

Profissão: Cientista de Dados



GLOSSÁRIO



Descritiva III – Gráficos interativos



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

- **Utilize o pacote Plotly**
- **Calcule o retorno de séries**
- **Calcule as diferenças**
- **Crie gráfico de Candlestick**



Utilize o pacote Plotly



Utilize o pacote Plotly

• Box plot

É um método para representar graficamente grupos de dados numéricos através de seus quartis.

• Express Line

É um comando do Plotly que permite criar gráficos de linhas.

• Plotly (PT)

É um pacote popular para criação de gráficos em Python, que também suporta outras linguagens como R, Julia e Matlab.



Calcule o retorno de séries



Calcule o retorno de séries

• Correlações Espúrias

São associações que não têm uma relação causal legítima e podem levar a mal-entendidos. Por exemplo, a correlação entre o consumo de mussarela per capita e o número de doutorados em engenharia civil.

• Função `pct_change()`

Função em Python usada para calcular o retorno de uma ação. Ela calcula a variação percentual entre o valor atual e um valor anterior.



Calcule as diferenças



Calcule as diferenças

• Down Sampling

É o processo de reduzir a frequência de uma série temporal. Por exemplo, passar de uma série diária para uma série mensal ou anual.

• Média Móvel

É uma média que se move conforme a janela de tempo escolhida. É uma forma de suavizar a série temporal, reduzindo a volatilidade e destacando tendências de longo prazo.

• Função 'ple'

Função do Python utilizada para agrupar uma série por um determinado período. Possui dois parâmetros importantes: 'right', que fecha a série à direita, e 'lab', que dá o nome para a série.

• Roling

Método do pandas utilizado para calcular a média móvel de uma série temporal.



Crie gráfico de Candlestick



Crie gráfico de Candlestick

● **Candlestick**

É um tipo de gráfico usado principalmente para descrever movimentos de preços de derivativos financeiros, ações e moedas em um determinado período.

● **Comando 'lock'**

No contexto do pandas, é um comando usado para selecionar dados específicos de um DataFrame.

● **Comando 'zp'**

É um comando usado para criar pares de dados.



Bons estudos!

