

# Banco de Questões para Prática de ETL e Data Visualization com Pandas

Aqui estão questões de **níveis básico, intermediário e avançado** para cada um dos três arquivos CSV (vendas, funcionários, produtos). Elas cobrem desafios comuns de ETL (Extração, Transformação e Carga) e incluem sugestões para visualizações de dados.

## 1. vendas.csv

### Básico

- Leia o arquivo `vendas.csv` em um DataFrame.
- Mostre as 5 primeiras linhas.
- Calcule o total de vendas (`preco_unitario` x `quantidade`) para cada linha e crie uma nova coluna chamada `total_venda`.
- Quantas vendas cada vendedor realizou?
- Qual produto mais aparece nas vendas?

### Intermediário

- Filtre apenas vendas feitas pela Joana e calcule o valor total vendido por ela.
- Agrupe as vendas por produto e mostre a soma total vendida de cada um.
- Crie uma tabela dinâmica que mostre a soma de vendas por vendedor e por produto.
- Gere um gráfico de barras mostrando o total vendido por cada vendedor.

### Avançado

- Junte o DataFrame de vendas com o de produtos (utilizando o campo `produto/nome_produto`) para obter a categoria de cada venda.
- Gere um relatório de vendas por categoria (quantidade, valor total).
- Crie um gráfico da evolução diária das vendas totais (soma de `total_venda` por data) utilizando um gráfico de linha.
- Identifique o produto mais rentável (maior soma de `total_venda`).
- Mostre, em um gráfico de setores (pizza), a participação de cada categoria no valor total vendido.

## 2. funcionarios.csv

### Básico

- Leia o arquivo `funcionarios.csv` em um DataFrame.
- Liste os funcionários do departamento "Vendas".
- Quantos funcionários existem em cada departamento?
- Qual o maior e o menor salário da empresa?

### Intermediário

- Calcule a média salarial por departamento.
- Adicione uma coluna "anos\_casa" calculando o tempo de empresa (em anos completos) de cada funcionário considerando a data atual.
- Quais funcionários possuem mais de 5 anos de empresa?
- Gere um gráfico de barras mostrando a quantidade de funcionários por departamento.

### Avançado

- Identifique o funcionário mais antigo e o mais novo (tempo de casa).
- Junte os dados de funcionários com vendas (pelo nome do vendedor) e calcule o valor total vendido por cada funcionário.
- Crie uma tabela mostrando, por departamento, o colaborador com o maior salário.
- Plote um gráfico de dispersão mostrando salário versus tempo de empresa para os funcionários de "Vendas".

### 3. produtos.csv

#### Básico

- Leia o arquivo `produtos.csv`.
- Liste todos os produtos da categoria "Vestuário".
- Qual produto tem o menor estoque?
- Quantos produtos diferentes existem em cada categoria?

#### Intermediário

- Calcule o valor total em estoque de cada produto ( $\text{estoque} \times \text{preço}$ ).
- Gere uma tabela com o valor total em estoque por categoria.
- Mostre todos os produtos com estoque inferior a 15 unidades.
- Crie um gráfico de barras mostrando o estoque por categoria.

#### Avançado

- Identifique quais produtos nunca foram vendidos, cruzando com o arquivo de vendas.
- Descubra o ticket médio dos produtos de cada categoria vendidos (valor médio de venda, cruzando com vendas).
- Plote um heatmap mostrando o estoque versus preço para todos os produtos de "Eletrônicos".
- Simule uma tabela de reposição, sugerindo reabastecimento para todos os produtos com estoque abaixo de 10 unidades.

## Dicas para Data Visualization

- Use `matplotlib` ou `seaborn` para gráficos de barra, linha, dispersão e pizza.
- Para tabelas dinâmicas (pivot table), utilize `pd.pivot_table`.
- Sempre comece validando os dados com `.head()`, `.info()` e `.describe()`.