

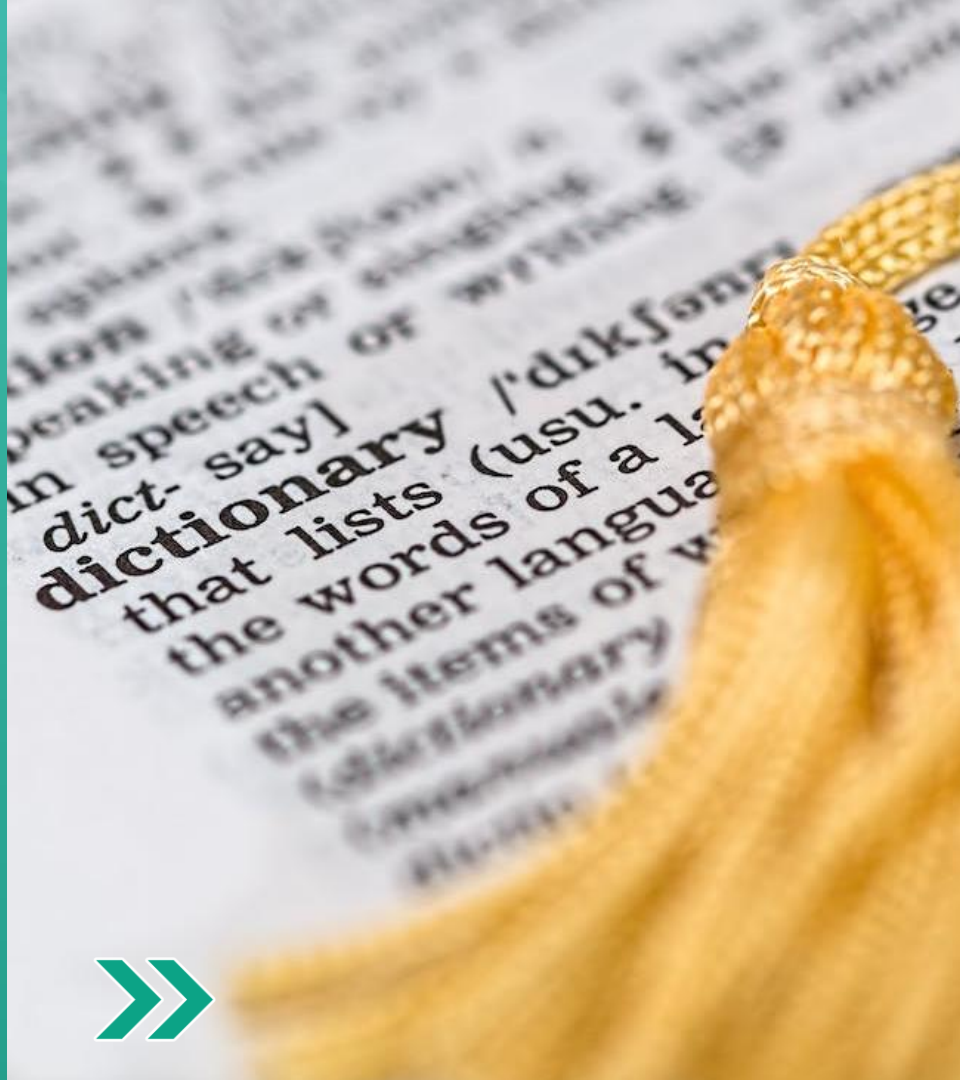


escola
britânica de
artes criativas
& tecnologia

Python para Análise de Dados



GLOSSÁRIO



Análise de Dados



Dica: para encontrar rapidamente a palavra que procura aperte o comando CTRL+F e digite o termo que deseja achar.

- **Conheça a Kaggle**
- **Explore dados**
- **Transforme e limpe dados**
- **Visualize dados**
- **Referência Bibliográfica**



Conheça a Kaggle



Conheça a Kaggle

Glossário

• Notebook

Os notebooks do Colab permitem combinar código executável e rich text em um só documento, além de imagens, HTML, LaTeX.

• Kaggle

Comunidade on-line de cientistas de dados e praticantes de aprendizado de máquina.

• Linguagem R

É uma linguagem de programação multi-paradigma orientada a objetos, programação funcional, dinâmica, fracamente tipada, voltada à manipulação, análise e visualização de dados.



Explore dados



Explore dados

Glossário

• Atributos categóricos

Podem ser ordinais ou nominais. Os nominais são aqueles que não remetem a uma ordem, por exemplo, "UF", "Gênero". Já dados categóricos ordinais remetem a uma ordenação, como tamanho de roupa e classificação.

• Drop

O método *drop()* remove a linha ou coluna especificada.

• Pandas.isna

Este método *pandas.isna()* é usado para detectar valores ausentes para um objeto semelhante a uma matriz.

• Schema

É uma biblioteca para validar as estruturas de dados do Python.



Explore dados

Glossário

● Variável resposta

Também chamada de variável dependente, é aquela que se quer entender e localiza-se no eixo Y do gráfico.

● Variável explicativa

Também chamada de independente ou preditora, é aquela que ajuda a explicar a variação da variável resposta e geralmente se localiza no eixo X do gráfico.

● Shape

A função `shape()` vai retornar o número de linhas e colunas de uma matriz.



Transforme e limpe dados



Transforme e limpe dados

Glossário

• Método `apply`

Os objetos passados para `pandas.apply()` são objetos Series cujo índice é do DataFrame (eixo = 0) ou as colunas do DataFrame (eixo = 1).

Por padrão (`result_type = None`), o tipo de retorno final é inferido do tipo de retorno da função aplicada.



Visualize dados



Visualize dados

Glossário

● Capitalize

O método *capitalize()* retorna uma string onde o primeiro caractere é maiúsculo, e o resto é minúsculo.

● Eixo

Os gráficos têm dois eixos usados para medir e categorizar dados: um eixo vertical (eixo dos valores ou o eixo y) e um eixo horizontal (eixo das categorias ou eixo x).

● Matplotlib

É uma biblioteca da linguagem de programação Python, utilizada para visualização de dados e plotagem gráfica.

● Plot

O método plot cria um gráfico de linhas.



Visualize dados

Glossário

● Rotate

A função *rotate()* é usada para rotacionar as imagens.

● Set

É uma coleção de itens desordenada, parcialmente imutável e que não pode conter elementos duplicados. Por ser parcialmente imutável, possui permissão de adição e remoção de elementos.

● Subplots

A função *subplot()* recebe três argumentos que descrevem o layout da figura.



Referências Bibliográficas

Para descobrir mais termos da área. Acesse:

- <https://www.w3schools.com/python/default.asp>
- <http://www.estruturas.ufpr.br/disciplinas/pos-graduacao/introducao-a-computacao-cientifica-com-python/introducao-python/3-1-plotagem-simples/>
- <https://matplotlib.org/stable/tutorials/introductory/pyplot.html>
- <https://acervolima.com/retorne-varias-colunas-usando-o-metodo-pandas-apply/>



Bons estudos!

