

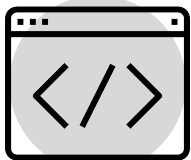
# Profissão: Cientista de Dados



# BOAS PRÁTICAS



# Combinação de modelos I



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



# Utilize o pacote Plotly

- Ao trabalhar com objetos de dados com índices múltiplos, é importante saber como referenciar os dados corretamente. Além disso, pode ser útil transformar o objeto em um objeto com índices simples para facilitar a manipulação dos dados.
- Ao usar o Plotly, explore a galeria de gráficos disponíveis. Isso pode ajudá-lo a encontrar o tipo de gráfico que melhor se adapta aos seus dados e às suas necessidades de análise.



# Calcule o retorno de séries

- Sempre verifique se as correlações observadas em seus dados têm uma relação causal legítima. Evite cair na armadilha das correlações espúrias.
- Utilize a função `pct_change()` em Python para calcular o retorno de uma ação. Esta função calcula a variação percentual entre o valor atual e um valor anterior.
- Ao analisar séries temporais, como preços de ações, é importante remover a tendência para evitar correlações espúrias. Uma maneira de fazer isso é calculando o retorno das ações, que é a diferença entre o preço de venda e o preço de compra, dividido pelo preço de compra.
- Ao visualizar os retornos de várias ações em um gráfico, esteja ciente de que, embora as ações geralmente tenham uma tendência de crescimento a longo prazo, os retornos diários são geralmente pequenos e não têm uma tendência clara.



# Calcule o retorno de séries

- Outra técnica para remover a tendência de séries é a diferença. A diferença entre os valores em uma série pode ser usada para remover a tendência. Em Python, você pode usar a função `diff()` para calcular a diferença entre os valores em uma série.
- Sempre questione seus resultados e verifique se eles fazem sentido no contexto do problema que você está tentando resolver. Lembre-se de que correlações espúrias podem ser enganosas.



# Suavize séries

- Utilize a função 'resample' do pandas para realizar o down sampling, que é útil quando se deseja analisar a série em um período maior.
- Ao usar a função 'resample', lembre-se de que os parâmetros 'right' e 'label' são importantes. 'Right' determina se a série é fechada à direita e 'label' dá o nome para a série.
- A média móvel é uma ferramenta útil para análise de séries temporais. Ela pode ser calculada usando o método 'rolling' na série de dados.
- Ao plotar gráficos de séries temporais, considere empilhar a média móvel para facilitar a visualização e comparação de diferentes séries.



# Crie gráfico de Candlestick

- Personalize o gráfico para torná-lo mais eficaz e fácil de entender. Isso pode incluir a definição do layout e a atualização do eixo X.
- Ao criar o gráfico, passe os parâmetros necessários através de um dicionário. Isso inclui o índice (data), valor de abertura, valor de fechamento, máximo, mínimo e o tipo do gráfico.
- Ao criar um gráfico de candlestick, é importante entender o que cada elemento representa. A linha fina mostra o mínimo e o máximo, enquanto a caixa representa a abertura e o fechamento do pregão. A cor da caixa indica se a bolsa terminou em alta (verde) ou em baixa (vermelha).





# Bons estudos!

