

Lista de Desafios de Estatística

1. Amostragem Aleatória Simples

- Desafio: Crie um DataFrame com 1000 registros de alunos, incluindo colunas para "ID", "Idade" e "Nota". Realize uma amostragem aleatória simples de 100 alunos e analise a média e a mediana das notas da amostra.

- Objetivo: Praticar a amostragem e entender as medidas de tendência central.

2. Boxplot e Identificação de Outliers

- Desafio: Use o DataFrame criado no desafio anterior e adicione uma coluna "Nota" com notas variando de 0 a 100. Plote um boxplot para as notas e identifique os outliers.

- Objetivo: Visualizar a distribuição das notas e entender o conceito de outliers.

3. Histograma para Frequência

- Desafio: Plote um histograma das idades dos alunos do DataFrame anterior, dividindo as idades em intervalos (bins) apropriados.

- Objetivo: Analisar a frequência das idades e entender a distribuição de dados.

4. Cálculo de Covariância

- Desafio: Adicione uma coluna "Horas de Estudo" ao DataFrame anterior. Calcule a covariância entre "Horas de Estudo" e "Nota".

- Objetivo: Compreender a relação entre duas variáveis.

5. Regressão Linear Simples

- Desafio: Com base nas colunas "Horas de Estudo" e "Nota", ajuste um modelo de regressão linear simples. Plote a linha de regressão junto com os dados.

- Objetivo: Aprender sobre modelagem e previsão com regressão.

6. Teste Qui-Quadrado

- Desafio: Crie um DataFrame com a preferência de cursos (por exemplo, "Matemática", "Física", "Química") entre alunos de diferentes idades. Realize um teste qui-quadrado para verificar se há associação entre a idade e a preferência de curso.

- Objetivo: Aplicar o teste qui-quadrado para dados categóricos.

7. K-Means para Agrupamento

- Desafio: Geração de um conjunto de dados com características de clientes (por exemplo, "Renda" e "Gastos"). Aplique o algoritmo K-Means para agrupar os clientes e visualize os grupos.

- Objetivo: Entender a técnica de agrupamento e suas aplicações.

8. Análise de Componentes Principais (PCA)

- Desafio: Use um conjunto de dados multivariados (como o conjunto de dados Iris). Aplique PCA e visualize os primeiros dois componentes principais.

- Objetivo: Reduzir a dimensionalidade dos dados e explorar suas variáveis.

9. Distribuição Binomial

- Desafio: Simule uma situação em que você realiza 10 tentativas de um experimento com uma probabilidade de sucesso de 0,3. Calcule e visualize a distribuição binomial para 0 a 10 sucessos.

- Objetivo: Entender e aplicar a distribuição binomial.

10. Análise de Séries Temporais

- Desafio: Crie uma série temporal simulada de vendas mensais de um produto ao longo de dois anos. Plote os dados e analise a tendência.

- Objetivo: Aprender a manipular e analisar dados de séries temporais.