

Profissão: Cientista de Dados



GLOSSÁRIO



Métodos de análise



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

- **Método chaining**
- **Use o método `.apply()`**
- **Explore dados com Pandas Profiling**
- **Copie e cole no Excel o método `.to_clipboard()`**
- **Funções de janela móvel (window functions)**



Method chaining (encadeamento de métodos)



Method chaining

● Variáveis Intermediárias

Variáveis usadas para armazenar resultados temporários durante o processamento de dados. Embora possam ser úteis para a depuração, o uso excessivo de variáveis intermediárias pode tornar o código mais difícil de ler e menos eficiente.

● Assign

Método usado para criar uma nova variável dentro do dataframe.



Use o método `.apply()`



Use o método `.apply()`

• **axis**

Parâmetro no método `.apply()` que determina se a função será aplicada ao longo das colunas (`axis=0`) ou das linhas (`axis=1`).



Método mágico `%timeit`

Método em Python usado para medir o tempo de execução de um pedaço de código. Pode ser usado para comparar a performance de diferentes métodos.



Explore dados com Pandas Profiling



Explore dados com Pandas Profiling

● Colunas constantes

Colunas em um conjunto de dados que têm o mesmo valor para todas as entradas. Essas colunas não são úteis para a análise, pois não fornecem nenhuma variação ou informação útil.

● Correlações fortes

Relações estatísticas entre duas ou mais variáveis onde a mudança em uma variável está associada à mudança em outra variável. Correlações fortes entre variáveis podem indicar redundância nos dados.



Copie e cole no Excel o método .to_clipboard()



Copie e cole no Excel o método `.to_clipboard()`

• `.to_clipboard()`

Método do Pandas que copia os dados de um DataFrame para a área de transferência, permitindo que sejam colados diretamente em outro programa, como o Excel, com o comando Ctrl+V.

• `.to_excel()`

Método do Pandas que salva os dados de um DataFrame diretamente em um arquivo Excel, eliminando a necessidade de copiar e colar manualmente.



Funções de janela móvel (window functions)



Funções de janela móvel (window functions)

● Agregação móvel

É uma técnica que envolve a aplicação de uma função de agregação, como a média, a um conjunto de dados em um intervalo móvel. Isso é útil para suavizar dados ao longo do tempo.

● Funções de janela móvel (window functions)

São funções que operam em um conjunto de linhas (uma "janela") de um DataFrame ou Series. Exemplos incluem a média móvel e a soma móvel.



Funções de janela móvel (window functions)

• Média móvel

É uma técnica que envolve o cálculo da média de um conjunto de dados em um intervalo móvel. Isso suaviza a curva dos dados, reduzindo o impacto de picos e outliers.

• Funções 'up' e 'down'

São funções usadas para aumentar ou diminuir a amostra de dados.

• Soma móvel

É uma técnica que envolve o cálculo da soma de um conjunto de dados em um intervalo móvel.

• Função 'shift'

É uma função que permite comparar um valor de dados com o valor anterior.



Bons estudos!

