

#### INSTRUÇÕES

1. A data de entrega deste trabalho prático é 6 de Junho, 2020.
2. Só serão aceites trabalhos entregues através da atividade respectiva no Moodle.
3. Os trabalhos serão realizados em grupo de, no máximo, dois elementos.
4. O trabalho prático deverá ser acompanhado de um relatório em PDF.
5. Cada grupo resolve um único enunciado. A escolha do enunciado é realizada através da soma dos números de aluno dos elementos do grupo. Se a soma dos números for ímpar, o grupo deve resolver o enunciado 1. Caso contrário, deverá resolver o enunciado 2.
6. Cada enunciado inclui dois grupos de ficheiros de teste. Um pequeno, e um grande. Só serão aceite trabalhos que consigam processar pelo menos o ficheiro de teste pequeno.
7. De acordo com o artigo 16º do Regulamento de Inscrição, Aprovação e Passagem de Ano da escola, *“A prática ou a tentativa de prática de qualquer fraude acarreta a anulação da prova em que tenha lugar, mediante decisão do docente e constitui infração disciplinar grave, sem prejuízo da responsabilidade civil ou criminal que ao caso couber.”*

## 1 International Movie Database

Foi construído um par de ficheiros que representa a relação entre atores ou outros membros da equipa de produção cinematográfica quando trabalharam no mesmo filme. Estes ficheiros têm a estrutura descrita de seguida.

O ficheiro `actors.txt`:

nm2192965	Mohamed Sharaf	M
nm0232931	José Donoso	?
nm0591877	Miou-Miou	F
nm0792209	Lisa Sheridan	F
nm0587256	Vera Miles	F
nm5457207	Navid Mohammadzadeh	M
nm2154515	Arurdoss	?

Cada linha é composta por três colunas. A primeira corresponde a um código único, identificativo da pessoa em causa. A segunda é o seu nome (podem existir nomes repetidos). A terceira refere-se ao género da pessoa, representado por um M para masculino, F para feminino, ou ? quando essa informação não está disponível na base de dados.

O ficheiro `co-actors.txt` inclui informação sobre o grafo de contracenação:

nm0000748	nm0933491
nm0000748	nm0293366
nm0000748	nm0804026
nm0000748	nm0789520
nm0000748	nm0414038
nm0000748	nm0006428
nm0000748	nm0439850

Cada linha do ficheiro contém dois identificadores, que correspondem a duas pessoas, indicando que essas duas pessoas trabalharam juntas em pelo menos um filme.

Construa o grafo (não orientado, não pesado) que represente a contracenação entre indivíduos, e guarde a informação sobre o seu nome e género.

1. Dado o nome de uma pessoa, indique se existe uma ou mais pessoas com esse mesmo nome. Para cada uma indique o seu identificador.
2. Dado o identificador de uma pessoa, indique quais os indivíduos com que contracenou.
3. Apresente todos os indivíduos que só contracenaram com mulheres (caso existam...).
4. Calcule quem contracenou com mais pessoas.
5. Dados dois identificadores de pessoas, indique se é possível construir um caminho entre elas. Caso seja possível, indique o nome de todos os indivíduos que fazem parte desse caminho. Tente obter o caminho mais curto possível.

## 2 Ibéria Interrail

Foi construído um par de ficheiros com informação relevante sobre cidades e o seu distanciamento. Estes ficheiros têm a estrutura descrita de seguida.

O ficheiro `worldcities.csv`:

```
city,city_ascii,lat,lng,country,iso2,iso3,admin_name,capital,population,id
Tokyo,Tokyo,35.6850,139.7514,Japan,JP,JPN,Tokyo,primary,35676000,1392685764
New York,New York,40.6943,-73.9249,United States,US,USA,New York,,19354922.0,1840034016
Mexico City,Mexico City,19.4424,-99.1310,Mexico,MX,MEX,Ciudad de México,primary,19028000,148...
Mumbai,Mumbai,19.0170,72.8570,India,IN,IND,Maharashtra,admin,18978000,1356226629
```

Cada linha é composta por onze colunas, sendo a primeira linha de títulos. A última coluna (`id`) é o identificador único da cidade. Recomenda-se o uso desta coluna uma vez que poderão existir cidades com o mesmo nome. A informação que não seja relevante para a execução do trabalho poderá ser ignorada.

O ficheiro `idadesIberia.txt` (e `idadesPT.txt`) inclui informação sobre o grafo de ligações entre cidades:

```
1620995356,Funchal,1620949401,Braga,1240.868652
1620995356,Funchal,1620680440,Viana do Castelo,1231.086914
1620995356,Funchal,1620619017,Lisbon,970.868286
```

Cada linha do ficheiro contém o identificador e nome da cidade de origem, identificador e nome da cidade de destino e a distância/custo da viagem entre elas. Atenção que podem não existir ligações em sentido contrário, e/ou terem custos diferentes. Assim, estes dados representam um grafo orientado e pesado.

Construa o grafo (orientado, pesado) que represente a distância/custo da viagem entre as duas cidades, e guarde a informação sobre o seu nome e país.

1. Dado o nome de uma cidade, indique se existe uma ou mais cidades com esse mesmo nome. Para cada uma indique o seu identificador.
2. Dado o identificador de uma cidade origem, indique quais as cidades para onde é possível viajar.
3. Apresente todas as cidades com mais de 50.000 habitantes de onde é possível viajar para outra cidade com menos de 30.000 habitantes (caso existam...).
4. Calcule as cidades de onde é possível partir para o maior número de destinos (caso exista mais que uma, liste todas).
5. Dados dois identificadores de cidades, indique se é possível construir um caminho entre elas. Caso seja possível, indique o nome de todas as cidades fazem parte desse caminho. Tente obter o caminho mais barato/curto possível.