball_Championship]
16: 77745 2008_Summer_Olympics
    [winner: 27566 Summer_Olympic_Games, pem_pos=0,
final_scores=0.17130284, final_local_scores=0.18149698,
pem=-0.4112264, similarity=11.742257|
    gt: 77745 2008_Summer_Olympics, pem_pos=2,
final_scores=-0.57364315, final_local_scores=-0.61491185,
pem=-3.4497817, similarity=4.4999657 ]
    [winner: 35825 2008, pem_pos=0, final_scores=-0.91748613,
final_local_scores=-0.84064305, pem=-0.29331458,
similarity=-19.401663|
    gt: 77745 2008_Summer_Olympics, pem_pos=12,
final_scores=-2.12466, final_local_scores=-2.3166845,
pem=-5.467841, similarity=-35.3098 ]
17: 17958777 2008_European_Women's_Handball_Championship
    [winner: 249510 UEFA_European_Championship, pem_pos=0,
final_scores=0.17838262, final_local_scores=0.17227829,
pem=-1.0168781, similarity=14.760782|
    RECALL MISS: 17958777
2008_European_Women's_Handball_Championship]
    [winner: 35825 2008, pem_pos=0, final_scores=-0.82005656,
final_local_scores=-0.73421603, pem=-0.29331458,
similarity=-16.225668|
    RECALL MISS: 17958777
2008_European_Women's_Handball_Championship]
1: 178143 Kristiansand
    [pem_pos=0, final_scores=0.40985513,
final_local_scores=0.4437713, pem=0.0, similarity=17.332733]
2: 21241 Norway
    [pem_pos=0, final_scores=0.5765663,
final_local_scores=0.5886436, pem=-0.605967,
similarity=24.951355]
3: 20981736 Győri_Audi_ETO_KC
    [pem_pos=0, final_scores=0.27787936,
final_local_scores=0.30718574, pem=0.0, similarity=13.256748]
4: 16367027 Kristine_Lunde-Borgersen
    [pem_pos=0, final_scores=0.64513046,
final_local_scores=0.692185, pem=0.0, similarity=24.745893]
5: 20763538 Aalborg_DH
    [pem_pos=0, final_scores=0.14842375,
final_local_scores=0.17376217, pem=0.0, similarity=9.275123]
6: 13275 Hungary
    [pem_pos=1, final_scores=0.6473603,
final_local_scores=0.67220354, pem=-1.0319586,

```

similarity=29.761559]
7: 20981736 Győri Audi ETO KC
[pem_pos=0, final_scores=0.29391688,
final_local_scores=0.32470426, pem=0.0, similarity=13.779535]
8: 18603746 Beijing
[pem_pos=0, final_scores=0.2858191,
final_local_scores=0.33105776, pem=-0.042161036,
similarity=14.198415]
1: 15730717 Katrine Lunde Haraldsen
[pem_pos=0, final_scores=0.06089668,
final_local_scores=0.06951304, pem=0.0, similarity=6.1641207]
2: 13275 Hungary
[pem_pos=1, final_scores=0.57978475,
final_local_scores=0.59838754, pem=-1.0319586,
similarity=27.558743]
3: 21241 Norway
[pem_pos=0, final_scores=0.5760649,
final_local_scores=0.5880959, pem=-0.605967,
similarity=24.93501]
4: 5140101 EHF_Champions_League
[pem_pos=0, final_scores=0.21675824,
final_local_scores=0.23223421, pem=0.0, similarity=11.020045]
5: 76972 Denmark
[pem_pos=0, final_scores=0.54140526,
final_local_scores=0.54856247, pem=-0.26762357,
similarity=21.915289]
6: 13275 Hungary
[pem_pos=0, final_scores=0.55858165,
final_local_scores=0.57522637, pem=-0.32577527,
similarity=23.027233]
7: 178143 Kristiansand
[pem_pos=0, final_scores=0.0995671,
final_local_scores=0.10482876, pem=0.0, similarity=7.2180123]
8: 44220 UEFA_Champions_League
[pem_pos=0, final_scores=0.08727674,
final_local_scores=0.0821998, pem=-0.7442114,
similarity=10.589856]
9: 2716449 Viborg_FF
[pem_pos=1, final_scores=0.2460901,
final_local_scores=0.25526786, pem=-1.0682886,
similarity=17.516935]
10: 16367027 Kristine Lunde-Borgersen
[pem_pos=0, final_scores=0.33935958,
final_local_scores=0.36275846, pem=0.0, similarity=14.915149]
11: 21241 Norway
[pem_pos=0, final_scores=0.53745806,
final_local_scores=0.5459238, pem=-0.605967,
similarity=23.676512]
12: 9239 Europe
[pem_pos=0, final_scores=0.34940684,
final_local_scores=0.42011052, pem=-0.25557294,
similarity=18.016495]

13: 27566 Summer_Olympic_Games
 [pem_pos=0, final_scores=0.17130284,
 final_local_scores=0.18149698, pem=-0.4112264,
 similarity=11.742257]
 14: 16367027 Kristine_Lunde-Borgersen
 [pem_pos=0, final_scores=0.466074,
 final_local_scores=0.51253945, pem=0.0, similarity=19.384912]
 15: 249510 UEFA_European_Championship
 [pem_pos=0, final_scores=0.17838262,
 final_local_scores=0.17227829, pem=-1.0168781,
 similarity=14.760782]

Sample: b'Jutta_Nardenbach'

[1 Jutta Nardenbach 1] -LRB- born 13 August 1968 -RRB- is a former [2 German 2] international 1 footballer . She plays the position of . Currently [3 Nardenbach 3] is player in the first team and coach of the youth teams at third tier FFC Montabaur . [4 Nardenbach 4] played in the for [1 1 TuS Ahrbach 1] , , [2 2 TSV Siegen 2] , [3 3 FC Rumeln-Kaldenhausen 3] , [4 4 SC 07 Bad Neuenahr 4] , 2 1 . [5 FFC Frankfurt 5] , and . She retired from the in 2004 , but is still playing for 1 . FFC Montabaur in the [6 Regionalliga 6] , where she is also part of the coaching staff . [7 Nardenbach 7] had appearances in 5 finals and won in , , each with [5 5 TSV Siegen 5] . In 1993 she also won the [8 DFB-Pokal 8] with [9 TSV Siegen 9] . In her time at [6 6 FFC Frankfurt 6] she won back to back 3 Doubles in 2001 and 2002 . Also in 2002 she won the with [10 Frankfurt 10] . [11 Jutta Nardenbach 11] had 59 appearances for [12 4 Germany 12] , scoring 4 times . Her first cap was against the [13 5 Netherlands 13] on 19 November 1986 . Her last game was against [14 6 Brazil 14] on 25 July 1996 . Her greatest successes were the 7 European Championships in 8 1989 and . At the [15 9 FIFA World Cup 15] in 10 1991

1: 18646898 TuS_Ahrbach
 2: 2731155 Sportfreunde_Siegen
 3: 14299116 FCR_2001_Duisburg
 4: 18731201 SC_07_Bad_Neuenahr
 5: 2731155 Sportfreunde_Siegen
 6: 7125149 1._FFC_Frankfurt

1: 10568 Association_football
 [winner: 16087792 DDR-Liga, pem_pos=6,
 final_scores=-1.5980904, final_local_scores=-1.7357845,
 pem=-2.302466, similarity=-35.18839]
 gt: 10568 Association_football, pem_pos=0,
 final_scores=-1.6653359, final_local_scores=-1.8260806,
 pem=0.0, similarity=-50.40417]
 2: 7125149 1._FFC_Frankfurt
 [winner: 7125149 1._FFC_Frankfurt, pem_pos=0,
 final_scores=0.38447344, final_local_scores=0.37131602,
 pem=0.0, similarity=15.170523]
 gt: 7125149 1._FFC_Frankfurt, pem_pos=0,

final_scores=0.38447344, final_local_scores=0.37131602,
 pem=0.0, similarity=15.170523]
 [winner: 1175 April_1, pem_pos=4,
 final_scores=-1.0059882, final_local_scores=-0.9257974,
 pem=-3.8010993, similarity=-2.8669598]
 RECALL MISS: 7125149 1._FFC_Frankfurt]
 [winner: 4721653 Zyphe_language, pem_pos=12,
 final_scores=-1.7535734, final_local_scores=-1.6720598,
 pem=-4.4188404, similarity=-21.777546]
 RECALL MISS: 7125149 1._FFC_Frankfurt]
 3: 88184 Double_(association_football)
 [winner: 18666965 Tennis_at_the_2008_Summer_Olympics_-_Women's_doubles, pem_pos=2, final_scores=-0.79563516,
 final_local_scores=-0.75292045, pem=-2.813311,
 similarity=-3.079713]
 RECALL MISS: 88184 Double_(association_football)]
 4: 662281 Germany_women's_national_football_team
 [winner: 250204 Germany_national_football_team,
 pem_pos=1, final_scores=0.22926979,
 final_local_scores=0.18185085, pem=-3.3613694,
 similarity=27.796154]
 gt: 662281 Germany_women's_national_football_team,
 pem_pos=4, final_scores=-0.28609037,
 final_local_scores=-0.38043994, pem=-4.8869205,
 similarity=19.31246]
 5: 5650619 Netherlands_women's_national_football_team
 [winner: 21148 Netherlands, pem_pos=0,
 final_scores=0.17220978, final_local_scores=0.20745087,
 pem=-0.24940795, similarity=11.63678]
 gt: 5650619 Netherlands_women's_national_football_team,
 pem_pos=9, final_scores=-0.38380367,
 final_local_scores=-0.4447642, pem=-4.4903617,
 similarity=15.236347]
 6: 2343221 Brazil_women's_national_football_team
 [winner: 149286 Brazil_national_football_team, pem_pos=1,
 final_scores=0.07853009, final_local_scores=0.054887056,
 pem=-2.5928838, similarity=19.828157]
 gt: 2343221 Brazil_women's_national_football_team,
 pem_pos=4, final_scores=-0.38002402,
 final_local_scores=-0.4326464, pem=-4.6023197,
 similarity=16.206812]
 7: 1985292 UEFA_Women's_Championship
 [winner: 1825838 European_Athletics_Championships,
 pem_pos=0, final_scores=-0.14127177,
 final_local_scores=-0.12876615, pem=-2.1033094,
 similarity=11.685198]
 RECALL MISS: 1985292 UEFA_Women's_Championship]
 8: 5785348 1989_European_Competition_for_Women's_Football
 [winner: 34847 1989, pem_pos=0, final_scores=-0.72163665,
 final_local_scores=-0.61042345, pem=-0.6980697,
 similarity=-10.330329]
 RECALL MISS: 5785348

1989_European_Competition_for_Women's_Football]

9: 244862 FIFA_Women's_World_Cup
 [winner: 11370 FIFA_World_Cup, pem_pos=0,
 final_scores=0.16343747, final_local_scores=0.15492576,
 pem=-0.06810418, similarity=9.083367]
 gt: 244862 FIFA_Women's_World_Cup, pem_pos=12,
 final_scores=-0.33185285, final_local_scores=-0.36651126,
 pem=-3.2328525, similarity=10.733041]

10: 335539 1991_FIFA_Women's_World_Cup
 [winner: 34703 1991, pem_pos=0, final_scores=-0.7789259,
 final_local_scores=-0.6637268, pem=-0.6661314,
 similarity=-12.094693]
 gt: 335539 1991_FIFA_Women's_World_Cup, pem_pos=16,
 final_scores=-1.5441225, final_local_scores=-1.693583,
 pem=-4.869703, similarity=-19.96798]

1: 18646898 TuS_Ahrbach
 [pem_pos=0, final_scores=0.15882914,
 final_local_scores=0.20666188, pem=0.0, similarity=10.256916]

2: 2731155 Sportfreunde_Siegen
 [pem_pos=0, final_scores=0.37176245,
 final_local_scores=0.33926633, pem=0.0, similarity=14.214097]

3: 14299116 FCR_2001_Duisburg
 [pem_pos=0, final_scores=0.4099185,
 final_local_scores=0.39337778, pem=0.0, similarity=15.82889]

4: 18731201 SC_07_Bad_Neuenahr
 [pem_pos=0, final_scores=0.27759892,
 final_local_scores=0.27795216, pem=0.0, similarity=12.38436]

5: 2731155 Sportfreunde_Siegen
 [pem_pos=0, final_scores=0.28549236,
 final_local_scores=0.24502936, pem=0.0, similarity=11.401877]

6: 7125149 1._FFC_Frankfurt
 [pem_pos=0, final_scores=0.2793284,
 final_local_scores=0.25646108, pem=0.0, similarity=11.743023]

1: 18727366 Jutta_Nardenbach
 [pem_pos=0, final_scores=0.5208048,
 final_local_scores=0.5332764, pem=0.0, similarity=20.003744]

2: 11867 Germany
 [pem_pos=1, final_scores=0.687952,
 final_local_scores=0.75753146, pem=-1.0465823,
 similarity=32.38744]

3: 18727366 Jutta_Nardenbach
 [pem_pos=0, final_scores=0.4421895,
 final_local_scores=0.44740105, pem=0.0, similarity=17.441053]

4: 18727366 Jutta_Nardenbach
 [pem_pos=0, final_scores=0.39140445,
 final_local_scores=0.3919261, pem=0.0, similarity=15.785569]

5: 7125149 1._FFC_Frankfurt
 [pem_pos=0, final_scores=0.38447344,
 final_local_scores=0.37131602, pem=0.0, similarity=15.170523]

6: 3476544 Regionalliga
 [pem_pos=0, final_scores=0.21445449,
 final_local_scores=0.19612846, pem=-0.0486752,

similarity=10.20728]
 7: 18727366 Jutta_Nardenbach
 [pem_pos=0, final_scores=0.45358008,
 final_local_scores=0.45984352, pem=0.0, similarity=17.81236]
 8: 675126 DFB-Pokal
 [pem_pos=0, final_scores=0.25060457,
 final_local_scores=0.2127589, pem=0.0, similarity=10.438863]
 9: 2731155 Sportfreunde_Siegen
 [pem_pos=0, final_scores=0.2734648,
 final_local_scores=0.23189104, pem=0.0, similarity=11.009804]
 10: 862147 Eintracht_Frankfurt
 [pem_pos=10, final_scores=0.34387594,
 final_local_scores=0.30248433, pem=-3.3683815,
 similarity=31.434233]
 11: 18727366 Jutta_Nardenbach
 [pem_pos=0, final_scores=0.8690552,
 final_local_scores=0.91368705, pem=0.0, similarity=31.355957]
 12: 250204 Germany_national_football_team
 [pem_pos=1, final_scores=0.22926979,
 final_local_scores=0.18185085, pem=-3.3613694,
 similarity=27.796154]
 13: 21148 Netherlands
 [pem_pos=0, final_scores=0.17220978,
 final_local_scores=0.20745087, pem=-0.24940795,
 similarity=11.63678]
 14: 149286 Brazil_national_football_team
 [pem_pos=1, final_scores=0.07853009,
 final_local_scores=0.054887056, pem=-2.5928838,
 similarity=19.828157]
 15: 11370 FIFA_World_Cup
 [pem_pos=0, final_scores=0.16343747,
 final_local_scores=0.15492576, pem=-0.06810418,
 similarity=9.083367]

Sample: b'Frank_Mortimer'

Frank Mortimer -LRB- born 23 February 1932 in [1 1 Streethou
 se 1] -- died in [2 2 Wakefield 2] -RRB- was an [1 1 English 1
] professional 2 rugby league footballer of the 1950s and
 '60s , playing at representative level for 3 Great Britain ,
 and [2 4 Yorkshire 2] , and at club level for [3 3 Wakefield T
 rinity 3] , and 5 Keighley , as a , or , i.e. number 1 , or ,
 3 or 4 . Frank Mortimer played and scored three s in 6 Great B
 ritain 's 21-10 victory over [3 7 Australia 3] in the first
 1956 [4 8 Ashes 9 Test Match 4] at [5 10 Central Park 5] , [6
 Wigan 6] on 17 November 1956 , and scored three in 11 Great Br
 itain 's 9-22 defeat to [7 12 Australia 7] in the second 1956
 [8 13 Ashes 14 Test Match 8] at [9 15 Odsal Stadium 9] , [10 B
 radford 10] on 1 December 1956 . Frank Mortimer was selected
 for 16 Yorkshire County XIII whilst at [4 4 Wakefield Trinity
 4]

1: 1219258 Streethouse
 2: 152098 Wakefield

3: 1095973 Wakefield_Trinity_Wildcats
4: 1095973 Wakefield_Trinity_Wildcats
1: 18803164 English_people
[winner: 9316 England, pem_pos=1, final_scores=0.5907472, final_local_scores=0.6073504, pem=-1.3527626, similarity=29.570793]
gt: 18803164 English_people, pem_pos=2, final_scores=-0.35185492, final_local_scores=-0.39348012, pem=-2.2084396, similarity=4.357319]
2: 25735 Rugby_league
[winner: 19774201 Rugby_Virginia, pem_pos=7, final_scores=-0.93325424, final_local_scores=-0.8188386, pem=-2.956099, similarity=-4.2703385]
gt: 25735 Rugby_league, pem_pos=14, final_scores=-1.9247414, final_local_scores=-2.1288278, pem=-3.3002021, similarity=-41.491745]
[winner: 1424274 League_(unit), pem_pos=2, final_scores=-1.9155958, final_local_scores=-1.9807272, pem=-1.8401146, similarity=-45.012306]
gt: 25735 Rugby_league, pem_pos=5, final_scores=-2.0963929, final_local_scores=-2.316331, pem=-3.5378149, similarity=-45.795036]
3: 951941 Great_Britain_national_rugby_league_team
[winner: 951941 Great_Britain_national_rugby_league_team, pem_pos=2, final_scores=-0.114543706, final_local_scores=-0.16168293, pem=-2.6635387, similarity=13.749509]
gt: 951941 Great_Britain_national_rugby_league_team, pem_pos=2, final_scores=-0.114543706, final_local_scores=-0.16168293, pem=-2.6635387, similarity=13.749509]
4: 493717 Rugby_League_War_of_the_Roses
[winner: 869945 Yorkshire_County_Cricket_Club, pem_pos=1, final_scores=0.25341624, final_local_scores=0.2261723, pem=-1.7779728, similarity=20.508038]
gt: 493717 Rugby_League_War_of_the_Roses, pem_pos=3, final_scores=-0.3408857, final_local_scores=-0.40098065, pem=-3.1132855, similarity=9.054181]
5: 3550456 Keighley_Cougars
[winner: 3550456 Keighley_Cougars, pem_pos=1, final_scores=-0.014383644, final_local_scores=-0.05314389, pem=-1.779445, similarity=12.180692]
gt: 3550456 Keighley_Cougars, pem_pos=1, final_scores=-0.014383644, final_local_scores=-0.05314389, pem=-1.779445, similarity=12.180692]
6: 951941 Great_Britain_national_rugby_league_team
[winner: 951941 Great_Britain_national_rugby_league_team, pem_pos=2, final_scores=-0.08475898, final_local_scores=-0.12914765, pem=-2.6635387, similarity=14.720427]
gt: 951941 Great_Britain_national_rugby_league_team, pem_pos=2, final_scores=-0.08475898,

final_local_scores=-0.12914765, pem=-2.6635387,
 similarity=14.720427]
 7: 951978 Australia_national_rugby_league_team
 [winner: 374833 Australia_national_cricket_team,
 pem_pos=2, final_scores=0.5959906,
 final_local_scores=0.60095334, pem=-3.4989398,
 similarity=41.05114]
 gt: 951978 Australia_national_rugby_league_team,
 pem_pos=5, final_scores=-0.310826,
 final_local_scores=-0.3777231, pem=-4.3727627,
 similarity=16.597466]
 8: 701471 The_Ashes_(rugby_league)
 [winner: 1132 The_Ashes, pem_pos=0,
 final_scores=0.31436306, final_local_scores=0.3122319,
 pem=0.0, similarity=13.407335]
 RECALL MISS: 701471 The_Ashes_(rugby_league)]
 9: 16926380 Test_match_(rugby_league)
 [winner: 1132 The_Ashes, pem_pos=0,
 final_scores=0.31436306, final_local_scores=0.3122319,
 pem=0.0, similarity=13.407335]
 RECALL MISS: 16926380 Test_match_(rugby_league)]
 10: 5691366 Central_Park_(Wigan)
 [winner: 8222430 Wigan, pem_pos=0,
 final_scores=0.13045938, final_local_scores=0.07491812,
 pem=-0.30239522, similarity=7.969892]
 RECALL MISS: 5691366 Central_Park_(Wigan)]
 [winner: 37536 Central_Park, pem_pos=0,
 final_scores=0.05168952, final_local_scores=0.1194579,
 pem=-0.114469826, similarity=8.27708]
 gt: 5691366 Central_Park_(Wigan), pem_pos=1,
 final_scores=-0.24512877, final_local_scores=-0.2937223,
 pem=-2.7907526, similarity=10.500999]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7930098, final_local_scores=-1.7860482,
 pem=-4.6728287, similarity=-23.79796]
 RECALL MISS: 5691366 Central_Park_(Wigan)]
 11: 951941 Great_Britain_national_rugby_league_team
 [winner: 951941 Great_Britain_national_rugby_league_team,
 pem_pos=2, final_scores=-0.064839795,
 final_local_scores=-0.10738897, pem=-2.6635387,
 similarity=15.36975]
 gt: 951941 Great_Britain_national_rugby_league_team,
 pem_pos=2, final_scores=-0.064839795,
 final_local_scores=-0.10738897, pem=-2.6635387,
 similarity=15.36975]
 12: 951978 Australia_national_rugby_league_team
 [winner: 374833 Australia_national_cricket_team,
 pem_pos=1, final_scores=0.5528161,
 final_local_scores=0.55379176, pem=-3.4989398,
 similarity=39.64374]
 gt: 951978 Australia_national_rugby_league_team,
 pem_pos=4, final_scores=-0.34147358,

```

final_local_scores=-0.41120094, pem=-4.3727627,
similarity=15.59842 ]
13: 701471 The_Ashes_(rugby_league)
    [winner: 1132 The_Ashes, pem_pos=0,
final_scores=0.28485006, final_local_scores=0.2799935,
pem=0.0, similarity=12.445277]
    RECALL MISS: 701471 The_Ashes_(rugby_league)]
14: 16926380 Test_match_(rugby_league)
    [winner: 1132 The_Ashes, pem_pos=0,
final_scores=0.28485006, final_local_scores=0.2799935,
pem=0.0, similarity=12.445277]
    RECALL MISS: 16926380 Test_match_(rugby_league)]
15: 3411441 Odsal_Stadium
    [winner: 3411441 Odsal_Stadium, pem_pos=0,
final_scores=0.2604156, final_local_scores=0.24241638,
pem=0.0, similarity=11.323901]
    gt: 3411441 Odsal_Stadium, pem_pos=0,
final_scores=0.2604156, final_local_scores=0.24241638,
pem=0.0, similarity=11.323901 ]
    [winner: 23861381 Bradford, pem_pos=0,
final_scores=0.24496742, final_local_scores=0.20295614,
pem=-0.6938952, similarity=13.91984]
    RECALL MISS: 3411441 Odsal_Stadium]
    [winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8243873, final_local_scores=-1.8203235,
pem=-4.6728287, similarity=-24.820797]
    RECALL MISS: 3411441 Odsal_Stadium]
16: 493717 Rugby_League_War_of_the_Roses
    [winner: 493717 Rugby_League_War_of_the_Roses, pem_pos=0,
final_scores=-0.02994056, final_local_scores=-0.06103596,
pem=0.0, similarity=2.2682781]
    gt: 493717 Rugby_League_War_of_the_Roses, pem_pos=0,
final_scores=-0.02994056, final_local_scores=-0.06103596,
pem=0.0, similarity=2.2682781 ]
1: 1219258 Streethouse
    [pem_pos=0, final_scores=0.0476083,
final_local_scores=0.11963138, pem=0.0, similarity=7.659752]
2: 152098 Wakefield
    [pem_pos=0, final_scores=0.3467738,
final_local_scores=0.34050873, pem=-0.35212237,
similarity=16.166069]
3: 1095973 Wakefield_Trinity_Wildcats
    [pem_pos=0, final_scores=0.2113695,
final_local_scores=0.18711853, pem=0.0, similarity=9.673703]
4: 1095973 Wakefield_Trinity_Wildcats
    [pem_pos=0, final_scores=0.24939059,
final_local_scores=0.22865078, pem=0.0, similarity=10.913108]
1: 9316 England
    [pem_pos=1, final_scores=0.5907472,
final_local_scores=0.6073504, pem=-1.3527626,
similarity=29.570793]
2: 869945 Yorkshire_County_Cricket_Club

```

[pem_pos=1, final_scores=0.25341624,
 final_local_scores=0.2261723, pem=-1.7779728,
 similarity=20.508038]
 3: 374833 Australia_national_cricket_team
 [pem_pos=2, final_scores=0.5959906,
 final_local_scores=0.60095334, pem=-3.4989398,
 similarity=41.05114]
 4: 1132 The_Ashes
 [pem_pos=0, final_scores=0.31436306,
 final_local_scores=0.3122319, pem=0.0, similarity=13.407335]
 5: 37536 Central_Park
 [pem_pos=0, final_scores=0.05168952,
 final_local_scores=0.1194579, pem=-0.114469826,
 similarity=8.27708]
 6: 8222430 Wigan
 [pem_pos=0, final_scores=0.13045938,
 final_local_scores=0.07491812, pem=-0.30239522,
 similarity=7.969892]
 7: 374833 Australia_national_cricket_team
 [pem_pos=1, final_scores=0.5528161,
 final_local_scores=0.55379176, pem=-3.4989398,
 similarity=39.64374]
 8: 1132 The_Ashes
 [pem_pos=0, final_scores=0.28485006,
 final_local_scores=0.2799935, pem=0.0, similarity=12.445277]
 9: 3411441 Odsal_Stadium
 [pem_pos=0, final_scores=0.2604156,
 final_local_scores=0.24241638, pem=0.0, similarity=11.323901]
 10: 23861381 Bradford
 [pem_pos=0, final_scores=0.24496742,
 final_local_scores=0.20295614, pem=-0.6938952,
 similarity=13.91984]

Sample: b'Ron_Ryder'

Ronald `` Ron `` [1 Ryder 1] is an [2 1 English 2]
 professional 2 rugby league footballer of the 1950s playing
 at representative level for [1 1 Great Britain 1] , and [3 3 E
 ngland 3] , and at club level for [2 2 Warrington 2] , as a ,
 i.e. number 3 or 4 . [4 Ron Ryder 4] won a cap for [5 4 Engla
 nd 5] while at [6 Warrington 6] in 1952 against [7 Other
 Nationalities 7] , and won a cap for [3 3 Great Britain 3]
 while at [8 Warrington 8] in 1952 against Australia . [9 Ron
 Ryder 9] played , i.e. number 4 , in [4 4 Warrington 4] 's
 18-4 victory over [5 5 Halifax 5] in the [6 6 Challenge Cup 6]
 final replay at [7 7 Odsal Stadium 7] , [8 8 Bradford 8] on
 Wednesday 5 May 1954 in front of a 102,569 + crowd . [10 Ron
 Ryder 10] had not played in the previous 4-4 draw with [9 9 Ha
 lifax 9] in the [10 10 Challenge Cup 10] final at [11 11 Wemb
 ley Stadium 11] on Saturday 1 May 1954 , had played in this
 match . [11 Ron Ryder 11] played , i.e. number 3 in [12 12 Wa
 rrrington 12] 's 5-28 defeat to [13 13 Wigan 13] in the 5 1949
 -- 50 [14 14 Lancashire Cup 14] final at [12 6 Station Road 12]

] , Swinton

1: 951941 Great_Britain_national_rugby_league_team
2: 852392 Warrington_Wolves
3: 951941 Great_Britain_national_rugby_league_team
4: 852392 Warrington_Wolves
5: 1196374 Halifax_RLFC
6: 21050680 Challenge_Cup
7: 3411441 Odsal_Stadium
8: 23861381 Bradford
9: 1196374 Halifax_RLFC
10: 21050680 Challenge_Cup
11: 143673 Wembley_Stadium_(1923)
12: 852392 Warrington_Wolves
13: 1095029 Wigan_Warriors
14: 1010713 Rugby_league_county_cups

1: 18803164 English_people

[winner: 9316 England, pem_pos=1,
final_scores=0.61241907, final_local_scores=0.67732126,
pem=-1.3527626, similarity=31.658863]

gt: 18803164 English_people, pem_pos=2,
final_scores=-0.4594688, final_local_scores=-0.45936608,
pem=-2.2084396, similarity=2.391151]

2: 25735 Rugby_league

[winner: 19774201 Rugby_Virginia, pem_pos=7,
final_scores=-0.91161805, final_local_scores=-0.7966585,
pem=-2.956099, similarity=-3.608441]

gt: 25735 Rugby_league, pem_pos=14,
final_scores=-1.7980511, final_local_scores=-2.0056913,
pem=-3.3002021, similarity=-37.817104]

[winner: 21211 National_Football_League, pem_pos=1,
final_scores=-1.845645, final_local_scores=-1.9795215,
pem=-1.6728458, similarity=-45.885963]

gt: 25735 Rugby_league, pem_pos=5,
final_scores=-1.9067954, final_local_scores=-2.1244779,
pem=-3.5378149, similarity=-40.06975]

3: 990422 England_national_rugby_league_team

[winner: 9316 England, pem_pos=0,
final_scores=0.10185392, final_local_scores=0.119606435,
pem=-0.19053856, similarity=8.695186]

gt: 990422 England_national_rugby_league_team,
pem_pos=5, final_scores=-0.38163543,
final_local_scores=-0.49807808, pem=-4.237946,
similarity=12.272679]

4: 990422 England_national_rugby_league_team

[winner: 691024 England_national_rugby_union_team,
pem_pos=3, final_scores=0.05399172,
final_local_scores=0.009678721, pem=-3.7874734,
similarity=24.975414]

gt: 990422 England_national_rugby_league_team,
pem_pos=5, final_scores=-0.27348912,
final_local_scores=-0.3778988, pem=-4.237946,
similarity=15.859069]

5: 24239445 1949-50_Northern_Rugby_Football_League_season
[winner: 21288296 1949_Baltimore_Colts_season, pem_pos=0,
final_scores=-0.8171779, final_local_scores=-0.707169,
pem=0.0, similarity=-17.01362]
RECALL MISS: 24239445 1949-
50_Northern_Rugby_Football_League_season]
[winner: 1347230 Qaf_(sura), pem_pos=11,
final_scores=-1.0502212, final_local_scores=-0.80192775,
pem=-3.7918332, similarity=0.77916765]
RECALL MISS: 24239445 1949-
50_Northern_Rugby_Football_League_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3309852, final_local_scores=-1.2063036,
pem=0.0, similarity=-31.908794]
RECALL MISS: 24239445 1949-
50_Northern_Rugby_Football_League_season]
6: 6170801 Station_Road,_Swinton
[winner: 6170801 Station_Road,_Swinton, pem_pos=0,
final_scores=0.3152147, final_local_scores=0.29347417,
pem=-0.5753641, similarity=15.976486]
gt: 6170801 Station_Road,_Swinton, pem_pos=0,
final_scores=0.3152147, final_local_scores=0.29347417,
pem=-0.5753641, similarity=15.976486]
[winner: 2000369 Swinton_Lions, pem_pos=0,
final_scores=-0.23462532, final_local_scores=-0.34260064,
pem=-1.1196827, similarity=-0.04516564]
gt: 6170801 Station_Road,_Swinton, pem_pos=12,
final_scores=-0.5613747, final_local_scores=-0.6700925,
pem=-3.0827026, similarity=0.85703224]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8383449, final_local_scores=-1.8222027,
pem=-4.6728287, similarity=-24.876877]
RECALL MISS: 6170801 Station_Road,_Swinton]
1: 951941 Great_Britain_national_rugby_league_team
[pem_pos=2, final_scores=0.08418484,
final_local_scores=0.016661048, pem=-2.6635387,
similarity=19.07165]
2: 852392 Warrington_Wolves
[pem_pos=1, final_scores=0.22935306,
final_local_scores=0.14925244, pem=-1.6982093,
similarity=17.77883]
3: 951941 Great_Britain_national_rugby_league_team
[pem_pos=2, final_scores=0.12646262,
final_local_scores=0.062843084, pem=-2.6635387,
similarity=20.449814]
4: 852392 Warrington_Wolves
[pem_pos=1, final_scores=0.34512728,
final_local_scores=0.27571818, pem=-1.6982093,
similarity=21.55282]
5: 1196374 Halifax_RLFC
[pem_pos=3, final_scores=0.26910108,
final_local_scores=0.2044757, pem=-2.7129183,

similarity=24.944946]
6: 21050680 Challenge_Cup
[pem_pos=0, final_scores=0.26644558,
final_local_scores=0.20993358, pem=-0.38845974,
similarity=12.467054]
7: 3411441 Odsal_Stadium
[pem_pos=0, final_scores=0.30000007,
final_local_scores=0.26354048, pem=0.0, similarity=11.954287]
8: 23861381 Bradford
[pem_pos=0, final_scores=0.07317366,
final_local_scores=0.06267841, pem=-0.6938952,
similarity=9.733672]
9: 1196374 Halifax_RLFC
[pem_pos=3, final_scores=0.39532596,
final_local_scores=0.34235716, pem=-2.7129183,
similarity=29.059605]
10: 21050680 Challenge_Cup
[pem_pos=0, final_scores=0.2708699,
final_local_scores=0.21476653, pem=-0.38845974,
similarity=12.6112795]
11: 143673 Wembley_Stadium_(1923)
[pem_pos=1, final_scores=0.29925847,
final_local_scores=0.26746082, pem=-0.27926707,
similarity=13.589975]
12: 852392 Warrington_Wolves
[pem_pos=1, final_scores=0.37182158,
final_local_scores=0.30487767, pem=-1.6982093,
similarity=22.422997]
13: 1095029 Wigan_Warriors
[pem_pos=1, final_scores=0.45715958,
final_local_scores=0.4095444, pem=-1.8027086,
similarity=26.114744]
14: 1010713 Rugby_league_county_cups
[pem_pos=0, final_scores=0.24664567,
final_local_scores=0.22195178, pem=0.0, similarity=10.713197]
1: 23000631 Ron_Ryder
[pem_pos=0, final_scores=0.113781735,
final_local_scores=0.094947115, pem=0.0, similarity=6.9231243]
2: 9316 England
[pem_pos=1, final_scores=0.61241907,
final_local_scores=0.67732126, pem=-1.3527626,
similarity=31.658863]
3: 9316 England
[pem_pos=0, final_scores=0.10185392,
final_local_scores=0.119606435, pem=-0.19053856,
similarity=8.695186]
4: 23000631 Ron_Ryder
[pem_pos=0, final_scores=0.36848837,
final_local_scores=0.37317538, pem=0.0, similarity=15.226011]
5: 691024 England_national_rugby_union_team
[pem_pos=3, final_scores=0.05399172,
final_local_scores=0.009678721, pem=-3.7874734,

similarity=24.975414]
 6: 852392 Warrington_Wolves
 [pem_pos=1, final_scores=0.21682124,
 final_local_scores=0.13556334, pem=-1.6982093,
 similarity=17.37032]
 7: 18195524 Other_Nationalities_rugby_league_team
 [pem_pos=0, final_scores=0.2318946,
 final_local_scores=0.18876562, pem=0.0, similarity=9.722856]
 8: 852392 Warrington_Wolves
 [pem_pos=1, final_scores=0.26985592,
 final_local_scores=0.19349566, pem=-1.6982093,
 similarity=19.099136]
 9: 23000631 Ron_Ryder
 [pem_pos=0, final_scores=0.32843035,
 final_local_scores=0.3294181, pem=0.0, similarity=13.920206]
 10: 23000631 Ron_Ryder
 [pem_pos=0, final_scores=0.33017045,
 final_local_scores=0.3313189, pem=0.0, similarity=13.976931]
 11: 23000631 Ron_Ryder
 [pem_pos=0, final_scores=0.2766459,
 final_local_scores=0.27285147, pem=0.0, similarity=12.232145]
 12: 6170801 Station_Road,_Swinton
 [pem_pos=0, final_scores=0.3152147,
 final_local_scores=0.29347417, pem=-0.5753641,
 similarity=15.976486]

Sample: b"Mary_0'Connor"

_ -LRB- sportsperson -RRB- Mary 0'Connor -LRB- born 1977 in [1 1 Killeagh 1] , [2 2 County Cork 2] -RRB- is an [1 1 Irish 1] sportsperson . A 2 dual player , she plays both 3 Ladies ' Gaelic football and 4 camogie . Despite her undoubted success in [2 Gaelic Games 2] , 0'Connor considers her finest achievement to have been the comprehensive victory she had with The Calf Taggers in the 2012 Clontarf Rugby Club Tag Rugby League . The Calf Taggers posted 8 wins from their 9 games and were victorious in the finals played on 20 August 2012 . 0'Connor was highly decorated and named as MVP following her try-scoring record throughout the entire tournament . At the time she said `` This is the pinnacle of my career , and I 'd like to thank our coach John Kelly for giving us the courage to believe in ourselves and having us ready for today 's game . This is his trophy as well . '' As a footballer Murphy has played with her local clubs 5 Donoughmore and and has been a member of the [3 6 Cork 3] senior inter-county team since the 1994 . She has won Five [4 7 All-Ireland 4] titles . As a camogie player 0'Connor plays for her local club [3 3 Killeagh 3] and has been a member of the [5 Cork 5] senior inter-county team since 1996 . Since then she has won seven [6 8 All-Ireland 6] titles . 0'Connor is one of the most decorated players in the history of 9 Gaelic games . Winner of [7 10 All Ireland 7] camogie medals in 1997 -LRB- when her last point of the first half off her left

side , scored as she was in full flight 50 yards out , is regarded as one of the best scores in the history of camogie - RRB- , 1998 , 2002 , 2005 , 2006 , 2008 and 2009 and awards in and . She holds six [8 Senior All-Ireland 8] and [9 11 National League 9] camogie medals as well as county , provincial and 12 All Ireland Club medals with [4 4 Granagh-Ballingarry 4] whom she played with while studying and working in [10 Limerick 10] . She was the overall winner of the 96/103fm Rochestown Park Hotel award following her performances in the [11 All-Ireland 11] and 13 National League finals in 2006 . As well as her two awards she was nominated for further awards in 2004 , 2008 and 2009 , and received an unofficial Lynchpin award in . She captained [12 Cork 12] to a fifth successive [13 14 All-Ireland 13] football title in 2009 having won an in 15 Ladies Football

1: 10741716 Killeagh

2: 46709 County_Cork

3: 6157166 Killeagh_GAA

4: 16335590 Granagh-Ballingarry_GAA

1: 775859 Irish_people

[winner: 14560 Republic_of_Ireland, pem_pos=3, final_scores=0.4819154, final_local_scores=0.48697644, pem=-2.2576945, similarity=30.899754|

gt: 775859 Irish_people, pem_pos=0, final_scores=-0.38815856, final_local_scores=-0.4457566, pem=-1.1828228, similarity=-2.7801788]

2: 6093216 Dual_player

[winner: 27566809

UCI_Mountain_Bike_&_Trials_World_Championships_-_Women's_dual, pem_pos=14, final_scores=-0.913873, final_local_scores=-0.80614096, pem=-2.635303, similarity=-5.635953|

RECALL MISS: 6093216 Dual_player]

[winner: 3391589 AFL_Players_Association, pem_pos=7, final_scores=-1.6686277, final_local_scores=-1.7845399, pem=-2.5584118, similarity=-35.251476|

RECALL MISS: 6093216 Dual_player]

3: 2911812 Ladies'_Gaelic_football

[winner: 37688 Gaelic_football, pem_pos=6, final_scores=-0.12432431, final_local_scores=-0.2207573, pem=-2.5592473, similarity=11.419458|

RECALL MISS: 2911812 Ladies'_Gaelic_football]

[winner: 33183 Woman, pem_pos=0, final_scores=-0.29093084, final_local_scores=-0.22914827, pem=-1.1424404, similarity=3.46424|

gt: 2911812 Ladies'_Gaelic_football, pem_pos=17, final_scores=-0.62842035, final_local_scores=-0.72325504, pem=-3.5955985, similarity=2.0597634]

[winner: 2302349 Vedanga, pem_pos=16, final_scores=-2.058371, final_local_scores=-2.0235593, pem=-4.75359, similarity=-30.44657|

RECALL MISS: 2911812 Ladies'_Gaelic_football]

[winner: 143183 Football_(ball), pem_pos=5,
final_scores=-2.2484624, final_local_scores=-2.3718708,
pem=-4.263315, similarity=-43.50707|
RECALL MISS: 2911812 Ladies'_Gaelic_football]

4: 73109 Camogie
[winner: 73109 Camogie, pem_pos=0,
final_scores=-0.6064964, final_local_scores=-0.7225113,
pem=-0.0008609424, similarity=-17.466784|
gt: 73109 Camogie, pem_pos=0, final_scores=-0.6064964,
final_local_scores=-0.7225113, pem=-0.0008609424,
similarity=-17.466784]

5: 34544702 Donoughmore_GAA
[winner: 4137926 Donoughmore, pem_pos=0,
final_scores=-0.035849154, final_local_scores=-0.022392556,
pem=0.0, similarity=3.4214745|
gt: 34544702 Donoughmore_GAA, pem_pos=4,
final_scores=-0.7805074, final_local_scores=-0.8208501,
pem=-5.0657544, similarity=7.14227]

6: 32532531 Cork_GAA
[winner: 8140815 Cork_(city), pem_pos=0,
final_scores=0.18636702, final_local_scores=0.14833796,
pem=-0.83752966, similarity=13.07103|
gt: 32532531 Cork_GAA, pem_pos=1,
final_scores=-0.07565981, final_local_scores=-0.16519146,
pem=-1.0852067, similarity=5.0615935]

7: 2911979 All-Ireland_Senior_Ladies'_Football_Championship
[winner: 780802 All-Ireland_Senior_Hurling_Championship,
pem_pos=0, final_scores=0.032502696,
final_local_scores=-0.038852915, pem=-0.83555937,
similarity=7.4741664|
gt: 2911979 All-Ireland_Senior_Ladies'_Football_Championship, pem_pos=11,
final_scores=-0.56786704, final_local_scores=-0.6110382,
pem=-2.7536645, similarity=0.82996875]

8: 2912562 All-Ireland_Senior_Camogie_Championship
[winner: 780802 All-Ireland_Senior_Hurling_Championship,
pem_pos=0, final_scores=0.16357385,
final_local_scores=0.111577764, pem=-0.83555937,
similarity=11.963318|
gt: 2912562 All-Ireland_Senior_Camogie_Championship,
pem_pos=7, final_scores=-0.21832307,
final_local_scores=-0.27906844, pem=-2.624855,
similarity=10.036124]

9: 879901 Gaelic_games
[winner: 37688 Gaelic_football, pem_pos=6,
final_scores=-0.11284913, final_local_scores=-0.20822242,
pem=-2.5592473, similarity=11.793525|
RECALL MISS: 879901 Gaelic_games]
[winner: 6980625 IPod_game, pem_pos=5,
final_scores=-1.5768816, final_local_scores=-1.5480359,
pem=-3.0674694, similarity=-25.425388|
RECALL MISS: 879901 Gaelic_games]

10: 2912562 All-Ireland_Senior_Camogie_Championship
 [winner: 885732 All-Ireland_Senior_Football_Championship,
 pem_pos=1, final_scores=0.13267778,
 final_local_scores=0.07730216, pem=-0.91213274,
 similarity=11.356883]
 gt: 2912562 All-Ireland_Senior_Camogie_Championship,
 pem_pos=2, final_scores=-0.022948816,
 final_local_scores=-0.06668162, pem=-1.5766481,
 similarity=10.673858]
 [winner: 16593817
 Art_competitions_at_the_1936_Summer_Olympics, pem_pos=1,
 final_scores=-0.8016128, final_local_scores=-0.71724504,
 pem=-1.3862944, similarity=-9.775426]
 RECALL MISS: 2912562 All-Ireland_Senior_Camogie_Championship]
 [winner: 73109 Camogie, pem_pos=0,
 final_scores=-1.1115314, final_local_scores=-1.2741852,
 pem=-0.0008609424, similarity=-33.929836]
 RECALL MISS: 2912562 All-Ireland_Senior_Camogie_Championship]
 11: 26268503 National_Camogie_League
 [winner: 64290 National_League, pem_pos=0,
 final_scores=0.04625444, final_local_scores=0.12214771,
 pem=-0.07431644, similarity=8.1389885]
 gt: 26268503 National_Camogie_League, pem_pos=10,
 final_scores=-0.40746868, final_local_scores=-0.4582425,
 pem=-3.39782, similarity=8.892715]
 12: 19888671 All-Ireland_Senior_Club_Camogie_Championship
 [winner: 19888671 All-Ireland_Senior_Club_Camogie_Championship, pem_pos=0,
 final_scores=-0.38157272, final_local_scores=-0.44510886,
 pem=0.0, similarity=-9.193223]
 gt: 19888671 All-Ireland_Senior_Club_Camogie_Championship, pem_pos=0,
 final_scores=-0.38157272, final_local_scores=-0.44510886,
 pem=0.0, similarity=-9.193223]
 13: 26268503 National_Camogie_League
 [winner: 64290 National_League, pem_pos=0,
 final_scores=-0.0547397, final_local_scores=-0.007707253,
 pem=-0.07431644, similarity=4.2638574]
 gt: 26268503 National_Camogie_League, pem_pos=10,
 final_scores=-0.51495427, final_local_scores=-0.5769716,
 pem=-3.39782, similarity=5.349603]
 14: 2911812 Ladies'_Gaelic_football
 [winner: 780802 All-Ireland_Senior_Hurling_Championship,
 pem_pos=0, final_scores=0.11806099,
 final_local_scores=0.061861917, pem=-0.83555937,
 similarity=10.479698]
 RECALL MISS: 2911812 Ladies'_Gaelic_football]
 [winner: 9119799 In_the_Shadow_of_the_Sun, pem_pos=8,
 final_scores=-1.0459007, final_local_scores=-0.90829027,
 pem=-2.944439, similarity=-7.003165]

RECALL MISS: 2911812 Ladies'_Gaelic_football]
 [winner: 143183 Football_(ball), pem_pos=5,
 final_scores=-2.3563938, final_local_scores=-2.4897692,
 pem=-4.263315, similarity=-47.02539]
 RECALL MISS: 2911812 Ladies'_Gaelic_football]
 15: 2911812 Ladies'_Gaelic_football
 [winner: 37688 Gaelic_football, pem_pos=0,
 final_scores=-0.11104491, final_local_scores=-0.20625159,
 pem=0.0, similarity=-2.0652459]
 gt: 2911812 Ladies'_Gaelic_football, pem_pos=1,
 final_scores=-0.2669527, final_local_scores=-0.3284065,
 pem=-0.2876821, similarity=-4.146133]
 1: 10741716 Killeagh
 [pem_pos=0, final_scores=0.3095116,
 final_local_scores=0.36134216, pem=0.0, similarity=14.872884]
 2: 46709 County_Cork
 [pem_pos=0, final_scores=0.3615626,
 final_local_scores=0.33699465, pem=0.0, similarity=14.146305]
 3: 6157166 Killeagh_GAA
 [pem_pos=1, final_scores=0.2230293,
 final_local_scores=0.20545405, pem=-0.16081539,
 similarity=11.095411]
 4: 16335590 Granagh-Ballingarry_GAA
 [pem_pos=0, final_scores=0.13626407,
 final_local_scores=0.16711476, pem=0.0, similarity=9.076751]
 1: 14560 Republic_of_Ireland
 [pem_pos=3, final_scores=0.4819154,
 final_local_scores=0.48697644, pem=-2.2576945,
 similarity=30.899754]
 2: 879901 Gaelic_games
 [pem_pos=0, final_scores=0.13809611,
 final_local_scores=0.10318436, pem=0.0, similarity=7.16894]
 3: 8140815 Cork_(city)
 [pem_pos=0, final_scores=0.18636702,
 final_local_scores=0.14833796, pem=-0.83752966,
 similarity=13.07103]
 4: 780802 All-Ireland_Senior_Hurling_Championship
 [pem_pos=0, final_scores=0.032502696,
 final_local_scores=-0.038852915, pem=-0.83555937,
 similarity=7.4741664]
 5: 8140815 Cork_(city)
 [pem_pos=0, final_scores=0.1586412,
 final_local_scores=0.11805174, pem=-0.83752966,
 similarity=12.167229]
 6: 780802 All-Ireland_Senior_Hurling_Championship
 [pem_pos=0, final_scores=0.16357385,
 final_local_scores=0.111577764, pem=-0.83555937,
 similarity=11.963318]
 7: 885732 All-Ireland_Senior_Football_Championship
 [pem_pos=1, final_scores=0.13267778,
 final_local_scores=0.07730216, pem=-0.91213274,
 similarity=11.356883]

8: 885732 All-Ireland_Senior_Football_Championship
 [pem_pos=0, final_scores=0.14209221,
 final_local_scores=0.08684045, pem=0.0, similarity=6.6812053]
 9: 64290 National_League
 [pem_pos=0, final_scores=0.04625444,
 final_local_scores=0.12214771, pem=-0.07431644,
 similarity=8.1389885]
 10: 56885 Limerick
 [pem_pos=0, final_scores=0.24662237,
 final_local_scores=0.21245492, pem=-0.62048864,
 similarity=13.804105]
 11: 780802 All-Ireland_Senior_Hurling_Championship
 [pem_pos=0, final_scores=0.07034804,
 final_local_scores=0.009742752, pem=-0.83555937,
 similarity=8.924358]
 12: 32532531 Cork_GAA
 [pem_pos=1, final_scores=0.07260005,
 final_local_scores=0.0055749714, pem=-1.0852067,
 similarity=10.157604]
 13: 780802 All-Ireland_Senior_Hurling_Championship
 [pem_pos=0, final_scores=0.11806099,
 final_local_scores=0.061861917, pem=-0.83555937,
 similarity=10.479698]

Sample: b'Trade_Lines'

_ -LRB- newspaper -RRB- ` Trade Lines ' is a 1 pennysaver -
 style free 2 weekly newspaper consisting exclusively of 3 cla
 ssified and 4 display advertising . It has been serving
 selected markets in [1 1 Berrien 1] and [2 2 Van Burren 2]
 counties in southwestern [3 3 Michigan 3] since 1949 . Trade
 Lines is distributed weekly each Monday to 44,000 households
 in the communities of [1 5 Baroda 1] , [4 4 Benton Harbor 4] ,
 [5 5 Berrien Springs 5] , [6 6 Berrien Center 6] , 6 Bridgman
 , 7 Coloma , [2 8 Eau Claire 2] , 9 Riverside , 10 St.
 Joseph , , 11 Sodus , 12 Stevensville , and [3 13 Watervliet 3
] in [4 Berrien County 4] ; and [5 14 Hartford 5] and [7 7 So
 uth Haven 7]

1: 95115 Berrien_County,_Michigan
 2: 95031 Van_Buren_County,_Michigan
 3: 18859 Michigan
 4: 117146 Benton_Harbor,_Michigan
 5: 117149 Berrien_Springs,_Michigan
 6: 117148 Berrien_Township,_Michigan
 7: 150537 South_Haven,_Michigan

1: 10656136 Pennysaver
 [winner: 10656136 Pennysaver, pem_pos=0,
 final_scores=-0.9311954, final_local_scores=-0.9294648,
 pem=0.0, similarity=-23.647371]
 gt: 10656136 Pennysaver, pem_pos=0,
 final_scores=-0.9311954, final_local_scores=-0.9294648,
 pem=0.0, similarity=-23.647371]
 2: 1618418 Weekly_newspaper

[winner: 27830504 Drita_(newspaper), pem_pos=13,
final_scores=-1.0406427, final_local_scores=-0.89559305,
pem=-3.1345718, similarity=-5.590282]
RECALL MISS: 1618418 Weekly_newspaper]
[winner: 1618418 Weekly_newspaper, pem_pos=2,
final_scores=-1.2564577, final_local_scores=-1.334094,
pem=-1.0135063, similarity=-30.21071]
gt: 1618418 Weekly_newspaper, pem_pos=2,
final_scores=-1.2564577, final_local_scores=-1.334094,
pem=-1.0135063, similarity=-30.21071]
3: 646525 Classified_advertising
[winner: 203211 Meteorite_classification, pem_pos=7,
final_scores=-1.4340096, final_local_scores=-1.3785088,
pem=-2.2858303, similarity=-24.617037]
gt: 646525 Classified_advertising, pem_pos=12,
final_scores=-1.9622226, final_local_scores=-2.041122,
pem=-3.247771, similarity=-39.15955]
4: 646645 Display_advertising
[winner: 646645 Display_advertising, pem_pos=0,
final_scores=-1.1071243, final_local_scores=-1.0712445,
pem=-0.16699702, similarity=-26.970201]
gt: 646645 Display_advertising, pem_pos=0,
final_scores=-1.1071243, final_local_scores=-1.0712445,
pem=-0.16699702, similarity=-26.970201]
5: 117143 Baroda,_Michigan
[winner: 515128 Vadodara, pem_pos=0,
final_scores=0.3336799, final_local_scores=0.4231211,
pem=-0.13705625, similarity=17.461823]
gt: 117143 Baroda,_Michigan, pem_pos=2,
final_scores=-0.29235792, final_local_scores=-0.29831612,
pem=-1.6106702, similarity=3.946439]
6: 117152 Bridgman,_Michigan
[winner: 117152 Bridgman,_Michigan, pem_pos=0,
final_scores=-0.054264963, final_local_scores=-0.045507118,
pem=-0.62954336, similarity=6.1552444]
gt: 117152 Bridgman,_Michigan, pem_pos=0,
final_scores=-0.054264963, final_local_scores=-0.045507118,
pem=-0.62954336, similarity=6.1552444]
7: 117156 Coloma,_Michigan
[winner: 745079 Coloma,_California, pem_pos=0,
final_scores=-0.10468279, final_local_scores=-0.08291133,
pem=-0.5440662, similarity=4.57419]
gt: 117156 Coloma,_Michigan, pem_pos=1,
final_scores=-0.28481755, final_local_scores=-0.31133592,
pem=-1.5029718, similarity=2.9722223]
8: 117158 Eau_Claire,_Michigan
[winner: 151296 Eau_Claire,_Wisconsin, pem_pos=0,
final_scores=0.26446897, final_local_scores=0.252352,
pem=-0.57503253, similarity=14.747516]
gt: 117158 Eau_Claire,_Michigan, pem_pos=4,
final_scores=-0.27857354, final_local_scores=-0.2859527,
pem=-1.8047962, similarity=5.3710775]

9: 117164 Hagar_Township,_Michigan
 [winner: 107942 Riverside,_California, pem_pos=0,
 final_scores=0.00875999, final_local_scores=-0.007050693,
 pem=-1.3697267, similarity=11.328093]
 gt: 117164 Hagar_Township,_Michigan, pem_pos=18,
 final_scores=-0.72388613, final_local_scores=-0.791011,
 pem=-3.8891306, similarity=1.634065]

10: 117180 St._Joseph,_Michigan
 [winner: 149369 St._Joseph,_Missouri, pem_pos=1,
 final_scores=0.023209319, final_local_scores=-0.023388833,
 pem=-1.4748399, similarity=11.412152]
 gt: 117180 St._Joseph,_Michigan, pem_pos=2,
 final_scores=-0.20087509, final_local_scores=-0.28135914,
 pem=-2.4217203, similarity=8.863087]

11: 117179 Sodus_Township,_Michigan
 [winner: 260169 Sodus,_New_York, pem_pos=1,
 final_scores=-0.06866966, final_local_scores=-0.09740269,
 pem=-0.603521, similarity=4.465063]
 gt: 117179 Sodus_Township,_Michigan, pem_pos=0,
 final_scores=-0.093017265, final_local_scores=-0.10913057,
 pem=-0.44911385, similarity=3.27539]

12: 117182 Stevensville,_Michigan
 [winner: 123589 Stevensville,_Montana, pem_pos=0,
 final_scores=-0.077975824, final_local_scores=-0.076166786,
 pem=-0.63204867, similarity=5.253923]
 gt: 117182 Stevensville,_Michigan, pem_pos=1,
 final_scores=-0.117828906, final_local_scores=-0.12711237,
 pem=-0.84815294, similarity=4.908814]

13: 117186 Watervliet,_Michigan
 [winner: 126099 Watervliet,_New_York, pem_pos=0,
 final_scores=0.1510794, final_local_scores=0.15301096,
 pem=-0.24169113, similarity=9.97022]
 gt: 117186 Watervliet,_Michigan, pem_pos=1,
 final_scores=-0.112693384, final_local_scores=-0.13329446,
 pem=-0.67051923, similarity=3.7583277]

14: 119207 Hartford,_Michigan
 [winner: 53678 Hartford,_Connecticut, pem_pos=0,
 final_scores=0.24441008, final_local_scores=0.23666209,
 pem=-0.5798223, similarity=14.305346]
 gt: 119207 Hartford,_Michigan, pem_pos=9,
 final_scores=-0.5965171, final_local_scores=-0.66095376,
 pem=-3.509788, similarity=3.4523082]

1: 95115 Berrien_County,_Michigan
 [pem_pos=1, final_scores=0.320427,
 final_local_scores=0.29167664, pem=-0.96312964,
 similarity=18.031574]

2: 95031 Van_Buren_County,_Michigan
 [pem_pos=0, final_scores=0.42242652,
 final_local_scores=0.42587727, pem=0.0, similarity=16.79874]

3: 18859 Michigan
 [pem_pos=0, final_scores=0.52009624,
 final_local_scores=0.49920845, pem=-0.24788691,

similarity=20.335136]
4: 117146 Benton_Harbor,_Michigan
[pem_pos=0, final_scores=0.2795959,
final_local_scores=0.24691641, pem=0.0, similarity=11.458191]
5: 117149 Berrien_Springs,_Michigan
[pem_pos=0, final_scores=0.28469306,
final_local_scores=0.2753798, pem=0.0, similarity=12.307595]
6: 117148 Berrien_Township,_Michigan
[pem_pos=0, final_scores=0.1465003,
final_local_scores=0.15789378, pem=0.0, similarity=8.8015785]
7: 150537 South_Haven,_Michigan
[pem_pos=0, final_scores=0.03982316,
final_local_scores=0.0016851127, pem=-0.49645925,
similarity=6.839823]
1: 515128 Vadodara
[pem_pos=0, final_scores=0.3336799,
final_local_scores=0.4231211, pem=-0.13705625,
similarity=17.461823]
2: 151296 Eau_Claire,_Wisconsin
[pem_pos=0, final_scores=0.26446897,
final_local_scores=0.252352, pem=-0.57503253,
similarity=14.747516]
3: 126099 Watervliet,_New_York
[pem_pos=0, final_scores=0.1510794,
final_local_scores=0.15301096, pem=-0.24169113,
similarity=9.97022]
4: 95115 Berrien_County,_Michigan
[pem_pos=0, final_scores=0.24419616,
final_local_scores=0.20840606, pem=0.0, similarity=10.308966]
5: 53678 Hartford,_Connecticut
[pem_pos=0, final_scores=0.24441008,
final_local_scores=0.23666209, pem=-0.5798223,
similarity=14.305346]

Sample: b'Peter_Scott'

_ -LRB- cricketer -RRB- Peter James Scott -LRB- born 20
November 1982 -RRB- is an [1 1 English 1] 2 cricket er .
Scott is a right-handed who bowls right-arm 3 off break . He
was born in [1 1 Nottingham 1] , [2 4 Nottinghamshire 2] .
Scott represented the [2 2 Nottinghamshire Cricket Board 2]
in a single [3 3 List A 3] match against [4 4 Bedfordshire 4]
in the 1st round of the 5 2001 [3 Cheltenham & Gloucester Tro
phy 3] . In his only [4 List A 4] match he took a single
wicket at a cost of 28 runs . In 2004 , he made his debut for
[5 5 Staffordshire 5] against [6 6 Bedfordshire 6] . Scott
played 2 further Championship matches for the county ,
against [7 7 Norfolk 7] in 2004 and [8 8 Hertfordshire 8] in
2005 . Scott joined [9 9 Herefordshire 9] in 2006 ,
representing the county in 6 Championship matches during that
season , the last of which came against [10 10 Cornwall 10] .
He also represented the county in the [11 11 MCCA Knockout Tr
ophy 11] , making his debut against [12 12 Northumberland 12]

. During the 2006 season , he played 2 further Trophy matches against [13 13 Cheshire 13] and [14 14 Shropshire 14]

1: 39470 Nottingham

2: 27793154 Nottinghamshire_Cricket_Board

3: 683459 List_A_cricket

4: 3691144 Bedfordshire_County_Cricket_Club

5: 3712165 Staffordshire_County_Cricket_Club

6: 3691144 Bedfordshire_County_Cricket_Club

7: 3690942 Norfolk_County_Cricket_Club

8: 3711946 Hertfordshire_County_Cricket_Club

9: 3702547 Herefordshire_County_Cricket_Club

10: 3702042 Cornwall_County_Cricket_Club

11: 7543416 MCCA_Knockout_Trophy

12: 3701987 Northumberland_County_Cricket_Club

13: 3712477 Cheshire_County_Cricket_Club

14: 3702447 Shropshire_County_Cricket_Club

1: 18803164 English_people

[winner: 9316 England, pem_pos=1, final_scores=0.5891632, final_local_scores=0.64894885, pem=-1.3527626, similarity=30.812172]

gt: 18803164 English_people, pem_pos=2, final_scores=-0.44324178, final_local_scores=-0.47246423, pem=-2.2084396, similarity=2.0002756]

2: 25675557 Cricket

[winner: 330091 Rhaphidophoridae, pem_pos=10, final_scores=-1.4454457, final_local_scores=-1.3987503, pem=-2.56456, similarity=-23.705307]

gt: 25675557 Cricket, pem_pos=0, final_scores=-1.6617601, final_local_scores=-1.8455372, pem=-0.07532314, similarity=-50.575176]

3: 711518 Off_break

[winner: 711518 Off_break, pem_pos=0, final_scores=-1.2669952, final_local_scores=-1.3976352, pem=0.0, similarity=-37.61851]

gt: 711518 Off_break, pem_pos=0, final_scores=-1.2669952, final_local_scores=-1.3976352, pem=0.0, similarity=-37.61851]

4: 68049 Nottinghamshire

[winner: 667341 Nottinghamshire_County_Cricket_Club, pem_pos=1, final_scores=0.3069541, final_local_scores=0.2435379, pem=-1.1067253, similarity=17.375914]

gt: 68049 Nottinghamshire, pem_pos=0, final_scores=0.019125089, final_local_scores=0.014126807, pem=-0.28438178, similarity=6.0577974]

5: 27734795 2001_Cheltenham_&_Gloucester_Trophy

[winner: 1899702 Friends_Provident_Trophy, pem_pos=0, final_scores=0.2438476, final_local_scores=0.19571885, pem=0.0, similarity=9.930354]

gt: 27734795 2001_Cheltenham_&_Gloucester_Trophy, pem_pos=1, final_scores=-0.08982684, final_local_scores=-0.14823973, pem=-1.537979,

similarity=8.029713]
[winner: 34551 2001, pem_pos=0, final_scores=-0.6289548,
final_local_scores=-0.39782894, pem=-0.5613713,
similarity=-4.7294717]
RECALL MISS: 27734795
2001_Cheltenham_&_Gloucester_Trophy]
1: 39470 Nottingham
[pem_pos=0, final_scores=0.42316633,
final_local_scores=0.46039397, pem=-0.29126778,
similarity=19.412746]
2: 27793154 Nottinghamshire_Cricket_Board
[pem_pos=0, final_scores=0.29058725,
final_local_scores=0.28789386, pem=0.0, similarity=12.68104]
3: 683459 List_A_cricket
[pem_pos=0, final_scores=0.16483693,
final_local_scores=0.10279384, pem=-0.020267159,
similarity=7.267502]
4: 3691144 Bedfordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.4703849,
final_local_scores=0.41738176, pem=-1.6135224,
similarity=25.319803]
5: 3712165 Staffordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.40576953,
final_local_scores=0.34921566, pem=-1.68797,
similarity=23.690449]
6: 3691144 Bedfordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.29806387,
final_local_scores=0.2291472, pem=-1.6135224,
similarity=19.702509]
7: 3690942 Norfolk_County_Cricket_Club
[pem_pos=6, final_scores=0.21317054,
final_local_scores=0.1340943, pem=-3.4736512,
similarity=26.98161]
8: 3711946 Hertfordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.51840883,
final_local_scores=0.46654075, pem=-2.0051975,
similarity=28.916796]
9: 3702547 Herefordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.18478684,
final_local_scores=0.11319205, pem=-1.3582021,
similarity=14.8537035]
10: 3702042 Cornwall_County_Cricket_Club
[pem_pos=2, final_scores=0.2151898,
final_local_scores=0.13861692, pem=-3.34357,
similarity=26.409172]
11: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.42373008,
final_local_scores=0.3775795, pem=0.0, similarity=15.357438]
12: 3701987 Northumberland_County_Cricket_Club
[pem_pos=1, final_scores=0.20606546,
final_local_scores=0.13255477, pem=-2.745579,
similarity=22.976297]

13: 3712477 Cheshire_County_Cricket_Club
 [pem_pos=2, final_scores=0.27110994,
 final_local_scores=0.20101431, pem=-2.629504,
 similarity=24.388033]
 14: 3702447 Shropshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.19529359,
 final_local_scores=0.12637562, pem=-2.279101,
 similarity=20.255117]
 1: 9316 England
 [pem_pos=1, final_scores=0.5891632,
 final_local_scores=0.64894885, pem=-1.3527626,
 similarity=30.812172]
 2: 667341 Nottinghamshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.3069541,
 final_local_scores=0.2435379, pem=-1.1067253,
 similarity=17.375914]
 3: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.2438476,
 final_local_scores=0.19571885, pem=0.0, similarity=9.930354]
 4: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.07171522,
 final_local_scores=0.0010724962, pem=-0.020267159,
 similarity=4.231934]

Sample: b'Chris_Rushworth'

Christopher Rushworth -LRB- born 11 July 1986 -RRB- is an [1
 1 English 1] 2 cricket er . Rushworth is a right-handed who
 bowls right-arm 3 fast-medium . He was born in [2 4 Sunderland
 2] , [1 1 County Durham 1] . Rushworth made his debut in 5 co
 unty cricket for [2 2 Northumberland 2] against [3 3 Suffolk 3
] in 2004 . He played a further [3 Minor Counties
 Championship 3] match for the county against 6 Norfolk in
 2005 . In 2004 , he also made his debut for [4 4 Durham 4] in
 a [5 5 List A 5] match against [4 7 Sri Lanka 4] A . He would
 n't feature for [5 Durham 5] again following this match until
 the 8 2010 season , when he made his 9 first-class debut
 against [6 6 Yorkshire 6] in the [6 10 County Championship 6]
 . He featured in further [7 List A 7] matches in the [8 11 Cl
 ydesdale Bank 40 8] . In the 12 2011 season , he made his 13 T
 wenty20 debut against [7 7 Warwickshire 7] in the . His
 cousin , [8 8 Phil Mustard 8] , is the current [9 Durham 9] 14
 wicket-keeper and 15 captain

1: 400297 County_Durham
 2: 3701987 Northumberland_County_Cricket_Club
 3: 3691249 Suffolk_County_Cricket_Club
 4: 1176371 Durham_County_Cricket_Club
 5: 683459 List_A_cricket
 6: 869945 Yorkshire_County_Cricket_Club
 7: 667292 Warwickshire_County_Cricket_Club
 8: 6327683 Phil_Mustard
 1: 18803164 English_people
 [winner: 9316 England, pem_pos=1,

final_scores=0.57115847, final_local_scores=0.6289328,
 pem=-1.3527626, similarity=30.214855|
 gt: 18803164 English_people, pem_pos=2,
 final_scores=-0.38165525, final_local_scores=-0.40277696,
 pem=-2.2084396, similarity=4.079883]
 2: 25675557 Cricket
 [winner: 330091 Rhabdophoridae, pem_pos=10,
 final_scores=-1.5061951, final_local_scores=-1.4704914,
 pem=-2.56456, similarity=-25.846205|
 gt: 25675557 Cricket, pem_pos=0,
 final_scores=-1.8136301, final_local_scores=-2.020934,
 pem=-0.07532314, similarity=-55.809376]
 3: 678695 Fast_bowling
 [winner: 678695 Fast_bowling, pem_pos=0,
 final_scores=-1.13025, final_local_scores=-1.2893271,
 pem=0.0, similarity=-34.38638|
 gt: 678695 Fast_bowling, pem_pos=0,
 final_scores=-1.13025, final_local_scores=-1.2893271,
 pem=0.0, similarity=-34.38638]
 4: 264853 Sunderland_Tyne_and_Wear
 [winner: 184474 Sunderland_A.F.C., pem_pos=0,
 final_scores=0.07675339, final_local_scores=0.08215357,
 pem=-0.39640337, similarity=8.697042|
 gt: 264853 Sunderland_Tyne_and_Wear, pem_pos=2,
 final_scores=-0.2390222, final_local_scores=-0.24185336,
 pem=-2.7942786, similarity=12.068048]
 5: 2228846 County_cricket
 [winner: 2228846 County_cricket, pem_pos=1,
 final_scores=-1.0112476, final_local_scores=-1.2027462,
 pem=-0.048115727, similarity=-31.54097|
 gt: 2228846 County_cricket, pem_pos=1,
 final_scores=-1.0112476, final_local_scores=-1.2027462,
 pem=-0.048115727, similarity=-31.54097]
 6: 3690942 Norfolk_County_Cricket_Club
 [winner: 3690942 Norfolk_County_Cricket_Club, pem_pos=6,
 final_scores=-0.03671284, final_local_scores=-0.11921835,
 pem=-3.4736512, similarity=19.422255|
 gt: 3690942 Norfolk_County_Cricket_Club, pem_pos=6,
 final_scores=-0.03671284, final_local_scores=-0.11921835,
 pem=-3.4736512, similarity=19.422255]
 7: 26123784 Sri_Lanka_A_cricket_team
 [winner: 397246 Sri_Lanka_national_cricket_team,
 pem_pos=1, final_scores=0.622747,
 final_local_scores=0.620364, pem=-2.2667584,
 similarity=34.929596|
 RECALL MISS: 26123784 Sri_Lanka_A_cricket_team]
 8: 26970049 2010_English_cricket_season
 [winner: 43226 2010, pem_pos=0, final_scores=-0.8909374,
 final_local_scores=-0.7752668, pem=-1.6731367,
 similarity=-9.947018|
 RECALL MISS: 26970049 2010_English_cricket_season]
 [winner: 690345 Season(sports), pem_pos=0,

final_scores=-1.567694, final_local_scores=-1.6563354,
 pem=-1.219945, similarity=-38.704384|
 RECALL MISS: 26970049 2010_English_cricket_season|
 9: 11040 First-class_cricket
 [winner: 11040 First-class_cricket, pem_pos=0,
 final_scores=-1.2497061, final_local_scores=-1.4376502,
 pem=-0.053819276, similarity=-38.51996|
 gt: 11040 First-class_cricket, pem_pos=0,
 final_scores=-1.2497061, final_local_scores=-1.4376502,
 pem=-0.053819276, similarity=-38.51996]
 10: 27064222 2010_County_Championship
 [winner: 683599 County_Championship, pem_pos=0,
 final_scores=0.25292116, final_local_scores=0.18810478,
 pem=0.0, similarity=9.703135|
 gt: 27064222 2010_County_Championship, pem_pos=5,
 final_scores=-0.826098, final_local_scores=-0.92505383,
 pem=-5.608046, similarity=6.9816866]
 11: 27495460 2010_Clydesdale_Bank_40
 [winner: 27194123 ECB_40, pem_pos=0,
 final_scores=0.056829825, final_local_scores=0.033016354,
 pem=0.0, similarity=5.074987|
 gt: 27495460 2010_Clydesdale_Bank_40, pem_pos=3,
 final_scores=-0.46910983, final_local_scores=-0.5074788,
 pem=-2.8195457, similarity=4.278663]
 12: 32516180 2011_English_cricket_season
 [winner: 36225 2011, pem_pos=0, final_scores=-0.96838945,
 final_local_scores=-0.8680358, pem=-1.4232541,
 similarity=-14.074329|
 RECALL MISS: 32516180 2011_English_cricket_season|
 [winner: 690345 Season(sports), pem_pos=0,
 final_scores=-1.5710406, final_local_scores=-1.659991,
 pem=-1.219945, similarity=-38.813477|
 RECALL MISS: 32516180 2011_English_cricket_season|
 13: 770690 Twenty20
 [winner: 770690 Twenty20, pem_pos=0,
 final_scores=-0.014235958, final_local_scores=-0.072963946,
 pem=-0.039191484, similarity=2.1254525|
 gt: 770690 Twenty20, pem_pos=0,
 final_scores=-0.014235958, final_local_scores=-0.072963946,
 pem=-0.039191484, similarity=2.1254525]
 14: 463313 Wicket-keeper
 [winner: 463313 Wicket-keeper, pem_pos=0,
 final_scores=-1.7963835, final_local_scores=-2.017941,
 pem=-0.004312502, similarity=-56.106224|
 gt: 463313 Wicket-keeper, pem_pos=0,
 final_scores=-1.7963835, final_local_scores=-2.017941,
 pem=-0.004312502, similarity=-56.106224]
 15: 896020 Captain_(cricket)
 [winner: 3319715 Australian_Army_officer_rank_insignia,
 pem_pos=11, final_scores=-1.3003103,
 final_local_scores=-1.2598424, pem=-2.7026498,
 similarity=-18.809063|

```

gt: 896020 Captain_(cricket), pem_pos=3,
final_scores=-1.4913304, final_local_scores=-1.7132449,
pem=-1.9929447, similarity=-36.19899 ]
1: 400297 County_Durham
[pem_pos=0, final_scores=0.21379794,
final_local_scores=0.24112362, pem=0.0, similarity=11.285322]
2: 3701987 Northumberland_County_Cricket_Club
[pem_pos=1, final_scores=0.20038946,
final_local_scores=0.14704722, pem=-2.745579,
similarity=23.408781]
3: 3691249 Suffolk_County_Cricket_Club
[pem_pos=4, final_scores=0.12221886,
final_local_scores=0.062203944, pem=-2.9750729,
similarity=22.124912]
4: 1176371 Durham_County_Cricket_Club
[pem_pos=2, final_scores=0.83024067,
final_local_scores=0.80321056, pem=-2.3952909,
similarity=41.085083]
5: 683459 List_A_cricket
[pem_pos=0, final_scores=0.32402068,
final_local_scores=0.26661992, pem=-0.020267159,
similarity=12.156399]
6: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.67548937,
final_local_scores=0.6312813, pem=-1.7779728,
similarity=32.597298]
7: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.72652215,
final_local_scores=0.68911755, pem=-1.1093602,
similarity=30.68723]
8: 6327683 Phil_Mustard
[pem_pos=0, final_scores=0.80482846,
final_local_scores=0.8660212, pem=0.0, similarity=29.933512]
1: 9316 England
[pem_pos=1, final_scores=0.57115847,
final_local_scores=0.6289328, pem=-1.3527626,
similarity=30.214855]
2: 184474 Sunderland_A.F.C.
[pem_pos=0, final_scores=0.07675339,
final_local_scores=0.08215357, pem=-0.39640337,
similarity=8.697042]
3: 7562952 Minor_Counties_Cricket_Championship
[pem_pos=0, final_scores=0.20732217,
final_local_scores=0.14875811, pem=0.0, similarity=8.528952]
4: 397246 Sri_Lanka_national_cricket_team
[pem_pos=1, final_scores=0.622747,
final_local_scores=0.620364, pem=-2.2667584,
similarity=34.929596]
5: 1176371 Durham_County_Cricket_Club
[pem_pos=2, final_scores=0.7936211,
final_local_scores=0.76320916, pem=-2.3952909,
similarity=39.89136]

```

6: 683599 County_Championship
 [pem_pos=0, final_scores=0.25292116,
 final_local_scores=0.18810478, pem=0.0, similarity=9.703135]
 7: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.14261337,
 final_local_scores=0.06845999, pem=-0.020267159,
 similarity=6.2429113]
 8: 27194123 ECB_40
 [pem_pos=0, final_scores=0.056829825,
 final_local_scores=0.033016354, pem=0.0, similarity=5.074987]
 9: 1176371 Durham_County_Cricket_Club
 [pem_pos=2, final_scores=0.7091306,
 final_local_scores=0.67091614, pem=-2.3952909,
 similarity=37.137154]

Sample: b'Mississippi_Delta_National_Heritage_Area'

Mississippi Delta National Heritage Area is a federally designated [1 1 National Heritage Area 1] that seeks to preserve and promote the landscape , culture and history of the [2 2 Mississippi Delta 2] in the northwestern portion of the [1 U.S. 1] state of [3 3 Mississippi 3] . The region is famous for 1 blues music and a unique culture that has had broad influence on music and literature in the [2 United States 2] and worldwide . The national heritage area comprises 2 Bolivar , 3 Carroll , [4 4 Coahoma 4] , , 4 Holmes , 5 Humphreys , 6 Issaquena , 7 Leflore , 8 Panola , 9 Quitman , 10 Sharkey , [5 5 Tallahatchie 5] , 11 Tate , [3 12 Tunica 3] , 13 Warren , [4 14 Washington 4] and 15 Yazoo counties . Mississippi Delta National Heritage Area was established by the 16 Omnibus Public Land Management Act of 2009

1: 1232306 National_Heritage_Area
 2: 400255 Mississippi_Delta
 3: 16949861 Mississippi
 4: 94839 Coahoma_County,_Mississippi
 5: 94745 Tallahatchie_County,_Mississippi
 1: 3352 Blues

[winner: 23488452 Upton_Blues_Festival, pem_pos=3,
 final_scores=-1.0138125, final_local_scores=-0.95848554,
 pem=-1.0985211, similarity=-18.539478]
 gt: 3352 Blues, pem_pos=0, final_scores=-1.2162402,
 final_local_scores=-1.3098507, pem=-0.021050163,
 similarity=-34.88437]
 2: 94850 Bolivar_County,_Mississippi
 [winner: 123054 Bolivar,_Missouri, pem_pos=0,
 final_scores=-0.041385338, final_local_scores=-0.0554218,
 pem=-1.341501, similarity=9.731107]
 gt: 94850 Bolivar_County,_Mississippi, pem_pos=1,
 final_scores=-0.10729045, final_local_scores=-0.15187237,
 pem=-1.7775733, similarity=9.224258]
 3: 94847 Carroll_County,_Mississippi
 [winner: 96282 Carroll_County,_Illinois, pem_pos=2,
 final_scores=-0.3549387, final_local_scores=-0.39982805,

pem=-2.7271717, similarity=6.9888296|
gt: 94847 Carroll County, Mississippi, pem_pos=9,
final_scores=-0.5285177, final_local_scores=-0.5990979,
pem=-3.4044673, similarity=4.7254586]
4: 94824 Holmes County, Mississippi
[winner: 27159 Sherlock_Holmes, pem_pos=0,
final_scores=-0.3567745, final_local_scores=-0.3053828,
pem=-1.6177502, similarity=3.7740583|
gt: 94824 Holmes County, Mississippi, pem_pos=4,
final_scores=-0.40182334, final_local_scores=-0.46151093,
pem=-2.9160643, similarity=6.1753163]
5: 94823 Humphreys County, Mississippi
[winner: 91726 Humphreys County, Tennessee, pem_pos=0,
final_scores=-0.17568722, final_local_scores=-0.19714889,
pem=-1.1096416, similarity=4.2408004|
gt: 94823 Humphreys County, Mississippi, pem_pos=1,
final_scores=-0.2002695, final_local_scores=-0.20499007,
pem=-1.4614282, similarity=5.9198737]
6: 94822 Issaquena County, Mississippi
[winner: 94822 Issaquena County, Mississippi, pem_pos=0,
final_scores=-0.0023607016, final_local_scores=-0.017686397,
pem=0.0, similarity=3.5619156|
gt: 94822 Issaquena County, Mississippi, pem_pos=0,
final_scores=-0.0023607016, final_local_scores=-0.017686397,
pem=0.0, similarity=3.5619156]
7: 94779 Leflore County, Mississippi
[winner: 94779 Leflore County, Mississippi, pem_pos=0,
final_scores=-0.012702137, final_local_scores=-0.026656307,
pem=0.0, similarity=3.2942355|
gt: 94779 Leflore County, Mississippi, pem_pos=0,
final_scores=-0.012702137, final_local_scores=-0.026656307,
pem=0.0, similarity=3.2942355]
8: 94766 Panola County, Mississippi
[winner: 94766 Panola County, Mississippi, pem_pos=0,
final_scores=-0.06542505, final_local_scores=-0.10121122,
pem=-0.740221, similarity=5.0948052|
gt: 94766 Panola County, Mississippi, pem_pos=0,
final_scores=-0.06542505, final_local_scores=-0.10121122,
pem=-0.740221, similarity=5.0948052]
9: 94755 Quitman County, Mississippi
[winner: 94755 Quitman County, Mississippi, pem_pos=0,
final_scores=-0.14706343, final_local_scores=-0.15538962,
pem=-1.2352453, similarity=6.170033|
gt: 94755 Quitman County, Mississippi, pem_pos=0,
final_scores=-0.14706343, final_local_scores=-0.15538962,
pem=-1.2352453, similarity=6.170033]
10: 94752 Sharkey County, Mississippi
[winner: 94752 Sharkey County, Mississippi, pem_pos=0,
final_scores=-0.21519393, final_local_scores=-0.23101966,
pem=-1.2896552, similarity=4.208971|
gt: 94752 Sharkey County, Mississippi, pem_pos=0,
final_scores=-0.21519393, final_local_scores=-0.23101966,

```

pem=-1.2896552, similarity=4.208971 ]
11: 94743 Tate_County,_Mississippi
    [winner: 66101 Tate, pem_pos=0, final_scores=-0.19295736,
final_local_scores=-0.15982679, pem=-0.59705365,
similarity=2.5670335]
    gt: 94743 Tate_County,_Mississippi, pem_pos=1,
final_scores=-0.3262654, final_local_scores=-0.35403818,
pem=-2.1343162, similarity=5.1312504 ]
12: 94740 Tunica_County,_Mississippi
    [winner: 122454 Tunica,_Mississippi, pem_pos=0,
final_scores=0.08792634, final_local_scores=0.08255172,
pem=-0.43936664, similarity=8.942565]
    gt: 94740 Tunica_County,_Mississippi, pem_pos=1,
final_scores=-0.01829563, final_local_scores=-0.026455134,
pem=-0.5424657, similarity=6.250252 ]
13: 94734 Warren_County,_Mississippi
    [winner: 129883 Warren,_Ohio, pem_pos=1,
final_scores=-0.16830821, final_local_scores=-0.20451215,
pem=-2.7520003, similarity=12.952468]
    RECALL MISS: 94734 Warren_County,_Mississippi]
14: 94732 Washington_County,_Mississippi
    [winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=0.24283962, final_local_scores=0.22669774,
pem=-1.7496481, similarity=20.369684]
    RECALL MISS: 94732 Washington_County,_Mississippi]
15: 94705 Yazoo_County,_Mississippi
    [winner: 186353 Yazoo_(band), pem_pos=0,
final_scores=-0.20301601, final_local_scores=-0.11524314,
pem=-0.1931624, similarity=1.7010748]
    gt: 94705 Yazoo_County,_Mississippi, pem_pos=1,
final_scores=-0.30565703, final_local_scores=-0.3500201,
pem=-1.7380952, similarity=3.0964465 ]
16: 21540981 Omnibus_Public_Land_Management_Act_of_2009
    [winner: 21540981
Omnibus_Public_Land_Management_Act_of_2009, pem_pos=0,
final_scores=-0.2602927, final_local_scores=-0.2565943,
pem=0.0, similarity=-3.567573]
    gt: 21540981 Omnibus_Public_Land_Management_Act_of_2009,
pem_pos=0, final_scores=-0.2602927,
final_local_scores=-0.2565943, pem=0.0, similarity=-3.567573 ]
1: 1232306 National_Heritage_Area
    [pem_pos=0, final_scores=0.063567385,
final_local_scores=0.13053271, pem=0.0, similarity=7.9850698]
2: 400255 Mississippi_Delta
    [pem_pos=0, final_scores=0.2586934,
final_local_scores=0.23159122, pem=0.0, similarity=11.000857]
3: 16949861 Mississippi
    [pem_pos=0, final_scores=0.30481535,
final_local_scores=0.24714944, pem=-0.31944752,
similarity=13.202351]
4: 94839 Coahoma_County,_Mississippi
    [pem_pos=0, final_scores=0.03509371,

```


final_local_scores=0.025209948, pem=0.0, similarity=4.842028]
5: 94745 Tallahatchie County, Mississippi
[pem_pos=0, final_scores=0.17412634,
final_local_scores=0.18091515, pem=0.0, similarity=9.488582]
1: 3434750 United States
[pem_pos=0, final_scores=0.5087128,
final_local_scores=0.5025324, pem=-0.14075686,
similarity=19.851742]
2: 3434750 United States
[pem_pos=0, final_scores=0.5150003,
final_local_scores=0.50940055, pem=-0.073457584,
similarity=19.690716]
3: 122454 Tunica, Mississippi
[pem_pos=0, final_scores=0.08792634,
final_local_scores=0.08255172, pem=-0.43936664,
similarity=8.942565]
4: 108956 Washington, D.C.
[pem_pos=0, final_scores=0.24283962,
final_local_scores=0.22669774, pem=-1.7496481,
similarity=20.369684]

Sample: b'Bastille_discography'

The 1 discography of 2 Bastille , a [1 British 1] 3 alternative rock band , consists of one studio album , two 4 extended play s , six 5 single s and five 6 music video s. Originally a solo project by Dan Smith , Bastille was formed in 2010 . The band released its debut single `` Flaws '' / `` Icarus '' in July 2011 on the 7 independent record label Young and Lost Club . An extended play titled Laura Palmer followed later that year . The critical success of the EP and extensive touring brought the group to the attention of [2 8 Virgin Records 2] , which signed them in December 2011 . Bastille released their debut studio album 9 Bad Blood in March 2013 . The album reached number one on the [3 10 United Kingdom 3] albums chart and the top ten of the [4 11 Irish 4] albums chart . In March 2013 , it was silver by the 12 British Phonographic Industry -LRB- BPI -RRB- . Five singles were released from the album : `` Overjoyed '' , '' 13 Bad Blood '' , '' 14 Flaws '' , '' 15 Pompeii '' and '' 16 Laura Palmer '' , with `` [5 Pompeii 5] '' peaking at number one in [6 17 Ireland 6] and [7 18 Scotland 7] , and number two in the 19 UK

1: 510969 Discography
[winner: 9262774 ThaMuseMeant_discography, pem_pos=9,
final_scores=-1.0293344, final_local_scores=-0.92113525,
pem=-2.8332133, similarity=-7.991347]
RECALL MISS: 510969 Discography]
2: 36888630 Bastille_(band)
[winner: 66186 Bastille, pem_pos=0,
final_scores=-0.121466234, final_local_scores=-0.10223579,
pem=-0.17315467, similarity=1.9804351]
gt: 36888630 Bastille_(band), pem_pos=11,
final_scores=-0.9547687, final_local_scores=-1.0075126,

pem=-4.7706847, similarity=-0.032745354]
3: 167409 Alternative_rock
[winner: 7572 Christian_alternative_rock, pem_pos=2,
final_scores=-1.1732664, final_local_scores=-1.208049,
pem=-1.0963926, similarity=-25.998524]
gt: 167409 Alternative_rock, pem_pos=0,
final_scores=-1.4031246, final_local_scores=-1.5289149,
pem=0.0, similarity=-41.53616]
4: 156702 Extended_play
[winner: 18177970 Generalized_quaternion_interpolation,
pem_pos=3, final_scores=-1.07723,
final_local_scores=-0.9744335, pem=-1.0948027,
similarity=-19.035618]
RECALL MISS: 156702 Extended_play]
[winner: 26744396 Play_(theatre), pem_pos=0,
final_scores=-1.4716566, final_local_scores=-1.5970728,
pem=-0.63150793, similarity=-40.135887]
RECALL MISS: 156702 Extended_play]
5: 156540 Single_(music)
[winner: 156540 Single_(music), pem_pos=0,
final_scores=-1.4398158, final_local_scores=-1.5727003,
pem=-0.12414119, similarity=-42.1677]
gt: 156540 Single_(music), pem_pos=0,
final_scores=-1.4398158, final_local_scores=-1.5727003,
pem=-0.12414119, similarity=-42.1677]
6: 90138 Music_video
[winner: 1406689 She's_the_One_(World_Party_song),
pem_pos=5, final_scores=-1.1492019,
final_local_scores=-1.196753, pem=-2.079155,
similarity=-20.317019]
gt: 90138 Music_video, pem_pos=0,
final_scores=-1.2254709, final_local_scores=-1.3490002,
pem=0.0, similarity=-36.167145]
7: 332442 Independent_record_label
[winner: 1570811 Independent_Record, pem_pos=0,
final_scores=-0.91072834, final_local_scores=-0.86193514,
pem=0.0, similarity=-21.63215]
RECALL MISS: 332442 Independent_record_label]
[winner: 7955122 Label_(control), pem_pos=1,
final_scores=-1.3931336, final_local_scores=-1.3428092,
pem=-2.002902, similarity=-25.0903]
RECALL MISS: 332442 Independent_record_label]
8: 93767 Virgin_Records
[winner: 213782 Virgin_Group, pem_pos=1,
final_scores=0.19849153, final_local_scores=0.24546921,
pem=-0.6912689, similarity=15.174232]
gt: 93767 Virgin_Records, pem_pos=0,
final_scores=0.16557525, final_local_scores=0.2037803,
pem=0.0, similarity=10.170923]
9: 38743187 Bad_Blood_(Bastille_album)
[winner: 12416465 Bad_Blood_(The_X-Files), pem_pos=1,
final_scores=-0.17604116, final_local_scores=-0.08794248,

pem=-1.3420758, similarity=8.763751|
 gt: 38743187 Bad_Blood_(Bastille_album), pem_pos=19,
 final_scores=-0.93556756, final_local_scores=-0.932964,
 pem=-5.0945168, similarity=3.9529788]
 10: 2465857 UK_Albums_Chart
 [winner: 31717 United_Kingdom, pem_pos=0,
 final_scores=0.3777082, final_local_scores=0.35174868,
 pem=-0.09933331, similarity=15.126785|
 RECALL MISS: 2465857 UK_Albums_Chart]
 11: 5784179 Irish_Albums_Chart
 [winner: 14560 Republic_of_Ireland, pem_pos=3,
 final_scores=0.4906627, final_local_scores=0.5090856,
 pem=-2.2576945, similarity=31.559536|
 RECALL MISS: 5784179 Irish_Albums_Chart]
 12: 3897803 British_Phonographic_Industry
 [winner: 3897803 British_Phonographic_Industry,
 pem_pos=0, final_scores=-0.0016652793,
 final_local_scores=0.03829664, pem=0.0, similarity=5.232561|
 gt: 3897803 British_Phonographic_Industry, pem_pos=0,
 final_scores=-0.0016652793, final_local_scores=0.03829664,
 pem=0.0, similarity=5.232561]
 13: 37421826 Bad_Blood_(Bastille_song)
 [winner: 12416465 Bad_Blood_(The_X-Files), pem_pos=1,
 final_scores=-0.25897163, final_local_scores=-0.17853142,
 pem=-1.3420758, similarity=6.060396|
 RECALL MISS: 37421826 Bad_Blood_(Bastille_song)]
 14: 37421846 Flaws_(song)
 [winner: 923214 Fane_Flaws, pem_pos=1,
 final_scores=-0.34953737, final_local_scores=-0.26594234,
 pem=-1.0986123, similarity=2.1278877|
 gt: 37421846 Flaws_(song), pem_pos=5,
 final_scores=-0.9655207, final_local_scores=-0.94486743,
 pem=-4.1526136, similarity=-1.5244567]
 15: 38697489 Pompeii_(Bastille_song)
 [winner: 21476593 Pompeii, pem_pos=0,
 final_scores=-0.0169902, final_local_scores=0.005414784,
 pem=0.0, similarity=4.251301|
 gt: 38697489 Pompeii_(Bastille_song), pem_pos=14,
 final_scores=-1.3049124, final_local_scores=-1.3475707,
 pem=-7.08322, similarity=2.3951693]
 16: 38760249 Laura_Palmer_(song)
 [winner: 1191222 Laura_Palmer, pem_pos=0,
 final_scores=0.012277409, final_local_scores=0.085176855,
 pem=0.0, similarity=6.6315603|
 gt: 38760249 Laura_Palmer_(song), pem_pos=3,
 final_scores=-1.1855406, final_local_scores=-1.1836438,
 pem=-6.2766433, similarity=2.900781]
 17: 3627762 Irish_Singles_Chart
 [winner: 14560 Republic_of_Ireland, pem_pos=1,
 final_scores=0.39182734, final_local_scores=0.401123,
 pem=-1.5040284, similarity=24.239162|
 gt: 3627762 Irish_Singles_Chart, pem_pos=14,

```

final_scores=-0.8646862, final_local_scores=-0.92860746,
pem=-4.849524, similarity=2.750679 ]
18: 26905588 Scottish_Singles_and_Albums_Charts
    [winner: 26994 Scotland, pem_pos=0,
final_scores=0.43955296, final_local_scores=0.44638962,
pem=-0.22905664, similarity=18.656515]
    RECALL MISS: 26905588 Scottish_Singles_and_Albums_Charts]
19: 3349146 UK_Singles_Chart
    [winner: 31717 United_Kingdom, pem_pos=0,
final_scores=-0.025260255, final_local_scores=-0.10034257,
pem=-0.19850598, similarity=2.174797]
    gt: 3349146 UK_Singles_Chart, pem_pos=1,
final_scores=-0.53532386, final_local_scores=-0.58699715,
pem=-2.8989732, similarity=2.3376138 ]
1: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4437471,
final_local_scores=0.42388618, pem=-0.48936236,
similarity=19.40055]
2: 213782 Virgin_Group
    [pem_pos=1, final_scores=0.19849153,
final_local_scores=0.24546921, pem=-0.6912689,
similarity=15.174232]
3: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.3777082,
final_local_scores=0.35174868, pem=-0.09933331,
similarity=15.126785]
4: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.4906627,
final_local_scores=0.5090856, pem=-2.2576945,
similarity=31.559536]
5: 21476593 Pompeii
    [pem_pos=0, final_scores=0.05599989,
final_local_scores=0.085145384, pem=0.0, similarity=6.630621]
6: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.39182734,
final_local_scores=0.401123, pem=-1.5040284,
similarity=24.239162]
7: 26994 Scotland
    [pem_pos=0, final_scores=0.43955296,
final_local_scores=0.44638962, pem=-0.22905664,
similarity=18.656515]

```

Sample: b'Wheeling_Creek'

_ -LRB- [1 Ohio 1] -RRB- [2 Wheeling Creek 2] is a 1 tributa
ry of the [1 1 Ohio River 1] , long , in eastern [2 2 Ohio 2]
in the [3 3 United States 3] . Via the [3 Ohio River 3] , it
is part of the 2 watershed of the [4 4 Mississippi River 4] ,
draining an area of on the portion of the [5 5 Allegheny Plat
eau 5] . It flows for its entire length in [6 6 Belmont County
6] ; its tributaries also drain small areas of south-eastern [4
3 Harrison County 4] and south-western [5 4 Jefferson County
5] . [6 Wheeling Creek 6] rises in just east of the community

of [7 5 Flushing 7] , and flows generally east through [8 6 Union 8] , [9 7 Wheeling 9] , 8 Richland , 9 Colerain , and Townships , past the communities of 10 Lafferty , 11 Bannock , [7 7 Fairpoint 7] , 12 Maynard , 13 Barton , 14 Blaine , [10 15 Lansing 10] , and 16 Brookside , to [11 17 Bridgeport 11] , where it flows into the [12 Ohio River 12] from the west , just upstream of the mouth of [8 8 West Virginia 8] 's [13 18 Wheeling Creek 13] on the opposite bank . The [9 9 National Road 9] -LRB- U.S. Route 40 -RRB- parallels the stream between [14 Bridgeport 14] and Blaine . The [10 10 United States Geological Survey 10] operates a 19 stream gauge on the creek downstream of Blaine , upstream of the creek 's mouth . Between 1984 and 2005 , the annual mean flow of the creek at the gauge was 115 -LRB- 3

1: 22388 Ohio_River

2: 22199 Ohio

3: 3434750 United_States

4: 19579 Mississippi_River

5: 609240 Allegheny_Plateau

6: 93137 Belmont_County,_Ohio

7: 9617861 Fairpoint,_Ohio

8: 32905 West_Virginia

9: 55562 National_Road

10: 23814944 United_States_Geological_Survey

1: 72465 Tributary

[winner: 72465 Tributary, pem_pos=1, final_scores=-1.1647446, final_local_scores=-1.1965752, pem=-0.012074053, similarity=-31.552814]

gt: 72465 Tributary, pem_pos=1, final_scores=-1.1647446, final_local_scores=-1.1965752, pem=-0.012074053, similarity=-31.552814]

2: 191162 Drainage_basin

[winner: 191162 Drainage_basin, pem_pos=0, final_scores=-1.3684946, final_local_scores=-1.4809852, pem=0.0, similarity=-40.10584]

gt: 191162 Drainage_basin, pem_pos=0, final_scores=-1.3684946, final_local_scores=-1.4809852, pem=0.0, similarity=-40.10584]

3: 93111 Harrison_County,_Ohio

[winner: 91041 Harrison_County,_West_Virginia, pem_pos=0, final_scores=0.2240266, final_local_scores=0.191861, pem=-1.1398624, similarity=16.013977]

gt: 93111 Harrison_County,_Ohio, pem_pos=3, final_scores=0.1415671, final_local_scores=0.11456224, pem=-1.4735548, similarity=15.521899]

4: 93064 Jefferson_County,_Ohio

[winner: 77868 Jefferson_County,_Alabama, pem_pos=0, final_scores=0.23394702, final_local_scores=0.21594757, pem=-1.6805408, similarity=19.673061]

gt: 93064 Jefferson_County,_Ohio, pem_pos=7, final_scores=-0.035387352, final_local_scores=-0.09769124, pem=-2.8259044, similarity=16.542118]

5: 128941 Flushing, _Ohio
 [winner: 267693 Flushing, _Queens, pem_pos=0,
 final_scores=0.18270914, final_local_scores=0.18502748,
 pem=-0.48217976, similarity=12.23347]
 gt: 128941 Flushing, _Ohio, pem_pos=4,
 final_scores=-0.30717975, final_local_scores=-0.3195018,
 pem=-2.4668987, similarity=7.9705215]

6: 9273687 Union_Township, _Belmont_County, _Ohio
 [winner: 481130 Union_(American_Civil_War), pem_pos=0,
 final_scores=0.07928951, final_local_scores=0.08751893,
 pem=-1.132797, similarity=12.861778]
 RECALL MISS: 9273687
 Union_Township, _Belmont_County, _Ohio]

7: 9275642 Wheeling_Township, _Belmont_County, _Ohio
 [winner: 151290 Wheeling, _West_Virginia, pem_pos=0,
 final_scores=0.31001127, final_local_scores=0.27064323,
 pem=-0.6108737, similarity=15.488272]
 gt: 9275642 Wheeling_Township, _Belmont_County, _Ohio,
 pem_pos=2, final_scores=-0.41461912,
 final_local_scores=-0.39481592, pem=-2.254051,
 similarity=4.5654974]

8: 9286418 Richland_Township, _Belmont_County, _Ohio
 [winner: 57637 Richland, _Washington, pem_pos=0,
 final_scores=-0.03251411, final_local_scores=-0.046016067,
 pem=-1.9799644, similarity=13.483856]
 gt: 9286418 Richland_Township, _Belmont_County, _Ohio,
 pem_pos=16, final_scores=-0.5837818,
 final_local_scores=-0.589065, pem=-3.632456,
 similarity=6.264698]

9: 9286389 Colerain_Township, _Belmont_County, _Ohio
 [winner: 234355 Colerain_Township, _Hamilton_County, _Ohio,
 pem_pos=0, final_scores=-0.030566916,
 final_local_scores=-0.023945183, pem=-1.1470125,
 similarity=9.612773]
 gt: 9286389 Colerain_Township, _Belmont_County, _Ohio,
 pem_pos=2, final_scores=-0.20927507,
 final_local_scores=-0.17672248, pem=-1.3297552,
 similarity=6.047377]

10: 9617843 Lafferty, _Ohio
 [winner: 9617843 Lafferty, _Ohio, pem_pos=0,
 final_scores=-0.12810588, final_local_scores=-0.077425495,
 pem=-0.8425417, similarity=6.3610544]
 gt: 9617843 Lafferty, _Ohio, pem_pos=0,
 final_scores=-0.12810588, final_local_scores=-0.077425495,
 pem=-0.8425417, similarity=6.3610544]

11: 9617879 Bannock, _Ohio
 [winner: 96360 Bannock_County, _Idaho, pem_pos=2,
 final_scores=0.018686816, final_local_scores=0.013341382,
 pem=-1.1474794, similarity=10.728018]
 gt: 9617879 Bannock, _Ohio, pem_pos=3,
 final_scores=-0.21640068, final_local_scores=-0.15174031,
 pem=-1.3137113, similarity=6.7056456]

12: 9617831 Maynard, _Ohio
[winner: 257281 Maynard, _Massachusetts, pem_pos=0,
final_scores=0.010119662, final_local_scores=0.033641398,
pem=-0.8960406, similarity=9.966448|
gt: 9617831 Maynard, _Ohio, pem_pos=3,
final_scores=-0.36749625, final_local_scores=-0.34060138,
pem=-2.2840471, similarity=6.3464913]

13: 9617876 Barton, _Ohio
[winner: 95746 Barton County, _Kansas, pem_pos=0,
final_scores=-0.24820533, final_local_scores=-0.27159834,
pem=-1.7467902, similarity=5.4839935|
gt: 9617876 Barton, _Ohio, pem_pos=10,
final_scores=-0.5468448, final_local_scores=-0.5489209,
pem=-3.4753926, similarity=6.608543]

14: 9617872 Blaine, _Ohio
[winner: 150566 Blaine, _Minnesota, pem_pos=1,
final_scores=-0.08160171, final_local_scores=-0.0629006,
pem=-2.1705818, similarity=14.016595|
gt: 9617872 Blaine, _Ohio, pem_pos=12,
final_scores=-0.37810522, final_local_scores=-0.35749716,
pem=-3.279365, similarity=11.254981]

15: 9617837 Lansing, _Ohio
[winner: 18538 Lansing, _Michigan, pem_pos=0,
final_scores=0.2056744, final_local_scores=0.18768331,
pem=-0.44571918, similarity=12.114448|
gt: 9617837 Lansing, _Ohio, pem_pos=7,
final_scores=-0.45114538, final_local_scores=-0.43002647,
pem=-3.1287937, similarity=8.271732]

16: 128940 Brookside, _Ohio
[winner: 104767 Brookside, _Alabama, pem_pos=1,
final_scores=-0.1644092, final_local_scores=-0.1504356,
pem=-1.7714326, similarity=9.233741|
gt: 128940 Brookside, _Ohio, pem_pos=4,
final_scores=-0.2936685, final_local_scores=-0.26631966,
pem=-2.329565, similarity=8.8107395]

17: 128939 Bridgeport, _Ohio
[winner: 23747575 Bridgeport, _Connecticut, pem_pos=0,
final_scores=0.18579738, final_local_scores=0.17283976,
pem=-0.7622919, similarity=13.393058|
gt: 128939 Bridgeport, _Ohio, pem_pos=5,
final_scores=-0.23786382, final_local_scores=-0.27311012,
pem=-2.9769523, similarity=12.128692]

18: 10119427 Wheeling_Creek_(West_Virginia)
[winner: 10119429 Wheeling_Creek_(Ohio), pem_pos=1,
final_scores=0.2164468, final_local_scores=0.2451899,
pem=-0.32378712, similarity=13.167473|
gt: 10119427 Wheeling_Creek_(West_Virginia), pem_pos=0,
final_scores=0.19738106, final_local_scores=0.20048177,
pem=0.0, similarity=10.072489]

19: 200161 Stream_gauge
[winner: 30860417 Needle_gauge_comparison_chart,
pem_pos=7, final_scores=-1.0635556,

```

final_local_scores=-0.9633927, pem=-2.360724,
similarity=-11.821861|
    RECALL MISS: 200161 Stream_gauge]
    [winner: 4877825 Bucks_Burn, pem_pos=3,
final_scores=-1.264446, final_local_scores=-1.211801,
pem=-1.7825414, similarity=-22.37911|
    RECALL MISS: 200161 Stream_gauge]
1: 22388 Ohio_River
    [pem_pos=0, final_scores=0.34570813,
final_local_scores=0.30041793, pem=0.0, similarity=13.054783]
2: 22199 Ohio
    [pem_pos=0, final_scores=0.5944413,
final_local_scores=0.5595888, pem=-0.19147073,
similarity=21.830206]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.478074,
final_local_scores=0.4899209, pem=-0.073457584,
similarity=19.109404]
4: 19579 Mississippi_River
    [pem_pos=0, final_scores=0.40835696,
final_local_scores=0.39265162, pem=0.0, similarity=15.80722]
5: 609240 Allegheny_Plateau
    [pem_pos=0, final_scores=0.05493532,
final_local_scores=0.063945666, pem=0.0, similarity=5.997979]
6: 93137 Belmont_County,_Ohio
    [pem_pos=0, final_scores=0.3204533,
final_local_scores=0.28812093, pem=0.0, similarity=12.687816]
7: 9617861 Fairpoint,_Ohio
    [pem_pos=0, final_scores=0.17826323,
final_local_scores=0.29118764, pem=0.0, similarity=12.779333]
8: 32905 West_Virginia
    [pem_pos=0, final_scores=0.53677434,
final_local_scores=0.49998182, pem=-0.30025622,
similarity=20.643007]
9: 55562 National_Road
    [pem_pos=0, final_scores=0.05518718,
final_local_scores=0.020730719, pem=0.0, similarity=4.708359]
10: 23814944 United_States_Geological_Survey
    [pem_pos=0, final_scores=0.16463287,
final_local_scores=0.19333342, pem=0.0, similarity=9.859168]
1: 22199 Ohio
    [pem_pos=0, final_scores=0.40416354,
final_local_scores=0.3517393, pem=-0.19147073,
similarity=15.6275625]
2: 10119429 Wheeling_Creek_(Ohio)
    [pem_pos=1, final_scores=0.15512519,
final_local_scores=0.17820534, pem=-0.32378712,
similarity=11.16852]
3: 22388 Ohio_River
    [pem_pos=0, final_scores=0.24576543,
final_local_scores=0.19124573, pem=0.0, similarity=9.796867]
4: 91041 Harrison_County,_West_Virginia

```


[pem_pos=0, final_scores=0.2240266,
 final_local_scores=0.191861, pem=-1.1398624,
 similarity=16.013977]
 5: 77868 Jefferson_County,_Alabama
 [pem_pos=0, final_scores=0.23394702,
 final_local_scores=0.21594757, pem=-1.6805408,
 similarity=19.673061]
 6: 10119427 Wheeling_Creek_(West_Virginia)
 [pem_pos=0, final_scores=0.08536939,
 final_local_scores=0.07812603, pem=0.0, similarity=6.4211493]
 7: 267693 Flushing,_Queens
 [pem_pos=0, final_scores=0.18270914,
 final_local_scores=0.18502748, pem=-0.48217976,
 similarity=12.23347]
 8: 481130 Union_(American_Civil_War)
 [pem_pos=0, final_scores=0.07928951,
 final_local_scores=0.08751893, pem=-1.132797,
 similarity=12.861778]
 9: 151290 Wheeling,_West_Virginia
 [pem_pos=0, final_scores=0.31001127,
 final_local_scores=0.27064323, pem=-0.6108737,
 similarity=15.488272]
 10: 18538 Lansing,_Michigan
 [pem_pos=0, final_scores=0.2056744,
 final_local_scores=0.18768331, pem=-0.44571918,
 similarity=12.114448]
 11: 23747575 Bridgeport,_Connecticut
 [pem_pos=0, final_scores=0.18579738,
 final_local_scores=0.17283976, pem=-0.7622919,
 similarity=13.393058]
 12: 22388 Ohio_River
 [pem_pos=0, final_scores=0.28993404,
 final_local_scores=0.23949322, pem=0.0, similarity=11.236669]
 13: 10119429 Wheeling_Creek_(Ohio)
 [pem_pos=1, final_scores=0.2164468,
 final_local_scores=0.2451899, pem=-0.32378712,
 similarity=13.167473]
 14: 23747575 Bridgeport,_Connecticut
 [pem_pos=0, final_scores=0.3389668,
 final_local_scores=0.3401541, pem=-0.7622919,
 similarity=18.386051]

Sample: b'Wolf_Prize'

The [1 Wolf Prize 1] is an international granted in [1 1 Isr
 ael 1] , that has been presented most years since 1978 to
 living scientists and artists for `` achievements in the
 interest of mankind and friendly relations among peoples ...
 irrespective of nationality , race , colour , religion , sex
 or political views . '' The prize is awarded in [2 2 Israel 2]
 by the 1 Wolf Foundation , founded by 2 Dr. [2 Ricardo Wolf 2]
 , a [3 German-born 3] inventor and former [3 3 Cuba 3] n
 ambassador to [4 Israel 4] . It is awarded in six fields : 3 A

griculture , 4 Chemistry , 5 Mathematics , [5 6 Medicine 5] , 7 Physics , and an 8 Arts prize that rotates between architecture , music , painting , and sculpture . Each prize consists of a diploma and [6 US\$ 6] 100,000 . The prize is described by the Foundation as being `` awarded annually '' , but is not in fact awarded every year : only six prizes were awarded in the 2000s -LRB- decade -RRB- in most fields , and four in [7 Physics 7] . The [8 Wolf Prizes 8] in physics and chemistry are often considered the most prestigious awards in those fields after the [9 Nobel Prize 9] . The prize in physics has gained a reputation for identifying future winners of the [10 Nobel Prize 10] -- from the 26 prizes awarded between 1978 and 2010 , fourteen winners have gone on to win the [11 Nobel Prize 11] , five of those in the following year . In medicine , the prize is probably the third most prestigious , after the [12 Nobel Prize 12] and the [4 4 Lasker Award 4] . Until the establishment of the [5 5 Abel Prize 5] , the [13 Wolf Prize 13] was probably the closest equivalent of a `` [14 Nobel Prize in Mathematics 14] '' , since the more prestigious [6 6 Fields Medal 6] was only awarded every four years to mathematicians under forty years old . The Prize in Agriculture has likewise been equated to a `` Nobel Prize in Agriculture '' . For past [15 Wolf Prize 15] laureates , see the separate lists for 9 Agriculture , 10 Chemistry , 11 Mathematics , [16 12 Medicine 16] , 13 Physics and 14 Arts

1: 9282173 Israel

2: 9282173 Israel

3: 5042481 Cuba

4: 717081 Lasker_Award

5: 159285 Abel_Prize

6: 10859 Fields_Medal

1: 12321305 Wolf_Foundation

[winner: 12321305 Wolf_Foundation, pem_pos=0, final_scores=0.02073969, final_local_scores=0.056628093, pem=0.0, similarity=5.7796082]

gt: 12321305 Wolf_Foundation, pem_pos=0, final_scores=0.02073969, final_local_scores=0.056628093, pem=0.0, similarity=5.7796082]

2: 12271740 Ricardo_Wolf

[winner: 12271740 Ricardo_Wolf, pem_pos=0, final_scores=0.12970392, final_local_scores=0.1947856, pem=0.0, similarity=9.902504]

gt: 12271740 Ricardo_Wolf, pem_pos=0, final_scores=0.12970392, final_local_scores=0.1947856, pem=0.0, similarity=9.902504]

[winner: 569138 Doctor_of_Divinity, pem_pos=1, final_scores=-1.5181558, final_local_scores=-1.5823212, pem=-1.294506, similarity=-36.090183]

RECALL MISS: 12271740 Ricardo_Wolf]

3: 714510 Wolf_Prize_in_Agriculture

[winner: 627 Agriculture, pem_pos=0,

final_scores=-0.2992171, final_local_scores=-0.29384804,
pem=-0.88757795, similarity=0.14748785|
gt: 714510 Wolf_Prize_in_Agriculture, pem_pos=3,
final_scores=-0.5743363, final_local_scores=-0.61345243,
pem=-2.0434456, similarity=-3.1043553]
4: 714440 Wolf_Prize_in_Chemistry
[winner: 5180 Chemistry, pem_pos=0,
final_scores=-0.04478067, final_local_scores=-0.07849766,
pem=-0.51907206, similarity=4.5699806|
gt: 714440 Wolf_Prize_in_Chemistry, pem_pos=6,
final_scores=-0.65920293, final_local_scores=-0.6868022,
pem=-2.7724245, similarity=-1.3289587]
5: 714278 Wolf_Prize_in_Mathematics
[winner: 18831 Mathematics, pem_pos=0,
final_scores=-0.011371687, final_local_scores=-0.024938084,
pem=-0.30250734, similarity=4.9905925|
gt: 714278 Wolf_Prize_in_Mathematics, pem_pos=3,
final_scores=-0.48414776, final_local_scores=-0.50536346,
pem=-2.4263039, similarity=2.2032754]
6: 714499 Wolf_Prize_in_Medicine
[winner: 18957 Medicine, pem_pos=0,
final_scores=0.034533635, final_local_scores=0.01886651,
pem=-0.5636483, similarity=7.7179337|
gt: 714499 Wolf_Prize_in_Medicine, pem_pos=6,
final_scores=-0.64606506, final_local_scores=-0.67896855,
pem=-3.23586, similarity=1.4250478]
7: 714321 Wolf_Prize_in_Physics
[winner: 22939 Physics, pem_pos=0,
final_scores=-0.11318527, final_local_scores=-0.15964176,
pem=-0.40543398, similarity=1.5304977|
gt: 714321 Wolf_Prize_in_Physics, pem_pos=3,
final_scores=-0.46050638, final_local_scores=-0.4832116,
pem=-2.152808, similarity=1.3770196]
8: 714281 Wolf_Prize_in_Arts
[winner: 29560452 The_arts, pem_pos=0,
final_scores=-0.32724577, final_local_scores=-0.36080423,
pem=-0.9358671, similarity=-1.5880146|
gt: 714281 Wolf_Prize_in_Arts, pem_pos=6,
final_scores=-0.62669355, final_local_scores=-0.61528087,
pem=-2.7273004, similarity=0.5599898]
9: 714510 Wolf_Prize_in_Agriculture
[winner: 627 Agriculture, pem_pos=0,
final_scores=-0.3389356, final_local_scores=-0.33723447,
pem=-0.88757795, similarity=-1.1472493|
gt: 714510 Wolf_Prize_in_Agriculture, pem_pos=3,
final_scores=-0.5802801, final_local_scores=-0.61994517,
pem=-2.0434456, similarity=-3.2981126]
10: 714440 Wolf_Prize_in_Chemistry
[winner: 5180 Chemistry, pem_pos=0,
final_scores=-0.029435158, final_local_scores=-0.06173502,
pem=-0.51907206, similarity=5.0702114|
gt: 714440 Wolf_Prize_in_Chemistry, pem_pos=6,

```

final_scores=-0.6711917, final_local_scores=-0.6998981,
pem=-2.7724245, similarity=-1.7197683 ]
11: 714278 Wolf_Prize_in_Mathematics
    [winner: 18831 Mathematics, pem_pos=0,
final_scores=0.0104019195, final_local_scores=-0.0011537373,
pem=-0.30250734, similarity=5.700365]
    gt: 714278 Wolf_Prize_in_Mathematics, pem_pos=3,
final_scores=-0.47719637, final_local_scores=-0.4977701,
pem=-2.4263039, similarity=2.4298775 ]
12: 714499 Wolf_Prize_in_Medicine
    [winner: 18957 Medicine, pem_pos=0,
final_scores=0.0908951, final_local_scores=0.08043286,
pem=-0.5636483, similarity=9.555197]
    gt: 714499 Wolf_Prize_in_Medicine, pem_pos=6,
final_scores=-0.5795146, final_local_scores=-0.6062723,
pem=-3.23586, similarity=3.5944488 ]
13: 714321 Wolf_Prize_in_Physics
    [winner: 22939 Physics, pem_pos=0,
final_scores=-0.068736106, final_local_scores=-0.11108779,
pem=-0.40543398, similarity=2.9794447]
    gt: 714321 Wolf_Prize_in_Physics, pem_pos=3,
final_scores=-0.42751348, final_local_scores=-0.44717187,
pem=-2.152808, similarity=2.4525173 ]
14: 714281 Wolf_Prize_in_Arts
    [winner: 29560452 The_arts, pem_pos=0,
final_scores=-0.2934668, final_local_scores=-0.32390586,
pem=-0.9358671, similarity=-0.48689324]
    gt: 714281 Wolf_Prize_in_Arts, pem_pos=6,
final_scores=-0.79431844, final_local_scores=-0.7983856,
pem=-2.7273004, similarity=-4.904223 ]
1: 9282173 Israel
    [pem_pos=0, final_scores=0.41090584,
final_local_scores=0.4495396, pem=-0.21279564,
similarity=18.662088]
2: 9282173 Israel
    [pem_pos=0, final_scores=0.4706757,
final_local_scores=0.5148291, pem=-0.21279564,
similarity=20.610456]
3: 5042481 Cuba
    [pem_pos=0, final_scores=0.41525376,
final_local_scores=0.4927041, pem=-0.23907764,
similarity=20.093128]
4: 717081 Lasker_Award
    [pem_pos=0, final_scores=0.17215298,
final_local_scores=0.19560751, pem=0.0, similarity=9.9270315]
5: 159285 Abel_Prize
    [pem_pos=0, final_scores=0.17679168,
final_local_scores=0.2065664, pem=0.0, similarity=10.254066]
6: 10859 Fields_Medal
    [pem_pos=0, final_scores=0.087228075,
final_local_scores=0.08080535, pem=0.0, similarity=6.501106]
1: 370712 Wolf_Prize

```

[pem_pos=0, final_scores=0.114829466,
final_local_scores=0.091249496, pem=0.0, similarity=6.81278]
2: 12271740 Ricardo_Wolf
[pem_pos=0, final_scores=0.12970392,
final_local_scores=0.1947856, pem=0.0, similarity=9.902504]
3: 11867 Germany
[pem_pos=0, final_scores=0.12711786,
final_local_scores=0.15041918, pem=0.0, similarity=8.578521]
4: 9282173 Israel
[pem_pos=0, final_scores=0.52665585,
final_local_scores=0.57597893, pem=-0.21279564,
similarity=22.43529]
5: 18957 Medicine
[pem_pos=0, final_scores=0.034533635,
final_local_scores=0.01886651, pem=-0.5636483,
similarity=7.7179337]
6: 18717338 United_States_dollar
[pem_pos=0, final_scores=0.35068738,
final_local_scores=0.4253736, pem=0.0, similarity=16.78371]
7: 22939 Physics
[pem_pos=0, final_scores=0.03176634,
final_local_scores=0.007701993, pem=-0.40543398,
similarity=6.5243692]
8: 370712 Wolf_Prize
[pem_pos=0, final_scores=0.068397924,
final_local_scores=0.0405301, pem=0.0, similarity=5.299212]
9: 21201 Nobel_Prize
[pem_pos=0, final_scores=0.21899353,
final_local_scores=0.15467569, pem=-0.15178943,
similarity=9.530999]
10: 21201 Nobel_Prize
[pem_pos=0, final_scores=0.15303038,
final_local_scores=0.08205603, pem=-0.15178943,
similarity=7.363883]
11: 21201 Nobel_Prize
[pem_pos=0, final_scores=0.18360849,
final_local_scores=0.11545795, pem=-0.15178943,
similarity=8.360663]
12: 21201 Nobel_Prize
[pem_pos=0, final_scores=0.23926602,
final_local_scores=0.17625535, pem=-0.15178943,
similarity=10.174979]
13: 370712 Wolf_Prize
[pem_pos=0, final_scores=0.11854164,
final_local_scores=0.09530449, pem=0.0, similarity=6.933789]
14: 1528075 Nobel_Prize_controversies
[pem_pos=0, final_scores=0.05299674,
final_local_scores=0.037427038, pem=0.0, similarity=5.20661]
15: 370712 Wolf_Prize
[pem_pos=0, final_scores=0.10888524,
final_local_scores=0.08475636, pem=0.0, similarity=6.619012]
16: 18957 Medicine

[pem_pos=0, final_scores=0.0908951,
final_local_scores=0.08043286, pem=-0.5636483,
similarity=9.555197]

Sample: b'Melville-Saltcoats'

Melville-Saltcoats is a 1 provincial 2 electoral district for the 3 Legislative Assembly of 4 Saskatchewan , [1 1 Canada 1] . Located in southeastern [1 Saskatchewan 1] , this constituency was created through the Representation Act , 1994 -LRB- [2 Saskatchewan 2] -RRB- by combining the district of [3 5 Saltcoats 3] with part of the constituency of 6 Melville . The largest centre in the riding is the city of 7 Melville -LRB- pop . 4,531 -RRB- . Melville is the in [4 Saskatchewan 4] . It is also a major transportation hub , with provincial 8 Highways 9 10 , 10 15 , and 11 47 running through the city -- as well as the [2 2 Canadian National Railway 2] mainline . Smaller centers in the district include the towns of 12 Langenburg , [3 3 Esterhazy 3] , [5 13 Saltcoats 5] and 14 Churchbridge ; and the villages of 15 Calder , 16 Grayson , 17 Spy Hill , 18 Neudorf and [6 19 Stockholm 6]

1: 5042916 Canada

2: 260799 Canadian_National_Railway

3: 3110683 Esterhazy,_Saskatchewan

1: 75763 Provinces_and_territories_of_Canada

[winner: 75763 Provinces_and_territories_of_Canada,
pem_pos=0, final_scores=-1.3728005,
final_local_scores=-1.5318279, pem=-0.7798718,
similarity=-37.38202]

gt: 75763 Provinces_and_territories_of_Canada,
pem_pos=0, final_scores=-1.3728005,
final_local_scores=-1.5318279, pem=-0.7798718,
similarity=-37.38202]

2: 667057 Electoral_district

[winner: 2198981 Electoral_districts_of_New_South_Wales,
pem_pos=2, final_scores=-0.9987764,
final_local_scores=-0.9750449, pem=-1.4789519,
similarity=-16.964802]

gt: 667057 Electoral_district, pem_pos=1,
final_scores=-1.4284718, final_local_scores=-1.5073636,
pem=-1.3563718, similarity=-33.516857]

3: 9013091 Legislative_Assembly_of_Saskatchewan

[winner: 9013091 Legislative_Assembly_of_Saskatchewan,
pem_pos=0, final_scores=-0.046147183,
final_local_scores=-0.029321201, pem=0.0,
similarity=3.2147098]

gt: 9013091 Legislative_Assembly_of_Saskatchewan,
pem_pos=0, final_scores=-0.046147183,
final_local_scores=-0.029321201, pem=0.0,
similarity=3.2147098]

4: 26840 Saskatchewan

[winner: 9013091 Legislative_Assembly_of_Saskatchewan,
pem_pos=0, final_scores=-0.046147183,

final_local_scores=-0.029321201, pem=0.0,
similarity=3.2147098|
RECALL MISS: 26840 Saskatchewan]

5: 28672742 Saltcoats_(provincial_electoral_district)
[winner: 506327 Saltcoats, pem_pos=0,
final_scores=0.19564326, final_local_scores=0.21861821,
pem=-0.108444534, similarity=11.203454|
gt: 28672742 Saltcoats_(provincial_electoral_district),
pem_pos=6, final_scores=-0.56749463,
final_local_scores=-0.5017879, pem=-3.2375245,
similarity=6.721523]

6: 9109290 Melville_(provincial_electoral_district)
[winner: 1034685 Melville, Saskatchewan, pem_pos=1,
final_scores=-0.029422179, final_local_scores=-0.07554218,
pem=-1.6235616, similarity=10.664564|
gt: 9109290 Melville_(provincial_electoral_district),
pem_pos=19, final_scores=-0.7075501,
final_local_scores=-0.6739621, pem=-3.9693713,
similarity=5.5633984]

7: 1034685 Melville, Saskatchewan
[winner: 1034685 Melville, Saskatchewan, pem_pos=1,
final_scores=0.013328418, final_local_scores=-0.028843671,
pem=-1.6235616, similarity=12.058141|
gt: 1034685 Melville, Saskatchewan, pem_pos=1,
final_scores=0.013328418, final_local_scores=-0.028843671,
pem=-1.6235616, similarity=12.058141]

8: 8620194 Roads_in_Saskatchewan
[winner: 2299280 Ontario_Highway_10, pem_pos=0,
final_scores=-0.18835528, final_local_scores=-0.15777358,
pem=0.0, similarity=-0.6185653|
RECALL MISS: 8620194 Roads_in_Saskatchewan]

9: 10528837 Saskatchewan_Highway_10
[winner: 2299280 Ontario_Highway_10, pem_pos=0,
final_scores=-0.18835528, final_local_scores=-0.15777358,
pem=0.0, similarity=-0.6185653|
RECALL MISS: 10528837 Saskatchewan_Highway_10]

10: 8742677 Saskatchewan_Highway_15
[winner: 15790 January_15, pem_pos=2,
final_scores=-0.9095679, final_local_scores=-0.81552577,
pem=-3.1689966, similarity=-3.0137072|
RECALL MISS: 8742677 Saskatchewan_Highway_15]

11: 12237985 Saskatchewan_Highway_47
[winner: 12237985 Saskatchewan_Highway_47, pem_pos=2,
final_scores=-0.9026004, final_local_scores=-0.9162329,
pem=-2.952993, similarity=-7.1936684|
gt: 12237985 Saskatchewan_Highway_47, pem_pos=2,
final_scores=-0.9026004, final_local_scores=-0.9162329,
pem=-2.952993, similarity=-7.1936684]

12: 7495447 Langenburg, Saskatchewan
[winner: 7495447 Langenburg, Saskatchewan, pem_pos=1,
final_scores=0.012943491, final_local_scores=0.03355047,
pem=-0.34229368, similarity=6.952372|

gt: 7495447 Langenburg, _Saskatchewan, pem_pos=1,
final_scores=0.012943491, final_local_scores=0.03355047,
pem=-0.34229368, similarity=6.952372]
13: 6264481 Saltcoats, _Saskatchewan
[winner: 506327 Saltcoats, pem_pos=0,
final_scores=0.11855261, final_local_scores=0.12916756,
pem=-0.108444534, similarity=8.534069]
gt: 6264481 Saltcoats, _Saskatchewan, pem_pos=1,
final_scores=-0.07403387, final_local_scores=-0.08694129,
pem=-1.0837123, similarity=7.3886075]
14: 13561417 Churchbridge, _Saskatchewan
[winner: 13561417 Churchbridge, _Saskatchewan, pem_pos=0,
final_scores=-0.03796205, final_local_scores=-0.024436615,
pem=0.0, similarity=3.3604755]
gt: 13561417 Churchbridge, _Saskatchewan, pem_pos=0,
final_scores=-0.03796205, final_local_scores=-0.024436615,
pem=0.0, similarity=3.3604755]
15: 12194026 Calder, _Saskatchewan
[winner: 30866045 Edmonton-Calder, pem_pos=2,
final_scores=-0.26432246, final_local_scores=-0.23968425,
pem=-2.3852634, similarity=9.90849]
gt: 12194026 Calder, _Saskatchewan, pem_pos=19,
final_scores=-0.60238546, final_local_scores=-0.6351601,
pem=-3.8091795, similarity=5.850179]
16: 15242339 Grayson, _Saskatchewan
[winner: 115084 Grayson, _Kentucky, pem_pos=0,
final_scores=-0.19516852, final_local_scores=-0.16973585,
pem=-1.4266951, similarity=6.7830467]
RECALL MISS: 15242339 Grayson, _Saskatchewan]
17: 25125830 Spy_Hill, _Saskatchewan
[winner: 25125830 Spy_Hill, _Saskatchewan, pem_pos=0,
final_scores=-0.048230812, final_local_scores=0.015846297,
pem=0.0, similarity=4.562598]
gt: 25125830 Spy_Hill, _Saskatchewan, pem_pos=0,
final_scores=-0.048230812, final_local_scores=0.015846297,
pem=0.0, similarity=4.562598]
18: 7718429 Neudorf, _Saskatchewan
[winner: 7718429 Neudorf, _Saskatchewan, pem_pos=1,
final_scores=0.01053603, final_local_scores=0.059023693,
pem=-0.42388186, similarity=8.156233]
gt: 7718429 Neudorf, _Saskatchewan, pem_pos=1,
final_scores=0.01053603, final_local_scores=0.059023693,
pem=-0.42388186, similarity=8.156233]
19: 25125889 Stockholm, _Saskatchewan
[winner: 26741 Stockholm, pem_pos=0,
final_scores=0.26777202, final_local_scores=0.3115566,
pem=-0.1047564, similarity=13.956865]
RECALL MISS: 25125889 Stockholm, _Saskatchewan]
1: 5042916 Canada
[pem_pos=0, final_scores=0.53619236,
final_local_scores=0.5423921, pem=-0.11662793,
similarity=20.910015]

2: 260799 Canadian_National_Railway
 [pem_pos=0, final_scores=0.30192822,
 final_local_scores=0.31095502, pem=0.0, similarity=13.369231]

3: 3110683 Esterhazy,_Saskatchewan
 [pem_pos=0, final_scores=0.15212451,
 final_local_scores=0.18084466, pem=-0.47893226,
 similarity=12.090986]

1: 26840 Saskatchewan
 [pem_pos=0, final_scores=0.49242705,
 final_local_scores=0.47057354, pem=-0.08728397,
 similarity=18.607231]

2: 26840 Saskatchewan
 [pem_pos=0, final_scores=0.1748284,
 final_local_scores=0.123645246, pem=-0.08728397,
 similarity=8.254197]

3: 506327 Saltcoats
 [pem_pos=0, final_scores=0.19564326,
 final_local_scores=0.21861821, pem=-0.108444534,
 similarity=11.203454]

4: 26840 Saskatchewan
 [pem_pos=0, final_scores=0.4105482,
 final_local_scores=0.38113332, pem=-0.08728397,
 similarity=15.938155]

5: 506327 Saltcoats
 [pem_pos=0, final_scores=0.11855261,
 final_local_scores=0.12916756, pem=-0.108444534,
 similarity=8.534069]

6: 26741 Stockholm
 [pem_pos=0, final_scores=0.26777202,
 final_local_scores=0.3115566, pem=-0.1047564,
 similarity=13.956865]

Sample: b'Division_of_Port_Adelaide'

The Division of [1 Port Adelaide 1] is an [2 1 Australian 2] electoral division in the 2 state of [1 1 South Australia 1] . It is located in the northern suburbs of [2 2 Adelaide 2] , covering the area around the [3 3 Barker Inlet 3] , part of the [4 4 Gulf St Vincent 4] . It stretches from [3 3 St Kilda 3] in the north down to , and is roughly bounded on the east by the [4 4 Gawler Central 4] railway line . It includes the suburbs of 5 Findon , 6 Woodville , [5 7 Hendon 5] , [6 8 Cro ydon 6] , [7 9 Exeter 7] , [8 10 North Haven 8] , 11 Wingfield , 12 Salisbury , 13 Buckland Park , [9 14 St Kilda 9] and [10 15 Port Adelaide 10] itself . The Division was named after the suburb of [11 Port Adelaide 11] , the working port of [12 Adelaide 12] . It was proclaimed at the redistribution of 11 May 1949 , and was first contested at the 16 1949 federal election . The seat is currently a very safe [5 5 Australian Lab or Party 5] seat with the closest the [13 ALP 13] has ever come to losing it was at the 17 1988 by-election , where [14 Labor 14] candidate [6 6 Rod Sawford 6] received 55.2 percent of the two-party vote . It currently stands , after the 18 201

0 vote , at 70.03 percent , making it the safest [15 Labor 15] seat in the state and the sixth-safest [16 Labor 16] seat in [17 Australia 17] , with a 20 percent swing required for the 19 Liberals to win it . [18 Port Adelaide 18] remains the only electorate in [19 South Australia 19] to have voted [20 Labor 20] at every federal election in its existence . A notable curiosity in recent years was that in the 20 1998 and 21 2001 federal elections , the seat was the only one in [21 Australia 21] where a 22 Communist Party candidate , Michael Perth , stood for election . This was the only occasion when the [22 Liberal Party 22] did not preference the [7 7 One Nation Party 7] last . He achieved less than 1 percent of the vote on each occasion . Sawford retired at the 23 2007 election , which saw [23 South Australian 23] Labor 's historically safe seat easily won by the newly endorsed [24 Labor 24] candidate , unionist and former head of the 24 Left state [25 Labor 25] faction 25 Mark Butler

1: 26716 South_Australia

2: 1148 Adelaide

3: 8613482 Barker_Inlet

4: 1251836 Gulf_St_Vincent

5: 1495 Australian_Labor_Party

6: 593376 Rod_Sawford

7: 83013 One_Nation_(Australia)

1: 984203 Divisions_of_the_Australian_House_of_Representatives
[winner: 4689264 Australia, pem_pos=0,
final_scores=0.5711932, final_local_scores=0.571812,
pem=-0.491009, similarity=23.823906]

RECALL MISS: 984203

Divisions_of_the_Australian_House_of_Representatives]
[winner: 1880929

Electoral_divisions_of_the_Northern_Territory, pem_pos=2,
final_scores=-0.9160335, final_local_scores=-0.91076624,
pem=-1.2819486, similarity=-16.117933]

gt: 984203

Divisions_of_the_Australian_House_of_Representatives,
pem_pos=3, final_scores=-1.3586605,
final_local_scores=-1.4851198, pem=-1.4548627,
similarity=-32.317455]

2: 4689096 States_and_territories_of_Australia

[winner: 229379 States_of_Brazil, pem_pos=0,
final_scores=-1.3079813, final_local_scores=-1.282754,
pem=0.0, similarity=-34.190224]

RECALL MISS: 4689096 States_and_territories_of_Australia]

3: 9031319 St_Kilda,_South_Australia

[winner: 92922 St_Kilda_Football_Club, pem_pos=0,
final_scores=0.19749935, final_local_scores=0.20274699,
pem=-0.48976773, similarity=12.80352]

gt: 9031319 St_Kilda,_South_Australia, pem_pos=8,
final_scores=-0.32961693, final_local_scores=-0.32977647,
pem=-2.407688, similarity=7.341907]

4: 3026256 Gawler_Central_railway_line

[winner: 3026256 Gawler_Central_railway_line, pem_pos=1,
 final_scores=0.27108902, final_local_scores=0.32914102,
 pem=-0.4840554, similarity=16.544306]
 gt: 3026256 Gawler_Central_railway_line, pem_pos=1,
 final_scores=0.27108902, final_local_scores=0.32914102,
 pem=-0.4840554, similarity=16.544306]
 [winner: 2335900 Color_Lines, pem_pos=1,
 final_scores=-1.2193702, final_local_scores=-1.1380169,
 pem=-1.974716, similarity=-19.132172]
 RECALL MISS: 3026256 Gawler_Central_railway_line]
 [winner: 19940036 Yarck_railway_station, pem_pos=14,
 final_scores=-1.7818333, final_local_scores=-1.8218424,
 pem=-4.538745, similarity=-25.5953]
 RECALL MISS: 3026256 Gawler_Central_railway_line]
 5: 12476797 Findon, South_Australia
 [winner: 2060949 Findon, West_Sussex, pem_pos=2,
 final_scores=-0.0154213905, final_local_scores=0.036471456,
 pem=-0.6385172, similarity=8.65045]
 gt: 12476797 Findon, South_Australia, pem_pos=1,
 final_scores=-0.07605958, final_local_scores=-0.049887456,
 pem=-0.6184664, similarity=5.9642878]
 6: 2950843 Woodville, South_Australia
 [winner: 136717 Woodville, Texas, pem_pos=1,
 final_scores=-0.28370607, final_local_scores=-0.23276135,
 pem=-1.9320654, similarity=7.6505218]
 gt: 2950843 Woodville, South_Australia, pem_pos=5,
 final_scores=-0.34588528, final_local_scores=-0.379307,
 pem=-2.6808064, similarity=7.3490777]
 7: 24011361 Hendon, South_Australia
 [winner: 94090 Hendon, pem_pos=0,
 final_scores=0.049419597, final_local_scores=0.038982004,
 pem=-0.5714235, similarity=8.360503]
 gt: 24011361 Hendon, South_Australia, pem_pos=6,
 final_scores=-0.41931254, final_local_scores=-0.41566896,
 pem=-2.6532276, similarity=6.1139855]
 8: 24011346 Croydon, South_Australia
 [winner: 491650 Croydon, pem_pos=0,
 final_scores=0.15915526, final_local_scores=0.14389288,
 pem=-0.5008817, similarity=11.107636]
 gt: 24011346 Croydon, South_Australia, pem_pos=16,
 final_scores=-0.58371806, final_local_scores=-0.59376067,
 pem=-3.6472366, similarity=6.204949]
 9: 3782000 Exeter, South_Australia
 [winner: 36614585 Exeter, pem_pos=0,
 final_scores=0.101948574, final_local_scores=0.08109367,
 pem=-0.56391937, similarity=9.576391]
 RECALL MISS: 3782000 Exeter, South_Australia]
 10: 3748383 North_Haven, South_Australia
 [winner: 530010 North_Haven, Connecticut, pem_pos=0,
 final_scores=0.15833522, final_local_scores=0.21960607,
 pem=-0.6041272, similarity=13.928534]
 gt: 3748383 North_Haven, South_Australia, pem_pos=4,

final_scores=-0.23648123, final_local_scores=-0.18922028,
pem=-1.585092, similarity=7.0629783]

11: 4360177 Wingfield, South_Australia
[winner: 4360177 Wingfield, South_Australia, pem_pos=1,
final_scores=-0.18014435, final_local_scores=-0.15480024,
pem=-1.5742475, similarity=8.031168]
gt: 4360177 Wingfield, South_Australia, pem_pos=1,
final_scores=-0.18014435, final_local_scores=-0.15480024,
pem=-1.5742475, similarity=8.031168]

12: 26419320 Salisbury, South_Australia
[winner: 72119 Salisbury, pem_pos=0,
final_scores=-0.019166544, final_local_scores=-0.047523722,
pem=-0.80015767, similarity=7.0228925]
gt: 26419320 Salisbury, South_Australia, pem_pos=7,
final_scores=-0.49441695, final_local_scores=-0.559343,
pem=-3.621498, similarity=7.0920725]

13: 27906123 Buckland_Park, South_Australia
[winner: 27906123 Buckland_Park, South_Australia,
pem_pos=0, final_scores=-0.05478023,
final_local_scores=0.039309993, pem=0.0,
similarity=5.2628016]
gt: 27906123 Buckland_Park, South_Australia, pem_pos=0,
final_scores=-0.05478023, final_local_scores=0.039309993,
pem=0.0, similarity=5.2628016]

14: 9031319 St_Kilda, South_Australia
[winner: 92922 St_Kilda_Football_Club, pem_pos=0,
final_scores=0.26393515, final_local_scores=0.27531803,
pem=-0.48976773, similarity=14.969185]
gt: 9031319 St_Kilda, South_Australia, pem_pos=8,
final_scores=-0.36269036, final_local_scores=-0.36590415,
pem=-2.407688, similarity=6.263785]

15: 1963141 Port_Adelaide
[winner: 24602 Port_Adelaide_Football_Club, pem_pos=0,
final_scores=0.18302463, final_local_scores=0.19398624,
pem=-0.52031404, similarity=12.708198]
gt: 1963141 Port_Adelaide, pem_pos=1,
final_scores=0.09157325, final_local_scores=0.04767549,
pem=-0.6625479, similarity=9.115483]

16: 10036999 Australian_federal_election, 1949
[winner: 21288296 1949_Baltimore_Colts_season, pem_pos=0,
final_scores=-0.7534327, final_local_scores=-0.61152875,
pem=0.0, similarity=-14.159524]
RECALL MISS: 10036999 Australian_federal_election, 1949]
[winner: 8203330 Swiss_federal_election, 2007, pem_pos=0,
final_scores=-0.9113172, final_local_scores=-0.85810083,
pem=-0.34350768, similarity=-19.64968]
RECALL MISS: 10036999 Australian_federal_election, 1949]

17: 24205681 Port_Adelaide_by-election, 1988
[winner: 24205765 Adelaide_by-election, 1988, pem_pos=0,
final_scores=-0.7469028, final_local_scores=-0.718607,
pem=-0.048790164, similarity=-17.089624]
gt: 24205681 Port_Adelaide_by-election, 1988, pem_pos=2,

final_scores=-1.1690478, final_local_scores=-1.1784163,
 pem=-1.4350846, similarity=-23.272366]
 18: 14449214 Australian_federal_election,_2010
 [winner: 43226 2010, pem_pos=0, final_scores=-0.9743661,
 final_local_scores=-0.86728096, pem=-1.6731367,
 similarity=-12.6929035]
 RECALL MISS: 14449214 Australian_federal_election,_2010]
 19: 18453 Liberal_Party_of_Australia
 [winner: 149536 Liberal_Party_of_Canada, pem_pos=0,
 final_scores=-0.104715094, final_local_scores=-0.10743007,
 pem=-1.3719091, similarity=8.34444]
 gt: 18453 Liberal_Party_of_Australia, pem_pos=2,
 final_scores=-0.1291589, final_local_scores=-0.20645347,
 pem=-2.037269, similarity=9.007716]
 20: 455788 Australian_federal_election,_1998
 [winner: 4669371 Juno_Awards_of_1998, pem_pos=0,
 final_scores=-0.72708344, final_local_scores=-0.61696,
 pem=0.0, similarity=-14.321604]
 RECALL MISS: 455788 Australian_federal_election,_1998]
 21: 454476 Australian_federal_election,_2001
 [winner: 34551 2001, pem_pos=0, final_scores=-0.55755705,
 final_local_scores=-0.34224665, pem=-0.5613713,
 similarity=-3.0707855]
 RECALL MISS: 454476 Australian_federal_election,_2001]
 [winner: 924152 Elections_in_Mexico, pem_pos=4,
 final_scores=-1.0147715, final_local_scores=-0.9562682,
 pem=-2.2547944, similarity=-12.185315]
 RECALL MISS: 454476 Australian_federal_election,_2001]
 22: 851181 Communist_Party_of_Australia_(current)
 [winner: 7271 Communist_Party_of_the_Soviet_Union,
 pem_pos=1, final_scores=9.7349286e-05,
 final_local_scores=0.0608685, pem=-1.814208,
 similarity=15.772097]
 RECALL MISS: 851181
 Communist_Party_of_Australia_(current)]
 23: 1056082 Australian_federal_election,_2007
 [winner: 36165 2007, pem_pos=0, final_scores=-0.60215265,
 final_local_scores=-0.49828812, pem=-0.24805899,
 similarity=-9.431215]
 RECALL MISS: 1056082 Australian_federal_election,_2007]
 [winner: 1442451 Elections_in_Rwanda, pem_pos=2,
 final_scores=-1.2610661, final_local_scores=-1.1885974,
 pem=-3.50929, similarity=-12.296338]
 gt: 1056082 Australian_federal_election,_2007,
 pem_pos=10, final_scores=-1.7347689,
 final_local_scores=-1.9375141, pem=-4.273487,
 similarity=-30.48968]
 24: 366524 Socialist_Left_(Australia)
 [winner: 1216632 Democratic_Left_Party_(Turkey),
 pem_pos=8, final_scores=-0.24721703,
 final_local_scores=-0.16282511, pem=-2.7027338,
 similarity=13.928573]

RECALL MISS: 366524 Socialist_Left_(Australia)]
25: 14484370 Mark_Butler
[winner: 14484370 Mark_Butler, pem_pos=0,
final_scores=-0.06820613, final_local_scores=-0.035736166,
pem=0.0, similarity=3.0232744]
gt: 14484370 Mark_Butler, pem_pos=0,
final_scores=-0.06820613, final_local_scores=-0.035736166,
pem=0.0, similarity=3.0232744]
1: 26716 South_Australia
[pem_pos=0, final_scores=0.5915695,
final_local_scores=0.5655279, pem=-0.05601515,
similarity=21.270813]
2: 1148 Adelaide
[pem_pos=0, final_scores=0.4322648,
final_local_scores=0.39045268, pem=-0.29570588,
similarity=17.349693]
3: 8613482 Barker_Inlet
[pem_pos=0, final_scores=0.14523052,
final_local_scores=0.23790708, pem=0.0, similarity=11.189335]
4: 1251836 Gulf_St_Vincent
[pem_pos=0, final_scores=0.25342172,
final_local_scores=0.31522232, pem=0.0, similarity=13.496575]
5: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.29506677,
final_local_scores=0.25431246, pem=0.0, similarity=11.678904]
6: 593376 Rod_Sawford
[pem_pos=0, final_scores=0.12440081,
final_local_scores=0.17868865, pem=0.0, similarity=9.422138]
7: 83013 One_Nation_(Australia)
[pem_pos=0, final_scores=0.3191753,
final_local_scores=0.34403425, pem=0.0, similarity=14.356381]
1: 1963141 Port_Adelaide
[pem_pos=1, final_scores=0.04950355,
final_local_scores=0.0012539327, pem=-0.6625479,
similarity=7.730171]
2: 4689264 Australia
[pem_pos=0, final_scores=0.5711932,
final_local_scores=0.571812, pem=-0.491009,
similarity=23.823906]
3: 92922 St_Kilda_Football_Club
[pem_pos=0, final_scores=0.19749935,
final_local_scores=0.20274699, pem=-0.48976773,
similarity=12.80352]
4: 3026256 Gawler_Central_railway_line
[pem_pos=1, final_scores=0.27108902,
final_local_scores=0.32914102, pem=-0.4840554,
similarity=16.544306]
5: 94090 Hendon
[pem_pos=0, final_scores=0.049419597,
final_local_scores=0.038982004, pem=-0.5714235,
similarity=8.360503]
6: 491650 Croydon

[pem_pos=0, final_scores=0.15915526,
final_local_scores=0.14389288, pem=-0.5008817,
similarity=11.107636]
7: 36614585 Exeter
[pem_pos=0, final_scores=0.101948574,
final_local_scores=0.08109367, pem=-0.56391937,
similarity=9.576391]
8: 530010 North_Haven,_Connecticut
[pem_pos=0, final_scores=0.15833522,
final_local_scores=0.21960607, pem=-0.6041272,
similarity=13.928534]
9: 92922 St_Kilda_Football_Club
[pem_pos=0, final_scores=0.26393515,
final_local_scores=0.27531803, pem=-0.48976773,
similarity=14.969185]
10: 24602 Port_Adelaide_Football_Club
[pem_pos=0, final_scores=0.18302463,
final_local_scores=0.19398624, pem=-0.52031404,
similarity=12.708198]
11: 1963141 Port_Adelaide
[pem_pos=1, final_scores=0.13888209,
final_local_scores=0.099353224, pem=-0.6625479,
similarity=10.65765]
12: 1148 Adelaide
[pem_pos=0, final_scores=0.41117758,
final_local_scores=0.3674181, pem=-0.29570588,
similarity=16.662296]
13: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.18214892,
final_local_scores=0.13096687, pem=-0.40684944,
similarity=10.210536]
14: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.30566967,
final_local_scores=0.26589453, pem=-0.2258703,
similarity=13.252853]
15: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.3537374,
final_local_scores=0.31840125, pem=-0.2258703,
similarity=14.819758]
16: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.3787158,
final_local_scores=0.34568632, pem=-0.2258703,
similarity=15.634]
17: 4689264 Australia
[pem_pos=0, final_scores=0.5301233,
final_local_scores=0.52694935, pem=-0.17168477,
similarity=20.74858]
18: 24602 Port_Adelaide_Football_Club
[pem_pos=0, final_scores=0.12252645,
final_local_scores=0.12790117, pem=-0.52031404,
similarity=10.736087]
19: 26716 South_Australia

```

    [pem_pos=0, final_scores=0.60111165,
final_local_scores=0.5759513, pem=-0.05601515,
similarity=21.581867]
20: 1495 Australian_Labor_Party
    [pem_pos=0, final_scores=0.22298704,
final_local_scores=0.17557636, pem=-0.2258703,
similarity=10.557579]
21: 4689264 Australia
    [pem_pos=0, final_scores=0.56195253,
final_local_scores=0.5617179, pem=-0.17168477,
similarity=21.786146]
22: 18453 Liberal_Party_of_Australia
    [pem_pos=1, final_scores=0.22768037,
final_local_scores=0.18932107, pem=-1.6500671,
similarity=18.712753]
23: 26716 South_Australia
    [pem_pos=0, final_scores=0.46638876,
final_local_scores=0.4269058, pem=0.0, similarity=16.829433]
24: 1495 Australian_Labor_Party
    [pem_pos=0, final_scores=0.2974916,
final_local_scores=0.25696123, pem=-0.2258703,
similarity=12.986266]
25: 1495 Australian_Labor_Party
    [pem_pos=0, final_scores=0.24510948,
final_local_scores=0.19974175, pem=-0.2258703,
similarity=11.278723]

```

Sample: b'Turkey_River'

_ -LRB- Iowa -RRB- The Turkey River is a tributary of the 1 upper [1 1 Mississippi River 1] . Its main branch rises in [1 2 Howard County 1] , Iowa , near the city of [2 2 Cresco 2] . The other counties it or its tributaries cover are [2 3 Chickasaw 2] , [3 3 Winneshiek 3] , 4 Fayette , 5 Clayton , [3 6 Delaware 3] , and [4 7 Dubuque 4] . Flowing from northwest to southeast , it flows through or near the cities of [4 4 Spillville 4] , [5 8 Fort Atkinson 5] , 9 Clermont , [6 10 Elgin 6] , [5 5 Elkader 5] , 11 Elkport , 12 Garber , and 13 Millville , where it is joined by the [7 Little Turkey River 7] before entering the [8 Mississippi 8] south of 14 Guttenberg and across from [9 15 Cassville 9] , [10 Wisconsin 10]

```

1: 19579 Mississippi_River
2: 113876 Cresco,_Iowa
3: 95905 Winneshiek_County,_Iowa
4: 114355 Spillville,_Iowa
5: 112903 Elkader,_Iowa

```

```

1: 4374200 Upper_Mississippi_River
    [winner: 540100 Upper_gastrointestinal_bleeding,
pem_pos=1, final_scores=-1.3158313,
final_local_scores=-1.2560809, pem=-1.637799,
similarity=-24.487637]
    gt: 4374200 Upper_Mississippi_River, pem_pos=13,
final_scores=-2.1866167, final_local_scores=-2.3971105,

```


pem=-3.3892846, similarity=-49.013386]
2: 95992 Howard_County,_Iowa
[winner: 47731 Howard_County,_Maryland, pem_pos=0,
final_scores=0.31023866, final_local_scores=0.3073256,
pem=-0.52799696, similarity=16.132252]
gt: 95992 Howard_County,_Iowa, pem_pos=6,
final_scores=-0.16361694, final_local_scores=-0.19933987,
pem=-2.5219548, similarity=11.855796]
[winner: 26810748 Iowa, pem_pos=0,
final_scores=0.0053563565, final_local_scores=-0.07537831,
pem=-0.3546882, similarity=3.7691233]
RECALL MISS: 95992 Howard_County,_Iowa]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7248226, final_local_scores=-1.6983306,
pem=-4.6728287, similarity=-21.18029]
RECALL MISS: 95992 Howard_County,_Iowa]
3: 96021 Chickasaw_County,_Iowa
[winner: 92809 Chickasaw, pem_pos=0,
final_scores=0.14341529, final_local_scores=0.15503639,
pem=-0.34369627, similarity=10.5853815]
gt: 96021 Chickasaw_County,_Iowa, pem_pos=12,
final_scores=-0.35406047, final_local_scores=-0.40317678,
pem=-2.9754243, similarity=8.238933]
4: 96005 Fayette_County,_Iowa
[winner: 91959 Fayette_County,_Pennsylvania, pem_pos=1,
final_scores=-0.062284052, final_local_scores=-0.11931151,
pem=-2.3120737, similarity=13.102636]
gt: 96005 Fayette_County,_Iowa, pem_pos=8,
final_scores=-0.26476282, final_local_scores=-0.31864086,
pem=-2.7793872, similarity=9.695574]
5: 96016 Clayton_County,_Iowa
[winner: 124840 Clayton,_Missouri, pem_pos=1,
final_scores=-0.13578439, final_local_scores=-0.18659198,
pem=-2.4868636, similarity=12.045387]
gt: 96016 Clayton_County,_Iowa, pem_pos=10,
final_scores=-0.46780097, final_local_scores=-0.5506922,
pem=-3.5371408, similarity=6.8914814]
6: 96011 Delaware_County,_Iowa
[winner: 7930 Delaware, pem_pos=0,
final_scores=0.3899302, final_local_scores=0.3704903,
pem=-0.5414915, similarity=18.090597]
RECALL MISS: 96011 Delaware_County,_Iowa]
7: 96007 Dubuque_County,_Iowa
[winner: 60796 Dubuque,_Iowa, pem_pos=0,
final_scores=0.20079385, final_local_scores=0.16989076,
pem=-0.19283246, similarity=10.208245]
gt: 96007 Dubuque_County,_Iowa, pem_pos=1,
final_scores=-0.13366675, final_local_scores=-0.20760907,
pem=-1.5713495, similarity=6.439487]
8: 114351 Fort_Atkinson,_Iowa
[winner: 139181 Fort_Atkinson,_Wisconsin, pem_pos=1,
final_scores=0.23333429, final_local_scores=0.21978378,

pem=-0.6707949, similarity=14.296387|
gt: 114351 Fort Atkinson, Iowa, pem_pos=2,
final_scores=-0.02577199, final_local_scores=-0.031018108,
pem=-1.0462121, similarity=8.853535]
9: 113771 Clermont, Iowa
[winner: 95434 Clermont County, Ohio, pem_pos=2,
final_scores=-0.16263285, final_local_scores=-0.23834822,
pem=-2.2316206, similarity=9.112826|
gt: 113771 Clermont, Iowa, pem_pos=10,
final_scores=-0.3430559, final_local_scores=-0.38043627,
pem=-2.7550738, similarity=7.719256]
10: 113772 Elgin, Iowa
[winner: 110972 Elgin, Illinois, pem_pos=0,
final_scores=0.095884666, final_local_scores=0.0667211,
pem=-1.0776671, similarity=11.941325|
gt: 113772 Elgin, Iowa, pem_pos=3,
final_scores=-0.613324, final_local_scores=-0.6561172,
pem=-3.0518885, similarity=1.10651]
11: 112904 Elkport, Iowa
[winner: 112904 Elkport, Iowa, pem_pos=0,
final_scores=-0.036916047, final_local_scores=-0.0066372305,
pem=0.0, similarity=3.8916447|
gt: 112904 Elkport, Iowa, pem_pos=0,
final_scores=-0.036916047, final_local_scores=-0.0066372305,
pem=0.0, similarity=3.8916447]
12: 112906 Garber, Iowa
[winner: 11755370 Milton C. Garber, pem_pos=2,
final_scores=-0.14572248, final_local_scores=-0.100313455,
pem=-1.5761967, similarity=9.667762|
gt: 112906 Garber, Iowa, pem_pos=0,
final_scores=-0.22592434, final_local_scores=-0.2090034,
pem=-0.84275603, similarity=2.4356728]
13: 112913 Millville, Iowa
[winner: 125146 Millville, New Jersey, pem_pos=0,
final_scores=-0.04696341, final_local_scores=-0.07184237,
pem=-1.5476592, similarity=10.362205|
gt: 112913 Millville, Iowa, pem_pos=5,
final_scores=-0.37390238, final_local_scores=-0.38693625,
pem=-2.262346, similarity=4.845753]
14: 112908 Guttenberg, Iowa
[winner: 125233 Guttenberg, New Jersey, pem_pos=1,
final_scores=0.029857561, final_local_scores=0.036592215,
pem=-0.8673669, similarity=9.898574|
gt: 112908 Guttenberg, Iowa, pem_pos=2,
final_scores=-0.06833632, final_local_scores=-0.10839948,
pem=-1.5062919, similarity=9.046306]
15: 260439 Cassville, Wisconsin
[winner: 33127 Wisconsin, pem_pos=0,
final_scores=0.43041682, final_local_scores=0.40434086,
pem=-0.21268336, similarity=17.312656|
RECALL MISS: 260439 Cassville, Wisconsin]
[winner: 122501 Cassville, Missouri, pem_pos=0,

final_scores=0.074502334, final_local_scores=0.07347241,
pem=-0.7912478, similarity=10.585204|
gt: 260439 Cassville,_Wisconsin, pem_pos=1,
final_scores=0.00577797, final_local_scores=-0.018394053,
pem=-1.3848889, similarity=11.072039]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7029312, final_local_scores=-1.6744175,
pem=-4.6728287, similarity=-20.466673|
RECALL MISS: 260439 Cassville,_Wisconsin]

1: 19579 Mississippi_River
[pem_pos=0, final_scores=0.395514,
final_local_scores=0.38213927, pem=0.0, similarity=15.49351]

2: 113876 Cresco,_Iowa
[pem_pos=0, final_scores=0.205649,
final_local_scores=0.21212426, pem=-0.092310786,
similarity=10.921925]

3: 95905 Winneshiek_County,_Iowa
[pem_pos=1, final_scores=0.17746983,
final_local_scores=0.15044886, pem=-0.0964603,
similarity=9.103973]

4: 114355 Spillville,_Iowa
[pem_pos=0, final_scores=0.052709296,
final_local_scores=0.08189088, pem=0.0, similarity=6.5335]

5: 112903 Elkader,_Iowa
[pem_pos=0, final_scores=0.07226251,
final_local_scores=0.057972163, pem=0.0, similarity=5.819718]

1: 47731 Howard_County,_Maryland
[pem_pos=0, final_scores=0.31023866,
final_local_scores=0.3073256, pem=-0.52799696,
similarity=16.132252]

2: 92809 Chickasaw
[pem_pos=0, final_scores=0.14341529,
final_local_scores=0.15503639, pem=-0.34369627,
similarity=10.5853815]

3: 7930 Delaware
[pem_pos=0, final_scores=0.3899302,
final_local_scores=0.3704903, pem=-0.5414915,
similarity=18.090597]

4: 60796 Dubuque,_Iowa
[pem_pos=0, final_scores=0.20079385,
final_local_scores=0.16989076, pem=-0.19283246,
similarity=10.208245]

5: 139181 Fort_Atkinson,_Wisconsin
[pem_pos=1, final_scores=0.23333429,
final_local_scores=0.21978378, pem=-0.6707949,
similarity=14.296387]

6: 110972 Elgin,_Illinois
[pem_pos=0, final_scores=0.095884666,
final_local_scores=0.0667211, pem=-1.0776671,
similarity=11.941325]

7: 10267784 Turkey_River_(Iowa)
[pem_pos=0, final_scores=0.047629103,

```

final_local_scores=0.092748895, pem=0.0, similarity=6.857525]
8: 19579 Mississippi_River
    [pem_pos=1, final_scores=0.2078164,
final_local_scores=0.18010986, pem=-2.3293974,
similarity=22.132175]
9: 122501 Cassville, Missouri
    [pem_pos=0, final_scores=0.074502334,
final_local_scores=0.07347241, pem=-0.7912478,
similarity=10.585204]
10: 33127 Wisconsin
    [pem_pos=0, final_scores=0.43041682,
final_local_scores=0.40434086, pem=-0.21268336,
similarity=17.312656]

```

Sample: b'Country_blues'

Country blues otherwise known as acoustic blues -LRB- also folk blues , rural blues , backwoods blues , or downhome blues -RRB- is a general term that refers to all the acoustic , mainly **1 guitar** - driven forms of the **2 blues** . It often incorporated elements of rural gospel , ragtime , hillbilly , and dixieland jazz . After blues ' birth in the **[1 1 Southern United States 1]** , it quickly spread throughout the country -LRB- and elsewhere -RRB- , giving birth to a host of regional styles . These include **[1 3 Memphis 1]** , **[2 4 Detroit 2]** , **[3 5 Chicago 3]** , **[4 6 Texas 4]** , **7 Piedmont** , **[5 8 Louisiana 5]** , **[6 9 West Coast 6]** , , **[7 10 St. Louis 7]** , **[8 11 East Coast 8]** , **12 Swamp** , **[9 13 New Orleans 9]** , **14 Delta** , **15 Hill country** and blues . When **[10 African-American 10]** musical tastes began to change in the early 1960s , moving toward **16 soul** and **17 rhythm and blues** music , country blues found renewed popularity as `` folk blues '' and was sold to a primarily white , college-age audience . Traditional artists like **[2 2 Big Bill Broonzy 2]** and **18 Sonny Boy Williamson II** reinvented themselves as folk blues artists , while Piedmont bluesmen like **19 Sonny Terry** and **20 Brownie McGhee** found great success on the **21 folk festival circuit**

1: 179553 Southern_United_States

2: 237960 Big_Bill_Broonzy

1: 11846 Guitar

```

[winner: 9611790 Epiphone_Supernova, pem_pos=8,
final_scores=-1.2425187, final_local_scores=-1.2185822,
pem=-2.0793161, similarity=-20.967567]

```

```

gt: 11846 Guitar, pem_pos=0, final_scores=-1.5791534,
final_local_scores=-1.7029488, pem=-0.13780104,
similarity=-45.980297 ]

```

2: 3352 Blues

```

[winner: 6928611 Nina_Simone_Sings_the_Blues, pem_pos=18,
final_scores=-1.1172875, final_local_scores=-1.045121,
pem=-3.6109178, similarity=-7.462048]

```

```

gt: 3352 Blues, pem_pos=0, final_scores=-1.6178029,
final_local_scores=-1.7912395, pem=-1.0090724,
similarity=-43.876953 ]

```

3: 171833 Memphis_blues
[winner: 48607 Memphis, _Tennessee, pem_pos=0,
final_scores=0.3968838, final_local_scores=0.3419837,
pem=-0.64647895, similarity=17.810843|
RECALL MISS: 171833 Memphis_blues]

4: 171924 Detroit_blues
[winner: 8687 Detroit, pem_pos=0,
final_scores=0.51092213, final_local_scores=0.47804862,
pem=-0.21700948, similarity=19.535769|
RECALL MISS: 171924 Detroit_blues]

5: 171903 Chicago_blues
[winner: 6886 Chicago, pem_pos=0, final_scores=0.6511561,
final_local_scores=0.6259478, pem=-0.20658566,
similarity=23.892689|
RECALL MISS: 171903 Chicago_blues]

6: 989150 Texas_blues
[winner: 29810 Texas, pem_pos=0, final_scores=0.73679453,
final_local_scores=0.7462882, pem=-0.16191302,
similarity=27.24095|
RECALL MISS: 989150 Texas_blues]

7: 172308 Piedmont_blues
[winner: 25061 Piedmont, pem_pos=0,
final_scores=-0.046131864, final_local_scores=-0.007653117,
pem=-0.6027155, similarity=7.138989|
gt: 172308 Piedmont_blues, pem_pos=8,
final_scores=-0.34000832, final_local_scores=-0.37374827,
pem=-3.1706498, similarity=10.178805]

8: 172220 Louisiana_blues
[winner: 18130 Louisiana, pem_pos=0,
final_scores=0.4895231, final_local_scores=0.45885122,
pem=-0.17745006, similarity=18.74775|
RECALL MISS: 172220 Louisiana_blues]

9: 172230 West_Coast_blues
[winner: 174579 West_Coast_of_the_United_States,
pem_pos=0, final_scores=0.051444724,
final_local_scores=-0.002462551, pem=-0.59504014,
similarity=7.2521462|
gt: 172230 West_Coast_blues, pem_pos=17,
final_scores=-0.46291322, final_local_scores=-0.45389652,
pem=-2.982977, similarity=6.766427]

10: 172217 St._Louis_blues_(music)
[winner: 27687 St._Louis, pem_pos=0,
final_scores=0.35501087, final_local_scores=0.3046592,
pem=-0.28531665, similarity=14.732947|
RECALL MISS: 172217 St._Louis_blues_(music)]

11: 3239076 East_Coast_blues
[winner: 89126 East_Coast_of_the_United_States,
pem_pos=0, final_scores=0.07164849,
final_local_scores=0.03960003, pem=-0.33207393,
similarity=7.0773263|
gt: 3239076 East_Coast_blues, pem_pos=13,
final_scores=-0.38875082, final_local_scores=-0.34039557,

pem=-2.6376746, similarity=8.275714]
12: 172221 Swamp_blues
[winner: 86359 Swamp, pem_pos=0,
final_scores=-0.17395496, final_local_scores=-0.11530235,
pem=-0.28581113, similarity=2.2031462]
gt: 172221 Swamp_blues, pem_pos=4,
final_scores=-0.59259284, final_local_scores=-0.572788,
pem=-2.468193, similarity=0.41899425]
13: 7286753 New_Orleans_blues
[winner: 53842 New_Orleans, pem_pos=0,
final_scores=0.43477792, final_local_scores=0.38263297,
pem=-0.12199242, similarity=16.171658]
RECALL MISS: 7286753 New_Orleans_blues]
14: 171907 Delta_blues
[winner: 77549 Delta_Air_Lines, pem_pos=0,
final_scores=-0.22489049, final_local_scores=-0.22492,
pem=-2.134003, similarity=8.982691]
gt: 171907 Delta_blues, pem_pos=18,
final_scores=-0.6034784, final_local_scores=-0.67181796,
pem=-3.8186269, similarity=4.8076115]
15: 32171692 Hill_country_blues
[winner: 1408983 Hill, pem_pos=0,
final_scores=-0.50430626, final_local_scores=-0.44518256,
pem=-2.1844907, similarity=2.6841772]
RECALL MISS: 32171692 Hill_country_blues]
[winner: 5247 Country_music, pem_pos=0,
final_scores=-1.9843439, final_local_scores=-2.1768014,
pem=-0.57029057, similarity=-57.76906]
RECALL MISS: 32171692 Hill_country_blues]
16: 62808 Soul_music
[winner: 62808 Soul_music, pem_pos=0,
final_scores=-1.2477283, final_local_scores=-1.3671608,
pem=0.0, similarity=-36.709095]
gt: 62808 Soul_music, pem_pos=0,
final_scores=-1.2477283, final_local_scores=-1.3671608,
pem=0.0, similarity=-36.709095]
17: 26168 Rhythm_and_blues
[winner: 23488452 Upton_Blues_Festival, pem_pos=3,
final_scores=-0.9907495, final_local_scores=-0.9372844,
pem=-1.0985211, similarity=-17.906796]
gt: 26168 Rhythm_and_blues, pem_pos=15,
final_scores=-2.6578882, final_local_scores=-2.9075344,
pem=-8.322789, similarity=-37.416286]
[winner: 26227 Rhythm, pem_pos=0,
final_scores=-1.1633711, final_local_scores=-1.1992824,
pem=-0.1975527, similarity=-30.62494]
gt: 26168 Rhythm_and_blues, pem_pos=12,
final_scores=-2.3377142, final_local_scores=-2.557793,
pem=-6.6461544, similarity=-36.097107]
18: 296929 Sonny_Boy_Williamson_II
[winner: 296929 Sonny_Boy_Williamson_II, pem_pos=0,
final_scores=-0.058999807, final_local_scores=-0.051014297,

pem=0.0, similarity=2.5673447|
gt: 296929 Sonny_Boy_Williamson_II, pem_pos=0,
final_scores=-0.058999807, final_local_scores=-0.051014297,
pem=0.0, similarity=2.5673447]
19: 758608 Sonny_Terry
[winner: 758608 Sonny_Terry, pem_pos=0,
final_scores=-0.08927767, final_local_scores=-0.08233025,
pem=0.0, similarity=1.632814|
gt: 758608 Sonny_Terry, pem_pos=0,
final_scores=-0.08927767, final_local_scores=-0.08233025,
pem=0.0, similarity=1.632814]
20: 4737 Brownie_McGhee
[winner: 4737 Brownie_McGhee, pem_pos=0,
final_scores=-0.0319847, final_local_scores=-0.01198408,
pem=0.0, similarity=3.7320843|
gt: 4737 Brownie_McGhee, pem_pos=0,
final_scores=-0.0319847, final_local_scores=-0.01198408,
pem=0.0, similarity=3.7320843]
21: 8928146 American_folk_music_revival
[winner: 8866648 The_Nodding_Folk, pem_pos=0,
final_scores=-0.7224191, final_local_scores=-0.56810087,
pem=0.0, similarity=-12.86355|
RECALL MISS: 8928146 American_folk_music_revival]
[winner: 1640015 Terrastock, pem_pos=1,
final_scores=-1.0348313, final_local_scores=-1.0005552,
pem=-0.117783025, similarity=-25.128326|
RECALL MISS: 8928146 American_folk_music_revival]
1: 179553 Southern_United_States
[pem_pos=0, final_scores=0.35396504,
final_local_scores=0.32427546, pem=0.0, similarity=13.76674]
2: 237960 Big_Bill_Broonzy
[pem_pos=0, final_scores=0.1382388,
final_local_scores=0.16629699, pem=0.0, similarity=9.052347]
1: 48607 Memphis,_Tennessee
[pem_pos=0, final_scores=0.3968838,
final_local_scores=0.3419837, pem=-0.64647895,
similarity=17.810843]
2: 8687 Detroit
[pem_pos=0, final_scores=0.51092213,
final_local_scores=0.47804862, pem=-0.21700948,
similarity=19.535769]
3: 6886 Chicago
[pem_pos=0, final_scores=0.6511561,
final_local_scores=0.6259478, pem=-0.20658566,
similarity=23.892689]
4: 29810 Texas
[pem_pos=0, final_scores=0.73679453,
final_local_scores=0.7462882, pem=-0.16191302,
similarity=27.24095]
5: 18130 Louisiana
[pem_pos=0, final_scores=0.4895231,
final_local_scores=0.45885122, pem=-0.17745006,

```

similarity=18.74775]
6: 174579 West_Coast_of_the_United_States
   [pem_pos=0, final_scores=0.051444724,
   final_local_scores=-0.002462551, pem=-0.59504014,
   similarity=7.2521462]
7: 27687 St._Louis
   [pem_pos=0, final_scores=0.35501087,
   final_local_scores=0.3046592, pem=-0.28531665,
   similarity=14.732947]
8: 89126 East_Coast_of_the_United_States
   [pem_pos=0, final_scores=0.07164849,
   final_local_scores=0.03960003, pem=-0.33207393,
   similarity=7.0773263]
9: 53842 New_Orleans
   [pem_pos=0, final_scores=0.43477792,
   final_local_scores=0.38263297, pem=-0.12199242,
   similarity=16.171658]
10: 2154 African_American
    [pem_pos=0, final_scores=0.27424353,
    final_local_scores=0.2274692, pem=0.0, similarity=10.877848]

```

Sample: b'Abbey_Park'

, _ Nottinghamshire Abbey Park in is a 1 housing estate that was constructed in the mid -- to late -- 70s by [1 1 Costain Homes 1] on land that were previously uncultivated fields and allotments . The Willow Tree was constructed at the same time . All the roads are named after famous 2 abbey s and are laid out in a tight , twisty manner that precludes fast driving . The estate is divided in two , with the majority of homes being constructed for private purchase , while a smaller number were constructed for 3 council 4 tenants . There is a mix of three - , four - and five-bedroom homes on the private-owned estate in 5 townhouse , 6 semi-detached , link detached , and detached configurations , with a small number of 7 bungalow s . This part of [1 Abbey Park 1] consists of 8 Bisham Drive , 9 Dale Close , 10 Newstead Drive , Drive , 11 Fountains Close , Close , and Drive . The houses are all arranged in 12 cul-de-sac s . The two through-roads on the estate are Rufford Way and Buckfast Way . The council-funded area consists of 13 Tewkesbury Close , 14 Hexham Close , 15 Bolton Close , and 16 Furness Close , and includes some warden-aided 17 accommodation for the elderly . Abbey Park is regarded as a pleasant and quiet area of [2 West Bridgford 2] , and 18 house prices

1: 3211635 Redrow_plc

1: 300425 Housing_estate

```

   [winner: 300425 Housing_estate, pem_pos=0,
   final_scores=-1.1086061, final_local_scores=-1.2215455,
   pem=-0.001540834, similarity=-32.355263]

```

```

   gt: 300425 Housing_estate, pem_pos=0,
   final_scores=-1.1086061, final_local_scores=-1.2215455,
   pem=-0.001540834, similarity=-32.355263 ]

```


2: 1315 Abbey
[winner: 1315 Abbey, pem_pos=0, final_scores=-0.9268973, final_local_scores=-0.9696683, pem=-0.7359849, similarity=-20.844723]
gt: 1315 Abbey, pem_pos=0, final_scores=-0.9268973, final_local_scores=-0.9696683, pem=-0.7359849, similarity=-20.844723]
3: 253412 Council_house
[winner: 253412 Council_house, pem_pos=0, final_scores=-1.1254909, final_local_scores=-1.2348418, pem=0.0, similarity=-32.760433]
gt: 253412 Council_house, pem_pos=0, final_scores=-1.1254909, final_local_scores=-1.2348418, pem=0.0, similarity=-32.760433]
4: 863165 Leasehold_estate
[winner: 253412 Council_house, pem_pos=0, final_scores=-1.1254909, final_local_scores=-1.2348418, pem=0.0, similarity=-32.760433]
RECALL MISS: 863165 Leasehold_estate]
5: 334977 Townhouse
[winner: 334977 Townhouse, pem_pos=0, final_scores=-1.0882366, final_local_scores=-1.1620514, pem=-0.021739965, similarity=-30.469997]
gt: 334977 Townhouse, pem_pos=0, final_scores=-1.0882366, final_local_scores=-1.1620514, pem=-0.021739965, similarity=-30.469997]
6: 356265 Semi-detached
[winner: 356265 Semi-detached, pem_pos=0, final_scores=-1.0317348, final_local_scores=-1.1293893, pem=-0.0010598551, similarity=-29.607756]
gt: 356265 Semi-detached, pem_pos=0, final_scores=-1.0317348, final_local_scores=-1.1293893, pem=-0.0010598551, similarity=-29.607756]
7: 323262 Bungalow
[winner: 1858116 California_bungalow, pem_pos=0, final_scores=-1.206397, final_local_scores=-1.2450588, pem=0.0, similarity=-33.065327]
gt: 323262 Bungalow, pem_pos=1, final_scores=-1.3130763, final_local_scores=-1.4022458, pem=-0.037490584, similarity=-37.55222]
8: 113540 Bisham_Abbey
[winner: 3275711 Bisham, pem_pos=0, final_scores=-0.6390211, final_local_scores=-0.67510396, pem=0.0, similarity=-16.056736]
gt: 113540 Bisham_Abbey, pem_pos=1, final_scores=-0.92684406, final_local_scores=-0.98105145, pem=-0.67951065, similarity=-21.491533]
9: 8787755 Dale_Abbey
[winner: 1522138 Dale,_Pembrokeshire, pem_pos=4, final_scores=-1.179414, final_local_scores=-1.2079966, pem=-3.038337, similarity=-15.436364]
RECALL MISS: 8787755 Dale_Abbey]

10: 145954 Newstead_Abbey
 [winner: 15097922 Shire_of_Newstead, pem_pos=1,
 final_scores=-0.9146202, final_local_scores=-0.86491233,
 pem=-1.0740353, similarity=-15.880227|
 gt: 145954 Newstead_Abbey, pem_pos=8,
 final_scores=-1.1865032, final_local_scores=-1.25513,
 pem=-2.429593, similarity=-20.153368]

11: 63417 Fountains_Abbey
 [winner: 63417 Fountains_Abbey, pem_pos=0,
 final_scores=-0.86012155, final_local_scores=-0.9109347,
 pem=0.0, similarity=-23.094395|
 gt: 63417 Fountains_Abbey, pem_pos=0,
 final_scores=-0.86012155, final_local_scores=-0.9109347,
 pem=0.0, similarity=-23.094395]

12: 390697 Cul-de-sac
 [winner: 390697 Cul-de-sac, pem_pos=1,
 final_scores=-1.3143014, final_local_scores=-1.4007587,
 pem=-0.02549796, similarity=-37.573063|
 gt: 390697 Cul-de-sac, pem_pos=1,
 final_scores=-1.3143014, final_local_scores=-1.4007587,
 pem=-0.02549796, similarity=-37.573063]

13: 72955 Tewkesbury_Abbey
 [winner: 308993 Tewkesbury, pem_pos=0,
 final_scores=-0.9890874, final_local_scores=-1.0991148,
 pem=-0.15949842, similarity=-27.842688|
 gt: 72955 Tewkesbury_Abbey, pem_pos=4,
 final_scores=-1.3507129, final_local_scores=-1.4445839,
 pem=-2.3078892, similarity=-26.468887]

14: 477919 Hexham_Abbey
 [winner: 3896461 Hexham_railway_station, pem_pos=3,
 final_scores=-0.9276341, final_local_scores=-0.927661,
 pem=-2.2206583, similarity=-11.517256|
 gt: 477919 Hexham_Abbey, pem_pos=6,
 final_scores=-1.3626542, final_local_scores=-1.4182513,
 pem=-2.266433, similarity=-25.908516]

15: 142156 Bolton_Abbey
 [winner: 5311910 Bolton_Council_election,_2004,
 pem_pos=16, final_scores=-1.1730839,
 final_local_scores=-1.1034307, pem=-4.1894255,
 similarity=-6.0561123|
 RECALL MISS: 142156 Bolton_Abbey]

16: 1362016 Furness_Abbey
 [winner: 784155 Furness, pem_pos=0,
 final_scores=-1.0483257, final_local_scores=-1.1281729,
 pem=-0.2890824, similarity=-28.00514|
 gt: 1362016 Furness_Abbey, pem_pos=5,
 final_scores=-1.4356678, final_local_scores=-1.549263,
 pem=-3.0254266, similarity=-25.690647]

17: 1563666 Retirement_home
 [winner: 26657251 Lighting_for_the_elderly, pem_pos=0,
 final_scores=-0.9420595, final_local_scores=-0.8351799,
 pem=0.0, similarity=-20.833721|

RECALL MISS: 1563666 Retirement_home]
 [winner: 1157448 Accommodation_reflex, pem_pos=0,
 final_scores=-1.0021551, final_local_scores=-0.89132905,
 pem=0.0, similarity=-22.509323]
 RECALL MISS: 1563666 Retirement_home]
 18: 4348258 Real_estate_pricing
 [winner: 72902 Market_price, pem_pos=0,
 final_scores=-1.0741683, final_local_scores=-1.045417,
 pem=0.0, similarity=-27.107615]
 RECALL MISS: 4348258 Real_estate_pricing]
 [winner: 16752663 House,_North_Carolina, pem_pos=16,
 final_scores=-1.3208945, final_local_scores=-1.2600839,
 pem=-4.8077354, similarity=-7.368492]
 RECALL MISS: 4348258 Real_estate_pricing]
 1: 3211635 Redrow_plc
 [pem_pos=0, final_scores=0.22306572,
 final_local_scores=0.32419232, pem=0.0, similarity=13.764258]
 1: 373781 Abbey_Park,_Leicester
 [pem_pos=1, final_scores=0.12574442,
 final_local_scores=0.20753694, pem=-0.49850512,
 similarity=12.993978]
 2: 771806 West_Bridgford
 [pem_pos=0, final_scores=0.27000695,
 final_local_scores=0.30899778, pem=0.0, similarity=13.310823]

Sample: b'Alessandro_Gramigni'

Alessandro Gramigni -LRB- born on December 29 , 1968 in -
 RRB- is a former 1 Grand Prix 2 motorcycle 3 road racer . He
 was the 4 1992 5 F.I.M. . After a succession of good results
 as an amateur , [1 Gramigni 1] made his Grand Prix debut in 6
 1990 riding an [1 1 Aprilia 1] in the 125cc Class , ending
 the season in ninth position . He won his first race at the 7
 1991 8 Czechoslovakian Grand Prix and finished the season
 ranked seventh in the championship with 90 points . In 9 1992
 , he edged out [2 Italian 2] [2 2 Fausto Gresini 2] to win
 the 125cc world championship , with wins in [3 10 Malaysia 3]
 and [4 11 Hungary 4] . [5 Gramigni 5] moved up to the 250cc
 class in 12 next season , first with the [3 3 Gilera 3] team
 before switching back to [6 Aprilia 6] . It was a
 disappointing season , as he scored only 2 points . He
 competed two more seasons in the 250 class , 13 1994 with the
 [7 Aprilia 7] , and in 14 1995 with [4 4 Honda 4] . After a
 two year sabbatical , he entered one race in the class at the
 15 1997 16 Malaysian Grand Prix before ending his Grand Prix
 career . He turned his attention to the [5 5 Superbike World C
 hampionship 5] from to and then in and . In 2004 he won the
 Italian Superbike Championship with the [6 6 Yamaha 6]

1: 75753 Aprilia
 2: 5585957 Fausto_Gresini
 3: 549741 Gilera
 4: 13729 Honda
 5: 13247557 Superbike_World_Championship

6: 332378 Yamaha_Motor_Company

1: 54221 Grand_Prix_motorcycle_racing

[winner: 54221 Grand_Prix_motorcycle_racing, pem_pos=1, final_scores=-0.1224881, final_local_scores=-0.18498938, pem=-1.5994394, similarity=7.2672606]

gt: 54221 Grand_Prix_motorcycle_racing, pem_pos=1, final_scores=-0.1224881, final_local_scores=-0.18498938, pem=-1.5994394, similarity=7.2672606]

2: 19876 Motorcycle

[winner: 2491735 Phelon_&_Moore, pem_pos=5, final_scores=-1.372218, final_local_scores=-1.4306595, pem=-1.7917595, similarity=-28.860153]

gt: 19876 Motorcycle, pem_pos=0, final_scores=-1.6447015, final_local_scores=-1.8381405, pem=-0.031461716, similarity=-50.592968]

3: 685673 Road_racing

[winner: 2619936 Racer_(Kennywood), pem_pos=2, final_scores=-0.8577262, final_local_scores=-0.78924215, pem=-1.4941926, similarity=-11.337195]

RECALL MISS: 685673 Road_racing]

[winner: 20324545 Jalan_Gunung_Brinchang, pem_pos=8, final_scores=-1.3513532, final_local_scores=-1.2800112, pem=-3.265966, similarity=-16.347542]

RECALL MISS: 685673 Road_racing]

4: 8046362 1992_Grand_Prix_motorcycle_racing_season

[winner: 34669 1992, pem_pos=0, final_scores=-0.9028079, final_local_scores=-0.77222145, pem=-0.76294684, similarity=-14.805892]

RECALL MISS: 8046362

1992_Grand_Prix_motorcycle_racing_season]

5: 664030 Fédération_Internationale_de_Motocyclisme

[winner: 664030

Fédération_Internationale_de_Motocyclisme, pem_pos=0, final_scores=0.009396628, final_local_scores=-0.013958886, pem=0.0, similarity=3.673152]

gt: 664030 Fédération_Internationale_de_Motocyclisme, pem_pos=0, final_scores=0.009396628, final_local_scores=-0.013958886, pem=0.0, similarity=3.673152]

6: 7999612 1990_Grand_Prix_motorcycle_racing_season

[winner: 4034936 1990_Wimbledon_Championships_-_Men's_Singles, pem_pos=0, final_scores=-0.724333, final_local_scores=-0.574401, pem=-1.3466852, similarity=-5.7280765]

RECALL MISS: 7999612

1990_Grand_Prix_motorcycle_racing_season]

7: 8030250 1991_Grand_Prix_motorcycle_racing_season

[winner: 34703 1991, pem_pos=0, final_scores=-1.1784055, final_local_scores=-1.0924971, pem=-0.6661314, similarity=-24.890053]

RECALL MISS: 8030250

1991_Grand_Prix_motorcycle_racing_season]

8: 8857587 1991_Czechoslovakian_motorcycle_Grand_Prix
[winner: 349568 Czechoslovakian_Grand_Prix, pem_pos=0,
final_scores=-0.079210564, final_local_scores=-0.045220286,
pem=0.0, similarity=2.7402494|
gt: 8857587 1991_Czechoslovakian_motorcycle_Grand_Prix,
pem_pos=1, final_scores=-0.30088758,
final_local_scores=-0.22169626, pem=-0.68885535,
similarity=1.2199572]
9: 8046362 1992_Grand_Prix_motorcycle_racing_season
[winner: 619184 Robert_P._Casey, pem_pos=1,
final_scores=-0.7827521, final_local_scores=-0.7041556,
pem=-0.49112052, similarity=-14.252906|
RECALL MISS: 8046362
1992_Grand_Prix_motorcycle_racing_season]
10: 8859851 1992_Malaysian_motorcycle_Grand_Prix
[winner: 3607937 Malaysia, pem_pos=0,
final_scores=0.5086731, final_local_scores=0.6413615,
pem=-0.20223404, similarity=24.328999|
RECALL MISS: 8859851
1992_Malaysian_motorcycle_Grand_Prix]
11: 8873404 1992_Hungarian_motorcycle_Grand_Prix
[winner: 13275 Hungary, pem_pos=0, final_scores=0.448034,
final_local_scores=0.5719851, pem=-0.32577527,
similarity=22.930508|
RECALL MISS: 8873404
1992_Hungarian_motorcycle_Grand_Prix]
12: 8048901 1993_Grand_Prix_motorcycle_racing_season
[winner: 1964091 Next_(2007_film), pem_pos=0,
final_scores=-1.0140584, final_local_scores=-0.95666736,
pem=-1.4509634, similarity=-16.568584|
RECALL MISS: 8048901
1993_Grand_Prix_motorcycle_racing_season]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.5313376, final_local_scores=-1.5813239,
pem=-1.219945, similarity=-36.465897|
RECALL MISS: 8048901
1993_Grand_Prix_motorcycle_racing_season]
13: 8064900 1994_Grand_Prix_motorcycle_racing_season
[winner: 18948337 1994, pem_pos=0,
final_scores=-0.89725626, final_local_scores=-0.83434737,
pem=-0.7539553, similarity=-16.708752|
RECALL MISS: 8064900
1994_Grand_Prix_motorcycle_racing_season]
14: 8066644 1995_Grand_Prix_motorcycle_racing_season
[winner: 5834357 Juno_Awards_of_1995, pem_pos=5,
final_scores=-1.1070254, final_local_scores=-1.015677,
pem=-1.8938538, similarity=-15.9210415|
RECALL MISS: 8066644
1995_Grand_Prix_motorcycle_racing_season]
15: 8083518 1997_Grand_Prix_motorcycle_racing_season
[winner: 2602335 Toronto_municipal_election,_1997,
pem_pos=0, final_scores=-1.1546848,

```

final_local_scores=-1.0982232, pem=0.0,
similarity=-28.683456]
    RECALL MISS: 8083518
1997_Grand_Prix_motorcycle_racing_season]
16: 17937153 1997_Malaysian_motorcycle_Grand_Prix
    [winner: 309599 Malaysian_Grand_Prix, pem_pos=0,
final_scores=-0.12282415, final_local_scores=-0.09715281,
pem=-0.5060316, similarity=3.942358]
    gt: 17937153 1997_Malaysian_motorcycle_Grand_Prix,
pem_pos=12, final_scores=-0.54023343,
final_local_scores=-0.5064968, pem=-2.4770737,
similarity=2.44555 ]
1: 75753 Aprilia
    [pem_pos=0, final_scores=0.31029344,
final_local_scores=0.2650867, pem=0.0, similarity=12.000429]
2: 5585957 Fausto_Gresini
    [pem_pos=0, final_scores=0.524357,
final_local_scores=0.5721107, pem=0.0, similarity=21.162638]
3: 549741 Gilera
    [pem_pos=0, final_scores=0.30238658,
final_local_scores=0.29848346, pem=0.0, similarity=12.997055]
4: 13729 Honda
    [pem_pos=0, final_scores=0.08533637,
final_local_scores=0.04126209, pem=-0.17204832,
similarity=6.256682]
5: 13247557 Superbike_World_Championship
    [pem_pos=0, final_scores=0.14457722,
final_local_scores=0.09184113, pem=-0.023684667,
similarity=6.959236]
6: 332378 Yamaha_Motor_Company
    [pem_pos=1, final_scores=0.33984226,
final_local_scores=0.314652, pem=-0.7066168,
similarity=17.322248]
1: 14085866 Alessandro_Gramigni
    [pem_pos=0, final_scores=0.35121715,
final_local_scores=0.39257294, pem=0.0, similarity=15.8048725]
2: 14532 Italy
    [pem_pos=0, final_scores=0.7238769,
final_local_scores=0.82302815, pem=-0.75948656,
similarity=32.78072]
3: 3607937 Malaysia
    [pem_pos=0, final_scores=0.5086731,
final_local_scores=0.6413615, pem=-0.20223404,
similarity=24.328999]
4: 13275 Hungary
    [pem_pos=0, final_scores=0.448034,
final_local_scores=0.5719851, pem=-0.32577527,
similarity=22.930508]
5: 14085866 Alessandro_Gramigni
    [pem_pos=0, final_scores=0.30631167,
final_local_scores=0.3435205, pem=0.0, similarity=14.341049]
6: 75753 Aprilia

```

[pem_pos=0, final_scores=0.45628744,
final_local_scores=0.4252699, pem=0.0, similarity=16.780615]
7: 75753 Aprilia

[pem_pos=0, final_scores=0.28851503,
final_local_scores=0.2412971, pem=0.0, similarity=11.2905]

Sample: b'Black_Lake_Bayou'

[1 Black Lake Bayou 1] is a waterway in 1 northwest [1 1 Louisiana 1] , , that extends from north of [2 2 Gibsland 2] and travels south to 2 Clarence . The watershed covers much of northwest [2 Louisiana 2] . The 3 bayou meanders its way through [3 3 Claiborne Parish 3] , 4 Webster , 5 Bienville , [3 6 Red River 3] and [4 7 Natchitoches 4] parishes . [5 Black Lake Bayou 5] empties into 8 Black Lake , which drains to [4 4 Saline Bayou 4] , a tributary of the . [6 Black Lake Bayou 6] runs near the towns of [5 5 Minden 5] , [6 6 Dubberly 6] , Fryburg , [7 7 Gibsland 7] , [7 9 Mount Lebanon 7] , [8 10 Jamestown 8] , 11 Castor and Ashland . [8 8 Kepler Lake 8] is also a part of the [9 Black Lake Bayou 9] watershed . [10 Black Lake Bayou 10] runs underneath Louisiana Highways 154 , 156 , 531 , 792 , 12 4 , 507 , and 155 . The bayou is choked by trees -- bridge overpasses are the only places where the bayou is easily accessed and usually where people fish . There is a crude boat launch at Nix Crossing , which is 2-3 miles north of 13 Castor and 3-4 miles south of 14 Jamestown . Nix Crossing is one of the widest and deepest parts of [11 Black Lake Bayou 11] . The

1: 18130 Louisiana

2: 115497 Gibsland, Louisiana

3: 97706 Claiborne_Parish, Louisiana

4: 24159027 Saline_Bayou

5: 115833 Minden, Louisiana

6: 115831 Dubberly, Louisiana

7: 115497 Gibsland, Louisiana

8: 18292017 Kepler_Lake

1: 922286 North_Louisiana

[winner: 6114879 Northwest_Historic_District, pem_pos=10,
final_scores=-1.2431343, final_local_scores=-1.1871347,
pem=-3.0471427, similarity=-14.765921]

RECALL MISS: 922286 North_Louisiana]

2: 115672 Clarence, Louisiana

[winner: 126403 Clarence, New_York, pem_pos=0,
final_scores=-0.17810407, final_local_scores=-0.1864352,
pem=-1.9442668, similarity=9.0993395]

gt: 115672 Clarence, Louisiana, pem_pos=12,
final_scores=-0.500471, final_local_scores=-0.51682395,
pem=-3.5571144, similarity=8.010796]

3: 633843 Bayou

[winner: 10360528 The_Bayou, pem_pos=0,
final_scores=-1.1645778, final_local_scores=-1.1388514,
pem=0.0, similarity=-29.895884]

RECALL MISS: 633843 Bayou]

4: 97657 Webster_Parish,_Louisiana
[winner: 97657 Webster_Parish,_Louisiana, pem_pos=7,
final_scores=-0.2184831, final_local_scores=-0.2743538,
pem=-3.345952, similarity=14.098255|
gt: 97657 Webster_Parish,_Louisiana, pem_pos=7,
final_scores=-0.2184831, final_local_scores=-0.2743538,
pem=-3.345952, similarity=14.098255]

5: 97713 Bienville_Parish,_Louisiana
[winner: 97713 Bienville_Parish,_Louisiana, pem_pos=1,
final_scores=0.0037930459, final_local_scores=-0.033824474,
pem=-0.69126844, similarity=6.839548|
gt: 97713 Bienville_Parish,_Louisiana, pem_pos=1,
final_scores=0.0037930459, final_local_scores=-0.033824474,
pem=-0.69126844, similarity=6.839548]

6: 97677 Red_River_Parish,_Louisiana
[winner: 150045 Red_River_of_the_North, pem_pos=1,
final_scores=0.25029767, final_local_scores=0.25717574,
pem=-1.4132268, similarity=19.449697|
gt: 97677 Red_River_Parish,_Louisiana, pem_pos=3,
final_scores=-0.16164583, final_local_scores=-0.19294223,
pem=-2.4828434, similarity=11.83402]

7: 97685 Natchitoches_Parish,_Louisiana
[winner: 115675 Natchitoches,_Louisiana, pem_pos=0,
final_scores=0.17994957, final_local_scores=0.1578443,
pem=-0.028531445, similarity=8.95526|
gt: 97685 Natchitoches_Parish,_Louisiana, pem_pos=1,
final_scores=-0.021053627, final_local_scores=-0.05862686,
pem=-0.7758661, similarity=6.5594506]

8: 337362 Black_Lake_(Louisiana)
[winner: 1851915 Black_Lake_(New_York), pem_pos=1,
final_scores=-0.014845207, final_local_scores=0.029119119,
pem=-1.1582506, similarity=11.257432|
gt: 337362 Black_Lake_(Louisiana), pem_pos=5,
final_scores=-0.060377494, final_local_scores=-0.051089317,
pem=-1.857, similarity=12.66376]

9: 115500 Mount_Lebanon,_Louisiana
[winner: 381121 Mount_Lebanon, pem_pos=0,
final_scores=0.14411949, final_local_scores=0.24242347,
pem=0.0, similarity=11.324113|
gt: 115500 Mount_Lebanon,_Louisiana, pem_pos=2,
final_scores=0.021399513, final_local_scores=0.04477574,
pem=-1.2579358, similarity=12.266761]

10: 115498 Jamestown,_Louisiana
[winner: 75804 Jamestown,_Virginia, pem_pos=0,
final_scores=0.08653198, final_local_scores=0.11212744,
pem=-1.0401132, similarity=13.092116|
RECALL MISS: 115498 Jamestown,_Louisiana]

11: 115496 Castor,_Louisiana
[winner: 115496 Castor,_Louisiana, pem_pos=4,
final_scores=-0.37174684, final_local_scores=-0.42107245,
pem=-2.7227316, similarity=6.330709|
gt: 115496 Castor,_Louisiana, pem_pos=4,

final_scores=-0.37174684, final_local_scores=-0.42107245,
pem=-2.7227316, similarity=6.330709]
12: 6043799 Louisiana_Highway_4
[winner: 19351 May_4, pem_pos=3, final_scores=-0.8583474,
final_local_scores=-0.6583802, pem=-3.5437841,
similarity=3.713983]

RECALL MISS: 6043799 Louisiana_Highway_4]
13: 115496 Castor, Louisiana
[winner: 115496 Castor, Louisiana, pem_pos=4,
final_scores=-0.38806486, final_local_scores=-0.43889746,
pem=-2.7227316, similarity=5.798775]

gt: 115496 Castor, Louisiana, pem_pos=4,
final_scores=-0.38806486, final_local_scores=-0.43889746,
pem=-2.7227316, similarity=5.798775]

14: 115498 Jamestown, Louisiana
[winner: 75804 Jamestown, Virginia, pem_pos=0,
final_scores=-0.14769484, final_local_scores=-0.156075,
pem=-1.0401132, similarity=5.088421]

RECALL MISS: 115498 Jamestown, Louisiana]

1: 18130 Louisiana
[pem_pos=0, final_scores=0.42229015,
final_local_scores=0.41041392, pem=-0.17745006,
similarity=17.302284]

2: 115497 Gibsland, Louisiana
[pem_pos=0, final_scores=0.17134099,
final_local_scores=0.17874181, pem=0.0, similarity=9.423725]

3: 97706 Claiborne Parish, Louisiana
[pem_pos=0, final_scores=0.29329818,
final_local_scores=0.29464996, pem=0.0, similarity=12.882655]

4: 24159027 Saline Bayou
[pem_pos=0, final_scores=0.091283664,
final_local_scores=0.15967539, pem=0.0, similarity=8.854745]

5: 115833 Minden, Louisiana
[pem_pos=1, final_scores=0.259355,
final_local_scores=0.26039797, pem=-1.3235818,
similarity=19.058352]

6: 115831 Dubberly, Louisiana
[pem_pos=0, final_scores=0.08209737,
final_local_scores=0.13509905, pem=0.0, similarity=8.121338]

7: 115497 Gibsland, Louisiana
[pem_pos=0, final_scores=0.07675992,
final_local_scores=0.075426355, pem=0.0, similarity=6.3405857]

8: 18292017 Kepler Lake
[pem_pos=0, final_scores=0.18734191,
final_local_scores=0.25016758, pem=0.0, similarity=11.555212]

1: 18291286 Black Lake Bayou
[pem_pos=0, final_scores=0.15138002,
final_local_scores=0.15744376, pem=0.0, similarity=8.788149]

2: 18130 Louisiana
[pem_pos=0, final_scores=0.45888942,
final_local_scores=0.45039308, pem=-0.17745006,
similarity=18.495342]

3: 150045 Red_River_of_the_North
 [pem_pos=1, final_scores=0.25029767,
 final_local_scores=0.25717574, pem=-1.4132268,
 similarity=19.449697]
 4: 115675 Natchitoches,_Louisiana
 [pem_pos=0, final_scores=0.17994957,
 final_local_scores=0.1578443, pem=-0.028531445,
 similarity=8.95526]
 5: 18291286 Black_Lake_Bayou
 [pem_pos=0, final_scores=0.11214183,
 final_local_scores=0.114582, pem=0.0, similarity=7.5090685]
 6: 18291286 Black_Lake_Bayou
 [pem_pos=0, final_scores=0.11005069,
 final_local_scores=0.11229776, pem=0.0, similarity=7.440902]
 7: 381121 Mount_Lebanon
 [pem_pos=0, final_scores=0.14411949,
 final_local_scores=0.24242347, pem=0.0, similarity=11.324113]
 8: 75804 Jamestown,_Virginia
 [pem_pos=0, final_scores=0.08653198,
 final_local_scores=0.11212744, pem=-1.0401132,
 similarity=13.092116]
 9: 18291286 Black_Lake_Bayou
 [pem_pos=0, final_scores=0.26371795,
 final_local_scores=0.2801559, pem=0.0, similarity=12.450124]
 10: 18291286 Black_Lake_Bayou
 [pem_pos=0, final_scores=0.18738912,
 final_local_scores=0.19677824, pem=0.0, similarity=9.961968]
 11: 18291286 Black_Lake_Bayou
 [pem_pos=0, final_scores=0.118827745,
 final_local_scores=0.12188536, pem=0.0, similarity=7.727015]

Sample: b'Dedi_I'

, _ Margrave_of_the_Saxon_Ostmark Dedi -LRB- or Dedo -RRB- -
 LRB- 1004 -- October 1075 -RRB- was the -LRB- also called [1 1
 Lower Lusatia 1] -RRB- from 1046 and a claimant for the title
 of from 1069 . He was the second son of and Matilda ,
 daughter of . Dedi inherited the [2 Ostmark 2] from its last
 dynast , the childless 2 Odo II , because he had married his
 sister , Oda -LRB- died before 1068 -RRB- . Oda was herself
 the widow of and mother of 3 William and 4 Otto , margraves
 of [3 Meissen 3] successively . When Otto his stepson died ,
 Dedi married his widow , Adela of [4 Louvain 4] , and in her
 name claimed the . Otto died in 1067 and was succeeded by [1 1
 Egbert I 1] , but Dedi married his widow in 1069 and
 rebelled . In claiming the [5 Meissen March 5] through his
 wife , Dedi was challenging the royal prerogative in the
 marches . With him in his revolt was 5 Albert II of [6 Ballen
 stedt 6] , who raided the monastery of [7 6 Nienburg 7] , a
 foundation of the family of Dedi 's first wife . Adela of [8
 Louvain 8] for her part aggressively supported her husband ,
 so much so that 7 Lambert of [9 Hersfeld 9] was compelled to
 call her a -LRB- rough meaning `` raging wife '' -RRB- . ,

one of the regents of the young king [10 8 Henry IV 10] , frustrated the [11 Thuringian 11] rebels and preserved the peace in [12 Thuringia 12] and [13 Meissen 13] . Dedi was confined to [14 Lusatia 14] , where he was succeeded by his eldest son , 9 Dedi II . Dedi left a daughter , [15 10 Adelai de 15] , by his first wife Oda , who married 11 Ernest , Margrave of [16 Austria 16] . By Adela of [17 Louvain 17] , Dedi had two sons : , who later ruled both [18 Lusatia 18] and [19 Meissen 19] , and Conrad , who died in battle with the [2 2 Wends 2]

1: 11642200 Egbert_I,_Margrave_of_Meissen

2: 33644 Wends

1: 11603057 March_of_Lusatia

[winner: 159283 Lusatia, pem_pos=0, final_scores=0.1606075, final_local_scores=0.120304555, pem=0.0, similarity=7.6798406]

gt: 11603057 March_of_Lusatia, pem_pos=2, final_scores=-0.14408168, final_local_scores=-0.1735892, pem=-1.0892664, similarity=4.833065]

2: 11779249 Odo_II,_Margrave_of_the_Saxon_Ostmark

[winner: 3137222 Odo_II,_Count_of_Blois, pem_pos=0, final_scores=0.01338236, final_local_scores=0.06818545, pem=-0.5389965, similarity=9.05565]

gt: 11779249 Odo_II,_Margrave_of_the_Saxon_Ostmark, pem_pos=1, final_scores=-0.37177446, final_local_scores=-0.25880158, pem=-1.0696247, similarity=2.1833432]

3: 4925633 William,_Margrave_of_Meissen

[winner: 72201 Prince_William,_Duke_of_Cambridge, pem_pos=0, final_scores=-0.053739935, final_local_scores=-0.019363761, pem=-1.8409588, similarity=13.52328]

RECALL MISS: 4925633 William,_Margrave_of_Meissen]

4: 11811105 Otto_I,_Margrave_of_Meissen

[winner: 193686 Otto,_King_of_Greece, pem_pos=1, final_scores=-0.06336805, final_local_scores=-0.0849348, pem=-2.0440273, similarity=12.670828]

RECALL MISS: 11811105 Otto_I,_Margrave_of_Meissen]

5: 10704782 Adalbert_II,_Count_of_Ballenstedt

[winner: 1042171 Ballenstedt, pem_pos=0, final_scores=0.23878835, final_local_scores=0.244885, pem=0.0, similarity=11.39757]

RECALL MISS: 10704782 Adalbert_II,_Count_of_Ballenstedt]

[winner: 1051570 Albert,_Somme, pem_pos=0, final_scores=-0.37240767, final_local_scores=-0.38127297, pem=-1.9531993, similarity=3.333569]

RECALL MISS: 10704782 Adalbert_II,_Count_of_Ballenstedt]

[winner: 26011508 Atsushi_Ii, pem_pos=5, final_scores=-0.8289163, final_local_scores=-0.5942098, pem=-2.1972246, similarity=-1.6938436]

RECALL MISS: 10704782 Adalbert_II,_Count_of_Ballenstedt]

6: 4185922 Nienburg,_Saxony-Anhalt

[winner: 1333670 Nienburg, Lower Saxony, pem_pos=0, final_scores=0.3810156, final_local_scores=0.43048042, pem=-0.26317766, similarity=18.367308]
gt: 4185922 Nienburg, Saxony-Anhalt, pem_pos=3, final_scores=-0.093630716, final_local_scores=-0.07488114, pem=-1.6634009, similarity=10.900943]
7: 1647641 Lambert_of_Hersfeld
[winner: 271043 Bad_Hersfeld, pem_pos=1, final_scores=0.13301341, final_local_scores=0.1231378, pem=-0.24702077, similarity=9.107728]
RECALL MISS: 1647641 Lambert_of_Hersfeld]
[winner: 1493071 Lambert_II_of_Spoleto, pem_pos=3, final_scores=-0.5221074, final_local_scores=-0.53064, pem=-2.777866, similarity=3.3608236]
gt: 1647641 Lambert_of_Hersfeld, pem_pos=17, final_scores=-0.74109155, final_local_scores=-0.7912795, pem=-4.174737, similarity=3.1792233]
8: 27485413 Henry_IV, Holy_Roman_Emperor
[winner: 57876 Henry_IV_of_France, pem_pos=1, final_scores=0.1491837, final_local_scores=0.13535649, pem=-1.0399671, similarity=13.784522]
gt: 27485413 Henry_IV, Holy_Roman_Emperor, pem_pos=2, final_scores=-0.009959236, final_local_scores=-0.050385714, pem=-1.5746015, similarity=11.149031]
9: 11827028 Dedi_II, Margrave_of_Lusatia
[winner: 11827028 Dedi_II, Margrave_of_Lusatia, pem_pos=0, final_scores=-0.11543098, final_local_scores=0.03739883, pem=0.0, similarity=5.2057686]
gt: 11827028 Dedi_II, Margrave_of_Lusatia, pem_pos=0, final_scores=-0.11543098, final_local_scores=0.03739883, pem=0.0, similarity=5.2057686]
10: 34954770 Adelaide_of_Eilenburg
[winner: 1148 Adelaide, pem_pos=0, final_scores=0.22282682, final_local_scores=0.2849296, pem=-0.29570588, similarity=14.200676]
RECALL MISS: 34954770 Adelaide_of_Eilenburg]
11: 1178564 Ernest, Margrave_of_Austria
[winner: 26964606 Austria, pem_pos=0, final_scores=0.57585305, final_local_scores=0.59872603, pem=-0.25262418, similarity=23.330702]
RECALL MISS: 1178564 Ernest, Margrave_of_Austria]
[winner: 196475 Margrave, pem_pos=1, final_scores=-0.23233877, final_local_scores=-0.3239068, pem=-0.02950915, similarity=-5.415837]
RECALL MISS: 1178564 Ernest, Margrave_of_Austria]
[winner: 1565385 Ernest, Elector_of_Saxony, pem_pos=3, final_scores=-0.27147716, final_local_scores=-0.31106192, pem=-2.604395, similarity=8.970109]
gt: 1178564 Ernest, Margrave_of_Austria, pem_pos=7, final_scores=-0.7017592, final_local_scores=-0.7173888, pem=-3.6370912, similarity=2.460466]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,

final_scores=-1.6145399, final_local_scores=-1.5681407,
pem=-4.6728287, similarity=-17.295164]
RECALL MISS: 1178564 Ernest,_Margrave_of_Austria]
1: 11642200 Egbert_I,_Margrave_of_Meissen
[pem_pos=0, final_scores=0.11723654,
final_local_scores=0.1913799, pem=0.0, similarity=9.800871]
2: 33644 Wends
[pem_pos=0, final_scores=0.10307191,
final_local_scores=0.1099329, pem=0.0, similarity=7.37033]
1: 159283 Lusatia
[pem_pos=0, final_scores=0.1606075,
final_local_scores=0.120304555, pem=0.0, similarity=7.6798406]
2: 312569 East_German_mark
[pem_pos=0, final_scores=0.051444754,
final_local_scores=0.14695382, pem=-0.21677566,
similarity=9.653968]
3: 40674 Meissen
[pem_pos=0, final_scores=0.4810745,
final_local_scores=0.44543356, pem=0.0, similarity=17.38234]
4: 76931 Leuven
[pem_pos=1, final_scores=0.4859602,
final_local_scores=0.4862573, pem=-0.087840684,
similarity=19.078293]
5: 3211873 Margraviate_of_Meissen
[pem_pos=0, final_scores=0.178946,
final_local_scores=0.15498999, pem=0.0, similarity=8.714923]
6: 1042171 Ballenstedt
[pem_pos=0, final_scores=0.23878835,
final_local_scores=0.244885, pem=0.0, similarity=11.39757]
7: 1333670 Nienburg,_Lower_Saxony
[pem_pos=0, final_scores=0.3810156,
final_local_scores=0.43048042, pem=-0.26317766,
similarity=18.367308]
8: 76931 Leuven
[pem_pos=1, final_scores=0.47675496,
final_local_scores=0.476202, pem=-0.087840684,
similarity=18.778223]
9: 271043 Bad_Hersfeld
[pem_pos=1, final_scores=0.13301341,
final_local_scores=0.1231378, pem=-0.24702077,
similarity=9.107728]
10: 57876 Henry_IV_of_France
[pem_pos=1, final_scores=0.1491837,
final_local_scores=0.13535649, pem=-1.0399671,
similarity=13.784522]
11: 31130 Thuringia
[pem_pos=0, final_scores=0.40794283,
final_local_scores=0.39073473, pem=-0.13528216,
similarity=16.485703]
12: 31130 Thuringia
[pem_pos=0, final_scores=0.31023735,
final_local_scores=0.2812555, pem=0.0, similarity=12.482939]

13: 40674 Meissen
 [pem_pos=0, final_scores=0.47169936,
 final_local_scores=0.43455577, pem=0.0, similarity=17.057724]
 14: 159283 Lusatia
 [pem_pos=0, final_scores=0.22224228,
 final_local_scores=0.18763122, pem=0.0, similarity=9.689003]
 15: 1148 Adelaide
 [pem_pos=0, final_scores=0.22282682,
 final_local_scores=0.2849296, pem=-0.29570588,
 similarity=14.200676]
 16: 26964606 Austria
 [pem_pos=0, final_scores=0.57585305,
 final_local_scores=0.59872603, pem=-0.25262418,
 similarity=23.330702]
 17: 76931 Leuven
 [pem_pos=1, final_scores=0.49349374,
 final_local_scores=0.49448657, pem=-0.087840684,
 similarity=19.32387]
 18: 159283 Lusatia
 [pem_pos=0, final_scores=0.21274905,
 final_local_scores=0.1772613, pem=0.0, similarity=9.379543]
 19: 40674 Meissen
 [pem_pos=0, final_scores=0.41212994,
 final_local_scores=0.36948526, pem=0.0, similarity=15.11589]

Sample: b'Dennis_Pilgrim'

Dennis Pilgrim -LRB- born 15 September 1972 -RRB- is a former [1 1 Bermudian 1] 1 cricket er . Pilgrim 's 2 batting and 3 bowling styles are unknown . Pilgrim made his debut for in the 4 2000 Americas Cricket Cup against 5 Argentina , with him making two further appearances in that competition against the and [2 2 Canada 2] . He made his [3 3 List A 3] debut for [1 Bermuda 1] in October 2000 against the [4 4 Leeward Islands 4] in the 2000/01 , with him making three further [2 List A 2] appearances in that tournament , against the [3 Cayman Islands 3] , the [5 5 Windward Islands 5] and [6 6 Guyana 6] . He scored a total of 20 runs in his four [4 List A 4] matches with a high score of 11 6 not out , while with the ball he also took a single wicket . In July 2001 , Pilgrim played in the 7 2001 [5 ICC Trophy 5] in [6 Canada 6] , making two appearances against and 8 Namibia . The following year he made four appearances in the 9 2002 [7 ICC Americas Championship 7] against the [8 Cayman Islands 8] , [9 Canada 9] , the [10 10 Bahamas 10] and the [11 11 United States 11]

1: 3460 Bermuda
 2: 1880843 Canada_national_cricket_team
 3: 683459 List_A_cricket
 4: 3226222 Leeward_Islands_cricket_team
 5: 3232006 Windward_Islands_cricket_team
 6: 3201481 Guyana_national_cricket_team

1: 25675557 Cricket
 [winner: 330091 Raphidophoridae, pem_pos=10,

```

final_scores=-1.5407736, final_local_scores=-1.5338175,
pem=-2.56456, similarity=-27.735985|
    gt: 25675557 Cricket, pem_pos=0,
final_scores=-2.2658126, final_local_scores=-2.5102944,
pem=-0.07532314, similarity=-70.41286 ]
2: 673526 Batting_(cricket)
    [winner: 15652278 Batting_(material), pem_pos=2,
final_scores=-0.9180748, final_local_scores=-0.81702155,
pem=-1.1616125, similarity=-13.974811|
    gt: 673526 Batting_(cricket), pem_pos=3,
final_scores=-1.4732323, final_local_scores=-1.6052263,
pem=-1.1802309, similarity=-37.395164 ]
3: 51878 Bowling_(cricket)
    [winner: 1042379 Bowled, pem_pos=3,
final_scores=-1.2675207, final_local_scores=-1.348897,
pem=-1.0816903, similarity=-30.281668|
    gt: 51878 Bowling_(cricket), pem_pos=4,
final_scores=-1.5849065, final_local_scores=-1.7255893,
pem=-1.7833481, similarity=-37.707188 ]
4: 28434668 2000_Americas_Cricket_Cup
    [winner: 28434668 2000_Americas_Cricket_Cup, pem_pos=0,
final_scores=-0.025711685, final_local_scores=0.0017693192,
pem=0.0, similarity=4.142513|
    gt: 28434668 2000_Americas_Cricket_Cup, pem_pos=0,
final_scores=-0.025711685, final_local_scores=0.0017693192,
pem=0.0, similarity=4.142513 ]
    [winner: 34548 2000, pem_pos=0, final_scores=-1.146415,
final_local_scores=-1.1103914, pem=-0.8862115,
similarity=-24.227226|
    RECALL MISS: 28434668 2000_Americas_Cricket_Cup]
5: 4367228 Argentina_national_cricket_team
    [winner: 454699 Argentina_national_football_team,
pem_pos=1, final_scores=-0.18801562,
final_local_scores=-0.19107607, pem=-2.3085454,
similarity=10.941851|
    gt: 4367228 Argentina_national_cricket_team, pem_pos=5,
final_scores=-0.31948054, final_local_scores=-0.39395714,
pem=-4.268207, similarity=15.54442 ]
6: 2430357 Not_out
    [winner: 2430357 Not_out, pem_pos=0,
final_scores=-0.97930753, final_local_scores=-1.097585,
pem=-0.008413082, similarity=-28.618662|
    gt: 2430357 Not_out, pem_pos=0,
final_scores=-0.97930753, final_local_scores=-1.097585,
pem=-0.008413082, similarity=-28.618662 ]
7: 7121453 2001_ICC_Trophy
    [winner: 1952482 ICC_World_Cup_Qualifier, pem_pos=0,
final_scores=0.3312946, final_local_scores=0.31355, pem=0.0,
similarity=13.446671|
    gt: 7121453 2001_ICC_Trophy, pem_pos=2,
final_scores=-0.119127065, final_local_scores=-0.15249793,
pem=-1.3528783, similarity=6.896033 ]

```

[winner: 34551 2001, pem_pos=0, final_scores=-0.5901143,
final_local_scores=-0.3809355, pem=-0.5613713,
similarity=-4.2253375]

RECALL MISS: 7121453 2001_ICC_Trophy]

8: 30864436 Namibia_national_cricket_team

[winner: 30864436 Namibia_national_cricket_team,
pem_pos=1, final_scores=-0.053966224,
final_local_scores=-0.08454636, pem=-2.6410618,
similarity=15.929186]

gt: 30864436 Namibia_national_cricket_team, pem_pos=1,
final_scores=-0.053966224, final_local_scores=-0.08454636,
pem=-2.6410618, similarity=15.929186]

9: 28484686 2002_ICC_Americas_Championship

[winner: 7115770 ICC_Americas_Championship, pem_pos=0,
final_scores=0.082735285, final_local_scores=0.058912516,
pem=0.0, similarity=5.8477798]

RECALL MISS: 28484686 2002_ICC_Americas_Championship]

[winner: 35502 2002, pem_pos=0, final_scores=-0.9978681,
final_local_scores=-0.9093702, pem=-0.5897454,
similarity=-19.84058]

RECALL MISS: 28484686 2002_ICC_Americas_Championship]

10: 6634183 Bahamas_national_cricket_team

[winner: 3451 The_Bahamas, pem_pos=0,
final_scores=0.21929263, final_local_scores=0.22252887,
pem=-0.35113433, similarity=12.639942]

gt: 6634183 Bahamas_national_cricket_team, pem_pos=2,
final_scores=-0.2616291, final_local_scores=-0.29054287,
pem=-2.5607586, similarity=9.345138]

11: 980450 United_States_national_cricket_team

[winner: 3434750 United_States, pem_pos=0,
final_scores=0.2654571, final_local_scores=0.29969817,
pem=-0.073457584, similarity=13.432778]

gt: 980450 United_States_national_cricket_team,
pem_pos=4, final_scores=-0.5343667,
final_local_scores=-0.6576171, pem=-5.176109,
similarity=12.613592]

1: 3460 Bermuda

[pem_pos=0, final_scores=0.08557226,
final_local_scores=0.067536175, pem=-0.26592073,
similarity=7.551245]

2: 1880843 Canada_national_cricket_team

[pem_pos=1, final_scores=0.17940165,
final_local_scores=0.1337238, pem=-4.476951,
similarity=32.426655]

3: 683459 List_A_cricket

[pem_pos=0, final_scores=0.09928523,
final_local_scores=0.04736063, pem=-0.020267159,
similarity=5.6132646]

4: 3226222 Leeward_Islands_cricket_team

[pem_pos=1, final_scores=0.18563284,
final_local_scores=0.15260005, pem=-1.3078154,
similarity=15.755707]

5: 3232006 Windward_Islands_cricket_team
 [pem_pos=1, final_scores=0.25188792,
 final_local_scores=0.21799889, pem=-1.1147723,
 similarity=16.65754]

6: 3201481 Guyana_national_cricket_team
 [pem_pos=2, final_scores=0.05451028,
 final_local_scores=0.010494888, pem=-2.8866508,
 similarity=20.100958]

1: 1829476 Bermuda_national_cricket_team
 [pem_pos=1, final_scores=0.31196088,
 final_local_scores=0.27608123, pem=-2.445982,
 similarity=25.630157]

2: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.30464643,
 final_local_scores=0.27168655, pem=-0.020267159,
 similarity=12.307597]

3: 2550024 Cayman_Islands_national_cricket_team
 [pem_pos=2, final_scores=0.040307954,
 final_local_scores=0.0025603771, pem=-2.2676213,
 similarity=16.497797]

4: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.29106253,
 final_local_scores=0.25684816, pem=-0.020267159,
 similarity=11.86479]

5: 1952482 ICC_World_Cup_Qualifier
 [pem_pos=0, final_scores=0.3312946,
 final_local_scores=0.31355, pem=0.0, similarity=13.446671]

6: 5042916 Canada
 [pem_pos=0, final_scores=0.7349164,
 final_local_scores=0.790517, pem=-0.11662793,
 similarity=28.314556]

7: 7115770 ICC_Americas_Championship
 [pem_pos=0, final_scores=0.082735285,
 final_local_scores=0.058912516, pem=0.0, similarity=5.8477798]

8: 5468 Cayman_Islands
 [pem_pos=0, final_scores=0.29311275,
 final_local_scores=0.29092804, pem=-0.47532094,
 similarity=15.356456]

9: 5042916 Canada
 [pem_pos=0, final_scores=0.31314558,
 final_local_scores=0.32979646, pem=-0.11662793,
 similarity=14.565738]

10: 3451 The_Bahamas
 [pem_pos=0, final_scores=0.21929263,
 final_local_scores=0.22252887, pem=-0.35113433,
 similarity=12.639942]

11: 3434750 United_States
 [pem_pos=0, final_scores=0.2654571,
 final_local_scores=0.29969817, pem=-0.073457584,
 similarity=13.432778]

Sample: b'Florida_Gulf_Coast_Eagles_men'

's _ basketball The [1 Florida Gulf Coast Eagles 1] men 's basketball team is the men 's basketball team that represents [1 1 Florida Gulf Coast University 1] in [2 1 Fort Myers 2] , [3 Florida 3] , [4 United States 4] . The school 's team currently competes in the [2 2 Atlantic Sun Conference 2] . In 2002 , Florida Gulf Coast became an independent member of . [5 Florida Gulf Coast 5] also started its men 's basketball team , with [3 3 Dave Balza 3] as head coach from the inaugural 2002 -- 03 to the 2010 -- 11 season . In 2006 , [6 Florida Gulf Coast 6] applied for status and became a transitory [7 Division I 7] effective in the 2007 -- 08 season . [8 Florida Gulf Coast 8] became a full Division I member on August 11 , 2011 . The Eagles men 's basketball team is best known for their run in the 2 2013 [9 NCAA Men 9] 's Division I Basketball Tournament , their first since moving up to [10 Division I 10] . In the Round of 64 , the Eagles defeated the # 2 seeded [11 3 Georgetown Hoyas 11] 78-68 . It was only the seventh time that a 15th seed has defeated a 2nd seed , and the second-highest margin of victory for one . Two days later in the Round of 32 , the Eagles defeated the # 7 seeded [12 4 San Diego State Aztecs 12] 81-71 , becoming the first time a 15th seed has ever advanced to the . The school was referred to as `` Dunk City '' or `` Florida Dunk Coast '' on many media outlets such as [13 ESPN 13] and [14 CBS 14] due to their high-flying athletic finishing style . The Eagles participated in their first [15 Division I 15] postseason tournament in 2013 , qualifying for the [16 NCAA Tournament 16] by winning the [17 Atlantic Sun Conference Championship 17] , beating top-seeded 5 Mercer in the conference tournament final . On March 22 , 2013 ; the Eagles , the 15th seed in the South Region , scored one of the biggest upsets in [18 NCAA Tournament 18] history , defeating second-seeded 6 Georgetown 78-68 . They were only the seventh # 15 seed to defeat a # 2 seed , the first being [19 Richmond 19] 's 1991 victory over Syracuse . Two days later , they defeated seventh seeded [20 7 San Diego State 20] to become the first # 15 seed ever to advance to the Sweet Sixteen . This is `` officially '' the third-deepest run by a first-time tournament participant in [21 NCAA 21] history , behind only [22 8 Indiana State 22] advancing all the way to the national championship game in 9 1979 and 10 Georgia reaching the [23 Final Four 23] in 11 1983 . advanced to the Sweet 16 in its first appearance as well , in 12 1972 , but that appearance was vacated due to [24 NCAA 24] sanctions . Coach [4 4 Andy Enfield 4] was hired by [25 13 USC 25] on April 1 . To replace him , the Eagles hired [5 5 Joe Dooley 5] , an assistant from [26 14 Kansas 26]

1: 1399555 Florida_Gulf_Coast_University

2: 672965 Atlantic_Sun_Conference

3: 22239452 Dave_Balza

4: 38893193 Andy_Enfield

5: 6616761 Joe_Dooley

1: 109328 Fort_Myers,_Florida
 [winner: 109328 Fort_Myers,_Florida, pem_pos=0,
 final_scores=0.47000152, final_local_scores=0.50188184,
 pem=0.0, similarity=19.066868|
 gt: 109328 Fort_Myers,_Florida, pem_pos=0,
 final_scores=0.47000152, final_local_scores=0.50188184,
 pem=0.0, similarity=19.066868]
 [winner: 18933066 Florida, pem_pos=0,
 final_scores=0.19788359, final_local_scores=0.21155775,
 pem=-0.15405534, similarity=11.240795|
 RECALL MISS: 109328 Fort_Myers,_Florida]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7956693, final_local_scores=-1.7637849,
 pem=-4.6728287, similarity=-23.133574|
 RECALL MISS: 109328 Fort_Myers,_Florida]

2: 24466611 2013_NCAA_Men's_Division_I_Basketball_Tournament
 [winner: 60706 National_Collegiate_Athletic_Association,
 pem_pos=0, final_scores=0.039570495,
 final_local_scores=-0.03638518, pem=0.0,
 similarity=3.0039067|
 RECALL MISS: 24466611
 2013_NCAA_Men's_Division_I_Basketball_Tournament]
 [winner: 46945 2013, pem_pos=0, final_scores=-0.8379611,
 final_local_scores=-0.6533955, pem=-1.490928,
 similarity=-7.301014|
 RECALL MISS: 24466611
 2013_NCAA_Men's_Division_I_Basketball_Tournament]
 [winner: 146996 Tournament, pem_pos=0,
 final_scores=-0.9995684, final_local_scores=-1.061228,
 pem=-0.68980515, similarity=-23.828178|
 RECALL MISS: 24466611
 2013_NCAA_Men's_Division_I_Basketball_Tournament]
 [winner: 522936 Fallacy_of_division, pem_pos=6,
 final_scores=-1.2323222, final_local_scores=-1.1119158,
 pem=-2.9699564, similarity=-12.940988|
 RECALL MISS: 24466611
 2013_NCAA_Men's_Division_I_Basketball_Tournament]
 [winner: 12713983 Basketball_(song), pem_pos=6,
 final_scores=-1.9765967, final_local_scores=-2.0011225,
 pem=-3.6918519, similarity=-35.550907|
 RECALL MISS: 24466611
 2013_NCAA_Men's_Division_I_Basketball_Tournament]
 [winner: 1415027 David_Berman_(musician), pem_pos=5,
 final_scores=-2.5645945, final_local_scores=-2.70734,
 pem=-3.0012724, similarity=-60.38133|
 RECALL MISS: 24466611
 2013_NCAA_Men's_Division_I_Basketball_Tournament]

3: 37267561 2012-13_Georgetown_Hoyas_men's_basketball_team
 [winner: 3493898 Georgetown_Hoyas, pem_pos=0,
 final_scores=0.10861643, final_local_scores=0.07989849,
 pem=0.0, similarity=6.4740434|
 gt: 37267561 2012-

13_Georgetown_Hoyas_men's_basketball_team, pem_pos=6,
final_scores=-1.0089018, final_local_scores=-1.04124,
pem=-5.8930244, similarity=5.064218]
4: 36237885 2012-
13_San_Diego_State_Aztecs_men's_basketball_team
[winner: 6767670 San_Diego_State_Aztecs, pem_pos=0,
final_scores=0.15094243, final_local_scores=0.13483128,
pem=0.0, similarity=8.113347]
gt: 36237885 2012-
13_San_Diego_State_Aztecs_men's_basketball_team, pem_pos=9,
final_scores=-1.0086645, final_local_scores=-1.0522552,
pem=-5.866468, similarity=4.591085]
5: 22012745 Mercer_Bears
[winner: 986092 Mercer_University, pem_pos=6,
final_scores=-0.36512077, final_local_scores=-0.43494698,
pem=-2.7965183, similarity=6.317929]
RECALL MISS: 22012745 Mercer_Bears]
6: 3493898 Georgetown_Hoyas
[winner: 101681 Georgetown_University, pem_pos=2,
final_scores=-0.3498559, final_local_scores=-0.41785032,
pem=-2.5124655, similarity=5.2834044]
gt: 3493898 Georgetown_Hoyas, pem_pos=8,
final_scores=-0.46994823, final_local_scores=-0.5577993,
pem=-3.0723305, similarity=4.1516795]
7: 23757153 San_Diego_State_Aztecs_men's_basketball
[winner: 286083 San_Diego_State_University, pem_pos=0,
final_scores=0.054670706, final_local_scores=0.010554671,
pem=-0.47265834, similarity=6.975075]
gt: 23757153 San_Diego_State_Aztecs_men's_basketball,
pem_pos=8, final_scores=-0.3623631,
final_local_scores=-0.42125475, pem=-2.921506,
similarity=7.406234]
8: 22445207 1978-
79_Indiana_State_Sycamores_men's_basketball_team
[winner: 380922 Indiana_State_University, pem_pos=0,
final_scores=0.09129174, final_local_scores=0.0646002,
pem=-0.38142183, similarity=8.0917425]
gt: 22445207 1978-
79_Indiana_State_Sycamores_men's_basketball_team, pem_pos=3,
final_scores=-0.19596854, final_local_scores=-0.17673145,
pem=-1.5396848, similarity=7.188739]
9: 4380947 1979_NCAA_Men's_Division_I_Basketball_Tournament
[winner: 34754 1979, pem_pos=0, final_scores=-0.5054708,
final_local_scores=-0.25280076, pem=-0.65922886,
similarity=0.13062291]
RECALL MISS: 4380947
1979_NCAA_Men's_Division_I_Basketball_Tournament]
10: 13209000 Georgia_Bulldogs_basketball
[winner: 1008603 Georgia_national_football_team,
pem_pos=4, final_scores=-0.13672161,
final_local_scores=-0.07908893, pem=-3.736231,
similarity=22.047747]

gt: 13209000 Georgia_Bulldogs_basketball, pem_pos=12,
 final_scores=-0.42476064, final_local_scores=-0.47783035,
 pem=-4.8529596, similarity=16.22145]
 11: 4381004 1983_NCAA_Men's_Division_I_Basketball_Tournament
 [winner: 34755 1983, pem_pos=0, final_scores=-0.6851489,
 final_local_scores=-0.5448249, pem=-0.6866735,
 similarity=-8.434711|
 RECALL MISS: 4381004
 1983_NCAA_Men's_Division_I_Basketball_Tournament]
 12: 4380850 1972_NCAA_Men's_Division_I_Basketball_Tournament
 [winner: 34671 1972, pem_pos=0, final_scores=-0.75908256,
 final_local_scores=-0.6790003, pem=-0.7308857,
 similarity=-12.198341|
 RECALL MISS: 4380850
 1972_NCAA_Men's_Division_I_Basketball_Tournament]
 13: 4396310 USC_Trojans_men's_basketball
 [winner: 32005 University_of_Southern_California,
 pem_pos=0, final_scores=0.13162275,
 final_local_scores=0.11151378, pem=-0.61051047,
 similarity=10.737557|
 gt: 4396310 USC_Trojans_men's_basketball, pem_pos=4,
 final_scores=-0.26649767, final_local_scores=-0.31490242,
 pem=-3.2481103, similarity=12.356126]
 14: 3258978 Kansas_Jayhawks_men's_basketball
 [winner: 16716 Kansas, pem_pos=0,
 final_scores=0.061764643, final_local_scores=0.06485987,
 pem=-0.3846871, similarity=8.117249|
 gt: 3258978 Kansas_Jayhawks_men's_basketball, pem_pos=4,
 final_scores=-0.36531013, final_local_scores=-0.445288,
 pem=-3.4145975, similarity=9.3705435]
 1: 1399555 Florida_Gulf_Coast_University
 [pem_pos=0, final_scores=0.23321865,
 final_local_scores=0.1919424, pem=0.0, similarity=9.8176565]
 2: 672965 Atlantic_Sun_Conference
 [pem_pos=0, final_scores=0.14514147,
 final_local_scores=0.12635663, pem=0.0, similarity=7.8604465]
 3: 22239452 Dave_Balza
 [pem_pos=0, final_scores=0.14661606,
 final_local_scores=0.20760813, pem=0.0, similarity=10.285153]
 4: 38893193 Andy_Enfield
 [pem_pos=0, final_scores=0.31146574,
 final_local_scores=0.33145007, pem=0.0, similarity=13.9808445]
 5: 6616761 Joe_Dooley
 [pem_pos=0, final_scores=0.20743398,
 final_local_scores=0.3144271, pem=0.0, similarity=13.472845]
 1: 1399555 Florida_Gulf_Coast_University
 [pem_pos=0, final_scores=0.23486967,
 final_local_scores=0.19374588, pem=0.0, similarity=9.871476]
 2: 109328 Fort_Myers,_Florida
 [pem_pos=0, final_scores=0.47000152,
 final_local_scores=0.50188184, pem=0.0, similarity=19.066868]
 3: 18933066 Florida

[pem_pos=0, final_scores=0.19788359,
final_local_scores=0.21155775, pem=-0.15405534,
similarity=11.240795]
4: 3434750 United_States
[pem_pos=0, final_scores=0.40364307,
final_local_scores=0.45301068, pem=-0.073457584,
similarity=18.007929]
5: 1399555 Florida_Gulf_Coast_University
[pem_pos=0, final_scores=0.066725716,
final_local_scores=0.010074168, pem=0.0, similarity=4.390346]
6: 1399555 Florida_Gulf_Coast_University
[pem_pos=0, final_scores=0.115844056,
final_local_scores=0.06372847, pem=0.0, similarity=5.9914975]
7: 669931 Division_I_(NCAA)
[pem_pos=1, final_scores=0.04547392,
final_local_scores=-0.035389334, pem=-0.7993879,
similarity=7.38082]
8: 1399555 Florida_Gulf_Coast_University
[pem_pos=0, final_scores=0.109610096,
final_local_scores=0.056918815, pem=0.0, similarity=5.788284]
9: 60706 National_Collegiate_Athletic_Association
[pem_pos=0, final_scores=0.039570495,
final_local_scores=-0.03638518, pem=0.0, similarity=3.0039067]
10: 669931 Division_I_(NCAA)
[pem_pos=1, final_scores=0.1089562,
final_local_scores=0.039077222, pem=-0.7993879,
similarity=9.60305]
11: 3493898 Georgetown_Hoyas
[pem_pos=0, final_scores=0.10861643,
final_local_scores=0.07989849, pem=0.0, similarity=6.4740434]
12: 6767670 San_Diego_State_Aztecs
[pem_pos=0, final_scores=0.15094243,
final_local_scores=0.13483128, pem=0.0, similarity=8.113347]
13: 77795 ESPN
[pem_pos=0, final_scores=0.1786596,
final_local_scores=0.1507689, pem=-0.09327439,
similarity=9.096198]
14: 37653 CBS
[pem_pos=0, final_scores=0.13100575,
final_local_scores=0.16137785, pem=-0.10726234,
similarity=9.488859]
15: 669931 Division_I_(NCAA)
[pem_pos=1, final_scores=0.11098628,
final_local_scores=0.03929168, pem=-0.7993879,
similarity=9.60945]
16: 507399 NCAA_Men's_Division_I_Basketball_Championship
[pem_pos=0, final_scores=0.22019611,
final_local_scores=0.17617589, pem=-0.54612035,
similarity=12.31704]
17: 672965 Atlantic_Sun_Conference
[pem_pos=0, final_scores=0.10367979,
final_local_scores=0.08106603, pem=0.0, similarity=6.508885]

18: 507399 NCAA_Men's_Division_I_Basketball_Championship
 [pem_pos=0, final_scores=0.12596436,
 final_local_scores=0.07324202, pem=-0.54612035,
 similarity=9.245288]
 19: 100040 Richmond_Football_Club
 [pem_pos=2, final_scores=0.10059138,
 final_local_scores=0.15146911, pem=-2.6048384,
 similarity=22.77537]
 20: 286083 San_Diego_State_University
 [pem_pos=0, final_scores=0.054670706,
 final_local_scores=0.010554671, pem=-0.47265834,
 similarity=6.975075]
 21: 60706 National_Collegiate_Athletic_Association
 [pem_pos=0, final_scores=0.22469772,
 final_local_scores=0.16932401, pem=-0.19246244,
 similarity=10.18932]
 22: 380922 Indiana_State_University
 [pem_pos=0, final_scores=0.09129174,
 final_local_scores=0.0646002, pem=-0.38142183,
 similarity=8.0917425]
 23: 8367801 Final_four
 [pem_pos=0, final_scores=0.101894036,
 final_local_scores=0.067466006, pem=-0.29987732,
 similarity=7.7338123]
 24: 60706 National_Collegiate_Athletic_Association
 [pem_pos=0, final_scores=0.110434756,
 final_local_scores=0.044509083, pem=-0.19246244,
 similarity=6.4645934]
 25: 32005 University_of_Southern_California
 [pem_pos=0, final_scores=0.13162275,
 final_local_scores=0.11151378, pem=-0.61051047,
 similarity=10.737557]
 26: 16716 Kansas
 [pem_pos=0, final_scores=0.061764643,
 final_local_scores=0.06485987, pem=-0.3846871,
 similarity=8.117249]

Sample: b'Houston_College_Classic'

The Houston College Classic is an annual six-team **1 college baseball** tournament held in **[1 1 Houston 1]** . The 12th annual tournament was held at **[2 2 Minute Maid Park 2]** in March 2012 . The tournament coincides with the **[3 3 Houston Astros 3]** ' Fan Fest . It was initially called Astros College Classic from 2001 through 2002 , and the Minute Maid Park College Classic from 2003 through 2007 . The Classic is seen as the start of the college baseball season in **[1 Texas 1]** . At the end of the tournament , a **2 Most Outstanding Player** and an All-Tournament Team are announced . The **[2 3 Houston Cougars 2]** and **[3 4 Rice Owls 3]** , **5 rivals** in **[4 4 Conference USA 4]** during conference play , compete in the tournament annually . The other four teams invited to the tournament alternate by year . Often , teams competing in the Classic are highly

ranked . In 2002 , the Classic involved three teams that ended the preceding season ranked in the top 25 of the / [5 5 ESPN 5] , [6 6 Baseball America 6] , and year-end polls : the [4 Rice Owls 4] , [5 6 Texas Tech Red Raiders 5] , and [6 7 Baylor Bears 6] . The 2005 Classic included four top ranked teams : the # 11 [7 8 Baylor Bears 7] , # 13 9 Texas A&M Aggies , # 19 [8 Rice Owls 8] , and # 23 [9 10 Oklahoma State Cowboys 9] . ranked all six teams competing in the 2007 Classic : they ranked the [10 Rice Owls 10] the top team in college baseball , and also ranked the # 8 [11 11 Vanderbilt Commodores 11] , # 14 [12 12 Arizona State Sun Devils 12] , # 21 [13 Baylor Bears 13] , and # 33 [14 Houston Cougars 14] . The [15 13 Texas Longhorns 15] and [16 Rice Owls 16] participated in the Classic when they won the [7 7 College World Series 7] championship in 2002 and 2003 , respectively . The [17 Texas Tech Red Raiders 17]

1: 13774 Houston
 2: 281603 Minute_Maid_Park
 3: 13894 Houston_Astros
 4: 78428 Conference_USA
 5: 77795 ESPN
 6: 634924 Baseball_America
 7: 536107 College_World_Series

1: 1370306 College_baseball

[winner: 5512063 NCAA_Division_I_Baseball_Championship,
 pem_pos=0, final_scores=-1.1526176,
 final_local_scores=-1.2970023, pem=0.0,
 similarity=-34.615425]

gt: 1370306 College_baseball, pem_pos=1,
 final_scores=-1.201772, final_local_scores=-1.3439915,
 pem=0.0, similarity=-36.017673]

2: 235000 Most_Valuable_Player

[winner: 1562112

NCAA_Basketball_Tournament_Most_Outstanding_Player,
 pem_pos=0, final_scores=-0.07193074,
 final_local_scores=-0.09207338, pem=-0.24130577,
 similarity=2.6543179]

gt: 235000 Most_Valuable_Player, pem_pos=3,
 final_scores=-0.39799523, final_local_scores=-0.4735716,
 pem=-1.4074113, similarity=-2.3888857]

3: 9685432 Houston_Cougars_baseball

[winner: 1941776 Houston_Cougars, pem_pos=0,
 final_scores=0.08492504, final_local_scores=0.028011948,
 pem=-0.37240267, similarity=6.950829]

gt: 9685432 Houston_Cougars_baseball, pem_pos=3,
 final_scores=-0.18125457, final_local_scores=-0.22184324,
 pem=-2.0113332, similarity=8.407411]

4: 17848400 Rice_Owls_baseball

[winner: 8886023 Rice_Owls, pem_pos=0,
 final_scores=0.17157622, final_local_scores=0.106394485,
 pem=-0.19374274, similarity=8.318339]

gt: 17848400 Rice_Owls_baseball, pem_pos=2,

final_scores=-0.031334385, final_local_scores=-0.07292125,
 pem=-1.4432188, similarity=9.762046]
 5: 16957975 Houston-Rice_rivalry
 [winner: 6082538
 Notre_Dame_Fighting_Irish_football_rivalries, pem_pos=13,
 final_scores=-1.549168, final_local_scores=-1.6851783,
 pem=-2.8088043, similarity=-30.924652]
 RECALL MISS: 16957975 Houston-Rice_rivalry]
 6: 17458573 Texas_Tech_Red_Raiders_baseball
 [winner: 3694376 Texas_Tech_Red_Raiders_and_Lady_Raiders,
 pem_pos=0, final_scores=0.20652737,
 final_local_scores=0.1409063, pem=-0.32595634,
 similarity=10.067239]
 gt: 17458573 Texas_Tech_Red_Raiders_baseball, pem_pos=3,
 final_scores=-0.101452604, final_local_scores=-0.13528088,
 pem=-1.8914113, similarity=10.33845]
 7: 21107627 Baylor_Bears_baseball
 [winner: 3694292 Baylor_Bears, pem_pos=0,
 final_scores=0.1423992, final_local_scores=0.07757664,
 pem=-0.24583499, similarity=7.741643]
 gt: 21107627 Baylor_Bears_baseball, pem_pos=3,
 final_scores=-0.13244872, final_local_scores=-0.19158095,
 pem=-1.7398746, similarity=7.834264]
 8: 21107627 Baylor_Bears_baseball
 [winner: 3694292 Baylor_Bears, pem_pos=0,
 final_scores=0.14900707, final_local_scores=0.08563036,
 pem=-0.24583499, similarity=7.981982]
 gt: 21107627 Baylor_Bears_baseball, pem_pos=3,
 final_scores=-0.095420435, final_local_scores=-0.15123934,
 pem=-1.7398746, similarity=9.038138]
 9: 18455992 Texas_A&M_Aggies_baseball
 [winner: 29810 Texas, pem_pos=0,
 final_scores=-0.98686886, final_local_scores=-1.0634131,
 pem=-0.16191302, similarity=-26.76415]
 RECALL MISS: 18455992 Texas_A&M_Aggies_baseball]
 [winner: 13366748 Glossary_of_Texas_A&M_University_terms,
 pem_pos=4, final_scores=-1.0671804,
 final_local_scores=-1.0555303, pem=-2.1361368,
 similarity=-15.792776]
 gt: 18455992 Texas_A&M_Aggies_baseball, pem_pos=7,
 final_scores=-1.408583, final_local_scores=-1.5662522,
 pem=-2.4661853, similarity=-29.238876]
 [winner: 7864249 A&M_Consolidated_High_School, pem_pos=1,
 final_scores=-1.6236377, final_local_scores=-1.7178655,
 pem=-0.6931472, similarity=-43.40538]
 RECALL MISS: 18455992 Texas_A&M_Aggies_baseball]
 10: 21111877 Oklahoma_State_Cowboys_baseball
 [winner: 3694342 Oklahoma_State_Cowboys_and_Cowgirls,
 pem_pos=0, final_scores=0.2924834,
 final_local_scores=0.25607893, pem=-0.22146514,
 similarity=12.935981]
 gt: 21111877 Oklahoma_State_Cowboys_baseball,

pem_pos=10, final_scores=-0.6144024,
final_local_scores=-0.68536377, pem=-5.09375,
similarity=11.337695]
11: 16638935 Vanderbilt_Commodores_baseball
[winner: 8168643 Vanderbilt_Commodores_football,
pem_pos=1, final_scores=0.11466353,
final_local_scores=0.059611484, pem=-0.71488714,
similarity=9.756306]
gt: 16638935 Vanderbilt_Commodores_baseball, pem_pos=2,
final_scores=-0.22453105, final_local_scores=-0.24744749,
pem=-1.5645188, similarity=5.2134824]
12: 5707227 Arizona_State_Sun_Devils_baseball
[winner: 4578009 Arizona_State_Sun_Devils, pem_pos=0,
final_scores=0.28156012, final_local_scores=0.2471382,
pem=-0.1781972, similarity=12.433874]
gt: 5707227 Arizona_State_Sun_Devils_baseball,
pem_pos=3, final_scores=-0.07153973,
final_local_scores=-0.115264595, pem=-2.1127071,
similarity=12.139217]
13: 10848103 Texas_Longhorns_baseball
[winner: 2834813 Texas_Longhorns, pem_pos=0,
final_scores=0.25188023, final_local_scores=0.18966937,
pem=-0.50125355, similarity=12.475719]
gt: 10848103 Texas_Longhorns_baseball, pem_pos=3,
final_scores=-0.20068996, final_local_scores=-0.26199496,
pem=-2.4782653, similarity=9.748453]
1: 13774 Houston
[pem_pos=0, final_scores=0.567614,
final_local_scores=0.6165052, pem=-0.3232624,
similarity=24.245409]
2: 281603 Minute_Maid_Park
[pem_pos=0, final_scores=0.29617995,
final_local_scores=0.32517383, pem=0.0, similarity=13.793549]
3: 13894 Houston_Astros
[pem_pos=0, final_scores=0.26210374,
final_local_scores=0.2701043, pem=-0.049621258,
similarity=12.4200115]
4: 78428 Conference_USA
[pem_pos=0, final_scores=0.120094165,
final_local_scores=0.07112993, pem=0.0, similarity=6.212372]
5: 77795 ESPN
[pem_pos=0, final_scores=0.16516401,
final_local_scores=0.139664, pem=-0.09327439,
similarity=8.764806]
6: 634924 Baseball_America
[pem_pos=0, final_scores=0.08361401,
final_local_scores=0.08499806, pem=0.0, similarity=6.6262245]
7: 536107 College_World_Series
[pem_pos=0, final_scores=0.3185916,
final_local_scores=0.28985187, pem=0.0, similarity=12.739471]
1: 29810 Texas
[pem_pos=0, final_scores=0.6367713,

final_local_scores=0.7192704, pem=-0.16191302,
similarity=26.434685]
2: 1941776 Houston_Cougars
[pem_pos=0, final_scores=0.08492504,
final_local_scores=0.028011948, pem=-0.37240267,
similarity=6.950829]
3: 8886023 Rice_Owls
[pem_pos=0, final_scores=0.17157622,
final_local_scores=0.106394485, pem=-0.19374274,
similarity=8.318339]
4: 8886023 Rice_Owls
[pem_pos=0, final_scores=0.20709528,
final_local_scores=0.14556512, pem=-0.19374274,
similarity=9.48727]
5: 3694376 Texas_Tech_Red_Raiders_and_Lady_Raiders
[pem_pos=0, final_scores=0.20652737,
final_local_scores=0.1409063, pem=-0.32595634,
similarity=10.067239]
6: 3694292 Baylor_Bears
[pem_pos=0, final_scores=0.1423992,
final_local_scores=0.07757664, pem=-0.24583499,
similarity=7.741643]
7: 3694292 Baylor_Bears
[pem_pos=0, final_scores=0.14900707,
final_local_scores=0.08563036, pem=-0.24583499,
similarity=7.981982]
8: 8886023 Rice_Owls
[pem_pos=0, final_scores=0.17418067,
final_local_scores=0.10954891, pem=-0.19374274,
similarity=8.412474]
9: 3694342 Oklahoma_State_Cowboys_and_Cowgirls
[pem_pos=0, final_scores=0.2924834,
final_local_scores=0.25607893, pem=-0.22146514,
similarity=12.935981]
10: 8886023 Rice_Owls
[pem_pos=0, final_scores=0.133821,
final_local_scores=0.06546211, pem=-0.19374274,
similarity=7.096836]
11: 8168643 Vanderbilt_Commodores_football
[pem_pos=1, final_scores=0.11466353,
final_local_scores=0.059611484, pem=-0.71488714,
similarity=9.756306]
12: 4578009 Arizona_State_Sun_Devils
[pem_pos=0, final_scores=0.28156012,
final_local_scores=0.2471382, pem=-0.1781972,
similarity=12.433874]
13: 3694292 Baylor_Bears
[pem_pos=0, final_scores=0.10626735,
final_local_scores=0.03810808, pem=-0.24583499,
similarity=6.5638227]
14: 1941776 Houston_Cougars
[pem_pos=0, final_scores=0.112254396,

```

final_local_scores=0.057865113, pem=-0.37240267,
similarity=7.841707]
15: 2834813 Texas_Longhorns
    [pem_pos=0, final_scores=0.25188023,
final_local_scores=0.18966937, pem=-0.50125355,
similarity=12.475719]
16: 8886023 Rice_Owls
    [pem_pos=0, final_scores=0.20342092,
final_local_scores=0.14148945, pem=-0.19374274,
similarity=9.3656435]
17: 3694376 Texas_Tech_Red_Raiders_and_Lady_Raiders
    [pem_pos=0, final_scores=0.2234423,
final_local_scores=0.1593833, pem=-0.32595634,
similarity=10.618629]

```

Sample: b'Maneer_Mirza'

Maneer Mohammed Mirza -LRB- born 1 April 1978 in [1 1 Birmin
gham 1] -RRB- is an [2 2 English 2] : a right-arm fast-medium
bowler and right-handed batsman who played for [3 3 Worcester
shire 3] . He is the younger brother of another former [1
Worcestershire 1] player , [4 4 Parvaz Mirza 4] , who died
aged just 24 in 1995 . After a number of second-team
appearances the previous year , [2 Mirza 2] made his 1 first-
class debut in [3 Worcestershire 3] 's drawn game against [4 2
Pakistan 4] A at [5 5 New Road 5] in July 3 1997 . He took
3-136 in his only innings , and two weeks later made his [6 6
County Championship 6] debut against [7 7 Kent 7] , taking
three wickets in the match . Having reached the first team he
remained there for the rest of the season , playing a total
of six first-class games and taking 19 wickets 4 at 32.63 ;
he also appeared in four [8 8 List A 8] games , but took just
one wicket in total . Still only 20 , [5 Mirza 5] 's career
was disastrously affected in July 5 1998 , when after a
handful of second-eleven games he suffered a 6 stress fracture
of the back and was unable to play again that season . He
returned for 7 1999 and 8 2000 , but never managed to force
his way back into the First XI , and eventually left [6
Worcestershire 6] altogether . He appeared once each for [9 9
Surrey 9] 's and [10 10 Leicestershire 10] 's second teams in
9 2002

```

1: 13882056 Birmingham
2: 9316 England
3: 1156428 Worcestershire_County_Cricket_Club
4: 2931611 Parvaz_Mirza
5: 1893040 New_Road,_Worcester
6: 683599 County_Championship
7: 1177719 Kent_County_Cricket_Club
8: 683459 List_A_cricket
9: 766575 Surrey_County_Cricket_Club
10: 1622318 Leicestershire_County_Cricket_Club

```

```

1: 11040 First-class_cricket
    [winner: 662844 First-class_citizen, pem_pos=3,

```

final_scores=-1.3602272, final_local_scores=-1.2833326,
 pem=-1.5935546, similarity=-25.541492|
 gt: 11040 First-class_cricket, pem_pos=0,
 final_scores=-1.4729049, final_local_scores=-1.6778905,
 pem=-0.053819276, similarity=-45.689213]
 2: 26123716 Pakistan_A_cricket_team
 [winner: 759295 Pakistan_national_cricket_team,
 pem_pos=1, final_scores=0.5776815,
 final_local_scores=0.571428, pem=-2.7556365,
 similarity=36.127842|
 RECALL MISS: 26123716 Pakistan_A_cricket_team]
 3: 7410667 1997_English_cricket_season
 [winner: 34601 1997, pem_pos=0, final_scores=-1.0899868,
 final_local_scores=-1.0097983, pem=-0.69383013,
 similarity=-22.271524|
 RECALL MISS: 7410667 1997_English_cricket_season]
 4: 875871 Bowling_average
 [winner: 875871 Bowling_average, pem_pos=0,
 final_scores=-0.8343716, final_local_scores=-0.95893264,
 pem=0.0, similarity=-24.52675|
 gt: 875871 Bowling_average, pem_pos=0,
 final_scores=-0.8343716, final_local_scores=-0.95893264,
 pem=0.0, similarity=-24.52675]
 5: 7635219 1998_English_cricket_season
 [winner: 34647 1998, pem_pos=0, final_scores=-1.1034423,
 final_local_scores=-1.0940952, pem=-0.6850515,
 similarity=-24.834856|
 RECALL MISS: 7635219 1998_English_cricket_season]
 6: 1173689 Stress_fracture
 [winner: 17234794 Fracture_(mineralogy), pem_pos=0,
 final_scores=-0.9960323, final_local_scores=-0.86214066,
 pem=-0.7096434, similarity=-17.779133|
 RECALL MISS: 1173689 Stress_fracture]
 [winner: 10417992 Muri_(Japanese_term), pem_pos=5,
 final_scores=-1.1103672, final_local_scores=-0.97555465,
 pem=-1.6091778, similarity=-16.271824|
 gt: 1173689 Stress_fracture, pem_pos=2,
 final_scores=-1.4914591, final_local_scores=-1.5139863,
 pem=-1.5985733, similarity=-32.397366]
 7: 7635102 1999_English_cricket_season
 [winner: 34553 1999, pem_pos=0, final_scores=-1.311618,
 final_local_scores=-1.3200731, pem=-0.6240096,
 similarity=-31.910439|
 RECALL MISS: 7635102 1999_English_cricket_season]
 8: 5672162 2000_English_cricket_season
 [winner: 432383 2000_United_States_Census, pem_pos=1,
 final_scores=-0.5051803, final_local_scores=-0.4376638,
 pem=-1.4367632, similarity=-1.157706|
 RECALL MISS: 5672162 2000_English_cricket_season]
 9: 7627600 2002_English_cricket_season
 [winner: 35502 2002, pem_pos=0, final_scores=-0.6216282,
 final_local_scores=-0.49447423, pem=-0.5897454,

```
similarity=-7.459257|
RECALL MISS: 7627600 2002_English_cricket_season]
1: 13882056 Birmingham
    [pem_pos=0, final_scores=0.50376433,
final_local_scores=0.5452167, pem=-0.40171203,
similarity=22.544638]
2: 9316 England
    [pem_pos=1, final_scores=0.5275431,
final_local_scores=0.5698376, pem=-1.3527626,
similarity=28.451336]
3: 1156428 Worcestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.55740374,
final_local_scores=0.505811, pem=-0.9164197,
similarity=24.167753]
4: 2931611 Parvaz_Mirza
    [pem_pos=0, final_scores=0.25091714,
final_local_scores=0.2783631, pem=0.0, similarity=12.396624]
5: 1893040 New_Road, Worcester
    [pem_pos=0, final_scores=0.48482198,
final_local_scores=0.46248764, pem=0.0, similarity=17.891266]
6: 683599 County_Championship
    [pem_pos=0, final_scores=0.29345912,
final_local_scores=0.23640269, pem=0.0, similarity=11.144441]
7: 1177719 Kent_County_Cricket_Club
    [pem_pos=1, final_scores=0.65407,
final_local_scores=0.6190252, pem=-2.5096242,
similarity=36.21039]
8: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.26034003,
final_local_scores=0.21837592, pem=-0.020267159,
similarity=10.7167015]
9: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.6482621,
final_local_scores=0.6113307, pem=-1.7952642,
similarity=32.095966]
10: 1622318 Leicestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6633889,
final_local_scores=0.63396484, pem=-1.0465667,
similarity=28.69988]
1: 1156428 Worcestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.69737715,
final_local_scores=0.65871066, pem=-0.9164197,
similarity=28.730585]
2: 2931611 Parvaz_Mirza
    [pem_pos=1, final_scores=0.28403193,
final_local_scores=0.31453598, pem=0.0, similarity=13.476094]
3: 1156428 Worcestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.72031635,
final_local_scores=0.6837682, pem=-0.9164197,
similarity=29.478354]
4: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.5776815,
```

```

final_local_scores=0.571428, pem=-2.7556365,
similarity=36.127842]
5: 2931611 Parvaz_Mirza
[pem_pos=1, final_scores=0.30872083,
final_local_scores=0.34150487, pem=0.0, similarity=14.280899]
6: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.573675,
final_local_scores=0.5235848, pem=-0.9164197,
similarity=24.69816]

```

Sample: b'Panhandle'

A panhandle is an [1 American 1] informal 1 geographic term for an elongated arm-like protrusion of a geo-political entity , such as a or a 2 sovereign state . The term ``panhandle `` derives from the analogous part of a , and its use is generally confined to the [1 1 United States 1] . A term used elsewhere is salient , derived from . While similar to a 3 peninsula in shape , a panhandle is not surrounded by water on three sides and connected to a 4 geographical 5 main land . Instead , it is delimited by a land 6 border on at least two sides and extends out from the larger geographical body of the administrative unit . The panhandle shape is the result of arbitrarily drawn international or boundaries , although the location of some administrative borders takes into account other considerations such as 7 economic ties or 8 topography . In the [2 United States 2] , a protrusion with a less elongated shape is informally called a 9 bootheel . † This definition of the Florida panhandle includes the following counties : 10 Bay , 11 Calhoun , [2 2 Escambia 2] , 12 Franklin , [3 13 Gadsden 3] , 14 Gulf , 15 Holmes , 16 Jackson , 17 Jefferson , 18 Leon , 19 Liberty , [4 20 Madison 4] , [3 3 Okaloosa 3] , [5 21 Santa Rosa 5] , 22 Taylor , [4 4 Wakulla 4] , 23 Walton , and [6 24 Washington 6] . Outside the [7 United States 7] , the term is not in common usage , with the arguable exception of the nearby [5 5 New Brunswick Panhandle 5]

```

1: 3434750 United_States
2: 72357 Escambia_County,_Florida
3: 73632 Okaloosa_County,_Florida
4: 75116 Wakulla_County,_Florida
5: 171250 Madawaska_County,_New_Brunswick
1: 1845092 Political_geography
[winner: 4951472 Georeference, pem_pos=1,
final_scores=-1.1414818, final_local_scores=-1.1018014,
pem=-1.164494, similarity=-22.457539]
gt: 1845092 Political_geography, pem_pos=8,
final_scores=-2.2274115, final_local_scores=-2.3687005,
pem=-5.2916284, similarity=-37.820335 ]
2: 23604120 Sovereign_state
[winner: 445321 Political_divisions_of_the_United_States,
pem_pos=0, final_scores=-1.3810967,
final_local_scores=-1.4596443, pem=-0.5713872,

```

```

similarity=-36.361694|
    RECALL MISS: 23604120 Sovereign_state]
    [winner: 266325 Sovereign_(British_coin), pem_pos=1,
    final_scores=-1.4637051, final_local_scores=-1.4893913,
    pem=-1.2852074, similarity=-33.36754|
        gt: 23604120 Sovereign_state, pem_pos=9,
    final_scores=-2.4117284, final_local_scores=-2.5946887,
    pem=-5.6624813, similarity=-42.54752 ]
3: 23260 Peninsula
    [winner: 8413400 Shawmut_Peninsula, pem_pos=0,
    final_scores=-1.1049031, final_local_scores=-1.1317825,
    pem=0.0, similarity=-29.684937|
    RECALL MISS: 23260 Peninsula]
4: 23263 Physical_geography
    [winner: 23263 Physical_geography, pem_pos=1,
    final_scores=-1.3648937, final_local_scores=-1.4397643,
    pem=-0.67928773, similarity=-35.181652|
        gt: 23263 Physical_geography, pem_pos=1,
    final_scores=-1.3648937, final_local_scores=-1.4397643,
    pem=-0.67928773, similarity=-35.181652 ]
5: 286960 Mainland
    [winner: 286960 Mainland, pem_pos=0,
    final_scores=-0.96356493, final_local_scores=-1.0135058,
    pem=-0.6685694, similarity=-22.519537|
        gt: 286960 Mainland, pem_pos=0,
    final_scores=-0.96356493, final_local_scores=-1.0135058,
    pem=-0.6685694, similarity=-22.519537 ]
6: 247613 Border
    [winner: 24411110 Ibrahim_Khalil_border, pem_pos=13,
    final_scores=-1.2618419, final_local_scores=-1.2302399,
    pem=-2.6358812, similarity=-18.28877|
        gt: 247613 Border, pem_pos=0, final_scores=-1.5826994,
    final_local_scores=-1.7161736, pem=-0.5865604,
    similarity=-43.934525 ]
7: 177512 Economic_geography
    [winner: 2110221 Tie_(engineering), pem_pos=0,
    final_scores=-1.0677118, final_local_scores=-0.99535525,
    pem=-0.6761832, similarity=-21.936483|
    RECALL MISS: 177512 Economic_geography]
    [winner: 10822780 2006-
07_economic_sanctions_against_the_Palestinian_National_Author
ity, pem_pos=9, final_scores=-1.3998325,
    final_local_scores=-1.407269, pem=-2.1972246,
    similarity=-25.957151|
    RECALL MISS: 177512 Economic_geography]
8: 79239 Topography
    [winner: 70713 Topographic_map, pem_pos=0,
    final_scores=-1.2350991, final_local_scores=-1.3127065,
    pem=0.0, similarity=-35.084064|
        gt: 79239 Topography, pem_pos=1,
    final_scores=-1.4305224, final_local_scores=-1.5405982,
    pem=-0.012113147, similarity=-41.81894 ]

```


9: 851935 Bootheel
[winner: 93407 Hidalgo County, New Mexico, pem_pos=1,
final_scores=-1.2995613, final_local_scores=-1.3841808,
pem=-0.6911748, similarity=-33.45829]
gt: 851935 Bootheel, pem_pos=0, final_scores=-1.3569803,
final_local_scores=-1.4759157, pem=0.0,
similarity=-39.954556]

10: 72004 Bay County, Florida
[winner: 72004 Bay County, Florida, pem_pos=2,
final_scores=-0.13632694, final_local_scores=-0.17382833,
pem=-2.5174596, similarity=12.592666]
gt: 72004 Bay County, Florida, pem_pos=2,
final_scores=-0.13632694, final_local_scores=-0.17382833,
pem=-2.5174596, similarity=12.592666]

11: 72011 Calhoun County, Florida
[winner: 52110 John C. Calhoun, pem_pos=0,
final_scores=-0.24222447, final_local_scores=-0.24766254,
pem=-1.9563721, similarity=7.3380246]
gt: 72011 Calhoun County, Florida, pem_pos=2,
final_scores=-0.33000785, final_local_scores=-0.34638843,
pem=-2.5573304, similarity=7.6599507]

12: 72362 Franklin County, Florida
[winner: 135422 Franklin, Tennessee, pem_pos=0,
final_scores=-0.30996794, final_local_scores=-0.36699638,
pem=-3.2882125, similarity=11.01962]
gt: 72362 Franklin County, Florida, pem_pos=18,
final_scores=-0.44110978, final_local_scores=-0.4776124,
pem=-4.098005, similarity=12.122394]

13: 72364 Gadsden County, Florida
[winner: 104690 Gadsden, Alabama, pem_pos=0,
final_scores=0.12171529, final_local_scores=0.0985457,
pem=-0.49837247, similarity=9.740741]
gt: 72364 Gadsden County, Florida, pem_pos=1,
final_scores=-0.078916356, final_local_scores=-0.10818258,
pem=-1.5537918, similarity=9.31109]

14: 72369 Gulf County, Florida
[winner: 21076367 Gulf of Mexico, pem_pos=1,
final_scores=-0.06703137, final_local_scores=-0.060537994,
pem=-1.568955, similarity=10.815359]
gt: 72369 Gulf County, Florida, pem_pos=3,
final_scores=-0.38168037, final_local_scores=-0.4114556,
pem=-2.0428784, similarity=2.9205477]

15: 72379 Holmes County, Florida
[winner: 93067 Holmes County, Ohio, pem_pos=5,
final_scores=-0.3785572, final_local_scores=-0.42578307,
pem=-3.002612, similarity=7.712169]
gt: 72379 Holmes County, Florida, pem_pos=2,
final_scores=-0.39514238, final_local_scores=-0.41172075,
pem=-2.6879237, similarity=6.4204903]

16: 72385 Jackson County, Florida
[winner: 16167 Jackson, Mississippi, pem_pos=0,
final_scores=-0.0744748, final_local_scores=-0.14717242,

pem=-2.1336284, similarity=11.300797|
gt: 72385 Jackson_County, Florida, pem_pos=16,
final_scores=-0.68752325, final_local_scores=-0.76449543,
pem=-4.2241497, similarity=4.247227]
17: 72388 Jefferson_County, Florida
[winner: 29922 Thomas_Jefferson, pem_pos=0,
final_scores=-0.34747994, final_local_scores=-0.38447225,
pem=-2.438518, similarity=5.877334|
gt: 72388 Jefferson_County, Florida, pem_pos=14,
final_scores=-0.57966137, final_local_scores=-0.6475027,
pem=-3.8746738, similarity=5.838018]
18: 73612 Leon_County, Florida
[winner: 73612 Leon_County, Florida, pem_pos=0,
final_scores=-0.28367752, final_local_scores=-0.33707651,
pem=-2.1715562, similarity=5.8399377|
gt: 73612 Leon_County, Florida, pem_pos=0,
final_scores=-0.28367752, final_local_scores=-0.33707651,
pem=-2.1715562, similarity=5.8399377]
19: 73618 Liberty_County, Florida
[winner: 456375 Liberty_University, pem_pos=14,
final_scores=-0.44753858, final_local_scores=-0.47777358,
pem=-3.8730142, similarity=10.89405|
gt: 73618 Liberty_County, Florida, pem_pos=10,
final_scores=-0.68888855, final_local_scores=-0.7287733,
pem=-3.7806487, similarity=2.9014204]
20: 73621 Madison_County, Florida
[winner: 57863 Madison, Wisconsin, pem_pos=0,
final_scores=0.12828474, final_local_scores=0.08671096,
pem=-1.5699619, similarity=15.215037|
gt: 73621 Madison_County, Florida, pem_pos=17,
final_scores=-0.65084505, final_local_scores=-0.7041299,
pem=-3.9533908, similarity=4.5762267]
21: 74651 Santa_Rosa_County, Florida
[winner: 108243 Santa_Rosa, California, pem_pos=0,
final_scores=0.13733535, final_local_scores=0.117494285,
pem=-0.85449636, similarity=12.242861|
gt: 74651 Santa_Rosa_County, Florida, pem_pos=2,
final_scores=-0.15202796, final_local_scores=-0.19444883,
pem=-2.6115823, similarity=12.489162]
22: 75111 Taylor_County, Florida
[winner: 119280 Taylor, Michigan, pem_pos=0,
final_scores=-0.37019747, final_local_scores=-0.40205544,
pem=-2.8038657, similarity=7.3394337|
gt: 75111 Taylor_County, Florida, pem_pos=4,
final_scores=-0.51421136, final_local_scores=-0.5501466,
pem=-3.597992, similarity=7.2386804]
23: 75118 Walton_County, Florida
[winner: 96683 Walton_County, Georgia, pem_pos=3,
final_scores=-0.32320476, final_local_scores=-0.3572697,
pem=-2.8238256, similarity=8.784475|
gt: 75118 Walton_County, Florida, pem_pos=2,
final_scores=-0.32624495, final_local_scores=-0.38013676,

pem=-2.7848344, similarity=7.890036]
24: 75122 Washington_County,_Florida
[winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=0.32149833, final_local_scores=0.28520274,
pem=-1.7496481, similarity=22.11559]
RECALL MISS: 75122 Washington_County,_Florida]
1: 3434750 United_States
[pem_pos=0, final_scores=0.53868204,
final_local_scores=0.50242, pem=-0.073457584,
similarity=19.482403]
2: 72357 Escambia_County,_Florida
[pem_pos=0, final_scores=0.15193824,
final_local_scores=0.14239022, pem=-0.28650078,
similarity=9.896956]
3: 73632 Okaloosa_County,_Florida
[pem_pos=0, final_scores=0.077796146,
final_local_scores=0.0842765, pem=0.0, similarity=6.6046915]
4: 75116 Wakulla_County,_Florida
[pem_pos=0, final_scores=0.11249523,
final_local_scores=0.11720091, pem=-0.24874793,
similarity=8.939952]
5: 171250 Madawaska_County,_New_Brunswick
[pem_pos=0, final_scores=0.13301642,
final_local_scores=0.22728083, pem=0.0, similarity=10.872226]
1: 3434750 United_States
[pem_pos=0, final_scores=0.49722856,
final_local_scores=0.45713836, pem=-0.12994486,
similarity=18.438295]
2: 3434750 United_States
[pem_pos=0, final_scores=0.49423003,
final_local_scores=0.4538629, pem=-0.073457584,
similarity=18.033361]
3: 104690 Gadsden,_Alabama
[pem_pos=0, final_scores=0.12171529,
final_local_scores=0.0985457, pem=-0.49837247,
similarity=9.740741]
4: 57863 Madison,_Wisconsin
[pem_pos=0, final_scores=0.12828474,
final_local_scores=0.08671096, pem=-1.5699619,
similarity=15.215037]
5: 108243 Santa_Rosa,_California
[pem_pos=0, final_scores=0.13733535,
final_local_scores=0.117494285, pem=-0.85449636,
similarity=12.242861]
6: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.32149833,
final_local_scores=0.28520274, pem=-1.7496481,
similarity=22.11559]
7: 3434750 United_States
[pem_pos=0, final_scores=0.5290392,
final_local_scores=0.49188668, pem=-0.073457584,
similarity=19.168066]

Sample: b'Paul_New'

Paul Michael New -LRB- born 30 September 1953 -RRB- is a former [1 1 English 1] . New was a right-handed who bowled right-arm 2 medium-fast . He was born at [1 1 Wokingham 1] , [2 2 Berkshire 2] . New made his debut for [3 3 Berkshire 3] in 1974 against [4 4 Wiltshire 4] . From 1974 to 1987 , he represented the county in 47 [2 Minor Counties Championship 2] matches , the last of which came in the 1987 Championship when [3 Berkshire 3] played [5 5 Cheshire 5] . New also played in the [6 6 MCCA Knockout Trophy 6] for [4 Berkshire 4] . His debut in that competition came in 1983 when [5 Berkshire 5] played [7 7 Norfolk 7] . From 1983 to 1987 , he represented the county in 9 Trophy matches , the last of which came when [6 Berkshire 6] played [8 8 Buckinghamshire 8] in the 1987 [7 MCCA Knockout Trophy 7] . Additionally , he also played matches for [8 Berkshire 8] . His [9 List-A 9] debut for the county came against [9 9 Yorkshire 9] in the 3 1 983 [10 NatWest Trophy 10] . From 1983 to 1986 , he represented the county in 4 [11 List-A 11] matches , with his final [12 List-A 12] match coming in the 4 1986 [13 NatWest Trophy 13] when [14 Berkshire 14] played [10 10 Gloucestershire 10] at [11 11 Courages Cricket Ground 11] in [12 12 Reading 12] . In his 4 matches , he took 6 wickets at a 5 bowling average

1: 194885 Wokingham

2: 64951 Berkshire

3: 3702129 Berkshire_County_Cricket_Club

4: 3702411 Wiltshire_County_Cricket_Club

5: 3712477 Cheshire_County_Cricket_Club

6: 7543416 MCCA_Knockout_Trophy

7: 3690942 Norfolk_County_Cricket_Club

8: 3702814 Buckinghamshire_County_Cricket_Club

9: 869945 Yorkshire_County_Cricket_Club

10: 1622322 Gloucestershire_County_Cricket_Club

11: 28160644 Courages_Cricket_Ground

12: 43980 Reading,_Berkshire

1: 18803164 English_people

[winner: 9316 England, pem_pos=1,
final_scores=0.32415432, final_local_scores=0.35114756,
pem=-1.3527626, similarity=21.92519]

gt: 18803164 English_people, pem_pos=2,
final_scores=-0.44724557, final_local_scores=-0.47870633,
pem=-2.2084396, similarity=1.8139992]

2: 689836 Seam_bowling

[winner: 678695 Fast_bowling, pem_pos=0,
final_scores=-0.89234704, final_local_scores=-1.0164106,
pem=0.0, similarity=-26.242008]

RECALL MISS: 689836 Seam_bowling]

3: 27448269 1983_NatWest_Trophy

[winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=0.34031814, final_local_scores=0.30002275,

pem=0.0, similarity=13.042991|
gt: 27448269 1983_NatWest_Trophy, pem_pos=19,
final_scores=-0.7749463, final_local_scores=-0.8789956,
pem=-6.0222363, similarity=10.608591]
[winner: 34755 1983, pem_pos=0, final_scores=-0.7708289,
final_local_scores=-0.60938394, pem=-0.6866735,
similarity=-10.361281|
RECALL MISS: 27448269 1983_NatWest_Trophy]
4: 27490135 1986_NatWest_Trophy
[winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=0.38506466, final_local_scores=0.34890148,
pem=0.0, similarity=14.501629|
gt: 27490135 1986_NatWest_Trophy, pem_pos=17,
final_scores=-0.76303166, final_local_scores=-0.85668135,
pem=-6.0222363, similarity=11.274492]
[winner: 34761 1986, pem_pos=0, final_scores=-0.8718695,
final_local_scores=-0.7408013, pem=-0.71220666,
similarity=-14.144185|
RECALL MISS: 27490135 1986_NatWest_Trophy]
5: 875871 Bowling_average
[winner: 875871 Bowling_average, pem_pos=0,
final_scores=-1.062958, final_local_scores=-1.2008555,
pem=0.0, similarity=-31.746208|
gt: 875871 Bowling_average, pem_pos=0,
final_scores=-1.062958, final_local_scores=-1.2008555,
pem=0.0, similarity=-31.746208]
1: 194885 Wokingham
[pem_pos=0, final_scores=0.1465811,
final_local_scores=0.18099242, pem=-0.23313677,
similarity=10.758722]
2: 64951 Berkshire
[pem_pos=0, final_scores=0.2524014,
final_local_scores=0.25879785, pem=-0.18451037,
similarity=12.816154]
3: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.4079451,
final_local_scores=0.33464155, pem=-2.301417,
similarity=26.591549]
4: 3702411 Wiltshire_County_Cricket_Club
[pem_pos=1, final_scores=0.33161938,
final_local_scores=0.27000743, pem=-2.690571,
similarity=26.779016]
5: 3712477 Cheshire_County_Cricket_Club
[pem_pos=2, final_scores=0.12020855,
final_local_scores=0.045497924, pem=-2.629504,
similarity=19.747114]
6: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.45887727,
final_local_scores=0.41505504, pem=0.0, similarity=16.475782]
7: 3690942 Norfolk_County_Cricket_Club
[pem_pos=6, final_scores=0.33274382,
final_local_scores=0.26683974, pem=-3.4736512,

similarity=30.942997]
8: 3702814 Buckinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.53096944,
final_local_scores=0.478684, pem=-1.6046519,
similarity=27.100946]
9: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6924355,
final_local_scores=0.6588747, pem=-1.7779728,
similarity=33.42074]
10: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.6917503,
final_local_scores=0.66009706, pem=-1.379793,
similarity=31.291857]
11: 28160644 Courages_Cricket_Ground
[pem_pos=0, final_scores=0.1663716,
final_local_scores=0.2014117, pem=0.0, similarity=10.10024]
12: 43980 Reading, Berkshire
[pem_pos=1, final_scores=0.15893386,
final_local_scores=0.16616017, pem=-1.2339983,
similarity=15.758939]
1: 9316 England
[pem_pos=1, final_scores=0.32415432,
final_local_scores=0.35114756, pem=-1.3527626,
similarity=21.92519]
2: 7562952 Minor_Counties_Cricket_Championship
[pem_pos=0, final_scores=0.31276828,
final_local_scores=0.24414274, pem=0.0, similarity=11.37542]
3: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.45286185,
final_local_scores=0.38370627, pem=-2.301417,
similarity=28.055738]
4: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.53951615,
final_local_scores=0.47836286, pem=-2.301417,
similarity=30.88048]
5: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.59992033,
final_local_scores=0.5443453, pem=-2.301417,
similarity=32.84953]
6: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.65662855,
final_local_scores=0.6062904, pem=-2.301417,
similarity=34.698093]
7: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.29101878,
final_local_scores=0.23169518, pem=0.0, similarity=11.003959]
8: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.5831127,
final_local_scores=0.52598554, pem=-2.301417,
similarity=32.301636]
9: 683459 List_A_cricket
[pem_pos=0, final_scores=0.2887866,

final_local_scores=0.2356767, pem=-0.07478003,
 similarity=11.529442]
 10: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.34031814,
 final_local_scores=0.30002275, pem=0.0, similarity=13.042991]
 11: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.26377118,
 final_local_scores=0.20835117, pem=-0.07478003,
 similarity=10.713993]
 12: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.14310132,
 final_local_scores=0.07653767, pem=-0.07478003,
 similarity=6.780415]
 13: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.38506466,
 final_local_scores=0.34890148, pem=0.0, similarity=14.501629]
 14: 3702129 Berkshire_County_Cricket_Club
 [pem_pos=2, final_scores=0.6426694,
 final_local_scores=0.5910421, pem=-2.301417,
 similarity=34.243053]

Sample: b'Roman_Catholic_Diocese_of_Superior'

The 1 Diocese of Superior -LRB- -RRB- encompasses the city
 of [1 1 Superior 1] and the counties of [1 2 Ashland 1] , 3 Ba
 rron , [2 2 Bayfield 2] , 4 Burnett , 5 Douglas , 6 Iron , 7 L
 incoln , 8 Oneida , 9 Price , 10 Polk , 11 Rusk , 12 Sawyer ,
 [2 13 St. Croix 2] , 14 Taylor , 15 Vilas , and 16 Washburn
 in northern [3 3 Wisconsin 3] , with an area of . Its 17 epis
 copal see is Superior , and the is its mother church . The
 diocese was established on May 3 , 1905 by [4 4 Pope Pius X 4]
 . It was created from the northern part of the and the
 northwestern part of the , with [5 5 Augustine Francis Schinn
 er 5] as the first bishop . It is a 18 suffragan diocese of
 the 19 Archdiocese of Milwaukee

1: 288141 Superior,_Wisconsin
 2: 90992 Bayfield_County,_Wisconsin
 3: 33127 Wisconsin
 4: 59205 Pope_Pius_X
 5: 5248051 Augustine_Francis_Schinner

1: 50549 Diocese
 [winner: 3590324 Roman_Catholic_Diocese_of_Superior,
 pem_pos=0, final_scores=-0.12308264,
 final_local_scores=-0.10979098, pem=0.0,
 similarity=0.81333095]
 RECALL MISS: 50549 Diocese]
 2: 90994 Ashland_County,_Wisconsin
 [winner: 115034 Ashland,_Kentucky, pem_pos=0,
 final_scores=0.16250308, final_local_scores=0.16578329,
 pem=-2.0963864, similarity=20.437492]
 gt: 90994 Ashland_County,_Wisconsin, pem_pos=15,
 final_scores=-0.29978472, final_local_scores=-0.35462272,
 pem=-3.2815113, similarity=11.352432]

3: 90993 Barron_County,_Wisconsin
[winner: 90993 Barron_County,_Wisconsin, pem_pos=1,
final_scores=-0.07075234, final_local_scores=-0.08285293,
pem=-1.6588719, similarity=10.638419|
gt: 90993 Barron_County,_Wisconsin, pem_pos=1,
final_scores=-0.07075234, final_local_scores=-0.08285293,
pem=-1.6588719, similarity=10.638419]
4: 90989 Burnett_County,_Wisconsin
[winner: 90989 Burnett_County,_Wisconsin, pem_pos=1,
final_scores=-0.3627947, final_local_scores=-0.4071181,
pem=-2.4509249, similarity=5.269007|
gt: 90989 Burnett_County,_Wisconsin, pem_pos=1,
final_scores=-0.3627947, final_local_scores=-0.4071181,
pem=-2.4509249, similarity=5.269007]
5: 90980 Douglas_County,_Wisconsin
[winner: 95728 Douglas_County,_Kansas, pem_pos=2,
final_scores=-0.2586282, final_local_scores=-0.3086561,
pem=-3.0914392, similarity=11.690526|
gt: 90980 Douglas_County,_Wisconsin, pem_pos=17,
final_scores=-0.5124141, final_local_scores=-0.5875629,
pem=-4.2144237, similarity=9.474354]
6: 90968 Iron_County,_Wisconsin
[winner: 14734 Iron, pem_pos=0, final_scores=-0.0497946,
final_local_scores=0.048714384, pem=-0.2294915,
similarity=6.791458|
gt: 90968 Iron_County,_Wisconsin, pem_pos=4,
final_scores=-0.50691926, final_local_scores=-0.55363035,
pem=-3.4454749, similarity=6.305307]
7: 90959 Lincoln_County,_Wisconsin
[winner: 17653 Lincoln,_Nebraska, pem_pos=3,
final_scores=-0.054986075, final_local_scores=-0.10353428,
pem=-2.526223, similarity=14.738036|
RECALL MISS: 90959 Lincoln_County,_Wisconsin]
8: 90949 Oneida_County,_Wisconsin
[winner: 55294 Oneida_County,_New_York, pem_pos=1,
final_scores=-0.03879656, final_local_scores=-0.0741536,
pem=-1.7837833, similarity=11.577312|
gt: 90949 Oneida_County,_Wisconsin, pem_pos=17,
final_scores=-0.4148013, final_local_scores=-0.47030863,
pem=-3.197256, similarity=7.4419403]
9: 90942 Price_County,_Wisconsin
[winner: 136961 Price,_Utah, pem_pos=1,
final_scores=-0.017443195, final_local_scores=-0.034778982,
pem=-2.3930426, similarity=16.065577|
gt: 90942 Price_County,_Wisconsin, pem_pos=3,
final_scores=-0.3057896, final_local_scores=-0.3535954,
pem=-3.3273137, similarity=11.632169]
10: 90944 Polk_County,_Wisconsin
[winner: 74644 Polk_County,_Florida, pem_pos=0,
final_scores=-0.035005644, final_local_scores=-0.060984164,
pem=-2.133336, similarity=13.871237|
gt: 90944 Polk_County,_Wisconsin, pem_pos=6,

final_scores=-0.24871968, final_local_scores=-0.28738225,
pem=-2.8277566, similarity=10.891434]
11: 90937 Rusk_County,_Wisconsin
[winner: 135672 Rusk,_Texas, pem_pos=1,
final_scores=-0.17075041, final_local_scores=-0.18893115,
pem=-1.5206742, similarity=6.7212934|
gt: 90937 Rusk_County,_Wisconsin, pem_pos=2,
final_scores=-0.17144907, final_local_scores=-0.19154054,
pem=-1.7509375, similarity=7.8956313]
12: 90935 Sawyer_County,_Wisconsin
[winner: 2231014 James_"Sawyer"_Ford, pem_pos=0,
final_scores=-0.28051662, final_local_scores=-0.15082549,
pem=-0.9384561, similarity=4.69225|
gt: 90935 Sawyer_County,_Wisconsin, pem_pos=1,
final_scores=-0.37915832, final_local_scores=-0.41843274,
pem=-2.6603806, similarity=6.0704083]
13: 90932 St._Croix_County,_Wisconsin
[winner: 229431 Saint_Croix,_U.S._Virgin_Islands,
pem_pos=0, final_scores=0.22862966,
final_local_scores=0.2572598, pem=-0.2547834,
similarity=13.1524105|
gt: 90932 St._Croix_County,_Wisconsin, pem_pos=1,
final_scores=-0.19219035, final_local_scores=-0.21974348,
pem=-1.565379, similarity=6.0449038]
14: 90930 Taylor_County,_Wisconsin
[winner: 119280 Taylor,_Michigan, pem_pos=0,
final_scores=-0.36572158, final_local_scores=-0.40021485,
pem=-2.8038657, similarity=7.3943605|
gt: 90930 Taylor_County,_Wisconsin, pem_pos=16,
final_scores=-0.6656055, final_local_scores=-0.7315205,
pem=-4.197761, similarity=5.0877614]
15: 90926 Vilas_County,_Wisconsin
[winner: 755014 William_Freeman_Vilas, pem_pos=2,
final_scores=-0.18369623, final_local_scores=-0.15870865,
pem=-1.2394258, similarity=6.0937204|
gt: 90926 Vilas_County,_Wisconsin, pem_pos=3,
final_scores=-0.2670905, final_local_scores=-0.30573338,
pem=-1.4348415, similarity=2.7689102]
16: 90923 Washburn_County,_Wisconsin
[winner: 115923 Washburn,_Maine, pem_pos=0,
final_scores=-0.22817615, final_local_scores=-0.24724361,
pem=-2.048049, similarity=7.849079|
gt: 90923 Washburn_County,_Wisconsin, pem_pos=8,
final_scores=-0.3487947, final_local_scores=-0.39075032,
pem=-2.6569374, similarity=6.8777823]
17: 199657 Episcopal_see
[winner: 19280748 Episcopal_Church_(United_States),
pem_pos=0, final_scores=-1.1132702,
final_local_scores=-1.2557523, pem=-0.049595453,
similarity=-33.11473|
gt: 199657 Episcopal_see, pem_pos=7,
final_scores=-1.5940902, final_local_scores=-1.7104685,

pem=-2.387467, similarity=-33.970665]
[winner: 11715007 Roman_Catholic_Diocese_of_Pavia,
pem_pos=2, final_scores=-1.1476629,
final_local_scores=-1.1779385, pem=-1.7917595,
similarity=-21.318455]
gt: 199657 Episcopal_see, pem_pos=0,
final_scores=-1.2385466, final_local_scores=-1.3220911,
pem=-0.6161582, similarity=-32.013363]
18: 6165764 Suffragan_diocese
[winner: 10065274 Roman_Catholic_Diocese_of_Victoria,
pem_pos=1, final_scores=-0.82708246,
final_local_scores=-0.8395757, pem=-0.6931472,
similarity=-17.19546]
RECALL MISS: 6165764 Suffragan_diocese]
[winner: 296341 Suffragan_bishop, pem_pos=0,
final_scores=-1.2100307, final_local_scores=-1.2740147,
pem=0.0, similarity=-33.92943]
gt: 6165764 Suffragan_diocese, pem_pos=1,
final_scores=-1.7565116, final_local_scores=-1.8114924,
pem=-4.117875, similarity=-27.57519]
19: 304718 Roman_Catholic_Archdiocese_of_Milwaukee
[winner: 304718 Roman_Catholic_Archdiocese_of_Milwaukee,
pem_pos=0, final_scores=-0.048052683,
final_local_scores=-0.037484504, pem=0.0,
similarity=2.9711006]
gt: 304718 Roman_Catholic_Archdiocese_of_Milwaukee,
pem_pos=0, final_scores=-0.048052683,
final_local_scores=-0.037484504, pem=0.0,
similarity=2.9711006]
1: 288141 Superior,_Wisconsin
[pem_pos=1, final_scores=0.19606964,
final_local_scores=0.15757158, pem=-1.5994453,
similarity=17.489996]
2: 90992 Bayfield_County,_Wisconsin
[pem_pos=1, final_scores=0.06512214,
final_local_scores=0.06857689, pem=-1.2023653,
similarity=12.674833]
3: 33127 Wisconsin
[pem_pos=0, final_scores=0.39649814,
final_local_scores=0.3716194, pem=-0.21268336,
similarity=16.336182]
4: 59205 Pope_Pius_X
[pem_pos=0, final_scores=0.12466203,
final_local_scores=0.19593844, pem=0.0, similarity=9.936907]
5: 5248051 Augustine_Francis_Schinner
[pem_pos=0, final_scores=0.07351442,
final_local_scores=0.19716087, pem=0.0, similarity=9.973387]
1: 115034 Ashland,_Kentucky
[pem_pos=0, final_scores=0.16250308,
final_local_scores=0.16578329, pem=-2.0963864,
similarity=20.437492]
2: 229431 Saint_Croix,_U.S._Virgin_Islands

[pem_pos=0, final_scores=0.22862966,
final_local_scores=0.2572598, pem=-0.2547834,
similarity=13.1524105]

Sample: b'Stewart_Hutton'

[1 Stewart Hutton 1] -LRB- born 13 November 1969 -RRB- is a former [2 1 English 2] 2 cricket er . [3 Hutton 3] was a left-handed who bowled right-arm 3 medium pace . He was born at [1 1 Stockton-on-Tees 1] , [2 2 County Durham 2] . [4 Hutton 4] made his debut for [3 3 Durham 3] in the 1991 , playing in three matches in the competition , against [4 4 Northumberland 4] , [5 5 Norfolk 5] and [6 6 Suffolk 6] . He recorded two in these matches . He was one of the few players retained by [5 Durham 5] who had played 4 minor counties cricket following their elevation to 5 first-class status for the 6 1992 season . He made his first-class debut midway through that season against [7 7 Essex 7] in the [8 8 County Championship 8] . He made eight first-class appearances in that season , scoring 406 runs at an 7 average of 27.06 . He made two , with his highest score of 78 coming in his second match against [9 9 Sussex 9] at [6 8 Horsham 6] . He also made his [10 10 List A 10] debut in that season against [11 11 Warwickshire 11] at [12 12 Edgbaston 12] in the [7 9 Sunday League 7] . He made two further appearances in that format in 1992 , against [13 13 Glamorgan 13] and [14 14 Worcestershire 14]

1: 327922 Stockton-on-Tees

2: 400297 County_Durham

3: 1176371 Durham_County_Cricket_Club

4: 3701987 Northumberland_County_Cricket_Club

5: 3690942 Norfolk_County_Cricket_Club

6: 3691249 Suffolk_County_Cricket_Club

7: 1622252 Essex_County_Cricket_Club

8: 683599 County_Championship

9: 1763691 Sussex_County_Cricket_Club

10: 683459 List_A_cricket

11: 667292 Warwickshire_County_Cricket_Club

12: 2641030 Edgbaston_Cricket_Ground

13: 1622320 Glamorgan_County_Cricket_Club

14: 1156428 Worcestershire_County_Cricket_Club

1: 18803164 English_people

[winner: 9316 England, pem_pos=1, final_scores=0.500013,
final_local_scores=0.53315336, pem=-1.3527626,
similarity=27.356606]

gt: 18803164 English_people, pem_pos=2,
final_scores=-0.4168328, final_local_scores=-0.45463997,
pem=-2.2084396, similarity=2.5321865]

2: 25675557 Cricket

[winner: 330091 Rhaphidophoridae, pem_pos=10,
final_scores=-1.4854531, final_local_scores=-1.4405856,
pem=-2.56456, similarity=-24.953756]

gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.8161716, final_local_scores=-2.0212922,

```

pem=-0.07532314, similarity=-55.820057 ]
3: 689836 Seam_bowling
  [winner: 689836 Seam_bowling, pem_pos=0,
  final_scores=-0.9387792, final_local_scores=-1.0631764,
  pem=-0.53094965, similarity=-24.750206]
  gt: 689836 Seam_bowling, pem_pos=0,
  final_scores=-0.9387792, final_local_scores=-1.0631764,
  pem=-0.53094965, similarity=-24.750206 ]
4: 770909 Minor_counties_of_English_and_Welsh_cricket
  [winner: 770909
  Minor_counties_of_English_and_Welsh_cricket, pem_pos=1,
  final_scores=-0.65025526, final_local_scores=-0.7907772,
  pem=-0.021804659, similarity=-19.39008]
  gt: 770909 Minor_counties_of_English_and_Welsh_cricket,
  pem_pos=1, final_scores=-0.65025526,
  final_local_scores=-0.7907772, pem=-0.021804659,
  similarity=-19.39008 ]
5: 11040 First-class_cricket
  [winner: 11040 First-class_cricket, pem_pos=0,
  final_scores=-1.217622, final_local_scores=-1.4003084,
  pem=-0.053819276, similarity=-37.405605]
  gt: 11040 First-class_cricket, pem_pos=0,
  final_scores=-1.217622, final_local_scores=-1.4003084,
  pem=-0.053819276, similarity=-37.405605 ]
6: 7635440 1992_English_cricket_season
  [winner: 19155267 44th_Primetime_Emmy_Awards, pem_pos=0,
  final_scores=-0.76306725, final_local_scores=-0.67944694,
  pem=0.0, similarity=-16.186338]
  RECALL MISS: 7635440 1992_English_cricket_season]
  [winner: 1340236 Eastertide, pem_pos=2,
  final_scores=-1.4186114, final_local_scores=-1.3989058,
  pem=-2.3723626, similarity=-24.755146]
  RECALL MISS: 7635440 1992_English_cricket_season]
7: 18933273 Batting_average
  [winner: 875871 Bowling_average, pem_pos=0,
  final_scores=-1.0591086, final_local_scores=-1.1961725,
  pem=0.0, similarity=-31.60646]
  gt: 18933273 Batting_average, pem_pos=1,
  final_scores=-1.3124027, final_local_scores=-1.4591126,
  pem=-0.6931472, similarity=-35.683678 ]
8: 11202135 Horsham_Cricket_Club
  [winner: 7386050 Horsham, pem_pos=0,
  final_scores=0.19657002, final_local_scores=0.20327604,
  pem=-0.3439661, similarity=12.026416]
  gt: 11202135 Horsham_Cricket_Club, pem_pos=10,
  final_scores=-0.20569262, final_local_scores=-0.22862598,
  pem=-2.3674474, similarity=10.141604 ]
9: 30702222 1992_Sunday_League
  [winner: 1177737 Pro40, pem_pos=0,
  final_scores=0.21202247, final_local_scores=0.18868831,
  pem=0.0, similarity=9.720549]
  gt: 30702222 1992_Sunday_League, pem_pos=2,

```

```
final_scores=-0.8105452, final_local_scores=-0.86330837,  
pem=-5.4574556, similarity=8.005362 ]  
1: 327922 Stockton-on-Tees  
    [pem_pos=0, final_scores=0.15282802,  
    final_local_scores=0.16140297, pem=0.0, similarity=8.9063]  
2: 400297 County_Durham  
    [pem_pos=0, final_scores=0.28491324,  
    final_local_scores=0.28860867, pem=0.0, similarity=12.702372]  
3: 1176371 Durham_County_Cricket_Club  
    [pem_pos=2, final_scores=0.5859945,  
    final_local_scores=0.5387047, pem=-2.3952909,  
    similarity=33.1917]  
4: 3701987 Northumberland_County_Cricket_Club  
    [pem_pos=1, final_scores=0.26954663,  
    final_local_scores=0.21974498, pem=-2.745579,  
    similarity=25.578226]  
5: 3690942 Norfolk_County_Cricket_Club  
    [pem_pos=6, final_scores=0.33376253,  
    final_local_scores=0.2830131, pem=-3.4736512,  
    similarity=31.425642]  
6: 3691249 Suffolk_County_Cricket_Club  
    [pem_pos=4, final_scores=0.2662487,  
    final_local_scores=0.21778923, pem=-2.9750729,  
    similarity=26.767887]  
7: 1622252 Essex_County_Cricket_Club  
    [pem_pos=1, final_scores=0.8953367,  
    final_local_scores=0.8748837, pem=-2.3459642,  
    similarity=42.955704]  
8: 683599 County_Championship  
    [pem_pos=0, final_scores=0.2617542,  
    final_local_scores=0.20250931, pem=0.0, similarity=10.132995]  
9: 1763691 Sussex_County_Cricket_Club  
    [pem_pos=1, final_scores=0.5960544,  
    final_local_scores=0.56532276, pem=-1.5566913,  
    similarity=29.425602]  
10: 683459 List_A_cricket  
    [pem_pos=0, final_scores=0.3090312,  
    final_local_scores=0.270068, pem=-0.020267159,  
    similarity=12.259296]  
11: 667292 Warwickshire_County_Cricket_Club  
    [pem_pos=1, final_scores=0.8221573,  
    final_local_scores=0.7932142, pem=-1.1093602,  
    similarity=33.793682]  
12: 2641030 Edgbaston_Cricket_Ground  
    [pem_pos=1, final_scores=0.45276302,  
    final_local_scores=0.43306524, pem=-0.4505704,  
    similarity=19.463516]  
13: 1622320 Glamorgan_County_Cricket_Club  
    [pem_pos=1, final_scores=0.775935,  
    final_local_scores=0.74528474, pem=-0.7903395,  
    similarity=30.628485]  
14: 1156428 Worcestershire_County_Cricket_Club
```

```

    [pem_pos=1, final_scores=0.56140095,
final_local_scores=0.51137245, pem=-0.9164197,
similarity=24.33372]
1: 35974933 Stewart_Hutton
    [pem_pos=0, final_scores=0.09938364,
final_local_scores=0.082562506, pem=0.0, similarity=6.5535426]
2: 9316 England
    [pem_pos=1, final_scores=0.500013,
final_local_scores=0.53315336, pem=-1.3527626,
similarity=27.356606]
3: 35974933 Stewart_Hutton
    [pem_pos=0, final_scores=0.30912226,
final_local_scores=0.31167006, pem=0.0, similarity=13.39057]
4: 35974933 Stewart_Hutton
    [pem_pos=0, final_scores=0.33086073,
final_local_scores=0.33541605, pem=0.0, similarity=14.099196]
5: 1176371 Durham_County_Cricket_Club
    [pem_pos=2, final_scores=0.8025119,
final_local_scores=0.7752171, pem=-2.3952909,
similarity=40.249702]
6: 7386050 Horsham
    [pem_pos=0, final_scores=0.19657002,
final_local_scores=0.20327604, pem=-0.3439661,
similarity=12.026416]
7: 1177737 Pro40
    [pem_pos=0, final_scores=0.21202247,
final_local_scores=0.18868831, pem=0.0, similarity=9.720549]

```

Sample: b'BLU-109_bomb'

The BLU-109 / B is a hardened penetration bomb used by the [1 1 United States Air Force 1] -LRB- BLU is an acronym for Bomb Live Unit -RRB- . As other '1 bunker buster s', it is intended to smash through concrete shelters and other hardened structures before exploding . In addition to the [1 US 1] , it is part of the armament of the air force of [2 2 Denmark 2] , [3 3 Egypt 3] , [4 4 France 4] , [5 5 Germany 5] , [6 6 Greece 6] , [7 7 Italy 7] , [8 8 Israel 8] , 9 the [9 Netherlands 9] , [10 10 Saudi Arabia 10] and [11 11 United Arab Emirates 11] . The BLU-109 / B has a steel casing about 1 inch -LRB- 25.4 mm -RRB- thick , filled with 530 lb -LRB- 240 kg -RRB- of 12 Tritonal . It has a delayed-action tail - 13 fuze . The BLU-109 entered service in 14 1985 . It is also used as the warhead of some marks of the 15 GBU-15 electro-optically guided bomb , the 16 GBU-27 Paveway III 17 laser-guided bomb , and the 18 AGM-130 rocket-boosted weapon . This weapon supposedly has the same penetration as the 19 Small Diameter Bomb

1: 32090 United_States_Air_Force

1: 167751 Bunker_buster

```

    [winner: 167751 Bunker_buster, pem_pos=1,
final_scores=-0.9973754, final_local_scores=-0.9877304,
pem=-0.10840933, similarity=-24.796585]

```

gt: 167751 Bunker_buster, pem_pos=1,
final_scores=-0.9973754, final_local_scores=-0.9877304,
pem=-0.10840933, similarity=-24.796585]
2: 548819 Royal_Danish_Air_Force
[winner: 76972 Denmark, pem_pos=0,
final_scores=0.51516265, final_local_scores=0.5040229,
pem=-0.26762357, similarity=20.58614]
RECALL MISS: 548819 Royal_Danish_Air_Force]
3: 377726 Egyptian_Air_Force
[winner: 8087628 Egypt, pem_pos=0, final_scores=0.586341,
final_local_scores=0.5690939, pem=-0.27480945,
similarity=22.567066]
RECALL MISS: 377726 Egyptian_Air_Force]
4: 451287 French_Air_Force
[winner: 5843419 France, pem_pos=0,
final_scores=0.7203796, final_local_scores=0.71198696,
pem=-0.16240834, similarity=26.220024]
RECALL MISS: 451287 French_Air_Force]
5: 17885 Luftwaffe
[winner: 11867 Germany, pem_pos=0,
final_scores=0.7715327, final_local_scores=0.77067536,
pem=-0.15529114, similarity=27.932701]
RECALL MISS: 17885 Luftwaffe]
6: 363254 Hellenic_Air_Force
[winner: 12108 Greece, pem_pos=0, final_scores=0.8817509,
final_local_scores=0.8971649, pem=-0.27607638,
similarity=32.36425]
RECALL MISS: 363254 Hellenic_Air_Force]
7: 244318 Italian_Air_Force
[winner: 14532 Italy, pem_pos=0, final_scores=0.7224595,
final_local_scores=0.73083687, pem=-0.20612453,
similarity=27.02028]
RECALL MISS: 244318 Italian_Air_Force]
8: 177619 Israeli_Air_Force
[winner: 9282173 Israel, pem_pos=0,
final_scores=0.5504661, final_local_scores=0.53835434,
pem=-0.21279564, similarity=21.312494]
RECALL MISS: 177619 Israeli_Air_Force]
9: 652937 Royal_Netherlands_Air_Force
[winner: 21148 Netherlands, pem_pos=0,
final_scores=0.48098868, final_local_scores=0.46411878,
pem=-0.24940795, similarity=19.296263]
RECALL MISS: 652937 Royal_Netherlands_Air_Force]
10: 570367 Royal_Saudi_Air_Force
[winner: 349303 Saudi_Arabia, pem_pos=0,
final_scores=0.5071222, final_local_scores=0.48881084,
pem=-0.25454172, similarity=20.06104]
RECALL MISS: 570367 Royal_Saudi_Air_Force]
11: 10815770 United_Arab_Emirates_Air_Force
[winner: 69328 United_Arab_Emirates, pem_pos=0,
final_scores=0.46998233, final_local_scores=0.45318323,
pem=-0.24986836, similarity=18.972427]

RECALL MISS: 10815770 United_Arab_Emirates_Air_Force]

12: 601254 Tritonal
[winner: 601254 Tritonal, pem_pos=0,
final_scores=-0.3103267, final_local_scores=-0.14442755,
pem=0.0, similarity=-0.22029325]
gt: 601254 Tritonal, pem_pos=0, final_scores=-0.3103267,
final_local_scores=-0.14442755, pem=0.0,
similarity=-0.22029325]

13: 19881804 Fuze
[winner: 683343 Fuse_(explosives), pem_pos=0,
final_scores=-0.8705104, final_local_scores=-0.85324985,
pem=0.0, similarity=-21.372965]
gt: 19881804 Fuze, pem_pos=2, final_scores=-1.150374,
final_local_scores=-1.1431258, pem=-1.6775757,
similarity=-20.900524]

14: 269018 1985_in_aviation
[winner: 34846 1985, pem_pos=0, final_scores=-0.7914199,
final_local_scores=-0.702266, pem=-0.6656445,
similarity=-13.247426]
RECALL MISS: 269018 1985_in_aviation]

15: 201675 GBU-15
[winner: 201675 GBU-15, pem_pos=0,
final_scores=-0.114794925, final_local_scores=0.0058615506,
pem=0.0, similarity=4.264633]
gt: 201675 GBU-15, pem_pos=0, final_scores=-0.114794925,
final_local_scores=0.0058615506, pem=0.0,
similarity=4.264633]

16: 201670 GBU-27_Paveway_III
[winner: 201670 GBU-27_Paveway_III, pem_pos=0,
final_scores=-0.13270804, final_local_scores=-0.019936278,
pem=0.0, similarity=3.4947746]
gt: 201670 GBU-27_Paveway_III, pem_pos=0,
final_scores=-0.13270804, final_local_scores=-0.019936278,
pem=0.0, similarity=3.4947746]

17: 306899 Laser-guided_bomb
[winner: 1451697 Laser_guidance, pem_pos=0,
final_scores=-0.999987, final_local_scores=-0.97898436,
pem=-0.30132484, similarity=-23.486483]
gt: 306899 Laser-guided_bomb, pem_pos=1,
final_scores=-1.194128, final_local_scores=-1.2165937,
pem=-0.33484757, similarity=-30.394917]
[winner: 47628 Bomb, pem_pos=0, final_scores=-1.0210141,
final_local_scores=-1.0876365, pem=-0.20826642,
similarity=-27.234941]
RECALL MISS: 306899 Laser-guided_bomb]

18: 2481646 AGM-130
[winner: 2481646 AGM-130, pem_pos=0,
final_scores=-0.14293587, final_local_scores=-0.044234164,
pem=0.0, similarity=2.7696772]
gt: 2481646 AGM-130, pem_pos=0,
final_scores=-0.14293587, final_local_scores=-0.044234164,
pem=0.0, similarity=2.7696772]

19: 30874902 Small_Diameter_Bomb
[winner: 30874902 Small_Diameter_Bomb, pem_pos=0,
final_scores=-0.16780654, final_local_scores=-0.071169585,
pem=0.0, similarity=1.9658706]
gt: 30874902 Small_Diameter_Bomb, pem_pos=0,
final_scores=-0.16780654, final_local_scores=-0.071169585,
pem=0.0, similarity=1.9658706]
1: 32090 United_States_Air_Force
[pem_pos=0, final_scores=0.2332307,
final_local_scores=0.255947, pem=0.0, similarity=11.727682]
1: 3434750 United_States
[pem_pos=0, final_scores=0.14752783,
final_local_scores=0.08729319, pem=-0.48755783,
similarity=9.346131]
2: 76972 Denmark
[pem_pos=0, final_scores=0.51516265,
final_local_scores=0.5040229, pem=-0.26762357,
similarity=20.58614]
3: 8087628 Egypt
[pem_pos=0, final_scores=0.586341,
final_local_scores=0.5690939, pem=-0.27480945,
similarity=22.567066]
4: 5843419 France
[pem_pos=0, final_scores=0.7203796,
final_local_scores=0.71198696, pem=-0.16240834,
similarity=26.220024]
5: 11867 Germany
[pem_pos=0, final_scores=0.7715327,
final_local_scores=0.77067536, pem=-0.15529114,
similarity=27.932701]
6: 12108 Greece
[pem_pos=0, final_scores=0.8817509,
final_local_scores=0.8971649, pem=-0.27607638,
similarity=32.36425]
7: 14532 Italy
[pem_pos=0, final_scores=0.7224595,
final_local_scores=0.73083687, pem=-0.20612453,
similarity=27.02028]
8: 9282173 Israel
[pem_pos=0, final_scores=0.5504661,
final_local_scores=0.53835434, pem=-0.21279564,
similarity=21.312494]
9: 21148 Netherlands
[pem_pos=0, final_scores=0.48098868,
final_local_scores=0.46411878, pem=-0.24940795,
similarity=19.296263]
10: 349303 Saudi_Arabia
[pem_pos=0, final_scores=0.5071222,
final_local_scores=0.48881084, pem=-0.25454172,
similarity=20.06104]
11: 69328 United_Arab_Emirates
[pem_pos=0, final_scores=0.46998233,

final_local_scores=0.45318323, pem=-0.24986836,
similarity=18.972427]

Sample: b'Rugby_union_in_Asia'

1 Rugby union is a growing sport in [1 1 Asia 1] , governed by the [2 2 Asian Rugby Football Union 2] -LRB- ARFU -RRB- . there are 28 [1 ARFU 1] member unions , of whom 15 are full members of the [3 3 International Rugby Board 3] -LRB- IRB -RRB- , and six further associate members of the IRB in [2 Asia 2] . The flagship tournament for promoting the sport in [3 Asia 3] is the [4 4 Asian Five Nations 4] , which launched in 2 2008 , and which most recently in 3 2011 saw the national teams of [5 5 Hong Kong 5] -LRB- 4 national team -RRB- , [6 6 Japan 6] -LRB- 5 national team -RRB- , [7 7 Kazakhstan 7] -LRB- 6 national team -RRB- , [8 8 Sri Lanka 8] -LRB- 7 national team -RRB- and the [4 8 UAE 4] , compete in the main tournament . The sport in [5 Asia 5] is currently dominated by [6 Japan 6] , the only [7 Asian 7] nation to be represented on the IRB 's executive council . The [8 Japanese 8] national team are so far the only [9 Asian 9] team to have competed in the [9 9 Rugby World Cup 9] , having appeared in all seven tournaments so far . Japan have been announced as the hosts of the 9 2019 [10 Rugby World Cup 10] , and will be the first [11 Asian 11] hosts of the tournament . Under the [12 Japanese 12] bid , [13 Hong Kong 13] and [10 10 Singapore 10] will also host matches . Sixteen [14 Asian 14] teams are full members of the [11 11 International Rugby Board 11] , though one -LRB- Guam -RRB- is arguably in [12 12 Oceania 12] , whilst the [15 10 Arabian Gulf 15] team -LSB- sic -RSB- represents five countries ; [13 13 Bahrain 13] , [14 14 Kuwait 14] , [15 15 Oman 15] , [16 16 Saudi Arabia 16] and the [17 17 United Arab Emirates 17] . The full list is : There are also six [16 Asian 16] associate members of the IRB : Like [17 Georgia 17] and [18 Russia 18] , and as in , [19 Armenia 19] , [20 Azerbaijan 20] and [21 Israel 21] play in [22 Europe 22] , though [23 Kazakhstan 23] does not . == 1991 World Cup == The 1991 World Cup was the first to operate a qualification system , the 11 1987 WC having been invitation-only . In the qualifying stage , [24 Asia 24] was combined with [18 18 Oceania 18] and given two berths . Only four countries took part ; [25 Japan 25] , [26 South Korea 26] , [19 19 Tonga 19] and [20 20 Western Samoa 20] -LRB- 12 Australia , [27 13 Fiji 27] and [21 21 New Zealand 21] were automatic qualifiers -RRB- . Played as a four-team round-robin , [28 Western Samoa 28] and Japan finished first and second respectively courtesy of Japan 's 28-16 win over [29 Tonga 29] , with the Japanese becoming the first Asian team to appear at the World Cup . Although they lost their first two games and did not qualify for the quarter-finals , Japan secured third spot in with a 52-8 win over 14 Zimbabwe . For the 15 1995 WC , [30 Asia 30] held its own tournament , with the winner qualifying for the finals . This time eight countries entered -LRB- [31 16 Hong K

ong 31] , [32 Japan 32] , [33 17 Malaysia 33] , [34 18 Singapore 34] , [35 South Korea 35] , [36 19 Sri Lanka 36] , [37 20 Taiwan 37] and [38 21 Thailand 38] -RRB- , and were split into two groups , with the winners of each group playing each other for the single qualification berth . The final play-off saw Japan beat South Korea 26-11 . The group stages showed the inequalities in the [39 Asian 39] game ; Japan won 103-9 against Malaysia , whilst Hong Kong thrashed Singapore 164-13 . The [40 World Cup 40] itself did little to showcase [41 Asian 41] rugby , as Japan lost every game , including a humiliating 145-17 defeat against the [42 All Blacks 42] . For the 22 1999 WC , [43 Asia 43] was given an extra qualification spot , albeit a repechage one . The same eight countries entered the qualification stages , but to avoid the scorelines witnessed in the previous competition , it was held on a three-stage basis ; Singapore , [44 Sri Lanka 44] and [45 Thailand 45] competing in the first round-robin -LRB- with [46 Sri Lanka 46] going through -RRB- ; [47 Malaysia 47] , [48 Sri Lanka 48] and Taiwan in the second round -LRB- with [49 Taiwan 49] progressing -RRB- ; and then [50 Hong Kong 50] , [51 Japan 51] , [52 South Korea 52] and Taiwan in the third round . Japan again emerged victorious , with a 134-6 win over Taiwan going some way to regain the pride lost in the finals in [53 South Africa 53] four years earlier . South Korea took part in the repechage , beating [54 23 Holland 54] 108-45 on aggregate in the first round , before losing 140-41 over two legs to [55 Tonga 55] . [56 Japan 56] once again lost all their games in the World Cup proper . [57 Asia 57] retained its single berth plus repechage spot -LRB- vs. [58 Oceania 58] -RRB- in the 24 2003 WC qualifiers . This time , eleven teams took part , with the [22 22 Arabian Gulf 22] , [59 25 China 59] and [60 26 Kazakhstan 60] joining the fray . Before [61 Japan 61] and [62 South Korea 62] entered the competition , the other nine teams were split into three round-robin pools , with the winners -LRB- [63 China 63] , [64 Hong Kong 64] and [65 Taiwan 65] -RRB- then playing another round-robin with the top team -LRB- Taiwan -RRB- going on to the final pool with Japan and South Korea , which would decide the automatic and repechage qualifying spots . Japan again emerged victorious , beating their previous score against Taiwan in a 155-3 win at home -LRB- their largest ever win -RRB- , followed by a 120-3 scoreline in the away match . [66 South Korea 66] again finished second , also thrashing [67 Taiwan 67] 119-7 , and repeated their previous repechage performance , losing 194-0 over two legs to the [68 Tongans 68] . Thirteen teams competed in the [69 Asian 69] qualifying tournament for the 27 2007 WC , [70 28 Guam 70] and [71 29 India 71] joining the competition for the first time . Another new format was tried ; this one more complex than previously ; The first round , played in 2005 , consisted of four `` divisions ' ' ; 1 , 2 , 3a and 3b . In the second round , the top two of Division One -LRB- 2005 -

RRB- and the winner of Division Two -LRB- 2005 -RRB- would form Division One -LRB- 2006 -RRB- , whilst the bottom team of Division One -LRB- 2005 -RRB- , the second -LRB- of three -RRB- in Division Two -LRB- 2005 -RRB- , and the winner of a play-off between the Division 3a -LRB- 2005 -RRB- and Division 3b -LRB- 2005 -RRB- champions would form Division Two -LRB- 2006 -RRB- . In the third round , the top two from Division One -LRB- 2006 -RRB- and the winner of Division Two -LRB- 2006 -RRB- would play a round-robin to decide the final placings . [72 Hong Kong 72] , Japan and [73 Korea 73] formed Division One -LRB- 2005 -RRB- , with [74 Hong Kong 74] finishing bottom and replaced by [75 Arabian Gulf 75] , who had won Division Two -LRB- 2005 -RRB- . Taiwan finished bottom of [76 Division Two 76] , and were replaced by [77 Sri Lanka 77] , winners of the 3a-3b play-off over [78 Kazakhstan 78] . In the second round , [79 Arabian Gulf 79] finished bottom of Division One -LRB- 2006 -RRB- , and were replaced in the final round by [80 Hong Kong 80] , who won Division Two -LRB- 2006 -RRB- . In the final group , Japan again emerged victorious , sealing their place at the [81 World Cup 81] with a comprehensive 54-0 defeat of [82 South Korea 82] , who went on to lose to the Tongans for the third successive time , this time 85-3 in a one-off match in [23 23 Wellington 23]

- 1: 689 Asia
- 2: 13292391 Asian_Rugby_Football_Union
- 3: 266858 International_Rugby_Board
- 4: 13435780 Asian_Five_Nations
- 5: 13404 Hong_Kong
- 6: 15573 Japan
- 7: 16642 Kazakhstan
- 8: 26750 Sri_Lanka
- 9: 25406 Rugby_World_Cup
- 10: 27318 Singapore
- 11: 266858 International_Rugby_Board
- 12: 22621 Oceania
- 13: 18933277 Bahrain
- 14: 7515890 Kuwait
- 15: 22316 Oman
- 16: 349303 Saudi_Arabia
- 17: 69328 United_Arab_Emirates
- 18: 22621 Oceania
- 19: 1074629 Tonga_national_rugby_union_team
- 20: 1048972 Samoa_national_rugby_union_team
- 21: 154957 New_Zealand_national_rugby_union_team
- 22: 7940119 Arabian_Gulf_rugby_union_team
- 23: 33804 Wellington

1: 25405 Rugby_union
 [winner: 181566 Rugby, Warwickshire, pem_pos=0,
 final_scores=-0.43915674, final_local_scores=-0.4233634,
 pem=-1.37056, similarity=-1.090976]
 gt: 25405 Rugby_union, pem_pos=3,

final_scores=-0.50957733, final_local_scores=-0.57165253,
pem=-2.0110612, similarity=-2.0330753]

[winner: 20918077

National_Register_of_Historic_Places_listings_in_Union_County
,_Pennsylvania, pem_pos=6, final_scores=-1.4111687,
final_local_scores=-1.3462718, pem=-4.1654887,
similarity=-13.433145|

RECALL MISS: 25405 Rugby_union]

2: 15832226 2008_Asian_Five_Nations

[winner: 35825 2008, pem_pos=0, final_scores=-0.5582472,
final_local_scores=-0.46313083, pem=-0.29331458,
similarity=-8.135944|

RECALL MISS: 15832226 2008_Asian_Five_Nations]

3: 27879275 2011_Asian_Five_Nations

[winner: 36225 2011, pem_pos=0, final_scores=-1.0374166,
final_local_scores=-0.9750604, pem=-1.4232541,
similarity=-17.268158|

RECALL MISS: 27879275 2011_Asian_Five_Nations]

4: 4921161 Hong_Kong_national_rugby_union_team

[winner: 302395 National_(brand), pem_pos=3,
final_scores=-1.5273668, final_local_scores=-1.5791135,
pem=-2.9526155, similarity=-26.977402|

RECALL MISS: 4921161

Hong_Kong_national_rugby_union_team]

[winner: 14708158 TEAMS_(cable_system), pem_pos=3,
final_scores=-1.7519885, final_local_scores=-1.7869089,
pem=-3.9775972, similarity=-27.604422|

RECALL MISS: 4921161 Hong_Kong_national_rugby_union_team]

5: 759561 Japan_national_rugby_union_team

[winner: 302395 National_(brand), pem_pos=3,
final_scores=-1.5119979, final_local_scores=-1.5623255,
pem=-2.9526155, similarity=-26.476416|

RECALL MISS: 759561 Japan_national_rugby_union_team]

[winner: 14708158 TEAMS_(cable_system), pem_pos=3,
final_scores=-1.8116152, final_local_scores=-1.852042,
pem=-3.9775972, similarity=-29.54812|

RECALL MISS: 759561 Japan_national_rugby_union_team]

6: 3039333 Kazakhstan_national_rugby_union_team

[winner: 302395 National_(brand), pem_pos=3,
final_scores=-1.5382255, final_local_scores=-1.590975,
pem=-2.9526155, similarity=-27.331375|

RECALL MISS: 3039333

Kazakhstan_national_rugby_union_team]

[winner: 14708158 TEAMS_(cable_system), pem_pos=3,
final_scores=-1.7969598, final_local_scores=-1.8360331,
pem=-3.9775972, similarity=-29.070387|

RECALL MISS: 3039333

Kazakhstan_national_rugby_union_team]

7: 8051528 Sri_Lanka_national_rugby_union_team

[winner: 302395 National_(brand), pem_pos=3,
final_scores=-1.5271312, final_local_scores=-1.5788562,
pem=-2.9526155, similarity=-26.969725|

RECALL MISS: 8051528
Sri_Lanka_national_rugby_union_team]
[winner: 14708158 TEAMS_(cable_system), pem_pos=3,
final_scores=-1.751743, final_local_scores=-1.7866406,
pem=-3.9775972, similarity=-27.596413]
RECALL MISS: 8051528 Sri_Lanka_national_rugby_union_team]
8: 31372669 United_Arab_Emirates_national_rugby_union_team
[winner: 69328 United_Arab_Emirates, pem_pos=0,
final_scores=0.31260282, final_local_scores=0.28267363,
pem=-0.1950423, similarity=13.585928]
RECALL MISS: 31372669
United_Arab_Emirates_national_rugby_union_team]
9: 18884664 2019_Rugby_World_Cup
[winner: 25406 Rugby_World_Cup, pem_pos=0,
final_scores=0.3385036, final_local_scores=0.31494564,
pem=-0.078096084, similarity=13.913018]
gt: 18884664 2019_Rugby_World_Cup, pem_pos=13,
final_scores=-1.3225857, final_local_scores=-1.4282508,
pem=-8.480944, similarity=7.588549]
[winner: 51391 2019, pem_pos=0, final_scores=-0.5189719,
final_local_scores=-0.29830596, pem=-0.6972276,
similarity=-1.0206996]
gt: 18884664 2019_Rugby_World_Cup, pem_pos=2,
final_scores=-1.7421881, final_local_scores=-1.8873723,
pem=-1.4049667, similarity=-44.59281]
10: 7940119 Arabian_Gulf_rugby_union_team
[winner: 7940119 Arabian_Gulf_rugby_union_team,
pem_pos=0, final_scores=0.1416562,
final_local_scores=0.16166782, pem=-0.49480718,
similarity=11.6050415]
gt: 7940119 Arabian_Gulf_rugby_union_team, pem_pos=0,
final_scores=0.1416562, final_local_scores=0.16166782,
pem=-0.49480718, similarity=11.6050415]
[winner: 14708158 TEAMS_(cable_system), pem_pos=3,
final_scores=-1.7667594, final_local_scores=-1.8030438,
pem=-3.9775972, similarity=-28.085918]
RECALL MISS: 7940119 Arabian_Gulf_rugby_union_team]
11: 380206 1987_Rugby_World_Cup
[winner: 224637 Ice_Hockey_World_Championships,
pem_pos=1, final_scores=-0.17118791,
final_local_scores=-0.17576952, pem=-1.283957,
similarity=5.8267574]
RECALL MISS: 380206 1987_Rugby_World_Cup]
[winner: 34760 1987, pem_pos=0, final_scores=-0.65649635,
final_local_scores=-0.5039835, pem=-0.72282183,
similarity=-7.019343]
gt: 380206 1987_Rugby_World_Cup, pem_pos=15,
final_scores=-1.765971, final_local_scores=-1.9379199,
pem=-4.6280804, similarity=-28.573458]
12: 620967 Australia_national_rugby_union_team
[winner: 374833 Australia_national_cricket_team,
pem_pos=1, final_scores=0.01101239,

final_local_scores=0.007556975, pem=-3.4989398,
similarity=23.343006|
gt: 620967 Australia_national_rugby_union_team,
pem_pos=3, final_scores=-0.09992157,
final_local_scores=-0.13030308, pem=-4.217898,
similarity=23.138794]
13: 1074616 Fiji_national_rugby_union_team
[winner: 1047965 Fiji_national_football_team, pem_pos=1,
final_scores=0.12834577, final_local_scores=0.15028474,
pem=-2.1074445, similarity=20.03512|
RECALL MISS: 1074616 Fiji_national_rugby_union_team]
14: 2703015 Zimbabwe_national_rugby_union_team
[winner: 759317 Zimbabwe_national_cricket_team,
pem_pos=1, final_scores=-0.025606915,
final_local_scores=-0.023256868, pem=-2.2776897,
similarity=15.782111|
gt: 2703015 Zimbabwe_national_rugby_union_team,
pem_pos=3, final_scores=-0.113537714,
final_local_scores=-0.12755102, pem=-3.235928,
similarity=17.880817]
15: 526056 1995_Rugby_World_Cup
[winner: 224637 Ice_Hockey_World_Championships,
pem_pos=1, final_scores=-0.13214102,
final_local_scores=-0.13311674, pem=-1.283957,
similarity=7.0996017|
RECALL MISS: 526056 1995_Rugby_World_Cup]
[winner: 34658 1995, pem_pos=0, final_scores=-1.4302802,
final_local_scores=-1.4386723, pem=-0.68201965,
similarity=-35.13421|
RECALL MISS: 526056 1995_Rugby_World_Cup]
16: 4921161 Hong_Kong_national_rugby_union_team
[winner: 13404 Hong_Kong, pem_pos=0,
final_scores=0.5108106, final_local_scores=0.4904554,
pem=-0.22402541, similarity=19.944166|
gt: 4921161 Hong_Kong_national_rugby_union_team,
pem_pos=6, final_scores=-0.31914735,
final_local_scores=-0.36626095, pem=-3.9262562,
similarity=14.511347]
17: 8081793 Malaysia_national_rugby_union_team
[winner: 3607937 Malaysia, pem_pos=0,
final_scores=0.18559961, final_local_scores=0.1408191,
pem=-0.20223404, similarity=9.391816|
gt: 8081793 Malaysia_national_rugby_union_team,
pem_pos=4, final_scores=-0.5156415,
final_local_scores=-0.5149145, pem=-3.953056,
similarity=10.220971]
18: 8082630 Singapore_national_rugby_union_team
[winner: 27318 Singapore, pem_pos=0,
final_scores=0.16305877, final_local_scores=0.11011341,
pem=-0.1922004, similarity=8.420933|
gt: 8082630 Singapore_national_rugby_union_team,
pem_pos=7, final_scores=-0.40874988,

final_local_scores=-0.44222096, pem=-4.146729,
similarity=13.443514]
19: 8051528 Sri_Lanka_national_rugby_union_team
[winner: 397246 Sri_Lanka_national_cricket_team,
pem_pos=1, final_scores=0.043576524,
final_local_scores=0.045195162, pem=-2.2667584,
similarity=17.765411]
gt: 8051528 Sri_Lanka_national_rugby_union_team,
pem_pos=4, final_scores=-0.2720156,
final_local_scores=-0.2925453, pem=-3.3168993,
similarity=13.397391]
20: 8082065 Chinese_Taipei_national_rugby_union_team
[winner: 25734 Taiwan, pem_pos=0, final_scores=0.1774408,
final_local_scores=0.14139345, pem=-0.077886656,
similarity=8.732736]
RECALL MISS: 8082065
Chinese_Taipei_national_rugby_union_team]
21: 8082794 Thailand_national_rugby_union_team
[winner: 30128 Thailand, pem_pos=0,
final_scores=0.06740497, final_local_scores=0.006094396,
pem=-0.24029273, similarity=5.578331]
gt: 8082794 Thailand_national_rugby_union_team,
pem_pos=6, final_scores=-0.30002838,
final_local_scores=-0.2756514, pem=-3.7642767,
similarity=16.334448]
22: 266016 1999_Rugby_World_Cup
[winner: 224637 Ice_Hockey_World_Championships,
pem_pos=1, final_scores=-0.18953863,
final_local_scores=-0.1958149, pem=-1.283957,
similarity=5.228564]
RECALL MISS: 266016 1999_Rugby_World_Cup]
[winner: 34553 1999, pem_pos=0, final_scores=-1.4093952,
final_local_scores=-1.487967, pem=-0.6240096,
similarity=-36.92073]
RECALL MISS: 266016 1999_Rugby_World_Cup]
23: 4863079 Netherlands_national_rugby_union_team
[winner: 9647657 Netherlands_national_football_team,
pem_pos=3, final_scores=0.04888986,
final_local_scores=0.062204897, pem=-3.631149,
similarity=25.692783]
RECALL MISS: 4863079
Netherlands_national_rugby_union_team]
24: 25279872 2003_Rugby_World_Cup
[winner: 224637 Ice_Hockey_World_Championships,
pem_pos=1, final_scores=-0.1808707,
final_local_scores=-0.18634649, pem=-1.283957,
similarity=5.51112]
RECALL MISS: 25279872 2003_Rugby_World_Cup]
[winner: 36163 2003, pem_pos=0, final_scores=-0.8149157,
final_local_scores=-0.69105774, pem=-0.50449735,
similarity=-13.789293]
RECALL MISS: 25279872 2003_Rugby_World_Cup]

25: 8081609 China_national_rugby_union_team
[winner: 5405 China, pem_pos=0, final_scores=0.4934718,
final_local_scores=0.47701204, pem=-0.15669827,
similarity=19.176853]
RECALL MISS: 8081609 China_national_rugby_union_team]

26: 3039333 Kazakhstan_national_rugby_union_team
[winner: 16642 Kazakhstan, pem_pos=0,
final_scores=0.294321, final_local_scores=0.27998465,
pem=-0.41320926, similarity=14.69211]
RECALL MISS: 3039333
Kazakhstan_national_rugby_union_team]

27: 375446 2007_Rugby_World_Cup
[winner: 224637 Ice_Hockey_World_Championships,
pem_pos=1, final_scores=-0.3055154,
final_local_scores=-0.3225019, pem=-1.283957,
similarity=1.4479706]
RECALL MISS: 375446 2007_Rugby_World_Cup]
[winner: 36165 2007, pem_pos=0, final_scores=-0.9567223,
final_local_scores=-0.91755736, pem=-0.24805899,
similarity=-21.943045]
RECALL MISS: 375446 2007_Rugby_World_Cup]

28: 8081254 Guam_national_rugby_union_team
[winner: 11974 Guam, pem_pos=0, final_scores=0.148441,
final_local_scores=0.15549734, pem=-0.26827332,
similarity=10.188975]
gt: 8081254 Guam_national_rugby_union_team, pem_pos=2,
final_scores=-0.25927472, final_local_scores=-0.23541775,
pem=-2.774191, similarity=12.150858]

29: 2210947 India_national_rugby_union_team
[winner: 14533 India, pem_pos=0, final_scores=0.28353715,
final_local_scores=0.27099717, pem=-0.13622136,
similarity=12.917602]
gt: 2210947 India_national_rugby_union_team, pem_pos=10,
final_scores=-0.5566886, final_local_scores=-0.60248554,
pem=-4.819714, similarity=12.320697]

1: 689 Asia
[pem_pos=0, final_scores=0.70303327,
final_local_scores=0.70742404, pem=-0.1494309,
similarity=26.013287]

2: 13292391 Asian_Rugby_Football_Union
[pem_pos=0, final_scores=0.05857216,
final_local_scores=0.08222841, pem=0.0, similarity=6.5435724]

3: 266858 International_Rugby_Board
[pem_pos=0, final_scores=0.14370869,
final_local_scores=0.13157564, pem=0.0, similarity=8.016192]

4: 13435780 Asian_Five_Nations
[pem_pos=0, final_scores=0.08085619,
final_local_scores=0.08219975, pem=0.0, similarity=6.5427175]

5: 13404 Hong_Kong
[pem_pos=0, final_scores=0.35554153,
final_local_scores=0.32084745, pem=-0.22402541,
similarity=14.882726]

6: 15573 Japan
[pem_pos=0, final_scores=0.37538487,
final_local_scores=0.3429676, pem=-0.13549899,
similarity=15.061415]
7: 16642 Kazakhstan
[pem_pos=0, final_scores=0.17297237,
final_local_scores=0.14742973, pem=-0.41320926,
similarity=10.736406]
8: 26750 Sri_Lanka
[pem_pos=0, final_scores=0.19762473,
final_local_scores=0.17384979, pem=-0.3027737,
similarity=10.924268]
9: 25406 Rugby_World_Cup
[pem_pos=0, final_scores=0.38352126,
final_local_scores=0.3701117, pem=-0.078096084,
similarity=15.559283]
10: 27318 Singapore
[pem_pos=0, final_scores=0.18156035,
final_local_scores=0.13032359, pem=-0.1922004,
similarity=9.024045]
11: 266858 International_Rugby_Board
[pem_pos=0, final_scores=0.24720563,
final_local_scores=0.2446303, pem=0.0, similarity=11.389969]
12: 22621 Oceania
[pem_pos=0, final_scores=0.50943804,
final_local_scores=0.49197137, pem=-0.26907927,
similarity=20.234415]
13: 18933277 Bahrain
[pem_pos=0, final_scores=0.3418743,
final_local_scores=0.33083457, pem=-0.42385748,
similarity=16.267479]
14: 7515890 Kuwait
[pem_pos=0, final_scores=0.33366543,
final_local_scores=0.32019117, pem=-0.3372397,
similarity=15.478819]
15: 22316 Oman
[pem_pos=0, final_scores=0.27294153,
final_local_scores=0.2545346, pem=-0.38991782,
similarity=13.805966]
16: 349303 Saudi_Arabia
[pem_pos=0, final_scores=0.39439702,
final_local_scores=0.38634676, pem=-0.25454172,
similarity=17.003307]
17: 69328 United_Arab_Emirates
[pem_pos=0, final_scores=0.35312515,
final_local_scores=0.32693812, pem=-0.24986836,
similarity=15.205021]
18: 22621 Oceania
[pem_pos=0, final_scores=0.068280086,
final_local_scores=0.010073349, pem=-0.26907927,
similarity=5.8536167]
19: 1074629 Tonga_national_rugby_union_team

[pem_pos=1, final_scores=0.24032448,
final_local_scores=0.24855095, pem=-2.4202409,
similarity=24.668615]
20: 1048972 Samoa_national_rugby_union_team
[pem_pos=1, final_scores=0.37027776,
final_local_scores=0.39346766, pem=-1.4189762,
similarity=23.548185]
21: 154957 New_Zealand_national_rugby_union_team
[pem_pos=4, final_scores=0.5778312,
final_local_scores=0.59906626, pem=-3.8192244,
similarity=42.73658]
22: 7940119 Arabian_Gulf_rugby_union_team
[pem_pos=0, final_scores=0.24055861,
final_local_scores=0.27265665, pem=-0.49480718,
similarity=14.91717]
23: 33804 Wellington
[pem_pos=0, final_scores=0.45692748,
final_local_scores=0.4859417, pem=-0.9265605,
similarity=23.629963]
1: 5374881 Auckland_Rugby_Football_Union
[pem_pos=1, final_scores=0.07825796,
final_local_scores=0.078884795, pem=-0.6931472,
similarity=10.213235]
2: 689 Asia
[pem_pos=0, final_scores=0.6042177,
final_local_scores=0.5994831, pem=-0.1494309,
similarity=22.792112]
3: 689 Asia
[pem_pos=0, final_scores=0.73516613,
final_local_scores=0.74252427, pem=-0.1494309,
similarity=27.060747]
4: 69328 United_Arab_Emirates
[pem_pos=0, final_scores=0.31260282,
final_local_scores=0.28267363, pem=-0.1950423,
similarity=13.585928]
5: 689 Asia
[pem_pos=0, final_scores=0.60828274,
final_local_scores=0.60392356, pem=-0.1494309,
similarity=22.924625]
6: 15573 Japan
[pem_pos=0, final_scores=0.15167932,
final_local_scores=0.09860331, pem=-0.13549899,
similarity=7.769097]
7: 689 Asia
[pem_pos=1, final_scores=0.3603775,
final_local_scores=0.33312464, pem=-1.9623682,
similarity=24.70248]
8: 15573 Japan
[pem_pos=0, final_scores=0.3205977,
final_local_scores=0.2842037, pem=-0.99076414,
similarity=17.958847]
9: 689 Asia

[pem_pos=1, final_scores=0.057766303,
final_local_scores=0.002567947, pem=-1.9623682,
similarity=14.838009]
10: 25406 Rugby_World_Cup
[pem_pos=0, final_scores=0.3385036,
final_local_scores=0.31494564, pem=-0.078096084,
similarity=13.913018]
11: 689 Asia
[pem_pos=1, final_scores=0.23687397,
final_local_scores=0.19821584, pem=-1.9623682,
similarity=20.676533]
12: 15573 Japan
[pem_pos=0, final_scores=0.71345407,
final_local_scores=0.7133396, pem=-0.99076414,
similarity=30.765118]
13: 13404 Hong_Kong
[pem_pos=0, final_scores=0.4237805,
final_local_scores=0.3953882, pem=-0.22402541,
similarity=17.10717]
14: 689 Asia
[pem_pos=1, final_scores=0.493712,
final_local_scores=0.4787724, pem=-1.9623682,
similarity=29.0489]
15: 7940119 Arabian_Gulf_rugby_union_team
[pem_pos=0, final_scores=0.1416562,
final_local_scores=0.16166782, pem=-0.49480718,
similarity=11.6050415]
16: 689 Asia
[pem_pos=1, final_scores=0.29655278,
final_local_scores=0.26340586, pem=-1.9623682,
similarity=22.621933]
17: 48768 Georgia (country)
[pem_pos=1, final_scores=0.16833924,
final_local_scores=0.15419573, pem=-1.3466994,
similarity=16.014782]
18: 25391 Russia
[pem_pos=0, final_scores=0.5875307,
final_local_scores=0.6157038, pem=-0.20299491,
similarity=23.567461]
19: 10918072 Armenia
[pem_pos=0, final_scores=0.39581728,
final_local_scores=0.4120102, pem=-0.4307794,
similarity=18.727564]
20: 746 Azerbaijan
[pem_pos=0, final_scores=0.2633165,
final_local_scores=0.26341993, pem=-0.33647978,
similarity=13.780519]
21: 9282173 Israel
[pem_pos=0, final_scores=0.3496096,
final_local_scores=0.3588924, pem=-0.21279564,
similarity=15.956994]
22: 9239 Europe

[pem_pos=0, final_scores=0.72823,
final_local_scores=0.7462382, pem=-0.086338006,
similarity=26.828468]
23: 16642 Kazakhstan
[pem_pos=0, final_scores=0.42725623,
final_local_scores=0.42519623, pem=-0.41320926,
similarity=19.025513]
24: 689 Asia
[pem_pos=0, final_scores=0.2752987,
final_local_scores=0.24018902, pem=-0.1494309,
similarity=12.070061]
25: 15573 Japan
[pem_pos=0, final_scores=0.112182006,
final_local_scores=0.0554585, pem=-0.13549899,
similarity=6.48157]
26: 27019 South_Korea
[pem_pos=0, final_scores=0.09880455,
final_local_scores=0.05725287, pem=-0.22449642,
similarity=7.019099]
27: 1047965 Fiji_national_football_team
[pem_pos=1, final_scores=0.12834577,
final_local_scores=0.15028474, pem=-2.1074445,
similarity=20.03512]
28: 1048972 Samoa_national_rugby_union_team
[pem_pos=1, final_scores=0.48436826,
final_local_scores=0.5178491, pem=-1.4189762,
similarity=27.259975]
29: 1074629 Tonga_national_rugby_union_team
[pem_pos=1, final_scores=0.2596798,
final_local_scores=0.26969367, pem=-2.4202409,
similarity=25.299557]
30: 689 Asia
[pem_pos=0, final_scores=0.33074296,
final_local_scores=0.30075344, pem=-0.1494309,
similarity=13.877424]
31: 13404 Hong_Kong
[pem_pos=0, final_scores=0.5108106,
final_local_scores=0.4904554, pem=-0.22402541,
similarity=19.944166]
32: 15573 Japan
[pem_pos=0, final_scores=0.55360556,
final_local_scores=0.53764665, pem=-0.13549899,
similarity=20.871025]
33: 3607937 Malaysia
[pem_pos=0, final_scores=0.18559961,
final_local_scores=0.1408191, pem=-0.20223404,
similarity=9.391816]
34: 27318 Singapore
[pem_pos=0, final_scores=0.16305877,
final_local_scores=0.11011341, pem=-0.1922004,
similarity=8.420933]
35: 27019 South_Korea

[pem_pos=0, final_scores=0.26263583,
final_local_scores=0.23621365, pem=-0.22449642,
similarity=12.359647]
36: 397246 Sri_Lanka_national_cricket_team
[pem_pos=1, final_scores=0.043576524,
final_local_scores=0.045195162, pem=-2.2667584,
similarity=17.765411]
37: 25734 Taiwan
[pem_pos=0, final_scores=0.1774408,
final_local_scores=0.14139345, pem=-0.077886656,
similarity=8.732736]
38: 30128 Thailand
[pem_pos=0, final_scores=0.06740497,
final_local_scores=0.006094396, pem=-0.24029273,
similarity=5.578331]
39: 689 Asia
[pem_pos=1, final_scores=0.38114637,
final_local_scores=0.35581154, pem=-1.9623682,
similarity=25.379503]
40: 7239 Cricket_World_Cup
[pem_pos=4, final_scores=0.03423436,
final_local_scores=0.015099347, pem=-3.1731894,
similarity=21.796604]
41: 689 Asia
[pem_pos=1, final_scores=0.34420758,
final_local_scores=0.31546152, pem=-1.9623682,
similarity=24.175377]
42: 154957 New_Zealand_national_rugby_union_team
[pem_pos=0, final_scores=0.42670172,
final_local_scores=0.43398023, pem=0.0, similarity=17.040548]
43: 689 Asia
[pem_pos=0, final_scores=0.16276927,
final_local_scores=0.11726773, pem=-0.1494309,
similarity=8.401844]
44: 397246 Sri_Lanka_national_cricket_team
[pem_pos=1, final_scores=0.34610826,
final_local_scores=0.37566507, pem=-2.2667584,
similarity=27.627293]
45: 1110063 Thailand_national_football_team
[pem_pos=1, final_scores=0.19596429,
final_local_scores=0.21276605, pem=-2.3524861,
similarity=23.232262]
46: 397246 Sri_Lanka_national_cricket_team
[pem_pos=1, final_scores=0.15390043,
final_local_scores=0.16570729, pem=-2.2667584,
similarity=21.361734]
47: 1205243 Malaysia_national_football_team
[pem_pos=1, final_scores=0.0688823,
final_local_scores=0.06644261, pem=-3.1013165,
similarity=22.937933]
48: 397246 Sri_Lanka_national_cricket_team
[pem_pos=1, final_scores=0.14177166,

final_local_scores=0.15335906, pem=-2.2667584,
similarity=20.993238]
49: 25734 Taiwan
[pem_pos=0, final_scores=0.14979763,
final_local_scores=0.11119747, pem=-0.077886656,
similarity=7.8316274]
50: 13404 Hong_Kong
[pem_pos=0, final_scores=0.2897421,
final_local_scores=0.24897164, pem=-0.22402541,
similarity=12.737809]
51: 15573 Japan
[pem_pos=0, final_scores=0.22133653,
final_local_scores=0.17469323, pem=-0.13549899,
similarity=10.039772]
52: 27019 South_Korea
[pem_pos=0, final_scores=0.075794,
final_local_scores=0.032117322, pem=-0.22449642,
similarity=6.2690043]
53: 17416221 South_Africa
[pem_pos=0, final_scores=0.39048332,
final_local_scores=0.38540667, pem=-0.26251382,
similarity=17.018608]
54: 9647657 Netherlands_national_football_team
[pem_pos=3, final_scores=0.04888986,
final_local_scores=0.062204897, pem=-3.631149,
similarity=25.692783]
55: 30158 Tonga
[pem_pos=0, final_scores=0.3885141,
final_local_scores=0.36877453, pem=-0.4741999,
similarity=17.673452]
56: 15573 Japan
[pem_pos=0, final_scores=0.25876796,
final_local_scores=0.21558139, pem=-0.13549899,
similarity=11.259956]
57: 689 Asia
[pem_pos=0, final_scores=0.43718457,
final_local_scores=0.41702473, pem=-0.1494309,
similarity=17.34719]
58: 22621 Oceania
[pem_pos=0, final_scores=0.26725972,
final_local_scores=0.22742835, pem=-0.26907927,
similarity=12.339923]
59: 5405 China
[pem_pos=0, final_scores=0.4934718,
final_local_scores=0.47701204, pem=-0.15669827,
similarity=19.176853]
60: 16642 Kazakhstan
[pem_pos=0, final_scores=0.294321,
final_local_scores=0.27998465, pem=-0.41320926,
similarity=14.69211]
61: 15573 Japan
[pem_pos=0, final_scores=0.5340505,

final_local_scores=0.5162857, pem=-0.13549899,
similarity=20.233572]
62: 27019 South_Korea
[pem_pos=0, final_scores=0.24807952,
final_local_scores=0.2203131, pem=-0.22449642,
similarity=11.885142]
63: 5405 China
[pem_pos=0, final_scores=0.38292992,
final_local_scores=0.35626185, pem=-0.15669827,
similarity=15.573426]
64: 13404 Hong_Kong
[pem_pos=0, final_scores=0.3859411,
final_local_scores=0.3540544, pem=-0.22402541,
similarity=15.873688]
65: 25734 Taiwan
[pem_pos=0, final_scores=0.08059709,
final_local_scores=0.0356064, pem=-0.077886656,
similarity=5.5758386]
66: 27019 South_Korea
[pem_pos=0, final_scores=0.21385191,
final_local_scores=0.18292466, pem=-0.22449642,
similarity=10.769397]
67: 25734 Taiwan
[pem_pos=0, final_scores=0.038635507,
final_local_scores=-0.011515275, pem=-0.077886656,
similarity=4.169634]
68: 30158 Tonga
[pem_pos=1, final_scores=0.25152576,
final_local_scores=0.21992767, pem=-0.42401847,
similarity=12.958672]
69: 689 Asia
[pem_pos=1, final_scores=0.55730003,
final_local_scores=0.5486862, pem=-1.9623682,
similarity=31.13527]
70: 11974 Guam
[pem_pos=0, final_scores=0.148441,
final_local_scores=0.15549734, pem=-0.26827332,
similarity=10.188975]
71: 14533 India
[pem_pos=0, final_scores=0.28353715,
final_local_scores=0.27099717, pem=-0.13622136,
similarity=12.917602]
72: 13404 Hong_Kong
[pem_pos=0, final_scores=0.15644301,
final_local_scores=0.10338749, pem=-0.22402541,
similarity=8.393287]
73: 1018627 South_Korea_national_football_team
[pem_pos=16, final_scores=0.12898462,
final_local_scores=0.10691512, pem=-3.8205078,
similarity=28.056787]
74: 1197359 Hong_Kong_national_football_team
[pem_pos=1, final_scores=0.10492711,

final_local_scores=0.094763696, pem=-3.4266143,
 similarity=25.55211]
 75: 7940119 Arabian_Gulf_rugby_union_team
 [pem_pos=0, final_scores=0.047873124,
 final_local_scores=0.059376597, pem=-0.49480718,
 similarity=8.552467]
 76: 563379 Football_League_Second_Division
 [pem_pos=0, final_scores=0.116979346,
 final_local_scores=0.15916103, pem=-0.19219752,
 similarity=9.884596]
 77: 397246 Sri_Lanka_national_cricket_team
 [pem_pos=1, final_scores=0.16466428,
 final_local_scores=0.17836577, pem=-2.2667584,
 similarity=21.739489]
 78: 1012044 Kazakhstan_national_football_team
 [pem_pos=1, final_scores=0.38834184,
 final_local_scores=0.4356076, pem=-2.2765565,
 similarity=29.469381]
 79: 7940119 Arabian_Gulf_rugby_union_team
 [pem_pos=0, final_scores=0.12079699,
 final_local_scores=0.13903484, pem=-0.49480718,
 similarity=10.929628]
 80: 1197359 Hong_Kong_national_football_team
 [pem_pos=1, final_scores=0.091731235,
 final_local_scores=0.08034921, pem=-3.4266143,
 similarity=25.121954]
 81: 11370 FIFA_World_Cup
 [pem_pos=0, final_scores=0.078939125,
 final_local_scores=0.03691052, pem=-1.001087,
 similarity=10.635263]
 82: 1018627 South_Korea_national_football_team
 [pem_pos=1, final_scores=0.06277095,
 final_local_scores=0.034586728, pem=-2.844636,
 similarity=20.591423]

Sample: b'Central_Lakes_State_Trail'

The Central Lakes State Trail is a paved recreational **1 rail trail** in central **[1 1 Minnesota 1]** , **[1 USA 1]** , running along a former **[2 2 Burlington Northern Railroad 2]** line . The trail is marked with mileposts every mile , corresponding with the mile markers of the former railroad line . Snowmobile use is allowed on the trail in winter , conditions permitting . The trail begins in **[3 3 Osakis 3]** at the western end of the **[4 4 Lake Wobegon Trail 4]** and runs parallel to **[2 2 Interstate 94 2]** before ending in **[5 5 Fergus Falls 5]** . This trail passes through the towns of **[3 3 Nelson 3]** , **[4 4 Alexandria 4]** , **5 Garfield** , **[5 6 Brandon 5]** , **[6 7 Evanston 6]** , Melby , **8 Ashby** , and **9 Dalton**
1: 19590 Minnesota
2: 305983 Burlington_Northern_Railroad
3: 150603 Osakis_Minnesota
4: 18428004 Lake_Wobegon_Trails

5: 121111 Fergus_Falls,_Minnesota

1: 1315222 Rail_trail

[winner: 8386232 Bicycle_and_motorcycle_geometry,
pem_pos=1, final_scores=-1.1673247,
final_local_scores=-1.1208723, pem=-1.3693422,
similarity=-21.912653]

gt: 1315222 Rail_trail, pem_pos=7,
final_scores=-1.5250851, final_local_scores=-1.6595321,
pem=-2.6347148, similarity=-31.106047]

[winner: 1037401 Picatinny_rail, pem_pos=13,
final_scores=-1.4412751, final_local_scores=-1.3964102,
pem=-3.3271766, similarity=-19.488247]

RECALL MISS: 1315222 Rail_trail]

2: 8711215 Interstate_94_in_Minnesota

[winner: 89876 Interstate_94, pem_pos=0,
final_scores=0.07471697, final_local_scores=0.064785615,
pem=0.0, similarity=6.023045]

gt: 8711215 Interstate_94_in_Minnesota, pem_pos=1,
final_scores=-0.37350023, final_local_scores=-0.3773408,
pem=-1.8324527, similarity=2.7942739]

3: 119964 Nelson,_Minnesota

[winner: 61978 Nelson,_New_Zealand, pem_pos=0,
final_scores=0.12662582, final_local_scores=0.1364164,
pem=-1.7804265, similarity=17.842888]

RECALL MISS: 119964 Nelson,_Minnesota]

4: 119938 Alexandria,_Minnesota

[winner: 18940583 Alexandria,_Virginia, pem_pos=1,
final_scores=0.03246598, final_local_scores=-0.000918746,
pem=-2.409911, similarity=17.167767]

gt: 119938 Alexandria,_Minnesota, pem_pos=4,
final_scores=-0.22021101, final_local_scores=-0.2846888,
pem=-3.3700972, similarity=13.921144]

5: 119950 Garfield,_Minnesota

[winner: 13022 Garfield, pem_pos=0,
final_scores=-0.06528045, final_local_scores=0.0056495667,
pem=-0.91356915, similarity=9.2264385]

RECALL MISS: 119950 Garfield,_Minnesota]

6: 119942 Brandon,_Minnesota

[winner: 248339 Brandon,_Manitoba, pem_pos=0,
final_scores=0.057476655, final_local_scores=0.03955707,
pem=-1.4873068, similarity=13.35838]

gt: 119942 Brandon,_Minnesota, pem_pos=18,
final_scores=-0.8585134, final_local_scores=-0.89467835,
pem=-4.369979, similarity=1.15535]

7: 119946 Evansville,_Minnesota

[winner: 101753 Evanston,_Illinois, pem_pos=0,
final_scores=0.302437, final_local_scores=0.28810763,
pem=-0.33685246, similarity=14.519276]

gt: 119946 Evansville,_Minnesota, pem_pos=9,
final_scores=-0.40560454, final_local_scores=-0.39726654,
pem=-2.3959804, similarity=5.264201]

8: 120107 Ashby,_Minnesota

[winner: 116830 Ashby, _Massachusetts, pem_pos=0,
final_scores=-0.099765405, final_local_scores=-0.070158556,
pem=-1.1168389, similarity=8.069586]
gt: 120107 Ashby, _Minnesota, pem_pos=3,
final_scores=-0.29909813, final_local_scores=-0.31075594,
pem=-2.4684732, similarity=8.240077]
9: 121075 Dalton, _Minnesota
[winner: 110437 Dalton, _Georgia, pem_pos=1,
final_scores=-0.42090046, final_local_scores=-0.4761486,
pem=-2.3826332, similarity=2.8376186]
gt: 121075 Dalton, _Minnesota, pem_pos=10,
final_scores=-0.7585086, final_local_scores=-0.7666169,
pem=-3.3127913, similarity=-0.77218974]
1: 19590 Minnesota
[pem_pos=0, final_scores=0.633549,
final_local_scores=0.6207428, pem=-0.21388707,
similarity=23.777067]
2: 305983 Burlington_Northern_Railroad
[pem_pos=0, final_scores=0.36462665,
final_local_scores=0.43176103, pem=0.0, similarity=16.974323]
3: 150603 Osakis, _Minnesota
[pem_pos=0, final_scores=0.16259201,
final_local_scores=0.19977963, pem=0.0, similarity=10.051536]
4: 18428004 Lake_Wobegon_Trails
[pem_pos=0, final_scores=0.030122682,
final_local_scores=0.13144165, pem=0.0, similarity=8.012194]
5: 121111 Fergus_Falls, _Minnesota
[pem_pos=0, final_scores=0.37880027,
final_local_scores=0.36164588, pem=0.0, similarity=14.8819475]
1: 3434750 United_States
[pem_pos=0, final_scores=0.339889,
final_local_scores=0.33747894, pem=-0.13038418,
similarity=14.869806]
2: 89876 Interstate_94
[pem_pos=0, final_scores=0.07471697,
final_local_scores=0.064785615, pem=0.0, similarity=6.023045]
3: 61978 Nelson, _New_Zealand
[pem_pos=0, final_scores=0.12662582,
final_local_scores=0.1364164, pem=-1.7804265,
similarity=17.842888]
4: 18940583 Alexandria, _Virginia
[pem_pos=1, final_scores=0.03246598,
final_local_scores=-0.000918746, pem=-2.409911,
similarity=17.167767]
5: 248339 Brandon, _Manitoba
[pem_pos=0, final_scores=0.057476655,
final_local_scores=0.03955707, pem=-1.4873068,
similarity=13.35838]
6: 101753 Evanston, _Illinois
[pem_pos=0, final_scores=0.302437,
final_local_scores=0.28810763, pem=-0.33685246,
similarity=14.519276]

Sample: b'Gregg_Brandon'

[1 Gregg Brandon 1] -LRB- born February 29 , 1956 -RRB- is the current 1 offensive coordinator of the [2 2 University of Wyoming Cowboys 2] . [3 Brandon 3] recently coached tight ends in the for the [1 1 Las Vegas Locomotives 1] . [4 Brandon 4] previously served as Offensive Coordinator for the [5 3 Virginia Cavaliers 5] for the duration of the 2009 season and head coach for the [6 4 Bowling Green Falcons 6] , but was fired following the conclusion of the 5 2008 season , after six years at the helm , which included three bowl appearances and shares of two division titles . Brandon had previously been an Assistant Coach and Offensive Coordinator for former Bowling Green head coach [2 2 Urban Meyer 2] , before [7 Meyer 7] left for the [8 6 University of Utah 8] in 7 2003 . [9 Brandon 9] is married to the former Robyn Mitchell and has two sons , Nicholas and Timothy . Brandon is also uncle to [10 Incarnate Word 10] standout , Paul Townsen . [11 Brandon 11] played football at 8 Air Academy High School in , [12 Colorado 12] . He went on to compete at the collegiate level as both a defensive back and wide receiver at [3 3 Colorado Mesa University 3] -LRB- 1974 -RRB- and [4 4 University of Northern Colorado 4] -LRB- 1975 -- 77 -RRB- . He graduated from [13 Northern Colorado 13] in 1978 with a bachelor 's degree in education . [14 Brandon 14] began his coaching career as head football coach at Ellicott -LRB- Colo. -RRB- High School . He spent three seasons there -LRB- 1978 -- 80 -RRB- before joining the college ranks at -LRB- under former 9 Washington State and current 10 UTEP coach [5 5 Mike Price 5] -RRB- , where he would coach the next six seasons -LRB- 1981 -- 86 -RRB- . He coached the tight ends and special teams in his first four seasons , the linebackers in his fifth year and the receivers and tight ends during his final year in [15 Ogden 15] . In 11 1987 was then named receivers coach at 12 Wyoming , where he would work the next four years -LRB- 1987 -- 90 -RRB- . During his tenure with head coach Paul Roach , the [16 Cowboys 16] posted a 35-15 record , including a 16-0 run in [6 6 Western Athletic Conference 6] play in 1987 and 1988 , and played in three bowls -LRB- 13 '87 and 14 '88 Holiday and 15 '90 Copper -RRB- . In 16 1991 , he returned to [17 Utah 17] where he coached the linebackers for one season -LRB- 1991 -RRB- at [18 17 Utah State 18] . The following year , [19 Brandon 19] joined the 18 Northwestern staff as receivers coach . He was also Northwestern 's recruiting coordinator for his last two seasons there -LRB- 1997 -- 98 -RRB- . In his seven seasons at Northwestern , Brandon assembled an extremely talented group of receivers , including 19 D'Wayne Bates , the second all-time leading receiver in [20 Big Ten 20] history . [21 Bates 21] was a semifinalist for the and an [22 All-America 22] candidate . [23 Brandon 23] would return to the West , spending two seasons at [24 20 Colorado 24] under [7 7 Gary Ba

rnett 7] , where he coached the receivers and kickoff return units both years and was the team 's passing-game coordinator in 2000 and recruiting coordinator in 1999 . He joined the Falcons as an assistant head coach and offensive coordinator following the 21 2000 season under head coach [8 8 Urban Meyer 8] . Following [25 Meyer 25] 's departure to 22 Utah in 23 2002 , [26 Gregg Brandon 26] was named the 16th head football coach in Bowling Green 's history . In his first season as Bowling Green 's head coach , [27 Brandon 27] lead the Falcons to an 11-3 record , a national ranking and an appearance in the 24 2003 25 MAC Championship Game , eventually losing to 26 Miami -LRB- OH -RRB- at 27 Doyt Perry Stadium . The Falcons would still be invited to a bowl game , defeating [28 Brandon 28] 's former Northwestern squad , 28-24 at the at [9 9 Ford Field 9] in . The following season , [29 Brandon 29] went 9-3 including taking the Falcons to a second consecutive bowl game , defeating 28 Memphis 52-35 at the 29 2004 GMAC Bowl in [30 30 Mobile 30] , [31 Alabama 31] . The Falcons would finish 6-5 in 31 2005 after losing a hard fought game against arch rival , Toledo in multiple overtimes , a game if the Falcons had won would 've gone to their second MAC Championship and could have gone to a bowl game for the third straight year under [32 Brandon 32] . Brandon would suffer his first and only losing season in , finishing 4-8 including losses to rivals 32 Kent State , Miami -LRB- OH -RRB- , and [33 Toledo 33] , as well as becoming the first team to lose to 33 Temple in 20-games . [34 Brandon 34] 's Falcons would put the prior season behind them as they would post an 8-5 record during their 34 2007 campaign earning a share of the MAC East Division title , but missing out on the opportunity to play in the MAC Championship game due to a tiebreaker . Brandon would lead the team to its third bowl under his tenure , losing to 63-7 in the 35 GMAC Bowl . Coming into the 36 2008 season , the Falcons were picked to be the beast of the [35 MAC East 35] and were projected to reach the MAC Championship Game as well as their second straight bowl game . However , the Falcons had a very disappointing campaign , going 6-6 including a 1-4 record at home . Following the Falcons double-overtime loss in their home finale against 37 Buffalo , [36 Brandon 36] showed irritation with the fans for their poor attendance at the game as it had championship implications . The day after winning their first road contest at arch rival Toledo since 38 1994 , [37 Brandon 37] was fired by athletic director Greg Christopher after six years as Bowling Green 's head coach and posting a 44-30 record . On December 18 , 2008 , [38 Brandon 38] accepted the position of offensive coordinator at the [39 University of Virginia 39] , replacing [10 10 Mike Groh 10] -LRB- son of then current Virginia head coach , [11 11 Al Groh 11]

1: 22031615 Las_Vegas_Locomotives

2: 1229075 Urban_Meyer

3: 4437312 Colorado_Mesa_University
4: 362136 University_of_Northern_Colorado
5: 2698421 Mike_Price
6: 79375 Western_Athletic_Conference
7: 3364537 Gary_Barnett
8: 1229075 Urban_Meyer
9: 752260 Ford_Field
10: 14065885 Mike_Groh
11: 1517670 Al_Groh
1: 908217 Offensive_coordinator
[winner: 5885946 Team_Role_Inventories, pem_pos=0,
final_scores=-1.1943964, final_local_scores=-1.1388245,
pem=-0.7472144, similarity=-25.831612|
RECALL MISS: 908217 Offensive_coordinator]
[winner: 8654568 Telus_Cup_-_Offensive, pem_pos=9,
final_scores=-1.4107145, final_local_scores=-1.3797736,
pem=-3.1763341, similarity=-19.81208|
RECALL MISS: 908217 Offensive_coordinator]
2: 535684 University_of_Wyoming
[winner: 10485435 Wyoming_Cowboys_and_Cowgirls,
pem_pos=0, final_scores=0.25719643,
final_local_scores=0.26277205, pem=0.0, similarity=11.931355|
RECALL MISS: 535684 University_of_Wyoming]
3: 7768993 Virginia_Cavaliers_football
[winner: 3485377 Virginia_Cavaliers, pem_pos=0,
final_scores=0.09771292, final_local_scores=0.06714666,
pem=-0.2621322, similarity=7.519018|
gt: 7768993 Virginia_Cavaliers_football, pem_pos=1,
final_scores=-0.021220222, final_local_scores=-0.0854646,
pem=-1.4949421, similarity=9.669006]
4: 9725713 Bowling_Green_Falcons_football
[winner: 9152125 Bowling_Green_Falcons, pem_pos=0,
final_scores=0.113464996, final_local_scores=0.08752824,
pem=-0.16813694, similarity=7.6160855|
gt: 9725713 Bowling_Green_Falcons_football, pem_pos=1,
final_scores=-0.057007685, final_local_scores=-0.11516556,
pem=-1.211941, similarity=7.243666]
5: 15128839 2008_Bowling_Green_Falcons_football_team
[winner: 35825 2008, pem_pos=0, final_scores=-0.6112023,
final_local_scores=-0.48718226, pem=-0.29331458,
similarity=-8.853686|
RECALL MISS: 15128839
2008_Bowling_Green_Falcons_football_team]
[winner: 690345 Season(sports), pem_pos=0,
final_scores=-1.4764364, final_local_scores=-1.588635,
pem=-1.219945, similarity=-36.68407|
RECALL MISS: 15128839
2008_Bowling_Green_Falcons_football_team]
6: 8044294 Utah_Utes_football
[winner: 32097 University_of_Utah, pem_pos=0,
final_scores=0.34872234, final_local_scores=0.35780862,
pem=0.0, similarity=14.767436|

gt: 8044294 Utah_Utes_football, pem_pos=1,
 final_scores=0.045103267, final_local_scores=-0.01542899,
 pem=-1.564084, similarity=12.135014]
 7: 4305778 2003_NCAA_Division_I-A_football_season
 [winner: 36163 2003, pem_pos=0, final_scores=-0.648923,
 final_local_scores=-0.47352502, pem=-0.50449735,
 similarity=-7.2976832]
 gt: 4305778 2003_NCAA_Division_I-A_football_season,
 pem_pos=16, final_scores=-1.3024249,
 final_local_scores=-1.4653227, pem=-5.171439,
 similarity=-11.515353]
 8: 3669211 Air_Academy_High_School
 [winner: 3669211 Air_Academy_High_School, pem_pos=0,
 final_scores=-0.011117235, final_local_scores=0.03802879,
 pem=0.0, similarity=5.224568]
 gt: 3669211 Air_Academy_High_School, pem_pos=0,
 final_scores=-0.011117235, final_local_scores=0.03802879,
 pem=0.0, similarity=5.224568]
 9: 7773465 Washington_State_Cougars_football
 [winner: 7773465 Washington_State_Cougars_football,
 pem_pos=1, final_scores=-0.0037876219,
 final_local_scores=-0.057314888, pem=-1.1912286,
 similarity=8.857409]
 gt: 7773465 Washington_State_Cougars_football,
 pem_pos=1, final_scores=-0.0037876219,
 final_local_scores=-0.057314888, pem=-1.1912286,
 similarity=8.857409]
 10: 13241567 UTEP_Miners_football
 [winner: 303903 University_of_Texas_at_El_Paso,
 pem_pos=0, final_scores=0.0037393123,
 final_local_scores=-0.04273285, pem=-0.3807725,
 similarity=4.8851795]
 gt: 13241567 UTEP_Miners_football, pem_pos=10,
 final_scores=-0.5264865, final_local_scores=-0.60450196,
 pem=-4.163337, similarity=8.691043]
 11: 4688573 1987_NCAA_Division_I-A_football_season
 [winner: 34760 1987, pem_pos=0, final_scores=-0.7216056,
 final_local_scores=-0.55271333, pem=-0.72282183,
 similarity=-8.473537]
 RECALL MISS: 4688573 1987_NCAA_Division_I-
 A_football_season]
 12: 13758243 Wyoming_Cowboys_football
 [winner: 33611 Wyoming, pem_pos=0,
 final_scores=-0.30522045, final_local_scores=-0.275868,
 pem=-0.4538281, similarity=-1.6747501]
 RECALL MISS: 13758243 Wyoming_Cowboys_football]
 13: 17569922 1987_Holiday_Bowl
 [winner: 17569922 1987_Holiday_Bowl, pem_pos=0,
 final_scores=-1.2381783, final_local_scores=-1.3036044,
 pem=0.0, similarity=-34.812443]
 gt: 17569922 1987_Holiday_Bowl, pem_pos=0,
 final_scores=-1.2381783, final_local_scores=-1.3036044,

pem=0.0, similarity=-34.812443]
 14: 17570401 1988_Holiday_Bowl
 [winner: 13279 Holiday, pem_pos=0,
 final_scores=-0.58548796, final_local_scores=-0.52433395,
 pem=-1.3911915, similarity=-3.991942]
 RECALL MISS: 17570401 1988_Holiday_Bowl]
 [winner: 17633715 1988_1_deild, pem_pos=1,
 final_scores=-1.1006192, final_local_scores=-0.9701316,
 pem=-2.0189111, similarity=-13.881796]
 RECALL MISS: 17570401 1988_Holiday_Bowl]
 15: 17595183 1990_Copper_Bowl
 [winner: 125293 Copper, pem_pos=0,
 final_scores=-0.59841204, final_local_scores=-0.4894378,
 pem=-0.15016231, similarity=-9.699481]
 gt: 17595183 1990_Copper_Bowl, pem_pos=5,
 final_scores=-0.78384095, final_local_scores=-0.7426082,
 pem=-2.9563775, similarity=-1.9939551]
 [winner: 17765157 1990_Úrvalsdeild, pem_pos=1,
 final_scores=-1.0325933, final_local_scores=-0.9220026,
 pem=-2.08451, similarity=-12.088795]
 RECALL MISS: 17595183 1990_Copper_Bowl]
 16: 3608976 1991_NCAA_Division_I-A_football_season
 [winner: 2457947 Luby's_massacre, pem_pos=0,
 final_scores=-0.92103493, final_local_scores=-0.89482754,
 pem=0.0, similarity=-22.613726]
 RECALL MISS: 3608976 1991_NCAA_Division_I-
 A_football_season]
 17: 15164948 Utah_State_Aggies_football
 [winner: 267513 Utah_State_University, pem_pos=0,
 final_scores=0.062033102, final_local_scores=0.031470492,
 pem=-0.13969474, similarity=5.788537]
 gt: 15164948 Utah_State_Aggies_football, pem_pos=3,
 final_scores=-0.20918019, final_local_scores=-0.2757852,
 pem=-1.9361018, similarity=6.388555]
 18: 5025856 Northwestern_Wildcats_football
 [winner: 5025856 Northwestern_Wildcats_football,
 pem_pos=1, final_scores=-0.076997206,
 final_local_scores=-0.13427791, pem=-1.4055009,
 similarity=7.7259245]
 gt: 5025856 Northwestern_Wildcats_football, pem_pos=1,
 final_scores=-0.076997206, final_local_scores=-0.13427791,
 pem=-1.4055009, similarity=7.7259245]
 19: 4326736 D'Wayne_Bates
 [winner: 4326736 D'Wayne_Bates, pem_pos=0,
 final_scores=-0.4433418, final_local_scores=-0.49409088,
 pem=0.0, similarity=-10.654944]
 gt: 4326736 D'Wayne_Bates, pem_pos=0,
 final_scores=-0.4433418, final_local_scores=-0.49409088,
 pem=0.0, similarity=-10.654944]
 20: 5740017 Colorado_Buffaloes_football
 [winner: 6669 Colorado_Rockies, pem_pos=11,
 final_scores=0.6239862, final_local_scores=0.68113,

pem=-4.482383, similarity=48.791885|
gt: 5740017 Colorado_Buffaloes_football, pem_pos=3,
final_scores=-0.33812737, final_local_scores=-0.43715814,
pem=-3.5544286, similarity=10.373577]
21: 4291375 2000_NCAA_Division_I-A_football_season
[winner: 34548 2000, pem_pos=0, final_scores=-0.5237812,
final_local_scores=-0.41872883, pem=-0.8862115,
similarity=-3.5866342|
RECALL MISS: 4291375 2000_NCAA_Division_I-
A_football_season]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.691545, final_local_scores=-1.8236086,
pem=-1.219945, similarity=-43.696156|
RECALL MISS: 4291375 2000_NCAA_Division_I-
A_football_season]
22: 8044294 Utah_Utes_football
[winner: 31716 Utah, pem_pos=0, final_scores=-0.2646613,
final_local_scores=-0.25879237, pem=-0.2611261,
similarity=-2.213123|
gt: 8044294 Utah_Utes_football, pem_pos=2,
final_scores=-0.3214497, final_local_scores=-0.41596663,
pem=-3.3862681, similarity=10.091491]
23: 4293387 2002_NCAA_Division_I-A_football_season
[winner: 35502 2002, pem_pos=0, final_scores=-0.8973539,
final_local_scores=-0.7678115, pem=-0.5897454,
similarity=-15.616186|
RECALL MISS: 4293387 2002_NCAA_Division_I-
A_football_season]
24: 4305778 2003_NCAA_Division_I-A_football_season
[winner: 36163 2003, pem_pos=0, final_scores=-0.7752328,
final_local_scores=-0.6114993, pem=-0.50449735,
similarity=-11.415111|
gt: 4305778 2003_NCAA_Division_I-A_football_season,
pem_pos=16, final_scores=-1.6146528,
final_local_scores=-1.8063843, pem=-5.171439,
similarity=-21.693308]
25: 3597670 MAC_Championship_Game
[winner: 3597670 MAC_Championship_Game, pem_pos=0,
final_scores=-0.07245009, final_local_scores=-0.06363895,
pem=-0.09856459, similarity=2.7266092|
gt: 3597670 MAC_Championship_Game, pem_pos=0,
final_scores=-0.07245009, final_local_scores=-0.06363895,
pem=-0.09856459, similarity=2.7266092]
26: 6512082 Miami_RedHawks_football
[winner: 53846 Miami, pem_pos=0, final_scores=-0.194858,
final_local_scores=-0.20682332, pem=-0.35396606,
similarity=-0.15738526|
RECALL MISS: 6512082 Miami_RedHawks_football]
[winner: 22199 Ohio, pem_pos=0, final_scores=-0.5192139,
final_local_scores=-0.5492377, pem=-0.9022944,
similarity=-7.3938203|
RECALL MISS: 6512082 Miami_RedHawks_football]

27: 1901875 Doyt_Perry_Stadium
[winner: 1901875 Doyt_Perry_Stadium, pem_pos=0,
final_scores=-0.041440293, final_local_scores=-0.03404724,
pem=0.0, similarity=3.0736754]
gt: 1901875 Doyt_Perry_Stadium, pem_pos=0,
final_scores=-0.041440293, final_local_scores=-0.03404724,
pem=0.0, similarity=3.0736754]
28: 23916969 Memphis_Tigers_football
[winner: 72886 Memphis_Grizzlies, pem_pos=4,
final_scores=-0.2994519, final_local_scores=-0.31841752,
pem=-3.4150865, similarity=13.15927]
RECALL MISS: 23916969 Memphis_Tigers_football]
29: 14739448 2004_GMAC_Bowl
[winner: 385350 GoDaddy_Bowl, pem_pos=0,
final_scores=0.013279453, final_local_scores=0.011220515,
pem=-0.090608984, similarity=4.917301]
gt: 14739448 2004_GMAC_Bowl, pem_pos=9,
final_scores=-0.54050153, final_local_scores=-0.51242137,
pem=-2.3025851, similarity=1.3198544]
[winner: 35524 2004, pem_pos=0, final_scores=-0.72478473,
final_local_scores=-0.58794117, pem=-0.4915334,
similarity=-10.782589]
RECALL MISS: 14739448 2004_GMAC_Bowl]
30: 20952 Mobile,_Alabama
[winner: 20952 Mobile,_Alabama, pem_pos=0,
final_scores=0.19254889, final_local_scores=0.22327378,
pem=-0.6921803, similarity=14.516831]
gt: 20952 Mobile,_Alabama, pem_pos=0,
final_scores=0.19254889, final_local_scores=0.22327378,
pem=-0.6921803, similarity=14.516831]
[winner: 303 Alabama, pem_pos=0, final_scores=0.18918787,
final_local_scores=0.23007628, pem=-0.32435498,
similarity=12.719542]
RECALL MISS: 20952 Mobile,_Alabama]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8156631, final_local_scores=-1.7756519,
pem=-4.6728287, similarity=-23.487713]
RECALL MISS: 20952 Mobile,_Alabama]
31: 4756185 2005_NCAA_Division_I-A_football_season
[winner: 35984 2005, pem_pos=0, final_scores=-0.6468522,
final_local_scores=-0.4916467, pem=-0.36604404,
similarity=-8.591401]
gt: 4756185 2005_NCAA_Division_I-A_football_season,
pem_pos=15, final_scores=-1.0914059,
final_local_scores=-1.243468, pem=-5.416904,
similarity=-3.55989]
32: 8210777 Kent_State_Golden_Flashes_football
[winner: 16986 Kent_State_University, pem_pos=1,
final_scores=-0.18433687, final_local_scores=-0.23294047,
pem=-1.1156918, similarity=3.2056098]
gt: 8210777 Kent_State_Golden_Flashes_football,
pem_pos=4, final_scores=-0.2855767,

final_local_scores=-0.36151677, pem=-2.2173264,
similarity=5.359497]
33: 3271212 Temple_Owls_football
[winner: 239930 Temple_University, pem_pos=3,
final_scores=-0.42381322, final_local_scores=-0.48061672,
pem=-2.6435006, similarity=4.122918|
gt: 3271212 Temple_Owls_football, pem_pos=5,
final_scores=-0.46348298, final_local_scores=-0.5578445,
pem=-3.1096358, similarity=4.353203]
34: 14238973 2007_Bowling_Green_Falcons_football_team
[winner: 36165 2007, pem_pos=0, final_scores=-0.65094066,
final_local_scores=-0.5197886, pem=-0.24805899,
similarity=-10.072833|
RECALL MISS: 14238973
2007_Bowling_Green_Falcons_football_team]
[winner: 1817986 Advertising_campaign, pem_pos=2,
final_scores=-1.9014885, final_local_scores=-1.9756138,
pem=-2.4250512, similarity=-41.678734|
RECALL MISS: 14238973
2007_Bowling_Green_Falcons_football_team]
35: 14738962 2007_GMAC_Bowl
[winner: 385350 GoDaddy_Bowl, pem_pos=0,
final_scores=0.008008316, final_local_scores=0.0054626167,
pem=-0.090608984, similarity=4.7454743|
gt: 14738962 2007_GMAC_Bowl, pem_pos=6,
final_scores=-0.3923296, final_local_scores=-0.35414964,
pem=-2.1853206, similarity=5.405294]
36: 15128839 2008_Bowling_Green_Falcons_football_team
[winner: 35825 2008, pem_pos=0, final_scores=-0.6551267,
final_local_scores=-0.535163, pem=-0.29331458,
similarity=-10.285528|
RECALL MISS: 15128839
2008_Bowling_Green_Falcons_football_team]
[winner: 690345 Season(sports), pem_pos=0,
final_scores=-1.5137761, final_local_scores=-1.6294229,
pem=-1.219945, similarity=-37.901268|
RECALL MISS: 15128839
2008_Bowling_Green_Falcons_football_team]
37: 19088221 2008_Buffalo_Bulls_football_team
[winner: 4315 Buffalo_Bills, pem_pos=12,
final_scores=-0.2333812, final_local_scores=-0.31595922,
pem=-3.9930825, similarity=16.375864|
RECALL MISS: 19088221 2008_Buffalo_Bulls_football_team]
38: 4145894 1994_NCAA_Division_I-A_football_season
[winner: 18948337 1994, pem_pos=0,
final_scores=-0.7221463, final_local_scores=-0.6173003,
pem=-0.7539553, similarity=-10.231633|
RECALL MISS: 4145894 1994_NCAA_Division_I-
A_football_season]
1: 22031615 Las_Vegas_Locomotives
[pem_pos=0, final_scores=0.23222105,
final_local_scores=0.22545269, pem=0.0, similarity=10.817671]

2: 1229075 Urban_Meyer
[pem_pos=0, final_scores=0.36020786,
final_local_scores=0.33189774, pem=0.0, similarity=13.994204]
3: 4437312 Colorado_Mesa_University
[pem_pos=0, final_scores=0.07484646,
final_local_scores=0.08075175, pem=0.0, similarity=6.499506]
4: 362136 University_of_Northern_Colorado
[pem_pos=0, final_scores=0.32523435,
final_local_scores=0.32425314, pem=0.0, similarity=13.766073]
5: 2698421 Mike_Price
[pem_pos=0, final_scores=0.29820567,
final_local_scores=0.29869464, pem=0.0, similarity=13.003357]
6: 79375 Western_Athletic_Conference
[pem_pos=0, final_scores=0.111350104,
final_local_scores=0.06962001, pem=-0.003959198,
similarity=6.1888437]
7: 3364537 Gary_Barnett
[pem_pos=0, final_scores=0.29181546,
final_local_scores=0.28356358, pem=0.0, similarity=12.551816]
8: 1229075 Urban_Meyer
[pem_pos=0, final_scores=0.28636676,
final_local_scores=0.2512375, pem=0.0, similarity=11.587141]
9: 752260 Ford_Field
[pem_pos=0, final_scores=0.14896445,
final_local_scores=0.15385514, pem=0.0, similarity=8.681057]
10: 14065885 Mike_Groh
[pem_pos=0, final_scores=0.35095638,
final_local_scores=0.38206846, pem=0.0, similarity=15.491397]
11: 1517670 Al_Groh
[pem_pos=0, final_scores=0.38671917,
final_local_scores=0.3925637, pem=0.0, similarity=15.804596]
1: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.1740656,
final_local_scores=0.1215238, pem=0.0, similarity=7.7162256]
2: 10485435 Wyoming_Cowboys_and_Cowgirls
[pem_pos=0, final_scores=0.25719643,
final_local_scores=0.26277205, pem=0.0, similarity=11.931355]
3: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.1698737,
final_local_scores=0.11694479, pem=0.0, similarity=7.579579]
4: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.13178162,
final_local_scores=0.07533498, pem=0.0, similarity=6.337859]
5: 3485377 Virginia_Cavaliers
[pem_pos=0, final_scores=0.09771292,
final_local_scores=0.06714666, pem=-0.2621322,
similarity=7.519018]
6: 9152125 Bowling_Green_Falcons
[pem_pos=0, final_scores=0.113464996,
final_local_scores=0.08752824, pem=-0.16813694,
similarity=7.6160855]
7: 1229075 Urban_Meyer

[pem_pos=0, final_scores=0.12940232,
final_local_scores=0.07977775, pem=0.0, similarity=6.47044]
8: 32097 University_of_Utah
[pem_pos=0, final_scores=0.34872234,
final_local_scores=0.35780862, pem=0.0, similarity=14.767436]
9: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.04405497,
final_local_scores=-0.023226708, pem=0.0,
similarity=3.3965816]
10: 3354114 University_of_the_Incarnate_Word
[pem_pos=0, final_scores=0.12271546,
final_local_scores=0.14534825, pem=0.0, similarity=8.427195]
11: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.10283135,
final_local_scores=0.0437112, pem=0.0, similarity=5.394142]
12: 5399 Colorado
[pem_pos=0, final_scores=0.3743946,
final_local_scores=0.44118822, pem=-0.26926568,
similarity=18.71996]
13: 362136 University_of_Northern_Colorado
[pem_pos=0, final_scores=0.20478983,
final_local_scores=0.19268584, pem=0.0, similarity=9.839843]
14: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.08753215,
final_local_scores=0.026999146, pem=0.0, similarity=4.8954215]
15: 137209 Ogden,_Utah
[pem_pos=0, final_scores=0.11549987,
final_local_scores=0.12993225, pem=-0.64864725,
similarity=11.494596]
16: 8121 Dallas_Cowboys
[pem_pos=0, final_scores=0.055684492,
final_local_scores=0.01417461, pem=-0.9397377,
similarity=9.623151]
17: 31716 Utah
[pem_pos=0, final_scores=0.044275984,
final_local_scores=0.085199386, pem=-0.2611261,
similarity=8.052277]
18: 267513 Utah_State_University
[pem_pos=0, final_scores=0.062033102,
final_local_scores=0.031470492, pem=-0.13969474,
similarity=5.788537]
19: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.038430408,
final_local_scores=-0.029370688, pem=0.0, similarity=3.213233]
20: 78358 Big_Ten_Conference
[pem_pos=0, final_scores=0.15866683,
final_local_scores=0.13371229, pem=0.0, similarity=8.079954]
21: 4326736 D'Wayne_Bates
[pem_pos=0, final_scores=0.048273906,
final_local_scores=0.0481209, pem=0.0, similarity=5.525737]
22: 735314 All-America
[pem_pos=0, final_scores=0.0988626,

final_local_scores=0.06997703, pem=-0.13139357,
similarity=6.8925056]
23: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.044470146,
final_local_scores=-0.022773184, pem=0.0,
similarity=3.4101157]
24: 6669 Colorado_Rockies
[pem_pos=11, final_scores=0.6239862,
final_local_scores=0.68113, pem=-4.482383,
similarity=48.791885]
25: 1229075 Urban_Meyer
[pem_pos=0, final_scores=0.12249951,
final_local_scores=0.07223749, pem=0.0, similarity=6.245424]
26: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.16409422,
final_local_scores=0.110631585, pem=0.0, similarity=7.39118]
27: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.05022843,
final_local_scores=-0.016483143, pem=0.0,
similarity=3.5978231]
28: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.2013274,
final_local_scores=0.15130317, pem=0.0, similarity=8.604901]
29: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.09009664,
final_local_scores=0.02980046, pem=0.0, similarity=4.979018]
30: 20952 Mobile,_Alabama
[pem_pos=0, final_scores=0.19254889,
final_local_scores=0.22327378, pem=-0.6921803,
similarity=14.516831]
31: 303 Alabama
[pem_pos=0, final_scores=0.18918787,
final_local_scores=0.23007628, pem=-0.32435498,
similarity=12.719542]
32: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.077749744,
final_local_scores=0.016313344, pem=0.0, similarity=4.5765357]
33: 30852 Toledo_Mud_Hens
[pem_pos=15, final_scores=0.07811876,
final_local_scores=0.110087395, pem=-4.032892,
similarity=29.306437]
34: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.09486137,
final_local_scores=0.03500521, pem=0.0, similarity=5.1343384]
35: 78433 Mid-American_Conference
[pem_pos=0, final_scores=0.107615486,
final_local_scores=0.064496174, pem=0.0, similarity=6.0144076]
36: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.202197,
final_local_scores=0.15225306, pem=0.0, similarity=8.633248]
37: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.11695416,

final_local_scores=0.059138224, pem=0.0, similarity=5.8545156]
38: 13651173 Gregg_Brandon
[pem_pos=0, final_scores=0.07301767,
final_local_scores=0.011144295, pem=0.0, similarity=4.422281]
39: 59801 University_of_Virginia
[pem_pos=0, final_scores=0.35675138,
final_local_scores=0.3863119, pem=0.0, similarity=15.618031]

Sample: b'Rally_Ireland'

[1 Rally Ireland 1] was a new addition to the [2 1 FIA 2 World Rally Championship 2] -LRB- WRC -RRB- calendar in 3 2007 . It was not part of the 4 2008 schedule , but returned as the First round of the championship in 5 2009 . The North-South event is the largest sporting occasion on the island of [3 Ireland 3] with over 250,000 spectators and a TV audience in 180 countries . As part of the [4 WRC 4] rotation [5 Rally Ireland 5] is not part of the 6 2010 season but was expected to figure in 2011 . However , on 19 February 2010 [1 1 North 0 ne Sport 1] the [6 WRC 6] promoter , announced that [7 Rally Ireland 7] would not be awarded a place on the 2011 calendar . The 2007 event took place between 15 and 18 November 2007 , starting with the Super Special Stage in the grounds of the [2 2 Stormont Parliament Buildings 2] on the outskirts of [3 3 Belfast 3] , [4 4 County Down 4] . The remainder of the event was based in the north-west of [8 Ireland 8] in the counties of [9 7 Sligo 9] , [5 5 Fermanagh 5] , [10 8 Donegal 10] , [6 6 Leitrim 6] , [7 7 Tyrone 7] , [8 8 Roscommon 8] and [9 9 Cavan 9] . [11 Rally Ireland 11] is organised by Motorsport Ireland . The event itself joined the [12 WRC 12] in 2007 when it was won by reigning World Champions [13 Sébastien Loeb 13] and [14 Daniel Elena 14] in a Citroën C4 . The rally is managed by Event Director John Naylor while the Clerk of the Course is Gordon Noble . A steering committee which is charged with developing the event is in place . This committee is chaired by Austin Frazer and includes representative of Motorsport Ireland -LRB- MI -RRB- and The Association of Northern Ireland Car Clubs -LRB- ANICC -RRB- . [15 Rally Ireland 15] received strong government support from both sides of the border . It continues a long tradition of cross-border rallying in [16 Ireland 16] , which dates back to the early years of the [10 10 Circuit of Ireland Rally 10] in the 1930s . The Rally was held on 9 asphalt conc rete , which was very wavy and slippery due to the humidly [17 Irish 17] weather , what makes this character being similar to the

- 1: 10819898 North_One_Television
- 2: 539931 Parliament_Buildings_(Northern_Ireland)
- 3: 5046 Belfast
- 4: 128263 County_Down
- 5: 5833 County_Fermanagh
- 6: 52817 County_Leitrim
- 7: 52821 County_Tyrone

8: 92633 County_Roscommon
9: 429660 County_Cavan
10: 11353404 Circuit_of_Ireland_Rally
1: 248019 Fédération_Internationale_de_l'Automobile
[winner: 234243 World_Rally_Championship, pem_pos=0,
final_scores=0.08529042, final_local_scores=0.09946346,
pem=0.0, similarity=7.057901|
RECALL MISS: 248019
Fédération_Internationale_de_l'Automobile]
2: 234243 World_Rally_Championship
[winner: 234243 World_Rally_Championship, pem_pos=0,
final_scores=0.08529042, final_local_scores=0.09946346,
pem=0.0, similarity=7.057901|
gt: 234243 World_Rally_Championship, pem_pos=0,
final_scores=0.08529042, final_local_scores=0.09946346,
pem=0.0, similarity=7.057901]
3: 8127993 2007_World_Rally_Championship_season
[winner: 36165 2007, pem_pos=0, final_scores=-0.6276328,
final_local_scores=-0.4983343, pem=-0.24805899,
similarity=-9.432593|
RECALL MISS: 8127993
2007_World_Rally_Championship_season]
4: 13909520 2008_World_Rally_Championship_season
[winner: 35825 2008, pem_pos=0, final_scores=-0.6107655,
final_local_scores=-0.4798594, pem=-0.29331458,
similarity=-8.6351595|
RECALL MISS: 13909520
2008_World_Rally_Championship_season]
5: 18811741 2009_World_Rally_Championship_season
[winner: 35983 2009, pem_pos=0, final_scores=-0.7978937,
final_local_scores=-0.67344344, pem=-0.9296576,
similarity=-10.951561|
RECALL MISS: 18811741
2009_World_Rally_Championship_season]
6: 18812209 2010_World_Rally_Championship_season
[winner: 43226 2010, pem_pos=0, final_scores=-0.88838506,
final_local_scores=-0.7989131, pem=-1.6731367,
similarity=-10.652671|
RECALL MISS: 18812209
2010_World_Rally_Championship_season]
7: 146777 County_Sligo
[winner: 584663 Sligo, pem_pos=0,
final_scores=0.42150968, final_local_scores=0.41404969,
pem=-0.3973797, similarity=18.606794|
gt: 146777 County_Sligo, pem_pos=1,
final_scores=0.1110125, final_local_scores=0.07960124,
pem=-1.2515068, similarity=13.271062]
8: 45576 County_Donegal
[winner: 996491 Donegal, pem_pos=0,
final_scores=0.12509535, final_local_scores=0.09189536,
pem=-0.9030535, similarity=11.742999|
gt: 45576 County_Donegal, pem_pos=1,

final_scores=0.10960512, final_local_scores=0.059947386,
pem=-1.0730512, similarity=11.714082]
9: 387797 Asphalt_concrete
[winner: 5241974 Asphalt_shingle, pem_pos=2,
final_scores=-0.88883144, final_local_scores=-0.82236916,
pem=-0.6824507, similarity=-16.74015]
gt: 387797 Asphalt_concrete, pem_pos=1,
final_scores=-1.0558676, final_local_scores=-1.0499105,
pem=-0.6308349, similarity=-23.811134]
[winner: 5371 Concrete, pem_pos=0,
final_scores=-1.0882016, final_local_scores=-1.0996218,
pem=-0.07044339, similarity=-28.342113]
RECALL MISS: 387797 Asphalt_concrete]
1: 10819898 North_One_Television
[pem_pos=0, final_scores=0.1230465,
final_local_scores=0.22538218, pem=0.0, similarity=10.815567]
2: 539931 Parliament_Buildings_(Northern_Ireland)
[pem_pos=0, final_scores=0.18700011,
final_local_scores=0.21244946, pem=0.0, similarity=10.429628]
3: 5046 Belfast
[pem_pos=0, final_scores=0.45096958,
final_local_scores=0.44083732, pem=-0.13156427,
similarity=17.960646]
4: 128263 County_Down
[pem_pos=0, final_scores=0.44596678,
final_local_scores=0.4330802, pem=0.0, similarity=17.013689]
5: 5833 County_Fermanagh
[pem_pos=0, final_scores=0.17369108,
final_local_scores=0.15128288, pem=-0.3936055,
similarity=10.744783]
6: 52817 County_Leitrim
[pem_pos=0, final_scores=0.20475851,
final_local_scores=0.18366855, pem=-0.46183294,
similarity=12.082268]
7: 52821 County_Tyrone
[pem_pos=0, final_scores=0.050270393,
final_local_scores=0.0006105751, pem=-1.1661593,
similarity=10.449689]
8: 92633 County_Roscommon
[pem_pos=0, final_scores=0.1481115,
final_local_scores=0.120339155, pem=-0.90064996,
similarity=12.578749]
9: 429660 County_Cavan
[pem_pos=1, final_scores=0.078786746,
final_local_scores=0.036997452, pem=-1.2261521,
similarity=11.861796]
10: 11353404 Circuit_of_Ireland_Rally
[pem_pos=0, final_scores=0.103219196,
final_local_scores=0.12371406, pem=0.0, similarity=7.7815876]
1: 5935156 Rally_Ireland
[pem_pos=0, final_scores=0.22586931,
final_local_scores=0.23367643, pem=-0.24207152,

similarity=12.379507]
2: 234243 World_Rally_Championship
[pem_pos=0, final_scores=0.08529042,
final_local_scores=0.09946346, pem=0.0, similarity=7.057901]
3: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.49975193,
final_local_scores=0.47738296, pem=-1.5040284,
similarity=26.514914]
4: 234243 World_Rally_Championship
[pem_pos=0, final_scores=0.09995128,
final_local_scores=0.10911472, pem=-0.42608443,
similarity=9.663027]
5: 5935156 Rally_Ireland
[pem_pos=0, final_scores=0.11592518,
final_local_scores=0.113579184, pem=-0.24207152,
similarity=8.795566]
6: 234243 World_Rally_Championship
[pem_pos=0, final_scores=0.07364683,
final_local_scores=0.08038111, pem=-0.42608443,
similarity=8.805559]
7: 5935156 Rally_Ireland
[pem_pos=0, final_scores=0.16043817,
final_local_scores=0.16220284, pem=-0.24207152,
similarity=10.246592]
8: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.46378803,
final_local_scores=0.44090945, pem=-1.5040284,
similarity=25.42647]
9: 584663 Sligo
[pem_pos=0, final_scores=0.42150968,
final_local_scores=0.41404969, pem=-0.3973797,
similarity=18.606794]
10: 996491 Donegal
[pem_pos=0, final_scores=0.12509535,
final_local_scores=0.09189536, pem=-0.9030535,
similarity=11.742999]
11: 5935156 Rally_Ireland
[pem_pos=0, final_scores=0.15076832,
final_local_scores=0.15164003, pem=-0.24207152,
similarity=9.931376]
12: 234243 World_Rally_Championship
[pem_pos=0, final_scores=0.118345484,
final_local_scores=0.12920761, pem=-0.42608443,
similarity=10.26264]
13: 944612 Sébastien_Loeb
[pem_pos=0, final_scores=0.33067697,
final_local_scores=0.384777, pem=0.0, similarity=15.572226]
14: 4295296 Daniel_Elena
[pem_pos=0, final_scores=0.48035377,
final_local_scores=0.5798627, pem=0.0, similarity=21.393972]
15: 5935156 Rally_Ireland
[pem_pos=0, final_scores=0.15999727,

final_local_scores=0.16172123, pem=-0.24207152,
 similarity=10.23222]
 16: 14560 Republic_of_Ireland
 [pem_pos=1, final_scores=0.5794448,
 final_local_scores=0.5644353, pem=-1.5040284,
 similarity=29.11273]
 17: 14560 Republic_of_Ireland
 [pem_pos=3, final_scores=0.62396747,
 final_local_scores=0.6130696, pem=-2.2576945,
 similarity=34.662624]

Sample: b'George_Clifford_Wilson'

[1 George Clifford Wilson 1] -LRB- 27 July 1902 - 18 May
 1957 -RRB- was an [1 1 English 1] who followed his father ,
 the more accomplished [2 2 George Alfred Wilson 2] , into a
 career as a right-arm 1 fast bowler -LRB- and right-handed
 batsman at the bottom of the 2 order -RRB- with [3 3 Worcest
 rshire 3] . Born in [4 4 Kidderminster 4] , [2 Wilson 2] made
 his 3 first-class debut in the 4 local derby with [5 5 Warwic
 kshire 5] at [6 6 Edgbaston 6] in May of the 5 1924 season ,
 but did little , bowling only three overs and making nought
 in his only innings . The match was notable both for a 6 hat-
 trick by [3 Worcestershire 3] 's [7 7 Charles Preece 7] and
 for the county 's desperate but ultimately successful
 struggle for a draw , having been reduced to 48/8 in their
 second innings when stumps were drawn . He also went
 wicketless in his second match , but took 2-7 in the second
 innings of his third . [4 Wilson 4] played a full part in the
 1924 [8 8 County Championship 8] campaign , playing in 20 of [
 5 Worcestershire 5] 's 24 games as well as one against the
 touring and finishing with 38 wickets at an 7 average of a
 little under 24 runs apiece . 8 The following summer brought
 him his first five-wicket haul : 5-67 against [9 9 Hampshire 9
] in August ; and 44 wickets in all . By some way [6 Wilson 6]
 's most successful season was his last , 9 1926 . He took 68
 first-class wickets in total , and five or more in an innings
 on seven occasions , this latter figure being second only to [
 10 10 Fred Root 10] in the [7 Worcestershire 7] side that
 year . In the space of one August week he took five wickets
 in each innings - his only ten-wicket match haul - against [11
 11 Derbyshire 11] on his home ground at [8 Kidderminster 8] ,
 then claimed his best innings figures of 8-81 in a drawn game
 against [12 12 Somerset 12] at [13 13 Weston-super-Mare 13] .
 [9 Wilson 9] played his final match in early September 1926
 against [14 14 Middlesex 14] , but bad weather meant that
 little play was possible and he did not take a wicket . He
 died in 10 Elswick , [15 15 Newcastle upon Tyne 15]

1: 9316 England
 2: 2763296 George_Alfred_Wilson
 3: 1156428 Worcestershire_County_Cricket_Club
 4: 39102 Kidderminster
 5: 667292 Warwickshire_County_Cricket_Club

6: 2641030 Edgbaston_Cricket_Ground
7: 18110950 Charles_Preece
8: 683599 County_Championship
9: 1919370 Hampshire_County_Cricket_Club
10: 1174056 Fred_Root
11: 1829984 Derbyshire_County_Cricket_Club
12: 1622178 Somerset_County_Cricket_Club
13: 330072 Weston-super-Mare
14: 822103 Middlesex_County_Cricket_Club
15: 70348 Newcastle_upon_Tyne
1: 678695 Fast_bowling
 [winner: 678695 Fast_bowling, pem_pos=0,
final_scores=-1.018996, final_local_scores=-1.1635473,
pem=-0.0193202, similarity=-30.52779|
 gt: 678695 Fast_bowling, pem_pos=0,
final_scores=-1.018996, final_local_scores=-1.1635473,
pem=-0.0193202, similarity=-30.52779]
2: 2298801 Batting_order_(cricket)
 [winner: 10932122 Order_(mouldings), pem_pos=18,
final_scores=-1.1783909, final_local_scores=-1.0774276,
pem=-2.8903718, similarity=-12.344585|
 RECALL MISS: 2298801 Batting_order_(cricket)]
3: 11040 First-class_cricket
 [winner: 662844 First-class_citizen, pem_pos=3,
final_scores=-1.2377255, final_local_scores=-1.1569278,
pem=-1.5935546, similarity=-21.769321|
 gt: 11040 First-class_cricket, pem_pos=0,
final_scores=-1.3902656, final_local_scores=-1.5759892,
pem=-0.053819276, similarity=-42.64827]
4: 416204 Local_derby
 [winner: 1718152 Local_variable, pem_pos=2,
final_scores=-1.4019932, final_local_scores=-1.3361833,
pem=-2.3668852, similarity=-22.913174|
 RECALL MISS: 416204 Local_derby]
 [winner: 875654 Bishop_of_Derby, pem_pos=13,
final_scores=-1.432225, final_local_scores=-1.4653988,
pem=-4.1165833, similarity=-17.25409|
 RECALL MISS: 416204 Local_derby]
5: 7611731 1924_English_cricket_season
 [winner: 34675 1924, pem_pos=0, final_scores=-0.6777738,
final_local_scores=-0.5576219, pem=-0.80803895,
similarity=-8.156595|
 RECALL MISS: 7611731 1924_English_cricket_season]
 [winner: 1340236 Eastertide, pem_pos=2,
final_scores=-1.6362733, final_local_scores=-1.6470804,
pem=-2.3723626, similarity=-32.16118|
 RECALL MISS: 7611731 1924_English_cricket_season]
6: 411541 Hat-trick
 [winner: 411541 Hat-trick, pem_pos=0,
final_scores=-1.4719653, final_local_scores=-1.5808673,
pem=-0.0025918067, similarity=-43.072426|
 gt: 411541 Hat-trick, pem_pos=0,

```

final_scores=-1.4719653, final_local_scores=-1.5808673,
pem=-0.0025918067, similarity=-43.072426 ]
7: 875871 Bowling_average
  [winner: 875871 Bowling_average, pem_pos=0,
final_scores=-1.0351126, final_local_scores=-1.1825438,
pem=0.0, similarity=-31.199749]
  gt: 875871 Bowling_average, pem_pos=0,
final_scores=-1.0351126, final_local_scores=-1.1825438,
pem=0.0, similarity=-31.199749 ]
8: 5671925 1925_English_cricket_season
  [winner: 7634850 2001_English_cricket_season, pem_pos=1,
final_scores=-0.8805035, final_local_scores=-0.9493768,
pem=-1.0986123, similarity=-18.26716]
  gt: 5671925 1925_English_cricket_season, pem_pos=5,
final_scores=-0.9159401, final_local_scores=-0.9992313,
pem=-1.0986123, similarity=-19.754917 ]
9: 7620860 1926_English_cricket_season
  [winner: 34609 1926, pem_pos=0, final_scores=-0.724252,
final_local_scores=-0.6012025, pem=-0.76264197,
similarity=-9.704001]
  RECALL MISS: 7620860 1926_English_cricket_season]
10: 2391822 Elswick,_Tyne_and_Wear
  [winner: 2391822 Elswick,_Tyne_and_Wear, pem_pos=0,
final_scores=-0.10677092, final_local_scores=-0.027715832,
pem=-0.20110148, similarity=4.3562384]
  gt: 2391822 Elswick,_Tyne_and_Wear, pem_pos=0,
final_scores=-0.10677092, final_local_scores=-0.027715832,
pem=-0.20110148, similarity=4.3562384 ]
1: 9316 England
  [pem_pos=1, final_scores=0.34281677,
final_local_scores=0.3714847, pem=-1.3527626,
similarity=22.532091]
2: 2763296 George_Alfred_Wilson
  [pem_pos=0, final_scores=0.28825247,
final_local_scores=0.2621052, pem=0.0, similarity=11.911455]
3: 1156428 Worcestershire_County_Cricket_Club
  [pem_pos=1, final_scores=0.5735386,
final_local_scores=0.5326175, pem=-0.9164197,
similarity=24.967714]
4: 39102 Kidderminster
  [pem_pos=0, final_scores=0.1432362,
final_local_scores=0.151137, pem=0.0, similarity=8.599942]
5: 667292 Warwickshire_County_Cricket_Club
  [pem_pos=1, final_scores=0.73888737,
final_local_scores=0.7178941, pem=-1.1093602,
similarity=31.545977]
6: 2641030 Edgbaston_Cricket_Ground
  [pem_pos=1, final_scores=0.37585467,
final_local_scores=0.34818518, pem=-0.4505704,
similarity=16.930525]
7: 18110950 Charles_Preece
  [pem_pos=0, final_scores=0.5870777,

```

final_local_scores=0.6675295, pem=0.0, similarity=24.010126]
8: 683599 County_Championship
[pem_pos=0, final_scores=0.20070316,
final_local_scores=0.13521233, pem=0.0, similarity=8.124719]
9: 1919370 Hampshire_County_Cricket_Club
[pem_pos=1, final_scores=0.71294594,
final_local_scores=0.69406, pem=-1.9232318,
similarity=35.26068]
10: 1174056 Fred_Root
[pem_pos=0, final_scores=0.6312987,
final_local_scores=0.68174535, pem=0.0, similarity=24.434355]
11: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.80274063,
final_local_scores=0.78840095, pem=-1.4828352,
similarity=35.681057]
12: 1622178 Somerset_County_Cricket_Club
[pem_pos=1, final_scores=0.64771813,
final_local_scores=0.6150546, pem=-2.4139037,
similarity=35.571354]
13: 330072 Weston-super-Mare
[pem_pos=0, final_scores=0.20515577,
final_local_scores=0.229379, pem=0.0, similarity=10.93484]
14: 822103 Middlesex_County_Cricket_Club
[pem_pos=1, final_scores=0.68635625,
final_local_scores=0.65354204, pem=-1.5695794,
similarity=32.128326]
15: 70348 Newcastle_upon_Tyne
[pem_pos=0, final_scores=0.05675365,
final_local_scores=0.07382597, pem=0.0, similarity=6.292827]
1: 2804802 George_Clifford_Wilson
[pem_pos=0, final_scores=0.04220812,
final_local_scores=0.017931178, pem=0.0, similarity=4.624815]
2: 2763296 George_Alfred_Wilson
[pem_pos=1, final_scores=0.6182347,
final_local_scores=0.62200063, pem=0.0, similarity=22.651453]
3: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.6506542,
final_local_scores=0.6168546, pem=-0.9164197,
similarity=27.481518]
4: 2763296 George_Alfred_Wilson
[pem_pos=1, final_scores=0.57377213,
final_local_scores=0.57343215, pem=0.0, similarity=21.202072]
5: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.6718606,
final_local_scores=0.6400193, pem=-0.9164197,
similarity=28.172798]
6: 2763296 George_Alfred_Wilson
[pem_pos=1, final_scores=0.39664167,
final_local_scores=0.37994397, pem=0.0, similarity=15.427999]
7: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.5115767,
final_local_scores=0.4649335, pem=-0.9164197,

similarity=22.94789]
8: 39102 Kidderminster
[pem_pos=0, final_scores=0.21956898,
final_local_scores=0.23451895, pem=0.0, similarity=11.088226]
9: 2763296 George_Alfred_Wilson
[pem_pos=1, final_scores=0.51195616,
final_local_scores=0.50590754, pem=0.0, similarity=19.187002]

Sample: b'Ulrike_Maier'

[1 Ulrike Maier 1] -LRB- 22 October 1967 -- 29 January 1994 -
RRB- was a 1 World Cup 2 alpine ski racer from [1 1 Austria 1]
, a two-time 3 World Champion in . Born in [2 2 Rauris 2] , [2
4 Salzburg 2] , where her father ran a ski school , [3 Maier 3
] won the gold medal at the 5 World Championships in both 6 19
89 and 7 1991 . She also took home the silver medal in the
1991 event . Her first of five World Cup wins came in
November 8 1992 and she had 21 podiums 59 top ten finishes in
her World Cup career . Two weeks prior to the 9 1994 Winter Ol
ympics , the women 's World Cup was in [3 3 Garmisch-
Partenkirchen 3] , [4 4 Germany 4] , in late January . The 10
downhill on the classic [4 Kandahar 4] course at [5 5 Garmisch
Classic 5] was held on Saturday , January 29 , following an
overnight snowfall . In a narrow part of the lower course
less than twenty seconds from the finish , [5 Maier 5] 's
right ski caught an inside edge at , possibly from a patch of
soft snow , and caused a crash which broke her neck . She
died of her injuries shortly after being evacuated to the
hospital in nearby [6 6 Murnau 6] . At age 26 , she had
thought to retire at the end of the 11 1994 season.However ,
she was reconsidering her decision in the days before the
fateful downhill run , planning to continue until in [7 7 Spa
in 7] . Unlike most other fatal skiing accidents , her crash
happened during a live television broadcast . [6 Maier 6] was
survived by her daughter Melanie -LRB- b. 1989 -RRB- and was
buried in her home village of [7 Rauris 7] , where thousands
attended her funeral . Teammate and close friend [8 8 Anita Wa
chter 8] wore [8 Maier 8] 's world championship medals in the
procession . Initially it was claimed that her death was
caused by hitting a wooden timing post . Based on that
claim , [9 Maier 9] 's fiancé Hubert Schweighofer criticized
the organizers of the race and filed suit against them .
However , several months later the court found that [10 Maier
10] actually did not hit the timing post with her head , but
probably broke her neck by crashing into a pile of snow on
the border of the race course . The court did not find any
negligence from the organizers and dismissed the suit .
Manslaughter charges against two 12 FIS

1: 26964606 Austria
2: 8627164 Rauris
3: 412433 Garmisch-Partenkirchen
4: 11867 Germany
5: 21582293 Garmisch_Classic

6: 1474848 Murnau_am_Staffelsee
7: 26667 Spain
8: 364378 Anita_Wachter
1: 183628 FIS_Alpine_Ski_World_Cup
[winner: 11370 FIFA_World_Cup, pem_pos=0,
final_scores=-0.17759201, final_local_scores=-0.17411502,
pem=-1.001087, similarity=4.33784|
gt: 183628 FIS_Alpine_Ski_World_Cup, pem_pos=2,
final_scores=-0.19298966, final_local_scores=-0.21670866,
pem=-2.7628422, similarity=12.64746]
2: 39576 Alpine_skiing
[winner: 679433 Ski_warfare, pem_pos=2,
final_scores=-0.93798876, final_local_scores=-0.9524373,
pem=-0.69268847, similarity=-20.565968|
gt: 39576 Alpine_skiing, pem_pos=1,
final_scores=-1.0253961, final_local_scores=-1.1073482,
pem=-0.65621436, similarity=-25.387169]
[winner: 2619936 Racer_(Kennywood), pem_pos=2,
final_scores=-0.9853953, final_local_scores=-0.9079934,
pem=-1.4941926, similarity=-14.880969|
RECALL MISS: 39576 Alpine_skiing]
[winner: 23160842 Alpine,_Mendocino_County,_California,
pem_pos=9, final_scores=-1.1750461,
final_local_scores=-1.0977352, pem=-3.0050297,
similarity=-12.327075|
gt: 39576 Alpine_skiing, pem_pos=13,
final_scores=-1.9402488, final_local_scores=-2.1066859,
pem=-3.601453, similarity=-39.192734]
3: 3226813 FIS_Alpine_World_Ski_Championships
[winner: 229356 World_Figure_Skating_Championships,
pem_pos=1, final_scores=-0.16469443,
final_local_scores=-0.1601536, pem=-1.8816465,
similarity=9.543096|
RECALL MISS: 3226813 FIS_Alpine_World_Ski_Championships]
4: 55859 Salzburg_(state)
[winner: 40335 Salzburg, pem_pos=0,
final_scores=0.4652825, final_local_scores=0.48909402,
pem=-0.11955855, similarity=19.335432|
gt: 55859 Salzburg_(state), pem_pos=1,
final_scores=-0.0815559, final_local_scores=-0.08239579,
pem=-1.7480965, similarity=11.137279]
5: 3226813 FIS_Alpine_World_Ski_Championships
[winner: 229356 World_Figure_Skating_Championships,
pem_pos=1, final_scores=-0.27599412,
final_local_scores=-0.2817316, pem=-2.3743477,
similarity=8.594352|
RECALL MISS: 3226813 FIS_Alpine_World_Ski_Championships]
6: 9073139 FIS_Alpine_World_Ski_Championships_1989
[winner: 34847 1989, pem_pos=0, final_scores=-0.8129906,
final_local_scores=-0.73208797, pem=-0.6980697,
similarity=-13.9610405|
RECALL MISS: 9073139

FIS_Alpine_World_Ski_Championships_1989]
7: 9125948 FIS_Alpine_World_Ski_Championships_1991
[winner: 34703 1991, pem_pos=0, final_scores=-0.78529286,
final_local_scores=-0.6816615, pem=-0.6661314,
similarity=-12.6299]
RECALL MISS: 9125948
FIS_Alpine_World_Ski_Championships_1991]
8: 4241285 1993_Alpine_Skiing_World_Cup
[winner: 34669 1992, pem_pos=0, final_scores=-0.8346614,
final_local_scores=-0.7008235, pem=-0.76294684,
similarity=-12.675235]
RECALL MISS: 4241285 1993_Alpine_Skiing_World_Cup]
9: 3309347 Alpine_skiing_at_the_1994_Winter_Olympics
[winner: 34069 Winter_Olympic_Games, pem_pos=0,
final_scores=-0.024747983, final_local_scores=-0.021273553,
pem=-0.28208634, similarity=4.9888973]
RECALL MISS: 3309347
Alpine_skiing_at_the_1994_Winter_Olympics]
[winner: 18948337 1994, pem_pos=0,
final_scores=-1.2115315, final_local_scores=-1.1649134,
pem=-0.7539553, similarity=-26.573498]
RECALL MISS: 3309347
Alpine_skiing_at_the_1994_Winter_Olympics]
10: 218183 Downhill
[winner: 21245534 The_Downhill, pem_pos=0,
final_scores=-0.6514785, final_local_scores=-0.54244447,
pem=0.0, similarity=-12.097911]
RECALL MISS: 218183 Downhill]
11: 4241273 1994_Alpine_Skiing_World_Cup
[winner: 18948337 1994, pem_pos=0,
final_scores=-0.87373793, final_local_scores=-0.79592526,
pem=-0.7539553, similarity=-15.562159]
RECALL MISS: 4241273 1994_Alpine_Skiing_World_Cup]
12: 175281 International_Ski_Federation
[winner: 403399 Islamic_Salvation_Front, pem_pos=2,
final_scores=0.00079362094, final_local_scores=0.11905247,
pem=-1.4075395, similarity=15.296894]
gt: 175281 International_Ski_Federation, pem_pos=0,
final_scores=-0.09181069, final_local_scores=-0.10224323,
pem=-0.31611037, similarity=2.7576284]
1: 26964606 Austria
[pem_pos=0, final_scores=0.6893695,
final_local_scores=0.72298723, pem=-0.25262418,
similarity=27.038904]
2: 8627164 Rauris
[pem_pos=0, final_scores=0.30844235,
final_local_scores=0.39222586, pem=0.0, similarity=15.794515]
3: 412433 Garmisch-Partenkirchen
[pem_pos=0, final_scores=0.32998663,
final_local_scores=0.31829667, pem=0.0, similarity=13.588321]
4: 11867 Germany
[pem_pos=0, final_scores=0.7799248,

```

final_local_scores=0.83081466, pem=-0.15529114,
similarity=29.727377]
5: 21582293 Garmisch_Classic
    [pem_pos=0, final_scores=0.109079406,
final_local_scores=0.15043816, pem=0.0, similarity=8.579087]
6: 1474848 Murnau_am_Staffelsee
    [pem_pos=2, final_scores=0.23312052,
final_local_scores=0.28350192, pem=-0.7694554,
similarity=16.734394]
7: 26667 Spain
    [pem_pos=0, final_scores=0.59483284,
final_local_scores=0.68364084, pem=-0.18857051,
similarity=25.516396]
8: 364378 Anita_Wachter
    [pem_pos=0, final_scores=0.62947863,
final_local_scores=0.7249608, pem=0.0, similarity=25.72399]
1: 4178529 Ulrike_Maier
    [pem_pos=0, final_scores=0.5500557,
final_local_scores=0.5602831, pem=0.0, similarity=20.809679]
2: 40335 Salzburg
    [pem_pos=0, final_scores=0.4652825,
final_local_scores=0.48909402, pem=-0.11955855,
similarity=19.335432]
3: 4178529 Ulrike_Maier
    [pem_pos=0, final_scores=0.42926884,
final_local_scores=0.4231221, pem=0.0, similarity=16.71652]
4: 17260 Kandahar
    [pem_pos=0, final_scores=0.21820007,
final_local_scores=0.33044815, pem=-0.11192558,
similarity=14.559613]
5: 4178529 Ulrike_Maier
    [pem_pos=0, final_scores=0.39248258,
final_local_scores=0.38293868, pem=0.0, similarity=15.517367]
6: 4178529 Ulrike_Maier
    [pem_pos=0, final_scores=0.3843575,
final_local_scores=0.3740633, pem=0.0, similarity=15.252509]
7: 8627164 Rauris
    [pem_pos=0, final_scores=0.316581,
final_local_scores=0.40111613, pem=0.0, similarity=16.059818]
8: 4178529 Ulrike_Maier
    [pem_pos=0, final_scores=0.34418595,
final_local_scores=0.33018196, pem=0.0, similarity=13.943001]
9: 4178529 Ulrike_Maier
    [pem_pos=0, final_scores=0.28314084,
final_local_scores=0.26871923, pem=0.0, similarity=12.10883]
10: 4178529 Ulrike_Maier
    [pem_pos=0, final_scores=0.34698993,
final_local_scores=0.3332449, pem=0.0, similarity=14.034405]

```

Sample: b'Iowa_Highway_7'

Iowa Highway 7 is a state **1 highway** that runs from east to west across the northwestern portion of **2 Iowa** . Highway 7 is

a fair sized highway in Iowa , at 74 miles -LRB- 119 kilometers -RRB- long . The western terminus of Iowa Highway 7 is at 3 Iowa Highway 3 north of 4 Aurelia . The eastern terminus of Iowa 7 is at near [1 1 Fort Dodge 1] . Iowa Highway 7 begins north of Aurelia at Iowa Highway 3 . It goes south through Aurelia , then turns southeast to go through 5 Alta . It continues southeast to [2 2 Storm Lake 2] , where it intersects 6 Iowa Highway 110 . After passing through [1 Storm Lake 1] , it intersects and goes south with [2 U.S. 2] 71 for . Iowa 7 then turns east , passing through 7 Newell and 8 Fonda , before turning south again at an intersection with 9 Iowa Highway 4 . Iowa 7 and Iowa 4 go south through 10 Pomeroy , before Iowa 7 turns east again . After passing through 11 Manson , it continues east and ends at U.S. Highway 169 in northwestern [3 3 Fort Dodge 3] . This is the fourth state highway to be designated Iowa Highway 7 . The first Iowa Highway 7 existed from 1919 to 1939 and was replaced by in eastern [3 Iowa 3] and 12 Iowa Highway 64 in western Iowa . The second Iowa Highway 7 , a short spur from to [4 13 Oakville 4] existed from 1955 to 1957 . The third Iowa Highway 7 existed from 1957 to 1969 when was redesignated . The current Iowa Highway 7 route was designated in 1969 . The eastern end was truncated in 1990 when

1: 114334 Fort_Dodge, _Iowa

2: 112805 Storm_Lake, _Iowa

3: 114334 Fort_Dodge, _Iowa

1: 48519 Highway

[winner: 332996 State_highway, pem_pos=0,
final_scores=-1.2275723, final_local_scores=-1.2826843,
pem=-0.024095139, similarity=-34.057114]

RECALL MISS: 48519 Highway]

2: 26810748 Iowa

[winner: 26810748 Iowa, pem_pos=0,
final_scores=-0.015728906, final_local_scores=-0.089554176,
pem=-0.3546882, similarity=3.3460872]
gt: 26810748 Iowa, pem_pos=0, final_scores=-0.015728906,
final_local_scores=-0.089554176, pem=-0.3546882,
similarity=3.3460872]

3: 4664862 Iowa_Highway_3

[winner: 4664862 Iowa_Highway_3, pem_pos=0,
final_scores=-0.123168364, final_local_scores=-0.051337384,
pem=0.0, similarity=2.557703]
gt: 4664862 Iowa_Highway_3, pem_pos=0,
final_scores=-0.123168364, final_local_scores=-0.051337384,
pem=0.0, similarity=2.557703]

4: 112870 Aurelia, _Iowa

[winner: 180484 Orillia, pem_pos=18,
final_scores=-0.25151688, final_local_scores=-0.28872955,
pem=-3.1354942, similarity=12.524752]
gt: 112870 Aurelia, _Iowa, pem_pos=2,
final_scores=-0.2713682, final_local_scores=-0.2855247,

```

pem=-1.7809485, similarity=5.254162 ]
5: 112798 Alta,_Iowa
   [winner: 178000 Alta,_Norway, pem_pos=0,
   final_scores=-0.05943042, final_local_scores=-0.015501231,
pem=-1.4639112, similarity=11.588102|
   gt: 112798 Alta,_Iowa, pem_pos=9,
   final_scores=-0.530092, final_local_scores=-0.57001424,
pem=-2.842181, similarity=2.5355735 ]
6: 9720824 Iowa_Highway_110
   [winner: 9720824 Iowa_Highway_110, pem_pos=0,
   final_scores=-0.27263272, final_local_scores=-0.1786,
pem=0.0, similarity=-1.2400672|
   gt: 9720824 Iowa_Highway_110, pem_pos=0,
   final_scores=-0.27263272, final_local_scores=-0.1786,
pem=0.0, similarity=-1.2400672 ]
7: 112802 Newell,_Iowa
   [winner: 5802952 Newell,_California, pem_pos=1,
   final_scores=-0.2722957, final_local_scores=-0.26514652,
pem=-1.9318647, similarity=6.682993|
   gt: 112802 Newell,_Iowa, pem_pos=7,
   final_scores=-0.3690641, final_local_scores=-0.3970922,
pem=-2.7163014, similarity=7.011359 ]
8: 114143 Fonda,_Iowa
   [winner: 63520 Jane_Fonda, pem_pos=2,
   final_scores=-0.04594387, final_local_scores=0.025467128,
pem=-1.4901745, similarity=12.953504|
   gt: 114143 Fonda,_Iowa, pem_pos=1,
   final_scores=-0.21861541, final_local_scores=-0.21790329,
pem=-1.0709971, similarity=3.4112928 ]
9: 5835725 Iowa_Highway_4
   [winner: 5835725 Iowa_Highway_4, pem_pos=0,
   final_scores=-0.2522172, final_local_scores=-0.17957471,
pem=0.0, similarity=-1.2691548|
   gt: 5835725 Iowa_Highway_4, pem_pos=0,
   final_scores=-0.2522172, final_local_scores=-0.17957471,
pem=0.0, similarity=-1.2691548 ]
10: 112824 Pomeroy,_Iowa
   [winner: 137909 Pomeroy,_Washington, pem_pos=0,
   final_scores=0.02222021, final_local_scores=0.055127144,
pem=-1.0792501, similarity=11.603947|
   gt: 112824 Pomeroy,_Iowa, pem_pos=5,
   final_scores=-0.35417482, final_local_scores=-0.32584968,
pem=-2.4096105, similarity=7.4695454 ]
11: 112823 Manson,_Iowa
   [winner: 3528860 Manson,_Washington, pem_pos=1,
   final_scores=-0.21732841, final_local_scores=-0.13790917,
pem=-1.7768338, similarity=9.636927|
   gt: 112823 Manson,_Iowa, pem_pos=4,
   final_scores=-0.5419901, final_local_scores=-0.55160826,
pem=-2.489428, similarity=1.1665187 ]
12: 6006532 Iowa_Highway_64
   [winner: 6006532 Iowa_Highway_64, pem_pos=0,

```

```

final_scores=-0.0925162, final_local_scores=-0.039258532,
pem=0.0, similarity=2.91816]
    gt: 6006532 Iowa_Highway_64, pem_pos=0,
final_scores=-0.0925162, final_local_scores=-0.039258532,
pem=0.0, similarity=2.91816 ]
13: 114010 Oakville,_Iowa
    [winner: 180181 Oakville,_Ontario, pem_pos=0,
final_scores=0.14082687, final_local_scores=0.15090212,
pem=-0.8904327, similarity=13.435245]
    gt: 114010 Oakville,_Iowa, pem_pos=7,
final_scores=-0.49611896, final_local_scores=-0.49249253,
pem=-2.5532022, similarity=3.2774642 ]
1: 114334 Fort_Dodge,_Iowa
    [pem_pos=0, final_scores=0.45722693,
final_local_scores=0.433464, pem=0.0, similarity=17.025143]
2: 112805 Storm_Lake,_Iowa
    [pem_pos=0, final_scores=0.4096514,
final_local_scores=0.40398276, pem=0.0, similarity=16.145365]
3: 114334 Fort_Dodge,_Iowa
    [pem_pos=0, final_scores=0.3904677,
final_local_scores=0.3582838, pem=0.0, similarity=14.781616]
1: 112805 Storm_Lake,_Iowa
    [pem_pos=0, final_scores=0.36761516,
final_local_scores=0.3580646, pem=0.0, similarity=14.775075]
2: 3434750 United_States
    [pem_pos=0, final_scores=0.33015466,
final_local_scores=0.33717775, pem=-0.14075686,
similarity=14.917228]
3: 26810748 Iowa
    [pem_pos=0, final_scores=0.06356089,
final_local_scores=-0.002942115, pem=-0.3546882,
similarity=5.9307637]
4: 180181 Oakville,_Ontario
    [pem_pos=0, final_scores=0.14082687,
final_local_scores=0.15090212, pem=-0.8904327,
similarity=13.435245]

```

Sample: b'Robert_Alexander'

-LRB- rugby_union_and_cricket -RRB- Captain Robert Alexander -LRB- 24 September 1910 , [1 1 Belfast 1] , [1 1 Ireland 1] -- 19 July 1943 , near [2 2 Catania 2] , [3 3 Sicily 3] -RRB- is former [2 Irish 2] 2 rugby union and 3 cricket player who represented [3 Ireland 3] at both sports during the 1930s . He also played rugby for both the [4 4 British Lions 4] and the . Alexander , an [5 5 RUC 5] officer , was killed in action during the while serving with the [6 6 Royal Inniskilling Fusiliers 6] . Between 4 1936 and 5 1939 Alexander made 11 appearances and scored 1 try for [7 7 Ireland 7] . He made his debut in a 6-3 win against [4 6 England 4] on 8 February 1936 at [8 8 Lansdowne Road 8] . On 27 February 7 1937 , again at Lansdowne , he scored his one try for [5 Ireland 5] in a 11-4 win against 8 Scotland . He made his

last appearance for Ireland on 11 March 1939 in a 7-0 defeat against 9 Wales at [9 9 Ravenhill 9] . In 1938 Alexander , together with fellow Ireland international 10 Paddy Mayne , was also a member of the [6 11 British 6] Lions squad that went on a 12 tour of [7 South Africa 7] . As a cricketer , Alexander was a right-handed batsman and a right-arm fast-medium bowler . He played once for 13 Ireland in a 14 first-class match against

1: 5046 Belfast

2: 44776 Catania

3: 27619 Sicily

4: 3914 British_and_Irish_Lions

5: 60995 Royal_Ulster_Constabulary

6: 2445066 Royal_Inniiskilling_Fusiliers

7: 955500 Ireland_national_rugby_union_team

8: 375226 Lansdowne_Road

9: 2021118 Ravenhill_Stadium

1: 21265 Northern_Ireland

[winner: 14560 Republic_of_Ireland, pem_pos=1, final_scores=0.35382563, final_local_scores=0.3222749, pem=-1.5040284, similarity=21.886179]

RECALL MISS: 21265 Northern_Ireland]

2: 25405 Rugby_union

[winner: 19774201 Rugby,_Virginia, pem_pos=7, final_scores=-0.9040964, final_local_scores=-0.79414135, pem=-2.956099, similarity=-3.5333226]

gt: 25405 Rugby_union, pem_pos=3, final_scores=-2.648316, final_local_scores=-2.9167478, pem=-2.0110612, similarity=-72.015396]

[winner: 481130 Union_(American_Civil_War), pem_pos=0, final_scores=-1.2362018, final_local_scores=-1.3136201, pem=-1.132797, similarity=-28.951002]

RECALL MISS: 25405 Rugby_union]

3: 25675557 Cricket

[winner: 330091 Rhaphidophoridae, pem_pos=10, final_scores=-1.3173035, final_local_scores=-1.2739174, pem=-2.56456, similarity=-19.98005]

gt: 25675557 Cricket, pem_pos=0, final_scores=-1.381319, final_local_scores=-1.507813, pem=-0.07532314, similarity=-40.49682]

4: 16175702 1936_Home_Nations_Championship

[winner: 34673 1936, pem_pos=0, final_scores=-0.8641409, final_local_scores=-0.7303553, pem=-0.8083407, similarity=-13.309665]

RECALL MISS: 16175702 1936_Home_Nations_Championship]

5: 16181610 1939_Home_Nations_Championship

[winner: 41239442 1939_WANFL_season, pem_pos=0, final_scores=-1.0359702, final_local_scores=-0.99242854, pem=0.0, similarity=-25.526335]

RECALL MISS: 16181610 1939_Home_Nations_Championship]

6: 691024 England_national_rugby_union_team

[winner: 759125 England_cricket_team, pem_pos=2,

final_scores=0.1299106, final_local_scores=0.13394505,
 pem=-3.3054674, similarity=26.062546|
 gt: 691024 England_national_rugby_union_team, pem_pos=3,
 final_scores=-0.118695274, final_local_scores=-0.16846669,
 pem=-3.7874734, similarity=19.6592]
 7: 16175767 1937_Home_Nations_Championship
 [winner: 34684 1937, pem_pos=0, final_scores=-0.80167377,
 final_local_scores=-0.70343876, pem=-0.6998789,
 similarity=-13.096253|
 RECALL MISS: 16175767 1937_Home_Nations_Championship]
 8: 758049 Scotland_national_rugby_union_team
 [winner: 758049 Scotland_national_rugby_union_team,
 pem_pos=2, final_scores=0.020484194,
 final_local_scores=-0.0061815977, pem=-3.3931751,
 similarity=22.357855|
 gt: 758049 Scotland_national_rugby_union_team,
 pem_pos=2, final_scores=0.020484194,
 final_local_scores=-0.0061815977, pem=-3.3931751,
 similarity=22.357855]
 9: 770930 Wales_national_rugby_union_team
 [winner: 770930 Wales_national_rugby_union_team,
 pem_pos=2, final_scores=-0.024313807,
 final_local_scores=-0.057609588, pem=-2.723462,
 similarity=17.181139|
 gt: 770930 Wales_national_rugby_union_team, pem_pos=2,
 final_scores=-0.024313807, final_local_scores=-0.057609588,
 pem=-2.723462, similarity=17.181139]
 10: 2318469 Paddy_Mayne
 [winner: 2318469 Paddy_Mayne, pem_pos=0,
 final_scores=-0.020782202, final_local_scores=0.01793696,
 pem=0.0, similarity=4.6249876|
 gt: 2318469 Paddy_Mayne, pem_pos=0,
 final_scores=-0.020782202, final_local_scores=0.01793696,
 pem=0.0, similarity=4.6249876]
 11: 3914 British_and_Irish_Lions
 [winner: 31717 United_Kingdom, pem_pos=0,
 final_scores=0.11009012, final_local_scores=0.0775381,
 pem=-0.48936236, similarity=9.064834|
 RECALL MISS: 3914 British_and_Irish_Lions]
 [winner: 36896 Lion, pem_pos=0, final_scores=-0.90017813,
 final_local_scores=-0.90816563, pem=-1.1969603,
 similarity=-16.502504|
 gt: 3914 British_and_Irish_Lions, pem_pos=3,
 final_scores=-1.1142377, final_local_scores=-1.272228,
 pem=-2.2812183, similarity=-21.47049]
 12: 20127598 1938_British_Lions_tour_to_South_Africa
 [winner: 17416221 South_Africa, pem_pos=0,
 final_scores=0.5515298, final_local_scores=0.59308183,
 pem=-0.26251382, similarity=23.21605|
 RECALL MISS: 20127598
 1938_British_Lions_tour_to_South_Africa]
 [winner: 236918 Concert, pem_pos=0,

final_scores=-1.4346807, final_local_scores=-1.5365314,
pem=-0.47050565, similarity=-39.20477|
RECALL MISS: 20127598
1938_British_Lions_tour_to_South_Africa]
13: 2217991 Ireland_cricket_team
[winner: 147575 Ireland, pem_pos=0,
final_scores=-0.052146927, final_local_scores=-0.14581239,
pem=-0.5653214, similarity=2.8126857|
gt: 2217991 Ireland_cricket_team, pem_pos=3,
final_scores=-0.080646574, final_local_scores=-0.10934788,
pem=-3.715907, similarity=21.034235]
14: 11040 First-class_cricket
[winner: 662844 First-class_citizen, pem_pos=3,
final_scores=-1.1890955, final_local_scores=-1.1059444,
pem=-1.5935546, similarity=-20.247875|
gt: 11040 First-class_cricket, pem_pos=0,
final_scores=-1.2822363, final_local_scores=-1.3925538,
pem=-0.053819276, similarity=-37.174194]
1: 5046 Belfast
[pem_pos=0, final_scores=0.3324594,
final_local_scores=0.29571617, pem=-0.13156427,
similarity=13.62994]
2: 44776 Catania
[pem_pos=0, final_scores=0.18217827,
final_local_scores=0.23905045, pem=-0.32144886,
similarity=12.971544]
3: 27619 Sicily
[pem_pos=0, final_scores=0.41063952,
final_local_scores=0.46423107, pem=-0.08753506,
similarity=18.419323]
4: 3914 British_and_Irish_Lions
[pem_pos=0, final_scores=0.14071383,
final_local_scores=0.10757619, pem=-0.11496338,
similarity=7.92519]
5: 60995 Royal_Ulster_Constabulary
[pem_pos=0, final_scores=0.1758412,
final_local_scores=0.20705023, pem=0.0, similarity=10.268505]
6: 2445066 Royal_Inniskilling_Fusiliers
[pem_pos=0, final_scores=0.15345423,
final_local_scores=0.17317566, pem=0.0, similarity=9.25762]
7: 955500 Ireland_national_rugby_union_team
[pem_pos=2, final_scores=0.033119366,
final_local_scores=0.0012691617, pem=-3.3784852,
similarity=22.500315]
8: 375226 Lansdowne_Road
[pem_pos=0, final_scores=0.34998882,
final_local_scores=0.35349753, pem=0.0, similarity=14.638784]
9: 2021118 Ravenhill_Stadium
[pem_pos=0, final_scores=0.18078826,
final_local_scores=0.18278447, pem=-0.15662159,
similarity=10.396099]
1: 14560 Republic_of_Ireland


```

    [pem_pos=1, final_scores=0.35382563,
final_local_scores=0.3222749, pem=-1.5040284,
similarity=21.886179]
2: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.5403485,
final_local_scores=0.5226633, pem=-2.2576945,
similarity=31.964722]
3: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.25874764,
final_local_scores=0.2184166, pem=-1.5040284,
similarity=18.78684]
4: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.1299106,
final_local_scores=0.13394505, pem=-3.3054674,
similarity=26.062546]
5: 2217991 Ireland_cricket_team
    [pem_pos=3, final_scores=0.058503702,
final_local_scores=0.05504954, pem=-3.715907,
similarity=25.940182]
6: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.11009012,
final_local_scores=0.0775381, pem=-0.48936236,
similarity=9.064834]
7: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5515298,
final_local_scores=0.59308183, pem=-0.26251382,
similarity=23.21605]

```

Sample: b'Exchange_Quay_Metrolink_station'

Exchange Quay Metrolink station is a stop on the Eccles Line of [1 1 Greater Manchester 1] 's light rail [2 2 Metrolink 2] system . It is located in the [3 3 Salford Quays 3] area , in [4 4 North West England 4] , and opened on 12 June 1999 as part of Phase 2 of the system 's expansion . Exchange Quay is located on the Eccles Line with trams towards 1 Eccles or [1 2 MediaCityUK 1] stopping every 12 minutes during the day Monday-Saturday , with trams running to [5 5 Manchester Piccadilly 5] via 3 Cornbrook every 6 minutes . Trams run every 12 minutes Monday-Saturday evenings and Sunday daytime and every 15 minutes Sunday evenings between [2 Manchester Piccadilly 2] and Eccles , which calls at [3 MediaCityUK 3] . Exchange Quay station is served by [6 6 Arriva North West 6] services 71 and 73 , which stops outside the station on [4 Ordsall Lane 4] . The service runs to [5 Clifton 5] via 4 Pendleton and [7 7 Swinton 7] and to Manchester 5 Shudehill Interchange . [6 Arriva 6] service 70 also serve the station on Sundays , which also runs to [7 Manchester 7] and to [8 Clifton 8] , although via [8 8 Eccles 8] . [9 Arriva 9] service 245 starts /ends its journeys at Exchange Quay and runs throughout the day Monday-Saturdays to [9 9 Altrincham 9] via [10 10 Trafford Centre 10] , [11 11 Flixton 11] , [12 12 Urmston 12] , [13 13 Stretford 13] and [14 14 Sale 14] The station is also served

by services stopping on Trafford Road . [15 15 First Greater Manchester 15] service 53 , which runs to Pendleton and to [10 Cheetham Hill 10] via [11 Rusholme 11] , [12 Gorton 12] and [13 Harpurhey 13] , while [14 Arriva 14] service 79 runs to [16 Swinton 16] via Ordsall and Pendleton and to [15 Stretford 15] via [16 6 Old Trafford 16]

1: 79819 Greater_Manchester
2: 253314 Manchester_Metrolink
3: 749917 Salford_Quays
4: 53214 North_West_England
5: 209875 Manchester_Piccadilly_station
6: 14300098 Arriva_North_West
7: 1806065 Swinton,_Greater_Manchester
8: 1742688 Eccles,_Greater_Manchester
9: 285284 Altrincham
10: 849601 Trafford_Centre
11: 3795509 Flixton,_Greater_Manchester
12: 1445300 Urmston
13: 1074030 Stretford
14: 367154 Sale,_Greater_Manchester
15: 4630689 First_Greater_Manchester
16: 1806065 Swinton,_Greater_Manchester

1: 9561941 Eccles_Interchange

[winner: 1742688 Eccles,_Greater_Manchester, pem_pos=0, final_scores=-0.02431272, final_local_scores=-0.10570538, pem=-1.037052, similarity=6.574904]

gt: 9561941 Eccles_Interchange, pem_pos=6, final_scores=-0.44767076, final_local_scores=-0.47367972, pem=-2.5348885, similarity=3.739284]

2: 16751797 MediaCityUK_Metrolink_station

[winner: 9409970 MediaCityUK, pem_pos=0, final_scores=0.037806198, final_local_scores=0.013049886, pem=-0.2876821, similarity=6.043607]

gt: 16751797 MediaCityUK_Metrolink_station, pem_pos=1, final_scores=-0.1977432, final_local_scores=-0.20850135, pem=-1.3862944, similarity=5.406501]

3: 9315155 Cornbrook_Metrolink_station

[winner: 9315155 Cornbrook_Metrolink_station, pem_pos=0, final_scores=-0.044866174, final_local_scores=-0.03318508, pem=-0.5567546, similarity=6.127122]

gt: 9315155 Cornbrook_Metrolink_station, pem_pos=0, final_scores=-0.044866174, final_local_scores=-0.03318508, pem=-0.5567546, similarity=6.127122]

4: 3886351 Pendleton,_Greater_Manchester

[winner: 130965 Pendleton,_Oregon, pem_pos=0, final_scores=-0.06554538, final_local_scores=-0.053765625, pem=-1.4213227, similarity=10.214613]

gt: 3886351 Pendleton,_Greater_Manchester, pem_pos=6, final_scores=-0.22732826, final_local_scores=-0.29909292, pem=-2.677043, similarity=9.722359]

5: 9316303 Shudehill_Interchange

[winner: 9316303 Shudehill_Interchange, pem_pos=0,

```

final_scores=-0.3248399, final_local_scores=-0.35848415,
pem=0.0, similarity=-6.6081686]
    gt: 9316303 Shudehill_Interchange, pem_pos=0,
final_scores=-0.3248399, final_local_scores=-0.35848415,
pem=0.0, similarity=-6.6081686 ]
6: 2428702 Old_Trafford, Greater_Manchester
    [winner: 660597 Old_Trafford_Cricket_Ground, pem_pos=1,
final_scores=0.06931238, final_local_scores=0.103358,
pem=-0.6437659, similarity=10.675021]
    gt: 2428702 Old_Trafford, Greater_Manchester, pem_pos=2,
final_scores=-0.10969062, final_local_scores=-0.16215828,
pem=-0.97548914, similarity=4.555448 ]
1: 79819 Greater_Manchester
    [pem_pos=0, final_scores=0.4004032,
final_local_scores=0.355426, pem=0.0, similarity=14.696334]
2: 253314 Manchester_Metrolink
    [pem_pos=1, final_scores=0.0802149,
final_local_scores=0.055362314, pem=-0.6763539,
similarity=9.419952]
3: 749917 Salford_Quays
    [pem_pos=0, final_scores=0.17897551,
final_local_scores=0.15667287, pem=0.0, similarity=8.765143]
4: 53214 North_West_England
    [pem_pos=0, final_scores=0.6214016,
final_local_scores=0.60943675, pem=0.0, similarity=22.276522]
5: 209875 Manchester_Piccadilly_station
    [pem_pos=0, final_scores=0.22672243,
final_local_scores=0.20207447, pem=0.0, similarity=10.120018]
6: 14300098 Arriva_North_West
    [pem_pos=0, final_scores=0.13805608,
final_local_scores=0.15006274, pem=0.0, similarity=8.567884]
7: 1806065 Swinton, Greater_Manchester
    [pem_pos=1, final_scores=0.13283534,
final_local_scores=0.07993828, pem=-1.2937785,
similarity=13.511]
8: 1742688 Eccles, Greater_Manchester
    [pem_pos=0, final_scores=0.07653724,
final_local_scores=0.007693127, pem=-1.037052,
similarity=9.958941]
9: 285284 Altrincham
    [pem_pos=0, final_scores=0.24512272,
final_local_scores=0.1941491, pem=-0.024761505,
similarity=10.018167]
10: 849601 Trafford_Centre
    [pem_pos=0, final_scores=0.31176394,
final_local_scores=0.29005104, pem=0.0, similarity=12.745415]
11: 3795509 Flixton, Greater_Manchester
    [pem_pos=0, final_scores=0.16770853,
final_local_scores=0.12724024, pem=-0.7864005,
similarity=12.163384]
12: 1445300 Urmston
    [pem_pos=0, final_scores=0.32901102,

```

final_local_scores=0.29949948, pem=0.0, similarity=13.027375]
13: 1074030 Stretford
[pem_pos=0, final_scores=0.23008491,
final_local_scores=0.18214539, pem=0.0, similarity=9.525294]
14: 367154 Sale, Greater_Manchester
[pem_pos=1, final_scores=0.049862936,
final_local_scores=-0.016679987, pem=-1.3014129,
similarity=10.669235]
15: 4630689 First_Greater_Manchester
[pem_pos=0, final_scores=0.143079,
final_local_scores=0.1289036, pem=0.0, similarity=7.936454]
16: 1806065 Swinton, Greater_Manchester
[pem_pos=1, final_scores=0.13932772,
final_local_scores=0.08703022, pem=-1.2937785,
similarity=13.722637]
1: 9409970 MediaCityUK
[pem_pos=0, final_scores=0.037806198,
final_local_scores=0.013049886, pem=-0.2876821,
similarity=6.043607]
2: 209875 Manchester_Piccadilly_station
[pem_pos=0, final_scores=0.17603181,
final_local_scores=0.1452403, pem=0.0, similarity=8.423973]
3: 9409970 MediaCityUK
[pem_pos=0, final_scores=0.16088395,
final_local_scores=0.14749363, pem=-0.2876821,
similarity=10.055677]
4: 9383681 Ordsall, Greater_Manchester
[pem_pos=0, final_scores=0.24462013,
final_local_scores=0.2590668, pem=0.0, similarity=11.820783]
5: 656995 Clifton, Bristol
[pem_pos=0, final_scores=0.1403264,
final_local_scores=0.11523619, pem=-2.1126766,
similarity=19.017654]
6: 171766 Arriva
[pem_pos=0, final_scores=0.23806016,
final_local_scores=0.26452166, pem=-0.032858714,
similarity=12.162257]
7: 20206 Manchester
[pem_pos=0, final_scores=0.409993,
final_local_scores=0.38056576, pem=-0.3209542,
similarity=17.191954]
8: 656995 Clifton, Bristol
[pem_pos=0, final_scores=0.12605868,
final_local_scores=0.09965089, pem=-2.1126766,
similarity=18.552557]
9: 171766 Arriva
[pem_pos=0, final_scores=0.28556758,
final_local_scores=0.31641632, pem=-0.032858714,
similarity=13.710897]
10: 2344566 Cheetham_Hill
[pem_pos=0, final_scores=0.20630716,
final_local_scores=0.18079758, pem=0.0, similarity=9.485073]

11: 439591 Rusholme
 [pem_pos=0, final_scores=0.23723419,
 final_local_scores=0.21246079, pem=0.0, similarity=10.429967]
 12: 1414171 Gorton
 [pem_pos=0, final_scores=0.16408817,
 final_local_scores=0.12096405, pem=-0.59120226,
 similarity=10.914571]
 13: 1179753 Harpurhey
 [pem_pos=0, final_scores=0.23344873,
 final_local_scores=0.21984783, pem=0.0, similarity=10.650411]
 14: 171766 Arriva
 [pem_pos=0, final_scores=0.27891362,
 final_local_scores=0.30914786, pem=-0.032858714,
 similarity=13.493992]
 15: 1074030 Stretford
 [pem_pos=0, final_scores=0.18522166,
 final_local_scores=0.13313907, pem=0.0, similarity=8.062848]
 16: 660597 Old_Trafford_Cricket_Ground
 [pem_pos=1, final_scores=0.06931238,
 final_local_scores=0.103358, pem=-0.6437659,
 similarity=10.675021]

Sample: b'Nebraska_Highway_11'

Nebraska Highway 11 is a highway in central and northern [1 1
 Nebraska 1] . Its southern terminus is at 1 Interstate 80
 south of [1 2 Wood River 1] . Its northern terminus is at the
 [2 2 South Dakota 2] border north of [2 3 Butte 2] . The
 highway goes through the eastern portion of the . Nebraska
 Highway 11 begins at Exit 300 of 4 Interstate 80 , which is
 also the exit for NE Spur 40D . It goes north through
 farmland into [3 5 Wood River 3] , where it meets . It
 continues north through [4 6 Cairo 4] , then meets 7 Nebraska
 Highway 58 . The two highways are paired together and go
 northeast to [5 8 Dannebrog 5] , where they separate . It
 proceeds north to 9 Elba , then turns northwest on an
 alignment that parallels the [6 North 10 Loup River 6] . It
 passes through 11 Ord and 12 Burwell , then turns north after
 passing 13 Nebraska Highway 91 . It continues north into
 areas of prairie and meets in 14 Atkinson . It goes north
 into farmland some more , briefly goes east , then turns
 north to go through [7 15 Butte 7] , where it meets 16 Nebras
 ka Highway 12 . After a brief concurrency with NE 12 , it
 turns north and goes into [3 3 South Dakota 3] , where the
 road continues as [8 17 South Dakota 8] Highway 43

1: 21647 Nebraska

2: 26746 South_Dakota

3: 26746 South_Dakota

1: 84143 Interstate_80

[winner: 84143 Interstate_80, pem_pos=0,
 final_scores=-0.07054566, final_local_scores=-0.072006606,
 pem=-0.33467698, similarity=3.7609177]

gt: 84143 Interstate_80, pem_pos=0,

```

final_scores=-0.07054566, final_local_scores=-0.072006606,
pem=-0.33467698, similarity=3.7609177 ]
2: 123948 Wood_River,_Nebraska
  [winner: 111606 Wood_River,_Illinois, pem_pos=0,
final_scores=0.24400331, final_local_scores=0.2768749,
pem=-0.7634926, similarity=16.504204]
  gt: 123948 Wood_River,_Nebraska, pem_pos=5,
final_scores=-0.1925593, final_local_scores=-0.16691336,
pem=-2.2031686, similarity=11.089859 ]
3: 123692 Butte,_Nebraska
  [winner: 257453 Butte,_Montana, pem_pos=0,
final_scores=0.11968784, final_local_scores=0.081992224,
pem=-0.6054963, similarity=9.829308]
  gt: 123692 Butte,_Nebraska, pem_pos=5,
final_scores=-0.32714534, final_local_scores=-0.31475967,
pem=-2.3077676, similarity=7.246656 ]
4: 84143 Interstate_80
  [winner: 84143 Interstate_80, pem_pos=0,
final_scores=-0.16337208, final_local_scores=-0.17340538,
pem=-0.33467698, similarity=0.7349761]
  gt: 84143 Interstate_80, pem_pos=0,
final_scores=-0.16337208, final_local_scores=-0.17340538,
pem=-0.33467698, similarity=0.7349761 ]
5: 123948 Wood_River,_Nebraska
  [winner: 111606 Wood_River,_Illinois, pem_pos=0,
final_scores=0.32072097, final_local_scores=0.3606773,
pem=-0.7634926, similarity=19.005033]
  gt: 123948 Wood_River,_Nebraska, pem_pos=5,
final_scores=-0.12951374, final_local_scores=-0.09804565,
pem=-2.2031686, similarity=13.145009 ]
6: 123944 Cairo,_Nebraska
  [winner: 6293 Cairo, pem_pos=0, final_scores=0.3095355,
final_local_scores=0.3556578, pem=-0.15211283,
similarity=15.530463]
  gt: 123944 Cairo,_Nebraska, pem_pos=15,
final_scores=-0.79370016, final_local_scores=-0.7717682,
pem=-3.6960857, similarity=1.1584997 ]
7: 16050775 Nebraska_Highway_58
  [winner: 16050775 Nebraska_Highway_58, pem_pos=0,
final_scores=-0.45910904, final_local_scores=-0.35007367,
pem=0.0, similarity=-6.3571835]
  gt: 16050775 Nebraska_Highway_58, pem_pos=0,
final_scores=-0.45910904, final_local_scores=-0.35007367,
pem=0.0, similarity=-6.3571835 ]
8: 123989 Dannebrog,_Nebraska
  [winner: 8752 Flag_of_Denmark, pem_pos=0,
final_scores=0.07340483, final_local_scores=0.15573972,
pem=-0.019684019, similarity=8.844341]
  gt: 123989 Dannebrog,_Nebraska, pem_pos=1,
final_scores=-0.20732906, final_local_scores=-0.16109002,
pem=-0.99792373, similarity=4.70933 ]
9: 123990 Elba,_Nebraska

```

[winner: 10479 Elba, pem_pos=0, final_scores=0.020782933, final_local_scores=0.073048085, pem=-0.29250675, similarity=7.8603106|
gt: 123990 Elba, Nebraska, pem_pos=7, final_scores=-0.5309024, final_local_scores=-0.508899, pem=-2.6251237, similarity=3.1789837]
10: 689531 Loup_River
[winner: 689531 Loup_River, pem_pos=0, final_scores=0.29739755, final_local_scores=0.37183118, pem=0.0, similarity=15.185898|
gt: 689531 Loup_River, pem_pos=0, final_scores=0.29739755, final_local_scores=0.37183118, pem=0.0, similarity=15.185898]
11: 124285 Ord, Nebraska
[winner: 124285 Ord, Nebraska, pem_pos=1, final_scores=-0.091219455, final_local_scores=-0.062862635, pem=-1.6162307, similarity=11.00308|
gt: 124285 Ord, Nebraska, pem_pos=1, final_scores=-0.091219455, final_local_scores=-0.062862635, pem=-1.6162307, similarity=11.00308]
12: 123933 Burwell, Nebraska
[winner: 3417105 Mahlon_Burwell, pem_pos=1, final_scores=-0.05293934, final_local_scores=0.004842192, pem=-1.2307835, similarity=10.927405|
gt: 123933 Burwell, Nebraska, pem_pos=2, final_scores=-0.3264016, final_local_scores=-0.30943587, pem=-1.8014464, similarity=4.6520753]
13: 16805998 Nebraska_Highway_91
[winner: 16805998 Nebraska_Highway_91, pem_pos=0, final_scores=-0.22548388, final_local_scores=-0.11657646, pem=0.0, similarity=0.6108387|
gt: 16805998 Nebraska_Highway_91, pem_pos=0, final_scores=-0.22548388, final_local_scores=-0.11657646, pem=0.0, similarity=0.6108387]
14: 123968 Atkinson, Nebraska
[winner: 124878 Atkinson, New_Hampshire, pem_pos=0, final_scores=-0.14013244, final_local_scores=-0.119205415, pem=-1.6299168, similarity=9.396127|
gt: 123968 Atkinson, Nebraska, pem_pos=8, final_scores=-0.5087864, final_local_scores=-0.5042804, pem=-3.345812, similarity=7.2360253]
15: 123692 Butte, Nebraska
[winner: 257453 Butte, Montana, pem_pos=0, final_scores=0.17878078, final_local_scores=0.14654225, pem=-0.6054963, similarity=11.7556095|
gt: 123692 Butte, Nebraska, pem_pos=5, final_scores=-0.3160579, final_local_scores=-0.3026483, pem=-2.3077676, similarity=7.6080823]
16: 11319650 Nebraska_Highway_12
[winner: 11319650 Nebraska_Highway_12, pem_pos=0, final_scores=-0.37338406, final_local_scores=-0.28898436, pem=0.0, similarity=-4.534157|

gt: 11319650 Nebraska_Highway_12, pem_pos=0,
final_scores=-0.37338406, final_local_scores=-0.28898436,
pem=0.0, similarity=-4.534157]
17: 10290230 South_Dakota_Highway_43
[winner: 26746 South_Dakota, pem_pos=0,
final_scores=0.10380821, final_local_scores=0.037855953,
pem=-0.23758878, similarity=6.5114546]
RECALL MISS: 10290230 South_Dakota_Highway_43]
[winner: 48519 Highway, pem_pos=0,
final_scores=-0.5796534, final_local_scores=-0.59491146,
pem=-0.30968767, similarity=-11.979503]
RECALL MISS: 10290230 South_Dakota_Highway_43]
[winner: 11421666 Small_nucleolar_RNA_R43, pem_pos=8,
final_scores=-1.1744341, final_local_scores=-1.0622411,
pem=-3.4938104, similarity=-8.6097975]
RECALL MISS: 10290230 South_Dakota_Highway_43]
1: 21647 Nebraska
[pem_pos=0, final_scores=0.47650164,
final_local_scores=0.44662976, pem=-0.4226844,
similarity=19.71666]
2: 26746 South_Dakota
[pem_pos=0, final_scores=0.43024015,
final_local_scores=0.39443326, pem=-0.23758878,
similarity=17.152433]
3: 26746 South_Dakota
[pem_pos=0, final_scores=0.41410094,
final_local_scores=0.37680358, pem=-0.23758878,
similarity=16.62633]
1: 111606 Wood_River,_Illinois
[pem_pos=0, final_scores=0.24400331,
final_local_scores=0.2768749, pem=-0.7634926,
similarity=16.504204]
2: 257453 Butte,_Montana
[pem_pos=0, final_scores=0.11968784,
final_local_scores=0.081992224, pem=-0.6054963,
similarity=9.829308]
3: 111606 Wood_River,_Illinois
[pem_pos=0, final_scores=0.32072097,
final_local_scores=0.3606773, pem=-0.7634926,
similarity=19.005033]
4: 6293 Cairo
[pem_pos=0, final_scores=0.3095355,
final_local_scores=0.3556578, pem=-0.15211283,
similarity=15.530463]
5: 8752 Flag_of_Denmark
[pem_pos=0, final_scores=0.07340483,
final_local_scores=0.15573972, pem=-0.019684019,
similarity=8.844341]
6: 689531 Loup_River
[pem_pos=0, final_scores=0.29739755,
final_local_scores=0.37183118, pem=0.0, similarity=15.185898]
7: 257453 Butte,_Montana


```
[pem_pos=0, final_scores=0.17878078,  
final_local_scores=0.14654225, pem=-0.6054963,  
similarity=11.7556095]  
8: 26746 South_Dakota  
[pem_pos=0, final_scores=0.10380821,  
final_local_scores=0.037855953, pem=-0.23758878,  
similarity=6.5114546]
```

Sample: b'Gmina_Jaraczewo'

Gmina Jaraczewo is a rural 1 gmina -LRB- administrative district -RRB- in 2 Jarocin County , [1 1 Greater Poland Voivodeship 1] , in west-central [2 2 Poland 2] . Its seat is the village of 3 Jaraczewo , which lies approximately west of [3 3 Jarocin 3] and south-east of the regional capital [4 4 Poznań 4] . The gmina covers an area of , and as of 2006 its total population is 8,281 . Gmina Jaraczewo contains the villages and settlements of 4 Bielejewo , 5 Brzostów , 6 Cerekwica , 7 Gola , 8 Góra , 9 Jaraczewo , 10 Łobez , 11 Łobzowiec , 12 Łowędice , 13 Łukaszewo , 14 Niedźwiady , 15 Nosków , 16 Nowa Cerekwica , 17 Panienka , 18 Parzęczew , 19 Poręba , 20 Rusko , 21 Strzyżewko , 22 Suchorzewko , 23 Wojciechowo and 24 Zalesie . Gmina Jaraczewo is bordered by the gminas of 25 Borek Wielkopolski , 26 Dolsk , 27 Jarocin , [1 28 Koźmin Wielkopolski 1] , 29 Książ Wielkopolski and 30 Nowe Miasto nad Wartą

1: 41498 Greater_Poland_Voivodeship

2: 22936 Poland

3: 390510 Jarocin

4: 23723 Poznań

1: 1234326 Gmina

```
[winner: 1234326 Gmina, pem_pos=0,  
final_scores=-0.49868834, final_local_scores=-0.56528103,  
pem=0.0, similarity=-12.779402]
```

```
gt: 1234326 Gmina, pem_pos=0, final_scores=-0.49868834,  
final_local_scores=-0.56528103, pem=0.0,  
similarity=-12.779402 ]
```

2: 5776135 Jarocin_County

```
[winner: 5776135 Jarocin_County, pem_pos=0,  
final_scores=-0.042884216, final_local_scores=-0.07296013,  
pem=0.0, similarity=1.9124371]
```

```
gt: 5776135 Jarocin_County, pem_pos=0,  
final_scores=-0.042884216, final_local_scores=-0.07296013,  
pem=0.0, similarity=1.9124371 ]
```

3: 20978373 Jaraczewo,_Jarocin_County

```
[winner: 15114655 Gmina_Jaraczewo, pem_pos=1,  
final_scores=-0.0207863, final_local_scores=-0.018763289,  
pem=-0.5070015, similarity=6.286932]
```

```
gt: 20978373 Jaraczewo,_Jarocin_County, pem_pos=0,  
final_scores=-0.024955168, final_local_scores=-0.04339412,  
pem=-0.19481303, similarity=3.8541694 ]
```

4: 20978338 Bielejewo,_Jarocin_County

```
[winner: 20978338 Bielejewo,_Jarocin_County, pem_pos=0,  
final_scores=-0.04834716, final_local_scores=-0.023589551,
```

pem=0.0, similarity=3.3857536|
gt: 20978338 Bielejewo, Jarocin County, pem_pos=0,
final_scores=-0.04834716, final_local_scores=-0.023589551,
pem=0.0, similarity=3.3857536]
5: 20978341 Brzostów
[winner: 20978341 Brzostów, pem_pos=0,
final_scores=-0.35525882, final_local_scores=-0.33600104,
pem=0.0, similarity=-5.9372277|
gt: 20978341 Brzostów, pem_pos=0,
final_scores=-0.35525882, final_local_scores=-0.33600104,
pem=0.0, similarity=-5.9372277]
6: 20978351 Cerekwica, Jarocin County
[winner: 14763382 Cerekwica, Lower Silesian Voivodeship,
pem_pos=1, final_scores=-0.34156138,
final_local_scores=-0.32189277, pem=-0.57187617,
similarity=-2.4062579|
gt: 20978351 Cerekwica, Jarocin County, pem_pos=2,
final_scores=-0.5377034, final_local_scores=-0.60043633,
pem=-0.9297456, similarity=-8.772405]
7: 20978354 Gola, Jarocin County
[winner: 2808335 Gola_people, pem_pos=0,
final_scores=-0.47784436, final_local_scores=-0.3565219,
pem=-1.0080004, similarity=-1.0679492|
gt: 20978354 Gola, Jarocin County, pem_pos=11,
final_scores=-0.85519505, final_local_scores=-0.910803,
pem=-2.6825976, similarity=-8.502084]
8: 20978367 Góra, Jarocin County
[winner: 664954 Góra, pem_pos=0,
final_scores=-0.35373378, final_local_scores=-0.39554492,
pem=-0.7992108, similarity=-3.367904|
gt: 20978367 Góra, Jarocin County, pem_pos=17,
final_scores=-0.9925211, final_local_scores=-1.103698,
pem=-4.325571, similarity=-5.323711]
9: 20978373 Jaraczewo, Jarocin County
[winner: 15114655 Gmina Jaraczewo, pem_pos=1,
final_scores=-0.25777176, final_local_scores=-0.2776339,
pem=-0.5070015, similarity=-1.4382837|
gt: 20978373 Jaraczewo, Jarocin County, pem_pos=0,
final_scores=-0.30631816, final_local_scores=-0.35074046,
pem=-0.19481303, similarity=-5.3176584]
10: 20978384 Łobez, Greater Poland Voivodeship
[winner: 4366789 Łobez, pem_pos=0,
final_scores=-0.25705463, final_local_scores=-0.20991424,
pem=-0.14767453, similarity=-1.3714696|
gt: 20978384 Łobez, Greater Poland Voivodeship,
pem_pos=2, final_scores=-0.78525484,
final_local_scores=-0.86836857, pem=-3.2446284,
similarity=-4.179343]
11: 20978391 Łobzowiec
[winner: 20978391 Łobzowiec, pem_pos=0,
final_scores=-0.25877562, final_local_scores=-0.23042823,
pem=0.0, similarity=-2.7867253|

gt: 20978391 Łobzowiec, pem_pos=0,
final_scores=-0.25877562, final_local_scores=-0.23042823,
pem=0.0, similarity=-2.7867253]
12: 20978412 Łowęcice, Greater_Poland_Voivodeship
[winner: 20978412 Łowęcice, Greater_Poland_Voivodeship,
pem_pos=1, final_scores=-0.34898,
final_local_scores=-0.40763566, pem=-0.8534898,
similarity=-3.433538]
gt: 20978412 Łowęcice, Greater_Poland_Voivodeship,
pem_pos=1, final_scores=-0.34898,
final_local_scores=-0.40763566, pem=-0.8534898,
similarity=-3.433538]
13: 20978443 Łukaszewo, Greater_Poland_Voivodeship
[winner: 18205496 Łukaszewo, Inowrocław_County,
pem_pos=1, final_scores=-0.22649278,
final_local_scores=-0.19830917, pem=-1.0414538,
similarity=3.8353596]
gt: 20978443 Łukaszewo, Greater_Poland_Voivodeship,
pem_pos=2, final_scores=-0.4146508,
final_local_scores=-0.4316629, pem=-1.307157,
similarity=-1.6834455]
14: 20978460 Niedźwiady, Jarocin_County
[winner: 22000123
Niedźwiady, West_Pomeranian_Voivodeship, pem_pos=0,
final_scores=-0.29390174, final_local_scores=-0.28271106,
pem=-1.6647006, similarity=4.7059517]
gt: 20978460 Niedźwiady, Jarocin_County, pem_pos=9,
final_scores=-0.73250324, final_local_scores=-0.7586553,
pem=-3.0065432, similarity=-2.200026]
15: 20978471 Nosków, Greater_Poland_Voivodeship
[winner: 20978471 Nosków, Greater_Poland_Voivodeship,
pem_pos=0, final_scores=-0.19541267,
final_local_scores=-0.2349883, pem=0.0,
similarity=-2.9228067]
gt: 20978471 Nosków, Greater_Poland_Voivodeship,
pem_pos=0, final_scores=-0.19541267,
final_local_scores=-0.2349883, pem=0.0,
similarity=-2.9228067]
16: 20978475 Nowa_Cerekwica
[winner: 20978475 Nowa_Cerekwica, pem_pos=0,
final_scores=0.011500284, final_local_scores=0.018061385,
pem=0.0, similarity=4.6287007]
gt: 20978475 Nowa_Cerekwica, pem_pos=0,
final_scores=0.011500284, final_local_scores=0.018061385,
pem=0.0, similarity=4.6287007]
17: 20978479 Panienka
[winner: 20978479 Panienka, pem_pos=0,
final_scores=-0.10843615, final_local_scores=-0.06027259,
pem=0.0, similarity=2.2910585]
gt: 20978479 Panienka, pem_pos=0,
final_scores=-0.10843615, final_local_scores=-0.06027259,
pem=0.0, similarity=2.2910585]

18: 20978489 Parzęczew, Greater Poland Voivodeship
[winner: 14932685 Gmina Parzęczew, pem_pos=1,
final_scores=-0.2843827, final_local_scores=-0.2701946,
pem=-1.1183174, similarity=2.108151|
gt: 20978489 Parzęczew, Greater Poland Voivodeship,
pem_pos=2, final_scores=-0.45398223,
final_local_scores=-0.5203245, pem=-2.1875157,
similarity=0.45824188]

19: 20978519 Poręba, Greater Poland Voivodeship
[winner: 9615612 Poręba, pem_pos=0,
final_scores=-0.1358199, final_local_scores=-0.0873667,
pem=-0.61579275, similarity=4.831293|
gt: 20978519 Poręba, Greater Poland Voivodeship,
pem_pos=7, final_scores=-0.88987136,
final_local_scores=-0.95120317, pem=-3.3552294,
similarity=-6.0498276]

20: 20978532 Rusko, Greater Poland Voivodeship
[winner: 21998527 Rusko, West Pomeranian Voivodeship,
pem_pos=2, final_scores=-0.24500547,
final_local_scores=-0.26162824, pem=-1.7454369,
similarity=5.7741613|
gt: 20978532 Rusko, Greater Poland Voivodeship,
pem_pos=4, final_scores=-0.4455539,
final_local_scores=-0.4991257, pem=-1.8754703,
similarity=-0.606094]

21: 20978537 Strzyżewko
[winner: 20978537 Strzyżewko, pem_pos=0,
final_scores=-0.21498337, final_local_scores=-0.17244822,
pem=0.0, similarity=-1.0564861|
gt: 20978537 Strzyżewko, pem_pos=0,
final_scores=-0.21498337, final_local_scores=-0.17244822,
pem=0.0, similarity=-1.0564861]

22: 20978539 Suchorzewko
[winner: 20978539 Suchorzewko, pem_pos=0,
final_scores=-0.18793927, final_local_scores=-0.16833583,
pem=0.0, similarity=-0.9337636|
gt: 20978539 Suchorzewko, pem_pos=0,
final_scores=-0.18793927, final_local_scores=-0.16833583,
pem=0.0, similarity=-0.9337636]

23: 20978549 Wojciechowo, Jarocin County
[winner: 21822672 Wojciechowo, Warmian-
Masurian Voivodeship, pem_pos=6, final_scores=-0.3283723,
final_local_scores=-0.28904206, pem=-1.6777875,
similarity=4.5881906|
gt: 20978549 Wojciechowo, Jarocin County, pem_pos=8,
final_scores=-0.55353, final_local_scores=-0.613208,
pem=-1.7528853, similarity=-4.677174]

24: 20978566 Zalesie, Jarocin County
[winner: 18325532 Zalesie, Nakło County, pem_pos=8,
final_scores=-0.6929758, final_local_scores=-0.720862,
pem=-3.702426, similarity=2.7121212|
RECALL MISS: 20978566 Zalesie, Jarocin County]

25: 15114444 Gmina_Borek_Wielkopolski
 [winner: 9317368 Borek_Wielkopolski, pem_pos=0,
 final_scores=0.02370514, final_local_scores=0.0649115,
 pem=0.0, similarity=6.0268016|
 gt: 15114444 Gmina_Borek_Wielkopolski, pem_pos=1,
 final_scores=-0.021853745, final_local_scores=-0.017502576,
 pem=-0.37906396, similarity=5.62881]

26: 15128719 Gmina_Dolsk
 [winner: 21935378 Dolsk, West_Pomeranian_Voivodeship,
 pem_pos=2, final_scores=-0.3001506,
 final_local_scores=-0.31632042, pem=-1.2182056,
 similarity=1.2748722|
 gt: 15128719 Gmina_Dolsk, pem_pos=1,
 final_scores=-0.3908774, final_local_scores=-0.39049447,
 pem=-1.1436913, similarity=-1.3438494]

27: 15114623 Gmina_Jarocin, Greater_Poland_Voivodeship
 [winner: 390510 Jarocin, pem_pos=0,
 final_scores=-0.047523916, final_local_scores=-0.1098055,
 pem=-0.1686428, similarity=1.7300035|
 gt: 15114623 Gmina_Jarocin, Greater_Poland_Voivodeship,
 pem_pos=4, final_scores=-0.36289847,
 final_local_scores=-0.39263266, pem=-1.8342848,
 similarity=2.3478968]

28: 15116016 Gmina_Koźmin_Wielkopolski
 [winner: 9424431 Koźmin_Wielkopolski, pem_pos=0,
 final_scores=0.07402466, final_local_scores=0.11788127,
 pem=-0.28636542, similarity=9.164825|
 gt: 15116016 Gmina_Koźmin_Wielkopolski, pem_pos=1,
 final_scores=-0.24062811, final_local_scores=-0.26976904,
 pem=-1.3902547, similarity=3.5996876]

29: 15128730 Gmina_Książ_Wielkopolski
 [winner: 9317694 Książ_Wielkopolski, pem_pos=0,
 final_scores=-0.022114426, final_local_scores=-0.0050605386,
 pem=-0.18781607, similarity=4.960069|
 gt: 15128730 Gmina_Książ_Wielkopolski, pem_pos=1,
 final_scores=-0.2385705, final_local_scores=-0.2522052,
 pem=-1.7647308, similarity=6.1602874]

30: 15119130 Gmina_Nowe_Miasto_nad_Wartą
 [winner: 21157765 Nowe_Miasto_nad_Wartą, pem_pos=0,
 final_scores=-0.36241835, final_local_scores=-0.38036695,
 pem=-0.2519385, similarity=-5.8911147|
 gt: 15119130 Gmina_Nowe_Miasto_nad_Wartą, pem_pos=1,
 final_scores=-0.5538308, final_local_scores=-0.59862596,
 pem=-1.5018964, similarity=-5.606934]

1: 41498 Greater_Poland_Voivodeship
 [pem_pos=0, final_scores=0.070929274,
 final_local_scores=0.075406015, pem=0.0, similarity=6.3399787]

2: 22936 Poland
 [pem_pos=0, final_scores=0.29863793,
 final_local_scores=0.33949402, pem=-0.15333635,
 similarity=15.054758]

3: 390510 Jarocin

[pem_pos=0, final_scores=0.20754834,
final_local_scores=0.19567448, pem=-0.1686428,
similarity=10.846136]
4: 23723 Poznań
[pem_pos=0, final_scores=0.36644644,
final_local_scores=0.37861425, pem=-0.056552384,
similarity=15.695859]
1: 9424431 Koźmin Wielkopolski
[pem_pos=0, final_scores=0.07402466,
final_local_scores=0.11788127, pem=-0.28636542,
similarity=9.164825]

Sample: b'1966'

-- 68_Liga_Leumit The 1966 -- 68 [1 1 Liga Leumit 1] season was the thirteenth in the league 's history , and is notable for the [2 2 Israel Football Association 2] 's decision to play it over two years as a play to combat corruption and increasing violence at matches . The sixteen teams played each other four times during the season , effectively combining two seasons into one , leading it to be known as the double season -LRB- , HaOna HaKfula -RRB- . At the end of the season [3 3 SK Nes Tziona 3] and [4 4 Hapoel Mahane Yehuda 4] were relegated to [5 5 Liga Alef 5] . Neither team has returned to the top flight since . They were replaced by [6 6 Hapoel Kfar Saba 6] and [7 7 Beitar Jerusalem 7] . As champions , [8 8 Maccabi Tel Aviv 8] entered the , which they won . [1 1 Maccabi Netanya 1] 's [9 9 Mordechai Spiegler 9] was the season 's top scorer with 38 goals - 15 during 1966 -- 67 and 23 in 1967 -- 68 . The season began with a protests from relegated players , with the uproar reaching as far as the [10 10 Knesset 10] . In an attempt to restore order to the game and solve the issues raised , the IFA decided to spread the league games over two years instead of one . The main objectives were to put an end to the riots on the field , reducing trouble at matches and improve the quality of play , as well as infusing new blood into teams by alleviating the immediate fear of relegation . The double season format involved all the 2 Maccabi -LRB- 3 Haifa , [11 1 1 Maccabi Jaffa 11] , [12 12 Netanya 12] , 4 Sha'arayim and [2 5 Tel Aviv 2] -RRB- and [3 6 Hapoel 3] -LRB- 7 Be'er Sheva , [4 8 Haifa 4] , 9 Jerusalem , 10 Mahane Yehuda , [13 13 Petah T ikva 13] , and [14 14 Tel Aviv 14]

1: 5695086 Liga_Leumit
2: 1265525 Israel_Football_Association
3: 6059253 Sektzia_Nes_Tziona_F.C.
4: 19444834 Hapoel_Mahane_Yehuda_F.C.
5: 7470766 Liga_Alef
6: 5130676 Hapoel_Kfar_Saba_F.C.
7: 1755537 Beitar_Jerusalem_F.C.
8: 861284 Maccabi_Tel_Aviv_F.C.
9: 2611690 Mordechai_Spiegler
10: 74656 Knesset

11: 11534927 Maccabi_Jaffa_F.C.
 12: 5046198 Maccabi_Netanya_F.C.
 13: 5166496 Hapoel_Petah_Tikva_F.C.
 14: 2301245 Hapoel_Tel_Aviv_F.C.
 1: 5046198 Maccabi_Netanya_F.C.
 [winner: 5046198 Maccabi_Netanya_F.C., pem_pos=0,
 final_scores=0.17582332, final_local_scores=0.10961418,
 pem=0.0, similarity=7.360819|
 gt: 5046198 Maccabi_Netanya_F.C., pem_pos=0,
 final_scores=0.17582332, final_local_scores=0.10961418,
 pem=0.0, similarity=7.360819]
 [winner: 1415027 David_Berman_(musician), pem_pos=5,
 final_scores=-2.0845554, final_local_scores=-2.1369908,
 pem=-3.0012724, similarity=-43.36097|
 RECALL MISS: 5046198 Maccabi_Netanya_F.C.]
 2: 7209569 Maccabi_World_Union
 [winner: 861284 Maccabi_Tel_Aviv_F.C., pem_pos=2,
 final_scores=-0.05805236, final_local_scores=-0.15633667,
 pem=-0.97941184, similarity=4.750509|
 gt: 7209569 Maccabi_World_Union, pem_pos=1,
 final_scores=-0.34554946, final_local_scores=-0.34727854,
 pem=-0.8720703, similarity=-1.5313174]
 3: 319872 Maccabi_Haifa_F.C.
 [winner: 319872 Maccabi_Haifa_F.C., pem_pos=7,
 final_scores=-0.02133052, final_local_scores=-0.1144461,
 pem=-2.5628016, similarity=14.611326|
 gt: 319872 Maccabi_Haifa_F.C., pem_pos=7,
 final_scores=-0.02133052, final_local_scores=-0.1144461,
 pem=-2.5628016, similarity=14.611326]
 4: 19623772 Maccabi_Sha'arayim_F.C.
 [winner: 19623772 Maccabi_Sha'arayim_F.C., pem_pos=1,
 final_scores=0.014940634, final_local_scores=-0.012986705,
 pem=-0.5877866, similarity=6.8986387|
 gt: 19623772 Maccabi_Sha'arayim_F.C., pem_pos=1,
 final_scores=0.014940634, final_local_scores=-0.012986705,
 pem=-0.5877866, similarity=6.8986387]
 5: 861284 Maccabi_Tel_Aviv_F.C.
 [winner: 2301245 Hapoel_Tel_Aviv_F.C., pem_pos=6,
 final_scores=0.22419117, final_local_scores=0.14770669,
 pem=-2.3020413, similarity=21.016434|
 RECALL MISS: 861284 Maccabi_Tel_Aviv_F.C.]
 6: 3056923 Hapoel
 [winner: 2301245 Hapoel_Tel_Aviv_F.C., pem_pos=1,
 final_scores=0.11961128, final_local_scores=0.03346902,
 pem=-1.3921366, similarity=12.65915|
 gt: 3056923 Hapoel, pem_pos=0,
 final_scores=-0.037250638, final_local_scores=-0.024721123,
 pem=0.0, similarity=3.3519852]
 7: 5834903 Hapoel_Be'er_Sheva_F.C.
 [winner: 312755 Beersheba, pem_pos=1,
 final_scores=-0.07527687, final_local_scores=-0.08988262,
 pem=-0.41640824, similarity=3.6719291|

gt: 5834903 Hapoel_Be'er_Sheva_F.C., pem_pos=5,
final_scores=-0.9812655, final_local_scores=-1.1467745,
pem=-6.660575, similarity=6.0889115]
8: 2869548 Hapoel_Haifa_F.C.
[winner: 319872 Maccabi_Haifa_F.C., pem_pos=7,
final_scores=0.1380039, final_local_scores=0.06359854,
pem=-2.5628016, similarity=19.924534]
RECALL MISS: 2869548 Hapoel_Haifa_F.C.]
9: 5834227 Hapoel_Jerusalem_F.C.
[winner: 16043 Jerusalem, pem_pos=0,
final_scores=-0.07136124, final_local_scores=-0.062511355,
pem=-0.13239029, similarity=2.9442084]
RECALL MISS: 5834227 Hapoel_Jerusalem_F.C.]
10: 19444834 Hapoel_Mahane_Yehuda_F.C.
[winner: 12870798 Mahane_Yehuda_Market, pem_pos=0,
final_scores=-0.033152536, final_local_scores=0.022668779,
pem=0.0, similarity=4.7661943]
gt: 19444834 Hapoel_Mahane_Yehuda_F.C., pem_pos=1,
final_scores=-0.6818298, final_local_scores=-0.7581793,
pem=-4.516339, similarity=6.024682]
1: 5695086 Liga_Leumit
[pem_pos=0, final_scores=0.17023908,
final_local_scores=0.102213874, pem=-0.10604504,
similarity=7.7166686]
2: 1265525 Israel_Football_Association
[pem_pos=0, final_scores=0.17794289,
final_local_scores=0.1508272, pem=0.0, similarity=8.590697]
3: 6059253 Sektzia_Nes_Tziona_F.C.
[pem_pos=0, final_scores=0.23009823,
final_local_scores=0.21949193, pem=0.0, similarity=10.63979]
4: 19444834 Hapoel_Mahane_Yehuda_F.C.
[pem_pos=0, final_scores=0.04649602,
final_local_scores=0.04941903, pem=0.0, similarity=5.5644755]
5: 7470766 Liga_Alef
[pem_pos=0, final_scores=0.14090575,
final_local_scores=0.10495114, pem=0.0, similarity=7.2216644]
6: 5130676 Hapoel_Kfar_Saba_F.C.
[pem_pos=0, final_scores=0.3096738,
final_local_scores=0.27279755, pem=0.0, similarity=12.230536]
7: 1755537 Beitar_Jerusalem_F.C.
[pem_pos=0, final_scores=0.47395772,
final_local_scores=0.42933607, pem=0.0, similarity=16.901958]
8: 861284 Maccabi_Tel_Aviv_F.C.
[pem_pos=0, final_scores=0.19821952,
final_local_scores=0.1284594, pem=-0.22945693,
similarity=9.1710205]
9: 2611690 Mordechai_Spiegler
[pem_pos=0, final_scores=0.43263668,
final_local_scores=0.47344285, pem=0.0, similarity=18.218191]
10: 74656 Knesset
[pem_pos=0, final_scores=0.123207435,
final_local_scores=0.15894908, pem=0.0, similarity=8.833071]

11: 11534927 Maccabi_Jaffa_F.C.
 [pem_pos=0, final_scores=0.24206172,
 final_local_scores=0.20751897, pem=0.0, similarity=10.282493]
 12: 5046198 Maccabi_Netanya_F.C.
 [pem_pos=3, final_scores=0.13568495,
 final_local_scores=0.065769106, pem=-1.097598,
 similarity=12.021302]
 13: 5166496 Hapoel_Petah_Tikva_F.C.
 [pem_pos=1, final_scores=0.27371943,
 final_local_scores=0.21514505, pem=-0.67840284,
 similarity=14.19933]
 14: 2301245 Hapoel_Tel_Aviv_F.C.
 [pem_pos=6, final_scores=0.1475151,
 final_local_scores=0.063949704, pem=-2.3020413,
 similarity=18.516958]
 1: 5046198 Maccabi_Netanya_F.C.
 [pem_pos=0, final_scores=0.17582332,
 final_local_scores=0.10961418, pem=0.0, similarity=7.360819]
 2: 2301245 Hapoel_Tel_Aviv_F.C.
 [pem_pos=6, final_scores=0.22419117,
 final_local_scores=0.14770669, pem=-2.3020413,
 similarity=21.016434]
 3: 2301245 Hapoel_Tel_Aviv_F.C.
 [pem_pos=1, final_scores=0.11961128,
 final_local_scores=0.03346902, pem=-1.3921366,
 similarity=12.65915]
 4: 319872 Maccabi_Haifa_F.C.
 [pem_pos=7, final_scores=0.1380039,
 final_local_scores=0.06359854, pem=-2.5628016,
 similarity=19.924534]

Sample: b'Rogier_Koordes'

[1 Rogier Koordes 1] -LRB- born 13 June 1972 -RRB- is a [2 1 Dutch 2] former 2 footballer who played in [3 England 3] for [1 1 Port Vale 1] between February 1997 and February 1999 . In between this spell he played in the [4 Netherlands 4] for [5 Telstar 5] , [2 2 Oss 2] , [3 3 Haarlem 3] , and [4 4 Ter Lee de 4] . [6 Koordes 6] began his career at [7 Telstar 7] , who finished the 3 1994 -- 95 season in 11th place in the [5 5 Eerste Divisie 5] , and then ninth place in 4 1995 -- 96 under [6 6 Cor Pot 6] 's stewardship . [8 Koordes 8] left the 5 TATA Steel Stadion when he was purchased by [7 7 English 7] 6 First Division club [8 8 Port Vale 8] in February 1997 for a # 75,000 fee . He played 13 league games in the remainder of the 7 1996 -- 97 season , but featured just nine times in the 8 1997 -- 98 campaign . He played 17 league and [9 9 League Cup 9] games in the 9 1998 -- 99 campaign , but left [10 10 Vale Park 10] in February 1999 after manager [11 11 John Rudge 11] was sacked and replaced by [12 12 Brian Horton 12] . After leaving 10 Burslem he returned to his native [9 Holland 9] to play for [13 13 Oss 13] . TOP finished 16th in the [10 Eerste Divisie 10] in 11 1998 -- 99 , and then in last place in 12 19

99 -- 2000 , before rising to tenth position in 13 2000 -- 01 . [11 Koordes 11] then switched to [14 14 Heini Otto 14] 's [1 5 15 HFC Haarlem 15] , and helped the club to 12th and 14th place finishes in 14 2001 -- 02 and 15 2002 -- 03 . He then left 16 Haarlem Stadion and ended up with [16 16 Ter Leede 16] . In January 2007 he joined Rijnvogels in the 17 Eerste Klasse

1: 203456 Port_Vale_F.C.
2: 1193665 FC_Oss
3: 2173100 HFC_Haarlem
4: 14147419 Ter_Leede
5: 1938252 Eerste_Divisie
6: 14910370 Cor_Pot
7: 9316 England
8: 203456 Port_Vale_F.C.
9: 144982 Football_League_Cup
10: 2817479 Vale_Park
11: 2844350 John_Rudge
12: 1996967 Brian_Horton
13: 1193665 FC_Oss
14: 8672689 Heini_Otto
15: 2173100 HFC_Haarlem
16: 14147419 Ter_Leede

1: 1656551 Dutch_people
[winner: 21148 Netherlands, pem_pos=0,
final_scores=0.39545584, final_local_scores=0.41723603,
pem=-0.7943452, similarity=20.860641]
gt: 1656551 Dutch_people, pem_pos=2,
final_scores=-0.43608588, final_local_scores=-0.4428102,
pem=-2.0349581, similarity=1.9417918]

2: 10568 Association_football
[winner: 18462732 Shane_Valenti, pem_pos=0,
final_scores=-1.0502986, final_local_scores=-1.0443194,
pem=0.0, similarity=-27.07486]
RECALL MISS: 10568 Association_football]

3: 21632969 1994-95_Eerste_Divisie
[winner: 34658 1995, pem_pos=5, final_scores=-0.6591265,
final_local_scores=-0.59420025, pem=-3.215059,
similarity=3.8415825]
RECALL MISS: 21632969 1994-95_Eerste_Divisie]
[winner: 18948337 1994, pem_pos=0,
final_scores=-1.0137484, final_local_scores=-0.966123,
pem=-0.7539553, similarity=-20.6412]
RECALL MISS: 21632969 1994-95_Eerste_Divisie]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3532909, final_local_scores=-1.2647095,
pem=0.0, similarity=-33.65174]
RECALL MISS: 21632969 1994-95_Eerste_Divisie]

4: 21632972 1995-96_Eerste_Divisie
[winner: 26089658 No._96_Squadron_RAF, pem_pos=0,
final_scores=-0.47696346, final_local_scores=-0.38217527,
pem=-2.1547496, similarity=4.4027033]
RECALL MISS: 21632972 1995-96_Eerste_Divisie]

[winner: 34658 1995, pem_pos=0, final_scores=-1.212566, final_local_scores=-1.1987488, pem=-0.68201965, similarity=-27.974415]
 RECALL MISS: 21632972 1995-96_Eerste_Divisie]
 [winner: 38340996 Methylene_bridge, pem_pos=0, final_scores=-1.4256951, final_local_scores=-1.3438001, pem=0.0, similarity=-36.01196]
 RECALL MISS: 21632972 1995-96_Eerste_Divisie]
 5: 3264939 TATA_Steel_Stadion
 [winner: 3264939 TATA_Steel_Stadion, pem_pos=0, final_scores=-0.018754676, final_local_scores=0.06230119, pem=0.0, similarity=5.9489045]
 gt: 3264939 TATA_Steel_Stadion, pem_pos=0, final_scores=-0.018754676, final_local_scores=0.06230119, pem=0.0, similarity=5.9489045]
 6: 28789929 Football_League_First_Division
 [winner: 28789929 Football_League_First_Division, pem_pos=0, final_scores=-0.26861304, final_local_scores=-0.33695847, pem=-0.5147791, similarity=-3.1663516]
 gt: 28789929 Football_League_First_Division, pem_pos=0, final_scores=-0.26861304, final_local_scores=-0.33695847, pem=-0.5147791, similarity=-3.1663516]
 7: 30501505 1996-97_Port_Vale_F.C._season
 [winner: 34636 1996, pem_pos=0, final_scores=-0.7274433, final_local_scores=-0.55167294, pem=-0.78063476, similarity=-8.128095]
 RECALL MISS: 30501505 1996-97_Port_Vale_F.C._season]
 [winner: 12254377 No._97_Squadron_RAF, pem_pos=2, final_scores=-0.7992443, final_local_scores=-0.74642277, pem=-2.8157518, similarity=-2.872535]
 RECALL MISS: 30501505 1996-97_Port_Vale_F.C._season]
 [winner: 38340996 Methylene_bridge, pem_pos=0, final_scores=-1.2546828, final_local_scores=-1.156995, pem=0.0, similarity=-30.437325]
 RECALL MISS: 30501505 1996-97_Port_Vale_F.C._season]
 8: 30506233 1997-98_Port_Vale_F.C._season
 [winner: 75433 Windows_98, pem_pos=0, final_scores=-0.8719332, final_local_scores=-0.8081634, pem=-1.4472766, similarity=-12.15698]
 RECALL MISS: 30506233 1997-98_Port_Vale_F.C._season]
 [winner: 2602335 Toronto_municipal_election,_1997, pem_pos=0, final_scores=-0.986096, final_local_scores=-0.9342847, pem=0.0, similarity=-23.791206]
 RECALL MISS: 30506233 1997-98_Port_Vale_F.C._season]
 [winner: 38340996 Methylene_bridge, pem_pos=0, final_scores=-1.186686, final_local_scores=-1.0827191, pem=0.0, similarity=-28.220783]
 RECALL MISS: 30506233 1997-98_Port_Vale_F.C._season]
 9: 30510072 1998-99_Port_Vale_F.C._season
 [winner: 4669371 Juno_Awards_of_1998, pem_pos=0,

final_scores=-0.78156465, final_local_scores=-0.6582986,
 pem=0.0, similarity=-15.555231|
 RECALL MISS: 30510072 1998-99_Port_Vale_F.C._season]
 [winner: 15325337 99_(1918_film), pem_pos=4,
 final_scores=-1.1106191, final_local_scores=-0.9744124,
 pem=-3.1113029, similarity=-8.068949|
 RECALL MISS: 30510072 1998-99_Port_Vale_F.C._season]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.2443997, final_local_scores=-1.1457624,
 pem=0.0, similarity=-30.102125|
 RECALL MISS: 30510072 1998-99_Port_Vale_F.C._season]
 10: 1266131 Burslem
 [winner: 1266131 Burslem, pem_pos=0,
 final_scores=-0.0062194765, final_local_scores=-0.0141083,
 pem=0.0, similarity=3.668693|
 gt: 1266131 Burslem, pem_pos=0,
 final_scores=-0.0062194765, final_local_scores=-0.0141083,
 pem=0.0, similarity=3.668693]
 11: 21632981 1998-99_Eerste_Divisie
 [winner: 15325337 99_(1918_film), pem_pos=4,
 final_scores=-1.0602137, final_local_scores=-0.9193521,
 pem=-3.1113029, similarity=-6.42584|
 RECALL MISS: 21632981 1998-99_Eerste_Divisie]
 [winner: 34647 1998, pem_pos=0, final_scores=-1.1148999,
 final_local_scores=-1.11045, pem=-0.6850515,
 similarity=-25.322916|
 RECALL MISS: 21632981 1998-99_Eerste_Divisie]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.2751248, final_local_scores=-1.1793249,
 pem=0.0, similarity=-31.103695|
 RECALL MISS: 21632981 1998-99_Eerste_Divisie]
 12: 21632982 1999-2000_Eerste_Divisie
 [winner: 432383 2000_United_States_Census, pem_pos=1,
 final_scores=-0.5114176, final_local_scores=-0.44343677,
 pem=-1.4367632, similarity=-1.329983|
 RECALL MISS: 21632982 1999-2000_Eerste_Divisie]
 [winner: 15723636 1999_Vladikavkaz_bombing, pem_pos=5,
 final_scores=-1.1677579, final_local_scores=-1.1133759,
 pem=-1.6269311, similarity=-20.288141|
 RECALL MISS: 21632982 1999-2000_Eerste_Divisie]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.3825557, final_local_scores=-1.2966769,
 pem=0.0, similarity=-34.60571|
 RECALL MISS: 21632982 1999-2000_Eerste_Divisie]
 13: 21632983 2000-01_Eerste_Divisie
 [winner: 432383 2000_United_States_Census, pem_pos=1,
 final_scores=-0.62225294, final_local_scores=-0.5645075,
 pem=-1.4367632, similarity=-4.9429755|
 RECALL MISS: 21632983 2000-01_Eerste_Divisie]
 [winner: 3453517 FabricLive.01, pem_pos=2,
 final_scores=-0.8200912, final_local_scores=-0.6297109,
 pem=-2.3963957, similarity=-1.6701453|

RECALL MISS: 21632983 2000-01_Eerste_Divisie]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.4190623, final_local_scores=-1.3365548,
pem=0.0, similarity=-35.79575]
RECALL MISS: 21632983 2000-01_Eerste_Divisie]
14: 21632986 2001-02_Eerste_Divisie
[winner: 34551 2001, pem_pos=0, final_scores=-0.6686946,
final_local_scores=-0.43030888, pem=-0.5613713,
similarity=-5.698738]
RECALL MISS: 21632986 2001-02_Eerste_Divisie]
[winner: 2385469 Federal_Firearms_License, pem_pos=2,
final_scores=-0.7472494, final_local_scores=-0.6759825,
pem=-2.2935443, similarity=-3.6103015]
RECALL MISS: 21632986 2001-02_Eerste_Divisie]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.384852, final_local_scores=-1.2991853,
pem=0.0, similarity=-34.680565]
RECALL MISS: 21632986 2001-02_Eerste_Divisie]
15: 21632987 2002-03_Eerste_Divisie
[winner: 35502 2002, pem_pos=0, final_scores=-0.78343374,
final_local_scores=-0.664358, pem=-0.5897454,
similarity=-12.528929]
RECALL MISS: 21632987 2002-03_Eerste_Divisie]
[winner: 3453736 FabricLive.03, pem_pos=2,
final_scores=-0.8305588, final_local_scores=-0.66527873,
pem=-2.3070445, similarity=-3.2174637]
RECALL MISS: 21632987 2002-03_Eerste_Divisie]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3619806, final_local_scores=-1.2742016,
pem=0.0, similarity=-33.935005]
RECALL MISS: 21632987 2002-03_Eerste_Divisie]
16: 4667360 Haarlem_Stadion
[winner: 4667360 Haarlem_Stadion, pem_pos=0,
final_scores=-0.06459972, final_local_scores=-0.009524509,
pem=0.0, similarity=3.8054824]
gt: 4667360 Haarlem_Stadion, pem_pos=0,
final_scores=-0.06459972, final_local_scores=-0.009524509,
pem=0.0, similarity=3.8054824]
17: 14282479 Eerste_Klasse
[winner: 14282479 Eerste_Klasse, pem_pos=0,
final_scores=-0.12776177, final_local_scores=-0.13411537,
pem=0.0, similarity=0.087442905]
gt: 14282479 Eerste_Klasse, pem_pos=0,
final_scores=-0.12776177, final_local_scores=-0.13411537,
pem=0.0, similarity=0.087442905]
1: 203456 Port_Vale_F.C.
[pem_pos=0, final_scores=0.47669,
final_local_scores=0.46979845, pem=0.0, similarity=18.109436]
2: 1193665 FC_Oss
[pem_pos=2, final_scores=0.2638805,
final_local_scores=0.24976444, pem=-1.5440304,
similarity=19.93986]

3: 2173100 HFC_Haarlem
[pem_pos=1, final_scores=0.12617455,
final_local_scores=0.08618519, pem=-1.6355026,
similarity=15.555767]
4: 14147419 Ter_Leede
[pem_pos=0, final_scores=0.23855202,
final_local_scores=0.27838117, pem=0.0, similarity=12.397162]
5: 1938252 Eerste_Divisie
[pem_pos=0, final_scores=0.19882236,
final_local_scores=0.1659756, pem=0.0, similarity=9.042756]
6: 14910370 Cor_Pot
[pem_pos=0, final_scores=0.34062904,
final_local_scores=0.380477, pem=0.0, similarity=15.443906]
7: 9316 England
[pem_pos=1, final_scores=0.3529958,
final_local_scores=0.4069273, pem=-1.3527626,
similarity=23.58977]
8: 203456 Port_Vale_F.C.
[pem_pos=0, final_scores=0.47496492,
final_local_scores=0.46791404, pem=0.0, similarity=18.053202]
9: 144982 Football_League_Cup
[pem_pos=0, final_scores=0.37706953,
final_local_scores=0.37681907, pem=-0.58769286,
similarity=18.530712]
10: 2817479 Vale_Park
[pem_pos=0, final_scores=0.4887159,
final_local_scores=0.5169637, pem=0.0, similarity=19.516941]
11: 2844350 John_Rudge
[pem_pos=0, final_scores=0.46272677,
final_local_scores=0.46867675, pem=0.0, similarity=18.075962]
12: 1996967 Brian_Horton
[pem_pos=0, final_scores=0.5244029,
final_local_scores=0.5642127, pem=0.0, similarity=20.926945]
13: 1193665 FC_Oss
[pem_pos=2, final_scores=0.34606117,
final_local_scores=0.33953434, pem=-1.5440304,
similarity=22.618773]
14: 8672689 Heini_Otto
[pem_pos=0, final_scores=0.7233947,
final_local_scores=0.7990436, pem=0.0, similarity=27.934767]
15: 2173100 HFC_Haarlem
[pem_pos=0, final_scores=0.3109455,
final_local_scores=0.28769168, pem=0.0, similarity=12.675007]
16: 14147419 Ter_Leede
[pem_pos=0, final_scores=0.20238627,
final_local_scores=0.23887557, pem=0.0, similarity=11.218236]
1: 21076674 Rogier_Koordes
[pem_pos=0, final_scores=0.19303884,
final_local_scores=0.19036901, pem=0.0, similarity=9.770704]
2: 21148 Netherlands
[pem_pos=0, final_scores=0.39545584,
final_local_scores=0.41723603, pem=-0.7943452,

```

similarity=20.860641]
3: 9316 England
    [pem_pos=0, final_scores=0.48402697,
final_local_scores=0.5500589, pem=-0.19053856,
similarity=21.540747]
4: 21148 Netherlands
    [pem_pos=0, final_scores=0.31276208,
final_local_scores=0.32690567, pem=-0.24940795,
similarity=15.2015505]
5: 1016387 SC_Telstar
    [pem_pos=2, final_scores=0.36312503,
final_local_scores=0.34141392, pem=-1.3744984,
similarity=21.75292]
6: 21076674 Rogier_Koordes
    [pem_pos=0, final_scores=0.30620748,
final_local_scores=0.3139885, pem=0.0, similarity=13.459756]
7: 1016387 SC_Telstar
    [pem_pos=2, final_scores=0.46535772,
final_local_scores=0.45308763, pem=-1.3744984,
similarity=25.085487]
8: 21076674 Rogier_Koordes
    [pem_pos=0, final_scores=0.31718427,
final_local_scores=0.32597905, pem=0.0, similarity=13.817578]
9: 21148 Netherlands
    [pem_pos=1, final_scores=0.06992687,
final_local_scores=0.06164509, pem=-2.7840104,
similarity=21.069206]
10: 1938252 Eerste_Divisie
    [pem_pos=0, final_scores=0.22649194,
final_local_scores=0.19620043, pem=0.0, similarity=9.944725]
11: 21076674 Rogier_Koordes
    [pem_pos=0, final_scores=0.36186993,
final_local_scores=0.37479132, pem=0.0, similarity=15.274233]

```

Sample: b'Bob_Coverdale'

Robin `` Bob '' [1 Coverdale 1] -LRB- birth registered
October → December in 1 Sculcoates district -RRB- is an [2 2 E
nglish 2] professional [1 1 Rugby League World Cup 1] winning
footballer of the 1950s and '60s , playing at representative
level for 3 Great Britain , and at club level for 4 Hull , [2
2 Wakefield Trinity 2] , and [3 3 Hull Kingston Rovers 3] ,
as a , i.e. number 8 or 10 , during the era of . [3 Bob
Coverdale 3] won caps for [4 4 Great Britain 4] while at [4
Hull 4] in the 5 1954 [5 Rugby League World Cup 5] against
Australia , [6 France 6] , [7 New Zealand 7] , and [8 France 8
] . [9 Bob Coverdale 9] played Right - , i.e. number 10 in
all four of [5 5 Great Britain 5] 's 6 1954 [10 Rugby League W
orld Cup 10] matches , including [6 6 Great Britain 6] 's
16-12 victory over [11 7 France 11] in the 8 1954 [12 Rugby Le
ague World Cup 12] final at [7 7 Parc des Princes 7] , [8 8 Pa
ris 8] on Saturday 13 November 1954 . [13 Bob Coverdale 13]
played in all three of [14 Hull 14] 's defeats in the 9 1953 -

- 54 , 10 1954 -- 55 , and 11 1955 -- 56 [9 9 Yorkshire Cup 9]

1: 12312312 Rugby_League_World_Cup
2: 1095973 Wakefield_Trinity_Wildcats
3: 514219 Hull_Kingston_Rovers
4: 951941 Great_Britain_national_rugby_league_team
5: 951941 Great_Britain_national_rugby_league_team
6: 951941 Great_Britain_national_rugby_league_team
7: 761302 Parc_des_Princes
8: 22989 Paris
9: 1010713 Rugby_league_county_cups

1: 28890286 Sculcoates
[winner: 28890286 Sculcoates, pem_pos=1,
final_scores=-0.048572674, final_local_scores=0.05644515,
pem=-0.15860501, similarity=6.6366677]
gt: 28890286 Sculcoates, pem_pos=1,
final_scores=-0.048572674, final_local_scores=0.05644515,
pem=-0.15860501, similarity=6.6366677]

2: 18803164 English_people
[winner: 9316 England, pem_pos=1, final_scores=0.5827705,
final_local_scores=0.6181838, pem=-1.3527626,
similarity=29.894083]
gt: 18803164 English_people, pem_pos=2,
final_scores=-0.40655172, final_local_scores=-0.41634923,
pem=-2.2084396, similarity=3.6748586]

3: 951941 Great_Britain_national_rugby_league_team
[winner: 13530298 Great_Britain, pem_pos=2,
final_scores=-0.034852877, final_local_scores=-0.03658071,
pem=-0.50213265, similarity=5.728747]
gt: 951941 Great_Britain_national_rugby_league_team,
pem_pos=5, final_scores=-0.09164162,
final_local_scores=-0.17328164, pem=-2.6635387,
similarity=13.40338]

4: 1095938 Hull_F.C.
[winner: 451358 Hull_City_A.F.C., pem_pos=11,
final_scores=-0.08802448, final_local_scores=-0.1291039,
pem=-3.7682815, similarity=20.729496]
gt: 1095938 Hull_F.C., pem_pos=1,
final_scores=-0.14996643, final_local_scores=-0.23262483,
pem=-2.238242, similarity=9.319632]

5: 377381 1954_Rugby_League_World_Cup
[winner: 12312312 Rugby_League_World_Cup, pem_pos=0,
final_scores=0.32020968, final_local_scores=0.2701482,
pem=0.0, similarity=12.151474]
gt: 377381 1954_Rugby_League_World_Cup, pem_pos=6,
final_scores=-0.92896837, final_local_scores=-1.047024,
pem=-7.179308, similarity=11.88662]
[winner: 34982 1954, pem_pos=0, final_scores=-0.57260704,
final_local_scores=-0.3393554, pem=-0.7808452,
similarity=-1.790971]
RECALL MISS: 377381 1954_Rugby_League_World_Cup]

6: 377381 1954_Rugby_League_World_Cup
[winner: 12312312 Rugby_League_World_Cup, pem_pos=0,

final_scores=0.3026368, final_local_scores=0.25095248,
 pem=0.0, similarity=11.578635|
 gt: 377381 1954_Rugby_League_World_Cup, pem_pos=6,
 final_scores=-0.9439108, final_local_scores=-1.0633464,
 pem=-7.179308, similarity=11.399529]
 [winner: 34982 1954, pem_pos=0, final_scores=-0.5945747,
 final_local_scores=-0.3633517, pem=-0.7808452,
 similarity=-2.5070693|
 RECALL MISS: 377381 1954_Rugby_League_World_Cup]
 7: 966005 France_national_rugby_league_team
 [winner: 250197 France_national_football_team, pem_pos=1,
 final_scores=0.20274673, final_local_scores=0.17520267,
 pem=-3.3831942, similarity=27.716446|
 gt: 966005 France_national_rugby_league_team, pem_pos=5,
 final_scores=-0.26174152, final_local_scores=-0.34815246,
 pem=-4.934728, similarity=20.535967]
 8: 377381 1954_Rugby_League_World_Cup
 [winner: 12312312 Rugby_League_World_Cup, pem_pos=0,
 final_scores=0.29529315, final_local_scores=0.24293068,
 pem=0.0, similarity=11.339249|
 gt: 377381 1954_Rugby_League_World_Cup, pem_pos=6,
 final_scores=-0.98542696, final_local_scores=-1.1086965,
 pem=-7.179308, similarity=10.046189]
 [winner: 34982 1954, pem_pos=0, final_scores=-0.57114035,
 final_local_scores=-0.33775327, pem=-0.7808452,
 similarity=-1.7431606|
 RECALL MISS: 377381 1954_Rugby_League_World_Cup]
 9: 18538177 1953-54_Northern_Rugby_Football_League_season
 [winner: 34865 1953, pem_pos=0, final_scores=-0.5014944,
 final_local_scores=-0.27858555, pem=-0.7062917,
 similarity=-0.38291207|
 RECALL MISS: 18538177 1953-
 54_Northern_Rugby_Football_League_season]
 [winner: 4984490 Now_That's_What_I_Call_Music!
 54(UK_series), pem_pos=4, final_scores=-1.1311569,
 final_local_scores=-1.0665963, pem=-3.2720842,
 similarity=-9.945544|
 RECALL MISS: 18538177 1953-
 54_Northern_Rugby_Football_League_season]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.329288, final_local_scores=-1.2208354,
 pem=0.0, similarity=-32.34245|
 RECALL MISS: 18538177 1953-
 54_Northern_Rugby_Football_League_season]
 10: 24318753 1954-55_Northern_Rugby_Football_League_season
 [winner: 34982 1954, pem_pos=0, final_scores=-0.55635566,
 final_local_scores=-0.32160324, pem=-0.7808452,
 similarity=-1.2612112|
 RECALL MISS: 24318753 1954-
 55_Northern_Rugby_Football_League_season]
 [winner: 35270 55, pem_pos=2, final_scores=-1.2082272,
 final_local_scores=-1.1498682, pem=-3.0374093,

similarity=-13.706746|
RECALL MISS: 24318753 1954-
55_Northern_Rugby_Football_League_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.2599955, final_local_scores=-1.1451437,
pem=0.0, similarity=-30.083662|
RECALL MISS: 24318753 1954-
55_Northern_Rugby_Football_League_season]
11: 24318840 1955-56_Northern_Rugby_Football_League_season
[winner: 34851 1955, pem_pos=0, final_scores=-0.4648481,
final_local_scores=-0.20233253, pem=-0.7263205,
similarity=2.0015502|
RECALL MISS: 24318840 1955-
56_Northern_Rugby_Football_League_season]
[winner: 30818211 FabricLive.56, pem_pos=18,
final_scores=-1.0968401, final_local_scores=-0.98391265,
pem=-3.8781977, similarity=-4.18196|
RECALL MISS: 24318840 1955-
56_Northern_Rugby_Football_League_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3009925, final_local_scores=-1.1899269,
pem=0.0, similarity=-31.420076|
RECALL MISS: 24318840 1955-
56_Northern_Rugby_Football_League_season]
1: 12312312 Rugby_League_World_Cup
[pem_pos=0, final_scores=0.25370175,
final_local_scores=0.19614255, pem=0.0, similarity=9.942998]
2: 1095973 Wakefield_Trinity_Wildcats
[pem_pos=0, final_scores=0.21626435,
final_local_scores=0.18137729, pem=0.0, similarity=9.502373]
3: 514219 Hull_Kingston_Rovers
[pem_pos=0, final_scores=0.14286189,
final_local_scores=0.0913434, pem=0.0, similarity=6.8155823]
4: 951941 Great_Britain_national_rugby_league_team
[pem_pos=5, final_scores=0.21933375,
final_local_scores=0.1708262, pem=-2.6635387,
similarity=23.672245]
5: 951941 Great_Britain_national_rugby_league_team
[pem_pos=2, final_scores=0.039413646,
final_local_scores=-0.030123651, pem=-2.6635387,
similarity=17.6755]
6: 951941 Great_Britain_national_rugby_league_team
[pem_pos=5, final_scores=0.15346883,
final_local_scores=0.09887874, pem=-2.6635387,
similarity=21.525188]
7: 761302 Parc_des_Princes
[pem_pos=0, final_scores=0.4957053,
final_local_scores=0.52567047, pem=0.0, similarity=19.776768]
8: 22989 Paris
[pem_pos=0, final_scores=0.5408806,
final_local_scores=0.6167627, pem=-0.115301244,
similarity=23.122168]

9: 1010713 Rugby_league_county_cups
[pem_pos=0, final_scores=0.06745763,
final_local_scores=0.06474444, pem=-0.13236055,
similarity=6.7416134]
1: 22886206 Bob_Coverdale
[pem_pos=0, final_scores=0.16150592,
final_local_scores=0.16162091, pem=0.0, similarity=8.912804]
2: 9316 England
[pem_pos=1, final_scores=0.5827705,
final_local_scores=0.6181838, pem=-1.3527626,
similarity=29.894083]
3: 22886206 Bob_Coverdale
[pem_pos=0, final_scores=0.4267748,
final_local_scores=0.45138687, pem=0.0, similarity=17.559998]
4: 1095938 Hull_F.C.
[pem_pos=1, final_scores=0.108869,
final_local_scores=0.05668506, pem=-2.238242,
similarity=17.953217]
5: 12312312 Rugby_League_World_Cup
[pem_pos=0, final_scores=0.32020968,
final_local_scores=0.2701482, pem=0.0, similarity=12.151474]
6: 5843419 France
[pem_pos=0, final_scores=0.15685146,
final_local_scores=0.17542872, pem=-0.16240834,
similarity=10.208057]
7: 154957 New_Zealand_national_rugby_union_team
[pem_pos=4, final_scores=0.4771387,
final_local_scores=0.46451342, pem=-3.8192244,
similarity=38.721256]
8: 5843419 France
[pem_pos=0, final_scores=0.18078883,
final_local_scores=0.20157668, pem=-0.16240834,
similarity=10.988365]
9: 22886206 Bob_Coverdale
[pem_pos=0, final_scores=0.47625875,
final_local_scores=0.50544053, pem=0.0, similarity=19.173067]
10: 12312312 Rugby_League_World_Cup
[pem_pos=0, final_scores=0.3026368,
final_local_scores=0.25095248, pem=0.0, similarity=11.578635]
11: 250197 France_national_football_team
[pem_pos=1, final_scores=0.20274673,
final_local_scores=0.17520267, pem=-3.3831942,
similarity=27.716446]
12: 12312312 Rugby_League_World_Cup
[pem_pos=0, final_scores=0.29529315,
final_local_scores=0.24293068, pem=0.0, similarity=11.339249]
13: 22886206 Bob_Coverdale
[pem_pos=0, final_scores=0.43248385,
final_local_scores=0.45762312, pem=0.0, similarity=17.7461]
14: 451358 Hull_City_A.F.C.
[pem_pos=11, final_scores=0.16888951,
final_local_scores=0.15872443, pem=-3.7682815,

similarity=29.318867]

Sample: b'Beeren_Island'

Beeren Island -LRB- Beeren Eylandt in the original -RRB- is an island in the [1 1 Hudson River 1] within the of , [2 2 New York 2] . It lies 11 miles south of the of [3 3 Albany 3] , and is at the border of four , [1 1 Albany 1] , 2 Greene -LRB- to the south -RRB- , 3 Columbia -LRB- to the southeast -RRB- , and 4 Rensselaer -LRB- to the east -RRB- . Beeren Island was once referred to as Mach-a-wa-meck , which may come from mashq -LRB- bear -RRB- and wamock -LRB- enough -RRB- , meaning `` place of many bears '' and would make it in agreement with the [2 Dutch 2] name . The island was claimed by the [4 4 Patroon 4] as the southern border of his of , it was from the start of the 5 New Netherland colony chosen as a fine site for fortifications to protect the patroon 's lands . In 1643 a fort named Rensselaer Stein was built and garrisoned , and included mounted cannon ; Nicolas Coorn was appointed watchmaster . Every ship passing the island had to lower its flag in deference to the patroon and pay five 6 sti ver s . A dispute arose between Coorn and the captain of the ship Good Hope when Coorn demanded that Captain Govert Loockermans lower his flag , and Loockermans replied that `` I lower my colors for no one except the [3 Prince of Orange 3] and the Lords my masters . '' At which point Coorn shot a cannonball through the mainsail of the Good Hope . The however refused to acknowledge the patroon 's exclusive right over the river , this led to tension between the two until the [5 5 English 5] conquered the entire [4 New Netherland 4] colony in 1664 . The island was the location of the first 7 wh ite child born along the [5 Hudson River 5] . The island has at various times also been occupied by a few [6 8 Native Amer ican 6] hunters , by the Coeymans family , a 9 quarantine station , and was even the site of fortune hunters looking for

1: 47911 Hudson_River

2: 8210131 New_York

3: 52106 Albany,_New_York

4: 2284768 Patroon

5: 9316 England

1: 50559 Albany_County,_New_York

[winner: 52106 Albany,_New_York, pem_pos=0,
final_scores=0.033741787, final_local_scores=-0.028245166,
pem=-1.175974, similarity=9.641951|

gt: 50559 Albany_County,_New_York, pem_pos=4,
final_scores=-0.5278687, final_local_scores=-0.6265199,
pem=-2.9701676, similarity=1.5453414]

2: 54338 Greene_County,_New_York

[winner: 54338 Greene_County,_New_York, pem_pos=0,
final_scores=-0.63962024, final_local_scores=-0.7190016,
pem=-2.5238388, similarity=-3.641701|

gt: 54338 Greene_County,_New_York, pem_pos=0,
final_scores=-0.63962024, final_local_scores=-0.7190016,

pem=-2.5238388, similarity=-3.641701]
 3: 54165 Columbia_County,_New_York
 [winner: 165108 Columbia_Records, pem_pos=0,
 final_scores=-0.28509885, final_local_scores=-0.25577682,
 pem=-1.516849, similarity=4.705683]
 RECALL MISS: 54165 Columbia_County,_New_York]
 4: 54453 Rensselaer_County,_New_York
 [winner: 54453 Rensselaer_County,_New_York, pem_pos=0,
 final_scores=-0.06295653, final_local_scores=-0.1003599,
 pem=-1.0937719, similarity=7.0428753]
 gt: 54453 Rensselaer_County,_New_York, pem_pos=0,
 final_scores=-0.06295653, final_local_scores=-0.1003599,
 pem=-1.0937719, similarity=7.0428753]
 5: 43451 New_Netherland
 [winner: 43451 New_Netherland, pem_pos=0,
 final_scores=0.028112814, final_local_scores=-0.037755534,
 pem=0.0, similarity=2.9630125]
 gt: 43451 New_Netherland, pem_pos=0,
 final_scores=0.028112814, final_local_scores=-0.037755534,
 pem=0.0, similarity=2.9630125]
 6: 3326732 Stiver
 [winner: 3326732 Stiver, pem_pos=0,
 final_scores=-0.98062986, final_local_scores=-0.91761357,
 pem=0.0, similarity=-23.293707]
 gt: 3326732 Stiver, pem_pos=0, final_scores=-0.98062986,
 final_local_scores=-0.91761357, pem=0.0,
 similarity=-23.293707]
 7: 166202 White_people
 [winner: 273285
 Race_and_ethnicity_in_the_United_States_Census, pem_pos=0,
 final_scores=-0.7061231, final_local_scores=-0.6566109,
 pem=-0.2843238, similarity=-13.95867]
 gt: 166202 White_people, pem_pos=3,
 final_scores=-1.5147173, final_local_scores=-1.6610291,
 pem=-3.3325114, similarity=-27.355991]
 8: 2402376 Indigenous_peoples_of_the_Americas
 [winner: 21217 Native_Americans_in_the_United_States,
 pem_pos=1, final_scores=0.07620405,
 final_local_scores=0.06090097, pem=-1.216356,
 similarity=12.5218525]
 gt: 2402376 Indigenous_peoples_of_the_Americas,
 pem_pos=2, final_scores=-0.030459642,
 final_local_scores=-0.04089701, pem=-1.7986634,
 similarity=12.650676]
 9: 25237 Quarantine
 [winner: 3001381 Quarantine_(Jim_Crace_novel), pem_pos=5,
 final_scores=-0.9621431, final_local_scores=-0.820448,
 pem=-1.7486272, similarity=-10.884785]
 gt: 25237 Quarantine, pem_pos=7, final_scores=-1.801105,
 final_local_scores=-1.9335701, pem=-1.8307467,
 similarity=-43.655994]
 1: 47911 Hudson_River

[pem_pos=0, final_scores=0.31461543,
 final_local_scores=0.27866888, pem=0.0, similarity=12.405748]
 2: 8210131 New_York
 [pem_pos=0, final_scores=0.20045374,
 final_local_scores=0.17122597, pem=-0.2889045,
 similarity=10.770545]
 3: 52106 Albany,_New_York
 [pem_pos=0, final_scores=0.080810204,
 final_local_scores=0.035457537, pem=-1.175974,
 similarity=11.542966]
 4: 2284768 Patroon
 [pem_pos=0, final_scores=0.1529861,
 final_local_scores=0.1670022, pem=0.0, similarity=9.073392]
 5: 9316 England
 [pem_pos=1, final_scores=0.32314396,
 final_local_scores=0.3315545, pem=-1.3527626,
 similarity=21.340494]
 1: 52106 Albany,_New_York
 [pem_pos=0, final_scores=0.033741787,
 final_local_scores=-0.028245166, pem=-1.175974,
 similarity=9.641951]
 2: 21148 Netherlands
 [pem_pos=0, final_scores=0.31316966,
 final_local_scores=0.3270619, pem=-0.7943452,
 similarity=18.169664]
 3: 100384 Willem-Alexander_of_the_Netherlands
 [pem_pos=0, final_scores=0.13460808,
 final_local_scores=0.1985785, pem=-0.29345012,
 similarity=11.611519]
 4: 43451 New_Netherland
 [pem_pos=0, final_scores=0.14047764,
 final_local_scores=0.09668669, pem=0.0, similarity=6.9750366]
 5: 47911 Hudson_River
 [pem_pos=0, final_scores=0.27631372,
 final_local_scores=0.2286984, pem=0.0, similarity=10.91453]
 6: 21217 Native_Americans_in_the_United_States
 [pem_pos=1, final_scores=0.07620405,
 final_local_scores=0.06090097, pem=-1.216356,
 similarity=12.5218525]

Sample: b'Gemaal_Hussain'

[1 Gemaal Maqsood Hussain 1] -LRB- born 10 October 1983 -
 RRB- is an [1 1 English 1] 1 cricket er who plays for [2 2 So
 merset County Cricket Club 2] . A right-arm 2 fast-medium
 paced bowler , he moved to [2 Somerset 2] from neighbouring
 county [3 3 Gloucestershire 3] at the close of the 3 2010 sea
 son . [3 Hussain 3] was born in [4 4 Waltham Forest 4] , [5 5
 London 5] on 10 October 1983 . He was raised in [4 Nottingham
 4] where he played club cricket for during 2002 , and later
 played in [5 Bradford 5] for four years while he studied [6 6
 Sports Science 6] at the [7 7 University of Leeds 7] . During
 the 2006 season , he played a number of games for the [8 8 Es

sex 8] second team , and the following year he appeared for a variety of county second teams , starting off at [9 9 Nottinghamshire 9] , before moving onto [10 10 Worcestershire 10] , and then back to [6 Essex 6] , followed by a single match for [7 Surrey 7] , and another for [8 Essex 8] . He played two second eleven matches in 2008 , both for [9 Worcestershire 9] , and was spotted playing in the [10 Birmingham League 10] by [11 11 Jack Russell 11] and [12 12 Stuart Barnes 12] , after which he was offered a trial at [11 Gloucestershire 11] . He made his debut for [12 Gloucestershire 12] during the 4 2009 Twenty20 Cup , claiming two wickets against [13 Worcestershire 13] . He finished the competition with 10 wickets , second amongst [14 Gloucestershire 14] bowlers . His first-class debut followed immediately after the [15 Twenty20 Cup 15] when he was selected for the [13 13 County Championship 13] match against [14 14 Kent 14] . He claimed two 5 tail-end wickets in the first innings of the match , which [16 Gloucestershire 16] lost by 76 runs . He did not appear for [17 Gloucestershire 17] again until the final match of the season , when he made his debut in [18 6 List A 18] cricket , taking the wickets of both of [19 Nottinghamshire 19] 's opening batsmen to help his side to a nine wicket victory . [20 Hussain 20] made his breakthrough for [21 Gloucestershire 21] in 2010 , appearing in all but one of their [22 County Championship 22] fixtures . He claimed his maiden five-wicket haul in bowler-friendly conditions at the [23 7 County Ground 23] in [15 15 Bristol 15] , taking five wickets for 36 against [16 16 Northamptonshire 16] . He picked up wickets regularly through the season , and claimed a second five-wicket haul against [17 17 Middlesex 17] in late April . [24 Hussain 24] finished the season with 67 first-class wickets , the second highest total in the [25 County Championship 25] , and the highest by an [26 England 26] qualified player . At the end of the season , [27 Hussain 27] rejected the offer of a new contract from [28 Gloucestershire 28] , stating that he wanted to play in the first division of the [29 County Championship 29] . He signed a three-year contract with [18 18 Somerset 18] in late October , joining his former [30 Gloucestershire 30] team-mate [19 19 Steven Kirby 19]

- 1: 9316 England
- 2: 1622178 Somerset_County_Cricket_Club
- 3: 1622322 Gloucestershire_County_Cricket_Club
- 4: 94458 London_Borough_of_Waltham_Forest
- 5: 17867 London
- 6: 1118651 Sports_science
- 7: 196911 University_of_Leeds
- 8: 1622252 Essex_County_Cricket_Club
- 9: 667341 Nottinghamshire_County_Cricket_Club
- 10: 1156428 Worcestershire_County_Cricket_Club
- 11: 1755522 Jack_Russell_(cricketer_and_artist)
- 12: 2358023 Stuart_Barnes
- 13: 683599 County_Championship

14: 1177719 Kent_County_Cricket_Club
15: 36741 Bristol
16: 530308 Northamptonshire_County_Cricket_Club
17: 822103 Middlesex_County_Cricket_Club
18: 1622178 Somerset_County_Cricket_Club
19: 10363998 Steven_Kirby
1: 25675557 Cricket
[winner: 330091 Raphidophoridae, pem_pos=10,
final_scores=-1.4120464, final_local_scores=-1.3590524,
pem=-2.56456, similarity=-22.520649|
gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.5325767, final_local_scores=-1.7210002,
pem=-0.07532314, similarity=-46.858746]
2: 678695 Fast_bowling
[winner: 678695 Fast_bowling, pem_pos=0,
final_scores=-1.1109427, final_local_scores=-1.2737792,
pem=-0.16673918, similarity=-33.015644|
gt: 678695 Fast_bowling, pem_pos=0,
final_scores=-1.1109427, final_local_scores=-1.2737792,
pem=-0.16673918, similarity=-33.015644]
3: 26970049 2010_English_cricket_season
[winner: 43226 2010, pem_pos=0, final_scores=-0.88961273,
final_local_scores=-0.77746814, pem=-1.6731367,
similarity=-10.012711|
RECALL MISS: 26970049 2010_English_cricket_season]
[winner: 1340236 Eastertide, pem_pos=2,
final_scores=-1.5897387, final_local_scores=-1.592896,
pem=-2.3723626, similarity=-30.5442|
RECALL MISS: 26970049 2010_English_cricket_season]
4: 20758829 2009_Twenty20_Cup
[winner: 771758 Twenty20_Cup, pem_pos=0,
final_scores=0.025324002, final_local_scores=-0.0013016313,
pem=0.0, similarity=4.0508695|
gt: 20758829 2009_Twenty20_Cup, pem_pos=1,
final_scores=-0.62734073, final_local_scores=-0.66136855,
pem=-3.5995927, similarity=3.9283006]
[winner: 35983 2009, pem_pos=0, final_scores=-0.85954285,
final_local_scores=-0.7751969, pem=-0.9296576,
similarity=-13.988087|
RECALL MISS: 20758829 2009_Twenty20_Cup]
5: 2298801 Batting_order_(cricket)
[winner: 2298801 Batting_order_(cricket), pem_pos=0,
final_scores=-1.2368971, final_local_scores=-1.4185815,
pem=0.0, similarity=-38.243587|
gt: 2298801 Batting_order_(cricket), pem_pos=0,
final_scores=-1.2368971, final_local_scores=-1.4185815,
pem=0.0, similarity=-38.243587]
6: 683459 List_A_cricket
[winner: 683459 List_A_cricket, pem_pos=0,
final_scores=0.12957565, final_local_scores=0.07748367,
pem=-0.020267159, similarity=6.512196|
gt: 683459 List_A_cricket, pem_pos=0,

final_scores=0.12957565, final_local_scores=0.07748367,
pem=-0.020267159, similarity=6.512196]
[winner: 330091 Rhaphidophoridae, pem_pos=10,
final_scores=-1.479542, final_local_scores=-1.4327812,
pem=-2.56456, similarity=-24.720861]
RECALL MISS: 683459 List_A_cricket]
7: 2642664 County_Cricket_Ground,_Bristol
[winner: 2646294 County_Ground,_Taunton, pem_pos=1,
final_scores=0.090020075, final_local_scores=0.039886445,
pem=-1.6581308, similarity=14.297176]
gt: 2642664 County_Cricket_Ground,_Bristol, pem_pos=3,
final_scores=0.013504371, final_local_scores=-0.01886642,
pem=-1.8513076, similarity=13.5944]
1: 9316 England
[pem_pos=1, final_scores=0.5642028,
final_local_scores=0.5949771, pem=-1.3527626,
similarity=29.201548]
2: 1622178 Somerset_County_Cricket_Club
[pem_pos=0, final_scores=0.42425543,
final_local_scores=0.35900983, pem=0.0, similarity=14.803283]
3: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.4493078,
final_local_scores=0.3836378, pem=-1.379793,
similarity=23.04176]
4: 94458 London_Borough_of_Waltham_Forest
[pem_pos=0, final_scores=0.19361113,
final_local_scores=0.22612613, pem=-0.17270607,
similarity=11.77697]
5: 17867 London
[pem_pos=0, final_scores=0.54444677,
final_local_scores=0.6015564, pem=-0.09086737,
similarity=22.535507]
6: 1118651 Sports_science
[pem_pos=0, final_scores=0.0397415,
final_local_scores=0.109099805, pem=-0.097543165,
similarity=7.8759236]
7: 196911 University_of_Leeds
[pem_pos=0, final_scores=0.29094094,
final_local_scores=0.33959085, pem=0.0, similarity=14.223782]
8: 1622252 Essex_County_Cricket_Club
[pem_pos=1, final_scores=0.7420481,
final_local_scores=0.70808804, pem=-2.3459642,
similarity=37.97819]
9: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.62185186,
final_local_scores=0.57674485, pem=-1.1067253,
similarity=27.319473]
10: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.54727715,
final_local_scores=0.49617052, pem=-0.9164197,
similarity=23.880064]
11: 1755522 Jack_Russell_(cricketer_and_artist)

[pem_pos=7, final_scores=0.11757605,
final_local_scores=0.13752508, pem=-1.9416747,
similarity=18.752865]
12: 2358023 Stuart_Barnes
[pem_pos=0, final_scores=0.16575913,
final_local_scores=0.20294243, pem=0.0, similarity=10.14592]
13: 683599 County_Championship
[pem_pos=0, final_scores=0.30677778,
final_local_scores=0.24108407, pem=0.0, similarity=11.2841425]
14: 1177719 Kent_County_Cricket_Club
[pem_pos=1, final_scores=0.66414005,
final_local_scores=0.62461334, pem=-2.5096242,
similarity=36.377148]
15: 36741 Bristol
[pem_pos=0, final_scores=0.44433665,
final_local_scores=0.4573124, pem=-0.43489033,
similarity=20.101828]
16: 530308 Northamptonshire_County_Cricket_Club
[pem_pos=1, final_scores=0.44956356,
final_local_scores=0.39318824, pem=-1.2069724,
similarity=22.386936]
17: 822103 Middlesex_County_Cricket_Club
[pem_pos=1, final_scores=0.7186187,
final_local_scores=0.67814004, pem=-1.5695794,
similarity=32.86238]
18: 1622178 Somerset_County_Cricket_Club
[pem_pos=1, final_scores=0.6015026,
final_local_scores=0.55262554, pem=-2.4139037,
similarity=33.708344]
19: 10363998 Steven_Kirby
[pem_pos=0, final_scores=0.4256081,
final_local_scores=0.4800533, pem=0.0, similarity=18.41546]
1: 28412046 Gemaal_Hussain
[pem_pos=0, final_scores=0.121654645,
final_local_scores=0.1012442, pem=0.0, similarity=7.1110415]
2: 1622178 Somerset_County_Cricket_Club
[pem_pos=1, final_scores=0.43057007,
final_local_scores=0.36569977, pem=-2.4139037,
similarity=28.130106]
3: 682925 Nasser_Hussain
[pem_pos=5, final_scores=0.17065613,
final_local_scores=0.14190504, pem=-2.447519,
similarity=21.634432]
4: 39470 Nottingham
[pem_pos=0, final_scores=0.5021494,
final_local_scores=0.51280016, pem=-0.29126778,
similarity=20.976652]
5: 23861381 Bradford
[pem_pos=0, final_scores=0.03407605,
final_local_scores=0.019897088, pem=-0.6938952,
similarity=8.456992]
6: 1622252 Essex_County_Cricket_Club

[pem_pos=1, final_scores=0.57297057,
final_local_scores=0.5233965, pem=-2.3459642,
similarity=32.466625]
7: 766575 Surrey_County_Cricket_Club
[pem_pos=1, final_scores=0.7735252,
final_local_scores=0.7407834, pem=-1.7952642,
similarity=35.95909]
8: 1622252 Essex_County_Cricket_Club
[pem_pos=1, final_scores=0.6438709,
final_local_scores=0.6008444, pem=-2.3459642,
similarity=34.777824]
9: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.6072052,
final_local_scores=0.5616328, pem=-0.9164197,
similarity=25.83359]
10: 3761368 West_Midlands_(Regional)_League
[pem_pos=0, final_scores=0.12500127,
final_local_scores=0.12864378, pem=0.0, similarity=7.9287]
11: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.41950417,
final_local_scores=0.35108185, pem=-1.379793,
similarity=22.070225]
12: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.57993704,
final_local_scores=0.5263304, pem=-1.379793,
similarity=27.299992]
13: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.7058918,
final_local_scores=0.66943294, pem=-0.9164197,
similarity=29.05056]
14: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.57765406,
final_local_scores=0.5238366, pem=-1.379793,
similarity=27.225573]
15: 771758 Twenty20_Cup
[pem_pos=0, final_scores=0.08785431,
final_local_scores=0.07050182, pem=0.0, similarity=6.193628]
16: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.661486,
final_local_scores=0.61541027, pem=-1.379793,
similarity=29.958313]
17: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.6423256,
final_local_scores=0.5944804, pem=-1.379793,
similarity=29.333723]
18: 683459 List_A_cricket
[pem_pos=0, final_scores=0.12957565,
final_local_scores=0.07748367, pem=-0.020267159,
similarity=6.512196]
19: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6928914,
final_local_scores=0.65434474, pem=-1.1067253,

```

similarity=29.63521]
20: 682925 Nasser_Hussain
    [pem_pos=5, final_scores=0.3491776,
    final_local_scores=0.3369126, pem=-2.447519,
    similarity=27.453846]
21: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.59725165,
    final_local_scores=0.545244, pem=-1.379793,
    similarity=27.86441]
22: 683599 County_Championship
    [pem_pos=0, final_scores=0.22599922,
    final_local_scores=0.15284577, pem=0.0, similarity=8.650935]
23: 2646294 County_Ground,_Taunton
    [pem_pos=1, final_scores=0.090020075,
    final_local_scores=0.039886445, pem=-1.6581308,
    similarity=14.297176]
24: 682925 Nasser_Hussain
    [pem_pos=5, final_scores=0.34803116,
    final_local_scores=0.3356603, pem=-2.447519,
    similarity=27.416475]
25: 683599 County_Championship
    [pem_pos=0, final_scores=0.28673714,
    final_local_scores=0.21919274, pem=0.0, similarity=10.630861]
26: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.59731233,
    final_local_scores=0.5645669, pem=-3.3054674,
    similarity=38.913162]
27: 682925 Nasser_Hussain
    [pem_pos=5, final_scores=0.2520758,
    final_local_scores=0.23084363, pem=-2.447519,
    similarity=24.288536]
28: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.5473331,
    final_local_scores=0.49071556, pem=-1.379793,
    similarity=26.237173]
29: 683599 County_Championship
    [pem_pos=0, final_scores=0.24524383,
    final_local_scores=0.17386758, pem=0.0, similarity=9.278269]
30: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.61555344,
    final_local_scores=0.56523585, pem=-1.379793,
    similarity=28.461008]

```

Sample: b'Tim_Hemp'

Timothy James Hemp -LRB- born 12 April 1974 -RRB- is a former [1 1 Bermudian 1] 2 cricket er . [2 Hemp 2] was a right-handed who bowled right-arm 3 medium pace . He was born at [3 4 Hamilton 3] , [1 1 Bermuda 1] . [4 Hemp 4] made his debut for [2 2 Wales Minor Counties 2] in the 1995 against [3 3 Dorset 3] . He made twelve further [5 Minor Counties Championship 5] appearances for the team , the last of which came against [4 4 Cheshire 4] in 1999 . His [5 5 MCCA Knockout

Trophy 5] debut for the team came in 1996 against [6 6 Cornwa
ll 6] . He represented the team in a total of five [6 MCCA
Knockout Trophy 6] matches , the last of which came against
the [7 7 Warwickshire Cricket Board 7] in 1999 . His only [8 8
List A 8] appearance for the team came in the 3rd round of
the 5 1999 [7 NatWest Trophy 7] against [9 9 Somerset 9] ,
scoring a single run before being dismissed by [10 10 Andrew C
addick 10] . His brother , 6 David , played [11 11 One Day In
ternational 11] and 7 Twenty20 cricket for , as well as
playing 8 first-class cricket for [12 12 Warwickshire 12] and
[13 13 Glamorgan 13]

- 1: 3460 Bermuda
- 2: 3712706 Wales_Minor_Counties_Cricket_Club
- 3: 3712026 Dorset_County_Cricket_Club
- 4: 3712477 Cheshire_County_Cricket_Club
- 5: 7543416 MCCA_Knockout_Trophy
- 6: 3702042 Cornwall_County_Cricket_Club
- 7: 27793461 Warwickshire_Cricket_Board
- 8: 683459 List_A_cricket
- 9: 1622178 Somerset_County_Cricket_Club
- 10: 896396 Andrew_Caddick
- 11: 381447 One_Day_International
- 12: 667292 Warwickshire_County_Cricket_Club
- 13: 1622320 Glamorgan_County_Cricket_Club

1: 68982 Demographics_of_Bermuda
[winner: 3460 Bermuda, pem_pos=0, final_scores=0.0455219,
final_local_scores=0.08570388, pem=-0.26592073,
similarity=8.093406]

gt: 68982 Demographics_of_Bermuda, pem_pos=5,
final_scores=-1.169477, final_local_scores=-1.1715086,
pem=-6.144901, similarity=2.5464852]

2: 25675557 Cricket

[winner: 330091 Raphidophoridae, pem_pos=10,
final_scores=-1.5308079, final_local_scores=-1.5030468,
pem=-2.56456, similarity=-26.817724]

gt: 25675557 Cricket, pem_pos=0,
final_scores=-2.2596262, final_local_scores=-2.5100622,
pem=-0.07532314, similarity=-70.40593]

3: 689836 Seam_bowling

[winner: 689836 Seam_bowling, pem_pos=0,
final_scores=-0.99504906, final_local_scores=-1.1247237,
pem=-0.53094965, similarity=-26.5869]

gt: 689836 Seam_bowling, pem_pos=0,
final_scores=-0.99504906, final_local_scores=-1.1247237,
pem=-0.53094965, similarity=-26.5869]

4: 727735 Hamilton,_Bermuda

[winner: 86654 Hamilton,_New_Zealand, pem_pos=1,
final_scores=0.032908186, final_local_scores=0.07311061,
pem=-2.40051, similarity=19.325827]

gt: 727735 Hamilton,_Bermuda, pem_pos=3,
final_scores=-0.40704626, final_local_scores=-0.37764433,
pem=-3.4693544, similarity=11.686941]

5: 27690999 1999_NatWest_Trophy
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.3363958, final_local_scores=0.29127318,
 pem=0.0, similarity=12.781885]
 gt: 27690999 1999_NatWest_Trophy, pem_pos=1,
 final_scores=-0.33472127, final_local_scores=-0.41676992,
 pem=-3.3480875, similarity=9.859887]
 [winner: 34553 1999, pem_pos=0, final_scores=-1.2919855,
 final_local_scores=-1.2842267, pem=-0.6240096,
 similarity=-30.840712]
 RECALL MISS: 27690999 1999_NatWest_Trophy]

6: 3326832 David_Hemp
 [winner: 8551 David, pem_pos=0, final_scores=-0.3690067,
 final_local_scores=-0.25217503, pem=-0.8369082,
 similarity=1.1155441]
 RECALL MISS: 3326832 David_Hemp]

7: 5599837 Twenty20_International
 [winner: 770690 Twenty20, pem_pos=0,
 final_scores=-0.051917568, final_local_scores=-0.114853084,
 pem=-0.039191484, similarity=0.8753973]
 gt: 5599837 Twenty20_International, pem_pos=4,
 final_scores=-0.7041042, final_local_scores=-0.80422825,
 pem=-3.6661954, similarity=0.02727919]

8: 11040 First-class_cricket
 [winner: 662844 First-class_citizen, pem_pos=3,
 final_scores=-1.2308089, final_local_scores=-1.1507912,
 pem=-1.5935546, similarity=-21.586193]
 gt: 11040 First-class_cricket, pem_pos=0,
 final_scores=-1.2359412, final_local_scores=-1.4207165,
 pem=-0.053819276, similarity=-38.01462]
 [winner: 330091 Rhaphidophoridae, pem_pos=10,
 final_scores=-1.368862, final_local_scores=-1.3261456,
 pem=-2.56456, similarity=-21.538643]
 RECALL MISS: 11040 First-class_cricket]

1: 3460 Bermuda
 [pem_pos=0, final_scores=0.12058495,
 final_local_scores=0.16769886, pem=-0.4078151,
 similarity=11.311943]

2: 3712706 Wales_Minor_Counties_Cricket_Club
 [pem_pos=0, final_scores=0.36924708,
 final_local_scores=0.32789436, pem=0.0, similarity=13.874735]

3: 3712026 Dorset_County_Cricket_Club
 [pem_pos=1, final_scores=0.36842054,
 final_local_scores=0.3212731, pem=-2.4537704,
 similarity=27.021128]

4: 3712477 Cheshire_County_Cricket_Club
 [pem_pos=2, final_scores=0.17772426,
 final_local_scores=0.12108418, pem=-2.629504,
 similarity=22.002758]

5: 7543416 MCCA_Knockout_Trophy
 [pem_pos=0, final_scores=0.38393277,
 final_local_scores=0.3413706, pem=0.0, similarity=14.276893]

6: 3702042 Cornwall_County_Cricket_Club
[pem_pos=2, final_scores=0.21103118,
final_local_scores=0.14706326, pem=-3.34357,
similarity=26.661228]
7: 27793461 Warwickshire_Cricket_Board
[pem_pos=0, final_scores=0.3059531,
final_local_scores=0.3035136, pem=0.0, similarity=13.147163]
8: 683459 List_A_cricket
[pem_pos=0, final_scores=0.24139805,
final_local_scores=0.17790282, pem=-0.020267159,
similarity=9.5089035]
9: 1622178 Somerset_County_Cricket_Club
[pem_pos=1, final_scores=0.6547635,
final_local_scores=0.6143017, pem=-2.4139037,
similarity=35.548885]
10: 896396 Andrew_Caddick
[pem_pos=0, final_scores=0.6492383,
final_local_scores=0.6973101, pem=0.0, similarity=24.898836]
11: 381447 One_Day_International
[pem_pos=1, final_scores=0.1345477,
final_local_scores=0.09822565, pem=-0.022472855,
similarity=7.143173]
12: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.65297073,
final_local_scores=0.6149309, pem=-1.1093602,
similarity=28.473352]
13: 1622320 Glamorgan_County_Cricket_Club
[pem_pos=1, final_scores=0.45769513,
final_local_scores=0.40397078, pem=-0.7903395,
similarity=20.442995]
1: 3460 Bermuda
[pem_pos=0, final_scores=0.0455219,
final_local_scores=0.08570388, pem=-0.26592073,
similarity=8.093406]
2: 3326832 David_Hemp
[pem_pos=2, final_scores=0.12910362,
final_local_scores=0.12209207, pem=-2.1604607,
similarity=19.482105]
3: 86654 Hamilton,_New_Zealand
[pem_pos=1, final_scores=0.032908186,
final_local_scores=0.07311061, pem=-2.40051,
similarity=19.325827]
4: 3326832 David_Hemp
[pem_pos=2, final_scores=0.24874263,
final_local_scores=0.25277954, pem=-2.1604607,
similarity=23.38208]
5: 7562952 Minor_Counties_Cricket_Championship
[pem_pos=0, final_scores=0.32320923,
final_local_scores=0.27280912, pem=0.0, similarity=12.230881]
6: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.285509,
final_local_scores=0.23385757, pem=0.0, similarity=11.068489]

7: 1899702 Friends_Provident_Trophy
[pem_pos=0, final_scores=0.3363958,
final_local_scores=0.29127318, pem=0.0, similarity=12.781885]

Sample: b'Gerwyn_Edwards'

William Gerwyn Edwards -LRB- born 22 June 1953 -RRB- is a former [1 1 Welsh 1] 2 cricket er . [2 Edwards 2] was a left-handed who bowled left-arm 3 medium pace . He was born in [1 1 Glanamman 1] , 4 Carmarthenshire . [3 Edwards 3] represented in the 5 1979 [4 ICC Trophy 4] , playing two matches against the [2 2 Netherlands 2] and 6 Israel . In 1988 , he made his debut for [3 3 Wales Minor Counties 3] in the against [4 4 Buckinghamshire 4] . From 1988 to 1995 , he represented the team in twenty-eight Championship matches , the last of which came against [5 5 Cornwall 5] . He played just the two [6 6 MCC Knockout Trophy 6] fixtures for the team , against [7 7 Staffordshire 7] and [8 8 Cumberland 8] in 1994 and 1995 respectively . It was also in 1994 that he made his [9 9 List A 9] debut for the team against [10 10 Sussex 10] in the 7 1993 [5 NatWest Trophy 5] 1st round , following that up the following season with his second and final appearance in that format against [11 11 Middlesex 11] . In his two [6 List A 6] matches he scored 33 runs at a 8 batting average of 33 , with a high score of 28 9 * . With the ball he took a single wicket at a 10 bowling average

1: 663562 Glanamman

2: 2164343 Netherlands_national_cricket_team

3: 3712706 Wales_Minor_Counties_Cricket_Club

4: 3702814 Buckinghamshire_County_Cricket_Club

5: 3702042 Cornwall_County_Cricket_Club

6: 7543416 MCCA_Knockout_Trophy

7: 3712165 Staffordshire_County_Cricket_Club

8: 3702699 Cumberland_County_Cricket_Club

9: 683459 List_A_cricket

10: 1763691 Sussex_County_Cricket_Club

11: 822103 Middlesex_County_Cricket_Club

1: 2030193 Welsh_people

[winner: 69894 Wales, pem_pos=1, final_scores=0.40250915,
final_local_scores=0.45613742, pem=-1.0776843,
similarity=23.562378]

gt: 2030193 Welsh_people, pem_pos=2,
final_scores=-0.2884282, final_local_scores=-0.2920214,
pem=-1.6894228, similarity=4.5625563]

2: 25675557 Cricket

[winner: 330091 Raphidophoridae, pem_pos=10,
final_scores=-1.5025644, final_local_scores=-1.4673717,
pem=-2.56456, similarity=-25.75311]

gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.9018493, final_local_scores=-2.1154003,
pem=-0.07532314, similarity=-58.628437]

3: 689836 Seam_bowling

[winner: 689836 Seam_bowling, pem_pos=0,


```

final_scores=-0.9874957, final_local_scores=-1.1111841,
pem=-0.53094965, similarity=-26.18285|
    gt: 689836 Seam_bowling, pem_pos=0,
final_scores=-0.9874957, final_local_scores=-1.1111841,
pem=-0.53094965, similarity=-26.18285 ]
4: 154250 Carmarthenshire
    [winner: 154250 Carmarthenshire, pem_pos=0,
final_scores=-0.010089025, final_local_scores=0.045931324,
pem=0.0, similarity=5.4603953|
    gt: 154250 Carmarthenshire, pem_pos=0,
final_scores=-0.010089025, final_local_scores=0.045931324,
pem=0.0, similarity=5.4603953 ]
5: 2675465 1979_ICC_Trophy
    [winner: 1952482 ICC_World_Cup_Qualifier, pem_pos=0,
final_scores=0.28834087, final_local_scores=0.29482713,
pem=0.0, similarity=12.887942|
    gt: 2675465 1979_ICC_Trophy, pem_pos=1,
final_scores=-0.075586006, final_local_scores=-0.08084606,
pem=-1.3137479, similarity=8.82147 ]
    [winner: 34754 1979, pem_pos=0, final_scores=-0.38441676,
final_local_scores=-0.1544362, pem=-0.65922886,
similarity=3.0660176|
    RECALL MISS: 2675465 1979_ICC_Trophy]
6: 4399971 Israel_national_cricket_team
    [winner: 9282173 Israel, pem_pos=0,
final_scores=-0.20963001, final_local_scores=-0.12490463,
pem=-0.21279564, similarity=1.5195254|
    gt: 4399971 Israel_national_cricket_team, pem_pos=7,
final_scores=-0.40958977, final_local_scores=-0.42340088,
pem=-4.876911, similarity=17.975986 ]
7: 27614039 1993_NatWest_Trophy
    [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=0.43272668, final_local_scores=0.39658612,
pem=0.0, similarity=15.924633|
    gt: 27614039 1993_NatWest_Trophy, pem_pos=3,
final_scores=-0.27801108, final_local_scores=-0.33604273,
pem=-3.670861, similarity=14.02424 ]
    [winner: 34598 1993, pem_pos=0, final_scores=-0.7017104,
final_local_scores=-0.5178966, pem=-0.7063994,
similarity=-7.5238442|
    RECALL MISS: 27614039 1993_NatWest_Trophy]
8: 18933273 Batting_average
    [winner: 15652278 Batting_(material), pem_pos=2,
final_scores=-1.040444, final_local_scores=-0.9375808,
pem=-1.1616125, similarity=-17.572538|
    gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.3081609, final_local_scores=-1.4512119,
pem=-0.29142123, similarity=-37.632553 ]
    [winner: 875871 Bowling_average, pem_pos=1,
final_scores=-1.2076042, final_local_scores=-1.3578224,
pem=-0.8657023, similarity=-31.722591|
    gt: 18933273 Batting_average, pem_pos=0,

```

final_scores=-1.3458402, final_local_scores=-1.4923708,
pem=-0.5012068, similarity=-37.71997]
9: 2430357 Not_out
[winner: 2430357 Not_out, pem_pos=0,
final_scores=-0.22935204, final_local_scores=-0.3069914,
pem=-0.3529179, similarity=-3.1523006]
gt: 2430357 Not_out, pem_pos=0,
final_scores=-0.22935204, final_local_scores=-0.3069914,
pem=-0.3529179, similarity=-3.1523006]
10: 875871 Bowling_average
[winner: 875871 Bowling_average, pem_pos=0,
final_scores=-1.0768214, final_local_scores=-1.2149622,
pem=0.0, similarity=-32.167187]
gt: 875871 Bowling_average, pem_pos=0,
final_scores=-1.0768214, final_local_scores=-1.2149622,
pem=0.0, similarity=-32.167187]
1: 663562 Glanamman
[pem_pos=0, final_scores=0.20528726,
final_local_scores=0.33273557, pem=0.0, similarity=14.019206]
2: 2164343 Netherlands_national_cricket_team
[pem_pos=2, final_scores=0.115847155,
final_local_scores=0.07441878, pem=-3.8293223,
similarity=27.134968]
3: 3712706 Wales_Minor_Counties_Cricket_Club
[pem_pos=0, final_scores=0.38124424,
final_local_scores=0.33841914, pem=0.0, similarity=14.188815]
4: 3702814 Buckinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.35019273,
final_local_scores=0.29527673, pem=-1.6046519,
similarity=21.627707]
5: 3702042 Cornwall_County_Cricket_Club
[pem_pos=2, final_scores=0.22674142,
final_local_scores=0.1615501, pem=-3.34357,
similarity=27.093544]
6: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.32824117,
final_local_scores=0.28217405, pem=0.0, similarity=12.510349]
7: 3712165 Staffordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.3062166,
final_local_scores=0.25307474, pem=-1.68797,
similarity=20.821411]
8: 3702699 Cumberland_County_Cricket_Club
[pem_pos=5, final_scores=0.06993152,
final_local_scores=0.005992353, pem=-3.2953885,
similarity=22.189371]
9: 683459 List_A_cricket
[pem_pos=0, final_scores=0.31865007,
final_local_scores=0.26145625, pem=-0.020267159,
similarity=12.002305]
10: 1763691 Sussex_County_Cricket_Club
[pem_pos=1, final_scores=0.6732637,
final_local_scores=0.6608663, pem=-1.5566913,

```

similarity=32.276814]
11: 822103 Middlesex_County_Cricket_Club
    [pem_pos=1, final_scores=0.6502642,
final_local_scores=0.6163856, pem=-1.5695794,
similarity=31.019505]
1: 69894 Wales
    [pem_pos=1, final_scores=0.40250915,
final_local_scores=0.45613742, pem=-1.0776843,
similarity=23.562378]
2: 30853526 Gerwyn_Edwards
    [pem_pos=0, final_scores=0.2622003,
final_local_scores=0.27686468, pem=0.0, similarity=12.351907]
3: 30853526 Gerwyn_Edwards
    [pem_pos=0, final_scores=0.21561123,
final_local_scores=0.22597319, pem=0.0, similarity=10.833203]
4: 1952482 ICC_World_Cup_Qualifier
    [pem_pos=0, final_scores=0.28834087,
final_local_scores=0.29482713, pem=0.0, similarity=12.887942]
5: 1899702 Friends_Provident_Trophy
    [pem_pos=0, final_scores=0.43272668,
final_local_scores=0.39658612, pem=0.0, similarity=15.924633]
6: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.08465473,
final_local_scores=0.0058518946, pem=-0.020267159,
similarity=4.374561]

```

Sample: b'Thomas_Pearsall'

-LRB- cricketer -RRB- Thomas A Pearsall -LRB- born 18 May 1943 -RRB- is a former [1 1 English 1] 2 cricket er . Pearsall was a left-handed who was a right-arm 3 bowler , but his bowling style is unknown . He was born in [1 1 West Bromwich 1] , [2 4 Staffordshire 2] . Pearsall made his debut for [2 2 Staffordshire 2] in the 1974 against [3 3 Shropshire 3] . Pearsall played 5 Minor counties cricket for [3 Staffordshire 3] from 1974 to 1981 , which included 29 [4 Minor Counties Championship 4] matches . In 1975 , he made his [4 4 List A 4] debut for [5 Staffordshire 5] against [5 5 Leicestershire 5] in the [6 6 Gillette Cup 6] . He made 2 further appearances in [7 List A 7] cricket for the county , against [6 6 Devon 6] in the 1st round of the 7 1978 [8 Gillette Cup 8] and [7 7 Sussex 7] in the 2nd round of the same competition . In his 3 [9 List A 9] matches for the county , he scored 49 runs at an 8 average of 16.33 , with a high score of 34 . He later made 2 [1 0 List A 10] appearances for the in the 9 1979 [11 Benson & Hedges Cup 11] against [8 8 Middlesex 8] and [9 9 Kent 9]

```

1: 263460 West_Bromwich
2: 3712165 Staffordshire_County_Cricket_Club
3: 3702447 Shropshire_County_Cricket_Club
4: 683459 List_A_cricket
5: 1622318 Leicestershire_County_Cricket_Club
6: 3691070 Devon_County_Cricket_Club
7: 1763691 Sussex_County_Cricket_Club

```

8: 822103 Middlesex_County_Cricket_Club
9: 1177719 Kent_County_Cricket_Club
1: 18803164 English_people
[winner: 9316 England, pem_pos=1,
final_scores=0.48231268, final_local_scores=0.5266687,
pem=-1.3527626, similarity=27.16309|
gt: 18803164 English_people, pem_pos=2,
final_scores=-0.4512692, final_local_scores=-0.48431343,
pem=-2.2084396, similarity=1.6466715]
2: 25675557 Cricket
[winner: 330091 Rhaphidophoridae, pem_pos=10,
final_scores=-1.4412142, final_local_scores=-1.3965054,
pem=-2.56456, similarity=-23.638317|
gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.7174383, final_local_scores=-1.9134173,
pem=-0.07532314, similarity=-52.600857]
3: 51878 Bowling_(cricket)
[winner: 51878 Bowling_(cricket), pem_pos=0,
final_scores=-1.3864517, final_local_scores=-1.5291488,
pem=-0.13362998, similarity=-40.81644|
gt: 51878 Bowling_(cricket), pem_pos=0,
final_scores=-1.3864517, final_local_scores=-1.5291488,
pem=-0.13362998, similarity=-40.81644]
4: 58733 Staffordshire
[winner: 3712165 Staffordshire_County_Cricket_Club,
pem_pos=1, final_scores=0.1709464,
final_local_scores=0.09352043, pem=-1.68797,
similarity=16.059992|
gt: 58733 Staffordshire, pem_pos=0,
final_scores=0.015169039, final_local_scores=0.010112122,
pem=-0.06319042, similarity=4.735118]
5: 770909 Minor_counties_of_English_and_Welsh_cricket
[winner: 897385 Morris_Minor, pem_pos=4,
final_scores=-0.4745585, final_local_scores=-0.4001799,
pem=-2.4046454, similarity=5.2243824|
RECALL MISS: 770909
Minor_counties_of_English_and_Welsh_cricket]
[winner: 480042 Counties_of_Liberia, pem_pos=4,
final_scores=-1.3195211, final_local_scores=-1.2319372,
pem=-2.8773274, similarity=-17.026394|
RECALL MISS: 770909
Minor_counties_of_English_and_Welsh_cricket]
6: 27364483 1975_Gillette_Cup
[winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=0.31070948, final_local_scores=0.26124188,
pem=-0.08377493, similarity=12.341273|
gt: 27364483 1975_Gillette_Cup, pem_pos=6,
final_scores=-0.49500048, final_local_scores=-0.5594331,
pem=-3.9435217, similarity=8.840597]
7: 27387455 1978_Gillette_Cup
[winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=0.39817566, final_local_scores=0.3567854,

pem=-0.08377493, similarity=15.192481|
gt: 27387455 1978_Gillette_Cup, pem_pos=12,
final_scores=-0.44891286, final_local_scores=-0.50790036,
pem=-4.2536764, similarity=12.065107]
[winner: 34753 1978, pem_pos=0, final_scores=-0.58818007,
final_local_scores=-0.37888697, pem=-0.7079138,
similarity=-3.367285|
RECALL MISS: 27387455 1978_Gillette_Cup]
8: 18933273 Batting_average
[winner: 875871 Bowling_average, pem_pos=0,
final_scores=-0.93120146, final_local_scores=-1.0639315,
pem=0.0, similarity=-27.660124|
gt: 18933273 Batting_average, pem_pos=1,
final_scores=-1.1956905, final_local_scores=-1.3322983,
pem=-0.6931472, similarity=-31.899286]
9: 37839460 1979_Benson_&_Hedges_Cup
[winner: 1895594 Benson_&_Hedges_Cup, pem_pos=0,
final_scores=0.3918032, final_local_scores=0.36615688,
pem=0.0, similarity=15.016565|
gt: 37839460 1979_Benson_&_Hedges_Cup, pem_pos=3,
final_scores=-0.5596134, final_local_scores=-0.60675395,
pem=-4.6821313, similarity=11.445123]
[winner: 34754 1979, pem_pos=0, final_scores=-0.3680351,
final_local_scores=-0.13820407, pem=-0.65922886,
similarity=3.5504167|
RECALL MISS: 37839460 1979_Benson_&_Hedges_Cup]
1: 263460 West_Bromwich
[pem_pos=0, final_scores=0.09764628,
final_local_scores=0.12944803, pem=0.0, similarity=7.9527]
2: 3712165 Staffordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.3305226,
final_local_scores=0.26849625, pem=-1.68797,
similarity=21.28162]
3: 3702447 Shropshire_County_Cricket_Club
[pem_pos=1, final_scores=0.2292646,
final_local_scores=0.1780954, pem=-2.279101,
similarity=21.79854]
4: 683459 List_A_cricket
[pem_pos=0, final_scores=0.28431273,
final_local_scores=0.22391099, pem=-0.020267159,
similarity=10.881879]
5: 1622318 Leicestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.80009323,
final_local_scores=0.7788258, pem=-1.0465667,
similarity=33.022823]
6: 3691070 Devon_County_Cricket_Club
[pem_pos=1, final_scores=0.5110853,
final_local_scores=0.4689957, pem=-3.0304468,
similarity=34.56552]
7: 1763691 Sussex_County_Cricket_Club
[pem_pos=1, final_scores=0.6643191,
final_local_scores=0.6390421, pem=-1.5566913,

similarity=31.625534]
8: 822103 Middlesex_County_Cricket_Club
[pem_pos=1, final_scores=0.7251262,
final_local_scores=0.6881205, pem=-1.5695794,
similarity=33.160217]
9: 1177719 Kent_County_Cricket_Club
[pem_pos=1, final_scores=0.40677834,
final_local_scores=0.34340784, pem=-2.5096242,
similarity=27.985415]
1: 9316 England
[pem_pos=1, final_scores=0.48231268,
final_local_scores=0.5266687, pem=-1.3527626,
similarity=27.16309]
2: 3712165 Staffordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.1709464,
final_local_scores=0.09352043, pem=-1.68797,
similarity=16.059992]
3: 3712165 Staffordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.33053118,
final_local_scores=0.26850566, pem=-1.68797,
similarity=21.2819]
4: 7562952 Minor_Counties_Cricket_Championship
[pem_pos=0, final_scores=0.32007152,
final_local_scores=0.2617489, pem=0.0, similarity=11.900823]
5: 3712165 Staffordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.43870938,
final_local_scores=0.38667387, pem=-1.68797,
similarity=24.808275]
6: 1899702 Friends_Provident_Trophy
[pem_pos=0, final_scores=0.31070948,
final_local_scores=0.26124188, pem=-0.08377493,
similarity=12.341273]
7: 683459 List_A_cricket
[pem_pos=0, final_scores=0.08753435,
final_local_scores=0.008960515, pem=-0.020267159,
similarity=4.4673285]
8: 1899702 Friends_Provident_Trophy
[pem_pos=0, final_scores=0.39817566,
final_local_scores=0.3567854, pem=-0.08377493,
similarity=15.192481]
9: 683459 List_A_cricket
[pem_pos=0, final_scores=0.06999533,
final_local_scores=-0.014230497, pem=-0.020267159,
similarity=3.7752624]
10: 683459 List_A_cricket
[pem_pos=0, final_scores=0.2015134,
final_local_scores=0.13346532, pem=-0.020267159,
similarity=8.1828]
11: 1895594 Benson_&_Hedges_Cup
[pem_pos=0, final_scores=0.3918032,
final_local_scores=0.36615688, pem=0.0, similarity=15.016565]

Sample: b'Joseph_Smith'

-LRB- cricketer -RRB- Joseph Abraham Smith -LRB- born 23 March 1946 -RRB- is a born former [1 1 English 1] 2 cricketer . Smith was a left-handed who bowled left-arm 3 fast-medium . He was born in [2 4 Kingston 2] , Jamaica . Smith made his debut for [1 1 Shropshire 1] in the 1968 against [2 2 Cambridgeshire 2] . Smith played 5 Minor counties cricket for [3 Shropshire 3] from 1968 to 1988 , which included 113 [4 Minor Counties Championship 4] appearances and 8 [3 3 MCCA Knockout Trophy 3] appearances . He made his [4 4 List A 4] debut against [5 5 Yorkshire 5] in the 6 1976 [5 Gillette Cup 5] . He made 2 further [6 List A 6] appearances for [7 Shropshire 7] , against [6 6 Surrey 6] in the 7 1978 [8 Gillette Cup 8] and [7 7 Northamptonshire 7] in the 8 1985 [9 NatWest Trophy 9] . In his 3 [10 List A 10] matches for the county , he took 5 wickets at an 9 average of 22.60 , with best figures of 3/40 . Playing for [11 Shropshire 11] allowed him to play for the , with Smith making 2 appearances for that team in the against [8 8 Worcestershire 8] and [9 9 Leicestershire 9]

1: 3702447 Shropshire_County_Cricket_Club

2: 1986069 Cambridgeshire_County_Cricket_Club

3: 7543416 MCCA_Knockout_Trophy

4: 683459 List_A_cricket

5: 869945 Yorkshire_County_Cricket_Club

6: 766575 Surrey_County_Cricket_Club

7: 530308 Northamptonshire_County_Cricket_Club

8: 1156428 Worcestershire_County_Cricket_Club

9: 1622318 Leicestershire_County_Cricket_Club

1: 18803164 English_people

[winner: 9316 England, pem_pos=1, final_scores=0.5389306, final_local_scores=0.5913343, pem=-1.3527626, similarity=29.092842]

gt: 18803164 English_people, pem_pos=2, final_scores=-0.47165495, final_local_scores=-0.5042645, pem=-2.2084396, similarity=1.0512929]

2: 25675557 Cricket

[winner: 330091 Rhabdophoridae, pem_pos=10, final_scores=-1.4378238, final_local_scores=-1.3975441, pem=-2.56456, similarity=-23.669317]

gt: 25675557 Cricket, pem_pos=0, final_scores=-1.6711276, final_local_scores=-1.864496, pem=-0.07532314, similarity=-51.14095]

3: 678695 Fast_bowling

[winner: 678695 Fast_bowling, pem_pos=0, final_scores=-1.2248247, final_local_scores=-1.3904889, pem=0.0, similarity=-37.40525]

gt: 678695 Fast_bowling, pem_pos=0, final_scores=-1.2248247, final_local_scores=-1.3904889, pem=0.0, similarity=-37.40525]

4: 37408 Kingston,_Jamaica

[winner: 37408 Kingston,_Jamaica, pem_pos=0, final_scores=0.08878811, final_local_scores=0.136383,

pem=-1.3675016, similarity=15.59634|
 gt: 37408 Kingston, Jamaica, pem_pos=0,
 final_scores=0.08878811, final_local_scores=0.136383,
 pem=-1.3675016, similarity=15.59634]
 [winner: 15660 Jamaica, pem_pos=0,
 final_scores=0.007647559, final_local_scores=0.061556756,
 pem=-0.33295843, similarity=7.7373695|
 RECALL MISS: 37408 Kingston, Jamaica]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7918845, final_local_scores=-1.7639048,
 pem=-4.6728287, similarity=-23.137157|
 RECALL MISS: 37408 Kingston, Jamaica]
 5: 770909 Minor_counties_of_English_and_Welsh_cricket
 [winner: 312613 Minor_League_Baseball, pem_pos=1,
 final_scores=-0.448587, final_local_scores=-0.43479818,
 pem=-2.282223, similarity=3.5255518|
 RECALL MISS: 770909
 Minor_counties_of_English_and_Welsh_cricket]
 [winner: 4569028 Counties_of_Uganda, pem_pos=18,
 final_scores=-1.3148443, final_local_scores=-1.2637641,
 pem=-3.1761155, similarity=-16.35132|
 RECALL MISS: 770909
 Minor_counties_of_English_and_Welsh_cricket]
 6: 27372075 1976_Gillette_Cup
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.3154593, final_local_scores=0.26224867,
 pem=-0.08377493, similarity=12.371318|
 gt: 27372075 1976_Gillette_Cup, pem_pos=13,
 final_scores=-0.5723343, final_local_scores=-0.63326365,
 pem=-4.454347, similarity=9.415294]
 [winner: 34661 1976, pem_pos=0, final_scores=-0.90884393,
 final_local_scores=-0.81480646, pem=-0.7313423,
 similarity=-16.248583|
 RECALL MISS: 27372075 1976_Gillette_Cup]
 7: 27387455 1978_Gillette_Cup
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.3881163, final_local_scores=0.34161544,
 pem=-0.08377493, similarity=14.73978|
 gt: 27387455 1978_Gillette_Cup, pem_pos=12,
 final_scores=-0.43300903, final_local_scores=-0.49194658,
 pem=-4.2536764, similarity=12.5412]
 [winner: 34753 1978, pem_pos=0, final_scores=-0.5773091,
 final_local_scores=-0.36671785, pem=-0.7079138,
 similarity=-3.0041342|
 RECALL MISS: 27387455 1978_Gillette_Cup]
 8: 27478442 1985_NatWest_Trophy
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.31752545, final_local_scores=0.2642425,
 pem=0.0, similarity=11.975236|
 gt: 27478442 1985_NatWest_Trophy, pem_pos=4,
 final_scores=-0.35631186, final_local_scores=-0.43219116,
 pem=-3.7196512, similarity=11.420309]

[winner: 10503893 The_1985, pem_pos=0,
final_scores=-0.75973475, final_local_scores=-0.6258571,
pem=0.0, similarity=-14.587112]
RECALL MISS: 27478442 1985_NatWest_Trophy]
9: 875871 Bowling_average
[winner: 875871 Bowling_average, pem_pos=0,
final_scores=-0.9595236, final_local_scores=-1.096312,
pem=0.0, similarity=-28.626425]
gt: 875871 Bowling_average, pem_pos=0,
final_scores=-0.9595236, final_local_scores=-1.096312,
pem=0.0, similarity=-28.626425]
1: 3702447 Shropshire_County_Cricket_Club
[pem_pos=1, final_scores=0.14964785,
final_local_scores=0.084909916, pem=-2.279101,
similarity=19.017698]
2: 1986069 Cambridgeshire_County_Cricket_Club
[pem_pos=1, final_scores=0.5069784,
final_local_scores=0.46450704, pem=-1.9373415,
similarity=28.487095]
3: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.3190747,
final_local_scores=0.26789013, pem=0.0, similarity=12.084088]
4: 683459 List_A_cricket
[pem_pos=0, final_scores=0.38291067,
final_local_scores=0.3251143, pem=-0.020267159,
similarity=13.901988]
5: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6765565,
final_local_scores=0.63529617, pem=-1.7779728,
similarity=32.71711]
6: 766575 Surrey_County_Cricket_Club
[pem_pos=1, final_scores=0.7131375,
final_local_scores=0.6772745, pem=-1.7952642,
similarity=34.06386]
7: 530308 Northamptonshire_County_Cricket_Club
[pem_pos=1, final_scores=0.43625498,
final_local_scores=0.380906, pem=-1.2069724,
similarity=22.02041]
8: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.67530304,
final_local_scores=0.63668007, pem=-0.9164197,
similarity=28.073149]
9: 1622318 Leicestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.5483603,
final_local_scores=0.50773615, pem=-1.0465667,
similarity=24.932964]
1: 9316 England
[pem_pos=1, final_scores=0.5389306,
final_local_scores=0.5913343, pem=-1.3527626,
similarity=29.092842]
2: 37408 Kingston,_Jamaica
[pem_pos=0, final_scores=0.08878811,

final_local_scores=0.136383, pem=-1.3675016,
 similarity=15.59634]
 3: 3702447 Shropshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.113658234,
 final_local_scores=0.04559672, pem=-2.279101,
 similarity=17.844515]
 4: 7562952 Minor_Counties_Cricket_Championship
 [pem_pos=0, final_scores=0.37993658,
 final_local_scores=0.32924885, pem=0.0, similarity=13.915155]
 5: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.3154593,
 final_local_scores=0.26224867, pem=-0.08377493,
 similarity=12.371318]
 6: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.2987134,
 final_local_scores=0.23314157, pem=-0.020267159,
 similarity=11.157338]
 7: 3702447 Shropshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.17861564,
 final_local_scores=0.11655283, pem=-2.279101,
 similarity=19.961987]
 8: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.3881163,
 final_local_scores=0.34161544, pem=-0.08377493,
 similarity=14.73978]
 9: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.31752545,
 final_local_scores=0.2642425, pem=0.0, similarity=11.975236]
 10: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.10561855,
 final_local_scores=0.022214815, pem=-0.020267159,
 similarity=4.862863]
 11: 3702447 Shropshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.17352326,
 final_local_scores=0.11099017, pem=-2.279101,
 similarity=19.795986]

Sample: b'Leading_Creek'

_-LRB- [1 Ohio 1] -RRB- Leading Creek is a 1 tributary of
 the [1 1 Ohio River 1] , long , in southeastern [2 2 Ohio 2]
 in the [3 3 United States 3] . Via the [2 Ohio River 2] , it
 is part of the 2 watershed of the [4 4 Mississippi River 4] ,
 draining an area of on the 3 unglaciated portion of the [5 5 A
 llegheny Plateau 5] . The creek 's headwaters are in southern
 [6 6 Athens County 6] and it flows for most of its length in
 western [7 7 Meigs County 7] ; its tributaries also drain a
 small area of northeastern [8 8 Gallia County 8] . Leading
 Creek rises northwest of [3 4 Albany 3] in 5 Lee Township in [
 4 Athens County 4] and initially flows southward into [5
 Meigs County 5] , through 6 Columbia Township , past the
 community of 7 Carpenter ; then southeastward through [6 8 Sa
 lem 6] , 9 Rutland , and 10 Salisbury townships , through the

community of 11 Langsville . It flows into the [7 Ohio River 7] at the south end of the village of [8 12 Middleport 8] . Little Leading Creek is a tributary of Leading Creek and flows for its entire length in [9 Meigs County 9] . It rises at in 13 Scipio Township and flows southward through the village of 14 Rutland

1: 22388 Ohio_River

2: 22199 Ohio

3: 3434750 United_States

4: 19579 Mississippi_River

5: 609240 Allegheny_Plateau

6: 93140 Athens_County,_Ohio

7: 93051 Meigs_County,_Ohio

8: 93118 Gallia_County,_Ohio

1: 72465 Tributary

[winner: 72465 Tributary, pem_pos=1, final_scores=-1.1300678, final_local_scores=-1.1599934, pem=-0.012074053, similarity=-30.461145]

gt: 72465 Tributary, pem_pos=1, final_scores=-1.1300678, final_local_scores=-1.1599934, pem=-0.012074053, similarity=-30.461145]

2: 191162 Drainage_basin

[winner: 191162 Drainage_basin, pem_pos=0, final_scores=-1.3543252, final_local_scores=-1.4673097, pem=0.0, similarity=-39.697735]

gt: 191162 Drainage_basin, pem_pos=0, final_scores=-1.3543252, final_local_scores=-1.4673097, pem=0.0, similarity=-39.697735]

3: 12463 Glacier

[winner: 603172 Unglaciaded_Allegheny_Plateau, pem_pos=0, final_scores=-0.57633644, final_local_scores=-0.5707543, pem=0.0, similarity=-12.942734]

gt: 12463 Glacier, pem_pos=1, final_scores=-1.0686312, final_local_scores=-1.1244876, pem=-0.71116567, similarity=-25.599812]

4: 128911 Albany,_Ohio

[winner: 52106 Albany,_New_York, pem_pos=0, final_scores=0.046634242, final_local_scores=0.018284753, pem=-1.175974, similarity=11.030496]

RECALL MISS: 128911 Albany,_Ohio]

5: 9292871 Lee_Township,_Athens_County,_Ohio

[winner: 117009 Lee_Township,_Allegan_County,_Michigan, pem_pos=2, final_scores=-0.24844241, final_local_scores=-0.22990401, pem=-1.766495, similarity=6.8353934]

gt: 9292871 Lee_Township,_Athens_County,_Ohio, pem_pos=5, final_scores=-0.29673818, final_local_scores=-0.25243318, pem=-1.9943272, similarity=7.4020643]

6: 9273698 Columbia_Township,_Meigs_County,_Ohio

[winner: 3889803 Columbia_Township,_Lorain_County,_Ohio, pem_pos=1, final_scores=-0.12112996,

final_local_scores=-0.13872898, pem=-1.7960792,
similarity=9.717122|
gt: 9273698 Columbia_Township,_Meigs_County,_Ohio,
pem_pos=11, final_scores=-0.45426756,
final_local_scores=-0.41425663, pem=-2.333222,
similarity=4.4158926]
7: 24920512 Carpenter,_Ohio
[winner: 22900888 Carpenter,_Delaware, pem_pos=3,
final_scores=-0.44853413, final_local_scores=-0.4116433,
pem=-2.8439317, similarity=7.271199|
RECALL MISS: 24920512 Carpenter,_Ohio]
8: 9273724 Salem_Township,_Meigs_County,_Ohio
[winner: 48970 Salem,_Oregon, pem_pos=0,
final_scores=0.11188577, final_local_scores=0.091199934,
pem=-1.8624039, similarity=16.939342|
RECALL MISS: 9273724 Salem_Township,_Meigs_County,_Ohio]
9: 9273717 Rutland_Township,_Meigs_County,_Ohio
[winner: 23571693 Rutland, pem_pos=0,
final_scores=-0.1131341, final_local_scores=-0.11920999,
pem=-0.7888901, similarity=4.822356|
gt: 9273717 Rutland_Township,_Meigs_County,_Ohio,
pem_pos=16, final_scores=-0.62034637,
final_local_scores=-0.6212704, pem=-3.2841563,
similarity=3.4095168]
10: 9273731 Salisbury_Township,_Meigs_County,_Ohio
[winner: 116627 Salisbury,_Maryland, pem_pos=2,
final_scores=-0.12245843, final_local_scores=-0.15285084,
pem=-2.6058002, similarity=13.699085|
RECALL MISS: 9273731
Salisbury_Township,_Meigs_County,_Ohio]
11: 9622807 Langsville,_Ohio
[winner: 9622807 Langsville,_Ohio, pem_pos=0,
final_scores=0.026249751, final_local_scores=0.12352738,
pem=0.0, similarity=7.7760167|
gt: 9622807 Langsville,_Ohio, pem_pos=0,
final_scores=0.026249751, final_local_scores=0.12352738,
pem=0.0, similarity=7.7760167]
12: 129603 Middleport,_Ohio
[winner: 126818 Middleport,_New_York, pem_pos=1,
final_scores=0.05707161, final_local_scores=0.04895012,
pem=-1.0345814, similarity=11.176697|
gt: 129603 Middleport,_Ohio, pem_pos=3,
final_scores=0.021610573, final_local_scores=0.010969996,
pem=-1.1546794, similarity=10.696405]
13: 9273736 Scipio_Township,_Meigs_County,_Ohio
[winner: 13749698
Scipio_Township,_LaPorte_County,_Indiana, pem_pos=2,
final_scores=-0.0856296, final_local_scores=-0.0051739216,
pem=-1.0422176, similarity=9.603054|
gt: 9273736 Scipio_Township,_Meigs_County,_Ohio,
pem_pos=1, final_scores=-0.17017846,
final_local_scores=-0.13186924, pem=-0.97930366,

```

similarity=5.4800773 ]
14: 17099445 Rutland,_Ohio
    [winner: 350747 Rutland_(city),_Vermont, pem_pos=1,
final_scores=-0.09343162, final_local_scores=-0.12664214,
pem=-2.1057184, similarity=11.7616825]
    gt: 17099445 Rutland,_Ohio, pem_pos=14,
final_scores=-0.44926718, final_local_scores=-0.4776643,
pem=-3.2411432, similarity=7.4610977 ]
1: 22388 Ohio_River
    [pem_pos=0, final_scores=0.33975917,
final_local_scores=0.29373085, pem=0.0, similarity=12.855227]
2: 22199 Ohio
    [pem_pos=0, final_scores=0.5974086,
final_local_scores=0.5632261, pem=-0.19147073,
similarity=21.938751]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.47597283,
final_local_scores=0.4996071, pem=-0.073457584,
similarity=19.398458]
4: 19579 Mississippi_River
    [pem_pos=0, final_scores=0.41771525,
final_local_scores=0.4108768, pem=0.0, similarity=16.351097]
5: 609240 Allegheny_Plateau
    [pem_pos=0, final_scores=0.051241532,
final_local_scores=0.053898945, pem=0.0, similarity=5.698165]
6: 93140 Athens_County,_Ohio
    [pem_pos=0, final_scores=0.2918474,
final_local_scores=0.25267226, pem=0.0, similarity=11.629957]
7: 93051 Meigs_County,_Ohio
    [pem_pos=0, final_scores=0.19560145,
final_local_scores=0.146162, pem=-0.0932314,
similarity=8.958486]
8: 93118 Gallia_County,_Ohio
    [pem_pos=0, final_scores=0.32990927,
final_local_scores=0.2965034, pem=0.0, similarity=12.937965]
1: 22199 Ohio
    [pem_pos=0, final_scores=0.4900067,
final_local_scores=0.44590586, pem=-0.19147073,
similarity=18.437681]
2: 22388 Ohio_River
    [pem_pos=0, final_scores=0.2451608,
final_local_scores=0.19039655, pem=0.0, similarity=9.771525]
3: 52106 Albany,_New_York
    [pem_pos=0, final_scores=0.046634242,
final_local_scores=0.018284753, pem=-1.175974,
similarity=11.030496]
4: 93140 Athens_County,_Ohio
    [pem_pos=0, final_scores=0.3431658,
final_local_scores=0.30872974, pem=0.0, similarity=13.302824]
5: 93051 Meigs_County,_Ohio
    [pem_pos=0, final_scores=0.2845326,
final_local_scores=0.24309927, pem=-0.0932314,

```

```

similarity=11.851287]
6: 48970 Salem,_Oregon
    [pem_pos=0, final_scores=0.11188577,
final_local_scores=0.091199934, pem=-1.8624039,
similarity=16.939342]
7: 22388 Ohio_River
    [pem_pos=0, final_scores=0.34895974,
final_local_scores=0.30378106, pem=0.0, similarity=13.155146]
8: 126818 Middleport,_New_York
    [pem_pos=1, final_scores=0.05707161,
final_local_scores=0.04895012, pem=-1.0345814,
similarity=11.176697]
9: 93051 Meigs_County,_Ohio
    [pem_pos=0, final_scores=0.21210836,
final_local_scores=0.16398683, pem=-0.0932314,
similarity=9.490414]

```

Sample: b'Paddy_Tuimavave'

[1 Paddy Tuimavave 1] is a former professional 1 rugby league player who represented and [2 2 Western Samoa 2] . His position of preference was at Fullback . He is related to [1 1 Evarn 1] and 3 Tony , who have both also played for the and [2 2 Paki Tuimavave 2] who also represented [3 Western Samoa 3] . A [3 3 Mount Albert Lions 3] and [4 4 Northcote Tigers 4] player in the [5 5 Auckland Rugby League 5] competition , . [4 Tuimavave 4] was part of the 1988 [5 4 Auckland 5] side that defeated [6 6 Great Britain 6] 30-14 at [7 7 Carlaw Park 7] . [6 Tuimavave 6] scored a try in that match . In 1991 [7 Tuimavave 7] played for [8 5 Canterbury 8] in the Lion Red Showdown competition against [9 6 Auckland 9] , [8 8 Western Suburbs 8] and [9 9 North Sydney 9] . [10 Tuimavave 10] also played for [10 10 Workington Town 10] in [11 England 11] . [12 Tuimavave 12] represented in two tests against 7 Papua New Guinea in 1990 . He also represented [13 8 Samoa 13] in the 9 19 86 , 10 1988 and 11 1992 [14 Pacific Cup 14]

```

1: 2926222 Evarn_Tuimavave
2: 16616057 Paki_Tuimavave
3: 6629901 Mount_Albert_Lions
4: 6727202 Northcote_Tigers
5: 6669663 Auckland_Rugby_League
6: 951941 Great_Britain_national_rugby_league_team
7: 5691334 Carlaw_Park
8: 1532351 Western_Suburbs_Magpies
9: 693234 North_Sydney_Bears
10: 3646900 Workington_Town

```

```

1: 25735 Rugby_league
    [winner: 19774201 Rugby,_Virginia, pem_pos=7,
final_scores=-0.92806435, final_local_scores=-0.8088774,
pem=-2.956099, similarity=-3.9730763]
    gt: 25735 Rugby_league, pem_pos=14,
final_scores=-1.9058337, final_local_scores=-2.1386845,
pem=-3.3002021, similarity=-41.78588 ]

```

[winner: 25735 Rugby_league, pem_pos=5,
final_scores=-1.8194104, final_local_scores=-2.0442803,
pem=-3.5378149, similarity=-37.676502]
gt: 25735 Rugby_league, pem_pos=5,
final_scores=-1.8194104, final_local_scores=-2.0442803,
pem=-3.5378149, similarity=-37.676502]
2: 2262678 Samoa_national_rugby_league_team
[winner: 1048972 Samoa_national_rugby_union_team,
pem_pos=1, final_scores=0.21310754,
final_local_scores=0.192806, pem=-1.4189762,
similarity=17.560041]
gt: 2262678 Samoa_national_rugby_league_team, pem_pos=2,
final_scores=0.13020922, final_local_scores=0.07019797,
pem=-1.5065966, similarity=14.377666]
3: 11761289 Tony_Tuimavave
[winner: 54741 Tony_Award, pem_pos=0,
final_scores=-0.36230093, final_local_scores=-0.28926355,
pem=-0.96209204, similarity=0.6895176]
RECALL MISS: 11761289 Tony_Tuimavave]
4: 14059712 Auckland_rugby_league_team
[winner: 702650 New_Zealand_Warriors, pem_pos=13,
final_scores=0.21166246, final_local_scores=0.17050195,
pem=-3.4653459, similarity=28.02292]
gt: 14059712 Auckland_rugby_league_team, pem_pos=2,
final_scores=-0.1262315, final_local_scores=-0.19949055,
pem=-2.8855133, similarity=13.828387]
5: 4900741 Canterbury_rugby_league_team
[winner: 516739 Canterbury-Bankstown_Bulldogs, pem_pos=9,
final_scores=0.04343863, final_local_scores=-0.010966718,
pem=-3.529538, similarity=22.956621]
gt: 4900741 Canterbury_rugby_league_team, pem_pos=11,
final_scores=-0.20966648, final_local_scores=-0.27797562,
pem=-3.5703843, similarity=15.21067]
6: 14059712 Auckland_rugby_league_team
[winner: 702650 New_Zealand_Warriors, pem_pos=13,
final_scores=0.24070238, final_local_scores=0.20222366,
pem=-3.4653459, similarity=28.96956]
gt: 14059712 Auckland_rugby_league_team, pem_pos=2,
final_scores=-0.072001904, final_local_scores=-0.14025295,
pem=-2.8855133, similarity=15.596155]
7: 990907 Papua_New_Guinea_national_rugby_league_team
[winner: 990907
Papua_New_Guinea_national_rugby_league_team, pem_pos=1,
final_scores=0.011519894, final_local_scores=-0.035235763,
pem=-2.7441988, similarity=17.961586]
gt: 990907 Papua_New_Guinea_national_rugby_league_team,
pem_pos=1, final_scores=0.011519894,
final_local_scores=-0.035235763, pem=-2.7441988,
similarity=17.961586]
8: 2262678 Samoa_national_rugby_league_team
[winner: 1048972 Samoa_national_rugby_union_team,
pem_pos=1, final_scores=0.13437004,

```

final_local_scores=0.1055789, pem=-2.411722,
similarity=20.355717|
    gt: 2262678 Samoa_national_rugby_league_team, pem_pos=4,
final_scores=0.04394059, final_local_scores=-0.026574016,
pem=-2.7627819, similarity=18.321129 ]
9: 30602246 1986_Pacific_Cup
    [winner: 34761 1986, pem_pos=0, final_scores=-0.5494735,
final_local_scores=-0.3892229, pem=-0.71220666,
similarity=-3.6523843]
    RECALL MISS: 30602246 1986_Pacific_Cup]
10: 30920661 1988_Pacific_Cup
    [winner: 34670 1988, pem_pos=0, final_scores=-0.62332404,
final_local_scores=-0.4551703, pem=-0.7886237,
similarity=-5.2048187]
    RECALL MISS: 30920661 1988_Pacific_Cup]
11: 30923953 1992_Pacific_Cup_(rugby_league)
    [winner: 2271679 Pacific_Cup, pem_pos=0,
final_scores=0.18360128, final_local_scores=0.22223997,
pem=0.0, similarity=10.721797]
    gt: 30923953 1992_Pacific_Cup_(rugby_league), pem_pos=2,
final_scores=-0.511885, final_local_scores=-0.56169176,
pem=-3.777675, similarity=7.871295 ]
    [winner: 34669 1992, pem_pos=0, final_scores=-0.7195647,
final_local_scores=-0.59171265, pem=-0.76294684,
similarity=-9.419149]
    RECALL MISS: 30923953 1992_Pacific_Cup_(rugby_league)]
1: 2926222 Evarn_Tuimavave
    [pem_pos=0, final_scores=0.31220907,
final_local_scores=0.28885904, pem=0.0, similarity=12.709843]
2: 16616057 Paki_Tuimavave
    [pem_pos=0, final_scores=0.47009283,
final_local_scores=0.4521821, pem=0.0, similarity=17.583729]
3: 6629901 Mount_Albert_Lions
    [pem_pos=0, final_scores=0.3078949,
final_local_scores=0.28985432, pem=0.0, similarity=12.739544]
4: 6727202 Northcote_Tigers
    [pem_pos=0, final_scores=0.22235598,
final_local_scores=0.19908258, pem=0.0, similarity=10.030734]
5: 6669663 Auckland_Rugby_League
    [pem_pos=0, final_scores=0.2590397,
final_local_scores=0.22722298, pem=0.0, similarity=10.8705]
6: 951941 Great_Britain_national_rugby_league_team
    [pem_pos=2, final_scores=0.4258235,
final_local_scores=0.4138263, pem=-2.6635387,
similarity=30.923851]
7: 5691334 Carlaw_Park
    [pem_pos=0, final_scores=0.3535437,
final_local_scores=0.34974512, pem=0.0, similarity=14.526805]
8: 1532351 Western_Suburbs_Magpies
    [pem_pos=0, final_scores=0.28904623,
final_local_scores=0.26821816, pem=-0.31132302,
similarity=13.7869005]

```


9: 693234 North_Sydney_Bears
[pem_pos=1, final_scores=0.3184268,
final_local_scores=0.29659742, pem=-1.09192,
similarity=18.878801]
10: 3646900 Workington_Town
[pem_pos=0, final_scores=0.2494372,
final_local_scores=0.25618893, pem=0.0, similarity=11.734901]
1: 24168298 Paddy_Tuimavave
[pem_pos=0, final_scores=0.32832462,
final_local_scores=0.28279176, pem=0.0, similarity=12.528783]
2: 1048972 Samoa_national_rugby_union_team
[pem_pos=1, final_scores=0.21310754,
final_local_scores=0.192806, pem=-1.4189762,
similarity=17.560041]
3: 1048972 Samoa_national_rugby_union_team
[pem_pos=1, final_scores=0.24017434,
final_local_scores=0.22237235, pem=-1.4189762,
similarity=18.44236]
4: 16616057 Paki_Tuimavave
[pem_pos=2, final_scores=0.6689623,
final_local_scores=0.6703338, pem=0.0, similarity=24.093811]
5: 702650 New_Zealand_Warriors
[pem_pos=13, final_scores=0.21166246,
final_local_scores=0.17050195, pem=-3.4653459,
similarity=28.02292]
6: 2926222 Evarn_Tuimavave
[pem_pos=0, final_scores=0.46985728,
final_local_scores=0.45806575, pem=0.0, similarity=17.759308]
7: 16616057 Paki_Tuimavave
[pem_pos=2, final_scores=0.4440685,
final_local_scores=0.42467153, pem=0.0, similarity=16.762758]
8: 516739 Canterbury-Bankstown_Bulldogs
[pem_pos=9, final_scores=0.04343863,
final_local_scores=-0.010966718, pem=-3.529538,
similarity=22.956621]
9: 702650 New_Zealand_Warriors
[pem_pos=13, final_scores=0.24070238,
final_local_scores=0.20222366, pem=-3.4653459,
similarity=28.96956]
10: 16616057 Paki_Tuimavave
[pem_pos=2, final_scores=0.30453986,
final_local_scores=0.27300197, pem=0.0, similarity=12.236636]
11: 9316 England
[pem_pos=0, final_scores=0.35661548,
final_local_scores=0.45564735, pem=-0.19053856,
similarity=18.723316]
12: 16616057 Paki_Tuimavave
[pem_pos=2, final_scores=0.47743303,
final_local_scores=0.46111715, pem=0.0, similarity=17.850368]
13: 1048972 Samoa_national_rugby_union_team
[pem_pos=1, final_scores=0.13437004,
final_local_scores=0.1055789, pem=-2.411722,

```
similarity=20.355717]
14: 2271679 Pacific_Cup
    [pem_pos=0, final_scores=0.18360128,
    final_local_scores=0.22223997, pem=0.0, similarity=10.721797]
```

Sample: b'Hunterdon_Plateau'

[1 Hunterdon Plateau 1] is a 1 plateau in western [1 1 Hunterdon County 1] , [2 2 New Jersey 2] . It borders the higher 2 Musconetcong Mountain to the northwest , the [3 3 Delaware River 3] to the west , 3 Amwell Valley to the south and the lower lying areas of the 4 Newark Basin to the east . The plateau 's edge follows a line from [2 5 Raven Rock 2] to 6 Flemington . From there it follows a curved path west of the [4 4 South Branch Raritan River 4] until it meets with the Musconetcongs . It is generally unmarred as a plateau , excluding some of the small valleys of tributaries along the [3 Delaware 3] , of Lopatcong Creek and 7 Wickecheoke Creek and Cakepoulin/Capoolong Creek Valley . Other features are Thatcher 's Hill and Sand Hill which form southeastward extensions to the plateau due to a valley of Walnut Brook and the curve of the plateau there . This also occurs between Lockatong Valley and the [4 Delaware River 4] . A small ridge called Barren Ridge runs through [5 8 Alexandria 5] and 9 Union Township rises slightly above the rest of the plateau . Another feature in [5 5 Holland Township 5] called Gravel Hill also rises above the rest of the plateau . There are numerous cliffs along the [6 Delaware 6] where the plateau meets it such as Milford Bluffs where grow , a rarity in [7 Hunterdon County 7] . The plateau includes 10 Milford , [8 New Jersey 8] , 11 Frenchtown , [9 New Jersey 9] , parts of [10 12 Kingwood 10] , , [6 6 Raritan Township 6] , 13 Flemington , [11 14 Franklin Township 11] , 15 Union Township , [12 16 Bethlehem 12] , 17 Alexandria and 18 Holland

1: 93441 Hunterdon_County,_New_Jersey

2: 21648 New_Jersey

3: 52616 Delaware_River

4: 2935654 South_Branch_Raritan_River

5: 125352 Holland_Township,_New_Jersey

6: 125363 Raritan_Township,_New_Jersey

1: 538005 Plateau

[winner: 538005 Plateau, pem_pos=1,
final_scores=-1.058664, final_local_scores=-1.1137319,
pem=-0.031117933, similarity=-28.977041]

gt: 538005 Plateau, pem_pos=1, final_scores=-1.058664,
final_local_scores=-1.1137319, pem=-0.031117933,
similarity=-28.977041]

2: 26008425 Musconetcong_Mountain

[winner: 26008425 Musconetcong_Mountain, pem_pos=0,
final_scores=0.0118107945, final_local_scores=0.084756,
pem=0.0, similarity=6.619001]

gt: 26008425 Musconetcong_Mountain, pem_pos=0,
final_scores=0.0118107945, final_local_scores=0.084756,

pem=0.0, similarity=6.619001]
3: 26003477 Amwell_Valley
[winner: 26003477 Amwell_Valley, pem_pos=0,
final_scores=-0.050206527, final_local_scores=-0.006992787,
pem=0.0, similarity=3.8810344]
gt: 26003477 Amwell_Valley, pem_pos=0,
final_scores=-0.050206527, final_local_scores=-0.006992787,
pem=0.0, similarity=3.8810344]
4: 10232894 Newark_Basin
[winner: 10232894 Newark_Basin, pem_pos=0,
final_scores=0.026584223, final_local_scores=0.09530155,
pem=0.0, similarity=6.9337015]
gt: 10232894 Newark_Basin, pem_pos=0,
final_scores=0.026584223, final_local_scores=0.09530155,
pem=0.0, similarity=6.9337015]
5: 5928687 Raven_Rock,_New_Jersey
[winner: 5453871 Raven_Rock_Mountain_Complex, pem_pos=0,
final_scores=0.18184458, final_local_scores=0.2912321,
pem=0.0, similarity=12.78066]
gt: 5928687 Raven_Rock,_New_Jersey, pem_pos=1,
final_scores=-0.17258622, final_local_scores=-0.12301274,
pem=-1.2500119, similarity=7.2165265]
6: 125341 Flemington,_New_Jersey
[winner: 125341 Flemington,_New_Jersey, pem_pos=0,
final_scores=-0.053605944, final_local_scores=-0.10641013,
pem=-1.0710335, similarity=6.738669]
gt: 125341 Flemington,_New_Jersey, pem_pos=0,
final_scores=-0.053605944, final_local_scores=-0.10641013,
pem=-1.0710335, similarity=6.738669]
7: 2902372 Wickecheoke_Creek
[winner: 2902372 Wickecheoke_Creek, pem_pos=0,
final_scores=-0.20886609, final_local_scores=-0.14087602,
pem=0.0, similarity=-0.114308864]
gt: 2902372 Wickecheoke_Creek, pem_pos=0,
final_scores=-0.20886609, final_local_scores=-0.14087602,
pem=0.0, similarity=-0.114308864]
8: 125246 Alexandria_Township,_New_Jersey
[winner: 3080 Alexandria, pem_pos=0,
final_scores=0.10655223, final_local_scores=0.15100569,
pem=-0.43061808, similarity=10.937792]
RECALL MISS: 125246 Alexandria_Township,_New_Jersey]
9: 125370 Union_Township,_Hunterdon_County,_New_Jersey
[winner: 125795 Union_Township,_Union_County,_New_Jersey,
pem_pos=1, final_scores=-0.093142256,
final_local_scores=-0.13875666, pem=-2.5798569,
similarity=13.9786005]
gt: 125370 Union_Township,_Hunterdon_County,_New_Jersey,
pem_pos=2, final_scores=-0.20058438,
final_local_scores=-0.24223474, pem=-2.681091,
similarity=11.441135]
10: 125361 Milford,_New_Jersey
[winner: 21648 New_Jersey, pem_pos=0,

final_scores=0.56423444, final_local_scores=0.52960134,
pem=-0.13504621, similarity=20.628475|
RECALL MISS: 125361 Milford,_New_Jersey]
[winner: 152159 Milford,_Connecticut, pem_pos=0,
final_scores=-0.034577385, final_local_scores=-0.07445505,
pem=-1.7518675, similarity=11.394754|
gt: 125361 Milford,_New_Jersey, pem_pos=10,
final_scores=-0.4272936, final_local_scores=-0.50144076,
pem=-3.3540568, similarity=7.365602]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7547762, final_local_scores=-1.7311056,
pem=-4.6728287, similarity=-22.15836|
RECALL MISS: 125361 Milford,_New_Jersey]
11: 125344 Frenchtown,_New_Jersey
[winner: 21648 New_Jersey, pem_pos=0,
final_scores=0.49171108, final_local_scores=0.4503805,
pem=-0.13504621, similarity=18.264368|
RECALL MISS: 125344 Frenchtown,_New_Jersey]
[winner: 125344 Frenchtown,_New_Jersey, pem_pos=1,
final_scores=-0.07126729, final_local_scores=-0.118763864,
pem=-1.5011909, similarity=8.709271|
gt: 125344 Frenchtown,_New_Jersey, pem_pos=1,
final_scores=-0.07126729, final_local_scores=-0.118763864,
pem=-1.5011909, similarity=8.709271]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7069626, final_local_scores=-1.6788764,
pem=-4.6728287, similarity=-20.59974|
RECALL MISS: 125344 Frenchtown,_New_Jersey]
12: 125354 Kingwood_Township,_New_Jersey
[winner: 322969 Kingwood,_Houston, pem_pos=0,
final_scores=0.18167429, final_local_scores=0.22329596,
pem=-0.18097745, similarity=11.7374935|
gt: 125354 Kingwood_Township,_New_Jersey, pem_pos=2,
final_scores=-0.050195202, final_local_scores=-0.06357092,
pem=-1.418875, similarity=9.908692]
13: 125341 Flemington,_New_Jersey
[winner: 125341 Flemington,_New_Jersey, pem_pos=0,
final_scores=-0.02333282, final_local_scores=-0.07334134,
pem=-1.0710335, similarity=7.7255077|
gt: 125341 Flemington,_New_Jersey, pem_pos=0,
final_scores=-0.02333282, final_local_scores=-0.07334134,
pem=-1.0710335, similarity=7.7255077]
14: 125342 Franklin_Township,_Hunterdon_County,_New_Jersey
[winner: 125712
Franklin_Township,_Somerset_County,_New_Jersey, pem_pos=0,
final_scores=0.15712841, final_local_scores=0.11985114,
pem=-1.5841758, similarity=16.281305|
gt: 125342
Franklin_Township,_Hunterdon_County,_New_Jersey, pem_pos=1,
final_scores=-0.008647293, final_local_scores=-0.03541392,
pem=-2.0533097, similarity=14.199108]
15: 125370 Union_Township,_Hunterdon_County,_New_Jersey

[winner: 125795 Union_Township,_Union_County,_New_Jersey,
pem_pos=1, final_scores=-0.054550514,
final_local_scores=-0.09660104, pem=-2.5798569,
similarity=15.2366085]

gt: 125370 Union_Township,_Hunterdon_County,_New_Jersey,
pem_pos=2, final_scores=-0.1908871,
final_local_scores=-0.23164192, pem=-2.681091,
similarity=11.757246]

16: 125249 Bethlehem_Township,_New_Jersey

[winner: 4312 Bethlehem, pem_pos=0,
final_scores=0.04733987, final_local_scores=0.08410224,
pem=-0.5095743, similarity=9.370636]

gt: 125249 Bethlehem_Township,_New_Jersey, pem_pos=19,
final_scores=-0.49661285, final_local_scores=-0.5702335,
pem=-3.4899187, similarity=6.051527]

17: 125246 Alexandria_Township,_New_Jersey

[winner: 3080 Alexandria, pem_pos=0,
final_scores=-0.020153001, final_local_scores=0.010437936,
pem=-0.43061808, similarity=6.74297]

RECALL MISS: 125246 Alexandria_Township,_New_Jersey]

18: 125352 Holland_Township,_New_Jersey

[winner: 13288 Holland, pem_pos=0,
final_scores=-0.17184523, final_local_scores=-0.16492552,
pem=-0.6316271, similarity=2.6028929]

RECALL MISS: 125352 Holland_Township,_New_Jersey]

1: 93441 Hunterdon_County,_New_Jersey

[pem_pos=0, final_scores=0.46647424,
final_local_scores=0.43598258, pem=0.0, similarity=17.100304]

2: 21648 New_Jersey

[pem_pos=0, final_scores=0.50267714,
final_local_scores=0.4623593, pem=-0.13504621,
similarity=18.62184]

3: 52616 Delaware_River

[pem_pos=0, final_scores=0.2114047,
final_local_scores=0.16987666, pem=0.0, similarity=9.159171]

4: 2935654 South_Branch_Raritan_River

[pem_pos=0, final_scores=0.1863349,
final_local_scores=0.21051261, pem=0.0, similarity=10.371829]

5: 125352 Holland_Township,_New_Jersey

[pem_pos=0, final_scores=0.06342192,
final_local_scores=0.06770143, pem=-0.627944,
similarity=9.524916]

6: 125363 Raritan_Township,_New_Jersey

[pem_pos=0, final_scores=0.3001352,
final_local_scores=0.31344882, pem=0.0, similarity=13.443651]

1: 26003649 Hunterdon_Plateau

[pem_pos=0, final_scores=0.06471722,
final_local_scores=0.112372816, pem=0.0, similarity=7.443142]

2: 5453871 Raven_Rock_Mountain_Complex

[pem_pos=0, final_scores=0.18184458,
final_local_scores=0.2912321, pem=0.0, similarity=12.78066]

3: 7930 Delaware

[pem_pos=0, final_scores=0.20363675,
 final_local_scores=0.14638615, pem=-0.5414915,
 similarity=11.402883]
 4: 52616 Delaware_River
 [pem_pos=0, final_scores=0.24637534,
 final_local_scores=0.20807678, pem=0.0, similarity=10.299139]
 5: 3080 Alexandria
 [pem_pos=0, final_scores=0.10655223,
 final_local_scores=0.15100569, pem=-0.43061808,
 similarity=10.937792]
 6: 7930 Delaware
 [pem_pos=0, final_scores=0.229097,
 final_local_scores=0.17419758, pem=-0.5414915,
 similarity=12.232832]
 7: 93441 Hunterdon_County,_New_Jersey
 [pem_pos=0, final_scores=0.43188035,
 final_local_scores=0.39819402, pem=0.0, similarity=15.972616]
 8: 21648 New_Jersey
 [pem_pos=0, final_scores=0.56423444,
 final_local_scores=0.52960134, pem=-0.13504621,
 similarity=20.628475]
 9: 21648 New_Jersey
 [pem_pos=0, final_scores=0.49171108,
 final_local_scores=0.4503805, pem=-0.13504621,
 similarity=18.264368]
 10: 322969 Kingwood,_Houston
 [pem_pos=0, final_scores=0.18167429,
 final_local_scores=0.22329596, pem=-0.18097745,
 similarity=11.7374935]
 11: 125712 Franklin_Township,_Somerset_County,_New_Jersey
 [pem_pos=0, final_scores=0.15712841,
 final_local_scores=0.11985114, pem=-1.5841758,
 similarity=16.281305]
 12: 4312 Bethlehem
 [pem_pos=0, final_scores=0.04733987,
 final_local_scores=0.08410224, pem=-0.5095743,
 similarity=9.370636]

Sample: b'Mississippi_Hills_National_Heritage_Area'

Mississippi Hills National Heritage Area is a federally
 designated [1 1 National Heritage Area 1] in the northeastern
 portion of the [1 U.S. 1] state of [2 2 Mississippi 2] . The
 designation commemorates the region 's impact on [2 American 2
] culture and its role in the [3 3 American Civil War 3] and
 the [3 American 3] civil rights movement . The national
 heritage area designation provides a unified marketing and
 promotion framework for the area . The national heritage area
 covers all of 1 Marshall , 2 Benton , [4 4 Tippah 4] , 3 Alcorn ,
 [4 4 Tishomingo 4] , [5 5 Lafayette 5] , 6 Union , [6 7 Pontotoc 6]
 , 8 Lee , [5 5 Itawamba 5] , 9 Calhoun , [7 10 Chickasaw 7] and
 11 Monroe counties , and portions of , 12 Tate ,
 13 Panola , [6 6 Yalobusha 6] , [8 14 Grenada 8] and 15 Mont

gomery counties to the east of 16 Interstate 55 and north of 17 Mississippi Highway 14 . Significant sites within the national heritage area include and the birthplace of [7 7 Elvis Presley 7] . Mississippi Hills National Heritage Area was established by the 18 Omnibus Public Land Management Act of 2009

1: 1232306 National_Heritage_Area

2: 16949861 Mississippi

3: 863 American_Civil_War

4: 94742 Tippah_County,_Mississippi

5: 94821 Itawamba_County,_Mississippi

6: 94725 Yalobusha_County,_Mississippi

7: 9288 Elvis_Presley

1: 94774 Marshall_County,_Mississippi

[winner: 136027 Marshall,_Texas, pem_pos=1, final_scores=-0.050594017, final_local_scores=-0.09234214, pem=-2.915105, similarity=17.186834]

gt: 94774 Marshall_County,_Mississippi, pem_pos=6, final_scores=-0.40655142, final_local_scores=-0.48541397, pem=-3.4187348, similarity=8.195603]

2: 94852 Benton_County,_Mississippi

[winner: 115335 Benton,_Kentucky, pem_pos=3, final_scores=-0.20962718, final_local_scores=-0.23602894, pem=-2.950948, similarity=13.093853]

gt: 94852 Benton_County,_Mississippi, pem_pos=8, final_scores=-0.4258644, final_local_scores=-0.4778431, pem=-3.2302656, similarity=7.3966093]

3: 94856 Alcorn_County,_Mississippi

[winner: 94856 Alcorn_County,_Mississippi, pem_pos=0, final_scores=-0.025034964, final_local_scores=-0.052644625, pem=-0.73905826, similarity=6.537806]

gt: 94856 Alcorn_County,_Mississippi, pem_pos=0, final_scores=-0.025034964, final_local_scores=-0.052644625, pem=-0.73905826, similarity=6.537806]

4: 94741 Tishomingo_County,_Mississippi

[winner: 130282 Tishomingo,_Oklahoma, pem_pos=0, final_scores=0.22836755, final_local_scores=0.2282038, pem=-0.22434364, similarity=12.119785]

gt: 94741 Tishomingo_County,_Mississippi, pem_pos=1, final_scores=0.078088954, final_local_scores=0.054372996, pem=-0.683598, similarity=9.429824]

5: 94811 Lafayette_County,_Mississippi

[winner: 115635 Lafayette,_Louisiana, pem_pos=0, final_scores=0.082363054, final_local_scores=0.06581378, pem=-1.5915291, similarity=14.708711]

gt: 94811 Lafayette_County,_Mississippi, pem_pos=8, final_scores=-0.33946002, final_local_scores=-0.39740166, pem=-3.5104885, similarity=11.321036]

6: 94738 Union_County,_Mississippi

[winner: 481130 Union_(American_Civil_War), pem_pos=0, final_scores=0.005519882, final_local_scores=-0.016984344, pem=-1.132797, similarity=9.743193]

RECALL MISS: 94738 Union_County,_Mississippi]

7: 94760 Pontotoc_County,_Mississippi
 [winner: 92054 Pontotoc_County,_Oklahoma, pem_pos=0,
 final_scores=0.072706565, final_local_scores=0.059517562,
 pem=-0.49032253, similarity=8.532286|
 gt: 94760 Pontotoc_County,_Mississippi, pem_pos=1,
 final_scores=-0.039961413, final_local_scores=-0.06681153,
 pem=-0.6155889, similarity=5.4435916]

8: 94780 Lee_County,_Mississippi
 [winner: 73610 Lee_County,_Florida, pem_pos=2,
 final_scores=-0.3819548, final_local_scores=-0.4335768,
 pem=-3.136036, similarity=8.205167|
 gt: 94780 Lee_County,_Mississippi, pem_pos=13,
 final_scores=-0.5845889, final_local_scores=-0.6729606,
 pem=-4.0134063, similarity=5.8327546]

9: 94849 Calhoun_County,_Mississippi
 [winner: 52110 John_C._Calhoun, pem_pos=0,
 final_scores=-0.21125482, final_local_scores=-0.23598853,
 pem=-1.9563721, similarity=7.686401|
 gt: 94849 Calhoun_County,_Mississippi, pem_pos=6,
 final_scores=-0.29573298, final_local_scores=-0.34397838,
 pem=-2.7722208, similarity=8.900479]

10: 94846 Chickasaw_County,_Mississippi
 [winner: 92809 Chickasaw, pem_pos=0,
 final_scores=0.04275398, final_local_scores=0.0115427375,
 pem=-0.34369627, similarity=6.303245|
 gt: 94846 Chickasaw_County,_Mississippi, pem_pos=1,
 final_scores=-0.1377282, final_local_scores=-0.18850207,
 pem=-1.8100746, similarity=8.307902]

11: 94773 Monroe_County,_Mississippi
 [winner: 115682 Monroe,_Louisiana, pem_pos=0,
 final_scores=-0.12707137, final_local_scores=-0.18560895,
 pem=-2.5014327, similarity=12.153952|
 RECALL MISS: 94773 Monroe_County,_Mississippi]

12: 94743 Tate_County,_Mississippi
 [winner: 66101 Tate, pem_pos=0, final_scores=-0.13365269,
 final_local_scores=-0.09667879, pem=-0.59705365,
 similarity=4.4514956|
 gt: 94743 Tate_County,_Mississippi, pem_pos=1,
 final_scores=-0.31598264, final_local_scores=-0.35200948,
 pem=-2.1343162, similarity=5.1917906]

13: 94766 Panola_County,_Mississippi
 [winner: 94766 Panola_County,_Mississippi, pem_pos=0,
 final_scores=0.02746801, final_local_scores=-0.008241266,
 pem=-0.740221, similarity=7.869214|
 gt: 94766 Panola_County,_Mississippi, pem_pos=0,
 final_scores=0.02746801, final_local_scores=-0.008241266,
 pem=-0.740221, similarity=7.869214]

14: 94828 Grenada_County,_Mississippi
 [winner: 17238662 Grenada, pem_pos=0,
 final_scores=0.154159, final_local_scores=0.17704025,
 pem=-0.4001768, similarity=11.549171|

gt: 94828 Grenada County, Mississippi, pem_pos=4,
 final_scores=-0.53143823, final_local_scores=-0.58984125,
 pem=-3.1639946, similarity=3.6939676]
 15: 57959 Montgomery County, Mississippi
 [winner: 57690 Montgomery, Alabama, pem_pos=0,
 final_scores=-0.06899463, final_local_scores=-0.13381162,
 pem=-1.8303614, similarity=10.050297]
 RECALL MISS: 57959 Montgomery County, Mississippi]
 16: 90711 Interstate_55
 [winner: 90711 Interstate_55, pem_pos=0,
 final_scores=-0.29122713, final_local_scores=-0.30065393,
 pem=0.0, similarity=-4.8824]
 gt: 90711 Interstate_55, pem_pos=0,
 final_scores=-0.29122713, final_local_scores=-0.30065393,
 pem=0.0, similarity=-4.8824]
 17: 6312959 Mississippi_Highway_14
 [winner: 6312959 Mississippi_Highway_14, pem_pos=0,
 final_scores=-0.14604451, final_local_scores=-0.059431076,
 pem=0.0, similarity=2.3161712]
 gt: 6312959 Mississippi_Highway_14, pem_pos=0,
 final_scores=-0.14604451, final_local_scores=-0.059431076,
 pem=0.0, similarity=2.3161712]
 18: 21540981 Omnibus_Public_Land_Management_Act_of_2009
 [winner: 21540981
 Omnibus_Public_Land_Management_Act_of_2009, pem_pos=0,
 final_scores=-0.26075214, final_local_scores=-0.2417747,
 pem=0.0, similarity=-3.1253262]
 gt: 21540981 Omnibus_Public_Land_Management_Act_of_2009,
 pem_pos=0, final_scores=-0.26075214,
 final_local_scores=-0.2417747, pem=0.0,
 similarity=-3.1253262]
 1: 1232306 National_Heritage_Area
 [pem_pos=0, final_scores=0.10445796,
 final_local_scores=0.18056485, pem=0.0, similarity=9.478128]
 2: 16949861 Mississippi
 [pem_pos=0, final_scores=0.26257825,
 final_local_scores=0.20640427, pem=-0.31944752,
 similarity=11.986433]
 3: 863 American_Civil_War
 [pem_pos=0, final_scores=0.47962403,
 final_local_scores=0.48225027, pem=0.0, similarity=18.481024]
 4: 94742 Tippah County, Mississippi
 [pem_pos=0, final_scores=0.09323172,
 final_local_scores=0.086355865, pem=0.0, similarity=6.666744]
 5: 94821 Itawamba County, Mississippi
 [pem_pos=0, final_scores=0.13814823,
 final_local_scores=0.12903348, pem=0.0, similarity=7.940329]
 6: 94725 Yalobusha County, Mississippi
 [pem_pos=0, final_scores=0.23046313,
 final_local_scores=0.21864226, pem=0.0, similarity=10.614434]
 7: 9288 Elvis_Presley
 [pem_pos=0, final_scores=0.27161705,

```

final_local_scores=0.3324447, pem=0.0, similarity=14.010526]
1: 3434750 United_States
    [pem_pos=0, final_scores=0.5351799,
final_local_scores=0.52785414, pem=-0.14075686,
similarity=20.607391]
2: 3434750 United_States
    [pem_pos=0, final_scores=0.53904134,
final_local_scores=0.5320722, pem=-0.12994486,
similarity=20.674469]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.5236915,
final_local_scores=0.51530474, pem=-0.12994486,
similarity=20.174095]
4: 130282 Tishomingo,_Oklahoma
    [pem_pos=0, final_scores=0.22836755,
final_local_scores=0.2282038, pem=-0.22434364,
similarity=12.119785]
5: 115635 Lafayette,_Louisiana
    [pem_pos=0, final_scores=0.082363054,
final_local_scores=0.06581378, pem=-1.5915291,
similarity=14.708711]
6: 92054 Pontotoc_County,_Oklahoma
    [pem_pos=0, final_scores=0.072706565,
final_local_scores=0.059517562, pem=-0.49032253,
similarity=8.532286]
7: 92809 Chickasaw
    [pem_pos=0, final_scores=0.04275398,
final_local_scores=0.0115427375, pem=-0.34369627,
similarity=6.303245]
8: 17238662 Grenada
    [pem_pos=0, final_scores=0.154159,
final_local_scores=0.17704025, pem=-0.4001768,
similarity=11.549171]

```

Sample: b'Chippenham_United_F_C'

Chippenham United F.C. was an [1 1 English 1] 1 football club based in [2 2 Chippenham 2] , [3 3 Wiltshire 3] . The team played in the [4 4 Western Football League 4] for twelve seasons after [5 5 World War II 5] , but left in 1960 . They also participated in the [6 6 FA Cup 6] for fourteen seasons . Joining the Western League Division Two for the 2 19 48 -- 49 season , [1 United 1] won the division at their first attempt , gaining promotion to [2 Division One 2] . The following season , they finished above their local rivals [7 7 Chippenham Town 7] for what was to be the only time in the club 's history . [3 United 3] 's best league finish was in 3 1950 -- 51 , when they finished fourth , with Town in third position . The following seasons were less successful , and they only escaped relegation in 4 1955 -- 56 due to expansion of [4 Division One 4] . However , the club was relegated in 5 1957 -- 58 after finishing bottom of the table . The club dropped out of the [5 Western League 5] at the end of the 6 19

59 -- 60 season , and disbanded a couple of years later . [6 Chippenham United 6] entered the [7 FA Cup 7] throughout the late 1940s and 1950s . During their first [8 Cup 8] run in 7 1 948 -- 49 they defeated 8 Radstock Town before losing 2 -- 1 at home to [8 8 Trowbridge Town 8] in the Preliminary Round . The following season they reached the First Qualifying Round , losing 3 -- 0 at 9 Clarendon . Their best efforts in [9 The FA Cup 9] were in 10 1953 -- 54 when they reached the Fourth Qualifying Round , losing 3 -- 1 at 11 Newport -LRB- 10 W -RRB- ; and in 12 1955 -- 56 when they defeated four clubs before losing 6 -- 2 to 13 Salisbury , again in the Fourth Qualifying Round . The club 's last attempt in the competition was in 14 1961 -- 62

- 1: 9316 England
- 2: 267323 Chippenham
- 3: 51231 Wiltshire
- 4: 636625 Western_Football_League
- 5: 32927 World_War_II
- 6: 11237 FA_Cup
- 7: 15238445 Chippenham_Town_F.C.
- 8: 3774632 Trowbridge_Town_F.C.
- 1: 10568 Association_football

[winner: 10568 Association_football, pem_pos=0, final_scores=-2.0999792, final_local_scores=-2.2843177, pem=-0.33534196, similarity=-62.25524]

gt: 10568 Association_football, pem_pos=0, final_scores=-2.0999792, final_local_scores=-2.2843177, pem=-0.33534196, similarity=-62.25524]

- 2: 36826775 1948-49_Western_Football_League

[winner: 34612 1948, pem_pos=0, final_scores=-0.67443806, final_local_scores=-0.52540475, pem=-0.67699766, similarity=-7.9077945]

RECALL MISS: 36826775 1948-49_Western_Football_League]

[winner: 26823822 Psalm_49, pem_pos=10, final_scores=-1.0193903, final_local_scores=-0.89595115, pem=-3.5746164, similarity=-3.2079372]

RECALL MISS: 36826775 1948-49_Western_Football_League]

[winner: 38340996 Methylene_bridge, pem_pos=0, final_scores=-1.2749382, final_local_scores=-1.1733241, pem=0.0, similarity=-30.924618]

RECALL MISS: 36826775 1948-49_Western_Football_League]

- 3: 36856080 1950-51_Western_Football_League

[winner: 19283873 1950, pem_pos=0, final_scores=-0.7995467, final_local_scores=-0.68486685, pem=-0.90795875, similarity=-11.410459]

RECALL MISS: 36856080 1950-51_Western_Football_League]

[winner: 35267 51, pem_pos=1, final_scores=-1.2140359, final_local_scores=-1.1112509, pem=-2.677695, similarity=-14.510508]

RECALL MISS: 36856080 1950-51_Western_Football_League]

[winner: 38340996 Methylene_bridge, pem_pos=0, final_scores=-1.3072461, final_local_scores=-1.2086155,

pem=0.0, similarity=-31.977785|
 RECALL MISS: 36856080 1950-51_Western_Football_League]

4: 36968771 1955-56_Western_Football_League
 [winner: 34851 1955, pem_pos=1, final_scores=-0.47126567,
 final_local_scores=-0.22340952, pem=-0.7263205,
 similarity=1.3725708|
 RECALL MISS: 36968771 1955-56_Western_Football_League]
 [winner: 8644438 Saskatchewan_Highway_56, pem_pos=0,
 final_scores=-0.8193926, final_local_scores=-0.6964194,
 pem=-2.7613366, similarity=-1.6762544|
 RECALL MISS: 36968771 1955-56_Western_Football_League]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.3284787, final_local_scores=-1.2318089,
 pem=0.0, similarity=-32.66992|
 RECALL MISS: 36968771 1955-56_Western_Football_League]

5: 37084457 1957-58_Western_Football_League
 [winner: 34606 1957, pem_pos=0, final_scores=-0.46945435,
 final_local_scores=-0.2369207, pem=-0.6960629,
 similarity=0.80482525|
 RECALL MISS: 37084457 1957-58_Western_Football_League]
 [winner: 11310 February_27, pem_pos=14,
 final_scores=-1.0420144, final_local_scores=-0.93135625,
 pem=-3.8107264, similarity=-2.9804928|
 RECALL MISS: 37084457 1957-58_Western_Football_League]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.3286173, final_local_scores=-1.2319603,
 pem=0.0, similarity=-32.674442|
 RECALL MISS: 37084457 1957-58_Western_Football_League]

6: 37092500 1959-60_Western_Football_League
 [winner: 11652207 No._60_Squadron_RAF, pem_pos=2,
 final_scores=-0.62479436, final_local_scores=-0.5697713,
 pem=-2.771336, similarity=2.157556|
 RECALL MISS: 37092500 1959-60_Western_Football_League]
 [winner: 34662 1959, pem_pos=0, final_scores=-0.6375455,
 final_local_scores=-0.41627347, pem=-0.6961271,
 similarity=-4.5470705|
 RECALL MISS: 37092500 1959-60_Western_Football_League]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.3587332, final_local_scores=-1.2648573,
 pem=0.0, similarity=-33.656155|
 RECALL MISS: 37092500 1959-60_Western_Football_League]

7: 19519056 1948-49_FA_Cup
 [winner: 34612 1948, pem_pos=0, final_scores=-0.7959636,
 final_local_scores=-0.6581529, pem=-0.67699766,
 similarity=-11.869263|
 RECALL MISS: 19519056 1948-49_FA_Cup]
 [winner: 11369938 Saskatchewan_Highway_49, pem_pos=1,
 final_scores=-1.1029236, final_local_scores=-0.9867404,
 pem=-2.8932111, similarity=-9.622857|
 RECALL MISS: 19519056 1948-49_FA_Cup]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.3406879, final_local_scores=-1.2451456,

pem=0.0, similarity=-33.067913|
RECALL MISS: 19519056 1948-49_FA_Cup]
8: 3574596 Radstock_Town_F.C.
[winner: 3574596 Radstock_Town_F.C., pem_pos=0,
final_scores=-0.029911757, final_local_scores=-0.034832202,
pem=0.0, similarity=3.0502505|
gt: 3574596 Radstock_Town_F.C., pem_pos=0,
final_scores=-0.029911757, final_local_scores=-0.034832202,
pem=0.0, similarity=3.0502505]
9: 13297081 Clandown_F.C.
[winner: 13297081 Clandown_F.C., pem_pos=1,
final_scores=-0.0306914, final_local_scores=0.020405143,
pem=-0.35667497, similarity=6.638297|
gt: 13297081 Clandown_F.C., pem_pos=1,
final_scores=-0.0306914, final_local_scores=0.020405143,
pem=-0.35667497, similarity=6.638297]
10: 19436146 1953-54_FA_Cup
[winner: 34865 1953, pem_pos=0, final_scores=-0.48030406,
final_local_scores=-0.25834358, pem=-0.7062917,
similarity=0.22114986|
RECALL MISS: 19436146 1953-54_FA_Cup]
[winner: 8800052
Pennsylvania_House_of_Representatives,_District_54,
pem_pos=17, final_scores=-1.1443701,
final_local_scores=-1.1044037, pem=-3.87822,
similarity=-7.777536|
RECALL MISS: 19436146 1953-54_FA_Cup]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.2644839, final_local_scores=-1.1619043,
pem=0.0, similarity=-30.583828|
RECALL MISS: 19436146 1953-54_FA_Cup]
11: 3239231 Newport_(IOW)_F.C.
[winner: 5296852 Iowa_Speedway, pem_pos=0,
final_scores=-0.2967165, final_local_scores=-0.2462387,
pem=-0.82198006, similarity=1.2115142|
RECALL MISS: 3239231 Newport_(IOW)_F.C.]
[winner: 242108 Newport_RFC, pem_pos=3,
final_scores=-0.4715683, final_local_scores=-0.524917,
pem=-3.385545, similarity=6.836263|
RECALL MISS: 3239231 Newport_(IOW)_F.C.]
12: 17054550 1955-56_FA_Cup
[winner: 34851 1955, pem_pos=1, final_scores=-0.45280284,
final_local_scores=-0.20324166, pem=-0.7263205,
similarity=1.9744201|
RECALL MISS: 17054550 1955-56_FA_Cup]
[winner: 2877574 No._56_Squadron_RAF, pem_pos=1,
final_scores=-0.77573097, final_local_scores=-0.7682734,
pem=-2.967447, similarity=-2.6996593|
RECALL MISS: 17054550 1955-56_FA_Cup]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3754052, final_local_scores=-1.2830689,
pem=0.0, similarity=-34.199623|

RECALL MISS: 17054550 1955-56_FA_Cup]
13: 3041062 Salisbury_City_F.C.
[winner: 72119 Salisbury, pem_pos=0,
final_scores=-0.70117056, final_local_scores=-0.77347773,
pem=-0.80015767, similarity=-14.641023]
RECALL MISS: 3041062 Salisbury_City_F.C.]
14: 19326484 1961-62_FA_Cup
[winner: 34659 1961, pem_pos=0, final_scores=-0.57517,
final_local_scores=-0.41851926, pem=-0.7041954,
similarity=-4.5702124]
RECALL MISS: 19326484 1961-62_FA_Cup]
[winner: 18250939 London_Buses_route_62, pem_pos=7,
final_scores=-0.9654913, final_local_scores=-0.93243116,
pem=-3.434474, similarity=-5.058688]
RECALL MISS: 19326484 1961-62_FA_Cup]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.2285886, final_local_scores=-1.1226941,
pem=0.0, similarity=-29.41372]
RECALL MISS: 19326484 1961-62_FA_Cup]
1: 9316 England
[pem_pos=1, final_scores=0.5095963,
final_local_scores=0.5442119, pem=-1.3527626,
similarity=27.686615]
2: 267323 Chippenham
[pem_pos=0, final_scores=0.15316899,
final_local_scores=0.16743028, pem=-0.14149046,
similarity=9.855614]
3: 51231 Wiltshire
[pem_pos=0, final_scores=0.35544288,
final_local_scores=0.39263606, pem=-0.030204631,
similarity=15.971014]
4: 636625 Western_Football_League
[pem_pos=0, final_scores=0.19460009,
final_local_scores=0.14315167, pem=0.0, similarity=8.361645]
5: 32927 World_War_II
[pem_pos=0, final_scores=0.19670464,
final_local_scores=0.27320027, pem=0.0, similarity=12.242554]
6: 11237 FA_Cup
[pem_pos=0, final_scores=0.39311105,
final_local_scores=0.3289308, pem=-0.16323638,
similarity=14.793368]
7: 15238445 Chippenham_Town_F.C.
[pem_pos=0, final_scores=0.11637877,
final_local_scores=0.0811954, pem=0.0, similarity=6.5127454]
8: 3774632 Trowbridge_Town_F.C.
[pem_pos=0, final_scores=0.1831583,
final_local_scores=0.15046626, pem=0.0, similarity=8.5799265]
1: 19961 Manchester_United_F.C.
[pem_pos=5, final_scores=0.5091267,
final_local_scores=0.48071045, pem=-3.3400002,
similarity=36.59851]
2: 28789929 Football_League_First_Division

[pem_pos=0, final_scores=0.31040806,
 final_local_scores=0.24360448, pem=-0.3421056,
 similarity=13.21978]
 3: 19961 Manchester_United_F.C.
 [pem_pos=5, final_scores=0.44126153,
 final_local_scores=0.406578, pem=-3.3400002,
 similarity=34.38625]
 4: 28789929 Football_League_First_Division
 [pem_pos=0, final_scores=0.3574986,
 final_local_scores=0.2966538, pem=-0.3421056,
 similarity=14.802877]
 5: 636625 Western_Football_League
 [pem_pos=0, final_scores=0.0932108,
 final_local_scores=0.032399267, pem=-0.5808582,
 similarity=8.215368]
 6: 36841906 Chippenham_United_F.C.
 [pem_pos=0, final_scores=0.061421677,
 final_local_scores=0.071353495, pem=0.0, similarity=6.2190437]
 7: 11237 FA_Cup
 [pem_pos=0, final_scores=0.4141944,
 final_local_scores=0.3519612, pem=-0.16323638,
 similarity=15.480641]
 8: 11237 FA_Cup
 [pem_pos=6, final_scores=0.06882201,
 final_local_scores=-0.028279483, pem=-3.1209865,
 similarity=20.218206]
 9: 11237 FA_Cup
 [pem_pos=0, final_scores=0.27935404,
 final_local_scores=0.20466858, pem=0.0, similarity=10.197432]

Sample: b'Colin_Evans'

-LRB- rugby -RRB- [1 Colin Evans 1] -LRB- born 20 November
 1936 in [1 1 Blaenavon 1] -- died in [2 2 Leeds 2] -RRB- was
 a [2 1 Welsh 2] 2 rugby union and professional 3 rugby league
 footballer of the 1960s and '70s , playing representative
 level rugby union -LRB- RU -RRB- for [3 4 Wales 3] , and at
 club level for [3 3 Blaenavon RFC 3] , [4 4 Tredegar RFC 4] ,
 [5 5 Newport RFC 5] , and [6 6 Pontypool RFC 6] , as a , i.e.
 number 9 , and playing representative level rugby league -
 LRB- RL -RRB- for [4 5 Wales 4] , and , and at club level for
 [5 6 Leeds 5] , and [7 7 Keighley 7] , as a , i.e. number 7 ,
 and at club level was assistant coach at [8 8 Keighley 8] . [6
 Colin Evans 6] won a cap for 7 Wales -LRB- RU -RRB- while at [
 7 Pontypool 7] in 1960 in the 6-14 defeat to [8 England 8] at
 [9 9 Twickenham Stadium 9] on Saturday 16 January 1960 , won
 2 caps for 8 Wales -LRB- RL -RRB- in 1963 -- 1969 while at [9
 Leeds 9] , and [10 Keighley 10] , and played for against [11 9
 France 11] . [12 Colin Evans 12] played , and scored a 10 try
 in [10 10 Leeds 10] ' 25-10 victory over [11 11 Warrington 11]
 in the 11 1960-61 [12 12 Rugby Football League Championship 12
] final at [13 13 Odsal Stadium 13] , [14 14 Bradford 14]
 1: 634778 Blaenavon

2: 8262427 Leeds
3: 17114910 Blaenavon RFC
4: 13550511 Tredegar RFC
5: 242108 Newport RFC
6: 2768109 Pontypool RFC
7: 3550456 Keighley Cougars
8: 3550456 Keighley Cougars
9: 871456 Twickenham Stadium
10: 1095024 Leeds Rhinos
11: 852392 Warrington Wolves
12: 1004374 Rugby Football League Championship
13: 3411441 Odsal Stadium
14: 23861381 Bradford
1: 2030193 Welsh_people
[winner: 69894 Wales, pem_pos=1, final_scores=0.6099862,
final_local_scores=0.6266744, pem=-1.0776843,
similarity=28.651543|
gt: 2030193 Welsh_people, pem_pos=2,
final_scores=-0.24460457, final_local_scores=-0.2902125,
pem=-1.6894228, similarity=4.6165366]
2: 25405 Rugby_union
[winner: 19774201 Rugby_Virginia, pem_pos=7,
final_scores=-0.9055419, final_local_scores=-0.79767114,
pem=-2.956099, similarity=-3.63866|
gt: 25405 Rugby_union, pem_pos=3,
final_scores=-2.1355655, final_local_scores=-2.3753293,
pem=-2.0110612, similarity=-55.858383]
[winner: 481130 Union_(American_Civil_War), pem_pos=0,
final_scores=-1.2544675, final_local_scores=-1.2918551,
pem=-1.132797, similarity=-28.301495|
RECALL MISS: 25405 Rugby_union]
3: 25735 Rugby_league
[winner: 19774201 Rugby_Virginia, pem_pos=7,
final_scores=-0.91731936, final_local_scores=-0.8105362,
pem=-2.956099, similarity=-4.022579|
gt: 25735 Rugby_league, pem_pos=14,
final_scores=-1.9212847, final_local_scores=-2.1523259,
pem=-3.3002021, similarity=-42.19297]
[winner: 1424274 League_(unit), pem_pos=2,
final_scores=-1.7526764, final_local_scores=-1.8064735,
pem=-1.8401146, similarity=-39.812225|
gt: 25735 Rugby_league, pem_pos=5,
final_scores=-1.9656858, final_local_scores=-2.2008274,
pem=-3.5378149, similarity=-42.34818]
4: 770930 Wales_national_rugby_union_team
[winner: 69894 Wales, pem_pos=0, final_scores=0.24604918,
final_local_scores=0.22912851, pem=-0.37222588,
similarity=12.951587|
gt: 770930 Wales_national_rugby_union_team, pem_pos=2,
final_scores=-0.15815611, final_local_scores=-0.24261919,
pem=-2.723462, similarity=11.660083]
5: 990439 Wales_national_rugby_league_team

[winner: 69894 Wales, pem_pos=0, final_scores=0.19610105, final_local_scores=0.17456776, pem=-0.37222588, similarity=11.323385]

gt: 990439 Wales_national_rugby_league_team, pem_pos=3, final_scores=-0.22969355, final_local_scores=-0.31888586, pem=-3.4042935, similarity=13.0866]

6: 1095024 Leeds_Rhinos

[winner: 7609513 Leeds_United_A.F.C., pem_pos=4, final_scores=0.20118593, final_local_scores=0.17414927, pem=-3.421221, similarity=27.891806]

gt: 1095024 Leeds_Rhinos, pem_pos=1, final_scores=0.06269367, final_local_scores=-0.012077361, pem=-2.7286434, similarity=18.568089]

7: 770930 Wales_national_rugby_union_team

[winner: 69894 Wales, pem_pos=0, final_scores=-0.024280101, final_local_scores=-0.072717614, pem=-0.37222588, similarity=3.9438965]

gt: 770930 Wales_national_rugby_union_team, pem_pos=2, final_scores=-0.10844521, final_local_scores=-0.18772909, pem=-2.723462, similarity=13.298113]

8: 990439 Wales_national_rugby_league_team

[winner: 580291 Wales_national_football_team, pem_pos=1, final_scores=0.012667999, final_local_scores=-0.05364874, pem=-2.4526134, similarity=15.82642]

gt: 990439 Wales_national_rugby_league_team, pem_pos=3, final_scores=-0.16648242, final_local_scores=-0.24835229, pem=-3.4042935, similarity=15.191463]

9: 966005 France_national_rugby_league_team

[winner: 5843419 France, pem_pos=0, final_scores=0.09260337, final_local_scores=0.14445448, pem=-0.16240834, similarity=9.283725]

gt: 966005 France_national_rugby_league_team, pem_pos=5, final_scores=-0.26030165, final_local_scores=-0.3114934, pem=-4.934728, similarity=21.629948]

10: 523137 Try

[winner: 30876210 Conversion_(gridiron_football), pem_pos=0, final_scores=-1.2075222, final_local_scores=-1.2624774, pem=0.0, similarity=-33.585133]

gt: 523137 Try, pem_pos=1, final_scores=-1.2164294, final_local_scores=-1.3919547, pem=-0.24407884, similarity=-36.12165]

11: 24319561 1960-61_Northern_Rugby_Football_League_season

[winner: 3081143 1960-61_Yugoslav_First_League, pem_pos=0, final_scores=-0.9918685, final_local_scores=-0.9598947, pem=-1.3065841, similarity=-17.450052]

gt: 24319561 1960-61_Northern_Rugby_Football_League_season, pem_pos=15, final_scores=-1.529233, final_local_scores=-1.6405632, pem=-3.7577488, similarity=-24.432747]

1: 634778 Blaenavon

[pem_pos=0, final_scores=0.19877817,
final_local_scores=0.23755515, pem=0.0, similarity=11.178832]
2: 8262427 Leeds
[pem_pos=0, final_scores=0.33517075,
final_local_scores=0.32332677, pem=-0.25452003,
similarity=15.122548]
3: 17114910 Blaenavon RFC
[pem_pos=0, final_scores=0.1533194,
final_local_scores=0.17749852, pem=0.0, similarity=9.386622]
4: 13550511 Tredegar RFC
[pem_pos=0, final_scores=0.17631616,
final_local_scores=0.18958324, pem=0.0, similarity=9.747255]
5: 242108 Newport RFC
[pem_pos=0, final_scores=0.24486212,
final_local_scores=0.19971281, pem=0.0, similarity=10.049541]
6: 2768109 Pontypool RFC
[pem_pos=0, final_scores=0.20723914,
final_local_scores=0.1687327, pem=0.0, similarity=9.125033]
7: 3550456 Keighley Cougars
[pem_pos=1, final_scores=0.12223633,
final_local_scores=0.06862524, pem=-1.779445,
similarity=15.814526]
8: 3550456 Keighley Cougars
[pem_pos=1, final_scores=0.14782204,
final_local_scores=0.09657374, pem=-1.779445,
similarity=16.648565]
9: 871456 Twickenham Stadium
[pem_pos=0, final_scores=0.27295738,
final_local_scores=0.23829752, pem=0.0, similarity=11.200986]
10: 1095024 Leeds Rhinos
[pem_pos=1, final_scores=0.26225972,
final_local_scores=0.2095531, pem=-2.7286434,
similarity=25.181982]
11: 852392 Warrington Wolves
[pem_pos=1, final_scores=0.4231431,
final_local_scores=0.38397437, pem=-1.6982093,
similarity=24.7834]
12: 1004374 Rugby Football League Championship
[pem_pos=0, final_scores=0.20155518,
final_local_scores=0.18691787, pem=0.0, similarity=9.667715]
13: 3411441 Odsal Stadium
[pem_pos=0, final_scores=0.28079212,
final_local_scores=0.26007122, pem=0.0, similarity=11.850757]
14: 23861381 Bradford
[pem_pos=0, final_scores=0.048539296,
final_local_scores=0.007916763, pem=-0.6938952,
similarity=8.099475]
1: 27519506 Colin_Evans_(rugby)
[pem_pos=1, final_scores=0.09079547,
final_local_scores=0.1029568, pem=-0.4135623,
similarity=9.411165]
2: 69894 Wales

[pem_pos=1, final_scores=0.6099862,
 final_local_scores=0.6266744, pem=-1.0776843,
 similarity=28.651543]
 3: 69894 Wales
 [pem_pos=0, final_scores=0.24604918,
 final_local_scores=0.22912851, pem=-0.37222588,
 similarity=12.951587]
 4: 69894 Wales
 [pem_pos=0, final_scores=0.19610105,
 final_local_scores=0.17456776, pem=-0.37222588,
 similarity=11.323385]
 5: 7609513 Leeds_United_A.F.C.
 [pem_pos=4, final_scores=0.20118593,
 final_local_scores=0.17414927, pem=-3.421221,
 similarity=27.891806]
 6: 27519506 Colin_Evans_(rugby)
 [pem_pos=1, final_scores=0.33308607,
 final_local_scores=0.3676225, pem=-0.4135623,
 similarity=17.309317]
 7: 2768109 Pontypool_RFC
 [pem_pos=2, final_scores=0.05284445,
 final_local_scores=-0.0052654296, pem=-1.2485871,
 similarity=10.722593]
 8: 691024 England_national_rugby_union_team
 [pem_pos=3, final_scores=0.14941926,
 final_local_scores=0.093327165, pem=-3.7874734,
 similarity=27.471651]
 9: 1095024 Leeds_Rhinos
 [pem_pos=1, final_scores=0.11620264,
 final_local_scores=0.05000791, pem=-2.7286434,
 similarity=20.420835]
 10: 3550456 Keighley_Cougars
 [pem_pos=1, final_scores=0.17618947,
 final_local_scores=0.12756082, pem=-1.779445,
 similarity=17.57328]
 11: 5843419 France
 [pem_pos=0, final_scores=0.09260337,
 final_local_scores=0.14445448, pem=-0.16240834,
 similarity=9.283725]
 12: 27519506 Colin_Evans_(rugby)
 [pem_pos=1, final_scores=0.42071384,
 final_local_scores=0.4633425, pem=-0.4135623,
 similarity=20.165792]

Sample: b'Rabbit_River'

The [1 Rabbit River 1] is a tributary of the [1 1 Kalamazoo R
 iver 1] within [2 2 Allegan County 2] in the [2 1 U.S. 2] sta
 te of [3 3 Michigan 3] . The river 's watershed covers of
 land and drains most of north-central and northeast [3
 Allegan County 3] as well as small portions of 2 Barry , [4 3
 Kent 4] and [5 4 Ottawa 5] counties . The river rises in
 farmland in the southeast portion of [4 4 Leighton Township 4]

and flows generally westward , traversing the northern portion of [5 5 Wayland Township 5] , the city of [6 5 Wayland 6] , the northern portion of [6 6 Hopkins Township 6] , the extreme southwest corner of [7 7 Dorr Township 7] , 6 Salem Township , [8 8 Overisel Township 8] , [9 9 Heath Township 9] , , and [10 10 Manlius Township 10] , where it merges with the [7 Kalamazoo 7] near . The [8 Little Rabbit River 8] is an tributary of the [9 Rabbit River 9] , with a watershed that drains sections of four townships in [10 Allegan 10] and [11 Kent 11] counties : [11 11 Byron Township 11] in [12 Kent County 12] , and 7 Leighton , 8 Dorr and 9 Salem townships in [13 Allegan County 13] . The upper reaches consist of two small branches : the Red Run Drain , which rises in the extreme western edge of [14 Leighton Township 14] just north of and the Dorr & Byron Drain , which rises in the southern portion of [15 Byron Township 15] . The branches converge west of

- 1: 472558 Kalamazoo_River
- 2: 95123 Allegan_County,_Michigan
- 3: 18859 Michigan
- 4: 117010 Leighton_Township,_Michigan
- 5: 117026 Wayland_Township,_Michigan
- 6: 117007 Hopkins_Township,_Michigan
- 7: 116999 Dorr_Township,_Michigan
- 8: 117017 Overisel_Township,_Michigan
- 9: 117005 Heath_Township,_Michigan
- 10: 117011 Manlius_Township,_Michigan
- 11: 117831 Byron_Township,_Michigan

1: 18618239 U.S._state

[winner: 3434750 United_States, pem_pos=0,
final_scores=0.236077, final_local_scores=0.25670567,
pem=-0.14075686, similarity=12.5157795]
gt: 18618239 U.S._state, pem_pos=6,
final_scores=-0.5021981, final_local_scores=-0.58079267,
pem=-3.9799929, similarity=8.401521]

[winner: 229379 States_of_Brazil, pem_pos=0,
final_scores=-1.651932, final_local_scores=-1.6682115,
pem=0.0, similarity=-45.693043]

RECALL MISS: 18618239 U.S._state]

2: 95118 Barry_County,_Michigan

[winner: 1251024 Barry,_Vale_of_Glamorgan, pem_pos=0,
final_scores=-0.12631208, final_local_scores=-0.10102844,
pem=-2.1395404, similarity=12.709976]
gt: 95118 Barry_County,_Michigan, pem_pos=4,
final_scores=-0.25584096, final_local_scores=-0.34091607,
pem=-3.245096, similarity=11.563434]

3: 95079 Kent_County,_Michigan

[winner: 16766 Kent, pem_pos=0, final_scores=0.22924326,
final_local_scores=0.24770609, pem=-0.53022504,
similarity=14.365203]

gt: 95079 Kent_County,_Michigan, pem_pos=13,
final_scores=-0.3360042, final_local_scores=-0.42618334,

```

pem=-4.535942, similarity=16.03871 ]
4: 95045 Ottawa_County,_Michigan
    [winner: 22219 Ottawa, pem_pos=0,
    final_scores=0.36293823, final_local_scores=0.40776116,
    pem=-0.34965137, similarity=18.159578]
    gt: 95045 Ottawa_County,_Michigan, pem_pos=10,
    final_scores=-0.23581336, final_local_scores=-0.3152442,
    pem=-3.826999, similarity=15.494012 ]
5: 117025 Wayland,_Michigan
    [winner: 116864 Wayland,_Massachusetts, pem_pos=1,
    final_scores=0.06163977, final_local_scores=0.059897304,
    pem=-1.1302838, similarity=12.023828]
    gt: 117025 Wayland,_Michigan, pem_pos=2,
    final_scores=-0.18403478, final_local_scores=-0.23786077,
    pem=-2.0179288, similarity=7.9652834 ]
6: 117019 Salem_Township,_Allegan_County,_Michigan
    [winner: 119232
    Salem_Township,_Washtenaw_County,_Michigan, pem_pos=2,
    final_scores=-0.06150289, final_local_scores=-0.09081012,
    pem=-2.2933638, similarity=13.851427]
    gt: 117019 Salem_Township,_Allegan_County,_Michigan,
    pem_pos=17, final_scores=-0.22289035,
    final_local_scores=-0.26113158, pem=-3.015127,
    similarity=12.693755 ]
7: 117010 Leighton_Township,_Michigan
    [winner: 12767807 Leighton,_Cheshire, pem_pos=0,
    final_scores=-0.1167295, final_local_scores=-0.07061967,
    pem=-1.5244246, similarity=10.272339]
    gt: 117010 Leighton_Township,_Michigan, pem_pos=2,
    final_scores=-0.2302136, final_local_scores=-0.29483557,
    pem=-1.9183651, similarity=5.7236 ]
8: 116999 Dorr_Township,_Michigan
    [winner: 116999 Dorr_Township,_Michigan, pem_pos=1,
    final_scores=-0.17828353, final_local_scores=-0.23890066,
    pem=-1.0680196, similarity=2.7684975]
    gt: 116999 Dorr_Township,_Michigan, pem_pos=1,
    final_scores=-0.17828353, final_local_scores=-0.23890066,
    pem=-1.0680196, similarity=2.7684975 ]
9: 117019 Salem_Township,_Allegan_County,_Michigan
    [winner: 48970 Salem,_Oregon, pem_pos=0,
    final_scores=0.016350284, final_local_scores=0.011557698,
    pem=-1.8624039, similarity=14.562659]
    RECALL MISS: 117019
    Salem_Township,_Allegan_County,_Michigan]
1: 472558 Kalamazoo_River
    [pem_pos=0, final_scores=0.36704195,
    final_local_scores=0.3971296, pem=0.0, similarity=15.940852]
2: 95123 Allegan_County,_Michigan
    [pem_pos=0, final_scores=0.32416755,
    final_local_scores=0.28444868, pem=0.0, similarity=12.578229]
3: 18859 Michigan
    [pem_pos=0, final_scores=0.45306665,

```

final_local_scores=0.44292712, pem=-0.24788691,
similarity=18.655588]
4: 117010 Leighton_Township,_Michigan
[pem_pos=0, final_scores=0.19704299,
final_local_scores=0.17868313, pem=0.0, similarity=9.421974]
5: 117026 Wayland_Township,_Michigan
[pem_pos=0, final_scores=0.2689185,
final_local_scores=0.26977983, pem=0.0, similarity=12.140481]
6: 117007 Hopkins_Township,_Michigan
[pem_pos=0, final_scores=0.19037063,
final_local_scores=0.1865974, pem=0.0, similarity=9.658152]
7: 116999 Dorr_Township,_Michigan
[pem_pos=0, final_scores=0.25908428,
final_local_scores=0.24560773, pem=0.0, similarity=11.419138]
8: 117017 Overisel_Township,_Michigan
[pem_pos=0, final_scores=0.12098996,
final_local_scores=0.1544356, pem=0.0, similarity=8.6983795]
9: 117005 Heath_Township,_Michigan
[pem_pos=0, final_scores=0.17641024,
final_local_scores=0.17383909, pem=-0.18232158,
similarity=10.268911]
10: 117011 Manlius_Township,_Michigan
[pem_pos=0, final_scores=0.28105867,
final_local_scores=0.2657321, pem=-0.19131714,
similarity=13.0601015]
11: 117831 Byron_Township,_Michigan
[pem_pos=0, final_scores=0.19480409,
final_local_scores=0.17156383, pem=-0.2856639,
similarity=10.763004]
1: 1261199 Rabbit_River
[pem_pos=0, final_scores=0.08399449,
final_local_scores=0.11162977, pem=0.0, similarity=7.420968]
2: 3434750 United_States
[pem_pos=0, final_scores=0.236077,
final_local_scores=0.25670567, pem=-0.14075686,
similarity=12.5157795]
3: 95123 Allegan_County,_Michigan
[pem_pos=0, final_scores=0.29012108,
final_local_scores=0.24725804, pem=0.0, similarity=11.468386]
4: 16766 Kent
[pem_pos=0, final_scores=0.22924326,
final_local_scores=0.24770609, pem=-0.53022504,
similarity=14.365203]
5: 22219 Ottawa
[pem_pos=0, final_scores=0.36293823,
final_local_scores=0.40776116, pem=-0.34965137,
similarity=18.159578]
6: 116864 Wayland,_Massachusetts
[pem_pos=1, final_scores=0.06163977,
final_local_scores=0.059897304, pem=-1.1302838,
similarity=12.023828]
7: 54353 Kalamazoo,_Michigan

[pem_pos=0, final_scores=0.2515276,
final_local_scores=0.26622206, pem=-0.16605529,
similarity=12.937345]
8: 1261199 Rabbit_River
[pem_pos=0, final_scores=0.0313458,
final_local_scores=0.054119095, pem=0.0, similarity=5.704735]
9: 1261199 Rabbit_River
[pem_pos=0, final_scores=0.0906748,
final_local_scores=0.11892703, pem=0.0, similarity=7.638733]
10: 116994 Allegan_Michigan
[pem_pos=1, final_scores=0.2068357,
final_local_scores=0.19279745, pem=-0.45552564,
similarity=12.320393]
11: 16766 Kent
[pem_pos=0, final_scores=0.21913491,
final_local_scores=0.2366642, pem=-0.53022504,
similarity=14.035691]
12: 95079 Kent_County_Michigan
[pem_pos=1, final_scores=0.24689911,
final_local_scores=0.21739995, pem=-0.8879609,
similarity=15.406231]
13: 95123 Allegan_County_Michigan
[pem_pos=0, final_scores=0.18848689,
final_local_scores=0.13623816, pem=0.0, similarity=8.155332]
14: 117010 Leighton_Township_Michigan
[pem_pos=0, final_scores=0.19106911,
final_local_scores=0.17215756, pem=0.0, similarity=9.227238]
15: 117831 Byron_Township_Michigan
[pem_pos=0, final_scores=0.16650192,
final_local_scores=0.140648, pem=-0.2856639,
similarity=9.840415]

Sample: b'Fred_J'

__ Hume_Award The Fred J. Hume Award is an annual award presented to the player deemed to be the most `` unsung hero '' for the [1 1 Vancouver Canucks 1] of the [2 2 National Hockey League 2] -LRB- NHL -RRB- . It is voted by the fans and presented at the [1 Canucks 1] ' the last home game of the regular season . The current holder of the award is forward [3 3 Jannik Hansen 3] , who won for the 1 2012 -- 13 [2 NHL 2] season . The Fred J. Hume Award was first presented after the [3 Canucks 3] ' inaugural season in 2 1970 -- 71 and was named after former , who was also owner of the [4 Canucks 4] while they were in the and an active campaigner to bring the [5 NHL 5] to [6 Vancouver 6] . Prior to being decided by a fan vote , the award was decided on by the Vancouver Canucks Booster Club before the organization dissolved in the 2000s . The most any [7 Canucks 7] player has won the award is twice , accomplished by [4 4 Hilliard Graves 4] -LRB- 3 1977 and 4 1978 -RRB- , [5 5 Rich Sutter 5] -LRB- 5 1988 and 6 1989 -RRB- , [6 6 Steve Bozek 6] -LRB- 7 1990 and 8 1991 -RRB- , [7 7 Martin Gelinis 7] -LRB- 9 1995 and 10 1996 -RRB- , and [8 8

Jannik Hansen 8] -LRB- and

1: 73133 Vancouver Canucks

2: 21809 National_Hockey_League

3: 6141803 Jannik_Hansen

4: 10812525 Hilliard_Graves

5: 3243690 Rich_Sutter

6: 10190961 Steve_Bozek

7: 667474 Martin_Gélinas

8: 6141803 Jannik_Hansen

1: 29020787 2012-13_NHL_season

[winner: 21809 National_Hockey_League, pem_pos=0,
final_scores=0.29158503, final_local_scores=0.22110835,
pem=-0.036005534, similarity=10.88383]

RECALL MISS: 29020787 2012-13_NHL_season]

[winner: 47374 2012, pem_pos=0, final_scores=-0.80575734,
final_local_scores=-0.6149019, pem=-1.4461981,
similarity=-6.3955374]

RECALL MISS: 29020787 2012-13_NHL_season]

[winner: 15816 June_13, pem_pos=7,
final_scores=-1.0835818, final_local_scores=-0.87160707,
pem=-3.3953273, similarity=-3.4564636]

RECALL MISS: 29020787 2012-13_NHL_season]

[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.2324594, final_local_scores=-1.1141253,
pem=0.0, similarity=-29.158009]

RECALL MISS: 29020787 2012-13_NHL_season]

[winner: 690345 Season(sports), pem_pos=0,
final_scores=-2.2131264, final_local_scores=-2.4621599,
pem=-1.219945, similarity=-62.751793]

RECALL MISS: 29020787 2012-13_NHL_season]

2: 2910492 1970-71_NHL_season

[winner: 34726 1970, pem_pos=0, final_scores=-0.71228284,
final_local_scores=-0.54946357, pem=-0.80837256,
similarity=-7.9113197]

RECALL MISS: 2910492 1970-71_NHL_season]

[winner: 12094424 Malaysia_Federal_Route_71, pem_pos=5,
final_scores=-1.226532, final_local_scores=-1.104753,
pem=-3.3220983, similarity=-10.812234]

RECALL MISS: 2910492 1970-71_NHL_season]

[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.2352507, final_local_scores=-1.1171743,
pem=0.0, similarity=-29.248995]

RECALL MISS: 2910492 1970-71_NHL_season]

3: 2950798 1976-77_NHL_season

[winner: 34549 1977, pem_pos=0, final_scores=-0.6213623,
final_local_scores=-0.4786209, pem=-0.68815714,
similarity=-6.450983]

RECALL MISS: 2950798 1976-77_NHL_season]

4: 2953815 1977-78_NHL_season

[winner: 34753 1978, pem_pos=0, final_scores=-0.6199447,
final_local_scores=-0.38377517, pem=-0.7079138,
similarity=-3.513158]

RECALL MISS: 2953815 1977-78_NHL_season]
5: 1941799 1987-88_NHL_season
[winner: 34670 1988, pem_pos=0, final_scores=-0.7088405,
final_local_scores=-0.5844411, pem=-0.7886237,
similarity=-9.062517]
RECALL MISS: 1941799 1987-88_NHL_season]
6: 2237071 1988-89_NHL_season
[winner: 34847 1989, pem_pos=0, final_scores=-1.0452073,
final_local_scores=-0.9768936, pem=-0.6980697,
similarity=-21.26653]
RECALL MISS: 2237071 1988-89_NHL_season]
7: 2107768 1989-90_NHL_season
[winner: 34635 1990, pem_pos=0, final_scores=-0.8132916,
final_local_scores=-0.6902806, pem=-0.79115707,
similarity=-12.207202]
RECALL MISS: 2107768 1989-90_NHL_season]
8: 2094062 1990-91_NHL_season
[winner: 34703 1991, pem_pos=0, final_scores=-1.0616119,
final_local_scores=-0.9693448, pem=-0.6661314,
similarity=-21.214943]
RECALL MISS: 2094062 1990-91_NHL_season]
9: 1882030 1994-95_NHL_season
[winner: 34658 1995, pem_pos=0, final_scores=-1.09864,
final_local_scores=-1.0617994, pem=-0.68201965,
similarity=-23.88757]
RECALL MISS: 1882030 1994-95_NHL_season]
10: 1880859 1995-96_NHL_season
[winner: 34636 1996, pem_pos=0, final_scores=-0.76558614,
final_local_scores=-0.5654667, pem=-0.78063476,
similarity=-8.539728]
RECALL MISS: 1880859 1995-96_NHL_season]
1: 73133 Vancouver_Canucks
[pem_pos=0, final_scores=0.26706624,
final_local_scores=0.18289715, pem=0.0, similarity=9.547729]
2: 21809 National_Hockey_League
[pem_pos=0, final_scores=0.23692368,
final_local_scores=0.16139916, pem=0.0, similarity=8.906186]
3: 6141803 Jannik_Hansen
[pem_pos=0, final_scores=0.4120863,
final_local_scores=0.39785135, pem=0.0, similarity=15.962391]
4: 10812525 Hilliard_Graves
[pem_pos=0, final_scores=0.21059106,
final_local_scores=0.28304216, pem=0.0, similarity=12.536256]
5: 3243690 Rich_Sutter
[pem_pos=0, final_scores=0.29276103,
final_local_scores=0.30972305, pem=0.0, similarity=13.332466]
6: 10190961 Steve_Bozek
[pem_pos=0, final_scores=0.34887105,
final_local_scores=0.35846573, pem=0.0, similarity=14.7870455]
7: 667474 Martin_Gélinas
[pem_pos=0, final_scores=0.54627174,
final_local_scores=0.5606563, pem=0.0, similarity=20.820816]

8: 6141803 Jannik_Hansen
 [pem_pos=0, final_scores=0.5119598,
 final_local_scores=0.506948, pem=0.0, similarity=19.218052]
 1: 73133 Vancouver_Canucks
 [pem_pos=0, final_scores=0.21561988,
 final_local_scores=0.1266998, pem=-0.14234254,
 similarity=8.644769]
 2: 21809 National_Hockey_League
 [pem_pos=0, final_scores=0.29158503,
 final_local_scores=0.22110835, pem=-0.036005534,
 similarity=10.88383]
 3: 73133 Vancouver_Canucks
 [pem_pos=0, final_scores=0.20664118,
 final_local_scores=0.11689195, pem=-0.14234254,
 similarity=8.352083]
 4: 73133 Vancouver_Canucks
 [pem_pos=0, final_scores=0.24974506,
 final_local_scores=0.1639764, pem=-0.14234254,
 similarity=9.757176]
 5: 21809 National_Hockey_League
 [pem_pos=0, final_scores=0.30345762,
 final_local_scores=0.2340774, pem=-0.036005534,
 similarity=11.270853]
 6: 32706 Vancouver
 [pem_pos=0, final_scores=0.2804147,
 final_local_scores=0.35081854, pem=-0.12074051,
 similarity=15.215444]
 7: 73133 Vancouver_Canucks
 [pem_pos=0, final_scores=0.18857755,
 final_local_scores=0.09716016, pem=-0.14234254,
 similarity=7.763247]

Sample: b'Andrew_Carter'

-LRB- cricketer -RRB- Andrew `` Andy '' Carter -LRB- born 27 August 1988 , 1 Lincoln -RRB- is an [1 2 English 1] 3 cricketer who currently plays for [1 1 Nottinghamshire County Cricket Club 1] . A right handed batsman and right hand medium pace bowler he made his first class debut for [2 2 Nottinghamshire 2] in a university match against [3 3 Oxford UCCE 3] in June 2009 . In 2010 , Carter signed a new two-year contract with [2 Nottinghamshire 2] that keeps him at the club until the end of 2012 . Carter made his debut as a 20 year old for [4 4 Nottinghamshire 4] against [3 Oxford UCCE 3] . He took just two wickets in the match claiming figures of 1-41 and 1-27 | in the first and second innings respectively . He made his [5 5 List A 5] debut on 9 August 2009 in a loss against [6 6 Essex 6] in a [7 7 Pro40 7] game at [8 8 Garons Park 8] . Carter took 3 wickets for 32 runs in the first innings to help dismiss [4 Essex 4] for 224 . Following Carter 's impressive beginning to his one-day career , he played in various further one-day games in 2009 . His best figures came on 7 September against [9 9 Sussex 9] at [5 4 Hove 5] . He

took 2 wickets for 44 in eight overs as [6 Nottinghamshire 6] fell to a heavy 152 run defeat . In 2010 , Carter was restricted to second XI and league cricket appearances which lead to [7 Nottinghamshire 7] agreeing a one-month loan deal with [10 10 Essex 10] . In only Carter 's second game for [8 Essex 8] he took 3 wickets in second innings of the match against [11 11 Yorkshire 11] , also taking 1 wicket in the fourth innings in a drawn match . In the following match , a one-day game against [12 12 Northamptonshire 12] , he took 2 wickets for 33 runs in a nine-wicket victory for [9 Essex 9] . In late July , Carter took his career best bowling figures of 7-121 in the [10 County Championship 10] against [1 3 13 Kent 13] , also achieving his first five-wicket haul of 5-40 in the same match . Despite this [11 Essex 11] lost the match by 99 runs . In March 2011 , Carter was selected and played for the [14 14 England Lions 14] . He played in one four-day match against the Winward Islands where he took 2 wickets for 30 runs in the second innings of the match as [12 England Lions 12] recorded a 258 run victory . Carter featured prominently in the [13 Nottinghamshire 13] 2011 [14 T20 14] campaign as he played seven matches . His best figures came against [15 15 Derbyshire 15] where he recorded 4 wickets for 20 runs in his four overs helping [15 Nottinghamshire 15] to a 38 run victory . He followed this up in his second match by claiming figures of 2-38 against [16 16 Yorkshire 16]

1: 667341 Nottinghamshire_County_Cricket_Club

2: 667341 Nottinghamshire_County_Cricket_Club

3: 1622255 Oxford_University_Cricket_Club

4: 667341 Nottinghamshire_County_Cricket_Club

5: 683459 List_A_cricket

6: 1622252 Essex_County_Cricket_Club

7: 1177737 Pro40

8: 16948617 Garons_Park

9: 1763691 Sussex_County_Cricket_Club

10: 1622252 Essex_County_Cricket_Club

11: 869945 Yorkshire_County_Cricket_Club

12: 530308 Northamptonshire_County_Cricket_Club

13: 1177719 Kent_County_Cricket_Club

14: 2421838 England_Lions_cricket_team

15: 1829984 Derbyshire_County_Cricket_Club

16: 869945 Yorkshire_County_Cricket_Club

1: 17880 Lincoln,_England

[winner: 17880 Lincoln,_England, pem_pos=0,
final_scores=-0.028886586, final_local_scores=-0.0133420825,
pem=-1.887664, similarity=13.956968]

gt: 17880 Lincoln,_England, pem_pos=0,
final_scores=-0.028886586, final_local_scores=-0.0133420825,
pem=-1.887664, similarity=13.956968]

2: 18803164 English_people

[winner: 9316 England, pem_pos=1, final_scores=0.5987568,
final_local_scores=0.6630859, pem=-1.3527626,

```

similarity=31.234049|
    gt: 18803164 English_people, pem_pos=2,
final_scores=-0.41548184, final_local_scores=-0.4413554,
pem=-2.2084396, similarity=2.9286244 ]
3: 25675557 Cricket
    [winner: 330091 Rhabdophoridae, pem_pos=10,
final_scores=-1.4282423, final_local_scores=-1.3811831,
pem=-2.56456, similarity=-23.181068]
    gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.6366656, final_local_scores=-1.8274894,
pem=-0.07532314, similarity=-50.0366 ]
4: 2642672 County_Cricket_Ground,_Hove
    [winner: 68203 Hove, pem_pos=0, final_scores=0.15787344,
final_local_scores=0.1712161, pem=-0.041051645,
similarity=9.422388]
    gt: 2642672 County_Cricket_Ground,_Hove, pem_pos=2,
final_scores=-0.067338765, final_local_scores=-0.14267445,
pem=-2.7078161, similarity=14.557548 ]
1: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=0, final_scores=0.51690185,
final_local_scores=0.448237, pem=0.0, similarity=17.466]
2: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.60550374,
final_local_scores=0.54502106, pem=-1.1067253,
similarity=26.372772]
3: 1622255 Oxford_University_Cricket_Club
    [pem_pos=0, final_scores=0.5394632,
final_local_scores=0.50442713, pem=0.0, similarity=19.142824]
4: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6495849,
final_local_scores=0.593173, pem=-1.1067253,
similarity=27.809723]
5: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.3000431,
final_local_scores=0.24981147, pem=-0.020267159,
similarity=11.654801]
6: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.9268921,
final_local_scores=0.8990298, pem=-2.3459642,
similarity=43.676273]
7: 1177737 Pro40
    [pem_pos=0, final_scores=0.4377665,
final_local_scores=0.42629546, pem=0.0, similarity=16.81122]
8: 16948617 Garons_Park
    [pem_pos=0, final_scores=0.12768988,
final_local_scores=0.19097486, pem=0.0, similarity=9.788784]
9: 1763691 Sussex_County_Cricket_Club
    [pem_pos=1, final_scores=0.59704745,
final_local_scores=0.5561376, pem=-1.5566913,
similarity=29.151499]
10: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.76745015,

```

final_local_scores=0.72486365, pem=-2.3459642,
similarity=38.47881]
11: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6759066,
final_local_scores=0.6250827, pem=-1.7779728,
similarity=32.41232]
12: 530308 Northamptonshire_County_Cricket_Club
[pem_pos=1, final_scores=0.43430984,
final_local_scores=0.37016767, pem=-1.2069724,
similarity=21.699957]
13: 1177719 Kent_County_Cricket_Club
[pem_pos=1, final_scores=0.6765761,
final_local_scores=0.6302362, pem=-2.5096242,
similarity=36.544945]
14: 2421838 England_Lions_cricket_team
[pem_pos=0, final_scores=0.20462202,
final_local_scores=0.18833885, pem=0.0, similarity=9.71012]
15: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6576272,
final_local_scores=0.6090924, pem=-1.4828352,
similarity=30.330132]
16: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.53097266,
final_local_scores=0.46676427, pem=-1.7779728,
similarity=27.687782]
1: 9316 England
[pem_pos=1, final_scores=0.5987568,
final_local_scores=0.6630859, pem=-1.3527626,
similarity=31.234049]
2: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6302284,
final_local_scores=0.57202905, pem=-1.1067253,
similarity=27.178745]
3: 1622255 Oxford_University_Cricket_Club
[pem_pos=0, final_scores=0.5831101,
final_local_scores=0.5521047, pem=0.0, similarity=20.565619]
4: 1622252 Essex_County_Cricket_Club
[pem_pos=1, final_scores=0.99754995,
final_local_scores=0.9762128, pem=-2.3459642,
similarity=45.97957]
5: 68203 Hove
[pem_pos=0, final_scores=0.15787344,
final_local_scores=0.1712161, pem=-0.041051645,
similarity=9.422388]
6: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.74543625,
final_local_scores=0.6978761, pem=-1.1067253,
similarity=30.93427]
7: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6687768,
final_local_scores=0.6141373, pem=-1.1067253,
similarity=28.435339]

8: 1622252 Essex_County_Cricket_Club
 [pem_pos=1, final_scores=0.84038454,
 final_local_scores=0.8045335, pem=-2.3459642,
 similarity=40.856316]
 9: 1622252 Essex_County_Cricket_Club
 [pem_pos=1, final_scores=0.84984344,
 final_local_scores=0.8148658, pem=-2.3459642,
 similarity=41.164654]
 10: 683599 County_Championship
 [pem_pos=0, final_scores=0.3185702,
 final_local_scores=0.25271636, pem=0.0, similarity=11.631273]
 11: 1622252 Essex_County_Cricket_Club
 [pem_pos=1, final_scores=0.9529262,
 final_local_scores=0.9274681, pem=-2.3459642,
 similarity=44.52493]
 12: 2421838 England_Lions_cricket_team
 [pem_pos=0, final_scores=0.42507154,
 final_local_scores=0.42914647, pem=0.0, similarity=16.8963]
 13: 667341 Nottinghamshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.6146454,
 final_local_scores=0.555007, pem=-1.1067253,
 similarity=26.670773]
 14: 770690 Twenty20
 [pem_pos=0, final_scores=0.1879146,
 final_local_scores=0.16280529, pem=-0.11869875,
 similarity=9.59365]
 15: 667341 Nottinghamshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.69866747,
 final_local_scores=0.6467883, pem=-1.1067253,
 similarity=29.40971]

Sample: b'Andrew_Hele'

Andrew Keith Hele -LRB- born 3 September 1967 -RRB- is a former [1 1 English 1] 2 cricket er . Hele was a right-handed who fielded primarily as a 3 wicket-keeper . He was born in [1 1 Torquay 1] , [2 2 Devon 2] . Hele made his debut for [3 3 Devon 3] in 1997 against [4 4 Berkshire 4] in the . From 1997 to 2001 , he represented [2 Devon 2] in 28 Championship matches , the last of which came against [5 5 Cheshire 5] . In 1998 , he made his debut [6 6 MCCA Knockout Trophy 6] appearance for the county against [7 7 Dorset 7] . From 1998 to 2004 , he represented the county in 20 Trophy matches , the last of which came against [8 8 Suffolk 8] . In the same season as he made his [3 MCCA Knockout Trophy 3] debut , [4 Hele 4] also made his [9 9 List A 9] debut for [5 Devon 5] against [10 10 Yorkshire 10] in the 1st round of the 4 1998 [6 NatWest Trophy 6] . From 1998 to 2001 , he played in 6 [7 List A 7] matches , the last of which came against [11 11 Shropshire 11] in the 1st round of the 5 2001 [8 Cheltenham & Gloucester Trophy 8] . In his 6 [9 List A 9] matches , he scored 40 runs at a 6 batting average of 13.33 , with a high score of 19 . Behind the stumps he took 8 and made 2

```

1: 145272 Torquay
2: 8166 Devon
3: 3691070 Devon_County_Cricket_Club
4: 3702129 Berkshire_County_Cricket_Club
5: 3712477 Cheshire_County_Cricket_Club
6: 7543416 MCCA_Knockout_Trophy
7: 3712026 Dorset_County_Cricket_Club
8: 3691249 Suffolk_County_Cricket_Club
9: 683459 List_A_cricket
10: 869945 Yorkshire_County_Cricket_Club
11: 3702447 Shropshire_County_Cricket_Club
1: 18803164 English_people
    [winner: 9316 England, pem_pos=1,
    final_scores=0.48182815, final_local_scores=0.52723473,
    pem=-1.3527626, similarity=27.179981]
    gt: 18803164 English_people, pem_pos=2,
    final_scores=-0.43296766, final_local_scores=-0.46358132,
    pem=-2.2084396, similarity=2.2653592 ]
2: 25675557 Cricket
    [winner: 330091 Raphidophoridae, pem_pos=10,
    final_scores=-1.4543431, final_local_scores=-1.4164591,
    pem=-2.56456, similarity=-24.233776]
    gt: 25675557 Cricket, pem_pos=0, final_scores=-1.749037,
    final_local_scores=-1.9446845, pem=-0.07532314,
    similarity=-53.533936 ]
3: 463313 Wicket-keeper
    [winner: 463313 Wicket-keeper, pem_pos=0,
    final_scores=-1.4808555, final_local_scores=-1.6668293,
    pem=-0.004312502, similarity=-45.62835]
    gt: 463313 Wicket-keeper, pem_pos=0,
    final_scores=-1.4808555, final_local_scores=-1.6668293,
    pem=-0.004312502, similarity=-45.62835 ]
4: 27683969 1998_NatWest_Trophy
    [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
    final_scores=0.3107316, final_local_scores=0.2616602,
    pem=0.0, similarity=11.898175]
    gt: 27683969 1998_NatWest_Trophy, pem_pos=5,
    final_scores=-0.56579775, final_local_scores=-0.6359606,
    pem=-3.8250117, similarity=5.912389 ]
    [winner: 4669371 Juno_Awards_of_1998, pem_pos=0,
    final_scores=-0.83130527, final_local_scores=-0.72275466,
    pem=0.0, similarity=-17.478727]
    RECALL MISS: 27683969 1998_NatWest_Trophy]
5: 27734795 2001_Cheltenham_&Gloucester_Trophy
    [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
    final_scores=0.28489977, final_local_scores=0.23344284,
    pem=0.0, similarity=11.056113]
    gt: 27734795 2001_Cheltenham_&Gloucester_Trophy,
    pem_pos=1, final_scores=-0.066603586,
    final_local_scores=-0.12010625, pem=-1.537979,
    similarity=8.869272 ]
    [winner: 34551 2001, pem_pos=0, final_scores=-0.60037756,

```

final_local_scores=-0.36985517, pem=-0.5613713,
similarity=-3.8946784]
RECALL MISS: 27734795
2001_Cheltenham_&_Gloucester_Trophy]
6: 18933273 Batting_average
[winner: 15652278 Batting_(material), pem_pos=2,
final_scores=-1.006069, final_local_scores=-0.9041029,
pem=-1.1616125, similarity=-16.573492]
gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.2882406, final_local_scores=-1.4255862,
pem=-0.29142123, similarity=-36.86783]
[winner: 875871 Bowling_average, pem_pos=1,
final_scores=-1.2292018, final_local_scores=-1.3845029,
pem=-0.8657023, similarity=-32.51879]
gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.3672278, final_local_scores=-1.5118678,
pem=-0.5012068, similarity=-38.301792]
1: 145272 Torquay
[pem_pos=0, final_scores=0.23435389,
final_local_scores=0.26065522, pem=0.0, similarity=11.868185]
2: 8166 Devon
[pem_pos=0, final_scores=0.21715818,
final_local_scores=0.2481173, pem=-0.1877896,
similarity=12.515257]
3: 3691070 Devon_County_Cricket_Club
[pem_pos=1, final_scores=0.2355005,
final_local_scores=0.15952003, pem=-3.0304468,
similarity=25.33015]
4: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.52859706,
final_local_scores=0.4839062, pem=-2.301417,
similarity=31.045904]
5: 3712477 Cheshire_County_Cricket_Club
[pem_pos=2, final_scores=0.20110719,
final_local_scores=0.13482025, pem=-2.629504,
similarity=22.41267]
6: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.3820606,
final_local_scores=0.33196378, pem=0.0, similarity=13.996174]
7: 3712026 Dorset_County_Cricket_Club
[pem_pos=1, final_scores=0.2804579,
final_local_scores=0.21287972, pem=-2.4537704,
similarity=23.786453]
8: 3691249 Suffolk_County_Cricket_Club
[pem_pos=4, final_scores=0.24800895,
final_local_scores=0.18302119, pem=-2.9750729,
similarity=25.73034]
9: 683459 List_A_cricket
[pem_pos=0, final_scores=0.34053653,
final_local_scores=0.28728968, pem=-0.020267159,
similarity=12.773226]
10: 869945 Yorkshire_County_Cricket_Club

[pem_pos=1, final_scores=0.697488,
 final_local_scores=0.66452366, pem=-1.7779728,
 similarity=33.589317]
 11: 3702447 Shropshire County Cricket Club
 [pem_pos=1, final_scores=0.29576522,
 final_local_scores=0.24483877, pem=-2.279101,
 similarity=23.790295]
 1: 9316 England
 [pem_pos=1, final_scores=0.48182815,
 final_local_scores=0.52723473, pem=-1.3527626,
 similarity=27.179981]
 2: 3691070 Devon County Cricket Club
 [pem_pos=1, final_scores=0.48392683,
 final_local_scores=0.43088806, pem=-3.0304468,
 similarity=33.428314]
 3: 7543416 MCCA Knockout Trophy
 [pem_pos=0, final_scores=0.36013955,
 final_local_scores=0.30801836, pem=0.0, similarity=13.281595]
 4: 3097330 George Hele
 [pem_pos=1, final_scores=0.08018117,
 final_local_scores=0.11358467, pem=-1.7089536,
 similarity=16.772861]
 5: 3691070 Devon County Cricket Club
 [pem_pos=1, final_scores=0.52948445,
 final_local_scores=0.4806528, pem=-3.0304468,
 similarity=34.91339]
 6: 1899702 Friends Provident Trophy
 [pem_pos=0, final_scores=0.3107316,
 final_local_scores=0.2616602, pem=0.0, similarity=11.898175]
 7: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.15491565,
 final_local_scores=0.084527075, pem=-0.020267159,
 similarity=6.7223854]
 8: 1899702 Friends Provident Trophy
 [pem_pos=0, final_scores=0.28489977,
 final_local_scores=0.23344284, pem=0.0, similarity=11.056113]
 9: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.12363942,
 final_local_scores=0.050362542, pem=-0.020267159,
 similarity=5.7028475]

Sample: b'Atiq-ul-Rehman'

Atiq-ur-Rehman Chishti -LRB- born 8 November 1981 -RRB- is a
 [1 1 Pakistani 1] born [2 2 English 2] 3 cricket er . Chishti
 is a right-handed 4 batsman who bowls right-arm 5 medium-fast
 . He was born in [1 1 Lahore 1] , . Chishti made his debut
 for [2 2 Shropshire 2] in 2007 against [3 3 Berkshire 3] .
 From 2007 to 2008 , he represented the county in 17
 Championship matches , the last of which came against [4 4 Ch
 eshire 4] . It was for [3 Shropshire 3] that he made his
 debut in the [5 5 MCCA Knockout Trophy 5] against [4 Cheshire
 4] in 2008 . From 2008 to 2009 , he represented the county in

7 Trophy matches , the last of which came against [6 6 Oxfordshire 6] . In 2010 , he joined [7 7 Herefordshire 7] . He made his [5 Minor Counties Championship 5] debut for the county against [8 8 Wales Minor Counties 8] . During the 2010 season , he played 4 further Championship matches , as well as 2 [6 MCCA Knockout Trophy 6] matches . Also in 2010 Chishti was called in to the [9 9 Unicorns 9] squad as a replacement for 6 Ed Young . He represented the team in 2 [10 10 List A 10] matches against [11 11 Surrey 11] and [12 12 Worcestershire 12] in the 7 2010 [7 Clydesdale Bank 40 7] . He joined [13 13 Staffordshire 13] in 2011 . In the past , Chishti has played Second XI cricket for the [14 14 Somerset Second XI 14] , the [15 15 Leicestershire Second XI 15] and the [16 16 Worcestershire Second XI 16]

1: 125315 Lahore

2: 3702447 Shropshire_County_Cricket_Club

3: 3702129 Berkshire_County_Cricket_Club

4: 3712477 Cheshire_County_Cricket_Club

5: 7543416 MCCA_Knockout_Trophy

6: 3712398 Oxfordshire_County_Cricket_Club

7: 3702547 Herefordshire_County_Cricket_Club

8: 3712706 Wales_Minor_Counties_Cricket_Club

9: 27193794 Unicorns_(cricket_team)

10: 683459 List_A_cricket

11: 766575 Surrey_County_Cricket_Club

12: 1156428 Worcestershire_County_Cricket_Club

13: 3712165 Staffordshire_County_Cricket_Club

14: 1622178 Somerset_County_Cricket_Club

15: 1622318 Leicestershire_County_Cricket_Club

16: 1156428 Worcestershire_County_Cricket_Club

1: 28171281 Pakistani_people

[winner: 23235 Pakistan, pem_pos=0, final_scores=0.3862754, final_local_scores=0.50088435, pem=-0.49826378, similarity=21.746737]

gt: 28171281 Pakistani_people, pem_pos=19, final_scores=-0.4133281, final_local_scores=-0.35339242, pem=-3.089457, similarity=10.344726]

2: 18803164 English_people

[winner: 9316 England, pem_pos=1, final_scores=0.47833198, final_local_scores=0.5324227, pem=-1.3527626, similarity=27.3348]

gt: 18803164 English_people, pem_pos=2, final_scores=-0.447999, final_local_scores=-0.4724364, pem=-2.2084396, similarity=2.001106]

3: 25675557 Cricket

[winner: 330091 Raphidophoridae, pem_pos=10, final_scores=-1.462848, final_local_scores=-1.4192417, pem=-2.56456, similarity=-24.316814]

gt: 25675557 Cricket, pem_pos=0, final_scores=-1.6962801, final_local_scores=-1.8851328, pem=-0.07532314, similarity=-51.756794]

4: 673526 Batting_(cricket)

```
[winner: 673526 Batting_(cricket), pem_pos=0,
final_scores=-1.1939144, final_local_scores=-1.3123126,
pem=0.0, similarity=-35.072315]
gt: 673526 Batting_(cricket), pem_pos=0,
final_scores=-1.1939144, final_local_scores=-1.3123126,
pem=0.0, similarity=-35.072315 ]
5: 689836 Seam_bowling
[winner: 678695 Fast_bowling, pem_pos=0,
final_scores=-0.8502706, final_local_scores=-0.97642696,
pem=0.0, similarity=-25.048815]
RECALL MISS: 689836 Seam_bowling]
6: 32912808 Ed_Young_(cricketer)
[winner: 4103142 Ed_Young_(illustrator), pem_pos=0,
final_scores=-0.2743658, final_local_scores=-0.15448698,
pem=0.0, similarity=-0.52048683]
gt: 32912808 Ed_Young_(cricketer), pem_pos=3,
final_scores=-0.4304338, final_local_scores=-0.45902184,
pem=-4.350278, similarity=14.049072 ]
7: 27495460 2010_Clydesdale_Bank_40
[winner: 27194123 ECB_40, pem_pos=0,
final_scores=0.07439263, final_local_scores=0.063623235,
pem=0.0, similarity=5.988357]
gt: 27495460 2010_Clydesdale_Bank_40, pem_pos=3,
final_scores=-0.41303974, final_local_scores=-0.44322237,
pem=-2.8195457, similarity=6.1962023 ]
[winner: 43226 2010, pem_pos=0, final_scores=-1.0593027,
final_local_scores=-0.96286124, pem=-1.6731367,
similarity=-15.545209]
RECALL MISS: 27495460 2010_Clydesdale_Bank_40]
1: 125315 Lahore
[pem_pos=0, final_scores=0.37408763,
final_local_scores=0.4774834, pem=0.0, similarity=18.33877]
2: 3702447 Shropshire_County_Cricket_Club
[pem_pos=1, final_scores=0.20368408,
final_local_scores=0.1377759, pem=-2.279101,
similarity=20.595325]
3: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.5217179,
final_local_scores=0.47400624, pem=-2.301417,
similarity=30.750471]
4: 3712477 Cheshire_County_Cricket_Club
[pem_pos=2, final_scores=0.23213737,
final_local_scores=0.16318575, pem=-2.629504,
similarity=23.259153]
5: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.4030735,
final_local_scores=0.35394156, pem=0.0, similarity=14.652035]
6: 3712398 Oxfordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.44051164,
final_local_scores=0.39051205, pem=-1.7495553,
similarity=25.257727]
7: 3702547 Herefordshire_County_Cricket_Club
```

[pem_pos=1, final_scores=0.35015398,
final_local_scores=0.29750705, pem=-1.3582021,
similarity=20.35403]
8: 3712706 Wales_Minor_Counties_Cricket_Club
[pem_pos=0, final_scores=0.32370472,
final_local_scores=0.27162233, pem=0.0, similarity=12.195465]
9: 27193794 Unicorns_(cricket_team)
[pem_pos=1, final_scores=0.24065758,
final_local_scores=0.23554155, pem=-0.8059427,
similarity=15.501585]
10: 683459 List_A_cricket
[pem_pos=0, final_scores=0.14234258,
final_local_scores=0.07327835, pem=-0.020267159,
similarity=6.386701]
11: 766575 Surrey_County_Cricket_Club
[pem_pos=1, final_scores=0.74257463,
final_local_scores=0.7065787, pem=-1.7952642,
similarity=34.938354]
12: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=0, final_scores=0.8235298,
final_local_scores=0.79188484, pem=0.0, similarity=27.721136]
13: 3712165 Staffordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.23306285,
final_local_scores=0.16360316, pem=-1.68797,
similarity=18.151402]
14: 1622178 Somerset_County_Cricket_Club
[pem_pos=0, final_scores=0.47869092,
final_local_scores=0.4167394, pem=0.0, similarity=16.526049]
15: 1622318 Leicestershire_County_Cricket_Club
[pem_pos=0, final_scores=0.5368729,
final_local_scores=0.48854607, pem=0.0, similarity=18.668901]
16: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=0, final_scores=0.35000193,
final_local_scores=0.27462757, pem=0.0, similarity=12.285148]
1: 23235 Pakistan
[pem_pos=0, final_scores=0.3862754,
final_local_scores=0.50088435, pem=-0.49826378,
similarity=21.746737]
2: 9316 England
[pem_pos=1, final_scores=0.47833198,
final_local_scores=0.5324227, pem=-1.3527626,
similarity=27.3348]
3: 3702447 Shropshire_County_Cricket_Club
[pem_pos=1, final_scores=0.15425621,
final_local_scores=0.08378345, pem=-2.279101,
similarity=18.984083]
4: 3712477 Cheshire_County_Cricket_Club
[pem_pos=2, final_scores=0.18005465,
final_local_scores=0.10629329, pem=-2.629504,
similarity=21.561369]
5: 7562952 Minor_Counties_Cricket_Championship
[pem_pos=0, final_scores=0.32226008,

```
final_local_scores=0.2603344, pem=0.0, similarity=11.858611]
6: 7543416 MCCA_Knockout_Trophy
  [pem_pos=0, final_scores=0.33042717,
final_local_scores=0.27458644, pem=0.0, similarity=12.28392]
7: 27194123 ECB_40
  [pem_pos=0, final_scores=0.07439263,
final_local_scores=0.063623235, pem=0.0, similarity=5.988357]
```

Sample: b'Czech_Republic_men'

's _ national_ice_hockey_team The [1 Czech 1] men 's national 1 ice hockey team is the national 2 ice hockey team of the [1 1 Czech Republic 1] . It is one of the most successful national ice hockey teams in the world , currently ranked third by the , behind [2 3 Russia 2] and [3 4 Finland 3] . It is controlled by the 5 Czech Ice Hockey Association . The [4 Czech Republic 4] has 72,075 players officially enrolled in organized hockey -LRB- 0.7 % of its population -RRB- . [5 Czech Republic 5] is a member of the so-called '' 6 Big Seven '' , the unofficial group of seven the strongest men 's ice hockey nations , along with [6 7 Canada 6] , [7 8 Finland 7] , [8 9 Russia 8] , [9 10 Slovakia 9] , [10 11 Sweden 10] and the [11 12 United States 11] . The [12 Czechs 12] won the gold medal at the 13 1998 Winter Olympics and won three straight gold medals at the from 14 1999 to 15 2001 . In the next three years , the team did not get a medal at the world championships -- not even home at the 16 2004 Men 's [13 World Ice Hockey Championships 13] held in [2 2 Prague 2] and [3 3 Ostrava 3] , thus keeping the `` world championship home ice curse '' alive . But the following year , the [14 Czechs 14] won gold at 17 the 2005 tournament , the only world championship where , due to the 18 2004 -- 05 [15 NHL 15] lockout , were available to participate . At the 19 2006 Winter Olympics , the [16 Czechs 16] won a bronze medal , defeating 3 -- 0 -LRB- -RRB- in the bronze medal game . At the 20 2006 Men 's [17 World Ice Hockey Championships 17] , the [18 Czechs 18] won silver , falling to [19 21 Sweden 19] in the final . [20 Czech Republic 20] won the 22 2010 World Championships in [21 Germany 21]

1: 5321 Czech_Republic

2: 23844 Prague

3: 464078 Ostrava

1: 14790 Ice_hockey

[winner: 14946 Ice, pem_pos=0, final_scores=-0.68291414, final_local_scores=-0.6584908, pem=-0.116236255, similarity=-14.928854]

gt: 14790 Ice_hockey, pem_pos=3, final_scores=-1.2164354, final_local_scores=-1.3466933, pem=-2.165202, similarity=-24.323599]

[winner: 14786334 Denver_Pioneers_men's_ice_hockey, pem_pos=1, final_scores=-1.3929499, final_local_scores=-1.4260504, pem=-1.0811191, similarity=-32.587185]

gt: 14790 Ice_hockey, pem_pos=12,
final_scores=-2.3256938, final_local_scores=-2.5583894,
pem=-5.1357985, similarity=-44.328457]
2: 14790 Ice_hockey
[winner: 14946 Ice, pem_pos=0, final_scores=-0.6658546,
final_local_scores=-0.6398558, pem=-0.116236255,
similarity=-14.37275]
gt: 14790 Ice_hockey, pem_pos=3,
final_scores=-1.1239384, final_local_scores=-1.2456543,
pem=-2.165202, similarity=-21.308393]
[winner: 14786334 Denver_Pioneers_men's_ice_hockey,
pem_pos=1, final_scores=-1.2070358,
final_local_scores=-1.2229676, pem=-1.0811191,
similarity=-26.526789]
gt: 14790 Ice_hockey, pem_pos=12,
final_scores=-2.0715265, final_local_scores=-2.2807503,
pem=-5.1357985, similarity=-36.043156]
3: 950313 Russia_men's_national_ice_hockey_team
[winner: 25391 Russia, pem_pos=0,
final_scores=0.68421775, final_local_scores=0.69332474,
pem=-0.20299491, similarity=25.883823]
gt: 950313 Russia_men's_national_ice_hockey_team,
pem_pos=3, final_scores=-0.51651436,
final_local_scores=-0.5508335, pem=-4.427518,
similarity=11.729272]
4: 950383 Finland_men's_national_ice_hockey_team
[winner: 10577 Finland, pem_pos=0,
final_scores=0.5576343, final_local_scores=0.51740175,
pem=-0.2445894, similarity=20.860128]
gt: 950383 Finland_men's_national_ice_hockey_team,
pem_pos=2, final_scores=-0.20146275,
final_local_scores=-0.22428131, pem=-3.8760931,
similarity=18.47551]
5: 16372406 Czech_Ice_Hockey_Association
[winner: 16372406 Czech_Ice_Hockey_Association,
pem_pos=0, final_scores=0.0029738396,
final_local_scores=0.08536103, pem=0.0, similarity=6.6370564]
gt: 16372406 Czech_Ice_Hockey_Association, pem_pos=0,
final_scores=0.0029738396, final_local_scores=0.08536103,
pem=0.0, similarity=6.6370564]
6: 16760449 International_Ice_Hockey_Federation
[winner: 415250 Big_Eight_Conference, pem_pos=0,
final_scores=-0.23250206, final_local_scores=-0.15463576,
pem=-0.35771066, similarity=1.4203597]
gt: 16760449 International_Ice_Hockey_Federation,
pem_pos=5, final_scores=-0.57373244,
final_local_scores=-0.623999, pem=-2.8478122,
similarity=0.95518345]
7: 744439 Canada_men's_national_ice_hockey_team
[winner: 5042916 Canada, pem_pos=0,
final_scores=0.6407105, final_local_scores=0.65521705,
pem=-0.11662793, similarity=24.276937]

gt: 744439 Canada_men's_national_ice_hockey_team,
pem_pos=2, final_scores=-0.6963391,
final_local_scores=-0.74901104, pem=-4.557522,
similarity=6.522241]
8: 950383 Finland_men's_national_ice_hockey_team
[winner: 10577 Finland, pem_pos=0,
final_scores=0.62773997, final_local_scores=0.59398156,
pem=-0.2445894, similarity=23.145424]
gt: 950383 Finland_men's_national_ice_hockey_team,
pem_pos=2, final_scores=-0.38563144,
final_local_scores=-0.42545763, pem=-3.8760931,
similarity=12.472007]
9: 950313 Russia_men's_national_ice_hockey_team
[winner: 25391 Russia, pem_pos=0, final_scores=0.724261,
final_local_scores=0.7370659, pem=-0.20299491,
similarity=27.189148]
gt: 950313 Russia_men's_national_ice_hockey_team,
pem_pos=3, final_scores=-0.63957345,
final_local_scores=-0.6852569, pem=-4.427518,
similarity=7.717809]
10: 944992 Slovakia_men's_national_ice_hockey_team
[winner: 26830 Slovakia, pem_pos=0,
final_scores=0.4399082, final_local_scores=0.38146454,
pem=-0.36602598, similarity=17.463882]
gt: 944992 Slovakia_men's_national_ice_hockey_team,
pem_pos=2, final_scores=-0.3746159,
final_local_scores=-0.40346566, pem=-3.4711561,
similarity=10.92618]
11: 747522 Sweden_men's_national_ice_hockey_team
[winner: 5058739 Sweden, pem_pos=0,
final_scores=0.5186723, final_local_scores=0.47744375,
pem=-0.2135508, similarity=19.498909]
gt: 747522 Sweden_men's_national_ice_hockey_team,
pem_pos=2, final_scores=-0.31384063,
final_local_scores=-0.35196972, pem=-4.1310015,
similarity=16.051262]
12: 945923 United_States_men's_national_ice_hockey_team
[winner: 3434750 United_States, pem_pos=0,
final_scores=0.45897263, final_local_scores=0.46149015,
pem=-0.073457584, similarity=18.260973]
gt: 945923 United_States_men's_national_ice_hockey_team,
pem_pos=2, final_scores=-0.5214056,
final_local_scores=-0.54713845, pem=-5.0140476,
similarity=15.029181]
13: 113189 1998_Winter_Olympics
[winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.058746114, final_local_scores=-0.071858406,
pem=-2.7333555, similarity=16.809727]
RECALL MISS: 113189 1998_Winter_Olympics]
[winner: 34061 Winter, pem_pos=0,
final_scores=-0.29655984, final_local_scores=-0.26972294,
pem=-0.8668251, similarity=0.75457275]

RECALL MISS: 113189 1998_Winter_Olympics]
 [winner: 4669371 Juno_Awards_of_1998, pem_pos=0,
 final_scores=-0.85428774, final_local_scores=-0.76626635,
 pem=0.0, similarity=-18.777203]
 RECALL MISS: 113189 1998_Winter_Olympics]
 14: 11141558 1999_Men's_World_Ice_Hockey_Championships
 [winner: 34553 1999, pem_pos=0, final_scores=-1.2179906,
 final_local_scores=-1.2638421, pem=-0.6240096,
 similarity=-30.232393]
 RECALL MISS: 11141558
 1999_Men's_World_Ice_Hockey_Championships]
 15: 8233028 2001_Men's_World_Ice_Hockey_Championships
 [winner: 34551 2001, pem_pos=0, final_scores=-0.6720528,
 final_local_scores=-0.47879773, pem=-0.5613713,
 similarity=-7.1457415]
 RECALL MISS: 8233028
 2001_Men's_World_Ice_Hockey_Championships]
 16: 1880717 2004_Men's_World_Ice_Hockey_Championships
 [winner: 224637 Ice_Hockey_World_Championships,
 pem_pos=0, final_scores=0.19982938,
 final_local_scores=0.1946978, pem=0.0, similarity=9.899883]
 gt: 1880717 2004_Men's_World_Ice_Hockey_Championships,
 pem_pos=4, final_scores=-0.7521838,
 final_local_scores=-0.79024947, pem=-4.8121843,
 similarity=6.676499]
 [winner: 17978594 2004_Churchill_Cup, pem_pos=0,
 final_scores=-0.42322645, final_local_scores=-0.35780546,
 pem=0.0, similarity=-6.587916]
 RECALL MISS: 1880717
 2004_Men's_World_Ice_Hockey_Championships]
 [winner: 1415027 David_Berman_(musician), pem_pos=5,
 final_scores=-2.2738848, final_local_scores=-2.3576424,
 pem=-3.0012724, similarity=-49.945652]
 RECALL MISS: 1880717
 2004_Men's_World_Ice_Hockey_Championships]
 17: 1716490 2005_Men's_World_Ice_Hockey_Championships
 [winner: 35984 2005, pem_pos=0, final_scores=-0.7330356,
 final_local_scores=-0.60012877, pem=-0.36604404,
 similarity=-11.828723]
 RECALL MISS: 1716490
 2005_Men's_World_Ice_Hockey_Championships]
 [winner: 8122672 Softball_at_the_2008_Summer_Olympics,
 pem_pos=0, final_scores=-1.2588396,
 final_local_scores=-1.2787504, pem=0.0, similarity=-34.07075]
 RECALL MISS: 1716490
 2005_Men's_World_Ice_Hockey_Championships]
 [winner: 1427502 Tournament_(graph_theory), pem_pos=3,
 final_scores=-1.415575, final_local_scores=-1.3749726,
 pem=-2.5068715, similarity=-23.309454]
 RECALL MISS: 1716490
 2005_Men's_World_Ice_Hockey_Championships]
 18: 931004 2004-05_NHL_lockout

[winner: 21809 National_Hockey_League, pem_pos=0,
 final_scores=0.18504353, final_local_scores=0.224175,
 pem=-0.036005534, similarity=10.975346|
 gt: 931004 2004-05_NHL_lockout, pem_pos=4,
 final_scores=-0.26832104, final_local_scores=-0.24908826,
 pem=-2.4602008, similarity=10.035377]
 [winner: 35524 2004, pem_pos=0, final_scores=-0.5774991,
 final_local_scores=-0.438003, pem=-0.4915334,
 similarity=-6.3081355|
 RECALL MISS: 931004 2004-05_NHL_lockout]
 [winner: 3423380 Fabric_05, pem_pos=4,
 final_scores=-0.9155519, final_local_scores=-0.73647666,
 pem=-2.7660456, similarity=-2.8460348|
 RECALL MISS: 931004 2004-05_NHL_lockout]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.2196453, final_local_scores=-1.1282277,
 pem=0.0, similarity=-29.57885|
 RECALL MISS: 931004 2004-05_NHL_lockout]
 [winner: 511550 Lockout_(industry), pem_pos=0,
 final_scores=-1.6666069, final_local_scores=-1.7313111,
 pem=-0.21046558, similarity=-46.431522|
 gt: 931004 2004-05_NHL_lockout, pem_pos=1,
 final_scores=-1.8323256, final_local_scores=-1.9651859,
 pem=-0.91970676, similarity=-49.553844]
 19: 77743 2006_Winter_Olympics
 [winner: 34069 Winter_Olympic_Games, pem_pos=0,
 final_scores=-0.056874588, final_local_scores=-0.086829126,
 pem=-0.28208634, similarity=3.0325878|
 gt: 77743 2006_Winter_Olympics, pem_pos=3,
 final_scores=-0.5799827, final_local_scores=-0.64908206,
 pem=-3.1767673, similarity=1.9955637]
 [winner: 36164 2006, pem_pos=0, final_scores=-0.5746863,
 final_local_scores=-0.42949116, pem=-0.3060146,
 similarity=-7.0630045|
 gt: 77743 2006_Winter_Olympics, pem_pos=16,
 final_scores=-1.9653988, final_local_scores=-2.1624386,
 pem=-5.475615, similarity=-30.664518]
 20: 1902647 2006_Men's_World_Ice_Hockey_Championships
 [winner: 224637 Ice_Hockey_World_Championships,
 pem_pos=0, final_scores=0.16718577,
 final_local_scores=0.15903959, pem=0.0, similarity=8.835772|
 RECALL MISS: 1902647
 2006_Men's_World_Ice_Hockey_Championships]
 [winner: 63455 Man_(Middle-earth), pem_pos=0,
 final_scores=-0.56092715, final_local_scores=-0.48540375,
 pem=-1.7409185, similarity=-0.9283172|
 RECALL MISS: 1902647
 2006_Men's_World_Ice_Hockey_Championships]
 [winner: 36164 2006, pem_pos=0, final_scores=-0.62983936,
 final_local_scores=-0.4897375, pem=-0.3060146,
 similarity=-8.860876|
 RECALL MISS: 1902647

2006_Men's_World_Ice_Hockey_Championships]
[winner: 1415027 David_Berman_(musician), pem_pos=5,
final_scores=-2.4537418, final_local_scores=-2.5541089,
pem=-3.0012724, similarity=-55.8086]
RECALL MISS: 1902647

2006_Men's_World_Ice_Hockey_Championships]
21: 747522 Sweden_men's_national_ice_hockey_team
[winner: 5058739 Sweden, pem_pos=0,
final_scores=0.37432522, final_local_scores=0.31976655,
pem=-0.2135508, similarity=14.793508]
gt: 747522 Sweden_men's_national_ice_hockey_team,
pem_pos=2, final_scores=-0.11748578,
final_local_scores=-0.13748187, pem=-4.1310015,
similarity=22.452007]

22: 17830059 2010_Men's_World_Ice_Hockey_Championships
[winner: 11867 Germany, pem_pos=0,
final_scores=0.56781036, final_local_scores=0.5434502,
pem=-0.15529114, similarity=21.151848]
RECALL MISS: 17830059

2010_Men's_World_Ice_Hockey_Championships]
[winner: 218926 IAAF_World_Championships_in_Athletics,
pem_pos=0, final_scores=-0.19175348,
final_local_scores=-0.1873172, pem=-2.1703603,
similarity=10.302551]
RECALL MISS: 17830059

2010_Men's_World_Ice_Hockey_Championships]
[winner: 43226 2010, pem_pos=0, final_scores=-1.0030166,
final_local_scores=-0.924166, pem=-1.6731367,
similarity=-14.390468]
RECALL MISS: 17830059

2010_Men's_World_Ice_Hockey_Championships]
1: 5321 Czech_Republic
[pem_pos=0, final_scores=0.57682806,
final_local_scores=0.5193388, pem=-0.27794555,
similarity=21.099329]
2: 23844 Prague
[pem_pos=0, final_scores=0.56362283,
final_local_scores=0.5567906, pem=-0.05392049,
similarity=20.998682]
3: 464078 Ostrava
[pem_pos=0, final_scores=0.49873716,
final_local_scores=0.48562515, pem=-0.1634837,
similarity=19.470785]
1: 5321 Czech_Republic
[pem_pos=1, final_scores=0.6665985,
final_local_scores=0.6173993, pem=-1.208016,
similarity=29.083517]
2: 25391 Russia
[pem_pos=0, final_scores=0.68421775,
final_local_scores=0.69332474, pem=-0.20299491,
similarity=25.883823]
3: 10577 Finland

[pem_pos=0, final_scores=0.5576343,
final_local_scores=0.51740175, pem=-0.2445894,
similarity=20.860128]
4: 5321 Czech_Republic
[pem_pos=0, final_scores=0.4901709,
final_local_scores=0.42467898, pem=-0.27794555,
similarity=18.274492]
5: 5321 Czech_Republic
[pem_pos=0, final_scores=0.5478564,
final_local_scores=0.48769164, pem=-0.27794555,
similarity=20.154915]
6: 5042916 Canada
[pem_pos=0, final_scores=0.6407105,
final_local_scores=0.65521705, pem=-0.11662793,
similarity=24.276937]
7: 10577 Finland
[pem_pos=0, final_scores=0.62773997,
final_local_scores=0.59398156, pem=-0.2445894,
similarity=23.145424]
8: 25391 Russia
[pem_pos=0, final_scores=0.724261,
final_local_scores=0.7370659, pem=-0.20299491,
similarity=27.189148]
9: 26830 Slovakia
[pem_pos=0, final_scores=0.4399082,
final_local_scores=0.38146454, pem=-0.36602598,
similarity=17.463882]
10: 5058739 Sweden
[pem_pos=0, final_scores=0.5186723,
final_local_scores=0.47744375, pem=-0.2135508,
similarity=19.498909]
11: 3434750 United_States
[pem_pos=0, final_scores=0.45897263,
final_local_scores=0.46149015, pem=-0.073457584,
similarity=18.260973]
12: 5321 Czech_Republic
[pem_pos=2, final_scores=0.20117591,
final_local_scores=0.10873523, pem=-2.2043104,
similarity=19.321972]
13: 224637 Ice_Hockey_World_Championships
[pem_pos=0, final_scores=0.19982938,
final_local_scores=0.1946978, pem=0.0, similarity=9.899883]
14: 151873 Czechs
[pem_pos=0, final_scores=0.16925918,
final_local_scores=0.11496264, pem=-0.024546154,
similarity=7.6539135]
15: 21809 National_Hockey_League
[pem_pos=0, final_scores=0.18504353,
final_local_scores=0.224175, pem=-0.036005534,
similarity=10.975346]
16: 151873 Czechs
[pem_pos=0, final_scores=0.16177069,

final_local_scores=0.106782615, pem=-0.024546154,
 similarity=7.4098053]
 17: 224637 Ice_Hockey_World_Championships
 [pem_pos=0, final_scores=0.16718577,
 final_local_scores=0.15903959, pem=0.0, similarity=8.835772]
 18: 151873 Czechs
 [pem_pos=0, final_scores=0.1803631,
 final_local_scores=0.127092, pem=-0.024546154,
 similarity=8.015878]
 19: 5058739 Sweden
 [pem_pos=0, final_scores=0.37432522,
 final_local_scores=0.31976655, pem=-0.2135508,
 similarity=14.793508]
 20: 5321 Czech_Republic
 [pem_pos=0, final_scores=0.31611985,
 final_local_scores=0.23390287, pem=-0.27794555,
 similarity=12.581352]
 21: 11867 Germany
 [pem_pos=0, final_scores=0.56781036,
 final_local_scores=0.5434502, pem=-0.15529114,
 similarity=21.151848]

Sample: b'House_of_Palatinate-Birkenfeld'

Palatinate-Birkenfeld -LRB- [1 1 German 1] : Pfalz-
 Birkenfeld -RRB- , later Palatinate-Zweibrücken-Birkenfeld ,
 was the name of a collateral line of the 2 Palatine 3 Wittels
 bachs . The [2 Counts Palatine 2] from this line initially
 ruled over only a relatively unimportant territory , namely
 the [3 Palatine 3] share of the ; however , their importance
 steadily grew . All living members of the [4 House of
 Wittelsbach 4] descend from [5 Palatinate-Birkenfeld 5] ,
 which thus became the parent branch of the 4 Kings of [6 Bava
 ria 6] . The patriarch of the line was . He was a son of
 Duke , who when he died in 1569 , bequeathed the [7 Palatine 7
] share of the Rear 5 County of [8 Sponheim 8] to his son
 Karl . The latter chose [1 1 Birkenfeld 1] as his residence
 and in 1595 , through the Treaty of [9 Kastellaun 9] with his
 joint ruler the [2 2 Margrave 2] of , [3 3 Edward Fortunatus 3
] , secured exclusive rule in the districts of [10 Birkenfeld
 10] and [11 Allenbach 11] -LRB- until 1671 -RRB- . Under
 him , Birkenfeld Castle underwent major expansion . Following
 Karl 's death in 1600 , he was succeeded by , who ruled until
 1669 . He was succeeded by , who however died in 1671 . The
 inheritance passed to [12 6 Christian II 12] , a son of [13
 Georg Wilhelm 13] 's younger brother , who had previously
 received the Lordship of from the holdings of [4 4 Zweibrücken
 4] and thus founded the [14 Palatinate-Birkenfeld-Bischweiler
 14] line . His son [15 7 Christian III 15] ended official
 residency at [16 Birkenfeld 16] in 1717 and resided primarily
 in [17 Bischweiler 17] , alternating with [18 Zweibrücken 18]
 after 1734 , in which year he succeeded to the . He was
 succeeded by his son [19 8 Christian IV 19] , who however was

9 married below his station , to the dancer Marianne Camasse , so that his sons were not entitled to inherit . After [20 Christian 20] 's death his nephew became ruler . Under his rule , the Rear County of [21 Sponheim 21] was finally partitioned . It was agreed that of the two joint rulers , the Margrave of [22 Baden 22] was to make the division and then the [23 Count Palatine 23] by [24 Rhine 24] was to choose his share . Possibly the hope on the [25 Baden 25] side was that Karl III August would decide on the portion in which [26 Birkenfeld 26] and the family seat were located ; nonetheless , Karl III August decided upon the [27 Mosel 27] region around , so that after 1776 the [28 Birkenfeld 28] region became the exclusive possession of [29 Baden 29] . His brother [5 5 Maximilian I Joseph 5] succeeded him after his death in 1795 . As a result of the wars of succession of the [6 6 French Revolution 6] he was the last Duke of [30 Zweibrücken 30] , becoming in 1799 [31 Palatine 31] and [32 Bavarian 32] Elector and in 1806 the first King of [33 Bavaria 33] . Prior to this , in 1645 , 10 John Charles , a younger brother of [34 11 Christian II 34] , had founded the collateral line of Palatinate-Birkenfeld-Gelnhausen . From 1799 on , his descendants were styled `` His Royal Highness 12 Duke in [35 Bavaria 35]

1: 1296944 Birkenfeld

2: 196475 Margrave

3: 21195350 Edward_Fortunatus

4: 58558 Zweibrücken

5: 295384 Maximilian_I_Joseph_of_Bavaria

6: 11188 French_Revolution

1: 11884 German_language

[winner: 11867 Germany, pem_pos=1, final_scores=0.22900127, final_local_scores=0.22204968, pem=-1.0465823, similarity=16.407597]

gt: 11884 German_language, pem_pos=0, final_scores=-0.10640159, final_local_scores=-0.12245491, pem=-0.981596, similarity=5.7734857]

2: 38846 Electoral_Palatinate

[winner: 159890 Palatine_Hill, pem_pos=1, final_scores=-0.87035286, final_local_scores=-0.9020298, pem=-1.3297042, similarity=-15.597518]

gt: 38846 Electoral_Palatinate, pem_pos=7, final_scores=-1.1413677, final_local_scores=-1.3393002, pem=-2.6429176, similarity=-21.505075]

3: 38848 House_of_Wittelsbach

[winner: 38848 House_of_Wittelsbach, pem_pos=0, final_scores=-0.8010484, final_local_scores=-0.9522055, pem=0.0, similarity=-24.325998]

gt: 38848 House_of_Wittelsbach, pem_pos=0, final_scores=-0.8010484, final_local_scores=-0.9522055, pem=0.0, similarity=-24.325998]

4: 3043377 Kingdom_of_Bavaria

[winner: 3764 Bavaria, pem_pos=1,

final_scores=0.54525685, final_local_scores=0.5191984,
 pem=-0.019504828, similarity=19.6897|
 gt: 3043377 Kingdom_of_Bavaria, pem_pos=2,
 final_scores=-0.13811882, final_local_scores=-0.23635408,
 pem=-2.4464288, similarity=10.340496]
 [winner: 80561 Kings_County,_California, pem_pos=2,
 final_scores=-0.87990594, final_local_scores=-0.8326823,
 pem=-1.9659233, similarity=-10.0681925|
 RECALL MISS: 3043377 Kingdom_of_Bavaria]
 5: 17645322 County_of_Sponheim
 [winner: 11855978 Sponheim, pem_pos=0,
 final_scores=0.082296535, final_local_scores=0.109714836,
 pem=-0.16061527, similarity=8.237273|
 gt: 17645322 County_of_Sponheim, pem_pos=3,
 final_scores=-0.18580163, final_local_scores=-0.23233908,
 pem=-1.6167793, similarity=5.948549]
 [winner: 382207 Powiat, pem_pos=0,
 final_scores=-1.1667044, final_local_scores=-1.2261508,
 pem=-0.39345494, similarity=-30.361403|
 RECALL MISS: 17645322 County_of_Sponheim]
 6: 21437566 Christian_II,_Count_Palatine_of_Zweibrücken-
 Birkenfeld
 [winner: 150846 Christian_II_of_Denmark, pem_pos=0,
 final_scores=0.35388893, final_local_scores=0.35105518,
 pem=0.0, similarity=14.5659|
 gt: 21437566 Christian_II,_Count_Palatine_of_Zweibrücken-
 Birkenfeld, pem_pos=2, final_scores=-0.20467262,
 final_local_scores=-0.1901887, pem=-1.6084604,
 similarity=7.1611605]
 7: 12581146 Christian_III,_Count_Palatine_of_Zweibrücken
 [winner: 150871 Christian_III_of_Denmark, pem_pos=0,
 final_scores=0.38532323, final_local_scores=0.37920094,
 pem=0.0, similarity=15.405825|
 gt: 12581146
 Christian_III,_Count_Palatine_of_Zweibrücken, pem_pos=1,
 final_scores=-0.06884821, final_local_scores=-0.05892156,
 pem=-1.2032946, similarity=8.875079]
 8: 12582368 Christian_IV,_Count_Palatine_of_Zweibrücken
 [winner: 150876 Christian_IV_of_Denmark, pem_pos=0,
 final_scores=0.19899367, final_local_scores=0.17580763,
 pem=0.0, similarity=9.3361635|
 gt: 12582368
 Christian_IV,_Count_Palatine_of_Zweibrücken, pem_pos=1,
 final_scores=-0.11741233, final_local_scores=-0.114669636,
 pem=-1.2933793, similarity=7.70134]
 9: 50213 Morganatic_marriage
 [winner: 13264900 WGMM, pem_pos=0,
 final_scores=-0.866629, final_local_scores=-0.7374741,
 pem=0.0, similarity=-17.917984|
 RECALL MISS: 50213 Morganatic_marriage]
 [winner: 1587943 Celestial_marriage, pem_pos=3,
 final_scores=-1.2828279, final_local_scores=-1.2928212,

pem=-1.3856965, similarity=-26.955013|
 RECALL MISS: 50213 Morganatic_marriage]
 10: 33368596 John_Charles,_Count_Palatine_of_Gelnhausen
 [winner: 480121 John_Charles, pem_pos=0,
 final_scores=-0.076292336, final_local_scores=0.027729556,
 pem=0.0, similarity=4.917218|
 gt: 33368596 John_Charles,_Count_Palatine_of_Gelnhausen,
 pem_pos=6, final_scores=-0.76999146,
 final_local_scores=-0.8515323, pem=-5.37412,
 similarity=7.9035926]
 11: 21437566 Christian_II,_Count_Palatine_of_Zweibrücken-
 Birkenfeld
 [winner: 150846 Christian_II_of_Denmark, pem_pos=0,
 final_scores=0.24530391, final_local_scores=0.23163098,
 pem=0.0, similarity=11.002044|
 gt: 21437566 Christian_II,_Count_Palatine_of_Zweibrücken-
 Birkenfeld, pem_pos=2, final_scores=-0.3082378,
 final_local_scores=-0.30503333, pem=-1.6084604,
 similarity=3.7339683]
 12: 2804342 Duke_in_Bavaria
 [winner: 3764 Bavaria, pem_pos=1, final_scores=0.2952783,
 final_local_scores=0.24613482, pem=-0.019504828,
 similarity=11.540937|
 gt: 2804342 Duke_in_Bavaria, pem_pos=17,
 final_scores=-0.64935905, final_local_scores=-0.72111994,
 pem=-2.944439, similarity=-1.417628]
 [winner: 148465 Stanisław_Leszczyński, pem_pos=0,
 final_scores=-0.6803743, final_local_scores=-0.7570859,
 pem=0.0, similarity=-18.50324|
 RECALL MISS: 2804342 Duke_in_Bavaria]
 1: 1296944 Birkenfeld
 [pem_pos=0, final_scores=0.16111557,
 final_local_scores=0.13361028, pem=-0.5507422,
 similarity=11.071932]
 2: 196475 Margrave
 [pem_pos=2, final_scores=0.17354028,
 final_local_scores=0.11537036, pem=-0.02950915,
 similarity=7.69307]
 3: 21195350 Edward_Fortunatus
 [pem_pos=0, final_scores=0.19036098,
 final_local_scores=0.24695528, pem=0.0, similarity=11.459351]
 4: 58558 Zweibrücken
 [pem_pos=0, final_scores=0.39360636,
 final_local_scores=0.37399793, pem=-0.14396068,
 similarity=16.033438]
 5: 295384 Maximilian_I_Joseph_of_Bavaria
 [pem_pos=0, final_scores=0.2791673,
 final_local_scores=0.23667654, pem=0.0, similarity=11.152613]
 6: 11188 French_Revolution
 [pem_pos=0, final_scores=0.18846337,
 final_local_scores=0.18079066, pem=0.0, similarity=9.484867]
 1: 11867 Germany

[pem_pos=1, final_scores=0.22900127,
final_local_scores=0.22204968, pem=-1.0465823,
similarity=16.407597]
2: 183760 Count_palatine
[pem_pos=0, final_scores=0.23909841,
final_local_scores=0.1968086, pem=0.0, similarity=9.962874]
3: 27773 Sophia_of_Hanover
[pem_pos=0, final_scores=0.071664855,
final_local_scores=0.06304881, pem=0.0, similarity=5.9712152]
4: 38848 House_of_Wittelsbach
[pem_pos=0, final_scores=0.23593377,
final_local_scores=0.1829252, pem=0.0, similarity=9.548566]
5: 20751523 House_of_Palatinate-Birkenfeld
[pem_pos=1, final_scores=0.03010033,
final_local_scores=0.017102033, pem=-0.028987534,
similarity=4.7577105]
6: 3764 Bavaria
[pem_pos=1, final_scores=0.54525685,
final_local_scores=0.5191984, pem=-0.019504828,
similarity=19.6897]
7: 27773 Sophia_of_Hanover
[pem_pos=0, final_scores=0.21129318,
final_local_scores=0.21570498, pem=0.0, similarity=10.52678]
8: 11855978 Sponheim
[pem_pos=0, final_scores=0.082296535,
final_local_scores=0.109714836, pem=-0.16061527,
similarity=8.237273]
9: 4264582 Kastellaun
[pem_pos=0, final_scores=0.17076282,
final_local_scores=0.19435892, pem=0.0, similarity=9.8897705]
10: 1296944 Birkenfeld
[pem_pos=0, final_scores=0.21294759,
final_local_scores=0.19022888, pem=-0.5507422,
similarity=12.761544]
11: 11880951 Allenbach
[pem_pos=0, final_scores=0.05009289,
final_local_scores=0.11375016, pem=0.0, similarity=7.484245]
12: 150846 Christian_II_of_Denmark
[pem_pos=0, final_scores=0.35388893,
final_local_scores=0.35105518, pem=0.0, similarity=14.5659]
13: 8113490 George_William,_Duke_of_Liegnitz
[pem_pos=0, final_scores=0.049263403,
final_local_scores=0.042186007, pem=-0.8298582,
similarity=9.861526]
14: 21437584 Palatinate-Birkenfeld-Bischweiler
[pem_pos=0, final_scores=0.059374705,
final_local_scores=0.09934835, pem=0.0, similarity=7.054466]
15: 150871 Christian_III_of_Denmark
[pem_pos=0, final_scores=0.38532323,
final_local_scores=0.37920094, pem=0.0, similarity=15.405825]
16: 1296944 Birkenfeld
[pem_pos=0, final_scores=0.1698306,

final_local_scores=0.14313012, pem=-0.5507422,
similarity=11.356023]
17: 3909917 Bischwiller
[pem_pos=0, final_scores=0.059741393,
final_local_scores=0.1652275, pem=0.0, similarity=9.0204315]
18: 58558 Zweibrücken
[pem_pos=0, final_scores=0.37933087,
final_local_scores=0.35840413, pem=-0.14396068,
similarity=15.568087]
19: 150876 Christian_IV_of_Denmark
[pem_pos=0, final_scores=0.19899367,
final_local_scores=0.17580763, pem=0.0, similarity=9.3361635]
20: 150876 Christian_IV_of_Denmark
[pem_pos=0, final_scores=0.11004208,
final_local_scores=0.07864155, pem=0.0, similarity=6.4365335]
21: 11855978 Sponheim
[pem_pos=0, final_scores=0.12741987,
final_local_scores=0.15900522, pem=-0.16061527,
similarity=9.708197]
22: 40684169 Baden
[pem_pos=0, final_scores=0.11342548,
final_local_scores=0.039093405, pem=-0.63090813,
similarity=8.687315]
23: 183760 Count_palatine
[pem_pos=0, final_scores=0.10293244,
final_local_scores=0.047473222, pem=0.0, similarity=5.5064087]
24: 25845 Rhine
[pem_pos=0, final_scores=0.27128255,
final_local_scores=0.25416285, pem=0.0, similarity=11.674439]
25: 40684169 Baden
[pem_pos=2, final_scores=0.08139114,
final_local_scores=0.0041007698, pem=-0.63090813,
similarity=7.6430645]
26: 1296944 Birkenfeld
[pem_pos=0, final_scores=0.12627594,
final_local_scores=0.09555328, pem=-0.5507422,
similarity=9.936235]
27: 143868 Moselle_(river)
[pem_pos=1, final_scores=0.380426,
final_local_scores=0.39020562, pem=-1.0370426,
similarity=21.373825]
28: 1296944 Birkenfeld
[pem_pos=0, final_scores=0.10205452,
final_local_scores=0.069095075, pem=-0.5507422,
similarity=9.146669]
29: 40684169 Baden
[pem_pos=2, final_scores=0.12474622,
final_local_scores=0.05145961, pem=-0.63090813,
similarity=9.056347]
30: 58558 Zweibrücken
[pem_pos=0, final_scores=0.283983,
final_local_scores=0.25425103, pem=-0.14396068,

```

similarity=12.459951]
31: 27773 Sophia_of_Hanover
    [pem_pos=0, final_scores=0.104479,
final_local_scores=0.09902671, pem=0.0, similarity=7.0448675]
32: 3764 Bavaria
    [pem_pos=0, final_scores=0.39485317,
final_local_scores=0.35490522, pem=-0.05476965,
similarity=14.978639]
33: 3764 Bavaria
    [pem_pos=1, final_scores=0.5067163,
final_local_scores=0.47709876, pem=-0.019504828,
similarity=18.433363]
34: 150846 Christian_II_of_Denmark
    [pem_pos=0, final_scores=0.24530391,
final_local_scores=0.23163098, pem=0.0, similarity=11.002044]
35: 3764 Bavaria
    [pem_pos=1, final_scores=0.2952783,
final_local_scores=0.24613482, pem=-0.019504828,
similarity=11.540937]

```

Sample: b'James_Barrow'

James Keith Barrow -LRB- born 16 December 1964 -RRB- is a former [1 1 English 1] . Barrow was a right-handed who bowled right-arm 2 fast-medium . He was born at [1 1 Haslemere 1] , [2 3 Surrey 2] . Barrow made his debut for [2 2 Berkshire 2] in 1990 against [3 3 Cornwall 3] . From 1990 to 2000 , he represented the county in 57 [3 Minor Counties Championship 3] matches , the last of which came in the 2000 Championship when [4 Berkshire 4] played [4 4 Oxfordshire 4] . Barrow also played in the [5 5 MCCA Knockout Trophy 5] for [5 Berkshire 5] . His debut in that competition came in 1992 when [6 Berkshire 6] played [6 6 Buckinghamshire 6] . From 1992 to 2000 , he represented the county in 20 Trophy matches , the last of which came when [7 Berkshire 7] played [7 7 Herefords hire 7] in the 2000 [8 MCCA Knockout Trophy 8] . Additionally , he also played matches for [9 Berkshire 9] . His [10 List-A 10] debut for the county came against [8 8 Kent 8] in the 4 1994 [11 NatWest Trophy 11] . From 1994 to 2001 , he represented the county in 9 matches , with his final [12 List-A 12] match coming when [13 Berkshire 13] played [9 9 Lincolnshire 9] in the 1st round of the 5 2002 [14 Cheltenham & Gloucester Trophy 14] , which was played in 2001 at [10 10 Lindum Sports Club Ground 10] , . In his 9 matches , he scored 21 runs at a 6 batting average of 7.00 , with a high score of 16 . With the ball he took 11 wickets at a 7 bowling average

```

1: 336911 Haslemere
2: 3702129 Berkshire_County_Cricket_Club
3: 3702042 Cornwall_County_Cricket_Club
4: 3712398 Oxfordshire_County_Cricket_Club
5: 7543416 MCCA_Knockout_Trophy
6: 3702814 Buckinghamshire_County_Cricket_Club
7: 3702547 Herefordshire_County_Cricket_Club

```

8: 1177719 Kent_County_Cricket_Club
 9: 3712244 Lincolnshire_County_Cricket_Club
 10: 28401320 Lindum_Sports_Club_Ground
 1: 18803164 English_people
 [winner: 9316 England, pem_pos=1,
 final_scores=0.31121868, final_local_scores=0.34522936,
 pem=-1.3527626, similarity=21.748579]
 gt: 18803164 English_people, pem_pos=2,
 final_scores=-0.4362503, final_local_scores=-0.46113938,
 pem=-2.2084396, similarity=2.3382318]
 2: 678695 Fast_bowling
 [winner: 678695 Fast_bowling, pem_pos=0,
 final_scores=-1.1835798, final_local_scores=-1.3303871,
 pem=0.0, similarity=-35.61169]
 gt: 678695 Fast_bowling, pem_pos=0,
 final_scores=-1.1835798, final_local_scores=-1.3303871,
 pem=0.0, similarity=-35.61169]
 3: 48189 Surrey
 [winner: 766575 Surrey_County_Cricket_Club, pem_pos=1,
 final_scores=0.25776798, final_local_scores=0.1831393,
 pem=-1.7952642, similarity=19.31788]
 gt: 48189 Surrey, pem_pos=0, final_scores=0.12807672,
 final_local_scores=0.13403302, pem=-0.28536296,
 similarity=9.641374]
 4: 27619367 1994_NatWest_Trophy
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.31731814, final_local_scores=0.2733494,
 pem=0.0, similarity=12.2470045]
 gt: 27619367 1994_NatWest_Trophy, pem_pos=10,
 final_scores=-0.51291317, final_local_scores=-0.59084046,
 pem=-4.076326, similarity=8.625549]
 [winner: 18948337 1994, pem_pos=0,
 final_scores=-0.9331207, final_local_scores=-0.8540284,
 pem=-0.7539553, similarity=-17.296072]
 RECALL MISS: 27619367 1994_NatWest_Trophy]
 5: 27764855 2002_Cheltenham_&Gloucester_Trophy
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.2574826, final_local_scores=0.20798814,
 pem=0.0, similarity=10.296494]
 gt: 27764855 2002_Cheltenham_&Gloucester_Trophy,
 pem_pos=3, final_scores=-0.3742863,
 final_local_scores=-0.44501904, pem=-3.295837,
 similarity=8.73273]
 [winner: 35502 2002, pem_pos=0, final_scores=-1.0676014,
 final_local_scores=-0.9643007, pem=-0.5897454,
 similarity=-21.479816]
 RECALL MISS: 27764855
 2002_Cheltenham_&Gloucester_Trophy]
 6: 18933273 Batting_average
 [winner: 15652278 Batting(material), pem_pos=2,
 final_scores=-1.0294446, final_local_scores=-0.92542803,
 pem=-1.1616125, similarity=-17.209875]

gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.3411313, final_local_scores=-1.4798768,
pem=-0.29142123, similarity=-38.48797]
[winner: 875871 Bowling_average, pem_pos=1,
final_scores=-1.2693092, final_local_scores=-1.4254684,
pem=-0.8657023, similarity=-33.741283]
gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.394173, final_local_scores=-1.5378168,
pem=-0.5012068, similarity=-39.076168]
7: 875871 Bowling_average
[winner: 875871 Bowling_average, pem_pos=0,
final_scores=-1.0300256, final_local_scores=-1.1640875,
pem=0.0, similarity=-30.64898]
gt: 875871 Bowling_average, pem_pos=0,
final_scores=-1.0300256, final_local_scores=-1.1640875,
pem=0.0, similarity=-30.64898]
1: 336911 Haslemere
[pem_pos=0, final_scores=0.12934123,
final_local_scores=0.18234235, pem=0.0, similarity=9.531173]
2: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.49258965,
final_local_scores=0.42571086, pem=-2.301417,
similarity=29.309238]
3: 3702042 Cornwall_County_Cricket_Club
[pem_pos=2, final_scores=0.13850488,
final_local_scores=0.055773437, pem=-3.34357,
similarity=23.936958]
4: 3712398 Oxfordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.31574714,
final_local_scores=0.25183848, pem=-1.7495553,
similarity=21.119429]
5: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.4570958,
final_local_scores=0.4104151, pem=0.0, similarity=16.337318]
6: 3702814 Buckinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.38844436,
final_local_scores=0.32165363, pem=-1.6046519,
similarity=22.414845]
7: 3702547 Herefordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6107746,
final_local_scores=0.5835588, pem=-1.3582021,
similarity=28.890387]
8: 1177719 Kent_County_Cricket_Club
[pem_pos=1, final_scores=0.7626414,
final_local_scores=0.73561305, pem=-2.5096242,
similarity=39.6896]
9: 3712244 Lincolnshire_County_Cricket_Club
[pem_pos=2, final_scores=0.44261116,
final_local_scores=0.3981188, pem=-2.1349053,
similarity=27.58032]
10: 28401320 Lindum_Sports_Club_Ground
[pem_pos=0, final_scores=0.1661282,

final_local_scores=0.16741282, pem=0.0, similarity=9.085646]
1: 9316 England
[pem_pos=1, final_scores=0.31121868,
final_local_scores=0.34522936, pem=-1.3527626,
similarity=21.748579]
2: 766575 Surrey_County_Cricket_Club
[pem_pos=1, final_scores=0.25776798,
final_local_scores=0.1831393, pem=-1.7952642,
similarity=19.31788]
3: 7562952 Minor_Counties_Cricket_Championship
[pem_pos=0, final_scores=0.3082279,
final_local_scores=0.23880136, pem=0.0, similarity=11.216022]
4: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.45617038,
final_local_scores=0.38592833, pem=-2.301417,
similarity=28.122047]
5: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.5462716,
final_local_scores=0.48435026, pem=-2.301417,
similarity=31.059156]
6: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.5633473,
final_local_scores=0.5030028, pem=-2.301417,
similarity=31.615786]
7: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.6762008,
final_local_scores=0.6262781, pem=-2.301417,
similarity=35.294567]
8: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.17962031,
final_local_scores=0.10731532, pem=0.0, similarity=7.292216]
9: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.5903756,
final_local_scores=0.53252715, pem=-2.301417,
similarity=32.49685]
10: 683459 List_A_cricket
[pem_pos=0, final_scores=0.2767167,
final_local_scores=0.22367978, pem=-0.07478003,
similarity=11.17143]
11: 1899702 Friends_Provident_Trophy
[pem_pos=0, final_scores=0.31731814,
final_local_scores=0.2733494, pem=0.0, similarity=12.2470045]
12: 683459 List_A_cricket
[pem_pos=0, final_scores=0.16432975,
final_local_scores=0.10091412, pem=-0.07478003,
similarity=7.507857]
13: 3702129 Berkshire_County_Cricket_Club
[pem_pos=2, final_scores=0.6782667,
final_local_scores=0.62853485, pem=-2.301417,
similarity=35.36191]
14: 1899702 Friends_Provident_Trophy
[pem_pos=0, final_scores=0.2574826,

final_local_scores=0.20798814, pem=0.0, similarity=10.296494]

Sample: b'Julius_Scriver'

Julius Scriver -LRB- February 5 , 1826 -- September 5 , 1907 -RRB- was a politician in [1 1 Quebec 1] , [2 2 Canada 2] . Born in , [3 3 Lower Canada 3] -LRB- now [4 4 Quebec 4] -RRB- , the son of and Lucretia Manning , he studied at the Workman 's School in [1 Montreal 1] and the [5 5 University of Vermont 5] . He became a 1 miller and 2 tanner in [2 Hemmingford 2] . In 1867 , he was elected to the [6 6 Legislative Assembly of Quebec 6] representing the provincial riding of . In a 1869 by-election , he was acclaimed as 3 Liberal Member of Parliament in the federal riding of [3 4 Huntingdon 3] . He was re-elected in 5 1872 -LRB- acclaimed -RRB- , 6 1874 , 7 1878 -LRB- acclaimed -RRB- , 8 1882 , 9 1887 -LRB- acclaimed -RRB- , 10 1891 , and 11 1896 . He died at [7 7 Westmount 7]

1: 7954867 Quebec

2: 5042916 Canada

3: 45945 Lower_Canada

4: 7954867 Quebec

5: 287800 University_of_Vermont

6: 517181 Legislative_Assembly_of_Quebec

7: 189886 Westmount,_Quebec

1: 2159644 Miller

[winner: 21088844 Alana_Miller, pem_pos=0, final_scores=-0.8479209, final_local_scores=-0.78848773, pem=0.0, similarity=-19.440334]

RECALL MISS: 2159644 Miller]

2: 140990 Tanning

[winner: 26184746 Väinö_Tanner's_cabinet, pem_pos=6, final_scores=-1.2730352, final_local_scores=-1.2094655, pem=-3.3775365, similarity=-13.635579]

gt: 140990 Tanning, pem_pos=13, final_scores=-1.7215528, final_local_scores=-1.8509583, pem=-3.944643, similarity=-29.694994]

3: 149536 Liberal_Party_of_Canada

[winner: 149536 Liberal_Party_of_Canada, pem_pos=0, final_scores=-0.23939915, final_local_scores=-0.30695656, pem=-1.2222316, similarity=1.5762028]

gt: 149536 Liberal_Party_of_Canada, pem_pos=0, final_scores=-0.23939915, final_local_scores=-0.30695656, pem=-1.2222316, similarity=1.5762028]

4: 1787071 Huntingdon_(electoral_district)

[winner: 250589 Huntingdon, pem_pos=0, final_scores=0.20436625, final_local_scores=0.21862078, pem=-0.16110621, similarity=11.489914]

gt: 1787071 Huntingdon_(electoral_district), pem_pos=10, final_scores=-0.564826, final_local_scores=-0.5604871, pem=-3.077146, similarity=4.0976586]

5: 488354 Canadian_federal_election, 1872

[winner: 34768 1872, pem_pos=0, final_scores=-0.80933124,

final_local_scores=-0.6914506, pem=-0.4293672,
similarity=-14.209587|
gt: 488354 Canadian_federal_election,_1872, pem_pos=4,
final_scores=-1.1632617, final_local_scores=-1.2442956,
pem=-2.9551756, similarity=-16.971842]
6: 711285 Canadian_federal_election,_1874
[winner: 34770 1874, pem_pos=0, final_scores=-0.8512741,
final_local_scores=-0.7254867, pem=-0.49628752,
similarity=-14.861368|
gt: 711285 Canadian_federal_election,_1874, pem_pos=4,
final_scores=-1.1242521, final_local_scores=-1.1900375,
pem=-2.9089265, similarity=-15.604183]
7: 1044962 Canadian_federal_election,_1878
[winner: 34938 1878, pem_pos=0, final_scores=-0.83958966,
final_local_scores=-0.6911445, pem=-0.4579571,
similarity=-14.044976|
gt: 1044962 Canadian_federal_election,_1878, pem_pos=4,
final_scores=-1.436379, final_local_scores=-1.5526533,
pem=-3.1359437, similarity=-25.190807]
8: 1053595 Canadian_federal_election,_1882
[winner: 34772 1882, pem_pos=0, final_scores=-0.80953777,
final_local_scores=-0.6831399, pem=-0.49682418,
similarity=-13.594738|
gt: 1053595 Canadian_federal_election,_1882, pem_pos=3,
final_scores=-1.0980324, final_local_scores=-1.1799033,
pem=-2.9665563, similarity=-14.988358]
9: 1053609 Canadian_federal_election,_1887
[winner: 34777 1887, pem_pos=0, final_scores=-1.0411577,
final_local_scores=-0.9453694, pem=-0.51339984,
similarity=-21.330048|
gt: 1053609 Canadian_federal_election,_1887, pem_pos=3,
final_scores=-1.236219, final_local_scores=-1.3153555,
pem=-3.0445802, similarity=-18.606216]
10: 524238 Canadian_federal_election,_1891
[winner: 34794 1891, pem_pos=0, final_scores=-0.9803683,
final_local_scores=-0.87412745, pem=-0.51282144,
similarity=-19.207191|
gt: 524238 Canadian_federal_election,_1891, pem_pos=2,
final_scores=-1.185429, final_local_scores=-1.2824674,
pem=-3.1970875, similarity=-16.795412]
11: 481060 Canadian_federal_election,_1896
[winner: 34668 1896, pem_pos=0, final_scores=-0.8131514,
final_local_scores=-0.6975419, pem=-0.73694056,
similarity=-12.718731|
gt: 481060 Canadian_federal_election,_1896, pem_pos=5,
final_scores=-1.1747774, final_local_scores=-1.2844803,
pem=-3.4991813, similarity=-15.212652]
1: 7954867 Quebec
[pem_pos=0, final_scores=0.45537835,
final_local_scores=0.41030747, pem=-0.089177474,
similarity=16.819067]
2: 5042916 Canada

[pem_pos=0, final_scores=0.5476436,
 final_local_scores=0.5498149, pem=-0.11662793,
 similarity=21.131527]
 3: 45945 Lower_Canada
 [pem_pos=0, final_scores=0.059014037,
 final_local_scores=0.015222415, pem=0.0, similarity=4.54398]
 4: 7954867 Quebec
 [pem_pos=0, final_scores=0.30392247,
 final_local_scores=0.24486497, pem=-0.089177474,
 similarity=11.881933]
 5: 287800 University_of_Vermont
 [pem_pos=0, final_scores=0.31845492,
 final_local_scores=0.3681947, pem=0.0, similarity=15.077377]
 6: 517181 Legislative_Assembly_of_Quebec
 [pem_pos=0, final_scores=0.12503977,
 final_local_scores=0.12364194, pem=0.0, similarity=7.779435]
 7: 189886 Westmount,_Quebec
 [pem_pos=0, final_scores=0.16585709,
 final_local_scores=0.13427088, pem=-0.37007916,
 similarity=10.109172]
 1: 7954681 Montreal
 [pem_pos=0, final_scores=0.47433418,
 final_local_scores=0.4560755, pem=-0.11023275,
 similarity=18.299377]
 2: 7417280 Hemmingford,_Quebec_(township)
 [pem_pos=1, final_scores=0.049955145,
 final_local_scores=0.11622575, pem=-1.0654339,
 similarity=13.352116]
 3: 250589 Huntingdon
 [pem_pos=0, final_scores=0.20436625,
 final_local_scores=0.21862078, pem=-0.16110621,
 similarity=11.489914]

Sample: b'Penske_PC-22'

The [1 Penske PC-22 1] was a 1 CART Penske Racing car which
 competed in the 2 1993 season . It raced in all sixteen
 events , scoring eight wins , three with [1 1 Emerson Fittipa
 ldi 1] and five with [2 2 Paul Tracy 2] . The most remarkable
 success of the [2 PC-22 2] was the 3 1993 [3 Indianapolis 500
 3] win by [4 Fittipaldi 4] , placed 2nd at the end of the
 season , missing the championship by only 8 points . The car
 was designed by as a radical departure from the basic concept
 of the previous [5 Penske 5] cars . The [6 PC-22 6] debuted
 at 4 1993 [7 Australian 7] FAI IndyCar Grand Prix , noted for
 the sensational debut win of [3 3 Nigel Mansell 3] ,
 qualifying 2nd with [8 Fittipaldi 8] and 3rd with [9 Tracy 9]
 . In the race , [10 Fittipaldi 10] scored a 2nd place
 finish , while [11 Tracy 11] retired for electrical
 problems . At [12 5 Phoenix 12] , the first oval race of the
 season , both Penskes went off , while in 6 Long Beach [13
 Tracy 13] scored the first win for the [14 PC-22 14] . In [15
 7 Indianapolis 15] , [16 Fittipaldi 16] qualified 9th -LRB-

220.150 mph -RRB- , and won the event exploiting [17 Mansell 17] 's inexperience on oval racing and [18 Andretti 18] 's handling problems , after their dominating runs . After two races , and , with only one podium finish , the [19 Penske 19] come back to victory in [20 8 Portland 20] , with [21 Fittipaldi 21] 1st and [22 Tracy 22] 3rd . Two 1-2 wins followed , in [23 9 Cleveland 23] and [24 10 Toronto 24] , with [25 Tracy 25] 1st and [26 Fittipaldi 26] 2nd in both occasions . [27 Nigel Mansell 27] scored two wins at [28 11 Michigan 28] and 12 New Hampshire as a major push for his championship bid , while [29 Fittipaldi 29] and [30 Tracy 30] at New Hampshire scored a 2nd and 3rd place respectively . [31 Tracy 31] won at [32 13 Road America 32] and [33 Fittipaldi 33] at , but [34 Mansell 34] clinched the title with his 5th season win at [35 14 Nazareth 35] . [36 Tracy 36] scored a 1st place at the final race of season in [37 15 Laguna Seca 37] . The highly successful 16 PC-23 that competed in the 17 1994 [38 IndyCar 38] season . was modeled after its predecessor , the PC-22 . The [39 Penske PC-23 39] was one of the most dominant open wheel race cars ever developed . It won both the 1994 [40 CART 40] season , and the 18 1994 [41 Indianapolis 500 41] with , together with [4 4 Emerson Fittipaldi 4] and [5 5 Paul Tracy 5]

1: 241678 Emerson_Fittipaldi

2: 350364 Paul_Tracy

3: 228615 Nigel_Mansell

4: 241678 Emerson_Fittipaldi

5: 350364 Paul_Tracy

1: 225290 Champ_Car

[winner: 225290 Champ_Car, pem_pos=0, final_scores=-0.9509632, final_local_scores=-1.1234496, pem=-0.031101579, similarity=-29.267128|

gt: 225290 Champ_Car, pem_pos=0, final_scores=-0.9509632, final_local_scores=-1.1234496, pem=-0.031101579, similarity=-29.267128]

2: 7512213 1993_IndyCar_season

[winner: 34598 1993, pem_pos=0, final_scores=-0.70928854, final_local_scores=-0.5065566, pem=-0.7063994, similarity=-7.185436|

RECALL MISS: 7512213 1993_IndyCar_season]

[winner: 27822865 2011_NASCAR_Sprint_Cup_Series, pem_pos=5, final_scores=-1.4707819, final_local_scores=-1.6261463, pem=-3.709535, similarity=-24.264706|

RECALL MISS: 7512213 1993_IndyCar_season]

3: 5222439 1993_Indianapolis_500

[winner: 159018 Indianapolis_500, pem_pos=0, final_scores=0.24358796, final_local_scores=0.19807357, pem=-0.024602843, similarity=10.134417|

RECALL MISS: 5222439 1993_Indianapolis_500]

[winner: 34598 1993, pem_pos=0, final_scores=-0.71598566, final_local_scores=-0.5138722, pem=-0.7063994,

similarity=-7.4037485|
 RECALL MISS: 5222439 1993_Indianapolis_500]
 4: 19065453 1993_Australian_FAI_IndyCar_Grand_Prix
 [winner: 4689264 Australia, pem_pos=0,
 final_scores=0.21746291, final_local_scores=0.33417404,
 pem=-0.491009, similarity=16.732315|
 RECALL MISS: 19065453
 1993_Australian_FAI_IndyCar_Grand_Prix]
 [winner: 13970439 Coop_Prix, pem_pos=1,
 final_scores=-0.69737166, final_local_scores=-0.54699016,
 pem=-0.6931472, similarity=-8.464122|
 RECALL MISS: 19065453
 1993_Australian_FAI_IndyCar_Grand_Prix]
 [winner: 34598 1993, pem_pos=0, final_scores=-0.71118224,
 final_local_scores=-0.50862515, pem=-0.7063994,
 similarity=-7.2471666|
 RECALL MISS: 19065453
 1993_Australian_FAI_IndyCar_Grand_Prix]
 [winner: 654087 Fédération_Aéronautique_Internationale,
 pem_pos=0, final_scores=-1.0283917,
 final_local_scores=-1.038261, pem=-0.6180443,
 similarity=-23.533047|
 RECALL MISS: 19065453
 1993_Australian_FAI_IndyCar_Grand_Prix]
 [winner: 96992 Grand_County,_Colorado, pem_pos=0,
 final_scores=-1.3292032, final_local_scores=-1.296725,
 pem=-1.487728, similarity=-26.51665|
 RECALL MISS: 19065453
 1993_Australian_FAI_IndyCar_Grand_Prix]
 [winner: 16099880 2010_IndyCar_Series_season, pem_pos=3,
 final_scores=-1.8225613, final_local_scores=-2.0118132,
 pem=-1.3862944, similarity=-48.407925|
 RECALL MISS: 19065453
 1993_Australian_FAI_IndyCar_Grand_Prix]
 5: 8648925 XM_Satellite_Radio_Indy_200
 [winner: 49121 Phoenix,_Arizona, pem_pos=0,
 final_scores=0.22364752, final_local_scores=0.3009316,
 pem=-0.90867877, similarity=18.011648|
 RECALL MISS: 8648925 XM_Satellite_Radio_Indy_200]
 6: 967558 Grand_Prix_of_Long_Beach
 [winner: 94240 Long_Beach,_California, pem_pos=0,
 final_scores=-0.8782895, final_local_scores=-0.9096838,
 pem=-0.4406212, similarity=-20.6609|
 gt: 967558 Grand_Prix_of_Long_Beach, pem_pos=4,
 final_scores=-1.0644438, final_local_scores=-1.2034271,
 pem=-2.7638814, similarity=-16.792536]
 7: 5222439 1993_Indianapolis_500
 [winner: 57707 Indianapolis, pem_pos=0,
 final_scores=0.46852177, final_local_scores=0.55787104,
 pem=-0.30184713, similarity=22.379189|
 RECALL MISS: 5222439 1993_Indianapolis_500]
 8: 8704841 Grand_Prix_of_Portland

[winner: 23503 Portland,_Oregon, pem_pos=0,
 final_scores=0.25785387, final_local_scores=0.3530814,
 pem=-0.66320467, similarity=18.232977]
 RECALL MISS: 8704841 Grand_Prix_of_Portland]
 9: 5199239 Grand_Prix_of_Cleveland
 [winner: 5951 Cleveland, pem_pos=0,
 final_scores=0.42678243, final_local_scores=0.5343282,
 pem=-0.57246673, similarity=23.148296]
 RECALL MISS: 5199239 Grand_Prix_of_Cleveland]
 10: 2975638 Honda_Indy_Toronto
 [winner: 64646 Toronto, pem_pos=0,
 final_scores=0.52745646, final_local_scores=0.6391832,
 pem=-0.12291032, similarity=23.83262]
 gt: 2975638 Honda_Indy_Toronto, pem_pos=10,
 final_scores=-0.68359995, final_local_scores=-0.77993333,
 pem=-4.4919505, similarity=5.2428718]
 11: 6734849 Firestone_Indy_400
 [winner: 18859 Michigan, pem_pos=0,
 final_scores=0.1434413, final_local_scores=0.23787597,
 pem=-0.24788691, similarity=12.536454]
 RECALL MISS: 6734849 Firestone_Indy_400]
 12: 1769428 New_Hampshire_Motor_Speedway
 [winner: 21134 New_Hampshire, pem_pos=0,
 final_scores=-0.02253826, final_local_scores=0.06900048,
 pem=-0.30258936, similarity=7.794353]
 gt: 1769428 New_Hampshire_Motor_Speedway, pem_pos=3,
 final_scores=-0.34392583, final_local_scores=-0.40853205,
 pem=-3.4053843, similarity=10.417311]
 13: 10613423 Champ_Car_Grand_Prix_of_Road_America
 [winner: 1776945 Road_America, pem_pos=0,
 final_scores=0.29358947, final_local_scores=0.2725003,
 pem=0.0, similarity=12.221665]
 gt: 10613423 Champ_Car_Grand_Prix_of_Road_America,
 pem_pos=2, final_scores=-0.5146246,
 final_local_scores=-0.5426427, pem=-3.8436792,
 similarity=8.7986965]
 14: 8628181 Firestone_Indy_225
 [winner: 21702 Nazareth, pem_pos=0,
 final_scores=0.14593081, final_local_scores=0.26970312,
 pem=-0.38387662, similarity=14.225773]
 gt: 8628181 Firestone_Indy_225, pem_pos=6,
 final_scores=-0.3571598, final_local_scores=-0.38384375,
 pem=-3.1681342, similarity=9.863856]
 15: 13234345 Monterey_Grand_Prix
 [winner: 676326 Mazda_Raceway_Laguna_Seca, pem_pos=0,
 final_scores=0.22869597, final_local_scores=0.1956057,
 pem=-0.039963335, similarity=10.144304]
 gt: 13234345 Monterey_Grand_Prix, pem_pos=2,
 final_scores=-0.26164204, final_local_scores=-0.26532573,
 pem=-2.0847936, similarity=7.5092955]
 16: 27747173 Penske_PC-23
 [winner: 27747173 Penske_PC-23, pem_pos=0,

```

final_scores=0.004939869, final_local_scores=-0.016050272,
pem=0.0, similarity=3.610741|
    gt: 27747173 Penske_PC-23, pem_pos=0,
final_scores=0.004939869, final_local_scores=-0.016050272,
pem=0.0, similarity=3.610741 ]
17: 19908651 1994_IndyCar_season
    [winner: 30446232 IndyCar, pem_pos=0,
final_scores=0.16864933, final_local_scores=0.11490241,
pem=-0.25514126, similarity=8.906128|
    RECALL MISS: 19908651 1994_IndyCar_season]
    [winner: 18948337 1994, pem_pos=0,
final_scores=-1.2880157, final_local_scores=-1.2295954,
pem=-0.7539553, similarity=-28.503744|
    RECALL MISS: 19908651 1994_IndyCar_season]
    [winner: 690345 Season(sports), pem_pos=0,
final_scores=-1.841968, final_local_scores=-1.9339781,
pem=-1.219945, similarity=-46.9898|
    RECALL MISS: 19908651 1994_IndyCar_season]
18: 5222433 1994_Indianapolis_500
    [winner: 159018 Indianapolis_500, pem_pos=0,
final_scores=0.24653871, final_local_scores=0.2012968,
pem=-0.024602843, similarity=10.230605|
    RECALL MISS: 5222433 1994_Indianapolis_500]
    [winner: 18948337 1994, pem_pos=0,
final_scores=-1.0934373, final_local_scores=-1.0170482,
pem=-0.7539553, similarity=-22.16091|
    RECALL MISS: 5222433 1994_Indianapolis_500]
1: 241678 Emerson_Fittipaldi
    [pem_pos=0, final_scores=0.5125469,
final_local_scores=0.45618087, pem=0.0, similarity=17.70306]
2: 350364 Paul_Tracy
    [pem_pos=0, final_scores=0.45488006,
final_local_scores=0.41735947, pem=0.0, similarity=16.544552]
3: 228615 Nigel_Mansell
    [pem_pos=0, final_scores=0.54507405,
final_local_scores=0.5115472, pem=0.0, similarity=19.355303]
4: 241678 Emerson_Fittipaldi
    [pem_pos=0, final_scores=0.50745314,
final_local_scores=0.45061672, pem=0.0, similarity=17.537014]
5: 350364 Paul_Tracy
    [pem_pos=0, final_scores=0.2640924,
final_local_scores=0.20895287, pem=0.0, similarity=10.325283]
1: 27747149 Penske_PC-22
    [pem_pos=0, final_scores=0.0414695,
final_local_scores=0.046946645, pem=0.0, similarity=5.4906945]
2: 27747149 Penske_PC-22
    [pem_pos=0, final_scores=0.09811051,
final_local_scores=0.10881832, pem=0.0, similarity=7.3370686]
3: 159018 Indianapolis_500
    [pem_pos=0, final_scores=0.24358796,
final_local_scores=0.19807357, pem=-0.024602843,
similarity=10.134417]

```

4: 241678 Emerson_Fittipaldi
[pem_pos=0, final_scores=0.4106639,
final_local_scores=0.34556168, pem=0.0, similarity=14.401963]
5: 2321219 Team_Penske
[pem_pos=0, final_scores=0.2480918,
final_local_scores=0.19540519, pem=0.0, similarity=9.920994]
6: 27747149 Penske_PC-22
[pem_pos=0, final_scores=0.09595342,
final_local_scores=0.10646203, pem=0.0, similarity=7.2667522]
7: 4689264 Australia
[pem_pos=0, final_scores=0.21746291,
final_local_scores=0.33417404, pem=-0.491009,
similarity=16.732315]
8: 241678 Emerson_Fittipaldi
[pem_pos=0, final_scores=0.4577325,
final_local_scores=0.39697695, pem=0.0, similarity=15.936296]
9: 350364 Paul_Tracy
[pem_pos=0, final_scores=0.7243542,
final_local_scores=0.71171904, pem=0.0, similarity=25.328829]
10: 241678 Emerson_Fittipaldi
[pem_pos=0, final_scores=0.545119,
final_local_scores=0.49243343, pem=0.0, similarity=18.784908]
11: 350364 Paul_Tracy
[pem_pos=0, final_scores=0.7045619,
final_local_scores=0.6900989, pem=0.0, similarity=24.68364]
12: 49121 Phoenix,_Arizona
[pem_pos=0, final_scores=0.22364752,
final_local_scores=0.3009316, pem=-0.90867877,
similarity=18.011648]
13: 350364 Paul_Tracy
[pem_pos=0, final_scores=0.29279572,
final_local_scores=0.24030688, pem=0.0, similarity=11.26095]
14: 27747149 Penske_PC-22
[pem_pos=0, final_scores=0.06406449,
final_local_scores=0.07162824, pem=0.0, similarity=6.2272425]
15: 57707 Indianapolis
[pem_pos=0, final_scores=0.46852177,
final_local_scores=0.55787104, pem=-0.30184713,
similarity=22.379189]
16: 241678 Emerson_Fittipaldi
[pem_pos=0, final_scores=0.43322033,
final_local_scores=0.3702011, pem=0.0, similarity=15.137252]
17: 228615 Nigel_Mansell
[pem_pos=0, final_scores=0.508231,
final_local_scores=0.47268867, pem=0.0, similarity=18.195686]
18: 271425 Mario_Andretti
[pem_pos=0, final_scores=0.32933635,
final_local_scores=0.27507076, pem=0.0, similarity=12.298373]
19: 2321219 Team_Penske
[pem_pos=0, final_scores=0.27885157,
final_local_scores=0.22900552, pem=0.0, similarity=10.923695]
20: 23503 Portland,_Oregon

[pem_pos=0, final_scores=0.25785387,
final_local_scores=0.3530814, pem=-0.66320467,
similarity=18.232977]
21: 241678 Emerson_Fittipaldi
[pem_pos=0, final_scores=0.36261553,
final_local_scores=0.29307607, pem=0.0, similarity=12.835688]
22: 350364 Paul_Tracy
[pem_pos=0, final_scores=0.6132694,
final_local_scores=0.5903757, pem=0.0, similarity=21.707703]
23: 5951 Cleveland
[pem_pos=0, final_scores=0.42678243,
final_local_scores=0.5343282, pem=-0.57246673,
similarity=23.148296]
24: 64646 Toronto
[pem_pos=0, final_scores=0.52745646,
final_local_scores=0.6391832, pem=-0.12291032,
similarity=23.83262]
25: 350364 Paul_Tracy
[pem_pos=0, final_scores=0.6094632,
final_local_scores=0.58621806, pem=0.0, similarity=21.58363]
26: 241678 Emerson_Fittipaldi
[pem_pos=0, final_scores=0.51140875,
final_local_scores=0.4556101, pem=0.0, similarity=17.686028]
27: 228615 Nigel_Mansell
[pem_pos=0, final_scores=0.58665395,
final_local_scores=0.55696696, pem=0.0, similarity=20.710718]
28: 18859 Michigan
[pem_pos=0, final_scores=0.1434413,
final_local_scores=0.23787597, pem=-0.24788691,
similarity=12.536454]
29: 241678 Emerson_Fittipaldi
[pem_pos=0, final_scores=0.51292175,
final_local_scores=0.45726287, pem=0.0, similarity=17.73535]
30: 350364 Paul_Tracy
[pem_pos=0, final_scores=0.5664372,
final_local_scores=0.5392186, pem=0.0, similarity=20.181072]
31: 350364 Paul_Tracy
[pem_pos=0, final_scores=0.64965034,
final_local_scores=0.6301164, pem=0.0, similarity=22.893642]
32: 1776945 Road_America
[pem_pos=0, final_scores=0.29358947,
final_local_scores=0.2725003, pem=0.0, similarity=12.221665]
33: 241678 Emerson_Fittipaldi
[pem_pos=0, final_scores=0.4992953,
final_local_scores=0.44237804, pem=0.0, similarity=17.291157]
34: 228615 Nigel_Mansell
[pem_pos=0, final_scores=0.65031534,
final_local_scores=0.6278943, pem=0.0, similarity=22.827332]
35: 21702 Nazareth
[pem_pos=0, final_scores=0.14593081,
final_local_scores=0.26970312, pem=-0.38387662,
similarity=14.225773]

36: 350364 Paul_Tracy
 [pem_pos=0, final_scores=0.66791564,
 final_local_scores=0.65006846, pem=0.0, similarity=23.489052]
 37: 676326 Mazda_Raceway_Laguna_Seca
 [pem_pos=0, final_scores=0.22869597,
 final_local_scores=0.1956057, pem=-0.039963335,
 similarity=10.144304]
 38: 30446232 IndyCar
 [pem_pos=0, final_scores=0.16864933,
 final_local_scores=0.11490241, pem=-0.25514126,
 similarity=8.906128]
 39: 27747173 Penske_PC-23
 [pem_pos=0, final_scores=0.08789872,
 final_local_scores=0.07758181, pem=0.0, similarity=6.4049087]
 40: 225290 Champ_Car
 [pem_pos=0, final_scores=0.26465803,
 final_local_scores=0.20577532, pem=-0.031101579,
 similarity=10.399594]
 41: 159018 Indianapolis_500
 [pem_pos=0, final_scores=0.24653871,
 final_local_scores=0.2012968, pem=-0.024602843,
 similarity=10.230605]

Sample: b'Progressive_Democratic_Party'

-LRB- Tunisia -RRB- The Progressive Democratic Party -
 LRB- , ; -RRB- , also referred to by its acronym PDP , was a 1
 secular 2 liberal 3 political party in [1 1 Tunisia 1] . It
 was founded under the name of Progressive Socialist Rally in
 1983 , gained legal recognition in 1988 and was renamed [1
 Progressive Democratic Party 1] in 2001 . Under the rule of [2
 2 Ben Ali 2] it was a legal opposition party , but subjected
 to political repression . After the [2 Tunisian 2] revolution
 it was one of the major left-leaning secular political
 forces . It was led by [3 3 Ahmed Najib Chebbi 3] and [4 4 Ma
 ya Jribi 4] . On 9 April 2012 , it merged into the [3 4 Repub
 lican Party 3] . In its beginnings , the Progressive
 Socialist Rally -LRB- now PDP -RRB- gathered a broad range of
 currents from [5 5 Marxists 5] and pro-democracy activists to
 5 progressive Muslims . During the [4 Ben Ali 4] rule , [5
 Najib Chebbi 5] and the [6 PDP 6] were harassed by the police
 for years , and verbally attacked by state-run media . After
 unsuccessfully participating in elections from 1989 to 1999 ,
 the party decided to boycott the elections of 6 2004 and 7 2009
 . Therefore it was unrepresented in the [7 8 Tunisian 7] parl
 iament . After the [8 Ben Ali 8] administration announced to
 force the party to move their head quarters from [9 Tunis
 9] , its leaders [10 Najib Chebbi 10] and [11 Maya Jribi 11]
 engaged in a 20-days hunger strike , which earned the party
 attention and prompted the administration to revoke their
 decision . Following the , shootings outside [12 PDP 12] 's
 headquarter were reported on 16 January 2011 . The next day ,
 on 17 January , party leader [13 Najib Chebbi 13] , was named

in the interim government . Ahead of the 9 Constituent Assembly election , the [14 PDP 14] evolved into a main exponent of the centre-left secular camp and rival of the Islamist [6 6 Ennahda Movement 6] . In the run-up to the elections , the Progressive [15 Democrats 15] have received quite an amount of financial support which enabled the party for a lavish campaign . Critics claim that an important part of the funding came from businesspeople close to the old [16 Ben Ali 16] power elite . In the 10 election for a constituent assembly , the [17 PDP 17] won 3.9 % of the popular vote and 16 of 217 seats in the [18 11 National Constituent Assembly 18] , putting it at the fifth place . As the party had categorically ruled out any collaboration with the victorious Islamist [7 7 Ennahda Movement 7] , the Progressive Democrats went into opposition and belong to the outspoken critics of the governing coalition of the [19 Islamists 19] with the secular 12 CPR and parties . After the electoral defeat , the [20 PDP 20] launched talks with other secularist and liberal parties to form a `` big party of the centre '' . The merger was realised on the [21 PDP 21] 's fifth congress on 9 April 2012 . The new party is called the [22 13 Republican Party 22] and comprises , in addition to the [23 PDP 23] , the 14 Afek Tounes

1: 30188 Tunisia

2: 390745 Zine_El_Abidine_Ben_Ali

3: 30411638 Ahmed_Najib_Chebbi

4: 33488944 Maya_Jribi

5: 1904053 Marxism

6: 30876073 Ennahda_Movement

7: 30876073 Ennahda_Movement

1: 27113 Secularism

[winner: 18675256 Hiloni, pem_pos=3,
final_scores=-1.0092825, final_local_scores=-0.9900047,
pem=-1.097609, similarity=-19.485033]

gt: 27113 Secularism, pem_pos=4,
final_scores=-1.4663278, final_local_scores=-1.600344,
pem=-1.7701793, similarity=-34.041233]

2: 19280734 Liberalism

[winner: 864886 Liberalism_by_country, pem_pos=1,
final_scores=-1.2441202, final_local_scores=-1.3071198,
pem=-0.4054651, similarity=-32.71237]

RECALL MISS: 19280734 Liberalism]

3: 23996 Political_party

[winner: 864886 Liberalism_by_country, pem_pos=1,
final_scores=-1.2441202, final_local_scores=-1.3071198,
pem=-0.4054651, similarity=-32.71237]

RECALL MISS: 23996 Political_party]

4: 35495084 Republican_Party_(Tunisia)

[winner: 32070 Republican_Party_(United_States),
pem_pos=0, final_scores=0.20549671,
final_local_scores=0.2169388, pem=-0.16706309,
similarity=11.472115]

RECALL MISS: 35495084 Republican_Party_(Tunisia)]
 5: 325137 Liberal_movements_within_Islam
 [winner: 19541 Muslim, pem_pos=0,
 final_scores=-0.004858345, final_local_scores=0.027496696,
 pem=-0.17009869, similarity=5.8352923]
 RECALL MISS: 325137 Liberal_movements_within_Islam]
 [winner: 2163978 Progressive_country, pem_pos=12,
 final_scores=-1.6046321, final_local_scores=-1.6086457,
 pem=-2.770291, similarity=-28.85021]
 RECALL MISS: 325137 Liberal_movements_within_Islam]
 6: 30156581 Tunisian_general_election,_2004
 [winner: 35524 2004, pem_pos=0, final_scores=-0.6011754,
 final_local_scores=-0.43721527, pem=-0.4915334,
 similarity=-6.284628]
 RECALL MISS: 30156581 Tunisian_general_election,_2004]
 7: 20464676 Tunisian_general_election,_2009
 [winner: 35983 2009, pem_pos=0, final_scores=-0.8422534,
 final_local_scores=-0.7387587, pem=-0.9296576,
 similarity=-12.900698]
 RECALL MISS: 20464676 Tunisian_general_election,_2009]
 8: 17501338 Parliament_of_Tunisia
 [winner: 30188 Tunisia, pem_pos=2,
 final_scores=0.5922928, final_local_scores=0.6441441,
 pem=-0.13068858, similarity=24.022963]
 RECALL MISS: 17501338 Parliament_of_Tunisia]
 [winner: 24406 Parliament, pem_pos=1,
 final_scores=-1.6601892, final_local_scores=-1.8053584,
 pem=-1.8263724, similarity=-39.853687]
 RECALL MISS: 17501338 Parliament_of_Tunisia]
 9: 31071064 Tunisian_Constituent_Assembly_election,_2011
 [winner: 266437 Constituent_assembly, pem_pos=0,
 final_scores=-0.11315565, final_local_scores=-0.12524709,
 pem=-0.8576672, similarity=5.016217]
 RECALL MISS: 31071064
 Tunisian_Constituent_Assembly_election,_2011]
 [winner: 1442451 Elections_in_Rwanda, pem_pos=2,
 final_scores=-1.1534563, final_local_scores=-1.1343874,
 pem=-3.50929, similarity=-10.678604]
 RECALL MISS: 31071064
 Tunisian_Constituent_Assembly_election,_2011]
 10: 31071064 Tunisian_Constituent_Assembly_election,_2011
 [winner: 2255685 First_German_Bundestag, pem_pos=0,
 final_scores=-0.6330969, final_local_scores=-0.43124583,
 pem=0.0, similarity=-8.779523]
 RECALL MISS: 31071064
 Tunisian_Constituent_Assembly_election,_2011]
 [winner: 13537419 Ecuadorian_Constituent_Assembly,
 pem_pos=0, final_scores=-1.144278,
 final_local_scores=-1.1896956, pem=0.0, similarity=-31.41318]
 RECALL MISS: 31071064
 Tunisian_Constituent_Assembly_election,_2011]
 11: 33526808 Constituent_Assembly_of_Tunisia

[winner: 192078 National_Constituent_Assembly, pem_pos=0, final_scores=0.056619868, final_local_scores=0.1121289, pem=0.0, similarity=7.435863]

gt: 33526808 Constituent_Assembly_of_Tunisia, pem_pos=6, final_scores=-0.8504706, final_local_scores=-0.9408123, pem=-6.0288806, similarity=8.799987]

12: 30510758 Congress_for_the_Republic

[winner: 66392 Cardiopulmonary_resuscitation, pem_pos=0, final_scores=-0.43708104, final_local_scores=-0.3615766, pem=-0.11111647, similarity=-6.096185]

RECALL MISS: 30510758 Congress_for_the_Republic]

13: 35495084 Republican_Party_(Tunisia)

[winner: 32070 Republican_Party_(United_States), pem_pos=0, final_scores=0.10682337, final_local_scores=0.10915315, pem=-0.16706309, similarity=8.255575]

RECALL MISS: 35495084 Republican_Party_(Tunisia)]

14: 33533363 Afek_Tounes

[winner: 33533363 Afek_Tounes, pem_pos=0, final_scores=0.0045350045, final_local_scores=0.06838532, pem=0.0, similarity=6.1304674]

gt: 33533363 Afek_Tounes, pem_pos=0, final_scores=0.0045350045, final_local_scores=0.06838532, pem=0.0, similarity=6.1304674]

1: 30188 Tunisia

[pem_pos=0, final_scores=0.39133555, final_local_scores=0.42108864, pem=-0.39291763, similarity=18.792585]

2: 390745 Zine_El_Abidine_Ben_Ali

[pem_pos=0, final_scores=0.32936704, final_local_scores=0.30869406, pem=0.0, similarity=13.30176]

3: 30411638 Ahmed_Najib_Chebbi

[pem_pos=0, final_scores=0.17900725, final_local_scores=0.19922507, pem=0.0, similarity=10.0349865]

4: 33488944 Maya_Jribi

[pem_pos=0, final_scores=0.36372447, final_local_scores=0.40866137, pem=0.0, similarity=16.284983]

5: 1904053 Marxism

[pem_pos=0, final_scores=0.11151181, final_local_scores=0.14507854, pem=0.0, similarity=8.419146]

6: 30876073 Ennahda_Movement

[pem_pos=0, final_scores=0.38169926, final_local_scores=0.36288518, pem=0.0, similarity=14.91893]

7: 30876073 Ennahda_Movement

[pem_pos=0, final_scores=0.23839442, final_local_scores=0.20634648, pem=0.0, similarity=10.247503]

1: 403170 Progressive_Democratic_Party_(Tunisia)

[pem_pos=1, final_scores=0.11926068, final_local_scores=0.13310194, pem=-0.82611233, similarity=12.554267]

2: 30188 Tunisia

[pem_pos=0, final_scores=0.5833611,

final_local_scores=0.6308474, pem=-0.13068858,
similarity=23.626162]
3: 32070 Republican_Party_(United_States)
[pem_pos=0, final_scores=0.20549671,
final_local_scores=0.2169388, pem=-0.16706309,
similarity=11.472115]
4: 390745 Zine_El_Abidine_Ben_Ali
[pem_pos=0, final_scores=0.37361258,
final_local_scores=0.35702556, pem=0.0, similarity=14.744068]
5: 30411638 Ahmed_Najib_Chebbi
[pem_pos=0, final_scores=0.2544114,
final_local_scores=0.28159267, pem=0.0, similarity=12.493]
6: 427291 People's_Democratic_Party_(Nigeria)
[pem_pos=1, final_scores=0.086910084,
final_local_scores=0.045168027, pem=-0.8260722,
similarity=9.929926]
7: 30188 Tunisia
[pem_pos=2, final_scores=0.5922928,
final_local_scores=0.6441441, pem=-0.13068858,
similarity=24.022963]
8: 390745 Zine_El_Abidine_Ben_Ali
[pem_pos=0, final_scores=0.38055283,
final_local_scores=0.3646068, pem=0.0, similarity=14.970306]
9: 57659 Tunis
[pem_pos=0, final_scores=0.26990753,
final_local_scores=0.29569176, pem=-0.032719024,
similarity=13.091676]
10: 30411638 Ahmed_Najib_Chebbi
[pem_pos=0, final_scores=0.25617486,
final_local_scores=0.2835189, pem=0.0, similarity=12.550483]
11: 33488944 Maya_Jribi
[pem_pos=0, final_scores=0.45427436,
final_local_scores=0.50757337, pem=0.0, similarity=19.236715]
12: 427291 People's_Democratic_Party_(Nigeria)
[pem_pos=1, final_scores=0.12664817,
final_local_scores=0.088575825, pem=-0.8260722,
similarity=11.225302]
13: 30411638 Ahmed_Najib_Chebbi
[pem_pos=0, final_scores=0.2550224,
final_local_scores=0.28226006, pem=0.0, similarity=12.512917]
14: 427291 People's_Democratic_Party_(Nigeria)
[pem_pos=1, final_scores=0.13880898,
final_local_scores=0.10548462, pem=-0.8260722,
similarity=11.729894]
15: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.12890832,
final_local_scores=0.1333203, pem=-0.18007085,
similarity=9.047509]
16: 390745 Zine_El_Abidine_Ben_Ali
[pem_pos=0, final_scores=0.31165034,
final_local_scores=0.28934127, pem=0.0, similarity=12.724234]
17: 427291 People's_Democratic_Party_(Nigeria)

[pem_pos=1, final_scores=0.13496305,
 final_local_scores=0.101283535, pem=-0.8260722,
 similarity=11.604525]
 18: 192078 National_Constituent_Assembly
 [pem_pos=0, final_scores=0.056619868,
 final_local_scores=0.1121289, pem=0.0, similarity=7.435863]
 19: 15012 Islamism
 [pem_pos=0, final_scores=0.055480257,
 final_local_scores=0.037741646, pem=0.0, similarity=5.215999]
 20: 427291 People's_Democratic_Party_(Nigeria)
 [pem_pos=1, final_scores=0.16933416,
 final_local_scores=0.13882875, pem=-0.8260722,
 similarity=12.72495]
 21: 427291 People's_Democratic_Party_(Nigeria)
 [pem_pos=1, final_scores=0.17561708,
 final_local_scores=0.14206696, pem=-0.8260722,
 similarity=12.821583]
 22: 32070 Republican_Party_(United_States)
 [pem_pos=0, final_scores=0.10682337,
 final_local_scores=0.10915315, pem=-0.16706309,
 similarity=8.255575]
 23: 427291 People's_Democratic_Party_(Nigeria)
 [pem_pos=1, final_scores=0.09738679,
 final_local_scores=0.05661224, pem=-0.8260722,
 similarity=10.271444]

Sample: b'Roger_Clitheroe'

Roger Ian Clitheroe -LRB- born 18 November 1966 -RRB- is a
 former [1 1 English 1] 2 cricket er . Clitheroe was a right-
 handed who fielded occasionally as a 3 wicket-keeper . He was
 born in 4 Radcliffe , [2 5 Lancashire 2] . Clitheroe made his
 6 first-class debut for [1 1 Cambridge University 1] against [2 2 Northamptonshire 2] in 1987 . He played 12 further first-
 class matches for [3 Cambridge University 3] , the last
 coming against [3 3 Oxford University 3] in 1991 . In his 13
 first-class matches for the University , he scored 291 runs
 at a 7 batting average of 15.31 , with a high score of 36 .
 Behind the he took 8 and 2 s . He later made his debut for [4 4 Wales Minor Counties 4] in the 1996 against [5 5 Devon 5] .
 He played 8 Minor counties cricket for [4 Wales Minor
 Counties 4] from 1996 to 2002 , which included 31 [5 Minor
 Counties Championship 5] matches and 12 [6 6 MCCA Knockout Tr
 ophy 6] matches . In 1998 , he made his [7 7 List A 7] debut
 for [6 Wales Minor Counties 6] against [8 8 Nottinghamshire 8]
 in the [7 9 NatWest Trophy 7] . He made 2 further [8 List A 8]
 appearances for the county , against [9 9 Norfolk 9] and [10 1
 0 Leicestershire 10] in the 10 2001 [9 Cheltenham & Gloucester
 Trophy 9] . In his 3 [10 List A 10] matches , he scored 37
 runs at a 11 batting average

1: 1622256 Cambridge_University_Cricket_Club
 2: 530308 Northamptonshire_County_Cricket_Club
 3: 1622255 Oxford_University_Cricket_Club

4: 3712706 Wales_Minor_Counties_Cricket_Club
 5: 3691070 Devon_County_Cricket_Club
 6: 7543416 MCCA_Knockout_Trophy
 7: 683459 List_A_cricket
 8: 667341 Nottinghamshire_County_Cricket_Club
 9: 3690942 Norfolk_County_Cricket_Club
 10: 1622318 Leicestershire_County_Cricket_Club
 1: 18803164 English_people
 [winner: 9316 England, pem_pos=1, final_scores=0.4881658,
 final_local_scores=0.52527755, pem=-1.3527626,
 similarity=27.121576|
 gt: 18803164 English_people, pem_pos=2,
 final_scores=-0.41395587, final_local_scores=-0.4526853,
 pem=-2.2084396, similarity=2.5905182]
 2: 25675557 Cricket
 [winner: 330091 Rhaphidophoridae, pem_pos=10,
 final_scores=-1.4712476, final_local_scores=-1.4234889,
 pem=-2.56456, similarity=-24.443562|
 gt: 25675557 Cricket, pem_pos=0,
 final_scores=-1.7479053, final_local_scores=-1.9508126,
 pem=-0.07532314, similarity=-53.716805]
 3: 463313 Wicket-keeper
 [winner: 463313 Wicket-keeper, pem_pos=0,
 final_scores=-1.5377159, final_local_scores=-1.733309,
 pem=-0.004312502, similarity=-47.612236|
 gt: 463313 Wicket-keeper, pem_pos=0,
 final_scores=-1.5377159, final_local_scores=-1.733309,
 pem=-0.004312502, similarity=-47.612236]
 4: 1737474 Radcliffe, Greater_Manchester
 [winner: 1737474 Radcliffe, Greater_Manchester,
 pem_pos=0, final_scores=-0.07363877,
 final_local_scores=-0.047165826, pem=-0.9188902,
 similarity=7.679258|
 gt: 1737474 Radcliffe, Greater_Manchester, pem_pos=0,
 final_scores=-0.07363877, final_local_scores=-0.047165826,
 pem=-0.9188902, similarity=7.679258]
 5: 50435 Lancashire
 [winner: 749919 Lancashire_County_Cricket_Club,
 pem_pos=1, final_scores=0.36703324,
 final_local_scores=0.2988457, pem=-1.7295884,
 similarity=22.413635|
 gt: 50435 Lancashire, pem_pos=0,
 final_scores=-0.11795016, final_local_scores=-0.14529756,
 pem=-0.20611861, similarity=0.8746489]
 6: 11040 First-class_cricket
 [winner: 662844 First-class_citizen, pem_pos=3,
 final_scores=-1.285476, final_local_scores=-1.218816,
 pem=-1.5935546, similarity=-23.616188|
 gt: 11040 First-class_cricket, pem_pos=0,
 final_scores=-1.3692462, final_local_scores=-1.579083,
 pem=-0.053819276, similarity=-42.740593]
 7: 18933273 Batting_average

[winner: 15652278 Batting_(material), pem_pos=2,
 final_scores=-1.0199902, final_local_scores=-0.9110166,
 pem=-1.1616125, similarity=-16.77981]
 gt: 18933273 Batting_average, pem_pos=0,
 final_scores=-1.2916608, final_local_scores=-1.4333694,
 pem=-0.29142123, similarity=-37.100094]
 [winner: 875871 Bowling_average, pem_pos=1,
 final_scores=-1.3609502, final_local_scores=-1.527673,
 pem=-0.8657023, similarity=-36.79127]
 gt: 18933273 Batting_average, pem_pos=0,
 final_scores=-1.4581554, final_local_scores=-1.6152394,
 pem=-0.5012068, similarity=-41.386612]
 8: 770909 Minor_counties_of_English_and_Welsh_cricket
 [winner: 14465683 William_T_Minor, pem_pos=3,
 final_scores=-0.47606742, final_local_scores=-0.4215444,
 pem=-2.3714392, similarity=4.406243]
 RECALL MISS: 770909
 Minor_counties_of_English_and_Welsh_cricket]
 [winner: 4569028 Counties_of_Uganda, pem_pos=18,
 final_scores=-1.3354352, final_local_scores=-1.2803423,
 pem=-3.1761155, similarity=-16.846046]
 RECALL MISS: 770909
 Minor_counties_of_English_and_Welsh_cricket]
 9: 27683969 1998_NatWest_Trophy
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.33650488, final_local_scores=0.2934565,
 pem=0.0, similarity=12.84704]
 gt: 27683969 1998_NatWest_Trophy, pem_pos=5,
 final_scores=-0.5283419, final_local_scores=-0.58920896,
 pem=-3.8250117, similarity=7.3075495]
 10: 27734795 2001_Cheltenham_&_Gloucester_Trophy
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.26960802, final_local_scores=0.22038186,
 pem=0.0, similarity=10.6663475]
 gt: 27734795 2001_Cheltenham_&_Gloucester_Trophy,
 pem_pos=1, final_scores=-0.103383094,
 final_local_scores=-0.15444073, pem=-1.537979,
 similarity=7.8446627]
 [winner: 34551 2001, pem_pos=0, final_scores=-0.62796724,
 final_local_scores=-0.4026872, pem=-0.5613713,
 similarity=-4.8744516]
 RECALL MISS: 27734795
 2001_Cheltenham_&_Gloucester_Trophy]
 11: 18933273 Batting_average
 [winner: 15652278 Batting_(material), pem_pos=2,
 final_scores=-0.9053006, final_local_scores=-0.7857357,
 pem=-1.1616125, similarity=-13.04118]
 gt: 18933273 Batting_average, pem_pos=0,
 final_scores=-1.1852467, final_local_scores=-1.3171282,
 pem=-0.29142123, similarity=-33.63122]
 [winner: 875871 Bowling_average, pem_pos=1,
 final_scores=-1.1731633, final_local_scores=-1.3225443,

pem=-0.8657023, similarity=-30.669823|
gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.293788, final_local_scores=-1.4356928,
pem=-0.5012068, similarity=-36.028584]
1: 1622256 Cambridge_University_Cricket_Club
[pem_pos=1, final_scores=0.25631016,
final_local_scores=0.19603014, pem=-1.3585418,
similarity=17.327606]
2: 530308 Northamptonshire_County_Cricket_Club
[pem_pos=1, final_scores=0.42543262,
final_local_scores=0.36522138, pem=-1.2069724,
similarity=21.552351]
3: 1622255 Oxford_University_Cricket_Club
[pem_pos=1, final_scores=0.20430596,
final_local_scores=0.14214194, pem=-1.5137167,
similarity=16.56334]
4: 3712706 Wales_Minor_Counties_Cricket_Club
[pem_pos=0, final_scores=0.3755539,
final_local_scores=0.33196476, pem=0.0, similarity=13.996203]
5: 3691070 Devon_County_Cricket_Club
[pem_pos=1, final_scores=0.25747997,
final_local_scores=0.19750863, pem=-3.0304468,
similarity=26.463806]
6: 7543416 MCCA_Knockout_Trophy
[pem_pos=0, final_scores=0.29856443,
final_local_scores=0.2520975, pem=0.0, similarity=11.612804]
7: 683459 List_A_cricket
[pem_pos=0, final_scores=0.21694012,
final_local_scores=0.15266255, pem=-0.020267159,
similarity=8.755684]
8: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6270288,
final_local_scores=0.5846247, pem=-1.1067253,
similarity=27.554625]
9: 3690942 Norfolk_County_Cricket_Club
[pem_pos=6, final_scores=0.3082785,
final_local_scores=0.25233614, pem=-3.4736512,
similarity=30.510181]
10: 1622318 Leicestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.77848583,
final_local_scores=0.7556613, pem=-1.0465667,
similarity=32.331547]
1: 9316 England
[pem_pos=1, final_scores=0.4881658,
final_local_scores=0.52527755, pem=-1.3527626,
similarity=27.121576]
2: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.36703324,
final_local_scores=0.2988457, pem=-1.7295884,
similarity=22.413635]
3: 1622256 Cambridge_University_Cricket_Club
[pem_pos=1, final_scores=0.20215048,

```

final_local_scores=0.13686895, pem=-1.3585418,
similarity=15.562118]
4: 3712706 Wales_Minor_Counties_Cricket_Club
    [pem_pos=0, final_scores=0.30727118,
final_local_scores=0.25737625, pem=0.0, similarity=11.770333]
5: 7562952 Minor_Counties_Cricket_Championship
    [pem_pos=0, final_scores=0.36976045,
final_local_scores=0.31499904, pem=0.0, similarity=13.489913]
6: 3712706 Wales_Minor_Counties_Cricket_Club
    [pem_pos=0, final_scores=0.4014455,
final_local_scores=0.36024737, pem=0.0, similarity=14.840213]
7: 1899702 Friends_Provident_Trophy
    [pem_pos=0, final_scores=0.33650488,
final_local_scores=0.2934565, pem=0.0, similarity=12.84704]
8: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.3060282,
final_local_scores=0.24997774, pem=-0.020267159,
similarity=11.659763]
9: 1899702 Friends_Provident_Trophy
    [pem_pos=0, final_scores=0.26960802,
final_local_scores=0.22038186, pem=0.0, similarity=10.6663475]
10: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.14300527,
final_local_scores=0.071899965, pem=-0.020267159,
similarity=6.345567]

```

Sample: b'South_Gippsland_Highway'

The [1 South Gippsland Highway 1] is a partially divided highway in [1 1 Victoria 1] , [2 2 Australia 2] which connects the city of [3 3 Melbourne 3] with the [2 South 1 Gippsland 2] region of [3 Victoria 3] , ending in the town of [4 4 Sale 4] . The highway begins with no route number , until the junction of Greens Road where it gains the Metropolitan Route 12 to Pound Road , before no route number is signed again until the [5 5 South Gippsland Freeway 5] / interchange where it becomes the M420 . The M420 continues through [6 6 Cranbourne 6] and until the [4 2 Bass Highway 4] turnoff , which is then designated A440 onwards to [5 Sale 5] . From the [6 Bass Highway 6] junction , the highway is undivided . The [7 South Gippsland Highway 7] is the gateway from [8 Melbourne 8] to many attractions including [9 3 Wilsons Promontory 9] and [7 7 Phillip Island 7] as well as being an important road for farmers in [10 Gippsland 10] . The highway commences in [8 8 Dandenong 8] , branching from the [9 9 Princes Highway 9] , and heads in a general south-east direction at first . It is mostly a 4 dual carriageway with two lanes in either direction , until the [11 5 Bass Highway 11] turnoff to , after which it is entirely an undivided highway . It passes through the 6 Cranbourne area , and then through or close to the [12 Gippsland 12] towns of [10 10 Too radin 10] , , 7 Lang Lang , , , 8 Foster , [13 9 Welshpool 1 3] and [11 11 Yarram 11] , before heading north to rejoin the

[14 Princes Highway 14] at 10 Sale

1: 4689460 Victoria_(Australia)
2: 4689264 Australia
3: 17306237 Melbourne
4: 80260 Sale,_Victoria
5: 1043548 South_Gippsland_Freeway
6: 611192 Cranbourne,_Victoria
7: 312060 Phillip_Island
8: 611262 Dandenong,_Victoria
9: 496486 Princes_Highway
10: 611248 Tooradin,_Victoria
11: 1567554 Yarram,_Victoria
1: 312055 Gippsland

[winner: 6153280 South_Gippsland, pem_pos=0,
final_scores=0.26058388, final_local_scores=0.2682527,
pem=-0.34897432, similarity=13.992685]

gt: 312055 Gippsland, pem_pos=4,
final_scores=0.24717514, final_local_scores=0.19702461,
pem=-1.5670587, similarity=18.49123]

2: 743468 Bass_Highway_(Victoria)

[winner: 786278 Bass_Highway,_Tasmania, pem_pos=0,
final_scores=0.32307315, final_local_scores=0.34596744,
pem=0.0, similarity=14.414071]

gt: 743468 Bass_Highway_(Victoria), pem_pos=1,
final_scores=0.045541212, final_local_scores=0.065877676,
pem=-0.37309983, similarity=8.084609]

3: 51150 Wilsons_Promontory_National_Park

[winner: 1089237 Wilsons_Promontory, pem_pos=0,
final_scores=0.3021136, final_local_scores=0.32785985,
pem=0.0, similarity=13.873705]

gt: 51150 Wilsons_Promontory_National_Park, pem_pos=1,
final_scores=0.08599652, final_local_scores=0.1335729,
pem=-0.5131485, similarity=10.866376]

4: 692197 Dual_carriageway

[winner: 20395887 Dual_narrative, pem_pos=0,
final_scores=-0.8240364, final_local_scores=-0.72226936,
pem=0.0, similarity=-17.464245]

RECALL MISS: 692197 Dual_carriageway]

[winner: 692197 Dual_carriageway, pem_pos=0,
final_scores=-1.0922118, final_local_scores=-1.189362,
pem=0.0, similarity=-31.403223]

gt: 692197 Dual_carriageway, pem_pos=0,
final_scores=-1.0922118, final_local_scores=-1.189362,
pem=0.0, similarity=-31.403223]

5: 743468 Bass_Highway_(Victoria)

[winner: 786278 Bass_Highway,_Tasmania, pem_pos=0,
final_scores=0.34902734, final_local_scores=0.37431842,
pem=0.0, similarity=15.260121]

gt: 743468 Bass_Highway_(Victoria), pem_pos=1,
final_scores=0.10236867, final_local_scores=0.12795302,
pem=-0.37309983, similarity=9.937061]

6: 611192 Cranbourne,_Victoria

[winner: 611192 Cranbourne, Victoria, pem_pos=2,
final_scores=0.0026590377, final_local_scores=-0.039080545,
pem=-0.975827, similarity=8.23017]
gt: 611192 Cranbourne, Victoria, pem_pos=2,
final_scores=0.0026590377, final_local_scores=-0.039080545,
pem=-0.975827, similarity=8.23017]
7: 2843413 Lang_Lang, Victoria
[winner: 2843413 Lang_Lang, Victoria, pem_pos=2,
final_scores=-0.0830185, final_local_scores=-0.05864975,
pem=-1.4594618, similarity=10.276267]
gt: 2843413 Lang_Lang, Victoria, pem_pos=2,
final_scores=-0.0830185, final_local_scores=-0.05864975,
pem=-1.4594618, similarity=10.276267]
8: 2645961 Foster, Victoria
[winner: 134326 Foster, Rhode_Island, pem_pos=0,
final_scores=-0.28863743, final_local_scores=-0.26680386,
pem=-2.4580972, similarity=9.495268]
gt: 2645961 Foster, Victoria, pem_pos=11,
final_scores=-0.45937055, final_local_scores=-0.5028557,
pem=-3.8940265, similarity=10.259817]
9: 28617201 Welshpool, Victoria
[winner: 480604 Welshpool, pem_pos=0,
final_scores=0.20826195, final_local_scores=0.25210646,
pem=0.0, similarity=11.613072]
gt: 28617201 Welshpool, Victoria, pem_pos=8,
final_scores=-0.5437022, final_local_scores=-0.5654897,
pem=-4.7011642, similarity=12.780034]
10: 80260 Sale, Victoria
[winner: 80260 Sale, Victoria, pem_pos=0,
final_scores=0.022096261, final_local_scores=-0.016854525,
pem=-1.232199, similarity=10.28763]
gt: 80260 Sale, Victoria, pem_pos=0,
final_scores=0.022096261, final_local_scores=-0.016854525,
pem=-1.232199, similarity=10.28763]
1: 4689460 Victoria_(Australia)
[pem_pos=0, final_scores=0.5275106,
final_local_scores=0.52051497, pem=-0.6703371,
similarity=23.268314]
2: 4689264 Australia
[pem_pos=0, final_scores=0.6046802,
final_local_scores=0.6406239, pem=-0.17168477,
similarity=24.140858]
3: 17306237 Melbourne
[pem_pos=0, final_scores=0.41806054,
final_local_scores=0.4129079, pem=-0.16841523,
similarity=17.327578]
4: 80260 Sale, Victoria
[pem_pos=0, final_scores=0.1929258,
final_local_scores=0.17645279, pem=-1.232199,
similarity=16.056305]
5: 1043548 South_Gippsland_Freeway
[pem_pos=0, final_scores=0.13062684,

final_local_scores=0.16092393, pem=0.0, similarity=8.892004]
6: 611192 Cranbourne, Victoria
[pem_pos=2, final_scores=0.10702644,
final_local_scores=0.081807, pem=-0.975827,
similarity=11.837696]
7: 312060 Phillip_Island
[pem_pos=0, final_scores=0.15039335,
final_local_scores=0.16804612, pem=-0.4078959,
similarity=11.322745]
8: 611262 Dandenong, Victoria
[pem_pos=0, final_scores=0.103538856,
final_local_scores=0.07297595, pem=-0.37683475,
similarity=8.316747]
9: 496486 Princes_Highway
[pem_pos=0, final_scores=0.4718371,
final_local_scores=0.44817376, pem=0.0, similarity=17.464111]
10: 611248 Tooradin, Victoria
[pem_pos=0, final_scores=0.113810286,
final_local_scores=0.17467308, pem=0.0, similarity=9.302306]
11: 1567554 Yarram, Victoria
[pem_pos=0, final_scores=0.101848796,
final_local_scores=0.10193567, pem=0.0, similarity=7.1316767]
1: 743479 South_Gippsland_Highway
[pem_pos=0, final_scores=0.28004164,
final_local_scores=0.29399568, pem=0.0, similarity=12.863131]
2: 6153280 South_Gippsland
[pem_pos=0, final_scores=0.26058388,
final_local_scores=0.2682527, pem=-0.34897432,
similarity=13.992685]
3: 4689460 Victoria_(Australia)
[pem_pos=0, final_scores=0.49942178,
final_local_scores=0.48983228, pem=-0.6703371,
similarity=22.352682]
4: 786278 Bass_Highway, Tasmania
[pem_pos=0, final_scores=0.32307315,
final_local_scores=0.34596744, pem=0.0, similarity=14.414071]
5: 80260 Sale, Victoria
[pem_pos=0, final_scores=0.21633182,
final_local_scores=0.20202032, pem=-1.232199,
similarity=16.819292]
6: 786278 Bass_Highway, Tasmania
[pem_pos=0, final_scores=0.2795629,
final_local_scores=0.29843915, pem=0.0, similarity=12.995732]
7: 743479 South_Gippsland_Highway
[pem_pos=0, final_scores=0.22017635,
final_local_scores=0.22860196, pem=0.0, similarity=10.911652]
8: 17306237 Melbourne
[pem_pos=0, final_scores=0.40450567,
final_local_scores=0.39810127, pem=-0.16841523,
similarity=16.885717]
9: 1089237 Wilsons_Promontory
[pem_pos=0, final_scores=0.3021136,

```

final_local_scores=0.32785985, pem=0.0, similarity=13.873705]
10: 312055 Gippsland
    [pem_pos=0, final_scores=0.2933051,
final_local_scores=0.24701819, pem=-0.1924336,
similarity=12.507712]
11: 786278 Bass_Highway,_Tasmania
    [pem_pos=0, final_scores=0.34902734,
final_local_scores=0.37431842, pem=0.0, similarity=15.260121]
12: 312055 Gippsland
    [pem_pos=0, final_scores=0.34765464,
final_local_scores=0.30638683, pem=-0.1924336,
similarity=14.279391]
13: 480604 Welshpool
    [pem_pos=0, final_scores=0.20826195,
final_local_scores=0.25210646, pem=0.0, similarity=11.613072]
14: 496486 Princes_Highway
    [pem_pos=0, final_scores=0.4075802,
final_local_scores=0.37798285, pem=0.0, similarity=15.369475]

```

Sample: b'Wesley_Brown_Field_House'

The [1 Wesley Brown Field House 1] is a 1 sports 2 arena at the in [2 3 Annapolis 2] , [3 Maryland 3] . It is located between the 7th Wing of [1 1 Bancroft Hall 1] and Santee Basin . The facility houses 4 physical education , , club sports , and 5 personal-fitness programs and 6 equipment . It is home to the 7 Midshipmen cross country and track and field teams , the sprint football team , the women 's lacrosse team , and sixteen club sports . It also serves as the space for the 8 football and women 's volleyball teams . There is also a centralized 9 sports-medicine facility . The building has a total room area of 5,800 sq. ft. , eight , and 300 10 lockers . The field house has a , , retractable Magic Carpet 11 AstroTurf . When the field is retracted , students can then use the 200-meter AstroTurf with a Mondo track surface and 12 hydraulically - controlled banked curves and three permanent 1 3 basketball court s . In four to six hours , the indoor track-and-field can be changed to an indoor football practice field , including target for placekicking practice . The 14 synthetic playing surface -LRB- with football-field yard-lines -RRB- is stored on a spool at the field house 's south end . The surface is put in place by nine winches and an 18-port air-blower that makes the 15 turf float across the field-house floor while being deployed and retracted . The 16 weight room is one of three `` strength and conditioning facilities '' at the academy . With , it serves the members of the following teams : men 's and women 's cross country , men 's and women 's soccer , men 's and women 's track , golf , sprint football , volleyball , water polo , and women 's lacrosse . The facility is named for the first [4 African American 4] to graduate from the Academy , retired Lt. Cmdr. 1 7 Wesley A. Brown , who graduated in 1949 . The was held on March 25 , 2006 . Brown wielded a shovel in the

groundbreaking . [5 Hensel Phelps Construction Company 5]
constructed the field house , which was completed in March
2008 . The U.S. 18 Naval Facilities Engineering Command -LRB-
NAVFAC -RRB- administered the contract for construction . The
building was dedicated on May 10 , 2008 . Brown participated
in the

1: 395874 Bancroft_Hall

1: 35544787 College_athletics

[winner: 1888136 Toledo_Sports_Arena, pem_pos=0,
final_scores=-1.0437471, final_local_scores=-1.025154,
pem=-0.2822249, similarity=-24.968145]

RECALL MISS: 35544787 College_athletics]

2: 180244 Arena

[winner: 1888136 Toledo_Sports_Arena, pem_pos=0,
final_scores=-1.0437471, final_local_scores=-1.025154,
pem=-0.2822249, similarity=-24.968145]

gt: 180244 Arena, pem_pos=3, final_scores=-1.322154,
final_local_scores=-1.3530085, pem=-1.8865807,
similarity=-26.027243]

3: 57839 Annapolis, Maryland

[winner: 57839 Annapolis, Maryland, pem_pos=0,
final_scores=0.32858598, final_local_scores=0.30225754,
pem=-0.22970897, similarity=14.358874]

gt: 57839 Annapolis, Maryland, pem_pos=0,
final_scores=0.32858598, final_local_scores=0.30225754,
pem=-0.22970897, similarity=14.358874]

[winner: 18858 Maryland, pem_pos=0,
final_scores=0.31155092, final_local_scores=0.30097538,
pem=-0.2529562, similarity=14.447034]

RECALL MISS: 57839 Annapolis, Maryland]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.795759, final_local_scores=-1.7771704,
pem=-4.6728287, similarity=-23.533024]

RECALL MISS: 57839 Annapolis, Maryland]

4: 217324 Physical_education

[winner: 179505 Physical_property, pem_pos=2,
final_scores=-1.4149803, final_local_scores=-1.3908792,
pem=-1.7335073, similarity=-27.989815]

RECALL MISS: 217324 Physical_education]

[winner: 18823453

Angelo_State_University_College_of_Education, pem_pos=3,
final_scores=-1.5047218, final_local_scores=-1.5190783,
pem=-3.356857, similarity=-22.9875]

RECALL MISS: 217324 Physical_education]

5: 432986 Physical_fitness

[winner: 432986 Physical_fitness, pem_pos=0,
final_scores=-1.5448745, final_local_scores=-1.6400545,
pem=0.0, similarity=-44.852787]

gt: 432986 Physical_fitness, pem_pos=0,
final_scores=-1.5448745, final_local_scores=-1.6400545,
pem=0.0, similarity=-44.852787]

6: 1802372 Sports_equipment

[winner: 24619464
Equipment_of_the_United_States_Air_Force, pem_pos=2,
final_scores=-0.7883179, final_local_scores=-0.6657356,
pem=-1.7513024, similarity=-6.25331|
RECALL MISS: 1802372 Sports_equipment]

7: 9954210 Navy_Midshipmen
[winner: 304197 Midshipman, pem_pos=1,
final_scores=-0.04483731, final_local_scores=-0.02722022,
pem=-0.5966363, similarity=6.5220084|
gt: 9954210 Navy_Midshipmen, pem_pos=0,
final_scores=-0.073011786, final_local_scores=-0.050154053,
pem=-0.5067956, similarity=5.349049]

8: 7305146 Navy_Midshipmen_football
[winner: 5730131 Nuclear_briefcase, pem_pos=0,
final_scores=-1.0150481, final_local_scores=-0.9563064,
pem=-0.044735882, similarity=-24.205097|
RECALL MISS: 7305146 Navy_Midshipmen_football]

9: 1365637 Sports_medicine
[winner: 1365637 Sports_medicine, pem_pos=0,
final_scores=-1.3313805, final_local_scores=-1.4074388,
pem=0.0, similarity=-37.911068|
gt: 1365637 Sports_medicine, pem_pos=0,
final_scores=-1.3313805, final_local_scores=-1.4074388,
pem=0.0, similarity=-37.911068]

10: 14202412 Locker_(cabinet)
[winner: 14202412 Locker_(cabinet), pem_pos=0,
final_scores=-1.158217, final_local_scores=-1.1704013,
pem=-0.4202802, similarity=-28.551847|
gt: 14202412 Locker_(cabinet), pem_pos=0,
final_scores=-1.158217, final_local_scores=-1.1704013,
pem=-0.4202802, similarity=-28.551847]

11: 66480 AstroTurf
[winner: 1156289 Artificial_turf, pem_pos=0,
final_scores=-0.8714743, final_local_scores=-0.89090353,
pem=0.0, similarity=-22.496626|
gt: 66480 AstroTurf, pem_pos=2,
final_scores=-0.96773547, final_local_scores=-1.0108986,
pem=-1.1817098, similarity=-19.651197]

12: 65424 Hydraulics
[winner: 1372353 Hydraulic_machinery, pem_pos=0,
final_scores=-1.0786155, final_local_scores=-1.0671177,
pem=0.0, similarity=-27.755209|
gt: 65424 Hydraulics, pem_pos=1,
final_scores=-1.1619098, final_local_scores=-1.1893542,
pem=-0.2859106, similarity=-29.848164]

13: 679837 Basketball_court
[winner: 679837 Basketball_court, pem_pos=0,
final_scores=-1.2220519, final_local_scores=-1.2847514,
pem=-0.00656814, similarity=-34.214115|
gt: 679837 Basketball_court, pem_pos=0,
final_scores=-1.2220519, final_local_scores=-1.2847514,
pem=-0.00656814, similarity=-34.214115]

14: 270081 Synthetic_fiber
 [winner: 24019437 Synthetic_musk, pem_pos=8,
 final_scores=-1.280363, final_local_scores=-1.1867464,
 pem=-2.0794415, similarity=-20.01684|
 gt: 270081 Synthetic_fiber, pem_pos=0,
 final_scores=-1.3570511, final_local_scores=-1.3540316,
 pem=-0.84031415, similarity=-31.747532]

15: 1156289 Artificial_turf
 [winner: 13835777 Gras, pem_pos=2,
 final_scores=-0.9372428, final_local_scores=-0.79927045,
 pem=-1.3834157, similarity=-12.238882|
 gt: 1156289 Artificial_turf, pem_pos=1,
 final_scores=-1.5120018, final_local_scores=-1.5905824,
 pem=-1.3261573, similarity=-36.16459]

16: 7367299 Health_club
 [winner: 44574 Weight_training, pem_pos=0,
 final_scores=-0.9397988, final_local_scores=-0.9766293,
 pem=-0.01047132, similarity=-24.99791|
 gt: 7367299 Health_club, pem_pos=3,
 final_scores=-1.8845546, final_local_scores=-2.017989,
 pem=-5.6629605, similarity=-25.33504]

17: 9396450 Wesley_A._Brown
 [winner: 17672410 Wesley_E._Brown, pem_pos=0,
 final_scores=-0.86040175, final_local_scores=-0.87067974,
 pem=0.0, similarity=-21.893106|
 gt: 9396450 Wesley_A._Brown, pem_pos=1,
 final_scores=-0.96824455, final_local_scores=-1.0390201,
 pem=-0.6931472, similarity=-23.147278]
 [winner: 17672410 Wesley_E._Brown, pem_pos=0,
 final_scores=-0.9787669, final_local_scores=-0.99997574,
 pem=0.0, similarity=-25.751558|
 gt: 9396450 Wesley_A._Brown, pem_pos=1,
 final_scores=-1.0084761, final_local_scores=-1.0829669,
 pem=-0.6931472, similarity=-24.458738]
 [winner: 13533133 Aeschynomene, pem_pos=1,
 final_scores=-1.1920378, final_local_scores=-1.0713875,
 pem=-2.0412204, similarity=-16.782152|
 RECALL MISS: 9396450 Wesley_A._Brown]

18: 6876917 Naval_Facilities_Engineering_Command
 [winner: 164858 Command_pattern, pem_pos=0,
 final_scores=-0.8493233, final_local_scores=-0.75518197,
 pem=-0.33437473, similarity=-16.628042|
 RECALL MISS: 6876917
 Naval_Facilities_Engineering_Command]
 [winner: 20518076 United_States_Navy, pem_pos=0,
 final_scores=-1.053799, final_local_scores=-1.168108,
 pem=0.0, similarity=-30.768957|
 RECALL MISS: 6876917
 Naval_Facilities_Engineering_Command]
 [winner: 41132 Telecommunications_facility, pem_pos=0,
 final_scores=-1.5111847, final_local_scores=-1.4664998,
 pem=-1.9597288, similarity=-29.016256|

RECALL MISS: 6876917
 Naval_Facilities_Engineering_Command]
 [winner: 16284578
 University_of_Illinois_at_Chicago_College_of_Engineering,
 pem_pos=7, final_scores=-1.6563913,
 final_local_scores=-1.6779435, pem=-3.1334255,
 similarity=-28.943413|
 RECALL MISS: 6876917
 Naval_Facilities_Engineering_Command]
 1: 395874 Bancroft_Hall
 [pem_pos=0, final_scores=0.17163222,
 final_local_scores=0.2382043, pem=0.0, similarity=11.198204]
 1: 26141631 Wesley_Brown_Field_House
 [pem_pos=0, final_scores=0.038548425,
 final_local_scores=0.09541564, pem=0.0, similarity=6.9371057]
 2: 57839 Annapolis,_Maryland
 [pem_pos=0, final_scores=0.32858598,
 final_local_scores=0.30225754, pem=-0.22970897,
 similarity=14.358874]
 3: 18858 Maryland
 [pem_pos=0, final_scores=0.31155092,
 final_local_scores=0.30097538, pem=-0.2529562,
 similarity=14.447034]
 4: 2154 African_American
 [pem_pos=1, final_scores=0.2348528,
 final_local_scores=0.25838673, pem=-0.7252736,
 similarity=15.74464]
 5: 17266331 Hensel_Phelps_Construction
 [pem_pos=0, final_scores=0.08320697,
 final_local_scores=0.19367921, pem=0.0, similarity=9.869487]

Sample: b'David_West'

-LRB- basketball -RRB- David Moorer West -LRB- born August
 29 , 1980 -RRB- is an [1 1 American 1] professional 1 basketb
 all player who currently plays for the [2 2 Indiana Pacers 2]
 of the . After growing up in his birthplace of [3 3 Teaneck 3]
 , [4 4 New Jersey 4] , West went to high school at 2 Garner Ma
 gnet High School in [1 3 Garner 1] , [2 North Carolina 2] and
 at [5 5 Hargrave Military Academy 5] in 4 Chatham , [3 Virgin
 ia 3] and won First Team All-State honors in [6 6 North Carol
 ina 6] . West attended and graduated from [7 7 Xavier Univers
 ity 7] in [8 8 Cincinnati 8] , [9 9 Ohio 9] . There , he was
 a four-year starter on the basketball team . West was the
 first XU player to get his number retired while still playing
 for the Musketeers , during which he led Xavier to a top 10
 national ranking and a # 3 seed in the [4 NCAA Tournament
 4] . West was named as a first-team [5 All-American 5] his
 senior year and the 2003 [6 Associated Press 6] and 5 United S
 tates Basketball Writers Association National Player of the
 Year , averaging 20.1 ppg and 11.8 rpg as a senior while
 leading XU to its second straight 26 -- 6 season . West led
 Xavier to 94 wins in his four seasons while finishing as

Xavier 's all-time blocked shot leader -LRB- 228 blocks -RRB- and second on XU 's all-time scoring -LRB- 2,132 points -RRB- and rebounding -LRB- 1,308 rebounds -RRB- lists . West was also named as the [7 Atlantic 10 Conference 7] Player of the Year three years in a row . As a senior , he was selected as the winner , recognizing him as the nation 's top senior men 's basketball player . In October 2007 , [10 10 Sports Illust rated 10] released [8 Sports Illustrated 8] : The Basketball Book , which featured West on its NCAA All-Decade Team for the decade beginning in 2000 . He was joined by 6 Connecticut 's [11 11 Emeka Okafor 11] and three 7 Duke players , [12 12 Jay Williams 12] , [13 13 J. J. Redick 13] and [14 14 Shane Battier 14] .

	33	33	29.4	.532
	.000	.667	9.1	1.7
	1.6	1.1	11.7	-
	8 2000	-- 01		
	29	29	33.7	.551
	.000	.740	10.9	2.0
	1.5	2.1	17.8	-
	9 2001	-- 02		
	32	32	34.2	.536
	.310	.768	9.8	1.6
	1.2	2.5	18.3	-
	32	32	36.5	.513
	.816	11.8	3.2	1.3
	1.6	20.1		

-RCB- West was drafted by the in the first round , eighteenth overall in the 2003 [9 NBA Draft 9] . After recording modest production in his first two seasons , in 10 2005 -- 06 West enjoyed a breakthrough , averaging 17.4 points and 7.4 rebounds with a . 512 FG % in 74 games . That year , he finished second in the voting for the [10 11 NBA 10] Most Improved Player Award , behind [15 15 Boris Diaw 15] of the [16 16 Phoenix Suns 16] . He also saw some time at 12 center , due to the team 's lack of a true big man . That season he demonstrated an ability to perform under pressure , as he converted three game-winning jump shots . In 13 2006 -- 07 , West ended the season averaging 18.3 points , 8.2 rebounds and 2.2 assists . Despite missing 30 games due to an 14 elbow injury , West recorded five games of 30 or more points and 13 games of 20 or more points . West again increased his statistical production in 15 2007 -- 08 , and was selected to his first [1 1 16 NBA All-Star Game 11] as a reserve , along with teammate [17 17 Chris Paul 17] . [12 Paul 12] later stated that he would rather not play with a power forward in the league other than West , because `` he 's not in it for the spotlight , all he cares about is getting buckets -LSB- scoring -RSB- . '' West reinforced his desire to be a team player in a 2008 postseason interview , saying that he does not care about being overlooked but would rather focus on `` play -LSB- ing -RSB- and be -LSB- ing -RSB- as effective as possible '' when he is out on the court . In 17 2008 -- 09 West repeated as an all-star and posted a new career high by scoring 21.0 points per game . On March 25 , 2011 , during

one of the ' last games of the season vs. the [18 18 Utah Jazz 18] , West came down awkwardly after a dunk . The event occurred with 22.3 left in the fourth quarter . West tore his 18 anterior cruciate ligament in his left knee . On June 27 , 2011 , West confirmed he was opting out of the final year of his contract to become a free agent . On December 11 , 2011 , West agreed to a 2 year deal with the [19 19 Indiana Pacers 19] , reportedly worth \$ 20 million . On January 12 , 2013 , West recorded his 1st career triple-double with 14 points , 12 rebounds , and 10 assists . At 6 ' 9 ' ' and , West is a 19 power forward . His game is fundamental for the most part , as he is not known for being a flamboyant player . West is known for his explosive scoring capabilities , with his fine midrange game and threatening inside presence . He is also an effective rebounder and consistent on the defensive end . West is an above average ball handler for his size , capable of converting with either hand . West is extremely interested in `` black history , philosophy and various societal issues '' and has been cited by Chris Paul as being the smartest [13 New Orleans Hornets 13] player . West graduated from Xavier with a bachelor of arts degree in Communications . He is also an accomplished musician who plays the 20 tuba . He and his wife , [14 Lesley 14] , have one daughter , Dasia , and one son , David Benjamin . West also loves to make mixtapes for his teammates , saying that it is a healthy habit that keeps him `` in the house and off -LSB- his -RSB- feet '' . West comes from a family of above-average height . His mother , Harriet , is 6 ft tall -LRB- 1.83 m -RRB- , while his father , Amos , is 6 ft 5 in -LRB- 1.96 m -RRB- . West has a Xavier `` X '' tattooed on his arm which displays , on either side , `` My life , my way '' . | - | align = `` left '' | | align = `` left '' | [15 21 New Orleans 15] | 71 | | 1 | | 13.1 | | .474 | | .000 | | .713 | | 4.2 | | .8 | | .4 | | .4 | | 3.8 | - | align = `` left '' | | align = `` left '' | [16 22 New Orleans 16] | 30 | | 8 | | 18.4 | | .436 | | .400 | | .680 | | 4.3 | | .8 | | .4 | | .5 | | 6.2 | - | align = `` left '' | | align = `` left '' | [17 23 New Orleans/Oklahoma City 17] | 74 | | 74 | | 34.1 | | .512 | | .273 | | .843 | | 7.4 | | 1.2 | | .8 | | .9 | | 17.1 | - | align = `` left '' | | align = `` left '' | [18 24 New Orleans/Oklahoma City 18] | 52 | | 52 | | 36.5 | | .476 | | .320 | | .824 | | 8.2 | | 2.2 | | .8 | | .7 | | 18.3 | - | align = `` left '' | | align = `` left '' | [19 25 New Orleans 19] | 76 | | 76 | | 37.8 | | .482 | | .240 | | .850 | | 8.9 | | 2.3 | | .8 | | 1.3 | | 20.6 | - | align = `` left '' | | align = `` left '' | [20 26 New Orleans 20] | 76 | | 76 | | 39.2 | | .472 | | .240 | | .884 | | 8.5 | | 2.3 | | .6 | | .9 | | 21.0 | - | align = `` left '' | | align = `` left '' | [21 27 New Orleans 21] | 81 | | 81 | | 36.4 | | .505 | | .259 | | .865 | | 7.5 | | 3.0 | | .9 | | .7 | | 19.0 | - | align = `` left '' | | align = `` left '' | [22 28 New Orleans 22] | 70 | | 70 | | 35.0 | | .508 | | .222 | | .807 | | 7.6 | | 2.3 | | 1.0 | | .9 | | 18.9 | - | align

```

= `` left '' | | align = `` left '' | 29 Indiana | 66 | | 66
| | 29.2 | | .487 | | .222 | | .807 | | 6.6 | | 2.1 | | .8 |
| .7 | | 12.8 | - | align = `` left '' | | align = `` left ''
| | 73 | | 73 | | 33.4 | | .498 | | .211 | | .768 | | 7.7 | |
2.9 | | 1.0 | | 1.0 | | 17.1 | - | align = `` center ''
colspan = `` 2 '' | Career | 669 | | 577 | | 32.2 | | .491 |
| .255 | | .828 | | 7.2 | | 2.1 | | .8 | | .8 | | 16.1 | - |
align = `` center '' colspan = `` 2 '' | All-Star | 2 | | 0 |
| 15.0 | | .545 | | .000 | | .000 | | 3.5 | | .5 | | .5 | | .
0 | | 6.0 | - | align = `` left '' | 30 2004 | align = ``
left '' | [23 31 New Orleans 23] | 7 | | 0 | | 15.9 | | .536
| | .000 | | .846 | | 4.3 | | 1.1 | | .3 | | .6 | | 5.9 | - |
align = `` left '' | 32 2008 | align = `` left '' | [24 33 New
Orleans 24] | 12 | | 12 | | 40.4 | | .466 | | .500 | | .891 |
| 8.5 | | 2.8 | | 1.1 | | 1.9 | | 21.2 | - | align = `` left
'' | 34 2009 | align = `` left '' | [25 35 New Orleans 25] |
5 | | 5 | | 35.6 | | .400 | | .000 | | .897 | | 7.4 | | 1.2 |
| 1.0 | | .4 | | 18.0 | - | align = `` left '' | 36 2012 |
align = `` left '' | | 11 | | 11 | | 37.8 | | .446 | | .000 |
| .818 | | 8.5 | | 2.0 | | .7 | | .5 | | 15.3 | - | align =
`` left '' | 37 2013 | align = `` left '' |

```

- 1: 3434750 United_States
- 2: 72875 Indiana_Pacers
- 3: 125011 Teaneck,_New_Jersey
- 4: 21648 New_Jersey
- 5: 5024359 Hargrave_Military_Academy
- 6: 21650 North_Carolina
- 7: 521840 Xavier_University
- 8: 18522615 Cincinnati
- 9: 22199 Ohio
- 10: 148956 Sports_Illustrated
- 11: 556331 Emeka_Okafor
- 12: 1794213 Jay_Williams_(basketball)
- 13: 1303234 J._J._Redick
- 14: 2109459 Shane_Battier
- 15: 2387461 Boris_Diaw
- 16: 72895 Phoenix_Suns
- 17: 4987149 Chris_Paul
- 18: 72890 Utah_Jazz
- 19: 72875 Indiana_Pacers

1: 3921 Basketball

[winner: 10088476 Sinbad_(comedian), pem_pos=0,
final_scores=-0.742592, final_local_scores=-0.7419298,
pem=0.0, similarity=-18.050953]

gt: 3921 Basketball, pem_pos=1, final_scores=-1.3587544,
final_local_scores=-1.4957857, pem=-0.05641887,
similarity=-40.240704]

2: 4008377 Garner_Magnet_High_School

[winner: 4008377 Garner_Magnet_High_School, pem_pos=0,
final_scores=-0.047372892, final_local_scores=0.03122963,
pem=0.0, similarity=5.0216675]

gt: 4008377 Garner_Magnet_High_School, pem_pos=0,

final_scores=-0.047372892, final_local_scores=0.03122963,
pem=0.0, similarity=5.0216675]

3: 128327 Garner, North_Carolina
[winner: 21650 North_Carolina, pem_pos=0,
final_scores=0.36766338, final_local_scores=0.3747232,
pem=-0.21680297, similarity=16.451208]
RECALL MISS: 128327 Garner, North_Carolina]
[winner: 128327 Garner, North_Carolina, pem_pos=0,
final_scores=0.04786168, final_local_scores=0.07019532,
pem=-0.90406543, similarity=11.100929]
gt: 128327 Garner, North_Carolina, pem_pos=0,
final_scores=0.04786168, final_local_scores=0.07019532,
pem=-0.90406543, similarity=11.100929]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7271423, final_local_scores=-1.6946142,
pem=-4.6728287, similarity=-21.069384]
RECALL MISS: 128327 Garner, North_Carolina]

4: 137713 Chatham, Virginia
[winner: 32432 Virginia, pem_pos=0,
final_scores=0.3820601, final_local_scores=0.3959949,
pem=-0.2023721, similarity=17.00752]
RECALL MISS: 137713 Chatham, Virginia]
[winner: 59899 Chatham, Kent, pem_pos=0,
final_scores=-0.06300336, final_local_scores=-0.042654216,
pem=-1.6528518, similarity=11.805291]
gt: 137713 Chatham, Virginia, pem_pos=1,
final_scores=-0.27658644, final_local_scores=-0.28389993,
pem=-2.35489, similarity=8.423831]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7324984, final_local_scores=-1.700465,
pem=-4.6728287, similarity=-21.24398]
RECALL MISS: 137713 Chatham, Virginia]

5: 3564194 United_States_Basketball_Writers_Association
[winner: 3564194
United_States_Basketball_Writers_Association, pem_pos=0,
final_scores=-0.44939372, final_local_scores=-0.4654867,
pem=0.0, similarity=-9.801339]
gt: 3564194
United_States_Basketball_Writers_Association, pem_pos=0,
final_scores=-0.44939372, final_local_scores=-0.4654867,
pem=0.0, similarity=-9.801339]

6: 239846 University_of_Connecticut
[winner: 6466 Connecticut, pem_pos=0,
final_scores=-0.13007094, final_local_scores=-0.16900605,
pem=-0.24788867, similarity=0.39429244]
gt: 239846 University_of_Connecticut, pem_pos=1,
final_scores=-0.3454618, final_local_scores=-0.40827703,
pem=-3.2877145, similarity=9.785014]

7: 53273 Duke_University
[winner: 8219053 Duke_Blue_Devils_men's_basketball,
pem_pos=2, final_scores=-0.05811791,
final_local_scores=-0.0945985, pem=-2.379933,

similarity=14.209151|
gt: 53273 Duke_University, pem_pos=1,
final_scores=-0.13099979, final_local_scores=-0.17213416,
pem=-1.6972626, similarity=8.182864]
8: 24348134 2000-01_NCAA_Division_I_men's_basketball_season
[winner: 34548 2000, pem_pos=0, final_scores=-0.44848555,
final_local_scores=-0.35460532, pem=-0.8862115,
similarity=-1.6730611|
RECALL MISS: 24348134 2000-
01_NCAA_Division_I_men's_basketball_season]
[winner: 34551 2001, pem_pos=3, final_scores=-0.85537875,
final_local_scores=-0.6409848, pem=-2.4587388,
similarity=-1.6675476|
RECALL MISS: 24348134 2000-
01_NCAA_Division_I_men's_basketball_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3908821, final_local_scores=-1.2951741,
pem=0.0, similarity=-34.560867|
RECALL MISS: 24348134 2000-
01_NCAA_Division_I_men's_basketball_season]
9: 24248450 2001-02_NCAA_Division_I_men's_basketball_season
[winner: 34551 2001, pem_pos=0, final_scores=-0.57286686,
final_local_scores=-0.3323834, pem=-0.5613713,
similarity=-2.7764463|
RECALL MISS: 24248450 2001-
02_NCAA_Division_I_men's_basketball_season]
[winner: 2385469 Federal_Firearms_License, pem_pos=2,
final_scores=-0.7408147, final_local_scores=-0.6920046,
pem=-2.2935443, similarity=-4.088435|
RECALL MISS: 24248450 2001-
02_NCAA_Division_I_men's_basketball_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3596494, final_local_scores=-1.2610571,
pem=0.0, similarity=-33.542747|
RECALL MISS: 24248450 2001-
02_NCAA_Division_I_men's_basketball_season]
10: 2498103 2005-06_NBA_season
[winner: 36164 2006, pem_pos=0, final_scores=-0.8323618,
final_local_scores=-0.7291575, pem=-1.542131,
similarity=-9.283454|
RECALL MISS: 2498103 2005-06_NBA_season]
[winner: 35984 2005, pem_pos=0, final_scores=-0.84183776,
final_local_scores=-0.7171938, pem=-0.36604404,
similarity=-15.322175|
RECALL MISS: 2498103 2005-06_NBA_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.274489, final_local_scores=-1.1680324,
pem=0.0, similarity=-30.766705|
RECALL MISS: 2498103 2005-06_NBA_season]
11: 2066890 NBA_Most_Improved_Player_Award
[winner: 22093 National_Basketball_Association,
pem_pos=0, final_scores=0.115782484,

final_local_scores=0.04330252, pem=0.0, similarity=5.3819466|
RECALL MISS: 2066890 NBA_Most_Improved_Player_Award]
[winner: 2066890 NBA_Most_Improved_Player_Award,
pem_pos=0, final_scores=-0.16303697,
final_local_scores=-0.16640945, pem=0.0,
similarity=-0.8762769|
gt: 2066890 NBA_Most_Improved_Player_Award, pem_pos=0,
final_scores=-0.16303697, final_local_scores=-0.16640945,
pem=0.0, similarity=-0.8762769]
12: 652569 Center_(basketball)
[winner: 123071 Center_Missouri, pem_pos=2,
final_scores=-1.3796746, final_local_scores=-1.3896828,
pem=-2.3863506, similarity=-24.403845|
gt: 652569 Center_(basketball), pem_pos=0,
final_scores=-1.4104921, final_local_scores=-1.5510314,
pem=-1.6317346, similarity=-33.322533]
13: 3395535 2006-07_NBA_season
[winner: 36164 2006, pem_pos=0, final_scores=-0.3759871,
final_local_scores=-0.2306374, pem=-0.3060146,
similarity=-1.1288122|
RECALL MISS: 3395535 2006-07_NBA_season]
[winner: 36165 2007, pem_pos=0, final_scores=-0.6937422,
final_local_scores=-0.58282185, pem=-1.6554711,
similarity=-4.3001447|
RECALL MISS: 3395535 2006-07_NBA_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3298218, final_local_scores=-1.2284751,
pem=0.0, similarity=-32.570435|
RECALL MISS: 3395535 2006-07_NBA_season]
14: 19595436 Elbow
[winner: 6336889 Peter_Elbow, pem_pos=5,
final_scores=-1.2575139, final_local_scores=-1.2619283,
pem=-1.9384128, similarity=-23.027355|
gt: 19595436 Elbow, pem_pos=1, final_scores=-1.4500337,
final_local_scores=-1.4551752, pem=-1.1764625,
similarity=-32.93783]
15: 5555179 2007-08_NBA_season
[winner: 36165 2007, pem_pos=0, final_scores=-0.6189595,
final_local_scores=-0.5011331, pem=-0.24805899,
similarity=-9.516113|
RECALL MISS: 5555179 2007-08_NBA_season]
[winner: 3454274 FabricLive.08, pem_pos=5,
final_scores=-0.82619894, final_local_scores=-0.6451164,
pem=-2.815669, similarity=0.15019807|
RECALL MISS: 5555179 2007-08_NBA_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.2820748, final_local_scores=-1.1763186,
pem=0.0, similarity=-31.013979|
RECALL MISS: 5555179 2007-08_NBA_season]
16: 4118244 2008_NBA_All-Star_Game
[winner: 278018 NBA_All-Star_Game, pem_pos=0,
final_scores=0.06671272, final_local_scores=0.023377076,

pem=0.0, similarity=4.7873316|
 gt: 4118244 2008_NBA_All-Star_Game, pem_pos=5,
 final_scores=-1.0243298, final_local_scores=-1.1301155,
 pem=-5.9285913, similarity=2.6054134]
 17: 17016050 2008-09_NBA_season
 [winner: 35825 2008, pem_pos=0, final_scores=-0.83412755,
 final_local_scores=-0.74098957, pem=-0.29331458,
 similarity=-16.427803|
 RECALL MISS: 17016050 2008-09_NBA_season]
 [winner: 3454417 FabricLive.09, pem_pos=4,
 final_scores=-1.0322652, final_local_scores=-0.89385736,
 pem=-2.6940393, similarity=-7.9341683|
 RECALL MISS: 17016050 2008-09_NBA_season]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.3546515, final_local_scores=-1.2555976,
 pem=0.0, similarity=-33.37982|
 RECALL MISS: 17016050 2008-09_NBA_season]
 18: 578923 Anterior_cruciate_ligament
 [winner: 5012564 Cruciate_anastomosis, pem_pos=0,
 final_scores=-0.90391403, final_local_scores=-0.7445962,
 pem=-0.86499745, similarity=-13.426532|
 gt: 578923 Anterior_cruciate_ligament, pem_pos=1,
 final_scores=-1.496345, final_local_scores=-1.6091647,
 pem=-0.9295359, similarity=-38.87601]
 [winner: 603770 Cooper's_ligaments, pem_pos=0,
 final_scores=-1.0522023, final_local_scores=-0.9397153,
 pem=0.0, similarity=-23.953266|
 gt: 578923 Anterior_cruciate_ligament, pem_pos=2,
 final_scores=-2.2760444, final_local_scores=-2.460868,
 pem=-5.1561775, similarity=-41.307396]
 [winner: 5296572 Anterior_tibial_recurrent_artery,
 pem_pos=13, final_scores=-1.1915494,
 final_local_scores=-1.0527043, pem=-4.082739,
 similarity=-5.1225147|
 gt: 578923 Anterior_cruciate_ligament, pem_pos=10,
 final_scores=-1.9541572, final_local_scores=-2.109255,
 pem=-3.896153, similarity=-37.66678]
 19: 284482 Power_forward_(basketball)
 [winner: 284482 Power_forward_(basketball), pem_pos=0,
 final_scores=-1.1630179, final_local_scores=-1.2543766,
 pem=0.0, similarity=-33.34339|
 gt: 284482 Power_forward_(basketball), pem_pos=0,
 final_scores=-1.1630179, final_local_scores=-1.2543766,
 pem=0.0, similarity=-33.34339]
 20: 30961 Tuba
 [winner: 3138384 Roman_tuba, pem_pos=1,
 final_scores=-1.1132199, final_local_scores=-1.0480652,
 pem=-0.6316987, similarity=-23.75137|
 gt: 30961 Tuba, pem_pos=0, final_scores=-1.2736726,
 final_local_scores=-1.3317909, pem=-0.097273715,
 similarity=-35.124596]
 21: 33114988 2003-04_New_Orleans_Hornets_season

[winner: 53842 New_Orleans, pem_pos=0,
 final_scores=0.25553066, final_local_scores=0.21567842,
 pem=-0.12199242, similarity=11.189401]
 RECALL MISS: 33114988 2003-04_New_Orleans_Hornets_season]
 22: 33115133 2004-05_New_Orleans_Hornets_season
 [winner: 53842 New_Orleans, pem_pos=0,
 final_scores=0.2646355, final_local_scores=0.22562405,
 pem=-0.12199242, similarity=11.486198]
 RECALL MISS: 33115133 2004-05_New_Orleans_Hornets_season]
 23: 18289217 2005-06_New_Orleans/Oklahoma_City_Hornets_season
 [winner: 72862 New_Orleans_Pelicans, pem_pos=0,
 final_scores=0.15142088, final_local_scores=0.11225882,
 pem=0.0, similarity=7.43974]
 RECALL MISS: 18289217 2005-06_New_Orleans/
 Oklahoma_City_Hornets_season]
 24: 28813868 2006-07_New_Orleans/Oklahoma_City_Hornets_season
 [winner: 72862 New_Orleans_Pelicans, pem_pos=0,
 final_scores=0.15163492, final_local_scores=0.11249262,
 pem=0.0, similarity=7.4467173]
 gt: 28813868 2006-07_New_Orleans/
 Oklahoma_City_Hornets_season, pem_pos=1,
 final_scores=-0.28404242, final_local_scores=-0.20314041,
 pem=-0.9315582, similarity=3.0935564]
 25: 11963536 2007-08_New_Orleans_Hornets_season
 [winner: 53842 New_Orleans, pem_pos=0,
 final_scores=0.27434433, final_local_scores=0.23622951,
 pem=-0.12199242, similarity=11.802686]
 RECALL MISS: 11963536 2007-08_New_Orleans_Hornets_season]
 26: 17288861 2008-09_New_Orleans_Hornets_season
 [winner: 53842 New_Orleans, pem_pos=0,
 final_scores=0.27193195, final_local_scores=0.23359436,
 pem=-0.12199242, similarity=11.724048]
 RECALL MISS: 17288861 2008-09_New_Orleans_Hornets_season]
 27: 23284757 2009-10_New_Orleans_Hornets_season
 [winner: 53842 New_Orleans, pem_pos=0,
 final_scores=0.25641793, final_local_scores=0.21664762,
 pem=-0.12199242, similarity=11.218324]
 RECALL MISS: 23284757 2009-10_New_Orleans_Hornets_season]
 28: 27723228 2010-11_New_Orleans_Hornets_season
 [winner: 53842 New_Orleans, pem_pos=0,
 final_scores=0.25034875, final_local_scores=0.21001798,
 pem=-0.12199242, similarity=11.020482]
 RECALL MISS: 27723228 2010-11_New_Orleans_Hornets_season]
 29: 32095450 2011-12_Indiana_Pacers_season
 [winner: 21883857 Indiana, pem_pos=0,
 final_scores=0.021577045, final_local_scores=-0.013248041,
 pem=-0.27230176, similarity=5.1751842]
 RECALL MISS: 32095450 2011-12_Indiana_Pacers_season]
 30: 5025170 2004_NBA_Playoffs
 [winner: 35524 2004, pem_pos=0, final_scores=-0.6076205,
 final_local_scores=-0.46132022, pem=-0.4915334,
 similarity=-7.0039673]

RECALL MISS: 5025170 2004_NBA_Playoffs]
31: 33114988 2003-04_New_Orleans_Hornets_season
[winner: 53842 New_Orleans, pem_pos=0,
final_scores=0.26458287, final_local_scores=0.22556657,
pem=-0.12199242, similarity=11.484483]
RECALL MISS: 33114988 2003-04_New_Orleans_Hornets_season]
32: 11967895 2008_NBA_Playoffs
[winner: 35825 2008, pem_pos=0, final_scores=-0.60056967,
final_local_scores=-0.48586303, pem=-0.29331458,
similarity=-8.814319]
RECALL MISS: 11967895 2008_NBA_Playoffs]
33: 11963536 2007-08_New_Orleans_Hornets_season
[winner: 53842 New_Orleans, pem_pos=0,
final_scores=0.29353154, final_local_scores=0.25718862,
pem=-0.12199242, similarity=12.428147]
RECALL MISS: 11963536 2007-08_New_Orleans_Hornets_season]
34: 16858598 2009_NBA_Playoffs
[winner: 35983 2009, pem_pos=0, final_scores=-0.8955446,
final_local_scores=-0.8005445, pem=-0.9296576,
similarity=-14.74451]
RECALL MISS: 16858598 2009_NBA_Playoffs]
35: 17288861 2008-09_New_Orleans_Hornets_season
[winner: 53842 New_Orleans, pem_pos=0,
final_scores=0.26767415, final_local_scores=0.22894332,
pem=-0.12199242, similarity=11.585252]
RECALL MISS: 17288861 2008-09_New_Orleans_Hornets_season]
36: 34286149 2012_NBA_Playoffs
[winner: 47374 2012, pem_pos=0, final_scores=-0.70255065,
final_local_scores=-0.49191234, pem=-1.4461981,
similarity=-2.7252822]
RECALL MISS: 34286149 2012_NBA_Playoffs]
37: 35600813 2013_NBA_Playoffs
[winner: 46945 2013, pem_pos=0, final_scores=-0.8791959,
final_local_scores=-0.69757503, pem=-1.490928,
similarity=-8.619419]
RECALL MISS: 35600813 2013_NBA_Playoffs]
1: 3434750 United_States
[pem_pos=0, final_scores=0.5374001,
final_local_scores=0.5679439, pem=-0.12994486,
similarity=21.744951]
2: 72875 Indiana_Pacers
[pem_pos=0, final_scores=0.28134793,
final_local_scores=0.25337288, pem=0.0, similarity=11.650865]
3: 125011 Teaneck,_New_Jersey
[pem_pos=0, final_scores=0.36460376,
final_local_scores=0.3915041, pem=0.0, similarity=15.772977]
4: 21648 New_Jersey
[pem_pos=0, final_scores=0.4822287,
final_local_scores=0.49536604, pem=-0.13504621,
similarity=19.606825]
5: 5024359 Hargrave_Military_Academy
[pem_pos=0, final_scores=0.17922966,

final_local_scores=0.1920546, pem=0.0, similarity=9.821005]
6: 21650 North_Carolina
[pem_pos=0, final_scores=0.374273,
final_local_scores=0.38194317, pem=-0.21680297,
similarity=16.666666]
7: 521840 Xavier_University
[pem_pos=0, final_scores=0.28258693,
final_local_scores=0.2655621, pem=0.0, similarity=12.014615]
8: 18522615 Cincinnati
[pem_pos=0, final_scores=0.49375385,
final_local_scores=0.48523766, pem=-0.46106797,
similarity=21.077532]
9: 22199 Ohio
[pem_pos=0, final_scores=0.5212098,
final_local_scores=0.54164976, pem=-0.19147073,
similarity=21.29487]
10: 148956 Sports_Illustrated
[pem_pos=0, final_scores=0.29827434,
final_local_scores=0.28838357, pem=0.0, similarity=12.695654]
11: 556331 Emeka_Okafor
[pem_pos=0, final_scores=0.49640715,
final_local_scores=0.52822745, pem=0.0, similarity=19.853073]
12: 1794213 Jay_Williams_(basketball)
[pem_pos=0, final_scores=0.21802293,
final_local_scores=0.23233354, pem=-0.5584368,
similarity=14.0598755]
13: 1303234 J._J._Redick
[pem_pos=0, final_scores=0.2414528,
final_local_scores=0.25786188, pem=0.0, similarity=11.784826]
14: 2109459 Shane_Battier
[pem_pos=0, final_scores=0.3878426,
final_local_scores=0.40443474, pem=0.0, similarity=16.158852]
15: 2387461 Boris_Diaw
[pem_pos=0, final_scores=0.499435,
final_local_scores=0.5863105, pem=0.0, similarity=21.586388]
16: 72895 Phoenix_Suns
[pem_pos=0, final_scores=0.24844496,
final_local_scores=0.21524242, pem=0.0, similarity=10.512976]
17: 4987149 Chris_Paul
[pem_pos=0, final_scores=0.04403214,
final_local_scores=0.015860528, pem=0.0, similarity=4.5630226]
18: 72890 Utah_Jazz
[pem_pos=0, final_scores=0.27044684,
final_local_scores=0.24407554, pem=0.0, similarity=11.373414]
19: 72875 Indiana_Pacers
[pem_pos=0, final_scores=0.27804846,
final_local_scores=0.24976873, pem=0.0, similarity=11.54331]
1: 128327 Garner,_North_Carolina
[pem_pos=0, final_scores=0.04786168,
final_local_scores=0.07019532, pem=-0.90406543,
similarity=11.100929]
2: 21650 North_Carolina

[pem_pos=0, final_scores=0.36766338,
final_local_scores=0.3747232, pem=-0.21680297,
similarity=16.451208]
3: 32432 Virginia
[pem_pos=0, final_scores=0.3820601,
final_local_scores=0.3959949, pem=-0.2023721,
similarity=17.00752]
4: 507399 NCAA_Men's_Division_I_Basketball_Championship
[pem_pos=0, final_scores=0.15338866,
final_local_scores=0.13623416, pem=-0.54612035,
similarity=11.1251]
5: 735314 All-America
[pem_pos=0, final_scores=0.031026408,
final_local_scores=0.005898088, pem=-0.22135477,
similarity=5.4694853]
6: 18935732 Associated_Press
[pem_pos=0, final_scores=0.31375474,
final_local_scores=0.31328508, pem=0.0, similarity=13.438765]
7: 79391 Atlantic_10_Conference
[pem_pos=0, final_scores=0.08787198,
final_local_scores=0.07590073, pem=-0.012717425,
similarity=6.4239016]
8: 148956 Sports_Illustrated
[pem_pos=0, final_scores=0.07722442,
final_local_scores=0.044019103, pem=0.0, similarity=5.403331]
9: 835946 NBA_draft
[pem_pos=0, final_scores=0.22686164,
final_local_scores=0.2140767, pem=0.0, similarity=10.4781885]
10: 22093 National_Basketball_Association
[pem_pos=0, final_scores=0.115782484,
final_local_scores=0.04330252, pem=0.0, similarity=5.3819466]
11: 278018 NBA_All-Star_Game
[pem_pos=0, final_scores=0.06671272,
final_local_scores=0.023377076, pem=0.0, similarity=4.7873316]
12: 4987149 Chris_Paul
[pem_pos=0, final_scores=0.06436138,
final_local_scores=0.038067117, pem=0.0, similarity=5.225712]
13: 72862 New_Orleans_Pelicans
[pem_pos=0, final_scores=0.15976815,
final_local_scores=0.12137693, pem=0.0, similarity=7.7118425]
14: 3372715 Lesley_University
[pem_pos=0, final_scores=0.060636416,
final_local_scores=0.11313999, pem=-0.02855616,
similarity=7.621329]
15: 53842 New_Orleans
[pem_pos=0, final_scores=0.25553066,
final_local_scores=0.21567842, pem=-0.12199242,
similarity=11.189401]
16: 53842 New_Orleans
[pem_pos=0, final_scores=0.2646355,
final_local_scores=0.22562405, pem=-0.12199242,
similarity=11.486198]

17: 72862 New_Orleans_Pelicans
 [pem_pos=0, final_scores=0.15142088,
 final_local_scores=0.11225882, pem=0.0, similarity=7.43974]
 18: 72862 New_Orleans_Pelicans
 [pem_pos=0, final_scores=0.15163492,
 final_local_scores=0.11249262, pem=0.0, similarity=7.4467173]
 19: 53842 New_Orleans
 [pem_pos=0, final_scores=0.27434433,
 final_local_scores=0.23622951, pem=-0.12199242,
 similarity=11.802686]
 20: 53842 New_Orleans
 [pem_pos=0, final_scores=0.27193195,
 final_local_scores=0.23359436, pem=-0.12199242,
 similarity=11.724048]
 21: 53842 New_Orleans
 [pem_pos=0, final_scores=0.25641793,
 final_local_scores=0.21664762, pem=-0.12199242,
 similarity=11.218324]
 22: 53842 New_Orleans
 [pem_pos=0, final_scores=0.25034875,
 final_local_scores=0.21001798, pem=-0.12199242,
 similarity=11.020482]
 23: 53842 New_Orleans
 [pem_pos=0, final_scores=0.26458287,
 final_local_scores=0.22556657, pem=-0.12199242,
 similarity=11.484483]
 24: 53842 New_Orleans
 [pem_pos=0, final_scores=0.29353154,
 final_local_scores=0.25718862, pem=-0.12199242,
 similarity=12.428147]
 25: 53842 New_Orleans
 [pem_pos=0, final_scores=0.26767415,
 final_local_scores=0.22894332, pem=-0.12199242,
 similarity=11.585252]

Sample: b'Malone_Area_Heritage_Museum'

The [1 Malone Area Heritage Museum 1] is a historical museum in 1 Malone , [2 Wisconsin 2] , [3 United States 3] . The museum documents the history of immigrants into the [4 2 Holy land 4] region of [5 Wisconsin 5] . The museum is affiliated with the [1 1 Wisconsin Historical Society 1] . It has been in existence since 2005 . It is open to the public on Thursday afternoons between 2:00 and 4:00 pm and on the second Sunday of each month . The museum consists of two buildings . Both were purchased from Malone 's co-op . One is the original train depot for Malone , which was moved across the street to its present location in early 2005 . Had the building not been moved , it would have been dismantled because it was in the way of large trucks . It documents the development and usage of the [2 2 Sheboygan 2] - [3 3 Fond du Lac 3] railroad . There is a piece of the original railroad that was dug up in a local field . The building is the final

remaining depot from that line . The other building is a more recently purchased agricultural shop which is being adapted for displaying items as of 2008 . Both buildings contain artifacts from activities common in the area , especially farming . Documentation and photography are displayed from the local communities of 3 St. Cloud , 4 Mt. Calvary and its train station at [4 4 Calvary Station 4] , [5 5 Charlesburg 5] , 5 Dotyville , 6 St. Peter , 7 St. Joe , 8 Garnet , [6 6 St. Anna 6] , 9 Marytown , , 10 Johnsburg , and 11 Silica

1: 5659889 Wisconsin_Historical_Society

2: 260692 Sheboygan,_Wisconsin

3: 260424 Fond_du_Lac,_Wisconsin

4: 27663972 Calvary,_Wisconsin

5: 25430091 Charlesburg,_Wisconsin

6: 14353653 St._Anna,_Wisconsin

1: 2917068 Malone,_Wisconsin

[winner: 33127 Wisconsin, pem_pos=0,
final_scores=0.5941046, final_local_scores=0.5759035,
pem=-0.21268336, similarity=22.432426]

RECALL MISS: 2917068 Malone,_Wisconsin]

[winner: 259790 Malone_(village),_New_York, pem_pos=1,
final_scores=-0.06004198, final_local_scores=-0.084180355,
pem=-2.2630258, similarity=13.884289]

gt: 2917068 Malone,_Wisconsin, pem_pos=2,
final_scores=-0.2503404, final_local_scores=-0.2534585,
pem=-2.3497348, similarity=9.304229]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8282086, final_local_scores=-1.8126357,
pem=-4.6728287, similarity=-24.591377]

RECALL MISS: 2917068 Malone,_Wisconsin]

2: 19324277 The_Holyland_(Wisconsin)

[winner: 33127 Wisconsin, pem_pos=0,
final_scores=0.57083815, final_local_scores=0.55048835,
pem=-0.21268336, similarity=21.673988]

RECALL MISS: 19324277 The_Holyland_(Wisconsin)]

[winner: 19324277 The_Holyland_(Wisconsin), pem_pos=2,
final_scores=0.20730661, final_local_scores=0.19985911,
pem=-0.9238953, similarity=15.078194]

gt: 19324277 The_Holyland_(Wisconsin), pem_pos=2,
final_scores=0.20730661, final_local_scores=0.19985911,
pem=-0.9238953, similarity=15.078194]

[winner: 2125190 Regions_of_Togo, pem_pos=18,
final_scores=-1.4175904, final_local_scores=-1.3655119,
pem=-3.5198548, similarity=-17.518368]

RECALL MISS: 19324277 The_Holyland_(Wisconsin)]

3: 139059 St._Cloud,_Wisconsin

[winner: 94524 St._Louis_County,_Missouri, pem_pos=0,
final_scores=-0.775642, final_local_scores=-0.88090855,
pem=0.0, similarity=-22.198355]

RECALL MISS: 139059 St._Cloud,_Wisconsin]

[winner: 19541494 Cloud_computing, pem_pos=0,
final_scores=-0.86903787, final_local_scores=-0.83071923,

pem=-0.60041815, similarity=-17.435438|
RECALL MISS: 139059 St._Cloud,_Wisconsin]

4: 139055 Mount_Calvary,_Wisconsin
[winner: 23793606 Calvary_(CRT_station), pem_pos=6,
final_scores=-0.7682782, final_local_scores=-0.6906385,
pem=-2.2755344, similarity=-4.145607|
gt: 139055 Mount_Calvary,_Wisconsin, pem_pos=8,
final_scores=-1.1433958, final_local_scores=-1.2551146,
pem=-2.3025851, similarity=-20.843592]
[winner: 1030974 Baruntse, pem_pos=1,
final_scores=-1.2196347, final_local_scores=-1.1494706,
pem=-1.3434429, similarity=-22.906927|
RECALL MISS: 139055 Mount_Calvary,_Wisconsin]

5: 32900140 Dotyville,_Wisconsin
[winner: 130489 Dotyville,_Oklahoma, pem_pos=0,
final_scores=-0.12853953, final_local_scores=-0.03606835,
pem=0.0, similarity=3.0133615|
gt: 32900140 Dotyville,_Wisconsin, pem_pos=1,
final_scores=-0.5917711, final_local_scores=-0.6118334,
pem=-3.8607297, similarity=6.826632]

6: 25559632 St._Peter,_Wisconsin
[winner: 31665644 Saint_Peter, pem_pos=0,
final_scores=-0.10567087, final_local_scores=-0.061428964,
pem=-0.23529279, similarity=3.536109|
gt: 25559632 St._Peter,_Wisconsin, pem_pos=2,
final_scores=-0.4994461, final_local_scores=-0.47647643,
pem=-2.6622052, similarity=4.3481913]

7: 23189413 St._Joe,_Wisconsin
[winner: 149369 St._Joseph,_Missouri, pem_pos=5,
final_scores=-0.090272024, final_local_scores=-0.13105643,
pem=-2.0794415, similarity=11.487055|
gt: 23189413 St._Joe,_Wisconsin, pem_pos=0,
final_scores=-0.19172588, final_local_scores=-0.18261735,
pem=-0.72635275, similarity=2.5900662]

8: 24347998 Garnet,_Wisconsin
[winner: 37506 Garnet, pem_pos=0,
final_scores=-0.18764848, final_local_scores=-0.0969524,
pem=-0.22185728, similarity=2.4029543|
gt: 24347998 Garnet,_Wisconsin, pem_pos=2,
final_scores=-0.42643115, final_local_scores=-0.39392263,
pem=-2.6211545, similarity=6.5885224]

9: 13635719 Marytown,_Wisconsin
[winner: 13635719 Marytown,_Wisconsin, pem_pos=0,
final_scores=0.012956455, final_local_scores=0.043689415,
pem=0.0, similarity=5.393492|
gt: 13635719 Marytown,_Wisconsin, pem_pos=0,
final_scores=0.012956455, final_local_scores=0.043689415,
pem=0.0, similarity=5.393492]

10: 11750465 Johnsburg,_Wisconsin
[winner: 111657 Johnsburg,_Illinois, pem_pos=0,
final_scores=0.024436101, final_local_scores=0.068691105,
pem=-0.51562643, similarity=8.943649|

```

gt: 11750465 Johnsborg, Wisconsin, pem_pos=1,
final_scores=-0.02844353, final_local_scores=-0.035525076,
pem=-0.5751087, similarity=6.1571045 ]
11: 25559832 Silica, Wisconsin
[winner: 25559832 Silica, Wisconsin, pem_pos=1,
final_scores=-0.3783276, final_local_scores=-0.3125376,
pem=-1.4479326, similarity=2.6370494]
gt: 25559832 Silica, Wisconsin, pem_pos=1,
final_scores=-0.3783276, final_local_scores=-0.3125376,
pem=-1.4479326, similarity=2.6370494 ]
1: 5659889 Wisconsin Historical Society
[pem_pos=0, final_scores=0.24647705,
final_local_scores=0.2967818, pem=0.0, similarity=12.946274]
2: 260692 Sheboygan, Wisconsin
[pem_pos=0, final_scores=0.26227403,
final_local_scores=0.23236209, pem=0.0, similarity=11.023861]
3: 260424 Fond du Lac, Wisconsin
[pem_pos=0, final_scores=0.17927961,
final_local_scores=0.14712492, pem=-1.0001285,
similarity=13.919068]
4: 27663972 Calvary, Wisconsin
[pem_pos=0, final_scores=0.21102194,
final_local_scores=0.24474949, pem=0.0, similarity=11.393526]
5: 25430091 Charlesburg, Wisconsin
[pem_pos=0, final_scores=0.13501619,
final_local_scores=0.15585414, pem=0.0, similarity=8.740711]
6: 14353653 St. Anna, Wisconsin
[pem_pos=0, final_scores=0.030367866,
final_local_scores=0.053218126, pem=-0.05036491,
similarity=5.9517403]
1: 20205124 Malone Area Heritage Museum
[pem_pos=0, final_scores=0.06197594,
final_local_scores=0.111017555, pem=0.0, similarity=7.402698]
2: 33127 Wisconsin
[pem_pos=0, final_scores=0.5941046,
final_local_scores=0.5759035, pem=-0.21268336,
similarity=22.432426]
3: 3434750 United States
[pem_pos=0, final_scores=0.4475249,
final_local_scores=0.48104602, pem=-0.073457584,
similarity=18.84456]
4: 19324277 The Holyland (Wisconsin)
[pem_pos=2, final_scores=0.20730661,
final_local_scores=0.19985911, pem=-0.9238953,
similarity=15.078194]
5: 33127 Wisconsin
[pem_pos=0, final_scores=0.57083815,
final_local_scores=0.55048835, pem=-0.21268336,
similarity=21.673988]

```

Sample: b'Petorca_Province'

Petorca Province -LRB- -RRB- is one of eight 1 province s of

the central [1 1 Chile 1] an 2 region of [1 3 Valparaíso 1] - LRB- V -RRB- . Its capital is the city of [2 2 La Ligua 2] . As a province , Petorca is a second-level 4 administrative di vision , governed by a provincial 5 governor who is appointed by the 6 president . The province comprises five 7 communes - LRB- [2 8 Spanish 2] : comunas -RRB- , each governed by a municipality consisting of an 9 alcalde and 10 municipal coun cil : The province spans an area of , the largest in the [3 Valparaíso Region 3] . According to the 2002 census , Petorca is the second least populous province in the region under [4 Isla de Pascua 4] -LRB- [5 Easter Island 5] -RRB- with a population of 70,610 . At that time , there were 50,289 people living in 11 urban area s , 20,321 people living in 12 rural area

1: 5489 Chile

2: 13439422 La_Ligua

1: 1605094 Provinces_of_Chile

[winner: 673434 Provinces_of_Burkina_Faso, pem_pos=3, final_scores=-1.1467044, final_local_scores=-1.1105661, pem=-2.634314, similarity=-14.725989]

RECALL MISS: 1605094 Provinces_of_Chile]

2: 299550 Regions_of_Chile

[winner: 2372980 Regions_of_Ivory_Coast, pem_pos=5, final_scores=-1.2125087, final_local_scores=-1.1778271, pem=-2.7300358, similarity=-16.212637]

gt: 299550 Regions_of_Chile, pem_pos=15, final_scores=-1.6181017, final_local_scores=-1.7691097, pem=-3.4150333, similarity=-30.132566]

3: 72848 Valparaíso_Region

[winner: 70566 Valparaíso, pem_pos=0, final_scores=0.30462217, final_local_scores=0.30446088, pem=-0.13340577, similarity=13.900915]

gt: 72848 Valparaíso_Region, pem_pos=1, final_scores=-0.2516601, final_local_scores=-0.33970416, pem=-3.0107968, similarity=10.325444]

4: 22432890 Administrative_divisions_of_Chile

[winner: 22432890 Administrative_divisions_of_Chile, pem_pos=2, final_scores=-1.0631931, final_local_scores=-1.1219802, pem=-1.6850263, similarity=-20.228981]

gt: 22432890 Administrative_divisions_of_Chile, pem_pos=2, final_scores=-1.0631931, final_local_scores=-1.1219802, pem=-1.6850263, similarity=-20.228981]

5: 56231 Governor

[winner: 101141 Governor_of_Maryland, pem_pos=0, final_scores=-1.1390052, final_local_scores=-1.143434, pem=0.0, similarity=-30.03264]

gt: 56231 Governor, pem_pos=4, final_scores=-1.8791767, final_local_scores=-2.0163207, pem=-3.218876, similarity=-38.576572]

6: 11885300 President_of_Chile

[winner: 6398179 The_President, pem_pos=0,
final_scores=-1.1116749, final_local_scores=-1.0320019,
pem=-0.3939714, similarity=-24.564804]
RECALL MISS: 11885300 President_of_Chile]

7: 3060770 Communes_of_Chile
[winner: 21844176 Communes_of_Cameroon, pem_pos=3,
final_scores=-1.164748, final_local_scores=-1.113277,
pem=-2.7258513, similarity=-14.309092]
gt: 3060770 Communes_of_Chile, pem_pos=4,
final_scores=-1.682521, final_local_scores=-1.8638597,
pem=-2.730428, similarity=-36.683083]

8: 26825 Spanish_language
[winner: 26667 Spain, pem_pos=1,
final_scores=0.052717045, final_local_scores=0.02693969,
pem=-1.2425895, similarity=11.651042]
gt: 26825 Spanish_language, pem_pos=0,
final_scores=-0.05325967, final_local_scores=-0.062736236,
pem=-0.6062411, similarity=5.514372]

9: 597450 Alcalde
[winner: 597450 Alcalde, pem_pos=0,
final_scores=-1.2007885, final_local_scores=-1.3040166,
pem=-0.06003342, similarity=-34.498272]
gt: 597450 Alcalde, pem_pos=0, final_scores=-1.2007885,
final_local_scores=-1.3040166, pem=-0.06003342,
similarity=-34.498272]

10: 3532493 Municipal_council
[winner: 19038 Municipality, pem_pos=0,
final_scores=-1.1260659, final_local_scores=-1.2111795,
pem=-0.19613224, similarity=-30.987703]
RECALL MISS: 3532493 Municipal_council]
[winner: 12666225 Executive_Council_of_Lower_Canada,
pem_pos=5, final_scores=-1.4921534,
final_local_scores=-1.4982533, pem=-3.009209,
similarity=-24.25661]
RECALL MISS: 3532493 Municipal_council]

11: 764593 Urban_area
[winner: 1269662 Urban_areas_in_Sweden, pem_pos=2,
final_scores=-1.023314, final_local_scores=-0.9638043,
pem=-1.760042, similarity=-15.100748]
gt: 764593 Urban_area, pem_pos=0,
final_scores=-1.2235923, final_local_scores=-1.3082459,
pem=-0.15587038, similarity=-34.103306]

12: 212614 Rural_area
[winner: 10420362 Area_(journal), pem_pos=6,
final_scores=-1.0864434, final_local_scores=-0.95683175,
pem=-2.3005102, similarity=-11.953523]
RECALL MISS: 212614 Rural_area]
[winner: 1256736 Rural_municipality, pem_pos=3,
final_scores=-1.1754014, final_local_scores=-1.1593075,
pem=-1.6088766, similarity=-21.757013]
gt: 212614 Rural_area, pem_pos=0,
final_scores=-1.3430178, final_local_scores=-1.454982,

```

pem=0.0, similarity=-39.329857 ]
1: 5489 Chile
    [pem_pos=0, final_scores=0.62824064,
final_local_scores=0.6265066, pem=-0.27404457,
similarity=24.276217]
2: 13439422 La_Ligua
    [pem_pos=0, final_scores=0.05548452,
final_local_scores=0.124412, pem=0.0, similarity=7.802415]
1: 70566 Valparaíso
    [pem_pos=0, final_scores=0.30462217,
final_local_scores=0.30446088, pem=-0.13340577,
similarity=13.900915]
2: 26667 Spain
    [pem_pos=1, final_scores=0.052717045,
final_local_scores=0.02693969, pem=-1.2425895,
similarity=11.651042]
3: 72848 Valparaíso_Region
    [pem_pos=0, final_scores=0.2840014,
final_local_scores=0.27261272, pem=0.0, similarity=12.22502]
4: 53418 Easter_Island
    [pem_pos=0, final_scores=0.14426313,
final_local_scores=0.11804071, pem=0.0, similarity=7.612283]
5: 53418 Easter_Island
    [pem_pos=0, final_scores=0.055762425,
final_local_scores=0.021367118, pem=0.0, similarity=4.72735]

```

Sample: b'Jack_Kennedy'

_ -LRB- hurler -RRB- Jack Kennedy -LRB- born 1982 -RRB- is an [1 1 Irish 1] 2 hurler who played as a left wing-forward for the [1 1 Waterford 1] senior team . [2 Kennedy 2] joined the team during the and was a regular member of the starting fifteen until his retirement during the . During that time he won two 3 Munster winners ' medals and one 4 National League winners ' medal . [3 Kennedy 3] ended up as an [4 All-Ireland 4] runner-up on one occasion . At club level Kennedy is a one-time 5 county club championship medalist with 6 De La Salle . He previously played with [2 2 Ballyduff Lower 2] . == Club == Kennedy began his club hurling career with [3 3 Ballyduff Lower 3] , however , he enjoyed little success with the club . In 2012 [5 Kennedy 5] transferred from Ballyduff Lower to the 7 De La Salle club in [6 Waterford 6] city . It was a successful as he won a 8 championship medal at the end of the year as De La Salle trounced 9 Dungarvan by 1-21 to 0-12 in the decider . [7 Kennedy 7] made his senior debut for [8 Waterford 8] in a National League game against [4 4 Dublin 4] in 2003 , however , he was unused substitute during the championship . By 2004 Kennedy made his championship debut in the Munster final against [5 5 Cork 5] . Described as the game that had everything [9 Waterford 9] beat [10 Cork 10] by 3-16 to 1-21 to win one of the greatest games of hurling ever played . It was Kennedy 's first 10 Munster medal . In 2007 Kennedy added a [6 6 National Hurling League 6] medal to his

collection when [11 Waterford 11] defeated [12 Kilkenny 12] by 0-20 to 0-18 in the final . He later claimed a second Munster medal as [13 Waterford 13] defeated [14 11 Limerick 14] by 3-17 to 1-14 in the provincial decider . While [15 Waterford 15] were viewed as possibly going on and winning the [16 All-Ireland 16] title for the first time in almost half a century , [17 Limerick 17] ambushed [18 Kennedy 18] 's side in the [19 All-Ireland 19] semi-final . 2008 began poorly for [20 Waterford 20] as the team lost their opening game to [7 7 Clare 7] as well as their manager [8 8 Justin Mc Carthy 8] . In spite of this poor start [21 Kennedy 21] 's side reached the [22 All-Ireland 22] final for the first time in forty-five years . [23 Kilkenny 23] provided the opposition and went on to trounce [24 Waterford 24] by 3-30 to 1-13 to claim a third [25 All-Ireland 25] title in-a-row . In the build up to [26 Waterford 26] 's [27 All-Ireland 27] quarter-final meeting with [9 9 Galway 9]

1: 2802645 Waterford_GAA
 2: 16264591 Ballyduff_Lower_GAA
 3: 16264591 Ballyduff_Lower_GAA
 4: 2759665 Dublin_GAA
 5: 32532531 Cork_GAA
 6: 2802698 National_Hurling_League
 7: 2794983 Clare_GAA
 8: 6184701 Justin_McCarthy_(hurler)
 9: 2810853 Galway_GAA

1: 775859 Irish_people
 [winner: 14560 Republic_of_Ireland, pem_pos=3, final_scores=0.40148467, final_local_scores=0.45202333, pem=-2.2576945, similarity=29.856684]
 gt: 775859 Irish_people, pem_pos=0, final_scores=-0.37719223, final_local_scores=-0.40140852, pem=-1.1828228, similarity=-1.4567434]
 2: 73105 Hurling
 [winner: 73105 Hurling, pem_pos=0, final_scores=-1.8239998, final_local_scores=-2.035423, pem=-0.061817452, similarity=-56.3152]
 gt: 73105 Hurling, pem_pos=0, final_scores=-1.8239998, final_local_scores=-2.035423, pem=-0.061817452, similarity=-56.3152]
 3: 3524830 Munster_Senior_Hurling_Championship
 [winner: 985954 Munster_Rugby, pem_pos=3, final_scores=-0.07145114, final_local_scores=-0.10070959, pem=-1.883883, similarity=11.329186]
 gt: 3524830 Munster_Senior_Hurling_Championship, pem_pos=5, final_scores=-0.15430081, final_local_scores=-0.23774883, pem=-2.548397, similarity=10.853393]
 4: 2802698 National_Hurling_League
 [winner: 64290 National_League, pem_pos=0, final_scores=-0.034419477, final_local_scores=0.012542337, pem=-0.07431644, similarity=4.8681455]

gt: 2802698 National_Hurling_League, pem_pos=8,
 final_scores=-0.363703, final_local_scores=-0.47344548,
 pem=-3.3032966, similarity=7.924996]
 5: 7085386 Waterford_Senior_Hurling_Championship
 [winner: 17212389 Kerry_Club_Football_Championship,
 pem_pos=0, final_scores=-0.65046984,
 final_local_scores=-0.66315407, pem=0.0,
 similarity=-15.700128]
 gt: 7085386 Waterford_Senior_Hurling_Championship,
 pem_pos=4, final_scores=-1.1336935,
 final_local_scores=-1.269186, pem=-2.5144656,
 similarity=-20.111277]
 6: 16096115 De_La_Salle_GAA
 [winner: 16096115 De_La_Salle_GAA, pem_pos=0,
 final_scores=-0.17882502, final_local_scores=-0.22718313,
 pem=-1.1841658, similarity=3.749793]
 gt: 16096115 De_La_Salle_GAA, pem_pos=0,
 final_scores=-0.17882502, final_local_scores=-0.22718313,
 pem=-1.1841658, similarity=3.749793]
 7: 16096115 De_La_Salle_GAA
 [winner: 16096115 De_La_Salle_GAA, pem_pos=0,
 final_scores=-0.17605934, final_local_scores=-0.22416204,
 pem=-1.1841658, similarity=3.8399484]
 gt: 16096115 De_La_Salle_GAA, pem_pos=0,
 final_scores=-0.17605934, final_local_scores=-0.22416204,
 pem=-1.1841658, similarity=3.8399484]
 8: 7085386 Waterford_Senior_Hurling_Championship
 [winner: 18178263 A_Championship, pem_pos=0,
 final_scores=-0.82980335, final_local_scores=-0.8932202,
 pem=0.0, similarity=-22.56576]
 RECALL MISS: 7085386
 Waterford_Senior_Hurling_Championship]
 9: 16194510 Dungarvan_GAA
 [winner: 704150 Dungarvan, pem_pos=0,
 final_scores=-0.111156926, final_local_scores=-0.0981158,
 pem=-0.04066234, similarity=1.38287]
 gt: 16194510 Dungarvan_GAA, pem_pos=1,
 final_scores=-0.13194942, final_local_scores=-0.116093576,
 pem=-1.0986123, similarity=6.5996733]
 10: 3524830 Munster_Senior_Hurling_Championship
 [winner: 985954 Munster_Rugby, pem_pos=3,
 final_scores=-0.083006635, final_local_scores=-0.11361879,
 pem=-1.883883, similarity=10.943951]
 gt: 3524830 Munster_Senior_Hurling_Championship,
 pem_pos=5, final_scores=-0.1806464,
 final_local_scores=-0.2662456, pem=-2.548397,
 similarity=10.002993]
 11: 2810846 Limerick_GAA
 [winner: 1234583 Limerick_F.C., pem_pos=4,
 final_scores=0.1880234, final_local_scores=0.14794546,
 pem=-2.6802762, similarity=23.08046]
 gt: 2810846 Limerick_GAA, pem_pos=1,

```

final_scores=0.11558856, final_local_scores=0.043862134,
pem=-1.520826, similarity=13.669135 ]
1: 2802645 Waterford_GAA
    [pem_pos=1, final_scores=0.07091065,
final_local_scores=-0.01047188, pem=-1.8432962,
similarity=13.801342]
2: 16264591 Ballyduff_Lower_GAA
    [pem_pos=0, final_scores=0.07832126,
final_local_scores=0.099150896, pem=0.0, similarity=7.0485735]
3: 16264591 Ballyduff_Lower_GAA
    [pem_pos=0, final_scores=0.07613604,
final_local_scores=0.09676385, pem=0.0, similarity=6.9773393]
4: 2759665 Dublin_GAA
    [pem_pos=1, final_scores=0.14377417,
final_local_scores=0.06444916, pem=-2.538983,
similarity=19.820389]
5: 32532531 Cork_GAA
    [pem_pos=1, final_scores=0.18353824,
final_local_scores=0.1130676, pem=-1.0852067,
similarity=13.365398]
6: 2802698 National_Hurling_League
    [pem_pos=0, final_scores=0.123091325,
final_local_scores=0.060592845, pem=0.0, similarity=5.8979244]
7: 2794983 Clare_GAA
    [pem_pos=0, final_scores=0.35119975,
final_local_scores=0.30044842, pem=-1.1644745,
similarity=19.388287]
8: 6184701 Justin_McCarthy_(hurler)
    [pem_pos=0, final_scores=0.34466404,
final_local_scores=0.3348368, pem=-0.59279925,
similarity=17.305645]
9: 2810853 Galway_GAA
    [pem_pos=1, final_scores=0.14986111,
final_local_scores=0.07625905, pem=-1.1811876,
similarity=12.788917]
1: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.40148467,
final_local_scores=0.45202333, pem=-2.2576945,
similarity=29.856684]
2: 11061816 Jack_Kennedy_(hurler)
    [pem_pos=4, final_scores=0.054251805,
final_local_scores=0.04220307, pem=-1.3938092,
similarity=12.928888]
3: 24271128 Jack_Kennedy_(footballer)
    [pem_pos=2, final_scores=0.17560767,
final_local_scores=0.17644224, pem=-1.3724864,
similarity=16.818895]
4: 780802 All-Ireland_Senior_Hurling_Championship
    [pem_pos=0, final_scores=0.20359631,
final_local_scores=0.15189284, pem=-0.83555937,
similarity=13.166401]
5: 24271128 Jack_Kennedy_(footballer)

```

[pem_pos=2, final_scores=0.0868447,
final_local_scores=0.07948217, pem=-1.3724864,
similarity=13.925413]
6: 92666 Waterford
[pem_pos=0, final_scores=0.15730567,
final_local_scores=0.17907384, pem=-0.8392663,
similarity=13.997694]
7: 24271128 Jack_Kennedy_(footballer)
[pem_pos=2, final_scores=0.045070216,
final_local_scores=0.03467104, pem=-1.3724864,
similarity=12.58816]
8: 2802645 Waterford_GAA
[pem_pos=1, final_scores=0.2211536,
final_local_scores=0.15476787, pem=-1.8432962,
similarity=18.732426]
9: 2802645 Waterford_GAA
[pem_pos=1, final_scores=0.25880778,
final_local_scores=0.19589931, pem=-1.8432962,
similarity=19.95987]
10: 32532531 Cork_GAA
[pem_pos=1, final_scores=0.19429623,
final_local_scores=0.12454039, pem=-1.0852067,
similarity=13.707769]
11: 1233201 Waterford_United_F.C.
[pem_pos=12, final_scores=0.29556006,
final_local_scores=0.25815517, pem=-3.4042218,
similarity=30.306265]
12: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.21189766,
final_local_scores=0.15037334, pem=-1.0248141,
similarity=14.150251]
13: 1233201 Waterford_United_F.C.
[pem_pos=12, final_scores=0.21348242,
final_local_scores=0.16849786, pem=-3.4042218,
similarity=27.630713]
14: 1234583 Limerick_F.C.
[pem_pos=4, final_scores=0.1880234,
final_local_scores=0.14794546, pem=-2.6802762,
similarity=23.08046]
15: 2802645 Waterford_GAA
[pem_pos=1, final_scores=0.11699338,
final_local_scores=0.040988624, pem=-1.8432962,
similarity=15.337027]
16: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.046529993,
final_local_scores=-0.0230705, pem=-0.83555937,
similarity=7.945145]
17: 1234583 Limerick_F.C.
[pem_pos=4, final_scores=0.23077829,
final_local_scores=0.19507688, pem=-2.6802762,
similarity=24.486954]
18: 24271128 Jack_Kennedy_(footballer)

[pem_pos=2, final_scores=0.08387475,
final_local_scores=0.07623795, pem=-1.3724864,
similarity=13.828599]
19: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.047971323,
final_local_scores=-0.021496072, pem=-0.83555937,
similarity=7.992129]
20: 2802645 Waterford_GAA
[pem_pos=1, final_scores=0.15030502,
final_local_scores=0.077376544, pem=-1.8432962,
similarity=16.422915]
21: 24271128 Jack_Kennedy_(footballer)
[pem_pos=2, final_scores=0.14133392,
final_local_scores=0.13900337, pem=-1.3724864,
similarity=15.701645]
22: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.084797874,
final_local_scores=0.020982787, pem=-0.83555937,
similarity=9.259783]
23: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.25207454,
final_local_scores=0.19467282, pem=-1.0248141,
similarity=15.472238]
24: 2802645 Waterford_GAA
[pem_pos=1, final_scores=0.28318816,
final_local_scores=0.22311085, pem=-1.8432962,
similarity=20.771917]
25: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.09854008,
final_local_scores=0.035994068, pem=-0.83555937,
similarity=9.707749]
26: 2802645 Waterford_GAA
[pem_pos=1, final_scores=0.15263186,
final_local_scores=0.08049792, pem=-1.8432962,
similarity=16.516062]
27: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.18846504,
final_local_scores=0.1342234, pem=-0.83555937,
similarity=12.639109]

Sample: b'Doug_Insole'

Douglas John Insole 1 CBE -LRB- born 18 April 1926 , 2 Clapton , [1 1 London 1] -RRB- is a former [1 3 English 1] 4 cricketer , who played for [2 2 Cambridge University 2] , [3 3 Essex 3] and in nine [2 5 Test 2] matches for , five of them on the , where he was vice-captain to [4 4 Peter May 4] . After retiring from playing , he was prominent in cricket administration , and served as chairman of the [3 England 3] selectors and as President of the [5 5 MCC 5] . [4 Insole 4] was captain of cricket at [6 6 Cambridge University 6] and went on to 6 captain [5 Essex 5] for many years . He scored 20,113 7 first-class runs for [6 Essex 6] , the ninth highest

aggreagate for the club . He played as a 8 wicket-keeper , , and as a . He was one of the 9 Wisden Cricketers of the Year in 1956 . He was President of the [7 7 MCC 7] for the twelve months , beginning on 1 October 2006 . For his many services to cricket , [7 Insole 7] was appointed a [8 8 CBE 8] in 1979 . [8 Insole 8] was chairman of selectors for the [9 9 England 9] in the 1960s , during which time [10 10 Geoffrey Boycott 10] was dropped in 1967 , after having scored 246 not out . [9 Boycott 9] admitted to still feeling aggrieved about this , over 40 years later , during commenting on the third [10 Test 10] between and [11 England 11] at [11 11 Napier 11] on 24 March 2008 and again during the 10 Test Match Special commentary of [12 England 12] against the [12 12 West Indies 12] at [13 13 Edgbaston 13] on 7 June 2012 , where he said that [13 Insole 13] `` should have spelt his name with an A ! '' [14 Insole 14] was a first team footballer for the amateur 11 Corinthian-Casuals F.C. , and played in the 1956 [14 14 FA Amateur Cup 14] final , before they lost to [15 15 Bishop Auckland 15]

- 1: 17867 London
- 2: 1622256 Cambridge_University_Cricket_Club
- 3: 1622252 Essex_County_Cricket_Club
- 4: 1351617 Peter_May
- 5: 307442 Marylebone_Cricket_Club
- 6: 25978572 University_of_Cambridge
- 7: 307442 Marylebone_Cricket_Club
- 8: 212182 Order_of_the_British_Empire
- 9: 759125 England_cricket_team
- 10: 413130 Geoffrey_Boycott
- 11: 171974 Napier,_New_Zealand
- 12: 3379941 West_Indies_cricket_team
- 13: 2641030 Edgbaston_Cricket_Ground
- 14: 627338 FA_Amateur_Cup
- 15: 1065593 Bishop_Auckland_F.C.

1: 212182 Order_of_the_British_Empire

[winner: 212182 Order_of_the_British_Empire, pem_pos=0, final_scores=-0.9087425, final_local_scores=-0.97188175, pem=0.0, similarity=-24.913177]

gt: 212182 Order_of_the_British_Empire, pem_pos=0, final_scores=-0.9087425, final_local_scores=-0.97188175, pem=0.0, similarity=-24.913177]

2: 1112504 Upper_Clapton

[winner: 10049 Eric_Clapton, pem_pos=0, final_scores=-0.002073422, final_local_scores=0.06021495, pem=-0.63749355, similarity=9.353436]

gt: 1112504 Upper_Clapton, pem_pos=3, final_scores=-0.2066002, final_local_scores=-0.20378062, pem=-1.9001088, similarity=8.341579]

3: 18803164 English_people

[winner: 9316 England, pem_pos=1, final_scores=0.53343797, final_local_scores=0.5500398, pem=-1.3527626, similarity=27.860533]

gt: 18803164 English_people, pem_pos=2,
final_scores=-0.38983166, final_local_scores=-0.44919345,
pem=-2.2084396, similarity=2.694722]
4: 25675557 Cricket
[winner: 330091 Rhabdophoridae, pem_pos=10,
final_scores=-1.4065295, final_local_scores=-1.3633802,
pem=-2.56456, similarity=-22.649797]
gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.5760505, final_local_scores=-1.7749825,
pem=-0.07532314, similarity=-48.46969]
5: 30863 Test_cricket
[winner: 30863 Test_cricket, pem_pos=0,
final_scores=0.27387524, final_local_scores=0.23107776,
pem=-0.23950954, similarity=12.288024]
gt: 30863 Test_cricket, pem_pos=0,
final_scores=0.27387524, final_local_scores=0.23107776,
pem=-0.23950954, similarity=12.288024]
[winner: 63904 Match, pem_pos=0, final_scores=-1.8084552,
final_local_scores=-1.891185, pem=-0.17114355,
similarity=-51.416313]
gt: 30863 Test_cricket, pem_pos=12,
final_scores=-2.9692175, final_local_scores=-3.3154285,
pem=-7.2545314, similarity=-55.398018]
6: 896020 Captain_(cricket)
[winner: 3319715 Australian_Army_officer_rank_insignia,
pem_pos=11, final_scores=-1.4874706,
final_local_scores=-1.4763947, pem=-2.7026498,
similarity=-25.271418]
gt: 896020 Captain_(cricket), pem_pos=3,
final_scores=-1.7707455, final_local_scores=-2.0128045,
pem=-1.9929447, similarity=-45.138443]
7: 11040 First-class_cricket
[winner: 662844 First-class_citizen, pem_pos=3,
final_scores=-1.2423499, final_local_scores=-1.167032,
pem=-1.5935546, similarity=-22.070852]
gt: 11040 First-class_cricket, pem_pos=0,
final_scores=-1.3047087, final_local_scores=-1.4859462,
pem=-0.053819276, similarity=-39.96121]
8: 463313 Wicket-keeper
[winner: 463313 Wicket-keeper, pem_pos=0,
final_scores=-1.3178852, final_local_scores=-1.4878702,
pem=-0.004312502, similarity=-40.28785]
gt: 463313 Wicket-keeper, pem_pos=0,
final_scores=-1.3178852, final_local_scores=-1.4878702,
pem=-0.004312502, similarity=-40.28785]
9: 612064 Wisden_Cricketers_of_the_Year
[winner: 1582694 Wisden_Cricketers'_Almanack, pem_pos=0,
final_scores=-0.011110529, final_local_scores=-0.068202384,
pem=0.0, similarity=2.0544178]
gt: 612064 Wisden_Cricketers_of_the_Year, pem_pos=1,
final_scores=-0.25188887, final_local_scores=-0.33776098,
pem=-1.602925, similarity=2.7272062]

[winner: 34341 Year, pem_pos=0, final_scores=-0.44606462, final_local_scores=-0.3960143, pem=-0.14641148, similarity=-6.9319353]

RECALL MISS: 612064 Wisden_Cricketers_of_the_Year]

[winner: 2217991 Ireland_cricket_team, pem_pos=0, final_scores=-1.3968078, final_local_scores=-1.5801268, pem=0.0, similarity=-43.064423]

RECALL MISS: 612064 Wisden_Cricketers_of_the_Year]

[winner: 10447732

Results_by_riding_of_the_Canadian_federal_election,_2008, pem_pos=12, final_scores=-1.445461, final_local_scores=-1.3244274, pem=-3.6198866, similarity=-15.748334]

RECALL MISS: 612064 Wisden_Cricketers_of_the_Year]

10: 646127 Test_Match_Special

[winner: 646127 Test_Match_Special, pem_pos=0, final_scores=-0.11646299, final_local_scores=-0.1554838, pem=0.0, similarity=-0.5502339]

gt: 646127 Test_Match_Special, pem_pos=0, final_scores=-0.11646299, final_local_scores=-0.1554838, pem=0.0, similarity=-0.5502339]

11: 1070250 Corinthian-Casuals_F.C.

[winner: 1070250 Corinthian-Casuals_F.C., pem_pos=0, final_scores=-0.6399329, final_local_scores=-0.6782015, pem=0.0, similarity=-16.149172]

gt: 1070250 Corinthian-Casuals_F.C., pem_pos=0, final_scores=-0.6399329, final_local_scores=-0.6782015, pem=0.0, similarity=-16.149172]

[winner: 21376985 Francis_Chan, pem_pos=1, final_scores=-1.3919289, final_local_scores=-1.4157374, pem=-0.2876821, similarity=-36.594257]

RECALL MISS: 1070250 Corinthian-Casuals_F.C.]

1: 17867 London

[pem_pos=0, final_scores=0.5172319, final_local_scores=0.55262434, pem=-0.09086737, similarity=21.075275]

2: 1622256 Cambridge_University_Cricket_Club

[pem_pos=1, final_scores=0.24381153, final_local_scores=0.20616317, pem=-1.3585418, similarity=17.629995]

3: 1622252 Essex_County_Cricket_Club

[pem_pos=1, final_scores=0.84563357, final_local_scores=0.85075796, pem=-2.3459642, similarity=42.235744]

4: 1351617 Peter_May

[pem_pos=0, final_scores=0.48356098, final_local_scores=0.50575316, pem=0.0, similarity=19.182396]

5: 307442 Marylebone_Cricket_Club

[pem_pos=0, final_scores=0.5496802, final_local_scores=0.51602817, pem=-0.18132487, similarity=20.475096]

6: 25978572 University_of_Cambridge

[pem_pos=0, final_scores=0.30974245,
final_local_scores=0.33068138, pem=-0.12541339,
similarity=14.639923]
7: 307442 Marylebone_Cricket_Club
[pem_pos=0, final_scores=0.46353817,
final_local_scores=0.42193115, pem=-0.18132487,
similarity=17.667053]
8: 212182 Order_of_the_British_Empire
[pem_pos=0, final_scores=0.030013546,
final_local_scores=0.06255804, pem=0.0, similarity=5.9565697]
9: 759125 England_cricket_team
[pem_pos=2, final_scores=0.13400994,
final_local_scores=0.04935229, pem=-3.3054674,
similarity=23.53813]
10: 413130 Geoffrey_Boycott
[pem_pos=0, final_scores=0.6948102,
final_local_scores=0.7111901, pem=0.0, similarity=25.313044]
11: 171974 Napier,_New_Zealand
[pem_pos=0, final_scores=0.1820334,
final_local_scores=0.21906143, pem=-0.5995366,
similarity=13.887317]
12: 3379941 West_Indies_cricket_team
[pem_pos=1, final_scores=0.24563377,
final_local_scores=0.1975584, pem=-1.1020415,
similarity=15.978322]
13: 2641030 Edgbaston_Cricket_Ground
[pem_pos=1, final_scores=0.19859912,
final_local_scores=0.15177661, pem=-0.4505704,
similarity=11.069302]
14: 627338 FA_Amateur_Cup
[pem_pos=0, final_scores=0.16417833,
final_local_scores=0.16698039, pem=0.0, similarity=9.072741]
15: 1065593 Bishop_Auckland_F.C.
[pem_pos=1, final_scores=0.20219432,
final_local_scores=0.22468749, pem=-0.7688232,
similarity=14.975816]
1: 9316 England
[pem_pos=1, final_scores=0.53343797,
final_local_scores=0.5500398, pem=-1.3527626,
similarity=27.860533]
2: 30863 Test_cricket
[pem_pos=0, final_scores=0.27387524,
final_local_scores=0.23107776, pem=-0.23950954,
similarity=12.288024]
3: 759125 England_cricket_team
[pem_pos=2, final_scores=0.37771755,
final_local_scores=0.31556576, pem=-3.3054674,
similarity=31.482471]
4: 3635871 Doug_Insole
[pem_pos=1, final_scores=0.5623116,
final_local_scores=0.5715741, pem=0.0, similarity=21.146624]
5: 1622252 Essex_County_Cricket_Club

```

    [pem_pos=1, final_scores=0.7229063,
final_local_scores=0.716697, pem=-2.3459642,
similarity=38.2351]
6: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.7838064,
final_local_scores=0.7832211, pem=-2.3459642,
similarity=40.22031]
7: 3635871 Doug_Insole
    [pem_pos=1, final_scores=0.4828285,
final_local_scores=0.48475087, pem=0.0, similarity=18.555645]
8: 3635871 Doug_Insole
    [pem_pos=1, final_scores=0.5813952,
final_local_scores=0.59242, pem=0.0, similarity=21.768707]
9: 413130 Geoffrey_Boycott
    [pem_pos=0, final_scores=0.59249705,
final_local_scores=0.603825, pem=0.0, similarity=22.109055]
10: 30863 Test_cricket
    [pem_pos=0, final_scores=0.0892656,
final_local_scores=0.02941981, pem=-0.23950954,
similarity=6.2701488]
11: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.35922796,
final_local_scores=0.29536873, pem=-3.3054674,
similarity=30.879751]
12: 9904 England_national_football_team
    [pem_pos=1, final_scores=0.16591795,
final_local_scores=0.14580524, pem=-2.776464,
similarity=23.539675]
13: 3635871 Doug_Insole
    [pem_pos=1, final_scores=0.2788555,
final_local_scores=0.26078025, pem=0.0, similarity=11.871916]
14: 3635871 Doug_Insole
    [pem_pos=1, final_scores=0.3457443,
final_local_scores=0.3338461, pem=0.0, similarity=14.052346]

```

Sample: b'Texas_State_Highway_110'

State Highway 110 or SH 110 is a 1 state highway in the [1 2 U.S. 1] state of [1 1 Texas 1] that runs from [2 2 Grand Saline 2] to [3 3 Rusk 3] . SH 110 begins at an intersection with and in downtown 3 Rusk and leaves the courthouse square north with US 84 , crossing on its way to a split on the northeast side of Rusk where US 84 goes off east and SH 110 turns north , out of town . The road passes 4 Ponta and [4 4 New Summerfield 4] before crossing the county line into [5 5 Smith County 5] as it enters [6 6 Troup 6] . After a brief downtown multiplex with 5 SH 135 , SH 110 leaves [2 Troup 2] going northwest through 6 Whitehouse on its way to [7 7 Tyler 7] . SH 110 swings around the downtown area to the south and west by way of multiplexes with 7 SH 64 , 8 SH 155 , and US 69 . After splitting with US 69 , SH 110 leaves [3 Tyler 3] toward the northwest . After several miles , SH 110 crosses 9 I-20 south of 10 Garden Valley . Just after Garden Valley , the

road multiplexes with FM 16 and goes west into [8 8 Van Zandt County 8] . SH 110 splits with FM 16 in 11 Van and turns back north . Approaching [9 9 Grand Saline 9] , SH 110 merges with FM 17 and travels the remaining several blocks multiplexed until it reaches 12 US 80 , the northern terminus of SH 110 . SH 110 was originally numbered in 1926 on a route from SH 26 north of [4 Nacogdoches 4] northwest to [10 10 Sacul 10] . This was a renumbering of a former portion of SH 40 , which was rerouted southward from in the 1926 plans . By 1933 , the route had been extended to include a route from [5 Rusk 5] northward to [6 Tyler 6] , along with its original route and a planned connection at [7 New Summerfield 7] . By 1936 , the original portion had been renumbered as SH 206 , leaving the routing from Rusk to [8 Tyler 8] , and an extension northwest to [9 Grand Saline 9] , completing its current routing . A section northward to 13 Alba was added in 1940 , but was transferred to FM 17

1: 29810 Texas

2: 136742 Grand_Saline,_Texas

3: 135672 Rusk,_Texas

4: 135671 New_Summerfield,_Texas

5: 91449 Smith_County,_Texas

6: 151134 Troup,_Texas

7: 136600 Tyler,_Texas

8: 91417 Van_Zandt_County,_Texas

9: 136742 Grand_Saline,_Texas

10: 18723249 Sacul,_Texas

1: 332996 State_highway

[winner: 332996 State_highway, pem_pos=0,
final_scores=-1.0721023, final_local_scores=-1.1353811,
pem=-0.024095139, similarity=-29.661291]

gt: 332996 State_highway, pem_pos=0,
final_scores=-1.0721023, final_local_scores=-1.1353811,
pem=-0.024095139, similarity=-29.661291]

2: 18618239 U.S._state

[winner: 3434750 United_States, pem_pos=0,
final_scores=0.4047094, final_local_scores=0.43041104,
pem=-0.14075686, similarity=17.699495]

gt: 18618239 U.S._state, pem_pos=6,
final_scores=-0.37546283, final_local_scores=-0.43239415,
pem=-3.9799929, similarity=12.830029]

[winner: 229379 States_of_Brazil, pem_pos=0,
final_scores=-1.797673, final_local_scores=-1.8437891,
pem=0.0, similarity=-50.93263]

RECALL MISS: 18618239 U.S._state]

3: 135672 Rusk,_Texas

[winner: 135672 Rusk,_Texas, pem_pos=1,
final_scores=-0.011544734, final_local_scores=-0.07181615,
pem=-1.5206742, similarity=10.216239]

gt: 135672 Rusk,_Texas, pem_pos=1,
final_scores=-0.011544734, final_local_scores=-0.07181615,
pem=-1.5206742, similarity=10.216239]

4: 28721402 Ponta,_Texas
[winner: 23324667 Victor_Ponta, pem_pos=0,
final_scores=-0.12649387, final_local_scores=-0.0093649775,
pem=0.0, similarity=3.8102431|
gt: 28721402 Ponta,_Texas, pem_pos=1,
final_scores=-0.39139313, final_local_scores=-0.3494002,
pem=-0.66782933, similarity=-2.7053258]
5: 7444940 Texas_State_Highway_135
[winner: 7444940 Texas_State_Highway_135, pem_pos=0,
final_scores=-0.49040765, final_local_scores=-0.473132,
pem=0.0, similarity=-10.02949|
gt: 7444940 Texas_State_Highway_135, pem_pos=0,
final_scores=-0.49040765, final_local_scores=-0.473132,
pem=0.0, similarity=-10.02949]
6: 136601 Whitehouse,_Texas
[winner: 125374 Whitehouse_Station,_New_Jersey,
pem_pos=5, final_scores=-0.07375117,
final_local_scores=-0.057837784, pem=-2.7896574,
similarity=17.53431|
gt: 136601 Whitehouse,_Texas, pem_pos=2,
final_scores=-0.101425216, final_local_scores=-0.13331062,
pem=-1.7904873, similarity=9.848407]
7: 3113832 Texas_State_Highway_64
[winner: 3113832 Texas_State_Highway_64, pem_pos=0,
final_scores=-0.1647493, final_local_scores=-0.13638294,
pem=0.0, similarity=0.019774094|
gt: 3113832 Texas_State_Highway_64, pem_pos=0,
final_scores=-0.1647493, final_local_scores=-0.13638294,
pem=0.0, similarity=0.019774094]
8: 3113772 Texas_State_Highway_155
[winner: 3113772 Texas_State_Highway_155, pem_pos=0,
final_scores=-0.27824363, final_local_scores=-0.25492668,
pem=0.0, similarity=-3.517808|
gt: 3113772 Texas_State_Highway_155, pem_pos=0,
final_scores=-0.27824363, final_local_scores=-0.25492668,
pem=0.0, similarity=-3.517808]
9: 10021978 Interstate_20_in_Texas
[winner: 88879 Interstate_20, pem_pos=0,
final_scores=-0.028636426, final_local_scores=-0.04443448,
pem=-0.54574674, similarity=5.731555|
gt: 10021978 Interstate_20_in_Texas, pem_pos=2,
final_scores=-0.21723151, final_local_scores=-0.21248627,
pem=-1.4031014, similarity=5.3789825]
10: 3118914 Garden_Valley,_Texas
[winner: 22720391
Garden_Valley,_El_Dorado_County,_California, pem_pos=0,
final_scores=-0.031376854, final_local_scores=0.07489307,
pem=-0.44992328, similarity=8.771424|
gt: 3118914 Garden_Valley,_Texas, pem_pos=3,
final_scores=-0.043864354, final_local_scores=0.04262255,
pem=-1.311539, similarity=12.494008]
11: 136743 Van,_Texas

[winner: 63679 Van, pem_pos=0, final_scores=-0.1914999,
final_local_scores=-0.12732193, pem=-0.6731859,
similarity=3.9510622]
gt: 136743 Van, Texas, pem_pos=6,
final_scores=-0.578357, final_local_scores=-0.65080464,
pem=-3.768375, similarity=5.1614122]
12: 11717904 U.S. Route 80 in Texas
[winner: 428913 U.S. Route 80, pem_pos=0,
final_scores=-0.14223707, final_local_scores=-0.15069422,
pem=0.0, similarity=-0.4073032]
gt: 11717904 U.S. Route 80 in Texas, pem_pos=1,
final_scores=-0.7950237, final_local_scores=-0.8323015,
pem=-3.5294518, similarity=-1.5541147]
13: 151233 Alba, Texas
[winner: 961156 Alba, Piedmont, pem_pos=1,
final_scores=-0.2927032, final_local_scores=-0.22715679,
pem=-1.4856683, similarity=5.390197]
gt: 151233 Alba, Texas, pem_pos=12,
final_scores=-0.55399394, final_local_scores=-0.62008107,
pem=-3.233408, similarity=3.1690326]
1: 29810 Texas
[pem_pos=0, final_scores=0.5151369,
final_local_scores=0.50166786, pem=-0.16191302,
similarity=19.94099]
2: 136742 Grand Saline, Texas
[pem_pos=0, final_scores=0.5183444,
final_local_scores=0.5268479, pem=0.0, similarity=19.811905]
3: 135672 Rusk, Texas
[pem_pos=1, final_scores=0.0771739,
final_local_scores=0.03219536, pem=-1.5206742,
similarity=13.320149]
4: 135671 New Summerfield, Texas
[pem_pos=0, final_scores=0.14854603,
final_local_scores=0.15378016, pem=0.0, similarity=8.67882]
5: 91449 Smith County, Texas
[pem_pos=0, final_scores=0.3507939,
final_local_scores=0.33108482, pem=-0.2291369,
similarity=15.216026]
6: 151134 Troup, Texas
[pem_pos=1, final_scores=0.08860256,
final_local_scores=0.069320396, pem=-0.9246328,
similarity=11.186668]
7: 136600 Tyler, Texas
[pem_pos=0, final_scores=0.13704513,
final_local_scores=0.07818557, pem=-1.0000982,
similarity=11.861616]
8: 91417 Van Zandt County, Texas
[pem_pos=0, final_scores=0.19107996,
final_local_scores=0.16845977, pem=0.0, similarity=9.116889]
9: 136742 Grand Saline, Texas
[pem_pos=0, final_scores=0.43622434,
final_local_scores=0.43714422, pem=0.0, similarity=17.134968]

10: 18723249 Sacul,_Texas
 [pem_pos=0, final_scores=0.058091924,
 final_local_scores=0.10435915, pem=0.0, similarity=7.203998]
 1: 3434750 United_States
 [pem_pos=0, final_scores=0.4047094,
 final_local_scores=0.43041104, pem=-0.14075686,
 similarity=17.699495]
 2: 151134 Troup,_Texas
 [pem_pos=1, final_scores=0.050835148,
 final_local_scores=0.029234275, pem=-0.9246328,
 similarity=9.990418]
 3: 136600 Tyler,_Texas
 [pem_pos=0, final_scores=0.20272534,
 final_local_scores=0.14993125, pem=-1.0000982,
 similarity=14.00265]
 4: 136419 Nacogdoches,_Texas
 [pem_pos=0, final_scores=0.25611168,
 final_local_scores=0.23002332, pem=0.0, similarity=10.954068]
 5: 135672 Rusk,_Texas
 [pem_pos=1, final_scores=0.0398338,
 final_local_scores=-0.01569289, pem=-1.5206742,
 similarity=11.891068]
 6: 136600 Tyler,_Texas
 [pem_pos=0, final_scores=0.29301393,
 final_local_scores=0.2485578, pem=-1.0000982,
 similarity=16.945864]
 7: 135671 New_Summerfield,_Texas
 [pem_pos=0, final_scores=0.18242185,
 final_local_scores=0.19078436, pem=0.0, similarity=9.783099]
 8: 136600 Tyler,_Texas
 [pem_pos=0, final_scores=0.1985219,
 final_local_scores=0.14533961, pem=-1.0000982,
 similarity=13.865626]
 9: 136742 Grand_Saline,_Texas
 [pem_pos=0, final_scores=0.4831336,
 final_local_scores=0.4883855, pem=0.0, similarity=18.66411]

Sample: b'Love'

's _ Welcome_at_Bolsover ` Love 's Welcome at [1 Bolsover 1]
 ' -LRB- alternative archaic spelling , ` Balsover ' -RRB- is
 the final 1 masque composed by [1 1 Ben Jonson 1] . It was
 performed on 30 July 1634 , three years before the poet 's
 death , and published in 2 1641 . The masque was not produced
 by the [2 3 Stuart Court 2] in one of the royal palaces
 around [3 London 3] , as many of [4 Jonson 4] 's notable
 early masques were . Rather it was staged by 4 William Cavend
 ish , 1st Duke of Newcastle -LRB- at the time , he was the [5
 Earl of Newcastle 5] -RRB- at [2 2 Bolsover Castle 2] in [3 3
 Derbyshire 3] , in honor of King [4 4 Charles I 4] and Queen [5
 5 Henrietta Maria 5] . [6 Newcastle 6] had put on a [7
 Jonson 7] masque for his royal visitors at [6 6 Welbeck 6] in
 [7 7 Nottinghamshire 7] the year before : 5 The [8 King 8] 's

Entertainment at Welbeck , performed on 21 May 6 1633 . It was such a success that the [9 King 9] requested another on his 1634 royal progress . The [10 Duke 10] spent between # 14,000 and # 15,000 on staging the Bolsover masque and providing for his royal guests and their attendants , which was more than double the # 4,000 to # 5,000 he 'd spent for the Welbeck entertainment the previous year . -LRB- 7 The Masque of Beauty , one of [11 Jonson 11] 's early Court masques , has cost # 4,000 to stage in 8 1608 , and was considered exorbitantly expensive at the time . -RRB- In Love 's Welcome , [12 Jonson 12] continued the mockery of [8 8 Inigo Jones 8] that he 'd practiced for two decades , starting -LRB- 9 1614 -RRB- and continuing through 10 The Masque of Augurs -LRB- 11 1622 -RRB- , 12 Neptune 's Triumph for the Return of Albion -LRB- 13 1624 -RRB- , 14 The Staple of News -LRB- 15 1626 -RRB- , and 16 A Tale of a Tub -LRB- 17 1633 -RRB- . In this masque , [13 Jones 13] is `` Colonel Iniquo Vitruvius . '' The masque was staged in what was called the `` little castle '' at [14 Bolsover 14] , a then-recent -LRB- Jacobean -RRB- construction . The pillared hall there was furnished with five brilliantly-colored paintings on the theme of The Senses ; [15 Jonson 15] alludes to the paintings in his text . The show was described by local witnesses as `` stupendous , '' more than adequate to establish [16 Newcastle 16] 's reputation as the greatest `` prince ... in all the northern quarter '' of the kingdom . Perhaps the most visually striking element in the masque lay in the two Cupids , Eros -LRB- Love -RRB- and Anteros -LRB- Love Returned -RRB- , who descended `` from the clouds '' bearing fronds of palms . The masque was published in 1641 in the 18 s econd folio collection of [17 Jonson 17] 's works

- 1: 48261 Ben_Jonson
- 2: 143019 Bolsover_Castle
- 3: 71070 Derbyshire
- 4: 7426 Charles_I_of_England
- 5: 20949403 Henrietta_Maria_of_France
- 6: 7598798 Welbeck_Abbey
- 7: 68049 Nottinghamshire
- 8: 165453 Inigo_Jones

1: 74909 Masque

[winner: 74909 Masque, pem_pos=0, final_scores=-1.1426446, final_local_scores=-1.2895808, pem=-0.013208394, similarity=-34.32212|
gt: 74909 Masque, pem_pos=0, final_scores=-1.1426446, final_local_scores=-1.2895808, pem=-0.013208394, similarity=-34.32212]

2: 226378 1641_in_literature

[winner: 35202 1641, pem_pos=0, final_scores=-0.8521693, final_local_scores=-0.7394999, pem=-0.3547, similarity=-16.049526|

gt: 226378 1641_in_literature, pem_pos=1, final_scores=-1.0167688, final_local_scores=-1.0443361,

pem=-1.328699, similarity=-19.849688]
3: 253174 House_of_Stuart
[winner: 253174 House_of_Stuart, pem_pos=0,
final_scores=0.27372956, final_local_scores=0.24285248,
pem=0.0, similarity=11.336915|
gt: 253174 House_of_Stuart, pem_pos=0,
final_scores=0.27372956, final_local_scores=0.24285248,
pem=0.0, similarity=11.336915]
4: 194444 William_Cavendish,_1st_Duke_of_Newcastle
[winner: 308985 Thomas_Pelham-
Holles,_1st_Duke_of_Newcastle, pem_pos=0,
final_scores=-0.048775777, final_local_scores=-0.09443374,
pem=-0.095273435, similarity=1.7897335|
gt: 194444 William_Cavendish,_1st_Duke_of_Newcastle,
pem_pos=3, final_scores=-0.25316274,
final_local_scores=-0.34572342, pem=-1.5616105,
similarity=2.2649174]
[winner: 233785
William_Cavendish,_1st_Duke_of_Devonshire, pem_pos=0,
final_scores=-0.09220235, final_local_scores=-0.10568802,
pem=-0.839545, similarity=5.501348|
gt: 194444 William_Cavendish,_1st_Duke_of_Newcastle,
pem_pos=4, final_scores=-0.16850859,
final_local_scores=-0.25325164, pem=-2.0115516,
similarity=7.4713087]
[winner: 695423 Cephalosporin, pem_pos=6,
final_scores=-1.4256345, final_local_scores=-1.3929329,
pem=-3.8507495, similarity=-16.537209|
RECALL MISS: 194444
William_Cavendish,_1st_Duke_of_Newcastle]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7339002, final_local_scores=-1.7363298,
pem=-4.6728287, similarity=-22.314264|
RECALL MISS: 194444
William_Cavendish,_1st_Duke_of_Newcastle]
5: 10099230 The_King's_Entertainment_at_Welbeck
[winner: 7426 Charles_I_of_England, pem_pos=0,
final_scores=0.14465435, final_local_scores=0.07848567,
pem=0.0, similarity=6.431882|
RECALL MISS: 10099230
The_King's_Entertainment_at_Welbeck]
[winner: 7598798 Welbeck_Abbey, pem_pos=2,
final_scores=0.024734154, final_local_scores=0.006502673,
pem=-0.9964024, similarity=9.702356|
RECALL MISS: 10099230
The_King's_Entertainment_at_Welbeck]
[winner: 9262 Entertainment, pem_pos=0,
final_scores=-0.41472092, final_local_scores=-0.3683853,
pem=-0.77327526, similarity=-2.6984475|
RECALL MISS: 10099230
The_King's_Entertainment_at_Welbeck]
[winner: 187801 Shopping_cart, pem_pos=8,

```

final_scores=-2.4464176, final_local_scores=-2.5349922,
pem=-3.4067376, similarity=-53.033142|
    RECALL MISS: 10099230
The_King's_Entertainment_at_Welbeck]
6: 226386 1633_in_literature
    [winner: 3731025 1633_in_art, pem_pos=3,
final_scores=-0.9132906, final_local_scores=-0.7480189,
pem=-1.5059526, similarity=-10.04306|
    gt: 226386 1633_in_literature, pem_pos=1,
final_scores=-1.213608, final_local_scores=-1.3034892,
pem=-1.2691121, similarity=-27.907375 ]
7: 9026957 The_Masque_of_Beauty
    [winner: 9026957 The_Masque_of_Beauty, pem_pos=0,
final_scores=-0.034581304, final_local_scores=0.02359879,
pem=0.0, similarity=4.7939477|
    gt: 9026957 The_Masque_of_Beauty, pem_pos=0,
final_scores=-0.034581304, final_local_scores=0.02359879,
pem=0.0, similarity=4.7939477 ]
8: 226487 1608_in_literature
    [winner: 10637925 1608_in_art, pem_pos=2,
final_scores=-0.657659, final_local_scores=-0.45649898,
pem=-1.2822334, similarity=-2.5601428|
    gt: 226487 1608_in_literature, pem_pos=1,
final_scores=-0.7834368, final_local_scores=-0.80293804,
pem=-1.2046945, similarity=-13.320245 ]
9: 226409 1614_in_literature
    [winner: 35102 1614, pem_pos=0, final_scores=-0.8655295,
final_local_scores=-0.72458917, pem=-0.23957103,
similarity=-16.230648|
    gt: 226409 1614_in_literature, pem_pos=1,
final_scores=-1.0023019, final_local_scores=-1.0281125,
pem=-1.2580211, similarity=-19.7499 ]
10: 9967127 The_Masque_of_Augurs
    [winner: 9967127 The_Masque_of_Augurs, pem_pos=0,
final_scores=-0.04688023, final_local_scores=-0.00669612,
pem=0.0, similarity=3.8898876|
    gt: 9967127 The_Masque_of_Augurs, pem_pos=0,
final_scores=-0.04688023, final_local_scores=-0.00669612,
pem=0.0, similarity=3.8898876 ]
11: 226398 1622_in_literature
    [winner: 34973 1622, pem_pos=0, final_scores=-0.7356359,
final_local_scores=-0.5865183, pem=-0.19878776,
similarity=-12.332124|
    gt: 226398 1622_in_literature, pem_pos=1,
final_scores=-0.90779907, final_local_scores=-0.9222072,
pem=-1.2010186, similarity=-16.899462 ]
12: 9747851 Neptune's_Triumph_for_the_Return_of_Albion
    [winner: 19003265 Neptune, pem_pos=0,
final_scores=-0.32136688, final_local_scores=-0.2391293,
pem=-0.3902986, similarity=-0.9238777|
    RECALL MISS: 9747851
Neptune's_Triumph_for_the_Return_of_Albion]

```

[winner: 18951002 Albion, pem_pos=0,
final_scores=-0.35042524, final_local_scores=-0.3672678,
pem=-1.3786598, similarity=0.62707645]
RECALL MISS: 9747851
Neptune's_Triumph_for_the_Return_of_Albion]
[winner: 365282 Triumph_Motor_Company, pem_pos=0,
final_scores=-0.37285078, final_local_scores=-0.32311398,
pem=-1.2883272, similarity=1.4534715]
RECALL MISS: 9747851
Neptune's_Triumph_for_the_Return_of_Albion]
[winner: 2089765 Nostoi, pem_pos=17,
final_scores=-0.70630807, final_local_scores=-0.6556572,
pem=-2.8903718, similarity=0.24188736]
RECALL MISS: 9747851
Neptune's_Triumph_for_the_Return_of_Albion]
[winner: 3349493 Male-
female_income_disparity_in_the_United_States, pem_pos=0,
final_scores=-1.2074959, final_local_scores=-1.2027745,
pem=-0.12731102, similarity=-31.11114]
RECALL MISS: 9747851
Neptune's_Triumph_for_the_Return_of_Albion]
[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.240289, final_local_scores=-2.3098278,
pem=-3.4067376, similarity=-46.31379]
RECALL MISS: 9747851
Neptune's_Triumph_for_the_Return_of_Albion]
13: 226396 1624_in_literature
[winner: 10441769 1624_in_art, pem_pos=2,
final_scores=-0.8459905, final_local_scores=-0.6454731,
pem=-0.98371303, similarity=-9.822907]
gt: 226396 1624_in_literature, pem_pos=1,
final_scores=-0.9931, final_local_scores=-1.0009911,
pem=-0.95662045, similarity=-20.57961]
14: 9010672 The_Staple_of_News
[winner: 9010672 The_Staple_of_News, pem_pos=0,
final_scores=-0.021345407, final_local_scores=0.021399736,
pem=0.0, similarity=4.728324]
gt: 9010672 The_Staple_of_News, pem_pos=0,
final_scores=-0.021345407, final_local_scores=0.021399736,
pem=0.0, similarity=4.728324]
15: 226394 1626_in_literature
[winner: 10445977 1626_in_art, pem_pos=2,
final_scores=-0.58409405, final_local_scores=-0.3837521,
pem=-1.0173962, similarity=-1.8294569]
gt: 226394 1626_in_literature, pem_pos=1,
final_scores=-0.6173489, final_local_scores=-0.6115548,
pem=-0.9911306, similarity=-8.770379]
16: 9015125 A_Tale_of_a_Tub_(play)
[winner: 294082 A_Tale_of_a_Tub, pem_pos=0,
final_scores=-0.08318178, final_local_scores=-0.07659939,
pem=0.0, similarity=1.8038343]
gt: 9015125 A_Tale_of_a_Tub_(play), pem_pos=1,

```

final_scores=-0.14767694, final_local_scores=-0.11835664,
pem=-0.6270854, similarity=3.9679024 ]
17: 226386 1633_in_literature
    [winner: 3731025 1633_in_art, pem_pos=3,
final_scores=-0.9647363, final_local_scores=-0.8042155,
pem=-1.5059526, similarity=-11.720076]
    gt: 226386 1633_in_literature, pem_pos=1,
final_scores=-0.9649272, final_local_scores=-1.0318432,
pem=-1.2691121, similarity=-19.800915 ]
18: 9395552 Ben_Jonson_folios
    [winner: 48261 Ben_Jonson, pem_pos=0,
final_scores=0.16621475, final_local_scores=0.108258024,
pem=0.0, similarity=7.3203483]
    RECALL MISS: 9395552 Ben_Jonson_folios]
    [winner: 9021612 Second_Folio, pem_pos=0,
final_scores=-0.8450989, final_local_scores=-0.87746,
pem=0.0, similarity=-22.095444]
    RECALL MISS: 9395552 Ben_Jonson_folios]
    [winner: 4438223 Caddy_Works, pem_pos=15,
final_scores=-1.1131071, final_local_scores=-1.010968,
pem=-3.0445223, similarity=-9.523002]
    RECALL MISS: 9395552 Ben_Jonson_folios]
    [winner: 12119816 Collection_(abstract_data_type),
pem_pos=1, final_scores=-1.325787,
final_local_scores=-1.2695777, pem=-1.3503652,
similarity=-26.453522]
    RECALL MISS: 9395552 Ben_Jonson_folios]
    [winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.0650558, final_local_scores=-2.1184123,
pem=-3.4067376, similarity=-40.601566]
    RECALL MISS: 9395552 Ben_Jonson_folios]
1: 48261 Ben_Jonson
    [pem_pos=0, final_scores=0.24860601,
final_local_scores=0.19507715, pem=0.0, similarity=9.911204]
2: 143019 Bolsover_Castle
    [pem_pos=0, final_scores=0.25768507,
final_local_scores=0.27745032, pem=0.0, similarity=12.369384]
3: 71070 Derbyshire
    [pem_pos=0, final_scores=0.35607558,
final_local_scores=0.37183452, pem=-0.34993473,
similarity=17.088995]
4: 7426 Charles_I_of_England
    [pem_pos=0, final_scores=0.28876942,
final_local_scores=0.23590946, pem=-0.19786721,
similarity=12.205754]
5: 20949403 Henrietta_Maria_of_France
    [pem_pos=0, final_scores=0.37615162,
final_local_scores=0.34714073, pem=0.0, similarity=14.449084]
6: 7598798 Welbeck_Abbey
    [pem_pos=2, final_scores=0.04838787,
final_local_scores=0.03234075, pem=-0.9964024,
similarity=10.473416]

```

7: 68049 Nottinghamshire
[pem_pos=0, final_scores=0.29765207,
final_local_scores=0.31231305, pem=-0.28438178,
similarity=14.95627]
8: 165453 Inigo_Jones
[pem_pos=0, final_scores=0.335208,
final_local_scores=0.32201433, pem=0.0, similarity=13.699263]
1: 350397 Bolsover
[pem_pos=0, final_scores=0.072534606,
final_local_scores=0.11893445, pem=-0.332106,
similarity=9.444999]
2: 253174 House_of_Stuart
[pem_pos=0, final_scores=0.27372956,
final_local_scores=0.24285248, pem=0.0, similarity=11.336915]
3: 17867 London
[pem_pos=0, final_scores=0.50940794,
final_local_scores=0.55649257, pem=-0.09086737,
similarity=21.190712]
4: 48261 Ben_Jonson
[pem_pos=0, final_scores=0.27875942,
final_local_scores=0.23119593, pem=0.0, similarity=10.98906]
5: 194444 William_Cavendish,_1st_Duke_of_Newcastle
[pem_pos=1, final_scores=0.16615896,
final_local_scores=0.11581361, pem=-0.16907631,
similarity=8.465285]
6: 194444 William_Cavendish,_1st_Duke_of_Newcastle
[pem_pos=1, final_scores=0.102370515,
final_local_scores=0.04758957, pem=0.0, similarity=5.5098805]
7: 48261 Ben_Jonson
[pem_pos=0, final_scores=0.30978096,
final_local_scores=0.26508224, pem=0.0, similarity=12.000296]
8: 7426 Charles_I_of_England
[pem_pos=0, final_scores=0.14465435,
final_local_scores=0.07848567, pem=0.0, similarity=6.431882]
9: 7426 Charles_I_of_England
[pem_pos=0, final_scores=0.038873777,
final_local_scores=-0.03983049, pem=0.0, similarity=2.9010916]
10: 8474 Arthur_Wellesley,_1st_Duke_of_Wellington
[pem_pos=0, final_scores=0.08497469,
final_local_scores=0.06511086, pem=0.0, similarity=6.032751]
11: 48261 Ben_Jonson
[pem_pos=0, final_scores=0.18523933,
final_local_scores=0.12903947, pem=0.0, similarity=7.940508]
12: 48261 Ben_Jonson
[pem_pos=0, final_scores=0.097672984,
final_local_scores=0.030205771, pem=0.0, similarity=4.991113]
13: 165453 Inigo_Jones
[pem_pos=0, final_scores=0.15879638,
final_local_scores=0.12868896, pem=0.0, similarity=7.9300485]
14: 350397 Bolsover
[pem_pos=0, final_scores=0.03403206,
final_local_scores=0.07687627, pem=-0.332106,

```

similarity=8.1898985]
15: 48261 Ben_Jonson
    [pem_pos=0, final_scores=0.18479712,
final_local_scores=0.12855643, pem=0.0, similarity=7.926093]
16: 194444 William_Cavendish,1st_Duke_of_Newcastle
    [pem_pos=1, final_scores=0.109837726,
final_local_scores=0.055746377, pem=0.0, similarity=5.753296]
17: 48261 Ben_Jonson
    [pem_pos=0, final_scores=0.16621475,
final_local_scores=0.108258024, pem=0.0, similarity=7.3203483]

```

Sample: b'Oriol_Lozano'

Oriol Lozano Farrán , simply known as Oriol -LRB- born 23 May 1981 -RRB- , is a [1 1 Spanish 1] professional 1 football er who plays for [2 2 FC Zestafoni 2] in [3 3 Georgia 3] , usually as a . [1 Oriol 1] was born in 2 Sudanell , [2 3 Lleid a 2] , [4 4 Catalonia 4] . During his early career he played with [5 5 UE Lleida 5] , [6 6 CD Onda 6] -LRB- in the 4 fourth division , loaned by [7 7 Villarreal CF 7] , who acted as Onda 's parent club -RRB- and 5 CE L'Hospitalet . In 6 2004 -- 05 he was promoted from [8 8 Racing de Santander 8] 's 7 B- team , making his [9 9 La Liga 9] debut on 19 September 2004 in a 1 -- 1 home draw against [3 Villarreal 3] . [4 Oriol 4] went on to become a relatively important defensive element for the 8 Cantabria ns in the subsequent four seasons , playing in a total of 75 games -LRB- 55 starts -RRB- . On 17 March 2007 he scored his only official goal for the club , in a 2 -- 0 away win against [10 10 Valencia CF 10] . [5 Oriol 5] appeared in 19 matches in the 9 2009 -- 10 season as [6 Racing 6] once again narrowly avoided relegation after finishing 16th -- he was 10 sent off twice , one of them in the last minute of the side 's 1 -- 0 success at [11 11 Sport ing de Gijón 11] . In June 2010 , the 29-year-old did not see his contract renewed and moved to [12 12 Greece 12] with 11 Ar is Thessaloniki FC

```

1: 26667 Spain
2: 6581398 FC_Zestafoni
3: 48768 Georgia_(country)
4: 6822 Catalonia
5: 32465511 UE_Lleida
6: 8313490 CD_Onda
7: 958154 Villarreal_CF
8: 7614075 Racing_de_Santander
9: 37981 La_Liga
10: 32780 Valencia_CF
11: 79943 Sporting_de_Gijón
12: 12108 Greece

```

```

1: 10568 Association_football
    [winner: 10568 Association_football, pem_pos=0,
final_scores=-1.6824996, final_local_scores=-1.8543549,
pem=0.0, similarity=-51.247936]
    gt: 10568 Association_football, pem_pos=0,

```

final_scores=-1.6824996, final_local_scores=-1.8543549,
pem=0.0, similarity=-51.247936]
2: 12237595 Sudanell
[winner: 12237595 Sudanell, pem_pos=0,
final_scores=-0.1299422, final_local_scores=0.033219025,
pem=0.0, similarity=5.0810347|
gt: 12237595 Sudanell, pem_pos=0,
final_scores=-0.1299422, final_local_scores=0.033219025,
pem=0.0, similarity=5.0810347]
3: 143464 Province_of_Lleida
[winner: 18729 Lleida, pem_pos=0,
final_scores=0.23457934, final_local_scores=0.26343778,
pem=-0.13350299, similarity=12.677232|
gt: 143464 Province_of_Lleida, pem_pos=1,
final_scores=-0.029798865, final_local_scores=0.01760383,
pem=-1.0116879, similarity=10.116762]
4: 30864930 Tercera_Divisi3n
[winner: 526932 Football_League_Fourth_Division,
pem_pos=0, final_scores=-1.0773827,
final_local_scores=-1.165479, pem=-0.027341865,
similarity=-30.541811|
gt: 30864930 Tercera_Divisi3n, pem_pos=3,
final_scores=-1.3676605, final_local_scores=-1.5631573,
pem=-2.4381227, similarity=-29.299131]
5: 4857293 CE_L'Hospitalet
[winner: 4857293 CE_L'Hospitalet, pem_pos=0,
final_scores=-0.01006186, final_local_scores=-0.05142338,
pem=0.0, similarity=2.5551367|
gt: 4857293 CE_L'Hospitalet, pem_pos=0,
final_scores=-0.01006186, final_local_scores=-0.05142338,
pem=0.0, similarity=2.5551367]
6: 30863613 2004-05_La_Liga
[winner: 35524 2004, pem_pos=0, final_scores=-0.6072981,
final_local_scores=-0.45085236, pem=-0.4915334,
similarity=-6.691586|
RECALL MISS: 30863613 2004-05_La_Liga]
[winner: 3921642 2005_in_association_football, pem_pos=3,
final_scores=-0.6267024, final_local_scores=-0.61770105,
pem=-2.7573876, similarity=0.65138507|
RECALL MISS: 30863613 2004-05_La_Liga]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3030746, final_local_scores=-1.2137525,
pem=0.0, similarity=-32.13108|
RECALL MISS: 30863613 2004-05_La_Liga]
7: 22064225 Racing_de_Santander_B
[winner: 9037074 Atl3tico_Madrid_B, pem_pos=1,
final_scores=-0.21613844, final_local_scores=-0.2802459,
pem=-0.6426777, similarity=-0.77840286|
RECALL MISS: 22064225 Racing_de_Santander_B]
8: 69605 Cantabria
[winner: 69605 Cantabria, pem_pos=0,
final_scores=-0.0030779243, final_local_scores=-0.05455835,

pem=0.0, similarity=2.461583|
gt: 69605 Cantabria, pem_pos=0,
final_scores=-0.0030779243, final_local_scores=-0.05455835,
pem=0.0, similarity=2.461583]
9: 22640051 2009-10_La_Liga
[winner: 13745451 ICD-10, pem_pos=0,
final_scores=-0.7535685, final_local_scores=-0.65079635,
pem=-1.1311272, similarity=-9.180103|
RECALL MISS: 22640051 2009-10_La_Liga]
[winner: 35983 2009, pem_pos=0, final_scores=-0.8657998,
final_local_scores=-0.7653134, pem=-0.9296576,
similarity=-13.693143|
RECALL MISS: 22640051 2009-10_La_Liga]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.2892848, final_local_scores=-1.1986892,
pem=0.0, similarity=-31.681562|
RECALL MISS: 22640051 2009-10_La_Liga]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.8361565, final_local_scores=-1.9535067,
pem=-1.219945, similarity=-47.57257|
RECALL MISS: 22640051 2009-10_La_Liga]
10: 2176922 Ejection_(sports)
[winner: 2176922 Ejection_(sports), pem_pos=0,
final_scores=-1.0514859, final_local_scores=-1.1514683,
pem=-0.12184974, similarity=-29.60976|
gt: 2176922 Ejection_(sports), pem_pos=0,
final_scores=-1.0514859, final_local_scores=-1.1514683,
pem=-0.12184974, similarity=-29.60976]
11: 25625594 Aris_F.C._(Thessaloniki)
[winner: 25625594 Aris_F.C._(Thessaloniki), pem_pos=0,
final_scores=-0.045759365, final_local_scores=-0.091195,
pem=0.0, similarity=1.3682722|
gt: 25625594 Aris_F.C._(Thessaloniki), pem_pos=0,
final_scores=-0.045759365, final_local_scores=-0.091195,
pem=0.0, similarity=1.3682722]
1: 26667 Spain
[pem_pos=1, final_scores=0.54741544,
final_local_scores=0.5736918, pem=-1.2425895,
similarity=27.967215]
2: 6581398 FC_Zestafoni
[pem_pos=0, final_scores=0.36904275,
final_local_scores=0.43934768, pem=0.0, similarity=17.200724]
3: 48768 Georgia_(country)
[pem_pos=1, final_scores=0.30632716,
final_local_scores=0.39339978, pem=-1.3466994,
similarity=23.153107]
4: 6822 Catalonia
[pem_pos=0, final_scores=0.39304882,
final_local_scores=0.38914967, pem=-0.05600393,
similarity=16.007273]
5: 32465511 UE_Lleida
[pem_pos=0, final_scores=0.36955905,

final_local_scores=0.3510084, pem=0.0, similarity=14.564504]
 6: 8313490 CD_Onda
 [pem_pos=0, final_scores=0.037707075,
 final_local_scores=0.077935055, pem=0.0, similarity=6.4154506]
 7: 958154 Villarreal_CF
 [pem_pos=0, final_scores=0.413356,
 final_local_scores=0.36978978, pem=0.0, similarity=15.124978]
 8: 7614075 Racing_de_Santander
 [pem_pos=0, final_scores=0.30734414,
 final_local_scores=0.25695106, pem=0.0, similarity=11.757646]
 9: 37981 La_Liga
 [pem_pos=0, final_scores=0.26181084,
 final_local_scores=0.1978645, pem=0.0, similarity=9.994385]
 10: 32780 Valencia_CF
 [pem_pos=0, final_scores=0.5024717,
 final_local_scores=0.46534246, pem=0.0, similarity=17.97646]
 11: 79943 Sporting_de_Gijón
 [pem_pos=0, final_scores=0.30728608,
 final_local_scores=0.25706607, pem=0.0, similarity=11.761077]
 12: 12108 Greece
 [pem_pos=0, final_scores=0.5901351,
 final_local_scores=0.6817649, pem=-0.27607638,
 similarity=25.936283]
 1: 10210196 Oriol_Lozano
 [pem_pos=0, final_scores=0.11613719,
 final_local_scores=0.1295489, pem=0.0, similarity=7.955711]
 2: 18729 Lleida
 [pem_pos=0, final_scores=0.23457934,
 final_local_scores=0.26343778, pem=-0.13350299,
 similarity=12.677232]
 3: 958154 Villarreal_CF
 [pem_pos=0, final_scores=0.4943201,
 final_local_scores=0.45823073, pem=-0.09685928,
 similarity=18.290968]
 4: 10210196 Oriol_Lozano
 [pem_pos=0, final_scores=0.14368607,
 final_local_scores=0.1588687, pem=0.0, similarity=8.830672]
 5: 10210196 Oriol_Lozano
 [pem_pos=0, final_scores=0.124926135,
 final_local_scores=0.13837633, pem=0.0, similarity=8.219138]
 6: 1404901 Racing_Club_de_Avellaneda
 [pem_pos=2, final_scores=0.19874017,
 final_local_scores=0.17601192, pem=-1.9044929,
 similarity=19.699188]

Sample: b'Juan_Cruz_Ochoa'

Juan Cruz Ochoa López -LRB- born 4 March 1979 in [1 1 Pamplo
 na 1] , [2 2 Navarre 2] -RRB- is a [3 3 Spanish 3] 1 football
 er who plays for [4 4 Orihuela CF 4] as a . After beginning
 professionally with [5 5 CD Calahorra 5] , [1 Ochoa 1] moved
 to [6 6 Deportivo Alavés 6] , appearing once for its first
 team in 2 2001 -- 02 . He could only appear in a further four

[7 7 La Liga 7] games during 3 the following campaign , being mainly registered with 4 its B side . In 5 2004 -- 05 [2 Ochoa 2] signed with [8 8 CD Numancia 8] , experiencing top flight relegation in his first year . After 6 a second , in which he also was heavily featured , he moved to another side in [9 9 Segunda División 9] , [10 10 Real Murcia 10] , playing 21 matches in 7 his second season for an eventual top level return . Having signed with the team the previous summer , [3 Ochoa 3] started 8 2011 -- 12 in the second level with [11 11 SD Huesca 11] , being released in the very last day of the winter 9 transfer window . On 20 September the 33-year old returned to active , signing for one season with [12 12 Orihuela CF 12] in [13 13 Segunda División B 13]

1: 59732 Pamplona

2: 59731 Navarre

3: 26667 Spain

4: 5800486 Orihuela_CF

5: 9543209 CD_Calahorra

6: 958234 Deportivo_Alavés

7: 37981 La_Liga

8: 82773 CD_Numancia

9: 5152226 Segunda_División

10: 1944449 Real_Murcia

11: 5590251 SD_Huesca

12: 5800486 Orihuela_CF

13: 30863062 Segunda_División_B

1: 10568 Association_football

[winner: 10568 Association_football, pem_pos=0, final_scores=-1.71905, final_local_scores=-1.8603444, pem=0.0, similarity=-51.426678]

gt: 10568 Association_football, pem_pos=0, final_scores=-1.71905, final_local_scores=-1.8603444, pem=0.0, similarity=-51.426678]

2: 6484617 2001-02_La_Liga

[winner: 34551 2001, pem_pos=0, final_scores=-0.62260944, final_local_scores=-0.3930776, pem=-0.5613713, similarity=-4.5876822]

RECALL MISS: 6484617 2001-02_La_Liga]

[winner: 12574894 Now_02_(Australian_series), pem_pos=0, final_scores=-0.8162733, final_local_scores=-0.6798733, pem=-1.8204205, similarity=-6.2993317]

RECALL MISS: 6484617 2001-02_La_Liga]

[winner: 38340996 Methylene_bridge, pem_pos=0, final_scores=-1.3023462, final_local_scores=-1.2115691, pem=0.0, similarity=-32.06592]

RECALL MISS: 6484617 2001-02_La_Liga]

3: 6484694 2002-03_La_Liga

[winner: 17867555 1993-94_Nationalliga_A, pem_pos=11, final_scores=-1.1512166, final_local_scores=-1.0891654, pem=-2.5513864, similarity=-14.538318]

RECALL MISS: 6484694 2002-03_La_Liga]

[winner: 17867399

Chuck_Baldwin_presidential_campaign_2008, pem_pos=13,
 final_scores=-1.7762198, final_local_scores=-1.7547069,
 pem=-4.0512786, similarity=-26.242758|
 RECALL MISS: 6484694 2002-03_La_Liga]

4: 22534943 Deportivo_Alavés_B
 [winner: 24309566 B_V_Kale, pem_pos=5,
 final_scores=-0.6894585, final_local_scores=-0.5127051,
 pem=-2.0794415, similarity=0.09789665|
 RECALL MISS: 22534943 Deportivo_Alavés_B]
 [winner: 9871300 Richard_Side, pem_pos=10,
 final_scores=-1.3123873, final_local_scores=-1.1887512,
 pem=-2.9957323, similarity=-15.093732|
 RECALL MISS: 22534943 Deportivo_Alavés_B]

5: 30863613 2004-05_La_Liga
 [winner: 35524 2004, pem_pos=0, final_scores=-0.58860165,
 final_local_scores=-0.42469114, pem=-0.4915334,
 similarity=-5.9108825|
 RECALL MISS: 30863613 2004-05_La_Liga]
 [winner: 3423380 Fabric_05, pem_pos=4,
 final_scores=-0.96358657, final_local_scores=-0.85087967,
 pem=-2.7660456, similarity=-6.260049|
 RECALL MISS: 30863613 2004-05_La_Liga]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.2544688, final_local_scores=-1.1592703,
 pem=0.0, similarity=-30.505226|
 RECALL MISS: 30863613 2004-05_La_Liga]

6: 30865849 2005-06_Segunda_División
 [winner: 36011930
 Maldivian_constitutional_referendum_1953, pem_pos=4,
 final_scores=-1.1830878, final_local_scores=-1.0810039,
 pem=-3.637586, similarity=-8.38784|
 gt: 30865849 2005-06_Segunda_División, pem_pos=5,
 final_scores=-1.213938, final_local_scores=-1.3714783,
 pem=-3.637586, similarity=-17.056177]

7: 6410212 2006-07_Segunda_División
 [winner: 30865815 2005-06_La_Liga, pem_pos=0,
 final_scores=-0.98462355, final_local_scores=-1.1283445,
 pem=0.0, similarity=-29.58234|
 RECALL MISS: 6410212 2006-07_Segunda_División]
 [winner: 690345 Season(sports), pem_pos=0,
 final_scores=-1.4127023, final_local_scores=-1.4765322,
 pem=-1.219945, similarity=-33.338703|
 RECALL MISS: 6410212 2006-07_Segunda_División]

8: 27666856 2011-12_Segunda_División
 [winner: 19497 May_12, pem_pos=3,
 final_scores=-0.9758658, final_local_scores=-0.7735683,
 pem=-3.2948976, similarity=-1.0769433|
 RECALL MISS: 27666856 2011-12_Segunda_División]
 [winner: 36225 2011, pem_pos=0, final_scores=-1.1182393,
 final_local_scores=-1.0041023, pem=-1.4232541,
 similarity=-18.134827|
 RECALL MISS: 27666856 2011-12_Segunda_División]

[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.1955116, final_local_scores=-1.0948685,
pem=0.0, similarity=-28.583347]
RECALL MISS: 27666856 2011-12_Segunda_División]

9: 2607843 Transfer_window
[winner: 1550386 Chaff_(countermeasure), pem_pos=1,
final_scores=-0.97811615, final_local_scores=-0.8643575,
pem=-0.6775151, similarity=-18.020004]
RECALL MISS: 2607843 Transfer_window]

[winner: 4079673 Optical_transfer_function, pem_pos=4,
final_scores=-1.4003085, final_local_scores=-1.3255424,
pem=-1.4980085, similarity=-27.320713]
RECALL MISS: 2607843 Transfer_window]

1: 59732 Pamplona
[pem_pos=0, final_scores=0.3262326,
final_local_scores=0.34772167, pem=-0.03809467,
similarity=14.673586]

2: 59731 Navarre
[pem_pos=0, final_scores=0.097612515,
final_local_scores=0.088493675, pem=-0.25764343,
similarity=8.131646]

3: 26667 Spain
[pem_pos=1, final_scores=0.37684423,
final_local_scores=0.41954958, pem=-1.2425895,
similarity=23.367304]

4: 5800486 Orihuela_CF
[pem_pos=0, final_scores=0.19491182,
final_local_scores=0.15782169, pem=0.0, similarity=8.799427]

5: 9543209 CD_Calahorra
[pem_pos=0, final_scores=0.1024525,
final_local_scores=0.10719408, pem=0.0, similarity=7.288598]

6: 958234 Deportivo_Alavés
[pem_pos=0, final_scores=0.26004177,
final_local_scores=0.19450131, pem=0.0, similarity=9.89402]

7: 37981 La_Liga
[pem_pos=0, final_scores=0.26226062,
final_local_scores=0.20112017, pem=0.0, similarity=10.09154]

8: 82773 CD_Numancia
[pem_pos=0, final_scores=0.25577396,
final_local_scores=0.19534945, pem=0.0, similarity=9.919331]

9: 5152226 Segunda_División
[pem_pos=0, final_scores=0.20137452,
final_local_scores=0.13982913, pem=-0.19970337,
similarity=9.348511]

10: 1944449 Real_Murcia
[pem_pos=0, final_scores=0.39708775,
final_local_scores=0.35365134, pem=0.0, similarity=14.643374]

11: 5590251 SD_Huesca
[pem_pos=0, final_scores=0.3321541,
final_local_scores=0.28554764, pem=0.0, similarity=12.611024]

12: 5800486 Orihuela_CF
[pem_pos=0, final_scores=0.17152031,

final_local_scores=0.13227001, pem=0.0, similarity=8.036914]
 13: 30863062 Segunda_División_B
 [pem_pos=0, final_scores=0.03127913,
 final_local_scores=-0.04253652, pem=-0.01971423,
 similarity=2.9275475]
 1: 1388604 Guillermo_Ochoa
 [pem_pos=2, final_scores=0.1536882,
 final_local_scores=0.13445908, pem=-1.3492571,
 similarity=15.43971]
 2: 1388604 Guillermo_Ochoa
 [pem_pos=2, final_scores=0.09413843,
 final_local_scores=0.072401255, pem=-1.3492571,
 similarity=13.587781]
 3: 1388604 Guillermo_Ochoa
 [pem_pos=2, final_scores=0.15312748,
 final_local_scores=0.13384658, pem=-1.3492571,
 similarity=15.4214325]

Sample: b'Colorado_State_Highway_94'

State Highway 94 is an long east -- west [1 Colorado 1]
 state highway that goes east beginning from just east of and
 ending at / in rural [1 1 Cheyenne County 1] west of [2 2 Wild
 Horse 2] . The highway serves [3 3 Schriever Air Force Base 3]
 and the towns of 1 Ellicott , 2 Yoder , and 3 Rush , where it
 then crosses through rural country . It also serves the 4 Pun
 kin Center area where it comes to a junction with 5 SH 71 .
 It then passes through more very rural country until it
 reaches a stop at US 40 / 6 287 west of the unincorporated
 town of [4 4 Wild Horse 4] . Throughout most of its length ,
 the highway runs due east -- west , pointing exactly at [5 5 P
 ikes Peak 5] . Under good conditions , the peak can be seen
 as far east as the rise between s 78 and 79 . `` Garbage Hill
 '' gets its name from the [2 Colorado Springs 2] 7 landfill
 located north of the highway . Originally , the entrance to
 the landfill was partway up the hill , but currently the
 entrance is from Blaney Road at the top of the hill . At the
 bottom of the hill , which is the eastern side of Jimmy Camp
 Creek valley , there are 8 wrecking yard s and two 9 motocross
 tracks . This hill is one of the most severe hills on Highway
 94 , having a . The route was established in the 1920s , with
 the eastern end swerving north then east , ending at 10 Boyero
 . By 1954 , the eastern end no longer swerved north , and by
 1970 the eastern end was extended to US 40 / 11 287 near 12 Ar
 oya

1: 97009 Cheyenne_County,_Colorado
 2: 8737357 Wild_Horse,_Colorado
 3: 1658649 Schriever_Air_Force_Base
 4: 8737357 Wild_Horse,_Colorado
 5: 26257090 Pikes_Peak
 1: 10655525 Ellicott,_Colorado
 [winner: 126236 Ellicott,_New_York, pem_pos=1,
 final_scores=-0.11720918, final_local_scores=-0.0854249,

pem=-1.1988889, similarity=8.060208|
gt: 10655525 Ellicott, Colorado, pem_pos=2,
final_scores=-0.17231941, final_local_scores=-0.14087692,
pem=-1.6436559, similarity=8.8241205]
2: 8737639 Yoder, Colorado
[winner: 653690 Yoder, pem_pos=0,
final_scores=-0.2127403, final_local_scores=-0.09008007,
pem=-1.1880018, similarity=7.862082|
gt: 8737639 Yoder, Colorado, pem_pos=3,
final_scores=-0.32460886, final_local_scores=-0.28578055,
pem=-1.87557, similarity=5.7610927]
3: 8725482 Rush, Colorado
[winner: 25432 Rush (band), pem_pos=0,
final_scores=-0.28177, final_local_scores=-0.1925424,
pem=-0.49376196, similarity=1.0290179|
gt: 8725482 Rush, Colorado, pem_pos=14,
final_scores=-0.6415738, final_local_scores=-0.65142673,
pem=-4.260393, similarity=7.81852]
4: 11893166 Punkin_Center, Colorado
[winner: 7554650 Punkin_Center, Arizona, pem_pos=0,
final_scores=0.016271308, final_local_scores=0.11130379,
pem=0.0, similarity=7.41124|
gt: 11893166 Punkin_Center, Colorado, pem_pos=1,
final_scores=-0.07806961, final_local_scores=-0.0071005076,
pem=-0.89658564, similarity=8.753592]
5: 18672136 Colorado_State_Highway_71
[winner: 18672136 Colorado_State_Highway_71, pem_pos=1,
final_scores=-0.50862104, final_local_scores=-0.48179698,
pem=-0.41275305, similarity=-8.043454|
gt: 18672136 Colorado_State_Highway_71, pem_pos=1,
final_scores=-0.50862104, final_local_scores=-0.48179698,
pem=-0.41275305, similarity=-8.043454]
6: 438343 U.S. Route 287
[winner: 20488849 SP-287, pem_pos=2,
final_scores=-0.75830036, final_local_scores=-0.5204108,
pem=-1.5830046, similarity=-2.8317566|
gt: 438343 U.S. Route 287, pem_pos=0,
final_scores=-0.9010361, final_local_scores=-0.98272765,
pem=-1.2405275, similarity=-18.490658]
7: 211285 Landfill
[winner: 211285 Landfill, pem_pos=0,
final_scores=-1.5588436, final_local_scores=-1.624651,
pem=-0.09809306, similarity=-43.85967|
gt: 211285 Landfill, pem_pos=0, final_scores=-1.5588436,
final_local_scores=-1.624651, pem=-0.09809306,
similarity=-43.85967]
8: 1686055 Wrecking_yard
[winner: 1404298 Wrecking (Soviet_crime), pem_pos=0,
final_scores=-1.1309986, final_local_scores=-1.0428445,
pem=-0.64737874, similarity=-23.5103|
RECALL MISS: 1686055 Wrecking_yard]
[winner: 1505481 Rail_yard, pem_pos=1,

final_scores=-1.3018159, final_local_scores=-1.3505101,
pem=-0.954412, similarity=-31.021961]
RECALL MISS: 1686055 Wrecking_yard]

9: 254052 Motocross
[winner: 7934048 Hawkstone_Park_Motocross_Circuit,
pem_pos=0, final_scores=-1.5564381,
final_local_scores=-1.5464513, pem=0.0, similarity=-42.05948]
gt: 254052 Motocross, pem_pos=1,
final_scores=-1.8505663, final_local_scores=-1.9428184,
pem=-0.00089536444, similarity=-53.88299]

10: 8737956 Boyero, Colorado
[winner: 8737956 Boyero, Colorado, pem_pos=1,
final_scores=-0.15226924, final_local_scores=-0.12374232,
pem=-0.10008345, similarity=0.9412646]
gt: 8737956 Boyero, Colorado, pem_pos=1,
final_scores=-0.15226924, final_local_scores=-0.12374232,
pem=-0.10008345, similarity=0.9412646]

11: 438343 U.S. Route 287
[winner: 20488849 SP-287, pem_pos=2,
final_scores=-0.7422694, final_local_scores=-0.50289935,
pem=-1.5830046, similarity=-2.3091812]
gt: 438343 U.S. Route 287, pem_pos=0,
final_scores=-0.999988, final_local_scores=-1.0908176,
pem=-1.2405275, similarity=-21.716278]

12: 29855197 Aroya, Colorado
[winner: 1190530 Arøya, pem_pos=0,
final_scores=-0.18512335, final_local_scores=-0.023852386,
pem=0.0, similarity=3.3779101]
gt: 29855197 Aroya, Colorado, pem_pos=1,
final_scores=-0.557715, final_local_scores=-0.54279816,
pem=-1.9195929, similarity=-1.6694214]

1: 97009 Cheyenne County, Colorado
[pem_pos=0, final_scores=0.15529357,
final_local_scores=0.17316467, pem=-0.42063507,
similarity=11.544771]

2: 8737357 Wild_Horse, Colorado
[pem_pos=1, final_scores=0.07859634,
final_local_scores=0.09749581, pem=-1.3274941,
similarity=14.218302]

3: 1658649 Schriever Air Force Base
[pem_pos=0, final_scores=0.08070524,
final_local_scores=0.13392735, pem=0.0, similarity=8.086372]

4: 8737357 Wild_Horse, Colorado
[pem_pos=1, final_scores=0.065569565,
final_local_scores=0.08431189, pem=-1.3274941,
similarity=13.824867]

5: 26257090 Pikes Peak
[pem_pos=0, final_scores=0.16441874,
final_local_scores=0.19716138, pem=-0.03459284,
similarity=10.161523]

1: 5399 Colorado
[pem_pos=0, final_scores=0.5569057,

final_local_scores=0.56857836, pem=-0.26926568,
similarity=22.521534]
2: 6250 Colorado_Springs,_Colorado
[pem_pos=0, final_scores=0.35785466,
final_local_scores=0.38103187, pem=0.0, similarity=15.4604645]

Sample: b'Flavius_Justus'

Titus Flavius Justus -LRB- , born 76 -RRB- was an aristocratic , wealthy 1 Roman 2 Jew . Justus was born and raised in [1 3 Rome 1] . He was the first son born to the [2 Roman Jewish 2] historian [1 1 Josephus 1] and his fourth wife , a distinguished unnamed [3 4 Greek 3] [4 Jewish 4] noblewoman from [2 2 Crete 2] . The parents of Justus met in [5 Rome 5] , and [6 Josephus 6] describes his mother , as one ' in character who excelled many women , as her subsequent life demonstrated ' . Justus had an older paternal half-brother , 5 Hyrcanus , and a younger brother , 6 Flavius Simonides Agrippa . His paternal uncle was 7 Matthias , while his paternal grandparents were 8 Matthias and his wife , an unnamed [7 Jewish 7] noblewoman . His paternal grandmother was an aristocratic woman who descended from royalty and of the former ruling . His paternal grandfather descended from the priestly order of the 9 Jehoiarib , which was the first of the twenty four-orders of priests in the [8 10 Temple 8] in [9 Jerusalem 9] and was a descendant of the [10 High 10] Priest Jonathon . Jonathon may have been [3 3 Alexander Jannaeus 3] , the [11 High Priest 11] and who governed [4 4 Judea 4] from 103 BC-76 BC . Justus ' 11 cognomen was a very common [5 5 Latin 5] name in the Jewish 12 Catacombs of [12 Rome 12] . [13 Josephus 13] gave his son a [14 Latin 14] name in preference to a Greek one , despite the fact that [15 Greek 15] was spoken more among [16 Roman Jews 16] than [17 Latin 17] . Justus was a contemporary of the ruling 13 Flavian dynasty and 14 Nerva -- Antonine dynasty of the [6 6 Roman Empire 6]

1: 16494 Josephus

2: 6591 Crete

3: 1606 Alexander_Jannaeus

4: 21491730 Judea

5: 17730 Latin

6: 25507 Roman_Empire

1: 1037008 Roman_citizenship

[winner: 521555 Ancient_Rome, pem_pos=0,
final_scores=-0.4908113, final_local_scores=-0.6143743,
pem=-1.0249887, similarity=-8.670394]

RECALL MISS: 1037008 Roman_citizenship]

2: 1263604 History_of_the_Jews_in_Italy

[winner: 25955086 Jews, pem_pos=0,
final_scores=-0.83287543, final_local_scores=-0.97465426,
pem=0.0, similarity=-24.995914]

RECALL MISS: 1263604 History_of_the_Jews_in_Italy]

3: 521555 Ancient_Rome

[winner: 25458 Rome, pem_pos=0, final_scores=0.5171837,

final_local_scores=0.5171261, pem=-0.2517334,
 similarity=20.89075|
 gt: 521555 Ancient_Rome, pem_pos=1,
 final_scores=-0.23676859, final_local_scores=-0.33686942,
 pem=-2.8942451, similarity=9.776213]
 4: 66540 Ancient_Greece
 [winner: 12108 Greece, pem_pos=1,
 final_scores=0.29508758, final_local_scores=0.3076408,
 pem=-2.3247876, similarity=25.912884|
 gt: 66540 Ancient_Greece, pem_pos=4,
 final_scores=-0.6002982, final_local_scores=-0.71134204,
 pem=-2.8341312, similarity=-1.7257059]
 [winner: 25955086 Jews, pem_pos=1,
 final_scores=0.18538593, final_local_scores=0.14208856,
 pem=-0.80476934, similarity=12.70638|
 RECALL MISS: 66540 Ancient_Greece]
 5: 32038207 Flavius_Hyrchanus
 [winner: 817525 John_Hyrchanus, pem_pos=1,
 final_scores=-0.026191562, final_local_scores=-0.06066502,
 pem=-0.034786116, similarity=2.4685202|
 gt: 32038207 Flavius_Hyrchanus, pem_pos=2,
 final_scores=-0.87438124, final_local_scores=-0.9082391,
 pem=-4.7621737, similarity=2.8834865]
 6: 32065187 Flavius_Simonides_Agrippa
 [winner: 32065187 Flavius_Simonides_Agrippa, pem_pos=0,
 final_scores=0.017136768, final_local_scores=0.041374415,
 pem=0.0, similarity=5.324408|
 gt: 32065187 Flavius_Simonides_Agrippa, pem_pos=0,
 final_scores=0.017136768, final_local_scores=0.041374415,
 pem=0.0, similarity=5.324408]
 7: 32016375 Matthias_(brother_of_Josephus)
 [winner: 151080 Matthias,_Holy_Roman_Emperor, pem_pos=1,
 final_scores=0.012839511, final_local_scores=0.015618861,
 pem=-1.0671463, similarity=10.359118|
 RECALL MISS: 32016375 Matthias_(brother_of_Josephus)]
 8: 31977155 Matthias_(father_of_Josephus)
 [winner: 151080 Matthias,_Holy_Roman_Emperor, pem_pos=1,
 final_scores=0.014617607, final_local_scores=0.017561167,
 pem=-1.0671463, similarity=10.417081|
 gt: 31977155 Matthias_(father_of_Josephus), pem_pos=10,
 final_scores=-0.6876914, final_local_scores=-0.6956045,
 pem=-4.9020953, similarity=9.989844]
 9: 7680132 Jehoiarib
 [winner: 7680132 Jehoiarib, pem_pos=0,
 final_scores=-0.080187604, final_local_scores=0.010213688,
 pem=0.0, similarity=4.39451|
 gt: 7680132 Jehoiarib, pem_pos=0,
 final_scores=-0.080187604, final_local_scores=0.010213688,
 pem=0.0, similarity=4.39451]
 10: 30526 Temple_in_Jerusalem
 [winner: 16043 Jerusalem, pem_pos=0,
 final_scores=0.5115747, final_local_scores=0.4914381,

```

pem=-0.13239029, similarity=19.475164|
  RECALL MISS: 30526 Temple_in_Jerusalem]
  [winner: 30526 Temple_in_Jerusalem, pem_pos=2,
final_scores=0.09639402, final_local_scores=0.05809474,
pem=-2.1650271, similarity=17.59713|
  gt: 30526 Temple_in_Jerusalem, pem_pos=2,
final_scores=0.09639402, final_local_scores=0.05809474,
pem=-2.1650271, similarity=17.59713 ]
11: 241700 Cognomen
  [winner: 241700 Cognomen, pem_pos=0,
final_scores=-1.0058719, final_local_scores=-1.091402,
pem=-0.023673315, similarity=-28.351164|
  gt: 241700 Cognomen, pem_pos=0, final_scores=-1.0058719,
final_local_scores=-1.091402, pem=-0.023673315,
similarity=-28.351164 ]
12: 177922 Catacombs_of_Rome
  [winner: 25458 Rome, pem_pos=0, final_scores=0.49694353,
final_local_scores=0.49501675, pem=-0.2517334,
similarity=20.230965|
  RECALL MISS: 177922 Catacombs_of_Rome]
  [winner: 19083084 Catacombs_(band), pem_pos=5,
final_scores=-0.63061374, final_local_scores=-0.46298304,
pem=-1.833251, similarity=0.24287891|
  gt: 177922 Catacombs_of_Rome, pem_pos=2,
final_scores=-0.8650034, final_local_scores=-0.951853,
pem=-1.4782517, similarity=-16.276516 ]
13: 532225 Flavian_dynasty
  [winner: 532225 Flavian_dynasty, pem_pos=2,
final_scores=-0.02413565, final_local_scores=-0.065496564,
pem=-1.1414793, similarity=8.342707|
  gt: 532225 Flavian_dynasty, pem_pos=2,
final_scores=-0.02413565, final_local_scores=-0.065496564,
pem=-1.1414793, similarity=8.342707 ]
  [winner: 27158897 Dynasty_(Kaskade_album), pem_pos=7,
final_scores=-1.30188, final_local_scores=-1.206789,
pem=-3.3888543, similarity=-13.494161|
  RECALL MISS: 532225 Flavian_dynasty]
14: 3008867 Nerva-Antonine_dynasty
  [winner: 23927371 Nerva, pem_pos=0,
final_scores=0.015416577, final_local_scores=0.008707896,
pem=0.0, similarity=4.349574|
  RECALL MISS: 3008867 Nerva-Antonine_dynasty]
  [winner: 3008867 Nerva-Antonine_dynasty, pem_pos=0,
final_scores=0.0052396506, final_local_scores=0.0077281147,
pem=-0.19997646, similarity=5.4078383|
  gt: 3008867 Nerva-Antonine_dynasty, pem_pos=0,
final_scores=0.0052396506, final_local_scores=0.0077281147,
pem=-0.19997646, similarity=5.4078383 ]
  [winner: 27158897 Dynasty_(Kaskade_album), pem_pos=7,
final_scores=-1.3043907, final_local_scores=-1.2095315,
pem=-3.3888543, similarity=-13.576002|
  RECALL MISS: 3008867 Nerva-Antonine_dynasty]

```

[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3201123, final_local_scores=-1.2453237,
pem=0.0, similarity=-33.073227]

RECALL MISS: 3008867 Nerva-Antonine_dynasty]

1: 16494 Josephus

[pem_pos=0, final_scores=0.3024469,
final_local_scores=0.25763556, pem=0.0, similarity=11.778072]

2: 6591 Crete

[pem_pos=0, final_scores=0.50368536,
final_local_scores=0.5075856, pem=-0.008379842,
similarity=19.28265]

3: 1606 Alexander_Jannaeus

[pem_pos=0, final_scores=0.2068687,
final_local_scores=0.1976141, pem=0.0, similarity=9.986912]

4: 21491730 Judea

[pem_pos=0, final_scores=0.33105296,
final_local_scores=0.29238534, pem=0.0, similarity=12.815075]

5: 17730 Latin

[pem_pos=0, final_scores=0.16527419,
final_local_scores=0.14521578, pem=-0.08185364,
similarity=8.868375]

6: 25507 Roman_Empire

[pem_pos=0, final_scores=0.16722946,
final_local_scores=0.095589995, pem=0.0, similarity=6.942309]

1: 25458 Rome

[pem_pos=0, final_scores=0.5171837,
final_local_scores=0.5171261, pem=-0.2517334,
similarity=20.89075]

2: 25507 Roman_Empire

[pem_pos=0, final_scores=0.27589124,
final_local_scores=0.21471232, pem=-0.6931472,
similarity=14.266599]

3: 12108 Greece

[pem_pos=1, final_scores=0.29508758,
final_local_scores=0.3076408, pem=-2.3247876,
similarity=25.912884]

4: 25955086 Jews

[pem_pos=1, final_scores=0.18538593,
final_local_scores=0.14208856, pem=-0.80476934,
similarity=12.70638]

5: 25458 Rome

[pem_pos=0, final_scores=0.53594786,
final_local_scores=0.53762305, pem=-0.2517334,
similarity=21.502422]

6: 16494 Josephus

[pem_pos=0, final_scores=0.32073814,
final_local_scores=0.27761596, pem=0.0, similarity=12.374327]

7: 25955086 Jews

[pem_pos=1, final_scores=0.40124625,
final_local_scores=0.37788314, pem=-0.80476934,
similarity=19.74296]

8: 30526 Temple_in_Jerusalem

[pem_pos=2, final_scores=0.09639402,
 final_local_scores=0.05809474, pem=-2.1650271,
 similarity=17.59713]
 9: 16043 Jerusalem
 [pem_pos=0, final_scores=0.5115747,
 final_local_scores=0.4914381, pem=-0.13239029,
 similarity=19.475164]
 10: 1095706 Jesus
 [pem_pos=0, final_scores=0.07372899,
 final_local_scores=0.031491444, pem=0.0, similarity=5.0294805]
 11: 1095706 Jesus
 [pem_pos=0, final_scores=0.18406998,
 final_local_scores=0.15359497, pem=0.0, similarity=8.673293]
 12: 25458 Rome
 [pem_pos=0, final_scores=0.49694353,
 final_local_scores=0.49501675, pem=-0.2517334,
 similarity=20.230965]
 13: 16494 Josephus
 [pem_pos=0, final_scores=0.30299044,
 final_local_scores=0.2582293, pem=0.0, similarity=11.79579]
 14: 17730 Latin
 [pem_pos=0, final_scores=0.24814491,
 final_local_scores=0.23573947, pem=-0.08185364,
 similarity=11.569782]
 15: 12108 Greece
 [pem_pos=1, final_scores=0.042557225,
 final_local_scores=0.03178978, pem=-2.3247876,
 similarity=17.68094]
 16: 26571478 History_of_the_Jews_in_the_Roman_Empire
 [pem_pos=0, final_scores=0.13078116,
 final_local_scores=0.11943331, pem=0.0, similarity=7.6538415]
 17: 17730 Latin
 [pem_pos=0, final_scores=0.31797463,
 final_local_scores=0.3120178, pem=-0.08185364,
 similarity=13.84608]

Sample: b'Michael_Youll'

[1 Michael Youll 1] -LRB- born 26 April 1939 -RRB- is a
 former [2 1 English 2] 2 cricket er . [3 Youll 3] was a right-
 handed who bowled right-arm 3 medium pace . He was born in 4 N
 ewcastle upon Tyne , 5 Northumberland . [4 Youll 4] made his 6
 first-class debut for [1 1 Warwickshire 1] , he against in
 1956 . [5 Youll 5] made 3 further appearances for [6
 Warwickshire 6] , the last of which came against the [2 2 Com
 bined Services 2] in 1957 . In his 4 first-class matches , he
 took 14 wickets at an 7 average of 21.57 , with a single .
 This came against [7 Scotland 7] in his debut match , with [8
 Youll 8] taking 5/99 in his first bowling innings in first-
 class cricket . He served in the [3 3 Royal Air Force 3]
 while at [9 Warwickshire 9] , but despite being a regular in
 the [10 Warwickshire Second XI 10] , he left the county at
 the end of the 8 1961 season having found his opportunities

limited . He joined [4 4 Northumberland 4] in 1962 , making his debut for the county in that season 's against the [5 5 Yorkshire Second XI 5] . [11 Youll 11] played 9 Minor counties cricket for [12 Northumberland 12] from 1962 to 1981 , which included over 100 [13 Minor Counties Championship 13] appearances . He made his [6 6 List A 6] debut for [14 Northumberland 14] against [7 7 Lincolnshire 7] in the 10 1971 [15 Gillette Cup 15] . He made 2 further [16 List A 16] appearances , against [8 8 Bedfordshire 8] and [9 9 Somerset 9] , both in the 11 1977 [17 Gillette Cup 17] . In his 3 [18 List A 18] matches , he scored 65 runs at a 12 batting average of 32.50 , with a high score of 36 13 not out . His son , 14 0 liver

- 1: 667292 Warwickshire_County_Cricket_Club
- 2: 6381746 Combined_Services_cricket_team
- 3: 25679 Royal_Air_Force
- 4: 3701987 Northumberland_County_Cricket_Club
- 5: 869945 Yorkshire_County_Cricket_Club
- 6: 683459 List_A_cricket
- 7: 3712244 Lincolnshire_County_Cricket_Club
- 8: 3691144 Bedfordshire_County_Cricket_Club
- 9: 1622178 Somerset_County_Cricket_Club

1: 18803164 English_people

[winner: 9316 England, pem_pos=1, final_scores=0.5039385, final_local_scores=0.55417794, pem=-1.3527626, similarity=27.98402]

gt: 18803164 English_people, pem_pos=2, final_scores=-0.41893768, final_local_scores=-0.44573137, pem=-2.2084396, similarity=2.7980375]

2: 25675557 Cricket

[winner: 330091 Rhaphidophoridae, pem_pos=10, final_scores=-1.459705, final_local_scores=-1.4160829, pem=-2.56456, similarity=-24.22255]

gt: 25675557 Cricket, pem_pos=0, final_scores=-1.8013052, final_local_scores=-1.9980025, pem=-0.07532314, similarity=-55.12505]

3: 689836 Seam_bowling

[winner: 678695 Fast_bowling, pem_pos=0, final_scores=-1.0012231, final_local_scores=-1.1468954, pem=0.0, similarity=-30.135933]

RECALL MISS: 689836 Seam_bowling]

[winner: 11531119 Joe_Pace, pem_pos=17, final_scores=-1.2262567, final_local_scores=-1.1877267, pem=-3.909537, similarity=-10.093755]

RECALL MISS: 689836 Seam_bowling]

4: 70348 Newcastle_upon_Tyne

[winner: 70348 Newcastle_upon_Tyne, pem_pos=0, final_scores=0.028193071, final_local_scores=0.057411373, pem=0.0, similarity=5.802983]

gt: 70348 Newcastle_upon_Tyne, pem_pos=0, final_scores=0.028193071, final_local_scores=0.057411373, pem=0.0, similarity=5.802983]

5: 59769 Northumberland
 [winner: 3701987 Northumberland_County_Cricket_Club,
 pem_pos=1, final_scores=0.0033082813,
 final_local_scores=-0.082417905, pem=-2.745579,
 similarity=16.561083]
 gt: 59769 Northumberland, pem_pos=0,
 final_scores=-0.010675743, final_local_scores=7.122755e-05,
 pem=-0.35131624, similarity=6.002351]

6: 11040 First-class_cricket
 [winner: 662844 First-class_citizen, pem_pos=3,
 final_scores=-1.2496186, final_local_scores=-1.1663165,
 pem=-1.5935546, similarity=-22.049498]
 gt: 11040 First-class_cricket, pem_pos=0,
 final_scores=-1.370907, final_local_scores=-1.5654764,
 pem=-0.053819276, similarity=-42.33455]

7: 875871 Bowling_average
 [winner: 875871 Bowling_average, pem_pos=0,
 final_scores=-0.9866486, final_local_scores=-1.1252048,
 pem=0.0, similarity=-29.488642]
 gt: 875871 Bowling_average, pem_pos=0,
 final_scores=-0.9866486, final_local_scores=-1.1252048,
 pem=0.0, similarity=-29.488642]

8: 7651333 1961_English_cricket_season
 [winner: 34659 1961, pem_pos=0, final_scores=-0.54831415,
 final_local_scores=-0.3876213, pem=-0.7041954,
 similarity=-3.6481562]
 RECALL MISS: 7651333 1961_English_cricket_season]
 [winner: 1340236 Eastertide, pem_pos=2,
 final_scores=-1.4985559, final_local_scores=-1.4723179,
 pem=-2.3723626, similarity=-26.945911]
 RECALL MISS: 7651333 1961_English_cricket_season]

9: 770909 Minor_counties_of_English_and_Welsh_cricket
 [winner: 897385 Morris_Minor, pem_pos=4,
 final_scores=-0.45630848, final_local_scores=-0.37704447,
 pem=-2.4046454, similarity=5.9147897]
 RECALL MISS: 770909
 Minor_counties_of_English_and_Welsh_cricket]
 [winner: 480042 Counties_of_Liberia, pem_pos=4,
 final_scores=-1.3099802, final_local_scores=-1.2125018,
 pem=-2.8773274, similarity=-16.446405]
 RECALL MISS: 770909
 Minor_counties_of_English_and_Welsh_cricket]

10: 27294383 1971_Gillette_Cup
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.33316696, final_local_scores=0.290414,
 pem=-0.08377493, similarity=13.211827]
 gt: 27294383 1971_Gillette_Cup, pem_pos=9,
 final_scores=-0.45359653, final_local_scores=-0.52069616,
 pem=-4.0866227, similarity=10.774791]
 [winner: 34748 1971, pem_pos=0, final_scores=-0.9683306,
 final_local_scores=-0.9020743, pem=-0.6167389,
 similarity=-19.47606]

RECALL MISS: 27294383 1971_Gillette_Cup]
11: 27376558 1977_Gillette_Cup
[winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=0.29038012, final_local_scores=0.24367589,
pem=-0.08377493, similarity=11.817069]
RECALL MISS: 27376558 1977_Gillette_Cup]
[winner: 34549 1977, pem_pos=0, final_scores=-0.56271076,
final_local_scores=-0.44911164, pem=-0.68815714,
similarity=-5.5703683]
RECALL MISS: 27376558 1977_Gillette_Cup]
12: 18933273 Batting_average
[winner: 15652278 Batting_(material), pem_pos=2,
final_scores=-1.0250734, final_local_scores=-0.92175746,
pem=-1.1616125, similarity=-17.100338]
gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.3078274, final_local_scores=-1.459435,
pem=-0.29142123, similarity=-37.877945]
[winner: 875871 Bowling_average, pem_pos=1,
final_scores=-1.2417233, final_local_scores=-1.403835,
pem=-0.8657023, similarity=-33.0957]
gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.3731922, final_local_scores=-1.5308361,
pem=-0.5012068, similarity=-38.86785]
13: 2430357 Not_out
[winner: 2430357 Not_out, pem_pos=0,
final_scores=-1.2255671, final_local_scores=-1.3996983,
pem=-0.008413082, similarity=-37.63432]
gt: 2430357 Not_out, pem_pos=0, final_scores=-1.2255671,
final_local_scores=-1.3996983, pem=-0.008413082,
similarity=-37.63432]
14: 31805948 Oliver_Youll
[winner: 1226463 Jackie_Oliver, pem_pos=8,
final_scores=-0.480958, final_local_scores=-0.41150644,
pem=-3.5670655, similarity=11.207796]
RECALL MISS: 31805948 Oliver_Youll]
1: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.68565696,
final_local_scores=0.64137566, pem=-1.1093602,
similarity=29.262516]
2: 6381746 Combined_Services_cricket_team
[pem_pos=0, final_scores=0.52199286,
final_local_scores=0.5240978, pem=0.0, similarity=19.729837]
3: 25679 Royal_Air_Force
[pem_pos=0, final_scores=0.041509613,
final_local_scores=0.09887038, pem=0.0, similarity=7.040202]
4: 3701987 Northumberland_County_Cricket_Club
[pem_pos=1, final_scores=0.1557451,
final_local_scores=0.087493956, pem=-2.745579,
similarity=21.631592]
5: 869945 Yorkshire_County_Cricket_Club
[pem_pos=0, final_scores=0.48625493,
final_local_scores=0.42432016, pem=0.0, similarity=16.752274]

6: 683459 List_A_cricket
[pem_pos=0, final_scores=0.22858368,
final_local_scores=0.16853938, pem=-0.020267159,
similarity=9.22948]
7: 3712244 Lincolnshire_County_Cricket_Club
[pem_pos=2, final_scores=0.43876857,
final_local_scores=0.40694565, pem=-2.1349053,
similarity=27.84373]
8: 3691144 Bedfordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.45909196,
final_local_scores=0.42386937, pem=-1.6135224,
similarity=25.513407]
9: 1622178 Somerset_County_Cricket_Club
[pem_pos=1, final_scores=0.6677136,
final_local_scores=0.62537897, pem=-2.4139037,
similarity=35.87945]
1: 32284245 Michael_Youll
[pem_pos=0, final_scores=0.19475923,
final_local_scores=0.17009622, pem=0.0, similarity=9.165724]
2: 9316 England
[pem_pos=1, final_scores=0.5039385,
final_local_scores=0.55417794, pem=-1.3527626,
similarity=27.98402]
3: 32284245 Michael_Youll
[pem_pos=0, final_scores=0.2119992,
final_local_scores=0.18892828, pem=0.0, similarity=9.72771]
4: 32284245 Michael_Youll
[pem_pos=0, final_scores=0.27601463,
final_local_scores=0.25885543, pem=0.0, similarity=11.814475]
5: 32284245 Michael_Youll
[pem_pos=0, final_scores=0.33397168,
final_local_scores=0.32216465, pem=0.0, similarity=13.703749]
6: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.68492216,
final_local_scores=0.640573, pem=-1.1093602,
similarity=29.238564]
7: 1859357 Scotland_national_cricket_team
[pem_pos=3, final_scores=0.18685402,
final_local_scores=0.123924255, pem=-3.70665,
similarity=27.9452]
8: 32284245 Michael_Youll
[pem_pos=0, final_scores=0.45214647,
final_local_scores=0.45125264, pem=0.0, similarity=17.555992]
9: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6252102,
final_local_scores=0.57534677, pem=-1.1093602,
similarity=27.292082]
10: 667292 Warwickshire_County_Cricket_Club
[pem_pos=0, final_scores=0.62514514,
final_local_scores=0.5752758, pem=0.0, similarity=21.25709]
11: 32284245 Michael_Youll
[pem_pos=0, final_scores=0.26570863,

final_local_scores=0.24759766, pem=0.0, similarity=11.478521]
 12: 3701987 Northumberland_Cricket_Club
 [pem_pos=1, final_scores=0.18739174,
 final_local_scores=0.1220631, pem=-2.745579,
 similarity=22.663204]
 13: 7562952 Minor_Counties_Cricket_Championship
 [pem_pos=0, final_scores=0.37127882,
 final_local_scores=0.3230793, pem=0.0, similarity=13.731043]
 14: 3701987 Northumberland_Cricket_Club
 [pem_pos=1, final_scores=0.28098708,
 final_local_scores=0.22430176, pem=-2.745579,
 similarity=25.71421]
 15: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.33316696,
 final_local_scores=0.290414, pem=-0.08377493,
 similarity=13.211827]
 16: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.28650463,
 final_local_scores=0.23180923, pem=-0.020267159,
 similarity=11.1175785]
 17: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.29038012,
 final_local_scores=0.24367589, pem=-0.08377493,
 similarity=11.817069]
 18: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.09608863,
 final_local_scores=0.023808688, pem=-0.020267159,
 similarity=4.9104276]

Sample: b'John_Moore'

-LRB- cricketer , _ born_1943 -RRB- John David Moore -LRB-
 19 May 1943 -- 17 November 2004 -RRB- was an [1 1 English 1] 2
 cricket er . Moore was a right-handed who bowled right-arm 3 m
 edium pace . He was born in [1 1 Newcastle-under-Lyme 1] , [2
 2 Staffordshire 2] . [2 Moore 2] made his debut for [3 3 Staf
 fordshire 3] in the 1961 against [4 4 Norfolk 4] . He played 4
 Minor counties cricket for [3 Staffordshire 3] from 1961 to
 1980 , making 123 [4 Minor Counties Championship 4]
 appearances . He made his [5 5 List A 5] debut against [6 6 Gl
 amorgan 6] in the 5 1971 [5 Gillette Cup 5] . He made 5
 further [6 List A 6] appearances for [7 Staffordshire 7] ,
 the last coming against [7 7 Sussex 7] in the 6 1978 [8 Gille
 tte Cup 8] . In his 6 [9 List A 9] matches for [10
 Staffordshire 10] , he scored 79 runs at an 7 average of
 15.80 , with a high score of 49 8 * . Playing for [11
 Staffordshire 11] allowed him to represent , making his debut
 for the team in the 9 1975 [12 Benson & Hedges Cup 12]
 against [8 8 Derbyshire 8] . He made 3 further [13 List A 13]
 appearances for the team , the last of which came against [9 9
 Lancashire 9] in the same competition . He scored 39 runs at
 an average of 13.00 , with a high score of 28 . He also made
 2 [14 List A 14] appearances for in the 10 1977 [15 Benson & H

edges Cup 15] against [10 10 Essex 10] and [11 11 Middlesex 11
1: 213585 Newcastle-under-Lyme
2: 58733 Staffordshire
3: 3712165 Staffordshire_County_Cricket_Club
4: 3690942 Norfolk_County_Cricket_Club
5: 683459 List_A_cricket
6: 1622320 Glamorgan_County_Cricket_Club
7: 1763691 Sussex_County_Cricket_Club
8: 1829984 Derbyshire_County_Cricket_Club
9: 749919 Lancashire_County_Cricket_Club
10: 1622252 Essex_County_Cricket_Club
11: 822103 Middlesex_County_Cricket_Club
1: 18803164 English_people
[winner: 9316 England, pem_pos=1,
final_scores=0.57958454, final_local_scores=0.6323169,
pem=-1.3527626, similarity=30.315844]
gt: 18803164 English_people, pem_pos=2,
final_scores=-0.41323447, final_local_scores=-0.44160825,
pem=-2.2084396, similarity=2.92108]
2: 25675557 Cricket
[winner: 330091 Rhaphidophoridae, pem_pos=10,
final_scores=-1.4626333, final_local_scores=-1.4181404,
pem=-2.56456, similarity=-24.283947]
gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.7043542, final_local_scores=-1.8982449,
pem=-0.07532314, similarity=-52.148083]
3: 689836 Seam_bowling
[winner: 689836 Seam_bowling, pem_pos=0,
final_scores=-0.9953647, final_local_scores=-1.117666,
pem=-0.53094965, similarity=-26.376284]
gt: 689836 Seam_bowling, pem_pos=0,
final_scores=-0.9953647, final_local_scores=-1.117666,
pem=-0.53094965, similarity=-26.376284]
4: 770909 Minor_counties_of_English_and_Welsh_cricket
[winner: 14465683 William_T_Minor, pem_pos=3,
final_scores=-0.4852121, final_local_scores=-0.42292893,
pem=-2.3714392, similarity=4.3649254]
RECALL MISS: 770909
Minor_counties_of_English_and_Welsh_cricket]
[winner: 4569028 Counties_of_Uganda, pem_pos=18,
final_scores=-1.3105419, final_local_scores=-1.2611611,
pem=-3.1761155, similarity=-16.273638]
RECALL MISS: 770909
Minor_counties_of_English_and_Welsh_cricket]
5: 27294383 1971_Gillette_Cup
[winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=0.34872538, final_local_scores=0.30215275,
pem=-0.08377493, similarity=13.562135]
gt: 27294383 1971_Gillette_Cup, pem_pos=9,
final_scores=-0.43957847, final_local_scores=-0.50493383,
pem=-4.0866227, similarity=11.245169]

[winner: 34748 1971, pem_pos=0, final_scores=-0.9612792,
final_local_scores=-0.89055896, pem=-0.6167389,
similarity=-19.132421]
RECALL MISS: 27294383 1971_Gillette_Cup]

6: 27387455 1978_Gillette_Cup
[winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=0.2846009, final_local_scores=0.23210654,
pem=-0.08377493, similarity=11.471816]
gt: 27387455 1978_Gillette_Cup, pem_pos=12,
final_scores=-0.5509905, final_local_scores=-0.6192908,
pem=-4.2536764, similarity=8.7409935]
[winner: 34753 1978, pem_pos=0, final_scores=-0.58095986,
final_local_scores=-0.36948764, pem=-0.7079138,
similarity=-3.0867903]
RECALL MISS: 27387455 1978_Gillette_Cup]

7: 18933273 Batting_average
[winner: 875871 Bowling_average, pem_pos=0,
final_scores=-0.9925503, final_local_scores=-1.1299347,
pem=0.0, similarity=-29.629791]
gt: 18933273 Batting_average, pem_pos=1,
final_scores=-1.263018, final_local_scores=-1.4051073,
pem=-0.6931472, similarity=-34.072052]

8: 2430357 Not_out
[winner: 2430357 Not_out, pem_pos=0,
final_scores=-0.06536527, final_local_scores=-0.13443568,
pem=-0.3529179, similarity=1.9971063]
gt: 2430357 Not_out, pem_pos=0,
final_scores=-0.06536527, final_local_scores=-0.13443568,
pem=-0.3529179, similarity=1.9971063]

9: 30982034 1975_Benson_&_Hedges_Cup
[winner: 1895594 Benson_&_Hedges_Cup, pem_pos=0,
final_scores=0.37889922, final_local_scores=0.34904507,
pem=0.0, similarity=14.505914]
gt: 30982034 1975_Benson_&_Hedges_Cup, pem_pos=2,
final_scores=-0.4770614, final_local_scores=-0.5213143,
pem=-3.988984, similarity=10.22537]
[winner: 34752 1975, pem_pos=0, final_scores=-0.9048734,
final_local_scores=-0.83276707, pem=-0.61872095,
similarity=-17.397017]
RECALL MISS: 30982034 1975_Benson_&_Hedges_Cup]

10: 37793759 1977_Benson_&_Hedges_Cup
[winner: 1895594 Benson_&_Hedges_Cup, pem_pos=0,
final_scores=0.38219315, final_local_scores=0.35264322,
pem=0.0, similarity=14.61329]
RECALL MISS: 37793759 1977_Benson_&_Hedges_Cup]
[winner: 34549 1977, pem_pos=0, final_scores=-0.67870104,
final_local_scores=-0.57441455, pem=-0.68815714,
similarity=-9.309657]
RECALL MISS: 37793759 1977_Benson_&_Hedges_Cup]

1: 213585 Newcastle-under-Lyme
[pem_pos=0, final_scores=0.14344679,
final_local_scores=0.17394069, pem=0.0, similarity=9.28045]

2: 58733 Staffordshire
[pem_pos=0, final_scores=0.1574264,
final_local_scores=0.16588274, pem=-0.06319042,
similarity=9.383624]
3: 3712165 Staffordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.322563,
final_local_scores=0.2608889, pem=-1.68797,
similarity=21.054602]
4: 3690942 Norfolk_County_Cricket_Club
[pem_pos=6, final_scores=0.24315397,
final_local_scores=0.17737406, pem=-3.4736512,
similarity=28.273165]
5: 683459 List_A_cricket
[pem_pos=0, final_scores=0.40024775,
final_local_scores=0.3482936, pem=-0.020267159,
similarity=14.593704]
6: 1622320 Glamorgan_County_Cricket_Club
[pem_pos=1, final_scores=0.80258614,
final_local_scores=0.777625, pem=-0.7903395,
similarity=31.593584]
7: 1763691 Sussex_County_Cricket_Club
[pem_pos=1, final_scores=0.6223205,
final_local_scores=0.59419423, pem=-1.5566913,
similarity=30.287184]
8: 1829984 Derbyshire_County_Cricket_Club
[pem_pos=1, final_scores=0.70163816,
final_local_scores=0.6647219, pem=-1.4828352,
similarity=31.990229]
9: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.5679049,
final_local_scores=0.5182109, pem=-1.7295884,
similarity=28.95993]
10: 1622252 Essex_County_Cricket_Club
[pem_pos=1, final_scores=0.8574843,
final_local_scores=0.833364, pem=-2.3459642,
similarity=41.716675]
11: 822103 Middlesex_County_Cricket_Club
[pem_pos=1, final_scores=0.53691345,
final_local_scores=0.48234499, pem=-1.5695794,
similarity=27.019466]
1: 9316 England
[pem_pos=1, final_scores=0.57958454,
final_local_scores=0.6323169, pem=-1.3527626,
similarity=30.315844]
2: 11037082 David_Moore_(cricket_coach)
[pem_pos=2, final_scores=0.08003397,
final_local_scores=0.13615534, pem=-1.526441,
similarity=16.453884]
3: 3712165 Staffordshire_County_Cricket_Club
[pem_pos=1, final_scores=0.30829352,
final_local_scores=0.24530163, pem=-1.68797,
similarity=20.589447]

4: 7562952 Minor_Counties_Cricket_Championship
 [pem_pos=0, final_scores=0.332987,
 final_local_scores=0.2780011, pem=0.0, similarity=12.38582]
 5: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.34872538,
 final_local_scores=0.30215275, pem=-0.08377493,
 similarity=13.562135]
 6: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.24590002,
 final_local_scores=0.17969215, pem=-0.020267159,
 similarity=9.562301]
 7: 3712165 Staffordshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.36265677,
 final_local_scores=0.30468526, pem=-1.68797,
 similarity=22.361572]
 8: 1899702 Friends_Provident_Trophy
 [pem_pos=0, final_scores=0.2846009,
 final_local_scores=0.23210654, pem=-0.08377493,
 similarity=11.471816]
 9: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.058117732,
 final_local_scores=-0.028382026, pem=-0.020267159,
 similarity=3.3529527]
 10: 3712165 Staffordshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.37733632,
 final_local_scores=0.32072046, pem=-1.68797,
 similarity=22.840096]
 11: 3712165 Staffordshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.37437457,
 final_local_scores=0.31748518, pem=-1.68797,
 similarity=22.743547]
 12: 1895594 Benson_&_Hedges_Cup
 [pem_pos=0, final_scores=0.37889922,
 final_local_scores=0.34904507, pem=0.0, similarity=14.505914]
 13: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.33216494,
 final_local_scores=0.27392352, pem=-0.020267159,
 similarity=12.374352]
 14: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.2976895,
 final_local_scores=0.23626429, pem=-0.020267159,
 similarity=11.250526]
 15: 1895594 Benson_&_Hedges_Cup
 [pem_pos=0, final_scores=0.38219315,
 final_local_scores=0.35264322, pem=0.0, similarity=14.61329]

Sample: b'Peter_Rochford'

Peter Rochford -LRB- 27 August 1928 -- 18 June 1992 -RRB-
 was an [1 1 English 1] who played 1 first-class cricket for [2
 2 Gloucestershire 2] between 1952 and 1957 . He was born in ,
 [3 3 Yorkshire 3] and died at [4 4 Stroud 4] , [1 2 Glouceste
 rshire 1] . A right-handed tail-end batsman and

wicketkeeper , Rochford played for [5 5 Yorkshire 5] 's second eleven in the in 1951 before joining [2 Gloucestershire 2] for 1952 . Unable to displace regular wicketkeeper [6 6 Andy Wilson 6] , who was a far better batsman , he played just two first-class games in 1952 and one in 1953 , and made only eight appearances in 1954 , when [3 Wilson 3] was 44 . In 1955 , however , Rochford took over as the regular wicketkeeper , playing in 30 matches and making 60 dismissals , and being awarded his county cap . His batting did not develop , however , and his highest score of the season was just 16 not out . There were a few signs of batting improvement in 1956 when , with 31 not out against [7 7 Oxford University 7] , Rochford made his highest score , and with [8 8 Bryan Wells 8] , and [9 9 Francis McHugh 9] in [4 Gloucestershire 4] 's side , Rochford often batted far higher in the order than his batting talent warranted . He again made 60 dismissals in the season . After a few matches in 1957 , he lost his place and with both [10 10 Bobby Etheridge 10] and [11 11 Barrie Meyer 11] available , he left the [5 Gloucestershire 5] staff at the end of the season . His obituary in 3 Wisden Cricketers ' [6 Almanack 6] in 1993 indicates that there was friction between Rochford and the [7 Gloucestershire 7] committee . From 1975 to 1977 , Rochford was on the 4 umpires ' list for first-class and [12 12 List A 12] matches in [8 England 8] and [9 Wales 9] . He also coached cricket and was a writer and journalist on cricket too . He collapsed and died suddenly in a

1: 9316 England

2: 1622322 Gloucestershire_County_Cricket_Club

3: 36637 Yorkshire

4: 324040 Stroud

5: 869945 Yorkshire_County_Cricket_Club

6: 11974371 Andy_Wilson_(cricketer)

7: 1622255 Oxford_University_Cricket_Club

8: 10080799 Bomber_Wells

9: 10750204 Francis_McHugh

10: 11698006 Bobby_Etheridge

11: 8027301 Barrie_Meyer

12: 683459 List_A_cricket

1: 11040 First-class_cricket

[winner: 662844 First-class_citizen, pem_pos=3, final_scores=-1.2622765, final_local_scores=-1.1846776, pem=-1.5935546, similarity=-22.597433]

gt: 11040 First-class_cricket, pem_pos=0, final_scores=-1.4063387, final_local_scores=-1.6073022, pem=-0.053819276, similarity=-43.582714]

[winner: 330091 Rhaphidophoridae, pem_pos=10, final_scores=-1.4090406, final_local_scores=-1.3542359, pem=-2.56456, similarity=-22.376913]

RECALL MISS: 11040 First-class_cricket]

2: 47751 Gloucestershire

[winner: 1622322 Gloucestershire_County_Cricket_Club,

```

pem_pos=1, final_scores=0.30390608,
final_local_scores=0.2289663, pem=-1.379793,
similarity=18.426054|
    gt: 47751 Gloucestershire, pem_pos=0,
final_scores=0.050922558, final_local_scores=0.020243093,
pem=-0.27548984, similarity=6.1919637 ]
3: 1582694 Wisden_Cricketers'_Almanack
    [winner: 1582694 Wisden_Cricketers'_Almanack, pem_pos=0,
final_scores=0.034888193, final_local_scores=-0.0064522475,
pem=0.0, similarity=3.897165|
    gt: 1582694 Wisden_Cricketers'_Almanack, pem_pos=0,
final_scores=0.034888193, final_local_scores=-0.0064522475,
pem=0.0, similarity=3.897165 ]
    [winner: 1582694 Wisden_Cricketers'_Almanack, pem_pos=0,
final_scores=-0.76535773, final_local_scores=-0.8805994,
pem=0.0, similarity=-22.189129|
    gt: 1582694 Wisden_Cricketers'_Almanack, pem_pos=0,
final_scores=-0.76535773, final_local_scores=-0.8805994,
pem=0.0, similarity=-22.189129 ]
    [winner: 2217991 Ireland_cricket_team, pem_pos=0,
final_scores=-1.456617, final_local_scores=-1.6442809,
pem=0.0, similarity=-44.978912|
    RECALL MISS: 1582694 Wisden_Cricketers'_Almanack]
    [winner: 2302349 Vedanga, pem_pos=16,
final_scores=-2.1790874, final_local_scores=-2.1724846,
pem=-4.75359, similarity=-34.890797|
    RECALL MISS: 1582694 Wisden_Cricketers'_Almanack]
4: 86701 Umpire_(cricket)
    [winner: 86701 Umpire_(cricket), pem_pos=0,
final_scores=-1.3796921, final_local_scores=-1.5217621,
pem=-0.16299397, similarity=-40.436317|
    gt: 86701 Umpire_(cricket), pem_pos=0,
final_scores=-1.3796921, final_local_scores=-1.5217621,
pem=-0.16299397, similarity=-40.436317 ]
    [winner: 2026886 Āstika_and_nāstika, pem_pos=6,
final_scores=-1.9841627, final_local_scores=-1.9663663,
pem=-4.060443, similarity=-32.509254|
    RECALL MISS: 86701 Umpire_(cricket)]
1: 9316 England
    [pem_pos=1, final_scores=0.359829,
final_local_scores=0.3477092, pem=-1.3527626,
similarity=21.822582]
2: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.47452015,
final_local_scores=0.41671538, pem=-1.379793,
similarity=24.02886]
3: 36637 Yorkshire
    [pem_pos=0, final_scores=0.28753954,
final_local_scores=0.25452194, pem=-0.31620604,
similarity=13.404733]
4: 324040 Stroud
    [pem_pos=0, final_scores=0.20449711,

```


final_local_scores=0.20892042, pem=-0.7699481,
similarity=14.511412]
5: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.5330239,
final_local_scores=0.4889472, pem=-1.7779728,
similarity=28.349764]
6: 11974371 Andy_Wilson_(cricketer)
[pem_pos=3, final_scores=0.20767467,
final_local_scores=0.25335038, pem=-1.1811988,
similarity=18.073736]
7: 1622255 Oxford_University_Cricket_Club
[pem_pos=1, final_scores=0.17226024,
final_local_scores=0.11760646, pem=-1.5137167,
similarity=15.831152]
8: 10080799 Bomber_Wells
[pem_pos=0, final_scores=0.38791853,
final_local_scores=0.45543903, pem=0.0, similarity=17.680922]
9: 10750204 Francis_McHugh
[pem_pos=0, final_scores=0.37399423,
final_local_scores=0.44776505, pem=0.0, similarity=17.451916]
10: 11698006 Bobby_Etheridge
[pem_pos=0, final_scores=0.5163776,
final_local_scores=0.5992747, pem=0.0, similarity=21.973265]
11: 8027301 Barrie_Meyer
[pem_pos=0, final_scores=0.4803269,
final_local_scores=0.5320743, pem=0.0, similarity=19.96787]
12: 683459 List_A_cricket
[pem_pos=0, final_scores=0.16593908,
final_local_scores=0.134767, pem=-0.020267159,
similarity=8.221644]
1: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.30390608,
final_local_scores=0.2289663, pem=-1.379793,
similarity=18.426054]
2: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.48476893,
final_local_scores=0.42791057, pem=-1.379793,
similarity=24.362947]
3: 11974371 Andy_Wilson_(cricketer)
[pem_pos=3, final_scores=0.20011322,
final_local_scores=0.24425346, pem=-1.1811988,
similarity=17.802267]
4: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.5841219,
final_local_scores=0.53643864, pem=-1.379793,
similarity=27.60164]
5: 1622322 Gloucestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.50274175,
final_local_scores=0.44754314, pem=-1.379793,
similarity=24.948822]
6: 1582694 Wisden_Cricketers'_Almanack
[pem_pos=0, final_scores=0.034888193,

final_local_scores=-0.0064522475, pem=0.0,
 similarity=3.897165]
 7: 1622322 Gloucestershire_County_Cricket_Club
 [pem_pos=1, final_scores=0.42280906,
 final_local_scores=0.3602289, pem=-1.379793,
 similarity=22.343191]
 8: 9316 England
 [pem_pos=0, final_scores=0.5422057,
 final_local_scores=0.546928, pem=-0.19053856,
 similarity=21.447313]
 9: 69894 Wales
 [pem_pos=0, final_scores=0.34302992,
 final_local_scores=0.3548227, pem=-0.37222588,
 similarity=16.702553]

Sample: b'Augustus_Simon_Frazer'

Augustus Simon Frazer -LRB- 1776 -- 1835 -RRB- , commanded
 the artillery at the [1 1 British 1] invasions of the [2 Río d
 e la Plata 2] -LRB- 1807 -RRB- and the Horse Artillery on [3
 Wellington 3] 's staff in the [1 1 Peninsular War 1] , and
 later during the . He was born 5 September 1776 at [4 Dunkirk
 4] , where his father , , was Assistant Commissary for the
 demolition of the fortress . He received his early education
 at the [5 High School of Edinburgh 5] , where he was a
 contemporary of [6 Lord Brougham 6] . He had only a sister ,
 Maria Ernestine , who married and lived in [7 Vienna 7] . In
 1809 , Simon Frazer married Emma Lynn , daughter of James
 Lynn , Esq , of [8 Woodbridge 8] in [9 Suffolk 9] . They had
 two sons . The elder , [10 Augustus Henry 10] , was born in
 August 1810 and obtained a commission in the Royal Artillery
 He retired from service at the rank of second captain and
 died unmarried while traveling in [11 Syria 11] in July
 1848 . The younger son , Andrew James , born in October
 1812 , obtained a Lieutenant 's [12 Commission 12] in the ,
 from which he retired , and he died , unmarried at [2 2 Ramsg
 ate 2] in July 1845 . Frazer joined the Royal Military
 Academy , at [13 Woolwich 13] , on 16 August 1790 as a
 gentleman cadet . On 18 September 1793 , he was gazetted a
 second lieutenant in the [14 Royal Artillery 14] , and at the
 end of the year joined the [15 Army 15] in [16 Flanders 16] ,
 under command of the Duke of York . In early 1794 , he was
 promoted to First Lieutenant , and attached with 2 field guns
 to the 3rd Foot Guards -LRB- Grenadiers -RRB- , where he
 served until the army returned to [17 England 17] in May
 1795 . During that time , he was present at several
 engagements , including the battles of [18 2 Tournay 18] and [19 3 Boxtel 19] . Upon return to [20 England 20] , he was
 appointed to the [21 Royal Horse Artillery 21] , and on 16
 July 1799 received the rank of Captain Lieutenant . From
 August to late November , he took part in the 4 Duke of York
 's expedition in [22 5 Holland 22] . In 1803 , Frazer
 received his rank of Captain , and in 1807 , embarked for [23

South America [23] to participate in General 's 6 expedition to [3 3 Buenos Aires 3] . Despite the disastrous results of this campaign , Frazer acquired experience and skill in his management of the transport and placement of the artillery . In 1811 , he was assigned to the 's army in the Peninsula , where he served at the battles of [24 7 Salamanca 24] and and the 8 Siege of [25 Burgos 25] . Later he was present at the 9 Battle of [26 Vitoria 26] , for which he received a special notice from Sir 10 Alexander Dickson , the overall commander of Artillery . Subsequently , he served at the blockade of Pappeluna , and the . At Pappeluna , 11 . Thomas Graham brought him to the notice of [27 Wellington 27] . At [28 San Sebastian 28] , he directed the fire from the flank , remarkable for its accuracy . He subsequently participated in the investment of [29 Bayonne 29] , at which he was wounded . At the conclusion of fighting after [30 Toulouse 30] , he returned to [31 England 31] , as a Lt. Colonel , and received the Peninsular Gold Cross , with one clasp for [32 Vittoria 32] , [33 San Sebastian 33] , [34 Nivelle 34] , [35 Nive 35] , and [36 Toulouse 36] , and was awarded . In 1815 , Frazer returned to command the Horse Artillery and was attached to [3 7 Wellington 37] 's Headquarters . He was present at the 12 Battle of [38 Quatre Bras 38] on 16 June , but his artillery did not actually arrive until after the fight . The artillery covered the retreat from [39 Quatra Bras 39] , delaying the [4 0 French 40] cavalry sufficiently to allow the infantry to withdraw . The Horse Artillery successfully dislodged the [41 French 41] from the woods surrounding Hougemont , without bringing the [42 Allied 42] troops under friendly fire . He successfully substituted 9-pounders for 6-pounders with many of the troops , and formed one troop with

1: 102485 Peninsular_War

2: 324352 Ramsgate

3: 255919 Buenos_Aires

1: 1220090 British_invasions_of_the_Río_de_la_Plata

[winner: 31717 United_Kingdom, pem_pos=0, final_scores=0.43300027, final_local_scores=0.43344563, pem=-0.48936236, similarity=19.685823]

RECALL MISS: 1220090

British_invasions_of_the_Río_de_la_Plata]

[winner: 193526 Río_de_la_Plata, pem_pos=0, final_scores=0.12842427, final_local_scores=0.13921475, pem=-0.11065125, similarity=8.845898]

gt: 1220090 British_invasions_of_the_Río_de_la_Plata, pem_pos=3, final_scores=-0.56945413, final_local_scores=-0.6425894, pem=-4.0839577, similarity=7.122761]

[winner: 14908473 Timur's_invasions_of_Georgia, pem_pos=2, final_scores=-1.368405, final_local_scores=-1.4002514, pem=-1.9081697, similarity=-27.319658]

gt: 1220090 British_invasions_of_the_Río_de_la_Plata,

pem_pos=8, final_scores=-1.5776966,
final_local_scores=-1.7457914, pem=-1.9459101,
similarity=-37.42602]
[winner: 10447732
Results_by_riding_of_the_Canadian_federal_election,_2008,
pem_pos=12, final_scores=-1.3929538,
final_local_scores=-1.287883, pem=-3.6198866,
similarity=-14.657782]
RECALL MISS: 1220090
British_invasions_of_the_Río_de_la_Plata]
2: 8643778 Battle_of_Tournay
[winner: 175362 Tournai, pem_pos=2,
final_scores=0.3031246, final_local_scores=0.2908681,
pem=-1.0160912, similarity=18.29546]
gt: 8643778 Battle_of_Tournay, pem_pos=0,
final_scores=0.17296521, final_local_scores=0.28396076,
pem=0.0, similarity=12.563668]
3: 9275450 Battle_of_Boxtel
[winner: 118709 Boxtel, pem_pos=0,
final_scores=0.1268449, final_local_scores=0.19178039,
pem=0.0, similarity=9.812822]
gt: 9275450 Battle_of_Boxtel, pem_pos=2,
final_scores=-0.037702158, final_local_scores=-0.019131824,
pem=-1.2783884, similarity=10.470856]
4: 209348 Prince_Frederick,_Duke_of_York_and_Albany
[winner: 125267 Prince_Andrew,_Duke_of_York, pem_pos=0,
final_scores=-0.10564722, final_local_scores=-0.12328644,
pem=-0.5877866, similarity=3.6070747]
gt: 209348 Prince_Frederick,_Duke_of_York_and_Albany,
pem_pos=7, final_scores=-0.39224195,
final_local_scores=-0.467712, pem=-1.8331008,
similarity=0.10094141]
5: 22608854 Anglo-Russian_invasion_of_Holland
[winner: 21148 Netherlands, pem_pos=1,
final_scores=0.10311706, final_local_scores=0.08211911,
pem=-2.7840104, similarity=21.680193]
RECALL MISS: 22608854 Anglo-Russian_invasion_of_Holland]
6: 1220090 British_invasions_of_the_Río_de_la_Plata
[winner: 992161 Narváez_expedition, pem_pos=1,
final_scores=-1.1664492, final_local_scores=-1.2219474,
pem=-1.6234654, similarity=-23.546978]
gt: 1220090 British_invasions_of_the_Río_de_la_Plata,
pem_pos=18, final_scores=-1.5796498,
final_local_scores=-1.747925, pem=-3.6318839,
similarity=-28.321104]
7: 157351 Battle_of_Salamanca
[winner: 23955286 Salamanca, pem_pos=0,
final_scores=0.20475648, final_local_scores=0.19109398,
pem=-0.45059586, similarity=12.242749]
gt: 157351 Battle_of_Salamanca, pem_pos=3,
final_scores=-0.31998897, final_local_scores=-0.35470384,
pem=-2.4156506, similarity=6.6413264]

8: 14066070 Siege_of_Burgos
[winner: 44730 Burgos, pem_pos=0,
final_scores=0.32486957, final_local_scores=0.32017812,
pem=-0.46851826, similarity=16.192343]
gt: 14066070 Siege_of_Burgos, pem_pos=7,
final_scores=-0.42497215, final_local_scores=-0.4484888,
pem=-3.3442516, similarity=8.892473]
[winner: 26905 Siege, pem_pos=0,
final_scores=-0.13772336, final_local_scores=-0.16082698,
pem=-0.75229585, similarity=3.3814166]
RECALL MISS: 14066070 Siege_of_Burgos]

9: 157766 Battle_of_Vitoria
[winner: 80735 Vitoria-Gasteiz, pem_pos=0,
final_scores=0.14841639, final_local_scores=0.16459563,
pem=-0.33011132, similarity=10.796772]
gt: 157766 Battle_of_Vitoria, pem_pos=1,
final_scores=-0.08817862, final_local_scores=-0.11908786,
pem=-1.0888104, similarity=6.457014]
[winner: 16649099 Mount_Battle, pem_pos=5,
final_scores=-0.9120561, final_local_scores=-0.8302401,
pem=-3.124515, similarity=-3.6947086]
RECALL MISS: 157766 Battle_of_Vitoria]

10: 2821734 Alexander_Dickson_(British_Army_officer)
[winner: 2821734
Alexander_Dickson_(British_Army_officer), pem_pos=0,
final_scores=0.014565513, final_local_scores=0.04701945,
pem=0.0, similarity=5.492867]
gt: 2821734 Alexander_Dickson_(British_Army_officer),
pem_pos=0, final_scores=0.014565513,
final_local_scores=0.04701945, pem=0.0, similarity=5.492867]

11: 1914508 Thomas_Graham,_1st_Baron_Lynedoch
[winner: 1914508 Thomas_Graham,_1st_Baron_Lynedoch,
pem_pos=1, final_scores=-0.010228723,
final_local_scores=-0.027858049, pem=-1.1688417,
similarity=9.614717]
gt: 1914508 Thomas_Graham,_1st_Baron_Lynedoch,
pem_pos=1, final_scores=-0.010228723,
final_local_scores=-0.027858049, pem=-1.1688417,
similarity=9.614717]
[winner: 249033 Multiplication_sign, pem_pos=0,
final_scores=-0.8708795, final_local_scores=-0.8114935,
pem=0.0, similarity=-20.126873]
RECALL MISS: 1914508 Thomas_Graham,_1st_Baron_Lynedoch]

12: 675526 Battle_of_Quatre_Bras
[winner: 675526 Battle_of_Quatre_Bras, pem_pos=0,
final_scores=0.06370132, final_local_scores=0.07466109,
pem=0.0, similarity=6.317749]
gt: 675526 Battle_of_Quatre_Bras, pem_pos=0,
final_scores=0.06370132, final_local_scores=0.07466109,
pem=0.0, similarity=6.317749]
[winner: 16649099 Mount_Battle, pem_pos=5,
final_scores=-0.8663863, final_local_scores=-0.7803528,

pem=-3.124515, similarity=-2.2059722|
RECALL MISS: 675526 Battle_of_Quatre_Bras]

1: 102485 Peninsular_War
[pem_pos=0, final_scores=0.22272073,
final_local_scores=0.1849261, pem=0.0, similarity=9.608276]

2: 324352 Ramsgate
[pem_pos=0, final_scores=0.3546604,
final_local_scores=0.3597714, pem=0.0, similarity=14.826009]

3: 255919 Buenos_Aires
[pem_pos=0, final_scores=0.35508233,
final_local_scores=0.35352886, pem=-0.12597719,
similarity=15.324802]

1: 31717 United_Kingdom
[pem_pos=0, final_scores=0.43300027,
final_local_scores=0.43344563, pem=-0.48936236,
similarity=19.685823]

2: 193526 Río_de_la_Plata
[pem_pos=0, final_scores=0.12842427,
final_local_scores=0.13921475, pem=-0.11065125,
similarity=8.845898]

3: 8474 Arthur_Wellesley,_1st_Duke_of_Wellington
[pem_pos=0, final_scores=0.16716616,
final_local_scores=0.12012747, pem=0.0, similarity=7.6745563]

4: 89922 Dunkirk
[pem_pos=0, final_scores=0.22081195,
final_local_scores=0.18902189, pem=-0.6066502,
similarity=13.029561]

5: 7326278 Royal_High_School,_Edinburgh
[pem_pos=0, final_scores=0.3086658,
final_local_scores=0.32300285, pem=0.0, similarity=13.728762]

6: 232016 Henry_Brougham,_1st_Baron_Brougham_and_Vaux
[pem_pos=0, final_scores=0.1412882,
final_local_scores=0.18066451, pem=0.0, similarity=9.481102]

7: 55866 Vienna
[pem_pos=0, final_scores=0.63303775,
final_local_scores=0.64531946, pem=-0.12517071,
similarity=24.02803]

8: 99648 Woodbridge,_Suffolk
[pem_pos=0, final_scores=0.27130967,
final_local_scores=0.28457025, pem=-0.5660508,
similarity=15.66013]

9: 27886 Suffolk
[pem_pos=0, final_scores=0.40103894,
final_local_scores=0.41101736, pem=-0.3275355,
similarity=18.136478]

10: 1235105 George_Augustus_Henry_Sala
[pem_pos=0, final_scores=0.08552514,
final_local_scores=0.1700601, pem=0.0, similarity=9.164645]

11: 7515849 Syria
[pem_pos=0, final_scores=0.47879136,
final_local_scores=0.5386052, pem=-0.26162508,
similarity=21.585526]

12: 9974 European_Commission
[pem_pos=0, final_scores=0.13379411,
final_local_scores=0.19391847, pem=-0.76247066,
similarity=14.023061]
13: 94398 Woolwich
[pem_pos=0, final_scores=0.22799487,
final_local_scores=0.21150926, pem=-0.22972013,
similarity=11.650825]
14: 83400 Royal_Artillery
[pem_pos=0, final_scores=0.1401961,
final_local_scores=0.11183621, pem=0.0, similarity=7.4271283]
15: 32087 United_States_Army
[pem_pos=0, final_scores=0.03743802,
final_local_scores=-0.003947586, pem=-1.138116,
similarity=10.16116]
16: 10878 Flanders
[pem_pos=0, final_scores=0.32887292,
final_local_scores=0.2909949, pem=-0.23052187,
similarity=14.027195]
17: 9316 England
[pem_pos=1, final_scores=0.5393631,
final_local_scores=0.534705, pem=-0.19053856,
similarity=21.082554]
18: 175362 Tournai
[pem_pos=2, final_scores=0.3031246,
final_local_scores=0.2908681, pem=-1.0160912,
similarity=18.29546]
19: 118709 Boxtel
[pem_pos=0, final_scores=0.1268449,
final_local_scores=0.19178039, pem=0.0, similarity=9.812822]
20: 9316 England
[pem_pos=1, final_scores=0.46897572,
final_local_scores=0.4577797, pem=-0.19053856,
similarity=18.786951]
21: 2173230 Royal_Horse_Artillery
[pem_pos=0, final_scores=0.06811331,
final_local_scores=0.08590926, pem=0.0, similarity=6.6534166]
22: 21148 Netherlands
[pem_pos=1, final_scores=0.10311706,
final_local_scores=0.08211911, pem=-2.7840104,
similarity=21.680193]
23: 26769 South_America
[pem_pos=0, final_scores=0.36523253,
final_local_scores=0.41803932, pem=0.0, similarity=16.56484]
24: 23955286 Salamanca
[pem_pos=0, final_scores=0.20475648,
final_local_scores=0.19109398, pem=-0.45059586,
similarity=12.242749]
25: 44730 Burgos
[pem_pos=0, final_scores=0.32486957,
final_local_scores=0.32017812, pem=-0.46851826,
similarity=16.192343]

26: 80735 Vitoria-Gasteiz
[pem_pos=0, final_scores=0.14841639,
final_local_scores=0.16459563, pem=-0.33011132,
similarity=10.796772]
27: 33804 Wellington
[pem_pos=0, final_scores=0.21207075,
final_local_scores=0.20091617, pem=-0.9265605,
similarity=15.124232]
28: 11178687 San Sebastián
[pem_pos=0, final_scores=0.4212454,
final_local_scores=0.4272117, pem=-0.33299246,
similarity=18.649427]
29: 4741 Bayonne
[pem_pos=0, final_scores=0.2720853,
final_local_scores=0.247421, pem=-0.35630593,
similarity=13.410895]
30: 45371 Toulouse
[pem_pos=0, final_scores=0.31750065,
final_local_scores=0.29490206, pem=-0.21018569,
similarity=14.033201]
31: 9316 England
[pem_pos=1, final_scores=0.5445939,
final_local_scores=0.5403811, pem=-0.19053856,
similarity=21.25194]
32: 157766 Battle_of_Vitoria
[pem_pos=1, final_scores=0.09367816,
final_local_scores=0.083666995, pem=-0.9315582,
similarity=11.652461]
33: 11178687 San Sebastián
[pem_pos=0, final_scores=0.441759,
final_local_scores=0.44961965, pem=-0.33299246,
similarity=19.318125]
34: 1300222 Robert_Nivelle
[pem_pos=2, final_scores=0.048783883,
final_local_scores=0.07041219, pem=-0.871267,
similarity=10.929038]
35: 8744414 Nive
[pem_pos=0, final_scores=0.06514491,
final_local_scores=0.16467103, pem=0.0, similarity=9.003825]
36: 45371 Toulouse
[pem_pos=0, final_scores=0.27209067,
final_local_scores=0.24529853, pem=-0.21018569,
similarity=12.552933]
37: 8474 Arthur_Wellesley,_1st_Duke_of_Wellington
[pem_pos=0, final_scores=0.29070765,
final_local_scores=0.2550778, pem=0.0, similarity=11.701743]
38: 675526 Battle_of_Quatre_Bras
[pem_pos=0, final_scores=0.06370132,
final_local_scores=0.07466109, pem=0.0, similarity=6.317749]
39: 675526 Battle_of_Quatre_Bras
[pem_pos=0, final_scores=0.11557816,
final_local_scores=0.13132864, pem=0.0, similarity=8.0088215]

40: 5843419 France
[pem_pos=0, final_scores=0.6323691,
final_local_scores=0.6360089, pem=-0.80223274,
similarity=27.432154]
41: 5843419 France
[pem_pos=0, final_scores=0.49638194,
final_local_scores=0.48746365, pem=-0.80223274,
similarity=22.999268]
42: 2198844 Allies_of_World_War_II
[pem_pos=0, final_scores=0.13766275,
final_local_scores=0.14552104, pem=-0.47402114,
similarity=11.010152]

Sample: b'Gran_Omar'

Omar Jose Navarro , known artistically as [1 Gran Omar 1] , is a [2 1 Puerto Rican 2] rapper , singer-songwriter , and record producer . He has received production and writing on credits on 2 various albums between 1996 and 2006 by [1 1 Ivy Queen 1] his former wife . These albums including 3 En Mi Imperio -LRB- 1996 -RRB- , 4 The Original Rude Girl -LRB- 1998 -RRB- , 5 Diva -LRB- 2003 -RRB- , 6 Real -LRB- 2004 -RRB- and 7 Flashback -LRB- 2005 -RRB- . These albums have been met with commercial and critical success within the Latin community . Several singles from these albums have featured [3 Gran Omar 3] including `` [4 Quiero Saber 4] '' and `` Guillaera '' among others . They met while in the all-male hip-hop group `` The Noise '' . He received his first single release writing credit on [5 Ivy Queen 5] 's '' 8 In The Zone '' featuring [2 2 Wyclef Jean 2] in 1998 . The song peaked at # 38 on the chart . He also received production credits on [6 Queen 6] 's third and fourth studio albums : 9 Diva -LRB- 2003 -RRB- and 10 Real -LRB- 2004 -RRB- . Diva peaked at # 24 on the chart . On the 11 Billboard Top Heatseekers -LRB- South Atlantic -RRB- chart , the album peaked at # 8 . On the Billboard Reggae Albums chart , the album peaked at # 4 . On the [7 Billboard Tropical Albums 7] chart , the album peaked at # 1 , spending four nonconsecutive weeks at the top with a total of 86 weeks on the chart , becoming the eighth best-selling Tropical Album of 2004 . While , Real , on the chart , peaked at # 25 , one position higher than Diva . On the [3 3 Reggae Albums 3] chart , the album peaked at # 4 , her highest peak on the chart with Diva reaching the same position , while spending a total of 17 consecutive weeks on the chart . On the [8 Billboard Tropical Albums 8] chart , the album peaked at # 6 , not being able to reach [9 Divas 9] peak position of # 1 . 12 Flashback , a retrospective of [10 Queen 10] 's previous hits along with a few new ones featured `` [11 Quiero Saber 11] '' which is credited to [12 Gran Omar 12] . In its first week of release , Flashback scanned 5000 sales ; however failed to debut on the [4 4 Billboard 200 4] . As of March 2007 , the album has scanned over 104,000 sales in the [13 United States 13] and [14 Puerto Rico 14]

alone . It was able to break their peak at # 24 on the [15 Billboard Top Latin Albums 15] chart when it reached # 10 . On [5 5 Billboard Top Heatseekers 5] chart , the album reached # 7 . It reached # 2 on both the [16 Billboard Top Heatseekers 16] for the Pacific and South Atlantic areas . It reached # 3 on the Billboard Latin Rhythm Albums becoming their debut on the newly instated chart . After this chart was instated , it was revealed that reggaetón albums could no longer appear on the Billboard Reggae Albums and [17 Billboard Tropical Albums 17] charts , deeming Flashback ineligible . Omar is credited with the following songs : '' [6 6 Baila Así 6] '' , '' 13 Dee Jay '' , the fourth single from Diva '' 14 Guillaera '' , '' 15 Matando '' , '' 16 Money Making '' , and the second single from 17 Tony Touch 's album The ReggaeTony Album : '' 18 Saca La Semilla

```
1: 1104730 Ivy_Queen
2: 302511 Wyclef_Jean
3: 1290194 Billboard_charts
4: 730350 Billboard_200
5: 2898379 Top_Heatseekers
6: 11702294 Real_(Ivy_Queen_album)
1: 82536 Puerto_Rican_people
  [winner: 23041 Puerto_Rico, pem_pos=0,
  final_scores=0.254237, final_local_scores=0.2934931,
  pem=-0.5221348, similarity=15.687581|
  gt: 82536 Puerto_Rican_people, pem_pos=1,
  final_scores=-0.007296577, final_local_scores=-0.008752868,
  pem=-0.83720803, similarity=8.381377 ]
2: 23911312 Ivy_Queen_discography
  [winner: 21565121 The_Police_discography, pem_pos=3,
  final_scores=-1.1239722, final_local_scores=-1.0636643,
  pem=-2.1882758, similarity=-15.751968|
  RECALL MISS: 23911312 Ivy_Queen_discography]
  [winner: 13652176 Various, pem_pos=0,
  final_scores=-1.4020778, final_local_scores=-1.4895351,
  pem=-1.5107073, similarity=-32.14552|
  RECALL MISS: 23911312 Ivy_Queen_discography]
3: 6583590 En_Mi_Imperio
  [winner: 6583590 En_Mi_Imperio, pem_pos=0,
  final_scores=-0.09201522, final_local_scores=-0.10629548,
  pem=0.0, similarity=0.91764385|
  gt: 6583590 En_Mi_Imperio, pem_pos=0,
  final_scores=-0.09201522, final_local_scores=-0.10629548,
  pem=0.0, similarity=0.91764385 ]
4: 11702108 The_Original_Rude_Girl
  [winner: 11702108 The_Original_Rude_Girl, pem_pos=0,
  final_scores=-0.0040697157, final_local_scores=-0.028737374,
  pem=0.0, similarity=3.2321324|
  gt: 11702108 The_Original_Rude_Girl, pem_pos=0,
  final_scores=-0.0040697157, final_local_scores=-0.028737374,
  pem=0.0, similarity=3.2321324 ]
5: 11702188 Diva_(Ivy_Queen_album)
```

[winner: 98032 Diva, pem_pos=0, final_scores=-0.28242373, final_local_scores=-0.30735022, pem=-1.1241698, similarity=1.0311791]
 gt: 11702188 Diva_(Ivy_Queen_album), pem_pos=7, final_scores=-0.6311259, final_local_scores=-0.76305705, pem=-2.7711737, similarity=-3.611357]
 6: 11702294 Real_(Ivy_Queen_album)
 [winner: 34201323 Brazilian_real, pem_pos=0, final_scores=-0.27902636, final_local_scores=-0.24212773, pem=-1.3368475, similarity=4.1341233]
 gt: 11702294 Real_(Ivy_Queen_album), pem_pos=17, final_scores=-0.6931295, final_local_scores=-0.81631446, pem=-3.4678988, similarity=-1.4117664]
 7: 11702377 Flashback_(Ivy_Queen_album)
 [winner: 1351691 Flashback_(Electric_Light_Orchestra_album), pem_pos=0, final_scores=-0.5557611, final_local_scores=-0.5400892, pem=-1.5402617, similarity=-3.651443]
 gt: 11702377 Flashback_(Ivy_Queen_album), pem_pos=9, final_scores=-0.60039014, final_local_scores=-0.71665365, pem=-2.4672263, similarity=-3.8795002]
 8: 37028846 In_the_Zone_(song)
 [winner: 365930 In_the_Zone, pem_pos=0, final_scores=0.010540709, final_local_scores=-0.02392663, pem=0.0, similarity=3.3756945]
 gt: 37028846 In_the_Zone_(song), pem_pos=2, final_scores=-0.5740096, final_local_scores=-0.65138245, pem=-3.394029, similarity=3.1084187]
 9: 11702188 Diva_(Ivy_Queen_album)
 [winner: 98032 Diva, pem_pos=0, final_scores=-0.1975263, final_local_scores=-0.21461266, pem=-1.1241698, similarity=3.7986531]
 gt: 11702188 Diva_(Ivy_Queen_album), pem_pos=7, final_scores=-0.5614307, final_local_scores=-0.6869256, pem=-2.7711737, similarity=-1.3394427]
 10: 11702294 Real_(Ivy_Queen_album)
 [winner: 34201323 Brazilian_real, pem_pos=0, final_scores=-0.28616655, final_local_scores=-0.2499273, pem=-1.3368475, similarity=3.9013686]
 gt: 11702294 Real_(Ivy_Queen_album), pem_pos=17, final_scores=-0.61618924, final_local_scores=-0.7322689, pem=-3.4678988, similarity=1.0963202]
 11: 2898379 Top_Heatseekers
 [winner: 2898379 Top_Heatseekers, pem_pos=0, final_scores=-0.088440716, final_local_scores=-0.15681197, pem=0.0, similarity=-0.5898689]
 gt: 2898379 Top_Heatseekers, pem_pos=0, final_scores=-0.088440716, final_local_scores=-0.15681197, pem=0.0, similarity=-0.5898689]
 12: 11702377 Flashback_(Ivy_Queen_album)
 [winner: 1351691 Flashback_(Electric_Light_Orchestra_album), pem_pos=0,

```

final_scores=-0.5240734, final_local_scores=-0.5054752,
pem=-1.5402617, similarity=-2.6184928|
    gt: 11702377 Flashback_(Ivy_Queen_album), pem_pos=9,
final_scores=-0.5810263, final_local_scores=-0.69550157,
pem=-2.4672263, similarity=-3.2482803 ]
13: 11702294 Real_(Ivy_Queen_album)
    [winner: 1096751 Dee_Jay, pem_pos=0,
final_scores=-0.2056773, final_local_scores=-0.114124805,
pem=-0.028573394, similarity=0.8393878|
    gt: 11702294 Real_(Ivy_Queen_album), pem_pos=2,
final_scores=-0.95051944, final_local_scores=-1.0974739,
pem=-5.6489744, similarity=2.058902 ]
14: 11299107 Guillaera
    [winner: 11299107 Guillaera, pem_pos=0,
final_scores=-0.28705233, final_local_scores=-0.11129734,
pem=0.0, similarity=0.7683784|
    gt: 11299107 Guillaera, pem_pos=0,
final_scores=-0.28705233, final_local_scores=-0.11129734,
pem=0.0, similarity=0.7683784 ]
15: 11702294 Real_(Ivy_Queen_album)
    [winner: 11702294 Real_(Ivy_Queen_album), pem_pos=0,
final_scores=-0.042682096, final_local_scores=-0.10579945,
pem=0.0, similarity=0.9324462|
    gt: 11702294 Real_(Ivy_Queen_album), pem_pos=0,
final_scores=-0.042682096, final_local_scores=-0.10579945,
pem=0.0, similarity=0.9324462 ]
16: 11702188 Diva_(Ivy_Queen_album)
    [winner: 11702188 Diva_(Ivy_Queen_album), pem_pos=0,
final_scores=-0.119366944, final_local_scores=-0.20403811,
pem=0.0, similarity=-1.9991912|
    gt: 11702188 Diva_(Ivy_Queen_album), pem_pos=0,
final_scores=-0.119366944, final_local_scores=-0.20403811,
pem=0.0, similarity=-1.9991912 ]
17: 979215 Tony_Touch
    [winner: 979215 Tony_Touch, pem_pos=0,
final_scores=-0.09695692, final_local_scores=-0.11710962,
pem=0.0, similarity=0.59492826|
    gt: 979215 Tony_Touch, pem_pos=0,
final_scores=-0.09695692, final_local_scores=-0.11710962,
pem=0.0, similarity=0.59492826 ]
18: 979215 Tony_Touch
    [winner: 979215 Tony_Touch, pem_pos=0,
final_scores=-0.01977472, final_local_scores=-0.03279978,
pem=0.0, similarity=3.110902|
    gt: 979215 Tony_Touch, pem_pos=0,
final_scores=-0.01977472, final_local_scores=-0.03279978,
pem=0.0, similarity=3.110902 ]
1: 1104730 Ivy_Queen
    [pem_pos=0, final_scores=0.32914907,
final_local_scores=0.28573403, pem=0.0, similarity=12.616587]
2: 302511 Wyclef_Jean
    [pem_pos=0, final_scores=0.23607646,

```

final_local_scores=0.21110371, pem=0.0, similarity=10.389468]
3: 1290194 Billboard_charts
[pem_pos=0, final_scores=0.17707275,
final_local_scores=0.1259127, pem=0.0, similarity=7.8471985]
4: 730350 Billboard_200
[pem_pos=0, final_scores=0.0703056,
final_local_scores=0.0151408315, pem=0.0,
similarity=4.5415454]
5: 2898379 Top_Heatseekers
[pem_pos=0, final_scores=0.16203506,
final_local_scores=0.123039454, pem=0.0, similarity=7.761456]
6: 11702294 Real_(Ivy_Queen_album)
[pem_pos=0, final_scores=0.17091681,
final_local_scores=0.13385403, pem=0.0, similarity=8.084185]
1: 39139456 Gran_Omar
[pem_pos=0, final_scores=0.044322237,
final_local_scores=0.0021709055, pem=0.0, similarity=4.154497]
2: 23041 Puerto_Rico
[pem_pos=0, final_scores=0.254237,
final_local_scores=0.2934931, pem=-0.5221348,
similarity=15.687581]
3: 39139456 Gran_Omar
[pem_pos=0, final_scores=0.104841575,
final_local_scores=0.06827909, pem=0.0, similarity=6.127297]
4: 11702188 Diva_(Ivy_Queen_album)
[pem_pos=0, final_scores=0.11470406,
final_local_scores=0.05676183, pem=-0.2876821,
similarity=7.348059]
5: 1104730 Ivy_Queen
[pem_pos=0, final_scores=0.31530607,
final_local_scores=0.2706126, pem=0.0, similarity=12.165333]
6: 1104730 Ivy_Queen
[pem_pos=0, final_scores=0.22037785,
final_local_scores=0.16686347, pem=0.0, similarity=9.069252]
7: 27615530 Tropical_Albums
[pem_pos=0, final_scores=0.08983083,
final_local_scores=0.06721583, pem=0.0, similarity=6.095567]
8: 27615530 Tropical_Albums
[pem_pos=0, final_scores=0.13992508,
final_local_scores=0.12193617, pem=0.0, similarity=7.7285314]
9: 580533 WWE_Diva
[pem_pos=0, final_scores=0.14633773,
final_local_scores=0.21250647, pem=0.0, similarity=10.43133]
10: 1104730 Ivy_Queen
[pem_pos=0, final_scores=0.18551557,
final_local_scores=0.1287817, pem=0.0, similarity=7.932816]
11: 11702188 Diva_(Ivy_Queen_album)
[pem_pos=0, final_scores=0.0855519,
final_local_scores=0.024917528, pem=-0.2876821,
similarity=6.3977613]
12: 39139456 Gran_Omar
[pem_pos=0, final_scores=0.13744967,

final_local_scores=0.103898466, pem=0.0, similarity=7.1902504]
 13: 3434750 United_States
 [pem_pos=0, final_scores=0.3693545,
 final_local_scores=0.4459517, pem=-0.073457584,
 similarity=17.797276]
 14: 23041 Puerto_Rico
 [pem_pos=0, final_scores=0.22810064,
 final_local_scores=0.26494312, pem=-0.2967188,
 similarity=13.609747]
 15: 17393449 Billboard_Top_Latin_Albums
 [pem_pos=0, final_scores=0.18622981,
 final_local_scores=0.14809653, pem=0.0, similarity=8.509209]
 16: 2898379 Top_Heatseekers
 [pem_pos=0, final_scores=0.15150456,
 final_local_scores=0.11153647, pem=0.0, similarity=7.418184]
 17: 27615530 Tropical_Albums
 [pem_pos=0, final_scores=0.15118174,
 final_local_scores=0.13423237, pem=0.0, similarity=8.095474]

Sample: b'Jessica_Mauboy_discography'

The 1 discography of [1 Australian 1] 2 R&B recording artist [1 1 Jessica Mauboy 1] consists of two studio album s , one 3 live album , fourteen 4 singles , sixteen 5 music video s and twenty-three . [2 Mauboy 2] was the runner-up on the 6 fourth season of [2 2 Australian Idol 2] in 2006 , and subsequently signed a contract with [3 3 Sony Music Australia 3] . She released her debut live album 7 The Journey in February 2007 , which features songs she performed as part of the top twelve on [3 Australian Idol 3] . The album debuted at number four on the [4 4 ARIA Albums Chart 4] and was 8 certified gold by the [5 5 Australian Recording Industry Association 5] - LRB- ARIA -RRB- , denoting shipments of 35,000 copies . [4 Mauboy 4] briefly became a member of the [5 Australian 5] pop girl group [6 6 Young Divas 6] later that year , before resuming her solo career early in 2008 . [6 Mauboy 6] released her debut studio album 9 Been Waiting in November 2008 , which peaked at number eleven on the [7 ARIA Albums Chart 7] and spent fifty-nine weeks in the top fifty . The album was certified double platinum , denoting shipments of 140,000 copies , and became the second highest-selling [8 Australian 8] album of 2009 . The album was preceded by the 10 lead single '' 11 Running Back '' , which features [9 American 9] rapper [7 7 Flo Rida 7] . The song peaked at number three on the [10 ARIA Singles Chart 10] and was certified double platinum . The second single '' 12 Burn '' became [11 Mauboy 11] 's first number-one song on the chart and was certified platinum . [12 Been Waiting 12] also yielded three other successful singles which were certified gold , the '' [13 13 Been Waiting 13] '' , '' 14 Because '' and '' 15 Up/Down '' . [14 Mauboy 14] 's second studio album 1 6 Get 'Em Girls was released in November 2010 , which debuted at number six and was certified gold . As a result , the

album was less commercially successful compared to its predecessor . The album 's title track ' ' 17 Get 'Em Girls ' ' which features [15 American 15] rapper [8 8 Snoop Dogg 8] , was released as the lead single and peaked at number nineteen . Subsequent singles ' ' 18 Saturday Night ' ' and ' ' 19 Inescapable ' ' both peaked within the top ten and received double platinum certifications . ' ' 20 Galaxy ' ' , the final single from Get 'Em Girls , features [16 Australian 16] recording artist [9 9 Stan Walker 9] . It became [17 Mauboy 17] 's first song to chart in [18 New Zealand 18] , where it reached the top forty and was certified gold by the [19 Recording Industry Association of New Zealand 19] -LRB- [20 RIANZ 20] -RRB- . To promote the film 21 The Sapphires -LRB- 2012 -RRB- , in which [21 Mauboy 21] portrays the character of Julie McCrae , she released ' ' 22 Gotcha ' ' as the lead single from the 23 soundtrack of the same name

1: 6699718 Jessica_Mauboy

2: 372646 Australian_Idol

3: 17452021 Sony_Music_Australia

4: 3002226 ARIA_Charts

5: 308882 Australian_Recording_Industry_Association

6: 5307421 Young_Divas

7: 13994263 Flo_Rida

8: 88923 Snoop_Dogg

9: 23116760 Stan_Walker

1: 510969 Discography

[winner: 9262774 ThaMuseMeant_discography, pem_pos=9, final_scores=-1.0366722, final_local_scores=-0.93221337, pem=-2.8332133, similarity=-8.3219385]

RECALL MISS: 510969 Discography]

2: 2529904 Contemporary_R&B

[winner: 26168 Rhythm_and_blues, pem_pos=0, final_scores=0.021238193, final_local_scores=-0.014931247, pem=-0.42558512, similarity=5.958533]

gt: 2529904 Contemporary_R&B, pem_pos=1, final_scores=-0.16320781, final_local_scores=-0.20915943, pem=-1.2034804, similarity=4.3926916]

3: 988516 Live_album

[winner: 988516 Live_album, pem_pos=0, final_scores=-0.9901878, final_local_scores=-1.0880854, pem=-0.010225603, similarity=-28.325317]

gt: 988516 Live_album, pem_pos=0, final_scores=-0.9901878, final_local_scores=-1.0880854, pem=-0.010225603, similarity=-28.325317]

4: 156540 Single_(music)

[winner: 156540 Single_(music), pem_pos=0, final_scores=-1.4711006, final_local_scores=-1.6649368, pem=-0.33920756, similarity=-43.75066]

gt: 156540 Single_(music), pem_pos=0, final_scores=-1.4711006, final_local_scores=-1.6649368, pem=-0.33920756, similarity=-43.75066]

5: 90138 Music_video

[winner: 1406689 She's_the_One_(World_Party_song),
 pem_pos=5, final_scores=-1.2434306,
 final_local_scores=-1.3408263, pem=-2.079155,
 similarity=-24.616451|
 gt: 90138 Music_video, pem_pos=0,
 final_scores=-1.3940167, final_local_scores=-1.5811937,
 pem=0.0, similarity=-43.096264]
 6: 17173375 Australian_Idol_(season_4)
 [winner: 31656
 Fourth_Amendment_to_the_United_States_Constitution,
 pem_pos=1, final_scores=-1.1799237,
 final_local_scores=-1.1244478, pem=-1.5976565,
 similarity=-20.777746|
 RECALL MISS: 17173375 Australian_Idol_(season_4)]
 [winner: 24873453 Season, pem_pos=1,
 final_scores=-1.5445892, final_local_scores=-1.5944078,
 pem=-1.3735892, similarity=-36.020805|
 RECALL MISS: 17173375 Australian_Idol_(season_4)]
 7: 12112003 The_Journey_(Jessica_Mauboy_album)
 [winner: 12730440 Grand_Theft_Auto_IV_soundtrack,
 pem_pos=2, final_scores=-0.39580262,
 final_local_scores=-0.3816172, pem=-1.9045819,
 similarity=3.0589077|
 gt: 12112003 The_Journey_(Jessica_Mauboy_album),
 pem_pos=3, final_scores=-0.49730235,
 final_local_scores=-0.51094556, pem=-1.9780673,
 similarity=-0.4008839]
 8: 308242 Music_recording_sales_certification
 [winner: 308159 RIAA_certification, pem_pos=0,
 final_scores=-0.9913394, final_local_scores=-1.1178449,
 pem=-0.7071209, similarity=-25.423576|
 gt: 308242 Music_recording_sales_certification,
 pem_pos=1, final_scores=-1.1746769,
 final_local_scores=-1.3332055, pem=-1.139748,
 similarity=-29.497671]
 [winner: 12240 Gold, pem_pos=0, final_scores=-1.0491619,
 final_local_scores=-1.0520585, pem=-0.22135684,
 similarity=-26.102037|
 gt: 308242 Music_recording_sales_certification,
 pem_pos=1, final_scores=-1.2453345,
 final_local_scores=-1.4103882, pem=-1.8607677,
 similarity=-27.879944]
 9: 19199205 Been_Waiting
 [winner: 19199205 Been_Waiting, pem_pos=0,
 final_scores=0.02560629, final_local_scores=-0.017458826,
 pem=0.0, similarity=3.5687068|
 gt: 19199205 Been_Waiting, pem_pos=0,
 final_scores=0.02560629, final_local_scores=-0.017458826,
 pem=0.0, similarity=3.5687068]
 10: 5381411 Lead_single
 [winner: 7787825 The_Lead, pem_pos=0,
 final_scores=-0.9725784, final_local_scores=-0.9474639,

pem=0.0, similarity=-24.1845|
RECALL MISS: 5381411 Lead_single]
[winner: 4391665 Can't_Help_Thinking_About_Me,
pem_pos=16, final_scores=-1.4643896,
final_local_scores=-1.4597461, pem=-3.9044814,
similarity=-18.238846|
RECALL MISS: 5381411 Lead_single]
11: 19198639 Running_Back_(Jessica_Mauboy_song)
[winner: 295079 Running_back, pem_pos=0,
final_scores=-0.079308495, final_local_scores=0.0048466027,
pem=0.0, similarity=4.2343454|
gt: 19198639 Running_Back_(Jessica_Mauboy_song),
pem_pos=1, final_scores=-0.14069192,
final_local_scores=-0.17772369, pem=-0.784036,
similarity=3.0497932]
12: 20052320 Burn_(Jessica_Mauboy_song)
[winner: 17591231 Burn_(energy_drink), pem_pos=3,
final_scores=-0.369043, final_local_scores=-0.32238033,
pem=-2.3649547, similarity=7.330232|
gt: 20052320 Burn_(Jessica_Mauboy_song), pem_pos=8,
final_scores=-0.5294759, final_local_scores=-0.5758619,
pem=-3.1917367, similarity=4.2620068]
13: 21460096 Been_Waiting_(song)
[winner: 19199205 Been_Waiting, pem_pos=0,
final_scores=0.03777887, final_local_scores=0.002103746,
pem=0.0, similarity=4.152493|
gt: 21460096 Been_Waiting_(song), pem_pos=1,
final_scores=-0.08510214, final_local_scores=-0.057916246,
pem=-0.3591954, similarity=4.3147373]
14: 22750443 Because_(Jessica_Mauboy_song) no overlap with
any filtered span
15: 23586193 Up/Down
[winner: 23586193 Up/Down, pem_pos=0,
final_scores=-0.22252777, final_local_scores=-0.2041682,
pem=0.0, similarity=-2.0030732|
gt: 23586193 Up/Down, pem_pos=0,
final_scores=-0.22252777, final_local_scores=-0.2041682,
pem=0.0, similarity=-2.0030732]
16: 29202144 Get_'Em_Girls
[winner: 29202144 Get_'Em_Girls, pem_pos=0,
final_scores=-0.3478577, final_local_scores=-0.38919008,
pem=-0.04800922, similarity=-7.2634134|
gt: 29202144 Get_'Em_Girls, pem_pos=0,
final_scores=-0.3478577, final_local_scores=-0.38919008,
pem=-0.04800922, similarity=-7.2634134]
17: 29298589 Get_'Em_Girls_(song)
[winner: 1723133 Girl, pem_pos=1,
final_scores=-0.37570578, final_local_scores=-0.34821883,
pem=-1.7719338, similarity=3.3342195|
RECALL MISS: 29298589 Get_'Em_Girls_(song)]
[winner: 3223469 Get_(divorce_document), pem_pos=0,
final_scores=-0.5904194, final_local_scores=-0.48472232,

```

pem=-0.6749649, similarity=-6.704804|
    RECALL MISS: 29298589 Get_'Em_Girls_(song)]
18: 29404953 Saturday_Night_(Jessica_Mauboy_song)
    [winner: 3387499 WCW_Saturday_Night, pem_pos=5,
final_scores=-0.24953224, final_local_scores=-0.19748834,
pem=-2.041324, similarity=9.297303|
    gt: 29404953 Saturday_Night_(Jessica_Mauboy_song),
pem_pos=8, final_scores=-0.48295933,
final_local_scores=-0.52867854, pem=-2.6992686,
similarity=2.9919329 ]
19: 19697266 Inescapable
    [winner: 83893 Escapology, pem_pos=0,
final_scores=-0.21774204, final_local_scores=-0.16502547,
pem=-0.45953238, similarity=1.6640322|
    gt: 19697266 Inescapable, pem_pos=1,
final_scores=-0.34385383, final_local_scores=-0.36469877,
pem=-1.2396909, similarity=-0.051993966 ]
20: 33609204 Galaxy_(song)
    [winner: 12558 Galaxy, pem_pos=0,
final_scores=-0.44745493, final_local_scores=-0.3493587,
pem=-0.6642233, similarity=-2.7236981|
    RECALL MISS: 33609204 Galaxy_(song)]
21: 35429216 The_Sapphires_(film)
    [winner: 20541429 The_Sapphires_(play), pem_pos=1,
final_scores=-0.111730486, final_local_scores=-0.048014626,
pem=-0.115715645, similarity=3.2861404|
    gt: 35429216 The_Sapphires_(film), pem_pos=2,
final_scores=-0.2671299, final_local_scores=-0.29703024,
pem=-1.8817856, similarity=5.459181 ]
22: 37020450 Gotcha_(song)
    [winner: 5143231 Gotcha_(video_game), pem_pos=0,
final_scores=-0.44679117, final_local_scores=-0.29103568,
pem=-0.3998471, similarity=-2.4209416|
    gt: 37020450 Gotcha_(song), pem_pos=12,
final_scores=-0.73574907, final_local_scores=-0.779219,
pem=-3.1640675, similarity=-1.9570444 ]
23: 36972180 The_Sapphires_(soundtrack)
    [winner: 6309935 Underground_(soundtrack), pem_pos=0,
final_scores=-0.592817, final_local_scores=-0.41513163,
pem=0.0, similarity=-8.298643|
    RECALL MISS: 36972180 The_Sapphires_(soundtrack)]
    [winner: 11651855
Kakusei_Heroism_(The_Hero_Without_a_"Name"), pem_pos=6,
final_scores=-0.87657267, final_local_scores=-0.7515509,
pem=-2.3025851, similarity=-5.8162494|
    RECALL MISS: 36972180 The_Sapphires_(soundtrack)]
    [winner: 1363640 300_(comics), pem_pos=0,
final_scores=-1.0481414, final_local_scores=-1.0183547,
pem=0.0, similarity=-26.30002|
    RECALL MISS: 36972180 The_Sapphires_(soundtrack)]
1: 6699718 Jessica_Mauboy
    [pem_pos=0, final_scores=0.32174176,

```

final_local_scores=0.26603296, pem=0.0, similarity=12.028667]
2: 372646 Australian_Idol
[pem_pos=0, final_scores=0.25437713,
final_local_scores=0.21554866, pem=0.0, similarity=10.522115]
3: 17452021 Sony_Music_Australia
[pem_pos=0, final_scores=0.16189422,
final_local_scores=0.14264315, pem=0.0, similarity=8.346469]
4: 3002226 ARIA_Charts
[pem_pos=0, final_scores=0.09002186,
final_local_scores=0.027957886, pem=0.0, similarity=4.924032]
5: 308882 Australian_Recording_Industry_Association
[pem_pos=0, final_scores=0.03780876,
final_local_scores=-0.007407874, pem=0.0,
similarity=3.8686473]
6: 5307421 Young_Divas
[pem_pos=0, final_scores=0.1678475,
final_local_scores=0.16521937, pem=0.0, similarity=9.020188]
7: 13994263 Flo_Rida
[pem_pos=0, final_scores=0.30620706,
final_local_scores=0.29631734, pem=0.0, similarity=12.932413]
8: 88923 Snoop_Dogg
[pem_pos=0, final_scores=0.27057374,
final_local_scores=0.2631171, pem=0.0, similarity=11.941652]
9: 23116760 Stan_Walker
[pem_pos=0, final_scores=0.08041422,
final_local_scores=0.032296464, pem=0.0, similarity=5.053504]
1: 4689264 Australia
[pem_pos=0, final_scores=0.52854526,
final_local_scores=0.5713563, pem=-0.491009,
similarity=23.810307]
2: 6699718 Jessica_Mauboy
[pem_pos=0, final_scores=0.23135044,
final_local_scores=0.1672942, pem=0.0, similarity=9.082106]
3: 372646 Australian_Idol
[pem_pos=0, final_scores=0.19595014,
final_local_scores=0.14916328, pem=0.0, similarity=8.541042]
4: 6699718 Jessica_Mauboy
[pem_pos=0, final_scores=0.247382,
final_local_scores=0.18480626, pem=0.0, similarity=9.6047]
5: 4689264 Australia
[pem_pos=0, final_scores=0.56472635,
final_local_scores=0.6108787, pem=-0.491009,
similarity=24.989735]
6: 6699718 Jessica_Mauboy
[pem_pos=0, final_scores=0.2348936,
final_local_scores=0.17116457, pem=0.0, similarity=9.197605]
7: 3002226 ARIA_Charts
[pem_pos=0, final_scores=0.115907505,
final_local_scores=0.056234002, pem=0.0, similarity=5.7678475]
8: 4689264 Australia
[pem_pos=0, final_scores=0.46569186,
final_local_scores=0.50269854, pem=-0.491009,

similarity=21.761423]
 9: 3434750 United_States
 [pem_pos=0, final_scores=0.45029688,
 final_local_scores=0.5037884, pem=-0.12994486,
 similarity=19.830425]
 10: 3002226 ARIA_Charts
 [pem_pos=0, final_scores=0.112574294,
 final_local_scores=0.052592993, pem=0.0, similarity=5.6591926]
 11: 6699718 Jessica_Mauboy
 [pem_pos=0, final_scores=0.21002351,
 final_local_scores=0.14399776, pem=0.0, similarity=8.386893]
 12: 19199205 Been_Waiting
 [pem_pos=0, final_scores=0.06119363,
 final_local_scores=0.027680814, pem=0.0, similarity=4.915764]
 13: 19199205 Been_Waiting
 [pem_pos=0, final_scores=0.03777887,
 final_local_scores=0.002103746, pem=0.0, similarity=4.152493]
 14: 6699718 Jessica_Mauboy
 [pem_pos=0, final_scores=0.25359392,
 final_local_scores=0.19159183, pem=0.0, similarity=9.807196]
 15: 3434750 United_States
 [pem_pos=0, final_scores=0.48557335,
 final_local_scores=0.54232264, pem=-0.12994486,
 similarity=20.980364]
 16: 4689264 Australia
 [pem_pos=0, final_scores=0.609363,
 final_local_scores=0.65963745, pem=-0.491009,
 similarity=26.444794]
 17: 6699718 Jessica_Mauboy
 [pem_pos=0, final_scores=0.26347226,
 final_local_scores=0.20238239, pem=0.0, similarity=10.129207]
 18: 4913064 New_Zealand
 [pem_pos=0, final_scores=0.51940596,
 final_local_scores=0.5604684, pem=-0.23795462,
 similarity=22.109241]
 19: 1588487 Recorded_Music_NZ
 [pem_pos=0, final_scores=0.12637012,
 final_local_scores=0.0875587, pem=0.0, similarity=6.702639]
 20: 1588487 Recorded_Music_NZ
 [pem_pos=0, final_scores=0.075586185,
 final_local_scores=0.03208497, pem=0.0, similarity=5.0471926]
 21: 6699718 Jessica_Mauboy
 [pem_pos=0, final_scores=0.28121263,
 final_local_scores=0.22176105, pem=0.0, similarity=10.707505]

Sample: b'Manti_National_Forest'

[1 Manti National Forest 1] was established as the [2 Manti
 Forest Reserve 2] by the 1 General Land Office in [1 1 Colora
 do 1] and [2 2 Utah 2] on May 29 , 1903 with . After the
 transfer of federal forests to the in 1905 , it became a [3 3
 National Forest 3] on March 4 , 1907 . On July 1 , 1915 2 Nebo
 National Forest was added . On November 11 , 1941 , [4 4 La Sa

l National Forest 4] was added . On August 28 , 1958 the name was changed to [5 5 Manti-La Sal National Forest 5] . [3 Manti 3] 's lands are on the west side of the [6 6 Colorado River 6] in central [4 Utah 4] , between [7 7 Manti 7] and 3 Price . In descending order of area , its lands are located in parts of [8 8 Sanpete 8] -LRB- 50.3 % -RRB- , 4 Emery -LRB- 28.96 % -RRB- , [5 5 Utah 5] -LRB- 12.59 % -RRB- , 6 Carbon -LRB- 4.11 % -RRB- , and 7 Sevier -LRB- 4.05 % -RRB- counties . As of 30 September 2008 its area was 735,358 acres -LRB- 2,975.9 km ² -RRB- , representing 57.87 % of the combined [6 Manti-La Sal National Forest 6] 's total area . There are local district offices in 8 Ephraim , 9 Ferron , and 10 Price

1: 5399 Colorado

2: 31716 Utah

3: 42653 United_States_National_Forest

4: 18897648 La_Sal_National_Forest

5: 2925149 Manti-La_Sal_National_Forest

6: 180968 Colorado_River

7: 137095 Manti,_Utah

8: 71075 Sanpete_County,_Utah

1: 883651 General_Land_Office

[winner: 883651 General_Land_Office, pem_pos=0, final_scores=0.014376238, final_local_scores=0.011234269, pem=0.0, similarity=4.424966|

gt: 883651 General_Land_Office, pem_pos=0, final_scores=0.014376238, final_local_scores=0.011234269, pem=0.0, similarity=4.424966]

2: 18898060 Nebo_National_Forest

[winner: 18898060 Nebo_National_Forest, pem_pos=0, final_scores=-0.0031686723, final_local_scores=0.051006913, pem=0.0, similarity=5.6118608|

gt: 18898060 Nebo_National_Forest, pem_pos=0, final_scores=-0.0031686723, final_local_scores=0.051006913, pem=0.0, similarity=5.6118608]

3: 136961 Price,_Utah

[winner: 136961 Price,_Utah, pem_pos=1, final_scores=-0.17160392, final_local_scores=-0.22634763, pem=-2.3930426, similarity=10.348785|

gt: 136961 Price,_Utah, pem_pos=1, final_scores=-0.17160392, final_local_scores=-0.22634763, pem=-2.3930426, similarity=10.348785]

4: 71056 Emery_County,_Utah

[winner: 71056 Emery_County,_Utah, pem_pos=1, final_scores=-0.2552505, final_local_scores=-0.3245125, pem=-1.7279739, similarity=3.8026032|

gt: 71056 Emery_County,_Utah, pem_pos=1, final_scores=-0.2552505, final_local_scores=-0.3245125, pem=-1.7279739, similarity=3.8026032]

5: 71080 Utah_County,_Utah

[winner: 31716 Utah, pem_pos=0, final_scores=0.056612, final_local_scores=-0.0075524747, pem=-0.2611261,

```

similarity=5.2843766|
    gt: 71080 Utah_County,_Utah, pem_pos=3,
final_scores=-0.59163374, final_local_scores=-0.6916426,
pem=-3.503183, similarity=2.5005713 ]
6: 71045 Carbon_County,_Utah
    [winner: 5299 Carbon, pem_pos=0,
final_scores=-0.44049796, final_local_scores=-0.348379,
pem=-0.8652096, similarity=-1.6014663]
    gt: 71045 Carbon_County,_Utah, pem_pos=5,
final_scores=-0.58812666, final_local_scores=-0.6685272,
pem=-2.739891, similarity=-0.96051776 ]
7: 71076 Sevier_County,_Utah
    [winner: 71076 Sevier_County,_Utah, pem_pos=2,
final_scores=-0.152766, final_local_scores=-0.19587384,
pem=-1.4575319, similarity=6.170732]
    gt: 71076 Sevier_County,_Utah, pem_pos=2,
final_scores=-0.152766, final_local_scores=-0.19587384,
pem=-1.4575319, similarity=6.170732 ]
8: 137090 Ephraim,_Utah
    [winner: 471326 Ephraim, pem_pos=0,
final_scores=-0.06396188, final_local_scores=0.021112114,
pem=-0.66764975, similarity=8.350524]
    gt: 137090 Ephraim,_Utah, pem_pos=2,
final_scores=-0.08524035, final_local_scores=-0.0966461,
pem=-1.8003013, similarity=10.995919 ]
9: 136997 Ferron,_Utah
    [winner: 1088160 Ferron, pem_pos=0,
final_scores=-0.10967545, final_local_scores=0.018208772,
pem=-0.054864094, similarity=4.9314585]
    gt: 136997 Ferron,_Utah, pem_pos=1,
final_scores=-0.27922398, final_local_scores=-0.27356833,
pem=-1.4473003, similarity=3.7965324 ]
10: 136961 Price,_Utah
    [winner: 136961 Price,_Utah, pem_pos=1,
final_scores=-0.39389527, final_local_scores=-0.46916717,
pem=-2.3930426, similarity=3.102566]
    gt: 136961 Price,_Utah, pem_pos=1,
final_scores=-0.39389527, final_local_scores=-0.46916717,
pem=-2.3930426, similarity=3.102566 ]
1: 5399 Colorado
    [pem_pos=0, final_scores=0.5015153,
final_local_scores=0.49939167, pem=-0.26926568,
similarity=20.456865]
2: 31716 Utah
    [pem_pos=0, final_scores=0.44105738,
final_local_scores=0.42157346, pem=-0.2611261,
similarity=18.09035]
3: 42653 United_States_National_Forest
    [pem_pos=0, final_scores=0.19796292,
final_local_scores=0.18048385, pem=0.0, similarity=9.475711]
4: 18897648 La_Sal_National_Forest
    [pem_pos=0, final_scores=0.14292188,

```

final_local_scores=0.17269605, pem=0.0, similarity=9.243308]
 5: 2925149 Manti-La_Sal_National_Forest
 [pem_pos=0, final_scores=0.21713595,
 final_local_scores=0.20802867, pem=0.0, similarity=10.297704]
 6: 180968 Colorado_River
 [pem_pos=0, final_scores=0.38371855,
 final_local_scores=0.37864995, pem=-0.04937945,
 similarity=15.657917]
 7: 137095 Manti,_Utah
 [pem_pos=0, final_scores=0.15631844,
 final_local_scores=0.1638085, pem=-0.1772828,
 similarity=9.942176]
 8: 71075 Sanpete_County,_Utah
 [pem_pos=0, final_scores=0.117638305,
 final_local_scores=0.101290986, pem=0.0, similarity=7.112438]
 1: 18897996 Manti_National_Forest
 [pem_pos=0, final_scores=0.08560042,
 final_local_scores=0.09198581, pem=0.0, similarity=6.834753]
 2: 18897996 Manti_National_Forest
 [pem_pos=0, final_scores=0.059633777,
 final_local_scores=0.06362119, pem=0.0, similarity=5.988296]
 3: 137095 Manti,_Utah
 [pem_pos=0, final_scores=0.0825396,
 final_local_scores=0.083216324, pem=-0.1772828,
 similarity=7.5371456]
 4: 31716 Utah
 [pem_pos=0, final_scores=0.56308687,
 final_local_scores=0.55487216, pem=-0.2611261,
 similarity=22.068249]
 5: 31716 Utah
 [pem_pos=0, final_scores=0.056612,
 final_local_scores=-0.0075524747, pem=-0.2611261,
 similarity=5.2843766]
 6: 2925149 Manti-La_Sal_National_Forest
 [pem_pos=0, final_scores=0.21606983,
 final_local_scores=0.20686412, pem=0.0, similarity=10.262951]

Sample: b'Green_Lane_railway_station'

Green Lane railway station serves the town of [1 1 Birkenhead
 1] , [1 1 Wirral 1] , [2 2 England 2] , located in the area
 of [2 2 Tranmere 2] . It is situated on the [3 3 Chester 3]
 and [4 4 Ellesmere Port 4] branches of the [5 5 Wirral Line 5]
 , part of the [6 6 Merseyrail 6] network . The station is
 located in a below ground level cutting . As initially the
 station accommodated steam trains it is partially open to the
 atmosphere above to aid ventilation . The station was opened
 in 1886 as the terminus of the [7 7 Mersey Railway 7] 's line
 from [8 8 Liverpool 8] via the [3 Mersey Railway Tunnel 3] .
 In 1891 the line was extended southwards to [4 3 Rock Ferry 4]
 , and was subsequently electrified in 1903 . It became part
 of the [9 9 British Rail 9] 's [5 Merseyrail 5] network in
 the 1970s . Further electrification in the 1980s and 1990s

allowed train services to be extended beyond [6 Rock Ferry 6]
; first to [10 10 Hooton 10] in 1985 , then to [7 4 Chester 7]
in 1993 and finally [8 5 Ellesmere Port 8] in 1994 . Trains
operate every 15 minutes between [9 Chester 9] and [10
Liverpool 10] on weekdays and Saturdays until late evening
when the service becomes half-hourly , as it is on Sundays .
Additionally there is a half-hourly service between [11
Liverpool 11] and [12 Ellesmere Port 12] all day , every
day . Northbound trains operate via station in [11 11 Birkenh
ead 11] and the to [12 12 Liverpool 12] . Southbound trains
all proceed as far as [13 13 Hooton 13] , where the lines to [13
6 Chester 13] and [14 7 Ellesmere Port 14]

1: 246155 Birkenhead

2: 9316 England

3: 153718 Chester

4: 311674 Ellesmere_Port

5: 2762567 Wirral_Line

6: 327883 Merseyrail

7: 1057496 Mersey_Railway

8: 18081 Liverpool

9: 4385 British_Rail

10: 3302255 Hooton_railway_station

11: 246155 Birkenhead

12: 18081 Liverpool

13: 3302255 Hooton_railway_station

1: 300824 Metropolitan_Borough_of_Wirral

[winner: 1526370 Wirral_Peninsula, pem_pos=1,
final_scores=0.30411476, final_local_scores=0.26633048,
pem=-1.0037606, similarity=17.496153]

gt: 300824 Metropolitan_Borough_of_Wirral, pem_pos=0,
final_scores=0.2884373, final_local_scores=0.281775,
pem=-0.63159513, similarity=15.933154]

2: 1113131 Tranmere,_Merseyside

[winner: 69219 Tranmere_Rovers_F.C., pem_pos=0,
final_scores=0.04478033, final_local_scores=0.067519575,
pem=-0.6002322, similarity=9.368788]

gt: 1113131 Tranmere,_Merseyside, pem_pos=1,
final_scores=-0.028790653, final_local_scores=-0.040243864,
pem=-0.87441665, similarity=7.6439695]

3: 4603115 Rock_Ferry_railway_station

[winner: 5551878 Rock_Ferry, pem_pos=0,
final_scores=0.14846976, final_local_scores=0.12683243,
pem=0.0, similarity=7.874645]

gt: 4603115 Rock_Ferry_railway_station, pem_pos=1,
final_scores=0.06608562, final_local_scores=0.0637117,
pem=-0.19845095, similarity=7.0702043]

4: 2908608 Chester_railway_station

[winner: 153718 Chester, pem_pos=0,
final_scores=0.2896605, final_local_scores=0.24562186,
pem=-0.548017, similarity=14.399761]

gt: 2908608 Chester_railway_station, pem_pos=8,
final_scores=-0.5124882, final_local_scores=-0.6059818,

pem=-4.1612525, similarity=8.635546]
5: 4127885 Ellesmere_Port_railway_station
[winner: 311674 Ellesmere_Port, pem_pos=0,
final_scores=0.31051105, final_local_scores=0.26631948,
pem=0.0, similarity=12.037217]
gt: 4127885 Ellesmere_Port_railway_station, pem_pos=1,
final_scores=0.03307666, final_local_scores=0.031129345,
pem=-1.0485171, similarity=10.720674]
6: 2908608 Chester_railway_station
[winner: 153718 Chester, pem_pos=0,
final_scores=0.26215565, final_local_scores=0.21557698,
pem=-0.548017, similarity=13.503161]
gt: 2908608 Chester_railway_station, pem_pos=8,
final_scores=-0.60602605, final_local_scores=-0.7081577,
pem=-4.1612525, similarity=5.5864124]
7: 4127885 Ellesmere_Port_railway_station
[winner: 311674 Ellesmere_Port, pem_pos=0,
final_scores=0.17803337, final_local_scores=0.122510046,
pem=0.0, similarity=7.7456574]
gt: 4127885 Ellesmere_Port_railway_station, pem_pos=1,
final_scores=-0.07385129, final_local_scores=-0.092485294,
pem=-1.0485171, similarity=7.0317664]
1: 246155 Birkenhead
[pem_pos=0, final_scores=0.29173416,
final_local_scores=0.24750617, pem=-0.1715908,
similarity=12.408929]
2: 9316 England
[pem_pos=0, final_scores=0.5323532,
final_local_scores=0.5482437, pem=-0.19053856,
similarity=21.486576]
3: 153718 Chester
[pem_pos=0, final_scores=0.33037657,
final_local_scores=0.29009795, pem=-0.548017,
similarity=15.727016]
4: 311674 Ellesmere_Port
[pem_pos=0, final_scores=0.3865838,
final_local_scores=0.34941742, pem=0.0, similarity=14.517026]
5: 2762567 Wirral_Line
[pem_pos=0, final_scores=0.3468622,
final_local_scores=0.3291358, pem=0.0, similarity=13.911782]
6: 327883 Merseyrail
[pem_pos=0, final_scores=0.24886538,
final_local_scores=0.21390048, pem=0.0, similarity=10.47293]
7: 1057496 Mersey_Railway
[pem_pos=0, final_scores=0.34963447,
final_local_scores=0.34589878, pem=0.0, similarity=14.412023]
8: 18081 Liverpool
[pem_pos=0, final_scores=0.35988814,
final_local_scores=0.3296262, pem=-0.56054556,
similarity=16.97475]
9: 4385 British_Rail
[pem_pos=0, final_scores=0.38077384,

final_local_scores=0.3854953, pem=0.0, similarity=15.593661]
10: 3302255 Hooton_railway_station
[pem_pos=0, final_scores=0.043994352,
final_local_scores=0.018308654, pem=-0.5862305,
similarity=7.8240924]
11: 246155 Birkenhead
[pem_pos=0, final_scores=0.27182055,
final_local_scores=0.22575355, pem=-0.1715908,
similarity=11.759787]
12: 18081 Liverpool
[pem_pos=0, final_scores=0.3955137,
final_local_scores=0.36854172, pem=-0.56054556,
similarity=18.136066]
13: 3302255 Hooton_railway_station
[pem_pos=0, final_scores=0.0566843,
final_local_scores=0.031858012, pem=-0.5862305,
similarity=8.228433]
1: 1526370 Wirral_Peninsula
[pem_pos=1, final_scores=0.30411476,
final_local_scores=0.26633048, pem=-1.0037606,
similarity=17.496153]
2: 69219 Tranmere_Rovers_F.C.
[pem_pos=0, final_scores=0.04478033,
final_local_scores=0.067519575, pem=-0.6002322,
similarity=9.368788]
3: 1057496 Mersey_Railway
[pem_pos=0, final_scores=0.45131034,
final_local_scores=0.4569642, pem=0.0, similarity=17.726437]
4: 5551878 Rock_Ferry
[pem_pos=0, final_scores=0.14846976,
final_local_scores=0.12683243, pem=0.0, similarity=7.874645]
5: 327883 Merseyrail
[pem_pos=0, final_scores=0.19677903,
final_local_scores=0.15700406, pem=0.0, similarity=8.775027]
6: 5551878 Rock_Ferry
[pem_pos=0, final_scores=0.2281482,
final_local_scores=0.21386904, pem=0.0, similarity=10.471992]
7: 153718 Chester
[pem_pos=0, final_scores=0.2896605,
final_local_scores=0.24562186, pem=-0.548017,
similarity=14.399761]
8: 311674 Ellesmere_Port
[pem_pos=0, final_scores=0.31051105,
final_local_scores=0.26631948, pem=0.0, similarity=12.037217]
9: 153718 Chester
[pem_pos=0, final_scores=0.40542728,
final_local_scores=0.3720795, pem=-0.548017,
similarity=18.17351]
10: 18081 Liverpool
[pem_pos=0, final_scores=0.34898537,
final_local_scores=0.31921935, pem=-0.56054556,
similarity=16.664188]

11: 18081 Liverpool
 [pem_pos=0, final_scores=0.29993182,
 final_local_scores=0.26413304, pem=-0.56054556,
 similarity=15.020304]
 12: 311674 Ellesmere_Port
 [pem_pos=0, final_scores=0.36130285,
 final_local_scores=0.3218018, pem=0.0, similarity=13.692921]
 13: 153718 Chester
 [pem_pos=0, final_scores=0.26215565,
 final_local_scores=0.21557698, pem=-0.548017,
 similarity=13.503161]
 14: 311674 Ellesmere_Port
 [pem_pos=0, final_scores=0.17803337,
 final_local_scores=0.122510046, pem=0.0, similarity=7.7456574]

Sample: b'Cyclone_Taylor_Trophy'

The [1 Cyclone Taylor Award 1] is the award given each year to the on the [1 1 Vancouver Canucks 1] -LRB- a [2 2 National Hockey League 2] team -RRB-. It is named after [3 3 Cyclone Taylor 3], a 1 professional 2 ice hockey 3 forward who led the [4 4 Vancouver Millionaires 4] to the [5 5 Stanley Cup 5] in 1915. The award was dedicated to him prior to the 1979-80 Canuck season, the season after his death on June 9, 1979, although an award for the Canucks MVP has existed since the team's 4 inauguration in 5 1970. Previously it was a Canucks MVP Award as selected by the fans while the other MVP award, the 6 President's Trophy was selected by and later [6 6 Canadian Airlines 6]. However after the 7 1995 -- 96 season, the [2 Cyclone Taylor Trophy 2] officially became the lone Canucks MVP award since the winners of each trophy was identical. The most prolific winner of the [3 Cyclone Taylor Trophy 3] is [7 7 Markus Naslund 7], who has been awarded five times -LRB- including four straight from 8 2001 to 9 2004 -RRB-, followed by [8 8 Trevor Linden 8] with four. The trophy's present holder is [9 9 Cory Schneider 9] -LRB- 10 20 13

1: 73133 Vancouver_Canucks
 2: 21809 National_Hockey_League
 3: 2025107 Cyclone_Taylor
 4: 614713 Vancouver_Millionaires
 5: 66968 Stanley_Cup
 6: 281157 Canadian_Airlines
 7: 440439 Markus_Naslund
 8: 540313 Trevor_Linden
 9: 5877043 Cory_Schneider

1: 290809 Professional
 [winner: 16786195]

Major_professional_sports_teams_of_the_United_States_and_Canada, pem_pos=0, final_scores=-1.204074,
 final_local_scores=-1.2792044, pem=-0.25423416,
 similarity=-32.70173]

RECALL MISS: 290809 Professional]

2: 14790 Ice_hockey
 [winner: 14946 Ice, pem_pos=0, final_scores=-0.63301873, final_local_scores=-0.55998194, pem=-0.116236255, similarity=-11.989155]
 gt: 14790 Ice_hockey, pem_pos=3, final_scores=-0.8478106, final_local_scores=-0.9265599, pem=-2.165202, similarity=-11.785978]
 [winner: 567665 Forward_(ice_hockey), pem_pos=0, final_scores=-1.1068281, final_local_scores=-1.185878, pem=0.0, similarity=-31.29925]
 RECALL MISS: 14790 Ice_hockey]

3: 567665 Forward_(ice_hockey)
 [winner: 567665 Forward_(ice_hockey), pem_pos=0, final_scores=-1.1068281, final_local_scores=-1.185878, pem=0.0, similarity=-31.29925]
 gt: 567665 Forward_(ice_hockey), pem_pos=0, final_scores=-1.1068281, final_local_scores=-1.185878, pem=0.0, similarity=-31.29925]

4: 226188 Inauguration
 [winner: 21204142 First_inauguration_of_Richard_Nixon, pem_pos=2, final_scores=-0.90932816, final_local_scores=-0.8570213, pem=-1.0986123, similarity=-15.5110855]
 gt: 226188 Inauguration, pem_pos=3, final_scores=-1.9588643, final_local_scores=-2.0651577, pem=-3.009456, similarity=-41.172825]

5: 2910492 1970-71_NHL_season
 [winner: 34726 1970, pem_pos=0, final_scores=-0.7175784, final_local_scores=-0.56470484, pem=-0.80837256, similarity=-8.366151]
 RECALL MISS: 2910492 1970-71_NHL_season]

6: 12016649 President's_Trophy_(Canucks_MVP)
 [winner: 619119 Trophy, pem_pos=0, final_scores=-0.49453136, final_local_scores=-0.49568862, pem=-1.336462, similarity=-3.434736]
 RECALL MISS: 12016649 President's_Trophy_(Canucks_MVP)]
 [winner: 6398179 The_President, pem_pos=0, final_scores=-0.81512517, final_local_scores=-0.771079, pem=-0.3939714, similarity=-16.778345]
 RECALL MISS: 12016649 President's_Trophy_(Canucks_MVP)]
 [winner: 1415027 David_Berman_(musician), pem_pos=5, final_scores=-2.4176867, final_local_scores=-2.4999568, pem=-3.0012724, similarity=-54.192596]
 RECALL MISS: 12016649 President's_Trophy_(Canucks_MVP)]

7: 11129464 1995-96_Vancouver_Canucks_season
 [winner: 26089658 No._96_Squadron_RAF, pem_pos=0, final_scores=-0.47835648, final_local_scores=-0.37241685, pem=-2.1547496, similarity=4.693915]
 RECALL MISS: 11129464 1995-96_Vancouver_Canucks_season]
 [winner: 5834357 Juno_Awards_of_1995, pem_pos=0, final_scores=-0.84682155, final_local_scores=-0.788847, pem=0.0, similarity=-19.451056]

RECALL MISS: 11129464 1995-96_Vancouver_Canucks_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3588662, final_local_scores=-1.2749078,
pem=0.0, similarity=-33.956078]
RECALL MISS: 11129464 1995-96_Vancouver_Canucks_season]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.678379, final_local_scores=-1.8618338,
pem=-1.219945, similarity=-44.836872]
RECALL MISS: 11129464 1995-96_Vancouver_Canucks_season]
8: 1673932 2000-01_NHL_season
[winner: 34551 2001, pem_pos=0, final_scores=-0.6426448,
final_local_scores=-0.40471822, pem=-0.5613713,
similarity=-4.935062]
RECALL MISS: 1673932 2000-01_NHL_season]
9: 576605 2003-04_NHL_season
[winner: 35524 2004, pem_pos=0, final_scores=-0.56775296,
final_local_scores=-0.42829645, pem=-0.4915334,
similarity=-6.0184727]
RECALL MISS: 576605 2003-04_NHL_season]
10: 29020787 2012-13_NHL_season
[winner: 46945 2013, pem_pos=0, final_scores=-0.81286925,
final_local_scores=-0.61601925, pem=-1.490928,
similarity=-6.185632]
RECALL MISS: 29020787 2012-13_NHL_season]
1: 73133 Vancouver_Canucks
[pem_pos=0, final_scores=0.2064436,
final_local_scores=0.15672082, pem=0.0, similarity=8.766575]
2: 21809 National_Hockey_League
[pem_pos=0, final_scores=0.31708843,
final_local_scores=0.28409454, pem=0.0, similarity=12.56766]
3: 2025107 Cyclone_Taylor
[pem_pos=0, final_scores=0.22964583,
final_local_scores=0.21235585, pem=0.0, similarity=10.426835]
4: 614713 Vancouver_Millionaires
[pem_pos=0, final_scores=0.077582374,
final_local_scores=0.05735776, pem=0.0, similarity=5.801383]
5: 66968 Stanley_Cup
[pem_pos=0, final_scores=0.29283595,
final_local_scores=0.24777487, pem=0.0, similarity=11.483809]
6: 281157 Canadian_Airlines
[pem_pos=0, final_scores=0.12425633,
final_local_scores=0.19255915, pem=0.0, similarity=9.836062]
7: 440439 Markus_Näslund
[pem_pos=0, final_scores=0.6827221,
final_local_scores=0.69170475, pem=0.0, similarity=24.731562]
8: 540313 Trevor_Linden
[pem_pos=0, final_scores=0.34186918,
final_local_scores=0.31490383, pem=0.0, similarity=13.487071]
9: 5877043 Cory_Schneider
[pem_pos=0, final_scores=0.28714877,
final_local_scores=0.27574798, pem=0.0, similarity=12.318583]
1: 9870036 Cyclone_Taylor_Trophy

[pem_pos=0, final_scores=0.07363595,
final_local_scores=0.04893607, pem=0.0, similarity=5.5500627]
2: 9870036 Cyclone_Taylor_Trophy
[pem_pos=0, final_scores=0.032062277,
final_local_scores=0.003523156, pem=0.0, similarity=4.194851]
3: 9870036 Cyclone_Taylor_Trophy
[pem_pos=0, final_scores=0.08373289,
final_local_scores=0.059965447, pem=0.0, similarity=5.8792014]

Sample: b'Province_of_Calatayud'

The Province of [1 Calatayud 1] was a of [1 1 Spain 1]
created in the 1 1822 territorial division of [2 Spain 2] -
LRB- 27 January 1822 -RRB- , during the [2 2 Trienio Liberal 2]
] of 1820 -- 1823 . It included the [3 3 Aragon 3] ese 2 coma
rca s of [4 4 Comunidad de Calatayud 4] , 3 Campo de Daroca ,
4 Aranda , the southern part of 5 Valdejalón -LRB- now in the
6 province of [3 Zaragoza 3] -RRB- , the northern part of the
-LRB- now in the [4 7 Teruel 4] -RRB- , as well as some 8 mun
icipalities that now fall in the provinces of [5 9 Soria 5]
and [6 10 Guadalajara 6] . It had a population of 105,947 ,
which constituted 0.9 percent of the [7 Spanish 7] population
of the time . Its capital was [5 5 Calatayud 5] . With the
restoration of 11 absolutism , this territorial division was
revoked 1 October 1823 . Although [6 6 Javier de Burgos 6] 's
12 1833 territorial division of [8 Spain 8] was very close to
that of 1833 , the province of [9 Calatayud 9] was not
recreated ; the other two major changes were the omission of
the provinces of [10 13 Villafranca del Bierzo 10] and 14 Ját
iva

1: 26667 Spain
2: 566318 Trienio_Liberal
3: 39443 Aragon
4: 23028793 Comunidad_de_Calatayud
5: 1369011 Calatayud
6: 4351294 Javier_de_Burgos

1: 25669911 1822_territorial_division_of_Spain
[winner: 26667 Spain, pem_pos=2, final_scores=0.5070808,
final_local_scores=0.49072546, pem=-0.18857051,
similarity=19.759415]

RECALL MISS: 25669911
1822_territorial_division_of_Spain]
[winner: 35020 1822, pem_pos=0, final_scores=-0.8967818,
final_local_scores=-0.7369365, pem=-0.32466814,
similarity=-16.136347]

gt: 25669911 1822_territorial_division_of_Spain,
pem_pos=8, final_scores=-1.2203861,
final_local_scores=-1.2823668, pem=-2.1962957,
similarity=-22.234869]

[winner: 7135137 Divisions_of_Malaysia, pem_pos=16,
final_scores=-1.3800788, final_local_scores=-1.3039939,
pem=-3.5447905, similarity=-15.5469475]

RECALL MISS: 25669911

1822_territorial_division_of_Spain]
 [winner: 501364 Territory_(animal), pem_pos=0,
 final_scores=-1.5073849, final_local_scores=-1.4912624,
 pem=-0.52021074, similarity=-37.58355]
 RECALL MISS: 25669911 1822_territorial_division_of_Spain]

2: 1555834 Comarca
 [winner: 592714 Comarcas_of_Spain, pem_pos=0,
 final_scores=-1.0093445, final_local_scores=-1.0838089,
 pem=0.0, similarity=-28.253307]
 gt: 1555834 Comarca, pem_pos=1, final_scores=-1.35583,
 final_local_scores=-1.5136311, pem=-0.8858003,
 similarity=-36.262936]

3: 30000098 Campo_de_Daroca
 [winner: 30000098 Campo_de_Daroca, pem_pos=0,
 final_scores=0.02287303, final_local_scores=0.101346344,
 pem=0.0, similarity=7.11409]
 gt: 30000098 Campo_de_Daroca, pem_pos=0,
 final_scores=0.02287303, final_local_scores=0.101346344,
 pem=0.0, similarity=7.11409]

4: 30994923 Aranda_(comarca)
 [winner: 22285685 Carlos_Aranda, pem_pos=3,
 final_scores=-0.23741019, final_local_scores=-0.17319468,
 pem=-1.5305026, similarity=7.2443495]
 gt: 30994923 Aranda_(comarca), pem_pos=1,
 final_scores=-0.3389273, final_local_scores=-0.27646467,
 pem=-1.2087805, similarity=2.412992]

5: 29402872 Valdejalón
 [winner: 29402872 Valdejalón, pem_pos=0,
 final_scores=-0.1359176, final_local_scores=-0.010041922,
 pem=-0.046520036, similarity=4.043025]
 gt: 29402872 Valdejalón, pem_pos=0,
 final_scores=-0.1359176, final_local_scores=-0.010041922,
 pem=-0.046520036, similarity=4.043025]

6: 143508 Province_of_Zaragoza
 [winner: 64203 Zaragoza, pem_pos=0,
 final_scores=0.19552435, final_local_scores=0.13524875,
 pem=-0.45408, similarity=10.595162]
 gt: 143508 Province_of_Zaragoza, pem_pos=2,
 final_scores=-0.2661441, final_local_scores=-0.3052389,
 pem=-1.9999385, similarity=5.8567514]
 [winner: 1531900 Arequipa_Province, pem_pos=3,
 final_scores=-1.1899931, final_local_scores=-1.1939967,
 pem=-1.0586069, similarity=-25.784664]
 RECALL MISS: 143508 Province_of_Zaragoza]

7: 143509 Province_of_Teruel
 [winner: 68299 Teruel, pem_pos=0,
 final_scores=0.12894557, final_local_scores=0.09294808,
 pem=-0.403703, similarity=9.058868]
 gt: 143509 Province_of_Teruel, pem_pos=1,
 final_scores=-0.03040722, final_local_scores=-0.043480173,
 pem=-0.72662306, similarity=6.743667]

8: 8767500 Municipalities_of_Spain

[winner: 13605167 Municipalities_of_Albania, pem_pos=12,
final_scores=-0.9003853, final_local_scores=-0.7764751,
pem=-3.0249755, similarity=-2.6315649]

gt: 8767500 Municipalities_of_Spain, pem_pos=14,
final_scores=-1.3998543, final_local_scores=-1.506654,
pem=-3.0713272, similarity=-24.169493]

9: 143645 Province_of_Soria

[winner: 616226 Soria, pem_pos=0,
final_scores=0.13933234, final_local_scores=0.095128536,
pem=-0.46040222, similarity=9.432278]

gt: 143645 Province_of_Soria, pem_pos=1,
final_scores=-0.109094456, final_local_scores=-0.101785794,
pem=-1.0967846, similarity=7.016707]

10: 143516 Province_of_Guadalajara

[winner: 73209 Guadalajara, pem_pos=0,
final_scores=0.2754565, final_local_scores=0.2895596,
pem=-0.39605975, similarity=14.884583]

gt: 143516 Province_of_Guadalajara, pem_pos=2,
final_scores=-0.24738088, final_local_scores=-0.26708758,
pem=-1.9054589, similarity=6.481469]

11: 67366 Absolute_monarchy

[winner: 49062 Moral_absolutism, pem_pos=1,
final_scores=-1.1777102, final_local_scores=-1.1172519,
pem=-0.8445612, similarity=-24.658457]

gt: 67366 Absolute_monarchy, pem_pos=0,
final_scores=-1.2576615, final_local_scores=-1.3457654,
pem=-0.099729486, similarity=-35.528267]

12: 25616199 1833_territorial_division_of_Spain

[winner: 26667 Spain, pem_pos=2, final_scores=0.6230293,
final_local_scores=0.61770916, pem=-0.18857051,
similarity=23.548862]

RECALL MISS: 25616199
1833_territorial_division_of_Spain]

[winner: 2281272 1833_in_archaeology, pem_pos=8,
final_scores=-0.76746094, final_local_scores=-0.5380371,
pem=-2.5955045, similarity=2.1483688]

RECALL MISS: 25616199
1833_territorial_division_of_Spain]

[winner: 7135137 Divisions_of_Malaysia, pem_pos=16,
final_scores=-1.3389828, final_local_scores=-1.2591028,
pem=-3.5447905, similarity=-14.207305]

RECALL MISS: 25616199
1833_territorial_division_of_Spain]

[winner: 501364 Territory_(animal), pem_pos=0,
final_scores=-1.4380925, final_local_scores=-1.415571,
pem=-0.52021074, similarity=-35.32476]

RECALL MISS: 25616199 1833_territorial_division_of_Spain]

13: 25671171 Province_of_Villafranca

[winner: 13156037 Villafranca_del_Bierzo, pem_pos=0,
final_scores=0.21540587, final_local_scores=0.25475237,
pem=0.0, similarity=11.692032]

gt: 25671171 Province_of_Villafranca, pem_pos=1,


```

final_scores=-0.0061115026, final_local_scores=0.07321727,
pem=-0.68278444, similarity=9.987751 ]
14: 25670837 Province_of_Játiva
    [winner: 4035136 Xàtiva, pem_pos=0,
final_scores=-0.07637985, final_local_scores=-0.08881638,
pem=-0.19237192, similarity=2.4854033]
    gt: 25670837 Province_of_Játiva, pem_pos=1,
final_scores=-0.47598496, final_local_scores=-0.44185156,
pem=-1.7429693, similarity=0.38251948 ]
1: 26667 Spain
    [pem_pos=2, final_scores=0.6109951,
final_local_scores=0.60456365, pem=-0.18857051,
similarity=23.156574]
2: 566318 Trienio Liberal
    [pem_pos=0, final_scores=0.0842673,
final_local_scores=0.09126541, pem=0.0, similarity=6.813255]
3: 39443 Aragon
    [pem_pos=0, final_scores=0.35417968,
final_local_scores=0.30150002, pem=-0.16680257,
similarity=13.994173]
4: 23028793 Comunidad_de_Calatayud
    [pem_pos=0, final_scores=0.11277489,
final_local_scores=0.1728827, pem=0.0, similarity=9.248878]
5: 1369011 Calatayud
    [pem_pos=0, final_scores=0.24946155,
final_local_scores=0.21799779, pem=0.0, similarity=10.5952015]
6: 4351294 Javier_de_Burgos
    [pem_pos=0, final_scores=0.09058957,
final_local_scores=0.1250023, pem=0.0, similarity=7.820031]
1: 1369011 Calatayud
    [pem_pos=0, final_scores=0.22078948,
final_local_scores=0.18667793, pem=0.0, similarity=9.660555]
2: 26667 Spain
    [pem_pos=2, final_scores=0.5070808,
final_local_scores=0.49072546, pem=-0.18857051,
similarity=19.759415]
3: 64203 Zaragoza
    [pem_pos=0, final_scores=0.19552435,
final_local_scores=0.13524875, pem=-0.45408,
similarity=10.595162]
4: 68299 Teruel
    [pem_pos=0, final_scores=0.12894557,
final_local_scores=0.09294808, pem=-0.403703,
similarity=9.058868]
5: 616226 Soria
    [pem_pos=0, final_scores=0.13933234,
final_local_scores=0.095128536, pem=-0.46040222,
similarity=9.432278]
6: 73209 Guadalajara
    [pem_pos=0, final_scores=0.2754565,
final_local_scores=0.2895596, pem=-0.39605975,
similarity=14.884583]

```

7: 26667 Spain
 [pem_pos=1, final_scores=0.52254736,
 final_local_scores=0.5079479, pem=-1.2425895,
 similarity=26.005287]
 8: 26667 Spain
 [pem_pos=2, final_scores=0.6230293,
 final_local_scores=0.61770916, pem=-0.18857051,
 similarity=23.548862]
 9: 1369011 Calatayud
 [pem_pos=0, final_scores=0.3096186,
 final_local_scores=0.2837102, pem=0.0, similarity=12.556191]
 10: 13156037 Villafranca_del_Bierzo
 [pem_pos=0, final_scores=0.21540587,
 final_local_scores=0.25475237, pem=0.0, similarity=11.692032]

Sample: b'Halsey_Beshears'

Halsey Beshears -LRB- born November 27 , 1971 , in [1 1 Wint
 er Park 1] , [2 Florida 2] -RRB- is a [1 1 Republican 1]
 member of the [2 2 Florida House of Representatives 2] ,
 representing the 7th District , which includes 2 Calhoun , 3 G
 ulf , 4 Lafayette , 5 Leon , 6 Liberty , [3 7 Madison 3] , 8 T
 aylor , and [3 3 Wakulla Counties 3] in the [4 4 Florida Panh
 andle 4] , since 2012 . Beshears 's family moved to 9 Montice
 llo and he graduated from Brookwood School , a private school
 in [4 10 Thomasville 4] , [5 Georgia 5] in 1989 . He later
 attended [5 5 Florida State University 5] and the [6 6 Univer
 sity of Florida 6] . Since 1997 , Beshears has served as an
 executive at Simpson Nurseries and the President of Total
 Landscape Supply , which serves wholesale clients across much
 of the [7 7 United States 7] . Following the reconfiguration
 of [8 8 Florida House of Representatives 8] districts in
 2012 , the 7th District was created , which sprawls across
 many counties in northern 11 Florida . Beshears opted to run
 in this district , and defeated Mike Williams , former 12 Sta
 te Representative Jamey Westbrook , and Don Curtis in the [9 9
 Republican 9] primary , winning 38 % of the vote . In the
 general election , Beshears faced [6 Robert Hill 6] , the [10
 10 Democratic 10] nominee and the former 13 Liberty County
 Administrator , Clerk of Court , and Public Schools
 Superintendent . Beshears was criticized by the [7 14 Tallaha
 ssee 7] Democrat for displaying a `` marked lack of interest
 '' and they noted that they `` just do n't get '' his
 popularity within the district . In the end , however ,
 Beshears swamped [8 Hill 8] , winning every county in the
 district except for [9 15 Liberty County 9] , which Hill
 carried comfortably . While serving in the [11 11 Florida Hou
 se of Representatives 11] , [10 Beshear 10] proposed
 legislation to protect `` agricultural lands from being
 subject to double regulations at both the local and state or
 federal level . '' When confronted with legislation that
 would `` outlaw all sales of various smoking devices defined
 as ` drug paraphernalia , ' ' Beshears took a strong stance

against it , declaring , `` The government can intrude in any business , and I just completely disagree with that . So if we 're going to outlaw smoking paraphernalia , why do n't we outlaw spoons , why do n't we outlaw aluminum cans , why do n't we outlaw anything that is used to smoke ? '' After a [12 12 University of Florida 12] study showed that the oyster population in 16 Apalachicola Bay

1: 32070 Republican_Party_(United_States)
2: 633627 Florida_House_of_Representatives
3: 75116 Wakulla_County,_Florida
4: 902849 Florida_Panhandle
5: 239297 Florida_State_University
6: 60611 University_of_Florida
7: 3434750 United_States
8: 633627 Florida_House_of_Representatives
9: 32070 Republican_Party_(United_States)
10: 5043544 Democratic_Party_(United_States)
11: 633627 Florida_House_of_Representatives
12: 60611 University_of_Florida

1: 109566 Winter_Park,_Florida
[winner: 109566 Winter_Park,_Florida, pem_pos=0, final_scores=0.47737545, final_local_scores=0.48059452, pem=-0.01514043, similarity=18.513947]
gt: 109566 Winter_Park,_Florida, pem_pos=0, final_scores=0.47737545, final_local_scores=0.48059452, pem=-0.01514043, similarity=18.513947]
[winner: 18933066 Florida, pem_pos=0, final_scores=0.047894105, final_local_scores=-0.0095321685, pem=-0.15405534, similarity=4.643031]
RECALL MISS: 109566 Winter_Park,_Florida]
[winner: 37537378 Admissions_(2011_film), pem_pos=12, final_scores=-1.7619554, final_local_scores=-1.7322876, pem=-4.6728287, similarity=-22.193638]
RECALL MISS: 109566 Winter_Park,_Florida]
2: 72011 Calhoun_County,_Florida
[winner: 52110 John_C._Calhoun, pem_pos=0, final_scores=-0.31721532, final_local_scores=-0.3697211, pem=-1.9563721, similarity=3.695553]
gt: 72011 Calhoun_County,_Florida, pem_pos=2, final_scores=-0.42029268, final_local_scores=-0.452035, pem=-2.5573304, similarity=4.507246]
3: 72369 Gulf_County,_Florida
[winner: 59218 Gulf_Oil, pem_pos=2, final_scores=-0.08564608, final_local_scores=-0.048428923, pem=-2.011698, similarity=13.584424]
gt: 72369 Gulf_County,_Florida, pem_pos=3, final_scores=-0.4163186, final_local_scores=-0.4554205, pem=-2.0428784, similarity=1.6085479]
4: 72386 Lafayette_County,_Florida
[winner: 115635 Lafayette,_Louisiana, pem_pos=0, final_scores=0.0222594, final_local_scores=-0.015686512, pem=-1.5915291, similarity=12.276579]

gt: 72386 Lafayette_County,_Florida, pem_pos=6,
 final_scores=-0.33668408, final_local_scores=-0.36372387,
 pem=-2.936742, similarity=9.205928]
 5: 73612 Leon_County,_Florida
 [winner: 73612 Leon_County,_Florida, pem_pos=0,
 final_scores=-0.22635011, final_local_scores=-0.29205132,
 pem=-2.1715562, similarity=7.183579]
 gt: 73612 Leon_County,_Florida, pem_pos=0,
 final_scores=-0.22635011, final_local_scores=-0.29205132,
 pem=-2.1715562, similarity=7.183579]
 6: 73618 Liberty_County,_Florida
 [winner: 122628 Liberty,_Missouri, pem_pos=1,
 final_scores=-0.40377283, final_local_scores=-0.44718903,
 pem=-2.801813, similarity=5.9813933]
 gt: 73618 Liberty_County,_Florida, pem_pos=10,
 final_scores=-0.6451765, final_local_scores=-0.6837625,
 pem=-3.7806487, similarity=4.2446322]
 7: 73621 Madison_County,_Florida
 [winner: 57863 Madison,_Wisconsin, pem_pos=0,
 final_scores=0.10681532, final_local_scores=0.0685952,
 pem=-1.5699619, similarity=14.674427]
 gt: 73621 Madison_County,_Florida, pem_pos=17,
 final_scores=-0.6072366, final_local_scores=-0.6617506,
 pem=-3.9533908, similarity=5.8409085]
 8: 75111 Taylor_County,_Florida
 [winner: 119280 Taylor,_Michigan, pem_pos=0,
 final_scores=-0.4142779, final_local_scores=-0.43799847,
 pem=-2.8038657, similarity=6.2668223]
 gt: 75111 Taylor_County,_Florida, pem_pos=4,
 final_scores=-0.5133827, final_local_scores=-0.5474501,
 pem=-3.597992, similarity=7.319149]
 9: 109285 Monticello,_Florida
 [winner: 52302 Monticello, pem_pos=0,
 final_scores=-0.048356414, final_local_scores=-0.044469878,
 pem=-0.7028594, similarity=6.584902]
 gt: 109285 Monticello,_Florida, pem_pos=5,
 final_scores=-0.37966913, final_local_scores=-0.43034562,
 pem=-2.853592, similarity=6.765617]
 10: 110375 Thomasville,_Georgia
 [winner: 48830 Georgia_(U.S._state), pem_pos=0,
 final_scores=0.32811755, final_local_scores=0.28857198,
 pem=-0.70047224, similarity=16.510553]
 RECALL MISS: 110375 Thomasville,_Georgia]
 [winner: 150829 Thomasville,_North_Carolina, pem_pos=0,
 final_scores=0.2264993, final_local_scores=0.24907598,
 pem=-0.8305334, similarity=16.039207]
 gt: 110375 Thomasville,_Georgia, pem_pos=1,
 final_scores=0.18713127, final_local_scores=0.17033231,
 pem=-1.1050901, similarity=15.182422]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.6818564, final_local_scores=-1.6447916,
 pem=-4.6728287, similarity=-19.582577]

RECALL MISS: 110375 Thomasville,_Georgia]

11: 18933066 Florida
[winner: 18933066 Florida, pem_pos=0,
final_scores=0.010436937, final_local_scores=-0.050448433,
pem=-0.15405534, similarity=3.422008|
gt: 18933066 Florida, pem_pos=0,
final_scores=0.010436937, final_local_scores=-0.050448433,
pem=-0.15405534, similarity=3.422008]

12: 633627 Florida_House_of_Representatives
[winner: 521625 Texas_House_of_Representatives,
pem_pos=0, final_scores=-0.49844062,
final_local_scores=-0.5910219, pem=-1.747669,
similarity=-4.043464|
gt: 633627 Florida_House_of_Representatives, pem_pos=6,
final_scores=-0.51149255, final_local_scores=-0.6187597,
pem=-2.5791028, similarity=-0.3497491]

13: 73618 Liberty_County,_Florida
[winner: 91518 Liberty_County,_Texas, pem_pos=0,
final_scores=-0.1648897, final_local_scores=-0.17475338,
pem=-0.5174323, similarity=1.6886007|
gt: 73618 Liberty_County,_Florida, pem_pos=2,
final_scores=-0.19028987, final_local_scores=-0.18686795,
pem=-0.8507097, similarity=3.139492]

14: 20500604 Tallahassee_Democrat
[winner: 57700 Tallahassee,_Florida, pem_pos=0,
final_scores=0.24163519, final_local_scores=0.18555096,
pem=-0.024764989, similarity=9.7616|
RECALL MISS: 20500604 Tallahassee_Democrat]
[winner: 5043544 Democratic_Party_(United_States),
pem_pos=0, final_scores=-0.20338629,
final_local_scores=-0.2802382, pem=-0.061652232,
similarity=-3.9378803|
RECALL MISS: 20500604 Tallahassee_Democrat]

15: 73618 Liberty_County,_Florida
[winner: 96745 Liberty_County,_Georgia, pem_pos=1,
final_scores=0.20358692, final_local_scores=0.2065903,
pem=-0.5235462, similarity=13.101905|
gt: 73618 Liberty_County,_Florida, pem_pos=2,
final_scores=0.05402799, final_local_scores=0.09898989,
pem=-0.8507097, similarity=11.67006]

16: 4016809 Apalachicola_Bay
[winner: 4016809 Apalachicola_Bay, pem_pos=0,
final_scores=-0.051979244, final_local_scores=-0.0061500072,
pem=0.0, similarity=3.9061844|
gt: 4016809 Apalachicola_Bay, pem_pos=0,
final_scores=-0.051979244, final_local_scores=-0.0061500072,
pem=0.0, similarity=3.9061844]

1: 32070 Republican_Party_(United_States)
[pem_pos=0, final_scores=0.43187314,
final_local_scores=0.41898823, pem=-0.1310191,
similarity=17.30566]

2: 633627 Florida_House_of_Representatives

[pem_pos=0, final_scores=0.14071815,
final_local_scores=0.09922163, pem=0.0, similarity=7.0506845]
3: 75116 Wakulla_County,_Florida
[pem_pos=0, final_scores=0.25673604,
final_local_scores=0.28174677, pem=0.0, similarity=12.497599]
4: 902849 Florida_Panhandle
[pem_pos=0, final_scores=0.094647065,
final_local_scores=0.110836074, pem=0.0, similarity=7.3972826]
5: 239297 Florida_State_University
[pem_pos=0, final_scores=0.22320299,
final_local_scores=0.21865475, pem=0.0, similarity=10.614807]
6: 60611 University_of_Florida
[pem_pos=0, final_scores=0.3089568,
final_local_scores=0.31411713, pem=0.0, similarity=13.463594]
7: 3434750 United_States
[pem_pos=0, final_scores=0.51339954,
final_local_scores=0.51703787, pem=-0.073457584,
similarity=19.918629]
8: 633627 Florida_House_of_Representatives
[pem_pos=0, final_scores=0.1090209,
final_local_scores=0.064597204, pem=0.0, similarity=6.017422]
9: 32070 Republican_Party_(United_States)
[pem_pos=0, final_scores=0.42541873,
final_local_scores=0.41193777, pem=-0.1310191,
similarity=17.09526]
10: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.48338813,
final_local_scores=0.4756412, pem=-0.076093495,
similarity=18.697603]
11: 633627 Florida_House_of_Representatives
[pem_pos=0, final_scores=0.18180852,
final_local_scores=0.1441066, pem=0.0, similarity=8.3901415]
12: 60611 University_of_Florida
[pem_pos=0, final_scores=0.18552344,
final_local_scores=0.17928496, pem=0.0, similarity=9.439934]
1: 109566 Winter_Park,_Florida
[pem_pos=0, final_scores=0.47737545,
final_local_scores=0.48059452, pem=-0.01514043,
similarity=18.513947]
2: 18933066 Florida
[pem_pos=0, final_scores=0.047894105,
final_local_scores=-0.0095321685, pem=-0.15405534,
similarity=4.643031]
3: 57863 Madison,_Wisconsin
[pem_pos=0, final_scores=0.10681532,
final_local_scores=0.0685952, pem=-1.5699619,
similarity=14.674427]
4: 150829 Thomasville,_North_Carolina
[pem_pos=0, final_scores=0.2264993,
final_local_scores=0.24907598, pem=-0.8305334,
similarity=16.039207]
5: 48830 Georgia_(U.S._state)

[pem_pos=0, final_scores=0.32811755,
 final_local_scores=0.28857198, pem=-0.70047224,
 similarity=16.510553]
 6: 1013237 Robert_Hill_(Australian_politician)
 [pem_pos=0, final_scores=0.064767554,
 final_local_scores=0.0967789, pem=-0.18092394,
 similarity=7.9616814]
 7: 57700 Tallahassee,_Florida
 [pem_pos=0, final_scores=0.24163519,
 final_local_scores=0.18555096, pem=-0.024764989,
 similarity=9.7616]
 8: 1013237 Robert_Hill_(Australian_politician)
 [pem_pos=0, final_scores=0.07348104,
 final_local_scores=0.106297046, pem=-0.18092394,
 similarity=8.245722]
 9: 96745 Liberty_County,_Georgia
 [pem_pos=1, final_scores=0.20358692,
 final_local_scores=0.2065903, pem=-0.5235462,
 similarity=13.101905]
 10: 23273510 Steve_Beshear
 [pem_pos=0, final_scores=0.28436077,
 final_local_scores=0.3159182, pem=0.0, similarity=13.517343]

Sample: b'Commonwealth_men'

The [1 Commonwealth 1] men , [2 Commonwealth 2] 's men , or
 Commonwealth Party were highly outspoken [3 British 3]
 Protestant religious , political , and economic reformers
 during the early 18th century . They were active in the
 movement called the [4 1 Country Party 4] . They promoted 2 re
 publicanism and had a great influence on [5 3 Republicanism 5]
 in the [6 United States 6] , but little impact in [7 Britain 7]
] . The most noted commonwealthmen were 4 John Trenchard and
 Thomas Gordon , who wrote the seminal work 5 Cato 's Letters
 between 1720 and 1723 . Other members include [8 Robert
 Crowley 8] , , [1 1 Thomas Beccon 1] , [2 2 Thomas Lever 2] ,
 and 6 John Hales . They condemned 7 corruption and lack of
 morality in [9 British 9] political life , theorizing that
 only 8 civic virtue could protect a country from despotism
 and ruin . Their criticism about 9 enclosure and the general
 material plight of the poor was particularly notable to early
 twentieth-century scholars like who saw in them a valuable
 though regrettably abortive form of that represented a
 preferable alternative to the view of [3 3 Max Weber 3] that [4 4 Protestantism 4] enabled and sustained the rise of 10 cap
 italism . On the other hand , it has been argued that the
 Commonwealthmen `` by no means stand against an
 individualistic or capitalistic spirit , and -- despite what -
 LSB- for example , historians [10 JGA Pocock 10] and Gordon
 Wood -RSB- have claimed -- are far from espousing classical
 virtue or the Aristotelian conception of man as zoon
 politikon -LSB- a political animal -RSB- . ' ' ' Although
 nearly all [11 British 11] politicians and thinkers rejected

the ideas of the commonwealth men , these writers had a powerful effect on [12 British 12] colonial [13 America 13] . It is estimated that half the private libraries in the held bound volumes of Cato 's Letters on their shelves . The Commonwealthman ideas of civic virtue , freedom , and government carefully regulated and controlled by the people were major principles in the 11 republicanism that became the dominant ideology of the [5 5 American Revolution 5]

1: 945776 Thomas_Becon

2: 21377575 Thomas_Lever

3: 19455 Max_Weber

4: 25814008 Protestantism

5: 1973 American_Revolution

1: 32891631 Country_Party_(Britain)

[winner: 21927 National_Party_of_Australia, pem_pos=0, final_scores=0.2921821, final_local_scores=0.33625045, pem=0.0, similarity=14.124097]

gt: 32891631 Country_Party_(Britain), pem_pos=7, final_scores=-0.6958519, final_local_scores=-0.7596388, pem=-5.2408648, similarity=9.921214]

2: 25755 Republicanism

[winner: 1283475 Republicanism_in_Canada, pem_pos=3, final_scores=-1.1268986, final_local_scores=-1.2250364, pem=-1.0860802, similarity=-26.561539]

gt: 25755 Republicanism, pem_pos=0, final_scores=-1.3294495, final_local_scores=-1.4966376, pem=-0.4223039, similarity=-38.276386]

3: 1764584 Republicanism_in_the_United_States

[winner: 3434750 United_States, pem_pos=0, final_scores=0.4995559, final_local_scores=0.50345415, pem=-0.073457584, similarity=19.513264]

RECALL MISS: 1764584 Republicanism_in_the_United_States]

[winner: 25755 Republicanism, pem_pos=0, final_scores=0.17531143, final_local_scores=0.16024649, pem=-0.34879944, similarity=10.768614]

gt: 1764584 Republicanism_in_the_United_States, pem_pos=2, final_scores=0.049258426, final_local_scores=0.037905663, pem=-0.49580377, similarity=7.9171515]

[winner: 2118516 German_federal_election,_March_1933, pem_pos=0, final_scores=-1.4760865, final_local_scores=-1.5549867, pem=-0.79592705, similarity=-37.985813]

RECALL MISS: 1764584 Republicanism_in_the_United_States]

4: 1789750 John_Trenchard_(writer)

[winner: 1789750 John_Trenchard_(writer), pem_pos=1, final_scores=0.018460408, final_local_scores=0.052639365, pem=-0.2678007, similarity=7.1169186]

gt: 1789750 John_Trenchard_(writer), pem_pos=1, final_scores=0.018460408, final_local_scores=0.052639365, pem=-0.2678007, similarity=7.1169186]

5: 376161 Cato's_Letters

[winner: 312698 Cato_the_Younger, pem_pos=0,
 final_scores=-0.12269811, final_local_scores=-0.079034105,
 pem=-1.1297225, similarity=7.8747845|
 RECALL MISS: 376161 Cato's_Letters]
 [winner: 18963870 Literature, pem_pos=3,
 final_scores=-0.37470335, final_local_scores=-0.41891238,
 pem=-1.8126833, similarity=1.4461863|
 RECALL MISS: 376161 Cato's_Letters]
 [winner: 1415027 David_Berman_(musician), pem_pos=5,
 final_scores=-2.4451668, final_local_scores=-2.5687191,
 pem=-3.0012724, similarity=-56.244602|
 RECALL MISS: 376161 Cato's_Letters]
 6: 1078624 John_Hales
 [winner: 1078624 John_Hales, pem_pos=0,
 final_scores=0.0077522844, final_local_scores=0.0536166,
 pem=-0.13595563, similarity=6.429087|
 gt: 1078624 John_Hales, pem_pos=0,
 final_scores=0.0077522844, final_local_scores=0.0536166,
 pem=-0.13595563, similarity=6.429087]
 7: 66220 Political_corruption
 [winner: 319947 Kwame_Kilpatrick, pem_pos=2,
 final_scores=-1.3150717, final_local_scores=-1.3751585,
 pem=-0.6931472, similarity=-33.17832|
 gt: 66220 Political_corruption, pem_pos=0,
 final_scores=-1.4614017, final_local_scores=-1.5829096,
 pem=-0.48016536, similarity=-40.536255]
 8: 2123148 Civic_virtue
 [winner: 49901 Virtue, pem_pos=0,
 final_scores=-0.7913808, final_local_scores=-0.8201611,
 pem=-0.04034033, similarity=-20.166153|
 RECALL MISS: 2123148 Civic_virtue]
 [winner: 2025952 Civic_Party_of_Montenegro, pem_pos=13,
 final_scores=-1.1070008, final_local_scores=-1.0308881,
 pem=-2.4849067, similarity=-13.160734|
 RECALL MISS: 2123148 Civic_virtue]
 9: 186721 Enclosure
 [winner: 186721 Enclosure, pem_pos=0,
 final_scores=-0.9844358, final_local_scores=-1.066097,
 pem=-0.44966057, similarity=-25.279425|
 gt: 186721 Enclosure, pem_pos=0,
 final_scores=-0.9844358, final_local_scores=-1.066097,
 pem=-0.44966057, similarity=-25.279425]
 10: 5416 Capitalism
 [winner: 21705978 Rhine_capitalism, pem_pos=0,
 final_scores=-1.0533117, final_local_scores=-1.0837924,
 pem=0.0, similarity=-28.252813|
 gt: 5416 Capitalism, pem_pos=1, final_scores=-1.1093445,
 final_local_scores=-1.2257435, pem=-0.026048161,
 similarity=-32.347267]
 11: 1764584 Republicanism_in_the_United_States
 [winner: 1283475 Republicanism_in_Canada, pem_pos=3,
 final_scores=-1.2290831, final_local_scores=-1.3366573,

pem=-1.0860802, similarity=-29.892532|
gt: 1764584 Republicanism_in_the_United_States,
pem_pos=1, final_scores=-1.2494466,
final_local_scores=-1.3951867, pem=-0.5150642,
similarity=-34.744442]
1: 945776 Thomas_Becon
[pem_pos=0, final_scores=0.031267688,
final_local_scores=0.11511266, pem=0.0, similarity=7.5249043]
2: 21377575 Thomas_Lever
[pem_pos=0, final_scores=0.17343979,
final_local_scores=0.25640413, pem=0.0, similarity=11.741323]
3: 19455 Max_Weber
[pem_pos=0, final_scores=0.29476374,
final_local_scores=0.3343411, pem=0.0, similarity=14.067119]
4: 25814008 Protestantism
[pem_pos=0, final_scores=0.21952777,
final_local_scores=0.1887733, pem=0.0, similarity=9.723084]
5: 1973 American_Revolution
[pem_pos=0, final_scores=0.27035254,
final_local_scores=0.23625192, pem=0.0, similarity=11.139941]
1: 21175158 Commonwealth_of_Nations
[pem_pos=0, final_scores=0.43040437,
final_local_scores=0.42461532, pem=-0.69133765,
similarity=20.520683]
2: 21175158 Commonwealth_of_Nations
[pem_pos=0, final_scores=0.2913087,
final_local_scores=0.27267444, pem=-0.69133765,
similarity=15.9864645]
3: 31717 United_Kingdom
[pem_pos=0, final_scores=0.22545461,
final_local_scores=0.18588278, pem=-0.48936236,
similarity=12.298055]
4: 21927 National_Party_of_Australia
[pem_pos=0, final_scores=0.2921821,
final_local_scores=0.33625045, pem=0.0, similarity=14.124097]
5: 25755 Republicanism
[pem_pos=0, final_scores=0.17531143,
final_local_scores=0.16024649, pem=-0.34879944,
similarity=10.768614]
6: 3434750 United_States
[pem_pos=0, final_scores=0.4995559,
final_local_scores=0.50345415, pem=-0.073457584,
similarity=19.513264]
7: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5359277,
final_local_scores=0.52502745, pem=-0.61114657,
similarity=23.081089]
8: 1667878 Robert_Crowley_(printer)
[pem_pos=0, final_scores=0.1157725,
final_local_scores=0.1693545, pem=0.0, similarity=9.143589]
9: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5590654,

```

final_local_scores=0.55030185, pem=-0.48936236,
similarity=23.173046]
10: 3510041 J._G._A._Pocock
    [pem_pos=0, final_scores=0.15189935,
final_local_scores=0.1829777, pem=0.0, similarity=9.550133]
11: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.538518,
final_local_scores=0.52785695, pem=-0.48936236,
similarity=22.503246]
12: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.39597613,
final_local_scores=0.37215155, pem=-0.48936236,
similarity=17.856686]
13: 3434750 United_States
    [pem_pos=0, final_scores=0.3290177,
final_local_scores=0.31716704, pem=-0.78483206,
similarity=17.822649]

```

Sample: b'Cave_Rock_Tunnel'

The [1 Cave Rock Tunnel 1] is a dual bore highway tunnel on 1 U.S. Route 50 -LRB- US 50 -RRB- along the eastern shore of [1 1 Lake Tahoe 1] approximately seven miles -LRB- 11.4 km -RRB- north of [2 2 Stateline 2] , in [2 2 Douglas County 2] , [3 3 Nevada 3] , . To the 3 Washoe Indian Tribe , [3 Cave Rock 3] is considered a sacred place and the tribe has placed restrictions on recreational activities in the vicinity of the tunnel . The tunnels carry through [4 Cave Rock 4] , a mountain along the eastern shore of [4 4 Lake Tahoe 4] . There are numerous small caves adjacent to the south portal of both tunnels which give the rock its name . The tunnel is located between the towns of [5 4 Zephyr Cove 5] and 5 Glenbrook along the US 50 corridor . This portion of US 50 is a [5 5 National Scenic Byway 5] , part of the . The westbound bore is long and features exposed rock ; the eastbound bore features a concrete liner and is long . The tunnels are an elevation of at approximately , about above the level of the lake . The tunnel dates back to the [6 6 Lincoln Highway 6] . Originally the [6 Lincoln Highway 6] was routed along a single lane and rock wall built in 1863 . Recognizing the inadequacy of the single lane road , efforts began to improve capacity on the primary road to [7 Lake Tahoe 7] . The first bore was constructed in 1931 , as part of a reconstruction of a section of the [8 Lincoln Highway 8] . Concerned about damaging [9 Cave Rock 9] , the project managers employed key people from the recently completed in what is now [7 7 Zion National Park 7] in [10 Utah 10] . The first traffic began flowing through the bore in mid-September of that year . Construction on the second bore began in 1957 , when US 50 was widened to four lanes , at a cost of just over \$ 450,000 -LRB- equivalent to \$ million today -RRB- . Coincidentally , both bores were constructed by [11 Utah 11] based construction companies . The [12 Cave Rock 12] area is

considered sacred by the [13 6 Washoe 13] tribe . The [14 Washoe 14] tribal leaders were not consulted about the construction of either bore , and were upset about the perceived desecration of their tribal lands . Within the last decade , the Washoe Tribe has had a larger influence on [15 Cave Rock 15] and its historic preservation . In 2007 , the Federal Government ruled on a precedent-setting case that has restricted activities around the tunnel , such as 7 rock climbing

1: 140899 Lake_Tahoe

2: 124409 Stateline,_Nevada

3: 21216 Nevada

4: 140899 Lake_Tahoe

5: 802772 National_Scenic_Byway

6: 80815 Lincoln_Highway

7: 18952631 Zion_National_Park

1: 5607575 U.S._Route_50_in_Nevada

[winner: 43951 United_States_Numbered_Highways,
pem_pos=0, final_scores=-0.15057841,
final_local_scores=-0.1859017, pem=-0.17589068,
similarity=-0.50144356|

RECALL MISS: 5607575 U.S._Route_50_in_Nevada]

[winner: 1347230 Qaf_(sura), pem_pos=11,
final_scores=-0.7905238, final_local_scores=-0.55574083,
pem=-3.7918332, similarity=8.125876|

RECALL MISS: 5607575 U.S._Route_50_in_Nevada]

2: 87960 Douglas_County,_Nevada

[winner: 56310 Douglas_County,_Colorado, pem_pos=0,
final_scores=0.25430983, final_local_scores=0.25520012,
pem=-1.262038, similarity=18.568552|

gt: 87960 Douglas_County,_Nevada, pem_pos=4,
final_scores=0.10185628, final_local_scores=0.075202495,
pem=-1.8117483, similarity=16.186474]

3: 705836 Washoe_people

[winner: 93593 Washoe_County,_Nevada, pem_pos=1,
final_scores=-1.0108105, final_local_scores=-1.1724503,
pem=-0.55221313, similarity=-27.895521|

gt: 705836 Washoe_people, pem_pos=2,
final_scores=-1.0848572, final_local_scores=-1.1708806,
pem=-0.87867063, similarity=-26.07335]

4: 379424 Zephyr_Cove,_Nevada

[winner: 124410 Zephyr_Cove-Round_Hill_Village,_Nevada,
pem_pos=0, final_scores=0.04890777,
final_local_scores=0.11624202, pem=-0.21511137,
similarity=8.7284155|

gt: 379424 Zephyr_Cove,_Nevada, pem_pos=1,
final_scores=-0.16819774, final_local_scores=-0.15237746,
pem=-1.6422278, similarity=8.473156]

5: 6820001 Glenbrook,_Nevada

[winner: 6274273 Glenbrook_(Stamford), pem_pos=0,
final_scores=-0.06569758, final_local_scores=-0.009512752,
pem=-1.2819371, similarity=10.777206|

gt: 6820001 Glenbrook, Nevada, pem_pos=6,
final_scores=-0.16486214, final_local_scores=-0.15901315,
pem=-1.9704413, similarity=10.060009]
6: 705836 Washoe_people
[winner: 93593 Washoe County, Nevada, pem_pos=1,
final_scores=0.083515, final_local_scores=0.028012305,
pem=-0.55221313, similarity=7.928677]
gt: 705836 Washoe_people, pem_pos=2,
final_scores=-0.18257251, final_local_scores=-0.18573657,
pem=-0.87867063, similarity=3.325311]
7: 929786 Rock_climbing
[winner: 55134 Rock_climbing_in_the_Peak_District,
pem_pos=2, final_scores=-0.91557676,
final_local_scores=-0.8706721, pem=-0.6926663,
similarity=-18.126053]
gt: 929786 Rock_climbing, pem_pos=0,
final_scores=-1.1491891, final_local_scores=-1.2391958,
pem=-0.037740305, similarity=-32.685123]
1: 140899 Lake_Tahoe
[pem_pos=0, final_scores=0.2700855,
final_local_scores=0.236821, pem=0.0, similarity=11.156924]
2: 124409 Stateline, Nevada
[pem_pos=0, final_scores=0.18419082,
final_local_scores=0.20227933, pem=-0.21874583,
similarity=11.315705]
3: 21216 Nevada
[pem_pos=0, final_scores=0.49849993,
final_local_scores=0.4919598, pem=-0.2665277,
similarity=20.220194]
4: 140899 Lake_Tahoe
[pem_pos=0, final_scores=0.28013438,
final_local_scores=0.24779785, pem=0.0, similarity=11.484495]
5: 802772 National_Scenic_Byway
[pem_pos=0, final_scores=0.22088306,
final_local_scores=0.22297016, pem=0.0, similarity=10.7435875]
6: 80815 Lincoln_Highway
[pem_pos=0, final_scores=0.4330191,
final_local_scores=0.4269523, pem=0.0, similarity=16.830822]
7: 18952631 Zion_National_Park
[pem_pos=0, final_scores=0.33718628,
final_local_scores=0.34013763, pem=0.0, similarity=14.240098]
1: 17393712 Cave_Rock_Tunnel
[pem_pos=0, final_scores=0.13106929,
final_local_scores=0.12638888, pem=0.0, similarity=7.8614087]
2: 56310 Douglas_County, Colorado
[pem_pos=0, final_scores=0.25430983,
final_local_scores=0.25520012, pem=-1.262038,
similarity=18.568552]
3: 17393712 Cave_Rock_Tunnel
[pem_pos=0, final_scores=0.15045695,
final_local_scores=0.14756697, pem=0.0, similarity=8.493405]
4: 17393712 Cave_Rock_Tunnel

[pem_pos=0, final_scores=0.15267296,
 final_local_scores=0.14998761, pem=0.0, similarity=8.565642]
 5: 124410 Zephyr_Cove-Round_Hill_Village,_Nevada
 [pem_pos=0, final_scores=0.04890777,
 final_local_scores=0.11624202, pem=-0.21511137,
 similarity=8.7284155]
 6: 80815 Lincoln_Highway
 [pem_pos=0, final_scores=0.29222357,
 final_local_scores=0.2731546, pem=0.0, similarity=12.241191]
 7: 140899 Lake_Tahoe
 [pem_pos=0, final_scores=0.22851549,
 final_local_scores=0.19141203, pem=0.0, similarity=9.801829]
 8: 80815 Lincoln_Highway
 [pem_pos=0, final_scores=0.4517333,
 final_local_scores=0.44739473, pem=0.0, similarity=17.440865]
 9: 17393712 Cave_Rock_Tunnel
 [pem_pos=0, final_scores=0.068663195,
 final_local_scores=0.058219716, pem=0.0, similarity=5.8271055]
 10: 31716 Utah
 [pem_pos=0, final_scores=0.5436268,
 final_local_scores=0.53442055, pem=-0.2611261,
 similarity=21.457932]
 11: 31716 Utah
 [pem_pos=0, final_scores=0.42669344,
 final_local_scores=0.40668863, pem=-0.2611261,
 similarity=17.646156]
 12: 17393712 Cave_Rock_Tunnel
 [pem_pos=0, final_scores=0.082260504,
 final_local_scores=0.0730727, pem=0.0, similarity=6.270348]
 13: 93593 Washoe_County,_Nevada
 [pem_pos=1, final_scores=0.083515,
 final_local_scores=0.028012305, pem=-0.55221313,
 similarity=7.928677]
 14: 93593 Washoe_County,_Nevada
 [pem_pos=1, final_scores=0.064536646,
 final_local_scores=0.007281333, pem=-0.55221313,
 similarity=7.3100233]
 15: 17393712 Cave_Rock_Tunnel
 [pem_pos=0, final_scores=0.04112287,
 final_local_scores=0.028136075, pem=0.0, similarity=4.9293494]

Sample: b'Barrett_Green'

Barrett Nicholas Green -LRB- born October 29 , 1977 -RRB- is
 a former professional [1 1 American 1] football 2 linebacker
 in the [1 1 National Football League 1] . He played 3 college
 football for [2 4 West Virginia 2] and was drafted in the
 second round -LRB- 50th overall -RRB- by the [2 2 Detroit Lio
 ns 2] in the 5 2000 [3 NFL Draft 3] . Green was born in [4 6 W
 est Palm Beach 4] , Florida and attended . At Suncoast , he
 was a wrestler and played baseball along with football .
 Following his senior season he played in the South Florida
 All-Star Game . As a true freshman in 1995 , Green 7 redshirt

ed but was still named Rookie of the Year due to his performance on the 8 scout team . In 1996 , Green was the back up and played in every game . For the season he recorded 12 9 tackles and a pass deflection . Green was elevated to the starting strong safety as a redshirt sophomore in 1997 and played in every game . He finished third on West Virginia with 71 tackles and also had three 10 sacks . Green added a fumble recovery , six pass deflections and four to his season totals . Against [3 3 Notre Dame 3] he had nine tackles which was his season high and he also had eight tackles against 11 Maryland . Green recorded the longest non-touchdown interception return in history with an 83 yard return against [4 4 Rutgers 4] . In 1998 , Green moved forward to the starting and added 84 more tackles than the previous season with a team -- leading 155 tackles and set a school record with 107 solo tackles . His 155 tackles was the seventh best in West Virginia school history . His season high in tackles was 19 against 12 Navy and 13 Miami -LRB- Fla. -RRB- . He added 15 tackles against 14 Syracuse and 14 in the [5 15 Boston College 5] game . In the 16 1999 Insight.com Bowl against 1 7 Missouri , he played with a cracked rib and yet he still recorded a team high 14 tackles . Following the season he was named First team [6 All Big East 6] . Green was named First team [7 All Big East 7] for a second straight year in 1999 a season in which he led the team once again with 126 tackles . Three sacks and seven tackles for a loss added to his season totals . He recorded 18 tackles against Navy , 17 against [8 1 8 Virginia Tech 8] and 15 against Syracuse . Green father was former 19 Bowling Green

1: 21211 National_Football_League

2: 8577 Detroit_Lions

3: 1245667 Notre_Dame_Fighting_Irish_football

4: 8176029 Rutgers_Scarlet_Knights_football

1: 18951490 American_football

[winner: 3434750 United_States, pem_pos=0, final_scores=0.41188127, final_local_scores=0.46041423, pem=-0.12994486, similarity=18.536053]

gt: 18951490 American_football, pem_pos=6, final_scores=-0.523885, final_local_scores=-0.63602716, pem=-5.032821, similarity=12.478658]

[winner: 143183 Football_(ball), pem_pos=5, final_scores=-2.6736376, final_local_scores=-2.847017, pem=-4.263315, similarity=-57.68638]

gt: 18951490 American_football, pem_pos=1, final_scores=-2.7567883, final_local_scores=-3.0770152, pem=-1.8467362, similarity=-77.69172]

2: 492475 Linebacker

[winner: 492475 Linebacker, pem_pos=0, final_scores=-1.4243937, final_local_scores=-1.6197486, pem=-0.0010181483, similarity=-44.24128]

gt: 492475 Linebacker, pem_pos=0, final_scores=-1.4243937, final_local_scores=-1.6197486,

pem=-0.0010181483, similarity=-44.24128]
3: 6771 College_football
[winner: 6771 College_football, pem_pos=0,
final_scores=-1.225926, final_local_scores=-1.427571,
pem=-0.0074130204, similarity=-38.471542|
gt: 6771 College_football, pem_pos=0,
final_scores=-1.225926, final_local_scores=-1.427571,
pem=-0.0074130204, similarity=-38.471542]
4: 6969836 West_Virginia_Mountaineers_football
[winner: 32905 West_Virginia, pem_pos=0,
final_scores=0.11775537, final_local_scores=0.15251288,
pem=-0.30025622, similarity=10.273841|
gt: 6969836 West_Virginia_Mountaineers_football,
pem_pos=2, final_scores=-0.17278625,
final_local_scores=-0.26237077, pem=-3.0787387,
similarity=13.002705]
5: 3495580 2000_NFL_Draft
[winner: 21211 National_Football_League, pem_pos=0,
final_scores=0.17496355, final_local_scores=0.111979485,
pem=0.0, similarity=7.431404|
RECALL MISS: 3495580 2000_NFL_Draft]
[winner: 34548 2000, pem_pos=0, final_scores=-1.1952405,
final_local_scores=-1.1594822, pem=-0.8862115,
similarity=-25.692194|
RECALL MISS: 3495580 2000_NFL_Draft]
6: 109634 West_Palm_Beach, Florida
[winner: 109634 West_Palm_Beach, Florida, pem_pos=0,
final_scores=0.46174848, final_local_scores=0.49303657,
pem=0.0, similarity=18.802908|
gt: 109634 West_Palm_Beach, Florida, pem_pos=0,
final_scores=0.46174848, final_local_scores=0.49303657,
pem=0.0, similarity=18.802908]
[winner: 18933066 Florida, pem_pos=0,
final_scores=-0.035503313, final_local_scores=-0.045969404,
pem=-0.15405534, similarity=3.5556712|
RECALL MISS: 109634 West_Palm_Beach, Florida]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7732246, final_local_scores=-1.7460907,
pem=-4.6728287, similarity=-22.605545|
RECALL MISS: 109634 West_Palm_Beach, Florida]
7: 3882910 Redshirt_(college_sports)
[winner: 3882910 Redshirt_(college_sports), pem_pos=0,
final_scores=-1.6789017, final_local_scores=-1.8790557,
pem=0.0, similarity=-51.985058|
gt: 3882910 Redshirt_(college_sports), pem_pos=0,
final_scores=-1.6789017, final_local_scores=-1.8790557,
pem=0.0, similarity=-51.985058]
8: 4841925 Scout_team
[winner: 23616516 The_Scout_(Buffalo_Bill_statue),
pem_pos=2, final_scores=-1.0293082,
final_local_scores=-0.9999121, pem=-1.3862944,
similarity=-18.210773|

RECALL MISS: 4841925 Scout_team]
 [winner: 11172882 1945_Cleveland_Rams_season, pem_pos=14,
 final_scores=-1.7023335, final_local_scores=-1.768913,
 pem=-4.3721313, similarity=-24.921852|
 RECALL MISS: 4841925 Scout_team]
 9: 10667 Tackle_(football_move)
 [winner: 5587977 Tackle_(American_and_Canadian_football),
 pem_pos=0, final_scores=-1.3072158,
 final_local_scores=-1.4891992, pem=0.0,
 similarity=-40.350964|
 gt: 10667 Tackle_(football_move), pem_pos=1,
 final_scores=-1.319556, final_local_scores=-1.4576035,
 pem=-0.03763886, similarity=-39.203396]
 10: 540877 Quarterback_sack
 [winner: 540877 Quarterback_sack, pem_pos=0,
 final_scores=-1.2214799, final_local_scores=-1.3629313,
 pem=-0.035632085, similarity=-36.389103|
 gt: 540877 Quarterback_sack, pem_pos=0,
 final_scores=-1.2214799, final_local_scores=-1.3629313,
 pem=-0.035632085, similarity=-36.389103]
 11: 6969824 Maryland_Terrapins_football
 [winner: 2566235 Russell_Maryland, pem_pos=9,
 final_scores=-0.2215901, final_local_scores=-0.22658211,
 pem=-4.371563, similarity=21.10129|
 gt: 6969824 Maryland_Terrapins_football, pem_pos=2,
 final_scores=-0.23288451, final_local_scores=-0.32565317,
 pem=-3.6678896, similarity=14.318128]
 12: 7305146 Navy_Midshipmen_football
 [winner: 20518076 United_States_Navy, pem_pos=0,
 final_scores=-0.12074758, final_local_scores=-0.10219163,
 pem=-1.0602508, similarity=6.805919|
 gt: 7305146 Navy_Midshipmen_football, pem_pos=2,
 final_scores=-0.15119879, final_local_scores=-0.21621653,
 pem=-2.65533, similarity=12.077477]
 13: 2517648 Miami_Hurricanes_football
 [winner: 53846 Miami, pem_pos=0, final_scores=-0.0681126,
 final_local_scores=-0.103812486, pem=-0.35396606,
 similarity=2.9166636|
 gt: 2517648 Miami_Hurricanes_football, pem_pos=3,
 final_scores=-0.3679735, final_local_scores=-0.46921065,
 pem=-3.6239817, similarity=9.795306]
 [winner: 84504 St._Petersburg_Florida, pem_pos=1,
 final_scores=-0.3994075, final_local_scores=-0.47324535,
 pem=-0.68339103, similarity=-6.3164854|
 RECALL MISS: 2517648 Miami_Hurricanes_football]
 14: 7958192 Syracuse_Orange_football
 [winner: 671171 Syracuse_Chiefs, pem_pos=9,
 final_scores=-0.059251666, final_local_scores=-0.05108714,
 pem=-3.6818287, similarity=22.587526|
 gt: 7958192 Syracuse_Orange_football, pem_pos=8,
 final_scores=-0.42591852, final_local_scores=-0.5257499,
 pem=-3.6644962, similarity=8.328387]

15: 7956873 Boston_College_Eagles_football
[winner: 239811 Boston_College, pem_pos=0,
final_scores=0.14424087, final_local_scores=0.09464328,
pem=-0.23712422, similarity=8.203575]
gt: 7956873 Boston_College_Eagles_football, pem_pos=1,
final_scores=-0.04030375, final_local_scores=-0.10631895,
pem=-1.7646937, similarity=10.513622]
16: 17527736 1999_Insight.com_Bowl
[winner: 335808 Buffalo_Wild_Wings_Bowl, pem_pos=0,
final_scores=-0.044022307, final_local_scores=-0.04323809,
pem=0.0, similarity=2.7994022]
gt: 17527736 1999_Insight.com_Bowl, pem_pos=1,
final_scores=-0.19946098, final_local_scores=-0.11983389,
pem=-0.91714144, similarity=5.5011883]
[winner: 34553 1999, pem_pos=0, final_scores=-1.4750413,
final_local_scores=-1.519794, pem=-0.6240096,
similarity=-37.87051]
RECALL MISS: 17527736 1999_Insight.com_Bowl]
17: 7957736 Missouri_Tigers_football
[winner: 19571 Missouri, pem_pos=0,
final_scores=-0.09182306, final_local_scores=-0.11808456,
pem=-0.27068004, similarity=2.0378344]
gt: 7957736 Missouri_Tigers_football, pem_pos=3,
final_scores=-0.26601565, final_local_scores=-0.3394501,
pem=-3.4871345, similarity=12.923425]
18: 15190346 Virginia_Tech_Hokies_football
[winner: 239897 Virginia_Tech, pem_pos=0,
final_scores=0.041298345, final_local_scores=-0.011424154,
pem=-0.6843117, similarity=7.470186]
gt: 15190346 Virginia_Tech_Hokies_football, pem_pos=2,
final_scores=-0.048899278, final_local_scores=-0.1075401,
pem=-1.640168, similarity=9.799991]
19: 9725713 Bowling_Green_Falcons_football
[winner: 275181 Bowling_Green_State_University,
pem_pos=4, final_scores=-0.28920752,
final_local_scores=-0.36894435, pem=-2.2925677,
similarity=5.547018]
gt: 9725713 Bowling_Green_Falcons_football, pem_pos=6,
final_scores=-0.31483704, final_local_scores=-0.37408635,
pem=-2.5489264, similarity=6.7876873]
1: 21211 National_Football_League
[pem_pos=0, final_scores=0.3641495,
final_local_scores=0.31412306, pem=0.0, similarity=13.463772]
2: 8577 Detroit_Lions
[pem_pos=0, final_scores=0.24227439,
final_local_scores=0.20449102, pem=-0.088858426,
similarity=10.675359]
3: 1245667 Notre_Dame_Fighting_Irish_football
[pem_pos=1, final_scores=0.15913264,
final_local_scores=0.106140405, pem=-1.3761536,
similarity=14.740891]
4: 8176029 Rutgers_Scarlet_Knights_football

[pem_pos=1, final_scores=0.16164128,
 final_local_scores=0.11558023, pem=-1.4275191,
 similarity=15.3019285]
 1: 3434750 United_States
 [pem_pos=0, final_scores=0.41188127,
 final_local_scores=0.46041423, pem=-0.12994486,
 similarity=18.536053]
 2: 32905 West_Virginia
 [pem_pos=0, final_scores=0.11775537,
 final_local_scores=0.15251288, pem=-0.30025622,
 similarity=10.273841]
 3: 21211 National_Football_League
 [pem_pos=0, final_scores=0.17496355,
 final_local_scores=0.111979485, pem=0.0, similarity=7.431404]
 4: 109634 West_Palm_Beach,Florida
 [pem_pos=0, final_scores=0.46174848,
 final_local_scores=0.49303657, pem=0.0, similarity=18.802908]
 5: 239811 Boston_College
 [pem_pos=0, final_scores=0.14424087,
 final_local_scores=0.09464328, pem=-0.23712422,
 similarity=8.203575]
 6: 78024 Big_East_Conference_(1979-2013)
 [pem_pos=0, final_scores=0.10921742,
 final_local_scores=0.07369143, pem=0.0, similarity=6.288812]
 7: 78024 Big_East_Conference_(1979-2013)
 [pem_pos=0, final_scores=0.18470864,
 final_local_scores=0.15615413, pem=0.0, similarity=8.749663]
 8: 239897 Virginia_Tech
 [pem_pos=0, final_scores=0.041298345,
 final_local_scores=-0.011424154, pem=-0.6843117,
 similarity=7.470186]

Sample: b'Electoral_district_of_Colton'

Colton is an of the **1 House of Assembly** in the **[1 1 Australia 1]** n state of **[2 2 South Australia 2]** . It is a 14.2 km ² urban electorate on **[3 3 Adelaide 3]** 's western beaches , taking in the suburbs of **2 Fulham** , **[4 4 Fulham Gardens 4]** , **3 Grange** , **[5 5 Henley Beach 5]** , **[1 4 Henley Beach South 1]** , **5 Kidman Park** as well as parts of **6 Findon** , **7 Lockleys** and **8 Seaton** . Colton was created in the 1991 electoral distribution as a nominal seat for the **[6 6 Australian Labor Party 6]** , replacing the abolished . It was first contested at the **9 1993 state election** , where it was won in a large swing to the **[7 7 Liberal Party 7]** by former Lord Mayor **10 Steve Condous** , recording 57.7 % of the two-party preferred vote . This was reduced in 1997 to 51 % , and upon **[2 Condous 2]** ' retirement , it was captured by **[8 8 Paul Caica 8]** for the **[3 ALP 3]** with a 5.6 % swing . The electoral district is named after Mary Colton , who arrived in **[9 9 Adelaide 9]** in 1839 and worked for the welfare of women and children . She was the President of the Women 's Suffrage League , and lived to see the introduction of equal voting rights for women in

1894 . Colton is the only current lower house seat in parliament to have been consistently won by the party that forms government , with only one other seat changing hands from [4 Liberal 4] to [5 Labor 5] when [6 Labor 6] formed government in 2002 , the seat of [7 11 Adelaide 7]

1: 4689264 Australia

2: 26716 South_Australia

3: 1148 Adelaide

4: 8758947 Fulham_Gardens,_South_Australia

5: 5662235 Henley_Beach,_South_Australia

6: 1495 Australian_Labor_Party

7: 18453 Liberal_Party_of_Australia

8: 3706488 Paul_Caica

9: 1148 Adelaide

1: 579453 South_Australian_House_of_Assembly

[winner: 363804 House_of_Assembly, pem_pos=0, final_scores=-0.078788534, final_local_scores=-0.07195921, pem=-1.255375, similarity=8.769231]

gt: 579453 South_Australian_House_of_Assembly, pem_pos=1, final_scores=-0.14096327, final_local_scores=-0.2098288, pem=-2.0483763, similarity=8.967392]

2: 7056640 Fulham,_South_Australia

[winner: 11228 Fulham_F.C., pem_pos=0, final_scores=-0.040693328, final_local_scores=0.03459625, pem=-0.3209339, similarity=6.8674226]

gt: 7056640 Fulham,_South_Australia, pem_pos=7, final_scores=-0.47012588, final_local_scores=-0.45547473, pem=-2.5972002, similarity=4.6214166]

3: 2531732 Grange,_South_Australia

[winner: 2531732 Grange,_South_Australia, pem_pos=3, final_scores=-0.2811321, final_local_scores=-0.3035993, pem=-2.4687304, similarity=8.455044]

gt: 2531732 Grange,_South_Australia, pem_pos=3, final_scores=-0.2811321, final_local_scores=-0.3035993, pem=-2.4687304, similarity=8.455044]

4: 13151720 Henley_Beach_South,_South_Australia

[winner: 1148 Adelaide, pem_pos=1, final_scores=0.37382573, final_local_scores=0.3344202, pem=-0.6931472, similarity=17.83892]

gt: 13151720 Henley_Beach_South,_South_Australia, pem_pos=0, final_scores=0.16217618, final_local_scores=0.21408248, pem=0.0, similarity=10.478361]

5: 8915481 Kidman_Park,_South_Australia

[winner: 8915481 Kidman_Park,_South_Australia, pem_pos=0, final_scores=-0.048765823, final_local_scores=-0.014646441, pem=0.0, similarity=3.652634]

gt: 8915481 Kidman_Park,_South_Australia, pem_pos=0, final_scores=-0.048765823, final_local_scores=-0.014646441, pem=0.0, similarity=3.652634]

6: 12476797 Findon,_South_Australia

[winner: 12476797 Findon,_South_Australia, pem_pos=1,

final_scores=-0.03521374, final_local_scores=-0.02345451,
pem=-0.6184664, similarity=6.7531|
gt: 12476797 Findon, South_Australia, pem_pos=1,
final_scores=-0.03521374, final_local_scores=-0.02345451,
pem=-0.6184664, similarity=6.7531]
7: 5677629 Lockleys, South_Australia
[winner: 5677629 Lockleys, South_Australia, pem_pos=0,
final_scores=0.01046209, final_local_scores=0.034173653,
pem=0.0, similarity=5.109523|
gt: 5677629 Lockleys, South_Australia, pem_pos=0,
final_scores=0.01046209, final_local_scores=0.034173653,
pem=0.0, similarity=5.109523]
8: 10150119 Seaton, South_Australia
[winner: 374524 Seaton, Devon, pem_pos=0,
final_scores=-0.08421962, final_local_scores=-0.03944719,
pem=-1.021337, similarity=8.46672|
gt: 10150119 Seaton, South_Australia, pem_pos=2,
final_scores=-0.4520892, final_local_scores=-0.44945472,
pem=-2.2766438, similarity=3.0578299]
9: 5338982 South_Australian_state_election, 1993
[winner: 34598 1993, pem_pos=0, final_scores=-0.6775032,
final_local_scores=-0.5047842, pem=-0.7063994,
similarity=-7.1325464|
RECALL MISS: 5338982
South_Australian_state_election, 1993]
[winner: 15563529 New_York_state_election, 1928,
pem_pos=5, final_scores=-1.0974846,
final_local_scores=-1.0975933, pem=-2.5788548,
similarity=-14.640445|
RECALL MISS: 5338982
South_Australian_state_election, 1993]
10: 5864573 Steve_Condous
[winner: 5864573 Steve_Condous, pem_pos=0,
final_scores=0.013995364, final_local_scores=0.047194988,
pem=0.0, similarity=5.4981055|
gt: 5864573 Steve_Condous, pem_pos=0,
final_scores=0.013995364, final_local_scores=0.047194988,
pem=0.0, similarity=5.4981055]
11: 3673541 Electoral_district_of_Adelaide
[winner: 1148 Adelaide, pem_pos=0,
final_scores=0.18978243, final_local_scores=0.13457662,
pem=-0.29570588, similarity=9.713842|
gt: 3673541 Electoral_district_of_Adelaide, pem_pos=9,
final_scores=-0.6240509, final_local_scores=-0.69729227,
pem=-4.244057, similarity=6.360965]
1: 4689264 Australia
[pem_pos=0, final_scores=0.53105766,
final_local_scores=0.54085827, pem=-0.17168477,
similarity=21.163652]
2: 26716 South_Australia
[pem_pos=0, final_scores=0.5810675,
final_local_scores=0.5756975, pem=-0.05601515,

similarity=21.574293]
3: 1148 Adelaide
[pem_pos=0, final_scores=0.44860816,
final_local_scores=0.4173044, pem=-0.29570588,
similarity=18.151003]
4: 8758947 Fulham_Gardens,_South_Australia
[pem_pos=0, final_scores=0.10574435,
final_local_scores=0.16780454, pem=0.0, similarity=9.097335]
5: 5662235 Henley_Beach,_South_Australia
[pem_pos=0, final_scores=0.29621178,
final_local_scores=0.32424283, pem=0.0, similarity=13.765766]
6: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.2598521,
final_local_scores=0.2093463, pem=0.0, similarity=10.337024]
7: 18453 Liberal_Party_of_Australia
[pem_pos=1, final_scores=0.19712292,
final_local_scores=0.14744547, pem=-1.6500671,
similarity=17.463102]
8: 3706488 Paul_Caica
[pem_pos=0, final_scores=0.23873793,
final_local_scores=0.2742478, pem=0.0, similarity=12.273814]
9: 1148 Adelaide
[pem_pos=0, final_scores=0.41058183,
final_local_scores=0.3757664, pem=-0.29570588,
similarity=16.911427]
1: 1148 Adelaide
[pem_pos=1, final_scores=0.37382573,
final_local_scores=0.3344202, pem=-0.6931472,
similarity=17.83892]
2: 5864573 Steve_Condous
[pem_pos=0, final_scores=0.102592155,
final_local_scores=0.14397353, pem=0.0, similarity=8.38617]
3: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.22186632,
final_local_scores=0.16785258, pem=-0.40684944,
similarity=11.311279]
4: 18453 Liberal_Party_of_Australia
[pem_pos=2, final_scores=0.042274967,
final_local_scores=-0.031305224, pem=-2.0174363,
similarity=14.126635]
5: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.37160975,
final_local_scores=0.33142453, pem=-0.2258703,
similarity=15.2084]
6: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.34397894,
final_local_scores=0.30124205, pem=-0.2258703,
similarity=14.307694]
7: 1148 Adelaide
[pem_pos=0, final_scores=0.18978243,
final_local_scores=0.13457662, pem=-0.29570588,
similarity=9.713842]

Sample: b'New_Manchester'

New Manchester or [1 The City 1] was an isolated mining village on the [1 1 Manchester Coalfield 1] north of [2 2 Mosley Common 2] in the [3 3 Tyldesley 3] township . It lies west of a boundary stone at [2 1 Ellenbrook 2] which marks the ancient boundary of the s of and , the boundary of 2 Eccles and [4 4 Leigh 4] es , [3 Tyldesley 3] , [5 5 Worsley 5] and [6 6 Little Hulton 6] townships and the metropolitan districts of [4 3 Wigan 4] and [5 4 Salford 5] . The route of the 5 Roman road from [6 Manchester 6] to [7 Wigan 7] and the [7 7 Tyldesley Loopline 7] passed south of the village . The 6 Lancashire and Yorkshire Railway 's [8 Manchester 8] to [9 Southport 9] line passed to the north . Coal was got in New Manchester where the coal seams of the [10 Manchester Coalfield 10] outcrop and are not far below the surface . The community was in existence by 1803 in houses provided by the [11 7 Duke of Bridgewater 11] who owned the mining rights . The men who came to work in the pits named the streets after places in their hometown , [8 8 Manchester 8] , Shude Hill , Gatley Row and [12 City Road 12] , the only road into the village . In the 1830s Shude Hill Pit had a steam winding engine . City Pit and Gatley Pit were in operation by 1838 linked to the [9 9 Worsley Navigable Levels 9] and a horse-drawn tramroad to Mathers Fold . The tramway was extended to link several pits to [10 10 Bridgewater Canal 10] at [11 11 Boothstown Basin 11] . These pits worked the -LRB- coal seam -RRB- at 262 feet and the at 360 feet and employed 64 workers in 1852 . They closed in 1877 . The area has been

1: 29500837 Manchester_Coalfield

2: 9357117 Mosley_Common

3: 3178968 Tyldesley

4: 450898 Leigh,_Greater_Manchester

5: 893196 Worsley

6: 9193680 Little_Hulton

7: 26581473 Tyldesley_Loopline

8: 20206 Manchester

9: 10186299 Worsley_Navigable_Levels

10: 313656 Bridgewater_Canal

11: 963324 Boothstown

1: 9206982 Ellenbrook,_Greater_Manchester

[winner: 3018932 Ellenbrook,_Western_Australia,

pem_pos=1, final_scores=0.11315532,

final_local_scores=0.1712465, pem=-0.30412927,

similarity=10.853952]

gt: 9206982 Ellenbrook,_Greater_Manchester, pem_pos=2,

final_scores=0.023626342, final_local_scores=0.0043902546,

pem=-0.8368875, similarity=8.771852]

2: 1742688 Eccles,_Greater_Manchester

[winner: 1742688 Eccles,_Greater_Manchester, pem_pos=0,

final_scores=-0.1014647, final_local_scores=-0.1857253,

pem=-1.037052, similarity=4.1869497]

gt: 1742688 Eccles,_Greater_Manchester, pem_pos=0,
final_scores=-0.1014647, final_local_scores=-0.1857253,
pem=-1.037052, similarity=4.1869497]
3: 635348 Metropolitan_Borough_of_Wigan
[winner: 8222430 Wigan, pem_pos=0,
final_scores=0.26847082, final_local_scores=0.22205535,
pem=-0.30239522, similarity=12.36076]
gt: 635348 Metropolitan_Borough_of_Wigan, pem_pos=3,
final_scores=-0.2311043, final_local_scores=-0.28926563,
pem=-2.5770032, similarity=9.471594]
4: 635388 City_of_Salford
[winner: 182231 Salford,_Greater_Manchester, pem_pos=0,
final_scores=0.16667284, final_local_scores=0.096484825,
pem=-1.0645502, similarity=12.758202]
gt: 635388 City_of_Salford, pem_pos=1,
final_scores=0.11101203, final_local_scores=0.06395711,
pem=-1.2130575, similarity=12.595117]
5: 13525 Roman_Britain
[winner: 25507 Roman_Empire, pem_pos=1,
final_scores=-0.0895471, final_local_scores=-0.042049304,
pem=-1.3445581, similarity=10.146794]
gt: 13525 Roman_Britain, pem_pos=4,
final_scores=-0.23416127, final_local_scores=-0.24096808,
pem=-3.0039637, similarity=13.234767]
6: 564105 Lancashire_and_Yorkshire_Railway
[winner: 564105 Lancashire_and_Yorkshire_Railway,
pem_pos=0, final_scores=-0.05756016,
final_local_scores=-0.09212561, pem=0.0,
similarity=1.3405011]
gt: 564105 Lancashire_and_Yorkshire_Railway, pem_pos=0,
final_scores=-0.05756016, final_local_scores=-0.09212561,
pem=0.0, similarity=1.3405011]
7: 30293242 Bridgewater_Collieries
[winner: 899190 Francis_Egerton,_3rd_Duke_of_Bridgewater,
pem_pos=0, final_scores=0.05766861,
final_local_scores=0.081434175, pem=0.0, similarity=6.519871]
gt: 30293242 Bridgewater_Collieries, pem_pos=4,
final_scores=-0.9671819, final_local_scores=-1.0694222,
pem=-5.961005, similarity=4.592894]
1: 29500837 Manchester_Coalfield
[pem_pos=0, final_scores=0.060309008,
final_local_scores=0.06697586, pem=0.0, similarity=6.088406]
2: 9357117 Mosley_Common
[pem_pos=0, final_scores=0.23427097,
final_local_scores=0.25745338, pem=0.0, similarity=11.772635]
3: 3178968 Tyldesley
[pem_pos=0, final_scores=0.30093437,
final_local_scores=0.25162947, pem=-0.01939993,
similarity=11.704338]
4: 450898 Leigh,_Greater_Manchester
[pem_pos=0, final_scores=0.114491865,
final_local_scores=0.040984437, pem=-1.1617748,

similarity=11.630682]
5: 893196 Worsley
[pem_pos=0, final_scores=0.12685327,
final_local_scores=0.061628968, pem=-0.2920191,
similarity=7.5168896]
6: 9193680 Little_Hulton
[pem_pos=0, final_scores=0.24853586,
final_local_scores=0.22628674, pem=0.0, similarity=10.842561]
7: 26581473 Tyldesley_Loopline
[pem_pos=0, final_scores=0.22898285,
final_local_scores=0.24029225, pem=0.0, similarity=11.260513]
8: 20206 Manchester
[pem_pos=0, final_scores=0.38149327,
final_local_scores=0.35184067, pem=-0.3209542,
similarity=16.33474]
9: 10186299 Worsley_Navigable_Levels
[pem_pos=0, final_scores=0.22758298,
final_local_scores=0.29287302, pem=0.0, similarity=12.829628]
10: 313656 Bridgewater_Canal
[pem_pos=0, final_scores=0.45898122,
final_local_scores=0.45037633, pem=0.0, similarity=17.52984]
11: 963324 Boothstown
[pem_pos=0, final_scores=0.34588534,
final_local_scores=0.34110385, pem=0.0, similarity=14.268932]
1: 6883 City_of_London
[pem_pos=1, final_scores=0.032281533,
final_local_scores=-0.007858381, pem=-1.1852708,
similarity=10.30089]
2: 3018932 Ellenbrook,_Western_Australia
[pem_pos=1, final_scores=0.11315532,
final_local_scores=0.1712465, pem=-0.30412927,
similarity=10.853952]
3: 3178968 Tyldesley
[pem_pos=0, final_scores=0.25885236,
final_local_scores=0.20566124, pem=-0.01939993,
similarity=10.332554]
4: 8222430 Wigan
[pem_pos=0, final_scores=0.26847082,
final_local_scores=0.22205535, pem=-0.30239522,
similarity=12.36076]
5: 182231 Salford,_Greater_Manchester
[pem_pos=0, final_scores=0.16667284,
final_local_scores=0.096484825, pem=-1.0645502,
similarity=12.758202]
6: 20206 Manchester
[pem_pos=0, final_scores=0.5120433,
final_local_scores=0.4944467, pem=-0.3209542,
similarity=20.590387]
7: 8222430 Wigan
[pem_pos=0, final_scores=0.23828276,
final_local_scores=0.1890795, pem=-0.30239522,
similarity=11.376695]

8: 20206 Manchester
 [pem_pos=0, final_scores=0.3723436,
 final_local_scores=0.341846, pem=-0.3209542,
 similarity=16.036478]
 9: 310742 Southport
 [pem_pos=0, final_scores=0.316989,
 final_local_scores=0.2822621, pem=-0.29562387,
 similarity=14.1206255]
 10: 29500837 Manchester_Coalfield
 [pem_pos=0, final_scores=0.19484,
 final_local_scores=0.21393055, pem=0.0, similarity=10.473827]
 11: 899190 Francis_Egerton,_3rd_Duke_of_Bridgewater
 [pem_pos=0, final_scores=0.05766861,
 final_local_scores=0.081434175, pem=0.0, similarity=6.519871]
 12: 234208 City_Road
 [pem_pos=0, final_scores=0.3764106,
 final_local_scores=0.42060083, pem=-0.19303977,
 similarity=17.691061]

Sample: b'Frank_S'

___ Pepper Frank Stuart Pepper -LRB- 1910 -- 1988 -RRB- was a [1 British 1] writer of 1 comics and 2 story paper s for [2 3 Amalgamated Press 2] , best known as the creator of '' 4 Roy o f the Rovers '' and '' '' . Born in [1 1 Ilford 1] , , on 8 February 1910 , on leaving school he joined the staff of 5 The Children 's Newspaper as an office boy in 1926 , under editor 6 Arthur Mee . He started by writing captions , then longer pieces , and by 1930 was selling articles to other papers on a freelance basis . He went freelance full-time in 1931 , writing for numerous newspapers and magazines . In the mid-1930s he started selling stories to the boys ' story papers , his first story being `` Snapshot Sammy '' for The Triumph . Editor 7 Reg Eves commissioned him to write a series about a boxing airman , `` Rockfist Rogan '' , for 8 Th e Champion in 1937 , which he went on to write , under the pseudonym Hal Wilton , for the next 22 years . Also for The Champion , he wrote football serial `` Danny of the Dazzlers '' under the pseudonym John Marshall , and `` Colwyn Dane '' as Mark Grimshaw . He also wrote `` The Adventures of [3 Beau Brummell 3] '' for 9 Knockout and `` The Return of Monte Cristo '' for , as John Morion . For The Children 's Newspaper , he wrote the adventures of twins `` Bill and Jill '' , starting in 1948 . In 1952 AP launched 10 Lion , a weekly adventure comic designed to compete with ' 11 Eagle , for which Pepper wrote the cover feature '' '' , a space hero created to rival Eagles '' [2 2 Dan Dare 2] '' , for twelve years . He also created the cover feature for [4 AP 4] 's sports comic Tiger , in 1953 : editor [3 3 Derek Birnage 3] requested a more realistic football series than `` Danny of the Dazzlers '' , and Pepper gave him '' 12 Roy of the Rovers '' , illustrated by [4 4 Joe Colquhoun 4] , who took over writing the strip , from Pepper 's outline , after four

episodes , using the pseudonym [5 Stewart Colwyn 5] . Other series he wrote scripts for included '' [5 5 Dan Dare 5] '' for 13 Eagle , '' 14 Jet-Ace Logan '' in and Tiger and `` The Spellbinder '' for 15 Lion . He retired from comics in 1983 , and started compiling collections of quotations , including Twentieth Century Quotations , Contemporary Biographical Quotations , Dictionary of Biographical Quotations and Twentieth Century Anecdotes . He died in [6 6 Cornwall 6]

1: 81476 Ilford

2: 169375 Dan_Dare

3: 32422459 Derek_Birnage

4: 1324405 Joe_Colquhoun

5: 169375 Dan_Dare

6: 5648 Cornwall

1: 491987 British_comics

[winner: 3175886 I._W._Publications, pem_pos=13, final_scores=-0.99802446, final_local_scores=-0.95767504, pem=-2.5649493, similarity=-10.5406275|

RECALL MISS: 491987 British_comics]

2: 3324250 Story_paper

[winner: 21730528 Electrical_insulation_paper, pem_pos=1, final_scores=-1.1659392, final_local_scores=-1.1088793, pem=-1.5966074, similarity=-20.318861|

RECALL MISS: 3324250 Story_paper]

[winner: 4799434 Two_Gallants_(short_story), pem_pos=6, final_scores=-1.1702477, final_local_scores=-1.1210442, pem=-2.3978953, similarity=-16.324352|

RECALL MISS: 3324250 Story_paper]

3: 2016739 Amalgamated_Press

[winner: 3591129 Fleetway_Publications, pem_pos=0, final_scores=0.15052266, final_local_scores=0.14976588, pem=-0.26495385, similarity=9.999886|

gt: 2016739 Amalgamated_Press, pem_pos=1, final_scores=-0.076775804, final_local_scores=-0.086339355, pem=-1.4577533, similarity=9.440663]

4: 999875 Roy_of_the_Rovers

[winner: 999875 Roy_of_the_Rovers, pem_pos=0, final_scores=-0.071089074, final_local_scores=-0.10907562, pem=0.0, similarity=0.83467877|

gt: 999875 Roy_of_the_Rovers, pem_pos=0, final_scores=-0.071089074, final_local_scores=-0.10907562, pem=0.0, similarity=0.83467877]

5: 5636361 The_Children's_Newspaper

[winner: 20259032 The_Children_(2008_film), pem_pos=0, final_scores=-0.22679147, final_local_scores=-0.12600547, pem=-0.59485793, similarity=3.5643883|

RECALL MISS: 5636361 The_Children's_Newspaper]

[winner: 1168956 Student_newspaper, pem_pos=1, final_scores=-0.3138482, final_local_scores=-0.31832266, pem=-0.53101766, similarity=-2.5219135|

RECALL MISS: 5636361 The_Children's_Newspaper]

[winner: 1415027 David_Berman_(musician), pem_pos=5,

final_scores=-2.6166987, final_local_scores=-2.7988358,
pem=-3.0012724, similarity=-63.111744|
RECALL MISS: 5636361 The_Children's_Newspaper]

6: 1371872 Arthur_Mee
[winner: 1371872 Arthur_Mee, pem_pos=0,
final_scores=0.019639298, final_local_scores=0.058303595,
pem=0.0, similarity=5.8296084|
gt: 1371872 Arthur_Mee, pem_pos=0,
final_scores=0.019639298, final_local_scores=0.058303595,
pem=0.0, similarity=5.8296084]

7: 33003744 Reg_Eves
[winner: 33003744 Reg_Eves, pem_pos=0,
final_scores=-0.054418445, final_local_scores=0.0060275197,
pem=0.0, similarity=4.269586|
gt: 33003744 Reg_Eves, pem_pos=0,
final_scores=-0.054418445, final_local_scores=0.0060275197,
pem=0.0, similarity=4.269586]

8: 9675253 The_Champion_(story_paper)
[winner: 4354190 The_Champion_(TV_series), pem_pos=1,
final_scores=-0.10786244, final_local_scores=0.008853197,
pem=-0.31593925, similarity=6.0720367|
gt: 9675253 The_Champion_(story_paper), pem_pos=3,
final_scores=-0.37602043, final_local_scores=-0.32599792,
pem=-2.757596, similarity=9.35752]

9: 653527 Knockout_(UK_comics)
[winner: 17163 Knockout, pem_pos=0,
final_scores=-0.105065405, final_local_scores=-0.025336161,
pem=-0.7417826, similarity=7.3675604|
gt: 653527 Knockout_(UK_comics), pem_pos=17,
final_scores=-0.6915746, final_local_scores=-0.7293039,
pem=-4.202116, similarity=5.177589]

10: 3452825 Lion_(comics)
[winner: 36896 Lion, pem_pos=0, final_scores=-0.22078377,
final_local_scores=-0.16781989, pem=-0.44021446,
similarity=1.4755874|
gt: 3452825 Lion_(comics), pem_pos=1,
final_scores=-0.3632409, final_local_scores=-0.4036408,
pem=-2.3300364, similarity=4.715367]

11: 634318 Eagle_(comic)
[winner: 634318 Eagle_(comic), pem_pos=1,
final_scores=-0.28917682, final_local_scores=-0.35967752,
pem=-2.6791112, similarity=7.9256415|
gt: 634318 Eagle_(comic), pem_pos=1,
final_scores=-0.28917682, final_local_scores=-0.35967752,
pem=-2.6791112, similarity=7.9256415]

12: 999875 Roy_of_the_Rovers
[winner: 999875 Roy_of_the_Rovers, pem_pos=0,
final_scores=-0.066647574, final_local_scores=-0.10422395,
pem=0.0, similarity=0.97946227|
gt: 999875 Roy_of_the_Rovers, pem_pos=0,
final_scores=-0.066647574, final_local_scores=-0.10422395,
pem=0.0, similarity=0.97946227]

```
13: 634318 Eagle_(comic)
    [winner: 634318 Eagle_(comic), pem_pos=1,
    final_scores=-0.3390842, final_local_scores=-0.41419375,
    pem=-2.6791112, similarity=6.298768]
    gt: 634318 Eagle_(comic), pem_pos=1,
    final_scores=-0.3390842, final_local_scores=-0.41419375,
    pem=-2.6791112, similarity=6.298768 ]
14: 3653767 Jet-Ace_Logan
    [winner: 3653767 Jet-Ace_Logan, pem_pos=0,
    final_scores=-0.0459895, final_local_scores=0.04864824,
    pem=0.0, similarity=5.5414734]
    gt: 3653767 Jet-Ace_Logan, pem_pos=0,
    final_scores=-0.0459895, final_local_scores=0.04864824,
    pem=0.0, similarity=5.5414734 ]
15: 3452825 Lion_(comics)
    [winner: 3452825 Lion_(comics), pem_pos=1,
    final_scores=-0.30887967, final_local_scores=-0.34425938,
    pem=-2.3300364, similarity=6.4874263]
    gt: 3452825 Lion_(comics), pem_pos=1,
    final_scores=-0.30887967, final_local_scores=-0.34425938,
    pem=-2.3300364, similarity=6.4874263 ]
1: 81476 Ilford
    [pem_pos=0, final_scores=0.1718419,
    final_local_scores=0.2012074, pem=-0.2926246,
    similarity=11.685481]
2: 169375 Dan_Dare
    [pem_pos=0, final_scores=0.21650778,
    final_local_scores=0.20710674, pem=0.0, similarity=10.270191]
3: 32422459 Derek_Birnage
    [pem_pos=0, final_scores=0.24569054,
    final_local_scores=0.29074615, pem=0.0, similarity=12.766158]
4: 1324405 Joe_Colquhoun
    [pem_pos=0, final_scores=0.13177894,
    final_local_scores=0.14938316, pem=0.0, similarity=8.547605]
5: 169375 Dan_Dare
    [pem_pos=0, final_scores=0.23873721,
    final_local_scores=0.23138902, pem=0.0, similarity=10.9948225]
6: 5648 Cornwall
    [pem_pos=0, final_scores=0.4120522,
    final_local_scores=0.47255772, pem=-0.2719962,
    similarity=19.670937]
1: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4533556,
    final_local_scores=0.49029583, pem=-0.48936236,
    similarity=21.382347]
2: 3591129 Fleetway_Publications
    [pem_pos=0, final_scores=0.15052266,
    final_local_scores=0.14976588, pem=-0.26495385,
    similarity=9.999886]
3: 194308 Beau_Brummell
    [pem_pos=0, final_scores=0.10155006,
    final_local_scores=0.13500988, pem=0.0, similarity=8.118677]
```

4: 18935732 Associated_Press
[pem_pos=1, final_scores=0.10181989,
final_local_scores=0.12591463, pem=-0.68645424,
similarity=11.580302]
5: 1324405 Joe_Colquhoun
[pem_pos=0, final_scores=0.15848683,
final_local_scores=0.17855749, pem=0.0, similarity=9.418224]

Sample: b'Aaron_Thomas'

_ -LRB- cricketer -RRB- Aaron Courteney Thomas -LRB- born 6
May 1985 -RRB- is a former [1 1 English 1] . Thomas is a
right-handed who plays primarily as a . He was born at [2 2 Ed
monton 2] , [1 1 London 1] . Thomas made his debut in for the
[2 2 Nottinghamshire Cricket Board 2] against [3 3 Bedfordshi
re 3] in the 3 2001 [3 Cheltenham & Gloucester Trophy 3] at ,
[4 4 Luton 4] , with the Board losing by 3 wickets . His
second and final [4 List-A 4] appearance for the [5 Board 5]
came in 1st round of the 4 2002 [6 Cheltenham & Gloucester Tr
ophy 6] against [5 5 Oxfordshire 5] at the [6 6 Christ Church
Ground 6] , [7 7 Oxford 7] . This round of the competition
was played in 2001 , with the Board losing by 5 wickets . In
his 2 [7 List-A 7] matches for the Board he batted in a
single innings , scoring an 5 unbeaten 7 . Behind the he took
a single and made a single . Thomas made a single 6 first-
class appearance for [8 8 Nottinghamshire 8] against [8 7 Ind
ia 8] in 2003 . Not required to bat in the match , Thomas
took a single catch behind the stumps . In local cricket he
currently represents the in the 8 Nottinghamshire Cricket Boa
rd Premier League

1: 17867 London
2: 27793154 Nottinghamshire_Cricket_Board
3: 3691144 Bedfordshire_County_Cricket_Club
4: 55639 Luton
5: 3712398 Oxfordshire_County_Cricket_Club
6: 28400320 Christ_Church_Ground
7: 22308 Oxford
8: 667341 Nottinghamshire_County_Cricket_Club

1: 18803164 English_people
[winner: 9316 England, pem_pos=1,
final_scores=0.42995948, final_local_scores=0.43587536,
pem=-1.3527626, similarity=24.453636]
gt: 18803164 English_people, pem_pos=2,
final_scores=-0.40736353, final_local_scores=-0.44960546,
pem=-2.2084396, similarity=2.682427]
2: 2140991 Edmonton,_London
[winner: 95405 Edmonton, pem_pos=0,
final_scores=0.20298897, final_local_scores=0.24908751,
pem=-0.27044967, similarity=12.993729]
gt: 2140991 Edmonton,_London, pem_pos=1,
final_scores=-0.2498058, final_local_scores=-0.27755785,
pem=-2.8274174, similarity=11.182769]
3: 27734795 2001_Cheltenham_&_Gloucester_Trophy

[winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.39863265, final_local_scores=0.3633592,
 pem=0.0, similarity=14.933076|
 gt: 27734795 2001_Cheltenham_&_Gloucester_Trophy,
 pem_pos=1, final_scores=0.05804901,
 final_local_scores=0.026702374, pem=-1.537979,
 similarity=13.250334]
 [winner: 34551 2001, pem_pos=0, final_scores=-0.6452723,
 final_local_scores=-0.42221448, pem=-0.5613713,
 similarity=-5.457185|
 RECALL MISS: 27734795
 2001_Cheltenham_&_Gloucester_Trophy]
 4: 27764855 2002_Cheltenham_&_Gloucester_Trophy
 [winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
 final_scores=0.29410416, final_local_scores=0.24490085,
 pem=0.0, similarity=11.398043|
 gt: 27764855 2002_Cheltenham_&_Gloucester_Trophy,
 pem_pos=3, final_scores=-0.36755586,
 final_local_scores=-0.43947676, pem=-3.295837,
 similarity=8.898123]
 [winner: 35502 2002, pem_pos=0, final_scores=-1.0958078,
 final_local_scores=-1.0054858, pem=-0.5897454,
 similarity=-22.708862|
 RECALL MISS: 27764855
 2002_Cheltenham_&_Gloucester_Trophy]
 5: 2430357 Not_out
 [winner: 2430357 Not_out, pem_pos=0,
 final_scores=-1.3465161, final_local_scores=-1.5252733,
 pem=-0.53147423, similarity=-38.537243|
 gt: 2430357 Not_out, pem_pos=0, final_scores=-1.3465161,
 final_local_scores=-1.5252733, pem=-0.53147423,
 similarity=-38.537243]
 6: 11040 First-class_cricket
 [winner: 662844 First-class_citizen, pem_pos=3,
 final_scores=-1.3332462, final_local_scores=-1.266175,
 pem=-1.5935546, similarity=-25.029478|
 gt: 11040 First-class_cricket, pem_pos=0,
 final_scores=-1.3421742, final_local_scores=-1.5348926,
 pem=-0.053819276, similarity=-41.421867]
 7: 23602828 India_A_cricket_team
 [winner: 407754 India_national_cricket_team, pem_pos=1,
 final_scores=0.4220047, final_local_scores=0.42010057,
 pem=-3.5152702, similarity=35.74294|
 RECALL MISS: 23602828 India_A_cricket_team]
 8: 19619537 Nottinghamshire_Cricket_Board_Premier_League
 [winner: 19619537
 Nottinghamshire_Cricket_Board_Premier_League, pem_pos=0,
 final_scores=-0.06544836, final_local_scores=-0.008340657,
 pem=0.0, similarity=3.8408113|
 gt: 19619537
 Nottinghamshire_Cricket_Board_Premier_League, pem_pos=0,
 final_scores=-0.06544836, final_local_scores=-0.008340657,

pem=0.0, similarity=3.8408113]

1: 17867 London

[pem_pos=0, final_scores=0.33551198,
final_local_scores=0.35736474, pem=-0.09086737,
similarity=15.248341]

2: 27793154 Nottinghamshire_Cricket_Board

[pem_pos=0, final_scores=0.27076113,
final_local_scores=0.27887127, pem=0.0, similarity=12.411788]

3: 3691144 Bedfordshire_County_Cricket_Club

[pem_pos=1, final_scores=0.44032878,
final_local_scores=0.42086434, pem=-1.6135224,
similarity=25.42373]

4: 55639 Luton

[pem_pos=0, final_scores=0.28166527,
final_local_scores=0.2800646, pem=-0.035978474,
similarity=12.643056]

5: 3712398 Oxfordshire_County_Cricket_Club

[pem_pos=1, final_scores=0.5135608,
final_local_scores=0.4974789, pem=-1.7495553,
similarity=28.449831]

6: 28400320 Christ_Church_Ground

[pem_pos=0, final_scores=0.29410362,
final_local_scores=0.3051879, pem=0.0, similarity=13.197129]

7: 22308 Oxford

[pem_pos=0, final_scores=0.11398463,
final_local_scores=0.11919546, pem=-0.68225896,
similarity=11.356973]

8: 667341 Nottinghamshire_County_Cricket_Club

[pem_pos=1, final_scores=0.60871655,
final_local_scores=0.57873994, pem=-1.1067253,
similarity=27.379011]

1: 9316 England

[pem_pos=1, final_scores=0.42995948,
final_local_scores=0.43587536, pem=-1.3527626,
similarity=24.453636]

2: 95405 Edmonton

[pem_pos=0, final_scores=0.20298897,
final_local_scores=0.24908751, pem=-0.27044967,
similarity=12.993729]

3: 1899702 Friends_Provident_Trophy

[pem_pos=0, final_scores=0.39863265,
final_local_scores=0.3633592, pem=0.0, similarity=14.933076]

4: 683459 List_A_cricket

[pem_pos=0, final_scores=0.22981529,
final_local_scores=0.17801872, pem=-0.07478003,
similarity=9.808812]

5: 24669195 Hampshire_Cricket_Board

[pem_pos=9, final_scores=0.0419202,
final_local_scores=0.015811682, pem=-3.583519,
similarity=24.049297]

6: 1899702 Friends_Provident_Trophy

[pem_pos=0, final_scores=0.29410416,


```

final_local_scores=0.24490085, pem=0.0, similarity=11.398043]
7: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.08351986,
final_local_scores=0.018213168, pem=-0.07478003,
similarity=5.0398955]
8: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.4220047,
final_local_scores=0.42010057, pem=-3.5152702,
similarity=35.74294]

```

Sample: b'Assembly_of_European_Wine-producing_Regions'

The Assembly of European Wine-producing Regions or Assemblée des Régions Européennes Viticoles -LRB- AREV -RRB- is an organisation of political and trade representatives of wine regions within the [1 1 European Union 1] -LRB- EU -RRB- and [1 Eastern Europe 1] . It lobbies the EU on matters affecting its members , and acts as a forum for discussion of policy and marketing . The European Wine Regions Conference -LRB- CERV -RRB- was created on 20 June 1988 under [2 Alsatian 2] law . It became the Assembly of European Wine Regions -LRB- AREV -RRB- in 1994 . It is financed by membership subscriptions , and is based in [2 2 Bordeaux 2] with an office in [3 3 Strasbourg 3] . The industry representatives are grouped together in the European Wine Trade Council -LRB- CEPV -RRB- , which drafts documents to be approved by the plenary council . [3 Alentejo 3] , [4 1 Alsace 4] , [5 2 Aquitaine 5] , [6 Arad 6] , [7 Aragon 7] , [8 Bacs-Kiskun 8] , [4 4 Baden-Württemberg 4] , Baranya , [5 5 Bavaria 5] , Borsod-Abauj-Zemplen , [9 3 Burgundy 9] , [10 Burgenland 10] , Castilla-La-Mancha , [11 Castilla y Leon 11] , , Centre , [12 Champagne-Ardenne 12] , [13 Constanta 13] , [14 Extremadura 14] , [15 Euskadi 15] , Franche-Comté , , , [6 6 Hesse 6] , , , , [16 4 Lombardia 16] , [17 Macedonia 17] , , [18 Madrid 18] , [19 Midi-Pyrenees 19] , [20 Murcia 20] , [21 Navarra 21] , [22 Niederösterreich 22] , [23 Norte 23] , , [24 Pays de Loire 24] , [25 Piedmont 25] , Provence-Alpes-Côte d'Azur , [7 7 Rhineland-Palatinate 7] , [26 Rhône-Alpes 26] , [27 Sicily 27] , , [28 Steiermark 28] , The , Tolna , [29 Trento 29] , [30 Valais 30] , [31 Valencia 31] , ,

```

1: 9317 European_Union
2: 4097 Bordeaux
3: 37407 Strasbourg
4: 66401 Baden-Württemberg
5: 3764 Bavaria
6: 38412 Hesse
7: 26239 Rhineland-Palatinate
1: 1083426 Alsace_wine

```

```

    [winner: 48129 Alsace, pem_pos=0,
final_scores=0.27414346, final_local_scores=0.23184949,
pem=0.0, similarity=11.008564]
    gt: 1083426 Alsace_wine, pem_pos=1,
final_scores=-0.11336757, final_local_scores=-0.08778295,

```

pem=-1.5746224, similarity=10.033135]
2: 2166602 Bordeaux_wine
[winner: 46378 Aquitaine, pem_pos=0,
final_scores=0.20151715, final_local_scores=0.17156157,
pem=0.0, similarity=9.209453]
gt: 2166602 Bordeaux_wine, pem_pos=4,
final_scores=-0.17251311, final_local_scores=-0.17071342,
pem=-1.608944, similarity=7.7449718]
3: 715978 Burgundy_wine
[winner: 85451 Burgundy, pem_pos=0,
final_scores=0.21832113, final_local_scores=0.17524752,
pem=-0.44129616, similarity=11.719285]
gt: 715978 Burgundy_wine, pem_pos=2,
final_scores=-0.25933394, final_local_scores=-0.25546655,
pem=-1.7732974, similarity=6.109549]
4: 14356821 Lombardia_(wine)
[winner: 43807 Lombardy, pem_pos=0,
final_scores=0.30266744, final_local_scores=0.26423123,
pem=0.0, similarity=11.9749]
gt: 14356821 Lombardia_(wine), pem_pos=1,
final_scores=-0.21850353, final_local_scores=-0.16541582,
pem=-0.46607432, similarity=1.6879593]
1: 9317 European_Union
[pem_pos=0, final_scores=0.3982761,
final_local_scores=0.4573633, pem=-0.005814796,
similarity=17.769968]
2: 4097 Bordeaux
[pem_pos=0, final_scores=0.3898391,
final_local_scores=0.37414092, pem=-0.40822297,
similarity=17.474804]
3: 37407 Strasbourg
[pem_pos=0, final_scores=0.39272505,
final_local_scores=0.37293738, pem=-0.16281788,
similarity=16.104338]
4: 66401 Baden-Württemberg
[pem_pos=0, final_scores=0.42247826,
final_local_scores=0.41646582, pem=-0.009141438,
similarity=16.567595]
5: 3764 Bavaria
[pem_pos=0, final_scores=0.40339327,
final_local_scores=0.3674304, pem=-0.019504828,
similarity=15.160639]
6: 38412 Hesse
[pem_pos=0, final_scores=0.18272828,
final_local_scores=0.15795892, pem=-0.1019326,
similarity=9.357847]
7: 26239 Rhineland-Palatinate
[pem_pos=0, final_scores=0.109990284,
final_local_scores=0.09424387, pem=0.0, similarity=6.9021378]
1: 37403 Eastern_Europe
[pem_pos=0, final_scores=0.20353012,
final_local_scores=0.21382654, pem=0.0, similarity=10.470723]

2: 48129 Alsace
[pem_pos=0, final_scores=0.17138897,
final_local_scores=0.11960587, pem=-0.22951354,
similarity=8.907122]
3: 1185527 Alentejo
[pem_pos=0, final_scores=0.18388294,
final_local_scores=0.17651904, pem=0.0, similarity=9.357393]
4: 48129 Alsace
[pem_pos=0, final_scores=0.27414346,
final_local_scores=0.23184949, pem=0.0, similarity=11.008564]
5: 46378 Aquitaine
[pem_pos=0, final_scores=0.20151715,
final_local_scores=0.17156157, pem=0.0, similarity=9.209453]
6: 22537901 Arad, Romania
[pem_pos=0, final_scores=0.23511182,
final_local_scores=0.22270775, pem=-0.6827664,
similarity=14.448746]
7: 39443 Aragon
[pem_pos=0, final_scores=0.3196174,
final_local_scores=0.29071355, pem=-0.16680257,
similarity=13.672283]
8: 993023 Bács-Kiskun County
[pem_pos=0, final_scores=0.18465199,
final_local_scores=0.24977854, pem=0.0, similarity=11.543603]
9: 85451 Burgundy
[pem_pos=0, final_scores=0.21832113,
final_local_scores=0.17524752, pem=-0.44129616,
similarity=11.719285]
10: 98935 Burgenland
[pem_pos=0, final_scores=0.16450436,
final_local_scores=0.14303795, pem=0.0, similarity=8.358251]
11: 70260 Castile and León
[pem_pos=0, final_scores=0.26681405,
final_local_scores=0.27636543, pem=0.0, similarity=12.337008]
12: 90435 Champagne-Ardenne
[pem_pos=0, final_scores=0.07427396,
final_local_scores=0.07241304, pem=0.0, similarity=6.2506623]
13: 265716 Constanța
[pem_pos=0, final_scores=0.28957063,
final_local_scores=0.30948493, pem=0.0, similarity=13.32536]
14: 68204 Extremadura
[pem_pos=0, final_scores=0.195397,
final_local_scores=0.18681684, pem=0.0, similarity=9.6647005]
15: 20514897 Basque Country_(autonomous_community)
[pem_pos=0, final_scores=0.1758445,
final_local_scores=0.16274354, pem=0.0, similarity=8.946304]
16: 43807 Lombardy
[pem_pos=0, final_scores=0.30266744,
final_local_scores=0.26423123, pem=0.0, similarity=11.9749]
17: 23564616 Republic_of_Macedonia
[pem_pos=0, final_scores=0.39196688,
final_local_scores=0.4051146, pem=-1.1139666,

similarity=22.237064]
18: 41188263 Madrid
[pem_pos=0, final_scores=0.46338463,
final_local_scores=0.47066408, pem=-0.13236846,
similarity=18.855108]
19: 198929 Midi-Pyrénées
[pem_pos=0, final_scores=0.035237834,
final_local_scores=0.034501433, pem=0.0, similarity=5.1193047]
20: 59727 Murcia
[pem_pos=0, final_scores=0.23068385,
final_local_scores=0.2113432, pem=-0.24673581,
similarity=11.738403]
21: 59731 Navarre
[pem_pos=0, final_scores=0.110080734,
final_local_scores=0.07598701, pem=-0.3183116,
similarity=8.088345]
22: 148988 Lower Austria
[pem_pos=0, final_scores=0.18504329,
final_local_scores=0.1601704, pem=0.0, similarity=8.869517]
23: 2528128 Norte Region, Portugal
[pem_pos=0, final_scores=0.057225212,
final_local_scores=0.052388266, pem=0.0, similarity=5.6530833]
24: 198748 Pays_de_la_Loire
[pem_pos=0, final_scores=0.20885591,
final_local_scores=0.21830499, pem=0.0, similarity=10.604369]
25: 25061 Piedmont
[pem_pos=0, final_scores=0.21345724,
final_local_scores=0.16150397, pem=-0.6027155,
similarity=12.186975]
26: 198828 Rhône-Alpes
[pem_pos=0, final_scores=0.1842569,
final_local_scores=0.16544998, pem=-0.009191365,
similarity=9.077054]
27: 27619 Sicily
[pem_pos=0, final_scores=0.45924342,
final_local_scores=0.48096597, pem=-0.08753506,
similarity=18.918726]
28: 224474 Styria
[pem_pos=0, final_scores=0.2224515,
final_local_scores=0.18699265, pem=0.0, similarity=9.669947]
29: 45882 Trento
[pem_pos=0, final_scores=0.23202659,
final_local_scores=0.20256624, pem=-0.13802086,
similarity=10.885272]
30: 198789 Valais
[pem_pos=0, final_scores=0.22732593,
final_local_scores=0.2085019, pem=0.0, similarity=10.311826]
31: 63861 Valencia
[pem_pos=0, final_scores=0.17079072,
final_local_scores=0.14747104, pem=-0.9876697,
similarity=13.861644]

Sample: b'Bering_Sea_Squadron'

The [1 Bering Sea Squadron 1] was a combined [1 1 American 1] and [2 2 British 2] naval station formed and disbanded in 1891 to suppress 1 seal 2 poaching in the [3 3 Bering Sea 3] . Eleven vessels were assigned to the squadron including eight [4 4 United States Navy 4] and 3 Revenue Cutter Service ships plus three from the [5 5 Royal Navy 5] . The [2 United States 2] and [3 United Kingdom 3] government banned 4 sealing in the [4 Bering Sea 4] in 1891 due to the near extinction of the animals so that year 5 Commander of the [5 American 5] navy was ordered to lead a force of two 6 sloops , 7 USS Mohican and 8 USS Marion , two 9 steamers , 10 USS Thetis and , plus the 11 gunboat s 12 USS Alert and 13 USS Ranger to 14 operate against poachers . The s 15 USRC Rush and were also placed under Cotton 's command . [6 British 6] forces sent the 16 cruiser 17 HMS Porpoise , the screw sloop 18 HMS Nymph and the steamer 19 HMS Pheasant . From the beginning of the operation on June 22 , 1891 to its end on October 5 , [7 American 7] and [8 British 8] marine s boarded and expelled dozens of merchant vessels mainly around the [6 6 Aleutian Islands 6] though only four 20 schooner s were found to be hunting seals illegally . Two of these vessels were taken over by marines and towed over 1,200 miles by the Al-Ki to , the other two were taken into [9 British 9] custody and sent to [7 7 Asia 7]

1: 3434750 United_States
2: 31717 United_Kingdom
3: 58055 Bering_Sea
4: 20518076 United_States_Navy
5: 26061 Royal_Navy
6: 26296222 Aleutian_Islands
7: 689 Asia

1: 60261 Pinniped

[winner: 2809136 Scaraboid_seal, pem_pos=16,
final_scores=-1.1729518, final_local_scores=-1.1158987,
pem=-2.944439, similarity=-13.198615]

RECALL MISS: 60261 Pinniped]

2: 52989 Poaching

[winner: 52988 Poaching_(cooking), pem_pos=1,
final_scores=-0.9674808, final_local_scores=-0.8930544,
pem=-0.54507697, similarity=-19.5966]

gt: 52989 Poaching, pem_pos=0, final_scores=-1.1022552,
final_local_scores=-1.1599212, pem=-0.08606558,
similarity=-30.05661]

3: 426234 United_States_Revenue_Cutter_Service

[winner: 426234 United_States_Revenue_Cutter_Service,
pem_pos=0, final_scores=-7.7441335e-05,
final_local_scores=0.011468694, pem=0.0,
similarity=4.4319615]

gt: 426234 United_States_Revenue_Cutter_Service,
pem_pos=0, final_scores=-7.7441335e-05,
final_local_scores=0.011468694, pem=0.0,

```
similarity=4.4319615 ]
4: 1844757 Seal_hunting
  [winner: 382516 Sealing_(Mormonism), pem_pos=2,
  final_scores=-1.3252817, final_local_scores=-1.3285406,
  pem=-1.0815595, similarity=-29.6749]
  gt: 1844757 Seal_hunting, pem_pos=0,
  final_scores=-1.334332, final_local_scores=-1.4713635,
  pem=0.0, similarity=-39.818707 ]
5: 226168 Commander
  [winner: 226168 Commander, pem_pos=0,
  final_scores=-0.69750136, final_local_scores=-0.77100074,
  pem=-0.38505897, similarity=-16.824476]
  gt: 226168 Commander, pem_pos=0,
  final_scores=-0.69750136, final_local_scores=-0.77100074,
  pem=-0.38505897, similarity=-16.824476 ]
6: 318237 Sloop-of-war
  [winner: 1220125 Black_Swan-class_sloop, pem_pos=2,
  final_scores=-1.0911775, final_local_scores=-1.1141684,
  pem=-0.67231315, similarity=-25.503149]
  gt: 318237 Sloop-of-war, pem_pos=0,
  final_scores=-1.1741623, final_local_scores=-1.2826781,
  pem=-0.18918715, similarity=-33.15913 ]
7: 3588714 USS_Mohican_(1883)
  [winner: 3588714 USS_Mohican_(1883), pem_pos=0,
  final_scores=-0.0022540838, final_local_scores=0.06397806,
  pem=0.0, similarity=5.9989457]
  gt: 3588714 USS_Mohican_(1883), pem_pos=0,
  final_scores=-0.0022540838, final_local_scores=0.06397806,
  pem=0.0, similarity=5.9989457 ]
8: 1168055 USS_Marion_(1839)
  [winner: 1168055 USS_Marion_(1839), pem_pos=0,
  final_scores=-0.08106233, final_local_scores=-0.034045056,
  pem=0.0, similarity=3.0737405]
  gt: 1168055 USS_Marion_(1839), pem_pos=0,
  final_scores=-0.08106233, final_local_scores=-0.034045056,
  pem=0.0, similarity=3.0737405 ]
9: 146635 Steamship
  [winner: 178690 Steamboat, pem_pos=0,
  final_scores=-1.2994975, final_local_scores=-1.4501402,
  pem=-0.27749527, similarity=-37.6763]
  gt: 146635 Steamship, pem_pos=3, final_scores=-1.692014,
  final_local_scores=-1.8692622, pem=-2.8469594,
  similarity=-36.210594 ]
10: 17626133 USS_Thetis_(1881)
  [winner: 17626133 USS_Thetis_(1881), pem_pos=0,
  final_scores=-0.0013199747, final_local_scores=0.04994379,
  pem=0.0, similarity=5.5801353]
  gt: 17626133 USS_Thetis_(1881), pem_pos=0,
  final_scores=-0.0013199747, final_local_scores=0.04994379,
  pem=0.0, similarity=5.5801353 ]
11: 274471 Gunboat
  [winner: 274471 Gunboat, pem_pos=0,
```

final_scores=-1.2037268, final_local_scores=-1.3229485,
 pem=-0.005174651, similarity=-35.36157|
 gt: 274471 Gunboat, pem_pos=0, final_scores=-1.2037268,
 final_local_scores=-1.3229485, pem=-0.005174651,
 similarity=-35.36157]
 12: 1467278 USS_Alert_(AS-4)
 [winner: 1467278 USS_Alert_(AS-4), pem_pos=0,
 final_scores=-0.056865856, final_local_scores=-0.0455539,
 pem=-0.44802472, similarity=5.166722|
 gt: 1467278 USS_Alert_(AS-4), pem_pos=0,
 final_scores=-0.056865856, final_local_scores=-0.0455539,
 pem=-0.44802472, similarity=5.166722]
 13: 2638415 USS_Nantucket_(IX-18)
 [winner: 402215 USS_Ranger_(CV-61), pem_pos=0,
 final_scores=-0.051780283, final_local_scores=-0.027865224,
 pem=-0.4830573, similarity=5.8851|
 gt: 2638415 USS_Nantucket_(IX-18), pem_pos=3,
 final_scores=-0.28613305, final_local_scores=-0.2935009,
 pem=-1.8761932, similarity=5.534091]
 14: 31730337 Bering_Sea_Anti-Poaching_Operations
 [winner: 14943085 Operate, pem_pos=0,
 final_scores=-1.1596768, final_local_scores=-1.1174306,
 pem=0.0, similarity=-29.256643|
 RECALL MISS: 31730337 Bering_Sea_Anti-
 Poaching_Operations]
 15: 25567654 USRC_Rush_(1885)
 [winner: 25567654 USRC_Rush_(1885), pem_pos=1,
 final_scores=-0.13754989, final_local_scores=-0.06425427,
 pem=-0.18232158, similarity=3.1637304|
 gt: 25567654 USRC_Rush_(1885), pem_pos=1,
 final_scores=-0.13754989, final_local_scores=-0.06425427,
 pem=-0.18232158, similarity=3.1637304]
 16: 7034 Cruiser
 [winner: 5734111 Bogatyr-class_cruiser, pem_pos=13,
 final_scores=-1.0578592, final_local_scores=-1.0340397,
 pem=-3.0445223, similarity=-10.211512|
 gt: 7034 Cruiser, pem_pos=0, final_scores=-1.1352911,
 final_local_scores=-1.2525144, pem=-0.15053521,
 similarity=-32.469177]
 17: 29012741 HMS_Porpoise_(1886)
 [winner: 21076424 HMS_Porpoise_(1799), pem_pos=0,
 final_scores=-0.049125314, final_local_scores=-0.021073632,
 pem=-0.6377971, similarity=6.9292736|
 gt: 29012741 HMS_Porpoise_(1886), pem_pos=4,
 final_scores=-0.2508906, final_local_scores=-0.20389894,
 pem=-1.5540878, similarity=6.456333]
 18: 2709921 HMS_Nymphe
 [winner: 2709921 HMS_Nymphe, pem_pos=0,
 final_scores=-0.25297368, final_local_scores=-0.12954704,
 pem=0.0, similarity=0.22377053|
 gt: 2709921 HMS_Nymphe, pem_pos=0,
 final_scores=-0.25297368, final_local_scores=-0.12954704,

pem=0.0, similarity=0.22377053]
19: 35880556 HMS_Pheasant
[winner: 13076222 HMS_Pheasant_(U49), pem_pos=0,
final_scores=-0.22182229, final_local_scores=-0.11907705,
pem=0.0, similarity=0.53621626]
RECALL MISS: 35880556 HMS_Pheasant]
20: 62413 Schooner
[winner: 62413 Schooner, pem_pos=0,
final_scores=-1.0923384, final_local_scores=-1.2114975,
pem=-0.008270214, similarity=-32.018818]
gt: 62413 Schooner, pem_pos=0, final_scores=-1.0923384,
final_local_scores=-1.2114975, pem=-0.008270214,
similarity=-32.018818]
1: 3434750 United_States
[pem_pos=0, final_scores=0.51157725,
final_local_scores=0.4847315, pem=-0.12994486,
similarity=19.261728]
2: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5480867,
final_local_scores=0.5222958, pem=-0.48936236,
similarity=22.33729]
3: 58055 Bering_Sea
[pem_pos=0, final_scores=0.24334632,
final_local_scores=0.2504562, pem=0.0, similarity=11.563826]
4: 20518076 United_States_Navy
[pem_pos=0, final_scores=0.28150296,
final_local_scores=0.26347145, pem=0.0, similarity=11.952227]
5: 26061 Royal_Navy
[pem_pos=0, final_scores=0.12764634,
final_local_scores=0.09767042, pem=0.0, similarity=7.004393]
6: 26296222 Aleutian_Islands
[pem_pos=0, final_scores=0.080662176,
final_local_scores=0.0625229, pem=0.0, similarity=5.955521]
7: 689 Asia
[pem_pos=0, final_scores=0.20721577,
final_local_scores=0.2134442, pem=-0.1494309,
similarity=11.271942]
1: 31730675 Bering_Sea_Squadron
[pem_pos=0, final_scores=0.07676296,
final_local_scores=0.15448785, pem=0.0, similarity=8.699939]
2: 3434750 United_States
[pem_pos=0, final_scores=0.4881553,
final_local_scores=0.45914656, pem=-0.073457584,
similarity=18.191036]
3: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5028149,
final_local_scores=0.47284317, pem=-0.09933331,
similarity=18.740486]
4: 58055 Bering_Sea
[pem_pos=0, final_scores=0.3088786,
final_local_scores=0.3220403, pem=0.0, similarity=13.700038]
5: 3434750 United_States


```

    [pem_pos=0, final_scores=0.56047016,
final_local_scores=0.53813964, pem=-0.12994486,
similarity=20.855534]
6: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.55507,
final_local_scores=0.5299239, pem=-0.48936236,
similarity=22.564928]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.60241085,
final_local_scores=0.58395344, pem=-0.12994486,
similarity=22.22271]
8: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5992035,
final_local_scores=0.57813317, pem=-0.48936236,
similarity=24.003588]
9: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.473723,
final_local_scores=0.44106466, pem=-0.48936236,
similarity=19.91319]

```

Sample: b'G-sharp_major'

G-sharp major is a **1 major scale** based on the musical **2 note 3 G-sharp** , consisting of the pitches **4 G , 5 A , , 6 C , 7 D** , and . Its **8 key signature** has six sharps and one . For clarity and simplicity , G-sharp major is usually notated as its enharmonic equivalent of **[1 9 A-flat 1] major** ; however , it does appear as a secondary key area in several works in sharp keys - most notably in the Prelude and **[2 Fugue 2]** in C-sharp major from **[1 1 Johann Sebastian Bach 1]** 's **[2 2 The Well-Tempered Clavier 2]** , Book 1 . The G-sharp minor prelude and fugue from the same set end with a **[3 10 Picardy 3] third** in G-sharp major . G-sharp major is tonicised briefly in several of **[3 3 Frédéric Chopin 3]** 's **11 nocturnes** in C-sharp minor . A section in the third **12 movement** of **[4 Chopin 4]** 's **13 Piano Concerto No. 1** is in G-sharp major , although the key signature has 4 sharps . The end of the **14 exposition** of the second movement **[4 4 Charles-Valentin Alkan 4]** 's **15 Grande sonate ` Les quatre âges '** , subtitled Quasi - **[5 5 Faust 5]** , is in G-sharp major , although written with a six-sharp key signature -LRB- the movement opens in **16 D-sharp minor** and ends in **17 F-sharp major** -RRB- . The final pages of **[6 6 A World Requiem 6]** by **18 John Foulds**

```

1: 9906294 Johann_Sebastian_Bach
2: 25496497 The_Well-Tempered_Clavier
3: 10823 Frédéric_Chopin
4: 439458 Charles-Valentin_Alkan
5: 84672 Faust
6: 12779036 A_World_Requiem

```

```

1: 20759 Major_scale
    [winner: 20759 Major_scale, pem_pos=1,
final_scores=-0.9887837, final_local_scores=-1.0723503,
pem=-0.00072702364, similarity=-27.907406]

```

gt: 20759 Major_scale, pem_pos=1,
 final_scores=-0.9887837, final_local_scores=-1.0723503,
 pem=-0.00072702364, similarity=-27.907406]
 2: 22026 Note
 [winner: 24885283 Note_verbale, pem_pos=3,
 final_scores=-1.0938504, final_local_scores=-1.0184107,
 pem=-1.3856297, similarity=-18.76642]
 gt: 22026 Note, pem_pos=1, final_scores=-1.3884326,
 final_local_scores=-1.4995813, pem=-0.96047807,
 similarity=-35.437553]
 3: 12665417 G#_(musical_note)
 [winner: 12665417 G#_(musical_note), pem_pos=0,
 final_scores=-0.11203122, final_local_scores=-0.037102178,
 pem=0.0, similarity=2.9825099]
 gt: 12665417 G#_(musical_note), pem_pos=0,
 final_scores=-0.11203122, final_local_scores=-0.037102178,
 pem=0.0, similarity=2.9825099]
 4: 12665417 G#_(musical_note)
 [winner: 12255 G, pem_pos=0, final_scores=-0.416826,
 final_local_scores=-0.3379531, pem=-1.9297645,
 similarity=4.4988785]
 RECALL MISS: 12665417 G#_(musical_note)]
 5: 12665496 A#_(musical_note) no overlap with any filtered
 span
 6: 12666226 C#_(musical_note)
 [winner: 6021 C_(programming_language), pem_pos=0,
 final_scores=-0.40930983, final_local_scores=-0.33872965,
 pem=-1.0105935, similarity=-0.52289087]
 RECALL MISS: 12666226 C#_(musical_note)]
 7: 12665654 D#_(musical_note) no overlap with any filtered
 span
 8: 17394 Key_signature
 [winner: 443212 Key_(instrument), pem_pos=1,
 final_scores=-0.83897483, final_local_scores=-0.820745,
 pem=-1.0603113, similarity=-14.636817]
 gt: 17394 Key_signature, pem_pos=8,
 final_scores=-1.457302, final_local_scores=-1.5705957,
 pem=-3.8801258, similarity=-21.679268]
 [winner: 252648 Diabolical_signature, pem_pos=4,
 final_scores=-1.1191969, final_local_scores=-1.0889001,
 pem=-1.5985636, similarity=-19.712004]
 RECALL MISS: 17394 Key_signature]
 9: 1969923 A-flat_major
 [winner: 1969923 A-flat_major, pem_pos=1,
 final_scores=0.03806539, final_local_scores=0.039609656,
 pem=-0.15684244, similarity=6.1246777]
 gt: 1969923 A-flat_major, pem_pos=1,
 final_scores=0.03806539, final_local_scores=0.039609656,
 pem=-0.15684244, similarity=6.1246777]
 [winner: 15707825 Major,_Saskatchewan, pem_pos=19,
 final_scores=-1.7212611, final_local_scores=-1.6703799,
 pem=-4.4332848, similarity=-21.648859]

RECALL MISS: 1969923 A-flat_major]
 10: 886963 Picardy_third
 [winner: 97048 Picardy, pem_pos=0,
 final_scores=0.11864458, final_local_scores=0.21024767,
 pem=0.0, similarity=10.363923]
 RECALL MISS: 886963 Picardy_third]
 [winner: 13719091 Third_(play), pem_pos=6,
 final_scores=-1.3718129, final_local_scores=-1.3820648,
 pem=-3.248335, similarity=-19.488903]
 RECALL MISS: 886963 Picardy_third]
 11: 23021759 Nocturnes_(Chopin)
 [winner: 20584637 Nocturnes_(book), pem_pos=1,
 final_scores=-0.9723616, final_local_scores=-0.9410192,
 pem=-1.2279081, similarity=-17.314621]
 gt: 23021759 Nocturnes_(Chopin), pem_pos=2,
 final_scores=-1.236561, final_local_scores=-1.3549712,
 pem=-2.2381852, similarity=-24.173727]
 12: 462250 Movement_(music)
 [winner: 2035515 Symphony_No._1_(Brahms), pem_pos=0,
 final_scores=-0.98742, final_local_scores=-1.0993066,
 pem=0.0, similarity=-28.715788]
 RECALL MISS: 462250 Movement_(music)]
 13: 1646154 Piano_Concerto_No._1_(Chopin)
 [winner: 306714 Piano_concerto, pem_pos=0,
 final_scores=-0.19979213, final_local_scores=-0.2802196,
 pem=-0.78851694, similarity=0.015478043]
 RECALL MISS: 1646154 Piano_Concerto_No._1_(Chopin)]
 [winner: 21460 November_1, pem_pos=9,
 final_scores=-0.9986924, final_local_scores=-0.88748384,
 pem=-3.916656, similarity=-1.0951914]
 RECALL MISS: 1646154 Piano_Concerto_No._1_(Chopin)]
 [winner: 564821 Numero_sign, pem_pos=1,
 final_scores=-1.0098294, final_local_scores=-1.0104449,
 pem=-1.0986123, similarity=-20.089552]
 RECALL MISS: 1646154 Piano_Concerto_No._1_(Chopin)]
 14: 14124258 Exposition_(music)
 [winner: 14124258 Exposition_(music), pem_pos=1,
 final_scores=-0.9727605, final_local_scores=-1.0442529,
 pem=-0.4484952, similarity=-24.63389]
 gt: 14124258 Exposition_(music), pem_pos=1,
 final_scores=-0.9727605, final_local_scores=-1.0442529,
 pem=-0.4484952, similarity=-24.63389]
 15: 13840456 Grande_sonate_'Les_quatre_âges'
 [winner: 237789 Les_Misérables_(musical), pem_pos=1,
 final_scores=-0.5283843, final_local_scores=-0.5555533,
 pem=-2.0576541, similarity=-1.2992636]
 RECALL MISS: 13840456 Grande_sonate_'Les_quatre_âges']
 [winner: 11751596 José_Antonio_Grande, pem_pos=7,
 final_scores=-0.5613488, final_local_scores=-0.42262277,
 pem=-2.9548862, similarity=7.5469365]
 RECALL MISS: 13840456 Grande_sonate_'Les_quatre_âges']
 [winner: 8982983 Quatre, pem_pos=0,

final_scores=-0.73191583, final_local_scores=-0.5855558,
 pem=0.0, similarity=-13.38444|
 RECALL MISS: 13840456 Grande_sonate_'Les_quatre_âges']
 [winner: 3912979 Piano_Sonata_No._10_(Beethoven),
 pem_pos=0, final_scores=-0.9343436,
 final_local_scores=-0.9911202, pem=0.0,
 similarity=-25.487291|
 RECALL MISS: 13840456 Grande_sonate_'Les_quatre_âges']
 [winner: 5110779 Grave_accent, pem_pos=0,
 final_scores=-1.5392754, final_local_scores=-1.6232378,
 pem=0.0, similarity=-44.350945|
 RECALL MISS: 13840456 Grande_sonate_'Les_quatre_âges']
 [winner: 2302349 Vedanga, pem_pos=16,
 final_scores=-1.8024225, final_local_scores=-1.8015535,
 pem=-4.75359, similarity=-23.821474|
 RECALL MISS: 13840456 Grande_sonate_'Les_quatre_âges']
 16: 2069247 D-sharp_minor
 [winner: 2069247 D-sharp_minor, pem_pos=1,
 final_scores=-0.21668036, final_local_scores=-0.21145372,
 pem=-0.6931472, similarity=1.5489541|
 gt: 2069247 D-sharp_minor, pem_pos=1,
 final_scores=-0.21668036, final_local_scores=-0.21145372,
 pem=-0.6931472, similarity=1.5489541]
 [winner: 417034 Minor_chord, pem_pos=1,
 final_scores=-1.2265419, final_local_scores=-1.2719429,
 pem=-1.52788, similarity=-25.558752|
 RECALL MISS: 2069247 D-sharp_minor]
 17: 1969903 F-sharp_major
 [winner: 2022271 F-sharp_minor, pem_pos=2,
 final_scores=-0.2360178, final_local_scores=-0.25536308,
 pem=-0.926762, similarity=1.5090449|
 gt: 1969903 F-sharp_major, pem_pos=1,
 final_scores=-0.25113493, final_local_scores=-0.28468382,
 pem=-0.926762, similarity=0.634055]
 [winner: 15707825 Major,_Saskatchewan, pem_pos=19,
 final_scores=-1.672567, final_local_scores=-1.6171889,
 pem=-4.4332848, similarity=-20.061539|
 RECALL MISS: 1969903 F-sharp_major]
 18: 2174412 John_Foulds
 [winner: 2174412 John_Foulds, pem_pos=0,
 final_scores=-0.011264816, final_local_scores=-0.016139641,
 pem=0.0, similarity=3.608074|
 gt: 2174412 John_Foulds, pem_pos=0,
 final_scores=-0.011264816, final_local_scores=-0.016139641,
 pem=0.0, similarity=3.608074]
 1: 9906294 Johann_Sebastian_Bach
 [pem_pos=0, final_scores=0.40936577,
 final_local_scores=0.35382414, pem=0.0, similarity=14.648531]
 2: 25496497 The_Well-Tempered_Clavier
 [pem_pos=0, final_scores=0.23122774,
 final_local_scores=0.19763848, pem=0.0, similarity=9.987639]
 3: 10823 Frédéric_Chopin

[pem_pos=0, final_scores=0.4244544,
 final_local_scores=0.37241924, pem=0.0, similarity=15.203446]
 4: 439458 Charles-Valentin_Alkan
 [pem_pos=0, final_scores=0.50914556,
 final_local_scores=0.50466174, pem=0.0, similarity=19.149826]
 5: 84672 Faust
 [pem_pos=0, final_scores=0.11025505,
 final_local_scores=0.10676558, pem=-0.84974116,
 similarity=11.896835]
 6: 12779036 A_World_Requiem
 [pem_pos=0, final_scores=0.04359494,
 final_local_scores=0.0782111, pem=0.0, similarity=6.423688]
 1: 1969923 A-flat_major
 [pem_pos=1, final_scores=0.03806539,
 final_local_scores=0.039609656, pem=-0.15684244,
 similarity=6.1246777]
 2: 10897 Fugue
 [pem_pos=0, final_scores=0.08092965,
 final_local_scores=0.022043675, pem=-0.040781412,
 similarity=4.9693155]
 3: 97048 Picardy
 [pem_pos=0, final_scores=0.11864458,
 final_local_scores=0.21024767, pem=0.0, similarity=10.363923]
 4: 10823 Frédéric_Chopin
 [pem_pos=0, final_scores=0.33140248,
 final_local_scores=0.27077422, pem=0.0, similarity=12.170156]

Sample: b'Hughie_Wilson'

Hughie ``Lalty`` Wilson -LRB- 18 March 1869 -- 7 April
 1940 -RRB- was a [1 1 Scottish 1] who played for [1 1 Sunderl
 and 1] and the [2 2 Scotland 2] national football team as an 3
 inside forward . Born in [2 2 Mauchline 2] , [3 3 Scotland 3]
 , Wilson came to prominence with , earning his first
 international cap while with that club . Wilson signed for [4
 4 Sunderland 4] from and made his debut for them on 13
 September 1890 against [5 5 Burnley 5] in a 3 -- 2 defeat at [6
 6 Newcastle Road 6] . Overall he made 258 league
 appearances scoring 46 goals while at the club , spanning
 from 1890 to 1899 . Wilson won with [3 Sunderland 3] in ,
 and . Wilson left [4 Sunderland 4] in 1899 to sign for [5 4 Be
 dminster 5] . A season later he moved to [7 7 Bristol City 7]
 , as a result of the merger between [6 Bedminster 6] and [7
 Bristol City 7] . Wilson then returned to [8 Scotland 8] in
 1901 with [8 8 Third Lanark 8] . He enjoyed success with
 Thirds , in , the [9 9 Scottish Cup 9] the , and reaching the
 final in . Wilson ended his playing days at [10 10 Kilmarnock
 10] . Wilson won his first cap for [11 11 Scotland 11]
 against [12 12 Wales 12] on 22 March 1890 in a 5 -- 0 win
 at . He went on to make 3 further caps for his country ,
 overall making 4 appearances scoring one goal . Wilson also
 represented the
 1: 184474 Sunderland_A.F.C.

2: 2961296 Mauchline
 3: 26994 Scotland
 4: 184474 Sunderland_A.F.C.
 5: 376725 Burnley_F.C.
 6: 16768374 Newcastle_Road_(stadium)
 7: 4392 Bristol_City_F.C.
 8: 783221 Third_Lanark_A.C.
 9: 883238 Scottish_Cup
 10: 405663 Kilmarnock_F.C.
 11: 578101 Scotland_national_football_team
 12: 580291 Wales_national_football_team
 1: 34850722 Scottish_people
 [winner: 26994 Scotland, pem_pos=0,
 final_scores=0.429282, final_local_scores=0.43822408,
 pem=-0.47951394, similarity=19.774864]
 gt: 34850722 Scottish_people, pem_pos=1,
 final_scores=-0.20019442, final_local_scores=-0.229369,
 pem=-1.5054524, similarity=5.4317684]
 2: 578101 Scotland_national_football_team
 [winner: 26994 Scotland, pem_pos=1,
 final_scores=0.20983903, final_local_scores=0.19851592,
 pem=-0.22905664, similarity=11.25947]
 gt: 578101 Scotland_national_football_team, pem_pos=2,
 final_scores=0.05103366, final_local_scores=-0.036486387,
 pem=-2.6956348, similarity=17.660168]
 [winner: 140801 Team, pem_pos=0, final_scores=-1.4186834,
 final_local_scores=-1.4752219, pem=-0.5173818,
 similarity=-37.120247]
 RECALL MISS: 578101 Scotland_national_football_team]
 [winner: 10568 Association_football, pem_pos=0,
 final_scores=-1.5678245, final_local_scores=-1.7190456,
 pem=-0.33534196, similarity=-45.3864]
 RECALL MISS: 578101 Scotland_national_football_team]
 [winner: 302395 National_(brand), pem_pos=3,
 final_scores=-1.7936176, final_local_scores=-1.8242798,
 pem=-2.9526155, similarity=-34.293655]
 RECALL MISS: 578101 Scotland_national_football_team]
 3: 428681 Forward_(association_football)
 [winner: 27307399 Inside_(White_Heart_album), pem_pos=9,
 final_scores=-1.0486073, final_local_scores=-0.91961724,
 pem=-2.4971073, similarity=-9.773843]
 RECALL MISS: 428681 Forward_(association_football)]
 [winner: 568962 Forward_chaining, pem_pos=6,
 final_scores=-1.3255998, final_local_scores=-1.2536404,
 pem=-2.2876325, similarity=-20.880919]
 gt: 428681 Forward_(association_football), pem_pos=0,
 final_scores=-1.6677573, final_local_scores=-1.8250105,
 pem=-0.7413364, similarity=-46.340736]
 4: 16143517 Bedminster_F.C.
 [winner: 4392 Bristol_City_F.C., pem_pos=4,
 final_scores=0.85659784, final_local_scores=0.84063584,
 pem=-1.50417, similarity=37.355873]

gt: 16143517 Bedminster_F.C., pem_pos=5,
final_scores=-0.40914732, final_local_scores=-0.44147602,
pem=-4.09527, similarity=13.185904]
1: 184474 Sunderland_A.F.C.
[pem_pos=0, final_scores=0.6274565,
final_local_scores=0.58732104, pem=-0.39640337,
similarity=23.772247]
2: 2961296 Mauchline
[pem_pos=0, final_scores=0.31210047,
final_local_scores=0.3676653, pem=0.0, similarity=15.061578]
3: 26994 Scotland
[pem_pos=1, final_scores=0.4477626,
final_local_scores=0.45841128, pem=-0.22905664,
similarity=19.015266]
4: 184474 Sunderland_A.F.C.
[pem_pos=0, final_scores=0.70032245,
final_local_scores=0.6669161, pem=-0.39640337,
similarity=26.147522]
5: 376725 Burnley_F.C.
[pem_pos=1, final_scores=0.32711178,
final_local_scores=0.26408, pem=-0.9914584,
similarity=17.362091]
6: 16768374 Newcastle_Road_(stadium)
[pem_pos=0, final_scores=0.031859264,
final_local_scores=0.0704623, pem=0.0, similarity=6.1924486]
7: 4392 Bristol_City_F.C.
[pem_pos=0, final_scores=0.39842033,
final_local_scores=0.3401465, pem=0.0, similarity=14.240363]
8: 783221 Third_Lanark_A.C.
[pem_pos=0, final_scores=0.41596895,
final_local_scores=0.38014543, pem=0.0, similarity=15.4340105]
9: 883238 Scottish_Cup
[pem_pos=0, final_scores=0.4277122,
final_local_scores=0.42150754, pem=-0.058939878,
similarity=16.988863]
10: 405663 Kilmarnock_F.C.
[pem_pos=1, final_scores=0.51730806,
final_local_scores=0.4814189, pem=-1.4326576,
similarity=26.247229]
11: 578101 Scotland_national_football_team
[pem_pos=2, final_scores=0.13149662,
final_local_scores=0.056083143, pem=-2.6956348,
similarity=20.422626]
12: 580291 Wales_national_football_team
[pem_pos=1, final_scores=0.13387422,
final_local_scores=0.06995031, pem=-2.4526134,
similarity=19.514862]
1: 26994 Scotland
[pem_pos=0, final_scores=0.429282,
final_local_scores=0.43822408, pem=-0.47951394,
similarity=19.774864]
2: 26994 Scotland

```

    [pem_pos=1, final_scores=0.20983903,
final_local_scores=0.19851592, pem=-0.22905664,
similarity=11.25947]
3: 184474 Sunderland_A.F.C.
    [pem_pos=0, final_scores=0.71743464,
final_local_scores=0.6856085, pem=-0.39640337,
similarity=26.705341]
4: 184474 Sunderland_A.F.C.
    [pem_pos=0, final_scores=0.69075906,
final_local_scores=0.6564695, pem=-0.39640337,
similarity=25.835777]
5: 4392 Bristol_City_F.C.
    [pem_pos=4, final_scores=0.85659784,
final_local_scores=0.84063584, pem=-1.50417,
similarity=37.355873]
6: 4392 Bristol_City_F.C.
    [pem_pos=4, final_scores=0.7931743,
final_local_scores=0.77135533, pem=-1.50417,
similarity=35.288403]
7: 4392 Bristol_City_F.C.
    [pem_pos=0, final_scores=0.43603224,
final_local_scores=0.38123178, pem=0.0, similarity=15.466431]
8: 26994 Scotland
    [pem_pos=0, final_scores=0.3174352,
final_local_scores=0.31604844, pem=-0.22905664,
similarity=14.766874]

```

Sample: b'Nidderdale_Way'

The [1 Nidderdale Way 1] is a circular long distance footpath in [1 1 Nidderdale 1] in [2 2 North Yorkshire 2] , [2 England 2] . The path is usually joined at 1 Ripley , linked by the frequent 2 route 36 bus to [3 3 Leeds 3] and [4 4 Harrogate 4] , or [5 5 Pateley Bridge 5] , the only town on the route . The Way follows one side of the dale on the outward leg , and returns on the other side , so that the outward and return legs are never far apart . From Ripley the Way follows a series of side valleys on the northern side of [3 Nidderdale 3] and then climbs to [6 6 Brimham Rocks 6] . It then descends to the hamlet of Smelthouses and climbs to [4 Blazefield 4] before descending to [5 Pateley Bridge 5] . From [6 Pateley 6] the path keeps close to the [7 7 River Nidd 7] , passing [7 3 Wath 7] , 4 Gouthwaite Reservoir and [8 8 Bouthwaite 8] to 5 Lofthouse . The path then takes a route high above the [8 Nidd 8] to reach the dam of [9 9 Scar House Reservoir 9] , where it crosses to the right or south bank of the river . From the dam the path climbs an old track to reach its highest point at , before descending to the village of [10 10 Middlesmoor 10] . It then crosses Stean Beck to reach [11 1 Stean 11] , and descends on the western side of the valley to [12 12 Ramsgill 12] . A short section of road walking follows , beside [9 Gouthwaite Reservoir 9] . The Way then ascends to [10 6 Heathfield 10] , and takes a long loop

around a side valley to reach [13 13 Bowerley 13] . The path then follows a high rocky section known as Guise Cliff , before descending to . It then follows the [11 Nidd 11] downstream to 7 Darley , [14 14 Birstwith 14] and [15 15 Hampsthwaite 15] , where it crosses the [12 Nidd 12] to reach 8 Clint . For the last few miles the path takes a gentle course past [13 9 Ripley Castle 13]

1: 3119850 Nidderdale
2: 143759 North_Yorkshire
3: 8262427 Leeds
4: 266660 Harrogate
5: 1854793 Pateley_Bridge
6: 1124057 Brimham_Rocks
7: 6027579 River_Nidd
8: 12881701 Bouthwaite
9: 6028001 Scar_House_Reservoir
10: 6751632 Middlesmoor
11: 8500719 Stean
12: 6752019 Ramsgill
13: 3231805 Bowerley
14: 3275656 Birstwith
15: 4824640 Hampsthwaite

1: 3311326 Ripley, North_Yorkshire
[winner: 227521 Ripley, Derbyshire, pem_pos=2,
final_scores=-0.1434646, final_local_scores=-0.1950612,
pem=-2.5742245, similarity=12.26773]
gt: 3311326 Ripley, North_Yorkshire, pem_pos=9,
final_scores=-0.42897865, final_local_scores=-0.44093046,
pem=-3.166151, similarity=8.149491]

2: 16496138 Transdev_Harrogate & District_route_36
[winner: 6333283 No. 36 Squadron RAF, pem_pos=0,
final_scores=-0.55525094, final_local_scores=-0.4579725,
pem=-2.8564863, similarity=5.956916]

RECALL MISS: 16496138
Transdev_Harrogate & District_route_36]
[winner: 27161403 Route (command), pem_pos=15,
final_scores=-1.1630214, final_local_scores=-1.0706571,
pem=-2.6855774, similarity=-13.256244]

RECALL MISS: 16496138
Transdev_Harrogate & District_route_36]

3: 8284822 Wath-in-Nidderdale
[winner: 2089418 Wath-upon-Deane, pem_pos=0,
final_scores=0.16410814, final_local_scores=0.15897173,
pem=-0.42177382, similarity=11.1274185]
gt: 8284822 Wath-in-Nidderdale, pem_pos=2,
final_scores=-0.19764076, final_local_scores=-0.19796193,
pem=-1.2963462, similarity=5.2318664]

4: 10092551 Gouthwaite_Reservoir
[winner: 10092551 Gouthwaite_Reservoir, pem_pos=0,
final_scores=-0.0035329908, final_local_scores=0.012697756,
pem=0.0, similarity=4.4686394]
gt: 10092551 Gouthwaite_Reservoir, pem_pos=0,

final_scores=-0.0035329908, final_local_scores=0.012697756,
pem=0.0, similarity=4.4686394]
5: 6752172 Lofthouse, North_Yorkshire
[winner: 5133351 Lofthouse, West_Yorkshire, pem_pos=1,
final_scores=-0.08875877, final_local_scores=-0.048792683,
pem=-0.51571745, similarity=5.438194]
gt: 6752172 Lofthouse, North_Yorkshire, pem_pos=2,
final_scores=-0.2442732, final_local_scores=-0.26581198,
pem=-1.8602698, similarity=6.2737904]
6: 32789506 Heathfield, North_Yorkshire
[winner: 2442752 Heathfield, East_Sussex, pem_pos=0,
final_scores=0.33987814, final_local_scores=0.36284667,
pem=-0.50992316, similarity=17.690823]
gt: 32789506 Heathfield, North_Yorkshire, pem_pos=12,
final_scores=-1.048012, final_local_scores=-1.088743,
pem=-6.134482, similarity=4.959717]
7: 577895 Darley, North_Yorkshire
[winner: 577894 Darley_Dale, pem_pos=9,
final_scores=-0.12928961, final_local_scores=-0.15898311,
pem=-2.989202, similarity=15.601084]
gt: 577895 Darley, North_Yorkshire, pem_pos=2,
final_scores=-0.35631132, final_local_scores=-0.34017298,
pem=-2.159385, similarity=5.681344]
8: 36062629 Clint, North_Yorkshire
[winner: 135822 Clint, Texas, pem_pos=1,
final_scores=-0.3172031, final_local_scores=-0.22275034,
pem=-1.2943714, similarity=4.4813914]
RECALL MISS: 36062629 Clint, North_Yorkshire]
9: 17635355 Ripley_Castle
[winner: 3311326 Ripley, North_Yorkshire, pem_pos=0,
final_scores=0.04269804, final_local_scores=0.08515982,
pem=0.0, similarity=6.631052]
gt: 17635355 Ripley_Castle, pem_pos=1,
final_scores=-0.2888168, final_local_scores=-0.2645486,
pem=-1.4604024, similarity=4.136951]
1: 3119850 Nidderdale
[pem_pos=0, final_scores=0.46375555,
final_local_scores=0.46650857, pem=0.0, similarity=18.01126]
2: 143759 North_Yorkshire
[pem_pos=0, final_scores=0.65879136,
final_local_scores=0.64626944, pem=0.0, similarity=23.375683]
3: 8262427 Leeds
[pem_pos=0, final_scores=0.2916438,
final_local_scores=0.3022628, pem=-0.25452003,
similarity=14.4939575]
4: 266660 Harrogate
[pem_pos=0, final_scores=0.278112,
final_local_scores=0.25077444, pem=0.0, similarity=11.573322]
5: 1854793 Pateley_Bridge
[pem_pos=0, final_scores=0.33308727,
final_local_scores=0.30467868, pem=0.0, similarity=13.181932]
6: 1124057 Brimham_Rocks

[pem_pos=0, final_scores=0.076875195,
final_local_scores=0.15747827, pem=0.0, similarity=8.789179]
7: 6027579 River_Nidd
[pem_pos=0, final_scores=0.39460653,
final_local_scores=0.3673821, pem=0.0, similarity=15.053127]
8: 12881701 Bouthwaite
[pem_pos=0, final_scores=0.101919815,
final_local_scores=0.16534296, pem=0.0, similarity=9.023876]
9: 6028001 Scar_House_Reservoir
[pem_pos=0, final_scores=0.2710836,
final_local_scores=0.33740225, pem=0.0, similarity=14.158469]
10: 6751632 Middlesmoor
[pem_pos=0, final_scores=0.22522788,
final_local_scores=0.2807525, pem=0.0, similarity=12.467928]
11: 8500719 Stean
[pem_pos=0, final_scores=0.30509102,
final_local_scores=0.36923474, pem=0.0, similarity=15.108414]
12: 6752019 Ramsgill
[pem_pos=0, final_scores=0.17420943,
final_local_scores=0.25894126, pem=0.0, similarity=11.817037]
13: 3231805 Bewerley
[pem_pos=0, final_scores=0.33724117,
final_local_scores=0.37434417, pem=0.0, similarity=15.26089]
14: 3275656 Birstwith
[pem_pos=0, final_scores=0.42908245,
final_local_scores=0.47489637, pem=0.0, similarity=18.261568]
15: 4824640 Hampsthwaite
[pem_pos=0, final_scores=0.16643669,
final_local_scores=0.19130385, pem=0.0, similarity=9.798601]
1: 38160430 Nidderdale_Way
[pem_pos=0, final_scores=0.11546032,
final_local_scores=0.15747628, pem=0.0, similarity=8.789119]
2: 9316 England
[pem_pos=0, final_scores=0.62644255,
final_local_scores=0.6718725, pem=-0.19053856,
similarity=25.175907]
3: 3119850 Nidderdale
[pem_pos=0, final_scores=0.22856762,
final_local_scores=0.20960152, pem=0.0, similarity=10.344641]
4: 1086090 Transdev_Blazefield
[pem_pos=0, final_scores=0.19583045,
final_local_scores=0.24717802, pem=0.0, similarity=11.465998]
5: 1854793 Pateley_Bridge
[pem_pos=0, final_scores=0.42753273,
final_local_scores=0.4078459, pem=0.0, similarity=16.260649]
6: 1854793 Pateley_Bridge
[pem_pos=0, final_scores=0.3070789,
final_local_scores=0.27626842, pem=0.0, similarity=12.334114]
7: 2089418 Wath-upon-Dearne
[pem_pos=0, final_scores=0.16410814,
final_local_scores=0.15897173, pem=-0.42177382,
similarity=11.1274185]

8: 6027579 River_Nidd
 [pem_pos=1, final_scores=0.21795483,
 final_local_scores=0.17320198, pem=-0.16276108,
 similarity=10.143526]
 9: 10092551 Gouthwaite_Reservoir
 [pem_pos=0, final_scores=0.07011436,
 final_local_scores=0.093146294, pem=0.0, similarity=6.8693843]
 10: 2442752 Heathfield,_East_Sussex
 [pem_pos=0, final_scores=0.33987814,
 final_local_scores=0.36284667, pem=-0.50992316,
 similarity=17.690823]
 11: 6027579 River_Nidd
 [pem_pos=1, final_scores=0.17384468,
 final_local_scores=0.12501833, pem=-0.16276108,
 similarity=8.705629]
 12: 6027579 River_Nidd
 [pem_pos=1, final_scores=0.3005401,
 final_local_scores=0.26341382, pem=-0.16276108,
 similarity=12.835627]
 13: 3311326 Ripley,_North_Yorkshire
 [pem_pos=0, final_scores=0.04269804,
 final_local_scores=0.08515982, pem=0.0, similarity=6.631052]

Sample: b'Philip_Threlfall'

Philip Walter Threlfall -LRB- born 11 February 1967 -RRB- is a former [1 1 English 1] 2 cricket er . Threlfall was a right-handed 3 batsman who bowled right-arm 4 medium-fast . He was born at [1 1 Barrow-in-Furness 1] , [2 5 Lancashire 2] . Threlfall played for [2 2 Cumberland 2] in 1987 , making a single appearance in the against [3 3 Cambridgeshire 3] and a single appearance in the [4 4 MCCA Knockout Trophy 4] against [5 5 Cheshire 5] . In 1988 , Threlfall played [3 Second XI 3] cricket for [6 6 Sussex 6] , with him also making his 6 first-class debut for the county in that season against [7 7 Somers et 7] at the [4 7 Recreation Ground 4] , [5 Bath 5] . He played regularly for the [6 Sussex Second XI 6] , but would only make two further first-class appearances for the county , against the touring [8 8 Sri Lankans 8] in 1990 and [7 8 Cambridge University 7] in 1991 . He never batted in his three first-class appearances , but with the ball he took 7 wickets at an 9 average of 18.57 , with best figures of 3/45 . He also made a single [9 9 List A 9] appearance for the county against the touring [10 10 Zimbabweans 10] in 1990 . He ended [8 Sussex 8] 's innings of 233/8 10 unbeaten on 17 , while in the [9 Zimbabweans 9] innings he took figures of 3/40 from 10 11 overs

1: 353068 Barrow-in-Furness
 2: 3702699 Cumberland_County_Cricket_Club
 3: 1986069 Cambridgeshire_County_Cricket_Club
 4: 7543416 MCCA_Knockout_Trophy
 5: 3712477 Cheshire_County_Cricket_Club
 6: 1763691 Sussex_County_Cricket_Club

7: 1622178 Somerset_County_Cricket_Club
 8: 397246 Sri_Lanka_national_cricket_team
 9: 683459 List_A_cricket
 10: 759317 Zimbabwe_national_cricket_team
 1: 18803164 English_people
 [winner: 9316 England, pem_pos=1, final_scores=0.5106135,
 final_local_scores=0.5414627, pem=-1.3527626,
 similarity=27.604572]
 gt: 18803164 English_people, pem_pos=2,
 final_scores=-0.41436684, final_local_scores=-0.4608621,
 pem=-2.2084396, similarity=2.3465059]
 2: 25675557 Cricket
 [winner: 330091 Rhabdophoridae, pem_pos=10,
 final_scores=-1.4534321, final_local_scores=-1.4135146,
 pem=-2.56456, similarity=-24.145905]
 gt: 25675557 Cricket, pem_pos=0,
 final_scores=-1.7512337, final_local_scores=-1.9656947,
 pem=-0.07532314, similarity=-54.16092]
 3: 673526 Batting_(cricket)
 [winner: 673526 Batting_(cricket), pem_pos=0,
 final_scores=-1.2096456, final_local_scores=-1.3420992,
 pem=0.0, similarity=-35.961205]
 gt: 673526 Batting_(cricket), pem_pos=0,
 final_scores=-1.2096456, final_local_scores=-1.3420992,
 pem=0.0, similarity=-35.961205]
 4: 689836 Seam_bowling
 [winner: 678695 Fast_bowling, pem_pos=0,
 final_scores=-0.8926992, final_local_scores=-1.0343618,
 pem=0.0, similarity=-26.777708]
 RECALL MISS: 689836 Seam_bowling]
 5: 50435 Lancashire
 [winner: 749919 Lancashire_County_Cricket_Club,
 pem_pos=1, final_scores=0.30701536,
 final_local_scores=0.23162103, pem=-1.7295884,
 similarity=20.407516]
 gt: 50435 Lancashire, pem_pos=0,
 final_scores=-0.038315967, final_local_scores=-0.06990248,
 pem=-0.20611861, similarity=3.1245885]
 6: 11040 First-class_cricket
 [winner: 662844 First-class_citizen, pem_pos=3,
 final_scores=-1.2382318, final_local_scores=-1.1576147,
 pem=-1.5935546, similarity=-21.789818]
 gt: 11040 First-class_cricket, pem_pos=0,
 final_scores=-1.2635152, final_local_scores=-1.4652293,
 pem=-0.053819276, similarity=-39.342976]
 7: 4674266 Recreation_Ground_(Bath)
 [winner: 41523 Bath,_Somerset, pem_pos=0,
 final_scores=0.16980396, final_local_scores=0.1767177,
 pem=-0.5799331, similarity=12.517088]
 gt: 4674266 Recreation_Ground_(Bath), pem_pos=18,
 final_scores=-0.39362174, final_local_scores=-0.46967518,
 pem=-3.7212925, similarity=10.3106365]

[winner: 4674266 Recreation_Ground_(Bath), pem_pos=1,
final_scores=0.11856796, final_local_scores=0.098989174,
pem=-1.00141, similarity=12.489571]

gt: 4674266 Recreation_Ground_(Bath), pem_pos=1,
final_scores=0.11856796, final_local_scores=0.098989174,
pem=-1.00141, similarity=12.489571]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7948585, final_local_scores=-1.783956,
pem=-4.6728287, similarity=-23.735521]

RECALL MISS: 4674266 Recreation_Ground_(Bath)]

8: 1622256 Cambridge_University_Cricket_Club

[winner: 25978572 University_of_Cambridge, pem_pos=0,
final_scores=0.14727615, final_local_scores=0.18282634,
pem=-0.12541339, similarity=10.227633]

gt: 1622256 Cambridge_University_Cricket_Club,
pem_pos=1, final_scores=0.13073958,
final_local_scores=0.059546098, pem=-1.3585418,
similarity=13.254649]

9: 875871 Bowling_average

[winner: 875871 Bowling_average, pem_pos=0,
final_scores=-1.0039725, final_local_scores=-1.1420817,
pem=0.0, similarity=-29.992281]

gt: 875871 Bowling_average, pem_pos=0,
final_scores=-1.0039725, final_local_scores=-1.1420817,
pem=0.0, similarity=-29.992281]

10: 2430357 Not_out

[winner: 2430357 Not_out, pem_pos=0,
final_scores=-0.88258857, final_local_scores=-1.0325825,
pem=-0.53147423, similarity=-23.834372]

gt: 2430357 Not_out, pem_pos=0,
final_scores=-0.88258857, final_local_scores=-1.0325825,
pem=-0.53147423, similarity=-23.834372]

11: 463787 Over_(cricket)

[winner: 463787 Over_(cricket), pem_pos=0,
final_scores=-1.0869309, final_local_scores=-1.2341003,
pem=0.0, similarity=-32.7383]

gt: 463787 Over_(cricket), pem_pos=0,
final_scores=-1.0869309, final_local_scores=-1.2341003,
pem=0.0, similarity=-32.7383]

1: 353068 Barrow-in-Furness

[pem_pos=0, final_scores=0.18045582,
final_local_scores=0.2351256, pem=0.0, similarity=11.10633]

2: 3702699 Cumberland_County_Cricket_Club

[pem_pos=5, final_scores=0.08521621,
final_local_scores=0.030752182, pem=-3.2953885,
similarity=22.928253]

3: 1986069 Cambridgeshire_County_Cricket_Club

[pem_pos=1, final_scores=0.5872056,
final_local_scores=0.5566687, pem=-1.9373415,
similarity=31.237381]

4: 7543416 MCCA_Knockout_Trophy

[pem_pos=0, final_scores=0.39726883,

final_local_scores=0.3711068, pem=0.0, similarity=15.164281]
5: 3712477 Cheshire_County_Cricket_Club
[pem_pos=2, final_scores=0.19018032,
final_local_scores=0.14160082, pem=-2.629504,
similarity=22.615015]
6: 1763691 Sussex_County_Cricket_Club
[pem_pos=1, final_scores=0.5945832,
final_local_scores=0.55090076, pem=-1.5566913,
similarity=28.99522]
7: 1622178 Somerset_County_Cricket_Club
[pem_pos=1, final_scores=0.66662246,
final_local_scores=0.6218033, pem=-2.4139037,
similarity=35.772747]
8: 397246 Sri_Lanka_national_cricket_team
[pem_pos=0, final_scores=0.60703206,
final_local_scores=0.6155567, pem=-0.17988552,
similarity=23.437399]
9: 683459 List_A_cricket
[pem_pos=0, final_scores=0.22834684,
final_local_scores=0.17319363, pem=-0.020267159,
similarity=9.368372]
10: 759317 Zimbabwe_national_cricket_team
[pem_pos=0, final_scores=0.74801975,
final_local_scores=0.7773806, pem=0.0, similarity=27.2883]
1: 9316 England
[pem_pos=1, final_scores=0.5106135,
final_local_scores=0.5414627, pem=-1.3527626,
similarity=27.604572]
2: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.30701536,
final_local_scores=0.23162103, pem=-1.7295884,
similarity=20.407516]
3: 7564410 Second_XI_Championship
[pem_pos=0, final_scores=0.1023985,
final_local_scores=0.051046357, pem=0.0, similarity=5.613038]
4: 4674266 Recreation_Ground_(Bath)
[pem_pos=1, final_scores=0.11856796,
final_local_scores=0.098989174, pem=-1.00141,
similarity=12.489571]
5: 41523 Bath,_Somerset
[pem_pos=0, final_scores=0.16980396,
final_local_scores=0.1767177, pem=-0.5799331,
similarity=12.517088]
6: 1763691 Sussex_County_Cricket_Club
[pem_pos=0, final_scores=0.5401502,
final_local_scores=0.49144095, pem=0.0, similarity=18.755291]
7: 25978572 University_of_Cambridge
[pem_pos=0, final_scores=0.14727615,
final_local_scores=0.18282634, pem=-0.12541339,
similarity=10.227633]
8: 1763691 Sussex_County_Cricket_Club
[pem_pos=1, final_scores=0.7020102,

final_local_scores=0.6682485, pem=-1.5566913,
similarity=32.49711]
9: 759317 Zimbabwe_national_cricket_team
[pem_pos=0, final_scores=0.71120256,
final_local_scores=0.73716336, pem=0.0, similarity=26.088139]

Sample: b'Roy_Vincent'

Roy Vincent -LRB- 6 February 1892 -- 5 June 1965 -RRB- was an [1 Australian 1] politician and a member of the [1 1 New South Wales Legislative Assembly 1] from 1922 until 1953 . He was a member of the `` True Blue '' faction of the 1 Progressive Party until it became the [2 2 Country Party 2] in 1927 . He was the party 's 2 Deputy Leader and Whip between 1950 and 1953 . He held ministerial rank as the Secretary of Mines and Minister for Forests between 1932 and 1941 . Vincent was born at [2 3 Glen Innes 2] , [3 New South Wales 3] and was the son of a newspaper editor . He was educated at [3 3 Uralla 3] and became a newspaper reporter but eventually edited and owned the Don Dorrigo Gazette in 4 Dorrigo . He became active in community organizations including the [4 4 New England New State Movement 4] of [5 5 Earle Page 5] , the 5 Returned and Services League of Australia and the [6 6 Aboriginal Protection Board 6] . Vincent served in the [7 7 First Australian Imperial Force 7] during [4 World War One 4] and was severely wounded and gassed in [5 France 5] . [6 Vincent 6] was elected to the parliament as a Progressive Party member for in the multi-member seat of 6 Oxley at the 7 1922 state election . He defeated the sitting member and his Progressive Party colleague . He moved to the seat of [7 8 Raleigh 7] as a member of the [8 Country Party 8] when [9 New South Wales 9] reverted to single member electorates at the 9 1927 election . He retained this seat for the next eight elections and retired at the 10 1953 election . The state election of 11 1932 saw the landslide defeat of [10 Labor 10] and the formation of a conservative coalition government by [8 8 Bert ram Stevens 8] . [11 Vincent 11] was the and Minister for Forrests throughout the premiership of [12 Stevens 12] and his successor [9 9 Alexander Mair 9] whose government was defeated at the 12 1941

1: 579433 New_South_Wales_Legislative_Assembly
2: 21927 National_Party_of_Australia
3: 613897 Uralla,_New_South_Wales
4: 18056083 New_England_New_State_Movement
5: 10357 Earle_Page
6: 2012063 Aboriginal_Protection_Board
7: 415475 First_Australian_Imperial_Force
8: 2770934 Bertram_Stevens_(politician)
9: 2770889 Alexander_Mair

1: 21453609 Progressive_Party_(1920)
[winner: 409359 Progressive_Party_(Iceland), pem_pos=1,
final_scores=-0.07770781, final_local_scores=-0.02360773,
pem=-1.9475602, similarity=13.976346]

gt: 21453609 Progressive_Party_(1920), pem_pos=18,
final_scores=-0.3828665, final_local_scores=-0.40839657,
pem=-3.3564174, similarity=10.155066]

2: 7133116 Leader_of_the_New_South_Wales_National_Party
[winner: 472615 Deputy_leader, pem_pos=0,
final_scores=-0.32016003, final_local_scores=-0.29785275,
pem=0.0, similarity=-4.798807]

gt: 7133116
Leader_of_the_New_South_Wales_National_Party, pem_pos=5,
final_scores=-1.1838964, final_local_scores=-1.280977,
pem=-5.8478827, similarity=-2.3355043]

3: 632836 Glen_Innes,_New_South_Wales
[winner: 21654 New_South_Wales, pem_pos=0,
final_scores=0.6484396, final_local_scores=0.661062,
pem=-0.051332146, similarity=24.096273]

RECALL MISS: 632836 Glen_Innes,_New_South_Wales]
[winner: 632836 Glen_Innes,_New_South_Wales, pem_pos=0,
final_scores=0.41493016, final_local_scores=0.4531927,
pem=-0.26889536, similarity=19.076181]

gt: 632836 Glen_Innes,_New_South_Wales, pem_pos=0,
final_scores=0.41493016, final_local_scores=0.4531927,
pem=-0.26889536, similarity=19.076181]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.9155983, final_local_scores=-1.8881211,
pem=-4.6728287, similarity=-26.844015]

RECALL MISS: 632836 Glen_Innes,_New_South_Wales]

4: 13827279 Dorrigo,_New_South_Wales
[winner: 13827279 Dorrigo,_New_South_Wales, pem_pos=0,
final_scores=0.029843852, final_local_scores=0.0685613,
pem=0.0, similarity=6.135719]

gt: 13827279 Dorrigo,_New_South_Wales, pem_pos=0,
final_scores=0.029843852, final_local_scores=0.0685613,
pem=0.0, similarity=6.135719]

5: 185528 Returned_and_Services_League_of_Australia
[winner: 185528
Returned_and_Services_League_of_Australia, pem_pos=0,
final_scores=-0.1851761, final_local_scores=-0.23183325,
pem=0.0, similarity=-2.8286538]

gt: 185528 Returned_and_Services_League_of_Australia,
pem_pos=0, final_scores=-0.1851761,
final_local_scores=-0.23183325, pem=0.0,
similarity=-2.8286538]

6: 4047731 Electoral_district_of_Oxley_(New_South_Wales)
[winner: 1838684 Oxley,_Queensland, pem_pos=1,
final_scores=-0.07897012, final_local_scores=-0.065179735,
pem=-1.5838208, similarity=10.757684]

gt: 4047731
Electoral_district_of_Oxley_(New_South_Wales), pem_pos=11,
final_scores=-0.4211235, final_local_scores=-0.46560395,
pem=-3.0553594, similarity=6.8106813]

7: 36358813 New_South_Wales_state_election,_1922
[winner: 36358813 New_South_Wales_state_election,_1922,

```

pem_pos=0, final_scores=-0.64172673,
final_local_scores=-0.66013676, pem=0.0,
similarity=-15.610085|
    gt: 36358813 New_South_Wales_state_election,_1922,
pem_pos=0, final_scores=-0.64172673,
final_local_scores=-0.66013676, pem=0.0,
similarity=-15.610085 ]
8: 10435310 Electoral_district_of_Raleigh
    [winner: 57845 Raleigh,_North_Carolina, pem_pos=0,
final_scores=0.2020383, final_local_scores=0.2479524,
pem=-0.35659322, similarity=13.428316|
    gt: 10435310 Electoral_district_of_Raleigh, pem_pos=6,
final_scores=-0.72645444, final_local_scores=-0.74010575,
pem=-3.5044794, similarity=1.061384 ]
9: 25980089 New_South_Wales_state_election,_1927
    [winner: 2228301 Manitoba_general_election,_1927,
pem_pos=0, final_scores=-0.7374391,
final_local_scores=-0.7808754, pem=-0.06479999,
similarity=-18.860773|
    gt: 25980089 New_South_Wales_state_election,_1927,
pem_pos=3, final_scores=-0.86187154,
final_local_scores=-0.92437273, pem=-1.5709716,
similarity=-14.952227 ]
10: 27854563 New_South_Wales_state_election,_1953
    [winner: 34865 1953, pem_pos=0, final_scores=-0.40257916,
final_local_scores=-0.20235156, pem=-0.7062917,
similarity=1.8920629|
    RECALL MISS: 27854563
New_South_Wales_state_election,_1953]
11: 26010113 New_South_Wales_state_election,_1932
    [winner: 34775 1932, pem_pos=0, final_scores=-0.8990925,
final_local_scores=-0.78856105, pem=-0.8542985,
similarity=-14.796715|
    RECALL MISS: 26010113
New_South_Wales_state_election,_1932]
12: 26614337 New_South_Wales_state_election,_1941
    [winner: 34632 1941, pem_pos=0, final_scores=-1.0391295,
final_local_scores=-0.9342827, pem=-0.6047071,
similarity=-20.502655|
    RECALL MISS: 26614337
New_South_Wales_state_election,_1941]
1: 579433 New_South_Wales_Legislative_Assembly
    [pem_pos=0, final_scores=0.28556192,
final_local_scores=0.24783677, pem=0.0, similarity=11.485657]
2: 21927 National_Party_of_Australia
    [pem_pos=0, final_scores=0.34481472,
final_local_scores=0.2977932, pem=0.0, similarity=12.976457]
3: 613897 Uralla,_New_South_Wales
    [pem_pos=0, final_scores=0.1950228,
final_local_scores=0.23595798, pem=0.0, similarity=11.131169]
4: 18056083 New_England_New_State_Movement
    [pem_pos=0, final_scores=0.24916159,

```

final_local_scores=0.25519165, pem=0.0, similarity=11.705141]
5: 10357 Earle_Page
[pem_pos=0, final_scores=0.31078154,
final_local_scores=0.29852837, pem=0.0, similarity=12.998395]
6: 2012063 Aboriginal_Protection_Board
[pem_pos=0, final_scores=0.1740294,
final_local_scores=0.20616049, pem=0.0, similarity=10.241953]
7: 415475 First_Australian_Imperial_Force
[pem_pos=0, final_scores=0.17823349,
final_local_scores=0.1977512, pem=0.0, similarity=9.991003]
8: 2770934 Bertram_Stevens_(politician)
[pem_pos=0, final_scores=0.27985138,
final_local_scores=0.26238945, pem=0.0, similarity=11.919937]
9: 2770889 Alexander_Mair
[pem_pos=0, final_scores=0.14878403,
final_local_scores=0.13236395, pem=0.0, similarity=8.039717]
1: 4689264 Australia
[pem_pos=0, final_scores=0.5212974,
final_local_scores=0.54600734, pem=-0.491009,
similarity=23.053844]
2: 632836 Glen_Innes,_New_South_Wales
[pem_pos=0, final_scores=0.41493016,
final_local_scores=0.4531927, pem=-0.26889536,
similarity=19.076181]
3: 21654 New_South_Wales
[pem_pos=0, final_scores=0.6484396,
final_local_scores=0.661062, pem=-0.051332146,
similarity=24.096273]
4: 4764461 World_War_I
[pem_pos=0, final_scores=0.37982225,
final_local_scores=0.42200792, pem=0.0, similarity=16.683271]
5: 5843419 France
[pem_pos=0, final_scores=0.590016,
final_local_scores=0.6909636, pem=-0.16240834,
similarity=25.592648]
6: 27716366 Roy_Vincent
[pem_pos=0, final_scores=0.0376537,
final_local_scores=0.035695598, pem=0.0, similarity=5.1549406]
7: 57845 Raleigh,_North_Carolina
[pem_pos=0, final_scores=0.2020383,
final_local_scores=0.2479524, pem=-0.35659322,
similarity=13.428316]
8: 21927 National_Party_of_Australia
[pem_pos=0, final_scores=0.38495559,
final_local_scores=0.34164104, pem=0.0, similarity=14.284963]
9: 21654 New_South_Wales
[pem_pos=0, final_scores=0.48465168,
final_local_scores=0.4821486, pem=-0.051332146,
similarity=18.757141]
10: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.3390649,
final_local_scores=0.29226258, pem=-0.2258703,

```

similarity=14.039729]
11: 27716366 Roy_Vincent
    [pem_pos=0, final_scores=0.036154702,
final_local_scores=0.03405817, pem=0.0, similarity=5.1060767]
12: 2770934 Bertram_Stevens_(politician)
    [pem_pos=0, final_scores=0.1841002,
final_local_scores=0.15779588, pem=0.0, similarity=8.798656]

```

Sample: b'Urla_Clashes'

Urla Clashes were a series of clashes in -LRB- now a district center of , Turkey -RRB- during the . After the , the advanced and occupied [1 1 Bornova 1] and [2 2 Karşıyaka 2] . After that it advanced to the 1 Karaburun Peninsula where [1 Urla 1] was situated . According to the population statistics from 1914 , 9,321 Turks and 24,711 [2 Greeks 2] were living in [3 Urla 3] . Incited by the [4 Greek 4] army 's presence , about 800 armed local started to besiege and attack Turkish villages in . The villages of Kuşçular , Kızılcaköy , Devederesi were looted and burnt down . Additionally , a [5 Greek 5] 2 Evzones 3 company , disembarking a 4 torpedo boat , went ashore in and started to occupy . At this time , the 173 . infantry 5 regiment of the [6 Ottoman 6] 56th 6 division was stationed in . The commander of the regiment was 7 Lieutenant colonel [7 Kazım Bey 7] . The regiment consisted of 100 men , but most of them were scattered around . [8 Kazım Bey 8] had only 25 soldiers and gendarmeries under his command when the attacks commenced . With these 25 men , he managed to repel the irregular [9 Greek 9] attacks . As soon as the local 8 Turkish population heard about these incidents , they entered the arms depot of the town and took the 9 rifle s and ammunition . With these rifles and ammunition they established the first [10 10 Kuva-yi Milliye 10] unit in 11 western [11 Anatolia 11] , consisting of 120 men . This militia unit joined [12 Kazım Bey 12] 's forces . The clashes lasted from 16 -- 17 May . A [13 Greek 13] battleship docked at the [14 Urla 14] pier and unloaded more [15 Greek 15] troops . The arrival of [16 Greek 16] troops , from [17 Smyrna 17] , accompanied by a [18 12 British 18] officer , ended the clashes at the evening of 17 May . The Turkish soldiers and gendarmeries were disarmed and arrested . Most of the 120 men strong 13 militia unit escaped to the interior regions . These clashes in marked the first armed resistance of Turkish irregular and regular units

1: 4171109 Bornova

2: 2054440 Karşıyaka

1: 11388985 Karaburun_Peninsula,_Turkey

```

    [winner: 11388985 Karaburun_Peninsula,_Turkey, pem_pos=0,
final_scores=-0.025507763, final_local_scores=0.0280658,
pem=-0.36693144, similarity=6.922683]

```

```

    gt: 11388985 Karaburun_Peninsula,_Turkey, pem_pos=0,
final_scores=-0.025507763, final_local_scores=0.0280658,
pem=-0.36693144, similarity=6.922683 ]

```

2: 2152244 Evzones
[winner: 2152244 Evzones, pem_pos=0,
final_scores=-0.31567648, final_local_scores=-0.2933169,
pem=0.0, similarity=-4.663449]
gt: 2152244 Evzones, pem_pos=0,
final_scores=-0.31567648, final_local_scores=-0.2933169,
pem=0.0, similarity=-4.663449]
3: 145418 Company_(military_unit)
[winner: 10929133 OAO_TMK, pem_pos=14,
final_scores=-1.7968167, final_local_scores=-1.8090487,
pem=-3.3672957, similarity=-31.584028]
gt: 145418 Company_(military_unit), pem_pos=1,
final_scores=-2.2179756, final_local_scores=-2.3911202,
pem=-2.0325165, similarity=-56.21294]
4: 31446 Torpedo_boat
[winner: 2364859 Mark_13_torpedo, pem_pos=0,
final_scores=-1.0792583, final_local_scores=-1.0392592,
pem=0.0, similarity=-26.923855]
RECALL MISS: 31446 Torpedo_boat]
[winner: 2274122 United_States_R-class_submarine,
pem_pos=4, final_scores=-1.1089547,
final_local_scores=-1.066143, pem=-1.3862944,
similarity=-20.187235]
RECALL MISS: 31446 Torpedo_boat]
5: 166653 Regiment
[winner: 3612822 Long's_Regiment, pem_pos=4,
final_scores=-1.2615298, final_local_scores=-1.208365,
pem=-1.9451132, similarity=-21.39248]
gt: 166653 Regiment, pem_pos=0, final_scores=-1.3954805,
final_local_scores=-1.4825737, pem=-0.004108632,
similarity=-40.1309]
6: 145415 Division_(military)
[winner: 7135137 Divisions_of_Malaysia, pem_pos=16,
final_scores=-1.4657624, final_local_scores=-1.4219973,
pem=-3.5447905, similarity=-19.0684]
gt: 145415 Division_(military), pem_pos=0,
final_scores=-1.7545625, final_local_scores=-1.8950055,
pem=-0.80402994, similarity=-48.088593]
7: 33576304 Lieutenant_colonel
[winner: 161405 Lieutenant, pem_pos=0,
final_scores=-0.40888613, final_local_scores=-0.38612917,
pem=-0.16377881, similarity=-6.542497]
gt: 33576304 Lieutenant_colonel, pem_pos=16,
final_scores=-1.1420883, final_local_scores=-1.18926,
pem=-4.2034535, similarity=-8.541142]
[winner: 7171000 Colonel_(Pakistan), pem_pos=8,
final_scores=-1.3696313, final_local_scores=-1.3356752,
pem=-3.478587, similarity=-16.852404]
RECALL MISS: 33576304 Lieutenant_colonel]
8: 2088822 Turkish_people
[winner: 29992 Turkish_language, pem_pos=0,
final_scores=-0.14085454, final_local_scores=-0.17780894,

pem=-0.7274845, similarity=2.7397134|
gt: 2088822 Turkish_people, pem_pos=2,
final_scores=-0.20266603, final_local_scores=-0.2751677,
pem=-1.8861854, similarity=6.135529]
9: 25456 Rifle
[winner: 25107948 The_Rifle, pem_pos=0,
final_scores=-0.82341415, final_local_scores=-0.7297568,
pem=0.0, similarity=-17.687685|
RECALL MISS: 25456 Rifle]
10: 28397257 Kuva-yi_Milliye
[winner: 3950041 Turkish_National_Movement, pem_pos=0,
final_scores=0.31974524, final_local_scores=0.34022027,
pem=-0.4439314, similarity=16.656733|
gt: 28397257 Kuva-yi_Milliye, pem_pos=1,
final_scores=0.1261702, final_local_scores=0.18870968,
pem=-1.0258529, similarity=15.299934]
11: 854 Anatolia
[winner: 854 Anatolia, pem_pos=0,
final_scores=0.41331542, final_local_scores=0.387304,
pem=0.0, similarity=15.647637|
gt: 854 Anatolia, pem_pos=0, final_scores=0.41331542,
final_local_scores=0.387304, pem=0.0, similarity=15.647637]
[winner: 65758 Western_(genre), pem_pos=1,
final_scores=-1.8394524, final_local_scores=-1.8774548,
pem=-1.6856036, similarity=-42.770706|
RECALL MISS: 854 Anatolia]
12: 19097669 British_people
[winner: 31717 United_Kingdom, pem_pos=0,
final_scores=0.45925128, final_local_scores=0.5015658,
pem=-0.48936236, similarity=21.718666|
gt: 19097669 British_people, pem_pos=1,
final_scores=-0.46611732, final_local_scores=-0.45191613,
pem=-2.2110078, similarity=2.6274385]
13: 20622 Militia
[winner: 2158227 Milicja_Obywatelska, pem_pos=7,
final_scores=-0.82941866, final_local_scores=-0.8206499,
pem=-2.396894, similarity=-7.365435|
gt: 20622 Militia, pem_pos=0, final_scores=-0.8493707,
final_local_scores=-0.90177625, pem=-0.13721985,
similarity=-22.074865]
1: 4171109 Bornova
[pem_pos=0, final_scores=0.26241714,
final_local_scores=0.30557477, pem=0.0, similarity=13.2086735]
2: 2054440 Karşıyaka
[pem_pos=0, final_scores=0.19876184,
final_local_scores=0.21598202, pem=-0.5268665,
similarity=13.400229]
1: 4116840 Urla, İzmir
[pem_pos=0, final_scores=0.24327014,
final_local_scores=0.23195341, pem=0.0, similarity=11.011665]
2: 42056 Greeks
[pem_pos=0, final_scores=0.12702344,

final_local_scores=0.08386682, pem=-0.48669693,
similarity=9.2392]
3: 4116840 Urla, İzmir
[pem_pos=0, final_scores=0.26652485,
final_local_scores=0.25735563, pem=0.0, similarity=11.769718]
4: 12108 Greece
[pem_pos=1, final_scores=0.8858836,
final_local_scores=0.8853572, pem=-2.3247876,
similarity=43.15309]
5: 12108 Greece
[pem_pos=1, final_scores=0.7404439,
final_local_scores=0.72648644, pem=-2.3247876,
similarity=38.412075]
6: 22278 Ottoman_Empire
[pem_pos=0, final_scores=0.29193902,
final_local_scores=0.24807376, pem=-0.2598744,
similarity=12.905966]
7: 3551768 Kâzım_Karabekir
[pem_pos=0, final_scores=0.20599024,
final_local_scores=0.2241795, pem=0.0, similarity=10.779676]
8: 3551768 Kâzım_Karabekir
[pem_pos=0, final_scores=0.27813464,
final_local_scores=0.3029863, pem=0.0, similarity=13.131429]
9: 12108 Greece
[pem_pos=1, final_scores=0.80193025,
final_local_scores=0.79365087, pem=-2.3247876,
similarity=40.416393]
10: 3950041 Turkish_National_Movement
[pem_pos=0, final_scores=0.31974524,
final_local_scores=0.34022027, pem=-0.4439314,
similarity=16.656733]
11: 854 Anatolia
[pem_pos=0, final_scores=0.41331542,
final_local_scores=0.387304, pem=0.0, similarity=15.647637]
12: 3551768 Kâzım_Karabekir
[pem_pos=0, final_scores=0.21281023,
final_local_scores=0.23162931, pem=0.0, similarity=11.001993]
13: 12108 Greece
[pem_pos=1, final_scores=0.8231295,
final_local_scores=0.816808, pem=-2.3247876,
similarity=41.10745]
14: 4116840 Urla, İzmir
[pem_pos=0, final_scores=0.16686456,
final_local_scores=0.14849192, pem=0.0, similarity=8.521008]
15: 12108 Greece
[pem_pos=1, final_scores=0.81311804,
final_local_scores=0.80587184, pem=-2.3247876,
similarity=40.781094]
16: 12108 Greece
[pem_pos=1, final_scores=0.9013743,
final_local_scores=0.9022785, pem=-2.3247876,
similarity=43.65806]

17: 50562 Smyrna
[pem_pos=0, final_scores=0.21945854,
final_local_scores=0.19470993, pem=-0.4654988,
similarity=12.431701]
18: 31717 United_Kingdom
[pem_pos=0, final_scores=0.45925128,
final_local_scores=0.5015658, pem=-0.48936236,
similarity=21.718666]

Sample: b'Ya'

`akov_Riftin [1 Ya'akov Riftin 1] -LRB- , born 16 March 1907 , died 14 May 1978 -RRB- was an [1 1 Israel 1] i politician who served as a member of the [2 2 Knesset 2] for [3 3 Mapam 3] between 1949 and 1965 . Born in near [4 4 Puławy 4] in the [5 5 Russian Empire 5] -LRB- today in [6 6 Poland 6] -RRB- , [2 Riftin 2] joined the [3 HaShomer 3] youth movement , and was later a member of [7 7 Hashomer Hatzair 7] , becoming one of the leaders of the movement in [4 Poland 4] . He made 1 aliyah to [8 8 Mandatory Palestine 8] in 1929 , and joined kibbutz [9 9 Ein Shemer 9] in 1931 . Riftin also joined the [10 10 Histadrut 10] trade union , becoming a member of its executive committee and co-ordinaing committee . He served as a member of the and of the [11 11 Jewish National Council 11] , as well as being a member of the [12 12 Yishuv 12] 's security committee . During 1947 and 1948 he was part of the [5 Jewish 5] delegation to the [13 13 United Nations 13] , and served as the political secretary of the [6 Israeli 6] delegation from independence -LRB- in 1948 -RRB- until 1954 . Despite being part of the [7 UN 7] delegation , he was elected to the [8 Knesset 8] in 2 1949 , and was re-elected in 3 1951 , 4 1955 , 5 1959 and 6 1961 . He lost his seat in the 7 1965 elections . A member of the left-wing group in [9 Mapam 9] , he left the party in 1968 when it formed 8 an alliance with the [14 14 Labor Party 14]

1: 9282173 Israel
2: 74656 Knesset
3: 41585618 Mapam
4: 386889 Puławy
5: 20611504 Russian_Empire
6: 22936 Poland
7: 40949046 Hashomer_Hatzair
8: 38584582 Mandatory_Palestine
9: 18676409 Ein_Shemer
10: 648340 Histadrut
11: 2524643 Jewish_National_Council
12: 430478 Yishuv
13: 31769 United_Nations
14: 40312187 Israeli_Labor_Party

1: 351513 Aliyah
[winner: 12058220 Yisrael_HaMithadeshet, pem_pos=1,
final_scores=-0.8180235, final_local_scores=-0.8312822,
pem=-1.0677794, similarity=-14.910656]

gt: 351513 Aliyah, pem_pos=0, final_scores=-1.2579844,
final_local_scores=-1.4494019, pem=-0.038090024,
similarity=-38.956192]
2: 7950885 Israeli_legislative_election, 1949
[winner: 34604 1949, pem_pos=0, final_scores=-0.90890723,
final_local_scores=-0.80731875, pem=-0.6436868,
similarity=-16.50182]
RECALL MISS: 7950885 Israeli_legislative_election, 1949]
3: 7966801 Israeli_legislative_election, 1951
[winner: 34602 1951, pem_pos=0, final_scores=-0.55702716,
final_local_scores=-0.3459177, pem=-0.7317484,
similarity=-2.2538]
RECALL MISS: 7966801 Israeli_legislative_election, 1951]
4: 7984520 Israeli_legislative_election, 1955
[winner: 34851 1955, pem_pos=0, final_scores=-0.38350946,
final_local_scores=-0.11762993, pem=-0.7263205,
similarity=4.529245]
RECALL MISS: 7984520 Israeli_legislative_election, 1955]
5: 8084868 Israeli_legislative_election, 1959
[winner: 34662 1959, pem_pos=0, final_scores=-0.5803796,
final_local_scores=-0.37806624, pem=-0.6961271,
similarity=-3.40689]
RECALL MISS: 8084868 Israeli_legislative_election, 1959]
6: 8330471 Israeli_legislative_election, 1961
[winner: 34659 1961, pem_pos=0, final_scores=-0.59990805,
final_local_scores=-0.4438952, pem=-0.7041954,
similarity=-5.3274813]
RECALL MISS: 8330471 Israeli_legislative_election, 1961]
7: 8454002 Israeli_legislative_election, 1965
[winner: 8454002 Israeli_legislative_election, 1965,
pem_pos=0, final_scores=-0.50928617,
final_local_scores=-0.5633562, pem=-0.10311921,
similarity=-12.161182]
gt: 8454002 Israeli_legislative_election, 1965,
pem_pos=0, final_scores=-0.50928617,
final_local_scores=-0.5633562, pem=-0.10311921,
similarity=-12.161182]
8: 9549928 Alignment_(political_party)
[winner: 15274590 Alliance, Indiana, pem_pos=12,
final_scores=-1.1450328, final_local_scores=-1.0392823,
pem=-3.4751482, similarity=-8.02615]
RECALL MISS: 9549928 Alignment_(political_party)]
1: 9282173 Israel
[pem_pos=0, final_scores=0.5282615,
final_local_scores=0.50035506, pem=-0.21279564,
similarity=20.178522]
2: 74656 Knesset
[pem_pos=0, final_scores=0.2640428,
final_local_scores=0.20956263, pem=0.0, similarity=10.34348]
3: 41585618 Mapam
[pem_pos=0, final_scores=0.19996674,
final_local_scores=0.16553548, pem=0.0, similarity=9.029622]

4: 386889 Puławy
[pem_pos=0, final_scores=0.31740886,
final_local_scores=0.3830042, pem=-0.23003705,
similarity=16.770298]
5: 20611504 Russian_Empire
[pem_pos=0, final_scores=0.19426928,
final_local_scores=0.191286, pem=0.0, similarity=9.798069]
6: 22936 Poland
[pem_pos=0, final_scores=0.4340108,
final_local_scores=0.4461357, pem=-0.15333635,
similarity=18.237158]
7: 40949046 Hashomer_Hatzair
[pem_pos=0, final_scores=0.22564076,
final_local_scores=0.19213143, pem=0.0, similarity=9.823298]
8: 38584582 Mandatory_Palestine
[pem_pos=0, final_scores=0.19203736,
final_local_scores=0.13230929, pem=0.0, similarity=8.038086]
9: 18676409 Ein_Shemer
[pem_pos=0, final_scores=0.122549966,
final_local_scores=0.13043568, pem=0.0, similarity=7.982174]
10: 648340 Histadrut
[pem_pos=0, final_scores=0.39527792,
final_local_scores=0.37938863, pem=0.0, similarity=15.411427]
11: 2524643 Jewish_National_Council
[pem_pos=0, final_scores=0.108211055,
final_local_scores=0.09098078, pem=0.0, similarity=6.804761]
12: 430478 Yishuv
[pem_pos=0, final_scores=0.28269404,
final_local_scores=0.25358114, pem=0.0, similarity=11.65708]
13: 31769 United_Nations
[pem_pos=0, final_scores=0.31708395,
final_local_scores=0.3368902, pem=0.0, similarity=14.143188]
14: 40312187 Israeli_Labor_Party
[pem_pos=1, final_scores=0.19835262,
final_local_scores=0.16876215, pem=-1.2490158,
similarity=15.918254]
1: 23467801 Ya'akov_Riftin
[pem_pos=0, final_scores=0.080051586,
final_local_scores=0.098308206, pem=0.0, similarity=7.023426]
2: 23467801 Ya'akov_Riftin
[pem_pos=0, final_scores=0.06476836,
final_local_scores=0.081613585, pem=0.0, similarity=6.525225]
3: 1833605 Hashomer
[pem_pos=0, final_scores=0.11704348,
final_local_scores=0.12692639, pem=0.0, similarity=7.87745]
4: 22936 Poland
[pem_pos=0, final_scores=0.44831032,
final_local_scores=0.4617557, pem=-0.15333635,
similarity=18.70329]
5: 25955086 Jews
[pem_pos=1, final_scores=0.24904846,
final_local_scores=0.23076105, pem=-0.80476934,

similarity=15.352543]
 6: 9282173 Israel
 [pem_pos=0, final_scores=0.54741925,
 final_local_scores=0.521282, pem=-0.34790626,
 similarity=21.537775]
 7: 31769 United_Nations
 [pem_pos=0, final_scores=0.05345504,
 final_local_scores=0.048915654, pem=-0.06880474,
 similarity=5.9236245]
 8: 74656 Knesset
 [pem_pos=0, final_scores=0.23941712,
 final_local_scores=0.18402067, pem=0.0, similarity=9.581257]
 9: 41585618 Mapam
 [pem_pos=0, final_scores=0.30809945,
 final_local_scores=0.28365406, pem=0.0, similarity=12.554516]

Sample: b'Shaka_Smart'

[1 Shaka Smart 1] -LRB- born April 8 , 1977 -RRB- is an [2 American 2] 1 college basketball coach , and the current head coach of the [1 1 Virginia Commonwealth University 1] . [3 Smart 3] has never won a conference championship . Before accepting the job at [4 VCU 4] , Smart served as an assistant coach for the [5 2 Akron Zips 5] , 3 Clemson Tigers , and [6 4 Florida Gators 6] men 's basketball programs . In high school , [7 Smart 7] was a three-year starter for 5 Oregon High School in [8 Wisconsin 8] . He was a second-team All-Badger Conference pick as a senior and by the end of his career was the all-time assists leader at Oregon for a career -LRB- 458 -RRB- , season -LRB- 201 -RRB- and single game -LRB- 20 -RRB- . After graduating from high school , [9 Smart 9] attended [2 2 Kenyon College 2] , a liberal arts school in [10 Ohio 10] . On the [11 ESPN 11] program 6 Pardon the Interruption , [12 Smart 12] credited his personal relationship with then head coach Bill Brown as the reason for his decision . As a member of the [13 Kenyon College 13] basketball team , he was an all-conference selection as a senior and is the school 's career assists leader -LRB- 542 -RRB- . [14 Smart 14] was named a member of the 1999 [15 USA 15] TODAY All-USA Academic Team . == Assistant coach == Smart began his coaching career in 1999 as an assistant at [3 3 California University of Pennsylvania 3] , where he also earned a Master 's Degree . Afterwards , he was hired as Director of Basketball Operations at the [4 4 University of Dayton 4] . He was then an assistant at the [5 5 University of Akron 5] for three years , [16 Clemson 16] for two , and Florida for one . VCU hired [17 Smart 17] to be the head coach in the spring of 2009 after the program 's previous coach , [18 7 Anthony Grant 18] , left to become the head coach of the [19 8 Alabama Crimson Tide 19] men 's basketball team . [20 Smart 20] 's hire made him the 10th-youngest head coach in [21 Division I 21] . In 9 his first season , he led the [22 Rams 22] to a 27 -- 10 season and a 10 CBI Championship after VCU swept 11 S

aint Louis in the championship best-of-three series . [23 Smart 23] 's second season began with forward 12 Larry Sanders declaring for the 2010 [24 NBA Draft 24] after his junior season . Sanders ' selection by the [6 6 Milwaukee Bucks 6] with the 15th pick made VCU the first school in the [7 7 Commonwealth of Virginia 7] to have a player selected in the first round of the [25 NBA Draft 25] in consecutive years , as the [8 8 Utah Jazz 8] had selected guard [9 9 Eric Maynor 9] with the 20th pick in the 2009 [26 NBA Draft 26] . The [27 Rams 27] under [28 Smart 28] went 23 -- 11 in the 2010 -- 2011 season . [29 Smart 29] led the [30 Rams 30] to their second consecutive 13 Colonial Athletic Association Championship Game , where they lost to [31 14 Old Dominion 31] . Despite not securing the automatic bid , [32 VCU 32] earned an at-large bid to the Southwest region of the 15 2011 [33 NC AA Tournament 33] , but were placed in the `` First Four `` against [34 16 USC 34] for a spot in the main 64-team tournament bracket . The [35 Rams 35] ' selection into the tournament was widely criticized . His coaching strategy , positive outlook , and patience helped VCU defeat [36 USC 36] in the First Four , and then upset the 6th-seeded [37 17 Georgetown Hoyas 37] and 3rd-seeded [38 18 Purdue Boilermakers 38] to advance to the 19 Sweet 16 in the [39 NCAA Tournament 39] for the first time in school history . [40 VCU 40] won 72 -- 71 against [41 20 Florida State University 41] in overtime to earn the school 's first spot in the 21 Elite Eight and subsequently upset the top-seeded [42 22 University of Kansas 42] 71 -- 61 for its first 23 Final Four appearance . But in the Final Four , the [43 Rams 43] lost to 24 Butler 70-62 . On April 4 , 2011 , [44 Smart 44] agreed to an 8-year contract extension with [45 VCU 45] , increasing his base pay from \$ 350,000 to \$ 1.2 million per year , prior to any performance bonuses . On January 19 , 2013 , [46 Shaka Smart 46] became the second youngest coach to win 100 games , in a 90-63 victory over Duquesne . [47 Smart 47] 's teams play an upbeat style of basketball known as `` havoc . `` [48 Smart 48] described his `` havoc `` defensive philosophy during his introductory press conference as , `` We are going to wreak havoc on our opponent 's psyche and their plan of attack . `` On the court , the `` havoc `` defensive mindset is visible through the heavy use of the and pressing after made baskets to disrupt opponents ' timing of offensive sets . In each of [49 Smart 49] 's three seasons , the [50 Rams 50] have consistently maintained one of the most efficient and disruptive defenses in the nation , ranking 31st in Steals Per Game and 17th in Turnover Margin his first season and 1st nationally in both categories his second and third seasons as head coach . Offensively , Smart-coached teams play uptempo , push the ball after misses , and have with what [51 Smart 51] describes as `` the freedom to make plays in the open court . `` [52 Smart 52] 's given name was bestowed in honor of 25 the famous Zulu warrior . He graduated magna cum laude from

Kenyon with a degree in History , researching and writing on issues related to race and the [10 10 Great Migration 10] his junior and senior years . [53 Smart 53] received an [54 NCAA 5 4] postgraduate scholarship and earned a master 's degree in social science at [11 11 California University of Pennsylvania 11] . He developed a love for quotations at Kenyon . [55 Smart 55] began writing down quotes into a digital document that is now over 110 pages long . He also likes nature documentaries featuring big cats . [56 Smart 56] has been married to Maya Payne , a professional writer and an alumna of both [57 Harvard University 57] and [12 12 Northwestern University 12] , since 2006 . Their child , Zora Sanae Smart , was born on September 25 , 2011 . [58 Smart 58] 's half-brother is the writer and professor , J.M. Tyree . He is a supporter of President [13 13 Barack Obama 13]

1: 172675 Virginia_Commonwealth_University

2: 185536 Kenyon_College

3: 1893665 California_University_of_Pennsylvania

4: 454830 University_of_Dayton

5: 597542 University_of_Akron

6: 72878 Milwaukee_Bucks

7: 32432 Virginia

8: 72890 Utah_Jazz

9: 10083159 Eric_Maynor

10: 375540 Great_Migration_(African_American)

11: 1893665 California_University_of_Pennsylvania

12: 174216 Northwestern_University

13: 534366 Barack_Obama

1: 768856 College_basketball

[winner: 10088476 Sinbad_(comedian), pem_pos=0, final_scores=-0.7978517, final_local_scores=-0.77802676, pem=0.0, similarity=-19.128157]

RECALL MISS: 768856 College_basketball]

[winner: 21634616 Old_Academy_(Munich), pem_pos=15, final_scores=-1.4554911, final_local_scores=-1.4339147, pem=-3.5136774, similarity=-19.593233]

RECALL MISS: 768856 College_basketball]

2: 9429001 Akron_Zips_men's_basketball

[winner: 9152167 Akron_Zips, pem_pos=0, final_scores=0.04015191, final_local_scores=0.03098537, pem=-0.1391739, similarity=5.7712274]

gt: 9429001 Akron_Zips_men's_basketball, pem_pos=5, final_scores=-0.4768905, final_local_scores=-0.53277284, pem=-3.0242686, similarity=4.6371517]

3: 14503740 Clemson_Tigers_men's_basketball

[winner: 6698897 Clemson_Tigers, pem_pos=0, final_scores=-0.018118367, final_local_scores=-0.05338891, pem=-0.5927408, similarity=5.7198977]

gt: 14503740 Clemson_Tigers_men's_basketball, pem_pos=13, final_scores=-0.46660268, final_local_scores=-0.5120447, pem=-3.166493, similarity=6.0291576]

4: 10244706 Florida_Gators_men's_basketball
 [winner: 965005 Florida_Gators, pem_pos=0,
 final_scores=0.058202967, final_local_scores=0.022431254,
 pem=-0.33751175, similarity=6.5945477]
 gt: 10244706 Florida_Gators_men's_basketball, pem_pos=2,
 final_scores=-0.27853334, final_local_scores=-0.35563797,
 pem=-2.4580379, similarity=6.843958]

5: 14059667 Oregon_High_School_(Wisconsin)
 [winner: 14059667 Oregon_High_School_(Wisconsin),
 pem_pos=0, final_scores=-0.18022756,
 final_local_scores=-0.105753735, pem=0.0,
 similarity=0.9338105]
 gt: 14059667 Oregon_High_School_(Wisconsin), pem_pos=0,
 final_scores=-0.18022756, final_local_scores=-0.105753735,
 pem=0.0, similarity=0.9338105]

6: 424637 Pardon_the_Interruption
 [winner: 424637 Pardon_the_Interruption, pem_pos=0,
 final_scores=-0.13049693, final_local_scores=-0.086074024,
 pem=0.0, similarity=1.5210924]
 gt: 424637 Pardon_the_Interruption, pem_pos=0,
 final_scores=-0.13049693, final_local_scores=-0.086074024,
 pem=0.0, similarity=1.5210924]

7: 1904584 Anthony_Grant
 [winner: 8434251 Anthony_Grant_(footballer), pem_pos=0,
 final_scores=0.060859725, final_local_scores=0.14981046,
 pem=-0.17651322, similarity=9.520263]
 gt: 1904584 Anthony_Grant, pem_pos=1,
 final_scores=-0.030583143, final_local_scores=-0.049530566,
 pem=-0.51840556, similarity=5.430792]

8: 9339018 Alabama_Crimson_Tide_men's_basketball
 [winner: 7751357 Alabama_Crimson_Tide_football,
 pem_pos=1, final_scores=0.17835854,
 final_local_scores=0.13694632, pem=-0.8202876,
 similarity=12.637317]
 gt: 9339018 Alabama_Crimson_Tide_men's_basketball,
 pem_pos=17, final_scores=-0.44264412,
 final_local_scores=-0.5060322, pem=-3.893674,
 similarity=10.163106]
 [winner: 408442 Rick_Pitino, pem_pos=0,
 final_scores=-1.0762868, final_local_scores=-1.21176,
 pem=0.0, similarity=-32.07162]
 RECALL MISS: 9339018
 Alabama_Crimson_Tide_men's_basketball]
 [winner: 7809948 Men's_colleges_in_the_United_States,
 pem_pos=9, final_scores=-1.1569536,
 final_local_scores=-1.172895, pem=-2.5649493,
 similarity=-16.963217]
 RECALL MISS: 9339018
 Alabama_Crimson_Tide_men's_basketball]
 [winner: 1415027 David_Berman_(musician), pem_pos=5,
 final_scores=-1.987177, final_local_scores=-2.0757704,
 pem=-3.0012724, similarity=-41.534027]

RECALL MISS: 9339018
 Alabama_Crimson_Tide_men's_basketball]
 9: 26726493 2009-10_VCU_Rams_men's_basketball_team
 [winner: 1098181 Ray_Guy, pem_pos=0,
 final_scores=-0.9086849, final_local_scores=-0.96418464,
 pem=0.0, similarity=-24.68348]
 RECALL MISS: 26726493 2009-
 10_VCU_Rams_men's_basketball_team]
 [winner: 690345 Season_(sports), pem_pos=0,
 final_scores=-1.5559975, final_local_scores=-1.685772,
 pem=-1.219945, similarity=-39.582832]
 RECALL MISS: 26726493 2009-
 10_VCU_Rams_men's_basketball_team]
 10: 26556818 2010_College_Basketball_Invitational
 [winner: 26556818 2010_College_Basketball_Invitational,
 pem_pos=0, final_scores=-0.13589405,
 final_local_scores=-0.12069951, pem=0.0,
 similarity=0.48779872]
 gt: 26556818 2010_College_Basketball_Invitational,
 pem_pos=0, final_scores=-0.13589405,
 final_local_scores=-0.12069951, pem=0.0,
 similarity=0.48779872]
 11: 26251363 2009-
 10_Saint_Louis_Billikens_men's_basketball_team
 [winner: 23408886 St._Louis_Rams, pem_pos=15,
 final_scores=-0.23697676, final_local_scores=-0.30739146,
 pem=-3.3747854, similarity=13.269147]
 RECALL MISS: 26251363 2009-
 10_Saint_Louis_Billikens_men's_basketball_team]
 12: 21758578 Larry_Sanders_(basketball)
 [winner: 21758578 Larry_Sanders_(basketball), pem_pos=1,
 final_scores=0.014058456, final_local_scores=0.02493444,
 pem=-0.22314353, similarity=6.0472956]
 gt: 21758578 Larry_Sanders_(basketball), pem_pos=1,
 final_scores=0.014058456, final_local_scores=0.02493444,
 pem=-0.22314353, similarity=6.0472956]
 13: 552381 Colonial_Athletic_Association
 [winner: 552381 Colonial_Athletic_Association, pem_pos=0,
 final_scores=-0.13519318, final_local_scores=-0.17187828,
 pem=0.0, similarity=-1.0394777]
 gt: 552381 Colonial_Athletic_Association, pem_pos=0,
 final_scores=-0.13519318, final_local_scores=-0.17187828,
 pem=0.0, similarity=-1.0394777]
 14: 31074745 2010-11_Old_Dominion_Monarchs_basketball_team
 [winner: 308123 Old_Dominion_University, pem_pos=0,
 final_scores=0.050003663, final_local_scores=0.0041092336,
 pem=-0.551419, similarity=7.2110434]
 gt: 31074745 2010-
 11_Old_Dominion_Monarchs_basketball_team, pem_pos=11,
 final_scores=-0.71619284, final_local_scores=-0.7085345,
 pem=-3.8836236, similarity=4.0653796]
 15: 5865138 2011_NCAA_Men's_Division_I_Basketball_Tournament

[winner: 507399
NCAA_Men's_Division_I_Basketball_Championship, pem_pos=0,
final_scores=0.16323726, final_local_scores=0.12323338,
pem=-0.54612035, similarity=10.73713|
gt: 5865138
2011_NCAA_Men's_Division_I_Basketball_Tournament, pem_pos=6,
final_scores=-0.34435833, final_local_scores=-0.3945471,
pem=-3.1404083, similarity=9.393668]
[winner: 36225 2011, pem_pos=0, final_scores=-1.3151462,
final_local_scores=-1.238512, pem=-1.4232541,
similarity=-25.13008|
RECALL MISS: 5865138
2011_NCAA_Men's_Division_I_Basketball_Tournament]
16: 31194016 2010-11_USC_Trojans_men's_basketball_team
[winner: 2514216 USC_Trojans_football, pem_pos=1,
final_scores=0.0876468, final_local_scores=0.04173413,
pem=-1.5579058, similarity=13.807277|
RECALL MISS: 31194016 2010-
11_USC_Trojans_men's_basketball_team]
17: 29844159 2010-11_Georgetown_Hoyas_men's_basketball_team
[winner: 3493898 Georgetown_Hoyas, pem_pos=0,
final_scores=0.058681503, final_local_scores=0.026498765,
pem=0.0, similarity=4.880489|
gt: 29844159 2010-
11_Georgetown_Hoyas_men's_basketball_team, pem_pos=14,
final_scores=-1.1062013, final_local_scores=-1.1601832,
pem=-6.5861716, similarity=5.284156]
18: 26724173 2010-11_Purdue_Boilermakers_men's_basketball_team
[winner: 6160171 Purdue_Boilermakers_football, pem_pos=1,
final_scores=0.11820041, final_local_scores=0.07965645,
pem=-1.1085298, similarity=12.495178|
gt: 26724173 2010-
11_Purdue_Boilermakers_men's_basketball_team, pem_pos=12,
final_scores=-0.6323352, final_local_scores=-0.6248323,
pem=-3.9252682, similarity=6.789689]
19: 507399 NCAA_Men's_Division_I_Basketball_Championship
[winner: 507399
NCAA_Men's_Division_I_Basketball_Championship, pem_pos=0,
final_scores=-0.095793486, final_local_scores=-0.16180035,
pem=-0.32055676, similarity=1.0045049|
gt: 507399
NCAA_Men's_Division_I_Basketball_Championship, pem_pos=0,
final_scores=-0.095793486, final_local_scores=-0.16180035,
pem=-0.32055676, similarity=1.0045049]
20: 31090994 2010-
11_Florida_State_Seminoles_men's_basketball_team
[winner: 239297 Florida_State_University, pem_pos=0,
final_scores=0.24068289, final_local_scores=0.20958298,
pem=0.0, similarity=10.344088|
gt: 31090994 2010-
11_Florida_State_Seminoles_men's_basketball_team, pem_pos=8,
final_scores=-1.4900016, final_local_scores=-1.596971,

pem=-9.071078, similarity=5.7628484]
 21: 735272 Elite_Eight
 [winner: 735272 Elite_Eight, pem_pos=0,
 final_scores=0.0064994246, final_local_scores=-0.018422037,
 pem=-0.08622118, similarity=4.0088468]
 gt: 735272 Elite_Eight, pem_pos=0,
 final_scores=0.0064994246, final_local_scores=-0.018422037,
 pem=-0.08622118, similarity=4.0088468]
 22: 26488212 2010-11_Kansas_Jayhawks_men's_basketball_team
 [winner: 163327 University_of_Kansas, pem_pos=0,
 final_scores=0.35823184, final_local_scores=0.3354148,
 pem=0.0, similarity=14.099159]
 gt: 26488212 2010-
 11_Kansas_Jayhawks_men's_basketball_team, pem_pos=7,
 final_scores=-1.4170288, final_local_scores=-1.5209382,
 pem=-9.0887375, similarity=8.127853]
 23: 507399 NCAA_Men's_Division_I_Basketball_Championship
 [winner: 8367801 Final_four, pem_pos=0,
 final_scores=0.011932537, final_local_scores=-0.032035902,
 pem=-0.29987732, similarity=4.7644773]
 gt: 507399
 NCAA_Men's_Division_I_Basketball_Championship, pem_pos=1,
 final_scores=-0.20989116, final_local_scores=-0.28643474,
 pem=-1.503558, similarity=3.7185116]
 24: 28315074 2010-11_Butler_Bulldogs_men's_basketball_team
 [winner: 10817026 Butler_Bulldogs_men's_basketball,
 pem_pos=10, final_scores=-0.5305052,
 final_local_scores=-0.6065087, pem=-3.9706943,
 similarity=7.583538]
 RECALL MISS: 28315074 2010-
 11_Butler_Bulldogs_men's_basketball_team]
 25: 28596 Shaka
 [winner: 61487 Zulu_language, pem_pos=0,
 final_scores=-0.013902292, final_local_scores=0.12422702,
 pem=-0.20988198, similarity=8.938266]
 RECALL MISS: 28596 Shaka]
 [winner: 3362636 Capture_of_HMS_Frolic, pem_pos=0,
 final_scores=-0.9131296, final_local_scores=-0.8270409,
 pem=0.0, similarity=-20.590837]
 RECALL MISS: 28596 Shaka]
 [winner: 10447732
 Results_by_riding_of_the_Canadian_federal_election,_2008,
 pem_pos=12, final_scores=-1.3532788,
 final_local_scores=-1.2237384, pem=-3.6198866,
 similarity=-12.743574]
 RECALL MISS: 28596 Shaka]
 [winner: 12695310 HMS_Warrior_(1781), pem_pos=16,
 final_scores=-1.7245908, final_local_scores=-1.7062867,
 pem=-3.5784001, similarity=-27.369389]
 RECALL MISS: 28596 Shaka]
 1: 172675 Virginia_Commonwealth_University
 [pem_pos=0, final_scores=0.2532475,

final_local_scores=0.22496393, pem=0.0, similarity=10.803085]
2: 185536 Kenyon_College
[pem_pos=0, final_scores=0.27468342,
final_local_scores=0.27098137, pem=0.0, similarity=12.176337]
3: 1893665 California_University_of_Pennsylvania
[pem_pos=0, final_scores=0.26151967,
final_local_scores=0.2554788, pem=0.0, similarity=11.71371]
4: 454830 University_of_Dayton
[pem_pos=0, final_scores=0.38451535,
final_local_scores=0.37993914, pem=0.0, similarity=15.427855]
5: 597542 University_of_Akron
[pem_pos=0, final_scores=0.2959116,
final_local_scores=0.2797497, pem=0.0, similarity=12.438002]
6: 72878 Milwaukee_Bucks
[pem_pos=0, final_scores=0.23741756,
final_local_scores=0.22378916, pem=0.0, similarity=10.768027]
7: 32432 Virginia
[pem_pos=0, final_scores=0.3240372,
final_local_scores=0.3741312, pem=0.0, similarity=15.254535]
8: 72890 Utah_Jazz
[pem_pos=0, final_scores=0.23687468,
final_local_scores=0.22274601, pem=0.0, similarity=10.736898]
9: 10083159 Eric_Maynor
[pem_pos=0, final_scores=0.34824318,
final_local_scores=0.3810035, pem=0.0, similarity=15.459617]
10: 375540 Great_Migration_(African_American)
[pem_pos=0, final_scores=0.09128432,
final_local_scores=0.18664855, pem=-0.014469738,
similarity=9.738367]
11: 1893665 California_University_of_Pennsylvania
[pem_pos=0, final_scores=0.29675466,
final_local_scores=0.29396766, pem=0.0, similarity=12.862294]
12: 174216 Northwestern_University
[pem_pos=0, final_scores=0.31211704,
final_local_scores=0.30971825, pem=0.0, similarity=13.332323]
13: 534366 Barack_Obama
[pem_pos=0, final_scores=0.12740533,
final_local_scores=0.17779681, pem=0.0, similarity=9.395524]
1: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.09512748,
final_local_scores=0.034858406, pem=0.0, similarity=5.129957]
2: 3434750 United_States
[pem_pos=0, final_scores=0.35788125,
final_local_scores=0.412744, pem=-0.12994486,
similarity=17.113478]
3: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.13740842,
final_local_scores=0.081043914, pem=0.0, similarity=6.508225]
4: 172675 Virginia_Commonwealth_University
[pem_pos=0, final_scores=0.102658674,
final_local_scores=0.06046854, pem=-0.09172689,
similarity=6.3930397]

5: 9152167 Akron_Zips
[pem_pos=0, final_scores=0.04015191,
final_local_scores=0.03098537, pem=-0.1391739,
similarity=5.7712274]
6: 965005 Florida_Gators
[pem_pos=0, final_scores=0.058202967,
final_local_scores=0.022431254, pem=-0.33751175,
similarity=6.5945477]
7: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.058643892,
final_local_scores=-0.006709695, pem=0.0,
similarity=3.8894823]
8: 33127 Wisconsin
[pem_pos=0, final_scores=0.216116,
final_local_scores=0.27390862, pem=-0.21268336,
similarity=13.420297]
9: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.11387141,
final_local_scores=0.055333287, pem=0.0, similarity=5.7409687]
10: 22199 Ohio
[pem_pos=0, final_scores=0.53058165,
final_local_scores=0.5901362, pem=-0.19147073,
similarity=22.7418]
11: 77795 ESPN
[pem_pos=0, final_scores=0.14996497,
final_local_scores=0.12472582, pem=-0.09327439,
similarity=8.319021]
12: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.21866427,
final_local_scores=0.16980356, pem=0.0, similarity=9.15699]
13: 185536 Kenyon_College
[pem_pos=0, final_scores=0.13496758,
final_local_scores=0.11836305, pem=0.0, similarity=7.621903]
14: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.16323139,
final_local_scores=0.10925156, pem=0.0, similarity=7.3499975]
15: 3434750 United_States
[pem_pos=0, final_scores=0.19663315,
final_local_scores=0.23660496, pem=-0.13038418,
similarity=11.859526]
16: 147456 Clemson_University
[pem_pos=0, final_scores=0.041418865,
final_local_scores=-0.021190211, pem=-0.85708696,
similarity=8.118327]
17: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.033362404,
final_local_scores=-0.03432586, pem=0.0, similarity=3.0653608]
18: 8434251 Anthony_Grant_(footballer)
[pem_pos=0, final_scores=0.060859725,
final_local_scores=0.14981046, pem=-0.17651322,
similarity=9.520263]
19: 7751357 Alabama_Crimson_Tide_football

[pem_pos=1, final_scores=0.17835854,
final_local_scores=0.13694632, pem=-0.8202876,
similarity=12.637317]
20: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.2161067,
final_local_scores=0.1670098, pem=0.0, similarity=9.073619]
21: 404595 NCAA_Division_I_Football_Championship
[pem_pos=0, final_scores=0.06873573,
final_local_scores=0.03818308, pem=-0.7377454,
similarity=9.241146]
22: 23408886 St._Louis_Rams
[pem_pos=0, final_scores=0.07266025,
final_local_scores=0.03314644, pem=-0.6623493,
similarity=8.680827]
23: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.09807639,
final_local_scores=0.03807962, pem=0.0, similarity=5.2260847]
24: 835946 NBA_draft
[pem_pos=0, final_scores=0.22202794,
final_local_scores=0.20096776, pem=0.0, similarity=10.086991]
25: 835946 NBA_draft
[pem_pos=0, final_scores=0.1839823,
final_local_scores=0.15940869, pem=0.0, similarity=8.8467865]
26: 835946 NBA_draft
[pem_pos=0, final_scores=0.21591054,
final_local_scores=0.19428545, pem=0.0, similarity=9.887578]
27: 23408886 St._Louis_Rams
[pem_pos=0, final_scores=0.09375571,
final_local_scores=0.056190014, pem=-0.6623493,
similarity=9.368493]
28: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.20008616,
final_local_scores=0.14950982, pem=0.0, similarity=8.551384]
29: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.11939897,
final_local_scores=0.061371297, pem=0.0, similarity=5.921155]
30: 23408886 St._Louis_Rams
[pem_pos=0, final_scores=0.16526623,
final_local_scores=0.1343044, pem=-0.6623493,
similarity=11.699582]
31: 308123 Old_Dominion_University
[pem_pos=0, final_scores=0.050003663,
final_local_scores=0.0041092336, pem=-0.551419,
similarity=7.2110434]
32: 172675 Virginia_Commonwealth_University
[pem_pos=0, final_scores=0.06073685,
final_local_scores=0.014696419, pem=-0.09172689,
similarity=5.0271087]
33: 507399 NCAA_Men's_Division_I_Basketball_Championship
[pem_pos=0, final_scores=0.16323726,
final_local_scores=0.12323338, pem=-0.54612035,
similarity=10.73713]

34: 2514216 USC_Trojans_football
[pem_pos=1, final_scores=0.0876468,
final_local_scores=0.04173413, pem=-1.5579058,
similarity=13.807277]
35: 23408886 St._Louis_Rams
[pem_pos=0, final_scores=0.14622961,
final_local_scores=0.113509804, pem=-0.6623493,
similarity=11.07903]
36: 2514216 USC_Trojans_football
[pem_pos=1, final_scores=0.108429596,
final_local_scores=0.06467804, pem=-1.5579058,
similarity=14.491969]
37: 3493898 Georgetown_Hoyas
[pem_pos=0, final_scores=0.058681503,
final_local_scores=0.026498765, pem=0.0, similarity=4.880489]
38: 6160171 Purdue_Boilermakers_football
[pem_pos=1, final_scores=0.11820041,
final_local_scores=0.07965645, pem=-1.1085298,
similarity=12.495178]
39: 507399 NCAA_Men's_Division_I_Basketball_Championship
[pem_pos=0, final_scores=0.19165231,
final_local_scores=0.1542725, pem=-0.54612035,
similarity=11.663399]
40: 172675 Virginia_Commonwealth_University
[pem_pos=0, final_scores=0.079480395,
final_local_scores=0.035170898, pem=-0.09172689,
similarity=5.638108]
41: 239297 Florida_State_University
[pem_pos=0, final_scores=0.24068289,
final_local_scores=0.20958298, pem=0.0, similarity=10.344088]
42: 163327 University_of_Kansas
[pem_pos=0, final_scores=0.35823184,
final_local_scores=0.3354148, pem=0.0, similarity=14.099159]
43: 23408886 St._Louis_Rams
[pem_pos=0, final_scores=0.072230145,
final_local_scores=0.032676607, pem=-0.6623493,
similarity=8.666806]
44: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.10389076,
final_local_scores=0.04443094, pem=0.0, similarity=5.415621]
45: 172675 Virginia_Commonwealth_University
[pem_pos=0, final_scores=0.06539713,
final_local_scores=0.019765958, pem=-0.09172689,
similarity=5.178394]
46: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.39069778,
final_local_scores=0.357724, pem=0.0, similarity=14.764911]
47: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.19705997,
final_local_scores=0.14620414, pem=0.0, similarity=8.452736]
48: 22228744 Shaka_Smart
[pem_pos=0, final_scores=0.3651219,

```

final_local_scores=0.32978627, pem=0.0, similarity=13.931193]
49: 22228744 Shaka_Smart
    [pem_pos=0, final_scores=0.3035376,
final_local_scores=0.26251477, pem=0.0, similarity=11.923677]
50: 23408886 St._Louis_Rams
    [pem_pos=0, final_scores=0.11813773,
final_local_scores=0.082823694, pem=-0.6623493,
similarity=10.163296]
51: 22228744 Shaka_Smart
    [pem_pos=0, final_scores=0.3037179,
final_local_scores=0.26271173, pem=0.0, similarity=11.929555]
52: 22228744 Shaka_Smart
    [pem_pos=0, final_scores=0.119629905,
final_local_scores=0.061623573, pem=0.0, similarity=5.9286833]
53: 22228744 Shaka_Smart
    [pem_pos=0, final_scores=0.25714308,
final_local_scores=0.21183577, pem=0.0, similarity=10.411315]
54: 60706 National_Collegiate_Athletic_Association
    [pem_pos=0, final_scores=0.118383154,
final_local_scores=0.08717722, pem=-0.19246244,
similarity=7.737896]
55: 22228744 Shaka_Smart
    [pem_pos=0, final_scores=0.17304377,
final_local_scores=0.11997011, pem=0.0, similarity=7.669861]
56: 22228744 Shaka_Smart
    [pem_pos=0, final_scores=0.14786719,
final_local_scores=0.092468515, pem=0.0, similarity=6.849158]
57: 18426501 Harvard_University
    [pem_pos=0, final_scores=0.3935138,
final_local_scores=0.42171437, pem=0.0, similarity=16.674511]
58: 22228744 Shaka_Smart
    [pem_pos=0, final_scores=0.065596715,
final_local_scores=0.0026005208, pem=0.0, similarity=4.167318]

```

Sample: b'2010_Marshall_Thundering_Herd_football_team'

The 2010 season was the 114th consecutive year of [1 1 Marsh
all University 2 Thundering Herd 1] 3 college 4 football .
The team competes in the eastern division of [1 1 Conference U
SA 1] of the [2 5 National Collegiate Athletic Association 2]
's . The 2010 season was the first for head coach . Marshall
finished the season 5 -- 7 , 4 -- 4 in [3 C-USA 3] play . In
the 6 2009 season under former head coach [2 2 Mark Snyder 2]
, the [4 Thundering Herd 4] finished with an overall record
of 7 -- 6 and 4 -- 4 within [5 Conference USA 5] . They
played in the 7 2009 [3 3 Little Caesars Pizza Bowl 3] ,
defeating the [6 8 Ohio University 9 Bobcats 6] 21 -- 17 . It
was Marshall first bowl game since the 10 2004 Fort Worth Bowl
, and the Herd 's first bowl win since the 11 2002 [7 GMAC Bo
wl 7] . [8 Snyder 8] is currently the defensive coordinator
of the [9 12 University of South Florida 13 Bulls 9] . : 14 Ro
se Bowl MVP [4 4 Terrelle Pryor 4] picked up where he left
off in [10 15 Pasadena 10] , [11 California 11] , with three

touchdown passes , [5 5 Brandon Saine 5] ran for 103 yards and two scores and Ohio State coasted past the error-prone [12 Thundering Herd 12] on Thursday night . It was a bleak debut for new Marshall coach [13 16 Doc Holliday 13] , who spent the last two seasons as an assistant at West Virginia . The [1 4 Thundering Herd 14] fumbled the opening kickoff and were down 14 -- 0 before running their first play in [15 Ohio State 15] territory . Ohio State 's defense limited Marshall to 44 yards on the ground and 199 total . [16 Brian Rolle 16] picked off [17 Brian Anderson 17] 's pass and returned it 30 yards for a touchdown late in the second quarter . [18 Anderson 18] , who completed 17 of 28 passes for 135 yards , was harassed most of the night . He had an interception and the Herd also lost two fumbles . External link : <http://espn.go.com/ncf/recap?gameId=302450194> [6 6 Geno Smith 6] rallied No. 23 West Virginia from 15 points down in the fourth quarter to force overtime and Tyler Bitancurt 's 20-yard field goal in the first extra session lifted the Mountaineers to a 24 -- 21 victory over Marshall on Friday night . West Virginia improved to 10 -- 0 all-time against Marshall . Except for their first meeting in 1911 , this one was the closest in the series between the state 's only [19 Bowl Subdivision 19] schools . Marshall coach [7 7 Doc Holliday 7] , the Mountaineers ' recruiting coordinator the past two seasons , brought inside knowledge of his former team . Former [20 WVU 20] offensive graduate assistant JaJuan Seider is Marshall 's running backs coach and recruiting coordinator , while three ex-graduate assistants at [21 WVU 21] have similar roles at Marshall . External link : <http://sports.espn.go.com/ncf/recap?gameId=302530276> Bowling Green returned two interceptions for touchdowns , scoring the first 21 points and the final 16 Saturday night in a 44 -- 28 win over Marshall . The [22 Thundering Herd 22] tied the game at 28 with 1:01 left in the third quarter on Martin Ward 's 1-yard run . The Falcons responded with Bryan Wright 's 27-yard field goal , and Aaron Pankratz threw a 31-yard touchdown pass to Tyrone Pronty with 9:21 left . External link : <http://espn.go.com/ncf/recap?gameId=302610189> [23 Brian Anderson 23] threw for three touchdowns and Marshall survived on the final play to hold off Ohio in a 24 -- 23 win on Saturday night . The [24 Thundering Herd 24] led 24 -- 17 when [25 Anderson 25] threw a 12-yard touchdown pass to Antavious Wilson with 8:21 left in the fourth quarter . [26 Anderson 26] , who finished with 295 yards passing , also had passing scores of 45 and 7 yards . Ohio 's Boo Jackson connected with Terrence McCrae on a 36-yard touchdown , cutting the lead to 24 -- 23 with no time left . The Bobcats elected to go for the two-point conversion , but Jackson 's pass sailed through the back of the end zone to secure Marshall 's win . External link : <http://espn.go.com/ncf/recap?gameId=302680276> Quarterback [27 Austin Davis 27] rushed for three touchdowns and threw for another as [28 Southern Mississippi 28] defeated Marshall 41

-- 16 Saturday night . [29 Southern Miss 29] built a 28 -- 0 halftime lead after Marshall gained just 46 yards on 23 first-half plays . The [30 Thundering Herd 30] finished with 180 yards of offense , the lowest total of any [31 Southern Miss 31] opponent since [8 8 Larry Fedora 8] took over after the 2007 season . Marshall defensive end [32 Vinny Curry 32] , the national sacks leader coming into Saturday , reached Davis twice to run his season total to eight in five games . [33 Southern Miss 33] ran its home winning streak to 11 games , and is off to its best start since 2004 when it started 4 -- 0 . External link : <http://espn.go.com/ncf/recap?gameId=302752572> Ronnie Weaver rushed for a career-high 150 yards , [34 Jeff Godfrey 34] ran for two touchdowns and Central Florida beat Marshall 35 -- 14 on Wednesday night . UCF broke open a close game in a steady rain with two touchdowns in the span of a minute in the third quarter and Marshall never recovered . The nation 's ninth-best defense limited Marshall 's running game to 37 yards . And on a soggy night , two Marshall quarterbacks could n't complete enough passes to overcome a big deficit . The game was stopped for more than an hour in the second quarter due to lightning . Only a few thousand fans returned to their seats when play resumed and halftime was canceled . External links : <http://espn.go.com/ncf/recap?gameId=302860276> Jon Williams ran for 111 yards and a touchdown to help [35 East Carolina 35] stay perfect in [36 Conference USA 36] by beating Marshall 37 -- 10 on Saturday . [9 9 Dominique Davis 9] threw for 208 yards and two scores for the [37 Pirates 37] , who are off to their best start in league play since going 5 -- 0 in 2001 . The two-time defending league champions have won 10 straight against [38 C-USA 38] opponents , including last season 's win against [39 Houston 39] in the title game . East Carolina shook off a sluggish start by scoring twice in the final 2 minutes of the first half to take a 20 -- 10 halftime lead . The [40 Pirates 40] then scored on their first three drives of the second half to blow the game open against the struggling [41 Thundering Herd 41] . Before the game , there was a moment of silence to honor the 75 victims of the 17 1970 plane crash that took place as the Marshall team was returning home from a loss at [42 East Carolina 42] . Next month marks the 40th anniversary of the crash . External link : <http://espn.go.com/ncf/recap?gameId=302960151> Tron Martinez gained a career-high 116 yards and scored on a 1-yard plunge with 1:47 remaining to give Marshall a 16 -- 12 victory over UTEP on Saturday . Martinez 's score , the first of his career , was set up by a personal foul penalty on the Miners . The [43 Thundering Herd 43] appeared to tie the score at 12 on Tyler Warner 's 19-yard field goal with just under 2 minutes to play . But Warner was knocked over on the attempt , and Marshall got the ball back . [44 Martinez 44] scored on the next play . UTEP led 12 -- 7 after Kris Adams caught a 23-yard TD pass from [10 10 Trevor Vittatoe 10] .

The Miners opted to go for the two-point conversion , but the play backfired . Donald Brown intercepted [45 Vittatoe 45] 's pass and returned it 100 yards for a two-point PAT return , making the score 12 -- 9 with 5:25 left . External link : <http://espn.go.com/ncf/recap?gameId=303030276> [46 Brian Anderson 46] threw for 304 yards and three touchdowns to lead Marshall to a 31 -- 17 win over Alabama-Birmingham on Saturday . [47 Anderson 47] , who completed 25 of 41 passes with one interception for the [48 Thundering Herd 48] , threw two first-quarter touchdown passes . The Blazers tied the game early in the fourth quarter on Josh Zahn 's 20-yard field goal . After the [49 Thundering Herd 49] 's Troy Evans recovered his own fumble in the end zone for 24 -- 17 , [50 Lee Smith 50] caught his second touchdown pass from [51 Anderson 51] with 5:02 to go . External link : <http://espn.go.com/ncf/recap?gameId=303100005> [52 Brian Anderson 52] threw three touchdown passes , all in the second half , as Marshall rallied to beat Memphis 28 -- 13 Saturday for its first three-game winning streak since 2006 . The [53 Thundering Herd 53] overcame a 10 -- 0 halftime deficit . Troy Evans started the scoring with a 13-yard touchdown run with 9:05 to go in the third quarter . After [54 Tigers 54] quarterback [55 Ryan Williams 55] fumbled a snap , Marshall took over on the Memphis 24 , setting up [56 Anderson 56] 's 6-yard scoring pass to Courtney Edmonson . Paulo Henriques kicked a 45-yard field goal to make it 21 -- 13 , then Anderson found Evans for a 62-yard touchdown . Marshall held the [57 Tigers 57] to 241 total yards , 77 in the second half . External link : <http://sports.espn.go.com/ncf/recap?gameId=303170276> [58 Zach Line 58] ran for 202 yards and a touchdown and [59 Kyle Padron 59] threw two TD passes and ran for a score as [60 Southern Methodist 60] defeated Marshall 31 -- 17 Saturday to become bowl eligible . [61 Brian Anderson 61] was 24 of 42 for 277 yards and a touchdown but also threw three interceptions for the [62 Thundering Herd 62] , who fell to 0 -- 7 all-time in [63 Texas 63] . Marshall was held to 50 rushing yards and had its three-game winning streak snapped . External link : <http://espn.go.com/ncf/recap?gameId=303242567> [64 Brian Anderson 64] threw for two touchdowns and Marshall jumped out to an early lead to beat Tulane 38 -- 23 on Saturday in the season finale for both teams . [65 Anderson 65] , who finished with 162 yards passing and one interception , connected with Courtney Edmonson on a 20-yard touchdown to give the [66 Thundering Herd 66] a 14 -- 7 lead in the second quarter . The [67 Thundering Herd 67] pushed their lead to 28 -- 7 after Kevin Perry scored off a blocked punt and Anderson threw a 19-yard touchdown pass to Tron Martinez . The [68 Green Wave 68] cut the margin to 14 points with 8 seconds left before halftime when [69 Ryan Grant 69] caught a 12-yard pass from Ryan Griffin . But Tulane pulled no closer after Tyler Warner 's 33-yard field goal and Rashad Jackson 's 84-yard interception

for a score for Marshall . External link : <http://espn.go.com/ncf/recap?gameId=303310276> The following players were selected in the 18 2011 NFL Draft

- 1: 78428 Conference_USA
- 2: 10567872 Mark_Snyder
- 3: 386136 Little_Caesars_Pizza_Bowl
- 4: 9918523 Terrelle_Pryor
- 5: 15335773 Brandon_Saine
- 6: 28833659 Geno_Smith
- 7: 24043868 Doc_Holliday_(American_football)
- 8: 13630435 Larry_Fedora
- 9: 20386281 Dominique_Davis
- 10: 19723837 Trevor_Vittatoe

1: 560365 Marshall_University

[winner: 8885940 Marshall_Thundering_Herd, pem_pos=0, final_scores=0.23704801, final_local_scores=0.22007892, pem=0.0, similarity=10.657307]

gt: 560365 Marshall_University, pem_pos=1, final_scores=0.10064511, final_local_scores=0.061927706, pem=-1.1856236, similarity=12.385365]

2: 8885940 Marshall_Thundering_Herd

[winner: 8885940 Marshall_Thundering_Herd, pem_pos=0, final_scores=0.23704801, final_local_scores=0.22007892, pem=0.0, similarity=10.657307]

gt: 8885940 Marshall_Thundering_Herd, pem_pos=0, final_scores=0.23704801, final_local_scores=0.22007892, pem=0.0, similarity=10.657307]

3: 6771 College_football

[winner: 6771 College_football, pem_pos=0, final_scores=-1.6436472, final_local_scores=-1.8514872, pem=-0.0074130204, similarity=-51.122047]

gt: 6771 College_football, pem_pos=0, final_scores=-1.6436472, final_local_scores=-1.8514872, pem=-0.0074130204, similarity=-51.122047]

4: 18951490 American_football

[winner: 6771 College_football, pem_pos=0, final_scores=-1.6436472, final_local_scores=-1.8514872, pem=-0.0074130204, similarity=-51.122047]

gt: 18951490 American_football, pem_pos=1, final_scores=-2.0871117, final_local_scores=-2.3177326, pem=-2.0672987, similarity=-53.83376]

5: 60706 National_Collegiate_Athletic_Association

[winner: 60706 National_Collegiate_Athletic_Association, pem_pos=0, final_scores=0.16848247, final_local_scores=0.15059298, pem=0.0, similarity=8.583708]

gt: 60706 National_Collegiate_Athletic_Association, pem_pos=0, final_scores=0.16848247, final_local_scores=0.15059298, pem=0.0, similarity=8.583708]

[winner: 1415027 David_Berman_(musician), pem_pos=5, final_scores=-2.2408042, final_local_scores=-2.3391194, pem=-3.0012724, similarity=-49.392887]

RECALL MISS: 60706

National_Collegiate_Athletic_Association]
6: 21035252 2009_Marshall_Thundering_Herd_football_team
[winner: 35983 2009, pem_pos=0, final_scores=-0.8811175,
final_local_scores=-0.7813709, pem=-0.9296576,
similarity=-14.17233|
RECALL MISS: 21035252
2009_Marshall_Thundering_Herd_football_team]
[winner: 690345 Season(sports), pem_pos=0,
final_scores=-1.7763933, final_local_scores=-1.9419272,
pem=-1.219945, similarity=-47.227016|
RECALL MISS: 21035252
2009_Marshall_Thundering_Herd_football_team]
7: 24226145 2009_Little_Caesars_Pizza_Bowl
[winner: 35983 2009, pem_pos=0, final_scores=-1.0415446,
final_local_scores=-0.956613, pem=-0.9296576,
similarity=-19.401905|
RECALL MISS: 24226145 2009_Little_Caesars_Pizza_Bowl]
8: 483329 Ohio_University
[winner: 7958366 Ohio_Bobcats_football, pem_pos=1,
final_scores=0.22941284, final_local_scores=0.21023166,
pem=-0.18232158, similarity=11.3549385|
gt: 483329 Ohio_University, pem_pos=2,
final_scores=0.06941025, final_local_scores=0.06316878,
pem=-1.0986123, similarity=11.94922]
9: 7958366 Ohio_Bobcats_football
[winner: 7958366 Ohio_Bobcats_football, pem_pos=1,
final_scores=0.22941284, final_local_scores=0.21023166,
pem=-0.18232158, similarity=11.3549385|
gt: 7958366 Ohio_Bobcats_football, pem_pos=1,
final_scores=0.22941284, final_local_scores=0.21023166,
pem=-0.18232158, similarity=11.3549385]
10: 16953913 2004_Fort_Worth_Bowl
[winner: 384584 Armed_Forces_Bowl, pem_pos=0,
final_scores=-0.014836609, final_local_scores=0.009987339,
pem=0.0, similarity=4.387755|
gt: 16953913 2004_Fort_Worth_Bowl, pem_pos=3,
final_scores=-0.38946664, final_local_scores=-0.33759654,
pem=-1.3862944, similarity=1.554043]
[winner: 35524 2004, pem_pos=0, final_scores=-0.7209571,
final_local_scores=-0.585704, pem=-0.4915334,
similarity=-10.715829|
RECALL MISS: 16953913 2004_Fort_Worth_Bowl]
11: 14739852 2002_GMAC_Bowl
[winner: 385350 GoDaddy_Bowl, pem_pos=0,
final_scores=0.038056657, final_local_scores=0.04269959,
pem=-0.090608984, similarity=5.8567|
gt: 14739852 2002_GMAC_Bowl, pem_pos=8,
final_scores=-0.32320258, final_local_scores=-0.2935744,
pem=-2.2276998, similarity=7.443445]
[winner: 35502 2002, pem_pos=0, final_scores=-0.99671304,
final_local_scores=-0.88436824, pem=-0.5897454,
similarity=-19.094473|

RECALL MISS: 14739852 2002_GMAC_Bowl]
 12: 3649046 University_of_South_Florida
 [winner: 6803709 South_Florida_Bulls_football, pem_pos=0,
 final_scores=0.26087767, final_local_scores=0.22965202,
 pem=0.0, similarity=10.942987]
 RECALL MISS: 3649046 University_of_South_Florida]
 13: 6803709 South_Florida_Bulls_football
 [winner: 6803709 South_Florida_Bulls_football, pem_pos=0,
 final_scores=0.26087767, final_local_scores=0.22965202,
 pem=0.0, similarity=10.942987]
 gt: 6803709 South_Florida_Bulls_football, pem_pos=0,
 final_scores=0.26087767, final_local_scores=0.22965202,
 pem=0.0, similarity=10.942987]
 14: 334276 Rose_Bowl_Game
 [winner: 1217500 Rose_Bowl_(stadium), pem_pos=0,
 final_scores=-0.07333304, final_local_scores=-0.08520212,
 pem=-0.65564966, similarity=5.1126366]
 gt: 334276 Rose_Bowl_Game, pem_pos=1,
 final_scores=-0.25697327, final_local_scores=-0.32554728,
 pem=-2.290435, similarity=6.8304744]
 15: 92408 Pasadena,_California
 [winner: 5407 California, pem_pos=0,
 final_scores=0.59632874, final_local_scores=0.7012776,
 pem=-0.08662026, similarity=25.48829]
 RECALL MISS: 92408 Pasadena,_California]
 [winner: 92408 Pasadena,_California, pem_pos=0,
 final_scores=0.33570236, final_local_scores=0.36973035,
 pem=-0.22124471, similarity=16.326366]
 gt: 92408 Pasadena,_California, pem_pos=0,
 final_scores=0.33570236, final_local_scores=0.36973035,
 pem=-0.22124471, similarity=16.326366]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.8917797, final_local_scores=-1.8626287,
 pem=-4.6728287, similarity=-26.083275]
 RECALL MISS: 92408 Pasadena,_California]
 16: 24043868 Doc_Holliday_(American_football)
 [winner: 298604 Doc_Holliday, pem_pos=0,
 final_scores=0.07600866, final_local_scores=0.15666276,
 pem=0.0, similarity=8.764842]
 gt: 24043868 Doc_Holliday_(American_football),
 pem_pos=1, final_scores=0.04463516,
 final_local_scores=0.058074668, pem=-1.2531835,
 similarity=12.637783]
 17: 4957196 Southern_Airways_Flight_932
 [winner: 34726 1970, pem_pos=0, final_scores=-0.7690731,
 final_local_scores=-0.6223136, pem=-0.80837256,
 similarity=-10.085312]
 RECALL MISS: 4957196 Southern_Airways_Flight_932]
 [winner: 84029 Plane_(geometry), pem_pos=0,
 final_scores=-1.1999246, final_local_scores=-1.1201491,
 pem=-0.40659222, similarity=-27.126663]
 gt: 4957196 Southern_Airways_Flight_932, pem_pos=13,

```

final_scores=-1.49085, final_local_scores=-1.5836401,
pem=-2.6370702, similarity=-28.828465 ]
[winner: 346953 PSA_Flight_182, pem_pos=0,
final_scores=-1.3723952, final_local_scores=-1.3914006,
pem=0.0, similarity=-37.432453]
RECALL MISS: 4957196 Southern_Airways_Flight_932]
18: 22399229 2011_NFL_Draft
[winner: 710502 National_Football_League_draft,
pem_pos=1, final_scores=0.021581396,
final_local_scores=-0.034062997, pem=-0.024888966,
similarity=3.2085552]
gt: 22399229 2011_NFL_Draft, pem_pos=6,
final_scores=-0.9833736, final_local_scores=-1.1077621,
pem=-5.9452753, similarity=3.363216 ]
[winner: 36225 2011, pem_pos=0, final_scores=-1.1841667,
final_local_scores=-1.0885832, pem=-1.4232541,
similarity=-20.655903]
RECALL MISS: 22399229 2011_NFL_Draft]
1: 78428 Conference_USA
[pem_pos=0, final_scores=0.14469998,
final_local_scores=0.11072484, pem=0.0, similarity=7.393963]
2: 10567872 Mark_Snyder
[pem_pos=0, final_scores=0.19228671,
final_local_scores=0.1806784, pem=0.0, similarity=9.481517]
3: 386136 Little_Caesars_Pizza_Bowl
[pem_pos=0, final_scores=0.07783948,
final_local_scores=0.07093233, pem=0.0, similarity=6.2064753]
4: 9918523 Terrelle_Pryor
[pem_pos=0, final_scores=0.4798563,
final_local_scores=0.5059502, pem=0.0, similarity=19.188276]
5: 15335773 Brandon_Saine
[pem_pos=0, final_scores=0.16007529,
final_local_scores=0.21949172, pem=0.0, similarity=10.639784]
6: 28833659 Geno_Smith
[pem_pos=0, final_scores=0.21827601,
final_local_scores=0.21064606, pem=0.0, similarity=10.375812]
7: 24043868 Doc_Holliday_(American_football)
[pem_pos=1, final_scores=0.049856648,
final_local_scores=0.063778356, pem=-1.2531835,
similarity=12.807993]
8: 13630435 Larry_Fedora
[pem_pos=0, final_scores=0.3327737,
final_local_scores=0.35577357, pem=0.0, similarity=14.706706]
9: 20386281 Dominique_Davis
[pem_pos=0, final_scores=0.13892375,
final_local_scores=0.13713714, pem=0.0, similarity=8.182158]
10: 19723837 Trevor_Vittatoe
[pem_pos=0, final_scores=0.36482793,
final_local_scores=0.4534366, pem=0.0, similarity=17.621166]
1: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.23704801,
final_local_scores=0.22007892, pem=0.0, similarity=10.657307]

```

2: 60706 National_Collegiate_Athletic_Association
[pem_pos=0, final_scores=0.16848247,
final_local_scores=0.15059298, pem=0.0, similarity=8.583708]
3: 78428 Conference_USA
[pem_pos=0, final_scores=0.10109587,
final_local_scores=0.06309399, pem=0.0, similarity=5.9725633]
4: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.13306116,
final_local_scores=0.10606398, pem=-0.07968813,
similarity=7.6882305]
5: 78428 Conference_USA
[pem_pos=0, final_scores=0.16938116,
final_local_scores=0.1376853, pem=0.0, similarity=8.198517]
6: 7958366 Ohio_Bobcats_football
[pem_pos=1, final_scores=0.22941284,
final_local_scores=0.21023166, pem=-0.18232158,
similarity=11.3549385]
7: 385350 GoDaddy_Bowl
[pem_pos=0, final_scores=0.038056657,
final_local_scores=0.04269959, pem=-0.090608984,
similarity=5.8567]
8: 10567872 Mark_Snyder
[pem_pos=0, final_scores=0.17774205,
final_local_scores=0.16479057, pem=0.0, similarity=9.007392]
9: 6803709 South_Florida_Bulls_football
[pem_pos=0, final_scores=0.26087767,
final_local_scores=0.22965202, pem=0.0, similarity=10.942987]
10: 92408 Pasadena_California
[pem_pos=0, final_scores=0.33570236,
final_local_scores=0.36973035, pem=-0.22124471,
similarity=16.326366]
11: 5407 California
[pem_pos=0, final_scores=0.59632874,
final_local_scores=0.7012776, pem=-0.08662026,
similarity=25.48829]
12: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.16764887,
final_local_scores=0.14421928, pem=-0.07968813,
similarity=8.82686]
13: 298604 Doc_Holliday
[pem_pos=0, final_scores=0.07600866,
final_local_scores=0.15666276, pem=0.0, similarity=8.764842]
14: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.18659477,
final_local_scores=0.1645413, pem=-0.07968813,
similarity=9.43331]
15: 22217 Ohio_State_University
[pem_pos=0, final_scores=0.042053625,
final_local_scores=0.035200983, pem=-0.9758306,
similarity=10.446899]
16: 31640116 Brian_Rolle
[pem_pos=0, final_scores=0.16184591,

final_local_scores=0.22643793, pem=0.0, similarity=10.847073]
17: 5224238 Brian_Anderson_(outfielder)
[pem_pos=2, final_scores=0.34264463,
final_local_scores=0.35682747, pem=-1.0207831,
similarity=20.289333]
18: 5224238 Brian_Anderson_(outfielder)
[pem_pos=2, final_scores=0.42756993,
final_local_scores=0.4495954, pem=-1.0207831,
similarity=23.057713]
19: 404595 NCAA_Division_I_Football_Championship
[pem_pos=0, final_scores=0.08817552,
final_local_scores=0.05515836, pem=-0.57807785,
similarity=8.879426]
20: 185979 West_Virginia_University
[pem_pos=0, final_scores=0.083350435,
final_local_scores=0.05850844, pem=-0.040671032,
similarity=6.056897]
21: 185979 West_Virginia_University
[pem_pos=0, final_scores=0.08574064,
final_local_scores=0.061119378, pem=-0.040671032,
similarity=6.134813]
22: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.17087243,
final_local_scores=0.14736703, pem=-0.07968813,
similarity=8.920795]
23: 5224238 Brian_Anderson_(outfielder)
[pem_pos=2, final_scores=0.46039468,
final_local_scores=0.48545152, pem=-1.0207831,
similarity=24.127733]
24: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.20734899,
final_local_scores=0.18758562, pem=-0.07968813,
similarity=10.120998]
25: 5224238 Brian_Anderson_(outfielder)
[pem_pos=2, final_scores=0.19255732,
final_local_scores=0.19287989, pem=-1.0207831,
similarity=15.396811]
26: 5224238 Brian_Anderson_(outfielder)
[pem_pos=2, final_scores=0.4082703,
final_local_scores=0.42856103, pem=-1.0207831,
similarity=22.430006]
27: 24161425 Austin_Davis
[pem_pos=0, final_scores=0.28990203,
final_local_scores=0.29987583, pem=0.0, similarity=13.038606]
28: 4998626 Southern_Miss_Golden_Eagles_and_Lady_Eagles
[pem_pos=0, final_scores=0.23553048,
final_local_scores=0.22252977, pem=-0.18565795,
similarity=11.740082]
29: 4998626 Southern_Miss_Golden_Eagles_and_Lady_Eagles
[pem_pos=1, final_scores=0.064937845,
final_local_scores=0.036183238, pem=-0.662015,
similarity=8.769633]

30: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.24550603,
final_local_scores=0.2292664, pem=-0.07968813,
similarity=11.364836]
31: 4998626 Southern_Miss_Golden_Eagles_and_Lady_Eagles
[pem_pos=1, final_scores=0.057164416,
final_local_scores=0.02769193, pem=-0.662015,
similarity=8.516235]
32: 33443440 Vinny_Curry
[pem_pos=0, final_scores=0.2148559,
final_local_scores=0.27738336, pem=0.0, similarity=12.367386]
33: 4998626 Southern_Miss_Golden_Eagles_and_Lady_Eagles
[pem_pos=1, final_scores=0.055430606,
final_local_scores=0.025798023, pem=-0.662015,
similarity=8.459718]
34: 29530896 Jeff_Godfrey
[pem_pos=0, final_scores=0.041912958,
final_local_scores=0.06464681, pem=0.0, similarity=6.018903]
35: 11939661 East_Carolina_Pirates_football
[pem_pos=1, final_scores=0.040345594,
final_local_scores=-0.008062571, pem=-1.2745516,
similarity=10.780319]
36: 78428 Conference_USA
[pem_pos=0, final_scores=0.12844683,
final_local_scores=0.09297077, pem=0.0, similarity=6.864146]
37: 24270 Pittsburgh_Pirates
[pem_pos=1, final_scores=0.6035792,
final_local_scores=0.604914, pem=-1.4743205,
similarity=30.159136]
38: 78428 Conference_USA
[pem_pos=0, final_scores=0.05956115,
final_local_scores=0.017723605, pem=0.0, similarity=4.618621]
39: 13894 Houston_Astros
[pem_pos=11, final_scores=0.23424868,
final_local_scores=0.20307231, pem=-4.650927,
similarity=35.44226]
40: 24270 Pittsburgh_Pirates
[pem_pos=1, final_scores=0.4260491,
final_local_scores=0.41098934, pem=-1.4743205,
similarity=24.372038]
41: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.08809294,
final_local_scores=0.056943014, pem=-0.07968813,
similarity=6.2223625]
42: 89488 East_Carolina_University
[pem_pos=0, final_scores=0.09262101,
final_local_scores=0.06600219, pem=-1.0250391,
similarity=11.633672]
43: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.13131128,
final_local_scores=0.104152486, pem=-0.07968813,
similarity=7.6311874]

44: 784152 Conchita_Martínez
[pem_pos=6, final_scores=0.049005255,
final_local_scores=0.17765522, pem=-3.2962992,
similarity=27.317087]
45: 19723837 Trevor_Vittatoe
[pem_pos=0, final_scores=0.10049744,
final_local_scores=0.16469571, pem=0.0, similarity=9.004561]
46: 5224238 Brian_Anderson_(outfielder)
[pem_pos=2, final_scores=0.34555858,
final_local_scores=0.36001047, pem=-1.0207831,
similarity=20.384321]
47: 5224238 Brian_Anderson_(outfielder)
[pem_pos=2, final_scores=0.31644982,
final_local_scores=0.32826105, pem=-1.0207831,
similarity=19.436855]
48: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.10999785,
final_local_scores=0.08087079, pem=-0.07968813,
similarity=6.936415]
49: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.12519868,
final_local_scores=0.097475395, pem=-0.07968813,
similarity=7.4319296]
50: 808613 Lee_Smith_(baseball)
[pem_pos=0, final_scores=0.23239748,
final_local_scores=0.2515408, pem=-0.47724175,
similarity=14.191507]
51: 1574319 Brian_Anderson_(pitcher)
[pem_pos=0, final_scores=0.11174126,
final_local_scores=0.09579161, pem=-0.44362736,
similarity=9.36084]
52: 5224238 Brian_Anderson_(outfielder)
[pem_pos=2, final_scores=0.4375913,
final_local_scores=0.46054226, pem=-1.0207831,
similarity=23.38439]
53: 8885940 Marshall_Thundering_Herd
[pem_pos=0, final_scores=0.120281145,
final_local_scores=0.092103735, pem=-0.07968813,
similarity=7.271629]
54: 8579 Detroit_Tigers
[pem_pos=0, final_scores=0.5930757,
final_local_scores=0.5924427, pem=-1.1469651,
similarity=28.00676]
55: 24763852 Ryan_Williams_(American_football)
[pem_pos=1, final_scores=0.118508354,
final_local_scores=0.15604904, pem=-0.16653766,
similarity=9.652185]
56: 1574319 Brian_Anderson_(pitcher)
[pem_pos=0, final_scores=0.089352354,
final_local_scores=0.07133514, pem=-0.44362736,
similarity=8.63101]
57: 8579 Detroit_Tigers

[pem_pos=0, final_scores=0.4529788,
 final_local_scores=0.43940806, pem=-1.1469651,
 similarity=23.439901]
 58: 39481419 Zach_Line
 [pem_pos=0, final_scores=0.03399737,
 final_local_scores=0.098554686, pem=0.0, similarity=7.0307813]
 59: 36854067 Kyle_Padron
 [pem_pos=0, final_scores=0.22930895,
 final_local_scores=0.27802458, pem=0.0, similarity=12.386521]
 60: 479109 Southern_Methodist_University
 [pem_pos=0, final_scores=0.08549796,
 final_local_scores=0.07623388, pem=-0.1988119,
 similarity=7.445854]
 61: 5224238 Brian_Anderson_(outfielder)
 [pem_pos=2, final_scores=0.36046654,
 final_local_scores=0.3763426, pem=-1.0207831,
 similarity=20.871702]
 62: 8885940 Marshall_Thundering_Herd
 [pem_pos=0, final_scores=0.039016858,
 final_local_scores=0.003334865, pem=-0.07968813,
 similarity=4.6225886]
 63: 29810 Texas
 [pem_pos=0, final_scores=0.32378715,
 final_local_scores=0.3881284, pem=-0.16191302,
 similarity=16.552746]
 64: 5224238 Brian_Anderson_(outfielder)
 [pem_pos=2, final_scores=0.35441792,
 final_local_scores=0.36968797, pem=-1.0207831,
 similarity=20.673117]
 65: 5224238 Brian_Anderson_(outfielder)
 [pem_pos=2, final_scores=0.4123187,
 final_local_scores=0.43293577, pem=-1.0207831,
 similarity=22.560558]
 66: 8885940 Marshall_Thundering_Herd
 [pem_pos=0, final_scores=0.07003771,
 final_local_scores=0.037220433, pem=-0.07968813,
 similarity=5.6338015]
 67: 8885940 Marshall_Thundering_Herd
 [pem_pos=0, final_scores=0.19228758,
 final_local_scores=0.17075983, pem=-0.07968813,
 similarity=9.618883]
 68: 2842113 Tulane_Green_Wave
 [pem_pos=0, final_scores=0.06195925,
 final_local_scores=0.03970632, pem=-0.4248462,
 similarity=7.585008]
 69: 10192225 Ryan_Grant
 [pem_pos=0, final_scores=0.20199428,
 final_local_scores=0.24157274, pem=0.0, similarity=11.298725]

Sample: b'Jackie_Tyrrell'

Jackie Tyrrell -LRB- born 19 June 1982 -RRB- is an [1 1 Irish
 1] 2 hurler who plays as a left corner-back for the [1 1 Kilk

enny 1] senior team . He joined the team during the and became a regular member of the starting fifteen over subsequent seasons . [2 Tyrrell 2] is regarded as one of the lynchpins of the [3 Kilkenny 3] defence . He has won six [2 2 All-Ireland 2] winners ' medals , seven [4 3 Leinster 4] winners ' medals and four 4 National League winners ' medal . The All-Ireland-winning captain of 2006 , he has ended up as an [5 All-Ireland 5] runner-up on one occasion . At club level [6 Tyrrell 6] is an [7 5 All-Ireland 7] medalist with [3 3 James Stephens 3] . In addition to this he has also won two [8 6 Leinster 8] medals , three 7 county hurling championship medals and two 8 county football championship medals . [9 Jackie Tyrrell 9] was born in [4 4 Kilkenny 4] in 1982 . He received his primary education locally before later attending , a famed hurling nursery in the city . It was here that [10 Tyrrell 10] 's hurling skills were developed and nurtured by the many hurling enthusiasts there . He won back-to-back [11 Leinster 11] colleges ' titles in 1999 and 2000 . This latter victory was later converted into an [12 9 All-Ireland 12] colleges title after a defeat of Gort Community College . [13 Tyrrell 13] later studied at the [5 5 Cork Institute of Technology 5] and at the [6 6 Limerick Institute of Technology 6] . It was at the latter institution that he won a [7 7 Fitzgibbon Cup 7] medal on a team trained by [8 8 Clare 8] 's [9 9 Davy Fitzgerald 9] . [14 Tyrrell 14] won a second [1 5 Fitzgibbon Cup 15] title two years after his first in 2007 . === Club === [16 Tyrrell 16] plays his club hurling and football with the famous [10 10 James Stephens 10] club . He has had much success at under-age levels , including championship medals at under-14 , under-16 , under-21 and junior . In 2003 [17 Tyrrell 17] won a 10 county football championship medal following a 2-7 to 0-6 defeat of 11 O'Loughlin Gaels . [18 Tyrrell 18] added a 12 county hurling championship medal to his collection in 2004 following a 2-16 to 3-12 defeat of 13 Young Ireland 's . He later collected a [19 14 Leinster 19] title , although the one-point defeat of [20 15 UCD 20] was in controversial circumstances . The [21 James Stephens 21] club subsequently qualified for the [22 All-Ireland 22] final with 16 Athenry providing the opposition . A victory for `` the village `` gave [23 Tyrrell 23] a coveted [24 17 All-Ireland 24] club medal . [25 James Stephens 25] retained the county championship in 2005 , with [26 Tyrrell 26] winning a second hurling medal following a 1-18 to 2-12 defeat pf . He later picked up a second [27 Leinster 27] club medal following a second consecutive triumph over [28 UCD 28] . [29 James Stephens 29] were subsequently trounced by eventual winners [11 11 Portumna 11] in the [30 All-Ireland 30] semi-final . After a six-year gap [31 Tyrrell 31] won a third hurling championship in 2011 after a thrilling draw and replay with [32 Ballyhale Shamrocks 32] . [33 Tyrrell 33] first came to prominence on the inter-county scene as a member of the [12 12 Kilkenny 12] under-21 team .

Success was slow in coming in this grade as [13 13 Wexford 13] dominated in the early years of the new century . In 2003 [34 Kilkenny 34] defeated [14 14 Dublin 14] by 0-12 to 1-4 giving [35 Tyrrell 35] a [36 18 Leinster 36] medal in that grade . He later lined out in the under-21 championship decider with [15 15 Galway 15] providing the opposition . [37 Kilkenny 37] won that game by 2-13 to 0-12 with [38 Tyrrell 38] collecting a coveted [39 19 All-Ireland 39] medal and collecting the cup as captain . In 2003 [40 Tyrrell 40] joined the [41 Kilkenny 41] senior team for the first time . He was a non-playing substitute that year as ' the Cats ' claimed the [16 16 All-Ireland 16] , [20 Leinster 20] and [21 National League 21] titles . [42 Tyrrell 42] made his competitive debut for [43 Kilkenny 43] in a National League game against [17 17 Waterford 17] in 2004 . He was a regular during that league campaign , however , he played no part in the subsequent championship . [44 Kilkenny 44] were back in form in 2005 , with [45 Tyrrell 45] winning his first National League medal following a huge 3-20 to 0-15 victory over [46 Clare 46] . [47 Kilkenny 47] later defeated [48 Wexford 48] by 0-22 to 1-16 in the provincial decider , giving Delaney a first [49 Leinster 49] winners ' medal . While a third successive [50 All-Ireland 50] showdown with [51 Cork 51] seemed likely , [52 Galway 52] defeated [53 Kilkenny 53] in the All-Ireland semi-final in one of the games of the decade . In 2006 [54 Tyrrell 54] was appointed captain for the year . He added a second consecutive National League medal to his collection following a 3-11 to 0-14 victory over [55 22 Limerick 55] . He later won his second [56 Leinster 56] medal following another facile victory over [57 Wexford 57] . The subsequent [58 All-Ireland 58] decider pitted [59 Kilkenny 59] against [60 23 Cork 60] . The Leesiders were aiming for a third [61 All-Ireland 61] victory in-a-row ; however , revenge was foremost in the minds of [62 Kilkenny 62] as it was [63 Cork 63] who denied their three-in-a-row bid in 2004 . On the day [64 Kilkenny 64] were far too strong for [65 Cork 65] as ' the Cats ' secured a 1-16 to 1-13 victory . It was Walsh 's first [66 All-Ireland 66] winners ' medal , while he also had the honour of lifting the cup as captain . [67 Tyrrell 67] collected a third [68 Leinster 68] medal in-a-row in 2007 following another facile victory in the provincial decider . [69 Kilkenny 69] later qualified for the [70 All-Ireland 70] final where , surprisingly , [71 24 Limerick 71] provided the opposition . [72 Kilkenny 72] got off to a flying start with [18 18 Eddie Brennan 18] and [19 19 Henry Shefflin 19] scoring two goals within the first ten minutes and went on to win the game by 2-19 to 1-15 . It was [73 Tyrrell 73] 's second consecutive [74 All-Ireland 74] medal . He was later presented with a coveted award . 2008 saw [75 Kilkenny 75] launch an all-out attack in an attempt to capture a third [76 All-Ireland 76] title in-a-row . [77 Tyrrell 77] collected a fourth [78 Leinster 78] winners ' medal after a 5 - - 21 to 0

-- 17 defeat of [79 Wexford 79] in the provincial final , before later lining out in the All-Ireland decider . [20 20 Waterford 20] provided the opposition , however , [80 Kilkenny 80] recorded an enormous 3 -- 30 to 1 -- 13 victory . It was a third All-Ireland winners ' medal for [81 Tyrrell 81] . A second All-Star award quickly followed . [82 Tyrrell 82] won his third National League winners ' medal in 2009 following a thrilling 4-17 to 2-26 defeat of [21 21 Tipperary 21] . A fifth [83 Leinster 83] medal quickly followed before [84 Kilkenny 84] set out to make history by becoming the second team in hurling history to win four successive [85 All-Ireland 85] finals . [22 22 Tipperary 22] were the opponents and , for much of the match , it looked as if they would pull off a shock and deny ` the Cats ' the title . Two quick goals in the space of a minute , one from a penalty by [86 Henry Shefflin 86] , sealed a 2 -- 22 to 0 -- 23 victory . It was [87 Tyrrell 87] 's fourth consecutive [88 All-Ireland 88] medal . He later won a third All-Star award . [89 Tyrrell 89] added a sixth [90 Leinster 90] winners ' medal to his collection in 2010 following a 1-19 to 1-12 defeat of new provincial rivals [91 Galway 91] . [92 Kilkenny 92] faced a number of injury worries going into an historic [93 All-Ireland 93] final , and ultimately failed in their ` drive for five ' as [94 Tipperary 94] won by 4-17 to 1-18 . In spite of the defeat Delaney later won a fourth successive All-Star award . In 2011 ` the Cats ' continued their provincial dominance with [95 Tyrrell 95] winning a seventh consecutive [96 Leinster 96] winners ' medal following a defeat of [23 23 Dublin 23] . [97 Kilkenny 97] later qualified for a sixth consecutive [98 All-Ireland 98] final . For the third year in succession [99 Tyrrell 99] 's side faced [100 Tipperary 100] , however , on this occasion [101 Kilkenny 101] were slight underdogs going up against the new champions . [102 Kilkenny 102] started quickly and never surrendered the lead in the 2-17 to 1-16 victory . It was Delaney 's fifth All-Ireland winners ' medal . 2012 began well for [103 Kilkenny 103] . A trouncing of old rivals [104 Cork 104] gave [105 Tyrrell 105] a fourth National League winners ' medal . After losing the [106 Leinster 106] final to [24 24 Galway 24] , both sides subsequently met in the All-Ireland decider and [107 Galway 107] nearly pulled off a victory courtesy of goals from [25 25 Joe Canning 25] and [26 26 Niall Burke 26]

1: 2795066 Kilkenny_GAA

2: 780802 All-Ireland_Senior_Hurling_Championship

3: 6837581 James_Stephens_GAA

4: 92621 Kilkenny

5: 542603 Cork_Institute_of_Technology

6: 537825 Limerick_Institute_of_Technology

7: 4560007 Fitzgibbon_Cup

8: 2794983 Clare_GAA

9: 5795114 Davy_Fitzgerald

10: 6837581 James_Stephens_GAA

```

11: 16715841 Portumna_GAA
12: 2795066 Kilkenny_GAA
13: 2810897 Wexford_GAA
14: 2759665 Dublin_GAA
15: 2810853 Galway_GAA
16: 780802 All-Ireland_Senior_Hurling_Championship
17: 2802645 Waterford_GAA
18: 5575719 Eddie_Brennan
19: 5287742 Henry_Shefflin
20: 2802645 Waterford_GAA
21: 2802637 Tipperary_GAA
22: 2802637 Tipperary_GAA
23: 2759665 Dublin_GAA
24: 2810853 Galway_GAA
25: 12949356 Joe_Canning
26: 35139982 Niall_Burke
1: 775859 Irish_people
    [winner: 14560 Republic_of_Ireland, pem_pos=3,
    final_scores=0.3304627, final_local_scores=0.3824578,
    pem=-2.2576945, similarity=27.78071|
    gt: 775859 Irish_people, pem_pos=0,
    final_scores=-0.34104228, final_local_scores=-0.35024363,
    pem=-1.1828228, similarity=0.07011868 ]
2: 73105 Hurling
    [winner: 73105 Hurling, pem_pos=0,
    final_scores=-1.835065, final_local_scores=-2.0468025,
    pem=-0.061817452, similarity=-56.654785|
    gt: 73105 Hurling, pem_pos=0, final_scores=-1.835065,
    final_local_scores=-2.0468025, pem=-0.061817452,
    similarity=-56.654785 ]
3: 3531363 Leinster_Senior_Hurling_Championship
    [winner: 985961 Leinster_Rugby, pem_pos=1,
    final_scores=0.107422605, final_local_scores=0.08209887,
    pem=-1.3755379, similarity=14.020096|
    gt: 3531363 Leinster_Senior_Hurling_Championship,
    pem_pos=11, final_scores=-0.11795074,
    final_local_scores=-0.19982168, pem=-2.6704001,
    similarity=12.648686 ]
4: 2802698 National_Hurling_League
    [winner: 64290 National_League, pem_pos=0,
    final_scores=-0.033826813, final_local_scores=0.011990011,
    pem=-0.07431644, similarity=4.851663|
    gt: 2802698 National_Hurling_League, pem_pos=8,
    final_scores=-0.34695876, final_local_scores=-0.4572574,
    pem=-3.3032966, similarity=8.40808 ]
5: 3341596 All-Ireland_Senior_Club_Hurling_Championship
    [winner: 780802 All-Ireland_Senior_Hurling_Championship,
    pem_pos=0, final_scores=0.16431175,
    final_local_scores=0.10153459, pem=-0.83555937,
    similarity=11.6636095|
    RECALL MISS: 3341596 All-
    Ireland_Senior_Club_Hurling_Championship]

```

6: 7501228 Leinster_Senior_Club_Hurling_Championship
[winner: 985961 Leinster_Rugby, pem_pos=1,
final_scores=0.14117165, final_local_scores=0.11898634,
pem=-1.3755379, similarity=15.120893]
gt: 7501228 Leinster_Senior_Club_Hurling_Championship,
pem_pos=5, final_scores=-0.13675697,
final_local_scores=-0.174146, pem=-2.1648939,
similarity=10.665878]
7: 6932812 Kilkenny_Senior_Hurling_Championship
[winner: 7086088 Wexford_Senior_Hurling_Championship,
pem_pos=1, final_scores=-0.587229,
final_local_scores=-0.6573477, pem=-0.6021754,
similarity=-12.25213]
gt: 6932812 Kilkenny_Senior_Hurling_Championship,
pem_pos=2, final_scores=-0.88291615,
final_local_scores=-0.9978498, pem=-1.658228,
similarity=-16.670414]
8: 7085557 Kilkenny_Senior_Football_Championship
[winner: 885732 All-Ireland_Senior_Football_Championship,
pem_pos=0, final_scores=-0.85382193,
final_local_scores=-0.99979687, pem=0.0,
similarity=-25.74622]
RECALL MISS: 7085557
Kilkenny_Senior_Football_Championship]
[winner: 382207 Powiat, pem_pos=0,
final_scores=-1.2281106, final_local_scores=-1.1831994,
pem=-0.39345494, similarity=-29.079647]
RECALL MISS: 7085557
Kilkenny_Senior_Football_Championship]
9: 5011213 Dr._Croke_Cup
[winner: 780802 All-Ireland_Senior_Hurling_Championship,
pem_pos=0, final_scores=0.1528322,
final_local_scores=0.088994935, pem=-0.83555937,
similarity=11.289401]
gt: 5011213 Dr._Croke_Cup, pem_pos=8,
final_scores=-0.38983864, final_local_scores=-0.4103055,
pem=-2.6813574, similarity=6.427017]
[winner: 10127998 Collegium_(ministry), pem_pos=2,
final_scores=-1.6173538, final_local_scores=-1.5661695,
pem=-1.6038731, similarity=-33.9258]
RECALL MISS: 5011213 Dr._Croke_Cup]
10: 7085557 Kilkenny_Senior_Football_Championship
[winner: 885732 All-Ireland_Senior_Football_Championship,
pem_pos=0, final_scores=-0.70852506,
final_local_scores=-0.84108204, pem=0.0,
similarity=-21.009853]
RECALL MISS: 7085557
Kilkenny_Senior_Football_Championship]
[winner: 93408 Harding_County,_New_Mexico, pem_pos=0,
final_scores=-0.80213386, final_local_scores=-0.70526767,
pem=0.0, similarity=-16.95688]
RECALL MISS: 7085557

Kilkenny_Senior_Football_Championship]
11: 23079478 O'Loughlin_Gaels_GAA
[winner: 23079478 O'Loughlin_Gaels_GAA, pem_pos=0,
final_scores=-0.11144562, final_local_scores=-0.14152843,
pem=0.0, similarity=-0.13377763]
gt: 23079478 O'Loughlin_Gaels_GAA, pem_pos=0,
final_scores=-0.11144562, final_local_scores=-0.14152843,
pem=0.0, similarity=-0.13377763]
12: 6932812 Kilkenny_Senior_Hurling_Championship
[winner: 7086088 Wexford_Senior_Hurling_Championship,
pem_pos=1, final_scores=-0.5072027,
final_local_scores=-0.5699311, pem=-0.6021754,
similarity=-9.643444]
gt: 6932812 Kilkenny_Senior_Hurling_Championship,
pem_pos=2, final_scores=-0.7976155,
final_local_scores=-0.90467185, pem=-1.658228,
similarity=-13.889797]
13: 13404771 Young_Irelands_GAA
[winner: 1943310 Young_Ireland, pem_pos=0,
final_scores=0.015516862, final_local_scores=0.08201648,
pem=-0.025664346, similarity=6.676815]
RECALL MISS: 13404771 Young_Irelands_GAA]
[winner: 33880936 English_possessive, pem_pos=0,
final_scores=-2.2596788, final_local_scores=-2.3353944,
pem=-0.7098606, similarity=-61.742783]
RECALL MISS: 13404771 Young_Irelands_GAA]
14: 7501228 Leinster_Senior_Club_Hurling_Championship
[winner: 985961 Leinster_Rugby, pem_pos=1,
final_scores=0.077520385, final_local_scores=0.049435228,
pem=-1.3755379, similarity=13.045348]
gt: 7501228 Leinster_Senior_Club_Hurling_Championship,
pem_pos=5, final_scores=-0.24991341,
final_local_scores=-0.29799515, pem=-2.1648939,
similarity=6.969972]
15: 7065111 UCD_GAA
[winner: 1234706 University_College_Dublin_A.F.C.,
pem_pos=0, final_scores=0.15514006,
final_local_scores=0.11857185, pem=-0.8758797,
similarity=12.391303]
gt: 7065111 UCD_GAA, pem_pos=1,
final_scores=-0.012580901, final_local_scores=-0.060609743,
pem=-1.1279844, similarity=8.415152]
16: 25648606 Athenry_GAA
[winner: 25648606 Athenry_GAA, pem_pos=1,
final_scores=0.0032768697, final_local_scores=-0.012871072,
pem=-1.0885377, similarity=9.625252]
gt: 25648606 Athenry_GAA, pem_pos=1,
final_scores=0.0032768697, final_local_scores=-0.012871072,
pem=-1.0885377, similarity=9.625252]
17: 3341596 All-Ireland_Senior_Club_Hurling_Championship
[winner: 780802 All-Ireland_Senior_Hurling_Championship,
pem_pos=0, final_scores=0.09899174,

final_local_scores=0.030182406, pem=-0.83555937,
 similarity=9.534318|
 RECALL MISS: 3341596 All-Ireland_Senior_Club_Hurling_Championship]
 18: 6195780 Leinster_Under-21_Hurling_Championship
 [winner: 985961 Leinster_Rugby, pem_pos=1,
 final_scores=0.16786595, final_local_scores=0.14817062,
 pem=-1.3755379, similarity=15.991809|
 gt: 6195780 Leinster_Under-21_Hurling_Championship,
 pem_pos=13, final_scores=-0.46496505,
 final_local_scores=-0.52434325, pem=-4.089484,
 similarity=10.681515]
 19: 6141965 All-Ireland_Under-21_Hurling_Championship
 [winner: 780802 All-Ireland_Senior_Hurling_Championship,
 pem_pos=0, final_scores=0.06572385,
 final_local_scores=-0.0067884773, pem=-0.83555937,
 similarity=8.431033|
 gt: 6141965 All-Ireland_Under-21_Hurling_Championship,
 pem_pos=3, final_scores=-0.32154238,
 final_local_scores=-0.4066107, pem=-2.3609734,
 similarity=4.7949786]
 20: 3531363 Leinster_Senior_Hurling_Championship
 [winner: 6087825 Brenton_Leister, pem_pos=2,
 final_scores=-0.2262338, final_local_scores=-0.1603744,
 pem=-1.609438, similarity=8.056194|
 gt: 3531363 Leinster_Senior_Hurling_Championship,
 pem_pos=6, final_scores=-0.4111876,
 final_local_scores=-0.52006453, pem=-4.0775375,
 similarity=10.744234]
 21: 2802698 National_Hurling_League
 [winner: 64290 National_League, pem_pos=0,
 final_scores=-0.04967551, final_local_scores=-0.0071751326,
 pem=-0.07431644, similarity=4.279737|
 gt: 2802698 National_Hurling_League, pem_pos=8,
 final_scores=-0.30678087, final_local_scores=-0.41332898,
 pem=-3.3032966, similarity=9.718992]
 22: 2810846 Limerick_GAA
 [winner: 1234583 Limerick_F.C., pem_pos=4,
 final_scores=0.28157276, final_local_scores=0.26514578,
 pem=-2.6802762, similarity=26.57795|
 gt: 2810846 Limerick_GAA, pem_pos=1,
 final_scores=0.11876194, final_local_scores=0.05095589,
 pem=-1.520826, similarity=13.880827]
 23: 32532531 Cork_GAA
 [winner: 7759007 Jack_Cork, pem_pos=5,
 final_scores=0.087552145, final_local_scores=0.112473786,
 pem=-3.552999, similarity=26.767916|
 gt: 32532531 Cork_GAA, pem_pos=1,
 final_scores=0.062330052, final_local_scores=-0.016239643,
 pem=-1.0852067, similarity=9.506613]
 24: 2810846 Limerick_GAA
 [winner: 1234583 Limerick_F.C., pem_pos=4,

```
final_scores=0.17101337, final_local_scores=0.14439487,  
pem=-2.6802762, similarity=22.974503]  
    gt: 2810846 Limerick_GAA, pem_pos=1,  
final_scores=0.05548291, final_local_scores=-0.018791258,  
pem=-1.520826, similarity=11.799433 ]  
1: 2795066 Kilkenny_GAA  
    [pem_pos=1, final_scores=0.1843292,  
final_local_scores=0.111410394, pem=-1.0248141,  
similarity=12.98752]  
2: 780802 All-Ireland_Senior_Hurling_Championship  
    [pem_pos=0, final_scores=0.053989217,  
final_local_scores=-0.019606784, pem=-0.83555937,  
similarity=8.04851]  
3: 6837581 James_Stephens_GAA  
    [pem_pos=1, final_scores=0.115096614,  
final_local_scores=0.09028819, pem=-0.7831714,  
similarity=11.0431]  
4: 92621 Kilkenny  
    [pem_pos=0, final_scores=0.053048298,  
final_local_scores=0.06672023, pem=-0.98183,  
similarity=11.420122]  
5: 542603 Cork_Institute_of_Technology  
    [pem_pos=0, final_scores=0.16242896,  
final_local_scores=0.18689737, pem=0.0, similarity=9.667103]  
6: 537825 Limerick_Institute_of_Technology  
    [pem_pos=0, final_scores=0.12790553,  
final_local_scores=0.15274084, pem=0.0, similarity=8.647804]  
7: 4560007 Fitzgibbon_Cup  
    [pem_pos=0, final_scores=0.13294469,  
final_local_scores=0.11704269, pem=0.0, similarity=7.5825]  
8: 2794983 Clare_GAA  
    [pem_pos=0, final_scores=0.07254173,  
final_local_scores=-0.00050009787, pem=-1.1644745,  
similarity=10.407382]  
9: 5795114 Davy_Fitzgerald  
    [pem_pos=0, final_scores=0.5838863,  
final_local_scores=0.5766894, pem=0.0, similarity=21.299276]  
10: 6837581 James_Stephens_GAA  
    [pem_pos=1, final_scores=0.1797627,  
final_local_scores=0.1609261, pem=-0.7831714,  
similarity=13.151077]  
11: 16715841 Portumna_GAA  
    [pem_pos=1, final_scores=0.067021355,  
final_local_scores=0.035630614, pem=-0.35438493,  
similarity=7.0802016]  
12: 2795066 Kilkenny_GAA  
    [pem_pos=1, final_scores=0.19264083,  
final_local_scores=0.1204896, pem=-1.0248141,  
similarity=13.258461]  
13: 2810897 Wexford_GAA  
    [pem_pos=1, final_scores=0.17291714,  
final_local_scores=0.10705638, pem=-1.1842172,
```

similarity=13.724446]
14: 2759665 Dublin_GAA
[pem_pos=1, final_scores=0.15838693,
final_local_scores=0.08209327, pem=-2.538983,
similarity=20.346926]
15: 2810853 Galway_GAA
[pem_pos=1, final_scores=0.17547093,
final_local_scores=0.102746904, pem=-1.1811876,
similarity=13.579368]
16: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.13535754,
final_local_scores=0.06990652, pem=-0.83555937,
similarity=10.719765]
17: 2802645 Waterford_GAA
[pem_pos=1, final_scores=0.14126547,
final_local_scores=0.07925248, pem=-1.8432962,
similarity=16.478895]
18: 5575719 Eddie_Brennan
[pem_pos=0, final_scores=0.41984683,
final_local_scores=0.422283, pem=0.0, similarity=16.69148]
19: 5287742 Henry_Shefflin
[pem_pos=0, final_scores=0.49723524,
final_local_scores=0.4947896, pem=0.0, similarity=18.85522]
20: 2802645 Waterford_GAA
[pem_pos=1, final_scores=0.19060723,
final_local_scores=0.13315088, pem=-1.8432962,
similarity=18.087332]
21: 2802637 Tipperary_GAA
[pem_pos=0, final_scores=0.20338796,
final_local_scores=0.13258591, pem=-0.63646483,
similarity=11.507536]
22: 2802637 Tipperary_GAA
[pem_pos=0, final_scores=0.26119906,
final_local_scores=0.19573575, pem=-0.63646483,
similarity=13.392053]
23: 2759665 Dublin_GAA
[pem_pos=1, final_scores=0.2103763,
final_local_scores=0.13888374, pem=-2.538983,
similarity=22.041666]
24: 2810853 Galway_GAA
[pem_pos=1, final_scores=0.10164051,
final_local_scores=0.022114158, pem=-1.1811876,
similarity=11.173125]
25: 12949356 Joe_Canning
[pem_pos=0, final_scores=0.353247,
final_local_scores=0.3356698, pem=0.0, similarity=14.106769]
26: 35139982 Niall_Burke
[pem_pos=0, final_scores=0.31546706,
final_local_scores=0.326652, pem=0.0, similarity=13.83766]
1: 14560 Republic_of_Ireland
[pem_pos=3, final_scores=0.3304627,
final_local_scores=0.3824578, pem=-2.2576945,

similarity=27.78071]
2: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.5376831,
final_local_scores=0.5133141, pem=0.0, similarity=19.40803]
3: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.20260297,
final_local_scores=0.13118339, pem=-1.0248141,
similarity=13.577585]
4: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.107422605,
final_local_scores=0.08209887, pem=-1.3755379,
similarity=14.020096]
5: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.20235415,
final_local_scores=0.14341193, pem=-0.83555937,
similarity=12.913313]
6: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.41720426,
final_local_scores=0.38170928, pem=0.0, similarity=15.480679]
7: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.16431175,
final_local_scores=0.10153459, pem=-0.83555937,
similarity=11.6636095]
8: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.14117165,
final_local_scores=0.11898634, pem=-1.3755379,
similarity=15.120893]
9: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.4493478,
final_local_scores=0.4171127, pem=0.0, similarity=16.537188]
10: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.35373503,
final_local_scores=0.31237876, pem=0.0, similarity=13.411718]
11: 3380379 Leinster_GAA
[pem_pos=2, final_scores=0.07148351,
final_local_scores=0.027502716, pem=-1.4623435,
similarity=12.8629]
12: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.1528322,
final_local_scores=0.088994935, pem=-0.83555937,
similarity=11.289401]
13: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.47166914,
final_local_scores=0.4412039, pem=0.0, similarity=17.256117]
14: 342015 Tyrrell_Racing
[pem_pos=1, final_scores=0.45529586,
final_local_scores=0.5499012, pem=-0.054408323,
similarity=20.795742]
15: 4560007 Fitzgibbon_Cup
[pem_pos=0, final_scores=0.16449855,
final_local_scores=0.15151054, pem=0.0, similarity=8.61109]
16: 342015 Tyrrell_Racing

[pem_pos=1, final_scores=0.29899883,
final_local_scores=0.37917042, pem=-0.054408323,
similarity=15.700796]
17: 342015 Tyrrell_Racing
[pem_pos=1, final_scores=0.42870063,
final_local_scores=0.52084994, pem=-0.054408323,
similarity=19.928795]
18: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.45791227,
final_local_scores=0.42617655, pem=0.0, similarity=16.80767]
19: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.077520385,
final_local_scores=0.049435228, pem=-1.3755379,
similarity=13.045348]
20: 1234706 University_College_Dublin_A.F.C.
[pem_pos=0, final_scores=0.15514006,
final_local_scores=0.11857185, pem=-0.8758797,
similarity=12.391303]
21: 6837581 James_Stephens_GAA
[pem_pos=1, final_scores=0.18816389,
final_local_scores=0.1701031, pem=-0.7831714,
similarity=13.424937]
22: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.12895252,
final_local_scores=0.062910005, pem=-0.83555937,
similarity=10.510975]
23: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.47553605,
final_local_scores=0.4454279, pem=0.0, similarity=17.38217]
24: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.09899174,
final_local_scores=0.030182406, pem=-0.83555937,
similarity=9.534318]
25: 6837581 James_Stephens_GAA
[pem_pos=1, final_scores=0.104464546,
final_local_scores=0.07867429, pem=-0.7831714,
similarity=10.696519]
26: 342015 Tyrrell_Racing
[pem_pos=1, final_scores=0.46584255,
final_local_scores=0.5614219, pem=-0.054408323,
similarity=21.139544]
27: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.06804781,
final_local_scores=0.039109647, pem=-1.3755379,
similarity=12.737212]
28: 1234706 University_College_Dublin_A.F.C.
[pem_pos=0, final_scores=0.30479306,
final_local_scores=0.28204498, pem=-0.8758797,
similarity=17.269667]
29: 6837581 James_Stephens_GAA
[pem_pos=1, final_scores=0.15267293,
final_local_scores=0.13133463, pem=-0.7831714,

similarity=12.268008]
30: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.059148148,
final_local_scores=-0.013971433, pem=-0.83555937,
similarity=8.21668]
31: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.31621772,
final_local_scores=0.27139685, pem=0.0, similarity=12.188736]
32: 7638195 Ballyhale_Shamrocks_GAA
[pem_pos=0, final_scores=0.26309198,
final_local_scores=0.22635266, pem=0.0, similarity=10.844528]
33: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.4828059,
final_local_scores=0.45336908, pem=0.0, similarity=17.61915]
34: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.18795334,
final_local_scores=0.1153692, pem=-1.0248141,
similarity=13.105659]
35: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.45197612,
final_local_scores=0.41969228, pem=0.0, similarity=16.614168]
36: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.16786595,
final_local_scores=0.14817062, pem=-1.3755379,
similarity=15.991809]
37: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.1971129,
final_local_scores=0.12518632, pem=-1.0248141,
similarity=13.398621]
38: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.50094086,
final_local_scores=0.4731788, pem=0.0, similarity=18.210312]
39: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.06572385,
final_local_scores=-0.0067884773, pem=-0.83555937,
similarity=8.431033]
40: 342015 Tyrrell_Racing
[pem_pos=1, final_scores=0.346888,
final_local_scores=0.43148208, pem=-0.054408323,
similarity=17.26188]
41: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.18733315,
final_local_scores=0.114691734, pem=-1.0248141,
similarity=13.085442]
42: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.49099344,
final_local_scores=0.46231276, pem=0.0, similarity=17.886047]
43: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.2803597,
final_local_scores=0.21612078, pem=-1.0248141,
similarity=16.112288]
44: 2795066 Kilkenny_GAA

[pem_pos=1, final_scores=0.23364814,
final_local_scores=0.16509551, pem=-1.0248141,
similarity=14.589591]
45: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.45200497,
final_local_scores=0.41972375, pem=0.0, similarity=16.615107]
46: 2794983 Clare_GAA
[pem_pos=0, final_scores=0.3530383,
final_local_scores=0.3065167, pem=-1.1644745,
similarity=19.569376]
47: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.25774956,
final_local_scores=0.19142264, pem=-1.0248141,
similarity=15.375245]
48: 2810897 Wexford_GAA
[pem_pos=1, final_scores=0.31171852,
final_local_scores=0.2586758, pem=-1.1842172,
similarity=18.249073]
49: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.06956531,
final_local_scores=0.040745527, pem=-1.3755379,
similarity=12.78603]
50: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.13747425,
final_local_scores=0.0722187, pem=-0.83555937,
similarity=10.788765]
51: 32532531 Cork_GAA
[pem_pos=1, final_scores=0.1052091,
final_local_scores=0.031399637, pem=-1.0852067,
similarity=10.928264]
52: 2810853 Galway_GAA
[pem_pos=1, final_scores=0.17688204,
final_local_scores=0.10428831, pem=-1.1811876,
similarity=13.625366]
53: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.2548532,
final_local_scores=0.18825883, pem=-1.0248141,
similarity=15.280831]
54: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.40414876,
final_local_scores=0.36744815, pem=0.0, similarity=15.0550995]
55: 1234583 Limerick_F.C.
[pem_pos=4, final_scores=0.28157276,
final_local_scores=0.26514578, pem=-2.6802762,
similarity=26.57795]
56: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.14226757,
final_local_scores=0.12016168, pem=-1.3755379,
similarity=15.155966]
57: 2810897 Wexford_GAA
[pem_pos=1, final_scores=0.28176695,
final_local_scores=0.22595829, pem=-1.1842172,

similarity=17.272717]
58: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.09747945,
final_local_scores=0.028530434, pem=-0.83555937,
similarity=9.48502]
59: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.15975033,
final_local_scores=0.08437337, pem=-1.0248141,
similarity=12.180681]
60: 7759007 Jack_Cork
[pem_pos=5, final_scores=0.087552145,
final_local_scores=0.112473786, pem=-3.552999,
similarity=26.767916]
61: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.053270534,
final_local_scores=-0.020391807, pem=-0.83555937,
similarity=8.025083]
62: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.15041511,
final_local_scores=0.07436441, pem=-1.0248141,
similarity=11.881994]
63: 32532531 Cork_GAA
[pem_pos=1, final_scores=0.12632956,
final_local_scores=0.05447054, pem=-1.0852067,
similarity=11.616746]
64: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.12214242,
final_local_scores=0.0434808, pem=-1.0248141,
similarity=10.960365]
65: 32532531 Cork_GAA
[pem_pos=1, final_scores=0.1157368,
final_local_scores=0.04289955, pem=-1.0852067,
similarity=11.271444]
66: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.096465155,
final_local_scores=0.027422503, pem=-0.83555937,
similarity=9.451957]
67: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.55219334,
final_local_scores=0.5291644, pem=0.0, similarity=19.881033]
68: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.15899001,
final_local_scores=0.13845018, pem=-1.3755379,
similarity=15.701731]
69: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.25455713,
final_local_scores=0.18793541, pem=-1.0248141,
similarity=15.27118]
70: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.07669942,
final_local_scores=0.0058314353, pem=-0.83555937,
similarity=8.807636]

71: 1234583 Limerick_F.C.
[pem_pos=4, final_scores=0.17101337,
final_local_scores=0.14439487, pem=-2.6802762,
similarity=22.974503]
72: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.1838011,
final_local_scores=0.11083354, pem=-1.0248141,
similarity=12.970305]
73: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.46301407,
final_local_scores=0.43174952, pem=0.0, similarity=16.97398]
74: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.07443555,
final_local_scores=0.0033584982, pem=-0.83555937,
similarity=8.733839]
75: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.26509947,
final_local_scores=0.19945127, pem=-1.0248141,
similarity=15.614836]
76: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.037328497,
final_local_scores=-0.03780608, pem=-0.83555937,
similarity=7.5054064]
77: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.50270516,
final_local_scores=0.47510606, pem=0.0, similarity=18.267826]
78: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.08730416,
final_local_scores=0.06012252, pem=-1.3755379,
similarity=13.364278]
79: 2810897 Wexford_GAA
[pem_pos=1, final_scores=0.24177341,
final_local_scores=0.18227142, pem=-1.1842172,
similarity=15.969012]
80: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.12695344,
final_local_scores=0.04854776, pem=-1.0248141,
similarity=11.111573]
81: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.46199238,
final_local_scores=0.4306335, pem=0.0, similarity=16.940676]
82: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.46159822,
final_local_scores=0.43020296, pem=0.0, similarity=16.927828]
83: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.15056331,
final_local_scores=0.12924531, pem=-1.3755379,
similarity=15.42704]
84: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.24419986,
final_local_scores=0.17662168, pem=-1.0248141,
similarity=14.933556]

85: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.08195038,
final_local_scores=0.0115673095, pem=-0.83555937,
similarity=8.9788065]
86: 5287742 Henry_Shefflin
[pem_pos=0, final_scores=0.48887306,
final_local_scores=0.4856552, pem=0.0, similarity=18.582632]
87: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.48406088,
final_local_scores=0.45474, pem=0.0, similarity=17.66006]
88: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.1506633,
final_local_scores=0.08662571, pem=-0.83555937,
similarity=11.2186985]
89: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.39560694,
final_local_scores=0.3581175, pem=0.0, similarity=14.776653]
90: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.043836728,
final_local_scores=0.012640953, pem=-1.3755379,
similarity=11.947333]
91: 2810853 Galway_GAA
[pem_pos=1, final_scores=0.24625693,
final_local_scores=0.18006983, pem=-1.1811876,
similarity=15.886838]
92: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.2656647,
final_local_scores=0.20006871, pem=-1.0248141,
similarity=15.633261]
93: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.12819283,
final_local_scores=0.06208016, pem=-0.83555937,
similarity=10.486211]
94: 2802637 Tipperary_GAA
[pem_pos=0, final_scores=0.36375177,
final_local_scores=0.307759, pem=-0.63646483,
similarity=16.73505]
95: 9582355 Jackie_Tyrrell
[pem_pos=0, final_scores=0.48384935,
final_local_scores=0.4545089, pem=0.0, similarity=17.653166]
96: 985961 Leinster_Rugby
[pem_pos=1, final_scores=0.13523932,
final_local_scores=0.11250618, pem=-1.3755379,
similarity=14.92751]
97: 2795066 Kilkenny_GAA
[pem_pos=1, final_scores=0.24690937,
final_local_scores=0.1795814, pem=-1.0248141,
similarity=15.021879]
98: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.06304185,
final_local_scores=-0.009718135, pem=-0.83555937,
similarity=8.343606]

99: 9582355 Jackie_Tyrrell
 [pem_pos=0, final_scores=0.39551127,
 final_local_scores=0.35801297, pem=0.0, similarity=14.773534]
 100: 2802637 Tipperary_GAA
 [pem_pos=0, final_scores=0.22036426,
 final_local_scores=0.15112993, pem=-0.63646483,
 similarity=12.060926]
 101: 2795066 Kilkenny_GAA
 [pem_pos=1, final_scores=0.16282238,
 final_local_scores=0.08791746, pem=-1.0248141,
 similarity=12.286444]
 102: 2795066 Kilkenny_GAA
 [pem_pos=1, final_scores=0.24030231,
 final_local_scores=0.17236418, pem=-1.0248141,
 similarity=14.806502]
 103: 2795066 Kilkenny_GAA
 [pem_pos=1, final_scores=0.18937181,
 final_local_scores=0.11673033, pem=-1.0248141,
 similarity=13.146278]
 104: 32532531 Cork_GAA
 [pem_pos=1, final_scores=0.11002825,
 final_local_scores=0.036595523, pem=-1.0852067,
 similarity=11.08332]
 105: 9582355 Jackie_Tyrrell
 [pem_pos=0, final_scores=0.43336695,
 final_local_scores=0.3993646, pem=0.0, similarity=16.00755]
 106: 985961 Leinster_Rugby
 [pem_pos=1, final_scores=0.13070352,
 final_local_scores=0.107551515, pem=-1.3755379,
 similarity=14.779654]
 107: 2810853 Galway_GAA
 [pem_pos=1, final_scores=0.16205676,
 final_local_scores=0.08810973, pem=-1.1811876,
 similarity=13.142565]

Sample: b'Hong_Kong_Family_Welfare_Society'

The [1 Hong Kong Family Welfare Society 1] -LRB- HKFWS ; -
 RRB- is a 1 non-governmental organization that runs on 2 subs
 idies donated by the government , donations from [1 1 The Com
 munity Chest of Hong Kong 1] and individual 3 sponsor ships .
 HKFWS focuses on social services through a family perspective
 that aims to enhance the 4 well-being of families and
 individuals in [2 Hong Kong 2] . Services such as counseling
 are used to help reconcile issues related to 5 family , 6 mar
 riage , and 7 friendship to help prevent family breakups and
 other societal problems . The HKFWS now has over 30 service
 centers and networks helping families across [3 Hong Kong
 3] . HKFWS 's official 8 mission statement is : `` With a
 family perspective , we are committed to delivering quality
 and professional social services to enhance the well-being of
 families and individuals in [4 Hong Kong 4] and to foster a
 caring 9 community . '' HKFWS pioneers service projects to

meet changing needs of the people and advocates a comprehensive 10 public policy for needy families . The `` [5 Hong Kong Family Welfare Society 5] '' was formally known as the Social Service Centre of the Churches . It was founded in 1938 , but only became a public agency on January 1 , 1949 . The Society started out by being heavily involved with relief work after [2 2 World War II 2] . After concerns for the growth of distressed families resonated with the members , the Society further expanded its scope . Today the bulk of their work revolves around counselling families with management difficulties or communication issues . In addition , the Society focuses on preventive measures for family problems , resulting in the creation of Family Life Education in 1971 . Other services like Home Help , School Social Work services , Clinical Psychological services and Foster Care came about in 1971 , 1979 and 1984 respectively . HKFWS services include : # Family resource centre in [3 3 Tuen Mun 3] # ` Drop-in ' services for single parents # Mental health services # Senior social work practice # Therapeutic group for abusive families # Emergency foster care # Volunteer service # Family aid services # Services for new arrivals from the 11 Mainland # Carers support service # Peer /divorce/family mediation service # Family retreat centre # Clinical psychological service # Youth employment # Community rehabilitation and pain relief service One programme since 1997 is the Integrated Elderly and Community Support Service to alleviate the problems associated with an . Between 2002 and 2004 , the Society made help accessible by piloting programs on Integrated Family Service Centres at its [4 4 West Kowloon 4] , [6 12 East Kowloon 6] -LRB- Shun Lee -RRB- and [7 13 New Territories 7] -LRB- Kwai Fong -RRB- Centres . For information on these centres : One of the methods of relieving tension in families is the use of mediation . This is vital due to the sharp increase in divorce rates in [8 Hong Kong 8] and the lack of communication between parents and children . Mediation is used in a bid to help salvage the family situation . It teaches individuals to reach a settlement acceptable to all parties . The Society also provides five types of services and programs namely ; Special and Innovative Services , Children and Youth , Elderly and Community , Family and Child Care . The Special and Innovative Services includes programmes for employment and fund assistance . HKFWS also has a subsidiary association called the where seminars and mediation services are available to the public . Due to the nature of the HKFWS as a , the Society relies largely on donations from the public and even the government . Important sources are the , [5 5 The Hong Kong Jockey Club Charities Trust 5] , civic minded companies , 14 charitable foundations and individual sponsorships . The HKFWS is made up of employees and a team of volunteers . The Society is composed of the following committees : A detailed list of the structure and committee

members is here : The Society encourages individuals to become volunteers . The function of the Corporate Volunteer Service is to link voluntary organizations and service users to volunteers . Over the years , many corporations , schools , non-profit organizations and government bodies have played their part in providing services to the HKFWS service users that consists of ethnic minorities , children , elderly , and foster care families . The organization seeks to advocate family volunteering so that family members will share experiences thus producing solidarity . At present , the revamped services operated by the Society includes five Integrated Family Service Centres , eight Integrated Home Care Teams -LRB- in 15 centres -RRB- , one Integrated Children and Youth Service Centre and one Neighbourhood Elderly Centre . The HKFWS and its branches have participated in campaigns co-organized by Ho-Sum that has allowed members to interact with the elderly , handicapped , young children , youths and volunteers . These campaigns are aimed at building relationships between family units and the [9 Hong Kong 9] society . To provide short-term financial assistance to children from families in need so that they can enjoy proper care , full development and have growth in physical , psychological , social and learning aspects . More information : The Society has a strong commitment to research . To date , 11 research studies have been conducted on the services of School , Family Casework , Integrated Family Service Centres , Child Protection Registry , Alternative to Violence Project , Peer 15 Mediation , and various problems involving 16 secondary school and 17 primary school

1: 1196650 The_Community_Chest_of_Hong_Kong

2: 32927 World_War_II

3: 2970235 Tuen_Mun

4: 198761 Kowloon

5: 323884 Hong_Kong_Jockey_Club

1: 46539 Non-governmental_organization

[winner: 23887135 Church_of_Christ_(Fettingite),

pem_pos=7, final_scores=-1.50272,

final_local_scores=-1.4989879, pem=-1.9454916,

similarity=-30.063194]

gt: 46539 Non-governmental_organization, pem_pos=2,

final_scores=-1.8500389, final_local_scores=-2.004321,

pem=-1.9305413, similarity=-45.22464]

[winner: 46539 Non-governmental_organization, pem_pos=0,

final_scores=-1.5782634, final_local_scores=-1.7074478,

pem=0.0, similarity=-46.863937]

gt: 46539 Non-governmental_organization, pem_pos=0,

final_scores=-1.5782634, final_local_scores=-1.7074478,

pem=0.0, similarity=-46.863937]

2: 146719 Subsidy

[winner: 146719 Subsidy, pem_pos=0,

final_scores=-1.1708211, final_local_scores=-1.2414088,

pem=-0.06159302, similarity=-32.62145|
gt: 146719 Subsidy, pem_pos=0, final_scores=-1.1708211,
final_local_scores=-1.2414088, pem=-0.06159302,
similarity=-32.62145]
3: 545607 Sponsor_(commercial)
[winner: 545607 Sponsor_(commercial), pem_pos=0,
final_scores=-1.3975295, final_local_scores=-1.5024233,
pem=0.0, similarity=-40.745598|
gt: 545607 Sponsor_(commercial), pem_pos=0,
final_scores=-1.3975295, final_local_scores=-1.5024233,
pem=0.0, similarity=-40.745598]
4: 47789 Quality_of_life
[winner: 990505 Mental_health, pem_pos=0,
final_scores=-1.6436137, final_local_scores=-1.7759879,
pem=0.0, similarity=-48.90931|
gt: 47789 Quality_of_life, pem_pos=1,
final_scores=-1.6504277, final_local_scores=-1.7553821,
pem=-0.36034116, similarity=-46.334797]
5: 10401954 Family
[winner: 56276 Family_(biology), pem_pos=0,
final_scores=-1.1979442, final_local_scores=-1.1491058,
pem=-0.29807892, similarity=-28.580896|
gt: 10401954 Family, pem_pos=1, final_scores=-1.8773067,
final_local_scores=-2.0029194, pem=-1.900655,
similarity=-45.345337]
6: 19728 Marriage
[winner: 19728 Marriage, pem_pos=0,
final_scores=-1.4313836, final_local_scores=-1.5155399,
pem=-0.14008214, similarity=-40.375233|
gt: 19728 Marriage, pem_pos=0, final_scores=-1.4313836,
final_local_scores=-1.5155399, pem=-0.14008214,
similarity=-40.375233]
7: 17157119 Friendship
[winner: 161744 Interpersonal_relationship, pem_pos=0,
final_scores=-1.4244657, final_local_scores=-1.4857707,
pem=0.0, similarity=-40.24865|
gt: 17157119 Friendship, pem_pos=1,
final_scores=-1.6730657, final_local_scores=-1.764648,
pem=-0.018108804, similarity=-48.472427]
8: 7623580 Mission_statement
[winner: 938656 Statement_(computer_science), pem_pos=0,
final_scores=-1.2068189, final_local_scores=-1.1622033,
pem=-0.37838498, similarity=-28.535036|
RECALL MISS: 7623580 Mission_statement]
[winner: 22673654 Mission_California, pem_pos=11,
final_scores=-1.2545049, final_local_scores=-1.1682587,
pem=-3.2960024, similarity=-12.849283|
gt: 7623580 Mission_statement, pem_pos=18,
final_scores=-1.9814842, final_local_scores=-2.1230855,
pem=-3.782471, similarity=-38.69773]
9: 5695 Community
[winner: 1324897 Residential_community, pem_pos=3,

final_scores=-1.4682918, final_local_scores=-1.5254252,
pem=-2.9847856, similarity=-25.200285|
gt: 5695 Community, pem_pos=0, final_scores=-1.5884434,
final_local_scores=-1.7243059, pem=-0.43085614,
similarity=-45.023952]
10: 153324 Public_policy
[winner: 12982164 Good_Neighbor_policy_(LDS_Church),
pem_pos=7, final_scores=-1.0801228,
final_local_scores=-0.99710155, pem=-1.9459101,
similarity=-15.083625|
gt: 153324 Public_policy, pem_pos=11,
final_scores=-1.9779476, final_local_scores=-2.1384823,
pem=-3.9743268, similarity=-38.113857]
[winner: 5687925 Public_(album), pem_pos=5,
final_scores=-1.3985697, final_local_scores=-1.3441598,
pem=-2.7873955, similarity=-20.864405|
RECALL MISS: 153324 Public_policy]
11: 55079 Mainland_China
[winner: 55079 Mainland_China, pem_pos=2,
final_scores=-0.118798, final_local_scores=-0.1928486,
pem=-1.3569025, similarity=5.713773|
gt: 55079 Mainland_China, pem_pos=2,
final_scores=-0.118798, final_local_scores=-0.1928486,
pem=-1.3569025, similarity=5.713773]
12: 5163958 Shun_Lee
[winner: 198761 Kowloon, pem_pos=0,
final_scores=0.1144665, final_local_scores=0.050481245,
pem=-0.2006707, similarity=6.6874523|
RECALL MISS: 5163958 Shun_Lee]
[winner: 5163958 Shun_Lee, pem_pos=0,
final_scores=-0.107989654, final_local_scores=-0.056330055,
pem=0.0, similarity=2.4087117|
gt: 5163958 Shun_Lee, pem_pos=0,
final_scores=-0.107989654, final_local_scores=-0.056330055,
pem=0.0, similarity=2.4087117]
13: 230930 New_Territories
[winner: 230930 New_Territories, pem_pos=0,
final_scores=0.2141171, final_local_scores=0.18302202,
pem=0.0, similarity=9.5514555|
gt: 230930 New_Territories, pem_pos=0,
final_scores=0.2141171, final_local_scores=0.18302202,
pem=0.0, similarity=9.5514555]
[winner: 741070 Kwai_Fong_Station, pem_pos=0,
final_scores=-0.13164532, final_local_scores=-0.087075345,
pem=0.0, similarity=1.4912109|
RECALL MISS: 230930 New_Territories]
14: 1176679 Charitable_organization
[winner: 693334 Foundation_(engineering), pem_pos=0,
final_scores=-1.2886184, final_local_scores=-1.3032551,
pem=-0.3597419, similarity=-32.845684|
RECALL MISS: 1176679 Charitable_organization]
[winner: 1931801 Principle_of_charity, pem_pos=3,

```

final_scores=-1.454204, final_local_scores=-1.4138119,
pem=-2.8031745, similarity=-22.857159|
    gt: 1176679 Charitable_organization, pem_pos=0,
final_scores=-1.5009046, final_local_scores=-1.6568086,
pem=0.0, similarity=-45.35276 ]
15: 20582 Mediation
    [winner: 20582 Mediation, pem_pos=0,
final_scores=-0.10605264, final_local_scores=-0.1014913,
pem=0.0, similarity=1.0610101|
    gt: 20582 Mediation, pem_pos=0,
final_scores=-0.10605264, final_local_scores=-0.1014913,
pem=0.0, similarity=1.0610101 ]
16: 42556 High_school
    [winner: 10678806 Klaipėda "Varpas" _secondary_school,
pem_pos=19, final_scores=-0.9894769,
final_local_scores=-0.8737524, pem=-2.944439,
similarity=-5.972486|
    gt: 42556 High_school, pem_pos=2, final_scores=-1.68321,
final_local_scores=-1.8134964, pem=-2.4730551,
similarity=-36.579773 ]
17: 11722032 Primary_education
    [winner: 4646029 Python_primary, pem_pos=11,
final_scores=-1.1661534, final_local_scores=-1.0974295,
pem=-2.4837437, similarity=-15.152788|
    gt: 11722032 Primary_education, pem_pos=1,
final_scores=-1.3758533, final_local_scores=-1.5023396,
pem=-1.2497656, similarity=-33.94668 ]
    [winner: 28022 School, pem_pos=0, final_scores=-1.30567,
final_local_scores=-1.4116848, pem=-0.2358994,
similarity=-36.75492|
    RECALL MISS: 11722032 Primary_education]
1: 1196650 The_Community_Chest_of_Hong_Kong
    [pem_pos=0, final_scores=0.1726963,
final_local_scores=0.20880327, pem=0.0, similarity=10.320819]
2: 32927 World_War_II
    [pem_pos=0, final_scores=0.37458748,
final_local_scores=0.42157924, pem=0.0, similarity=16.670479]
3: 2970235 Tuen_Mun
    [pem_pos=0, final_scores=0.09294145,
final_local_scores=0.08274937, pem=-0.6333592,
similarity=10.003425]
4: 198761 Kowloon
    [pem_pos=1, final_scores=0.09683557,
final_local_scores=0.035658807, pem=-0.5757712,
similarity=8.284976]
5: 323884 Hong_Kong_Jockey_Club
    [pem_pos=0, final_scores=0.15048002,
final_local_scores=0.13738799, pem=0.0, similarity=8.189645]
1: 7566750 Hong_Kong_Family_Welfare_Society
    [pem_pos=0, final_scores=0.13367544,
final_local_scores=0.15858975, pem=0.0, similarity=8.822348]
2: 13404 Hong_Kong

```


[pem_pos=0, final_scores=0.5028259,
 final_local_scores=0.4578011, pem=-0.22402541,
 similarity=18.969696]
 3: 13404 Hong_Kong
 [pem_pos=0, final_scores=0.45859343,
 final_local_scores=0.4094838, pem=-0.22402541,
 similarity=17.527811]
 4: 13404 Hong_Kong
 [pem_pos=0, final_scores=0.5038337,
 final_local_scores=0.45890194, pem=-0.22402541,
 similarity=19.002548]
 5: 7566750 Hong_Kong_Family_Welfare_Society
 [pem_pos=0, final_scores=0.10002421,
 final_local_scores=0.12183088, pem=0.0, similarity=7.725389]
 6: 198761 Kowloon
 [pem_pos=0, final_scores=0.1144665,
 final_local_scores=0.050481245, pem=-0.2006707,
 similarity=6.6874523]
 7: 230930 New_Territories
 [pem_pos=0, final_scores=0.2141171,
 final_local_scores=0.18302202, pem=0.0, similarity=9.5514555]
 8: 13404 Hong_Kong
 [pem_pos=0, final_scores=0.49922413,
 final_local_scores=0.45386666, pem=-0.22402541,
 similarity=18.852285]
 9: 13404 Hong_Kong
 [pem_pos=0, final_scores=0.50599366,
 final_local_scores=0.46126133, pem=-0.22402541,
 similarity=19.072956]

Sample: b'CA_Saint-Etienne_Loire_Sud_Rugby'

Club Athlétique de Saint-Étienne Loire Sud Rugby , CASE
 Loire Sud Rugby , or simply CASE , is a [1 1 French 1] 1 rugby
 union club from [2 2 Saint-Étienne 2] in the [3 3 Loire 3] 2 d
 épartement . They are currently competing in 3 Fédérale 1 ,
 having been relegated from [4 4 Rugby Pro D2 4] at the end of
 the 4 2010 -- 11 season . The club was founded in 1908 under
 the name Stade Forézien Universitaire . They went through a
 difficult time in the 1960s , and had to merge with the CASE
 omnisport structure in 1975 at the request of the
 municipality . It did not help turn back the tide , however ,
 and poor results and fewer and fewer licensed players forced
 the club to withdraw its senior team from all competitions in
 1991 . The next year , the club was restarted in 4ème Série ,
 the lowest [1 French 1] rugby division , and was able to
 climb the ladder , gaining promotion to 5 Fédérale 3 in
 2004 , one year after merging with the neighbouring club
 SEMUR and adopting its current name . In 2007 , [2 Saint-
 Etienne 2] were promoted to 6 Fédérale 2 . A mere two years
 later , they finished 13th overall in the championship , with
 the top 12 clubs being promoted to Fédérale 1 . However ,
 declined their promotion , sending [3 Saint-Étienne 3] to the

top amateur division for the first time . For their first season in Fédérale 1 , [4 Saint-Étienne 4] signed the coach from 7 Arras , and the 8 international fullback [5 5 Alexandre Péclier 5] from [5 9 Lyon 5] . They had an impressive start with two away wins in their first two games , and on November 7 , they took on 10 Béziers in 11 Andrézieux-Bouthéon 's . The 3,000 people in the stadium were not the only ones to witness the locals ' 15 -- 9 victory though , as the match was broadcast live on [6 6 Eurosport 6] , a first for a 12 Fédérale 1 game . This run of good form allowed them to clinch the fourth place in their group and thus to qualify for the Trophée Jean Prat . In the second phase , they finished second behind [7 7 Tyrosse 7] and qualified for the Round of 16 where they faced [6 13 Nice 6] . They prevailed 19 -- 18 thanks to a last-second drop-goal by [7 Péclier 7] and faced 1 4 Chalon-sur-Saône , one of the favorites for promotion , in the quarters . [8 Saint-Étienne 8] stunned them by winning both legs -LRB- 12 -- 3 at home , 24 -- 20 away -RRB- , inflicting them their only home defeat of the season in the process . In the semis , they took on [8 8 Bourg-en-Bresse 8] . The popular support for the team was such that the municipality allowed the first leg to be played in [9 9 Stade Geoffroy-Guichard 9] . Over 16,000 people came to see the locals take the slenderest of leads into the second leg with a 23 -- 22 victory . Another shock 19 -- 16 win in [9 Bourg-en-Bresse 9] sent [10 Saint-Étienne 10] in the Jean Prat Trophy final , but much more importantly gave them a ticket to professional rugby and the [10 10 Pro D2 10] , after only one year spent in Fédérale 1 , and perhaps more impressively only three years after leaving Fédérale 3 . They fell in the final to 15 Carcassonne

1: 5843419 France

2: 214482 Saint-Étienne

3: 31464821 Loire

4: 5992163 Rugby_Pro_D2

5: 10120820 Alexandre_Péclier

6: 629239 Eurosport

7: 8332924 Union_Sportive_Tyrosse_Rugby_Côte_Sud

8: 17965716 Union_Sportive_Bressane

9: 3387365 Stade_Geoffroy-Guichard

10: 5992163 Rugby_Pro_D2

1: 25405 Rugby_union

[winner: 19774201 Rugby_Virginia, pem_pos=7, final_scores=-0.96575636, final_local_scores=-0.819801, pem=-2.956099, similarity=-4.299058]

gt: 25405 Rugby_union, pem_pos=3, final_scores=-2.3532212, final_local_scores=-2.5809233, pem=-2.0110612, similarity=-61.99372]

[winner: 481130 Union_(American_Civil_War), pem_pos=0, final_scores=-1.4707228, final_local_scores=-1.4919705, pem=-1.132797, similarity=-34.273335]

RECALL MISS: 25405 Rugby_union]

2: 38523 Departments_of_France
[winner: 38523 Departments_of_France, pem_pos=0,
final_scores=-1.5027506, final_local_scores=-1.5878055,
pem=-0.05971194, similarity=-42.96885|
gt: 38523 Departments_of_France, pem_pos=0,
final_scores=-1.5027506, final_local_scores=-1.5878055,
pem=-0.05971194, similarity=-42.96885]
3: 5991818 Fédérale_1
[winner: 5991818 Fédérale_1, pem_pos=0,
final_scores=-0.16368267, final_local_scores=-0.22163907,
pem=0.0, similarity=-2.524439|
gt: 5991818 Fédérale_1, pem_pos=0,
final_scores=-0.16368267, final_local_scores=-0.22163907,
pem=0.0, similarity=-2.524439]
4: 27702503 2010-11_Rugby_Pro_D2_season
[winner: 19452 May_11, pem_pos=2,
final_scores=-0.7947425, final_local_scores=-0.5424653,
pem=-3.271949, similarity=5.694833|
RECALL MISS: 27702503 2010-11_Rugby_Pro_D2_season]
[winner: 43226 2010, pem_pos=0, final_scores=-0.9975369,
final_local_scores=-0.8806697, pem=-1.6731367,
similarity=-13.09245|
RECALL MISS: 27702503 2010-11_Rugby_Pro_D2_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.2694834, final_local_scores=-1.1704717,
pem=0.0, similarity=-30.839493|
RECALL MISS: 27702503 2010-11_Rugby_Pro_D2_season]
[winner: 690345 Season(sports), pem_pos=0,
final_scores=-1.4009665, final_local_scores=-1.4903698,
pem=-1.219945, similarity=-33.751644|
RECALL MISS: 27702503 2010-11_Rugby_Pro_D2_season]
5: 10137842 Fédérale_3
[winner: 10137842 Fédérale_3, pem_pos=0,
final_scores=-0.02936475, final_local_scores=-0.03259495,
pem=0.0, similarity=3.1170146|
gt: 10137842 Fédérale_3, pem_pos=0,
final_scores=-0.02936475, final_local_scores=-0.03259495,
pem=0.0, similarity=3.1170146]
6: 10137757 Fédérale_2
[winner: 10137757 Fédérale_2, pem_pos=0,
final_scores=-0.10986096, final_local_scores=-0.12760767,
pem=0.0, similarity=0.28164533|
gt: 10137757 Fédérale_2, pem_pos=0,
final_scores=-0.10986096, final_local_scores=-0.12760767,
pem=0.0, similarity=0.28164533]
7: 24542134 Rugby_Club_d'Arras
[winner: 105054 Arras, pem_pos=0,
final_scores=-0.21344802, final_local_scores=-0.22638692,
pem=-0.13882738, similarity=-1.9111593|
gt: 24542134 Rugby_Club_d'Arras, pem_pos=5,
final_scores=-0.41040462, final_local_scores=-0.39033443,
pem=-2.0676122, similarity=3.6853514]

8: 955656 France_national_rugby_union_team
[winner: 20954423 The_International_(play), pem_pos=2,
final_scores=-0.9989661, final_local_scores=-0.91283244,
pem=-1.4805043, similarity=-15.099816|
RECALL MISS: 955656 France_national_rugby_union_team]

9: 6981472 Lyon_OU
[winner: 979521 Olympique_Lyonnais, pem_pos=1,
final_scores=0.28965026, final_local_scores=0.24929872,
pem=-1.9193003, similarity=21.966738|
gt: 6981472 Lyon_OU, pem_pos=10,
final_scores=-0.69199276, final_local_scores=-0.78201675,
pem=-4.4806156, similarity=5.119057]

10: 1722990 AS_Béziers_Hérault
[winner: 16480998 AS_Béziers_Hérault_(football),
pem_pos=1, final_scores=-0.28525335,
final_local_scores=-0.3045538, pem=-2.3735886,
similarity=7.909163|
gt: 1722990 AS_Béziers_Hérault, pem_pos=2,
final_scores=-0.33293754, final_local_scores=-0.36971477,
pem=-2.7115715, similarity=7.802634]

11: 5494088 Andrézieux-Bouthéon
[winner: 5494088 Andrézieux-Bouthéon, pem_pos=0,
final_scores=-0.080347106, final_local_scores=-0.042036414,
pem=0.0, similarity=2.8352625|
gt: 5494088 Andrézieux-Bouthéon, pem_pos=0,
final_scores=-0.080347106, final_local_scores=-0.042036414,
pem=0.0, similarity=2.8352625]

12: 5991818 Fédérale_1
[winner: 5991818 Fédérale_1, pem_pos=0,
final_scores=0.020983502, final_local_scores=-0.019919336,
pem=0.0, similarity=3.4952803|
gt: 5991818 Fédérale_1, pem_pos=0,
final_scores=0.020983502, final_local_scores=-0.019919336,
pem=0.0, similarity=3.4952803]

13: 8328448 Rugby_Nice_Côte_d'Azur_Université-Racing
[winner: 704478 OGC_Nice, pem_pos=1,
final_scores=0.7655211, final_local_scores=0.7744055,
pem=-2.1390686, similarity=38.832104|
gt: 8328448 Rugby_Nice_Côte_d'Azur_Université-Racing,
pem_pos=14, final_scores=-0.5390988,
final_local_scores=-0.5932541, pem=-3.7032585,
similarity=6.524724]

14: 27697874 RC_Chalon
[winner: 3462850 Élan_Chalon, pem_pos=2,
final_scores=-0.35064483, final_local_scores=-0.37072784,
pem=-3.4187198, similarity=11.617985|
gt: 27697874 RC_Chalon, pem_pos=4,
final_scores=-1.0086797, final_local_scores=-1.0844269,
pem=-7.6083746, similarity=13.103776]

15: 10196332 US_Carcassonne
[winner: 10196332 US_Carcassonne, pem_pos=3,
final_scores=-0.37176356, final_local_scores=-0.401284,

pem=-2.321024, similarity=4.736686|
gt: 10196332 US_Carcassonne, pem_pos=3,
final_scores=-0.37176356, final_local_scores=-0.401284,
pem=-2.321024, similarity=4.736686]
1: 5843419 France
[pem_pos=0, final_scores=0.70707124,
final_local_scores=0.75788325, pem=-0.80223274,
similarity=31.069126]
2: 214482 Saint-Étienne
[pem_pos=0, final_scores=0.0714591,
final_local_scores=0.05787912, pem=-0.6656834,
similarity=9.437031]
3: 31464821 Loire
[pem_pos=0, final_scores=0.07350446,
final_local_scores=0.11629549, pem=-0.06706437,
similarity=7.924908]
4: 5992163 Rugby_Pro_D2
[pem_pos=0, final_scores=0.19659488,
final_local_scores=0.16528481, pem=0.0, similarity=9.022141]
5: 10120820 Alexandre_Péclier
[pem_pos=0, final_scores=0.67513895,
final_local_scores=0.7517581, pem=0.0, similarity=26.523674]
6: 629239 Eurosport
[pem_pos=0, final_scores=0.034161434,
final_local_scores=0.06809147, pem=0.0, similarity=6.121698]
7: 8332924 Union_Sportive_Tyrosse_Rugby_Côte_Sud
[pem_pos=0, final_scores=0.25806582,
final_local_scores=0.32158428, pem=0.0, similarity=13.686429]
8: 17965716 Union_Sportive_Bressane
[pem_pos=1, final_scores=0.2768861,
final_local_scores=0.30967018, pem=-0.6608374,
similarity=16.924625]
9: 3387365 Stade_Geoffroy-Guichard
[pem_pos=0, final_scores=0.2582758,
final_local_scores=0.27529222, pem=0.0, similarity=12.304982]
10: 5992163 Rugby_Pro_D2
[pem_pos=0, final_scores=0.15325768,
final_local_scores=0.12131119, pem=0.0, similarity=7.7098804]
1: 5843419 France
[pem_pos=0, final_scores=0.62303025,
final_local_scores=0.66608113, pem=-0.80223274,
similarity=28.329569]
2: 1145121 AS_Saint-Étienne
[pem_pos=1, final_scores=0.51506156,
final_local_scores=0.4784026, pem=-0.7714895,
similarity=22.56168]
3: 1145121 AS_Saint-Étienne
[pem_pos=1, final_scores=0.5427414,
final_local_scores=0.5086387, pem=-0.88346994,
similarity=24.072954]
4: 1145121 AS_Saint-Étienne
[pem_pos=1, final_scores=0.59097236,

final_local_scores=0.5613237, pem=-0.88346994,
 similarity=25.645178]
 5: 979521 Olympique_Lyonnais
 [pem_pos=1, final_scores=0.28965026,
 final_local_scores=0.24929872, pem=-1.9193003,
 similarity=21.966738]
 6: 704478 OGC_Nice
 [pem_pos=1, final_scores=0.7655211,
 final_local_scores=0.7744055, pem=-2.1390686,
 similarity=38.832104]
 7: 10120820 Alexandre_Péclier
 [pem_pos=0, final_scores=0.25094157,
 final_local_scores=0.28838694, pem=0.0, similarity=12.695754]
 8: 1145121 AS_Saint-Étienne
 [pem_pos=1, final_scores=0.5315191,
 final_local_scores=0.49638003, pem=-0.88346994,
 similarity=23.70713]
 9: 263841 Bourg-en-Bresse
 [pem_pos=0, final_scores=0.1282854,
 final_local_scores=0.13479981, pem=0.0, similarity=8.112408]
 10: 1145121 AS_Saint-Étienne
 [pem_pos=1, final_scores=0.47581977,
 final_local_scores=0.43553692, pem=-0.88346994,
 similarity=21.89145]

Sample: b'Rapp_Road_Community_Historic_District'

The [1 Rapp Road Community Historic District 1] is located
 in the [2 1 Pine Bush 2] area of [1 1 Albany 1] , [2 2 New Yo
 rk 2] . It is a residential neighborhood . In 2002 it was
 listed on the [3 3 National Register of Historic Places 3] .
 It was established in the 1920s by an [4 4 African American 4]
 minister who had moved north from [3 Mississippi 3] ,
 originally to [4 Albany 4] 's . He was followed by other
 members of his . Neither he nor they liked urban life much ,
 and eventually he bought the land along [5 Rapp Road 5] where
 they all moved . Half of the original purchase was 2 taken by
 the state for road projects in the 1970s . The remaining
 half , today 's , has many of the original buildings . Most
 of the original families ' descendants still live there . It
 is a rare intact example of a 3 chain migration community
 from the [5 5 Great Migration 5] . ' [6 Rapp Road 6] is
 located in the long , narrow western protrusion of [7 Albany 7
] known as the [8 4 Pine Bush 8] . The portion of the street
 on which the district is located lies between Pine Lane and
 the South Frontage Road of [6 6 Washington Avenue Extension 6]
 . It is just north of the boundary between the city and the
 town of [7 7 Guilderland 7] . To the southeast is 5 Crossgates
 Mall . Wooded lands on the east and west serve as a buffer
 between it and the mall and various other commercial and
 office developments in those directions . It is further
 isolated from any other residential neighborhoods by
 Interstates 6 87 and 7 90 , both part of the [8 8 New York St

ate Thruway 8] , to the north and east . The land is generally level . The district is formed by the 27 8 lots that remain of the two original purchases that created the neighborhood , on both sides of the road . There are 21 buildings on those lots , all but two of which are 9 contributing properties . One property has two stone piers marking its driveway , both of which are considered contributing objects . The buildings are all 10 wood frame residences , a mix of cottages and . The Reverend Louis W. Parson and his wife migrated to [9 Albany 9] from [10 11 Shubuta 10] , [11 Mississippi 11] in 1927 . He founded the First Church of God in Christ in [12 Albany 12] . In four trips to [13 Mississippi 13] , Parson encouraged friends and family to move to [14 Albany 14] and join the church . Many friends and family did , during the 1930s and 1940s . Some were s who owed money to their landlords . Eventually the majority of Shubuta moved to [15 Albany 15] , part of the 400,000 [9 9 African American 9] s who left the 12 South during what has been called the [10 10 Great Migration 10] around [16 World War I 16] seeking better opportunities . Despite pressure from the city government to not bring in anymore unemployed people due to the [11 11 Depression 11] , Parson continued doing so throughout the 1930s . But many of the migrants did not like life in the . There were bars , brothels and gambling houses in the neighborhood , affronts to their religious values that they did not want their children to be near . Many had also come from rural areas , and found city life hard to adjust to . Some went back to [17 Mississippi 17] . Parson found two parcels of undeveloped land in the [18 13 Pine Bush 18] for the community to move to between 1930 and 1933 . They saved money and built , by themselves , houses similar to those they had lived in in [19 Mississippi 19] . After the purchases were fully 14 closed in 1942 , two years after Parson 's death , families began building their own houses with help from friends and family over the next decades . Many began growing crops and raising animals on the land , becoming . During the [20 World War II 20] years , especially , building materials were hard to obtain , and thus one family found they could complete a small shotgun house but not the larger house that they had started . At its peak , 23 families lived on [21 Rapp Road 21] . The neighborhood remained intact until 1971 , when the state planned the [12 12 Washington Avenue Extension 12] to better connect the city and its growing western suburbs . This required 15 eminent domain proceedings against the northern parcel . Most of the homeowners moved away . One who lived at what is today the middle of the extension chose instead to have her house moved , and today it is at the north end of the street , 8 [22 Rapp Road 22] , 300 feet -LRB- 90 m -RRB- south of its original location . The community has remained intact despite the large-scale development of the surrounding area . Descendants of the original homeowners have returned

to raise their families . Every year the community holds a ''

16 family reunion

1: 52106 Albany,_New_York

2: 8210131 New_York

3: 64065 National_Register_of_Historic_Places

4: 2154 African_American

5: 375540 Great_Migration_(African_American)

6: 8498414 New_York_State_Route_910D

7: 126084 Guilderland,_New_York

8: 632092 New_York_State_Thruway

9: 2154 African_American

10: 375540 Great_Migration_(African_American)

11: 19283335 Great_Depression

12: 8498414 New_York_State_Route_910D

1: 2031138 Albany_Pine_Bush

[winner: 126951 Pine_Bush,_New_York, pem_pos=0,
final_scores=0.3198712, final_local_scores=0.35729548,
pem=0.0, similarity=14.752123]

gt: 2031138 Albany_Pine_Bush, pem_pos=1,
final_scores=0.23537187, final_local_scores=0.26269132,
pem=-0.8457769, similarity=16.528412]

2: 144089 Eminent_domain

[winner: 19729978 Talent_(comics), pem_pos=5,
final_scores=-0.9465464, final_local_scores=-0.8567564,
pem=-1.9459101, similarity=-10.895444]

RECALL MISS: 144089 Eminent_domain]

3: 3176113 Chain_migration

[winner: 3176113 Chain_migration, pem_pos=0,
final_scores=-0.9459979, final_local_scores=-0.9726962,
pem=0.0, similarity=-24.93748]

gt: 3176113 Chain_migration, pem_pos=0,
final_scores=-0.9459979, final_local_scores=-0.9726962,
pem=0.0, similarity=-24.93748]

4: 2031138 Albany_Pine_Bush

[winner: 126951 Pine_Bush,_New_York, pem_pos=0,
final_scores=0.19536977, final_local_scores=0.22129661,
pem=0.0, similarity=10.6936455]

gt: 2031138 Albany_Pine_Bush, pem_pos=1,
final_scores=0.076942965, final_local_scores=0.08963178,
pem=-0.8457769, similarity=11.363971]

5: 4001780 Crossgates_Mall

[winner: 4001780 Crossgates_Mall, pem_pos=0,
final_scores=-0.042306498, final_local_scores=0.018972188,
pem=0.0, similarity=4.655881]

gt: 4001780 Crossgates_Mall, pem_pos=0,
final_scores=-0.042306498, final_local_scores=0.018972188,
pem=0.0, similarity=4.655881]

6: 93830 Interstate_87

[winner: 4598927 Malaysia_Federal_Route_87, pem_pos=4,
final_scores=-0.99376565, final_local_scores=-0.89892155,
pem=-3.0130494, similarity=-6.3504677]

gt: 93830 Interstate_87, pem_pos=12,

final_scores=-1.4242992, final_local_scores=-1.5519218,
pem=-3.4933624, similarity=-23.225285]
7: 10137133 Interstate_90_in_New_York
[winner: 89866 Interstate_90, pem_pos=1,
final_scores=-0.4862224, final_local_scores=-0.5502707,
pem=-2.4818695, similarity=1.16533|
RECALL MISS: 10137133 Interstate_90_in_New_York]
8: 6841967 Land_lot
[winner: 6841967 Land_lot, pem_pos=0,
final_scores=-1.3676269, final_local_scores=-1.5091238,
pem=0.0, similarity=-40.94555|
gt: 6841967 Land_lot, pem_pos=0,
final_scores=-1.3676269, final_local_scores=-1.5091238,
pem=0.0, similarity=-40.94555]
9: 9596931 Contributing_property
[winner: 25309702
Contributing_Properties_in_the_Historic_Union_Ave_Commercial_
District, pem_pos=1, final_scores=-0.7504587,
final_local_scores=-0.6829285, pem=-0.69071114,
similarity=-12.53404|
gt: 9596931 Contributing_property, pem_pos=0,
final_scores=-1.0678152, final_local_scores=-1.1659648,
pem=0.0, similarity=-30.705002]
10: 699422 Timber_framing
[winner: 6587903 Tonewood, pem_pos=4,
final_scores=-0.9143526, final_local_scores=-0.8469364,
pem=-1.6092613, similarity=-12.4331455|
RECALL MISS: 699422 Timber_framing]
[winner: 41172 Frame_(networking), pem_pos=1,
final_scores=-1.1207614, final_local_scores=-1.0462828,
pem=-1.5288478, similarity=-18.819338|
RECALL MISS: 699422 Timber_framing]
11: 122250 Shubuta,_Mississippi
[winner: 16949861 Mississippi, pem_pos=0,
final_scores=0.3254583, final_local_scores=0.28841063,
pem=-0.31944752, similarity=14.433667|
RECALL MISS: 122250 Shubuta,_Mississippi]
[winner: 122250 Shubuta,_Mississippi, pem_pos=0,
final_scores=0.22672118, final_local_scores=0.25276715,
pem=0.0, similarity=11.632789|
gt: 122250 Shubuta,_Mississippi, pem_pos=0,
final_scores=0.22672118, final_local_scores=0.25276715,
pem=0.0, similarity=11.632789]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8074273, final_local_scores=-1.7889225,
pem=-4.6728287, similarity=-23.883732|
RECALL MISS: 122250 Shubuta,_Mississippi]
12: 179553 Southern_United_States
[winner: 179553 Southern_United_States, pem_pos=0,
final_scores=-0.015420243, final_local_scores=-0.06879377,
pem=-1.3313899, similarity=9.277075|
gt: 179553 Southern_United_States, pem_pos=0,

final_scores=-0.015420243, final_local_scores=-0.06879377,
 pem=-1.3313899, similarity=9.277075]
 13: 2031138 Albany_Pine_Bush
 [winner: 126951 Pine_Bush, New_York, pem_pos=0,
 final_scores=0.17291017, final_local_scores=0.19676289,
 pem=0.0, similarity=9.96151]
 gt: 2031138 Albany_Pine_Bush, pem_pos=1,
 final_scores=0.055914387, final_local_scores=0.06666128,
 pem=-0.8457769, similarity=10.678485]
 14: 1543554 Closing_(real_estate)
 [winner: 47279 Closed_set, pem_pos=2,
 final_scores=-1.2602991, final_local_scores=-1.208698,
 pem=-1.5285214, similarity=-23.667912]
 RECALL MISS: 1543554 Closing_(real_estate)]
 15: 144089 Eminent_domain
 [winner: 8667766 Oregon_Ballot_Measure_39_(2006),
 pem_pos=0, final_scores=-0.7248811,
 final_local_scores=-0.60531616, pem=0.0,
 similarity=-13.97413]
 gt: 144089 Eminent_domain, pem_pos=1,
 final_scores=-1.0243558, final_local_scores=-1.1207972,
 pem=-0.0087238895, similarity=-29.30967]
 16: 1004097 Family_reunion
 [winner: 56276 Family_(biology), pem_pos=0,
 final_scores=-1.0563444, final_local_scores=-1.01737,
 pem=-0.29807892, similarity=-24.64964]
 RECALL MISS: 1004097 Family_reunion]
 [winner: 17503871 Reunion_(2001_film), pem_pos=14,
 final_scores=-1.1033733, final_local_scores=-0.9977338,
 pem=-3.1024413, similarity=-8.813094]
 RECALL MISS: 1004097 Family_reunion]
 1: 52106 Albany, New_York
 [pem_pos=0, final_scores=0.15648125,
 final_local_scores=0.0989625, pem=-1.175974,
 similarity=13.438081]
 2: 8210131 New_York
 [pem_pos=0, final_scores=0.2107542,
 final_local_scores=0.17236981, pem=-0.2889045,
 similarity=10.804679]
 3: 64065 National_Register_of_Historic_Places
 [pem_pos=0, final_scores=0.3228665,
 final_local_scores=0.32930744, pem=0.0, similarity=13.9169035]
 4: 2154 African_American
 [pem_pos=1, final_scores=0.2868845,
 final_local_scores=0.2730841, pem=-0.7252736,
 similarity=16.183239]
 5: 375540 Great_Migration_(African_American)
 [pem_pos=0, final_scores=0.030443981,
 final_local_scores=0.033092618, pem=-0.014469738,
 similarity=5.1559515]
 6: 8498414 New_York_State_Route_910D
 [pem_pos=0, final_scores=0.04624705,

final_local_scores=0.10467672, pem=0.0, similarity=7.213475]
7: 126084 Guilderland, New_York
[pem_pos=0, final_scores=0.30891508,
final_local_scores=0.30935067, pem=0.0, similarity=13.321354]
8: 632092 New_York_State_Thruway
[pem_pos=0, final_scores=0.091444954,
final_local_scores=0.093109205, pem=0.0, similarity=6.8682775]
9: 2154 African_American
[pem_pos=1, final_scores=0.33135813,
final_local_scores=0.32166475, pem=-0.7252736,
similarity=17.632982]
10: 375540 Great_Migration_(African_American)
[pem_pos=0, final_scores=0.18953408,
final_local_scores=0.20687437, pem=-0.014469738,
similarity=10.341946]
11: 19283335 Great_Depression
[pem_pos=0, final_scores=0.08331515,
final_local_scores=0.074563086, pem=-0.43945006,
similarity=8.704622]
12: 8498414 New_York_State_Route_910D
[pem_pos=0, final_scores=0.045550868,
final_local_scores=0.10391627, pem=0.0, similarity=7.1907816]
1: 22487448 Rapp_Road_Community_Historic_District
[pem_pos=0, final_scores=0.10566114,
final_local_scores=0.098842606, pem=0.0, similarity=7.0393734]
2: 126951 Pine_Bush, New_York
[pem_pos=0, final_scores=0.3198712,
final_local_scores=0.35729548, pem=0.0, similarity=14.752123]
3: 16949861 Mississippi
[pem_pos=0, final_scores=0.23339973,
final_local_scores=0.18937889, pem=-0.31944752,
similarity=11.478362]
4: 52106 Albany, New_York
[pem_pos=0, final_scores=0.06985007,
final_local_scores=0.004331127, pem=-1.175974,
similarity=10.614092]
5: 22487448 Rapp_Road_Community_Historic_District
[pem_pos=0, final_scores=0.15730964,
final_local_scores=0.15526077, pem=0.0, similarity=8.723004]
6: 22487448 Rapp_Road_Community_Historic_District
[pem_pos=0, final_scores=0.076213256,
final_local_scores=0.066675276, pem=0.0, similarity=6.0794363]
7: 52106 Albany, New_York
[pem_pos=0, final_scores=0.08062114,
final_local_scores=0.016096875, pem=-1.175974,
similarity=10.965205]
8: 126951 Pine_Bush, New_York
[pem_pos=0, final_scores=0.19536977,
final_local_scores=0.22129661, pem=0.0, similarity=10.6936455]
9: 52106 Albany, New_York
[pem_pos=0, final_scores=0.19491325,
final_local_scores=0.14094365, pem=-1.175974,

```

similarity=14.690883]
10: 122250 Shubuta, Mississippi
    [pem_pos=0, final_scores=0.22672118,
final_local_scores=0.25276715, pem=0.0, similarity=11.632789]
11: 16949861 Mississippi
    [pem_pos=0, final_scores=0.3254583,
final_local_scores=0.28841063, pem=-0.31944752,
similarity=14.433667]
12: 52106 Albany, New York
    [pem_pos=0, final_scores=0.1752897,
final_local_scores=0.11950791, pem=-1.175974,
similarity=14.051196]
13: 16949861 Mississippi
    [pem_pos=0, final_scores=0.26642776,
final_local_scores=0.22392872, pem=-0.31944752,
similarity=12.5093975]
14: 52106 Albany, New York
    [pem_pos=0, final_scores=0.13404255,
final_local_scores=0.07445164, pem=-1.175974,
similarity=12.706628]
15: 52106 Albany, New York
    [pem_pos=0, final_scores=0.14753948,
final_local_scores=0.089195, pem=-1.175974,
similarity=13.146599]
16: 4764461 World War I
    [pem_pos=0, final_scores=0.37355655,
final_local_scores=0.4091413, pem=0.0, similarity=16.299305]
17: 16949861 Mississippi
    [pem_pos=0, final_scores=0.2084174,
final_local_scores=0.16056126, pem=-0.31944752,
similarity=10.618386]
18: 126951 Pine Bush, New York
    [pem_pos=0, final_scores=0.17291017,
final_local_scores=0.19676289, pem=0.0, similarity=9.96151]
19: 16949861 Mississippi
    [pem_pos=0, final_scores=0.24335812,
final_local_scores=0.19872868, pem=-0.31944752,
similarity=11.757379]
20: 32927 World War II
    [pem_pos=0, final_scores=0.43916827,
final_local_scores=0.47562134, pem=0.0, similarity=18.283203]
21: 22487448 Rapp Road Community Historic District
    [pem_pos=0, final_scores=0.1791756,
final_local_scores=0.17914602, pem=0.0, similarity=9.435787]
22: 22487448 Rapp Road Community Historic District
    [pem_pos=0, final_scores=0.141405,
final_local_scores=0.13788736, pem=0.0, similarity=8.204547]

```

Sample: b'4769_Castalia'

The asteroid 4769 Castalia -LRB- ; previously known by the 1989 PB -RRB- was the first 1 asteroid to be modeled by 2 radar imaging . It is an , and . It was discovered on August 9 ,

1989 , by 3 Eleanor F. Helin -LRB- -RRB- on photographic plates taken at 4 Palomar Observatory . It is named after 5 Castalia , a 6 nymph in [1 7 Greek 1] mythology . On 25 August 1989 Castalia passed -LRB- within eleven 8 lunar distance s -RRB- of [2 Earth 2] , allowing it to be observed with 9 radar from the [1 1 Arecibo Observatory 1] by 10 Scott Hudson -LRB- [2 2 Washington State University 2] -RRB- and [3 3 Steven J. Ostro 3] -LRB- 11 JPL -RRB- . The data allowed Hudson et al. to produce a three-dimensional 12 model of the object . During the 1989 passage Castalia peaked at an 13 apparent magnitude of 12 . Castalia has a peanut shape , suggesting two approximately 800-meter-diameter pieces held together by their weak mutual gravity . Since then radar observations of other asteroids have found other 14 contact binaries

1: 159191 Arecibo_Observatory

2: 228600 Washington_State_University

3: 522065 Steven_J._Ostro

1: 791 Asteroid

[winner: 47264 Asteroid_belt, pem_pos=2, final_scores=-0.8585998, final_local_scores=-0.85604936, pem=-1.788938, similarity=-11.727984]

gt: 791 Asteroid, pem_pos=0, final_scores=-0.928358, final_local_scores=-1.0026505, pem=-0.018797128, similarity=-25.729155]

2: 1086531 Radar_imaging

[winner: 1086531 Radar_imaging, pem_pos=0, final_scores=-0.78102124, final_local_scores=-0.8009402, pem=-0.018692123, similarity=-19.71029]

gt: 1086531 Radar_imaging, pem_pos=0, final_scores=-0.78102124, final_local_scores=-0.8009402, pem=-0.018692123, similarity=-19.71029]

3: 590308 Eleanor_F._Helin

[winner: 590308 Eleanor_F._Helin, pem_pos=0, final_scores=0.024777249, final_local_scores=0.044089124, pem=0.0, similarity=5.4054203]

gt: 590308 Eleanor_F._Helin, pem_pos=0, final_scores=0.024777249, final_local_scores=0.044089124, pem=0.0, similarity=5.4054203]

4: 392143 Palomar_Observatory

[winner: 392143 Palomar_Observatory, pem_pos=0, final_scores=0.02495636, final_local_scores=0.058598578, pem=0.0, similarity=5.8384113]

gt: 392143 Palomar_Observatory, pem_pos=0, final_scores=0.02495636, final_local_scores=0.058598578, pem=0.0, similarity=5.8384113]

5: 46978 Castalia

[winner: 129167 Castalia,_Ohio, pem_pos=1, final_scores=-0.23801382, final_local_scores=-0.14106198, pem=-1.2463183, similarity=6.657815]

gt: 46978 Castalia, pem_pos=0, final_scores=-0.25044423, final_local_scores=-0.12061102, pem=-0.4563048, similarity=2.971896]

6: 22058 Nymph
[winner: 99748 Apsara, pem_pos=0,
final_scores=-0.9764665, final_local_scores=-0.9099981,
pem=0.0, similarity=-23.066446|
gt: 22058 Nymph, pem_pos=1, final_scores=-0.98024935,
final_local_scores=-0.99909693, pem=-0.27546817,
similarity=-24.227293]
7: 23416994 Greek_mythology
[winner: 12108 Greece, pem_pos=1,
final_scores=0.46512085, final_local_scores=0.59136295,
pem=-2.3247876, similarity=34.37972|
gt: 23416994 Greek_mythology, pem_pos=5,
final_scores=-0.76878613, final_local_scores=-0.76077664,
pem=-3.2765498, similarity=-0.7949919]
[winner: 5873281 Māui_(Māori_mythology), pem_pos=4,
final_scores=-1.1772398, final_local_scores=-1.1099916,
pem=-1.9459101, similarity=-18.452486|
gt: 23416994 Greek_mythology, pem_pos=1,
final_scores=-1.3895817, final_local_scores=-1.4633148,
pem=-1.7787589, similarity=-29.905354]
8: 899115 Lunar_distance_(astronomy)
[winner: 19331 Moon, pem_pos=0, final_scores=-0.99991035,
final_local_scores=-1.1334114, pem=-0.32505962,
similarity=-27.965816|
RECALL MISS: 899115 Lunar_distance_(astronomy)]
[winner: 18182884 Distance_(EP), pem_pos=7,
final_scores=-1.1738678, final_local_scores=-1.0391519,
pem=-2.1893663, similarity=-15.014537|
RECALL MISS: 899115 Lunar_distance_(astronomy)]
9: 264337 Radar_astronomy
[winner: 9354788 SHARAD, pem_pos=2,
final_scores=-1.1516814, final_local_scores=-1.1924105,
pem=-1.6091789, similarity=-22.743225|
gt: 264337 Radar_astronomy, pem_pos=9,
final_scores=-1.9469724, final_local_scores=-2.1360042,
pem=-5.7572975, similarity=-28.343834]
10: 522076 Scott_Hudson
[winner: 522076 Scott_Hudson, pem_pos=0,
final_scores=-0.09528558, final_local_scores=-0.032386266,
pem=0.0, similarity=3.1232421|
gt: 522076 Scott_Hudson, pem_pos=0,
final_scores=-0.09528558, final_local_scores=-0.032386266,
pem=0.0, similarity=3.1232421]
11: 16459 Jet_Propulsion_Laboratory
[winner: 16459 Jet_Propulsion_Laboratory, pem_pos=0,
final_scores=-0.010554433, final_local_scores=-0.062208563,
pem=0.0, similarity=2.2332854|
gt: 16459 Jet_Propulsion_Laboratory, pem_pos=0,
final_scores=-0.010554433, final_local_scores=-0.062208563,
pem=0.0, similarity=2.2332854]
12: 20590 Mathematical_model
[winner: 19916559 Atomic_nucleus, pem_pos=0,

```

final_scores=-1.0505266, final_local_scores=-1.0979207,
pem=0.0, similarity=-28.67443|
    RECALL MISS: 20590 Mathematical_model]
13: 1962 Apparent_magnitude
    [winner: 1962 Apparent_magnitude, pem_pos=0,
final_scores=-1.0469126, final_local_scores=-1.0997283,
pem=-0.0031082607, similarity=-28.71147|
    gt: 1962 Apparent_magnitude, pem_pos=0,
final_scores=-1.0469126, final_local_scores=-1.0997283,
pem=-0.0031082607, similarity=-28.71147 ]
14: 7066200 Contact_binary_(asteroid)
    [winner: 185175 Contact_binary, pem_pos=1,
final_scores=-0.5811985, final_local_scores=-0.47181672,
pem=-0.52324814, similarity=-7.144734|
    gt: 7066200 Contact_binary_(asteroid), pem_pos=0,
final_scores=-0.7552385, final_local_scores=-0.7324794,
pem=0.0, similarity=-17.768932 ]
1: 159191 Arecibo_Observatory
    [pem_pos=0, final_scores=0.1373009,
final_local_scores=0.1341401, pem=0.0, similarity=8.092721]
2: 228600 Washington_State_University
    [pem_pos=0, final_scores=0.09298308,
final_local_scores=0.15644705, pem=0.0, similarity=8.758405]
3: 522065 Steven_J._Ostro
    [pem_pos=0, final_scores=0.08908318,
final_local_scores=0.12808943, pem=0.0, similarity=7.912157]
1: 12108 Greece
    [pem_pos=1, final_scores=0.46512085,
final_local_scores=0.59136295, pem=-2.3247876,
similarity=34.37972]
2: 9228 Earth
    [pem_pos=0, final_scores=0.039902195,
final_local_scores=0.050044313, pem=-0.07783105,
similarity=6.0063925]

```

Sample: b'William_Wood'

, _ 1st_Baron_Hatherley William Page Wood , 1st Baron Hatherley , 1 QC -LRB- 29 November 1801 -- 10 July 1881 -RRB- was a [1 1 British 1] lawyer and statesman who served as a 2 L iberal 3 Lord Chancellor between 1868 and 1872 in [2 2 William Ewart Gladstone 2] 's first ministry . [1 Wood 1] was born in [3 3 London 3] , the second son of 4 Sir [2 Matthew Wood 2] , 1st Baronet , a [3 London 3] 5 alderman and 6 Lord Mayor who became famous for befriending 7 Queen [4 Caroline 4] and braving [4 4 George IV 4] . 8 Sir [5 Evelyn Wood 5] and 9 Kat harine O'Shea were his nephew and niece respectively . He was educated at [6 10 Winchester 6] , from which he was expelled after a revolt against the headmaster , [5 5 Geneva University 5] , and [7 11 Trinity College 7] , [8 Cambridge 8] , where he became a fellow after being 24th 12 wrangler in 1824 . [9 Wood 9] entered 13 Lincoln 's Inn , and was in 1824 , studying conveyancing in John Tyrrell 's chambers . He soon

obtained a good practice as an equity draughtsman and before 1
4 parliamentary committees . In 1845 he became a 15 Queen 's C
ounsel , and in 1847 was elected to parliament for the city
of [10 16 Oxford 10] as a Liberal . In 1849 he was appointed
Vice-Chancellor of the [11 County Palatine of Lancaster 11] ,
and in 1851 was made 17 Solicitor General for [12 England 12]
and Wales and knighted , vacating the former position in
1852 . When his party returned to power in 1853 , he was
raised to the bench as a 18 Vice-Chancellor . In 1868 he was
made a 19 Lord Justice of Appeal , but before the end of the
year was selected by [13 Gladstone 13] to be [6 6 Lord Chance
llor 6] and was raised to the peerage as Baron Hatherley , of
[14 Down Hatherley 14] in the [15 County of Gloucester 15] .
He retired in 1872 owing to failing eyesight , but sat
occasionally as a law lord . Wood married Charlotte ,
daughter of 20 Edward Moor , in 1830 . They had no children .
Charlotte 's death in 1878 was a great blow to [16 Wood 16] ,
from which he never recovered , and he died in [17 London 17]
on 10 July 1881 , aged 79 . Both are buried in the churchyard
in 21 Great Bealings

1: 31717 United_Kingdom

2: 33954 William_Ewart_Gladstone

3: 17867 London

4: 40225 George_IV_of_the_United_Kingdom

5: 342348 University_of_Geneva

6: 226197 Lord_Chancellor

1: 90072 Queen's_Counsel

[winner: 90072 Queen's_Counsel, pem_pos=0,
final_scores=-0.36618167, final_local_scores=-0.41903713,
pem=-0.39452267, similarity=-6.2697153]

gt: 90072 Queen's_Counsel, pem_pos=0,
final_scores=-0.36618167, final_local_scores=-0.41903713,
pem=-0.39452267, similarity=-6.2697153]

2: 4482 Liberal_Party_(UK)

[winner: 149536 Liberal_Party_of_Canada, pem_pos=2,
final_scores=-0.13159898, final_local_scores=-0.115353405,
pem=-1.2222316, similarity=7.2940226]

gt: 4482 Liberal_Party_(UK), pem_pos=3,
final_scores=-0.19940719, final_local_scores=-0.25098452,
pem=-1.7910957, similarity=6.3400908]

3: 226197 Lord_Chancellor

[winner: 226197 Lord_Chancellor, pem_pos=0,
final_scores=-0.07060526, final_local_scores=-0.12953323,
pem=0.0, similarity=0.22418289]

gt: 226197 Lord_Chancellor, pem_pos=0,
final_scores=-0.07060526, final_local_scores=-0.12953323,
pem=0.0, similarity=0.22418289]

4: 29926481 Sir_Matthew_Wood,_1st_Baronet

[winner: 170968 Matthew_Wood_(sound_editor), pem_pos=0,
final_scores=0.11734323, final_local_scores=0.2604498,
pem=-0.026418537, similarity=12.005723]

gt: 29926481 Sir_Matthew_Wood,_1st_Baronet, pem_pos=5,

final_scores=-0.110942796, final_local_scores=-0.09378028,
 pem=-2.8805194, similarity=16.955835]
 [winner: 79912 Baronet, pem_pos=0,
 final_scores=0.017275706, final_local_scores=-0.03132792,
 pem=-0.3796905, similarity=5.219641|
 RECALL MISS: 29926481 Sir_Matthew_Wood,_1st_Baronet]
 [winner: 19786922 Sir,_Tulkarm, pem_pos=16,
 final_scores=-1.1117029, final_local_scores=-1.0073113,
 pem=-3.1354942, similarity=-8.919165|
 RECALL MISS: 29926481 Sir_Matthew_Wood,_1st_Baronet]
 [winner: 406194 William_Wood,_1st_Baron_Hatherley,
 pem_pos=0, final_scores=-1.1512963,
 final_local_scores=-1.2472546, pem=0.0,
 similarity=-33.130856|
 RECALL MISS: 29926481 Sir_Matthew_Wood,_1st_Baronet]
 5: 286925 Alderman
 [winner: 286925 Alderman, pem_pos=0,
 final_scores=-1.1866726, final_local_scores=-1.3184035,
 pem=-0.03405784, similarity=-35.068867|
 gt: 286925 Alderman, pem_pos=0, final_scores=-1.1866726,
 final_local_scores=-1.3184035, pem=-0.03405784,
 similarity=-35.068867]
 6: 40339 Lord_Mayor_of_London
 [winner: 40339 Lord_Mayor_of_London, pem_pos=1,
 final_scores=-0.15792425, final_local_scores=-0.22014982,
 pem=-1.0694871, similarity=3.3360405|
 gt: 40339 Lord_Mayor_of_London, pem_pos=1,
 final_scores=-0.15792425, final_local_scores=-0.22014982,
 pem=-1.0694871, similarity=3.3360405]
 7: 48913 Caroline_of_Brunswick
 [winner: 292511 Caroline,_Princess_of_Hanover,
 pem_pos=11, final_scores=0.102857485,
 final_local_scores=0.17026162, pem=-3.6799214,
 similarity=29.182644|
 gt: 48913 Caroline_of_Brunswick, pem_pos=18,
 final_scores=-0.27667975, final_local_scores=-0.3102702,
 pem=-4.1385627, similarity=17.336779]
 [winner: 42010 Queen_(band), pem_pos=0,
 final_scores=-1.2448883, final_local_scores=-1.2926164,
 pem=-0.43144405, similarity=-32.138283|
 RECALL MISS: 48913 Caroline_of_Brunswick]
 8: 1189274 Evelyn_Wood_(British_Army_officer)
 [winner: 1189274 Evelyn_Wood_(British_Army_officer),
 pem_pos=1, final_scores=0.3742236,
 final_local_scores=0.38293338, pem=-0.40969086,
 similarity=17.74517|
 gt: 1189274 Evelyn_Wood_(British_Army_officer),
 pem_pos=1, final_scores=0.3742236,
 final_local_scores=0.38293338, pem=-0.40969086,
 similarity=17.74517]
 [winner: 226158 Sir, pem_pos=0, final_scores=-0.93473554,
 final_local_scores=-1.01748, pem=-0.31541774,

similarity=-24.55863|
 RECALL MISS: 1189274 Evelyn_Wood_(British_Army_officer)]

9: 504827 Katharine_O'Shea
 [winner: 43337 Katharine_Hepburn, pem_pos=0,
 final_scores=-0.10684948, final_local_scores=-0.082561046,
 pem=-0.695599, similarity=5.408702|
 gt: 504827 Katharine_O'Shea, pem_pos=7,
 final_scores=-0.8022811, final_local_scores=-0.8585279,
 pem=-4.158595, similarity=1.0846152]
 [winner: 4360368 O'Shea, pem_pos=2,
 final_scores=-1.3333933, final_local_scores=-1.2716706,
 pem=-1.7460895, similarity=-24.363966|
 gt: 504827 Katharine_O'Shea, pem_pos=9,
 final_scores=-2.4956658, final_local_scores=-2.7082937,
 pem=-3.3885336, similarity=-58.3038]

10: 50460 Winchester_College
 [winner: 43412 Winchester, pem_pos=0,
 final_scores=0.54468304, final_local_scores=0.52242035,
 pem=-0.40002757, similarity=21.85519|
 gt: 50460 Winchester_College, pem_pos=2,
 final_scores=-0.0400794, final_local_scores=-0.11999726,
 pem=-3.2583508, similarity=18.228174]

11: 31803 Trinity_College,_Cambridge
 [winner: 36995 Cambridge, pem_pos=0,
 final_scores=0.26116675, final_local_scores=0.22210076,
 pem=-0.751733, similarity=14.805684|
 gt: 31803 Trinity_College,_Cambridge, pem_pos=15,
 final_scores=-0.45693445, final_local_scores=-0.5698358,
 pem=-4.3138967, similarity=10.544318]
 [winner: 142298 Trinity_College,_Dublin, pem_pos=0,
 final_scores=0.048006788, final_local_scores=-0.0059433877,
 pem=-1.2422744, similarity=10.668032|
 gt: 31803 Trinity_College,_Cambridge, pem_pos=1,
 final_scores=0.02996312, final_local_scores=-0.043481365,
 pem=-1.2609318, similarity=9.649285]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.6350775, final_local_scores=-1.6436448,
 pem=-4.6728287, similarity=-19.548353|
 RECALL MISS: 31803 Trinity_College,_Cambridge]

12: 639371 Wrangler_(University_of_Cambridge)
 [winner: 13483523 Wrangler_(TV_series), pem_pos=4,
 final_scores=-1.2504591, final_local_scores=-1.1916277,
 pem=-2.2935119, similarity=-18.99836|
 gt: 639371 Wrangler_(University_of_Cambridge),
 pem_pos=2, final_scores=-1.3726536,
 final_local_scores=-1.4827945, pem=-1.8356787,
 similarity=-30.177122]

13: 207880 Lincoln's_Inn
 [winner: 24578 Pub, pem_pos=1, final_scores=0.0065398663,
 final_local_scores=-0.01254683, pem=-0.54729533,
 similarity=6.6915674|
 RECALL MISS: 207880 Lincoln's_Inn]

[winner: 17880 Lincoln_England, pem_pos=0,
final_scores=-0.24551366, final_local_scores=-0.31892565,
pem=-1.887664, similarity=4.837745]
RECALL MISS: 207880 Lincoln's_Inn]

[winner: 1415027 David_Berman_(musician), pem_pos=5,
final_scores=-2.5099607, final_local_scores=-2.6331599,
pem=-3.0012724, similarity=-58.167645]
RECALL MISS: 207880 Lincoln's_Inn]

14: 378032 Parliament_of_Great_Britain
[winner: 14152577 Kenyan_parliamentary_election,_2007,
pem_pos=19, final_scores=-1.5438004,
final_local_scores=-1.5770082, pem=-3.7269008,
similarity=-22.703892]
RECALL MISS: 378032 Parliament_of_Great_Britain]

15: 90072 Queen's_Counsel
[winner: 605309 Counsel, pem_pos=0,
final_scores=-0.008054331, final_local_scores=0.0065637827,
pem=-0.354545, similarity=6.21366]
RECALL MISS: 90072 Queen's_Counsel]
[winner: 48913 Caroline_of_Brunswick, pem_pos=0,
final_scores=-0.14436622, final_local_scores=-0.16802396,
pem=0.0, similarity=-0.9244573]
RECALL MISS: 90072 Queen's_Counsel]
[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.3047931, final_local_scores=-2.3866513,
pem=-3.4067376, similarity=-48.606354]
RECALL MISS: 90072 Queen's_Counsel]

16: 2757984 Oxford_(UK_Parliament_constituency)
[winner: 22308 Oxford, pem_pos=0, final_scores=0.3602808,
final_local_scores=0.31388605, pem=-0.68225896,
similarity=17.16693]
RECALL MISS: 2757984 Oxford_(UK_Parliament_constituency)]

17: 591120 Solicitor_General_for_England_and_Wales
[winner: 9316 England, pem_pos=0,
final_scores=0.20834391, final_local_scores=0.16451901,
pem=-0.19053856, similarity=10.035466]
RECALL MISS: 591120
Solicitor_General_for_England_and_Wales]
[winner: 69894 Wales, pem_pos=0,
final_scores=-0.12437877, final_local_scores=-0.17385691,
pem=-0.37222588, similarity=0.9256983]
RECALL MISS: 591120
Solicitor_General_for_England_and_Wales]
[winner: 731240 Solicitor_General_of_India, pem_pos=2,
final_scores=-0.56668466, final_local_scores=-0.52649426,
pem=-1.7193832, similarity=-2.271652]
gt: 591120 Solicitor_General_for_England_and_Wales,
pem_pos=1, final_scores=-0.64376706,
final_local_scores=-0.72974104, pem=-1.4649674,
similarity=-9.720496]

18: 2707477 Chancellor_of_the_High_Court
[winner: 2681197 Chancellor_(education), pem_pos=0,

```
final_scores=-0.03979005, final_local_scores=-0.06974615,
pem=0.0, similarity=2.0083487|
    gt: 2707477 Chancellor_of_the_High_Court, pem_pos=5,
final_scores=-0.83450323, final_local_scores=-0.90389574,
pem=-3.963654, similarity=-1.3293719 ]
19: 3740873 Lord_Justice_of_Appeal
    [winner: 3740873 Lord_Justice_of_Appeal, pem_pos=0,
final_scores=-0.04285942, final_local_scores=-0.054292396,
pem=0.0, similarity=2.4695196|
    gt: 3740873 Lord_Justice_of_Appeal, pem_pos=0,
final_scores=-0.04285942, final_local_scores=-0.054292396,
pem=0.0, similarity=2.4695196 ]
20: 13360738 Edward_Moor
    [winner: 13360738 Edward_Moor, pem_pos=0,
final_scores=-0.008648738, final_local_scores=0.04693438,
pem=0.0, similarity=5.4903283|
    gt: 13360738 Edward_Moor, pem_pos=0,
final_scores=-0.008648738, final_local_scores=0.04693438,
pem=0.0, similarity=5.4903283 ]
21: 15148456 Great_Bealings
    [winner: 15148456 Great_Bealings, pem_pos=0,
final_scores=-0.034181237, final_local_scores=0.015150964,
pem=0.0, similarity=4.5418477|
    gt: 15148456 Great_Bealings, pem_pos=0,
final_scores=-0.034181237, final_local_scores=0.015150964,
pem=0.0, similarity=4.5418477 ]
1: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5459748,
final_local_scores=0.5661313, pem=-0.48936236,
similarity=23.64543]
2: 33954 William_Ewart_Gladstone
    [pem_pos=0, final_scores=0.2645492,
final_local_scores=0.24277613, pem=0.0, similarity=11.334637]
3: 17867 London
    [pem_pos=0, final_scores=0.5815718,
final_local_scores=0.5806693, pem=-0.09086737,
similarity=21.912193]
4: 40225 George_IV_of_the_United_Kingdom
    [pem_pos=0, final_scores=0.32374674,
final_local_scores=0.3023802, pem=0.0, similarity=13.113341]
5: 342348 University_of_Geneva
    [pem_pos=0, final_scores=0.30531347,
final_local_scores=0.37355655, pem=0.0, similarity=15.237385]
6: 226197 Lord_Chancellor
    [pem_pos=0, final_scores=0.082186624,
final_local_scores=0.0423522, pem=0.0, similarity=5.353587]
1: 1189274 Evelyn_Wood_(British_Army_officer)
    [pem_pos=4, final_scores=0.055324927,
final_local_scores=0.03529887, pem=0.0, similarity=5.1431017]
2: 170968 Matthew_Wood_(sound_editor)
    [pem_pos=0, final_scores=0.11734323,
final_local_scores=0.2604498, pem=-0.026418537,
```

similarity=12.005723]
3: 17867 London
[pem_pos=0, final_scores=0.43200523,
final_local_scores=0.41729045, pem=-0.09086737,
similarity=17.036642]
4: 292511 Caroline,_Princess_of_Hanover
[pem_pos=11, final_scores=0.102857485,
final_local_scores=0.17026162, pem=-3.6799214,
similarity=29.182644]
5: 1189274 Evelyn_Wood_(British_Army_officer)
[pem_pos=1, final_scores=0.3742236,
final_local_scores=0.38293338, pem=-0.40969086,
similarity=17.74517]
6: 43412 Winchester
[pem_pos=0, final_scores=0.54468304,
final_local_scores=0.52242035, pem=-0.40002757,
similarity=21.85519]
7: 142298 Trinity_College,_Dublin
[pem_pos=0, final_scores=0.048006788,
final_local_scores=-0.0059433877, pem=-1.2422744,
similarity=10.668032]
8: 36995 Cambridge
[pem_pos=0, final_scores=0.26116675,
final_local_scores=0.22210076, pem=-0.751733,
similarity=14.805684]
9: 1189274 Evelyn_Wood_(British_Army_officer)
[pem_pos=4, final_scores=0.06171976,
final_local_scores=0.042232394, pem=0.0, similarity=5.350012]
10: 22308 Oxford
[pem_pos=0, final_scores=0.3602808,
final_local_scores=0.31388605, pem=-0.68225896,
similarity=17.16693]
11: 50435 Lancashire
[pem_pos=0, final_scores=0.30917794,
final_local_scores=0.28135732, pem=0.0, similarity=12.485976]
12: 9316 England
[pem_pos=0, final_scores=0.20834391,
final_local_scores=0.16451901, pem=-0.19053856,
similarity=10.035466]
13: 33954 William_Ewart_Gladstone
[pem_pos=0, final_scores=0.29159516,
final_local_scores=0.27782133, pem=0.0, similarity=12.380456]
14: 19779813 Down_Hatherley
[pem_pos=0, final_scores=0.07850893,
final_local_scores=0.16625074, pem=0.0, similarity=9.050966]
15: 47751 Gloucestershire
[pem_pos=0, final_scores=0.3635044,
final_local_scores=0.34443673, pem=0.0, similarity=14.368392]
16: 1189274 Evelyn_Wood_(British_Army_officer)
[pem_pos=4, final_scores=0.27621126,
final_local_scores=0.28225535, pem=0.0, similarity=12.512775]
17: 17867 London

[pem_pos=0, final_scores=0.5651747,
final_local_scores=0.5627579, pem=-0.09086737,
similarity=21.377682]

Sample: b"Jeanne_d'Evreux"

Jeanne d'Évreux -LRB- 1310 -- 4 March 1371 -RRB- was the third wife of King 1 Charles IV of [1 France 1] , daughter of his uncle and [1 1 Margaret of Artois 1] . She bore no male heir , thus `` causing '' the end of the 2 direct line of the [2 3 Capetian 2] dynasty . Because she was his first cousin , the couple required papal permission to marry from 4 Pope [3 John XXII 3] . [4 Jeanne 4] bore [5 Charles 5] three daughters , [6 Jeanne 6] , Marie and 5 Blanche . Two of [7 Jeanne 7] 's remarkable possessions survive -- her and a statue of the [8 6 Virgin 8] and Child . The Book of Hours , known as the 7 Hours of Jeanne d'Evreux , is in [2 2 The Cloisters 2] collection of the [3 3 Metropolitan Museum of Art 3] in [9 New York 9] . It was commissioned from the artist [4 4 Jean Pucelle 4] between 1324 and 1328 , probably as a gift from her husband [10 Charles IV 10] . The book contains the usual prayers of the 8 Canonical hours as arranged for the laity along with the notable inclusion of the office dedicated to [11 9 St Louis 11] , her great-grandfather . The small statue of the Virgin and Child -LRB- and 10 enamel , 69 cm high -RRB- , which [12 Jeanne 12] left to the monastery of outside [13 Paris 13] , is in the . Jeanne d'Évreux died on 4 March 1371 in her 11 château at [5 5 Brie-Comte-Robert 5] , in the region , some twenty miles south-east of [14 Paris 14] . She was buried at the [6 6 Basilica of St Denis 6]

1: 21806207 Margaret_of_Artois

2: 299976 The_Cloisters

3: 37535 Metropolitan_Museum_of_Art

4: 2443414 Jean_Pucelle

5: 5477238 Brie-Comte-Robert

6: 66437 Basilica_of_St_Denis

1: 78500 Charles_IV_of_France

[winner: 5843419 France, pem_pos=0,
final_scores=0.2104979, final_local_scores=0.21624863,
pem=-0.16240834, similarity=11.426206]

RECALL MISS: 78500 Charles_IV_of_France]

[winner: 38895 Charles_IV, Holy_Roman_Emperor, pem_pos=0,
final_scores=-0.34603566, final_local_scores=-0.444628,
pem=0.0, similarity=-9.178873]

gt: 78500 Charles_IV_of_France, pem_pos=3,
final_scores=-0.6121505, final_local_scores=-0.7323507,
pem=-1.6658691, similarity=-8.705837]

2: 1680861 House_of_Capet

[winner: 719514 Direct_Line, pem_pos=0,
final_scores=-1.1233704, final_local_scores=-1.1014946,
pem=0.0, similarity=-28.781084]

RECALL MISS: 1680861 House_of_Capet]

3: 6279 Capetian_dynasty

[winner: 1680861 House_of_Capet, pem_pos=0,
final_scores=0.085447654, final_local_scores=0.043562308,
pem=0.0, similarity=5.389699]

gt: 6279 Capetian_dynasty, pem_pos=1,
final_scores=0.02206926, final_local_scores=-0.014665559,
pem=-0.27329335, similarity=5.138275]

[winner: 27158897 Dynasty_(Kaskade_album), pem_pos=7,
final_scores=-1.3075103, final_local_scores=-1.207423,
pem=-3.3888543, similarity=-13.513078]

RECALL MISS: 6279 Capetian_dynasty]

4: 37670 Pope_John_XXII

[winner: 37670 Pope_John_XXII, pem_pos=0,
final_scores=0.17676897, final_local_scores=0.15170541,
pem=0.0, similarity=8.616905]

gt: 37670 Pope_John_XXII, pem_pos=0,
final_scores=0.17676897, final_local_scores=0.15170541,
pem=0.0, similarity=8.616905]

[winner: 23056 Pope, pem_pos=0, final_scores=-0.6562234,
final_local_scores=-0.774882, pem=-0.30024973,
similarity=-17.401508]

RECALL MISS: 37670 Pope_John_XXII]

5: 19942421 Blanche_of_France,_Duchess_of_Orléans

[winner: 1357858 Blanche_Devereaux, pem_pos=0,
final_scores=-0.14889114, final_local_scores=-0.08297217,
pem=-1.4923283, similarity=9.729171]

RECALL MISS: 19942421

Blanche_of_France,_Duchess_of_Orléans]

6: 21402951 Virgin_of_Jeanne_d'Evreux

[winner: 73513 Mary_(mother_of_Jesus), pem_pos=0,
final_scores=0.14073072, final_local_scores=0.148157,
pem=0.0, similarity=8.511013]

RECALL MISS: 21402951 Virgin_of_Jeanne_d'Evreux]

[winner: 128987 Child, pem_pos=0,
final_scores=-0.2384535, final_local_scores=-0.19682832,
pem=-0.6613066, similarity=1.8122514]

RECALL MISS: 21402951 Virgin_of_Jeanne_d'Evreux]

7: 11310306 Hours_of_Jeanne_d'Evreux

[winner: 585820 Jeanne_d'Évreux, pem_pos=0,
final_scores=0.023335889, final_local_scores=-0.0008728355,
pem=0.0, similarity=4.063666]

RECALL MISS: 11310306 Hours_of_Jeanne_d'Evreux]

[winner: 1951124 Liturgy_of_the_Hours, pem_pos=2,
final_scores=-0.37111598, final_local_scores=-0.37854937,
pem=-1.5732348, similarity=1.3485404]

RECALL MISS: 11310306 Hours_of_Jeanne_d'Evreux]

8: 203688 Canonical_hours

[winner: 27857492 Biblical_canon, pem_pos=1,
final_scores=-0.08754283, final_local_scores=-0.039759815,
pem=-1.0716839, similarity=8.731184]

RECALL MISS: 203688 Canonical_hours]

[winner: 80513 Horae, pem_pos=1, final_scores=-1.0505385,
final_local_scores=-1.0255713, pem=-0.6917485,

similarity=-22.753544|
gt: 203688 Canonical_hours, pem_pos=9,
final_scores=-2.201099, final_local_scores=-2.346713,
pem=-4.962145, similarity=-38.955967]
9: 18549 Louis_IX_of_France
[winner: 27687 St._Louis, pem_pos=0,
final_scores=0.17479305, final_local_scores=0.21750423,
pem=-0.092236795, similarity=11.08207|
gt: 18549 Louis_IX_of_France, pem_pos=1,
final_scores=-0.027504683, final_local_scores=-0.11038497,
pem=-2.2484844, similarity=13.023212]
10: 314629 Vitreous_enamel
[winner: 314629 Vitreous_enamel, pem_pos=0,
final_scores=-0.8205156, final_local_scores=-0.8454082,
pem=-0.19492835, similarity=-20.078903|
gt: 314629 Vitreous_enamel, pem_pos=0,
final_scores=-0.8205156, final_local_scores=-0.8454082,
pem=-0.19492835, similarity=-20.078903]
11: 11040959 Château_de_Brie-Comte-Robert
[winner: 61914 Château, pem_pos=0,
final_scores=-1.0889027, final_local_scores=-1.1976526,
pem=-0.044732515, similarity=-31.407366|
gt: 11040959 Château_de_Brie-Comte-Robert, pem_pos=13,
final_scores=-2.3012276, final_local_scores=-2.5122073,
pem=-7.5719886, similarity=-29.701912]
1: 21806207 Margaret_of_Artois
[pem_pos=0, final_scores=0.081212685,
final_local_scores=0.15530214, pem=0.0, similarity=8.724238]
2: 299976 The_Cloisters
[pem_pos=0, final_scores=0.23992999,
final_local_scores=0.2873588, pem=0.0, similarity=12.665072]
3: 37535 Metropolitan_Museum_of_Art
[pem_pos=0, final_scores=0.48358274,
final_local_scores=0.5341471, pem=0.0, similarity=20.029728]
4: 2443414 Jean_Pucelle
[pem_pos=0, final_scores=0.045298263,
final_local_scores=0.11195257, pem=0.0, similarity=7.430601]
5: 5477238 Brie-Comte-Robert
[pem_pos=0, final_scores=0.13391413,
final_local_scores=0.20929858, pem=0.0, similarity=10.3356]
6: 66437 Basilica_of_St_Denis
[pem_pos=0, final_scores=0.20838378,
final_local_scores=0.1721569, pem=0.0, similarity=9.227219]
1: 5843419 France
[pem_pos=0, final_scores=0.2104979,
final_local_scores=0.21624863, pem=-0.16240834,
similarity=11.426206]
2: 1680861 House_of_Capet
[pem_pos=0, final_scores=0.085447654,
final_local_scores=0.043562308, pem=0.0, similarity=5.389699]
3: 37670 Pope_John_XXII
[pem_pos=0, final_scores=0.17676897,

final_local_scores=0.15170541, pem=0.0, similarity=8.616905]
 4: 585820 Jeanne_d'Évreux
 [pem_pos=0, final_scores=0.070874825,
 final_local_scores=0.057011798, pem=0.0, similarity=5.7910585]
 5: 149439 Charles_IV_of_Spain
 [pem_pos=2, final_scores=0.22058158,
 final_local_scores=0.19150665, pem=-0.33441955,
 similarity=11.623279]
 6: 585820 Jeanne_d'Évreux
 [pem_pos=0, final_scores=0.13041376,
 final_local_scores=0.12197107, pem=0.0, similarity=7.729573]
 7: 585820 Jeanne_d'Évreux
 [pem_pos=0, final_scores=0.1823083,
 final_local_scores=0.17873597, pem=0.0, similarity=9.423551]
 8: 73513 Mary_(mother_of_Jesus)
 [pem_pos=0, final_scores=0.14073072,
 final_local_scores=0.148157, pem=0.0, similarity=8.511013]
 9: 645042 New_York_City
 [pem_pos=1, final_scores=0.3295613,
 final_local_scores=0.39068025, pem=-1.7906687,
 similarity=25.486326]
 10: 78500 Charles_IV_of_France
 [pem_pos=0, final_scores=0.3659045,
 final_local_scores=0.35232505, pem=0.0, similarity=14.603795]
 11: 27687 St._Louis
 [pem_pos=0, final_scores=0.17479305,
 final_local_scores=0.21750423, pem=-0.092236795,
 similarity=11.08207]
 12: 585820 Jeanne_d'Évreux
 [pem_pos=0, final_scores=0.14979519,
 final_local_scores=0.1432203, pem=0.0, similarity=8.363692]
 13: 22989 Paris
 [pem_pos=0, final_scores=0.7124369,
 final_local_scores=0.7443379, pem=-0.115301244,
 similarity=26.929268]
 14: 22989 Paris
 [pem_pos=0, final_scores=0.66792214,
 final_local_scores=0.69571227, pem=-0.115301244,
 similarity=25.47818]

Sample: b'M-79'

_ -LRB- Michigan_highway -RRB- M-79 is an east -- west 1 sta
 te trunkline highway in the 2 central portion of [1 1 Lower Pe
 ninsula 1] of the [1 U.S. 1] state of [2 2 Michigan 2] . The
 western terminus is about southeast of [2 3 Hastings 2] at
 the junction with 4 M-37 and the eastern terminus is in
 downtown [3 5 Charlotte 3] at the junction with 6 M-50 and -
 LRB- BL I-69 -RRB- . It passes through and [4 7 Nashville
 4] , where there is a junction with 8 M-66 . The entire
 highway is undivided surface road . It has no direct access
 with 9 Interstate 69 -LRB- I-69 -RRB- , although a sign for
 the highway is located on southbound I-69 at exit 61 . The

highway was first designated in 1919 between [5 Hastings 5] and [3 3 Battle Creek 3] . It was later moved to run to [6 Charlotte 6] . A section of M-79 was designated as M-214 in the 1930s . M-214 would later be decommissioned as a highway designation and the M-79 designation was reapplied to the roadway . M-79 starts south of [7 Hastings 7] at a three-way intersection with M-37 in . The trunkline runs east and parallel to a section of the [4 4 Thornapple River 4] through wooded terrain that is interspersed with farm fields . South of Thornapple Lake , M-79 angles southeasterly moving from Quimby Road to [8 Scott Road 8] running parallel to the [9 Thornapple River 9] again into the community of [10 Nashville 10] . M-79 meets M-66 and turns south along the latter highway 's route along Main and Durkee streets through the community . South of town in Maple Grove Township , M-79 separates from M-66 and turns east along [11 Lawrence Road 11] . Outside of [12 Charlotte 12] , M-79 curves southeast to transition to [13 Lawrence Avenue 13] through town . The eastern terminus is at an intersection with Cochran Avenue , which carries BL I-69 and M-50 . M-79 was formed as a state trunkline by July 1 , 1919 along part of its present routing . At the time , it ran between [14 Hastings 14] and [1 5 Nashville 15] as it does today . From [16 Nashville 16] , the trunkline turned south and west to end at the contemporary 10 M-17 in [5 5 Battle Creek 5] . By 1927 , the southernmost section of M-79 was truncated when 11 M-78 was extended to [17 Battle Creek 17] . In late 1930 , the southernmost section is shortened once again as a new 12 M-14 designation replaced M-79 south of [18 Nashville 18] . An eastward extension of M-79 at the same time carries the highway designation to [6 6 Vermontville 6] . M-79 was extended easterly to [19 Charlotte 19] in 1934 at the same time that the section of the trunkline between [20 Nashville 2 0] and [21 Vermontville 21] was redesignated M-214 . M-79 was rerouted along [22 Assyria 22] and Lawrence roads at the time to connect the sections of M-79 on either side of M-214 . By 1941 , M-214 was shortened to a connector route in downtown [2 3 Nashville 23] running between M-79 and 13 M-66

1: 18317 Lower_Peninsula_of_Michigan

2: 18859 Michigan

3: 23278051 Battle_Creek,_Michigan

4: 8476321 Thornapple_River

5: 23278051 Battle_Creek,_Michigan

6: 117420 Vermontville,_Michigan

1: 23554903 Michigan_State_Trunkline_Highway_System

[winner: 23554903

Michigan_State_Trunkline_Highway_System, pem_pos=0,

final_scores=-1.2136585, final_local_scores=-1.2911019,

pem=0.0, similarity=-34.439346]

gt: 23554903 Michigan_State_Trunkline_Highway_System,

pem_pos=0, final_scores=-1.2136585,

final_local_scores=-1.2911019, pem=0.0,

similarity=-34.439346]
[winner: 1586542 Goulburn_Valley_Highway, pem_pos=17,
final_scores=-1.4865983, final_local_scores=-1.5165126,
pem=-2.8903718, similarity=-25.447754|
RECALL MISS: 23554903
Michigan_State_Trunkline_Highway_System]
2: 6483240 Central_Michigan
[winner: 12593551 The_Central, pem_pos=0,
final_scores=-1.0853558, final_local_scores=-1.0547252,
pem=0.0, similarity=-27.38539|
RECALL MISS: 6483240 Central_Michigan]
[winner: 876106 Parashah, pem_pos=2,
final_scores=-1.2974248, final_local_scores=-1.2715163,
pem=-1.0014485, similarity=-28.408834|
RECALL MISS: 6483240 Central_Michigan]
3: 117088 Hastings,_Michigan
[winner: 13830 Hastings, pem_pos=0,
final_scores=0.07389949, final_local_scores=0.05104485,
pem=-0.5115382, similarity=8.394817|
gt: 117088 Hastings,_Michigan, pem_pos=16,
final_scores=-0.4354689, final_local_scores=-0.4986291,
pem=-4.2751775, similarity=12.458706]
4: 491919 M-37_(Michigan_highway)
[winner: 491919 M-37_(Michigan_highway), pem_pos=0,
final_scores=-0.013159409, final_local_scores=0.019407555,
pem=0.0, similarity=4.6688733|
gt: 491919 M-37_(Michigan_highway), pem_pos=0,
final_scores=-0.013159409, final_local_scores=0.019407555,
pem=0.0, similarity=4.6688733]
5: 117401 Charlotte,_Michigan
[winner: 57447 Charlotte,_North_Carolina, pem_pos=0,
final_scores=0.34224087, final_local_scores=0.30262226,
pem=-0.73618495, similarity=17.124052|
gt: 117401 Charlotte,_Michigan, pem_pos=8,
final_scores=-0.5065565, final_local_scores=-0.6034337,
pem=-4.0852966, similarity=8.298527]
6: 843037 M-50_(Michigan_highway)
[winner: 843037 M-50_(Michigan_highway), pem_pos=0,
final_scores=-0.06430136, final_local_scores=-0.0620495,
pem=0.0, similarity=2.238032|
gt: 843037 M-50_(Michigan_highway), pem_pos=0,
final_scores=-0.06430136, final_local_scores=-0.0620495,
pem=0.0, similarity=2.238032]
7: 117095 Nashville,_Michigan
[winner: 22018 Nashville,_Tennessee, pem_pos=0,
final_scores=0.36876446, final_local_scores=0.33896565,
pem=-0.34690806, similarity=16.091665|
gt: 117095 Nashville,_Michigan, pem_pos=17,
final_scores=-0.60095406, final_local_scores=-0.65076226,
pem=-3.9149144, similarity=5.9595804]
8: 497685 M-66_(Michigan_highway)
[winner: 497685 M-66_(Michigan_highway), pem_pos=0,

```

final_scores=-0.11316295, final_local_scores=-0.10080461,
pem=0.0, similarity=1.0815022|
    gt: 497685 M-66_(Michigan_highway), pem_pos=0,
final_scores=-0.11316295, final_local_scores=-0.10080461,
pem=0.0, similarity=1.0815022 ]
9: 12826921 Interstate_69_in_Michigan
    [winner: 43950 Interstate_Highway_System, pem_pos=0,
final_scores=-0.22882172, final_local_scores=-0.2540035,
pem=-0.8229691, similarity=0.9851756|
    RECALL MISS: 12826921 Interstate_69_in_Michigan]
    [winner: 13496458 Psalm_69, pem_pos=19,
final_scores=-1.08341, final_local_scores=-0.92498577,
pem=-4.0260296, similarity=-1.6195322|
    RECALL MISS: 12826921 Interstate_69_in_Michigan]
10: 1603713 M-17_(Michigan_highway)
    [winner: 1603713 M-17_(Michigan_highway), pem_pos=0,
final_scores=-0.10162346, final_local_scores=-0.06389492,
pem=-0.3988499, similarity=4.351969|
    gt: 1603713 M-17_(Michigan_highway), pem_pos=0,
final_scores=-0.10162346, final_local_scores=-0.06389492,
pem=-0.3988499, similarity=4.351969 ]
11: 497711 M-78_(Michigan_highway)
    [winner: 497711 M-78_(Michigan_highway), pem_pos=0,
final_scores=-0.05100228, final_local_scores=0.0049117357,
pem=0.0, similarity=4.236289|
    gt: 497711 M-78_(Michigan_highway), pem_pos=0,
final_scores=-0.05100228, final_local_scores=0.0049117357,
pem=0.0, similarity=4.236289 ]
12: 844813 M-14_(Michigan_highway)
    [winner: 844813 M-14_(Michigan_highway), pem_pos=1,
final_scores=-0.3542873, final_local_scores=-0.35736823,
pem=-0.47909147, similarity=-3.9694946|
    gt: 844813 M-14_(Michigan_highway), pem_pos=1,
final_scores=-0.3542873, final_local_scores=-0.35736823,
pem=-0.47909147, similarity=-3.9694946 ]
13: 497685 M-66_(Michigan_highway)
    [winner: 497685 M-66_(Michigan_highway), pem_pos=0,
final_scores=-0.342058, final_local_scores=-0.3508377,
pem=0.0, similarity=-6.3799844|
    gt: 497685 M-66_(Michigan_highway), pem_pos=0,
final_scores=-0.342058, final_local_scores=-0.3508377,
pem=0.0, similarity=-6.3799844 ]
1: 18317 Lower_Peninsula_of_Michigan
    [pem_pos=0, final_scores=0.38027072,
final_local_scores=0.4032222, pem=0.0, similarity=16.122667]
2: 18859 Michigan
    [pem_pos=0, final_scores=0.49788898,
final_local_scores=0.46659088, pem=-0.24788691,
similarity=19.361763]
3: 23278051 Battle_Creek,_Michigan
    [pem_pos=0, final_scores=0.3103488,
final_local_scores=0.27367908, pem=-0.06125184,

```

similarity=12.589939]
4: 8476321 Thornapple_River
[pem_pos=0, final_scores=0.16211821,
final_local_scores=0.2179386, pem=0.0, similarity=10.593435]
5: 23278051 Battle_Creek,_Michigan
[pem_pos=0, final_scores=0.40299278,
final_local_scores=0.37487853, pem=-0.06125184,
similarity=15.609934]
6: 117420 Vermontville,_Michigan
[pem_pos=0, final_scores=0.06955738,
final_local_scores=0.095653385, pem=-0.46038085,
similarity=9.447824]
1: 3434750 United_States
[pem_pos=0, final_scores=0.51752424,
final_local_scores=0.52156526, pem=-0.14075686,
similarity=20.41972]
2: 13830 Hastings
[pem_pos=0, final_scores=0.07389949,
final_local_scores=0.05104485, pem=-0.5115382,
similarity=8.394817]
3: 57447 Charlotte,_North_Carolina
[pem_pos=0, final_scores=0.34224087,
final_local_scores=0.30262226, pem=-0.73618495,
similarity=17.124052]
4: 22018 Nashville,_Tennessee
[pem_pos=0, final_scores=0.36876446,
final_local_scores=0.33896565, pem=-0.34690806,
similarity=16.091665]
5: 13830 Hastings
[pem_pos=0, final_scores=0.23684703,
final_local_scores=0.2290403, pem=-0.5115382,
similarity=13.706556]
6: 57447 Charlotte,_North_Carolina
[pem_pos=0, final_scores=0.18874742,
final_local_scores=0.13495404, pem=-0.73618495,
similarity=12.120499]
7: 13830 Hastings
[pem_pos=0, final_scores=0.21088342,
final_local_scores=0.20067897, pem=-0.5115382,
similarity=12.860198]
8: 1702529 Scott_Road_Station
[pem_pos=0, final_scores=0.09339701,
final_local_scores=0.17657956, pem=0.0, similarity=9.3592]
9: 8476321 Thornapple_River
[pem_pos=0, final_scores=0.19529127,
final_local_scores=0.25417516, pem=0.0, similarity=11.674807]
10: 22018 Nashville,_Tennessee
[pem_pos=0, final_scores=0.37871128,
final_local_scores=0.34983099, pem=-0.34690806,
similarity=16.415907]
11: 215563 U.S._Route_206
[pem_pos=0, final_scores=0.3503117,

final_local_scores=0.42237723, pem=0.0, similarity=16.694292]
 12: 57447 Charlotte, North Carolina
 [pem_pos=0, final_scores=0.30896026,
 final_local_scores=0.2662683, pem=-0.73618495,
 similarity=16.039179]
 13: 1646333 Lawrence Avenue
 [pem_pos=0, final_scores=0.3196622,
 final_local_scores=0.35372093, pem=0.0, similarity=14.645452]
 14: 13830 Hastings
 [pem_pos=0, final_scores=0.23789524,
 final_local_scores=0.23018527, pem=-0.5115382,
 similarity=13.740725]
 15: 22018 Nashville, Tennessee
 [pem_pos=0, final_scores=0.4095158,
 final_local_scores=0.3834803, pem=-0.34690806,
 similarity=17.42007]
 16: 22018 Nashville, Tennessee
 [pem_pos=0, final_scores=0.35968047,
 final_local_scores=0.3290428, pem=-0.34690806,
 similarity=15.7955475]
 17: 23278051 Battle Creek, Michigan
 [pem_pos=0, final_scores=0.36961204,
 final_local_scores=0.33841512, pem=-0.06125184,
 similarity=14.521791]
 18: 22018 Nashville, Tennessee
 [pem_pos=0, final_scores=0.28977793,
 final_local_scores=0.2526849, pem=-0.34690806,
 similarity=13.516874]
 19: 57447 Charlotte, North Carolina
 [pem_pos=0, final_scores=0.24446563,
 final_local_scores=0.19581771, pem=-0.73618495,
 similarity=13.936792]
 20: 22018 Nashville, Tennessee
 [pem_pos=0, final_scores=0.31948626,
 final_local_scores=0.2851367, pem=-0.34690806,
 similarity=14.485301]
 21: 117420 Vermontville, Michigan
 [pem_pos=0, final_scores=0.07259704,
 final_local_scores=0.09897375, pem=-0.46038085,
 similarity=9.546909]
 22: 2085 Assyria
 [pem_pos=0, final_scores=0.23675309,
 final_local_scores=0.3566952, pem=0.0, similarity=14.734209]
 23: 22018 Nashville, Tennessee
 [pem_pos=0, final_scores=0.33865112,
 final_local_scores=0.30607134, pem=-0.34690806,
 similarity=15.110033]

Sample: b'Johnny_Moss'

[1 Johnny Moss 1] -LRB- May 14 , 1907 -- December 16 , 1995 -
 RRB- was a 1 gambler and professional 2 poker player . He was
 the first winner of the 3 World Series of Poker -LRB- WSOP -

RRB- Main Event , at the time a 4 cash game event in which he was awarded the title by the vote of his peers in 1970 , He also twice won the current 5 tournament format of the [2 WSOP Main Event 2] in 1971 and 1974 . He was one of the charter inductees into the [1 1 Poker Hall of Fame 1] . [3 Moss 3] was born in 6 Marshall , [4 Texas 4] in 1907 and grew up in , [2 2 Texas 2] where he learned how to gamble as a young boy . A group of cheaters taught him how to cheat in games , but [5 Moss 5] put this knowledge to good use . As a he was hired by a local 7 saloon to watch over games and make sure they were played fairly . While he was keeping games safe from cheaters , he was also learning the strategy behind playing poker . Two years later he became a , a player who travels the country looking for gambling action . In 1949 , [6 Moss 6] played with in a five month long `` heads up `` poker marathon set up by [3 3 Benny Binion 3] , winning between \$ 2 and \$ 4 million . At the conclusion of the game , Nick the [7 Greek 7] uttered what has become one of the most famous poker quotes ever : `` Mr. Moss , I have to let you go . `` This game is often cited as the inspiration behind the [8 WSOP 8] . This game became the foundation for Al Alvarez 's book The Biggest Game in Town and is one of the best known stories in poker . Despite being one of the best known poker stories , a soon to be released book , Showgirl Stories , by [9 Steve Fischer 9] claims the game never took place . According to [10 Fischer 10] , there were no stories or reports of this tournament until six years after Nick 's death . [11 Binion 11] never spoke of the game even when providing a detailed history of [12 Las Vegas 12] and avoided answering questions about the game by saying , `` Well , my memory ai n't what it used to be . `` While Nick the [13 Greek 13] was often covered by the national media , there are no news reports in any [14 Vegas 14] or national source . [15 Fischer 15] says that nearly every version of the story is virtually identical to the version first told by [16 Moss 16] beginning around 1971 . The story is consistently said to have taken place in 1949 at the [17 Horseshoe Casino 17] , a casino that did not exist for another year and a half . [18 Fischer 18] also points out that during the time that [19 Binion 19] supposedly set up the game , he was fighting off a request from [20 Texas 20] to have him extradited . Because of his past , [21 Binion 21] lost his license to run a gambling establishment in 1948 and did not regain it until April 13 , 1950 . He was not granted a license to open the [22 Horseshoe Casino 22] until December 5 , 1952 . [23 Fischer 23] believes that the notion of [24 Binion 24] sponsoring a poker game , in front of a window , of a casino that has n't opened , while fighting extradition is `` absurd . `` [25 Moss 25] won the 8 1970 , 9 1971 , and 10 1974 [4 4 World Series of Poker 4] main events . For the 1970 [26 WSOP 26] , [27 Moss 27] was actually elected champion by his peers and received a silver cup as his prize . A -LRB- possibly

apocryphal -RRB- story about that election which has appeared in print several times has every one of the eight players voting for himself as the best player , and that it was only when the players were asked to vote for the second best player that [28 Moss 28] emerged . He played at every [29 WSOP 29] from 1970 through 11 the 1995 [30 World Series 30] , and during his career he won nine [31 WSOP 31] bracelets , placing him fourth overall , behind [5 5 Johnny Chan 5] and [6 6 Doyle Brunson 6] -LRB- 10 each -RRB- and [7 7 Phil Hellmuth 7]

1: 947231 Poker_Hall_of_Fame
2: 29810 Texas
3: 979605 Benny_Binion
4: 33116 World_Series_of_Poker
5: 923414 Johnny_Chan
6: 255519 Doyle_Brunson
7: 207599 Phil_Hellmuth

1: 11921 Gambling

[winner: 11921 Gambling, pem_pos=0,
final_scores=-1.2707682, final_local_scores=-1.3939629,
pem=-0.025101153, similarity=-37.372417|
gt: 11921 Gambling, pem_pos=0, final_scores=-1.2707682,
final_local_scores=-1.3939629, pem=-0.025101153,
similarity=-37.372417]

2: 23014 Poker

[winner: 8648661 Adam_Schoenfeld, pem_pos=0,
final_scores=-0.78515357, final_local_scores=-0.79306906,
pem=0.0, similarity=-19.57705|
RECALL MISS: 23014 Poker]

3: 33116 World_Series_of_Poker

[winner: 33116 World_Series_of_Poker, pem_pos=0,
final_scores=-0.35196882, final_local_scores=-0.4607546,
pem=0.0, similarity=-9.660124|
gt: 33116 World_Series_of_Poker, pem_pos=0,
final_scores=-0.35196882, final_local_scores=-0.4607546,
pem=0.0, similarity=-9.660124]

4: 9333505 Cash_game

[winner: 21402710 The_Game_(mind_game), pem_pos=3,
final_scores=-1.144746, final_local_scores=-1.1169362,
pem=-1.0411923, similarity=-23.579725|
RECALL MISS: 9333505 Cash_game]

[winner: 16376403 Cash_(Chinese_coin), pem_pos=2,
final_scores=-1.3761129, final_local_scores=-1.3544345,
pem=-1.5984011, similarity=-27.63696|
RECALL MISS: 9333505 Cash_game]

5: 141837 Poker_tournament

[winner: 13405688 V-League_4th_Season_2nd_Conference,
pem_pos=0, final_scores=-1.0041648,
final_local_scores=-0.9818066, pem=0.0,
similarity=-25.209354|

RECALL MISS: 141837 Poker_tournament]

6: 136027 Marshall,_Texas

[winner: 29810 Texas, pem_pos=0, final_scores=0.61126345,
final_local_scores=0.7013816, pem=-0.16191302,
similarity=25.900848|

RECALL MISS: 136027 Marshall,_Texas]

[winner: 136027 Marshall,_Texas, pem_pos=1,
final_scores=-0.017604634, final_local_scores=0.03115189,
pem=-2.915105, similarity=20.872143|

gt: 136027 Marshall,_Texas, pem_pos=1,
final_scores=-0.017604634, final_local_scores=0.03115189,
pem=-2.915105, similarity=20.872143]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8512068, final_local_scores=-1.8413792,
pem=-4.6728287, similarity=-25.449146|

RECALL MISS: 136027 Marshall,_Texas]

7: 272207 Bar_(establishment)

[winner: 381325 Sedan_(automobile), pem_pos=0,
final_scores=-1.1834922, final_local_scores=-1.1652865,
pem=0.0, similarity=-30.684761|

gt: 272207 Bar_(establishment), pem_pos=1,
final_scores=-1.460088, final_local_scores=-1.5694036,
pem=-1.0619726, similarity=-36.969254]

8: 2195921 1970_World_Series_of_Poker

[winner: 34726 1970, pem_pos=0, final_scores=-0.71350163,
final_local_scores=-0.55424476, pem=-0.80837256,
similarity=-8.054001|

RECALL MISS: 2195921 1970_World_Series_of_Poker]

9: 3883586 1971_World_Series_of_Poker

[winner: 34748 1971, pem_pos=0, final_scores=-0.88052887,
final_local_scores=-0.79870826, pem=-0.6167389,
similarity=-16.391413|

RECALL MISS: 3883586 1971_World_Series_of_Poker]

10: 3750706 1974_World_Series_of_Poker

[winner: 34654 1974, pem_pos=0, final_scores=-0.6377812,
final_local_scores=-0.45365342, pem=-0.70012724,
similarity=-5.640809|

RECALL MISS: 3750706 1974_World_Series_of_Poker]

11: 3696471 1995_World_Series_of_Poker

[winner: 23489161 World_Series, pem_pos=0,
final_scores=0.08569722, final_local_scores=0.090525195,
pem=-0.38314086, similarity=8.874745|

RECALL MISS: 3696471 1995_World_Series_of_Poker]

[winner: 34658 1995, pem_pos=0, final_scores=-0.8248188,
final_local_scores=-0.78483737, pem=-0.68201965,
similarity=-15.62247|

RECALL MISS: 3696471 1995_World_Series_of_Poker]

[winner: 218110 Plymouth_Colony, pem_pos=0,
final_scores=-1.58933, final_local_scores=-1.6185136,
pem=-0.4578331, similarity=-41.720192|

RECALL MISS: 3696471 1995_World_Series_of_Poker]

1: 947231 Poker_Hall_of_Fame

[pem_pos=0, final_scores=0.08336814,
final_local_scores=0.057445407, pem=0.0, similarity=5.8039985]

2: 29810 Texas
[pem_pos=0, final_scores=0.5097266,
final_local_scores=0.59046805, pem=-0.16191302,
similarity=22.590965]
3: 979605 Benny_Binion
[pem_pos=0, final_scores=0.3404395,
final_local_scores=0.3308316, pem=0.0, similarity=13.962387]
4: 33116 World_Series_of_Poker
[pem_pos=0, final_scores=0.30338407,
final_local_scores=0.2583563, pem=0.0, similarity=11.799581]
5: 923414 Johnny_Chan
[pem_pos=0, final_scores=0.3770575,
final_local_scores=0.39212227, pem=0.0, similarity=15.791424]
6: 255519 Doyle_Brunson
[pem_pos=0, final_scores=0.5839183,
final_local_scores=0.5983256, pem=0.0, similarity=21.944942]
7: 207599 Phil_Hellmuth
[pem_pos=0, final_scores=0.42911947,
final_local_scores=0.43120956, pem=0.0, similarity=16.957867]
1: 979276 Johnny_Moss
[pem_pos=0, final_scores=0.18909936,
final_local_scores=0.14356089, pem=0.0, similarity=8.373857]
2: 33116 World_Series_of_Poker
[pem_pos=0, final_scores=0.2095307,
final_local_scores=0.15583578, pem=0.0, similarity=8.740164]
3: 979276 Johnny_Moss
[pem_pos=0, final_scores=0.51005554,
final_local_scores=0.49415672, pem=0.0, similarity=18.836334]
4: 29810 Texas
[pem_pos=0, final_scores=0.61126345,
final_local_scores=0.7013816, pem=-0.16191302,
similarity=25.900848]
5: 979276 Johnny_Moss
[pem_pos=0, final_scores=0.34047014,
final_local_scores=0.3089104, pem=0.0, similarity=13.308215]
6: 979276 Johnny_Moss
[pem_pos=0, final_scores=0.5589476,
final_local_scores=0.5475639, pem=0.0, similarity=20.430113]
7: 12108 Greece
[pem_pos=2, final_scores=0.10459228,
final_local_scores=0.21480757, pem=-2.3247876,
similarity=23.142555]
8: 33116 World_Series_of_Poker
[pem_pos=0, final_scores=0.22998635,
final_local_scores=0.18013415, pem=0.0, similarity=9.465275]
9: 2146971 Steve_Fisher
[pem_pos=0, final_scores=0.24696611,
final_local_scores=0.29359862, pem=0.0, similarity=12.851281]
10: 2146971 Steve_Fisher
[pem_pos=0, final_scores=0.2617244,
final_local_scores=0.3097198, pem=0.0, similarity=13.33237]
11: 979605 Benny_Binion

[pem_pos=0, final_scores=0.30192107,
final_local_scores=0.28875607, pem=0.0, similarity=12.70677]
12: 47737 Las_Vegas
[pem_pos=0, final_scores=0.31817377,
final_local_scores=0.313837, pem=-0.49559593,
similarity=16.150362]
13: 12108 Greece
[pem_pos=2, final_scores=0.12026735,
final_local_scores=0.2319302, pem=-2.3247876,
similarity=23.653528]
14: 47737 Las_Vegas
[pem_pos=0, final_scores=0.30968803,
final_local_scores=0.30456766, pem=-0.4826669,
similarity=15.803436]
15: 2146971 Steve_Fisher
[pem_pos=0, final_scores=0.15965728,
final_local_scores=0.19822702, pem=0.0, similarity=10.005202]
16: 979276 Johnny_Moss
[pem_pos=0, final_scores=0.36692506,
final_local_scores=0.33780837, pem=0.0, similarity=14.1705885]
17: 3612525 Binion's_Gambling_Hall_and_Hotel
[pem_pos=0, final_scores=0.062487736,
final_local_scores=0.03330508, pem=-0.31328344,
similarity=6.7872868]
18: 2146971 Steve_Fisher
[pem_pos=0, final_scores=0.32957155,
final_local_scores=0.3838325, pem=0.0, similarity=15.544041]
19: 979605 Benny_Binion
[pem_pos=0, final_scores=0.12966372,
final_local_scores=0.10059106, pem=0.0, similarity=7.091551]
20: 29810 Texas
[pem_pos=0, final_scores=0.6897295,
final_local_scores=0.7870939, pem=-0.16191302,
similarity=28.458672]
21: 979605 Benny_Binion
[pem_pos=0, final_scores=0.13830267,
final_local_scores=0.11002779, pem=0.0, similarity=7.373162]
22: 3612525 Binion's_Gambling_Hall_and_Hotel
[pem_pos=0, final_scores=0.08230488,
final_local_scores=0.054952294, pem=-0.31328344,
similarity=7.433283]
23: 2146971 Steve_Fisher
[pem_pos=0, final_scores=0.14471225,
final_local_scores=0.18190181, pem=0.0, similarity=9.518025]
24: 979605 Benny_Binion
[pem_pos=0, final_scores=0.17978112,
final_local_scores=0.15533671, pem=0.0, similarity=8.72527]
25: 979276 Johnny_Moss
[pem_pos=0, final_scores=0.52371204,
final_local_scores=0.5090744, pem=0.0, similarity=19.281507]
26: 33116 World_Series_of_Poker
[pem_pos=0, final_scores=0.15082343,

```

final_local_scores=0.091707006, pem=0.0, similarity=6.8264327]
27: 979276 Johnny_Moss
    [pem_pos=0, final_scores=0.37701505,
final_local_scores=0.34883022, pem=0.0, similarity=14.499502]
28: 979276 Johnny_Moss
    [pem_pos=0, final_scores=0.40968496,
final_local_scores=0.38451713, pem=0.0, similarity=15.56447]
29: 33116 World_Series_of_Poker
    [pem_pos=0, final_scores=0.18835281,
final_local_scores=0.13270214, pem=0.0, similarity=8.049809]
30: 23489161 World_Series
    [pem_pos=0, final_scores=0.08569722,
final_local_scores=0.090525195, pem=-0.38314086,
similarity=8.874745]
31: 33116 World_Series_of_Poker
    [pem_pos=0, final_scores=0.17532341,
final_local_scores=0.11846951, pem=0.0, similarity=7.6250796]

```

Sample: b'Donald_Hogarth'

Donald McDonald Hogarth -LRB- June 15 , 1879 -- June 27 , 1950 -RRB- was an [1 1 Ontario 1] politician and mining financier . Donald Hogarth was born in in 1879 , the son of [1 William Hogarth 1] , and was educated in [2 2 Mattawa 2] . As an associate of real estate promoter and politician [3 3 John James Carrick 3] , he moved to [2 1 Port Arthur 2] in February 1905 . He was elected to the [4 4 Legislative Assembly of Ontario 4] for [3 2 Port Arthur 3] riding in [5 5 Northern Ontario 5] as a in the 3 December 1911 provincial election . He was re-elected in 4 1914 and 5 1919 , serving until May 1923 . Re-entering provincial politics , he was elected as an Independent-Conservative in December 1926 and re-elected as a [4 Conservative 4] in October 1929 , ending his political career in May 1934 . During his long political career he focused his efforts on the development of the mining and 6 pulp and paper industry . While still a member of the assembly , Don Hogarth enlisted in 1914 for [6 6 World War I 6] and rose rapidly through the [5 Canadian 5] military ranks , from lieutenant to captain to major in 1915 in charge of military supplies and transport in [6 London 6] . In January 1917 he was appointed a and made director of supply and transport for the [7 Canadian 7] forces . He was awarded the DSO -LRB- Distinguished Service Order -RRB- in June 1917 , the year he became acting 7 quartermaster - general of the [7 7 Canadian Expeditionary Force 7] . In February 1918 he was appointed Quartermaster-General with the rank of , and left the military in 1919 with the rank of . As a mining financier and political operator , he was associated with the Little Long Lac gold mine near . His greatest venture was the development of the at 8 Steep Rock Lake near [8 8 Atikokan 8] . He was inducted into the [9 9 Canadian Mining Hall of Fame 9] . He died in [10 10 Toronto 10]

1: 22218 Ontario

2: 2080240 Mattawa,_Ontario
3: 13139509 John_James_Carrick
4: 661462 Legislative_Assembly_of_Ontario
5: 211003 Northern_Ontario
6: 4764461 World_War_I
7: 1182044 Canadian_Expeditinary_Force
8: 1926018 Atikokan
9: 5629558 Canadian_Mining_Hall_of_Fame
10: 64646 Toronto

1: 479703 Port_Arthur,_Ontario

[winner: 265537 Port_Arthur,_Tasmania, pem_pos=2,
final_scores=0.17416655, final_local_scores=0.23609653,
pem=-1.4185004, similarity=18.84933|

gt: 479703 Port_Arthur,_Ontario, pem_pos=3,
final_scores=-0.074022904, final_local_scores=-0.13384634,
pem=-1.8565243, similarity=10.191539]

2: 1793223 Port_Arthur_(electoral_district)

[winner: 265537 Port_Arthur,_Tasmania, pem_pos=2,
final_scores=0.21742068, final_local_scores=0.28334513,
pem=-1.4185004, similarity=20.259323|

gt: 1793223 Port_Arthur_(electoral_district), pem_pos=5,
final_scores=-0.224333, final_local_scores=-0.22833121,
pem=-2.1498647, similarity=8.967153]

3: 1666927 Ontario_general_election,_1911

[winner: 1666927 Ontario_general_election,_1911,
pem_pos=0, final_scores=-0.54771096,
final_local_scores=-0.5099102, pem=0.0,
similarity=-11.127025|

gt: 1666927 Ontario_general_election,_1911, pem_pos=0,
final_scores=-0.54771096, final_local_scores=-0.5099102,
pem=0.0, similarity=-11.127025]

4: 1666947 Ontario_general_election,_1914

[winner: 19283920 1914, pem_pos=0,
final_scores=-1.1733176, final_local_scores=-1.1064014,
pem=-0.7211613, similarity=-25.005726|

RECALL MISS: 1666947 Ontario_general_election,_1914]

5: 1667045 Ontario_general_election,_1919

[winner: 34665 1919, pem_pos=0, final_scores=-0.9395587,
final_local_scores=-0.8304854, pem=-0.5577421,
similarity=-17.66054|

RECALL MISS: 1667045 Ontario_general_election,_1919]

6: 293465 Pulp_and_paper_industry

[winner: 13018743 Champion_International_Paper,
pem_pos=0, final_scores=-0.77137184,
final_local_scores=-0.72870666, pem=0.0,
similarity=-17.656347|

gt: 293465 Pulp_and_paper_industry, pem_pos=1,
final_scores=-1.0388526, final_local_scores=-1.0614097,
pem=-0.055030085, similarity=-27.285606]

[winner: 17653927 Pulp_(juice), pem_pos=5,
final_scores=-0.92420304, final_local_scores=-0.78486145,
pem=-2.0563924, similarity=-8.149136|

RECALL MISS: 293465 Pulp_and_paper_industry]
7: 1071259 Quartermaster
[winner: 973994 Quartermaster_sergeant_instructor,
pem_pos=3, final_scores=-1.2133517,
final_local_scores=-1.1973181, pem=-1.0986123,
similarity=-25.666225]
gt: 1071259 Quartermaster, pem_pos=0,
final_scores=-1.312093, final_local_scores=-1.4215384,
pem=-0.07179593, similarity=-37.941387]
8: 33454100 Steep_Rock_Lake
[winner: 33454100 Steep_Rock_Lake, pem_pos=0,
final_scores=-0.10129544, final_local_scores=0.011382848,
pem=0.0, similarity=4.4294]
gt: 33454100 Steep_Rock_Lake, pem_pos=0,
final_scores=-0.10129544, final_local_scores=0.011382848,
pem=0.0, similarity=4.4294]
1: 22218 Ontario
[pem_pos=0, final_scores=0.4659474,
final_local_scores=0.4372638, pem=-0.10088341,
similarity=17.687157]
2: 2080240 Mattawa,_Ontario
[pem_pos=0, final_scores=0.15588833,
final_local_scores=0.16252276, pem=-0.35459015,
similarity=10.868032]
3: 13139509 John_James_Carrick
[pem_pos=0, final_scores=0.083150074,
final_local_scores=0.1326797, pem=0.0, similarity=8.04914]
4: 661462 Legislative_Assembly_of_Ontario
[pem_pos=0, final_scores=0.04665725,
final_local_scores=0.029455185, pem=0.0, similarity=4.968714]
5: 211003 Northern_Ontario
[pem_pos=1, final_scores=0.36830145,
final_local_scores=0.36157966, pem=-0.07787235,
similarity=15.3034525]
6: 4764461 World_War_I
[pem_pos=0, final_scores=0.26042384,
final_local_scores=0.2685526, pem=0.0, similarity=12.103858]
7: 1182044 Canadian_Expeditionary_Force
[pem_pos=0, final_scores=0.15419455,
final_local_scores=0.15850165, pem=0.0, similarity=8.819718]
8: 1926018 Atikokan
[pem_pos=0, final_scores=0.13198303,
final_local_scores=0.14060655, pem=0.0, similarity=8.285693]
9: 5629558 Canadian_Mining_Hall_of_Fame
[pem_pos=0, final_scores=0.06308006,
final_local_scores=0.1194742, pem=0.0, similarity=7.6550617]
10: 64646 Toronto
[pem_pos=0, final_scores=0.22803883,
final_local_scores=0.19385567, pem=-0.12291032,
similarity=10.543159]
1: 103020 William_Hogarth
[pem_pos=0, final_scores=0.17649452,

final_local_scores=0.25451788, pem=0.0, similarity=11.685034]
 2: 265537 Port_Arthur,_Tasmania
 [pem_pos=2, final_scores=0.17416655,
 final_local_scores=0.23609653, pem=-1.4185004,
 similarity=18.84933]
 3: 265537 Port_Arthur,_Tasmania
 [pem_pos=2, final_scores=0.21742068,
 final_local_scores=0.28334513, pem=-1.4185004,
 similarity=20.259323]
 4: 32113 Conservative_Party_(UK)
 [pem_pos=0, final_scores=0.045046613,
 final_local_scores=0.05203271, pem=-0.55914843,
 similarity=8.683208]
 5: 5042916 Canada
 [pem_pos=0, final_scores=0.57843024,
 final_local_scores=0.5675506, pem=-0.41303828,
 similarity=23.27272]
 6: 17867 London
 [pem_pos=0, final_scores=0.5423967,
 final_local_scores=0.5625958, pem=-0.09086737,
 similarity=21.372843]
 7: 5042916 Canada
 [pem_pos=0, final_scores=0.5614361,
 final_local_scores=0.5489871, pem=-0.41303828,
 similarity=22.718748]

Sample: b'Luc_Alphand'

[1 Luc Alphand 1] -LRB- born 6 August 1965 -RRB- is a former
 1 World Cup 2 alpine ski racer from [1 1 France 1] . He
 specialized in the speed events and later became a 3 race car
 driver . Born in [2 2 Briançon 2] -LRB- 4 Hautes-Alpes -
 RRB- , [2 Alphand 2] made his World Cup debut in 1984 . In 5 1
 997 he won the World Cup overall title by collecting points
 only in the two speed disciplines , downhill and super G -- a
 unique achievement in World Cup history . The previous year
 he won a bronze medal at the 6 1996 World Championships at [3
 7 Sierra Nevada 3] , [3 3 Spain 3] . In the village of
 Chantemerle -LRB- which neighbours his home town of [4 4 Bria
 nçon 4] in the region -RRB- an expert ski run has been named
 in his honour . He retired from competitive skiing in 1997
 and started a career in auto racing . First in the [5 5 Nissan
 Micra Stars Cup 5] -LRB- 1997 -- 1998 -RRB- , then in the [4 8
 European Le Mans Series 4] -LRB- 2001 -RRB- , the [6 6 FIA GT
 Championship 6] -LRB- 2002 -RRB- , and the -LRB- 2002 -RRB- .
 He won the 2006 [7 7 Dakar Rally 7] , in which he had
 finished runner-up a year earlier . With this victory , he
 was the first ex-skier to win the [5 Paris-Dakar 5] . He
 recently purchased two [6 Corvette 6] race cars from [8 8 Pra
 tt & Miller 8] for use in the and 9 24 Hours of Le Mans

1: 5843419 France
 2: 445836 Briançon
 3: 26667 Spain

4: 445836 Briançon
 5: 750514 Nissan_Micra
 6: 2345752 FIA_GT_Championship
 7: 25136 Dakar_Rally
 8: 13059391 Pratt_Miller
 1: 183628 FIS_Alpine_Ski_World_Cup
 [winner: 11370 FIFA_World_Cup, pem_pos=0,
 final_scores=-0.13993046, final_local_scores=-0.1269604,
 pem=-1.001087, similarity=5.745028|
 gt: 183628 FIS_Alpine_Ski_World_Cup, pem_pos=2,
 final_scores=-0.1487231, final_local_scores=-0.16836989,
 pem=-2.7628422, similarity=14.089986]
 2: 39576 Alpine_skiing
 [winner: 679433 Ski_warfare, pem_pos=2,
 final_scores=-0.9441317, final_local_scores=-0.951153,
 pem=-0.69268847, similarity=-20.527641|
 gt: 39576 Alpine_skiing, pem_pos=1,
 final_scores=-1.0237478, final_local_scores=-1.0931126,
 pem=-0.65621436, similarity=-24.96235]
 [winner: 2619936 Racer_(Kennywood), pem_pos=2,
 final_scores=-0.9669305, final_local_scores=-0.89965004,
 pem=-1.4941926, similarity=-14.631988|
 RECALL MISS: 39576 Alpine_skiing]
 [winner: 23160842 Alpine_Mendocino_County_California,
 pem_pos=9, final_scores=-1.1510377,
 final_local_scores=-1.0740728, pem=-3.0050297,
 similarity=-11.620947|
 gt: 39576 Alpine_skiing, pem_pos=13,
 final_scores=-1.9906975, final_local_scores=-2.1493583,
 pem=-3.601453, similarity=-40.46616]
 3: 1022 Auto_racing
 [winner: 1724720 Car_Driver, pem_pos=0,
 final_scores=-0.93611634, final_local_scores=-0.848262,
 pem=0.0, similarity=-21.224117|
 RECALL MISS: 1022 Auto_racing]
 [winner: 7885197
 2004_Deutsche_Tourenwagen_Masters_season, pem_pos=14,
 final_scores=-1.3627342, final_local_scores=-1.398004,
 pem=-2.944439, similarity=-21.6172|
 RECALL MISS: 1022 Auto_racing]
 4: 83190 Hautes-Alpes
 [winner: 83190 Hautes-Alpes, pem_pos=0,
 final_scores=-0.4845001, final_local_scores=-0.48402604,
 pem=0.0, similarity=-10.354589|
 gt: 83190 Hautes-Alpes, pem_pos=0,
 final_scores=-0.4845001, final_local_scores=-0.48402604,
 pem=0.0, similarity=-10.354589]
 5: 4241220 1997_Alpine_Skiing_World_Cup
 [winner: 34601 1997, pem_pos=0, final_scores=-1.0414052,
 final_local_scores=-0.9705064, pem=-0.69383013,
 similarity=-21.09898|
 RECALL MISS: 4241220 1997_Alpine_Skiing_World_Cup]

6: 9079425 FIS_Alpine_World_Ski_Championships_1996
 [winner: 218926 IAAF_World_Championships_in_Athletics,
 pem_pos=0, final_scores=-0.22456013,
 final_local_scores=-0.21159199, pem=-2.1703603,
 similarity=9.578143]
 RECALL MISS: 9079425
 FIS_Alpine_World_Ski_Championships_1996]
 [winner: 34636 1996, pem_pos=0, final_scores=-0.75006306,
 final_local_scores=-0.55644286, pem=-0.78063476,
 similarity=-8.270439]
 RECALL MISS: 9079425
 FIS_Alpine_World_Ski_Championships_1996]

7: 30735525 Sierra_Nevada_Ski_Station
 [winner: 50413 Sierra_Nevada_(U.S.), pem_pos=0,
 final_scores=0.307046, final_local_scores=0.44098735,
 pem=-0.06485869, similarity=17.602367]
 gt: 30735525 Sierra_Nevada_Ski_Station, pem_pos=12,
 final_scores=-0.77481914, final_local_scores=-0.7620582,
 pem=-6.074975, similarity=14.385036]

8: 7027797 2001_European_Le_Mans_Series_season
 [winner: 895712 European_Le_Mans_Series, pem_pos=0,
 final_scores=0.127272, final_local_scores=0.12181279,
 pem=0.0, similarity=7.7248497]
 gt: 7027797 2001_European_Le_Mans_Series_season,
 pem_pos=1, final_scores=-0.2841069,
 final_local_scores=-0.29514313, pem=-2.2454267,
 similarity=7.4930325]

9: 1401596 24_Hours_of_Le_Mans
 [winner: 1401596 24_Hours_of_Le_Mans, pem_pos=0,
 final_scores=-0.113553226, final_local_scores=-0.17647532,
 pem=-0.05606659, similarity=-0.8717637]
 gt: 1401596 24_Hours_of_Le_Mans, pem_pos=0,
 final_scores=-0.113553226, final_local_scores=-0.17647532,
 pem=-0.05606659, similarity=-0.8717637]

1: 5843419 France
 [pem_pos=0, final_scores=0.6645071,
 final_local_scores=0.72168446, pem=-0.16240834,
 similarity=26.509418]

2: 445836 Briançon
 [pem_pos=0, final_scores=0.13606988,
 final_local_scores=0.1537273, pem=-0.21047623,
 similarity=9.821845]

3: 26667 Spain
 [pem_pos=0, final_scores=0.60673803,
 final_local_scores=0.684304, pem=-0.18857051,
 similarity=25.536184]

4: 445836 Briançon
 [pem_pos=0, final_scores=0.38814688,
 final_local_scores=0.4290831, pem=-0.21047623,
 similarity=18.039011]

5: 750514 Nissan_Micra
 [pem_pos=0, final_scores=0.10534774,

final_local_scores=0.16144586, pem=0.0, similarity=8.907579]
 6: 2345752 FIA_GT_Championship
 [pem_pos=0, final_scores=0.15755914,
 final_local_scores=0.14995155, pem=-0.035444517,
 similarity=8.757319]
 7: 25136 Dakar_Rally
 [pem_pos=0, final_scores=0.054565117,
 final_local_scores=0.0036465675, pem=-0.014799432,
 similarity=4.2790155]
 8: 13059391 Pratt_&_Miller
 [pem_pos=0, final_scores=0.20441292,
 final_local_scores=0.22156736, pem=0.0, similarity=10.701725]
 1: 1616581 Luc_Alphand
 [pem_pos=0, final_scores=0.13297342,
 final_local_scores=0.105489284, pem=0.0, similarity=7.2377234]
 2: 1616581 Luc_Alphand
 [pem_pos=1, final_scores=0.3830282,
 final_local_scores=0.39414316, pem=0.0, similarity=15.85173]
 3: 50413 Sierra_Nevada_(U.S.)
 [pem_pos=0, final_scores=0.307046,
 final_local_scores=0.44098735, pem=-0.06485869,
 similarity=17.602367]
 4: 895712 European_Le_Mans_Series
 [pem_pos=0, final_scores=0.127272,
 final_local_scores=0.12181279, pem=0.0, similarity=7.7248497]
 5: 25136 Dakar_Rally
 [pem_pos=0, final_scores=0.31835955,
 final_local_scores=0.2918019, pem=0.0, similarity=12.797664]
 6: 221005 Chevrolet_Corvette
 [pem_pos=0, final_scores=0.19989966,
 final_local_scores=0.1981186, pem=-0.56036633,
 similarity=13.049326]

Sample: b'Electoral_district_of_Murray-Darling'

[1 Murray-Darling 1] is an of the 1 Legislative Assembly in
 the [1 1 Australia 1] n state of [2 2 New South Wales 2] . It
 is represented by 2 John Williams of the [3 3 National Party o
 f Australia 3] . It includes [4 4 Urana Shire 4] , [5 5 Jeril
 derie Shire 5] , [6 6 Berrigan Shire 6] , [7 7 Murray Shire 7]
 , 3 Deniliquin Council , [8 8 Conargo Shire 8] , 4 Wakool Shi
 re , [9 9 Hay Shire 9] , [10 10 Balranald Shire 10] , [11 11 W
 entworth Shire 11] , part of [12 12 Carrathool Shire 12] -
 LRB- including and -RRB- , [13 13 Central Darling Shire 13] ,
 the 5 City of Broken Hill and the . 6 Murray was a single
 member electorate from 1859 to 1920 . From 1920 to 1927 , it
 was merged with [2 7 Albury 2] , 8 Corowa and [3 9 Wagga Wagga
 3] , to create a four-member electorate , with voting by 10 pr
 oportional representation . It was recreated as a single-
 member electorate in 1927 . In 1999 , [4 11 Broken Hill 4]
 1: 4689264 Australia
 2: 21654 New_South_Wales
 3: 21927 National_Party_of_Australia

4: 7836936 Urana_Shire
 5: 1940257 Jerilderie_Shire
 6: 7467895 Berrigan_Shire
 7: 5040539 Murray_Shire
 8: 3982813 Conargo_Shire
 9: 7848553 Hay_Shire
 10: 7849288 Balranald_Shire
 11: 7851209 Wentworth_Shire
 12: 5101397 Carrathool_Shire
 13: 5101666 Central_Darling_Shire
 1: 579433 New_South_Wales_Legislative_Assembly
 [winner: 476068 California_State_Assembly, pem_pos=0,
 final_scores=-0.32659906, final_local_scores=-0.30978185,
 pem=-0.633089, similarity=-1.7119592]
 RECALL MISS: 579433
 New_South_Wales_Legislative_Assembly]
 [winner: 13020463 Legislative_Branch_of_Colombia,
 pem_pos=1, final_scores=-1.0394877,
 final_local_scores=-1.0012846, pem=-2.1536694,
 similarity=-14.07863]
 RECALL MISS: 579433 New_South_Wales_Legislative_Assembly]
 2: 13940236 John_Williams_(New_South_Wales_state_politician)
 [winner: 154554 John_Williams, pem_pos=0,
 final_scores=-0.48487926, final_local_scores=-0.3859122,
 pem=-0.33463192, similarity=-5.6068964]
 gt: 13940236
 John_Williams_(New_South_Wales_state_politician), pem_pos=17,
 final_scores=-0.8471709, final_local_scores=-0.8581962,
 pem=-4.3015866, similarity=1.8721237]
 3: 7846044 Deniliquin_Council
 [winner: 7846044 Deniliquin_Council, pem_pos=0,
 final_scores=0.0072146803, final_local_scores=-0.013727747,
 pem=0.0, similarity=3.6800497]
 gt: 7846044 Deniliquin_Council, pem_pos=0,
 final_scores=0.0072146803, final_local_scores=-0.013727747,
 pem=0.0, similarity=3.6800497]
 4: 5114278 Wakool_Shire
 [winner: 5114278 Wakool_Shire, pem_pos=0,
 final_scores=0.0005493313, final_local_scores=-0.033016317,
 pem=0.0, similarity=3.1044402]
 gt: 5114278 Wakool_Shire, pem_pos=0,
 final_scores=0.0005493313, final_local_scores=-0.033016317,
 pem=0.0, similarity=3.1044402]
 5: 7782272 City_of_Broken_Hill
 [winner: 150204 Broken_Hill, New_South_Wales, pem_pos=0,
 final_scores=-0.048569277, final_local_scores=-0.08605356,
 pem=-0.022847641, similarity=1.6459521]
 gt: 7782272 City_of_Broken_Hill, pem_pos=8,
 final_scores=-0.7886047, final_local_scores=-0.88262975,
 pem=-4.1390214, similarity=0.25892314]
 [winner: 2318702 City_of, pem_pos=0,
 final_scores=-0.90347785, final_local_scores=-0.7930127,

pem=0.0, similarity=-19.575367|
 RECALL MISS: 7782272 City_of_Broken_Hill]

6: 23261115 Electoral_district_of_Murray_(New_South_Wales)
 [winner: 19811 Murray_River, pem_pos=3,
 final_scores=-0.17650549, final_local_scores=-0.22063392,
 pem=-3.2232785, similarity=15.034246|
 gt: 23261115
 Electoral_district_of_Murray_(New_South_Wales), pem_pos=5,
 final_scores=-0.64795804, final_local_scores=-0.6761277,
 pem=-3.4696038, similarity=2.7809598]

7: 2525757 Electoral_district_of_Albury
 [winner: 59047 Albury, pem_pos=0,
 final_scores=0.123328224, final_local_scores=0.108867586,
 pem=-0.33708054, similarity=9.171635|
 gt: 2525757 Electoral_district_of_Albury, pem_pos=1,
 final_scores=-0.30703282, final_local_scores=-0.35286087,
 pem=-1.788152, similarity=3.2838886]

8: 10139463 Electoral_district_of_Corowa
 [winner: 2413567 Corowa, pem_pos=0,
 final_scores=0.028610602, final_local_scores=0.02197291,
 pem=-0.17408814, similarity=5.692147|
 gt: 10139463 Electoral_district_of_Corowa, pem_pos=1,
 final_scores=-0.28279042, final_local_scores=-0.25429976,
 pem=-0.77046776, similarity=0.69082457]

9: 4047689 Electoral_district_of_Wagga_Wagga
 [winner: 349313 Wagga_Wagga, pem_pos=0,
 final_scores=0.35771555, final_local_scores=0.3553113,
 pem=0.0, similarity=14.692911|
 gt: 4047689 Electoral_district_of_Wagga_Wagga,
 pem_pos=1, final_scores=-0.010479733,
 final_local_scores=0.0072219074, pem=-0.8719583,
 similarity=9.047073]

10: 23298 Proportional_representation
 [winner: 4415141 Wold's_theorem, pem_pos=8,
 final_scores=-1.1237241, final_local_scores=-1.0155387,
 pem=-2.0768921, similarity=-14.921527|
 gt: 23298 Proportional_representation, pem_pos=6,
 final_scores=-1.6349874, final_local_scores=-1.76368,
 pem=-2.0617309, similarity=-37.329998]
 [winner: 81863 Proportionality_(mathematics), pem_pos=0,
 final_scores=-1.2075422, final_local_scores=-1.15816,
 pem=0.0, similarity=-30.47209|
 gt: 23298 Proportional_representation, pem_pos=1,
 final_scores=-1.3286915, final_local_scores=-1.4290984,
 pem=-0.46681154, similarity=-36.01884]

11: 10136005 Electoral_district_of_Broken_Hill
 [winner: 150204 Broken_Hill, New_South_Wales, pem_pos=0,
 final_scores=0.073986396, final_local_scores=0.059900522,
 pem=-0.022847641, similarity=6.001513|
 gt: 10136005 Electoral_district_of_Broken_Hill,
 pem_pos=1, final_scores=-0.23299973,
 final_local_scores=-0.2358002, pem=-1.2270454,

similarity=3.7258284]

1: 4689264 Australia

[pem_pos=0, final_scores=0.5204184,
final_local_scores=0.5644742, pem=-0.17168477,
similarity=21.868399]

2: 21654 New_South_Wales

[pem_pos=0, final_scores=0.69137734,
final_local_scores=0.7147025, pem=-0.051332146,
similarity=25.697014]

3: 21927 National_Party_of_Australia

[pem_pos=0, final_scores=0.23756997,
final_local_scores=0.23555487, pem=0.0, similarity=11.119141]

4: 7836936 Urana_Shire

[pem_pos=0, final_scores=0.2271593,
final_local_scores=0.22777206, pem=0.0, similarity=10.886886]

5: 1940257 Jerilderie_Shire

[pem_pos=0, final_scores=0.16875027,
final_local_scores=0.1642645, pem=0.0, similarity=8.9916935]

6: 7467895 Berrigan_Shire

[pem_pos=0, final_scores=0.19319455,
final_local_scores=0.18762314, pem=0.0, similarity=9.688762]

7: 5040539 Murray_Shire

[pem_pos=0, final_scores=0.18143438,
final_local_scores=0.17633978, pem=0.0, similarity=9.352043]

8: 3982813 Conargo_Shire

[pem_pos=0, final_scores=0.2605092,
final_local_scores=0.26980525, pem=0.0, similarity=12.14124]

9: 7848553 Hay_Shire

[pem_pos=0, final_scores=0.077203915,
final_local_scores=0.06572203, pem=0.0, similarity=6.0509896]

10: 7849288 Balranald_Shire

[pem_pos=0, final_scores=0.09050588,
final_local_scores=0.082706794, pem=0.0, similarity=6.5578485]

11: 7851209 Wentworth_Shire

[pem_pos=0, final_scores=0.0689552,
final_local_scores=0.054485008, pem=0.0, similarity=5.7156544]

12: 5101397 Carrathool_Shire

[pem_pos=0, final_scores=0.14859663,
final_local_scores=0.15273312, pem=0.0, similarity=8.647573]

13: 5101666 Central_Darling_Shire

[pem_pos=0, final_scores=0.13627191,
final_local_scores=0.13658589, pem=0.0, similarity=8.165709]

1: 517404 Murray-Darling_basin

[pem_pos=1, final_scores=0.09166591,
final_local_scores=0.10515222, pem=-0.047644015,
similarity=7.4867606]

2: 59047 Albury

[pem_pos=0, final_scores=0.123328224,
final_local_scores=0.108867586, pem=-0.33708054,
similarity=9.171635]

3: 349313 Wagga_Wagga

[pem_pos=0, final_scores=0.35771555,

final_local_scores=0.3553113, pem=0.0, similarity=14.692911]
4: 150204 Broken_Hill, New_South_Wales
[pem_pos=0, final_scores=0.073986396,
final_local_scores=0.059900522, pem=-0.022847641,
similarity=6.001513]

Sample: b'Shorkot'

[1 Shorkot 1] -LRB- -RRB- , -LRB- -RRB- , is a city in the ,
[1 1 Pakistan 1] . It is also a capital city of [2 2 Shorkot T
ehsil 2] in [2 1 Jhang 2] district . It is located at 30 ° 30
'N 72 ° 24 ' E with an altitude of 131 metres -LRB- 433 ft -
RRB- . The city is famous for the tombs of s [3 3 Sultan Bahu
3] , Shah Mehmood Ghazi -LRB- also known as `` Ghazi Pir '' -
RRB- and Syed Mahboob Alam Shah Gilani . The tomb of the
latter is in the centre of the town . He was sent to the [3
Shorkot 3] by the 2 Mughal Emperor [4 3 Shah Alam 4] , to
spread [4 4 Islam 4] . faced bitter resistance in [5 Shorkot 5
] when fighting the 4 Malhi in 325 BC . Dray horses were
taken as presents from 5 Queen Victoria to [6 6 Ranjit Singh 6
] from [5 5 Multan 5] to [6 6 Lahore 6] , they stopped near [7
Shorkot 7] and visited the shrine of Pir Syed Mahboob Alam
Shah , found it to be a dusty shrine attended by an old [8
Syed 8] . There on the [7 7 Bhirr 7] -LRB- mound -RRB- , they
found some old coins with the [9 7 Greek 9] word Bazileous
stamped on it and conjectured that these were the coins left
behind by the [10 8 Greeks 10] . Scattered shards of pottery
are still to be found for miles around the old [11 Bhirr 11]
showing that it was an extensive settlement . [12 Bhirr 12]
is the historical castle covered some acres of land , with a
200 feet high tower located near [13 Shorkot 13] . The
original settlements were the part of the old civilisation .
The 9 Thal town is a mound , built with the material from its
southern side . This mound 's surface has three layers of
soot showing that it has been destroyed three times . The [8 8
Chenab 8] river was lied near it but later moved away to the
west . Darbar Maan Baap shrine is located in [14 Shorkot
14] , were the parents of Sultan Bahu . Darbar Maan Baap is a
notable place in [15 Shorkot 15] where many people come from
far areas to reciting [16 10 Fateh 16] on shrine . In [9 9 Mu
harram 9] there is great crowd of people found there . The 11
Sarfraz Rafiqi Airbase

1: 23235 Pakistan

2: 18896748 Shorkot_Tehsil

3: 1444386 Sultan_Bahu

4: 6037917 Islam

5: 433843 Multan

6: 125315 Lahore

7: 30531541 Bhirr

8: 1100160 Chenab_River

9: 444123 Muharram

1: 2792918 Jhang_District

[winner: 2792911 Jhang, pem_pos=0,

```

final_scores=0.44459277, final_local_scores=0.44880253,
pem=0.0, similarity=17.482876|
    gt: 2792918 Jhang_District, pem_pos=1,
final_scores=0.006293997, final_local_scores=-0.014599919,
pem=-1.4979004, similarity=11.799837 ]
2: 432324 Mughal_emperors
    [winner: 24217897 Mughal_Empire, pem_pos=0,
final_scores=0.00753133, final_local_scores=-0.041990206,
pem=-0.1372593, similarity=3.5830789|
    gt: 432324 Mughal_emperors, pem_pos=14,
final_scores=-1.0451128, final_local_scores=-1.1801469,
pem=-6.272158, similarity=2.9807427 ]
    [winner: 11962817 Emperor_(Stephen_Baxter), pem_pos=15,
final_scores=-1.4635423, final_local_scores=-1.4132359,
pem=-4.276575, similarity=-14.827385|
    RECALL MISS: 432324 Mughal_emperors]
3: 1934790 Shah_Alam_II
    [winner: 113668 Shah_Alam, pem_pos=0,
final_scores=0.26503402, final_local_scores=0.33117345,
pem=0.0, similarity=13.9725895|
    gt: 1934790 Shah_Alam_II, pem_pos=1,
final_scores=0.12194757, final_local_scores=0.15131944,
pem=-1.3736767, similarity=16.075655 ]
4: 1942619 Malhi
    [winner: 1942619 Malhi, pem_pos=0,
final_scores=-0.20112662, final_local_scores=-0.15698072,
pem=0.0, similarity=-0.594905|
    gt: 1942619 Malhi, pem_pos=0, final_scores=-0.20112662,
final_local_scores=-0.15698072, pem=0.0,
similarity=-0.594905 ]
5: 47923 Queen_Victoria
    [winner: 4689460 Victoria_(Australia), pem_pos=0,
final_scores=-0.048189312, final_local_scores=0.010228127,
pem=-0.6703371, similarity=8.040338|
    gt: 47923 Queen_Victoria, pem_pos=2,
final_scores=-0.18989049, final_local_scores=-0.1673181,
pem=-2.8769097, similarity=14.74169 ]
    [winner: 42010 Queen_(band), pem_pos=0,
final_scores=-1.3086383, final_local_scores=-1.2958133,
pem=-0.43144405, similarity=-32.233685|
    gt: 47923 Queen_Victoria, pem_pos=9,
final_scores=-1.6981616, final_local_scores=-1.820092,
pem=-4.022664, similarity=-28.349586 ]
6: 6208485 Ranjit_Singh_of_Bharatpur
    [winner: 519237 Ranjit_Singh, pem_pos=0,
final_scores=0.23553427, final_local_scores=0.22202241,
pem=0.0, similarity=10.715304|
    gt: 6208485 Ranjit_Singh_of_Bharatpur, pem_pos=1,
final_scores=-0.15098627, final_local_scores=-0.06833628,
pem=-0.69166946, similarity=5.811828 ]
7: 11887 Greek_language
    [winner: 12108 Greece, pem_pos=1,

```

```

final_scores=0.15583102, final_local_scores=0.22223574,
pem=-2.3247876, similarity=23.364227]
    gt: 11887 Greek_language, pem_pos=0,
final_scores=-0.31151026, final_local_scores=-0.24659988,
pem=-0.6261292, similarity=0.13566864 ]
8: 42056 Greeks
    [winner: 12108 Greece, pem_pos=2,
final_scores=0.11392777, final_local_scores=0.18596971,
pem=-2.5869315, similarity=23.707556]
    gt: 42056 Greeks, pem_pos=0, final_scores=-0.025113195,
final_local_scores=0.019797117, pem=-0.48669693,
similarity=7.327232 ]
9: 6013019 Thal_Desert
    [winner: 542024 Thal,_Styria, pem_pos=1,
final_scores=0.016166762, final_local_scores=0.11099215,
pem=-0.96370983, similarity=12.642744]
    gt: 6013019 Thal_Desert, pem_pos=0,
final_scores=-0.033838406, final_local_scores=-0.0067369193,
pem=-0.8005853, similarity=8.242377 ]
10: 1100352 Al-Fatiha
    [winner: 11104 Fatah, pem_pos=0, final_scores=0.1535586,
final_local_scores=0.22879583, pem=-0.718465,
similarity=14.824561]
    gt: 1100352 Al-Fatiha, pem_pos=17,
final_scores=-0.66156346, final_local_scores=-0.66491425,
pem=-3.6888795, similarity=4.3080454 ]
11: 6171483 PAF_Base_Rafiqui
    [winner: 6171483 PAF_Base_Rafiqui, pem_pos=0,
final_scores=-0.08659437, final_local_scores=-0.0027553886,
pem=0.0, similarity=4.007487]
    gt: 6171483 PAF_Base_Rafiqui, pem_pos=0,
final_scores=-0.08659437, final_local_scores=-0.0027553886,
pem=0.0, similarity=4.007487 ]
1: 23235 Pakistan
    [pem_pos=0, final_scores=0.55337936,
final_local_scores=0.5670044, pem=-0.21944217,
similarity=22.203615]
2: 18896748 Shorkot_Tehsil
    [pem_pos=0, final_scores=0.049220487,
final_local_scores=0.17480442, pem=0.0, similarity=9.306226]
3: 1444386 Sultan_Bahu
    [pem_pos=0, final_scores=0.100984946,
final_local_scores=0.1332049, pem=0.0, similarity=8.064813]
4: 6037917 Islam
    [pem_pos=0, final_scores=0.39850157,
final_local_scores=0.4319268, pem=-0.12106096,
similarity=17.637617]
5: 433843 Multan
    [pem_pos=0, final_scores=0.41741067,
final_local_scores=0.37551558, pem=0.0, similarity=15.295847]
6: 125315 Lahore
    [pem_pos=0, final_scores=0.47733492,

```


final_local_scores=0.4615535, pem=0.0, similarity=17.86339]
7: 30531541 Bhirr
[pem_pos=0, final_scores=0.04208468,
final_local_scores=0.08419122, pem=0.0, similarity=6.6021466]
8: 1100160 Chenab River
[pem_pos=0, final_scores=0.2575485,
final_local_scores=0.26529953, pem=0.0, similarity=12.00678]
9: 444123 Muharram
[pem_pos=0, final_scores=0.20810418,
final_local_scores=0.22330573, pem=0.0, similarity=10.753601]
1: 4760476 Shorkot
[pem_pos=0, final_scores=0.24401133,
final_local_scores=0.22059155, pem=0.0, similarity=10.672605]
2: 2792911 Jhang
[pem_pos=0, final_scores=0.44459277,
final_local_scores=0.44880253, pem=0.0, similarity=17.482876]
3: 4760476 Shorkot
[pem_pos=0, final_scores=0.32140917,
final_local_scores=0.30513692, pem=0.0, similarity=13.195607]
4: 113668 Shah Alam
[pem_pos=0, final_scores=0.26503402,
final_local_scores=0.33117345, pem=0.0, similarity=13.9725895]
5: 4760476 Shorkot
[pem_pos=0, final_scores=0.3068387,
final_local_scores=0.28922087, pem=0.0, similarity=12.72064]
6: 519237 Ranjit Singh
[pem_pos=0, final_scores=0.23553427,
final_local_scores=0.22202241, pem=0.0, similarity=10.715304]
7: 4760476 Shorkot
[pem_pos=0, final_scores=0.35559386,
final_local_scores=0.34247845, pem=0.0, similarity=14.309953]
8: 599998 Sayyid
[pem_pos=0, final_scores=0.16695933,
final_local_scores=0.15059441, pem=-0.05230848,
similarity=8.868213]
9: 12108 Greece
[pem_pos=1, final_scores=0.15583102,
final_local_scores=0.22223574, pem=-2.3247876,
similarity=23.364227]
10: 12108 Greece
[pem_pos=2, final_scores=0.11392777,
final_local_scores=0.18596971, pem=-2.5869315,
similarity=23.707556]
11: 30531541 Bhirr
[pem_pos=0, final_scores=0.14740793,
final_local_scores=0.19924086, pem=0.0, similarity=10.035458]
12: 30531541 Bhirr
[pem_pos=0, final_scores=0.18583499,
final_local_scores=0.2412166, pem=0.0, similarity=11.288097]
13: 4760476 Shorkot
[pem_pos=0, final_scores=0.353356,
final_local_scores=0.34003398, pem=0.0, similarity=14.237005]

14: 4760476 Shorkot
[pem_pos=0, final_scores=0.38552707,
final_local_scores=0.375176, pem=0.0, similarity=15.285714]
15: 4760476 Shorkot
[pem_pos=0, final_scores=0.32093924,
final_local_scores=0.3046236, pem=0.0, similarity=13.180289]
16: 11104 Fatah
[pem_pos=0, final_scores=0.1535586,
final_local_scores=0.22879583, pem=-0.718465,
similarity=14.824561]

Sample: b'Source_of_the_Nile'

_ -LRB- board_game -RRB- ` Source of the Nile ' is a 1 board game by Ross Maker and 2 David Wesely . It was released by in 1977 and re-released by 3 Avalon Hill in 1979 . Set in the 19th century , the object of the game is to explore the interior of [1 1 Africa 1] , make an important discovery , and report it back to [2 2 Europe 2] an 4 civilization . The 5 explorer s must outfit their 6 expedition s and choose a method of travel , be it 7 canoe , 8 camel or 9 foot . At the outset of the game the map of the interior of [1 Africa 1] is blank . As explorers enter each unexplored 10 hex , a card is drawn to determine the 11 terrain and events , if any , that befall the explorers . 12 Crayon s are used to draw terrain on the map as it is discovered . The explorers must confront 1 3 dangers such as 14 starvation , 15 disease , 16 river catar acts and hostile natives . The game is , in fact , highly lethal to explorers , but if a player loses an explorer he or she may create a new one who re-enters the game at a 17 coastal 18 port

1: 5334607 Africa

2: 9239 Europe

1: 3401 Board_game

[winner: 21402710 The_Game_(mind_game), pem_pos=2,
final_scores=-1.1336929, final_local_scores=-1.1265674,
pem=-1.0411923, similarity=-23.867138]

RECALL MISS: 3401 Board_game]

[winner: 4662671 American_Board_of_Medical_Specialties,
pem_pos=2, final_scores=-1.2635689,
final_local_scores=-1.222188, pem=-1.3367633,
similarity=-25.113285]

gt: 3401 Board_game, pem_pos=6, final_scores=-1.9777281,
final_local_scores=-2.153161, pem=-2.4430273,
similarity=-46.879345]

2: 5027624 David_Wesely

[winner: 5027624 David_Wesely, pem_pos=0,
final_scores=-0.040976748, final_local_scores=0.09463973,
pem=0.0, similarity=6.9139514]

gt: 5027624 David_Wesely, pem_pos=0,
final_scores=-0.040976748, final_local_scores=0.09463973,
pem=0.0, similarity=6.9139514]

3: 68069 Avalon_Hill

[winner: 68069 Avalon_Hill, pem_pos=0,
final_scores=-0.14060643, final_local_scores=-0.11337522,
pem=0.0, similarity=0.70637023|
gt: 68069 Avalon_Hill, pem_pos=0,
final_scores=-0.14060643, final_local_scores=-0.11337522,
pem=0.0, similarity=0.70637023]
4: 7299 Colonialism
[winner: 6259 Civilization_(video_game), pem_pos=0,
final_scores=-1.3069547, final_local_scores=-1.3702977,
pem=-0.59028786, similarity=-33.59263|
RECALL MISS: 7299 Colonialism]
5: 219878 Exploration
[winner: 4047060 The_Northwest_Explorer, pem_pos=1,
final_scores=-0.79663527, final_local_scores=-0.68168944,
pem=-0.4466232, similarity=-13.824453|
RECALL MISS: 219878 Exploration]
6: 219878 Exploration
[winner: 992161 Narváez_expedition, pem_pos=1,
final_scores=-1.4848133, final_local_scores=-1.544513,
pem=-1.6234654, similarity=-33.172977|
gt: 219878 Exploration, pem_pos=0,
final_scores=-1.8726274, final_local_scores=-2.0635052,
pem=-1.5692027, similarity=-48.95583]
7: 99960 Canoe
[winner: 4630959 Māori_migration_canoes, pem_pos=2,
final_scores=-1.1536324, final_local_scores=-1.1153326,
pem=-1.3227102, similarity=-22.000933|
gt: 99960 Canoe, pem_pos=0, final_scores=-1.4561148,
final_local_scores=-1.5837419, pem=-0.05534773,
similarity=-42.87132]
8: 6598 Camel
[winner: 22314808 Camel_(color), pem_pos=5,
final_scores=-1.2006098, final_local_scores=-1.1334395,
pem=-2.4334888, similarity=-16.500694|
gt: 6598 Camel, pem_pos=0, final_scores=-1.458665,
final_local_scores=-1.5863631, pem=-0.8722684,
similarity=-38.506996]
9: 46779 Human-powered_transport
[winner: 202485 Foot_(prosody), pem_pos=2,
final_scores=-1.2962793, final_local_scores=-1.278944,
pem=-1.0231506, similarity=-28.512476|
gt: 46779 Human-powered_transport, pem_pos=18,
final_scores=-2.4907715, final_local_scores=-2.6171412,
pem=-8.927978, similarity=-25.459269]
10: 1720758 Hex_map
[winner: 1341701 Hex_(Doctor_Who), pem_pos=4,
final_scores=-1.0965471, final_local_scores=-1.0591742,
pem=-2.204492, similarity=-15.529786|
RECALL MISS: 1720758 Hex_map]
11: 79239 Topography
[winner: 586839 Terrain, pem_pos=0,
final_scores=-1.3003386, final_local_scores=-1.3697345,

pem=-0.021533424, similarity=-36.668797|
gt: 79239 Topography, pem_pos=1,
final_scores=-2.4228237, final_local_scores=-2.629743,
pem=-5.072671, similarity=-46.80109]
12: 46415 Crayon
[winner: 9844043 Crayon_(band), pem_pos=1,
final_scores=-0.28329903, final_local_scores=-0.15750733,
pem=-0.5132917, similarity=2.18074|
gt: 46415 Crayon, pem_pos=0, final_scores=-0.28819358,
final_local_scores=-0.24479005, pem=-0.11839899,
similarity=-2.5714386]
13: 24462958 Risk
[winner: 15706646 Dangers, pem_pos=0,
final_scores=-0.7782858, final_local_scores=-0.6426005,
pem=0.0, similarity=-15.086767|
gt: 24462958 Risk, pem_pos=4, final_scores=-1.961562,
final_local_scores=-2.0757627, pem=-3.6463199,
similarity=-38.025936]
14: 215891 Starvation
[winner: 14045789 Hunger_Plan, pem_pos=1,
final_scores=-1.3369612, final_local_scores=-1.3412092,
pem=-0.6927946, similarity=-32.16712|
gt: 215891 Starvation, pem_pos=0,
final_scores=-1.4272571, final_local_scores=-1.4770412,
pem=-0.17869559, similarity=-39.01637]
15: 8072 Disease
[winner: 8072 Disease, pem_pos=0, final_scores=-1.161376,
final_local_scores=-1.2146454, pem=-0.15189487,
similarity=-31.331697|
gt: 8072 Disease, pem_pos=0, final_scores=-1.161376,
final_local_scores=-1.2146454, pem=-0.15189487,
similarity=-31.331697]
16: 2465387 Cataracts_of_the_Nile
[winner: 88931 Cataract, pem_pos=0,
final_scores=-1.2245635, final_local_scores=-1.2416666,
pem=-0.09956762, similarity=-32.42263|
gt: 2465387 Cataracts_of_the_Nile, pem_pos=3,
final_scores=-2.057176, final_local_scores=-2.1938217,
pem=-4.5559797, similarity=-36.602173]
[winner: 18842395 River, pem_pos=0,
final_scores=-1.3847909, final_local_scores=-1.4540737,
pem=-0.12563813, similarity=-38.619507|
RECALL MISS: 2465387 Cataracts_of_the_Nile]
17: 5236 Coast
[winner: 23716943 Cilentan_Coast, pem_pos=7,
final_scores=-1.2044687, final_local_scores=-1.1341581,
pem=-1.9440129, similarity=-19.183983|
gt: 5236 Coast, pem_pos=0, final_scores=-1.2772499,
final_local_scores=-1.3888304, pem=-0.09349413,
similarity=-36.847317]
18: 97169 Port
[winner: 11397167 Parrsboro_Harbour, pem_pos=17,

```

final_scores=-1.1897143, final_local_scores=-1.1011546,
pem=-3.4518807, similarity=-9.99907|
    gt: 97169 Port, pem_pos=0, final_scores=-1.4152217,
final_local_scores=-1.5141594, pem=-0.7182568,
similarity=-37.189835 ]
1: 5334607 Africa
    [pem_pos=0, final_scores=0.66511875,
final_local_scores=0.66546774, pem=-0.13197383,
similarity=24.666292]
2: 9239 Europe
    [pem_pos=0, final_scores=0.544367,
final_local_scores=0.5570416, pem=-0.086338006,
similarity=21.182465]
1: 5334607 Africa
    [pem_pos=0, final_scores=0.654579,
final_local_scores=0.6539546, pem=-0.13197383,
similarity=24.322718]

```

Sample: b'Jimmy_Rooney'

James `` Jimmy '' [1 Rooney 1] -LRB- born 10 December 1945 in [1 1 Dundee 1] , [2 Scotland 2] -RRB- is a former player . He was a member of the squad in [3 West Germany 3] . [4 Rooney 4] was a left sided midfielder who played first team 1 football with [2 2 Lochee United 2] , [3 3 Peterborough United 3] and 2 Montrose F.C. in the [5 United Kingdom 5] , before emigrating to in the late 1960s . He was an instant success there with [6 Ukrainian 6] Essendon Lions before transferring to [7 Sydney 7] to join the club in 1970 , and later playing at , in the . In 1977 , with APIA not joining the newly formed [4 4 National Soccer League 4] , [8 Rooney 8] joined for two seasons , before moving back to Victoria where he played with and He made his [9 Australian 9] debut against the [10 3 England 10] in 1971 , going on play 99 times for his country in a career spanning a decade . He was Australia 's first choice midfielder in the campaign , earning himself many admirers for his hard tackling and ball distribution . The Jimmy Rooney Medal , awarded from 1992 onwards to the man of the match in the 4 Victorian Premier League Grand Final , was named in his honour . He has recently settled to a managerial role of Victorian senior team , Eastern Lions , and also hosts a junior training program . He is also currently the head coach of the first eleven soccer team at 5 St Kevin 's College , [11 Melbourne 11]

```

1: 8828 Dundee
2: 5962688 Lochee_United_F.C.
3: 451132 Peterborough_United_F.C.
4: 529785 National_Soccer_League
1: 10568 Association_football
    [winner: 10568 Association_football, pem_pos=0,
final_scores=-1.5691899, final_local_scores=-1.7292061,
pem=-0.33534196, similarity=-45.6896]
    gt: 10568 Association_football, pem_pos=0,

```

final_scores=-1.5691899, final_local_scores=-1.7292061,
pem=-0.33534196, similarity=-45.6896]

2: 880408 Montrose_F.C.
[winner: 880408 Montrose_F.C., pem_pos=1,
final_scores=-0.64504313, final_local_scores=-0.73237085,
pem=-2.1839786, similarity=-5.888879]
gt: 880408 Montrose_F.C., pem_pos=1,
final_scores=-0.64504313, final_local_scores=-0.73237085,
pem=-2.1839786, similarity=-5.888879]

[winner: 588062 Bournemouth_F.C., pem_pos=0,
final_scores=-1.2727269, final_local_scores=-1.3792968,
pem=0.0, similarity=-37.07125]
RECALL MISS: 880408 Montrose_F.C.]

3: 46136 The_Football_Association
[winner: 9904 England_national_football_team, pem_pos=1,
final_scores=0.1839187, final_local_scores=0.13166237,
pem=-2.776464, similarity=23.117624]
gt: 46136 The_Football_Association, pem_pos=7,
final_scores=-0.47571027, final_local_scores=-0.56356174,
pem=-4.7043276, similarity=12.854773]

4: 2161864 Victorian_Premier_League
[winner: 2161864 Victorian_Premier_League, pem_pos=0,
final_scores=-0.069132194, final_local_scores=-0.08625501,
pem=0.0, similarity=1.5156915]
gt: 2161864 Victorian_Premier_League, pem_pos=0,
final_scores=-0.069132194, final_local_scores=-0.08625501,
pem=0.0, similarity=1.5156915]

5: 1677080 St_Kevin's_College,_Melbourne
[winner: 17306237 Melbourne, pem_pos=0,
final_scores=0.24340715, final_local_scores=0.23165792,
pem=-0.16841523, similarity=11.9187155]
RECALL MISS: 1677080 St_Kevin's_College,_Melbourne]
[winner: 1651417 Kevin_of_Glenda_lough, pem_pos=0,
final_scores=-0.032560885, final_local_scores=0.07147315,
pem=0.0, similarity=6.2226143]
RECALL MISS: 1677080 St_Kevin's_College,_Melbourne]
[winner: 5689 College, pem_pos=0,
final_scores=-0.36559376, final_local_scores=-0.3758654,
pem=-0.85246265, similarity=-2.4910362]
RECALL MISS: 1677080 St_Kevin's_College,_Melbourne]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.6352093, final_local_scores=-1.5984747,
pem=-4.6728287, similarity=-18.20039]
RECALL MISS: 1677080 St_Kevin's_College,_Melbourne]
[winner: 1415027 David_Berman_(musician), pem_pos=5,
final_scores=-2.5979416, final_local_scores=-2.7282522,
pem=-3.0012724, similarity=-61.00539]
RECALL MISS: 1677080 St_Kevin's_College,_Melbourne]

1: 8828 Dundee
[pem_pos=0, final_scores=0.15621762,
final_local_scores=0.15587252, pem=-0.59776163,
similarity=11.991981]

2: 5962688 Lochee_United_F.C.
[pem_pos=0, final_scores=0.14432131,
final_local_scores=0.18686423, pem=0.0, similarity=9.666114]
3: 451132 Peterborough_United_F.C.
[pem_pos=0, final_scores=0.4385175,
final_local_scores=0.44020587, pem=0.0, similarity=17.226334]
4: 529785 National_Soccer_League
[pem_pos=0, final_scores=0.12641947,
final_local_scores=0.11688441, pem=0.0, similarity=7.5777764]
1: 6762719 Jimmy_Rooney
[pem_pos=0, final_scores=0.051028654,
final_local_scores=0.039337277, pem=0.0, similarity=5.2636156]
2: 26994 Scotland
[pem_pos=0, final_scores=0.26431513,
final_local_scores=0.2584804, pem=-0.22905664,
similarity=13.048929]
3: 250204 Germany_national_football_team
[pem_pos=1, final_scores=0.18894605,
final_local_scores=0.18738878, pem=-1.8682308,
similarity=19.841497]
4: 199445 Wayne_Rooney
[pem_pos=1, final_scores=0.44648004,
final_local_scores=0.46670914, pem=-0.5937984,
similarity=21.246412]
5: 31717 United_Kingdom
[pem_pos=0, final_scores=0.36020643,
final_local_scores=0.3773874, pem=-0.09933331,
similarity=15.891896]
6: 31750 Ukraine
[pem_pos=1, final_scores=0.160861,
final_local_scores=0.17844796, pem=-1.0458583,
similarity=15.102497]
7: 27862 Sydney
[pem_pos=0, final_scores=0.32755417,
final_local_scores=0.33143106, pem=-0.1443489,
similarity=14.765268]
8: 199445 Wayne_Rooney
[pem_pos=1, final_scores=0.38034505,
final_local_scores=0.3944667, pem=-0.5937984,
similarity=19.090553]
9: 4689264 Australia
[pem_pos=0, final_scores=0.19626506,
final_local_scores=0.2050727, pem=-0.491009,
similarity=12.879675]
10: 9904 England_national_football_team
[pem_pos=1, final_scores=0.1839187,
final_local_scores=0.13166237, pem=-2.776464,
similarity=23.117624]
11: 17306237 Melbourne
[pem_pos=0, final_scores=0.24340715,
final_local_scores=0.23165792, pem=-0.16841523,
similarity=11.9187155]

Sample: b'Enticho'

-LRB- worda -RRB- [1 Enticho 1] -LRB- also known as Ahiferom -RRB- is one of 1 worda s in the [1 1 Tigray Region 1] of 2 Ethiopia . Part of the [2 2 Mehakelegnaw Zone 2] , [2 Enticho 2] is bordered on the south by [3 3 Werie Lehe 3] , on the southwest by [3 3 Adwa 3] , on the west by [4 4 Mereb L ehe 4] , on the north by [5 5 Eritrea 5] , and on the east by the 4 Misraqawi -LRB- Eastern -RRB- Zone . The administrative center of this worda is [6 6 Enticho 6] ; other towns in [4 Enticho 4] include , and . Ancient monastery of [7 7 Debre Da mo 7] is also located in [5 Enticho 5] . Prominent high points in this worda include , which R.S. Whiteway identifies as the location of the 5 Battle of [6 Baçente 6] , where [8 8 Cristóvão da Gama 8] in 1542 enjoyed his first victory over his [7 Moslem 7] foe . Based on the 2007 national census conducted by the [8 6 Central Statistical Age ncy of Ethiopia 8] -LRB- CSA -RRB- , this worda has a total population of 173,700 , an increase of 32.43 % over the 1994 census , of whom 84,014 are men and 89,686 women ; 23,421 or 13.48 % are urban inhabitants . With an area of 2,367.84 square kilometers , [9 Enticho 9] has a population density of 73.36 , which is greater than the Zone average of 56.29 persons per square kilometer . A total of 38,934 households were counted in this worda , resulting in an average of 4.46 persons to a household , and 37,483 housing units . The majority of the inhabitants said they practiced , with 97.76 % reporting that as their religion , while 2.22 % of the population were [10 7 Muslim 10] . The 1994 national census reported a total population for this worda of 105,814 , of whom 51,871 were men and 53,943 were women ; 12,375 or 11.7 % of its population were urban dwellers . The two largest ethnic groups reported in Abergele were the [9 9 Tigrayan 9] -LRB- 99.5 % -RRB- , and the [10 10 Amhara 10] -LRB- 0.3 % -RRB- ; all other ethnic groups made up 0.2 % of the population . 8 Tigrinya was spoken as a first language by 99.44 % , and 0.44 % spoke [11 11 Amharic 11] ; the remaining 0.12 % spoke all other primary languages reported . 99.32 % of the population practiced , and 0.41 % were [11 9 Muslim 11] . Concerning 10 education , 13.64 % of the population were considered literate , which is less than the Zone average of 15.71 % ; 21.31 % of children aged 7-12 were in primary school ; 1.71 % of the children aged 13-14 were in junior secondary school ; 2.62 % of the inhabitants aged 15-18 were in senior secondary school . Concerning 11 sanitary conditions , about 90 % of the urban houses and 29 % of all houses had access to safe drinking water at the time of the census ; 29 % of the urban and 6 % of the total had toilet facilities . A sample enumeration performed by the CSA in 2001 interviewed 34,455 farmers in this worda , who held an average of 0.67 hectares of land . Of the 23,196 hectares of private land surveyed , 85.86 % was in cultivation , 0.85 % pasture , 8.32

% fallow , 0.62 % **12 woodland** , and 4.35 % was devoted to other uses . For the land under cultivation in this woreda , 72.68 % was planted in cereals , 7.76 % in pulses , and 0.48 % in oilseeds ; the amount in vegetables is missing . The area planted in fruit trees was 42 hectares , while the area planted in

- 1: 357350 Tigray_Region
- 2: 4973073 Mehakelegnaw_Zone
- 3: 7368351 Werie_Lehe
- 4: 7368252 Mereb_Lehe
- 5: 17238590 Eritrea
- 6: 11702297 Enticho
- 7: 2189735 Debre_Damo
- 8: 3808781 Cristóvão_da_Gama
- 9: 4663597 Tigray-Tigrinya_people
- 10: 355173 Amhara_people
- 11: 95135 Amharic_language

- 1: 16847257 Districts_of_Ethiopia

[winner: 16847257 Districts_of_Ethiopia, pem_pos=0, final_scores=-0.67231846, final_local_scores=-0.7182324, pem=-0.006476348, similarity=-17.308556|

gt: 16847257 Districts_of_Ethiopia, pem_pos=0, final_scores=-0.67231846, final_local_scores=-0.7182324, pem=-0.006476348, similarity=-17.308556]

- 2: 187749 Ethiopia

[winner: 9686 Ethiopian_cuisine, pem_pos=0, final_scores=-0.7629858, final_local_scores=-0.7591502, pem=0.0, similarity=-18.564844|

RECALL MISS: 187749 Ethiopia]

- 3: 7368366 Adwa_(woreda)

[winner: 848823 Adwa, pem_pos=0, final_scores=0.21765386, final_local_scores=0.20046315, pem=0.0, similarity=10.071933|

gt: 7368366 Adwa_(woreda), pem_pos=4, final_scores=-0.47752428, final_local_scores=-0.49638855, pem=-2.5423176, similarity=3.1020062]

- 4: 4972944 Misraqawi_Zone

[winner: 4972944 Misraqawi_Zone, pem_pos=0, final_scores=-0.034747794, final_local_scores=-0.026236564, pem=0.0, similarity=3.3067615|

gt: 4972944 Misraqawi_Zone, pem_pos=0, final_scores=-0.034747794, final_local_scores=-0.026236564, pem=0.0, similarity=3.3067615]

[winner: 605222 Zone_(band), pem_pos=0, final_scores=-0.68937856, final_local_scores=-0.5921152, pem=-1.6527174, similarity=-4.5924525|

RECALL MISS: 4972944 Misraqawi_Zone]

[winner: 1404486 Eastern_Province_(Kenya), pem_pos=15, final_scores=-1.5236706, final_local_scores=-1.6031189, pem=-3.994853, similarity=-22.025917|

RECALL MISS: 4972944 Misraqawi_Zone]

- 5: 3277663 Battle_of_Baçente

[winner: 3277663 Battle_of_Baçente, pem_pos=0,

final_scores=0.09865533, final_local_scores=0.19118917,
 pem=0.0, similarity=9.795179|
 gt: 3277663 Battle_of_Baçente, pem_pos=0,
 final_scores=0.09865533, final_local_scores=0.19118917,
 pem=0.0, similarity=9.795179]
 [winner: 4181 Battle, pem_pos=0, final_scores=-0.4238438,
 final_local_scores=-0.43626937, pem=-1.4437126,
 similarity=-1.0783021|
 RECALL MISS: 3277663 Battle_of_Baçente]
 6: 4242861 Central_Statistical_Agency_(Ethiopia)
 [winner: 4242861 Central_Statistical_Agency_(Ethiopia),
 pem_pos=0, final_scores=0.06789629,
 final_local_scores=0.0643201, pem=0.0, similarity=6.009153|
 gt: 4242861 Central_Statistical_Agency_(Ethiopia),
 pem_pos=0, final_scores=0.06789629,
 final_local_scores=0.0643201, pem=0.0, similarity=6.009153]
 7: 1274248 Islam_in_Ethiopia
 [winner: 19541 Muslim, pem_pos=0,
 final_scores=0.075418696, final_local_scores=0.06657076,
 pem=-0.22736222, similarity=7.3127484|
 RECALL MISS: 1274248 Islam_in_Ethiopia]
 8: 95129 Tigrinya_language
 [winner: 95129 Tigrinya_language, pem_pos=0,
 final_scores=-0.003943205, final_local_scores=0.005798638,
 pem=0.0, similarity=4.262756|
 gt: 95129 Tigrinya_language, pem_pos=0,
 final_scores=-0.003943205, final_local_scores=0.005798638,
 pem=0.0, similarity=4.262756]
 9: 1274248 Islam_in_Ethiopia
 [winner: 19541 Muslim, pem_pos=0,
 final_scores=0.068329886, final_local_scores=0.058827296,
 pem=-0.22736222, similarity=7.081668|
 RECALL MISS: 1274248 Islam_in_Ethiopia]
 10: 11378216 Education_in_Ethiopia
 [winner: 18823453
 Angelo_State_University_College_of_Education, pem_pos=3,
 final_scores=-1.645139, final_local_scores=-1.631043,
 pem=-3.356857, similarity=-26.328747|
 RECALL MISS: 11378216 Education_in_Ethiopia]
 11: 11844406 Water_supply_and_sanitation_in_Ethiopia
 [winner: 16496127 Conditions_(magazine), pem_pos=1,
 final_scores=-1.0444366, final_local_scores=-0.95581764,
 pem=-0.37142345, similarity=-22.413933|
 RECALL MISS: 11844406
 Water_supply_and_sanitation_in_Ethiopia]
 [winner: 10831865 Swimming_pool_sanitation, pem_pos=1,
 final_scores=-1.6037259, final_local_scores=-1.5569487,
 pem=-2.5758784, similarity=-28.364717|
 RECALL MISS: 11844406
 Water_supply_and_sanitation_in_Ethiopia]
 12: 16012762 Forestry_in_Ethiopia
 [winner: 10294380

Woodland_Township,_Fulton_County,_Illinois, pem_pos=7,
final_scores=-1.1512042, final_local_scores=-1.0650526,
pem=-2.9717824, similarity=-11.532567]

RECALL MISS: 16012762 Forestry_in_Ethiopia]

- 1: 357350 Tigray_Region
[pem_pos=0, final_scores=0.33140498,
final_local_scores=0.315188, pem=0.0, similarity=13.495551]
- 2: 4973073 Mehakelegnaw_Zone
[pem_pos=0, final_scores=0.28198498,
final_local_scores=0.30612332, pem=0.0, similarity=13.225043]
- 3: 7368351 Werie_Lehe
[pem_pos=0, final_scores=0.10485844,
final_local_scores=0.14634743, pem=0.0, similarity=8.457012]
- 4: 7368252 Mereb_Lehe
[pem_pos=0, final_scores=0.045618042,
final_local_scores=0.07151736, pem=0.0, similarity=6.2239337]
- 5: 17238590 Eritrea
[pem_pos=0, final_scores=0.27204442,
final_local_scores=0.27286005, pem=-0.3232545,
similarity=13.990309]
- 6: 11702297 Enticho
[pem_pos=0, final_scores=0.24697055,
final_local_scores=0.25869903, pem=0.0, similarity=11.809808]
- 7: 2189735 Debre_Damo
[pem_pos=0, final_scores=0.1515501,
final_local_scores=0.22943375, pem=0.0, similarity=10.936474]
- 8: 3808781 Cristóvão_da_Gama
[pem_pos=0, final_scores=0.0954601,
final_local_scores=0.13996741, pem=0.0, similarity=8.26662]
- 9: 4663597 Tigray-Tigrinya_people
[pem_pos=0, final_scores=0.1943485,
final_local_scores=0.18214947, pem=0.0, similarity=9.525416]
- 10: 355173 Amhara_people
[pem_pos=0, final_scores=0.053653434,
final_local_scores=0.057771355, pem=0.0, similarity=5.8137255]
- 11: 95135 Amharic_language
[pem_pos=0, final_scores=0.30218464,
final_local_scores=0.31845996, pem=0.0, similarity=13.593193]
- 1: 11702297 Enticho
[pem_pos=0, final_scores=0.07574265,
final_local_scores=0.07165857, pem=0.0, similarity=6.2281475]
- 2: 11702297 Enticho
[pem_pos=0, final_scores=0.15783577,
final_local_scores=0.16133282, pem=0.0, similarity=8.904206]
- 3: 848823 Adwa
[pem_pos=0, final_scores=0.21765386,
final_local_scores=0.20046315, pem=0.0, similarity=10.071933]
- 4: 11702297 Enticho
[pem_pos=0, final_scores=0.34443218,
final_local_scores=0.365161, pem=0.0, similarity=14.986845]
- 5: 11702297 Enticho
[pem_pos=0, final_scores=0.27239925,

final_local_scores=0.28647602, pem=0.0, similarity=12.638729]
6: 3277663 Battle_of_Bağente
[pem_pos=0, final_scores=0.09865533,
final_local_scores=0.19118917, pem=0.0, similarity=9.795179]
7: 6037917 İslam
[pem_pos=1, final_scores=0.43967122,
final_local_scores=0.47669643, pem=-1.2150909,
similarity=24.923138]
8: 4242861 Central_Statistical_Agency_(Ethiopia)
[pem_pos=0, final_scores=0.06789629,
final_local_scores=0.0643201, pem=0.0, similarity=6.009153]
9: 11702297 Enticho
[pem_pos=0, final_scores=0.069685474,
final_local_scores=0.065042034, pem=0.0, similarity=6.030697]
10: 19541 Muslim
[pem_pos=0, final_scores=0.075418696,
final_local_scores=0.06657076, pem=-0.22736222,
similarity=7.3127484]
11: 19541 Muslim
[pem_pos=0, final_scores=0.068329886,
final_local_scores=0.058827296, pem=-0.22736222,
similarity=7.081668]

Sample: b'Fresia'

, _ Chile Fresia is a and 1 commune in [1 1 Llanquihue Province 1] , [2 2 Los Lagos Region 2] , [3 3 Chile 3] . According to the 2002 2 census of the [1 3 National Statistics Institute 1] , Fresia spans an area of and has 12,804 inhabitants -LRB- 6,580 men and 6,224 women -RRB- . Of these , 6,144 -LRB- 48 % -RRB- lived in 4 urban area s and 6,660 -LRB- 52 % -RRB- in 5 rural area s . The population fell by 1.6 % -LRB- 209 persons -RRB- between the 1992 and 2002 censuses . As a commune , Fresia is a third-level administered by a 6 municipal council , headed by an 7 alcalde who is directly elected every four years . The 2008-2012 alcalde is Bernardo Espinoza Villalobos -LRB- 8 PS -RRB- . Within the 9 electoral divisions of [2 Chile 2] , Fresia is represented in the 10 Chamber of Deputies by Fidel Espinoza -LRB- PS -RRB- and Carlos Recondo -LRB- 11 UDI -RRB- as part of the 56th electoral district , -LRB- together with 12 Puyehue , 13 Río Negro , 14 Purranque , 15 Puerto Octay , 16 Frutillar , 17 Llanquihue , [4 4 Puerto Varas 4] and 18 Los Muermos -RRB- . The commune is represented in the 19 Senate by Camilo Escalona Medina -LRB- PS -RRB- and Carlos Kuschel Silva -LRB- 20 RN

1: 2201189 Llanquihue_Province

2: 585718 Los_Lagos_Region

3: 5489 Chile

4: 2242900 Puerto_Varas

1: 3060770 Communes_of_Chile

[winner: 8277118 Communes_of_Algeria, pem_pos=3,
final_scores=-1.0338826, final_local_scores=-0.98538953,
pem=-3.2370336, similarity=-7.712785]

gt: 3060770 Communes_of_Chile, pem_pos=8,
final_scores=-1.6184554, final_local_scores=-1.8096178,
pem=-3.809452, similarity=-29.196499]
2: 6889 Census
[winner: 6889 Census, pem_pos=0, final_scores=-1.3592983,
final_local_scores=-1.480242, pem=-0.05546598,
similarity=-39.782032]
gt: 6889 Census, pem_pos=0, final_scores=-1.3592983,
final_local_scores=-1.480242, pem=-0.05546598,
similarity=-39.782032]
3: 29015808 National_Statistics_Institute_(Chile)
[winner: 4070894
Instituto_Nacional_de_Estadística_y_Geografía, pem_pos=2,
final_scores=0.18949084, final_local_scores=0.20814133,
pem=-0.67166406, similarity=13.953679]
gt: 29015808 National_Statistics_Institute_(Chile),
pem_pos=0, final_scores=0.0908014,
final_local_scores=0.12011191, pem=-0.053524505,
similarity=7.9651666]
4: 764593 Urban_area
[winner: 1269662 Urban_areas_in_Sweden, pem_pos=2,
final_scores=-1.0060916, final_local_scores=-0.9464961,
pem=-1.760042, similarity=-14.584237]
gt: 764593 Urban_area, pem_pos=0,
final_scores=-1.2881488, final_local_scores=-1.3711329,
pem=-0.15587038, similarity=-35.979977]
5: 212614 Rural_area
[winner: 10420362 Area_(journal), pem_pos=6,
final_scores=-1.2045774, final_local_scores=-1.0803704,
pem=-2.3005102, similarity=-15.640167]
RECALL MISS: 212614 Rural_area]
[winner: 3909779 Penrith_Rural_District, pem_pos=1,
final_scores=-1.3435714, final_local_scores=-1.3282857,
pem=-1.6043973, similarity=-26.82402]
gt: 212614 Rural_area, pem_pos=0,
final_scores=-1.6374035, final_local_scores=-1.7677703,
pem=0.0, similarity=-48.66408]
6: 3532493 Municipal_council
[winner: 19038 Municipality, pem_pos=0,
final_scores=-1.049777, final_local_scores=-1.1369112,
pem=-0.19613224, similarity=-28.771387]
RECALL MISS: 3532493 Municipal_council]
[winner: 12666225 Executive_Council_of_Lower_Canada,
pem_pos=5, final_scores=-1.433614,
final_local_scores=-1.4463019, pem=-3.009209,
similarity=-22.706276]
RECALL MISS: 3532493 Municipal_council]
7: 597450 Alcalde
[winner: 597450 Alcalde, pem_pos=0,
final_scores=-1.2002343, final_local_scores=-1.3071713,
pem=-0.06003342, similarity=-34.592415]
gt: 597450 Alcalde, pem_pos=0, final_scores=-1.2002343,

final_local_scores=-1.3071713, pem=-0.06003342,
 similarity=-34.592415]
 8: 405910 Socialist_Party_of_Chile
 [winner: 39565408 Socialist_Party_(France), pem_pos=1,
 final_scores=-0.5861349, final_local_scores=-0.5757023,
 pem=-1.8423814, similarity=-3.0712364]
 gt: 405910 Socialist_Party_of_Chile, pem_pos=4,
 final_scores=-0.70520425, final_local_scores=-0.7321522,
 pem=-2.926659, similarity=-1.8435438]
 9: 2683116 Electoral_divisions_of_Chile
 [winner: 5489 Chile, pem_pos=0, final_scores=0.6505557,
 final_local_scores=0.6649039, pem=-0.27404457,
 similarity=25.422068]
 RECALL MISS: 2683116 Electoral_divisions_of_Chile]
 [winner: 31703793 Municipal_Units_of_Tirana, pem_pos=15,
 final_scores=-1.1459138, final_local_scores=-1.0252472,
 pem=-4.145923, similarity=-3.9595335]
 RECALL MISS: 2683116 Electoral_divisions_of_Chile]
 [winner: 948687 Roman_Catholic_Diocese_of_Mainz,
 pem_pos=7, final_scores=-1.4542828,
 final_local_scores=-1.4737473, pem=-2.0709488,
 similarity=-28.6277]
 RECALL MISS: 2683116 Electoral_divisions_of_Chile]
 10: 716259 Chamber_of_Deputies_of_Chile
 [winner: 411807 Chamber_of_Deputies_(Mexico), pem_pos=0,
 final_scores=-0.020051032, final_local_scores=0.013050824,
 pem=-1.5470045, similarity=12.892027]
 gt: 716259 Chamber_of_Deputies_of_Chile, pem_pos=2,
 final_scores=-0.09408875, final_local_scores=-0.062357336,
 pem=-1.9696342, similarity=12.940022]
 11: 526689 Independent_Democratic_Union
 [winner: 526689 Independent_Democratic_Union, pem_pos=0,
 final_scores=-0.14569265, final_local_scores=-0.132325,
 pem=-0.39522702, similarity=2.2901769]
 gt: 526689 Independent_Democratic_Union, pem_pos=0,
 final_scores=-0.14569265, final_local_scores=-0.132325,
 pem=-0.39522702, similarity=2.2901769]
 12: 11988863 Puyehue,_Chile
 [winner: 22559991 Puyehue-Cordón_Caulle, pem_pos=0,
 final_scores=-0.11410871, final_local_scores=-0.08008455,
 pem=-0.41259703, similarity=3.9435973]
 gt: 11988863 Puyehue,_Chile, pem_pos=2,
 final_scores=-0.1423038, final_local_scores=-0.16409476,
 pem=-1.0219333, similarity=4.75023]
 13: 2476583 Río_Negro,_Chile
 [winner: 315251 Río_Negro_Province, pem_pos=0,
 final_scores=0.018871948, final_local_scores=0.011553854,
 pem=-1.0573694, similarity=10.184642]
 gt: 2476583 Río_Negro,_Chile, pem_pos=6,
 final_scores=-0.3265372, final_local_scores=-0.3796567,
 pem=-3.2853656, similarity=10.626328]
 14: 7404196 Purranque

```
[winner: 7404196 Purranque, pem_pos=0,
final_scores=-0.17251946, final_local_scores=-0.17523368,
pem=0.0, similarity=-1.1396095|
gt: 7404196 Purranque, pem_pos=0,
final_scores=-0.17251946, final_local_scores=-0.17523368,
pem=0.0, similarity=-1.1396095 ]
15: 10158951 Puerto_Octay
[winner: 10158951 Puerto_Octay, pem_pos=0,
final_scores=0.00490357, final_local_scores=-0.013609797,
pem=0.0, similarity=3.6835694|
gt: 10158951 Puerto_Octay, pem_pos=0,
final_scores=0.00490357, final_local_scores=-0.013609797,
pem=0.0, similarity=3.6835694 ]
16: 1949922 Frutillar
[winner: 1949922 Frutillar, pem_pos=0,
final_scores=-0.06815015, final_local_scores=-0.06365183,
pem=0.0, similarity=2.1902153|
gt: 1949922 Frutillar, pem_pos=0,
final_scores=-0.06815015, final_local_scores=-0.06365183,
pem=0.0, similarity=2.1902153 ]
17: 20086376 Llanquihue,_Chile
[winner: 2201189 Llanquihue_Province, pem_pos=0,
final_scores=-0.113040626, final_local_scores=-0.15998596,
pem=-0.47215647, similarity=1.8830733|
gt: 20086376 Llanquihue,_Chile, pem_pos=2,
final_scores=-0.27338254, final_local_scores=-0.31253618,
pem=-0.6805195, similarity=-1.536219 ]
18: 22506536 Los_Muermos
[winner: 22506536 Los_Muermos, pem_pos=0,
final_scores=0.0028518587, final_local_scores=0.013544351,
pem=0.0, similarity=4.493903|
gt: 22506536 Los_Muermos, pem_pos=0,
final_scores=0.0028518587, final_local_scores=0.013544351,
pem=0.0, similarity=4.493903 ]
19: 683411 Senate_of_Chile
[winner: 16685964 Roman_Senate, pem_pos=2,
final_scores=-0.5590683, final_local_scores=-0.4912034,
pem=-2.525485, similarity=3.165204|
gt: 683411 Senate_of_Chile, pem_pos=9,
final_scores=-0.6694031, final_local_scores=-0.7033989,
pem=-3.724167, similarity=3.351487 ]
20: 1898078 National_Renewal_(Chile)
[winner: 1898078 National_Renewal_(Chile), pem_pos=2,
final_scores=-0.5288023, final_local_scores=-0.5414278,
pem=-1.8687539, similarity=-1.9049985|
gt: 1898078 National_Renewal_(Chile), pem_pos=2,
final_scores=-0.5288023, final_local_scores=-0.5414278,
pem=-1.8687539, similarity=-1.9049985 ]
1: 2201189 Llanquihue_Province
[pem_pos=0, final_scores=0.17679016,
final_local_scores=0.17752224, pem=0.0, similarity=9.387331]
2: 585718 Los_Lagos_Region
```

[pem_pos=0, final_scores=0.3079282,
final_local_scores=0.32068193, pem=0.0, similarity=13.659501]
3: 5489 Chile
[pem_pos=0, final_scores=0.4626236,
final_local_scores=0.4596166, pem=-0.27404457,
similarity=19.295885]
4: 2242900 Puerto_Varas
[pem_pos=0, final_scores=0.0698439,
final_local_scores=0.03349489, pem=0.0, similarity=5.0892673]
1: 4070894 Instituto_Nacional_de_Estadística_y_Geografía
[pem_pos=2, final_scores=0.18949084,
final_local_scores=0.20814133, pem=-0.67166406,
similarity=13.953679]
2: 5489 Chile
[pem_pos=0, final_scores=0.6505557,
final_local_scores=0.6649039, pem=-0.27404457,
similarity=25.422068]

Sample: b'NTFS-3G'

NTFS-3G is an 1 open source 2 cross-platform implementation of the 3 Microsoft Windows 4 NTFS file system with read-write support . [1 NTFS-3G 1] often uses the 5 FUSE , so it can run unmodified on many different 6 operating system s . It is runnable on [1 1 Linux 1] , [2 2 FreeBSD 2] , [3 3 NetBSD 3] , [2 7 OpenSolaris 2] , [4 4 BeOS 4] , 8 QNX , , 9 Nucleus , 10 VxWorks , 11 Haiku , [5 5 MorphOS 5] , and . It is licensed under either the [6 6 GNU General Public License 6] or a proprietary license . It is a partial 12 fork of 13 ntfsprogs and is under active maintenance and development . [3 NTFS-3G 3] was introduced by one of the senior Linux NTFS developers , Szabolcs Szakacsits -LRB- -RRB- , in July 2006 . The first stable version was released on 2007-02-21 as version 1.0 . The developers of [4 NTFS-3G 4] later formed a company , 14 Tuxera Inc. . , to further develop the code . [5 NTFS-3G 5] is now the free `` community edition '' , while Tuxera NTFS is the proprietary version . [6 NTFS-3G 6] supports all operations for writing files : files of any size can be created , modified , renamed , moved , or deleted on NTFS partitions . is supported , but there is no support for 15 encryption . Support to modify 16 access control list s and s is available . NTFS partitions are 17 mounted using the 18 Filesystem in Userspace -LRB- FUSE -RRB- interface . According to its 19 man page , [7 NTFS-3G 7] supports and . NTFS-3G supports partial NTFS 20 journaling , so if an unexpected computer failure leaves the file system in an inconsistent state , the volume can be repaired . As of 2009 , a volume having an unclean journal file is recovered and mounted by default . The `norecover` mount option can be used to disable this behavior . Benchmarks show that the driver 's performance via 21 FUSE

1: 6097297 Linux
2: 7580554 FreeBSD


```
3: 28901365 NetBSD
4: 4801 BeOS
5: 263521 MorphOS
6: 18938683 GNU_General_Public_License
1: 18938758 Open_source
    [winner: 18938758 Open_source, pem_pos=0,
final_scores=-0.7645392, final_local_scores=-0.91021794,
pem=-0.16215573, similarity=-22.191177|
    gt: 18938758 Open_source, pem_pos=0,
final_scores=-0.7645392, final_local_scores=-0.91021794,
pem=-0.16215573, similarity=-22.191177 ]
2: 81193 Cross-platform
    [winner: 81193 Cross-platform, pem_pos=0,
final_scores=-0.94461745, final_local_scores=-1.0912294,
pem=-0.0029796006, similarity=-28.458544|
    gt: 81193 Cross-platform, pem_pos=0,
final_scores=-0.94461745, final_local_scores=-1.0912294,
pem=-0.0029796006, similarity=-28.458544 ]
3: 18890 Microsoft_Windows
    [winner: 18890 Microsoft_Windows, pem_pos=0,
final_scores=-0.7422347, final_local_scores=-0.8894409,
pem=0.0, similarity=-22.452978|
    gt: 18890 Microsoft_Windows, pem_pos=0,
final_scores=-0.7422347, final_local_scores=-0.8894409,
pem=0.0, similarity=-22.452978 ]
4: 39184 NTFS
    [winner: 39184 NTFS, pem_pos=0, final_scores=-0.2593125,
final_local_scores=-0.33281153, pem=0.0,
similarity=-5.8420477|
    gt: 39184 NTFS, pem_pos=0, final_scores=-0.2593125,
final_local_scores=-0.33281153, pem=0.0,
similarity=-5.8420477 ]
5: 1219082 Filesystem_in_Userspace
    [winner: 1219082 Filesystem_in_Userspace, pem_pos=0,
final_scores=-0.10695346, final_local_scores=-0.14034463,
pem=-0.07386334, similarity=0.3032293|
    gt: 1219082 Filesystem_in_Userspace, pem_pos=0,
final_scores=-0.10695346, final_local_scores=-0.14034463,
pem=-0.07386334, similarity=0.3032293 ]
6: 22194 Operating_system
    [winner: 20825 Monolithic_kernel, pem_pos=2,
final_scores=-0.6927126, final_local_scores=-0.7732349,
pem=-0.6931472, similarity=-15.215716|
    gt: 22194 Operating_system, pem_pos=0,
final_scores=-0.811681, final_local_scores=-0.9697887,
pem=-0.030303603, similarity=-24.68592 ]
7: 46145 Solaris_(operating_system)
    [winner: 1178658 OpenSolaris, pem_pos=0,
final_scores=0.09911178, final_local_scores=0.048333764,
pem=-0.010893154, similarity=5.591327|
    gt: 46145 Solaris_(operating_system), pem_pos=4,
final_scores=-1.2920496, final_local_scores=-1.4841716,
```

pem=-8.495357, similarity=5.998144]
8: 65184 QNX
[winner: 65184 QNX, pem_pos=0, final_scores=-0.15430814,
final_local_scores=-0.2182577, pem=0.0, similarity=-2.423532|
gt: 65184 QNX, pem_pos=0, final_scores=-0.15430814,
final_local_scores=-0.2182577, pem=0.0, similarity=-2.423532]
9: 903339 Nucleus_RTOS
[winner: 6235 Cell_nucleus, pem_pos=0,
final_scores=-0.28929186, final_local_scores=-0.18415627,
pem=-0.46576488, similarity=1.1270242|
gt: 903339 Nucleus_RTOS, pem_pos=9,
final_scores=-0.48891604, final_local_scores=-0.52272844,
pem=-2.6396475, similarity=2.8452668]
10: 157105 VxWorks
[winner: 157105 VxWorks, pem_pos=0,
final_scores=-0.20311642, final_local_scores=-0.23645984,
pem=0.0, similarity=-2.9667203|
gt: 157105 VxWorks, pem_pos=0, final_scores=-0.20311642,
final_local_scores=-0.23645984, pem=0.0,
similarity=-2.9667203]
11: 554281 Haiku_(operating_system)
[winner: 13217 Haiku, pem_pos=0,
final_scores=-0.22783037, final_local_scores=-0.13188905,
pem=-0.3153545, similarity=1.8688273|
gt: 554281 Haiku_(operating_system), pem_pos=4,
final_scores=-0.27648282, final_local_scores=-0.35481632,
pem=-1.5840362, similarity=2.1155217]
12: 301506 Fork_(software_development)
[winner: 1040555 Garden_fork, pem_pos=6,
final_scores=-1.0193429, final_local_scores=-0.9411891,
pem=-1.7911415, similarity=-14.256741|
gt: 301506 Fork_(software_development), pem_pos=0,
final_scores=-1.2249707, final_local_scores=-1.3991442,
pem=-0.21426767, similarity=-36.498318]
13: 17577141 Ntfsprogs
[winner: 17577141 Ntfsprogs, pem_pos=0,
final_scores=-0.5838171, final_local_scores=-0.58722913,
pem=0.0, similarity=-13.434376|
gt: 17577141 Ntfsprogs, pem_pos=0,
final_scores=-0.5838171, final_local_scores=-0.58722913,
pem=0.0, similarity=-13.434376]
14: 38914995 Tuxera
[winner: 38914995 Tuxera, pem_pos=0,
final_scores=-0.8767217, final_local_scores=-0.96375597,
pem=0.0, similarity=-24.670687|
gt: 38914995 Tuxera, pem_pos=0, final_scores=-0.8767217,
final_local_scores=-0.96375597, pem=0.0,
similarity=-24.670687]
[winner: 23752548 Friends_of_Seagate_Inc., pem_pos=1,
final_scores=-1.1189739, final_local_scores=-1.09063,
pem=-1.0986123, similarity=-22.482439|
RECALL MISS: 38914995 Tuxera]

[winner: 4721653 Zyphe_language, pem_pos=12,
final_scores=-1.7162603, final_local_scores=-1.6305473,
pem=-4.4188404, similarity=-20.538729]
RECALL MISS: 38914995 Tuxera]

15: 628391 Encrypting_File_System
[winner: 10294 Encryption, pem_pos=0,
final_scores=-0.8040007, final_local_scores=-0.87556094,
pem=-0.05624306, similarity=-21.732914]
gt: 628391 Encrypting_File_System, pem_pos=7,
final_scores=-1.6599736, final_local_scores=-1.8129635,
pem=-5.9413853, similarity=-17.702553]

16: 61589 Access_control_list
[winner: 61589 Access_control_list, pem_pos=0,
final_scores=-0.7757208, final_local_scores=-0.8559433,
pem=0.0, similarity=-21.453342]
gt: 61589 Access_control_list, pem_pos=0,
final_scores=-0.7757208, final_local_scores=-0.8559433,
pem=0.0, similarity=-21.453342]

17: 5176987 Mount_(computing)
[winner: 5176987 Mount_(computing), pem_pos=0,
final_scores=-0.8318303, final_local_scores=-0.9082565,
pem=-0.39612266, similarity=-20.860294]
gt: 5176987 Mount_(computing), pem_pos=0,
final_scores=-0.8318303, final_local_scores=-0.9082565,
pem=-0.39612266, similarity=-20.860294]

18: 1219082 Filesystem_in_Userspace
[winner: 1010280 File_system, pem_pos=0,
final_scores=-0.22700657, final_local_scores=-0.31105977,
pem=0.0, similarity=-5.1929307]
RECALL MISS: 1219082 Filesystem_in_Userspace]
[winner: 93467 User_space, pem_pos=0,
final_scores=-0.33097392, final_local_scores=-0.40459624,
pem=0.0, similarity=-7.984246]
RECALL MISS: 1219082 Filesystem_in_Userspace]

19: 85332 Man_page
[winner: 85332 Man_page, pem_pos=0,
final_scores=-0.95393175, final_local_scores=-1.072304,
pem=0.0, similarity=-27.909975]
gt: 85332 Man_page, pem_pos=0, final_scores=-0.95393175,
final_local_scores=-1.072304, pem=0.0, similarity=-27.909975]

20: 23455133 Journaling_file_system
[winner: 23455133 Journaling_file_system, pem_pos=0,
final_scores=-0.9694529, final_local_scores=-1.0744469,
pem=0.0, similarity=-27.973928]
gt: 23455133 Journaling_file_system, pem_pos=0,
final_scores=-0.9694529, final_local_scores=-1.0744469,
pem=0.0, similarity=-27.973928]

21: 1219082 Filesystem_in_Userspace
[winner: 1219082 Filesystem_in_Userspace, pem_pos=0,
final_scores=-0.03799571, final_local_scores=-0.06501879,
pem=-0.07386334, similarity=2.5511029]
gt: 1219082 Filesystem_in_Userspace, pem_pos=0,

```

final_scores=-0.03799571, final_local_scores=-0.06501879,
pem=-0.07386334, similarity=2.5511029 ]
1: 6097297 Linux
    [pem_pos=0, final_scores=0.14393835,
final_local_scores=0.06656891, pem=0.0, similarity=6.076262]
2: 7580554 FreeBSD
    [pem_pos=0, final_scores=0.13651113,
final_local_scores=0.06471628, pem=0.0, similarity=6.0209756]
3: 28901365 NetBSD
    [pem_pos=0, final_scores=0.1395158,
final_local_scores=0.08430718, pem=0.0, similarity=6.6056075]
4: 4801 BeOS
    [pem_pos=0, final_scores=0.07288878,
final_local_scores=0.022366181, pem=-0.00836175,
similarity=4.802637]
5: 263521 MorphOS
    [pem_pos=0, final_scores=0.15976708,
final_local_scores=0.11118485, pem=0.0, similarity=7.4076905]
6: 18938683 GNU_General_Public_License
    [pem_pos=0, final_scores=0.17813967,
final_local_scores=0.123793304, pem=-0.008635326,
similarity=7.8309116]
1: 8072127 NTFS-3G
    [pem_pos=0, final_scores=0.099100366,
final_local_scores=0.059808373, pem=0.0, similarity=5.874514]
2: 1178658 OpenSolaris
    [pem_pos=0, final_scores=0.09911178,
final_local_scores=0.048333764, pem=-0.010893154,
similarity=5.591327]
3: 8072127 NTFS-3G
    [pem_pos=0, final_scores=0.037108287,
final_local_scores=-0.023603924, pem=0.0,
similarity=3.3853247]
4: 8072127 NTFS-3G
    [pem_pos=0, final_scores=0.049925223,
final_local_scores=-0.009603366, pem=0.0,
similarity=3.8031294]
5: 8072127 NTFS-3G
    [pem_pos=0, final_scores=0.038937226,
final_local_scores=-0.021606058, pem=0.0, similarity=3.444945]
6: 8072127 NTFS-3G
    [pem_pos=0, final_scores=0.06280343,
final_local_scores=0.020159468, pem=0.0, similarity=4.691312]
7: 8072127 NTFS-3G
    [pem_pos=0, final_scores=0.098795876,
final_local_scores=0.059475765, pem=0.0, similarity=5.8645883]

```

Sample: b'The_Great_White_Hope'

-LRB- film -RRB- [1 The Great White Hope 1] ' is a 1970 1
biographical 2 romantic 3 drama film written and 4 adapted fr
om the [1 1 Howard Sackler 1] 5 play of the same name . The
film was directed by [2 2 Martin Ritt 2] , starring [3 3 James

Earl Jones 3] , [4 4 Jane Alexander 4] , [5 5 Chester Morris 5] , [6 6 Hal Holbrook 6] , [7 7 Beah Richards 7] and [8 8 Moses Gunn 8] . [2 Jones 2] and [3 Alexander 3] , who also appeared in the same roles in the stage versions , both received 6 Best Actor and 7 Actress nominations for their performances . Set between 1910 and 1915 , the story follows [4 Jack Jefferson 4] -LRB- [9 9 James Earl Jones 9] ; patterned after real-life boxer [10 10 Jack Johnson 10] -RRB- going on a hot streak of victories in the boxing ring as he defeats every white boxer around . Soon the press and other whites announce the search for a `` great white hope '' , a boxer who will defeat [5 Jefferson 5] for the Heavyweight title . Meanwhile , [6 Jefferson 6] prepares for a few more matches , but he lets his guard down by courting the beautiful -LRB- and very white -RRB- Eleanor Bachman -LRB- [11 11 Jane Alexander 11] -RRB- , and when everyone , including [7 Jack 7] 's black `` wife '' , discover this , the tensions grow to fever pitch . [8 Jack 8] 's close black friends become scared over his pushing the envelope of success and the white authorities conspire to frame him with unlawful sexual relations with Eleanor and thereby take away his title . It leads to jealousy , a run from the law , and finally , disaster . The film opened to positive responses from both audiences and critics . They especially loved the performances of both [12 12 James Earl Jones 12] and [13 13 Jane Alexander 13] , who were in the original stage play : they won for the play . [9 Jones 9] would get bigger roles after this film , and [10 Alexander 10] made a debut here . [11 Jones 11] later contributed commentary to a documentary about [12 Jack Johnson 12] that would sum up this film , saying : `` To know the story of [13 Jack Johnson 13] is to know that it is a study in 8 hubris

- 1: 6901831 Howard_Sackler
- 2: 623662 Martin_Ritt
- 3: 18622049 James_Earl_Jones
- 4: 217809 Jane_Alexander
- 5: 142787 Chester_Morris
- 6: 870619 Hal_Holbrook
- 7: 1500049 Beah_Richards
- 8: 2802489 Moses_Gunn
- 9: 18622049 James_Earl_Jones
- 10: 71965 Jack_Johnson_(boxer)
- 11: 217809 Jane_Alexander
- 12: 18622049 James_Earl_Jones
- 13: 217809 Jane_Alexander

1: 781998 Biographical_film

[winner: 781998 Biographical_film, pem_pos=0, final_scores=-1.1367462, final_local_scores=-1.2576499, pem=0.0, similarity=-33.441067]

gt: 781998 Biographical_film, pem_pos=0, final_scores=-1.1367462, final_local_scores=-1.2576499, pem=0.0, similarity=-33.441067]

2: 522822 Romance_film
 [winner: 6301878 Romantic_(song), pem_pos=5,
 final_scores=-1.1985798, final_local_scores=-1.1914084,
 pem=-2.1598196, similarity=-19.718855|
 gt: 522822 Romance_film, pem_pos=4,
 final_scores=-1.4086162, final_local_scores=-1.4988751,
 pem=-2.1321309, similarity=-29.044847]

3: 8089 Drama_film
 [winner: 8089 Drama_film, pem_pos=0,
 final_scores=-0.81835926, final_local_scores=-0.8416368,
 pem=0.0, similarity=-21.026407|
 gt: 8089 Drama_film, pem_pos=0,
 final_scores=-0.81835926, final_local_scores=-0.8416368,
 pem=0.0, similarity=-21.026407]

4: 1851712 Film_adaptation
 [winner: 1293541 The_Silence_(The_Twilight_Zone),
 pem_pos=1, final_scores=-1.039253,
 final_local_scores=-1.0042148, pem=-1.1664349,
 similarity=-19.534805|
 gt: 1851712 Film_adaptation, pem_pos=0,
 final_scores=-1.2261292, final_local_scores=-1.3631406,
 pem=-0.37320423, similarity=-34.559578]

5: 161916 The_Great_White_Hope
 [winner: 26440260 First_Lady_(play), pem_pos=0,
 final_scores=-0.767492, final_local_scores=-0.7292049,
 pem=0.0, similarity=-17.671215|
 RECALL MISS: 161916 The_Great_White_Hope]
 [winner: 26517843 Carlos_Trillo_Name, pem_pos=9,
 final_scores=-0.775517, final_local_scores=-0.6640164,
 pem=-2.3025851, similarity=-3.2040453|
 RECALL MISS: 161916 The_Great_White_Hope]

6: 23245410 Academy_Award_for_Best_Actor
 [winner: 23245410 Academy_Award_for_Best_Actor,
 pem_pos=0, final_scores=-0.17309262,
 final_local_scores=-0.22591977, pem=-0.9233848,
 similarity=2.3693264|
 gt: 23245410 Academy_Award_for_Best_Actor, pem_pos=0,
 final_scores=-0.17309262, final_local_scores=-0.22591977,
 pem=-0.9233848, similarity=2.3693264]

7: 23245556 Academy_Award_for_Best_Actress
 [winner: 21504235 Actor, pem_pos=0,
 final_scores=-0.25659928, final_local_scores=-0.27154058,
 pem=-0.23982608, similarity=-2.7093883|
 gt: 23245556 Academy_Award_for_Best_Actress, pem_pos=1,
 final_scores=-0.4221074, final_local_scores=-0.4874783,
 pem=-1.9148145, similarity=-0.0445521]

8: 14282 Hubris
 [winner: 14282 Hubris, pem_pos=0,
 final_scores=-0.9557244, final_local_scores=-0.9088973,
 pem=-0.0037958452, similarity=-23.012953|
 gt: 14282 Hubris, pem_pos=0, final_scores=-0.9557244,
 final_local_scores=-0.9088973, pem=-0.0037958452,

similarity=-23.012953]

1: 6901831 Howard_Sackler

[pem_pos=0, final_scores=0.16426753,
final_local_scores=0.20382899, pem=0.0, similarity=10.172377]

2: 623662 Martin_Ritt

[pem_pos=0, final_scores=0.25800258,
final_local_scores=0.26240206, pem=0.0, similarity=11.920314]

3: 18622049 James_Earl_Jones

[pem_pos=0, final_scores=0.34279805,
final_local_scores=0.30627617, pem=0.0, similarity=13.229605]

4: 217809 Jane_Alexander

[pem_pos=0, final_scores=0.06930722,
final_local_scores=0.026872516, pem=0.0, similarity=4.8916426]

5: 142787 Chester_Morris

[pem_pos=0, final_scores=0.121503666,
final_local_scores=0.124845564, pem=0.0, similarity=7.8153534]

6: 870619 Hal_Holbrook

[pem_pos=0, final_scores=0.20230602,
final_local_scores=0.17332703, pem=0.0, similarity=9.262137]

7: 1500049 Beah_Richards

[pem_pos=0, final_scores=0.12363191,
final_local_scores=0.13942239, pem=0.0, similarity=8.250355]

8: 2802489 Moses_Gunn

[pem_pos=0, final_scores=0.11092688,
final_local_scores=0.13348141, pem=0.0, similarity=8.073065]

9: 18622049 James_Earl_Jones

[pem_pos=0, final_scores=0.18901251,
final_local_scores=0.13828886, pem=0.0, similarity=8.216528]

10: 71965 Jack_Johnson_(boxer)

[pem_pos=1, final_scores=0.17719553,
final_local_scores=0.1255168, pem=-0.85534924,
similarity=12.486907]

11: 217809 Jane_Alexander

[pem_pos=0, final_scores=0.088155344,
final_local_scores=0.047461227, pem=0.0, similarity=5.5060506]

12: 18622049 James_Earl_Jones

[pem_pos=0, final_scores=0.34072596,
final_local_scores=0.30401278, pem=0.0, similarity=13.162061]

13: 217809 Jane_Alexander

[pem_pos=0, final_scores=0.1896206,
final_local_scores=0.15884963, pem=0.0, similarity=8.830103]

1: 161916 The_Great_White_Hope

[pem_pos=0, final_scores=0.055234298,
final_local_scores=0.03757623, pem=0.0, similarity=5.2110624]

2: 18622049 James_Earl_Jones

[pem_pos=0, final_scores=0.31832832,
final_local_scores=0.27954674, pem=0.0, similarity=12.431945]

3: 217809 Jane_Alexander

[pem_pos=0, final_scores=0.15106462,
final_local_scores=0.11618006, pem=0.0, similarity=7.5567575]

4: 71965 Jack_Johnson_(boxer)

[pem_pos=0, final_scores=0.34244055,

final_local_scores=0.30602187, pem=0.0, similarity=13.222016]
 5: 71965 Jack_Johnson_(boxer)
 [pem_pos=0, final_scores=0.43175524,
 final_local_scores=0.4035846, pem=0.0, similarity=16.133482]
 6: 71965 Jack_Johnson_(boxer)
 [pem_pos=0, final_scores=0.41208917,
 final_local_scores=0.38210237, pem=0.0, similarity=15.49241]
 7: 71965 Jack_Johnson_(boxer)
 [pem_pos=0, final_scores=0.22410582,
 final_local_scores=0.18124652, pem=0.0, similarity=9.49847]
 8: 71965 Jack_Johnson_(boxer)
 [pem_pos=0, final_scores=0.36462706,
 final_local_scores=0.3347446, pem=0.0, similarity=14.07916]
 9: 18622049 James_Earl_Jones
 [pem_pos=0, final_scores=0.44751674,
 final_local_scores=0.42066544, pem=0.0, similarity=16.64321]
 10: 217809 Jane_Alexander
 [pem_pos=0, final_scores=0.27116537,
 final_local_scores=0.24737188, pem=0.0, similarity=11.471784]
 11: 18622049 James_Earl_Jones
 [pem_pos=0, final_scores=0.25849134,
 final_local_scores=0.21418387, pem=0.0, similarity=10.481387]
 12: 71965 Jack_Johnson_(boxer)
 [pem_pos=0, final_scores=0.26334614,
 final_local_scores=0.22308019, pem=0.0, similarity=10.746871]
 13: 71965 Jack_Johnson_(boxer)
 [pem_pos=0, final_scores=0.3514915,
 final_local_scores=0.3193656, pem=0.0, similarity=13.620219]

Sample: b'Peire_Pelet'

[1 Peire Pelet 1] -LRB- died 1303 -RRB- was the 1 **conseigneur**
 of [1 1 Alès 1] in the [2 2 Languedoc 2] . He was married to
 Delfina -LRB- Delphine -RRB- , a sister of [2 2 Henry 2] II of
 [3 Rodez 3] . He is the `senher d'Alest ' -LRB- the lord of [4
 4 Alès 4] -RRB- referred to as a participant in the 3 **torneya**
men Senhe n'Enric , us reys un ric avar along with his
 brother-in-law [5 Henry 5] and the 4 **troubadour** [3 3 Guiraut R
 iquier 3] . [6 Peire 6] was descended from a Raymond Pelet
 who took part in the [4 4 First Crusade 4] . His father was
 Bernard Pelet , who died in September 1252 , leaving his
 entire heritage to his eldest son , William , under the
 tutelage of Bernard de Barre , Guillaume de Pontils , and
 Jean de Bossoles . In 1253 William , with his tutors and his
 younger brothers , including [7 Peire 7] , received the 5 **hom**
age of their father 's vassals : Hugues de Melet , Pierre de
 Spinasson , and many others , in the presence of Pierre
 Gaucelin , Bérenger de la Fare , Arnaud d'Arsac , and Pierre
 Gaucelin de Follaquier . On the death of [8 William 8]
 without heirs [9 Peire 9] succeeded him . He married Delfina
 sometime before April 1276 , when her brother released a sum
 of money left her by their father , [10 6 Henry I 10] : a
 hundred marks of silver or 5,000 . In 1274 7 **James I** of [11 Ar

agon 11] renounced the homage and owed by [12 Peire 12] for his fief since his great-great-grandfather , , , had rendered it to [13 James 13] 's grandfather , 8 Alfonso II . [14 Peire 14] then began a war with , , over the county of [15 Melgueil 15] , which at that time was held by the bishop from the . In 1276 the combatants were brought before the episcopal court of the . [16 Peire 16] lost the case and was ordered to pay 1,000 in compensation . Joseph Angalde placed the aforementioned torneyamen around 1280 -- 1 , before the death of [17 Peire 17] , which he believed occurred in 1282 ; after the return of [18 Riquier 18] from 9 Castile in 1279 ; and after the succession of [19 Henry 19] , whom he believed was referred to as coms -LRB- count -RRB- in the torneyamen , in 1274 . This dating has been revised in light of [20 Anglade 20] 's errors -LRB- Peire died in 1303 and [21 Henry 21] is referred to only as senher , lord -RRB- . Since all three debaters call on the adjudication of the , probably

1: 521553 Alès

2: 83900 Languedoc

3: 986793 Guiraut_Riquier

4: 106128 First_Crusade

1: 69304 Lord

[winner: 69304 Lord, pem_pos=0, final_scores=-1.3846593, final_local_scores=-1.5481429, pem=0.0, similarity=-42.10996|

gt: 69304 Lord, pem_pos=0, final_scores=-1.3846593, final_local_scores=-1.5481429, pem=0.0, similarity=-42.10996]

2: 14611489 Henry_II_of_Rodez

[winner: 1188347 Rodez, pem_pos=0, final_scores=0.42451394, final_local_scores=0.4572984, pem=0.0, similarity=17.73641|

RECALL MISS: 14611489 Henry_II_of_Rodez]

[winner: 14179 Henry_I_of_England, pem_pos=0, final_scores=0.22810899, final_local_scores=0.18469754, pem=-0.43118137, similarity=11.946287|

RECALL MISS: 14611489 Henry_II_of_Rodez]

[winner: 26011508 Atsushi_Ii, pem_pos=5, final_scores=-0.7767972, final_local_scores=-0.5236273, pem=-2.1972246, similarity=0.41248035|

RECALL MISS: 14611489 Henry_II_of_Rodez]

3: 18943222 Torneyamen

[winner: 18943222 Torneyamen, pem_pos=0, final_scores=-0.8757681, final_local_scores=-0.8012891, pem=0.0, similarity=-19.822351|

gt: 18943222 Torneyamen, pem_pos=0, final_scores=-0.8757681, final_local_scores=-0.8012891, pem=0.0, similarity=-19.822351]

4: 63788 Troubadour

[winner: 63788 Troubadour, pem_pos=0, final_scores=-1.517076, final_local_scores=-1.6639683, pem=-0.01333596, similarity=-45.493896|

gt: 63788 Troubadour, pem_pos=0, final_scores=-1.517076, final_local_scores=-1.6639683, pem=-0.01333596,

```

similarity=-45.493896 ]
5: 274341 Homage
    [winner: 379963 Prussian_Homage, pem_pos=0,
final_scores=-1.092232, final_local_scores=-1.1178825,
pem=0.0, similarity=-29.270128]
    gt: 274341 Homage, pem_pos=1, final_scores=-1.3902373,
final_local_scores=-1.5083196, pem=-0.09785581,
similarity=-40.3894 ]
6: 14758849 Henry_I_of_Rodez
    [winner: 14179 Henry_I_of_England, pem_pos=0,
final_scores=0.19089065, final_local_scores=0.14404216,
pem=-0.43118137, similarity=10.73305]
    RECALL MISS: 14758849 Henry_I_of_Rodez]
7: 63332 James_I_of_Aragon
    [winner: 39443 Aragon, pem_pos=0, final_scores=0.1562768,
final_local_scores=0.15833545, pem=-0.16680257,
similarity=9.721857]
    RECALL MISS: 63332 James_I_of_Aragon]
    [winner: 269055 James_VI_and_I, pem_pos=0,
final_scores=0.009396657, final_local_scores=-0.019467331,
pem=-0.08140422, similarity=3.9514577]
    gt: 63332 James_I_of_Aragon, pem_pos=2,
final_scores=-0.29992974, final_local_scores=-0.37879437,
pem=-1.7306066, similarity=2.1970413 ]
8: 70643 Alfonso_II_of_Aragon
    [winner: 70643 Alfonso_II_of_Aragon, pem_pos=1,
final_scores=-0.0032311976, final_local_scores=-0.04358618,
pem=-0.6713525, similarity=6.439934]
    gt: 70643 Alfonso_II_of_Aragon, pem_pos=1,
final_scores=-0.0032311976, final_local_scores=-0.04358618,
pem=-0.6713525, similarity=6.439934 ]
9: 750274 Kingdom_of_Castile
    [winner: 4711133 Crown_of_Castile, pem_pos=0,
final_scores=0.028276011, final_local_scores=-0.010579392,
pem=-0.6885, similarity=7.5181727]
    gt: 750274 Kingdom_of_Castile, pem_pos=1,
final_scores=0.0106488615, final_local_scores=-0.03361903,
pem=-0.9579114, similarity=8.295725 ]
1: 521553 Alès
    [pem_pos=0, final_scores=0.28923565,
final_local_scores=0.33532697, pem=-0.45544896,
similarity=16.573341]
2: 83900 Languedoc
    [pem_pos=0, final_scores=0.42334914,
final_local_scores=0.40376526, pem=0.0, similarity=16.138874]
3: 986793 Guiraut_Riquier
    [pem_pos=0, final_scores=0.13300548,
final_local_scores=0.2013484, pem=0.0, similarity=10.098351]
4: 106128 First_Crusade
    [pem_pos=0, final_scores=0.29606944,
final_local_scores=0.29473856, pem=0.0, similarity=12.8853]
1: 21376617 Peire_Pelet

```

[pem_pos=0, final_scores=0.037026033,
final_local_scores=0.014810726, pem=0.0, similarity=4.5316944]

2: 14179 Henry_I_of_England
[pem_pos=0, final_scores=0.22810899,
final_local_scores=0.18469754, pem=-0.43118137,
similarity=11.946287]

3: 1188347 Rodez
[pem_pos=0, final_scores=0.42451394,
final_local_scores=0.4572984, pem=0.0, similarity=17.73641]

4: 521553 Alès
[pem_pos=0, final_scores=0.29998386,
final_local_scores=0.34706777, pem=-0.45544896,
similarity=16.92371]

5: 14611489 Henry_II_of_Rodez
[pem_pos=0, final_scores=0.3287341,
final_local_scores=0.37348515, pem=0.0, similarity=15.235255]

6: 21376617 Peire_Pelet
[pem_pos=0, final_scores=0.1436729,
final_local_scores=0.13130623, pem=0.0, similarity=8.008153]

7: 21376617 Peire_Pelet
[pem_pos=0, final_scores=0.06487192,
final_local_scores=0.04522814, pem=0.0, similarity=5.4394107]

8: 188885 William_Whiston
[pem_pos=0, final_scores=0.040375218,
final_local_scores=0.08797823, pem=0.0, similarity=6.715159]

9: 21376617 Peire_Pelet
[pem_pos=0, final_scores=0.10431115,
final_local_scores=0.08830951, pem=0.0, similarity=6.7250447]

10: 14179 Henry_I_of_England
[pem_pos=0, final_scores=0.19089065,
final_local_scores=0.14404216, pem=-0.43118137,
similarity=10.73305]

11: 39443 Aragon
[pem_pos=0, final_scores=0.1562768,
final_local_scores=0.15833545, pem=-0.16680257,
similarity=9.721857]

12: 21376617 Peire_Pelet
[pem_pos=0, final_scores=0.034985676,
final_local_scores=0.012581944, pem=0.0, similarity=4.4651833]

13: 63332 James_I_of_Aragon
[pem_pos=0, final_scores=0.1903777,
final_local_scores=0.16301575, pem=0.0, similarity=8.954428]

14: 21376617 Peire_Pelet
[pem_pos=0, final_scores=0.067814216,
final_local_scores=0.048442155, pem=0.0, similarity=5.5353236]

15: 1102499 Mauguio
[pem_pos=0, final_scores=0.12255846,
final_local_scores=0.20640385, pem=0.0, similarity=10.249215]

16: 21376617 Peire_Pelet
[pem_pos=0, final_scores=0.10805924,
final_local_scores=0.09240374, pem=0.0, similarity=6.8472247]

17: 21376617 Peire_Pelet

[pem_pos=0, final_scores=0.09832664,
final_local_scores=0.08177234, pem=0.0, similarity=6.5299625]
18: 1803302 Richarius
[pem_pos=1, final_scores=0.043725803,
final_local_scores=0.12947142, pem=0.0, similarity=7.953398]
19: 14179 Henry_I_of_England
[pem_pos=1, final_scores=0.32705408,
final_local_scores=0.29504025, pem=-0.43118137,
similarity=15.239135]
20: 11237285 Henry_Anglade
[pem_pos=2, final_scores=0.04095368,
final_local_scores=0.109849155, pem=-1.3862944,
similarity=14.906715]
21: 14179 Henry_I_of_England
[pem_pos=1, final_scores=0.2104807,
final_local_scores=0.1684885, pem=-0.43118137,
similarity=11.462578]

Sample: b'New_York_Yankees'

-LRB- 1936_AFL -RRB- The [1 New York Yankees 1] of the was the second professional [2 1 American 2] football team competing under that name . It is unrelated to the [3 2 Yankees 3] of the -LRB- and the [1 1 National Football League 1] -RRB- , the [4 3 Yankees 4] of the , the [5 4 Yankees 5] of the -LRB- then also called the [6 American Football League 6] -RRB- and the -LRB- later -RRB- 5 Yankees of the [2 2 All-America Football Conference 2] . The [7 Yankees 7] played their home games in [8 6 Yankee Stadium 8] and in . [3 3 Jack McBride 3] was the team 's head coach throughout its existence ; [9 Yankees 9] ' president James Bush served as president of the second [10 American Football League 10] in 1936 . Formed by signing [4 4 National Football League 4] players , primarily s , the 1936 [11 Yankees 11] -LRB- including back and tackle [5 5 Jess Quatse 5] -RRB- battled the [12 7 Boston Shamrocks 12] and the [13 Cleveland Rams 13] for the league title -LRB- which [14 Boston 14] won -RRB- . New York Giants-turned-Yankees also included end and back 8 St u Clancy . The 1937 edition of the [15 Yankees 15] was never truly in the running as league newcomers [6 6 Los Angeles Bulldogs 6] were the only [16 AFL 16] team with a winning record as [17 Los Angeles 17] dominated the league with an unbeaten , untied 8-0-0 record -LRB- the [18 Yankees 18] finished in third place for the second straight year , this time with 2-3-1 -RRB- . While there was no `` official ' ' all-AFL team declared for the 1936 season , guard was the sole [19 Yankee 19] to be named to the All-League team in 1937 . Other [20 Yankees 20] stars include [7 7 Dean Nesmith 7] , , 9 Al Rose , and . The complete domination of the [21 AFL 21] by the Bulldogs decimated attendance in [22 10 Yankee Stadium 22] as [23 LA 23] played the second half of its season on the -LRB- while the [24 Yankees 24] averaged roughly 14,000 in attendance in 1936 , the number was cut to about 5000 in 1937

-RRB- . With all the [25 AFL 25] clubs -LRB- except [26 Los Angeles 26] -RRB- deeply in the red , the second [27 American Football League 27] -- and the [28 Yankees 28] -- closed up shop at the end of the 1937 season . Some sources indicate the [29 Yankees 29] of the -LRB- 1940 -RRB- a continuation of the 1936-1937 [30 Yankees 30] , citing a common head coach -LRB- [8 8 Jack McBride 8] -RRB- and the continuing existences of two other [31 AFL II 31] teams , the -LRB- which also competed in the third [32 AFL 32] -RRB- and the [9 9 Los Angeles Bulldogs 9] -LRB- which became a founding member of the 11 Pacific Coast Professional Football League in 1940 -- after competing with [33 Cincinnati 33] in the [34 American Professional Football League 34]

1: 21211 National_Football_League

2: 73669 All-America_Football_Conference

3: 10158373 Jack_McBride

4: 21211 National_Football_League

5: 15474568 Jess_Quatse

6: 16972343 Los_Angeles_Bulldogs

7: 37015348 Dean_Nesmith

8: 10158373 Jack_McBride

9: 16972343 Los_Angeles_Bulldogs

1: 18951490 American_football

[winner: 3434750 United_States, pem_pos=0,
final_scores=0.341739, final_local_scores=0.41132456,
pem=-0.12994486, similarity=17.07112]

gt: 18951490 American_football, pem_pos=6,
final_scores=-0.54561126, final_local_scores=-0.6363586,
pem=-5.032821, similarity=12.468767]

[winner: 143183 Football_(ball), pem_pos=5,
final_scores=-2.7564702, final_local_scores=-2.9407609,
pem=-4.263315, similarity=-60.483887]

gt: 18951490 American_football, pem_pos=1,
final_scores=-2.798875, final_local_scores=-3.0992467,
pem=-1.8467362, similarity=-78.35515]

2: 239756 New_York_Yankees_(NFL)

[winner: 4848143 New_York_Yankees, pem_pos=0,
final_scores=0.46670884, final_local_scores=0.4092217,
pem=-0.4294726, similarity=18.637243]

RECALL MISS: 239756 New_York_Yankees_(NFL)]

3: 17144071 New_York_Yankees_(1940_AFL)

[winner: 4848143 New_York_Yankees, pem_pos=0,
final_scores=0.4680193, final_local_scores=0.41065317,
pem=-0.4294726, similarity=18.679962]

RECALL MISS: 17144071 New_York_Yankees_(1940_AFL)]

4: 17144071 New_York_Yankees_(1940_AFL)

[winner: 4848143 New_York_Yankees, pem_pos=0,
final_scores=0.381822, final_local_scores=0.3164957,
pem=-0.4294726, similarity=15.870114]

RECALL MISS: 17144071 New_York_Yankees_(1940_AFL)]

5: 2314091 New_York_Yankees_(AAFC)

[winner: 4848143 New_York_Yankees, pem_pos=0,

```

final_scores=-0.53203994, final_local_scores=-0.6830037,
pem=-0.4294726, similarity=-13.956941|
    RECALL MISS: 2314091 New_York_Yankees_(AAFC)]
6: 40019 Yankee_Stadium_(1923)
    [winner: 22146637 Yankee_Stadium, pem_pos=0,
final_scores=0.34178478, final_local_scores=0.33480558,
pem=-0.025903488, similarity=14.221847|
    gt: 40019 Yankee_Stadium_(1923), pem_pos=1,
final_scores=0.20325483, final_local_scores=0.16626734,
pem=-0.64733, similarity=12.571743 ]
7: 6307332 Boston_Shamrocks_(AFL)
    [winner: 15805527 American_Football_League_(1936),
pem_pos=0, final_scores=0.22619967,
final_local_scores=0.20068029, pem=0.0, similarity=10.078413|
    gt: 6307332 Boston_Shamrocks_(AFL), pem_pos=2,
final_scores=-0.26008987, final_local_scores=-0.26823285,
pem=-1.9387417, similarity=6.6282883 ]
8: 29912390 Stu_Clancy
    [winner: 29912390 Stu_Clancy, pem_pos=0,
final_scores=-0.18718658, final_local_scores=-0.103212625,
pem=0.0, similarity=1.0096422|
    gt: 29912390 Stu_Clancy, pem_pos=0,
final_scores=-0.18718658, final_local_scores=-0.103212625,
pem=0.0, similarity=1.0096422 ]
9: 20921117 Al_Rose
    [winner: 20921117 Al_Rose, pem_pos=0,
final_scores=-0.09921309, final_local_scores=-0.034450322,
pem=0.0, similarity=3.0616467|
    gt: 20921117 Al_Rose, pem_pos=0,
final_scores=-0.09921309, final_local_scores=-0.034450322,
pem=0.0, similarity=3.0616467 ]
10: 40019 Yankee_Stadium_(1923)
    [winner: 22146637 Yankee_Stadium, pem_pos=0,
final_scores=0.3088525, final_local_scores=0.29883206,
pem=-0.025903488, similarity=13.148325|
    gt: 40019 Yankee_Stadium_(1923), pem_pos=1,
final_scores=0.16868989, final_local_scores=0.12851039,
pem=-0.64733, similarity=11.445001 ]
11: 16790451 Pacific_Coast_Professional_Football_League
    [winner: 16790451
Pacific_Coast_Professional_Football_League, pem_pos=0,
final_scores=0.028047785, final_local_scores=0.01113753,
pem=0.0, similarity=4.422079|
    gt: 16790451 Pacific_Coast_Professional_Football_League,
pem_pos=0, final_scores=0.028047785,
final_local_scores=0.01113753, pem=0.0, similarity=4.422079 ]
1: 21211 National_Football_League
    [pem_pos=0, final_scores=0.39312965,
final_local_scores=0.3458725, pem=0.0, similarity=14.411238]
2: 73669 All-America_Football_Conference
    [pem_pos=0, final_scores=0.30312723,
final_local_scores=0.27112174, pem=0.0, similarity=12.180527]

```

3: 10158373 Jack_McBride
[pem_pos=0, final_scores=0.06323515,
final_local_scores=0.114922136, pem=0.0, similarity=7.5192184]
4: 21211 National_Football_League
[pem_pos=0, final_scores=0.4268784,
final_local_scores=0.38273793, pem=0.0, similarity=15.511376]
5: 15474568 Jess_Quatse
[pem_pos=0, final_scores=0.27777338,
final_local_scores=0.33623502, pem=0.0, similarity=14.123636]
6: 16972343 Los_Angeles_Bulldogs
[pem_pos=0, final_scores=0.16840674,
final_local_scores=0.15677595, pem=0.0, similarity=8.76822]
7: 37015348 Dean_Nesmith
[pem_pos=0, final_scores=0.035074756,
final_local_scores=0.16388792, pem=0.0, similarity=8.980455]
8: 10158373 Jack_McBride
[pem_pos=0, final_scores=0.054413125,
final_local_scores=0.10528539, pem=0.0, similarity=7.231639]
9: 16972343 Los_Angeles_Bulldogs
[pem_pos=0, final_scores=0.14840217,
final_local_scores=0.134924, pem=0.0, similarity=8.116114]
1: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.5354957,
final_local_scores=0.48436087, pem=-0.14222674,
similarity=19.317457]
2: 3434750 United_States
[pem_pos=0, final_scores=0.341739,
final_local_scores=0.41132456, pem=-0.12994486,
similarity=17.07112]
3: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.46670884,
final_local_scores=0.4092217, pem=-0.4294726,
similarity=18.637243]
4: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.4680193,
final_local_scores=0.41065317, pem=-0.4294726,
similarity=18.679962]
5: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.381822,
final_local_scores=0.3164957, pem=-0.4294726,
similarity=15.870114]
6: 2357 American_Football_League
[pem_pos=0, final_scores=0.43442708,
final_local_scores=0.3982343, pem=0.0, similarity=15.973819]
7: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.4093421,
final_local_scores=0.34655723, pem=-0.4294726,
similarity=16.76721]
8: 22146637 Yankee_Stadium
[pem_pos=0, final_scores=0.34178478,
final_local_scores=0.33480558, pem=-0.025903488,
similarity=14.221847]

9: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.48088402,
final_local_scores=0.42470592, pem=-0.4294726,
similarity=19.099325]
10: 2357 American_Football_League
[pem_pos=0, final_scores=0.47548395,
final_local_scores=0.44308275, pem=0.0, similarity=17.312185]
11: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.38241315,
final_local_scores=0.31714144, pem=-0.4294726,
similarity=15.889385]
12: 15805527 American_Football_League_(1936)
[pem_pos=0, final_scores=0.22619967,
final_local_scores=0.20068029, pem=0.0, similarity=10.078413]
13: 23408886 St._Louis_Rams
[pem_pos=0, final_scores=0.27939206,
final_local_scores=0.24225697, pem=-0.0891782,
similarity=11.804109]
14: 4339 Boston_Red_Sox
[pem_pos=5, final_scores=0.44010395,
final_local_scores=0.39202565, pem=-4.418931,
similarity=39.819374]
15: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.43011075,
final_local_scores=0.3692438, pem=-0.4294726,
similarity=17.444223]
16: 4732785 Australian_Football_League
[pem_pos=0, final_scores=0.16736959,
final_local_scores=0.20268762, pem=-0.73298496,
similarity=14.124402]
17: 18213 Los_Angeles_Dodgers
[pem_pos=15, final_scores=0.40190285,
final_local_scores=0.34313864, pem=-4.7524934,
similarity=40.174454]
18: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.465555,
final_local_scores=0.4079613, pem=-0.4294726,
similarity=18.59963]
19: 4848143 New_York_Yankees
[pem_pos=1, final_scores=0.20922993,
final_local_scores=0.12796512, pem=-2.2829223,
similarity=20.323334]
20: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.50425524,
final_local_scores=0.45023543, pem=-0.4294726,
similarity=19.861174]
21: 2357 American_Football_League
[pem_pos=1, final_scores=0.13031732,
final_local_scores=0.0663133, pem=-1.0971715,
similarity=12.035224]
22: 22146637 Yankee_Stadium
[pem_pos=0, final_scores=0.3088525,

final_local_scores=0.29883206, pem=-0.025903488,
 similarity=13.148325]
 23: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.099917725,
 final_local_scores=0.17476758, pem=-1.2080293,
 similarity=15.8745775]
 24: 4848143 New_York_Yankees
 [pem_pos=0, final_scores=0.46097332,
 final_local_scores=0.40295655, pem=-0.4294726,
 similarity=18.450277]
 25: 4732785 Australian_Football_League
 [pem_pos=0, final_scores=0.18355049,
 final_local_scores=0.22036278, pem=-0.73298496,
 similarity=14.651863]
 26: 18213 Los_Angeles_Dodgers
 [pem_pos=15, final_scores=0.15522452,
 final_local_scores=0.07380766, pem=-4.7524934,
 similarity=32.137077]
 27: 2357 American_Football_League
 [pem_pos=0, final_scores=0.49136454,
 final_local_scores=0.46042985, pem=0.0, similarity=17.829859]
 28: 4848143 New_York_Yankees
 [pem_pos=0, final_scores=0.41013354,
 final_local_scores=0.3474218, pem=-0.4294726,
 similarity=16.79301]
 29: 4848143 New_York_Yankees
 [pem_pos=0, final_scores=0.45296204,
 final_local_scores=0.3942054, pem=-0.4294726,
 similarity=18.189125]
 30: 4848143 New_York_Yankees
 [pem_pos=0, final_scores=0.47149998,
 final_local_scores=0.41445524, pem=-0.4294726,
 similarity=18.79342]
 31: 15805527 American_Football_League_(1936)
 [pem_pos=0, final_scores=0.15980156,
 final_local_scores=0.12802067, pem=0.0, similarity=7.910105]
 32: 4732785 Australian_Football_League
 [pem_pos=0, final_scores=0.09312205,
 final_local_scores=0.12126899, pem=-0.73298496,
 similarity=11.694707]
 33: 6671 Cincinnati_Reds
 [pem_pos=4, final_scores=0.36437196,
 final_local_scores=0.30852252, pem=-3.771247,
 similarity=33.805267]
 34: 21211 National_Football_League
 [pem_pos=1, final_scores=0.22808646,
 final_local_scores=0.16607529, pem=-0.6931472,
 similarity=12.815173]

Sample: b'Doug_Melvin'

_ -LRB- rower -RRB- Douglas V Melvin -LRB- born 1928 -RRB-
 is a [1 British 1] rower who twice won the [1 1 Wingfield Scu

lls 1] , the amateur sculling championship of the [2 2 River Thames 2] . Melvin was born at [2 1 Lancaster 2] . He took a job there with the 2 Electricity Board , and was working there and rowing for 3 John O'Gaunt Rowing Club when he won the [3 3 Wingfield Sculls 3] in 1955 . He was encouraged by Eric Phelps to join [4 4 London Rowing Club 4] and was able to transfer his Electricity Board job to [5 5 Wandsworth 5] in order to do so . This gave him the opportunity to train alongside other notable scullers at the club including 4 Farn Carpmael , [6 6 John Marsden 6] , [7 7 Tony Fox 7] and [8 8 John Pinches 8] . In the [9 9 Diamond Challenge Sculls 9] in 1955 he reached the semi-final where he met [10 10 Teodor Kocerka 10] , the eventual winner . He went on to represent [3 Great Britain 3] in the single sculls at the 5 European Championships later the same year . In 1956 he was selected as the spare man for the [4 6 Great Britain 4] team for the 7 1956 Summer Olympics in . After his boat and blades were shipped off to the Games he was de-selected in favour of someone more inclined to 8 rowing rather than 9 sculling . In 1957 he won the 10 Scullers Head of the River Race

1: 9594440 Wingfield_Sculls
 2: 49031 River_Thames
 3: 9594440 Wingfield_Sculls
 4: 3977632 London_Rowing_Club
 5: 94365 Wandsworth
 6: 19377088 John_Marsden_(rower)
 7: 23653980 Tony_Fox
 8: 23613446 John_Pinches
 9: 12963154 Diamond_Challenge_Sculls
 10: 20727947 Teodor_Kocerka

1: 87578 Lancaster,_Lancashire

[winner: 82388 Lancaster,_Pennsylvania, pem_pos=1, final_scores=0.16419469, final_local_scores=0.20575747, pem=-2.1130877, similarity=21.721226]

gt: 87578 Lancaster,_Lancashire, pem_pos=0, final_scores=0.13607933, final_local_scores=0.16925985, pem=-1.2549295, similarity=15.965266]

2: 340733 Central_Electricity_Generating_Board

[winner: 2655363 Public_electricity_suppliers, pem_pos=0, final_scores=-0.06107837, final_local_scores=-0.0126966685, pem=0.0, similarity=3.710819]

gt: 340733 Central_Electricity_Generating_Board, pem_pos=1, final_scores=-0.0704446, final_local_scores=-0.05510796, pem=-0.68441355, similarity=6.167129]

3: 26861176 Lancaster_John_O'_Gaunt_Rowing_Club

[winner: 26861176 Lancaster_John_O'_Gaunt_Rowing_Club, pem_pos=0, final_scores=-0.19550811, final_local_scores=-0.16661294, pem=0.0, similarity=-0.88234943]

gt: 26861176 Lancaster_John_O'_Gaunt_Rowing_Club, pem_pos=0, final_scores=-0.19550811,

final_local_scores=-0.16661294, pem=0.0,
 similarity=-0.88234943]
 4: 23655064 Farn_Carpmael
 [winner: 23655064 Farn_Carpmael, pem_pos=0,
 final_scores=-0.09975603, final_local_scores=-0.039307,
 pem=0.0, similarity=2.9167137|
 gt: 23655064 Farn_Carpmael, pem_pos=0,
 final_scores=-0.09975603, final_local_scores=-0.039307,
 pem=0.0, similarity=2.9167137]
 5: 1152427 European_Rowing_Championships
 [winner: 1825838 European_Athletics_Championships,
 pem_pos=0, final_scores=-0.21076222,
 final_local_scores=-0.198522, pem=-2.1033094,
 similarity=9.603544|
 RECALL MISS: 1152427 European_Rowing_Championships]
 6: 7755971 Rowing_at_the_1956_Summer_Olympics
 [winner: 13530298 Great_Britain, pem_pos=0,
 final_scores=0.1216353, final_local_scores=0.0801025,
 pem=-0.50213265, similarity=9.210807|
 RECALL MISS: 7755971 Rowing_at_the_1956_Summer_Olympics]
 [winner: 14708158 TEAMS_(cable_system), pem_pos=3,
 final_scores=-1.857702, final_local_scores=-1.8807881,
 pem=-3.9775972, similarity=-30.405964|
 RECALL MISS: 7755971 Rowing_at_the_1956_Summer_Olympics]
 7: 83870 1956_Summer_Olympics
 [winner: 27566 Summer_Olympic_Games, pem_pos=0,
 final_scores=0.003579244, final_local_scores=0.021952927,
 pem=-0.4112264, similarity=6.9811454|
 RECALL MISS: 83870 1956_Summer_Olympics]
 [winner: 34672 1956, pem_pos=0, final_scores=-0.6396378,
 final_local_scores=-0.44870463, pem=-0.7774988,
 similarity=-5.072369|
 gt: 83870 1956_Summer_Olympics, pem_pos=6,
 final_scores=-1.5608342, final_local_scores=-1.7108402,
 pem=-4.0945764, similarity=-24.698227]
 8: 23172664 Sweep_(rowing)
 [winner: 7742635 Rowing_at_the_1964_Summer_Olympics,
 pem_pos=9, final_scores=-1.3851944,
 final_local_scores=-1.4049773, pem=-3.4014323,
 similarity=-19.340092|
 RECALL MISS: 23172664 Sweep_(rowing)]
 9: 927319 Sculling
 [winner: 927319 Sculling, pem_pos=1,
 final_scores=-1.3517557, final_local_scores=-1.4929998,
 pem=-0.09729888, similarity=-39.93525|
 gt: 927319 Sculling, pem_pos=1, final_scores=-1.3517557,
 final_local_scores=-1.4929998, pem=-0.09729888,
 similarity=-39.93525]
 10: 1489859 Scullers_Head_of_the_River_Race
 [winner: 1489859 Scullers_Head_of_the_River_Race,
 pem_pos=0, final_scores=-0.10298072,
 final_local_scores=-0.062281944, pem=0.0,

```
similarity=2.2310956|
    gt: 1489859 Scullers_Head_of_the_River_Race, pem_pos=0,
final_scores=-0.10298072, final_local_scores=-0.062281944,
pem=0.0, similarity=2.2310956 ]
1: 9594440 Wingfield_Sculls
    [pem_pos=0, final_scores=0.27242297,
final_local_scores=0.26522627, pem=0.0, similarity=12.004594]
2: 49031 River_Thames
    [pem_pos=0, final_scores=0.29603153,
final_local_scores=0.3085678, pem=0.0, similarity=13.297992]
3: 9594440 Wingfield_Sculls
    [pem_pos=0, final_scores=0.14594059,
final_local_scores=0.12706351, pem=0.0, similarity=7.8815417]
4: 3977632 London_Rowing_Club
    [pem_pos=0, final_scores=0.081215724,
final_local_scores=0.0826761, pem=0.0, similarity=6.5569324]
5: 94365 Wandsworth
    [pem_pos=0, final_scores=0.12235065,
final_local_scores=0.10727695, pem=-0.33546227,
similarity=9.115367]
6: 19377088 John_Marsden_(rower)
    [pem_pos=2, final_scores=0.15672158,
final_local_scores=0.19664165, pem=-1.2891831,
similarity=16.96867]
7: 23653980 Tony_Fox
    [pem_pos=0, final_scores=0.111484155,
final_local_scores=0.14834902, pem=0.0, similarity=8.516744]
8: 23613446 John_Pinches
    [pem_pos=0, final_scores=0.19599505,
final_local_scores=0.22974813, pem=0.0, similarity=10.945855]
9: 12963154 Diamond_Challenge_Sculls
    [pem_pos=0, final_scores=0.14217867,
final_local_scores=0.15925989, pem=0.0, similarity=8.842345]
10: 20727947 Teodor_Kocerka
    [pem_pos=0, final_scores=0.46008343,
final_local_scores=0.5694102, pem=0.0, similarity=21.082048]
1: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.49196565,
final_local_scores=0.50003356, pem=-0.48936236,
similarity=21.672941]
2: 82388 Lancaster,_Pennsylvania
    [pem_pos=1, final_scores=0.16419469,
final_local_scores=0.20575747, pem=-2.1130877,
similarity=21.721226]
3: 13530298 Great_Britain
    [pem_pos=0, final_scores=0.08655594,
final_local_scores=0.04178363, pem=-0.50213265,
similarity=8.067295]
4: 13530298 Great_Britain
    [pem_pos=0, final_scores=0.1216353,
final_local_scores=0.0801025, pem=-0.50213265,
similarity=9.210807]
```

Sample: b'Hassanine_Sebei'

[1 Hassanine Sebei_1] , , -LRB- born 21 April 1984 in 1 Taje
rouine -RRB- is a [1 1 Tunisia 1] n who specialises in the 2 2
0 kilometres race walk . He was the 2010 [2 3 African 2] cham
pion and competed at the 4 2008 [3 Summer Olympics 3] . He
won the 5 silver medal at the 6 2004 African Championships in
Athletics . The following year he took 7 gold medal s at the 8
2005 Islamic Solidarity Games and the [4 Pan Arab
Championships 4] , as well as 9 bronze medal s at both the [5
10 Jeux de la Francophonie 5] and the African Race Walking
Championships . He made his debut in the [2 2 IAAF World Race
Walking Cup 2] the following year and finished 57th overall .
He missed out on the medals at the 11 2006 African Championsh
ips in Athletics , taking fourth place in the event . [6
Sebei 6] took part in the 12 2007 World Championships in Athl
etics and came in 24th place , the second best [7 African 7]
performer . Coached by [3 3 Hatem Ghoula 3] , he succeeded
his tutor to the 20,000 m track walk title at the 13 2007 [8 P
an Arab Games 8] in November . At the 14 2008 African Champio
nships in Athletics in 15 Addis Ababa he came in third place
and he represented [9 Tunisia 9] at the later that year ,
where he was 34th in the 20 km walk . In 2009 he was fourth
at the 16 Mediterranean Games , had a top twenty finish at
the 17 World Championships and ended the year with a silver
medal behind at the [10 18 Jeux de la Francophonie 10] . At
the 19 2010 [11 IAAF World Race Walking Cup 11] he had his
best global finish by coming in eleventh place . He won his
first continental title at the 20 2010 African Championships i
n Athletics , setting a personal best and

1: 30188 Tunisia

2: 4589004 IAAF_World_Race_Walking_Cup

3: 7074651 Hatem_Ghoula

1: 26028553 Tajerouine

[winner: 26028553 Tajerouine, pem_pos=0,
final_scores=-0.18383095, final_local_scores=0.01406315,
pem=0.0, similarity=4.5093856]

gt: 26028553 Tajerouine, pem_pos=0,
final_scores=-0.18383095, final_local_scores=0.01406315,
pem=0.0, similarity=4.5093856]

2: 23873342 20_kilometres_race_walk

[winner: 251442 Road_running, pem_pos=0,
final_scores=-0.843708, final_local_scores=-0.93238485,
pem=0.0, similarity=-23.73451]

gt: 23873342 20_kilometres_race_walk, pem_pos=3,
final_scores=-1.2335962, final_local_scores=-1.3401473,
pem=-3.0445223, similarity=-19.346365]

[winner: 1294089 London_Outer_Orbital_Path, pem_pos=2,
final_scores=-0.94422, final_local_scores=-0.8610588,
pem=-0.6931472, similarity=-17.836557]

RECALL MISS: 23873342 20_kilometres_race_walk]

[winner: 7885197

2004_Deutsche_Tourenwagen_Masters_season, pem_pos=14,
final_scores=-1.3701197, final_local_scores=-1.3533449,
pem=-2.944439, similarity=-20.284483|
RECALL MISS: 23873342 20_kilometres_race_walk]

3: 6278087 African_Championships_in_Athletics
[winner: 5334607 Africa, pem_pos=0,
final_scores=0.07989009, final_local_scores=0.114433885,
pem=-0.91165566, similarity=12.462373|
RECALL MISS: 6278087 African_Championships_in_Athletics]
[winner: 5209263 Champion (role_variant), pem_pos=1,
final_scores=-1.2095305, final_local_scores=-1.1453292,
pem=-2.3057463, similarity=-17.550192|
RECALL MISS: 6278087 African_Championships_in_Athletics]

4: 77745 2008_Summer_Olympics
[winner: 27566 Summer_Olympic_Games, pem_pos=0,
final_scores=0.14749782, final_local_scores=0.13684452,
pem=-0.4112264, similarity=10.4097395|
gt: 77745 2008_Summer_Olympics, pem_pos=2,
final_scores=-0.54077774, final_local_scores=-0.6045225,
pem=-3.4497817, similarity=4.8100038]
[winner: 35825 2008, pem_pos=0, final_scores=-0.77882105,
final_local_scores=-0.682548, pem=-0.29331458,
similarity=-14.68379|
gt: 77745 2008_Summer_Olympics, pem_pos=12,
final_scores=-1.7375141, final_local_scores=-1.9212697,
pem=-5.467841, similarity=-23.509836]

5: 526647 Silver_medal
[winner: 375211 Medal, pem_pos=0,
final_scores=-0.69906473, final_local_scores=-0.72823554,
pem=-0.28203553, similarity=-16.108536|
RECALL MISS: 526647 Silver_medal]
[winner: 27119 Silver, pem_pos=0,
final_scores=-1.0948858, final_local_scores=-1.1001983,
pem=-0.12949902, similarity=-28.03816|
gt: 526647 Silver_medal, pem_pos=2,
final_scores=-1.2905545, final_local_scores=-1.4322453,
pem=-2.3577166, similarity=-25.829714]

6: 7110318 2004_African_Championships_in_Athletics
[winner: 7110318 2004_African_Championships_in_Athletics,
pem_pos=0, final_scores=-0.19333935,
final_local_scores=-0.2048052, pem=0.0,
similarity=-2.0220823|
gt: 7110318 2004_African_Championships_in_Athletics,
pem_pos=0, final_scores=-0.19333935,
final_local_scores=-0.2048052, pem=0.0,
similarity=-2.0220823]

7: 526488 Gold_medal
[winner: 12240 Gold, pem_pos=0, final_scores=-1.1275082,
final_local_scores=-1.1436906, pem=-0.22135684,
similarity=-28.836525|
gt: 526488 Gold_medal, pem_pos=2,
final_scores=-1.3524327, final_local_scores=-1.5044172,

pem=-2.463969, similarity=-27.405651]
[winner: 375211 Medal, pem_pos=0, final_scores=-1.255066,
final_local_scores=-1.3355825, pem=-0.28203553,
similarity=-34.232983]
RECALL MISS: 526488 Gold_medal]
8: 19531349 Athletics_at_the_2005_Islamic_Solidarity_Games
[winner: 22477672 Islamic_Solidarity_Games, pem_pos=0,
final_scores=-0.00726448, final_local_scores=0.053854793,
pem=0.0, similarity=5.6968474]
gt: 19531349
Athletics_at_the_2005_Islamic_Solidarity_Games, pem_pos=1,
final_scores=-0.16103238, final_local_scores=-0.14712045,
pem=-0.5196363, similarity=2.525208]
[winner: 35984 2005, pem_pos=0, final_scores=-0.93439317,
final_local_scores=-0.83577067, pem=-0.36604404,
similarity=-18.860746]
RECALL MISS: 19531349
Athletics_at_the_2005_Islamic_Solidarity_Games]
9: 526648 Bronze_medal
[winner: 3548094 Brazil_at_the_1996_Summer_Olympics,
pem_pos=7, final_scores=-1.0863688,
final_local_scores=-1.08739, pem=-2.0775077,
similarity=-17.062359]
gt: 526648 Bronze_medal, pem_pos=0,
final_scores=-1.2337692, final_local_scores=-1.3704455,
pem=-0.023749426, similarity=-36.67796]
10: 28218044 Athletics_at_the_2005_Jeux_de_la_Francophonie
[winner: 1627067 Jeux_de_la_Francophonie, pem_pos=0,
final_scores=0.24044852, final_local_scores=0.22328311,
pem=-0.18819143, similarity=11.7763405]
gt: 28218044
Athletics_at_the_2005_Jeux_de_la_Francophonie, pem_pos=8,
final_scores=-0.6713988, final_local_scores=-0.72671974,
pem=-4.9528046, similarity=9.337068]
11: 6278392 2006_African_Championships_in_Athletics
[winner: 6278392 2006_African_Championships_in_Athletics,
pem_pos=0, final_scores=-0.1166008,
final_local_scores=-0.10066035, pem=0.0,
similarity=1.0858074]
gt: 6278392 2006_African_Championships_in_Athletics,
pem_pos=0, final_scores=-0.1166008,
final_local_scores=-0.10066035, pem=0.0,
similarity=1.0858074]
12: 3517973 2007_World_Championships_in_Athletics
[winner: 218926 IAAF_World_Championships_in_Athletics,
pem_pos=0, final_scores=0.0050622374,
final_local_scores=-0.02916672, pem=0.0,
similarity=3.2193198]
gt: 3517973 2007_World_Championships_in_Athletics,
pem_pos=2, final_scores=-0.40726393,
final_local_scores=-0.44599128, pem=-2.0854733,
similarity=2.121572]

[winner: 36165 2007, pem_pos=0, final_scores=-1.0300606,
final_local_scores=-0.9690641, pem=-0.24805899,
similarity=-23.480108]

RECALL MISS: 3517973

2007_World_Championships_in_Athletics]

13: 28015813 Athletics_at_the_2007_Pan_Arab_Games

[winner: 3747037 Pan_Arab_Games, pem_pos=0,
final_scores=0.05247931, final_local_scores=0.03196992,
pem=-0.86277163, similarity=9.735645]

gt: 28015813 Athletics_at_the_2007_Pan_Arab_Games,
pem_pos=11, final_scores=-0.7763571,
final_local_scores=-0.8292703, pem=-5.2186913,
similarity=7.7226863]

[winner: 36165 2007, pem_pos=0, final_scores=-0.96233314,
final_local_scores=-0.89508206, pem=-0.24805899,
similarity=-21.272337]

RECALL MISS: 28015813

Athletics_at_the_2007_Pan_Arab_Games]

14: 7546635 2008_African_Championships_in_Athletics

[winner: 7546635 2008_African_Championships_in_Athletics,
pem_pos=0, final_scores=-0.13493288,
final_local_scores=-0.13651854, pem=0.0,
similarity=0.015727282]

gt: 7546635 2008_African_Championships_in_Athletics,
pem_pos=0, final_scores=-0.13493288,
final_local_scores=-0.13651854, pem=0.0,
similarity=0.015727282]

15: 52593 Addis_Ababa

[winner: 52593 Addis_Ababa, pem_pos=0,
final_scores=-0.002520293, final_local_scores=-0.0057591945,
pem=0.0, similarity=3.9178472]

gt: 52593 Addis_Ababa, pem_pos=0,
final_scores=-0.002520293, final_local_scores=-0.0057591945,
pem=0.0, similarity=3.9178472]

16: 23397137 Athletics_at_the_2009_Mediterranean_Games

[winner: 19006 Mediterranean_Sea, pem_pos=0,
final_scores=-0.060311496, final_local_scores=-0.023052245,
pem=-0.20345205, similarity=4.508192]

RECALL MISS: 23397137

Athletics_at_the_2009_Mediterranean_Games]

[winner: 27566 Summer_Olympic_Games, pem_pos=11,
final_scores=-0.62573224, final_local_scores=-0.7215982,
pem=-3.8328784, similarity=3.3995702]

RECALL MISS: 23397137

Athletics_at_the_2009_Mediterranean_Games]

17: 3517978 2009_World_Championships_in_Athletics

[winner: 218926 IAAF_World_Championships_in_Athletics,
pem_pos=0, final_scores=-0.12446442,
final_local_scores=-0.17065492, pem=-2.1703603,
similarity=10.799787]

gt: 3517978 2009_World_Championships_in_Athletics,
pem_pos=19, final_scores=-0.4850495,

final_local_scores=-0.55165195, pem=-3.947866,
 similarity=9.096429]
 18: 24547331 Athletics_at_the_2009_Jeux_de_la_Francophonie
 [winner: 1627067 Jeux_de_la_Francophonie, pem_pos=0,
 final_scores=0.19770391, final_local_scores=0.17659113,
 pem=-0.18819143, similarity=10.382958]
 gt: 24547331
 Athletics_at_the_2009_Jeux_de_la_Francophonie, pem_pos=4,
 final_scores=-0.3949516, final_local_scores=-0.40443945,
 pem=-3.0700731, similarity=8.715968]
 19: 25656607 2010_IAAF_World_Race_Walking_Cup
 [winner: 4589004 IAAF_World_Race_Walking_Cup, pem_pos=0,
 final_scores=0.08980425, final_local_scores=0.12001234,
 pem=0.0, similarity=7.6711206]
 gt: 25656607 2010_IAAF_World_Race_Walking_Cup,
 pem_pos=2, final_scores=-0.21826556,
 final_local_scores=-0.1937455, pem=-1.7917595,
 similarity=8.051827]
 [winner: 43226 2010, pem_pos=0, final_scores=-1.0002891,
 final_local_scores=-0.88706, pem=-1.6731367,
 similarity=-13.28315]
 RECALL MISS: 25656607 2010_IAAF_World_Race_Walking_Cup]
 20: 20816182 2010_African_Championships_in_Athletics
 [winner: 6278087 African_Championships_in_Athletics,
 pem_pos=0, final_scores=-0.18039934,
 final_local_scores=-0.21139252, pem=0.0,
 similarity=-2.2186613]
 RECALL MISS: 20816182
 2010_African_Championships_in_Athletics]
 [winner: 43226 2010, pem_pos=0, final_scores=-1.032774,
 final_local_scores=-0.9225449, pem=-1.6731367,
 similarity=-14.34209]
 RECALL MISS: 20816182
 2010_African_Championships_in_Athletics]
 1: 30188 Tunisia
 [pem_pos=0, final_scores=0.46598077,
 final_local_scores=0.47038442, pem=-0.39291763,
 similarity=20.26367]
 2: 4589004 IAAF_World_Race_Walking_Cup
 [pem_pos=0, final_scores=0.0991445,
 final_local_scores=0.13021517, pem=0.0, similarity=7.975593]
 3: 7074651 Hatem_Ghoula
 [pem_pos=0, final_scores=0.41000217,
 final_local_scores=0.44200116, pem=0.0, similarity=17.27991]
 1: 30222726 Hassanine_Sebei
 [pem_pos=0, final_scores=0.4509521,
 final_local_scores=0.49219233, pem=0.0, similarity=18.777714]
 2: 5334607 Africa
 [pem_pos=0, final_scores=0.07989009,
 final_local_scores=0.114433885, pem=-0.91165566,
 similarity=12.462373]
 3: 27566 Summer_Olympic_Games

[pem_pos=0, final_scores=0.14749782,
 final_local_scores=0.13684452, pem=-0.4112264,
 similarity=10.4097395]
 4: 39491575 Arab_Athletics_Championships
 [pem_pos=0, final_scores=0.04207538,
 final_local_scores=0.05195014, pem=0.0, similarity=5.640009]
 5: 1627067 Jeux_de_la_Francophonie
 [pem_pos=0, final_scores=0.24044852,
 final_local_scores=0.22328311, pem=-0.18819143,
 similarity=11.7763405]
 6: 30222726 Hassanine_Sebei
 [pem_pos=1, final_scores=0.4544893,
 final_local_scores=0.4960562, pem=0.0, similarity=18.893019]
 7: 5334607 Africa
 [pem_pos=0, final_scores=0.39575654,
 final_local_scores=0.45946997, pem=-0.91165566,
 similarity=22.75894]
 8: 3747037 Pan_Arab_Games
 [pem_pos=0, final_scores=0.05247931,
 final_local_scores=0.03196992, pem=-0.86277163,
 similarity=9.735645]
 9: 30188 Tunisia
 [pem_pos=0, final_scores=0.3539493,
 final_local_scores=0.34800705, pem=-0.39291763,
 similarity=16.611685]
 10: 1627067 Jeux_de_la_Francophonie
 [pem_pos=0, final_scores=0.19770391,
 final_local_scores=0.17659113, pem=-0.18819143,
 similarity=10.382958]
 11: 4589004 IAAF_World_Race_Walking_Cup
 [pem_pos=0, final_scores=0.08980425,
 final_local_scores=0.12001234, pem=0.0, similarity=7.6711206]

Sample: b'Andrew_Procter'

_ -LRB- cricketer -RRB- Andrew John Procter -LRB- born 7
 September 1968 -RRB- is an [1 1 English 1] 2 cricket er . [2
 Procter 2] was a right-handed who bowled right-arm 3 off break
 . He was born in [1 1 Preston 1] , [3 4 Lancashire 3] . [4
 Procter 4] first played for [2 2 Devon 2] in 1996 against [3 3
 Herefordshire 3] in the . It would be another four years
 before [5 Procter 5] represented [6 Devon 6] again ,
 returning to the county for the 5 2000 season . During this
 season he made both his [4 4 MCCA Knockout Trophy 4] and [5 5
 List A 5] debuts . Over the next seven seasons , he
 represented [7 Devon 7] in 40 Championship matches , 27
 Trophy matches , and 10 [8 List A 8] matches . In [9 List A 9]
 cricket , all of his appearances came for [10 Devon 10] at a
 time when there were permitted to take part in the domestic 6
 one-day tournament . He played his final [11 List A 11] match
 against [6 6 Essex 6] in the 1st round of the 7 2005 [12 Chel
 tenham & Gloucester Trophy 12] at [7 7 The Maer Ground 7] in [
 8 8 Exmouth 8] . In his 10 [13 List A 13] matches , he scored

114 runs at a 8 batting average of 28.50 , with a high score of 29 9 * . With the ball he took 14 wickets at a 10 bowling average

1: 88918 Preston, Lancashire
2: 3691070 Devon_County_Cricket_Club
3: 3702547 Herefordshire_County_Cricket_Club
4: 7543416 MCCA_Knockout_Trophy
5: 683459 List_A_cricket
6: 1622252 Essex_County_Cricket_Club
7: 25655931 The_Maer_Ground
8: 373409 Exmouth

1: 18803164 English_people
[winner: 9316 England, pem_pos=1,
final_scores=0.53296775, final_local_scores=0.5940222,
pem=-1.3527626, similarity=29.173056]
gt: 18803164 English_people, pem_pos=2,
final_scores=-0.45052707, final_local_scores=-0.47863504,
pem=-2.2084396, similarity=1.816126]

2: 25675557 Cricket
[winner: 330091 Rhabdophoridae, pem_pos=10,
final_scores=-1.4207075, final_local_scores=-1.3746018,
pem=-2.56456, similarity=-22.984673]
gt: 25675557 Cricket, pem_pos=0,
final_scores=-1.6660153, final_local_scores=-1.8567686,
pem=-0.07532314, similarity=-50.91035]

3: 711518 Off_break
[winner: 711518 Off_break, pem_pos=0,
final_scores=-1.3463044, final_local_scores=-1.4952886,
pem=0.0, similarity=-40.53268]
gt: 711518 Off_break, pem_pos=0,
final_scores=-1.3463044, final_local_scores=-1.4952886,
pem=0.0, similarity=-40.53268]

4: 50435 Lancashire
[winner: 749919 Lancashire_County_Cricket_Club,
pem_pos=1, final_scores=0.26863408,
final_local_scores=0.19794351, pem=-1.7295884,
similarity=19.402514]
gt: 50435 Lancashire, pem_pos=0,
final_scores=-0.033605143, final_local_scores=-0.041912302,
pem=-0.20611861, similarity=3.9598713]

5: 5672162 2000_English_cricket_season
[winner: 432383 2000_United_States_Census, pem_pos=1,
final_scores=-0.50335217, final_local_scores=-0.42671618,
pem=-1.4367632, similarity=-0.8310071]
RECALL MISS: 5672162 2000_English_cricket_season]

[winner: 1340236 Eastertide, pem_pos=2,
final_scores=-1.4612042, final_local_scores=-1.4410343,
pem=-2.3723626, similarity=-26.012346]
RECALL MISS: 5672162 2000_English_cricket_season]

6: 1899702 Friends_Provident_Trophy
[winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=-1.0707864, final_local_scores=-1.246561,

pem=0.0, similarity=-33.110157|
gt: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=-1.0707864, final_local_scores=-1.246561,
pem=0.0, similarity=-33.110157]
7: 23752831 2005_Cheltenham_&_Gloucester_Trophy
[winner: 1899702 Friends_Provident_Trophy, pem_pos=0,
final_scores=0.30414408, final_local_scores=0.26058668,
pem=0.0, similarity=11.866139|
gt: 23752831 2005_Cheltenham_&_Gloucester_Trophy,
pem_pos=2, final_scores=-0.17941944,
final_local_scores=-0.23385194, pem=-2.8258333,
similarity=12.478426]
[winner: 35984 2005, pem_pos=0, final_scores=-1.035663,
final_local_scores=-0.91515064, pem=-0.36604404,
similarity=-21.229603|
RECALL MISS: 23752831
2005_Cheltenham_&_Gloucester_Trophy]
8: 18933273 Batting_average
[winner: 15652278 Batting_(material), pem_pos=2,
final_scores=-1.0039521, final_local_scores=-0.8991238,
pem=-1.1616125, similarity=-16.424904|
gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.2823999, final_local_scores=-1.4252708,
pem=-0.29142123, similarity=-36.858414]
[winner: 875871 Bowling_average, pem_pos=1,
final_scores=-1.1894588, final_local_scores=-1.3471856,
pem=-0.8657023, similarity=-31.405169|
gt: 18933273 Batting_average, pem_pos=0,
final_scores=-1.3335139, final_local_scores=-1.4811051,
pem=-0.5012068, similarity=-37.383778]
9: 2430357 Not_out
[winner: 2430357 Not_out, pem_pos=0,
final_scores=-0.26846302, final_local_scores=-0.35621995,
pem=-0.3529179, similarity=-4.621378|
gt: 2430357 Not_out, pem_pos=0,
final_scores=-0.26846302, final_local_scores=-0.35621995,
pem=-0.3529179, similarity=-4.621378]
10: 875871 Bowling_average
[winner: 875871 Bowling_average, pem_pos=0,
final_scores=-1.0249405, final_local_scores=-1.1674743,
pem=0.0, similarity=-30.750051|
gt: 875871 Bowling_average, pem_pos=0,
final_scores=-1.0249405, final_local_scores=-1.1674743,
pem=0.0, similarity=-30.750051]
1: 88918 Preston, Lancashire
[pem_pos=0, final_scores=0.0364912,
final_local_scores=0.068051934, pem=-1.5624295,
similarity=14.617253]
2: 3691070 Devon_County_Cricket_Club
[pem_pos=1, final_scores=0.33908063,
final_local_scores=0.27672607, pem=-3.0304468,
similarity=28.827812]

3: 3702547 Herefordshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.45237905,
 final_local_scores=0.42194355, pem=-1.3582021,
 similarity=24.067465]

4: 7543416 MCCA_Knockout_Trophy
 [pem_pos=0, final_scores=0.35716945,
 final_local_scores=0.31981784, pem=0.0, similarity=13.633716]

5: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.28912753,
 final_local_scores=0.23068729, pem=-0.020267159,
 similarity=11.084098]

6: 1622252 Essex_County_Cricket_Club
 [pem_pos=1, final_scores=0.9098583,
 final_local_scores=0.8920663, pem=-2.3459642,
 similarity=43.468468]

7: 25655931 The_Maer_Ground
 [pem_pos=0, final_scores=0.27818233,
 final_local_scores=0.31732208, pem=0.0, similarity=13.559237]

8: 373409 Exmouth
 [pem_pos=0, final_scores=0.13916562,
 final_local_scores=0.18848029, pem=-0.42244935,
 similarity=12.011685]

1: 9316 England
 [pem_pos=1, final_scores=0.53296775,
 final_local_scores=0.5940222, pem=-1.3527626,
 similarity=29.173056]

2: 1479445 Mike_Procter
 [pem_pos=3, final_scores=0.49403363,
 final_local_scores=0.51506907, pem=-1.9902581,
 similarity=30.283733]

3: 749919 Lancashire_County_Cricket_Club
 [pem_pos=1, final_scores=0.26863408,
 final_local_scores=0.19794351, pem=-1.7295884,
 similarity=19.402514]

4: 1479445 Mike_Procter
 [pem_pos=3, final_scores=0.6554588,
 final_local_scores=0.69140154, pem=-1.9902581,
 similarity=35.545845]

5: 1479445 Mike_Procter
 [pem_pos=3, final_scores=0.61526775,
 final_local_scores=0.6474989, pem=-1.9902581,
 similarity=34.235703]

6: 3691070 Devon_County_Cricket_Club
 [pem_pos=1, final_scores=0.13017486,
 final_local_scores=0.048528254, pem=-3.0304468,
 similarity=22.017935]

7: 3691070 Devon_County_Cricket_Club
 [pem_pos=1, final_scores=0.5044816,
 final_local_scores=0.4574015, pem=-3.0304468,
 similarity=34.21953]

8: 683459 List_A_cricket
 [pem_pos=0, final_scores=0.3845393,

```

final_local_scores=0.33491018, pem=-0.020267159,
similarity=14.194317]
9: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.038370594,
final_local_scores=-0.047286294, pem=-0.020267159,
similarity=2.7888117]
10: 3691070 Devon_County_Cricket_Club
    [pem_pos=1, final_scores=0.3194046,
final_local_scores=0.255233, pem=-3.0304468,
similarity=28.186417]
11: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.23552857,
final_local_scores=0.17213863, pem=-0.020267159,
similarity=9.336889]
12: 1899702 Friends_Provident_Trophy
    [pem_pos=0, final_scores=0.30414408,
final_local_scores=0.26058668, pem=0.0, similarity=11.866139]
13: 683459 List_A_cricket
    [pem_pos=0, final_scores=0.09125753,
final_local_scores=0.0145444125, pem=-0.020267159,
similarity=4.633963]

```

Sample: b'Sturgeon_House'

The Sturgeon House is a 1 saltbox house dating from around 1838 in 2 Fairview , [1 3 Erie County 1] in the [2 4 U.S. 2] s tate of [1 1 Pennsylvania 1] . It was listed on the [2 2 Nati onal Register of Historic Places 2] in 1980 . The Sturgeon House is operated as 5 museum by the Fairview Area Historical Society . The Sturgeon House is located at the intersection of Water Street and Avonia Road -LRB- 6 Pennsylvania Route 98 -RRB- . The house is a rare example of a 7 saltbox house in 8 northwestern [3 Pennsylvania 3] , as well as a house with a recessed side 9 porch . The house is constructed of `` stone foundation walls and topped with heavy timber 10 sills which are notched to receive log form beams . '' The roof of the house is slanted at an angle of 30 degrees . The front door is framed by a 11 cornice and 12 Federal - style 13 entablature and 14 pilaster . Jeremiah and William Sturgeon were the earliest settlers of what is now 15 Fairview Township , who acquired land from the in 1797 . The Sturgeons operated '' 16 coach stops and taverns '' in the area for travelers after the first road was cut from [4 Erie 4] to [3 3 Cleveland 3] , [4 4 Ohio 4] in 1805 . They founded the community of [5 Sturgeonville 5] , which eventually became 17 Fairview . The house was built around 1838 by Samuel C. Sturgeon and is one of several built by the Sturgeons . The southern addition of the house was possibly constructed in the early 1850s . The Sturgeon House was sold to the Fairview Area Historical Society in 1979 by a descendant of the Jeremiah Sturgeon . The house was listed on the [5 5 National Register of Historic Places 5]

1: 23332 Pennsylvania

2: 64065 National_Register_of_Historic_Places
 3: 5951 Cleveland
 4: 22199 Ohio
 5: 64065 National_Register_of_Historic_Places
 1: 2104881 Saltbox
 [winner: 2104881 Saltbox, pem_pos=0,
 final_scores=-0.8201825, final_local_scores=-0.87930983,
 pem=0.0, similarity=-22.150646]
 gt: 2104881 Saltbox, pem_pos=0, final_scores=-0.8201825,
 final_local_scores=-0.87930983, pem=0.0,
 similarity=-22.150646]
 2: 12799960 Fairview, Erie County, Pennsylvania
 [winner: 135693 Fairview, Texas, pem_pos=11,
 final_scores=-0.03555061, final_local_scores=0.011742532,
 pem=-3.3622322, similarity=22.724474]
 RECALL MISS: 12799960
 Fairview, Erie County, Pennsylvania]
 3: 91960 Erie County, Pennsylvania
 [winner: 54169 Erie County, New York, pem_pos=0,
 final_scores=0.28622168, final_local_scores=0.26497668,
 pem=-0.47740197, similarity=14.593332]
 gt: 91960 Erie County, Pennsylvania, pem_pos=1,
 final_scores=0.25743264, final_local_scores=0.22769874,
 pem=-0.71024585, similarity=14.747125]
 4: 18618239 U.S. state
 [winner: 3434750 United States, pem_pos=0,
 final_scores=0.3915288, final_local_scores=0.39363343,
 pem=-0.14075686, similarity=16.601978]
 gt: 18618239 U.S. state, pem_pos=6,
 final_scores=-0.40787917, final_local_scores=-0.5042324,
 pem=-3.9799929, similarity=10.686233]
 [winner: 229379 States of Brazil, pem_pos=0,
 final_scores=-1.7187102, final_local_scores=-1.7569337,
 pem=0.0, similarity=-48.340694]
 RECALL MISS: 18618239 U.S. state]
 5: 9650166 Historic_house_museum
 [winner: 37585 Museum, pem_pos=0,
 final_scores=-0.87679213, final_local_scores=-0.92368406,
 pem=-0.115303785, similarity=-22.847822]
 gt: 9650166 Historic_house_museum, pem_pos=7,
 final_scores=-1.2773294, final_local_scores=-1.379664,
 pem=-3.8846362, similarity=-15.956957]
 6: 7124637 Pennsylvania_Route_98
 [winner: 7124637 Pennsylvania_Route_98, pem_pos=0,
 final_scores=-0.30499288, final_local_scores=-0.23456869,
 pem=0.0, similarity=-2.9102848]
 gt: 7124637 Pennsylvania_Route_98, pem_pos=0,
 final_scores=-0.30499288, final_local_scores=-0.23456869,
 pem=0.0, similarity=-2.9102848]
 7: 2104881 Saltbox
 [winner: 2104881 Saltbox, pem_pos=0,
 final_scores=-0.7370529, final_local_scores=-0.7885033,

pem=0.0, similarity=-19.440798|
 gt: 2104881 Saltbox, pem_pos=0, final_scores=-0.7370529,
 final_local_scores=-0.7885033, pem=0.0,
 similarity=-19.440798]
 8: 1608659 Northwestern_Pennsylvania
 [winner: 23332 Pennsylvania, pem_pos=0,
 final_scores=0.49914604, final_local_scores=0.45007622,
 pem=-0.11810702, similarity=18.16317|
 RECALL MISS: 1608659 Northwestern_Pennsylvania]
 [winner: 23211731 Northwestern_California, pem_pos=12,
 final_scores=-1.2620797, final_local_scores=-1.2654555,
 pem=-3.3955035, similarity=-15.208726|
 gt: 1608659 Northwestern_Pennsylvania, pem_pos=17,
 final_scores=-1.6997387, final_local_scores=-1.8470013,
 pem=-3.5936375, similarity=-31.485731]
 9: 1251142 Porch
 [winner: 9245549 Vestibule_(architecture), pem_pos=0,
 final_scores=-1.0822508, final_local_scores=-1.1087124,
 pem=0.0, similarity=-28.996477|
 gt: 1251142 Porch, pem_pos=1, final_scores=-1.1024108,
 final_local_scores=-1.1306605, pem=-0.020505866,
 similarity=-29.539942]
 10: 8907461 Sill_plate
 [winner: 1105295 Sill_(geology), pem_pos=0,
 final_scores=-0.90341544, final_local_scores=-0.8533344,
 pem=0.0, similarity=-21.375488|
 gt: 8907461 Sill_plate, pem_pos=1,
 final_scores=-1.2045538, final_local_scores=-1.1753793,
 pem=-2.7994678, similarity=-15.762011]
 11: 358820 Cornice
 [winner: 358820 Cornice, pem_pos=0,
 final_scores=-1.0707984, final_local_scores=-1.0915532,
 pem=-0.06297184, similarity=-28.14196|
 gt: 358820 Cornice, pem_pos=0, final_scores=-1.0707984,
 final_local_scores=-1.0915532, pem=-0.06297184,
 similarity=-28.14196]
 12: 1691932 Federal_architecture
 [winner: 195149 Federal_government_of_the_United_States,
 pem_pos=1, final_scores=-0.035345316,
 final_local_scores=-0.023592949, pem=-1.574776,
 similarity=11.94953|
 gt: 1691932 Federal_architecture, pem_pos=0,
 final_scores=-0.2016821, final_local_scores=-0.19331148,
 pem=-1.2528081, similarity=5.133878]
 13: 613749 Entablature
 [winner: 613749 Entablature, pem_pos=0,
 final_scores=-0.9592475, final_local_scores=-0.9175313,
 pem=0.0, similarity=-23.291252|
 gt: 613749 Entablature, pem_pos=0,
 final_scores=-0.9592475, final_local_scores=-0.9175313,
 pem=0.0, similarity=-23.291252]
 14: 145380 Pilaster

[winner: 145380 Pilaster, pem_pos=0,
 final_scores=-0.96090937, final_local_scores=-0.942283,
 pem=0.0, similarity=-24.02989]
 gt: 145380 Pilaster, pem_pos=0,
 final_scores=-0.96090937, final_local_scores=-0.942283,
 pem=0.0, similarity=-24.02989]
 15: 132483 Fairview_Township,_Erie_County,_Pennsylvania
 [winner: 132483
 Fairview_Township,_Erie_County,_Pennsylvania, pem_pos=0,
 final_scores=0.007962421, final_local_scores=0.01611057,
 pem=-1.4839778, similarity=12.640587]
 gt: 132483 Fairview_Township,_Erie_County,_Pennsylvania,
 pem_pos=0, final_scores=0.007962421,
 final_local_scores=0.01611057, pem=-1.4839778,
 similarity=12.640587]
 16: 214764 Stagecoach
 [winner: 13640744 Coach_(carriage), pem_pos=7,
 final_scores=-1.4113748, final_local_scores=-1.4656451,
 pem=-2.182298, similarity=-27.780386]
 gt: 214764 Stagecoach, pem_pos=11,
 final_scores=-1.9190015, final_local_scores=-2.098914,
 pem=-4.036971, similarity=-36.59239]
 17: 12799960 Fairview,_Erie_County,_Pennsylvania
 [winner: 136881 Fairview,_Bergen_County,_New_Jersey,
 pem_pos=0, final_scores=-0.22797976,
 final_local_scores=-0.2554146, pem=-2.7451415,
 similarity=11.396139]
 RECALL MISS: 12799960
 Fairview,_Erie_County,_Pennsylvania]
 1: 23332 Pennsylvania
 [pem_pos=0, final_scores=0.47650367,
 final_local_scores=0.42534292, pem=-0.11810702,
 similarity=17.425077]
 2: 64065 National_Register_of_Historic_Places
 [pem_pos=0, final_scores=0.38782734,
 final_local_scores=0.3840189, pem=0.0, similarity=15.5496025]
 3: 5951 Cleveland
 [pem_pos=0, final_scores=0.58254856,
 final_local_scores=0.58425325, pem=-0.57246673,
 similarity=24.638159]
 4: 22199 Ohio
 [pem_pos=0, final_scores=0.64289165,
 final_local_scores=0.608841, pem=-0.19147073,
 similarity=23.29999]
 5: 64065 National_Register_of_Historic_Places
 [pem_pos=0, final_scores=0.27821398,
 final_local_scores=0.26428297, pem=0.0, similarity=11.976444]
 1: 54169 Erie_County,_New_York
 [pem_pos=0, final_scores=0.28622168,
 final_local_scores=0.26497668, pem=-0.47740197,
 similarity=14.593332]
 2: 3434750 United_States

[pem_pos=0, final_scores=0.3915288,
 final_local_scores=0.39363343, pem=-0.14075686,
 similarity=16.601978]
 3: 23332 Pennsylvania
 [pem_pos=0, final_scores=0.49914604,
 final_local_scores=0.45007622, pem=-0.11810702,
 similarity=18.16317]
 4: 132482 Erie, Pennsylvania
 [pem_pos=0, final_scores=0.2885973,
 final_local_scores=0.25011274, pem=-1.0382025,
 similarity=17.199482]
 5: 30084004 Sturgeonville, Virginia
 [pem_pos=0, final_scores=0.037054792,
 final_local_scores=0.13636029, pem=0.0, similarity=8.158976]

Sample: b'Ubaoner'

Ubaoner is the name of a 1 fictitious 2 ancient [1 Egypt 1]
 ian 3 magician appearing in the second chapter of a story
 told in the legendary [1 1 Westcar Papyrus 1] . He is said to
 have worked 4 wonders during the reign of king -LRB- 5 pharaoh
 -RRB- 6 Nebka -LRB- -RRB- . Ubaoner appears only in the
 second story of the [2 Westcar Papyrus 2] -- there is no or 7
 historical evidence that he existed . Nevertheless he is an
 object of great interest for , since his magical trick is
 connected to later cultural perceptions of the personality of
 king Nebka . In the story , [3 Ubaoner 3] is described as a
 high lector priest . According to the [4 Westcar Papyrus 4] ,
 Ubaoner is betrayed by his own wife several times . She
 secretly loves a commoner who resides at [5 8 Memphis 5] and
 sends him expensive clothes and other material . The commoner
 comes to visit [6 Ubaoner 6] 's wife and says to her : `` Is
 n't there a 9 pavilion at the lake ? '' The wife says : ``
 See , let 's spend some time in there . '' She calls for [7
 Ubaoner 7] 's 10 caretaker and orders him : `` Prepare the
 pavilion with all good things . '' And so Ubaoner 's wife and
 the commoner spend their time in the pavilion drinking
 together , until they were finished . Later that evening the
 commoner is picked up by his maidservant , who was waiting at
 the lake . [8 Ubaoner 8] 's caretaker was waiting there , too
 and he says : `` I will go to [9 Ubaoner 9] . '' Two days
 later the caretaker tells everything to Ubaoner and makes the
 comment : `` See , it was your wife who spent a day in your
 pavilion ; she was together with that commoner you know . It
 was 11 adultery what he did , two times , at your lake . ''
 Thereupon , Ubaoner says to the caretaker : `` Bring me a
 pair of -LSB- ... -RSB- made of 12 ebony and 13 dja'am . I
 will create -LSB- ... -RSB- and send it as a harbinger . ''
 He creates a small figurine of 14 wax in the shape of a 15 cr
 ocodile with a length of seven fingers . He bewitches the
 figurine and gives it to his caretaker with the words : ``
 When he comes to purify himself in the lake , this commoner ,
 then you may throw this crocodile made of wax after him . ''

The caretaker does as ordered and waits until the next day . The next morning Ubaoner 's wife orders the caretaker : `` Let the pavilion at the lake be prepared , for I will sit in it . '' Thereupon she spends the whole day in the pavilion together with the commoner . Later that evening the commoner leaves the pavilion to take a bath in the lake . The caretaker quickly throws the wax figurine after the stranger and when the idol touches the water it becomes a real , living crocodile 3.70 m in length . The animal devours the commoner and disappears in the depths of the lake for seven days . During these seven days [10 Ubaoner 10] is received by pharaoh Nebka for an important audience . After the audience Ubaoner invites Nebka to visit his house with the words : `` May thy majesty proceed and see the wonder that has happened in the time of thy majesty -LSB- ... -RSB- a commoner . '' . Nebka and [11 Ubaoner 11] walk to the lake where [12 Ubaoner 1 2] orders the crocodile to come out of the water and to release the commoner . When king Nebka sees that he says : `` This crocodile is dangerous ! '' But [13 Ubaoner 13] bends down and touches the crocodile and immediately it becomes a figurine of wax again . Then [14 Ubaoner 14] gives a report to Nebka about the affairs . Nebka tells the crocodile : `` Take away what is yours ! '' and the animal grabs the commoner and then disappears . The wife of Ubaoner is brought to Nebka , too and the pharaoh sentence her to death . She is brought to a place east of the palace and burnt alive . Her ash is thrown into the [2 2 Nile 2] . [3 3 Adolf Erman 3] and [4 4 Kurt Heinrich Sethe 4] once considered the stories of [15 Papyrus Westcar 15] as mere 16 folklore . They saw literary figures like 17 Djadjaemankh and the other heroes of the [16 Westcar Papyrus 16] as a pure fiction , created only for entertainment , since there is no archeological evidence of them . Modern Egyptologists like and [5 5 Miriam Lichtheim 5] deny this view and argue , that Sethe and Erman may have just failed to see the profundity of such novels . They hold that the story of Ubaoner describes a classical case of 18 adultery through the wife and a typical sort of punishment for that during the [6 6 Old Kingdom 6] -LRB- in this case it is the -RRB- . Furthermore it throws a positive light on the personality of king Nebka . The king is depicted as a strict but lawful 19 judge : He punishes mischief and 20 unethical

1: 1507443 Westcar_Papyrus

2: 21244 Nile

3: 1693558 Adolf_Erman

4: 5818971 Kurt_Heinrich_Sethe

5: 13198062 Miriam_Lichtheim

6: 152500 Old_Kingdom_of_Egypt

1: 18949461 Fiction

[winner: 1082841 Fictitious_force, pem_pos=4, final_scores=-1.3803707, final_local_scores=-1.3757179, pem=-1.828317, similarity=-27.021782]

gt: 18949461 Fiction, pem_pos=2,

final_scores=-1.6214383, final_local_scores=-1.7129467,
pem=-0.76949304, similarity=-42.843407]
2: 874 Ancient_Egypt
[winner: 8087628 Egypt, pem_pos=0,
final_scores=0.31294876, final_local_scores=0.36989528,
pem=-0.27480945, similarity=16.622583]
gt: 874 Ancient_Egypt, pem_pos=1,
final_scores=-0.34318107, final_local_scores=-0.3945648,
pem=-2.924479, similarity=8.2188835]
[winner: 1064603 Ancient_woodland, pem_pos=5,
final_scores=-1.5962788, final_local_scores=-1.6144013,
pem=-2.0350637, similarity=-33.02025]
gt: 874 Ancient_Egypt, pem_pos=3,
final_scores=-1.689358, final_local_scores=-1.8735754,
pem=-1.8627092, similarity=-41.691814]
3: 53223 Magician_(paranormal)
[winner: 2203198 Magician_(video_game), pem_pos=3,
final_scores=-0.99265766, final_local_scores=-0.95469373,
pem=-1.3853766, similarity=-16.86636]
gt: 53223 Magician_(paranormal), pem_pos=4,
final_scores=-1.2032564, final_local_scores=-1.2963183,
pem=-2.012839, similarity=-23.648878]
4: 52041 Miracle
[winner: 6260914 Joe_Arpaio, pem_pos=0,
final_scores=-1.1125318, final_local_scores=-1.0631449,
pem=0.0, similarity=-27.63665]
RECALL MISS: 52041 Miracle]
5: 23294 Pharaoh
[winner: 23294 Pharaoh, pem_pos=0,
final_scores=-0.88579583, final_local_scores=-1.0000625,
pem=-0.021238824, similarity=-25.638647]
gt: 23294 Pharaoh, pem_pos=0, final_scores=-0.88579583,
final_local_scores=-1.0000625, pem=-0.021238824,
similarity=-25.638647]
6: 8397233 Nebka
[winner: 1098346 Sanakht, pem_pos=0,
final_scores=-0.11810902, final_local_scores=-0.06735392,
pem=-0.06301298, similarity=2.4224122]
gt: 8397233 Nebka, pem_pos=1, final_scores=-0.4844099,
final_local_scores=-0.4895115, pem=-2.7957559,
similarity=4.685469]
7: 10772350 History
[winner: 8643851 English_historical_school_of_economics,
pem_pos=13, final_scores=-1.5918514,
final_local_scores=-1.5837939, pem=-2.7725887,
similarity=-28.096087]
gt: 10772350 History, pem_pos=0,
final_scores=-1.8513086, final_local_scores=-1.9686105,
pem=-0.5710359, similarity=-51.55217]
8: 96603 Memphis,_Egypt
[winner: 48607 Memphis,_Tennessee, pem_pos=0,
final_scores=0.19981487, final_local_scores=0.35589328,

pem=-0.64647895, similarity=18.225931|
gt: 96603 Memphis_Egypt, pem_pos=1,
final_scores=-0.15401907, final_local_scores=-0.17169186,
pem=-2.1332347, similarity=10.566948]
9: 1358820 Pavilion
[winner: 1358820 Pavilion, pem_pos=0,
final_scores=-1.2057182, final_local_scores=-1.2275345,
pem=0.0, similarity=-32.542366|
gt: 1358820 Pavilion, pem_pos=0,
final_scores=-1.2057182, final_local_scores=-1.2275345,
pem=0.0, similarity=-32.542366]
10: 10621039 Property_caretaker
[winner: 10621039 Property_caretaker, pem_pos=0,
final_scores=-1.3655014, final_local_scores=-1.358144,
pem=-0.90955675, similarity=-31.493704|
gt: 10621039 Property_caretaker, pem_pos=0,
final_scores=-1.3655014, final_local_scores=-1.358144,
pem=-0.90955675, similarity=-31.493704]
11: 58809 Adultery
[winner: 23892113 Thou_shalt_not_commit_adultery,
pem_pos=0, final_scores=-0.8635929,
final_local_scores=-0.8785066, pem=0.0,
similarity=-22.126677|
gt: 58809 Adultery, pem_pos=1, final_scores=-1.1908606,
final_local_scores=-1.2535396, pem=-0.016930522,
similarity=-33.22634]
12: 184073 Ebony
[winner: 389489 Ebony_(magazine), pem_pos=0,
final_scores=-0.8697574, final_local_scores=-0.82698905,
pem=-0.6426758, similarity=-17.094318|
gt: 184073 Ebony, pem_pos=1, final_scores=-0.9308372,
final_local_scores=-0.92842466, pem=-0.8221297,
similarity=-19.14546]
13: 56226 Electrum
[winner: 20063724 Colored_gold, pem_pos=0,
final_scores=-1.2341987, final_local_scores=-1.2469699,
pem=-0.6931472, similarity=-29.352917|
gt: 56226 Electrum, pem_pos=1, final_scores=-1.3617266,
final_local_scores=-1.4054439, pem=-0.6931472,
similarity=-34.082092]
14: 58256 Wax
[winner: 596482 Hair_wax, pem_pos=1,
final_scores=-0.6230945, final_local_scores=-0.5015354,
pem=-1.0628463, similarity=-5.097182|
gt: 58256 Wax, pem_pos=0, final_scores=-0.62625164,
final_local_scores=-0.595143, pem=-0.09161498,
similarity=-13.172325]
15: 1284973 Nile_crocodile
[winner: 2025062 Crocodylus, pem_pos=3,
final_scores=-1.3115054, final_local_scores=-1.3012137,
pem=-1.381939, similarity=-27.225899|
gt: 1284973 Nile_crocodile, pem_pos=1,

final_scores=-1.4442213, final_local_scores=-1.4824469,
pem=-1.3402308, similarity=-32.861076]
16: 11303 Folklore
[winner: 18619209 Bloody_Mary_(folklore), pem_pos=0,
final_scores=-0.79873186, final_local_scores=-0.7418441,
pem=0.0, similarity=-18.048395]
gt: 11303 Folklore, pem_pos=1, final_scores=-1.1581517,
final_local_scores=-1.2429924, pem=-0.09392009,
similarity=-32.49291]
17: 31895210 Djadjaemankh
[winner: 31895210 Djadjaemankh, pem_pos=0,
final_scores=-0.007412553, final_local_scores=0.02384238,
pem=0.0, similarity=4.801217]
gt: 31895210 Djadjaemankh, pem_pos=0,
final_scores=-0.007412553, final_local_scores=0.02384238,
pem=0.0, similarity=4.801217]
18: 58809 Adultery
[winner: 23892113 Thou_shalt_not_commit_adultery,
pem_pos=0, final_scores=-0.8243317,
final_local_scores=-0.8356196, pem=0.0,
similarity=-20.846844]
gt: 58809 Adultery, pem_pos=1, final_scores=-1.1132607,
final_local_scores=-1.1687734, pem=-0.016930522,
similarity=-30.696747]
19: 44719 Judge
[winner: 6061994
The_Honourable_Mr_Justice_Stephen_Rawley, pem_pos=7,
final_scores=-1.2344292, final_local_scores=-1.2071092,
pem=-2.6123815, similarity=-17.726294]
gt: 44719 Judge, pem_pos=0, final_scores=-1.3639724,
final_local_scores=-1.3843722, pem=-0.11628761,
similarity=-36.590324]
20: 9258 Ethics
[winner: 18934464 Embrace_extend_and_extinguish,
pem_pos=1, final_scores=-1.2578582,
final_local_scores=-1.1886902, pem=-1.716497,
similarity=-22.048595]
gt: 9258 Ethics, pem_pos=0, final_scores=-1.3015271,
final_local_scores=-1.3481691, pem=-0.43278924,
similarity=-33.78877]
1: 1507443 Westcar_Papyrus
[pem_pos=0, final_scores=0.1433826,
final_local_scores=0.0826181, pem=0.0, similarity=6.555202]
2: 21244 Nile
[pem_pos=0, final_scores=0.3640442,
final_local_scores=0.41799903, pem=-0.08206034,
similarity=17.009893]
3: 1693558 Adolf_Erman
[pem_pos=0, final_scores=0.24955778,
final_local_scores=0.31355312, pem=0.0, similarity=13.446764]
4: 5818971 Kurt_Heinrich_Sethe
[pem_pos=0, final_scores=0.12865861,

final_local_scores=0.18338889, pem=0.0, similarity=9.562403]
5: 13198062 Miriam_Lichtheim
[pem_pos=0, final_scores=0.14098446,
final_local_scores=0.21430722, pem=0.0, similarity=10.485067]
6: 152500 Old_Kingdom_of_Egypt
[pem_pos=0, final_scores=0.13950907,
final_local_scores=0.14439684, pem=0.0, similarity=8.398803]
1: 8087628 Egypt
[pem_pos=0, final_scores=0.31294876,
final_local_scores=0.36989528, pem=-0.27480945,
similarity=16.622583]
2: 1507443 Westcar_Papyrus
[pem_pos=0, final_scores=0.08955698,
final_local_scores=0.023821786, pem=0.0, similarity=4.8006024]
3: 31896730 Ubaoner
[pem_pos=0, final_scores=0.05606623,
final_local_scores=-0.002143383, pem=0.0, similarity=4.02575]
4: 1507443 Westcar_Papyrus
[pem_pos=0, final_scores=0.14809023,
final_local_scores=0.08776048, pem=0.0, similarity=6.7086606]
5: 48607 Memphis_Tennessee
[pem_pos=0, final_scores=0.19981487,
final_local_scores=0.35589328, pem=-0.64647895,
similarity=18.225931]
6: 31896730 Ubaoner
[pem_pos=0, final_scores=0.08955066,
final_local_scores=0.043064207, pem=0.0, similarity=5.3748345]
7: 31896730 Ubaoner
[pem_pos=0, final_scores=0.04873453,
final_local_scores=-0.010152146, pem=0.0,
similarity=3.7867527]
8: 31896730 Ubaoner
[pem_pos=0, final_scores=0.0386806,
final_local_scores=-0.021134548, pem=0.0,
similarity=3.4590158]
9: 31896730 Ubaoner
[pem_pos=0, final_scores=0.07461737,
final_local_scores=0.026751876, pem=0.0, similarity=4.8880424]
10: 31896730 Ubaoner
[pem_pos=0, final_scores=0.08457787,
final_local_scores=0.037632212, pem=0.0, similarity=5.2127333]
11: 31896730 Ubaoner
[pem_pos=0, final_scores=0.23063742,
final_local_scores=0.19718006, pem=0.0, similarity=9.973959]
12: 31896730 Ubaoner
[pem_pos=0, final_scores=0.05424525,
final_local_scores=0.004498422, pem=0.0, similarity=4.2239547]
13: 31896730 Ubaoner
[pem_pos=0, final_scores=0.17364763,
final_local_scores=0.13492739, pem=0.0, similarity=8.116216]
14: 31896730 Ubaoner
[pem_pos=0, final_scores=0.03662367,

final_local_scores=-0.023381434, pem=0.0,
similarity=3.3919642]
15: 1507443 Westcar_Papyrus
[pem_pos=0, final_scores=0.24064021,
final_local_scores=0.18885726, pem=0.0, similarity=9.72559]
16: 1507443 Westcar_Papyrus
[pem_pos=0, final_scores=0.23479794,
final_local_scores=0.18247548, pem=0.0, similarity=9.535145]

Sample: b'Daniel_Levy'

_ -LRB- politician -RRB- Sir [1 Daniel Levy 1] -LRB- 30
November 1872 -- 30 May 1937 -RRB- was an [2 Australian 2]
politician . He was born in [1 1 London 1] to tailor [3
Joseph Levy 3] and Esther , née Cohen . He arrived in [2 2 Sy
dney 2] in 1880 and attended Crown Street Superior School , [3
3 Sydney Grammar School 3] and the [4 4 University of Sydney 4
] , graduating with a first in [4 1 Bachelor of Arts 4] -LRB-
Honours -RRB- in 1893 and a in 1895 , in which year he was
called to the Bar . In 1902 he was admitted as a solicitor .
He entered the [5 5 New South Wales Legislative Assembly 5]
in 1901 as the 2 Liberal member for 3 Sydney-Fitzroy ,
transferring to [5 4 Darlinghurst 5] in 1904 -LRB- he would
represent [6 5 Sydney 6] for the period of 6 proportional rep
resentation from 1920 to 1927 , 7 Paddington from 1927 to
1930 , and 8 Woollahra thereafter -RRB- . In 1919 , Levy ,
despite being a 9 Nationalist , accepted the position of 10 Sp
eaker of the Legislative Assembly from 11 James Dooley 's [6 6
Labor 6] government , provoking discontent within his own
party . He resigned with the fall of the government in 1920 ,
was re-appointed by [7 7 George Fuller 7] for his seven hours
as premier , and continued the position when [7 Dooley 7]
resumed the premiership later the same day . Becoming known
as a scrupulously independent Speaker , Levy advocated the [8
British 8] model of Speakership in which the Speaker 's seat
was uncontested , and made efforts to have this practice
adopted by statute . He was Speaker until 1925 , served again
from 1927 to 1930 , and served briefly in [8 8 Bertram Stevens
8] ' 1932 emergency cabinet as Minister for Justice and
Attorney-General . Resuming the Speakership in 1932 , he held
the position until his death in 1937 at [9 9 Darling Point 9]

1: 17867 London

2: 27862 Sydney

3: 417362 Sydney_Grammar_School

4: 31928 University_of_Sydney

5: 579433 New_South_Wales_Legislative_Assembly

6: 1495 Australian_Labor_Party

7: 3150155 George_Fuller_(Australian_politician)

8: 2770934 Bertram_Stevens_(politician)

9: 2648858 Darling_Point,_New_South_Wales

1: 156992 Bachelor_of_Arts

[winner: 156992 Bachelor_of_Arts, pem_pos=0,
final_scores=0.075702295, final_local_scores=0.092357695,


```

pem=0.0, similarity=6.845851|
    gt: 156992 Bachelor_of_Arts, pem_pos=0,
final_scores=0.075702295, final_local_scores=0.092357695,
pem=0.0, similarity=6.845851 ]
[winner: 460615
British_undergraduate_degree_classification, pem_pos=2,
final_scores=-0.37933785, final_local_scores=-0.41559872,
pem=-1.3862944, similarity=-0.77369773|
    RECALL MISS: 156992 Bachelor_of_Arts]
2: 4352739 Liberal_Reform_Party_(Australia)
    [winner: 149536 Liberal_Party_of_Canada, pem_pos=0,
final_scores=-0.002326578, final_local_scores=-0.0087222755,
pem=-1.2222316, similarity=10.476109|
    RECALL MISS: 4352739 Liberal_Reform_Party_(Australia)]
3: 10807575 Electoral_district_of_Sydney-Fitzroy
    [winner: 10807575 Electoral_district_of_Sydney-Fitzroy,
pem_pos=0, final_scores=-0.02846782,
final_local_scores=0.049302638, pem=0.0, similarity=5.561002|
    gt: 10807575 Electoral_district_of_Sydney-Fitzroy,
pem_pos=0, final_scores=-0.02846782,
final_local_scores=0.049302638, pem=0.0, similarity=5.561002 ]
4: 10829395 Electoral_district_of_Darlinghurst
    [winner: 240037 Darlinghurst, New_South_Wales, pem_pos=0,
final_scores=0.07939704, final_local_scores=0.0510509,
pem=0.0, similarity=5.6131735|
    gt: 10829395 Electoral_district_of_Darlinghurst,
pem_pos=1, final_scores=-0.16329733,
final_local_scores=-0.10158976, pem=-0.35342422,
similarity=2.9800477 ]
5: 8856503 Electoral_district_of_Sydney
    [winner: 27862 Sydney, pem_pos=0,
final_scores=0.19450109, final_local_scores=0.15110463,
pem=-0.1443489, similarity=9.383968|
    gt: 8856503 Electoral_district_of_Sydney, pem_pos=11,
final_scores=-1.0128344, final_local_scores=-1.0941737,
pem=-5.191942, similarity=-0.32802296 ]
6: 23298 Proportional_representation
    [winner: 4415141 Wold's_theorem, pem_pos=8,
final_scores=-1.0943005, final_local_scores=-0.9772911,
pem=-2.0768921, similarity=-13.780141|
    gt: 23298 Proportional_representation, pem_pos=6,
final_scores=-1.5483042, final_local_scores=-1.6588838,
pem=-2.0617309, similarity=-34.202675 ]
    [winner: 81863 Proportionality_(mathematics), pem_pos=0,
final_scores=-1.1018376, final_local_scores=-1.0390408,
pem=0.0, similarity=-26.917337|
    gt: 23298 Proportional_representation, pem_pos=1,
final_scores=-1.1578867, final_local_scores=-1.2324121,
pem=-0.46681154, similarity=-30.14933 ]
7: 10680241 Electoral_district_of_Paddington_(New_South_Wales)
    [winner: 94211 Paddington, pem_pos=0,
final_scores=-0.03643675, final_local_scores=-0.07246575,

```

pem=-0.5713622, similarity=5.034347|
gt: 10680241
Electoral_district_of_Paddington_(New_South_Wales),
pem_pos=4, final_scores=-0.51027894,
final_local_scores=-0.5230781, pem=-2.0735204,
similarity=-0.24385414]
8: 4632947 Electoral_district_of_Woollahra
[winner: 1477064 Woollahra, New_South_Wales, pem_pos=0,
final_scores=-0.016857669, final_local_scores=-0.03853377,
pem=-0.23478514, similarity=4.2165866|
gt: 4632947 Electoral_district_of_Woollahra, pem_pos=2,
final_scores=-0.36965835, final_local_scores=-0.35121682,
pem=-1.113162, similarity=-0.3377494]
9: 52536 Nationalist_Party_of_Australia
[winner: 52536 Nationalist_Party_of_Australia, pem_pos=0,
final_scores=-0.05714996, final_local_scores=-0.08958705,
pem=-0.7300673, similarity=5.3864765|
gt: 52536 Nationalist_Party_of_Australia, pem_pos=0,
final_scores=-0.05714996, final_local_scores=-0.08958705,
pem=-0.7300673, similarity=5.3864765]
10: 11845332
Speaker_of_the_New_South_Wales_Legislative_Assembly
[winner: 579433 New_South_Wales_Legislative_Assembly,
pem_pos=1, final_scores=-0.24286573,
final_local_scores=-0.32266888, pem=-1.9991245,
similarity=5.33218|
RECALL MISS: 11845332
Speaker_of_the_New_South_Wales_Legislative_Assembly]
[winner: 4047486 Speaker_of_the_Dewan_Rakyat, pem_pos=1,
final_scores=-0.9008533, final_local_scores=-0.8484333,
pem=0.0, similarity=-21.229229|
RECALL MISS: 11845332
Speaker_of_the_New_South_Wales_Legislative_Assembly]
11: 3150165 James_Dooley_(politician)
[winner: 3150165 James_Dooley_(politician), pem_pos=1,
final_scores=0.02225785, final_local_scores=0.01756835,
pem=-0.77538526, similarity=8.830653|
gt: 3150165 James_Dooley_(politician), pem_pos=1,
final_scores=0.02225785, final_local_scores=0.01756835,
pem=-0.77538526, similarity=8.830653]
1: 17867 London
[pem_pos=0, final_scores=0.5689643,
final_local_scores=0.57544523, pem=-0.09086737,
similarity=21.756296]
2: 27862 Sydney
[pem_pos=0, final_scores=0.41384786,
final_local_scores=0.3907076, pem=-0.1443489,
similarity=16.5342]
3: 417362 Sydney_Grammar_School
[pem_pos=0, final_scores=0.2933455,
final_local_scores=0.27019432, pem=0.0, similarity=12.15285]
4: 31928 University_of_Sydney

[pem_pos=0, final_scores=0.37347585,
 final_local_scores=0.35325742, pem=0.0, similarity=14.6316185]
 5: 579433 New_South_Wales_Legislative_Assembly
 [pem_pos=0, final_scores=0.21410035,
 final_local_scores=0.18608713, pem=0.0, similarity=9.642924]
 6: 1495 Australian_Labor_Party
 [pem_pos=0, final_scores=0.22616617,
 final_local_scores=0.19645995, pem=-0.2258703,
 similarity=11.180787]
 7: 3150155 George_Fuller_(Australian_politician)
 [pem_pos=0, final_scores=0.31492108,
 final_local_scores=0.33888596, pem=-0.03275912,
 similarity=14.380895]
 8: 2770934 Bertram_Stevens_(politician)
 [pem_pos=0, final_scores=0.29416317,
 final_local_scores=0.31079364, pem=0.0, similarity=13.364415]
 9: 2648858 Darling_Point,_New_South_Wales
 [pem_pos=0, final_scores=0.18619095,
 final_local_scores=0.20322159, pem=0.0, similarity=10.15425]
 1: 1883750 Daniel_Levy_(businessman)
 [pem_pos=0, final_scores=0.18455063,
 final_local_scores=0.2834495, pem=-0.2343586,
 similarity=13.82289]
 2: 4689264 Australia
 [pem_pos=0, final_scores=0.61380213,
 final_local_scores=0.6433929, pem=-0.491009,
 similarity=25.960024]
 3: 2646407 Joseph_Moses_Levy
 [pem_pos=0, final_scores=0.103057966,
 final_local_scores=0.19736391, pem=0.0, similarity=9.979445]
 4: 156992 Bachelor_of_Arts
 [pem_pos=0, final_scores=0.075702295,
 final_local_scores=0.092357695, pem=0.0, similarity=6.845851]
 5: 240037 Darlinghurst,_New_South_Wales
 [pem_pos=0, final_scores=0.07939704,
 final_local_scores=0.0510509, pem=0.0, similarity=5.6131735]
 6: 27862 Sydney
 [pem_pos=0, final_scores=0.19450109,
 final_local_scores=0.15110463, pem=-0.1443489,
 similarity=9.383968]
 7: 3150165 James_Doolley_(politician)
 [pem_pos=0, final_scores=0.04823251,
 final_local_scores=0.04594171, pem=0.0, similarity=5.4607053]
 8: 31717 United_Kingdom
 [pem_pos=0, final_scores=0.5609606,
 final_local_scores=0.58857083, pem=-0.48936236,
 similarity=24.31507]

Sample: b'Pawnee_Rangers'

Pawnee Rangers is the fourth episode of the fourth season of
 the [1 1 NBC 1] television series [2 2 Parks and Recreation 2]
 . 1 Ron -LRB- [3 3 Nick Offerman 3] -RRB- is the leader of a

- like group called the Pawnee Rangers , with 2 Andy -LRB- [4 4 Chris Pratt 4] -RRB- as his assistant . Girls are not allowed to join the Rangers , so 3 Leslie created the girls-only Pawnee Goddesses , with 4 Ann -LRB- [5 5 Rashida Jones 5] -RRB- , who is not fitting in too well , and 5 April -LRB- [6 6 Aubrey Plaza 6] -RRB- , whose morbid tattoo on her arm gets her the Goddesses ' approval , as her assistants . Leslie is determined to prove that her group is better than Ron 's to avenge the past rejection , so she arranges a weekend camping trip for the Goddesses at the same camping ground as the Rangers . Meanwhile , after noticing that 6 Ben -LRB- [1 7 Adam Scott 1] -RRB- is feeling down , 8 Donna -LRB- 9 Retta -RRB- and 10 Tom -LRB- [7 7 Aziz Ansari 7] -RRB- invite him on their annual `` Treat Yo self '' trip , where they spend a day spending extravagantly on things they do n't need . 11 Jerry -LRB- 12 Jim O'Heir -RRB- is left alone in the office , so 13 Chris -LRB- [8 8 Rob Lowe 8] -RRB- gives him the day off . Jerry invites Chris to have lunch with him , which [2 Chris 2] initially declines , but after Jerry 's attractive daughter Millicent `` Millie '' Gergich -LRB- 14 Sarah Wright

1: 21780 NBC

2: 20803357 Parks_and_Recreation

3: 3008793 Nick_Offerman

4: 1916589 Chris_Pratt

5: 2388997 Rashida_Jones

6: 20213373 Aubrey_Plaza

7: 3744603 Aziz_Ansari

8: 222445 Rob_Lowe

1: 25885950 Ron_Swanson

[winner: 25885950 Ron_Swanson, pem_pos=4, final_scores=-0.582487, final_local_scores=-0.67566025, pem=-2.9806836, similarity=0.13608482]

gt: 25885950 Ron_Swanson, pem_pos=4, final_scores=-0.582487, final_local_scores=-0.67566025, pem=-2.9806836, similarity=0.13608482]

2: 25886118 Andy_Dwyer

[winner: 8019836 Andy_Bernard, pem_pos=0, final_scores=-0.41518253, final_local_scores=-0.4777813, pem=-2.0718658, similarity=1.0988939]

gt: 25886118 Andy_Dwyer, pem_pos=4, final_scores=-0.5304265, final_local_scores=-0.6236186, pem=-3.1096735, similarity=2.3905776]

3: 24721389 Leslie_Knope

[winner: 24721389 Leslie_Knope, pem_pos=4, final_scores=-0.43589735, final_local_scores=-0.5174653, pem=-3.0304983, similarity=5.1278377]

gt: 24721389 Leslie_Knope, pem_pos=4, final_scores=-0.43589735, final_local_scores=-0.5174653, pem=-3.0304983, similarity=5.1278377]

4: 25885777 Ann_Perkins

[winner: 25885777 Ann_Perkins, pem_pos=1, final_scores=-0.40876442, final_local_scores=-0.47708154,

pem=-2.5960343, similarity=3.970286|
gt: 25885777 Ann_Perkins, pem_pos=1,
final_scores=-0.40876442, final_local_scores=-0.47708154,
pem=-2.5960343, similarity=3.970286]
5: 25886056 April_Ludgate
[winner: 1004 April, pem_pos=0, final_scores=-0.565976,
final_local_scores=-0.4312958, pem=-0.46492952,
similarity=-6.252655|
gt: 25886056 April_Ludgate, pem_pos=3,
final_scores=-0.8093365, final_local_scores=-0.9245915,
pem=-3.6588426, similarity=-3.604587]
6: 31086824 Ben_Wyatt_(Parks_and_Recreation)
[winner: 2110759 Ben_(album), pem_pos=1,
final_scores=-0.4182026, final_local_scores=-0.35111225,
pem=-1.9163477, similarity=4.0332193|
gt: 31086824 Ben_Wyatt_(Parks_and_Recreation),
pem_pos=13, final_scores=-0.7506956,
final_local_scores=-0.8548511, pem=-4.358192,
similarity=2.2797744]
7: 1999425 Adam_Scott_(actor)
[winner: 1080294 Adam_Scott_(golfer), pem_pos=1,
final_scores=0.24697898, final_local_scores=0.37822622,
pem=-0.053057656, similarity=15.665273|
gt: 1999425 Adam_Scott_(actor), pem_pos=0,
final_scores=0.14276488, final_local_scores=0.091735974,
pem=0.0, similarity=6.827297]
8: 27363309 Donna_Meagle
[winner: 976300 Donna_Moss, pem_pos=4,
final_scores=-0.33860952, final_local_scores=-0.3493799,
pem=-2.3120928, similarity=6.23704|
gt: 27363309 Donna_Meagle, pem_pos=12,
final_scores=-0.5795499, final_local_scores=-0.6293778,
pem=-4.122284, similarity=7.725446]
9: 25108286 Retta
[winner: 25108286 Retta, pem_pos=0,
final_scores=-0.042610094, final_local_scores=-0.06866498,
pem=-0.15886348, similarity=2.9045372|
gt: 25108286 Retta, pem_pos=0,
final_scores=-0.042610094, final_local_scores=-0.06866498,
pem=-0.15886348, similarity=2.9045372]
10: 25885657 Tom_Haverford
[winner: 31460 Tom_Cruise, pem_pos=0,
final_scores=-0.5837952, final_local_scores=-0.65860945,
pem=-1.7660737, similarity=-5.960323|
gt: 25885657 Tom_Haverford, pem_pos=8,
final_scores=-0.70743275, final_local_scores=-0.8123353,
pem=-3.7369194, similarity=0.1699556]
11: 27363212 Jerry_Gergich
[winner: 2091529 Jerry_Seinfeld_(character), pem_pos=0,
final_scores=-0.26959962, final_local_scores=-0.31022257,
pem=-1.3140358, similarity=1.9779831|
gt: 27363212 Jerry_Gergich, pem_pos=13,

```
final_scores=-0.81849843, final_local_scores=-0.89855164,  
pem=-4.3255506, similarity=0.798157 ]  
12: 26729620 Jim_0'Heir  
    [winner: 3721286 Jim_Halpert, pem_pos=0,  
final_scores=-0.7841278, final_local_scores=-0.86863595,  
pem=-1.3848933, similarity=-14.300852]  
    RECALL MISS: 26729620 Jim_0'Heir]  
    [winner: 26729620 Jim_0'Heir, pem_pos=0,  
final_scores=-1.9470099, final_local_scores=-2.1618934,  
pem=0.0, similarity=-60.425495]  
    gt: 26729620 Jim_0'Heir, pem_pos=0,  
final_scores=-1.9470099, final_local_scores=-2.1618934,  
pem=0.0, similarity=-60.425495 ]  
13: 31090092 Chris_Traeger  
    [winner: 1916589 Chris_Pratt, pem_pos=0,  
final_scores=-0.13631135, final_local_scores=-0.22497658,  
pem=0.0, similarity=-2.624037]  
    gt: 31090092 Chris_Traeger, pem_pos=13,  
final_scores=-0.7102736, final_local_scores=-0.8063449,  
pem=-4.1706667, similarity=2.7075062 ]  
14: 951798 Sarah_Wright  
    [winner: 951798 Sarah_Wright, pem_pos=0,  
final_scores=-0.044687197, final_local_scores=-0.044400297,  
pem=0.0, similarity=2.7647195]  
    gt: 951798 Sarah_Wright, pem_pos=0,  
final_scores=-0.044687197, final_local_scores=-0.044400297,  
pem=0.0, similarity=2.7647195 ]  
1: 21780 NBC  
    [pem_pos=0, final_scores=0.33562773,  
final_local_scores=0.34889746, pem=-0.06403453,  
similarity=14.849739]  
2: 20803357 Parks_and_Recreation  
    [pem_pos=1, final_scores=0.14136909,  
final_local_scores=0.11336231, pem=-0.039120883,  
similarity=7.685416]  
3: 3008793 Nick_Offerman  
    [pem_pos=0, final_scores=0.3385244,  
final_local_scores=0.31415585, pem=0.0, similarity=13.46475]  
4: 1916589 Chris_Pratt  
    [pem_pos=0, final_scores=0.13512413,  
final_local_scores=0.08196083, pem=0.0, similarity=6.5355873]  
5: 2388997 Rashida_Jones  
    [pem_pos=0, final_scores=0.20747332,  
final_local_scores=0.16592994, pem=0.0, similarity=9.041393]  
6: 20213373 Aubrey_Plaza  
    [pem_pos=0, final_scores=0.20830362,  
final_local_scores=0.17394114, pem=0.0, similarity=9.280463]  
7: 3744603 Aziz_Ansari  
    [pem_pos=0, final_scores=0.33743906,  
final_local_scores=0.30863687, pem=0.0, similarity=13.300053]  
8: 222445 Rob_Lowe  
    [pem_pos=0, final_scores=0.26712054,
```

```
final_local_scores=0.22630328, pem=0.0, similarity=10.843054]
1: 1080294 Adam_Scott_(golfer)
  [pem_pos=1, final_scores=0.24697898,
final_local_scores=0.37822622, pem=-0.053057656,
similarity=15.665273]
2: 1916589 Chris_Pratt
  [pem_pos=0, final_scores=0.06347604,
final_local_scores=0.0036961734, pem=0.0, similarity=4.200014]
```

Sample: b'Rio_Verde'

, _ Chile Río Verde is a [1 1 Chile 1] an 1 commune and settlement in [2 2 Magallanes Province 2] and . The commune is located in the area surrounding [3 3 Seno Skyring 3] , a sea inlet north of [4 4 Riesco Island 4] . According to the 2002 2 census of the [1 3 National Statistics Institute 1] , Río Verde spans an area of and has 358 inhabitants -LRB- 295 men and 63 women -RRB- , making the commune an entirely 4 rural area . The population grew by 6.9 % -LRB- 23 persons -RRB- between the 1992 and 2002 censuses . As a commune , Río Verde is a third-level administered by a 5 municipal council , headed by an 6 alcalde who is directly elected every four years . The 2008-2012 alcalde is Tatiana Vásquez Barrientos -LRB- 7 UDI -RRB- . Within the 8 electoral divisions of [2 Chile 2] , [3 Río Verde 3] is represented in the 9 Chamber of Deputies by Carolina Goic -LRB- -RRB- and Miodrag Marinovic -LRB- 10 Ind. . -RRB- as part of the 60th electoral district , which includes the entire . The commune is represented in the 11 Senate by Carlos Bianchi Chelech -LRB- Ind. -RRB- and Pedro Muñoz Aburto -LRB- 12 PS

1: 5489 Chile

2: 10173097 Magallanes_Province

3: 32070048 Seno_Skyring

4: 1346139 Riesco_Island

1: 3060770 Communes_of_Chile

[winner: 8277118 Communes_of_Algeria, pem_pos=3, final_scores=-1.0174497, final_local_scores=-0.9966226, pem=-3.2370336, similarity=-8.048001]

gt: 3060770 Communes_of_Chile, pem_pos=8, final_scores=-1.658514, final_local_scores=-1.8440056, pem=-3.809452, similarity=-30.222702]

2: 6889 Census

[winner: 6889 Census, pem_pos=0, final_scores=-1.3724471, final_local_scores=-1.4904125, pem=-0.05546598, similarity=-40.085537]

gt: 6889 Census, pem_pos=0, final_scores=-1.3724471, final_local_scores=-1.4904125, pem=-0.05546598, similarity=-40.085537]

3: 29015808 National_Statistics_Institute_(Chile)

[winner: 4070894

Instituto_Nacional_de_Estadística_y_Geografía, pem_pos=2, final_scores=0.11702941, final_local_scores=0.16051799, pem=-0.67166406, similarity=12.532503]

gt: 29015808 National_Statistics_Institute_(Chile),
 pem_pos=0, final_scores=0.021006599,
 final_local_scores=0.053246453, pem=-0.053524505,
 similarity=5.9697676]
 4: 212614 Rural_area
 [winner: 3909779 Penrith_Rural_District, pem_pos=1,
 final_scores=-1.1544425, final_local_scores=-1.106162,
 pem=-1.6043973, similarity=-20.195404]
 gt: 212614 Rural_area, pem_pos=0,
 final_scores=-1.5825955, final_local_scores=-1.6986268,
 pem=0.0, similarity=-46.6007]
 [winner: 19746394 Area_(nightclub), pem_pos=4,
 final_scores=-1.1703645, final_local_scores=-1.0558274,
 pem=-2.289687, similarity=-14.96661]
 RECALL MISS: 212614 Rural_area]
 5: 3532493 Municipal_council
 [winner: 19038 Municipality, pem_pos=0,
 final_scores=-1.0682284, final_local_scores=-1.1503761,
 pem=-0.19613224, similarity=-29.173208]
 RECALL MISS: 3532493 Municipal_council]
 [winner: 794347 Jedi_Council, pem_pos=6,
 final_scores=-1.4792751, final_local_scores=-1.3800738,
 pem=-3.1592648, similarity=-19.913864]
 RECALL MISS: 3532493 Municipal_council]
 6: 597450 Alcalde
 [winner: 125956 Alcalde,_New_Mexico, pem_pos=1,
 final_scores=-1.2617215, final_local_scores=-1.2800109,
 pem=-1.2197001, similarity=-27.475449]
 gt: 597450 Alcalde, pem_pos=0, final_scores=-1.2746001,
 final_local_scores=-1.3408165, pem=-0.06003342,
 similarity=-35.59646]
 7: 526689 Independent_Democratic_Union
 [winner: 526689 Independent_Democratic_Union, pem_pos=0,
 final_scores=-0.24075907, final_local_scores=-0.16413671,
 pem=-0.39522702, similarity=1.3408521]
 gt: 526689 Independent_Democratic_Union, pem_pos=0,
 final_scores=-0.24075907, final_local_scores=-0.16413671,
 pem=-0.39522702, similarity=1.3408521]
 8: 2683116 Electoral_divisions_of_Chile
 [winner: 5489 Chile, pem_pos=0, final_scores=0.64453095,
 final_local_scores=0.67613196, pem=-0.27404457,
 similarity=25.757135]
 RECALL MISS: 2683116 Electoral_divisions_of_Chile]
 [winner: 31703793 Municipal_Units_of_Tirana, pem_pos=15,
 final_scores=-1.1496234, final_local_scores=-1.0229826,
 pem=-4.145923, similarity=-3.891955]
 RECALL MISS: 2683116 Electoral_divisions_of_Chile]
 [winner: 948687 Roman_Catholic_Diocese_of_Mainz,
 pem_pos=7, final_scores=-1.382327,
 final_local_scores=-1.3665423, pem=-2.0709488,
 similarity=-25.428493]
 RECALL MISS: 2683116 Electoral_divisions_of_Chile]

9: 716259 Chamber_of_Deputies_of_Chile
 [winner: 411807 Chamber_of_Deputies_(Mexico), pem_pos=0,
 final_scores=-0.094394594, final_local_scores=-0.027270406,
 pem=-1.5470045, similarity=11.688761]
 gt: 716259 Chamber_of_Deputies_of_Chile, pem_pos=2,
 final_scores=-0.15433577, final_local_scores=-0.077522606,
 pem=-1.9696342, similarity=12.487461]

10: 2172378 Independent_(politician)
 [winner: 21883857 Indiana, pem_pos=3,
 final_scores=-0.506902, final_local_scores=-0.48946595,
 pem=-1.5513303, similarity=-2.080552]
 gt: 2172378 Independent_(politician), pem_pos=0,
 final_scores=-0.6417812, final_local_scores=-0.5616332,
 pem=-0.23172729, similarity=-11.410373]
 [winner: 330842 Achang_people, pem_pos=1,
 final_scores=-1.6399748, final_local_scores=-1.6470585,
 pem=-2.4264104, similarity=-31.866602]
 RECALL MISS: 2172378 Independent_(politician)]

11: 683411 Senate_of_Chile
 [winner: 16685964 Roman_Senate, pem_pos=2,
 final_scores=-0.5430304, final_local_scores=-0.45276248,
 pem=-2.525485, similarity=4.312357]
 gt: 683411 Senate_of_Chile, pem_pos=9,
 final_scores=-0.6765993, final_local_scores=-0.65011793,
 pem=-3.724167, similarity=4.9414964]

12: 405910 Socialist_Party_of_Chile
 [winner: 39565408 Socialist_Party_(France), pem_pos=1,
 final_scores=-0.58553064, final_local_scores=-0.5211546,
 pem=-1.8423814, similarity=-1.4434223]
 gt: 405910 Socialist_Party_of_Chile, pem_pos=4,
 final_scores=-0.7136347, final_local_scores=-0.66806495,
 pem=-2.926659, similarity=0.068948835]

1: 5489 Chile
 [pem_pos=0, final_scores=0.64217514,
 final_local_scores=0.6735586, pem=-0.27404457,
 similarity=25.680342]

2: 10173097 Magallanes_Province
 [pem_pos=0, final_scores=0.2517712,
 final_local_scores=0.27908736, pem=0.0, similarity=12.418237]

3: 32070048 Seno_Skyring
 [pem_pos=0, final_scores=0.042341784,
 final_local_scores=0.12629724, pem=0.0, similarity=7.8586745]

4: 1346139 Riesco_Island
 [pem_pos=0, final_scores=0.15874939,
 final_local_scores=0.22048327, pem=0.0, similarity=10.6693735]

1: 4070894 Instituto_Nacional_de_Estadística_y_Geografía
 [pem_pos=2, final_scores=0.11702941,
 final_local_scores=0.16051799, pem=-0.67166406,
 similarity=12.532503]

2: 5489 Chile
 [pem_pos=0, final_scores=0.64453095,
 final_local_scores=0.67613196, pem=-0.27404457,

similarity=25.757135]
3: 25988731 Río_Verde_Canton
[pem_pos=0, final_scores=0.042594418,
final_local_scores=0.09277862, pem=-0.1533611,
similarity=7.6924133]

Sample: b'Bulgarian_Black_Sea_Coast'

The Bulgarian 1 Black Sea Coast -LRB- , Chernomorie -RRB- covers the entire eastern bound of [1 1 Bulgaria 1] stretching from the [1 2 Romanian 1] [2 Black Sea 2] resorts in the north to in the south , along 378 km of coastline . White and golden sandy 3 beach es occupy approximately 130 km of the 378 km long 4 coast . The region is an important center of tourism during the summer season -LRB- May -- October -RRB- , drawing millions of foreign and local tourists alike and constituting one of the country 's most popular tourist destinations . Prior to 1989 the [3 Bulgarian 3] [4 Black Sea 4] coast was internationally known as the Red Riviera . Since the 5 fall of the [5 Iron Curtain 5] , however , its nickname has been changed to the [6 Bulgarian Riviera 6] . The [7 Bulgarian 7] [8 Black Sea Coast 8] has a 6 humid subtropical climate -LRB- 7 Köppen climate classification Cfa -RRB- , with considerable 8 maritime and 9 continental influences . The area 's average air temperature in the summer is about 28 °C , with the average water temperature at 26 °C . There are more than 240 hours of 10 sunshine in May and September and more than 300 hours in July and August . The [2 2 Balkan Mountains 2] cross the country reaching to the edge of the [9 Black Sea 9] at 11 Cape Emine , dividing the coastline into a southern and northern part . Parts of [10 Bulgaria 10] 's northern [11 Black Sea Coast 11] feature 12 rocky 13 headlands where the sea abuts 14 cliff s up to 70 metres in height . The southern coast is known for its wide 15 sandy beaches . The two largest cities and main seaports on the [12 Bulgarian Riviera 12] are [3 3 Varna 3] -LRB- third largest in the country -RRB- and [4 4 Burgas 4] -LRB- fourth largest in the country -RRB- . [13 Varna 13] is located on the northern part of the coast and [14 Burgas 14] is located on the southern coast . The two cities ' international airports , [5 5 Varna Airport 5] and [6 6 Burgas Airport 6] , are the main hubs servicing the region . In addition , the and 16 A2 motorways , currently in construction , would make the trip from the capital [7 7 Sofia 7] to the coast substantially easier and faster , while the 17 A3

1: 3415 Bulgaria
2: 364825 Balkan_Mountains
3: 86762 Varna
4: 86758 Burgas
5: 963533 Varna_Airport
6: 963591 Burgas_Airport
7: 57644 Sofia
1: 3386 Black_Sea

[winner: 1943653 Bulgarian_Black_Sea_Coast, pem_pos=0,
final_scores=0.0129580945, final_local_scores=-0.041566655,
pem=-0.047628075, similarity=3.1082897]

RECALL MISS: 3386 Black_Sea]

2: 2333918 Romanian_Black_Sea_resorts

[winner: 25445 Romania, pem_pos=1,
final_scores=0.5720773, final_local_scores=0.5889979,
pem=-1.0382323, similarity=27.312654]

RECALL MISS: 2333918 Romanian_Black_Sea_resorts]

[winner: 3386 Black_Sea, pem_pos=0,
final_scores=0.4414447, final_local_scores=0.4127276,
pem=0.0, similarity=16.406328]

gt: 2333918 Romanian_Black_Sea_resorts, pem_pos=12,
final_scores=-1.1526953, final_local_scores=-1.2059188,
pem=-7.069145, similarity=6.545802]

[winner: 471772 Resort, pem_pos=1,
final_scores=-1.7969946, final_local_scores=-1.9414091,
pem=-0.0061335526, similarity=-53.812454]

RECALL MISS: 2333918 Romanian_Black_Sea_resorts]

3: 40630 Beach

[winner: 23885719 Sandy_Beach, New_York, pem_pos=1,
final_scores=-0.9673515, final_local_scores=-0.89578784,
pem=-0.60737646, similarity=-19.339375]

RECALL MISS: 40630 Beach]

4: 5236 Coast

[winner: 23716943 Cilentan_Coast, pem_pos=7,
final_scores=-1.2691276, final_local_scores=-1.2681391,
pem=-1.9440129, similarity=-23.182243]

gt: 5236 Coast, pem_pos=0, final_scores=-1.6207744,
final_local_scores=-1.7618315, pem=-0.09349413,
similarity=-47.978416]

5: 4584893 Revolutions_of_1989

[winner: 18953051 Iron_Curtain, pem_pos=0,
final_scores=0.20056365, final_local_scores=0.22970623,
pem=0.0, similarity=10.944605]

RECALL MISS: 4584893 Revolutions_of_1989]

[winner: 20448910 Granada_War, pem_pos=0,
final_scores=-1.0276543, final_local_scores=-1.0239568,
pem=0.0, similarity=-26.4672]

RECALL MISS: 4584893 Revolutions_of_1989]

6: 19792392 Humid_subtropical_climate

[winner: 19792392 Humid_subtropical_climate, pem_pos=0,
final_scores=-0.9798667, final_local_scores=-0.977643,
pem=-0.0026874263, similarity=-25.07049]

gt: 19792392 Humid_subtropical_climate, pem_pos=0,
final_scores=-0.9798667, final_local_scores=-0.977643,
pem=-0.0026874263, similarity=-25.07049]

[winner: 6000991 Climates_(film), pem_pos=1,
final_scores=-1.1660851, final_local_scores=-1.1442666,
pem=-0.6931472, similarity=-26.288044]

RECALL MISS: 19792392 Humid_subtropical_climate]

7: 484254 Köppen_climate_classification

[winner: 484254 Köppen_climate_classification, pem_pos=0, final_scores=-0.16881537, final_local_scores=-0.11143084, pem=-0.011456118, similarity=0.8266945]

gt: 484254 Köppen_climate_classification, pem_pos=0, final_scores=-0.16881537, final_local_scores=-0.11143084, pem=-0.011456118, similarity=0.8266945]

[winner: 25469046

Climate,_Community_&_Biodiversity_Alliance, pem_pos=0, final_scores=-1.4216751, final_local_scores=-1.3736153, pem=-0.6931472, similarity=-33.132267]

RECALL MISS: 484254 Köppen_climate_classification]

[winner: 7512482 Attribute-value_system, pem_pos=9, final_scores=-1.4512936, final_local_scores=-1.4003453, pem=-2.6390574, similarity=-23.347782]

RECALL MISS: 484254 Köppen_climate_classification]

8: 560047 Oceanic_climate

[winner: 560047 Oceanic_climate, pem_pos=2, final_scores=-1.4345034, final_local_scores=-1.5096982, pem=-1.5940117, similarity=-32.29421]

gt: 560047 Oceanic_climate, pem_pos=2, final_scores=-1.4345034, final_local_scores=-1.5096982, pem=-1.5940117, similarity=-32.29421]

9: 1225918 Humid_continental_climate

[winner: 25853611 Climate_categories_in_viticulture, pem_pos=0, final_scores=-0.9637539, final_local_scores=-0.96468955, pem=-0.22314353, similarity=-23.485058]

RECALL MISS: 1225918 Humid_continental_climate]

10: 27979 Sunlight

[winner: 8238831 Sunshine,_New_South_Wales, pem_pos=6, final_scores=-1.1520092, final_local_scores=-1.0776364, pem=-3.3929029, similarity=-9.617972]

gt: 27979 Sunlight, pem_pos=1, final_scores=-1.5433707, final_local_scores=-1.5813401, pem=-2.3701515, similarity=-30.211374]

11: 1172949 Cape_Emine

[winner: 1172949 Cape_Emine, pem_pos=0, final_scores=0.02598302, final_local_scores=0.10940482, pem=0.0, similarity=7.354571]

gt: 1172949 Cape_Emine, pem_pos=0, final_scores=0.02598302, final_local_scores=0.10940482, pem=0.0, similarity=7.354571]

12: 199661 Rock_(geology)

[winner: 45772 Rocky, pem_pos=0, final_scores=-1.3593824, final_local_scores=-1.3611238, pem=-0.4167528, similarity=-34.26257]

RECALL MISS: 199661 Rock_(geology)]

13: 275124 Headlands_and_bays

[winner: 275124 Headlands_and_bays, pem_pos=0, final_scores=-1.0434455, final_local_scores=-1.08177, pem=-0.41127065, similarity=-25.955906]

gt: 275124 Headlands_and_bays, pem_pos=0,

final_scores=-1.0434455, final_local_scores=-1.08177,
pem=-0.41127065, similarity=-25.955906]
14: 181897 Cliff
[winner: 15859588
Cliff_Township,_Custer_County,_Nebraska, pem_pos=5,
final_scores=-1.262431, final_local_scores=-1.1704175,
pem=-3.270561, similarity=-13.052061]
gt: 181897 Cliff, pem_pos=0, final_scores=-1.493451,
final_local_scores=-1.588614, pem=-0.5649818,
similarity=-40.245235]
15: 18994037 Sand
[winner: 18994037 Sand, pem_pos=0,
final_scores=-0.97732884, final_local_scores=-1.0049517,
pem=-0.019828038, similarity=-25.792221]
gt: 18994037 Sand, pem_pos=0, final_scores=-0.97732884,
final_local_scores=-1.0049517, pem=-0.019828038,
similarity=-25.792221]
16: 4944050 Hemus_motorway
[winner: 928750 A2_motorway_(Switzerland), pem_pos=0,
final_scores=-0.6064408, final_local_scores=-0.56449175,
pem=0.0, similarity=-12.755847]
RECALL MISS: 4944050 Hemus_motorway]
17: 6769601 Cherno_More_motorway
[winner: 174223 A3_road, pem_pos=0,
final_scores=-0.7108978, final_local_scores=-0.71050745,
pem=-1.6441958, similarity=-8.171854]
RECALL MISS: 6769601 Cherno_More_motorway]
1: 3415 Bulgaria
[pem_pos=1, final_scores=0.6954227,
final_local_scores=0.6799075, pem=-0.31727514,
similarity=26.1049]
2: 364825 Balkan_Mountains
[pem_pos=0, final_scores=0.23255445,
final_local_scores=0.20909142, pem=0.0, similarity=10.329418]
3: 86762 Varna
[pem_pos=0, final_scores=0.27199417,
final_local_scores=0.23086748, pem=-0.1519721,
similarity=11.805707]
4: 86758 Burgas
[pem_pos=0, final_scores=0.23873644,
final_local_scores=0.17993739, pem=0.0, similarity=9.459404]
5: 963533 Varna_Airport
[pem_pos=0, final_scores=0.1989374,
final_local_scores=0.2612521, pem=0.0, similarity=11.885997]
6: 963591 Burgas_Airport
[pem_pos=0, final_scores=0.17654185,
final_local_scores=0.21445724, pem=0.0, similarity=10.489545]
7: 57644 Sofia
[pem_pos=0, final_scores=0.57658273,
final_local_scores=0.5645944, pem=-0.10680699,
similarity=21.519169]
1: 25445 Romania

[pem_pos=1, final_scores=0.5720773,
 final_local_scores=0.5889979, pem=-1.0382323,
 similarity=27.312654]
 2: 3386 Black_Sea
 [pem_pos=0, final_scores=0.4414447,
 final_local_scores=0.4127276, pem=0.0, similarity=16.406328]
 3: 3415 Bulgaria
 [pem_pos=1, final_scores=0.50262064,
 final_local_scores=0.4693005, pem=-1.1560022,
 similarity=24.381096]
 4: 3386 Black_Sea
 [pem_pos=0, final_scores=0.37542063,
 final_local_scores=0.34060633, pem=0.0, similarity=14.254086]
 5: 18953051 Iron_Curtain
 [pem_pos=0, final_scores=0.20056365,
 final_local_scores=0.22970623, pem=0.0, similarity=10.944605]
 6: 1943653 Bulgarian_Black_Sea_Coast
 [pem_pos=0, final_scores=0.08478956,
 final_local_scores=0.044433266, pem=0.0, similarity=5.4156904]
 7: 3415 Bulgaria
 [pem_pos=1, final_scores=0.39810795,
 final_local_scores=0.35513633, pem=-1.1560022,
 similarity=20.974209]
 8: 1943653 Bulgarian_Black_Sea_Coast
 [pem_pos=0, final_scores=0.08435382,
 final_local_scores=0.043957278, pem=-0.047628075,
 similarity=5.660495]
 9: 3386 Black_Sea
 [pem_pos=0, final_scores=0.4177181,
 final_local_scores=0.3868099, pem=0.0, similarity=15.632891]
 10: 3415 Bulgaria
 [pem_pos=1, final_scores=0.710365,
 final_local_scores=0.6962297, pem=-0.31727514,
 similarity=26.591988]
 11: 1943653 Bulgarian_Black_Sea_Coast
 [pem_pos=0, final_scores=0.13244735,
 final_local_scores=0.09649216, pem=-0.047628075,
 similarity=7.2282405]
 12: 1943653 Bulgarian_Black_Sea_Coast
 [pem_pos=0, final_scores=0.16896875,
 final_local_scores=0.13638625, pem=0.0, similarity=8.159751]
 13: 86762 Varna
 [pem_pos=0, final_scores=0.3006608,
 final_local_scores=0.2621814, pem=-0.1519721,
 similarity=12.740177]
 14: 86758 Burgas
 [pem_pos=0, final_scores=0.37295353,
 final_local_scores=0.32654914, pem=0.0, similarity=13.834591]

Sample: b'Colombo_West_Electoral_District'

[1 Colombo West 1] electoral district was an 1 electoral dis-
 trict of [1 1 Sri Lanka 1] between 2 July 1977 and 3 February

1989 . The district was named after the city of [2 2 Colombo 2] in [3 3 Colombo District 3] , 4 Western Province . The district was created by the division of the 5 Colombo South Electoral District into [4 4 Colombo East 4] and [2 Colombo West 2] electoral districts in July 1977 . The 6 1978 Constitution of Sri Lanka introduced the 7 proportional representation electoral system for 8 electing 9 members of 10 Parliament . The existing 160 mainly 11 single-member electoral districts were replaced with 22 12 electoral districts. [3 Colombo West 3] electoral district was replaced by the [4 13 Colombo 4] multi-member electoral district at the 14 1989 general elections

1: 26750 Sri_Lanka

2: 56636 Colombo

3: 3973477 Colombo_District

4: 24244335 Colombo_East_Electoral_District

1: 667057 Electoral_district

[winner: 149917 Verden_(district), pem_pos=0, final_scores=-1.01193, final_local_scores=-0.9387446, pem=0.0, similarity=-23.9243]

RECALL MISS: 667057 Electoral_district]

[winner: 948687 Roman_Catholic_Diocese_of_Mainz, pem_pos=7, final_scores=-1.4032903, final_local_scores=-1.4056337, pem=-2.0709488, similarity=-26.59506]

gt: 667057 Electoral_district, pem_pos=15, final_scores=-2.2011957, final_local_scores=-2.385485, pem=-5.7451377, similarity=-35.85496]

2: 6050096 Sri_Lankan_parliamentary_election,_1977

[winner: 6050096 Sri_Lankan_parliamentary_election,_1977, pem_pos=0, final_scores=-0.37272, final_local_scores=-0.39533696, pem=0.0, similarity=-7.7079306]

gt: 6050096 Sri_Lankan_parliamentary_election,_1977, pem_pos=0, final_scores=-0.37272, final_local_scores=-0.39533696, pem=0.0, similarity=-7.7079306]

3: 6049990 Sri_Lankan_parliamentary_election,_1989

[winner: 22846351 Kenneth_Eather, pem_pos=6, final_scores=-0.49543715, final_local_scores=-0.44420013, pem=-2.0794415, similarity=2.142222]

RECALL MISS: 6049990

Sri_Lankan_parliamentary_election,_1989]

[winner: 34847 1989, pem_pos=0, final_scores=-1.2125962, final_local_scores=-1.1273863, pem=-0.6980697, similarity=-25.757534]

RECALL MISS: 6049990

Sri_Lankan_parliamentary_election,_1989]

4: 262187 Western_Province,_Sri_Lanka

[winner: 262187 Western_Province,_Sri_Lanka, pem_pos=3, final_scores=-0.018319681, final_local_scores=-0.06602797, pem=-1.4717305, similarity=10.122806]

gt: 262187 Western_Province, Sri_Lanka, pem_pos=3,
 final_scores=-0.018319681, final_local_scores=-0.06602797,
 pem=-1.4717305, similarity=10.122806]
 5: 24244155 Colombo_South_Electoral_District
 [winner: 24244155 Colombo_South_Electoral_District,
 pem_pos=0, final_scores=0.020345673,
 final_local_scores=0.023647293, pem=0.0,
 similarity=4.7953954]
 gt: 24244155 Colombo_South_Electoral_District,
 pem_pos=0, final_scores=0.020345673,
 final_local_scores=0.023647293, pem=0.0,
 similarity=4.7953954]
 6: 5308117 Constitution_of_Sri_Lanka
 [winner: 5308117 Constitution_of_Sri_Lanka, pem_pos=0,
 final_scores=-0.13164572, final_local_scores=-0.11461155,
 pem=0.0, similarity=0.6694756]
 gt: 5308117 Constitution_of_Sri_Lanka, pem_pos=0,
 final_scores=-0.13164572, final_local_scores=-0.11461155,
 pem=0.0, similarity=0.6694756]
 [winner: 34753 1978, pem_pos=0, final_scores=-0.71843773,
 final_local_scores=-0.5039027, pem=-0.7079138,
 similarity=-7.098004]
 RECALL MISS: 5308117 Constitution_of_Sri_Lanka]
 7: 23298 Proportional_representation
 [winner: 4415141 Wold's_theorem, pem_pos=8,
 final_scores=-1.1966091, final_local_scores=-1.0662844,
 pem=-2.0768921, similarity=-16.435879]
 gt: 23298 Proportional_representation, pem_pos=6,
 final_scores=-1.5987892, final_local_scores=-1.7175345,
 pem=-2.0617309, similarity=-35.952927]
 [winner: 81863 Proportionality_(mathematics), pem_pos=0,
 final_scores=-1.2068963, final_local_scores=-1.1232095,
 pem=0.0, similarity=-29.4291]
 gt: 23298 Proportional_representation, pem_pos=1,
 final_scores=-1.3525656, final_local_scores=-1.4485726,
 pem=-0.46681154, similarity=-36.599987]
 8: 925514 Elections_in_Sri_Lanka
 [winner: 925514 Elections_in_Sri_Lanka, pem_pos=0,
 final_scores=-0.9428234, final_local_scores=-0.99050933,
 pem=0.0, similarity=-25.46906]
 gt: 925514 Elections_in_Sri_Lanka, pem_pos=0,
 final_scores=-0.9428234, final_local_scores=-0.99050933,
 pem=0.0, similarity=-25.46906]
 9: 98066 Member_of_Parliament
 [winner: 9090953 Gammaproteobacteria, pem_pos=0,
 final_scores=-1.1663285, final_local_scores=-1.0361732,
 pem=-1.5855416, similarity=-18.209341]
 gt: 98066 Member_of_Parliament, pem_pos=1,
 final_scores=-1.5907747, final_local_scores=-1.6998215,
 pem=-1.7353394, similarity=-37.199306]
 10: 571660 Parliament_of_Sri_Lanka
 [winner: 24406 Parliament, pem_pos=1,


```

final_scores=-0.5123124, final_local_scores=-0.4998821,
pem=-1.8263724, similarity=-0.89566875|
    gt: 571660 Parliament_of_Sri_Lanka, pem_pos=19,
final_scores=-0.6881277, final_local_scores=-0.786396,
pem=-4.4426966, similarity=4.782162 ]
11: 10880 Plurality_voting_system
    [winner: 10880 Plurality_voting_system, pem_pos=0,
final_scores=-1.2751738, final_local_scores=-1.3195984,
pem=0.0, similarity=-35.289738|
    gt: 10880 Plurality_voting_system, pem_pos=0,
final_scores=-1.2751738, final_local_scores=-1.3195984,
pem=0.0, similarity=-35.289738 ]
12: 667057 Electoral_district
    [winner: 2198981 Electoral_districts_of_New_South_Wales,
pem_pos=2, final_scores=-1.1890165,
final_local_scores=-1.2114053, pem=-1.4789519,
similarity=-24.018269|
    gt: 667057 Electoral_district, pem_pos=1,
final_scores=-1.7097101, final_local_scores=-1.8486114,
pem=-1.3563718, similarity=-43.700375 ]
13: 24923476 Colombo_Electoral_District
    [winner: 56636 Colombo, pem_pos=0,
final_scores=0.40016246, final_local_scores=0.36829364,
pem=-0.19393882, similarity=16.135|
    gt: 24923476 Colombo_Electoral_District, pem_pos=8,
final_scores=-0.69406843, final_local_scores=-0.7230161,
pem=-3.6970882, similarity=2.618812 ]
    [winner: 2198981 Electoral_districts_of_New_South_Wales,
pem_pos=2, final_scores=-1.0227764,
final_local_scores=-1.0298133, pem=-1.4789519,
similarity=-18.599197|
    RECALL MISS: 24923476 Colombo_Electoral_District]
    [winner: 15605478 Plurality-at-large_voting, pem_pos=0,
final_scores=-1.6211681, final_local_scores=-1.6940591,
pem=0.0, similarity=-46.464394|
    RECALL MISS: 24923476 Colombo_Electoral_District]
14: 6049990 Sri_Lankan_parliamentary_election,_1989
    [winner: 6049990 Sri_Lankan_parliamentary_election,_1989,
pem_pos=0, final_scores=-0.6688155,
final_local_scores=-0.7312185, pem=0.0,
similarity=-17.731306|
    gt: 6049990 Sri_Lankan_parliamentary_election,_1989,
pem_pos=0, final_scores=-0.6688155,
final_local_scores=-0.7312185, pem=0.0,
similarity=-17.731306 ]
1: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.36403555,
final_local_scores=0.36843747, pem=-0.3027737,
similarity=16.731152]
2: 56636 Colombo
    [pem_pos=0, final_scores=0.47154582,
final_local_scores=0.4462691, pem=-0.19393882,

```

similarity=18.461943]
 3: 3973477 Colombo_District
 [pem_pos=0, final_scores=0.222456,
 final_local_scores=0.24064705, pem=0.0, similarity=11.271101]
 4: 24244335 Colombo_East_Electoral_District
 [pem_pos=0, final_scores=0.03944148,
 final_local_scores=0.015378371, pem=0.0, similarity=4.548634]
 1: 56636 Colombo
 [pem_pos=1, final_scores=0.25105643,
 final_local_scores=0.19335693, pem=-0.6546809,
 similarity=13.420126]
 2: 56636 Colombo
 [pem_pos=1, final_scores=0.16940115,
 final_local_scores=0.10416095, pem=-0.6546809,
 similarity=10.75834]
 3: 56636 Colombo
 [pem_pos=1, final_scores=0.14631854,
 final_local_scores=0.078946695, pem=-0.6546809,
 similarity=10.005897]
 4: 56636 Colombo
 [pem_pos=0, final_scores=0.40016246,
 final_local_scores=0.36829364, pem=-0.19393882,
 similarity=16.135]

Sample: b'Jacob_de_Jager'

[1 Jacob de Jager 1] -LRB- 16 January 1923 -- 25 February
 2004 -RRB- was a 1 general authority of [1 1 The Church of Je
 sus Christ of Latter-day Saints 1] -LRB- LDS Church -RRB- .
 He was the first person to serve as a general authority . De
 Jager was born in the [2 2 Netherlands 2] and joined the
 church while living in [3 3 Toronto 3] , [4 4 Canada 4] . His
 work took him to [5 5 Australia 5] , [6 6 Indonesia 6] , [7 7
 Mexico 7] and [8 8 Canada 8] . His first in the church was
 hymnal coordinator which he filled honorably . He also served
 as 2 Elders 3 Quorum president , 4 branch president ,
 counselor to 5 mission president and before his call to the .
 In 1993 de Jager was given general authority 6 emeritus
 status and following his call to the Seventy he served as 7 bi
 shop of a 8 ward

1: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints
 2: 21148 Netherlands
 3: 64646 Toronto
 4: 5042916 Canada
 5: 4689264 Australia
 6: 14579 Indonesia
 7: 3966054 Mexico
 8: 5042916 Canada

1: 657147 General_authority
 [winner: 2534853 Franciszek_Kleeberg, pem_pos=1,
 final_scores=-1.339475, final_local_scores=-1.2996895,
 pem=-0.4054651, similarity=-32.49063]
 RECALL MISS: 657147 General_authority]

[winner: 3677766 Auctoritas, pem_pos=3,
final_scores=-1.4529564, final_local_scores=-1.4525473,
pem=-1.0978392, similarity=-33.286972]
RECALL MISS: 657147 General_authority]

2: 1471081 Elder_(Latter_Day_Saints)
[winner: 1471081 Elder_(Latter_Day_Saints), pem_pos=1,
final_scores=-0.5680165, final_local_scores=-0.5507746,
pem=-1.1027532, similarity=-6.349555]
gt: 1471081 Elder_(Latter_Day_Saints), pem_pos=1,
final_scores=-0.5680165, final_local_scores=-0.5507746,
pem=-1.1027532, similarity=-6.349555]

3: 1495670 Quorum_(Latter_Day_Saints)
[winner: 5985284 Korem, pem_pos=0,
final_scores=-0.60037357, final_local_scores=-0.5103138,
pem=0.0, similarity=-11.139069]
gt: 1495670 Quorum_(Latter_Day_Saints), pem_pos=1,
final_scores=-0.9050626, final_local_scores=-0.88425297,
pem=-0.6931472, similarity=-18.528717]

4: 1100647 Branch_president
[winner: 1100647 Branch_president, pem_pos=0,
final_scores=-1.3984901, final_local_scores=-1.487617,
pem=0.0, similarity=-40.30375]
gt: 1100647 Branch_president, pem_pos=0,
final_scores=-1.3984901, final_local_scores=-1.487617,
pem=0.0, similarity=-40.30375]

5: 2047259 Mission_president
[winner: 22673654 Mission_California, pem_pos=11,
final_scores=-1.1035595, final_local_scores=-1.017841,
pem=-3.2960024, similarity=-8.360519]
RECALL MISS: 2047259 Mission_president]
[winner: 24113 President_of_the_United_States, pem_pos=0,
final_scores=-1.65101, final_local_scores=-1.7777746,
pem=-0.99314094, similarity=-43.561768]
RECALL MISS: 2047259 Mission_president]

6: 185793 Emeritus
[winner: 20469341 Emeritus_(album), pem_pos=0,
final_scores=-0.670498, final_local_scores=-0.5659586,
pem=0.0, similarity=-12.799621]
gt: 185793 Emeritus, pem_pos=1, final_scores=-1.0187973,
final_local_scores=-1.0929058, pem=-0.07016907,
similarity=-28.143183]

7: 1236579 Bishop_(Latter_Day_Saints)
[winner: 15592443 Roman_Catholic_Diocese_of_Cuenca,
pem_pos=0, final_scores=-0.9340952,
final_local_scores=-0.90967506, pem=0.0,
similarity=-23.056805]
RECALL MISS: 1236579 Bishop_(Latter_Day_Saints)]

8: 436529 Ward_(LDS_Church)
[winner: 3726883 Alan_Ward, pem_pos=0,
final_scores=-1.0529363, final_local_scores=-1.0004793,
pem=0.0, similarity=-25.766586]
RECALL MISS: 436529 Ward_(LDS_Church)]

1: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints
 [pem_pos=0, final_scores=0.14072214,
 final_local_scores=0.15190163, pem=0.0, similarity=8.622761]
 2: 21148 Netherlands
 [pem_pos=0, final_scores=0.41252762,
 final_local_scores=0.40035844, pem=-0.24940795,
 similarity=17.393526]
 3: 64646 Toronto
 [pem_pos=0, final_scores=0.5517546,
 final_local_scores=0.5599384, pem=-0.12291032,
 similarity=21.467796]
 4: 5042916 Canada
 [pem_pos=0, final_scores=0.73459536,
 final_local_scores=0.7175445, pem=-0.11662793,
 similarity=26.136913]
 5: 4689264 Australia
 [pem_pos=0, final_scores=0.5238885,
 final_local_scores=0.5055313, pem=-0.17168477,
 similarity=20.109425]
 6: 14579 Indonesia
 [pem_pos=0, final_scores=0.6084346,
 final_local_scores=0.6343453, pem=-0.23151328,
 similarity=24.278847]
 7: 3966054 Mexico
 [pem_pos=0, final_scores=0.8003983,
 final_local_scores=0.8331627, pem=-0.18258883,
 similarity=29.945896]
 8: 5042916 Canada
 [pem_pos=0, final_scores=0.68266505,
 final_local_scores=0.6608185, pem=-0.11662793,
 similarity=24.444096]
 1: 12573377 Jacob_de_Jager
 [pem_pos=0, final_scores=0.0448675,
 final_local_scores=0.13290158, pem=0.0, similarity=8.055761]

Sample: b'Kazuo_Aoki'

, was a bureaucrat and cabinet minister in the [1 1 Empire of Japan 1] , serving as 1 Minister of [1 Finance 1] , and 2 Minister of [2 Greater East Asia 2] . Aoki was born to a farming family in 3 Sarashina District , -LRB- now part of the city of [2 2 Nagano 2] -RRB- , and was trained as a lawyer , graduating from the Law School of in 1916 . On graduation , he entered the 4 Ministry of Finance . Rising rapidly through the ranks , Aoki became chief of the Financial Bureau under [3 3 Takahashi Korekiyo 3] , who made use of his legal background to have Aoki draft a Foreign Exchange Management Act , which was passed by the in 1933 . Up until that time , [3 Japan 3] had not attempted to implement comprehensive state control over foreign exchange . Aoki followed up on this law with the Rice Control Act -LRB- 1933 -RRB- and the Petroleum Control Act -LRB- 1934 -RRB- , which set the stage for increasing state control over strategic sectors of the

economy . He was also on the committee which drafted the 5 National Service Draft Ordinance , which placed the 4 Japanese 4] economy on a 6 war economy footing after the start of the [4 4 Second Sino-Japanese War 4] . 7 Prime Minister [5 5 Fumimaro Konoe 5] asked Aoki to become deputy director of the in 1937 , and he became its chairman in 1939 . The same year , Aoki was nominated to a seat in the 8 Upper House in the [5 Diet 5] . Under Prime Minister [6 9 Abe 6] , Aoki was Minister of [7 Finance 7] in 1939 , while retaining his post as chairman of the Planning Board . After the fall of the [8 Abe 8] administration , Aoki was assigned as a special envoy to the [6 6 Reorganized National Government of China 6] to guide economic policy . He was recalled to [9 Japan 9] under the to the newly-created cabinet post of 10 Minister of [10 Greater East Asia 10] , in which position he oversaw the [7 7 Greater East Asia Conference 7] . After the end of [11 World War II 11] , Aoki was arrested -LRB- as were all former government members -RRB- by the and held in 11 Sugamo Prison on charges of 12 war crime s. However , he was released in 1948 without coming to trial . Afterwards , Aoki established a private legal practice . In 1953 , he ran for a seat in the with the support of the 13 Liberal Party -LRB- Japan , 1945 -RRB- on a nation-wide ticket . He was subsequently reelected to the same seat three more times as a member of the [12 14 Liberal Democratic Party 12] -LRB- [13 LDP 13] -RRB- . He joined a right-wing faction within the [14 LDP 14] in 1960 which was adamantly opposed to [15 Japan 15] 's normalization of relations with the and supported [16 Japan 16] 's continued recognition of the on [8 8 Taiwan 8] . He was also a strong supporter of building a nation-wide network of highways in [17 Japan 17] , especially the [9 9 Chūō Expressway 9] . After his retirement from politics , Aoki joined the Board of Directors of 15 Shin-etsu Broadcasting , and in 1968 was one of the founders of the 16 Nagano Broadcasting Systems . In 1971 , he was awarded the 17 Order of the Rising Sun

1: 183897 Empire_of_Japan

2: 6792932 Nagano, Nagano

3: 174133 Takahashi_Korekiyo

4: 69980 Second_Sino-Japanese_War

5: 17401 Fumimaro_Konoe

6: 1840005 Reorganized_National_Government_of_China

7: 16881276 Greater_East_Asia_Conference

8: 25734 Taiwan

9: 971312 Chūō_Expressway

1: 1326934 Minister_of_Finance_(Japan)

[winner: 354563 Finance_minister, pem_pos=12, final_scores=0.20904385, final_local_scores=0.23245168, pem=-3.2094803, similarity=28.480188]

gt: 1326934 Minister_of_Finance_(Japan), pem_pos=17, final_scores=-0.37280023, final_local_scores=-0.38398564, pem=-3.2837787, similarity=10.488514]

[winner: 2634434 Shinzō_Abe, pem_pos=0,

final_scores=-1.270005, final_local_scores=-1.4346056,
 pem=0.0, similarity=-38.72178|
 RECALL MISS: 1326934 Minister_of_Finance_(Japan)]
 2: 1618940 Ministry_of_Greater_East_Asia
 [winner: 244149 Greater_East_Asia_Co-Prosperity_Sphere,
 pem_pos=0, final_scores=0.19724949,
 final_local_scores=0.19576758, pem=-0.12648417,
 similarity=10.619649|
 gt: 1618940 Ministry_of_Greater_East_Asia, pem_pos=2,
 final_scores=-0.7108192, final_local_scores=-0.7434245,
 pem=-4.6151204, similarity=7.0021873]
 [winner: 2634434 Shinzō_Abe, pem_pos=0,
 final_scores=-1.1016766, final_local_scores=-1.2507324,
 pem=0.0, similarity=-33.23464|
 RECALL MISS: 1618940 Ministry_of_Greater_East_Asia]
 3: 624572 Sarashina_District,_Nagano
 [winner: 624572 Sarashina_District,_Nagano, pem_pos=0,
 final_scores=-0.07500696, final_local_scores=0.029188827,
 pem=0.0, similarity=4.960766|
 gt: 624572 Sarashina_District,_Nagano, pem_pos=0,
 final_scores=-0.07500696, final_local_scores=0.029188827,
 pem=0.0, similarity=4.960766]
 4: 870389 Ministry_of_Finance_(Japan)
 [winner: 870389 Ministry_of_Finance_(Japan), pem_pos=3,
 final_scores=-0.027903482, final_local_scores=-0.04706952,
 pem=-2.2930157, similarity=15.15484|
 gt: 870389 Ministry_of_Finance_(Japan), pem_pos=3,
 final_scores=-0.027903482, final_local_scores=-0.04706952,
 pem=-2.2930157, similarity=15.15484]
 5: 15254732 National_Mobilization_Law
 [winner: 15254732 National_Mobilization_Law, pem_pos=0,
 final_scores=-0.10497923, final_local_scores=-0.07133597,
 pem=0.0, similarity=1.9609052|
 gt: 15254732 National_Mobilization_Law, pem_pos=0,
 final_scores=-0.10497923, final_local_scores=-0.07133597,
 pem=0.0, similarity=1.9609052]
 6: 562666 War_economy
 [winner: 562666 War_economy, pem_pos=0,
 final_scores=-0.8456814, final_local_scores=-0.8874118,
 pem=0.0, similarity=-22.392424|
 gt: 562666 War_economy, pem_pos=0,
 final_scores=-0.8456814, final_local_scores=-0.8874118,
 pem=0.0, similarity=-22.392424]
 7: 24833 Prime_Minister_of_Japan
 [winner: 2634434 Shinzō_Abe, pem_pos=0,
 final_scores=-0.36584872, final_local_scores=-0.44695234,
 pem=0.0, similarity=-9.248237|
 RECALL MISS: 24833 Prime_Minister_of_Japan]
 [winner: 2634434 Shinzō_Abe, pem_pos=0,
 final_scores=-1.0032741, final_local_scores=-1.1432426,
 pem=0.0, similarity=-30.026924|
 RECALL MISS: 24833 Prime_Minister_of_Japan]

8: 30639531 House_of_Peers_(Japan)
[winner: 538644 Upper_house, pem_pos=0,
final_scores=-0.031848222, final_local_scores=0.021276146,
pem=0.0, similarity=4.7246356|
gt: 30639531 House_of_Peers_(Japan), pem_pos=7,
final_scores=-0.8049168, final_local_scores=-0.90911055,
pem=-4.96695, similarity=3.9710865]
9: 174160 Nobuyuki_Abe
[winner: 2634434 Shinzō_Abe, pem_pos=6,
final_scores=0.047012374, final_local_scores=0.0104801655,
pem=-3.3039014, similarity=22.36959|
gt: 174160 Nobuyuki_Abe, pem_pos=9, final_scores=-0.166,
final_local_scores=-0.19003159, pem=-3.4293873,
similarity=17.068333]
10: 1618940 Ministry_of_Greater_East_Asia
[winner: 244149 Greater_East_Asia_Co-Prosperity_Sphere,
pem_pos=0, final_scores=0.19007306,
final_local_scores=0.18792844, pem=-0.12648417,
similarity=10.385713|
gt: 1618940 Ministry_of_Greater_East_Asia, pem_pos=2,
final_scores=-0.7427214, final_local_scores=-0.7782728,
pem=-4.6151204, similarity=5.962242]
[winner: 2634434 Shinzō_Abe, pem_pos=0,
final_scores=-1.1391163, final_local_scores=-1.2916296,
pem=0.0, similarity=-34.455086|
RECALL MISS: 1618940 Ministry_of_Greater_East_Asia]
11: 483209 Sugamo_Prison
[winner: 483209 Sugamo_Prison, pem_pos=0,
final_scores=-0.0022689253, final_local_scores=0.029297099,
pem=0.0, similarity=4.963997|
gt: 483209 Sugamo_Prison, pem_pos=0,
final_scores=-0.0022689253, final_local_scores=0.029297099,
pem=0.0, similarity=4.963997]
12: 48626 War_crime
[winner: 48626 War_crime, pem_pos=0,
final_scores=-0.9853721, final_local_scores=-1.0591729,
pem=-0.031808812, similarity=-27.345139|
gt: 48626 War_crime, pem_pos=0, final_scores=-0.9853721,
final_local_scores=-1.0591729, pem=-0.031808812,
similarity=-27.345139]
13: 24203178 Liberal_Party_(Japan,_1945)
[winner: 406013 Colombian_Liberal_Party, pem_pos=12,
final_scores=-0.14089374, final_local_scores=-0.09480345,
pem=-3.7501097, similarity=21.654268|
RECALL MISS: 24203178 Liberal_Party_(Japan,_1945)]
[winner: 15573 Japan, pem_pos=0,
final_scores=-0.16724715, final_local_scores=-0.24630567,
pem=-0.13549899, similarity=-2.5236754|
RECALL MISS: 24203178 Liberal_Party_(Japan,_1945)]
[winner: 34624 1945, pem_pos=0, final_scores=-0.79362637,
final_local_scores=-0.7137378, pem=-0.45536914,
similarity=-14.733278|

RECALL MISS: 24203178 Liberal_Party_(Japan,_1945)]
 [winner: 5481852 Xi_baryon, pem_pos=11,
 final_scores=-1.9072516, final_local_scores=-1.8968561,
 pem=-4.6728287, similarity=-27.104685]
 RECALL MISS: 24203178 Liberal_Party_(Japan,_1945)]
 14: 164231 Liberal_Democratic_Party_(Japan)
 [winner: 164231 Liberal_Democratic_Party_(Japan),
 pem_pos=0, final_scores=0.32610053,
 final_local_scores=0.31068408, pem=-0.41367084,
 similarity=15.610752]
 gt: 164231 Liberal_Democratic_Party_(Japan), pem_pos=0,
 final_scores=0.32610053, final_local_scores=0.31068408,
 pem=-0.41367084, similarity=15.610752]
 [winner: 164231 Liberal_Democratic_Party_(Japan),
 pem_pos=0, final_scores=0.10692127,
 final_local_scores=0.07126406, pem=-0.76990485,
 similarity=10.403236]
 gt: 164231 Liberal_Democratic_Party_(Japan), pem_pos=0,
 final_scores=0.10692127, final_local_scores=0.07126406,
 pem=-0.76990485, similarity=10.403236]
 15: 7392554 Shin-etsu_Broadcasting
 [winner: 7392554 Shin-etsu_Broadcasting, pem_pos=0,
 final_scores=-0.072364286, final_local_scores=0.043848753,
 pem=0.0, similarity=5.3982472]
 gt: 7392554 Shin-etsu_Broadcasting, pem_pos=0,
 final_scores=-0.072364286, final_local_scores=0.043848753,
 pem=0.0, similarity=5.3982472]
 16: 7217448 Nagano_Broadcasting_Systems
 [winner: 7217448 Nagano_Broadcasting_Systems, pem_pos=0,
 final_scores=-0.08204216, final_local_scores=0.014458507,
 pem=0.0, similarity=4.5211835]
 gt: 7217448 Nagano_Broadcasting_Systems, pem_pos=0,
 final_scores=-0.08204216, final_local_scores=0.014458507,
 pem=0.0, similarity=4.5211835]
 17: 1193966 Order_of_the_Rising_Sun
 [winner: 1193966 Order_of_the_Rising_Sun, pem_pos=0,
 final_scores=-0.097501904, final_local_scores=-0.083409145,
 pem=-0.013818712, similarity=1.675766]
 gt: 1193966 Order_of_the_Rising_Sun, pem_pos=0,
 final_scores=-0.097501904, final_local_scores=-0.083409145,
 pem=-0.013818712, similarity=1.675766]
 1: 183897 Empire_of_Japan
 [pem_pos=0, final_scores=0.15396287,
 final_local_scores=0.118394464, pem=0.0, similarity=7.62284]
 2: 6792932 Nagano,_Nagano
 [pem_pos=0, final_scores=0.14047854,
 final_local_scores=0.17824754, pem=-0.45359597,
 similarity=11.875701]
 3: 174133 Takahashi_Korekiyo
 [pem_pos=0, final_scores=0.3735023,
 final_local_scores=0.40871537, pem=0.0, similarity=16.286594]
 4: 69980 Second_Sino-Japanese_War

[pem_pos=0, final_scores=0.3156914,
final_local_scores=0.30206254, pem=0.0, similarity=13.103862]
5: 17401 Fumimaro_Konoe
[pem_pos=0, final_scores=0.5224581,
final_local_scores=0.5244153, pem=0.0, similarity=19.739311]
6: 1840005 Reorganized_National_Government_of_China
[pem_pos=0, final_scores=0.17402653,
final_local_scores=0.171444, pem=0.0, similarity=9.205944]
7: 16881276 Greater_East_Asia_Conference
[pem_pos=0, final_scores=0.20394753,
final_local_scores=0.24803719, pem=0.0, similarity=11.491637]
8: 25734 Taiwan
[pem_pos=0, final_scores=0.40039152,
final_local_scores=0.38976783, pem=-0.077886656,
similarity=16.144722]
9: 971312 Chūō_Expressway
[pem_pos=0, final_scores=0.034936085,
final_local_scores=0.1333946, pem=0.0, similarity=8.070474]
1: 354563 Finance_minister
[pem_pos=12, final_scores=0.20904385,
final_local_scores=0.23245168, pem=-3.2094803,
similarity=28.480188]
2: 244149 Greater_East_Asia_Co-Prosperity_Sphere
[pem_pos=0, final_scores=0.19724949,
final_local_scores=0.19576758, pem=-0.12648417,
similarity=10.619649]
3: 15573 Japan
[pem_pos=0, final_scores=0.709457,
final_local_scores=0.7166356, pem=-0.13549899,
similarity=26.212412]
4: 15573 Japan
[pem_pos=0, final_scores=0.7294155,
final_local_scores=0.7384371, pem=-0.99076414,
similarity=31.514078]
5: 163847 National_Diet
[pem_pos=0, final_scores=0.20883451,
final_local_scores=0.173946, pem=-0.6489993,
similarity=12.809967]
6: 2634434 Shinzō_Abe
[pem_pos=6, final_scores=0.047012374,
final_local_scores=0.0104801655, pem=-3.3039014,
similarity=22.36959]
7: 354563 Finance_minister
[pem_pos=12, final_scores=0.20967339,
final_local_scores=0.23313934, pem=-3.2094803,
similarity=28.50071]
8: 2634434 Shinzō_Abe
[pem_pos=6, final_scores=0.055150762,
final_local_scores=0.019370139, pem=-3.3039014,
similarity=22.634886]
9: 15573 Japan
[pem_pos=0, final_scores=0.57636005,

final_local_scores=0.5712474, pem=-0.13549899,
 similarity=21.87374]
 10: 244149 Greater_East_Asia_Co-Prosperity_Sphere
 [pem_pos=0, final_scores=0.19007306,
 final_local_scores=0.18792844, pem=-0.12648417,
 similarity=10.385713]
 11: 32927 World_War_II
 [pem_pos=0, final_scores=0.39895362,
 final_local_scores=0.41552693, pem=0.0, similarity=16.489864]
 12: 164231 Liberal_Democratic_Party_(Japan)
 [pem_pos=0, final_scores=0.32610053,
 final_local_scores=0.31068408, pem=-0.41367084,
 similarity=15.610752]
 13: 164231 Liberal_Democratic_Party_(Japan)
 [pem_pos=0, final_scores=0.10692127,
 final_local_scores=0.07126406, pem=-0.76990485,
 similarity=10.403236]
 14: 164231 Liberal_Democratic_Party_(Japan)
 [pem_pos=0, final_scores=0.4376341,
 final_local_scores=0.4325176, pem=-0.76990485,
 similarity=21.183764]
 15: 15573 Japan
 [pem_pos=0, final_scores=0.6378297,
 final_local_scores=0.63839364, pem=-0.13549899,
 similarity=23.877516]
 16: 15573 Japan
 [pem_pos=0, final_scores=0.62819105,
 final_local_scores=0.62786484, pem=-0.13549899,
 similarity=23.563316]
 17: 15573 Japan
 [pem_pos=0, final_scores=0.63557035,
 final_local_scores=0.6359256, pem=-0.13549899,
 similarity=23.803865]

Sample: b'Law'

_ & _ Order _ -LRB- season_16 -RRB- The 16th season of [1 1 L
 aw & Order 1] premiered on [2 2 NBC 2] September 21 , 2005 ,
 and concluded on May 17 , 2006 . [3 3 Jesse L. Martin 3]
 returned to his role of 1 Detective [1 Ed Green 1] in the
 season premiere episode , he left at the end of 2 last season
 to film in the movie 3 Rent . Mid-season , [2 Law & Order 2]
 's future on [3 NBC 3] was in the air , unknown if the show
 would be renewed or canceled after this season . Prior to the
 end of the season , [4 4 Annie Parisse 4] who portrayed [5 5 A
 DA Borgia 5] , quit the series , wanting out out of her
 contract . A show insider said to , `` She saw the writing on
 the wall , they never treated her very well . They were
 always complaining about her hair . And they also thought she
 looked too young next to [6 6 Sam Waterson 6] -LRB- sic -
 RRB- . '' Pariesse 's ADA [4 Alex Borgia 4] was brutally
 killed off in the season finale episode ' ' ' ' . It was
 announced that [7 7 Dennis Farina 7] -LRB- -RRB- wished to

depart the cast weeks after the season finale episode aired . At the time Farina wished to pursue other offers and projects being developed by his production company . [8 8 Dick Wolf 8] said in a statement that he respected [5 Farina 5] 's decision and looked forward to working with him again . There were `` a lot '' of firings , including a portion of the camera crew . `` None of the original writer-producers are there anymore , '' a source said to [6 FOX News 6] . `` They 've all been replaced by people from [9 9 L.A. 9] who do n't get the show . '' [7 People 7] also had issues with [8 Dennis Farina 8] coming in to replace [10 10 Jerry Orbach 10] in the 4 previous season . 5 Walon Green was show runner/executive producer this season with [11 11 Nicholas Wootton 11] , over 6 Matthew Penn , and [12 12 Peter Jankowski 12] . Green stepped down as show runner/EP at the end of the season and Wootton took over in the 7 17th season

1: 173629 Law_&_Order
2: 21780 NBC
3: 1107510 Jesse_L._Martin
4: 1342164 Annie_Parisse
5: 4022490 Alexandra_Borgia
6: 168518 Sam_Waterston
7: 809289 Dennis_Farina
8: 1069577 Dick_Wolf
9: 18110 Los_Angeles
10: 584551 Jerry_Orbach
11: 16326968 Nicholas_Wootton
12: 7657288 Peter_Jankowski

1: 2675595 Ed_Green

[winner: 2675595 Ed_Green, pem_pos=0,
final_scores=0.07837145, final_local_scores=0.09197466,
pem=0.0, similarity=6.83442|

gt: 2675595 Ed_Green, pem_pos=0,
final_scores=0.07837145, final_local_scores=0.09197466,
pem=0.0, similarity=6.83442]

[winner: 9912458 Detective_(1954_film), pem_pos=9,
final_scores=-0.7946764, final_local_scores=-0.6827545,
pem=-3.0910425, similarity=0.524525|

RECALL MISS: 2675595 Ed_Green]

2: 2618251 Law_&_Order_(season_15)

[winner: 927001 Last, pem_pos=0, final_scores=-1.2044399,
final_local_scores=-1.1082081, pem=-1.3263657,
similarity=-21.768444|

RECALL MISS: 2618251 Law_&_Order_(season_15)]

[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.3919225, final_local_scores=-1.420018,
pem=-1.219945, similarity=-31.652208|

RECALL MISS: 2618251 Law_&_Order_(season_15)]

3: 1505118 Rent_(film)

[winner: 26436 Rent_(musical), pem_pos=0,
final_scores=-0.1905178, final_local_scores=-0.23494267,
pem=-0.12806667, similarity=-2.2249987|

```

    gt: 1505118 Rent_(film), pem_pos=1,
    final_scores=-0.4121493, final_local_scores=-0.47010463,
    pem=-1.1739928, similarity=-3.554792 ]
4: 2618251 Law_&_Order_(season_15)
    [winner: 2550090 This_Year's_Girl, pem_pos=1,
    final_scores=-0.85729575, final_local_scores=-0.8637584,
    pem=0.0, similarity=-21.68656]
    RECALL MISS: 2618251 Law_&_Order_(season_15)]
    [winner: 690345 Season_(sports), pem_pos=0,
    final_scores=-1.4702872, final_local_scores=-1.5056195,
    pem=-1.219945, similarity=-34.20673]
    RECALL MISS: 2618251 Law_&_Order_(season_15)]
5: 2593121 Walon_Green
    [winner: 2593121 Walon_Green, pem_pos=0,
    final_scores=-0.07878493, final_local_scores=-0.08088706,
    pem=0.0, similarity=1.6758819]
    gt: 2593121 Walon_Green, pem_pos=0,
    final_scores=-0.07878493, final_local_scores=-0.08088706,
    pem=0.0, similarity=1.6758819 ]
6: 32853412 Matthew_Penn
    [winner: 32853412 Matthew_Penn, pem_pos=0,
    final_scores=-0.09338395, final_local_scores=-0.064373635,
    pem=0.0, similarity=2.1686752]
    gt: 32853412 Matthew_Penn, pem_pos=0,
    final_scores=-0.09338395, final_local_scores=-0.064373635,
    pem=0.0, similarity=2.1686752 ]
7: 6714495 Law_&_Order_(season_17)
    [winner: 6714495 Law_&_Order_(season_17), pem_pos=5,
    final_scores=-0.9367827, final_local_scores=-0.9446905,
    pem=-2.7080503, similarity=-9.374939]
    gt: 6714495 Law_&_Order_(season_17), pem_pos=5,
    final_scores=-0.9367827, final_local_scores=-0.9446905,
    pem=-2.7080503, similarity=-9.374939 ]
1: 173629 Law_&_Order
    [pem_pos=0, final_scores=0.29015094,
    final_local_scores=0.2309624, pem=0.0, similarity=10.982092]
2: 21780 NBC
    [pem_pos=0, final_scores=0.48011762,
    final_local_scores=0.44706923, pem=-0.06403453,
    similarity=17.779379]
3: 1107510 Jesse_L._Martin
    [pem_pos=0, final_scores=0.22944234,
    final_local_scores=0.21555969, pem=0.0, similarity=10.522444]
4: 1342164 Annie_Parisse
    [pem_pos=0, final_scores=0.18962012,
    final_local_scores=0.19368005, pem=0.0, similarity=9.869512]
5: 4022490 Alexandra_Borgia
    [pem_pos=0, final_scores=0.058197126,
    final_local_scores=0.1176894, pem=0.0, similarity=7.601799]
6: 168518 Sam_Waterston
    [pem_pos=0, final_scores=0.06821011,
    final_local_scores=0.032369018, pem=0.0, similarity=5.055669]

```

7: 809289 Dennis_Farina
 [pem_pos=0, final_scores=0.14541315,
 final_local_scores=0.111982554, pem=0.0, similarity=7.4314957]
 8: 1069577 Dick_Wolf
 [pem_pos=0, final_scores=0.2643978,
 final_local_scores=0.23434925, pem=0.0, similarity=11.083162]
 9: 18110 Los_Angeles
 [pem_pos=0, final_scores=0.41654032,
 final_local_scores=0.4335696, pem=0.0, similarity=17.028296]
 10: 584551 Jerry_Orbach
 [pem_pos=0, final_scores=0.20528392,
 final_local_scores=0.18221474, pem=0.0, similarity=9.527364]
 11: 16326968 Nicholas_Wootton
 [pem_pos=0, final_scores=0.040616766,
 final_local_scores=0.05471848, pem=0.0, similarity=5.7226214]
 12: 7657288 Peter_Jankowski
 [pem_pos=0, final_scores=0.05769591,
 final_local_scores=0.11658117, pem=0.0, similarity=7.5687275]
 1: 2675595 Ed_Green
 [pem_pos=0, final_scores=0.07837145,
 final_local_scores=0.09197466, pem=0.0, similarity=6.83442]
 2: 173629 Law_&_Order
 [pem_pos=0, final_scores=0.25356883,
 final_local_scores=0.19100195, pem=0.0, similarity=9.789592]
 3: 21780 NBC
 [pem_pos=0, final_scores=0.50446194,
 final_local_scores=0.47366178, pem=-0.06403453,
 similarity=18.572954]
 4: 4022490 Alexandra_Borgia
 [pem_pos=0, final_scores=0.101925656,
 final_local_scores=0.16545618, pem=0.0, similarity=9.027255]
 5: 809289 Dennis_Farina
 [pem_pos=0, final_scores=0.036484763,
 final_local_scores=-0.01557634, pem=0.0, similarity=3.624884]
 6: 11121 Fox_News_Channel
 [pem_pos=0, final_scores=0.17384331,
 final_local_scores=0.18393788, pem=0.0, similarity=9.578786]
 7: 507970 People_(magazine)
 [pem_pos=0, final_scores=0.077428,
 final_local_scores=0.0746693, pem=-0.5865796,
 similarity=9.507905]
 8: 809289 Dennis_Farina
 [pem_pos=0, final_scores=0.07378723,
 final_local_scores=0.0337421, pem=0.0, similarity=5.0966444]

Sample: b'Love.'

_ & _ Life _ -LRB- Mary_J __ Blige_album -RRB- ` Love & Life
 ' is the sixth studio album by American 1 R&B singer-
 songwriter [1 1 Mary J. Blige 1] , released on August 26 ,
 2003 by [2 2 Geffen Records 2] . This album marks the first
 collaboration with former one-time producer 2 Sean `` [1 Diddy
 1] ' ' Combs since he produced her groundbreaking debut , 3 Wh

at 's the 411 ? -LRB- 1992 -RRB- and her critically acclaimed second album 4 My Life -LRB- 1994 -RRB- , which both went on to sell more than three million copies worldwide , going 3x by the [3 3 Recording Industry Association of America 3] -LRB- RIAA -RRB- . The album received mixed to positive reviews from most music critics , feeling that [2 Blige 2] and [3 Combs 3] failed to recreate the `` magic '' and the essence that was shown in 5 her first and 6 second album . However , the album was commercially successful worldwide . It became [4 Blige 4] 's second number-one album on the [4 4 U S Billboard 200 4] , selling 285,298 copies in its first week sales . The album also debuted in the top ten on the charts in countries such as [5 7 Canada 5] , , the [6 8 United Kingdom 6] , with similar charting in and , where it charted in top twenty . It also opened at the top spot on the 9 R&B / Hip-Hop Albums chart respectively , making this [7 Blige 7] 's sixth album to top the latter chart . The album earned [8 Blige 8] two nominations at the for and for '' 10 Ooh ! '' . The album was eventually certified by the for shipping one million copies , and more than two million copies worldwide . Love & Life received a score of 72 out of 100 from [5 5 Metacritic 5] based on generally favorable reviews . [9 Allmusic 9] gave it three-and-a-half stars out of five and said the album `` beamed with joy '' , and [6 6 Rolling Stone 6]

1: 214210 Mary_J._Blige

2: 54975 Geffen_Records

3: 18949819 Recording_Industry_Association_of_America

4: 730350 Billboard_200

5: 805870 Metacritic

6: 25441 Rolling_Stone

1: 2529904 Contemporary_R&B

[winner: 26168 Rhythm_and_blues, pem_pos=0, final_scores=-0.03088905, final_local_scores=-0.106425, pem=-0.42558512, similarity=3.2281768]

gt: 2529904 Contemporary_R&B, pem_pos=1, final_scores=-0.1843374, final_local_scores=-0.26401696, pem=-1.2034804, similarity=2.7556338]

2: 152447 Sean_Combs

[winner: 152447 Sean_Combs, pem_pos=0, final_scores=0.25170308, final_local_scores=0.19706723, pem=0.0, similarity=9.9705925]

gt: 152447 Sean_Combs, pem_pos=0, final_scores=0.25170308, final_local_scores=0.19706723, pem=0.0, similarity=9.9705925]

[winner: 171410 Sean_Lennon, pem_pos=0, final_scores=-0.19363931, final_local_scores=-0.23735687, pem=-1.5740016, similarity=5.5661755]

gt: 152447 Sean_Combs, pem_pos=15, final_scores=-0.5193416, final_local_scores=-0.65084535, pem=-4.655769, similarity=9.985985]

[winner: 395604 Ray_Combs, pem_pos=0, final_scores=-0.26740348, final_local_scores=-0.25949764,

pem=-1.1549822, similarity=2.6267588|
 gt: 152447 Sean_Combs, pem_pos=8,
 final_scores=-0.4402249, final_local_scores=-0.56442237,
 pem=-3.0460322, similarity=3.8110201]
 [winner: 59349 Quotation_mark, pem_pos=0,
 final_scores=-2.0627618, final_local_scores=-2.1256928,
 pem=0.0, similarity=-59.345203]
 RECALL MISS: 152447 Sean_Combs]
 3: 215772 What's_the_411?
 [winner: 215772 What's_the_411?, pem_pos=0,
 final_scores=-0.42044914, final_local_scores=-0.43684047,
 pem=0.0, similarity=-8.946478|
 gt: 215772 What's_the_411?, pem_pos=0,
 final_scores=-0.42044914, final_local_scores=-0.43684047,
 pem=0.0, similarity=-8.946478]
 4: 5549830 My_Life_(Mary_J._Blige_album)
 [winner: 4234464 My_Life_(Bill_Clinton_autobiography),
 pem_pos=0, final_scores=-0.253922,
 final_local_scores=-0.21625926, pem=-1.231726,
 similarity=4.334423|
 gt: 5549830 My_Life_(Mary_J._Blige_album), pem_pos=2,
 final_scores=-0.34713417, final_local_scores=-0.39164492,
 pem=-1.6674998, similarity=1.4703702]
 5: 215772 What's_the_411?
 [winner: 14914 Industrial_Revolution, pem_pos=0,
 final_scores=-1.4013392, final_local_scores=-1.4203293,
 pem=0.0, similarity=-38.29575|
 RECALL MISS: 215772 What's_the_411?]
 6: 5549830 My_Life_(Mary_J._Blige_album)
 [winner: 1939689 Nattens_madriral, pem_pos=2,
 final_scores=-1.0854872, final_local_scores=-1.0349383,
 pem=-2.6505456, similarity=-12.380832|
 RECALL MISS: 5549830 My_Life_(Mary_J._Blige_album)]
 7: 9331428 Canadian_Albums_Chart
 [winner: 5042916 Canada, pem_pos=0,
 final_scores=0.64625025, final_local_scores=0.7350934,
 pem=-0.11662793, similarity=26.660606|
 RECALL MISS: 9331428 Canadian_Albums_Chart]
 8: 2465857 UK_Albums_Chart
 [winner: 31717 United_Kingdom, pem_pos=0,
 final_scores=0.40258163, final_local_scores=0.44435734,
 pem=-0.09933331, similarity=17.890413|
 RECALL MISS: 2465857 UK_Albums_Chart]
 9: 8614434 Top_R&B/Hip-Hop_Albums
 [winner: 26168 Rhythm_and_blues, pem_pos=0,
 final_scores=-0.22561969, final_local_scores=-0.31913865,
 pem=-0.42558512, similarity=-3.1196227|
 gt: 8614434 Top_R&B/Hip-Hop_Albums, pem_pos=3,
 final_scores=-0.37602824, final_local_scores=-0.46002346,
 pem=-1.5949938, similarity=-0.9644801]
 [winner: 528282 Album, pem_pos=0,
 final_scores=-1.0232416, final_local_scores=-1.1775379,

pem=-0.85669935, similarity=-26.391504|
RECALL MISS: 8614434 Top_R&B/Hip-Hop_Albums]
[winner: 3555227 Hip_hop_production, pem_pos=3,
final_scores=-1.4206278, final_local_scores=-1.6051121,
pem=-1.6084737, similarity=-35.0629|
RECALL MISS: 8614434 Top_R&B/Hip-Hop_Albums]
[winner: 59352 Slash_(punctuation), pem_pos=0,
final_scores=-1.517988, final_local_scores=-1.5188694,
pem=-0.015267457, similarity=-41.153355|
RECALL MISS: 8614434 Top_R&B/Hip-Hop_Albums]
10: 3312781 Ooh!
[winner: 3312781 Ooh!, pem_pos=0,
final_scores=-0.18886364, final_local_scores=-0.14401019,
pem=0.0, similarity=-0.20783815|
gt: 3312781 Ooh!, pem_pos=0, final_scores=-0.18886364,
final_local_scores=-0.14401019, pem=0.0,
similarity=-0.20783815]
[winner: 3243047 !_(album), pem_pos=3,
final_scores=-1.3962048, final_local_scores=-1.4147575,
pem=-2.4849067, similarity=-24.616163|
RECALL MISS: 3312781 Ooh!]
1: 214210 Mary_J._Blige
[pem_pos=0, final_scores=0.3598175,
final_local_scores=0.30008575, pem=0.0, similarity=13.04487]
2: 54975 Geffen_Records
[pem_pos=0, final_scores=0.22087924,
final_local_scores=0.18365437, pem=0.0, similarity=9.570326]
3: 18949819 Recording_Industry_Association_of_America
[pem_pos=0, final_scores=0.1632406,
final_local_scores=0.1316422, pem=0.0, similarity=8.018179]
4: 730350 Billboard_200
[pem_pos=0, final_scores=0.187146,
final_local_scores=0.14714149, pem=0.0, similarity=8.480708]
5: 805870 Metacritic
[pem_pos=0, final_scores=0.035631016,
final_local_scores=0.08561742, pem=0.0, similarity=6.6447077]
6: 25441 Rolling_Stone
[pem_pos=0, final_scores=0.2545116,
final_local_scores=0.22973979, pem=0.0, similarity=10.945606]
1: 152447 Sean_Combs
[pem_pos=0, final_scores=0.25170308,
final_local_scores=0.19706723, pem=0.0, similarity=9.9705925]
2: 214210 Mary_J._Blige
[pem_pos=0, final_scores=0.32952386,
final_local_scores=0.2669945, pem=0.0, similarity=12.057362]
3: 152447 Sean_Combs
[pem_pos=0, final_scores=0.29689032,
final_local_scores=0.24806911, pem=-0.0139488345,
similarity=11.568446]
4: 214210 Mary_J._Blige
[pem_pos=0, final_scores=0.31085503,
final_local_scores=0.2466017, pem=0.0, similarity=11.448799]

5: 5042916 Canada
 [pem_pos=0, final_scores=0.64625025,
 final_local_scores=0.7350934, pem=-0.11662793,
 similarity=26.660606]
 6: 31717 United_Kingdom
 [pem_pos=0, final_scores=0.40258163,
 final_local_scores=0.44435734, pem=-0.09933331,
 similarity=17.890413]
 7: 214210 Mary_J._Blige
 [pem_pos=0, final_scores=0.34972197,
 final_local_scores=0.28905794, pem=0.0, similarity=12.715778]
 8: 214210 Mary_J._Blige
 [pem_pos=0, final_scores=0.3000102,
 final_local_scores=0.23475537, pem=0.0, similarity=11.095282]
 9: 697535 AllMusic
 [pem_pos=0, final_scores=0.09011443,
 final_local_scores=0.08838625, pem=0.0, similarity=6.727335]

Sample: b'Marcus_Thomas_Pius_Gilbert'

Marcus Thomas Pius Gilbert -LRB- also known as Tom Gilbert ,
 and publishing as MTP Gilbert -RRB- is an . He received a BA
 at , , and a D.Phil -LRB- **1 Doctor of Philosophy** -RRB- in the
 Zoology Dept and at **[1 1 New College 1]** under Alan Cooper .
 Subsequently , he was a post-doctoral fellow at the
 Department of Ecology and Evolutionary Biology , at the **[2 2 U
 niversity of Arizona 2]** . In 2005 , he became an Assistant
 Professor at the **[3 3 University of Copenhagen 3]** , where he
 is currently Professor of Palaeogenomics at the **[4 4 Universi
 ty of Copenhagen 4]** 's **[1 2 Natural History Museum of Denmark
 1]** , and a member of the **[5 5 Centre for Geogenetics 5]** at .
 He is also Adjunct Professor of **[6 6 Murdoch University 6]** -
 LRB- Perth , **[2 Australia 2]** -RRB- , and an Associate Editor
 of the Journals , , and . He is notable for his work on **3 anc
 ient DNA** , having several publications in journals such as **[3
 4 Science 3]** , **[4 5 Nature 4]** and . He was involved in the
 sequencing of DNA from Pre - **6 clovis 7 native american s** .
 He is currently coordinator of the EUROFAST **8 Marie Curie Act
 ions** Training Network , an interdisciplinary research program
 of 13 PhD students based at 10 international institutions ,
 aimed at applying the latest in research techniques to
 helping contribute to modern understanding of the Trans -LRB-
 www.eurofast.eu -RRB- . Outside of bioarchaeology , his other
 interests include phylogenomics with particular focus on and
 the **9 giant squid**

1: 128354 New_College,_Oxford
 2: 32015 University_of_Arizona
 3: 176767 University_of_Copenhagen
 4: 176767 University_of_Copenhagen
 5: 34492330 Centre_for_Geogenetics
 6: 285107 Murdoch_University

1: 21031297 Doctor_of_Philosophy
 [winner: 21031297 Doctor_of_Philosophy, pem_pos=0,

final_scores=-0.34290951, final_local_scores=-0.43374065,
pem=0.0, similarity=-8.853973|
gt: 21031297 Doctor_of_Philosophy, pem_pos=0,
final_scores=-0.34290951, final_local_scores=-0.43374065,
pem=0.0, similarity=-8.853973]
2: 34923725 Natural_History_Museum_of_Denmark
[winner: 4206906
University_of_Copenhagen_Botanical_Garden, pem_pos=1,
final_scores=0.05729626, final_local_scores=0.124691874,
pem=0.0, similarity=7.810767|
gt: 34923725 Natural_History_Museum_of_Denmark,
pem_pos=0, final_scores=-0.008666098,
final_local_scores=0.07892527, pem=0.0, similarity=6.445]
3: 3285684 Ancient_DNA
[winner: 7955 DNA, pem_pos=0, final_scores=-0.09327726,
final_local_scores=-0.03854856, pem=-0.054993372,
similarity=3.2384095|
RECALL MISS: 3285684 Ancient_DNA]
[winner: 148363 Ancient_Greek, pem_pos=4,
final_scores=-1.5523947, final_local_scores=-1.6363492,
pem=-1.8986653, similarity=-34.416977|
gt: 3285684 Ancient_DNA, pem_pos=7,
final_scores=-1.652827, final_local_scores=-1.7322791,
pem=-2.0689092, similarity=-36.3539]
4: 26700 Science
[winner: 193513 Science_(journal), pem_pos=0,
final_scores=0.12834947, final_local_scores=0.14016798,
pem=-1.060998, similarity=14.042477|
gt: 26700 Science, pem_pos=1, final_scores=0.022217497,
final_local_scores=-0.005628854, pem=-1.0814888,
similarity=9.803041]
5: 21830 Nature
[winner: 43427 Nature_(journal), pem_pos=0,
final_scores=0.2760666, final_local_scores=0.29030132,
pem=-0.4031761, similarity=14.945417|
gt: 21830 Nature, pem_pos=13, final_scores=-0.62655336,
final_local_scores=-0.6483638, pem=-4.4107924,
similarity=8.7278185]
6: 696591 Clovis_culture
[winner: 555204 Clovis_IV, pem_pos=10,
final_scores=-1.1680766, final_local_scores=-1.0363178,
pem=-2.8903718, similarity=-11.117786|
gt: 696591 Clovis_culture, pem_pos=0,
final_scores=-1.2003092, final_local_scores=-1.236016,
pem=-1.0005511, similarity=-27.35432]
7: 2402376 Indigenous_peoples_of_the_Americas
[winner: 3434750 United_States, pem_pos=0,
final_scores=-0.5970291, final_local_scores=-0.6660487,
pem=-0.12994486, similarity=-15.079847|
RECALL MISS: 2402376 Indigenous_peoples_of_the_Americas]
[winner: 2038320 Native_plant, pem_pos=0,
final_scores=-1.1947501, final_local_scores=-1.1665957,

pem=-0.7661968, similarity=-26.557133|
 gt: 2402376 Indigenous_peoples_of_the_Americas,
 pem_pos=4, final_scores=-1.8205779,
 final_local_scores=-1.9568534, pem=-2.436607,
 similarity=-41.05605]
 8: 10454621 Marie_Curie_Actions
 [winner: 15325219 Marie,_Alpes-Maritimes, pem_pos=1,
 final_scores=-0.9055342, final_local_scores=-0.7230141,
 pem=-2.1231313, similarity=-5.940552|
 RECALL MISS: 10454621 Marie_Curie_Actions]
 [winner: 24513955 Curie_(Martian_crater), pem_pos=7,
 final_scores=-1.1531601, final_local_scores=-1.0479848,
 pem=-2.5649493, similarity=-13.23565|
 RECALL MISS: 10454621 Marie_Curie_Actions]
 [winner: 3101692 Tensor_veli_palatini_muscle, pem_pos=11,
 final_scores=-1.6040037, final_local_scores=-1.5404155,
 pem=-4.2754025, similarity=-18.629057|
 RECALL MISS: 10454621 Marie_Curie_Actions]
 9: 215419 Giant_squid
 [winner: 3439284 GIANT_Company_Software, pem_pos=11,
 final_scores=-1.0450671, final_local_scores=-0.9838986,
 pem=-3.096045, similarity=-8.435008|
 RECALL MISS: 215419 Giant_squid]
 [winner: 205620 Squid_(software), pem_pos=1,
 final_scores=-1.0903492, final_local_scores=-1.057202,
 pem=-1.0346134, similarity=-21.832914|
 gt: 215419 Giant_squid, pem_pos=5,
 final_scores=-2.212644, final_local_scores=-2.3483672,
 pem=-5.9476943, similarity=-33.645763]
 1: 128354 New_College,_Oxford
 [pem_pos=0, final_scores=0.03575091,
 final_local_scores=0.04121025, pem=-0.33538523,
 similarity=7.143386]
 2: 32015 University_of_Arizona
 [pem_pos=0, final_scores=0.46852946,
 final_local_scores=0.46555865, pem=0.0, similarity=17.982912]
 3: 176767 University_of_Copenhagen
 [pem_pos=0, final_scores=0.41044837,
 final_local_scores=0.38892835, pem=0.0, similarity=15.696111]
 4: 176767 University_of_Copenhagen
 [pem_pos=0, final_scores=0.3180914,
 final_local_scores=0.28804237, pem=0.0, similarity=12.685472]
 5: 34492330 Centre_for_Geogenetics
 [pem_pos=0, final_scores=0.08525728,
 final_local_scores=0.15470886, pem=0.0, similarity=8.706534]
 6: 285107 Murdoch_University
 [pem_pos=0, final_scores=0.14229707,
 final_local_scores=0.14651343, pem=0.0, similarity=8.461966]
 1: 4206906 University_of_Copenhagen_Botanical_Garden
 [pem_pos=1, final_scores=0.05729626,
 final_local_scores=0.124691874, pem=0.0, similarity=7.810767]
 2: 4689264 Australia

[pem_pos=0, final_scores=0.5006521,
 final_local_scores=0.5791654, pem=-0.17168477,
 similarity=22.306812]
 3: 193513 Science_(journal)
 [pem_pos=0, final_scores=0.12834947,
 final_local_scores=0.14016798, pem=-1.060998,
 similarity=14.042477]
 4: 43427 Nature_(journal)
 [pem_pos=0, final_scores=0.2760666,
 final_local_scores=0.29030132, pem=-0.4031761,
 similarity=14.945417]

Sample: b'R'

___ F ___ Bayford Robert Frederic Bayford 1 OBE 2 KC -LRB- 24
 September 1871 -- 5 June 1951 -RRB- was an [1 1 English 1] 3 b
 arrister , born in [2 2 Westminster 2] , [1 London 1] .
 Bayford was the eldest son of , also an eminent barrister .
 He was educated at [3 3 Eton College 3] and [2 4 Trinity Hall
 2] , [3 Cambridge 3] , rowing in the 1893 and graduating in
 1895 . He was by the [4 4 Inner Temple 4] the same year and
 specialised in 5 probate and 6 divorce cases . He in 1919 and
 became a 7 Benchler in 1925 . From 1929 to 1944 he was 8 Recor
 der of [5 5 Portsmouth 5] and from 1938 to 1947 he was deputy
 chairman of the [6 6 Hampshire 6] Quarter Sessions . During
 the he served as a 9 Divisional Commander with the [7 7 Metro
 politan Special Constabulary 7] , and for this he was
 appointed 10 Officer of the Order of the British Empire

1: 9316 England
 2: 37432 Westminster
 3: 53228 Eton_College
 4: 207885 Inner_Temple
 5: 24207 Portsmouth
 6: 13861 Hampshire
 7: 4787684 Metropolitan_Special_Constabulary
 1: 212182 Order_of_the_British_Empire

[winner: 159539 Off-budget_enterprise, pem_pos=3,
 final_scores=-1.151449, final_local_scores=-1.1093245,
 pem=-1.9459101, similarity=-18.43258]
 gt: 212182 Order_of_the_British_Empire, pem_pos=0,
 final_scores=-1.1824301, final_local_scores=-1.3077598,
 pem=0.0, similarity=-34.936443]

2: 90072 Queen's_Counsel
 [winner: 90072 Queen's_Counsel, pem_pos=0,
 final_scores=-0.40189022, final_local_scores=-0.45191,
 pem=-0.35013807, similarity=-7.492077]
 gt: 90072 Queen's_Counsel, pem_pos=0,
 final_scores=-0.40189022, final_local_scores=-0.45191,
 pem=-0.35013807, similarity=-7.492077]

3: 4848 Barrister
 [winner: 4848 Barrister, pem_pos=1,
 final_scores=-1.2706237, final_local_scores=-1.431884,
 pem=-0.01358637, similarity=-38.566677]

gt: 4848 Barrister, pem_pos=1, final_scores=-1.2706237,
final_local_scores=-1.431884, pem=-0.01358637,
similarity=-38.566677]

4: 149918 Trinity_Hall_Cambridge
[winner: 149918 Trinity_Hall_Cambridge, pem_pos=0,
final_scores=0.3168698, final_local_scores=0.29532543,
pem=0.0, similarity=12.902813]
gt: 149918 Trinity_Hall_Cambridge, pem_pos=0,
final_scores=0.3168698, final_local_scores=0.29532543,
pem=0.0, similarity=12.902813]
[winner: 36995 Cambridge, pem_pos=0,
final_scores=0.29466558, final_local_scores=0.24110669,
pem=-0.751733, similarity=15.372858]
RECALL MISS: 149918 Trinity_Hall_Cambridge]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.6512344, final_local_scores=-1.6553307,
pem=-4.6728287, similarity=-19.897087]
RECALL MISS: 149918 Trinity_Hall_Cambridge]

5: 1043001 Probate
[winner: 18012669 Probate_court, pem_pos=0,
final_scores=-1.0713416, final_local_scores=-1.1532581,
pem=0.0, similarity=-30.32581]
gt: 1043001 Probate, pem_pos=1, final_scores=-1.1826246,
final_local_scores=-1.2798057, pem=-0.03727866,
similarity=-33.89951]

6: 40147 Divorce
[winner: 3223469 Get_(divorce_document), pem_pos=1,
final_scores=-1.2676433, final_local_scores=-1.2686386,
pem=-0.68084836, similarity=-30.066437]
gt: 40147 Divorce, pem_pos=0, final_scores=-1.4933847,
final_local_scores=-1.592191, pem=-0.021743195,
similarity=-43.306202]

7: 6003690 Bencher
[winner: 6003690 Bencher, pem_pos=0,
final_scores=-0.08002631, final_local_scores=-0.08389586,
pem=0.0, similarity=1.5860931]
gt: 6003690 Bencher, pem_pos=0,
final_scores=-0.08002631, final_local_scores=-0.08389586,
pem=0.0, similarity=1.5860931]

8: 2083796 Recorder_(judge)
[winner: 2083796 Recorder_(judge), pem_pos=0,
final_scores=-0.28435838, final_local_scores=-0.3246189,
pem=0.0, similarity=-5.597563]
gt: 2083796 Recorder_(judge), pem_pos=0,
final_scores=-0.28435838, final_local_scores=-0.3246189,
pem=0.0, similarity=-5.597563]

9: 1242107 Police_division
[winner: 1242107 Police_division, pem_pos=0,
final_scores=-0.16386218, final_local_scores=-0.10888606,
pem=0.0, similarity=0.8403355]
gt: 1242107 Police_division, pem_pos=0,
final_scores=-0.16386218, final_local_scores=-0.10888606,

pem=0.0, similarity=0.8403355]
10: 212182 Order_of_the_British_Empire
[winner: 212182 Order_of_the_British_Empire, pem_pos=0,
final_scores=-0.031156436, final_local_scores=-0.05016806,
pem=0.0, similarity=2.592598]
gt: 212182 Order_of_the_British_Empire, pem_pos=0,
final_scores=-0.031156436, final_local_scores=-0.05016806,
pem=0.0, similarity=2.592598]
[winner: 36301328 Officer_(armed_forces), pem_pos=0,
final_scores=-0.41689873, final_local_scores=-0.4868031,
pem=-1.20684, similarity=-3.874479]
RECALL MISS: 212182 Order_of_the_British_Empire]
[winner: 10447732
Results_by_riding_of_the_Canadian_federal_election,_2008,
pem_pos=12, final_scores=-1.4270074,
final_local_scores=-1.3127842, pem=-3.6198866,
similarity=-15.400879]
RECALL MISS: 212182 Order_of_the_British_Empire]
1: 9316 England
[pem_pos=1, final_scores=0.5319885,
final_local_scores=0.506908, pem=-1.3527626,
similarity=26.573393]
2: 37432 Westminster
[pem_pos=0, final_scores=0.31831408,
final_local_scores=0.26569226, pem=-1.0392612,
similarity=17.670164]
3: 53228 Eton_College
[pem_pos=0, final_scores=0.29911786,
final_local_scores=0.24815375, pem=0.0, similarity=11.495115]
4: 207885 Inner_Temple
[pem_pos=0, final_scores=0.15604775,
final_local_scores=0.1205478, pem=0.0, similarity=7.6870995]
5: 24207 Portsmouth
[pem_pos=0, final_scores=0.30541664,
final_local_scores=0.2757002, pem=-0.75559247,
similarity=16.426188]
6: 13861 Hampshire
[pem_pos=0, final_scores=0.19899796,
final_local_scores=0.15008336, pem=-0.3150438,
similarity=10.281757]
7: 4787684 Metropolitan_Special_Constabulary
[pem_pos=0, final_scores=0.057552293,
final_local_scores=0.12891957, pem=0.0, similarity=7.93693]
1: 17867 London
[pem_pos=0, final_scores=0.6862622,
final_local_scores=0.68819386, pem=-0.09086737,
similarity=25.120941]
2: 149918 Trinity_Hall,_Cambridge
[pem_pos=0, final_scores=0.3168698,
final_local_scores=0.29532543, pem=0.0, similarity=12.902813]
3: 36995 Cambridge
[pem_pos=0, final_scores=0.29466558,

final_local_scores=0.24110669, pem=-0.751733,
similarity=15.372858]

Sample: b'Roland_Hyatt'

[1 Roland Shane Hyatt 1] -LRB- born 30 December 1961 in -
RRB- , is a former 1 cricket er who played 2 first-class cric
ket for from 1983 to 1986 . He played for 3 Clarence District
Cricket Club in the [1 1 Tasmanian Grade Cricket 1]
competition . He was an who bowled right arm 4 off spin . He
made his first-class debut against at the [2 2 Melbourne Cric
ket Ground 2] on 11 November 1983 , in a match in which a
remarkable 227 by [3 3 David Boon 3] gave [2 Tasmania 2]
first innings points , despite in ending in a draw . Despite
taking an excellent catch , his bowling did not impress , and
he was dropped for the next match . He was picked again to
take on 5 New South Wales at the [4 4 SCG 4] on what was
expected to be a spinning wicket , but again failed to take a
wicket . [3 Hyatt 3] soon learned his most effective
deliveries were drifting balls and quickening sliders , two
delivery techniques he quickly learned to capitalise on . He
took his first [4 Sheffield Shield 4] wicket against ,
dismissing [5 5 Greg Shipperd 5] with a drifting ball . [5
Hyatt 5] never proved to be a prolific wicket-taker , and his
best bowling of 3/30 came by cleaning up the 6 New South Wales
tail-order . His prize first-class wicket came in that
innings when he claimed the wicket of the 7 test cricket
player [6 6 Bob Holland 6] . Likewise [6 Hyatt 6] 's batting
never proved to be effective at first-class level , although
he did have a good run of form during the 1985-86 season . In
that season he made his highest first-class score of 80 ,
attempting to save a match against . In a match against
towards the end of that season he was n't bowled at all ,
giving the impression then [7 Tasmanian 7] [7 7 Mark Ray 7]
had lost faith in him . He played his last match against at 8
Devonport before being dropped from the side , and returning
to 9 grade cricket

1: 8660035 Tasmanian_Grade_Cricket

2: 19765 Melbourne_Cricket_Ground

3: 610312 David_Boon

4: 773497 Sydney_Cricket_Ground

5: 10010275 Greg_Shipperd

6: 2952826 Bob_Holland

7: 16350768 Mark_Ray

1: 25675557 Cricket

[winner: 330091 Rhaphidophoridae, pem_pos=10,
final_scores=-1.3159246, final_local_scores=-1.2734036,
pem=-2.56456, similarity=-19.964716]

gt: 25675557 Cricket, pem_pos=0, final_scores=-1.391334,
final_local_scores=-1.535238, pem=-0.07532314,
similarity=-41.315235]

2: 11040 First-class_cricket

[winner: 11040 First-class_cricket, pem_pos=0,

final_scores=-1.2470627, final_local_scores=-1.3705842,
 pem=-0.00035632044, similarity=-36.809315|
 gt: 11040 First-class_cricket, pem_pos=0,
 final_scores=-1.2470627, final_local_scores=-1.3705842,
 pem=-0.00035632044, similarity=-36.809315]
 3: 8670853 Clarence_District_Cricket_Club
 [winner: 8670853 Clarence_District_Cricket_Club,
 pem_pos=0, final_scores=-0.03159146,
 final_local_scores=0.0089235455, pem=0.0,
 similarity=4.3560095|
 gt: 8670853 Clarence_District_Cricket_Club, pem_pos=0,
 final_scores=-0.03159146, final_local_scores=0.0089235455,
 pem=0.0, similarity=4.3560095]
 4: 22499 Off_spin
 [winner: 22499 Off_spin, pem_pos=0,
 final_scores=-0.9244979, final_local_scores=-1.0449088,
 pem=-0.27103484, similarity=-25.618517|
 gt: 22499 Off_spin, pem_pos=0, final_scores=-0.9244979,
 final_local_scores=-1.0449088, pem=-0.27103484,
 similarity=-25.618517]
 5: 1327417 New_South_Wales_cricket_team
 [winner: 3187751 New_South_Wales_rugby_league_team,
 pem_pos=1, final_scores=-0.05480139,
 final_local_scores=-0.07485223, pem=-2.4022121,
 similarity=14.919576|
 RECALL MISS: 1327417 New_South_Wales_cricket_team]
 6: 1327417 New_South_Wales_cricket_team
 [winner: 21654 New_South_Wales, pem_pos=0,
 final_scores=-0.04631357, final_local_scores=-0.015082449,
 pem=-0.051332146, similarity=3.9187746|
 RECALL MISS: 1327417 New_South_Wales_cricket_team]
 7: 30863 Test_cricket
 [winner: 27484820 The_Test_(short_story), pem_pos=0,
 final_scores=-0.8761552, final_local_scores=-0.75344896,
 pem=-0.4980725, similarity=-15.6861105|
 RECALL MISS: 30863 Test_cricket]
 [winner: 330091 Rhabdophoridae, pem_pos=10,
 final_scores=-1.4240122, final_local_scores=-1.3914728,
 pem=-2.56456, similarity=-23.488138|
 gt: 30863 Test_cricket, pem_pos=14,
 final_scores=-2.3398025, final_local_scores=-2.6001468,
 pem=-5.040464, similarity=-46.093025]
 8: 239257 Devonport,_Tasmania
 [winner: 21455887 Devonport_Football_Club, pem_pos=8,
 final_scores=-0.28890595, final_local_scores=-0.2886493,
 pem=-2.3048298, similarity=8.009866|
 gt: 239257 Devonport,_Tasmania, pem_pos=1,
 final_scores=-0.36341995, final_local_scores=-0.38599992,
 pem=-1.1094006, similarity=-1.3962014]
 9: 8660035 Tasmanian_Grade_Cricket
 [winner: 4704581 Sydney_Grade_Cricket, pem_pos=0,
 final_scores=-1.0027552, final_local_scores=-1.126401,


```

pem=-0.19706385, similarity=-28.452673|
  gt: 8660035 Tasmanian_Grade_Cricket, pem_pos=3,
final_scores=-1.2337253, final_local_scores=-1.3789718,
pem=-1.0626802, similarity=-31.282537 ]
1: 8660035 Tasmanian_Grade_Cricket
  [pem_pos=0, final_scores=0.10146813,
final_local_scores=0.09405188, pem=0.0, similarity=6.8964086]
2: 19765 Melbourne_Cricket_Ground
  [pem_pos=0, final_scores=0.32908314,
final_local_scores=0.31205556, pem=0.0, similarity=13.402073]
3: 610312 David_Boon
  [pem_pos=0, final_scores=0.750259,
final_local_scores=0.77778715, pem=0.0, similarity=27.300432]
4: 773497 Sydney_Cricket_Ground
  [pem_pos=0, final_scores=0.3155375,
final_local_scores=0.29259768, pem=0.0, similarity=12.821411]
5: 10010275 Greg_Shipperd
  [pem_pos=0, final_scores=0.60574216,
final_local_scores=0.6455108, pem=0.0, similarity=23.353043]
6: 2952826 Bob_Holland
  [pem_pos=0, final_scores=0.5460773,
final_local_scores=0.5916269, pem=0.0, similarity=21.745039]
7: 16350768 Mark_Ray
  [pem_pos=0, final_scores=0.34651226,
final_local_scores=0.45481724, pem=0.0, similarity=17.662367]
1: 24310182 Roland_Hyatt
  [pem_pos=0, final_scores=0.393637,
final_local_scores=0.3914293, pem=0.0, similarity=15.770744]
2: 613023 Tasmania_cricket_team
  [pem_pos=1, final_scores=0.32463157,
final_local_scores=0.272128, pem=-2.4291463,
similarity=25.42063]
3: 24310182 Roland_Hyatt
  [pem_pos=0, final_scores=0.34958196,
final_local_scores=0.3399725, pem=0.0, similarity=14.23517]
4: 1585656 Sheffield_Shield
  [pem_pos=0, final_scores=0.38326854,
final_local_scores=0.3437418, pem=0.0, similarity=14.347654]
5: 24310182 Roland_Hyatt
  [pem_pos=0, final_scores=0.59303474,
final_local_scores=0.60924107, pem=0.0, similarity=22.270681]
6: 24310182 Roland_Hyatt
  [pem_pos=0, final_scores=0.59124255,
final_local_scores=0.60728335, pem=0.0, similarity=22.21226]
7: 613023 Tasmania_cricket_team
  [pem_pos=2, final_scores=0.15658896,
final_local_scores=0.09039748, pem=-1.7128749,
similarity=16.102234]

```

Sample: b'Scotch_and_Soda'

Scotch and Soda is a 1 magic effect involving a 2 copper 3 coin and a 4 silver coin which appear to transpose in the

spectator 's hands . The effect relies on simple 5 gimmick and the tricks performed with it are 6 self-working and require almost no skill . The trick is named after the 7 cock tail Scotch and soda ; the copper coin represents the ' ' ' and the silver coin represents the ' ' 8 soda ' ' . The effect is usually performed as follows : The magician displays two coins of almost equal size , one 9 copper and one 10 silver . The silver coin is most often a [1 1 U.S. 1] and the copper coin is usually either an [1 11 English 1] or a [2 2 Mexican 2] 12 centavo . The magician stacks the coins and places them into the spectator 's hand . He then asks her to place her hands behind her back and put one coin in each hand , remarking that the silver coin is slightly larger than the copper coin , making them easy to tell apart . The magician asks to see the silver coin which the spectator produces . When the spectator opens her other hand , the copper coin has become a . The copper coin can then be made to appear wherever the magician desires , such as in the spectator 's pocket or under an object across the room . Scotch and Soda is a popular trick that can be purchased at many magic stores . A number of books exclusively on the subject describe different effects that can be achieved with the [2 Scotch 2] and Soda gimmick . Gin and tonic is a version of the trick using a and a . The modern version that uses a 20 Centavo and a [3 Kennedy 3] half is often attributed to Richard Himber . In the early 1900s , [4 America 4] 's oldest magic company , 13 Martinka

1: 3434750 United_States

2: 3966054 Mexico

1: 25152964 Magic_(illusion)

[winner: 25152964 Magic_(illusion), pem_pos=0, final_scores=-0.97129977, final_local_scores=-1.0555196, pem=0.0, similarity=-27.409096]

gt: 25152964 Magic_(illusion), pem_pos=0, final_scores=-0.97129977, final_local_scores=-1.0555196, pem=0.0, similarity=-27.409096]

2: 125293 Copper

[winner: 16376403 Cash_(Chinese_coin), pem_pos=0, final_scores=-0.9821404, final_local_scores=-0.96798563, pem=0.0, similarity=-24.79691]

RECALL MISS: 125293 Copper]

3: 7558 Coin

[winner: 16376403 Cash_(Chinese_coin), pem_pos=0, final_scores=-0.9821404, final_local_scores=-0.96798563, pem=0.0, similarity=-24.79691]

RECALL MISS: 7558 Coin]

4: 27119 Silver

[winner: 3749317 Silver_coin, pem_pos=0, final_scores=-1.1140072, final_local_scores=-1.1622229, pem=0.0, similarity=-30.593338]

RECALL MISS: 27119 Silver]

5: 232791 Gimmick

[winner: 27875694 Movie_gimmick, pem_pos=2,
final_scores=-1.43592, final_local_scores=-1.387671,
pem=-3.6605046, similarity=-17.414762]
gt: 232791 Gimmick, pem_pos=1, final_scores=-1.532158,
final_local_scores=-1.6207383, pem=-0.9453368,
similarity=-39.135464]
6: 2370017 Self-working_magic
[winner: 2370017 Self-working_magic, pem_pos=0,
final_scores=-1.0994117, final_local_scores=-1.0404475,
pem=0.0, similarity=-26.959312]
gt: 2370017 Self-working_magic, pem_pos=0,
final_scores=-1.0994117, final_local_scores=-1.0404475,
pem=0.0, similarity=-26.959312]
7: 7599 Cocktail
[winner: 2832107 Gimlet_(cocktail), pem_pos=1,
final_scores=-0.81344575, final_local_scores=-0.7582969,
pem=-0.68437725, similarity=-14.81763]
gt: 7599 Cocktail, pem_pos=0, final_scores=-1.2125418,
final_local_scores=-1.3242948, pem=-0.22320151,
similarity=-34.21608]
8: 240561 Carbonated_water
[winner: 27061 Soft_drink, pem_pos=0,
final_scores=-0.5916033, final_local_scores=-0.6430986,
pem=-0.6392334, similarity=-11.625381]
gt: 240561 Carbonated_water, pem_pos=1,
final_scores=-0.6951795, final_local_scores=-0.7352229,
pem=-1.0941036, similarity=-11.900899]
9: 125293 Copper
[winner: 230641 Copper_interconnect, pem_pos=7,
final_scores=-1.1655693, final_local_scores=-1.0931277,
pem=-1.9456949, similarity=-17.950407]
gt: 125293 Copper, pem_pos=0, final_scores=-1.246788,
final_local_scores=-1.309478, pem=-0.017616155,
similarity=-34.891926]
10: 27119 Silver
[winner: 27119 Silver, pem_pos=0,
final_scores=-1.1754906, final_local_scores=-1.2410254,
pem=-0.12949902, similarity=-32.24073]
gt: 27119 Silver, pem_pos=0, final_scores=-1.1754906,
final_local_scores=-1.2410254, pem=-0.12949902,
similarity=-32.24073]
11: 9316 England
[winner: 8569916 English_language, pem_pos=0,
final_scores=0.2964393, final_local_scores=0.32104912,
pem=-0.640266, similarity=17.152325]
gt: 9316 England, pem_pos=1, final_scores=-0.04556489,
final_local_scores=-0.08532539, pem=-1.3527626,
similarity=8.899966]
12: 814147 Centavo
[winner: 1951903 East_Timor_centavo_coins, pem_pos=0,
final_scores=-0.81175643, final_local_scores=-0.72544456,
pem=0.0, similarity=-17.559]

```

gt: 814147 Centavo, pem_pos=1, final_scores=-0.9074983,
final_local_scores=-0.8097951, pem=-0.45675844,
similarity=-17.592264 ]
13: 5510392 Martinka
[ winner: 5510392 Martinka, pem_pos=0,
final_scores=-0.14893779, final_local_scores=-0.07613054,
pem=0.0, similarity=1.8178258 ]
gt: 5510392 Martinka, pem_pos=0,
final_scores=-0.14893779, final_local_scores=-0.07613054,
pem=0.0, similarity=1.8178258 ]
1: 3434750 United_States
[pem_pos=0, final_scores=0.473525,
final_local_scores=0.45274276, pem=-0.14075686,
similarity=18.365917]
2: 3966054 Mexico
[pem_pos=0, final_scores=0.4262824,
final_local_scores=0.45827353, pem=-0.3968012,
similarity=19.923376]
1: 8569916 English_language
[pem_pos=0, final_scores=0.2964393,
final_local_scores=0.32104912, pem=-0.640266,
similarity=17.152325]
2: 28896 Scotch_whisky
[pem_pos=0, final_scores=0.056434438,
final_local_scores=0.06494382, pem=-0.37716895,
similarity=8.07887]
3: 5119376 John_F._Kennedy
[pem_pos=0, final_scores=0.105848715,
final_local_scores=0.12744987, pem=-1.3076366,
similarity=15.004203]
4: 3434750 United_States
[pem_pos=0, final_scores=0.5588786,
final_local_scores=0.5459786, pem=-0.78483206,
similarity=24.650843]

```

Sample: b'Sherwin_Campbell'

Sherwin Legay Campbell -LRB- born 1 November 1970 -RRB- is a who played 52 [1 1 Tests 1] and 90 [2 2 One Day International 2] s for the [3 3 West Indies 3] . He has played a total of 177 games between and , scoring more than 10,000 first class runs with 26 1 centuries . Four of those were in [1 Test 1] cricket , with a highest score of 208 against in the first [2 Test 2] of the series , played at his home ground in [4 4 Bridgetown 4] . Batting for thirteen hours , he hit 30 fours and by the time he left the score was 458 for 8 - well ahead of [3 New Zealand 3] 's first innings total of 195 . He also hit 29 not out in the second innings as the [4 West Indies 4] won by ten wickets . However , after hitting 170 against [5 New Zealand 5] in the first [6 Test 6] at [5 5 Hamilton 5] in he failed to hit a century in his next 33 innings , and only passed 50 five times - in a run yielding 701 runs at 21.24 . 79 and 54 in the last [7 Test 7] of the [6 6 Frank Worrell Tr

ophy 6] against could not prevent him being dropped after the tour , and after missing 10 [8 Tests 8] he returned for the first [9 Test 9] against in [7 7 Sharjah 7] . He made six and 20 , and did not return to international cricket again . He continued to play first class cricket for , and has not announced his retirement , although he was left out of the squad to play for the first match of the . He now plays cricket for [10 Heywood Cricket Club 10] in 2 Heywood , near 3 Rochdale

1: 30863 Test_cricket
2: 381447 One_Day_International
3: 3379941 West_Indies_cricket_team
4: 56630 Bridgetown
5: 86654 Hamilton,_New_Zealand
6: 1176812 Frank_Worrell_Trophy
7: 15674115 Sharjah_(city)

1: 1671458 Century_(cricket)
[winner: 34741 6th_century, pem_pos=1,
final_scores=-0.9234584, final_local_scores=-0.85268766,
pem=-0.6931472, similarity=-17.586746]
gt: 1671458 Century_(cricket), pem_pos=2,
final_scores=-1.3052146, final_local_scores=-1.4896519,
pem=-0.928203, similarity=-35.31676]
2: 1219535 Heywood,_Greater_Manchester
[winner: 1219535 Heywood,_Greater_Manchester, pem_pos=0,
final_scores=0.025294527, final_local_scores=0.08773719,
pem=-1.0117869, similarity=12.21022]
gt: 1219535 Heywood,_Greater_Manchester, pem_pos=0,
final_scores=0.025294527, final_local_scores=0.08773719,
pem=-1.0117869, similarity=12.21022]

3: 635367 Rochdale
[winner: 635367 Rochdale, pem_pos=0,
final_scores=-0.032799244, final_local_scores=-0.011447206,
pem=-0.7956193, similarity=8.074807]
gt: 635367 Rochdale, pem_pos=0,
final_scores=-0.032799244, final_local_scores=-0.011447206,
pem=-0.7956193, similarity=8.074807]

1: 30863 Test_cricket
[pem_pos=0, final_scores=0.47866637,
final_local_scores=0.41049272, pem=-0.11903022,
similarity=16.98694]
2: 381447 One_Day_International
[pem_pos=1, final_scores=0.2047476,
final_local_scores=0.12899223, pem=-0.022472855,
similarity=8.061309]
3: 3379941 West_Indies_cricket_team
[pem_pos=1, final_scores=0.32916147,
final_local_scores=0.26181734, pem=-1.1020415,
similarity=17.895937]
4: 56630 Bridgetown
[pem_pos=0, final_scores=0.26277262,
final_local_scores=0.2777087, pem=-0.10175923,

similarity=12.930477]
5: 86654 Hamilton, New Zealand
[pem_pos=1, final_scores=0.14437206,
final_local_scores=0.16560546, pem=-2.40051,
similarity=22.086058]
6: 1176812 Frank_Worrell_Trophy
[pem_pos=0, final_scores=0.29211926,
final_local_scores=0.34043953, pem=0.0, similarity=14.249107]
7: 15674115 Sharjah_(city)
[pem_pos=0, final_scores=0.12086026,
final_local_scores=0.15065214, pem=-0.425711,
similarity=10.900556]
1: 30863 Test_cricket
[pem_pos=0, final_scores=0.3682561,
final_local_scores=0.29030108, pem=-0.23950954,
similarity=14.0553665]
2: 30863 Test_cricket
[pem_pos=0, final_scores=0.5282096,
final_local_scores=0.4653306, pem=-0.23950954,
similarity=19.278597]
3: 478115 New_Zealand_national_cricket_team
[pem_pos=1, final_scores=0.48274302,
final_local_scores=0.43153155, pem=-3.142959,
similarity=34.059376]
4: 3379941 West_Indies_cricket_team
[pem_pos=1, final_scores=0.51314884,
final_local_scores=0.46279556, pem=-1.1020415,
similarity=23.893528]
5: 478115 New_Zealand_national_cricket_team
[pem_pos=1, final_scores=0.49548167,
final_local_scores=0.4454466, pem=-3.142959,
similarity=34.47463]
6: 30863 Test_cricket
[pem_pos=0, final_scores=0.52887493,
final_local_scores=0.46605736, pem=-0.23950954,
similarity=19.300283]
7: 30863 Test_cricket
[pem_pos=0, final_scores=0.31127435,
final_local_scores=0.22805724, pem=-0.23950954,
similarity=12.1978855]
8: 30863 Test_cricket
[pem_pos=0, final_scores=0.48163903,
final_local_scores=0.41373986, pem=-0.11903022,
similarity=17.083841]
9: 30863 Test_cricket
[pem_pos=0, final_scores=0.39947337,
final_local_scores=0.3244012, pem=-0.23950954,
similarity=15.072983]
10: 451214 Heywood_Cricket_Club
[pem_pos=0, final_scores=0.13699453,
final_local_scores=0.17694393, pem=0.0, similarity=9.370072]

Sample: b'Simon_L'

___ Adler Simon Louis Adler -LRB- August 30 , 1867 in [1 1 Seneca Falls 1] , [2 2 Seneca County 2] , [3 New York 3] -- May 23 , 1934 -RRB- was a [4 3 United States 4] federal judge . He graduated from [1 1 Cornell University 1] in 1889 , and from [2 2 Harvard Law School 2] in 1892 . He practiced law in [5 4 Rochester 5] , [6 New York 6] . [7 Adler 7] was a member of the 5 New York State Assembly -LRB- Monroe Co. , 2nd D. -RRB- in 6 1911 , 7 1912 , 8 1913 , 9 1914 , , , , , , , , , , , and ; and was Majority Leader from 1916 to 1926 . On May 19 , 1927 , Adler received a 10 recess appointment from President [3 3 Calvin Coolidge 3] to a new seat on the [8 11 United States District Court 8] for the Western District of New York created by 44 Stat . 1370 . Formally nominated on December 6 , 1927 , he was confirmed by the 12 United States Senate

1: 7954422 Cornell_University

2: 158984 Harvard_Law_School

3: 6195 Calvin_Coolidge

1: 260042 Seneca_Falls_(hamlet),_New_York

[winner: 260043 Seneca_Falls,_New_York, pem_pos=0, final_scores=0.39978284, final_local_scores=0.4177397, pem=-0.070279814, similarity=16.938091]

gt: 260042 Seneca_Falls_(hamlet),_New_York, pem_pos=1, final_scores=0.3330319, final_local_scores=0.35326982, pem=-0.10198058, similarity=15.186574]

2: 56940 Seneca_County,_New_York

[winner: 8210131 New_York, pem_pos=0, final_scores=0.25140738, final_local_scores=0.18364614, pem=-0.2889045, similarity=11.141188]

RECALL MISS: 56940 Seneca_County,_New_York]

[winner: 93030 Seneca_County,_Ohio, pem_pos=0, final_scores=0.1264884, final_local_scores=0.17653325, pem=-0.20632973, similarity=10.479871]

gt: 56940 Seneca_County,_New_York, pem_pos=1, final_scores=0.08749895, final_local_scores=0.1014858, pem=-0.7116392, similarity=10.988256]

[winner: 37537378 Admissions_(2011_film), pem_pos=12, final_scores=-1.7177888, final_local_scores=-1.7064795, pem=-4.6728287, similarity=-21.42347]

RECALL MISS: 56940 Seneca_County,_New_York]

3: 834368 United_States_federal_judge

[winner: 3434750 United_States, pem_pos=0, final_scores=0.46088248, final_local_scores=0.45187092, pem=-0.073457584, similarity=17.973917]

RECALL MISS: 834368 United_States_federal_judge]

[winner: 6277728 James_L._Oakes, pem_pos=13, final_scores=-1.2587407, final_local_scores=-1.3223674, pem=-2.6390574, similarity=-21.020763]

gt: 834368 United_States_federal_judge, pem_pos=1, final_scores=-1.3977486, final_local_scores=-1.5431697, pem=-2.563578, similarity=-28.020414]

[winner: 1691932 Federal_architecture, pem_pos=12, final_scores=-1.5916439, final_local_scores=-1.6575463, pem=-3.3951852, similarity=-26.91123]
 RECALL MISS: 834368 United_States_federal_judge]

4: 126641 Rochester, New_York
 [winner: 126641 Rochester, New_York, pem_pos=0, final_scores=0.33951634, final_local_scores=0.28684145, pem=-1.0925732, similarity=18.591217]
 gt: 126641 Rochester, New_York, pem_pos=0, final_scores=0.33951634, final_local_scores=0.28684145, pem=-1.0925732, similarity=18.591217]
 [winner: 8210131 New_York, pem_pos=0, final_scores=0.3133965, final_local_scores=0.25135988, pem=-0.2889045, similarity=13.1619005]
 RECALL MISS: 126641 Rochester, New_York]

[winner: 37537378 Admissions_(2011_film), pem_pos=12, final_scores=-1.675556, final_local_scores=-1.6603465, pem=-4.6728287, similarity=-20.046766]
 RECALL MISS: 126641 Rochester, New_York]

5: 581663 New_York_State_Assembly
 [winner: 581663 New_York_State_Assembly, pem_pos=0, final_scores=-0.6514204, final_local_scores=-0.74327606, pem=0.0, similarity=-18.091127]
 gt: 581663 New_York_State_Assembly, pem_pos=0, final_scores=-0.6514204, final_local_scores=-0.74327606, pem=0.0, similarity=-18.091127]

6: 39099705 134th_New_York_State_Legislature
 [winner: 34605 1911, pem_pos=0, final_scores=-1.012427, final_local_scores=-0.9036789, pem=-0.6630881, similarity=-19.271893]
 RECALL MISS: 39099705 134th_New_York_State_Legislature]

7: 39281426 135th_New_York_State_Legislature
 [winner: 34616 1912, pem_pos=0, final_scores=-0.7461456, final_local_scores=-0.6195767, pem=-0.75507396, similarity=-10.293482]
 RECALL MISS: 39281426 135th_New_York_State_Legislature]

8: 39501535 136th_New_York_State_Legislature
 [winner: 34617 1913, pem_pos=0, final_scores=-0.9531842, final_local_scores=-0.82886195, pem=-0.64449716, similarity=-17.140305]
 RECALL MISS: 39501535 136th_New_York_State_Legislature]

9: 39515509 137th_New_York_State_Legislature
 [winner: 19283920 1914, pem_pos=0, final_scores=-1.0385349, final_local_scores=-0.9615398, pem=-0.7211613, similarity=-20.682766]
 RECALL MISS: 39515509 137th_New_York_State_Legislature]

10: 435607 Recess_appointment
 [winner: 2071378 Recess, County_Galway, pem_pos=1, final_scores=-0.9065521, final_local_scores=-0.7894681, pem=-0.99537104, similarity=-14.056608]
 RECALL MISS: 435607 Recess_appointment]
 [winner: 435607 Recess_appointment, pem_pos=6,

final_scores=-1.4154942, final_local_scores=-1.5116282,
pem=-2.8157609, similarity=-25.707739|
gt: 435607 Recess_appointment, pem_pos=6,
final_scores=-1.4154942, final_local_scores=-1.5116282,
pem=-2.8157609, similarity=-25.707739]
11: 873559
United_States_District_Court_for_the_Western_District_of_New_York
[winner: 70235 United_States_district_court, pem_pos=0,
final_scores=0.16064493, final_local_scores=0.17768535,
pem=0.0, similarity=9.392199|
RECALL MISS: 873559
United_States_District_Court_for_the_Western_District_of_New_York]
[winner: 873559
United_States_District_Court_for_the_Western_District_of_New_York, pem_pos=0, final_scores=-0.22020371,
final_local_scores=-0.20279294, pem=0.0,
similarity=-1.9620328|
gt: 873559
United_States_District_Court_for_the_Western_District_of_New_York, pem_pos=0, final_scores=-0.22020371,
final_local_scores=-0.20279294, pem=0.0,
similarity=-1.9620328]
[winner: 3349493 Male-
female_income_disparity_in_the_United_States, pem_pos=0,
final_scores=-1.6910732, final_local_scores=-1.7700739,
pem=-0.12731102, similarity=-48.040485|
RECALL MISS: 873559
United_States_District_Court_for_the_Western_District_of_New_York]
12: 24909346 United_States_Senate
[winner: 24909346 United_States_Senate, pem_pos=0,
final_scores=0.024036393, final_local_scores=-0.0019713193,
pem=0.0, similarity=4.0308847|
gt: 24909346 United_States_Senate, pem_pos=0,
final_scores=0.024036393, final_local_scores=-0.0019713193,
pem=0.0, similarity=4.0308847]
1: 7954422 Cornell_University
[pem_pos=0, final_scores=0.400302,
final_local_scores=0.38943553, pem=0.0, similarity=15.711246]
2: 158984 Harvard_Law_School
[pem_pos=0, final_scores=0.336253,
final_local_scores=0.35740137, pem=0.0, similarity=14.755282]
3: 6195 Calvin_Coolidge
[pem_pos=0, final_scores=0.32188207,
final_local_scores=0.3230664, pem=0.0, similarity=13.7306595]
1: 260043 Seneca_Falls,_New_York
[pem_pos=0, final_scores=0.39978284,
final_local_scores=0.4177397, pem=-0.070279814,
similarity=16.938091]
2: 93030 Seneca_County,_Ohio

[pem_pos=0, final_scores=0.1264884,
 final_local_scores=0.17653325, pem=-0.20632973,
 similarity=10.479871]
 3: 8210131 New_York
 [pem_pos=0, final_scores=0.25140738,
 final_local_scores=0.18364614, pem=-0.2889045,
 similarity=11.141188]
 4: 3434750 United_States
 [pem_pos=0, final_scores=0.46088248,
 final_local_scores=0.45187092, pem=-0.073457584,
 similarity=17.973917]
 5: 126641 Rochester,_New_York
 [pem_pos=0, final_scores=0.33951634,
 final_local_scores=0.28684145, pem=-1.0925732,
 similarity=18.591217]
 6: 8210131 New_York
 [pem_pos=0, final_scores=0.3133965,
 final_local_scores=0.25135988, pem=-0.2889045,
 similarity=13.1619005]
 7: 21677962 Simon_L._Adler
 [pem_pos=0, final_scores=0.09191568,
 final_local_scores=0.17596263, pem=0.0, similarity=9.340789]
 8: 70235 United_States_district_court
 [pem_pos=0, final_scores=0.16064493,
 final_local_scores=0.17768535, pem=0.0, similarity=9.392199]

Sample: b'Steep_Falls'

, _Maine Steep Falls is a 1 census-designated place -LRB-
 CDP -RRB- in the 2 town of [1 1 Standish 1] in [1 3 Cumberland
 County 1] , [2 2 Maine 2] , [3 3 United States 3] . The
 population was 1,139 at the 4 2010 census . It is part of the
 [2 5 Portland 2] -- [4 4 South Portland 4] -- [5 5 Biddeford 5
] , [3 Maine 3] . Steep Falls is located in the western
 corner of the town of [4 Standish 4] along the [6 6 Saco River
 6] at . 6 Maine State Route 11 crosses the [5 Saco River 5]
 into the community , heading southwest towards [7 7 Limington
 7] and [6 7 Limerick 6] and north towards [7 8 Naples 7] . 9 M
 aine State Route 113 coincides with Route 11 through the
 center of [8 Steep Falls 8] , then heads southeast towards [9
 10 Standish 9] village and northwest towards [8 8 Hiram 8]
 and [9 9 Fryeburg 9] . According to the [10 10 United States C
 ensus Bureau 10]

1: 115951 Standish,_Maine
 2: 19977 Maine
 3: 3434750 United_States
 4: 115950 South_Portland,_Maine
 5: 116267 Biddeford,_Maine
 6: 1112876 Saco_River
 7: 116278 Limington,_Maine
 8: 116080 Hiram,_Maine
 9: 259268 Fryeburg,_Maine
 10: 57070 United_States_Census_Bureau

1: 269771 Census-designated_place
 [winner: 269771 Census-designated_place, pem_pos=0, final_scores=-0.8527604, final_local_scores=-0.9334541, pem=-0.025317784, similarity=-23.628736|
 gt: 269771 Census-designated_place, pem_pos=0, final_scores=-0.8527604, final_local_scores=-0.9334541, pem=-0.025317784, similarity=-23.628736]
 [winner: 23110064 Zelah, Judea, pem_pos=6, final_scores=-1.1850766, final_local_scores=-1.1301274, pem=-1.7917595, similarity=-19.89168|
 RECALL MISS: 269771 Census-designated_place]

2: 3400990 New_England_town
 [winner: 2813519 Types_of_inhabited_localities_in_Russia, pem_pos=2, final_scores=-1.1408868, final_local_scores=-1.0695401, pem=-2.3084521, similarity=-15.273777|
 gt: 3400990 New_England_town, pem_pos=1, final_scores=-1.5773368, final_local_scores=-1.7344749, pem=-1.9336947, similarity=-37.154743]

3: 95498 Cumberland_County, Maine
 [winner: 93447 Cumberland_County, New_Jersey, pem_pos=1, final_scores=0.23892172, final_local_scores=0.21511513, pem=-1.3867015, similarity=18.050276|
 gt: 95498 Cumberland_County, Maine, pem_pos=2, final_scores=0.23175831, final_local_scores=0.1912083, pem=-1.7433397, similarity=19.276302]

4: 3448729 2010_United_States_Census
 [winner: 3448729 2010_United_States_Census, pem_pos=0, final_scores=-1.036038, final_local_scores=-1.1679444, pem=0.0, similarity=-30.764082|
 gt: 3448729 2010_United_States_Census, pem_pos=0, final_scores=-1.036038, final_local_scores=-1.1679444, pem=0.0, similarity=-30.764082]

5: 44591 Portland, Maine
 [winner: 23503 Portland, Oregon, pem_pos=0, final_scores=0.42617935, final_local_scores=0.4439991, pem=-0.66320467, similarity=20.946142|
 gt: 44591 Portland, Maine, pem_pos=1, final_scores=-0.1024864, final_local_scores=-0.1697542, pem=-1.9638829, similarity=9.70381]

6: 4638960 Maine_State_Route_11
 [winner: 4638960 Maine_State_Route_11, pem_pos=0, final_scores=-0.24849752, final_local_scores=-0.2541339, pem=0.0, similarity=-3.49415|
 gt: 4638960 Maine_State_Route_11, pem_pos=0, final_scores=-0.24849752, final_local_scores=-0.2541339, pem=0.0, similarity=-3.49415]

7: 116277 Limerick, Maine
 [winner: 56885 Limerick, pem_pos=0, final_scores=0.12807204, final_local_scores=0.15482023, pem=-0.62048864, similarity=12.084171|
 gt: 116277 Limerick, Maine, pem_pos=6,

```

final_scores=-0.4030909, final_local_scores=-0.4832012,
pem=-3.0363002, similarity=6.1818967 ]
8: 115943 Naples,_Maine
  [winner: 55880 Naples, pem_pos=0,
final_scores=0.18327586, final_local_scores=0.26925266,
pem=-0.1811944, similarity=13.110112]
  gt: 115943 Naples,_Maine, pem_pos=3,
final_scores=-0.30567068, final_local_scores=-0.37254724,
pem=-3.0591774, similarity=9.608442 ]
9: 24113848 Maine_State_Route_113
  [winner: 332996 State_highway, pem_pos=0,
final_scores=-0.2579016, final_local_scores=-0.2520644,
pem=-0.2668789, similarity=-1.9810624]
  RECALL MISS: 24113848 Maine_State_Route_113]
  [winner: 19977 Maine, pem_pos=0, final_scores=-0.5737,
final_local_scores=-0.7009547, pem=-0.2749413,
similarity=-15.332999]
  RECALL MISS: 24113848 Maine_State_Route_113]
  [winner: 20488584 SP-113, pem_pos=19,
final_scores=-1.1422385, final_local_scores=-0.9270588,
pem=-3.1888444, similarity=-6.2341394]
  gt: 24113848 Maine_State_Route_113, pem_pos=13,
final_scores=-1.385812, final_local_scores=-1.4608612,
pem=-2.636882, similarity=-25.165522 ]
10: 36136895 Standish_(CDP),_Maine
  [winner: 115951 Standish,_Maine, pem_pos=0,
final_scores=0.37156624, final_local_scores=0.32801265,
pem=-0.9469492, similarity=19.027922]
  RECALL MISS: 36136895 Standish_(CDP),_Maine]
  [winner: 53509 Village, pem_pos=0,
final_scores=-1.3002867, final_local_scores=-1.2561085,
pem=-0.05490723, similarity=-33.096478]
  RECALL MISS: 36136895 Standish_(CDP),_Maine]
1: 115951 Standish,_Maine
  [pem_pos=0, final_scores=0.37577128,
final_local_scores=0.33319438, pem=-0.9469492,
similarity=19.182554]
2: 19977 Maine
  [pem_pos=0, final_scores=0.16095062,
final_local_scores=0.106468976, pem=-0.2749413,
similarity=8.762134]
3: 3434750 United_States
  [pem_pos=0, final_scores=0.42976612,
final_local_scores=0.45114475, pem=-0.073457584,
similarity=17.952246]
4: 115950 South_Portland,_Maine
  [pem_pos=0, final_scores=0.19791631,
final_local_scores=0.1699026, pem=0.0, similarity=9.1599455]
5: 116267 Biddeford,_Maine
  [pem_pos=0, final_scores=0.398943,
final_local_scores=0.37530637, pem=0.0, similarity=15.289603]
6: 1112876 Saco_River

```

[pem_pos=0, final_scores=0.19361071,
final_local_scores=0.19611907, pem=0.0, similarity=9.942297]
7: 116278 Limington, Maine
[pem_pos=1, final_scores=0.39938676,
final_local_scores=0.38978636, pem=-0.10848068,
similarity=16.31165]
8: 116080 Hiram, Maine
[pem_pos=0, final_scores=0.07437278,
final_local_scores=0.035760194, pem=-1.2902242,
similarity=12.173308]
9: 259268 Fryeburg, Maine
[pem_pos=0, final_scores=0.359006,
final_local_scores=0.342194, pem=0.0, similarity=14.301464]
10: 57070 United_States_Census_Bureau
[pem_pos=0, final_scores=0.082806006,
final_local_scores=0.051280558, pem=0.0, similarity=5.620027]
1: 93447 Cumberland_County, New_Jersey
[pem_pos=1, final_scores=0.23892172,
final_local_scores=0.21511513, pem=-1.3867015,
similarity=18.050276]
2: 23503 Portland, Oregon
[pem_pos=0, final_scores=0.42617935,
final_local_scores=0.4439991, pem=-0.66320467,
similarity=20.946142]
3: 19977 Maine
[pem_pos=0, final_scores=0.17008953,
final_local_scores=0.11645186, pem=-0.2749413,
similarity=9.060042]
4: 115951 Standish, Maine
[pem_pos=0, final_scores=0.36429542,
final_local_scores=0.32065874, pem=-0.9469492,
similarity=18.808466]
5: 1112876 Saco_River
[pem_pos=0, final_scores=0.1563646,
final_local_scores=0.15543333, pem=0.0, similarity=8.728153]
6: 56885 Limerick
[pem_pos=0, final_scores=0.12807204,
final_local_scores=0.15482023, pem=-0.62048864,
similarity=12.084171]
7: 55880 Naples
[pem_pos=0, final_scores=0.18327586,
final_local_scores=0.26925266, pem=-0.1811944,
similarity=13.110112]
8: 115951 Standish, Maine
[pem_pos=0, final_scores=0.31609035,
final_local_scores=0.26841712, pem=-0.3184537,
similarity=13.831615]
9: 115951 Standish, Maine
[pem_pos=0, final_scores=0.37156624,
final_local_scores=0.32801265, pem=-0.9469492,
similarity=19.027922]

Sample: b'Teddy_Holland'

[1 Teddy Holland 1] -LRB- born 1948 in [1 1 Ballinascarty 1] , [2 2 County Cork 2] -RRB- is an [2 1 Irish 2] former 2 Gaelic football and former player . He played football with his local clubs [3 3 Ballinascorthy 3] and 3 St. Finbarr 's and was a member of the [3 4 Cork 3] senior inter-county team from 1969 until 1970 . [4 Holland 4] is best known in recent times for his short-lived managerial appointment of the [5 Cork 5] senior football team during a period when they were on . [6 Holland 6] played his early club football with his local [4 4 Ballinascarty 4] side . He also lined out with the [5 5 Carbery 5] division in the county senior championship before moving to the famous [7 Cork 7] club 5 St Finbarr 's club in the early 1970s . It was with ` the Barr 's ' that [8 Holland 8] enjoyed much success . It was here that he won his first [6 6 Cork Senior Football Championship 6] medal in 1976 . He won two further county medals in 1979 and in 1980 . The latter victory was converted into a [7 7 Munster Senior Club Football Championship 7] title as well as an [8 8 All-Ireland Senior Club Football Championship 8] title . [9 Holland 9] won a fourth county medal in 1982 . [10 Holland 10] 's time off the field with [9 9 Cork GAA 9] began as a selector with the county senior football team in 1993 . It was a great era for [10 10 Cork 10] as the team captured three successive 6 Munster titles , however , the team failed to capture an [11 All-Ireland 11] crown . [12 Holland 12] later moved onto the Cork under-21 team where he worked as a selector again . In 1998 [13 Holland 13] took over as manager of the [14 Cork 14] minor football team . His first season in charge brought a surprise loss to [15 7 Limerick 15] . The following year saw [16 Cork 16] win the [17 Munster 17] title before going on to be defeated by [11 11 Mayo 11] in the [18 All-Ireland 18] semi-final . [19 Holland 19] 's team retained their provincial title in 2000 before going on to win the [20 8 All-Ireland 20] title . He was appointed manager of the [21 Cork 21] football team in 2007 while the players were on protesting about the process for picking selectors . The players refused to play under [22 Holland 22] as he was appointed during the strike . The player strike lasted until February 18 , 2008 when he resigned as manager as one of the conditions of resolving the strike without ever actually working with the players . He was succeeded as team manager by [12 12 Conor Counihan 12]

1: 16331371 Ballinascarty

2: 46709 County_Cork

3: 23640683 Ballinascorthy_GAA

4: 23640683 Ballinascorthy_GAA

5: 11154395 Carbery_GAA

6: 7084096 Cork_Senior_Football_Championship

7: 8280565 Munster_Senior_Club_Football_Championship

8: 3331546 All-Ireland_Senior_Club_Football_Championship

9: 32532531 Cork_GAA

10: 32532531 Cork_GAA
11: 2818470 Mayo_GAA
12: 15845927 Conór_Counihan
1: 775859 Irish_people
[winner: 14560 Republic_of_Ireland, pem_pos=3,
final_scores=0.41853207, final_local_scores=0.45817167,
pem=-2.2576945, similarity=30.040161|
gt: 775859 Irish_people, pem_pos=0,
final_scores=-0.3749821, final_local_scores=-0.40199098,
pem=-1.1828228, similarity=-1.4741249]
2: 37688 Gaelic_football
[winner: 37688 Gaelic_football, pem_pos=6,
final_scores=-0.22792356, final_local_scores=-0.30581143,
pem=-2.5592473, similarity=8.881274|
gt: 37688 Gaelic_football, pem_pos=6,
final_scores=-0.22792356, final_local_scores=-0.30581143,
pem=-2.5592473, similarity=8.881274]
[winner: 4324269 Mike_Mayock, pem_pos=0,
final_scores=-1.5363805, final_local_scores=-1.5916333,
pem=-0.23922965, similarity=-42.106834|
gt: 37688 Gaelic_football, pem_pos=13,
final_scores=-2.5313866, final_local_scores=-2.821995,
pem=-4.4640217, similarity=-55.848198]
3: 8330555 St._Finbarr's_GAA
[winner: 1483333 Finbarr_of_Cork, pem_pos=0,
final_scores=-0.011923492, final_local_scores=0.073010996,
pem=0.0, similarity=6.2685065|
RECALL MISS: 8330555 St._Finbarr's_GAA]
[winner: 1777106 Vince_Welnick, pem_pos=3,
final_scores=-2.387701, final_local_scores=-2.4741359,
pem=-2.8006017, similarity=-54.51333|
RECALL MISS: 8330555 St._Finbarr's_GAA]
4: 32532531 Cork_GAA
[winner: 8140815 Cork_(city), pem_pos=0,
final_scores=0.13149486, final_local_scores=0.12430984,
pem=-0.83752966, similarity=12.353983|
gt: 32532531 Cork_GAA, pem_pos=1,
final_scores=0.08697711, final_local_scores=0.014481455,
pem=-1.0852067, similarity=10.423391]
5: 8330555 St._Finbarr's_GAA
[winner: 27889293 Finbarr_Delaney, pem_pos=2,
final_scores=-0.96316916, final_local_scores=-0.9989409,
pem=-1.609438, similarity=-16.968302|
RECALL MISS: 8330555 St._Finbarr's_GAA]
[winner: 463683 Stump_(cricket), pem_pos=1,
final_scores=-1.427153, final_local_scores=-1.4937816,
pem=-1.7897967, similarity=-30.754517|
RECALL MISS: 8330555 St._Finbarr's_GAA]
[winner: 1415027 David_Berman_(musician), pem_pos=5,
final_scores=-2.396475, final_local_scores=-2.4774904,
pem=-3.0012724, similarity=-53.522152|
RECALL MISS: 8330555 St._Finbarr's_GAA]

6: 5518247 Munster_Senior_Football_Championship
[winner: 985954 Munster_Rugby, pem_pos=3,
final_scores=-0.024109349, final_local_scores=-0.05794224,
pem=-1.883883, similarity=12.60545]
gt: 5518247 Munster_Senior_Football_Championship,
pem_pos=10, final_scores=-0.23426993,
final_local_scores=-0.2958322, pem=-3.1944976,
similarity=12.633665]
7: 2810846 Limerick_GAA
[winner: 1234583 Limerick_F.C., pem_pos=4,
final_scores=0.39661676, final_local_scores=0.38234884,
pem=-2.6802762, similarity=30.075521]
gt: 2810846 Limerick_GAA, pem_pos=1,
final_scores=0.11272992, final_local_scores=0.04975784,
pem=-1.520826, similarity=13.845075]
8: 1589212 All-Ireland
[winner: 780802 All-Ireland_Senior_Hurling_Championship,
pem_pos=0, final_scores=0.07505016,
final_local_scores=0.025283411, pem=-0.83555937,
similarity=9.388123]
gt: 1589212 All-Ireland, pem_pos=10,
final_scores=-0.37871063, final_local_scores=-0.40717626,
pem=-2.6961806, similarity=6.6010103]
1: 16331371 Ballinascarty
[pem_pos=0, final_scores=0.030729935,
final_local_scores=0.14025289, pem=0.0, similarity=8.275139]
2: 46709 County_Cork
[pem_pos=0, final_scores=0.28256267,
final_local_scores=0.3035835, pem=0.0, similarity=13.14925]
3: 23640683 Ballinascorthy_GAA
[pem_pos=0, final_scores=0.21571623,
final_local_scores=0.22380334, pem=0.0, similarity=10.768451]
4: 23640683 Ballinascorthy_GAA
[pem_pos=1, final_scores=0.061515436,
final_local_scores=0.05546376, pem=-0.68016,
similarity=9.443678]
5: 11154395 Carbery_GAA
[pem_pos=0, final_scores=0.16286851,
final_local_scores=0.17931497, pem=-0.067718156,
similarity=9.809092]
6: 7084096 Cork_Senior_Football_Championship
[pem_pos=0, final_scores=0.10813032,
final_local_scores=0.086508125, pem=0.0, similarity=6.671288]
7: 8280565 Munster_Senior_Club_Football_Championship
[pem_pos=0, final_scores=0.08601622,
final_local_scores=0.09800409, pem=0.0, similarity=7.0143504]
8: 3331546 All-Ireland_Senior_Club_Football_Championship
[pem_pos=0, final_scores=0.17730735,
final_local_scores=0.13955986, pem=0.0, similarity=8.254457]
9: 32532531 Cork_GAA
[pem_pos=0, final_scores=0.25042886,
final_local_scores=0.19301513, pem=0.0, similarity=9.849669]

10: 32532531 Cork_GAA
[pem_pos=1, final_scores=0.20697962,
final_local_scores=0.14380556, pem=-1.0852067,
similarity=14.2826805]
11: 2818470 Mayo_GAA
[pem_pos=0, final_scores=0.36483938,
final_local_scores=0.32990313, pem=-1.3491615,
similarity=21.271631]
12: 15845927 Conor_Counihan
[pem_pos=0, final_scores=0.13488911,
final_local_scores=0.10480416, pem=0.0, similarity=7.217278]
1: 14351740 Teddy_Holland
[pem_pos=0, final_scores=0.10593234,
final_local_scores=0.06726922, pem=0.0, similarity=6.097161]
2: 14560 Republic_of_Ireland
[pem_pos=3, final_scores=0.41853207,
final_local_scores=0.45817167, pem=-2.2576945,
similarity=30.040161]
3: 8140815 Cork_(city)
[pem_pos=0, final_scores=0.13149486,
final_local_scores=0.12430984, pem=-0.83752966,
similarity=12.353983]
4: 9647657 Netherlands_national_football_team
[pem_pos=4, final_scores=0.3119921,
final_local_scores=0.3062396, pem=-3.631149,
similarity=32.975266]
5: 32532531 Cork_GAA
[pem_pos=1, final_scores=0.13754077,
final_local_scores=0.06970206, pem=-1.0852067,
similarity=12.071284]
6: 9647657 Netherlands_national_football_team
[pem_pos=4, final_scores=0.23419122,
final_local_scores=0.22125399, pem=-3.631149,
similarity=30.439127]
7: 8140815 Cork_(city)
[pem_pos=0, final_scores=0.23043294,
final_local_scores=0.23635867, pem=-0.83752966,
similarity=15.697744]
8: 9647657 Netherlands_national_football_team
[pem_pos=4, final_scores=0.5838561,
final_local_scores=0.60320973, pem=-3.631149,
similarity=41.83745]
9: 14351740 Teddy_Holland
[pem_pos=0, final_scores=0.4334085,
final_local_scores=0.4240023, pem=0.0, similarity=16.742786]
10: 14351740 Teddy_Holland
[pem_pos=0, final_scores=0.25512874,
final_local_scores=0.22925875, pem=0.0, similarity=10.931252]
11: 780802 All-Ireland_Senior_Hurling_Championship
[pem_pos=0, final_scores=0.03796865,
final_local_scores=-0.018346116, pem=-0.83555937,
similarity=8.08613]

12: 9647657 Netherlands_national_football_team
 [pem_pos=4, final_scores=0.36395234,
 final_local_scores=0.36299825, pem=-3.631149,
 similarity=34.669056]
 13: 14351740 Teddy_Holland
 [pem_pos=0, final_scores=0.2441382,
 final_local_scores=0.21725324, pem=0.0, similarity=10.572983]
 14: 32532531 Cork_GAA
 [pem_pos=1, final_scores=0.089531764,
 final_local_scores=0.017259538, pem=-1.0852067,
 similarity=10.506295]
 15: 1234583 Limerick_F.C.
 [pem_pos=4, final_scores=0.39661676,
 final_local_scores=0.38234884, pem=-2.6802762,
 similarity=30.075521]
 16: 32532531 Cork_GAA
 [pem_pos=1, final_scores=0.22350334,
 final_local_scores=0.16185525, pem=-1.0852067,
 similarity=14.82132]
 17: 985954 Munster_Rugby
 [pem_pos=1, final_scores=0.058182076,
 final_local_scores=0.035650313, pem=-1.883883,
 similarity=15.3984375]
 18: 780802 All-Ireland_Senior_Hurling_Championship
 [pem_pos=0, final_scores=0.055798545,
 final_local_scores=0.0042539686, pem=-0.83555937,
 similarity=8.760562]
 19: 9647657 Netherlands_national_football_team
 [pem_pos=4, final_scores=0.39895582,
 final_local_scores=0.40123427, pem=-3.631149,
 similarity=35.810097]
 20: 780802 All-Ireland_Senior_Hurling_Championship
 [pem_pos=0, final_scores=0.07505016,
 final_local_scores=0.025283411, pem=-0.83555937,
 similarity=9.388123]
 21: 32532531 Cork_GAA
 [pem_pos=1, final_scores=0.11848505,
 final_local_scores=0.048886597, pem=-1.0852067,
 similarity=11.4501095]
 22: 9647657 Netherlands_national_football_team
 [pem_pos=4, final_scores=0.22897933,
 final_local_scores=0.2155608, pem=-3.631149,
 similarity=30.269228]

Sample: b'Wurttemberger'

The Württemberger , Baden-Württemberger or [1 Württemberg 1]
 is a [1 1 Warmblood 1] horse originating in [2 2 Germany 2] .
 They are primarily riding 1 horse s , and are selectively
 bred for 2 dressage and 3 show jumping . The 4 breed registry
 studbook was founded in 1895 , and around 1896 , the 5 Marbach
 stud began concentrating on the original Württemberger type ,
 which was a multi-purpose horse used for both 6 draft and

under-saddle work . The original horses were produced by introducing [2 7 Arabian 2] , [3 3 Trakehner 3] , 8 Anglo-Norman , 9 Friesian , [3 10 Spanish 3] , , and 11 Suffolk Punch blood to local 12 warmblood lines . An Anglo-Norman stallion named Faust was a significant early influence on the breed , and contributed to the original shape of the breed . This type lasted until the mid-1950s , when Marbach recognized , and began breeding a lighter , more athletic horse for modern competition . The Trakehner stallion Julmond -LRB- 1943 -- 1965 -RRB- was the foundation stud of new type , and helped to shape the modern Baden Württemberger . Additions of , , , [4 4 Thoroughbred 4] and additional Trakehner blood helped to further refine the breed into its current form . The best Württemberger horses have excellent temperaments , combining a sensible attitude with lively and free action . These traits may have originally been installed by the use of draft horse and [4 Arabian 4] blood . They are tough and economical to feed when compared to [5 Thoroughbreds 5] . The ideal Württemberger stands around high , and is usually 13 bay , 14 chestnut , 15 brown , or 16 black in 17 color

1: 92431 Warmblood

2: 11867 Germany

3: 31052 Trakehner

4: 18842022 Thoroughbred

1: 13645 Horse

[winner: 13645 Horse, pem_pos=0, final_scores=-1.4563093, final_local_scores=-1.6199279, pem=-0.1558871, similarity=-43.40443]

gt: 13645 Horse, pem_pos=0, final_scores=-1.4563093, final_local_scores=-1.6199279, pem=-0.1558871, similarity=-43.40443]

2: 48929 Dressage

[winner: 2380256 Equestrian_at_the_1952_Summer_Olympics, pem_pos=0, final_scores=-1.1338762, final_local_scores=-1.1533961, pem=0.0, similarity=-30.329927]

gt: 48929 Dressage, pem_pos=1, final_scores=-1.269335, final_local_scores=-1.3931589, pem=-0.0015838762, similarity=-37.47632]

3: 36871 Show_jumping

[winner: 43609 Jumping, pem_pos=0, final_scores=-1.1113843, final_local_scores=-1.1243529, pem=-0.4093905, similarity=-27.236893]

gt: 36871 Show_jumping, pem_pos=1, final_scores=-1.3928864, final_local_scores=-1.5309258, pem=-0.6180137, similarity=-38.235313]

[winner: 1676359 Show_(animal), pem_pos=4, final_scores=-1.6063426, final_local_scores=-1.6831172, pem=-3.1693556, similarity=-28.902409]

RECALL MISS: 36871 Show_jumping]

4: 1235059 Breed_registry

[winner: 21957462 The_Breed_Motorcycle_Club, pem_pos=0,
final_scores=-1.0106887, final_local_scores=-0.9556703,
pem=-0.3750016, similarity=-22.390078]
RECALL MISS: 1235059 Breed_registry]

[winner: 1164933 Windows_Registry, pem_pos=0,
final_scores=-1.0412791, final_local_scores=-0.97396696,
pem=-0.07797931, similarity=-24.55134]
gt: 1235059 Breed_registry, pem_pos=1,
final_scores=-1.7087286, final_local_scores=-1.8815336,
pem=-1.1869378, similarity=-45.604248]

5: 15077573 Marbach_stud
[winner: 1256326 Marbach_am_Neckar, pem_pos=0,
final_scores=-0.1111877, final_local_scores=-0.03691967,
pem=-0.48564446, similarity=5.6289663]
RECALL MISS: 15077573 Marbach_stud]

[winner: 5708497 Wall_stud, pem_pos=5,
final_scores=-1.511224, final_local_scores=-1.5084345,
pem=-3.1495166, similarity=-23.797422]
RECALL MISS: 15077573 Marbach_stud]

6: 517512 Draft_horse
[winner: 9223115 Draft_(hull), pem_pos=0,
final_scores=-1.2199912, final_local_scores=-1.222661,
pem=-1.3245776, similarity=-25.193668]
gt: 517512 Draft_horse, pem_pos=5,
final_scores=-1.6558008, final_local_scores=-1.8317926,
pem=-2.2862294, similarity=-38.14176]

7: 292690 Arabian_horse
[winner: 47858 Arabian_Peninsula, pem_pos=0,
final_scores=0.065824375, final_local_scores=0.066352025,
pem=-0.9173084, similarity=11.058255]
gt: 292690 Arabian_horse, pem_pos=1,
final_scores=-0.018694088, final_local_scores=-0.062345594,
pem=-1.1521257, similarity=8.494635]

8: 19634991 Anglo-Norman_horse
[winner: 11292933 Anglo-Norman_language, pem_pos=1,
final_scores=-0.22342274, final_local_scores=-0.20764405,
pem=-0.59488094, similarity=1.1282554]
gt: 19634991 Anglo-Norman_horse, pem_pos=2,
final_scores=-0.2805328, final_local_scores=-0.2915008,
pem=-1.8624369, similarity=5.5189686]

9: 538413 Friesian_horse
[winner: 538413 Friesian_horse, pem_pos=0,
final_scores=0.022341743, final_local_scores=0.016671121,
pem=0.0, similarity=4.5872126]
gt: 538413 Friesian_horse, pem_pos=0,
final_scores=0.022341743, final_local_scores=0.016671121,
pem=0.0, similarity=4.5872126]

10: 1133097 Iberian_horse
[winner: 26667 Spain, pem_pos=1, final_scores=0.39129084,
final_local_scores=0.4301628, pem=-1.2425895,
similarity=23.684023]
RECALL MISS: 1133097 Iberian_horse]

11: 929458 Suffolk_Punch
 [winner: 929458 Suffolk_Punch, pem_pos=0,
 final_scores=-0.006300077, final_local_scores=0.0044754744,
 pem=0.0, similarity=4.22327|
 gt: 929458 Suffolk_Punch, pem_pos=0,
 final_scores=-0.006300077, final_local_scores=0.0044754744,
 pem=0.0, similarity=4.22327]

12: 92431 Warmblood
 [winner: 92431 Warmblood, pem_pos=0,
 final_scores=-1.0064068, final_local_scores=-1.1172875,
 pem=0.0, similarity=-29.252375|
 gt: 92431 Warmblood, pem_pos=0, final_scores=-1.0064068,
 final_local_scores=-1.1172875, pem=0.0,
 similarity=-29.252375]

13: 913743 Bay_(horse)
 [winner: 16819546 Lazarev_Bay, pem_pos=17,
 final_scores=-1.2221675, final_local_scores=-1.1864347,
 pem=-2.9935458, similarity=-15.036501|
 gt: 913743 Bay_(horse), pem_pos=2,
 final_scores=-1.3768728, final_local_scores=-1.4743195,
 pem=-1.5024016, similarity=-31.736628]

14: 3746113 Chestnut_(coat)
 [winner: 67733 Chestnut, pem_pos=1,
 final_scores=-1.2290546, final_local_scores=-1.3029726,
 pem=-0.6847083, similarity=-31.07004|
 gt: 3746113 Chestnut_(coat), pem_pos=0,
 final_scores=-1.2820408, final_local_scores=-1.3615787,
 pem=-0.5853683, similarity=-33.359184]

15: 913743 Bay_(horse)
 [winner: 22891336 Brown_California, pem_pos=3,
 final_scores=-1.4576246, final_local_scores=-1.4227729,
 pem=-3.3700302, similarity=-20.041922|
 RECALL MISS: 913743 Bay_(horse)]

16: 6736865 Black_(horse)
 [winner: 273285
 Race_and_ethnicity_in_the_United_States_Census, pem_pos=0,
 final_scores=-1.1341847, final_local_scores=-1.1066685,
 pem=-1.1573607, similarity=-22.641571|
 RECALL MISS: 6736865 Black_(horse)]

17: 5596470 Equine_coat_color
 [winner: 180870 Color_charge, pem_pos=2,
 final_scores=-1.1814514, final_local_scores=-1.1443756,
 pem=-2.0534868, similarity=-18.893553|
 RECALL MISS: 5596470 Equine_coat_color]

1: 92431 Warmblood
 [pem_pos=0, final_scores=0.070648625,
 final_local_scores=0.08082561, pem=0.0, similarity=6.5017104]

2: 11867 Germany
 [pem_pos=0, final_scores=0.76480776,
 final_local_scores=0.8101738, pem=-0.15529114,
 similarity=29.111414]

3: 31052 Trakehner

[pem_pos=0, final_scores=0.03912525,
final_local_scores=0.0397927, pem=0.0, similarity=5.2772064]
4: 18842022 Thoroughbred
[pem_pos=0, final_scores=0.23935299,
final_local_scores=0.2639008, pem=-0.02425665,
similarity=12.096951]
1: 40981147 Württemberg
[pem_pos=0, final_scores=0.105233476,
final_local_scores=0.12176409, pem=-0.2705484,
similarity=9.194681]
2: 47858 Arabian_Peninsula
[pem_pos=0, final_scores=0.065824375,
final_local_scores=0.066352025, pem=-0.9173084,
similarity=11.058255]
3: 26667 Spain
[pem_pos=1, final_scores=0.39129084,
final_local_scores=0.4301628, pem=-1.2425895,
similarity=23.684023]
4: 47858 Arabian_Peninsula
[pem_pos=0, final_scores=0.06638871,
final_local_scores=0.07953866, pem=-0.9173084,
similarity=11.451771]
5: 18842022 Thoroughbred
[pem_pos=0, final_scores=0.1421415,
final_local_scores=0.15771204, pem=-0.04130381,
similarity=9.020772]

Sample: b"Chateau_d'Oiron"

The 1 Château d'Oiron is located in 2 Oiron , in the [1 1 Deux-Sèvres 1] of western [2 2 France 2] . It has its origins in the 15th century war with the [1 English 1] for control of [2 France 2] when a victorious [3 3 Charles VII 3] of [4 France 4] gave the domain and great forest of Oiron to [3 3 Guillaume Gouffier 3] who became governor of [4 4 Touraine 4] . This château is the background for [5 5 Charles Perrault 5] 's fairy tale , 4 Puss in Boots . 5 King [5 Louis XIV 5] 's mistress , was one of the residents in the place . Today , Oiron is only a short drive from the royal châteaux in the [6 6 Val de Loire 6] , but in the 15th century it was considered far removed from the seat of power at the royal domains . Nonetheless , [6 Guillaume Gouffier 6] built a magnificent château and his offspring updated and improved it . In 1538 , his daughter-in-law , , was responsible for the construction of a collegiate church adjacent to the château . In 1551 , [7 6 Henry II 7] and his entire court were guests of [7 7 Claude Gouffier 7] who had been granted the title Marquis de Caravaz . [8 8 Claude Gouffier 8] served as the model for [8 Charles Perrault 8] 's `` Marquis de Carabas `` in his story , Puss in Boots . Two generations later , another [9 Gouffier 9] was exiled from the king 's court by [9 9 Cardinal Richelieu 9] in 1620 . In the mid 17th century , Charlotte Gouffier became enamored with the renowned intellect , [10 10

Blaise Pascal [10] -LRB- 1623 -- 62 -RRB- , who spent considerable time at the [10 Château 10] d'Oiron . After Pascal died , Charlotte Gouffier married 7 Francois d'Aubusson , the duc de La Feuillade , who enhanced the castle with his wealth and connections to [11 11 Louis XIV 11] . With the renovations , the castle ended up with a main building and two long projecting wings , one of which is a [12 12 Renaissance 12] structure built over a cloister . One of the galleries contains one of the most prestigious works of art from the [11 French Renaissance 11] period . The [12 Duc de La Feuillade 12] 's son sold the château to [13 8 Louis XIV 13] 's mistress , who lived there for the rest of her life . Her son had little interest in the property as he preferred to be much closer to the royal court so in 1736 he sold the château to the . After that , the château went into severe decline and in 1793 was ransacked by 9 Revolutionaries . For many years the château lay abandoned until the government of [14 France 14] took possession just before [13 13 World War II 13] eventually converting it to a museum . Recognized world wide , the museum is dedicated to contemporary art . In 1993 , The Year of Solar Burns was commissioned by the for permanent installation in the Château d'Oiron . The château has been listed as a 10 monument historique

1: 90607 Deux-Sèvres

2: 5843419 France

3: 308077 Guillaume_Gouffier,_seigneur_de_Bonnivet

4: 79849 Touraine

5: 87644 Charles_Perrault

6: 81870 Loire_Valley

7: 25718950 Claude_Gouffier

8: 25718950 Claude_Gouffier

9: 85254 Cardinal_Richelieu

10: 4068 Blaise_Pascal

11: 18553 Louis_XIV_of_France

12: 25532 Renaissance

13: 32927 World_War_II

1: 61914 Château

[winner: 61914 Château, pem_pos=0, final_scores=-0.09027678, final_local_scores=-0.14027886, pem=-0.5944852, similarity=3.136415]

gt: 61914 Château, pem_pos=0, final_scores=-0.09027678, final_local_scores=-0.14027886, pem=-0.5944852, similarity=3.136415]

2: 1521639 Oiron

[winner: 89451 Château_d'Oiron, pem_pos=1, final_scores=-0.13802335, final_local_scores=-0.13290444, pem=-0.6714465, similarity=3.7750094]

gt: 1521639 Oiron, pem_pos=0, final_scores=-0.23133299, final_local_scores=-0.09513357, pem=0.0, similarity=1.2507374]

3: 76701 Charles_VII_of_France

[winner: 5843419 France, pem_pos=1,

final_scores=0.44757718, final_local_scores=0.43138754,
 pem=-0.16240834, similarity=17.846378|
 gt: 76701 Charles_VII_of_France, pem_pos=0,
 final_scores=-0.13461457, final_local_scores=-0.21389195,
 pem=0.0, similarity=-2.2932494]
 [winner: 76701 Charles_VII_of_France, pem_pos=0,
 final_scores=0.33902353, final_local_scores=0.3113134,
 pem=0.0, similarity=13.379926|
 gt: 76701 Charles_VII_of_France, pem_pos=0,
 final_scores=0.33902353, final_local_scores=0.3113134,
 pem=0.0, similarity=13.379926]
 4: 23194446 Puss_in_Boots
 [winner: 23194446 Puss_in_Boots, pem_pos=0,
 final_scores=-0.11015575, final_local_scores=-0.039955385,
 pem=-0.54895705, similarity=5.8826785|
 gt: 23194446 Puss_in_Boots, pem_pos=0,
 final_scores=-0.11015575, final_local_scores=-0.039955385,
 pem=-0.54895705, similarity=5.8826785]
 5: 18553 Louis_XIV_of_France
 [winner: 18553 Louis_XIV_of_France, pem_pos=0,
 final_scores=0.3242014, final_local_scores=0.25294292,
 pem=0.0, similarity=11.638034|
 gt: 18553 Louis_XIV_of_France, pem_pos=0,
 final_scores=0.3242014, final_local_scores=0.25294292,
 pem=0.0, similarity=11.638034]
 [winner: 497665 Ubu_Roi, pem_pos=3,
 final_scores=-1.1830525, final_local_scores=-1.2644634,
 pem=-1.0986123, similarity=-27.669977|
 gt: 18553 Louis_XIV_of_France, pem_pos=10,
 final_scores=-1.3653582, final_local_scores=-1.5941124,
 pem=-2.7080503, similarity=-28.754984]
 6: 75910 Henry_II_of_France
 [winner: 26042562 Henry_II_of_England, pem_pos=0,
 final_scores=0.10013868, final_local_scores=0.068448454,
 pem=-0.49336484, similarity=8.815347|
 gt: 75910 Henry_II_of_France, pem_pos=2,
 final_scores=0.0040840358, final_local_scores=-0.06943309,
 pem=-1.8793458, similarity=12.237865]
 7: 30901068 François_d'Aubusson_de_La_Feuillade
 [winner: 15575846 La_Feuillade, pem_pos=0,
 final_scores=-0.09717938, final_local_scores=0.023375511,
 pem=0.0, similarity=4.787285|
 gt: 30901068 François_d'Aubusson_de_La_Feuillade,
 pem_pos=1, final_scores=-1.0548186,
 final_local_scores=-1.101531, pem=-6.5308776,
 similarity=6.7337604]
 [winner: 50012 Francis_I_of_France, pem_pos=3,
 final_scores=-0.39012545, final_local_scores=-0.51295197,
 pem=-2.6797566, similarity=3.3551357|
 RECALL MISS: 30901068
 François_d'Aubusson_de_La_Feuillade]
 [winner: 26273795 Duc_de_Piney, pem_pos=3,

final_scores=-1.2812443, final_local_scores=-1.3343859,
 pem=-2.3978953, similarity=-22.690895|
 RECALL MISS: 30901068
 François_d'Aubusson_de_La_Feuillade]
 [winner: 717213 .de, pem_pos=1, final_scores=-1.4876993,
 final_local_scores=-1.4717417, pem=-1.897993,
 similarity=-29.508415|
 RECALL MISS: 30901068
 François_d'Aubusson_de_La_Feuillade]
 [winner: 28260 Sonnet, pem_pos=0, final_scores=-1.691063,
 final_local_scores=-1.808481, pem=0.0, similarity=-49.878967|
 RECALL MISS: 30901068
 François_d'Aubusson_de_La_Feuillade]
 8: 18553 Louis_XIV_of_France
 [winner: 18553 Louis_XIV_of_France, pem_pos=0,
 final_scores=0.423194, final_local_scores=0.3620992, pem=0.0,
 similarity=14.895475|
 gt: 18553 Louis_XIV_of_France, pem_pos=0,
 final_scores=0.423194, final_local_scores=0.3620992, pem=0.0,
 similarity=14.895475]
 [winner: 187801 Shopping_cart, pem_pos=8,
 final_scores=-2.2736847, final_local_scores=-2.335067,
 pem=-3.4067376, similarity=-47.06698|
 RECALL MISS: 18553 Louis_XIV_of_France]
 9: 11188 French_Revolution
 [winner: 4141439
 Revolutionary_movement_for_Indian_independence, pem_pos=0,
 final_scores=-0.045938924, final_local_scores=0.056058496,
 pem=-0.22743031, similarity=6.9994116|
 gt: 11188 French_Revolution, pem_pos=1,
 final_scores=-0.06642152, final_local_scores=-0.14788896,
 pem=-1.0377641, similarity=5.3199334]
 10: 3608419 Monument_historique
 [winner: 1616263 Tropaeum_Traiani, pem_pos=0,
 final_scores=-0.712048, final_local_scores=-0.6616927,
 pem=0.0, similarity=-15.656517|
 RECALL MISS: 3608419 Monument_historique]
 [winner: 5986847 Polish_Chess_Championship, pem_pos=5,
 final_scores=-1.1213846, final_local_scores=-1.0924256,
 pem=-2.4981518, similarity=-14.925108|
 RECALL MISS: 3608419 Monument_historique]
 1: 90607 Deux-Sèvres
 [pem_pos=0, final_scores=0.25199693,
 final_local_scores=0.29970163, pem=-0.092858076,
 similarity=13.5383835]
 2: 5843419 France
 [pem_pos=1, final_scores=0.49330598,
 final_local_scores=0.47917986, pem=-0.16240834,
 similarity=19.272596]
 3: 308077 Guillaume_Gouffier,_seigneur_de_Bonnivet
 [pem_pos=0, final_scores=0.3398388,
 final_local_scores=0.3898024, pem=0.0, similarity=15.722194]

4: 79849 Touraine
[pem_pos=0, final_scores=0.43578738,
final_local_scores=0.43055266, pem=0.0, similarity=16.938263]
5: 87644 Charles_Perrault
[pem_pos=0, final_scores=0.2720847,
final_local_scores=0.30774838, pem=0.0, similarity=13.273539]
6: 81870 Loire_Valley
[pem_pos=0, final_scores=0.4721474,
final_local_scores=0.47482717, pem=0.0, similarity=18.259502]
7: 25718950 Claude_Gouffier
[pem_pos=0, final_scores=0.19408129,
final_local_scores=0.23051351, pem=0.0, similarity=10.968696]
8: 25718950 Claude_Gouffier
[pem_pos=0, final_scores=0.20316254,
final_local_scores=0.2404334, pem=0.0, similarity=11.264725]
9: 85254 Cardinal_Richelieu
[pem_pos=0, final_scores=0.17335351,
final_local_scores=0.11676192, pem=0.0, similarity=7.574122]
10: 4068 Blaise_Pascal
[pem_pos=0, final_scores=0.08353911,
final_local_scores=0.11757344, pem=0.0, similarity=7.5983386]
11: 18553 Louis_XIV_of_France
[pem_pos=0, final_scores=0.4692819,
final_local_scores=0.41244322, pem=0.0, similarity=16.39784]
12: 25532 Renaissance
[pem_pos=0, final_scores=0.18993993,
final_local_scores=0.18440688, pem=-0.26284593,
similarity=11.022179]
13: 32927 World_War_II
[pem_pos=0, final_scores=0.45470518,
final_local_scores=0.49567562, pem=0.0, similarity=18.881662]
1: 9316 England
[pem_pos=1, final_scores=0.34153295,
final_local_scores=0.37637502, pem=-1.3527626,
similarity=22.678026]
2: 5843419 France
[pem_pos=1, final_scores=0.5942059,
final_local_scores=0.59057355, pem=-0.16240834,
similarity=22.596807]
3: 76701 Charles_VII_of_France
[pem_pos=0, final_scores=0.33902353,
final_local_scores=0.3113134, pem=0.0, similarity=13.379926]
4: 5843419 France
[pem_pos=1, final_scores=0.44757718,
final_local_scores=0.43138754, pem=-0.16240834,
similarity=17.846378]
5: 18553 Louis_XIV_of_France
[pem_pos=0, final_scores=0.3242014,
final_local_scores=0.25294292, pem=0.0, similarity=11.638034]
6: 308077 Guillaume_Gouffier,_seigneur_de_Bonnivet
[pem_pos=0, final_scores=0.33340228,
final_local_scores=0.38277143, pem=0.0, similarity=15.512376]

7: 26042562 Henry_II_of_England
 [pem_pos=0, final_scores=0.10013868,
 final_local_scores=0.068448454, pem=-0.49336484,
 similarity=8.815347]
 8: 87644 Charles_Perrault
 [pem_pos=0, final_scores=0.17047529,
 final_local_scores=0.1967555, pem=0.0, similarity=9.961289]
 9: 308077 Guillaume_Gouffier,_seigneur_de_Bonnivet
 [pem_pos=0, final_scores=0.08112724,
 final_local_scores=0.10764377, pem=0.0, similarity=7.3020177]
 10: 61914 Château
 [pem_pos=0, final_scores=0.261347,
 final_local_scores=0.2506668, pem=-0.5944852,
 similarity=14.803013]
 11: 1398990 French_Renaissance
 [pem_pos=0, final_scores=0.09974508,
 final_local_scores=0.079606324, pem=0.0, similarity=6.4653244]
 12: 30901068 François_d'Aubusson_de_La_Feuillade
 [pem_pos=0, final_scores=0.16461082,
 final_local_scores=0.2380946, pem=0.0, similarity=11.194931]
 13: 18553 Louis_XIV_of_France
 [pem_pos=0, final_scores=0.423194,
 final_local_scores=0.3620992, pem=0.0, similarity=14.895475]
 14: 5843419 France
 [pem_pos=1, final_scores=0.5989246,
 final_local_scores=0.5967116, pem=-0.16240834,
 similarity=22.779978]

Sample: b'Malcolm_Azania'

Malcolm Azania -LRB- born 1969 -RRB- , also known as Minister Faust -LRB- for the literary reference , see 1 Faust -RRB- , is a [1 1 Canadian 1] teacher , writer , community activist , radio host and political aspirant . He appeared on [2 2 CBC Television 2] 's in 2003 as one of three contestants in the `` Political Animal '' segment . He appeared alongside [3 3 Bridget Pastoor 3] , who would later become a 2 MLA for the [4 4 Alberta Liberal Party 4] in in 2004 . A resident of [5 5 Edmonton 5] , [6 6 Alberta 6] , Azania ran as a candidate for the in the electoral district of [1 3 Edmonton 1] -- Strathcona in the 4 2004 federal election . The riding and Mr. Azania were profiled by in its 5 Campaign Politics TV show . Azania finished third , behind both 6 Conservative incumbent 7 Member of Parliament -LRB- MP -RRB- [7 7 Rahim Jaffer 7] and 8 Liberal challenger and former [2 Alberta 2] 9 Member of the Legislative Assembly [8 8 Debby Carlson 8] . Also as Minister Faust , Azania wrote a comical and satiric 10 science fiction novel , 11 The Coyote Kings of the Space-Age Bachelor Pad , released internationally by [9 9 Random House 9] in August 2004 . In 2007 , he published his second novel , 12 From the Notebooks of Dr. Brain

1: 5042916 Canada

2: 1288819 CBC_Television

3: 8599579 Bridget_Pastoor
 4: 482693 Alberta_Liberal_Party
 5: 95405 Edmonton
 6: 717 Alberta
 7: 385673 Rahim_Jaffer
 8: 1159856 Debby_Carlson
 9: 182947 Random_House
 1: 84672 Faust
 [winner: 712909 Malcolm_Azania, pem_pos=0,
 final_scores=-0.3681808, final_local_scores=-0.39824286,
 pem=0.0, similarity=-7.794648]
 gt: 84672 Faust, pem_pos=1, final_scores=-0.9124394,
 final_local_scores=-0.90812373, pem=-0.84974116,
 similarity=-18.389486]
 2: 1864012 Member_of_the_Legislative_Assembly
 [winner: 1864012 Member_of_the_Legislative_Assembly,
 pem_pos=0, final_scores=-0.36885735,
 final_local_scores=-0.375909, pem=-1.337187,
 similarity=0.14366952]
 gt: 1864012 Member_of_the_Legislative_Assembly,
 pem_pos=0, final_scores=-0.36885735,
 final_local_scores=-0.375909, pem=-1.337187,
 similarity=0.14366952]
 3: 562215 Edmonton-Strathcona
 [winner: 95405 Edmonton, pem_pos=0,
 final_scores=0.4188419, final_local_scores=0.39050484,
 pem=-0.27044967, similarity=17.213903]
 RECALL MISS: 562215 Edmonton-Strathcona]
 [winner: 1418242 Strathcona_County, pem_pos=3,
 final_scores=-0.23198707, final_local_scores=-0.24138895,
 pem=-2.075476, similarity=8.172946]
 gt: 562215 Edmonton-Strathcona, pem_pos=18,
 final_scores=-0.56884515, final_local_scores=-0.5960686,
 pem=-3.0910425, similarity=3.1114054]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.3267057, final_local_scores=-1.2260053,
 pem=0.0, similarity=-32.496735]
 RECALL MISS: 562215 Edmonton-Strathcona]
 4: 312153 Canadian_federal_election,_2004
 [winner: 312153 Canadian_federal_election,_2004,
 pem_pos=0, final_scores=-0.49435425,
 final_local_scores=-0.5717487, pem=0.0,
 similarity=-12.972408]
 gt: 312153 Canadian_federal_election,_2004, pem_pos=0,
 final_scores=-0.49435425, final_local_scores=-0.5717487,
 pem=0.0, similarity=-12.972408]
 5: 3762019 Campaign_Politics
 [winner: 3762019 Campaign_Politics, pem_pos=0,
 final_scores=-0.57942814, final_local_scores=-0.52423185,
 pem=0.0, similarity=-11.554411]
 gt: 3762019 Campaign_Politics, pem_pos=0,
 final_scores=-0.57942814, final_local_scores=-0.52423185,

pem=0.0, similarity=-11.554411]

6: 308647 Conservative_Party_of_Canada
 [winner: 32113 Conservative_Party_(UK), pem_pos=0,
 final_scores=-0.11657846, final_local_scores=-0.11853221,
 pem=-0.55914843, similarity=3.5932114]
 gt: 308647 Conservative_Party_of_Canada, pem_pos=1,
 final_scores=-0.25941896, final_local_scores=-0.3320834,
 pem=-2.1512978, similarity=5.878774]

7: 98066 Member_of_Parliament
 [winner: 98066 Member_of_Parliament, pem_pos=0,
 final_scores=-0.33304554, final_local_scores=-0.32720178,
 pem=-0.04316567, similarity=-5.439899]
 gt: 98066 Member_of_Parliament, pem_pos=0,
 final_scores=-0.33304554, final_local_scores=-0.32720178,
 pem=-0.04316567, similarity=-5.439899]

8: 149536 Liberal_Party_of_Canada
 [winner: 149536 Liberal_Party_of_Canada, pem_pos=0,
 final_scores=-0.12534863, final_local_scores=-0.16826461,
 pem=-1.2222316, similarity=5.715047]
 gt: 149536 Liberal_Party_of_Canada, pem_pos=0,
 final_scores=-0.12534863, final_local_scores=-0.16826461,
 pem=-1.2222316, similarity=5.715047]

9: 1864012 Member_of_the_Legislative_Assembly
 [winner: 222426 Legislative_Assembly, pem_pos=0,
 final_scores=-0.77623415, final_local_scores=-0.8120555,
 pem=-1.839396, similarity=-10.140723]
 RECALL MISS: 1864012 Member_of_the_Legislative_Assembly]
 [winner: 682629 Element_(mathematics), pem_pos=6,
 final_scores=-1.096766, final_local_scores=-1.0266507,
 pem=-2.3025851, similarity=-14.025775]
 gt: 1864012 Member_of_the_Legislative_Assembly,
 pem_pos=1, final_scores=-1.3214881,
 final_local_scores=-1.4165134, pem=-2.0809023,
 similarity=-26.865604]
 [winner: 10447732

Results_by_riding_of_the_Canadian_federal_election,_2008,
 pem_pos=12, final_scores=-1.2886336,
 final_local_scores=-1.2282925, pem=-3.6198866,
 similarity=-12.879477]
 RECALL MISS: 1864012 Member_of_the_Legislative_Assembly]

10: 26787 Science_fiction
 [winner: 6921949 Fiction_(magazine), pem_pos=6,
 final_scores=-1.1176525, final_local_scores=-1.0741735,
 pem=-2.076133, similarity=-16.675432]
 RECALL MISS: 26787 Science_fiction]
 [winner: 430822 Doctor_of_Science, pem_pos=7,
 final_scores=-1.5369761, final_local_scores=-1.5871565,
 pem=-2.0794415, similarity=-31.965878]
 RECALL MISS: 26787 Science_fiction]

11: 1171234 The_Coyote_Kings_of_the_Space-Age_Bachelor_Pad
 [winner: 1171234 The_Coyote_Kings_of_the_Space-
 Age_Bachelor_Pad, pem_pos=0, final_scores=-0.15294424,

final_local_scores=-0.061792642, pem=0.0,
similarity=2.2456973]
gt: 1171234 The_Coyote_Kings_of_the_Space-
Age_Bachelor_Pad, pem_pos=0, final_scores=-0.15294424,
final_local_scores=-0.061792642, pem=0.0,
similarity=2.2456973]
12: 10648228 From_the_Notebooks_of_Dr._Brain
[winner: 3502080 Notebooks_of_Henry_James, pem_pos=0,
final_scores=-0.51866215, final_local_scores=-0.43784687,
pem=0.0, similarity=-8.976511]
RECALL MISS: 10648228 From_the_Notebooks_of_Dr._Brain]
[winner: 15568949 Brain,_Côte-d'Or, pem_pos=1,
final_scores=-0.8225666, final_local_scores=-0.6175643,
pem=-1.6839364, similarity=-5.182131]
RECALL MISS: 10648228 From_the_Notebooks_of_Dr._Brain]
[winner: 4720231 Mac_(Green_Wing), pem_pos=6,
final_scores=-0.8855035, final_local_scores=-0.8354004,
pem=-1.7917595, similarity=-11.096435]
RECALL MISS: 10648228 From_the_Notebooks_of_Dr._Brain]
1: 5042916 Canada
[pem_pos=0, final_scores=0.50363207,
final_local_scores=0.49487782, pem=-0.41303828,
similarity=21.10402]
2: 1288819 CBC_Television
[pem_pos=0, final_scores=0.25116235,
final_local_scores=0.26693422, pem=0.0, similarity=12.055562]
3: 8599579 Bridget_Pastoor
[pem_pos=0, final_scores=0.23674159,
final_local_scores=0.3141762, pem=0.0, similarity=13.465358]
4: 482693 Alberta_Liberal_Party
[pem_pos=0, final_scores=0.31761283,
final_local_scores=0.33384547, pem=0.0, similarity=14.052327]
5: 95405 Edmonton
[pem_pos=0, final_scores=0.54422873,
final_local_scores=0.5274708, pem=-0.27044967,
similarity=21.30124]
6: 717 Alberta
[pem_pos=0, final_scores=0.45054418,
final_local_scores=0.43255854, pem=-0.06637208,
similarity=17.359064]
7: 385673 Rahim_Jaffer
[pem_pos=0, final_scores=0.2738924,
final_local_scores=0.33995214, pem=0.0, similarity=14.234563]
8: 1159856 Debby_Carlson
[pem_pos=0, final_scores=0.17925988,
final_local_scores=0.26441482, pem=0.0, similarity=11.980378]
9: 182947 Random_House
[pem_pos=0, final_scores=0.23841272,
final_local_scores=0.30598986, pem=0.0, similarity=13.221061]
1: 95405 Edmonton
[pem_pos=0, final_scores=0.4188419,
final_local_scores=0.39050484, pem=-0.27044967,

```
similarity=17.213903]
2: 717 Alberta
    [pem_pos=0, final_scores=0.33557945,
final_local_scores=0.30697706, pem=-0.06637208,
similarity=13.611463]
```

Sample: b'Messier_49'

[1 Messier 1] 49 -LRB- also known as M 49 or NGC 4472 -RRB- is an 1 elliptical galaxy located about away in the 2 equatorial 3 constellation of 4 Virgo . This galaxy was discovered by [2 French 2] astronomer [1 1 Charles Messier 1] on February 19 , 1771 . [3 Messier 3] 49 is positioned 4.1 ° west-southwest of the star 5 Epsilon Virginis . As an elliptical galaxy , [4 Messier 4] 49 has the physical form of a 6 radio galaxy , but it only has the radio emission of a normal galaxy . From the detected radio emission , the core region is emitting roughly 1053 erg -LRB- 1046 7 J or 1022 YJ -RRB- of energy . The 8 nucleus of this galaxy is emitting X-rays , suggests the likely presence of a 9 supermassive black hole with an estimated mass of , or 565 million times the mass of the Sun . X-ray emissions shows a structure to the north of [5 Messier 5] 49 that resembles a 10 bow shock . To the southwest of the core , the luminous outline of the galaxy can be traced out to a distance of 260 kpc . The only 1 1 supernova event observed within this galaxy is , discovered in June 1969 . This galaxy has a large collection of globular clusters , estimated at about 5,900 . However , this count is far exceeded by the 13,450 globular clusters orbiting the supergiant elliptical galaxy [6 12 Messier 6] 87 . On average , the globular clusters of M 49 are about 10 billion years old . Between 2000 -- 2009 , strong evidence for was discovered in an M 49 cluster . A second candidate was announced in 2011 . [7 Messier 7] 49 was the first member of the 13 Virgo Cluster of galaxies to be discovered . It is the most luminous member of that cluster and more luminous than any galaxy closer to the Earth . This galaxy forms part of the smaller Virgo B subcluster located 4.5 ° away from the dynamic center of the [8 Virgo Cluster 8] , centered on [9 14 Messier 9] 87 . [10 Messier 10] 49 is gravitationally interacting with the 15 dwarf 16 irregular galaxy

1: 7246 Charles_Messier

1: 207619 Elliptical_galaxy

```
[winner: 207619 Elliptical_galaxy, pem_pos=0,
final_scores=-0.96817476, final_local_scores=-1.0316024,
pem=-0.00062914606, similarity=-26.691936]
```

```
gt: 207619 Elliptical_galaxy, pem_pos=0,
final_scores=-0.96817476, final_local_scores=-1.0316024,
pem=-0.00062914606, similarity=-26.691936 ]
```

2: 153681 Celestial_equator

```
[winner: 1055334 Equatorial_mount, pem_pos=1,
final_scores=-0.98831886, final_local_scores=-0.99344355,
pem=-0.6256887, similarity=-22.154032]
```

gt: 153681 Celestial_equator, pem_pos=4,
final_scores=-1.339809, final_local_scores=-1.4293756,
pem=-2.4211614, similarity=-25.399057]

3: 5267 Constellation
[winner: 5267 Constellation, pem_pos=0,
final_scores=-1.1092261, final_local_scores=-1.1988862,
pem=-0.14364368, similarity=-30.906286]
gt: 5267 Constellation, pem_pos=0,
final_scores=-1.1092261, final_local_scores=-1.1988862,
pem=-0.14364368, similarity=-30.906286]

4: 32407 Virgo_(constellation)
[winner: 32407 Virgo_(constellation), pem_pos=1,
final_scores=-0.13733436, final_local_scores=-0.14772451,
pem=-0.69043946, similarity=3.4360356]
gt: 32407 Virgo_(constellation), pem_pos=1,
final_scores=-0.13733436, final_local_scores=-0.14772451,
pem=-0.69043946, similarity=3.4360356]

5: 2679233 Epsilon_Virginis
[winner: 2679233 Epsilon_Virginis, pem_pos=0,
final_scores=-0.04385826, final_local_scores=0.05458522,
pem=0.0, similarity=5.7186446]
gt: 2679233 Epsilon_Virginis, pem_pos=0,
final_scores=-0.04385826, final_local_scores=0.05458522,
pem=0.0, similarity=5.7186446]

6: 211168 Radio_galaxy
[winner: 211168 Radio_galaxy, pem_pos=0,
final_scores=-0.9190015, final_local_scores=-0.97137934,
pem=0.0, similarity=-24.898184]
gt: 211168 Radio_galaxy, pem_pos=0,
final_scores=-0.9190015, final_local_scores=-0.97137934,
pem=0.0, similarity=-24.898184]

7: 16327 Joule
[winner: 372347 Jacksonian_democracy, pem_pos=4,
final_scores=-0.69368297, final_local_scores=-0.57855904,
pem=-2.7404957, similarity=1.7275978]
gt: 16327 Joule, pem_pos=3, final_scores=-0.803894,
final_local_scores=-0.75923085, pem=-2.566722,
similarity=-4.6090164]

8: 177092 Active_galactic_nucleus
[winner: 3899771 Comet_nucleus, pem_pos=3,
final_scores=-1.2367554, final_local_scores=-1.3030171,
pem=-1.5899647, similarity=-26.148443]
gt: 177092 Active_galactic_nucleus, pem_pos=2,
final_scores=-1.3407315, final_local_scores=-1.4504442,
pem=-1.514572, similarity=-30.957958]

9: 215706 Supermassive_black_hole
[winner: 4650 Black_hole, pem_pos=0,
final_scores=-0.78941005, final_local_scores=-0.8275454,
pem=-0.037721857, similarity=-20.400755]
gt: 215706 Supermassive_black_hole, pem_pos=4,
final_scores=-1.6415489, final_local_scores=-1.7566831,
pem=-5.260577, similarity=-19.725378]

[winner: 215706 Supermassive_black_hole, pem_pos=0,
final_scores=-0.93997025, final_local_scores=-0.9903151,
pem=0.0, similarity=-25.463264|
gt: 215706 Supermassive_black_hole, pem_pos=0,
final_scores=-0.93997025, final_local_scores=-0.9903151,
pem=0.0, similarity=-25.463264]
10: 1020537 Bow_shock
[winner: 1020537 Bow_shock, pem_pos=0,
final_scores=-0.9973302, final_local_scores=-1.026972,
pem=-0.17994747, similarity=-25.578598|
gt: 1020537 Bow_shock, pem_pos=0,
final_scores=-0.9973302, final_local_scores=-1.026972,
pem=-0.17994747, similarity=-25.578598]
11: 27680 Supernova
[winner: 14774765 Supernova/Karma, pem_pos=3,
final_scores=-1.1707416, final_local_scores=-1.0840027,
pem=-1.0986123, similarity=-22.284666|
gt: 27680 Supernova, pem_pos=0, final_scores=-1.2936894,
final_local_scores=-1.3918078, pem=-0.030021273,
similarity=-37.281345]
12: 214077 Messier_87
[winner: 7246 Charles_Messier, pem_pos=0,
final_scores=0.20458241, final_local_scores=0.10586682,
pem=0.0, similarity=7.24899|
RECALL MISS: 214077 Messier_87]
[winner: 14439860 GPR87, pem_pos=16,
final_scores=-1.1376643, final_local_scores=-0.9860327,
pem=-3.8215787, similarity=-4.5531306|
RECALL MISS: 214077 Messier_87]
13: 214067 Virgo_Cluster
[winner: 214067 Virgo_Cluster, pem_pos=0,
final_scores=-0.082747206, final_local_scores=-0.084038146,
pem=0.0, similarity=1.5818471|
gt: 214067 Virgo_Cluster, pem_pos=0,
final_scores=-0.082747206, final_local_scores=-0.084038146,
pem=0.0, similarity=1.5818471]
14: 214077 Messier_87
[winner: 7246 Charles_Messier, pem_pos=0,
final_scores=0.1942135, final_local_scores=0.09226319,
pem=0.0, similarity=6.8430305|
RECALL MISS: 214077 Messier_87]
[winner: 14439860 GPR87, pem_pos=16,
final_scores=-1.2126964, final_local_scores=-1.0679939,
pem=-3.8215787, similarity=-6.999012|
RECALL MISS: 214077 Messier_87]
15: 964475 Dwarf_galaxy
[winner: 207621 Irregular_galaxy, pem_pos=0,
final_scores=-0.72346395, final_local_scores=-0.7521833,
pem=0.0, similarity=-18.356937|
RECALL MISS: 964475 Dwarf_galaxy]
16: 207621 Irregular_galaxy
[winner: 207621 Irregular_galaxy, pem_pos=0,

```

final_scores=-0.72346395, final_local_scores=-0.7521833,
pem=0.0, similarity=-18.356937|
    gt: 207621 Irregular_galaxy, pem_pos=0,
final_scores=-0.72346395, final_local_scores=-0.7521833,
pem=0.0, similarity=-18.356937 ]
    [winner: 12558 Galaxy, pem_pos=0,
final_scores=-0.9112407, final_local_scores=-0.9903066,
pem=-0.20692216, similarity=-24.337738|
    RECALL MISS: 207621 Irregular_galaxy]
1: 7246 Charles_Messier
    [pem_pos=0, final_scores=0.33430183,
final_local_scores=0.2452884, pem=0.0, similarity=11.409608]
1: 7246 Charles_Messier
    [pem_pos=0, final_scores=0.12552206,
final_local_scores=0.01722823, pem=0.0, similarity=4.6038375]
2: 5843419 France
    [pem_pos=0, final_scores=0.5491944,
final_local_scores=0.70147437, pem=-0.80223274,
similarity=29.385775]
3: 7246 Charles_Messier
    [pem_pos=0, final_scores=0.14340417,
final_local_scores=0.03676173, pem=0.0, similarity=5.186756]
4: 7246 Charles_Messier
    [pem_pos=0, final_scores=0.09627099,
final_local_scores=-0.016337477, pem=0.0, similarity=3.60217]
5: 7246 Charles_Messier
    [pem_pos=0, final_scores=0.13731362,
final_local_scores=0.03010872, pem=0.0, similarity=4.9882174]
6: 7246 Charles_Messier
    [pem_pos=0, final_scores=0.20458241,
final_local_scores=0.10586682, pem=0.0, similarity=7.24899]
7: 7246 Charles_Messier
    [pem_pos=0, final_scores=0.10293017,
final_local_scores=-0.009063348, pem=0.0,
similarity=3.8192444]
8: 214067 Virgo_Cluster
    [pem_pos=0, final_scores=0.054035798,
final_local_scores=0.0914032, pem=0.0, similarity=6.8173666]
9: 7246 Charles_Messier
    [pem_pos=0, final_scores=0.1942135,
final_local_scores=0.09226319, pem=0.0, similarity=6.8430305]
10: 7246 Charles_Messier
    [pem_pos=0, final_scores=0.1681615,
final_local_scores=0.06380536, pem=0.0, similarity=5.993792]

```

Sample: b'EMD_E8'

The E8 was a , passenger-train locomotive built by ' 1 Elect
ro-Motive Division -LRB- EMD -RRB- of [1 2 La Grange 1] , [2 I
llinois 2] . 450 3 cab versions , or E8As , were built from
August 1949 to January 1954 , 447 for the [3 U.S. 3] and 3
for [4 Canada 4] . 46 E8Bs were built from December 1949 to
January 1954 , all for the [5 U.S. 5] . The 2,250 hp came

from two 4 12 cylinder model 567B engines , each driving a to power the two 5 traction motor s on one truck . The E8 was the ninth model in the line of passenger diesels of similar design known as . In profile the front of the nose of 6 E7 , E8 , and 7 E9 units is less slanted than earlier [6 EMD 6] units , and E7/8/9s -LRB- and their four axle cousins , the F-unit series -RRB- have been nicknamed `` 8 bulldog nose '' units . Earlier E-unit locomotives were nicknamed `` shovel nose '' units or `` slant nose '' units . After passenger trains were canceled on the [7 Erie Lackawanna 7] in 1970 , the E8s were re-gearred for freight and were very reliable for the EL . These units were on freight trains until the early years of [8 Consolidated Railroad Corporation 8] -LRB- '' [1 1 Conrail 1] '' -RRB- . It is estimated that 58 E8s have survived into preservation . The former [9 NYC 9] 4085 , preserved at the , was the lead locomotive on the final eastbound [2 2 20th Century Limited 2] . Another surviving E8 is operated by the [10 9 Midland Railway 10] , in [11 10 Baldwin City 11] , [12 Kansas 12] . Privately owned , this unit is ex - E8A # 652 and is used for special events . [13 NYC 13] 4096 , recently restored , is also on display . [14 New York Central 14] 4097 , privately owned , is on display at Merli Mfg Co , [15 Duanesburg 15] , [16 New York 16] . [3 3 Southern Railway 3] E8A unit # 6900 is operational at the [4 4 North Carolina Transportation Museum 4] in Spencer , [17 North Carolina 17] . It pulled the [18 Southern Crescent 18] . Southern # 6901 is on display in [19 Duluth 19] , [20 GA 20] , at the 11 Southeastern Railway Museum . Another [21 Southern Railway 21] E8 # 6913 is being restored at the 12 Southern Appalachia Railway Museum in [22 Oak Ridge 22] , [23 TN 23] for their Southern excursion train . Yet another , Southern # 6914 , is under restoration at [5 5 Tennessee Valley Railroad Museum 5] . Of the units owned by [24 Conrail 24] , three were saved after their freight-service retirement and went on to be refurbished by the 13 Juniata Locomotive Shops in [25 Altoona 25] , [26 PA 26] for use as [27 Conrail 27] 's Office Car Special -LRB- OCS -RRB- until the merger of 1999 . One unit went to [28 CSX 28] , and two were sold off to Bennett Levin , CEO of the 14 Juniata Terminal Company , where they have been overhauled and painted as twin [29 Pennsylvania Railroad 29] E8 's . Another , the former EL 833 , was purchased by the [30 New York 30] and Greenwood Lake Railroad . The unit was repainted in its original livery as [31 Erie 31] 833 , and was on display for a while on the turntable at [32 Port Jervis 32] , [33 NY 33] . In June 2008 , two authentic New York Central E8 's units -LRB- 4080A & 4068A -RRB- were brought to the [6 6 Medina Railroad Museum 6]

1: 1399844 Conrail

2: 1014037 20th_Century_Limited

3: 1095495 Southern_Railway_(U.S.)

4: 2058124 North_Carolina_Transportation_Museum

5: 1419364 Tennessee_Valley_Railroad_Museum
6: 9350965 Medina_Railroad_Museum
1: 1968562 Electro-Motive_Diesel
[winner: 1968562 Electro-Motive_Diesel, pem_pos=0,
final_scores=-0.21401985, final_local_scores=-0.21779819,
pem=0.0, similarity=-2.4098194]
gt: 1968562 Electro-Motive_Diesel, pem_pos=0,
final_scores=-0.21401985, final_local_scores=-0.21779819,
pem=0.0, similarity=-2.4098194]
2: 111002 La_Grange,_Illinois
[winner: 14849 Illinois, pem_pos=0,
final_scores=0.5228346, final_local_scores=0.4953397,
pem=-0.16995399, similarity=19.795874]
RECALL MISS: 111002 La_Grange,_Illinois]
[winner: 115390 La_Grange,_Kentucky, pem_pos=3,
final_scores=0.14574866, final_local_scores=0.15965766,
pem=-1.7714915, similarity=18.487862]
gt: 111002 La_Grange,_Illinois, pem_pos=1,
final_scores=0.13995932, final_local_scores=0.13242558,
pem=-1.3922368, similarity=15.612756]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.9343617, final_local_scores=-1.9198363,
pem=-4.6728287, similarity=-27.79046]
RECALL MISS: 111002 La_Grange,_Illinois]
3: 1323819 A_unit
[winner: 1323819 A_unit, pem_pos=2,
final_scores=-1.1439465, final_local_scores=-1.1534407,
pem=-1.2816924, similarity=-23.361216]
gt: 1323819 A_unit, pem_pos=2, final_scores=-1.1439465,
final_local_scores=-1.1534407, pem=-1.2816924,
similarity=-23.361216]
4: 32660 V12_engine
[winner: 200289 Flat-twelve_engine, pem_pos=1,
final_scores=-0.8797918, final_local_scores=-0.80492294,
pem=0.0, similarity=-19.930794]
gt: 32660 V12_engine, pem_pos=0,
final_scores=-0.9161113, final_local_scores=-0.8975575,
pem=0.0, similarity=-22.695192]
5: 1081613 Traction_motor
[winner: 1081613 Traction_motor, pem_pos=0,
final_scores=-0.92728484, final_local_scores=-0.9489554,
pem=0.0, similarity=-24.22901]
gt: 1081613 Traction_motor, pem_pos=0,
final_scores=-0.92728484, final_local_scores=-0.9489554,
pem=0.0, similarity=-24.22901]
6: 1099476 EMD_E7
[winner: 649115 E7_(mathematics), pem_pos=0,
final_scores=-0.44018024, final_local_scores=-0.2944765,
pem=-0.8323443, similarity=-0.17163606]
gt: 1099476 EMD_E7, pem_pos=2, final_scores=-0.47574455,
final_local_scores=-0.46466187, pem=-2.0406556,
similarity=1.3206785]

7: 1280214 EMD_E9
 [winner: 1280214 EMD_E9, pem_pos=0,
 final_scores=-0.40760565, final_local_scores=-0.36757416,
 pem=-1.0015384, similarity=-1.432911|
 gt: 1280214 EMD_E9, pem_pos=0, final_scores=-0.40760565,
 final_local_scores=-0.36757416, pem=-1.0015384,
 similarity=-1.432911]

8: 1882674 Bulldog_nose
 [winner: 1882674 Bulldog_nose, pem_pos=0,
 final_scores=-0.9927611, final_local_scores=-0.9569045,
 pem=0.0, similarity=-24.466225|
 gt: 1882674 Bulldog_nose, pem_pos=0,
 final_scores=-0.9927611, final_local_scores=-0.9569045,
 pem=0.0, similarity=-24.466225]

9: 1142060 Midland_Railway_(Kansas)
 [winner: 188429 Midland_Railway, pem_pos=0,
 final_scores=0.2375928, final_local_scores=0.25427145,
 pem=0.0, similarity=11.67768|
 gt: 1142060 Midland_Railway_(Kansas), pem_pos=1,
 final_scores=-0.2677847, final_local_scores=-0.2279594,
 pem=-2.330628, similarity=9.961267]

10: 114534 Baldwin_City,_Kansas
 [winner: 16716 Kansas, pem_pos=0,
 final_scores=0.37149566, final_local_scores=0.35288268,
 pem=-0.3846871, similarity=16.712423|
 RECALL MISS: 114534 Baldwin_City,_Kansas]
 [winner: 114534 Baldwin_City,_Kansas, pem_pos=0,
 final_scores=0.32012284, final_local_scores=0.35030374,
 pem=0.0, similarity=14.543475|
 gt: 114534 Baldwin_City,_Kansas, pem_pos=0,
 final_scores=0.32012284, final_local_scores=0.35030374,
 pem=0.0, similarity=14.543475]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7669716, final_local_scores=-1.7369881,
 pem=-4.6728287, similarity=-22.333902|
 RECALL MISS: 114534 Baldwin_City,_Kansas]

11: 22613940 Southeastern_Railway_Museum
 [winner: 22613940 Southeastern_Railway_Museum, pem_pos=0,
 final_scores=-0.05337827, final_local_scores=0.0244136,
 pem=0.0, similarity=4.8182635|
 gt: 22613940 Southeastern_Railway_Museum, pem_pos=0,
 final_scores=-0.05337827, final_local_scores=0.0244136,
 pem=0.0, similarity=4.8182635]

12: 25074579 Southern_Appalachia_Railway_Museum
 [winner: 25074579 Southern_Appalachia_Railway_Museum,
 pem_pos=0, final_scores=-0.042123765,
 final_local_scores=0.031075776, pem=0.0, similarity=5.017076|
 gt: 25074579 Southern_Appalachia_Railway_Museum,
 pem_pos=0, final_scores=-0.042123765,
 final_local_scores=0.031075776, pem=0.0, similarity=5.017076]

13: 19202340 Altoona_Works
 [winner: 19202340 Altoona_Works, pem_pos=0,

```

final_scores=0.01890479, final_local_scores=0.036486775,
pem=0.0, similarity=5.178551]
    gt: 19202340 Altoona_Works, pem_pos=0,
final_scores=0.01890479, final_local_scores=0.036486775,
pem=0.0, similarity=5.178551 ]
14: 8124002 Juniata_Terminal_Company
    [winner: 8124002 Juniata_Terminal_Company, pem_pos=0,
final_scores=-0.030408606, final_local_scores=0.05823566,
pem=0.0, similarity=5.8275814]
    gt: 8124002 Juniata_Terminal_Company, pem_pos=0,
final_scores=-0.030408606, final_local_scores=0.05823566,
pem=0.0, similarity=5.8275814 ]
1: 1399844 Conrail
    [pem_pos=0, final_scores=0.15039472,
final_local_scores=0.122465014, pem=0.0, similarity=7.7443132]
2: 1014037 20th_Century_Limited
    [pem_pos=0, final_scores=0.055266365,
final_local_scores=0.056168735, pem=0.0, similarity=5.7659]
3: 1095495 Southern_Railway_(U.S.)
    [pem_pos=1, final_scores=0.056900933,
final_local_scores=-0.0018505156, pem=-0.71496606,
similarity=7.922587]
4: 2058124 North_Carolina_Transportation_Museum
    [pem_pos=0, final_scores=0.22508888,
final_local_scores=0.27436116, pem=0.0, similarity=12.277197]
5: 1419364 Tennessee_Valley_Railroad_Museum
    [pem_pos=0, final_scores=0.18070604,
final_local_scores=0.20923164, pem=0.0, similarity=10.333603]
6: 9350965 Medina_Railroad_Museum
    [pem_pos=0, final_scores=0.046210185,
final_local_scores=0.12820879, pem=0.0, similarity=7.915719]
1: 115390 La_Grange,_Kentucky
    [pem_pos=3, final_scores=0.14574866,
final_local_scores=0.15965766, pem=-1.7714915,
similarity=18.487862]
2: 14849 Illinois
    [pem_pos=0, final_scores=0.5228346,
final_local_scores=0.4953397, pem=-0.16995399,
similarity=19.795874]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.48244375,
final_local_scores=0.48427504, pem=-0.14075686,
similarity=19.306904]
4: 5042916 Canada
    [pem_pos=0, final_scores=0.63158244,
final_local_scores=0.68007845, pem=-0.11662793,
similarity=25.01885]
5: 3434750 United_States
    [pem_pos=0, final_scores=0.49893576,
final_local_scores=0.50229007, pem=-0.14075686,
similarity=19.844507]
6: 1968562 Electro-Motive_Diesel

```

[pem_pos=0, final_scores=0.14113359,
final_local_scores=0.17749983, pem=0.0, similarity=9.386662]
7: 814023 Erie_Lackawanna_Railway
[pem_pos=0, final_scores=0.20431848,
final_local_scores=0.20770693, pem=0.0, similarity=10.288102]
8: 1399844 Conrail
[pem_pos=0, final_scores=0.09266774,
final_local_scores=0.05940707, pem=0.0, similarity=5.8625383]
9: 645042 New_York_City
[pem_pos=0, final_scores=0.18088253,
final_local_scores=0.13932917, pem=-0.38631868,
similarity=10.348434]
10: 188429 Midland_Railway
[pem_pos=0, final_scores=0.2375928,
final_local_scores=0.25427145, pem=0.0, similarity=11.67768]
11: 114534 Baldwin_City,_Kansas
[pem_pos=0, final_scores=0.32012284,
final_local_scores=0.35030374, pem=0.0, similarity=14.543475]
12: 16716 Kansas
[pem_pos=0, final_scores=0.37149566,
final_local_scores=0.35288268, pem=-0.3846871,
similarity=16.712423]
13: 645042 New_York_City
[pem_pos=0, final_scores=0.22235923,
final_local_scores=0.18463618, pem=-0.38631868,
similarity=11.700486]
14: 345656 New_York_Central_Railroad
[pem_pos=0, final_scores=0.1196564,
final_local_scores=0.06768593, pem=0.0, similarity=6.1095963]
15: 260023 Duaneburg,_New_York
[pem_pos=0, final_scores=0.34069997,
final_local_scores=0.36386102, pem=0.0, similarity=14.948052]
16: 8210131 New_York
[pem_pos=0, final_scores=0.18464683,
final_local_scores=0.14152303, pem=-0.2889045,
similarity=9.88415]
17: 21650 North_Carolina
[pem_pos=0, final_scores=0.49079603,
final_local_scores=0.48279226, pem=-0.21680297,
similarity=19.676205]
18: 573891 Crescent_(train)
[pem_pos=0, final_scores=0.08579357,
final_local_scores=0.09160298, pem=0.0, similarity=6.8233285]
19: 63503 Duluth,_Minnesota
[pem_pos=0, final_scores=0.42146343,
final_local_scores=0.4204719, pem=-0.2524106,
similarity=18.010082]
20: 48830 Georgia_(U.S._state)
[pem_pos=0, final_scores=0.16080527,
final_local_scores=0.12099588, pem=-1.3101573,
similarity=14.82531]
21: 1095495 Southern_Railway_(U.S.)

[pem_pos=1, final_scores=0.06872921,
 final_local_scores=0.014770418, pem=-0.71496606,
 similarity=8.418589]
 22: 151055 Oak_Ridge,_Tennessee
 [pem_pos=0, final_scores=0.15886171,
 final_local_scores=0.16618863, pem=-0.72964156,
 similarity=13.017017]
 23: 30395 Tennessee
 [pem_pos=0, final_scores=0.25762147,
 final_local_scores=0.22589797, pem=-0.81725425,
 similarity=15.275314]
 24: 1399844 Conrail
 [pem_pos=0, final_scores=0.19299124,
 final_local_scores=0.16899526, pem=0.0, similarity=9.132869]
 25: 131425 Altoona,_Pennsylvania
 [pem_pos=0, final_scores=0.19252522,
 final_local_scores=0.15283316, pem=-0.7864993,
 similarity=12.927665]
 26: 23332 Pennsylvania
 [pem_pos=0, final_scores=0.33364588,
 final_local_scores=0.28309304, pem=-0.9631348,
 similarity=17.77545]
 27: 1399844 Conrail
 [pem_pos=0, final_scores=0.1327794,
 final_local_scores=0.10322295, pem=0.0, similarity=7.1700916]
 28: 287679 CSX_Transportation
 [pem_pos=0, final_scores=0.15803991,
 final_local_scores=0.122608215, pem=0.0, similarity=7.7485867]
 29: 320965 Pennsylvania_Railroad
 [pem_pos=0, final_scores=0.27420795,
 final_local_scores=0.2207815, pem=0.0, similarity=10.678273]
 30: 645042 New_York_City
 [pem_pos=1, final_scores=0.17577837,
 final_local_scores=0.13375366, pem=-1.7906687,
 similarity=17.819124]
 31: 132482 Erie,_Pennsylvania
 [pem_pos=0, final_scores=0.050185695,
 final_local_scores=-0.016443431, pem=-1.0382025,
 similarity=9.244915]
 32: 126952 Port_Jervis,_New_York
 [pem_pos=0, final_scores=0.38211256,
 final_local_scores=0.37875456, pem=0.0, similarity=15.392505]
 33: 8210131 New_York
 [pem_pos=1, final_scores=0.21135207,
 final_local_scores=0.1673769, pem=-0.3855658,
 similarity=11.18134]

Sample: b'X_Corps'

_ -LRB- [1 Union_Army 1] -RRB- X Corps was a 1 corps of the [1 1 Union_Army 1] during the [2 2 American Civil War 2] . It served during operations in [3 3 South Carolina 3] in the [2 Department of the South 2] , and later in 's [4 4 Army of the

James 4] , during the [3 2 Bermuda Hundred 3] and Campaigns . The corps was officially created on September 13 , 1862 , to consist of the majority of [4 Union 4] troops operating in [5 South Carolina 5] and eastern [6 Georgia 6] . -LRB- Other troops in Florida were officially under its command but were not effectively . -RRB- The corps was initially commanded by [5 5 Ormsby M. Mitchel 5] , who died in October 1862 . He was succeeded by [6 6 John Milton Brannan 6] , [7 7 David Hunter 7] , [8 8 Quincy Adams Gillmore 8] , [9 9 David B. Birney 9] and . The corps took part in most of the operations against [10 10 Charleston 10] from 1862 -- 63 , including attacks on [11 11 James Island 11] and [12 12 Morris Island 12] and the . Other elements of the corps based in Florida took part in the disastrous 3 Battle of [7 Olustee 7] . In early 1864 , the corps , now commanded by [8 Gillmore 8] , was transferred to the [13 13 Army of the James 13] . It took part in the [9 Bermuda Hundred 9] operations and played a principal role in the disastrous 4 Drewry 's Bluff action . It also took part in the attack on [14 14 Cold Harbor 14] in conjunction with units of the [15 15 Army of the Potomac 15] , and the corps played a major role in the early stages of the . In December 1864 , the corps was disbanded ; its white contingent went to the new , while its black units went to the . A detachment of former X Corps troops took part in the successful attack on [10 5 Fort Fisher 10] , [11 North Carolina 11] , in early 1865 . The X Corps was `` revived '' , under the command of 6 Alfred Terry , in March 1865 . It was attached to the

- 1: 360126 Union_Army
- 2: 863 American_Civil_War
- 3: 27956 South_Carolina
- 4: 1464983 Army_of_the_James
- 5: 936503 Ormsby_M._Mitchel
- 6: 12695429 John_Milton_Brannan
- 7: 224424 David_Hunter
- 8: 8908909 Quincy_Adams_Gillmore
- 9: 1616983 David_B._Birney
- 10: 61024 Charleston,_South_Carolina
- 11: 2384324 James_Island_(South_Carolina)
- 12: 327720 Morris_Island
- 13: 1464983 Army_of_the_James
- 14: 48965 Battle_of_Cold_Harbor
- 15: 222590 Army_of_the_Potomac

1: 166674 Corps

[winner: 38188326 Belgian_Forces_in_Germany, pem_pos=0, final_scores=-1.0129085, final_local_scores=-1.0065229, pem=0.0, similarity=-25.946938]

RECALL MISS: 166674 Corps]

2: 1661378 Bermuda_Hundred_Campaign

[winner: 8914115 Bermuda_Hundred,_Virginia, pem_pos=1, final_scores=0.21244518, final_local_scores=0.18741572, pem=-0.06752448, similarity=10.04978]

gt: 1661378 Bermuda_Hundred_Campaign, pem_pos=0,

```

final_scores=0.199578, final_local_scores=0.17377314,
pem=0.0, similarity=9.27545 ]
3: 192998 Battle_of_Olustee
  [winner: 192998 Battle_of_Olustee, pem_pos=0,
final_scores=0.15293242, final_local_scores=0.13450533,
pem=0.0, similarity=8.103621]
  gt: 192998 Battle_of_Olustee, pem_pos=0,
final_scores=0.15293242, final_local_scores=0.13450533,
pem=0.0, similarity=8.103621 ]
  [winner: 16649099 Mount_Battle, pem_pos=5,
final_scores=-0.9587433, final_local_scores=-0.88455755,
pem=-3.124515, similarity=-5.31565]
  RECALL MISS: 192998 Battle_of_Olustee]
4: 3974795 Battle_of_Proctor's_Creek
  [winner: 492150 Bluff, New_Zealand, pem_pos=0,
final_scores=-0.07480347, final_local_scores=0.011975348,
pem=-0.7790896, similarity=8.683891]
  RECALL MISS: 3974795 Battle_of_Proctor's_Creek]
  [winner: 1225147 CSS_Drewry, pem_pos=0,
final_scores=-0.12062718, final_local_scores=-0.030396692,
pem=-0.23280548, similarity=4.4486475]
  RECALL MISS: 3974795 Battle_of_Proctor's_Creek]
  [winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.4919636, final_local_scores=-2.5727565,
pem=-3.4067376, similarity=-54.160103]
  RECALL MISS: 3974795 Battle_of_Proctor's_Creek]
5: 140460 Second_Battle_of_Fort_Fisher
  [winner: 1536460 Fort_Fisher, pem_pos=0,
final_scores=0.22265999, final_local_scores=0.19267848,
pem=0.0, similarity=9.839623]
  gt: 140460 Second_Battle_of_Fort_Fisher, pem_pos=3,
final_scores=-0.031229928, final_local_scores=-0.075883105,
pem=-1.2872831, similarity=8.825655 ]
6: 396831 Alfred_Terry
  [winner: 396831 Alfred_Terry, pem_pos=0,
final_scores=0.0092957765, final_local_scores=-0.032890588,
pem=0.0, similarity=3.1081922]
  gt: 396831 Alfred_Terry, pem_pos=0,
final_scores=0.0092957765, final_local_scores=-0.032890588,
pem=0.0, similarity=3.1081922 ]
1: 360126 Union_Army
  [pem_pos=0, final_scores=0.37453288,
final_local_scores=0.32560113, pem=-0.0130147785,
similarity=13.877076]
2: 863 American_Civil_War
  [pem_pos=0, final_scores=0.50245535,
final_local_scores=0.4713934, pem=0.0, similarity=18.157032]
3: 27956 South_Carolina
  [pem_pos=0, final_scores=0.3769377,
final_local_scores=0.3663866, pem=-0.27085266,
similarity=16.496357]
4: 1464983 Army_of_the_James

```

[pem_pos=0, final_scores=0.06522353,
final_local_scores=0.031411186, pem=0.0, similarity=5.0270853]
5: 936503 Ormsby_M_Mitchel
[pem_pos=0, final_scores=0.056339666,
final_local_scores=0.05293329, pem=0.0, similarity=5.669348]
6: 12695429 John_Milton_Brannan
[pem_pos=0, final_scores=0.049109414,
final_local_scores=0.030143991, pem=0.0, similarity=4.9892697]
7: 224424 David_Hunter
[pem_pos=0, final_scores=0.16192035,
final_local_scores=0.13549331, pem=0.0, similarity=8.133103]
8: 8908909 Quincy_Adams_Gillmore
[pem_pos=0, final_scores=0.14277907,
final_local_scores=0.104553, pem=0.0, similarity=7.209783]
9: 1616983 David_B_Birney
[pem_pos=0, final_scores=0.16455571,
final_local_scores=0.15767351, pem=0.0, similarity=8.795005]
10: 61024 Charleston_South_Carolina
[pem_pos=0, final_scores=0.29824507,
final_local_scores=0.2588596, pem=-0.98667276,
similarity=17.180279]
11: 2384324 James_Island_(South_Carolina)
[pem_pos=0, final_scores=0.06762357,
final_local_scores=0.07381034, pem=-0.7227017,
similarity=10.222525]
12: 327720 Morris_Island
[pem_pos=0, final_scores=0.21710919,
final_local_scores=0.2168298, pem=0.0, similarity=10.560347]
13: 1464983 Army_of_the_James
[pem_pos=0, final_scores=0.058284238,
final_local_scores=0.023831084, pem=0.0, similarity=4.80088]
14: 48965 Battle_of_Cold_Harbor
[pem_pos=0, final_scores=0.3232153,
final_local_scores=0.30975962, pem=0.0, similarity=13.333558]
15: 222590 Army_of_the_Potomac
[pem_pos=0, final_scores=0.11363973,
final_local_scores=0.052314356, pem=0.0, similarity=5.6508775]
1: 360126 Union_Army
[pem_pos=0, final_scores=0.08567889,
final_local_scores=0.0071119517, pem=0.0,
similarity=4.3019476]
2: 1620299 X_Corps_(Union_Army)
[pem_pos=1, final_scores=0.048240945,
final_local_scores=0.024282813, pem=0.0, similarity=4.8143606]
3: 8914115 Bermuda_Hundred_Virginia
[pem_pos=1, final_scores=0.21244518,
final_local_scores=0.18741572, pem=-0.06752448,
similarity=10.04978]
4: 481130 Union_(American_Civil_War)
[pem_pos=0, final_scores=0.032351986,
final_local_scores=-0.046296895, pem=-1.132797,
similarity=8.868447]

5: 27956 South_Carolina
 [pem_pos=0, final_scores=0.32180482,
 final_local_scores=0.30616227, pem=-0.27085266,
 similarity=14.699144]
 6: 48830 Georgia_(U.S._state)
 [pem_pos=0, final_scores=0.27102035,
 final_local_scores=0.2617506, pem=-0.70047224,
 similarity=15.71015]
 7: 192998 Battle_of_Olustee
 [pem_pos=0, final_scores=0.15293242,
 final_local_scores=0.13450533, pem=0.0, similarity=8.103621]
 8: 8908909 Quincy_Adams_Gillmore
 [pem_pos=0, final_scores=0.25288302,
 final_local_scores=0.23015288, pem=0.0, similarity=10.957933]
 9: 8914115 Bermuda_Hundred,_Virginia
 [pem_pos=1, final_scores=0.1771131,
 final_local_scores=0.14882076, pem=-0.06752448,
 similarity=8.898029]
 10: 1536460 Fort_Fisher
 [pem_pos=0, final_scores=0.22265999,
 final_local_scores=0.19267848, pem=0.0, similarity=9.839623]
 11: 21650 North_Carolina
 [pem_pos=0, final_scores=0.39520139,
 final_local_scores=0.393193, pem=-0.21680297,
 similarity=17.002384]

Sample: b'Nembrionic'

Nembrionic was a [1 1 Dutch 1] 1 death metal band . They were formed in 1991 under the name Nembrionic Hammerdeath as a 2 black metal band ; they changed their name to Nembrionic in 1993 . The band released three full-lengths on 3 Displeased Records and one full-length together with [2 2 Osdorp Posse 2] . They split up in 1999 . The band started in April 1991 , influenced by 4 Venom , 5 Possessed , and [1 6 Terrorizer 1] , and recorded a demo in that same year . Their 1992 EP sold 1800 copies and garnered the band a record deal with [2 Dutch 2] metal label 7 Displeased Records . They moved toward 8 grindcore and 9 death metal on 10 Tempter , a split with 11 Consolation . The band played about 30 shows were played to promote the album and went on tour with 12 At the Gates and 13 Consolation . Hammerdeath was dropped from the name in 1995 , and their next album , 14 Psycho One Hundred , led to shows at open air festivals such as Liberation Day , Mundial , and Lowlands , and even a show in [3 3 Ahoy Rotterdam 3] . In 1996 , they joined with [3 Osdorp Posse 3] to record 15 Briljant , Hard en Geslepen , which charted in the [4 Netherlands 4] . Nembrionic played at [4 4 Dynamo Open Air 4] and [5 5 Pinkpop Festival 5] , and even opened for [6 6 Slayer 6] in [7 7 Paradiso 7] , June 1996 . Their second to last release was 16 Bloodcult

1: 21148 Netherlands
 2: 2377223 Osdorp_Posse

3: 4256663 Ahoy_Rotterdam
4: 1068850 Dynamo_Open_Air
5: 292403 Pinkpop_Festival
6: 73575 Slayer
7: 4448294 Paradiso_(Amsterdam)
1: 8233 Death_metal
[winner: 209048 Swedish_death_metal, pem_pos=0,
final_scores=-0.89746815, final_local_scores=-0.9402254,
pem=0.0, similarity=-23.968489|
gt: 8233 Death_metal, pem_pos=1,
final_scores=-1.0799602, final_local_scores=-1.1841087,
pem=-0.006190767, similarity=-31.212788]
2: 4874 Black_metal
[winner: 4874 Black_metal, pem_pos=0,
final_scores=-0.7243852, final_local_scores=-0.78801936,
pem=-0.009154551, similarity=-19.376572|
gt: 4874 Black_metal, pem_pos=0,
final_scores=-0.7243852, final_local_scores=-0.78801936,
pem=-0.009154551, similarity=-19.376572]
3: 4465306 Displeased_Records
[winner: 4465306 Displeased_Records, pem_pos=0,
final_scores=-0.041750297, final_local_scores=-0.011988029,
pem=0.0, similarity=3.7319663|
gt: 4465306 Displeased_Records, pem_pos=0,
final_scores=-0.041750297, final_local_scores=-0.011988029,
pem=0.0, similarity=3.7319663]
4: 299508 Venom_(band)
[winner: 299508 Venom_(band), pem_pos=1,
final_scores=-0.11270149, final_local_scores=-0.1282575,
pem=-1.2398694, similarity=7.0048556|
gt: 299508 Venom_(band), pem_pos=1,
final_scores=-0.11270149, final_local_scores=-0.1282575,
pem=-1.2398694, similarity=7.0048556]
5: 400134 Possessed_(band)
[winner: 400134 Possessed_(band), pem_pos=0,
final_scores=-0.23905908, final_local_scores=-0.23189592,
pem=-0.51352954, similarity=-0.03787068|
gt: 400134 Possessed_(band), pem_pos=0,
final_scores=-0.23905908, final_local_scores=-0.23189592,
pem=-0.51352954, similarity=-0.03787068]
6: 875875 Terrorizer
[winner: 6976638 Terrorizer_(magazine), pem_pos=0,
final_scores=0.033838823, final_local_scores=0.057755485,
pem=0.0, similarity=5.813252|
gt: 875875 Terrorizer, pem_pos=1,
final_scores=-0.084891155, final_local_scores=-0.01573772,
pem=-0.07102191, similarity=4.006296]
7: 4465306 Displeased_Records
[winner: 4465306 Displeased_Records, pem_pos=0,
final_scores=-0.05209294, final_local_scores=-0.023285799,
pem=0.0, similarity=3.3948183|
gt: 4465306 Displeased_Records, pem_pos=0,

final_scores=-0.05209294, final_local_scores=-0.023285799,
pem=0.0, similarity=3.3948183]
8: 13137 Grindcore
[winner: 13137 Grindcore, pem_pos=0,
final_scores=-1.1993546, final_local_scores=-1.3046594,
pem=-0.0076511395, similarity=-34.802315]
gt: 13137 Grindcore, pem_pos=0, final_scores=-1.1993546,
final_local_scores=-1.3046594, pem=-0.0076511395,
similarity=-34.802315]
9: 8233 Death_metal
[winner: 209048 Swedish_death_metal, pem_pos=0,
final_scores=-0.846131, final_local_scores=-0.8841474,
pem=0.0, similarity=-22.29501]
gt: 8233 Death_metal, pem_pos=1,
final_scores=-1.0218037, final_local_scores=-1.1205816,
pem=-0.006190767, similarity=-29.317015]
10: 4571704 Tempter_(album)
[winner: 7040454 Temptation, pem_pos=1,
final_scores=-0.22301713, final_local_scores=-0.13072936,
pem=-0.25423416, similarity=1.571053]
gt: 4571704 Tempter_(album), pem_pos=0,
final_scores=-0.35855392, final_local_scores=-0.2401208,
pem=0.0, similarity=-3.0759704]
11: 41245169 Consolation
[winner: 41245169 Consolation, pem_pos=1,
final_scores=-0.14884216, final_local_scores=-0.030344963,
pem=-0.1264139, similarity=3.8716173]
gt: 41245169 Consolation, pem_pos=1,
final_scores=-0.14884216, final_local_scores=-0.030344963,
pem=-0.1264139, similarity=3.8716173]
12: 586983 At_the_Gates
[winner: 586983 At_the_Gates, pem_pos=0,
final_scores=-0.2765187, final_local_scores=-0.27946663,
pem=0.0, similarity=-4.250129]
gt: 586983 At_the_Gates, pem_pos=0,
final_scores=-0.2765187, final_local_scores=-0.27946663,
pem=0.0, similarity=-4.250129]
13: 41245169 Consolation
[winner: 7963793 USS_Consolation_(AH-15), pem_pos=0,
final_scores=-0.17493212, final_local_scores=-0.035255253,
pem=0.0, similarity=3.0376258]
gt: 41245169 Consolation, pem_pos=1,
final_scores=-0.22408149, final_local_scores=-0.1125325,
pem=-0.1264139, similarity=1.4189773]
14: 4571726 Psycho_One_Hundred
[winner: 4571726 Psycho_One_Hundred, pem_pos=0,
final_scores=-0.2697696, final_local_scores=-0.1785605,
pem=0.0, similarity=-1.2388885]
gt: 4571726 Psycho_One_Hundred, pem_pos=0,
final_scores=-0.2697696, final_local_scores=-0.1785605,
pem=0.0, similarity=-1.2388885]
15: 9509421 Briljant,_Hard_en_Geslepen

[winner: 9509421 Briljant, Hard_en_Geslepen, pem_pos=0,
final_scores=-0.13718228, final_local_scores=-0.06460162,
pem=0.0, similarity=2.1618717]
gt: 9509421 Briljant, Hard_en_Geslepen, pem_pos=0,
final_scores=-0.13718228, final_local_scores=-0.06460162,
pem=0.0, similarity=2.1618717]
[winner: 24951685 Hard_(song), pem_pos=0,
final_scores=-0.6306029, final_local_scores=-0.6183461,
pem=-1.1078234, similarity=-8.338451]
RECALL MISS: 9509421 Briljant, Hard_en_Geslepen]
[winner: 17464004 Parietal_operculum, pem_pos=19,
final_scores=-1.4932288, final_local_scores=-1.4274259,
pem=-4.352231, similarity=-14.839412]
RECALL MISS: 9509421 Briljant, Hard_en_Geslepen]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.989817, final_local_scores=-1.9719923,
pem=-4.6728287, similarity=-29.3469]
RECALL MISS: 9509421 Briljant, Hard_en_Geslepen]
16: 4571798 Bloodcult
[winner: 4571798 Bloodcult, pem_pos=0,
final_scores=-0.4279495, final_local_scores=-0.3019321,
pem=0.0, similarity=-4.9205427]
gt: 4571798 Bloodcult, pem_pos=0,
final_scores=-0.4279495, final_local_scores=-0.3019321,
pem=0.0, similarity=-4.9205427]
1: 21148 Netherlands
[pem_pos=0, final_scores=0.36503738,
final_local_scores=0.35742158, pem=-0.7943452,
similarity=19.075659]
2: 2377223 Osdorp_Posse
[pem_pos=0, final_scores=0.17331676,
final_local_scores=0.20385316, pem=0.0, similarity=10.173098]
3: 4256663 Ahoy_Rotterdam
[pem_pos=0, final_scores=0.25706416,
final_local_scores=0.30527824, pem=0.0, similarity=13.199824]
4: 1068850 Dynamo_Open_Air
[pem_pos=0, final_scores=0.1176215,
final_local_scores=0.14275283, pem=0.0, similarity=8.349742]
5: 292403 Pinkpop_Festival
[pem_pos=0, final_scores=0.07912372,
final_local_scores=0.07468949, pem=0.0, similarity=6.3185964]
6: 73575 Slayer
[pem_pos=0, final_scores=0.27949232,
final_local_scores=0.311024, pem=-0.06576258,
similarity=13.728917]
7: 4448294 Paradiso_(Amsterdam)
[pem_pos=0, final_scores=0.12202646,
final_local_scores=0.10601309, pem=-0.39445245,
similarity=9.398449]
1: 6976638 Terrorizer_(magazine)
[pem_pos=0, final_scores=0.033838823,
final_local_scores=0.057755485, pem=0.0, similarity=5.813252]

2: 21148 Netherlands
 [pem_pos=0, final_scores=0.5256788,
 final_local_scores=0.5328979, pem=-0.7943452,
 similarity=24.312222]
 3: 2377223 Osdorp_Posse
 [pem_pos=0, final_scores=0.27158254,
 final_local_scores=0.3111936, pem=0.0, similarity=13.37635]
 4: 21148 Netherlands
 [pem_pos=0, final_scores=0.14591022,
 final_local_scores=0.11805844, pem=-0.24940795,
 similarity=8.969131]

Sample: b'Giovanni_Battista_Landolina'

[1 Giovanni Battista Landolina 1] , '' di S. Alfano '' , was a [1 1 Sicilian 1] landowner and 1 intellectual instrumental in having the city of [2 2 Noto 2] removed from its former site on 2 Mount Alveria to a more level location following the 3 earthquake in 1693 centred on the 4 Val di [2 Noto 2] . He is commemorated by a piazza in the city named in his honour . Ultimate responsibility for the re-building of [3 Sicily 3] at this time was however in the hands of the [3 3 Spanish 3] 5 viceroy , 6 Duke of Camastra ; but Landolina successfully argued his case for the rebuilding to be some 10 km . from the former site . He is credited with having designed the city with the help of three local 7 architect s . It is based on three parallel streets , interconnected by a series of narrower streets at 8 right angle s , thus producing a 9 grid pattern , providing s to the three 10 piazza s , each with their own church , the largest having a 11 cathedral . Landolina 's town planning 12 philosophy was based on the [4 4 Baroque 4] system , in which the town was divided according to social rank and position , the 13 aristocracy were given the highest sites , the 14 church the town centre , to reflect their position at the centre of one 's life , and the poor the of the town where no-one else wanted to live . Later the architects [5 5 Giovanni Battista Vaccarini 5] , and 15 Rosario Gagliardi designed many buildings in the city . Landolina rule was broken in 1730 when his son Francesco Landolina , Marchese di S. Alfano , built his new house the `` Palazzo Landolina '' , in the town centre next to the new cathedral . Today , [4 Noto 4] is a great attraction owing to the many examples of [6 6 Sicilian Baroque 6]

1: 27619 Sicily
 2: 157906 Noto
 3: 26667 Spain
 4: 3957 Baroque
 5: 2397816 Giovanni_Battista_Vaccarini
 6: 2372757 Sicilian_Baroque
 1: 227107 Intellectual

[winner: 832518 Intellectualism, pem_pos=1,
 final_scores=-0.9869108, final_local_scores=-0.97345805,

pem=-0.64358115, similarity=-21.460323|
 gt: 227107 Intellectual, pem_pos=0,
 final_scores=-1.0449432, final_local_scores=-1.0732177,
 pem=-0.24564637, similarity=-26.601383]
 2: 2908410 Mount_Alveria
 [winner: 2908410 Mount_Alveria, pem_pos=0,
 final_scores=-0.06341179, final_local_scores=-0.0075971633,
 pem=0.0, similarity=3.8629985|
 gt: 2908410 Mount_Alveria, pem_pos=0,
 final_scores=-0.06341179, final_local_scores=-0.0075971633,
 pem=0.0, similarity=3.8629985]
 3: 14807319 1693_Sicily_earthquake
 [winner: 10630050 1693_in_art, pem_pos=3,
 final_scores=-0.686163, final_local_scores=-0.5185492,
 pem=-1.5129857, similarity=-3.1569777|
 gt: 14807319 1693_Sicily_earthquake, pem_pos=2,
 final_scores=-1.0494998, final_local_scores=-1.1673305,
 pem=-1.4988189, similarity=-22.594944]
 [winner: 10106 Earthquake, pem_pos=0,
 final_scores=-1.2451073, final_local_scores=-1.3170753,
 pem=-0.3929322, similarity=-33.07761|
 RECALL MISS: 14807319 1693_Sicily_earthquake]
 4: 2896611 Val_di_Noto
 [winner: 157906 Noto, pem_pos=0, final_scores=0.19976626,
 final_local_scores=0.1446136, pem=-0.29883283,
 similarity=10.030371|
 RECALL MISS: 2896611 Val_di_Noto]
 [winner: 176556 Aichi_D3A, pem_pos=0,
 final_scores=-0.8274495, final_local_scores=-0.740776,
 pem=-1.25673, similarity=-11.182227|
 RECALL MISS: 2896611 Val_di_Noto]
 [winner: 295459 Haleakalā, pem_pos=2,
 final_scores=-1.6039566, final_local_scores=-1.6047902,
 pem=-2.2553322, similarity=-31.53558|
 RECALL MISS: 2896611 Val_di_Noto]
 5: 180708 Viceroy
 [winner: 9587063 Icon_Brickell, pem_pos=9,
 final_scores=-1.0929939, final_local_scores=-1.0262809,
 pem=-2.1972246, similarity=-14.587708|
 gt: 180708 Viceroy, pem_pos=0, final_scores=-1.1223044,
 final_local_scores=-1.1961124, pem=-0.44570243,
 similarity=-29.180866]
 6: 2447063 Giuseppe_Lanza
 [winner: 2447063 Giuseppe_Lanza, pem_pos=0,
 final_scores=-0.093053, final_local_scores=-0.04960236,
 pem=0.0, similarity=2.6094797|
 gt: 2447063 Giuseppe_Lanza, pem_pos=0,
 final_scores=-0.093053, final_local_scores=-0.04960236,
 pem=0.0, similarity=2.6094797]
 7: 1170 Architect
 [winner: 19131789 Thomas_Farnolls_Pritchard, pem_pos=0,
 final_scores=-0.84090406, final_local_scores=-0.8014031,

pem=0.0, similarity=-19.825754|
 RECALL MISS: 1170 Architect]

8: 76956 Right_angle
 [winner: 20753867 Thread_angle, pem_pos=3,
 final_scores=-0.9138385, final_local_scores=-0.7822219,
 pem=-1.0975384, similarity=-13.284765|
 RECALL MISS: 76956 Right_angle]
 [winner: 5595237 Right_ventricular_hypertrophy,
 pem_pos=5, final_scores=-1.2515997,
 final_local_scores=-1.1693177, pem=-2.3725495,
 similarity=-17.902765|
 gt: 76956 Right_angle, pem_pos=15,
 final_scores=-1.6935073, final_local_scores=-1.7068117,
 pem=-3.8855858, similarity=-25.71453]

9: 129332 Grid_plan
 [winner: 530340 Control_grid, pem_pos=5,
 final_scores=-0.9441975, final_local_scores=-0.85121065,
 pem=-1.4983205, similarity=-13.164011|
 gt: 129332 Grid_plan, pem_pos=1,
 final_scores=-1.1352557, final_local_scores=-1.1723807,
 pem=-1.1315072, similarity=-24.743156]

10: 291718 Piazza
 [winner: 291718 Piazza, pem_pos=0,
 final_scores=-0.6964283, final_local_scores=-0.7520486,
 pem=-1.0710682, similarity=-12.528282|
 gt: 291718 Piazza, pem_pos=0, final_scores=-0.6964283,
 final_local_scores=-0.7520486, pem=-1.0710682,
 similarity=-12.528282]

11: 45883 Cathedral
 [winner: 13182282 Évry_Cathedral, pem_pos=0,
 final_scores=-0.8944283, final_local_scores=-0.87924683,
 pem=0.0, similarity=-22.148766|
 RECALL MISS: 45883 Cathedral]

12: 13692155 Philosophy
 [winner: 13498997 Constructivism_(psychological_school),
 pem_pos=7, final_scores=-1.3116155,
 final_local_scores=-1.2397561, pem=-1.9444226,
 similarity=-22.333008|
 gt: 13692155 Philosophy, pem_pos=0,
 final_scores=-1.5830644, final_local_scores=-1.6721141,
 pem=-0.12607375, similarity=-45.1239]

13: 37671 Aristocracy
 [winner: 181397 Patrician_(ancient_Rome), pem_pos=1,
 final_scores=-1.1216112, final_local_scores=-1.1974182,
 pem=-0.6925146, similarity=-27.877632|
 gt: 37671 Aristocracy, pem_pos=2,
 final_scores=-1.4963084, final_local_scores=-1.6258976,
 pem=-2.3076596, similarity=-31.8809]

14: 6325 Church_(building)
 [winner: 6325 Church_(building), pem_pos=0,
 final_scores=-1.1079743, final_local_scores=-1.1832683,
 pem=-0.35553187, similarity=-29.287933|

```

gt: 6325 Church_(building), pem_pos=0,
final_scores=-1.1079743, final_local_scores=-1.1832683,
pem=-0.35553187, similarity=-29.287933 ]
15: 2397739 Rosario_Gagliardi
[winner: 2397739 Rosario_Gagliardi, pem_pos=0,
final_scores=0.0026432723, final_local_scores=0.03131534,
pem=0.0, similarity=5.024225]
gt: 2397739 Rosario_Gagliardi, pem_pos=0,
final_scores=0.0026432723, final_local_scores=0.03131534,
pem=0.0, similarity=5.024225 ]
1: 27619 Sicily
[pem_pos=0, final_scores=0.46338338,
final_local_scores=0.44890517, pem=-0.66251653,
similarity=21.088806]
2: 157906 Noto
[pem_pos=0, final_scores=0.24936919,
final_local_scores=0.19879723, pem=-0.29883283,
similarity=11.647318]
3: 26667 Spain
[pem_pos=1, final_scores=0.525939,
final_local_scores=0.61152905, pem=-1.2425895,
similarity=29.096355]
4: 3957 Baroque
[pem_pos=0, final_scores=0.0887153,
final_local_scores=0.08190994, pem=-0.4497743,
similarity=8.980012]
5: 2397816 Giovanni_Battista_Vaccarini
[pem_pos=0, final_scores=0.17264624,
final_local_scores=0.20541081, pem=0.0, similarity=10.219582]
6: 2372757 Sicilian_Baroque
[pem_pos=0, final_scores=0.034392998,
final_local_scores=-0.0008584112, pem=0.0,
similarity=4.0640965]
1: 2911538 Giovanni_Battista_Landolina
[pem_pos=0, final_scores=0.11804937,
final_local_scores=0.16005597, pem=0.0, similarity=8.866102]
2: 157906 Noto
[pem_pos=0, final_scores=0.19976626,
final_local_scores=0.1446136, pem=-0.29883283,
similarity=10.030371]
3: 27619 Sicily
[pem_pos=0, final_scores=0.48396075,
final_local_scores=0.4713828, pem=-0.08753506,
similarity=18.632744]
4: 157906 Noto
[pem_pos=0, final_scores=0.23826711,
final_local_scores=0.18666992, pem=-0.29883283,
similarity=11.285415]

```

Sample: b'Human_image_synthesis'

Human image synthesis can be applied to make believable and even renditions of 1 human s , moving or still . A

synthesized human image is constructed by first 2 sampling the target by means of and 3 photography and creating a 4 3D based on the samples , applying 5 statistical information and 6 approximation s. Synthesis with an 7 actor and suitable 8 al gorithm s is applied using powerful 9 computer s . The actor 's part in the synthesis is to take care of mimicking human 10 expressions in still picture synthesizing and also human movement in synthesizing . Algorithms are needed to simulate laws of 11 physics and 12 physiology and to map the model and its appearance and movements accordingly . Many films using have featured synthetic images of humans or human-like characters . The bonus 13 DVD documents and depicts the process in some detail and the techniques used , including 14 facial and 15 limb al 16 motion capture

1: 682482 Human

[winner: 11772883 André_Bloch_(mathematician), pem_pos=0, final_scores=-1.2110194, final_local_scores=-1.0502675, pem=0.0, similarity=-27.252363]

RECALL MISS: 682482 Human]

2: 201605 Sampling_(signal_processing)

[winner: 37864 Nyquist-Shannon_sampling_theorem, pem_pos=2, final_scores=-1.3952553, final_local_scores=-1.2515173, pem=-1.0957491, similarity=-27.299202]

gt: 201605 Sampling_(signal_processing), pem_pos=4, final_scores=-1.7329547, final_local_scores=-1.6204026, pem=-2.3753495, similarity=-31.348814]

3: 23604 Photography

[winner: 2573710 Documentary_photography, pem_pos=2, final_scores=-1.4268345, final_local_scores=-1.2860126, pem=-1.3810041, similarity=-26.777353]

gt: 23604 Photography, pem_pos=0, final_scores=-1.5482851, final_local_scores=-1.418679, pem=-0.042842545, similarity=-38.013515]

4: 10175073 3D_computer_graphics

[winner: 10175073 3D_computer_graphics, pem_pos=0, final_scores=-0.53120255, final_local_scores=-0.3076709, pem=-0.7520788, similarity=-1.0018787]

gt: 10175073 3D_computer_graphics, pem_pos=0, final_scores=-0.53120255, final_local_scores=-0.3076709, pem=-0.7520788, similarity=-1.0018787]

5: 26685 Statistics

[winner: 18075640

Information_Minister_of_the_Palestinian_National_Authority, pem_pos=4, final_scores=-1.2228094, final_local_scores=-1.0631464, pem=-2.5649493, similarity=-13.688104]

RECALL MISS: 26685 Statistics]

[winner: 26685 Statistics, pem_pos=0, final_scores=-1.5209479, final_local_scores=-1.3888173, pem=-0.10773423, similarity=-36.769493]

gt: 26685 Statistics, pem_pos=0,

final_scores=-1.5209479, final_local_scores=-1.3888173,
 pem=-0.10773423, similarity=-36.769493]
 6: 336271 Approximation
 [winner: 563105 Approximation_algorithm, pem_pos=0,
 final_scores=-1.0381455, final_local_scores=-0.86142915,
 pem=0.0, similarity=-21.617052]
 gt: 336271 Approximation, pem_pos=1,
 final_scores=-1.6050161, final_local_scores=-1.480649,
 pem=-0.14525814, similarity=-39.305866]
 7: 21504235 Actor
 [winner: 8722094 Tom_Wright_(Australian_actor),
 pem_pos=0, final_scores=-1.1376697,
 final_local_scores=-0.97014403, pem=-1.0258529,
 similarity=-19.282572]
 RECALL MISS: 21504235 Actor]
 8: 775 Algorithm
 [winner: 10237 Exponentiation_by_squaring, pem_pos=0,
 final_scores=-1.2944157, final_local_scores=-1.1413653,
 pem=0.0, similarity=-29.970901]
 gt: 775 Algorithm, pem_pos=1, final_scores=-1.4225074,
 final_local_scores=-1.281286, pem=-0.10567855,
 similarity=-33.57172]
 9: 7878457 Computer
 [winner: 7272050 Rangekeeper, pem_pos=10,
 final_scores=-1.4039974, final_local_scores=-1.2610667,
 pem=-2.4849067, similarity=-20.029722]
 gt: 7878457 Computer, pem_pos=0,
 final_scores=-1.4570258, final_local_scores=-1.3189921,
 pem=-0.25448227, similarity=-33.887733]
 10: 242760 Facial_expression
 [winner: 938708 Expression_(computer_science), pem_pos=0,
 final_scores=-1.2977372, final_local_scores=-1.1449935,
 pem=0.0, similarity=-30.07918]
 gt: 242760 Facial_expression, pem_pos=3,
 final_scores=-2.0743928, final_local_scores=-1.9933717,
 pem=-3.79355, similarity=-34.76656]
 11: 22939 Physics
 [winner: 538206 Cartoon_physics, pem_pos=5,
 final_scores=-1.3180413, final_local_scores=-1.1671727,
 pem=-1.607689, similarity=-21.998184]
 gt: 22939 Physics, pem_pos=0, final_scores=-1.527723,
 final_local_scores=-1.3962181, pem=-0.07662393,
 similarity=-37.159523]
 12: 23597 Physiology
 [winner: 52502 Nobel_Prize_in_Physiology_or_Medicine,
 pem_pos=0, final_scores=-1.4137679,
 final_local_scores=-1.2717395, pem=0.0,
 similarity=-33.861526]
 gt: 23597 Physiology, pem_pos=1,
 final_scores=-1.6268622, final_local_scores=-1.5045125,
 pem=-0.09582976, similarity=-40.286804]
 13: 11014498 DVD

```

[winner: 11014498 DVD, pem_pos=0,
final_scores=-0.4567098, final_local_scores=-0.22629888,
pem=0.0, similarity=-2.6634972]
gt: 11014498 DVD, pem_pos=0, final_scores=-0.4567098,
final_local_scores=-0.22629888, pem=0.0,
similarity=-2.6634972 ]
14: 439102 Facial
[winner: 439102 Facial, pem_pos=1,
final_scores=-1.176098, final_local_scores=-1.0121212,
pem=-0.70433676, similarity=-22.283712]
gt: 439102 Facial, pem_pos=1, final_scores=-1.176098,
final_local_scores=-1.0121212, pem=-0.70433676,
similarity=-22.283712 ]
15: 13280372 Limb(anatomy)
[winner: 17492224 Congenital_limb_deformities, pem_pos=3,
final_scores=-1.068348, final_local_scores=-0.89442074,
pem=-2.367436, similarity=-9.727101]
gt: 13280372 Limb(anatomy), pem_pos=0,
final_scores=-1.3722976, final_local_scores=-1.2264395,
pem=0.0, similarity=-32.50969 ]
16: 396968 Motion_capture
[winner: 20580 Motion(physics), pem_pos=0,
final_scores=-1.1232282, final_local_scores=-0.95436895,
pem=-0.20888695, similarity=-23.2546]
RECALL MISS: 396968 Motion_capture]
[winner: 150643 Glossary_of_chess, pem_pos=0,
final_scores=-1.3414555, final_local_scores=-1.192749,
pem=-1.127043, similarity=-25.375261]
RECALL MISS: 396968 Motion_capture]

```

Sample: b'Gender_binary'

The gender binary , also referred to as gender binarism -LRB- sometimes shortened to just binarism -RRB- , is the classification of 1 sex and 2 gender into 3 two distinct , opposite and disconnected forms of 4 masculine and 5 feminine . It is one general type of a 6 gender system . As one of the core principles of 7 genderism , it can describe a 8 social boundary that discourages people from crossing or mixing 9 gender roles , or from identifying with third -LRB- or more -RRB- forms of 10 gender expression altogether . It can also represent some of the 11 prejudice s which 12 stigmatize 13 intersex and 14 transgender people , especially those that are 15 genderqueer - identified -- individuals who may not always fit neatly into the gender binary . The term describes the system in which a society splits people into 16 male and 17 female 18 gender roles , 19 gender identities and attributes . Gender role is one aspect of a gender binary . Many societies have used the gender binary to divide and organize people , though the ways this happens differs between societies . A universal aspect of the gender binaries is that women give birth . Gender binaries exist as a means of bringing order , though some , such as [1 1 Riki Wilchins 1] in GenderQueer :

Voices from Beyond the Sexual Binary , argue that gender binaries divide and polarize society . Certain notable religions are often used as authorities for the justification and description . [2 2 Islam 2] , for example , teaches that mothers are the primary care givers to their children and [3 3 Catholics 3] only allow males to serve as their priests . Exceptions have widely existed to the gender binary in the form of specific transgender identities . Besides the biological identification of intersex individuals , elements of the both or neither sexes have been taken by people biologically female and male such as 20 Two-Spirit ed [1 21 Native Americans 1] and 22 hijra of [4 4 India 4] . In the contemporary 23 West , transgender people break the gender binary in the form of 24 genderqueer . s have a unique place in relation to the gender binary because in many cases their gender expression transitions from one side of the gender binary to the other but still conforms to the gender binary itself . Many scholars have contested the existence of a clear gender binary . There is an increasing amount of research that illustrates that the evidence for dividing humans into the two distinct categories of men and women is problematic and a self-fulfilling prophecy . For example , [5 5 Judith Lorber 5] explains the problem of failing to question dividing people into these two groups `` even though they often find more significant within-group differences than between-group differences . '' [2 Lorber 2] argues that this corroborates the fact that the gender binary is quite arbitrary and leads to false expectations of both genders . Instead , there is growing support for the possibility of utilizing additional categories that compare people without `` prior assumptions about who is like whom . '' By allowing for a more fluid approach to gender , people will be better able to identify themselves however they choose , and scholarly research will find different similarities and differences . A further issue with the gender binary is the insistence that men are and women are . This reduces options for people to act outside of their 25 gender role

1: 1164715 Riki_Wilchins

2: 6037917 Islam

3: 606848 Catholic_Church

4: 14533 India

5: 36948008 Judith_Lorber

1: 26805 Sex

[winner: 6116892 Sex_and_gender_distinction, pem_pos=0, final_scores=-1.048292, final_local_scores=-1.1236002, pem=0.0, similarity=-29.440762]

RECALL MISS: 26805 Sex]

2: 38076 Gender

[winner: 6116892 Sex_and_gender_distinction, pem_pos=0, final_scores=-1.048292, final_local_scores=-1.1236002, pem=0.0, similarity=-29.440762]

RECALL MISS: 38076 Gender]

3: 1058899 Binary_opposition
[winner: 47285 Distinct, pem_pos=0,
final_scores=-1.0090996, final_local_scores=-0.9488872,
pem=-0.6663, similarity=-20.603533|
RECALL MISS: 1058899 Binary_opposition]
[winner: 5827481 Thornycroft_M-class_destroyer,
pem_pos=6, final_scores=-1.3929043,
final_local_scores=-1.3197997, pem=-3.1354942,
similarity=-18.244438|
RECALL MISS: 1058899 Binary_opposition]

4: 240058 Masculinity
[winner: 240058 Masculinity, pem_pos=1,
final_scores=-0.6470793, final_local_scores=-0.7008391,
pem=-0.16093658, similarity=-15.949525|
gt: 240058 Masculinity, pem_pos=1,
final_scores=-0.6470793, final_local_scores=-0.7008391,
pem=-0.16093658, similarity=-15.949525]

5: 240059 Femininity
[winner: 240059 Femininity, pem_pos=0,
final_scores=-1.1290897, final_local_scores=-1.2209301,
pem=-0.43592775, similarity=-29.974634|
gt: 240059 Femininity, pem_pos=0,
final_scores=-1.1290897, final_local_scores=-1.2209301,
pem=-0.43592775, similarity=-29.974634]

6: 29880756 Gender_systems
[winner: 6813680 .sys, pem_pos=15,
final_scores=-1.1346775, final_local_scores=-1.0196083,
pem=-3.4339871, similarity=-7.6628714|
RECALL MISS: 29880756 Gender_systems]
[winner: 38076 Gender, pem_pos=0,
final_scores=-1.1723691, final_local_scores=-1.321003,
pem=-0.26921824, similarity=-33.867596|
RECALL MISS: 29880756 Gender_systems]

7: 963393 Genderism
[winner: 963393 Genderism, pem_pos=1,
final_scores=-0.95099497, final_local_scores=-0.9023987,
pem=-2.1972246, similarity=-10.890816|
gt: 963393 Genderism, pem_pos=1,
final_scores=-0.95099497, final_local_scores=-0.9023987,
pem=-2.1972246, similarity=-10.890816]

8: 61117 Taboo
[winner: 7275586 Boundary_marker, pem_pos=0,
final_scores=-1.1823065, final_local_scores=-1.21598,
pem=0.0, similarity=-32.19756|
RECALL MISS: 61117 Taboo]
[winner: 1696316 Social_actions, pem_pos=5,
final_scores=-1.6265134, final_local_scores=-1.7536418,
pem=-2.6473815, similarity=-33.845585|
RECALL MISS: 61117 Taboo]

9: 43094 Gender_role
[winner: 43094 Gender_role, pem_pos=0,
final_scores=-0.878073, final_local_scores=-0.9867421,

pem=-0.018132841, similarity=-25.25803|
gt: 43094 Gender_role, pem_pos=0,
final_scores=-0.878073, final_local_scores=-0.9867421,
pem=-0.018132841, similarity=-25.25803]
10: 162025 Gender_identity
[winner: 938708 Expression_(computer_science), pem_pos=0,
final_scores=-1.0848271, final_local_scores=-1.0408332,
pem=-0.47098094, similarity=-24.409557|
RECALL MISS: 162025 Gender_identity]
[winner: 38076 Gender, pem_pos=0,
final_scores=-1.2294357, final_local_scores=-1.3833394,
pem=-0.26921824, similarity=-35.72784|
gt: 162025 Gender_identity, pem_pos=3,
final_scores=-1.4542929, final_local_scores=-1.6041319,
pem=-1.569396, similarity=-35.246166]
11: 77182 Prejudice
[winner: 77182 Prejudice, pem_pos=1,
final_scores=-1.2060175, final_local_scores=-1.3211963,
pem=-0.05326471, similarity=-35.04776|
gt: 77182 Prejudice, pem_pos=1, final_scores=-1.2060175,
final_local_scores=-1.3211963, pem=-0.05326471,
similarity=-35.04776]
12: 2649767 Social_stigma
[winner: 2649767 Social_stigma, pem_pos=0,
final_scores=-1.279631, final_local_scores=-1.3871958,
pem=0.0, similarity=-37.30698|
gt: 2649767 Social_stigma, pem_pos=0,
final_scores=-1.279631, final_local_scores=-1.3871958,
pem=0.0, similarity=-37.30698]
13: 26652964 Intersex
[winner: 26652964 Intersex, pem_pos=0,
final_scores=-0.9866232, final_local_scores=-1.0838691,
pem=-0.0072579877, similarity=-28.215633|
gt: 26652964 Intersex, pem_pos=0,
final_scores=-0.9866232, final_local_scores=-1.0838691,
pem=-0.0072579877, similarity=-28.215633]
14: 19904525 Transgender
[winner: 19904525 Transgender, pem_pos=0,
final_scores=-1.5083538, final_local_scores=-1.6831152,
pem=-0.011912266, similarity=-46.07302|
gt: 19904525 Transgender, pem_pos=0,
final_scores=-1.5083538, final_local_scores=-1.6831152,
pem=-0.011912266, similarity=-46.07302]
15: 264989 Genderqueer
[winner: 264989 Genderqueer, pem_pos=0,
final_scores=-0.98321223, final_local_scores=-1.0319891,
pem=0.0, similarity=-26.706902|
gt: 264989 Genderqueer, pem_pos=0,
final_scores=-0.98321223, final_local_scores=-1.0319891,
pem=0.0, similarity=-26.706902]
16: 1822282 Male
[winner: 1025265 Gender_of_connectors_and_fasteners,

pem_pos=0, final_scores=-1.0318555,
final_local_scores=-0.98481905, pem=-0.4906229,
similarity=-22.631168|
RECALL MISS: 1822282 Male]

17: 20913864 Female
[winner: 1025265 Gender_of_connectors_and_fasteners,
pem_pos=0, final_scores=-1.0318555,
final_local_scores=-0.98481905, pem=-0.4906229,
similarity=-22.631168|
RECALL MISS: 20913864 Female]

18: 43094 Gender_role
[winner: 76749 Character_(arts), pem_pos=0,
final_scores=-1.0779173, final_local_scores=-1.0885359,
pem=0.0, similarity=-28.39437|
gt: 43094 Gender_role, pem_pos=4,
final_scores=-1.827344, final_local_scores=-2.0236764,
pem=-3.7818756, similarity=-35.7344]
[winner: 38076 Gender, pem_pos=0,
final_scores=-1.1888779, final_local_scores=-1.3390362,
pem=-0.26921824, similarity=-34.40575|
gt: 43094 Gender_role, pem_pos=1,
final_scores=-1.3520334, final_local_scores=-1.5044718,
pem=-1.4983028, similarity=-32.658722]

19: 162025 Gender_identity
[winner: 14962 Identity_element, pem_pos=2,
final_scores=-1.3157405, final_local_scores=-1.3037491,
pem=-1.0859271, similarity=-28.911322|
RECALL MISS: 162025 Gender_identity]
[winner: 38076 Gender, pem_pos=0,
final_scores=-1.3933667, final_local_scores=-1.5624092,
pem=-0.26921824, similarity=-41.07164|
gt: 162025 Gender_identity, pem_pos=3,
final_scores=-1.5716332, final_local_scores=-1.7323084,
pem=-1.569396, similarity=-39.07121]

20: 23929145 Two-Spirit
[winner: 23929145 Two-Spirit, pem_pos=0,
final_scores=-0.016607508, final_local_scores=-0.03514467,
pem=0.0, similarity=3.0409257|
gt: 23929145 Two-Spirit, pem_pos=0,
final_scores=-0.016607508, final_local_scores=-0.03514467,
pem=0.0, similarity=3.0409257]

21: 2402376 Indigenous_peoples_of_the_Americas
[winner: 21217 Native_Americans_in_the_United_States,
pem_pos=0, final_scores=0.20382367,
final_local_scores=0.22504044, pem=-0.29015967,
similarity=12.383302|
gt: 2402376 Indigenous_peoples_of_the_Americas,
pem_pos=1, final_scores=0.087814406,
final_local_scores=0.082491174, pem=-1.0635396,
similarity=12.335108]

22: 301626 Hijra_(South_Asia)
[winner: 301626 Hijra_(South_Asia), pem_pos=0,

```

final_scores=-0.71669066, final_local_scores=-0.77300835,
pem=-0.27908552, similarity=-17.460688|
    gt: 301626 Hijra_(South_Asia), pem_pos=0,
final_scores=-0.71669066, final_local_scores=-0.77300835,
pem=-0.27908552, similarity=-17.460688 ]
23: 21208200 Western_world
    [winner: 21208200 Western_world, pem_pos=0,
final_scores=-0.16076533, final_local_scores=-0.21814217,
pem=-1.6371275, similarity=6.482869|
    gt: 21208200 Western_world, pem_pos=0,
final_scores=-0.16076533, final_local_scores=-0.21814217,
pem=-1.6371275, similarity=6.482869 ]
24: 264989 Genderqueer
    [winner: 264989 Genderqueer, pem_pos=0,
final_scores=-0.940502, final_local_scores=-0.98533463,
pem=0.0, similarity=-25.314638|
    gt: 264989 Genderqueer, pem_pos=0,
final_scores=-0.940502, final_local_scores=-0.98533463,
pem=0.0, similarity=-25.314638 ]
25: 43094 Gender_role
    [winner: 76749 Character_(arts), pem_pos=0,
final_scores=-1.0558625, final_local_scores=-1.0644443,
pem=0.0, similarity=-27.675426|
    gt: 43094 Gender_role, pem_pos=4,
final_scores=-1.7843279, final_local_scores=-1.9766879,
pem=-3.7818756, similarity=-34.33217 ]
    [winner: 38076 Gender, pem_pos=0,
final_scores=-1.3150445, final_local_scores=-1.4768541,
pem=-0.26921824, similarity=-38.518505|
    gt: 43094 Gender_role, pem_pos=1,
final_scores=-1.4658387, final_local_scores=-1.6287868,
pem=-1.4983028, similarity=-36.368534 ]
1: 1164715 Riki_Wilchins
    [pem_pos=0, final_scores=0.41596997,
final_local_scores=0.52179325, pem=0.0, similarity=19.661064]
2: 6037917 Islam
    [pem_pos=0, final_scores=0.32507557,
final_local_scores=0.33150595, pem=-0.12106096,
similarity=14.640861]
3: 606848 Catholic_Church
    [pem_pos=0, final_scores=0.18641306,
final_local_scores=0.19060585, pem=-0.38803694,
similarity=11.887977]
4: 14533 India
    [pem_pos=0, final_scores=0.5411929,
final_local_scores=0.6105918, pem=-0.13622136,
similarity=23.051785]
5: 36948008 Judith_Lorber
    [pem_pos=0, final_scores=0.23852243,
final_local_scores=0.28244135, pem=0.0, similarity=12.518326]
1: 21217 Native_Americans_in_the_United_States
    [pem_pos=0, final_scores=0.20382367,

```

```
final_local_scores=0.22504044, pem=-0.29015967,  
similarity=12.383302]  
2: 36948008 Judith_Lorber  
[pem_pos=0, final_scores=0.08317749,  
final_local_scores=0.112750605, pem=0.0, similarity=7.454416]
```

Sample: b'Steve_Durbano'

Harry Steven `` Mental Case '' [1 Durbano 1] -LRB- December 12 , 1951 in [1 1 Toronto 1] , [2 2 Ontario 2] -- November 10 , 2002 in [3 3 Yellowknife 3] , [4 4 Northwest Territories 4] -RRB- was a professional 1 ice hockey player noted for his villainous behavior on the ice and his larger-than-life persona off it . Born in [2 Toronto 2] , [3 Durbano 3] rose through minor hockey to a place on the [5 5 Toronto Marlboros 5] of the [4 2 Ontario Hockey Association 4] . He was drafted by the [6 6 New York Rangers 6] in the first round of the 3 19 71 [5 NHL Amateur Draft 5] . He never played a game for the [6 Rangers 6] , though , and was traded to the [7 7 St. Louis Blues 7] where he began his career . [7 Durbano 7] quickly gained a reputation for being one of the toughest players in the league , amassing 1,411 major league-career , including an NHL-leading 370 PIM in the 4 1975 -- 76 season . He played 220 [8 8 NHL 8] games over the course of his career for St. Louis as well as the [9 9 Pittsburgh Penguins 9] , [10 10 Kansas City Scouts 10] and [8 5 Colorado Rockies 8] , along with 45 games in the [11 11 World Hockey Association 11] for the 6 Birmingham Bulls . In 1983 , shortly after he retired , [9 Durbano 9] was implicated in a scheme to import half a million dollars worth of 7 cocaine into [10 Canada 10] and was sentenced to seven years in prison . After his release , Durbano still could n't keep straight - he was arrested once for shoplifting , and , in 1995 , he was sent back to prison after he attempted to recruit an undercover police officer into an in . He moved to the [12 12 Northwest Territories 12] late in life in an effort to find peace , but died there of 8 liver cancer

```
1: 64646 Toronto  
2: 22218 Ontario  
3: 101392 Yellowknife  
4: 21186 Northwest_Territories  
5: 1034626 Toronto_Marlboros  
6: 66881 New_York_Rangers  
7: 30519527 St._Louis_Blues  
8: 21809 National_Hockey_League  
9: 66964 Pittsburgh_Penguins  
10: 665626 Kansas_City_Scouts  
11: 240143 World_Hockey_Association  
12: 21186 Northwest_Territories
```

```
1: 14790 Ice_hockey  
[winner: 14946 Ice, pem_pos=0, final_scores=-0.7208982,  
final_local_scores=-0.67018396, pem=-0.116236255,  
similarity=-15.2778015]
```

gt: 14790 Ice_hockey, pem_pos=3,
 final_scores=-1.0658029, final_local_scores=-1.1842163,
 pem=-2.165202, similarity=-19.474955]
 [winner: 6904528 Pakistan_national_field_hockey_team,
 pem_pos=0, final_scores=-1.4912661,
 final_local_scores=-1.5631976, pem=0.0, similarity=-42.55922]
 gt: 14790 Ice_hockey, pem_pos=5,
 final_scores=-1.9119126, final_local_scores=-2.1084626,
 pem=-2.5494452, similarity=-44.966736]
 2: 218742 Ontario_Hockey_League
 [winner: 1458074 Ontario_Hockey_Association, pem_pos=0,
 final_scores=0.17048977, final_local_scores=0.14125118,
 pem=-0.20795111, similarity=9.4358]
 gt: 218742 Ontario_Hockey_League, pem_pos=1,
 final_scores=-0.06623037, final_local_scores=-0.1294967,
 pem=-1.6968997, similarity=9.453278]
 3: 1213511 1971_NHL_Amateur_Draft
 [winner: 858982 NHL_Entry_Draft, pem_pos=0,
 final_scores=0.071023926, final_local_scores=0.04114628,
 pem=0.0, similarity=5.3176]
 RECALL MISS: 1213511 1971_NHL_Amateur_Draft]
 [winner: 34748 1971, pem_pos=0, final_scores=-1.105464,
 final_local_scores=-1.0246708, pem=-0.6167389,
 similarity=-23.134588]
 RECALL MISS: 1213511 1971_NHL_Amateur_Draft]
 4: 2945254 1975-76_NHL_season
 [winner: 34752 1975, pem_pos=0, final_scores=-0.7204512,
 final_local_scores=-0.60554445, pem=-0.61872095,
 similarity=-10.61624]
 RECALL MISS: 2945254 1975-76_NHL_season]
 [winner: 21598459 KRT76, pem_pos=16,
 final_scores=-0.9033748, final_local_scores=-0.7023239,
 pem=-3.951379, similarity=4.6191792]
 RECALL MISS: 2945254 1975-76_NHL_season]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.2599657, final_local_scores=-1.1550262,
 pem=0.0, similarity=-30.378574]
 RECALL MISS: 2945254 1975-76_NHL_season]
 [winner: 690345 Season(sports), pem_pos=0,
 final_scores=-1.404655, final_local_scores=-1.5728197,
 pem=-1.219945, similarity=-36.212112]
 RECALL MISS: 2945254 1975-76_NHL_season]
 5: 6017954 Colorado_Rockies_(NHL)
 [winner: 6669 Colorado_Rockies, pem_pos=0,
 final_scores=0.3242563, final_local_scores=0.34273618,
 pem=-0.10158596, similarity=14.870085]
 gt: 6017954 Colorado_Rockies_(NHL), pem_pos=1,
 final_scores=-0.08407399, final_local_scores=-0.1536157,
 pem=-1.4611784, similarity=7.4516296]
 6: 1955053 Birmingham_Bulls
 [winner: 1955053 Birmingham_Bulls, pem_pos=0,
 final_scores=-0.028116122, final_local_scores=-0.049305692,

pem=0.0, similarity=2.6183326|
gt: 1955053 Birmingham_Bulls, pem_pos=0,
final_scores=-0.028116122, final_local_scores=-0.049305692,
pem=0.0, similarity=2.6183326]
7: 7701 Cocaine
[winner: 14079207 Cocaine_dependence, pem_pos=0,
final_scores=-1.1094594, final_local_scores=-1.0871964,
pem=0.0, similarity=-28.354397|
gt: 7701 Cocaine, pem_pos=1, final_scores=-1.3772265,
final_local_scores=-1.4154532, pem=-0.014357626,
similarity=-38.072155]
8: 237585 Hepatocellular_carcinoma
[winner: 263759 Liver_tumor, pem_pos=1,
final_scores=-1.0200835, final_local_scores=-0.87342525,
pem=-0.8778436, similarity=-17.201189|
gt: 237585 Hepatocellular_carcinoma, pem_pos=0,
final_scores=-1.2222466, final_local_scores=-1.2272282,
pem=0.0, similarity=-32.533222]
1: 64646 Toronto
[pem_pos=0, final_scores=0.537451,
final_local_scores=0.5545755, pem=-0.12291032,
similarity=21.307758]
2: 22218 Ontario
[pem_pos=0, final_scores=0.36969608,
final_local_scores=0.37651354, pem=-0.10088341,
similarity=15.874248]
3: 101392 Yellowknife
[pem_pos=0, final_scores=0.29340285,
final_local_scores=0.33426782, pem=-0.10625713,
similarity=14.642775]
4: 21186 Northwest_Territories
[pem_pos=0, final_scores=0.3027857,
final_local_scores=0.3460263, pem=0.0, similarity=14.415828]
5: 1034626 Toronto_Marlboros
[pem_pos=0, final_scores=0.14456196,
final_local_scores=0.123798996, pem=0.0, similarity=7.7841215]
6: 66881 New_York_Rangers
[pem_pos=0, final_scores=0.27634257,
final_local_scores=0.22608835, pem=0.0, similarity=10.83664]
7: 30519527 St._Louis_Blues
[pem_pos=0, final_scores=0.19298129,
final_local_scores=0.14142728, pem=-0.11159538,
similarity=8.917058]
8: 21809 National_Hockey_League
[pem_pos=0, final_scores=0.25890142,
final_local_scores=0.1921359, pem=-0.036005534,
similarity=10.019235]
9: 66964 Pittsburgh_Penguins
[pem_pos=0, final_scores=0.26532286,
final_local_scores=0.22103998, pem=0.0, similarity=10.6859865]
10: 665626 Kansas_City_Scouts
[pem_pos=0, final_scores=0.20587458,

```

final_local_scores=0.17966864, pem=0.0, similarity=9.451384]
11: 240143 World_Hockey_Association
    [pem_pos=0, final_scores=0.28530335,
final_local_scores=0.24941903, pem=0.0, similarity=11.532874]
12: 21186 Northwest_Territories
    [pem_pos=0, final_scores=0.3722964,
final_local_scores=0.42195618, pem=0.0, similarity=16.681726]
1: 5785148 Steve_Durbano
    [pem_pos=0, final_scores=0.12804209,
final_local_scores=0.14773566, pem=0.0, similarity=8.49844]
2: 64646 Toronto
    [pem_pos=0, final_scores=0.48093265,
final_local_scores=0.4928378, pem=-0.12291032,
similarity=19.465382]
3: 5785148 Steve_Durbano
    [pem_pos=0, final_scores=0.13915218,
final_local_scores=0.15987173, pem=0.0, similarity=8.860604]
4: 1458074 Ontario_Hockey_Association
    [pem_pos=0, final_scores=0.17048977,
final_local_scores=0.14125118, pem=-0.20795111,
similarity=9.4358]
5: 858982 NHL_Entry_Draft
    [pem_pos=0, final_scores=0.071023926,
final_local_scores=0.04114628, pem=0.0, similarity=5.3176]
6: 30857 Texas_Rangers_(baseball)
    [pem_pos=1, final_scores=0.25109273,
final_local_scores=0.25806937, pem=-2.6930556,
similarity=26.436274]
7: 5785148 Steve_Durbano
    [pem_pos=0, final_scores=0.21037139,
final_local_scores=0.23766789, pem=0.0, similarity=11.182197]
8: 6669 Colorado_Rockies
    [pem_pos=0, final_scores=0.3242563,
final_local_scores=0.34273618, pem=-0.10158596,
similarity=14.870085]
9: 5785148 Steve_Durbano
    [pem_pos=0, final_scores=0.1432365,
final_local_scores=0.16433325, pem=0.0, similarity=8.993745]
10: 5042916 Canada
    [pem_pos=0, final_scores=0.6646432,
final_local_scores=0.7046433, pem=-0.11662793,
similarity=25.751915]

```

Sample: b'Diabolique'

-LRB- band -RRB- Diabolique is a [1 Swedish 1] 1 gothic me
tal band , formed in 1995 after the dissolution of 2 Liers in
Wait . The group is heavily influenced by [1 1 The Sisters of
Mercy 1] , [2 2 Black Sabbath 2] , and [3 3 Fields of the Nep
hilim 3] , and in sound are similar to the styles of 3 Charon
 , 4 Tiamat , 5 Moonspell , and [4 4 Type O Negative 4] . The
band 's current and former line-up has consisted of many
musicians from progenital [5 5 Gothenburg 5] 6 melodic death m

etal bands ; [6 6 Kristian Wåhlin 6] , Johan Osterberg , and [2 7 Alf Svensson 2] were in 8 Grotesque -LRB- a predecessor of 9 At the Gates , who Alf would eventually join -RRB- ; and 10 Daniel Svensson briefly played drums with the group before going on to major success with [7 7 In Flames 7] . The band 's most recent release was The Green Goddess -LRB- 2001 -RRB- . Various sources differ as to band 's status . However , [3 11 Hans Nilsson 3] 's full-time involvement with 12 Dimension Zero and [4 Wåhlin 4] 's extensive contributions as an album cover artist and musician for other [8 8 Scandina

via 8] n-area 13 extreme metal

1: 162144 The_Sisters_of_Mercy

2: 4314 Black_Sabbath

3: 174383 Fields_of_the_Nephilim

4: 554883 Type_0_Negative

5: 11861 Gothenburg

6: 2819782 Kristian_Wåhlin

7: 737240 In_Flames

8: 26740 Scandinavia

1: 3203201 Gothic_metal

[winner: 3203201 Gothic_metal, pem_pos=0, final_scores=-0.7396818, final_local_scores=-0.8525573, pem=0.0, similarity=-21.352297]

gt: 3203201 Gothic_metal, pem_pos=0, final_scores=-0.7396818, final_local_scores=-0.8525573, pem=0.0, similarity=-21.352297]

2: 7457615 Liers_in_Wait

[winner: 7457615 Liers_in_Wait, pem_pos=0, final_scores=-0.05895315, final_local_scores=-0.04565724, pem=0.0, similarity=2.7272098]

gt: 7457615 Liers_in_Wait, pem_pos=0, final_scores=-0.05895315, final_local_scores=-0.04565724, pem=0.0, similarity=2.7272098]

3: 1079447 Charon_(band)

[winner: 52838 Charon_(moon), pem_pos=1, final_scores=-0.31600425, final_local_scores=-0.21895544, pem=-0.7947314, similarity=1.8775188]

gt: 1079447 Charon_(band), pem_pos=2, final_scores=-0.38029382, final_local_scores=-0.380485, pem=-1.8126502, similarity=2.5927548]

4: 888656 Tiamat_(band)

[winner: 181591 Tiamat, pem_pos=0, final_scores=-0.067505926, final_local_scores=0.025726557, pem=-0.21400948, similarity=6.021261]

gt: 888656 Tiamat_(band), pem_pos=1, final_scores=-0.08231558, final_local_scores=-0.110764794, pem=-0.59298015, similarity=4.008989]

5: 422741 Moonspell

[winner: 422741 Moonspell, pem_pos=0, final_scores=0.026099965, final_local_scores=7.343292e-05, pem=0.0, similarity=4.091904]

gt: 422741 Moonspell, pem_pos=0,

final_scores=0.026099965, final_local_scores=7.343292e-05,
pem=0.0, similarity=4.091904]
6: 633174 Melodic_death_metal
[winner: 209048 Swedish_death_metal, pem_pos=0,
final_scores=-0.67074573, final_local_scores=-0.74496984,
pem=0.0, similarity=-18.141674]
RECALL MISS: 633174 Melodic_death_metal]
[winner: 37735 Melody, pem_pos=0,
final_scores=-1.2666221, final_local_scores=-1.3593032,
pem=-0.12246316, similarity=-35.808636]
gt: 633174 Melodic_death_metal, pem_pos=2,
final_scores=-1.8335905, final_local_scores=-2.058378,
pem=-3.7981608, similarity=-36.6814]
7: 8839941 Alf_Svensson_(guitarist)
[winner: 302209 Alf_Svensson, pem_pos=0,
final_scores=0.08305226, final_local_scores=0.19966477,
pem=0.0, similarity=10.048108]
gt: 8839941 Alf_Svensson_(guitarist), pem_pos=1,
final_scores=-0.09855126, final_local_scores=-0.067563556,
pem=-0.16454941, similarity=2.9683268]
8: 2819665 Grotesque_(band)
[winner: 2819665 Grotesque_(band), pem_pos=3,
final_scores=-0.5139753, final_local_scores=-0.5255527,
pem=-1.5334953, similarity=-3.2544408]
gt: 2819665 Grotesque_(band), pem_pos=3,
final_scores=-0.5139753, final_local_scores=-0.5255527,
pem=-1.5334953, similarity=-3.2544408]
9: 586983 At_the_Gates
[winner: 586983 At_the_Gates, pem_pos=0,
final_scores=-0.06970602, final_local_scores=-0.12107473,
pem=0.0, similarity=0.47660157]
gt: 586983 At_the_Gates, pem_pos=0,
final_scores=-0.06970602, final_local_scores=-0.12107473,
pem=0.0, similarity=0.47660157]
10: 1911547 Daniel_Svensson
[winner: 1911547 Daniel_Svensson, pem_pos=0,
final_scores=0.024435654, final_local_scores=0.10560371,
pem=0.0, similarity=7.2411385]
gt: 1911547 Daniel_Svensson, pem_pos=0,
final_scores=0.024435654, final_local_scores=0.10560371,
pem=0.0, similarity=7.2411385]
11: 9415654 Hans_Nilsson_(musician)
[winner: 22799527 Hans_Nilsson_(canoeist), pem_pos=0,
final_scores=0.2507242, final_local_scores=0.44799477,
pem=0.0, similarity=17.45877]
gt: 9415654 Hans_Nilsson_(musician), pem_pos=1,
final_scores=0.0028837472, final_local_scores=0.051038176,
pem=-0.37469345, similarity=7.650435]
12: 2098028 Dimension_Zero_(Swedish_band)
[winner: 2098028 Dimension_Zero_(Swedish_band),
pem_pos=0, final_scores=-0.024506092,
final_local_scores=-0.029410645, pem=0.0,

```

similarity=3.2120407|
    gt: 2098028 Dimension_Zero_(Swedish_band), pem_pos=0,
final_scores=-0.024506092, final_local_scores=-0.029410645,
pem=0.0, similarity=3.2120407 ]
13: 633178 Extreme_metal
    [winner: 633178 Extreme_metal, pem_pos=0,
final_scores=-0.94765615, final_local_scores=-1.0750498,
pem=-0.009245085, similarity=-27.94164]
    gt: 633178 Extreme_metal, pem_pos=0,
final_scores=-0.94765615, final_local_scores=-1.0750498,
pem=-0.009245085, similarity=-27.94164 ]
1: 162144 The_Sisters_of_Mercy
    [pem_pos=0, final_scores=0.2480809,
final_local_scores=0.2509172, pem=0.0, similarity=11.577582]
2: 4314 Black_Sabbath
    [pem_pos=0, final_scores=0.26510936,
final_local_scores=0.24693474, pem=0.0, similarity=11.458738]
3: 174383 Fields_of_the_Nephilim
    [pem_pos=0, final_scores=0.078859076,
final_local_scores=0.0691023, pem=0.0, similarity=6.1518636]
4: 554883 Type_0_Negative
    [pem_pos=0, final_scores=0.05377631,
final_local_scores=0.026386023, pem=0.0, similarity=4.8771243]
5: 11861 Gothenburg
    [pem_pos=0, final_scores=0.6770198,
final_local_scores=0.7400172, pem=0.0, similarity=26.173302]
6: 2819782 Kristian_Wåhlin
    [pem_pos=0, final_scores=0.1496677,
final_local_scores=0.14684448, pem=0.0, similarity=8.471845]
7: 737240 In_Flames
    [pem_pos=0, final_scores=0.069928154,
final_local_scores=0.028082296, pem=0.0, similarity=4.927745]
8: 26740 Scandinavia
    [pem_pos=0, final_scores=0.0518872,
final_local_scores=0.06722979, pem=-0.08773311,
similarity=6.5730906]
1: 5058739 Sweden
    [pem_pos=0, final_scores=0.6848542,
final_local_scores=0.7411348, pem=-0.63135606,
similarity=29.640066]
2: 302209 Alf_Svensson
    [pem_pos=0, final_scores=0.08305226,
final_local_scores=0.19966477, pem=0.0, similarity=10.048108]
3: 22799527 Hans_Nilsson_(canoeist)
    [pem_pos=0, final_scores=0.2507242,
final_local_scores=0.44799477, pem=0.0, similarity=17.45877]
4: 2819782 Kristian_Wåhlin
    [pem_pos=0, final_scores=0.09006359,
final_local_scores=0.08173604, pem=0.0, similarity=6.5288796]

```

Sample: b'Krasi'

, _ Thalassa_Kai_T ' _ Agori_Mou `` [1 Krasi 1] , thalassa

kai t ' agori mou '' -LRB- [2 1 Greek 2] : '' Κρασί , θάλασσα
 και τ ' αγόρι μου '' ; [3 English 3] translation : `` Wine ,
 sea and my boyfriend '' -RRB- was the [1 1 Greek 1] entry in
 the [4 2 Eurovision Song Contest 4] 1974 , performed in [5 3 G
 reek 5] by 4 Marinella . The song is a typical [6 Greek 6]
 traditional 5 bouzouki tune , with Marinella singing that all
 she needs in life is `` wine , the sea and -LSB- her -RSB-
 boyfriend '' . Composed and conducted by 6 Giorgos Katsaros
 with lyrics by 7 Pythagoras , `` Krasi , thalassa kai t '
 agori mou '' was the first song to represent [7 8 Greece 7] at
 the [8 Eurovision Song Contest 8] . Popular [9 Greek 9]
 singer 9 Marinella was selected to sing the song at the [10 10
 Eurovision Song Contest 10] 1974 which took place in [2 2 Bri
 ghton 2] , [3 3 England 3] on April 6 , 1974 . The song was
 presented fifth out of the seventeen entries of the night and
 followed [4 4 Norway 4] 's [5 5 Anne-Karine Strøm 5] with '' 1
 1 The First Day of Love '' while preceding [6 6 Israel 6] 's 1
 2 Poogy with '' 13 Natati La Khayay '' . At the close of
 voting , the performance had received seven points , placing
 it eleventh in the field of 17 . `` Krasi , Thalassa Kai T '
 Agori Mou '' was succeeded as the [11 Greek 11]
 representative at the 14 1976 Contest by [7 7 Mariza Koch 7]
 with '' '' . [12 Greece 12] had pulled out of the 15 1975 Con
 test for `` unknown reasons '' according to the 16 EBU , but
 it was later discovered that the withdrawal was in protest of
 Turkey 's debut and its 17 invasion of [13 Cyprus 13]

1: 12108 Greece

2: 49796 Brighton

3: 9316 England

4: 21241 Norway

5: 25219107 Anne-Karine_Strøm

6: 9282173 Israel

7: 23670710 Mariza_Koch

1: 715909 Greek_alphabet

[winner: 12108 Greece, pem_pos=1,
 final_scores=0.29358816, final_local_scores=0.23292875,
 pem=-2.3247876, similarity=23.683329]

gt: 715909 Greek_alphabet, pem_pos=6,
 final_scores=-0.8763432, final_local_scores=-0.9270535,
 pem=-3.4867885, similarity=-4.6137156]

2: 192545 Eurovision_Song_Contest_1974

[winner: 9954 Eurovision_Song_Contest, pem_pos=0,
 final_scores=0.12871097, final_local_scores=0.10978496,
 pem=-0.1344273, similarity=8.0969515]

RECALL MISS: 192545 Eurovision_Song_Contest_1974]

[winner: 34654 1974, pem_pos=0, final_scores=-0.5852305,
 final_local_scores=-0.4320386, pem=-0.70012724,
 similarity=-4.99578]

RECALL MISS: 192545 Eurovision_Song_Contest_1974]

3: 11887 Greek_language

[winner: 12108 Greece, pem_pos=1,
 final_scores=0.45307356, final_local_scores=0.40714228,

pem=-2.3247876, similarity=28.882208|
gt: 11887 Greek_language, pem_pos=0,
final_scores=-0.36227232, final_local_scores=-0.38300923,
pem=-0.6261292, similarity=-3.9350586]
4: 748098 Marinella
[winner: 748098 Marinella, pem_pos=0,
final_scores=-0.17577834, final_local_scores=-0.17123988,
pem=0.0, similarity=-1.0204268|
gt: 748098 Marinella, pem_pos=0,
final_scores=-0.17577834, final_local_scores=-0.17123988,
pem=0.0, similarity=-1.0204268]
5: 208837 Bouzouki
[winner: 208837 Bouzouki, pem_pos=0,
final_scores=-1.1580032, final_local_scores=-1.2224753,
pem=0.0, similarity=-32.391388|
gt: 208837 Bouzouki, pem_pos=0, final_scores=-1.1580032,
final_local_scores=-1.2224753, pem=0.0,
similarity=-32.391388]
6: 4230193 Giorgos_Katsaros
[winner: 4230193 Giorgos_Katsaros, pem_pos=0,
final_scores=-0.19234833, final_local_scores=-0.07670135,
pem=0.0, similarity=1.8007916|
gt: 4230193 Giorgos_Katsaros, pem_pos=0,
final_scores=-0.19234833, final_local_scores=-0.07670135,
pem=0.0, similarity=1.8007916]
7: 10861801 Pythagoras_Papastamatiou
[winner: 23275 Pythagoras, pem_pos=0,
final_scores=-0.08776967, final_local_scores=-0.03459879,
pem=0.0, similarity=3.057216|
gt: 10861801 Pythagoras_Papastamatiou, pem_pos=4,
final_scores=-0.43152404, final_local_scores=-0.3903775,
pem=-1.6067936, similarity=1.1780635]
8: 3867528 Greece_in_the_Eurovision_Song_Contest
[winner: 12108 Greece, pem_pos=0, final_scores=0.8960094,
final_local_scores=0.89098245, pem=-0.27607638,
similarity=32.179752|
gt: 3867528 Greece_in_the_Eurovision_Song_Contest,
pem_pos=8, final_scores=-0.8352369,
final_local_scores=-0.90417284, pem=-4.3479457,
similarity=0.7521956]
[winner: 9954 Eurovision_Song_Contest, pem_pos=0,
final_scores=0.11598174, final_local_scores=0.09588021,
pem=-0.1344273, similarity=7.6820064|
RECALL MISS: 3867528
Greece_in_the_Eurovision_Song_Contest]
[winner: 1928089 Conservation_status, pem_pos=1,
final_scores=-1.6904916, final_local_scores=-1.6959722,
pem=-0.117783025, similarity=-45.880962|
RECALL MISS: 3867528
Greece_in_the_Eurovision_Song_Contest]
9: 748098 Marinella
[winner: 748098 Marinella, pem_pos=0,

final_scores=-0.19133776, final_local_scores=-0.18823618,
 pem=0.0, similarity=-1.52763|
 gt: 748098 Marinella, pem_pos=0,
 final_scores=-0.19133776, final_local_scores=-0.18823618,
 pem=0.0, similarity=-1.52763]
 10: 192545 Eurovision_Song_Contest_1974
 [winner: 9954 Eurovision_Song_Contest, pem_pos=0,
 final_scores=0.1384349, final_local_scores=0.120406866,
 pem=-0.1344273, similarity=8.413931|
 RECALL MISS: 192545 Eurovision_Song_Contest_1974]
 [winner: 34654 1974, pem_pos=0, final_scores=-0.5877854,
 final_local_scores=-0.4348294, pem=-0.70012724,
 similarity=-5.079063|
 RECALL MISS: 192545 Eurovision_Song_Contest_1974]
 11: 9534795 The_First_Day_of_Love
 [winner: 9534795 The_First_Day_of_Love, pem_pos=0,
 final_scores=-0.039057866, final_local_scores=0.02888307,
 pem=0.0, similarity=4.9516416|
 gt: 9534795 The_First_Day_of_Love, pem_pos=0,
 final_scores=-0.039057866, final_local_scores=0.02888307,
 pem=0.0, similarity=4.9516416]
 12: 2949125 Kaveret
 [winner: 2949125 Kaveret, pem_pos=0,
 final_scores=-0.27905744, final_local_scores=-0.21444423,
 pem=0.0, similarity=-2.3097308|
 gt: 2949125 Kaveret, pem_pos=0,
 final_scores=-0.27905744, final_local_scores=-0.21444423,
 pem=0.0, similarity=-2.3097308]
 13: 6805316 Natati_La_Khayay
 [winner: 6805316 Natati_La_Khayay, pem_pos=0,
 final_scores=-0.060226843, final_local_scores=0.012882233,
 pem=0.0, similarity=4.4741445|
 gt: 6805316 Natati_La_Khayay, pem_pos=0,
 final_scores=-0.060226843, final_local_scores=0.012882233,
 pem=0.0, similarity=4.4741445]
 14: 193442 Eurovision_Song_Contest_1976
 [winner: 9954 Eurovision_Song_Contest, pem_pos=5,
 final_scores=-0.2395028, final_local_scores=-0.301915,
 pem=-2.0794415, similarity=6.3882947|
 RECALL MISS: 193442 Eurovision_Song_Contest_1976]
 [winner: 34661 1976, pem_pos=0, final_scores=-1.1682216,
 final_local_scores=-1.1299438, pem=-0.7313423,
 similarity=-25.652914|
 RECALL MISS: 193442 Eurovision_Song_Contest_1976]
 15: 193441 Eurovision_Song_Contest_1975
 [winner: 9954 Eurovision_Song_Contest, pem_pos=5,
 final_scores=-0.32896888, final_local_scores=-0.39964306,
 pem=-2.0794415, similarity=3.4718935|
 RECALL MISS: 193441 Eurovision_Song_Contest_1975]
 [winner: 34752 1975, pem_pos=0, final_scores=-0.8513031,
 final_local_scores=-0.7677653, pem=-0.61872095,
 similarity=-15.457233|

RECALL MISS: 193441 Eurovision_Song_Contest_1975]
16: 10380 European_Broadcasting_Union
[winner: 10380 European_Broadcasting_Union, pem_pos=0,
final_scores=-0.102087855, final_local_scores=-0.10825488,
pem=0.0, similarity=0.8591713]
gt: 10380 European_Broadcasting_Union, pem_pos=0,
final_scores=-0.102087855, final_local_scores=-0.10825488,
pem=0.0, similarity=0.8591713]
17: 1958873 Turkish_invasion_of_Cyprus
[winner: 5593 Cyprus, pem_pos=0, final_scores=0.595156,
final_local_scores=0.57246006, pem=-0.3752463,
similarity=23.21371]

RECALL MISS: 1958873 Turkish_invasion_of_Cyprus]
[winner: 11014528 Byzantine-Sasanian_wars, pem_pos=0,
final_scores=-0.82658416, final_local_scores=-0.7948708,
pem=-0.31309178, similarity=-17.928175]

RECALL MISS: 1958873 Turkish_invasion_of_Cyprus]
1: 12108 Greece
[pem_pos=1, final_scores=0.7512086,
final_local_scores=0.73280966, pem=-2.3247876,
similarity=38.600773]
2: 49796 Brighton
[pem_pos=0, final_scores=0.3546589,
final_local_scores=0.40128398, pem=-0.61380595,
similarity=19.4028]
3: 9316 England
[pem_pos=0, final_scores=0.5594425,
final_local_scores=0.59996605, pem=-0.19053856,
similarity=23.030075]
4: 21241 Norway
[pem_pos=0, final_scores=0.71489275,
final_local_scores=0.7528036, pem=-0.22668563,
similarity=27.787626]
5: 25219107 Anne-Karine_Strøm
[pem_pos=0, final_scores=0.23937325,
final_local_scores=0.32646227, pem=0.0, similarity=13.831998]
6: 9282173 Israel
[pem_pos=0, final_scores=0.6115894,
final_local_scores=0.6306254, pem=-0.21279564,
similarity=24.06605]
7: 23670710 Mariza_Koch
[pem_pos=0, final_scores=0.24167077,
final_local_scores=0.33774263, pem=0.0, similarity=14.168627]
1: 6305002 Krasi_Thalassa_Kai_T'_Agori_Mou
[pem_pos=0, final_scores=0.033698663,
final_local_scores=0.07278679, pem=0.0, similarity=6.261816]
2: 12108 Greece
[pem_pos=1, final_scores=0.29358816,
final_local_scores=0.23292875, pem=-2.3247876,
similarity=23.683329]
3: 8569916 English_language
[pem_pos=0, final_scores=0.21995775,

```

final_local_scores=0.23381144, pem=-0.640266,
similarity=14.548979]
4: 9954 Eurovision_Song_Contest
    [pem_pos=0, final_scores=0.12871097,
final_local_scores=0.10978496, pem=-0.1344273,
similarity=8.0969515]
5: 12108 Greece
    [pem_pos=1, final_scores=0.45307356,
final_local_scores=0.40714228, pem=-2.3247876,
similarity=28.882208]
6: 12108 Greece
    [pem_pos=1, final_scores=0.72105145,
final_local_scores=0.6998675, pem=-2.3247876,
similarity=37.61771]
7: 12108 Greece
    [pem_pos=0, final_scores=0.8960094,
final_local_scores=0.89098245, pem=-0.27607638,
similarity=32.179752]
8: 9954 Eurovision_Song_Contest
    [pem_pos=0, final_scores=0.11598174,
final_local_scores=0.09588021, pem=-0.1344273,
similarity=7.6820064]
9: 12108 Greece
    [pem_pos=1, final_scores=0.7393689,
final_local_scores=0.7198765, pem=-2.3247876,
similarity=38.21482]
10: 9954 Eurovision_Song_Contest
    [pem_pos=0, final_scores=0.1384349,
final_local_scores=0.120406866, pem=-0.1344273,
similarity=8.413931]
11: 12108 Greece
    [pem_pos=1, final_scores=0.70265925,
final_local_scores=0.6797768, pem=-2.3247876,
similarity=37.018166]
12: 12108 Greece
    [pem_pos=0, final_scores=0.51883346,
final_local_scores=0.478975, pem=-0.27607638,
similarity=19.884628]
13: 5593 Cyprus
    [pem_pos=0, final_scores=0.595156,
final_local_scores=0.57246006, pem=-0.3752463,
similarity=23.21371]

```

Sample: b'Iemasa_Tokugawa'

1 Prince was a [1 1 Japan 1] ese political figure of the [2 2 Taishō 2] and early [1 2 Shōwa 1] period s . He was the 17th hereditary head of the former 3 shogunal branch of the [2 4 Tokugawa 2] clan and the final President of the [3 3 House of Peers 3] in the . [3 Tokugawa Iemasa 3] was born in what is now the [4 4 Sendagaya 4] district of [4 Tokyo 4] , as the eldest son of [5 5 Tokugawa Iesato 5] . He graduated from the law school of in 1909 , and accepted a post in the diplomatic

corps of 5 Foreign Ministry the same year . In 1924 , he was appointed 6 Consul-general to the [5 Japanese 5] consulate in [6 6 Sydney 6] , [7 7 Australia 7] . In 1929 , he was appointed 7 Envoy to [8 8 Canada 8] and from 1937-1939 served as the [6 Japanese 6] ambassador to 8 Turkey . In 1940 , on the death of his father , he inherited the title of kōshaku - LRB- [7 公爵 7] , prince -RRB- under the 9 kazoku peerage system , and a seat as a member of the 10 House of Peers of the . On 19 June 1946 , he served as the President of the [8 House of Peers 8] , a post which he held until 2 May 1947 , when the 11 Allied occupation authorities authorized the current [9 9 Constitution of Japan 9] abolishing the [9 House of Peers 9] . He died of at his home in [10 12 Shibuya 10] , [11 Tokyo 11] on February 18 , 1963 , and was posthumously awarded the 13 Order of the Rising Sun with Paulownia Flowers , 1st class . His grave is located at the [12 Yanaka Cemetery 12] in [13 Tokyo 13] . He was succeeded as head of the [14 Tokugawa 14] clan by 14 Tokugawa Tsunenari

- 1: 15573 Japan
- 2: 193353 Taishō_period
- 3: 30639531 House_of_Peers_(Japan)
- 4: 4160218 Sendagaya
- 5: 7743234 Tokugawa_Iesato
- 6: 27862 Sydney
- 7: 4689264 Australia
- 8: 5042916 Canada
- 9: 62699 Constitution_of_Japan
- 1: 51511 Prince

[winner: 57317 Prince_(musician), pem_pos=0, final_scores=-0.3907147, final_local_scores=-0.3183077, pem=-0.46589175, similarity=-2.8756318]
 gt: 51511 Prince, pem_pos=1, final_scores=-0.51645803, final_local_scores=-0.53453755, pem=-1.5037805, similarity=-3.684161]

2: 261240 Shōwa_period
 [winner: 261240 Shōwa_period, pem_pos=0, final_scores=0.21465443, final_local_scores=0.19989583, pem=-0.4586625, similarity=12.549281]
 gt: 261240 Shōwa_period, pem_pos=0, final_scores=0.21465443, final_local_scores=0.19989583, pem=-0.4586625, similarity=12.549281]

[winner: 86350 Period_(periodic_table), pem_pos=1, final_scores=-1.5076268, final_local_scores=-1.4947824, pem=-1.742594, similarity=-31.041077]
 RECALL MISS: 261240 Shōwa_period]

3: 84507 Tokugawa_shogunate
 [winner: 84507 Tokugawa_shogunate, pem_pos=0, final_scores=-1.0689232, final_local_scores=-1.2131624, pem=0.0, similarity=-32.113472]
 gt: 84507 Tokugawa_shogunate, pem_pos=0, final_scores=-1.0689232, final_local_scores=-1.2131624, pem=0.0, similarity=-32.113472]

[winner: 133634
Branch_Township,_Schuylkill_County,_Pennsylvania, pem_pos=2,
final_scores=-1.3461422, final_local_scores=-1.3168488,
pem=-2.2740836, similarity=-22.840862|
RECALL MISS: 84507 Tokugawa_shogunate]

4: 2019824 Tokugawa_clan
[winner: 7743234 Tokugawa_Iesato, pem_pos=2,
final_scores=0.2792477, final_local_scores=0.2898945,
pem=0.0, similarity=12.740743|
gt: 2019824 Tokugawa_clan, pem_pos=4,
final_scores=-0.054302722, final_local_scores=-0.05138558,
pem=-1.5366385, similarity=10.912744]
[winner: 3362578 Rajput_clans, pem_pos=9,
final_scores=-1.5964957, final_local_scores=-1.6015975,
pem=-2.6033757, similarity=-29.547592|
RECALL MISS: 2019824 Tokugawa_clan]

5: 868429 Ministry_of_Foreign_Affairs_(Japan)
[winner: 317153 Foreign_minister, pem_pos=0,
final_scores=-0.012200847, final_local_scores=-0.0013264716,
pem=-0.35490593, similarity=5.9801617|
gt: 868429 Ministry_of_Foreign_Affairs_(Japan),
pem_pos=1, final_scores=-0.24632846,
final_local_scores=-0.29323328, pem=-2.1448066,
similarity=7.0028396]

6: 3684583 Consul_(representative)
[winner: 3684583 Consul_(representative), pem_pos=0,
final_scores=0.025884554, final_local_scores=0.033245057,
pem=0.0, similarity=5.081812|
gt: 3684583 Consul_(representative), pem_pos=0,
final_scores=0.025884554, final_local_scores=0.033245057,
pem=0.0, similarity=5.081812]

7: 18002418 Envoy_(title)
[winner: 865314 GMC_Envoy, pem_pos=1,
final_scores=-0.361439, final_local_scores=-0.26112348,
pem=-0.80613405, similarity=0.6811499|
gt: 18002418 Envoy_(title), pem_pos=4,
final_scores=-0.4913281, final_local_scores=-0.5033592,
pem=-1.6598319, similarity=-1.9051068]

8: 11125639 Turkey
[winner: 11125639 Turkey, pem_pos=0,
final_scores=-0.46333882, final_local_scores=-0.47787002,
pem=-0.20984955, similarity=-9.029687|
gt: 11125639 Turkey, pem_pos=0,
final_scores=-0.46333882, final_local_scores=-0.47787002,
pem=-0.20984955, similarity=-9.029687]

9: 277405 Kazoku
[winner: 277405 Kazoku, pem_pos=0,
final_scores=-0.16936955, final_local_scores=-0.190162,
pem=0.0, similarity=-1.5851004|
gt: 277405 Kazoku, pem_pos=0, final_scores=-0.16936955,
final_local_scores=-0.190162, pem=0.0, similarity=-1.5851004]

10: 30639531 House_of_Peers_(Japan)

[winner: 30639531 House_of_Peers_(Japan), pem_pos=1,
final_scores=-0.6448419, final_local_scores=-0.75487876,
pem=-0.4318518, similarity=-16.088898]
gt: 30639531 House_of_Peers_(Japan), pem_pos=1,
final_scores=-0.6448419, final_local_scores=-0.75487876,
pem=-0.4318518, similarity=-16.088898]
11: 87995 Occupation_of_Japan
[winner: 2198844 Allies_of_World_War_II, pem_pos=0,
final_scores=0.014996633, final_local_scores=0.04209605,
pem=-0.47402114, similarity=7.9237437]
RECALL MISS: 87995 Occupation_of_Japan]
[winner: 4856563 Military_Intelligence_(Pakistan),
pem_pos=1, final_scores=-1.2479321,
final_local_scores=-1.2573254, pem=-1.0962858,
similarity=-27.469612]
RECALL MISS: 87995 Occupation_of_Japan]
[winner: 816214 Military_occupation, pem_pos=0,
final_scores=-1.7819921, final_local_scores=-1.9285173,
pem=-1.1190172, similarity=-47.375698]
gt: 87995 Occupation_of_Japan, pem_pos=3,
final_scores=-1.8839988, final_local_scores=-2.0885806,
pem=-2.7400806, similarity=-43.336716]
12: 147739 Shibuya,_Tokyo
[winner: 30057 Tokyo, pem_pos=0, final_scores=0.50223726,
final_local_scores=0.4908405, pem=-0.09283106,
similarity=19.242203]
RECALL MISS: 147739 Shibuya,_Tokyo]
[winner: 147739 Shibuya,_Tokyo, pem_pos=0,
final_scores=0.16316698, final_local_scores=0.18030137,
pem=0.0, similarity=9.470265]
gt: 147739 Shibuya,_Tokyo, pem_pos=0,
final_scores=0.16316698, final_local_scores=0.18030137,
pem=0.0, similarity=9.470265]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7702935, final_local_scores=-1.7407539,
pem=-4.6728287, similarity=-22.446285]
RECALL MISS: 147739 Shibuya,_Tokyo]
13: 1193966 Order_of_the_Rising_Sun
[winner: 1193966 Order_of_the_Rising_Sun, pem_pos=0,
final_scores=0.0039342493, final_local_scores=0.0066459626,
pem=-0.013818712, similarity=4.36319]
gt: 1193966 Order_of_the_Rising_Sun, pem_pos=0,
final_scores=0.0039342493, final_local_scores=0.0066459626,
pem=-0.013818712, similarity=4.36319]
14: 7749526 Tokugawa_Tsunenari
[winner: 7749526 Tokugawa_Tsunenari, pem_pos=0,
final_scores=-0.27926576, final_local_scores=-0.22356072,
pem=0.0, similarity=-2.5817847]
gt: 7749526 Tokugawa_Tsunenari, pem_pos=0,
final_scores=-0.27926576, final_local_scores=-0.22356072,
pem=0.0, similarity=-2.5817847]
1: 15573 Japan

[pem_pos=0, final_scores=0.68391347,
final_local_scores=0.6922548, pem=-0.13549899,
similarity=25.48484]
2: 193353 Taishō_period
[pem_pos=0, final_scores=0.20649652,
final_local_scores=0.1910581, pem=-0.5147257,
similarity=12.5904255]
3: 30639531 House_of_Peers_(Japan)
[pem_pos=1, final_scores=0.2176799,
final_local_scores=0.19466114, pem=-0.4318518,
similarity=12.247267]
4: 4160218 Sendagaya
[pem_pos=0, final_scores=0.095066175,
final_local_scores=0.17822787, pem=0.0, similarity=9.408388]
5: 7743234 Tokugawa_Iesato
[pem_pos=0, final_scores=0.29771268,
final_local_scores=0.30666935, pem=0.0, similarity=13.241338]
6: 27862 Sydney
[pem_pos=0, final_scores=0.47689945,
final_local_scores=0.52349675, pem=-0.1443489,
similarity=20.496893]
7: 4689264 Australia
[pem_pos=0, final_scores=0.67481244,
final_local_scores=0.73613155, pem=-0.17168477,
similarity=26.990995]
8: 5042916 Canada
[pem_pos=0, final_scores=0.5789512,
final_local_scores=0.64016664, pem=-0.11662793,
similarity=23.827803]
9: 62699 Constitution_of_Japan
[pem_pos=0, final_scores=0.06541653,
final_local_scores=0.06424305, pem=0.0, similarity=6.0068536]
1: 261240 Shōwa_period
[pem_pos=0, final_scores=0.21465443,
final_local_scores=0.19989583, pem=-0.4586625,
similarity=12.549281]
2: 7743234 Tokugawa_Iesato
[pem_pos=2, final_scores=0.2792477,
final_local_scores=0.2898945, pem=0.0, similarity=12.740743]
3: 7749437 Iemasa_Tokugawa
[pem_pos=0, final_scores=0.1800449,
final_local_scores=0.19762486, pem=0.0, similarity=9.987233]
4: 30057 Tokyo
[pem_pos=0, final_scores=0.5213101,
final_local_scores=0.5116747, pem=-0.09283106,
similarity=19.863935]
5: 15573 Japan
[pem_pos=1, final_scores=0.83019894,
final_local_scores=0.8574595, pem=-0.99076414,
similarity=35.065945]
6: 15573 Japan
[pem_pos=0, final_scores=0.7017668,

final_local_scores=0.71716684, pem=-0.99076414,
 similarity=30.879332]
 7: 233815 Andrew_Cavendish,_11th_Duke_of_Devonshire
 [pem_pos=0, final_scores=0.17443965,
 final_local_scores=0.25835603, pem=0.0, similarity=11.799572]
 8: 30639531 House_of_Peers_(Japan)
 [pem_pos=1, final_scores=0.1282144,
 final_local_scores=0.09693369, pem=-0.4318518,
 similarity=9.330885]
 9: 30639531 House_of_Peers_(Japan)
 [pem_pos=1, final_scores=0.13801293,
 final_local_scores=0.10763708, pem=-0.4318518,
 similarity=9.650295]
 10: 147739 Shibuya,_Tokyo
 [pem_pos=0, final_scores=0.16316698,
 final_local_scores=0.18030137, pem=0.0, similarity=9.470265]
 11: 30057 Tokyo
 [pem_pos=0, final_scores=0.50223726,
 final_local_scores=0.4908405, pem=-0.09283106,
 similarity=19.242203]
 12: 15530508 Yanaka_Cemetery
 [pem_pos=0, final_scores=0.030689314,
 final_local_scores=0.10264915, pem=0.0, similarity=7.1529684]
 13: 30057 Tokyo
 [pem_pos=0, final_scores=0.47103083,
 final_local_scores=0.45675224, pem=-0.09283106,
 similarity=18.224941]
 14: 7743234 Tokugawa_Iesato
 [pem_pos=2, final_scores=0.3027355,
 final_local_scores=0.3155514, pem=0.0, similarity=13.506396]

Sample: b'Frank_A'

__ Moore [1 Frank A. Moore 1] -LRB- November 5 , 1844 --
 September 25 , 1918 -RRB- was an [1 1 American 1] politician
 and judge in the state of [2 Oregon 2] . He was the 17th 1 Ch
 ief Justice of the [2 2 Oregon Supreme Court 2] . He is the
 only person to serve as chief justice for [3 3 Oregon 3] 's
 highest court on four different occasions , and he spent 26
 years overall on the bench . A native of [4 4 Maine 4] , he
 was also a two-time [3 Republican 3] member of the [5 5 Oregon
 State Senate 5] . Moore was born on November 5 , 1844 , in [4
 2 Ellsworth 4] , [5 Maine 5] , to Heard L. Moore and
 Bathshaba Moore -LRB- née Higgins -RRB- . [6 Moore 6]
 received his education in the public schools in [6 6 Maine 6]
 and at the state normal school in [7 3 Iowa Falls 7] , [8 Iowa
 8] . After his education [9 Moore 9] became a county
 superintendent for schools in [10 4 Hardin County 10] ,
 Iowa , serving from 1871 to 1875 . He also read law in [11 5 E
 ldora 11] , [12 Iowa 12] , at this time under the tutelage of
 [7 7 Enoch W. Eastman 7] , the then lieutenant governor of [13
 Iowa 13] . [14 Moore 14] passed the 6 bar in [15 Iowa 15] in
 1874 , and practiced law in [16 Eldora 16] . In 1877 , Moore

moved to [17 Oregon 17] and was practicing law in [18 7 St. He
lens 18] , [19 Oregon 19] , and was admitted to the [20
Oregon 20] bar in January 1879 . He remained in private
practice there until 1884 when he was elected as judge for [21
8 Columbia County 21] . In 1888 , [22 Moore 22] was elected
as a state senator to the [23 Oregon 23] legislature . During
the 1889 legislature he represented 9 Columbia , [24 10 Tilla
mook 24] , and [25 11 Washington 25] counties . He won re-
election in 1890 to the [26 Moore 26] was elected to the [8 8
Oregon Supreme Court 8] in 1892 . He won re-election to
additional six-year terms in 1898 , 1904 , 1910 , and 1916 . [27
Moore 27] then served as chief justice four times : 1896
to 1898 , 1902 to 1905 , 1909 to 1911 , and 1915 to 1917 . [28
Moore 28] died while in office on September 25 , 1918 , and
was replaced by [9 9 Conrad P. Olson 9] . Frank A. Moore was
married to Emma Shuntaffer on April 15 , 1866 . They had
three children together . Moore was buried at [10 10 City View
Cemetery 10] in [29 12 Salem 29] , [30 Oregon 30] . [31 Moore
31] was a [32 Republican 32] . He was a Mason and a member of
the 13 Benevolent and Protective Order of Elks

1: 3434750 United_States

2: 527090 Oregon_Supreme_Court

3: 26811621 Oregon

4: 19977 Maine

5: 5423153 Oregon_State_Senate

6: 19977 Maine

7: 36676886 Enoch_W._Eastman

8: 527090 Oregon_Supreme_Court

9: 11639456 Conrad_P._Olson

10: 33316737 City_View_Cemetery

1: 356035 Chief_Justice

[winner: 356035 Chief_Justice, pem_pos=0,
final_scores=-0.12784204, final_local_scores=-0.12858407,
pem=-1.0018724, similarity=5.7008457]

gt: 356035 Chief_Justice, pem_pos=0,
final_scores=-0.12784204, final_local_scores=-0.12858407,
pem=-1.0018724, similarity=5.7008457]

2: 115984 Ellsworth,_Maine

[winner: 19977 Maine, pem_pos=0, final_scores=0.2358775,
final_local_scores=0.1910018, pem=-0.2749413,
similarity=11.284761]

RECALL MISS: 115984 Ellsworth,_Maine]

[winner: 115984 Ellsworth,_Maine, pem_pos=1,
final_scores=0.22202943, final_local_scores=0.23200649,
pem=-1.8791751, similarity=21.232494]

gt: 115984 Ellsworth,_Maine, pem_pos=1,
final_scores=0.22202943, final_local_scores=0.23200649,
pem=-1.8791751, similarity=21.232494]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.6555665, final_local_scores=-1.6504316,
pem=-4.6728287, similarity=-19.75089]

RECALL MISS: 115984 Ellsworth,_Maine]

3: 113848 Iowa_Falls,_Iowa
 [winner: 113848 Iowa_Falls,_Iowa, pem_pos=0,
 final_scores=0.21902479, final_local_scores=0.24783447,
 pem=0.0, similarity=11.485588|
 gt: 113848 Iowa_Falls,_Iowa, pem_pos=0,
 final_scores=0.21902479, final_local_scores=0.24783447,
 pem=0.0, similarity=11.485588]
 [winner: 26810748 Iowa, pem_pos=1,
 final_scores=0.10694672, final_local_scores=0.043895975,
 pem=-0.3546882, similarity=7.328506|
 RECALL MISS: 113848 Iowa_Falls,_Iowa]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7887056, final_local_scores=-1.795866,
 pem=-4.6728287, similarity=-24.090942|
 RECALL MISS: 113848 Iowa_Falls,_Iowa]

4: 95995 Hardin_County,_Iowa
 [winner: 93112 Hardin_County,_Ohio, pem_pos=2,
 final_scores=0.15427594, final_local_scores=0.17320862,
 pem=-0.947461, similarity=14.411044|
 gt: 95995 Hardin_County,_Iowa, pem_pos=5,
 final_scores=-0.12910166, final_local_scores=-0.12815294,
 pem=-2.0624485, similarity=11.48129]
 [winner: 26810748 Iowa, pem_pos=1,
 final_scores=-0.081468254, final_local_scores=-0.16377437,
 pem=-0.3546882, similarity=1.1312087|
 RECALL MISS: 95995 Hardin_County,_Iowa]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7857752, final_local_scores=-1.792665,
 pem=-4.6728287, similarity=-23.995417|
 RECALL MISS: 95995 Hardin_County,_Iowa]

5: 113846 Eldora,_Iowa
 [winner: 113846 Eldora,_Iowa, pem_pos=0,
 final_scores=0.19362842, final_local_scores=0.225258,
 pem=-0.8078988, similarity=15.20534|
 gt: 113846 Eldora,_Iowa, pem_pos=0,
 final_scores=0.19362842, final_local_scores=0.225258,
 pem=-0.8078988, similarity=15.20534]
 [winner: 26810748 Iowa, pem_pos=1,
 final_scores=0.13826986, final_local_scores=0.07811175,
 pem=-0.3546882, similarity=8.349573|
 RECALL MISS: 113846 Eldora,_Iowa]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.6832399, final_local_scores=-1.6806607,
 pem=-4.6728287, similarity=-20.652983|
 RECALL MISS: 113846 Eldora,_Iowa]

6: 268188 Bar_(law)
 [winner: 4098163 State_bar_association, pem_pos=5,
 final_scores=-1.4234469, final_local_scores=-1.5198798,
 pem=-1.9226918, similarity=-30.810633|
 gt: 268188 Bar_(law), pem_pos=1,
 final_scores=-1.5308024, final_local_scores=-1.6955161,
 pem=-1.0096852, similarity=-41.017044]

7: 130711 St._Helens,_Oregon
[winner: 26811621 Oregon, pem_pos=1,
final_scores=0.58603233, final_local_scores=0.58477455,
pem=-0.28686622, similarity=23.100574|
RECALL MISS: 130711 St._Helens,_Oregon]
[winner: 130711 St._Helens,_Oregon, pem_pos=2,
final_scores=0.06023918, final_local_scores=0.063004375,
pem=-1.6306924, similarity=14.837848|
gt: 130711 St._Helens,_Oregon, pem_pos=2,
final_scores=0.06023918, final_local_scores=0.063004375,
pem=-1.6306924, similarity=14.837848]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.913497, final_local_scores=-1.9321816,
pem=-4.6728287, similarity=-28.15887|
RECALL MISS: 130711 St._Helens,_Oregon]

8: 92033 Columbia_County,_Oregon
[winner: 96799 Columbia_County,_Georgia, pem_pos=3,
final_scores=0.09371032, final_local_scores=0.06578061,
pem=-1.501265, similarity=14.21685|
gt: 92033 Columbia_County,_Oregon, pem_pos=5,
final_scores=-0.045913264, final_local_scores=-0.064434886,
pem=-1.5332949, similarity=10.505143]

9: 92033 Columbia_County,_Oregon
[winner: 57851 Columbia,_South_Carolina, pem_pos=2,
final_scores=-0.044384763, final_local_scores=-0.09801042,
pem=-2.4337058, similarity=14.399757|
RECALL MISS: 92033 Columbia_County,_Oregon]

10: 91996 Tillamook_County,_Oregon
[winner: 130953 Tillamook,_Oregon, pem_pos=0,
final_scores=0.13141964, final_local_scores=0.14375788,
pem=-0.595012, similarity=11.615502|
gt: 91996 Tillamook_County,_Oregon, pem_pos=1,
final_scores=-0.04712467, final_local_scores=-0.05374469,
pem=-1.2616086, similarity=9.346688]

11: 48966 Washington_County,_Oregon
[winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=0.34040576, final_local_scores=0.31952846,
pem=-1.7496481, similarity=23.139938|
gt: 48966 Washington_County,_Oregon, pem_pos=11,
final_scores=-0.52608055, final_local_scores=-0.62003237,
pem=-4.2138076, similarity=8.50205]

12: 48970 Salem,_Oregon
[winner: 26811621 Oregon, pem_pos=1,
final_scores=0.5767651, final_local_scores=0.57465154,
pem=-0.28686622, similarity=22.798483|
RECALL MISS: 48970 Salem,_Oregon]
[winner: 48970 Salem,_Oregon, pem_pos=0,
final_scores=0.16252114, final_local_scores=0.11024511,
pem=-1.8624039, similarity=17.507689|
gt: 48970 Salem,_Oregon, pem_pos=0,
final_scores=0.16252114, final_local_scores=0.11024511,
pem=-1.8624039, similarity=17.507689]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.781868, final_local_scores=-1.7883968,
 pem=-4.6728287, similarity=-23.868044]
 RECALL MISS: 48970 Salem, Oregon]

13: 3087337 Benevolent_and_Protective_Order_of_Elks
 [winner: 3087337 Benevolent_and_Protective_Order_of_Elks,
 pem_pos=0, final_scores=-0.114778236,
 final_local_scores=-0.13205369, pem=0.0,
 similarity=0.14896755]
 gt: 3087337 Benevolent_and_Protective_Order_of_Elks,
 pem_pos=0, final_scores=-0.114778236,
 final_local_scores=-0.13205369, pem=0.0,
 similarity=0.14896755]

1: 3434750 United States
 [pem_pos=0, final_scores=0.47373754,
 final_local_scores=0.48830473, pem=-0.12994486,
 similarity=19.36836]

2: 527090 Oregon_Supreme_Court
 [pem_pos=0, final_scores=0.24025758,
 final_local_scores=0.23293793, pem=0.0, similarity=11.041045]

3: 26811621 Oregon
 [pem_pos=1, final_scores=0.5471683,
 final_local_scores=0.54232156, pem=-0.28686622,
 similarity=21.833693]

4: 19977 Maine
 [pem_pos=0, final_scores=0.30229956,
 final_local_scores=0.26355776, pem=-0.2749413,
 similarity=13.449976]

5: 5423153 Oregon_State_Senate
 [pem_pos=0, final_scores=0.14983736,
 final_local_scores=0.12713969, pem=0.0, similarity=7.883815]

6: 19977 Maine
 [pem_pos=0, final_scores=0.16951121,
 final_local_scores=0.11850673, pem=-0.2749413,
 similarity=9.121364]

7: 36676886 Enoch_W._Eastman
 [pem_pos=0, final_scores=0.03960766,
 final_local_scores=0.12313709, pem=0.0, similarity=7.7643695]

8: 527090 Oregon_Supreme_Court
 [pem_pos=1, final_scores=0.1990047,
 final_local_scores=0.18906248, pem=0.0, similarity=9.731714]

9: 11639456 Conrad_P._Olson
 [pem_pos=0, final_scores=0.08032386,
 final_local_scores=0.11692211, pem=0.0, similarity=7.578902]

10: 33316737 City_View_Cemetery
 [pem_pos=0, final_scores=0.1941713,
 final_local_scores=0.26430917, pem=0.0, similarity=11.977225]

1: 11163050 Frank_A._Moore
 [pem_pos=0, final_scores=0.26889175,
 final_local_scores=0.28845274, pem=0.0, similarity=12.697719]

2: 26811621 Oregon
 [pem_pos=1, final_scores=0.50302494,

final_local_scores=0.49410164, pem=-0.28686622,
similarity=20.394714]
3: 32070 Republican_Party_(United_States)
[pem_pos=1, final_scores=0.41796035,
final_local_scores=0.403041, pem=-0.1310191,
similarity=16.829763]
4: 115984 Ellsworth,_Maine
[pem_pos=1, final_scores=0.22202943,
final_local_scores=0.23200649, pem=-1.8791751,
similarity=21.232494]
5: 19977 Maine
[pem_pos=0, final_scores=0.2358775,
final_local_scores=0.1910018, pem=-0.2749413,
similarity=11.284761]
6: 442312 A._Harry_Moore
[pem_pos=4, final_scores=0.03089805,
final_local_scores=0.019475713, pem=-1.0986123,
similarity=10.645331]
7: 113848 Iowa_Falls,_Iowa
[pem_pos=0, final_scores=0.21902479,
final_local_scores=0.24783447, pem=0.0, similarity=11.485588]
8: 26810748 Iowa
[pem_pos=1, final_scores=0.10694672,
final_local_scores=0.043895975, pem=-0.3546882,
similarity=7.328506]
9: 442312 A._Harry_Moore
[pem_pos=4, final_scores=0.124758914,
final_local_scores=0.12239367, pem=-1.0986123,
similarity=13.716608]
10: 93112 Hardin_County,_Ohio
[pem_pos=2, final_scores=0.15427594,
final_local_scores=0.17320862, pem=-0.947461,
similarity=14.411044]
11: 113846 Eldora,_Iowa
[pem_pos=0, final_scores=0.19362842,
final_local_scores=0.225258, pem=-0.8078988,
similarity=15.20534]
12: 26810748 Iowa
[pem_pos=1, final_scores=0.13826986,
final_local_scores=0.07811175, pem=-0.3546882,
similarity=8.349573]
13: 26810748 Iowa
[pem_pos=0, final_scores=0.09994255,
final_local_scores=0.036398083, pem=-0.3546882,
similarity=7.104754]
14: 442312 A._Harry_Moore
[pem_pos=4, final_scores=0.23430987,
final_local_scores=0.24206144, pem=-1.0986123,
similarity=17.287733]
15: 26810748 Iowa
[pem_pos=1, final_scores=0.074051306,
final_local_scores=0.0079627335, pem=-0.3546882,

similarity=6.2561865]
16: 113846 Eldora, Iowa
[pem_pos=0, final_scores=0.14967988,
final_local_scores=0.17725089, pem=-0.8078988,
similarity=13.772713]
17: 26811621 Oregon
[pem_pos=1, final_scores=0.511304,
final_local_scores=0.5031453, pem=-0.28686622,
similarity=20.664595]
18: 130711 St. Helens, Oregon
[pem_pos=2, final_scores=0.06023918,
final_local_scores=0.063004375, pem=-1.6306924,
similarity=14.837848]
19: 26811621 Oregon
[pem_pos=1, final_scores=0.58603233,
final_local_scores=0.58477455, pem=-0.28686622,
similarity=23.100574]
20: 26811621 Oregon
[pem_pos=1, final_scores=0.46827012,
final_local_scores=0.4570526, pem=-0.28686622,
similarity=19.289097]
21: 96799 Columbia County, Georgia
[pem_pos=3, final_scores=0.09371032,
final_local_scores=0.06578061, pem=-1.501265,
similarity=14.21685]
22: 442312 A. Harry Moore
[pem_pos=4, final_scores=0.08853392,
final_local_scores=0.08247428, pem=-1.0986123,
similarity=12.525334]
23: 26811621 Oregon
[pem_pos=1, final_scores=0.5530475,
final_local_scores=0.54965895, pem=-0.28686622,
similarity=22.052654]
24: 130953 Tillamook, Oregon
[pem_pos=0, final_scores=0.13141964,
final_local_scores=0.14375788, pem=-0.595012,
similarity=11.615502]
25: 108956 Washington, D.C.
[pem_pos=0, final_scores=0.34040576,
final_local_scores=0.31952846, pem=-1.7496481,
similarity=23.139938]
26: 442312 A. Harry Moore
[pem_pos=4, final_scores=0.15420876,
final_local_scores=0.15417391, pem=-1.0986123,
similarity=14.664995]
27: 442312 A. Harry Moore
[pem_pos=4, final_scores=0.07515718,
final_local_scores=0.06780268, pem=-1.0986123,
similarity=12.087504]
28: 442312 A. Harry Moore
[pem_pos=4, final_scores=0.12545879,
final_local_scores=0.12276894, pem=-1.0986123,

```

similarity=13.727808]
29: 48970 Salem,_Oregon
    [pem_pos=0, final_scores=0.16252114,
final_local_scores=0.11024511, pem=-1.8624039,
similarity=17.507689]
30: 26811621 Oregon
    [pem_pos=1, final_scores=0.5767651,
final_local_scores=0.57465154, pem=-0.28686622,
similarity=22.798483]
31: 442312 A._Harry_Moore
    [pem_pos=4, final_scores=0.063605055,
final_local_scores=0.055203155, pem=-1.0986123,
similarity=11.71151]
32: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.45936745,
final_local_scores=0.4469748, pem=-0.1310191,
similarity=18.140835]

```

Sample: b'Genesis_Group'

The Genesis Group is a of [1 1 The Church of Jesus Christ of Latter-day Saints 1] for [2 2 African American 2] members and their families . It was first organized in [3 3 Salt Lake City 3] , [4 4 Utah 4] in 1971 to provide members an organization where they could affiliate with fellow [1 African American 1] members . The group was led by Ruffin Bridgeforth from 1971 through 1978 . Shortly after the church 's June 8 , 1978 announcement of the 1 revelation extending the 2 priesthood to all worthy male members of the church , the group was dissolved . The Genesis Group was reorganized in the 1990s , based on a perception that [2 African Americans 2] still had unique issues and could benefit from a chance to affiliate with one another , especially since many were the only members of [5 5 Africa 5] n descent in their local 3 wards and even in their 4 stakes . Leaders of the group include 5 Da rius Gray -LRB- 1997 -- 2003 -RRB- and Don Harwell -LRB- 2003 -- present -RRB- . 6 Seventy President [6 6 Ronald A. Rasband 6] is currently the [3 LDS 3] 7 general authority responsible for overseeing the group . Whereas when the group was first organized it was a potential resource at least in theory for issues relating to black members throughout the [4 United States 4] , under its current existence it is under the [5 Utah Salt Lake City Area 5] with designation to be a resource throughout all three 8 Areas in [6 Utah 6] , a multi-area system facilitated in part by having one man preside over all three areas . Genesis Groups formed in other parts of the [7 United States 7] would either be under the local stake or more likely have either the member of the Presidency of the Seventy assigned to that area or an as the priesthood advisor over the group . This would be roughly similar to how members of an [8 LDS 8] interact with the wards or branches they are assigned to . Other Genesis groups have existed in [9 9 Washi ngton 9] , [10 D.C. 10] and presently exist in [11 10 Hatties

burg 11] , [12 Mississippi 12] , [7 7 Cincinnati 7] and [13 11 Columbus 13] , [14 Ohio 14] , [8 8 Los Angeles 8] and [15 12 Oakland 15] , [16 California 16] , [9 9 Houston 9] , [10 10 Texas 10] and [17 13 Rochester 17] , [18 New York 18]

1: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints

2: 2154 African_American

3: 53837 Salt_Lake_City

4: 31716 Utah

5: 5334607 Africa

6: 12191584 Ronald_A._Rasband

7: 18522615 Cincinnati

8: 18110 Los_Angeles

9: 13774 Houston

10: 29810 Texas

1: 10242200 Revelation_(Latter_Day_Saints)

[winner: 2987672 The_Revelation_(Little_novel),
pem_pos=1, final_scores=-0.87912196,
final_local_scores=-0.80816084, pem=-0.5288441,
similarity=-17.151482|

RECALL MISS: 10242200 Revelation_(Latter_Day_Saints)]

2: 50405 Clergy

[winner: 72346 Kohen, pem_pos=0, final_scores=-1.1922754,
final_local_scores=-1.2077708, pem=-0.037740305,
similarity=-31.747343|

RECALL MISS: 50405 Clergy]

3: 436529 Ward_(LDS_Church)

[winner: 27943750 Ward_(Vietnam), pem_pos=7,
final_scores=-1.3439504, final_local_scores=-1.266757,
pem=-2.1428587, similarity=-22.059645|

gt: 436529 Ward_(LDS_Church), pem_pos=3,
final_scores=-1.7070186, final_local_scores=-1.878299,
pem=-1.7102716, similarity=-42.66175]

4: 436512 Stake_(Latter_Day_Saints)

[winner: 436512 Stake_(Latter_Day_Saints), pem_pos=0,
final_scores=-1.1792674, final_local_scores=-1.3102036,
pem=-0.04044906, similarity=-34.78941|

gt: 436512 Stake_(Latter_Day_Saints), pem_pos=0,
final_scores=-1.1792674, final_local_scores=-1.3102036,
pem=-0.04044906, similarity=-34.78941]

5: 16998994 Darius_Gray

[winner: 16998994 Darius_Gray, pem_pos=0,
final_scores=-0.13116339, final_local_scores=-0.0973091,
pem=0.0, similarity=1.1858153|

gt: 16998994 Darius_Gray, pem_pos=0,
final_scores=-0.13116339, final_local_scores=-0.0973091,
pem=0.0, similarity=1.1858153]

6: 532256 Seventy_(Latter_Day_Saints)

[winner: 532256 Seventy_(Latter_Day_Saints), pem_pos=0,
final_scores=-0.24352358, final_local_scores=-0.22206813,
pem=0.0, similarity=-2.537243|

gt: 532256 Seventy_(Latter_Day_Saints), pem_pos=0,
final_scores=-0.24352358, final_local_scores=-0.22206813,

pem=0.0, similarity=-2.537243]
7: 657147 General_authority
[winner: 3677766 Auctoritas, pem_pos=3,
final_scores=-1.1502322, final_local_scores=-1.1414676,
pem=-1.0978392, similarity=-24.003733]
RECALL MISS: 657147 General_authority]
[winner: 148537 General_officer, pem_pos=0,
final_scores=-1.4748405, final_local_scores=-1.5682297,
pem=-0.36330885, similarity=-40.73366]
RECALL MISS: 657147 General_authority]
8: 12054839 Area_(LDS_Church)
[winner: 9602 Edinburgh, pem_pos=2,
final_scores=-0.0450584, final_local_scores=-0.08521102,
pem=-1.2325603, similarity=8.249701]
gt: 12054839 Area_(LDS_Church), pem_pos=0,
final_scores=-0.19279854, final_local_scores=-0.20535058,
pem=-0.6392334, similarity=1.4378924]
9: 108956 Washington,_D.C.
[winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=0.4435467, final_local_scores=0.39901847,
pem=-1.7496481, similarity=25.512077]
gt: 108956 Washington,_D.C., pem_pos=0,
final_scores=0.4435467, final_local_scores=0.39901847,
pem=-1.7496481, similarity=25.512077]
[winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=0.08515708, final_local_scores=0.0075323135,
pem=-0.15729709, similarity=5.169898]
gt: 108956 Washington,_D.C., pem_pos=0,
final_scores=0.08515708, final_local_scores=0.0075323135,
pem=-0.15729709, similarity=5.169898]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7986805, final_local_scores=-1.7807593,
pem=-4.6728287, similarity=-23.640125]
RECALL MISS: 108956 Washington,_D.C.]
10: 129294 Hattiesburg,_Mississippi
[winner: 129294 Hattiesburg,_Mississippi, pem_pos=0,
final_scores=0.34935087, final_local_scores=0.3416877,
pem=0.0, similarity=14.286356]
gt: 129294 Hattiesburg,_Mississippi, pem_pos=0,
final_scores=0.34935087, final_local_scores=0.3416877,
pem=0.0, similarity=14.286356]
[winner: 16949861 Mississippi, pem_pos=0,
final_scores=0.3088221, final_local_scores=0.27214694,
pem=-0.31944752, similarity=13.948326]
RECALL MISS: 129294 Hattiesburg,_Mississippi]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.755578, final_local_scores=-1.7336762,
pem=-4.6728287, similarity=-22.235073]
RECALL MISS: 129294 Hattiesburg,_Mississippi]
11: 5950 Columbus,_Ohio
[winner: 22199 Ohio, pem_pos=0, final_scores=0.5152733,
final_local_scores=0.48518825, pem=-0.19147073,

```

similarity=19.609945|
  RECALL MISS: 5950 Columbus,_Ohio]
  [winner: 5950 Columbus,_Ohio, pem_pos=0,
final_scores=0.23149745, final_local_scores=0.17653078,
pem=-1.0754056, similarity=15.205967]
  gt: 5950 Columbus,_Ohio, pem_pos=0,
final_scores=0.23149745, final_local_scores=0.17653078,
pem=-1.0754056, similarity=15.205967 ]
  [winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7864933, final_local_scores=-1.7674465,
pem=-4.6728287, similarity=-23.24285]
  RECALL MISS: 5950 Columbus,_Ohio]
12: 50548 Oakland,_California
  [winner: 5407 California, pem_pos=0,
final_scores=0.4581136, final_local_scores=0.42668933,
pem=-0.08662026, similarity=17.29403]
  RECALL MISS: 50548 Oakland,_California]
  [winner: 50548 Oakland,_California, pem_pos=0,
final_scores=0.16057123, final_local_scores=0.1089797,
pem=-0.4997394, similarity=10.0595455]
  gt: 50548 Oakland,_California, pem_pos=0,
final_scores=0.16057123, final_local_scores=0.1089797,
pem=-0.4997394, similarity=10.0595455 ]
  [winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8900506, final_local_scores=-1.8805673,
pem=-4.6728287, similarity=-26.618597]
  RECALL MISS: 50548 Oakland,_California]
13: 126641 Rochester,_New_York
  [winner: 126641 Rochester,_New_York, pem_pos=0,
final_scores=0.26630205, final_local_scores=0.22230381,
pem=-1.0925732, similarity=16.665285]
  gt: 126641 Rochester,_New_York, pem_pos=0,
final_scores=0.26630205, final_local_scores=0.22230381,
pem=-1.0925732, similarity=16.665285 ]
  [winner: 8210131 New_York, pem_pos=0,
final_scores=0.106303915, final_local_scores=0.034306556,
pem=-0.2889045, similarity=6.6845965]
  RECALL MISS: 126641 Rochester,_New_York]
  [winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.6820481, final_local_scores=-1.6533561,
pem=-4.6728287, similarity=-19.838158]
  RECALL MISS: 126641 Rochester,_New_York]
1: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints
  [pem_pos=0, final_scores=0.23611547,
final_local_scores=0.2022711, pem=0.0, similarity=10.125886]
2: 2154 African_American
  [pem_pos=1, final_scores=0.36284924,
final_local_scores=0.32268173, pem=-0.7252736,
similarity=17.66333]
3: 53837 Salt_Lake_City
  [pem_pos=0, final_scores=0.52987456,
final_local_scores=0.5076029, pem=-0.03391632,

```

similarity=19.422037]
4: 31716 Utah
[pem_pos=0, final_scores=0.45713627,
final_local_scores=0.43511158, pem=-0.2611261,
similarity=18.494354]
5: 5334607 Africa
[pem_pos=0, final_scores=0.3899147,
final_local_scores=0.4363945, pem=-0.13197383,
similarity=17.83029]
6: 12191584 Ronald_A._Rasband
[pem_pos=0, final_scores=0.14525978,
final_local_scores=0.1975382, pem=0.0, similarity=9.984647]
7: 18522615 Cincinnati
[pem_pos=0, final_scores=0.35642475,
final_local_scores=0.3095276, pem=-0.46106797,
similarity=15.833993]
8: 18110 Los_Angeles
[pem_pos=0, final_scores=0.30057973,
final_local_scores=0.26055896, pem=-0.100652285,
similarity=12.412675]
9: 13774 Houston
[pem_pos=0, final_scores=0.3867584,
final_local_scores=0.34644222, pem=-0.3232624,
similarity=16.186192]
10: 29810 Texas
[pem_pos=0, final_scores=0.6260158,
final_local_scores=0.6137839, pem=-0.16191302,
similarity=23.286757]
1: 2154 African_American
[pem_pos=1, final_scores=0.40248197,
final_local_scores=0.36618918, pem=-0.7252736,
similarity=18.96168]
2: 2154 African_American
[pem_pos=0, final_scores=0.344755,
final_local_scores=0.30291653, pem=0.0, similarity=13.129346]
3: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints
[pem_pos=0, final_scores=0.19868587,
final_local_scores=0.16138494, pem=-0.3500127,
similarity=10.809185]
4: 3434750 United_States
[pem_pos=0, final_scores=0.5757217,
final_local_scores=0.54772395, pem=-0.073457584,
similarity=20.834362]
5: 12054839 Area_(LDS_Church)
[pem_pos=0, final_scores=0.08380006,
final_local_scores=0.10527946, pem=0.0, similarity=7.231462]
6: 31716 Utah
[pem_pos=0, final_scores=0.3432153,
final_local_scores=0.31067023, pem=-0.2611261,
similarity=14.780776]
7: 3434750 United_States
[pem_pos=0, final_scores=0.49490488,

final_local_scores=0.45944387, pem=-0.073457584,
 similarity=18.19991]
 8: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints
 [pem_pos=0, final_scores=0.18767224,
 final_local_scores=0.14935425, pem=-0.3500127,
 similarity=10.450165]
 9: 108956 Washington,_D.C.
 [pem_pos=0, final_scores=0.4435467,
 final_local_scores=0.39901847, pem=-1.7496481,
 similarity=25.512077]
 10: 108956 Washington,_D.C.
 [pem_pos=0, final_scores=0.08515708,
 final_local_scores=0.0075323135, pem=-0.15729709,
 similarity=5.169898]
 11: 129294 Hattiesburg,_Mississippi
 [pem_pos=0, final_scores=0.34935087,
 final_local_scores=0.3416877, pem=0.0, similarity=14.286356]
 12: 16949861 Mississippi
 [pem_pos=0, final_scores=0.3088221,
 final_local_scores=0.27214694, pem=-0.31944752,
 similarity=13.948326]
 13: 5950 Columbus,_Ohio
 [pem_pos=0, final_scores=0.23149745,
 final_local_scores=0.17653078, pem=-1.0754056,
 similarity=15.205967]
 14: 22199 Ohio
 [pem_pos=0, final_scores=0.5152733,
 final_local_scores=0.48518825, pem=-0.19147073,
 similarity=19.609945]
 15: 50548 Oakland,_California
 [pem_pos=0, final_scores=0.16057123,
 final_local_scores=0.1089797, pem=-0.4997394,
 similarity=10.0595455]
 16: 5407 California
 [pem_pos=0, final_scores=0.4581136,
 final_local_scores=0.42668933, pem=-0.08662026,
 similarity=17.29403]
 17: 126641 Rochester,_New_York
 [pem_pos=0, final_scores=0.26630205,
 final_local_scores=0.22230381, pem=-1.0925732,
 similarity=16.665285]
 18: 8210131 New_York
 [pem_pos=0, final_scores=0.106303915,
 final_local_scores=0.034306556, pem=-0.2889045,
 similarity=6.6845965]

Sample: b'George_Waddell'

George Barr Waddell -LRB- 29 November 1888 -- 17 September
 1966 -RRB- was a [1 Scottish 1] professional 1 footballer who
 played in the [1 1 Scottish Football League 1] for [2 2 Range
 rs 2] and in 2 the [2 Football League 2] in [3 England 3] for
 [3 3 Bradford City 3] , [4 4 Preston North End 4] , [5 5 Oldh

am Athletic 5] , [6 6 Birmingham 6] and [7 7 New Brighton 7] . He played as a 3 wing half . Waddell was born in [8 8 Lesmahagow 8] , [9 9 Lanarkshire 9] . He began his football career in 4 junior football before joining [10 10 Rangers 10] in 1909 . He made his debut in on 26 April 1909 in a 3 - - 2 defeat to 5 Queen 's Park . Over the next four seasons he played 21 league games for [4 Rangers 4] , and spent time at [11 11 Kilburn 11] . In June 1914 , [12 12 Bradford City 12] paid # 1,000 for Waddell 's services . He played 23 [13 13 First Division 13] games either side of the before joining [14 14 Preston North End 14] for a fee of # 1,750 in September 1920 . After two seasons and 51 [5 First Division 5] games he had brief spells with a succession of clubs : [15 15 Oldham Athletic 15] , who paid # 250 for his services , [16 16 Birmingham 16] for # 325 , to [6 Scotland 6] with [17 17 Hamilton Academical 17] , back to [7 England 7] with [18 18 New Brighton 18] , and then [19 19 Wolverhampton Wanderers 19] , but without appearing for their league side . Waddell then had various 6 player-coach roles , with , [20 20 Chorley 20] and [21 21 Fraserburgh 21] , and had spells as 7 reserve team coach and assistant trainer with former club [8 Preston North End 8] , and trainer with , before finishing his football career with Ribblesdale Motors . Waddell died in 8 Sible Hedingham , [22 22 Essex 22]

1: 60918 Scottish_Football_League
2: 69713 Rangers_F.C.
3: 431144 Bradford_City_A.F.C.
4: 429092 Preston_North_End_F.C.
5: 451131 Oldham_Athletic_A.F.C.
6: 185196 Birmingham_City_F.C.
7: 2246991 New_Brighton_A.F.C.
8: 604447 Lesmahagow
9: 164890 Lanarkshire
10: 69713 Rangers_F.C.
11: 405663 Kilburn_F.C.
12: 431144 Bradford_City_A.F.C.
13: 28789929 Football_League_First_Division
14: 429092 Preston_North_End_F.C.
15: 451131 Oldham_Athletic_A.F.C.
16: 185196 Birmingham_City_F.C.
17: 309502 Hamilton_Academical_F.C.
18: 2246991 New_Brighton_A.F.C.
19: 217730 Wolverhampton_Wanderers_F.C.
20: 3667991 Chorley_F.C.
21: 880430 Fraserburgh_F.C.
22: 40100758 Essex

1: 10568 Association_football

[winner: 10568 Association_football, pem_pos=0,
final_scores=-1.3433129, final_local_scores=-1.4704537,
pem=0.0, similarity=-39.79156]

gt: 10568 Association_football, pem_pos=0,
final_scores=-1.3433129, final_local_scores=-1.4704537,

```

pem=0.0, similarity=-39.79156 ]
2: 14419588 The_Football_League
  [winner: 14419588 The_Football_League, pem_pos=0,
  final_scores=0.29634398, final_local_scores=0.22883242,
  pem=-0.12384478, similarity=11.592016]
  gt: 14419588 The_Football_League, pem_pos=0,
  final_scores=0.29634398, final_local_scores=0.22883242,
  pem=-0.12384478, similarity=11.592016 ]
  [winner: 2118516 German_federal_election,_March_1933,
  pem_pos=0, final_scores=-1.5881468,
  final_local_scores=-1.5943499, pem=-0.79592705,
  similarity=-39.160492]
  RECALL MISS: 14419588 The_Football_League]
3: 13293657 Wing_half
  [winner: 13293657 Wing_half, pem_pos=0,
  final_scores=-1.2104386, final_local_scores=-1.3980329,
  pem=0.0, similarity=-37.630375]
  gt: 13293657 Wing_half, pem_pos=0,
  final_scores=-1.2104386, final_local_scores=-1.3980329,
  pem=0.0, similarity=-37.630375 ]
4: 421401 Scottish_Junior_Football_Association
  [winner: 454394 Porsche_Junior, pem_pos=1,
  final_scores=-0.9851017, final_local_scores=-0.83787334,
  pem=-2.1720424, similarity=-9.102195]
  gt: 421401 Scottish_Junior_Football_Association,
  pem_pos=5, final_scores=-1.4453554,
  final_local_scores=-1.6374025, pem=-2.9652731,
  similarity=-28.648024 ]
  [winner: 16200 Jack_Kemp, pem_pos=0,
  final_scores=-1.278847, final_local_scores=-1.3246322,
  pem=0.0, similarity=-35.439953]
  RECALL MISS: 421401 Scottish_Junior_Football_Association]
5: 426311 Queen's_Park_F.C.
  [winner: 426311 Queen's_Park_F.C., pem_pos=0,
  final_scores=-0.55034643, final_local_scores=-0.68259484,
  pem=-0.7745507, similarity=-12.068152]
  gt: 426311 Queen's_Park_F.C., pem_pos=0,
  final_scores=-0.55034643, final_local_scores=-0.68259484,
  pem=-0.7745507, similarity=-12.068152 ]
6: 4113951 Player-coach
  [winner: 4113951 Player-coach, pem_pos=0,
  final_scores=-1.2193117, final_local_scores=-1.3704164,
  pem=-0.010582093, similarity=-36.7487]
  gt: 4113951 Player-coach, pem_pos=0,
  final_scores=-1.2193117, final_local_scores=-1.3704164,
  pem=-0.010582093, similarity=-36.7487 ]
7: 3752681 Reserve_team
  [winner: 13701187
  Reserve_Township,_Parke_County,_Indiana, pem_pos=9,
  final_scores=-1.2230185, final_local_scores=-1.18312,
  pem=-2.7167182, similarity=-16.443008]
  RECALL MISS: 3752681 Reserve_team]

```

[winner: 14708158 TEAMS_(cable_system), pem_pos=3,
 final_scores=-1.6844648, final_local_scores=-1.6689348,
 pem=-3.9775972, similarity=-24.083841]
 RECALL MISS: 3752681 Reserve_team]
 8: 13696562 Sible_Hedingham
 [winner: 13696562 Sible_Hedingham, pem_pos=0,
 final_scores=-0.030129924, final_local_scores=0.07197948,
 pem=0.0, similarity=6.2377243]
 gt: 13696562 Sible_Hedingham, pem_pos=0,
 final_scores=-0.030129924, final_local_scores=0.07197948,
 pem=0.0, similarity=6.2377243]
 1: 60918 Scottish_Football_League
 [pem_pos=0, final_scores=0.33197433,
 final_local_scores=0.3027084, pem=0.0, similarity=13.123135]
 2: 69713 Rangers_F.C.
 [pem_pos=0, final_scores=0.6037789,
 final_local_scores=0.576552, pem=-0.30191645,
 similarity=22.937042]
 3: 431144 Bradford_City_A.F.C.
 [pem_pos=0, final_scores=0.412265,
 final_local_scores=0.3545144, pem=0.0, similarity=14.669129]
 4: 429092 Preston_North_End_F.C.
 [pem_pos=0, final_scores=0.5356366,
 final_local_scores=0.47802174, pem=0.0, similarity=18.354836]
 5: 451131 Oldham_Athletic_A.F.C.
 [pem_pos=0, final_scores=0.41207045,
 final_local_scores=0.35247186, pem=0.0, similarity=14.608176]
 6: 185196 Birmingham_City_F.C.
 [pem_pos=2, final_scores=0.6746997,
 final_local_scores=0.6334972, pem=-3.102209,
 similarity=39.864826]
 7: 2246991 New_Brighton_A.F.C.
 [pem_pos=0, final_scores=0.35130352,
 final_local_scores=0.3104235, pem=-1.2933958,
 similarity=20.387056]
 8: 604447 Lesmahagow
 [pem_pos=0, final_scores=0.064720735,
 final_local_scores=0.113921344, pem=0.0, similarity=7.489353]
 9: 164890 Lanarkshire
 [pem_pos=0, final_scores=0.082837775,
 final_local_scores=0.08540852, pem=0.0, similarity=6.6384735]
 10: 69713 Rangers_F.C.
 [pem_pos=0, final_scores=0.658058,
 final_local_scores=0.63584363, pem=-0.30191645,
 similarity=24.706423]
 11: 405663 Kilburn_F.C.
 [pem_pos=1, final_scores=0.4326573,
 final_local_scores=0.39230114, pem=-1.4326576,
 similarity=23.587776]
 12: 431144 Bradford_City_A.F.C.
 [pem_pos=0, final_scores=0.3041613,
 final_local_scores=0.23593694, pem=0.0, similarity=11.130542]

13: 28789929 Football_League_First_Division
[pem_pos=0, final_scores=0.1960027,
final_local_scores=0.12909341, pem=-0.5147791,
similarity=10.741567]
14: 429092 Preston_North_End_F.C.
[pem_pos=0, final_scores=0.45108938,
final_local_scores=0.38566673, pem=0.0, similarity=15.598777]
15: 451131 Oldham_Athletic_A.F.C.
[pem_pos=0, final_scores=0.41518193,
final_local_scores=0.35587066, pem=0.0, similarity=14.709603]
16: 185196 Birmingham_City_F.C.
[pem_pos=2, final_scores=0.5166625,
final_local_scores=0.46086544, pem=-3.102209,
similarity=34.71315]
17: 309502 Hamilton_Academical_F.C.
[pem_pos=0, final_scores=0.44616777,
final_local_scores=0.40574598, pem=0.0, similarity=16.197983]
18: 2246991 New_Brighton_A.F.C.
[pem_pos=0, final_scores=0.2737456,
final_local_scores=0.2257033, pem=-1.2933958,
similarity=17.858837]
19: 217730 Wolverhampton_Wanderers_F.C.
[pem_pos=0, final_scores=0.39051467,
final_local_scores=0.32434598, pem=0.0, similarity=13.768844]
20: 3667991 Chorley_F.C.
[pem_pos=1, final_scores=0.17282663,
final_local_scores=0.121827036, pem=-1.7396293,
similarity=17.18565]
21: 880430 Fraserburgh_F.C.
[pem_pos=1, final_scores=0.18516953,
final_local_scores=0.19546363, pem=-1.0449705,
similarity=15.60545]
22: 40100758 Essex
[pem_pos=0, final_scores=0.22481914,
final_local_scores=0.23026079, pem=-0.576716,
similarity=14.097425]
1: 26994 Scotland
[pem_pos=0, final_scores=0.5022191,
final_local_scores=0.53363144, pem=-0.47951394,
similarity=22.622011]
2: 14419588 The_Football_League
[pem_pos=0, final_scores=0.29634398,
final_local_scores=0.22883242, pem=-0.12384478,
similarity=11.592016]
3: 9316 England
[pem_pos=1, final_scores=0.40086502,
final_local_scores=0.41860408, pem=-0.19053856,
similarity=17.617872]
4: 69713 Rangers_F.C.
[pem_pos=0, final_scores=0.61824936,
final_local_scores=0.5923587, pem=-0.30191645,
similarity=23.408747]

5: 28789929 Football_League_First_Division
 [pem_pos=0, final_scores=0.2697215,
 final_local_scores=0.20962, pem=-0.5147791,
 similarity=13.144641]
 6: 26994 Scotland
 [pem_pos=0, final_scores=0.28720278,
 final_local_scores=0.29875883, pem=-0.22905664,
 similarity=14.250918]
 7: 9316 England
 [pem_pos=0, final_scores=0.42078662,
 final_local_scores=0.44036543, pem=-0.19053856,
 similarity=18.267275]
 8: 429092 Preston_North_End_F.C.
 [pem_pos=0, final_scores=0.56635886,
 final_local_scores=0.5115811, pem=0.0, similarity=19.356314]

Sample: b'Alexander_MacDonnell'

, _ 3rd_Earl_of_Antrim Alexander MacDonnell , 3rd Earl of Antrim -LRB- 1615 -- 10 December 1699 -RRB- was a peer and military commander in [1 1 Ireland 1] . He was the son of [2 2 Randal MacDonnell 2] , 1st Earl of Antrim and his wife , Alice O'Neill -LRB- Alice was the daughter of -RRB- . After coming of age , MacDonnell spend three years abroad in [3 Europe 3] before returning to [4 Ireland 4] just before the outbreak of the 3 Irish Rebellion of 1641 . [5 MacDonnell 5] was commander of a regiment of [6 Irish 6] against [7 Cromwell 7] in 1641 , and soon took his place among the [8 Irish 8] leadership . Unlike his brother , [1 1 Randal 1] , he urged a conciliatory approach to the [9 English 9] under [2 2 Oliver Cromwell 2] . He remained in command of a regiment until the Confederation 4 surrendered to [10 Cromwell 10] in 1652 . His lands were confiscated and 5 distributed amongst [1 1 Cromwellian 11] soldiers . By 1656 , MacDonnell was living in [12 England 12] , and at was restored to his lands . In 1680 , MacDonnell was appointed 6 custos rotulorum for Antrim , and he succeeded to the title of 7 Earl of Antrim on 3 February 1682 on the death of his brother . He was invested as a 8 Privy Counsellor and shortly after the accession of King [3 3 James II 3] , as the King followed a policy of replacing [13 Protestants 13] with [14 Catholics 14] across [1 5 Ireland 15] . At the outbreak of the [4 4 Glorious Revolution 4] , MacDonnell adhered to [16 James 16] and presided over the unsuccessful 9 siege of [17 Derry 17] . On [18 James 18] 's defeat , MacDonnell 's lands and titles were , again , declared forfeit . In 1697 he was restored to his title . He married , firstly , Lady Elizabeth Annesley , daughter of [19 10 Arthur Annesley 19] , 1st Earl of [20 Anglesey 20] . She died childless in 1669 . He married , secondly , Helena Burke , daughter of Sir John Burke of , after 1672 . They had one son ,

1: 2718846

Randal_MacDonnell,_1st_Marquess_of_Antrim_(1645_creation)

2: 22413 Oliver_Cromwell
3: 5243711 James_II_of_England
4: 12466 Glorious_Revolution
1: 147575 Ireland
[winner: 14560 Republic_of_Ireland, pem_pos=1,
final_scores=0.43864137, final_local_scores=0.4716584,
pem=-1.5040284, similarity=26.344082]
gt: 147575 Ireland, pem_pos=0,
final_scores=-0.019648135, final_local_scores=-0.06882865,
pem=-0.5653214, similarity=5.1100345]
2: 2718997 Randal_MacDonnell,_1st_Earl_of_Antrim
[winner: 2718846
Randal_MacDonnell,_1st_Marquess_of_Antrim_(1645_creation),
pem_pos=0, final_scores=0.23305546,
final_local_scores=0.22143757, pem=0.0, similarity=10.697852]
gt: 2718997 Randal_MacDonnell,_1st_Earl_of_Antrim,
pem_pos=1, final_scores=0.07256673,
final_local_scores=0.076762915, pem=-0.095310144,
similarity=6.8987827]
[winner: 465369 Earl_of_Antrim, pem_pos=0,
final_scores=-0.10382706, final_local_scores=-0.14183775,
pem=0.0, similarity=-0.14300838]
RECALL MISS: 2718997
Randal_MacDonnell,_1st_Earl_of_Antrim]
[winner: 1071666 Arthur_Annesley,_1st_Earl_of_Anglesey,
pem_pos=0, final_scores=-1.0891793,
final_local_scores=-1.2132838, pem=0.0,
similarity=-32.117096]
RECALL MISS: 2718997
Randal_MacDonnell,_1st_Earl_of_Antrim]
3: 1460319 Irish_Rebellion_of_1641
[winner: 1460319 Irish_Rebellion_of_1641, pem_pos=0,
final_scores=-0.19507678, final_local_scores=-0.27412945,
pem=-0.2189152, similarity=-2.9003623]
gt: 1460319 Irish_Rebellion_of_1641, pem_pos=0,
final_scores=-0.19507678, final_local_scores=-0.27412945,
pem=-0.2189152, similarity=-2.9003623]
[winner: 35202 1641, pem_pos=0, final_scores=-0.7713476,
final_local_scores=-0.6571799, pem=-0.3547,
similarity=-13.592932]
RECALL MISS: 1460319 Irish_Rebellion_of_1641]
4: 1485171 Cromwellian_conquest_of_Ireland
[winner: 22413 Oliver_Cromwell, pem_pos=0,
final_scores=0.1274272, final_local_scores=0.05603984,
pem=0.0, similarity=5.7620535]
gt: 1485171 Cromwellian_conquest_of_Ireland, pem_pos=18,
final_scores=-0.6166063, final_local_scores=-0.7024669,
pem=-3.6950116, similarity=3.2207477]
[winner: 9987467 Peace_of_Travendal, pem_pos=12,
final_scores=-1.4360538, final_local_scores=-1.5076256,
pem=-2.580217, similarity=-26.869219]
RECALL MISS: 1485171 Cromwellian_conquest_of_Ireland]

5: 1701154 Act_for_the_Settlement_of_Ireland_1652
 [winner: 22413 Oliver_Cromwell, pem_pos=0,
 final_scores=0.046325877, final_local_scores=-0.035269707,
 pem=-0.17073841, similarity=3.9656963]
 RECALL MISS: 1701154
 Act_for_the_Settlement_of_Ireland_1652]
 [winner: 301950 Distributed_element_model, pem_pos=3,
 final_scores=-1.1926082, final_local_scores=-1.1006849,
 pem=-1.5071591, similarity=-20.560755]
 RECALL MISS: 1701154
 Act_for_the_Settlement_of_Ireland_1652]
 [winner: 5731456 Amongst_Women, pem_pos=0,
 final_scores=-1.218382, final_local_scores=-1.2711661,
 pem=-0.6359887, similarity=-30.385813]
 RECALL MISS: 1701154
 Act_for_the_Settlement_of_Ireland_1652]
 [winner: 34349293 Croatian_Home_Guard, pem_pos=14,
 final_scores=-1.2515231, final_local_scores=-1.1895499,
 pem=-2.6390574, similarity=-17.057226]
 RECALL MISS: 1701154
 Act_for_the_Settlement_of_Ireland_1652]
 6: 2611959 Custos_rotulorum
 [winner: 4321189 Custos_Rotulorum_of_Berkshire,
 pem_pos=0, final_scores=-0.5594545,
 final_local_scores=-0.4573114, pem=-0.6931472,
 similarity=-5.7879286]
 RECALL MISS: 2611959 Custos_rotulorum]
 [winner: 17453757 Custos, pem_pos=0,
 final_scores=-1.0438033, final_local_scores=-1.0509441,
 pem=-0.1554849, similarity=-26.427004]
 gt: 2611959 Custos_rotulorum, pem_pos=1,
 final_scores=-1.3411292, final_local_scores=-1.4652338,
 pem=-1.9951004, similarity=-28.78612]
 7: 465369 Earl_of_Antrim
 [winner: 465369 Earl_of_Antrim, pem_pos=0,
 final_scores=0.005190119, final_local_scores=-0.022753038,
 pem=0.0, similarity=3.410717]
 gt: 465369 Earl_of_Antrim, pem_pos=0,
 final_scores=0.005190119, final_local_scores=-0.022753038,
 pem=0.0, similarity=3.410717]
 8: 3713657 Privy_Council_of_Ireland
 [winner: 2958004 Queen-in-Council, pem_pos=0,
 final_scores=-0.09553085, final_local_scores=-0.06847893,
 pem=-0.574459, similarity=5.1701627]
 gt: 3713657 Privy_Council_of_Ireland, pem_pos=1,
 final_scores=-0.29225445, final_local_scores=-0.3564424,
 pem=-1.3184307, similarity=0.6225921]
 9: 1064035 Siege_of_Derry
 [winner: 9055 Derry, pem_pos=0, final_scores=0.28669733,
 final_local_scores=0.27111754, pem=-0.3054378,
 similarity=13.841419]
 gt: 1064035 Siege_of_Derry, pem_pos=11,

final_scores=-0.47096837, final_local_scores=-0.53382045,
pem=-3.3737128, similarity=6.5062203]
[winner: 20434447 Siege_of_Constantinople_(1235),
pem_pos=0, final_scores=-0.8997851,
final_local_scores=-0.86954474, pem=0.0,
similarity=-21.859236]
RECALL MISS: 1064035 Siege_of_Derry]
10: 1071666 Arthur_Annesley,_1st_Earl_of_Anglesey
[winner: 18932156 Anglesey, pem_pos=0,
final_scores=0.24287353, final_local_scores=0.24050117,
pem=-0.03841769, similarity=11.475669]
RECALL MISS: 1071666
Arthur_Annesley,_1st_Earl_of_Anglesey]
[winner: 1071666 Arthur_Annesley,_1st_Earl_of_Anglesey,
pem_pos=0, final_scores=0.117473796,
final_local_scores=0.110933304, pem=-0.26495385,
similarity=8.841044]
gt: 1071666 Arthur_Annesley,_1st_Earl_of_Anglesey,
pem_pos=0, final_scores=0.117473796,
final_local_scores=0.110933304, pem=-0.26495385,
similarity=8.841044]
[winner: 1071666 Arthur_Annesley,_1st_Earl_of_Anglesey,
pem_pos=2, final_scores=-0.4973306,
final_local_scores=-0.566779, pem=0.0, similarity=-12.824104]
gt: 1071666 Arthur_Annesley,_1st_Earl_of_Anglesey,
pem_pos=2, final_scores=-0.4973306,
final_local_scores=-0.566779, pem=0.0,
similarity=-12.824104]
[winner: 2718997 Randal_MacDonnell,_1st_Earl_of_Antrim,
pem_pos=0, final_scores=-1.1260513,
final_local_scores=-1.2408712, pem=0.0,
similarity=-32.940357]
RECALL MISS: 1071666
Arthur_Annesley,_1st_Earl_of_Anglesey]
1: 2718846
Randal_MacDonnell,_1st_Marquess_of_Antrim_(1645_creation)
[pem_pos=1, final_scores=0.17176975,
final_local_scores=0.15632415, pem=0.0, similarity=8.754737]
2: 22413 Oliver_Cromwell
[pem_pos=0, final_scores=0.22757526,
final_local_scores=0.16543639, pem=0.0, similarity=9.026665]
3: 5243711 James_II_of_England
[pem_pos=0, final_scores=0.33478254,
final_local_scores=0.27363393, pem=-0.12727182,
similarity=12.9476185]
4: 12466 Glorious_Revolution
[pem_pos=0, final_scores=0.22066455,
final_local_scores=0.1672937, pem=0.0, similarity=9.08209]
1: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.43864137,
final_local_scores=0.4716584, pem=-1.5040284,
similarity=26.344082]

2: 2718846
Randal_MacDonnell,_1st_Marquess_of_Antrim_(1645_creation)
[pem_pos=0, final_scores=0.23305546,
final_local_scores=0.22143757, pem=0.0, similarity=10.697852]
3: 9239 Europe
[pem_pos=0, final_scores=0.46499234,
final_local_scores=0.55247265, pem=-0.086338006,
similarity=21.04612]
4: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.41075188,
final_local_scores=0.4411934, pem=-1.5040284,
similarity=25.434946]
5: 2718846
Randal_MacDonnell,_1st_Marquess_of_Antrim_(1645_creation)
[pem_pos=1, final_scores=0.14238562,
final_local_scores=0.12101784, pem=0.0, similarity=7.7011266]
6: 14560 Republic_of_Ireland
[pem_pos=3, final_scores=0.5838245,
final_local_scores=0.63024896, pem=-2.2576945,
similarity=35.175293]
7: 22413 Oliver_Cromwell
[pem_pos=0, final_scores=0.16937779,
final_local_scores=0.10186449, pem=0.0, similarity=7.1295524]
8: 14560 Republic_of_Ireland
[pem_pos=3, final_scores=0.5179262,
final_local_scores=0.55826503, pem=-2.2576945,
similarity=33.02715]
9: 9316 England
[pem_pos=1, final_scores=0.3832639,
final_local_scores=0.38711733, pem=-1.3527626,
similarity=22.998598]
10: 22413 Oliver_Cromwell
[pem_pos=0, final_scores=0.1274272,
final_local_scores=0.05603984, pem=0.0, similarity=5.7620535]
11: 22413 Oliver_Cromwell
[pem_pos=0, final_scores=0.046325877,
final_local_scores=-0.035269707, pem=-0.17073841,
similarity=3.9656963]
12: 9316 England
[pem_pos=0, final_scores=0.51440555,
final_local_scores=0.53036964, pem=-0.19053856,
similarity=20.953178]
13: 25814008 Protestantism
[pem_pos=0, final_scores=0.22164236,
final_local_scores=0.20979977, pem=0.0, similarity=10.350556]
14: 606848 Catholic_Church
[pem_pos=0, final_scores=0.24523963,
final_local_scores=0.23723191, pem=-0.38803694,
similarity=13.279391]
15: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.42729908,
final_local_scores=0.45926863, pem=-1.5040284,

```

similarity=25.974346]
16: 5243711 James_II_of_England
    [pem_pos=0, final_scores=0.27294862,
final_local_scores=0.20978364, pem=0.0, similarity=10.350076]
17: 9055 Derry
    [pem_pos=0, final_scores=0.28669733,
final_local_scores=0.27111754, pem=-0.3054378,
similarity=13.841419]
18: 5243711 James_II_of_England
    [pem_pos=0, final_scores=0.24033977,
final_local_scores=0.17416343, pem=0.0, similarity=9.287097]
19: 1071666 Arthur_Annesley,_1st_Earl_of_Anglesey
    [pem_pos=0, final_scores=0.117473796,
final_local_scores=0.110933304, pem=-0.26495385,
similarity=8.841044]
20: 18932156 Anglesey
    [pem_pos=0, final_scores=0.24287353,
final_local_scores=0.24050117, pem=-0.03841769,
similarity=11.475669]

```

Sample: b'Marksville_culture'

The [1 Marksville 1] culture was an 1 archaeological culture in the lower valley , 2 Yazoo valley , and [2 3 Tensas 2] valley areas of [1 1 Louisiana 1] , [2 2 Mississippi 2] , [3 3 Missouri 3] , [4 4 Arkansas 4] and extended eastward along the [3 Gulf Coast 3] to the [5 5 Mobile Bay 5] area , from 100 BCE to 400 CE . This culture takes its name from the 4 Marksville Prehistoric Indian Site in [4 5 Avoyelles Parish 4] , [5 Louisiana 5] . Marksville Culture was contemporaneous with the 6 Hopewell cultures of [6 6 Ohio 6] and [7 7 Illinois 7] . It evolved from the earlier and into the [6 7 Baytown 6] and 8 Troyville culture s and later the [8 8 Coles Creek 8] and 9 Plum Bayou cultures . It is considered ancestral to the [7 10 Natchez 7] and 11 Taensa Peoples . The Hopewell tradition was a widely dispersed set of related populations , which were connected by a common network of trade routes , known as the . The [8 Marksville 8] culture was a southern manifestation of this network . Settlements were large and usually located on terraces of major streams . Evidence from excavations of burial mounds from this period suggest they were constructed for persons of high social status , and contained grave goods of imported exotic materials such as 12 copper panpipes , earspools , bracelets and beads , rare minerals , stone platform pipes , 13 mica figurines , marine shells , freshwater pearls , and 14 greenstone celts . The pipes had flat bases with a hole for a stem , and a bowl in the center . Animal figurines on the platform are n't unusual , with the bowl being located in the animal 's back . The high status leaders organized community life , and officiated at burial ceremonies , an important part of the Marksville Culture . The mounds were constructed in stages over many years , with the first stage being a flat , low

platform . The ceremonies may have been held years apart and those who died between ceremonies were temporarily stored in other locations , then gathered up and buried together . Although made from local clay , [9 Marksville 9] pottery was similar in design and decoration to pottery found in [10 Illinois 10] and [11 Ohio 11] . A typical vessel was three to five inches tall and three to seven inches in diameter and often decorated with geometric and effigy designs , usually stylized birds . This decorated pottery was made primarily for ceremonial uses , with other plainer utilitarian ware for daily use . [12 Marksville 12] pottery influenced [13 Santa Rosa 13] pottery , a defining character of the contemporary 15 Santa Rosa-Swift Creek culture , located to the east of the [14 Marksville 14] culture area along the Gulf coast . The [15 Marksville 15] culture was preceded by the , and was eventually succeeded by the 16 Troyville culture in southeastern and eastern [16 Louisiana 16] and western [17 Mississippi 17] and [18 17 Baytown 18] culture

1: 18130 Louisiana

2: 16949861 Mississippi

3: 19571 Missouri

4: 1930 Arkansas

5: 36530 Mobile_Bay

6: 22199 Ohio

7: 14849 Illinois

8: 19250964 Coles_Creek_culture

1: 642871 Archaeological_culture

[winner: 8125431 Saskatchewan_Archaeological_Society,
pem_pos=2, final_scores=-1.259771,
final_local_scores=-1.2356019, pem=-1.0979562,
similarity=-26.81226]

RECALL MISS: 642871 Archaeological_culture]

[winner: 19159508 Culture, pem_pos=0,
final_scores=-1.5910233, final_local_scores=-1.6734045,
pem=-0.34725943, similarity=-43.959564]

gt: 642871 Archaeological_culture, pem_pos=4,
final_scores=-1.8038727, final_local_scores=-1.9071658,
pem=-3.0768232, similarity=-36.091675]

2: 296815 Yazoo_River

[winner: 296815 Yazoo_River, pem_pos=3,
final_scores=-0.038229793, final_local_scores=-0.05576554,
pem=-1.9354336, similarity=12.950747]

gt: 296815 Yazoo_River, pem_pos=3,
final_scores=-0.038229793, final_local_scores=-0.05576554,
pem=-1.9354336, similarity=12.950747]

3: 1785256 Tensas_River

[winner: 97663 Tensas_Parish, Louisiana, pem_pos=0,
final_scores=0.20790039, final_local_scores=0.17063221,
pem=0.0, similarity=9.181719]

gt: 1785256 Tensas_River, pem_pos=1,
final_scores=-0.007182017, final_local_scores=0.0242351,
pem=-0.6241007, similarity=8.206893]

4: 15511127 Marksville_Prehistoric_Indian_Site
 [winner: 15511127 Marksville_Prehistoric_Indian_Site,
 pem_pos=0, final_scores=-0.12536687,
 final_local_scores=-0.039033726, pem=0.0,
 similarity=2.9248686|
 gt: 15511127 Marksville_Prehistoric_Indian_Site,
 pem_pos=0, final_scores=-0.12536687,
 final_local_scores=-0.039033726, pem=0.0,
 similarity=2.9248686]

5: 97715 Avoyelles_Parish,_Louisiana
 [winner: 18130 Louisiana, pem_pos=0,
 final_scores=0.41979396, final_local_scores=0.3562283,
 pem=-0.17745006, similarity=15.685277|
 RECALL MISS: 97715 Avoyelles_Parish,_Louisiana]
 [winner: 97715 Avoyelles_Parish,_Louisiana, pem_pos=0,
 final_scores=0.37651265, final_local_scores=0.35343677,
 pem=0.0, similarity=14.636971|
 gt: 97715 Avoyelles_Parish,_Louisiana, pem_pos=0,
 final_scores=0.37651265, final_local_scores=0.35343677,
 pem=0.0, similarity=14.636971]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7812811, final_local_scores=-1.7462897,
 pem=-4.6728287, similarity=-22.611485|
 RECALL MISS: 97715 Avoyelles_Parish,_Louisiana]

6: 313110 Hopewell_tradition
 [winner: 77669 Hopewell,_Virginia, pem_pos=1,
 final_scores=-0.19588423, final_local_scores=-0.23110348,
 pem=-2.4775558, similarity=10.666457|
 gt: 313110 Hopewell_tradition, pem_pos=0,
 final_scores=-0.34490752, final_local_scores=-0.3490409,
 pem=-2.0350084, similarity=4.7403283]
 [winner: 541900 Microbiological_culture, pem_pos=1,
 final_scores=-1.4937221, final_local_scores=-1.4719844,
 pem=-1.0307838, similarity=-34.231674|
 RECALL MISS: 313110 Hopewell_tradition]

7: 19442337 Baytown_culture
 [winner: 151127 Baytown,_Texas, pem_pos=0,
 final_scores=0.34739417, final_local_scores=0.3385375,
 pem=0.0, similarity=14.192348|
 gt: 19442337 Baytown_culture, pem_pos=1,
 final_scores=-0.23861656, final_local_scores=-0.18487687,
 pem=-0.8517815, similarity=3.2047386]

8: 33491589 Troyville_culture
 [winner: 33491589 Troyville_culture, pem_pos=0,
 final_scores=-0.18383797, final_local_scores=-0.11525783,
 pem=0.0, similarity=0.6501895|
 gt: 33491589 Troyville_culture, pem_pos=0,
 final_scores=-0.18383797, final_local_scores=-0.11525783,
 pem=0.0, similarity=0.6501895]
 [winner: 1106830 Cell_culture, pem_pos=5,
 final_scores=-1.8215518, final_local_scores=-1.8404534,
 pem=-3.1298656, similarity=-33.81239|

RECALL MISS: 33491589 Troyville_culture]

9: 33224238 Plum_Bayou_culture
[winner: 33224238 Plum_Bayou_culture, pem_pos=0,
final_scores=0.010551855, final_local_scores=0.08607276,
pem=0.0, similarity=6.6582956|
gt: 33224238 Plum_Bayou_culture, pem_pos=0,
final_scores=0.010551855, final_local_scores=0.08607276,
pem=0.0, similarity=6.6582956]

10: 1537516 Natchez_people
[winner: 122205 Natchez, Mississippi, pem_pos=0,
final_scores=0.13686214, final_local_scores=0.06950866,
pem=-0.18184498, similarity=7.152891|
gt: 1537516 Natchez_people, pem_pos=1,
final_scores=-0.19890611, final_local_scores=-0.21195547,
pem=-1.4231377, similarity=5.5037827]

11: 3799783 Taensa
[winner: 3799783 Taensa, pem_pos=0,
final_scores=-0.6976647, final_local_scores=-0.7116486,
pem=0.0, similarity=-17.1473|
gt: 3799783 Taensa, pem_pos=0, final_scores=-0.6976647,
final_local_scores=-0.7116486, pem=0.0, similarity=-17.1473]

12: 125293 Copper
[winner: 230641 Copper_interconnect, pem_pos=7,
final_scores=-1.2215356, final_local_scores=-1.1357772,
pem=-1.9456949, similarity=-19.223156|
gt: 125293 Copper, pem_pos=0, final_scores=-1.3212641,
final_local_scores=-1.3596041, pem=-0.017616155,
similarity=-36.387787]

13: 21061 Mica
[winner: 21061 Mica, pem_pos=0, final_scores=-0.7881574,
final_local_scores=-0.771555, pem=-0.0020422915,
similarity=-18.923922|
gt: 21061 Mica, pem_pos=0, final_scores=-0.7881574,
final_local_scores=-0.771555, pem=-0.0020422915,
similarity=-18.923922]

14: 13766386 Greenstone_(archaeology)
[winner: 2636840 Greenschist, pem_pos=1,
final_scores=-1.0058094, final_local_scores=-0.948798,
pem=-0.7464557, similarity=-20.164968|
gt: 13766386 Greenstone_(archaeology), pem_pos=3,
final_scores=-1.2091991, final_local_scores=-1.2032387,
pem=-1.1361951, similarity=-25.638523]

15: 22148459 Santa_Rosa-Swift_Creek_culture
[winner: 22148459 Santa_Rosa-Swift_Creek_culture,
pem_pos=0, final_scores=-0.2508293,
final_local_scores=-0.15529805, pem=0.0,
similarity=-0.54469097|
gt: 22148459 Santa_Rosa-Swift_Creek_culture, pem_pos=0,
final_scores=-0.2508293, final_local_scores=-0.15529805,
pem=0.0, similarity=-0.54469097]
[winner: 1106830 Cell_culture, pem_pos=5,
final_scores=-1.7948097, final_local_scores=-1.8112416,

```

pem=-3.1298656, similarity=-32.94065|
  RECALL MISS: 22148459 Santa_Rosa-Swift_Creek_culture]
16: 33491589 Troyville_culture
  [winner: 33491589 Troyville_culture, pem_pos=0,
  final_scores=-0.174659, final_local_scores=-0.1052312,
  pem=0.0, similarity=0.9494039|
  gt: 33491589 Troyville_culture, pem_pos=0,
  final_scores=-0.174659, final_local_scores=-0.1052312,
  pem=0.0, similarity=0.9494039 ]
  [winner: 1106830 Cell_culture, pem_pos=5,
  final_scores=-1.8558439, final_local_scores=-1.8779123,
  pem=-3.1298656, similarity=-34.930237|
  RECALL MISS: 33491589 Troyville_culture]
17: 19442337 Baytown_culture
  [winner: 151127 Baytown, Texas, pem_pos=0,
  final_scores=0.29888862, final_local_scores=0.28555253,
  pem=0.0, similarity=12.61117|
  gt: 19442337 Baytown_culture, pem_pos=1,
  final_scores=-0.4047309, final_local_scores=-0.36633155,
  pem=-0.8517815, similarity=-2.2102308 ]
  [winner: 541900 Microbiological_culture, pem_pos=2,
  final_scores=-1.4858736, final_local_scores=-1.4634111,
  pem=-3.0016317, similarity=-23.258053|
  RECALL MISS: 19442337 Baytown_culture]
1: 18130 Louisiana
  [pem_pos=0, final_scores=0.46851224,
  final_local_scores=0.40944558, pem=-0.17745006,
  similarity=17.273386]
2: 16949861 Mississippi
  [pem_pos=0, final_scores=0.28162855,
  final_local_scores=0.20919773, pem=-0.31944752,
  similarity=12.069796]
3: 19571 Missouri
  [pem_pos=0, final_scores=0.44961107,
  final_local_scores=0.4062106, pem=-0.27068004,
  similarity=17.683847]
4: 1930 Arkansas
  [pem_pos=0, final_scores=0.4596762,
  final_local_scores=0.41589433, pem=-0.32007957,
  similarity=18.241472]
5: 36530 Mobile_Bay
  [pem_pos=0, final_scores=0.18736587,
  final_local_scores=0.18653408, pem=0.0, similarity=9.656261]
6: 22199 Ohio
  [pem_pos=0, final_scores=0.6348135,
  final_local_scores=0.6233074, pem=-0.19147073,
  similarity=23.731697]
7: 14849 Illinois
  [pem_pos=0, final_scores=0.5213904,
  final_local_scores=0.4942084, pem=-0.16995399,
  similarity=19.762114]
8: 19250964 Coles_Creek_culture

```

[pem_pos=0, final_scores=0.09534274,
final_local_scores=0.17429128, pem=0.0, similarity=9.290913]
1: 115488 Marksville, Louisiana
[pem_pos=0, final_scores=0.10431366,
final_local_scores=0.06293422, pem=0.0, similarity=5.9677954]
2: 97663 Tensas Parish, Louisiana
[pem_pos=0, final_scores=0.20790039,
final_local_scores=0.17063221, pem=0.0, similarity=9.181719]
3: 11969 Gulf Coast of the United States
[pem_pos=0, final_scores=0.07390158,
final_local_scores=0.041662857, pem=-0.19244206,
similarity=6.3795457]
4: 97715 Avoyelles Parish, Louisiana
[pem_pos=0, final_scores=0.37651265,
final_local_scores=0.35343677, pem=0.0, similarity=14.636971]
5: 18130 Louisiana
[pem_pos=0, final_scores=0.41979396,
final_local_scores=0.3562283, pem=-0.17745006,
similarity=15.685277]
6: 151127 Baytown, Texas
[pem_pos=0, final_scores=0.34739417,
final_local_scores=0.3385375, pem=0.0, similarity=14.192348]
7: 122205 Natchez, Mississippi
[pem_pos=0, final_scores=0.13686214,
final_local_scores=0.06950866, pem=-0.18184498,
similarity=7.152891]
8: 115488 Marksville, Louisiana
[pem_pos=0, final_scores=0.1791629,
final_local_scores=0.14469564, pem=0.0, similarity=8.40772]
9: 115488 Marksville, Louisiana
[pem_pos=0, final_scores=0.12297116,
final_local_scores=0.08331469, pem=0.0, similarity=6.5759892]
10: 14849 Illinois
[pem_pos=0, final_scores=0.57619995,
final_local_scores=0.5540796, pem=-0.16995399,
similarity=21.54879]
11: 22199 Ohio
[pem_pos=0, final_scores=0.6574132,
final_local_scores=0.6479941, pem=-0.19147073,
similarity=24.468397]
12: 115488 Marksville, Louisiana
[pem_pos=0, final_scores=0.046873674,
final_local_scores=-0.0054096133, pem=0.0,
similarity=3.9282792]
13: 108243 Santa Rosa, California
[pem_pos=0, final_scores=0.102810875,
final_local_scores=0.10931423, pem=-0.85449636,
similarity=11.998752]
14: 115488 Marksville, Louisiana
[pem_pos=0, final_scores=0.11138578,
final_local_scores=0.070659444, pem=0.0, similarity=6.198332]
15: 115488 Marksville, Louisiana

[pem_pos=0, final_scores=0.08213435,
 final_local_scores=0.03870669, pem=0.0, similarity=5.2447977]
 16: 18130 Louisiana
 [pem_pos=0, final_scores=0.46132833,
 final_local_scores=0.40159827, pem=-0.17745006,
 similarity=17.039207]
 17: 16949861 Mississippi
 [pem_pos=0, final_scores=0.24255781,
 final_local_scores=0.1665189, pem=-0.31944752,
 similarity=10.796174]
 18: 151127 Baytown, Texas
 [pem_pos=0, final_scores=0.29888862,
 final_local_scores=0.28555253, pem=0.0, similarity=12.61117]

Sample: b'Cody_monoplane'

The Cody IV monoplane was a single-engined **1 monoplane** designed and built by the born but **[1 British 1]** based aviation pioneer in 1912 . It was intended for entry into the **2 1912 [2 British Military Aeroplane Competition 2]** , but was wrecked in a crash before the start of the competition . In December 1911 the **[3 British 3 War Office 3]** announced a competition for a Military aeroplane capable of carrying a pilot and observer for the recently established **[1 1 Royal Flying Corps 1]** . First prize was # 4,000 , with the **[4 War Office 4]** having the option to purchase any of the prize winning machines . The **[5 American 5]** showman and aviation pioneer **[6 Samuel Cody 6]** , who had developed a system of man-carrying kites from 1901 , built his first aircraft , the **4 British Army Aeroplane No 1** at the **[2 2 Army Balloon Factory 2]** at **[3 3 Farnborough 3]** in 1908 , making its first flight , recognised as the first powered controlled flight in the **[7 United Kingdom 7]** on 16 October 1908 . In order to compete for the prize money offered by the **[8 War Office 8]** , **[9 Cody 9]** decided to design and build a new **5 monoplane** , while also entering his existing **6 biplane** that had finished fourth in the 1911 **[4 4 Daily Mail Circuit of Britain Air Race 4]** . **[10 Cody 10]** 's previous designs were **7 pusher 8 canard 9 biplane** s , but the new aircraft was a **10 tractor** high-winged monoplane powered by a single 120 hp -LRB- 89 kW -RRB- **11 Austro-Daimler** engine which had been fitted to an **[5 5 Etrich Taube 5]** which had crashed during the 1911 **[11 Circuit of Britain 11]** , which **[12 Cody 12]** had purchased after the race . It had a deep fuselage that accommodated the crew of two side-by-side . The tail arrangement was unusual , with two **12 elevators** and two rudders carried on four **13 bamboo** booms that led back from the fuselage , the all moving tail surfaces forming a box shape . Fabric was stretched between the tail booms giving a cruciform shape as seen in cross-section . Lateral control was by means of **14 wing warping** , and the aircraft was fitted with a nosewheel undercarriage . The monoplane made its maiden flight on 21 June 1912 , but on 8 July a cow ran into the path of the aircraft during

landing . The resulting crash badly damaged the aircraft and killed the cow , although [13 Cody 13] was not badly injured . As the [14 Circuit of Britain 14] aircraft had already been wrecked in a crash , [15 Cody 15] used the engine from the monoplane together with what parts he could salvage from the biplane to construct a new aircraft , the [16 15 Cody V 16] biplane

1: 233670 Royal_Flying_Corps

2: 250131 Royal_Aircraft_Establishment

3: 232477 Farnborough, Hampshire

4: 19104883 Daily_Mail_Circuit_of_Britain_Air_Race

5: 603400 Etrich_Taube

1: 75127 Monoplane

[winner: 75127 Monoplane, pem_pos=0,
final_scores=-0.75032556, final_local_scores=-0.7753606,
pem=-0.00019468772, similarity=-19.047535|

gt: 75127 Monoplane, pem_pos=0,
final_scores=-0.75032556, final_local_scores=-0.7753606,
pem=-0.00019468772, similarity=-19.047535]

2: 32897971 1912_British_Military_Aeroplane_Competition

[winner: 32897971
1912_British_Military_Aeroplane_Competition, pem_pos=0,
final_scores=0.084046856, final_local_scores=0.07998259,
pem=0.0, similarity=6.476553|

gt: 32897971
1912_British_Military_Aeroplane_Competition, pem_pos=0,
final_scores=0.084046856, final_local_scores=0.07998259,
pem=0.0, similarity=6.476553]

[winner: 34616 1912, pem_pos=0, final_scores=-1.0214727,
final_local_scores=-0.9025517, pem=-0.75507396,
similarity=-18.73802|

RECALL MISS: 32897971
1912_British_Military_Aeroplane_Competition]

3: 274472 War_Office

[winner: 274472 War_Office, pem_pos=0,
final_scores=0.13095315, final_local_scores=0.10698347,
pem=0.0, similarity=7.282313|

gt: 274472 War_Office, pem_pos=0,
final_scores=0.13095315, final_local_scores=0.10698347,
pem=0.0, similarity=7.282313]

4: 12386543 British_Army_Aeroplane_No_1

[winner: 12386543 British_Army_Aeroplane_No_1, pem_pos=0,
final_scores=-0.38063765, final_local_scores=-0.4181959,
pem=0.0, similarity=-8.390087|

gt: 12386543 British_Army_Aeroplane_No_1, pem_pos=0,
final_scores=-0.38063765, final_local_scores=-0.4181959,
pem=0.0, similarity=-8.390087]

5: 75127 Monoplane

[winner: 75127 Monoplane, pem_pos=0,
final_scores=-0.8574472, final_local_scores=-0.8923747,
pem=-0.00019468772, similarity=-22.539469|
gt: 75127 Monoplane, pem_pos=0, final_scores=-0.8574472,

```

final_local_scores=-0.8923747, pem=-0.00019468772,
similarity=-22.539469 ]
6: 36040686 Cody_Circuit_of_Britain_biplane
  [winner: 49201 Biplane, pem_pos=0,
final_scores=-1.319394, final_local_scores=-1.4460118,
pem=-0.00053622795, similarity=-39.059246]
  gt: 36040686 Cody_Circuit_of_Britain_biplane, pem_pos=2,
final_scores=-2.468935, final_local_scores=-2.662266,
pem=-8.629807, similarity=-28.427382 ]
7: 1026521 Pusher_configuration
  [winner: 1026521 Pusher_configuration, pem_pos=0,
final_scores=-0.96819943, final_local_scores=-1.016633,
pem=0.0, similarity=-26.248644]
  gt: 1026521 Pusher_configuration, pem_pos=0,
final_scores=-0.96819943, final_local_scores=-1.016633,
pem=0.0, similarity=-26.248644 ]
8: 9214592 Canard_(aeronautics)
  [winner: 9214439 Antisemitic_canard, pem_pos=1,
final_scores=-1.0671991, final_local_scores=-1.0487107,
pem=-0.6406417, similarity=-23.721998]
  gt: 9214592 Canard_(aeronautics), pem_pos=0,
final_scores=-1.0862606, final_local_scores=-1.1341643,
pem=0.0, similarity=-29.756012 ]
9: 49201 Biplane
  [winner: 49201 Biplane, pem_pos=0,
final_scores=-1.1188073, final_local_scores=-1.2269013,
pem=-0.00053622795, similarity=-32.520554]
  gt: 49201 Biplane, pem_pos=0, final_scores=-1.1188073,
final_local_scores=-1.2269013, pem=-0.00053622795,
similarity=-32.520554 ]
10: 1028296 Tractor_configuration
  [winner: 1028296 Tractor_configuration, pem_pos=1,
final_scores=-1.0531667, final_local_scores=-1.1031804,
pem=-0.9505074, similarity=-23.662384]
  gt: 1028296 Tractor_configuration, pem_pos=1,
final_scores=-1.0531667, final_local_scores=-1.1031804,
pem=-0.9505074, similarity=-23.662384 ]
11: 12101901 Austro-Daimler_6
  [winner: 862162 Austro-Daimler, pem_pos=0,
final_scores=-0.01952158, final_local_scores=0.04854016,
pem=0.0, similarity=5.538248]
  gt: 12101901 Austro-Daimler_6, pem_pos=1,
final_scores=-1.156409, final_local_scores=-1.1226963,
pem=-5.9454207, similarity=2.9183402 ]
12: 950992 Elevator_(aircraft)
  [winner: 950992 Elevator_(aircraft), pem_pos=0,
final_scores=-0.980782, final_local_scores=-0.9661895,
pem=-0.15455846, similarity=-23.902796]
  gt: 950992 Elevator_(aircraft), pem_pos=0,
final_scores=-0.980782, final_local_scores=-0.9661895,
pem=-0.15455846, similarity=-23.902796 ]
13: 39029 Bamboo

```

[winner: 87497 Bambusa, pem_pos=1,
final_scores=-0.8792212, final_local_scores=-0.73279274,
pem=-0.69300616, similarity=-14.009609]
gt: 39029 Bamboo, pem_pos=0, final_scores=-1.1428695,
final_local_scores=-1.1198068, pem=-0.033041302,
similarity=-29.14787]

14: 1647947 Wing_warping

[winner: 1647947 Wing_warping, pem_pos=0,
final_scores=-0.8977725, final_local_scores=-0.90354705,
pem=0.0, similarity=-22.873934]

gt: 1647947 Wing_warping, pem_pos=0,
final_scores=-0.8977725, final_local_scores=-0.90354705,
pem=0.0, similarity=-22.873934]

15: 27040074 Cody_V_biplane

[winner: 27040074 Cody_V_biplane, pem_pos=0,
final_scores=0.044519022, final_local_scores=0.06394112,
pem=0.0, similarity=5.9978437]

gt: 27040074 Cody_V_biplane, pem_pos=0,
final_scores=0.044519022, final_local_scores=0.06394112,
pem=0.0, similarity=5.9978437]

[winner: 49201 Biplane, pem_pos=0,
final_scores=-1.4816164, final_local_scores=-1.6232152,
pem=-0.00053622795, similarity=-44.347347]

RECALL MISS: 27040074 Cody_V_biplane]

1: 233670 Royal_Flying_Corps

[pem_pos=0, final_scores=0.032718077,
final_local_scores=0.00094901025, pem=0.0,
similarity=4.1180334]

2: 250131 Royal_Aircraft_Establishment

[pem_pos=0, final_scores=0.14460514,
final_local_scores=0.104279086, pem=0.0, similarity=7.2016087]

3: 232477 Farnborough, Hampshire

[pem_pos=0, final_scores=0.24447893,
final_local_scores=0.24009219, pem=-0.5432498,
similarity=14.208819]

4: 19104883 Daily_Mail_Circuit_of_Britain_Air_Race

[pem_pos=0, final_scores=0.15115301,
final_local_scores=0.1458632, pem=0.0, similarity=8.442562]

5: 603400 Etrich-Taube

[pem_pos=0, final_scores=0.17339776,
final_local_scores=0.21370727, pem=0.0, similarity=10.467164]

1: 31717 United_Kingdom

[pem_pos=0, final_scores=0.48644167,
final_local_scores=0.5189124, pem=-0.48936236,
similarity=22.236322]

2: 32897971 1912_British_Military_Aeroplane_Competition

[pem_pos=0, final_scores=0.084046856,
final_local_scores=0.07998259, pem=0.0, similarity=6.476553]

3: 274472 War_Office

[pem_pos=0, final_scores=0.13095315,
final_local_scores=0.10698347, pem=0.0, similarity=7.282313]

4: 274472 War_Office

[pem_pos=0, final_scores=0.06868319,
final_local_scores=0.038963005, pem=0.0, similarity=5.2524467]
5: 3434750 United_States
[pem_pos=0, final_scores=0.46969467,
final_local_scores=0.5330577, pem=-0.12994486,
similarity=20.703878]
6: 350278 Samuel_Franklin_Cody
[pem_pos=0, final_scores=0.19927166,
final_local_scores=0.14806888, pem=0.0, similarity=8.508384]
7: 31717 United_Kingdom
[pem_pos=0, final_scores=0.30075443,
final_local_scores=0.31607726, pem=-0.09933331,
similarity=14.062279]
8: 274472 War_Office
[pem_pos=0, final_scores=0.27002913,
final_local_scores=0.25890282, pem=0.0, similarity=11.815889]
9: 350278 Samuel_Franklin_Cody
[pem_pos=0, final_scores=0.1273716,
final_local_scores=0.06952898, pem=0.0, similarity=6.1645966]
10: 350278 Samuel_Franklin_Cody
[pem_pos=0, final_scores=0.25140995,
final_local_scores=0.20342952, pem=0.0, similarity=10.160456]
11: 19104883 Daily_Mail_Circuit_of_Britain_Air_Race
[pem_pos=0, final_scores=0.12933435,
final_local_scores=0.1220296, pem=-0.117783025,
similarity=8.371841]
12: 350278 Samuel_Franklin_Cody
[pem_pos=0, final_scores=0.19119696,
final_local_scores=0.13924849, pem=0.0, similarity=8.245166]
13: 350278 Samuel_Franklin_Cody
[pem_pos=0, final_scores=0.25642174,
final_local_scores=0.20890415, pem=0.0, similarity=10.32383]
14: 19104883 Daily_Mail_Circuit_of_Britain_Air_Race
[pem_pos=0, final_scores=0.097517505,
final_local_scores=0.08727454, pem=-0.117783025,
similarity=7.3346815]
15: 350278 Samuel_Franklin_Cody
[pem_pos=0, final_scores=0.22068296,
final_local_scores=0.1714575, pem=0.0, similarity=9.2063465]
16: 27040074 Cody_V_biplane
[pem_pos=0, final_scores=0.044519022,
final_local_scores=0.06394112, pem=0.0, similarity=5.9978437]

Sample: b'Battle_of_Vila_Velha'

The Battle of Vila Velha or Battle of Vila Velha de Ródão took place in October 1762 when a [1 1 British 1] - [1 1 Portuguese 1] force led by 2 John Burgoyne and 3 Charles Lee surprised and recaptured the town of 4 Vila Velha de Ródão from [2 2 Spanish 2] invaders during the 5 Seven Years ' War as part of the . Burgoyne , who 6 took the [2 Spanish 2] base at Valencia de Alcántara two months earlier then marched against forces preparing to cross the [3 River 7 Tagus 3]

into [3 3 Alentejo 3] . On the 3rd October 1762 in anticipation of a [4 Spanish 4] offensive across the 8 Zêzere River against the [5 Portuguese 5] headquarters at [4 4 Abrantes 4] instructed 9 George Townshend to march to the country , marching along the left bank of the Zêzere river , to make a junction with 10 Lord George Lennox 's forces and to threaten the [6 Spanish 6] lines of communication with 11 Almeida and [5 5 Ciudad Rodrigo 5] by advancing on 12 Belmonte and . This new march was promptly executed and Townshend 's [7 Portuguese 7] soldiers enduring the greatest privations ; but his men successfully attacked a [8 French 8] force escorting a convoy near , capturing a large quantity of supplies . The same day , the [9 Spanish 9] force , who had taken [10 Vila Velha 10] on the 2nd , advanced on , leaving behind eight artillery pieces guarded by 200 grenadiers and 100 horse . General Burgoyne , who was in charge of the defence of the south bank of the [6 6 Tagus 6] in this area , noticed that only a small force was guarding the [11 Spanish 11] battery at [12 Vila Velha 12] and ordered lieutenant-colonel [13 Charles Lee 13] to take the head of a detachment -LRB- 100 [14 Portuguese 14] 13 grenadier s , 200 men of the and 50 men of the -RRB- , to pass the [15 Tagus 15] and to attack this position . A [16 Portuguese 16] grenadier of the 2nd Cascais Infantry courageously crossed the [17 Tagus 17] with a rope to facilitate the passage of a barge , sacrificing his life in this action . On 7 October , Lee 's detachment , using surprise , launched a night attack and surprised the [18 Spanish 18] camp at Vila Velha . Although they were entrenched surprise helped the [19 British 19] and [20 Portuguese 20] to manage to overpower and disperse the [21 Spanish 21] horse and foot inflicting considerable loss . As well as 250 [22 Spanish 22] killed wounded or captured , 6 guns and 60 artillery mules were acquired . Lee 's main target however was the artillery depot and this was mostly burned , some of it being taken . This was achieved at a cost of only 1 man killed and 10 wounded . The town was held as [23 Portuguese 23] and [24 British 24] reinforcements arrived the next day . The taking of [25 Vila Velha 25] effectively ended the [26 Spanish 26] and [27 French 27] invasion campaign By the middle of October the Franco-Spanish army decided to retreat back into [28 Spain 28] with Lippe and Townshend pursuing them . The [29 Spanish 29] were forced to leave their sick behind and by 24 October the [30 Spanish 30] and [31 French 31] army were back in [32 Spain 32] . There were a few skirmishes along the border as both sides went into Winter quarters but the [33 Spanish 33] would try once more in November at 14 Olivença and but they were repelled . On 22 November the [34 Spanish 34] commander 15 Count of Aranda

1: 23033 Portugal

2: 26667 Spain

3: 1185527 Alentejo

4: 24718975 Abrantes

5: 1930312 Ciudad_Rodrigo

6: 64205 Tagus

1: 13530298 Great_Britain

[winner: 31717 United_Kingdom, pem_pos=0,
final_scores=0.4310462, final_local_scores=0.4395495,
pem=-0.48936236, similarity=19.867975|

gt: 13530298 Great_Britain, pem_pos=2,
final_scores=-0.35832483, final_local_scores=-0.4422023,
pem=-2.6425416, similarity=5.2640667]

2: 163208 John_Burgoyne

[winner: 163208 John_Burgoyne, pem_pos=0,
final_scores=-0.03459926, final_local_scores=0.0070275664,
pem=0.0, similarity=4.2994294|

gt: 163208 John_Burgoyne, pem_pos=0,
final_scores=-0.03459926, final_local_scores=0.0070275664,
pem=0.0, similarity=4.2994294]

3: 52285 Charles_Lee_(general)

[winner: 52285 Charles_Lee_(general), pem_pos=0,
final_scores=0.022018656, final_local_scores=0.09350881,
pem=-0.16957131, similarity=7.802357|

gt: 52285 Charles_Lee_(general), pem_pos=0,
final_scores=0.022018656, final_local_scores=0.09350881,
pem=-0.16957131, similarity=7.802357]

4: 2052395 Vila_Velha_de_Ródão_Municipality

[winner: 2052395 Vila_Velha_de_Ródão_Municipality,
pem_pos=0, final_scores=0.024086788,
final_local_scores=0.1299164, pem=-0.45050585,
similarity=10.416598|

gt: 2052395 Vila_Velha_de_Ródão_Municipality, pem_pos=0,
final_scores=0.024086788, final_local_scores=0.1299164,
pem=-0.45050585, similarity=10.416598]

5: 19039354 Seven_Years'_War

[winner: 19039354 Seven_Years'_War, pem_pos=0,
final_scores=-0.52118003, final_local_scores=-0.56460977,
pem=0.0, similarity=-12.759368|

gt: 19039354 Seven_Years'_War, pem_pos=0,
final_scores=-0.52118003, final_local_scores=-0.56460977,
pem=0.0, similarity=-12.759368]

6: 30706695 Battle_of_Valencia_de_Alcántara

[winner: 26667 Spain, pem_pos=2, final_scores=0.5911957,
final_local_scores=0.53404367, pem=-1.2425895,
similarity=26.784037|

RECALL MISS: 30706695 Battle_of_Valencia_de_Alcántara]

[winner: 1965146 Valencia_de_Alcántara, pem_pos=0,
final_scores=-0.1171336, final_local_scores=-0.071078,
pem=-0.4202802, similarity=4.254153|

gt: 30706695 Battle_of_Valencia_de_Alcántara, pem_pos=1,
final_scores=-0.25115818, final_local_scores=-0.17276552,
pem=-1.0696247, similarity=4.75083]

[winner: 16069942 Battle_of_Thouars, pem_pos=1,
final_scores=-1.1189581, final_local_scores=-1.0832036,
pem=-0.68726486, similarity=-24.497787|

RECALL MISS: 30706695 Battle_of_Valencia_de_Alcántara]
[winner: 140459 Base_(chemistry), pem_pos=0,
final_scores=-1.1945546, final_local_scores=-1.1321709,
pem=-0.75725174, similarity=-25.578474]

RECALL MISS: 30706695 Battle_of_Valencia_de_Alcántara]
7: 64205 Tagus

[winner: 64205 Tagus, pem_pos=0, final_scores=0.35377002,
final_local_scores=0.37273014, pem=0.0, similarity=15.212724]

gt: 64205 Tagus, pem_pos=0, final_scores=0.35377002,
final_local_scores=0.37273014, pem=0.0, similarity=15.212724]

8: 2832319 Zêzere_River

[winner: 2832319 Zêzere_River, pem_pos=0,
final_scores=0.02906935, final_local_scores=0.12566283,
pem=0.0, similarity=7.8397427]

gt: 2832319 Zêzere_River, pem_pos=0,
final_scores=0.02906935, final_local_scores=0.12566283,
pem=0.0, similarity=7.8397427]

9: 622283 George_Townshend,_1st_Marquess_Townshend

[winner: 4072829 George_Townshend_(Bahá'í), pem_pos=0,
final_scores=-0.09664652, final_local_scores=0.01999849,
pem=-0.14770058, similarity=5.4897265]

gt: 622283 George_Townshend,_1st_Marquess_Townshend,
pem_pos=1, final_scores=-0.3929624,
final_local_scores=-0.32102942, pem=-1.2820709,
similarity=1.4816556]

10: 4601212 Lord_George_Lennox

[winner: 4601212 Lord_George_Lennox, pem_pos=0,
final_scores=-0.113368094, final_local_scores=-0.011586249,
pem=0.0, similarity=3.743956]

gt: 4601212 Lord_George_Lennox, pem_pos=0,
final_scores=-0.113368094, final_local_scores=-0.011586249,
pem=0.0, similarity=3.743956]

11: 24737419 Almeida,_Portugal

[winner: 24737419 Almeida,_Portugal, pem_pos=3,
final_scores=-0.33472857, final_local_scores=-0.31618083,
pem=-2.114299, similarity=6.1521335]

gt: 24737419 Almeida,_Portugal, pem_pos=3,
final_scores=-0.33472857, final_local_scores=-0.31618083,
pem=-2.114299, similarity=6.1521335]

12: 4238819 Belmonte_Municipality

[winner: 4238819 Belmonte_Municipality, pem_pos=1,
final_scores=-0.16466537, final_local_scores=-0.13484868,
pem=-1.526117, similarity=8.3648205]

gt: 4238819 Belmonte_Municipality, pem_pos=1,
final_scores=-0.16466537, final_local_scores=-0.13484868,
pem=-1.526117, similarity=8.3648205]

13: 71349 Grenadier

[winner: 9223962 Violet-eared_Waxbill, pem_pos=8,
final_scores=-1.0160204, final_local_scores=-0.9075641,
pem=-2.3978953, similarity=-9.953683]

gt: 71349 Grenadier, pem_pos=0, final_scores=-1.1028842,
final_local_scores=-1.1644497, pem=-0.39638236,

```

similarity=-28.504196 ]
14: 305819 Olivenza
    [winner: 305819 Olivenza, pem_pos=1,
final_scores=-0.09991978, final_local_scores=-0.07967459,
pem=-1.1638945, similarity=8.041504]
    gt: 305819 Olivenza, pem_pos=1,
final_scores=-0.09991978, final_local_scores=-0.07967459,
pem=-1.1638945, similarity=8.041504 ]
15: 2114946 Pedro_Pablo_Abarca_de_Bolea,_10th_Count_of_Aranda
    [winner: 2114946
Pedro_Pablo_Abarca_de_Bolea,_10th_Count_of_Aranda, pem_pos=0,
final_scores=-0.07001999, final_local_scores=-0.014480501,
pem=0.0, similarity=3.6575859]
    gt: 2114946
Pedro_Pablo_Abarca_de_Bolea,_10th_Count_of_Aranda, pem_pos=0,
final_scores=-0.07001999, final_local_scores=-0.014480501,
pem=0.0, similarity=3.6575859 ]
1: 23033 Portugal
    [pem_pos=1, final_scores=0.5079793,
final_local_scores=0.45478076, pem=-1.127839,
similarity=23.794641]
2: 26667 Spain
    [pem_pos=2, final_scores=0.71635836,
final_local_scores=0.67076486, pem=-1.2425895,
similarity=30.86407]
3: 1185527 Alentejo
    [pem_pos=0, final_scores=0.30257052,
final_local_scores=0.29739153, pem=0.0, similarity=12.96447]
4: 24718975 Abrantes
    [pem_pos=0, final_scores=0.12761088,
final_local_scores=0.15993387, pem=0.0, similarity=8.862458]
5: 1930312 Ciudad_Rodrigo
    [pem_pos=0, final_scores=0.1341977,
final_local_scores=0.17821449, pem=0.0, similarity=9.407989]
6: 64205 Tagus
    [pem_pos=0, final_scores=0.32363802,
final_local_scores=0.33981553, pem=0.0, similarity=14.230486]
1: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4310462,
final_local_scores=0.4395495, pem=-0.48936236,
similarity=19.867975]
2: 26667 Spain
    [pem_pos=2, final_scores=0.5911957,
final_local_scores=0.53404367, pem=-1.2425895,
similarity=26.784037]
3: 64205 Tagus
    [pem_pos=0, final_scores=0.35377002,
final_local_scores=0.37273014, pem=0.0, similarity=15.212724]
4: 26667 Spain
    [pem_pos=2, final_scores=0.69500357,
final_local_scores=0.647438, pem=-1.2425895,
similarity=30.16795]

```


5: 23033 Portugal
[pem_pos=1, final_scores=0.5323171,
final_local_scores=0.48136616, pem=-1.127839,
similarity=24.588003]
6: 26667 Spain
[pem_pos=2, final_scores=0.6530015,
final_local_scores=0.6015571, pem=-1.2425895,
similarity=28.79877]
7: 23033 Portugal
[pem_pos=1, final_scores=0.5472451,
final_local_scores=0.49767268, pem=-1.127839,
similarity=25.074623]
8: 5843419 France
[pem_pos=0, final_scores=0.6807918,
final_local_scores=0.6737999, pem=-0.80223274,
similarity=28.559912]
9: 26667 Spain
[pem_pos=2, final_scores=0.5798668,
final_local_scores=0.5216686, pem=-1.2425895,
similarity=26.41474]
10: 630943 Vila_Velha
[pem_pos=0, final_scores=0.07713996,
final_local_scores=0.14211679, pem=0.0, similarity=8.330761]
11: 26667 Spain
[pem_pos=2, final_scores=0.6513986,
final_local_scores=0.59980625, pem=-1.2425895,
similarity=28.746523]
12: 630943 Vila_Velha
[pem_pos=0, final_scores=0.053907648,
final_local_scores=0.116739005, pem=0.0, similarity=7.5734377]
13: 52285 Charles_Lee_(general)
[pem_pos=0, final_scores=0.07729195,
final_local_scores=0.1538865, pem=-0.16957131,
similarity=9.604148]
14: 23033 Portugal
[pem_pos=1, final_scores=0.47521377,
final_local_scores=0.41898942, pem=-1.127839,
similarity=22.726557]
15: 64205 Tagus
[pem_pos=0, final_scores=0.34546512,
final_local_scores=0.3636583, pem=0.0, similarity=14.942002]
16: 23033 Portugal
[pem_pos=1, final_scores=0.5022748,
final_local_scores=0.44854945, pem=-1.127839,
similarity=23.608688]
17: 64205 Tagus
[pem_pos=0, final_scores=0.43548757,
final_local_scores=0.4619941, pem=0.0, similarity=17.87654]
18: 26667 Spain
[pem_pos=2, final_scores=0.6292357,
final_local_scores=0.57559663, pem=-1.2425895,
similarity=28.02406]

19: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5169422,
final_local_scores=0.5333778, pem=-0.48936236,
similarity=22.668]
20: 23033 Portugal
[pem_pos=1, final_scores=0.37471265,
final_local_scores=0.3092072, pem=-1.127839,
similarity=19.450438]
21: 26667 Spain
[pem_pos=2, final_scores=0.68074995,
final_local_scores=0.6318681, pem=-1.2425895,
similarity=29.703314]
22: 26667 Spain
[pem_pos=2, final_scores=0.621695,
final_local_scores=0.5673595, pem=-1.2425895,
similarity=27.778248]
23: 23033 Portugal
[pem_pos=1, final_scores=0.46472383,
final_local_scores=0.40753073, pem=-1.127839,
similarity=22.384607]
24: 31717 United_Kingdom
[pem_pos=0, final_scores=0.42410868,
final_local_scores=0.43197137, pem=-0.48936236,
similarity=19.641829]
25: 630943 Vila_Velha
[pem_pos=0, final_scores=0.049849376,
final_local_scores=0.11230597, pem=0.0, similarity=7.441147]
26: 26667 Spain
[pem_pos=2, final_scores=0.606726,
final_local_scores=0.55100816, pem=-1.2425895,
similarity=27.29029]
27: 5843419 France
[pem_pos=0, final_scores=0.6465029,
final_local_scores=0.6363445, pem=-0.80223274,
similarity=27.44217]
28: 26667 Spain
[pem_pos=0, final_scores=0.6586434,
final_local_scores=0.607126, pem=-0.18857051,
similarity=23.23304]
29: 26667 Spain
[pem_pos=2, final_scores=0.56447476,
final_local_scores=0.50485516, pem=-1.2425895,
similarity=25.912992]
30: 26667 Spain
[pem_pos=2, final_scores=0.5588697,
final_local_scores=0.49873245, pem=-1.2425895,
similarity=25.730278]
31: 5843419 France
[pem_pos=0, final_scores=0.62343025,
final_local_scores=0.61114115, pem=-0.80223274,
similarity=26.69005]
32: 26667 Spain

[pem_pos=0, final_scores=0.6176092,
final_local_scores=0.56230235, pem=-0.18857051,
similarity=21.895412]
33: 26667 Spain
[pem_pos=2, final_scores=0.5143407,
final_local_scores=0.45009124, pem=-1.2425895,
similarity=24.278728]
34: 26667 Spain
[pem_pos=2, final_scores=0.6341879,
final_local_scores=0.58100605, pem=-1.2425895,
similarity=28.185488]

Sample: b'Taifa_of_Badajoz'

The Taifa of Badajoz -LRB- from -RRB- was a medieval [1 1 Muslim 1] kingdom in what is now parts of [1 1 Portugal 1] and [2 Spain 2] and centred on the city of [2 2 Badajoz 2] which exists today as the first city of [3 3 Extremadura 3] , in [4 Spain 4] . It rose , like the other [5 5 Taifa 5] kingdoms of the [6 6 Iberian Peninsula 6] , after the fragmentation of [7 7 Al-Andalus 7] -LRB- the -RRB- in the late 10th and early 11th centuries , and ruled by the [3 2 Aftasid 3] dynasty . It was created by Sabur , a former slave of perhaps Slavic origin . The taifa 's control extended over most of the ancient [4 3 Lusitania 4] , including [5 4 Mérida 5] and [8 8 Lisbon 8] . Sabur was succeeded in 1022 by his vizier , , who founded a dynasty . Sabur 's sons fled to [6 Lisbon 6] , where they created the short-lived 5 taifa of [7 Lisbon 7] , which was soon reconquered by [8 Badajoz 8] . In 1055 [9 Badajoz 9] became a tributary of the , losing significant parts of its territory , south of the [10 Mondego 10] river -LRB- south of [11 Coimbra 11] -RRB- . The 6 Abbadids of [12 Seville 12] also conquered parts of their territory . After the death of Abdallah 's son , [13 Abu Bakr 13] , a civil war broke out between the latter 's sons , Yahya and Abu , the former being victorious . His troops fought alongside the against the Christian army in the 7 battle of [14 Sagradas 14] -LRB- 1086 -RRB- , which occurred not far from [15 Badajoz 15] . However , after the [16 Almoravid 16] victory , Yahya , who feared their increasing power , allied with . In 1094 the [17 Almoravids 17] occupied [18 Badajoz 18] and [19 Yahya 19] was killed together with two of his sons . A surviving son fled first to [9 9 Montánchez 9] and then to Alfonso 's court . After the taifa 's original territory had been ruled by various kingdoms -LRB- Almoravids , s , [20 8 Portugal 20]

1: 23033 Portugal
2: 58480 Badajoz
3: 68204 Extremadura
4: 26667 Spain
5: 453117 Taifa
6: 14883 Iberian_Peninsula
7: 355643 Al-Andalus
8: 18091 Lisbon

9: 11572905 Montánchez

1: 45127 Moors

[winner: 19541 Muslim, pem_pos=0,
final_scores=0.109887525, final_local_scores=0.106111765,
pem=-0.22736222, similarity=8.492731|
gt: 45127 Moors, pem_pos=10, final_scores=-0.21902047,
final_local_scores=-0.29306424, pem=-3.4653811,
similarity=14.189374]

2: 2166509 Aftasid_dynasty

[winner: 2166509 Aftasid_dynasty, pem_pos=0,
final_scores=0.112631485, final_local_scores=0.18112102,
pem=0.0, similarity=9.494725|
gt: 2166509 Aftasid_dynasty, pem_pos=0,
final_scores=0.112631485, final_local_scores=0.18112102,
pem=0.0, similarity=9.494725]

[winner: 27158897 Dynasty (Kaskade_album), pem_pos=7,
final_scores=-1.4576738, final_local_scores=-1.3609858,
pem=-3.3888543, similarity=-18.095701|

RECALL MISS: 2166509 Aftasid_dynasty]

3: 18577 Lusitania

[winner: 436113 RMS_Lusitania, pem_pos=0,
final_scores=0.1759079, final_local_scores=0.2704555,
pem=-0.40260512, similarity=14.350074|
gt: 18577 Lusitania, pem_pos=1, final_scores=0.15098314,
final_local_scores=0.14298204, pem=-0.49355015,
similarity=11.040585]

4: 68683 Mérida, Spain

[winner: 68682 Mérida, Yucatán, pem_pos=0,
final_scores=0.11778234, final_local_scores=0.102483466,
pem=-1.0951699, similarity=13.103728|
gt: 68683 Mérida, Spain, pem_pos=1,
final_scores=0.030371353, final_local_scores=-0.023710132,
pem=-1.578716, similarity=11.967459]

5: 19267694 Taifa_of_Lisbon

[winner: 18091 Lisbon, pem_pos=0, final_scores=0.5048661,
final_local_scores=0.49057943, pem=-0.09749677,
similarity=19.259785|

RECALL MISS: 19267694 Taifa_of_Lisbon]

[winner: 453117 Taifa, pem_pos=0,
final_scores=-1.1053572, final_local_scores=-1.231771,
pem=0.0, similarity=-32.66879|

RECALL MISS: 19267694 Taifa_of_Lisbon]

6: 15020925 Taifa_of_Seville

[winner: 37770 Seville, pem_pos=0,
final_scores=0.40173298, final_local_scores=0.35152495,
pem=-0.12539493, similarity=15.261835|
gt: 15020925 Taifa_of_Seville, pem_pos=6,
final_scores=-0.4703101, final_local_scores=-0.52412456,
pem=-2.7871034, similarity=3.605492]

[winner: 49865 Abbadid_dynasty, pem_pos=0,
final_scores=0.009156391, final_local_scores=0.018819124,
pem=0.0, similarity=4.6513133|

RECALL MISS: 15020925 Taifa_of_Seville]

7: 13650140 Battle_of_Sagrajas
 [winner: 13650140 Battle_of_Sagrajas, pem_pos=0,
 final_scores=0.052525952, final_local_scores=0.066180244,
 pem=0.0, similarity=6.0646634]
 gt: 13650140 Battle_of_Sagrajas, pem_pos=0,
 final_scores=0.052525952, final_local_scores=0.066180244,
 pem=0.0, similarity=6.0646634]
 [winner: 14659603 Battle_of_Beth_Horon_(66), pem_pos=0,
 final_scores=-0.97429675, final_local_scores=-0.9712123,
 pem=0.0, similarity=-24.8932]

RECALL MISS: 13650140 Battle_of_Sagrajas]

8: 9352542 Kingdom_of_Portugal
 [winner: 23033 Portugal, pem_pos=0,
 final_scores=0.5846875, final_local_scores=0.58659285,
 pem=-0.30942732, similarity=23.277527]
 gt: 9352542 Kingdom_of_Portugal, pem_pos=2,
 final_scores=-0.551944, final_local_scores=-0.64697874,
 pem=-3.9057226, similarity=6.022504]

1: 23033 Portugal
 [pem_pos=0, final_scores=0.58290684,
 final_local_scores=0.58464783, pem=-0.30942732,
 similarity=23.219484]

2: 58480 Badajoz
 [pem_pos=0, final_scores=0.26479113,
 final_local_scores=0.21925002, pem=-0.34104064,
 similarity=12.487203]

3: 68204 Extremadura
 [pem_pos=0, final_scores=0.27942675,
 final_local_scores=0.25839123, pem=0.0, similarity=11.800623]

4: 26667 Spain
 [pem_pos=0, final_scores=0.63013446,
 final_local_scores=0.6301355, pem=-0.18857051,
 similarity=23.919687]

5: 453117 Taifa
 [pem_pos=0, final_scores=0.12379949,
 final_local_scores=0.11714277, pem=0.0, similarity=7.5854864]

6: 14883 Iberian_Peninsula
 [pem_pos=0, final_scores=0.2523945,
 final_local_scores=0.24520317, pem=0.0, similarity=11.407064]

7: 355643 Al-Andalus
 [pem_pos=0, final_scores=0.28911835,
 final_local_scores=0.27133295, pem=0.0, similarity=12.18683]

8: 18091 Lisbon
 [pem_pos=0, final_scores=0.5594743,
 final_local_scores=0.5502305, pem=-0.09749677,
 similarity=21.039892]

9: 11572905 Montánchez
 [pem_pos=0, final_scores=0.103965744,
 final_local_scores=0.15007862, pem=0.0, similarity=8.568358]

1: 19541 Muslim
 [pem_pos=0, final_scores=0.109887525,

final_local_scores=0.106111765, pem=-0.22736222,
similarity=8.492731]
2: 26667 Spain
[pem_pos=0, final_scores=0.58416164,
final_local_scores=0.57991713, pem=-0.18857051,
similarity=22.421072]
3: 2166509 Aftasid_dynasty
[pem_pos=0, final_scores=0.112631485,
final_local_scores=0.18112102, pem=0.0, similarity=9.494725]
4: 436113 RMS_Lusitania
[pem_pos=0, final_scores=0.1759079,
final_local_scores=0.2704555, pem=-0.40260512,
similarity=14.350074]
5: 68682 Mérida, Yucatán
[pem_pos=0, final_scores=0.11778234,
final_local_scores=0.102483466, pem=-1.0951699,
similarity=13.103728]
6: 18091 Lisbon
[pem_pos=0, final_scores=0.6088571,
final_local_scores=0.60417384, pem=-0.09749677,
similarity=22.649668]
7: 18091 Lisbon
[pem_pos=0, final_scores=0.5048661,
final_local_scores=0.49057943, pem=-0.09749677,
similarity=19.259785]
8: 58480 Badajoz
[pem_pos=0, final_scores=0.27581698,
final_local_scores=0.23129404, pem=-0.34104064,
similarity=12.846621]
9: 58480 Badajoz
[pem_pos=0, final_scores=0.21151279,
final_local_scores=0.1608223, pem=-0.34104064,
similarity=10.743603]
10: 879881 Mondego_River
[pem_pos=0, final_scores=0.16886453,
final_local_scores=0.20450133, pem=0.0, similarity=10.192441]
11: 378109 Coimbra
[pem_pos=0, final_scores=0.28700286,
final_local_scores=0.25399145, pem=-0.026842317,
similarity=11.815297]
12: 37770 Seville
[pem_pos=0, final_scores=0.40173298,
final_local_scores=0.35152495, pem=-0.12539493,
similarity=15.261835]
13: 1715 Abu_Bakr
[pem_pos=0, final_scores=0.22540091,
final_local_scores=0.2801339, pem=0.0, similarity=12.449468]
14: 13650140 Battle_of_Sagrajas
[pem_pos=0, final_scores=0.052525952,
final_local_scores=0.066180244, pem=0.0, similarity=6.0646634]
15: 58480 Badajoz
[pem_pos=0, final_scores=0.3111506,

```

final_local_scores=0.2698907, pem=-0.34104064,
similarity=13.998422]
16: 1649 Almoravid_dynasty
    [pem_pos=0, final_scores=0.20687,
final_local_scores=0.1788168, pem=0.0, similarity=9.425962]
17: 1649 Almoravid_dynasty
    [pem_pos=0, final_scores=0.23153953,
final_local_scores=0.2057645, pem=0.0, similarity=10.230136]
18: 58480 Badajoz
    [pem_pos=0, final_scores=0.28000772,
final_local_scores=0.23587182, pem=-0.34104064,
similarity=12.983231]
19: 16125 John_the_Baptist
    [pem_pos=0, final_scores=0.0597281,
final_local_scores=0.07579246, pem=-0.63666075,
similarity=9.813771]
20: 23033 Portugal
    [pem_pos=0, final_scores=0.5846875,
final_local_scores=0.58659285, pem=-0.30942732,
similarity=23.277527]

```

Sample: b'2011_Valencian_Community_motorcycle_Grand_Prix'

The 2011 [1 Valencian Community 1] motorcycle Grand Prix - LRB- formally the Gran Premio Generali de la Comunitat Valenciana -RRB- was the last round of the 1 2011 Grand Prix m otorcycle racing season . It took place on the weekend of November 4 -- 6 , 2011 at the . It was the first race after the death of [1 1 Marco Simoncelli 1] in 2 Sepang . It was the final ever 125cc race before [2 Moto3 2] -LRB- 250cc - RRB- replaced it in 3 2012 . On Saturday , [2 2 Stefan Bradl 2] officially became Moto2 World Champion after his only rival , [3 3 Marc Márquez 3] , could not go on the track during the free practices and the qualifying session , because of injuries sustained in [3 4 Malaysia 3] . [4 4 Colin Edwards 4] had an injured shoulder and was replaced by 5 2011 [5 5 AMA Pro Superbike 5] champion [6 6 Josh Hayes 6] on the 6 Yamaha 7 Tech 3 and [7 7 Jorge Lorenzo 7] was replaced for the second time by test-rider [8 8 Katsuyuki Nakasuga 8] . [9 9 Loris Capirossi 9] has now retired from 8 motorcycle racing following the race . Although [10 10 Gresini Racing 10] originally pulled out of the [4 Valencia 4] race after [5 Simoncelli 5] 's fatal accident , it was later decided that the team would race . As tribute to [6 Marco Simoncelli 6] , [7 Loris Capirossi 7] , in his final [8 MotoGP 8] race , participated wearing [9 Simoncelli 9] 's racing number 58 instead of his usual racing number 65 , although on the official reports he was still listed as # 65 . All riders from [10 MotoGP 10] , [11 Moto2 11] and the 125cc races participated in a parade lap on Sunday morning after the warm-up session with 1993 500cc World Champion [11 11 Kevin Schwantz 11] riding [12 Simoncelli 12] 's bike . [12 12 Valentino Rossi 12]

```

1: 14084208 Marco_Simoncelli
2: 16743573 Stefan_Bradl
3: 16908657 Marc_Márquez
4: 2224069 Colin_Edwards
5: 1182972 AMA_Superbike_Championship
6: 8461621 Josh_Hayes
7: 5934401 Jorge_Lorenzo
8: 35431084 Katsuyuki_Nakasuga
9: 2721014 Loris_Capirossi
10: 5586054 Gresini_Racing
11: 1861396 Kevin_Schwantz
12: 180306 Valentino_Rossi
1: 26781017 2011_Grand_Prix_motorcycle_racing_season
    [winner: 54221 Grand_Prix_motorcycle_racing, pem_pos=1,
    final_scores=-0.0756392, final_local_scores=-0.14842418,
    pem=-1.5994394, similarity=8.358439|
    RECALL MISS: 26781017
2011_Grand_Prix_motorcycle_racing_season]
    [winner: 36225 2011, pem_pos=0, final_scores=-1.1834015,
    final_local_scores=-1.1002133, pem=-1.4232541,
    similarity=-21.002972|
    RECALL MISS: 26781017
2011_Grand_Prix_motorcycle_racing_season]
    [winner: 2446333 2005-06_A1_Grand_Prix_season, pem_pos=0,
    final_scores=-1.2891139, final_local_scores=-1.4031193,
    pem=0.0, similarity=-37.782166|
    RECALL MISS: 26781017
2011_Grand_Prix_motorcycle_racing_season]
    [winner: 2732661 Police_motorcycle, pem_pos=1,
    final_scores=-1.5650454, final_local_scores=-1.6005497,
    pem=-1.7721045, similarity=-34.036903|
    RECALL MISS: 26781017
2011_Grand_Prix_motorcycle_racing_season]
2: 944336 Sepang_International_Circuit
    [winner: 468544 Sepang, pem_pos=0,
    final_scores=0.0264913, final_local_scores=0.09403238,
    pem=-0.02650924, similarity=7.039988|
    gt: 944336 Sepang_International_Circuit, pem_pos=1,
    final_scores=-0.039090335, final_local_scores=-0.06312992,
    pem=-0.8799167, similarity=6.9909143 ]
3: 27948565 2012_Grand_Prix_motorcycle_racing_season
    [winner: 47374 2012, pem_pos=0, final_scores=-0.7054245,
    final_local_scores=-0.5193306, pem=-1.4461981,
    similarity=-3.5434992|
    RECALL MISS: 27948565
2012_Grand_Prix_motorcycle_racing_season]
4: 31308466 2011_Malaysian_motorcycle_Grand_Prix
    [winner: 3607937 Malaysia, pem_pos=0,
    final_scores=0.5604248, final_local_scores=0.7102911,
    pem=-0.20223404, similarity=26.385996|
    RECALL MISS: 31308466
2011_Malaysian_motorcycle_Grand_Prix]

```


5: 29545573
2011_AMA_Pro_American_Superbike_Championship_season
[winner: 36225 2011, pem_pos=0, final_scores=-1.1221441,
final_local_scores=-1.033299, pem=-1.4232541,
similarity=-19.006111]
RECALL MISS: 29545573
2011_AMA_Pro_American_Superbike_Championship_season]
6: 332378 Yamaha_Motor_Company
[winner: 14562867 Tech_3, pem_pos=0,
final_scores=-0.06067066, final_local_scores=-0.08962482,
pem=0.0, similarity=1.4151294]
RECALL MISS: 332378 Yamaha_Motor_Company]
7: 14562867 Tech_3
[winner: 14562867 Tech_3, pem_pos=0,
final_scores=-0.06067066, final_local_scores=-0.08962482,
pem=0.0, similarity=1.4151294]
gt: 14562867 Tech_3, pem_pos=0,
final_scores=-0.06067066, final_local_scores=-0.08962482,
pem=0.0, similarity=1.4151294]
8: 7501246 Motorcycle_racing
[winner: 2732661 Police_motorcycle, pem_pos=1,
final_scores=-1.0775735, final_local_scores=-1.0680608,
pem=-1.7721045, similarity=-18.14637]
gt: 7501246 Motorcycle_racing, pem_pos=8,
final_scores=-2.1252325, final_local_scores=-2.3523057,
pem=-5.7287316, similarity=-34.95404]
[winner: 1900663 Canoe_racing, pem_pos=6,
final_scores=-1.3379167, final_local_scores=-1.3764732,
pem=-2.4332824, similarity=-23.754427]
RECALL MISS: 7501246 Motorcycle_racing]
1: 14084208 Marco_Simoncelli
[pem_pos=0, final_scores=0.56866306,
final_local_scores=0.5461505, pem=0.0, similarity=20.387934]
2: 16743573 Stefan_Bradl
[pem_pos=0, final_scores=0.47480136,
final_local_scores=0.4980091, pem=0.0, similarity=18.951298]
3: 16908657 Marc_Márquez
[pem_pos=0, final_scores=0.50752336,
final_local_scores=0.5179261, pem=0.0, similarity=19.54566]
4: 2224069 Colin_Edwards
[pem_pos=0, final_scores=0.43782753,
final_local_scores=0.41362637, pem=0.0, similarity=16.43315]
5: 1182972 AMA_Superbike_Championship
[pem_pos=0, final_scores=0.2853443,
final_local_scores=0.30512533, pem=0.0, similarity=13.195261]
6: 8461621 Josh_Hayes
[pem_pos=0, final_scores=0.29033285,
final_local_scores=0.32616785, pem=0.0, similarity=13.823213]
7: 5934401 Jorge_Lorenzo
[pem_pos=0, final_scores=0.30399853,
final_local_scores=0.27056086, pem=0.0, similarity=12.163789]
8: 35431084 Katsuyuki_Nakasuga

[pem_pos=0, final_scores=0.41599107,
final_local_scores=0.484865, pem=0.0, similarity=18.559052]
9: 2721014 Loris_Capirossi
[pem_pos=0, final_scores=0.55861896,
final_local_scores=0.5580742, pem=0.0, similarity=20.74376]
10: 5586054 Gresini_Racing
[pem_pos=0, final_scores=0.18459766,
final_local_scores=0.16690782, pem=0.0, similarity=9.070575]
11: 1861396 Kevin_Schwantz
[pem_pos=0, final_scores=0.68391734,
final_local_scores=0.7117359, pem=0.0, similarity=25.329332]
12: 180306 Valentino_Rossi
[pem_pos=0, final_scores=0.4634797,
final_local_scores=0.43618542, pem=0.0, similarity=17.106356]
1: 20514168 Valencian_Community
[pem_pos=0, final_scores=0.0845993,
final_local_scores=0.18408978, pem=0.0, similarity=9.583319]
2: 54221 Grand_Prix_motorcycle_racing
[pem_pos=0, final_scores=0.17926799,
final_local_scores=0.13604, pem=-0.18232158,
similarity=9.140911]
3: 3607937 Malaysia
[pem_pos=0, final_scores=0.5604248,
final_local_scores=0.7102911, pem=-0.20223404,
similarity=26.385996]
4: 63861 Valencia
[pem_pos=0, final_scores=0.16183643,
final_local_scores=0.21760228, pem=-0.9876697,
similarity=15.9545]
5: 14084208 Marco_Simoncelli
[pem_pos=0, final_scores=0.33805686,
final_local_scores=0.29262668, pem=0.0, similarity=12.822277]
6: 14084208 Marco_Simoncelli
[pem_pos=0, final_scores=0.45155925,
final_local_scores=0.41823244, pem=0.0, similarity=16.570602]
7: 2721014 Loris_Capirossi
[pem_pos=0, final_scores=0.432719,
final_local_scores=0.4205476, pem=0.0, similarity=16.639692]
8: 54221 Grand_Prix_motorcycle_racing
[pem_pos=0, final_scores=0.12610729,
final_local_scores=0.07797001, pem=-0.026200628,
similarity=6.5589767]
9: 14084208 Marco_Simoncelli
[pem_pos=0, final_scores=0.2845772,
final_local_scores=0.23582989, pem=0.0, similarity=11.127347]
10: 54221 Grand_Prix_motorcycle_racing
[pem_pos=0, final_scores=0.25795346,
final_local_scores=0.22199193, pem=-0.026200628,
similarity=10.856877]
11: 20935975 2010_Grand_Prix_motorcycle_racing_season
[pem_pos=0, final_scores=0.15890168,
final_local_scores=0.13384804, pem=-0.24116208,

```
similarity=9.395483]
12: 14084208 Marco_Simoncelli
    [pem_pos=0, final_scores=0.31436825,
final_local_scores=0.2667504, pem=0.0, similarity=12.0500765]
```

Sample: b'Rio_Grande_Association'

The [1 Rio Grande Association 1] was a Class-D minor baseball league that lasted for less than one season , 1915 . The league initially consisted of six teams based in [1 1 Texas 1] , [2 2 Arizona 2] and [3 3 New Mexico 3] : the [4 4 Albuquerque Dukes 4] , , , , 1 Phoenix Senators and . By May 24 , the Miners and Farmers had disbanded and on July 6 , the league disbanded . The league was founded by [2 2 John McCloskey 2] , who also founded the [5 5 Texas League 5] . Each team that appeared in the [3 Rio Grande Association 3] was the first professional baseball team to come out of their respective cities , often predating their successors by many years . For example , after its appearance in the Association , [6 6 Phoenix 6] and [7 7 Tucson 7] did not again have a professional team until 1928 ; [8 8 El Paso 8] until 1930 ; [9 9 Albuquerque 9] until 1932 ; 3 Douglas until 1948 and [10 10 Las Cruces 10] until 2010 . With a record of 38-21 , the [4 Senators 4] were the de facto league champions , finishing in first place . The league featured multiple notable players and managers , including those who would attain 4 major league experience and those who had major league experience . They include [11 11 Kitty Brashear 11] , [12 12 Frank Huelsman 12] -LRB- who led the league with ten home runs -RRB- , [13 13 Rudy Kallio 13] , 5 Stoney McGlynn and 6 Herb Hall

```
1: 29810 Texas
2: 21883824 Arizona
3: 21649 New_Mexico
4: 1398193 Albuquerque_Dukes
5: 70292 Texas_League
6: 49121 Phoenix,_Arizona
7: 57180 Tucson,_Arizona
8: 135823 El_Paso,_Texas
9: 51278 Albuquerque,_New_Mexico
10: 125874 Las_Cruces,_New_Mexico
11: 30612207 Kitty_Brashear
12: 10077606 Frank_Huelsman
13: 15931458 Rudy_Kallio
```

```
1: 28328847 Phoenix_Senators
    [winner: 28328847 Phoenix_Senators, pem_pos=0,
final_scores=0.019311264, final_local_scores=0.053375438,
pem=0.0, similarity=5.6825423]
```

```
    gt: 28328847 Phoenix_Senators, pem_pos=0,
final_scores=0.019311264, final_local_scores=0.053375438,
pem=0.0, similarity=5.6825423 ]
```

```
2: 19760714 John_McCloskey_(baseball_manager)
    [winner: 245244 John_McCloskey, pem_pos=0,
```

```

final_scores=0.17974435, final_local_scores=0.28007057,
pem=0.0, similarity=12.447577|
    gt: 19760714 John_McCloskey_(baseball_manager),
pem_pos=1, final_scores=0.16028525,
final_local_scores=0.245296, pem=-0.70774597,
similarity=15.258669 ]
3: 106575 Douglas,_Arizona
    [winner: 201793 Douglas_Aircraft_Company, pem_pos=0,
final_scores=-0.21125323, final_local_scores=-0.18461868,
pem=-2.0593867, similarity=9.779589|
    gt: 106575 Douglas,_Arizona, pem_pos=4,
final_scores=-0.63047487, final_local_scores=-0.6928283,
pem=-3.327517, similarity=1.5098883 ]
4: 38776 Major_League_Baseball
    [winner: 685987 Major_League_(film), pem_pos=1,
final_scores=-1.128817, final_local_scores=-1.1867268,
pem=-0.65972805, similarity=-27.736881|
    gt: 38776 Major_League_Baseball, pem_pos=0,
final_scores=-1.2192512, final_local_scores=-1.3921857,
pem=-0.47792417, similarity=-34.85686 ]
5: 11718901 Stoney_McGlynn
    [winner: 11718901 Stoney_McGlynn, pem_pos=0,
final_scores=-0.08226207, final_local_scores=0.01085034,
pem=0.0, similarity=4.413509|
    gt: 11718901 Stoney_McGlynn, pem_pos=0,
final_scores=-0.08226207, final_local_scores=0.01085034,
pem=0.0, similarity=4.413509 ]
6: 34160739 Herb_Hall_(baseball)
    [winner: 16678272 Herb_Hall, pem_pos=0,
final_scores=-0.34147978, final_local_scores=-0.27265954,
pem=0.0, similarity=-4.0469913|
    gt: 34160739 Herb_Hall_(baseball), pem_pos=1,
final_scores=-0.90559196, final_local_scores=-0.8934086,
pem=-4.4886365, similarity=1.8385212 ]
1: 29810 Texas
    [pem_pos=0, final_scores=0.721888,
final_local_scores=0.7621469, pem=-0.16191302,
similarity=27.714205]
2: 21883824 Arizona
    [pem_pos=0, final_scores=0.45034498,
final_local_scores=0.48092997, pem=-0.22599088,
similarity=19.670597]
3: 21649 New_Mexico
    [pem_pos=0, final_scores=0.40876096,
final_local_scores=0.43996978, pem=-0.20017311,
similarity=18.307861]
4: 1398193 Albuquerque_Dukes
    [pem_pos=0, final_scores=0.35752362,
final_local_scores=0.36362052, pem=0.0, similarity=14.940875]
5: 70292 Texas_League
    [pem_pos=0, final_scores=0.31612498,
final_local_scores=0.3085198, pem=0.0, similarity=13.296559]

```

6: 49121 Phoenix, Arizona
 [pem_pos=0, final_scores=0.20181511,
 final_local_scores=0.16545323, pem=-0.90867877,
 similarity=13.968703]
 7: 57180 Tucson, Arizona
 [pem_pos=0, final_scores=0.21059035,
 final_local_scores=0.16634074, pem=0.0, similarity=9.053653]
 8: 135823 El_Paso, Texas
 [pem_pos=0, final_scores=0.17479981,
 final_local_scores=0.14058554, pem=-0.27809596,
 similarity=9.797396]
 9: 51278 Albuquerque, New_Mexico
 [pem_pos=0, final_scores=0.2507053,
 final_local_scores=0.22133693, pem=-0.10911326,
 similarity=11.288223]
 10: 125874 Las_Cruces, New_Mexico
 [pem_pos=0, final_scores=0.18892972,
 final_local_scores=0.18359274, pem=0.0, similarity=9.568486]
 11: 30612207 Kitty_Brashear
 [pem_pos=0, final_scores=0.11696763,
 final_local_scores=0.23269233, pem=0.0, similarity=11.033716]
 12: 10077606 Frank_Huelsman
 [pem_pos=0, final_scores=0.34404653,
 final_local_scores=0.44432718, pem=0.0, similarity=17.349321]
 13: 15931458 Rudy_Kallio
 [pem_pos=0, final_scores=0.18840788,
 final_local_scores=0.30688822, pem=0.0, similarity=13.2478695]
 1: 34299113 Rio_Grande_Association
 [pem_pos=0, final_scores=0.072390124,
 final_local_scores=0.09960951, pem=0.0, similarity=7.062259]
 2: 245244 John_McCloskey
 [pem_pos=0, final_scores=0.17974435,
 final_local_scores=0.28007057, pem=0.0, similarity=12.447577]
 3: 34299113 Rio_Grande_Association
 [pem_pos=0, final_scores=0.18786232,
 final_local_scores=0.22574538, pem=0.0, similarity=10.826406]
 4: 20050 Minnesota_Twins
 [pem_pos=3, final_scores=0.050674275,
 final_local_scores=0.013514519, pem=-2.594315,
 similarity=18.601301]

Sample: b'Men_at_Work'

-LRB- season_1 -RRB- The first season of the TBS sitcom ' 1
 Men at Work ' premiered on May 24 , 2012 and ended July 12 ,
 2012 . A total of ten episodes were produced . Season one
 regular cast members include [1 1 Danny Masterson 1] , 2 Mich
 ael Cassidy , 3 Adam Busch , [2 2 Meredith Hagner 2] and [3 3
 James Lesure 3] . On January 6 , 2012 , [1 TBS 1] picked up
 Men at Work for a first season of 10 episodes set to premiere
 on May 24 , 2012 at 10 pm . Production on the first season
 began in April 2012 . Guest stars for season one of Men at
 Work include , [4 4 Amy Smart 4] as Lisa , Milo 's ex-

girlfriend ; Julian Morris as Damien ; [5 5 J. K. Simmons 5]
as P.J. Jordan , Amy 's father and owner of the magazine ; [6
6 Melissa Claire Egan 6] as Jessica , Milo 's date ; [7 7 Eth
an Suplee 7] as Dan ; [8 8 Stacy Keibler 8] as Keri ; [9 9 Ka
thy Najimy 9] as Sasha Ryan , a well-known sex blogger ; [10 1
0 Josh Hopkins 10] as Ryan , a friend of Amy 's who is
blind ; [11 11 Kevin Pollak 11] as Pavel ; [12 12 Christopher
Masterson 12] as a hotel employee ; [13 13 Wilmer Valderrama 1
3] as Eri Ricaldo ; [14 14 Alex Breckenridge 14] as Katelyn ,
a young woman who keeps blowing off her dates with Tyler ; 4 J
oel David Moore as Doug , a Full Steam Magazine editor ; [15 1
5 Laura Prepon 15] as Hannah , Neal 's friend who is set up
with Milo ; [16 16 William Baldwin 16] as Shepard Peters , a
famous , free-spirited , sexually driven photographer , and [1
7 17 Richard Riehle 17]

1: 947500 Danny_Masterson
2: 17406451 Meredith_Hagner
3: 1989233 James_Lesure
4: 814271 Amy_Smart
5: 1086198 J._K._Simmons
6: 5950296 Melissa_Claire_Egan
7: 1931312 Ethan_Suplee
8: 307284 Stacy_Keibler
9: 452048 Kathy_Najimy
10: 3138026 Josh_Hopkins
11: 524332 Kevin_Pollak
12: 947516 Christopher_Masterson
13: 497170 Wilmer_Valderrama
14: 4359350 Alex_Breckenridge
15: 753300 Laura_Prepon
16: 182535 William_Baldwin
17: 3005349 Richard_Riehle

1: 34305662 Men_at_Work_(TV_series)
[winner: 20451 Men_at_Work, pem_pos=0,
final_scores=-0.1715082, final_local_scores=-0.13530307,
pem=0.0, similarity=0.051999554]
gt: 34305662 Men_at_Work_(TV_series), pem_pos=3,
final_scores=-0.89382845, final_local_scores=-0.9747738,
pem=-5.0731983, similarity=2.5893595]
2: 2331466 Michael_Cassidy_(actor)
[winner: 2331466 Michael_Cassidy_(actor), pem_pos=1,
final_scores=-0.087190896, final_local_scores=-0.08926219,
pem=-0.3667178, similarity=3.4202194]
gt: 2331466 Michael_Cassidy_(actor), pem_pos=1,
final_scores=-0.087190896, final_local_scores=-0.08926219,
pem=-0.3667178, similarity=3.4202194]
3: 345562 Adam_Busch
[winner: 345562 Adam_Busch, pem_pos=0,
final_scores=-0.046319947, final_local_scores=-0.041204825,
pem=0.0, similarity=2.8600788]
gt: 345562 Adam_Busch, pem_pos=0,
final_scores=-0.046319947, final_local_scores=-0.041204825,

pem=0.0, similarity=2.8600788]
4: 1402228 Joel_David_Moore
[winner: 1402228 Joel_David_Moore, pem_pos=0,
final_scores=0.022432193, final_local_scores=0.0015868098,
pem=0.0, similarity=4.1370664]
gt: 1402228 Joel_David_Moore, pem_pos=0,
final_scores=0.022432193, final_local_scores=0.0015868098,
pem=0.0, similarity=4.1370664]
1: 947500 Danny_Masterson
[pem_pos=0, final_scores=0.1463023,
final_local_scores=0.1089385, pem=0.0, similarity=7.340655]
2: 17406451 Meredith_Hagner
[pem_pos=0, final_scores=0.16163246,
final_local_scores=0.19802663, pem=0.0, similarity=9.999223]
3: 1989233 James_Lesure
[pem_pos=0, final_scores=0.21750058,
final_local_scores=0.25759593, pem=0.0, similarity=11.776889]
4: 814271 Amy_Smart
[pem_pos=0, final_scores=0.13577549,
final_local_scores=0.119712144, pem=0.0, similarity=7.662162]
5: 1086198 J._K._Simmons
[pem_pos=0, final_scores=0.07929604,
final_local_scores=0.06272328, pem=0.0, similarity=5.9615006]
6: 5950296 Melissa_Claire_Egan
[pem_pos=0, final_scores=0.09294753,
final_local_scores=0.11561689, pem=0.0, similarity=7.5399513]
7: 1931312 Ethan_Suplee
[pem_pos=0, final_scores=0.20111994,
final_local_scores=0.21779409, pem=0.0, similarity=10.589123]
8: 307284 Stacy_Keibler
[pem_pos=0, final_scores=0.30588245,
final_local_scores=0.34707418, pem=0.0, similarity=14.447099]
9: 452048 Kathy_Najimy
[pem_pos=0, final_scores=0.19106208,
final_local_scores=0.17551431, pem=0.0, similarity=9.32741]
10: 3138026 Josh_Hopkins
[pem_pos=0, final_scores=0.08769311,
final_local_scores=0.0789918, pem=0.0, similarity=6.4469857]
11: 524332 Kevin_Pollak
[pem_pos=0, final_scores=0.3026635,
final_local_scores=0.28156632, pem=0.0, similarity=12.492213]
12: 947516 Christopher_Masterson
[pem_pos=0, final_scores=0.07637964,
final_local_scores=0.074953556, pem=0.0, similarity=6.3264766]
13: 497170 Wilmer_Valderrama
[pem_pos=0, final_scores=0.21920316,
final_local_scores=0.19716355, pem=0.0, similarity=9.973467]
14: 4359350 Alex_Breckenridge
[pem_pos=0, final_scores=0.21836914,
final_local_scores=0.24169981, pem=0.0, similarity=11.302518]
15: 753300 Laura_Prepon
[pem_pos=0, final_scores=0.29703277,

final_local_scores=0.28200296, pem=0.0, similarity=12.505243]
 16: 182535 William_Baldwin
 [pem_pos=0, final_scores=0.20519714,
 final_local_scores=0.18728009, pem=0.0, similarity=9.678524]
 17: 3005349 Richard_Riehle
 [pem_pos=0, final_scores=0.12076782,
 final_local_scores=0.13836512, pem=0.0, similarity=8.218804]
 1: 31234796 TBS_(TV_network)
 [pem_pos=0, final_scores=0.121704265,
 final_local_scores=0.11769885, pem=-0.67694145,
 similarity=11.283395]

Sample: b'King_Diamond_discography'

The 1 discography of 2 King Diamond , a [1 Danish 1] 3 heavy metal band , consists of twelve studio releases , two 4 live albums , five 5 compilations , six 6 singles , and four 7 music video s. [2 King Diamond 2] was formed in 1985 , after the dissolution of the group [1 1 Mercyful Fate 1] , by vocalist [2 2 King Diamond 2] , guitarists 8 Andy LaRocque and 9 Michael Denner , bassist 10 Timi Hansen , and drummer [3 3 Mikkey Dee 3] . The following year , the band released their debut album 11 Fatal Portrait , which charted at number 33 in [4 4 Sweden 4] . [3 King Diamond 3] 's second studio album , 12 Abigail , was released on October 21 , 1987 , and reached number 123 in the [5 5 US 5] , number 39 in [4 Sweden 4] and number 68 in the [6 6 Netherlands 6] . Following some line-up changes , the group released the album 13 `` Them `` in 1988 , which peaked at number 38 in [5 Sweden 5] , number 65 in the [6 Netherlands 6] , and at number 89 in the [7 US 7] , making `` Them `` [8 King Diamond 8] 's highest charting album in [9 North America 9] . The following year , the band released the follow-up album 14 Conspiracy , which charted at number 111 in [10 North America 10] , number 41 in [11 Sweden 11] and at number 64 in the [12 Netherlands 12] . In 1990 , after more line-up changes , [13 King Diamond 13] released the album 15 The Eye , which only charted at number 179 in the [14 US 14] , which makes The Eye [15 King Diamond 15] 's lowest charting album in [16 North America 16] . After [17 Mercyful Fate 17] was reformed in 1993 , [18 King Diamond 18] remained inactive until 1995 , when the band released the album 16 The Spider 's Lullabye with the line-up of [19 King Diamond 19] , [20 Andy LaRocque 20] , guitarist Herb Simonsen , bassist 17 Chris Estes and drummer Darrin Anthony . The album went on to reach number 31 in [7 7 Finland 7] . The Spider 's Lullabye was followed by 18 The Graveyard -LRB- # 23 in [21 Finland 21] -RRB- and 19 Voodoo -LRB- # 27 in [22 Finland 22] , # 55 in the [23 Netherlands 23] -RRB- in 1996 and 1998 respectively . In 2000 , [24 King Diamond 24] released the album 20 House of God , which peaked at number 60 in [25 Sweden 25] . After the release of 2002 's , which peaked at number 42 in [26 Sweden 26] and at number 24 in [27 Finland 27] , [28 King Diamond 28] 's line-up has remained stable to this day , consisting of [29

King Diamond 29] , Andy LaRocque , bassist [8 8 Hal Patino 8]
 , guitarist [9 9 Mike Wead 9] and drummer Matt Thompson . In
 2003 , the band released 21 The Puppet Master , which reached
 number 36 in [30 Sweden 30] . In 2007 , [31 King Diamond 31]
 released their 12th studio album 22 Give Me Your Soul ... Please

1: 164678 Mercyful_Fate
2: 164694 King_Diamond
3: 2181006 Mikkey_Dee
4: 5058739 Sweden
5: 3434750 United_States
6: 21148 Netherlands
7: 10577 Finland
8: 10627090 Hal_Patino
9: 2534555 Mike_Wead

1: 510969 Discography

[winner: 9262774 ThaMuseMeant_discography, pem_pos=9,
final_scores=-0.9843866, final_local_scores=-0.90570396,
pem=-2.8332133, similarity=-7.530844]

RECALL MISS: 510969 Discography]

2: 1857489 King_Diamond_(band)

[winner: 164694 King_Diamond, pem_pos=0,
final_scores=-0.010073826, final_local_scores=-0.06808741,
pem=0.0, similarity=2.057849]

gt: 1857489 King_Diamond_(band), pem_pos=1,
final_scores=-0.101540804, final_local_scores=-0.15441902,
pem=-0.2568413, similarity=0.87828445]

3: 13869 Heavy_metal_music

[winner: 19042 Metal, pem_pos=0, final_scores=-0.8478363,
final_local_scores=-0.8568706, pem=-0.45522028,
similarity=-19.005455]

gt: 13869 Heavy_metal_music, pem_pos=1,
final_scores=-0.9118002, final_local_scores=-1.0494165,
pem=-0.93149173, similarity=-22.161373]

[winner: 21444840 Heavy_chain, pem_pos=10,
final_scores=-1.3049347, final_local_scores=-1.2294734,
pem=-3.8633578, similarity=-11.590685]

gt: 13869 Heavy_metal_music, pem_pos=1,
final_scores=-1.6614995, final_local_scores=-1.8683491,
pem=-1.4358722, similarity=-43.85705]

4: 988516 Live_album

[winner: 21565121 The_Police_discography, pem_pos=3,
final_scores=-1.1387229, final_local_scores=-1.0831189,
pem=-2.1882758, similarity=-16.332531]

RECALL MISS: 988516 Live_album]

[winner: 9925211 Anna_Vissi_Live, pem_pos=10,
final_scores=-1.2164544, final_local_scores=-1.1581719,
pem=-3.8937864, similarity=-9.297432]

gt: 988516 Live_album, pem_pos=6,
final_scores=-1.742614, final_local_scores=-1.9033458,
pem=-3.8485236, similarity=-31.78105]

5: 504309 Compilation_album

[winner: 504309 Compilation_album, pem_pos=0,
final_scores=-1.369446, final_local_scores=-1.5160592,
pem=0.0, similarity=-41.152515|
gt: 504309 Compilation_album, pem_pos=0,
final_scores=-1.369446, final_local_scores=-1.5160592,
pem=0.0, similarity=-41.152515]
6: 156540 Single_(music)
[winner: 156540 Single_(music), pem_pos=0,
final_scores=-1.1710606, final_local_scores=-1.303216,
pem=-0.33920756, similarity=-32.956192|
gt: 156540 Single_(music), pem_pos=0,
final_scores=-1.1710606, final_local_scores=-1.303216,
pem=-0.33920756, similarity=-32.956192]
7: 90138 Music_video
[winner: 4298091 Four_Music, pem_pos=0,
final_scores=-1.1435641, final_local_scores=-1.1364275,
pem=0.0, similarity=-29.823551|
RECALL MISS: 90138 Music_video]
[winner: 9960899 Crichton_Leprechaun, pem_pos=12,
final_scores=-1.3721842, final_local_scores=-1.3120742,
pem=-2.8332133, similarity=-19.657742|
gt: 90138 Music_video, pem_pos=1,
final_scores=-1.7095308, final_local_scores=-1.8915305,
pem=-1.8401091, similarity=-42.35053]
8: 1347877 Andy_LaRocque
[winner: 1347877 Andy_LaRocque, pem_pos=0,
final_scores=0.0017060488, final_local_scores=0.009531066,
pem=0.0, similarity=4.374139|
gt: 1347877 Andy_LaRocque, pem_pos=0,
final_scores=0.0017060488, final_local_scores=0.009531066,
pem=0.0, similarity=4.374139]
9: 2147127 Michael_Denner
[winner: 2147127 Michael_Denner, pem_pos=0,
final_scores=-0.031257346, final_local_scores=0.030053332,
pem=0.0, similarity=4.986564|
gt: 2147127 Michael_Denner, pem_pos=0,
final_scores=-0.031257346, final_local_scores=0.030053332,
pem=0.0, similarity=4.986564]
10: 9132193 Timi_Hansen
[winner: 9132193 Timi_Hansen, pem_pos=0,
final_scores=-0.106741294, final_local_scores=-0.052725673,
pem=0.0, similarity=2.5162737|
gt: 9132193 Timi_Hansen, pem_pos=0,
final_scores=-0.106741294, final_local_scores=-0.052725673,
pem=0.0, similarity=2.5162737]
11: 1178604 Fatal_Portrait
[winner: 1178604 Fatal_Portrait, pem_pos=0,
final_scores=-0.2903452, final_local_scores=-0.23181547,
pem=0.0, similarity=-2.828123|
gt: 1178604 Fatal_Portrait, pem_pos=0,
final_scores=-0.2903452, final_local_scores=-0.23181547,
pem=0.0, similarity=-2.828123]

12: 1618986 Abigail_(album)
 [winner: 102745 Abigail_Adams, pem_pos=0,
 final_scores=-0.23848256, final_local_scores=-0.16507453,
 pem=-0.8793044, similarity=3.9453535|
 gt: 1618986 Abigail_(album), pem_pos=3,
 final_scores=-0.5649956, final_local_scores=-0.5255862,
 pem=-2.4873402, similarity=1.931715]

13: 30863063 Them_(King_Diamond_album)
 [winner: 59349 Quotation_mark, pem_pos=0,
 final_scores=-1.9660491, final_local_scores=-2.010955,
 pem=0.0, similarity=-55.9212|
 RECALL MISS: 30863063 Them_(King_Diamond_album)]

14: 193248 Conspiracy_(King_Diamond_album)
 [winner: 3406102 Conspiracy_(criminal), pem_pos=0,
 final_scores=-0.4724571, final_local_scores=-0.40864143,
 pem=-1.3528501, similarity=-0.7479536|
 gt: 193248 Conspiracy_(King_Diamond_album), pem_pos=5,
 final_scores=-0.7544233, final_local_scores=-0.74261755,
 pem=-2.45766, similarity=-4.706338]

15: 1618994 The_Eye_(King_Diamond_album)
 [winner: 3511710 The_Eye_(2002_film), pem_pos=0,
 final_scores=-0.06688577, final_local_scores=0.042658567,
 pem=-0.38755548, similarity=7.470316|
 gt: 1618994 The_Eye_(King_Diamond_album), pem_pos=3,
 final_scores=-0.43465558, final_local_scores=-0.3742836,
 pem=-1.9127628, similarity=3.3222446]

16: 1618971 The_Spider's_Lullabye
 [winner: 19532082 Lullabye_(Goodnight,_My_Angel),
 pem_pos=0, final_scores=-0.14940049,
 final_local_scores=-0.04164464, pem=0.0,
 similarity=2.8469539|
 RECALL MISS: 1618971 The_Spider's_Lullabye]
 [winner: 649269 Spider_(pulp_fiction), pem_pos=0,
 final_scores=-0.15027942, final_local_scores=-0.09306918,
 pem=-0.2631951, similarity=2.7436388|
 RECALL MISS: 1618971 The_Spider's_Lullabye]
 [winner: 1777106 Vince_Welnick, pem_pos=3,
 final_scores=-2.149366, final_local_scores=-2.3161585,
 pem=-2.8006017, similarity=-49.798973|
 RECALL MISS: 1618971 The_Spider's_Lullabye]

17: 5128298 Chris_Estes
 [winner: 5128298 Chris_Estes, pem_pos=0,
 final_scores=-0.049358726, final_local_scores=0.010417715,
 pem=0.0, similarity=4.4005985|
 gt: 5128298 Chris_Estes, pem_pos=0,
 final_scores=-0.049358726, final_local_scores=0.010417715,
 pem=0.0, similarity=4.4005985]

18: 1618964 The_Graveyard_(album)
 [winner: 21901145 The_Graveyard_(video_game), pem_pos=0,
 final_scores=-0.26221558, final_local_scores=-0.12976493,
 pem=0.0, similarity=0.21726862|
 gt: 1618964 The_Graveyard_(album), pem_pos=4,

```

final_scores=-0.9026529, final_local_scores=-0.8646273,
pem=-4.2904596, similarity=1.6196961 ]
19: 578354 Voodoo_(King_Diamond_album)
    [winner: 10950407 Louisiana_Voodoo, pem_pos=1,
final_scores=-0.43321463, final_local_scores=-0.39949265,
pem=-1.568844, similarity=0.6996728]
    gt: 578354 Voodoo_(King_Diamond_album), pem_pos=6,
final_scores=-0.5847332, final_local_scores=-0.58359087,
pem=-2.3730764, similarity=-0.42064375 ]
20: 2150995 House_of_God_(album)
    [winner: 876239 The_House_of_God, pem_pos=0,
final_scores=-0.041644946, final_local_scores=0.08042678,
pem=-0.4932813, similarity=9.172349]
    gt: 2150995 House_of_God_(album), pem_pos=1,
final_scores=-0.15714878, final_local_scores=-0.11231563,
pem=-0.49812397, similarity=3.446866 ]
21: 2151023 The_Puppet_Master_(album)
    [winner: 2151023 The_Puppet_Master_(album), pem_pos=0,
final_scores=-0.22721258, final_local_scores=-0.18168774,
pem=0.0, similarity=-1.3322119]
    gt: 2151023 The_Puppet_Master_(album), pem_pos=0,
final_scores=-0.22721258, final_local_scores=-0.18168774,
pem=0.0, similarity=-1.3322119 ]
22: 11327385 Give_Me_Your_Soul...Please
    [winner: 3277627 Please_(U2_song), pem_pos=0,
final_scores=-0.309704, final_local_scores=-0.28633606,
pem=-0.4267834, similarity=-2.134212]
    RECALL MISS: 11327385 Give_Me_Your_Soul...Please]
    [winner: 2845393 Give_(album), pem_pos=0,
final_scores=-0.65542656, final_local_scores=-0.5841818,
pem=-0.3792768, similarity=-11.280871]
    RECALL MISS: 11327385 Give_Me_Your_Soul...Please]
    [winner: 62808 Soul_music, pem_pos=0,
final_scores=-1.0551373, final_local_scores=-1.1734757,
pem=-0.631794, similarity=-27.493347]
    RECALL MISS: 11327385 Give_Me_Your_Soul...Please]
    [winner: 27874167 ..."Let_Me_Sing", pem_pos=0,
final_scores=-1.564025, final_local_scores=-1.6332633,
pem=0.0, similarity=-44.65012]
    RECALL MISS: 11327385 Give_Me_Your_Soul...Please]
1: 164678 Mercyful_Fate
    [pem_pos=0, final_scores=0.18606292,
final_local_scores=0.16863242, pem=0.0, similarity=9.122041]
2: 164694 King_Diamond
    [pem_pos=0, final_scores=0.13377316,
final_local_scores=0.094286576, pem=0.0, similarity=6.9034123]
3: 2181006 Mikkey_Dee
    [pem_pos=0, final_scores=0.10315709,
final_local_scores=0.11274324, pem=0.0, similarity=7.454196]
4: 5058739 Sweden
    [pem_pos=0, final_scores=0.6846186,
final_local_scores=0.70273376, pem=-0.2135508,

```

similarity=26.222012]
5: 3434750 United_States
[pem_pos=0, final_scores=0.10349043,
final_local_scores=0.10142876, pem=-0.48755783,
similarity=9.767964]
6: 21148 Netherlands
[pem_pos=0, final_scores=0.47279304,
final_local_scores=0.4976079, pem=-0.24940795,
similarity=20.295645]
7: 10577 Finland
[pem_pos=0, final_scores=0.7027847,
final_local_scores=0.74042183, pem=-0.2445894,
similarity=27.515493]
8: 10627090 Hal_Patino
[pem_pos=0, final_scores=0.049584106,
final_local_scores=0.12916765, pem=0.0, similarity=7.944333]
9: 2534555 Mike_Wead
[pem_pos=0, final_scores=0.06334217,
final_local_scores=0.14503834, pem=0.0, similarity=8.417946]
1: 76972 Denmark
[pem_pos=0, final_scores=0.5577523,
final_local_scores=0.57908565, pem=-0.65021485,
similarity=24.906752]
2: 164694 King_Diamond
[pem_pos=0, final_scores=0.104387745,
final_local_scores=0.062187478, pem=0.0, similarity=5.9455113]
3: 164694 King_Diamond
[pem_pos=0, final_scores=0.07356079,
final_local_scores=0.026867405, pem=0.0, similarity=4.89149]
4: 5058739 Sweden
[pem_pos=0, final_scores=0.58760655,
final_local_scores=0.59676284, pem=-0.2135508,
similarity=23.059628]
5: 5058739 Sweden
[pem_pos=0, final_scores=0.61448437,
final_local_scores=0.6261228, pem=-0.2135508,
similarity=23.935787]
6: 21148 Netherlands
[pem_pos=0, final_scores=0.45960903,
final_local_scores=0.48320633, pem=-0.24940795,
similarity=19.865873]
7: 3434750 United_States
[pem_pos=0, final_scores=0.21714143,
final_local_scores=0.22557521, pem=-0.48755783,
similarity=13.472743]
8: 164694 King_Diamond
[pem_pos=0, final_scores=0.15928994,
final_local_scores=0.12215981, pem=0.0, similarity=7.7352057]
9: 21139 North_America
[pem_pos=0, final_scores=0.15104492,
final_local_scores=0.18719783, pem=-0.011697312,
similarity=9.739681]

10: 21139 North_America
[pem_pos=0, final_scores=0.17004494,
final_local_scores=0.2079525, pem=-0.011697312,
similarity=10.359042]
11: 5058739 Sweden
[pem_pos=0, final_scores=0.62685883,
final_local_scores=0.63964, pem=-0.2135508,
similarity=24.339167]
12: 21148 Netherlands
[pem_pos=0, final_scores=0.45981103,
final_local_scores=0.483427, pem=-0.24940795,
similarity=19.872458]
13: 164694 King_Diamond
[pem_pos=0, final_scores=0.15295993,
final_local_scores=0.11524519, pem=0.0, similarity=7.5288596]
14: 3434750 United_States
[pem_pos=0, final_scores=0.08835827,
final_local_scores=0.08489917, pem=-0.48755783,
similarity=9.274689]
15: 164694 King_Diamond
[pem_pos=0, final_scores=0.066270754,
final_local_scores=0.018904135, pem=0.0, similarity=4.65385]
16: 21139 North_America
[pem_pos=0, final_scores=0.10977273,
final_local_scores=0.14211422, pem=-0.011697312,
similarity=8.394297]
17: 164678 Mercyful_Fate
[pem_pos=0, final_scores=0.21054189,
final_local_scores=0.19385868, pem=0.0, similarity=9.874843]
18: 164694 King_Diamond
[pem_pos=0, final_scores=0.112987086,
final_local_scores=0.06993462, pem=0.0, similarity=6.1767015]
19: 164694 King_Diamond
[pem_pos=0, final_scores=0.11818431,
final_local_scores=0.07561183, pem=0.0, similarity=6.346121]
20: 1347877 Andy_LaRocque
[pem_pos=0, final_scores=0.039413825,
final_local_scores=0.05072108, pem=0.0, similarity=5.603331]
21: 10577 Finland
[pem_pos=0, final_scores=0.5181696,
final_local_scores=0.5387578, pem=-0.2445894,
similarity=21.497437]
22: 10577 Finland
[pem_pos=0, final_scores=0.5316841,
final_local_scores=0.55352044, pem=-0.2445894,
similarity=21.937983]
23: 21148 Netherlands
[pem_pos=0, final_scores=0.324067,
final_local_scores=0.33514726, pem=-0.24940795,
similarity=15.4474945]
24: 164694 King_Diamond
[pem_pos=0, final_scores=0.111990765,

```

final_local_scores=0.068846315, pem=0.0, similarity=6.144224]
25: 5058739 Sweden
    [pem_pos=0, final_scores=0.63117903,
final_local_scores=0.6443592, pem=-0.2135508,
similarity=24.479998]
26: 5058739 Sweden
    [pem_pos=0, final_scores=0.587635,
final_local_scores=0.5967939, pem=-0.2135508,
similarity=23.060555]
27: 10577 Finland
    [pem_pos=0, final_scores=0.73293215,
final_local_scores=0.77335334, pem=-0.2445894,
similarity=28.498234]
28: 164694 King_Diamond
    [pem_pos=0, final_scores=0.16759284,
final_local_scores=0.12958312, pem=0.0, similarity=7.9567313]
29: 164694 King_Diamond
    [pem_pos=0, final_scores=0.17089851,
final_local_scores=0.1348404, pem=0.0, similarity=8.11362]
30: 5058739 Sweden
    [pem_pos=0, final_scores=0.67045254,
final_local_scores=0.6872595, pem=-0.2135508,
similarity=25.76023]
31: 164694 King_Diamond
    [pem_pos=0, final_scores=0.13543604,
final_local_scores=0.09445672, pem=0.0, similarity=6.9084897]

```

Sample: b'OK_Hotel'

The [1 OK Hotel 1] was a bar and music venue located under the viaduct in [1 1 Seattle 1] 's [2 2 Pioneer Square 2] district . The club 's 15-year-plus life span came to an end with the of February 28 , 2001 , which damaged numerous buildings in the historic district . Most widely recognized as a prominent location in the movie 1 Singles -LRB- 1992 -RRB- , the [2 OK Hotel 2] was one of numerous active [3 Seattle 3] rock venues during the celebrated local scene of the late eighties . The first musical group to play live at the venue was local band Seers of Bavaria , featuring future bass player for pop-punk band 2 Flop Paul Schurr on lead vocals . The club 's debut headliner was Vexed , also from [4 Seattle 4] . Local bands such as , 3 Tad , [3 3 Mother Love Bone 3] , [4 4 Soundgarden 4] and [5 5 Nirvana 5] -LRB- including the first live performance of '' [6 6 Smells Like Teen Spirit 6] '' -RRB- played at the [5 OK Hotel 5] , as well as touring 4 punk , and 5 jazz artists from around the country . It is mentioned in Thanx , the last track on 6 Sublime 's debut album . It was also where 7 Queens of the Stone Age played their first live show . In its final years , the [6 OK Hotel 6] had become a treasured home for [7 Seattle 7] 's lively creative music scene , featuring such artists as [7 7 Bill Frisell 7] , [8 8 Wayne Horvitz 8] with 9 Zony Mash and 10 Ponga , [9 Robin Holcomb 9] , 11 Medicine Hat , , 12 the Livi

ng Daylights , 13 Amy Denio

1: 11388236 Seattle

2: 295027 Pioneer_Square,_Seattle

3: 169609 Mother_Love_Bone

4: 40321 Soundgarden

5: 21231 Nirvana_(band)

6: 59993 Smells_Like_Teen_Spirit

7: 376072 Bill_Frisell

1: 577663 Singles_(1992_film)

[winner: 156540 Single_(music), pem_pos=0,
final_scores=-0.58289397, final_local_scores=-0.65779525,
pem=-1.1366256, similarity=-9.359061|

gt: 577663 Singles_(1992_film), pem_pos=1,
final_scores=-0.6407644, final_local_scores=-0.7031749,
pem=-2.0612242, similarity=-5.68517]

2: 4353309 Flop_(band)

[winner: 4353309 Flop_(band), pem_pos=0,
final_scores=-0.3219745, final_local_scores=-0.3450997,
pem=0.0, similarity=-6.2087502|

gt: 4353309 Flop_(band), pem_pos=0,
final_scores=-0.3219745, final_local_scores=-0.3450997,
pem=0.0, similarity=-6.2087502]

3: 820886 Tad_(band)

[winner: 3848490 Tad_Lincoln, pem_pos=1,
final_scores=-0.0671708, final_local_scores=0.028350547,
pem=-0.94802105, similarity=10.091236|

gt: 820886 Tad_(band), pem_pos=3,
final_scores=-0.42727977, final_local_scores=-0.46902144,
pem=-2.2586803, similarity=2.3762326]

4: 23037 Punk_rock

[winner: 5078840 Punk_(fireworks), pem_pos=4,
final_scores=-1.1136115, final_local_scores=-1.0392613,
pem=-2.2692502, similarity=-14.583385|

gt: 23037 Punk_rock, pem_pos=0, final_scores=-1.258892,
final_local_scores=-1.4173331, pem=-0.1681528,
similarity=-37.291897]

5: 15613 Jazz

[winner: 15613 Jazz, pem_pos=0, final_scores=-0.6438046,
final_local_scores=-0.74141455, pem=-0.030054808,
similarity=-17.872133|

gt: 15613 Jazz, pem_pos=0, final_scores=-0.6438046,
final_local_scores=-0.74141455, pem=-0.030054808,
similarity=-17.872133]

6: 20647155 Sublime_(band)

[winner: 20647155 Sublime_(band), pem_pos=0,
final_scores=-0.0719665, final_local_scores=-0.110108845,
pem=-0.057310108, similarity=1.1155066|

gt: 20647155 Sublime_(band), pem_pos=0,
final_scores=-0.0719665, final_local_scores=-0.110108845,
pem=-0.057310108, similarity=1.1155066]

7: 159119 Queens_of_the_Stone_Age

[winner: 29219 Stone_Age, pem_pos=0,

final_scores=-0.054512054, final_local_scores=0.059234932,
pem=0.0, similarity=5.8574014|

RECALL MISS: 159119 Queens_of_the_Stone_Age]

[winner: 45579 Queens, pem_pos=0,
final_scores=-0.92096865, final_local_scores=-1.0083591,
pem=-0.11920574, similarity=-25.353474|

gt: 159119 Queens_of_the_Stone_Age, pem_pos=11,
final_scores=-1.1557887, final_local_scores=-1.2826622,
pem=-3.6382337, similarity=-14.402204]

[winner: 10447732

Results_by_riding_of_the_Canadian_federal_election,_2008,
pem_pos=12, final_scores=-1.7112517,
final_local_scores=-1.6009295, pem=-3.6198866,
similarity=-23.999708|

RECALL MISS: 159119 Queens_of_the_Stone_Age]

8: 867479 Wayne_Horvitz

[winner: 23318218 Wayne_L._Horvitz, pem_pos=1,
final_scores=0.14702292, final_local_scores=0.26287276,
pem=-0.6931472, similarity=15.703801|

gt: 867479 Wayne_Horvitz, pem_pos=0,
final_scores=0.10418554, final_local_scores=0.12958965,
pem=0.0, similarity=7.9569263]

9: 27300369 Zony_Mash

[winner: 27300369 Zony_Mash, pem_pos=0,
final_scores=-0.16546112, final_local_scores=-0.046525434,
pem=0.0, similarity=2.701301|

gt: 27300369 Zony_Mash, pem_pos=0,
final_scores=-0.16546112, final_local_scores=-0.046525434,
pem=0.0, similarity=2.701301]

10: 9532840 Ponga_(band)

[winner: 2705623 Ponga,_Asturias, pem_pos=0,
final_scores=-0.33337316, final_local_scores=-0.18016373,
pem=-0.15341023, similarity=-0.4524634|

gt: 9532840 Ponga_(band), pem_pos=1,
final_scores=-0.33737093, final_local_scores=-0.30748993,
pem=-0.9617531, similarity=0.14376245]

11: 242729 Medicine_Hat

[winner: 242729 Medicine_Hat, pem_pos=0,
final_scores=-0.0046821237, final_local_scores=0.035692796,
pem=-0.035786577, similarity=5.34947|

gt: 242729 Medicine_Hat, pem_pos=0,
final_scores=-0.0046821237, final_local_scores=0.035692796,
pem=-0.035786577, similarity=5.34947]

12: 566713 The_Living_Daylight

[winner: 9522171 Living_Daylight, pem_pos=0,
final_scores=-0.11592774, final_local_scores=0.0028720498,
pem=0.0, similarity=4.1754208|

gt: 566713 The_Living_Daylight, pem_pos=2,
final_scores=-0.5895023, final_local_scores=-0.598814,
pem=-4.158883, similarity=8.836563]

13: 753219 Amy_Denio

[winner: 753219 Amy_Denio, pem_pos=0,

```
final_scores=-0.22497523, final_local_scores=-0.18595836,  
pem=0.0, similarity=-1.4596552|  
    gt: 753219 Amy_Denio, pem_pos=0,  
final_scores=-0.22497523, final_local_scores=-0.18595836,  
pem=0.0, similarity=-1.4596552 ]  
1: 11388236 Seattle  
    [pem_pos=0, final_scores=0.55753595,  
final_local_scores=0.5408762, pem=-0.1054736,  
similarity=20.804121]  
2: 295027 Pioneer_Square,_Seattle  
    [pem_pos=0, final_scores=0.30078053,  
final_local_scores=0.36313504, pem=0.0, similarity=14.926387]  
3: 169609 Mother_Love_Bone  
    [pem_pos=0, final_scores=0.075655,  
final_local_scores=0.06638205, pem=0.0, similarity=6.070686]  
4: 40321 Soundgarden  
    [pem_pos=0, final_scores=0.041316494,  
final_local_scores=0.0063850433, pem=0.0,  
similarity=4.2802553]  
5: 21231 Nirvana_(band)  
    [pem_pos=0, final_scores=0.12395273,  
final_local_scores=0.09331003, pem=-0.21961989,  
similarity=8.068597]  
6: 59993 Smells_Like_Teen_Spirit  
    [pem_pos=0, final_scores=0.039379105,  
final_local_scores=0.036687165, pem=0.0, similarity=5.184531]  
7: 376072 Bill_Frisell  
    [pem_pos=0, final_scores=0.19694082,  
final_local_scores=0.21329242, pem=0.0, similarity=10.454784]  
1: 2469617 OK_Hotel  
    [pem_pos=0, final_scores=0.13239916,  
final_local_scores=0.115856856, pem=0.0, similarity=7.547113]  
2: 2469617 OK_Hotel  
    [pem_pos=0, final_scores=0.1297686,  
final_local_scores=0.112983346, pem=0.0, similarity=7.4613614]  
3: 11388236 Seattle  
    [pem_pos=0, final_scores=0.5598605,  
final_local_scores=0.5434154, pem=-0.1054736,  
similarity=20.879896]  
4: 11388236 Seattle  
    [pem_pos=0, final_scores=0.49297363,  
final_local_scores=0.4703517, pem=-0.1054736,  
similarity=18.699528]  
5: 2469617 OK_Hotel  
    [pem_pos=0, final_scores=0.1923645,  
final_local_scores=0.18135989, pem=0.0, similarity=9.501854]  
6: 2469617 OK_Hotel  
    [pem_pos=0, final_scores=0.2116385,  
final_local_scores=0.2024138, pem=0.0, similarity=10.130144]  
7: 11388236 Seattle  
    [pem_pos=0, final_scores=0.4914906,  
final_local_scores=0.4687317, pem=-0.1054736,
```

```
similarity=18.651184]
8: 23318218 Wayne_L._Horvitz
   [pem_pos=1, final_scores=0.14702292,
final_local_scores=0.26287276, pem=-0.6931472,
similarity=15.703801]
9: 35874752 Robin_Holcomb
   [pem_pos=0, final_scores=0.06366755,
final_local_scores=0.1301657, pem=0.0, similarity=7.9741173]
```

Sample: b'Penn_State_Lady_Lions_basketball'

The [1 1 Penn State 1] women 's basketball team represents and plays its home games in the [1 1 Bryce Jordan Center 1] . In 2013 , the [2 Lady Lions 2] became just the twelfth program in NCAA Division I history to reach 850 wins . [3 Penn State 3] has won 7 regular season titles and the first 2 tournament titles in 1995 and 1996 . The [4 Lady Lions 4] have 24 tournament appearances as of 2013 , the most in the . The team 's best post-season finish came in 2000 when the [5 Lady Lions 5] reached the Final Four before losing to eventual champion . The [6 Lady Lions 6] captured the 2 WNIT title in 1998 defeating [7 3 Baylor 7] 59-56 in [8 4 Waco 8] , [9 Texas 9] . Notable alumni include 5 WBCA First Team All-Americans , Susan Robinson , [2 2 Helen Darling 2] , and [3 3 Kelly Mazzante 3] . [4 4 ESPN 4] correspondent [5 5 Lisa Salters 5] is the shortest player in [10 Lady Lions 10] history at 5 ' -2 ' ' . Head Coach [6 6 Coquese Washington 6] is in her sixth year coaching the [11 Lady Lions 11] . Annually , the [12 Lady Lions 12] don pink jerseys in support of several organizations that fight 6 breast cancer in what is now known as the `` Pink Zone at [13 Penn State 13] ' ' game . The [14 Lady Lions 14] were the first [15 Division I 15] team in the nation to wear pink jerseys , a growing trend in athletics . Then-head coach [7 7 Rene Portland 7] developed the idea in 2006 with money from the Conference , and the first game - LRB- termed the `` Think Pink ' ' game -RRB- occurred in February 2007 against

1: 1874597 Bryce_Jordan_Center

2: 2456898 Helen_Darling

3: 5700404 Kelly_Mazzante

4: 77795 ESPN

5: 4298028 Lisa_Salters

6: 5782031 Coquese_Washington

7: 4445300 Rene_Portland

1: 5382830 Penn_State_Nittany_Lions

[winner: 1209509 Pennsylvania_State_University,

pem_pos=0, final_scores=0.18605413,

final_local_scores=0.15652359, pem=-0.49028125,

similarity=11.426914]

gt: 5382830 Penn_State_Nittany_Lions, pem_pos=2,
final_scores=-0.14273241, final_local_scores=-0.2607472,
pem=-2.2703366, similarity=8.654941]

2: 4402725 Women's_National_Invitation_Tournament

[winner: 4402725 Women's_National_Invitation_Tournament,
pem_pos=0, final_scores=-0.067029074,
final_local_scores=-0.06820388, pem=0.0, similarity=2.054373]
gt: 4402725 Women's_National_Invitation_Tournament,
pem_pos=0, final_scores=-0.067029074,
final_local_scores=-0.06820388, pem=0.0, similarity=2.054373]
3: 22212779 Baylor_Lady_Bears_basketball
[winner: 606467 Don_Baylor, pem_pos=4,
final_scores=0.22302096, final_local_scores=0.31029862,
pem=-2.5183601, similarity=27.044876]
gt: 22212779 Baylor_Lady_Bears_basketball, pem_pos=6,
final_scores=-0.64852095, final_local_scores=-0.6740596,
pem=-2.8242264, similarity=-0.666988]
4: 136365 Waco, Texas
[winner: 29810 Texas, pem_pos=0, final_scores=0.37265307,
final_local_scores=0.4381433, pem=-0.16191302,
similarity=18.045292]
RECALL MISS: 136365 Waco, Texas]
[winner: 136365 Waco, Texas, pem_pos=0,
final_scores=0.34576404, final_local_scores=0.37113923,
pem=-0.42231324, similarity=17.461853]
gt: 136365 Waco, Texas, pem_pos=0,
final_scores=0.34576404, final_local_scores=0.37113923,
pem=-0.42231324, similarity=17.461853]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8964969, final_local_scores=-1.8814917,
pem=-4.6728287, similarity=-26.646181]
RECALL MISS: 136365 Waco, Texas]
5: 24565855 Women's_Basketball_Coaches_Association
[winner: 24565855 Women's_Basketball_Coaches_Association,
pem_pos=0, final_scores=-0.059103653,
final_local_scores=-0.075630605, pem=0.0,
similarity=1.832745]
gt: 24565855 Women's_Basketball_Coaches_Association,
pem_pos=0, final_scores=-0.059103653,
final_local_scores=-0.075630605, pem=0.0,
similarity=1.832745]
6: 70547 Breast_cancer
[winner: 70547 Breast_cancer, pem_pos=0,
final_scores=-1.0177126, final_local_scores=-1.0284843,
pem=0.0, similarity=-26.602314]
gt: 70547 Breast_cancer, pem_pos=0,
final_scores=-1.0177126, final_local_scores=-1.0284843,
pem=0.0, similarity=-26.602314]
1: 1874597 Bryce_Jordan_Center
[pem_pos=0, final_scores=0.11924501,
final_local_scores=0.14215332, pem=0.0, similarity=8.331852]
2: 2456898 Helen_Darling
[pem_pos=0, final_scores=0.17788021,
final_local_scores=0.23281267, pem=0.0, similarity=11.037308]
3: 5700404 Kelly_Mazzante
[pem_pos=0, final_scores=0.35023338,

final_local_scores=0.4104517, pem=0.0, similarity=16.338411]
4: 77795 ESPN
[pem_pos=0, final_scores=0.17734341,
final_local_scores=0.15853897, pem=-0.09327439,
similarity=9.328073]
5: 4298028 Lisa_Salters
[pem_pos=0, final_scores=0.07546018,
final_local_scores=0.1218811, pem=0.0, similarity=7.7268877]
6: 5782031 Coquese_Washington
[pem_pos=0, final_scores=0.18009724,
final_local_scores=0.20605433, pem=0.0, similarity=10.238785]
7: 4445300 Rene_Portland
[pem_pos=0, final_scores=0.20517229,
final_local_scores=0.23917475, pem=0.0, similarity=11.227164]
1: 1209509 Pennsylvania_State_University
[pem_pos=0, final_scores=0.18605413,
final_local_scores=0.15652359, pem=-0.49028125,
similarity=11.426914]
2: 5382830 Penn_State_Nittany_Lions
[pem_pos=0, final_scores=0.1970358,
final_local_scores=0.11593631, pem=0.0, similarity=7.5494833]
3: 1209509 Pennsylvania_State_University
[pem_pos=0, final_scores=0.11771472,
final_local_scores=0.08187316, pem=-0.49028125,
similarity=9.199197]
4: 5382830 Penn_State_Nittany_Lions
[pem_pos=0, final_scores=0.22061013,
final_local_scores=0.14168766, pem=0.0, similarity=8.317955]
5: 5382830 Penn_State_Nittany_Lions
[pem_pos=0, final_scores=0.23251085,
final_local_scores=0.1546874, pem=0.0, similarity=8.7058935]
6: 5382830 Penn_State_Nittany_Lions
[pem_pos=0, final_scores=0.22127686,
final_local_scores=0.142416, pem=0.0, similarity=8.33969]
7: 606467 Don_Baylor
[pem_pos=4, final_scores=0.22302096,
final_local_scores=0.31029862, pem=-2.5183601,
similarity=27.044876]
8: 136365 Waco,_Texas
[pem_pos=0, final_scores=0.34576404,
final_local_scores=0.37113923, pem=-0.42231324,
similarity=17.461853]
9: 29810 Texas
[pem_pos=0, final_scores=0.37265307,
final_local_scores=0.4381433, pem=-0.16191302,
similarity=18.045292]
10: 5382830 Penn_State_Nittany_Lions
[pem_pos=0, final_scores=0.2889592,
final_local_scores=0.21634862, pem=0.0, similarity=10.545987]
11: 5382830 Penn_State_Nittany_Lions
[pem_pos=0, final_scores=0.27410692,
final_local_scores=0.20012477, pem=0.0, similarity=10.061835]

12: 5382830 Penn_State_Nittany_Lions
 [pem_pos=0, final_scores=0.2765774,
 final_local_scores=0.20282346, pem=0.0, similarity=10.142369]
 13: 1209509 Pennsylvania_State_University
 [pem_pos=0, final_scores=0.19390775,
 final_local_scores=0.16510248, pem=-0.49028125,
 similarity=11.682926]
 14: 5382830 Penn_State_Nittany_Lions
 [pem_pos=0, final_scores=0.26204938,
 final_local_scores=0.18695375, pem=0.0, similarity=9.668786]
 15: 669931 Division_I_(NCAA)
 [pem_pos=1, final_scores=0.04978256,
 final_local_scores=-0.0024681091, pem=-0.7993879,
 similarity=8.363255]

Sample: b'Tom_Baxter'

_ -LRB- footballer born 1903 -RRB- Thomas William `` Tom ''
 Baxter -LRB- 2 February 1903 -- 21 August 1987 -RRB- was an [1
 1 English 1] 2 footballer who played on the 3 left-wing for
 Welbeck Colliery , [1 1 Newark Town 1] , [2 2 Worksop Town 2]
 , [3 3 Mansfield Town 3] , [4 4 Wolverhampton Wanderers 4] , [5
 5 Port Vale 5] , [6 6 Margate 6] , [7 7 Carlisle United
 7] , and [8 8 Distillery 8] . He helped the `` [2 Valiants 2]
 '' to win the [9 9 Third Division North 9] title in 1929 --
 30 . Baxter played amateur football for Welbeck Colliery , [10
 10 Newark Town 10] and [11 11 Worksop Town 11] -LRB- [12 12 Mi
 dland League 12] -RRB- . He began his professional career
 with his hometown club [13 13 Mansfield Town 13] . He moved
 to [14 14 Second Division 14] club [15 15 Wolverhampton Wande
 rers 15] in late 1927 , and made his `` [3 Wolves 3] '' debut
 on 17 December 1927 in a 2 -- 0 defeat to [16 16 Chelsea 16]
 at [17 17 Stamford Bridge 17] . He remained in the team for
 the remainder of the 4 1927 -- 28 season and played the
 majority of the 5 1928 -- 29 campaign . An attacking player ,
 he scored 15 goals in a total of 53 appearances for the
 Midlanders . He left [18 18 Molineux 18] to join [19 19 Port V
 ale 19] in August 1929 . He was a regular member of the `` [4
 Valiants 4] '' 6 1929 -- 30 [20 20 Third Division North 20]
 winning side , claiming nine goals in 43 appearances . He
 lost his first team place in December 1930 , and featured in
 just 16 games of the 7 1930 -- 31 season , as the Vale
 achieved a high of fifth in the [5 Second Division 5] . He
 left [21 21 The Old Recreation Ground 21] at the end of the
 campaign and returned to [6 Mansfield 6] . He later spent two
 spells with [22 22 Margate 22] and also played for [23 23 Car
 lisle United 23] and [24 24 Distillery 24] -LRB- [7 Irish
 League 7]

1: 3943524 Newark_Town_F.C.
 2: 3552498 Worksop_Town_F.C.
 3: 451164 Mansfield_Town_F.C.
 4: 217730 Wolverhampton_Wanderers_F.C.
 5: 203456 Port_Vale_F.C.

6: 451203 Margate_F.C.
 7: 451157 Carlisle_United_F.C.
 8: 1343502 Lisburn_Distillery_F.C.
 9: 580901 Football_League_Third_Division_North
 10: 3943524 Newark_Town_F.C.
 11: 3552498 Worksop_Town_F.C.
 12: 3761630 Midland_Football_League
 13: 451164 Mansfield_Town_F.C.
 14: 563379 Football_League_Second_Division
 15: 217730 Wolverhampton_Wanderers_F.C.
 16: 7473 Chelsea_F.C.
 17: 666456 Stamford_Bridge_(stadium)
 18: 262164 Molineux_Stadium
 19: 203456 Port_Vale_F.C.
 20: 580901 Football_League_Third_Division_North
 21: 20665164 The_Old_Recreation_Ground
 22: 451203 Margate_F.C.
 23: 451157 Carlisle_United_F.C.
 24: 1343502 Lisburn_Distillery_F.C.
 1: 18803164 English_people
 [winner: 9316 England, pem_pos=1, final_scores=0.4385814,
 final_local_scores=0.4974236, pem=-1.3527626,
 similarity=26.290358]
 gt: 18803164 English_people, pem_pos=2,
 final_scores=-0.40763783, final_local_scores=-0.40633526,
 pem=-2.2084396, similarity=3.9736958]
 2: 10568 Association_football
 [winner: 10568 Association_football, pem_pos=0,
 final_scores=-1.439574, final_local_scores=-1.5645549,
 pem=0.0, similarity=-42.599724]
 gt: 10568 Association_football, pem_pos=0,
 final_scores=-1.439574, final_local_scores=-1.5645549,
 pem=0.0, similarity=-42.599724]
 3: 548981 Midfielder
 [winner: 484772 Ginger_Group, pem_pos=2,
 final_scores=-1.42636, final_local_scores=-1.4245293,
 pem=-1.0982366, similarity=-32.448704]
 gt: 548981 Midfielder, pem_pos=9,
 final_scores=-2.5512087, final_local_scores=-2.7644997,
 pem=-6.5874248, similarity=-42.58503]
 4: 16971415 1927-28_Football_League
 [winner: 34939 1927, pem_pos=0, final_scores=-0.94524163,
 final_local_scores=-0.8217801, pem=-0.7312665,
 similarity=-16.457104]
 RECALL MISS: 16971415 1927-28_Football_League]
 [winner: 1639 April_28, pem_pos=0,
 final_scores=-0.9538844, final_local_scores=-0.82031345,
 pem=-2.8589544, similarity=-4.8426385]
 RECALL MISS: 16971415 1927-28_Football_League]
 [winner: 38340996 Methylene_bridge, pem_pos=0,
 final_scores=-1.2201132, final_local_scores=-1.0993531,
 pem=0.0, similarity=-28.717176]

RECALL MISS: 16971415 1927-28_Football_League]
5: 16971422 1928-29_Football_League
[winner: 19355 May_29, pem_pos=2,
final_scores=-0.7501909, final_local_scores=-0.5481416,
pem=-2.9972408, similarity=4.0315366]
RECALL MISS: 16971422 1928-29_Football_League]
[winner: 34674 1928, pem_pos=0, final_scores=-1.080074,
final_local_scores=-1.002531, pem=-0.8454088,
similarity=-21.230349]
RECALL MISS: 16971422 1928-29_Football_League]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.316748, final_local_scores=-1.204912,
pem=0.0, similarity=-31.867262]
RECALL MISS: 16971422 1928-29_Football_League]
6: 29915019 1929-30_Port_Vale_F.C._season
[winner: 19654 May_30, pem_pos=2,
final_scores=-0.99461174, final_local_scores=-0.8515165,
pem=-2.998547, similarity=-5.0146728]
RECALL MISS: 29915019 1929-30_Port_Vale_F.C._season]
[winner: 2500584 1929_NFL_season, pem_pos=3,
final_scores=-1.1716125, final_local_scores=-1.1845467,
pem=-3.6330893, similarity=-11.502217]
RECALL MISS: 29915019 1929-30_Port_Vale_F.C._season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.2882658, final_local_scores=-1.1737995,
pem=0.0, similarity=-30.938808]
RECALL MISS: 29915019 1929-30_Port_Vale_F.C._season]
7: 29958413 1930-31_Port_Vale_F.C._season
[winner: 19653 May_31, pem_pos=1,
final_scores=-0.82675326, final_local_scores=-0.63546455,
pem=-2.705463, similarity=-0.16108906]
RECALL MISS: 29958413 1930-31_Port_Vale_F.C._season]
[winner: 34911 1930, pem_pos=0, final_scores=-0.97909844,
final_local_scores=-0.8404341, pem=-0.94197285,
similarity=-15.867921]
RECALL MISS: 29958413 1930-31_Port_Vale_F.C._season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3121464, final_local_scores=-1.1998854,
pem=0.0, similarity=-31.717257]
RECALL MISS: 29958413 1930-31_Port_Vale_F.C._season]
1: 3943524 Newark_Town_F.C.
[pem_pos=0, final_scores=0.22693585,
final_local_scores=0.21174487, pem=0.0, similarity=10.408603]
2: 3552498 Worksop_Town_F.C.
[pem_pos=0, final_scores=0.34148508,
final_local_scores=0.29388416, pem=0.0, similarity=12.859802]
3: 451164 Mansfield_Town_F.C.
[pem_pos=0, final_scores=0.5671089,
final_local_scores=0.51358116, pem=0.0, similarity=19.415998]
4: 217730 Wolverhampton_Wanderers_F.C.
[pem_pos=0, final_scores=0.45418864,
final_local_scores=0.38515192, pem=0.0, similarity=15.583415]

5: 203456 Port_Vale_F.C.
[pem_pos=0, final_scores=0.58011156,
final_local_scores=0.5279976, pem=0.0, similarity=19.846214]
6: 451203 Margate_F.C.
[pem_pos=1, final_scores=0.45115465,
final_local_scores=0.411184, pem=-1.220816,
similarity=22.99925]
7: 451157 Carlisle_United_F.C.
[pem_pos=0, final_scores=0.5584083,
final_local_scores=0.5039591, pem=0.0, similarity=19.128859]
8: 1343502 Lisburn_Distillery_F.C.
[pem_pos=0, final_scores=0.45323056,
final_local_scores=0.4350729, pem=0.0, similarity=17.073156]
9: 580901 Football_League_Third_Division_North
[pem_pos=0, final_scores=0.39153582,
final_local_scores=0.33313942, pem=0.0, similarity=14.031258]
10: 3943524 Newark_Town_F.C.
[pem_pos=0, final_scores=0.24928994,
final_local_scores=0.23616335, pem=0.0, similarity=11.137299]
11: 3552498 Worksop_Town_F.C.
[pem_pos=0, final_scores=0.27298927,
final_local_scores=0.21906286, pem=0.0, similarity=10.626986]
12: 3761630 Midland_Football_League
[pem_pos=0, final_scores=0.19742917,
final_local_scores=0.13303259, pem=0.0, similarity=8.05967]
13: 451164 Mansfield_Town_F.C.
[pem_pos=0, final_scores=0.4963197,
final_local_scores=0.43625468, pem=0.0, similarity=17.108423]
14: 563379 Football_League_Second_Division
[pem_pos=0, final_scores=0.18536334,
final_local_scores=0.095206946, pem=-0.48164165,
similarity=9.55012]
15: 217730 Wolverhampton_Wanderers_F.C.
[pem_pos=0, final_scores=0.42221397,
final_local_scores=0.35022452, pem=0.0, similarity=14.541111]
16: 7473 Chelsea_F.C.
[pem_pos=0, final_scores=0.7131034,
final_local_scores=0.7050569, pem=-0.51707125,
similarity=27.94193]
17: 666456 Stamford_Bridge_(stadium)
[pem_pos=0, final_scores=0.512108,
final_local_scores=0.4989829, pem=0.0, similarity=18.980358]
18: 262164 Molineux_Stadium
[pem_pos=0, final_scores=0.36195147,
final_local_scores=0.30612794, pem=0.0, similarity=13.225182]
19: 203456 Port_Vale_F.C.
[pem_pos=0, final_scores=0.532808,
final_local_scores=0.47632563, pem=0.0, similarity=18.30422]
20: 580901 Football_League_Third_Division_North
[pem_pos=0, final_scores=0.43246478,
final_local_scores=0.3778481, pem=0.0, similarity=15.365454]
21: 20665164 The_Old_Recreation_Ground

[pem_pos=0, final_scores=0.23620282,
 final_local_scores=0.19378403, pem=0.0, similarity=9.872615]
 22: 451203 Margate_F.C.
 [pem_pos=1, final_scores=0.4393931,
 final_local_scores=0.39833623, pem=-1.220816,
 similarity=22.615849]
 23: 451157 Carlisle_United_F.C.
 [pem_pos=0, final_scores=0.43555474,
 final_local_scores=0.36976022, pem=0.0, similarity=15.124096]
 24: 1343502 Lisburn_Distillery_F.C.
 [pem_pos=0, final_scores=0.3505345,
 final_local_scores=0.32289302, pem=0.0, similarity=13.725485]
 1: 9316 England
 [pem_pos=1, final_scores=0.4385814,
 final_local_scores=0.4974236, pem=-1.3527626,
 similarity=26.290358]
 2: 203456 Port_Vale_F.C.
 [pem_pos=1, final_scores=0.0747091,
 final_local_scores=-0.025495589, pem=-2.7725887,
 similarity=18.406643]
 3: 217730 Wolverhampton_Wanderers_F.C.
 [pem_pos=0, final_scores=0.61932564,
 final_local_scores=0.565539, pem=-0.29574642,
 similarity=22.574839]
 4: 203456 Port_Vale_F.C.
 [pem_pos=1, final_scores=0.12776558,
 final_local_scores=0.033878386, pem=-2.7725887,
 similarity=20.17848]
 5: 563379 Football_League_Second_Division
 [pem_pos=0, final_scores=0.19290347,
 final_local_scores=0.103443384, pem=-0.48164165,
 similarity=9.795912]
 6: 451164 Mansfield_Town_F.C.
 [pem_pos=16, final_scores=0.38952094,
 final_local_scores=0.31959325, pem=-4.267106,
 similarity=36.8322]
 7: 4262952 NIFL_Premiership
 [pem_pos=0, final_scores=0.15123911,
 final_local_scores=0.12817359, pem=0.0, similarity=7.9146686]

Sample: b'World_Without_Superman'

`` World Without Superman `` is a [1 1 Superman 1] 1 comic bo
 ok 2 story arc published by [2 2 DC Comics 2] . It takes
 place in [3 3 Action Comics 3] written by [4 4 Greg Rucka 4]
 with art by and [1 3 Superman 1] written by with art by . The
 story deals with Metropolis dealing with a world without
 Superman , who has gone to live on [2 New Krypton 2] to keep
 General [3 Zod 3] in check . As a result , the two [4
 Superman 4] series , [5 Action Comics 5] and [6 Superman 6]
 star Nightwing & Flamebird and [7 Mon-El 7] respectively .
 When [8 Superman 8] leaves Earth for [9 New Krypton 9] , he
 appoints , newly freed from the [5 5 Phantom Zone 5] , to

take his place as guardian of 4 Metropolis . [10 Mon-El 10] assumes the secret identity of [11 Johnathan Kent 11] as a tribute to Clark 's adoptive father , posing as Clark 's cousin . The series directly leads into the Codename Patriot crossover . Kryptonians [12 5 Chris Kent 12] -LRB- comics - RRB- and [6 6 Thara Ak-Var 6] become the superheroes [7 7 Nightwing 7] and [8 8 Flamebird 8] on Earth and are being hunted by . The mission of the two superheroes is to hunt down [13 Zod 13] 's sleeper agents on Earth and return them to the [14 Phantom Zone 14] . They start by hunting down sleeper agent Tor-Ann , who was secretly posing as a human in [15 Australia 15] . Next , they are confronted by Ursa , who sadistically slashes [16 Flamebird 16] with a [17 Kryptonite 17] knife , but is taken down by [18 Nightwing 18] before Ursa can kill Thara . Chris and Thara survive the encounter , and then go see Chris 's adoptive mother , Lois . Lois sees to it that Thara is cured from the [19 Kryptonite 19] attack with the help of . The next mission of the Kryptonian duo deals with taking down Nadira Var-Em and Az-Rel , who are described as Krypton 's equivalent of Bonnie and Clyde . This ends with [20 Nightwing 20] and [21 Flamebird 21] losing the track of the two criminals , who are held captive by Sam Lane

1: 28381 Superman

2: 9105 DC_Comics

3: 81284 Action_Comics

4: 634686 Greg_Rucka

5: 233625 Phantom_Zone

6: 21936099 Thara_Ak-Var

7: 335590 Nightwing

8: 3179032 Flamebird

1: 6231 Comic_book

[winner: 3284403 Belphegor_(comics), pem_pos=6,
final_scores=-1.2946631, final_local_scores=-1.291378,
pem=-3.048396, similarity=-17.869928]

gt: 6231 Comic_book, pem_pos=1, final_scores=-1.7951977,
final_local_scores=-1.9751158, pem=-1.9606276,
similarity=-44.189484]

[winner: 9199046 When_Prophecy_Fails, pem_pos=10,
final_scores=-1.5444638, final_local_scores=-1.5612304,
pem=-4.3429904, similarity=-18.88266]

RECALL MISS: 6231 Comic_book]

2: 529844 Story_arc

[winner: 4799434 Two_Gallants_(short_story), pem_pos=6,
final_scores=-1.0648309, final_local_scores=-0.9807689,
pem=-2.3978953, similarity=-12.138261]

RECALL MISS: 529844 Story_arc]

[winner: 3082731 Arc_(geometry), pem_pos=0,
final_scores=-1.1335073, final_local_scores=-1.1086247,
pem=-0.4334168, similarity=-26.636873]

gt: 529844 Story_arc, pem_pos=6,
final_scores=-1.6068081, final_local_scores=-1.7870586,
pem=-2.7352626, similarity=-34.364895]

```
3: 3215786 Superman_(comic_book)
  [winner: 28381 Superman, pem_pos=0,
  final_scores=0.0882159, final_local_scores=0.0015880316,
  pem=-0.21335615, similarity=5.297367]
  gt: 3215786 Superman_(comic_book), pem_pos=1,
  final_scores=-0.40284586, final_local_scores=-0.48934016,
  pem=-2.7411907, similarity=4.3938475 ]
4: 112176 Metropolis_(comics)
  [winner: 112176 Metropolis_(comics), pem_pos=0,
  final_scores=-0.13917625, final_local_scores=-0.21292415,
  pem=-1.1092179, similarity=3.7677312]
  gt: 112176 Metropolis_(comics), pem_pos=0,
  final_scores=-0.13917625, final_local_scores=-0.21292415,
  pem=-1.1092179, similarity=3.7677312 ]
5: 8340183 Chris_Kent_(comics)
  [winner: 8340183 Chris_Kent_(comics), pem_pos=0,
  final_scores=0.25762975, final_local_scores=0.23324773,
  pem=0.0, similarity=11.05029]
  gt: 8340183 Chris_Kent_(comics), pem_pos=0,
  final_scores=0.25762975, final_local_scores=0.23324773,
  pem=0.0, similarity=11.05029 ]
  [winner: 3175886 I._W._Publications, pem_pos=13,
  final_scores=-0.95506626, final_local_scores=-0.9104783,
  pem=-2.5649493, similarity=-9.132183]
  RECALL MISS: 8340183 Chris_Kent_(comics)]
1: 28381 Superman
  [pem_pos=0, final_scores=0.07962377,
  final_local_scores=-0.010010138, pem=-0.21335615,
  similarity=4.9512544]
2: 9105 DC_Comics
  [pem_pos=0, final_scores=0.2598166,
  final_local_scores=0.24249107, pem=0.0, similarity=11.32613]
3: 81284 Action_Comics
  [pem_pos=0, final_scores=0.13314445,
  final_local_scores=0.09586978, pem=0.0, similarity=6.9506583]
4: 634686 Greg_Rucka
  [pem_pos=0, final_scores=0.18903099,
  final_local_scores=0.19556186, pem=0.0, similarity=9.925669]
5: 233625 Phantom_Zone
  [pem_pos=0, final_scores=0.18943219,
  final_local_scores=0.15975904, pem=0.0, similarity=8.857242]
6: 21936099 Thara_Ak-Var
  [pem_pos=0, final_scores=0.29740757,
  final_local_scores=0.2860293, pem=0.0, similarity=12.625398]
7: 335590 Nightwing
  [pem_pos=0, final_scores=0.14028282,
  final_local_scores=0.07312484, pem=-0.2273638,
  similarity=7.5083437]
8: 3179032 Flamebird
  [pem_pos=0, final_scores=0.14910917,
  final_local_scores=0.12284836, pem=-0.24177234,
  similarity=9.070548]
```

1: 28381 Superman
[pem_pos=0, final_scores=0.0882159,
final_local_scores=0.0015880316, pem=-0.21335615,
similarity=5.297367]
2: 563279 Kandor
[pem_pos=0, final_scores=0.18589063,
final_local_scores=0.1575219, pem=-0.0126583865,
similarity=8.859319]
3: 864458 General_Zod
[pem_pos=0, final_scores=0.11256941,
final_local_scores=0.0598803, pem=0.0, similarity=5.8766603]
4: 28381 Superman
[pem_pos=0, final_scores=0.09695996,
final_local_scores=0.011139587, pem=-0.21335615,
similarity=5.5824046]
5: 81284 Action_Comics
[pem_pos=0, final_scores=0.15381654,
final_local_scores=0.11845088, pem=0.0, similarity=7.624523]
6: 28381 Superman
[pem_pos=0, final_scores=0.10863505,
final_local_scores=0.02389285, pem=-0.21335615,
similarity=5.9629874]
7: 238071 Lar_Gand
[pem_pos=0, final_scores=0.090324745,
final_local_scores=0.038634017, pem=0.0, similarity=5.242629]
8: 28381 Superman
[pem_pos=0, final_scores=0.103430375,
final_local_scores=0.01820755, pem=-0.21335615,
similarity=5.7933264]
9: 563279 Kandor
[pem_pos=0, final_scores=0.120665595,
final_local_scores=0.08627346, pem=-0.0126583865,
similarity=6.7331233]
10: 238071 Lar_Gand
[pem_pos=0, final_scores=0.06630011,
final_local_scores=0.012390733, pem=0.0, similarity=4.459477]
11: 1624138 Jonathan_and_Martha_Kent
[pem_pos=0, final_scores=0.31698328,
final_local_scores=0.29779646, pem=0.0, similarity=12.976553]
12: 8340183 Chris_Kent_(comics)
[pem_pos=0, final_scores=0.25762975,
final_local_scores=0.23324773, pem=0.0, similarity=11.05029]
13: 864458 General_Zod
[pem_pos=0, final_scores=0.076790795,
final_local_scores=0.02079761, pem=0.0, similarity=4.7103553]
14: 233625 Phantom_Zone
[pem_pos=0, final_scores=0.17924796,
final_local_scores=0.14863434, pem=0.0, similarity=8.525258]
15: 4689264 Australia
[pem_pos=0, final_scores=0.5546375,
final_local_scores=0.7231461, pem=-0.17168477,
similarity=26.603483]

16: 3179032 Flamebird
 [pem_pos=0, final_scores=0.091662005,
 final_local_scores=0.058743298, pem=-0.24177234,
 similarity=7.1575255]
 17: 17390 Kryptonite
 [pem_pos=0, final_scores=0.15184163,
 final_local_scores=0.11593363, pem=-0.2047345,
 similarity=8.662782]
 18: 335590 Nightwing
 [pem_pos=0, final_scores=0.16635515,
 final_local_scores=0.10287519, pem=-0.2273638,
 similarity=8.396153]
 19: 17390 Kryptonite
 [pem_pos=0, final_scores=0.070170805,
 final_local_scores=0.026720643, pem=-0.2047345,
 similarity=6.0004883]
 20: 335590 Nightwing
 [pem_pos=0, final_scores=0.09202646,
 final_local_scores=0.020412117, pem=-0.2273638,
 similarity=5.935291]
 21: 3179032 Flamebird
 [pem_pos=0, final_scores=0.12233411,
 final_local_scores=0.09417686, pem=-0.24177234,
 similarity=8.214933]

Sample: b'Bill_Schulz'

William Dawes `` Bill ' ' Schulz -LRB- born August 14 , 1975 -
 RRB- is a regular panelist , writer , and producer on [1 1 Fox
 News Channel 1] 's late night show , , since its debut on
 February 5 , 2007 . Serving as host [1 1 Greg Gutfeld 1] 's
 `` repulsive sidekick , ' ' who is routinely the target of [2
 Gutfeld 2] 's running gags , Schulz often looks directly into
 the camera -LRB- even when he is not being talked to -RRB-
 with his signature 2 `` crazy-eyed look , ' ' along with
 frequently 3 waving to the television viewing audience .
 Schulz provides the voice for an of [2 2 The New York Times 2]
 newspaper , named `` Pinch ' ' -LRB- a reference to [3 4 Arthur
 Ochs Sulzberger 3] , Jr. . , current publisher of the [4
 Times 4] -RRB- . Schulz is also a freelance writer and a
 former senior editor of . Schulz was born in [5 5 Lake Forest
 5] , [6 Illinois 6] . He attended high school in [3 3 Illinois
 3] at [4 4 Lake Forest Academy 4] , and in 1998 received a BA
 in Print Journalism with a minor in Theater graduating from [5
 5 Emerson College 5] . In April 2009 , Schulz made a notable
 cameo on [6 6 Glenn Beck 6] 's TV program , . [7 Beck 7]
 described Schulz as the `` average [8 American 8] ' ' and
 pretended to pour gasoline -LRB- actually water -RRB- on him
 from a red container during [9 Beck 9] 's monologue about the
 effects of President [7 7 Barack Obama 7] 's policies . [10
 Beck 10] discussed setting Schulz on fire . During the April
 11 , 2013 episode of 6 The Five , [8 8 Dana Perino 8] , aka `
 Tiny D Day-P ' , responded to [9 9 Jay-Z 9] 's [11 Cuba 11]

rap track with her own rap , which was written by [12 Bill Schulz 12] . [13 Perino 13] s Schulz ' lyrics : Well my name is Tiny-D , and I 'm here to say , I bust funky-fresh rhymes in a major way . So I 'm white like 7 Casper , got a dog named Jasper , and if you do n't think fears me , go ahead and ask her . So if you love , stick with 8 Jay-Z

1: 5755576 Greg_Gutfeld
2: 30680 The_New_York_Times
3: 14849 Illinois
4: 2608961 Lake_Forest_Academy
5: 401950 Emerson_College
6: 418944 Glenn_Beck
7: 534366 Barack_Obama
8: 6689554 Dana_Perino
9: 162870 Jay-Z

1: 11121 Fox_News_Channel

[winner: 11121 Fox_News_Channel, pem_pos=0,
final_scores=0.50630766, final_local_scores=0.48254305,
pem=0.0, similarity=18.48976|

gt: 11121 Fox_News_Channel, pem_pos=0,
final_scores=0.50630766, final_local_scores=0.48254305,
pem=0.0, similarity=18.48976]

[winner: 1415027 David_Berman_(musician), pem_pos=5,
final_scores=-2.1724527, final_local_scores=-2.355486,
pem=-3.0012724, similarity=-49.881298|

RECALL MISS: 11121 Fox_News_Channel]

2: 242760 Facial_expression

[winner: 4471136 Look_(Beth_Nielsen_Chapman_album),
pem_pos=5, final_scores=-1.1330082,
final_local_scores=-1.0410438, pem=-2.7927067,
similarity=-11.789933|

RECALL MISS: 242760 Facial_expression]

3: 403676 Gesture

[winner: 5135140 Waving_(dance), pem_pos=0,
final_scores=-1.0714182, final_local_scores=-0.9995636,
pem=0.0, similarity=-25.739258|

gt: 403676 Gesture, pem_pos=1, final_scores=-2.2201827,
final_local_scores=-2.321892, pem=-5.9738097,
similarity=-32.713673]

4: 2190157 Arthur_Ochs_Sulzberger,_Jr.

[winner: 2171164 Arthur_Ochs_Sulzberger, pem_pos=0,
final_scores=0.23572205, final_local_scores=0.31056583,
pem=0.0, similarity=13.357616|

gt: 2190157 Arthur_Ochs_Sulzberger,_Jr., pem_pos=1,
final_scores=0.17049627, final_local_scores=0.20593774,
pem=-0.6931472, similarity=14.004748]

[winner: 16861 Kurt_Vonnegut, pem_pos=5,
final_scores=-0.26210344, final_local_scores=-0.28861228,
pem=-2.3999593, similarity=8.528299|

RECALL MISS: 2190157 Arthur_Ochs_Sulzberger,_Jr.]

[winner: 4721653 Zyphe_language, pem_pos=12,
final_scores=-1.601348, final_local_scores=-1.5581424,

pem=-4.4188404, similarity=-18.378025|
 RECALL MISS: 2190157 Arthur_Ochs_Sulzberger,_Jr.]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-2.052768, final_local_scores=-2.058325,
 pem=-4.6728287, similarity=-31.923244|
 RECALL MISS: 2190157 Arthur_Ochs_Sulzberger,_Jr.]
 5: 111471 Lake_Forest,_Illinois
 [winner: 14849 Illinois, pem_pos=0,
 final_scores=0.55492693, final_local_scores=0.5730835,
 pem=-0.16995399, similarity=22.115904|
 RECALL MISS: 111471 Lake_Forest,_Illinois]
 [winner: 111471 Lake_Forest,_Illinois, pem_pos=0,
 final_scores=0.24808733, final_local_scores=0.2900951,
 pem=-0.7013507, similarity=16.560783|
 gt: 111471 Lake_Forest,_Illinois, pem_pos=0,
 final_scores=0.24808733, final_local_scores=0.2900951,
 pem=-0.7013507, similarity=16.560783]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7906253, final_local_scores=-1.7719738,
 pem=-4.6728287, similarity=-23.37795|
 RECALL MISS: 111471 Lake_Forest,_Illinois]
 6: 32267907 The_Five_(TV_program)
 [winner: 32267907 The_Five_(TV_program), pem_pos=0,
 final_scores=-0.0017748177, final_local_scores=0.0141854435,
 pem=-0.79094493, similarity=8.814317|
 gt: 32267907 The_Five_(TV_program), pem_pos=0,
 final_scores=-0.0017748177, final_local_scores=0.0141854435,
 pem=-0.79094493, similarity=8.814317]
 7: 21132349 Casper_the_Friendly_Ghost
 [winner: 21132349 Casper_the_Friendly_Ghost, pem_pos=1,
 final_scores=-0.3335217, final_local_scores=-0.332385,
 pem=-1.2110755, similarity=0.75669885|
 gt: 21132349 Casper_the_Friendly_Ghost, pem_pos=1,
 final_scores=-0.3335217, final_local_scores=-0.332385,
 pem=-1.2110755, similarity=0.75669885]
 8: 162870 Jay-Z
 [winner: 162870 Jay-Z, pem_pos=0,
 final_scores=-0.13873972, final_local_scores=-0.16042648,
 pem=0.0, similarity=-0.6977333|
 gt: 162870 Jay-Z, pem_pos=0, final_scores=-0.13873972,
 final_local_scores=-0.16042648, pem=0.0,
 similarity=-0.6977333]
 1: 5755576 Greg_Gutfeld
 [pem_pos=0, final_scores=0.4239927,
 final_local_scores=0.45026594, pem=0.0, similarity=17.526546]
 2: 30680 The_New_York_Times
 [pem_pos=0, final_scores=0.5946594,
 final_local_scores=0.60640454, pem=0.0, similarity=22.186033]
 3: 14849 Illinois
 [pem_pos=0, final_scores=0.452529,
 final_local_scores=0.46122938, pem=-0.16995399,
 similarity=18.777954]

4: 2608961 Lake_Forest_Academy
[pem_pos=0, final_scores=0.23102395,
final_local_scores=0.28302756, pem=0.0, similarity=12.53582]
5: 401950 Emerson_College
[pem_pos=0, final_scores=0.29505426,
final_local_scores=0.30051634, pem=0.0, similarity=13.05772]
6: 418944 Glenn_Beck
[pem_pos=0, final_scores=0.19829176,
final_local_scores=0.16817525, pem=0.0, similarity=9.1083975]
7: 534366 Barack_Obama
[pem_pos=0, final_scores=0.33061218,
final_local_scores=0.3232268, pem=0.0, similarity=13.735446]
8: 6689554 Dana_Perino
[pem_pos=0, final_scores=0.5083799,
final_local_scores=0.5497802, pem=0.0, similarity=20.49625]
9: 162870 Jay-Z
[pem_pos=0, final_scores=0.11384334,
final_local_scores=0.124376595, pem=0.0, similarity=7.8013587]
1: 11121 Fox_News_Channel
[pem_pos=0, final_scores=0.50630766,
final_local_scores=0.48254305, pem=0.0, similarity=18.48976]
2: 5755576 Greg_Gutfeld
[pem_pos=0, final_scores=0.18300505,
final_local_scores=0.18702355, pem=0.0, similarity=9.670869]
3: 2171164 Arthur_Ochs_Sulzberger
[pem_pos=0, final_scores=0.23572205,
final_local_scores=0.31056583, pem=0.0, similarity=13.357616]
4: 39127 The_Times
[pem_pos=0, final_scores=0.19724698,
final_local_scores=0.19088906, pem=-0.25275695,
similarity=11.160755]
5: 111471 Lake_Forest,_Illinois
[pem_pos=0, final_scores=0.24808733,
final_local_scores=0.2900951, pem=-0.7013507,
similarity=16.560783]
6: 14849 Illinois
[pem_pos=0, final_scores=0.55492693,
final_local_scores=0.5730835, pem=-0.16995399,
similarity=22.115904]
7: 418944 Glenn_Beck
[pem_pos=0, final_scores=0.44051248,
final_local_scores=0.43655026, pem=0.0, similarity=17.117243]
8: 3434750 United_States
[pem_pos=0, final_scores=0.352266,
final_local_scores=0.35826674, pem=-0.12994486,
similarity=15.487767]
9: 418944 Glenn_Beck
[pem_pos=0, final_scores=0.27017438,
final_local_scores=0.24602923, pem=0.0, similarity=11.431716]
10: 418944 Glenn_Beck
[pem_pos=0, final_scores=0.3543635,
final_local_scores=0.33799303, pem=0.0, similarity=14.176099]

11: 5042481 Cuba
 [pem_pos=0, final_scores=0.26777488,
 final_local_scores=0.33406356, pem=-0.23907764,
 similarity=15.358976]
 12: 11670747 Bill_Schulz
 [pem_pos=0, final_scores=0.085832104,
 final_local_scores=0.111660704, pem=0.0, similarity=7.421891]
 13: 6689554 Dana_Perino
 [pem_pos=0, final_scores=0.2730888,
 final_local_scores=0.2927604, pem=0.0, similarity=12.826267]

Sample: b'Brian_Tamberlin'

[1 Brian Tamberlin 1] , 1 BA , 2 LLB , 3 LLM , 4 QC is a 5 re
 tired 6 jurist , barrister , commentator and 7 justice in the
 [2 8 Australia 2] 's 9 federal court . [3 Tamberlin 3]
 received both is 10 bachelor 's of 11 arts and 12 law from
 the [1 1 University of Sydney 1] . He later received his in
 law from [2 2 Harvard Law School 2] . [4 Tamberlin 4] was
 called to the [3 3 New South Wales 3] bar in 1967
 specializing in 13 town planning law , 14 commercial law and 1
 5 administrative law and , and became a 16 Queen 's Counsel
 in 1981 . [5 Tamberlin 5] was appointed a federal court
 justice in 1994 . [6 Tamberlin 6] engaged in lecturing
 initiatives in legal areas such as 17 federal constitutional l
 aw , 18 administrative law and 19 intellectual property in
 several countries such as the [4 4 United States 4] , [5 5 Vi
 etnam 5] and [6 6 Thailand 6] . He traveled to [7 7 Indonesia
 7] in [8 8 AusAID 8] programs operated by the Law Faculty for
 [7 Indonesian 7] judges , with 20 professor 21 Philip Lewis Gr
 iffiths . [8 Tamberlin 8] has also served as a justice in the
 22 Supreme Court of the [9 Australian Capital Territory 9] .
 Following [10 Tamberlin 10] 's retirement from the federal
 court of [11 Australia 11] in March 2009 , he became the
 Deputy President of the 23 Administrative Appeals Tribunal ,
 an independent body which reviews the decisions of the . On
 27 February 2009 , A ceremony marking the retirement of
 Justice [12 Tamberlin 12] from the federal court was held .
 Recently , [13 Tamberlin 13] headed a royal commission-style
 inquiry that was investigating the former [14 24 Labor 14] go
 vernment 's

1: 31928 University_of_Sydney
 2: 158984 Harvard_Law_School
 3: 21654 New_South_Wales
 4: 3434750 United_States
 5: 202354 Vietnam
 6: 30128 Thailand
 7: 14579 Indonesia
 8: 3053483 AusAID

1: 156992 Bachelor_of_Arts
 [winner: 156992 Bachelor_of_Arts, pem_pos=0,
 final_scores=-0.21607275, final_local_scores=-0.22604378,
 pem=-0.2940791, similarity=-1.0566363]

gt: 156992 Bachelor_of_Arts, pem_pos=0,
final_scores=-0.21607275, final_local_scores=-0.22604378,
pem=-0.2940791, similarity=-1.0566363]
2: 299903 Bachelor_of_Laws
[winner: 299903 Bachelor_of_Laws, pem_pos=0,
final_scores=-0.16829309, final_local_scores=-0.17904659,
pem=0.0, similarity=-1.2533945]
gt: 299903 Bachelor_of_Laws, pem_pos=0,
final_scores=-0.16829309, final_local_scores=-0.17904659,
pem=0.0, similarity=-1.2533945]
3: 464522 Master_of_Laws
[winner: 464522 Master_of_Laws, pem_pos=0,
final_scores=-0.16946845, final_local_scores=-0.18413715,
pem=0.0, similarity=-1.405307]
gt: 464522 Master_of_Laws, pem_pos=0,
final_scores=-0.16946845, final_local_scores=-0.18413715,
pem=0.0, similarity=-1.405307]
4: 90072 Queen's_Counsel
[winner: 90072 Queen's_Counsel, pem_pos=0,
final_scores=-0.18493567, final_local_scores=-0.19727126,
pem=-0.39452267, similarity=0.34822088]
gt: 90072 Queen's_Counsel, pem_pos=0,
final_scores=-0.18493567, final_local_scores=-0.19727126,
pem=-0.39452267, similarity=0.34822088]
5: 26495 Retirement
[winner: 26495 Retirement, pem_pos=1,
final_scores=-1.1683799, final_local_scores=-1.2697415,
pem=-0.36344215, similarity=-31.825453]
gt: 26495 Retirement, pem_pos=1,
final_scores=-1.1683799, final_local_scores=-1.2697415,
pem=-0.36344215, similarity=-31.825453]
6: 51481 Jurist
[winner: 1636627 Federal_judiciary_of_the_United_States,
pem_pos=0, final_scores=-1.3649379,
final_local_scores=-1.4560761, pem=0.0,
similarity=-39.362507]
gt: 51481 Jurist, pem_pos=1, final_scores=-1.3866411,
final_local_scores=-1.476888, pem=-0.06042996,
similarity=-39.65494]
7: 44719 Judge
[winner: 21826428 James_Clark_McReynolds, pem_pos=0,
final_scores=-1.053318, final_local_scores=-1.0777149,
pem=0.0, similarity=-28.071451]
gt: 44719 Judge, pem_pos=9, final_scores=-2.0346766,
final_local_scores=-2.2234976, pem=-3.6982958,
similarity=-42.151978]
8: 4689264 Australia
[winner: 4689264 Australia, pem_pos=0,
final_scores=0.56537926, final_local_scores=0.5826751,
pem=-0.17168477, similarity=22.411549]
gt: 4689264 Australia, pem_pos=0,
final_scores=0.56537926, final_local_scores=0.5826751,

pem=-0.17168477, similarity=22.411549]
 [winner: 1415027 David_Berman_(musician), pem_pos=5,
 final_scores=-1.8843372, final_local_scores=-1.9040124,
 pem=-3.0012724, similarity=-36.408424|
 RECALL MISS: 4689264 Australia]
 9: 4688076 Federal_Court_of_Australia
 [winner: 1441887
 United_States_District_Court_for_the_Western_District_of_Mich
 igan, pem_pos=2, final_scores=-1.0180002,
 final_local_scores=-0.9953628, pem=-1.5739895,
 similarity=-17.0543|
 RECALL MISS: 4688076 Federal_Court_of_Australia]
 10: 188874 Bachelor's_degree
 [winner: 9405067 The_Bachelors_(novel), pem_pos=14,
 final_scores=-1.1826104, final_local_scores=-1.0935408,
 pem=-2.7080503, similarity=-13.816927|
 gt: 188874 Bachelor's_degree, pem_pos=1,
 final_scores=-1.7260219, final_local_scores=-1.8785963,
 pem=-0.9350501, similarity=-46.8864]
 [winner: 187801 Shopping_cart, pem_pos=8,
 final_scores=-2.2129948, final_local_scores=-2.2878697,
 pem=-3.4067376, similarity=-45.658512|
 RECALL MISS: 188874 Bachelor's_degree]
 11: 156992 Bachelor_of_Arts
 [winner: 48991 Arts_and_Crafts_movement, pem_pos=0,
 final_scores=-1.1222552, final_local_scores=-1.1388118,
 pem=0.0, similarity=-29.894701|
 RECALL MISS: 156992 Bachelor_of_Arts]
 12: 299903 Bachelor_of_Laws
 [winner: 18949668 Law, pem_pos=0,
 final_scores=-1.4677317, final_local_scores=-1.6022036,
 pem=-0.18675159, similarity=-42.70765|
 gt: 299903 Bachelor_of_Laws, pem_pos=2,
 final_scores=-1.5995972, final_local_scores=-1.7425289,
 pem=-3.4596343, similarity=-29.096796]
 13: 45346 Urban_planning
 [winner: 2813519 Types_of_inhabited_localities_in_Russia,
 pem_pos=2, final_scores=-1.1543823,
 final_local_scores=-1.0557675, pem=-2.3084521,
 similarity=-14.862773|
 RECALL MISS: 45346 Urban_planning]
 [winner: 4466133 Town_and_Country_Planning_Act_1990,
 pem_pos=0, final_scores=-1.2561576,
 final_local_scores=-1.2920585, pem=-0.22314353,
 similarity=-33.254402|
 gt: 45346 Urban_planning, pem_pos=2,
 final_scores=-1.8820611, final_local_scores=-2.0464554,
 pem=-2.3025851, similarity=-44.458786]
 14: 77444 Commercial_law
 [winner: 1621419 Ensign-Bickford_Company, pem_pos=11,
 final_scores=-1.0819663, final_local_scores=-1.0069481,
 pem=-2.4842746, similarity=-12.449756|

RECALL MISS: 77444 Commercial_law]
[winner: 18949668 Law, pem_pos=0,
final_scores=-1.4476666, final_local_scores=-1.5802855,
pem=-0.18675159, similarity=-42.053577]
RECALL MISS: 77444 Commercial_law]
15: 2560 Administrative_law
[winner: 2560 Administrative_law, pem_pos=0,
final_scores=-1.1381437, final_local_scores=-1.2399187,
pem=-0.10300483, similarity=-32.35178]
gt: 2560 Administrative_law, pem_pos=0,
final_scores=-1.1381437, final_local_scores=-1.2399187,
pem=-0.10300483, similarity=-32.35178]
16: 90072 Queen's_Counsel
[winner: 605309 Counsel, pem_pos=0,
final_scores=-0.019559994, final_local_scores=0.006195724,
pem=-0.354545, similarity=6.2026763]
RECALL MISS: 90072 Queen's_Counsel]
[winner: 42010 Queen_(band), pem_pos=0,
final_scores=-0.48535162, final_local_scores=-0.40847826,
pem=-0.43144405, similarity=-5.753834]
RECALL MISS: 90072 Queen's_Counsel]
[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.539378, final_local_scores=-2.6443937,
pem=-3.4067376, similarity=-56.297897]
RECALL MISS: 90072 Queen's_Counsel]
17: 5919 Constitutional_law
[winner: 232687 Federal_law, pem_pos=0,
final_scores=-1.0886707, final_local_scores=-1.1249942,
pem=0.0, similarity=-29.482357]
RECALL MISS: 5919 Constitutional_law]
18: 2560 Administrative_law
[winner: 2560 Administrative_law, pem_pos=0,
final_scores=-1.3043588, final_local_scores=-1.4214835,
pem=-0.10300483, similarity=-37.77004]
gt: 2560 Administrative_law, pem_pos=0,
final_scores=-1.3043588, final_local_scores=-1.4214835,
pem=-0.10300483, similarity=-37.77004]
19: 14724 Intellectual_property
[winner: 14724 Intellectual_property, pem_pos=0,
final_scores=-1.2761366, final_local_scores=-1.3531075,
pem=-0.017415497, similarity=-36.195]
gt: 14724 Intellectual_property, pem_pos=0,
final_scores=-1.2761366, final_local_scores=-1.3531075,
pem=-0.017415497, similarity=-36.195]
20: 221054 Legal_education
[winner: 531186 Regius_Professor, pem_pos=0,
final_scores=-1.0682595, final_local_scores=-1.0919778,
pem=0.0, similarity=-28.497082]
RECALL MISS: 221054 Legal_education]
21: 5983935 Philip_Lewis_Griffiths
[winner: 5983935 Philip_Lewis_Griffiths, pem_pos=0,
final_scores=-0.19385315, final_local_scores=-0.13237642,

pem=0.0, similarity=0.13933678|
gt: 5983935 Philip_Lewis_Griffiths, pem_pos=0,
final_scores=-0.19385315, final_local_scores=-0.13237642,
pem=0.0, similarity=0.13933678]
22: 1970938 Supreme_Court_of_the_Australian_Capital_Territory
[winner: 1944 Australian_Capital_Territory, pem_pos=0,
final_scores=0.35492176, final_local_scores=0.3808837,
pem=0.0, similarity=15.456041|
RECALL MISS: 1970938
Supreme_Court_of_the_Australian_Capital_Territory]
[winner: 809976 Supreme_Court_of_Pennsylvania, pem_pos=0,
final_scores=-0.26867205, final_local_scores=-0.22124249,
pem=0.0, similarity=-2.512604|
RECALL MISS: 1970938
Supreme_Court_of_the_Australian_Capital_Territory]
23: 3824051 Administrative_Appeals_Tribunal
[winner: 3824051 Administrative_Appeals_Tribunal,
pem_pos=0, final_scores=-0.070059136,
final_local_scores=-0.039176553, pem=0.0,
similarity=2.9206064|
gt: 3824051 Administrative_Appeals_Tribunal, pem_pos=0,
final_scores=-0.070059136, final_local_scores=-0.039176553,
pem=0.0, similarity=2.9206064]
24: 1495 Australian_Labor_Party
[winner: 1495 Australian_Labor_Party, pem_pos=0,
final_scores=0.27251226, final_local_scores=0.32469857,
pem=-0.2258703, similarity=15.007684|
gt: 1495 Australian_Labor_Party, pem_pos=0,
final_scores=0.27251226, final_local_scores=0.32469857,
pem=-0.2258703, similarity=15.007684]
[winner: 12229 Government, pem_pos=0,
final_scores=-1.2969228, final_local_scores=-1.4196153,
pem=-0.3899827, similarity=-36.153652|
RECALL MISS: 1495 Australian_Labor_Party]
[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-1.8844109, final_local_scores=-1.9289417,
pem=-3.4067376, similarity=-34.947388|
RECALL MISS: 1495 Australian_Labor_Party]
1: 31928 University_of_Sydney
[pem_pos=0, final_scores=0.35136455,
final_local_scores=0.3745311, pem=0.0, similarity=15.266467]
2: 158984 Harvard_Law_School
[pem_pos=0, final_scores=0.2982489,
final_local_scores=0.35524064, pem=0.0, similarity=14.690803]
3: 21654 New_South_Wales
[pem_pos=0, final_scores=0.4695403,
final_local_scores=0.49966872, pem=-0.051332146,
similarity=19.279976]
4: 3434750 United_States
[pem_pos=0, final_scores=0.48147172,
final_local_scores=0.5153361, pem=-0.073457584,
similarity=19.867844]

5: 202354 Vietnam
[pem_pos=0, final_scores=0.506217,
final_local_scores=0.5626625, pem=-0.39026603,
similarity=23.003012]
6: 30128 Thailand
[pem_pos=0, final_scores=0.66246593,
final_local_scores=0.7222936, pem=-0.24029273,
similarity=26.951143]
7: 14579 Indonesia
[pem_pos=0, final_scores=0.5432629,
final_local_scores=0.5895965, pem=-0.23151328,
similarity=22.943455]
8: 3053483 AusAID
[pem_pos=0, final_scores=0.11242251,
final_local_scores=0.16532865, pem=0.0, similarity=9.02345]
1: 32073918 Brian_Tamberlin
[pem_pos=0, final_scores=0.28426373,
final_local_scores=0.26323456, pem=0.0, similarity=11.945157]
2: 4689264 Australia
[pem_pos=0, final_scores=0.56537926,
final_local_scores=0.5826751, pem=-0.17168477,
similarity=22.411549]
3: 32073918 Brian_Tamberlin
[pem_pos=0, final_scores=0.18398957,
final_local_scores=0.15370029, pem=0.0, similarity=8.676436]
4: 32073918 Brian_Tamberlin
[pem_pos=0, final_scores=0.081663,
final_local_scores=0.04192406, pem=0.0, similarity=5.3408103]
5: 32073918 Brian_Tamberlin
[pem_pos=0, final_scores=0.062528506,
final_local_scores=0.021022543, pem=0.0, similarity=4.7170677]
6: 32073918 Brian_Tamberlin
[pem_pos=0, final_scores=0.20075522,
final_local_scores=0.17201424, pem=0.0, similarity=9.22296]
7: 14579 Indonesia
[pem_pos=1, final_scores=0.34521306,
final_local_scores=0.3732571, pem=-0.74208045,
similarity=19.263998]
8: 32073918 Brian_Tamberlin
[pem_pos=0, final_scores=0.1618071,
final_local_scores=0.12946934, pem=0.0, similarity=7.9533362]
9: 1944 Australian_Capital_Territory
[pem_pos=0, final_scores=0.35492176,
final_local_scores=0.3808837, pem=0.0, similarity=15.456041]
10: 32073918 Brian_Tamberlin
[pem_pos=0, final_scores=0.25977623,
final_local_scores=0.23648569, pem=0.0, similarity=11.146917]
11: 4689264 Australia
[pem_pos=0, final_scores=0.3353936,
final_local_scores=0.33145064, pem=-0.17168477,
similarity=14.91451]
12: 32073918 Brian_Tamberlin

[pem_pos=0, final_scores=0.22908212,
final_local_scores=0.20295706, pem=0.0, similarity=10.146357]
13: 32073918 Brian_Tamberlin
[pem_pos=0, final_scores=0.25858325,
final_local_scores=0.23518252, pem=0.0, similarity=11.108028]
14: 1495 Australian_Labor_Party
[pem_pos=0, final_scores=0.27251226,
final_local_scores=0.32469857, pem=-0.2258703,
similarity=15.007684]

Sample: b'Harry_Hooper'

-LRB- footballer_born_1910 -RRB- Harry Reed Hooper -LRB-
16 December 1910 -- 24 March 1970 -RRB- was an [1 English 1]
professional 1 footballer who played as a full-back . He
started his career with [2 Nelson 2] before joining [1 1 Shef
field United 1] in 1930 . During his time with [3 Sheffield
United 3] , he captained the side in the 2 1936 [4 FA Cup Fin
al 4] . After 17 years with the club , he moved to before
retiring in 1949 . Between 1957 and 1962 , he was the manager
of [2 2 Halifax Town 2] . [5 Harry Hooper 5] was born in the
town of [6 3 Burnley 6] , [3 3 Lancashire 3] , on 16 December
1910 . Before becoming a professional footballer , he worked
in a tailor 's shop . Following his retirement from football
he lived in 4 Halifax , [7 West Yorkshire 7] , and died there
on 24 March 1970 , at the age of 59 . His son , also named [8
Harry 8] , was also a footballer and represented four
different league clubs . He was playing for Nelson Tradesmen
when he was spotted by [4 4 Football League Third Division No
rth 4] side [9 Nelson 9] , who signed him as an amateur in
October 1928 . He made his senior debut on 20 October in the
0 -- 4 defeat away at [5 5 Carlisle United 5] . [10 Hooper 10]
was awarded a professional contract the following month , but
found it difficult to break into the first-team . He returned
to the starting line-up in place of [6 6 Clem Rigg 6] for the
2 -- 3 loss to [7 7 Darlington 7] on 5 January 1929 and
retained his place for the next ten matches . [11 Hooper 11]
found his first-team opportunities even more limited in the
1929 -- 30 season following the signing of [8 8 Billy Fairhur
st 8] from [9 9 Southport 9] . After making eight further
league appearances for [12 Nelson 12] , [13 Hooper 13] was
signed by [10 10 First Division 10] side [11 11 Sheffield Uni
ted 11] . He moved along with half-back for a combined
transfer fee of # 750 . [14 Hooper 14] went on to play for [15
Sheffield United 15] for 17 years , scoring 10 goals in 269
league matches for the club . He captained the side as they
lost 0 -- 1 to [12 12 Arsenal 12] in the 5 1936 [16 FA Cup Fi
nal 16] . During the [17 Second World War 17] , he appeared
as a guest player for [13 13 Portsmouth 13] , and [14 14 Hudd
ersfield Town 14] . In July 1947 , [18 Hooper 18] signed for [
19 Hartlepoons United 19] on a permanent basis and went on to
make 66 league appearances for them , scoring 4 goals . [20
Hooper 20] retired from professional football in November

1949 in order to become a 6 trainer at [15 15 West Ham United 15] . In October 1957 , he was appointed 7 manager of [16 16 Halifax Town 16]

1: 314933 Sheffield_United_F.C.
2: 451200 Halifax_Town_A.F.C.
3: 50435 Lancashire
4: 580901 Football_League_Third_Division_North
5: 451157 Carlisle_United_F.C.
6: 24481319 Clem_Rigg
7: 451159 Darlington_F.C.
8: 30110933 Billy_Fairhurst
9: 1239131 Southport_F.C.
10: 28789929 Football_League_First_Division
11: 314933 Sheffield_United_F.C.
12: 2174 Arsenal_F.C.
13: 564301 Portsmouth_F.C.
14: 236211 Huddersfield_Town_F.C.
15: 46417 West_Ham_United_F.C.
16: 451200 Halifax_Town_A.F.C.

1: 10568 Association_football
[winner: 751310 Ashley_Cole, pem_pos=2,
final_scores=-1.2132818, final_local_scores=-1.3205357,
pem=-0.4855078, similarity=-32.67744]
gt: 10568 Association_football, pem_pos=1,
final_scores=-1.4607791, final_local_scores=-1.5910802,
pem=0.0, similarity=-43.391293]
2: 9165866 1936_FA_Cup_Final
[winner: 2706210 FA_Cup_Final, pem_pos=0,
final_scores=0.22365196, final_local_scores=0.18921024,
pem=-0.37451902, similarity=11.772817]
RECALL MISS: 9165866 1936_FA_Cup_Final]
[winner: 34673 1936, pem_pos=0, final_scores=-0.8407614,
final_local_scores=-0.6860294, pem=-0.8083407,
similarity=-11.986891]
RECALL MISS: 9165866 1936_FA_Cup_Final]
3: 311737 Burnley
[winner: 376725 Burnley_F.C., pem_pos=1,
final_scores=0.20974319, final_local_scores=0.12558314,
pem=-0.9914584, similarity=13.22907]
gt: 311737 Burnley, pem_pos=0,
final_scores=-0.006818235, final_local_scores=-0.029791817,
pem=-0.53772724, similarity=6.12491]
4: 18539471 Halifax,_West_Yorkshire
[winner: 149139 West_Yorkshire, pem_pos=0,
final_scores=0.34424573, final_local_scores=0.39083314,
pem=0.0, similarity=15.752954]
RECALL MISS: 18539471 Halifax,_West_Yorkshire]
[winner: 18539471 Halifax,_West_Yorkshire, pem_pos=2,
final_scores=-0.002374813, final_local_scores=-0.0049056113,
pem=-1.6685851, similarity=13.017344]
gt: 18539471 Halifax,_West_Yorkshire, pem_pos=2,
final_scores=-0.002374813, final_local_scores=-0.0049056113,

```

pem=-1.6685851, similarity=13.017344 ]
  [winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7943226, final_local_scores=-1.7684116,
pem=-4.6728287, similarity=-23.27165]
  RECALL MISS: 18539471 Halifax,_West_Yorkshire]
5: 9165866 1936_FA_Cup_Final
  [winner: 2706210 FA_Cup_Final, pem_pos=0,
final_scores=0.23640056, final_local_scores=0.20313615,
pem=-0.37451902, similarity=12.188393]
  RECALL MISS: 9165866 1936_FA_Cup_Final]
  [winner: 34673 1936, pem_pos=0, final_scores=-0.8326441,
final_local_scores=-0.67716247, pem=-0.8083407,
similarity=-11.722282]
  RECALL MISS: 9165866 1936_FA_Cup_Final]
6: 626967 Coach_(sport)
  [winner: 39083773 Horse_trainer, pem_pos=0,
final_scores=-1.0854841, final_local_scores=-1.1329565,
pem=-0.26937997, similarity=-28.255041]
  gt: 626967 Coach_(sport), pem_pos=14,
final_scores=-2.3186471, final_local_scores=-2.5269077,
pem=-6.016854, similarity=-38.597664 ]
7: 5468399 Manager_(association_football)
  [winner: 394333 Wadih_el-Hage, pem_pos=0,
final_scores=-1.1727417, final_local_scores=-1.1402483,
pem=0.0, similarity=-29.937572]
  gt: 5468399 Manager_(association_football), pem_pos=4,
final_scores=-1.4976263, final_local_scores=-1.6400702,
pem=-2.228848, similarity=-32.73243 ]
1: 314933 Sheffield_United_F.C.
  [pem_pos=0, final_scores=0.316454,
final_local_scores=0.23892042, pem=0.0, similarity=11.219575]
2: 451200 Halifax_Town_A.F.C.
  [pem_pos=0, final_scores=0.4430719,
final_local_scores=0.3976776, pem=-0.07232743,
similarity=16.350534]
3: 50435 Lancashire
  [pem_pos=0, final_scores=0.11688475,
final_local_scores=0.11554217, pem=-0.20611861,
similarity=8.658627]
4: 580901 Football_League_Third_Division_North
  [pem_pos=0, final_scores=0.33327883,
final_local_scores=0.28770566, pem=0.0, similarity=12.675424]
5: 451157 Carlisle_United_F.C.
  [pem_pos=0, final_scores=0.49927133,
final_local_scores=0.44936067, pem=0.0, similarity=17.49953]
6: 24481319 Clem_Rigg
  [pem_pos=0, final_scores=0.47766805,
final_local_scores=0.539593, pem=0.0, similarity=20.192244]
7: 451159 Darlington_F.C.
  [pem_pos=1, final_scores=0.5829713,
final_local_scores=0.53952986, pem=-1.8695662,
similarity=30.357351]

```

8: 30110933 Billy_Fairhurst
[pem_pos=0, final_scores=0.28944403,
final_local_scores=0.34565014, pem=0.0, similarity=14.404602]
9: 1239131 Southport_F.C.
[pem_pos=1, final_scores=0.66914684,
final_local_scores=0.6358137, pem=-1.7938718,
similarity=32.819016]
10: 28789929 Football_League_First_Division
[pem_pos=0, final_scores=0.18996175,
final_local_scores=0.12697977, pem=-0.5147791,
similarity=10.678492]
11: 314933 Sheffield_United_F.C.
[pem_pos=0, final_scores=0.29934675,
final_local_scores=0.22023338, pem=0.0, similarity=10.661917]
12: 2174 Arsenal_F.C.
[pem_pos=0, final_scores=0.57402045,
final_local_scores=0.54040605, pem=-0.15820743,
similarity=21.076864]
13: 564301 Portsmouth_F.C.
[pem_pos=1, final_scores=0.6319551,
final_local_scores=0.5891847, pem=-1.2302613,
similarity=28.362513]
14: 236211 Huddersfield_Town_F.C.
[pem_pos=0, final_scores=0.4335428,
final_local_scores=0.37013113, pem=0.0, similarity=15.135164]
15: 46417 West_Ham_United_F.C.
[pem_pos=0, final_scores=0.5005534,
final_local_scores=0.45130378, pem=0.0, similarity=17.557518]
16: 451200 Halifax_Town_A.F.C.
[pem_pos=0, final_scores=0.37258178,
final_local_scores=0.3203179, pem=-0.07232743,
similarity=14.0419655]
1: 9316 England
[pem_pos=1, final_scores=0.57049805,
final_local_scores=0.62974054, pem=-1.3527626,
similarity=30.238958]
2: 2205875 Nelson_F.C.
[pem_pos=2, final_scores=0.33896673,
final_local_scores=0.29072577, pem=-2.787361,
similarity=27.923653]
3: 314933 Sheffield_United_F.C.
[pem_pos=0, final_scores=0.3433181,
final_local_scores=0.26826543, pem=0.0, similarity=12.095288]
4: 2706210 FA_Cup_Final
[pem_pos=0, final_scores=0.22365196,
final_local_scores=0.18921024, pem=-0.37451902,
similarity=11.772817]
5: 8682107 Harry_Hooper_(footballer,_born_1933)
[pem_pos=1, final_scores=0.39070296,
final_local_scores=0.38670653, pem=-0.93501675,
similarity=20.714575]
6: 376725 Burnley_F.C.

[pem_pos=1, final_scores=0.20974319,
final_local_scores=0.12558314, pem=-0.9914584,
similarity=13.22907]
7: 149139 West_Yorkshire
[pem_pos=0, final_scores=0.34424573,
final_local_scores=0.39083314, pem=0.0, similarity=15.752954]
8: 8682107 Harry_Hooper_(footballer,_born_1933)
[pem_pos=2, final_scores=0.31763744,
final_local_scores=0.30618805, pem=-0.93501675,
similarity=18.311743]
9: 2205875 Nelson_F.C.
[pem_pos=2, final_scores=0.2405409,
final_local_scores=0.18321055, pem=-2.787361,
similarity=24.715183]
10: 8682107 Harry_Hooper_(footballer,_born_1933)
[pem_pos=1, final_scores=0.50260955,
final_local_scores=0.50818646, pem=-0.93501675,
similarity=24.339777]
11: 8682107 Harry_Hooper_(footballer,_born_1933)
[pem_pos=1, final_scores=0.57367176,
final_local_scores=0.5858111, pem=-0.93501675,
similarity=26.656252]
12: 2205875 Nelson_F.C.
[pem_pos=2, final_scores=0.33049577,
final_local_scores=0.28147256, pem=-2.787361,
similarity=27.647518]
13: 8682107 Harry_Hooper_(footballer,_born_1933)
[pem_pos=1, final_scores=0.6164784,
final_local_scores=0.6325708, pem=-0.93501675,
similarity=28.051655]
14: 8682107 Harry_Hooper_(footballer,_born_1933)
[pem_pos=1, final_scores=0.4169547,
final_local_scores=0.41462147, pem=-0.93501675,
similarity=21.547611]
15: 314933 Sheffield_United_F.C.
[pem_pos=0, final_scores=0.31715393,
final_local_scores=0.239685, pem=0.0, similarity=11.242392]
16: 2706210 FA_Cup_Final
[pem_pos=0, final_scores=0.23640056,
final_local_scores=0.20313615, pem=-0.37451902,
similarity=12.188393]
17: 32927 World_War_II
[pem_pos=0, final_scores=0.18280147,
final_local_scores=0.2814057, pem=0.0, similarity=12.48742]
18: 8682107 Harry_Hooper_(footballer,_born_1933)
[pem_pos=1, final_scores=0.47466248,
final_local_scores=0.47765845, pem=-0.93501675,
similarity=23.42876]
19: 451130 Hartlepool_United_F.C.
[pem_pos=0, final_scores=0.4114501,
final_local_scores=0.35190028, pem=0.0, similarity=14.591119]
20: 8682107 Harry_Hooper_(footballer,_born_1933)

[pem_pos=1, final_scores=0.46929324,
final_local_scores=0.47215718, pem=-0.93501675,
similarity=23.264591]

Sample: b'Jiri_Tranovsky'

[1 Jiří Třanovský 1] -LRB- , , -RRB- -LRB- 9 April 1592 , [1
1 Teschen 1] , [2 2 Silesia 2] -- 29 May 1637 , 1 Liptovský
Sv . Mikuláš , [3 3 Upper Hungary 3] -RRB- , was a 2 hymn
writer from the [4 4 Cieszyn Silesia 4] . He was sometimes
called the father of [2 Slovak 2] hymnody and the '' [5 5 Lut
her 5] of the [3 Slavs 3] . '' His name is sometimes to
George . Třanovský was born in [6 6 Teschen 6] , and studied
at [7 7 Guben 7] and . In 1607 , he was admitted to the where
[8 8 Martin Luther 8] had taught less than a century
earlier . He traveled in [9 9 Bohemia 9] and [10 10 Silesia 10
] in 1612 and became a teacher at St. Nicholas Gymnasium in [1
1 11 Prague 11] . Later , he became 3 rector of a school in [1
2 12 Holešov 12] , [13 13 Moravia 13] . In 1616 he was
ordained a priest in 4 Meziříčí and served as a pastor for
four years . The persecution of in [4 Bohemia 4] under 5 Ferd
inand II forced him into exile . After an imprisonment in
1623 and the death of two children from 6 plague the
following year , Třanovský received a call to be pastor to a
church in [14 14 Bielitz 14] , [15 15 Teschen Silesia 15] .
He also became personal chaplain to Count Kasper Illehazy in
1627 . Třanovský was a lover of poetry and hymns . He issued
several collections of hymns , the first being the 7 Latin
Odorum Sacrarum sive Hymnorum Libri III in 1629 , but his
most important and most famous work was Cithara Sanctorum -
LRB- Lyre of the Saints -RRB- , written in [5 Czech 5] ,
which appeared in 1636 in [16 16 Levoča 16] . This latter
volume has formed the basis of [6 Czech 6] and [7 Slovak 7]
Lutheran hymnody to the present day . In addition to hymn
collections , Třanovský translated the [17 17 Augsburg Confes
sion 17] in 1620 into [8 Czech 8] . These two latter works
together with [9 8 Bible 9] of Kralice are the pillars that
supported the Slovak Reformation . From 1631 until 1637 ,
Třanovský was pastor at a church in [18 18 Liptovský Svätý Mi
kuláš 18] in present-day [19 19 Slovakia 19] . He died on 29
May of that year and was buried in an unmarked grave at his
church . He was forty-six years old . Třanovský is
commemorated on 29 May in the [20 20 Evangelical Lutheran Chu
rch in America 20] and the 9 Evangelical Lutheran Church in Ca
nada

- 1: 353922 Cieszyn
- 2: 28399 Silesia
- 3: 2130666 Upper_Hungary
- 4: 789474 Cieszyn_Silesia
- 5: 7567080 Martin_Luther
- 6: 353922 Cieszyn
- 7: 703939 Guben
- 8: 7567080 Martin_Luther

9: 4345 Bohemia
 10: 28399 Silesia
 11: 23844 Prague
 12: 5363219 Holešov
 13: 20216 Moravia
 14: 74603 Bielsko-Biała
 15: 789474 Cieszyn_Silesia
 16: 609312 Levoča
 17: 23883384 Augsburg_Confession
 18: 1060406 Liptovský_Mikuláš
 19: 26830 Slovakia
 20: 248255 Evangelical_Lutheran_Church_in_America
 1: 1060406 Liptovský_Mikuláš
 [winner: 155823 Sievert, pem_pos=1,
 final_scores=-1.174267, final_local_scores=-1.1436448,
 pem=-0.85181254, similarity=-25.40664|
 RECALL MISS: 1060406 Liptovský_Mikuláš]
 [winner: 4721653 Zyphe_language, pem_pos=12,
 final_scores=-1.6299142, final_local_scores=-1.5995896,
 pem=-4.4188404, similarity=-19.614893|
 RECALL MISS: 1060406 Liptovský_Mikuláš]
 2: 13756 Hymn
 [winner: 1571479 Adam-on-di-Ahman_(hymn), pem_pos=0,
 final_scores=-0.740315, final_local_scores=-0.6886494,
 pem=0.0, similarity=-16.460958|
 RECALL MISS: 13756 Hymn]
 3: 66772 Rector_(academia)
 [winner: 36115576 Rector_(ecclesiastical), pem_pos=0,
 final_scores=-1.2160928, final_local_scores=-1.2583501,
 pem=-0.7380934, similarity=-29.448097|
 gt: 66772 Rector_(academia), pem_pos=1,
 final_scores=-1.4422053, final_local_scores=-1.5773396,
 pem=-1.7644953, similarity=-33.385647]
 4: 4445175 Valašské_Meziříčí
 [winner: 24076116 Meziříčí_(Tábor_District), pem_pos=0,
 final_scores=-0.041549847, final_local_scores=0.08171326,
 pem=-0.06280091, similarity=6.869721|
 gt: 4445175 Valašské_Meziříčí, pem_pos=1,
 final_scores=-0.18902837, final_local_scores=-0.18159983,
 pem=-2.799022, similarity=13.891928]
 5: 151083 Ferdinand_II,_Holy_Roman_Emperor
 [winner: 151083 Ferdinand_II,_Holy_Roman_Emperor,
 pem_pos=0, final_scores=0.029013678,
 final_local_scores=0.017415643, pem=-0.632278,
 similarity=8.047856|
 gt: 151083 Ferdinand_II,_Holy_Roman_Emperor, pem_pos=0,
 final_scores=0.029013678, final_local_scores=0.017415643,
 pem=-0.632278, similarity=8.047856]
 6: 16392927 Bubonic_plague
 [winner: 4746 Plague_(disease), pem_pos=0,
 final_scores=-1.0145398, final_local_scores=-1.0831265,
 pem=-0.75807434, similarity=-24.110416|

gt: 16392927 Bubonic_plague, pem_pos=1,
final_scores=-1.1312213, final_local_scores=-1.1945949,
pem=-1.2438496, similarity=-24.795132]
7: 17730 Latin
[winner: 17730 Latin, pem_pos=0, final_scores=-0.3367108,
final_local_scores=-0.36285308, pem=-0.08185364,
similarity=-6.2934136]
gt: 17730 Latin, pem_pos=0, final_scores=-0.3367108,
final_local_scores=-0.36285308, pem=-0.08185364,
similarity=-6.2934136]
8: 2342832 Bible_of_Kralice
[winner: 3390 Bible, pem_pos=0, final_scores=0.51410604,
final_local_scores=0.5688234, pem=0.0, similarity=21.064537]
RECALL MISS: 2342832 Bible_of_Kralice]
[winner: 24129151 Kralice_nad_Āslavou, pem_pos=0,
final_scores=-0.15723081, final_local_scores=-0.066536635,
pem=-0.6931472, similarity=5.873569]
RECALL MISS: 2342832 Bible_of_Kralice]
9: 848904 Evangelical_Lutheran_Church_in_Canada
[winner: 848904 Evangelical_Lutheran_Church_in_Canada,
pem_pos=0, final_scores=-0.037301287,
final_local_scores=0.0066517293, pem=0.0,
similarity=4.2882137]
gt: 848904 Evangelical_Lutheran_Church_in_Canada,
pem_pos=0, final_scores=-0.037301287,
final_local_scores=0.0066517293, pem=0.0,
similarity=4.2882137]
1: 353922 Cieszyn
[pem_pos=1, final_scores=0.115305886,
final_local_scores=0.071800336, pem=-0.6864312,
similarity=9.965298]
2: 28399 Silesia
[pem_pos=0, final_scores=0.24787395,
final_local_scores=0.18229267, pem=-0.040829465,
similarity=9.751727]
3: 2130666 Upper_Hungary
[pem_pos=0, final_scores=0.088599995,
final_local_scores=0.07243386, pem=0.0, similarity=6.2512836]
4: 789474 Cieszyn_Silesia
[pem_pos=0, final_scores=0.41559654,
final_local_scores=0.39456755, pem=-0.009901085,
similarity=15.91824]
5: 7567080 Martin_Luther
[pem_pos=0, final_scores=0.11485247,
final_local_scores=0.10879432, pem=0.0, similarity=7.3363523]
6: 353922 Cieszyn
[pem_pos=1, final_scores=0.24832426,
final_local_scores=0.2200042, pem=-0.6864312,
similarity=14.387997]
7: 703939 Guben
[pem_pos=0, final_scores=0.33305377,
final_local_scores=0.35850233, pem=0.0, similarity=14.788137]

8: 7567080 Martin_Luther
[pem_pos=0, final_scores=0.45107037,
final_local_scores=0.48321766, pem=0.0, similarity=18.509892]
9: 4345 Bohemia
[pem_pos=0, final_scores=0.36673003,
final_local_scores=0.30635834, pem=-0.08606065,
similarity=13.7000675]
10: 28399 Silesia
[pem_pos=0, final_scores=0.3804707,
final_local_scores=0.32713446, pem=-0.040829465,
similarity=14.074095]
11: 23844 Prague
[pem_pos=0, final_scores=0.506401,
final_local_scores=0.49584246, pem=-0.05392049,
similarity=19.179869]
12: 5363219 Holešov
[pem_pos=0, final_scores=0.36116272,
final_local_scores=0.41662955, pem=-0.017178318,
similarity=16.616188]
13: 20216 Moravia
[pem_pos=0, final_scores=0.35502744,
final_local_scores=0.29109254, pem=-0.003610146,
similarity=12.796127]
14: 74603 Bielsko-Biała
[pem_pos=0, final_scores=0.6018866,
final_local_scores=0.62955916, pem=0.0, similarity=22.877014]
15: 789474 Cieszyn_Silesia
[pem_pos=0, final_scores=0.28191584,
final_local_scores=0.24854177, pem=0.0, similarity=11.506695]
16: 609312 Levoča
[pem_pos=0, final_scores=0.26076025,
final_local_scores=0.2566353, pem=-0.040454958,
similarity=11.968223]
17: 23883384 Augsburg_Confession
[pem_pos=0, final_scores=0.16069798,
final_local_scores=0.1886518, pem=0.0, similarity=9.719459]
18: 1060406 Liptovský_Mikuláš
[pem_pos=0, final_scores=0.39629823,
final_local_scores=0.41117322, pem=0.0, similarity=16.359941]
19: 26830 Slovakia
[pem_pos=0, final_scores=0.5467618,
final_local_scores=0.5479031, pem=-0.36602598,
similarity=22.430742]
20: 248255 Evangelical_Lutheran_Church_in_America
[pem_pos=0, final_scores=0.1779183,
final_local_scores=0.20794344, pem=0.0, similarity=10.29516]
1: 5549725 Jiří_Třanovský
[pem_pos=0, final_scores=0.068403885,
final_local_scores=0.10467759, pem=0.0, similarity=7.213501]
2: 26830 Slovakia
[pem_pos=1, final_scores=0.56201833,
final_local_scores=0.5645686, pem=-1.0527612,

similarity=26.662643]
 3: 29440 Slavs
 [pem_pos=0, final_scores=0.23195802,
 final_local_scores=0.22203988, pem=0.0, similarity=10.715826]
 4: 4345 Bohemia
 [pem_pos=0, final_scores=0.444687,
 final_local_scores=0.391132, pem=-0.08606065,
 similarity=16.229883]
 5: 5321 Czech_Republic
 [pem_pos=1, final_scores=0.17271234,
 final_local_scores=0.13376987, pem=-1.208016,
 similarity=14.651052]
 6: 5321 Czech_Republic
 [pem_pos=1, final_scores=0.566729,
 final_local_scores=0.5623842, pem=-1.208016,
 similarity=27.441757]
 7: 26830 Slovakia
 [pem_pos=1, final_scores=0.2661097,
 final_local_scores=0.2413334, pem=-1.0527612,
 similarity=17.016663]
 8: 5321 Czech_Republic
 [pem_pos=1, final_scores=0.31877977,
 final_local_scores=0.29153734, pem=-1.208016,
 similarity=19.359148]
 9: 3390 Bible
 [pem_pos=0, final_scores=0.51410604,
 final_local_scores=0.5688234, pem=0.0, similarity=21.064537]

Sample: b'Manfred_Seissler'

Manfred `` Manny `` Seissler -LRB- [1 German 1] : Manfred
 Seißler -RRB- -LRB- born August 8 , 1939 in [2 Germany 2] -
 RRB- is a former U.S.-German 1 player who began his career in
 the lower [3 German 3] divisions before moving to the [4
 United States 4] . He played one season in the [5 2 National P
 rofessional Soccer League 5] and seven in the . He also
 earned one with the [6 3 U.S. 6] national team in 1973 . [7
 Seissler 7] began his career in 1960 with the [1 1 KSV Hessen
 Kassel 1] junior team . In 1961 , he moved to the first team
 which competed in the , a regional [8 German 8] league which
 preceded the creation of the . In 1963 , he moved to -LRB-
 Second Division -RRB- club 4 1 . [9 FC Pforzheim 9] . In
 1965 , he moved to [2 2 SV Eintracht Trier 05 2] of the . In
 1967 , he left [10 Germany 10] for the [11 United States 11]
 where he signed with the 5 Pittsburgh Phantoms of the [12 6 Na
 tional Professional Soccer League 12] . The NPSL merged with
 the [3 3 United Soccer Association 3] in 1968 to form the -
 LRB- NASL -RRB- . The Phantoms did not enter the new league
 and [13 Seissler 13] signed with the [4 4 Kansas City Spurs 4]
 , playing with the [14 Spurs 14] until they folded following
 the 1970 season . He then transferred to the [5 5 Rochester La
 ncers 5] . In 1973 , the Lancers traded [15 Seissler 15] to
 the [6 6 Montreal Olympique 6] but it folded at the end of

the season and [16 Seissler 16] returned to the Lancers for the 1974 season . He played only one game , then retired . During his early years in the [17 NASL 17] , [18 Seissler 18] was one of the league 's top forwards , being named a first team All Star in 1969 , 1970 and 1971 . [19 Seissler 19] earned one with the [20 7 U.S. 20] national team in a 4-0 loss to [7 7 Poland 7] on March 20 , 1973 . He started the game , but came off for [8 8 Siegfried Stritzl 8]

1: 3949816 KSV_Hessen_Kassel

2: 2735454 SV_Eintracht_Trier_05

3: 1727696 United_Soccer_Association

4: 1899371 Kansas_City_Spurs

5: 1954095 Rochester_Lancers

6: 1955743 Montreal_Olympique

7: 771860 Poland_national_football_team

8: 12025156 Siegfried_Stritzl

1: 607501 Player_(game)

[winner: 153914 Cassette_deck, pem_pos=8,
final_scores=-1.8831695, final_local_scores=-1.9066417,
pem=-2.561675, similarity=-38.87749|

gt: 607501 Player_(game), pem_pos=0,
final_scores=-2.0295594, final_local_scores=-2.1865644,
pem=-0.6836065, similarity=-57.444168]

2: 11238708 National_Professional_Soccer_League_(1967)

[winner: 578812 National_Professional_Soccer_League_(1984
-2001), pem_pos=0, final_scores=0.168626,
final_local_scores=0.18201041, pem=0.0, similarity=9.521267|

gt: 11238708 National_Professional_Soccer_League_(1967),
pem_pos=1, final_scores=0.117844656,
final_local_scores=0.103238076, pem=-0.5409579,
similarity=10.112356]

3: 582488 United_States_men's_national_soccer_team

[winner: 3434750 United_States, pem_pos=0,
final_scores=0.4298094, final_local_scores=0.46265548,
pem=-0.14075686, similarity=18.661732|

gt: 582488 United_States_men's_national_soccer_team,
pem_pos=12, final_scores=-0.34261352,
final_local_scores=-0.41988903, pem=-4.0913,
similarity=13.808512]

[winner: 140801 Team, pem_pos=0, final_scores=-1.4889678,
final_local_scores=-1.5933838, pem=-0.5173818,
similarity=-40.64643|

RECALL MISS: 582488

United_States_men's_national_soccer_team]

[winner: 302395 National_(brand), pem_pos=3,
final_scores=-1.5275557, final_local_scores=-1.5408621,
pem=-2.9526155, similarity=-25.8359|

RECALL MISS: 582488

United_States_men's_national_soccer_team]

4: 7116239 1._FC_Pforzheim

[winner: 7116239 1._FC_Pforzheim, pem_pos=0,
final_scores=0.12870227, final_local_scores=0.15961677,

```

pem=0.0, similarity=8.852996|
  gt: 7116239 1._FC_Pforzheim, pem_pos=0,
final_scores=0.12870227, final_local_scores=0.15961677,
pem=0.0, similarity=8.852996 ]
  [winner: 1175 April_1, pem_pos=4,
final_scores=-1.0891451, final_local_scores=-1.0204325,
pem=-3.8010993, similarity=-5.6910577|
  RECALL MISS: 7116239 1._FC_Pforzheim]
  [winner: 4721653 Zyphe_language, pem_pos=12,
final_scores=-1.6711869, final_local_scores=-1.6051269,
pem=-4.4188404, similarity=-19.78013|
  RECALL MISS: 7116239 1._FC_Pforzheim]
5: 6350141 Pittsburgh_Phantoms_(NPSL)
  [winner: 17360124 Pittsburgh_Phantoms_(RHI), pem_pos=0,
final_scores=-0.029537395, final_local_scores=0.008902118,
pem=-0.1022245, similarity=4.9112825|
  gt: 6350141 Pittsburgh_Phantoms_(NPSL), pem_pos=2,
final_scores=-0.22459105, final_local_scores=-0.2182863,
pem=-1.4710596, similarity=5.5754647 ]
6: 11238708 National_Professional_Soccer_League_(1967)
  [winner: 578812 National_Professional_Soccer_League_(1984
-2001), pem_pos=0, final_scores=0.1292096,
final_local_scores=0.13895398, pem=0.0, similarity=8.236377|
  gt: 11238708 National_Professional_Soccer_League_(1967),
pem_pos=1, final_scores=0.080187395,
final_local_scores=0.06210324, pem=-0.5409579,
similarity=8.88481 ]
7: 582488 United_States_men's_national_soccer_team
  [winner: 3434750 United_States, pem_pos=0,
final_scores=0.3375042, final_local_scores=0.3618261,
pem=-0.14075686, similarity=15.652782|
  gt: 582488 United_States_men's_national_soccer_team,
pem_pos=12, final_scores=-0.24035518,
final_local_scores=-0.30818737, pem=-4.0913,
similarity=17.141912 ]
  [winner: 302395 National_(brand), pem_pos=3,
final_scores=-1.6174914, final_local_scores=-1.6391032,
pem=-2.9526155, similarity=-28.767614|
  RECALL MISS: 582488
United_States_men's_national_soccer_team]
  [winner: 140801 Team, pem_pos=0, final_scores=-1.6271096,
final_local_scores=-1.7442827, pem=-0.5173818,
similarity=-45.14956|
  RECALL MISS: 582488
United_States_men's_national_soccer_team]
1: 3949816 KSV_Hessen_Kassel
  [pem_pos=0, final_scores=0.30702037,
final_local_scores=0.3411375, pem=0.0, similarity=14.269937]
2: 2735454 SV_Eintracht_Trier_05
  [pem_pos=0, final_scores=0.24743311,
final_local_scores=0.26534975, pem=0.0, similarity=12.008279]
3: 1727696 United_Soccer_Association

```

[pem_pos=0, final_scores=0.15457986,
final_local_scores=0.16479042, pem=0.0, similarity=9.007388]
4: 1899371 Kansas_City_Spurs
[pem_pos=0, final_scores=0.12033068,
final_local_scores=0.13162464, pem=0.0, similarity=8.017654]
5: 1954095 Rochester_Lancers
[pem_pos=0, final_scores=0.13566037,
final_local_scores=0.13115674, pem=0.0, similarity=8.003692]
6: 1955743 Montreal_Olympique
[pem_pos=0, final_scores=0.04188843,
final_local_scores=0.06819381, pem=0.0, similarity=6.124752]
7: 771860 Poland_national_football_team
[pem_pos=1, final_scores=0.269795,
final_local_scores=0.2774542, pem=-3.319454,
similarity=30.421206]
8: 12025156 Siegfried_Stritzl
[pem_pos=0, final_scores=0.34605628,
final_local_scores=0.42342776, pem=0.0, similarity=16.725641]
1: 11867 Germany
[pem_pos=1, final_scores=0.4388693,
final_local_scores=0.4608255, pem=-1.0465823,
similarity=23.533144]
2: 11867 Germany
[pem_pos=1, final_scores=0.67846364,
final_local_scores=0.7225459, pem=-0.15529114,
similarity=26.496422]
3: 11867 Germany
[pem_pos=1, final_scores=0.62157995,
final_local_scores=0.66040903, pem=-1.0465823,
similarity=29.489113]
4: 3434750 United_States
[pem_pos=0, final_scores=0.39622307,
final_local_scores=0.42596757, pem=-0.073457584,
similarity=17.200909]
5: 578812 National_Professional_Soccer_League_(1984-2001)
[pem_pos=0, final_scores=0.168626,
final_local_scores=0.18201041, pem=0.0, similarity=9.521267]
6: 3434750 United_States
[pem_pos=0, final_scores=0.4298094,
final_local_scores=0.46265548, pem=-0.14075686,
similarity=18.661732]
7: 14731830 Manfred_Seissler
[pem_pos=0, final_scores=0.2471989,
final_local_scores=0.24616551, pem=0.0, similarity=11.435782]
8: 11867 Germany
[pem_pos=1, final_scores=0.6380146,
final_local_scores=0.6783615, pem=-1.0465823,
similarity=30.02485]
9: 7116239 1._FC_Pforzheim
[pem_pos=0, final_scores=0.12870227,
final_local_scores=0.15961677, pem=0.0, similarity=8.852996]
10: 11867 Germany

```

    [pem_pos=0, final_scores=0.5709196,
final_local_scores=0.60507035, pem=-0.15529114,
similarity=22.990717]
11: 3434750 United_States
    [pem_pos=0, final_scores=0.4076845,
final_local_scores=0.43848747, pem=-0.073457584,
similarity=17.574528]
12: 578812 National_Professional_Soccer_League_(1984-2001)
    [pem_pos=0, final_scores=0.1292096,
final_local_scores=0.13895398, pem=0.0, similarity=8.236377]
13: 14731830 Manfred_Seissler
    [pem_pos=0, final_scores=0.25235325,
final_local_scores=0.2517959, pem=0.0, similarity=11.603805]
14: 68198 Tottenham_Hotspur_F.C.
    [pem_pos=0, final_scores=0.42021757,
final_local_scores=0.44141513, pem=-0.24605535,
similarity=18.600508]
15: 14731830 Manfred_Seissler
    [pem_pos=0, final_scores=0.23207323,
final_local_scores=0.22964302, pem=0.0, similarity=10.9427185]
16: 14731830 Manfred_Seissler
    [pem_pos=0, final_scores=0.21575837,
final_local_scores=0.21182153, pem=0.0, similarity=10.41089]
17: 34063194 North_American_Soccer_League
    [pem_pos=0, final_scores=0.21862619,
final_local_scores=0.18934438, pem=0.0, similarity=9.740127]
18: 14731830 Manfred_Seissler
    [pem_pos=0, final_scores=0.16809915,
final_local_scores=0.15976104, pem=0.0, similarity=8.857301]
19: 14731830 Manfred_Seissler
    [pem_pos=0, final_scores=0.121842906,
final_local_scores=0.109233156, pem=0.0, similarity=7.349448]
20: 3434750 United_States
    [pem_pos=0, final_scores=0.3375042,
final_local_scores=0.3618261, pem=-0.14075686,
similarity=15.652782]

```

Sample: b'Marillier_shot'

The [1 Marillier 1] shot is a modern shot in 1 cricket which involves using the bat as a ramp to flick a ball backwards over the 's shoulder for a 2 boundary . It is a rare , risky and unorthodox shot but when successfully used can be devastating . The shot was pioneered by [2 Zimbabwean 2] batsman [1 1 Dougie Marillier 1] , after whom the shot is named . In a triangular tournament in [2 2 Australia 2] with , and , [3 Zimbabwe 3] played their final match with [4 Australia 4] and [5 Marillier 5] got a chance in the team . He could hardly have had a more testing experience , as a fine [6 Zimbabwe 6] batting performance after [7 Australia 7] scored over 300 meant that he came in at number seven needing to score 15 to win the match in the final over , which was to be bowled by [3 3 Glenn McGrath 3] . [8 Marillier 8] moved

across to the first and third balls he received from [9 McGrath 9] and flicked them over his shoulder to 3 fine leg for boundaries , reviving hopes of an incredible [10 Zimbabwe 10] victory . But he was just unable to complete the job , and his team lost by one run . His two courageous and unorthodox boundary strokes made him famous , with the shot becoming known as the [11 Marillier 11] shot . [12 Marillier 12] continued to do reasonably well for the national side . In 2002 he `` Marilliered '' [13 Zimbabwe 13] to a famous win in [14 4 India 14] in an [4 4 One Day International 4] at [15 Faridabad 15] , [16 India 16] with 56 5 not out at the death , although this time he used the shot against [5 5 Zaheer Khan 5] . Other notable exponents of the [17 Marillier 17] shot include the 6 wicket-keeper [6 6 Brendon McCullum 6] , [18 7 Western Australia 18] 's former wicket-keeper [19 Ryan Campbell 19] , the batsman [7 7 Robin Uthappa 7] and the batsman [8 8 Tillakaratne Dilshan 8] . [20 Dilshan 20] 's success with a similar shot in the 8 2009 [21 ICC World Twenty20 21] tournament has led [22 Dilshan 22] 's shot being titled the '' [9 9 Dilscoop 9]

1: 2873728 Dougie_Marillier

2: 4689264 Australia

3: 423488 Glenn_McGrath

4: 381447 One_Day_International

5: 679977 Zaheer_Khan

6: 2283247 Brendon_McCullum

7: 4635830 Robin_Uthappa

8: 894383 Tillakaratne_Dilshan

9: 23329245 Dilscoop

1: 25675557 Cricket

[winner: 330091 Rhabdophoridae, pem_pos=10, final_scores=-1.37151, final_local_scores=-1.3352191, pem=-2.56456, similarity=-21.809416]

gt: 25675557 Cricket, pem_pos=0, final_scores=-1.5151813, final_local_scores=-1.6886971, pem=-0.07532314, similarity=-45.894756]

2: 711720 Boundary_(cricket)

[winner: 7275586 Boundary_marker, pem_pos=0, final_scores=-1.2088434, final_local_scores=-1.1723468, pem=0.0, similarity=-30.895452]

gt: 711720 Boundary_(cricket), pem_pos=2, final_scores=-1.7803228, final_local_scores=-1.9858818, pem=-1.365021, similarity=-47.749763]

3: 65951 Fielding_(cricket)

[winner: 65951 Fielding_(cricket), pem_pos=0, final_scores=-1.0059144, final_local_scores=-1.1310389, pem=0.0, similarity=-29.662743]

gt: 65951 Fielding_(cricket), pem_pos=0, final_scores=-1.0059144, final_local_scores=-1.1310389, pem=0.0, similarity=-29.662743]

4: 14533 India

[winner: 407754 India_national_cricket_team, pem_pos=1,

```

final_scores=0.1903253, final_local_scores=0.11579126,
pem=-3.5152702, similarity=26.661741|
    gt: 14533 India, pem_pos=0, final_scores=0.12969671,
final_local_scores=0.20386168, pem=-0.13622136,
similarity=10.914145 ]
5: 2430357 Not_out
    [winner: 2430357 Not_out, pem_pos=0,
final_scores=-1.0907052, final_local_scores=-1.2500293,
pem=-0.008413082, similarity=-33.167904|
    gt: 2430357 Not_out, pem_pos=0, final_scores=-1.0907052,
final_local_scores=-1.2500293, pem=-0.008413082,
similarity=-33.167904 ]
6: 463313 Wicket-keeper
    [winner: 463313 Wicket-keeper, pem_pos=0,
final_scores=-1.2860174, final_local_scores=-1.4490018,
pem=-0.004312502, similarity=-39.12794|
    gt: 463313 Wicket-keeper, pem_pos=0,
final_scores=-1.2860174, final_local_scores=-1.4490018,
pem=-0.004312502, similarity=-39.12794 ]
7: 1327443 Western_Australia_cricket_team
    [winner: 1327443 Western_Australia_cricket_team,
pem_pos=1, final_scores=0.26960373,
final_local_scores=0.25094354, pem=-2.1796863,
similarity=23.43184|
    gt: 1327443 Western_Australia_cricket_team, pem_pos=1,
final_scores=0.26960373, final_local_scores=0.25094354,
pem=-2.1796863, similarity=23.43184 ]
    [winner: 1415027 David_Berman_(musician), pem_pos=5,
final_scores=-2.0874698, final_local_scores=-2.12039,
pem=-3.0012724, similarity=-42.865566|
    RECALL MISS: 1327443 Western_Australia_cricket_team]
8: 8258541 2009_ICC_World_Twenty20
    [winner: 8258172 ICC_World_Twenty20, pem_pos=0,
final_scores=0.32542235, final_local_scores=0.3139296,
pem=-0.06018666, similarity=13.785302|
    gt: 8258541 2009_ICC_World_Twenty20, pem_pos=2,
final_scores=0.23167117, final_local_scores=0.19704744,
pem=-1.0498221, similarity=15.679098 ]
    [winner: 35983 2009, pem_pos=0, final_scores=-1.044258,
final_local_scores=-0.9766661, pem=-0.9296576,
similarity=-20.000328|
    RECALL MISS: 8258541 2009_ICC_World_Twenty20]
1: 2873728 Dougie_Marillier
    [pem_pos=0, final_scores=0.696961,
final_local_scores=0.69359577, pem=0.0, similarity=24.787994]
2: 4689264 Australia
    [pem_pos=0, final_scores=0.21454541,
final_local_scores=0.29482463, pem=-0.17168477,
similarity=13.821516]
3: 423488 Glenn_McGrath
    [pem_pos=0, final_scores=0.8478945,
final_local_scores=0.8543746, pem=0.0, similarity=29.585955]

```

4: 381447 One_Day_International
[pem_pos=1, final_scores=0.3165155,
final_local_scores=0.26015016, pem=-0.022472855,
similarity=11.975324]
5: 679977 Zaheer_Khan
[pem_pos=0, final_scores=0.61528546,
final_local_scores=0.6157603, pem=0.0, similarity=22.465229]
6: 2283247 Brendon_McCullum
[pem_pos=0, final_scores=0.7963727,
final_local_scores=0.8034184, pem=0.0, similarity=28.065321]
7: 4635830 Robin_Uthappa
[pem_pos=0, final_scores=0.8110649,
final_local_scores=0.8659846, pem=0.0, similarity=29.93242]
8: 894383 Tillakaratne_Dilshan
[pem_pos=0, final_scores=0.61010873,
final_local_scores=0.6103187, pem=0.0, similarity=22.302841]
9: 23329245 Dilscoop
[pem_pos=0, final_scores=0.12631063,
final_local_scores=0.17279458, pem=-0.14884561,
similarity=10.055694]
1: 2873728 Dougie_Marillier
[pem_pos=0, final_scores=0.262366,
final_local_scores=0.21886677, pem=0.0, similarity=10.621134]
2: 759317 Zimbabwe_national_cricket_team
[pem_pos=1, final_scores=0.7523703,
final_local_scores=0.7281112, pem=-0.92012954,
similarity=30.821812]
3: 759317 Zimbabwe_national_cricket_team
[pem_pos=1, final_scores=0.87141114,
final_local_scores=0.8581452, pem=-2.2776897,
similarity=42.084908]
4: 374833 Australia_national_cricket_team
[pem_pos=1, final_scores=0.72125715,
final_local_scores=0.69597185, pem=-3.4989398,
similarity=43.88668]
5: 2873728 Dougie_Marillier
[pem_pos=0, final_scores=0.6056922,
final_local_scores=0.5948774, pem=0.0, similarity=21.842041]
6: 759317 Zimbabwe_national_cricket_team
[pem_pos=1, final_scores=0.7167599,
final_local_scores=0.68921226, pem=-2.2776897,
similarity=37.04361]
7: 374833 Australia_national_cricket_team
[pem_pos=1, final_scores=0.908521,
final_local_scores=0.9005292, pem=-3.4989398,
similarity=49.99108]
8: 2873728 Dougie_Marillier
[pem_pos=0, final_scores=0.5626791,
final_local_scores=0.54691327, pem=0.0, similarity=20.410696]
9: 423488 Glenn_McGrath
[pem_pos=0, final_scores=0.85220236,
final_local_scores=0.8627669, pem=0.0, similarity=29.8364]

10: 759317 Zimbabwe_national_cricket_team
 [pem_pos=1, final_scores=0.7014271,
 final_local_scores=0.6724635, pem=-2.2776897,
 similarity=36.543793]
 11: 2873728 Dougie_Marillier
 [pem_pos=0, final_scores=0.2764899,
 final_local_scores=0.23429501, pem=0.0, similarity=11.081543]
 12: 2873728 Dougie_Marillier
 [pem_pos=0, final_scores=0.52498883,
 final_local_scores=0.5057424, pem=0.0, similarity=19.182074]
 13: 759317 Zimbabwe_national_cricket_team
 [pem_pos=1, final_scores=0.67687476,
 final_local_scores=0.6456438, pem=-2.2776897,
 similarity=35.74344]
 14: 407754 India_national_cricket_team
 [pem_pos=1, final_scores=0.1903253,
 final_local_scores=0.11579126, pem=-3.5152702,
 similarity=26.661741]
 15: 1180794 Faridabad
 [pem_pos=0, final_scores=0.15511675,
 final_local_scores=0.2603878, pem=0.0, similarity=11.860205]
 16: 14533 India
 [pem_pos=0, final_scores=0.34296995,
 final_local_scores=0.43550754, pem=-0.13622136,
 similarity=17.82692]
 17: 2873728 Dougie_Marillier
 [pem_pos=0, final_scores=0.5701458,
 final_local_scores=0.5550695, pem=0.0, similarity=20.654095]
 18: 1327443 Western_Australia_cricket_team
 [pem_pos=1, final_scores=0.26960373,
 final_local_scores=0.25094354, pem=-2.1796863,
 similarity=23.43184]
 19: 2906540 Ryan_Campbell_(cricketer)
 [pem_pos=0, final_scores=0.6487178,
 final_local_scores=0.7071268, pem=-0.13353139,
 similarity=25.917952]
 20: 894383 Tillakaratne_Dilshan
 [pem_pos=0, final_scores=0.51457673,
 final_local_scores=0.5065435, pem=0.0, similarity=19.205982]
 21: 8258172 ICC_World_Twenty20
 [pem_pos=0, final_scores=0.32542235,
 final_local_scores=0.3139296, pem=-0.06018666,
 similarity=13.785302]
 22: 894383 Tillakaratne_Dilshan
 [pem_pos=0, final_scores=0.54449433,
 final_local_scores=0.53864497, pem=0.0, similarity=20.163954]

Sample: b'Oregon_Route_10'

Oregon Route 10 is an [1 1 Oregon 1] 1 state highway which
 serves [2 2 Portland 2] and some of its western 2 suburb s.
 OR 10 begins as Naito Parkway in downtown [1 Portland 1] ,
 starting where Naito Parkway interchanges with . It heads

south out of downtown , multiplexed with 3 Oregon Route 99W .
 After passing under the [3 3 Portland Aerial Tram 3] , Naito
 Parkway ends at an interchange with Barbur Boulevard ; the
 two routes continue south out of [2 Portland 2] on Barbur .
 OR 10 separates from OR 99W a few miles south of downtown ,
 and proceeds along through the [3 Portland 3] neighborhood of
 4 Hillsdale . Along here , it is a surface street , which
 cuts through the southern part of [4 5 Portland 4] 's [5 West
 Hills 5] . It separates from [6 Capitol Highway 6] in
 Hillsdale , which continues unnumbered , and becomes the [7
 Beaverton-Hillsdale Highway 7] , highway number 40 , a
 surface street with frequent traffic signals . The highway
 continues west into 6 Washington County . In the community of
 [4 4 Raleigh Hills 4] , OR 10 intersects with 7 Oregon Route 2
 10 -LRB- locally known as [8 Scholls Ferry Road 8] -RRB- ,
 which heads southwest towards 8 Progress , [5 5 Tigard 5] ,
 and [6 6 Scholls 6] . OR 10 continues west into [7 7 Beaverton
 7] , where it interchanges with 9 Oregon Route 217 , a
 freeway . West of that interchange , the street name changes
 to [9 Farmington Road 9] , the eastern part of which is not a
 state highway , and comes a block parallel with 10 Oregon Rou
 te 8 in front of 11 Beaverton High School . The portion of OR
 10 from downtown [10 Beaverton 10] to the intersection with 12
 Oregon Route 219 was once known as . OR 8 and 10 do not
 intersect , but it is not uncommon for commuters to use the
 frontage road for the OR 217 interchange or another surface
 street to change between the routes . As [11 Farmington Road 1
 1] , OR 10 leaves [12 Beaverton 12] and cuts across half-
 developed suburbia to 13 Farmington and its intersection with
 OR 219 . This final section comprises the [13 Farmington
 Highway 13] , highway number 142 . : Note : OR 10 uses
 several different highways -LRB- see -RRB- ; the milepoint
 column shows mileage along these highways , not adjusted for

1: 26811621 Oregon

2: 23503 Portland,_Oregon

3: 587562 Portland_Aerial_Tram

4: 131012 Raleigh_Hills,_Oregon

5: 48964 Tigard,_Oregon

6: 6269465 Scholls,_Oregon

7: 42160 Beaverton,_Oregon

1: 332996 State_highway

[winner: 332996 State_highway, pem_pos=0,
 final_scores=-1.0497422, final_local_scores=-1.1153142,
 pem=-0.024095139, similarity=-29.062456]

gt: 332996 State_highway, pem_pos=0,
 final_scores=-1.0497422, final_local_scores=-1.1153142,
 pem=-0.024095139, similarity=-29.062456]

2: 28908 Suburb

[winner: 7960265 Lake_Manchester,_Queensland, pem_pos=8,
 final_scores=-1.0678176, final_local_scores=-0.9879115,
 pem=-2.7070017, similarity=-10.670441]

gt: 28908 Suburb, pem_pos=0, final_scores=-1.4866291,

final_local_scores=-1.6221714, pem=-0.039078053,
similarity=-44.10661]
3: 3735022 Oregon_Route_99W
[winner: 3735022 Oregon_Route_99W, pem_pos=0,
final_scores=-0.017533854, final_local_scores=-0.053759083,
pem=0.0, similarity=2.4854348]
gt: 3735022 Oregon_Route_99W, pem_pos=0,
final_scores=-0.017533854, final_local_scores=-0.053759083,
pem=0.0, similarity=2.4854348]
4: 2924784 Hillsdale,_Portland,_Oregon
[winner: 124984 Hillsdale,_New_Jersey, pem_pos=0,
final_scores=-0.044579178, final_local_scores=-0.029631138,
pem=-2.002263, similarity=14.094079]
gt: 2924784 Hillsdale,_Portland,_Oregon, pem_pos=5,
final_scores=-0.21734485, final_local_scores=-0.20660067,
pem=-2.424167, similarity=11.107334]
5: 6549343 Tualatin_Mountains
[winner: 23503 Portland,_Oregon, pem_pos=0,
final_scores=0.47513336, final_local_scores=0.44863528,
pem=-0.66320467, similarity=21.084496]
RECALL MISS: 6549343 Tualatin_Mountains]
[winner: 611873 West_Hills,_Los_Angeles, pem_pos=0,
final_scores=0.1644315, final_local_scores=0.18988484,
pem=-0.60480255, similarity=13.045266]
gt: 6549343 Tualatin_Mountains, pem_pos=3,
final_scores=0.05904676, final_local_scores=0.036029935,
pem=-1.5588119, similarity=13.641979]
[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.0468526, final_local_scores=-2.120146,
pem=-3.4067376, similarity=-40.653305]
RECALL MISS: 6549343 Tualatin_Mountains]
6: 48966 Washington_County,_Oregon
[winner: 48966 Washington_County,_Oregon, pem_pos=1,
final_scores=-0.020132735, final_local_scores=-0.08337876,
pem=-2.0204854, similarity=12.589238]
gt: 48966 Washington_County,_Oregon, pem_pos=1,
final_scores=-0.020132735, final_local_scores=-0.08337876,
pem=-2.0204854, similarity=12.589238]
7: 2210840 Oregon_Route_210
[winner: 1911992 State_highways_in_Oregon, pem_pos=0,
final_scores=-0.090553835, final_local_scores=-0.06290756,
pem=0.0, similarity=2.2124257]
RECALL MISS: 2210840 Oregon_Route_210]
[winner: 8644468 Saskatchewan_Highway_210, pem_pos=1,
final_scores=-1.0048261, final_local_scores=-0.95700043,
pem=-1.8114557, similarity=-14.618111]
gt: 2210840 Oregon_Route_210, pem_pos=9,
final_scores=-1.1391938, final_local_scores=-1.2337162,
pem=-2.4163866, similarity=-19.586151]
8: 6269943 Progress,_Oregon
[winner: 178206 Progress_(spacecraft), pem_pos=0,
final_scores=-0.30327535, final_local_scores=-0.1679376,

pem=-0.5162685, similarity=1.8856679|
gt: 6269943 Progress, Oregon, pem_pos=5,
final_scores=-0.30873695, final_local_scores=-0.3061944,
pem=-2.795777, similarity=10.156133]
9: 1094662 Oregon_Route_217
[winner: 1911992 State_highways_in_Oregon, pem_pos=0,
final_scores=-0.049519613, final_local_scores=-0.018083908,
pem=0.0, similarity=3.5500531|
RECALL MISS: 1094662 Oregon_Route_217]
[winner: 14881761 ZNF217, pem_pos=7,
final_scores=-1.1002519, final_local_scores=-0.9196281,
pem=-2.7972267, similarity=-8.14207|
RECALL MISS: 1094662 Oregon_Route_217]
10: 1094633 Oregon_Route_8
[winner: 1911992 State_highways_in_Oregon, pem_pos=0,
final_scores=-0.1105151, final_local_scores=-0.0847122,
pem=0.0, similarity=1.5617319|
RECALL MISS: 1094633 Oregon_Route_8]
[winner: 19353 May_8, pem_pos=2, final_scores=-1.0032563,
final_local_scores=-0.83233565, pem=-3.3406982,
similarity=-2.5816064|
RECALL MISS: 1094633 Oregon_Route_8]
11: 5256010 Beaverton_High_School
[winner: 5256010 Beaverton_High_School, pem_pos=0,
final_scores=-0.005659282, final_local_scores=0.046380118,
pem=0.0, similarity=5.4737883|
gt: 5256010 Beaverton_High_School, pem_pos=0,
final_scores=-0.005659282, final_local_scores=0.046380118,
pem=0.0, similarity=5.4737883]
12: 2221462 Oregon_Route_219
[winner: 1911992 State_highways_in_Oregon, pem_pos=0,
final_scores=-0.0732134, final_local_scores=-0.043965757,
pem=0.0, similarity=2.777687|
RECALL MISS: 2221462 Oregon_Route_219]
[winner: 15217559 ZNF219, pem_pos=6,
final_scores=-0.72674924, final_local_scores=-0.49983433,
pem=-2.3566995, similarity=1.9897578|
gt: 2221462 Oregon_Route_219, pem_pos=4,
final_scores=-0.9740725, final_local_scores=-1.0645459,
pem=-1.6051989, similarity=-18.949137]
13: 25503370 Farmington, Oregon
[winner: 125969 Farmington, New_Mexico, pem_pos=2,
final_scores=0.004947379, final_local_scores=-0.0049292743,
pem=-2.3123665, similarity=16.517622|
RECALL MISS: 25503370 Farmington, Oregon]
1: 26811621 Oregon
[pem_pos=0, final_scores=0.5997315,
final_local_scores=0.60270447, pem=-0.28686622,
similarity=23.63564]
2: 23503 Portland, Oregon
[pem_pos=0, final_scores=0.46448272,
final_local_scores=0.437001, pem=-0.66320467,

similarity=20.737305]
3: 587562 Portland_Aerial_Tram
[pem_pos=0, final_scores=0.14320515,
final_local_scores=0.17342699, pem=0.0, similarity=9.2651205]
4: 131012 Raleigh_Hills,_Oregon
[pem_pos=0, final_scores=0.3403179,
final_local_scores=0.36237624, pem=0.0, similarity=14.903743]
5: 48964 Tigard,_Oregon
[pem_pos=0, final_scores=0.42823178,
final_local_scores=0.40520436, pem=0.0, similarity=16.18182]
6: 6269465 Scholls,_Oregon
[pem_pos=0, final_scores=0.17144577,
final_local_scores=0.20810017, pem=0.0, similarity=10.299837]
7: 42160 Beaverton,_Oregon
[pem_pos=0, final_scores=0.37009758,
final_local_scores=0.33579302, pem=-0.36361748,
similarity=16.087854]
1: 23503 Portland,_Oregon
[pem_pos=0, final_scores=0.4836563,
final_local_scores=0.45794523, pem=-0.66320467,
similarity=21.362322]
2: 23503 Portland,_Oregon
[pem_pos=0, final_scores=0.49513704,
final_local_scores=0.47048616, pem=-0.66320467,
similarity=21.736568]
3: 23503 Portland,_Oregon
[pem_pos=0, final_scores=0.50481325,
final_local_scores=0.48105603, pem=-0.66320467,
similarity=22.051994]
4: 23503 Portland,_Oregon
[pem_pos=0, final_scores=0.47513336,
final_local_scores=0.44863528, pem=-0.66320467,
similarity=21.084496]
5: 611873 West_Hills,_Los_Angeles
[pem_pos=0, final_scores=0.1644315,
final_local_scores=0.18988484, pem=-0.60480255,
similarity=13.045266]
6: 3735022 Oregon_Route_99W
[pem_pos=0, final_scores=0.14946733,
final_local_scores=0.13853258, pem=0.0, similarity=8.223802]
7: 1094577 Oregon_Route_10
[pem_pos=0, final_scores=0.059357718,
final_local_scores=0.06681654, pem=0.0, similarity=6.0836515]
8: 2210840 Oregon_Route_210
[pem_pos=0, final_scores=0.24805598,
final_local_scores=0.2952155, pem=0.0, similarity=12.899532]
9: 1094577 Oregon_Route_10
[pem_pos=0, final_scores=0.20173614,
final_local_scores=0.22234333, pem=0.0, similarity=10.724881]
10: 42160 Beaverton,_Oregon
[pem_pos=0, final_scores=0.3376277,
final_local_scores=0.30032462, pem=-0.36361748,

similarity=15.029407]
11: 1094577 Oregon_Route_10
[pem_pos=0, final_scores=0.20960866,
final_local_scores=0.23094288, pem=0.0, similarity=10.981509]
12: 42160 Beaverton, Oregon
[pem_pos=0, final_scores=0.29447103,
final_local_scores=0.2531825, pem=-0.36361748,
similarity=13.622592]
13: 1094577 Oregon_Route_10
[pem_pos=0, final_scores=0.20268978,
final_local_scores=0.22338504, pem=0.0, similarity=10.755968]

Sample: b'Party_of_New_Forces'

Parti des forces nouvelles -LRB- PFN -RRB- or [1 Party of New Forces 1] was a 1 political party formed in November 1974 from the Comité faire front , a group of anti- [1 1 Jean-Marie Le Pen 1] dissidents who had split from the [2 2 National Front 2] -LRB- FN -RRB- . The group included amongst its early members most of the membership of , which had dissolved not long before the formation of the [2 PFN 2] , Alain Robert -LRB- the founder of [3 3 Occident 3] and the [4 4 Groupe Union Défense 4] or GUD -RRB- , the academic Pascal Gauchon , the journalists and [5 5 Roland Gaucher 5] and the draughtsman Jack Marchal . A youth movement , Front de la jeunesse , was formed , although the party was also closely linked to GUD . Positioned on the far right , the [3 PFN 3] also sought links with the more mainstream right and joined former members of the 2 Organisation de l'armée secrète in campaigning for [4 3 Valéry Giscard 4] d'Estaing in 1974 . The group also launched its own well-produced journal , Initiative nationale , organised protests against the 1977 visit to [6 6 Paris 6] by [7 7 Leonid Brezhnev 7] -LRB- on the pretext of his support for the [8 8 Polisario Front 8] , which had taken [5 French 5] hostages -RRB- and in 1979 launched the Eurodroite alliance with the [9 9 Italian Social Movement 9] , [10 10 Fuerza Nueva 10] and the [11 11 Belgian 11] [6 PFN 6] . The party ran for the 4 1979 [7 European 7] elections under the name Union française pour l'Eurodroite -LRB- led by [12 12 Jean-Louis Tixier-Vignancour 12] -RRB- , winning 1.3 % of the vote . They would go on to have members elected to town councils in 1983 as part of the [13 13 Rally for the Republic 13] -LRB- RPR -RRB- - [14 14 Union for French Democracy 14] -LRB- UDF -RRB- list . Following the 1979 [8 European 8] elections , [9 Roland Gaucher 9] , who had been in charge of Initiative nationale , left the [10 PFN 10] along with François Brigneau to join again the [15 15 National Front 15] -LRB- FN -RRB- . The party also endured failure , notably in 1981 when they were unable to secure the 500 signatures necessary to get Pascal Gauchon as a candidate for the presidency . Following this set-back leadership fell into the hands of young members Roland Hélie , Didier Lecerf , Jack Marchal and Olivier Cazal , with former leaders such as [16 16 Hervé Novelli 16]

and [11 Alain Robert 11] leaving to join the 5 National Centre of Independents and Peasants . The party then became involved in 6 anti-communism activities , occupying [12 7 French 12] Communist Party ministries and joining [13 RPR 13] supporters in breaking up a rally by communist ex-servicemen in a move that provoked scandal for the [14 RPR 14] . The party itself split in 1986 with a [15 European 15] group known as splitting from a tendency rechristened Natrope -LRB- Nationalistes européens -RRB- , which was close to the [17 17 Nouvelle Droite 17] ideas of [18 18 Alain de Benoist 18] and 8 GRECE

1: 49955 Jean-Marie_Le_Pen
2: 67119 National_Front_(France)
3: 320275 Occident_(movement)
4: 320279 Groupe_Union_Défense
5: 320311 Roland_Gaucher
6: 22989 Paris
7: 50151 Leonid_Brezhnev
8: 201574 Polisario_Front
9: 33643342 Italian_Social_Movement
10: 5037074 New_Force_(Spain)
11: 3343 Belgium
12: 4522984 Jean-Louis_Tixier-Vignancour
13: 323430 Rally_for_the_Republic
14: 323466 Union_for_French_Democracy
15: 67119 National_Front_(France)
16: 11205713 Hervé_Novelli
17: 1490234 Nouvelle_Droite
18: 652176 Alain_de_Benoist

1: 23996 Political_party
[winner: 20965184 Ukrainian_Radical_Party, pem_pos=5, final_scores=-1.4094133, final_local_scores=-1.487592, pem=-1.6091769, similarity=-31.552044]
gt: 23996 Political_party, pem_pos=0, final_scores=-1.4965285, final_local_scores=-1.6491237, pem=-0.070212804, similarity=-44.7416]
2: 62191 Organisation_de_l'armée_secrète
[winner: 62191 Organisation_de_l'armée_secrète, pem_pos=0, final_scores=-0.17962645, final_local_scores=-0.22175345, pem=0.0, similarity=-2.5278525]
gt: 62191 Organisation_de_l'armée_secrète, pem_pos=0, final_scores=-0.17962645, final_local_scores=-0.22175345, pem=0.0, similarity=-2.5278525]
3: 79474 Valéry_Giscard_d'Estaing
[winner: 79474 Valéry_Giscard_d'Estaing, pem_pos=0, final_scores=0.6836811, final_local_scores=0.69899756, pem=0.0, similarity=24.949194]
gt: 79474 Valéry_Giscard_d'Estaing, pem_pos=0, final_scores=0.6836811, final_local_scores=0.69899756, pem=0.0, similarity=24.949194]
[winner: 280518 Charles_Hector,_comte_d'Estaing,

pem_pos=0, final_scores=-1.7186933,
final_local_scores=-1.8254633, pem=0.0, similarity=-50.38575|
RECALL MISS: 79474 Valéry_Giscard_d'Estaing]

4: 9737190 European_Parliament_election,_1979
[winner: 9239 Europe, pem_pos=0, final_scores=0.18310274,
final_local_scores=0.2305628, pem=-0.25557294,
similarity=12.360013|
RECALL MISS: 9737190 European_Parliament_election,_1979]
[winner: 34754 1979, pem_pos=0, final_scores=-0.48136127,
final_local_scores=-0.21149975, pem=-0.65922886,
similarity=1.3631275|
gt: 9737190 European_Parliament_election,_1979,
pem_pos=13, final_scores=-1.7011929,
final_local_scores=-1.8692908, pem=-4.600358,
similarity=-26.676191]
[winner: 1356050 Elections_in_Bermuda, pem_pos=19,
final_scores=-1.4916228, final_local_scores=-1.4801593,
pem=-3.5263605, similarity=-20.904299|
RECALL MISS: 9737190 European_Parliament_election,_1979]

5: 30863436 National_Centre_of_Independents_and_Peasants
[winner: 30863436
National_Centre_of_Independents_and_Peasants, pem_pos=0,
final_scores=-0.056189418, final_local_scores=-0.07123183,
pem=0.0, similarity=1.9640132|
gt: 30863436
National_Centre_of_Independents_and_Peasants, pem_pos=0,
final_scores=-0.056189418, final_local_scores=-0.07123183,
pem=0.0, similarity=1.9640132]

6: 18940658 Anti-communism
[winner: 18940658 Anti-communism, pem_pos=0,
final_scores=-1.1214702, final_local_scores=-1.2649248,
pem=-0.003099412, similarity=-33.64131|
gt: 18940658 Anti-communism, pem_pos=0,
final_scores=-1.1214702, final_local_scores=-1.2649248,
pem=-0.003099412, similarity=-33.64131]

7: 319197 French_Communist_Party
[winner: 5843419 France, pem_pos=0,
final_scores=0.32091743, final_local_scores=0.31778228,
pem=-0.80223274, similarity=17.935637|
RECALL MISS: 319197 French_Communist_Party]
[winner: 37008 Communist_party, pem_pos=0,
final_scores=0.02572526, final_local_scores=-0.0047018975,
pem=-1.2514052, similarity=10.754735|
gt: 319197 French_Communist_Party, pem_pos=6,
final_scores=-0.2712292, final_local_scores=-0.37956846,
pem=-3.1580987, similarity=9.936865]

8: 3274373
Groupement_de_recherche_et_d'études_pour_la_civilisation_euro
péenne
[winner: 31724247 Grece,_Pezë, pem_pos=0,
final_scores=-0.23894104, final_local_scores=-0.029468149,
pem=0.0, similarity=3.2103245|

RECALL MISS: 3274373

Groupement_de_recherche_et_d'études_pour_la_civilisation_euro
péenne]

- 1: 49955 Jean-Marie_Le_Pen
[pem_pos=0, final_scores=0.49429506,
final_local_scores=0.47817868, pem=0.0, similarity=18.359518]
- 2: 67119 National_Front_(France)
[pem_pos=1, final_scores=0.03864415,
final_local_scores=-0.022050142, pem=-1.2167886,
similarity=10.048779]
- 3: 320275 Occident_(movement)
[pem_pos=1, final_scores=0.14839612,
final_local_scores=0.18216187, pem=-0.97700524,
similarity=14.838893]
- 4: 320279 Groupe_Union_Défense
[pem_pos=0, final_scores=0.26340407,
final_local_scores=0.2930774, pem=0.0, similarity=12.835728]
- 5: 320311 Roland_Gaucher
[pem_pos=0, final_scores=0.585908,
final_local_scores=0.61716807, pem=0.0, similarity=22.507238]
- 6: 22989 Paris
[pem_pos=0, final_scores=0.70079786,
final_local_scores=0.7423115, pem=-0.115301244,
similarity=26.868793]
- 7: 50151 Leonid_Brezhnev
[pem_pos=0, final_scores=0.5834057,
final_local_scores=0.64034265, pem=0.0, similarity=23.198814]
- 8: 201574 Polisario_Front
[pem_pos=0, final_scores=0.3314187,
final_local_scores=0.382891, pem=0.0, similarity=15.515944]
- 9: 33643342 Italian_Social_Movement
[pem_pos=0, final_scores=0.15170817,
final_local_scores=0.16166574, pem=0.0, similarity=8.914141]
- 10: 5037074 New_Force_(Spain)
[pem_pos=0, final_scores=0.16340964,
final_local_scores=0.20751381, pem=0.0, similarity=10.282339]
- 11: 3343 Belgium
[pem_pos=0, final_scores=0.19727059,
final_local_scores=0.20324382, pem=-0.12403572,
similarity=10.82944]
- 12: 4522984 Jean-Louis_Tixier-Vignancour
[pem_pos=0, final_scores=0.46129227,
final_local_scores=0.5117821, pem=0.0, similarity=19.362312]
- 13: 323430 Rally_for_the_Republic
[pem_pos=0, final_scores=0.09609412,
final_local_scores=0.06044124, pem=0.0, similarity=5.8934]
- 14: 323466 Union_for_French_Democracy
[pem_pos=0, final_scores=0.14583229,
final_local_scores=0.116719246, pem=0.0, similarity=7.572848]
- 15: 67119 National_Front_(France)
[pem_pos=1, final_scores=0.053404436,
final_local_scores=-0.005926788, pem=-1.2167886,

similarity=10.529932]
16: 11205713 Hervé Novelli
[pem_pos=0, final_scores=0.34342438,
final_local_scores=0.42313308, pem=0.0, similarity=16.716848]
17: 1490234 Nouvelle Droite
[pem_pos=0, final_scores=0.2554283,
final_local_scores=0.29120833, pem=0.0, similarity=12.77995]
18: 652176 Alain de Benoist
[pem_pos=0, final_scores=0.56366843,
final_local_scores=0.61205065, pem=0.0, similarity=22.354525]
1: 10932302 Party_of_New_Forces
[pem_pos=0, final_scores=0.23292379,
final_local_scores=0.23037165, pem=0.0, similarity=10.964462]
2: 10932302 Party_of_New_Forces
[pem_pos=0, final_scores=0.25184256,
final_local_scores=0.25103757, pem=0.0, similarity=11.581175]
3: 10932302 Party_of_New_Forces
[pem_pos=0, final_scores=0.30022663,
final_local_scores=0.3038898, pem=0.0, similarity=13.158391]
4: 79474 Valéry Giscard d'Estaing
[pem_pos=0, final_scores=0.6836811,
final_local_scores=0.69899756, pem=0.0, similarity=24.949194]
5: 5843419 France
[pem_pos=0, final_scores=0.6543588,
final_local_scores=0.6820163, pem=-0.80223274,
similarity=28.805107]
6: 10932302 Party_of_New_Forces
[pem_pos=0, final_scores=0.21624576,
final_local_scores=0.21215346, pem=0.0, similarity=10.420795]
7: 9239 Europe
[pem_pos=0, final_scores=0.18310274,
final_local_scores=0.2305628, pem=-0.25557294,
similarity=12.360013]
8: 9239 Europe
[pem_pos=0, final_scores=0.15851901,
final_local_scores=0.20370883, pem=-0.25557294,
similarity=11.558636]
9: 320311 Roland Gaucher
[pem_pos=0, final_scores=0.51227427,
final_local_scores=0.5367344, pem=0.0, similarity=20.106937]
10: 10932302 Party_of_New_Forces
[pem_pos=0, final_scores=0.35876328,
final_local_scores=0.36783218, pem=0.0, similarity=15.06656]
11: 384286 Alain Robert
[pem_pos=0, final_scores=0.09465708,
final_local_scores=0.18730173, pem=0.0, similarity=9.67917]
12: 5843419 France
[pem_pos=0, final_scores=0.32091743,
final_local_scores=0.31778228, pem=-0.80223274,
similarity=17.935637]
13: 323430 Rally_for_the_Republic
[pem_pos=0, final_scores=0.33313173,

```

final_local_scores=0.31936875, pem=0.0, similarity=13.620314]
14: 323430 Rally_for_the_Republic
    [pem_pos=0, final_scores=0.191299,
final_local_scores=0.1644381, pem=0.0, similarity=8.996874]
15: 9239 Europe
    [pem_pos=0, final_scores=0.21535964,
final_local_scores=0.26579857, pem=-0.25557294,
similarity=13.411517]

```

Sample: b'River_Bride'

The [1 River Bride 1] is a 1 tributary of the [1 1 Munster Bl
ackwater 1] in [2 2 Ireland 2] . -LRB- It should not be
confused with the [2 River Bride 2] in [3 Dorset 3] , [4
England 4] . -RRB- Rising in [3 3 County Cork 3] 's , it
flows eastward , passing through the counties of and before
reaching the [5 Munster Blackwater 5] , approximately inland
from . The famous [6 English 6] poet [7 Edmund Spenser 7] is
reputed to have written part of his poem the ` Faerie Queen '
on the banks of The Bride in the [8 Conna 8] area . 2 Kilcrea
Friary is on the southern bank . The villages of [4 4 Carrigr
ohane 4] , [5 5 Castlelyons 5] , [6 6 Conna 6] , [7 7 Farran 7
] , [8 8 Killumney 8] , and 3 Ovens are situated on or near
the river . The river valley runs through the 4 Barony of Bar
rymore and 5 Imokilly . The river is crossed at , [9 9 County
Waterford 9] by the 6 N72 road

```

1: 3863197 Munster_Blackwater
2: 14560 Republic_of_Ireland
3: 46709 County_Cork
4: 35019114 Carrigrohane
5: 10522982 Castlelyons
6: 7125947 Conna
7: 18452873 Farran
8: 25232097 Killumney
9: 74190 County_Waterford
1: 72465 Tributary

```

```

    [winner: 72465 Tributary, pem_pos=1,
final_scores=-1.1790901, final_local_scores=-1.1695669,
pem=-0.012074053, similarity=-30.746836]
    gt: 72465 Tributary, pem_pos=1, final_scores=-1.1790901,
final_local_scores=-1.1695669, pem=-0.012074053,
similarity=-30.746836 ]
2: 1618790 Kilcrea_Friary
    [winner: 1618790 Kilcrea_Friary, pem_pos=0,
final_scores=-0.024285048, final_local_scores=0.019604012,
pem=0.0, similarity=4.6747355]
    gt: 1618790 Kilcrea_Friary, pem_pos=0,
final_scores=-0.024285048, final_local_scores=0.019604012,
pem=0.0, similarity=4.6747355 ]
3: 556000 Ovens,_County_Cork
    [winner: 1794439 Ovens_River, pem_pos=1,
final_scores=-0.06847224, final_local_scores=-0.012316793,
pem=-1.0096895, similarity=9.213003]

```

gt: 556000 Ovens, County_Cork, pem_pos=2,
final_scores=-0.14190285, final_local_scores=-0.15308462,
pem=-1.1998072, similarity=6.0461]

4: 26997293 Barony_of_Barrymore

[winner: 26997293 Barony_of_Barrymore, pem_pos=0,
final_scores=-0.11483, final_local_scores=-0.08974384,
pem=0.0, similarity=1.4115778]

gt: 26997293 Barony_of_Barrymore, pem_pos=0,
final_scores=-0.11483, final_local_scores=-0.08974384,
pem=0.0, similarity=1.4115778]

5: 34196866 Imokilly

[winner: 25060627 Imokilly_GAA, pem_pos=0,
final_scores=-0.18168676, final_local_scores=-0.12597412,
pem=0.0, similarity=0.33039406]

gt: 34196866 Imokilly, pem_pos=1,
final_scores=-0.48691958, final_local_scores=-0.49145797,
pem=-2.5455313, similarity=3.2666218]

6: 6117118 N72_road_(Ireland)

[winner: 6117118 N72_road_(Ireland), pem_pos=0,
final_scores=0.021571204, final_local_scores=0.08396424,
pem=0.0, similarity=6.5953736]

gt: 6117118 N72_road_(Ireland), pem_pos=0,
final_scores=0.021571204, final_local_scores=0.08396424,
pem=0.0, similarity=6.5953736]

[winner: 20324545 Jalan_Gunung_Brinchang, pem_pos=8,
final_scores=-1.2729642, final_local_scores=-1.1745577,
pem=-3.265966, similarity=-13.200603]

RECALL MISS: 6117118 N72_road_(Ireland)]

1: 3863197 Munster_Blackwater

[pem_pos=0, final_scores=0.23590384,
final_local_scores=0.23214158, pem=0.0, similarity=11.017281]

2: 14560 Republic_of_Ireland

[pem_pos=1, final_scores=0.45563895,
final_local_scores=0.4630689, pem=-1.5040284,
similarity=26.087751]

3: 46709 County_Cork

[pem_pos=0, final_scores=0.38675052,
final_local_scores=0.35054463, pem=0.0, similarity=14.550664]

4: 35019114 Carrigrohane

[pem_pos=0, final_scores=0.12539272,
final_local_scores=0.1656658, pem=0.0, similarity=9.033511]

5: 10522982 Castlelyons

[pem_pos=0, final_scores=0.24671756,
final_local_scores=0.25371838, pem=0.0, similarity=11.661176]

6: 7125947 Conna

[pem_pos=0, final_scores=0.08773036,
final_local_scores=0.100637525, pem=0.0, similarity=7.0929375]

7: 18452873 Farran

[pem_pos=0, final_scores=0.08414452,
final_local_scores=0.096393034, pem=0.0, similarity=6.9662733]

8: 25232097 Killumney

[pem_pos=0, final_scores=0.24331717,

final_local_scores=0.3395752, pem=0.0, similarity=14.223314]
 9: 74190 County_Waterford
 [pem_pos=0, final_scores=0.34404337,
 final_local_scores=0.3088507, pem=0.0, similarity=13.306434]
 1: 35021282 River_Bride
 [pem_pos=0, final_scores=0.087173626,
 final_local_scores=0.11519641, pem=0.0, similarity=7.5274034]
 2: 35021282 River_Bride
 [pem_pos=0, final_scores=0.14598106,
 final_local_scores=0.1794346, pem=0.0, similarity=9.444399]
 3: 37589 Dorset
 [pem_pos=0, final_scores=0.48974258,
 final_local_scores=0.49893844, pem=-0.1353917,
 similarity=19.715311]
 4: 9316 England
 [pem_pos=0, final_scores=0.61282295,
 final_local_scores=0.6520476, pem=-0.19053856,
 similarity=24.584291]
 5: 3863197 Munster_Blackwater
 [pem_pos=0, final_scores=0.28984457,
 final_local_scores=0.29106367, pem=0.0, similarity=12.775633]
 6: 9316 England
 [pem_pos=1, final_scores=0.43918878,
 final_local_scores=0.46237856, pem=-1.3527626,
 similarity=25.244543]
 7: 9536 Edmund_Spenser
 [pem_pos=0, final_scores=0.14367674,
 final_local_scores=0.19914198, pem=0.0, similarity=10.032507]
 8: 7125947 Conna
 [pem_pos=0, final_scores=0.23863493,
 final_local_scores=0.26547784, pem=0.0, similarity=12.012101]

Sample: b'St'

__ Michael_the_Archangel_Church _ -LRB- Cleveland , _ Ohio -
 RRB- St. Michael the Archangel is a 1 church located at 3114
 Scranton Road in the 2 Clark-Fulton neighborhood on the west
 side of [1 1 Cleveland 1] , [2 2 Ohio 2] . The church is
 named in honor of and was completed in 1892 . The 3 congregat
 ion was founded in 1881 as a of [1 4 Ohio City 1] , [2 Clevel
 and 2] 's to serve a growing [3 3 German 3] 5 population on
 west side . The original church and 6 school 7 building was
 constructed in 1883 but burned on June 29 , 1891 , while the
 new building was under construction . The current church ,
 which had its 8 cornerstone laid in 1889 , was completed five
 years ahead of schedule in 1892 . The church building is
 considered a good example of [3 9 High Victorian Gothic 3] ar
 chitecture with its two 10 tower s of unequal 11 height , the
 taller of which rises 232 feet , and archways . An 12 archite
 cturally notable 13 school adjacent to the 14 church began
 construction in 1906 and was completed in 1907 . For many
 years the church was one of the most costly and artistically
 notable churches in the [4 4 Cleveland Diocese 4] . The

reached its peak size in the 1950 , when only 25 % of the 15 parish ioners were of [5 5 German 5] 16 descent . The first 17 Spanish 18 Mass was said in 1971 for the growing number of 19 Hispanic parishioners . The congregation now is mostly . The church now offers [4 Masses 4] in both [6 6 English 6] and [5 Spanish 5] . An annual 20 Good Friday 21 Procession begins or ends at St. Michael 's . St Michael 's and [6 La Sagrada Familia 6] church alternate every year to begin or end at one of the two churches . There is a stop at [7 St Patrick 7] 's on Bridge Avenue to hear the first part of Mass before continuing on the journey to either St Michael 's or [8 La Sagrada Familia 8] . The 22 parish has become a center of in [7 7 Cleveland 7]

1: 5951 Cleveland

2: 22199 Ohio

3: 11867 Germany

4: 4347620 Roman_Catholic_Diocese_of_Cleveland

5: 11867 Germany

6: 8569916 English_language

7: 5951 Cleveland

1: 6325 Church_(building)

[winner: 6325 Church_(building), pem_pos=0, final_scores=-1.2317668, final_local_scores=-1.3916597, pem=-0.35553187, similarity=-35.50675]

gt: 6325 Church_(building), pem_pos=0, final_scores=-1.2317668, final_local_scores=-1.3916597, pem=-0.35553187, similarity=-35.50675]

2: 32468717 Clark-Fulton

[winner: 32468717 Clark-Fulton, pem_pos=0, final_scores=-0.08508085, final_local_scores=0.040265515, pem=0.0, similarity=5.291316]

gt: 32468717 Clark-Fulton, pem_pos=0, final_scores=-0.08508085, final_local_scores=0.040265515, pem=0.0, similarity=5.291316]

3: 1934020 Congregation_(Catholic)

[winner: 29283909 The_Congregation_(album), pem_pos=1, final_scores=-0.80237937, final_local_scores=-0.6938647, pem=-1.0055218, similarity=-11.14841]

RECALL MISS: 1934020 Congregation_(Catholic)]

4: 3123011 Ohio_City,_Cleveland

[winner: 5951 Cleveland, pem_pos=0, final_scores=0.47399884, final_local_scores=0.461766, pem=-0.57246673, similarity=20.982895]

RECALL MISS: 3123011 Ohio_City,_Cleveland]

[winner: 3123011 Ohio_City,_Cleveland, pem_pos=0, final_scores=0.12817548, final_local_scores=0.16963097, pem=0.0, similarity=9.151839]

gt: 3123011 Ohio_City,_Cleveland, pem_pos=0, final_scores=0.12817548, final_local_scores=0.16963097, pem=0.0, similarity=9.151839]

[winner: 37537378 Admissions_(2011_film), pem_pos=12, final_scores=-1.7847583, final_local_scores=-1.7752807,

pem=-4.6728287, similarity=-23.476631|
RECALL MISS: 3123011 Ohio_City,_Cleveland|
[winner: 1415027 David_Berman_(musician), pem_pos=5,
final_scores=-2.1446934, final_local_scores=-2.2589011,
pem=-3.0012724, similarity=-46.999016|
RECALL MISS: 3123011 Ohio_City,_Cleveland|
5: 22949 Population
[winner: 12702333 Population_(album), pem_pos=1,
final_scores=-1.0557535, final_local_scores=-0.99724305,
pem=-0.6897683, similarity=-21.918941|
gt: 22949 Population, pem_pos=0,
final_scores=-1.9538308, final_local_scores=-2.0834937,
pem=-0.26559433, similarity=-56.641552]
6: 28022 School
[winner: 28022 School, pem_pos=0,
final_scores=-1.3937321, final_local_scores=-1.5506439,
pem=-0.2358994, similarity=-40.901737|
gt: 28022 School, pem_pos=0, final_scores=-1.3937321,
final_local_scores=-1.5506439, pem=-0.2358994,
similarity=-40.901737]
7: 45995 Building
[winner: 17765490 Coca-
Cola_Bottling_Company_Building_(Columbia,_Missouri),
pem_pos=14, final_scores=-1.0745243,
final_local_scores=-1.0102663, pem=-2.7080503,
similarity=-11.33185|
gt: 45995 Building, pem_pos=0, final_scores=-1.5244782,
final_local_scores=-1.6753936, pem=-0.24166258,
similarity=-44.593174]
8: 452955 Cornerstone
[winner: 16596945 Operation_Cornerstone, pem_pos=4,
final_scores=-0.7661896, final_local_scores=-0.61247796,
pem=-1.7110246, similarity=-4.8830338|
gt: 452955 Cornerstone, pem_pos=2,
final_scores=-1.5608519, final_local_scores=-1.7326045,
pem=-1.6188121, similarity=-38.81131]
9: 521613 Victorian_architecture
[winner: 311821 Gothic_Revival_architecture, pem_pos=0,
final_scores=0.27733475, final_local_scores=0.2900379,
pem=0.0, similarity=12.745023|
gt: 521613 Victorian_architecture, pem_pos=1,
final_scores=0.10294239, final_local_scores=0.11211863,
pem=-0.66101587, similarity=11.030264]
[winner: 6206939 Sikh_architecture, pem_pos=19,
final_scores=-1.3932941, final_local_scores=-1.3786209,
pem=-3.3900714, similarity=-18.615349|
RECALL MISS: 521613 Victorian_architecture|
10: 214505 Tower
[winner: 26900577 Tower_(typeface), pem_pos=18,
final_scores=-1.2720933, final_local_scores=-1.2316961,
pem=-3.9470859, similarity=-11.20169|
gt: 214505 Tower, pem_pos=0, final_scores=-1.6039201,

final_local_scores=-1.736762, pem=-1.4821498,
similarity=-39.678566]
11: 239972 Height
[winner: 27411333 Height_(album), pem_pos=2,
final_scores=-0.9917037, final_local_scores=-0.90171856,
pem=-1.379667, similarity=-15.316523]
gt: 239972 Height, pem_pos=6, final_scores=-1.6243007,
final_local_scores=-1.6738453, pem=-2.3014774,
similarity=-33.345375]
12: 21296224 Architecture
[winner: 48364 Software_architecture, pem_pos=0,
final_scores=-1.2373953, final_local_scores=-1.2168872,
pem=-0.50226444, similarity=-29.493238]
gt: 21296224 Architecture, pem_pos=1,
final_scores=-1.7158136, final_local_scores=-1.8983142,
pem=-1.2809339, similarity=-45.593853]
13: 28022 School
[winner: 28022 School, pem_pos=0,
final_scores=-1.4237765, final_local_scores=-1.583463,
pem=-0.2358994, similarity=-41.881126]
gt: 28022 School, pem_pos=0, final_scores=-1.4237765,
final_local_scores=-1.583463, pem=-0.2358994,
similarity=-41.881126]
14: 6325 Church_(building)
[winner: 564181 The_Church_(band), pem_pos=0,
final_scores=-1.3112195, final_local_scores=-1.3211405,
pem=0.0, similarity=-35.33576]
RECALL MISS: 6325 Church_(building)]
15: 23623 Parish
[winner: 23623 Parish, pem_pos=0,
final_scores=-1.1077183, final_local_scores=-1.1922212,
pem=-0.2306374, similarity=-30.234306]
gt: 23623 Parish, pem_pos=0, final_scores=-1.1077183,
final_local_scores=-1.1922212, pem=-0.2306374,
similarity=-30.234306]
16: 3218648 Cultural_heritage
[winner: 24555509 Descent_(song), pem_pos=3,
final_scores=-1.3050699, final_local_scores=-1.2319808,
pem=-2.6019576, similarity=-18.5252]
RECALL MISS: 3218648 Cultural_heritage]
17: 26825 Spanish_language
[winner: 26825 Spanish_language, pem_pos=0,
final_scores=-0.4382686, final_local_scores=-0.5023112,
pem=-0.6062411, similarity=-7.603421]
gt: 26825 Spanish_language, pem_pos=0,
final_scores=-0.4382686, final_local_scores=-0.5023112,
pem=-0.6062411, similarity=-7.603421]
18: 56371 Mass_(liturgy)
[winner: 19048 Mass, pem_pos=1, final_scores=-0.59048337,
final_local_scores=-0.5707422, pem=-1.0387344,
similarity=-7.293573]
gt: 56371 Mass_(liturgy), pem_pos=0,

final_scores=-0.65858954, final_local_scores=-0.73714685,
pem=-0.6321225, similarity=-14.470638]
19: 56120 Hispanic
[winner: 26667 Spain, pem_pos=2,
final_scores=-0.025392592, final_local_scores=-0.03067252,
pem=-2.391836, similarity=16.181559]
gt: 56120 Hispanic, pem_pos=1,
final_scores=-0.058742896, final_local_scores=-0.08648148,
pem=-1.3790978, similarity=9.008681]
20: 58009 Good_Friday
[winner: 2630391 Good, West_Virginia, pem_pos=3,
final_scores=-0.8882395, final_local_scores=-0.82514554,
pem=-2.2099226, similarity=-8.516374]
gt: 58009 Good_Friday, pem_pos=16,
final_scores=-1.574906, final_local_scores=-1.7359571,
pem=-4.094783, similarity=-25.446644]
[winner: 10189895 Friday_(soundtrack), pem_pos=9,
final_scores=-1.6478635, final_local_scores=-1.6208608,
pem=-3.360129, similarity=-26.007103]
gt: 58009 Good_Friday, pem_pos=3,
final_scores=-2.8655553, final_local_scores=-3.1457953,
pem=-3.0196342, similarity=-73.36585]
21: 1686010 Procession
[winner: 4536806 Procession_(album), pem_pos=1,
final_scores=-0.8681799, final_local_scores=-0.7596522,
pem=-1.1940823, similarity=-12.086219]
gt: 1686010 Procession, pem_pos=0,
final_scores=-0.9548768, final_local_scores=-1.0151751,
pem=-0.2923076, similarity=-24.61552]
22: 23623 Parish
[winner: 1104809 Meopham, pem_pos=0,
final_scores=-0.79124385, final_local_scores=-0.8410059,
pem=0.0, similarity=-21.007582]
RECALL MISS: 23623 Parish]
1: 5951 Cleveland
[pem_pos=0, final_scores=0.61140865,
final_local_scores=0.6118653, pem=-0.57246673,
similarity=25.462156]
2: 22199 Ohio
[pem_pos=0, final_scores=0.51137173,
final_local_scores=0.50496346, pem=-0.19147073,
similarity=20.200077]
3: 11867 Germany
[pem_pos=1, final_scores=0.49149197,
final_local_scores=0.52181524, pem=-1.0465823,
similarity=25.353197]
4: 4347620 Roman_Catholic_Diocese_of_Cleveland
[pem_pos=0, final_scores=0.21420972,
final_local_scores=0.2562391, pem=0.0, similarity=11.736399]
5: 11867 Germany
[pem_pos=1, final_scores=0.3937739,
final_local_scores=0.4150731, pem=-1.0465823,

similarity=22.167799]
6: 8569916 English_language
[pem_pos=0, final_scores=0.48112243,
final_local_scores=0.5208898, pem=-0.640266,
similarity=23.115969]
7: 5951 Cleveland
[pem_pos=0, final_scores=0.32634658,
final_local_scores=0.30047834, pem=-0.57246673,
similarity=16.169748]
1: 3123011 Ohio_City,_Cleveland
[pem_pos=0, final_scores=0.12817548,
final_local_scores=0.16963097, pem=0.0, similarity=9.151839]
2: 5951 Cleveland
[pem_pos=0, final_scores=0.47399884,
final_local_scores=0.461766, pem=-0.57246673,
similarity=20.982895]
3: 311821 Gothic_Revival_architecture
[pem_pos=0, final_scores=0.27733475,
final_local_scores=0.2900379, pem=0.0, similarity=12.745023]
4: 56371 Mass_(liturgy)
[pem_pos=0, final_scores=0.07372989,
final_local_scores=0.079016075, pem=-0.09591032,
similarity=6.9692855]
5: 26825 Spanish_language
[pem_pos=0, final_scores=0.21363376,
final_local_scores=0.2253105, pem=-0.6062411,
similarity=14.110261]
6: 59545 Sagrada_Família
[pem_pos=0, final_scores=0.087616846,
final_local_scores=0.13263956, pem=0.0, similarity=8.047942]
7: 8161655 Saint_Patrick
[pem_pos=0, final_scores=0.08779566,
final_local_scores=0.10507733, pem=0.0, similarity=7.22543]
8: 59545 Sagrada_Família
[pem_pos=0, final_scores=0.14926104,
final_local_scores=0.19997644, pem=0.0, similarity=10.057409]

Sample: b'Tommy_Cairns'

[1 Tommy Cairns 1] -LRB- 30 October 1890 -- December 1967 -
RRB- was a [1 1 Scottish 1] 1 footballer who played for [2 2 B
ristol City 2] , [3 3 Peebles Rovers 3] , [4 4 St. Johnstone 4
] , [5 5 Rangers 5] , [6 6 Bradford City 6] and [7 7 Scotland
7] . Born in , [8 8 Lanarkshire 8] , [2 Cairns 2] ' made his
name in 2 Junior football , helping Burnbank Athletic to the [9 9 Scottish Junior Cup 9] in 1911 . He also appeared for [10
10 Larkhall Thistle 10] before earning an opportunity in the
senior leagues . His first professional club was [3 Bristol
City 3] where he played for two seasons from 1911 . He joined
[4 Peebles Rovers 4] in 1913 and had a short spell with [5
St. Johnstone 5] before joining [6 Rangers 6] in November
1913 . He made his debut against [11 11 Hamilton Academical 11
] on 27 December 1913 and played a total of nine games in his

first season as [7 Rangers 7] finished second to rivals [12 12 Celtic 12] in [8 Division One 8] . [9 Cairns 9] was a regular in the [10 Rangers 10] team the , missing just one league game . [11 Cairns 11] continued to be a regular in the team for the next two seasons but it was in that he won his first [12 League 12] title . [13 Cairns 13] was an ever present in [14 Rangers 14] ' first [15 League 15] title for five years , and scored 11 goals . [16 Cairns 16] enjoyed further [17 League 17] title successes in , , , , and before joining [18 Bradford City 18] in 1927 . His final [19 Rangers 19] appearance was in a 2 -- 1 win over 3 Queen 's Park on 1 March 1927 . In his time with Bradford , he made 135 appearances and won a [20 Division Three North 20] medal in 1929 . He retired from football in 1932 and later worked as a scout for [13 13 Arsenal 13] . [21 Cairns 21] was also at international level , making eight [22 Scotland 22] appearances . He made his international debut in a [14 14 British Home Championship 14] match against [15 15 Wales 15] on 26 February 1920 and he scored in a 1 -- 1 draw . During his international career , he was never on the losing team , winning six games and drawing two . His final international appearance was on 4 April 1925 in a 2 -- 0 win over [23 4 England 23] at [16 16 Hampden Park 16] . [24 Cairns 24] also represented the

1: 26994 Scotland
2: 4392 Bristol_City_F.C.
3: 2488160 Peebles_Rovers_F.C.
4: 393756 St._Johnstone_F.C.
5: 69713 Rangers_F.C.
6: 431144 Bradford_City_A.F.C.
7: 578101 Scotland_national_football_team
8: 164890 Lanarkshire
9: 560696 Scottish_Junior_Cup
10: 252116 Larkhall_Thistle_F.C.
11: 309502 Hamilton_Academical_F.C.
12: 69708 Celtic_F.C.
13: 2174 Arsenal_F.C.
14: 762761 British_Home_Championship
15: 580291 Wales_national_football_team
16: 544840 Hampden_Park

1: 10568 Association_football
[winner: 16087792 DDR-Liga, pem_pos=6,
final_scores=-1.9465075, final_local_scores=-2.039271,
pem=-2.302466, similarity=-44.245037]
gt: 10568 Association_football, pem_pos=0,
final_scores=-2.0056481, final_local_scores=-2.191236,
pem=0.0, similarity=-61.301136]
2: 421401 Scottish_Junior_Football_Association
[winner: 421401 Scottish_Junior_Football_Association,
pem_pos=5, final_scores=-0.3226114,
final_local_scores=-0.41960785, pem=-2.9652731,
similarity=7.6933947]

gt: 421401 Scottish_Junior_Football_Association,
 pem_pos=5, final_scores=-0.3226114,
 final_local_scores=-0.41960785, pem=-2.9652731,
 similarity=7.6933947]
 3: 426311 Queen's_Park_F.C.
 [winner: 426311 Queen's_Park_F.C., pem_pos=0,
 final_scores=-0.5295102, final_local_scores=-0.67174256,
 pem=-0.7745507, similarity=-11.744298]
 gt: 426311 Queen's_Park_F.C., pem_pos=0,
 final_scores=-0.5295102, final_local_scores=-0.67174256,
 pem=-0.7745507, similarity=-11.744298]
 4: 9904 England_national_football_team
 [winner: 759125 England_cricket_team, pem_pos=2,
 final_scores=0.2716394, final_local_scores=0.30284935,
 pem=-3.3054674, similarity=31.102987]
 gt: 9904 England_national_football_team, pem_pos=1,
 final_scores=0.014592603, final_local_scores=-0.0635587,
 pem=-2.776464, similarity=17.291838]
 1: 26994 Scotland
 [pem_pos=0, final_scores=0.4249333,
 final_local_scores=0.463745, pem=-0.47951394,
 similarity=20.53646]
 2: 4392 Bristol_City_F.C.
 [pem_pos=0, final_scores=0.49699926,
 final_local_scores=0.45461679, pem=0.0, similarity=17.656385]
 3: 2488160 Peebles_Rovers_F.C.
 [pem_pos=0, final_scores=0.23876135,
 final_local_scores=0.21976304, pem=0.0, similarity=10.647881]
 4: 393756 St._Johnstone_F.C.
 [pem_pos=0, final_scores=0.51446736,
 final_local_scores=0.46738362, pem=0.0, similarity=18.037373]
 5: 69713 Rangers_F.C.
 [pem_pos=0, final_scores=0.59513015,
 final_local_scores=0.55676275, pem=-0.30191645,
 similarity=22.346493]
 6: 431144 Bradford_City_A.F.C.
 [pem_pos=0, final_scores=0.4326828,
 final_local_scores=0.38672984, pem=0.0, similarity=15.630503]
 7: 578101 Scotland_national_football_team
 [pem_pos=1, final_scores=0.31828398,
 final_local_scores=0.25374496, pem=-2.6956348,
 similarity=26.32125]
 8: 164890 Lanarkshire
 [pem_pos=0, final_scores=0.12455432,
 final_local_scores=0.1425707, pem=0.0, similarity=8.344307]
 9: 560696 Scottish_Junior_Cup
 [pem_pos=0, final_scores=0.19155066,
 final_local_scores=0.17829677, pem=0.0, similarity=9.410444]
 10: 252116 Larkhall_Thistle_F.C.
 [pem_pos=0, final_scores=0.07901041,
 final_local_scores=0.0738167, pem=0.0, similarity=6.2925506]
 11: 309502 Hamilton_Academical_F.C.

[pem_pos=0, final_scores=0.4076503,
final_local_scores=0.3573674, pem=0.0, similarity=14.754269]
12: 69708 Celtic_F.C.
[pem_pos=0, final_scores=0.6343298,
final_local_scores=0.60706645, pem=-0.9785444,
similarity=27.527262]
13: 2174 Arsenal_F.C.
[pem_pos=0, final_scores=0.5905706,
final_local_scores=0.56975603, pem=-0.15820743,
similarity=21.952726]
14: 762761 British_Home_Championship
[pem_pos=0, final_scores=0.14226355,
final_local_scores=0.11067033, pem=-0.1535597,
similarity=8.227418]
15: 580291 Wales_national_football_team
[pem_pos=1, final_scores=0.150524,
final_local_scores=0.0890151, pem=-2.4526134,
similarity=20.083794]
16: 544840 Hampden_Park
[pem_pos=0, final_scores=0.24727266,
final_local_scores=0.2138642, pem=0.0, similarity=10.471848]
1: 6699437 Tommy_Cairns
[pem_pos=0, final_scores=0.2612397,
final_local_scores=0.24461174, pem=0.0, similarity=11.389415]
2: 6699437 Tommy_Cairns
[pem_pos=0, final_scores=0.40215307,
final_local_scores=0.39853817, pem=0.0, similarity=15.982887]
3: 4392 Bristol_City_F.C.
[pem_pos=0, final_scores=0.35634065,
final_local_scores=0.30096862, pem=0.0, similarity=13.071217]
4: 2488160 Peebles_Rovers_F.C.
[pem_pos=0, final_scores=0.139508,
final_local_scores=0.111343846, pem=0.0, similarity=7.4124355]
5: 393756 St._Johnstone_F.C.
[pem_pos=0, final_scores=0.5715768,
final_local_scores=0.52976704, pem=0.0, similarity=19.899017]
6: 69713 Rangers_F.C.
[pem_pos=0, final_scores=0.6352022,
final_local_scores=0.6005354, pem=-0.30191645,
similarity=23.652756]
7: 69713 Rangers_F.C.
[pem_pos=0, final_scores=0.8130408,
final_local_scores=0.794797, pem=-0.30191645,
similarity=29.44991]
8: 28789929 Football_League_First_Division
[pem_pos=0, final_scores=0.28384697,
final_local_scores=0.23794407, pem=-0.3421056,
similarity=13.050862]
9: 6699437 Tommy_Cairns
[pem_pos=0, final_scores=0.47746426,
final_local_scores=0.48080426, pem=0.0, similarity=18.437872]
10: 69713 Rangers_F.C.

[pem_pos=0, final_scores=0.52941746,
final_local_scores=0.48498154, pem=-0.30191645,
similarity=20.204397]
11: 6699437 Tommy_Cairns
[pem_pos=0, final_scores=0.5224035,
final_local_scores=0.5298935, pem=0.0, similarity=19.902792]
12: 14419588 The_Football_League
[pem_pos=0, final_scores=0.079670265,
final_local_scores=0.007348746, pem=-1.7772934,
similarity=13.974213]
13: 6699437 Tommy_Cairns
[pem_pos=0, final_scores=0.44457287,
final_local_scores=0.44487536, pem=0.0, similarity=17.36568]
14: 69713 Rangers_F.C.
[pem_pos=0, final_scores=0.64472324,
final_local_scores=0.6109356, pem=-0.30191645,
similarity=23.96312]
15: 14419588 The_Football_League
[pem_pos=0, final_scores=0.07865666,
final_local_scores=0.00624156, pem=-1.7772934,
similarity=13.941172]
16: 6699437 Tommy_Cairns
[pem_pos=0, final_scores=0.52994365,
final_local_scores=0.53813, pem=0.0, similarity=20.148584]
17: 14419588 The_Football_League
[pem_pos=0, final_scores=0.11954813,
final_local_scores=0.05090925, pem=-1.7772934,
similarity=15.274145]
18: 431144 Bradford_City_A.F.C.
[pem_pos=0, final_scores=0.34742135,
final_local_scores=0.2935946, pem=0.0, similarity=12.851161]
19: 69713 Rangers_F.C.
[pem_pos=0, final_scores=0.6442631,
final_local_scores=0.6104329, pem=-0.30191645,
similarity=23.948118]
20: 580901 Football_League_Third_Division_North
[pem_pos=0, final_scores=0.45025682,
final_local_scores=0.42469192, pem=0.0, similarity=16.763367]
21: 6699437 Tommy_Cairns
[pem_pos=0, final_scores=0.47072184,
final_local_scores=0.47343916, pem=0.0, similarity=18.218082]
22: 578101 Scotland_national_football_team
[pem_pos=1, final_scores=0.25609297,
final_local_scores=0.18581069, pem=-2.6956348,
similarity=24.293957]
23: 759125 England_cricket_team
[pem_pos=2, final_scores=0.2716394,
final_local_scores=0.30284935, pem=-3.3054674,
similarity=31.102987]
24: 6699437 Tommy_Cairns
[pem_pos=0, final_scores=0.32129365,
final_local_scores=0.3102116, pem=0.0, similarity=13.347046]

Sample: b'Nether_Poppleton_Tithebarn'

Nether Poppleton Tithebarn is a 1 tithe barn at [1 Manor Farm 1] in the village of [1 1 Nether Poppleton 1] in the 2 unitary authority of in the . Research by 3 dendrochronologists has shown that the tithe , which was built on the site of an old nunnery , is at least 450 years old . The building is often referred to as `` [2 Rupert 2] 's Barn '' because [2 2 Prince Rupert of the Rhine 2] is said to have stationed part of his 4 army in the barn before the 5 Battle of [3 Marston Moor 3] in 1644 during the [3 3 English Civil War 3] . There is also an historical connection with the of King [4 4 Charles II 4] because in 1660 , and two hundred [5 5 Yorkshire 5] gentlemen gathered at the barn before marching into [6 6 York 6] for the proclamation of the restored [4 6 King 4] . The area surrounding the barn forms part of a and after nearby 7 Saint 8 Everilda 's 9 parish church the tithe barn is the oldest , largest and most important building in both Poppletons . The Friends of the Poppleton Tithe Barn -LRB- number 1060767 -RRB- was founded in 1997 to preserve and maintain the [5 Tithe Barn 5] for the benefit of the local community and future generations . It was restored in 1999 and completed in the year 2000 . 10 The Duke of York

1: 153606 Nether_Poppleton

2: 153673 Prince_Rupert_of_the_Rhine

3: 9709 English_Civil_War

4: 46688 Charles_II_of_England

5: 36637 Yorkshire

6: 34361 York

1: 147609 Tithe_barn

[winner: 147609 Tithe_barn, pem_pos=0,
final_scores=-0.95582896, final_local_scores=-1.0719409,
pem=-0.009803991, similarity=-27.845825|

gt: 147609 Tithe_barn, pem_pos=0,
final_scores=-0.95582896, final_local_scores=-1.0719409,
pem=-0.009803991, similarity=-27.845825]

2: 53123 Unitary_authority

[winner: 225999 Unitary_operator, pem_pos=2,
final_scores=-0.9561787, final_local_scores=-0.8526456,
pem=-1.2892056, similarity=-14.344032|

gt: 53123 Unitary_authority, pem_pos=3,
final_scores=-1.0910695, final_local_scores=-1.1904778,
pem=-1.2938021, similarity=-24.400623]

[winner: 3677766 Auctoritas, pem_pos=3,
final_scores=-1.1806576, final_local_scores=-1.201787,
pem=-1.0978392, similarity=-25.803787|

RECALL MISS: 53123 Unitary_authority]

3: 37800 Dendrochronology

[winner: 37800 Dendrochronology, pem_pos=0,
final_scores=-1.3247622, final_local_scores=-1.4067526,
pem=0.0, similarity=-37.890587|

gt: 37800 Dendrochronology, pem_pos=0,

final_scores=-1.3247622, final_local_scores=-1.4067526,
pem=0.0, similarity=-37.890587]
4: 1376 Army
[winner: 181149 New_Model_Army, pem_pos=2,
final_scores=-1.2215883, final_local_scores=-1.363776,
pem=-0.6931472, similarity=-32.838642]
RECALL MISS: 1376 Army]
5: 68771 Battle_of_Marston_Moor
[winner: 68771 Battle_of_Marston_Moor, pem_pos=0,
final_scores=0.28335488, final_local_scores=0.26701936,
pem=0.0, similarity=12.058104]
gt: 68771 Battle_of_Marston_Moor, pem_pos=0,
final_scores=0.28335488, final_local_scores=0.26701936,
pem=0.0, similarity=12.058104]
[winner: 17340550 King's_Head_Mill, Battle, pem_pos=2,
final_scores=-0.83726156, final_local_scores=-0.8073802,
pem=-2.7770143, similarity=-4.902288]
RECALL MISS: 68771 Battle_of_Marston_Moor]
6: 19012 Monarch
[winner: 46688 Charles_II_of_England, pem_pos=0,
final_scores=0.12539975, final_local_scores=0.062978506,
pem=0.0, similarity=5.969117]
gt: 19012 Monarch, pem_pos=2, final_scores=-0.28381166,
final_local_scores=-0.2939408, pem=-0.9897082,
similarity=0.7001197]
7: 28436 Saint
[winner: 1361511 GWR_2900_Class, pem_pos=1,
final_scores=-1.1273862, final_local_scores=-1.1280127,
pem=-2.7852771, similarity=-14.425664]
gt: 28436 Saint, pem_pos=0, final_scores=-1.2641286,
final_local_scores=-1.3505676, pem=-0.31810653,
similarity=-34.484]
8: 2245106 Everilda
[winner: 2245106 Everilda, pem_pos=0,
final_scores=-0.56344265, final_local_scores=-0.53529894,
pem=0.0, similarity=-11.884674]
gt: 2245106 Everilda, pem_pos=0,
final_scores=-0.56344265, final_local_scores=-0.53529894,
pem=0.0, similarity=-11.884674]
9: 149937 Parish_church
[winner: 149937 Parish_church, pem_pos=0,
final_scores=-1.0381418, final_local_scores=-1.0945773,
pem=-0.06794557, similarity=-28.205156]
gt: 149937 Parish_church, pem_pos=0,
final_scores=-1.0381418, final_local_scores=-1.0945773,
pem=-0.06794557, similarity=-28.205156]
10: 125267 Prince_Andrew, Duke_of_York
[winner: 34361 York, pem_pos=0, final_scores=-0.2036881,
final_local_scores=-0.30647546, pem=-0.7238267,
similarity=-1.1198438]
RECALL MISS: 125267 Prince_Andrew, Duke_of_York]
[winner: 58255 Duke, pem_pos=0, final_scores=-0.94418603,


```

final_local_scores=-1.0272026, pem=-1.0947025,
similarity=-20.610897]
    RECALL MISS: 125267 Prince_Andrew,_Duke_of_York]
1: 153606 Nether_Poppleton
    [pem_pos=0, final_scores=0.47723895,
final_local_scores=0.51590556, pem=0.0, similarity=19.485363]
2: 153673 Prince_Rupert_of_the_Rhine
    [pem_pos=0, final_scores=0.25444072,
final_local_scores=0.21298209, pem=0.0, similarity=10.445523]
3: 9709 English_Civil_War
    [pem_pos=0, final_scores=0.28075904,
final_local_scores=0.22791457, pem=0.0, similarity=10.891138]
4: 46688 Charles_II_of_England
    [pem_pos=0, final_scores=0.28348243,
final_local_scores=0.22943604, pem=-0.16132857,
similarity=11.813871]
5: 36637 Yorkshire
    [pem_pos=0, final_scores=0.3107416,
final_local_scores=0.2914745, pem=-0.31620604,
similarity=14.50747]
6: 34361 York
    [pem_pos=1, final_scores=0.39838475,
final_local_scores=0.36006474, pem=-0.7238267,
similarity=18.771046]
1: 31185132 Manor_Farm,_Ruislip
    [pem_pos=0, final_scores=0.3878768,
final_local_scores=0.42745233, pem=0.0, similarity=16.845743]
2: 153673 Prince_Rupert_of_the_Rhine
    [pem_pos=0, final_scores=0.13443305,
final_local_scores=0.08189194, pem=0.0, similarity=6.5335317]
3: 68771 Battle_of_Marston_Moor
    [pem_pos=0, final_scores=0.28335488,
final_local_scores=0.26701936, pem=0.0, similarity=12.058104]
4: 46688 Charles_II_of_England
    [pem_pos=0, final_scores=0.12539975,
final_local_scores=0.062978506, pem=0.0, similarity=5.969117]
5: 147609 Tithe_barn
    [pem_pos=0, final_scores=0.22837411,
final_local_scores=0.23619232, pem=0.0, similarity=11.138163]

```

Sample: b'Solemn_League_and_Covenant'

The Solemn League and Covenant was an agreement between the [1 Scottish 1 Covenanter 1] s and the leaders of the English 2 Parliamentarians . It was agreed to in 1643 , during the [1 1 First English Civil War 1] . The leaders of the embattled , faced with the threat of [2 3 Irish 2] Catholic troops joining with the 4 Royalist army , requested the aid of the Scots . The Covenanters promised their aid against the `papists ' , on condition that the [3 Scottish 3] system of church government was adopted in [4 England 4] . This was acceptable to the majority of the [5 English 5] 5 Long Parliament , as many of them were presbyterians , while others

preferred allying with the [6 Scots 6] to losing the [7 Civil War 7] . After some haggling a document called `` The Solemn League and Covenant '' was drawn up . This was in effect a treaty between the [8 English 8] Parliament and [9 Scotland 9] for the preservation of the in [10 Scotland 10] , the reformation of religion in [11 England 11] and [12 Ireland 12] `` according to the word of [13 God 13] and the example of the best reformed churches '' , and the extirpation of and . It did not explicitly mention presbyterianism , and included some ambiguous formulations which left the door open to the , another strong faction on the [14 English Parliamentary 14] side , particularly in the parliamentary armies . It was subscribed to by many in [15 England 15] , [16 Scotland 16] , and [17 Ireland 17] , approved by the [18 English 18] 6 Long Parliament , and , with some slight modifications , by the Westminster . This agreement meant that the Covenanters sent another army south to [19 England 19] to fight on the Parliamentary side in the [20 First English Civil War 20] . Not all those on the parliamentary side were happy with this arrangement and some , like [2 2 John Lilburne 2] , chose to leave the parliamentary armies rather than take the oath prescribed in the Act enforcing the [21 Solemn League and Covenant 21] . After the Royalists had lost the [22 First Civil War 22] , [3 3 Charles I 3] was able to enter into an '' 7 Engagement '' with the majority of the Covenanters in which they agreed to support him in the [4 4 Second English Civil War 4] against their mutual enemy the English Independents , in return for [23 Charles 23] imposing presbyterianism for three years on [24 England 24] . In 1648 the Royalists and the Covenanters were defeated at the 8 Battle of [25 Preston 25] , and [26 Charles 26] was executed in January 1649 . The Scottish Covenants persuaded the exiled [27 9 Charles II 27] of [28 England 28] to agree to the terms of the Solemn League and Covenant in the 10 Treaty of [29 Breda 2 9] -LRB- 1650 -RRB- . However the defeat of the Royalist and [30 Scottish 30] army at the 11 Battle of [31 Worcester 31] in 1651 ended the relevance of the Solemn League and Covenant as the power of the Presbyterians was broken on both sides of the border . After the the [32 English 32] Parliament passed the 12 Sedition Act 1661

1: 1063738 First_English_Civil_War

2: 141624 John_Lilburne

3: 7426 Charles_I_of_England

4: 1057632 Second_English_Civil_War

1: 27238622 Covenanter

[winner: 27238622 Covenanter, pem_pos=0, final_scores=0.09756677, final_local_scores=0.06199135, pem=0.0, similarity=5.9396586]

gt: 27238622 Covenanter, pem_pos=0, final_scores=0.09756677, final_local_scores=0.06199135, pem=0.0, similarity=5.9396586]

2: 69093 Roundhead

[winner: 69093 Roundhead, pem_pos=0,
final_scores=-0.06619592, final_local_scores=-0.122834,
pem=-0.20386027, similarity=1.5327255]
gt: 69093 Roundhead, pem_pos=0,
final_scores=-0.06619592, final_local_scores=-0.122834,
pem=-0.20386027, similarity=1.5327255]
3: 545025 Irish_Catholic
[winner: 14560 Republic_of_Ireland, pem_pos=3,
final_scores=0.15378504, final_local_scores=0.14785692,
pem=-2.2576945, similarity=20.779753]
RECALL MISS: 545025 Irish_Catholic]
[winner: 606848 Catholic_Church, pem_pos=1,
final_scores=-0.4200508, final_local_scores=-0.48968205,
pem=-0.9041281, similarity=-5.606587]
RECALL MISS: 545025 Irish_Catholic]
4: 484563 Cavalier
[winner: 484563 Cavalier, pem_pos=0,
final_scores=-0.18747284, final_local_scores=-0.26502457,
pem=-0.32696095, similarity=-2.0410843]
gt: 484563 Cavalier, pem_pos=0,
final_scores=-0.18747284, final_local_scores=-0.26502457,
pem=-0.32696095, similarity=-2.0410843]
5: 18065 Long_Parliament
[winner: 18065 Long_Parliament, pem_pos=0,
final_scores=-0.04329379, final_local_scores=-0.062037207,
pem=0.0, similarity=2.238399]
gt: 18065 Long_Parliament, pem_pos=0,
final_scores=-0.04329379, final_local_scores=-0.062037207,
pem=0.0, similarity=2.238399]
6: 18065 Long_Parliament
[winner: 18065 Long_Parliament, pem_pos=0,
final_scores=-0.06609531, final_local_scores=-0.08694441,
pem=0.0, similarity=1.4951184]
gt: 18065 Long_Parliament, pem_pos=0,
final_scores=-0.06609531, final_local_scores=-0.08694441,
pem=0.0, similarity=1.4951184]
7: 1775062 Engagers
[winner: 1775062 Engagers, pem_pos=2,
final_scores=-0.31230438, final_local_scores=-0.30026206,
pem=-1.4739476, similarity=3.1448505]
gt: 1775062 Engagers, pem_pos=2,
final_scores=-0.31230438, final_local_scores=-0.30026206,
pem=-1.4739476, similarity=3.1448505]
8: 23789144 Battle_of_Preston_(1648)
[winner: 88918 Preston,_Lancashire, pem_pos=0,
final_scores=0.1927815, final_local_scores=0.15571424,
pem=-1.5624295, similarity=17.23327]
RECALL MISS: 23789144 Battle_of_Preston_(1648)]
[winner: 4659200 Battle_of_"The_Saw", pem_pos=5,
final_scores=-0.9577244, final_local_scores=-0.8890812,
pem=-1.7917595, similarity=-12.698376]
RECALL MISS: 23789144 Battle_of_Preston_(1648)]

9: 46688 Charles_II_of_England
 [winner: 9316 England, pem_pos=3,
 final_scores=0.38669813, final_local_scores=0.33768475,
 pem=-0.19053856, similarity=15.203077|
 RECALL MISS: 46688 Charles_II_of_England]
 [winner: 46688 Charles_II_of_England, pem_pos=0,
 final_scores=0.2073391, final_local_scores=0.14666253,
 pem=-0.16132857, similarity=9.343745|
 gt: 46688 Charles_II_of_England, pem_pos=0,
 final_scores=0.2073391, final_local_scores=0.14666253,
 pem=-0.16132857, similarity=9.343745]

10: 1853137 Treaty_of_Breda_(1650)
 [winner: 1098285 Breda, pem_pos=0,
 final_scores=0.28895628, final_local_scores=0.31266963,
 pem=-0.08950387, similarity=13.907134|
 gt: 1853137 Treaty_of_Breda_(1650), pem_pos=15,
 final_scores=-0.5122402, final_local_scores=-0.5372622,
 pem=-3.2538385, similarity=5.751617]
 [winner: 13343676 1650_in_England, pem_pos=1,
 final_scores=-0.4928751, final_local_scores=-0.34924293,
 pem=-1.3776877, similarity=1.1596876|
 RECALL MISS: 1853137 Treaty_of_Breda_(1650)]
 [winner: 149028 Tribal_Voice, pem_pos=5,
 final_scores=-0.86843294, final_local_scores=-0.77442473,
 pem=-2.0368252, similarity=-7.944093|
 RECALL MISS: 1853137 Treaty_of_Breda_(1650)]

11: 70181 Battle_of_Worcester
 [winner: 58681 Worcester, pem_pos=0,
 final_scores=0.4146924, final_local_scores=0.39105523,
 pem=-0.59997296, similarity=19.022327|
 gt: 70181 Battle_of_Worcester, pem_pos=15,
 final_scores=-0.37057406, final_local_scores=-0.43911192,
 pem=-3.740713, similarity=11.328315]
 [winner: 48965 Battle_of_Cold_Harbor, pem_pos=0,
 final_scores=-0.95654774, final_local_scores=-0.9800062,
 pem=-0.3140309, similarity=-23.447878|
 RECALL MISS: 70181 Battle_of_Worcester]

12: 4470471 Sedition_Act_1661
 [winner: 573930 Sedition_Act_of_1918, pem_pos=0,
 final_scores=-0.18673214, final_local_scores=-0.13341963,
 pem=0.0, similarity=0.10820516|
 gt: 4470471 Sedition_Act_1661, pem_pos=4,
 final_scores=-1.2185433, final_local_scores=-1.2848732,
 pem=-7.044905, similarity=4.05782]
 [winner: 34588 1661, pem_pos=0, final_scores=-0.8346946,
 final_local_scores=-0.72300524, pem=-0.25815487,
 similarity=-16.08232|
 RECALL MISS: 4470471 Sedition_Act_1661]

1: 1063738 First_English_Civil_War
 [pem_pos=0, final_scores=0.19984318,
 final_local_scores=0.16323552, pem=0.0, similarity=8.960986]

2: 141624 John_Lilburne

[pem_pos=0, final_scores=0.1601051,
final_local_scores=0.18072361, pem=0.0, similarity=9.482866]
3: 7426 Charles_I_of_England
[pem_pos=0, final_scores=0.1871997,
final_local_scores=0.12681031, pem=-0.19786721,
similarity=8.950018]
4: 1057632 Second_English_Civil_War
[pem_pos=0, final_scores=0.14094664,
final_local_scores=0.123662114, pem=0.0, similarity=7.7800374]
1: 27238622 Covenanters
[pem_pos=0, final_scores=0.09756677,
final_local_scores=0.06199135, pem=0.0, similarity=5.9396586]
2: 14560 Republic_of_Ireland
[pem_pos=3, final_scores=0.15378504,
final_local_scores=0.14785692, pem=-2.2576945,
similarity=20.779753]
3: 26994 Scotland
[pem_pos=0, final_scores=0.41671056,
final_local_scores=0.38901025, pem=-0.47951394,
similarity=18.306227]
4: 9316 England
[pem_pos=4, final_scores=0.5637391,
final_local_scores=0.5310751, pem=-0.19053856,
similarity=20.974232]
5: 9316 England
[pem_pos=1, final_scores=0.17638133,
final_local_scores=0.10794562, pem=-1.3527626,
similarity=14.667559]
6: 26994 Scotland
[pem_pos=2, final_scores=0.08289899,
final_local_scores=0.024371803, pem=-2.0904498,
similarity=16.185207]
7: 863 American_Civil_War
[pem_pos=0, final_scores=0.18611659,
final_local_scores=0.17097902, pem=-0.48123205,
similarity=11.809082]
8: 9316 England
[pem_pos=1, final_scores=0.08085157,
final_local_scores=0.003593862, pem=-1.3527626,
similarity=11.553494]
9: 26994 Scotland
[pem_pos=0, final_scores=0.49762774,
final_local_scores=0.47739995, pem=-0.22905664,
similarity=19.581924]
10: 26994 Scotland
[pem_pos=0, final_scores=0.5190247,
final_local_scores=0.5007729, pem=-0.22905664,
similarity=20.27942]
11: 9316 England
[pem_pos=2, final_scores=0.46500963,
final_local_scores=0.4232282, pem=-0.19053856,
similarity=17.755865]

12: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.30449277,
final_local_scores=0.31521383, pem=-1.5040284,
similarity=21.675463]
13: 5042765 God
[pem_pos=0, final_scores=0.5305709,
final_local_scores=0.6392477, pem=-0.18657415,
similarity=24.18076]
14: 378033 Parliament_of_England
[pem_pos=0, final_scores=0.038094833,
final_local_scores=0.0071677268, pem=0.0, similarity=4.303612]
15: 9316 England
[pem_pos=4, final_scores=0.5016713,
final_local_scores=0.4625299, pem=-0.19053856,
similarity=18.928705]
16: 26994 Scotland
[pem_pos=0, final_scores=0.52215475,
final_local_scores=0.50419194, pem=-0.22905664,
similarity=20.381453]
17: 14560 Republic_of_Ireland
[pem_pos=3, final_scores=0.33137017,
final_local_scores=0.34457335, pem=-1.5040284,
similarity=22.551609]
18: 9316 England
[pem_pos=1, final_scores=0.20490275,
final_local_scores=0.13910094, pem=-1.3527626,
similarity=15.597296]
19: 9316 England
[pem_pos=3, final_scores=0.5903137,
final_local_scores=0.5601038, pem=-0.19053856,
similarity=21.840506]
20: 1063738 First_English_Civil_War
[pem_pos=0, final_scores=0.22143994,
final_local_scores=0.18682668, pem=0.0, similarity=9.664993]
21: 176294 Solemn_League_and_Covenant
[pem_pos=0, final_scores=0.037752196,
final_local_scores=0.040205985, pem=-0.027919194,
similarity=5.4413686]
22: 1063738 First_English_Civil_War
[pem_pos=0, final_scores=0.2159444,
final_local_scores=0.18349344, pem=-0.13564064,
similarity=10.303158]
23: 46688 Charles_II_of_England
[pem_pos=0, final_scores=0.2653963,
final_local_scores=0.2148315, pem=0.0, similarity=10.500713]
24: 9316 England
[pem_pos=4, final_scores=0.53705794,
final_local_scores=0.50193, pem=-0.19053856,
similarity=20.104483]
25: 88918 Preston,_Lancashire
[pem_pos=0, final_scores=0.1927815,
final_local_scores=0.15571424, pem=-1.5624295,

similarity=17.23327]
 26: 46688 Charles_II_of_England
 [pem_pos=0, final_scores=0.24664746,
 final_local_scores=0.19435132, pem=0.0, similarity=9.889544]
 27: 46688 Charles_II_of_England
 [pem_pos=0, final_scores=0.2073391,
 final_local_scores=0.14666253, pem=-0.16132857,
 similarity=9.343745]
 28: 9316 England
 [pem_pos=3, final_scores=0.38669813,
 final_local_scores=0.33768475, pem=-0.19053856,
 similarity=15.203077]
 29: 1098285 Breda
 [pem_pos=0, final_scores=0.28895628,
 final_local_scores=0.31266963, pem=-0.08950387,
 similarity=13.907134]
 30: 26994 Scotland
 [pem_pos=0, final_scores=0.4118055,
 final_local_scores=0.3836522, pem=-0.47951394,
 similarity=18.146332]
 31: 58681 Worcester
 [pem_pos=0, final_scores=0.4146924,
 final_local_scores=0.39105523, pem=-0.59997296,
 similarity=19.022327]
 32: 9316 England
 [pem_pos=1, final_scores=0.22796462,
 final_local_scores=0.16429254, pem=-1.3527626,
 similarity=16.349064]

Sample: b'Bioneers'

Bioneers , under its parent foundation , [1 Collective
 Heritage Institute 1] , is a [1 1 New Mexico 1] - based
 nonprofit organization promoting practical and innovative
 solutions to global 1 environmental and bio-cultural
 challenges . The organization operates within a philosophy
 that recognizes and cultivates the value and wisdom of the
 natural world . Official Programs include Moonrise Women 's
 Leadership , Restorative Food Systems , -LRB- Indigenous
 Forums , -RRB- Education for Action , and the award-winning
 Dreaming New Mexico community resilience program . Bioneers
 is also widely recognized for producing innovative media
 covering subjects such as environmentalism , rights of
 nature , social justice , sustainability and permaculture .
 Bioneers Radio is broadcast on local radio stations across
 the [2 U.S. 2] , as well as having segments featured on
 national [3 NPR 3] stations . The organization also produces
 the annual National Bioneers Conference each October in [2 2 S
 an Rafael 2] , [3 3 California 3] . The history of National
 Bioneers Conference presenters includes [4 4 Michael Pollan 4]
 , Dr. [5 5 Andrew Weil 5] , [6 6 Gloria Steinem 6] , [7 7 Jane
 Goodall 7] , [8 8 Philippe Cousteau 8] , [9 9 Eve Ensler 9] ,
 [10 10 Bill McKibben 10] , [11 11 Paul Hawken 11] , and

more . Plenary -LRB- Keynote -RRB- sessions from the national conference are also webcast to Beaming Bioneers satellite conferences held simultaneously in various locations throughout the [12 12 United States 12] and [13 13 Canada 13] . Bioneers | Collective Heritage Institute was founded in 1990 . 2 Bioneer -LRB- root : `` biological pioneer '' -RRB- is a 3 neologism coined by founder [14 14 Kenny Ausubel 14] . It describes individuals and groups working in diverse disciplines who have crafted creative solutions to various 4 environmental and socio-cultural problems rooted in shared core values , including , -LRB- -RRB- thinking , a view of all life as interdependent , and sustainable mutual aid . The first National Bioneers Conference , organized by Co-Founders [15 15 Kenny Ausubel 15] and [16 16 Nina Simons 16] , took place in 1990 . Currently the conference takes place annually in [4 5 San Rafael 4] , [5 California 5] , on the third weekend in October . The national conference highlights the work of scientific and social innovators and helps support , nurture and propagate their ideas and models . Conference speakers come from fields : 6 environmental and socio-political activism ; `` green '' 7 biology , 8 chemistry , design , 9 architecture and 10 urban planning ; 11 organic and `` beyond organic '' farming and gardening ; 12 indigenous perspectives ; 13 biodiversity , 14 bioremediation , and 15 wildland preservation ; 16 alternative energy ; engaged spirituality , literature and the arts ; 17 holistic and `` ecological '' medicine ; 18 ethnobotany ; 19 socially responsible entrepreneurship , business and philanthropy ; the 20 environmental justice , women 's and youth movements ; 21 independent media ; etc. . In many cases the technological or social solutions to problems showcased are founded on emulation of natural

1: 21649 New_Mexico
2: 107733 San_Rafael,_California
3: 5407 California
4: 1875303 Michael_Pollan
5: 1178580 Andrew_Weil
6: 299078 Gloria_Steinem
7: 45397 Jane_Goodall
8: 3894042 Philippe_Cousteau
9: 967570 Eve_Ensler
10: 367037 Bill_McKibben
11: 1230468 Paul_Hawken
12: 3434750 United_States
13: 5042916 Canada
14: 38524267 Kenny_Ausubel
15: 38524267 Kenny_Ausubel
16: 38678288 Nina_Simons

1: 558685 Natural_environment
[winner: 2364800 Environmental_impact_assessment,
pem_pos=0, final_scores=-1.0914677,
final_local_scores=-1.1231949, pem=0.0,

similarity=-29.428663|
RECALL MISS: 558685 Natural_environment]

2: 10165060 Bioneer
[winner: 10165060 Bioneer, pem_pos=0,
final_scores=-0.27126014, final_local_scores=-0.22697607,
pem=0.0, similarity=-2.6837058|
gt: 10165060 Bioneer, pem_pos=0,
final_scores=-0.27126014, final_local_scores=-0.22697607,
pem=0.0, similarity=-2.6837058]

3: 153907 Neologism
[winner: 12891693 Neonism, pem_pos=0,
final_scores=-0.82444316, final_local_scores=-0.75807273,
pem=-0.12340098, similarity=-17.861616|
gt: 153907 Neologism, pem_pos=1,
final_scores=-1.4019272, final_local_scores=-1.4776413,
pem=-2.1950552, similarity=-28.069002]

4: 9632 Ecosystem
[winner: 2364800 Environmental_impact_assessment,
pem_pos=0, final_scores=-1.2073736,
final_local_scores=-1.2498045, pem=0.0,
similarity=-33.206947|
RECALL MISS: 9632 Ecosystem]

5: 107733 San_Rafael,_California
[winner: 5407 California, pem_pos=0,
final_scores=0.59321535, final_local_scores=0.6059026,
pem=-0.08662026, similarity=22.64211|
RECALL MISS: 107733 San_Rafael,_California]
[winner: 107733 San_Rafael,_California, pem_pos=0,
final_scores=0.37369007, final_local_scores=0.36958873,
pem=-0.319724, similarity=16.857687|
gt: 107733 San_Rafael,_California, pem_pos=0,
final_scores=0.37369007, final_local_scores=0.36958873,
pem=-0.319724, similarity=16.857687]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8176335, final_local_scores=-1.7979851,
pem=-4.6728287, similarity=-24.154175|
RECALL MISS: 107733 San_Rafael,_California]

6: 1866009 Environmentalism
[winner: 2364800 Environmental_impact_assessment,
pem_pos=0, final_scores=-1.1137514,
final_local_scores=-1.1475365, pem=0.0,
similarity=-30.155066|
RECALL MISS: 1866009 Environmentalism]

7: 9127632 Biology
[winner: 9202982 Biolab, pem_pos=0,
final_scores=-0.9378878, final_local_scores=-0.84629786,
pem=0.0, similarity=-21.165503|
gt: 9127632 Biology, pem_pos=1, final_scores=-1.4032066,
final_local_scores=-1.5166395, pem=-0.043245334,
similarity=-40.93466]

8: 5180 Chemistry
[winner: 1052154 Interpersonal_attraction, pem_pos=2,

```

final_scores=-1.3418152, final_local_scores=-1.3639123,
pem=-0.69172484, similarity=-32.850445|
    gt: 5180 Chemistry, pem_pos=0, final_scores=-1.5172626,
final_local_scores=-1.6190696, pem=-0.051642966,
similarity=-43.94571 ]
9: 21296224 Architecture
    [winner: 21296224 Architecture, pem_pos=0,
final_scores=-1.2773949, final_local_scores=-1.3761864,
pem=-0.33441302, similarity=-35.159843|
    gt: 21296224 Architecture, pem_pos=0,
final_scores=-1.2773949, final_local_scores=-1.3761864,
pem=-0.33441302, similarity=-35.159843 ]
10: 45346 Urban_planning
    [winner: 8586385 Systematization_(Romania), pem_pos=0,
final_scores=-0.931311, final_local_scores=-0.8697267,
pem=0.0, similarity=-21.864668|
    gt: 45346 Urban_planning, pem_pos=1,
final_scores=-1.197122, final_local_scores=-1.3036406,
pem=-0.013276101, similarity=-34.74132 ]
11: 72754 Organic_farming
    [winner: 13153591 Soil_organic_matter, pem_pos=11,
final_scores=-1.0929589, final_local_scores=-1.0359364,
pem=-2.1917012, similarity=-14.905886|
    gt: 72754 Organic_farming, pem_pos=2,
final_scores=-1.1686982, final_local_scores=-1.2696872,
pem=-1.6796521, similarity=-24.666079 ]
    [winner: 13153591 Soil_organic_matter, pem_pos=11,
final_scores=-1.1667203, final_local_scores=-1.1165094,
pem=-2.1917012, similarity=-17.31034|
    gt: 72754 Organic_farming, pem_pos=2,
final_scores=-1.2640723, final_local_scores=-1.373869,
pem=-1.6796521, similarity=-27.775068 ]
    [winner: 9288716 Beyond_(Dinosaur_Jr._album), pem_pos=1,
final_scores=-1.1824864, final_local_scores=-1.1708295,
pem=-1.6161708, similarity=-22.061192|
    RECALL MISS: 72754 Organic_farming]
    [winner: 627 Agriculture, pem_pos=0,
final_scores=-1.3315246, final_local_scores=-1.4219463,
pem=-0.115239374, similarity=-37.717308|
    gt: 72754 Organic_farming, pem_pos=7,
final_scores=-2.1736038, final_local_scores=-2.367394,
pem=-5.7964153, similarity=-35.036236 ]
    [winner: 59349 Quotation_mark, pem_pos=0,
final_scores=-1.7918457, final_local_scores=-1.806865,
pem=0.0, similarity=-49.83074|
    RECALL MISS: 72754 Organic_farming]
12: 45281 Indigenous_peoples
    [winner: 45281 Indigenous_peoples, pem_pos=0,
final_scores=-1.2801793, final_local_scores=-1.3747299,
pem=-0.9612204, similarity=-31.707703|
    gt: 45281 Indigenous_peoples, pem_pos=0,
final_scores=-1.2801793, final_local_scores=-1.3747299,

```

pem=-0.9612204, similarity=-31.707703]

13: 45086 Biodiversity
 [winner: 8621972 Guatemala_Biodiversity, pem_pos=0,
 final_scores=-0.9947976, final_local_scores=-0.9363758,
 pem=0.0, similarity=-23.85361|
 gt: 45086 Biodiversity, pem_pos=1,
 final_scores=-1.323668, final_local_scores=-1.4197438,
 pem=-0.021447008, similarity=-38.161644]

14: 434188 Bioremediation
 [winner: 434188 Bioremediation, pem_pos=0,
 final_scores=-1.0514591, final_local_scores=-1.0495614,
 pem=-0.0056022275, similarity=-27.200827|
 gt: 434188 Bioremediation, pem_pos=0,
 final_scores=-1.0514591, final_local_scores=-1.0495614,
 pem=-0.0056022275, similarity=-27.200827]

15: 113261 Nature_reserve
 [winner: 396275 Wilderness, pem_pos=0,
 final_scores=-1.1667595, final_local_scores=-1.2584922,
 pem=-0.05837198, similarity=-33.14877|
 RECALL MISS: 113261 Nature_reserve]
 [winner: 1404387 Digital_preservation, pem_pos=0,
 final_scores=-1.2571907, final_local_scores=-1.2848997,
 pem=-0.6534256, similarity=-30.700825|
 RECALL MISS: 113261 Nature_reserve]

16: 193939 Alternative_energy
 [winner: 193939 Alternative_energy, pem_pos=0,
 final_scores=-0.91048545, final_local_scores=-0.98656696,
 pem=-0.033738136, similarity=-25.16794|
 gt: 193939 Alternative_energy, pem_pos=0,
 final_scores=-0.91048545, final_local_scores=-0.98656696,
 pem=-0.033738136, similarity=-25.16794]

17: 337667 Holistic_health
 [winner: 337667 Holistic_health, pem_pos=0,
 final_scores=-0.9245696, final_local_scores=-0.98564553,
 pem=0.0, similarity=-25.323915|
 gt: 337667 Holistic_health, pem_pos=0,
 final_scores=-0.9245696, final_local_scores=-0.98564553,
 pem=0.0, similarity=-25.323915]

18: 571941 Ethnobotany
 [winner: 571941 Ethnobotany, pem_pos=0,
 final_scores=-1.0539304, final_local_scores=-1.139864,
 pem=0.0, similarity=-29.9261|
 gt: 571941 Ethnobotany, pem_pos=0,
 final_scores=-1.0539304, final_local_scores=-1.139864,
 pem=0.0, similarity=-29.9261]

19: 398356 Corporate_social_responsibility
 [winner: 1159092 Social_responsibility, pem_pos=0,
 final_scores=-1.0033914, final_local_scores=-1.0690244,
 pem=0.0, similarity=-27.812109|
 gt: 398356 Corporate_social_responsibility, pem_pos=1,
 final_scores=-1.4569025, final_local_scores=-1.5687637,
 pem=-2.2583368, similarity=-30.44414]

20: 1860870 Environmental_justice
[winner: 1860870 Environmental_justice, pem_pos=0,
final_scores=-1.0407917, final_local_scores=-1.123838,
pem=-0.02023991, similarity=-29.337784]
gt: 1860870 Environmental_justice, pem_pos=0,
final_scores=-1.0407917, final_local_scores=-1.123838,
pem=-0.02023991, similarity=-29.337784]
21: 3313367 Independent_media
[winner: 673184 Alternative_media, pem_pos=0,
final_scores=-0.8441063, final_local_scores=-0.9198343,
pem=0.0, similarity=-23.359978]
gt: 3313367 Independent_media, pem_pos=3,
final_scores=-1.1026937, final_local_scores=-1.1417627,
pem=-2.7880929, similarity=-14.8206835]
1: 21649 New_Mexico
[pem_pos=0, final_scores=0.423643,
final_local_scores=0.4597935, pem=-0.20017311,
similarity=18.89944]
2: 107733 San_Rafael,_California
[pem_pos=0, final_scores=0.376082,
final_local_scores=0.3722015, pem=-0.319724,
similarity=16.935658]
3: 5407 California
[pem_pos=0, final_scores=0.6543167,
final_local_scores=0.6726466, pem=-0.08662026,
similarity=24.633884]
4: 1875303 Michael_Pollan
[pem_pos=0, final_scores=0.3078596,
final_local_scores=0.36613166, pem=0.0, similarity=15.015812]
5: 1178580 Andrew_Weil
[pem_pos=0, final_scores=0.28974807,
final_local_scores=0.3564356, pem=0.0, similarity=14.726462]
6: 299078 Gloria_Steinem
[pem_pos=0, final_scores=0.26690906,
final_local_scores=0.28483176, pem=0.0, similarity=12.589661]
7: 45397 Jane_Goodall
[pem_pos=0, final_scores=0.32611936,
final_local_scores=0.34806484, pem=0.0, similarity=14.476662]
8: 3894042 Philippe_Cousteau
[pem_pos=0, final_scores=0.19252168,
final_local_scores=0.28161073, pem=0.0, similarity=12.493539]
9: 967570 Eve_Ensler
[pem_pos=0, final_scores=0.26682824,
final_local_scores=0.29303935, pem=0.0, similarity=12.834592]
10: 367037 Bill_McKibben
[pem_pos=0, final_scores=0.27909815,
final_local_scores=0.29028416, pem=0.0, similarity=12.752371]
11: 1230468 Paul_Hawken
[pem_pos=0, final_scores=0.16474698,
final_local_scores=0.18779781, pem=0.0, similarity=9.6939745]
12: 3434750 United_States
[pem_pos=0, final_scores=0.4648857,

```

final_local_scores=0.47347116, pem=-0.073457584,
similarity=18.618511]
13: 5042916 Canada
    [pem_pos=0, final_scores=0.6110285,
final_local_scores=0.66415673, pem=-0.11662793,
similarity=24.543715]
14: 38524267 Kenny_Ausubel
    [pem_pos=0, final_scores=0.33135122,
final_local_scores=0.3530307, pem=0.0, similarity=14.624853]
15: 38524267 Kenny_Ausubel
    [pem_pos=0, final_scores=0.23232664,
final_local_scores=0.24486142, pem=0.0, similarity=11.396866]
16: 38678288 Nina_Simons
    [pem_pos=0, final_scores=0.051876828,
final_local_scores=0.092019826, pem=0.0, similarity=6.835768]
1: 320021 Bioneers
    [pem_pos=0, final_scores=0.16808538,
final_local_scores=0.1843574, pem=0.0, similarity=9.591306]
2: 3434750 United_States
    [pem_pos=0, final_scores=0.45246083,
final_local_scores=0.4598989, pem=-0.14075686,
similarity=18.57947]
3: 29697232 NPR
    [pem_pos=0, final_scores=0.10934372,
final_local_scores=0.09293945, pem=0.0, similarity=6.8632116]
4: 107733 San_Rafael,_California
    [pem_pos=0, final_scores=0.37369007,
final_local_scores=0.36958873, pem=-0.319724,
similarity=16.857687]
5: 5407 California
    [pem_pos=0, final_scores=0.59321535,
final_local_scores=0.6059026, pem=-0.08662026,
similarity=22.64211]

```

Sample: b'Ficoll'

[1 Ficoll 1] is a neutral , highly branched , high-mass , 1 polysaccharide which dissolves readily in 2 aqueous 3 solution s. [2 Ficoll 2] radii range from 2-7 nm . It is prepared by reaction of the polysaccharide with 4 epichlorohydrin . Ficoll is a registered trademark owned by [3 GE Healthcare 3] companies . Ficoll is part of Ficoll-Paque which is used in 5 biology 6 laboratories to separate 7 blood to its components - LRB- s , s etc -RRB- . Ficoll-Paque is normally placed at the bottom of a conical tube , and blood is then slowly layered above Ficoll-Paque . After being 8 centrifuge d , the following layers will be visible in the conical tube , from top to bottom : plasma and other constituents , a layer of mono-nuclear cells called 9 buffy coat -LRB- / MNC -RRB- , Ficoll-Paque , and erythrocytes & which should be present in pellet form . This separation allows easy harvest of PBMC 's . Note that some 10 red blood cell trapping -LRB- presence of erythrocytes & granulocytes -RRB- may occur in the PBMC or

Ficoll-Paque layer . Major may sometimes occur in the PBMC layer . Ethylene diamine tetra-acetate -LRB- -RRB- and heparin are commonly used in conjunction with Ficoll-Paque -LRB- TM -RRB- to prevent clotting . At one time , fairly early in the research for diabetes treatment by transplantation , [4 Ficoll 4] was used in an attempt to separate 11 islets of Langerhans from enzyme -LRB- collagenase -RRB- digested human pancreatic tissue . The theory was that separated islets could then be used for transplantation into patients with . In practice , using [5 Ficoll 5] led to difficulty obtaining a pure enough preparation of islets , with enough islet recovery from the original tissue , for safe and effective islet transplantation in humans . Because Ficoll-Paque layering is a very slow process , devices that aid in the most time-consuming step , the overlay , have been developed . One such a product is SepMate-50 , a specialized tube containing a porous insert that forms a physical barrier between the Ficoll-Paque and blood sample . This allows the blood sample to be rapidly pipetted onto the insert , avoiding the need for overlaying it directly onto Ficoll-Paque . The SepMate insert also reduces the duration of the centrifugation step , and after centrifugation , the top layer containing plasma and PBMCs can be poured into a new tube . Other devices include a column containing a porous high-density 12 polyethylene barrier or `` frit . '' These products allow blood to be layered on much more quickly without mixing polysaccharide and blood . An example of such a product is the `` Accuspin System Histopaque-1077 '' sold by [6 Sigma Aldrich 6] . The cost of this product is significantly higher than basic [7 Ficoll 7] and Ficoll-Paque products . It is also possible to have the Ficoll-Paque separating system included in the 13 Vacutainer

1: 23978 Polysaccharide

[winner: 23978 Polysaccharide, pem_pos=0, final_scores=-1.0420513, final_local_scores=-1.1205004, pem=-0.002124837, similarity=-29.3367|

gt: 23978 Polysaccharide, pem_pos=0, final_scores=-1.0420513, final_local_scores=-1.1205004, pem=-0.002124837, similarity=-29.3367]

2: 33306 Water

[winner: 106240 Aqueous_solution, pem_pos=0, final_scores=-0.903479, final_local_scores=-0.95987886, pem=-0.0022329567, similarity=-24.542843|

gt: 33306 Water, pem_pos=1, final_scores=-2.2643695, final_local_scores=-2.423231, pem=-6.511002, similarity=-32.81649]

3: 28729 Solution

[winner: 106240 Aqueous_solution, pem_pos=0, final_scores=-0.903479, final_local_scores=-0.95987886, pem=-0.0022329567, similarity=-24.542843|

RECALL MISS: 28729 Solution]

4: 5233716 Epichlorohydrin
[winner: 5233716 Epichlorohydrin, pem_pos=0,
final_scores=-0.64452827, final_local_scores=-0.6312989,
pem=0.0, similarity=-14.749505|
gt: 5233716 Epichlorohydrin, pem_pos=0,
final_scores=-0.64452827, final_local_scores=-0.6312989,
pem=0.0, similarity=-14.749505]

5: 9127632 Biology
[winner: 9202982 Biolab, pem_pos=0,
final_scores=-0.9564054, final_local_scores=-0.94245124,
pem=0.0, similarity=-24.034912|
gt: 9127632 Biology, pem_pos=1, final_scores=-1.5643202,
final_local_scores=-1.6257868, pem=-0.043245334,
similarity=-44.19183]

6: 173351 Laboratory
[winner: 173351 Laboratory, pem_pos=1,
final_scores=-1.278602, final_local_scores=-1.3295491,
pem=-0.030312207, similarity=-35.42184|
gt: 173351 Laboratory, pem_pos=1,
final_scores=-1.278602, final_local_scores=-1.3295491,
pem=-0.030312207, similarity=-35.42184]

7: 3997 Blood
[winner: 3997 Blood, pem_pos=0, final_scores=-1.1028166,
final_local_scores=-1.1842692, pem=-0.08232216,
similarity=-30.803562|
gt: 3997 Blood, pem_pos=0, final_scores=-1.1028166,
final_local_scores=-1.1842692, pem=-0.08232216,
similarity=-30.803562]

8: 106284 Centrifuge
[winner: 12285825 Laboratory_centrifuge, pem_pos=0,
final_scores=-0.8035272, final_local_scores=-0.8030416,
pem=0.0, similarity=-19.87465|
gt: 106284 Centrifuge, pem_pos=1,
final_scores=-1.2389228, final_local_scores=-1.3063557,
pem=-0.025683608, similarity=-34.754875]

9: 3832048 Buffy_coat
[winner: 3832048 Buffy_coat, pem_pos=0,
final_scores=-0.8920689, final_local_scores=-0.8907425,
pem=0.0, similarity=-22.49182|
gt: 3832048 Buffy_coat, pem_pos=0,
final_scores=-0.8920689, final_local_scores=-0.8907425,
pem=0.0, similarity=-22.49182]

10: 67158 Red_blood_cell
[winner: 67158 Red_blood_cell, pem_pos=0,
final_scores=-0.6734057, final_local_scores=-0.72944665,
pem=0.0, similarity=-17.67843|
gt: 67158 Red_blood_cell, pem_pos=0,
final_scores=-0.6734057, final_local_scores=-0.72944665,
pem=0.0, similarity=-17.67843]
[winner: 4230 Cell_(biology), pem_pos=0,
final_scores=-1.0689969, final_local_scores=-1.1606648,
pem=-0.15427125, similarity=-29.707886|

RECALL MISS: 67158 Red_blood_cell]
 11: 199453 Islets_of_Langerhans
 [winner: 1552139 Langerhans_cell, pem_pos=1,
 final_scores=-0.07700272, final_local_scores=0.03686686,
 pem=-0.44628713, similarity=7.6168723]
 RECALL MISS: 199453 Islets_of_Langerhans]
 [winner: 199453 Islets_of_Langerhans, pem_pos=0,
 final_scores=-1.2925872, final_local_scores=-1.3679636,
 pem=0.0, similarity=-36.733047]
 gt: 199453 Islets_of_Langerhans, pem_pos=0,
 final_scores=-1.2925872, final_local_scores=-1.3679636,
 pem=0.0, similarity=-36.733047]
 12: 77385 Polyethylene
 [winner: 1911597 High-density_polyethylene, pem_pos=0,
 final_scores=-0.89717424, final_local_scores=-0.9217685,
 pem=0.0, similarity=-23.417698]
 gt: 77385 Polyethylene, pem_pos=1,
 final_scores=-0.91922784, final_local_scores=-0.95699775,
 pem=-0.017064385, similarity=-24.37621]
 13: 250556 Vacutainer
 [winner: 250556 Vacutainer, pem_pos=0,
 final_scores=-0.35064545, final_local_scores=-0.3212666,
 pem=0.0, similarity=-5.497523]
 gt: 250556 Vacutainer, pem_pos=0,
 final_scores=-0.35064545, final_local_scores=-0.3212666,
 pem=0.0, similarity=-5.497523]
 1: 1414700 Ficoll
 [pem_pos=0, final_scores=0.116263226,
 final_local_scores=0.027099282, pem=0.0, similarity=4.8984094]
 2: 1414700 Ficoll
 [pem_pos=0, final_scores=0.12406553,
 final_local_scores=0.035622135, pem=0.0, similarity=5.1527486]
 3: 877606 GE_Healthcare
 [pem_pos=0, final_scores=0.12274964,
 final_local_scores=0.27543026, pem=0.0, similarity=12.309101]
 4: 1414700 Ficoll
 [pem_pos=0, final_scores=0.10678284,
 final_local_scores=0.016743392, pem=0.0, similarity=4.5893693]
 5: 1414700 Ficoll
 [pem_pos=0, final_scores=0.14897116,
 final_local_scores=0.06282772, pem=0.0, similarity=5.9646173]
 6: 2887549 Sigma-Aldrich
 [pem_pos=0, final_scores=0.26022387,
 final_local_scores=0.3752669, pem=0.0, similarity=15.288426]
 7: 1414700 Ficoll
 [pem_pos=0, final_scores=0.09851684,
 final_local_scores=0.007714063, pem=0.0, similarity=4.319916]

Sample: b'James_Motluk'

James Motluk -LRB- born 26 April 1964 in [1 1 Hamilton 1] , [1 1 Canada 1] -RRB- is a [2 Canadian 2] filmmaker of [3 2 Ukrainian 3] descent . After studying philosophy at [2 2 Trent Un

iversity 2] he travelled to [3 3 Toronto 3] where he struggled to break into the film industry working for the [4 4 Canadian Broadcasting Corporation 4] as an assistant director on a television show called 3 Seeing Things . In 1990 he wrote and directed his first feature film called 4 Nasty Burgers which was released to theaters in 1993 . Nasty Burgers gained him entry into the [5 5 Writers Guild of Canada 5] . For much of the 1990s he worked as a writer for television . His credits included a season on the hit show 5 Road to [4 Av onlea 4] . In 2000 he released a documentary funded by [5 American 5] film director [6 6 Michael Moore 6] called 6 Life Under Mike which took a critical and biased look at then [7 7 Mike Harris 7] . In 2001 the film earned him a Media Human Rights Award from the [6 League for Human Rights 6] of . He released a second documentary to theaters in 2003 called 7 Wh ose University Is It ? . This work used [7 Trent University 7] as a case study and argued that any corporate connections with higher learning are negative - though notably ignoring that the entire [8 Trent University 8] campus was donated by , and its library by the 8 Bata Shoes Corporation . Most recently he released a documentary called 9 Jajo 's Secret which revealed the internment of [9 Ukrainians 9] by the [10 Canadian 10] government during [11 World War One 11] . This movie was broadcast on [12 OMNI TV 12] in [13 Canada 13] and screened in [14 New York 14] at [15 Columbia University 15] in 2011 . His movies tend to sympathize with the left , presenting an anti-capital view of the world , and encouraging grass roots social change . He produces independently through the [16 Toronto 16] based production company . He is currently listed as a member of

1: 5042916 Canada

2: 261859 Trent_University

3: 64646 Toronto

4: 66730 Canadian_Broadcasting_Corporation

5: 3798387 Writers_Guild_of_Canada

6: 74461 Michael_Moore

7: 293593 Mike_Harris

1: 14288 Hamilton,_Ontario

[winner: 86654 Hamilton,_New_Zealand, pem_pos=1, final_scores=0.1126764, final_local_scores=0.11826292, pem=-2.40051, similarity=20.673262]

gt: 14288 Hamilton,_Ontario, pem_pos=0, final_scores=0.016445503, final_local_scores=-0.047254264, pem=-1.8266565, similarity=12.6131935]

2: 1015202 Ukrainian_Canadian

[winner: 31750 Ukraine, pem_pos=1, final_scores=0.49932617, final_local_scores=0.539629, pem=-1.0458583, similarity=25.880857]

gt: 1015202 Ukrainian_Canadian, pem_pos=3, final_scores=-0.77620703, final_local_scores=-0.859714, pem=-2.862655, similarity=-5.998304]

3: 869569 Seeing_Things_(TV_series)

[winner: 869569 Seeing_Things_(TV_series), pem_pos=1,
final_scores=-0.12892261, final_local_scores=-0.06685208,
pem=-0.3919849, similarity=4.2263885]
gt: 869569 Seeing_Things_(TV_series), pem_pos=1,
final_scores=-0.12892261, final_local_scores=-0.06685208,
pem=-0.3919849, similarity=4.2263885]
4: 4667021 Nasty_Burgers
[winner: 4667021 Nasty_Burgers, pem_pos=0,
final_scores=-0.20453893, final_local_scores=-0.0636552,
pem=0.0, similarity=2.190115]
gt: 4667021 Nasty_Burgers, pem_pos=0,
final_scores=-0.20453893, final_local_scores=-0.0636552,
pem=0.0, similarity=2.190115]
5: 455108 Road_to_Avonlea
[winner: 6800475 Avonlea, Saskatchewan, pem_pos=0,
final_scores=0.07577647, final_local_scores=0.17564029,
pem=0.0, similarity=9.331169]
gt: 455108 Road_to_Avonlea, pem_pos=1,
final_scores=-0.22746423, final_local_scores=-0.19731459,
pem=-1.1986957, similarity=4.720145]
[winner: 3649389 Road_to_..., pem_pos=0,
final_scores=-0.92238826, final_local_scores=-0.8884292,
pem=0.0, similarity=-22.422787]
RECALL MISS: 455108 Road_to_Avonlea]
6: 5304805 Life_Under_Mike
[winner: 5304805 Life_Under_Mike, pem_pos=0,
final_scores=-0.07339877, final_local_scores=-0.01462917,
pem=0.0, similarity=3.6531494]
gt: 5304805 Life_Under_Mike, pem_pos=0,
final_scores=-0.07339877, final_local_scores=-0.01462917,
pem=0.0, similarity=3.6531494]
7: 37631352 Whose_University_Is_It?
[winner: 37631352 Whose_University_Is_It?, pem_pos=0,
final_scores=-0.009571746, final_local_scores=0.0597239,
pem=0.0, similarity=5.871993]
gt: 37631352 Whose_University_Is_It?, pem_pos=0,
final_scores=-0.009571746, final_local_scores=0.0597239,
pem=0.0, similarity=5.871993]
8: 264855 Bata_Shoes
[winner: 27238019 Bata_(village), pem_pos=7,
final_scores=-1.1563206, final_local_scores=-1.0319773,
pem=-2.795813, similarity=-11.502481]
gt: 264855 Bata_Shoes, pem_pos=1,
final_scores=-1.594181, final_local_scores=-1.6851997,
pem=-1.2842451, similarity=-39.216084]
[winner: 10794777 Shoes_(album), pem_pos=5,
final_scores=-1.6592796, final_local_scores=-1.640106,
pem=-2.6419954, similarity=-30.486736]
gt: 264855 Bata_Shoes, pem_pos=10,
final_scores=-3.0834484, final_local_scores=-3.3119981,
pem=-2.7080503, similarity=-80.02012]
9: 36094349 Jajo's_Secret

```
[winner: 36094349 Jajo's_Secret, pem_pos=0,
final_scores=-0.40067604, final_local_scores=-0.34230733,
pem=0.0, similarity=-6.125421]
gt: 36094349 Jajo's_Secret, pem_pos=0,
final_scores=-0.40067604, final_local_scores=-0.34230733,
pem=0.0, similarity=-6.125421 ]
1: 5042916 Canada
[pem_pos=0, final_scores=0.63471586,
final_local_scores=0.6123214, pem=-0.11662793,
similarity=22.996845]
2: 261859 Trent_University
[pem_pos=0, final_scores=0.3093158,
final_local_scores=0.28868267, pem=0.0, similarity=12.704579]
3: 64646 Toronto
[pem_pos=0, final_scores=0.5807138,
final_local_scores=0.54586357, pem=-0.12291032,
similarity=21.047777]
4: 66730 Canadian_Broadcasting_Corporation
[pem_pos=0, final_scores=0.39661634,
final_local_scores=0.37729728, pem=0.0, similarity=15.349017]
5: 3798387 Writers_Guild_of_Canada
[pem_pos=0, final_scores=0.179547,
final_local_scores=0.22581616, pem=0.0, similarity=10.828518]
6: 74461 Michael_Moore
[pem_pos=0, final_scores=0.19882493,
final_local_scores=0.22322306, pem=-0.010127032,
similarity=10.806207]
7: 293593 Mike_Harris
[pem_pos=0, final_scores=0.2614631,
final_local_scores=0.29097664, pem=0.0, similarity=12.773036]
1: 86654 Hamilton,_New_Zealand
[pem_pos=1, final_scores=0.1126764,
final_local_scores=0.11826292, pem=-2.40051,
similarity=20.673262]
2: 5042916 Canada
[pem_pos=0, final_scores=0.5658806,
final_local_scores=0.5371293, pem=-0.41303828,
similarity=22.364887]
3: 31750 Ukraine
[pem_pos=1, final_scores=0.49932617,
final_local_scores=0.539629, pem=-1.0458583,
similarity=25.880857]
4: 6800475 Avonlea,_Saskatchewan
[pem_pos=0, final_scores=0.07577647,
final_local_scores=0.17564029, pem=0.0, similarity=9.331169]
5: 3434750 United_States
[pem_pos=0, final_scores=0.5525945,
final_local_scores=0.5493194, pem=-0.12994486,
similarity=21.18916]
6: 26292671 Human_Rights_League_(France)
[pem_pos=0, final_scores=0.09174679,
final_local_scores=0.20332396, pem=0.0, similarity=10.157306]
```

7: 261859 Trent_University
 [pem_pos=0, final_scores=0.22942357,
 final_local_scores=0.20141259, pem=0.0, similarity=10.100266]
 8: 261859 Trent_University
 [pem_pos=0, final_scores=0.19050132,
 final_local_scores=0.15665475, pem=0.0, similarity=8.764603]
 9: 31750 Ukraine
 [pem_pos=1, final_scores=0.2578103,
 final_local_scores=0.27580953, pem=-1.8523102,
 similarity=22.39357]
 10: 5042916 Canada
 [pem_pos=1, final_scores=0.6442649,
 final_local_scores=0.62712663, pem=-0.41303828,
 similarity=25.050589]
 11: 4764461 World_War_I
 [pem_pos=0, final_scores=0.36086023,
 final_local_scores=0.39260554, pem=0.0, similarity=15.805845]
 12: 1302610 Omni_Television
 [pem_pos=0, final_scores=0.12506501,
 final_local_scores=0.1482197, pem=0.0, similarity=8.512884]
 13: 5042916 Canada
 [pem_pos=0, final_scores=0.51247305,
 final_local_scores=0.4787897, pem=-0.11662793,
 similarity=19.011993]
 14: 645042 New_York_City
 [pem_pos=1, final_scores=0.43638378,
 final_local_scores=0.4207211, pem=-1.7906687,
 similarity=26.382805]
 15: 6310 Columbia_University
 [pem_pos=0, final_scores=0.38027978,
 final_local_scores=0.3895684, pem=0.0, similarity=15.71521]
 16: 64646 Toronto
 [pem_pos=0, final_scores=0.526382,
 final_local_scores=0.4865144, pem=-0.12291032,
 similarity=19.276678]

Sample: b'Spondylosoma'

` Spondylosoma ' -LRB- meaning `` vertebra body '' -RRB- is a 1 genus of 2 archosaur of uncertain affinities from the 3 late Ladinian - age 4 Middle Triassic 5 Lower Santa Maria Formation in 6 Geopark of 7 Paleorrota , [1 1 Brazil 1] . Recent studies have gone back and forth on its identity , suggesting 8 rauisuchia n or 9 basal 10 saurischia n 11 dinosaur . If it was a dinosaur , it would be the oldest . [2 2 Friedrich von Huene 2] based the genus on a fragmentary postcranial 12 skeleton held at the [3 3 University of Tübingen 3] . This skeleton includes two , two cervical e , four dorsal vertebrae , three 13 sacral vertebrae , 14 scapula e , part of a 15 humerus , part of a 16 femur , and part of a 17 pubis . At the time , he thought it was a . With the discovery of basal dinosaur 18 Staurikosaurus , Spondylosoma drew attention as a possible relative . Authors have gone back and

forth on the question , considering it either as a basal dinosaur , or as a ' ' 19 thecodont ' ' or other basal archosaur . The two most recent papers to address it illustrate this clearly : in 2000 , 20 Peter Galton claimed that it lacks dinosaurian 21 characteristics and was probably a 22 rauisuchia n , whereas in 2004 disputed this and included Spondylosoma as a possible basal dinosaur similar to the

1: 3383 Brazil

2: 1347582 Friedrich_von_Huene

3: 262301 University_of_Tübingen

1: 38493 Genus

[winner: 12516072 Pseudochirops, pem_pos=7,
final_scores=-1.0301996, final_local_scores=-0.95111173,
pem=-2.197119, similarity=-12.345083]

gt: 38493 Genus, pem_pos=0, final_scores=-1.1676481,
final_local_scores=-1.2514291, pem=-0.034435946,
similarity=-33.06816]

2: 492039 Archosaur

[winner: 492039 Archosaur, pem_pos=0,
final_scores=-1.14412, final_local_scores=-1.215888, pem=0.0,
similarity=-32.19481]

gt: 492039 Archosaur, pem_pos=0, final_scores=-1.14412,
final_local_scores=-1.215888, pem=0.0, similarity=-32.19481]

3: 3029165 Ladinian

[winner: 3029165 Ladinian, pem_pos=0,
final_scores=-0.08609226, final_local_scores=-0.022783808,
pem=0.0, similarity=3.4097986]

gt: 3029165 Ladinian, pem_pos=0,
final_scores=-0.08609226, final_local_scores=-0.022783808,
pem=0.0, similarity=3.4097986]

[winner: 4899498 Late_(rapper), pem_pos=3,
final_scores=-1.1087112, final_local_scores=-0.9927178,
pem=-1.8690101, similarity=-15.371]

RECALL MISS: 3029165 Ladinian]

4: 3029063 Middle_Triassic

[winner: 3029063 Middle_Triassic, pem_pos=0,
final_scores=-0.8488929, final_local_scores=-0.902363,
pem=0.0, similarity=-22.838598]

gt: 3029063 Middle_Triassic, pem_pos=0,
final_scores=-0.8488929, final_local_scores=-0.902363,
pem=0.0, similarity=-22.838598]

5: 12503145 Santa_Maria_Formation

[winner: 12503145 Santa_Maria_Formation, pem_pos=0,
final_scores=-0.21203957, final_local_scores=-0.17411023,
pem=0.0, similarity=-1.106084]

gt: 12503145 Santa_Maria_Formation, pem_pos=0,
final_scores=-0.21203957, final_local_scores=-0.17411023,
pem=0.0, similarity=-1.106084]

6: 898161 Geopark

[winner: 898161 Geopark, pem_pos=0,
final_scores=-0.7250928, final_local_scores=-0.76228523,

```

pem=0.0, similarity=-18.6584|
    gt: 898161 Geopark, pem_pos=0, final_scores=-0.7250928,
final_local_scores=-0.76228523, pem=0.0, similarity=-18.6584 ]
7: 14198658 Paleorrota
    [winner: 14198658 Paleorrota, pem_pos=0,
final_scores=-0.109848574, final_local_scores=-0.07689417,
pem=0.0, similarity=1.7950375|
    gt: 14198658 Paleorrota, pem_pos=0,
final_scores=-0.109848574, final_local_scores=-0.07689417,
pem=0.0, similarity=1.7950375 ]
8: 1654012 Raurisuchia
    [winner: 1654012 Raurisuchia, pem_pos=0,
final_scores=-1.0020655, final_local_scores=-1.0467434,
pem=0.0, similarity=-27.147196|
    gt: 1654012 Raurisuchia, pem_pos=0,
final_scores=-1.0020655, final_local_scores=-1.0467434,
pem=0.0, similarity=-27.147196 ]
9: 11287648 Basal_(phylogenetics)
    [winner: 696132 Basal-cell_carcinoma, pem_pos=0,
final_scores=-1.0622541, final_local_scores=-1.0221949,
pem=0.0, similarity=-26.41462|
    gt: 11287648 Basal_(phylogenetics), pem_pos=1,
final_scores=-1.4738399, final_local_scores=-1.5996864,
pem=-0.122070394, similarity=-42.98428 ]
10: 6832171 Saurischia
    [winner: 6832171 Saurischia, pem_pos=0,
final_scores=-0.99580234, final_local_scores=-1.0364693,
pem=0.0, similarity=-26.8406|
    gt: 6832171 Saurischia, pem_pos=0,
final_scores=-0.99580234, final_local_scores=-1.0364693,
pem=0.0, similarity=-26.8406 ]
11: 8311 Dinosaur
    [winner: 8311 Dinosaur, pem_pos=0,
final_scores=-1.026852, final_local_scores=-1.1111104,
pem=-0.045856673, similarity=-28.818666|
    gt: 8311 Dinosaur, pem_pos=0, final_scores=-1.026852,
final_local_scores=-1.1111104, pem=-0.045856673,
similarity=-28.818666 ]
12: 27609 Skeleton
    [winner: 4741774 Postcrania, pem_pos=0,
final_scores=-0.9266313, final_local_scores=-0.94985086,
pem=0.0, similarity=-24.255732|
    RECALL MISS: 27609 Skeleton]
13: 232244 Sacrum
    [winner: 232244 Sacrum, pem_pos=0,
final_scores=-1.0913696, final_local_scores=-1.115611,
pem=0.0, similarity=-29.202343|
    gt: 232244 Sacrum, pem_pos=0, final_scores=-1.0913696,
final_local_scores=-1.115611, pem=0.0, similarity=-29.202343 ]
14: 211386 Scapula
    [winner: 211386 Scapula, pem_pos=0,
final_scores=-1.373205, final_local_scores=-1.4188898,

```

```

pem=0.0, similarity=-38.25279|
    gt: 211386 Scapula, pem_pos=0, final_scores=-1.373205,
final_local_scores=-1.4188898, pem=0.0, similarity=-38.25279 ]
15: 227807 Humerus
    [winner: 227807 Humerus, pem_pos=0,
final_scores=-1.2637506, final_local_scores=-1.3198016,
pem=-0.0015752198, similarity=-35.287235|
    gt: 227807 Humerus, pem_pos=0, final_scores=-1.2637506,
final_local_scores=-1.3198016, pem=-0.0015752198,
similarity=-35.287235 ]
16: 188497 Femur
    [winner: 188497 Femur, pem_pos=0,
final_scores=-1.3388896, final_local_scores=-1.4372315,
pem=-0.008099065, similarity=-38.756096|
    gt: 188497 Femur, pem_pos=0, final_scores=-1.3388896,
final_local_scores=-1.4372315, pem=-0.008099065,
similarity=-38.756096 ]
17: 3963126 Pubis_(bone)
    [winner: 3963126 Pubis_(bone), pem_pos=0,
final_scores=-1.2246034, final_local_scores=-1.250449,
pem=-0.047420617, similarity=-32.9683|
    gt: 3963126 Pubis_(bone), pem_pos=0,
final_scores=-1.2246034, final_local_scores=-1.250449,
pem=-0.047420617, similarity=-32.9683 ]
18: 1091923 Staurikosaurus
    [winner: 1091923 Staurikosaurus, pem_pos=0,
final_scores=-0.21536756, final_local_scores=-0.17884439,
pem=0.0, similarity=-1.2473607|
    gt: 1091923 Staurikosaurus, pem_pos=0,
final_scores=-0.21536756, final_local_scores=-0.17884439,
pem=0.0, similarity=-1.2473607 ]
19: 6512371 Thecodontia
    [winner: 6512371 Thecodontia, pem_pos=0,
final_scores=-1.0626535, final_local_scores=-1.0906985,
pem=0.0, similarity=-28.458904|
    gt: 6512371 Thecodontia, pem_pos=0,
final_scores=-1.0626535, final_local_scores=-1.0906985,
pem=0.0, similarity=-28.458904 ]
20: 4741936 Peter_Galton
    [winner: 4741936 Peter_Galton, pem_pos=0,
final_scores=-0.045320287, final_local_scores=0.014522225,
pem=0.0, similarity=4.523085|
    gt: 4741936 Peter_Galton, pem_pos=0,
final_scores=-0.045320287, final_local_scores=0.014522225,
pem=0.0, similarity=4.523085 ]
21: 1006006 Synapomorphy
    [winner: 47264 Asteroid_belt, pem_pos=2,
final_scores=-1.4149607, final_local_scores=-1.3950412,
pem=-3.085209, similarity=-20.76325|
    RECALL MISS: 1006006 Synapomorphy]
22: 1654012 Rauisuchia
    [winner: 1654012 Rauisuchia, pem_pos=0,

```

```

final_scores=-1.0775563, final_local_scores=-1.1292056,
pem=0.0, similarity=-29.608034]
    gt: 1654012 Rauisuchia, pem_pos=0,
final_scores=-1.0775563, final_local_scores=-1.1292056,
pem=0.0, similarity=-29.608034 ]
1: 3383 Brazil
    [pem_pos=0, final_scores=0.45299125,
final_local_scores=0.5700406, pem=-0.21313342,
similarity=22.259912]
2: 1347582 Friedrich_von_Huene
    [pem_pos=0, final_scores=0.03815119,
final_local_scores=0.046344593, pem=0.0, similarity=5.4727283]
3: 262301 University_of_Tübingen
    [pem_pos=0, final_scores=0.31073958,
final_local_scores=0.4178834, pem=0.0, similarity=16.560186]

```

Sample: b'Live_Nation_UK'

Live Nation UK is the [1 1 United Kingdom 1] subsidiary of , formerly the entertainment division of , a major media conglomerate but now a separate and independent corporation worldwide . As 1 Clear Channel UK , Live Nation UK acquired the organisation and owns 51 % of [2 2 Academy Music Group 2] , giving it a stake in most major and medium-sized [3 3 London 3] music venues . It has since merged with ticket sales and distribution company , 2 Ticketmaster to become [4 4 Live Nation Entertainment 4] . A large majority of the UK 's biggest regional theatres were owned/managed by Live Nation UK until November 2009 when the majority were sold to [5 5 Ambassador Theatre Group 5] , including the [6 6 Sunderland Empire Theatre 6] , 3 Palace Theatre , [1 Manchester 1] , 4 Grimsby Auditorium , in [7 7 Birmingham 7] , [2 5 Lyceum Theatre 2] , [3 London 3] , [8 8 Edinburgh Playhouse 8] , [9 9 Liverpool Empire 9] , [10 10 New Theatre Oxford 10] and [11 11 Bristol Hippodrome 11] , including [4 6 London 4] 's [5 West End 5] venues : [12 12 Apollo Victoria Theatre 12] and [13 13 Lyceum Theatre 13] . In addition , the [14 14 Dominion Theatre 14] , [6 London 6] was managed by [7 Live Nation 7] until October 2009 when it was bought outright by the majority owner

```

1: 31717 United_Kingdom
2: 3309848 Academy_Music_Group
3: 17867 London
4: 23039062 Live_Nation_Entertainment
5: 8638721 Ambassador_Theatre_Group
6: 1483231 Sunderland_Empire_Theatre
7: 13882056 Birmingham
8: 8108496 Edinburgh_Playhouse
9: 1044571 Liverpool_Empire_Theatre
10: 5284718 New_Theatre_Oxford
11: 2357067 Bristol_Hippodrome
12: 15156375 Apollo_Victoria_Theatre
13: 2492291 Lyceum_Theatre,_London
14: 173849 Dominion_Theatre

```


1: 1756264 Clear_Channel_UK
[winner: 1756264 Clear_Channel_UK, pem_pos=0,
final_scores=-0.15854861, final_local_scores=-0.0932115,
pem=0.0, similarity=1.3080958|
gt: 1756264 Clear_Channel_UK, pem_pos=0,
final_scores=-0.15854861, final_local_scores=-0.0932115,
pem=0.0, similarity=1.3080958]
2: 545768 Ticketmaster
[winner: 545768 Ticketmaster, pem_pos=0,
final_scores=0.014245018, final_local_scores=0.03149928,
pem=0.0, similarity=5.029714|
gt: 545768 Ticketmaster, pem_pos=0,
final_scores=0.014245018, final_local_scores=0.03149928,
pem=0.0, similarity=5.029714]
3: 1518627 Palace_Theatre,_Manchester
[winner: 20206 Manchester, pem_pos=0,
final_scores=0.25687474, final_local_scores=0.21664274,
pem=-0.3209542, similarity=12.300163|
RECALL MISS: 1518627 Palace_Theatre,_Manchester]
[winner: 230266 Palace_Theatre,_London, pem_pos=1,
final_scores=-0.12739812, final_local_scores=-0.18564047,
pem=-1.0512383, similarity=4.2666283|
gt: 1518627 Palace_Theatre,_Manchester, pem_pos=2,
final_scores=-0.23051243, final_local_scores=-0.29387134,
pem=-2.0115747, similarity=6.2592616]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7925146, final_local_scores=-1.7891107,
pem=-4.6728287, similarity=-23.88935|
RECALL MISS: 1518627 Palace_Theatre,_Manchester]
4: 5468681 Grimsby_Auditorium
[winner: 5468681 Grimsby_Auditorium, pem_pos=0,
final_scores=-0.0082214475, final_local_scores=0.029635638,
pem=0.0, similarity=4.9740996|
gt: 5468681 Grimsby_Auditorium, pem_pos=0,
final_scores=-0.0082214475, final_local_scores=0.029635638,
pem=0.0, similarity=4.9740996]
5: 2492291 Lyceum_Theatre,_London
[winner: 17867 London, pem_pos=0,
final_scores=0.32145315, final_local_scores=0.28872603,
pem=-0.09086737, similarity=13.200025|
RECALL MISS: 2492291 Lyceum_Theatre,_London]
[winner: 2492291 Lyceum_Theatre,_London, pem_pos=0,
final_scores=0.16818695, final_local_scores=0.1194759,
pem=-0.1543445, similarity=8.494461|
gt: 2492291 Lyceum_Theatre,_London, pem_pos=0,
final_scores=0.16818695, final_local_scores=0.1194759,
pem=-0.1543445, similarity=8.494461]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8421229, final_local_scores=-1.8433003,
pem=-4.6728287, similarity=-25.506477|
RECALL MISS: 2492291 Lyceum_Theatre,_London]
6: 1154952 West_End_theatre

[winner: 17867 London, pem_pos=0, final_scores=0.605413,
final_local_scores=0.5989091, pem=-0.09086737,
similarity=22.456505]

gt: 1154952 West_End_theatre, pem_pos=11,
final_scores=-0.6878048, final_local_scores=-0.8253591,
pem=-5.2367783, similarity=7.937765]

[winner: 1154952 West_End_theatre, pem_pos=0,
final_scores=0.07880728, final_local_scores=0.018473417,
pem=-0.5811406, similarity=7.8013296]

gt: 1154952 West_End_theatre, pem_pos=0,
final_scores=0.07880728, final_local_scores=0.018473417,
pem=-0.5811406, similarity=7.8013296]

[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-1.8726133, final_local_scores=-1.9477379,
pem=-3.4067376, similarity=-35.508305]

RECALL MISS: 1154952 West_End_theatre]

1: 31717 United_Kingdom

[pem_pos=0, final_scores=0.36993974,
final_local_scores=0.3804794, pem=-0.09933331,
similarity=15.984168]

2: 3309848 Academy_Music_Group

[pem_pos=0, final_scores=0.24895768,
final_local_scores=0.28177887, pem=0.0, similarity=12.498556]

3: 17867 London

[pem_pos=0, final_scores=0.5696901,
final_local_scores=0.55988723, pem=-0.09086737,
similarity=21.292015]

4: 23039062 Live_Nation_Entertainment

[pem_pos=0, final_scores=0.12539719,
final_local_scores=0.109886885, pem=0.0, similarity=7.3689566]

5: 8638721 Ambassador_Theatre_Group

[pem_pos=0, final_scores=0.26725596,
final_local_scores=0.26634908, pem=0.0, similarity=12.0381]

6: 1483231 Sunderland_Empire_Theatre

[pem_pos=0, final_scores=0.2162899,
final_local_scores=0.2047346, pem=0.0, similarity=10.199402]

7: 13882056 Birmingham

[pem_pos=0, final_scores=0.3559938,
final_local_scores=0.3314405, pem=-0.40171203,
similarity=16.16513]

8: 8108496 Edinburgh_Playhouse

[pem_pos=0, final_scores=0.10545416,
final_local_scores=0.10449025, pem=0.0, similarity=7.2079105]

9: 1044571 Liverpool_Empire_Theatre

[pem_pos=0, final_scores=0.22014724,
final_local_scores=0.20314604, pem=0.0, similarity=10.151996]

10: 5284718 New_Theatre_Oxford

[pem_pos=0, final_scores=0.13645135,
final_local_scores=0.11934394, pem=0.0, similarity=7.651174]

11: 2357067 Bristol_Hippodrome

[pem_pos=0, final_scores=0.09764938,
final_local_scores=0.08269879, pem=0.0, similarity=6.55761]

12: 15156375 Apollo_Victoria_Theatre
 [pem_pos=0, final_scores=0.07876293,
 final_local_scores=0.0811875, pem=0.0, similarity=6.51251]
 13: 2492291 Lyceum_Theatre,_London
 [pem_pos=0, final_scores=0.21065871,
 final_local_scores=0.16800582, pem=-0.1543445,
 similarity=9.942691]
 14: 173849 Dominion_Theatre
 [pem_pos=0, final_scores=0.2219473,
 final_local_scores=0.19118282, pem=0.0, similarity=9.79499]
 1: 20206 Manchester
 [pem_pos=0, final_scores=0.25687474,
 final_local_scores=0.21664274, pem=-0.3209542,
 similarity=12.300163]
 2: 2492291 Lyceum_Theatre,_London
 [pem_pos=0, final_scores=0.16818695,
 final_local_scores=0.1194759, pem=-0.1543445,
 similarity=8.494461]
 3: 17867 London
 [pem_pos=0, final_scores=0.32145315,
 final_local_scores=0.28872603, pem=-0.09086737,
 similarity=13.200025]
 4: 17867 London
 [pem_pos=0, final_scores=0.605413,
 final_local_scores=0.5989091, pem=-0.09086737,
 similarity=22.456505]
 5: 1154952 West_End_theatre
 [pem_pos=0, final_scores=0.07880728,
 final_local_scores=0.018473417, pem=-0.5811406,
 similarity=7.8013296]
 6: 17867 London
 [pem_pos=0, final_scores=0.345944,
 final_local_scores=0.31547856, pem=-0.09086737,
 similarity=13.998374]
 7: 23039062 Live_Nation_Entertainment
 [pem_pos=0, final_scores=0.11201899,
 final_local_scores=0.0952732, pem=0.0, similarity=6.932855]

Sample: b'Tony_Blanco'

Tony Hemisphere Blanco -LRB- born November 10 , 1980 in [1 1 S
 an Juan de la Maguana 1] , [2 2 Dominican Republic 2] -RRB-
 is a professional 1 baseball player -LRB- mainly playing as 2
 first baseman , 3 third baseman and 4 outfielder -RRB- , He
 last played in [3 3 Major League Baseball 3] for the [4 4 Was
 hington Nationals 4] . During , his rookie season , he 5 hit .
 177 with one 6 home run and seven [1 7 RBI 1] . Blanco began
 his career in the [2 Boston Red Sox 2] farm system as a third
 basemen . In , he was traded to the [5 5 Cincinnati Reds 5] .
 As a top prospect in the [3 Reds 3] organization in , Blanco
 made a name for himself while playing for the of the [6 6 Car
 olina League 6] . Blanco was picked up by the [4 Nationals 4]
 as a choice in 2005 and started the year playing for the [5

Potomac Nationals 5] . Blanco was invited to 8 spring training for the big league club and quickly made it to the roster early in 2005 , playing in 56 games . He has not yet returned to the majors . In , Blanco played for the Double-A [7 7 Tulsa Drillers 7] in the [8 8 Colorado Rockies 8] organization batting .323 with 23 home runs and 88 [6 RBI 6] . After the 2008 season , Blanco played in the for the [9 9 Estrellas Orientales 9] . He was then acquired by the [7 Chunichi Dragons 7] , who released their cleanup hitter [10 10 Tyrone Woods 10] due to his high wage , shortly after that . His first year in [8 NPB 8] was very successful , and he became known renown for his incredible power . Examples of his power include a grounded home run in 9 Nagoya Dome -LRB- first in dome history -RRB- , hitting the light unit above the advertisement board in [11 11 Tokyo Dome 11]

1: 9117911 San_Juan_de_la_Maguana

2: 8060 Dominican_Republic

3: 38776 Major_League_Baseball

4: 885151 Washington_Nationals

5: 6671 Cincinnati_Reds

6: 70295 Carolina_League

7: 672600 Tulsa_Drillers

8: 6669 Colorado_Rockies

9: 7530168 Estrellas_Orientales

10: 10888494 Tyrone_Woods

11: 680412 Tokyo_Dome

1: 3850 Baseball

[winner: 1087714 Mike_Hampton, pem_pos=1, final_scores=-1.2252496, final_local_scores=-1.3598249, pem=-0.3090048, similarity=-34.80976]

RECALL MISS: 3850 Baseball]

2: 581933 First_baseaman

[winner: 581933 First_baseaman, pem_pos=0, final_scores=-1.3934668, final_local_scores=-1.6044929, pem=0.0, similarity=-43.791553]

gt: 581933 First_baseaman, pem_pos=0, final_scores=-1.3934668, final_local_scores=-1.6044929, pem=0.0, similarity=-43.791553]

3: 305185 Third_baseaman

[winner: 305185 Third_baseaman, pem_pos=0, final_scores=-1.2265804, final_local_scores=-1.4185271, pem=0.0, similarity=-38.241966]

gt: 305185 Third_baseaman, pem_pos=0, final_scores=-1.2265804, final_local_scores=-1.4185271, pem=0.0, similarity=-38.241966]

4: 528703 Outfielder

[winner: 528703 Outfielder, pem_pos=0, final_scores=-1.429633, final_local_scores=-1.6277971, pem=-0.0142888175, similarity=-44.409298]

gt: 528703 Outfielder, pem_pos=0, final_scores=-1.429633, final_local_scores=-1.6277971, pem=-0.0142888175, similarity=-44.409298]

5: 18933273 Batting_average
[winner: 10389717 Another_Step_(Closer_to_You),
pem_pos=10, final_scores=-1.4111226,
final_local_scores=-1.3761437, pem=-2.397191,
similarity=-23.940859]
gt: 18933273 Batting_average, pem_pos=4,
final_scores=-1.943012, final_local_scores=-2.1457713,
pem=-2.0488129, similarity=-48.802616]
6: 14148 Home_run
[winner: 22013347 Run_(waterfalls), pem_pos=18,
final_scores=-1.2048682, final_local_scores=-1.134093,
pem=-3.1146128, similarity=-12.816139]
RECALL MISS: 14148 Home_run]
[winner: 3580035 Home_(Angel), pem_pos=17,
final_scores=-1.5627996, final_local_scores=-1.5419362,
pem=-3.9339054, similarity=-20.531551]
RECALL MISS: 14148 Home_run]
7: 236469 Run_batted_in
[winner: 300747 Run_(baseball), pem_pos=2,
final_scores=0.15938555, final_local_scores=0.11443442,
pem=-1.0986123, similarity=13.479089]
gt: 236469 Run_batted_in, pem_pos=0,
final_scores=-0.10126521, final_local_scores=-0.17199807,
pem=-0.041762907, similarity=-0.8159393]
8: 18952605 Spring_training
[winner: 17721852 To_Spring, pem_pos=0,
final_scores=-0.95904744, final_local_scores=-0.84710336,
pem=0.0, similarity=-21.18954]
RECALL MISS: 18952605 Spring_training]
[winner: 11242284 Discens, pem_pos=7,
final_scores=-1.1889392, final_local_scores=-1.0959945,
pem=-2.07863, similarity=-17.313036]
RECALL MISS: 18952605 Spring_training]
9: 1130633 Nagoya_Dome
[winner: 1130633 Nagoya_Dome, pem_pos=0,
final_scores=0.0037944466, final_local_scores=0.06675807,
pem=0.0, similarity=6.081907]
gt: 1130633 Nagoya_Dome, pem_pos=0,
final_scores=0.0037944466, final_local_scores=0.06675807,
pem=0.0, similarity=6.081907]
1: 9117911 San_Juan_de_la_Maguana
[pem_pos=0, final_scores=0.22045682,
final_local_scores=0.34991047, pem=0.0, similarity=14.531739]
2: 8060 Dominican_Republic
[pem_pos=0, final_scores=0.2523921,
final_local_scores=0.3003154, pem=-0.24480942,
similarity=14.383036]
3: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.41747338,
final_local_scores=0.36716986, pem=0.0, similarity=15.046794]
4: 885151 Washington_Nationals
[pem_pos=0, final_scores=0.40827447,

```

final_local_scores=0.3498406, pem=-0.090107284,
similarity=15.019671]
5: 6671 Cincinnati_Reds
    [pem_pos=0, final_scores=0.55337346,
final_local_scores=0.5079885, pem=-0.106723964,
similarity=19.829485]
6: 70295 Carolina_League
    [pem_pos=0, final_scores=0.3667434,
final_local_scores=0.35416976, pem=0.0, similarity=14.658845]
7: 672600 Tulsa_Drillers
    [pem_pos=0, final_scores=0.28786844,
final_local_scores=0.27794144, pem=0.0, similarity=12.38404]
8: 6669 Colorado_Rockies
    [pem_pos=0, final_scores=0.4542578,
final_local_scores=0.40461886, pem=-0.10158596,
similarity=16.716787]
9: 7530168 Estrellas_Orientales
    [pem_pos=0, final_scores=0.11035715,
final_local_scores=0.1358552, pem=0.0, similarity=8.143903]
10: 10888494 Tyrone_Woods
    [pem_pos=0, final_scores=0.41420525,
final_local_scores=0.4629473, pem=0.0, similarity=17.904984]
11: 680412 Tokyo_Dome
    [pem_pos=0, final_scores=0.05627255,
final_local_scores=0.09838621, pem=0.0, similarity=7.025754]
1: 300747 Run_(baseball)
    [pem_pos=2, final_scores=0.15938555,
final_local_scores=0.11443442, pem=-1.0986123,
similarity=13.479089]
2: 4339 Boston_Red_Sox
    [pem_pos=0, final_scores=0.52412707,
final_local_scores=0.48292708, pem=-0.08447198,
similarity=18.960592]
3: 6671 Cincinnati_Reds
    [pem_pos=1, final_scores=0.39554745,
final_local_scores=0.33558762, pem=-1.0258155,
similarity=19.682861]
4: 885151 Washington_Nationals
    [pem_pos=0, final_scores=0.3503548,
final_local_scores=0.28657216, pem=-0.64318085,
similarity=16.139315]
5: 1654881 Potomac_Nationals
    [pem_pos=0, final_scores=0.31331474,
final_local_scores=0.30626172, pem=0.0, similarity=13.229174]
6: 300747 Run_(baseball)
    [pem_pos=2, final_scores=0.15687408,
final_local_scores=0.11169101, pem=-1.0986123,
similarity=13.39722]
7: 495027 Chunichi_Dragons
    [pem_pos=0, final_scores=0.2930358,
final_local_scores=0.28093073, pem=0.0, similarity=12.473247]
8: 147794 Nippon_Professional_Baseball

```

[pem_pos=0, final_scores=0.24063061,
final_local_scores=0.21611837, pem=0.0, similarity=10.539116]

Sample: b'Janene_Higgins'

Janene Higgins is a 1 graphic designer and 2 video artist based in . Her work spans several genres , from video performance in the experimental music scene to videos for [1 1 Saks Fifth Avenue 1] , and the design of several hundred CD packages for such record labels as 3 Sony , , [2 2 PolyGram 2] , and a wide variety of independents . Higgins moved to [1 New York City 1] in the late 80s to design magazines , working for such publications as [3 3 Esquire 3] , Working Woman , and Fame , culminating in three years as a designer for [4 4 Vanity Fair 4] magazine . In the mid-90 's , she began designing CD covers and videos for friends in the 4 downtown music scene , which led to an ongoing series of music/video performance collaborations with some of [2 New York 2] 's preeminent composers and improvisors of new music . These include duos with 5 Elliott Sharp , 6 Ikue Mori , 7 Mari Kimura , [5 5 Alan Licht 5] , , , , and 8 Zeena Parkins . In addition to video performance , she has directed several short experimental works , often with sound design by notable avant-garde artists including Sharp , Parkins , [6 6 Christian Marclay 6] and 9 John Duncan . Her work has been performed and exhibited at The New York Video Festival at ; 10 documenta in ; [7 7 The Kitchen 7] , [3 NYC 3] ; [4 11 Musée 4] d'art contemporain de Lyon ; City of Women festival , [5 Slovenia 5] ; The 12 Chelsea Art Museum , [6 NYC 6] ; HiTeca Festival in [7 Porto 7] , [8 Portugal 8] ; [8 8 Art Institute of Chicago 8] ; Experimenta Festival in [9 Buenos Aires 9] ; The Hamburg Short Film Festival ; and at The 13 Impakt Festival

1: 29483 Saks_Fifth_Avenue
2: 423772 PolyGram
3: 244002 Esquire_(magazine)
4: 286542 Vanity_Fair_(magazine)
5: 6702046 Alan_Licht
6: 309302 Christian_Marclay
7: 964958 The_Kitchen
8: 8758450 Art_Institute_of_Chicago

1: 1571880 Graphic_designer

[winner: 2490135 When_the_Wind_Blows_(comics), pem_pos=0,
final_scores=-0.8894927, final_local_scores=-0.8427999,
pem=0.0, similarity=-21.061117]

RECALL MISS: 1571880 Graphic_designer]

[winner: 6978016 Rachel_Roy, pem_pos=0,
final_scores=-0.920204, final_local_scores=-0.94039947,
pem=0.0, similarity=-23.973682]

RECALL MISS: 1571880 Graphic_designer]

2: 32622 Video_art

[winner: 32622 Video_art, pem_pos=0,
final_scores=-1.042764, final_local_scores=-1.164114,
pem=0.0, similarity=-30.649769]

gt: 32622 Video_art, pem_pos=0, final_scores=-1.042764,
final_local_scores=-1.164114, pem=0.0, similarity=-30.649769]

3: 26989 Sony
[winner: 26989 Sony, pem_pos=0,
final_scores=-0.0076197535, final_local_scores=0.0109063685,
pem=-0.18419, similarity=5.4168344]
gt: 26989 Sony, pem_pos=0, final_scores=-0.0076197535,
final_local_scores=0.0109063685, pem=-0.18419,
similarity=5.4168344]

4: 487233 Downtown_music
[winner: 13769835 Shichiten_Hakki_Shijou_Shugi!,
pem_pos=2, final_scores=-0.7242074,
final_local_scores=-0.5728591, pem=-1.0986123,
similarity=-7.031121]
RECALL MISS: 487233 Downtown_music]
[winner: 487233 Downtown_music, pem_pos=0,
final_scores=-0.9779864, final_local_scores=-1.0289218,
pem=0.0, similarity=-26.615366]
gt: 487233 Downtown_music, pem_pos=0,
final_scores=-0.9779864, final_local_scores=-1.0289218,
pem=0.0, similarity=-26.615366]

5: 486882 Elliott_Sharp
[winner: 486882 Elliott_Sharp, pem_pos=0,
final_scores=-0.10221194, final_local_scores=-0.06762811,
pem=0.0, similarity=2.0715554]
gt: 486882 Elliott_Sharp, pem_pos=0,
final_scores=-0.10221194, final_local_scores=-0.06762811,
pem=0.0, similarity=2.0715554]

6: 408331 Ikue_Mori
[winner: 408331 Ikue_Mori, pem_pos=0,
final_scores=-0.15546392, final_local_scores=-0.10661153,
pem=0.0, similarity=0.9082123]
gt: 408331 Ikue_Mori, pem_pos=0,
final_scores=-0.15546392, final_local_scores=-0.10661153,
pem=0.0, similarity=0.9082123]

7: 16733369 Mari_Kimura
[winner: 16733369 Mari_Kimura, pem_pos=0,
final_scores=-0.063386604, final_local_scores=-0.016648397,
pem=0.0, similarity=3.5928915]
gt: 16733369 Mari_Kimura, pem_pos=0,
final_scores=-0.063386604, final_local_scores=-0.016648397,
pem=0.0, similarity=3.5928915]

8: 708292 Zeena_Parkins
[winner: 708292 Zeena_Parkins, pem_pos=0,
final_scores=-0.00426504, final_local_scores=0.055000976,
pem=0.0, similarity=5.731052]
gt: 708292 Zeena_Parkins, pem_pos=0,
final_scores=-0.00426504, final_local_scores=0.055000976,
pem=0.0, similarity=5.731052]

9: 5292602 John_Duncan_(artist)
[winner: 1089396 Jimmy_Duncan_(U.S._politician),
pem_pos=0, final_scores=-0.12436183,


```

final_local_scores=-0.036944807, pem=-1.0822316,
similarity=8.87255|
    gt: 5292602 John_Duncan_(artist), pem_pos=1,
final_scores=-0.20501758, final_local_scores=-0.1922842,
pem=-1.3764124, similarity=5.836714 ]
10: 243987 Documenta
    [winner: 243987 Documenta, pem_pos=0,
final_scores=-1.086942, final_local_scores=-1.1859934,
pem=0.0, similarity=-31.302694|
    gt: 243987 Documenta, pem_pos=0, final_scores=-1.086942,
final_local_scores=-1.1859934, pem=0.0,
similarity=-31.302694 ]
11: 27692237 Musée_d'art_contemporain_de_Lyon
    [winner: 53919 Musée_d'Orsay, pem_pos=0,
final_scores=0.1618961, final_local_scores=0.1881578,
pem=0.0, similarity=9.704717|
    RECALL MISS: 27692237 Musée_d'art_contemporain_de_Lyon]
    [winner: 8638634 Lyon, pem_pos=0,
final_scores=-0.4177469, final_local_scores=-0.4778858,
pem=-0.4769171, similarity=-7.5778036|
    RECALL MISS: 27692237 Musée_d'art_contemporain_de_Lyon]
    [winner: 3522809 Palästinalied, pem_pos=2,
final_scores=-1.6518682, final_local_scores=-1.6532512,
pem=-3.6810567, similarity=-25.228437|
    RECALL MISS: 27692237 Musée_d'art_contemporain_de_Lyon]
12: 16713459 Chelsea_Art_Museum
    [winner: 16713459 Chelsea_Art_Museum, pem_pos=0,
final_scores=-0.5386169, final_local_scores=-0.55551106,
pem=0.0, similarity=-12.487845|
    gt: 16713459 Chelsea_Art_Museum, pem_pos=0,
final_scores=-0.5386169, final_local_scores=-0.55551106,
pem=0.0, similarity=-12.487845 ]
13: 24499730 Impakt_Festival
    [winner: 24499730 Impakt_Festival, pem_pos=0,
final_scores=-0.91608137, final_local_scores=-0.9255872,
pem=0.0, similarity=-23.531654|
    gt: 24499730 Impakt_Festival, pem_pos=0,
final_scores=-0.91608137, final_local_scores=-0.9255872,
pem=0.0, similarity=-23.531654 ]
1: 29483 Saks_Fifth_Avenue
    [pem_pos=0, final_scores=0.3265834,
final_local_scores=0.36313224, pem=0.0, similarity=14.926303]
2: 423772 PolyGram
    [pem_pos=0, final_scores=0.14316033,
final_local_scores=0.1638161, pem=0.0, similarity=8.978312]
3: 244002 Esquire_(magazine)
    [pem_pos=0, final_scores=0.3196376,
final_local_scores=0.31888363, pem=-0.046910003,
similarity=13.860941]
4: 286542 Vanity_Fair_(magazine)
    [pem_pos=0, final_scores=0.36290067,
final_local_scores=0.38269395, pem=-0.26357266,

```

similarity=16.943413]
 5: 6702046 Alan_Licht
 [pem_pos=0, final_scores=0.09319578,
 final_local_scores=0.1680153, pem=0.0, similarity=9.103624]
 6: 309302 Christian_Marclay
 [pem_pos=0, final_scores=0.22533472,
 final_local_scores=0.29304338, pem=0.0, similarity=12.834712]
 7: 964958 The_Kitchen
 [pem_pos=0, final_scores=0.1589333,
 final_local_scores=0.15255529, pem=0.0, similarity=8.642267]
 8: 8758450 Art_Institute_of_Chicago
 [pem_pos=0, final_scores=0.3903684,
 final_local_scores=0.39713985, pem=0.0, similarity=15.941158]
 1: 645042 New_York_City
 [pem_pos=0, final_scores=0.6434316,
 final_local_scores=0.61001277, pem=0.0, similarity=22.29371]
 2: 645042 New_York_City
 [pem_pos=1, final_scores=0.45449656,
 final_local_scores=0.40363002, pem=-1.7906687,
 similarity=25.872772]
 3: 645042 New_York_City
 [pem_pos=0, final_scores=0.31169707,
 final_local_scores=0.24764326, pem=-0.38631868,
 similarity=13.580743]
 4: 53919 Musée_d'Orsay
 [pem_pos=0, final_scores=0.1618961,
 final_local_scores=0.1881578, pem=0.0, similarity=9.704717]
 5: 27338 Slovenia
 [pem_pos=0, final_scores=0.32338154,
 final_local_scores=0.41363013, pem=-0.27845925,
 similarity=17.947565]
 6: 645042 New_York_City
 [pem_pos=0, final_scores=0.310893,
 final_local_scores=0.24676493, pem=-0.38631868,
 similarity=13.554532]
 7: 100715 Porto
 [pem_pos=0, final_scores=0.34083766,
 final_local_scores=0.3761124, pem=-0.42562416,
 similarity=17.62827]
 8: 23033 Portugal
 [pem_pos=0, final_scores=0.52663434,
 final_local_scores=0.60788804, pem=-0.30942732,
 similarity=23.913017]
 9: 255919 Buenos_Aires
 [pem_pos=0, final_scores=0.38947457,
 final_local_scores=0.39387095, pem=-0.12597719,
 similarity=16.52869]

Sample: b'Jablanica'

_ -LRB- river -RRB- The Jablanica -LRB- [1 1 Serbian 1] :
 Јабланица , -RRB- is an river in southern [1 1 Serbia 1] . A
 left tributary of the 2 South -LRB- or Južna -RRB- Morava

river , it gives its name to the region of [2 Jablanica 2] and to modern [3 Serbia 3] 's [2 2 Jablanica District 2] , with the region contributing about one third of the district 's area . The [4 Jablanica 4] originates from the 3 Goljak mountain , near the village of Grbavce , on the administrative border of the [3 3 Kosovo 3] province . The area is rich in thermal springs , so several 4 spas are located near the river : Stara Banja , Ravna Banja and Sijarina with popular 5 Sijarinska Banja . At the village of Maćedonce Retkocersko the [5 Jablanica 5] receives the left tributary Čokotinska reka -LRB- Cyrillic : Чокотинска река -RRB- , turns southeast and the region of [6 Jablanica 6] begins from that point . The upper [7 Jablanica 7] region is made of narrow river valley on the southern slopes of the and 6 Radan mountains , in the westernmost corner of [8 Jablanica District 8] and near the border of the in [4 4 Kosovo 4] . The river flows eastward , passing regional center of 7 Medveđa and the villages of Rujkovac and Šilovo , before reaching the town of [5 5 Lebane 5] where the [9 Jablanica 9] enters the lower part of its valley and flows for the next 48 km in low Field of [10 Leskovac 10] , part of the depression of [11 Leskovac 11] , which itself is part of the composite valley of the [12 Južna Morava 12] . In the region , near Lebane , a famed archeological find of -LRB- Iustiniana Prima -RRB- is located . The [13 Jablanica 13] flows next to the villages of Ždeglovo , Vranovce , Bošnjace , Turekovac and Stopanje , reaches 8 Vinarce , the northern suburb of the city of [6 6 Leskovac 6] , and turns north . First it flows parallel to the 9 Veternica river , to which it is connected by canal at the village of [14 Zalužje 14] , then parallel to the [15 Južna Morava 15] . After the villages of Dupljane and 10 Pečenjevce it turns east and flows into the [16 Južna Morava 16] . From Pečenjevce , a parallel flow -LRB- canal -RRB- begins , connecting the [17 Jablanica 17] and s , some 15 km to the north . Through [18 Južna Morava 18] , the [19 Jablanica 19] belongs to the [7 7 Black Sea 7] 11 drainage basin , draining an area of 895 km ² itself . It 's not navigable and has an average discharge of 6 m ³ / s , which in rainy years grows much higher , so the river floods its valley causing lots of material damage . The name of the river , [20 Jablanica 20] , in [21 Serbian 21] means the

1: 29265 Serbia

2: 696682 Jablanica_District

3: 17391 Kosovo

4: 17391 Kosovo

5: 3723050 Lebane

6: 283915 Leskovac

7: 3386 Black_Sea

1: 75595 Serbian_language

[winner: 29265 Serbia, pem_pos=1,
final_scores=0.23330678, final_local_scores=0.1917029,
pem=-1.2504034, similarity=16.6104]

gt: 75595 Serbian_language, pem_pos=0,
final_scores=-0.193885, final_local_scores=-0.23547232,
pem=-0.7628811, similarity=1.2114153]
2: 4956000 South_Morava
[winner: 960242 Morava_(river), pem_pos=0,
final_scores=-0.08444783, final_local_scores=-0.08344105,
pem=-0.088232785, similarity=2.0794892]
gt: 4956000 South_Morava, pem_pos=7,
final_scores=-0.8786876, final_local_scores=-0.98938423,
pem=-5.422303, similarity=4.051842]
[winner: 179553 Southern_United_States, pem_pos=0,
final_scores=-0.40121824, final_local_scores=-0.34784567,
pem=-1.3313899, similarity=0.94960946]
RECALL MISS: 4956000 South_Morava]
3: 10106183 Goljak
[winner: 10106183 Goljak, pem_pos=0,
final_scores=0.01685132, final_local_scores=0.1801906,
pem=0.0, similarity=9.46696]
gt: 10106183 Goljak, pem_pos=0, final_scores=0.01685132,
final_local_scores=0.1801906, pem=0.0, similarity=9.46696]
4: 4220024 Destination_spa
[winner: 1192598 Spas, pem_pos=4,
final_scores=-1.3675528, final_local_scores=-1.3525138,
pem=-2.2659464, similarity=-23.949425]
gt: 4220024 Destination_spa, pem_pos=3,
final_scores=-1.5631237, final_local_scores=-1.6532481,
pem=-2.201027, similarity=-33.27698]
5: 7394175 Sijarinska_Banja
[winner: 7394175 Sijarinska_Banja, pem_pos=0,
final_scores=-0.089240596, final_local_scores=-0.04928746,
pem=0.0, similarity=2.6188767]
gt: 7394175 Sijarinska_Banja, pem_pos=0,
final_scores=-0.089240596, final_local_scores=-0.04928746,
pem=0.0, similarity=2.6188767]
6: 7498984 Radan_(mountain)
[winner: 7498984 Radan_(mountain), pem_pos=0,
final_scores=0.00689207, final_local_scores=0.07343227,
pem=0.0, similarity=6.2810783]
gt: 7498984 Radan_(mountain), pem_pos=0,
final_scores=0.00689207, final_local_scores=0.07343227,
pem=0.0, similarity=6.2810783]
7: 3557885 Medveda
[winner: 3557885 Medveda, pem_pos=0,
final_scores=0.00032486022, final_local_scores=0.009904519,
pem=-0.01584821, similarity=4.4714684]
gt: 3557885 Medveda, pem_pos=0,
final_scores=0.00032486022, final_local_scores=0.009904519,
pem=-0.01584821, similarity=4.4714684]
8: 27394299 Vinarce
[winner: 27394299 Vinarce, pem_pos=0,
final_scores=-0.23253268, final_local_scores=-0.10247171,
pem=0.0, similarity=1.0317528]

gt: 27394299 Vinarce, pem_pos=0,
final_scores=-0.23253268, final_local_scores=-0.10247171,
pem=0.0, similarity=1.0317528]
9: 5403136 Veternica
[winner: 5403136 Veternica, pem_pos=0,
final_scores=-0.1224138, final_local_scores=-0.06820578,
pem=0.0, similarity=2.0543163]
gt: 5403136 Veternica, pem_pos=0,
final_scores=-0.1224138, final_local_scores=-0.06820578,
pem=0.0, similarity=2.0543163]
10: 6549867 Pečenjevce
[winner: 6549867 Pečenjevce, pem_pos=0,
final_scores=-0.013576627, final_local_scores=0.05559601,
pem=0.0, similarity=5.748809]
gt: 6549867 Pečenjevce, pem_pos=0,
final_scores=-0.013576627, final_local_scores=0.05559601,
pem=0.0, similarity=5.748809]
11: 191162 Drainage_basin
[winner: 182732 Sedimentary_basin, pem_pos=3,
final_scores=-1.3540775, final_local_scores=-1.3750293,
pem=-1.9469279, similarity=-26.356209]
gt: 191162 Drainage_basin, pem_pos=0,
final_scores=-1.365071, final_local_scores=-1.4309559,
pem=-0.7217738, similarity=-34.687748]
[winner: 46596 Drainage, pem_pos=1,
final_scores=-1.7589054, final_local_scores=-1.8338768,
pem=-0.08378179, similarity=-50.18121]
gt: 191162 Drainage_basin, pem_pos=0,
final_scores=-1.7815049, final_local_scores=-1.8858466,
pem=0.0, similarity=-52.187714]
1: 29265 Serbia
[pem_pos=0, final_scores=0.47118783,
final_local_scores=0.4515518, pem=-0.34077588,
similarity=19.418112]
2: 696682 Jablanica_District
[pem_pos=0, final_scores=0.13384853,
final_local_scores=0.17391562, pem=0.0, similarity=9.279702]
3: 17391 Kosovo
[pem_pos=0, final_scores=0.35985315,
final_local_scores=0.34395528, pem=-0.3821891,
similarity=16.432428]
4: 17391 Kosovo
[pem_pos=0, final_scores=0.3474024,
final_local_scores=0.3303547, pem=-0.3821891,
similarity=16.02656]
5: 3723050 Lebane
[pem_pos=0, final_scores=0.06382491,
final_local_scores=0.0887173, pem=0.0, similarity=6.737214]
6: 283915 Leskovac
[pem_pos=0, final_scores=0.35408854,
final_local_scores=0.32832772, pem=0.0, similarity=13.887667]
7: 3386 Black_Sea

[pem_pos=0, final_scores=0.2794718,
final_local_scores=0.32603076, pem=0.0, similarity=13.819121]
1: 29265 Serbia
[pem_pos=1, final_scores=0.23330678,
final_local_scores=0.1917029, pem=-1.2504034,
similarity=16.6104]
2: 1896341 Jablanica,_Bosnia_and_Herzegovina
[pem_pos=0, final_scores=0.24462466,
final_local_scores=0.20297033, pem=-0.6719055,
similarity=13.800678]
3: 29265 Serbia
[pem_pos=0, final_scores=0.49085635,
final_local_scores=0.4730367, pem=-0.34077588,
similarity=20.059265]
4: 1896341 Jablanica,_Bosnia_and_Herzegovina
[pem_pos=0, final_scores=0.06866504,
final_local_scores=0.010761157, pem=-0.6719055,
similarity=8.064774]
5: 1896341 Jablanica,_Bosnia_and_Herzegovina
[pem_pos=0, final_scores=0.16405062,
final_local_scores=0.114955425, pem=-0.6719055,
similarity=11.174138]
6: 1896341 Jablanica,_Bosnia_and_Herzegovina
[pem_pos=0, final_scores=0.25555438,
final_local_scores=0.21490937, pem=-0.6719055,
similarity=14.156964]
7: 1896341 Jablanica,_Bosnia_and_Herzegovina
[pem_pos=0, final_scores=0.17732726,
final_local_scores=0.12945813, pem=-0.6719055,
similarity=11.606928]
8: 696682 Jablanica_District
[pem_pos=0, final_scores=0.13621424,
final_local_scores=0.17649981, pem=0.0, similarity=9.356819]
9: 1896341 Jablanica,_Bosnia_and_Herzegovina
[pem_pos=0, final_scores=0.14878608,
final_local_scores=0.09828122, pem=-0.6719055,
similarity=10.676547]
10: 283915 Leskovac
[pem_pos=0, final_scores=0.41678637,
final_local_scores=0.39681554, pem=0.0, similarity=15.931479]
11: 283915 Leskovac
[pem_pos=0, final_scores=0.37541896,
final_local_scores=0.35162798, pem=0.0, similarity=14.5829935]
12: 4956000 South_Morava
[pem_pos=0, final_scores=0.16573574,
final_local_scores=0.15886435, pem=0.0, similarity=8.830542]
13: 1896341 Jablanica,_Bosnia_and_Herzegovina
[pem_pos=0, final_scores=0.05325277,
final_local_scores=-0.010778636, pem=-0.6719055,
similarity=7.4219837]
14: 29017458 Zalužje_(Nevesinje)
[pem_pos=0, final_scores=0.063611194,

final_local_scores=0.1921649, pem=-0.034486163,
 similarity=10.011838]
 15: 4956000 South_Morava
 [pem_pos=0, final_scores=0.16643159,
 final_local_scores=0.15962449, pem=0.0, similarity=8.853226]
 16: 4956000 South_Morava
 [pem_pos=0, final_scores=0.19907843,
 final_local_scores=0.19528615, pem=0.0, similarity=9.917441]
 17: 1896341 Jablanica, Bosnia_and_Herzegovina
 [pem_pos=0, final_scores=0.20929871,
 final_local_scores=0.1643821, pem=-0.6719055,
 similarity=12.64913]
 18: 4956000 South_Morava
 [pem_pos=0, final_scores=0.2029338,
 final_local_scores=0.19949758, pem=0.0, similarity=10.0431185]
 19: 1896341 Jablanica, Bosnia_and_Herzegovina
 [pem_pos=0, final_scores=0.20518263,
 final_local_scores=0.15988588, pem=-0.6719055,
 similarity=12.514953]
 20: 1896341 Jablanica, Bosnia_and_Herzegovina
 [pem_pos=0, final_scores=0.13223414,
 final_local_scores=0.08020075, pem=-0.6719055,
 similarity=10.13699]
 21: 29265 Serbia
 [pem_pos=1, final_scores=0.27578467,
 final_local_scores=0.23810351, pem=-1.2504034,
 similarity=17.995085]

Sample: b'Frank_Coombs'

Frank Leslie Coombs -LRB- December 27 , 1853 -- October 5 ,
 1934 -RRB- was a from [1 1 California 1] . Frank Leslie
 Coombs was born in [2 2 Napa 2] , [1 1 Napa County 1] , [2 Ca
 lifornia 2] , the son of [3 3 Nathan Coombs 3] and Maria
 Isabel Gordon . He attended the public schools in [3
 California 3] and in . He graduated from the law department
 of [4 4 Columbian University 4] -LRB- now 2 George Washington
 University Law School -RRB- , [4 3 Washington 4] , [5 D.C. 5]
 , in 1875 . He was admitted to the bar in 1875 and commenced
 practice in [6 Napa 6] . He was the of [7 Napa County 7] from
 1880-1885 . Coombs was a member of the [5 5 California State A
 ssembly 5] from 1887 to 1893 and from 1897 to 1899 , each
 time representing [8 Napa County 8] , and served as speaker
 in 1891 and again in 1897 . On the death of , he was
 appointed [9 4 United States 9] Minister to [10 Japan 10] and
 served from June 1892 to August 1893 . He was the State 5 Lib
 rarian of [11 California 11] from April 1 , 1898 to April 1 ,
 1899 . He was the 6 United States Attorney for the 7 Northern
 District of California from April 1 , 1899 to March 1 ,
 1901 . Coombs was elected as a [6 6 Republican 6] to the 8 Fi
 fty-seventh Congress -LRB- March 4 , 1901 - March 3 , 1903 -
 RRB- . He was an unsuccessful candidate for reelection in
 1902 . He resumed the practice of law in [12 Napa 12] , and

was again a member of the [13 State Assembly 13] from 1921 to 1931 , representing [14 Napa 14] and 9 Lake

1: 5407 California

2: 107794 Napa,_California

3: 24855464 Nathan_Coombs

4: 268451 George_Washington_University

5: 476068 California_State_Assembly

6: 32070 Republican_Party_(United_States)

1: 82026 Napa_County,_California

[winner: 5407 California, pem_pos=0,
final_scores=0.7279815, final_local_scores=0.7045231,
pem=-0.08662026, similarity=25.585142]

RECALL MISS: 82026 Napa_County,_California]

[winner: 82026 Napa_County,_California, pem_pos=0,
final_scores=0.33548504, final_local_scores=0.2966628,
pem=0.0, similarity=12.942723]

gt: 82026 Napa_County,_California, pem_pos=0,
final_scores=0.33548504, final_local_scores=0.2966628,
pem=0.0, similarity=12.942723]

[winner: 5481852 Xi_baryon, pem_pos=11,
final_scores=-1.8970478, final_local_scores=-1.8694985,
pem=-4.6728287, similarity=-26.28828]

RECALL MISS: 82026 Napa_County,_California]

2: 1878671 George_Washington_University_Law_School

[winner: 1878671 George_Washington_University_Law_School,
pem_pos=0, final_scores=-0.35509652,
final_local_scores=-0.3710097, pem=0.0,
similarity=-6.9819565]

gt: 1878671 George_Washington_University_Law_School,
pem_pos=0, final_scores=-0.35509652,
final_local_scores=-0.3710097, pem=0.0,
similarity=-6.9819565]

3: 108956 Washington,_D.C.

[winner: 108956 Washington,_D.C., pem_pos=1,
final_scores=0.3217643, final_local_scores=0.29316854,
pem=-1.7496481, similarity=22.353306]

gt: 108956 Washington,_D.C., pem_pos=1,
final_scores=0.3217643, final_local_scores=0.29316854,
pem=-1.7496481, similarity=22.353306]

[winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=0.13099588, final_local_scores=0.081064925,
pem=-0.15729709, similarity=7.364258]

gt: 108956 Washington,_D.C., pem_pos=0,
final_scores=0.13099588, final_local_scores=0.081064925,
pem=-0.15729709, similarity=7.364258]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8076434, final_local_scores=-1.7819014,
pem=-4.6728287, similarity=-23.67421]

RECALL MISS: 108956 Washington,_D.C.]

4: 2317760 United_States_Ambassador_to_Japan

[winner: 15573 Japan, pem_pos=0, final_scores=0.52887255,
final_local_scores=0.61387384, pem=-0.13549899,

similarity=23.145796|
 RECALL MISS: 2317760 United_States_Ambassador_to_Japan]
 [winner: 3434750 United_States, pem_pos=0,
 final_scores=0.4786629, final_local_scores=0.47057974,
 pem=-0.073457584, similarity=18.532225|
 RECALL MISS: 2317760 United_States_Ambassador_to_Japan]
 [winner: 739237 Minister_(government), pem_pos=0,
 final_scores=-1.1587257, final_local_scores=-1.2138283,
 pem=-0.49105036, similarity=-29.462936|
 RECALL MISS: 2317760 United_States_Ambassador_to_Japan]
 5: 250641 Librarian
 [winner: 6870497 United_States_Senate_Librarian,
 pem_pos=2, final_scores=-1.1382811,
 final_local_scores=-1.0962849, pem=-1.486541,
 similarity=-20.541573|
 gt: 250641 Librarian, pem_pos=0,
 final_scores=-2.2270494, final_local_scores=-2.429021,
 pem=-0.4152604, similarity=-66.13886]
 6: 443584 United_States_Attorney
 [winner: 443584 United_States_Attorney, pem_pos=0,
 final_scores=-0.0025879592, final_local_scores=-0.0055204034,
 pem=0.0, similarity=3.9249732|
 gt: 443584 United_States_Attorney, pem_pos=0,
 final_scores=-0.0025879592, final_local_scores=-0.0055204034,
 pem=0.0, similarity=3.9249732]
 7: 891692
 United_States_District_Court_for_the_Northern_District_of_California
 [winner: 891692
 United_States_District_Court_for_the_Northern_District_of_California,
 pem_pos=0, final_scores=-0.021842286,
 final_local_scores=0.02744171, pem=0.0, similarity=4.9086285|
 gt: 891692
 United_States_District_Court_for_the_Northern_District_of_California,
 pem_pos=0, final_scores=-0.021842286,
 final_local_scores=0.02744171, pem=0.0, similarity=4.9086285]
 8: 2117822 57th_United_States_Congress
 [winner: 31756 United_States_Congress, pem_pos=0,
 final_scores=-0.04238698, final_local_scores=-0.07618551,
 pem=-0.3349529, similarity=3.6377113|
 RECALL MISS: 2117822 57th_United_States_Congress]
 [winner: 2117822 57th_United_States_Congress, pem_pos=0,
 final_scores=-0.82188064, final_local_scores=-0.8465533,
 pem=0.0, similarity=-21.173126|
 gt: 2117822 57th_United_States_Congress, pem_pos=0,
 final_scores=-0.82188064, final_local_scores=-0.8465533,
 pem=0.0, similarity=-21.173126]
 9: 80563 Lake_County,_California
 [winner: 18842431 Lake, pem_pos=0,
 final_scores=-0.58297557, final_local_scores=-0.5668504,
 pem=-1.5881573, similarity=-4.189587|
 gt: 80563 Lake_County,_California, pem_pos=4,

```
final_scores=-0.65779716, final_local_scores=-0.74063295,  
pem=-2.9945033, similarity=-1.7276776 ]  
1: 5407 California  
    [pem_pos=0, final_scores=0.6627877,  
final_local_scores=0.6333087, pem=-0.08662026,  
similarity=23.459963]  
2: 107794 Napa,_California  
    [pem_pos=0, final_scores=0.33581656,  
final_local_scores=0.29494998, pem=-0.37185588,  
similarity=14.913818]  
3: 24855464 Nathan_Coombs  
    [pem_pos=0, final_scores=0.1735272,  
final_local_scores=0.22826004, pem=0.0, similarity=10.901448]  
4: 268451 George_Washington_University  
    [pem_pos=0, final_scores=0.115716085,  
final_local_scores=0.12133357, pem=0.0, similarity=7.710549]  
5: 476068 California_State_Assembly  
    [pem_pos=0, final_scores=0.190315,  
final_local_scores=0.1725893, pem=0.0, similarity=9.240122]  
6: 32070 Republican_Party_(United_States)  
    [pem_pos=0, final_scores=0.41872942,  
final_local_scores=0.44289976, pem=-0.1310191,  
similarity=18.019228]  
1: 82026 Napa_County,_California  
    [pem_pos=0, final_scores=0.33548504,  
final_local_scores=0.2966628, pem=0.0, similarity=12.942723]  
2: 5407 California  
    [pem_pos=0, final_scores=0.7279815,  
final_local_scores=0.7045231, pem=-0.08662026,  
similarity=25.585142]  
3: 5407 California  
    [pem_pos=0, final_scores=0.76675147,  
final_local_scores=0.7468734, pem=-0.08662026,  
similarity=26.848959]  
4: 108956 Washington,_D.C.  
    [pem_pos=1, final_scores=0.3217643,  
final_local_scores=0.29316854, pem=-1.7496481,  
similarity=22.353306]  
5: 108956 Washington,_D.C.  
    [pem_pos=0, final_scores=0.13099588,  
final_local_scores=0.081064925, pem=-0.15729709,  
similarity=7.364258]  
6: 107794 Napa,_California  
    [pem_pos=0, final_scores=0.19769959,  
final_local_scores=0.14407817, pem=-0.37185588,  
similarity=10.411503]  
7: 82026 Napa_County,_California  
    [pem_pos=0, final_scores=0.29790467,  
final_local_scores=0.25561196, pem=0.0, similarity=11.717684]  
8: 82026 Napa_County,_California  
    [pem_pos=0, final_scores=0.22913168,  
final_local_scores=0.18048796, pem=0.0, similarity=9.475834]
```

9: 3434750 United_States
 [pem_pos=0, final_scores=0.4786629,
 final_local_scores=0.47057974, pem=-0.073457584,
 similarity=18.532225]
 10: 15573 Japan
 [pem_pos=0, final_scores=0.52887255,
 final_local_scores=0.61387384, pem=-0.13549899,
 similarity=23.145796]
 11: 5407 California
 [pem_pos=0, final_scores=0.035491124,
 final_local_scores=-0.055539586, pem=-0.08662026,
 similarity=2.9033554]
 12: 107794 Napa,_California
 [pem_pos=0, final_scores=0.16140787,
 final_local_scores=0.10320699, pem=-0.37185588,
 similarity=9.191825]
 13: 476068 California_State_Assembly
 [pem_pos=0, final_scores=0.10510568,
 final_local_scores=0.079511076, pem=-0.3172562,
 similarity=8.187771]
 14: 107794 Napa,_California
 [pem_pos=0, final_scores=0.17657436,
 final_local_scores=0.1197741, pem=-0.37185588,
 similarity=9.686221]

Sample: b'Jiri_Tranovsky'

[1 Jiří Třanovský 1] -LRB- , , -RRB- -LRB- 9 April 1592 , [1
 1 Teschen 1] , [2 2 Silesia 2] -- 29 May 1637 , 1 Liptovský
 Sv . Mikuláš , [3 3 Upper Hungary 3] -RRB- , was a 2 hymn
 writer from the [4 4 Cieszyn Silesia 4] . He was sometimes
 called the father of [2 Slovak 2] hymnody and the '' [5 5 Lut
 her 5] of the [3 Slavs 3] . '' His name is sometimes to
 George . Třanovský was born in [6 6 Teschen 6] , and studied
 at [7 7 Guben 7] and . In 1607 , he was admitted to the where
 [8 8 Martin Luther 8] had taught less than a century
 earlier . He traveled in [9 9 Bohemia 9] and [10 10 Silesia 10
] in 1612 and became a teacher at St. Nicholas Gymnasium in [1
 1 11 Prague 11] . Later , he became rector of a school in [12
 12 Holešov 12] , [13 13 Moravia 13] . In 1616 he was ordained
 a priest in 3 Meziříčí and served as a pastor for four
 years . The persecution of in [4 Bohemia 4] under 4 Ferdinand
 II forced him into exile . After an imprisonment in 1623 and
 the death of two children from 5 plague the following year ,
 Třanovský received a call to be pastor to a church in [14 14 B
 ielitz 14] , [15 15 Teschen Silesia 15] . He also became
 personal chaplain to Count Kasper Illehazy in 1627 .
 Třanovský was a lover of poetry and hymns . He issued several
 collections of hymns , the first being the 6 Latin Odarum
 Sacrarum sive Hymnorum Libri III in 1629 , but his most
 important and most famous work was Cithara Sanctorum -LRB-
 Lyre of the Saints -RRB- , written in [5 Czech 5] , which
 appeared in 1636 in [16 16 Levoča 16] . This latter volume

has formed the basis of [6 Czech 6] and [7 Slovak 7] Lutheran hymnody to the present day . In addition to hymn collections , Třanovský translated the [17 17 Augsburg Confession 17] in 1620 into [8 Czech 8] . These two latter works together with [9 7 Bible 9] of Kralice are the pillars that supported the Slovak Reformation . From 1631 until 1637 , Třanovský was pastor at a church in [18 18 Liptovský Svätý Mikuláš 18] in present-day [19 19 Slovakia 19] . He died on 29 May of that year and was buried in an unmarked grave at his church . He was forty-six years old . Třanovský is commemorated on 29 May in the [20 20 Evangelical Lutheran Church in America 20] and the 8 Evangelical Lutheran Church in Canada

1: 353922 Cieszyn
2: 28399 Silesia
3: 2130666 Upper_Hungary
4: 789474 Cieszyn_Silesia
5: 7567080 Martin_Luther
6: 353922 Cieszyn
7: 703939 Guben
8: 7567080 Martin_Luther
9: 4345 Bohemia
10: 28399 Silesia
11: 23844 Prague
12: 5363219 Holešov
13: 20216 Moravia
14: 74603 Bielsko-Biała
15: 789474 Cieszyn_Silesia
16: 609312 Levoča
17: 23883384 Augsburg_Confession
18: 1060406 Liptovský_Mikuláš
19: 26830 Slovakia
20: 248255 Evangelical_Lutheran_Church_in_America

1: 1060406 Liptovský_Mikuláš

[winner: 155823 Sievert, pem_pos=1,
final_scores=-1.1742786, final_local_scores=-1.1436574,
pem=-0.85181254, similarity=-25.407017]

RECALL MISS: 1060406 Liptovský_Mikuláš]

[winner: 4721653 Zyphe_language, pem_pos=12,
final_scores=-1.6299151, final_local_scores=-1.5995905,
pem=-4.4188404, similarity=-19.61492]

RECALL MISS: 1060406 Liptovský_Mikuláš]

2: 13756 Hymn

[winner: 1571479 Adam-on-di-Ahman_(hymn), pem_pos=0,
final_scores=-0.7403154, final_local_scores=-0.68864983,
pem=0.0, similarity=-16.460972]

RECALL MISS: 13756 Hymn]

3: 4445175 Valašské_Meziříčí

[winner: 24076116 Meziříčí_(Tábor_District), pem_pos=0,
final_scores=-0.041694626, final_local_scores=0.08155513,
pem=-0.06280091, similarity=6.8650017]

gt: 4445175 Valašské_Meziříčí, pem_pos=1,

```

final_scores=-0.18934172, final_local_scores=-0.18194214,
pem=-2.799022, similarity=13.881713 ]
4: 151083 Ferdinand_II,_Holy_Roman_Emperor
   [winner: 151083 Ferdinand_II,_Holy_Roman_Emperor,
pem_pos=0, final_scores=0.028968528,
final_local_scores=0.01736632, pem=-0.632278,
similarity=8.046385]
   gt: 151083 Ferdinand_II,_Holy_Roman_Emperor, pem_pos=0,
final_scores=0.028968528, final_local_scores=0.01736632,
pem=-0.632278, similarity=8.046385 ]
5: 16392927 Bubonic_plague
   [winner: 4746 Plague_(disease), pem_pos=0,
final_scores=-1.0145406, final_local_scores=-1.0831273,
pem=-0.75807434, similarity=-24.110435]
   gt: 16392927 Bubonic_plague, pem_pos=1,
final_scores=-1.1312217, final_local_scores=-1.1945953,
pem=-1.2438496, similarity=-24.79515 ]
6: 17730 Latin
   [winner: 17730 Latin, pem_pos=0, final_scores=-0.3368508,
final_local_scores=-0.363006, pem=-0.08185364,
similarity=-6.297977]
   gt: 17730 Latin, pem_pos=0, final_scores=-0.3368508,
final_local_scores=-0.363006, pem=-0.08185364,
similarity=-6.297977 ]
7: 2342832 Bible_of_Kralice
   [winner: 3390 Bible, pem_pos=0, final_scores=0.51038194,
final_local_scores=0.5647554, pem=0.0, similarity=20.94314]
   RECALL MISS: 2342832 Bible_of_Kralice]
   [winner: 24129151 Kralice_nad_Oslavou, pem_pos=0,
final_scores=-0.15737998, final_local_scores=-0.06669958,
pem=-0.6931472, similarity=5.8687067]
   RECALL MISS: 2342832 Bible_of_Kralice]
8: 848904 Evangelical_Lutheran_Church_in_Canada
   [winner: 848904 Evangelical_Lutheran_Church_in_Canada,
pem_pos=0, final_scores=-0.037301242,
final_local_scores=0.006651774, pem=0.0, similarity=4.288215]
   gt: 848904 Evangelical_Lutheran_Church_in_Canada,
pem_pos=0, final_scores=-0.037301242,
final_local_scores=0.006651774, pem=0.0, similarity=4.288215 ]
1: 353922 Cieszyn
   [pem_pos=1, final_scores=0.114716604,
final_local_scores=0.071156636, pem=-0.6864312,
similarity=9.946089]
2: 28399 Silesia
   [pem_pos=0, final_scores=0.24773382,
final_local_scores=0.18213958, pem=-0.040829465,
similarity=9.747158]
3: 2130666 Upper_Hungary
   [pem_pos=0, final_scores=0.088538036,
final_local_scores=0.07236616, pem=0.0, similarity=6.249264]
4: 789474 Cieszyn_Silesia
   [pem_pos=0, final_scores=0.41541094,

```

final_local_scores=0.39436483, pem=-0.009901085,
similarity=15.9121895]
5: 7567080 Martin_Luther
[pem_pos=0, final_scores=0.11450963,
final_local_scores=0.108419806, pem=0.0, similarity=7.3251762]
6: 353922 Cieszyn
[pem_pos=1, final_scores=0.24806689,
final_local_scores=0.21972305, pem=-0.6864312,
similarity=14.379606]
7: 703939 Guben
[pem_pos=0, final_scores=0.3327759,
final_local_scores=0.3581988, pem=0.0, similarity=14.779079]
8: 7567080 Martin_Luther
[pem_pos=0, final_scores=0.45104975,
final_local_scores=0.48319507, pem=0.0, similarity=18.509218]
9: 4345 Bohemia
[pem_pos=0, final_scores=0.3659243,
final_local_scores=0.30547822, pem=-0.08606065,
similarity=13.673803]
10: 28399 Silesia
[pem_pos=0, final_scores=0.37906814,
final_local_scores=0.32560235, pem=-0.040829465,
similarity=14.028374]
11: 23844 Prague
[pem_pos=0, final_scores=0.5059412,
final_local_scores=0.49534017, pem=-0.05392049,
similarity=19.164879]
12: 5363219 Holešov
[pem_pos=0, final_scores=0.3601793,
final_local_scores=0.4155553, pem=-0.017178318,
similarity=16.584131]
13: 20216 Moravia
[pem_pos=0, final_scores=0.3545491,
final_local_scores=0.29057, pem=-0.003610146,
similarity=12.780534]
14: 74603 Bielsko-Biała
[pem_pos=0, final_scores=0.5996792,
final_local_scores=0.6271479, pem=0.0, similarity=22.805058]
15: 789474 Cieszyn_Silesia
[pem_pos=0, final_scores=0.28154898,
final_local_scores=0.24814099, pem=0.0, similarity=11.494735]
16: 609312 Levoča
[pem_pos=0, final_scores=0.25793183,
final_local_scores=0.25354573, pem=-0.040454958,
similarity=11.876023]
17: 23883384 Augsburg_Confession
[pem_pos=0, final_scores=0.16058616,
final_local_scores=0.18852967, pem=0.0, similarity=9.715815]
18: 1060406 Liptovský_Mikuláš
[pem_pos=0, final_scores=0.39627683,
final_local_scores=0.41114986, pem=0.0, similarity=16.359245]
19: 26830 Slovakia

[pem_pos=0, final_scores=0.54591995,
 final_local_scores=0.5469836, pem=-0.36602598,
 similarity=22.403301]
 20: 248255 Evangelical_Lutheran_Church_in_America
 [pem_pos=0, final_scores=0.17791493,
 final_local_scores=0.20793977, pem=0.0, similarity=10.295051]
 1: 5549725 Jiří_Třanovský
 [pem_pos=0, final_scores=0.06840269,
 final_local_scores=0.10467629, pem=0.0, similarity=7.2134624]
 2: 26830 Slovakia
 [pem_pos=1, final_scores=0.5619238,
 final_local_scores=0.56446534, pem=-1.0527612,
 similarity=26.659563]
 3: 29440 Slavs
 [pem_pos=0, final_scores=0.23177509,
 final_local_scores=0.22184005, pem=0.0, similarity=10.709863]
 4: 4345 Bohemia
 [pem_pos=0, final_scores=0.44331282,
 final_local_scores=0.38963085, pem=-0.08606065,
 similarity=16.185087]
 5: 5321 Czech_Republic
 [pem_pos=1, final_scores=0.17165329,
 final_local_scores=0.132613, pem=-1.208016,
 similarity=14.6165285]
 6: 5321 Czech_Republic
 [pem_pos=1, final_scores=0.56348747,
 final_local_scores=0.5588434, pem=-1.208016,
 similarity=27.336092]
 7: 26830 Slovakia
 [pem_pos=1, final_scores=0.26470852,
 final_local_scores=0.23980281, pem=-1.0527612,
 similarity=16.970985]
 8: 5321 Czech_Republic
 [pem_pos=1, final_scores=0.31629294,
 final_local_scores=0.2888208, pem=-1.208016,
 similarity=19.27808]
 9: 3390 Bible
 [pem_pos=0, final_scores=0.51038194,
 final_local_scores=0.5647554, pem=0.0, similarity=20.94314]

Sample: b'John_Burton'

_ -LRB- political_agent -RRB- John Burton was a [1 1 Labour Party 1] councillor in [2 2 County Durham 2] , [3 3 England 3] . He was a member of [4 4 Sedgefield Borough Council 4] , representing the of [5 5 Fishburn 5] and [6 6 Old Trimdon 6] . However , he is most notable for being the 1 constituency agent of former [1 British 1] 2 Prime Minister [7 7 Tony Blair 7] , who represented the 3 Sedgefield constituency from 1983 until standing down as Prime Minister in 2007 . [2 Burton 2] was one of the first to spot the potential of the young [3 Tony Blair 3] . When the new Sedgefield constituency was created for the 4 1983 general election , which included [4

Burton 4] 's home of [8 8 Trimdon 8] , the local [5 Labour Party 5] found themselves needing to appoint a candidate for the forthcoming election at short notice . [6 Blair 6] , then a [9 9 London 9] - based 5 barrister , identified [7 Burton 7] as a branch secretary willing to consider his candidature and travelled to Trimdon -LRB- [8 Burton 8] 's power base -RRB- to put his case . [9 Blair 9] secured the nomination , was elected an 6 MP and began his political rise . [10 Burton 10] has remained one of [11 Blair 11] 's inner circle . In the Sedgefield Council Register of Members ' Interests [12 Burton 12] also lists himself as Political Assistant to Rt. . Hon. [13 Tony Blair 13] , MP Myrobella House , Farfield Terrace , [14 Trimdon Station 14] , Co. . [15 Durham 15] , TS29 6DU . The Sunday newspaper [10 10 The Observer 10] reported in December 2007 that [16 Burton 16] plans to write a book about [17 Tony Blair 17] . According to Sedgefield Council 's Register of Members ' Interests , [18 Burton 18] is a member of the 7 Transport and General Workers ' Union , a , and a governor of Trimdon College . He was formerly a 8 physical education teacher . He is , unlike the [11 11 Newcastle 11] - supporting [19 Blair 19] , a devoted fan of [20 9 Sunderland 20] A.F.C.

1: 19279158 Labour_Party_(UK)

2: 400297 County_Durham

3: 9316 England

4: 515698 Sedgefield_(borough)

5: 1606680 Fishburn

6: 1391460 Trimdon

7: 3301347 Tony_Blair

8: 1391460 Trimdon

9: 17867 London

10: 191086 The_Observer

11: 57802 Newcastle_United_F.C.

1: 2928464 Election_agent

[winner: 2928464 Election_agent, pem_pos=0, final_scores=-1.0755473, final_local_scores=-1.1651108, pem=0.0, similarity=-30.67952]

gt: 2928464 Election_agent, pem_pos=0, final_scores=-1.0755473, final_local_scores=-1.1651108, pem=0.0, similarity=-30.67952]

2: 24150 Prime_Minister_of_the_United_Kingdom

[winner: 2806649 Prime_(film), pem_pos=4, final_scores=-0.7252668, final_local_scores=-0.66442543, pem=-2.1258695, similarity=-4.1772585]

RECALL MISS: 24150 Prime_Minister_of_the_United_Kingdom]

[winner: 739237 Minister_(government), pem_pos=0, final_scores=-1.0085369, final_local_scores=-1.1119945, pem=-0.49105036, similarity=-26.424011]

RECALL MISS: 24150 Prime_Minister_of_the_United_Kingdom]

3: 1789868 Sedgefield_(UK_Parliament_constituency)

[winner: 349709 Sedgefield, pem_pos=0, final_scores=0.016624019, final_local_scores=0.018827677,

pem=-0.5523662, similarity=7.6554213|
 gt: 1789868 Sedgefield_(UK_Parliament_constituency),
 pem_pos=1, final_scores=-0.13138613,
 final_local_scores=-0.14190349, pem=-0.9671507,
 similarity=5.1145453]
 [winner: 12174399
 Nigerian_National_Assembly_delegation_from_Katsina,
 pem_pos=14, final_scores=-1.3648463,
 final_local_scores=-1.2882652, pem=-3.2927792,
 similarity=-16.448046|
 RECALL MISS: 1789868
 Sedgefield_(UK_Parliament_constituency)]
 4: 80330 United_Kingdom_general_election,_1983
 [winner: 3301347 Tony_Blair, pem_pos=0,
 final_scores=-0.7218689, final_local_scores=-0.87597424,
 pem=-0.33647218, similarity=-20.221317|
 gt: 80330 United_Kingdom_general_election,_1983,
 pem_pos=5, final_scores=-1.3035035,
 final_local_scores=-1.4502409, pem=-3.5228868,
 similarity=-20.030357]
 [winner: 12795 General_election, pem_pos=0,
 final_scores=-0.9526939, final_local_scores=-1.0208548,
 pem=-0.5947298, similarity=-23.140398|
 RECALL MISS: 80330 United_Kingdom_general_election,_1983]
 5: 4848 Barrister
 [winner: 219487 Bar_association, pem_pos=0,
 final_scores=-1.1697806, final_local_scores=-1.2466764,
 pem=0.0, similarity=-33.1136|
 gt: 4848 Barrister, pem_pos=1, final_scores=-1.1810856,
 final_local_scores=-1.311667, pem=-0.01358637,
 similarity=-34.979164]
 6: 98066 Member_of_Parliament
 [winner: 98066 Member_of_Parliament, pem_pos=0,
 final_scores=-0.2406496, final_local_scores=-0.28891286,
 pem=-0.41824368, similarity=-2.257549|
 gt: 98066 Member_of_Parliament, pem_pos=0,
 final_scores=-0.2406496, final_local_scores=-0.28891286,
 pem=-0.41824368, similarity=-2.257549]
 7: 359312 Transport_and_General_Workers'_Union
 [winner: 359312 Transport_and_General_Workers'_Union,
 pem_pos=0, final_scores=-0.32170606,
 final_local_scores=-0.35837537, pem=0.0,
 similarity=-6.6049232|
 gt: 359312 Transport_and_General_Workers'_Union,
 pem_pos=0, final_scores=-0.32170606,
 final_local_scores=-0.35837537, pem=0.0,
 similarity=-6.6049232]
 8: 217324 Physical_education
 [winner: 217324 Physical_education, pem_pos=0,
 final_scores=-1.28809, final_local_scores=-1.3784263,
 pem=-0.00019361463, similarity=-37.044224|
 gt: 217324 Physical_education, pem_pos=0,

```

final_scores=-1.28809, final_local_scores=-1.3784263,
pem=-0.00019361463, similarity=-37.044224 ]
9: 184474 Sunderland_A.F.C.
    [winner: 184474 Sunderland_A.F.C., pem_pos=0,
final_scores=0.04198663, final_local_scores=0.081094876,
pem=-0.39640337, similarity=8.665449]
    gt: 184474 Sunderland_A.F.C., pem_pos=0,
final_scores=0.04198663, final_local_scores=0.081094876,
pem=-0.39640337, similarity=8.665449 ]
    [winner: 787518 Air_Force_Cross_(United_Kingdom),
pem_pos=0, final_scores=-1.1320852,
final_local_scores=-1.1634037, pem=0.0,
similarity=-30.628572]
    RECALL MISS: 184474 Sunderland_A.F.C.]
1: 19279158 Labour_Party_(UK)
    [pem_pos=0, final_scores=0.15546726,
final_local_scores=0.10768205, pem=-0.49561852,
similarity=9.99841]
2: 400297 County_Durham
    [pem_pos=0, final_scores=0.4149192,
final_local_scores=0.44572133, pem=0.0, similarity=17.390926]
3: 9316 England
    [pem_pos=0, final_scores=0.6666782,
final_local_scores=0.7091265, pem=-0.19053856,
similarity=26.28764]
4: 515698 Sedgefield_(borough)
    [pem_pos=0, final_scores=0.1495925,
final_local_scores=0.2602113, pem=0.0, similarity=11.854937]
5: 1606680 Fishburn
    [pem_pos=0, final_scores=0.11263533,
final_local_scores=0.20862424, pem=0.0, similarity=10.315476]
6: 1391460 Trimdon
    [pem_pos=0, final_scores=0.10895805,
final_local_scores=0.16975033, pem=0.0, similarity=9.155401]
7: 3301347 Tony_Blair
    [pem_pos=0, final_scores=0.38865346,
final_local_scores=0.34006804, pem=0.0, similarity=14.238022]
8: 1391460 Trimdon
    [pem_pos=0, final_scores=0.0937808,
final_local_scores=0.15317148, pem=0.0, similarity=8.660655]
9: 17867 London
    [pem_pos=0, final_scores=0.40703332,
final_local_scores=0.42757314, pem=-0.09086737,
similarity=17.343498]
10: 191086 The_Observer
    [pem_pos=0, final_scores=0.24537851,
final_local_scores=0.29333055, pem=0.0, similarity=12.843282]
11: 57802 Newcastle_United_F.C.
    [pem_pos=2, final_scores=0.30305505,
final_local_scores=0.3587646, pem=-2.4885974,
similarity=28.329344]
1: 31717 United_Kingdom

```

[pem_pos=0, final_scores=0.5238077,
final_local_scores=0.5517462, pem=-0.48936236,
similarity=23.216148]
2: 5749478 John_L._Burton
[pem_pos=0, final_scores=0.06557511,
final_local_scores=0.07796322, pem=-0.2852129,
similarity=7.9673233]
3: 3301347 Tony_Blair
[pem_pos=0, final_scores=0.42034274,
final_local_scores=0.37468386, pem=0.0, similarity=15.271027]
4: 5749478 John_L._Burton
[pem_pos=0, final_scores=0.18474011,
final_local_scores=0.20813286, pem=-0.2852129,
similarity=11.851845]
5: 19279158 Labour_Party_(UK)
[pem_pos=0, final_scores=0.10325952,
final_local_scores=0.05065304, pem=-0.49561852,
similarity=8.296551]
6: 3301347 Tony_Blair
[pem_pos=0, final_scores=0.83741516,
final_local_scores=0.8327734, pem=0.0, similarity=28.941332]
7: 5749478 John_L._Burton
[pem_pos=0, final_scores=0.24180023,
final_local_scores=0.2704624, pem=-0.2852129,
similarity=13.711882]
8: 5749478 John_L._Burton
[pem_pos=0, final_scores=0.171156,
final_local_scores=0.19329429, pem=-0.2852129,
similarity=11.409032]
9: 3301347 Tony_Blair
[pem_pos=0, final_scores=0.8189632,
final_local_scores=0.8126175, pem=0.0, similarity=28.33984]
10: 5749478 John_L._Burton
[pem_pos=0, final_scores=0.32786733,
final_local_scores=0.3643145, pem=-0.2852129,
similarity=16.512617]
11: 3301347 Tony_Blair
[pem_pos=0, final_scores=0.6569969,
final_local_scores=0.63569385, pem=0.0, similarity=23.060085]
12: 5749478 John_L._Burton
[pem_pos=0, final_scores=0.24701305,
final_local_scores=0.2761566, pem=-0.2852129,
similarity=13.881808]
13: 3301347 Tony_Blair
[pem_pos=0, final_scores=0.29882902,
final_local_scores=0.24120414, pem=0.0, similarity=11.287725]
14: 1391460 Trimdon
[pem_pos=0, final_scores=0.0817558,
final_local_scores=0.14003602, pem=0.0, similarity=8.268667]
15: 68489 Durham
[pem_pos=0, final_scores=0.19901206,
final_local_scores=0.1925014, pem=-1.0101248,

```

similarity=15.327554]
16: 5749478 John_L._Burton
    [pem_pos=0, final_scores=0.1395839,
final_local_scores=0.15880656, pem=-0.2852129,
similarity=10.379849]
17: 3301347 Tony_Blair
    [pem_pos=0, final_scores=0.37160093,
final_local_scores=0.3214408, pem=0.0, similarity=13.682147]
18: 5749478 John_L._Burton
    [pem_pos=0, final_scores=0.17016666,
final_local_scores=0.1922136, pem=-0.2852129,
similarity=11.376781]
19: 3301347 Tony_Blair
    [pem_pos=0, final_scores=0.6390069,
final_local_scores=0.6160426, pem=0.0, similarity=22.473654]
20: 184474 Sunderland_A.F.C.
    [pem_pos=0, final_scores=0.04198663,
final_local_scores=0.081094876, pem=-0.39640337,
similarity=8.665449]

```

Sample: b'Satavahana_Express'

Satavahana Express -LRB- ; [1 Hindi 1] : - शातवाहना एक्सप्रेस -RRB- , is an [2 Intercity Express 2] running between [3 Vijayawada 3] and [1 1 Secunderabad 1] in the [2 2 Andhra Pradesh 2] state . The 1 Vijayawada Division of of [3 3 Indian Railways 3] administers this train . The train covers in just 5 hours and 35 minutes . Due to the increasing demand , [4 4 Mahabubabad 4] and [5 5 Madhira 5] were added as stops to the train . The train halts at [6 6 Kazipet 6] , [7 7 Warangal 7] , [8 8 Mahabubabad 8] , [9 9 Khammam 9] and 2 Madhira before reaching 3 Vijayawada Junction.This is the one of Intercity expresses that depart from [4 Vijayawada 4] early in the morning . The other are and express . The rake composition is SLR , UR , UR , UR , UR , PC , D1 , D2 , D3 , D4 , D5 , C1 , UR , UR , UR , SLR making a total of 18 coaches . This train is hauled by a WAP-7 locomotive of the [10 10 Lallaguda 10] shed . The train is called in honour of the which ruled [5 4 Andhra 5] and parts of [11 11 Maharashtra 11] , [12 12 Madhya Pradesh 12] , [13 13 Vidarbha 13] , [14 14 Karnataka 14] and [1 5 15 Goa 15] . Satavahana Express starts from [6 Vijayawada Junction 6] at 06:10 [7 IST 7] and reaches [8 Secunderabad Junction 8] at 11:45 [9 IST 9] . On its return journey it starts from [10 Secunderabad Junction 10] at 16:15 [11 IST 11] and reaches [12 Vijayawada Junction 12] at 21:50 [13 IST 13] . It has an average delay of 25 minutes to reach [16 16 Secunderabad 16] and delay of 20 minutes to reach [14 Vijayawada 14]

```

1: 1013659 Secunderabad
2: 2377 Andhra_Pradesh
3: 655887 Indian_Railways
4: 7400466 Mahabubabad
5: 3447423 Madhira

```

6: 3639399 Kazipet
7: 995245 Warangal
8: 7400466 Mahabubabad
9: 505838 Khammam
10: 10251024 Lallaguda
11: 20629 Maharashtra
12: 47945 Madhya_Pradesh
13: 638748 Vidarbha
14: 16880 Karnataka
15: 40010153 Goa
16: 1013659 Secunderabad
1: 750441 Vijayawada
[winner: 7462818 Vijayawada_Airport, pem_pos=3,
final_scores=-1.0549381, final_local_scores=-1.1237314,
pem=-1.5956699, similarity=-20.76717]
gt: 750441 Vijayawada, pem_pos=0,
final_scores=-1.0564029, final_local_scores=-1.2539756,
pem=0.0, similarity=-33.33142]
2: 3447423 Madhira
[winner: 3447423 Madhira, pem_pos=0,
final_scores=0.027680144, final_local_scores=0.038479984,
pem=0.0, similarity=5.2380323]
gt: 3447423 Madhira, pem_pos=0,
final_scores=0.027680144, final_local_scores=0.038479984,
pem=0.0, similarity=5.2380323]
3: 750441 Vijayawada
[winner: 750441 Vijayawada, pem_pos=0,
final_scores=-1.0235637, final_local_scores=-1.2181039,
pem=0.0, similarity=-32.260937]
gt: 750441 Vijayawada, pem_pos=0,
final_scores=-1.0235637, final_local_scores=-1.2181039,
pem=0.0, similarity=-32.260937]
4: 990249 Coastal_Andhra
[winner: 2377 Andhra_Pradesh, pem_pos=0,
final_scores=0.17908706, final_local_scores=0.1135329,
pem=-0.24926202, similarity=8.833286]
gt: 990249 Coastal_Andhra, pem_pos=2,
final_scores=-0.22599795, final_local_scores=-0.24992773,
pem=-1.9261807, similarity=7.106241]
1: 1013659 Secunderabad
[pem_pos=0, final_scores=0.37815833,
final_local_scores=0.34009823, pem=0.0, similarity=14.238922]
2: 2377 Andhra_Pradesh
[pem_pos=0, final_scores=0.32706147,
final_local_scores=0.27517244, pem=0.0, similarity=12.301408]
3: 655887 Indian_Railways
[pem_pos=0, final_scores=0.1977294,
final_local_scores=0.17612255, pem=0.0, similarity=9.345561]
4: 7400466 Mahabubabad
[pem_pos=0, final_scores=0.21411888,
final_local_scores=0.18820438, pem=0.0, similarity=9.706107]
5: 3447423 Madhira

[pem_pos=0, final_scores=0.0869606,
final_local_scores=0.10323484, pem=0.0, similarity=7.1704464]
6: 3639399 Kazipet
[pem_pos=0, final_scores=0.22018664,
final_local_scores=0.20724902, pem=0.0, similarity=10.274437]
7: 995245 Warangal
[pem_pos=0, final_scores=0.34584647,
final_local_scores=0.32256877, pem=-0.058167208,
similarity=14.032131]
8: 7400466 Mahabubabad
[pem_pos=0, final_scores=0.2777604,
final_local_scores=0.25772303, pem=0.0, similarity=11.780683]
9: 505838 Khammam
[pem_pos=0, final_scores=0.251805,
final_local_scores=0.21958676, pem=0.0, similarity=10.64262]
10: 10251024 Lallaguda
[pem_pos=0, final_scores=0.08690141,
final_local_scores=0.13820973, pem=0.0, similarity=8.214167]
11: 20629 Maharashtra
[pem_pos=0, final_scores=0.3227536,
final_local_scores=0.27528876, pem=0.0, similarity=12.304878]
12: 47945 Madhya_Pradesh
[pem_pos=0, final_scores=0.33239907,
final_local_scores=0.28921762, pem=0.0, similarity=12.720544]
13: 638748 Vidarbha
[pem_pos=0, final_scores=0.24723803,
final_local_scores=0.22318342, pem=0.0, similarity=10.749951]
14: 16880 Karnataka
[pem_pos=0, final_scores=0.28842962,
final_local_scores=0.23711297, pem=0.0, similarity=11.165637]
15: 40010153 Goa
[pem_pos=0, final_scores=0.27667183,
final_local_scores=0.26514378, pem=-0.10114117,
similarity=12.552154]
16: 1013659 Secunderabad
[pem_pos=0, final_scores=0.30922157,
final_local_scores=0.2647115, pem=0.0, similarity=11.989232]
1: 13652 Hindi
[pem_pos=0, final_scores=0.10476525,
final_local_scores=0.09868415, pem=0.0, similarity=7.0346446]
2: 186716 Intercity-Express
[pem_pos=0, final_scores=0.09813224,
final_local_scores=0.17696893, pem=-0.3087355,
similarity=11.04977]
3: 750441 Vijayawada
[pem_pos=0, final_scores=0.3483699,
final_local_scores=0.28232402, pem=0.0, similarity=12.514825]
4: 750441 Vijayawada
[pem_pos=0, final_scores=0.28914815,
final_local_scores=0.21763325, pem=0.0, similarity=10.584323]
5: 2377 Andhra_Pradesh
[pem_pos=0, final_scores=0.17908706,

final_local_scores=0.1135329, pem=-0.24926202,
 similarity=8.833286]
 6: 750441 Vijayawada
 [pem_pos=1, final_scores=0.27489382,
 final_local_scores=0.2014727, pem=-0.6931472,
 similarity=13.871502]
 7: 513283 Indian_Standard_Time
 [pem_pos=0, final_scores=0.20983173,
 final_local_scores=0.21859077, pem=-0.30001742,
 similarity=12.244438]
 8: 14286283 Secunderabad_railway_station
 [pem_pos=0, final_scores=0.20596229,
 final_local_scores=0.16770673, pem=0.0, similarity=9.094417]
 9: 513283 Indian_Standard_Time
 [pem_pos=0, final_scores=0.20246087,
 final_local_scores=0.21053922, pem=-0.30001742,
 similarity=12.004165]
 10: 1013659 Secunderabad
 [pem_pos=1, final_scores=0.23358522,
 final_local_scores=0.18217412, pem=-0.5389965,
 similarity=12.457299]
 11: 513283 Indian_Standard_Time
 [pem_pos=0, final_scores=0.14824913,
 final_local_scores=0.15132114, pem=-0.30001742,
 similarity=10.236979]
 12: 750441 Vijayawada
 [pem_pos=1, final_scores=0.22843607,
 final_local_scores=0.15072471, pem=-0.6931472,
 similarity=12.35708]
 13: 513283 Indian_Standard_Time
 [pem_pos=0, final_scores=0.16891794,
 final_local_scores=0.17389867, pem=-0.30001742,
 similarity=10.910737]
 14: 750441 Vijayawada
 [pem_pos=0, final_scores=0.14415579,
 final_local_scores=0.059251115, pem=0.0, similarity=5.8578844]

Sample: b'Joseph_J'

__ Cannon Joseph Jenne Cannon -LRB- May 22 , 1877 --
 November 5 , 1945 -RRB- was a [1 1 Utah 1] politician and
 newspaper editor and was a leader in [2 2 The Church of Jesus
 Christ of Latter-day Saints 2] -LRB- LDS Church -RRB- . He
 was a member of the prominent [1 1 Cannon 1] political family
 . As a young man , [2 Cannon 2] served as a for the [3 LDS
 Church 3] in [3 3 Europe 3] . He accompanied [4 LDS Church 4]
 2 apostle [4 4 Francis M. Lyman 4] in offering prayers in and
 [5 5 Moscow 5] which dedicated [6 6 Russia 6] for the
 preaching of [7 7 Mormonism 7] in August 1903 . Lyman and [5
 Cannon 5] also similarly dedicated [8 8 Finland 8] in 1903 .
 In the 1908 election , [6 Cannon 6] was elected as a member
 of the [9 9 Utah House of Representatives 9] from . He served
 one term , from 1909 to 1911 . [7 Cannon 7] was not formally

associated with any political party . From 1931 to 1934 , [8 Cannon 8] was the editor of the [10 10 Deseret News 10] , a [1 1 11 Salt Lake City 11] newspaper owned by the [9 LDS Church 9] . His tenure ended when the [10 LDS Church 10] asked him to become the 3 president of the church 's [11 British 11] 4 Mission . [12 Cannon 12] served in this capacity for three years , until 1937 . Immediately following his return to [13 Utah 13] , [14 Cannon 14] was asked to become the first assistant to 5 George Q. Morris , the general superintendent of the church 's . [15 Cannon 15] served in this capacity until his death from 6 pancreatic cancer

1: 31716 Utah

2: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints

3: 9239 Europe

4: 1783358 Francis_M._Lyman

5: 19004 Moscow

6: 25391 Russia

7: 21023 Mormonism

8: 10577 Finland

9: 7123926 Utah_House_of_Representatives

10: 1046373 Deseret_News

11: 53837 Salt_Lake_City

1: 3024243 Cannon_family

[winner: 13510620 Joseph_J._Cannon, pem_pos=0, final_scores=0.14240755, final_local_scores=0.11652532, pem=0.0, similarity=7.567061]

RECALL MISS: 3024243 Cannon_family]

[winner: 1648510 Stevenson_family, pem_pos=1, final_scores=-1.1058549, final_local_scores=-1.0862467, pem=-1.0240437, similarity=-22.757149]

RECALL MISS: 3024243 Cannon_family]

2: 363753 Apostle_(Latter_Day_Saints)

[winner: 1671932 Seventy_disciples, pem_pos=3, final_scores=-1.6661098, final_local_scores=-1.7383049, pem=-1.5857146, similarity=-39.161407]

gt: 363753 Apostle_(Latter_Day_Saints), pem_pos=0, final_scores=-1.7522428, final_local_scores=-1.9413829, pem=-0.4959507, similarity=-51.14797]

3: 2047259 Mission_president

[winner: 6398179 The_President, pem_pos=0, final_scores=-0.9221572, final_local_scores=-0.8487086, pem=-0.3939714, similarity=-19.094969]

RECALL MISS: 2047259 Mission_president]

4: 5810892 Mission_(LDS_Church)

[winner: 22673654 Mission, California, pem_pos=11, final_scores=-1.1115706, final_local_scores=-1.0398118, pem=-3.2960024, similarity=-9.016171]

gt: 5810892 Mission_(LDS_Church), pem_pos=1, final_scores=-1.3697537, final_local_scores=-1.5152113, pem=-1.8259888, similarity=-31.1972]

5: 1072725 George_Q._Morris

[winner: 1072725 George_Q._Morris, pem_pos=0,


```

final_scores=-0.15801772, final_local_scores=-0.13247325,
pem=0.0, similarity=0.13644704|
    gt: 1072725 George_Q._Morris, pem_pos=0,
final_scores=-0.15801772, final_local_scores=-0.13247325,
pem=0.0, similarity=0.13644704 ]
6: 363559 Pancreatic_cancer
    [winner: 5616156 Pancreatic_veins, pem_pos=2,
final_scores=-1.0695001, final_local_scores=-0.9201593,
pem=-2.3215952, similarity=-10.744481|
    gt: 363559 Pancreatic_cancer, pem_pos=0,
final_scores=-1.2073425, final_local_scores=-1.25734,
pem=0.0, similarity=-33.431816 ]
    [winner: 105219 Cancer, pem_pos=0,
final_scores=-1.0953666, final_local_scores=-1.1367044,
pem=-0.023738135, similarity=-29.702724|
    RECALL MISS: 363559 Pancreatic_cancer]
1: 31716 Utah
    [pem_pos=0, final_scores=0.4459409,
final_local_scores=0.45202512, pem=-0.2611261,
similarity=18.999088]
2: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints
    [pem_pos=0, final_scores=0.2776624,
final_local_scores=0.22261146, pem=0.0, similarity=10.7328825]
3: 9239 Europe
    [pem_pos=0, final_scores=0.5160384,
final_local_scores=0.5906304, pem=-0.086338006,
similarity=22.184822]
4: 1783358 Francis_M._Lyman
    [pem_pos=0, final_scores=0.2137913,
final_local_scores=0.23705661, pem=0.0, similarity=11.163955]
5: 19004 Moscow
    [pem_pos=0, final_scores=0.437518,
final_local_scores=0.49691296, pem=-0.106582455,
similarity=19.4982]
6: 25391 Russia
    [pem_pos=0, final_scores=0.77533764,
final_local_scores=0.8712048, pem=-0.20299491,
similarity=31.192118]
7: 21023 Mormonism
    [pem_pos=0, final_scores=0.25734705,
final_local_scores=0.25654858, pem=-0.3136064,
similarity=13.451076]
8: 10577 Finland
    [pem_pos=0, final_scores=0.55205864,
final_local_scores=0.6245553, pem=-0.2445894,
similarity=24.057804]
9: 7123926 Utah_House_of_Representatives
    [pem_pos=0, final_scores=0.08342664,
final_local_scores=0.10613124, pem=0.0, similarity=7.2568808]
10: 1046373 Deseret_News
    [pem_pos=0, final_scores=0.15046774,
final_local_scores=0.17606995, pem=0.0, similarity=9.343991]

```

11: 53837 Salt_Lake_City
[pem_pos=0, final_scores=0.35566735,
final_local_scores=0.34120837, pem=-0.03391632,
similarity=14.456493]
1: 13510620 Joseph_J._Cannon
[pem_pos=0, final_scores=0.14240755,
final_local_scores=0.11652532, pem=0.0, similarity=7.567061]
2: 13510620 Joseph_J._Cannon
[pem_pos=0, final_scores=0.14020343,
final_local_scores=0.11411765, pem=0.0, similarity=7.4952116]
3: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints
[pem_pos=0, final_scores=0.2733627,
final_local_scores=0.2179147, pem=0.0, similarity=10.592722]
4: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints
[pem_pos=0, final_scores=0.3111841,
final_local_scores=0.2592288, pem=0.0, similarity=11.825617]
5: 13510620 Joseph_J._Cannon
[pem_pos=0, final_scores=0.03947352,
final_local_scores=-0.0031246394, pem=0.0,
similarity=3.9964676]
6: 13510620 Joseph_J._Cannon
[pem_pos=0, final_scores=0.2004415,
final_local_scores=0.17991859, pem=0.0, similarity=9.458842]
7: 13510620 Joseph_J._Cannon
[pem_pos=0, final_scores=0.12534867,
final_local_scores=0.09789108, pem=0.0, similarity=7.0109777]
8: 13510620 Joseph_J._Cannon
[pem_pos=0, final_scores=0.14538403,
final_local_scores=0.119776666, pem=0.0, similarity=7.664088]
9: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints
[pem_pos=0, final_scores=0.33718807,
final_local_scores=0.28763422, pem=0.0, similarity=12.673292]
10: 5935 The_Church_of_Jesus_Christ_of_Latter-day_Saints
[pem_pos=0, final_scores=0.27546966,
final_local_scores=0.22021624, pem=0.0, similarity=10.661405]
11: 31717 United_Kingdom
[pem_pos=0, final_scores=0.15694754,
final_local_scores=0.16809765, pem=-0.48936236,
similarity=11.767311]
12: 13510620 Joseph_J._Cannon
[pem_pos=0, final_scores=0.17594661,
final_local_scores=0.15316164, pem=0.0, similarity=8.660362]
13: 31716 Utah
[pem_pos=0, final_scores=0.49500954,
final_local_scores=0.5056251, pem=-0.2611261,
similarity=20.59862]
14: 13510620 Joseph_J._Cannon
[pem_pos=0, final_scores=0.09820293,
final_local_scores=0.06823848, pem=0.0, similarity=6.1260853]
15: 13510620 Joseph_J._Cannon
[pem_pos=0, final_scores=0.07973145,
final_local_scores=0.048061207, pem=0.0, similarity=5.5239553]

Sample: b'Jan-Erasmus_Quellinus'

[1 Jan-Erasmus Quellinus 1] -LRB- 1 Antwerp , 1634 -- [1 1 Mechelen 1] , 11 March 1715 -RRB- was a of large 2 altarpiece s and 3 histories influenced by [2 2 Paolo Veronese 2] and his father . Born into a family of artists that included his grandfather 4 Erasmus Quellinus I , father , and uncle , [2 Jan-Erasmus 2] 's early career was shaped by a trip to [3 3 Italy 3] . Drawings from this period show careful study of the late [3 5 Italian 3] Renaissance painter [4 Veronese 4] . Time in and around [4 4 Venice 4] in 1660 -- 1661 also introduced him the architecture of that informed decorative motifs used in his later compositions . Quellinus also spent time in [5 5 Rome 5] , where he was a member of the 6 Bentvueghels known as `` Cederboom '' -LRB- 7 cedar 8 tree -RRB- . He was back in [5 Antwerp 5] in 1661 , when he became a master in the city 's . Shortly afterwards he married the daughter of [6 9 David Teniers 6] the Younger , Cornelia . Quellinus painted large altarpieces in [7 Antwerp 7] , including several works for the city 's St. Michael 's 10 Abbey , and for churches and abbeys in [6 6 Brabant 6] . In 1680 he travelled to [7 7 Vienna 7] , where he was 11 court painter to [8 8 Emperor Leopold I 8] . Among the works made in that capacity were fifteen ceiling paintings depicting the Life of [9 9 Charles V 9]

1: 156169 Mechelen

2: 236906 Paolo_Veronese

3: 14532 Italy

4: 32616 Venice

5: 25458 Rome

6: 4589770 Duchy_of_Brabant

7: 55866 Vienna

8: 148788 Leopold_I,_Holy_Roman_Emperor

9: 70716 Charles_V,_Holy_Roman_Emperor

1: 32149462 Antwerp

[winner: 32149462 Antwerp, pem_pos=0, final_scores=-0.079659715, final_local_scores=-0.15976158, pem=-0.11904648, similarity=-0.030497849]

gt: 32149462 Antwerp, pem_pos=0, final_scores=-0.079659715, final_local_scores=-0.15976158, pem=-0.11904648, similarity=-0.030497849]

2: 195811 Altarpiece

[winner: 195811 Altarpiece, pem_pos=1, final_scores=-1.050554, final_local_scores=-1.1795568, pem=-0.015390082, similarity=-31.02692]

gt: 195811 Altarpiece, pem_pos=1, final_scores=-1.050554, final_local_scores=-1.1795568, pem=-0.015390082, similarity=-31.02692]

3: 14051 History_painting

[winner: 752522 Histories_(Herodotus), pem_pos=0, final_scores=-1.2847925, final_local_scores=-1.3503823, pem=-0.39351246, similarity=-34.068405]

RECALL MISS: 14051 History_painting]

4: 14171282 Erasmus_Quellinus_I
 [winner: 14171282 Erasmus_Quellinus_I, pem_pos=0,
 final_scores=-0.04637499, final_local_scores=0.009228811,
 pem=0.0, similarity=4.365119|
 gt: 14171282 Erasmus_Quellinus_I, pem_pos=0,
 final_scores=-0.04637499, final_local_scores=0.009228811,
 pem=0.0, similarity=4.365119]

5: 11335226 Italian_Renaissance_painting
 [winner: 14532 Italy, pem_pos=0, final_scores=0.39248997,
 final_local_scores=0.3919797, pem=-0.75948656,
 similarity=19.917374|
 RECALL MISS: 11335226 Italian_Renaissance_painting]

[winner: 25532 Renaissance, pem_pos=0,
 final_scores=-0.10879643, final_local_scores=-0.17722666,
 pem=-0.26284593, similarity=0.23031273|
 RECALL MISS: 11335226 Italian_Renaissance_painting]

6: 10082112 Bentvueghels
 [winner: 10082112 Bentvueghels, pem_pos=0,
 final_scores=-0.028944805, final_local_scores=-0.019491516,
 pem=0.0, similarity=3.508047|
 gt: 10082112 Bentvueghels, pem_pos=0,
 final_scores=-0.028944805, final_local_scores=-0.019491516,
 pem=0.0, similarity=3.508047]

7: 10766148 Cedar_wood
 [winner: 15073875 Cedar,_Indiana, pem_pos=6,
 final_scores=-1.2832279, final_local_scores=-1.1511486,
 pem=-2.93761, similarity=-14.287676|
 RECALL MISS: 10766148 Cedar_wood]

8: 18955875 Tree
 [winner: 2140770 Family_tree_of_Abu_Bakr, pem_pos=7,
 final_scores=-0.9477745, final_local_scores=-0.81296045,
 pem=-2.3937075, similarity=-7.1532946|
 gt: 18955875 Tree, pem_pos=0, final_scores=-1.4066275,
 final_local_scores=-1.4024158, pem=-0.33544275,
 similarity=-35.93698]

9: 850430 David_Teniers_the_Younger
 [winner: 850430 David_Teniers_the_Younger, pem_pos=0,
 final_scores=0.33002186, final_local_scores=0.34382945,
 pem=-0.3294792, similarity=16.142029|
 gt: 850430 David_Teniers_the_Younger, pem_pos=0,
 final_scores=0.33002186, final_local_scores=0.34382945,
 pem=-0.3294792, similarity=16.142029]
 [winner: 79451 William_Pitt_the_Younger, pem_pos=5,
 final_scores=-0.50138956, final_local_scores=-0.5253063,
 pem=-2.4765384, similarity=1.8813262|
 RECALL MISS: 850430 David_Teniers_the_Younger]

10: 1315 Abbey
 [winner: 1315 Abbey, pem_pos=0, final_scores=-0.10306792,
 final_local_scores=-0.12750854, pem=-0.7359849,
 similarity=4.2870045|
 gt: 1315 Abbey, pem_pos=0, final_scores=-0.10306792,

```

final_local_scores=-0.12750854, pem=-0.7359849,
similarity=4.2870045 ]
11: 3273207 Court_painter
    [winner: 3172708 Painterwork, pem_pos=6,
final_scores=-0.99770105, final_local_scores=-0.8958697,
pem=-1.7917595, similarity=-12.900959]
    RECALL MISS: 3273207 Court_painter]
    [winner: 5237052 Alter_Hof, pem_pos=6,
final_scores=-1.3817135, final_local_scores=-1.437309,
pem=-2.759927, similarity=-23.793543]
    RECALL MISS: 3273207 Court_painter]
1: 156169 Mechelen
    [pem_pos=0, final_scores=0.40356725,
final_local_scores=0.39614785, pem=-0.034399055,
similarity=16.098623]
2: 236906 Paolo_Veronese
    [pem_pos=0, final_scores=0.2736178,
final_local_scores=0.25204542, pem=0.0, similarity=11.611251]
3: 14532 Italy
    [pem_pos=0, final_scores=0.60461706,
final_local_scores=0.6236963, pem=-0.20612453,
similarity=23.822992]
4: 32616 Venice
    [pem_pos=0, final_scores=0.5039436,
final_local_scores=0.4711252, pem=-0.18918262,
similarity=19.177832]
5: 25458 Rome
    [pem_pos=0, final_scores=0.56976765,
final_local_scores=0.54831904, pem=-0.2517334,
similarity=21.821611]
6: 4589770 Duchy_of_Brabant
    [pem_pos=0, final_scores=0.21827866,
final_local_scores=0.20074543, pem=-0.414512,
similarity=12.3345375]
7: 55866 Vienna
    [pem_pos=0, final_scores=0.654473,
final_local_scores=0.657771, pem=-0.12517071,
similarity=24.399609]
8: 148788 Leopold_I,_Holy_Roman_Emperor
    [pem_pos=0, final_scores=0.20298858,
final_local_scores=0.19168359, pem=0.0, similarity=9.809934]
9: 70716 Charles_V,_Holy_Roman_Emperor
    [pem_pos=0, final_scores=0.24236213,
final_local_scores=0.2229341, pem=0.0, similarity=10.742511]
1: 14170887 Jan-Erasmus_Quellinus
    [pem_pos=0, final_scores=0.2102878,
final_local_scores=0.23713443, pem=0.0, similarity=11.166277]
2: 14170887 Jan-Erasmus_Quellinus
    [pem_pos=0, final_scores=0.035284862,
final_local_scores=0.045970306, pem=0.0, similarity=5.4615583]
3: 14532 Italy
    [pem_pos=0, final_scores=0.39248997,

```

final_local_scores=0.3919797, pem=-0.75948656,
 similarity=19.917374]
 4: 236906 Paolo_Veronese
 [pem_pos=0, final_scores=0.215367,
 final_local_scores=0.1961984, pem=0.0, similarity=9.944665]
 5: 32149462 Antwerp
 [pem_pos=0, final_scores=0.43790513,
 final_local_scores=0.41094834, pem=-0.11904648,
 similarity=17.000624]
 6: 850430 David_Teniers_the_Younger
 [pem_pos=0, final_scores=0.33002186,
 final_local_scores=0.34382945, pem=-0.3294792,
 similarity=16.142029]
 7: 32149462 Antwerp
 [pem_pos=0, final_scores=0.40592313,
 final_local_scores=0.37601286, pem=-0.11904648,
 similarity=15.958081]

Sample: b'Jeremy_Davis'

Jeremy Davis is an [1 1 American 1] 1 bassist for the band [2
 2 Paramore 2] . In 2002 , at the age of 16 , he was living in
 2 Franklin , [1 Tennessee 1] , where he played in a funk
 cover band called [2 The Factory 2] where he met [3 3 Hayley W
 illiams 3] . Through Williams , Davis met the other members ,
 brothers [3 3 Josh Farro 3] and [4 4 Zac Farro 4] . In 2005 ,
 [5 John Janick 5] , founder of record label [4 4 Fueled by Ra
 men 4] , signed a contract with them . Davis also stated that
 his favorite bands are [5 5 Deftones 5] , [6 6 Death Cab for C
 utie 6] , 5 Thrice , [7 7 Sigur Rós 7] , 6 Paper Route and 7 M
 ew . On September 30 , 2011 , he married [6 British 6]
 actress Kathryn Camsey . Paramore was created in 8 Franklin ,
 [7 Tennessee 7] in 2004 by the two brothers [8 8 Josh Farro 8]
 -LRB- 9 lead guitar / -RRB- and [8 10 Zac Farro 8] -LRB- -
 RRB- . [9 11 Taylor York 9] was also a part of the band from
 the very beginning , but his parents wanted him to finish
 school first . Later , they asked 12 Hayley Williams -LRB- / 1
 3 keyboards -RRB- to join the band , and , through Hayley ,
 Jeremy Davis -LRB- 14 bass guitar -RRB- joined as well .
 Prior to forming [10 Paramore 10] , the other members of what
 was soon to be [11 Paramore 11] had been `` edgy about the
 whole female thing '' of having Williams as singer , but as
 they were good friends she began writing with them and
 eventually became a member . The band was eventually signed
 to a deal on 15 Fueled by Ramen . However , before they began
 recording , [12 Davis 12] left the band for unknown reasons .
 The song [9 9 All We Know 9] is about this . For this time ,
 Davis was replaced by John Hembree . The band released their
 first album , [10 10 All We Know Is Falling 10] , and one
 EP , 16 The Summer Tic , without him . However , he rejoined
 soon afterwards and was present on the band 's second album ,
 17 Riot ! . Davis also plays bass on the live albums 18 The Fi
 nal Riot ! and 19 Live in the UK . The band 's third album , 2

0 Brand New Eyes , was released on September 29 , 2009 .
Their fourth album , [13 21 Paramore 13] -LRB- album -RRB-

- 1: 3434750 United_States
- 2: 3565012 Paramore
- 3: 14227477 Josh_Farro
- 4: 991149 Fueled_by_Ramen
- 5: 103959 Deftones
- 6: 594504 Death_Cab_for_Cutie
- 7: 148705 Sigur_Rós
- 8: 14227477 Josh_Farro
- 9: 8779687 All_We_Know
- 10: 4435775 All_We_Know_Is_Falling
- 1: 60001 Bassist

[winner: 3916 Bass_guitar, pem_pos=1,
final_scores=-1.3051772, final_local_scores=-1.4492266,
pem=-0.5640744, similarity=-36.09058]

gt: 60001 Bassist, pem_pos=0, final_scores=-1.3957283,
final_local_scores=-1.5671535, pem=-0.33387414,
similarity=-40.86161]

- 2: 135422 Franklin, Tennessee

[winner: 30395 Tennessee, pem_pos=0,
final_scores=0.4551767, final_local_scores=0.530469,
pem=-0.27977526, similarity=21.441427]

RECALL MISS: 135422 Franklin, Tennessee]

[winner: 135422 Franklin, Tennessee, pem_pos=0,
final_scores=-0.24093448, final_local_scores=-0.25530693,
pem=-3.2882125, similarity=14.352655]

gt: 135422 Franklin, Tennessee, pem_pos=0,
final_scores=-0.24093448, final_local_scores=-0.25530693,
pem=-3.2882125, similarity=14.352655]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7373004, final_local_scores=-1.6967168,
pem=-4.6728287, similarity=-21.132133]

RECALL MISS: 135422 Franklin, Tennessee]

- 3: 9726633 Hayley_Williams

[winner: 3565012 Paramore, pem_pos=0,
final_scores=0.24214564, final_local_scores=0.16307145,
pem=0.0, similarity=8.95609]

gt: 9726633 Hayley_Williams, pem_pos=1,
final_scores=-0.07799798, final_local_scores=-0.14178452,
pem=-1.0011808, similarity=5.303157]

- 4: 9899554 Zac_Farro

[winner: 3565012 Paramore, pem_pos=0,
final_scores=0.37061435, final_local_scores=0.30277765,
pem=0.0, similarity=13.125202]

gt: 9899554 Zac_Farro, pem_pos=1,
final_scores=0.13365297, final_local_scores=0.11953926,
pem=-0.6283233, similarity=11.073922]

- 5: 617354 Thrice

[winner: 617354 Thrice, pem_pos=0,
final_scores=-0.027335584, final_local_scores=-0.075087026,
pem=0.0, similarity=1.8489664]

gt: 617354 Thrice, pem_pos=0, final_scores=-0.027335584,
 final_local_scores=-0.075087026, pem=0.0,
 similarity=1.8489664]
 6: 16850931 Paper_Route_(band)
 [winner: 16850931 Paper_Route_(band), pem_pos=0,
 final_scores=0.013794884, final_local_scores=-0.0005775243,
 pem=0.0, similarity=4.0724783]
 gt: 16850931 Paper_Route_(band), pem_pos=0,
 final_scores=0.013794884, final_local_scores=-0.0005775243,
 pem=0.0, similarity=4.0724783]
 7: 506187 Mew_(band)
 [winner: 428082 Mew_(Pokémon), pem_pos=1,
 final_scores=-0.31838456, final_local_scores=-0.22602656,
 pem=-0.57983243, similarity=0.4978486]
 gt: 506187 Mew_(band), pem_pos=0,
 final_scores=-0.4265045, final_local_scores=-0.4880895,
 pem=-0.4527097, similarity=-8.013946]
 8: 135422 Franklin,_Tennessee
 [winner: 30395 Tennessee, pem_pos=0,
 final_scores=0.42177457, final_local_scores=0.49398226,
 pem=-0.27977526, similarity=20.35259]
 RECALL MISS: 135422 Franklin,_Tennessee]
 [winner: 135422 Franklin,_Tennessee, pem_pos=0,
 final_scores=-0.09558672, final_local_scores=-0.09653658,
 pem=-3.2882125, similarity=19.09068]
 gt: 135422 Franklin,_Tennessee, pem_pos=0,
 final_scores=-0.09558672, final_local_scores=-0.09653658,
 pem=-3.2882125, similarity=19.09068]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7615991, final_local_scores=-1.7232594,
 pem=-4.6728287, similarity=-21.924213]
 RECALL MISS: 135422 Franklin,_Tennessee]
 9: 440525 Lead_guitar
 [winner: 9611790 Epiphone_Supernova, pem_pos=8,
 final_scores=-1.1625111, final_local_scores=-1.17078,
 pem=-2.0793161, similarity=-19.541048]
 gt: 440525 Lead_guitar, pem_pos=2,
 final_scores=-1.7906601, final_local_scores=-1.9874778,
 pem=-1.951711, similarity=-44.606876]
 [winner: 13007240 Lead_(engineering), pem_pos=6,
 final_scores=-1.2395583, final_local_scores=-1.1188011,
 pem=-2.1831918, similarity=-17.425005]
 gt: 440525 Lead_guitar, pem_pos=1,
 final_scores=-1.6463103, final_local_scores=-1.8297977,
 pem=-1.8187729, similarity=-40.624332]
 10: 9899554 Zac_Farro
 [winner: 3565012 Paramore, pem_pos=0,
 final_scores=0.22005315, final_local_scores=0.13831243,
 pem=0.0, similarity=8.217232]
 gt: 9899554 Zac_Farro, pem_pos=1,
 final_scores=0.05675651, final_local_scores=0.03554155,
 pem=-0.6283233, similarity=8.567264]

11: 21066232 Taylor_York
[winner: 3565012 Paramore, pem_pos=1,
final_scores=0.15169035, final_local_scores=0.06426276,
pem=-0.022356918, similarity=6.1290226|
gt: 21066232 Taylor_York, pem_pos=0,
final_scores=-0.02998656, final_local_scores=-0.030986972,
pem=0.0, similarity=3.165]
12: 9726633 Hayley_Williams
[winner: 3565012 Paramore, pem_pos=0,
final_scores=-0.09853196, final_local_scores=-0.2116029,
pem=0.0, similarity=-2.2249393|
gt: 9726633 Hayley_Williams, pem_pos=1,
final_scores=-0.29287153, final_local_scores=-0.37700453,
pem=-1.0011808, similarity=-1.7162777]
13: 49842 Keyboard_instrument
[winner: 49842 Keyboard_instrument, pem_pos=0,
final_scores=-1.0863109, final_local_scores=-1.2084839,
pem=-0.2511502, similarity=-30.608065|
gt: 49842 Keyboard_instrument, pem_pos=0,
final_scores=-1.0863109, final_local_scores=-1.2084839,
pem=-0.2511502, similarity=-30.608065]
14: 3916 Bass_guitar
[winner: 9611790 Epiphone_Supernova, pem_pos=8,
final_scores=-1.0544698, final_local_scores=-1.0527613,
pem=-2.0793161, similarity=-16.01914|
RECALL MISS: 3916 Bass_guitar]
[winner: 3916 Bass_guitar, pem_pos=0,
final_scores=-1.295162, final_local_scores=-1.4382865,
pem=-0.38971284, similarity=-36.712307|
gt: 3916 Bass_guitar, pem_pos=0, final_scores=-1.295162,
final_local_scores=-1.4382865, pem=-0.38971284,
similarity=-36.712307]
15: 991149 Fueled_by_Ramen
[winner: 991149 Fueled_by_Ramen, pem_pos=0,
final_scores=-0.01916638, final_local_scores=-0.075781055,
pem=0.0, similarity=1.828255|
gt: 991149 Fueled_by_Ramen, pem_pos=0,
final_scores=-0.01916638, final_local_scores=-0.075781055,
pem=0.0, similarity=1.828255]
16: 5712259 The_Summer_Tic_EP
[winner: 5712259 The_Summer_Tic_EP, pem_pos=0,
final_scores=-0.20343077, final_local_scores=-0.14503577,
pem=0.0, similarity=-0.23844379|
gt: 5712259 The_Summer_Tic_EP, pem_pos=0,
final_scores=-0.20343077, final_local_scores=-0.14503577,
pem=0.0, similarity=-0.23844379]
17: 10047267 Riot!
[winner: 2633338 Riot_(band), pem_pos=1,
final_scores=-0.53226167, final_local_scores=-0.60595435,
pem=-1.2112795, similarity=-7.406047|
gt: 10047267 Riot!, pem_pos=8, final_scores=-0.7429598,
final_local_scores=-0.8288446, pem=-2.670878,

```

similarity=-6.1200147 ]
    [winner: 3243047 !_(album), pem_pos=3,
final_scores=-1.2995096, final_local_scores=-1.330524,
pem=-2.4849067, similarity=-22.102468]
    RECALL MISS: 10047267 Riot!]
18: 20116308 The_Final_Riot!
    [winner: 20116308 The_Final_Riot!, pem_pos=0,
final_scores=-0.21142703, final_local_scores=-0.17328683,
pem=0.0, similarity=-1.0815114]
    gt: 20116308 The_Final_Riot!, pem_pos=0,
final_scores=-0.21142703, final_local_scores=-0.17328683,
pem=0.0, similarity=-1.0815114 ]
    [winner: 3243047 !_(album), pem_pos=3,
final_scores=-1.4650238, final_local_scores=-1.511323,
pem=-2.4849067, similarity=-27.497873]
    RECALL MISS: 20116308 The_Final_Riot!]
19: 15719923 Live_in_the_UK_2008
    [winner: 31717 United_Kingdom, pem_pos=0,
final_scores=-0.22158064, final_local_scores=-0.22846724,
pem=-0.19850598, similarity=-1.6486988]
    RECALL MISS: 15719923 Live_in_the_UK_2008]
    [winner: 147871 Live_(band), pem_pos=0,
final_scores=-0.29619855, final_local_scores=-0.34887064,
pem=-1.2837523, similarity=0.6599624]
    RECALL MISS: 15719923 Live_in_the_UK_2008]
    [winner: 2118516 German_federal_election,_March_1933,
pem_pos=0, final_scores=-1.7034998,
final_local_scores=-1.718415, pem=-0.79592705,
similarity=-42.862843]
    RECALL MISS: 15719923 Live_in_the_UK_2008]
20: 23388310 Brand_New_Eyes
    [winner: 23388310 Brand_New_Eyes, pem_pos=0,
final_scores=0.017806813, final_local_scores=0.0005361289,
pem=0.0, similarity=4.105712]
    gt: 23388310 Brand_New_Eyes, pem_pos=0,
final_scores=0.017806813, final_local_scores=0.0005361289,
pem=0.0, similarity=4.105712 ]
21: 37852474 Paramore_(album)
    [winner: 3565012 Paramore, pem_pos=0,
final_scores=0.094428316, final_local_scores=-0.00082311034,
pem=0.0, similarity=4.06515]
    gt: 37852474 Paramore_(album), pem_pos=6,
final_scores=-1.0299523, final_local_scores=-1.1435871,
pem=-5.923085, similarity=2.1734514 ]
    [winner: 528282 Album, pem_pos=0,
final_scores=-1.1212332, final_local_scores=-1.2754972,
pem=-0.11110695, similarity=-33.36945]
    RECALL MISS: 37852474 Paramore_(album)]
1: 3434750 United_States
    [pem_pos=0, final_scores=0.17764817,
final_local_scores=0.21903351, pem=-0.12994486,
similarity=11.33277]

```

2: 3565012 Paramore
[pem_pos=0, final_scores=0.2238354,
final_local_scores=0.14307028, pem=0.0, similarity=8.359216]
3: 14227477 Josh_Farro
[pem_pos=0, final_scores=0.31237262,
final_local_scores=0.3175797, pem=0.0, similarity=13.566924]
4: 991149 Fueled_by_Ramen
[pem_pos=0, final_scores=0.07788779,
final_local_scores=0.037377536, pem=0.0, similarity=5.205133]
5: 103959 Deftones
[pem_pos=0, final_scores=0.2570135,
final_local_scores=0.2253446, pem=0.0, similarity=10.8144455]
6: 594504 Death_Cab_for_Cutie
[pem_pos=0, final_scores=0.24108367,
final_local_scores=0.20403707, pem=0.0, similarity=10.178586]
7: 148705 Sigur_Rós
[pem_pos=0, final_scores=0.20971839,
final_local_scores=0.20384431, pem=-0.0017096702,
similarity=10.182131]
8: 14227477 Josh_Farro
[pem_pos=0, final_scores=0.24978898,
final_local_scores=0.24921662, pem=0.0, similarity=11.526834]
9: 8779687 All_We_Know
[pem_pos=0, final_scores=0.057948574,
final_local_scores=0.11880693, pem=0.0, similarity=7.635149]
10: 4435775 All_We_Know_Is_Falling
[pem_pos=0, final_scores=0.03462781,
final_local_scores=0.041455626, pem=0.0, similarity=5.3268313]
1: 30395 Tennessee
[pem_pos=0, final_scores=0.4551767,
final_local_scores=0.530469, pem=-0.27977526,
similarity=21.441427]
2: 936533 The_Factory
[pem_pos=0, final_scores=0.13093598,
final_local_scores=0.1867724, pem=-0.095310144,
similarity=10.181685]
3: 3565012 Paramore
[pem_pos=0, final_scores=0.24214564,
final_local_scores=0.16307145, pem=0.0, similarity=8.95609]
4: 3565012 Paramore
[pem_pos=0, final_scores=0.37061435,
final_local_scores=0.30277765, pem=0.0, similarity=13.125202]
5: 991149 Fueled_by_Ramen
[pem_pos=0, final_scores=0.30105287,
final_local_scores=0.28115144, pem=0.0, similarity=12.479833]
6: 31717 United_Kingdom
[pem_pos=0, final_scores=0.44603437,
final_local_scores=0.5131439, pem=-0.48936236,
similarity=22.064178]
7: 30395 Tennessee
[pem_pos=0, final_scores=0.42177457,
final_local_scores=0.49398226, pem=-0.27977526,

```

similarity=20.35259]
8: 3565012 Paramore
    [pem_pos=0, final_scores=0.22005315,
final_local_scores=0.13831243, pem=0.0, similarity=8.217232]
9: 3565012 Paramore
    [pem_pos=1, final_scores=0.15169035,
final_local_scores=0.06426276, pem=-0.022356918,
similarity=6.1290226]
10: 3565012 Paramore
    [pem_pos=0, final_scores=0.04180537,
final_local_scores=-0.05830569, pem=0.0, similarity=2.3497548]
11: 3565012 Paramore
    [pem_pos=0, final_scores=0.051544383,
final_local_scores=-0.047667302, pem=0.0,
similarity=2.6672256]
12: 3565012 Paramore
    [pem_pos=1, final_scores=0.07248883,
final_local_scores=-0.024788663, pem=-0.018660422,
similarity=3.451448]
13: 3565012 Paramore
    [pem_pos=0, final_scores=0.094428316,
final_local_scores=-0.00082311034, pem=0.0,
similarity=4.06515]

```

Sample: b'June_Preisser'

[1 June Preisser 1] -LRB- June 26 , 1920 -- September 19 , 1984 -RRB- was an [2 American 2] actress , briefly popular in musical films during the late 1930s and early 1940s , many of which capitalized on her skills as an acrobat . Born in [1 1 New Orleans 1] , [2 2 Louisiana 2] , [3 Preisser 3] was one of six children , and was an underweight baby . Her parents sent her to an athletic club at an early age , in an attempt to build her strength . There she , and her sister Cherry , learnt acrobatics . Their mother was keen to have them follow a career in showbusiness , especially when their father died suddenly , leaving the family with few options to make a living . When Preisser was nine years old an actor noticed the two sisters performing acrobatics on a sidewalk near their home , and his interest in them eventually led to them working in 1 vaudeville , and later for the [3 3 Ziegfeld Folies 3] in 1934 and 1936 . The [4 Preisser 4] sisters were successful in the [5 United States 5] , and also performed in [6 Europe 6] , most notably for . Cherry retired in 1938 following her marriage , and June was signed to a film contract by . Her first film Dancing Co-Ed -LRB- 1939 -RRB- provided only a small part , but her next film , [7 2 Babes in Arms 7] -LRB- 1939 -RRB- , gave her a significant role opposite [4 4 Mickey Rooney 4] and [5 5 Judy Garland 5] . She performed with [8 Rooney 8] and [9 Garland 9] again in 3 Strike Up the Band -LRB- 1940 -RRB- , and with [10 Rooney 10] in two `` [11 Andy Hardy 11] '' films , Judge Hardy and Son -LRB- 1939 -RRB- and [12 4 Andy Hardy 12] 's Private Secretary

-LRB- 1941 -RRB- . Gallant Sons -LRB- 1940 -RRB- placed her in a comedic murder mystery , and she played her first lead role opposite [6 6 Jimmy Lydon 6] in -LRB- 1941 -RRB- , and followed this with 5 Sweater Girl -LRB- 1942 -RRB- , opposite [7 7 Eddie Bracken 7] . She continued her career following her marriage in 1942 to J. Moss Terry , and the birth of a son , Ricky . By this time , [13 MGM 13] had little interest in promoting her , and she left to work for [8 8 Monogram Pictures 8] . She continued to appear in musical comedies over the next few years , and played the character `` Dodie Rogers '' in seven `` high school '' comedy films with [9 9 Frankie Darro 9] and [10 10 Noel Neill 10] from 1946 to 1948 . Her final film was -LRB- 1948 -RRB- , and after appearing in a [11 11 Los Angeles 11] theater production of 6 Annie Get Your Gun

1: 53842 New_Orleans

2: 18130 Louisiana

3: 264637 Ziegfeld_Follies

4: 87624 Mickey_Rooney

5: 58719 Judy_Garland

6: 11469494 Jimmy_Lydon

7: 149731 Eddie_Bracken

8: 1588593 Monogram_Pictures

9: 1972886 Frankie_Darro

10: 2062226 Noel_Neill

11: 18110 Los_Angeles

1: 48235 Vaudeville

[winner: 48235 Vaudeville, pem_pos=0, final_scores=-0.6283689, final_local_scores=-0.7503873, pem=-0.01125163, similarity=-18.242153]

gt: 48235 Vaudeville, pem_pos=0, final_scores=-0.6283689, final_local_scores=-0.7503873, pem=-0.01125163, similarity=-18.242153]

2: 9112052 Babes_in_Arms_(film)

[winner: 145622 Babes_in_Arms, pem_pos=0, final_scores=0.12873988, final_local_scores=0.12635529, pem=0.0, similarity=7.8604064]

gt: 9112052 Babes_in_Arms_(film), pem_pos=1, final_scores=-0.02671367, final_local_scores=-0.046795793, pem=-0.51860774, similarity=5.5135026]

3: 20847232 Strike_Up_the_Band_(film)

[winner: 11930055 Strike_Up_the_Band_(song), pem_pos=1, final_scores=-0.18850929, final_local_scores=-0.13659003, pem=-0.59939224, similarity=3.2731822]

gt: 20847232 Strike_Up_the_Band_(film), pem_pos=0, final_scores=-0.19990066, final_local_scores=-0.20086485, pem=-0.19522493, similarity=-0.8428315]

4: 17617782 Andy_Hardy's_Private_Secretary

[winner: 661150 Andy_Hardy, pem_pos=0, final_scores=0.19251744, final_local_scores=0.19147018, pem=0.0, similarity=9.803565]

RECALL MISS: 17617782 Andy_Hardy's_Private_Secretary]

[winner: 2244472 Private_Secretary, pem_pos=0,

```

final_scores=-0.27545518, final_local_scores=-0.23690549,
pem=-0.06609008, similarity=-2.6206112|
    RECALL MISS: 17617782 Andy_Hardy's_Private_Secretary]
    [winner: 1415027 David_Berman_(musician), pem_pos=5,
final_scores=-2.4635336, final_local_scores=-2.6469462,
pem=-3.0012724, similarity=-58.579056|
    RECALL MISS: 17617782 Andy_Hardy's_Private_Secretary]
5: 34543746 Sweater_Girl_(film)
    [winner: 1921764 Sweater_girl, pem_pos=0,
final_scores=-0.024531066, final_local_scores=0.061149254,
pem=-0.09015111, similarity=6.404784|
    gt: 34543746 Sweater_Girl_(film), pem_pos=1,
final_scores=-0.45976475, final_local_scores=-0.43279636,
pem=-2.451005, similarity=4.5031524 ]
6: 286348 Annie_Get_Your_Gun_(musical)
    [winner: 286348 Annie_Get_Your_Gun_(musical), pem_pos=0,
final_scores=-0.010864541, final_local_scores=-0.06600616,
pem=0.0, similarity=2.1199574|
    gt: 286348 Annie_Get_Your_Gun_(musical), pem_pos=0,
final_scores=-0.010864541, final_local_scores=-0.06600616,
pem=0.0, similarity=2.1199574 ]
1: 53842 New_Orleans
    [pem_pos=0, final_scores=0.5691679,
final_local_scores=0.6143916, pem=-0.12199242,
similarity=23.087797]
2: 18130 Louisiana
    [pem_pos=0, final_scores=0.44948006,
final_local_scores=0.5234734, pem=-0.17745006,
similarity=20.676203]
3: 264637 Ziegfeld_Follies
    [pem_pos=0, final_scores=0.2510904,
final_local_scores=0.21179181, pem=0.0, similarity=10.410003]
4: 87624 Mickey_Rooney
    [pem_pos=0, final_scores=0.29179066,
final_local_scores=0.22665554, pem=0.0, similarity=10.853566]
5: 58719 Judy_Garland
    [pem_pos=0, final_scores=0.37699956,
final_local_scores=0.32436877, pem=0.0, similarity=13.769525]
6: 11469494 Jimmy_Lydon
    [pem_pos=0, final_scores=0.08557825,
final_local_scores=0.07086575, pem=0.0, similarity=6.2044883]
7: 149731 Eddie_Bracken
    [pem_pos=0, final_scores=0.36505622,
final_local_scores=0.37418777, pem=0.0, similarity=15.256222]
8: 1588593 Monogram_Pictures
    [pem_pos=0, final_scores=0.29853684,
final_local_scores=0.2856105, pem=0.0, similarity=12.6129]
9: 1972886 Frankie_Darro
    [pem_pos=0, final_scores=0.17830865,
final_local_scores=0.1939817, pem=0.0, similarity=9.878514]
10: 2062226 Noel_Neill
    [pem_pos=0, final_scores=0.14328144,

```

```

final_local_scores=0.15632597, pem=0.0, similarity=8.754791]
11: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.5713039,
final_local_scores=0.5856378, pem=-0.100652285,
similarity=22.113676]
1: 15566344 June_Preisser
    [pem_pos=0, final_scores=0.0567462,
final_local_scores=0.042699367, pem=0.0, similarity=5.3639474]
2: 3434750 United_States
    [pem_pos=0, final_scores=0.52174324,
final_local_scores=0.57748675, pem=-0.12994486,
similarity=22.02973]
3: 15566344 June_Preisser
    [pem_pos=0, final_scores=0.037145063,
final_local_scores=0.021288112, pem=0.0, similarity=4.7249928]
4: 15566344 June_Preisser
    [pem_pos=0, final_scores=0.0529965,
final_local_scores=0.03860338, pem=0.0, similarity=5.241715]
5: 3434750 United_States
    [pem_pos=0, final_scores=0.47344613,
final_local_scores=0.5247295, pem=-0.073457584,
similarity=20.14816]
6: 9239 Europe
    [pem_pos=0, final_scores=0.4983474,
final_local_scores=0.6236378, pem=-0.086338006,
similarity=23.169828]
7: 145622 Babes_in_Arms
    [pem_pos=0, final_scores=0.12873988,
final_local_scores=0.12635529, pem=0.0, similarity=7.8604064]
8: 58719 Judy_Garland
    [pem_pos=1, final_scores=0.4263546,
final_local_scores=0.37865102, pem=0.0, similarity=15.389415]
9: 58719 Judy_Garland
    [pem_pos=0, final_scores=0.6449489,
final_local_scores=0.6207854, pem=0.0, similarity=22.615189]
10: 58719 Judy_Garland
    [pem_pos=1, final_scores=0.35818464,
final_local_scores=0.30396935, pem=0.0, similarity=13.160765]
11: 661150 Andy_Hardy
    [pem_pos=0, final_scores=0.20038505,
final_local_scores=0.20006433, pem=0.0, similarity=10.060032]
12: 661150 Andy_Hardy
    [pem_pos=0, final_scores=0.19251744,
final_local_scores=0.19147018, pem=0.0, similarity=9.803565]
13: 58819 Metro-Goldwyn-Mayer
    [pem_pos=0, final_scores=0.34661442,
final_local_scores=0.2964594, pem=-0.020185718,
similarity=13.046426]

```

Sample: b'Sopwith_1919_Schneider_Cup_Seaplane'

The Sopwith Schneider of 1919 was a [1 British 1] racing 1 se
aplane . It was a single seat 2 biplane intended to compete

in the 1919 [1 1 Schneider Trophy 1] . After this race was abandoned due to fog , the Schneider was rebuilt into a landplane racer as the Sopwith Rainbow , being destroyed in a crash in 1923 . In 1919 , it became possible to restart the [2 2 Schneider Trophy 2] races for 3 seaplane s , which had not been held since 1914 owing to the [3 3 First World War 3] . As it was last won by the development of the Sopwith Tabloid , the race was organised by the British 4 Royal Aero Club , and was planned to be held at [4 4 Bournemouth 4] on 10 September that year . In order to compete in the 1919 race , the [5 5 Sopwith Aviation Company 5] designed a small 5 float plane , powered by the new 6 Cosmos Jupiter 7 radial engine . It was of all-wooden construction , with single bay wings . The Sopwith entry , 8 registration G-EAKI , along with entries from 9 Supermarine -LRB- the 10 Sea Lion -RRB- and [6 6 Fairey 6] -LRB- the 11 Fairey III -RRB- was one of three [2 British 2] entries to compete in the race -LRB- the [3 12 Avro 3] 539 was eliminated prior to race-day -RRB- . On the day of the race , the weather was poor , with thick fog . The Sopwith , flown by [7 7 Harry Hawker 7] together with the Fairey entry , abandoned the race owing to the fog , while the [4 Supermarine 4] aircraft hit debris following alighting to try to find where it was on the course , sinking when it tried to land again . The only aircraft to complete the race , the [8 8 Italian 8] 13 SIAI S. 13 was disqualified as the pilot consistently missed one of the turning points , the race being declared void . In 1920 , the [5 Schneider Cup 5] racer was rebuilt as a landplane with a 14 tailwheel undercarriage , powered by a 320 hp -LRB- 239 kW -RRB- 15 ABC Dragonfly as the Jupiter engine which powered it in 1919 was unavailable , and known as the Sopwith Rainbow . It was entered into the 1920 , but was disqualified . It was rebuilt again in 1922 by 16 H.G. Hawker Engineering , -LRB- Sopwith Aviation having gone into 17 receivership

- 1: 56291 Schneider_Trophy
- 2: 56291 Schneider_Trophy
- 3: 4764461 World_War_I
- 4: 165530 Bournemouth
- 5: 382421 Sopwith_Aviation_Company
- 6: 508042 Fairey_Aviation_Company
- 7: 471396 Harry_Hawker
- 8: 14532 Italy

1: 291731 Seaplane

[winner: 3648322 Ocean_Falls_Water_Aerodrome, pem_pos=2, final_scores=-1.0628912, final_local_scores=-1.0406251, pem=-1.0978429, similarity=-20.994377]

gt: 291731 Seaplane, pem_pos=0, final_scores=-1.1481733, final_local_scores=-1.2976327, pem=-0.008760688, similarity=-34.586594]

2: 49201 Biplane

[winner: 49201 Biplane, pem_pos=0, final_scores=-1.1549373, final_local_scores=-1.2936757,

pem=-0.00053622795, similarity=-34.51323|
 gt: 49201 Biplane, pem_pos=0, final_scores=-1.1549373,
 final_local_scores=-1.2936757, pem=-0.00053622795,
 similarity=-34.51323]

3: 291731 Seaplane
 [winner: 3648322 Ocean_Falls_Water_Aerodrome, pem_pos=2,
 final_scores=-0.95830923, final_local_scores=-0.9263852,
 pem=-1.0978429, similarity=-17.58523|
 gt: 291731 Seaplane, pem_pos=0, final_scores=-1.0027863,
 final_local_scores=-1.1388195, pem=-0.008760688,
 similarity=-29.847292]

4: 330770 Royal_Aero_Club
 [winner: 330770 Royal_Aero_Club, pem_pos=0,
 final_scores=-0.31739622, final_local_scores=-0.3701383,
 pem=0.0, similarity=-6.955952|
 gt: 330770 Royal_Aero_Club, pem_pos=0,
 final_scores=-0.31739622, final_local_scores=-0.3701383,
 pem=0.0, similarity=-6.955952]

5: 583442 Floatplane
 [winner: 582772 Fairey_Seafox, pem_pos=1,
 final_scores=-1.0845791, final_local_scores=-1.1283975,
 pem=-0.69167006, similarity=-25.822508|
 gt: 583442 Floatplane, pem_pos=0,
 final_scores=-1.2773204, final_local_scores=-1.4182789,
 pem=-0.07359641, similarity=-37.83433]

6: 146920 Bristol_Jupiter
 [winner: 146920 Bristol_Jupiter, pem_pos=0,
 final_scores=-0.006123334, final_local_scores=0.031061307,
 pem=0.0, similarity=5.0166445|
 gt: 146920 Bristol_Jupiter, pem_pos=0,
 final_scores=-0.006123334, final_local_scores=0.031061307,
 pem=0.0, similarity=5.0166445]

7: 146924 Radial_engine
 [winner: 9640 Engine, pem_pos=0, final_scores=-1.1860857,
 final_local_scores=-1.2869663, pem=-0.6065824,
 similarity=-31.017242|
 gt: 146924 Radial_engine, pem_pos=10,
 final_scores=-1.5297369, final_local_scores=-1.656364,
 pem=-3.5225887, similarity=-26.183102]
 [winner: 690495 Radial_artery, pem_pos=2,
 final_scores=-1.3028578, final_local_scores=-1.1974111,
 pem=-1.6580532, similarity=-22.626667|
 gt: 146924 Radial_engine, pem_pos=0,
 final_scores=-1.5490887, final_local_scores=-1.6775029,
 pem=-0.80782056, similarity=-41.577263]

8: 2367169 Aircraft_registration
 [winner: 2367169 Aircraft_registration, pem_pos=0,
 final_scores=-1.4322653, final_local_scores=-1.5580776,
 pem=-0.3926665, similarity=-40.27105|
 gt: 2367169 Aircraft_registration, pem_pos=0,
 final_scores=-1.4322653, final_local_scores=-1.5580776,
 pem=-0.3926665, similarity=-40.27105]

9: 27548 Supermarine
[winner: 27548 Supermarine, pem_pos=0,
final_scores=0.008372918, final_local_scores=-0.031802267,
pem=0.0, similarity=3.1406698|
gt: 27548 Supermarine, pem_pos=0,
final_scores=0.008372918, final_local_scores=-0.031802267,
pem=0.0, similarity=3.1406698]
10: 16060145 Supermarine_Sea_Lion_I
[winner: 22338 Operation_Sea_Lion, pem_pos=1,
final_scores=-0.14066483, final_local_scores=-0.14130679,
pem=-1.2313182, similarity=6.568937|
gt: 16060145 Supermarine_Sea_Lion_I, pem_pos=7,
final_scores=-0.39502278, final_local_scores=-0.37502766,
pem=-1.8680931, similarity=3.0571203]
11: 984177 Fairey_III
[winner: 984177 Fairey_III, pem_pos=0,
final_scores=0.0018391162, final_local_scores=-0.005214587,
pem=0.0, similarity=3.9340992|
gt: 984177 Fairey_III, pem_pos=0,
final_scores=0.0018391162, final_local_scores=-0.005214587,
pem=0.0, similarity=3.9340992]
12: 16716846 Avro_539
[winner: 49812 Avro, pem_pos=0, final_scores=0.13002865,
final_local_scores=0.097491324, pem=-0.07870271,
similarity=7.4270463|
RECALL MISS: 16716846 Avro_539]
[winner: 35768 539, pem_pos=0, final_scores=-1.0882953,
final_local_scores=-0.987094, pem=-0.5319874,
similarity=-22.47411|
gt: 16716846 Avro_539, pem_pos=1,
final_scores=-1.3004595, final_local_scores=-1.3268945,
pem=-1.3414736, similarity=-28.212326]
13: 20913740 SIAI_S.13
[winner: 1239743 SIAI-Marchetti, pem_pos=1,
final_scores=-0.10791819, final_local_scores=-0.12158366,
pem=-0.47445795, similarity=3.04159|
RECALL MISS: 20913740 SIAI_S.13]
[winner: 4239171 S_(South_Korean_band), pem_pos=0,
final_scores=-0.859994, final_local_scores=-0.78569114,
pem=-0.96774405, similarity=-14.094135|
RECALL MISS: 20913740 SIAI_S.13]
[winner: 15816 June_13, pem_pos=7,
final_scores=-1.0055832, final_local_scores=-0.8167382,
pem=-3.3953273, similarity=-1.8190663|
RECALL MISS: 20913740 SIAI_S.13]
14: 1049599 Conventional_landing_gear
[winner: 1049599 Conventional_landing_gear, pem_pos=0,
final_scores=-0.8795453, final_local_scores=-0.9066162,
pem=0.0, similarity=-22.965523|
gt: 1049599 Conventional_landing_gear, pem_pos=0,
final_scores=-0.8795453, final_local_scores=-0.9066162,
pem=0.0, similarity=-22.965523]

15: 15986878 ABC_Dragonfly
[winner: 15986878 ABC_Dragonfly, pem_pos=0,
final_scores=-0.10492599, final_local_scores=-0.06251267,
pem=0.0, similarity=2.22421]
gt: 15986878 ABC_Dragonfly, pem_pos=0,
final_scores=-0.10492599, final_local_scores=-0.06251267,
pem=0.0, similarity=2.22421]
16: 5933791 Hawker_Aircraft
[winner: 5933791 Hawker_Aircraft, pem_pos=0,
final_scores=-0.009868607, final_local_scores=-0.03510327,
pem=0.0, similarity=3.0421615]
gt: 5933791 Hawker_Aircraft, pem_pos=0,
final_scores=-0.009868607, final_local_scores=-0.03510327,
pem=0.0, similarity=3.0421615]
17: 660248 Receivership
[winner: 2935092 Official_Receiver, pem_pos=0,
final_scores=-1.1541319, final_local_scores=-1.1714535,
pem=0.0, similarity=-30.868793]
gt: 660248 Receivership, pem_pos=1,
final_scores=-1.2213098, final_local_scores=-1.2953587,
pem=-0.019312786, similarity=-34.461346]
1: 56291 Schneider_Trophy
[pem_pos=0, final_scores=0.1362675,
final_local_scores=0.101609394, pem=0.0, similarity=7.12194]
2: 56291 Schneider_Trophy
[pem_pos=0, final_scores=0.24464543,
final_local_scores=0.21999583, pem=0.0, similarity=10.654827]
3: 4764461 World_War_I
[pem_pos=0, final_scores=0.29752797,
final_local_scores=0.30001995, pem=0.0, similarity=13.042907]
4: 165530 Bournemouth
[pem_pos=0, final_scores=0.30490184,
final_local_scores=0.33469766, pem=-0.46563098,
similarity=16.609932]
5: 382421 Sopwith_Aviation_Company
[pem_pos=0, final_scores=0.104948714,
final_local_scores=0.074277684, pem=0.0, similarity=6.3063073]
6: 508042 Fairey_Aviation_Company
[pem_pos=0, final_scores=0.09732689,
final_local_scores=0.060947686, pem=-0.46459824,
similarity=8.435071]
7: 471396 Harry_Hawker
[pem_pos=0, final_scores=0.19523902,
final_local_scores=0.21782464, pem=0.0, similarity=10.5900345]
8: 14532 Italy
[pem_pos=0, final_scores=0.655791,
final_local_scores=0.72807616, pem=-0.75948656,
similarity=29.947165]
1: 31717 United_Kingdom
[pem_pos=0, final_scores=0.4359259,
final_local_scores=0.44136757, pem=-0.48936236,
similarity=19.92223]

2: 31717 United_Kingdom
 [pem_pos=0, final_scores=0.4458546,
 final_local_scores=0.4522131, pem=-0.48936236,
 similarity=20.245882]
 3: 49812 Avro
 [pem_pos=0, final_scores=0.13002865,
 final_local_scores=0.097491324, pem=-0.07870271,
 similarity=7.4270463]
 4: 27548 Supermarine
 [pem_pos=0, final_scores=0.044700786,
 final_local_scores=0.021913022, pem=0.0, similarity=4.7436414]
 5: 56291 Schneider_Trophy
 [pem_pos=0, final_scores=0.24049698,
 final_local_scores=0.2154643, pem=0.0, similarity=10.519597]

Sample: b'Madarihat'

[1 Madarihat 1] is a small village located in the [2 1 Alipurduar 2] subdivision of [3 2 Jalpaiguri 3] district , [1 1 West Bengal 1] , [2 2 India 2] . This village is situated on the outskirts of [3 3 Jaldapara National Park 3] which is famous for [4 3 Indian 4] rhinoceros . This is one of the main tourist spot in [5 North Bengal 5] surrounded by tea gardens , forests , hill and small rivers and also having a great diversity in the population. Madarihat is also world famous for the local wooden furniture . The local train station , [6 Madarihat 6] , is connected by the newly converted . The train line leads through the forests and tea gardens of [4 4 North Bengal 4] . This train line is also connected to the major cities of [7 India 7] like [5 5 Delhi 5] , [6 6 Kolkata 6] , [7 7 Ranchi 7] and [8 8 Patna 8] . However , the important stations are [9 9 Birpara 9] , [10 10 Hasimara 10] both of which are 20 km away and [11 11 New Alipurduar 11] -LRB- 50 km away -RRB- , -LRB- 80 km -RRB- and -LRB- 140 km -RRB- . [12 12 Bagdogra Airport 12] is the nearest airport . -LRB- 140 km -RRB- It is located on the NH 31 -LRB- Siliguri-Hasimara -RRB- and is connected with [8 Siliguri 8] and other places in [9 North Bengal 9] . 4 North Bengal State Transport Corporation

1: 34040 West_Bengal
 2: 14533 India
 3: 2014790 Jaldapara_National_Park
 4: 2550666 North_Bengal
 5: 37756 Delhi
 6: 47905 Kolkata
 7: 315858 Ranchi
 8: 18691354 Patna
 9: 3566300 Birpara
 10: 4602450 Hasimara
 11: 2452038 Alipurduar
 12: 1533201 Bagdogra_Airport

1: 20785983 Alipurduar_subdivision
 [winner: 2452038 Alipurduar, pem_pos=0,

```

final_scores=0.33071303, final_local_scores=0.31163624,
pem=0.0, similarity=13.38956|
    gt: 20785983 Alipurduar_subdivision, pem_pos=1,
final_scores=-0.65797126, final_local_scores=-0.7224033,
pem=-4.067316, similarity=4.6504555 ]
    [winner: 353021 Homeomorphism(graph_theory), pem_pos=6,
final_scores=-1.6513531, final_local_scores=-1.5932069,
pem=-2.900422, similarity=-27.681814|
    RECALL MISS: 20785983 Alipurduar_subdivision]
2: 2934596 Jalpaiguri_district
    [winner: 2186661 Jalpaiguri, pem_pos=0,
final_scores=0.33382034, final_local_scores=0.31022167,
pem=0.0, similarity=13.347346|
    gt: 2934596 Jalpaiguri_district, pem_pos=1,
final_scores=0.04753293, final_local_scores=0.011269137,
pem=-1.1294467, similarity=10.568113 ]
    [winner: 13974770 Districts_of_Burma, pem_pos=15,
final_scores=-1.3620236, final_local_scores=-1.3155715,
pem=-3.551226, similarity=-15.857448|
    RECALL MISS: 2934596 Jalpaiguri_district]
3: 802149 Indian_rhinoceros
    [winner: 14533 India, pem_pos=0, final_scores=0.38440353,
final_local_scores=0.3643378, pem=-0.618294,
similarity=18.32466|
    RECALL MISS: 802149 Indian_rhinoceros]
    [winner: 802149 Indian_rhinoceros, pem_pos=0,
final_scores=-1.6202953, final_local_scores=-1.716783,
pem=0.0, similarity=-47.142517|
    gt: 802149 Indian_rhinoceros, pem_pos=0,
final_scores=-1.6202953, final_local_scores=-1.716783,
pem=0.0, similarity=-47.142517 ]
4: 5695601 North_Bengal_State_Transport_Corporation
    [winner: 5695601
North_Bengal_State_Transport_Corporation, pem_pos=0,
final_scores=-0.14267254, final_local_scores=-0.106201075,
pem=0.0, similarity=0.92046106|
    gt: 5695601 North_Bengal_State_Transport_Corporation,
pem_pos=0, final_scores=-0.14267254,
final_local_scores=-0.106201075, pem=0.0,
similarity=0.92046106 ]
1: 34040 West_Bengal
    [pem_pos=0, final_scores=0.4816094,
final_local_scores=0.43389738, pem=0.0, similarity=17.038076]
2: 14533 India
    [pem_pos=0, final_scores=0.5950926,
final_local_scores=0.59448355, pem=-0.13622136,
similarity=22.571081]
3: 2014790 Jaldapara_National_Park
    [pem_pos=0, final_scores=0.13598315,
final_local_scores=0.19702235, pem=0.0, similarity=9.969253]
4: 2550666 North_Bengal
    [pem_pos=0, final_scores=0.3000825,

```

final_local_scores=0.2746305, pem=-0.022622399,
similarity=12.408258]
5: 37756 Delhi
[pem_pos=0, final_scores=0.28037387,
final_local_scores=0.23843932, pem=-0.14447327,
similarity=11.990887]
6: 47905 Kolkata
[pem_pos=0, final_scores=0.46389085,
final_local_scores=0.43958133, pem=0.0, similarity=17.207697]
7: 315858 Ranchi
[pem_pos=0, final_scores=0.27190655,
final_local_scores=0.24703026, pem=0.0, similarity=11.461589]
8: 18691354 Patna
[pem_pos=0, final_scores=0.2798916,
final_local_scores=0.22922853, pem=0.0, similarity=10.930349]
9: 3566300 Birpara
[pem_pos=0, final_scores=0.28742898,
final_local_scores=0.31317586, pem=0.0, similarity=13.435505]
10: 4602450 Hasimara
[pem_pos=0, final_scores=0.18532662,
final_local_scores=0.24838015, pem=0.0, similarity=11.501872]
11: 2452038 Alipurduar
[pem_pos=0, final_scores=0.2829885,
final_local_scores=0.25950447, pem=0.0, similarity=11.833844]
12: 1533201 Bagdogra Airport
[pem_pos=0, final_scores=0.111039236,
final_local_scores=0.14799288, pem=0.0, similarity=8.506116]
1: 23252848 Madarihat
[pem_pos=0, final_scores=0.05760856,
final_local_scores=0.07596251, pem=-0.47398502,
similarity=8.93419]
2: 2452038 Alipurduar
[pem_pos=0, final_scores=0.33071303,
final_local_scores=0.31163624, pem=0.0, similarity=13.38956]
3: 2186661 Jalpaiguri
[pem_pos=0, final_scores=0.33382034,
final_local_scores=0.31022167, pem=0.0, similarity=13.347346]
4: 14533 India
[pem_pos=0, final_scores=0.38440353,
final_local_scores=0.3643378, pem=-0.618294,
similarity=18.32466]
5: 2550666 North_Bengal
[pem_pos=0, final_scores=0.30366164,
final_local_scores=0.2785402, pem=-0.022622399,
similarity=12.524932]
6: 23252848 Madarihat
[pem_pos=0, final_scores=0.124032184,
final_local_scores=0.14852023, pem=-0.47398502,
similarity=11.099457]
7: 14533 India
[pem_pos=0, final_scores=0.5264472,
final_local_scores=0.5194989, pem=-0.13622136,

```
similarity=20.33339]
8: 1129032 Siliguri
   [pem_pos=0, final_scores=0.33579284,
final_local_scores=0.3028778, pem=0.0, similarity=13.128191]
9: 2550666 North_Bengal
   [pem_pos=0, final_scores=0.29656053,
final_local_scores=0.2707833, pem=-0.022622399,
similarity=12.293451]
```

Sample: b'A_Trip_Down_Memory_Lane'

' A Trip Down Memory Lane ' is a 1965 1 experimental 2 collage film by [1 1 Arthur Lipsett 1] , created by editing together images and sound clips from over fifty years of 3 newsreel footage . The film was intended as an audiovisual tour of the 4 post-war 5 technocracy . In his notes for producer [2 2 Donald Brittain 2] about the film , [1 Lipsett 1] wrote that `` as science grows , religious belief seems to have diminished ... The new machines -LRB- of every description -RRB- are now invested with spiritual qualities . They have become ritualistic implements . '' [2 Lipsett 2] envisioned his film as a kind of cinematic 6 time capsule for future generations , and sub-titled the film accordingly , as `` Additional Material for a Time-Capsule . '' A Trip Down Memory Lane combines footage from a , religious procession , failed airflight , automotive and science experiments , , 7 skyscraper construction , military paraphernalia , [3 3 John D. Rockefeller 3] and scenes of leisure , [4 4 Richard Nixon 4] and scenes of war , and , and a . To create the film , [3 Lipsett 3] travelled from [5 5 Montreal 5] to [6 6 New York City 6] to obtain stock footage in which he had reduced to . He then drew new

```
1: 1479968 Arthur_Lipsett
2: 2107708 Donald_Brittain
3: 186300 John_D._Rockefeller
4: 25473 Richard_Nixon
5: 7954681 Montreal
6: 645042 New_York_City
```

```
1: 1056084 Experimental_film
   [winner: 6890038 Experimental_musical_instrument,
pem_pos=7, final_scores=-1.2529799,
final_local_scores=-1.2840016, pem=-2.1770387,
similarity=-22.388378]
   gt: 1056084 Experimental_film, pem_pos=11,
final_scores=-1.5812131, final_local_scores=-1.7277215,
pem=-3.1181169, similarity=-30.512138 ]
2: 34884755 Collage_film
   [winner: 34884755 Collage_film, pem_pos=0,
final_scores=-0.8569555, final_local_scores=-0.9130456,
pem=0.0, similarity=-23.157389]
   gt: 34884755 Collage_film, pem_pos=0,
final_scores=-0.8569555, final_local_scores=-0.9130456,
pem=0.0, similarity=-23.157389 ]
```

3: 170644 Newsreel
[winner: 73979 Hindenburg_disaster_newsreel_footage,
pem_pos=0, final_scores=-0.70448124,
final_local_scores=-0.6846319, pem=0.0,
similarity=-16.341068]
gt: 170644 Newsreel, pem_pos=1,
final_scores=-0.91897005, final_local_scores=-1.0052652,
pem=-0.001502629, similarity=-25.901236]
4: 4008941 Post-war
[winner: 4008941 Post-war, pem_pos=0,
final_scores=-1.2653275, final_local_scores=-1.3384988,
pem=-0.2788568, similarity=-34.337296]
gt: 4008941 Post-war, pem_pos=0,
final_scores=-1.2653275, final_local_scores=-1.3384988,
pem=-0.2788568, similarity=-34.337296]
5: 1325116 Technocracy
[winner: 1324774 Technocracy_movement, pem_pos=0,
final_scores=-0.99641335, final_local_scores=-1.0104053,
pem=0.0, similarity=-26.062798]
gt: 1325116 Technocracy, pem_pos=1,
final_scores=-1.3344938, final_local_scores=-1.3804483,
pem=-0.17901878, similarity=-36.132084]
6: 358837 Time_capsule
[winner: 3053322 Capsule_(pharmacy), pem_pos=2,
final_scores=-1.1162554, final_local_scores=-1.0742681,
pem=-0.894762, similarity=-23.102737]
RECALL MISS: 358837 Time_capsule]
[winner: 31600 Time_(magazine), pem_pos=0,
final_scores=-1.3481319, final_local_scores=-1.4712374,
pem=-0.31107646, similarity=-38.123264]
RECALL MISS: 358837 Time_capsule]
7: 29485 Skyscraper
[winner: 3475655 Skyscraper_design_and_construction,
pem_pos=0, final_scores=-0.74567837,
final_local_scores=-0.68472815, pem=0.0,
similarity=-16.34394]
RECALL MISS: 29485 Skyscraper]
1: 1479968 Arthur_Lipsett
[pem_pos=0, final_scores=0.25118583,
final_local_scores=0.22496966, pem=0.0, similarity=10.803256]
2: 2107708 Donald_Brittain
[pem_pos=0, final_scores=0.21721815,
final_local_scores=0.2902486, pem=0.0, similarity=12.75131]
3: 186300 John_D._Rockefeller
[pem_pos=0, final_scores=0.09233566,
final_local_scores=0.17860708, pem=0.0, similarity=9.419704]
4: 25473 Richard_Nixon
[pem_pos=0, final_scores=0.19764958,
final_local_scores=0.26723316, pem=0.0, similarity=12.064484]
5: 7954681 Montreal
[pem_pos=0, final_scores=0.50087863,
final_local_scores=0.56524533, pem=-0.11023275,


```

similarity=21.557224]
6: 645042 New_York_City
    [pem_pos=0, final_scores=0.5494421,
final_local_scores=0.62155974, pem=0.0, similarity=22.638296]
1: 1479968 Arthur_Lipsett
    [pem_pos=1, final_scores=0.16485928,
final_local_scores=0.13067105, pem=0.0, similarity=7.9891977]
2: 1479968 Arthur_Lipsett
    [pem_pos=1, final_scores=0.2126218,
final_local_scores=0.18284431, pem=0.0, similarity=9.546152]
3: 1479968 Arthur_Lipsett
    [pem_pos=1, final_scores=0.20648597,
final_local_scores=0.17614186, pem=0.0, similarity=9.346137]

```

Sample: b'2010_NASCAR_Canadian_Tire_Series_season'

The 2010 [1 NASCAR Canadian Tire Series 1] was the fourth racing season since the buy out of the [1 1 CASCAR Super Series 1] . In 2010 many procedures in the series remained the same , with all of the races airing on [2 2 TSN 2] , in tape-delayed one hour segments excluding [2 Montreal 2] and the [3 Toronto Indy 3] . [3 3 TSN 3] and RDS provided live coverage of the NAPA Autopro 100 from the [4 4 Circuit Gilles Villeneuve 4] on Sunday , August 29 and July 25 from the streets of [5 5 Toronto 5] . Major changes to the series included former [4 CASCAR 4] champion [6 6 Dave Whitlock 6] and his team closing down leaving many looking for rides . [5 Mark Dille 5] joined up with [7 7 Scott Steckly 7] to run full-time backed by [8 8 Canadian Tire 8] and [9 9 Dodge 9] and [6 Pierre Bourque 6] drove the 00 1 Aaron 's dream machine full-time . Last year 's champion [10 10 Andrew Ranger 10] did not compete full-time as his sponsors 2 Tide and 3 Wal-Mart backed out of a deal and he focused on the 4 NASCAR 5 K&N Pro Series East in [7 America 7] and elected to run only part-time in the [8 Canadian Tire Series 8] . The series kicked off at [11 11 Delaware Speedway 11] with [12 12 D. J. Kennington 12] defending his win in a close battle with [13 13 J. R. Fitzpatrick 13] who would end up winning the race at . [9 Kennington 9] won the third race of the year after leading only the last lap passing both [14 14 Kerry Micks 14] and [15 15 Scott Steckly 15] en route to victory at 6 Autodrome Saint-Eustache . [10 Ranger 10] led the majority of the race en route to victory in [11 Toronto 11] while [12 Fitzpatrick 12] won in [13 Edmonton 13] starting the western road swing . [14 Kennington 14] led the majority of the race at [16 16 Motoplex Speedway 16] but was overtaken by [17 17 Scott Steckly 17] late in the race for the win . Jason Hathaway finished second but was demoted to ninth for illegal tire changes . [15 Kennington 15] picked up where he left off , earning his third victory of the season at [18 18 Auto Clearing Motor Speedway 18] , from pole position . As a result , [16 Kennington 16] returned to the top of the championship standings . [17 Andrew Ranger 17] won in [18 7 Trois-Rivières 18] , giving [19

Fitzpatrick 19] the points lead by one point over [20
Kennington 20] following the race . [21 8 Don Thomson 21] , Jr
. . got back to his old ways by winning at the Mosport Oval .
At [22 Montreal 22] , [23 Andrew Ranger 23] bumped [24 Jason
Bowles 24] in the last turn for his fourth win of the year ,
both [25 Kennington 25] and [26 Fitzpatrick 26] had
problems . [27 Kennington 27] went on to win the next two
races at Riverside and [19 19 Barrie Speedway 19] to take a
lead of 40 points over [28 Fitzpatrick 28] , into the final
round of the season at [20 20 Kawartha Speedway 20] . At [29
Kawartha 29] , 9 Pete Shepherd III

1: 880534 CASCAR_Super_Series
2: 354723 The_Sports_Network
3: 354723 The_Sports_Network
4: 947540 Circuit_Gilles_Villeneuve
5: 64646 Toronto
6: 23240453 Dave_Whitlock
7: 19206314 Scott_Steckly
8: 177850 Canadian_Tire
9: 42969 Dodge
10: 2613675 Andrew_Ranger
11: 7674303 Delaware_Speedway
12: 12160023 D._J._Kennington
13: 12382592 J._R._Fitzpatrick
14: 23240270 Kerry_Micks
15: 19206314 Scott_Steckly
16: 12264205 Motoplex_Speedway
17: 19206314 Scott_Steckly
18: 10332537 Auto_Clearing_Motor_Speedway
19: 8516283 Barrie_Speedway
20: 12105396 Kawartha_Speedway

1: 1236852 Aaron's,_Inc.

[winner: 1006 Aaron, pem_pos=0, final_scores=-0.34927973,
final_local_scores=-0.20949747, pem=-0.21499276,
similarity=-0.9929454]

RECALL MISS: 1236852 Aaron's,_Inc.]

[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.4163632, final_local_scores=-2.527074,
pem=-3.4067376, similarity=-52.796852]

RECALL MISS: 1236852 Aaron's,_Inc.]

2: 733700 Tide_(brand)

[winner: 75128 Norfolk_Tides, pem_pos=8,
final_scores=-0.18382733, final_local_scores=-0.13352805,
pem=-2.0794415, similarity=11.413297]

gt: 733700 Tide_(brand), pem_pos=1,
final_scores=-0.31397808, final_local_scores=-0.30314556,
pem=-0.8898574, similarity=-0.11757229]

3: 33589 Walmart

[winner: 33589 Walmart, pem_pos=0,
final_scores=0.021657243, final_local_scores=0.08046147,
pem=0.0, similarity=6.490844]

gt: 33589 Walmart, pem_pos=0, final_scores=0.021657243,

final_local_scores=0.08046147, pem=0.0, similarity=6.490844]
4: 21114 NASCAR
[winner: 21114 NASCAR, pem_pos=0,
final_scores=-0.6842663, final_local_scores=-0.8052396,
pem=0.0, similarity=-19.940243]
gt: 21114 NASCAR, pem_pos=0, final_scores=-0.6842663,
final_local_scores=-0.8052396, pem=0.0,
similarity=-19.940243]
5: 4864615 K&N_Pro_Series_East
[winner: 27300604 Ferrari_Maranello_Series, pem_pos=9,
final_scores=-1.0428424, final_local_scores=-0.91310704,
pem=-3.295837, similarity=-5.23595]
RECALL MISS: 4864615 K&N_Pro_Series_East]
[winner: 52924 East_of_England, pem_pos=2,
final_scores=-1.067254, final_local_scores=-1.0550627,
pem=-2.5119424, similarity=-13.735128]
RECALL MISS: 4864615 K&N_Pro_Series_East]
[winner: 2998388 K&N_Engineering, pem_pos=0,
final_scores=-1.2565542, final_local_scores=-1.2775953,
pem=0.0, similarity=-34.036278]
RECALL MISS: 4864615 K&N_Pro_Series_East]
[winner: 17982774 Photomatix_Pro, pem_pos=14,
final_scores=-1.4295237, final_local_scores=-1.3649092,
pem=-3.0445223, similarity=-20.085312]
RECALL MISS: 4864615 K&N_Pro_Series_East]
6: 8516543 Autodrome_Saint-Eustache
[winner: 8516543 Autodrome_Saint-Eustache, pem_pos=0,
final_scores=-0.007350266, final_local_scores=0.005694747,
pem=0.0, similarity=4.2596555]
gt: 8516543 Autodrome_Saint-Eustache, pem_pos=0,
final_scores=-0.007350266, final_local_scores=0.005694747,
pem=0.0, similarity=4.2596555]
7: 11095234 Circuit_Trois-Rivières
[winner: 247035 Trois-Rivières, pem_pos=0,
final_scores=0.20011379, final_local_scores=0.2550859,
pem=-0.32429853, similarity=13.46557]
gt: 11095234 Circuit_Trois-Rivières, pem_pos=4,
final_scores=-0.50672334, final_local_scores=-0.54705095,
pem=-3.595043, similarity=7.315024]
8: 12065059 Don_Thomson,_Jr.
[winner: 12065059 Don_Thomson,_Jr., pem_pos=0,
final_scores=0.14877899, final_local_scores=0.16373086,
pem=0.0, similarity=8.975768]
gt: 12065059 Don_Thomson,_Jr., pem_pos=0,
final_scores=0.14877899, final_local_scores=0.16373086,
pem=0.0, similarity=8.975768]
[winner: 768194 Al_Unser,_Jr., pem_pos=1,
final_scores=-0.16059859, final_local_scores=-0.18649411,
pem=-1.068153, similarity=4.333139]
RECALL MISS: 12065059 Don_Thomson,_Jr.]
[winner: 330842 Achang_people, pem_pos=1,
final_scores=-1.6164988, final_local_scores=-1.5762892,

pem=-2.4264104, similarity=-29.754705|
RECALL MISS: 12065059 Don_Thomson,_Jr.]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-2.1576447, final_local_scores=-2.149284,
pem=-4.6728287, similarity=-34.637634|
RECALL MISS: 12065059 Don_Thomson,_Jr.]
9: 8356186 Peter_Shepherd_(racing_driver)
[winner: 8356186 Peter_Shepherd_(racing_driver),
pem_pos=0, final_scores=-0.05778359,
final_local_scores=-0.048685506, pem=0.0,
similarity=2.6368403|
gt: 8356186 Peter_Shepherd_(racing_driver), pem_pos=0,
final_scores=-0.05778359, final_local_scores=-0.048685506,
pem=0.0, similarity=2.6368403]
1: 880534 CASCAR_Super_Series
[pem_pos=0, final_scores=0.08147065,
final_local_scores=0.09786309, pem=0.0, similarity=7.010143]
2: 354723 The_Sports_Network
[pem_pos=0, final_scores=0.038768455,
final_local_scores=0.061729133, pem=-0.17330989,
similarity=6.8743196]
3: 354723 The_Sports_Network
[pem_pos=0, final_scores=0.099681094,
final_local_scores=0.12826693, pem=-0.17330989,
similarity=8.85994]
4: 947540 Circuit_Gilles_Villeneuve
[pem_pos=0, final_scores=0.15241067,
final_local_scores=0.12689799, pem=0.0, similarity=7.876602]
5: 64646 Toronto
[pem_pos=0, final_scores=0.49582785,
final_local_scores=0.54382235, pem=-0.12291032,
similarity=20.986862]
6: 23240453 Dave_Whitlock
[pem_pos=0, final_scores=0.102892235,
final_local_scores=0.13450453, pem=0.0, similarity=8.103597]
7: 19206314 Scott_Steckly
[pem_pos=0, final_scores=0.19014691,
final_local_scores=0.17683429, pem=0.0, similarity=9.3668]
8: 177850 Canadian_Tire
[pem_pos=0, final_scores=0.17612232,
final_local_scores=0.21845675, pem=0.0, similarity=10.608898]
9: 42969 Dodge
[pem_pos=0, final_scores=0.10005067,
final_local_scores=0.07415111, pem=-0.3009799,
similarity=7.9393053]
10: 2613675 Andrew_Ranger
[pem_pos=0, final_scores=0.3584826,
final_local_scores=0.34515604, pem=0.0, similarity=14.389857]
11: 7674303 Delaware_Speedway
[pem_pos=0, final_scores=0.16929297,
final_local_scores=0.17125988, pem=0.0, similarity=9.200449]
12: 12160023 D._J._Kennington

[pem_pos=0, final_scores=0.3172115,
final_local_scores=0.29295307, pem=0.0, similarity=12.832017]
13: 12382592 J._R._Fitzpatrick
[pem_pos=0, final_scores=0.16616784,
final_local_scores=0.13773024, pem=0.0, similarity=8.199858]
14: 23240270 Kerry_Micks
[pem_pos=0, final_scores=0.12306611,
final_local_scores=0.17050198, pem=0.0, similarity=9.177832]
15: 19206314 Scott_Steckly
[pem_pos=0, final_scores=0.2546569,
final_local_scores=0.24730167, pem=0.0, similarity=11.469688]
16: 12264205 Motoplex_Speedway
[pem_pos=0, final_scores=0.032082364,
final_local_scores=0.06917465, pem=0.0, similarity=6.154022]
17: 19206314 Scott_Steckly
[pem_pos=0, final_scores=0.19798689,
final_local_scores=0.18539828, pem=0.0, similarity=9.622368]
18: 10332537 Auto_Clearing_Motor_Speedway
[pem_pos=0, final_scores=0.07248519,
final_local_scores=0.10339531, pem=0.0, similarity=7.1752353]
19: 8516283 Barrie_Speedway
[pem_pos=0, final_scores=0.11359407,
final_local_scores=0.1251066, pem=0.0, similarity=7.8231435]
20: 12105396 Kawartha_Speedway
[pem_pos=0, final_scores=0.08444725,
final_local_scores=0.11599001, pem=0.0, similarity=7.5510864]
1: 6966833 NASCAR_Canadian_Tire_Series
[pem_pos=0, final_scores=0.17776431,
final_local_scores=0.13476253, pem=0.0, similarity=8.111296]
2: 7954681 Montreal
[pem_pos=0, final_scores=0.55112356,
final_local_scores=0.6062317, pem=-0.11023275,
similarity=22.780338]
3: 2975638 Honda_Indy_Toronto
[pem_pos=0, final_scores=0.07825662,
final_local_scores=0.04837595, pem=0.0, similarity=5.5333476]
4: 658603 CASCAR
[pem_pos=0, final_scores=0.17495738,
final_local_scores=0.19828516, pem=0.0, similarity=10.006938]
5: 23240344 Mark_Dilley
[pem_pos=0, final_scores=0.20858555,
final_local_scores=0.26880193, pem=0.0, similarity=12.111299]
6: 28218630 Pierre_Bourque_(journalist)
[pem_pos=0, final_scores=0.19983794,
final_local_scores=0.2809882, pem=0.0, similarity=12.474961]
7: 3434750 United_States
[pem_pos=0, final_scores=0.09425993,
final_local_scores=0.1552043, pem=-0.78483206,
similarity=12.989357]
8: 6966833 NASCAR_Canadian_Tire_Series
[pem_pos=0, final_scores=0.21047132,
final_local_scores=0.17048997, pem=0.0, similarity=9.177474]

9: 12160023 D._J._Kennington
[pem_pos=0, final_scores=0.41964978,
final_local_scores=0.40485132, pem=0.0, similarity=16.171284]
10: 2613675 Andrew_Ranger
[pem_pos=0, final_scores=0.2843384,
final_local_scores=0.26416475, pem=0.0, similarity=11.972916]
11: 64646 Toronto
[pem_pos=0, final_scores=0.39931983,
final_local_scores=0.438402, pem=-0.12291032,
similarity=17.84091]
12: 12382592 J._R._Fitzpatrick
[pem_pos=0, final_scores=0.10485999,
final_local_scores=0.070915565, pem=0.0, similarity=6.205975]
13: 95405 Edmonton
[pem_pos=0, final_scores=0.33091408,
final_local_scores=0.35300508, pem=-0.27044967,
similarity=16.094835]
14: 12160023 D._J._Kennington
[pem_pos=0, final_scores=0.44024277,
final_local_scores=0.42734605, pem=0.0, similarity=16.842571]
15: 12160023 D._J._Kennington
[pem_pos=0, final_scores=0.42458814,
final_local_scores=0.41024572, pem=0.0, similarity=16.332264]
16: 12160023 D._J._Kennington
[pem_pos=0, final_scores=0.4305144,
final_local_scores=0.41671932, pem=0.0, similarity=16.525448]
17: 2613675 Andrew_Ranger
[pem_pos=0, final_scores=0.35215896,
final_local_scores=0.33824846, pem=0.0, similarity=14.183722]
18: 247035 Trois-Rivières
[pem_pos=0, final_scores=0.20011379,
final_local_scores=0.2550859, pem=-0.32429853,
similarity=13.46557]
19: 20004657 John_Fitzpatrick_(racing_driver)
[pem_pos=1, final_scores=0.08523743,
final_local_scores=0.116894126, pem=-1.6531818,
similarity=16.568327]
20: 12160023 D._J._Kennington
[pem_pos=0, final_scores=0.31967026,
final_local_scores=0.29563895, pem=0.0, similarity=12.9121685]
21: 12065059 Don_Thomson,_Jr.
[pem_pos=0, final_scores=0.14877899,
final_local_scores=0.16373086, pem=0.0, similarity=8.975768]
22: 7954681 Montreal
[pem_pos=0, final_scores=0.5908753,
final_local_scores=0.6496545, pem=-0.11023275,
similarity=24.076162]
23: 2613675 Andrew_Ranger
[pem_pos=0, final_scores=0.32026058,
final_local_scores=0.30340433, pem=0.0, similarity=13.143904]
24: 31936040 Jason_Bowles
[pem_pos=0, final_scores=0.21428756,

final_local_scores=0.24002206, pem=0.0, similarity=11.25245]
 25: 12160023 D._J._Kennington
 [pem_pos=0, final_scores=0.39783078,
 final_local_scores=0.38101745, pem=0.0, similarity=15.460034]
 26: 1444074 Sean_Fitzpatrick
 [pem_pos=11, final_scores=0.1252319,
 final_local_scores=0.22080642, pem=-3.317121,
 similarity=28.718037]
 27: 12160023 D._J._Kennington
 [pem_pos=0, final_scores=0.4395582,
 final_local_scores=0.4265983, pem=0.0, similarity=16.820257]
 28: 12382592 J._R._Fitzpatrick
 [pem_pos=0, final_scores=0.07852785,
 final_local_scores=0.04215172, pem=0.0, similarity=5.3476043]
 29: 386793 Kawartha_Lakes
 [pem_pos=0, final_scores=0.061009303,
 final_local_scores=0.14811537, pem=0.0, similarity=8.509771]

Sample: b'Ruby_B'

DeMesme [1 Ruby Butler DeMesme 1] was [1 1 United States 1]
 [1 Assistant Secretary of the Air Force for Manpower , Reser
 ve Affairs , Installations and Environment from 1998 to
 2001 . [2 Ruby B. DeMesme 2] was educated at 2 St. Augustine '
 s College in [3 3 Raleigh 3] , [4 North Carolina 4] ,
 receiving a in 1969 . She later earned an from the [2 2 Unive
 rsity of North Carolina at Chapel Hill 2] . From 1969 to
 1980 , DeMesme worked as an executive administrator in the [5
 4 Cumberland County 5] , North Carolina Department of 5 Social
 Services . In 1980 , she joined the [3 3 United States Depart
 ment of the Army 3] and spent the next decade working in a
 number of positions dealing with . From January 1989 to April
 1990 , she was a 6 United States Office of Personnel Manageme
 nt Congressional Fellow in the Office of 7 Sen. [4 4 John Gle
 nn 4] -LRB- 8 D -- -RRB- . She then returned to the [6
 Department of the Army 6] . In 1991 , DeMesme moved to the [7
 9 United States Department of the Air Force 7] , becoming
 Assistant Deputy Assistant Secretary of the Air Force for
 Health Affairs , and Morale , Welfare , Recreation and
 Services Program . In February 1993 , she became Acting
 Deputy Assistant Secretary of the Air Force for Force
 Management and Personnel , being promoted to full Deputy
 Assistant Secretary of the [8 Air Force 8] for Force
 Management and Personnel in May 1994 . On July 6 , 1998 , 10 P
 resident of the United States [5 5 Bill Clinton 5] announced
 his intent to nominate DeMesme to be 11 Assistant Secretary of
 the [9 Air Force 9] for Manpower , Reserve Affairs , Installa
 tions and Environment . She was the first [10 African
 American 10] woman to be appointed to this position . DeMesme
 retired from government service in 2001 , joining [6 6 IBM 6]
 as Senior Defense Fellow and , later becoming Director of
 Human Capital Solutions & Defense Practices . In 2005 , she
 joined [11 BearingPoint 11] 's Public Services Business

Transformation Team . In May 2009 , she joined 12 Deloitte

1: 3434750 United_States

2: 89510 University_of_North_Carolina_at_Chapel_Hill

3: 798348 United_States_Department_of_the_Army

4: 58702 John_Glenn

5: 3356 Bill_Clinton

6: 40379651 IBM

1: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs)

[winner: 32090 United_States_Air_Force, pem_pos=0, final_scores=-0.07411061, final_local_scores=-0.13855018, pem=-0.8379956, similarity=4.5122495|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs)]

[winner: 25658639

Assistant_Secretary_of_the_Navy_(Installations_and_Environment), pem_pos=0, final_scores=-0.3199292, final_local_scores=-0.2875175, pem=0.0, similarity=-4.4903827|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs)]

[winner: 935342 Manpower_Inc., pem_pos=0, final_scores=-0.4203066, final_local_scores=-0.4110117, pem=-0.7000777, similarity=-4.368565|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs)]

[winner: 1253416 Air_Force_Reserve_Command, pem_pos=0, final_scores=-0.64905065, final_local_scores=-0.73334366, pem=-1.2089005, similarity=-11.220536|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs)]

[winner: 214196 Secretary, pem_pos=0, final_scores=-1.0127794, final_local_scores=-1.1090748, pem=-1.1345513, similarity=-22.837423|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs)]

[winner: 169228 Affair, pem_pos=0, final_scores=-1.1909881, final_local_scores=-1.2128899, pem=-0.07647034, similarity=-31.68948|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs)]

[winner: 417934 Office_Assistant, pem_pos=0, final_scores=-1.4327686, final_local_scores=-1.4546685, pem=-0.7444405, similarity=-35.272114|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs))]

[winner: 10447732

Results_by_riding_of_the_Canadian_federal_election,_2008,
pem_pos=12, final_scores=-1.443124,
final_local_scores=-1.342524, pem=-3.6198866,
similarity=-16.288374|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs))]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8189741, final_local_scores=-1.8088939,
pem=-4.6728287, similarity=-24.47972|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs))]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.9290615, final_local_scores=-1.9291477,
pem=-4.6728287, similarity=-28.068333|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affairs))]

2: 3698454 St._Augustine's_University

[winner: 5689 College, pem_pos=0,
final_scores=-0.069328174, final_local_scores=-0.07958618,
pem=-0.85246265, similarity=6.3505273|

RECALL MISS: 3698454 St._Augustine's_University]

[winner: 109790 St._Augustine,_Florida, pem_pos=1,
final_scores=-0.3447727, final_local_scores=-0.37550324,
pem=-0.85844827, similarity=-2.4476779|

RECALL MISS: 3698454 St._Augustine's_University]

[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.45786, final_local_scores=-2.5641432,
pem=-3.4067376, similarity=-53.903065|

RECALL MISS: 3698454 St._Augustine's_University]

3: 57845 Raleigh,_North_Carolina

[winner: 21650 North_Carolina, pem_pos=0,
final_scores=0.5020995, final_local_scores=0.5005045,
pem=-0.21680297, similarity=20.204773|

RECALL MISS: 57845 Raleigh,_North_Carolina]

[winner: 57845 Raleigh,_North_Carolina, pem_pos=0,
final_scores=0.42631125, final_local_scores=0.40958875,
pem=-0.35659322, similarity=18.25187|

gt: 57845 Raleigh,_North_Carolina, pem_pos=0,
final_scores=0.42631125, final_local_scores=0.40958875,
pem=-0.35659322, similarity=18.25187]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7026236, final_local_scores=-1.6817987,
pem=-4.6728287, similarity=-20.686941|

RECALL MISS: 57845 Raleigh,_North_Carolina]

4: 93320 Cumberland_County,_North_Carolina

[winner: 3456672 Cumberland_County,_Pennsylvania,

pem_pos=0, final_scores=0.046067253,
 final_local_scores=0.067389846, pem=-1.3470802,
 similarity=13.426392|
 gt: 93320 Cumberland_County,_North_Carolina, pem_pos=4,
 final_scores=-0.06142676, final_local_scores=-0.06698236,
 pem=-1.9901336, similarity=12.913482]
 [winner: 21650 North_Carolina, pem_pos=0,
 final_scores=-0.061431184, final_local_scores=-0.12329047,
 pem=-0.21680297, similarity=1.5894877|
 RECALL MISS: 93320 Cumberland_County,_North_Carolina]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.8425689, final_local_scores=-1.8346677,
 pem=-4.6728287, similarity=-25.248857|
 RECALL MISS: 93320 Cumberland_County,_North_Carolina]
 5: 146717 Social_work
 [winner: 146717 Social_work, pem_pos=0,
 final_scores=-0.29020494, final_local_scores=-0.31861657,
 pem=0.0, similarity=-5.4184413|
 gt: 146717 Social_work, pem_pos=0,
 final_scores=-0.29020494, final_local_scores=-0.31861657,
 pem=0.0, similarity=-5.4184413]
 6: 1086289 United_States_Office_of_Personnel_Management
 [winner: 1086289
 United_States_Office_of_Personnel_Management, pem_pos=0,
 final_scores=-0.796673, final_local_scores=-0.87333757,
 pem=0.0, similarity=-21.972422|
 gt: 1086289
 United_States_Office_of_Personnel_Management, pem_pos=0,
 final_scores=-0.796673, final_local_scores=-0.87333757,
 pem=0.0, similarity=-21.972422]
 7: 24909346 United_States_Senate
 [winner: 24909346 United_States_Senate, pem_pos=0,
 final_scores=-0.40095812, final_local_scores=-0.4703168,
 pem=-0.020029303, similarity=-9.8365555|
 gt: 24909346 United_States_Senate, pem_pos=0,
 final_scores=-0.40095812, final_local_scores=-0.4703168,
 pem=-0.020029303, similarity=-9.8365555]
 8: 5043544 Democratic_Party_(United_States) no overlap with
 any filtered span
 9: 798400 United_States_Department_of_the_Air_Force
 [winner: 32090 United_States_Air_Force, pem_pos=0,
 final_scores=0.047937438, final_local_scores=-0.0052312165,
 pem=0.0, similarity=3.9336028|
 gt: 798400 United_States_Department_of_the_Air_Force,
 pem_pos=1, final_scores=-0.58455265,
 final_local_scores=-0.6345099, pem=-3.1354942,
 similarity=2.2059765]
 10: 24113 President_of_the_United_States
 [winner: 3434750 United_States, pem_pos=0,
 final_scores=-0.45055443, final_local_scores=-0.5414759,
 pem=-0.073457584, similarity=-11.669533|
 RECALL MISS: 24113 President_of_the_United_States]

[winner: 1961717 Supreme_Court_of_South_Australia,
pem_pos=0, final_scores=-1.0347346,
final_local_scores=-1.0935847, pem=0.0,
similarity=-28.545034|

RECALL MISS: 24113 President_of_the_United_States]

[winner: 24113 President_of_the_United_States, pem_pos=0,
final_scores=-1.1371475, final_local_scores=-1.2864885,
pem=-0.99314094, similarity=-28.900812|

gt: 24113 President_of_the_United_States, pem_pos=0,
final_scores=-1.1371475, final_local_scores=-1.2864885,
pem=-0.99314094, similarity=-28.900812]

11: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affa
irs)

[winner: 32090 United_States_Air_Force, pem_pos=0,
final_scores=0.14382054, final_local_scores=0.10758692,
pem=-0.8379956, similarity=11.857471|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affa
irs))]

[winner: 25658639

Assistant_Secretary_of_the_Navy_(Installations_and_Environmen
t), pem_pos=0, final_scores=-0.30153406,
final_local_scores=-0.2674236, pem=0.0,
similarity=-3.8907409|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affa
irs))]

[winner: 935342 Manpower_Inc., pem_pos=0,
final_scores=-0.40127105, final_local_scores=-0.39021826,
pem=-0.7000777, similarity=-3.7480469|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affa
irs))]

[winner: 3672039 Assistant_Secretary, pem_pos=0,
final_scores=-0.51752055, final_local_scores=-0.5165579,
pem=0.0, similarity=-11.325404|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affa
irs))]

[winner: 1253416 Air_Force_Reserve_Command, pem_pos=0,
final_scores=-0.898415, final_local_scores=-1.0057364,
pem=-1.2089005, similarity=-19.34928|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affa
irs))]

[winner: 169228 Affair, pem_pos=0,
final_scores=-1.3505433, final_local_scores=-1.3871796,
pem=-0.07647034, similarity=-36.890633|

RECALL MISS: 31053544

Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affa
irs))]

[winner: 10447732
 Results_by_riding_of_the_Canadian_federal_election,_2008,
 pem_pos=12, final_scores=-1.3657029,
 final_local_scores=-1.2579532, pem=-3.6198866,
 similarity=-13.764614|
 RECALL MISS: 31053544
 Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affa
 irs)]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.8679212, final_local_scores=-1.8623612,
 pem=-4.6728287, similarity=-26.075287|
 RECALL MISS: 31053544
 Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affa
 irs)]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-2.0129654, final_local_scores=-2.0208,
 pem=-4.6728287, similarity=-30.803425|
 RECALL MISS: 31053544
 Assistant_Secretary_of_the_Air_Force_(Manpower_&_Reserve_Affa
 irs)]
 12: 586015 Deloitte
 [winner: 586015 Deloitte, pem_pos=0,
 final_scores=-0.39761937, final_local_scores=-0.41134045,
 pem=0.0, similarity=-8.185507|
 gt: 586015 Deloitte, pem_pos=0,
 final_scores=-0.39761937, final_local_scores=-0.41134045,
 pem=0.0, similarity=-8.185507]
 1: 3434750 United_States
 [pem_pos=0, final_scores=0.43731833,
 final_local_scores=0.43729997, pem=-0.073457584,
 similarity=17.539091]
 2: 89510 University_of_North_Carolina_at_Chapel_Hill
 [pem_pos=0, final_scores=0.43018097,
 final_local_scores=0.45251083, pem=0.0, similarity=17.593538]
 3: 798348 United_States_Department_of_the_Army
 [pem_pos=0, final_scores=0.13369994,
 final_local_scores=0.12414861, pem=-0.0037105891,
 similarity=7.8147335]
 4: 58702 John_Glenn
 [pem_pos=0, final_scores=0.30671364,
 final_local_scores=0.32619548, pem=-0.040958993,
 similarity=14.046778]
 5: 3356 Bill_Clinton
 [pem_pos=0, final_scores=0.2965203,
 final_local_scores=0.30396748, pem=0.0, similarity=13.160708]
 6: 40379651 IBM
 [pem_pos=0, final_scores=0.25681216,
 final_local_scores=0.30358705, pem=0.0, similarity=13.149356]
 1: 31062949 Ruby_B._DeMesme
 [pem_pos=0, final_scores=0.2191975,
 final_local_scores=0.23991624, pem=0.0, similarity=11.249292]
 2: 31062949 Ruby_B._DeMesme

[pem_pos=0, final_scores=0.11551817,
 final_local_scores=0.12666234, pem=0.0, similarity=7.86957]
 3: 57845 Raleigh, North Carolina
 [pem_pos=0, final_scores=0.42631125,
 final_local_scores=0.40958875, pem=-0.35659322,
 similarity=18.25187]
 4: 21650 North Carolina
 [pem_pos=0, final_scores=0.5020995,
 final_local_scores=0.5005045, pem=-0.21680297,
 similarity=20.204773]
 5: 3456672 Cumberland County, Pennsylvania
 [pem_pos=0, final_scores=0.046067253,
 final_local_scores=0.067389846, pem=-1.3470802,
 similarity=13.426392]
 6: 798348 United States Department of the Army
 [pem_pos=1, final_scores=0.11775334,
 final_local_scores=0.10996431, pem=-0.008387416,
 similarity=7.4168797]
 7: 32090 United States Air Force
 [pem_pos=0, final_scores=0.047937438,
 final_local_scores=-0.0052312165, pem=0.0,
 similarity=3.9336028]
 8: 32090 United States Air Force
 [pem_pos=0, final_scores=0.08598678,
 final_local_scores=0.044412315, pem=-0.8379956,
 similarity=9.972216]
 9: 32090 United States Air Force
 [pem_pos=0, final_scores=0.14382054,
 final_local_scores=0.10758692, pem=-0.8379956,
 similarity=11.857471]
 10: 2154 African American
 [pem_pos=1, final_scores=0.20194446,
 final_local_scores=0.18969125, pem=-0.7252736,
 similarity=13.69463]
 11: 201806 BearingPoint
 [pem_pos=0, final_scores=0.13892908,
 final_local_scores=0.20894334, pem=0.0, similarity=10.324999]

Sample: b'Colombophis'

`Colombophis` -LRB- '' **1 snake** of [1 Colombia 1] '' in [2 2
 Greek 2] -RRB- is an extinct genus of snakes of the **3 clade 4**
Alethinophidia, a group of `` primitive '' snakes . The **5 ge**
nus was first recognized in the formation Villavieja in the
 town of Los Mangos , part of the known fossil fauna of [3 6 La
 Venta 3] in the -LRB- [1 1 Colombia 1] -RRB- , in the middle
 Miocene . With the remains of a **7 fossil** snake was erected
 the species *Colombophis portai* in 1977 , based on forty
 fragmentary . These vertebrae are characterized by a low
 neural spine , and subdivided paradiapophysis and thin
 zygosphen . The vertebrae are medium to large , so the snake
 would measure about 1.77 meters long , similar in size to the
 current **8 Boa constrictor** . Remains found in the [4 Brazilian

4] in the [2 2 Purus River 2] area -LRB- 9 Late Miocene - RRB- , with vertebrae shorter than wide , thick zygosphen and high neural thorn , were classified as a new species , Colombophis spinosus ; fossil remains found in [3 3 Venezuela 3] -LRB- 10 Urumaco area -RRB- and [5 Colombia 5] -LRB- La Venta -RRB- also seem to belong to this new species rather than C. portai . Traditionally , Colombophis was taken as a representative of the family primitive 11 Aniliidae , 12 fossil and primitive snakes that still retained some traces of lizard as , the false coral . However , retention of primitive traits that identify Colombophis and other members of the Anilioidea -LRB- including features already found in the 13 Cretaceous snake 14 Dinilysia patagonica -RRB- indicate that this group is and that for now , Colombophis remains as an Alethinophidia of uncertain classification . Moreover , their remains indicate that the ecosystems of [4 4 South America 4] from the mid-Miocene were humid tropical environments as reinforces the finding of other reptile faunas of [6 Urumaco 6] , [7 La Venta 7] and Solimões as and

1: 5222 Colombia

2: 23895 Purus_River

3: 32374 Venezuela

4: 26769 South_America

1: 29370 Snake

[winner: 18766454 Snake_Island_(Victoria), pem_pos=11, final_scores=-1.434327, final_local_scores=-1.4396343, pem=-3.8359885, similarity=-18.011143]

RECALL MISS: 29370 Snake]

2: 11887 Greek_language

[winner: 12108 Greece, pem_pos=1, final_scores=0.36452967, final_local_scores=0.4107908, pem=-2.3247876, similarity=28.991087]

gt: 11887 Greek_language, pem_pos=0, final_scores=-0.4550013, final_local_scores=-0.4373997, pem=-0.6261292, similarity=-5.558178]

3: 6682 Clade

[winner: 6682 Clade, pem_pos=0, final_scores=-1.5086042, final_local_scores=-1.6058404, pem=-0.038815293, similarity=-43.62068]

gt: 6682 Clade, pem_pos=0, final_scores=-1.5086042, final_local_scores=-1.6058404, pem=-0.038815293, similarity=-43.62068]

4: 12764603 Alethinophidia

[winner: 12764603 Alethinophidia, pem_pos=0, final_scores=-0.37313348, final_local_scores=-0.20770723, pem=0.0, similarity=-2.1086853]

gt: 12764603 Alethinophidia, pem_pos=0, final_scores=-0.37313348, final_local_scores=-0.20770723, pem=0.0, similarity=-2.1086853]

5: 38493 Genus

[winner: 7721104 The_genus_Banksia_L.f._(Proteaceae), pem_pos=0, final_scores=-0.97707677,

```

final_local_scores=-0.9078916, pem=0.0,
similarity=-23.003582|
    RECALL MISS: 38493 Genus]
6: 10040828 La_Venta_(Colombia)
    [winner: 336434 La_Venta, pem_pos=0,
final_scores=0.09989767, final_local_scores=0.10248344,
pem=0.0, similarity=7.148023|
    gt: 10040828 La_Venta_(Colombia), pem_pos=2,
final_scores=-0.11959496, final_local_scores=-0.0931246,
pem=-1.0181272, similarity=6.847423 ]
7: 10958 Fossil
    [winner: 367492 Trace_fossil, pem_pos=0,
final_scores=-1.2312133, final_local_scores=-1.3045919,
pem=0.0, similarity=-34.84191|
    gt: 10958 Fossil, pem_pos=1, final_scores=-1.3947134,
final_local_scores=-1.5306842, pem=-0.021992672,
similarity=-41.469357 ]
8: 30863385 Boa_constrictor
    [winner: 20705199 Boliviana_de_Aviación, pem_pos=15,
final_scores=-0.4009226, final_local_scores=-0.34543243,
pem=-3.0445223, similarity=10.337905|
    RECALL MISS: 30863385 Boa_constrictor]
    [winner: 2984918 Constrictor_(comics), pem_pos=0,
final_scores=-1.4393059, final_local_scores=-1.3920038,
pem=0.0, similarity=-37.450455|
    RECALL MISS: 30863385 Boa_constrictor]
9: 9869128 Late_Miocene
    [winner: 9869128 Late_Miocene, pem_pos=0,
final_scores=-0.20321971, final_local_scores=-0.18481387,
pem=-0.04367506, similarity=-1.1879898|
    gt: 9869128 Late_Miocene, pem_pos=0,
final_scores=-0.20321971, final_local_scores=-0.18481387,
pem=-0.04367506, similarity=-1.1879898 ]
10: 4613845 Urumaco
    [winner: 4613845 Urumaco, pem_pos=0,
final_scores=-0.017717123, final_local_scores=0.028595895,
pem=0.0, similarity=4.9430714|
    gt: 4613845 Urumaco, pem_pos=0,
final_scores=-0.017717123, final_local_scores=0.028595895,
pem=0.0, similarity=4.9430714 ]
11: 1076177 Aniliidae
    [winner: 1076177 Aniliidae, pem_pos=0,
final_scores=-0.33680803, final_local_scores=-0.29360723,
pem=0.0, similarity=-4.672112|
    gt: 1076177 Aniliidae, pem_pos=0,
final_scores=-0.33680803, final_local_scores=-0.29360723,
pem=0.0, similarity=-4.672112 ]
12: 3072173 Fossorial
    [winner: 3072173 Fossorial, pem_pos=0,
final_scores=-0.9654857, final_local_scores=-0.96639633,
pem=0.0, similarity=-24.749481|
    gt: 3072173 Fossorial, pem_pos=0,

```

```
final_scores=-0.9654857, final_local_scores=-0.96639633,  
pem=0.0, similarity=-24.749481 ]  
13: 5615 Cretaceous  
    [winner: 5615 Cretaceous, pem_pos=0,  
final_scores=-0.055333897, final_local_scores=-0.07229768,  
pem=-0.0036390396, similarity=1.9519956]  
    gt: 5615 Cretaceous, pem_pos=0,  
final_scores=-0.055333897, final_local_scores=-0.07229768,  
pem=-0.0036390396, similarity=1.9519956 ]  
14: 6004630 Dinilysia  
    [winner: 6004630 Dinilysia, pem_pos=0,  
final_scores=-0.2313255, final_local_scores=-0.12847869,  
pem=0.0, similarity=0.2556526]  
    gt: 6004630 Dinilysia, pem_pos=0,  
final_scores=-0.2313255, final_local_scores=-0.12847869,  
pem=0.0, similarity=0.2556526 ]  
1: 5222 Colombia  
    [pem_pos=0, final_scores=0.27120703,  
final_local_scores=0.2229163, pem=-0.22328302,  
similarity=11.956228]  
2: 23895 Purus_River  
    [pem_pos=0, final_scores=0.21094106,  
final_local_scores=0.2633266, pem=0.0, similarity=11.947904]  
3: 32374 Venezuela  
    [pem_pos=0, final_scores=0.44724017,  
final_local_scores=0.42023444, pem=-0.21226841,  
similarity=17.784697]  
4: 26769 South_America  
    [pem_pos=0, final_scores=0.30512303,  
final_local_scores=0.2627412, pem=0.0, similarity=11.930434]  
1: 5222 Colombia  
    [pem_pos=0, final_scores=0.37894136,  
final_local_scores=0.34059972, pem=-0.22328302,  
similarity=15.468135]  
2: 12108 Greece  
    [pem_pos=1, final_scores=0.36452967,  
final_local_scores=0.4107908, pem=-2.3247876,  
similarity=28.991087]  
3: 336434 La_Venta  
    [pem_pos=0, final_scores=0.09989767,  
final_local_scores=0.10248344, pem=0.0, similarity=7.148023]  
4: 3383 Brazil  
    [pem_pos=0, final_scores=0.5658133,  
final_local_scores=0.5583772, pem=-0.2733083,  
similarity=22.239096]  
5: 5222 Colombia  
    [pem_pos=0, final_scores=0.64269406,  
final_local_scores=0.6287095, pem=-0.22328302,  
similarity=24.065907]  
6: 4613845 Urumaco  
    [pem_pos=0, final_scores=0.052011266,  
final_local_scores=0.10476358, pem=0.0, similarity=7.2160673]
```


7: 336434 La_Venta

[pem_pos=0, final_scores=0.06712599,
final_local_scores=0.06668538, pem=0.0, similarity=6.0797377]

Sample: b'Wijnand_van_der_Sanden'

[1 Wijnand Antonius Bernardus van der Sanden 1] -LRB- born 1953 in [1 1 Geldrop 1] -RRB- is a [2 2 Dutch 2] 1 archaeologist and 2 prehistorian . After graduating he studied 3 history of art s and 4 classical archaeology at the [3 3 Radboud University Nijmegen 3] , later he studied 5 prehistory at the [4 4 University of Groningen 4] where he gained his doctorate in 1990 with a thesis on [2 Dutch 2] 6 bog body finds from the to the Mens en moeras : veenlijken in Nederland van de bronstijd tot en met de Romeinse tijd . From 1987 to 1997 van der Sanden was 7 curator of the archaeological department of the [5 5 Drents Museum 5] in [6 6 Assen 6] . Since 1997 he has been the provincial archaeologist of the 8 province [7 7 Drenthe 7] . Van der Sanden was a member of the Werkgroep Hunebedden -LRB- Woprking group Megalithics -RRB- from 1988 to 1999 and since 1988 has been 9 editor-in-chief of the archaeological section of the . Since 1987 van der Sanden has been working on scientific studies relating to archaeological bog finds in the province of [3 Drenthe 3] and continued his studies in this area , which was also the subject of his doctorate 10 thesis . In 1995 , as a guest curator , he organised an 11 special exhibition at in [8 8 Denmark 8] on North-West European bog body finds and he edited the catalogue Vereeuwigd in het veen . -LRB- 12 English title Through nature to eternity . -RRB- Since 1990 van der Sanden has been working on the scientific verification of find records and archaeological finds of bog bodies in the [4 Netherlands 4] and [5 Germany 5] , especially from the 13 natchlass of the [6 German 6] scientist [9 9 Alfred Dieck 9] . [7 Wijnand van der Sanden 7] has also published several works on 14 Neolithic

1: 2132535 Geldrop

2: 21148 Netherlands

3: 956504 Radboud_University_Nijmegen

4: 255460 University_of_Groningen

5: 26137690 Drents_Museum

6: 73526 Assen

7: 57158 Drenthe

8: 76972 Denmark

9: 31496407 Alfred_Dieck

1: 18951655 Archaeology

[winner: 1480596 Biblical_archaeology, pem_pos=0,
final_scores=-1.2691653, final_local_scores=-1.3143122,
pem=0.0, similarity=-35.131985]

gt: 18951655 Archaeology, pem_pos=1,
final_scores=-1.5653721, final_local_scores=-1.7166157,
pem=-0.002445168, similarity=-47.12422]

2: 18994022 Prehistory

[winner: 18994022 Prehistory, pem_pos=0,
 final_scores=-1.3143852, final_local_scores=-1.4253137,
 pem=0.0, similarity=-38.44449|
 gt: 18994022 Prehistory, pem_pos=0,
 final_scores=-1.3143852, final_local_scores=-1.4253137,
 pem=0.0, similarity=-38.44449]
 3: 18985263 History_of_art
 [winner: 21207536 Art_history, pem_pos=0,
 final_scores=-0.99588317, final_local_scores=-1.0875964,
 pem=0.0, similarity=-28.366331|
 gt: 18985263 History_of_art, pem_pos=1,
 final_scores=-1.0342616, final_local_scores=-1.1127899,
 pem=-0.027446007, similarity=-28.968903]
 4: 606810 Classical_archaeology
 [winner: 606810 Classical_archaeology, pem_pos=0,
 final_scores=-0.7926056, final_local_scores=-0.8478563,
 pem=0.0, similarity=-21.21201|
 gt: 606810 Classical_archaeology, pem_pos=0,
 final_scores=-0.7926056, final_local_scores=-0.8478563,
 pem=0.0, similarity=-21.21201]
 5: 18994022 Prehistory
 [winner: 18994022 Prehistory, pem_pos=0,
 final_scores=-1.4458953, final_local_scores=-1.5689685,
 pem=-0.015107148, similarity=-42.64928|
 gt: 18994022 Prehistory, pem_pos=0,
 final_scores=-1.4458953, final_local_scores=-1.5689685,
 pem=-0.015107148, similarity=-42.64928]
 6: 102411 Bog_body
 [winner: 6532966 Bog-wood, pem_pos=0,
 final_scores=-1.2618006, final_local_scores=-1.2473278,
 pem=0.0, similarity=-33.133038|
 gt: 102411 Bog_body, pem_pos=9, final_scores=-3.0012271,
 final_local_scores=-3.2353206, pem=-7.7186856,
 similarity=-50.48329]
 [winner: 4903251 Body_of_humerus, pem_pos=14,
 final_scores=-1.4073551, final_local_scores=-1.3259468,
 pem=-3.3404412, similarity=-17.313345|
 RECALL MISS: 102411 Bog_body]
 7: 391167 Curator
 [winner: 391167 Curator, pem_pos=0,
 final_scores=-1.5477828, final_local_scores=-1.6785073,
 pem=-0.020542735, similarity=-45.888577|
 gt: 391167 Curator, pem_pos=0, final_scores=-1.5477828,
 final_local_scores=-1.6785073, pem=-0.020542735,
 similarity=-45.888577]
 8: 57951 Provinces_of_the_Netherlands
 [winner: 13776679 Flag_of_Flevoland, pem_pos=0,
 final_scores=-0.8134219, final_local_scores=-0.71037084,
 pem=0.0, similarity=-17.109169|
 RECALL MISS: 57951 Provinces_of_the_Netherlands]
 9: 577429 Editor-in-chief
 [winner: 577429 Editor-in-chief, pem_pos=0,

```

final_scores=-1.232043, final_local_scores=-1.2647169,
pem=-0.21463469, similarity=-32.484745|
    gt: 577429 Editor-in-chief, pem_pos=0,
final_scores=-1.232043, final_local_scores=-1.2647169,
pem=-0.21463469, similarity=-32.484745 ]
10: 300215 Thesis
    [winner: 3012337 Thesis_statement, pem_pos=1,
final_scores=-1.4061487, final_local_scores=-1.3903158,
pem=-1.035337, similarity=-31.769764|
    gt: 300215 Thesis, pem_pos=4, final_scores=-1.8432552,
final_local_scores=-1.9883192, pem=-1.7726059,
similarity=-45.605988 ]
11: 29815796 Exhibition
    [winner: 106203 Special_pleading, pem_pos=8,
final_scores=-1.35564, final_local_scores=-1.2564113,
pem=-2.3526614, similarity=-20.609966|
    RECALL MISS: 29815796 Exhibition]
    [winner: 6439791
Vincent_van_Gogh's_display_at_Les_XX,_1890, pem_pos=11,
final_scores=-1.3937045, final_local_scores=-1.4064837,
pem=-2.2988, similarity=-25.381332|
    gt: 29815796 Exhibition, pem_pos=0,
final_scores=-1.5867164, final_local_scores=-1.6885362,
pem=-1.1008902, similarity=-40.312763 ]
12: 8569916 English_language
    [winner: 8569916 English_language, pem_pos=0,
final_scores=-0.15140823, final_local_scores=-0.13773426,
pem=-0.640266, similarity=3.4613135|
    gt: 8569916 English_language, pem_pos=0,
final_scores=-0.15140823, final_local_scores=-0.13773426,
pem=-0.640266, similarity=3.4613135 ]
13: 3124337 Nachlass
    [winner: 3124327 Literary_executor, pem_pos=0,
final_scores=-0.780552, final_local_scores=-0.7521215,
pem=-0.0045249164, similarity=-18.330486|
    gt: 3124337 Nachlass, pem_pos=1,
final_scores=-1.5750356, final_local_scores=-1.6031885,
pem=-5.4004226, similarity=-14.384295 ]
14: 21189 Neolithic
    [winner: 21189 Neolithic, pem_pos=0,
final_scores=-0.07542537, final_local_scores=-0.030205593,
pem=-0.022485228, similarity=3.310596|
    gt: 21189 Neolithic, pem_pos=0,
final_scores=-0.07542537, final_local_scores=-0.030205593,
pem=-0.022485228, similarity=3.310596 ]
1: 2132535 Geldrop
    [pem_pos=0, final_scores=0.041604415,
final_local_scores=0.08292638, pem=0.0, similarity=6.5644016]
2: 21148 Netherlands
    [pem_pos=0, final_scores=0.49802607,
final_local_scores=0.47038162, pem=-0.7943452,
similarity=22.446611]

```

3: 956504 Radboud_University_Nijmegen
 [pem_pos=0, final_scores=0.32225186,
 final_local_scores=0.34490395, pem=0.0, similarity=14.382335]
 4: 255460 University_of_Groningen
 [pem_pos=0, final_scores=0.43001968,
 final_local_scores=0.44950587, pem=0.0, similarity=17.503864]
 5: 26137690 Drents_Museum
 [pem_pos=0, final_scores=0.118386135,
 final_local_scores=0.18266293, pem=0.0, similarity=9.540739]
 6: 73526 Assen
 [pem_pos=0, final_scores=0.29006702,
 final_local_scores=0.3109549, pem=-0.20065941,
 similarity=14.460444]
 7: 57158 Drenthe
 [pem_pos=0, final_scores=0.115296796,
 final_local_scores=0.09100349, pem=0.0, similarity=6.8054385]
 8: 76972 Denmark
 [pem_pos=0, final_scores=0.40255195,
 final_local_scores=0.39979196, pem=-0.26762357,
 similarity=17.475681]
 9: 31496407 Alfred_Dieck
 [pem_pos=0, final_scores=0.23459204,
 final_local_scores=0.28609973, pem=0.0, similarity=12.6275]
 1: 37414165 Wijnand_van_der_Sanden
 [pem_pos=0, final_scores=0.116751954,
 final_local_scores=0.12729022, pem=0.0, similarity=7.888307]
 2: 21148 Netherlands
 [pem_pos=0, final_scores=0.46197158,
 final_local_scores=0.4309975, pem=-0.7943452,
 similarity=21.27131]
 3: 57158 Drenthe
 [pem_pos=0, final_scores=0.18438096,
 final_local_scores=0.16646746, pem=0.0, similarity=9.057434]
 4: 21148 Netherlands
 [pem_pos=0, final_scores=0.451914,
 final_local_scores=0.42001116, pem=-0.24940795,
 similarity=17.980003]
 5: 11867 Germany
 [pem_pos=0, final_scores=0.8022571,
 final_local_scores=0.8310076, pem=-0.15529114,
 similarity=29.733135]
 6: 11867 Germany
 [pem_pos=1, final_scores=0.48758805,
 final_local_scores=0.48727936, pem=-1.0465823,
 similarity=24.322578]
 7: 37414165 Wijnand_van_der_Sanden
 [pem_pos=0, final_scores=0.044845447,
 final_local_scores=0.048743263, pem=0.0, similarity=5.544309]

Sample: b'Central_Appalachian_pine-oak_rocky_woodland'
 Central Appalachian pine-oak rocky woodland is a forest
 system found in [1 1 Maine 1] , 1 Vermont , [2 2 New Hampshire

2] , [3 3 Massachusetts 3] , [4 4 Connecticut 4] , [1 2 New York 1] , [5 5 New Jersey 5] , [6 6 Pennsylvania 6] , [7 7 Ohio 7] , [8 8 Kentucky 8] , [9 9 Virginia 9] , [10 10 Maryland 10] , and [11 11 West Virginia 11] . It occurs on lower-elevation hilltops , 3 outcrop s , and rocky slopes in the central [12 12 Appalachian Mountains 12] , the [13 13 Allegheny Plateau 13] , and the northern [2 4 Piedmont 2] . These forests have a patchy or open aspect . Pitch pine -LRB- [3 5 Pinus 3] rigida -RRB- and [4 Virginia 4] pine -LRB- [5 6 Pinus 5] virginiana -RRB- are common within their respective ranges . These pines are often mixed with dry-site 7 oak s such as chestnut oak -LRB- 8 Quercus prinus -RRB- , bear oak -LRB- [6 9 Quercus 6] ilicifolia -RRB- , northern red oak -LRB- [7 10 Quercus 7] rubra -RRB- , and scarlet oak -LRB- 11 Quercus coccinea -RRB- . Sprouts of chestnut -LRB- -RRB- can also be found . In the northeast , eastern red-cedar -LRB- [8 12 Juniperus 8] virginiana -RRB- or hophornbeam -LRB- 13 Ostrya virginiana -RRB- are sometimes important . In the understory , some areas have a fairly well-developed 14 heath shrub layer , others a layer , the latter particularly common under deciduous trees such as oaks . In the north , this system transitions to the , where 15 spruce begins to appear . In the south it transitions to , where table mountain pine -LRB- -RRB- becomes prominent . It differs from the 16 Central Appalachian dry oak-pine forest

1: 19977 Maine
 2: 21134 New_Hampshire
 3: 1645518 Massachusetts
 4: 6466 Connecticut
 5: 21648 New_Jersey
 6: 23332 Pennsylvania
 7: 22199 Ohio
 8: 16846 Kentucky
 9: 32432 Virginia
 10: 18858 Maryland
 11: 32905 West_Virginia
 12: 85023 Appalachian_Mountains
 13: 609240 Allegheny_Plateau

1: 32578 Vermont

[winner: 32578 Vermont, pem_pos=0, final_scores=-0.10251482, final_local_scores=-0.18132742, pem=-0.2893125, similarity=0.25186747]
 gt: 32578 Vermont, pem_pos=0, final_scores=-0.10251482, final_local_scores=-0.18132742, pem=-0.2893125, similarity=0.25186747]

2: 8210131 New_York

[winner: 645042 New_York_City, pem_pos=1, final_scores=0.15760733, final_local_scores=0.14653635, pem=-1.7906687, similarity=18.200584]
 gt: 8210131 New_York, pem_pos=0, final_scores=0.02755855, final_local_scores=-0.019969098, pem=-0.2889045, similarity=5.064903]

3: 2436357 Outcrop
[winner: 2436357 Outcrop, pem_pos=0,
final_scores=-1.1936709, final_local_scores=-1.285538,
pem=0.0, similarity=-34.273304]
gt: 2436357 Outcrop, pem_pos=0, final_scores=-1.1936709,
final_local_scores=-1.285538, pem=0.0, similarity=-34.273304]
4: 427778 Piedmont_(United_States)
[winner: 25061 Piedmont, pem_pos=0,
final_scores=0.14584963, final_local_scores=0.17833668,
pem=-0.6027155, similarity=12.689296]
gt: 427778 Piedmont_(United_States), pem_pos=1,
final_scores=-0.09202886, final_local_scores=-0.15235457,
pem=-1.4792805, similarity=7.587706]
5: 507788 Pinus_rigida
[winner: 39389 Pine, pem_pos=0, final_scores=0.09764175,
final_local_scores=0.09681323, pem=-0.0019518696,
similarity=6.9894276]
RECALL MISS: 507788 Pinus_rigida]
6: 512381 Pinus_virginiana
[winner: 39389 Pine, pem_pos=0, final_scores=0.106974855,
final_local_scores=0.10700825, pem=-0.0019518696,
similarity=7.293667]
RECALL MISS: 512381 Pinus_virginiana]
[winner: 493120 Magnolia_virginiana, pem_pos=0,
final_scores=-1.4420512, final_local_scores=-1.5088019,
pem=0.0, similarity=-40.935947]
RECALL MISS: 512381 Pinus_virginiana]
7: 39696 Oak
[winner: 39696 Oak, pem_pos=0, final_scores=-0.4877028,
final_local_scores=-0.5589669, pem=-0.08284905,
similarity=-12.140428]
gt: 39696 Oak, pem_pos=0, final_scores=-0.4877028,
final_local_scores=-0.5589669, pem=-0.08284905,
similarity=-12.140428]
8: 66693 Quercus_prinus
[winner: 39696 Oak, pem_pos=0, final_scores=-0.003858164,
final_local_scores=-0.03044013, pem=-0.04793055,
similarity=3.4419725]
RECALL MISS: 66693 Quercus_prinus]
9: 20002273 Quercus_ilicifolia
[winner: 39696 Oak, pem_pos=0, final_scores=0.055133715,
final_local_scores=0.04509285, pem=-0.04793055,
similarity=5.6960273]
RECALL MISS: 20002273 Quercus_ilicifolia]
10: 66687 Quercus_rubra
[winner: 39696 Oak, pem_pos=0, final_scores=0.03817956,
final_local_scores=0.026573002, pem=-0.04793055,
similarity=5.143358]
RECALL MISS: 66687 Quercus_rubra]
[winner: 21040317 Granulosis_rubra_nasi, pem_pos=9,
final_scores=-0.9474, final_local_scores=-0.8072999,
pem=-2.3978953, similarity=-6.9615993]

gt: 66687 *Quercus_rubra*, pem_pos=2,
 final_scores=-1.5315496, final_local_scores=-1.6564968,
 pem=-2.3978953, similarity=-32.30333]
 11: 2604649 *Quercus_coccinea*
 [winner: 39696 Oak, pem_pos=0, final_scores=0.009242222,
 final_local_scores=-0.016129956, pem=-0.04793055,
 similarity=3.8690166]
 RECALL MISS: 2604649 *Quercus_coccinea*
 [winner: 3209075 *Banksia_coccinea*, pem_pos=0,
 final_scores=-1.6538709, final_local_scores=-1.7548156,
 pem=0.0, similarity=-48.27748]
 RECALL MISS: 2604649 *Quercus_coccinea*
 12: 446150 *Juniperus_virginiana*
 [winner: 170104 Juniper, pem_pos=0,
 final_scores=0.12826829, final_local_scores=0.16878638,
 pem=-0.008518757, similarity=9.172961]
 RECALL MISS: 446150 *Juniperus_virginiana*
 [winner: 493120 *Magnolia_virginiana*, pem_pos=0,
 final_scores=-1.4437882, final_local_scores=-1.5106993,
 pem=0.0, similarity=-40.992565]
 RECALL MISS: 446150 *Juniperus_virginiana*
 13: 6881276 *Ostrya_virginiana*
 [winner: 875188 *Ostrya*, pem_pos=0,
 final_scores=-0.09692849, final_local_scores=-0.013262697,
 pem=0.0, similarity=3.6939275]
 RECALL MISS: 6881276 *Ostrya_virginiana*
 [winner: 493120 *Magnolia_virginiana*, pem_pos=0,
 final_scores=-1.4383316, final_local_scores=-1.5047388,
 pem=0.0, similarity=-40.814693]
 RECALL MISS: 6881276 *Ostrya_virginiana*
 14: 461082 Heath
 [winner: 17640767 Heath_railway_station, pem_pos=7,
 final_scores=-1.0277297, final_local_scores=-0.95891047,
 pem=-3.1261127, similarity=-7.5258007]
 gt: 461082 Heath, pem_pos=0, final_scores=-1.2288777,
 final_local_scores=-1.3261065, pem=-1.653199,
 similarity=-26.493595]
 15: 170718 Spruce
 [winner: 170718 Spruce, pem_pos=0,
 final_scores=-1.0344007, final_local_scores=-1.1146483,
 pem=-0.10825555, similarity=-28.584906]
 gt: 170718 Spruce, pem_pos=0, final_scores=-1.0344007,
 final_local_scores=-1.1146483, pem=-0.10825555,
 similarity=-28.584906]
 16: 37538189 Central_Appalachian_dry_oak-pine_forest
 [winner: 37538189 Central_Appalachian_dry_oak-
 pine_forest, pem_pos=0, final_scores=-0.34422797,
 final_local_scores=-0.3058033, pem=0.0,
 similarity=-5.0360675]
 gt: 37538189 Central_Appalachian_dry_oak-pine_forest,
 pem_pos=0, final_scores=-0.34422797,
 final_local_scores=-0.3058033, pem=0.0,

similarity=-5.0360675]

1: 19977 Maine

[pem_pos=0, final_scores=0.037048742,
final_local_scores=-0.036179326, pem=-0.2749413,
similarity=4.5052233]

2: 21134 New_Hampshire

[pem_pos=0, final_scores=0.38184905,
final_local_scores=0.34322628, pem=-0.30258936,
similarity=15.9777975]

3: 1645518 Massachusetts

[pem_pos=0, final_scores=0.47525215,
final_local_scores=0.45680785, pem=-0.1423749,
similarity=18.496027]

4: 6466 Connecticut

[pem_pos=0, final_scores=0.30160332,
final_local_scores=0.2560809, pem=-0.24788867,
similarity=13.079735]

5: 21648 New_Jersey

[pem_pos=0, final_scores=0.38867867,
final_local_scores=0.34716785, pem=-0.13504621,
similarity=15.184296]

6: 23332 Pennsylvania

[pem_pos=0, final_scores=0.36827284,
final_local_scores=0.3170548, pem=-0.11810702,
similarity=14.193545]

7: 22199 Ohio

[pem_pos=0, final_scores=0.54576546,
final_local_scores=0.5142445, pem=-0.19147073,
similarity=20.477041]

8: 16846 Kentucky

[pem_pos=0, final_scores=0.409743,
final_local_scores=0.37738496, pem=-0.28237545,
similarity=16.887236]

9: 32432 Virginia

[pem_pos=0, final_scores=0.51105136,
final_local_scores=0.47476363, pem=-0.2023721,
similarity=19.358137]

10: 18858 Maryland

[pem_pos=0, final_scores=0.3802266,
final_local_scores=0.33632344, pem=-0.2529562,
similarity=15.501891]

11: 32905 West_Virginia

[pem_pos=0, final_scores=0.44153124,
final_local_scores=0.41273165, pem=-0.30025622,
similarity=18.039288]

12: 85023 Appalachian_Mountains

[pem_pos=0, final_scores=0.27726573,
final_local_scores=0.24166137, pem=0.0, similarity=11.301371]

13: 609240 Allegheny_Plateau

[pem_pos=0, final_scores=0.1313722,
final_local_scores=0.16148508, pem=0.0, similarity=8.90875]

1: 645042 New_York_City

[pem_pos=1, final_scores=0.15760733,
 final_local_scores=0.14653635, pem=-1.7906687,
 similarity=18.200584]
 2: 25061 Piedmont
 [pem_pos=0, final_scores=0.14584963,
 final_local_scores=0.17833668, pem=-0.6027155,
 similarity=12.689296]
 3: 39389 Pine
 [pem_pos=0, final_scores=0.09764175,
 final_local_scores=0.09681323, pem=-0.0019518696,
 similarity=6.9894276]
 4: 32432 Virginia
 [pem_pos=0, final_scores=0.26605207,
 final_local_scores=0.20713913, pem=-0.2023721,
 similarity=11.371689]
 5: 39389 Pine
 [pem_pos=0, final_scores=0.106974855,
 final_local_scores=0.10700825, pem=-0.0019518696,
 similarity=7.293667]
 6: 39696 Oak
 [pem_pos=0, final_scores=0.055133715,
 final_local_scores=0.04509285, pem=-0.04793055,
 similarity=5.6960273]
 7: 39696 Oak
 [pem_pos=0, final_scores=0.03817956,
 final_local_scores=0.026573002, pem=-0.04793055,
 similarity=5.143358]
 8: 170104 Juniper
 [pem_pos=0, final_scores=0.12826829,
 final_local_scores=0.16878638, pem=-0.008518757,
 similarity=9.172961]

Sample: b'1994_Winter_Olympics_opening_ceremony'

The 1994 Winter Olympics **1 opening ceremony** took place on 12
 February 1994 at **2 Lyngårdsbakken** in **[1 1 Lillehammer 1]** , **[2**
2 Norway 2] . Artistic content was made to present a range of
[1 Norwegian 1] culture , included **3 Sami 4 joik** , **[2 5 Telem**
ark 2] skiing , fiddlers and folk dancing , simulations of
 traditional weddings and their processions , and **6 vetter**
 from **7 Norse mythology** . After speeches by Heiberg and **[3 IOC**
3] president [3 3 Juan Antonio Samaranch 3] , the games were
 officially declared opened by **8 King [4 Harald V 4]** . The was
 to be carried by a skier down the skijump before lighting the
 cauldron . Originally this task had rested upon **[4 4 Ole Gunn**
ar Fidjestøl 4] , but after he was injured in a practice
 jump , his back-up received the honour . The cauldron was lit
 by . The s were issued by **[5 5 Vegard Ulvang 5]** for the
 athletes and **[6 6 Kari Kåring 6]** for the officials . In the
 day of the opening ceremonies , stole **[7 7 Edvard Munch 7]** 's
 masterpiece **[8 8 The Scream 8]**

1: 64014 Lillehammer

2: 21241 Norway

```

3: 77787 Juan_Antonio_Samaranch
4: 3142354 Ole_Gunnar_Fidjestøl
5: 1248742 Vegard_Ulvang
6: 17959908 Kari_Kåring
7: 9779 Edvard_Munch
8: 212420 The_Scream
1: 4363383 Opening_ceremony
    [winner: 1670957 Moro-Naba_Ceremony, pem_pos=1,
final_scores=-0.98540187, final_local_scores=-0.8861639,
pem=-1.09773, similarity=-16.385557]
    RECALL MISS: 4363383 Opening_ceremony]
    [winner: 645748 Opening_(morphology), pem_pos=9,
final_scores=-1.4592317, final_local_scores=-1.3814635,
pem=-2.5904672, similarity=-23.04855]
    gt: 4363383 Opening_ceremony, pem_pos=19,
final_scores=-2.8430603, final_local_scores=-3.0827072,
pem=-5.1161957, similarity=-60.08175 ]
2: 4220684 Lysgårdsbakken
    [winner: 4220684 Lysgårdsbakken, pem_pos=0,
final_scores=-0.093312785, final_local_scores=-0.043613993,
pem=0.0, similarity=2.7881844]
    gt: 4220684 Lysgårdsbakken, pem_pos=0,
final_scores=-0.093312785, final_local_scores=-0.043613993,
pem=0.0, similarity=2.7881844 ]
3: 27122 Sami_people
    [winner: 27122 Sami_people, pem_pos=0,
final_scores=0.02055116, final_local_scores=0.009686649,
pem=-0.8177901, similarity=8.826052]
    gt: 27122 Sami_people, pem_pos=0,
final_scores=0.02055116, final_local_scores=0.009686649,
pem=-0.8177901, similarity=8.826052 ]
4: 432864 Joik
    [winner: 432864 Joik, pem_pos=0, final_scores=-1.4490507,
final_local_scores=-1.5107338, pem=-0.026712772,
similarity=-40.848328]
    gt: 432864 Joik, pem_pos=0, final_scores=-1.4490507,
final_local_scores=-1.5107338, pem=-0.026712772,
similarity=-40.848328 ]
5: 175092 Telemark_skiing
    [winner: 175076 Telemark, pem_pos=0,
final_scores=0.24332292, final_local_scores=0.25044373,
pem=-0.12048148, similarity=12.218651]
    gt: 175092 Telemark_skiing, pem_pos=3,
final_scores=-0.4102825, final_local_scores=-0.4068937,
pem=-2.631811, similarity=6.2593904 ]
    [winner: 2955619 Ski_(driving_stunt), pem_pos=4,
final_scores=-1.0853015, final_local_scores=-1.0117526,
pem=-1.6015035, similarity=-17.39378]
    RECALL MISS: 175092 Telemark_skiing]
6: 1812610 Vættir
    [winner: 842981 Helmut_Vetter, pem_pos=2,
final_scores=-0.77054965, final_local_scores=-0.6499354,

```

```

pem=-2.1972246, similarity=-3.356808|
    gt: 1812610 Vættir, pem_pos=5, final_scores=-1.1881,
final_local_scores=-1.1721613, pem=-2.1972246,
similarity=-18.941069 ]
7: 23417070 Norse_mythology
    [winner: 32610 Vikings, pem_pos=0,
final_scores=-0.0031151474, final_local_scores=0.025380641,
pem=-0.7229763, similarity=8.778779|
    gt: 23417070 Norse_mythology, pem_pos=2,
final_scores=-0.19497459, final_local_scores=-0.19275315,
pem=-1.4024369, similarity=5.964244 ]
    [winner: 5873281 Māui (Māori_mythology), pem_pos=4,
final_scores=-1.4477971, final_local_scores=-1.4298108,
pem=-1.9459101, similarity=-27.996532|
    RECALL MISS: 23417070 Norse_mythology]
8: 66069 Harald_V_of_Norway
    [winner: 66069 Harald_V_of_Norway, pem_pos=0,
final_scores=0.12901352, final_local_scores=0.14521399,
pem=0.0, similarity=8.423188|
    gt: 66069 Harald_V_of_Norway, pem_pos=0,
final_scores=0.12901352, final_local_scores=0.14521399,
pem=0.0, similarity=8.423188 ]
    [winner: 19012 Monarch, pem_pos=0,
final_scores=-1.2093207, final_local_scores=-1.2860448,
pem=-0.9897082, similarity=-28.906244|
    RECALL MISS: 66069 Harald_V_of_Norway]
1: 64014 Lillehammer
    [pem_pos=0, final_scores=0.4807781,
final_local_scores=0.46244162, pem=0.0, similarity=17.889893]
2: 21241 Norway
    [pem_pos=0, final_scores=0.7432826,
final_local_scores=0.75775063, pem=-0.22668563,
similarity=27.935255]
3: 77787 Juan_Antonio_Samaranch
    [pem_pos=0, final_scores=0.27405977,
final_local_scores=0.33523628, pem=0.0, similarity=14.093832]
4: 3142354 Ole_Gunnar_Fidjestøl
    [pem_pos=0, final_scores=0.13403545,
final_local_scores=0.23449543, pem=0.0, similarity=11.087524]
5: 1248742 Vegard_Ulvang
    [pem_pos=0, final_scores=0.33832377,
final_local_scores=0.4245264, pem=0.0, similarity=16.758427]
6: 17959908 Kari_Kåring
    [pem_pos=0, final_scores=0.1793376,
final_local_scores=0.3045072, pem=0.0, similarity=13.176815]
7: 9779 Edvard_Munch
    [pem_pos=0, final_scores=0.325792,
final_local_scores=0.38013643, pem=0.0, similarity=15.433742]
8: 212420 The_Scream
    [pem_pos=0, final_scores=0.06938623,
final_local_scores=0.111092284, pem=0.0, similarity=7.404928]
1: 21241 Norway

```

[pem_pos=0, final_scores=0.48564273,
final_local_scores=0.47631818, pem=-0.605967,
similarity=21.59934]
2: 175076 Telemark
[pem_pos=0, final_scores=0.24332292,
final_local_scores=0.25044373, pem=-0.12048148,
similarity=12.218651]
3: 7971634 Indian_Oil_Corporation
[pem_pos=2, final_scores=0.38914168,
final_local_scores=0.5597529, pem=-1.36746,
similarity=28.230316]
4: 66069 Harald_V_of_Norway
[pem_pos=0, final_scores=0.12901352,
final_local_scores=0.14521399, pem=0.0, similarity=8.423188]

Sample: b'Children_in_Need_Rocks_Manchester'

Children in Need Rocks Manchester was a charity music concert held at the [1 1 Manchester Arena 1] in [2 2 Manchester 2] , [1 England 1] , on 17 November 2011 . The concert was organised by [3 3 Take That 3] singer and [2 1 The X Factor 2] judge [4 4 Gary Barlow 4] as one of a series of events to raise money for [3 2 Children in Need 3] 2011 . It became the second Children in Need Rocks concert organised by [4 Barlow 4] , after the 3 Children in Need Rocks the [5 Royal Albert Hall 5] in 2009 . The concert was broadcast on [5 5 BBC One 5] , [6 BBC One HD 6] and [6 6 BBC Radio 1 6] on Thursday 17 November 2011 , the day before the 4 official appeal telethon . Short films of projects being funded by the charity were shown at various points throughout the show , often featuring one of the celebrities meeting a child or group of children . The event was hosted by radio and television presenters [7 7 Chris Moyles 7] and [8 8 Fearne Cotton 8] , along with former [9 9 Doctor Who 9] actor [10 10 David Tennant 10] . The acts performing at the concert included [7 Canadian 7] [11 11 Michael Bublé 11] , [12 12 Jessie J 12] , [13 13 Coldplay 13] , [14 14 James Morrison 14] , [8 Barlow 8] 's fellow [9 The X Factor 9] judges , [10 5 Tulisa Contostavlos 10] and [15 15 Kelly Rowland 15] , [11 Barlow 11] himself and [16 16 Lady Gaga 16] . [12 Barlow 12] revealed on 6 The Chris Moyles Show when announcing the concert that he had personally contacted the acts he wanted to perform . Tickets for the concert cost between # 55 and # 95 , and the 12,000 tickets sold out within 10 minutes of going on sale on 16 September 2011 . Some tickets were also bought by radio stations and television programmes to be auctioned , with one pair of tickets being sold to a [17 17 BBC Radio Jersey 17]

1: 511678 Manchester_Arena
2: 20206 Manchester
3: 371370 Take_That
4: 755424 Gary_Barlow
5: 212157 BBC_One
6: 4348 BBC_Radio_1

7: 341294 Chris_Moyles
 8: 21931784 Fearne_Cotton
 9: 8209 Doctor_Who
 10: 1672146 David_Tennant
 11: 621503 Michael_Bubl  
 12: 29427669 Jessie_J
 13: 80103 Coldplay
 14: 4683946 James_Morrison_(singer)
 15: 261035 Kelly_Rowland
 16: 17782843 Lady_Gaga
 17: 569498 BBC_Radio_Jersey
 1: 31774825 The_X_Factor_(UK_series_8)
 [winner: 14724189 The_X_Factor_(UK_TV_series), pem_pos=0,
 final_scores=0.1319985, final_local_scores=0.056289643,
 pem=-0.43565995, similarity=8.138695]
 gt: 31774825 The_X_Factor_(UK_series_8), pem_pos=16,
 final_scores=-0.7396886, final_local_scores=-0.86677474,
 pem=-5.61073, similarity=8.735446]
 2: 33793527 Children_in_Need_2011
 [winner: 1073488 Children_in_Need, pem_pos=0,
 final_scores=0.046651885, final_local_scores=-0.01879131,
 pem=0.0, similarity=3.5289428]
 gt: 33793527 Children_in_Need_2011, pem_pos=7,
 final_scores=-1.3938313, final_local_scores=-1.5278816,
 pem=-7.9218984, similarity=1.5751891]
 [winner: 36225 2011, pem_pos=0, final_scores=-1.1023595,
 final_local_scores=-1.0134237, pem=-1.4232541,
 similarity=-18.412996]
 RECALL MISS: 33793527 Children_in_Need_2011]
 3: 33828926 Children_in_Need_Rocks_the_Royal_Albert_Hall
 [winner: 179238 Royal_Albert_Hall, pem_pos=0,
 final_scores=0.19089155, final_local_scores=0.15990749,
 pem=0.0, similarity=8.861671]
 RECALL MISS: 33828926
 Children_in_Need_Rocks_the_Royal_Albert_Hall]
 [winner: 8867388 Islam_and_children, pem_pos=0,
 final_scores=-0.57669973, final_local_scores=-0.44911087,
 pem=-1.049245, similarity=-3.606693]
 RECALL MISS: 33828926
 Children_in_Need_Rocks_the_Royal_Albert_Hall]
 [winner: 196054 Rocks_(Aerosmith_album), pem_pos=0,
 final_scores=-0.63653696, final_local_scores=-0.6581158,
 pem=-0.5548053, similarity=-12.532659]
 RECALL MISS: 33828926
 Children_in_Need_Rocks_the_Royal_Albert_Hall]
 [winner: 24489875 Need_(album), pem_pos=2,
 final_scores=-0.95709395, final_local_scores=-0.9234349,
 pem=-1.5179106, similarity=-15.212793]
 RECALL MISS: 33828926
 Children_in_Need_Rocks_the_Royal_Albert_Hall]
 4: 33793527 Children_in_Need_2011
 [winner: 16181295 Appeal_(motion), pem_pos=2,

```

final_scores=-1.115533, final_local_scores=-1.0817058,
pem=-1.3665628, similarity=-20.758966|
    RECALL MISS: 33793527 Children_in_Need_2011]
    [winner: 10549754 Anne_de_Chantaine, pem_pos=11,
final_scores=-1.284297, final_local_scores=-1.2447622,
pem=-2.3025851, similarity=-20.534658|
    RECALL MISS: 33793527 Children_in_Need_2011]
    [winner: 4151405 ITV_Telethon, pem_pos=0,
final_scores=-1.287696, final_local_scores=-1.3811274,
pem=0.0, similarity=-37.12588|
    RECALL MISS: 33793527 Children_in_Need_2011]
5: 22731525 Tulisa_Contostavlos
    [winner: 14292104 N-Dubz, pem_pos=1,
final_scores=0.38055438, final_local_scores=0.36338967,
pem=-0.23072718, similarity=16.188715|
    gt: 22731525 Tulisa_Contostavlos, pem_pos=0,
final_scores=0.373892, final_local_scores=0.362824, pem=0.0,
similarity=14.917105 ]
6: 1925261 The_Chris_Moyles_Show
    [winner: 1925261 The_Chris_Moyles_Show, pem_pos=0,
final_scores=-0.15761258, final_local_scores=-0.2178708,
pem=0.0, similarity=-2.4119864|
    gt: 1925261 The_Chris_Moyles_Show, pem_pos=0,
final_scores=-0.15761258, final_local_scores=-0.2178708,
pem=0.0, similarity=-2.4119864 ]
1: 511678 Manchester_Arena
    [pem_pos=0, final_scores=0.25914085,
final_local_scores=0.26904953, pem=0.0, similarity=12.118688]
2: 20206 Manchester
    [pem_pos=0, final_scores=0.4428851,
final_local_scores=0.4732532, pem=-0.3209542,
similarity=19.957932]
3: 371370 Take_That
    [pem_pos=0, final_scores=0.24399517,
final_local_scores=0.18768963, pem=0.0, similarity=9.690746]
4: 755424 Gary_Barlow
    [pem_pos=0, final_scores=0.27820665,
final_local_scores=0.2098028, pem=0.0, similarity=10.350647]
5: 212157 BBC_One
    [pem_pos=0, final_scores=0.15767647,
final_local_scores=0.1262596, pem=0.0, similarity=7.857551]
6: 4348 BBC_Radio_1
    [pem_pos=0, final_scores=0.05950047,
final_local_scores=-0.014233701, pem=0.0,
similarity=3.6649508]
7: 341294 Chris_Moyles
    [pem_pos=0, final_scores=0.31563073,
final_local_scores=0.30008557, pem=0.0, similarity=13.044865]
8: 21931784 Fearne_Cotton
    [pem_pos=0, final_scores=0.2902481,
final_local_scores=0.26195067, pem=0.0, similarity=11.906843]
9: 8209 Doctor_Who

```

[pem_pos=0, final_scores=0.17233758,
final_local_scores=0.18159881, pem=-0.026654126,
similarity=9.653933]
10: 1672146 David_Tennant
[pem_pos=0, final_scores=0.25697815,
final_local_scores=0.26283106, pem=0.0, similarity=11.933116]
11: 621503 Michael_Bubl 
[pem_pos=0, final_scores=0.30636406,
final_local_scores=0.27256402, pem=-0.028789124,
similarity=12.380127]
12: 29427669 Jessie_J
[pem_pos=0, final_scores=0.29341954,
final_local_scores=0.24255547, pem=0.0, similarity=11.328052]
13: 80103 Coldplay
[pem_pos=0, final_scores=0.23193638,
final_local_scores=0.18176913, pem=0.0, similarity=9.514067]
14: 4683946 James_Morrison_(singer)
[pem_pos=0, final_scores=0.26550692,
final_local_scores=0.2319799, pem=-0.47383493,
similarity=13.589245]
15: 261035 Kelly_Rowland
[pem_pos=0, final_scores=0.38582116,
final_local_scores=0.35840878, pem=0.0, similarity=14.785346]
16: 17782843 Lady_Gaga
[pem_pos=0, final_scores=0.36243725,
final_local_scores=0.33511007, pem=0.0, similarity=14.090066]
17: 569498 BBC_Radio_Jersey
[pem_pos=0, final_scores=0.034071133,
final_local_scores=0.14094007, pem=0.0, similarity=8.295646]
1: 9316 England
[pem_pos=0, final_scores=0.52432716,
final_local_scores=0.60077935, pem=-0.19053856,
similarity=23.054346]
2: 14724189 The_X_Factor_(UK_TV_series)
[pem_pos=0, final_scores=0.1319985,
final_local_scores=0.056289643, pem=-0.43565995,
similarity=8.138695]
3: 1073488 Children_in_Need
[pem_pos=0, final_scores=0.046651885,
final_local_scores=-0.01879131, pem=0.0, similarity=3.5289428]
4: 755424 Gary_Barlow
[pem_pos=0, final_scores=0.52894807,
final_local_scores=0.48369974, pem=0.0, similarity=18.524279]
5: 179238 Royal_Albert_Hall
[pem_pos=0, final_scores=0.19089155,
final_local_scores=0.15990749, pem=0.0, similarity=8.861671]
6: 212157 BBC_One
[pem_pos=0, final_scores=0.05449097,
final_local_scores=0.013545126, pem=0.0, similarity=4.4939265]
7: 5042916 Canada
[pem_pos=0, final_scores=0.60053974,
final_local_scores=0.7026142, pem=-0.41303828,

```

similarity=27.303288]
8: 755424 Gary_Barlow
   [pem_pos=0, final_scores=0.5054655,
   final_local_scores=0.45804858, pem=0.0, similarity=17.758797]
9: 14724189 The_X_Factor_(UK_TV_series)
   [pem_pos=0, final_scores=0.1383933,
   final_local_scores=0.06327501, pem=-0.43565995,
   similarity=8.347152]
10: 14292104 N-Dubz
    [pem_pos=1, final_scores=0.38055438,
    final_local_scores=0.36338967, pem=-0.23072718,
    similarity=16.188715]
11: 755424 Gary_Barlow
    [pem_pos=0, final_scores=0.52404314,
    final_local_scores=0.47834182, pem=0.0, similarity=18.364388]
12: 755424 Gary_Barlow
    [pem_pos=0, final_scores=0.4938889,
    final_local_scores=0.44540292, pem=0.0, similarity=17.381424]

```

Sample: b'MV_Tustumena'

M/V Tustumena ' is a mainline **1 ferry** vessel for the . Tustumena was constructed in 1963 in [1 1 Sturgeon Bay 1] , [2 2 Wisconsin 2] and refurbished in 1969 in . As the only mainline ferry in South-central [1 Alaska 1] and the [3 3 Aleutian Chain 3] , it principally runs between [4 4 Kodiak 4] , [5 5 Seldovia 5] , [6 6 Port Lions 6] , and **2 Homer** with Homer providing a road link to the other communities on the [7 7 Kenai Peninsula 7] and [8 8 Kodiak Island 8] area . The only interruptions from this schedule occur when making a voyage out the [2 Aleutian Chain 2] -LRB- the [3 Aleutian Chain 3] run consists of the communities of [9 9 Akutan 9] , [10 10 Chignik 10] , [11 11 Cold Bay 11] , [12 12 False Pass 12] , [13 13 King Cove 13] , [14 14 Sand Point 14] , and [15 15 Unalaska 15] /

```

1: 529459 Sturgeon_Bay,_Wisconsin
2: 33127 Wisconsin
3: 26296222 Aleutian_Islands
4: 105505 Kodiak,_Alaska
5: 105489 Seldovia,_Alaska
6: 105510 Port_Lions,_Alaska
7: 413542 Kenai_Peninsula
8: 615988 Kodiak_Island
9: 105201 Akutan,_Alaska
10: 105513 Chignik,_Alaska
11: 105202 Cold_Bay,_Alaska
12: 105203 False_Pass,_Alaska
13: 105204 King_Cove,_Alaska
14: 105206 Sand_Point,_Alaska
15: 105213 Unalaska,_Alaska

```

```

1: 50771 Ferry
   [winner: 23544400 Ferry_River, pem_pos=12,
   final_scores=-1.0646694, final_local_scores=-0.9760017,

```



```
pem=-3.47156, similarity=-6.15724|
  gt: 50771 Ferry, pem_pos=0, final_scores=-1.2046535,
final_local_scores=-1.3190405, pem=-0.71013147,
similarity=-31.411278 ]
2: 105472 Homer, Alaska
  [winner: 105472 Homer, Alaska, pem_pos=2,
final_scores=-0.12782459, final_local_scores=-0.20257789,
pem=-3.135527, similarity=15.095866|
  gt: 105472 Homer, Alaska, pem_pos=2,
final_scores=-0.12782459, final_local_scores=-0.20257789,
pem=-3.135527, similarity=15.095866 ]
1: 529459 Sturgeon_Bay, Wisconsin
  [pem_pos=0, final_scores=0.3421805,
final_local_scores=0.3879205, pem=-0.004405293,
similarity=15.68999]
2: 33127 Wisconsin
  [pem_pos=0, final_scores=0.4285248,
final_local_scores=0.4938541, pem=-0.21268336,
similarity=19.98391]
3: 26296222 Aleutian_Islands
  [pem_pos=0, final_scores=0.3260097,
final_local_scores=0.28971392, pem=0.0, similarity=12.735354]
4: 105505 Kodiak, Alaska
  [pem_pos=0, final_scores=0.33581197,
final_local_scores=0.31023645, pem=-0.36096606,
similarity=15.310778]
5: 105489 Seldovia, Alaska
  [pem_pos=0, final_scores=0.20226623,
final_local_scores=0.17617625, pem=0.0, similarity=9.347163]
6: 105510 Port_Lions, Alaska
  [pem_pos=0, final_scores=0.11798705,
final_local_scores=0.13490608, pem=0.0, similarity=8.11558]
7: 413542 Kenai_Peninsula
  [pem_pos=0, final_scores=0.32325488,
final_local_scores=0.31366456, pem=0.0, similarity=13.4500885]
8: 615988 Kodiak_Island
  [pem_pos=0, final_scores=0.25571865,
final_local_scores=0.22362712, pem=0.0, similarity=10.763192]
9: 105201 Akutan, Alaska
  [pem_pos=0, final_scores=0.22205763,
final_local_scores=0.21373999, pem=0.0, similarity=10.468141]
10: 105513 Chignik, Alaska
  [pem_pos=0, final_scores=0.17340516,
final_local_scores=0.17827737, pem=0.0, similarity=9.409865]
11: 105202 Cold_Bay, Alaska
  [pem_pos=0, final_scores=0.24128018,
final_local_scores=0.23102853, pem=0.0, similarity=10.984065]
12: 105203 False_Pass, Alaska
  [pem_pos=0, final_scores=0.12144469,
final_local_scores=0.11832273, pem=0.0, similarity=7.620699]
13: 105204 King_Cove, Alaska
  [pem_pos=0, final_scores=0.22447129,
```

```

final_local_scores=0.21911529, pem=0.0, similarity=10.628551]
14: 105206 Sand_Point, Alaska
    [pem_pos=0, final_scores=0.068951264,
final_local_scores=0.042229176, pem=-0.60738033,
similarity=8.652945]
15: 105213 Unalaska, Alaska
    [pem_pos=0, final_scores=0.3222453,
final_local_scores=0.29184926, pem=0.0, similarity=12.799077]
1: 624 Alaska
    [pem_pos=0, final_scores=0.52049154,
final_local_scores=0.5005507, pem=-0.18647984,
similarity=20.04125]
2: 26296222 Aleutian_Islands
    [pem_pos=0, final_scores=0.1831408,
final_local_scores=0.1336514, pem=0.0, similarity=8.078137]
3: 26296222 Aleutian_Islands
    [pem_pos=0, final_scores=0.23518758,
final_local_scores=0.19050461, pem=0.0, similarity=9.774751]

```

Sample: b'Nelsonic_Industries'

Nelsonic Industries is the name of an electronics manufacturing and development company that operated from -LRB- [1 1 New York City 1] -RRB- in the early 1980s and throughout the 1990s when it was acquired by the watch-manufacturer , M.Z. Berger . Nelsonic produced numerous toy-themed wrist-watches during their existence , often targeting younger audiences with likenesses of characters from popular franchises such as [2 2 Barbie 2] , the [3 3 Ghostbusters 3] , [4 4 Mario 4] , and . Nelsonic became notable during the early mid-1980s for being the first electronics company in the [1 United States 1] to produce game-watches -LRB- multi-purpose electronic devices capable of functioning as both a time-piece and as a typically electronic game -RRB- . For a period subsequent to its purchase by M.Z. Berger , Nelsonic operated as a subsidiary division of its parent company and game-watches were produced that bore the Nelsonic mark . This practice ended as M.Z. Berger shifted focus to more traditional and higher-end timepieces . Today the original Nelsonic Game Watch line has entered the secondary market and individual Game Watches have become highly sought-after collectibles that often fetch high prices on [2 eBay 2] and other online auction websites . Throughout its existence , Nelsonic produced pop-culture-themed wrist-watches for children and young adults . The chronograph digital watches , typically made of molded plastic , invariably featured an alarm and utilized display-screens to display the time for their wearers . In time the company began manufacturing multi-purpose units that used the [3 LCD 3] screen to combine time display functions with simple video game functions . These simple video games were variations on the theme of the 1 Calculator watch , and the patent covering the Game Watch line is in fact classified as an `` electronic calculator watch

structure . '' Becoming the first electronics manufacturer in the [4 United States 4] to produce game-watches , Nelsonic earned a large share in this specialized market and was able to earn the attention of large video game companies at events such as the . Having caught the popular attention and with goodwill at a high-point , Nelsonic was able to obtain licensing from several big-name video game companies such as [5 5 Sega 5] , [6 6 Nintendo 6] , [7 7 Midway Games 7] , and Mylstar Electronics . With roots in the toy market , Nelsonic was also able to obtain similar licenses to produce [5 LCD 5] versions of popular electronic toys like [6 2 Milton Bradley 6] 's 3 Simon as well as to produce original [7 LCD 7] games for non-game toy franchises like [8 8 Barbie 8] , 4 G.I. Joe , and [9 9 Power Rangers 9] and even for film and TV franchises such as [10 10 Ghostbusters 10] . So well-known became the Nelsonic brand that it was even able to license its own original games -LRB- such as the Nelsonic version of the public-domain 5 blackjack which was licensed to [11 11 Caesars Palace 11] -RRB- . As time progressed , Nelsonic experimented with higher-end products such as metal watches and increasingly complex game-watch designs . In 1990 the watch-making giant , M.Z. Berger , made a large bid and was able to successfully purchase the company . For a period of nearly a decade after this acquisition , M.Z. Berger continued to use Nelsonic as a subsidiary branch and to employ the [8 Nelsonic 8] mark in the release of game-watches and the production of new re-releases of popular models from the 1980s and early 1990s . By the end of the 1990s , however , public interest had waned -LRB- quite possibly due to the rise in popularity of more advanced s and , eventually , of other portable computing devices , such as s and -RRB- and this practice came to an end as M.Z. Berger shifted markets to target higher end consumers more exclusively . Today the Nelsonic mark is still in use for traditional watches and is not used in connection with game-watches . It is still a subsidiary of M.Z. Berger and as recently as 2007 it was listed by the

1: 645042 New_York_City

2: 164328 Barbie

3: 205013 Ghostbusters

4: 81361 Mario

5: 28222625 Sega

6: 21197 Nintendo

7: 430266 Midway_Games

8: 164328 Barbie

9: 173486 Power_Rangers

10: 205013 Ghostbusters

11: 839916 Caesars_Palace

1: 2867523 Calculator_watch

[winner: 7593 Calculator, pem_pos=0,
final_scores=-0.2575831, final_local_scores=-0.27738163,
pem=-0.7833525, similarity=0.07208438]

RECALL MISS: 2867523 Calculator_watch]
[winner: 9214462 Watch_(Unix), pem_pos=12,
final_scores=-1.3744633, final_local_scores=-1.334171,
pem=-2.6390574, similarity=-21.373005]
RECALL MISS: 2867523 Calculator_watch]
2: 154681 Milton_Bradley_Company
[winner: 7576369 Milton_Bradley, pem_pos=0,
final_scores=0.035293773, final_local_scores=0.03186226,
pem=-0.4746027, similarity=7.621509]
gt: 154681 Milton_Bradley_Company, pem_pos=1,
final_scores=0.002404198, final_local_scores=-0.032572232,
pem=-0.5586286, similarity=6.155602]
3: 593761 Simon_(game)
[winner: 593761 Simon_(game), pem_pos=0,
final_scores=-0.3670106, final_local_scores=-0.39298025,
pem=-0.7344016, similarity=-3.6438117]
gt: 593761 Simon_(game), pem_pos=0,
final_scores=-0.3670106, final_local_scores=-0.39298025,
pem=-0.7344016, similarity=-3.6438117]
4: 28713688 G.I._Joe
[winner: 5845030 G.I._(Annoyed_Grunt), pem_pos=2,
final_scores=-0.86853796, final_local_scores=-0.8392729,
pem=-1.2733762, similarity=-14.031048]
RECALL MISS: 28713688 G.I._Joe]
[winner: 835997 Joe_(singer), pem_pos=0,
final_scores=-1.1388986, final_local_scores=-1.1888602,
pem=-1.3273427, similarity=-24.169952]
RECALL MISS: 28713688 G.I._Joe]
5: 3981 Blackjack
[winner: 12841199 Baton_(law_enforcement), pem_pos=0,
final_scores=-1.0502976, final_local_scores=-1.0671544,
pem=0.0, similarity=-27.756304]
gt: 3981 Blackjack, pem_pos=1, final_scores=-1.4688256,
final_local_scores=-1.6029112, pem=-0.09315106,
similarity=-43.237785]
1: 645042 New_York_City
[pem_pos=0, final_scores=0.44913292,
final_local_scores=0.48262513, pem=0.0, similarity=18.492208]
2: 164328 Barbie
[pem_pos=0, final_scores=0.14484273,
final_local_scores=0.12151033, pem=0.0, similarity=7.7158237]
3: 205013 Ghostbusters
[pem_pos=0, final_scores=0.074953064,
final_local_scores=0.05659081, pem=0.0, similarity=5.778496]
4: 81361 Mario
[pem_pos=0, final_scores=0.033074334,
final_local_scores=0.02060543, pem=-0.22842951,
similarity=5.946855]
5: 28222625 Sega
[pem_pos=0, final_scores=0.08081211,
final_local_scores=0.024274543, pem=0.0, similarity=4.8141136]
6: 21197 Nintendo

[pem_pos=0, final_scores=0.18409239,
 final_local_scores=0.13209951, pem=0.0, similarity=8.031826]
 7: 430266 Midway_Games
 [pem_pos=0, final_scores=0.16930155,
 final_local_scores=0.1495916, pem=0.0, similarity=8.553824]
 8: 164328 Barbie
 [pem_pos=0, final_scores=0.16107206,
 final_local_scores=0.13923842, pem=0.0, similarity=8.244865]
 9: 173486 Power_Rangers
 [pem_pos=0, final_scores=0.07869591,
 final_local_scores=0.06254397, pem=0.0, similarity=5.95615]
 10: 205013 Ghostbusters
 [pem_pos=0, final_scores=0.08256216,
 final_local_scores=0.0649026, pem=0.0, similarity=6.026536]
 11: 839916 Caesars_Palace
 [pem_pos=0, final_scores=0.21100406,
 final_local_scores=0.24906811, pem=0.0, similarity=11.522402]
 1: 3434750 United_States
 [pem_pos=0, final_scores=0.4306609,
 final_local_scores=0.47158402, pem=-0.073457584,
 similarity=18.562195]
 2: 130495 EBay
 [pem_pos=1, final_scores=0.14520071,
 final_local_scores=0.1406933, pem=-0.023997083,
 similarity=8.418781]
 3: 17932 Liquid-crystal_display
 [pem_pos=0, final_scores=0.12087546,
 final_local_scores=0.12825096, pem=0.0, similarity=7.9169774]
 4: 3434750 United_States
 [pem_pos=0, final_scores=0.45148504,
 final_local_scores=0.49433118, pem=-0.073457584,
 similarity=19.241016]
 5: 17932 Liquid-crystal_display
 [pem_pos=0, final_scores=0.0436251,
 final_local_scores=0.043866694, pem=0.0, similarity=5.3987827]
 6: 7576369 Milton_Bradley
 [pem_pos=0, final_scores=0.035293773,
 final_local_scores=0.03186226, pem=-0.4746027,
 similarity=7.621509]
 7: 17932 Liquid-crystal_display
 [pem_pos=0, final_scores=0.05811219,
 final_local_scores=0.059691638, pem=0.0, similarity=5.8710303]
 8: 26608947 Nelsonic_Industries
 [pem_pos=0, final_scores=0.036713824,
 final_local_scores=0.03752123, pem=0.0, similarity=5.209421]

Sample: b'Steadfastness_and_Confrontation_Front'

The Steadfastness and Confrontation Front -LRB- -RRB- was a
 political initiative by a number of governments in the
 1970s , related to the . The Steadfastness and Confrontation
 Front was formed in 1977 by the governments of [1 1 Libya 1] ,
 [2 2 Algeria 2] , [3 3 Syria 3] and [4 4 South Yemen 4] , as

well as by the 1 Palestine Liberation Organization -LRB- PLO - RRB- . It was intended as a protest and a show of position , after 2 President [5 5 Anwar Sadat 5] of [6 6 Egypt 6] had travelled to to meet [7 7 Israel 7] i 3 Prime Minister [8 8 Menachem Begin 8] , and begin peace negotiations that would eventually lead to the [9 9 Camp David Accords 9] . This [1 Egyptian 1] initiative was widely seen in the [2 4 Arab 2] world as an abandonment of agreed-upon principles of non-recognition of [3 Israel 3] , and as breaking the [4 Arab 4] alliance against [5 Israel 5] . It was generally condemned as treachery . The front affirmed its rejection of resolutions and , and reiterated the unwillingness to recognize [6 Israel 6] or negotiate with it for a . It also called for a 5 boycott of [7 Egypt 7] , condemned every [8 Arab 8] government who did not take part of the Front , and emphasized the ties between [9 Syria 9] and the [10 Palestinians 10] . Still , interestingly , it did not explicitly call for [11 Israel 11] 's destruction , but instead repeated the of the PLO , calling for a [12 Palestinian 12] state on `` any part of [13 Palestinian 13] land ... as an interim aim of the Palestinian 6 Revolution '' . This had been interpreted , at least inside the PLO , as a step towards a 7 two-state solution , and was highly controversial among . The countries of the Front also recognized the [10 10 Sahrawi Arab Democratic Republic 10] as the sovereign entity of the [11 11 Western Sahara 11] territory . The Steadfastness and Confrontation Front should not be confused with the [12 12 Rejectionist Front 12] , which was a gathering of radical [14 Palestinian 14] factions outside the framework of the PLO . They had left the PLO after [13 13 Fatah 13] pushed through the [15 Ten Point Program 15] in 1974 . In fact , the Steadfastness and Confrontation Front and its conflict with [16 Egypt 16] - the most influential [17 Arab 17] state - helped bring together the fractured PLO , and temporarily heal the rift between the main organization under

1: 17633 Libya

2: 358 Algeria

3: 7515849 Syria

4: 718749 South_Yemen

5: 49522 Anwar_Sadat

6: 8087628 Egypt

7: 9282173 Israel

8: 20767 Menachem_Begin

9: 181901 Camp_David_Accords

10: 38762388 Sahrawi_Arab_Democratic_Republic

11: 33220 Western_Sahara

12: 3008136 Rejectionist_Front

13: 11104 Fatah

1: 24324 Palestine_Liberation_Organization

[winner: 24324 Palestine_Liberation_Organization,

pem_pos=0, final_scores=-0.03363301,

final_local_scores=-0.11704985, pem=0.0, similarity=0.596712]

gt: 24324 Palestine_Liberation_Organization, pem_pos=0,
final_scores=-0.03363301, final_local_scores=-0.11704985,
pem=0.0, similarity=0.596712]
2: 335341 President_of_Egypt
[winner: 24113 President_of_the_United_States, pem_pos=0,
final_scores=-1.5719141, final_local_scores=-1.6926146,
pem=-0.99314094, similarity=-41.020428]
RECALL MISS: 335341 President_of_Egypt]
3: 24898 Prime_Minister_of_Israel
[winner: 23666 Prime_number, pem_pos=2,
final_scores=-0.64401567, final_local_scores=-0.5696924,
pem=-2.391252, similarity=0.09295766]
RECALL MISS: 24898 Prime_Minister_of_Israel]
[winner: 739237 Minister_(government), pem_pos=0,
final_scores=-1.0211729, final_local_scores=-1.090908,
pem=-0.49105036, similarity=-25.79475]
RECALL MISS: 24898 Prime_Minister_of_Israel]
4: 159433 Arab_world
[winner: 159433 Arab_world, pem_pos=14,
final_scores=0.3531987, final_local_scores=0.3084643,
pem=-3.8052368, similarity=33.988373]
gt: 159433 Arab_world, pem_pos=14,
final_scores=0.3531987, final_local_scores=0.3084643,
pem=-3.8052368, similarity=33.988373]
[winner: 49819 Gross_world_product, pem_pos=0,
final_scores=-1.6636499, final_local_scores=-1.7004054,
pem=-1.9928869, similarity=-35.816143]
RECALL MISS: 159433 Arab_world]
5: 18591996 Boycott
[winner: 18591996 Boycott, pem_pos=0,
final_scores=-0.9917589, final_local_scores=-1.0807633,
pem=0.0, similarity=-28.16242]
gt: 18591996 Boycott, pem_pos=0,
final_scores=-0.9917589, final_local_scores=-1.0807633,
pem=0.0, similarity=-28.16242]
6: 25964 Revolution
[winner: 25964 Revolution, pem_pos=0,
final_scores=-1.1144246, final_local_scores=-1.2271187,
pem=-1.8644581, similarity=-22.390745]
gt: 25964 Revolution, pem_pos=0,
final_scores=-1.1144246, final_local_scores=-1.2271187,
pem=-1.8644581, similarity=-22.390745]
7: 40863732 Two-state_solution
[winner: 40863732 Two-state_solution, pem_pos=0,
final_scores=-1.0548786, final_local_scores=-1.1753745,
pem=-0.24512246, similarity=-29.652792]
gt: 40863732 Two-state_solution, pem_pos=0,
final_scores=-1.0548786, final_local_scores=-1.1753745,
pem=-0.24512246, similarity=-29.652792]
[winner: 28729 Solution, pem_pos=0,
final_scores=-1.2596192, final_local_scores=-1.2444134,
pem=-0.20737508, similarity=-31.91833]

```
gt: 40863732 Two-state_solution, pem_pos=2,
final_scores=-1.2928123, final_local_scores=-1.4352808,
pem=-1.3515586, similarity=-31.391943 ]
1: 17633 Libya
    [pem_pos=0, final_scores=0.4347118,
final_local_scores=0.39906102, pem=-0.3054436,
similarity=17.65954]
2: 358 Algeria
    [pem_pos=0, final_scores=0.5701335,
final_local_scores=0.583311, pem=-0.33694586,
similarity=23.32924]
3: 7515849 Syria
    [pem_pos=0, final_scores=0.5176798,
final_local_scores=0.48492527, pem=-0.26162508,
similarity=19.983608]
4: 718749 South_Yemen
    [pem_pos=0, final_scores=0.4605559,
final_local_scores=0.4628253, pem=0.0, similarity=17.901342]
5: 49522 Anwar_Sadat
    [pem_pos=0, final_scores=0.58315104,
final_local_scores=0.59145343, pem=0.0, similarity=21.739862]
6: 8087628 Egypt
    [pem_pos=0, final_scores=0.56217414,
final_local_scores=0.5272735, pem=-0.27480945,
similarity=21.319061]
7: 9282173 Israel
    [pem_pos=1, final_scores=0.7055653,
final_local_scores=0.6791752, pem=-0.21279564,
similarity=25.51487]
8: 20767 Menachem_Begin
    [pem_pos=0, final_scores=0.5558135,
final_local_scores=0.59592354, pem=0.0, similarity=21.87326]
9: 181901 Camp_David_Accords
    [pem_pos=0, final_scores=0.2017902,
final_local_scores=0.18824917, pem=0.0, similarity=9.707443]
10: 38762388 Sahrawi_Arab_Democratic_Republic
    [pem_pos=0, final_scores=0.28518742,
final_local_scores=0.28048298, pem=-0.015920518,
similarity=12.546463]
11: 33220 Western_Sahara
    [pem_pos=0, final_scores=0.10877983,
final_local_scores=0.07172905, pem=-0.039486464,
similarity=6.4449844]
12: 3008136 Rejectionist_Front
    [pem_pos=0, final_scores=0.18946056,
final_local_scores=0.23714292, pem=0.0, similarity=11.166531]
13: 11104 Fatah
    [pem_pos=0, final_scores=0.34247988,
final_local_scores=0.32309118, pem=0.0, similarity=13.731399]
1: 8087628 Egypt
    [pem_pos=0, final_scores=0.43587327,
final_local_scores=0.389309, pem=-1.052751,
```


similarity=21.432493]
2: 159433 Arab_world
[pem_pos=14, final_scores=0.3531987,
final_local_scores=0.3084643, pem=-3.8052368,
similarity=33.988373]
3: 9282173 Israel
[pem_pos=1, final_scores=0.5437619,
final_local_scores=0.50242954, pem=-0.21279564,
similarity=20.240429]
4: 159433 Arab_world
[pem_pos=14, final_scores=0.23950537,
final_local_scores=0.18427157, pem=-3.8052368,
similarity=30.282213]
5: 9282173 Israel
[pem_pos=1, final_scores=0.575743,
final_local_scores=0.53736407, pem=-0.21279564,
similarity=21.282944]
6: 9282173 Israel
[pem_pos=1, final_scores=0.55854994,
final_local_scores=0.51858324, pem=-0.21279564,
similarity=20.722488]
7: 8087628 Egypt
[pem_pos=0, final_scores=0.51727694,
final_local_scores=0.47823012, pem=-0.27480945,
similarity=19.85551]
8: 159433 Arab_world
[pem_pos=14, final_scores=0.46318358,
final_local_scores=0.42860603, pem=-3.8052368,
similarity=37.573643]
9: 7515849 Syria
[pem_pos=0, final_scores=0.72492963,
final_local_scores=0.7113142, pem=-0.26162508,
similarity=26.739506]
10: 241405 State_of_Palestine
[pem_pos=5, final_scores=0.42612588,
final_local_scores=0.39827955, pem=-1.7909594,
similarity=25.714685]
11: 9282173 Israel
[pem_pos=0, final_scores=0.56735206,
final_local_scores=0.5281982, pem=-0.21279564,
similarity=21.009415]
12: 1114732 Palestine
[pem_pos=1, final_scores=0.13403724,
final_local_scores=0.05872631, pem=-1.9100745,
similarity=16.229506]
13: 241405 State_of_Palestine
[pem_pos=9, final_scores=0.2050368,
final_local_scores=0.15677333, pem=-2.8508017,
similarity=24.271246]
14: 241405 State_of_Palestine
[pem_pos=9, final_scores=0.4766724,
final_local_scores=0.45349395, pem=-2.8508017,

similarity=33.12598]
 15: 24324 Palestine_Liberation_Organization
 [pem_pos=0, final_scores=0.25834876,
 final_local_scores=0.20722234, pem=-0.6143663,
 similarity=13.61466]
 16: 8087628 Egypt
 [pem_pos=0, final_scores=0.59147674,
 final_local_scores=0.5592822, pem=-0.27480945,
 similarity=22.274265]
 17: 159433 Arab_world
 [pem_pos=14, final_scores=0.33028334,
 final_local_scores=0.28343272, pem=-3.8052368,
 similarity=33.24138]

Sample: b'Hockley_Valley_Provincial_Nature_Reserve'

Hockley Valley Provincial Nature Reserve is a 1 nature reser
 ve located on the [1 1 Niagara Escarpment 1] within the
 township of [1 2 Mono 1] , [2 2 Ontario 2] , [3 3 Canada 3] .
 It is managed by [4 4 Ontario Parks 4] . Approximately 80 %
 of the Cannings Falls 3 Area of Natural and Scientific Intere
 st is within the borders of the park ; land acquisitions are
 planned to encompass it entirely . The [5 5 Bruce Trail 5]
 links the reserve to other protected areas of the [2 Niagara
 Escarpment 2] . The 4 topography of the area was greatly
 influenced by processes , followed by 5 fluvial erosion . The
 reserve is characterized by a 6 till 7 moraine and other
 glacial features , such as 8 kame s and 9 kettles . A deeply
 cut tributary valley to the [6 6 Nottawasaga River 6] bisects
 the reserve through its centre . The [3 Niagara Escarpment 3]
 's 10 dolostone 11 caprock -LRB- the Amabel Formation -RRB-
 can be seen in the upper parts of the gorge ; rapids and
 waterfalls expose stratigraphic succession down to the .
 Within the reserve , 417 species of 12 vascular plant s ,
 many of which are rare , have been identified . This includes
 the regionally rare and the , as well as the nationally
 endangered 13 Butternut , population of which has dwindled
 due to the disease . Fauna includes the locally rare 14 North
 ern Long-eared Bat , the nationally vulnerable 15 Louisiana Wa
 terthrush and the 16 Northern Brook Lamprey , as well as the
 nationally threatened . Recreation in the park is several
 limited . Due to the park 's designation as a nature
 reserve , permitted activities are limited to 17 hiking , and

1: 174573 Niagara_Escarpment

2: 22218 Ontario

3: 5042916 Canada

4: 2154113 Ontario_Parks

5: 174562 Bruce_Trail

6: 588262 Nottawasaga_River

1: 113261 Nature_reserve

[winner: 21192765 Protected_areas_of_Turkmenistan,
 pem_pos=0, final_scores=-0.768828,
 final_local_scores=-0.70446384, pem=0.0,

```

similarity=-16.932894|
    gt: 113261 Nature_reserve, pem_pos=1,
final_scores=-1.0419887, final_local_scores=-1.1392384,
pem=-0.032286346, similarity=-29.731855 ]
2: 386688 Mono_Ontario
    [winner: 82024 Mono_County_California, pem_pos=2,
final_scores=0.103118524, final_local_scores=0.14645848,
pem=-1.9120501, similarity=18.858353|
    gt: 386688 Mono_Ontario, pem_pos=8,
final_scores=-0.13329715, final_local_scores=-0.13867393,
pem=-2.6731987, similarity=14.488676 ]
3: 12482336 Area_of_Natural_and_Scientific_Interest
    [winner: 12482336
Area_of_Natural_and_Scientific_Interest, pem_pos=0,
final_scores=-0.49399024, final_local_scores=-0.47941825,
pem=0.0, similarity=-10.217084|
    gt: 12482336 Area_of_Natural_and_Scientific_Interest,
pem_pos=0, final_scores=-0.49399024,
final_local_scores=-0.47941825, pem=0.0,
similarity=-10.217084 ]
4: 79239 Topography
    [winner: 70713 Topographic_map, pem_pos=0,
final_scores=-1.302272, final_local_scores=-1.3944395,
pem=0.0, similarity=-37.52314|
    gt: 79239 Topography, pem_pos=1,
final_scores=-1.4882677, final_local_scores=-1.6089222,
pem=-0.012113147, similarity=-43.857864 ]
5: 286837 Fluvial
    [winner: 286837 Fluvial, pem_pos=0,
final_scores=-1.3892233, final_local_scores=-1.4772823,
pem=-0.02423027, similarity=-39.863567|
    gt: 286837 Fluvial, pem_pos=0, final_scores=-1.3892233,
final_local_scores=-1.4772823, pem=-0.02423027,
similarity=-39.863567 ]
6: 44545 Till
    [winner: 44545 Till, pem_pos=0, final_scores=-0.9430156,
final_local_scores=-1.009798, pem=-0.09682539,
similarity=-25.518122|
    gt: 44545 Till, pem_pos=0, final_scores=-0.9430156,
final_local_scores=-1.009798, pem=-0.09682539,
similarity=-25.518122 ]
7: 44299 Moraine
    [winner: 44299 Moraine, pem_pos=1,
final_scores=-0.9171007, final_local_scores=-1.0031443,
pem=-0.0013080445, similarity=-25.838999|
    gt: 44299 Moraine, pem_pos=1, final_scores=-0.9171007,
final_local_scores=-1.0031443, pem=-0.0013080445,
similarity=-25.838999 ]
8: 631191 Kame
    [winner: 23135525 Zimmerman_Kame, pem_pos=1,
final_scores=-0.4175467, final_local_scores=-0.3589118,
pem=0.0, similarity=-6.620931|

```

gt: 631191 Kame, pem_pos=0, final_scores=-0.46266317,
 final_local_scores=-0.4838966, pem=0.0,
 similarity=-10.350727]
 9: 882162 Kettle_landform
 [winner: 953864 Kettle, pem_pos=0,
 final_scores=-0.88096493, final_local_scores=-0.84693813,
 pem=-0.55586904, similarity=-18.161709]
 gt: 882162 Kettle_landform, pem_pos=1,
 final_scores=-0.9567345, final_local_scores=-1.0566378,
 pem=-1.0807546, similarity=-21.565155]
 10: 1020491 Dolostone
 [winner: 1020491 Dolostone, pem_pos=0,
 final_scores=-0.8585791, final_local_scores=-0.8806711,
 pem=-0.13931872, similarity=-21.433632]
 gt: 1020491 Dolostone, pem_pos=0,
 final_scores=-0.8585791, final_local_scores=-0.8806711,
 pem=-0.13931872, similarity=-21.433632]
 11: 13382972 Caprock
 [winner: 4609108 Caprock_Texas, pem_pos=2,
 final_scores=-0.7919896, final_local_scores=-0.72835636,
 pem=-0.51134497, similarity=-14.86512]
 gt: 13382972 Caprock, pem_pos=1,
 final_scores=-0.9106393, final_local_scores=-0.9419622,
 pem=-0.46317345, similarity=-21.501509]
 12: 66966 Vascular_plant
 [winner: 3094833 Vascular_tissue, pem_pos=0,
 final_scores=-0.94644153, final_local_scores=-0.941072,
 pem=-0.7628168, similarity=-19.845436]
 gt: 66966 Vascular_plant, pem_pos=1,
 final_scores=-1.191411, final_local_scores=-1.2585115,
 pem=-0.97258854, similarity=-28.17769]
 [winner: 3296070 Trenton_Engine_Plant, pem_pos=8,
 final_scores=-1.2682623, final_local_scores=-1.254113,
 pem=-2.0794415, similarity=-22.027191]
 RECALL MISS: 66966 Vascular_plant]
 13: 1271076 Juglans_cinerea
 [winner: 1271076 Juglans_cinerea, pem_pos=1,
 final_scores=-0.13991266, final_local_scores=-0.10749613,
 pem=-0.43746784, similarity=3.2608323]
 gt: 1271076 Juglans_cinerea, pem_pos=1,
 final_scores=-0.13991266, final_local_scores=-0.10749613,
 pem=-0.43746784, similarity=3.2608323]
 14: 8429960 Northern_long-eared_myotis
 [winner: 8429960 Northern_long-eared_myotis, pem_pos=1,
 final_scores=-0.1255247, final_local_scores=-0.009977311,
 pem=-0.64185387, similarity=7.2824707]
 gt: 8429960 Northern_long-eared_myotis, pem_pos=1,
 final_scores=-0.1255247, final_local_scores=-0.009977311,
 pem=-0.64185387, similarity=7.2824707]
 15: 418961 Louisiana_Waterthrush
 [winner: 418961 Louisiana_Waterthrush, pem_pos=0,
 final_scores=-0.06821494, final_local_scores=0.014992461,

```

pem=0.0, similarity=4.537118|
    gt: 418961 Louisiana_Waterthrush, pem_pos=0,
final_scores=-0.06821494, final_local_scores=0.014992461,
pem=0.0, similarity=4.537118 ]
16: 11759880 Northern_Brook_Lamprey
    [winner: 11759880 Northern_Brook_Lamprey, pem_pos=0,
final_scores=-0.016611218, final_local_scores=0.117563665,
pem=0.0, similarity=7.5980473|
    gt: 11759880 Northern_Brook_Lamprey, pem_pos=0,
final_scores=-0.016611218, final_local_scores=0.117563665,
pem=0.0, similarity=7.5980473 ]
17: 147313 Hiking
    [winner: 147313 Hiking, pem_pos=0,
final_scores=-1.1429005, final_local_scores=-1.2607882,
pem=-0.022422442, similarity=-33.412785|
    gt: 147313 Hiking, pem_pos=0, final_scores=-1.1429005,
final_local_scores=-1.2607882, pem=-0.022422442,
similarity=-33.412785 ]
1: 174573 Niagara_Escarpment
    [pem_pos=0, final_scores=0.3340904,
final_local_scores=0.29977837, pem=0.0, similarity=13.035698]
2: 22218 Ontario
    [pem_pos=0, final_scores=0.5858932,
final_local_scores=0.6180207, pem=-0.10088341,
similarity=23.081305]
3: 5042916 Canada
    [pem_pos=0, final_scores=0.61177576,
final_local_scores=0.6769192, pem=-0.11662793,
similarity=24.924574]
4: 2154113 Ontario_Parks
    [pem_pos=0, final_scores=0.10505764,
final_local_scores=0.1208545, pem=0.0, similarity=7.696252]
5: 174562 Bruce_Trail
    [pem_pos=0, final_scores=0.21900715,
final_local_scores=0.22516298, pem=0.0, similarity=10.809026]
6: 588262 Nottawasaga_River
    [pem_pos=0, final_scores=0.15284641,
final_local_scores=0.22394624, pem=0.0, similarity=10.772716]
1: 82024 Mono_County,_California
    [pem_pos=2, final_scores=0.103118524,
final_local_scores=0.14645848, pem=-1.9120501,
similarity=18.858353]
2: 174573 Niagara_Escarpment
    [pem_pos=0, final_scores=0.3497581,
final_local_scores=0.31689292, pem=0.0, similarity=13.54643]
3: 174573 Niagara_Escarpment
    [pem_pos=0, final_scores=0.32050753,
final_local_scores=0.28494117, pem=0.0, similarity=12.592926]

```

Sample: b'Alfred_Conkling_Coxe'

, Sr [1 Alfred Conkling Coxe 1] , Sr. -LRB- May 20 ,
1847 -- April 15 , 1923 -RRB- was longtime a 1 federal judge

in [2 2 New York 2] . [3 Coxe 3] was born in 3 Auburn , [4 New York 4] . His legal career began with private practice in [1 1 Utica 1] from 1868 . In 1870 he entered the firm of Conkling , Holmes & Coxe , of [5 Utica 5] , composed of [2 2 Roscoe Conkling 2] , then , ex-Judge [6 Sidney T. Holmes 6] , and Mr. Coxe . He continued in private practice to 1882 . He also served as manager of a state hospital in [7 Utica 7] 1880 to 1882 . In 1882 , 4 President [3 3 Chester A. Arthur 3] nominated Coxe as judge of the [8 5 United States District Court 8] for the Northern District of New York . Coxe served as a trial-level judge of that court for twenty years , until 1902 , when President [4 4 Theodore Roosevelt 4] promoted him to an appellate position on the [9 6 United States Court of Appeals 9] for the Second Circuit , headquartered in . Coxe served on that court for 15 years , retiring from the bench in 1917 . He died in 1923 . Coxe was the grandson of [5 5 Alfred Conkling 5] , who served as a [10 7 U.S. 10] Representative from upstate [11 New York 11] and a judge in the Northern District , and nephew of [6 6 Roscoe Conkling 6] , who was a [12 Congressman 12] and Senator from [13 New York 13] and boss of the state 's [14 Republican 14] political machine . He was also the nephew of [7 7 Arthur Cleveland Coxe 7] , the [15 Episcopal 15] bishop of Western New York , and grandson of abolitionist minister 8 Samuel Hanson Cox . [16 Coxe 16] 's son , [17 9 Alfred Conkling Coxe 17] , Jr. . , also became a federal judge , serving on the [18 10 United States District Court 18] for the [19 Southern District of New York 19] from 1929 to 1957 . Another son , Howard Coxe , was a newspaperman and novelist , and his grandson [8 8 Louis O. Coxe 8] was a poet and playwright best known for writing the [20 Broadway 20] version of 11 Billy Budd

1: 85435 Utica, New York

2: 988494 Roscoe_Conkling

3: 21490963 Chester_A._Arthur

4: 30535 Theodore_Roosevelt

5: 21807388 Alfred_Conkling

6: 988494 Roscoe_Conkling

7: 2612230 Arthur_Cleveland_Coxe

8: 1682758 Louis_O._Coxe

1: 834368 United_States_federal_judge

[winner: 834368 United_States_federal_judge, pem_pos=0, final_scores=-0.93464845, final_local_scores=-1.0562316, pem=-0.3375109, similarity=-25.59491]

gt: 834368 United_States_federal_judge, pem_pos=0, final_scores=-0.93464845, final_local_scores=-1.0562316, pem=-0.3375109, similarity=-25.59491]

2: 8210131 New_York

[winner: 645042 New_York_City, pem_pos=2, final_scores=0.41752762, final_local_scores=0.39016902, pem=-1.7906687, similarity=25.47107]

gt: 8210131 New_York, pem_pos=1, final_scores=-0.059576422, final_local_scores=-0.13783191,

pem=-0.2889045, similarity=1.5476414]
3: 126182 Auburn, New_York
[winner: 21131695 Robert_F._Kennedy, pem_pos=0,
final_scores=0.30738813, final_local_scores=0.2879943,
pem=0.0, similarity=12.684037]
RECALL MISS: 126182 Auburn, New_York]
[winner: 151274 Auburn, Washington, pem_pos=3,
final_scores=-0.18696916, final_local_scores=-0.16808408,
pem=-2.5687284, similarity=13.042893]
gt: 126182 Auburn, New_York, pem_pos=10,
final_scores=-0.24036966, final_local_scores=-0.30832586,
pem=-3.1743135, similarity=12.151065]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.6300913, final_local_scores=-1.6231079,
pem=-4.6728287, similarity=-18.935497]
RECALL MISS: 126182 Auburn, New_York]
4: 24113 President_of_the_United_States
[winner: 30535 Theodore_Roosevelt, pem_pos=0,
final_scores=-1.3259375, final_local_scores=-1.5125062,
pem=0.0, similarity=-41.04649]
gt: 24113 President_of_the_United_States, pem_pos=1,
final_scores=-1.5207441, final_local_scores=-1.7154312,
pem=-0.99314094, similarity=-41.70132]
5: 882426
United_States_District_Court_for_the_Northern_District_of_New_York
[winner: 70235 United_States_district_court, pem_pos=0,
final_scores=0.32325006, final_local_scores=0.32076585,
pem=0.0, similarity=13.662005]
RECALL MISS: 882426
United_States_District_Court_for_the_Northern_District_of_New_York]
[winner: 882426
United_States_District_Court_for_the_Northern_District_of_New_York, pem_pos=0, final_scores=-0.14701155,
final_local_scores=-0.17073956, pem=0.0,
similarity=-1.0054963]
gt: 882426
United_States_District_Court_for_the_Northern_District_of_New_York, pem_pos=0, final_scores=-0.14701155,
final_local_scores=-0.17073956, pem=0.0,
similarity=-1.0054963]
[winner: 3349493 Male-
female_income_disparity_in_the_United_States, pem_pos=0,
final_scores=-1.5274539, final_local_scores=-1.5871611,
pem=-0.12731102, similarity=-42.582]
RECALL MISS: 882426
United_States_District_Court_for_the_Northern_District_of_New_York]
6: 428320
United_States_Court_of_Appeals_for_the_Second_Circuit
[winner: 70228 United_States_courts_of_appeals,

pem_pos=1, final_scores=0.2589715,
final_local_scores=0.27899817, pem=-0.74575967,
similarity=16.471132|
RECALL MISS: 428320
United_States_Court_of_Appeals_for_the_Second_Circuit]
[winner: 428320
United_States_Court_of_Appeals_for_the_Second_Circuit,
pem_pos=0, final_scores=0.005254343,
final_local_scores=-0.012753844, pem=0.0,
similarity=3.709113|
gt: 428320
United_States_Court_of_Appeals_for_the_Second_Circuit,
pem_pos=0, final_scores=0.005254343,
final_local_scores=-0.012753844, pem=0.0,
similarity=3.709113]
[winner: 3349493 Male-
female_income_disparity_in_the_United_States, pem_pos=0,
final_scores=-1.6722462, final_local_scores=-1.7453246,
pem=-0.12731102, similarity=-47.301914|
RECALL MISS: 428320
United_States_Court_of_Appeals_for_the_Second_Circuit]
7: 19468510 United_States_House_of_Representatives
[winner: 3434750 United_States, pem_pos=0,
final_scores=0.44492888, final_local_scores=0.45909792,
pem=-0.14075686, similarity=18.555569|
gt: 19468510 United_States_House_of_Representatives,
pem_pos=15, final_scores=-0.55723536,
final_local_scores=-0.6607312, pem=-4.1090565,
similarity=6.7178645]
[winner: 19468510 United_States_House_of_Representatives,
pem_pos=0, final_scores=-0.9605747,
final_local_scores=-1.101866, pem=-0.47166955,
similarity=-26.227154|
gt: 19468510 United_States_House_of_Representatives,
pem_pos=0, final_scores=-0.9605747,
final_local_scores=-1.101866, pem=-0.47166955,
similarity=-26.227154]
8: 2612205 Samuel_Hanson_Cox
[winner: 2612205 Samuel_Hanson_Cox, pem_pos=0,
final_scores=0.018508598, final_local_scores=0.083003014,
pem=0.0, similarity=6.5666885|
gt: 2612205 Samuel_Hanson_Cox, pem_pos=0,
final_scores=0.018508598, final_local_scores=0.083003014,
pem=0.0, similarity=6.5666885]
9: 21807720 Alfred_Conkling_Coxe,_Jr.
[winner: 9782117 Alfred_Conkling_Coxe,_Sr., pem_pos=0,
final_scores=0.23003705, final_local_scores=0.2378774,
pem=0.0, similarity=11.188449|
RECALL MISS: 21807720 Alfred_Conkling_Coxe,_Jr.]
[winner: 22829226 Junior_(education), pem_pos=0,
final_scores=-0.46562546, final_local_scores=-0.43395406,
pem=-0.32146513, similarity=-7.1121645|

RECALL MISS: 21807720 Alfred_Conkling_Coxe,_Jr.]

[winner: 4721653 Zyphe_language, pem_pos=12,
final_scores=-1.5274593, final_local_scores=-1.4762473,
pem=-4.4188404, similarity=-15.934108]

RECALL MISS: 21807720 Alfred_Conkling_Coxe,_Jr.]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-2.0505824, final_local_scores=-2.0824306,
pem=-4.6728287, similarity=-32.642597]

RECALL MISS: 21807720 Alfred_Conkling_Coxe,_Jr.]

10: 747093

United_States_District_Court_for_the_Southern_District_of_New_York

[winner: 70235 United_States_district_court, pem_pos=0,
final_scores=0.3640831, final_local_scores=0.3682912,
pem=0.0, similarity=15.080257]

gt: 747093

United_States_District_Court_for_the_Southern_District_of_New_York, pem_pos=2, final_scores=-0.70222217,
final_local_scores=-0.7901493, pem=-5.6048827,
similarity=10.990307]

[winner: 747093

United_States_District_Court_for_the_Southern_District_of_New_York, pem_pos=0, final_scores=0.08806713,
final_local_scores=0.07206936, pem=0.0, similarity=6.2404065]

gt: 747093

United_States_District_Court_for_the_Southern_District_of_New_York, pem_pos=0, final_scores=0.08806713,
final_local_scores=0.07206936, pem=0.0,
similarity=6.2404065]

[winner: 3349493 Male-

female_income_disparity_in_the_United_States, pem_pos=0,
final_scores=-1.4104875, final_local_scores=-1.459393,
pem=-0.12731102, similarity=-38.769146]

RECALL MISS: 747093

United_States_District_Court_for_the_Southern_District_of_New_York]

11: 438653 Billy_Budd

[winner: 438653 Billy_Budd, pem_pos=0,
final_scores=0.006581679, final_local_scores=0.071246505,
pem=-0.056298614, similarity=6.5220113]

gt: 438653 Billy_Budd, pem_pos=0,

final_scores=0.006581679, final_local_scores=0.071246505,
pem=-0.056298614, similarity=6.5220113]

1: 85435 Utica,_New_York

[pem_pos=0, final_scores=0.29033816,
final_local_scores=0.26125956, pem=-0.821632,
similarity=16.354382]

2: 988494 Roscoe_Conkling

[pem_pos=0, final_scores=0.11290397,
final_local_scores=0.104701936, pem=0.0, similarity=7.2142277]

3: 21490963 Chester_A._Arthur

[pem_pos=0, final_scores=0.45861334,

final_local_scores=0.44878638, pem=0.0, similarity=17.482393]
4: 30535 Theodore_Roosevelt
[pem_pos=0, final_scores=0.33388376,
final_local_scores=0.30440295, pem=0.0, similarity=13.173704]
5: 21807388 Alfred_Conkling
[pem_pos=1, final_scores=0.33411264,
final_local_scores=0.36152917, pem=0.0, similarity=14.878465]
6: 988494 Roscoe_Conkling
[pem_pos=0, final_scores=0.30594832,
final_local_scores=0.31557354, pem=0.0, similarity=13.507057]
7: 2612230 Arthur_Cleveland_Coxe
[pem_pos=0, final_scores=0.08306949,
final_local_scores=0.13446733, pem=0.0, similarity=8.102487]
8: 1682758 Louis_O._Coxe
[pem_pos=0, final_scores=0.17148106,
final_local_scores=0.21216601, pem=0.0, similarity=10.42117]
1: 9782117 Alfred_Conkling_Coxe,_Sr.
[pem_pos=0, final_scores=0.15870292,
final_local_scores=0.15995571, pem=0.0, similarity=8.863111]
2: 645042 New_York_City
[pem_pos=2, final_scores=0.41752762,
final_local_scores=0.39016902, pem=-1.7906687,
similarity=25.47107]
3: 1682758 Louis_O._Coxe
[pem_pos=2, final_scores=0.060973898,
final_local_scores=0.09145373, pem=0.0, similarity=6.818875]
4: 21131695 Robert_F._Kennedy
[pem_pos=0, final_scores=0.30738813,
final_local_scores=0.2879943, pem=0.0, similarity=12.684037]
5: 85435 Utica,_New_York
[pem_pos=0, final_scores=0.3567286,
final_local_scores=0.333781, pem=-0.821632,
similarity=18.518566]
6: 11594003 Sidney_T._Holmes
[pem_pos=0, final_scores=0.051636204,
final_local_scores=0.12873065, pem=0.0, similarity=7.9312925]
7: 85435 Utica,_New_York
[pem_pos=0, final_scores=0.20730256,
final_local_scores=0.1705558, pem=-0.821632,
similarity=13.647601]
8: 70235 United_States_district_court
[pem_pos=0, final_scores=0.32325006,
final_local_scores=0.32076585, pem=0.0, similarity=13.662005]
9: 70228 United_States_courts_of_appeals
[pem_pos=1, final_scores=0.2589715,
final_local_scores=0.27899817, pem=-0.74575967,
similarity=16.471132]
10: 3434750 United_States
[pem_pos=0, final_scores=0.44492888,
final_local_scores=0.45909792, pem=-0.14075686,
similarity=18.555569]
11: 21131695 Robert_F._Kennedy

[pem_pos=0, final_scores=0.25834042,
 final_local_scores=0.23441711, pem=0.0, similarity=11.085187]
 12: 19468510 United_States_House_of_Representatives
 [pem_pos=1, final_scores=0.05182509,
 final_local_scores=0.008443534, pem=-0.7877172,
 similarity=8.625414]
 13: 645042 New_York_City
 [pem_pos=1, final_scores=0.39562428,
 final_local_scores=0.36605442, pem=-1.7906687,
 similarity=24.751442]
 14: 32070 Republican_Party_(United_States)
 [pem_pos=0, final_scores=0.41547078,
 final_local_scores=0.4098193, pem=-0.1310191,
 similarity=17.032042]
 15: 19280748 Episcopal_Church_(United_States)
 [pem_pos=0, final_scores=0.3296005,
 final_local_scores=0.31882825, pem=-0.049595453,
 similarity=13.873892]
 16: 9782117 Alfred_Conkling_Coxe,_Sr.
 [pem_pos=0, final_scores=0.034043834,
 final_local_scores=0.02255787, pem=0.0, similarity=4.7628846]
 17: 9782117 Alfred_Conkling_Coxe,_Sr.
 [pem_pos=0, final_scores=0.23003705,
 final_local_scores=0.2378774, pem=0.0, similarity=11.188449]
 18: 70235 United_States_district_court
 [pem_pos=0, final_scores=0.3640831,
 final_local_scores=0.3682912, pem=0.0, similarity=15.080257]
 19: 747093
 United_States_District_Court_for_the_Southern_District_of_New
 York
 [pem_pos=0, final_scores=0.08806713,
 final_local_scores=0.07206936, pem=0.0, similarity=6.2404065]
 20: 725252 Broadway_theatre
 [pem_pos=0, final_scores=0.23652564,
 final_local_scores=0.24841785, pem=-0.28891575,
 similarity=13.074166]

Sample: b'Andy_Flynn'

-LRB- footballer -RRB- Andrew `` Andy '' Flynn -LRB- 1895
 -- ? -RRB- was an [1 English 1] 1 footballer who played as
 a . He played for , [1 1 Exeter City 1] , [2 2 York City 2]
 and . Born in [3 3 Sheffield 3] , 2 West Riding of [2 Yorkshi
 re 2] , Flynn played for of the [4 4 Midland League 4] before
 joining 3 Third Division South 4 Exeter City in 1922 . He
 made his league debut for Exeter against [5 5 Southend United
 5] on 9 December , and scored his first and only goal for the
 club with the winner in a 1 -- 0 victory away at [6 6 Swindon
 Town 6] 16 February 1924 . However , [3 Flynn 3] failed to
 command a regular place in the Exeter team during his four-
 season stay , finishing his time at the club with 34 league
 appearances and one goal . [4 Flynn 4] returned to playing [5
 Midland League 5] football after signing for [7 7 York City 7]

in August 1926 , making his debut on 28 August in a 1 - - 1 draw at home to 5 Loughborough Corinthians . He was ever-present for York in the 6 1926 - - 27 season , appearing in all of York 's 38 [6 Midland League 6] matches . He also played in York 's seven [8 8 FA Cup 8] games that season , in a run that saw the club reach the 7 first round for the first time in its history , eventually being knocked out in the second round away at [9 9 Grimsby Town 9] on 11 December . [7 Flynn 7] missed two games the following season , against [10 10 Scarborough 10] and [8 Mexborough Athletic 8] , in which he made 48 appearances in all competitions . He left York following the end of the season , having made 93 appearances in all competitions , joining as a player and a trainer .

Flynn played as a

- 1: 347780 Exeter_City_F.C.
- 2: 451170 York_City_F.C.
- 3: 88885 Sheffield
- 4: 3761630 Midland_Football_League
- 5: 451168 Southend_United_F.C.
- 6: 74291 Swindon_Town_F.C.
- 7: 451170 York_City_F.C.
- 8: 11237 FA_Cup
- 9: 451129 Grimsby_Town_F.C.
- 10: 946449 Scarborough_F.C.

1: 10568 Association_football

[winner: 10568 Association_football, pem_pos=0, final_scores=-1.4763049, final_local_scores=-1.6011019, pem=0.0, similarity=-43.69036]

gt: 10568 Association_football, pem_pos=0, final_scores=-1.4763049, final_local_scores=-1.6011019, pem=0.0, similarity=-43.69036]

2: 208741 West_Riding_of_Yorkshire

[winner: 36637 Yorkshire, pem_pos=0, final_scores=0.09554164, final_local_scores=0.091777876, pem=-0.31620604, similarity=8.548125]

gt: 208741 West_Riding_of_Yorkshire, pem_pos=19, final_scores=-0.43320453, final_local_scores=-0.4739854, pem=-3.4173226, similarity=8.528974]

[winner: 208741 West_Riding_of_Yorkshire, pem_pos=0, final_scores=-0.060668543, final_local_scores=-0.07490629, pem=0.0, similarity=1.8543599]

gt: 208741 West_Riding_of_Yorkshire, pem_pos=0, final_scores=-0.060668543, final_local_scores=-0.07490629, pem=0.0, similarity=1.8543599]

3: 580907 Football_League_Third_Division_South

[winner: 580907 Football_League_Third_Division_South, pem_pos=0, final_scores=-0.7003674, final_local_scores=-0.84576184, pem=0.0, similarity=-21.149508]

gt: 580907 Football_League_Third_Division_South, pem_pos=0, final_scores=-0.7003674, final_local_scores=-0.84576184, pem=0.0,

```

similarity=-21.149508 ]
4: 347780 Exeter_City_F.C.
  [winner: 347780 Exeter_City_F.C., pem_pos=0,
  final_scores=-0.87889093, final_local_scores=-1.0603731,
  pem=0.0, similarity=-27.553934]
  gt: 347780 Exeter_City_F.C., pem_pos=0,
  final_scores=-0.87889093, final_local_scores=-1.0603731,
  pem=0.0, similarity=-27.553934 ]
5: 34924063 Loughborough_Corinthians_F.C.
  [winner: 34924063 Loughborough_Corinthians_F.C.,
  pem_pos=0, final_scores=-0.043277413,
  final_local_scores=-0.011948541, pem=0.0,
  similarity=3.7331448]
  gt: 34924063 Loughborough_Corinthians_F.C., pem_pos=0,
  final_scores=-0.043277413, final_local_scores=-0.011948541,
  pem=0.0, similarity=3.7331448 ]
6: 5707944 1926-27_in_English_football
  [winner: 34609 1926, pem_pos=0, final_scores=-0.8200462,
  final_local_scores=-0.6895535, pem=-0.76264197,
  similarity=-12.340573]
  RECALL MISS: 5707944 1926-27_in_English_football]
  [winner: 19624 May_27, pem_pos=2,
  final_scores=-0.9912644, final_local_scores=-0.8120284,
  pem=-3.0017982, similarity=-3.8185892]
  RECALL MISS: 5707944 1926-27_in_English_football]
  [winner: 38340996 Methylene_bridge, pem_pos=0,
  final_scores=-1.2343111, final_local_scores=-1.1208216,
  pem=0.0, similarity=-29.35784]
  RECALL MISS: 5707944 1926-27_in_English_football]
7: 19999550 1926-27_FA_Cup
  [winner: 7730957 1993_NBA_Playoffs, pem_pos=0,
  final_scores=-1.1010796, final_local_scores=-1.0930812,
  pem=0.0, similarity=-28.53001]
  RECALL MISS: 19999550 1926-27_FA_Cup]
1: 347780 Exeter_City_F.C.
  [pem_pos=0, final_scores=0.41228276,
  final_local_scores=0.35311687, pem=0.0, similarity=14.627424]
2: 451170 York_City_F.C.
  [pem_pos=0, final_scores=0.35685295,
  final_local_scores=0.28528216, pem=0.0, similarity=12.603102]
3: 88885 Sheffield
  [pem_pos=0, final_scores=0.37673455,
  final_local_scores=0.39015985, pem=-0.25810194,
  similarity=17.13646]
4: 3761630 Midland_Football_League
  [pem_pos=0, final_scores=0.2875548,
  final_local_scores=0.22724456, pem=0.0, similarity=10.871143]
5: 451168 Southend_United_F.C.
  [pem_pos=0, final_scores=0.4690501,
  final_local_scores=0.4165309, pem=0.0, similarity=16.519827]
6: 74291 Swindon_Town_F.C.
  [pem_pos=0, final_scores=0.3948186,

```

```

final_local_scores=0.33538544, pem=0.0, similarity=14.098284]
7: 451170 York_City_F.C.
  [pem_pos=0, final_scores=0.35872513,
final_local_scores=0.28732732, pem=0.0, similarity=12.664133]
8: 11237 FA_Cup
  [pem_pos=0, final_scores=0.44098252,
final_local_scores=0.39837825, pem=-0.16323638,
similarity=16.86582]
9: 451129 Grimsby_Town_F.C.
  [pem_pos=0, final_scores=0.37711692,
final_local_scores=0.30990532, pem=0.0, similarity=13.337906]
10: 946449 Scarborough_F.C.
  [pem_pos=2, final_scores=0.73647946,
final_local_scores=0.7143641, pem=-1.8160111,
similarity=35.283512]
1: 9316 England
  [pem_pos=1, final_scores=0.5454505,
final_local_scores=0.59539205, pem=-1.3527626,
similarity=29.213932]
2: 36637 Yorkshire
  [pem_pos=0, final_scores=0.09554164,
final_local_scores=0.091777876, pem=-0.31620604,
similarity=8.548125]
3: 2393334 Brian_Flynn
  [pem_pos=11, final_scores=0.11305483,
final_local_scores=0.066515744, pem=-3.4775658,
similarity=24.986217]
4: 2393334 Brian_Flynn
  [pem_pos=11, final_scores=0.18209971,
final_local_scores=0.14193678, pem=-3.4775658,
similarity=27.236933]
5: 3761630 Midland_Football_League
  [pem_pos=0, final_scores=0.22686706,
final_local_scores=0.16095245, pem=0.0, similarity=8.892855]
6: 3761630 Midland_Football_League
  [pem_pos=0, final_scores=0.28420407,
final_local_scores=0.22358441, pem=0.0, similarity=10.761918]
7: 4814657 Michael_Flynn_(footballer)
  [pem_pos=13, final_scores=0.1509716,
final_local_scores=0.12282914, pem=-3.660579,
similarity=27.661978]
8: 11065643 Mexborough_Town_A.F.C.
  [pem_pos=0, final_scores=0.16931541,
final_local_scores=0.1509704, pem=0.0, similarity=8.594971]

```

Sample: b'Chetco_people'

The Chetco -LRB- 1 Chetco-Tolowa : chit-dee-ni , chit-dee-ne -RRB- are a tribe of [1 1 Native Americans 1] who originally lived along the lower [2 2 Chetco River 2] in [3 3 Curry County 3] in the [1 2 U.S. 1] state of [4 4 Oregon 4] . The name Chetco comes from the word meaning `` close to the mouth of the stream '' in their own language , which is part of the [2

3 Athapascan 2] languages . Although they were once one of the largest tribes on the Pacific coast of [3 Oregon 3] , `` the last known full-blooded Chetco '' died in 1940 , and as of 2009 only 40 or so descendants of the tribe remain . Some Chetco people are enrolled in 4 Cher-Ae Heights Indian Community of the Trinidad Rancheria , located in [4 5 Humboldt County 4] , [5 California 5] , and [5 5 Confederated Tribes of Siletz Indians 5] in [6 Oregon 6] . The 6 Chetco language is a member of the [7 7 Athapascan 7] languages , which also includes most native languages in [6 6 Alaska 6] , the and [8 8 Navajo 8] language s in the southwest [9 United States 9] , and the languages spoken by the and tribes in [10 Oregon 10] . The name Chetco comes from the word Cheti in their own language , meaning `` close to the mouth of the stream '' . The nine villages of the tribe on the [11 Chetco River 11] were named Chettanne , Chettannene -LRB- twin villages at the mouth of the river -RRB- , Khuniliikhwut , Nakwutthume , Nukhwuchutun , Setthattun , Siskhaslitun , Tachukhaslitun , and Thlcharghilitun . The Chetco are believed to have come to coastal [12 Oregon 12] between 3000 and 1000 years ago . They had nine villages on the lower of the [13 Chetco River 13] , with their principal villages at the mouth , where the river flows into the 9 Pacific Ocean . The Chetco territory extended a short distance on either side of the river , along the Pacific coast from in the north to the [7 7 Winchuck River 7] in the south . They were the most populous of the 12 coastal tribes in southern [14 Oregon 14] . The Chetco were 10 hunter-gatherer s with a diet based on hunting deer and elk , gathering 11 acorn s and 12 mussel s , and fishing . They used 13 dugout canoes on the ocean and river and worked with stone tools . The Chetco cooked on open fires or with simple pots , and were culturally very similar to the tribe to the south , `` who shared the same customs regulating social relationships and frequently intermarried '' . The tribe is thought to have had perhaps one thousand members at its peak , but its numbers declined after [15 European-American 15] settlers came into contact with the Chetco in the 19th century . Settlers destroyed the Chetco villages in 1853 and the surviving members of the tribe were forcibly removed to the [8 8 Siletz Reservation 8] in [16 14 Tillamook County 16] , [17 Oregon 17] -LRB- in 1879 the land the reservation is on became part of 15 Lincoln County

1: 21217 Native_Americans_in_the_United_States

2: 708307 Chetco_River

3: 92029 Curry_County,_Oregon

4: 26811621 Oregon

5: 4548581 Confederated_Tribes_of_Siletz_Indians

6: 624 Alaska

7: 31501230 Winchuck_River

8: 327207 Siletz_Reservation

1: 4033483 Tolowa_language

[winner: 4033483 Tolowa_language, pem_pos=0,

```

final_scores=-0.22588395, final_local_scores=-0.20939033,
pem=0.0, similarity=-2.1589122|
    gt: 4033483 Tolowa_language, pem_pos=0,
final_scores=-0.22588395, final_local_scores=-0.20939033,
pem=0.0, similarity=-2.1589122 ]
2: 18618239 U.S._state
    [winner: 3434750 United_States, pem_pos=0,
final_scores=0.44578898, final_local_scores=0.44534248,
pem=-0.14075686, similarity=18.145079|
    gt: 18618239 U.S._state, pem_pos=6,
final_scores=-0.3370306, final_local_scores=-0.43369997,
pem=-3.9799929, similarity=12.79106 ]
    [winner: 229379 States_of_Brazil, pem_pos=0,
final_scores=-1.5474039, final_local_scores=-1.5775836,
pem=0.0, similarity=-42.98853|
    RECALL MISS: 18618239 U.S._state]
3: 337846 Athabaskan_languages
    [winner: 337846 Athabaskan_languages, pem_pos=0,
final_scores=0.109429225, final_local_scores=0.1120618,
pem=0.0, similarity=7.4338603|
    gt: 337846 Athabaskan_languages, pem_pos=0,
final_scores=0.109429225, final_local_scores=0.1120618,
pem=0.0, similarity=7.4338603 ]
    [winner: 3552038 Languages_of_Egypt, pem_pos=1,
final_scores=-1.6226313, final_local_scores=-1.6556735,
pem=-1.3716635, similarity=-37.85957|
    RECALL MISS: 337846 Athabaskan_languages]
4: 36250480 Cher-
Ae_Heights_Indian_Community_of_the_Trinidad_Rancheria
    [winner: 36250480 Cher-
Ae_Heights_Indian_Community_of_the_Trinidad_Rancheria,
pem_pos=0, final_scores=-0.004483953,
final_local_scores=0.038819015, pem=0.0, similarity=5.24815|
    gt: 36250480 Cher-
Ae_Heights_Indian_Community_of_the_Trinidad_Rancheria,
pem_pos=0, final_scores=-0.004483953,
final_local_scores=0.038819015, pem=0.0, similarity=5.24815 ]
    [winner: 36250480 Cher-
Ae_Heights_Indian_Community_of_the_Trinidad_Rancheria,
pem_pos=0, final_scores=-0.024232373,
final_local_scores=0.017246857, pem=0.0,
similarity=4.6043935|
    gt: 36250480 Cher-
Ae_Heights_Indian_Community_of_the_Trinidad_Rancheria,
pem_pos=0, final_scores=-0.024232373,
final_local_scores=0.017246857, pem=0.0,
similarity=4.6043935 ]
    [winner: 10447732
Results_by_riding_of_the_Canadian_federal_election,_2008,
pem_pos=13, final_scores=-1.3265189,
final_local_scores=-1.2201114, pem=-3.6198866,
similarity=-12.635341|

```


RECALL MISS: 36250480 Cher-
 Ae_Heights_Indian_Community_of_the_Trinidad_Rancheria]
 5: 80544 Humboldt_County,_California
 [winner: 5407 California, pem_pos=0,
 final_scores=0.6316523, final_local_scores=0.6304811,
 pem=-0.08662026, similarity=23.375582]
 RECALL MISS: 80544 Humboldt_County,_California]
 [winner: 80544 Humboldt_County,_California, pem_pos=0,
 final_scores=0.36834955, final_local_scores=0.33969703,
 pem=0.0, similarity=14.22695]
 gt: 80544 Humboldt_County,_California, pem_pos=0,
 final_scores=0.36834955, final_local_scores=0.33969703,
 pem=0.0, similarity=14.22695]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.6960824, final_local_scores=-1.6465931,
 pem=-4.6728287, similarity=-19.63634]
 RECALL MISS: 80544 Humboldt_County,_California]
 6: 4033483 Tolowa_language
 [winner: 708307 Chetco_River, pem_pos=1,
 final_scores=-0.01760386, final_local_scores=-0.048167527,
 pem=-0.3173578, similarity=4.378139]
 RECALL MISS: 4033483 Tolowa_language]
 [winner: 18910 Markup_language, pem_pos=4,
 final_scores=-2.0967574, final_local_scores=-2.1471117,
 pem=-3.3860023, similarity=-41.57077]
 RECALL MISS: 4033483 Tolowa_language]
 7: 337846 Athabaskan_languages
 [winner: 337846 Athabaskan_languages, pem_pos=0,
 final_scores=0.06236495, final_local_scores=0.060651243,
 pem=0.0, similarity=5.899667]
 gt: 337846 Athabaskan_languages, pem_pos=0,
 final_scores=0.06236495, final_local_scores=0.060651243,
 pem=0.0, similarity=5.899667]
 [winner: 3552038 Languages_of_Egypt, pem_pos=1,
 final_scores=-1.6629671, final_local_scores=-1.6997342,
 pem=-1.3716635, similarity=-39.17443]
 RECALL MISS: 337846 Athabaskan_languages]
 8: 176526 Navajo_language
 [winner: 3007285 Navajo_people, pem_pos=0,
 final_scores=0.11310245, final_local_scores=0.101061255,
 pem=-1.0083177, similarity=12.588971]
 gt: 176526 Navajo_language, pem_pos=1,
 final_scores=-0.09430045, final_local_scores=-0.08751139,
 pem=-1.1293374, similarity=7.6197114]
 [winner: 18910 Markup_language, pem_pos=4,
 final_scores=-2.0225286, final_local_scores=-2.0660279,
 pem=-3.3860023, similarity=-39.151073]
 RECALL MISS: 176526 Navajo_language]
 9: 23070 Pacific_Ocean
 [winner: 3211921 Spanish_East_Indies, pem_pos=2,
 final_scores=-0.13516347, final_local_scores=-0.08058496,
 pem=-1.0973934, similarity=7.6526933]

gt: 23070 Pacific_Ocean, pem_pos=0,
 final_scores=-0.13963205, final_local_scores=-0.17725523,
 pem=-0.007366304, similarity=-1.1598775]
 10: 210098 Hunter-gatherer
 [winner: 210098 Hunter-gatherer, pem_pos=0,
 final_scores=-1.2121181, final_local_scores=-1.3018384,
 pem=0.0, similarity=-34.75974]
 gt: 210098 Hunter-gatherer, pem_pos=0,
 final_scores=-1.2121181, final_local_scores=-1.3018384,
 pem=0.0, similarity=-34.75974]
 11: 922939 Acorn
 [winner: 21168485 Acorn, _Virginia, pem_pos=6,
 final_scores=-0.8820045, final_local_scores=-0.82961017,
 pem=-2.3745632, similarity=-7.7542663]
 gt: 922939 Acorn, pem_pos=3, final_scores=-1.3415428,
 final_local_scores=-1.4188416, pem=-1.8217509,
 similarity=-28.344383]
 12: 144156 Mussel
 [winner: 2572321 Perna_canalicula, pem_pos=0,
 final_scores=-0.80670863, final_local_scores=-0.78705895,
 pem=0.0, similarity=-19.397696]
 gt: 144156 Mussel, pem_pos=1, final_scores=-1.047349,
 final_local_scores=-1.0900695, pem=-0.006779713,
 similarity=-28.403267]
 13: 490476 Dugout_(boat)
 [winner: 992676 Dugout_(shelter), pem_pos=1,
 final_scores=-1.1561732, final_local_scores=-1.2341399,
 pem=-0.63115144, similarity=-29.307186]
 gt: 490476 Dugout_(boat), pem_pos=2,
 final_scores=-1.3379192, final_local_scores=-1.4205062,
 pem=-0.643963, similarity=-34.799057]
 [winner: 490476 Dugout_(boat), pem_pos=0,
 final_scores=-1.2960379, final_local_scores=-1.3747573,
 pem=0.0, similarity=-36.935787]
 gt: 490476 Dugout_(boat), pem_pos=0,
 final_scores=-1.2960379, final_local_scores=-1.3747573,
 pem=0.0, similarity=-36.935787]
 14: 91996 Tillamook_County,_Oregon
 [winner: 26811621 Oregon, pem_pos=0,
 final_scores=0.476281, final_local_scores=0.42902732,
 pem=-0.28686622, similarity=18.452766]
 RECALL MISS: 91996 Tillamook_County,_Oregon]
 [winner: 91996 Tillamook_County,_Oregon, pem_pos=0,
 final_scores=0.2642703, final_local_scores=0.2638869,
 pem=0.0, similarity=11.964624]
 gt: 91996 Tillamook_County,_Oregon, pem_pos=0,
 final_scores=0.2642703, final_local_scores=0.2638869,
 pem=0.0, similarity=11.964624]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.8217435, final_local_scores=-1.7838588,
 pem=-4.6728287, similarity=-23.732622]
 RECALL MISS: 91996 Tillamook_County,_Oregon]

15: 92004 Lincoln_County,_Oregon
[winner: 93600 Lincoln_County,_Nevada, pem_pos=0,
final_scores=-0.020640776, final_local_scores=-0.041590303,
pem=-1.6845158, similarity=12.009234]
gt: 92004 Lincoln_County,_Oregon, pem_pos=1,
final_scores=-0.17775361, final_local_scores=-0.23405188,
pem=-2.2334604, similarity=9.251043]

1: 21217 Native_Americans_in_the_United_States
[pem_pos=0, final_scores=0.2295387,
final_local_scores=0.19364223, pem=-0.29015967,
similarity=11.446317]

2: 708307 Chetco_River
[pem_pos=0, final_scores=0.28034675,
final_local_scores=0.2843674, pem=0.0, similarity=12.575804]

3: 92029 Curry_County,_Oregon
[pem_pos=0, final_scores=0.3468185,
final_local_scores=0.33813545, pem=-0.009064326,
similarity=14.229642]

4: 26811621 Oregon
[pem_pos=0, final_scores=0.47090983,
final_local_scores=0.42316014, pem=-0.28686622,
similarity=18.277678]

5: 4548581 Confederated_Tribes_of_Siletz_Indians
[pem_pos=0, final_scores=0.38407725,
final_local_scores=0.41865718, pem=0.0, similarity=16.583279]

6: 624 Alaska
[pem_pos=0, final_scores=0.639796,
final_local_scores=0.6311699, pem=-0.18647984,
similarity=23.939188]

7: 31501230 Winchuck_River
[pem_pos=0, final_scores=0.14707665,
final_local_scores=0.21775389, pem=0.0, similarity=10.587923]

8: 327207 Siletz_Reservation
[pem_pos=0, final_scores=0.21100084,
final_local_scores=0.24257162, pem=0.0, similarity=11.328534]

1: 3434750 United_States
[pem_pos=0, final_scores=0.44578898,
final_local_scores=0.44534248, pem=-0.14075686,
similarity=18.145079]

2: 337846 Athabaskan_languages
[pem_pos=0, final_scores=0.109429225,
final_local_scores=0.1120618, pem=0.0, similarity=7.4338603]

3: 26811621 Oregon
[pem_pos=0, final_scores=0.421359,
final_local_scores=0.3690334, pem=-0.28686622,
similarity=16.662428]

4: 80544 Humboldt_County,_California
[pem_pos=0, final_scores=0.36834955,
final_local_scores=0.33969703, pem=0.0, similarity=14.22695]

5: 5407 California
[pem_pos=0, final_scores=0.6316523,
final_local_scores=0.6304811, pem=-0.08662026,

similarity=23.375582]
 6: 26811621 Oregon
 [pem_pos=0, final_scores=0.6348025,
 final_local_scores=0.602188, pem=-0.28686622,
 similarity=23.620226]
 7: 337846 Athabaskan_languages
 [pem_pos=0, final_scores=0.06236495,
 final_local_scores=0.060651243, pem=0.0, similarity=5.899667]
 8: 3007285 Navajo_people
 [pem_pos=0, final_scores=0.11310245,
 final_local_scores=0.101061255, pem=-1.0083177,
 similarity=12.588971]
 9: 3434750 United_States
 [pem_pos=0, final_scores=0.5002301,
 final_local_scores=0.5048111, pem=-0.073457584,
 similarity=19.553757]
 10: 26811621 Oregon
 [pem_pos=0, final_scores=0.58201736,
 final_local_scores=0.54452825, pem=-0.28686622,
 similarity=21.899544]
 11: 708307 Chetco_River
 [pem_pos=0, final_scores=0.15073298,
 final_local_scores=0.14278406, pem=0.0, similarity=8.350674]
 12: 26811621 Oregon
 [pem_pos=0, final_scores=0.5161888,
 final_local_scores=0.47262055, pem=-0.28686622,
 similarity=19.753675]
 13: 708307 Chetco_River
 [pem_pos=0, final_scores=0.20675014,
 final_local_scores=0.20397434, pem=0.0, similarity=10.176714]
 14: 26811621 Oregon
 [pem_pos=0, final_scores=0.5764662,
 final_local_scores=0.5384644, pem=-0.28686622,
 similarity=21.718588]
 15: 3434750 United_States
 [pem_pos=2, final_scores=0.13578711,
 final_local_scores=0.107303664, pem=-1.0986123,
 similarity=13.266293]
 16: 91996 Tillamook_County,_Oregon
 [pem_pos=0, final_scores=0.2642703,
 final_local_scores=0.2638869, pem=0.0, similarity=11.964624]
 17: 26811621 Oregon
 [pem_pos=0, final_scores=0.476281,
 final_local_scores=0.42902732, pem=-0.28686622,
 similarity=18.452766]

Sample: b'Claudiopolis'

_ -LRB- [1 Cilicia 1] -RRB- **Claudiopolis** -LRB- [2 1 **Greek** 2]
 : -RRB- also called **Ninica** and **Ninica** **Claudiopolis** , was an
 ancient city of [1 1 **Cilicia** 1] . mentions [2 2 **Silifke** 2]
 and **Claudiopolis** as cities of [3 **Cilicia** 3] , or of the
 country drained by the ; and **Claudiopolis** was a colony of [3 3

Claudius Caesar 3] . It is described by [4 2 Theophanes 4] of [5 Byzantium 5] as situated in a plain between the two [4 4 Taurus Mountains 4] , a description which exactly , corresponds to the position of the basin of the [6 Calycadnus 6] . Claudiopoli s may therefore be represented by , which is higher up the valley than [7 Seleucia 7] , and near the junction of the northern and western branches of the [8 Calycadnus 8] . It is also the place to which the pass over the northern Taurus leads from . [5 5 Pliny 5] -LRB- v. 24 -RRB- mentions a 3 Claudiopoli s of [6 6 Cappadocia 6] , and 4 Ptolemy -LRB- v. 7 -RRB- has a Claudiopoli s in 5 Cataonia . Both these passages and those of [9 Ammianus 9] and [10 Theophanes 10] are cited to prove that there is a Claudiopoli s in Cataonia , though it is manifest that the passage in [11 Ammianus 11] at least can only apply to a town in the valley of the [12 Calycadnus 12] in . The two Tauri of [13 Theophanes 13] might mean the Taurus and . But 6 Hierocles places Claudiopoli s in [7 7 Isauria 7] , a description which can not apply to the Claudiopoli s -LRB- es -RRB- of [14 Pliny 14] and [15 Ptolemy 15] . The city apparently received the [16 Roman 16] 7 colony

1: 75462 Cilicia

2: 3298437 Silifke

3: 6140 Claudius

4: 596736 Taurus_Mountains

5: 44920 Pliny_the_Elder

6: 79642 Cappadocia

7: 218914 Isauria

1: 11887 Greek_language

[winner: 12108 Greece, pem_pos=1, final_scores=0.41557926, final_local_scores=0.45250332, pem=-2.3247876, similarity=30.23587]

gt: 11887 Greek_language, pem_pos=0, final_scores=-0.38530803, final_local_scores=-0.42174587, pem=-0.6261292, similarity=-5.091037]

2: 28196796 Theophanes_of_Byzantium

[winner: 272031 Theophanes_the_Confessor, pem_pos=0, final_scores=0.31764084, final_local_scores=0.30348986, pem=0.0, similarity=13.146456]

gt: 28196796 Theophanes_of_Byzantium, pem_pos=10, final_scores=-1.084357, final_local_scores=-1.1160331, pem=-6.2186003, similarity=4.6027765]

[winner: 3719 Byzantium, pem_pos=0, final_scores=0.3115573, final_local_scores=0.26825857, pem=-0.099208966, similarity=12.634598]

RECALL MISS: 28196796 Theophanes_of_Byzantium]

3: 9640309 Claudiopoli s_(Cappadocia)

[winner: 1834318 Bo lu, pem_pos=1, final_scores=-0.2885015, final_local_scores=-0.26502782, pem=-1.0589339, similarity=1.9394004]

gt: 9640309 Claudiopoli s_(Cappadocia), pem_pos=2, final_scores=-0.45573658, final_local_scores=-0.40916407,

pem=-1.4120193, similarity=-0.44177967]
4: 23979 Ptolemy
[winner: 23979 Ptolemy, pem_pos=0,
final_scores=-0.011332378, final_local_scores=-0.057365842,
pem=-0.18743706, similarity=3.3971138]
gt: 23979 Ptolemy, pem_pos=0, final_scores=-0.011332378,
final_local_scores=-0.057365842, pem=-0.18743706,
similarity=3.3971138]
5: 2506884 Cataonia
[winner: 2506884 Cataonia, pem_pos=0,
final_scores=-0.07424274, final_local_scores=-0.038819626,
pem=-0.030213784, similarity=3.095565]
gt: 2506884 Cataonia, pem_pos=0,
final_scores=-0.07424274, final_local_scores=-0.038819626,
pem=-0.030213784, similarity=3.095565]
6: 4335432 Hierocles_(author_of_Synecdemus)
[winner: 4335432 Hierocles_(author_of_Synecdemus),
pem_pos=0, final_scores=0.01125215,
final_local_scores=0.01209712, pem=-0.5042051,
similarity=7.192661]
gt: 4335432 Hierocles_(author_of_Synecdemus), pem_pos=0,
final_scores=0.01125215, final_local_scores=0.01209712,
pem=-0.5042051, similarity=7.192661]
7: 6916 Colony
[winner: 161296 Colony_(biology), pem_pos=2,
final_scores=-1.5808876, final_local_scores=-1.5921147,
pem=-2.462081, similarity=-30.032986]
gt: 6916 Colony, pem_pos=0, final_scores=-1.8028082,
final_local_scores=-1.9546666, pem=-0.42839065,
similarity=-51.91178]
1: 75462 Cilicia
[pem_pos=0, final_scores=0.48080337,
final_local_scores=0.4381165, pem=0.0, similarity=17.163982]
2: 3298437 Silifke
[pem_pos=0, final_scores=0.20875297,
final_local_scores=0.22485343, pem=0.0, similarity=10.7997875]
3: 6140 Claudius
[pem_pos=0, final_scores=0.28063798,
final_local_scores=0.26210168, pem=0.0, similarity=11.91135]
4: 596736 Taurus_Mountains
[pem_pos=0, final_scores=0.11265944,
final_local_scores=0.11424595, pem=0.0, similarity=7.49904]
5: 44920 Pliny_the_Elder
[pem_pos=0, final_scores=0.050218686,
final_local_scores=0.0061519444, pem=-0.049081847,
similarity=4.540214]
6: 79642 Cappadocia
[pem_pos=0, final_scores=0.31741947,
final_local_scores=0.264508, pem=0.0, similarity=11.983159]
7: 218914 Isauria
[pem_pos=0, final_scores=0.123594984,
final_local_scores=0.11295748, pem=0.0, similarity=7.4605894]

1: 75462 Cilicia
[pem_pos=0, final_scores=0.2586636,
final_local_scores=0.19589144, pem=0.0, similarity=9.935504]
2: 12108 Greece
[pem_pos=1, final_scores=0.41557926,
final_local_scores=0.45250332, pem=-2.3247876,
similarity=30.23587]
3: 75462 Cilicia
[pem_pos=0, final_scores=0.52702886,
final_local_scores=0.4886108, pem=0.0, similarity=18.670834]
4: 272031 Theophanes_the_Confessor
[pem_pos=0, final_scores=0.31764084,
final_local_scores=0.30348986, pem=0.0, similarity=13.146456]
5: 3719 Byzantium
[pem_pos=0, final_scores=0.3115573,
final_local_scores=0.26825857, pem=-0.099208966,
similarity=12.634598]
6: 1992718 Göksu
[pem_pos=0, final_scores=0.14792453,
final_local_scores=0.21058416, pem=0.0, similarity=10.373964]
7: 1985695 Seleucia
[pem_pos=0, final_scores=0.1683624,
final_local_scores=0.17321548, pem=0.0, similarity=9.258808]
8: 1992718 Göksu
[pem_pos=0, final_scores=0.030051395,
final_local_scores=0.08182569, pem=0.0, similarity=6.5315547]
9: 1767 Ammianus_Marcellinus
[pem_pos=0, final_scores=0.30328953,
final_local_scores=0.273342, pem=0.0, similarity=12.246784]
10: 272031 Theophanes_the_Confessor
[pem_pos=0, final_scores=0.25055677,
final_local_scores=0.23021078, pem=0.0, similarity=10.9596615]
11: 1767 Ammianus_Marcellinus
[pem_pos=0, final_scores=0.20699544,
final_local_scores=0.16815537, pem=0.0, similarity=9.107804]
12: 1992718 Göksu
[pem_pos=0, final_scores=0.21466254,
final_local_scores=0.2834853, pem=0.0, similarity=12.5494795]
13: 272031 Theophanes_the_Confessor
[pem_pos=0, final_scores=0.2922209,
final_local_scores=0.27572244, pem=0.0, similarity=12.317821]
14: 44920 Pliny_the_Elder
[pem_pos=0, final_scores=0.06933071,
final_local_scores=0.027028948, pem=-0.049081847,
similarity=5.163225]
15: 23979 Ptolemy
[pem_pos=0, final_scores=0.13714243,
final_local_scores=0.11186977, pem=-0.18743706,
similarity=8.447442]
16: 521555 Ancient_Rome
[pem_pos=0, final_scores=0.07111795,
final_local_scores=0.0041914433, pem=-1.0249887,

similarity=9.788841]

Sample: b'Daniel_Bovet'

Daniel Bovet -LRB- 23 March 1907 -- 8 April 1992 -RRB- was a Swiss-born [1 1 Italian 1] 1 pharmacologist who won the 1957 2 Nobel Prize in Physiology or Medicine for his discovery of drugs that block the actions of specific 3 neurotransmitter s . He is best known for his discovery in 1937 of s , which block the neurotransmitter 4 histamine and are used in 5 allergy medication . His other research included work on 6 chemotherapy , 7 sulfa drug s , the 8 sympathetic nervous system , the pharmacology of 9 curare , and other neuropharmacological interests . In 1965 , [1 Bovet 1] led a study team which concluded that smoking of tobacco cigarettes increased users ' intelligence . He told [2 2 The New York Times 2] that the object was not to `` create geniuses , but only -LSB- to -RSB- put the less-endowed individual in a position to reach a satisfactory mental and intellectual development . [2 Bovet 2] was born in [3 3 Fleurier 3] , [4 4 Switzerland 4] . He was a . He graduated from the [5 5 University of Geneva 5] in 1927 and received his doctorate in 1929 . Beginning in 1929 until 1947 he worked at the [6 6 Pasteur Institute 6] in [7 7 Paris 7] . He then moved in 1947 to the 10 Istituto Superiore di Sanità -LRB- Superior Institute of Health -RRB- in [8 8 Rome 8] . In 1964 , he became a professor in at the [9 9 University of Sassari 9] in [3 Italy 3] . From 1969 to 1971 , he was the head of the [4 National Research Council 4] in [5 Rome 5] before stepping down to become a professor at the

1: 14532 Italy

2: 30680 The_New_York_Times

3: 6301466 Fleurier

4: 26748 Switzerland

5: 342348 University_of_Geneva

6: 866902 Pasteur_Institute

7: 22989 Paris

8: 25458 Rome

9: 4894028 University_of_Sassari

1: 24354 Pharmacology

[winner: 5992 Traditional_Chinese_medicine, pem_pos=0, final_scores=-1.1140062, final_local_scores=-1.1513839, pem=0.0, similarity=-30.269876]

gt: 24354 Pharmacology, pem_pos=1, final_scores=-1.3062842, final_local_scores=-1.4309995, pem=-0.07083635, similarity=-38.22895]

2: 52502 Nobel_Prize_in_Physiology_or_Medicine

[winner: 21201 Nobel_Prize, pem_pos=0, final_scores=-0.0055547208, final_local_scores=-0.024895892, pem=-0.15178943, similarity=4.172224]

gt: 52502 Nobel_Prize_in_Physiology_or_Medicine, pem_pos=5, final_scores=-0.31182057, final_local_scores=-0.33617097, pem=-1.7501515, similarity=3.5752966]

[winner: 18957 Medicine, pem_pos=0,
final_scores=-0.5924325, final_local_scores=-0.65991575,
pem=-0.5636483, similarity=-12.538282]
gt: 52502 Nobel_Prize_in_Physiology_or_Medicine,
pem_pos=2, final_scores=-0.88976383,
final_local_scores=-0.9674862, pem=-2.8691354,
similarity=-9.1792]
[winner: 880497 Linoleic_acid, pem_pos=0,
final_scores=-1.2041136, final_local_scores=-1.1537411,
pem=0.0, similarity=-30.340223]
RECALL MISS: 52502 Nobel_Prize_in_Physiology_or_Medicine]
3: 21865 Neurotransmitter
[winner: 21865 Neurotransmitter, pem_pos=0,
final_scores=-1.058732, final_local_scores=-1.0541217,
pem=-0.0074249697, similarity=-27.327003]
gt: 21865 Neurotransmitter, pem_pos=0,
final_scores=-1.058732, final_local_scores=-1.0541217,
pem=-0.0074249697, similarity=-27.327003]
4: 395877 Histamine
[winner: 395877 Histamine, pem_pos=0,
final_scores=-0.90342873, final_local_scores=-0.83711624,
pem=-0.007939, similarity=-20.848331]
gt: 395877 Histamine, pem_pos=0,
final_scores=-0.90342873, final_local_scores=-0.83711624,
pem=-0.007939, similarity=-20.848331]
5: 55313 Allergy
[winner: 55313 Allergy, pem_pos=0,
final_scores=-1.2411798, final_local_scores=-1.2759366,
pem=-0.017777046, similarity=-33.890106]
gt: 55313 Allergy, pem_pos=0, final_scores=-1.2411798,
final_local_scores=-1.2759366, pem=-0.017777046,
similarity=-33.890106]
6: 7172 Chemotherapy
[winner: 7172 Chemotherapy, pem_pos=0,
final_scores=-1.1660546, final_local_scores=-1.2151289,
pem=-0.0044940813, similarity=-32.147717]
gt: 7172 Chemotherapy, pem_pos=0,
final_scores=-1.1660546, final_local_scores=-1.2151289,
pem=-0.0044940813, similarity=-32.147717]
7: 499685 Sulfonamide_(medicine)
[winner: 499685 Sulfonamide_(medicine), pem_pos=0,
final_scores=-0.7016205, final_local_scores=-0.67787737,
pem=0.0, similarity=-16.1395]
gt: 499685 Sulfonamide_(medicine), pem_pos=0,
final_scores=-0.7016205, final_local_scores=-0.67787737,
pem=0.0, similarity=-16.1395]
[winner: 803469 ATC_code_N, pem_pos=9,
final_scores=-0.9699911, final_local_scores=-0.7699283,
pem=-2.9199536, similarity=-3.007321]
RECALL MISS: 499685 Sulfonamide_(medicine)]
8: 193753 Sympathetic_nervous_system
[winner: 193753 Sympathetic_nervous_system, pem_pos=0,

```
final_scores=-0.86150414, final_local_scores=-0.7943014,
pem=-0.0021337967, similarity=-19.60222|
    gt: 193753 Sympathetic_nervous_system, pem_pos=0,
final_scores=-0.86150414, final_local_scores=-0.7943014,
pem=-0.0021337967, similarity=-19.60222 ]
9: 1841740 Curare
    [winner: 1841740 Curare, pem_pos=0,
final_scores=-0.8716403, final_local_scores=-0.84439224,
pem=0.0, similarity=-21.108635|
    gt: 1841740 Curare, pem_pos=0, final_scores=-0.8716403,
final_local_scores=-0.84439224, pem=0.0,
similarity=-21.108635 ]
10: 25164846 Istituto_Superiore_di_Sanità
    [winner: 25164846 Istituto_Superiore_di_Sanità,
pem_pos=0, final_scores=-0.39186734,
final_local_scores=-0.38060784, pem=0.0,
similarity=-7.268384|
    gt: 25164846 Istituto_Superiore_di_Sanità, pem_pos=0,
final_scores=-0.39186734, final_local_scores=-0.38060784,
pem=0.0, similarity=-7.268384 ]
1: 14532 Italy
    [pem_pos=0, final_scores=0.7092436,
final_local_scores=0.745789, pem=-0.75948656,
similarity=30.47575]
2: 30680 The_New_York_Times
    [pem_pos=0, final_scores=0.4142015,
final_local_scores=0.50702393, pem=0.0, similarity=19.220318]
3: 6301466 Fleurier
    [pem_pos=0, final_scores=0.14460747,
final_local_scores=0.2481071, pem=0.0, similarity=11.493724]
4: 26748 Switzerland
    [pem_pos=0, final_scores=0.6875791,
final_local_scores=0.72199684, pem=-0.18598008,
similarity=26.646927]
5: 342348 University_of_Geneva
    [pem_pos=0, final_scores=0.48692125,
final_local_scores=0.49770337, pem=0.0, similarity=18.942175]
6: 866902 Pasteur_Institute
    [pem_pos=0, final_scores=0.32507223,
final_local_scores=0.3340671, pem=0.0, similarity=14.058942]
7: 22989 Paris
    [pem_pos=0, final_scores=0.73740786,
final_local_scores=0.7526042, pem=-0.115301244,
similarity=27.17595]
8: 25458 Rome
    [pem_pos=0, final_scores=0.48553842,
final_local_scores=0.493124, pem=-0.2517334,
similarity=20.174482]
9: 4894028 University_of_Sassari
    [pem_pos=0, final_scores=0.24324326,
final_local_scores=0.3000231, pem=0.0, similarity=13.043001]
1: 1408625 Daniel_Bovet
```

[pem_pos=0, final_scores=0.13699593,
 final_local_scores=0.16261742, pem=0.0, similarity=8.942541]
 2: 1408625 Daniel_Bovet
 [pem_pos=0, final_scores=0.24384217,
 final_local_scores=0.2793307, pem=0.0, similarity=12.425498]
 3: 14532 Italy
 [pem_pos=0, final_scores=0.60163075,
 final_local_scores=0.6282383, pem=-0.20612453,
 similarity=23.958536]
 4: 37638 United_States_National_Research_Council
 [pem_pos=0, final_scores=0.22041579,
 final_local_scores=0.28347462, pem=0.0, similarity=12.549161]
 5: 25458 Rome
 [pem_pos=0, final_scores=0.4986167,
 final_local_scores=0.50741, pem=-0.2517334,
 similarity=20.600803]

Sample: b'William_Carr_Lane'

William Carr Lane -LRB- December 1 , 1789January 6 , 1863 -
 RRB- was a 1 doctor and the first 2 Mayor of , [1 1 Missouri 1
] , serving from 1823 to 1829 and 1837 to 1840 . He was also
 the Governor of [2 2 New Mexico Territory 2] from 1852 to
 1853 . Born in [1 3 Fayette County 1] , [2 Pennsylvania 2] to
 Presley Carr Lane and Sarah `` [3 Sallie 3] `` Stephenson , [4
 Lane 4] attended college in [3 3 Pennsylvania 3] and studied 4
 medicine in [5 5 Louisville 5] , [6 Kentucky 6] . He entered
 the , and was appointed 6 post surgeon at [7 7 Fort Harrison 7
] on the [4 4 Wabash River 4] north of [8 8 Terre Haute 8] , [9
 Indiana 9] in 1816 . He resigned from the army in 1819 to
 enter private practice . He married February 26 , 1818 in [10
 Vincennes 10] , [11 Indiana 11] to Miss Mary Ewing ,
 daughter of [12 Nathaniel Ewing 12] and Ann Breeding . Their
 children were Anne Ewing Lane -LRB- 1819 -- 1904 -RRB- ,
 Sarah L. Lane -LRB- 1821 -- 1887 -RRB- , and Victor Carr Lane
 -LRB- 1831 -- 1848 -RRB- . [13 Lane 13] served as [14 St.
 Louis 14] 's first mayor from 1823 to 1829 , when the city 's
 population was around 4,000 . He oversaw the first public
 health system in the city , free public schools , and street
 improvements , including the paving of [15 Main Street 15] . [16
 Lane 16] helped erect the city 's first . He was also
 instrumental in beautifying the city with fountains and
 greenery . The was adopted , and election procedures were
 written . Perhaps the most memorable event in his term was a
 visit by in 1825 , and a ball given in his honor . [17 Lane 17
] served again as mayor from 1837 to 1840 . In 1852 ,
 President [5 5 Millard Fillmore 5] appointed him to be the
 new governor of the [6 6 New Mexico Territory 6]

1: 19571 Missouri
 2: 440641 New_Mexico_Territory
 3: 23332 Pennsylvania
 4: 427204 Wabash_River
 5: 19729548 Millard_Fillmore

6: 440641 New_Mexico_Territory

1: 23315 Physician

[winner: 23315 Physician, pem_pos=0,
final_scores=-1.3169754, final_local_scores=-1.4092853,
pem=-0.18676473, similarity=-36.950512|
gt: 23315 Physician, pem_pos=0, final_scores=-1.3169754,
final_local_scores=-1.4092853, pem=-0.18676473,
similarity=-36.950512]

2: 1018942 Mayor_of_St._Louis

[winner: 101146 Mayor, pem_pos=0,
final_scores=-0.6899908, final_local_scores=-0.74869657,
pem=-0.28877318, similarity=-16.682491|

RECALL MISS: 1018942 Mayor_of_St._Louis]

3: 91959 Fayette_County,_Pennsylvania

[winner: 23332 Pennsylvania, pem_pos=0,
final_scores=0.56042916, final_local_scores=0.5325574,
pem=-0.11810702, similarity=20.624573|

RECALL MISS: 91959 Fayette_County,_Pennsylvania]

[winner: 91959 Fayette_County,_Pennsylvania, pem_pos=0,
final_scores=0.13501935, final_local_scores=0.109884605,
pem=-1.2421032, similarity=14.123639|

gt: 91959 Fayette_County,_Pennsylvania, pem_pos=0,
final_scores=0.13501935, final_local_scores=0.109884605,
pem=-1.2421032, similarity=14.123639]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7953987, final_local_scores=-1.7870471,
pem=-4.6728287, similarity=-23.82777|

RECALL MISS: 91959 Fayette_County,_Pennsylvania]

4: 18957 Medicine

[winner: 7518132 Medicinal_plants_of_the_American_West,
pem_pos=13, final_scores=-0.9910032,
final_local_scores=-0.8769384, pem=-2.5649493,
similarity=-8.131287|

gt: 18957 Medicine, pem_pos=0, final_scores=-1.2485694,
final_local_scores=-1.2896702, pem=-0.09248193,
similarity=-33.893692]

5: 58592 Louisville,_Kentucky

[winner: 58592 Louisville,_Kentucky, pem_pos=0,
final_scores=0.43110734, final_local_scores=0.40936774,
pem=-0.72280806, similarity=20.236807|

gt: 58592 Louisville,_Kentucky, pem_pos=0,
final_scores=0.43110734, final_local_scores=0.40936774,
pem=-0.72280806, similarity=20.236807]

[winner: 16846 Kentucky, pem_pos=0,
final_scores=0.41399783, final_local_scores=0.38280714,
pem=-0.28237545, similarity=17.049044|

RECALL MISS: 58592 Louisville,_Kentucky]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7976049, final_local_scores=-1.7894571,
pem=-4.6728287, similarity=-23.899685|

RECALL MISS: 58592 Louisville,_Kentucky]

6: 9104734 Surgeon

[winner: 1168304 Surgical_oncology, pem_pos=0,
final_scores=-1.2320842, final_local_scores=-1.2368991,
pem=0.0, similarity=-32.821823|
gt: 9104734 Surgeon, pem_pos=1, final_scores=-1.548827,
final_local_scores=-1.6558068, pem=-0.46885663,
similarity=-42.773148]
[winner: 18440640 Post,_Oregon, pem_pos=16,
final_scores=-1.4378992, final_local_scores=-1.453645,
pem=-3.9696014, similarity=-17.702644|
RECALL MISS: 9104734 Surgeon]
7: 12134724 Fort_Harrison
[winner: 1353048 Siege_of_Fort_Harrison, pem_pos=1,
final_scores=0.1311379, final_local_scores=0.17958537,
pem=-0.5389965, similarity=12.380046|
gt: 12134724 Fort_Harrison, pem_pos=0,
final_scores=0.045373455, final_local_scores=0.11059308,
pem=-0.45306793, similarity=9.853885]
8: 100584 Terre_Haute,_Indiana
[winner: 21883857 Indiana, pem_pos=0,
final_scores=0.4896261, final_local_scores=0.44616157,
pem=-0.27230176, similarity=18.884884|
RECALL MISS: 100584 Terre_Haute,_Indiana]
[winner: 100584 Terre_Haute,_Indiana, pem_pos=0,
final_scores=0.38513964, final_local_scores=0.36606407,
pem=0.0, similarity=15.013795|
gt: 100584 Terre_Haute,_Indiana, pem_pos=0,
final_scores=0.38513964, final_local_scores=0.36606407,
pem=0.0, similarity=15.013795]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.6828002, final_local_scores=-1.6640503,
pem=-4.6728287, similarity=-20.157297|
RECALL MISS: 100584 Terre_Haute,_Indiana]
9: 112391 Vincennes,_Indiana
[winner: 21883857 Indiana, pem_pos=0,
final_scores=0.5421367, final_local_scores=0.5035214,
pem=-0.27230176, similarity=20.596615|
RECALL MISS: 112391 Vincennes,_Indiana]
[winner: 112391 Vincennes,_Indiana, pem_pos=1,
final_scores=0.2642789, final_local_scores=0.22383979,
pem=-0.88693506, similarity=15.592829|
gt: 112391 Vincennes,_Indiana, pem_pos=1,
final_scores=0.2642789, final_local_scores=0.22383979,
pem=-0.88693506, similarity=15.592829]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.6479155, final_local_scores=-1.6259441,
pem=-4.6728287, similarity=-19.020134|
RECALL MISS: 112391 Vincennes,_Indiana]
1: 19571 Missouri
[pem_pos=0, final_scores=0.47557682,
final_local_scores=0.4372173, pem=-0.27068004,
similarity=18.60915]
2: 440641 New_Mexico_Territory

[pem_pos=0, final_scores=0.38509357,
final_local_scores=0.37916666, pem=0.0, similarity=15.404802]
3: 23332 Pennsylvania
[pem_pos=0, final_scores=0.49848622,
final_local_scores=0.46489424, pem=-0.11810702,
similarity=18.605368]
4: 427204 Wabash_River
[pem_pos=0, final_scores=0.36575317,
final_local_scores=0.37438822, pem=0.0, similarity=15.262204]
5: 19729548 Millard_Fillmore
[pem_pos=0, final_scores=0.41338915,
final_local_scores=0.42592525, pem=0.0, similarity=16.800173]
6: 440641 New_Mexico_Territory
[pem_pos=0, final_scores=0.2998613,
final_local_scores=0.28606337, pem=0.0, similarity=12.626414]
1: 91959 Fayette_County,_Pennsylvania
[pem_pos=0, final_scores=0.13501935,
final_local_scores=0.109884605, pem=-1.2421032,
similarity=14.123639]
2: 23332 Pennsylvania
[pem_pos=0, final_scores=0.56042916,
final_local_scores=0.5325574, pem=-0.11810702,
similarity=20.624573]
3: 277347 Sallie_Mae
[pem_pos=0, final_scores=0.054117247,
final_local_scores=0.14363647, pem=0.0, similarity=8.376112]
4: 3465017 William_Carr_Lane
[pem_pos=0, final_scores=0.11268453,
final_local_scores=0.11700091, pem=0.0, similarity=7.581253]
5: 58592 Louisville,_Kentucky
[pem_pos=0, final_scores=0.43110734,
final_local_scores=0.40936774, pem=-0.72280806,
similarity=20.236807]
6: 16846 Kentucky
[pem_pos=0, final_scores=0.41399783,
final_local_scores=0.38280714, pem=-0.28237545,
similarity=17.049044]
7: 1353048 Siege_of_Fort_Harrison
[pem_pos=1, final_scores=0.1311379,
final_local_scores=0.17958537, pem=-0.5389965,
similarity=12.380046]
8: 100584 Terre_Haute,_Indiana
[pem_pos=0, final_scores=0.38513964,
final_local_scores=0.36606407, pem=0.0, similarity=15.013795]
9: 21883857 Indiana
[pem_pos=0, final_scores=0.4896261,
final_local_scores=0.44616157, pem=-0.27230176,
similarity=18.884884]
10: 112391 Vincennes,_Indiana
[pem_pos=1, final_scores=0.2642789,
final_local_scores=0.22383979, pem=-0.88693506,
similarity=15.592829]

11: 21883857 Indiana
 [pem_pos=0, final_scores=0.5421367,
 final_local_scores=0.5035214, pem=-0.27230176,
 similarity=20.596615]
 12: 21809779 Nathaniel_Ewing
 [pem_pos=0, final_scores=0.07730268,
 final_local_scores=0.17595181, pem=0.0, similarity=9.340466]
 13: 3465017 William_Carr_Lane
 [pem_pos=0, final_scores=0.102720246,
 final_local_scores=0.10611643, pem=0.0, similarity=7.2564387]
 14: 27687 St._Louis
 [pem_pos=0, final_scores=0.31229752,
 final_local_scores=0.27958867, pem=-0.28531665,
 similarity=13.984793]
 15: 23523337 Main_Street
 [pem_pos=0, final_scores=0.036532566,
 final_local_scores=0.07943413, pem=-0.831929,
 similarity=10.984345]
 16: 3465017 William_Carr_Lane
 [pem_pos=0, final_scores=0.08024387,
 final_local_scores=0.08156441, pem=0.0, similarity=6.5237575]
 17: 3465017 William_Carr_Lane
 [pem_pos=0, final_scores=0.24675892,
 final_local_scores=0.2634568, pem=0.0, similarity=11.951789]

Sample: b'Yakov_Malkiel'

[1 Yakov Malkiel 1] -LRB- July 22 , 1914 -- April 24 , 1998 -
 RRB- was a -LRB- [1 1 Russia 1] n-born -RRB- Romance 1 etymol
 ogist and 2 philologist . His specialty was the development
 of words , roots , prefixes , and suffixes in modern [2 3 Rom
 ance 2] languages , particularly [3 4 Spanish 3] . He was the
 founder of the journal . [4 Malkiel 4] was born in [2 2 Kiev 2
] to a [3 3 Russian-Jewish 3] family , and was brought up and
 educated in [4 4 Berlin 4] , after the [5 5 Russian Civil War
 5] . Despite an early interest in literature , he ended up
 studying 5 linguistics at the [6 6 Humboldt University of Ber
 lin 6] , then known as the [5 Friedrich-Wilhelms-Universität 5
] . Being a [6 Jew 6] in 1930s [7 Germany 7] was an obstacle
 to his education , but one he was able to overcome ; his
 family finally emigrated to the [8 United States 8] in 1940 .
 After two years unemployed in [9 New York 9] , [10 Malkiel 10]
 accepted a one-term appointment at the [7 7 University of Wyo
 ming 7] in [8 8 Laramie 8] . In 1943 , he was offered an
 initially temporary position at the [11 6 University of Calif
 ornia 11] , [12 Berkeley 12] , which later was converted to a
 permanent professorship ; [13 Malkiel 13] remained there
 until his retirement in 1983 , teaching in the departments of
 [14 Spanish 14] and -LRB- later -RRB- Linguistics . He
 married fellow [15 Berkeley 15] professor [9 9 María Rosa Lida
 de Malkiel 9] , a philologist and literary critic from [10 10
 Argentina 10] , in 1948 . During a period when etymology was
 receding from prominence in linguistics , [16 Malkiel 16] was

both one of its chief champions and most rigorous theorists . Best known for his work on the role of sound in the development of suffixes , [17 Malkiel 17] coined the term to describe the influence in sound words tend to have over the development of their opposites , when s occur in pairs . A major secondary interest was in the history of his field , explored in the pages of [18 Romance Philology 18] and in his last book , Etymology . His work in all fields was characterized by a doggedly comprehensive use of evidence ; of his book , Development of the Latin Suffixes - antia and - entia in [19 Romance Languages 19] , influential linguist [11 11 Leo Spitzer 11] said in a review , `` No one can fail to be impressed by this outstanding example of

1: 25391 Russia

2: 585629 Kiev

3: 210710 History_of_the_Jews_in_Russia

4: 3354 Berlin

5: 26295 Russian_Civil_War

6: 308234 Humboldt_University_of_Berlin

7: 535684 University_of_Wyoming

8: 140050 Laramie, Wyoming

9: 31751240 María_Rosa_Lida_de_Malkiel

10: 18951905 Argentina

11: 8662917 Leo_Spitzer

1: 888207 Etymology

[winner: 888207 Etymology, pem_pos=0, final_scores=-2.0222652, final_local_scores=-2.185981, pem=0.0, similarity=-61.144318]

gt: 888207 Etymology, pem_pos=0, final_scores=-2.0222652, final_local_scores=-2.185981, pem=0.0, similarity=-61.144318]

2: 23193 Philology

[winner: 23193 Philology, pem_pos=0, final_scores=-1.3688197, final_local_scores=-1.555188, pem=-0.011182733, similarity=-42.259384]

gt: 23193 Philology, pem_pos=0, final_scores=-1.3688197, final_local_scores=-1.555188, pem=-0.011182733, similarity=-42.259384]

3: 25401 Romance_languages

[winner: 25401 Romance_languages, pem_pos=0, final_scores=0.07221566, final_local_scores=0.07275204, pem=-1.1368215, similarity=12.442991]

gt: 25401 Romance_languages, pem_pos=0, final_scores=0.07221566, final_local_scores=0.07275204, pem=-1.1368215, similarity=12.442991]

[winner: 3552038 Languages_of_Egypt, pem_pos=1, final_scores=-1.3339012, final_local_scores=-1.3472679, pem=-1.3716635, similarity=-28.656128]

RECALL MISS: 25401 Romance_languages]

4: 26825 Spanish_language

[winner: 26667 Spain, pem_pos=1, final_scores=0.19347085, final_local_scores=0.21942592, pem=-1.2425895,


```

similarity=17.395214|
    gt: 26825 Spanish_language, pem_pos=0,
    final_scores=0.051559374, final_local_scores=0.051508933,
    pem=-0.6062411, similarity=8.923676 ]
5: 22760983 Linguistics
    [winner: 22760983 Linguistics, pem_pos=0,
    final_scores=-1.1661887, final_local_scores=-1.3152623,
    pem=-0.028691405, similarity=-35.004307|
    gt: 22760983 Linguistics, pem_pos=0,
    final_scores=-1.1661887, final_local_scores=-1.3152623,
    pem=-0.028691405, similarity=-35.004307 ]
6: 31922 University_of_California,_Berkeley
    [winner: 31921 University_of_California, pem_pos=0,
    final_scores=0.4399693, final_local_scores=0.4443543,
    pem=-0.094431244, similarity=17.863663|
    gt: 31922 University_of_California,_Berkeley, pem_pos=1,
    final_scores=0.06631012, final_local_scores=0.026198566,
    pem=-2.24399, similarity=17.074696 ]
    [winner: 4860 Berkeley,_California, pem_pos=0,
    final_scores=0.31457746, final_local_scores=0.29555696,
    pem=-0.7355298, similarity=16.909647|
    gt: 31922 University_of_California,_Berkeley, pem_pos=1,
    final_scores=-0.014442354, final_local_scores=-0.06840727,
    pem=-1.8697176, similarity=12.216119 ]
    [winner: 37537378 Admissions_(2011_film), pem_pos=12,
    final_scores=-1.6921707, final_local_scores=-1.6608963,
    pem=-4.6728287, similarity=-20.063173|
    RECALL MISS: 31922 University_of_California,_Berkeley]
1: 25391 Russia
    [pem_pos=0, final_scores=0.4129172,
    final_local_scores=0.4667474, pem=-0.20299491,
    similarity=19.122303]
2: 585629 Kiev
    [pem_pos=0, final_scores=0.44693393,
    final_local_scores=0.4709226, pem=-0.032242518,
    similarity=18.318321]
3: 210710 History_of_the_Jews_in_Russia
    [pem_pos=0, final_scores=0.3180272,
    final_local_scores=0.33548012, pem=-0.18583651,
    similarity=15.111717]
4: 3354 Berlin
    [pem_pos=0, final_scores=0.5411008,
    final_local_scores=0.5502611, pem=-0.15483057,
    similarity=21.352594]
5: 26295 Russian_Civil_War
    [pem_pos=0, final_scores=0.1911758,
    final_local_scores=0.24184847, pem=0.0, similarity=11.306953]
6: 308234 Humboldt_University_of_Berlin
    [pem_pos=0, final_scores=0.36978167,
    final_local_scores=0.3703546, pem=0.0, similarity=15.141832]
7: 535684 University_of_Wyoming
    [pem_pos=0, final_scores=0.40282464,

```

final_local_scores=0.44576514, pem=0.0, similarity=17.392233]
8: 140050 Laramie, Wyoming
[pem_pos=0, final_scores=0.23461814,
final_local_scores=0.2734978, pem=-0.46636093,
similarity=14.787576]
9: 31751240 María_Rosa_Lida_de_Malkiel
[pem_pos=0, final_scores=0.16386388,
final_local_scores=0.17226484, pem=0.0, similarity=9.230439]
10: 18951905 Argentina
[pem_pos=0, final_scores=0.6018155,
final_local_scores=0.65676194, pem=-0.30003807,
similarity=25.320452]
11: 8662917 Leo_Spitzer
[pem_pos=0, final_scores=0.15344761,
final_local_scores=0.22996667, pem=0.0, similarity=10.952377]
1: 18738051 Yakov_Malkiel
[pem_pos=0, final_scores=0.22760864,
final_local_scores=0.26354787, pem=0.0, similarity=11.954507]
2: 25401 Romance_languages
[pem_pos=0, final_scores=0.07221566,
final_local_scores=0.07275204, pem=-1.1368215,
similarity=12.442991]
3: 26667 Spain
[pem_pos=1, final_scores=0.19347085,
final_local_scores=0.21942592, pem=-1.2425895,
similarity=17.395214]
4: 31751240 María_Rosa_Lida_de_Malkiel
[pem_pos=0, final_scores=0.08284195,
final_local_scores=0.08293007, pem=0.0, similarity=6.564512]
5: 308234 Humboldt_University_of_Berlin
[pem_pos=0, final_scores=0.38950473,
final_local_scores=0.39189905, pem=0.0, similarity=15.784762]
6: 25955086 Jews
[pem_pos=0, final_scores=0.43235075,
final_local_scores=0.44416565, pem=0.0, similarity=17.344501]
7: 11867 Germany
[pem_pos=0, final_scores=0.5361969,
final_local_scores=0.56203604, pem=-0.15529114,
similarity=21.706488]
8: 3434750 United_States
[pem_pos=0, final_scores=0.472067,
final_local_scores=0.48045367, pem=-0.073457584,
similarity=18.826883]
9: 645042 New_York_City
[pem_pos=1, final_scores=0.5183382,
final_local_scores=0.5195889, pem=-1.7906687,
similarity=29.333216]
10: 695839 Burton_Malkiel
[pem_pos=2, final_scores=0.07127701,
final_local_scores=0.13297898, pem=-0.06453852,
similarity=8.409041]
11: 31921 University_of_California

[pem_pos=0, final_scores=0.4399693,
 final_local_scores=0.4443543, pem=-0.094431244,
 similarity=17.863663]
 12: 4860 Berkeley, California
 [pem_pos=0, final_scores=0.31457746,
 final_local_scores=0.29555696, pem=-0.7355298,
 similarity=16.909647]
 13: 695839 Burton Malkiel
 [pem_pos=2, final_scores=0.09406234,
 final_local_scores=0.15718248, pem=-0.06453852,
 similarity=9.131322]
 14: 26825 Spanish language
 [pem_pos=0, final_scores=0.17952184,
 final_local_scores=0.18744478, pem=-0.6062411,
 similarity=12.980272]
 15: 4860 Berkeley, California
 [pem_pos=0, final_scores=0.48911017,
 final_local_scores=0.4886129, pem=-0.7355298,
 similarity=22.67082]
 16: 695839 Burton Malkiel
 [pem_pos=2, final_scores=0.19415645,
 final_local_scores=0.26652008, pem=-0.06453852,
 similarity=12.394174]
 17: 31751240 María Rosa Lida de Malkiel
 [pem_pos=0, final_scores=0.20372836,
 final_local_scores=0.21564487, pem=0.0, similarity=10.524986]
 18: 23193 Philology
 [pem_pos=1, final_scores=0.2608375,
 final_local_scores=0.23163927, pem=-0.9162907,
 similarity=15.985222]
 19: 25401 Romance languages
 [pem_pos=0, final_scores=0.23975007,
 final_local_scores=0.25575793, pem=0.0, similarity=11.722039]

Sample: b'Maxwell_Hunter'

Maxwell White Hunter II -LRB- March 11 , 1922 -- November
 10 , 2001 -RRB- was a prominent [1 American 1] aerospace
 engineer . He worked on the design of the [2 1 Douglas B-42 2]
 and [3 B-43 3] bombers , the [4 Honest John 4] , Nike-Ajax ,
 and Nike-Zeus missiles , the [5 Thor IRBM 5] , and on parts
 of the [1 1 Strategic Defense Initiative 1] . In later years
 he worked on space-launch vehicles and was a proponent of
 Single-Stage-to-Orbit -LRB- SSTO -RRB- designs . He was
 honored in 1995 by the [2 2 National Space Society 2] for
 lifelong contributions to the technology of spaceflight . [6
 Maxwell Hunter 6] was born in [7 2 Hollidaysburg 7] , [8 Penn
 sylvania 8] and graduated from in that state with a degree in
 physics and mathematics . In 1944 , he earned a Master 's
 Degree in aeronautical engineering from the 3 Massachusetts In
 stitute of Technology -LRB- MIT -RRB- . After earning his MS
 from [9 MIT 9] , [10 Hunter 10] went to work for the [11
 Douglas Aircraft Company 11] . He became their chief designer

in aeronautics , working on modifications of the B-42 and [12 B-43 12] bombers . Later he was promoted to chief engineer for space systems . In this position he oversaw the production of military missiles including the Honest John artillery rocket , the Nike anti-aircraft missile , and the Sparrow air-to-air missile . [13 Douglas Aircraft 13] began work on the 4 Thor , an early -LRB- IRBM -RRB- , in late 1955 and put [14 Hunter 14] in charge of that program . He assembled a small team of crack designers and brought the development to success in a short time . Thor became the basis of the [3 3 McDonnell Douglas 3] , whose descendants are still boosting payloads into [15 Earth 15] orbit today . After working for [16 Douglas Aircraft 16] from 1944 to 1961 , [17 Hunter 17] joined the staff of the , a group that advised the [18 U.S. 18] president on space policy . Following several years in that position , he went back to the aerospace industry in 1965 , spending over two decades with Lockheed Missiles and Space Company -LRB- LMSC -RRB- . -LRB- LMSC was a division of [4 4 Lockheed Corporation 4] , which merged with [5 5 Martin Marietta 5] in 1995 to become . -RRB- At [19 LMSC 19] , [20 Hunter 20] worked on a variety of projects including parts of the mammoth [6 6 Strategic Defense Initiative 6] . [21 SDI 21] gave him a chance at realizing a dream he had nurtured since the 1960s : the construction of a 5 Single-stage-to-orbit -LRB- SSTO -RRB- vehicle . In 1984 , [22 Hunter 22] proposed a vehicle he called the X-rocket , based on those earlier SSTO designs . The design was reviewed by [23 LMSC 23] 's Astronautics Division -LRB- AD -RRB- , which judged it worth pursuing , and by its Missile Systems Division -LRB- MSD -RRB- , which did not . [24 LMSC 24] then declined to support further work on the X-rocket , and [25 Hunter 25] retired . As an independent consultant , he renamed the concept SSX -LRB- for Space Ship Experimental -RRB- , and began to refine it . In December 1988 the ad hoc was briefed on the concept by [26 Hunter 26] and others . The general concept was endorsed by the Council and by High Frontier , Inc. , a Washington-based group headed by Lt. Gen. [7 7 Daniel O. Graham 7] that lobbied for [27 SDI 27] programs . Working together , [28 Hunter 28] and High Frontier convinced SDI management and other national officials that a study should be initiated to determine once and for all the feasibility of SSTO vehicles for military missions . This led to a [29 U.S. Department of Defense 29] -LRB- DoD -RRB- program to design a single-stage vehicle that could replace military satellites on short notice during a national emergency . This vehicle was known as the [30 Delta Clipper 30] . [31 6 McDonnell Douglas Corporation 31] was contracted to build and test fly a one-third-scale , suborbital testbed called the 7 DC-X , or [32 Delta Clipper Experimental 32] . While not capable of spaceflight , the [33 DC-X 33] incorporated many of the technologies needed for an SSTO vehicle , including highly automated systems enabling a

quick turnaround -LRB- just twenty-six hours -RRB- between launches . It made eight successful test flights between August 18 , 1993 , and July 7 , 1995 , and then was taken over by [34 NASA 34] and flown four times as the [35 DC-XA 35] between May 18 and July 31 , 1996 . It was damaged on its last flight when one landing strut failed to deploy and the vehicle tipped over at landing . Funding to make repairs and continue flight tests was not made available . [36 Max Hunter 36] continued to advise on space policy issues and advocate the building of SSTO vehicles for airline-like access to space , a goal he was sure could be attained . He chaired the Rules Committee for the [8 8 Ansari X Prize 8] , announced in 1996 , a \$ 10 million prize for the first reusable craft to reach space twice within two weeks without substantial refit . In October 2000 , he agreed to serve as Chairman of the Advisory Board for the Undersea Hotel Project of the Cala Corporation . He died at [37 Stanford Hospital 37] , [38 Stanford 38] , [39 California 39] on the evening of November 10 , 2001 after a lingering illness . [40 Max Hunter 40] was [9 9 Phi Beta Kappa 9] , 8 Tau Beta Pi , and a Fellow of the 9 American Institute of Aeronautics and Astronautics -LRB- AIAA -RRB- , the 10 American Astronautical Society -LRB- AAS -RRB- , and the [10 10 British Interplanetary Society 10] -LRB- BIS -RRB- . He was also a member of the 11 International Academy of Astronautics -LRB- IAA -RRB- and an honorary member of the . In 1995 , he received the Wernher von Braun Memorial Award of the [11 11 National Space Society 11] for `` lifelong contributions to the fields of rockets , missiles and spaceflight . '' He was also the recipient of a [41 NASA Public Service Medal 41] for `` the definition and promotion of the [42 Space Shuttle 42] and its utilization . '' The 12 S pace Frontier Foundation

- 1: 29186 Strategic_Defense_Initiative
- 2: 317304 National_Space_Society
- 3: 154809 McDonnell_Douglas
- 4: 50042 Lockheed_Corporation
- 5: 497478 Martin_Marietta
- 6: 29186 Strategic_Defense_Initiative
- 7: 2087022 Daniel_O._Graham
- 8: 328602 Ansari_X_Prize
- 9: 166504 Phi_Beta_Kappa_Society
- 10: 1462668 British_Interplanetary_Society
- 11: 317304 National_Space_Society

1: 201793 Douglas_Aircraft_Company
 [winner: 559959 Douglas_XB-42_Mixmaster, pem_pos=0,
 final_scores=0.09151189, final_local_scores=0.13221505,
 pem=0.0, similarity=8.035274]

RECALL MISS: 201793 Douglas_Aircraft_Company]

2: 131435 Hollidaysburg,_Pennsylvania
 [winner: 23332 Pennsylvania, pem_pos=0,
 final_scores=0.5062311, final_local_scores=0.5848828,
 pem=-0.11810702, similarity=22.186066]

RECALL MISS: 131435 Hollidaysburg,_Pennsylvania]
[winner: 131435 Hollidaysburg,_Pennsylvania, pem_pos=0,
final_scores=0.2699653, final_local_scores=0.37740403,
pem=0.0, similarity=15.352202]
gt: 131435 Hollidaysburg,_Pennsylvania, pem_pos=0,
final_scores=0.2699653, final_local_scores=0.37740403,
pem=0.0, similarity=15.352202]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.6931434, final_local_scores=-1.6691282,
pem=-4.6728287, similarity=-20.308828]
RECALL MISS: 131435 Hollidaysburg,_Pennsylvania]
3: 18879 Massachusetts_Institute_of_Technology
[winner: 18879 Massachusetts_Institute_of_Technology,
pem_pos=0, final_scores=-0.052566916,
final_local_scores=-0.08279834, pem=0.0,
similarity=1.6188453]
gt: 18879 Massachusetts_Institute_of_Technology,
pem_pos=0, final_scores=-0.052566916,
final_local_scores=-0.08279834, pem=0.0,
similarity=1.6188453]
4: 400451 PGM-17_Thor
[winner: 26288569 Thor_(Marvel_Comics), pem_pos=1,
final_scores=-0.17348614, final_local_scores=-0.120692715,
pem=-0.97401845, similarity=5.7848654]
gt: 400451 PGM-17_Thor, pem_pos=3,
final_scores=-0.44415474, final_local_scores=-0.491556,
pem=-3.1763976, similarity=6.694446]
5: 29398 Single-stage-to-orbit
[winner: 29398 Single-stage-to-orbit, pem_pos=0,
final_scores=-0.29857433, final_local_scores=-0.35697109,
pem=0.0, similarity=-6.563016]
gt: 29398 Single-stage-to-orbit, pem_pos=0,
final_scores=-0.29857433, final_local_scores=-0.35697109,
pem=0.0, similarity=-6.563016]
6: 154809 McDonnell_Douglas
[winner: 154809 McDonnell_Douglas, pem_pos=0,
final_scores=0.10896815, final_local_scores=0.052853554,
pem=0.0, similarity=5.6669683]
gt: 154809 McDonnell_Douglas, pem_pos=0,
final_scores=0.10896815, final_local_scores=0.052853554,
pem=0.0, similarity=5.6669683]
7: 585987 McDonnell_Douglas_DC-X
[winner: 585987 McDonnell_Douglas_DC-X, pem_pos=0,
final_scores=0.0047911853, final_local_scores=-0.04116392,
pem=0.0, similarity=2.8612995]
gt: 585987 McDonnell_Douglas_DC-X, pem_pos=0,
final_scores=0.0047911853, final_local_scores=-0.04116392,
pem=0.0, similarity=2.8612995]
8: 554021 Tau_Beta_Pi
[winner: 554021 Tau_Beta_Pi, pem_pos=0,
final_scores=0.009976909, final_local_scores=0.04427834,
pem=0.0, similarity=5.411067]

gt: 554021 Tau_Beta_Pi, pem_pos=0,
 final_scores=0.009976909, final_local_scores=0.04427834,
 pem=0.0, similarity=5.411067]
 9: 704658 American_Institute_of_Aeronautics_and_Astronautics
 [winner: 704658
 American_Institute_of_Aeronautics_and_Astronautics,
 pem_pos=0, final_scores=-0.2595001,
 final_local_scores=-0.305314, pem=0.0, similarity=-5.0214663|
 gt: 704658
 American_Institute_of_Aeronautics_and_Astronautics,
 pem_pos=0, final_scores=-0.2595001,
 final_local_scores=-0.305314, pem=0.0, similarity=-5.0214663]
 10: 2828383 American_Astronautical_Society
 [winner: 2828383 American_Astronautical_Society,
 pem_pos=0, final_scores=0.013109252,
 final_local_scores=0.026550427, pem=0.0,
 similarity=4.8820305|
 gt: 2828383 American_Astronautical_Society, pem_pos=0,
 final_scores=0.013109252, final_local_scores=0.026550427,
 pem=0.0, similarity=4.8820305]
 11: 5640397 International_Academy_of_Astronautics
 [winner: 5640397 International_Academy_of_Astronautics,
 pem_pos=0, final_scores=-0.086958945,
 final_local_scores=-0.07465698, pem=0.0,
 similarity=1.8617997|
 gt: 5640397 International_Academy_of_Astronautics,
 pem_pos=0, final_scores=-0.086958945,
 final_local_scores=-0.07465698, pem=0.0,
 similarity=1.8617997]
 12: 2062926 Space_Frontier_Foundation
 [winner: 2062926 Space_Frontier_Foundation, pem_pos=0,
 final_scores=-0.11173053, final_local_scores=-0.11221493,
 pem=0.0, similarity=0.7409955|
 gt: 2062926 Space_Frontier_Foundation, pem_pos=0,
 final_scores=-0.11173053, final_local_scores=-0.11221493,
 pem=0.0, similarity=0.7409955]
 1: 29186 Strategic_Defense_Initiative
 [pem_pos=0, final_scores=0.24814142,
 final_local_scores=0.22458595, pem=0.0, similarity=10.791805]
 2: 317304 National_Space_Society
 [pem_pos=0, final_scores=0.23038004,
 final_local_scores=0.25719616, pem=0.0, similarity=11.764959]
 3: 154809 McDonnell_Douglas
 [pem_pos=0, final_scores=0.06262033,
 final_local_scores=-0.00045348704, pem=0.0,
 similarity=4.07618]
 4: 50042 Lockheed_Corporation
 [pem_pos=0, final_scores=0.25236833,
 final_local_scores=0.20186329, pem=0.0, similarity=10.113716]
 5: 497478 Martin_Marietta
 [pem_pos=0, final_scores=0.35123378,
 final_local_scores=0.33922842, pem=0.0, similarity=14.212966]

6: 29186 Strategic_Defense_Initiative
[pem_pos=0, final_scores=0.23114593,
final_local_scores=0.20602095, pem=0.0, similarity=10.237789]
7: 2087022 Daniel_O._Graham
[pem_pos=0, final_scores=0.27906364,
final_local_scores=0.28524002, pem=0.0, similarity=12.601844]
8: 328602 Ansari_X_Prize
[pem_pos=0, final_scores=0.24236558,
final_local_scores=0.25589597, pem=0.0, similarity=11.726159]
9: 166504 Phi_Beta_Kappa_Society
[pem_pos=0, final_scores=0.11930172,
final_local_scores=0.1747134, pem=0.0, similarity=9.30351]
10: 1462668 British_Interplanetary_Society
[pem_pos=0, final_scores=0.09772368,
final_local_scores=0.115804434, pem=0.0, similarity=7.5455484]
11: 317304 National_Space_Society
[pem_pos=0, final_scores=0.20697401,
final_local_scores=0.23162863, pem=0.0, similarity=11.001973]
1: 3434750 United_States
[pem_pos=0, final_scores=0.5522118,
final_local_scores=0.61663646, pem=-0.12994486,
similarity=23.198036]
2: 559959 Douglas_XB-42_Mixmaster
[pem_pos=0, final_scores=0.09151189,
final_local_scores=0.13221505, pem=0.0, similarity=8.035274]
3: 6337344 Douglas_XB-43_Jetmaster
[pem_pos=0, final_scores=0.1487041,
final_local_scores=0.18733811, pem=0.0, similarity=9.680256]
4: 1097486 MGR-1_Honest_John
[pem_pos=0, final_scores=0.057858214,
final_local_scores=0.085760325, pem=0.0, similarity=6.648972]
5: 400451 PGM-17_Thor
[pem_pos=0, final_scores=0.102694824,
final_local_scores=0.1108066, pem=0.0, similarity=7.396403]
6: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.20683913,
final_local_scores=0.16401538, pem=0.0, similarity=8.984259]
7: 131435 Hollidaysburg,_Pennsylvania
[pem_pos=0, final_scores=0.2699653,
final_local_scores=0.37740403, pem=0.0, similarity=15.352202]
8: 23332 Pennsylvania
[pem_pos=0, final_scores=0.5062311,
final_local_scores=0.5848828, pem=-0.11810702,
similarity=22.186066]
9: 18879 Massachusetts_Institute_of_Technology
[pem_pos=0, final_scores=0.15273513,
final_local_scores=0.14581901, pem=-0.16236198,
similarity=9.324192]
10: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.16235141,
final_local_scores=0.1154193, pem=0.0, similarity=7.5340548]
11: 201793 Douglas_Aircraft_Company

[pem_pos=0, final_scores=0.32364708,
final_local_scores=0.3003191, pem=0.0, similarity=13.051834]
12: 6337344 Douglas_XB-43_Jetmaster
[pem_pos=0, final_scores=0.090889856,
final_local_scores=0.12418482, pem=0.0, similarity=7.7956357]
13: 201793 Douglas_Aircraft_Company
[pem_pos=0, final_scores=0.2559479,
final_local_scores=0.22636807, pem=0.0, similarity=10.844988]
14: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.124974325,
final_local_scores=0.07459052, pem=0.0, similarity=6.315643]
15: 9228 Earth
[pem_pos=0, final_scores=0.0875064,
final_local_scores=0.15477407, pem=-0.07783105,
similarity=9.131738]
16: 201793 Douglas_Aircraft_Company
[pem_pos=0, final_scores=0.34653223,
final_local_scores=0.32531768, pem=0.0, similarity=13.797841]
17: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.08020775,
final_local_scores=0.025689855, pem=0.0, similarity=4.8563495]
18: 3434750 United_States
[pem_pos=0, final_scores=0.4848305,
final_local_scores=0.5430326, pem=-0.14075686,
similarity=21.060345]
19: 25687104 Lockheed_Missiles_and_Space_Company
[pem_pos=0, final_scores=0.10187121,
final_local_scores=0.13277903, pem=0.0, similarity=8.052104]
20: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.15818433,
final_local_scores=0.110867426, pem=0.0, similarity=7.398218]
21: 29186 Strategic_Defense_Initiative
[pem_pos=1, final_scores=0.14243923,
final_local_scores=0.10912235, pem=-0.89878803,
similarity=12.233891]
22: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.11845599,
final_local_scores=0.06747022, pem=0.0, similarity=6.103159]
23: 25687104 Lockheed_Missiles_and_Space_Company
[pem_pos=0, final_scores=0.058643147,
final_local_scores=0.08555892, pem=0.0, similarity=6.642962]
24: 25687104 Lockheed_Missiles_and_Space_Company
[pem_pos=0, final_scores=0.050567195,
final_local_scores=0.07673718, pem=0.0, similarity=6.3797035]
25: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.15387468,
final_local_scores=0.10615976, pem=0.0, similarity=7.257732]
26: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.11215596,
final_local_scores=0.06058839, pem=0.0, similarity=5.8977914]
27: 29186 Strategic_Defense_Initiative
[pem_pos=1, final_scores=0.12379293,

final_local_scores=0.0887541, pem=-0.89878803,
similarity=11.626061]
28: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.15010361,
final_local_scores=0.102040455, pem=0.0, similarity=7.134804]
29: 7279897 United_States_Department_of_Defense
[pem_pos=0, final_scores=0.16803978,
final_local_scores=0.14446515, pem=0.0, similarity=8.400841]
30: 585987 McDonnell_Douglas_DC-X
[pem_pos=0, final_scores=0.1877362,
final_local_scores=0.16225144, pem=0.0, similarity=8.93162]
31: 154809 McDonnell_Douglas
[pem_pos=0, final_scores=0.10896815,
final_local_scores=0.052853554, pem=0.0, similarity=5.6669683]
32: 585987 McDonnell_Douglas_DC-X
[pem_pos=0, final_scores=0.1968439,
final_local_scores=0.17220026, pem=0.0, similarity=9.228512]
33: 585987 McDonnell_Douglas_DC-X
[pem_pos=0, final_scores=0.06442572,
final_local_scores=0.027553484, pem=0.0, similarity=4.911964]
34: 18426568 NASA
[pem_pos=0, final_scores=0.15435313,
final_local_scores=0.1068887, pem=0.0, similarity=7.2794847]
35: 585987 McDonnell_Douglas_DC-X
[pem_pos=0, final_scores=0.15593575,
final_local_scores=0.12751427, pem=0.0, similarity=7.8949933]
36: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.2017213,
final_local_scores=0.15842491, pem=0.0, similarity=8.817429]
37: 11594875 Stanford_Hospital_and_Clinics
[pem_pos=0, final_scores=0.053925887,
final_local_scores=0.14126289, pem=-0.004618924,
similarity=8.330398]
38: 26977 Stanford_University
[pem_pos=0, final_scores=0.12409286,
final_local_scores=0.14446622, pem=-0.6528699,
similarity=11.951282]
39: 5407 California
[pem_pos=0, final_scores=0.49495763,
final_local_scores=0.56110805, pem=-0.08662026,
similarity=21.305351]
40: 5926342 Maxwell_Hunter
[pem_pos=0, final_scores=0.09917377,
final_local_scores=0.046407342, pem=0.0, similarity=5.474601]
41: 22153223 NASA_Exceptional_Public_Service_Medal
[pem_pos=0, final_scores=0.034760162,
final_local_scores=0.14073372, pem=0.0, similarity=8.289488]
42: 28189 Space_Shuttle
[pem_pos=0, final_scores=0.17472242,
final_local_scores=0.1443237, pem=-0.043386377,
similarity=8.632563]

Sample: b'Jacques_Thibaud'

Jacques Thibaud -LRB- 27 September 1880 1 September 1953 - RRB- was a [1 1 French 1] 1 violin ist . [1 Thibaud 1] was born in [2 2 Bordeaux 2] and studied the violin with his father before entering the at the age of thirteen . In 1896 he jointly won the conservatory 's violin prize with [3 3 Pierre Monteux 3] -LRB- who later became a famous 2 conductor - RRB- . He had to rebuild his technique after being injured in [4 4 World War I 4] . In 1943 he and [5 5 Marguerite Long 5] established the [6 6 Marguerite Long-Jacques Thibaud International Competition 6] for violinists and pianists , which takes place each year in [7 7 Paris 7] . From 2011 , it has included singers and is now known as the [2 Long-Thibaud-Crespin Competition 2] , in honour of the soprano [8 8 Régine Crespin 8] . [3 Thibaud 3] was noted not only for his work as a soloist , but also for his performances of 3 chamber music , particularly in a 4 piano trio with the pianist [9 9 Alfred Cortot 9] and cellist [10 10 Pablo Casals 10] . He undertook concert tours with pianist [11 11 Yves Nat 11] and [12 12 George Enescu 12] . He was a friend of violinist [13 13 Eugène Ysaÿe 13] , who dedicated his 5 2nd [4 Sonata 4] for solo violin to him . Among his students were 6 Manuel Quiroga -LRB- the dedicatee of [5 Ysaÿe 5] 's 6th solo sonata -RRB- , [14 14 Eric Rosenblith 14] and 7 Joan Field . On 1 September 1953 , [6 Thibaud 6] and 41 others died in the crash of an 8 Air France 9 Lockheed Constellation airliner . The aircraft , registration F-BAZZ , was on final approach to when it crashed into Mt Cimet in the [7 10 French 7] [8 Alps 8] . His 1720 [15 15 Stradivarius 15] violin was destroyed in the crash . The accident investigation established ' ' 11 controlled flight into terrain -LRB- CFIT -RRB-

1: 5843419 France

2: 4097 Bordeaux

3: 163531 Pierre_Monteux

4: 4764461 World_War_I

5: 2691394 Marguerite_Long

6: 31234889 Long-Thibaud-Crespin_Competition

7: 22989 Paris

8: 2878369 Régine_Crespin

9: 312802 Alfred_Cortot

10: 49073 Pablo_Casals

11: 10097232 Yves_Nat

12: 319074 George_Enescu

13: 274754 Eugène_Ysaÿe

14: 3870001 Eric_Rosenblith

15: 153773 Stradivarius

1: 32427 Violin

[winner: 2842544 Violin_Concerto_(Dvořák), pem_pos=8, final_scores=-1.2937081, final_local_scores=-1.3153017, pem=-2.0790691, similarity=-23.855211]

gt: 32427 Violin, pem_pos=0, final_scores=-1.3290414, final_local_scores=-1.4980705, pem=-0.020752236,

similarity=-40.502846]
 2: 199162 Conducting
 [winner: 18847878 Conductor_of_an_abelian_variety,
 pem_pos=4, final_scores=-1.2019515,
 final_local_scores=-1.1232497, pem=-1.6088006,
 similarity=-20.681389|
 gt: 199162 Conducting, pem_pos=0,
 final_scores=-1.5630652, final_local_scores=-1.7599323,
 pem=-0.27369234, similarity=-46.941795]
 3: 82991 Chamber_music
 [winner: 5489867 Mycenaean_chamber_tomb, pem_pos=6,
 final_scores=-1.0293487, final_local_scores=-0.9217172,
 pem=-1.9459101, similarity=-12.834004|
 gt: 82991 Chamber_music, pem_pos=0,
 final_scores=-1.1964636, final_local_scores=-1.3673873,
 pem=-0.28337866, similarity=-35.174793]
 [winner: 22515916 .music, pem_pos=1,
 final_scores=-1.2247603, final_local_scores=-1.1900988,
 pem=-1.9571527, similarity=-20.781908|
 RECALL MISS: 82991 Chamber_music]
 4: 174981 Piano_trio
 [winner: 629285 Pianos, pem_pos=1,
 final_scores=-0.93189186, final_local_scores=-0.9231693,
 pem=-1.3240921, similarity=-16.258883|
 RECALL MISS: 174981 Piano_trio]
 [winner: 166508 String_trio, pem_pos=4,
 final_scores=-1.3091062, final_local_scores=-1.4047866,
 pem=-1.7902174, similarity=-28.096443|
 gt: 174981 Piano_trio, pem_pos=3,
 final_scores=-1.3780315, final_local_scores=-1.5228212,
 pem=-1.7598661, similarity=-31.78388]
 5: 30010630 Six_Sonatas_for_solo_violin_(Ysaÿe)
 [winner: 44118 Sonata, pem_pos=0,
 final_scores=0.07215868, final_local_scores=0.050976127,
 pem=-0.28629988, similarity=7.167886|
 RECALL MISS: 30010630
 Six_Sonatas_for_solo_violin_(Ysaÿe)]
 [winner: 32427 Violin, pem_pos=0,
 final_scores=-0.96271855, final_local_scores=-1.0979185,
 pem=-0.020752236, similarity=-28.561512|
 RECALL MISS: 30010630
 Six_Sonatas_for_solo_violin_(Ysaÿe)]
 [winner: 377234 Solo_(music), pem_pos=0,
 final_scores=-1.2117146, final_local_scores=-1.3488433,
 pem=-0.01683729, similarity=-36.0709|
 RECALL MISS: 30010630
 Six_Sonatas_for_solo_violin_(Ysaÿe)]
 [winner: 6307967
 2nd_Guards_Infantry_Division_(German_Empire), pem_pos=16,
 final_scores=-1.3960744, final_local_scores=-1.3556144,
 pem=-3.9846575, similarity=-14.695337|
 RECALL MISS: 30010630

Six_Sonatas_for_solo_violin_(Ysaÿe)]
6: 38180306 Manuel_Quiroga_(violinist)
[winner: 38180306 Manuel_Quiroga_(violinist), pem_pos=1,
final_scores=-0.011149585, final_local_scores=-0.047168374,
pem=-1.3609765, similarity=10.083316]
gt: 38180306 Manuel_Quiroga_(violinist), pem_pos=1,
final_scores=-0.011149585, final_local_scores=-0.047168374,
pem=-1.3609765, similarity=10.083316]
7: 30832797 Joan_Field
[winner: 30832797 Joan_Field, pem_pos=0,
final_scores=0.000892207, final_local_scores=0.011143699,
pem=0.0, similarity=4.422263]
gt: 30832797 Joan_Field, pem_pos=0,
final_scores=0.000892207, final_local_scores=0.011143699,
pem=0.0, similarity=4.422263]
8: 14998012 Air_France
[winner: 14998012 Air_France, pem_pos=0,
final_scores=-0.039499715, final_local_scores=0.018102497,
pem=0.0, similarity=4.6299276]
gt: 14998012 Air_France, pem_pos=0,
final_scores=-0.039499715, final_local_scores=0.018102497,
pem=0.0, similarity=4.6299276]
9: 274115 Lockheed_Constellation
[winner: 274115 Lockheed_Constellation, pem_pos=0,
final_scores=-0.67807776, final_local_scores=-0.6397464,
pem=-0.018028632, similarity=-14.903553]
gt: 274115 Lockheed_Constellation, pem_pos=0,
final_scores=-0.67807776, final_local_scores=-0.6397464,
pem=-0.018028632, similarity=-14.903553]
10: 961317 French_Alps
[winner: 5843419 France, pem_pos=0,
final_scores=0.5423774, final_local_scores=0.57937574,
pem=-0.80223274, similarity=25.742107]
RECALL MISS: 961317 French_Alps]
[winner: 981 Alps, pem_pos=0, final_scores=0.27088648,
final_local_scores=0.35478184, pem=-0.014871542,
similarity=14.757984]
gt: 961317 French_Alps, pem_pos=11,
final_scores=-1.0533226, final_local_scores=-1.111453,
pem=-7.3471923, similarity=10.876911]
11: 490866 Controlled_flight_into_terrain
[winner: 490866 Controlled_flight_into_terrain,
pem_pos=0, final_scores=-0.33487105,
final_local_scores=-0.2155495, pem=0.0, similarity=-2.342714]
gt: 490866 Controlled_flight_into_terrain, pem_pos=0,
final_scores=-0.33487105, final_local_scores=-0.2155495,
pem=0.0, similarity=-2.342714]
[winner: 1640288 Scientific_control, pem_pos=0,
final_scores=-1.1225402, final_local_scores=-1.0723557,
pem=0.0, similarity=-27.911524]
RECALL MISS: 490866 Controlled_flight_into_terrain]
[winner: 586839 Terrain, pem_pos=0,

final_scores=-1.1539011, final_local_scores=-1.0862843,
pem=-0.021533424, similarity=-28.210075]
RECALL MISS: 490866 Controlled_flight_into_terrain]
[winner: 24064732 Flight_(Lessing_story), pem_pos=5,
final_scores=-1.2503345, final_local_scores=-1.1598032,
pem=-2.9761064, similarity=-14.336598]
RECALL MISS: 490866 Controlled_flight_into_terrain]
1: 5843419 France
[pem_pos=0, final_scores=0.7345941,
final_local_scores=0.7893433, pem=-0.80223274,
similarity=32.007957]
2: 4097 Bordeaux
[pem_pos=0, final_scores=0.40584487,
final_local_scores=0.41597885, pem=-0.40822297,
similarity=18.723331]
3: 163531 Pierre_Monteux
[pem_pos=0, final_scores=0.089029506,
final_local_scores=0.05227834, pem=0.0, similarity=5.6498027]
4: 4764461 World_War_I
[pem_pos=0, final_scores=0.3244241,
final_local_scores=0.3812763, pem=0.0, similarity=15.467758]
5: 2691394 Marguerite_Long
[pem_pos=0, final_scores=0.3456567,
final_local_scores=0.3372358, pem=0.0, similarity=14.153502]
6: 31234889 Long-Thibaud-Crespin_Competition
[pem_pos=0, final_scores=0.20873968,
final_local_scores=0.22607923, pem=0.0, similarity=10.836368]
7: 22989 Paris
[pem_pos=0, final_scores=0.64034426,
final_local_scores=0.64022344, pem=-0.115301244,
similarity=23.822283]
8: 2878369 Régine_Crespin
[pem_pos=0, final_scores=0.4352677,
final_local_scores=0.46368003, pem=0.0, similarity=17.92685]
9: 312802 Alfred_Cortot
[pem_pos=0, final_scores=0.64817286,
final_local_scores=0.6474892, pem=0.0, similarity=23.412083]
10: 49073 Pablo_Casals
[pem_pos=0, final_scores=0.3233143,
final_local_scores=0.28352606, pem=0.0, similarity=12.550696]
11: 10097232 Yves_Nat
[pem_pos=0, final_scores=0.56504774,
final_local_scores=0.60363424, pem=0.0, similarity=22.103363]
12: 319074 George_Enescu
[pem_pos=0, final_scores=0.33925188,
final_local_scores=0.31144553, pem=0.0, similarity=13.383869]
13: 274754 Eugène_Ysaÿe
[pem_pos=0, final_scores=0.6125043,
final_local_scores=0.60757, pem=0.0, similarity=22.220814]
14: 3870001 Eric_Rosenblith
[pem_pos=0, final_scores=0.087724134,
final_local_scores=0.15760484, pem=0.0, similarity=8.792955]

15: 153773 Stradivarius
 [pem_pos=0, final_scores=0.14913307,
 final_local_scores=0.14307591, pem=0.0, similarity=8.359384]
 1: 312283 Jacques_Thibaud
 [pem_pos=0, final_scores=0.08794318,
 final_local_scores=0.04554014, pem=0.0, similarity=5.4487214]
 2: 31234889 Long-Thibaud-Crespin_Competition
 [pem_pos=0, final_scores=0.10546668,
 final_local_scores=0.11326915, pem=0.0, similarity=7.4698906]
 3: 312283 Jacques_Thibaud
 [pem_pos=0, final_scores=0.13253503,
 final_local_scores=0.094249964, pem=0.0, similarity=6.90232]
 4: 44118 Sonata
 [pem_pos=0, final_scores=0.07215868,
 final_local_scores=0.050976127, pem=-0.28629988,
 similarity=7.167886]
 5: 274754 Eugène_Ysaÿe
 [pem_pos=0, final_scores=0.22106145,
 final_local_scores=0.18652594, pem=0.0, similarity=9.656019]
 6: 312283 Jacques_Thibaud
 [pem_pos=0, final_scores=0.07313581,
 final_local_scores=0.029365331, pem=0.0, similarity=4.966033]
 7: 5843419 France
 [pem_pos=0, final_scores=0.5423774,
 final_local_scores=0.57937574, pem=-0.80223274,
 similarity=25.742107]
 8: 981 Alps
 [pem_pos=0, final_scores=0.27088648,
 final_local_scores=0.35478184, pem=-0.014871542,
 similarity=14.757984]

Sample: b'Nucleoid'

The nucleoid -LRB- meaning **1 nucleus** - like -RRB- is an irregularly-shaped region within the **2 cell** of a **3 prokaryote** that contains all or most of the **4 genetic material** . In contrast to the **5 nucleus** of a **6 eukaryotic** cell , it is not surrounded by a **7 nuclear membrane** . The **8 genome** of prokaryotic organisms generally is a circular , double-stranded piece of **9 DNA** , of which multiple copies may exist at any time . The length of a genome widely varies , but generally is at least a few million base pairs . A genophore is the DNA of a prokaryote . It is commonly referred to as a prokaryotic **10 chromosome** . The term `` chromosome '' is misleading for a genophore because the genophore lacks **11 chromatin** . The genophore is compacted through a mechanism known as , whereas a chromosome is additionally compacted via chromatin . The genophore is circular in most prokaryotes , and linear in very few . The circular nature of the genophore allows replication to occur without . Genophores are generally of a much smaller size than Eukaryotic chromosomes . A genophore can be as small as 580,073 base pairs -LRB- **12 Mycoplasma genitalium** -RRB- . Many eukaryotes -

LRB- such as plants and animals -RRB- carry genophores in organelles such as and . These organelles are very similar to true prokaryotes . The nucleoid can be clearly visualized on an at high **13 magnification** , where , although its appearance may differ , it is clearly visible against the **14 cytosol** . Sometimes even strands of what is thought to be DNA are visible . By staining with the **15 Feulgen stain** , which specifically stains DNA , the nucleoid can also be seen under a light microscope . The DNA-intercalating stains **16 DAPI** and **17 ethidium bromide** are widely used for of nucleoids . Experimental evidence suggests that the nucleoid is largely composed of DNA , about 60 % , with a small amount of **18 RNA** and **19 protein** . The latter two constituents are likely to be mainly **20 messenger RNA** and the **21 transcription factor** proteins found regulating the bacterial genome . Proteins helping to maintain the supercoiled structure of the nucleic acid are known as nucleoid proteins or nucleoid-associated proteins and are distinct from **22 histone**

1: 6235 Cell_nucleus

[winner: 6235 Cell_nucleus, pem_pos=0, final_scores=-0.9535731, final_local_scores=-1.0712276, pem=-0.37133834, similarity=-25.858461]

gt: 6235 Cell_nucleus, pem_pos=0, final_scores=-0.9535731, final_local_scores=-1.0712276, pem=-0.37133834, similarity=-25.858461]

2: 4230 Cell_(biology)

[winner: 4230 Cell_(biology), pem_pos=0, final_scores=-1.118193, final_local_scores=-1.254796, pem=-0.15427125, similarity=-32.516953]

gt: 4230 Cell_(biology), pem_pos=0, final_scores=-1.118193, final_local_scores=-1.254796, pem=-0.15427125, similarity=-32.516953]

3: 19172225 Prokaryote

[winner: 19172225 Prokaryote, pem_pos=0, final_scores=-1.0065476, final_local_scores=-1.1061518, pem=0.0, similarity=-28.920063]

gt: 19172225 Prokaryote, pem_pos=0, final_scores=-1.0065476, final_local_scores=-1.1061518, pem=0.0, similarity=-28.920063]

4: 12388 Genome

[winner: 1558261 Inorganic_compounds_by_element, pem_pos=9, final_scores=-0.91290563, final_local_scores=-0.85659766, pem=-2.5714664, similarity=-7.4888353]

RECALL MISS: 12388 Genome]

[winner: 40254 Genetic_algorithm, pem_pos=2, final_scores=-1.1814897, final_local_scores=-1.2260132, pem=-1.315508, similarity=-25.343029]

RECALL MISS: 12388 Genome]

5: 6235 Cell_nucleus

[winner: 6235 Cell_nucleus, pem_pos=0, final_scores=-1.0583339, final_local_scores=-1.1856627,


```

pem=-0.37133834, similarity=-29.27343|
    gt: 6235 Cell_nucleus, pem_pos=0,
    final_scores=-1.0583339, final_local_scores=-1.1856627,
pem=-0.37133834, similarity=-29.27343 ]
6: 24536543 Eukaryote
    [winner: 24536543 Eukaryote, pem_pos=0,
    final_scores=-0.7460351, final_local_scores=-0.8453643,
pem=-0.060584784, similarity=-20.808174|
    gt: 24536543 Eukaryote, pem_pos=0,
    final_scores=-0.7460351, final_local_scores=-0.8453643,
pem=-0.060584784, similarity=-20.808174 ]
7: 12840258 Nuclear_membrane
    [winner: 12840258 Nuclear_membrane, pem_pos=0,
    final_scores=-0.7544929, final_local_scores=-0.8028456,
pem=0.0, similarity=-19.868801|
    gt: 12840258 Nuclear_membrane, pem_pos=0,
    final_scores=-0.7544929, final_local_scores=-0.8028456,
pem=0.0, similarity=-19.868801 ]
8: 12388 Genome
    [winner: 12388 Genome, pem_pos=0,
    final_scores=-1.2919182, final_local_scores=-1.4476228,
pem=-0.026028708, similarity=-38.968693|
    gt: 12388 Genome, pem_pos=0, final_scores=-1.2919182,
    final_local_scores=-1.4476228, pem=-0.026028708,
    similarity=-38.968693 ]
9: 7955 DNA
    [winner: 7955 DNA, pem_pos=0, final_scores=-0.16765504,
    final_local_scores=-0.23612776, pem=-0.054993372,
    similarity=-2.657748|
    gt: 7955 DNA, pem_pos=0, final_scores=-0.16765504,
    final_local_scores=-0.23612776, pem=-0.054993372,
    similarity=-2.657748 ]
10: 6438 Chromosome
    [winner: 49399 XY_sex-determination_system, pem_pos=0,
    final_scores=-1.1834755, final_local_scores=-1.2256792,
pem=0.0, similarity=-32.487|
    gt: 6438 Chromosome, pem_pos=1, final_scores=-1.3141614,
    final_local_scores=-1.4552429, pem=-0.023342757,
    similarity=-39.210697 ]
11: 6933 Chromatin
    [winner: 6933 Chromatin, pem_pos=0,
    final_scores=-1.0828391, final_local_scores=-1.1708126,
pem=0.0, similarity=-30.849669|
    gt: 6933 Chromatin, pem_pos=0, final_scores=-1.0828391,
    final_local_scores=-1.1708126, pem=0.0,
    similarity=-30.849669 ]
12: 20219 Mycoplasma_genitalium
    [winner: 219939 Mycoplasma, pem_pos=1,
    final_scores=0.010049328, final_local_scores=0.040993884,
pem=-0.09880582, similarity=5.8503737|
    RECALL MISS: 20219 Mycoplasma_genitalium]
13: 603273 Magnification

```

[winner: 3214848 Make_a_mountain_out_of_a_molehill,
 pem_pos=2, final_scores=-0.79291826,
 final_local_scores=-0.7333158, pem=-0.6931472,
 similarity=-14.024451|
 gt: 603273 Magnification, pem_pos=0,
 final_scores=-0.9378018, final_local_scores=-0.904729,
 pem=-0.29243463, similarity=-21.318901]
 14: 6781 Cytosol
 [winner: 6781 Cytosol, pem_pos=0,
 final_scores=-1.0181131, final_local_scores=-1.0925138,
 pem=0.0, similarity=-28.51308|
 gt: 6781 Cytosol, pem_pos=0, final_scores=-1.0181131,
 final_local_scores=-1.0925138, pem=0.0, similarity=-28.51308]
 15: 3248757 Feulgen_stain
 [winner: 3059556 Robert_Feulgen, pem_pos=0,
 final_scores=-0.008159816, final_local_scores=0.12415883,
 pem=0.0, similarity=7.7948604|
 gt: 3248757 Feulgen_stain, pem_pos=1,
 final_scores=-0.17220676, final_local_scores=-0.12523907,
 pem=0.0, similarity=0.35232913]
 [winner: 1750229 Wood_stain, pem_pos=0,
 final_scores=-1.1314045, final_local_scores=-1.1018829,
 pem=0.0, similarity=-28.792673|
 RECALL MISS: 3248757 Feulgen_stain]
 16: 1503867 DAPI
 [winner: 1503867 DAPI, pem_pos=0,
 final_scores=-0.22216913, final_local_scores=-0.209923,
 pem=0.0, similarity=-2.1748083|
 gt: 1503867 DAPI, pem_pos=0, final_scores=-0.22216913,
 final_local_scores=-0.209923, pem=0.0, similarity=-2.1748083]
 17: 363891 Ethidium_bromide
 [winner: 363891 Ethidium_bromide, pem_pos=0,
 final_scores=-0.6853924, final_local_scores=-0.72360516,
 pem=0.0, similarity=-17.504108|
 gt: 363891 Ethidium_bromide, pem_pos=0,
 final_scores=-0.6853924, final_local_scores=-0.72360516,
 pem=0.0, similarity=-17.504108]
 18: 25758 RNA
 [winner: 25758 RNA, pem_pos=0, final_scores=-0.19744536,
 final_local_scores=-0.24720737, pem=-0.051015966,
 similarity=-3.0100157|
 gt: 25758 RNA, pem_pos=0, final_scores=-0.19744536,
 final_local_scores=-0.24720737, pem=-0.051015966,
 similarity=-3.0100157]
 19: 23634 Protein
 [winner: 23634 Protein, pem_pos=0,
 final_scores=-1.1046664, final_local_scores=-1.2285883,
 pem=-0.04669233, similarity=-32.319897|
 gt: 23634 Protein, pem_pos=0, final_scores=-1.1046664,
 final_local_scores=-1.2285883, pem=-0.04669233,
 similarity=-32.319897]
 20: 20232 Messenger_RNA

[winner: 25758 RNA, pem_pos=0, final_scores=-0.116176754, final_local_scores=-0.15843377, pem=-0.051015966, similarity=-0.3608337]
 gt: 20232 Messenger_RNA, pem_pos=9, final_scores=-1.0072325, final_local_scores=-1.107727, pem=-5.3808064, similarity=0.2945893]
 [winner: 21026473 Mercury_Messenger_concept, pem_pos=3, final_scores=-1.1945505, final_local_scores=-1.0552197, pem=-2.399456, similarity=-14.351532]
 RECALL MISS: 20232 Messenger_RNA]
 21: 31474 Transcription_factor
 [winner: 891398 Factor_(programming_language), pem_pos=0, final_scores=-0.977206, final_local_scores=-0.9111708, pem=-0.15148865, similarity=-22.277622]
 RECALL MISS: 31474 Transcription_factor]
 [winner: 167544 Transcription_(genetics), pem_pos=1, final_scores=-1.1784178, final_local_scores=-1.277044, pem=-0.39599004, similarity=-31.866373]
 gt: 31474 Transcription_factor, pem_pos=5, final_scores=-1.9138436, final_local_scores=-2.0705242, pem=-5.0193276, similarity=-30.402971]
 22: 14029 Histone
 [winner: 14029 Histone, pem_pos=0, final_scores=-0.78684676, final_local_scores=-0.85114604, pem=-0.0004760564, similarity=-21.307594]
 gt: 14029 Histone, pem_pos=0, final_scores=-0.78684676, final_local_scores=-0.85114604, pem=-0.0004760564, similarity=-21.307594]

Sample: b'Mondo_2000'

` [1 Mondo 2000 1] ' was a glossy 1 cyberculture 2 magazine published in [1 1 California 1] during the 1980s and 1990s . It covered 3 cyberpunk topics such as 4 virtual reality and 5 smart drugs . It was a more anarchic and subversive prototype for the later-founded magazine . [2 Mondo 2000 2] originated as High Frontiers in 1984 , edited by [2 2 R. U. Sirius 2] -LRB- pseudonym for [3 Ken Goffman 3] -RRB- . He was succeeded as Editor-in-Chief by 6 anthropologist Alison Bailey Kennedy , a.k.a. `` Queen Mu '' and `` Alison Wonderland '' . [4 Sirius 4] was joined by hacker 7 Jude Milhon -LRB- a.k.a. St. Jude -RRB- as editor and the magazine was renamed Reality Hackers in 1988 to better reflect its drugs and computers theme . It changed title again to [5 Mondo 2000 5] in 1989 . Art director and photographer Bart Nagel , a pioneer in 8 collage , created the publication 's elegantly 9 surrealist aesthetic . [6 R. U. Sirius 6] left at the beginning of 1993 , at about the same time as the launch of [7 Wired 7] . The magazine continued until 1998 , with the last issue being # 17 . Along with the print version of [3 3 Boing Boing 3] -- with which [8 Mondo 2000 8] shared several writers , including [4 4 Mark Frauenfelder 4] , [5 5 Richard Kadrey 5] , [6 6 Gareth Branwyn 6] , and [7 7 Jon Lebkowsky 7] -- [9

Mondo 2000 9] helped develop what was to become the 10 cyberpunk 11 subculture . Writers included [8 8 William Gibson 8] , Nan C. Druid -LRB- pseudonym for Maerian Morris -RRB- , [9 9 Rudy Rucker 9] , [10 10 Bruce Sterling 10] , and 12 Robert Anton Wilson

1: 5407 California
2: 709285 R._U._Sirius
3: 984234 Boing_Boing
4: 1198480 Mark_Frauenfelder
5: 8875843 Richard_Kadrey
6: 3441216 Gareth_Branwyn
7: 2827736 Jon_Lebkowsky
8: 33119 William_Gibson
9: 254812 Rudy_Rucker
10: 3838 Bruce_Sterling

1: 429909 Cyberculture

[winner: 429909 Cyberculture, pem_pos=0,
final_scores=-0.9637535, final_local_scores=-1.0099802,
pem=0.0, similarity=-26.05011]

gt: 429909 Cyberculture, pem_pos=0,
final_scores=-0.9637535, final_local_scores=-1.0099802,
pem=0.0, similarity=-26.05011]

2: 21001 Magazine

[winner: 896497 Magazine_(firearms), pem_pos=0,
final_scores=-1.0828108, final_local_scores=-1.0079287,
pem=-0.62660754, similarity=-22.5813]

gt: 21001 Magazine, pem_pos=1, final_scores=-1.1347814,
final_local_scores=-1.2364173, pem=-0.85497606,
similarity=-28.15795]

3: 5703 Cyberpunk

[winner: 5703 Cyberpunk, pem_pos=0,
final_scores=-1.2485579, final_local_scores=-1.382678,
pem=-0.043043267, similarity=-36.93808]

gt: 5703 Cyberpunk, pem_pos=0, final_scores=-1.2485579,
final_local_scores=-1.382678, pem=-0.043043267,
similarity=-36.93808]

4: 32612 Virtual_reality

[winner: 247365 Simulated_reality, pem_pos=0,
final_scores=-0.7109185, final_local_scores=-0.68018305,
pem=0.0, similarity=-16.208305]

gt: 32612 Virtual_reality, pem_pos=1,
final_scores=-0.7499414, final_local_scores=-0.78395504,
pem=-0.0077801016, similarity=-19.26276]

5: 685457 Nootropic

[winner: 685457 Nootropic, pem_pos=0,
final_scores=-0.93499243, final_local_scores=-0.8474126,
pem=-0.2876821, similarity=-19.634308]

gt: 685457 Nootropic, pem_pos=0,
final_scores=-0.93499243, final_local_scores=-0.8474126,
pem=-0.2876821, similarity=-19.634308]

6: 342908 Anthropologist

[winner: 569 Anthropology, pem_pos=0,

final_scores=-1.6602412, final_local_scores=-1.75332,
pem=-0.34495884, similarity=-46.356915|
gt: 342908 Anthropologist, pem_pos=1,
final_scores=-1.8466587, final_local_scores=-1.9441974,
pem=-1.264204, similarity=-47.054077]
7: 299670 Jude_Milhon
[winner: 299670 Jude_Milhon, pem_pos=0,
final_scores=-0.014975518, final_local_scores=0.064181834,
pem=0.0, similarity=6.005027|
gt: 299670 Jude_Milhon, pem_pos=0,
final_scores=-0.014975518, final_local_scores=0.064181834,
pem=0.0, similarity=6.005027]
8: 18938195 Collage
[winner: 18938195 Collage, pem_pos=1,
final_scores=-1.0486835, final_local_scores=-1.1007128,
pem=-0.0072125997, similarity=-28.71853|
gt: 18938195 Collage, pem_pos=1,
final_scores=-1.0486835, final_local_scores=-1.1007128,
pem=-0.0072125997, similarity=-28.71853]
9: 28766 Surrealism
[winner: 1592244 Surreal_humour, pem_pos=0,
final_scores=-1.3189952, final_local_scores=-1.4042611,
pem=0.0, similarity=-37.816242|
gt: 28766 Surrealism, pem_pos=1,
final_scores=-1.3521167, final_local_scores=-1.42992,
pem=-0.013966907, similarity=-38.505993]
10: 5703 Cyberpunk
[winner: 5703 Cyberpunk, pem_pos=0,
final_scores=-1.2280107, final_local_scores=-1.3602333,
pem=-0.043043267, similarity=-36.268288|
gt: 5703 Cyberpunk, pem_pos=0, final_scores=-1.2280107,
final_local_scores=-1.3602333, pem=-0.043043267,
similarity=-36.268288]
11: 150349 Subculture
[winner: 150349 Subculture, pem_pos=0,
final_scores=-1.3514615, final_local_scores=-1.4372263,
pem=-0.093592905, similarity=-38.291016|
gt: 150349 Subculture, pem_pos=0,
final_scores=-1.3514615, final_local_scores=-1.4372263,
pem=-0.093592905, similarity=-38.291016]
12: 26056 Robert_Anton_Wilson
[winner: 26056 Robert_Anton_Wilson, pem_pos=0,
final_scores=-0.04914759, final_local_scores=-0.060403496,
pem=0.0, similarity=2.287152|
gt: 26056 Robert_Anton_Wilson, pem_pos=0,
final_scores=-0.04914759, final_local_scores=-0.060403496,
pem=0.0, similarity=2.287152]
1: 5407 California
[pem_pos=0, final_scores=0.47599053,
final_local_scores=0.6000925, pem=-0.08662026,
similarity=22.468725]
2: 709285 R._U._Sirius

```
[pem_pos=0, final_scores=0.12087633,
final_local_scores=0.076443285, pem=0.0, similarity=6.370933]
3: 984234 Boing_Boing
[pem_pos=0, final_scores=0.048080787,
final_local_scores=0.018327698, pem=0.0, similarity=4.636648]
4: 1198480 Mark_Frauenfelder
[pem_pos=0, final_scores=0.1923192,
final_local_scores=0.22372428, pem=0.0, similarity=10.766091]
5: 8875843 Richard_Kadrey
[pem_pos=0, final_scores=0.06988697,
final_local_scores=0.09615509, pem=0.0, similarity=6.9591727]
6: 3441216 Gareth_Branwyn
[pem_pos=0, final_scores=0.2207499,
final_local_scores=0.2570673, pem=0.0, similarity=11.761114]
7: 2827736 Jon_Lebkowsky
[pem_pos=0, final_scores=0.1851577,
final_local_scores=0.18868712, pem=0.0, similarity=9.720512]
8: 33119 William_Gibson
[pem_pos=0, final_scores=0.07160632,
final_local_scores=0.058289975, pem=-0.0041210214,
similarity=5.8516126]
9: 254812 Rudy_Rucker
[pem_pos=0, final_scores=0.1264671,
final_local_scores=0.15624598, pem=0.0, similarity=8.752404]
10: 3838 Bruce_Sterling
[pem_pos=0, final_scores=0.11463414,
final_local_scores=0.11412862, pem=0.0, similarity=7.495538]
1: 1108228 Mondo_2000
[pem_pos=0, final_scores=0.15157051,
final_local_scores=0.09234324, pem=0.0, similarity=6.8454194]
2: 1108228 Mondo_2000
[pem_pos=0, final_scores=0.14622335,
final_local_scores=0.08650227, pem=0.0, similarity=6.671113]
3: 709285 R._U._Sirius
[pem_pos=0, final_scores=0.16222863,
final_local_scores=0.1216144, pem=0.0, similarity=7.718929]
4: 709285 R._U._Sirius
[pem_pos=0, final_scores=0.1215774,
final_local_scores=0.077209115, pem=0.0, similarity=6.393787]
5: 1108228 Mondo_2000
[pem_pos=0, final_scores=0.06718658,
final_local_scores=0.00016662478, pem=0.0,
similarity=4.0946856]
6: 709285 R._U._Sirius
[pem_pos=0, final_scores=0.16151334,
final_local_scores=0.12083307, pem=0.0, similarity=7.695613]
7: 65411 Wired_(magazine)
[pem_pos=0, final_scores=0.38137037,
final_local_scores=0.3680312, pem=0.0, similarity=15.072498]
8: 1108228 Mondo_2000
[pem_pos=0, final_scores=0.16840951,
final_local_scores=0.11073728, pem=0.0, similarity=7.3943343]
```

9: 1108228 Mondo_2000

[pem_pos=0, final_scores=0.11927767,
final_local_scores=0.05706823, pem=0.0, similarity=5.7927427]

Sample: b'Omicron2_Canis_Majoris'

Omicron2 Canis Majoris -LRB- o ² CMa , o ² Canis Majoris -RRB- is a 1 star in the 2 constellation [1 1 Canis Major 1] . Since 1943 , the 3 spectrum of this star has served as one of the stable anchor points by which other stars are classified . It has an of 3.043 , making it one of the brighter members of the constellation . The distance to this star is roughly 2,800 s -LRB- 800 4 parsec s -RRB- , with a 34 % 5 margin of error . This is a massive 6 supergiant star with a 7 stellar classification of B3 Ia , indicating that , at the age of around 7 million years , it has exhausted the supply of hydrogen at its core and is now undergoing 8 nuclear fusion of helium to generate energy . It has about 21 times the mass of the Sun and 65 times the Sun 's radius . In all likelihood , it will end its life as a 9 Type II supernova . Omicron2 Canis Majoris is one of the known , as it radiates about 220,000 times as much luminosity as the [1 Sun 1] from its 10 outer envelope at a temperature of 15,500 K . At this heat , the star is glowing with the blue-white hue of a B-type star . This star is classified as an 11 Alpha Cygni-type 12 variable star that undergoes periodic non-radial pulsations , which cause its brightness to cycle from magnitude +2.93 to +3.08 over a 24.44 day interval . It is losing mass from its 13 stellar wind at the rate of around times the mass of the [2 Sun 2] per year , or the equivalent of the [3 Sun 3] 's mass every 500 million years . While this star lies in the field of view of the 14 open cluster named Collinder 121 , it is unlikely to be a member . In fact , its optical neighbor , o1 [4 Canis Majoris 4] has a much higher likelihood of 23.1 % based upon its 15 proper motion being a closer match to the motion of the cluster . Although they are located near each other on the 16 celestial sphere , o1 CMa and o2 CMa are most likely not to each other as they appear to lie many light years apart . In the catalogue of stars in the Calendarium of Al Achsasi Al Mouakket , this star was designated Thanih al Adzari -LRB- تاني العذاري - taanii al-aḏārii -RRB- , which was translated into [2 2 Latin 2] as Secunda Virginum , meaning the second virgin . This star , along with 17 ε CMa -LRB- Adhara -RRB- , 18 δ CMa -LRB- Wezen -RRB- and 19 η CMa

1: 6366 Canis_Major

2: 17730 Latin

1: 26808 Star

[winner: 1239848 Star_(glyph), pem_pos=8,
final_scores=-1.0293008, final_local_scores=-0.9661134,
pem=-3.1183283, similarity=-7.7830844]

gt: 26808 Star, pem_pos=0, final_scores=-1.0536964,
final_local_scores=-1.2156184, pem=-0.1941345,

```

similarity=-31.13103 ]
2: 5267 Constellation
  [winner: 5267 Constellation, pem_pos=0,
  final_scores=-1.198437, final_local_scores=-1.3628342,
  pem=-0.14364368, similarity=-35.79882|
  gt: 5267 Constellation, pem_pos=0,
  final_scores=-1.198437, final_local_scores=-1.3628342,
  pem=-0.14364368, similarity=-35.79882 ]
3: 29329 Spectrum
  [winner: 13576645 Eigendecomposition_of_a_matrix,
  pem_pos=5, final_scores=-1.2020103,
  final_local_scores=-1.207381, pem=-1.7880713,
  similarity=-22.217133|
  gt: 29329 Spectrum, pem_pos=0, final_scores=-1.4024982,
  final_local_scores=-1.5576136, pem=-0.8907548,
  similarity=-37.54852 ]
4: 23335 Parsec
  [winner: 23335 Parsec, pem_pos=0,
  final_scores=-1.1423104, final_local_scores=-1.285495,
  pem=0.0, similarity=-34.272022|
  gt: 23335 Parsec, pem_pos=0, final_scores=-1.1423104,
  final_local_scores=-1.285495, pem=0.0, similarity=-34.272022 ]
5: 277379 Margin_of_error
  [winner: 37085 Software_bug, pem_pos=0,
  final_scores=-0.91823727, final_local_scores=-0.8873085,
  pem=0.0, similarity=-22.389341|
  gt: 277379 Margin_of_error, pem_pos=18,
  final_scores=-1.8426478, final_local_scores=-1.9919405,
  pem=-6.001415, similarity=-22.71714 ]
  [winner: 1151882 Profit_margin, pem_pos=1,
  final_scores=-1.326778, final_local_scores=-1.2688506,
  pem=-0.5617299, similarity=-30.720545|
  RECALL MISS: 277379 Margin_of_error]
6: 170853 Supergiant
  [winner: 170853 Supergiant, pem_pos=0,
  final_scores=-0.80788726, final_local_scores=-0.90300775,
  pem=-0.009708832, similarity=-22.805042|
  gt: 170853 Supergiant, pem_pos=0,
  final_scores=-0.80788726, final_local_scores=-0.90300775,
  pem=-0.009708832, similarity=-22.805042 ]
7: 28927 Stellar_classification
  [winner: 2145671 Hyundai_Stellar, pem_pos=1,
  final_scores=-1.0978671, final_local_scores=-1.066361,
  pem=-0.8665292, similarity=-23.020304|
  RECALL MISS: 28927 Stellar_classification]
  [winner: 4910705 Navy_Enlisted_Classification, pem_pos=1,
  final_scores=-1.4061625, final_local_scores=-1.3495889,
  pem=-2.0769815, similarity=-24.889763|
  RECALL MISS: 28927 Stellar_classification]
8: 21544 Nuclear_fusion
  [winner: 21544 Nuclear_fusion, pem_pos=0,
  final_scores=-1.0069094, final_local_scores=-1.0833976,

```


pem=-0.09132205, similarity=-27.744408|
gt: 21544 Nuclear_fusion, pem_pos=0,
final_scores=-1.0069094, final_local_scores=-1.0833976,
pem=-0.09132205, similarity=-27.744408]
9: 11009033 Type_II_supernova
[winner: 11009033 Type_II_supernova, pem_pos=0,
final_scores=-0.17227489, final_local_scores=-0.1830779,
pem=-0.7846767, similarity=2.8934968|
gt: 11009033 Type_II_supernova, pem_pos=0,
final_scores=-0.17227489, final_local_scores=-0.1830779,
pem=-0.7846767, similarity=2.8934968]
[winner: 14774765 Supernova/Karma, pem_pos=3,
final_scores=-1.1143605, final_local_scores=-1.0335585,
pem=-1.0986123, similarity=-20.779308|
gt: 11009033 Type_II_supernova, pem_pos=1,
final_scores=-1.274179, final_local_scores=-1.3867407,
pem=-1.0729846, similarity=-31.45834]
10: 535458 Stellar_atmosphere
[winner: 535458 Stellar_atmosphere, pem_pos=0,
final_scores=-0.9244209, final_local_scores=-1.0075922,
pem=0.0, similarity=-25.978848|
gt: 535458 Stellar_atmosphere, pem_pos=0,
final_scores=-0.9244209, final_local_scores=-1.0075922,
pem=0.0, similarity=-25.978848]
11: 4974134 Alpha_Cygni_variable
[winner: 4974134 Alpha_Cygni_variable, pem_pos=0,
final_scores=-0.07315409, final_local_scores=-0.026926078,
pem=0.0, similarity=3.286185|
gt: 4974134 Alpha_Cygni_variable, pem_pos=0,
final_scores=-0.07315409, final_local_scores=-0.026926078,
pem=0.0, similarity=3.286185]
12: 63025 Variable_star
[winner: 1239848 Star_(glyph), pem_pos=8,
final_scores=-1.1199924, final_local_scores=-1.0651801,
pem=-3.1183283, similarity=-10.739434|
RECALL MISS: 63025 Variable_star]
[winner: 2712792 BY_Draconis_variable, pem_pos=3,
final_scores=-1.1918641, final_local_scores=-1.282237,
pem=-1.3862944, similarity=-26.635914|
gt: 63025 Variable_star, pem_pos=2,
final_scores=-1.4727285, final_local_scores=-1.6513197,
pem=-1.1615772, similarity=-38.872128]
13: 183934 Stellar_wind
[winner: 183934 Stellar_wind, pem_pos=0,
final_scores=-0.7949478, final_local_scores=-0.85276526,
pem=-0.08029993, similarity=-20.92182|
gt: 183934 Stellar_wind, pem_pos=0,
final_scores=-0.7949478, final_local_scores=-0.85276526,
pem=-0.08029993, similarity=-20.92182]
14: 22669 Open_cluster
[winner: 22669 Open_cluster, pem_pos=0,
final_scores=-0.9404885, final_local_scores=-1.0649946,

```

pem=-0.0009478649, similarity=-27.686697|
  gt: 22669 Open_cluster, pem_pos=0,
final_scores=-0.9404885, final_local_scores=-1.0649946,
pem=-0.0009478649, similarity=-27.686697 ]
15: 52015 Proper_motion
  [winner: 52015 Proper_motion, pem_pos=0,
final_scores=-0.89004165, final_local_scores=-1.0116471,
pem=0.0, similarity=-26.099855|
  gt: 52015 Proper_motion, pem_pos=0,
final_scores=-0.89004165, final_local_scores=-1.0116471,
pem=0.0, similarity=-26.099855 ]
16: 48239 Celestial_sphere
  [winner: 27859 Sphere, pem_pos=0,
final_scores=-0.8644592, final_local_scores=-0.93454635,
pem=-0.07567089, similarity=-23.387503|
  gt: 48239 Celestial_sphere, pem_pos=11,
final_scores=-1.9770558, final_local_scores=-2.2070184,
pem=-6.261318, similarity=-27.722094 ]
  [winner: 2482801 Celestial_Recordings, pem_pos=12,
final_scores=-1.100367, final_local_scores=-0.9755029,
pem=-2.7725887, similarity=-9.94347|
  RECALL MISS: 48239 Celestial_sphere]
17: 1522271 Epsilon_Canis_Majoris
  [winner: 6366 Canis_Major, pem_pos=0,
final_scores=-0.47563297, final_local_scores=-0.5940193,
pem=0.0, similarity=-13.637009|
  RECALL MISS: 1522271 Epsilon_Canis_Majoris]
  [winner: 14762594 CEBPE, pem_pos=0,
final_scores=-0.9578232, final_local_scores=-0.87250423,
pem=-1.5931553, similarity=-13.283728|
  RECALL MISS: 1522271 Epsilon_Canis_Majoris]
18: 1522349 Delta_Canis_Majoris
  [winner: 6366 Canis_Major, pem_pos=0,
final_scores=-0.40198272, final_local_scores=-0.51356757,
pem=0.0, similarity=-11.236167|
  RECALL MISS: 1522349 Delta_Canis_Majoris]
  [winner: 15805229 CD3D, pem_pos=1,
final_scores=-0.775082, final_local_scores=-0.6508931,
pem=-1.8190212, similarity=-5.4421153|
  RECALL MISS: 1522349 Delta_Canis_Majoris]
19: 158142 Eta_Canis_Majoris
  [winner: 158142 Eta_Canis_Majoris, pem_pos=0,
final_scores=-0.49964434, final_local_scores=-0.4912173,
pem=0.0, similarity=-10.569191|
  gt: 158142 Eta_Canis_Majoris, pem_pos=0,
final_scores=-0.49964434, final_local_scores=-0.4912173,
pem=0.0, similarity=-10.569191 ]
1: 6366 Canis_Major
  [pem_pos=0, final_scores=0.08159034,
final_local_scores=0.03429483, pem=0.0, similarity=5.113139]
2: 17730 Latin
  [pem_pos=0, final_scores=0.2303618,

```

```

final_local_scores=0.2993689, pem=-0.08185364,
similarity=13.468611]
1: 26751 Sun
    [pem_pos=0, final_scores=0.05237417,
final_local_scores=-0.035456523, pem=-0.23268287,
similarity=4.296985]
2: 26751 Sun
    [pem_pos=0, final_scores=0.10531609,
final_local_scores=0.035847306, pem=-0.23268287,
similarity=6.424834]
3: 26751 Sun
    [pem_pos=0, final_scores=0.08060835,
final_local_scores=0.008857861, pem=-0.23268287,
similarity=5.619415]
4: 6366 Canis_Major
    [pem_pos=1, final_scores=0.083768144,
final_local_scores=0.049879998, pem=-0.38747153,
similarity=7.6853623]

```

Sample: b'Andrea_Giganti'

Andrea Giganti -LRB- 1731 -- 1787 -RRB- was a [1 1 Sicilian 1] 1 architect of the [1 2 Sicilian 1] Baroque era . He was born in [2 2 Trapani 2] in 1731 ; in his youth he studied architecture under -LRB- 1684 -- 1754 -RRB- . Around 1751 Giganti came under the patronage of , , with whom he went to [3 3 Palermo 3] where he studied for the 3 priest hood and was ed . After ordination he entered the household of the [2 Sicilian 2] 4 aristocrat the Prince of 5 Scordia , where he seems to have been employed as both 6 confessor and private architect , working on the various houses owned by the Scordia family . Other [3 Sicilian 3] buildings in the baroque style designed by Giganti include Villa Galetti at [4 4 Bagheria 4] , Villa Ventimiglia , at , the church of S. Paolo dei Giardinieri , and the 7 dais and at the church of San Salvatore in [4 Palermo 4] . A trained 8 engineer , Giganti was also responsible for a number of 9 bridge s. Towards the end of his life he began to forsake the baroque style in favour of a more simple 10 neoclassical

```

1: 27619 Sicily
2: 44778 Trapani
3: 38881 Palermo
4: 44797 Bagheria
1: 1170 Architect

```

```

[winner: 2497175 Systems_architect, pem_pos=0,
final_scores=-1.0746021, final_local_scores=-0.9868608,
pem=0.0, similarity=-25.360182]
gt: 1170 Architect, pem_pos=1, final_scores=-1.1408386,
final_local_scores=-1.1869802, pem=-0.07108142,
similarity=-30.945591 ]
2: 2372757 Sicilian_Baroque
    [winner: 27619 Sicily, pem_pos=0, final_scores=0.4714548,
final_local_scores=0.4178959, pem=-0.66251653,

```

similarity=20.16343|
gt: 2372757 Sicilian_Baroque, pem_pos=8,
final_scores=-0.6210806, final_local_scores=-0.7036645,
pem=-3.0239794, similarity=-0.46416843]
[winner: 3957 Baroque, pem_pos=0,
final_scores=-0.057291865, final_local_scores=-0.08031717,
pem=-0.4497743, similarity=4.138831|
gt: 2372757 Sicilian_Baroque, pem_pos=9,
final_scores=-0.58755434, final_local_scores=-0.66234404,
pem=-3.4154756, similarity=2.8979316]
3: 23707 Priest
[winner: 13390365 The_Priest, pem_pos=0,
final_scores=-0.9423347, final_local_scores=-0.8001444,
pem=0.0, similarity=-19.78819|
RECALL MISS: 23707 Priest]
4: 20750674 Aristocracy_(class)
[winner: 4592678 Aristocrat_(fashion), pem_pos=2,
final_scores=-1.2744911, final_local_scores=-1.1933928,
pem=-3.996889, similarity=-9.787806|
gt: 20750674 Aristocracy_(class), pem_pos=0,
final_scores=-1.2955, final_local_scores=-1.3920112, pem=0.0,
similarity=-37.450676]
5: 6816834 Scordia
[winner: 6816834 Scordia, pem_pos=0,
final_scores=-0.24655895, final_local_scores=-0.12284526,
pem=0.0, similarity=0.4237653|
gt: 6816834 Scordia, pem_pos=0,
final_scores=-0.24655895, final_local_scores=-0.12284526,
pem=0.0, similarity=0.4237653]
6: 1206096 Confessor
[winner: 17501477 Confessor_of_the_Faith, pem_pos=0,
final_scores=-1.2173344, final_local_scores=-1.221586,
pem=0.0, similarity=-32.36485|
gt: 1206096 Confessor, pem_pos=1,
final_scores=-1.4238938, final_local_scores=-1.5328317,
pem=-0.040726963, similarity=-41.43156]
7: 2646133 Dais
[winner: 2646133 Dais, pem_pos=0,
final_scores=-0.9678189, final_local_scores=-0.96229476,
pem=-0.0012307585, similarity=-24.620388|
gt: 2646133 Dais, pem_pos=0, final_scores=-0.9678189,
final_local_scores=-0.96229476, pem=-0.0012307585,
similarity=-24.620388]
8: 38223 Engineer
[winner: 38223 Engineer, pem_pos=0,
final_scores=-1.2299231, final_local_scores=-1.2720907,
pem=-0.24685611, similarity=-32.52957|
gt: 38223 Engineer, pem_pos=0, final_scores=-1.2299231,
final_local_scores=-1.2720907, pem=-0.24685611,
similarity=-32.52957]
9: 3397 Bridge
[winner: 2702169 Bridging_(networking), pem_pos=4,

```

final_scores=-1.4263451, final_local_scores=-1.3604631,
pem=-2.6445367, similarity=-22.127821|
    gt: 3397 Bridge, pem_pos=0, final_scores=-1.6398252,
final_local_scores=-1.7024951, pem=-0.9226701,
similarity=-41.69851 ]
10: 2682331 Neoclassical_architecture
    [winner: 2682331 Neoclassical_architecture, pem_pos=0,
final_scores=-0.65220594, final_local_scores=-0.65379465,
pem=-0.14328104, similarity=-14.641639|
    gt: 2682331 Neoclassical_architecture, pem_pos=0,
final_scores=-0.65220594, final_local_scores=-0.65379465,
pem=-0.14328104, similarity=-14.641639 ]
1: 27619 Sicily
    [pem_pos=0, final_scores=0.5754024,
final_local_scores=0.5314429, pem=-0.66251653,
similarity=23.551897]
2: 44778 Trapani
    [pem_pos=0, final_scores=0.5324019,
final_local_scores=0.5218177, pem=-0.10367105,
similarity=20.225573]
3: 38881 Palermo
    [pem_pos=0, final_scores=0.38983256,
final_local_scores=0.33475113, pem=-0.5341467,
similarity=16.984127]
4: 44797 Bagheria
    [pem_pos=0, final_scores=0.13088726,
final_local_scores=0.14501455, pem=0.0, similarity=8.417236]
1: 27619 Sicily
    [pem_pos=0, final_scores=0.4714548,
final_local_scores=0.4178959, pem=-0.66251653,
similarity=20.16343]
2: 27619 Sicily
    [pem_pos=0, final_scores=0.60220575,
final_local_scores=0.56072146, pem=-0.66251653,
similarity=24.425627]
3: 27619 Sicily
    [pem_pos=0, final_scores=0.5395306,
final_local_scores=0.49225837, pem=-0.66251653,
similarity=22.382553]
4: 38881 Palermo
    [pem_pos=0, final_scores=0.42396045,
final_local_scores=0.37203062, pem=-0.5341467,
similarity=18.096622]

```

Sample: b'Julian_Knowles'

[1 Julian Knowles 1] is an [2 Australian 2] composer and performer , specialising in new and emerging technologies . His creative work spans the fields of composition for theatre , dance , film and television , electronic music , sound and new media arts , popular music and record production . Since the mid-1980s , he has established himself as a prominent artist in the area of electronic and new

music , achieving significant critical recognition for his performances and recordings . Although working primarily as a solo artist , he has been a member of the experimental sound art ensemble since the early 90s and has collaborated with a wide range of [3 Australian 3] and international artists . In recent years , his work has been presented at venues and events such as the [1 1 San Francisco Museum of Modern Art 1] , in [4 New York City 4] , the , The 1 Melbourne International Film Festival , 2 What Is Music , 3 Electrofringe , the [2 2 Sydney Opera House 2] , the 4 International Computer Music Conference , and a range of major arts venues . [5 Knowles 5] has an extensive background in the popular music area , having been involved in the [6 Australian 6] and UK independent music scenes from the mid-1980s onwards as a performer , programmer and producer in bands such as 5 The Shinking Violets -LRB- guitar , keyboards , production -RRB- , 6 Even As We Speak -LRB- guitar , keyboards , production -RRB- and Dutiful Daughters -LRB- bass guitar -RRB- ; Even As We Speak achieved multiple top 10s in the UK independent charts -LRB- NME/Melody Maker -RRB- during the 1990s . The band were one of only five [7 Australian 7] bands ever to record radio sessions with [8 British 8] DJ , [3 3 John Peel 3] , on [4 4 BBC Radio 1 4] . He also appeared as a guest artist on releases by 7 Big Heavy Stuff and 8 Swirl . In 1994 [9 Knowles 9] took up a position on the foundation staff of a new music program at the [5 5 University of Western Sydney 5] . In this role he established and steered a new major in music technology and played a key role in the establishment of the [10 Electronic Arts 10] program . In 2002 , he was appointed to the position of Head of the School of Contemporary Arts . In 2005 he was appointed Professor of Music and Head , School of Music and Drama at the [6 6 University of Wollongong 6] . In 2007 he took up a position as Professor in Music at the Creative Industries Faculty at [7 7 Queensland University of Technology 7] where he is the co-director of the Independent Music Project and a researcher within the . Since 2007 , [11 Knowles 11] has been Chair of , the peak body for the music industry in [12 Queensland 12] and a director of the [8 8 Music Council of Australia 8] . From 2003-2007 , he was a member of the Music Board of the [9 9 Australia Council for the Arts 9]

1: 148118 San_Francisco_Museum_of_Modern_Art

2: 28222 Sydney_Opera_House

3: 16036 John_Peel

4: 4348 BBC_Radio_1

5: 304175 University_of_Western_Sydney

6: 461935 University_of_Wollongong

7: 198332 Queensland_University_of_Technology

8: 23887242 Music_Council_of_Australia

9: 3779377 Australia_Council_for_the_Arts

1: 957551 Melbourne_International_Film_Festival

[winner: 1530814 Edinburgh_International_Film_Festival,

pem_pos=1, final_scores=-0.5375833,
 final_local_scores=-0.5876027, pem=-1.5040774,
 similarity=-5.2661157|
 RECALL MISS: 957551
 Melbourne_International_Film_Festival]
 [winner: 485363 HMAS_Melbourne_(R21), pem_pos=0,
 final_scores=-0.6340365, final_local_scores=-0.59513557,
 pem=0.0, similarity=-13.670319|
 RECALL MISS: 957551
 Melbourne_International_Film_Festival]
 2: 1459621 What_Is_Music
 [winner: 1459621 What_Is_Music, pem_pos=0,
 final_scores=-0.17049012, final_local_scores=-0.114213005,
 pem=0.0, similarity=0.6813691|
 gt: 1459621 What_Is_Music, pem_pos=0,
 final_scores=-0.17049012, final_local_scores=-0.114213005,
 pem=0.0, similarity=0.6813691]
 3: 3682125 Electrofringe
 [winner: 3682125 Electrofringe, pem_pos=0,
 final_scores=-0.12736544, final_local_scores=-0.050743468,
 pem=0.0, similarity=2.5754266|
 gt: 3682125 Electrofringe, pem_pos=0,
 final_scores=-0.12736544, final_local_scores=-0.050743468,
 pem=0.0, similarity=2.5754266]
 4: 483877 International_Computer_Music_Conference
 [winner: 483877 International_Computer_Music_Conference,
 pem_pos=0, final_scores=-0.048039183,
 final_local_scores=0.02315098, pem=0.0, similarity=4.7805843|
 gt: 483877 International_Computer_Music_Conference,
 pem_pos=0, final_scores=-0.048039183,
 final_local_scores=0.02315098, pem=0.0, similarity=4.7805843]
 5: 1924225 The_Shrinking_Violets
 [winner: 1924225 The_Shrinking_Violets, pem_pos=0,
 final_scores=-0.36564416, final_local_scores=-0.2688152,
 pem=0.0, similarity=-3.9322684|
 gt: 1924225 The_Shrinking_Violets, pem_pos=0,
 final_scores=-0.36564416, final_local_scores=-0.2688152,
 pem=0.0, similarity=-3.9322684]
 6: 1042222 Even_As_We_Speak
 [winner: 1042222 Even_As_We_Speak, pem_pos=0,
 final_scores=-0.1440642, final_local_scores=-0.08032253,
 pem=0.0, similarity=1.6927285|
 gt: 1042222 Even_As_We_Speak, pem_pos=0,
 final_scores=-0.1440642, final_local_scores=-0.08032253,
 pem=0.0, similarity=1.6927285]
 7: 2979533 Big_Heavy_Stuff
 [winner: 2979533 Big_Heavy_Stuff, pem_pos=0,
 final_scores=-0.05396256, final_local_scores=0.008877754,
 pem=0.0, similarity=4.354643|
 gt: 2979533 Big_Heavy_Stuff, pem_pos=0,
 final_scores=-0.05396256, final_local_scores=0.008877754,
 pem=0.0, similarity=4.354643]

```

8: 1121036 Swirl_(band)
  [winner: 1121036 Swirl_(band), pem_pos=0,
final_scores=-0.05132307, final_local_scores=0.025514394,
pem=-0.21585637, similarity=6.024974]
  gt: 1121036 Swirl_(band), pem_pos=0,
final_scores=-0.05132307, final_local_scores=0.025514394,
pem=-0.21585637, similarity=6.024974 ]
1: 148118 San_Francisco_Museum_of_Modern_Art
  [pem_pos=0, final_scores=0.31230527,
final_local_scores=0.36569327, pem=0.0, similarity=15.002729]
2: 28222 Sydney_Opera_House
  [pem_pos=0, final_scores=0.34379786,
final_local_scores=0.3413739, pem=0.0, similarity=14.276991]
3: 16036 John_Peel
  [pem_pos=0, final_scores=0.2294376,
final_local_scores=0.25509906, pem=0.0, similarity=11.702377]
4: 4348 BBC_Radio_1
  [pem_pos=0, final_scores=0.058462128,
final_local_scores=0.05629109, pem=0.0, similarity=5.7695513]
5: 304175 University_of_Western_Sydney
  [pem_pos=0, final_scores=0.41224378,
final_local_scores=0.42414814, pem=0.0, similarity=16.747139]
6: 461935 University_of_Wollongong
  [pem_pos=0, final_scores=0.35180765,
final_local_scores=0.37126523, pem=0.0, similarity=15.169007]
7: 198332 Queensland_University_of_Technology
  [pem_pos=0, final_scores=0.25141937,
final_local_scores=0.23990682, pem=0.0, similarity=11.249011]
8: 23887242 Music_Council_of_Australia
  [pem_pos=0, final_scores=0.13033883,
final_local_scores=0.19860628, pem=0.0, similarity=10.0165205]
9: 3779377 Australia_Council_for_the_Arts
  [pem_pos=0, final_scores=0.041374817,
final_local_scores=0.007638663, pem=0.0, similarity=4.317666]
1: 3289842 Julian_Knowles
  [pem_pos=0, final_scores=0.07070531,
final_local_scores=0.05738403, pem=0.0, similarity=5.802167]
2: 4689264 Australia
  [pem_pos=0, final_scores=0.631523,
final_local_scores=0.6428253, pem=-0.491009,
similarity=25.943085]
3: 4689264 Australia
  [pem_pos=0, final_scores=0.56878036,
final_local_scores=0.5742885, pem=-0.491009,
similarity=23.89781]
4: 645042 New_York_City
  [pem_pos=0, final_scores=0.5730836,
final_local_scores=0.58543736, pem=0.0, similarity=21.560331]
5: 3289842 Julian_Knowles
  [pem_pos=0, final_scores=0.10895263,
final_local_scores=0.09916341, pem=0.0, similarity=7.048947]
6: 4689264 Australia

```


[pem_pos=0, final_scores=0.6100885,
 final_local_scores=0.6194114, pem=-0.491009,
 similarity=25.244368]
 7: 4689264 Australia
 [pem_pos=0, final_scores=0.48737288,
 final_local_scores=0.4853632, pem=-0.491009,
 similarity=21.2441]
 8: 31717 United_Kingdom
 [pem_pos=0, final_scores=0.50021803,
 final_local_scores=0.51307386, pem=-0.48936236,
 similarity=22.06209]
 9: 3289842 Julian_Knowles
 [pem_pos=0, final_scores=0.14621972,
 final_local_scores=0.13783243, pem=0.0, similarity=8.202908]
 10: 262933 Electronic_Arts
 [pem_pos=0, final_scores=0.09044962,
 final_local_scores=0.1204395, pem=0.0, similarity=7.6838675]
 11: 3289842 Julian_Knowles
 [pem_pos=0, final_scores=0.097799286,
 final_local_scores=0.086980104, pem=0.0, similarity=6.685373]
 12: 59051 Queensland
 [pem_pos=0, final_scores=0.36875427,
 final_local_scores=0.40361488, pem=-0.083265886,
 similarity=16.587198]

Sample: b'Kilometre_Zero'

_ -LRB- Bucharest -RRB- The Kilometre Zero monument -LRB- [1
 1 Romanian 1] : Kilometrul Zero -RRB- located in central [1 1
 Bucharest 1] , [2 2 Romania 2] , in front of Saint George 's
 Church , was created by in 1938 . The distances from [2
 Bucharest 2] to other cities in [3 Romania 3] are measured
 from this monument . It is divided into eight sections , each
 representing a [4 Romanian 4] historical province : [5 2 Munt
 enia 5] , , [3 3 Bessarabia 3] , [4 4 Moldavia 4] , , [5 5 Tr
 ansylvania 5] , [6 6 Banat 6] and [7 7 Oltenia 7] . Among the
 cities inscribed on it are also , [8 8 Orhei 8] , [9 9 Tighina
 9] , which are currently in the , as well as [10 10 Silistra 1
 0] and [11 11 Dobrich 11] -LRB- Bazargic -RRB- in [12 12 Bulg
 aria 12] , which were part of 3 Greater Romania

1: 36877 Bucharest
 2: 25445 Romania
 3: 38872 Bessarabia
 4: 46007 Moldavia
 5: 30853 Transylvania
 6: 271319 Banat
 7: 172232 Oltenia
 8: 544978 Orhei
 9: 449688 Bender, Moldova
 10: 296915 Silistra
 11: 298104 Dobrich
 12: 3415 Bulgaria
 1: 25534 Romanian_language

[winner: 25445 Romania, pem_pos=1,
final_scores=0.10606824, final_local_scores=0.027888283,
pem=-1.0382323, similarity=10.568024]
gt: 25534 Romanian_language, pem_pos=0,
final_scores=-0.35727733, final_local_scores=-0.433582,
pem=-0.80176646, similarity=-4.4891086]
2: 268039 Muntenia
[winner: 46026 Wallachia, pem_pos=0,
final_scores=0.40987873, final_local_scores=0.38285136,
pem=0.0, similarity=15.514762]
gt: 268039 Muntenia, pem_pos=1,
final_scores=0.112836555, final_local_scores=0.108207196,
pem=-0.0010444013, similarity=7.324511]
3: 392816 Greater_Romania
[winner: 392816 Greater_Romania, pem_pos=0,
final_scores=-0.16667338, final_local_scores=-0.20972937,
pem=0.0, similarity=-2.1690297]
gt: 392816 Greater_Romania, pem_pos=0,
final_scores=-0.16667338, final_local_scores=-0.20972937,
pem=0.0, similarity=-2.1690297]
1: 36877 Bucharest
[pem_pos=0, final_scores=0.5910664,
final_local_scores=0.5704508, pem=-0.01781679,
similarity=21.209993]
2: 25445 Romania
[pem_pos=0, final_scores=0.6367868,
final_local_scores=0.6076177, pem=-0.21241479,
similarity=23.37738]
3: 38872 Bessarabia
[pem_pos=0, final_scores=0.3756551,
final_local_scores=0.34525064, pem=0.0, similarity=14.39268]
4: 46007 Moldavia
[pem_pos=0, final_scores=0.3870818,
final_local_scores=0.33447552, pem=0.0, similarity=14.07113]
5: 30853 Transylvania
[pem_pos=0, final_scores=0.3008967,
final_local_scores=0.2504291, pem=-0.067388475,
similarity=11.929485]
6: 271319 Banat
[pem_pos=0, final_scores=0.14336418,
final_local_scores=0.087868184, pem=-0.003813974,
similarity=6.7326155]
7: 172232 Oltenia
[pem_pos=0, final_scores=0.22240691,
final_local_scores=0.18091658, pem=0.0, similarity=9.488625]
8: 544978 Orhei
[pem_pos=0, final_scores=0.07758896,
final_local_scores=0.11935046, pem=0.0, similarity=7.6513686]
9: 449688 Bender, Moldova
[pem_pos=0, final_scores=0.39113265,
final_local_scores=0.3875304, pem=0.0, similarity=15.654393]
10: 296915 Silistra

[pem_pos=0, final_scores=0.2877304,
 final_local_scores=0.2777996, pem=0.0, similarity=12.379807]
 11: 298104 Dobrich
 [pem_pos=0, final_scores=0.20412521,
 final_local_scores=0.20229939, pem=0.0, similarity=10.12673]
 12: 3415 Bulgaria
 [pem_pos=0, final_scores=0.6567034,
 final_local_scores=0.6591562, pem=-0.31727514,
 similarity=25.485641]
 1: 25445 Romania
 [pem_pos=1, final_scores=0.10606824,
 final_local_scores=0.027888283, pem=-1.0382323,
 similarity=10.568024]
 2: 36877 Bucharest
 [pem_pos=0, final_scores=0.51110244,
 final_local_scores=0.48310232, pem=-0.01781679,
 similarity=18.60334]
 3: 25445 Romania
 [pem_pos=0, final_scores=0.68439615,
 final_local_scores=0.6596237, pem=-0.21241479,
 similarity=24.929344]
 4: 25445 Romania
 [pem_pos=1, final_scores=0.60161763,
 final_local_scores=0.5692007, pem=-1.0382323,
 similarity=26.721867]
 5: 46026 Wallachia
 [pem_pos=0, final_scores=0.40987873,
 final_local_scores=0.38285136, pem=0.0, similarity=15.514762]

Sample: b'Clara_Nordstrom'

[1 Clara Nordström 1] , maiden name and pseudonym for Clara
 Elisabet von Vegesack , -LRB- [1 1 Karlskrona 1] , [2 2 Sweden
 2] , January 18 , 1886 - [3 3 Mindelheim 3] , [4 4 Germany 4]
 , February 7 , 1962 -RRB- was a [2 German 2] writer and
 translator of [3 Swedish 3] descent . With the themes of her
 writing and her [4 Swedish 4] maiden name she profited from
 the [5 German 5] interest for [6 Scandinavian 6] writers .
 Born the daughter of a physician and a peasant woman in [7
 Karlskrona 7] and brought up in [5 5 Växjö 5] -LRB- [8 Sweden
 8] -RRB- , she was bed-ridden owing to illness up to her
 twelfth year . It was only after that , from about 1897 ,
 that she started to frequent various private schools in [9
 Växjö 9] . In 1903 , she went to 1 Hildesheim -LRB- Germany -
 RRB- and shortly afterwards to [6 6 Braunschweig 6] -LRB- [10
 Germany 10] -RRB- in order to learn the [11 German 11]
 language . On April 17th 1905 she married , in [12 Växjö
 12] , the son of her teacher , 15 years her senior , and in
 1906 gave birth to their son [13 Gustav Adolf 13] . She got
 divorced in 1909 after she had been left by her husband .
 Nordström returned to [14 Växjö 14] for a short time and in
 the same year moved to [7 7 Berlin 7] to become a
 photographer . After three years of instruction and practical

training , she had to give up her profession for health reasons . In 1912 she went to [8 8 Munich 8] to become a writer . It was there that in 1914 she met Siegfried von Vegesack , whom she married on February 16th in [9 9 Stockholm 9] . In 1916 she moved with her husband to [15 Berlin 15] , where in April 1917 her daughter Isabel was born -LRB- who died in 2005 -RRB- . Because of an ailment of Siegfried von Vegesack 's , the family in 1917 moved to a farm near [10 10 Dingolfing 10] and later to Großwalding near [11 11 Deggendorf 11] . In 1918 they acquired a corn tower near [12 12 Regen 12] , which they refurbished into a residential tower . In 1920 Karin their second daughter was born but died only a few days later . In 1923 their son Gotthard was born , to be killed in action in the [16 Second World War 16] in 1944 . In the same year Nordström published her first novel `` Tomtelilla '' both in [17 Germany 17] and in [18 Sweden 18] . With her mother 's death an important source of income had run dry . Therefore , Nordström opened up , in the tower , a place for artists and writers to live . In these years the couple started gradually to grow apart . In 1929 the family moved to [19 Switzerland 19] . Shortly afterwards [20 Clara Nordström 20] moved with her children to [13 13 Stuttgart 13] and got a divorce in 1935 at Vegesack 's wish . In that year she started to read from her works all over [21 Germany 21] . From her [22 German 22] base , she also wrote articles in the [23 Swedish 23] press . In 1936 she returned for a short period to the residential tower in Weißenstein near Regen and in 1938/39 she built a house in Baiersbronn in the [14 14 Black Forest 14] . In 1944 she was called to [15 15 Königsberg 15] to read from her texts for the radio station run by the [24 German 24] transmitting in [25 Swedish 25] , but in 1944 had to flee to [16 16 Hamburg 16] . Throughout her life she again and again had to struggle with severe ailments and therefore intensely questioned her faith , which is what the characters in her books do . In 1948 she converted from [26 Protestantism 26] to [27 Catholicism 27] . Round about 1950 she again moved to [28 Stuttgart 28] and took orders -LRB- of St. Benedict -RRB- in the convent of 2 Neresheim

- 1: 1018885 Karlskrona
- 2: 5058739 Sweden
- 3: 3340882 Mindelheim
- 4: 11867 Germany
- 5: 1019408 Växjö
- 6: 51032 Braunschweig
- 7: 3354 Berlin
- 8: 19058 Munich
- 9: 26741 Stockholm
- 10: 5453118 Dingolfing
- 11: 1395440 Deggendorf
- 12: 2539663 Regen
- 13: 28565 Stuttgart
- 14: 3385 Black_Forest

15: 15413504 Königsberg
16: 13467 Hamburg
1: 153231 Hildesheim
 [winner: 153231 Hildesheim, pem_pos=0,
final_scores=-0.32171118, final_local_scores=-0.40740025,
pem=0.0, similarity=-8.067924]
 gt: 153231 Hildesheim, pem_pos=0,
final_scores=-0.32171118, final_local_scores=-0.40740025,
pem=0.0, similarity=-8.067924]
2: 5947109 Neresheim
 [winner: 5947109 Neresheim, pem_pos=0,
final_scores=-0.23419857, final_local_scores=-0.16405314,
pem=0.0, similarity=-0.8059601]
 gt: 5947109 Neresheim, pem_pos=0,
final_scores=-0.23419857, final_local_scores=-0.16405314,
pem=0.0, similarity=-0.8059601]
1: 1018885 Karlskrona
 [pem_pos=0, final_scores=0.2999478,
final_local_scores=0.31345758, pem=0.0, similarity=13.4439125]
2: 5058739 Sweden
 [pem_pos=0, final_scores=0.75250727,
final_local_scores=0.7410285, pem=-0.2135508,
similarity=27.364803]
3: 3340882 Mindelheim
 [pem_pos=1, final_scores=0.09672172,
final_local_scores=0.14324123, pem=-0.040166024,
similarity=8.582746]
4: 11867 Germany
 [pem_pos=0, final_scores=0.76466995,
final_local_scores=0.73009604, pem=-0.15529114,
similarity=26.721733]
5: 1019408 Växjö
 [pem_pos=0, final_scores=0.29351443,
final_local_scores=0.30768648, pem=-0.15867667,
similarity=14.1346]
6: 51032 Braunschweig
 [pem_pos=0, final_scores=0.117148146,
final_local_scores=0.053026095, pem=-0.12703769,
similarity=6.362968]
7: 3354 Berlin
 [pem_pos=0, final_scores=0.54844654,
final_local_scores=0.5112891, pem=-0.15483057,
similarity=20.189594]
8: 19058 Munich
 [pem_pos=0, final_scores=0.60493666,
final_local_scores=0.5773232, pem=-0.119245164,
similarity=21.966663]
9: 26741 Stockholm
 [pem_pos=0, final_scores=0.76811403,
final_local_scores=0.7615038, pem=-0.1047564,
similarity=27.384188]
10: 5453118 Dingolfing

[pem_pos=0, final_scores=0.37321496,
final_local_scores=0.41192693, pem=0.0, similarity=16.382435]
11: 1395440 Deggendorf
[pem_pos=0, final_scores=0.2928328,
final_local_scores=0.30362245, pem=0.0, similarity=13.150413]
12: 2539663 Regen
[pem_pos=1, final_scores=0.07517664,
final_local_scores=0.12923992, pem=-0.68120277,
similarity=11.650976]
13: 28565 Stuttgart
[pem_pos=0, final_scores=0.5561809,
final_local_scores=0.523092, pem=-0.15141529,
similarity=20.523241]
14: 3385 Black_Forest
[pem_pos=0, final_scores=0.4939986,
final_local_scores=0.49729443, pem=0.0, similarity=18.92997]
15: 15413504 Königsberg
[pem_pos=0, final_scores=0.31626374,
final_local_scores=0.29236296, pem=-0.21226496,
similarity=13.968737]
16: 13467 Hamburg
[pem_pos=0, final_scores=0.51489633,
final_local_scores=0.47922564, pem=-0.27202022,
similarity=19.87005]
1: 6855791 Clara_Nordström
[pem_pos=0, final_scores=0.085257754,
final_local_scores=0.14110458, pem=0.0, similarity=8.300555]
2: 11867 Germany
[pem_pos=1, final_scores=0.5409635,
final_local_scores=0.4857307, pem=-1.0465823,
similarity=24.276365]
3: 5058739 Sweden
[pem_pos=0, final_scores=0.4787501,
final_local_scores=0.44199044, pem=-0.63135606,
similarity=20.713001]
4: 5058739 Sweden
[pem_pos=0, final_scores=0.52224904,
final_local_scores=0.48950642, pem=-0.63135606,
similarity=22.130974]
5: 11867 Germany
[pem_pos=2, final_scores=0.8544888,
final_local_scores=0.82917565, pem=-1.0465823,
similarity=34.525448]
6: 26740 Scandinavia
[pem_pos=0, final_scores=0.23427747,
final_local_scores=0.21252716, pem=-0.27464768,
similarity=11.925524]
7: 1018885 Karlskrona
[pem_pos=0, final_scores=0.36312503,
final_local_scores=0.38246912, pem=0.0, similarity=15.503355]
8: 5058739 Sweden
[pem_pos=0, final_scores=0.48989064,

final_local_scores=0.45415974, pem=-0.2135508,
similarity=18.804068]
9: 1019408 Växjö
[pem_pos=0, final_scores=0.36598814,
final_local_scores=0.38685298, pem=-0.15867667,
similarity=16.497086]
10: 11867 Germany
[pem_pos=0, final_scores=0.22171961,
final_local_scores=0.13700533, pem=-0.15529114,
similarity=9.022722]
11: 11884 German_language
[pem_pos=1, final_scores=0.20636491,
final_local_scores=0.1776393, pem=-0.981596,
similarity=14.728896]
12: 1019408 Växjö
[pem_pos=0, final_scores=0.4176342,
final_local_scores=0.44326842, pem=-0.15867667,
similarity=18.180634]
13: 40869264 Gustavus_Adolphus_of_Sweden
[pem_pos=0, final_scores=0.12549959,
final_local_scores=0.13380235, pem=-0.6048546,
similarity=11.371936]
14: 1019408 Växjö
[pem_pos=0, final_scores=0.24072216,
final_local_scores=0.25001898, pem=-0.15867667,
similarity=12.413687]
15: 3354 Berlin
[pem_pos=0, final_scores=0.70410436,
final_local_scores=0.6813217, pem=-0.15483057,
similarity=25.263704]
16: 32927 World_War_II
[pem_pos=0, final_scores=0.36638504,
final_local_scores=0.36779213, pem=0.0, similarity=15.065364]
17: 11867 Germany
[pem_pos=0, final_scores=0.72462255,
final_local_scores=0.6863504, pem=-0.15529114,
similarity=25.416275]
18: 5058739 Sweden
[pem_pos=0, final_scores=0.7067587,
final_local_scores=0.6910552, pem=-0.2135508,
similarity=25.873499]
19: 26748 Switzerland
[pem_pos=0, final_scores=0.62224025,
final_local_scores=0.60810983, pem=-0.18598008,
similarity=23.248312]
20: 6855791 Clara_Nordström
[pem_pos=0, final_scores=0.0975012,
final_local_scores=0.1544787, pem=0.0, similarity=8.699665]
21: 11867 Germany
[pem_pos=0, final_scores=0.66114044,
final_local_scores=0.6170058, pem=-0.15529114,
similarity=23.346895]

22: 11867 Germany
 [pem_pos=2, final_scores=0.5309445,
 final_local_scores=0.47575277, pem=-1.0465823,
 similarity=23.978601]
 23: 5058739 Sweden
 [pem_pos=0, final_scores=0.63415295,
 final_local_scores=0.6117444, pem=-0.63135606,
 similarity=25.7788]
 24: 11867 Germany
 [pem_pos=2, final_scores=0.6336587,
 final_local_scores=0.58795244, pem=-1.0465823,
 similarity=27.326864]
 25: 5058739 Sweden
 [pem_pos=0, final_scores=0.48998874,
 final_local_scores=0.45426697, pem=-0.63135606,
 similarity=21.079357]
 26: 25814008 Protestantism
 [pem_pos=0, final_scores=0.094435945,
 final_local_scores=0.09222503, pem=0.0, similarity=6.841892]
 27: 606848 Catholic_Church
 [pem_pos=1, final_scores=0.03474392,
 final_local_scores=0.050579816, pem=-0.660428,
 similarity=9.190625]
 28: 28565 Stuttgart
 [pem_pos=0, final_scores=0.4983284,
 final_local_scores=0.45989692, pem=-0.15141529,
 similarity=18.637375]

Sample: b'Matt_Higgins'

-LRB- ice_hockey -RRB- Matt Higgins -LRB- born October
 29 , 1977 -RRB- is a [1 1 Canadian 1] professional 1 ice hock
 ey who currently plays for [2 2 Slovenia 2] n based team [3 3
 HDD Olimpija Ljubljana 3] of the [4 4 Erste Bank Eishockey Li
 ga 4] -LRB- EBEL -RRB- . Matt Higgins was selected 18th
 overall in the 2 1996 [1 NHL Entry Draft 1] by the . He
 played in 57 regular season [2 NHL 2] games throughout the 3 1
 997 -- 98 [3 NHL 3] season to the 4 2000 -- 01 [4 NHL 4] seas
 on . He began his hockey career with the 5 Western Hockey Lea
 gue 6 Moose Jaw Warriors in 1993 where he played until he was
 drafted . He then played for the 7 American Hockey League 8 Fr
 edericton Canadiens and 9 Quebec Citadelles . He played
 through two seasons for the [5 5 Bridgeport Sound Tigers 5]
 before moving to [6 6 Europe 6] to play for the 10 Deutsche Ei
 shockey Liga 11 Iserlohn Roosters . He currently plays for
 the [7 7 HDD Olimpija Ljubljana 7]

1: 5042916 Canada
 2: 27338 Slovenia
 3: 11724485 HDD_Olimpija_Ljubljana
 4: 9760394 Austrian_Hockey_League
 5: 664236 Bridgeport_Sound_Tigers
 6: 9239 Europe
 7: 11724485 HDD_Olimpija_Ljubljana

1: 14790 Ice_hockey
[winner: 14946 Ice, pem_pos=0, final_scores=-0.74159044, final_local_scores=-0.6906713, pem=-0.116236255, similarity=-15.889185]
gt: 14790 Ice_hockey, pem_pos=3, final_scores=-1.0115256, final_local_scores=-1.1667902, pem=-2.165202, similarity=-18.954933]
[winner: 14790 Ice_hockey, pem_pos=1, final_scores=-1.2804953, final_local_scores=-1.4605987, pem=-0.8225384, similarity=-35.024372]
gt: 14790 Ice_hockey, pem_pos=1, final_scores=-1.2804953, final_local_scores=-1.4605987, pem=-0.8225384, similarity=-35.024372]
2: 2840500 1996_NHL_Entry_Draft
[winner: 858982 NHL_Entry_Draft, pem_pos=0, final_scores=0.091090485, final_local_scores=0.06001556, pem=0.0, similarity=5.880697]
RECALL MISS: 2840500 1996_NHL_Entry_Draft]
[winner: 34636 1996, pem_pos=0, final_scores=-0.7337299, final_local_scores=-0.5534564, pem=-0.78063476, similarity=-8.181317]
RECALL MISS: 2840500 1996_NHL_Entry_Draft]
3: 1903294 1997-98_NHL_season
[winner: 21809 National_Hockey_League, pem_pos=0, final_scores=0.08831917, final_local_scores=0.009707987, pem=-0.036005534, similarity=4.5752225]
RECALL MISS: 1903294 1997-98_NHL_season]
[winner: 2602335 Toronto_municipal_election,_1997, pem_pos=0, final_scores=-1.0481259, final_local_scores=-1.008955, pem=0.0, similarity=-26.019518]
RECALL MISS: 1903294 1997-98_NHL_season]
[winner: 75433 Windows_98, pem_pos=0, final_scores=-1.1480738, final_local_scores=-1.1071274, pem=-1.4472766, similarity=-21.078661]
RECALL MISS: 1903294 1997-98_NHL_season]
[winner: 38340996 Methylene_bridge, pem_pos=0, final_scores=-1.2079742, final_local_scores=-1.1056668, pem=0.0, similarity=-28.905588]
RECALL MISS: 1903294 1997-98_NHL_season]
[winner: 690345 Season_(sports), pem_pos=0, final_scores=-1.9103427, final_local_scores=-2.1348076, pem=-1.219945, similarity=-52.982952]
RECALL MISS: 1903294 1997-98_NHL_season]
4: 1673932 2000-01_NHL_season
[winner: 21809 National_Hockey_League, pem_pos=0, final_scores=0.27728623, final_local_scores=0.21612582, pem=-0.036005534, similarity=10.735142]
RECALL MISS: 1673932 2000-01_NHL_season]
[winner: 432383 2000_United_States_Census, pem_pos=1, final_scores=-0.4857319, final_local_scores=-0.43071103, pem=-1.4367632, similarity=-0.9502213]
RECALL MISS: 1673932 2000-01_NHL_season]

[winner: 3453517 FabricLive.01, pem_pos=2,
final_scores=-0.7905298, final_local_scores=-0.59470075,
pem=-2.3963957, similarity=-0.6253717|
RECALL MISS: 1673932 2000-01_NHL_season]
[winner: 38340996 Methylene_bridge, pem_pos=0,
final_scores=-1.3370082, final_local_scores=-1.2466168,
pem=0.0, similarity=-33.11182|
RECALL MISS: 1673932 2000-01_NHL_season]
[winner: 690345 Season_(sports), pem_pos=0,
final_scores=-1.8949299, final_local_scores=-2.1179714,
pem=-1.219945, similarity=-52.480526|
RECALL MISS: 1673932 2000-01_NHL_season]
5: 218746 Western_Hockey_League
[winner: 218746 Western_Hockey_League, pem_pos=0,
final_scores=-0.9122657, final_local_scores=-1.0449829,
pem=-0.149761, similarity=-26.28024|
gt: 218746 Western_Hockey_League, pem_pos=0,
final_scores=-0.9122657, final_local_scores=-1.0449829,
pem=-0.149761, similarity=-26.28024]
6: 1893905 Moose_Jaw_Warriors
[winner: 1893905 Moose_Jaw_Warriors, pem_pos=0,
final_scores=-0.33302546, final_local_scores=-0.38113275,
pem=0.0, similarity=-7.2840486|
gt: 1893905 Moose_Jaw_Warriors, pem_pos=0,
final_scores=-0.33302546, final_local_scores=-0.38113275,
pem=0.0, similarity=-7.2840486]
7: 207085 American_Hockey_League
[winner: 207085 American_Hockey_League, pem_pos=0,
final_scores=-0.7782212, final_local_scores=-0.91358745,
pem=-0.0011107532, similarity=-23.167519|
gt: 207085 American_Hockey_League, pem_pos=0,
final_scores=-0.7782212, final_local_scores=-0.91358745,
pem=-0.0011107532, similarity=-23.167519]
8: 2808917 Fredericton_Canadiens
[winner: 2808917 Fredericton_Canadiens, pem_pos=0,
final_scores=-0.39307642, final_local_scores=-0.4307386,
pem=0.0, similarity=-8.764386|
gt: 2808917 Fredericton_Canadiens, pem_pos=0,
final_scores=-0.39307642, final_local_scores=-0.4307386,
pem=0.0, similarity=-8.764386]
9: 1031323 Quebec_Citadelles
[winner: 1031323 Quebec_Citadelles, pem_pos=0,
final_scores=-0.02076742, final_local_scores=0.012453169,
pem=0.0, similarity=4.4613404|
gt: 1031323 Quebec_Citadelles, pem_pos=0,
final_scores=-0.02076742, final_local_scores=0.012453169,
pem=0.0, similarity=4.4613404]
10: 521823 Deutsche_Eishockey_Liga
[winner: 521823 Deutsche_Eishockey_Liga, pem_pos=0,
final_scores=-0.7919485, final_local_scores=-0.88571537,
pem=0.0, similarity=-22.3418|
gt: 521823 Deutsche_Eishockey_Liga, pem_pos=0,

```

final_scores=-0.7919485, final_local_scores=-0.88571537,
pem=0.0, similarity=-22.3418 ]
11: 3155950 Iserlohn_Roosters
    [winner: 3155950 Iserlohn_Roosters, pem_pos=0,
final_scores=-0.7194321, final_local_scores=-0.7976597,
pem=0.0, similarity=-19.714043]
    gt: 3155950 Iserlohn_Roosters, pem_pos=0,
final_scores=-0.7194321, final_local_scores=-0.7976597,
pem=0.0, similarity=-19.714043 ]
1: 5042916 Canada
    [pem_pos=0, final_scores=0.51897067,
final_local_scores=0.5695348, pem=-0.41303828,
similarity=23.331932]
2: 27338 Slovenia
    [pem_pos=0, final_scores=0.37482256,
final_local_scores=0.45055705, pem=-0.27845925,
similarity=19.049538]
3: 11724485 HDD_Olimpija_Ljubljana
    [pem_pos=0, final_scores=0.1715266,
final_local_scores=0.16467345, pem=0.0, similarity=9.003897]
4: 9760394 Austrian_Hockey_League
    [pem_pos=0, final_scores=0.26330465,
final_local_scores=0.26131198, pem=0.0, similarity=11.887784]
5: 664236 Bridgeport_Sound_Tigers
    [pem_pos=0, final_scores=0.11811112,
final_local_scores=0.10027285, pem=0.0, similarity=7.0820546]
6: 9239 Europe
    [pem_pos=0, final_scores=0.51977545,
final_local_scores=0.63159794, pem=-0.086338006,
similarity=23.407375]
7: 11724485 HDD_Olimpija_Ljubljana
    [pem_pos=0, final_scores=0.035913035,
final_local_scores=0.01653622, pem=0.0, similarity=4.5831866]
1: 858982 NHL_Entry_Draft
    [pem_pos=0, final_scores=0.091090485,
final_local_scores=0.06001556, pem=0.0, similarity=5.880697]
2: 21809 National_Hockey_League
    [pem_pos=0, final_scores=0.29419142,
final_local_scores=0.23459214, pem=-0.036005534,
similarity=11.286214]
3: 21809 National_Hockey_League
    [pem_pos=0, final_scores=0.08831917,
final_local_scores=0.009707987, pem=-0.036005534,
similarity=4.5752225]
4: 21809 National_Hockey_League
    [pem_pos=0, final_scores=0.27728623,
final_local_scores=0.21612582, pem=-0.036005534,
similarity=10.735142]

```

Sample: b'Todd_Wider'

[1 Todd Wider 1] is an [1 1 American 1] and [2 Emmy Award 2]
winning 1 film producer based in [3 2 New York 3] , who is

active in 3 documentary filmmaking . He graduated from [2 2 Ward Melville High School 2] in 1982 , [3 3 Princeton University 3] in 1986 , and 4 Columbia College of Physicians and Surgeons in 1990 . As a surgeon , he was active in helping the passage of the Women 's Health and Cancer Act of 1998 , federal legislation signed into law by President [4 4 Bill Clinton 4] , mandating insurance coverage for 5 breast cancer reconstruction . He also was a volunteer surgeon for Victims Services , an organization providing surgery to victims of abuse , and was a volunteer surgeon at after the in [4 New York City 4] . As a film producer , [5 Wider 5] has produced , with his brother Jedd , 6 Beyond Conviction -LRB- 2006 -RRB- , a documentary about restorative justice in the [6 Pennsylvania 6] prison system that premiered on [7 MSNBC 7] , was featured on [8 Oprah 8] , and won a number of awards . [9 Wider 9] also produced 7 What Would Jesus Buy ? -LRB- 2007 -RRB- , a [5 5 Morgan Spurlock 5] film about consumerism in the [6 6 United States 6] that focuses on the performance artist . He is also an executive producer of 8 Taxi to the Dark Side -LRB- 2007 -RRB- , a documentary about prisoner abuse directed by [7 7 Alex Gibney 7] that won the [10 Academy Award 10] for Best Documentary of 2008 . Taxi to the Dark Side also won an [11 Emmy 11] for Best Documentary of 2008 . Wider also produced The Untyings -LRB- 2006 -RRB- , a film about 9 exorcism in [8 8 Romania 8] and A Dream in Doubt -LRB- 2006 -RRB- , directed by Tami Yeager , a documentary about intolerance in the wake of the September 11 attacks that focuses on a [9 9 Sikh 9] family in [10 10 Arizona 10] . In addition , he produced 10 Kicking It -LRB- 2007 -RRB- , a documentary about the Homeless World Cup , an international tournament for the homeless football league , which premiered at the [11 11 Sundance Film Festival 11] in 2008 , and features [12 12 Colin Farrell 12] and the music of [13 13 U2 13] . He also produced A Time to Stir which screened as a work-in-progress at the 2008 and concerns the [14 14 Columbia University 14]

1: 3434750 United_States

2: 2189889 Ward_Melville_High_School

3: 23922 Princeton_University

4: 3356 Bill_Clinton

5: 723301 Morgan_Spurlock

6: 3434750 United_States

7: 4210630 Alex_Gibney

8: 25445 Romania

9: 29405 Sikh

10: 21883824 Arizona

11: 23943448 Sundance_Film_Festival

12: 206705 Colin_Farrell

13: 52780 U2

14: 6310 Columbia_University

1: 21555729 Film

[winner: 23721339 House_of_the_Wolf_Man, pem_pos=0,

final_scores=-0.793215, final_local_scores=-0.74628985,
 pem=0.0, similarity=-18.181065|
 RECALL MISS: 21555729 Film]

2: 8210131 New_York
 [winner: 645042 New_York_City, pem_pos=1,
 final_scores=0.45834845, final_local_scores=0.43132222,
 pem=-1.7906687, similarity=26.699163|
 gt: 8210131 New_York, pem_pos=0,
 final_scores=-0.09118207, final_local_scores=-0.16179419,
 pem=-0.2889045, similarity=0.8325591]

3: 8088 Documentary_film
 [winner: 20184441 Jack_Johnson_(film), pem_pos=13,
 final_scores=-1.1417705, final_local_scores=-1.1342902,
 pem=-2.8903718, similarity=-14.041475|
 gt: 8088 Documentary_film, pem_pos=0,
 final_scores=-1.3621817, final_local_scores=-1.5457232,
 pem=-0.07851275, similarity=-41.610786]

4: 1170596
 Columbia_University_College_of_Physicians_and_Surgeons
 [winner: 1170596
 Columbia_University_College_of_Physicians_and_Surgeons,
 pem_pos=0, final_scores=-0.075943425,
 final_local_scores=-0.069384366, pem=0.0,
 similarity=2.019145|
 gt: 1170596
 Columbia_University_College_of_Physicians_and_Surgeons,
 pem_pos=0, final_scores=-0.075943425,
 final_local_scores=-0.069384366, pem=0.0,
 similarity=2.019145]

5: 70547 Breast_cancer
 [winner: 4489 Breast, pem_pos=1, final_scores=-1.2283348,
 final_local_scores=-1.2189753, pem=-0.3169917,
 similarity=-30.563087|
 gt: 70547 Breast_cancer, pem_pos=0,
 final_scores=-1.3144155, final_local_scores=-1.4031696,
 pem=0.0, similarity=-37.78367]

[winner: 8263081 Nasopharynx_cancer, pem_pos=5,
 final_scores=-1.2352449, final_local_scores=-1.1883347,
 pem=-1.609438, similarity=-22.620192|
 gt: 70547 Breast_cancer, pem_pos=10,
 final_scores=-2.4231884, final_local_scores=-2.6143355,
 pem=-6.7711987, similarity=-37.10444]

6: 7481141 Beyond_Conviction
 [winner: 7481141 Beyond_Conviction, pem_pos=0,
 final_scores=-0.1417911, final_local_scores=-0.023358755,
 pem=0.0, similarity=3.392641|
 gt: 7481141 Beyond_Conviction, pem_pos=0,
 final_scores=-0.1417911, final_local_scores=-0.023358755,
 pem=0.0, similarity=3.392641]

7: 6047259 What_Would_Jesus_Buy?
 [winner: 6047259 What_Would_Jesus_Buy?, pem_pos=0,
 final_scores=-9.75579e-05, final_local_scores=0.07672538,

pem=0.0, similarity=6.379351|
gt: 6047259 What_Would_Jesus_Buy?, pem_pos=0,
final_scores=-9.75579e-05, final_local_scores=0.07672538,
pem=0.0, similarity=6.379351]
8: 11046944 Taxi_to_the_Dark_Side
[winner: 11046944 Taxi_to_the_Dark_Side, pem_pos=0,
final_scores=-0.13514465, final_local_scores=-0.11535786,
pem=0.0, similarity=0.6472044|
gt: 11046944 Taxi_to_the_Dark_Side, pem_pos=0,
final_scores=-0.13514465, final_local_scores=-0.11535786,
pem=0.0, similarity=0.6472044]
9: 3700127 Exorcism
[winner: 3700127 Exorcism, pem_pos=0,
final_scores=-0.69650394, final_local_scores=-0.67887384,
pem=-0.0032791556, similarity=-16.151405|
gt: 3700127 Exorcism, pem_pos=0,
final_scores=-0.69650394, final_local_scores=-0.67887384,
pem=-0.0032791556, similarity=-16.151405]
10: 19649217 Kicking_It
[winner: 19649217 Kicking_It, pem_pos=0,
final_scores=-0.09573808, final_local_scores=-0.053860694,
pem=0.0, similarity=2.4824023|
gt: 19649217 Kicking_It, pem_pos=0,
final_scores=-0.09573808, final_local_scores=-0.053860694,
pem=0.0, similarity=2.4824023]
1: 3434750 United_States
[pem_pos=0, final_scores=0.43087012,
final_local_scores=0.4282149, pem=-0.12994486,
similarity=17.575161]
2: 2189889 Ward_Melville_High_School
[pem_pos=0, final_scores=0.051377848,
final_local_scores=0.13260013, pem=0.0, similarity=8.046765]
3: 23922 Princeton_University
[pem_pos=0, final_scores=0.3497408,
final_local_scores=0.37749112, pem=0.0, similarity=15.3548]
4: 3356 Bill_Clinton
[pem_pos=0, final_scores=0.48167074,
final_local_scores=0.50124496, pem=0.0, similarity=19.047863]
5: 723301 Morgan_Spurlock
[pem_pos=0, final_scores=0.21982978,
final_local_scores=0.22654164, pem=0.0, similarity=10.850167]
6: 3434750 United_States
[pem_pos=0, final_scores=0.5380507,
final_local_scores=0.54529345, pem=-0.073457584,
similarity=20.761831]
7: 4210630 Alex_Gibney
[pem_pos=0, final_scores=0.24655114,
final_local_scores=0.27740085, pem=0.0, similarity=12.367908]
8: 25445 Romania
[pem_pos=0, final_scores=0.55002445,
final_local_scores=0.64445823, pem=-0.21241479,
similarity=24.476776]

9: 29405 Sikh
[pem_pos=0, final_scores=0.059482798,
final_local_scores=0.13820776, pem=-0.047280814,
similarity=8.471229]
10: 21883824 Arizona
[pem_pos=0, final_scores=0.5569262,
final_local_scores=0.62840956, pem=-0.22599088,
similarity=24.071682]
11: 23943448 Sundance_Film_Festival
[pem_pos=0, final_scores=0.33106273,
final_local_scores=0.30130592, pem=0.0, similarity=13.081283]
12: 206705 Colin_Farrell
[pem_pos=0, final_scores=0.16518317,
final_local_scores=0.1772787, pem=0.0, similarity=9.380063]
13: 52780 U2
[pem_pos=0, final_scores=0.3079114,
final_local_scores=0.33222315, pem=0.0, similarity=14.003914]
14: 6310 Columbia_University
[pem_pos=0, final_scores=0.28049535,
final_local_scores=0.28068313, pem=0.0, similarity=12.4658575]
1: 7511294 Todd_Wider
[pem_pos=0, final_scores=0.14314424,
final_local_scores=0.17330542, pem=0.0, similarity=9.261493]
2: 151921 Emmy_Award
[pem_pos=0, final_scores=0.2862692,
final_local_scores=0.2534156, pem=0.0, similarity=11.652141]
3: 645042 New_York_City
[pem_pos=1, final_scores=0.45834845,
final_local_scores=0.43132222, pem=-1.7906687,
similarity=26.699163]
4: 645042 New_York_City
[pem_pos=0, final_scores=0.6855406,
final_local_scores=0.6794952, pem=0.0, similarity=24.367205]
5: 7511294 Todd_Wider
[pem_pos=0, final_scores=0.13102107,
final_local_scores=0.16006267, pem=0.0, similarity=8.8663025]
6: 23332 Pennsylvania
[pem_pos=0, final_scores=0.5894692,
final_local_scores=0.6202003, pem=-0.11810702,
similarity=23.24001]
7: 159846 MSNBC
[pem_pos=0, final_scores=0.20815115,
final_local_scores=0.20613047, pem=0.0, similarity=10.241057]
8: 20395872 Oprah_Winfrey
[pem_pos=0, final_scores=0.3350863,
final_local_scores=0.33632216, pem=0.0, similarity=14.126237]
9: 7511294 Todd_Wider
[pem_pos=0, final_scores=0.05820386,
final_local_scores=0.08052096, pem=0.0, similarity=6.492619]
10: 324 Academy_Awards
[pem_pos=0, final_scores=0.15964513,
final_local_scores=0.12848225, pem=-0.04133388,

```
similarity=8.14866]
11: 151921 Emmy_Award
    [pem_pos=0, final_scores=0.12470357,
final_local_scores=0.07692973, pem=-0.041731462,
similarity=6.612392]
```

Sample: b'Tony_Drake'

[1 Tony Drake 1] is an , best known for his appearances in the short-lived 1 Wrestling Society X . He currently competes on the 2 independent circuit using the ring name [2 Babi Slymm 2] . A and competitor in and track at high school , [3 Drake 3] began to train as a wrestler at the invitation of a friend operating a small [4 Californian 4] wrestling promotion in [5 Oakland 5] . [6 Drake 6] has competed for numerous organisations in the [7 California 7] area , including [1 1 Xtreme Pro Wrestling 1] under the ring name [8 Saleem Jihad 8] , Ultimate Pro Wrestling , [9 NWA Pro Wrestling 9] and [2 2 Pro Wrestling Guerrilla 2] . [10 Drake 10] received his greatest exposure competing for [3 3 MTV 3] 's 3 Wrestling Society X , in which he formed a team with [11 Ruckus 11] named Keepin ' It Gangsta . On March 29 , 2004 , [12 Drake 12] won the inaugural when he defeated [4 4 Adam Pearce 4] at the in . [13 Drake 13] later feuded at AWS with both [5 5 Al Katrazz 5] and the Plague . On January 26 , 2007 , [14 Drake 14] joined [6 6 Naruki Doi 6] 's Muscle Outlaw ' z faction in the [15 Japanese 15] [7 7 Dragon Gate 7] promotion . Using the name [16 Muscle Gang 16] , he became part of a tag team with fellow faction member [8 8 Cyber Kong 8] . The team was short-lived , however , as [17 Kong 17] soon left the Outlaw ' z to join the 4 New Hazard

```
1: 2239039 Xtreme_Pro_Wrestling
2: 1428870 Pro_Wrestling_Guerrilla
3: 18856 MTV
4: 5924097 Adam_Pearce
5: 20691330 Al_Katrazz
6: 8943483 Naruki_Doi
7: 2267485 Dragon_Gate
8: 11857324 Cyber_Kong
```

```
1: 4047459 Wrestling_Society_X
```

```
[winner: 4047459 Wrestling_Society_X, pem_pos=0,
final_scores=-0.08740017, final_local_scores=-0.12116572,
pem=0.0, similarity=0.4738862]
```

```
gt: 4047459 Wrestling_Society_X, pem_pos=0,
final_scores=-0.08740017, final_local_scores=-0.12116572,
pem=0.0, similarity=0.4738862 ]
```

```
2: 885498 Independent_circuit
```

```
[winner: 2313000 The_Independent_(Acadiana), pem_pos=9,
final_scores=-1.1288614, final_local_scores=-1.0000825,
pem=-2.9957323, similarity=-9.463488]
```

```
RECALL MISS: 885498 Independent_circuit]
```

```
[winner: 346001 Telecommunication_circuit, pem_pos=2,
final_scores=-1.3644266, final_local_scores=-1.2936902,
```


pem=-1.8410275, similarity=-24.504786|
RECALL MISS: 885498 Independent_circuit]
3: 4047459 Wrestling_Society_X
[winner: 4047459 Wrestling_Society_X, pem_pos=0,
final_scores=-0.041403905, final_local_scores=-0.07092178,
pem=0.0, similarity=1.9732656|
gt: 4047459 Wrestling_Society_X, pem_pos=0,
final_scores=-0.041403905, final_local_scores=-0.07092178,
pem=0.0, similarity=1.9732656]
4: 11189801 New_Hazard
[winner: 11189801 New_Hazard, pem_pos=0,
final_scores=-0.02297309, final_local_scores=0.014966056,
pem=0.0, similarity=4.5363297|
gt: 11189801 New_Hazard, pem_pos=0,
final_scores=-0.02297309, final_local_scores=0.014966056,
pem=0.0, similarity=4.5363297]
1: 2239039 Xtreme_Pro_Wrestling
[pem_pos=0, final_scores=0.13937421,
final_local_scores=0.13071385, pem=0.0, similarity=7.990475]
2: 1428870 Pro_Wrestling_Guerrilla
[pem_pos=0, final_scores=0.16202907,
final_local_scores=0.13378489, pem=0.0, similarity=8.082121]
3: 18856 MTV
[pem_pos=0, final_scores=0.2492763,
final_local_scores=0.27322862, pem=-0.04743405,
similarity=12.501353]
4: 5924097 Adam_Pearce
[pem_pos=0, final_scores=0.21151029,
final_local_scores=0.19987029, pem=0.0, similarity=10.054241]
5: 20691330 Al_Katrass
[pem_pos=0, final_scores=0.24142675,
final_local_scores=0.26676482, pem=0.0, similarity=12.050508]
6: 8943483 Naruki_Doi
[pem_pos=0, final_scores=0.2538051,
final_local_scores=0.28191584, pem=0.0, similarity=12.502644]
7: 2267485 Dragon_Gate
[pem_pos=0, final_scores=0.16929166,
final_local_scores=0.15407911, pem=-0.008928643,
similarity=8.736297]
8: 11857324 Cyber_Kong
[pem_pos=0, final_scores=0.19671218,
final_local_scores=0.23146784, pem=0.0, similarity=10.997175]
1: 9201241 Babi_Slymm
[pem_pos=0, final_scores=0.105823204,
final_local_scores=0.055850565, pem=0.0, similarity=5.7564054]
2: 9201241 Babi_Slymm
[pem_pos=0, final_scores=0.28995055,
final_local_scores=0.25698167, pem=0.0, similarity=11.758558]
3: 9201241 Babi_Slymm
[pem_pos=0, final_scores=0.10416339,
final_local_scores=0.050410837, pem=0.0, similarity=5.594073]
4: 5407 California

[pem_pos=0, final_scores=0.4476974,
 final_local_scores=0.53924394, pem=-0.093090415,
 similarity=20.688068]
 5: 50548 Oakland, California
 [pem_pos=0, final_scores=0.42212647,
 final_local_scores=0.4858585, pem=-0.4997394,
 similarity=21.30636]
 6: 9201241 Babi_Slymm
 [pem_pos=0, final_scores=0.097434506,
 final_local_scores=0.043060526, pem=0.0, similarity=5.374725]
 7: 5407 California
 [pem_pos=0, final_scores=0.48800856,
 final_local_scores=0.5832777, pem=-0.08662026,
 similarity=21.966938]
 8: 9201241 Babi_Slymm
 [pem_pos=0, final_scores=0.19033141,
 final_local_scores=0.14816293, pem=0.0, similarity=8.51119]
 9: 305917 National_Wrestling_Alliance
 [pem_pos=0, final_scores=0.20947538,
 final_local_scores=0.18411466, pem=0.0, similarity=9.584062]
 10: 21466444 Drake_(rapper)
 [pem_pos=1, final_scores=0.20014654,
 final_local_scores=0.21509954, pem=-0.46896923,
 similarity=13.05904]
 11: 5544008 Ruckus_(wrestler)
 [pem_pos=0, final_scores=0.04724823,
 final_local_scores=0.025137156, pem=-0.573264,
 similarity=7.957354]
 12: 21466444 Drake_(rapper)
 [pem_pos=1, final_scores=0.1086614,
 final_local_scores=0.11516592, pem=-0.46896923,
 similarity=10.076821]
 13: 9201241 Babi_Slymm
 [pem_pos=0, final_scores=0.24596618,
 final_local_scores=0.20530882, pem=0.0, similarity=10.216537]
 14: 9201241 Babi_Slymm
 [pem_pos=0, final_scores=0.16855554,
 final_local_scores=0.120749444, pem=0.0, similarity=7.6931176]
 15: 15573 Japan
 [pem_pos=0, final_scores=0.36676574,
 final_local_scores=0.46506685, pem=-0.99076414,
 similarity=23.356165]
 16: 9201241 Babi_Slymm
 [pem_pos=0, final_scores=0.04396434,
 final_local_scores=-0.017085709, pem=0.0,
 similarity=3.5798414]
 17: 11857324 Cyber_Kong
 [pem_pos=0, final_scores=0.10958259,
 final_local_scores=0.13629201, pem=0.0, similarity=8.156939]

Sample: b'Furanocoumarin'

Furanocoumarins , or furocoumarins , are a class of 1 organic

chemical compounds produced by a variety of plants . They are biosynthesized partly through the and the 2 mevalonate pathway , which is biosynthesized by a coupling of 3 dimethylallyl pyrophosphate -LRB- DMAPP -RRB- and 4 7-hydroxycoumarin -LRB- umbelliferone -RRB- . The chemical structure of furanocoumarins consists of a 5 furan ring fused with 6 coumarin . The furan may be fused in different ways producing several 7 isomers . The compounds that form the core structure of the two most common isomers are 8 psoralen and 9 angelicin . Derivatives of these two core structures are referred to respectively as linear and angular furanocoumarins . Many furanocoumarins are toxic and are produced by plants as a defense mechanism against various types of predators ranging from 10 insects to 11 mammals . This class of 12 phytochemical is responsible for the 13 phytophotodermatitis seen in exposure to the juices of the wild 14 parsnip and the . There are also reports of these chemicals in gas plants -LRB- -RRB- . Furanocoumarins have other biological effects as well . For example , in humans , 15 bergamottin and are responsible for the ' ' grapefruit juice effect

1: 22203 Organic_compound

[winner: 21347411 Chemical_compound, pem_pos=0, final_scores=-1.2598767, final_local_scores=-1.1036367, pem=-0.033664096, similarity=-28.661938]

gt: 22203 Organic_compound, pem_pos=7, final_scores=-2.1494157, final_local_scores=-2.075323, pem=-4.844468, similarity=-31.497097]

[winner: 13153591 Soil_organic_matter, pem_pos=11, final_scores=-1.3596603, final_local_scores=-1.212635, pem=-2.1917012, similarity=-20.178925]

gt: 22203 Organic_compound, pem_pos=0, final_scores=-1.4787501, final_local_scores=-1.3427227, pem=-1.6342998, similarity=-27.09223]

[winner: 14423824

Chemical_tests_in_mushroom_identification, pem_pos=11, final_scores=-1.4449179, final_local_scores=-1.3057661, pem=-2.3978953, similarity=-21.836824]

RECALL MISS: 22203 Organic_compound]

2: 619938 Mevalonate_pathway

[winner: 10941208 Mevalonate_kinase, pem_pos=0, final_scores=-0.7354391, final_local_scores=-0.53076833, pem=-0.08612919, similarity=-11.281089]

RECALL MISS: 619938 Mevalonate_pathway]

[winner: 20941 Metabolic_pathway, pem_pos=0, final_scores=-1.3421308, final_local_scores=-1.1934867, pem=-0.5229653, similarity=-28.68234]

RECALL MISS: 619938 Mevalonate_pathway]

3: 1476233 Dimethylallyl_pyrophosphate

[winner: 1476233 Dimethylallyl_pyrophosphate, pem_pos=0, final_scores=-1.0230805, final_local_scores=-0.8449727, pem=0.0, similarity=-21.12596]

gt: 1476233 Dimethylallyl_pyrophosphate, pem_pos=0,
final_scores=-1.0230805, final_local_scores=-0.8449727,
pem=0.0, similarity=-21.12596]
4: 2397690 Umbelliferone
[winner: 2397690 Umbelliferone, pem_pos=0,
final_scores=-0.7491989, final_local_scores=-0.54579884,
pem=0.0, similarity=-12.198013]
gt: 2397690 Umbelliferone, pem_pos=0,
final_scores=-0.7491989, final_local_scores=-0.54579884,
pem=0.0, similarity=-12.198013]
5: 546505 Furan
[winner: 546505 Furan, pem_pos=0, final_scores=-1.154,
final_local_scores=-0.9879825, pem=0.0,
similarity=-25.393656]
gt: 546505 Furan, pem_pos=0, final_scores=-1.154,
final_local_scores=-0.9879825, pem=0.0,
similarity=-25.393656]
6: 1116938 Coumarin
[winner: 1116938 Coumarin, pem_pos=0,
final_scores=-1.0777811, final_local_scores=-0.9047249,
pem=0.0, similarity=-22.909082]
gt: 1116938 Coumarin, pem_pos=0,
final_scores=-1.0777811, final_local_scores=-0.9047249,
pem=0.0, similarity=-22.909082]
7: 20648024 Isomer
[winner: 20648024 Isomer, pem_pos=0,
final_scores=-1.1143582, final_local_scores=-0.94467986,
pem=-0.03369239, similarity=-23.918194]
gt: 20648024 Isomer, pem_pos=0, final_scores=-1.1143582,
final_local_scores=-0.94467986, pem=-0.03369239,
similarity=-23.918194]
8: 2594155 Psoralen
[winner: 2594155 Psoralen, pem_pos=0,
final_scores=-0.88381195, final_local_scores=-0.69284314,
pem=0.0, similarity=-16.58611]
gt: 2594155 Psoralen, pem_pos=0,
final_scores=-0.88381195, final_local_scores=-0.69284314,
pem=0.0, similarity=-16.58611]
9: 30962869 Angelicin
[winner: 30962869 Angelicin, pem_pos=0,
final_scores=-0.53608906, final_local_scores=-0.31300864,
pem=0.0, similarity=-5.251089]
gt: 30962869 Angelicin, pem_pos=0,
final_scores=-0.53608906, final_local_scores=-0.31300864,
pem=0.0, similarity=-5.251089]
10: 23366462 Insect
[winner: 20812800 Paul_Insect, pem_pos=2,
final_scores=-0.8046396, final_local_scores=-0.6063593,
pem=-1.0986123, similarity=-8.030834]
gt: 23366462 Insect, pem_pos=0, final_scores=-1.0807796,
final_local_scores=-0.90800023, pem=-0.0069361115,
similarity=-22.969105]

11: 18838 Mammal
 [winner: 18838 Mammal, pem_pos=0,
 final_scores=-1.0567362, final_local_scores=-0.88173664,
 pem=-0.0016966542, similarity=-22.21384|
 gt: 18838 Mammal, pem_pos=0, final_scores=-1.0567362,
 final_local_scores=-0.88173664, pem=-0.0016966542,
 similarity=-22.21384]
 12: 472231 Phytochemical
 [winner: 472231 Phytochemical, pem_pos=0,
 final_scores=-1.1483848, final_local_scores=-0.9818488,
 pem=-0.002166829, similarity=-25.19883|
 gt: 472231 Phytochemical, pem_pos=0,
 final_scores=-1.1483848, final_local_scores=-0.9818488,
 pem=-0.002166829, similarity=-25.19883]
 13: 17672651 Phytophotodermatitis
 [winner: 17672651 Phytophotodermatitis, pem_pos=0,
 final_scores=-0.80585045, final_local_scores=-0.60768205,
 pem=0.0, similarity=-14.044731|
 gt: 17672651 Phytophotodermatitis, pem_pos=0,
 final_scores=-0.80585045, final_local_scores=-0.60768205,
 pem=0.0, similarity=-14.044731]
 14: 335507 Parsnip
 [winner: 335507 Parsnip, pem_pos=0,
 final_scores=-1.2636073, final_local_scores=-1.1077118,
 pem=0.0, similarity=-28.966616|
 gt: 335507 Parsnip, pem_pos=0, final_scores=-1.2636073,
 final_local_scores=-1.1077118, pem=0.0,
 similarity=-28.966616]
 15: 10033995 Bergamottin
 [winner: 10033995 Bergamottin, pem_pos=0,
 final_scores=-0.8113928, final_local_scores=-0.6137362,
 pem=0.0, similarity=-14.2254|
 gt: 10033995 Bergamottin, pem_pos=0,
 final_scores=-0.8113928, final_local_scores=-0.6137362,
 pem=0.0, similarity=-14.2254]

Sample: b'Vijaya_Dasa'

Vijaya Dasa -LRB- -RRB- -LRB- 1682 -- 1755 -RRB- -LRB- or Sri Vijaya Dasa -RRB- was a prominent saint from the [1 1 Har idasa 1] tradition of [1 1 Karnataka 1] , [2 2 India 2] in the 18th century , and a scholar of the 2 Dvaita philosophical tradition . Along with contemporary haridasa saints such as [2 Gopala Dasa 2] , Helevankatte Giriamma , 3 J agannatha Dasa and Prasanna Venkata Dasa , he propagated the virtues of the philosophy of across [3 South India 3] through devotional songs called devaranama written in the [4 4 Kannada 4] language . An integral part of Kannada devotional literature , these compositions in praise of the [5 Hindu 5] god [3 3 Vishnu 3] are called dasara padagalu -LRB- compositions of the dasas -RRB- . These compositions can be more specifically categorized as keertanas , suladis , ugabhogas , and simply padas . They were easy to sing to the

accompaniment of a musical instrument and dealt with 5 bhakti -LRB- devotion -RRB- and the virtues of a pious life . Vijaya Dasa was born in a poor [4 4 Brahmin 4] family in [5 5 Cheekalaparvi 5] in [6 6 Manvi 6] taluk of [6 6 Raichur 6] district , [7 Karnataka 7] state . His parents were Sri Srinivasappa and Kusamma . He left home at a young age and studied [7 7 Sanskrit 7] in [8 8 Varanasi 8] for four years . Later when his marriage was troubled by poverty , he went back to [8 Varanasi 8] where he became a scholar . One night , he had a dream in which the 16th century [9 9 Carnatic 9] composer and wandering saint [10 10 Purandara Dasa 10] initiated him into the [9 Haridasa 9] tradition and gave him the -LRB- ankita nama -RRB- Vijaya Vittala . From that day he was called Vijaya Dasa -LRB- dasa lit , slave to [10 God 10] -RRB- , and dedicated his life to spreading 7 Dvaita teachings . His 25,000 extant compositions earned him the title Dasa Shrestha -LRB- noble among the dasas -RRB- . His compositions which comprise of use of many [11 Sanskrit 11] words come under the category of [12 Kalasha 12] and Urasu creations and are considered an important component of [13 8 Kannada 13] literature -LRB- Kannada Sahitya -RRB- . His purported miracles include calming the [14 Ganges 14] , entering it without getting wet , preventing a woman from committing suicide , resurrecting his son , and making a uneducated man speak difficult [15 Sanskrit 15] proficiently . He is among the group credited with starting the practice of singing devotional songs while walking up the [16 9 Tirumala 16] hills in modern [11 11 Andhra Pradesh 11] . These hills are the location of the [12 12 Tirumala Venkateswara Temple 12] , an important pilgrim locations for [17 Vaishnava 10 Hindu 17]

1: 16880 Karnataka

2: 14533 India

3: 19334491 Vishnu

4: 244747 Brahmin

5: 37926822 Cheekalaparvi

6: 3357190 Manvi

7: 27698 Sanskrit

8: 37613554 Varanasi

9: 164188 Carnatic_music

10: 510331 Purandara_Dasa

11: 2377 Andhra_Pradesh

12: 7865082 Tirumala_Venkateswara_Temple

1: 5559287 Haridasa

[winner: 17732773 Haridasa_Thakur, pem_pos=0, final_scores=0.16697358, final_local_scores=0.16754505, pem=0.0, similarity=9.089591]

gt: 5559287 Haridasa, pem_pos=1, final_scores=0.07564677, final_local_scores=0.04643275, pem=-0.0023696462, similarity=5.488245]

2: 1064133 Dvaita

[winner: 1064133 Dvaita, pem_pos=0, final_scores=-0.050781265, final_local_scores=-0.06320488,

pem=0.0, similarity=2.2035534|
gt: 1064133 Dvaita, pem_pos=0,
final_scores=-0.050781265, final_local_scores=-0.06320488,
pem=0.0, similarity=2.2035534]
3: 21155310 Jagannatha_Dasa
[winner: 21155310 Jagannatha_Dasa, pem_pos=0,
final_scores=-0.041438743, final_local_scores=-0.046899036,
pem=0.0, similarity=2.6901522|
gt: 21155310 Jagannatha_Dasa, pem_pos=0,
final_scores=-0.041438743, final_local_scores=-0.046899036,
pem=0.0, similarity=2.6901522]
4: 37445 Kannada_language
[winner: 37445 Kannada_language, pem_pos=0,
final_scores=0.15691216, final_local_scores=0.10768928,
pem=-0.025031967, similarity=7.4395037|
gt: 37445 Kannada_language, pem_pos=0,
final_scores=0.15691216, final_local_scores=0.10768928,
pem=-0.025031967, similarity=7.4395037]
[winner: 18910 Markup_language, pem_pos=4,
final_scores=-1.7726346, final_local_scores=-1.8155379,
pem=-3.3860023, similarity=-31.67595|
RECALL MISS: 37445 Kannada_language]
5: 175567 Bhakti
[winner: 175567 Bhakti, pem_pos=0,
final_scores=-1.0597523, final_local_scores=-1.2163486,
pem=-0.07180528, similarity=-31.818064|
gt: 175567 Bhakti, pem_pos=0, final_scores=-1.0597523,
final_local_scores=-1.2163486, pem=-0.07180528,
similarity=-31.818064]
6: 1142301 Raichur_district
[winner: 4763118 Raichur, pem_pos=0,
final_scores=0.29427165, final_local_scores=0.31077507,
pem=0.0, similarity=13.363861|
gt: 1142301 Raichur_district, pem_pos=1,
final_scores=-0.028847337, final_local_scores=-0.045950413,
pem=-1.0301862, similarity=8.320774]
[winner: 13974770 Districts_of_Burma, pem_pos=15,
final_scores=-1.2757962, final_local_scores=-1.1813874,
pem=-3.551226, similarity=-11.853128|
RECALL MISS: 1142301 Raichur_district]
7: 1064133 Dvaita
[winner: 1064133 Dvaita, pem_pos=0,
final_scores=-0.040747896, final_local_scores=-0.052244946,
pem=0.0, similarity=2.5306196|
gt: 1064133 Dvaita, pem_pos=0,
final_scores=-0.040747896, final_local_scores=-0.052244946,
pem=0.0, similarity=2.5306196]
8: 860171 Kannada_literature
[winner: 37445 Kannada_language, pem_pos=0,
final_scores=0.047150865, final_local_scores=-0.016536653,
pem=-0.025031967, similarity=3.7323542|
gt: 860171 Kannada_literature, pem_pos=9,

final_scores=-0.5743861, final_local_scores=-0.65873075,
pem=-3.901045, similarity=5.6463623]
[winner: 3451925 Pali_literature, pem_pos=12,
final_scores=-1.7881502, final_local_scores=-1.9042861,
pem=-2.8903718, similarity=-37.01969]
RECALL MISS: 860171 Kannada_literature]
9: 222409 Tirumala
[winner: 7865082 Tirumala_Venkateswara_Temple, pem_pos=0,
final_scores=0.25961423, final_local_scores=0.22119522,
pem=0.0, similarity=10.690619]
gt: 222409 Tirumala, pem_pos=2,
final_scores=-0.30189735, final_local_scores=-0.36551884,
pem=-2.9979358, similarity=9.485142]
10: 13677 Hindu
[winner: 7866009 Vaishnavism, pem_pos=0,
final_scores=0.18011759, final_local_scores=0.11224456,
pem=0.0, similarity=7.4393144]
RECALL MISS: 13677 Hindu]
1: 16880 Karnataka
[pem_pos=0, final_scores=0.4166559,
final_local_scores=0.37669408, pem=0.0, similarity=15.331017]
2: 14533 India
[pem_pos=0, final_scores=0.61273485,
final_local_scores=0.6222705, pem=-0.13622136,
similarity=23.4003]
3: 19334491 Vishnu
[pem_pos=0, final_scores=0.0320379,
final_local_scores=-0.027695723, pem=-0.0047303685,
similarity=3.2889419]
4: 244747 Brahmin
[pem_pos=0, final_scores=0.17652802,
final_local_scores=0.108075395, pem=0.0, similarity=7.3148985]
5: 37926822 Cheekalaparvi
[pem_pos=0, final_scores=0.14925222,
final_local_scores=0.3038122, pem=0.0, similarity=13.1560755]
6: 3357190 Manvi
[pem_pos=0, final_scores=0.31455392,
final_local_scores=0.3579951, pem=0.0, similarity=14.773001]
7: 27698 Sanskrit
[pem_pos=0, final_scores=0.3160109,
final_local_scores=0.2810506, pem=0.0, similarity=12.476823]
8: 37613554 Varanasi
[pem_pos=0, final_scores=0.34732527,
final_local_scores=0.29759735, pem=0.0, similarity=12.970612]
9: 164188 Carnatic_music
[pem_pos=0, final_scores=0.12766479,
final_local_scores=0.101071194, pem=-0.010047376,
similarity=7.160518]
10: 510331 Purandara_Dasa
[pem_pos=0, final_scores=0.19160752,
final_local_scores=0.17842445, pem=0.0, similarity=9.414254]
11: 2377 Andhra_Pradesh

[pem_pos=0, final_scores=0.34944612,
final_local_scores=0.31082237, pem=0.0, similarity=13.3652725]
12: 7865082 Tirumala Venkateswara Temple
[pem_pos=0, final_scores=0.20025103,
final_local_scores=0.15634996, pem=0.0, similarity=8.755507]
1: 17732773 Haridasa Thakur
[pem_pos=0, final_scores=0.16697358,
final_local_scores=0.16754505, pem=0.0, similarity=9.089591]
2: 14138620 Gopala Dasa
[pem_pos=0, final_scores=0.1825255,
final_local_scores=0.2175326, pem=0.0, similarity=10.58132]
3: 678583 South India
[pem_pos=0, final_scores=0.3660788,
final_local_scores=0.3104627, pem=0.0, similarity=13.35454]
4: 37445 Kannada language
[pem_pos=0, final_scores=0.15691216,
final_local_scores=0.10768928, pem=-0.025031967,
similarity=7.4395037]
5: 13677 Hindu
[pem_pos=0, final_scores=0.12605213,
final_local_scores=0.052557677, pem=-0.37618497,
similarity=7.7038913]
6: 4763118 Raichur
[pem_pos=0, final_scores=0.29427165,
final_local_scores=0.31077507, pem=0.0, similarity=13.363861]
7: 16880 Karnataka
[pem_pos=0, final_scores=0.34104288,
final_local_scores=0.29144314, pem=0.0, similarity=12.786958]
8: 37613554 Varanasi
[pem_pos=0, final_scores=0.36546165,
final_local_scores=0.31740856, pem=0.0, similarity=13.561817]
9: 17732773 Haridasa Thakur
[pem_pos=0, final_scores=0.10418151,
final_local_scores=0.098954245, pem=0.0, similarity=7.042705]
10: 5042765 God
[pem_pos=0, final_scores=0.51623976,
final_local_scores=0.5720658, pem=-0.18657415,
similarity=22.175915]
11: 27698 Sanskrit
[pem_pos=0, final_scores=0.14456077,
final_local_scores=0.093767405, pem=0.0, similarity=6.8879194]
12: 437071 Nuristani people
[pem_pos=1, final_scores=0.039064184,
final_local_scores=0.09858203, pem=-0.6871948,
similarity=10.76867]
13: 37445 Kannada language
[pem_pos=0, final_scores=0.047150865,
final_local_scores=-0.016536653, pem=-0.025031967,
similarity=3.7323542]
14: 12448 Ganges
[pem_pos=0, final_scores=0.2697422,
final_local_scores=0.24625632, pem=-0.03463801,

similarity=11.62686]
15: 27698 Sanskrit
[pem_pos=0, final_scores=0.19237702,
final_local_scores=0.1459994, pem=0.0, similarity=8.446627]
16: 7865082 Tirumala Venkateswara Temple
[pem_pos=0, final_scores=0.25961423,
final_local_scores=0.22119522, pem=0.0, similarity=10.690619]
17: 7866009 Vaishnavism
[pem_pos=0, final_scores=0.18011759,
final_local_scores=0.11224456, pem=0.0, similarity=7.4393144]

Sample: b'Hittin'

' the Trail for Hallelujah Land ' Hittin ' the Trail for Hallelujah Land ' is a [1 1 Merrie Melodies 1] 1 animated cartoon directed by -LRB- uncredited -RRB- , produced by [2 2 Leon Schlesinger Productions 2] , and released to theatres on November 28 , 1931 by and . The minimal storyline centers on the plucky 's efforts to rescue his girlfriend and a doglike 2 Uncle Tom from perilous predicaments and villains . The short 's stereotypical portrayal of black characters prompted [3 3 United Artists 3] to withhold it from syndication in 1968 , making it one of the infamous [4 4 Censored Eleven 4] . It was the second [5 5 Merrie Melodies 5] film to be released . Hittin ' the Trail for Hallelujah Land has a rudimentary plot , unlike most of the [1 Merrie Melodies 1] of the time , which barely have any plot at all . The cartoon stars the [6 6 Mickey Mouse 6] - esque , his girlfriend Fluffy , and a canine 3 Uncle Tom . The film opens with a singing steamboat dancing down a river . On the deck , three 4 blackface caricatures play the song for which the short is named on the harmonica , banjo , and bones . Meanwhile , Uncle Tom drives Fluffy toward the boat by donkey cart . The scene shifts to Piggy the riverboat captain in a sequence reminiscent of [2 Disney 2] 's 1928 film [7 7 Steamboat Willie 7] . Fluffy joins the frolicking steamboat passengers and reunites with her boyfriend , but during the revelry , Piggy falls overboard . The pig has a run in with an alligator , but he makes it back to the boat . Meanwhile , Uncle Tom 's donkey bucks him into a cemetery . There , in a variation on a stock gag featuring a superstitious black man , he is scared by three dancing skeletons reminiscent of those in [3 Disney 3] 's 1929 short 5 The Skeleton Dance . Tom escapes to the middle of the river , but a shoddy boat leaves him stranded and drowning . Piggy saves the day but not before a 6 vaudeville villain kidnaps his porcine paramour . Piggy captures the villain on a passing mail hook , leaving the villain tortured over a buzz saw . This short marks the second and last appearance of the characters Piggy and Fluffy . 7 Hittin ' the Trail for Hallelujah Land was released in [4 American 4] theaters on November 28 , 1931 by 8 Warner Bros. . . The cartoon 's copyright expired in 1965 , making it go into public domain . However , the cartoon has been withheld from

distribution since 1968 . At that time , [8 8 United Artists 8] owned the rights to most [5 Looney Tunes 5] and [6 Merrie Melodies 6] cartoons . Hittin ' the Trail for [7 Hallelujah 7] land and ten other cartoons were deemed to feature racist depictions of [8 African Americans 8] that were too integral to the films for simple cuts to make them palatable for modern audiences . The cartoon has never been released on laserdisc , home video , or DVD . These eleven cartoons make up the so-called 9 Censored Eleven

1: 199608 Merrie_Melodies

2: 1555666 Warner_Bros._Cartoons

3: 174319 United_Artists

4: 205061 Censored_Eleven

5: 199608 Merrie_Melodies

6: 20859 Mickey_Mouse

7: 18943962 Steamboat_Willie

8: 174319 United_Artists

1: 593 Animation

[winner: 684885 The_Inspector, pem_pos=0, final_scores=-1.1353046, final_local_scores=-1.2158284, pem=0.0, similarity=-32.193035]

gt: 593 Animation, pem_pos=3, final_scores=-2.1296623, final_local_scores=-2.3712587, pem=-2.4212937, similarity=-53.506004]

2: 51174 Uncle_Tom

[winner: 51174 Uncle_Tom, pem_pos=0, final_scores=-0.060222864, final_local_scores=-0.05614479, pem=0.0, similarity=2.4142404]

gt: 51174 Uncle_Tom, pem_pos=0, final_scores=-0.060222864, final_local_scores=-0.05614479, pem=0.0, similarity=2.4142404]

3: 51174 Uncle_Tom

[winner: 51174 Uncle_Tom, pem_pos=0, final_scores=-0.01848656, final_local_scores=0.008335218, pem=0.0, similarity=4.3384523]

gt: 51174 Uncle_Tom, pem_pos=0, final_scores=-0.01848656, final_local_scores=0.008335218, pem=0.0, similarity=4.3384523]

4: 89708 Blackface

[winner: 89708 Blackface, pem_pos=0, final_scores=-1.1229092, final_local_scores=-1.2632875, pem=-0.0023178477, similarity=-33.596703]

gt: 89708 Blackface, pem_pos=0, final_scores=-1.1229092, final_local_scores=-1.2632875, pem=-0.0023178477, similarity=-33.596703]

5: 3490551 The_Skeleton_Dance

[winner: 3490551 The_Skeleton_Dance, pem_pos=0, final_scores=-0.16343847, final_local_scores=-0.15008914, pem=0.0, similarity=-0.38924652]

gt: 3490551 The_Skeleton_Dance, pem_pos=0, final_scores=-0.16343847, final_local_scores=-0.15008914, pem=0.0, similarity=-0.38924652]

6: 48235 Vaudeville
 [winner: 48235 Vaudeville, pem_pos=0,
 final_scores=-0.76420474, final_local_scores=-0.87277156,
 pem=0.0, similarity=-21.955532|
 gt: 48235 Vaudeville, pem_pos=0,
 final_scores=-0.76420474, final_local_scores=-0.87277156,
 pem=0.0, similarity=-21.955532]

7: 11894702 Hittin'_the_Trail_for_Hallelujah_Land
 [winner: 20555890 Trail,_British_Columbia, pem_pos=2,
 final_scores=-0.19138598, final_local_scores=-0.11206129,
 pem=-1.4930917, similarity=8.865245|
 RECALL MISS: 11894702
 Hittin'_the_Trail_for_Hallelujah_Land]
 [winner: 334927 Hittin, pem_pos=0,
 final_scores=-0.28243953, final_local_scores=-0.19020256,
 pem=0.0, similarity=-1.5863112|
 RECALL MISS: 11894702
 Hittin'_the_Trail_for_Hallelujah_Land]
 [winner: 8462526 Hallelujah!(gospel_song), pem_pos=11,
 final_scores=-0.90861255, final_local_scores=-0.87295955,
 pem=-3.073107, similarity=-5.2491083|
 RECALL MISS: 11894702
 Hittin'_the_Trail_for_Hallelujah_Land]
 [winner: 16488452 Land_(song), pem_pos=10,
 final_scores=-1.1016558, final_local_scores=-1.0380912,
 pem=-3.1821997, similarity=-9.583703|
 RECALL MISS: 11894702
 Hittin'_the_Trail_for_Hallelujah_Land]
 [winner: 2302349 Vedanga, pem_pos=16,
 final_scores=-2.2484722, final_local_scores=-2.2214782,
 pem=-4.75359, similarity=-36.352867|
 RECALL MISS: 11894702
 Hittin'_the_Trail_for_Hallelujah_Land]

8: 34052 Warner_Bros.
 [winner: 727428 Jerry_Zucker_(film_director), pem_pos=0,
 final_scores=-0.7734842, final_local_scores=-0.8224892,
 pem=0.0, similarity=-20.455006|
 RECALL MISS: 34052 Warner_Bros.]
 [winner: 1354262 Karl_Warner, pem_pos=10,
 final_scores=-1.2490206, final_local_scores=-1.1799817,
 pem=-3.476092, similarity=-12.219762|
 gt: 34052 Warner_Bros., pem_pos=4,
 final_scores=-1.4356804, final_local_scores=-1.6633663,
 pem=-2.9466906, similarity=-29.523893]
 [winner: 4721653 Zyphe_language, pem_pos=12,
 final_scores=-2.038411, final_local_scores=-2.0258124,
 pem=-4.4188404, similarity=-32.334225|
 RECALL MISS: 34052 Warner_Bros.]

9: 205061 Censored_Eleven
 [winner: 205061 Censored_Eleven, pem_pos=0,
 final_scores=-0.108073965, final_local_scores=-0.15359046,
 pem=0.0, similarity=-0.4937326|

gt: 205061 Censored_Eleven, pem_pos=0,
final_scores=-0.108073965, final_local_scores=-0.15359046,
pem=0.0, similarity=-0.4937326]

1: 199608 Merrie_Melodies
[pem_pos=0, final_scores=0.12375562,
final_local_scores=0.082356855, pem=0.0, similarity=6.5474057]

2: 1555666 Warner_Bros._Cartoons
[pem_pos=0, final_scores=0.20189907,
final_local_scores=0.1643619, pem=0.0, similarity=8.994599]

3: 174319 United_Artists
[pem_pos=0, final_scores=0.32039976,
final_local_scores=0.29003495, pem=0.0, similarity=12.744934]

4: 205061 Censored_Eleven
[pem_pos=0, final_scores=0.030652717,
final_local_scores=-0.002052635, pem=0.0, similarity=4.028458]

5: 199608 Merrie_Melodies
[pem_pos=0, final_scores=0.09470062,
final_local_scores=0.05061868, pem=0.0, similarity=5.600275]

6: 20859 Mickey_Mouse
[pem_pos=0, final_scores=0.16302536,
final_local_scores=0.118952364, pem=0.0, similarity=7.6394887]

7: 18943962 Steamboat_Willie
[pem_pos=0, final_scores=0.05863397,
final_local_scores=0.04267302, pem=0.0, similarity=5.363161]

8: 174319 United_Artists
[pem_pos=0, final_scores=0.26288956,
final_local_scores=0.2272138, pem=0.0, similarity=10.870226]

1: 199608 Merrie_Melodies
[pem_pos=0, final_scores=0.041874245,
final_local_scores=-0.017227411, pem=0.0,
similarity=3.5756125]

2: 37398 The_Walt_Disney_Company
[pem_pos=0, final_scores=0.30630314,
final_local_scores=0.26521224, pem=-0.24767768,
similarity=13.351085]

3: 37398 The_Walt_Disney_Company
[pem_pos=0, final_scores=0.33016247,
final_local_scores=0.29127496, pem=-0.24767768,
similarity=14.128848]

4: 3434750 United_States
[pem_pos=0, final_scores=0.36439127,
final_local_scores=0.42934364, pem=-0.12994486,
similarity=17.608845]

5: 50287 Looney_Tunes
[pem_pos=0, final_scores=0.22022565,
final_local_scores=0.1616062, pem=0.0, similarity=8.912364]

6: 199608 Merrie_Melodies
[pem_pos=0, final_scores=0.18286024,
final_local_scores=0.1469197, pem=0.0, similarity=8.47409]

7: 2186423 Hallelujah_(Leonard_Cohen_song)
[pem_pos=0, final_scores=0.05803798,
final_local_scores=0.09832482, pem=-0.607928,

similarity=10.329928]
8: 2154 African_American
[pem_pos=0, final_scores=0.28391927,
final_local_scores=0.31211817, pem=0.0, similarity=13.403942]

Sample: b'John_Wertheim'

John V. Wertheim -LRB- born 12 February 1968 -RRB- is an [1 1 American 1] 1 lawyer and 2 politician who served as Chairman of the 3 Democratic Party -LRB- United States -RRB- of [2 2 New Mexico 2] from 2004 until 2007 . During that time , he also served on the Executive Committee of the [3 3 Democratic National Committee 3] . In 1996 , he was the [1 Democratic 1] nominee for the [4 4 United States House of Representatives 4] in against the incumbent Representative of the [2 4 Republican Party 2] -LRB- [3 United States 3] -RRB- . The treasurer for his congressional campaign was [5 5 Arvind A. Raichur 5] . Wertheim is a graduate of [6 6 Yale University 6] , where the [7 7 Yale Political Union 7] elected him President and then Speaker . Representing the [4 Yale Debate Association 4] , he and partner Matt Wolf won the 1990 5 World Universities Debating Championship held at the [8 8 Glasgow University Union 8] in [5 Scotland 5] , becoming the first [6 Americans 6] to do so . In 1990 , he and debate partner [9 9 Austan Goolsbee 9] placed second at the [10 10 APDA National Debating Championship 10] . At [7 Yale 7] , he was also a member of the secret society [11 11 Skull and Bones 11] . In his senior year of high school , 1986 , Wertheim won dual championships at the 6 National Speech and Debate Tournament , sponsored by the [12 1 2 National Forensic League 12] , in both 7 Foreign Extemporaneous Speaking and . Wertheim received his law degree from the [13 13 University of New Mexico School of Law 13]

1: 3434750 United_States

2: 21649 New_Mexico

3: 182450 Democratic_National_Committee

4: 19468510 United_States_House_of_Representatives

5: 26775848 Arvind_A._Raichur

6: 34273 Yale_University

7: 30871917 Yale_Political_Union

8: 645984 Glasgow_University_Union

9: 4388185 Austan_Goolsbee

10: 552857 American_Parliamentary_Debate_Association

11: 174638 Skull_and_Bones

12: 183860 National_Forensic_League

13: 18262051 University_of_New_Mexico_School_of_Law

1: 17541 Lawyer

[winner: 9327970 Guantanamo_Bay_attorneys, pem_pos=3,
final_scores=-1.0497139, final_local_scores=-1.0072074,
pem=-1.0986123, similarity=-19.99294]

gt: 17541 Lawyer, pem_pos=0, final_scores=-1.476031,
final_local_scores=-1.6591294, pem=-0.03994547,
similarity=-45.204792]

2: 55959 Politician

[winner: 55959 Politician, pem_pos=0,
final_scores=-1.2732644, final_local_scores=-1.3910215,
pem=-0.0811182, similarity=-36.980007|
gt: 55959 Politician, pem_pos=0,
final_scores=-1.2732644, final_local_scores=-1.3910215,
pem=-0.0811182, similarity=-36.980007]
3: 5043544 Democratic_Party_(United_States)
[winner: 5043544 Democratic_Party_(United_States),
pem_pos=0, final_scores=-0.10123131,
final_local_scores=-0.17586805, pem=-0.33323824,
similarity=0.6536609|
gt: 5043544 Democratic_Party_(United_States), pem_pos=0,
final_scores=-0.10123131, final_local_scores=-0.17586805,
pem=-0.33323824, similarity=0.6536609]
[winner: 3434750 United_States, pem_pos=0,
final_scores=-0.13082801, final_local_scores=-0.19723547,
pem=-0.073457584, similarity=-1.3967131|
RECALL MISS: 5043544 Democratic_Party_(United_States)]
4: 32070 Republican_Party_(United_States)
[winner: 3434750 United_States, pem_pos=0,
final_scores=0.23806386, final_local_scores=0.21425277,
pem=-0.073457584, similarity=10.882917|
RECALL MISS: 32070 Republican_Party_(United_States)]
[winner: 32070 Republican_Party_(United_States),
pem_pos=0, final_scores=0.23374684,
final_local_scores=0.19815022, pem=-0.16706309,
similarity=10.911426|
gt: 32070 Republican_Party_(United_States), pem_pos=0,
final_scores=0.23374684, final_local_scores=0.19815022,
pem=-0.16706309, similarity=10.911426]
5: 980824 World_Universities_Debating_Championship
[winner: 980824 World_Universities_Debating_Championship,
pem_pos=0, final_scores=0.020270154,
final_local_scores=0.055830166, pem=0.0,
similarity=5.7557964|
gt: 980824 World_Universities_Debating_Championship,
pem_pos=0, final_scores=0.020270154,
final_local_scores=0.055830166, pem=0.0,
similarity=5.7557964]
6: 449700 National_Speech_and_Debate_Tournament
[winner: 449700 National_Speech_and_Debate_Tournament,
pem_pos=0, final_scores=-0.32937157,
final_local_scores=-0.3143013, pem=0.0,
similarity=-5.2896647|
gt: 449700 National_Speech_and_Debate_Tournament,
pem_pos=0, final_scores=-0.32937157,
final_local_scores=-0.3143013, pem=0.0,
similarity=-5.2896647]
7: 447528 Foreign_Extemporaneous_Speaking
[winner: 447528 Foreign_Extemporaneous_Speaking,
pem_pos=0, final_scores=-0.24575482,
final_local_scores=-0.14486772, pem=0.0,

```

similarity=-0.23342857|
    gt: 447528 Foreign_Extemporaneous_Speaking, pem_pos=0,
final_scores=-0.24575482, final_local_scores=-0.14486772,
pem=0.0, similarity=-0.23342857 ]
1: 3434750 United_States
    [pem_pos=0, final_scores=0.62173766,
final_local_scores=0.6333581, pem=-0.12994486,
similarity=23.697044]
2: 21649 New_Mexico
    [pem_pos=0, final_scores=0.42140448,
final_local_scores=0.46811587, pem=-0.20017311,
similarity=19.147797]
3: 182450 Democratic_National_Committee
    [pem_pos=0, final_scores=0.39295655,
final_local_scores=0.3828925, pem=0.0, similarity=15.515989]
4: 19468510 United_States_House_of_Representatives
    [pem_pos=0, final_scores=0.15895395,
final_local_scores=0.12007341, pem=0.0, similarity=7.6729426]
5: 26775848 Arvind_A._Raichur
    [pem_pos=0, final_scores=0.094671175,
final_local_scores=0.16363296, pem=0.0, similarity=8.972847]
6: 34273 Yale_University
    [pem_pos=0, final_scores=0.40770835,
final_local_scores=0.39633214, pem=0.0, similarity=15.917054]
7: 30871917 Yale_Political_Union
    [pem_pos=0, final_scores=0.1590368,
final_local_scores=0.19054607, pem=0.0, similarity=9.775988]
8: 645984 Glasgow_University_Union
    [pem_pos=0, final_scores=0.13182272,
final_local_scores=0.16258496, pem=0.0, similarity=8.941572]
9: 4388185 Austan_Goolsbee
    [pem_pos=0, final_scores=0.32921463,
final_local_scores=0.3772241, pem=0.0, similarity=15.346832]
10: 552857 American_Parliamentary_Debate_Association
    [pem_pos=0, final_scores=0.10277693,
final_local_scores=0.1283406, pem=0.0, similarity=7.9196525]
11: 174638 Skull_and_Bones
    [pem_pos=1, final_scores=0.21490364,
final_local_scores=0.20745796, pem=-0.007452653,
similarity=10.321201]
12: 183860 National_Forensic_League
    [pem_pos=0, final_scores=0.1498905,
final_local_scores=0.15569839, pem=0.0, similarity=8.736063]
13: 18262051 University_of_New_Mexico_School_of_Law
    [pem_pos=0, final_scores=0.12799774,
final_local_scores=0.1591517, pem=0.0, similarity=8.839117]
1: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.4822362,
final_local_scores=0.4688729, pem=-0.076093495,
similarity=18.495625]
2: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.23374684,

```


final_local_scores=0.19815022, pem=-0.16706309,
similarity=10.911426]
3: 3434750 United_States
[pem_pos=0, final_scores=0.23806386,
final_local_scores=0.21425277, pem=-0.073457584,
similarity=10.882917]
4: 27765362 Yale_Debate_Association
[pem_pos=0, final_scores=0.13334851,
final_local_scores=0.15143749, pem=0.0, similarity=8.60891]
5: 26994 Scotland
[pem_pos=0, final_scores=0.47072095,
final_local_scores=0.5209089, pem=-0.22905664,
similarity=20.88032]
6: 3434750 United_States
[pem_pos=0, final_scores=0.24417804,
final_local_scores=0.22093159, pem=-0.5581965,
similarity=13.718311]
7: 34273 Yale_University
[pem_pos=0, final_scores=0.33641958,
final_local_scores=0.31846002, pem=-0.3216871,
similarity=15.34258]

Sample: b'Rainer_Polak'

[1 Rainer Polak 1] is an ethnomusicologist and 1 djembe drummer who has researched in the field of and written in the field of 2 ethnomusicology . [2 Rainer Polak 2] studied 3 social anthropology , , [3 4 Bambara 3] language , and 5 History of Africa from 1989 to 1996 at -LRB- 6 Germany -RRB- , and 7 jenbe music performance from 1991 until today in [1 1 Bamako 1] -LRB- 8 Mali -RRB- . All of his studies and work in [2 2 Bamako 2] were accomplished with the help of locally and traditionally minded drummers whose playing is presented in his book and the corresponding CD . Polak worked as a professional 9 jenbe player in [3 3 Bamako 3] for one year in 1997/98 , performing at well over a hundred traditional weddings , spirit possession dances and other celebrations on the basis of being hired by the late , most of the times , and occasionally by the late [4 4 Yamadu Bani Dunbia 4] , by , and by [5 5 Drissa Kone 5] . The ethnomusicological dissertation and book he wrote on that experience won the academic prize of the in 2003/04 . Polak ranks as an outstanding 10 jenbe soloist in [6 6 Germany 6] . As a teacher he has specialized in giving focused classes on , and master-classes in 11 jenbe

1: 56606 Bamako

2: 56606 Bamako

3: 56606 Bamako

4: 16742083 Yamadu_Bani_Dunbia

5: 13923075 Drissa_Kone

6: 11867 Germany

1: 326011 Djembe

[winner: 326011 Djembe, pem_pos=0,

final_scores=-0.92193794, final_local_scores=-0.9856305,
 pem=0.0, similarity=-25.323467|
 gt: 326011 Djembe, pem_pos=0, final_scores=-0.92193794,
 final_local_scores=-0.9856305, pem=0.0,
 similarity=-25.323467]
 2: 80077 Ethnomusicology
 [winner: 80077 Ethnomusicology, pem_pos=0,
 final_scores=-1.0761853, final_local_scores=-1.1647692,
 pem=0.0, similarity=-30.66932|
 gt: 80077 Ethnomusicology, pem_pos=0,
 final_scores=-1.0761853, final_local_scores=-1.1647692,
 pem=0.0, similarity=-30.66932]
 3: 33661613 Social_anthropology
 [winner: 33661613 Social_anthropology, pem_pos=0,
 final_scores=-1.1948996, final_local_scores=-1.2607007,
 pem=-0.0008506816, similarity=-33.527485|
 gt: 33661613 Social_anthropology, pem_pos=0,
 final_scores=-1.1948996, final_local_scores=-1.2607007,
 pem=-0.0008506816, similarity=-33.527485]
 4: 4565 Bambara_language
 [winner: 4565 Bambara_language, pem_pos=0,
 final_scores=0.08570026, final_local_scores=0.123344064,
 pem=-0.06515657, similarity=8.124878|
 gt: 4565 Bambara_language, pem_pos=0,
 final_scores=0.08570026, final_local_scores=0.123344064,
 pem=-0.06515657, similarity=8.124878]
 [winner: 10939 Formal_language, pem_pos=3,
 final_scores=-1.827584, final_local_scores=-1.8604,
 pem=-3.3610234, similarity=-33.150566|
 RECALL MISS: 4565 Bambara_language]
 5: 14099 History_of_Africa
 [winner: 14099 History_of_Africa, pem_pos=0,
 final_scores=0.0042491704, final_local_scores=0.03395684,
 pem=0.0, similarity=5.1030526|
 gt: 14099 History_of_Africa, pem_pos=0,
 final_scores=0.0042491704, final_local_scores=0.03395684,
 pem=0.0, similarity=5.1030526]
 6: 11867 Germany
 [winner: 11867 Germany, pem_pos=0,
 final_scores=-0.69275796, final_local_scores=-0.7725807,
 pem=-0.15529114, similarity=-18.121138|
 gt: 11867 Germany, pem_pos=0, final_scores=-0.69275796,
 final_local_scores=-0.7725807, pem=-0.15529114,
 similarity=-18.121138]
 7: 326011 Djembe
 [winner: 326011 Djembe, pem_pos=0,
 final_scores=-1.3589563, final_local_scores=-1.4630067,
 pem=0.0, similarity=-39.569324|
 gt: 326011 Djembe, pem_pos=0, final_scores=-1.3589563,
 final_local_scores=-1.4630067, pem=0.0,
 similarity=-39.569324]
 8: 19127 Mali

[winner: 19127 Mali, pem_pos=0, final_scores=-0.23466684,
 final_local_scores=-0.2985014, pem=-0.37842536,
 similarity=-2.7602282]
 gt: 19127 Mali, pem_pos=0, final_scores=-0.23466684,
 final_local_scores=-0.2985014, pem=-0.37842536,
 similarity=-2.7602282]
 9: 326011 Djembe
 [winner: 326011 Djembe, pem_pos=0,
 final_scores=-1.1344622, final_local_scores=-1.2177811,
 pem=0.0, similarity=-32.2513]
 gt: 326011 Djembe, pem_pos=0, final_scores=-1.1344622,
 final_local_scores=-1.2177811, pem=0.0, similarity=-32.2513]
 10: 326011 Djembe
 [winner: 326011 Djembe, pem_pos=0,
 final_scores=-1.2674749, final_local_scores=-1.3630772,
 pem=0.0, similarity=-36.58723]
 gt: 326011 Djembe, pem_pos=0, final_scores=-1.2674749,
 final_local_scores=-1.3630772, pem=0.0, similarity=-36.58723]
 11: 326011 Djembe
 [winner: 326011 Djembe, pem_pos=0,
 final_scores=-1.2045952, final_local_scores=-1.2943907,
 pem=0.0, similarity=-34.537487]
 gt: 326011 Djembe, pem_pos=0, final_scores=-1.2045952,
 final_local_scores=-1.2943907, pem=0.0,
 similarity=-34.537487]
 1: 56606 Bamako
 [pem_pos=0, final_scores=0.29164982,
 final_local_scores=0.26507872, pem=0.0, similarity=12.000191]
 2: 56606 Bamako
 [pem_pos=0, final_scores=0.35861945,
 final_local_scores=0.33823287, pem=0.0, similarity=14.183256]
 3: 56606 Bamako
 [pem_pos=0, final_scores=0.4077086,
 final_local_scores=0.39185524, pem=0.0, similarity=15.783455]
 4: 16742083 Yamadu_Bani_Dunbia
 [pem_pos=0, final_scores=0.09693365,
 final_local_scores=0.16744092, pem=0.0, similarity=9.086484]
 5: 13923075 Drissa_Kone
 [pem_pos=0, final_scores=0.06547035,
 final_local_scores=0.1270805, pem=0.0, similarity=7.8820486]
 6: 11867 Germany
 [pem_pos=0, final_scores=0.6304203,
 final_local_scores=0.7033405, pem=-0.15529114,
 similarity=25.923294]
 1: 16760510 Rainer_Polak
 [pem_pos=0, final_scores=0.05615507,
 final_local_scores=0.0810011, pem=0.0, similarity=6.5069475]
 2: 16760510 Rainer_Polak
 [pem_pos=0, final_scores=0.1531853,
 final_local_scores=0.1869919, pem=0.0, similarity=9.669924]
 3: 4565 Bambara_language
 [pem_pos=0, final_scores=0.08570026,

final_local_scores=0.123344064, pem=-0.06515657,
similarity=8.124878]

Sample: b'Hagop_Sandaldjian'

[1 Hagop Sandaldjian 1] -LRB- 1931 -- 1990 -RRB- was an Egyptian-born [1 1 Armenian American 1] musician and microminiature 1 sculptor , best known for his tiny art pieces displayed at the [2 2 Museum of Jurassic Technology 2] in . Sandaldjian 's creations included a carving of [3 3 Mount Ararat 3] on a grain of rice ; a 2 crucifix in which a minute golden figure of [4 4 Jesus 4] hangs upon a 3 cross made from a bisected strand of Sandaldjian 's own hair ; and recreations of [2 Disney 2] figures -LRB- 4 Snow White and the Seven Dwarfs or [5 5 Mickey Mouse 5] , for example -RRB- or historical figures -LRB- such as [6 6 Napoleon 6] or [7 7 Pope John Paul II 7] -RRB- presented in the eye or on the tip of a needle . Sandaldjian was born in [8 8 Alexandria 8] , [9 9 Egypt 9] , and went to the [10 10 Soviet Union 10] to study music in [11 11 Yerevan 11] , [12 12 Armenia 12] and in [13 13 Moscow 13] . He became a violinist , and taught music at a conservatory in [3 Yerevan 3] . In the early 1970s he learned about the art of microminiature from one of his students , Eduard Ghazaryan . Sandaldjian emigrated to the [4 United States 4] in 1980 , but was required to leave behind his first collection of 18 miniature works . Over the next decade , he produced another 33 miniatures . Working extremely slowly under a microscope , Sandaldjian employed self-made tools such as sharpened needles tipped with ruby or diamond dust , compiling his sculptures out of minuscule materials such as dust , lint , and hair . Even among the few practitioners of microminiature art , Sandaldjian was unusual in that he painted his work , using a single sharpened strand of hair as a brush . Sandaldjian would time his motions to come between his heartbeats , thus maximizing his control of his fingers . Those who saw Sandaldjian at work said that they could not tell when his hands moved . As recounted in [14 14 Lawrence Weschler 14] 's book 5 Mr. Wilson 's Cabinet of Wonder , [5 Museum of Jurassic Technology 5] founder 6 David Hildebrand Wilson learned about Sandaldjian from a museum visitor , and began visiting him at his home in [6 7 Montebello 6] , [7 California 7]

1: 4277629 Armenian_American

2: 20448 Museum_of_Jurassic_Technology

3: 313870 Mount_Ararat

4: 1095706 Jesus

5: 20859 Mickey_Mouse

6: 69880 Napoleon

7: 23805 Pope_John_Paul_II

8: 3080 Alexandria

9: 8087628 Egypt

10: 26779 Soviet_Union

11: 34352 Yerevan

12: 10918072 Armenia
 13: 19004 Moscow
 14: 2387576 Lawrence_Weschler
 1: 26714 Sculpture
 [winner: 26714 Sculpture, pem_pos=1,
 final_scores=-1.177151, final_local_scores=-1.2996104,
 pem=-0.018946731, similarity=-34.59022|
 gt: 26714 Sculpture, pem_pos=1, final_scores=-1.177151,
 final_local_scores=-1.2996104, pem=-0.018946731,
 similarity=-34.59022]
 2: 7324 Crucifix
 [winner: 16822415 Altar_crucifix, pem_pos=0,
 final_scores=-0.9108185, final_local_scores=-0.86313033,
 pem=0.0, similarity=-21.667818|
 gt: 7324 Crucifix, pem_pos=1, final_scores=-1.1667912,
 final_local_scores=-1.2295258, pem=-0.011329513,
 similarity=-32.54018]
 3: 7303 Cross
 [winner: 7303 Cross, pem_pos=0, final_scores=-1.2567399,
 final_local_scores=-1.3157468, pem=-0.590119,
 similarity=-31.965635|
 gt: 7303 Cross, pem_pos=0, final_scores=-1.2567399,
 final_local_scores=-1.3157468, pem=-0.590119,
 similarity=-31.965635]
 4: 76361 Snow_White_and_the_Seven_Dwarfs_(1937_film)
 [winner: 160108 Snow_White, pem_pos=0,
 final_scores=-0.080884054, final_local_scores=-0.041162744,
 pem=0.0, similarity=2.8613346|
 gt: 76361 Snow_White_and_the_Seven_Dwarfs_(1937_film),
 pem_pos=1, final_scores=-0.12816446,
 final_local_scores=-0.10615015, pem=-0.45798308,
 similarity=3.4125638]
 [winner: 160108 Snow_White, pem_pos=0,
 final_scores=-0.18003297, final_local_scores=-0.14946787,
 pem=-0.3951778, similarity=1.778332|
 gt: 76361 Snow_White_and_the_Seven_Dwarfs_(1937_film),
 pem_pos=1, final_scores=-0.33111554,
 final_local_scores=-0.3278434, pem=-1.455418,
 similarity=2.2210011]
 [winner: 27870556 Alberto_Guzik, pem_pos=1,
 final_scores=-1.2761738, final_local_scores=-1.2961841,
 pem=-0.6931472, similarity=-30.821566|
 RECALL MISS: 76361
 Snow_White_and_the_Seven_Dwarfs_(1937_film)]
 5: 24183459 Mr._Wilson's_Cabinet_of_Wonder
 [winner: 63568 Dennis_the_Menace_(U.S._comics),
 pem_pos=0, final_scores=-0.192213,
 final_local_scores=-0.1302387, pem=-0.6128472,
 similarity=3.535889|
 RECALL MISS: 24183459 Mr._Wilson's_Cabinet_of_Wonder]
 [winner: 16232423 Wonder_(emotion), pem_pos=1,
 final_scores=-0.75788975, final_local_scores=-0.7012352,

pem=-1.3543878, similarity=-9.471171|
 RECALL MISS: 24183459 Mr._Wilson's_Cabinet_of_Wonder]
 [winner: 7323222 Reinfeldt_Cabinet, pem_pos=0,
 final_scores=-0.9112261, final_local_scores=-0.88503176,
 pem=0.0, similarity=-22.3214|
 RECALL MISS: 24183459 Mr._Wilson's_Cabinet_of_Wonder]
 [winner: 1415027 David_Berman_(musician), pem_pos=5,
 final_scores=-2.4644046, final_local_scores=-2.6213195,
 pem=-3.0012724, similarity=-57.814304|
 RECALL MISS: 24183459 Mr._Wilson's_Cabinet_of_Wonder]
 6: 16243664 David_Hildebrand_Wilson
 [winner: 16243664 David_Hildebrand_Wilson, pem_pos=0,
 final_scores=-0.04787995, final_local_scores=0.062940955,
 pem=0.0, similarity=5.9679966|
 gt: 16243664 David_Hildebrand_Wilson, pem_pos=0,
 final_scores=-0.04787995, final_local_scores=0.062940955,
 pem=0.0, similarity=5.9679966]
 7: 107661 Montebello,_California
 [winner: 5407 California, pem_pos=0,
 final_scores=0.5032758, final_local_scores=0.5238857,
 pem=-0.08662026, similarity=20.194563|
 RECALL MISS: 107661 Montebello,_California]
 [winner: 107661 Montebello,_California, pem_pos=0,
 final_scores=0.15764157, final_local_scores=0.1943891,
 pem=-0.7027307, similarity=13.71223|
 gt: 107661 Montebello,_California, pem_pos=0,
 final_scores=0.15764157, final_local_scores=0.1943891,
 pem=-0.7027307, similarity=13.71223]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7218527, final_local_scores=-1.7028446,
 pem=-4.6728287, similarity=-21.314997|
 RECALL MISS: 107661 Montebello,_California]
 1: 4277629 Armenian_American
 [pem_pos=0, final_scores=0.16269664,
 final_local_scores=0.1581797, pem=0.0, similarity=8.810111]
 2: 20448 Museum_of_Jurassic_Technology
 [pem_pos=0, final_scores=0.17258106,
 final_local_scores=0.23635146, pem=0.0, similarity=11.142912]
 3: 313870 Mount_Ararat
 [pem_pos=0, final_scores=0.09231891,
 final_local_scores=0.12821528, pem=0.0, similarity=7.9159126]
 4: 1095706 Jesus
 [pem_pos=0, final_scores=0.44886273,
 final_local_scores=0.5061037, pem=-0.0733903,
 similarity=19.591965]
 5: 20859 Mickey_Mouse
 [pem_pos=0, final_scores=0.09678762,
 final_local_scores=0.14503014, pem=0.0, similarity=8.417702]
 6: 69880 Napoleon
 [pem_pos=0, final_scores=0.39068115,
 final_local_scores=0.4210962, pem=-0.08754977,
 similarity=17.132172]

7: 23805 Pope_John_Paul_II
[pem_pos=0, final_scores=0.05764763,
final_local_scores=0.056781128, pem=0.0, similarity=5.784175]
8: 3080 Alexandria
[pem_pos=0, final_scores=0.286614,
final_local_scores=0.29317477, pem=-0.43061808,
similarity=15.1804]
9: 8087628 Egypt
[pem_pos=0, final_scores=0.5826183,
final_local_scores=0.5995953, pem=-0.27480945,
similarity=23.47729]
10: 26779 Soviet_Union
[pem_pos=0, final_scores=0.36654097,
final_local_scores=0.37348723, pem=-0.075323656,
similarity=15.644939]
11: 34352 Yerevan
[pem_pos=0, final_scores=0.3985923,
final_local_scores=0.3901164, pem=0.0, similarity=15.731564]
12: 10918072 Armenia
[pem_pos=0, final_scores=0.4506331,
final_local_scores=0.45010304, pem=-0.4307794,
similarity=19.86433]
13: 19004 Moscow
[pem_pos=0, final_scores=0.6193876,
final_local_scores=0.6320003, pem=-0.106582455,
similarity=23.529476]
14: 2387576 Lawrence_Weschler
[pem_pos=0, final_scores=0.12996994,
final_local_scores=0.235551, pem=0.0, similarity=11.119024]
1: 16766459 Hagop_Sandaldjian
[pem_pos=0, final_scores=0.043772206,
final_local_scores=0.12371892, pem=0.0, similarity=7.781732]
2: 37398 The_Walt_Disney_Company
[pem_pos=0, final_scores=0.23163556,
final_local_scores=0.2779144, pem=-0.24767768,
similarity=13.730143]
3: 34352 Yerevan
[pem_pos=0, final_scores=0.49016744,
final_local_scores=0.4901483, pem=0.0, similarity=18.716717]
4: 3434750 United_States
[pem_pos=0, final_scores=0.4690377,
final_local_scores=0.46369654, pem=-0.073457584,
similarity=18.326817]
5: 20448 Museum_of_Jurassic_Technology
[pem_pos=0, final_scores=0.20901357,
final_local_scores=0.27614844, pem=0.0, similarity=12.330533]
6: 107661 Montebello,_California
[pem_pos=0, final_scores=0.15764157,
final_local_scores=0.1943891, pem=-0.7027307,
similarity=13.71223]
7: 5407 California
[pem_pos=0, final_scores=0.5032758,

final_local_scores=0.5238857, pem=-0.08662026,
similarity=20.194563]

Sample: b'Evelyn_Fanshawe'

Sir Evelyn Dalrymple Fanshawe 1 CB , 2 CBE -LRB- 1895 --
1979 -RRB- was a [1 British 1] and the Director of the in the
from 1948-1952 . A grandson on his mother 's side of [1 1 Sir
Evelyn Wood 1] , he was born to Lt. Gen. Sir [2 2 Hew Dalrymp
le Fanshawe 2] and Anna Paulina Mary Wood in 1895 . He
married Marie Harari in 1920 . He was commissioned into the
in 1914 and saw service in [3 3 France 3] , [4 4 Palestine 4]
 , [5 5 Mesopotamia 5] , , [6 6 Russia 6] and [7 7 Syria 7]
during [8 8 World War I 8] ; among his assignments during
this period was 3 Aide-de-camp to his father who was General
Officer Commanding the British Cavalry Corps -LRB- 1915 -
RRB- . From 1915-1919 he was seconded to the [9 9 Royal Flying
Corps 9] . He returned to his regiment as 4 Adjutant in
1919 . In 1939 he was appointed Commander of the 5 20th Armou
red Brigade and following promotion to Major General he held
the post of the Armoured Training and Commander of the 6 Royal
Armoured Corps Training Establishment from 1942 to the end of
the War . [2 Fanshawe 2] retired from the [3 Army 3] in 1945
whereupon he became the 7 UN Relief and Rehabilitation
Administration Director in the -LRB- 1945 -- 1948 -RRB- .
Subsequently he was Director of the in the [4 British Zone 4]
of [5 Germany 5] from 1948-1952 . In 1952 he was attached to
the Dominion Countries UN Organisation Mission , and in 1960
was 8 High Sheriff of Northamptonshire

1: 1189274 Evelyn_Wood_(British_Army_officer)

2: 29779045 Hew_Dalrymple_Fanshawe

3: 5843419 France

4: 1114732 Palestine

5: 20189 Mesopotamia

6: 25391 Russia

7: 7515849 Syria

8: 4764461 World_War_I

9: 233670 Royal_Flying_Corps

1: 311455 Order_of_the_Bath

[winner: 2886317 CB_postcode_area, pem_pos=4,
final_scores=-1.2250526, final_local_scores=-1.127944,
pem=-3.3588254, similarity=-11.304568]

gt: 311455 Order_of_the_Bath, pem_pos=0,
final_scores=-1.3292316, final_local_scores=-1.4433849,
pem=-0.41609186, similarity=-36.721]

2: 212182 Order_of_the_British_Empire

[winner: 212182 Order_of_the_British_Empire, pem_pos=0,
final_scores=-0.2282359, final_local_scores=-0.24309935,
pem=0.0, similarity=-3.1648564]

gt: 212182 Order_of_the_British_Empire, pem_pos=0,
final_scores=-0.2282359, final_local_scores=-0.24309935,
pem=0.0, similarity=-3.1648564]

3: 239481 Aide-de-camp

[winner: 239481 Aide-de-camp, pem_pos=0, final_scores=-0.09398036, final_local_scores=-0.13338605, pem=0.0, similarity=0.10920724|
gt: 239481 Aide-de-camp, pem_pos=0, final_scores=-0.09398036, final_local_scores=-0.13338605, pem=0.0, similarity=0.10920724]

4: 534754 Adjutant
[winner: 751581 Adjutant_general, pem_pos=0, final_scores=-0.121235356, final_local_scores=-0.16263337, pem=0.0, similarity=-0.76359123|
gt: 534754 Adjutant, pem_pos=1, final_scores=-0.19764726, final_local_scores=-0.2541818, pem=-0.08121905, similarity=-3.0538974]

5: 5451893 20th_Armoured_Brigade_(United_Kingdom)
[winner: 1824150 Armoured_Brigade_(Finnish_Army), pem_pos=0, final_scores=-0.007090792, final_local_scores=0.057719216, pem=0.0, similarity=5.8121696|
RECALL MISS: 5451893
20th_Armoured_Brigade_(United_Kingdom)]

[winner: 2336235 Hilbert's_twentieth_problem, pem_pos=17, final_scores=-1.4366103, final_local_scores=-1.3624146, pem=-3.68571, similarity=-16.52399|
RECALL MISS: 5451893
20th_Armoured_Brigade_(United_Kingdom)]

6: 83399 Royal_Armoured_Corps
[winner: 16716500 Royal,_Beaufort_County,_North_Carolina, pem_pos=1, final_scores=-0.8322434, final_local_scores=-0.70352846, pem=-2.3211868, similarity=-4.2820053|
RECALL MISS: 83399 Royal_Armoured_Corps]

[winner: 24947367 Armoured_Multirole_Carrier, pem_pos=3, final_scores=-1.1735798, final_local_scores=-1.1278709, pem=-1.0986123, similarity=-23.59378|
gt: 83399 Royal_Armoured_Corps, pem_pos=10, final_scores=-2.7939913, final_local_scores=-3.084194, pem=-6.0935698, similarity=-54.811]

[winner: 13761964 Corps,_Isère, pem_pos=2, final_scores=-1.7106185, final_local_scores=-1.7591677, pem=-1.7153218, similarity=-39.07917|
RECALL MISS: 83399 Royal_Armoured_Corps]

7: 32268
United_Nations_Relief_and_Rehabilitation_Administration
[winner: 5235745 Steffan_Cravos, pem_pos=1, final_scores=-1.0076798, final_local_scores=-0.9355607, pem=-0.6931472, similarity=-20.059843|
RECALL MISS: 32268
United_Nations_Relief_and_Rehabilitation_Administration]

[winner: 14290232
United_States_lightship_Relief_(WAL-605), pem_pos=11, final_scores=-1.3409634, final_local_scores=-1.2659361, pem=-2.3978953, similarity=-20.648216|

RECALL MISS: 32268
United_Nations_Relief_and_Rehabilitation_Administration]
[winner: 14220245 Occupational_rehabilitation, pem_pos=7,
final_scores=-1.6352245, final_local_scores=-1.6076431,
pem=-3.7044976, similarity=-23.739933]
RECALL MISS: 32268
United_Nations_Relief_and_Rehabilitation_Administration]
8: 3520208 High_Sheriff_of_Northamptonshire
[winner: 66584 Northamptonshire, pem_pos=0,
final_scores=0.004866287, final_local_scores=0.0055289716,
pem=-0.30746016, similarity=5.9267244]
gt: 3520208 High_Sheriff_of_Northamptonshire, pem_pos=6,
final_scores=-0.5242211, final_local_scores=-0.5150833,
pem=-2.4072304, similarity=1.8094932]
[winner: 1539792 High_sheriff, pem_pos=0,
final_scores=-0.5294286, final_local_scores=-0.5579718,
pem=-0.080371104, similarity=-12.124207]
RECALL MISS: 3520208 High_Sheriff_of_Northamptonshire]
1: 1189274 Evelyn_Wood_(British_Army_officer)
[pem_pos=0, final_scores=0.044907495,
final_local_scores=0.04598117, pem=0.0, similarity=5.4618826]
2: 29779045 Hew_Dalrymple_Fanshawe
[pem_pos=0, final_scores=0.13332997,
final_local_scores=0.15326536, pem=0.0, similarity=8.663457]
3: 5843419 France
[pem_pos=0, final_scores=0.60891575,
final_local_scores=0.6244603, pem=-0.16240834,
similarity=23.608055]
4: 1114732 Palestine
[pem_pos=0, final_scores=0.22489838,
final_local_scores=0.18853575, pem=-0.53610015,
similarity=12.631392]
5: 20189 Mesopotamia
[pem_pos=0, final_scores=0.27886426,
final_local_scores=0.3055045, pem=-0.030018875,
similarity=13.369824]
6: 25391 Russia
[pem_pos=0, final_scores=0.6347672,
final_local_scores=0.7048174, pem=-0.20299491,
similarity=26.226788]
7: 7515849 Syria
[pem_pos=0, final_scores=0.5434109,
final_local_scores=0.5746694, pem=-0.26162508,
similarity=22.661753]
8: 4764461 World_War_I
[pem_pos=0, final_scores=0.31761426,
final_local_scores=0.26353434, pem=0.0, similarity=11.954103]
9: 233670 Royal_Flying_Corps
[pem_pos=0, final_scores=0.10118948,
final_local_scores=0.0819127, pem=0.0, similarity=6.534151]
1: 31717 United_Kingdom
[pem_pos=0, final_scores=0.45786452,

```

final_local_scores=0.450709, pem=-0.48936236,
similarity=20.200996]
2: 29779045 Hew_Dalrymple_Fanshawe
   [pem_pos=0, final_scores=0.09034519,
final_local_scores=0.10631102, pem=0.0, similarity=7.262246]
3: 32087 United_States_Army
   [pem_pos=0, final_scores=0.2151329,
final_local_scores=0.18680435, pem=-1.138116,
similarity=15.853579]
4: 2949977 Allied-occupied_Germany
   [pem_pos=0, final_scores=0.22650723,
final_local_scores=0.23745131, pem=0.0, similarity=11.175734]
5: 11867 Germany
   [pem_pos=0, final_scores=0.60577077,
final_local_scores=0.6344235, pem=-0.15529114,
similarity=23.866673]

```

Sample: b'Appollo'

-LRB- dog -RRB- Appollo was a 1 search and rescue dog who served with the K-9 unit of the . He was awarded the [1 1 Dickin Medal 1] , the animals ' equivalent of the [2 2 Victoria Cross 2] , in recognition of the work done by all search and rescue dogs following the [1 2 September 11] 11 attacks . Appollo and his handler were working at the [2 3 World Trade Center 2] site soon after the attacks . Appollo was a [3 3 German Shepherd 3] born around 1992 , who was in service with the K-9 unit of the -LRB- NYPD -RRB- . In 1994 , he graduated from the NYPD Canine Special Operations Division , and was one of the first dogs to learn 4 search and rescue . Appollo passed Type-II training in 5 Florida in 1997 , and Type-I in [4 4 Indianapolis 4] in 1999 . He was also part of the first NYPD K-9 team to pass them for 6 Urban Search and Rescue New York Task Force 1 . Appollo and his handler Peter Davis also worked in the [5 5 Dominican Republic 5] after a hurricane . Appollo died in November 2006 . Appollo and his handler , Peter Davis , were called in to assist with the rescue operations after the [3 7 September 3] 11 terror attacks . They arrived at the [4 8 World Trade Center 4] site fifteen minutes after the attack , making Appollo the first search and rescue dog to arrive at the site after the collapse of the [6 6 World Trade Center 6] . At one point , Appollo was almost killed by flames and falling debris . However , he survived , having been drenched after falling into a pool of water just before this incident . [5 Appollo 5] started working again as soon as Davis had brushed the debris off him . Appollo received the [7 7 Dickin Medal 7] , the animal equivalent of the [8 8 Victoria Cross 8] , on behalf of all the search and rescue dogs who participated in the rescue operations at the [6 World Trade Center 6] site and the [7 Pentagon 7] . He received the award along with guide dogs 9 Roselle and Salty , who rescued their owners from the [8 World Trade Center 8] . The citation for the award was as follows :

Appollo received the 10 American Kennel Club Ace award in 2001 . He was also honoured for his work at the 11 Westminster Kennel Club Dog Show

- 1: 766519 Dickin_Medal
- 2: 17491404 Victoria_Cross
- 3: 79289 German_Shepherd
- 4: 57707 Indianapolis
- 5: 8060 Dominican_Republic
- 6: 11395198 World_Trade_Center
- 7: 766519 Dickin_Medal
- 8: 17491404 Victoria_Cross

1: 3342400 Search_and_rescue_dog
[winner: 16685370 Rescue_California, pem_pos=1, final_scores=-0.9156771, final_local_scores=-0.82949436, pem=-1.6762859, similarity=-11.54815]
gt: 3342400 Search_and_rescue_dog, pem_pos=19, final_scores=-1.6752002, final_local_scores=-1.8010478, pem=-3.4816246, similarity=-30.723528]
[winner: 14648008 Search_of_persons, pem_pos=3, final_scores=-1.0345672, final_local_scores=-0.9674805, pem=-1.3862944, similarity=-17.242949]
RECALL MISS: 3342400 Search_and_rescue_dog]
[winner: 4269567 Dog, pem_pos=0, final_scores=-1.3797858, final_local_scores=-1.44171, pem=-0.090513, similarity=-38.44157]

RECALL MISS: 3342400 Search_and_rescue_dog]
2: 5058690 September_11_attacks
[winner: 5058690 September_11_attacks, pem_pos=0, final_scores=0.17932837, final_local_scores=0.1355556, pem=0.0, similarity=8.134962]
gt: 5058690 September_11_attacks, pem_pos=0, final_scores=0.17932837, final_local_scores=0.1355556, pem=0.0, similarity=8.134962]
[winner: 19452 May_11, pem_pos=2, final_scores=-0.7886639, final_local_scores=-0.5311905, pem=-3.271949, similarity=6.0312963]
RECALL MISS: 5058690 September_11_attacks]
[winner: 9594284 Timeline_of_al-Qaeda_attacks, pem_pos=2, final_scores=-0.9450589, final_local_scores=-0.8876391, pem=-1.0432429, similarity=-16.725891]
gt: 5058690 September_11_attacks, pem_pos=0, final_scores=-1.3551024, final_local_scores=-1.547261, pem=0.0, similarity=-42.083645]

3: 67173 World_Trade_Center_site
[winner: 11395198 World_Trade_Center, pem_pos=0, final_scores=0.34814817, final_local_scores=0.3135547, pem=-0.09772106, similarity=13.978233]
gt: 67173 World_Trade_Center_site, pem_pos=7, final_scores=-0.23322685, final_local_scores=-0.31014404, pem=-3.3089428, similarity=12.828942]
[winner: 2348804 New_Kalabsha, pem_pos=7, final_scores=-1.4176266, final_local_scores=-1.407551,

pem=-2.1972246, similarity=-25.965567|
RECALL MISS: 67173 World_Trade_Center_site]

4: 172599 Search_and_rescue
[winner: 16685370 Rescue_California, pem_pos=1,
final_scores=-0.9194836, final_local_scores=-0.83365244,
pem=-1.6762859, similarity=-11.6722355|
RECALL MISS: 172599 Search_and_rescue]
[winner: 14648008 Search_of_persons, pem_pos=3,
final_scores=-1.061631, final_local_scores=-0.99704355,
pem=-1.3862944, similarity=-18.12517|
RECALL MISS: 172599 Search_and_rescue]

5: 18933066 Florida
[winner: 18933066 Florida, pem_pos=0,
final_scores=-0.102514446, final_local_scores=-0.089610636,
pem=-0.15405534, similarity=2.2533298|
gt: 18933066 Florida, pem_pos=0,
final_scores=-0.102514446, final_local_scores=-0.089610636,
pem=-0.15405534, similarity=2.2533298]

6: 6773163 Urban_Search_and_Rescue_New_York_Task_Force_1
[winner: 6773163
Urban_Search_and_Rescue_New_York_Task_Force_1, pem_pos=0,
final_scores=-0.18839212, final_local_scores=-0.09846727,
pem=0.0, similarity=1.1512532|
gt: 6773163
Urban_Search_and_Rescue_New_York_Task_Force_1, pem_pos=0,
final_scores=-0.18839212, final_local_scores=-0.09846727,
pem=0.0, similarity=1.1512532]

7: 5058690 September_11_attacks
[winner: 5058690 September_11_attacks, pem_pos=0,
final_scores=0.216408, final_local_scores=0.17605942,
pem=0.0, similarity=9.3436775|
gt: 5058690 September_11_attacks, pem_pos=0,
final_scores=0.216408, final_local_scores=0.17605942,
pem=0.0, similarity=9.3436775]
[winner: 19452 May_11, pem_pos=2,
final_scores=-0.7797933, final_local_scores=-0.5215007,
pem=-3.271949, similarity=6.32046|
RECALL MISS: 5058690 September_11_attacks]
[winner: 9594284 Timeline_of_al-Qaeda_attacks, pem_pos=2,
final_scores=-0.83816105, final_local_scores=-0.7708695,
pem=-1.0432429, similarity=-13.241253|
gt: 5058690 September_11_attacks, pem_pos=0,
final_scores=-1.1999649, final_local_scores=-1.3777969,
pem=0.0, similarity=-37.026497]
[winner: 31333334 Terror, pem_pos=3,
final_scores=-1.2224078, final_local_scores=-1.2299025,
pem=-1.3416064, similarity=-25.317165|
RECALL MISS: 5058690 September_11_attacks]

8: 67173 World_Trade_Center_site
[winner: 11395198 World_Trade_Center, pem_pos=0,
final_scores=0.34801638, final_local_scores=0.3134107,
pem=-0.09772106, similarity=13.973936|

gt: 67173 World_Trade_Center_site, pem_pos=7,
final_scores=-0.23376097, final_local_scores=-0.31072748,
pem=-3.3089428, similarity=12.811531]
[winner: 2348804 New_Kalabsha, pem_pos=7,
final_scores=-1.5292748, final_local_scores=-1.5295098,
pem=-2.1972246, similarity=-29.605064]
RECALL MISS: 67173 World_Trade_Center_site]
9: 35411002 Salty_and_Roselle
[winner: 125788 Roselle, New_Jersey, pem_pos=1,
final_scores=-0.23259746, final_local_scores=-0.16731009,
pem=-0.96099615, similarity=4.3228927]
RECALL MISS: 35411002 Salty_and_Roselle]
[winner: 3245628 Railway_engines_(Thomas_&_Friends),
pem_pos=0, final_scores=-0.40032917,
final_local_scores=-0.35427344, pem=-0.87035364,
similarity=-1.7493944]
RECALL MISS: 35411002 Salty_and_Roselle]
10: 84627 American_Kennel_Club
[winner: 84627 American_Kennel_Club, pem_pos=0,
final_scores=-1.0265881, final_local_scores=-1.0634985,
pem=0.0, similarity=-27.647205]
gt: 84627 American_Kennel_Club, pem_pos=0,
final_scores=-1.0265881, final_local_scores=-1.0634985,
pem=0.0, similarity=-27.647205]
11: 921387 Westminster_Kennel_Club_Dog_Show
[winner: 921387 Westminster_Kennel_Club_Dog_Show,
pem_pos=0, final_scores=-0.03863068,
final_local_scores=0.02648069, pem=0.0, similarity=4.8799496]
gt: 921387 Westminster_Kennel_Club_Dog_Show, pem_pos=0,
final_scores=-0.03863068, final_local_scores=0.02648069,
pem=0.0, similarity=4.8799496]
1: 766519 Dickin_Medal
[pem_pos=0, final_scores=0.03606157,
final_local_scores=0.08999023, pem=0.0, similarity=6.775201]
2: 17491404 Victoria_Cross
[pem_pos=0, final_scores=0.056437716,
final_local_scores=0.08473633, pem=0.0, similarity=6.618414]
3: 79289 German_Shepherd
[pem_pos=0, final_scores=0.16653161,
final_local_scores=0.19930723, pem=0.0, similarity=10.037438]
4: 57707 Indianapolis
[pem_pos=0, final_scores=0.32971668,
final_local_scores=0.37873155, pem=-0.30184713,
similarity=17.03331]
5: 8060 Dominican_Republic
[pem_pos=0, final_scores=0.19042741,
final_local_scores=0.28065318, pem=-0.24480942,
similarity=13.796275]
6: 11395198 World_Trade_Center
[pem_pos=0, final_scores=0.3151502,
final_local_scores=0.27552742, pem=-0.09772106,
similarity=12.843423]

7: 766519 Dickin_Medal
 [pem_pos=0, final_scores=0.05296071,
 final_local_scores=0.108449966, pem=0.0, similarity=7.326076]
 8: 17491404 Victoria_Cross
 [pem_pos=0, final_scores=0.097388074,
 final_local_scores=0.12946838, pem=0.0, similarity=7.9533076]
 1: 5058690 September_11_attacks
 [pem_pos=0, final_scores=0.17932837,
 final_local_scores=0.1355556, pem=0.0, similarity=8.134962]
 2: 11395198 World_Trade_Center
 [pem_pos=0, final_scores=0.34814817,
 final_local_scores=0.3135547, pem=-0.09772106,
 similarity=13.978233]
 3: 5058690 September_11_attacks
 [pem_pos=0, final_scores=0.216408,
 final_local_scores=0.17605942, pem=0.0, similarity=9.3436775]
 4: 11395198 World_Trade_Center
 [pem_pos=0, final_scores=0.34801638,
 final_local_scores=0.3134107, pem=-0.09772106,
 similarity=13.973936]
 5: 21130589 Appollo_(dog)
 [pem_pos=0, final_scores=0.037499025,
 final_local_scores=0.042445242, pem=0.0, similarity=5.356364]
 6: 11395198 World_Trade_Center
 [pem_pos=0, final_scores=0.32173866,
 final_local_scores=0.2847063, pem=-0.09772106,
 similarity=13.117339]
 7: 20740978 The_Pentagon
 [pem_pos=0, final_scores=0.32183248,
 final_local_scores=0.31479934, pem=-0.14635207,
 similarity=14.279839]
 8: 11395198 World_Trade_Center
 [pem_pos=0, final_scores=0.2990216,
 final_local_scores=0.25989133, pem=-0.09772106,
 similarity=12.376811]

Sample: b'Statue_of_Europe'

The Statue of Europe -LRB- also referred to as Unity in Peace -RRB- is a sculpture symbolising peace through [1 1 European 1] integration , while at the same time aiming to demonstrate the motto of the [1 1 European Union 1] -LRB- EU -RRB- , 2 United in Diversity . It is located in the garden of [2 2 Convent Van Maerlant 2] -LRB- the library of the [3 3 European Commission 3] -RRB- in the crossroads of the street Van Maerlant and `` chaussée d'Etterbeek '' , in the [4 4 European Quarter 4] of [5 5 Brussels 5] , [6 6 Belgium 6] . Made out of resin , the statue measures more than 5 meters and weighs nearly 800 kg . It was inaugurated on 9 December 2003 by [7 7 Neil Kinnock 7] and [8 8 Viviane Reding 8] , who respectively were 3 Vice-President of the [2 European Commission 2] and 4 Commissioner for Culture at the time . 2003 was the 5 European Year of People with Disabilities , and the 6 Ea

stern enlargement of the EU was awaited . The statue was manufactured , modelled , polished and painted by 7 visually impaired children under the supervision of by [9 9 French 9] artist [10 10 Bernard Romain 10] , who wanted to demonstrate that 8 disability was not a disqualifying factor . Tightly entwined , the differently coloured arms symbolise the culturally diverse peoples of [3 Europe 3] , collectively lifting a sphere redolent of the 9 EU flag , surmounted by the

1: 9317 European_Union

2: 18994554 Convent_Van_Maerlant

3: 9974 European_Commission

4: 9981557 Brussels_and_the_European_Union

5: 3708 Brussels

6: 3343 Belgium

7: 21978 Neil_Kinnock

8: 969052 Viviane_Reding

9: 5843419 France

10: 35582420 Bernard_Romain

1: 683704 European_integration

[winner: 9239 Europe, pem_pos=0, final_scores=0.21572666, final_local_scores=0.20671666, pem=-0.25557294, similarity=11.648396]

RECALL MISS: 683704 European_integration]

[winner: 24328894 Application_Integration_Architecture, pem_pos=12, final_scores=-1.0477914, final_local_scores=-0.93991774, pem=-3.3346288, similarity=-5.825078]

gt: 683704 European_integration, pem_pos=10, final_scores=-2.1483057, final_local_scores=-2.3931766, pem=-3.0941603, similarity=-50.500923]

2: 1759956 Motto_of_the_European_Union

[winner: 1759956 Motto_of_the_European_Union, pem_pos=0, final_scores=-0.15038365, final_local_scores=-0.14684235, pem=0.0, similarity=-0.2923556]

gt: 1759956 Motto_of_the_European_Union, pem_pos=0, final_scores=-0.15038365, final_local_scores=-0.14684235, pem=0.0, similarity=-0.2923556]

3: 10066430 Vice-President_of_the_European_Commission

[winner: 9974 European_Commission, pem_pos=0, final_scores=0.47607273, final_local_scores=0.4603331, pem=0.0, similarity=17.826971]

RECALL MISS: 10066430 Vice-President_of_the_European_Commission]

[winner: 52113 Vice_president, pem_pos=0, final_scores=-0.49091882, final_local_scores=-0.493418, pem=-1.4391059, similarity=-2.8087838]

gt: 10066430 Vice-President_of_the_European_Commission, pem_pos=8, final_scores=-0.69782233, final_local_scores=-0.7451977, pem=-2.4965036, similarity=-4.5720987]

[winner: 1961717 Supreme_Court_of_South_Australia, pem_pos=0, final_scores=-1.2252805,

final_local_scores=-1.2570379, pem=0.0,
 similarity=-33.422806|
 RECALL MISS: 10066430 Vice-
 President_of_the_European_Commission]
 4: 9759828
 European_Commissioner_for_Education,_Culture,_Multilingualism
 _and_Youth
 [winner: 9759828
 European_Commissioner_for_Education,_Culture,_Multilingualism
 _and_Youth, pem_pos=0, final_scores=-0.3305317,
 final_local_scores=-0.33358926, pem=0.0,
 similarity=-5.8652563|
 gt: 9759828
 European_Commissioner_for_Education,_Culture,_Multilingualism
 _and_Youth, pem_pos=0, final_scores=-0.3305317,
 final_local_scores=-0.33358926, pem=0.0,
 similarity=-5.8652563]
 5: 985031 European_Year_of_People_with_Disabilities
 [winner: 985031
 European_Year_of_People_with_Disabilities, pem_pos=0,
 final_scores=-0.05044906, final_local_scores=0.061699554,
 pem=0.0, similarity=5.9309506|
 gt: 985031 European_Year_of_People_with_Disabilities,
 pem_pos=0, final_scores=-0.05044906,
 final_local_scores=0.061699554, pem=0.0,
 similarity=5.9309506]
 6: 21133193 2004_enlargement_of_the_European_Union
 [winner: 37403 Eastern_Europe, pem_pos=1,
 final_scores=-0.23687576, final_local_scores=-0.2998006,
 pem=-2.7159638, similarity=9.912898|
 RECALL MISS: 21133193
 2004_enlargement_of_the_European_Union]
 [winner: 9317 European_Union, pem_pos=0,
 final_scores=-0.24361874, final_local_scores=-0.34054774,
 pem=-0.39703724, similarity=-3.913761|
 RECALL MISS: 21133193
 2004_enlargement_of_the_European_Union]
 [winner: 1961717 Supreme_Court_of_South_Australia,
 pem_pos=0, final_scores=-1.311105,
 final_local_scores=-1.3507881, pem=0.0, similarity=-36.2205|
 RECALL MISS: 21133193
 2004_enlargement_of_the_European_Union]
 [winner: 6063358 Enlargement_of_the_United_Nations,
 pem_pos=1, final_scores=-1.5154544,
 final_local_scores=-1.5477808, pem=-1.4485426,
 similarity=-34.221752|
 RECALL MISS: 21133193
 2004_enlargement_of_the_European_Union]
 7: 2158298 Visual_impairment
 [winner: 682337 Low_vision, pem_pos=0,
 final_scores=-0.73768735, final_local_scores=-0.6942372,
 pem=0.0, similarity=-16.627708|

gt: 2158298 Visual_impairment, pem_pos=1,
 final_scores=-0.8954279, final_local_scores=-0.88665926,
 pem=-0.15003121, similarity=-21.554075]
 8: 56285 Disability
 [winner: 149699 Paralympic_association_football,
 pem_pos=5, final_scores=-1.093008,
 final_local_scores=-1.0501803, pem=-1.609438,
 similarity=-18.497389]
 gt: 56285 Disability, pem_pos=0,
 final_scores=-1.1769376, final_local_scores=-1.2538161,
 pem=-0.09660623, similarity=-32.801304]
 9: 10025 Flag_of_Europe
 [winner: 9317 European_Union, pem_pos=0,
 final_scores=-0.25575978, final_local_scores=-0.35381,
 pem=-0.39703724, similarity=-4.309533]
 RECALL MISS: 10025 Flag_of_Europe]
 [winner: 11424 Flag, pem_pos=0, final_scores=-1.3790913,
 final_local_scores=-1.469351, pem=-0.08931922,
 similarity=-39.27292]
 gt: 10025 Flag_of_Europe, pem_pos=11,
 final_scores=-1.9610322, final_local_scores=-2.1345701,
 pem=-4.810413, similarity=-33.450344]
 1: 9317 European_Union
 [pem_pos=0, final_scores=0.434227,
 final_local_scores=0.41003466, pem=-0.005814796,
 similarity=16.357588]
 2: 18994554 Convent_Van_Maerlant
 [pem_pos=0, final_scores=0.107054755,
 final_local_scores=0.17011774, pem=0.0, similarity=9.166366]
 3: 9974 European_Commission
 [pem_pos=0, final_scores=0.31682163,
 final_local_scores=0.2863755, pem=0.0, similarity=12.635729]
 4: 9981557 Brussels_and_the_European_Union
 [pem_pos=0, final_scores=0.2709461,
 final_local_scores=0.29508865, pem=-0.24994177,
 similarity=14.25497]
 5: 3708 Brussels
 [pem_pos=0, final_scores=0.614371,
 final_local_scores=0.60272163, pem=-0.025440974,
 similarity=22.214481]
 6: 3343 Belgium
 [pem_pos=0, final_scores=0.60258687,
 final_local_scores=0.60425454, pem=-0.24180195,
 similarity=23.43683]
 7: 21978 Neil_Kinnock
 [pem_pos=0, final_scores=0.27766776,
 final_local_scores=0.3498776, pem=0.0, similarity=14.530758]
 8: 969052 Viviane_Reding
 [pem_pos=0, final_scores=0.6795172,
 final_local_scores=0.75528556, pem=0.0, similarity=26.62894]
 9: 5843419 France
 [pem_pos=0, final_scores=0.7207611,

final_local_scores=0.7434854, pem=-0.80223274,
 similarity=30.639465]
 10: 35582420 Bernard_Romain
 [pem_pos=0, final_scores=0.084102675,
 final_local_scores=0.16285428, pem=0.0, similarity=8.94961]
 1: 9239 Europe
 [pem_pos=0, final_scores=0.21572666,
 final_local_scores=0.20671666, pem=-0.25557294,
 similarity=11.648396]
 2: 9974 European_Commission
 [pem_pos=0, final_scores=0.47607273,
 final_local_scores=0.4603331, pem=0.0, similarity=17.826971]
 3: 9239 Europe
 [pem_pos=0, final_scores=0.48703092,
 final_local_scores=0.50307536, pem=-0.086338006,
 similarity=19.572006]

Sample: b'Jose_Evangelista'

José Evangelista -LRB- born 5 August 1943 -RRB- is a [1
 Spanish 1] 1 composer and who is based in [2 Montreal 2] , [3
 Canada 3] . A member of the [1 1 Canadian League of Composers
 1] , the [2 2 Sociedad General de Autores y Editores 2] , and
 an associate of the [3 3 Canadian Music Centre 3] ,
 Evangelista is known for his commitment to 2 contemporary cla
 ssical music and non-Western music . In 1974 he was awarded
 first prize in the Confederación Española de Cajas de Ahorros
 Contest in [4 Madrid 4] for his En guise de fête . In 1978 he
 helped found , a concert society dedicated to 3 avant-garde mu
 sic . In 1982 he was awarded a special prize from [5 Spain 5]
 's Ministry of Culture for his work Vision , and in 1988 he
 won first prize in the 4 choral works competition of St Mary
 Magdalene 's Church in [6 Toronto 6] for his O quam suavis
 est. . Since 1972 he has taught on the music faculty of the [4
 4 Université de Montréal 4] . Among his notable pupils is
 composer [5 5 Analia Llugdar 5] . Born in , Evangelista began
 his professional training at the [6 6 University of Valencia 6
] where he studied 5 computer science for seven years and
 earned a degree in that subject in 1967 . At the same time he
 pursued music courses at the and earned a premier prix in
 there in 1967 . His most influential teacher at that school
 was [7 7 Vicente Asencio 7] who taught his 6 harmony ,
 composition , and 7 orchestration classes . In 1969
 Evangelista moved to [7 Canada 7] and settled in the city of [8
 8 Montreal 8] . From 1970-1973 he was a pupil of [8 8 André Pr
 évost 8] at the [9 9 Université de Montréal 9] where he
 earned a [10 10 Master of Music 10] in composition . In 1974
 he studied contemporary music at the for a semester , later
 returning there for similar studies in 1984 when he was named
 a composer-in-residence at the [11 11 Darmstädter Ferienkurse
 11] . In the Fall of 1974 he entered the graduate music
 program at [12 12 McGill University 12] where he studied
 composition with 8 Bruce Mather and earned a [13 13 Doctor of

Music 13] in 1981 . Evangelista has been active in [9 Canada 9] with the Traditions musicales du monde , a concert society dedicated to promoting non-Western music which he helped found . He developed an interest in the music and culture of [14 14 Southeast Asia 14] and a number of his compositions reflect that interest . In order to pursue studies in this music , he lived in [10 Indonesia 10] during the summers of 1976 and 1980 and in [11 Burma 11] during the summer of 1986 where he studied the and [12 Burmese 12] piano . He was made composer-in-residence at the

1: 22327634 Canadian_League_of_Composers

2: 2899579 Sociedad_General_de_Autores_y_Editores

3: 5895384 Canadian_Music_Centre

4: 7955348 Université_de_Montréal

5: 26861091 Analia_Llúgdar

6: 4449809 University_of_Valencia

7: 26865915 Vicente_Asencio

8: 4264013 André_Prévost_(composer)

9: 7955348 Université_de_Montréal

10: 626525 Master_of_Music

11: 5477430 Darmstädter_Ferienkurse

12: 7954643 McGill_University

13: 2625291 Doctor_of_Music

14: 28741 Southeast_Asia

1: 198778 Composer

[winner: 2066576 Joseph_Holbrooke, pem_pos=3, final_scores=-1.2104918, final_local_scores=-1.243748, pem=-1.3861996, similarity=-25.487839]

gt: 198778 Composer, pem_pos=0, final_scores=-1.2467148, final_local_scores=-1.3853562, pem=-0.023134483, similarity=-37.126266]

2: 330217 Contemporary_classical_music

[winner: 2518639 Contemporary_French_literature, pem_pos=0, final_scores=-0.9808883, final_local_scores=-0.93154305, pem=-1.3518424, similarity=-16.35786]

gt: 330217 Contemporary_classical_music, pem_pos=4, final_scores=-1.3008158, final_local_scores=-1.4405177, pem=-2.1464176, similarity=-27.225657]

[winner: 1175388 American_classical_music, pem_pos=2, final_scores=-1.0262121, final_local_scores=-1.0584269, pem=-0.69302726, similarity=-23.727066]

gt: 330217 Contemporary_classical_music, pem_pos=5, final_scores=-2.0788455, final_local_scores=-2.2903972, pem=-6.1383266, similarity=-30.879133]

3: 276127 Avant-garde_music

[winner: 276127 Avant-garde_music, pem_pos=1, final_scores=-0.88134944, final_local_scores=-0.924593, pem=-0.01666703, similarity=-23.411346]

gt: 276127 Avant-garde_music, pem_pos=1, final_scores=-0.88134944, final_local_scores=-0.924593, pem=-0.01666703, similarity=-23.411346]

4: 43866 Choir
[winner: 276390 Chloral_hydrate, pem_pos=0,
final_scores=-0.9713779, final_local_scores=-0.91624224,
pem=0.0, similarity=-23.252783]
gt: 43866 Choir, pem_pos=1, final_scores=-1.3894668,
final_local_scores=-1.5435748, pem=-0.02112869,
similarity=-41.858734]
[winner: 4438223 Caddy_Works, pem_pos=15,
final_scores=-1.2052572, final_local_scores=-1.1518579,
pem=-3.0445223, similarity=-13.727437]
RECALL MISS: 43866 Choir]

5: 5323 Computer_science
[winner: 8508082 Range_(computer_programming), pem_pos=2,
final_scores=-0.99510866, final_local_scores=-0.8798043,
pem=-0.69258565, similarity=-18.399015]
gt: 5323 Computer_science, pem_pos=0,
final_scores=-1.096282, final_local_scores=-1.170404,
pem=-0.03386557, similarity=-30.653309]

6: 40258 Harmony
[winner: 27464771 Harmony, West_Virginia, pem_pos=6,
final_scores=-1.1576903, final_local_scores=-1.043966,
pem=-3.1100695, similarity=-10.151275]
gt: 40258 Harmony, pem_pos=0, final_scores=-1.7897645,
final_local_scores=-1.9384742, pem=-1.3779649,
similarity=-46.264637]

7: 47398 Orchestration
[winner: 679737 Orchestration_(computing), pem_pos=1,
final_scores=-1.2455615, final_local_scores=-1.1720593,
pem=-2.4458542, similarity=-17.58594]
gt: 47398 Orchestration, pem_pos=0,
final_scores=-1.3814445, final_local_scores=-1.4869838,
pem=-0.1217845, similarity=-39.62257]

8: 8768710 Bruce_Mather
[winner: 8768710 Bruce_Mather, pem_pos=0,
final_scores=0.014992908, final_local_scores=0.06962924,
pem=0.0, similarity=6.167588]
gt: 8768710 Bruce_Mather, pem_pos=0,
final_scores=0.014992908, final_local_scores=0.06962924,
pem=0.0, similarity=6.167588]

1: 22327634 Canadian_League_of_Composers
[pem_pos=0, final_scores=0.12641059,
final_local_scores=0.17714432, pem=0.0, similarity=9.376053]

2: 2899579 Sociedad_General_de_Autores_y_Editores
[pem_pos=0, final_scores=0.066086754,
final_local_scores=0.13107646, pem=0.0, similarity=8.001296]

3: 5895384 Canadian_Music_Centre
[pem_pos=0, final_scores=0.07405974,
final_local_scores=0.09581472, pem=0.0, similarity=6.949015]

4: 7955348 Université_de_Montréal
[pem_pos=0, final_scores=0.48227823,
final_local_scores=0.47557193, pem=-0.0005798829,
similarity=18.284882]

5: 26861091 Analia_Llugdar
[pem_pos=0, final_scores=0.06751834,
final_local_scores=0.1667144, pem=0.0, similarity=9.064803]
6: 4449809 University_of_Valencia
[pem_pos=0, final_scores=0.4008053,
final_local_scores=0.44258887, pem=0.0, similarity=17.297447]
7: 26865915 Vicente_Ascencio
[pem_pos=0, final_scores=0.10015957,
final_local_scores=0.20066863, pem=0.0, similarity=10.078065]
8: 4264013 André_Prévost_(composer)
[pem_pos=0, final_scores=0.28998476,
final_local_scores=0.3713568, pem=-0.35667497,
similarity=17.111395]
9: 7955348 Université_de_Montréal
[pem_pos=0, final_scores=0.5106339,
final_local_scores=0.5065462, pem=-0.0005798829,
similarity=19.209215]
10: 626525 Master_of_Music
[pem_pos=0, final_scores=0.072180614,
final_local_scores=0.097059235, pem=0.0, similarity=6.986154]
11: 5477430 Darmstädter_Ferienkurse
[pem_pos=0, final_scores=0.16838293,
final_local_scores=0.24647179, pem=0.0, similarity=11.444922]
12: 7954643 McGill_University
[pem_pos=0, final_scores=0.29607272,
final_local_scores=0.27061278, pem=0.0, similarity=12.165338]
13: 2625291 Doctor_of_Music
[pem_pos=0, final_scores=0.06293444,
final_local_scores=0.08413997, pem=0.0, similarity=6.6006174]
14: 28741 Southeast_Asia
[pem_pos=0, final_scores=0.24511947,
final_local_scores=0.30286592, pem=0.0, similarity=13.127836]
1: 26667 Spain
[pem_pos=1, final_scores=0.58765715,
final_local_scores=0.61691004, pem=-1.2425895,
similarity=29.256933]
2: 7954681 Montreal
[pem_pos=0, final_scores=0.7261743,
final_local_scores=0.7210586, pem=-0.11023275,
similarity=26.207003]
3: 5042916 Canada
[pem_pos=0, final_scores=0.7203386,
final_local_scores=0.72681856, pem=-0.11662793,
similarity=26.41367]
4: 41188263 Madrid
[pem_pos=0, final_scores=0.66448104,
final_local_scores=0.6803687, pem=-0.13236846,
similarity=25.113111]
5: 26667 Spain
[pem_pos=0, final_scores=0.63078374,
final_local_scores=0.6640192, pem=-0.18857051,
similarity=24.930847]

6: 64646 Toronto
 [pem_pos=0, final_scores=0.59194213,
 final_local_scores=0.57849956, pem=-0.12291032,
 similarity=22.0217]
 7: 5042916 Canada
 [pem_pos=0, final_scores=0.7316848,
 final_local_scores=0.7392126, pem=-0.11662793,
 similarity=26.783531]
 8: 7954681 Montreal
 [pem_pos=0, final_scores=0.6310207,
 final_local_scores=0.6171178, pem=-0.11023275,
 similarity=23.105202]
 9: 5042916 Canada
 [pem_pos=0, final_scores=0.73669994,
 final_local_scores=0.7446909, pem=-0.11662793,
 similarity=26.947016]
 10: 14579 Indonesia
 [pem_pos=0, final_scores=0.47964036,
 final_local_scores=0.5398058, pem=-0.23151328,
 similarity=21.457602]
 11: 19457 Burma
 [pem_pos=0, final_scores=0.5170887,
 final_local_scores=0.5905549, pem=-0.20192546,
 similarity=22.811152]
 12: 19457 Burma
 [pem_pos=1, final_scores=0.39680004,
 final_local_scores=0.45915776, pem=-1.2286767,
 similarity=24.47363]

Sample: b'May_Peterson_Thompson'

May Esther Peterson Thompson -LRB- October 7 , 1880 --
 October 8 , 1952 -RRB- was an opera singer for the . She was
 born on October 7 , 1880 in [1 1 Oshkosh 1] , [2 Wisconsin 2]
 as May Esther Peterson . She was one of nine children of a [3
 Methodist 3] minister . She began singing in 1884 at church
 meetings and later joined with her sister , Clara Peterson ,
 an organist , to give recitals and concerts . She studied at [1
 1 The Music Conservatory of Chicago College of Performing Ar
 ts 1] and in 1917 joined the . She debuted as Michaela in [2 2
 Carmen 2] on November 29 , 1917 . She toured [4 Europe 4]
 with Frederick Delzell -LRB- piano -RRB- and sang with 2 Opér
 a-Comique in [3 3 Paris 3] . She married on June 9 , 1924 ,
 in [5 3 Bronxville 5] , [6 New York 6] . Afterward they
 traveled to [7 4 Amarillo 7] , [8 Texas 8] to a reception
 held in the ballroom of the , which Thompson owned and had
 built . She retired from the after her marriage , but she
 continued doing concert tours . In 1932 , after her husband
 was appointed to the [4 4 Railroad Commission of Texas 4] ,
 they moved to [9 5 Austin 9] , [10 Texas 10] . On October 1 ,
 1952 she had a 6 cerebral hemorrhage at their 7 summer house
 in [11 8 Estes Park 11] , [12 Colorado 12] , and she was in a
 coma . She was flown back to [13 Austin 13] , where she died

at the on October 8 , 1952 never regaining consciousness .
She was buried in the [5 5 Texas State Cemetery 5] in [6 6 Austin 6]

1: 14244679

The_Music_Conservatory_of_Chicago_College_of_Performing_Arts

2: 37914 Carmen

3: 22989 Paris

4: 2250414 Railroad_Commission_of_Texas

5: 3277500 Texas_State_Cemetery

6: 1998 Austin,_Texas

1: 260786 Oshkosh,_Wisconsin

[winner: 33127 Wisconsin, pem_pos=0,
final_scores=0.5910099, final_local_scores=0.5815506,
pem=-0.21268336, similarity=22.600948]

RECALL MISS: 260786 Oshkosh,_Wisconsin]

[winner: 260786 Oshkosh,_Wisconsin, pem_pos=0,
final_scores=0.42083818, final_local_scores=0.43106437,
pem=-0.229028, similarity=18.199022]

gt: 260786 Oshkosh,_Wisconsin, pem_pos=0,
final_scores=0.42083818, final_local_scores=0.43106437,
pem=-0.229028, similarity=18.199022]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8229342, final_local_scores=-1.805006,
pem=-4.6728287, similarity=-24.363697]

RECALL MISS: 260786 Oshkosh,_Wisconsin]

2: 537627 Opéra-Comique

[winner: 537627 Opéra-Comique, pem_pos=0,
final_scores=-0.0058457702, final_local_scores=0.042951792,
pem=-0.010010108, similarity=5.4259167]

gt: 537627 Opéra-Comique, pem_pos=0,
final_scores=-0.0058457702, final_local_scores=0.042951792,
pem=-0.010010108, similarity=5.4259167]

3: 127546 Bronxville,_New_York

[winner: 127546 Bronxville,_New_York, pem_pos=0,
final_scores=0.34815198, final_local_scores=0.35569823,
pem=0.0, similarity=14.704457]

gt: 127546 Bronxville,_New_York, pem_pos=0,
final_scores=0.34815198, final_local_scores=0.35569823,
pem=0.0, similarity=14.704457]

[winner: 645042 New_York_City, pem_pos=1,
final_scores=0.24758135, final_local_scores=0.19087273,
pem=-1.7906687, similarity=19.52367]

RECALL MISS: 127546 Bronxville,_New_York]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7232654, final_local_scores=-1.6961331,
pem=-4.6728287, similarity=-21.114714]

RECALL MISS: 127546 Bronxville,_New_York]

4: 151232 Amarillo,_Texas

[winner: 29810 Texas, pem_pos=0, final_scores=0.73403686,
final_local_scores=0.72702944, pem=-0.16191302,
similarity=26.66623]

RECALL MISS: 151232 Amarillo,_Texas]

[winner: 151232 Amarillo,_Texas, pem_pos=0,
 final_scores=0.5162489, final_local_scores=0.5022654,
 pem=-0.17648862, similarity=20.038088|
 gt: 151232 Amarillo,_Texas, pem_pos=0,
 final_scores=0.5162489, final_local_scores=0.5022654,
 pem=-0.17648862, similarity=20.038088]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.8081722, final_local_scores=-1.7888808,
 pem=-4.6728287, similarity=-23.882492|
 RECALL MISS: 151232 Amarillo,_Texas]
 5: 1998 Austin,_Texas
 [winner: 29810 Texas, pem_pos=0, final_scores=0.7962772,
 final_local_scores=0.79501754, pem=-0.16191302,
 similarity=28.69513|
 RECALL MISS: 1998 Austin,_Texas]
 [winner: 1998 Austin,_Texas, pem_pos=0,
 final_scores=0.5206009, final_local_scores=0.48439616,
 pem=-0.5521731, similarity=21.547865|
 gt: 1998 Austin,_Texas, pem_pos=0,
 final_scores=0.5206009, final_local_scores=0.48439616,
 pem=-0.5521731, similarity=21.547865]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.8069351, final_local_scores=-1.7875295,
 pem=-4.6728287, similarity=-23.842163|
 RECALL MISS: 1998 Austin,_Texas]
 6: 2959528 Cerebral_hemorrhage
 [winner: 146315 Bleeding, pem_pos=0,
 final_scores=-1.138475, final_local_scores=-1.1099746,
 pem=-0.09400854, similarity=-28.52291|
 gt: 2959528 Cerebral_hemorrhage, pem_pos=1,
 final_scores=-1.8522947, final_local_scores=-2.0143182,
 pem=-2.6669388, similarity=-41.51833]
 [winner: 487841 Cerebrum, pem_pos=1,
 final_scores=-1.2002195, final_local_scores=-1.131105,
 pem=-0.7324413, similarity=-25.681587|
 gt: 2959528 Cerebral_hemorrhage, pem_pos=8,
 final_scores=-2.2241135, final_local_scores=-2.4204738,
 pem=-4.691348, similarity=-42.62977]
 7: 1565816 Summer_house
 [winner: 571415 Ice_stock_sport, pem_pos=2,
 final_scores=-1.2052566, final_local_scores=-1.1699858,
 pem=-0.6931472, similarity=-27.055555|
 RECALL MISS: 1565816 Summer_house]
 [winner: 16752663 House,_North_Carolina, pem_pos=16,
 final_scores=-1.3939744, final_local_scores=-1.352426,
 pem=-4.8077354, similarity=-10.124166|
 RECALL MISS: 1565816 Summer_house]
 8: 108553 Estes_Park,_Colorado
 [winner: 5399 Colorado, pem_pos=0,
 final_scores=0.59632015, final_local_scores=0.5990367,
 pem=-0.26926568, similarity=23.430471|
 RECALL MISS: 108553 Estes_Park,_Colorado]

[winner: 108553 Estes_Park, Colorado, pem_pos=0,
final_scores=0.35687888, final_local_scores=0.38607967,
pem=0.0, similarity=15.611101|
gt: 108553 Estes_Park, Colorado, pem_pos=0,
final_scores=0.35687888, final_local_scores=0.38607967,
pem=0.0, similarity=15.611101]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8044943, final_local_scores=-1.7848632,
pem=-4.6728287, similarity=-23.762596|
RECALL MISS: 108553 Estes_Park, Colorado]

1: 14244679

The_Music_Conservatory_of_Chicago_College_of_Performing_Arts
[pem_pos=0, final_scores=0.10906534,
final_local_scores=0.20806584, pem=0.0, similarity=10.298813]

2: 37914 Carmen

[pem_pos=0, final_scores=0.13717614,
final_local_scores=0.21101305, pem=-0.22672436,
similarity=11.619726]

3: 22989 Paris

[pem_pos=0, final_scores=0.6621738,
final_local_scores=0.69386977, pem=-0.115301244,
similarity=25.423197]

4: 2250414 Railroad_Commission_of_Texas

[pem_pos=0, final_scores=0.05972673,
final_local_scores=0.10773611, pem=0.0, similarity=7.3047733]

5: 3277500 Texas_State_Cemetery

[pem_pos=0, final_scores=0.21676816,
final_local_scores=0.26776415, pem=0.0, similarity=12.080329]

6: 1998 Austin, Texas

[pem_pos=0, final_scores=0.10415731,
final_local_scores=0.029494792, pem=-0.5521731,
similarity=7.9726996]

1: 260786 Oshkosh, Wisconsin

[pem_pos=0, final_scores=0.42083818,
final_local_scores=0.43106437, pem=-0.229028,
similarity=18.199022]

2: 33127 Wisconsin

[pem_pos=0, final_scores=0.5910099,
final_local_scores=0.5815506, pem=-0.21268336,
similarity=22.600948]

3: 20119 Methodism

[pem_pos=0, final_scores=0.31522638,
final_local_scores=0.3272751, pem=0.0, similarity=13.856255]

4: 9239 Europe

[pem_pos=0, final_scores=0.56763726,
final_local_scores=0.63907754, pem=-0.086338006,
similarity=23.63058]

5: 127546 Bronxville, New York

[pem_pos=0, final_scores=0.34815198,
final_local_scores=0.35569823, pem=0.0, similarity=14.704457]

6: 645042 New_York_City

[pem_pos=1, final_scores=0.24758135,

final_local_scores=0.19087273, pem=-1.7906687,
 similarity=19.52367]
 7: 151232 Amarillo, _Texas
 [pem_pos=0, final_scores=0.5162489,
 final_local_scores=0.5022654, pem=-0.17648862,
 similarity=20.038088]
 8: 29810 Texas
 [pem_pos=0, final_scores=0.73403686,
 final_local_scores=0.72702944, pem=-0.16191302,
 similarity=26.66623]
 9: 1998 Austin, _Texas
 [pem_pos=0, final_scores=0.5206009,
 final_local_scores=0.48439616, pem=-0.5521731,
 similarity=21.547865]
 10: 29810 Texas
 [pem_pos=0, final_scores=0.7962772,
 final_local_scores=0.79501754, pem=-0.16191302,
 similarity=28.69513]
 11: 108553 Estes_Park, _Colorado
 [pem_pos=0, final_scores=0.35687888,
 final_local_scores=0.38607967, pem=0.0, similarity=15.611101]
 12: 5399 Colorado
 [pem_pos=0, final_scores=0.59632015,
 final_local_scores=0.5990367, pem=-0.26926568,
 similarity=23.430471]
 13: 1998 Austin, _Texas
 [pem_pos=0, final_scores=0.489129,
 final_local_scores=0.45001787, pem=-0.5521731,
 similarity=20.521946]

Sample: b'Matt_Kohler'

Matthew Robert `` Matt '' Kohler -LRB- born January 12 ,
 1979 -RRB- is an [1 American 1] 1 actor who has starred in
 several film and television productions including , 2 Ocean 's
 Thirteen , 3 Nip/Tuck and [2 4 Betty White 2] 's Off Their Ro
 ckers . Born and raised in [3 5 Port Washington 3] , [4 Wisco
 nsin 4] , approximately 25 miles north of [1 1 Milwaukee 1]
 aside [2 2 Lake Michigan 2] . He has a twin brother , James
 Michael Kohler , who is a fitness model and the 2007 Mr.
 Natural California . Prior to his acting career he earned a
 Bachelor 's Degree in Kinesiology from the and a -LRB- MBA -
 RRB- from [5 6 California State University 5] , [6 Channel Is
 lands 6] . He followed his entrepreneurial spirit by creating
 and opening a tanning salon in 2003 named Baja Tan located in
 [7 7 Fox Point 7] , [8 Wisconsin 8] . During this time he
 modeled as a fashion , commercial and fitness model . Most
 notably working with photographer [3 3 Mario Testino 3] for 8
 VMAN magazine . Matt began his acting career with the desire
 to inspire people through his work , so he studied acting
 rigorously . He recently launched his own production
 company , Silver Hero Entertainment , and is producing and
 developing several projects . His first being a live-action

Trailer for [9 9 Captain Planet 9] and the Planetees in which he played [10 10 Captain Planet 10] and . The Trailer has gone viral , being featured and reviewed on over 100 websites , [4 4 Entertainment Weekly 4] stated ' ' ` Given the choice between a studio remake and this quasi-gritty indie take , I 'd go with the latter . ' ' Matt will produce and star in the feature film titled Stripped . Matt will play the lead Mike ` Sep ' Sepowitz , the hero and calm in the storm . Stripped is a highly stylized neo-noir crime thriller in the narrative vein of 11 The Town -LRB- 2010 film -RRB- meets [5 5 Reservoir Dogs 5] . Matt will produce and star in a thriller film titled Hackathon , a 12 whodunit horror/thriller . It misleads the audience at first -- the film starts off as an upbeat , fast-paced ensemble about a group of intelligent [6 6 Silicon Valley 6] programmer types trying to come up with the next ` big idea ' ' -- then quickly takes a violent turn towards the horror genre when things go horribly wrong and people start dying/disappearing ... hence the double meaning of the title . Matt will play MK , a trust fund , best described as a 13 Winklevoss twins

1: 53117 Milwaukee

2: 17948 Lake_Michigan

3: 3006902 Mario_Testino

4: 541239 Entertainment_Weekly

5: 297111 Reservoir_Dogs

6: 26976 Silicon_Valley

1: 21504235 Actor

[winner: 21504235 Actor, pem_pos=0, final_scores=-1.5764469, final_local_scores=-1.7015779, pem=-0.032278158, similarity=-46.513233|
gt: 21504235 Actor, pem_pos=0, final_scores=-1.5764469, final_local_scores=-1.7015779, pem=-0.032278158, similarity=-46.513233]

2: 3723947 Ocean's_Thirteen

[winner: 468353 Thirteen_(film), pem_pos=0, final_scores=-0.13156164, final_local_scores=-0.102114506, pem=-0.42535043, similarity=3.355534|

RECALL MISS: 3723947 Ocean's_Thirteen]

[winner: 93433 Ocean_County,_New_Jersey, pem_pos=1, final_scores=-0.4531836, final_local_scores=-0.4629655, pem=-1.9708861, similarity=0.9918843|

RECALL MISS: 3723947 Ocean's_Thirteen]

[winner: 1415027 David_Berman_(musician), pem_pos=5, final_scores=-2.1572838, final_local_scores=-2.3187745, pem=-3.0012724, similarity=-48.78575|

RECALL MISS: 3723947 Ocean's_Thirteen]

3: 339955 Nip/Tuck

[winner: 339955 Nip/Tuck, pem_pos=0, final_scores=0.014543965, final_local_scores=0.029302225, pem=0.0, similarity=4.96415|

gt: 339955 Nip/Tuck, pem_pos=0, final_scores=0.014543965, final_local_scores=0.029302225,

pem=0.0, similarity=4.96415]
 4: 34383297 Betty_White's_Off_Their_Rockers
 [winner: 415045 Betty_White, pem_pos=0,
 final_scores=0.2604801, final_local_scores=0.2831336,
 pem=0.0, similarity=12.538984]
 RECALL MISS: 34383297 Betty_White's_Off_Their_Rockers]
 [winner: 39174309 Off_Their_Rockers_(British_TV_series),
 pem_pos=0, final_scores=-0.18218216,
 final_local_scores=-0.051929966, pem=0.0,
 similarity=2.5400193]
 RECALL MISS: 34383297 Betty_White's_Off_Their_Rockers]
 [winner: 1415027 David_Berman_(musician), pem_pos=5,
 final_scores=-2.1500294, final_local_scores=-2.3108501,
 pem=-3.0012724, similarity=-48.549274]
 RECALL MISS: 34383297 Betty_White's_Off_Their_Rockers]
 5: 260603 Port_Washington,_Wisconsin
 [winner: 33127 Wisconsin, pem_pos=0,
 final_scores=0.4669674, final_local_scores=0.47221053,
 pem=-0.21268336, similarity=19.338022]
 RECALL MISS: 260603 Port_Washington,_Wisconsin]
 [winner: 126777 Port_Washington,_New_York, pem_pos=0,
 final_scores=0.17767458, final_local_scores=0.1909968,
 pem=-0.6346485, similarity=13.240756]
 gt: 260603 Port_Washington,_Wisconsin, pem_pos=1,
 final_scores=0.10448752, final_local_scores=0.13645068,
 pem=-1.0824642, similarity=14.048282]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.8331761, final_local_scores=-1.8359323,
 pem=-4.6728287, similarity=-25.286596]
 RECALL MISS: 260603 Port_Washington,_Wisconsin]
 6: 700101 California_State_University,_Channel_Islands
 [winner: 56704 California_State_University, pem_pos=0,
 final_scores=0.23223417, final_local_scores=0.25663713,
 pem=0.0, similarity=11.748277]
 gt: 700101 California_State_University,_Channel_Islands,
 pem_pos=15, final_scores=-0.4127657,
 final_local_scores=-0.38975215, pem=-2.8903718,
 similarity=8.177023]
 [winner: 5643 Channel_Islands, pem_pos=0,
 final_scores=0.20336954, final_local_scores=0.23303846,
 pem=-0.15505362, similarity=11.887251]
 gt: 700101 California_State_University,_Channel_Islands,
 pem_pos=3, final_scores=-0.23374543,
 final_local_scores=-0.20058548, pem=-2.0253844,
 similarity=9.118197]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.8537109, final_local_scores=-1.8583634,
 pem=-4.6728287, similarity=-25.95599]
 RECALL MISS: 700101
 California_State_University,_Channel_Islands]
 7: 139412 Fox_Point,_Wisconsin
 [winner: 33127 Wisconsin, pem_pos=0,

final_scores=0.5874221, final_local_scores=0.603789,
 pem=-0.21268336, similarity=23.264585|
 RECALL MISS: 139412 Fox_Point,_Wisconsin]
 [winner: 139412 Fox_Point,_Wisconsin, pem_pos=0,
 final_scores=0.25080055, final_local_scores=0.32707608,
 pem=-0.21292198, similarity=15.008219|
 gt: 139412 Fox_Point,_Wisconsin, pem_pos=0,
 final_scores=0.25080055, final_local_scores=0.32707608,
 pem=-0.21292198, similarity=15.008219]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7862259, final_local_scores=-1.7846463,
 pem=-4.6728287, similarity=-23.756119|
 RECALL MISS: 139412 Fox_Point,_Wisconsin]
 8: 22912600 VMAN
 [winner: 5121371 V_(American_magazine), pem_pos=0,
 final_scores=-0.024507448, final_local_scores=0.054482564,
 pem=0.0, similarity=5.7155814|
 gt: 22912600 VMAN, pem_pos=1, final_scores=-0.265257,
 final_local_scores=-0.17037132, pem=-1.2656664,
 similarity=5.8883834]
 9: 492218 Captain_Planet_and_the_Planeteers
 [winner: 492218 Captain_Planet_and_the_Planeteers,
 pem_pos=1, final_scores=0.040132925,
 final_local_scores=0.10323763, pem=-0.013612653,
 similarity=7.2445574|
 gt: 492218 Captain_Planet_and_the_Planeteers, pem_pos=1,
 final_scores=0.040132925, final_local_scores=0.10323763,
 pem=-0.013612653, similarity=7.2445574]
 [winner: 492218 Captain_Planet_and_the_Planeteers,
 pem_pos=0, final_scores=-0.100568876,
 final_local_scores=-0.06680712, pem=0.0,
 similarity=2.0960553|
 gt: 492218 Captain_Planet_and_the_Planeteers, pem_pos=0,
 final_scores=-0.100568876, final_local_scores=-0.06680712,
 pem=0.0, similarity=2.0960553]
 [winner: 27870556 Alberto_Guzik, pem_pos=1,
 final_scores=-1.5656413, final_local_scores=-1.6067872,
 pem=-0.6931472, similarity=-40.090584|
 RECALL MISS: 492218 Captain_Planet_and_the_Planeteers]
 10: 534496 Captain_Planet
 [winner: 492218 Captain_Planet_and_the_Planeteers,
 pem_pos=1, final_scores=0.089451775,
 final_local_scores=0.15711096, pem=-0.013612653,
 similarity=8.852245|
 gt: 534496 Captain_Planet, pem_pos=2,
 final_scores=-1.0520415, final_local_scores=-1.041503,
 pem=-5.7889404, similarity=4.490343]
 11: 24228929 The_Town_(2010_film)
 [winner: 24228929 The_Town_(2010_film), pem_pos=0,
 final_scores=0.013845727, final_local_scores=0.049017712,
 pem=-0.04529459, similarity=5.7988186|
 gt: 24228929 The_Town_(2010_film), pem_pos=0,

```

final_scores=0.013845727, final_local_scores=0.049017712,
pem=-0.04529459, similarity=5.7988186 ]
[winner: 30026222 Macbeth_(2010_TV_drama), pem_pos=3,
final_scores=-0.8349404, final_local_scores=-0.7517846,
pem=-1.3805963, similarity=-10.837143]
RECALL MISS: 24228929 The_Town_(2010_film)]
12: 69923 Whodunit
[winner: 69923 Whodunit, pem_pos=0,
final_scores=-1.372969, final_local_scores=-1.4771607,
pem=0.0, similarity=-39.99171]
gt: 69923 Whodunit, pem_pos=0, final_scores=-1.372969,
final_local_scores=-1.4771607, pem=0.0, similarity=-39.99171 ]
13: 26800687 Winklevoss_twins
[winner: 21797439 Cameron_Winklevoss, pem_pos=0,
final_scores=0.0029310435, final_local_scores=0.09179197,
pem=0.0, similarity=6.8289685]
gt: 26800687 Winklevoss_twins, pem_pos=2,
final_scores=-0.5774543, final_local_scores=-0.54622644,
pem=-3.7336934, similarity=8.093631 ]
[winner: 13274566 Twins_(1925_film), pem_pos=11,
final_scores=-1.4199637, final_local_scores=-1.3925884,
pem=-4.009591, similarity=-15.663122]
RECALL MISS: 26800687 Winklevoss_twins]
1: 53117 Milwaukee
[pem_pos=0, final_scores=0.1892597,
final_local_scores=0.16887784, pem=-0.28296277,
similarity=10.66816]
2: 17948 Lake_Michigan
[pem_pos=0, final_scores=0.3400861,
final_local_scores=0.37258238, pem=0.0, similarity=15.208315]
3: 3006902 Mario_Testino
[pem_pos=0, final_scores=0.38707352,
final_local_scores=0.4733044, pem=0.0, similarity=18.21406]
4: 541239 Entertainment_Weekly
[pem_pos=0, final_scores=0.19832407,
final_local_scores=0.24153873, pem=0.0, similarity=11.29771]
5: 297111 Reservoir_Dogs
[pem_pos=0, final_scores=0.23405252,
final_local_scores=0.27285564, pem=0.0, similarity=12.232269]
6: 26976 Silicon_Valley
[pem_pos=0, final_scores=0.31739068,
final_local_scores=0.34785882, pem=0.0, similarity=14.470514]
1: 3434750 United_States
[pem_pos=0, final_scores=0.5671058,
final_local_scores=0.5826863, pem=-0.12994486,
similarity=22.184896]
2: 415045 Betty_White
[pem_pos=0, final_scores=0.2604801,
final_local_scores=0.2831336, pem=0.0, similarity=12.538984]
3: 126777 Port_Washington,_New_York
[pem_pos=0, final_scores=0.17767458,
final_local_scores=0.1909968, pem=-0.6346485,

```

similarity=13.240756]
 4: 33127 Wisconsin
 [pem_pos=0, final_scores=0.4669674,
 final_local_scores=0.47221053, pem=-0.21268336,
 similarity=19.338022]
 5: 56704 California_State_University
 [pem_pos=0, final_scores=0.23223417,
 final_local_scores=0.25663713, pem=0.0, similarity=11.748277]
 6: 5643 Channel_Islands
 [pem_pos=0, final_scores=0.20336954,
 final_local_scores=0.23303846, pem=-0.15505362,
 similarity=11.887251]
 7: 139412 Fox_Point,_Wisconsin
 [pem_pos=0, final_scores=0.25080055,
 final_local_scores=0.32707608, pem=-0.21292198,
 similarity=15.008219]
 8: 33127 Wisconsin
 [pem_pos=0, final_scores=0.5874221,
 final_local_scores=0.603789, pem=-0.21268336,
 similarity=23.264585]
 9: 492218 Captain_Planet_and_the_Planeteers
 [pem_pos=1, final_scores=0.040132925,
 final_local_scores=0.10323763, pem=-0.013612653,
 similarity=7.2445574]
 10: 492218 Captain_Planet_and_the_Planeteers
 [pem_pos=1, final_scores=0.089451775,
 final_local_scores=0.15711096, pem=-0.013612653,
 similarity=8.852245]

Sample: b'Gabriel_Bouvery'

[1 Gabriel Bouvery 1] -LRB- died 1572 -RRB- was a [2 French 2]
], successor to who died 12 April 1540 . intervened with the
 chapter of [1 1 Angers Cathedral 1] , imposing the nomination
 of [3 Bouvery 3] . The king had taken the 1534 1 Affair of the
 Placards badly , as far as his attitude to toleration of [4
 Protestants 4] went . [5 Bouvery 5] was the nephew of [2 2 Gu
 illaume Poyet 2] , a courtier close to the king and son of
 who had been mayor of [3 3 Angers 3] . It was a political
 appointment aimed at royal control in the [4 4 Catholic Church
 4] , in a successor to the late Jean V Olivier . In 1553
 Bouvert consecrated the new church of in [5 5 Saumur 5] ,
 founded by [6 Olivier 6] . He worked to implement the
 decisions of the [6 6 Council of Trent 6] , which he
 attended . Through his uncle , [7 Bouvery 7] became patron of
 [7 7 Guillaume Postel 7] ; he was patron also of [8 8 Jean Bo
 din 8] . On the 1560 accession of [8 2 Charles IX 8] of [9 Fr
 ance 9] , religious war in [10 France 10] came into the
 open . [11 Bouvery 11] put in place an Angevin Catholic
 League , in 1567 : it comprised 59 of the nobility , with
 others including , the . The diocese was spared the worst of
 the violence . Fighting continued to the 3 Peace of [12 Longj
 umeau 12] -LRB- 1568 -RRB- and 4 Peace of Saint-Germain-en-

Laye -LRB- 1570 -RRB- . Bouvery died on 10 February 1572 . He was succeeded by , . Fresh violence in 1572 followed shortly on the of August ; the governor of [9 9 Anjou 9] was then in [13 Paris 13] , but entrusted plans to attack [14 Protestants 14] in [15 Saumur 15] and [16 Angers 16] to , baron of 5 Monts oreau

1: 6593882 Angers_Cathedral

2: 2368755 Guillaume_Poyet

3: 82357 Angers

4: 606848 Catholic_Church

5: 74393 Saumur

6: 6354 Council_of_Trent

7: 575356 Guillaume_Postel

8: 272095 Jean_Bodin

9: 39594 Anjou

1: 843118 Affair_of_the_Placards

[winner: 843118 Affair_of_the_Placards, pem_pos=0, final_scores=-0.180737, final_local_scores=-0.15315571, pem=0.0, similarity=-0.48075897]

gt: 843118 Affair_of_the_Placards, pem_pos=0, final_scores=-0.180737, final_local_scores=-0.15315571, pem=0.0, similarity=-0.48075897]

2: 75908 Charles_IX_of_France

[winner: 5843419 France, pem_pos=0, final_scores=0.5653232, final_local_scores=0.56190264, pem=-0.16240834, similarity=21.741209]

RECALL MISS: 75908 Charles_IX_of_France]

[winner: 75908 Charles_IX_of_France, pem_pos=0, final_scores=0.29659194, final_local_scores=0.2748704, pem=0.0, similarity=12.292394]

gt: 75908 Charles_IX_of_France, pem_pos=0, final_scores=0.29659194, final_local_scores=0.2748704, pem=0.0, similarity=12.292394]

3: 7726893 Peace_of_Longjumeau

[winner: 4457864 Longjumeau, pem_pos=0, final_scores=0.05764763, final_local_scores=0.1107879, pem=0.0, similarity=7.395845]

RECALL MISS: 7726893 Peace_of_Longjumeau]

[winner: 2788270 The_Peace!, pem_pos=0, final_scores=-1.0700009, final_local_scores=-0.9461506, pem=0.0, similarity=-24.145308]

RECALL MISS: 7726893 Peace_of_Longjumeau]

4: 4288521 Peace_of_Saint-Germain-en-Laye

[winner: 75952 Saint-Germain-en-Laye, pem_pos=0, final_scores=-0.10024053, final_local_scores=-0.15525335, pem=0.0, similarity=-0.5433568]

RECALL MISS: 4288521 Peace_of_Saint-Germain-en-Laye]

[winner: 749287 Peace_River, pem_pos=1, final_scores=-1.1526021, final_local_scores=-1.1262393, pem=-1.9109402, similarity=-19.127523]

RECALL MISS: 4288521 Peace_of_Saint-Germain-en-Laye]

5: 3239325 Montsoreau

[winner: 3239325 Montsoreau, pem_pos=0,
final_scores=-0.09083842, final_local_scores=0.07619245,
pem=0.0, similarity=6.3634477]

gt: 3239325 Montsoreau, pem_pos=0,
final_scores=-0.09083842, final_local_scores=0.07619245,
pem=0.0, similarity=6.3634477]

1: 6593882 Angers_Cathedral

[pem_pos=0, final_scores=0.09731786,
final_local_scores=0.11238323, pem=0.0, similarity=7.443453]

2: 2368755 Guillaume_Poyet

[pem_pos=0, final_scores=0.1581705,
final_local_scores=0.252732, pem=0.0, similarity=11.631741]

3: 82357 Angers

[pem_pos=0, final_scores=0.54151875,
final_local_scores=0.5219453, pem=-0.17387784,
similarity=20.611177]

4: 606848 Catholic_Church

[pem_pos=0, final_scores=0.19739674,
final_local_scores=0.1745196, pem=0.0, similarity=9.297726]

5: 74393 Saumur

[pem_pos=0, final_scores=0.3935929,
final_local_scores=0.3846209, pem=0.0, similarity=15.567569]

6: 6354 Council_of_Trent

[pem_pos=0, final_scores=0.22762822,
final_local_scores=0.23982838, pem=0.0, similarity=11.24667]

7: 575356 Guillaume_Postel

[pem_pos=0, final_scores=0.3059414,
final_local_scores=0.33994046, pem=0.0, similarity=14.234214]

8: 272095 Jean_Bodin

[pem_pos=0, final_scores=0.28723007,
final_local_scores=0.32898664, pem=0.0, similarity=13.9073305]

9: 39594 Anjou

[pem_pos=0, final_scores=0.19607358,
final_local_scores=0.15960515, pem=-0.23354693,
similarity=10.122714]

1: 33160828 Gabriel_Bouvery

[pem_pos=0, final_scores=0.031166658,
final_local_scores=0.028811246, pem=0.0, similarity=4.949498]

2: 5843419 France

[pem_pos=0, final_scores=0.5689848,
final_local_scores=0.5659024, pem=-0.80223274,
similarity=25.340036]

3: 33160828 Gabriel_Bouvery

[pem_pos=0, final_scores=0.05335675,
final_local_scores=0.053050563, pem=0.0, similarity=5.6728477]

4: 25814008 Protestantism

[pem_pos=0, final_scores=0.28797942,
final_local_scores=0.27770448, pem=0.0, similarity=12.376968]

5: 33160828 Gabriel_Bouvery

[pem_pos=0, final_scores=0.114069924,
final_local_scores=0.11937049, pem=0.0, similarity=7.651966]

6: 43675 Laurence_Olivier

[pem_pos=0, final_scores=0.091006264,
 final_local_scores=0.16483271, pem=-0.6573949,
 similarity=12.583666]
 7: 33160828 Gabriel_Bouvery
 [pem_pos=0, final_scores=0.033169612,
 final_local_scores=0.030999199, pem=0.0, similarity=5.014791]
 8: 75908 Charles_IX_of_France
 [pem_pos=0, final_scores=0.29659194,
 final_local_scores=0.2748704, pem=0.0, similarity=12.292394]
 9: 5843419 France
 [pem_pos=0, final_scores=0.5653232,
 final_local_scores=0.56190264, pem=-0.16240834,
 similarity=21.741209]
 10: 5843419 France
 [pem_pos=1, final_scores=0.58208257,
 final_local_scores=0.5873345, pem=-0.16240834,
 similarity=22.500147]
 11: 33160828 Gabriel_Bouvery
 [pem_pos=0, final_scores=0.07807307,
 final_local_scores=0.08004938, pem=0.0, similarity=6.478546]
 12: 4457864 Longjumeau
 [pem_pos=0, final_scores=0.05764763,
 final_local_scores=0.1107879, pem=0.0, similarity=7.395845]
 13: 22989 Paris
 [pem_pos=0, final_scores=0.7502019,
 final_local_scores=0.764841, pem=-0.115301244,
 similarity=27.54112]
 14: 25814008 Protestantism
 [pem_pos=0, final_scores=0.14705308,
 final_local_scores=0.12376392, pem=0.0, similarity=7.7830753]
 15: 74393 Saumur
 [pem_pos=0, final_scores=0.28483647,
 final_local_scores=0.26582107, pem=0.0, similarity=12.022344]
 16: 82357 Angers
 [pem_pos=0, final_scores=0.45522374,
 final_local_scores=0.4276811, pem=-0.17387784,
 similarity=17.798145]

Sample: b'Lubov_Egorova'

[1 Lubov Egorova 1] , Princess Nikita Troubetska -LRB-
 August 8 , 1880 -- August 18 , 1972 -RRB- was a [2 Russian 2]
 ballerina who danced with the [1 1 Imperial Ballet 1] and the
 [2 2 Ballets Russes 2] . [3 Lubov Egorova 3] was born in , [3
 3 Russia 3] . She studied ballet at the [4 Imperial Theatre
 School 4] in [5 St. Petersburg 5] with [4 4 Ekaterina Vazem 4]
 , [5 5 Enrico Cecchetti 5] and [6 6 Anna Johansson 6] . After
 graduating in 1898 , she started work as a in the [7 7 Imperi
 al Ballet 7] at and became a in 1914 . A role as [6 Myrtha 6]
 in [8 8 Giselle 8] brought her to the attention of [9 9 Sergei
 Diaghilev 9] who cast her in the role of Princess Florine in
 '' 1 The Sleeping Beauty '' in 1918 , where she danced with [1
 0 10 Vaslav Nijinsky 10] . She also went on to dance other

roles with the [11 11 Ballets Russes 11] . [7 Egorova 7] 's farewell performance in 1917 at the was in [12 12 Swan Lake 12] . However , she continued to dance , and in 1921 she interpreted the role of Aurora in [8 Diaghilev 8] 's Sleeping Princess production in [9 London 9] . After retiring from the stage , she taught as head of the [10 Ballet Russe 10] school in [11 Paris 11] from 1923-1968 , and founded the Ballets de la Jeunesse company in 1937 . She received the Chevalier de l'Ordre des arts et lettres in 1964 . Notable students included [13 13 Serge Lifar 13] , [14 14 Anton Dolin 14] , and 2 Yvonne Mounsey

1: 407551 Mariinsky_Ballet
2: 4802982 Ballets_Russes
3: 25391 Russia
4: 22086526 Ekaterina_Vazem
5: 2604881 Enrico_Cecchetti
6: 34516601 Anna_Johansson
7: 407551 Mariinsky_Ballet
8: 454072 Giselle
9: 28607 Sergei_Diaghilev
10: 63148 Vaslav_Nijinsky
11: 4802982 Ballets_Russes
12: 277945 Swan_Lake
13: 2328615 Serge_Lifar
14: 674708 Anton_Dolin

1: 2518724 The_Sleeping_Beauty_(ballet)
[winner: 2518724 The_Sleeping_Beauty_(ballet), pem_pos=0, final_scores=-0.026062712, final_local_scores=-0.06146764, pem=0.0, similarity=2.255396]

gt: 2518724 The_Sleeping_Beauty_(ballet), pem_pos=0, final_scores=-0.026062712, final_local_scores=-0.06146764, pem=0.0, similarity=2.255396]

2: 37209266 Yvonne_Mounsey
[winner: 37209266 Yvonne_Mounsey, pem_pos=0, final_scores=-0.041559815, final_local_scores=-0.031681083, pem=0.0, similarity=3.1442862]

gt: 37209266 Yvonne_Mounsey, pem_pos=0, final_scores=-0.041559815, final_local_scores=-0.031681083, pem=0.0, similarity=3.1442862]

1: 407551 Mariinsky_Ballet
[pem_pos=0, final_scores=0.289527, final_local_scores=0.22605547, pem=0.0, similarity=10.835659]

2: 4802982 Ballets_Russes
[pem_pos=0, final_scores=0.39751214, final_local_scores=0.34922582, pem=0.0, similarity=14.511308]

3: 25391 Russia
[pem_pos=0, final_scores=0.6797794, final_local_scores=0.79136646, pem=-0.20299491, similarity=28.809584]

4: 22086526 Ekaterina_Vazem
[pem_pos=0, final_scores=0.13977902, final_local_scores=0.176929, pem=0.0, similarity=9.369627]

5: 2604881 Enrico_Cecchetti
[pem_pos=0, final_scores=0.2937104,
final_local_scores=0.29007185, pem=0.0, similarity=12.746036]
6: 34516601 Anna_Johansson
[pem_pos=0, final_scores=0.033354953,
final_local_scores=0.073928714, pem=0.0, similarity=6.295893]
7: 407551 Mariinsky_Ballet
[pem_pos=0, final_scores=0.34976912,
final_local_scores=0.29186085, pem=0.0, similarity=12.799423]
8: 454072 Giselle
[pem_pos=0, final_scores=0.06764291,
final_local_scores=0.02647607, pem=-0.03032179,
similarity=5.0447063]
9: 28607 Sergei_Diaghilev
[pem_pos=0, final_scores=0.3959759,
final_local_scores=0.3462583, pem=0.0, similarity=14.422751]
10: 63148 Vaslav_Nijinsky
[pem_pos=0, final_scores=0.3909883,
final_local_scores=0.3607784, pem=0.0, similarity=14.85606]
11: 4802982 Ballets_Russes
[pem_pos=0, final_scores=0.3140204,
final_local_scores=0.25747877, pem=0.0, similarity=11.773393]
12: 277945 Swan_Lake
[pem_pos=0, final_scores=0.092546895,
final_local_scores=0.045514494, pem=-0.13664111,
similarity=6.191032]
13: 2328615 Serge_Lifar
[pem_pos=0, final_scores=0.500099,
final_local_scores=0.5092692, pem=0.0, similarity=19.287321]
14: 674708 Anton_Dolin
[pem_pos=0, final_scores=0.35747832,
final_local_scores=0.37252426, pem=0.0, similarity=15.20658]
1: 34515786 Lubov_Egorova
[pem_pos=0, final_scores=0.23952834,
final_local_scores=0.2301898, pem=0.0, similarity=10.959036]
2: 25391 Russia
[pem_pos=1, final_scores=0.65987974,
final_local_scores=0.7696291, pem=-1.0515022,
similarity=32.77521]
3: 34515786 Lubov_Egorova
[pem_pos=0, final_scores=0.33097076,
final_local_scores=0.33007672, pem=0.0, similarity=13.93986]
4: 917386 Mariinsky_Theatre
[pem_pos=0, final_scores=0.26838905,
final_local_scores=0.21197245, pem=-0.18232158,
similarity=11.406887]
5: 24320051 Saint_Petersburg
[pem_pos=0, final_scores=0.37443346,
final_local_scores=0.37543803, pem=-0.29571477,
similarity=16.901674]
6: 454072 Giselle
[pem_pos=0, final_scores=0.06902765,

final_local_scores=0.027988672, pem=0.0, similarity=4.9249506]
 7: 34515786 Lubov_Egorova
 [pem_pos=0, final_scores=0.18760301,
 final_local_scores=0.17346925, pem=0.0, similarity=9.266381]
 8: 28607 Sergei_Diaghilev
 [pem_pos=0, final_scores=0.2406074,
 final_local_scores=0.17654186, pem=0.0, similarity=9.358074]
 9: 17867 London
 [pem_pos=0, final_scores=0.53549856,
 final_local_scores=0.5925176, pem=-0.09086737,
 similarity=22.26577]
 10: 4802982 Ballets_Russes
 [pem_pos=0, final_scores=0.21048151,
 final_local_scores=0.1443783, pem=0.0, similarity=8.39825]
 11: 22989 Paris
 [pem_pos=0, final_scores=0.6752188,
 final_local_scores=0.7091311, pem=-0.115301244,
 similarity=25.878626]

Sample: b'Parkway_Center_Mall'

Parkway Center Mall was an enclosed 1 shopping mall in .
 Opened in 1981 , the mall closed in 2013 after losing [1 K-
 Mart 1] , its last 2 anchor store . This has left [2 Giant
 Eagle 2] supermarket as the only store still in operation at
 the mall , having recently signed a five-year lease . The [1 1
 Giant Eagle 1] is accessible through its own entrance doors ,
 not through the mall . Kossman Development opened Parkway
 Center Mall in November 1981 . Original anchor stores
 included [2 2 Gold Circle 2] , [3 3 Giant Eagle 3] , [4 4 Zay
 re 4] , [5 5 Thrift Drug 5] , and a David Weis . Other major
 tenants included 3 Chi-Chi 's , , , [6 6 Payless ShoeSource 6]
 , and . An exit from 4 Interstate 376 was built specifically
 to service the mall . [3 Gold Circle 3] became [7 7 Kmart 7]
 in 1986 , and shortly thereafter the foundation shifted
 resulting a crack in the floor that ran the length of the
 store . 8 inch wide steel plates were used to cover it . [4
 Zayre 4] closed in 1989 along with three other stores in the
 area . fDavid Weis closed in 1990 , and its space became [8 8
 Phar-Mor 8] a year later . Also in 1990 , the former [5 Zayre
 5] became a 5 Kaufmann 's clearance center . By 1992 , the
 clearance store was closed , and one-third of its space was
 converted to [9 9 CompUSA 9] . The rest later became [10 10 Sy
 ms Corporation 10] . In 1997 , [6 Thrift Drug 6] became [11 11
 Eckerd 11] when the latter bought out the chain . In 1999 ,
 the arcade and 6 food court were closed , with the latter
 later converted to a dance studio and a drivers ' license
 office . [7 Syms 7] and [8 Phar-Mor 8] both closed in 2002 ,
 and [9 CompUSA 9] moved to [12 12 Robinson Town Centre 12] .
 By 2006 , the center was largely vacant , with all tenants
 consolidated to the second floor after the first and third
 floors were closed to traffic . The mall building also had
 foundation issues that caused it to shake as the building

settled on the 7 fill dirt on which it was built . Some of the vibrations were caused by truck traffic passing by on the nearby Interstate . [10 Kmart 10] announced its closure in late 2012 , with its last day of operation January 6 , 2013 . After this closure , the last remaining tenants were shuttered or relocated within the month , including [13 13 GNC 13]

1: 1592195 Giant_Eagle
2: 4510801 Gold_Circle
3: 1592195 Giant_Eagle
4: 1563341 Zayre
5: 6904627 Thrift_Drug
6: 2176734 Payless_ShoeSource
7: 242080 Kmart
8: 3594370 Phar-Mor
9: 399531 CompUSA
10: 5263183 Syms_Corporation
11: 768223 Eckerd_Corporation
12: 5302593 Robinson_Town_Centre
13: 1843625 GNC_(store)

1: 235723 Shopping_mall
[winner: 235723 Shopping_mall, pem_pos=0,
final_scores=-0.8156049, final_local_scores=-0.88242745,
pem=0.0, similarity=-22.243683|
gt: 235723 Shopping_mall, pem_pos=0,
final_scores=-0.8156049, final_local_scores=-0.88242745,
pem=0.0, similarity=-22.243683]
[winner: 4910616 Shopping_(film), pem_pos=1,
final_scores=-1.2824916, final_local_scores=-1.2386174,
pem=-1.0934756, similarity=-26.926613|
gt: 235723 Shopping_mall, pem_pos=5,
final_scores=-2.0304565, final_local_scores=-2.2094684,
pem=-5.5320673, similarity=-31.760986]
2: 3484891 Anchor_store
[winner: 3484891 Anchor_store, pem_pos=0,
final_scores=-1.0326937, final_local_scores=-1.1750875,
pem=0.0, similarity=-30.977238|
gt: 3484891 Anchor_store, pem_pos=0,
final_scores=-1.0326937, final_local_scores=-1.1750875,
pem=0.0, similarity=-30.977238]
3: 2028226 Chi-Chi's
[winner: 23877066 Chi-Chi_Olivo, pem_pos=1,
final_scores=-0.13581379, final_local_scores=0.058263183,
pem=-0.65314186, similarity=9.38029|
RECALL MISS: 2028226 Chi-Chi's]
[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.331725, final_local_scores=-2.4784,
pem=-3.4067376, similarity=-51.344315|
RECALL MISS: 2028226 Chi-Chi's]
4: 862530 Interstate_376
[winner: 43950 Interstate_Highway_System, pem_pos=0,
final_scores=-0.21424966, final_local_scores=-0.18490845,

pem=-0.8229691, similarity=3.0471096|
RECALL MISS: 862530 Interstate_376]
[winner: 16077660 Saskatchewan_Highway_376, pem_pos=0,
final_scores=-0.8107064, final_local_scores=-0.67897433,
pem=-0.71631426, similarity=-12.276808|
gt: 862530 Interstate_376, pem_pos=9,
final_scores=-1.5545592, final_local_scores=-1.6394622,
pem=-4.2690616, similarity=-21.619286]
5: 1552913 Kaufmann's
[winner: 5227437 Walter_Kaufmann_(physicist), pem_pos=0,
final_scores=-0.22288565, final_local_scores=-0.049300894,
pem=-1.267531, similarity=9.511507|
RECALL MISS: 1552913 Kaufmann's]
[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.1501849, final_local_scores=-2.2800953,
pem=-3.4067376, similarity=-45.42651|
RECALL MISS: 1552913 Kaufmann's]
6: 13465541 Food_court
[winner: 13465541 Food_court, pem_pos=0,
final_scores=-1.0387771, final_local_scores=-1.1654887,
pem=0.0, similarity=-30.690794|
gt: 13465541 Food_court, pem_pos=0,
final_scores=-1.0387771, final_local_scores=-1.1654887,
pem=0.0, similarity=-30.690794]
7: 2614524 Fill_dirt
[winner: 13510596 Dirt_cake, pem_pos=8,
final_scores=-1.0849783, final_local_scores=-0.97907835,
pem=-2.847508, similarity=-9.642745|
RECALL MISS: 2614524 Fill_dirt]
[winner: 2877773 Fill_(music), pem_pos=0,
final_scores=-1.2935636, final_local_scores=-1.2537196,
pem=0.0, similarity=-33.323776|
gt: 2614524 Fill_dirt, pem_pos=1,
final_scores=-1.3078315, final_local_scores=-1.2778313,
pem=-0.6020534, similarity=-30.76926]
1: 1592195 Giant_Eagle
[pem_pos=0, final_scores=0.32104266,
final_local_scores=0.28641015, pem=0.0, similarity=12.636764]
2: 4510801 Gold_Circle
[pem_pos=0, final_scores=0.12373687,
final_local_scores=0.14255941, pem=0.0, similarity=8.34397]
3: 1592195 Giant_Eagle
[pem_pos=0, final_scores=0.3285063,
final_local_scores=0.294563, pem=0.0, similarity=12.88006]
4: 1563341 Zayre
[pem_pos=0, final_scores=0.24100326,
final_local_scores=0.20678258, pem=0.0, similarity=10.260517]
5: 6904627 Thrift_Drug
[pem_pos=0, final_scores=0.04336013,
final_local_scores=0.046514645, pem=0.0, similarity=5.4778028]
6: 2176734 Payless_ShoeSource
[pem_pos=0, final_scores=0.04531227,

final_local_scores=0.026542604, pem=0.0, similarity=4.8817973]
7: 242080 Kmart
[pem_pos=0, final_scores=0.28284335,
final_local_scores=0.22799265, pem=0.0, similarity=10.893469]
8: 3594370 Phar-Mor
[pem_pos=0, final_scores=0.15402763,
final_local_scores=0.12905875, pem=0.0, similarity=7.9410834]
9: 399531 CompUSA
[pem_pos=0, final_scores=0.06569184,
final_local_scores=0.053025767, pem=0.0, similarity=5.6721077]
10: 5263183 Syms_Corporation
[pem_pos=0, final_scores=0.29020602,
final_local_scores=0.33520883, pem=0.0, similarity=14.093013]
11: 768223 Eckerd_Corporation
[pem_pos=1, final_scores=0.19069253,
final_local_scores=0.16798064, pem=-0.56485415,
similarity=12.1743555]
12: 5302593 Robinson_Town_Centre
[pem_pos=0, final_scores=0.08546959,
final_local_scores=0.14203748, pem=0.0, similarity=8.328395]
13: 1843625 GNC_(store)
[pem_pos=0, final_scores=0.12593426,
final_local_scores=0.104308814, pem=0.0, similarity=7.202496]
1: 242080 Kmart
[pem_pos=0, final_scores=0.43982762,
final_local_scores=0.40139925, pem=0.0, similarity=16.068268]
2: 1592195 Giant_Eagle
[pem_pos=0, final_scores=0.32369357,
final_local_scores=0.28930584, pem=0.0, similarity=12.723176]
3: 4510801 Gold_Circle
[pem_pos=0, final_scores=0.10665415,
final_local_scores=0.12389913, pem=0.0, similarity=7.7871103]
4: 1563341 Zayre
[pem_pos=0, final_scores=0.24452274,
final_local_scores=0.21062711, pem=0.0, similarity=10.375246]
5: 1563341 Zayre
[pem_pos=0, final_scores=0.19792823,
final_local_scores=0.15972966, pem=0.0, similarity=8.856364]
6: 6904627 Thrift_Drug
[pem_pos=0, final_scores=0.05934365,
final_local_scores=0.06397422, pem=0.0, similarity=5.9988313]
7: 5263183 Syms_Corporation
[pem_pos=0, final_scores=0.14911126,
final_local_scores=0.18108428, pem=0.0, similarity=9.4936285]
8: 3594370 Phar-Mor
[pem_pos=0, final_scores=0.13711958,
final_local_scores=0.110589266, pem=0.0, similarity=7.3899174]
9: 399531 CompUSA
[pem_pos=0, final_scores=0.13561715,
final_local_scores=0.12940854, pem=0.0, similarity=7.9515214]
10: 242080 Kmart
[pem_pos=0, final_scores=0.3343497,

final_local_scores=0.28425553, pem=0.0, similarity=12.572465]

Sample: b'Frances_Carpenter'

Frances Aretta Carpenter -LRB- April 30 , 1890 -- November 2 , 1972 -RRB- was an [1 1 American 1] 1 photographer and 2 author . She was the daughter of , and worked closely with , [2 2 Frank G. Carpenter 2] . Frances Aretta Carpenter was born in [1 3 Washington 1] , [2 D.C. 2] in 1890 . Her mother was Joanna Condict . Her father was [3 3 Frank G. Carpenter 3] . In 1912 she graduated from [3 Smith College 3] . She photographed subjects for her fathers books . She traveled with him from 1908 until 1915 . They traveled throughout the world . On April 6 , 1920 , Carpenter married William Chapin Huntington . He worked at the 4 Embassy of the [4 United States 4] , [5 Paris 5] and the [4 4 United States Foreign Service 4] from 1920 until 1961 . Carpenter traveled with [6 Huntington 6] . W. Chapin Huntington died in 1958 . The couple had two children : Joanna Huntington Noel and Edith Chapin Huntington Williams . In the early 1960s she visited [5 5 Canada 5] and the . In 1964 she visited [6 6 Africa 6] and traveled throughout the continent by car . In 1966 they visited [7 7 Japan 7] and [8 8 Korea 8] . She edited work for her father , including Carp 's [7 Washington 7] , a memoir of his about working in [8 Washington 8] , [9 D.C. 9] . It was published in 1960 . She wrote s and work about women artists . Carpenter was a fellow for the [9 9 Royal Geographical Society 9] . She was president of the Smith College Alumnae Association . She was also on their [10 Board of Trustees 10] from 1936 until 1944 . From 1960 until 1930 she was on the Board of Counselors . She also served as vice president of the from 1939 until 1942 . She was also a member of the 5 Sulgrave Club , the [10 10 Cosmos Club 10] , the and the [11 11 Cosmopolitan Club 11] . She died on November 2 , 1972 . Carpentered is interred at [12 12 Oak Hill Cemetery 12] , [11 Washington 11] , [12 D.C. 12] . The Frances Carpenter Papers are held in the collection of the [13 Sophia Smith Collection 13] at [14 Smith College 14] . The [13 13 Library of Congress 13] has a collection of approximately 7,000 6 negatives

1: 3434750 United_States

2: 4557557 Frank_G._Carpenter

3: 4557557 Frank_G._Carpenter

4: 472939 United_States_Foreign_Service

5: 5042916 Canada

6: 5334607 Africa

7: 15573 Japan

8: 16749 Korea

9: 395918 Royal_Geographical_Society

10: 3219596 Cosmos_Club

11: 16076757 Cosmopolitan_Club_(New_York)

12: 2926443 Oak_Hill_Cemetery_(Washington,_D.C.)

13: 18944081 Library_of_Congress

1: 84303 Photographer

[winner: 84303 Photographer, pem_pos=0,
final_scores=-1.6360984, final_local_scores=-1.8066223,
pem=-0.1463651, similarity=-49.027546]

gt: 84303 Photographer, pem_pos=0,
final_scores=-1.6360984, final_local_scores=-1.8066223,
pem=-0.1463651, similarity=-49.027546]

2: 914 Author

[winner: 10653999 Author_(music), pem_pos=14,
final_scores=-0.926248, final_local_scores=-0.79401296,
pem=-2.6390574, similarity=-5.2536125]

gt: 914 Author, pem_pos=0, final_scores=-1.2335751,
final_local_scores=-1.3673918, pem=-0.07582591,
similarity=-36.303635]

3: 108956 Washington, D.C.

[winner: 108956 Washington, D.C., pem_pos=0,
final_scores=0.50403684, final_local_scores=0.44952935,
pem=-1.7496481, similarity=27.019423]

gt: 108956 Washington, D.C., pem_pos=0,
final_scores=0.50403684, final_local_scores=0.44952935,
pem=-1.7496481, similarity=27.019423]

[winner: 108956 Washington, D.C., pem_pos=0,
final_scores=0.14920129, final_local_scores=0.061925396,
pem=-0.15729709, similarity=6.793096]

gt: 108956 Washington, D.C., pem_pos=0,
final_scores=0.14920129, final_local_scores=0.061925396,
pem=-0.15729709, similarity=6.793096]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8086321, final_local_scores=-1.8031039,
pem=-4.6728287, similarity=-24.306934]

RECALL MISS: 108956 Washington, D.C.]

4: 4627835 Embassy_of_the_United_States,_Paris

[winner: 22989 Paris, pem_pos=0, final_scores=0.5865783,
final_local_scores=0.60589755, pem=-0.115301244,
similarity=22.79793]

RECALL MISS: 4627835

Embassy_of_the_United_States,_Paris]

[winner: 3434750 United_States, pem_pos=0,
final_scores=0.5196722, final_local_scores=0.4960119,
pem=-0.073457584, similarity=19.291172]

RECALL MISS: 4627835

Embassy_of_the_United_States,_Paris]

[winner: 8970 Diplomatic_mission, pem_pos=0,
final_scores=-0.17892289, final_local_scores=-0.19571048,
pem=-0.306012, similarity=-0.08653736]

RECALL MISS: 4627835

Embassy_of_the_United_States,_Paris]

[winner: 1961717 Supreme_Court_of_South_Australia,
pem_pos=0, final_scores=-1.2133304,
final_local_scores=-1.2674847, pem=0.0, similarity=-33.73456]

RECALL MISS: 4627835

Embassy_of_the_United_States,_Paris]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.808016, final_local_scores=-1.8024309,
pem=-4.6728287, similarity=-24.28685]
RECALL MISS: 4627835 Embassy_of_the_United_States,_Paris]
5: 24087019 Sulgrave_Club
[winner: 24087019 Sulgrave_Club, pem_pos=0,
final_scores=-0.004737258, final_local_scores=0.07102935,
pem=0.0, similarity=6.2093706]
gt: 24087019 Sulgrave_Club, pem_pos=0,
final_scores=-0.004737258, final_local_scores=0.07102935,
pem=0.0, similarity=6.2093706]
6: 407044 Negative_(photography)
[winner: 407044 Negative_(photography), pem_pos=0,
final_scores=-1.1484998, final_local_scores=-1.175626,
pem=0.0, similarity=-30.993313]
gt: 407044 Negative_(photography), pem_pos=0,
final_scores=-1.1484998, final_local_scores=-1.175626,
pem=0.0, similarity=-30.993313]
1: 3434750 United_States
[pem_pos=0, final_scores=0.61656743,
final_local_scores=0.6018552, pem=-0.12994486,
similarity=22.756935]
2: 4557557 Frank_G._Carpenter
[pem_pos=0, final_scores=0.08663373,
final_local_scores=0.16290644, pem=0.0, similarity=8.951166]
3: 4557557 Frank_G._Carpenter
[pem_pos=0, final_scores=0.16230021,
final_local_scores=0.24556056, pem=0.0, similarity=11.417729]
4: 472939 United_States_Foreign_Service
[pem_pos=0, final_scores=0.24541758,
final_local_scores=0.24234188, pem=0.0, similarity=11.321678]
5: 5042916 Canada
[pem_pos=0, final_scores=0.6919791,
final_local_scores=0.71993804, pem=-0.11662793,
similarity=26.20834]
6: 5334607 Africa
[pem_pos=0, final_scores=0.60587037,
final_local_scores=0.67005616, pem=-0.13197383,
similarity=24.80322]
7: 15573 Japan
[pem_pos=0, final_scores=0.6582016,
final_local_scores=0.7232322, pem=-0.13549899,
similarity=26.40927]
8: 16749 Korea
[pem_pos=0, final_scores=0.3939913,
final_local_scores=0.44109064, pem=-0.27902022,
similarity=18.770094]
9: 395918 Royal_Geographical_Society
[pem_pos=0, final_scores=0.16960742,
final_local_scores=0.17558858, pem=0.0, similarity=9.329626]
10: 3219596 Cosmos_Club
[pem_pos=0, final_scores=0.119233176,

final_local_scores=0.12962192, pem=0.0, similarity=7.9578896]
11: 16076757 Cosmopolitan_Club_(New_York)
[pem_pos=0, final_scores=0.038478866,
final_local_scores=0.12535882, pem=0.0, similarity=7.83067]
12: 2926443 Oak_Hill_Cemetery_(Washington,_D.C.)
[pem_pos=0, final_scores=0.18962322,
final_local_scores=0.24930769, pem=-0.016699152,
similarity=11.620364]
13: 18944081 Library_of_Congress
[pem_pos=0, final_scores=0.29336226,
final_local_scores=0.28612456, pem=0.0, similarity=12.628241]
1: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.50403684,
final_local_scores=0.44952935, pem=-1.7496481,
similarity=27.019423]
2: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.14920129,
final_local_scores=0.061925396, pem=-0.15729709,
similarity=6.793096]
3: 247957 Smith_College
[pem_pos=0, final_scores=0.41330743,
final_local_scores=0.40867978, pem=0.0, similarity=16.285532]
4: 3434750 United_States
[pem_pos=0, final_scores=0.5196722,
final_local_scores=0.4960119, pem=-0.073457584,
similarity=19.291172]
5: 22989 Paris
[pem_pos=0, final_scores=0.5865783,
final_local_scores=0.60589755, pem=-0.115301244,
similarity=22.79793]
6: 93552 Huntington,_West_Virginia
[pem_pos=0, final_scores=0.055884525,
final_local_scores=0.036828697, pem=-1.8056281,
similarity=15.00804]
7: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.22371961,
final_local_scores=0.14332533, pem=-1.7496481,
similarity=17.881685]
8: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.35528678,
final_local_scores=0.28704244, pem=-1.7496481,
similarity=22.17049]
9: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.15247224,
final_local_scores=0.0654984, pem=-0.15729709,
similarity=6.8997216]
10: 4822 Board_of_directors
[pem_pos=0, final_scores=0.15195386,
final_local_scores=0.14943713, pem=-0.59147835,
similarity=11.765766]
11: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.36090946,

final_local_scores=0.2931844, pem=-1.7496481,
 similarity=22.353779]
 12: 108956 Washington,_D.C.
 [pem_pos=0, final_scores=0.07108192,
 final_local_scores=-0.025762044, pem=-0.15729709,
 similarity=4.176328]
 13: 9618008 Sophia_Smith_Collection
 [pem_pos=0, final_scores=0.11848657,
 final_local_scores=0.17146066, pem=0.0, similarity=9.206441]
 14: 247957 Smith_College
 [pem_pos=0, final_scores=0.3473649,
 final_local_scores=0.3366476, pem=0.0, similarity=14.135949]

Sample: b'2009_Superfinalen'

The 2009 Superfinalen was the first edition of **1 Superfinalen**
 , an annual **2 football** match contested by the winners of the
 previous season 's **[1 1 Tippeligaen 1]** and **[2 2 Norwegian Cup**
2] competitions . It was played on 8 March 2009 , at **[3 3 Tel**
enor Arena 3] in **[4 4 Bærum 4]** , between **[5 5 Stabæk 5]** and **[6**
6 Vålerenga 6] , winners of the and **3 2008 [1 Norwegian Footb**
all Cup 1] respectively . A collaboration between the **[7 7 Fo**
otball Association of Norway 7] , **[8 8 Norsk Toppfotball 8]** ,
 the incumbent league and cup champions , and the broadcaster **[**
9 9 TV 2 9] , the match was seen as a test of the viability
 of a in **[2 Norway 2]** . Modeled on the **[10 10 FA Community Shi**
eld 10] , the Superfinalen is intended to be a season
 opener , with the net proceeds of each match going to
 charity . In 2009 , these proceeds were donated to **[11 11 UNI**
CEF 11] . The game was a repeat of the **4 2008 Norwegian Footb**
all Cup Final , in which **[3 Vålerenga 3]** had defeated **[4**
Stabæk 4] 4 -- 1 . This time , however , it was **[5 Stabæk 5]**
 who won the match 3 -- 1 after goals by **[12 12 Daniel Nannskog**
12] , **[13 13 Daigo Kobayashi 13]** and **[14 14 Pálmi Rafn Pálmas**
on 14] . **[15 15 Mohammed Abdellaoue 15]** scored for the losing
 side . The game also served as the official opening match of **[**
16 16 Telenor Arena 16]

1: 603971 Tippeligaen
 2: 603997 Norwegian_Football_Cup
 3: 7752300 Telenor_Arena
 4: 73092 Bærum
 5: 318926 Stabæk_Fotball
 6: 337678 Vålerenga_Fotball
 7: 678005 Football_Association_of_Norway
 8: 6498795 Norsk_Toppfotball
 9: 384208 TV_2_(Norway)
 10: 291075 FA_Community_Shield
 11: 19698110 UNICEF
 12: 2140716 Daniel_Nannskog
 13: 17177603 Daigo_Kobayashi
 14: 18381569 Pálmi_Rafn_Pálmason
 15: 6054878 Mohammed_Abdellaoue
 16: 7752300 Telenor_Arena

1: 21969292 Superfinalen
[winner: 21969292 Superfinalen, pem_pos=0,
final_scores=-0.17139873, final_local_scores=-0.16428412,
pem=0.0, similarity=-0.81285316]
gt: 21969292 Superfinalen, pem_pos=0,
final_scores=-0.17139873, final_local_scores=-0.16428412,
pem=0.0, similarity=-0.81285316]
2: 10568 Association_football
[winner: 10568 Association_football, pem_pos=0,
final_scores=-1.4301335, final_local_scores=-1.5578547,
pem=0.0, similarity=-42.39978]
gt: 10568 Association_football, pem_pos=0,
final_scores=-1.4301335, final_local_scores=-1.5578547,
pem=0.0, similarity=-42.39978]
3: 17263188 2008_Norwegian_Football_Cup
[winner: 603997 Norwegian_Football_Cup, pem_pos=0,
final_scores=0.28524423, final_local_scores=0.24515113,
pem=0.0, similarity=11.405512]
RECALL MISS: 17263188 2008_Norwegian_Football_Cup]
[winner: 35825 2008, pem_pos=0, final_scores=-1.0354234,
final_local_scores=-0.95048636, pem=-0.29331458,
similarity=-22.679605]
RECALL MISS: 17263188 2008_Norwegian_Football_Cup]
4: 19635039 2008_Norwegian_Football_Cup_Final
[winner: 19635378 2000_Norwegian_Football_Cup_Final,
pem_pos=0, final_scores=-0.063446686,
final_local_scores=-0.012334898, pem=0.0,
similarity=3.721615]
RECALL MISS: 19635039 2008_Norwegian_Football_Cup_Final]
[winner: 35825 2008, pem_pos=0, final_scores=-1.0420606,
final_local_scores=-0.95773655, pem=-0.29331458,
similarity=-22.895966]
RECALL MISS: 19635039 2008_Norwegian_Football_Cup_Final]
1: 603971 Tippeligaen
[pem_pos=0, final_scores=0.3304718,
final_local_scores=0.2782287, pem=0.0, similarity=12.392612]
2: 603997 Norwegian_Football_Cup
[pem_pos=0, final_scores=0.34707075,
final_local_scores=0.31268725, pem=0.0, similarity=13.420924]
3: 7752300 Telenor_Arena
[pem_pos=0, final_scores=0.24381976,
final_local_scores=0.2474738, pem=0.0, similarity=11.474825]
4: 73092 Bærum
[pem_pos=0, final_scores=0.4482532,
final_local_scores=0.4921248, pem=-0.051573426,
similarity=19.056162]
5: 318926 Stabæk_Fotball
[pem_pos=0, final_scores=0.31127518,
final_local_scores=0.2604076, pem=-0.10377823,
similarity=12.425158]
6: 337678 Vålerenga_Fotball
[pem_pos=0, final_scores=0.5012827,

final_local_scores=0.46234846, pem=-0.357853,
similarity=19.833174]
7: 678005 Football_Association_of_Norway
[pem_pos=0, final_scores=0.29940754,
final_local_scores=0.29691216, pem=0.0, similarity=12.950164]
8: 6498795 Norsk_Toppfotball
[pem_pos=0, final_scores=0.10971229,
final_local_scores=0.16274095, pem=0.0, similarity=8.946227]
9: 384208 TV_2_(Norway)
[pem_pos=0, final_scores=0.1445571,
final_local_scores=0.1524575, pem=-0.21610944,
similarity=9.814586]
10: 291075 FA_Community_Shield
[pem_pos=0, final_scores=0.21462075,
final_local_scores=0.25753722, pem=0.0, similarity=11.775137]
11: 19698110 UNICEF
[pem_pos=0, final_scores=0.038789555,
final_local_scores=0.13531628, pem=0.0, similarity=8.127821]
12: 2140716 Daniel_Nannskog
[pem_pos=0, final_scores=0.68396026,
final_local_scores=0.7350753, pem=0.0, similarity=26.025826]
13: 17177603 Daigo_Kobayashi
[pem_pos=0, final_scores=0.6965315,
final_local_scores=0.7815, pem=0.0, similarity=27.411232]
14: 18381569 Pálmi_Rafn_Pálmason
[pem_pos=0, final_scores=0.2987389,
final_local_scores=0.3649103, pem=0.0, similarity=14.979363]
15: 6054878 Mohammed_Abdellaoue
[pem_pos=0, final_scores=0.5485429,
final_local_scores=0.5996347, pem=0.0, similarity=21.984009]
16: 7752300 Telenor_Arena
[pem_pos=0, final_scores=0.13849632,
final_local_scores=0.13242397, pem=0.0, similarity=8.041509]
1: 603997 Norwegian_Football_Cup
[pem_pos=0, final_scores=0.28524423,
final_local_scores=0.24515113, pem=0.0, similarity=11.405512]
2: 21241 Norway
[pem_pos=1, final_scores=0.5182232,
final_local_scores=0.5675173, pem=-0.22668563,
similarity=22.258312]
3: 337678 Vålerenga_Fotball
[pem_pos=0, final_scores=0.43671316,
final_local_scores=0.39181608, pem=-0.357853,
similarity=17.728346]
4: 318926 Stabæk_Fotball
[pem_pos=0, final_scores=0.3713494,
final_local_scores=0.32602954, pem=-0.10377823,
similarity=14.383448]
5: 318926 Stabæk_Fotball
[pem_pos=0, final_scores=0.27091563,
final_local_scores=0.21632087, pem=-0.10377823,
similarity=11.109522]

Sample: b'Alfred_Schickel'

Alfred Schickel -LRB- born 18 June 1933 in 1 Ústí nad Labem/
Aussig an der [1 Elbe 1] , [1 1 Czechoslovakia 1] -RRB- is a [2
German 2] 2 revisionist historian and Scientific Director
of the 3 Zeitgeschichtliche Forschungsstelle [3 Ingolstadt 3]
. Schickel attended [2 2 Kolleg St. Blasien 2] , a school ,
from 1947 till graduating in 1954 and studied history and
philosophy in [3 3 Munich 3] . After finishing his studies he
taught at a private school in [4 4 Ingolstadt 4] . In 1966
Schickel earned a doctorate in with a dissertation on 4 Roman
legal history . His research has since focused on 5 contempor
ary history , particularly the history of [4 Central Europe 4]
. From 1974 till 1995 he was head of a educational
institution , the Katholisches Bildungswerk Ingolstadt .
Schickel , a , experienced the and came as a refugee to [5
West Germany 5] . Nominated by the [6 Bavarian 6] government
he was decorated with the in 1989 . The award was criticized
by the leading [7 German 7] newsmagazine '' [5 5 Der Spiegel 5
] '' as Schickel was allegedly working towards an exculpation
of [8 Nazism 8] . [6 6 Heinz Nawratil 6] wrote a `` A `
Diatribes ' in Honor of Dr. [9 Alfred Schickel 9] '' published
in the [10 6 Journal of Historical Review 10] . Schickel has
written an 7 obituary for [7 7 Hellmut Diwald 7] , published
in the far-right 8 National Zeitung

1: 5322 Czechoslovakia

2: 9388891 Kolleg_St._Blasien

3: 19058 Munich

4: 208251 Ingolstadt

5: 210137 Der_Spiegel

6: 24095536 Heinz_Nawratil

7: 25376770 Hellmut_Diwald

1: 248718 Ústí_nad_Labem

[winner: 50759 Elbe, pem_pos=0, final_scores=0.2511078,
final_local_scores=0.2634609, pem=0.0, similarity=11.951912]

RECALL MISS: 248718 Ústí_nad_Labem]

[winner: 23858801 Ústí (Přerov_District), pem_pos=1,
final_scores=-0.51138365, final_local_scores=-0.38885847,
pem=-1.0548553, similarity=-1.7781317]

gt: 248718 Ústí_nad_Labem, pem_pos=3,
final_scores=-0.98926836, final_local_scores=-1.0666158,
pem=-3.800293, similarity=-7.073645]

[winner: 8338601 Nad, County_Cork, pem_pos=0,
final_scores=-0.8750237, final_local_scores=-0.74857914,
pem=0.0, similarity=-18.249382]

RECALL MISS: 248718 Ústí_nad_Labem]

[winner: 7454672 Der_(Sumer), pem_pos=0,
final_scores=-1.1166667, final_local_scores=-1.0849941,
pem=-0.39436623, similarity=-26.14405]

RECALL MISS: 248718 Ústí_nad_Labem]

2: 13293 Historical_revisionism

[winner: 13293 Historical_revisionism, pem_pos=0,

final_scores=-1.1050727, final_local_scores=-1.1848891,
 pem=0.0, similarity=-31.269743|
 gt: 13293 Historical_revisionism, pem_pos=0,
 final_scores=-1.1050727, final_local_scores=-1.1848891,
 pem=0.0, similarity=-31.269743]
 3: 16306159 Zeitgeschichtliche_Forschungsstelle_Ingolstadt
 [winner: 208251 Ingolstadt, pem_pos=0,
 final_scores=0.16977291, final_local_scores=0.12804237,
 pem=0.0, similarity=7.910753|
 RECALL MISS: 16306159
 Zeitgeschichtliche_Forschungsstelle_Ingolstadt]
 4: 26364 Roman_law
 [winner: 521555 Ancient_Rome, pem_pos=0,
 final_scores=-0.06326769, final_local_scores=-0.034206346,
 pem=-1.0249887, similarity=8.642975|
 gt: 26364 Roman_law, pem_pos=14,
 final_scores=-0.7996283, final_local_scores=-0.84862214,
 pem=-4.983691, similarity=5.8672237]
 [winner: 457101 Samuel_Nicholas, pem_pos=2,
 final_scores=-1.219765, final_local_scores=-1.2125971,
 pem=-2.1118069, similarity=-20.61227|
 RECALL MISS: 26364 Roman_law]
 [winner: 866993 Legal_psychology, pem_pos=3,
 final_scores=-1.8857657, final_local_scores=-1.9376264,
 pem=-2.2870862, similarity=-41.29539|
 RECALL MISS: 26364 Roman_law]
 5: 10689543 Contemporary_history
 [winner: 2518639 Contemporary_French_literature,
 pem_pos=0, final_scores=-0.9598003,
 final_local_scores=-0.9111119, pem=-1.3518424,
 similarity=-15.748154|
 gt: 10689543 Contemporary_history, pem_pos=9,
 final_scores=-1.4868832, final_local_scores=-1.5989397,
 pem=-2.3493586, similarity=-30.849659]
 [winner: 5548817 Brothers_Hildebrandt, pem_pos=1,
 final_scores=-1.2869619, final_local_scores=-1.2613473,
 pem=-2.0858314, similarity=-22.208334|
 RECALL MISS: 10689543 Contemporary_history]
 6: 1478156 Journal_of_Historical_Review
 [winner: 65662 Institute_for_Historical_Review,
 pem_pos=1, final_scores=0.15426825,
 final_local_scores=0.23557708, pem=-0.6270854,
 similarity=14.529991|
 gt: 1478156 Journal_of_Historical_Review, pem_pos=0,
 final_scores=0.0510671, final_local_scores=0.16627142,
 pem=0.0, similarity=9.051583]
 7: 379881 Obituary
 [winner: 23773124 Knut_Hamsun's_obituary_of_Adolf_Hitler,
 pem_pos=0, final_scores=-1.1505678,
 final_local_scores=-1.20773, pem=0.0, similarity=-31.951363|
 RECALL MISS: 379881 Obituary]
 8: 30863971 National_Zeitung

```

[winner: 30863971 National_Zeitung, pem_pos=0,
final_scores=-0.04459448, final_local_scores=0.059330896,
pem=0.0, similarity=5.8602653]
gt: 30863971 National_Zeitung, pem_pos=0,
final_scores=-0.04459448, final_local_scores=0.059330896,
pem=0.0, similarity=5.8602653 ]
1: 5322 Czechoslovakia
[pem_pos=0, final_scores=0.2984882,
final_local_scores=0.2719124, pem=-0.34588256,
similarity=14.085085]
2: 9388891 Kolleg_St._Blasien
[pem_pos=0, final_scores=0.2250366,
final_local_scores=0.27447018, pem=0.0, similarity=12.280451]
3: 19058 Munich
[pem_pos=0, final_scores=0.5713411,
final_local_scores=0.55052716, pem=-0.119245164,
similarity=21.167015]
4: 208251 Ingolstadt
[pem_pos=0, final_scores=0.35530633,
final_local_scores=0.33070943, pem=0.0, similarity=13.958742]
5: 210137 Der_Spiegel
[pem_pos=0, final_scores=0.42459476,
final_local_scores=0.45103997, pem=0.0, similarity=17.549644]
6: 24095536 Heinz_Nawratil
[pem_pos=0, final_scores=0.3120461,
final_local_scores=0.3982668, pem=0.0, similarity=15.974788]
7: 25376770 Hellmut_Diwald
[pem_pos=0, final_scores=0.3699721,
final_local_scores=0.45094955, pem=0.0, similarity=17.546947]
1: 50759 Elbe
[pem_pos=0, final_scores=0.2511078,
final_local_scores=0.2634609, pem=0.0, similarity=11.951912]
2: 11867 Germany
[pem_pos=1, final_scores=0.8257579,
final_local_scores=0.8174464, pem=-1.0465823,
similarity=34.175423]
3: 208251 Ingolstadt
[pem_pos=0, final_scores=0.16977291,
final_local_scores=0.12804237, pem=0.0, similarity=7.910753]
4: 5188 Central_Europe
[pem_pos=0, final_scores=0.29824203,
final_local_scores=0.30548796, pem=0.0, similarity=13.206083]
5: 33166 West_Germany
[pem_pos=0, final_scores=0.3230813,
final_local_scores=0.2826164, pem=-0.2963801,
similarity=14.135311]
6: 3764 Bavaria
[pem_pos=0, final_scores=0.58955,
final_local_scores=0.57305974, pem=-0.05476965,
similarity=21.488804]
7: 11867 Germany
[pem_pos=2, final_scores=0.6417426,

```

```

final_local_scores=0.6217591, pem=-1.0465823,
similarity=28.335724]
8: 31045316 Nazism
    [pem_pos=0, final_scores=0.32889348,
final_local_scores=0.30051336, pem=0.0, similarity=13.057631]
9: 25307455 Alfred_Schickel
    [pem_pos=0, final_scores=0.22813533,
final_local_scores=0.27709097, pem=0.0, similarity=12.35866]
10: 65662 Institute_for_Historical_Review
    [pem_pos=1, final_scores=0.15426825,
final_local_scores=0.23557708, pem=-0.6270854,
similarity=14.529991]

```

Sample: b'Empower_MediaMarketing'

Empower MediaMarketing is an independently-owned and award-winning headquartered in offering across-the-board **1 media planning** , , digital creative and development , , and **2 analytics** for clients including , **[1 1 Meijer 1]** , **3 Stanley Steemer** , , **[2 2 Godiva Chocolatier 2]** , Bush Brothers & Co. , Michaels Arts & Crafts Stores . Ms. Mary Beth Price , former and buyer for **[3 3 Leo Burnett 3]** , and , founded Media That Works -- now Empower MediaMarketing -- in **[1 Cincinnati 1]** in 1985 . Ms. Price stepped down as in 1997 but remains on Empower 's board of directors along with her husband , Bill Price . After leaving Empower , Ms. Price was the Richard A. Forsythe Chair of Entrepreneurship at **[4 4 Miami University 4]** in **[2 4 Oxford 2]** , **[3 Ohio 3]** and earned multiple awards : McCall 's / Ad Week Best Woman in Advertising Award , Cincinnati Business Person of the Year , Forbes/Advertising Age Media Maven , Ad Club of Cincinnati Silver Medalist , Leadership Cincinnati participant and

1: 907506 Meijer

2: 1437812 Godiva_Chocolatier

3: 609192 Leo_Burnett

4: 295150 Miami_University

1: 16234440 Media_planning

```

[winner: 16234440 Media_planning, pem_pos=0,
final_scores=-1.146188, final_local_scores=-1.1598053,
pem=0.0, similarity=-30.521193]

```

```

gt: 16234440 Media_planning, pem_pos=0,
final_scores=-1.146188, final_local_scores=-1.1598053,
pem=0.0, similarity=-30.521193 ]

```

2: 487132 Analytics

```

[winner: 48398 Philosophical_analysis, pem_pos=0,
final_scores=-1.1127517, final_local_scores=-1.0673302,
pem=0.0, similarity=-27.76155]

```

```

gt: 487132 Analytics, pem_pos=1,
final_scores=-1.2779764, final_local_scores=-1.3179781,
pem=-0.22309503, similarity=-34.028156 ]

```

3: 12502479 Stanley_Steemer

```

[winner: 12502479 Stanley_Steemer, pem_pos=0,
final_scores=-0.12684725, final_local_scores=-0.02814585,

```

```

pem=0.0, similarity=3.2497847|
  gt: 12502479 Stanley_Steemer, pem_pos=0,
final_scores=-0.12684725, final_local_scores=-0.02814585,
pem=0.0, similarity=3.2497847 ]
4: 128966 Oxford,_Ohio
  [winner: 22199 Ohio, pem_pos=0, final_scores=0.520482,
final_local_scores=0.554121, pem=-0.19147073,
similarity=21.667036]
  RECALL MISS: 128966 Oxford,_Ohio]
  [winner: 22308 Oxford, pem_pos=0,
final_scores=0.047761306, final_local_scores=0.0674573,
pem=-0.68225896, similarity=9.813004]
  gt: 128966 Oxford,_Ohio, pem_pos=4,
final_scores=-0.4856643, final_local_scores=-0.5411054,
pem=-3.856499, similarity=8.91429 ]
  [winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.621228, final_local_scores=-1.5959024,
pem=-4.6728287, similarity=-18.12363]
  RECALL MISS: 128966 Oxford,_Ohio]
1: 907506 Meijer
  [pem_pos=0, final_scores=0.17305295,
final_local_scores=0.19383052, pem=-0.07487176,
similarity=10.281167]
2: 1437812 Godiva_Chocolatier
  [pem_pos=0, final_scores=0.12523781,
final_local_scores=0.18710893, pem=0.0, similarity=9.673417]
3: 609192 Leo_Burnett
  [pem_pos=0, final_scores=0.1259676,
final_local_scores=0.19181716, pem=0.0, similarity=9.81392]
4: 295150 Miami_University
  [pem_pos=0, final_scores=0.16831104,
final_local_scores=0.19302896, pem=0.0, similarity=9.850082]
1: 18522615 Cincinnati
  [pem_pos=0, final_scores=0.5151891,
final_local_scores=0.5195917, pem=-0.46106797,
similarity=22.102724]
2: 22308 Oxford
  [pem_pos=0, final_scores=0.047761306,
final_local_scores=0.0674573, pem=-0.68225896,
similarity=9.813004]
3: 22199 Ohio
  [pem_pos=0, final_scores=0.520482,
final_local_scores=0.554121, pem=-0.19147073,
similarity=21.667036]

```

Sample: b'Mike_Smith'

```

_ -LRB- jazz_saxophonist -RRB- Michael `` Mike '' Smith -
LRB- born March 3 , 1957 -RRB- is an [1 1 American 1] 1 jazz
saxophonist who has released albums as leader on [2 2 Delmark
Records 2] , and his own labels Fastrax and Underground
Labs . Smith graduated in 1980 from the [3 3 University of No
rth Texas 3] and played in the famed 2 One O'Clock Lab Band .

```

He studied with notable saxophone professor [4 4 Jim Riggs 4] . That same year , he moved to [5 5 Chicago 5] . He toured with the [6 6 Maynard Ferguson 6] and [7 7 Buddy Rich 7] bands in the early 80s . While he was with [8 8 Buddy Rich 8] , he toured with [9 9 Frank Sinatra 9] . Since 1981 , he has worked for both [10 10 Frank Sinatra 10] and as featured saxophonist and music contractor . In 1985 , he won a [11 11 National Endowment for the Arts 11] award . In 1993 , he won the 3 Cannonball Adderley Award for 4 alto saxophone . He has recorded and performed with [12 12 Tony Bennett 12] , [1 5 Harry Connick 1] , Jr. . , [13 13 Nancy Wilson 13] , [14 14 Art Farmer 14] and [15 15 Nat Adderley 15] . Mike has had a weekly spot at Andy 's [2 Jazz Club 2] in [3 Chicago 3] for over thirty years . Currently , he is professor of 6 saxophone at [16 16 Roosevelt University 16] , [17 17 Chicago College of Performing Arts 17] . In addition , he is a clinician and performing artist for

- 1: 3434750 United_States
- 2: 31077527 Delmark_Records
- 3: 179976 University_of_North_Texas
- 4: 2667572 Jim_Riggs
- 5: 6886 Chicago
- 6: 165604 Maynard_Ferguson
- 7: 416796 Buddy_Rich
- 8: 416796 Buddy_Rich
- 9: 11181 Frank_Sinatra
- 10: 11181 Frank_Sinatra
- 11: 282752 National_Endowment_for_the_Arts
- 12: 232669 Tony_Bennett
- 13: 618543 Nancy_Wilson_(jazz_singer)
- 14: 310414 Art_Farmer
- 15: 782487 Nat_Adderley
- 16: 2084233 Roosevelt_University
- 17: 5364726 Chicago_College_of_Performing_Arts

1: 15613 Jazz

[winner: 15613 Jazz, pem_pos=0, final_scores=-1.2158183, final_local_scores=-1.3904669, pem=-0.030054808, similarity=-37.241154]

gt: 15613 Jazz, pem_pos=0, final_scores=-1.2158183, final_local_scores=-1.3904669, pem=-0.030054808, similarity=-37.241154]

2: 3807061 One_O'Clock_Lab_Band

[winner: 3807061 One_O'Clock_Lab_Band, pem_pos=0, final_scores=-0.021815985, final_local_scores=-0.038447678, pem=0.0, similarity=2.9423575]

gt: 3807061 One_O'Clock_Lab_Band, pem_pos=0, final_scores=-0.021815985, final_local_scores=-0.038447678, pem=0.0, similarity=2.9423575]

3: 172119 Cannonball_Adderley

[winner: 172119 Cannonball_Adderley, pem_pos=0, final_scores=-0.69739926, final_local_scores=-0.8198288, pem=0.0, similarity=-20.375614]

gt: 172119 Cannonball_Adderley, pem_pos=0,
 final_scores=-0.69739926, final_local_scores=-0.8198288,
 pem=0.0, similarity=-20.375614]
 4: 556372 Alto_saxophone
 [winner: 6370131 For_Alto, pem_pos=0,
 final_scores=-0.8503521, final_local_scores=-0.8111263,
 pem=0.0, similarity=-20.115913]
 RECALL MISS: 556372 Alto_saxophone]
 [winner: 556372 Alto_saxophone, pem_pos=0,
 final_scores=-1.1109167, final_local_scores=-1.254216,
 pem=0.0, similarity=-33.338593]
 gt: 556372 Alto_saxophone, pem_pos=0,
 final_scores=-1.1109167, final_local_scores=-1.254216,
 pem=0.0, similarity=-33.338593]
 5: 13743 Harry_Connick,_Jr.
 [winner: 13743 Harry_Connick,_Jr., pem_pos=1,
 final_scores=0.12225194, final_local_scores=0.06960909,
 pem=-1.2395113, similarity=12.907642]
 gt: 13743 Harry_Connick,_Jr., pem_pos=1,
 final_scores=0.12225194, final_local_scores=0.06960909,
 pem=-1.2395113, similarity=12.907642]
 [winner: 16861 Kurt_Vonnegut, pem_pos=5,
 final_scores=-0.3659836, final_local_scores=-0.39197522,
 pem=-2.3999593, similarity=5.4437428]
 RECALL MISS: 13743 Harry_Connick,_Jr.]
 [winner: 4721653 Zyphe_language, pem_pos=12,
 final_scores=-1.6934079, final_local_scores=-1.6345577,
 pem=-4.4188404, similarity=-20.658411]
 RECALL MISS: 13743 Harry_Connick,_Jr.]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-2.051708, final_local_scores=-2.0531921,
 pem=-4.6728287, similarity=-31.770067]
 RECALL MISS: 13743 Harry_Connick,_Jr.]
 6: 26795 Saxophone
 [winner: 556372 Alto_saxophone, pem_pos=0,
 final_scores=-1.007278, final_local_scores=-1.1410065,
 pem=0.0, similarity=-29.960196]
 gt: 26795 Saxophone, pem_pos=1, final_scores=-1.1304072,
 final_local_scores=-1.2896855, pem=-0.0026145754,
 similarity=-34.382854]
 1: 3434750 United_States
 [pem_pos=0, final_scores=0.5380287,
 final_local_scores=0.59435564, pem=-0.12994486,
 similarity=22.533133]
 2: 31077527 Delmark_Records
 [pem_pos=0, final_scores=0.09580292,
 final_local_scores=0.07476556, pem=0.0, similarity=6.3208666]
 3: 179976 University_of_North_Texas
 [pem_pos=0, final_scores=0.3960626,
 final_local_scores=0.423279, pem=0.0, similarity=16.721203]
 4: 2667572 Jim_Riggs
 [pem_pos=0, final_scores=0.038373515,

final_local_scores=0.043181404, pem=0.0, similarity=5.378332]
5: 6886 Chicago
[pem_pos=0, final_scores=0.6291488,
final_local_scores=0.6442283, pem=-0.20658566,
similarity=24.438213]
6: 165604 Maynard_Ferguson
[pem_pos=0, final_scores=0.15262793,
final_local_scores=0.10796644, pem=0.0, similarity=7.311647]
7: 416796 Buddy_Rich
[pem_pos=0, final_scores=0.09414606,
final_local_scores=0.0400033, pem=0.0, similarity=5.283491]
8: 416796 Buddy_Rich
[pem_pos=0, final_scores=0.16591488,
final_local_scores=0.11839986, pem=0.0, similarity=7.6230006]
9: 11181 Frank_Sinatra
[pem_pos=0, final_scores=0.46507972,
final_local_scores=0.43348145, pem=0.0, similarity=17.025663]
10: 11181 Frank_Sinatra
[pem_pos=0, final_scores=0.44292414,
final_local_scores=0.4070058, pem=0.0, similarity=16.235577]
11: 282752 National_Endowment_for_the_Arts
[pem_pos=0, final_scores=0.09796731,
final_local_scores=0.0955843, pem=0.0, similarity=6.942139]
12: 232669 Tony_Bennett
[pem_pos=0, final_scores=0.3217346,
final_local_scores=0.27814555, pem=0.0, similarity=12.390131]
13: 618543 Nancy_Wilson_(jazz_singer)
[pem_pos=0, final_scores=0.15010737,
final_local_scores=0.10393348, pem=-0.2875037,
similarity=8.754786]
14: 310414 Art_Farmer
[pem_pos=0, final_scores=0.12235828,
final_local_scores=0.087331966, pem=0.0, similarity=6.695873]
15: 782487 Nat_Adderley
[pem_pos=0, final_scores=0.11277036,
final_local_scores=0.08832511, pem=0.0, similarity=6.7255106]
16: 2084233 Roosevelt_University
[pem_pos=0, final_scores=0.30142134,
final_local_scores=0.31003305, pem=0.0, similarity=13.341718]
17: 5364726 Chicago_College_of_Performing_Arts
[pem_pos=0, final_scores=0.27011508,
final_local_scores=0.321915, pem=0.0, similarity=13.696299]
1: 13743 Harry_Connick,_Jr.
[pem_pos=1, final_scores=0.12225194,
final_local_scores=0.06960909, pem=-1.2395113,
similarity=12.907642]
2: 28698868 The_Fast_Show
[pem_pos=0, final_scores=0.1439635,
final_local_scores=0.17535031, pem=0.0, similarity=9.3225155]
3: 6886 Chicago
[pem_pos=0, final_scores=0.6150332,
final_local_scores=0.62880915, pem=-0.20658566,

similarity=23.978077]

Sample: b'Putin'

's _ rynda [1 Putin 1] 's rynda -LRB- -RRB- is an 1 internet meme which refers to a 2 LiveJournal user post , picked up by the radio station [1 1 Echo of Moscow 1] , and forwarded to the 3 Prime Minister of [2 Russia 2] [2 2 Vladimir Putin 2] . The post , written on August 1 , 2010 with obscene words by the [3 Russian 3] user top-lap , who introduced himself as Alexander , criticised the 4 fire safety condition in his native village during the 5 2010 [4 Russian 4] wildfires and emphasized the disappearance of the village 's old fire alarm bell , which top-lap called `` rynda '' -LRB- a 6 ship 's bell in [5 Russian 5] -RRB- . Top-lap asked to support his appeal , directed to the authorities of [3 3 Kalyazinsky District 3] in [6 Russia 6] 's [4 4 Tver Oblast 4] . The post was spotted by the [7 Echo of Moscow 7] 's editor-in-chief [5 5 Alexei Venediktov 5] on August 1 . According to [8 Venediktov 8] , he `` saw that this message contains not only a money request or criticism of authorities , but a proposal to organize a voluntary fire guard if discounts are offered '' and `` it appeared as something new '' to him . Venediktov forwarded the post with his accompanying commentary on August 2 . [9 Putin 9] 's reply was e-mailed two days later . The rynda was ultimately installed , although top-lap claimed he was visited by the authorities who took his 7 USB flash drive s and , that reportedly stored nothing but video games . In August , 2010 the question `` what is rynda '' became the top query on [6 6 Yandex 6] search engine , briefly surpassing the query `` love '' . [10 Putin 10] 's rynda also became a subject of s and is mentioned on the November page of the erotic calendar 8 Happy Birthday , [11 Mr. Putin 11] ! During the 9 2010 [12 Russian 12] wildfires local officials were accused of taking inadequate measures to prevent fires and respond effectively . A 10 state of emergency was declared near about 500 towns and villages . The then [13 Russian 13] Minister of Emergency Situations [7 7 Sergey Shoygu 7] outlined a lack of volunteer firefighting units and poor enforcement of safety rules at 11 dacha

1: 1867521 Echo_of_Moscow

2: 32817 Vladimir_Putin

3: 32658499 Kalyazinsky_District

4: 463952 Tver_Oblast

5: 24752858 Alexei_Venediktov

6: 3099107 Yandex

7: 857985 Sergey_Shoygu

1: 1616492 Internet_meme

[winner: 1616492 Internet_meme, pem_pos=0,
final_scores=-1.130007, final_local_scores=-1.165729,
pem=-0.32970145, similarity=-28.905]

gt: 1616492 Internet_meme, pem_pos=0,
final_scores=-1.130007, final_local_scores=-1.165729,

pem=-0.32970145, similarity=-28.905]
2: 92407 LiveJournal
[winner: 92407 LiveJournal, pem_pos=0,
final_scores=-0.00667952, final_local_scores=0.03770542,
pem=0.0, similarity=5.214918]
gt: 92407 LiveJournal, pem_pos=0,
final_scores=-0.00667952, final_local_scores=0.03770542,
pem=0.0, similarity=5.214918]
3: 339344 Prime_Minister_of_Russia
[winner: 25391 Russia, pem_pos=0,
final_scores=0.58255404, final_local_scores=0.56252164,
pem=-0.20299491, similarity=21.980398]
RECALL MISS: 339344 Prime_Minister_of_Russia]
[winner: 1647874 Prime_Minister_of_Kyrgyzstan, pem_pos=0,
final_scores=-0.70341796, final_local_scores=-0.70440406,
pem=0.0, similarity=-16.931108]
RECALL MISS: 339344 Prime_Minister_of_Russia]
4: 987507 Fire_safety
[winner: 5186748 The_Fire_(Seinfeld), pem_pos=1,
final_scores=-1.1239259, final_local_scores=-1.0505545,
pem=-0.5175043, similarity=-24.44666]
RECALL MISS: 987507 Fire_safety]
[winner: 30857927 Safety_(firearms), pem_pos=3,
final_scores=-1.367966, final_local_scores=-1.3461206,
pem=-1.2871779, similarity=-29.081335]
RECALL MISS: 987507 Fire_safety]
5: 28224526 2010_Russian_wildfires
[winner: 25391 Russia, pem_pos=1, final_scores=0.7917126,
final_local_scores=0.7909956, pem=-1.0515022,
similarity=33.41283]
RECALL MISS: 28224526 2010_Russian_wildfires]
[winner: 43226 2010, pem_pos=0, final_scores=-1.0274211,
final_local_scores=-0.9537314, pem=-1.6731367,
similarity=-15.272758]
RECALL MISS: 28224526 2010_Russian_wildfires]
[winner: 24085206 2009_Greek_forest_fires, pem_pos=4,
final_scores=-1.1839259, final_local_scores=-1.1879928,
pem=-1.3862944, similarity=-23.823475]
gt: 28224526 2010_Russian_wildfires, pem_pos=2,
final_scores=-1.2718222, final_local_scores=-1.3724649,
pem=-1.3333358, similarity=-29.616488]
6: 274908 Ship's_bell
[winner: 518342 SS_Atlantic_Conveyor, pem_pos=0,
final_scores=-1.055452, final_local_scores=-1.0067341,
pem=-0.36524028, similarity=-23.967005]
RECALL MISS: 274908 Ship's_bell]
[winner: 138707 Bell,_Wisconsin, pem_pos=12,
final_scores=-1.5760887, final_local_scores=-1.5485816,
pem=-4.0786037, similarity=-19.942972]
RECALL MISS: 274908 Ship's_bell]
[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-1.9856281, final_local_scores=-2.0568864,

pem=-3.4067376, similarity=-38.76551|
RECALL MISS: 274908 Ship's_bell]

7: 400414 USB_flash_drive
[winner: 32073 USB, pem_pos=0, final_scores=-0.09251954,
final_local_scores=0.017741874, pem=0.0, similarity=4.619166|
gt: 400414 USB_flash_drive, pem_pos=3,
final_scores=-0.76220065, final_local_scores=-0.7342244,
pem=-4.226956, similarity=5.165837]
[winner: 25008295 Drive_(Béla_Fleck_album), pem_pos=6,
final_scores=-1.3099668, final_local_scores=-1.2109783,
pem=-2.9816263, similarity=-15.833745|
RECALL MISS: 400414 USB_flash_drive]
[winner: 20947 Adobe_Flash, pem_pos=0,
final_scores=-1.6777942, final_local_scores=-1.750742,
pem=-0.4169662, similarity=-45.88839|
RECALL MISS: 400414 USB_flash_drive]

8: 33900431 Happy_Birthday,_Mr._Putin!
[winner: 32817 Vladimir_Putin, pem_pos=0,
final_scores=0.3595764, final_local_scores=0.32622308,
pem=0.0, similarity=13.824861|
RECALL MISS: 33900431 Happy_Birthday,_Mr._Putin!]
[winner: 87328 Happy_Birthday_to_You, pem_pos=0,
final_scores=-0.2187908, final_local_scores=-0.12780109,
pem=-0.66365856, similarity=3.8849518|
RECALL MISS: 33900431 Happy_Birthday,_Mr._Putin!]
[winner: 3243047 !_(album), pem_pos=3,
final_scores=-1.6472218, final_local_scores=-1.6190445,
pem=-2.4849067, similarity=-30.712498|
RECALL MISS: 33900431 Happy_Birthday,_Mr._Putin!]
[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.7458451, final_local_scores=-1.6915674,
pem=-4.6728287, similarity=-20.97846|
RECALL MISS: 33900431 Happy_Birthday,_Mr._Putin!]

9: 28224526 2010_Russian_wildfires
[winner: 25391 Russia, pem_pos=1, final_scores=0.8672429,
final_local_scores=0.873501, pem=-1.0515022,
similarity=35.874958|
RECALL MISS: 28224526 2010_Russian_wildfires]
[winner: 43226 2010, pem_pos=0, final_scores=-1.0218911,
final_local_scores=-0.9476908, pem=-1.6731367,
similarity=-15.092493|
RECALL MISS: 28224526 2010_Russian_wildfires]
[winner: 24085206 2009_Greek_forest_fires, pem_pos=4,
final_scores=-1.3163304, final_local_scores=-1.3326247,
pem=-1.3862944, similarity=-28.139582|
gt: 28224526 2010_Russian_wildfires, pem_pos=2,
final_scores=-1.4744898, final_local_scores=-1.5938485,
pem=-1.3333358, similarity=-36.223015]

10: 298608 State_of_emergency
[winner: 229379 States_of_Brazil, pem_pos=0,
final_scores=-1.135222, final_local_scores=-1.1047698,
pem=0.0, similarity=-28.878822|

RECALL MISS: 298608 State_of_emergency]
[winner: 146334 Emergency, pem_pos=0,
final_scores=-1.1514193, final_local_scores=-1.1615314,
pem=-1.1955357, similarity=-24.071196]
gt: 298608 State_of_emergency, pem_pos=10,
final_scores=-1.3857288, final_local_scores=-1.451478,
pem=-2.8992596, similarity=-23.458662]
11: 350046 Dacha
[winner: 350046 Dacha, pem_pos=0,
final_scores=-0.8650535, final_local_scores=-0.9561013,
pem=-0.0005412971, similarity=-24.439314]
gt: 350046 Dacha, pem_pos=0, final_scores=-0.8650535,
final_local_scores=-0.9561013, pem=-0.0005412971,
similarity=-24.439314]
1: 1867521 Echo_of_Moscow
[pem_pos=0, final_scores=0.6330697,
final_local_scores=0.685928, pem=0.0, similarity=24.559172]
2: 32817 Vladimir_Putin
[pem_pos=0, final_scores=0.44012183,
final_local_scores=0.41420674, pem=0.0, similarity=16.450468]
3: 32658499 Kalyazinsky_District
[pem_pos=0, final_scores=0.04599209,
final_local_scores=0.14170223, pem=0.0, similarity=8.31839]
4: 463952 Tver_Oblast
[pem_pos=0, final_scores=0.3392002,
final_local_scores=0.3628496, pem=0.0, similarity=14.917869]
5: 24752858 Alexei_Venediktov
[pem_pos=0, final_scores=0.17611398,
final_local_scores=0.23791891, pem=0.0, similarity=11.189688]
6: 3099107 Yandex
[pem_pos=0, final_scores=0.15808187,
final_local_scores=0.21763375, pem=0.0, similarity=10.584338]
7: 857985 Sergey_Shoygu
[pem_pos=0, final_scores=0.093780145,
final_local_scores=0.11239277, pem=0.0, similarity=7.4437375]
1: 32817 Vladimir_Putin
[pem_pos=0, final_scores=0.34634787,
final_local_scores=0.31177297, pem=0.0, similarity=13.3936405]
2: 25391 Russia
[pem_pos=0, final_scores=0.58255404,
final_local_scores=0.56252164, pem=-0.20299491,
similarity=21.980398]
3: 25391 Russia
[pem_pos=1, final_scores=0.7543877,
final_local_scores=0.75022376, pem=-1.0515022,
similarity=32.196117]
4: 25391 Russia
[pem_pos=1, final_scores=0.7917126,
final_local_scores=0.7909956, pem=-1.0515022,
similarity=33.41283]
5: 25391 Russia
[pem_pos=1, final_scores=0.4795733,

final_local_scores=0.45003086, pem=-1.0515022,
 similarity=23.237764]
 6: 25391 Russia
 [pem_pos=0, final_scores=0.83512145,
 final_local_scores=0.8384132, pem=-0.20299491,
 similarity=30.21355]
 7: 1867521 Echo_of_Moscow
 [pem_pos=0, final_scores=0.5551582,
 final_local_scores=0.6008215, pem=0.0, similarity=22.019424]
 8: 24752858 Alexei_Venediktov
 [pem_pos=0, final_scores=0.03539966,
 final_local_scores=0.084209904, pem=0.0, similarity=6.6027045]
 9: 32817 Vladimir_Putin
 [pem_pos=0, final_scores=0.43518955,
 final_local_scores=0.40881896, pem=0.0, similarity=16.289686]
 10: 32817 Vladimir_Putin
 [pem_pos=0, final_scores=0.44847035,
 final_local_scores=0.42332625, pem=0.0, similarity=16.722612]
 11: 32817 Vladimir_Putin
 [pem_pos=0, final_scores=0.3595764,
 final_local_scores=0.32622308, pem=0.0, similarity=13.824861]
 12: 25391 Russia
 [pem_pos=1, final_scores=0.8672429,
 final_local_scores=0.873501, pem=-1.0515022,
 similarity=35.874958]
 13: 25391 Russia
 [pem_pos=1, final_scores=0.7034531,
 final_local_scores=0.69458556, pem=-1.0515022,
 similarity=30.535765]

Sample: b'Richard_of_Salerno'

Richard of [1 Salerno 1] -LRB- c. 1060 -- 1114 -RRB- , who
 is not to be confused with his homonym cousin 1 Richard of Ha
 uteville , was a participant in the [1 1 First Crusade 1] and
 regent of the 2 County of [2 Edessa 2] from 1104 to 1108 .
 Richard was born around 1060 , the third son of 3 William of t
 he Principate , a Norman count , and Mary , daughter of
 Guido , the [3 4 Lombard 3] duke of [4 Sorrento 4] . He was
 also one of the many nephews of [2 2 Robert Guiscard 2] and [5
 5 Roger 5] I of [6 Sicily 6] and in his early life [7 Richard
 7] participated with his two famous uncles to the conquest of
 [8 Sicily 8] . In 1097 , [9 Richard 9] joined his cousins , [1
 0 6 Bohemund 10] of [11 Taranto 11] and [3 3 Tancred 3] , on
 the [4 4 First Crusade 4] . Richard and Tancred were notable
 for being among the few Crusaders who could speak [5 5 Arabic
 5] , an ability doubtlessly learned during the wars in [12
 Sicily 12] , which had a strong [13 Arab 13] presence . [6 6 A
 nna Komnene 6] relates that when [14 Richard 14] crossed the [
 7 7 Adriatic Sea 7] , his ship was attacked and captured by
 the [15 7 Byzantine 15] fleet , who had mistaken him for a
 pirate . He was soon released and joined the main [16
 Crusader 16] army marching through [8 8 Bulgaria 8] and [9 9 H

ungary 9] . Along with Tancred , Richard refused to swear an oath of fidelity to the [17 Byzantine 17] emperor [10 10 Alex ius I Comnenus 10] , preferring to cross the [18 Bosphorus 18] in secret . [19 Richard 19] was one of the commanders at the 8 Battle of Dorylaeum in summer 1097 . 1098 , [20 Richard 20] and Tancred joined [21 Bohemund 21] at the 9 siege of [22 Ant ioch 22] . Richard was among those captured with [23 Bohemund 23] in 1100 when ambushed by the 10 Danishmends at the 11 Bat tle of [24 Melitene 24] . From there Richard was sent to the emperor [25 Alexius 25] , who imprisoned him in [11 11 Consta ntinople 11] before he was finally released in 1103 . His cousin [26 Tancred 26] then appointed him governor of [12 12 E dessa 12] in the winter of 1104 , a city which he ruled until 1108 . He was bitterly hated by the citizens of [27 Edessa 27] for being ruthless and greedy . During this time , [28 Richard 28] also acted as a diplomat , traveling to [29 France 29] and [30 Italy 30] and arranging the marriage of [31 Bohemund 31] to the princess Constance of [32 France 32] . He was a witness of the 1108 12 Treaty of [33 Devol 33] . He also participated in the ultimately disastrous campaigns [34 Bohemund 34] waged against the emperor [35 Alexius 35] , but may have been secretly plotting with [36 Alexius 36] against [37 Bohemund 37] . After [38 Bohemund 38] 's death in 1111 , [3 9 Richard 39] retired to [13 13 Marash 13] , where he died in the great earthquake of November 29 , 1114 . By his wife Altrude , daughter of 13 Emma of Hauteville and 14 Odo the Go od Marquis , he had a son , 15 Roger of [40 Salerno 40] , who was the regent of the 16 Principality of [41 Antioch 41] . His daughter Maria married [42 17 Joscelin I 42] , Count of [4 3 Edessa 43]

- 1: 106128 First_Crusade
- 2: 157902 Robert_Guiscard
- 3: 1127415 Tancred,_Prince_of_Galilee
- 4: 106128 First_Crusade
- 5: 803 Arabic_language
- 6: 157653 Anna_Komnene
- 7: 23275478 Adriatic_Sea
- 8: 3415 Bulgaria
- 9: 13275 Hungary
- 10: 1613 Alexios_I_Komnenos
- 11: 5646 Constantinople
- 12: 285454 Edessa
- 13: 2731392 Kahramanmaraş

1: 25783730 Richard_of_Hauteville
 [winner: 18097500 Emma_of_Hauteville, pem_pos=0,
 final_scores=-0.09307033, final_local_scores=-0.03173256,
 pem=0.0, similarity=3.14275]

RECALL MISS: 25783730 Richard_of_Hauteville]
 [winner: 2844614 Richard_of_Salerno, pem_pos=0,
 final_scores=-0.4806223, final_local_scores=-0.5428388,
 pem=0.0, similarity=-12.109679]

RECALL MISS: 25783730 Richard_of_Hauteville]

2: 255790 County_of_Edessa
[winner: 285454 Edessa, pem_pos=1,
final_scores=0.15709843, final_local_scores=0.11109565,
pem=-0.12033821, similarity=8.059446]
gt: 255790 County_of_Edessa, pem_pos=3,
final_scores=-0.19911034, final_local_scores=-0.26622763,
pem=-1.4260596, similarity=3.9000833]
[winner: 382207 Powiat, pem_pos=0,
final_scores=-0.5810241, final_local_scores=-0.552491,
pem=-0.39345494, similarity=-10.258051]
RECALL MISS: 255790 County_of_Edessa]

3: 4333752 William_of_the_Principate
[winner: 4333752 William_of_the_Principate, pem_pos=0,
final_scores=-0.012991682, final_local_scores=0.024643883,
pem=0.0, similarity=4.8251357]
gt: 4333752 William_of_the_Principate, pem_pos=0,
final_scores=-0.012991682, final_local_scores=0.024643883,
pem=0.0, similarity=4.8251357]

4: 18011 Lombards
[winner: 43807 Lombardy, pem_pos=3,
final_scores=0.07260524, final_local_scores=0.056584716,
pem=-2.0284312, similarity=16.809238]
gt: 18011 Lombards, pem_pos=0, final_scores=0.046933964,
final_local_scores=-0.014993653, pem=-0.86339796,
similarity=8.337564]

5: 157904 Roger_I_of_Sicily
[winner: 27619 Sicily, pem_pos=0, final_scores=0.3833174,
final_local_scores=0.36301374, pem=-0.08753506,
similarity=15.398796]
RECALL MISS: 157904 Roger_I_of_Sicily]
[winner: 1210109 Roger_of_Salerno, pem_pos=0,
final_scores=0.15792732, final_local_scores=0.19973776,
pem=0.0, similarity=10.050286]
gt: 157904 Roger_I_of_Sicily, pem_pos=18,
final_scores=-0.5776218, final_local_scores=-0.6843835,
pem=-4.0336113, similarity=5.6017494]

6: 157674 Bohemond_I_of_Antioch
[winner: 56148 Taranto, pem_pos=0,
final_scores=0.34257412, final_local_scores=0.3399177,
pem=-0.28114104, similarity=15.762423]
RECALL MISS: 157674 Bohemond_I_of_Antioch]
[winner: 157674 Bohemond_I_of_Antioch, pem_pos=0,
final_scores=0.14181606, final_local_scores=0.08390285,
pem=0.0, similarity=6.593541]
gt: 157674 Bohemond_I_of_Antioch, pem_pos=0,
final_scores=0.14181606, final_local_scores=0.08390285,
pem=0.0, similarity=6.593541]

7: 2134878 Byzantine_navy
[winner: 16972981 Byzantine_Empire, pem_pos=1,
final_scores=0.21854849, final_local_scores=0.15486327,
pem=-0.3418526, similarity=10.570189]
RECALL MISS: 2134878 Byzantine_navy]

[winner: 16705512 Fleet_railway_station_(Lincolnshire),
pem_pos=4, final_scores=-1.162365,
final_local_scores=-1.0515392, pem=-2.7385583,
similarity=-12.397606|
RECALL MISS: 2134878 Byzantine_navy]

8: 398093 Battle_of_Dorylaeum_(1097)
[winner: 308704 Dorylaeum, pem_pos=0,
final_scores=0.021406606, final_local_scores=0.045757815,
pem=0.0, similarity=5.4552174|
gt: 398093 Battle_of_Dorylaeum_(1097), pem_pos=1,
final_scores=-0.5984093, final_local_scores=-0.6475705,
pem=-3.4318345, similarity=3.4277675]
[winner: 4181 Battle, pem_pos=0, final_scores=-0.8237315,
final_local_scores=-0.89478153, pem=-1.4437126,
similarity=-14.761218|
RECALL MISS: 398093 Battle_of_Dorylaeum_(1097)]

9: 1119248 Siege_of_Antioch
[winner: 36900 Antioch, pem_pos=0,
final_scores=0.28980905, final_local_scores=0.23782897,
pem=-0.36888143, similarity=13.193039|
gt: 1119248 Siege_of_Antioch, pem_pos=4,
final_scores=-0.26203775, final_local_scores=-0.3068637,
pem=-2.8799932, similarity=10.594139]
[winner: 17095521 The_Castle_(Kadare), pem_pos=1,
final_scores=-0.8303464, final_local_scores=-0.7472531,
pem=-0.39713177, similarity=-16.050146|
RECALL MISS: 1119248 Siege_of_Antioch]

10: 1308260 Danishmends
[winner: 1308260 Danishmends, pem_pos=0,
final_scores=-0.046856776, final_local_scores=-0.032471813,
pem=0.0, similarity=3.1206892|
gt: 1308260 Danishmends, pem_pos=0,
final_scores=-0.046856776, final_local_scores=-0.032471813,
pem=0.0, similarity=3.1206892]

11: 16230308 Battle_of_Melitene
[winner: 1291890 Malatya, pem_pos=0,
final_scores=0.27743345, final_local_scores=0.27520996,
pem=0.0, similarity=12.302527|
gt: 16230308 Battle_of_Melitene, pem_pos=1,
final_scores=-0.36242366, final_local_scores=-0.3549817,
pem=-3.0072045, similarity=9.849997]
[winner: 16649099 Mount_Battle, pem_pos=5,
final_scores=-0.97796625, final_local_scores=-0.88595897,
pem=-3.124515, similarity=-5.3574715|
RECALL MISS: 16230308 Battle_of_Melitene]

12: 750464 Treaty_of_Devol
[winner: 20371105 Devol_(Albania), pem_pos=1,
final_scores=0.15863071, final_local_scores=0.1914832,
pem=-0.7641447, similarity=13.959492|
gt: 750464 Treaty_of_Devol, pem_pos=3,
final_scores=-0.03138207, final_local_scores=-0.07134387,
pem=-1.7917595, similarity=11.704536]

[winner: 149028 Tribal_Voice, pem_pos=5,
final_scores=-0.9509726, final_local_scores=-0.820303,
pem=-2.0368252, similarity=-9.313193]
RECALL MISS: 750464 Treaty_of_Devol]

13: 18097500 Emma_of_Hauteville
[winner: 18097500 Emma_of_Hauteville, pem_pos=0,
final_scores=-0.094808474, final_local_scores=-0.033749692,
pem=0.0, similarity=3.0825548]
gt: 18097500 Emma_of_Hauteville, pem_pos=0,
final_scores=-0.094808474, final_local_scores=-0.033749692,
pem=0.0, similarity=3.0825548]

14: 33710749 Odo_the_Good_Marquis
[winner: 33710749 Odo_the_Good_Marquis, pem_pos=0,
final_scores=-0.15284568, final_local_scores=-0.089703366,
pem=0.0, similarity=1.4127856]
gt: 33710749 Odo_the_Good_Marquis, pem_pos=0,
final_scores=-0.15284568, final_local_scores=-0.089703366,
pem=0.0, similarity=1.4127856]

15: 1210109 Roger_of_Salerno
[winner: 62415 Salerno, pem_pos=1,
final_scores=0.38111007, final_local_scores=0.39291376,
pem=-0.15453224, similarity=16.655415]
RECALL MISS: 1210109 Roger_of_Salerno]
[winner: 157904 Roger_I_of_Sicily, pem_pos=0,
final_scores=-0.27745646, final_local_scores=-0.35669875,
pem=0.0, similarity=-6.5548897]
RECALL MISS: 1210109 Roger_of_Salerno]

16: 256387 Principality_of_Antioch
[winner: 36900 Antioch, pem_pos=0,
final_scores=0.15758948, final_local_scores=0.09339921,
pem=-0.36888143, similarity=8.882966]
gt: 256387 Principality_of_Antioch, pem_pos=5,
final_scores=-0.35152918, final_local_scores=-0.44614393,
pem=-2.961368, similarity=6.8802686]
[winner: 77322 Principality, pem_pos=0,
final_scores=-0.82976, final_local_scores=-0.92525226,
pem=-0.14635165, similarity=-22.725777]
gt: 256387 Principality_of_Antioch, pem_pos=8,
final_scores=-1.0353777, final_local_scores=-1.1933699,
pem=-2.1933174, similarity=-19.595219]

17: 296046 Joscelin_I,_Count_of_Edessa
[winner: 296046 Joscelin_I,_Count_of_Edessa, pem_pos=0,
final_scores=0.32568175, final_local_scores=0.33810288,
pem=0.0, similarity=14.179378]
gt: 296046 Joscelin_I,_Count_of_Edessa, pem_pos=0,
final_scores=0.32568175, final_local_scores=0.33810288,
pem=0.0, similarity=14.179378]
[winner: 285454 Edessa, pem_pos=0,
final_scores=0.1949255, final_local_scores=0.15241599,
pem=-0.12033821, similarity=9.292528]
RECALL MISS: 296046 Joscelin_I,_Count_of_Edessa]
[winner: 196470 Count, pem_pos=0,

final_scores=-0.68156266, final_local_scores=-0.7573371,
pem=-0.092834726, similarity=-18.005888]

RECALL MISS: 296046 Joscelin_I,_Count_of_Edessa]

[winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.5312855, final_local_scores=-1.5053303,
pem=-4.6728287, similarity=-15.420774]

RECALL MISS: 296046 Joscelin_I,_Count_of_Edessa]

1: 106128 First_Crusade

[pem_pos=0, final_scores=0.33468336,
final_local_scores=0.29131946, pem=0.0, similarity=12.783267]

2: 157902 Robert_Guiscard

[pem_pos=0, final_scores=0.4505185,
final_local_scores=0.43360132, pem=0.0, similarity=17.029242]

3: 1127415 Tancred,_Prince_of_Galilee

[pem_pos=0, final_scores=0.04701461,
final_local_scores=0.015130103, pem=-1.0901932,
similarity=10.469865]

4: 106128 First_Crusade

[pem_pos=0, final_scores=0.2865355,
final_local_scores=0.23872522, pem=0.0, similarity=11.21375]

5: 803 Arabic_language

[pem_pos=0, final_scores=0.21474694,
final_local_scores=0.25825053, pem=-0.084880665,
similarity=12.2580185]

6: 157653 Anna_Komnene

[pem_pos=0, final_scores=0.12552838,
final_local_scores=0.121533364, pem=0.0, similarity=7.716511]

7: 23275478 Adriatic_Sea

[pem_pos=0, final_scores=0.19277619,
final_local_scores=0.1869238, pem=0.0, similarity=9.667892]

8: 3415 Bulgaria

[pem_pos=0, final_scores=0.56849796,
final_local_scores=0.64756745, pem=-0.31727514,
similarity=25.139809]

9: 13275 Hungary

[pem_pos=0, final_scores=0.47128457,
final_local_scores=0.54593164, pem=-0.32577527,
similarity=22.153019]

10: 1613 Alexios_I_Komnenos

[pem_pos=0, final_scores=0.14590065,
final_local_scores=0.092846066, pem=0.0, similarity=6.8604245]

11: 5646 Constantinople

[pem_pos=0, final_scores=0.35885793,
final_local_scores=0.30185342, pem=-0.077381566,
similarity=13.518434]

12: 285454 Edessa

[pem_pos=1, final_scores=0.37848347,
final_local_scores=0.3533163, pem=-0.12033821,
similarity=15.287795]

13: 2731392 Kahramanmaraş

[pem_pos=0, final_scores=0.30223614,
final_local_scores=0.31724364, pem=0.0, similarity=13.556896]

1: 62415 Salerno
[pem_pos=1, final_scores=0.16743265,
final_local_scores=0.16009116, pem=-0.15453224,
similarity=9.707522]
2: 285454 Edessa
[pem_pos=1, final_scores=0.15709843,
final_local_scores=0.11109565, pem=-0.12033821,
similarity=8.059446]
3: 43807 Lombardy
[pem_pos=3, final_scores=0.07260524,
final_local_scores=0.056584716, pem=-2.0284312,
similarity=16.809238]
4: 172580 Sorrento
[pem_pos=0, final_scores=0.39060277,
final_local_scores=0.4143567, pem=-0.33544636,
similarity=18.279152]
5: 1210109 Roger_of_Salerno
[pem_pos=0, final_scores=0.15792732,
final_local_scores=0.19973776, pem=0.0, similarity=10.050286]
6: 27619 Sicily
[pem_pos=0, final_scores=0.3833174,
final_local_scores=0.36301374, pem=-0.08753506,
similarity=15.398796]
7: 2844614 Richard_of_Salerno
[pem_pos=0, final_scores=0.046300367,
final_local_scores=0.035949126, pem=0.0, similarity=5.1625066]
8: 27619 Sicily
[pem_pos=1, final_scores=0.480766,
final_local_scores=0.46971232, pem=-0.08753506,
similarity=18.582893]
9: 2844614 Richard_of_Salerno
[pem_pos=0, final_scores=0.11638348,
final_local_scores=0.11250427, pem=0.0, similarity=7.447065]
10: 157674 Bohemond_I_of_Antioch
[pem_pos=0, final_scores=0.14181606,
final_local_scores=0.08390285, pem=0.0, similarity=6.593541]
11: 56148 Taranto
[pem_pos=0, final_scores=0.34257412,
final_local_scores=0.3399177, pem=-0.28114104,
similarity=15.762423]
12: 27619 Sicily
[pem_pos=1, final_scores=0.43917352,
final_local_scores=0.42427891, pem=-0.08753506,
similarity=17.22707]
13: 159433 Arab_world
[pem_pos=14, final_scores=0.16866378,
final_local_scores=0.22207046, pem=-3.8052368,
similarity=31.41021]
14: 2844614 Richard_of_Salerno
[pem_pos=0, final_scores=0.047143742,
final_local_scores=0.03687039, pem=0.0, similarity=5.189999]
15: 16972981 Byzantine_Empire

[pem_pos=1, final_scores=0.21854849,
final_local_scores=0.15486327, pem=-0.3418526,
similarity=10.570189]
16: 4412145 Crusades
[pem_pos=0, final_scores=0.032399163,
final_local_scores=-0.04420165, pem=-0.60334986,
similarity=6.051758]
17: 16972981 Byzantine_Empire
[pem_pos=0, final_scores=0.18361486,
final_local_scores=0.11616254, pem=-0.3418526,
similarity=9.415282]
18: 3705 Bosphorus
[pem_pos=0, final_scores=0.42636722,
final_local_scores=0.4454847, pem=0.0, similarity=17.383865]
19: 2844614 Richard_of_Salerno
[pem_pos=0, final_scores=0.055839643,
final_local_scores=0.04636936, pem=0.0, similarity=5.473467]
20: 2844614 Richard_of_Salerno
[pem_pos=0, final_scores=0.06170009,
final_local_scores=0.052770987, pem=0.0, similarity=5.6645045]
21: 157674 Bohemond_I_of_Antioch
[pem_pos=0, final_scores=0.29458672,
final_local_scores=0.25078157, pem=0.0, similarity=11.573535]
22: 36900 Antioch
[pem_pos=0, final_scores=0.28980905,
final_local_scores=0.23782897, pem=-0.36888143,
similarity=13.193039]
23: 157674 Bohemond_I_of_Antioch
[pem_pos=0, final_scores=0.29560512,
final_local_scores=0.251894, pem=0.0, similarity=11.606732]
24: 1291890 Malatya
[pem_pos=0, final_scores=0.27743345,
final_local_scores=0.27520996, pem=0.0, similarity=12.302527]
25: 1613 Alexios_I_Komnenos
[pem_pos=0, final_scores=0.16892003,
final_local_scores=0.11799121, pem=0.0, similarity=7.610806]
26: 1127415 Tancred, Prince_of_Galilee
[pem_pos=0, final_scores=0.082767114,
final_local_scores=0.054777056, pem=-1.0901932,
similarity=11.653009]
27: 285454 Edessa
[pem_pos=1, final_scores=0.32081658,
final_local_scores=0.29032397, pem=-0.12033821,
similarity=13.407977]
28: 2844614 Richard_of_Salerno
[pem_pos=0, final_scores=0.06609659,
final_local_scores=0.057573512, pem=0.0, similarity=5.8078213]
29: 5843419 France
[pem_pos=1, final_scores=0.5493939,
final_local_scores=0.6023771, pem=-0.16240834,
similarity=22.949049]
30: 14532 Italy

[pem_pos=0, final_scores=0.57884914,
 final_local_scores=0.6264734, pem=-0.20612453,
 similarity=23.905867]
 31: 157674 Bohemond_I_of_Antioch
 [pem_pos=0, final_scores=0.26130378,
 final_local_scores=0.214425, pem=0.0, similarity=10.488583]
 32: 5843419 France
 [pem_pos=0, final_scores=0.5396234,
 final_local_scores=0.5873119, pem=-0.16240834,
 similarity=22.499474]
 33: 20371105 Devol_(Albania)
 [pem_pos=1, final_scores=0.15863071,
 final_local_scores=0.1914832, pem=-0.7641447,
 similarity=13.959492]
 34: 157674 Bohemond_I_of_Antioch
 [pem_pos=0, final_scores=0.2638442,
 final_local_scores=0.21719998, pem=0.0, similarity=10.571394]
 35: 1613 Alexios_I_Komnenos
 [pem_pos=0, final_scores=0.1790001,
 final_local_scores=0.12900215, pem=0.0, similarity=7.939395]
 36: 1613 Alexios_I_Komnenos
 [pem_pos=0, final_scores=0.06650625,
 final_local_scores=0.006119743, pem=0.0, similarity=4.2723384]
 37: 157674 Bohemond_I_of_Antioch
 [pem_pos=0, final_scores=0.26638937,
 final_local_scores=0.21998024, pem=0.0, similarity=10.654362]
 38: 157674 Bohemond_I_of_Antioch
 [pem_pos=0, final_scores=0.30944544,
 final_local_scores=0.26701245, pem=0.0, similarity=12.057897]
 39: 2844614 Richard_of_Salerno
 [pem_pos=0, final_scores=0.09240471,
 final_local_scores=0.08631112, pem=0.0, similarity=6.6654086]
 40: 62415 Salerno
 [pem_pos=1, final_scores=0.38111007,
 final_local_scores=0.39291376, pem=-0.15453224,
 similarity=16.655415]
 41: 36900 Antioch
 [pem_pos=0, final_scores=0.15758948,
 final_local_scores=0.09339921, pem=-0.36888143,
 similarity=8.882966]
 42: 296046 Joscelin_I,_Count_of_Edessa
 [pem_pos=0, final_scores=0.32568175,
 final_local_scores=0.33810288, pem=0.0, similarity=14.179378]
 43: 285454 Edessa
 [pem_pos=0, final_scores=0.1949255,
 final_local_scores=0.15241599, pem=-0.12033821,
 similarity=9.292528]

Sample: b'Yellowback_stingaree'

The yellowback stingaree , *Urolophus sufflavus* , is a
 locally abundant but little-known species of 1 stingray in
 the family 2 Urolophidae . It is almost 3 endemic to [1 1 New

South Wales 1] , with a range from northward , extending only barely into [2 2 Queensland 2] -LRB- [3 3 Stradbroke Island 3] -RRB- . It inhabits soft-substrate 4 habitat s and has been reported from depths of 45-300 meters -LRB- 150-980 ft -RRB- , though it is most common on the outer 5 continental shelf at depths of 100-160 meters -LRB- 330-520 ft -RRB- . This species attains a maximum length of 42 cm -LRB- 17 in -RRB- . It has a flattened disk about as wide as long , with rounded corners and straight anterior margins . There is a 6 skirt -like nasal curtain in front of the 7 mouth , without lateral lobes . The tail is short and stout , measuring 64-76 % the length of the disk and bearing a serrated spine . The tail ends in a small ; there are no 8 dorsal fin s or lateral folds . The skin is devoid of s . The coloration is uniformly yellowish above , sometimes with an ill-defined brown stripe running down the back . The yellowback stingaree is likely with low 9 fecundity , as in other stingarees . Males mature at a length of 23 cm -LRB- 9 in -RRB- . This species shares the southern extent of its range with the -LRB- U. cruciatus -RRB- . The two species apparently 10 hybridize , highly unusual for , and produce offspring that are intermediate in color pattern . In a 2007 study of 388 fishes , these two species were the only two that could not be distinguished on the basis of their 11 cytochrome c s , attesting to a close evolutionary relationship . Almost the entire range of the yellowback stingaree is under pressure from [1 Australian 1] [2 Commonwealth 2] and State-managed 12 commercial fisheries . This species is caught as 13 bycatch in s and s ; though it is generally discarded , survival post-capture is believed to be low , and in addition stingarees tend to abort gestating young if captured and handled . From 1966-67 to 1996-97 , the capture rate of stingarees in trawl surveys on the [3 New South Wales 3] upper slope , including the yellowback stingaree , declined by some 65 % . Trawl surveys off [4 4 Sydney 4] found a similar decline of 45 % . These negative trends and this species ' restricted distribution has led it to be assessed as 14 Vulnerable by the

1: 21654 New_South_Wales

2: 59051 Queensland

3: 1081368 Stradbroke_Island

4: 27862 Sydney

1: 54195 Stingray

[winner: 3465201 Stingray_Pipeline, pem_pos=19, final_scores=-1.1066161, final_local_scores=-0.9783775, pem=-3.0910425, similarity=-8.297456]

gt: 54195 Stingray, pem_pos=1, final_scores=-1.1957855, final_local_scores=-1.2541852, pem=-1.2216048, similarity=-26.694399]

2: 4533443 Urolophidae

[winner: 22121236 Urotrygonidae, pem_pos=0, final_scores=-0.170995, final_local_scores=-0.096607685, pem=-0.06150808, similarity=1.5412375]

gt: 4533443 Urolophidae, pem_pos=1,
 final_scores=-0.56198674, final_local_scores=-0.520605,
 pem=-2.8191836, similarity=3.8849788]
 3: 937971 Endemism
 [winner: 2118623 Endemic_birds_of_New_Zealand, pem_pos=0,
 final_scores=-0.9186353, final_local_scores=-0.8499235,
 pem=0.0, similarity=-21.273699]
 gt: 937971 Endemism, pem_pos=1, final_scores=-1.1043397,
 final_local_scores=-1.1781292, pem=-0.040974516,
 similarity=-30.845188]
 4: 1596317 Habitat
 [winner: 1596317 Habitat, pem_pos=0,
 final_scores=-1.493295, final_local_scores=-1.591733,
 pem=-0.006358471, similarity=-43.376198]
 gt: 1596317 Habitat, pem_pos=0, final_scores=-1.493295,
 final_local_scores=-1.591733, pem=-0.006358471,
 similarity=-43.376198]
 5: 416651 Continental_shelf
 [winner: 13556749 Outer_Continental_Shelf, pem_pos=0,
 final_scores=-0.8981507, final_local_scores=-0.86186045,
 pem=0.0, similarity=-21.629923]
 gt: 416651 Continental_shelf, pem_pos=1,
 final_scores=-1.1903048, final_local_scores=-1.2748666,
 pem=-0.65528077, similarity=-30.39133]
 6: 231676 Skirt
 [winner: 231676 Skirt, pem_pos=0,
 final_scores=-1.1224445, final_local_scores=-1.1238325,
 pem=-0.00065616437, similarity=-29.444122]
 gt: 231676 Skirt, pem_pos=0, final_scores=-1.1224445,
 final_local_scores=-1.1238325, pem=-0.00065616437,
 similarity=-29.444122]
 7: 9293603 Mouth
 [winner: 9293603 Mouth, pem_pos=0,
 final_scores=-1.2491996, final_local_scores=-1.262919,
 pem=-0.15602832, similarity=-32.749798]
 gt: 9293603 Mouth, pem_pos=0, final_scores=-1.2491996,
 final_local_scores=-1.262919, pem=-0.15602832,
 similarity=-32.749798]
 8: 294244 Dorsal_fin
 [winner: 7625042 Dorsal_pancreatic_artery, pem_pos=5,
 final_scores=-1.077754, final_local_scores=-0.96184015,
 pem=-2.7235835, similarity=-9.802245]
 gt: 294244 Dorsal_fin, pem_pos=2,
 final_scores=-1.5309193, final_local_scores=-1.5856714,
 pem=-1.1255463, similarity=-37.108986]
 [winner: 11036 Fin, pem_pos=0, final_scores=-1.1262611,
 final_local_scores=-1.1003425, pem=-0.29828316,
 similarity=-27.124594]
 RECALL MISS: 294244 Dorsal_fin]
 9: 900798 Fecundity
 [winner: 900798 Fecundity, pem_pos=0,
 final_scores=-1.227575, final_local_scores=-1.2176957,

pem=0.0, similarity=-32.248753|
gt: 900798 Fecundity, pem_pos=0, final_scores=-1.227575,
final_local_scores=-1.2176957, pem=0.0,
similarity=-32.248753]
10: 41244 Hybrid_(biology)
[winner: 41244 Hybrid_(biology), pem_pos=0,
final_scores=-1.3614814, final_local_scores=-1.4397705,
pem=-0.108051, similarity=-38.288315|
gt: 41244 Hybrid_(biology), pem_pos=0,
final_scores=-1.3614814, final_local_scores=-1.4397705,
pem=-0.108051, similarity=-38.288315]
11: 92388 Cytochrome_c
[winner: 8599171 C_R_Irani, pem_pos=10,
final_scores=-0.8453479, final_local_scores=-0.75954837,
pem=-2.5649493, similarity=-4.628132|
RECALL MISS: 92388 Cytochrome_c|
[winner: 6925 Cytochrome, pem_pos=0,
final_scores=-1.0679125, final_local_scores=-0.98615986,
pem=-0.007352974, similarity=-25.299276|
gt: 92388 Cytochrome_c, pem_pos=1,
final_scores=-1.3420328, final_local_scores=-1.3199463,
pem=-0.6931472, similarity=-31.530678]
12: 1028691 Commercial_fishing
[winner: 1621419 Ensign-Bickford_Company, pem_pos=11,
final_scores=-1.0317333, final_local_scores=-0.95663553,
pem=-2.4842746, similarity=-10.948329|
gt: 1028691 Commercial_fishing, pem_pos=13,
final_scores=-1.7573245, final_local_scores=-1.8730202,
pem=-3.0225344, similarity=-35.36793]
[winner: 1459316 Lobster_fishing, pem_pos=2,
final_scores=-1.365618, final_local_scores=-1.4410748,
pem=-0.68021494, similarity=-35.21572|
gt: 1028691 Commercial_fishing, pem_pos=14,
final_scores=-2.5820637, final_local_scores=-2.7739227,
pem=-7.599652, similarity=-37.36158]
13: 363336 Bycatch
[winner: 363336 Bycatch, pem_pos=0,
final_scores=-1.1937033, final_local_scores=-1.224922,
pem=-0.04573194, similarity=-32.215706|
gt: 363336 Bycatch, pem_pos=0, final_scores=-1.1937033,
final_local_scores=-1.224922, pem=-0.04573194,
similarity=-32.215706]
14: 25986983 Vulnerable_species
[winner: 25986983 Vulnerable_species, pem_pos=0,
final_scores=-0.25654376, final_local_scores=-0.21708393,
pem=-0.16373333, similarity=-1.4980975|
gt: 25986983 Vulnerable_species, pem_pos=0,
final_scores=-0.25654376, final_local_scores=-0.21708393,
pem=-0.16373333, similarity=-1.4980975]
1: 21654 New_South_Wales
[pem_pos=0, final_scores=0.6075066,
final_local_scores=0.56781197, pem=-0.051332146,


```

similarity=21.313507]
2: 59051 Queensland
    [pem_pos=0, final_scores=0.40765387,
final_local_scores=0.3571728, pem=-0.083265886,
similarity=15.201274]
3: 1081368 Stradbroke_Island
    [pem_pos=0, final_scores=0.07463975,
final_local_scores=0.12892693, pem=0.0, similarity=7.9371495]
4: 27862 Sydney
    [pem_pos=0, final_scores=0.40415508,
final_local_scores=0.40061098, pem=-0.1443489,
similarity=16.829737]
1: 4689264 Australia
    [pem_pos=0, final_scores=0.65634614,
final_local_scores=0.6141016, pem=-0.491009,
similarity=25.085913]
2: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.14069657,
final_local_scores=0.16806191, pem=-0.69133765,
similarity=12.864617]
3: 21654 New_South_Wales
    [pem_pos=0, final_scores=0.6498632,
final_local_scores=0.61408013, pem=-0.051332146,
similarity=22.69424]

```

Sample: b'Theramine'

Theramine is a 1 Food and Drug Administration -LRB- FDA - RRB- regulated 2 medical food designed to address the increased nutritional requirements associated with 3 chronic pain syndromes and 4 low back pain . Theramine promotes the production of the that help manage and improve the sensory response to 5 pain and 6 inflammation . Theramine is available by 7 prescription only . Theramine is manufactured according to [1 FDA 1] current 8 Good Manufacturing Practices -LRB- cGMP -RRB- . All of the ingredients in Theramine are classified as 9 generally recognized as safe -LRB- GRAS -RRB- by the [2 FDA 2] . Theramine is a specialized formula that consists of : , , HCL , , , , 10 Griffonia Seed -LRB- 95 % 5HTP -RRB- , 11 Whey Protein 12 Hydrolysate , 13 Grape Seed Extract -LRB- 85 % 14 Polyphenols -RRB- , 15 Cinnamon , and Cocoa Extract -LRB- 6 % 16 Theobromine

```

1: 11632 Food_and_Drug_Administration
    [winner: 11632 Food_and_Drug_Administration, pem_pos=0,
final_scores=-0.28568643, final_local_scores=-0.26933306,
pem=-0.0007439638, similarity=-3.943677]
    gt: 11632 Food_and_Drug_Administration, pem_pos=0,
final_scores=-0.28568643, final_local_scores=-0.26933306,
pem=-0.0007439638, similarity=-3.943677 ]
2: 12571217 Medical_food
    [winner: 12571217 Medical_food, pem_pos=0,
final_scores=-1.1007248, final_local_scores=-1.097629,
pem=0.0, similarity=-28.665724]

```

gt: 12571217 Medical_food, pem_pos=0,
final_scores=-1.1007248, final_local_scores=-1.097629,
pem=0.0, similarity=-28.665724]
3: 21081803 Chronic_pain_syndrome
[winner: 3547305 Pain_compliance, pem_pos=3,
final_scores=-1.0093719, final_local_scores=-0.8985119,
pem=-1.0985076, similarity=-16.749819]
RECALL MISS: 21081803 Chronic_pain_syndrome]
[winner: 277102 Syndrome, pem_pos=0,
final_scores=-1.1584316, final_local_scores=-1.1081591,
pem=-0.15345997, similarity=-28.145424]
RECALL MISS: 21081803 Chronic_pain_syndrome]
[winner: 4421272 Chronic_(medicine), pem_pos=0,
final_scores=-1.3519841, final_local_scores=-1.3602173,
pem=-0.07160068, similarity=-36.112507]
RECALL MISS: 21081803 Chronic_pain_syndrome]
4: 618631 Low_back_pain
[winner: 3547305 Pain_compliance, pem_pos=3,
final_scores=-0.91846526, final_local_scores=-0.7992103,
pem=-1.0985076, similarity=-13.786461]
RECALL MISS: 618631 Low_back_pain]
[winner: 18735077 Low_Islands, pem_pos=2,
final_scores=-1.0580928, final_local_scores=-0.9172865,
pem=-2.283167, similarity=-10.867729]
RECALL MISS: 618631 Low_back_pain]
[winner: 2776139 Ralph-Johan_Back, pem_pos=18,
final_scores=-1.2415444, final_local_scores=-1.1670618,
pem=-3.637586, similarity=-10.955982]
RECALL MISS: 618631 Low_back_pain]
5: 24373 Pain
[winner: 3547305 Pain_compliance, pem_pos=3,
final_scores=-0.9309864, final_local_scores=-0.8128877,
pem=-1.0985076, similarity=-14.194623]
gt: 24373 Pain, pem_pos=0, final_scores=-1.152236,
final_local_scores=-1.1079092, pem=-0.050350305,
similarity=-28.698694]
6: 70425 Inflammation
[winner: 70425 Inflammation, pem_pos=0,
final_scores=-0.9953085, final_local_scores=-0.93903214,
pem=-0.014396689, similarity=-23.854588]
gt: 70425 Inflammation, pem_pos=0,
final_scores=-0.9953085, final_local_scores=-0.93903214,
pem=-0.014396689, similarity=-23.854588]
7: 304263 Medical_prescription
[winner: 304263 Medical_prescription, pem_pos=0,
final_scores=-1.2676913, final_local_scores=-1.2925861,
pem=-0.105360545, similarity=-33.910667]
gt: 304263 Medical_prescription, pem_pos=0,
final_scores=-1.2676913, final_local_scores=-1.2925861,
pem=-0.105360545, similarity=-33.910667]
8: 614085 Good_manufacturing_practice
[winner: 614085 Good_manufacturing_practice, pem_pos=0,

final_scores=-0.06276327, final_local_scores=0.020398244,
 pem=0.0, similarity=4.698437|
 gt: 614085 Good_manufacturing_practice, pem_pos=0,
 final_scores=-0.06276327, final_local_scores=0.020398244,
 pem=0.0, similarity=4.698437]
 9: 811905 Generally_recognized_as_safe
 [winner: 811905 Generally_recognized_as_safe, pem_pos=0,
 final_scores=-0.99628013, final_local_scores=-0.9961526,
 pem=0.0, similarity=-25.637466|
 gt: 811905 Generally_recognized_as_safe, pem_pos=0,
 final_scores=-0.99628013, final_local_scores=-0.9961526,
 pem=0.0, similarity=-25.637466]
 10: 23661409 Griffonia
 [winner: 2665069 Griffonia_simplicifolia, pem_pos=0,
 final_scores=-0.74705094, final_local_scores=-0.61129177,
 pem=-0.24625114, similarity=-12.813301|
 gt: 23661409 Griffonia, pem_pos=1,
 final_scores=-0.9499207, final_local_scores=-0.8195458,
 pem=-1.5220037, similarity=-12.090276]
 11: 1525498 Whey_protein
 [winner: 1525498 Whey_protein, pem_pos=0,
 final_scores=-0.8308252, final_local_scores=-0.7601559,
 pem=0.0, similarity=-18.594856|
 gt: 1525498 Whey_protein, pem_pos=0,
 final_scores=-0.8308252, final_local_scores=-0.7601559,
 pem=0.0, similarity=-18.594856]
 12: 33441448 Hydrolysate
 [winner: 5329879 Hydrolyzed_collagen, pem_pos=0,
 final_scores=-0.5270846, final_local_scores=-0.3964889,
 pem=-0.044451755, similarity=-7.5005713|
 gt: 33441448 Hydrolysate, pem_pos=1,
 final_scores=-1.0112666, final_local_scores=-0.9316288,
 pem=-3.1354942, similarity=-6.6606445]
 13: 2058099 Grape_seed_extract
 [winner: 2058099 Grape_seed_extract, pem_pos=0,
 final_scores=-0.39388496, final_local_scores=-0.25724566,
 pem=0.0, similarity=-3.5870109|
 gt: 2058099 Grape_seed_extract, pem_pos=0,
 final_scores=-0.39388496, final_local_scores=-0.25724566,
 pem=0.0, similarity=-3.5870109]
 14: 362892 Polyphenol
 [winner: 362892 Polyphenol, pem_pos=0,
 final_scores=-0.49023044, final_local_scores=-0.3848569,
 pem=0.0, similarity=-7.395185|
 gt: 362892 Polyphenol, pem_pos=0,
 final_scores=-0.49023044, final_local_scores=-0.3848569,
 pem=0.0, similarity=-7.395185]
 15: 53469 Cinnamon
 [winner: 53469 Cinnamon, pem_pos=0,
 final_scores=-0.18157287, final_local_scores=-0.039066777,
 pem=-0.091629155, similarity=3.422176|
 gt: 53469 Cinnamon, pem_pos=0, final_scores=-0.18157287,

```

final_local_scores=-0.039066777, pem=-0.091629155,
similarity=3.422176 ]
16: 31128 Theobromine
    [winner: 31128 Theobromine, pem_pos=0,
final_scores=-0.6196081, final_local_scores=-0.5035186,
pem=0.0, similarity=-10.936286]
    gt: 31128 Theobromine, pem_pos=0,
final_scores=-0.6196081, final_local_scores=-0.5035186,
pem=0.0, similarity=-10.936286 ]
1: 3709121 FDA_(trade_union)
    [pem_pos=0, final_scores=0.3925879,
final_local_scores=0.41686547, pem=0.0, similarity=16.52981]
2: 3709121 FDA_(trade_union)
    [pem_pos=0, final_scores=0.3880031,
final_local_scores=0.4118573, pem=0.0, similarity=16.380356]

```

Sample: b'Whitesands_Bay'

-LRB- [1 Pembrokeshire 1] -RRB- [2 Whitesands Bay 2] -LRB-
-RRB- shown as [3 Whitesand Bay 3] on some maps , is an award-
winning , standard , wide sandy 1 beach in [1 1 St Brides Bay
1] in the [2 2 Pembrokeshire Coast National Park 2] , [3 3 Un
ited Kingdom 3] . The beach is located about two miles west
of the small 2 city of and about one mile south of and has
been described as the best 3 surfing beach in [4
Pembrokeshire 4] and one of the best tourist beaches in the
world . The area to the north east of the bay is dominated by
a large rocky outcrop , at its highest point , known as Carn
Llidi - which could mean Cairn of the Gates or Cairn of
Wrath . The [4 4 Pembrokeshire Coast Path 4] passes alongside
the bay , giving access in the north to the secluded bays of
Porthlleuog and Porthmelgan -LRB- which are only accessible
on foot -RRB- , and the rugged coastal scenery of [5 St
Davids Head 5] . To the south , the coastal path leads to
Porthselau and St. Justinian 's , with views across the [6
Ramsey Sound 6] to [5 5 Ramsey Island 5] . There are a number
of s , stone hut circles and [6 6 British Iron Age 6] 4 field
system s and enclosures to be seen in the vicinity of Carn
Llidi and [7 St Davids Head 7] . It is said that had his
vision to convert [8 Ireland 8] to [7 7 Christianity 7] here
and set sail from the bay in the fifth century . The site of
a Celtic chapel , dedicated to [9 St Patrick 9] , is located
under a mound by the car park just to the east of the bay ,
at what is thought to have been the embarkation point for 5 pi
lgrim s to . At very low tide and after heavy storms the
remains of an ancient , 6 submerged forest can be seen on the
beach , consisting of stumps of birch , fir , hazel and oak
trees . The remains of animals have been found in these
deposits including parts of an 7 aurochs

```

1: 634995 St_Brides_Bay
2: 579827 Pembrokeshire_Coast_National_Park
3: 31717 United_Kingdom
4: 113249 Pembrokeshire_Coast_Path

```

5: 613991 Ramsey_Island
 6: 1299531 British_Iron_Age
 7: 5211 Christianity
 1: 40630 Beach
 [winner: 23885719 Sandy_Beach, New_York, pem_pos=1,
 final_scores=-0.9786251, final_local_scores=-0.94224024,
 pem=-0.60737646, similarity=-20.725607|
 RECALL MISS: 40630 Beach]
 2: 5391 City
 [winner: 807717 Rafsanjan, pem_pos=1,
 final_scores=-0.97742385, final_local_scores=-0.9132872,
 pem=-0.41729954, similarity=-20.89526|
 RECALL MISS: 5391 City]
 3: 28198 Surfing
 [winner: 33139 World_Wide_Web, pem_pos=0,
 final_scores=-1.3943633, final_local_scores=-1.4135721,
 pem=0.0, similarity=-38.094097|
 gt: 28198 Surfing, pem_pos=1, final_scores=-1.6658872,
 final_local_scores=-1.7830002, pem=-0.15044633,
 similarity=-48.30042]
 4: 10066222 Field_system
 [winner: 10066222 Field_system, pem_pos=0,
 final_scores=-0.7305964, final_local_scores=-0.7693963,
 pem=0.0, similarity=-18.870607|
 gt: 10066222 Field_system, pem_pos=0,
 final_scores=-0.7305964, final_local_scores=-0.7693963,
 pem=0.0, similarity=-18.870607]
 5: 61829 Pilgrim
 [winner: 2303527 Pilgrim_(Timothy_Findley_novel),
 pem_pos=4, final_scores=-1.1928226,
 final_local_scores=-1.1123633, pem=-2.281098,
 similarity=-16.70046|
 gt: 61829 Pilgrim, pem_pos=0, final_scores=-1.3347924,
 final_local_scores=-1.4517195, pem=-0.85725284,
 similarity=-34.570614]
 6: 20795362 Submerged_forest
 [winner: 20795362 Submerged_forest, pem_pos=0,
 final_scores=-0.77777445, final_local_scores=-0.7741839,
 pem=0.0, similarity=-19.01348|
 gt: 20795362 Submerged_forest, pem_pos=0,
 final_scores=-0.77777445, final_local_scores=-0.7741839,
 pem=0.0, similarity=-19.01348]
 7: 2494 Aurochs
 [winner: 2494 Aurochs, pem_pos=0,
 final_scores=-1.0345712, final_local_scores=-1.045877,
 pem=-0.004926099, similarity=-27.094551|
 gt: 2494 Aurochs, pem_pos=0, final_scores=-1.0345712,
 final_local_scores=-1.045877, pem=-0.004926099,
 similarity=-27.094551]
 1: 634995 St_Brides_Bay
 [pem_pos=0, final_scores=0.10951643,
 final_local_scores=0.13272724, pem=0.0, similarity=8.050558]

2: 579827 Pembrokeshire_Coast_National_Park
 [pem_pos=0, final_scores=0.3371681,
 final_local_scores=0.32977515, pem=0.0, similarity=13.930861]

3: 31717 United_Kingdom
 [pem_pos=0, final_scores=0.31628984,
 final_local_scores=0.3649215, pem=-0.09933331,
 similarity=15.519889]

4: 113249 Pembrokeshire_Coast_Path
 [pem_pos=0, final_scores=0.30231327,
 final_local_scores=0.3042207, pem=0.0, similarity=13.168265]

5: 613991 Ramsey_Island
 [pem_pos=0, final_scores=0.2038884,
 final_local_scores=0.2029813, pem=0.0, similarity=10.147079]

6: 1299531 British_Iron_Age
 [pem_pos=0, final_scores=0.14003251,
 final_local_scores=0.1341523, pem=-0.0038314452,
 similarity=8.113921]

7: 5211 Christianity
 [pem_pos=0, final_scores=0.13712172,
 final_local_scores=0.18400133, pem=-0.1579441,
 similarity=10.439604]

1: 113250 Pembrokeshire
 [pem_pos=0, final_scores=0.20257546,
 final_local_scores=0.16147369, pem=0.0, similarity=8.90841]

2: 13859182 Whitesands_Bay_(Pembrokeshire)
 [pem_pos=0, final_scores=0.15752439,
 final_local_scores=0.18897632, pem=0.0, similarity=9.729143]

3: 6219027 Whitesand_Bay
 [pem_pos=0, final_scores=0.16810216,
 final_local_scores=0.2506756, pem=0.0, similarity=11.570373]

4: 113250 Pembrokeshire
 [pem_pos=0, final_scores=0.3801179,
 final_local_scores=0.35541183, pem=0.0, similarity=14.69591]

5: 8838126 St_David's_Head
 [pem_pos=0, final_scores=0.3181824,
 final_local_scores=0.3394943, pem=0.0, similarity=14.2209]

6: 613991 Ramsey_Island
 [pem_pos=0, final_scores=0.19262181,
 final_local_scores=0.18904364, pem=0.0, similarity=9.731153]

7: 8838126 St_David's_Head
 [pem_pos=0, final_scores=0.30298084,
 final_local_scores=0.32288888, pem=0.0, similarity=13.725361]

8: 14560 Republic_of_Ireland
 [pem_pos=1, final_scores=0.26857513,
 final_local_scores=0.29779682, pem=-1.5040284,
 similarity=21.155704]

9: 8161655 Saint_Patrick
 [pem_pos=0, final_scores=0.19208615,
 final_local_scores=0.22084033, pem=0.0, similarity=10.680029]

Sample: b'AutoTrack'

AutoTrack is a vehicle **1 swept path analysis** software

program used for analysing the movements of steered and wheeled vehicles including , , , 2 aircraft and other more specialist vehicles such as trucks , and access platforms . AutoTrack was the worlds first 3 swept path analysis software program , originally being jointly developed by TRL -LRB- the UK 's [1 1 Transport Research Laboratory 1] -RRB- and [1 British 1] engineering consultants , Travers Morgan -LRB- acquired by [2 Symonds 2] in 1995 who are now part of [3 Capita Symonds 3] , in the form of TRACK . The term track refers to the tracking of a vehicles simulated movements in relation to geometry , based upon vehicle dimensions , 4 chassis and 5 steering specification . AutoTrack is generally used by transportation engineers , architects and planners for the analysis and design of , , and other facilities to check that provision has been made for the space and 6 geometry required to manouvre specified design . A design vehicle may be a real vehicle modelled within the software 's computer environment but will often be a virtual vehicle that does not exist in real life , but rather is indicative of the type and configuration of vehicle that the final design is expected to accommodate . Design vehicles are commonly specified by the relevant governing body that also controls the specification for which the design must conform . The AutoTrack range is divided into modules that service the 7 Road , and industries and operates within several different systems such as [2 2 AutoCAD 2] by [3 3 Autodesk 3] , by [4 4 Bentley Systems 4] , Bricscad by Bricsys -LRB- a former member of the -RRB- and standalone in [5 5 Microsoft Windows 5] . When the software was first developed it initially ran as a [6 6 DOS 6] based application and when it was ported to operate within 8 AutoCAD

1: 2567865 Transport_Research_Laboratory

2: 2753 AutoCAD

3: 180584 Autodesk

4: 198672 Bentley_Systems

5: 18890 Microsoft_Windows

6: 21304364 DOS

1: 21775090 Swept_path_analysis

[winner: 21775090 Swept_path_analysis, pem_pos=0, final_scores=-0.7271836, final_local_scores=-0.70445865, pem=0.0, similarity=-16.93274]

gt: 21775090 Swept_path_analysis, pem_pos=0, final_scores=-0.7271836, final_local_scores=-0.70445865, pem=0.0, similarity=-16.93274]

2: 849 Aircraft

[winner: 18312727 Short_Type_827, pem_pos=8, final_scores=-1.138172, final_local_scores=-1.0970871, pem=-2.0794415, similarity=-17.341227]

gt: 849 Aircraft, pem_pos=0, final_scores=-1.293948, final_local_scores=-1.3767083, pem=-0.061057825, similarity=-36.661964]

3: 21775090 Swept_path_analysis

```
[winner: 21775090 Swept_path_analysis, pem_pos=0,
final_scores=-0.7167558, final_local_scores=-0.69306785,
pem=0.0, similarity=-16.592813]
gt: 21775090 Swept_path_analysis, pem_pos=0,
final_scores=-0.7167558, final_local_scores=-0.69306785,
pem=0.0, similarity=-16.592813 ]
4: 183330 Chassis
[winner: 1351303 Type_Four_platform, pem_pos=3,
final_scores=-0.89852, final_local_scores=-0.84429306,
pem=-1.0986123, similarity=-15.13125]
gt: 183330 Chassis, pem_pos=0, final_scores=-1.1294436,
final_local_scores=-1.18609, pem=-0.022715928,
similarity=-31.182047 ]
5: 264373 Steering
[winner: 662176 Renault_Laguna, pem_pos=0,
final_scores=-1.3963492, final_local_scores=-1.4524512,
pem=-0.25946823, similarity=-37.843304]
gt: 264373 Steering, pem_pos=2, final_scores=-1.4161777,
final_local_scores=-1.4601545, pem=-2.7889633,
similarity=-24.31739 ]
6: 18973446 Geometry
[winner: 25197704 Galois_geometry, pem_pos=2,
final_scores=-0.9175527, final_local_scores=-0.8207259,
pem=-1.386011, similarity=-12.865043]
gt: 18973446 Geometry, pem_pos=0,
final_scores=-1.0121888, final_local_scores=-1.0038595,
pem=-0.08519087, similarity=-25.404179 ]
7: 25897 Road
[winner: 25897 Road, pem_pos=0, final_scores=-0.23328449,
final_local_scores=-0.18671373, pem=-0.71068734,
similarity=2.3826313]
gt: 25897 Road, pem_pos=0, final_scores=-0.23328449,
final_local_scores=-0.18671373, pem=-0.71068734,
similarity=2.3826313 ]
8: 2753 AutoCAD
[winner: 2753 AutoCAD, pem_pos=0,
final_scores=-0.11958712, final_local_scores=-0.18357943,
pem=-0.0105610695, similarity=-1.3312308]
gt: 2753 AutoCAD, pem_pos=0, final_scores=-0.11958712,
final_local_scores=-0.18357943, pem=-0.0105610695,
similarity=-1.3312308 ]
1: 2567865 Transport_Research_Laboratory
[pem_pos=0, final_scores=0.13277225,
final_local_scores=0.22100177, pem=0.0, similarity=10.684847]
2: 2753 AutoCAD
[pem_pos=0, final_scores=0.08351387,
final_local_scores=0.058812156, pem=-0.0105610695,
similarity=5.902218]
3: 180584 Autodesk
[pem_pos=0, final_scores=0.1308897,
final_local_scores=0.10265821, pem=0.0, similarity=7.153239]
4: 198672 Bentley_Systems
```


[pem_pos=0, final_scores=0.22632731,
final_local_scores=0.26327175, pem=0.0, similarity=11.946267]
5: 18890 Microsoft_Windows
[pem_pos=0, final_scores=0.23005845,
final_local_scores=0.22672722, pem=0.0, similarity=10.855705]
6: 21304364 DOS
[pem_pos=0, final_scores=0.03858237,
final_local_scores=0.007794112, pem=-0.18606299,
similarity=5.334144]
1: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5177869,
final_local_scores=0.58645827, pem=-0.48936236,
similarity=24.252026]
2: 928264 Andrew_Symonds
[pem_pos=0, final_scores=0.06517352,
final_local_scores=0.17690164, pem=-0.9552046,
similarity=14.563362]
3: 2018145 Capita
[pem_pos=1, final_scores=0.40632874,
final_local_scores=0.47560382, pem=-0.30538166,
similarity=19.943392]

Sample: b'Douglas_Dickinson'

Douglas Povah Dickinson 1 CB 2 DSO 3 OBE 4 MC -LRB- 6
November 1886 -- 8 January 1949 -RRB- was a [1 1 British Army
1] officer who commanded the [2 2 East Africa Force 2] at the
start of [3 3 World War II 3] . Dickinson was into the [4 4 We
lch Regiment 4] in 1906 and served in [5 5 World War I 5] in [6
6 France 6] and [7 7 Belgium 7] . He was appointed Deputy
Assistant Quartermaster-General at the [1 5 Staff College 1] ,
[2 Camberley 2] in 1925 , Inspector of the Iraq Army in 6 Kur
distan in 1930 and Inspector with [3 British 3] Military
Mission attached to the Iraq Army in 1932 . He went on to be
of the 1st Battalion of [4 The 7 Welch Regiment 4] in 1934 ,
Commandant of the 8 Nigeria Regiment in 1936 and Inspector-
General of the African Colonial Forces early in 1939 .
Dickinson served in [8 8 World War II 8] as General Officer
Commanding the [9 9 East Africa Force 9] from September 1939
and as Chief of Staff of [10 10 Western Command 10]

1: 4887 British_Army
2: 25744566 East_Africa_Command
3: 32927 World_War_II
4: 1318097 Welch_Regiment
5: 4764461 World_War_I
6: 5843419 France
7: 3343 Belgium
8: 32927 World_War_II
9: 25744566 East_Africa_Command
10: 24643659 Western_Command_(United_Kingdom)
1: 311455 Order_of_the_Bath

[winner: 311455 Order_of_the_Bath, pem_pos=0,
final_scores=-1.130471, final_local_scores=-1.2495914,

pem=-0.41609186, similarity=-30.93781|
 gt: 311455 Order_of_the_Bath, pem_pos=0,
 final_scores=-1.130471, final_local_scores=-1.2495914,
 pem=-0.41609186, similarity=-30.93781]
 2: 355648 Distinguished_Service_Order
 [winner: 355648 Distinguished_Service_Order, pem_pos=0,
 final_scores=-0.57801807, final_local_scores=-0.63513005,
 pem=0.0, similarity=-14.863834|
 gt: 355648 Distinguished_Service_Order, pem_pos=0,
 final_scores=-0.57801807, final_local_scores=-0.63513005,
 pem=0.0, similarity=-14.863834]
 3: 212182 Order_of_the_British_Empire
 [winner: 9740786 Obe_District, pem_pos=5,
 final_scores=-1.472115, final_local_scores=-1.4281139,
 pem=-1.9459101, similarity=-27.945892|
 gt: 212182 Order_of_the_British_Empire, pem_pos=0,
 final_scores=-1.7818502, final_local_scores=-1.9547639,
 pem=0.0, similarity=-54.24434]
 4: 488249 Military_Cross
 [winner: 488249 Military_Cross, pem_pos=0,
 final_scores=-0.96461457, final_local_scores=-1.0656421,
 pem=-1.0457042, similarity=-22.02447|
 gt: 488249 Military_Cross, pem_pos=0,
 final_scores=-0.96461457, final_local_scores=-1.0656421,
 pem=-1.0457042, similarity=-22.02447]
 5: 3641295 Staff_College,_Camberley
 [winner: 249779 Camberley, pem_pos=0,
 final_scores=0.24473123, final_local_scores=0.26162478,
 pem=0.0, similarity=11.897119|
 gt: 3641295 Staff_College,_Camberley, pem_pos=2,
 final_scores=0.0092886835, final_local_scores=-0.03992012,
 pem=-1.095134, similarity=8.853926]
 [winner: 2390045 Staff_college, pem_pos=0,
 final_scores=0.1427782, final_local_scores=0.1428847,
 pem=0.0, similarity=8.353678|
 gt: 3641295 Staff_College,_Camberley, pem_pos=1,
 final_scores=0.14159028, final_local_scores=0.119817585,
 pem=-0.89301306, similarity=12.521652]
 [winner: 37537378 Admissions_(2011_film), pem_pos=12,
 final_scores=-1.7349241, final_local_scores=-1.7209275,
 pem=-4.6728287, similarity=-21.854624|
 RECALL MISS: 3641295 Staff_College,_Camberley]
 6: 80777 Kurdistan
 [winner: 80777 Kurdistan, pem_pos=1,
 final_scores=-0.004342869, final_local_scores=0.017654061,
 pem=-1.8834355, similarity=14.85896|
 gt: 80777 Kurdistan, pem_pos=1,
 final_scores=-0.004342869, final_local_scores=0.017654061,
 pem=-1.8834355, similarity=14.85896]
 7: 1318097 Welch_Regiment
 [winner: 1318097 Welch_Regiment, pem_pos=0,
 final_scores=0.10948028, final_local_scores=0.06854586,

pem=0.0, similarity=6.135258|
gt: 1318097 Welch_Regiment, pem_pos=0,
final_scores=0.10948028, final_local_scores=0.06854586,
pem=0.0, similarity=6.135258]
8: 3873905 Nigeria_Regiment
[winner: 3873905 Nigeria_Regiment, pem_pos=0,
final_scores=0.003218934, final_local_scores=0.04513721,
pem=0.0, similarity=5.4366975|
gt: 3873905 Nigeria_Regiment, pem_pos=0,
final_scores=0.003218934, final_local_scores=0.04513721,
pem=0.0, similarity=5.4366975]
1: 4887 British_Army
[pem_pos=0, final_scores=0.21475469,
final_local_scores=0.14454734, pem=0.0, similarity=8.403294]
2: 25744566 East_Africa_Command
[pem_pos=0, final_scores=0.110942766,
final_local_scores=0.12712601, pem=0.0, similarity=7.8834066]
3: 32927 World_War_II
[pem_pos=0, final_scores=0.47216827,
final_local_scores=0.4518034, pem=0.0, similarity=17.572428]
4: 1318097 Welch_Regiment
[pem_pos=0, final_scores=0.13300501,
final_local_scores=0.09424308, pem=0.0, similarity=6.9021144]
5: 4764461 World_War_I
[pem_pos=0, final_scores=0.37528974,
final_local_scores=0.33254412, pem=0.0, similarity=14.013493]
6: 5843419 France
[pem_pos=0, final_scores=0.6044611,
final_local_scores=0.6481517, pem=-0.16240834,
similarity=24.315054]
7: 3343 Belgium
[pem_pos=0, final_scores=0.57995266,
final_local_scores=0.6191865, pem=-0.24180195,
similarity=23.882431]
8: 32927 World_War_II
[pem_pos=0, final_scores=0.41549903,
final_local_scores=0.38990086, pem=0.0, similarity=15.725133]
9: 25744566 East_Africa_Command
[pem_pos=0, final_scores=0.12418865,
final_local_scores=0.14159513, pem=0.0, similarity=8.315194]
10: 24643659 Western_Command_(United_Kingdom)
[pem_pos=0, final_scores=0.08265893,
final_local_scores=0.0908944, pem=0.0, similarity=6.802183]
1: 2390045 Staff_college
[pem_pos=0, final_scores=0.1427782,
final_local_scores=0.1428847, pem=0.0, similarity=8.353678]
2: 249779 Camberley
[pem_pos=0, final_scores=0.24473123,
final_local_scores=0.26162478, pem=0.0, similarity=11.897119]
3: 31717 United_Kingdom
[pem_pos=0, final_scores=0.14025529,
final_local_scores=0.11536321, pem=-0.48936236,

similarity=10.19361]
4: 1318097 Welch_Regiment
[pem_pos=0, final_scores=0.10948028,
final_local_scores=0.06854586, pem=0.0, similarity=6.135258]

Sample: b'Goh_Seng_Choo_Gallery'

The Goh Seng Choo Gallery is located within the [1 1 National Museum of Singapore 1] . The gallery hosts the collection of [2 2 William Farquhar 2] `s natural history drawings . The gallery is named after the museum 's chairman , Mr G.K. Goh 's late father Mr Goh Seng Choo in his honor . The collection comprises 477 paintings in total , 70 in the main gallery whereas the rest are displayed in the Founding of Singapore section of the [3 3 Singapore History Gallery 3] . This is due to space constraint as many of the paintings occupy the galleries . Prior to his tenure as [1 Singapore 1] 's first Resident and Commandant , [4 4 William Farquhar 4] served as 1 Resident and Commandant of [2 Melaka 2] from 1803 to 1818 . During this period , he made several important botanical and zoological discoveries , and amassed a sizable collection of natural history drawings . Painted by unidentified [5 5 Chinese 5] artists , [3 Farquhar 3] had commissioned these which depict the diverse flora and fauna of the [6 6 Malay Peninsula 6] to the 2 Royal Asiatic Society of Great Britain and Ireland . These 477 drawings have since found their way back to [7 7 Singapore 7] through Mr G.K. Goh , who purchased the paintings at a [4 3 Sotheby 4] 's auction held in [8 8 London 8] for 4 \$ 3 millon . The complete set was donated to the [9 9 National Museum of Singapore 9]

1: 6244699 National_Museum_of_Singapore

2: 1909127 William_Farquhar

3: 32366483 Singapore_History_Gallery

4: 1909127 William_Farquhar

5: 5405 China

6: 20403 Malay_Peninsula

7: 27318 Singapore

8: 17867 London

9: 6244699 National_Museum_of_Singapore

1: 8414383 Dutch_Malacca

[winner: 33588423 Malacca, pem_pos=0,
final_scores=0.2582193, final_local_scores=0.24675626,
pem=-0.20780921, similarity=12.583511]

RECALL MISS: 8414383 Dutch_Malacca]

[winner: 491489 Resident_(title), pem_pos=0,
final_scores=-0.21627428, final_local_scores=-0.23194517,
pem=-0.22167425, similarity=-1.6264948]

RECALL MISS: 8414383 Dutch_Malacca]

[winner: 1090246 Commandant, pem_pos=0,
final_scores=-0.69424725, final_local_scores=-0.71581477,
pem=-0.56724876, similarity=-14.186842]

RECALL MISS: 8414383 Dutch_Malacca]

2: 777291 Royal_Asiatic_Society_of_Great_Britain_and_Ireland

[winner: 14560 Republic_of_Ireland, pem_pos=1,
final_scores=-0.08088277, final_local_scores=-0.053516895,
pem=-1.5040284, similarity=10.6718025]

RECALL MISS: 777291

Royal_Asiatic_Society_of_Great_Britain_and_Ireland]

[winner: 24523313 Royal_Asiatic_Society_of_Sri_Lanka,
pem_pos=0, final_scores=-0.79175776,
final_local_scores=-0.8157018, pem=0.0,
similarity=-20.252455]

gt: 777291

Royal_Asiatic_Society_of_Great_Britain_and_Ireland,
pem_pos=1, final_scores=-0.97632766,
final_local_scores=-1.0558064, pem=-0.0019047508,
similarity=-27.407299]

[winner: 11091761 Great_Island,_Falkland_Islands,
pem_pos=2, final_scores=-1.1623389,
final_local_scores=-1.0644617, pem=-2.7648776,
similarity=-12.640113]

RECALL MISS: 777291

Royal_Asiatic_Society_of_Great_Britain_and_Ireland]

[winner: 31717 United_Kingdom, pem_pos=0,
final_scores=-1.4301114, final_local_scores=-1.5869331,
pem=-0.61114657, similarity=-39.94403]

RECALL MISS: 777291

Royal_Asiatic_Society_of_Great_Britain_and_Ireland]

3: 541267 Sotheby's

[winner: 541267 Sotheby's, pem_pos=0,
final_scores=0.18973516, final_local_scores=0.2376194,
pem=-0.014445427, similarity=11.259306]

gt: 541267 Sotheby's, pem_pos=0,
final_scores=0.18973516, final_local_scores=0.2376194,
pem=-0.014445427, similarity=11.259306]

[winner: 187801 Shopping_cart, pem_pos=8,
final_scores=-2.2506444, final_local_scores=-2.3479564,
pem=-3.4067376, similarity=-47.45162]

RECALL MISS: 541267 Sotheby's]

4: 476174 Singapore_dollar

[winner: 476174 Singapore_dollar, pem_pos=0,
final_scores=-0.9647303, final_local_scores=-1.0605085,
pem=0.0, similarity=-27.557978]

gt: 476174 Singapore_dollar, pem_pos=0,
final_scores=-0.9647303, final_local_scores=-1.0605085,
pem=0.0, similarity=-27.557978]

1: 6244699 National_Museum_of_Singapore

[pem_pos=0, final_scores=0.30694187,
final_local_scores=0.29705355, pem=0.0, similarity=12.954383]

2: 1909127 William_Farquhar

[pem_pos=0, final_scores=0.19083457,
final_local_scores=0.17367876, pem=0.0, similarity=9.272634]

3: 32366483 Singapore_History_Gallery

[pem_pos=0, final_scores=0.09711878,
final_local_scores=0.15377617, pem=0.0, similarity=8.6787]

4: 1909127 William_Farquhar
 [pem_pos=0, final_scores=0.23213993,
 final_local_scores=0.21879858, pem=0.0, similarity=10.619099]
 5: 5405 China
 [pem_pos=1, final_scores=0.4384901,
 final_local_scores=0.47485405, pem=-1.4196125,
 similarity=25.98038]
 6: 20403 Malay_Peninsula
 [pem_pos=0, final_scores=0.111820325,
 final_local_scores=0.12762403, pem=0.0, similarity=7.8982687]
 7: 27318 Singapore
 [pem_pos=0, final_scores=0.5147828,
 final_local_scores=0.49682593, pem=-0.1922004,
 similarity=19.961206]
 8: 17867 London
 [pem_pos=0, final_scores=0.46876162,
 final_local_scores=0.49874407, pem=-0.09086737,
 similarity=19.46738]
 9: 6244699 National_Museum_of_Singapore
 [pem_pos=0, final_scores=0.21223073,
 final_local_scores=0.19359604, pem=0.0, similarity=9.867004]
 1: 27318 Singapore
 [pem_pos=1, final_scores=0.48260075,
 final_local_scores=0.46167195, pem=-0.1922004,
 similarity=18.91214]
 2: 33588423 Malacca
 [pem_pos=0, final_scores=0.2582193,
 final_local_scores=0.24675626, pem=-0.20780921,
 similarity=12.583511]
 3: 1909127 William_Farquhar
 [pem_pos=0, final_scores=0.12407677,
 final_local_scores=0.100755975, pem=0.0, similarity=7.0964723]
 4: 541267 Sotheby's
 [pem_pos=0, final_scores=0.18973516,
 final_local_scores=0.2376194, pem=-0.014445427,
 similarity=11.259306]

Sample: b'Hama_Yumi'

The is a 1 sacred bow used in 1103 2 C.E. in [1 1 Japan 1] .
 The Bow is said to be one of the oldest and most sacred [1
 Japanese 1] weapons ; the first Emperor Jimmu is always
 depicted carrying a bow . According to legend , at that
 time , the [2 Imperial Palace 2] was taken over by an 3 evil 4
 demon , which caused the 5 Emperor to fall ill with great
 anxiety and suffering . When the Imperial High Priests tried
 and failed in their efforts to destroy the demon and dispel
 the Imperial household of its influence , they were at a
 loss . Finally , an 6 archer , [2 2 Yorimasu Minamoto 2] ,
 was summoned to the [3 Imperial Palace 3] in the hopes of
 slaying the demon with his 7 bow and arrow , riding the
 palace of this plague . With a steady hand and a 8 virtuous
 heart , [4 Yorimasu Minamoto 4] vanquished the demon with the

first arrow , and his bow was declared to be a Hama Yumi ; an `` Evil-Destroying Bow '' , -LRB- and the first arrow a Hama Ya ; a `` Evil-Destroying Arrow '' -RRB- . Since then , Hama Yumi have been used in and [3 3 Shinto 3] 9 ritual s of purification . -LRB- For example -- Shihobarai -- the Purification of the Four Directions . In [5 Japan 5] , it is universally believed that merely the twanging of the 10 Bowst ring will frighten ghosts and evil spirits from the house . A 11 Miko will carry a [6 Hama Yumi 6] and a set of Hama Ya as part of their religious regalia , while back in the , they were used quite literally in defense of the temple . As a result , Hama Ya are sold even today at shrines as , -LRB- good-luck charms -RRB- ; smaller replicas have been placed in shrines and people 's homes . It is believed that even one Hamaya which has been blessed by a [7 Shinto 7] Priest carries great spiritual power , will bring protection against the forces of evil , and for purification , and they are also believed to have the ability to attract vast good fortune . Hama Ya and Hama Yumi are often given as gifts to celebrate the first New Year of a male baby 's life . Hama-yumi replicas are scale versions of the Sacred Japanese Bow , coated with , wrapped in fine 12 Rattan and accented in 13 Gold leaf . They are displayed in a stand , along with two arrows tipped with Yanone , -LRB- traditional warrior tips -RRB- , one representing 14 male and the other 15 female

1: 15573 Japan

2: 189396 Minamoto_clan

3: 28272 Shinto

1: 12178732 Sacred

[winner: 12178732 Sacred, pem_pos=0, final_scores=-1.1819907, final_local_scores=-1.2820723, pem=-0.18464987, similarity=-33.165726]

gt: 12178732 Sacred, pem_pos=0, final_scores=-1.1819907, final_local_scores=-1.2820723, pem=-0.18464987, similarity=-33.165726]

2: 6088 Common_Era

[winner: 6088 Common_Era, pem_pos=0, final_scores=-0.17807905, final_local_scores=-0.16625085, pem=0.0, similarity=-0.871544]

gt: 6088 Common_Era, pem_pos=0, final_scores=-0.17807905, final_local_scores=-0.16625085, pem=0.0, similarity=-0.871544]

3: 9897 Evil

[winner: 773618 Evil_demon, pem_pos=0, final_scores=-0.9190305, final_local_scores=-0.8418252, pem=-0.24294616, similarity=-19.71085]

RECALL MISS: 9897 Evil]

4: 8280 Demon

[winner: 773618 Evil_demon, pem_pos=0, final_scores=-0.9190305, final_local_scores=-0.8418252, pem=-0.24294616, similarity=-19.71085]

RECALL MISS: 8280 Demon]

5: 10111 Emperor
[winner: 99598 Qin_Shi_Huang, pem_pos=0,
final_scores=0.016170487, final_local_scores=0.029524252,
pem=0.0, similarity=4.9707756|
gt: 10111 Emperor, pem_pos=1, final_scores=-0.21952164,
final_local_scores=-0.22764191, pem=-1.0488753,
similarity=3.0003715]

6: 2007 Archery
[winner: 5073140 Archer_(automobile), pem_pos=7,
final_scores=-1.2774161, final_local_scores=-1.166616,
pem=-3.4947934, similarity=-11.719207|
gt: 2007 Archery, pem_pos=13, final_scores=-2.272109,
final_local_scores=-2.437343, pem=-3.8420644,
similarity=-47.751717]

7: 219810 Bow_and_arrow
[winner: 219810 Bow_and_arrow, pem_pos=0,
final_scores=-1.041098, final_local_scores=-1.1021732,
pem=0.0, similarity=-28.801334|
gt: 219810 Bow_and_arrow, pem_pos=0,
final_scores=-1.041098, final_local_scores=-1.1021732,
pem=0.0, similarity=-28.801334]

8: 49901 Virtue
[winner: 49901 Virtue, pem_pos=0,
final_scores=-0.9555921, final_local_scores=-1.0082608,
pem=-0.053147472, similarity=-25.709778|
gt: 49901 Virtue, pem_pos=0, final_scores=-0.9555921,
final_local_scores=-1.0082608, pem=-0.053147472,
similarity=-25.709778]

9: 102920 Ritual
[winner: 102920 Ritual, pem_pos=0,
final_scores=-1.5483143, final_local_scores=-1.6733651,
pem=-0.04409594, similarity=-45.607037|
gt: 102920 Ritual, pem_pos=0, final_scores=-1.5483143,
final_local_scores=-1.6733651, pem=-0.04409594,
similarity=-45.607037]

10: 10618229 Bowstring
[winner: 10618229 Bowstring, pem_pos=2,
final_scores=-0.26401114, final_local_scores=-0.17986567,
pem=-0.8430873, similarity=3.3070025|
gt: 10618229 Bowstring, pem_pos=2,
final_scores=-0.26401114, final_local_scores=-0.17986567,
pem=-0.8430873, similarity=3.3070025]

11: 590270 Miko
[winner: 22813199 Ladislav_Miko, pem_pos=5,
final_scores=-0.21271752, final_local_scores=-0.05662182,
pem=-2.3025851, similarity=14.921821|
gt: 590270 Miko, pem_pos=0, final_scores=-0.2151958,
final_local_scores=-0.24732783, pem=0.0,
similarity=-3.2910428]

12: 208612 Rattan
[winner: 208612 Rattan, pem_pos=0,
final_scores=-0.101130635, final_local_scores=-0.047983557,

pem=-0.11700207, similarity=3.2940633|
gt: 208612 Rattan, pem_pos=0, final_scores=-0.101130635,
final_local_scores=-0.047983557, pem=-0.11700207,
similarity=3.2940633]
13: 2927992 Gold_leaf
[winner: 12240 Gold, pem_pos=0, final_scores=-0.16305539,
final_local_scores=-0.10438694, pem=-0.774066,
similarity=5.1840897|
RECALL MISS: 2927992 Gold_leaf]
[winner: 1312625 Metal_leaf, pem_pos=1,
final_scores=-1.06784, final_local_scores=-1.0424345,
pem=-0.67525595, similarity=-23.346464|
gt: 2927992 Gold_leaf, pem_pos=19,
final_scores=-2.5075722, final_local_scores=-2.6695096,
pem=-8.345218, similarity=-30.191185]
14: 1822282 Male
[winner: 1822282 Male, pem_pos=0,
final_scores=-0.80243415, final_local_scores=-0.8165323,
pem=0.0, similarity=-20.27724|
gt: 1822282 Male, pem_pos=0, final_scores=-0.80243415,
final_local_scores=-0.8165323, pem=0.0, similarity=-20.27724]
15: 20913864 Female
[winner: 2539056 Female_tank, pem_pos=4,
final_scores=-1.0386511, final_local_scores=-0.94696826,
pem=-1.9446163, similarity=-13.594582|
gt: 20913864 Female, pem_pos=0, final_scores=-1.1752405,
final_local_scores=-1.2309537, pem=-0.05172588,
similarity=-32.36311]
1: 15573 Japan
[pem_pos=0, final_scores=0.6528282,
final_local_scores=0.6935618, pem=-0.13549899,
similarity=25.523844]
2: 189396 Minamoto_clan
[pem_pos=0, final_scores=0.29683894,
final_local_scores=0.30025724, pem=0.0, similarity=13.049988]
3: 28272 Shinto
[pem_pos=0, final_scores=0.21903382,
final_local_scores=0.17942354, pem=0.0, similarity=9.444069]
1: 15573 Japan
[pem_pos=0, final_scores=0.60549766,
final_local_scores=0.64186037, pem=-0.99076414,
similarity=28.632036]
2: 266152 Tokyo_Imperial_Palace
[pem_pos=0, final_scores=0.21453618,
final_local_scores=0.19194132, pem=-0.36942884,
similarity=11.826636]
3: 266152 Tokyo_Imperial_Palace
[pem_pos=0, final_scores=0.23222615,
final_local_scores=0.21126494, pem=-0.36942884,
similarity=12.403292]
4: 189396 Minamoto_clan
[pem_pos=0, final_scores=0.29115003,

```
final_local_scores=0.29404303, pem=0.0, similarity=12.864544]
5: 15573 Japan
    [pem_pos=0, final_scores=0.5277892,
final_local_scores=0.5569756, pem=-0.13549899,
similarity=21.44784]
6: 990816 Hama_Yumi
    [pem_pos=0, final_scores=0.060957983,
final_local_scores=0.0908941, pem=0.0, similarity=6.802174]
7: 28272 Shinto
    [pem_pos=0, final_scores=0.06325774,
final_local_scores=0.009261847, pem=0.0, similarity=4.366105]
```