



**Universidade Federal do Rio de Janeiro**  
**Centro de Ciências da Saúde**  
**Instituto de Biofísica Carlos Chagas Filho**

Disciplina: CFB017 - Programação para Biociências

Professor: Dr. Vitor Lima Coelho

**Exercícios 1**

**1. QUESTÃO 1:**

**Importe a biblioteca pandas (veja os vídeos de configuração do ambiente de trabalho) e leia a planilha contendo os dados relativos às mortes por tuberculose obtidos da Organização Mundial de Saúde (“WHO POP TB some.xls”):**

Country	Population (1000s)	TB deaths
Angola	21472	6900
Brazil	200362	4400
China	1393337	41000
Equatorial Guinea	757	67
Guinea-Bissau	1704	1200
India	1252140	240000
Mozambique	25834	18000
Portugal	10608	140
Russian Federation	142834	17000
Sao Tome and Principe	193	18
South Africa	52776	25000
Timor-Leste	1133	990



# IBCCF

INSTITUTO DE BIOFÍSICA  
CARLOS CHAGAS FILHO



UFRJ  
faz 100  
ANOS

1920 | 2020

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

Código:

```
# importação da biblioteca para o script.  
import pandas  
# para o .read_excel poder rodar, o IDE exigiu o xlrd, um pacote do  
pandas para leitura de tabelas em excel (.xls ou .xlsx)  
import xlrd  
# leitura da tabela em formato excel.  
WHO_data = pandas.read_excel('/home/carol/Documents/CFB017/WHO.xls')  
# imprimir toda a tabela.  
print(WHO_data)
```

Output:

Population (1000s)	TB deaths
21472	6900
200362	4400
1393337	41000
757	67
1704	1200
1252140	240000
25834	18000
10608	140
142834	17000
193	18
52776	25000
1133	990



## 2. QUESTÃO 2:

**Calcule o total de mortes:**

Código:

```
# importação da biblioteca para o script.  
import pandas  
# para o .read_excel poder rodar, o IDE exigiu o xlrd, um pacote do  
pandas para leitura de tabelas em excel (.xls ou .xlsx)  
import xlrd  
# leitura da tabela em formato excel.  
WHO_data = pandas.read_excel('/home/carol/Documents/CFB017/WHO.xls')  
# criação de variável  
# seleção da coluna que se quer somar.  
# .sum() realiza a soma e retorna o total.  
total_deaths = WHO_data['TB deaths'].sum()  
print(total_deaths)
```

Output:

> 354715



### 3. QUESTÃO 3:

Calcule o maior e menor número de mortes em um único país:

Código:

```
# importação da biblioteca para o script.
import pandas
# para o .read_excel poder rodar, o IDE exigiu o xlrd, um pacote do
pandas para leitura de tabelas em excel (.xls ou .xlsx)
import xlrd
# leitura da tabela em formato excel.
WHO_data = pandas.read_excel('/home/carol/Documents/CFB017/WHO.xls')
# deixar a tabela em ordem crescente pela terceira coluna.
# a função .sort_values('Nome da Coluna') realiza essa organização.
ascending_deaths = WHO_data.sort_values('TB deaths')
# seleção da primeira e da ultima linha com .iloc[].
# primeira linha: 0, última linha: -1
first_and_last = ascending_deaths.iloc[[0, -1]]
# seleção das colunas que queremos imprimir: [['coluna 1', 'coluna 2']]
# além de deletar a coluna com index, podemos deletar o cabeçalho da
tabela com header=None
print(first_and_last[['Country', 'TB      deaths']].to_string(index=False,
header=None))
```

Output:

```
> Sao Tome and Principe      18
> India                     240000
```



#### 4. QUESTÃO 4:

**Algumas métricas estatísticas podem nos dar a dimensão do nosso problema. Calcule o número médio e a mediana de mortes por tuberculose.**

Código:

```
# importação da biblioteca para o script.  
import pandas  
# para o .read_excel poder rodar, o IDE exigiu o xlrd, um pacote do  
pandas para leitura de tabelas em excel (.xls ou .xlsx)  
import xlrd  
# leitura da tabela em formato excel.  
WHO_data = pandas.read_excel('/home/carol/Documents/CFB017/WHO.xls')  
# cálculo da média pela função .mean().  
média = WHO_data['TB deaths'].mean()  
# cálculo da mediana pela função .median().  
mediana = WHO_data['TB deaths'].median()  
print('Média:', média, 'Mediana:', mediana)
```

Output:

```
> Média: 29559.583333333332 Mediana: 5650.0
```



## 5. QUESTÃO 5:

Para ver a distribuição do número de óbitos, ordene a tabela de maneira crescente:

Código:

```
# importação da biblioteca para o script.  
import pandas  
# para o .read_excel poder rodar, o IDE exigiu o xlrd, um pacote do  
# pandas para leitura de tabelas em excel (.xls ou .xlsx)  
import xlrd  
# leitura da tabela em formato excel.  
WHO_data = pandas.read_excel('/home/carol/Documents/CFB017/WHO.xls')  
# deixar a tabela em ordem crescente pela terceira coluna.  
# a função .sort_values('Nome da Coluna') realiza essa organização.  
ascending_deaths = WHO_data.sort_values('TB deaths')  
# printar sem o index (primeira coluna que indica os números das  
# linhas).  
# a função .to_string transforma a variável tabular em strings e  
# retira a indexação.  
print(ascending_deaths.to_string(index=False))
```

Output:

Country	Population (1000s)	TB deaths
Sao Tome and Principe	193	18
Equatorial Guinea	757	67
Portugal	10608	140
Timor-Leste	1133	990
Guínea-Bissau	1704	1200
Brazil	200362	4400
Angola	21472	6900
Russian Federation	142834	17000
Mozambique	25834	18000
South Africa	52776	25000
China	1393337	41000
India	1252140	240000



## 6. QUESTÃO 6:

Como os países possuem números populacionais variáveis, devemos normalizar a quantidade de óbitos pelo total de habitantes. Portanto, vamos normalizar a taxa de morte por 100.000 habitantes:

$$X \cdot 1000/100.000 = X/100$$

Código:

```
# importação da biblioteca para o script.
import pandas
# para o .read_excel poder rodar, o IDE exigiu o xlrd, um pacote do
pandas para leitura de tabelas em excel (.xls ou .xlsx)
import xlrd
# leitura da tabela em formato excel.
WHO_data = pandas.read_excel('/home/carol/Documents/CFB017/WHO.xls')
# normalização por 100.000 habitantes.
# [Population (1000s)] . 1000/100.000 = [Population (1000s)]/100
pop_normalizada = WHO_data['Population (1000s)']/100
# calculo da taxa de morte por país.
# geração de uma nova coluna na tabela existente.
WHO_data['taxa_morte'] = WHO_data['TB deaths']/pop_normalizada
# impressão de apenas duas colunas: Países e a taxa de mortes.
# pode-se usar .to_csv também, e utilizar sep='\t' para separar as
colunas por tabulação.
print(WHO_data[['Country', 'taxa_morte']].to_string(index=False))
```



# IBCCF

INSTITUTO DE BIOFÍSICA  
CARLOS CHAGAS FILHO

Output:



UFRJ  
faz 100  
ANOS

1920 | 2020

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

Country	taxa morte
Angola	32.134873
Brazil	2.196025
China	2.942576
Equatorial Guinea	8.850727
Guinea-Bissau	70.422535
India	19.167186
Mozambique	69.675621
Portugal	1.319759
Russian Federation	11.901928
Sao Tome and Principe	9.326425
South Africa	47.370017
Timor-Leste	87.378641





## 7. QUESTÃO 7:

**Calcule o total e a média de óbitos por tuberculose nos países pertencentes aos BRICS.**

BRICS = Brasil, Rússia, Índia, China, África do Sul (South Africa)

Código:

```
#importação da biblioteca para o script.
import pandas
# para o .read_excel poder rodar, o IDE exigiu o xlrd, um pacote do
pandas para leitura de tabelas em excel (.xls ou .xlsx)
import xlrd
# leitura da tabela em formato excel.
WHO_data = pandas.read_excel('/home/carol/Documents/CFB017/WHO.xls')
# Selecionando apenas os países da BRICS.
BRICS = WHO_data.iloc[[1, 2, 5, 8, 10]]
# calcular a soma e colocar numa variável.
total_BRICS = BRICS['TB deaths'].sum()
# calculo da média. a função .sum() já deixa a variável como um número
ao invés de uma string.
media_BRICS = total_BRICS/5
# imprimir total.
print('Total de Mortes: ',total_BRICS)
# imprimir média.
print('Média de Mortes: ',media_BRICS)
```

Output:

```
> Total de Mortes: 327400
> Média de Mortes: 65480.0
```