

Tarefa 3 - Fake News

Prazo de entrega: 11/10/2020

[Voltar para lista de tarefas](#)

Introdução



Fonte: [UOL](#)

No ano de 2018, o fenômeno das chamadas *fake news* ganhou força no Brasil e proliferou-se nos anos seguintes. A pandemia do coronavírus mostrou-se um ambiente fértil, pois pouco se sabe sobre a doença e muitas pessoas recorrem à internet desesperadas para encontrar respostas.

Nesta tarefa, você estará a cargo de analisar uma série de termos buscados na internet, para tentar encontrar a mais recente *fake news*. Para cada termo, há um histórico de relevância que mostra como o interesse por esse termo variou em um certo período. Assim, um termo considerado relevante por muito tempo pode ser um indicativo de *fake news*, ou de algum evento surpreendente.

Como existem milhões de termos buscados a cada minuto, é imprescindível ter alguma descrição resumida do histórico de relevância de cada um deles. Essa descrição permite comparar os vários termos, oferecendo uma visão global da tendência de pesquisas. Entre as maneiras de descrever resumidamente o histórico de relevância de um termo em um determinado período, está o uso de estatísticas:

- **Máximo:** Indica quanto o termo foi relevante no momento de maior interesse;
- **Mínimo:** Indica quanto o termo foi relevante no momento de maior desinteresse;
- **Média:** Indica a relevância típica do termo considerando todo o período;
- **Desvio-Padrão:** Indica quão abruptas foram as variações de interesse pelo termo.

Por exemplo, o [desvio-padrão](#) indica quão disperso estão os dados em relação à media e um valor de desvio-padrão alto pode ser um bom indicativo de atividade de robôs. Como se sabe, *bots* são ferramentas muito utilizadas para disseminação de *fake news* pois, através do falso engajamento, promovem-se os termos específicos para seu fim. Assim, nesta tarefa você deverá avaliar a qual categoria pertencente cada termo:

- **Bot:** Média ≥ 60 e desvio-padrão > 15 ;
- **Evento surpreendente:** Média ≥ 60 e desvio-padrão ≤ 15 ;
- **Evento normal:** Média < 60 , máximo ≥ 80 e mínimo > 20 ;
- **Evento local:** Média < 60 , máximo ≥ 80 e mínimo ≤ 20 ;
- **Evento irrelevante:** Média < 60 e máximo < 80 .

Critério

O seu programa deverá se chamar **fakenews.c**. Além disso, nesta tarefa, é vedado todo e qualquer tipo de dado não-primitivo (e.g., structs, vetores, etc.) alocado de maneira estática. Esta tarefa será avaliada automaticamente com o uso da ferramenta **Valgrind**.

Dicas

Um termo não possui mais do que 25 caracteres, e sua relevância é medida por porcentagem (*i.e* [0.0,100.0]). Sugere-se o uso de **double** para evitar problemas decorrentes de precisão. Além disso, é garantido que sempre haverá pelo menos um termo a ser lido.

Entrada

A entrada é constituída por várias linhas, como exemplificado abaixo. Na primeira, constam dois números inteiros, N e M , que indicam o número de termos extraídos e a quantidade de dias no histórico, respectivamente. Então, seguem $2N$ linhas cujas informações se estruturam na seguinte forma: para um termo em uma linha i , o seu histórico de M dias constará na linha seguinte (*i.e.*, $i+1$).

Exemplo de entrada

```
16 13
MarielleFranco
39.37846851847752 42.7713666163944 29.569761893653926 30.460029739590173 58.958
JamalKhashoggi
65.63422571794575 61.969402302072766 28.879193218139484 30.43416290926318 41.68
Netflix
48.878682350092696 44.78368373154002 54.563955893880845 64.7122444144054 78.13
JadeBarbosa
53.86282179529801 45.72901471852829 68.62930294129958 66.00484093884023 33.711
EdSheeran
49.02763684295281 30.877373642921388 73.27909661635366 83.59435486361276 58.36
Thanos
41.34706851645633 41.15108587689882 81.45850765417418 73.10845211102516 22.125
ReiLeao
38.224849082578174 80.83201609940504 74.05361712975201 76.0639217821535 26.237
DeuOnda
61.40961019356536 63.66933729589508 66.44757743857534 92.77001253385019 85.310
Anitta
27.30082896243401 50.1853228703274 85.25609747133964 29.605810964621217 22.645
RockinRio
29.70501368123857 65.58236001048928 18.252203903536767 42.20675067471663 52.06
RodaViva
69.3991260254192 40.56543255218003 67.79628926626106 56.600428953073646 49.946
FIFA16
83.43978997320441 43.71024204081357 38.71343167046329 41.16464836357171 57.955
JeanineAnez
53.15053114751034 85.86568930372208 90.9483785015327 46.99941071622334 44.6437
Russia
62.169146078575395 73.98985862052027 41.771628751853754 80.02339592803712 75.3
Facebook
63.97111214647395 73.96442979666216 70.0046658776831 88.63632544509134 32.8636
Franca
58.31430820453798 86.30167186926755 44.13655208009486 76.2901054084678 44.2267
```

Saída

A saída também contém várias linhas (ilustrado abaixo). Para cada termo, deve-se imprimir as estatísticas calculadas a partir de seu histórico na seguinte ordem: máximo, mínimo, média e desvio-padrão. Após imprimir os dados particulares de cada termo, deve-se imprimir as categorias, e o número de elementos que ela contém, e os termos que as constituem.

Exemplo de saída

```
MarielleFranco 90.04 29.57 59.37 19.78
JamalKhashoggi 65.63 22.77 45.69 15.55
Netflix 79.54 18.41 53.38 18.89
JadeBarbosa 80.01 28.19 57.95 16.07
EdSheeran 93.96 29.22 60.52 23.66
Thanos 81.62 15.87 44.62 21.44
ReiLeao 80.83 26.24 61.41 16.15
DeuOnda 92.77 36.02 63.43 17.07
Anitta 85.26 22.65 46.88 18.44
RockinRio 82.35 10.79 44.18 20.93
RodaViva 90.01 40.57 65.40 15.07
FIFA16 84.21 38.33 61.04 17.25
JeanineAnez 97.54 39.49 67.72 20.69
Russia 91.66 33.07 59.28 17.43
Facebook 92.16 25.55 64.95 23.91
Franca 86.30 24.69 56.28 16.60

RESULTADO:
Bot (7): EdSheeran ReiLeao DeuOnda RodaViva FIFA16 JeanineAnez Facebook
Surpreendente (0):
Normal (5): MarielleFranco JadeBarbosa Anitta Russia Franca
Local(2): Thanos RockinRio
Irrelevante (2): JamalKhashoggi Netflix
```