**Flávia Vitória de Queiroz**

**Projeto de Análise e Visualização de Dados de Vendas Online**

**2025**

**Análise e Visualização de Dados de Vendas Online**

Este projeto tem como objetivo simular um cenário real de negócios, no qual é necessário coletar, limpar, analisar e apresentar dados de vendas de uma loja online de roupas, proporcionando insights valiosos para a tomada de decisões.

**Objetivo Geral**

O objetivo principal é construir um fluxo de trabalho que:

1. **Gerar dados de vendas**.
2. **Limpeza e preparação** dos dados para análise.
3. **Análise dos dados** para extrair insights valiosos.
4. **Visualização dos resultados** de forma clara.

**Súmario**

Etapas de Desenvolvimento ...............................................................................4

Simulação de Dados..................................................................................4

Limpeza e Preparação de Dados...............................................................6

Análise de Dados.......................................................................................9

Visualização dos Dados...........................................................................10

Conclusão..........................................................................................................13

Anexo.................................................................................................................14

**ETAPAS DE DESENVOLVIMENTO**

1. Coleta/Simulação de Dados (Python)

**Ferramenta:** Python (random, datetime)

Como não havia dados reais disponíveis para o protótipo, utilizou-se um *loop for i in range* para automatizar a criação de uma lista simulando dados típicos de uma loja de roupas, incluindo data da venda, ID da venda, produto vendido, quantidade e valor do produto (figura 1).

**Figura 1:** Código de Automatização da Criação da Lista

Tela de computador com texto preto sobre fundo branco

O conteúdo gerado por IA pode estar incorreto.

**Fonte:** Do autor.

Para realizar a seleção aleatória de itens padrão dentro da categoria de produtos, foi criada a lista **“nomes\_produtos”**, conforme ilustrado na Figura 2.

**Figuras 2:** Lista do Nome dos Produtos

Texto

O conteúdo gerado por IA pode estar incorreto.

**Fonte:** Do autor.

Para automatizar a inserção de datas dentro do loop for i in range, foi criada a função **“gerar\_data\_aleatoria”,** que gera datas aleatórias com até 6 meses de diferença em relação à data atual.

**Figura 3:** Código da def **“gerar\_data\_aleatoria”**

Texto

O conteúdo gerado por IA pode estar incorreto.

**Fonte:** Do autor.

Por fim, para se ter uma visualização limpa e precisa das informações a saída foi formatada, como pode-se visualizar-se na figura 4. E pode-se observar a saída do código na figura 5.

**Figura 4:** Código da Formatação da Saída

Texto

O conteúdo gerado por IA pode estar incorreto.**Fonte:** Do autor.

]

**Figura 5:** Saída Formatada

Texto

O conteúdo gerado por IA pode estar incorreto.

**Fonte:** Do autor.

Para garantir que a saída fosse salva e os dados pudessem ser reutilizados posteriormente, foi criado um arquivo no formato CSV.

**Figura 6:** Salvando os Dados em um Arquivo .csv

Tela de computador com texto preto sobre fundo branco

O conteúdo gerado por IA pode estar incorreto.

**Fonte:** Do autor.

1. Limpeza e Preparação de Dados

**Ferramentas:** Python (Pandas).

Primeiramente, importei a biblioteca pandas, utilizando o as pd. O arquivo foi aberto para leitura, e em seguida realizou-se a limpeza dos dados, removendo registros duplicados. Depois, as colunas que representavam datas foram convertidas para o formato datetime, enquanto as colunas numéricas foram convertidas para os tipos inteiros (int) ou de ponto flutuante (float), conforme ilustrado na figura 7.

**Figura 7:** Limpeza e Conversão dos Dados

Texto

O conteúdo gerado por IA pode estar incorreto.

**Fonte:** Do autor.

Em seguida, foram criadas colunas adicionais para aprimorar a visualização dos dados, com o objetivo de obter resultados mais precisos (Figura 8 e 9). Para refinar ainda mais a análise, os produtos foram classificados em categorias: roupas, calçados e outros (Figura 10).

**Figura 8:** Colunas Alternativas 1

Texto

O conteúdo gerado por IA pode estar incorreto.

**Fonte:** Do autor.

**Figura 9:** Colunas Alternativas 2

Texto

O conteúdo gerado por IA pode estar incorreto.

**Fonte:** Do autor.

**Figura 10:** Código de Categorização

Texto

O conteúdo gerado por IA pode estar incorreto.

**Fonte:** Do autor.

Em seguidas, as novas colunas foram salvas em .csv para serem utilizadas na visualização de dados.

**Figura 11:** Colunas Salvas em .csv

Tela de computador com texto preto sobre fundo branco

O conteúdo gerado por IA pode estar incorreto.

**Fonte:** Do autor.

1. Análise dos dados

Nessa etapa do projeto, o objetivo foi responder a perguntas básicas de análise de dados, como: *“Qual produto teve o maior número de vendas nos últimos seis meses?”* e *“Qual categoria obteve mais vendas no mesmo período?”*.

Para alcançar essas respostas, foi necessário utilizar as bibliotecas Matplotlib, Pandas e Seaborn.

Com o uso dessas ferramentas, foram desenvolvidos os códigos responsáveis pela geração dos gráficos de quantidade de vendas por produto (Figura 12), valor médio dos produtos (Figura 13) e valor dos produtos por categoria (Figura 14).

**Figura 12:** Código Para a Criação do Gráfico Quantidade de Vendas por Produto

A computer screen with text

AI-generated content may be incorrect.

**Fonte:** Do autor.

**Figura 13:** Código Para a Criação do Gráfico Valor por Produto Durante os Últimos 6 Meses

A screen shot of a computer code

AI-generated content may be incorrect.

**Fonte:** Do autor.

**Figura 14:** Código Para a Criação do Gráfico Valor dos Produtos por Categoria

A computer screen with colorful text

AI-generated content may be incorrect.

**Fonte:** Do autor.

1. Visualização dos resultados

Os resultados foram visualizados por meio dos gráficos obtidos a partir dos códigos apresentados nas Figuras 12, 13 e 14.

O Gráfico 1 representa a quantidade de vendas por produto, em que o eixo X indica a quantidade vendida e o eixo Y apresenta os diferentes produtos. As barras azuis correspondem aos valores registrados na tabela original, enquanto as barras cinzas indicam uma possível taxa de erro.

**Gráfico 1:** Quantidade de Vendas por Produto

A graph of a bar chart

AI-generated content may be incorrect.

**Fonte:** Do autor.

O Gráfico 1 representa a quantidade de vendas por produto, em que o eixo X indica a quantidade e o eixo Y apresenta os diferentes produtos.  
As barras azuis correspondem aos valores registrados na tabela inicial, enquanto as barras cinzas indicam uma possível taxa de erro.  
Observa-se que as meias foram os produtos mais vendidos da loja, enquanto as calças apresentaram o menor número de vendas. Além disso, os itens saia e meia registraram as maiores taxas de erro.

A Figura 16 ilustra o valor médio de cada produto vendido nos últimos seis meses, sendo o eixo X correspondente ao preço e o eixo Y aos produtos.

**Gráfico 2:** Valor por Produto Durante os Últimos 6 Meses

A graph of blue and black bars

AI-generated content may be incorrect.

**Fonte:** Do autor.

O Gráfico 2 mostra que os produtos mais caros são os bonés e os blazers. Entretanto, os blazers não registraram altos índices de vendas, enquanto os bonés apresentaram um volume médio de vendas.

A categoria com o maior número de vendas foi a de calçados, ocupando o terceiro, quarto e quinto lugar entre os itens com maiores valores (gráfico 3).

Por fim, a figura 17 simboliza o último gráfico, valor dos produtos por categoria. No qual o eixo x corresponde ao preço e o y a categoria.

**Gráfico 3:** Valor dos Produtos vendidos por Categoria

A graph with blue and black bars

AI-generated content may be incorrect.

**Fonte:** Do autor.

Dessa forma, podemos concluir que, para aumentar o lucro, é necessário investir em estratégias que incentivem a venda de produtos com maior valor agregado, especialmente aqueles que já apresentam boa aceitação pelo público, como os calçados e bonés.

**CONCLUSÃO**

Em conclusão, o projeto permitiu compreender o comportamento das vendas da loja nos últimos seis meses, identificando os produtos mais e menos vendidos, bem como as categorias com maior impacto no faturamento.  
A análise revelou que as meias foram os produtos com maior volume de vendas, enquanto as calças apresentaram o menor desempenho. Além disso, foi possível observar que os bonés e blazers estão entre os itens de maior valor médio, mas com diferenças significativas em seu volume de vendas.

O uso das bibliotecas Pandas, Matplotlib e Seaborn foi essencial para organizar, tratar e visualizar os dados de forma eficiente, permitindo a geração de insights relevantes para a tomada de decisões.

Dessa forma, conclui-se que, para aumentar o lucro e otimizar o desempenho das vendas, recomenda-se investir em estratégias de marketing voltadas a produtos de maior valor agregado, além de avaliar oportunidades de melhoria nos itens com baixo volume de vendas.

**ANEXO**

Os códigos para a elaboração desse projeto podendo ser acessados por meio do seguinte link: *https://github.com/FlaviaVitoriadeQueiroz/-Prototipagem-An-lise-e-Visualiza-o-de-Dados-de-Vendas-Online.*