Solução do teste

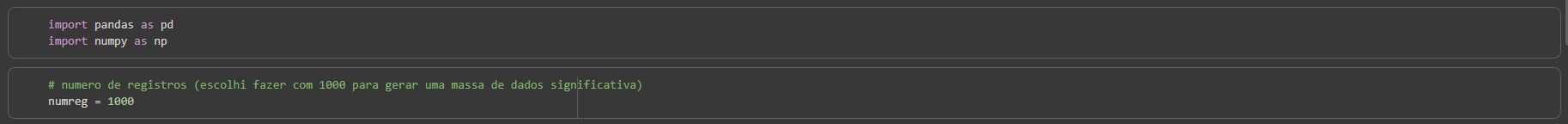
Parte1

Escolhi Python para fazer a parte 1 do teste, por ter maior experiencia com essa linguagem de programação, utilizei o google Colab para desenvolver o código e segui os seguintes passos.

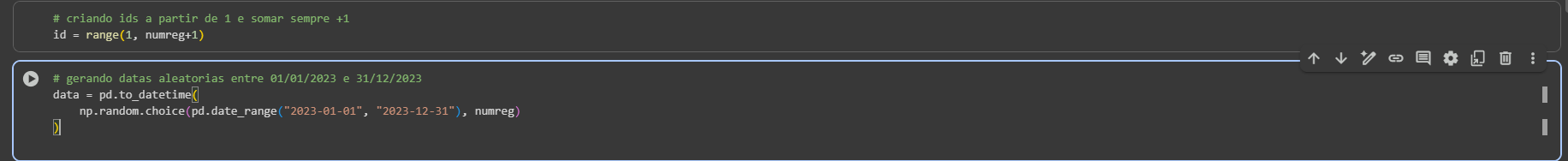
(Em vermelho nesse documento incluo as orientações recebidas para a realização do mesmo)

#### Limpeza e Análise de Dados de Vendas

Crie um script para simular um dataset de vendas com pelo menos 50 registros, contendo as colunas: ID, Data, Produto, Categoria, Quantidade, Preço. O período dos dados deve ser de 01/01/2023 a 31/12/2023.



Importei as bibliotecas necessárias e para gerar uma quantidade de dados maior e mais relevante, escolhi trabalhar com 1000 registros.



Criei as colunas id e data aleatória, obedecendo a regra de que o período deveria ser entre 01/01/2023 a 31/12/2023.

Texto

O conteúdo gerado por IA pode estar incorreto.

Criei uma lista de produtos e atribui a eles uma categoria e um preço.

Interface gráfica do usuário, Aplicativo

O conteúdo gerado por IA pode estar incorreto.

Garanti que a escolha de produtos fosse aleatória para preencher a base de dados e a quantidade de cada venda também fosse aletaória, podendo ser ente 1 e 50.

Associe a categoria e o preço aos produtos.

Por fim gerei a base de dados.

Interface gráfica do usuário, Texto, Aplicativo

O conteúdo gerado por IA pode estar incorreto.

Salvei como data1.csv essa base gerada, mas notei que não poderia fazer nenhum tratamento nela, porque ela não possuia valores nulos ou linhas duplicadas, então simulei alguns desses problemas nas colunas preço e produto inclui alguns valores NaN e aleatoriamente gerei 10 linhas duplicadas e salvei o arquivo como dataasertratado.csv. Para poder cumprir com as instruções do documento de:

* + Realize a limpeza dos dados, incluindo:
    - Tratamento de valores faltantes
    - Remoção de duplicatas
    - Conversão de tipos de dados, se necessário

Texto

O conteúdo gerado por IA pode estar incorreto.

Interface gráfica do usuário, Aplicativo

O conteúdo gerado por IA pode estar incorreto.

A partir dai apaguei as linhas duplicatas, e removi também as linhas que haviam registros nulos (achei que remover os nulos nesse caso seria a melhor abordagem, por ser uma quantidade de dados pequena com nulos, se fossem mais registros talvez tivesse tido uma outra abordagem, como preenche-los com a média, por exemplo)

Por fim, salvei esse dataset limpo e tratado como data\_clean.csv, conforme a orientação do documento.

* + Salve o dataset limpo em um arquivo data\_clean.csv.

Interface gráfica do usuário, Aplicativo

O conteúdo gerado por IA pode estar incorreto.

Carreguei a base limpa e tratada no python. Para iniciar as seguintes etapas solicitadas no documento

* + Utilizando o dataset limpo, realize as seguintes análises:
    - Calcule o total de vendas (Quantidade \* Preço) por produto.
    - Identifique o produto com o maior número de vendas totais.

Interface gráfica do usuário

O conteúdo gerado por IA pode estar incorreto.

Criei uma coluna Total de Vendas que multiplica a quantidade de produtos pelo Preco.

Interface gráfica do usuário, Texto, Aplicativo

O conteúdo gerado por IA pode estar incorreto.

Salvei novamente a base de dados, agora incluindo a coluna Total de Vendas como data\_clean\_comtotaldevendas.csv

Agrupei os produtos e somei o Total de Vendas de cada um, ordenei em ordem decrescente para sabermos quais foram os produtos que mais faturaram em vendas.

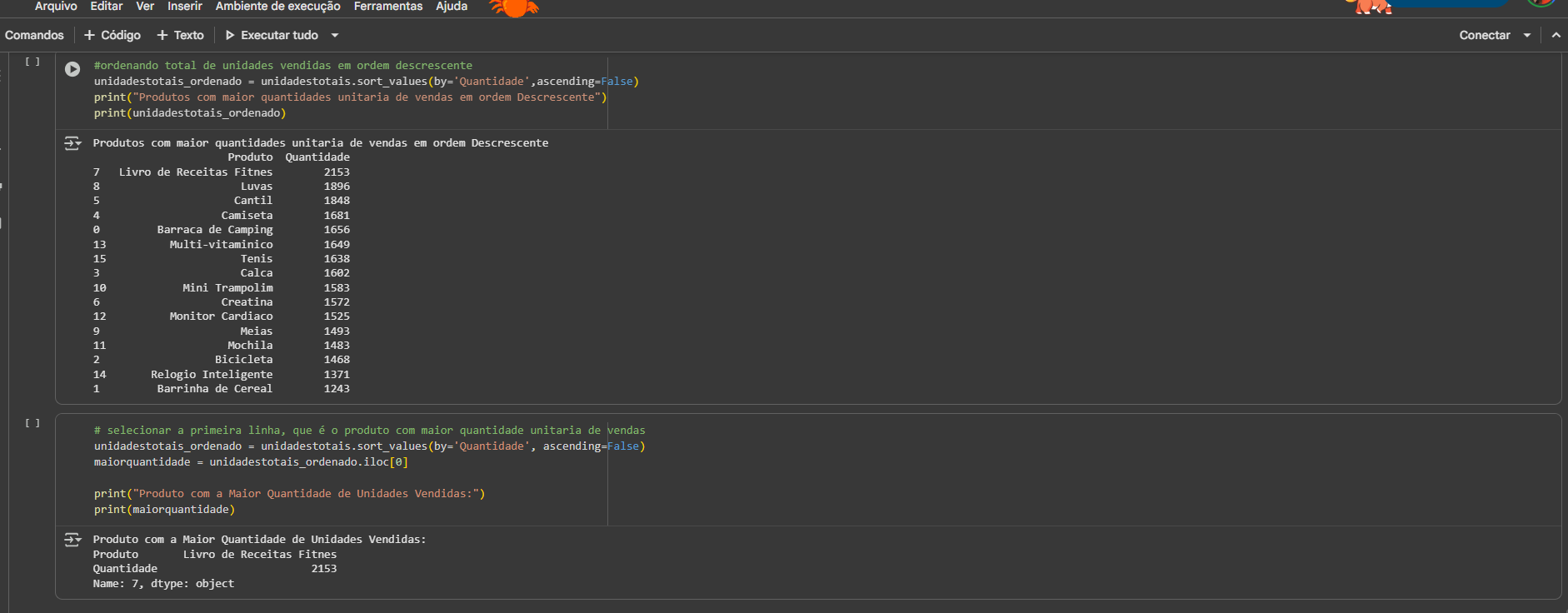
Texto

O conteúdo gerado por IA pode estar incorreto.

Selecionei apenas a primeira linha da lista, para evidenciar o produto mais lucrativo, que como podemos observar foram as bicicletas.

Depois fiquei curiosa para saber qual produto é o mais vendido em quantidade de unidades.

Agrupei e somei a quantidade de cada produto vendido e posteriormente ordenei por ordem descrescente a lista de quantidade de produtos mais vendidos.



Por fim, selecionei apenas a primeira linha que é o produto mais vendido e com isso descobrimos que o livro de receitas é o lider de vendas em quantidade.

O código dessa primeira etapa está salvo como teste\_quod.ipynb e teste\_quod.py.

A próxima etapa solicitada era:

#### Análise Exploratória de Dados de Vendas

* + Com o mesmo dataset simulado da tarefa anterior, crie visualizações/gráficos e com eles:
    - Crie um gráfico de linha mostrando a tendência de vendas ao longo do tempo, analisando os dados mensalmente.

Uma imagem contendo Interface gráfica do usuário

O conteúdo gerado por IA pode estar incorreto.

Importei as bibliotecas necessárias em um novo notebook, li o dataset com a coluna extra Total de Vendas.

Interface gráfica do usuário, Aplicativo

O conteúdo gerado por IA pode estar incorreto.

Tratei a coluna de data para conseguir separar por meses e fazer a análise temporal, somando as vendas de cada mês.

Gráfico, Gráfico de linhas

O conteúdo gerado por IA pode estar incorreto.

Criei um gráfico de linhas, no qual podemos visualizar a quantidade de vendas por mês.

Gráfico

O conteúdo gerado por IA pode estar incorreto.

Criei um gráfico de barras de quantidade de cada produto vendido no mês.

Gráfico, Gráfico de barras

O conteúdo gerado por IA pode estar incorreto.

Criei um gráfico de Quantidade de produtos mais vendidos no ano.

Interface gráfica do usuário

O conteúdo gerado por IA pode estar incorreto.

Por fim, um gráfico de quais os produtos mais lucrativos durante o ano.

* + - Identifique e descreva pelo menos dois padrões ou insights interessantes que você observou nos dados.

**Insights que pude tirar desses gráficos:**

Evolução das vendas ao longo do ano  
O mês de setembro foi o grande destaque, apresentando um pico acentuado de vendas. Esse aumento pode estar relacionado à chegada da primavera, à realização de promoções ou até ao lançamento de produtos aguardados.  
Por outro lado, os meses de fevereiro e março tiveram desempenho fraco, o que sugere a possibilidade de reduzir o estoque nesse período para diminuir custos de armazenamento.  
Julho também registrou queda, possivelmente por ser um mês mais frio e de férias, fatores que podem impactar negativamente as vendas, principalmente considerando que o negócio é focado em produtos esportivos.

Produtos mais vendidos em quantidade x Produtos mais lucrativos  
Alguns itens de menor valor unitário, como o livro de receitas e as luvas, estão entre os mais vendidos em quantidade, talvez por serem mais acessíveis e até utilizados como opções de presente.  
Por outro lado, as bicicletas se destacam como nosso principal produto em termos de lucratividade, sendo o verdadeiro carro-chefe da operação. Os relógios inteligentes também apresentam participação expressiva no valor total de vendas.  
A análise por mês mostra que as barracas de camping têm picos em janeiro, abril e julho, o que evidencia sazonalidade e aponta a necessidade de reforçar o estoque nesses períodos. Já as bicicletas mantêm um bom desempenho durante todo o ano, confirmando seu papel como produto estratégico para o negócio.

Solução do teste

Parte2 SQL

Para as consultas em SQL utilizei o MySQL Workbench, seguem as consultas efetuadas.

Com base nas tabelas geradas no exercício anterior, escreva consultas SQL para:

* Listar o nome do produto, categoria e a soma total de vendas (Quantidade \* Preço) para cada produto. Ordene o resultado pelo valor total de vendas em ordem decrescente.

Interface gráfica do usuário, Texto, Aplicativo, Email

O conteúdo gerado por IA pode estar incorreto.

Utilizando a tabela data\_clean.csv selecionei o produto e a categoria, multipliquei a quantidade pelo preço e agrupei pelos produtos, utilizei o format para deixar apenas duas casas decimais.

* Identificar os produtos que venderam menos no mês de junho de 2024.

Interface gráfica do usuário, Texto, Aplicativo, Email

O conteúdo gerado por IA pode estar incorreto.

No mês de Junho de 2024 não temos dados, pois a base de dados foi gerada de 01/01/2023 a 31/12/2023.

Mas fiz a consulta para os dados de Junho de 2023.

Selecionei o produto, somei o valor total de vendas formatando para duas casas decimais, filtrei o período do mês 06 de 2023, agrupei por produtos e organizei do menor para o maior valor mostrando os 10 produtos com menor valor de vendas no mês em questão.

**Salve suas consultas e a explicação da lógica em um arquivo chamado consultas\_sql.sql.**

As consultas estão salvas como consultas\_sql.sql.

**Análise Dashboard**

Utilizei a base de dados gerada para fazer um dashboard no Looker Studio, como atividade adicional, para uma análise mais aprofundada, o link desse dashboard para a consulta é:

<https://lookerstudio.google.com/reporting/03fd562f-f4d0-4d37-baf4-b1289cad6972>

O dashboard é interativo e permite a filtragem por Mês e Produto, o que será crucial para análises mais aprofundadas, os gráficos tem interação uns com os outros o que facilita visualizações especificas.

Aqui estão alguns prints dele e insights extraídos:

Interface gráfica do usuário, Aplicativo

O conteúdo gerado por IA pode estar incorreto.

**Visão Geral das Vendas Anuais**

* Valor Total de Vendas: O faturamento anual atingiu R$ 9.038.574,52.
* Total de Itens Vendidos: Foram vendidos 25.861 itens ao longo do ano.

**Destaques por Categoria**

As vendas em três categorias representaram mais de 78% do faturamento total, com destaque para:

1. Equipamento Esportivo: Categoria líder, respondendo por 45,9% do total de vendas.
2. Eletrônicos: Segunda categoria mais representativa, com 19,6% da participação.
3. Lazer: Terceira categoria mais importante, totalizando 12,8% das vendas.

**Principais Produtos Vendidos**

A análise do Total de Vendas por Produto demonstra uma forte concentração em dois itens de alto valor:

* Bicicleta: Produto que domina o volume de vendas.
* Relógio Inteligente: Segundo produto mais vendido em termos de valor

Outros produtos de destaque que contribuíram para as vendas incluem Barraca de Camping, Tênis, Calça, Creatina e Livro de Receitas.

**Recomendações**

Manter o foco e investir em estoques e marketing para Equipamento Esportivo e Eletrônicos, já que são os pilares de faturamento.

Gráfico, Gráfico de linhas

O conteúdo gerado por IA pode estar incorreto.

**Análise da Evolução Mensal (Tendências e Sazonalidade)**

A análise da evolução mensal revela um ano com forte sazonalidade e um desempenho de vendas notável no segundo semestre:

* Queda Inicial: O ano começou com uma queda, saindo de aproximadamente R$ 678 mil em Janeiro para o ponto mais baixo do ano em Fevereiro.
* Recuperação e Estabilidade: As vendas se recuperaram continuamente de Março a Junho, atingindo um patamar estável.
* Pico de Vendas: O desempenho atingiu seu ponto máximo em Setembro, ultrapassando a marca de R$ 1,1 milhão (o pico mais alto do ano).
* Performance Final: Após o pico de Setembro, houve uma desaceleração em Outubro e Novembro, mas o ano terminou com uma forte recuperação em Dezembro.

Conclusão: O terceiro trimestre (Julho-Setembro) foi o período de maior crescimento e rentabilidade. O pico em Setembro e a forte recuperação em Dezembro sugerem a importância das vendas sazonais de final de ano e/ou promoções específicas.

Uma imagem contendo Interface gráfica do usuário

O conteúdo gerado por IA pode estar incorreto.

**Desempenho Detalhado por Produto: Volume vs. Receita**

Esta análise compara o volume de vendas (Unidades Vendidas - barras roxas) com o valor total de vendas (Receita - linha branca) para os principais produtos, revelando diferenças significativas:

1. Produtos de Alto Valor

Estes são os produtos que geram a maior receita, mesmo que seu volume de unidades vendidas não seja o maior:

* Bicicleta: É, de longe, o produto de maior receita, apesar de ter um volume de unidades vendidas inferior ao segundo colocado. Isso indica um ticket médio muito elevado.
* Relógio Inteligente: Segue o mesmo padrão, sendo o segundo em receita, mas com um volume de unidades abaixo da média dos outros produtos.

2. Produtos de Alto Volume

Estes produtos representam a maioria das transações, mas contribuem menos individualmente para o faturamento total, indicando um ticket médio baixo:

* Livro de Receitas: É o produto com o maior volume de unidades vendidas mas sua receita é muito baixa. Isso sugere que é um produto de preço baixo.
* Luvas, Camiseta, Cantil, Meias, Barra de Cereal, etc.: A grande maioria dos itens, como Luvas, Camiseta, Creatina e Tênis, possuem volumes de unidades altas, mas suas receitas individuais são modestas. Eles são cruciais para o volume de transações, mas não para o valor total.

Conclusão e Implicações

* O faturamento da empresa é dominado por poucos produtos de alto valor (Bicicleta e Relógio Inteligente), que representam a maior parte da receita total.
* A maioria dos itens vendidos é de baixo valor (ex: Livro de Receitas, Tênis, Roupas), sendo importante para a base de clientes e o fluxo de caixa, mas menos impactante no resultado final.

Gráfico

O conteúdo gerado por IA pode estar incorreto.

**Análise Estratégica de Produtos: Preço vs. Volume (Gráfico de Bolhas)**

O gráfico de bolhas separa claramente os produtos em quatro grupos estratégicos, confirmando o domínio do faturamento por itens de alto valor:

1. Pilares de Receita

Estes produtos têm um preço unitário elevado, e o tamanho de suas bolhas (Valor Total de Vendas) é significativamente maior, confirmando sua importância para o faturamento:

* Bicicleta (azul escuro, à esquerda): Tem o Preço Unitário mais alto e o maior Valor Total de Vendas, mas sua quantidade vendida é moderada. É o principal gerador de receita.
* Relógio Inteligente (vermelho, à esquerda): Possui o segundo maior preço unitário e uma bolha grande, confirmando sua posição como o segundo principal gerador de receita.

2. Motores de Volume (Alto Volume, Baixo Preço Unitário)

Estes produtos têm um preço unitário baixo e são vendidos em grandes quantidades. Contribuem para o volume de transações, mas menos para o faturamento:

* Mini Trampolim, Monitor Cardíaco, Tênis, Calça, Luvas, Multi-vitamínico: Todos concentrados no canto superior esquerdo (baixa faixa de preço e alta quantidade).

3. Oportunidades de Margem Média

Estes produtos se destacam por um equilíbrio entre preço e quantidade.

* Barraca de Camping (amarelo, no centro): Tem um preço unitário intermediário R$ 680,00 e uma quantidade vendida alta, resultando em um Valor Total de Vendas significativo, o que o torna um produto de margem eficiente.
* Cantil (azul claro, à direita): Apesar de ter um preço unitário menor, o volume de vendas altíssimo e o tamanho da bolha indicam que ele é um item crucial, provavelmente com um ótimo custo-benefício percebido pelo cliente.

Síntese da Estratégia de Produtos

A empresa possui um modelo de vendas misto: o faturamento é impulsionado por itens premium (Bicicleta e Relógio), enquanto o volume de clientes é construído em torno de itens de baixo e médio preço (Acessórios, Vestuário e Suplementos).

Gráfico

O conteúdo gerado por IA pode estar incorreto.

**Análise da Participação Mensal de Produtos**

O gráfico de barras empilhadas revela como a participação dos produtos flutua mensalmente:

1. Sazonalidade dos Principais Produtos (Bicicleta e Relógio Inteligente)

* Bicicleta: A participação da Bicicleta é consistentemente alta, dominando o mix em quase todos os meses, especialmente no início e no final do ano. Ela representa consistentemente 40% a 55% do total de vendas mensais, confirmando ser o principal pilar de faturamento.
* Relógio Inteligente: O Relógio Inteligente mantém uma participação forte, geralmente entre 15% e 25%.

2. Destaques de Produtos Secundários por Período

Observa-se que produtos específicos ganham maior relevância em certos trimestres:

* 1º Trimestre : O produto Barraca de Camping demonstra uma participação significativa, sugerindo compras relacionadas a férias ou início de ano.
* 2º Trimestre : Produtos como Mini Trampolim e Monitor Cardíaco aumentam sua fatia, possivelmente devido a promoções de primavera ou aumento de atividades internas.
* 2º Semestre : O produto Calça e outros itens de vestuário e acessórios Luvas, Mochila, por exemplo, têm uma participação notável e estável, indicando uma demanda contínua ou relacionada a mudanças sazonais, com a chegada do final do ano, podemos supor que esses produtos são muito vendidos para serem presenteados.

### **Conclusão**

O desempenho de vendas em 2023 foi sólido, fortemente concentrado nas categorias de Equipamento Esportivo e Eletrônicos. A sustentação do faturamento depende criticamente da performance constante da Bicicleta e do Relógio Inteligente. A gestão de estoque e marketing deve ser otimizada para maximizar as vendas nos picos de Setembro e Dezembro e buscar estratégias de recuperação mais rápidas nos meses de baixa (Fevereiro, Julho).

### Parte 3: Interpretação de Resultados

* Com base nas análises realizadas nas Partes 1 e 2, escreva um breve relatório (máximo de 300 palavras) resumindo os principais insights encontrados e sugerindo possíveis ações baseadas nesses insights.

**Salve este relatório como relatorio\_insights.md ou relatorio\_insights.pdf.**

O relatório foi salvo com o nome solicitado.