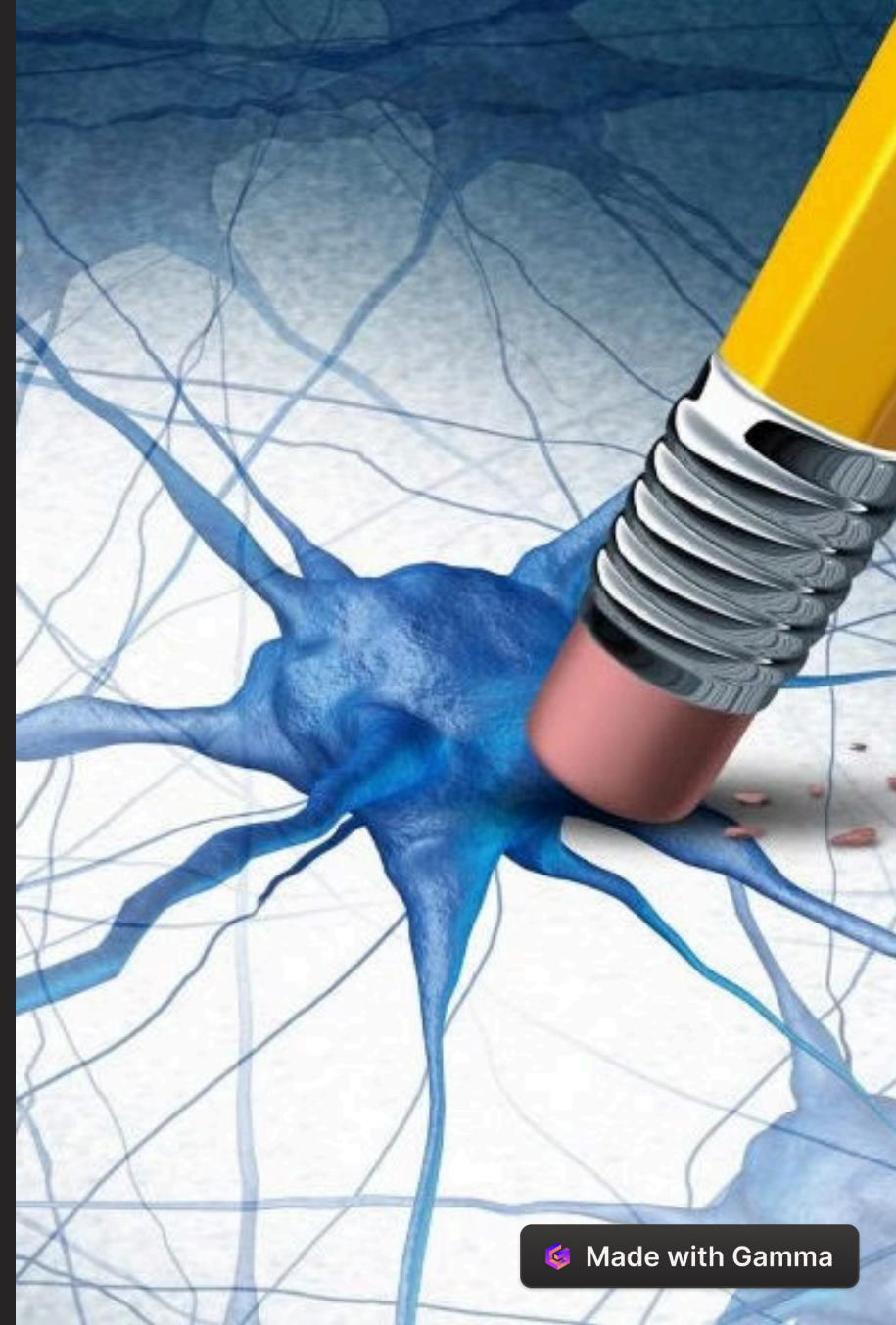


Introdução ao Alzheimer

O Alzheimer é uma doença neurodegenerativa que afeta milhões de pessoas em todo o mundo. É a forma mais comum de demência, causando problemas de memória, cognição e comportamento. Neste slide iremos mostrar um Datset com o registro de óbitos por Alzheimer em vários países no ano de 2019.

Por: Jefferson Martins, Amadeu Seabra, Ryan Gabriel, Gustavo Junqueira.



Dados sobre a doença no mundo

Prevalência Global

Estima-se que existam cerca de 50 milhões de pessoas com Alzheimer no mundo.

Mortes em 2019

Em 2019, o Alzheimer foi a causa de mais de 1,5 milhão de mortes em todo o mundo.

Aumento Projetado

Prevê-se que o número de pessoas com Alzheimer dobre a cada 20 anos, chegando a 152 milhões em 2050.

Custos Econômicos

Os custos diretos e indiretos relacionados ao Alzheimer atingem centenas de bilhões de dólares por ano em todo o mundo.

20
nº de anos
que você
pode ter a
doença
antes dos
sintomas
aparecerem



6ª
causa de
mortes nos
Estados
Unidos



2/3
são
mulheres



O dataset está delimitado por ponto e vírgula (;). O conteúdo está sendo lido como uma única coluna. Este dataset é sobre Alzheimer e contém 106 entradas e 4 colunas, que são:

- Entidade: Nome do país ou região.
- Código: Código do país ou região.
- Ano: Ano em que os dados foram coletados.
- Óbitos: Taxa de óbitos por Alzheimer (presumivelmente por 100.000 habitantes).

Tabela do Dataset

Este código lê um arquivo CSV chamado 'dataset-alzheimer.csv' usando a biblioteca pandas, carregando em um DataFrame. Ele usa ';' como separador de campo e a codificação ISO 8859-1 para interpretar os caracteres, e salva isso em uma variável chamada 'dataset'

```
dataset = pd.read_csv('dataset-alzheimer.csv', sep=';', encoding='iso-8859-1')
dataset
```

	Entidade	Código	Ano	Óbitos
0	Afghanistan	AFG	2019	30.79
1	African Region (WHO)	RSA	2019	25.62
2	Albania	ALB	2019	24.00
3	Algeria	DZA	2019	26.97
4	American Samoa	ASM	2019	23.03
...
101	Kuwait	KWT	2019	25.08
102	Kyrgyzstan	KGZ	2019	23.47
103	Laos	LAO	2019	22.35
104	Latin America & Caribbean (WB)	LAC	2019	23.65
105	Latvia	LVA	2019	21.34

106 rows × 4 columns

Comandos

Este comando retorna um objeto Index contendo os nomes das colunas, e **dataset.columns.str.strip()** aplica o método **strip()** a cada nome de coluna para remover os espaços em branco. Em seguida, os nomes das colunas são atualizados no DataFrame usando **dataset.columns =** O **print(dataset.columns)** exibe os nomes das colunas após a remoção dos espaços em branco.

```
dataset.columns = dataset.columns.str.strip()
print(dataset.columns)
```

```
Index(['Entidade', 'Código', 'Ano', 'Óbitos'], dtype='object')
```

Esses 6 comandos retornam a Soma, Mediana, Mínimo, Máximo, Média e o Desvio Padrão de Óbitos registrados.

```
[ ] mediana_obtios = dataset ['Óbitos'].median()
    print("Mediana de registro de óbitos:", mediana_obtios)
```

```
⇒ A mediana de registro de óbitos é: 23.03
```

```
[ ] soma_obtios = dataset ['Óbitos'].sum()
    print("Soma do registro de óbitos:", soma_obtios)
```

```
⇒ A soma do registro de óbitos é: 2511.24
```

```
[ ] min_obtios = dataset ['Óbitos'].min()
    print("Menor registro de óbitos:", min_obtios)
```

```
⇒ O menor registro de óbitos é: 18.23
```

```
[ ] max_obtios = dataset ['Óbitos'].max()
    print("Maior registro de óbitos:", max_obtios)
```

```
⇒ O maior registro de óbitos é: 33.29
```

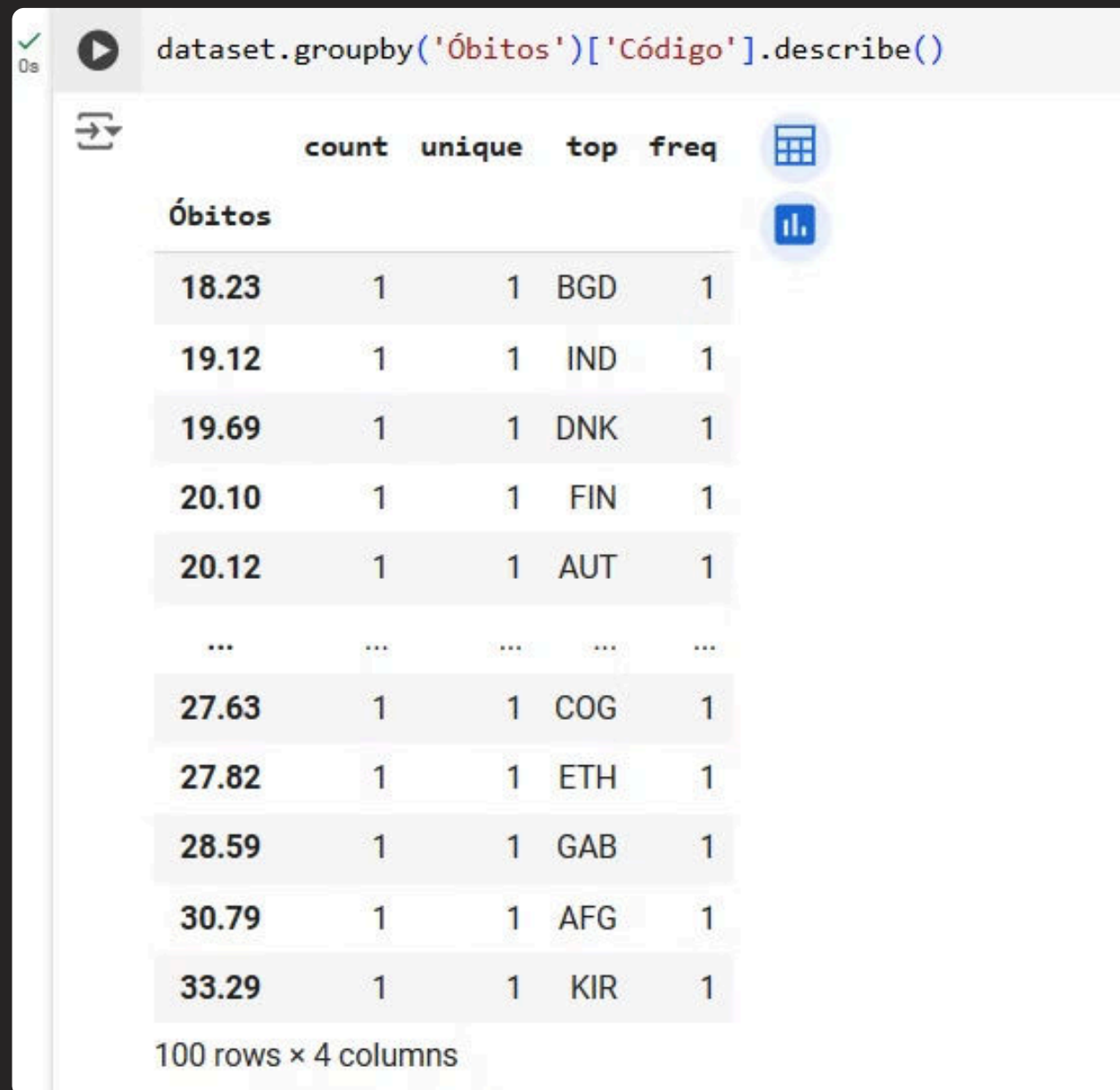
```
[ ] media_obtios = dataset ['Óbitos'].mean()
    print("Média do registro de óbitos:", media_obtios)
```

```
⇒ A média do registro de óbitos é: 23.475566037735852
```

```
[ ] desvio_obtios = dataset ['Óbitos'].std()
    print("Desvio padrão do registro de óbitos:", desvio_obtios)
```

```
⇒ O desvio padrão do registro de óbitos é: 2.61708334817186
```

Esse comando de GroupBy retorna as estatísticas descritivas da coluna óbitos com o seu Código (to)



The screenshot shows a Jupyter Notebook interface. At the top, a code cell contains the command `dataset.groupby('Óbitos')['Código'].describe()`. Below the code cell, the output is displayed as a table. The table has five columns: 'Óbitos', 'count', 'unique', 'top', and 'freq'. The rows represent different death counts, with the first row being 18.23 and the last row being 33.29. Each row shows a count of 1, a unique value of 1, a top value (e.g., BGD, IND, DNK, FIN, AUT, COG, ETH, GAB, AFG, KIR), and a frequency of 1. The table is truncated, showing only 100 rows out of 106. To the right of the table, there are two icons: a calendar icon and a bar chart icon.

Óbitos	count	unique	top	freq
18.23	1	1	BGD	1
19.12	1	1	IND	1
19.69	1	1	DNK	1
20.10	1	1	FIN	1
20.12	1	1	AUT	1
...
27.63	1	1	COG	1
27.82	1	1	ETH	1
28.59	1	1	GAB	1
30.79	1	1	AFG	1
33.29	1	1	KIR	1

100 rows x 4 columns

❗ E possível que algumas linhas de óbitos tenha alguns valores iguais aos de outras linhas, por isso são mostrado apenas 100 rows, ao invés das 106 rows.

Gráficos

Gráfico de Dispersão

O gráfico apresentado mostra os 30 países com o menor número de óbitos por Alzheimer, com base em dados agrupados.



Gráfico de Colunas

O gráfico apresentado mostra 15 registro óbitos entre 20 a 30, não mostra todos os registros entre 20 a 30 registros de óbitos, pois o gráfico e dividido em somente 15 bins (15 colunas).

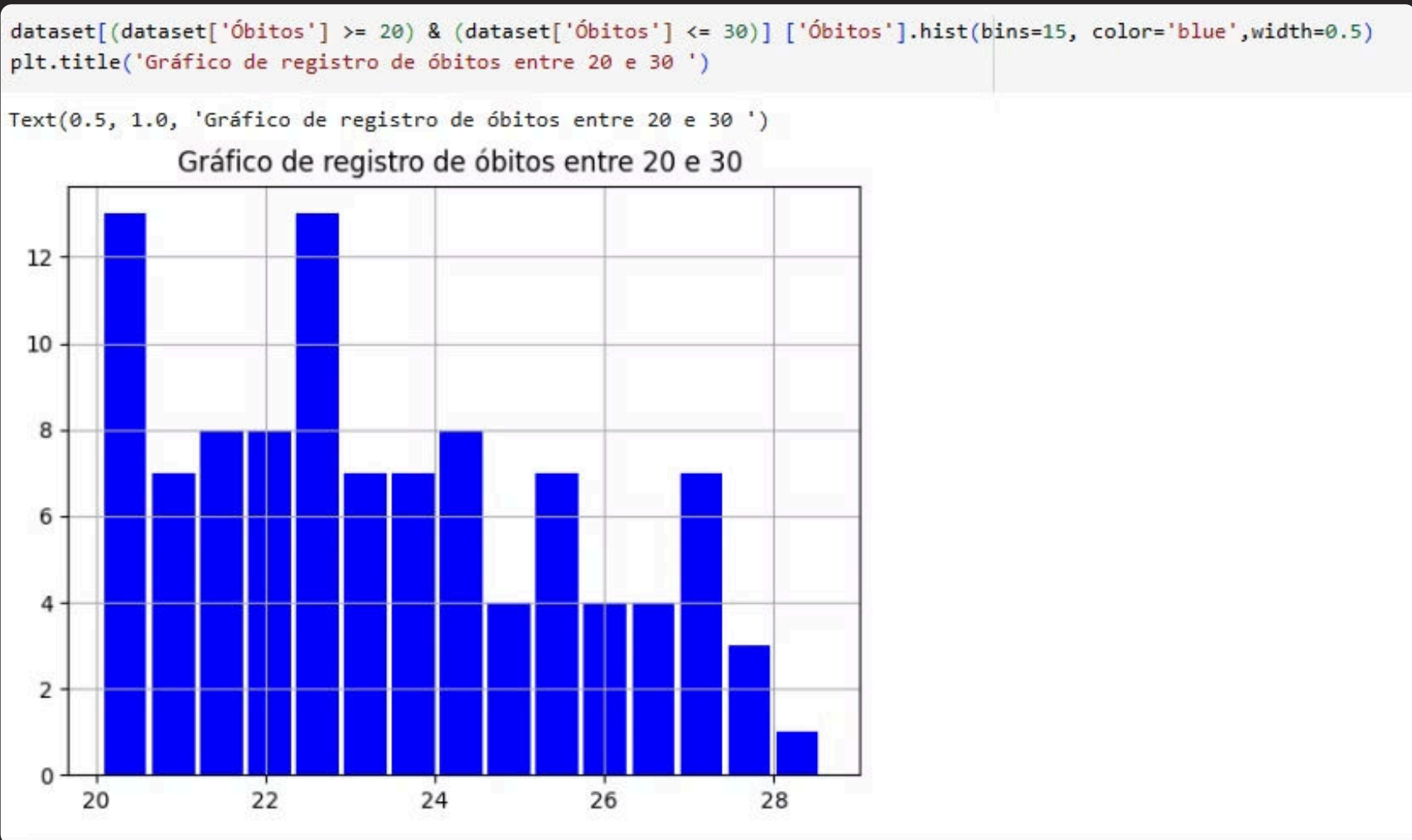


Gráfico de Pizza

Este gráfico mostra a quantidade de registros óbitos em relação a média(23.475...), a parte Laranja represente os registros abaixo da média e o Azul representa os registros acima da média

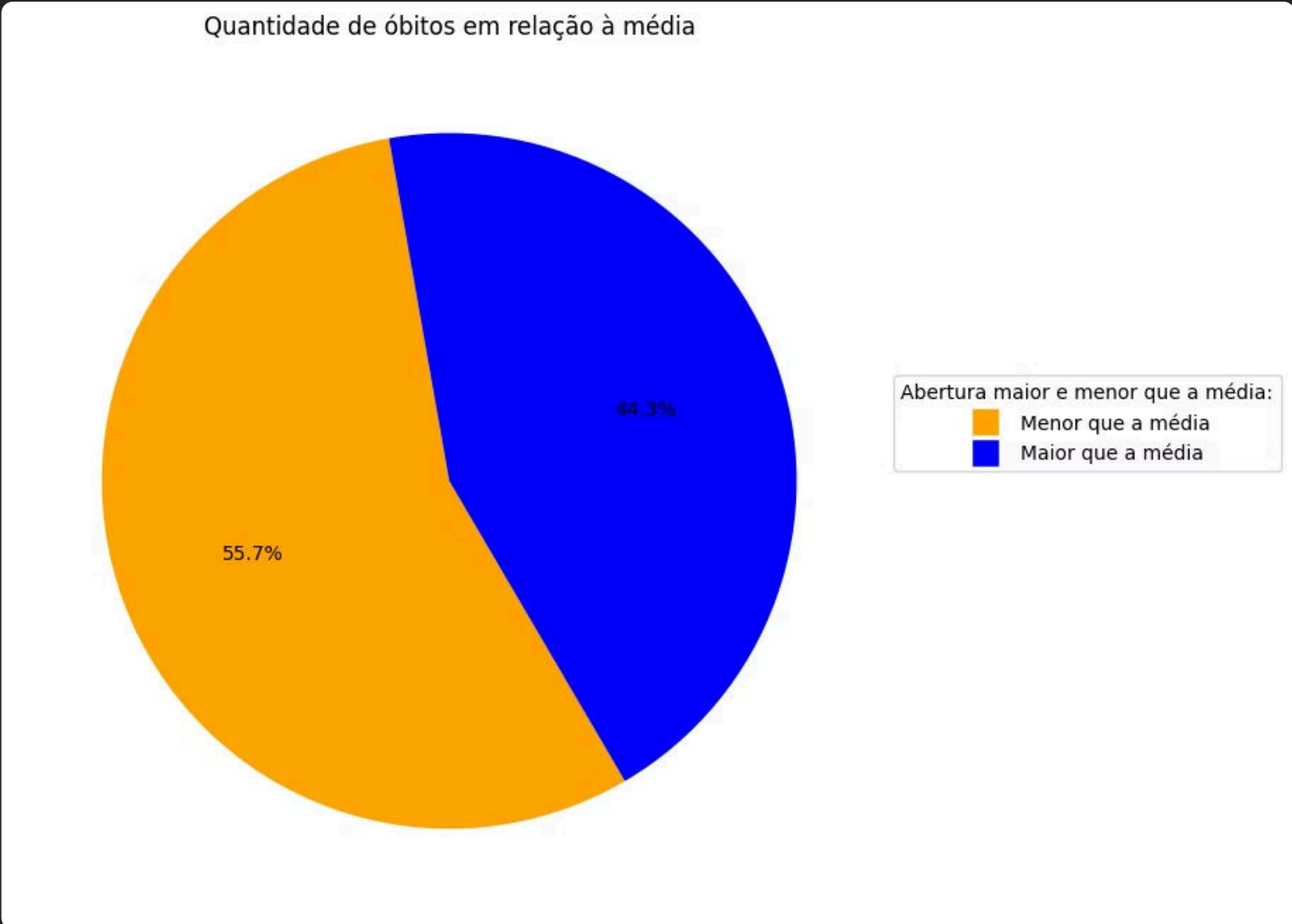


Gráfico de Colunas

Este gráfico mostra os 20 países com o registros de óbitos de forma decrescente

