

Unifavip Wyden

**Relatório Técnico de Desenvolvimento rápido de Aplicações
Python:Painel Interativo de Perfil de Compra de Clientes**

Nomes dos integrantes: Micael José de Lima - 202303811969;
Enrico Cassimiro Vertano Germâncio - 202303505728; Fernando
Torres da Silva - 202302229999

Nome do professor : Kayo Henrique

2024
Caruaru/Pernambuco

Sumário

1.	Introdução.....	3
2.	Relevância.....	3
3.	Atributos e Descrição.....	4
4.	Contextualizando o problema.....	5
5.	Etapas do Projeto.....	5
	5.1 Pré-processamento.....	5
	5.1.Painel Interativo.....	6
6.	Conclusão.....	8
7.	Referências e link úteis.....	9

Projeto: Painel Interativo de Perfil de Compra de Clientes

1. Introdução

O seguinte projeto trata-se do desenvolvimento de um painel interativo construído em linguagem Python com as bibliotecas Plotly e Streamlit para visualização de dados e as bibliotecas pandas e Numpy para manipulação de dados. O dataset utilizado neste projeto foi baixado na página referenciada nos links uteis ao final deste documento e tem o registro de 3900 clientes com 18 colunas de uma loja de roupas dos Estados Unidos, que nesse projeto denominamos a loja com o nome “Clothing Store” e criamos uma logo para representá-la e fazer parte do painel interativo, com as visualizações, filtros e cartões, ambos, com o intuito de facilitar a tomada de decisões informadas e entregar um projeto visualmente elegante. O projeto seguiu etapas de pré-processamento de dados em um arquivo jupyter notebook e construção do painel interativo com duas páginas em arquivos py, além de discussões e planejamento sobre como entregar um projeto útil para os tomadores de decisão.



2. Relevância

Esse projeto é relevante academicamente por permitir que apliquemos os conhecimentos teóricos e práticos adquiridos dentro e fora da sala de aula em um projeto que se assemelha a resolução de problemas reais. Se alinhando com o objetivo da formação acadêmica em preparar o aluno para o mercado de trabalho, assim

desenvolvendo soft skills importantes como trabalho em equipe, cumprimento de prazos, comunicação, pensamento analítico, raciocínio lógico, negócios, dentre várias outras habilidades. Além de hard skills como python usado para desenvolver o projeto e github utilizado para versionamento do código e armazenamento na nuvem. Todos esses, valiosos para se trabalhar no mercado de tecnologia e conseguir demonstrar isso é importante, por isso projetos como esse são fundamentais, ainda mais para iniciantes que estão na universidade atrás da sua primeira oportunidade na área e precisam demonstrar que conseguem lidar com o ambiente dinâmico do mercado de trabalho.

3. Atributos e Descrição

Atributos	Descrição
Customer ID	Identificador exclusivo para cada cliente
Age	Idade do cliente
Gender	Gênero do cliente (Masculino/Feminino)
Item Purchased	O item comprado pelo cliente
Category	Categoria do item adquirido
Purchase Amount (USD)	O valor da compra em USD
Location	Local onde a compra foi realizada
Size	Tamanho do item adquirido
Color	Cor do item adquirido
Season	Temporada durante a qual a compra foi feita
Review Rating	Classificação dada pelo cliente para o item comprado
Subscription Status	Indica se o cliente possui uma assinatura (Sim/Não)
Shipping Type	Tipo de envio escolhido pelo cliente
Discount Applied	Indica se foi aplicado desconto na compra (Sim/Não)

Atributos	Descrição
Promo Code Used	Indica se um código promocional foi usado para a compra(Sim/Não)
Previous Purchases	A contagem total de transações concluídas pelo cliente na loja, excluindo a transação em andamento
Payment Method	Método de pagamento preferido do cliente
Frequency of Purchases	Frequência com que o cliente realiza compras (ex: Semanal, Quinzenal, Mensal).

4. Contextualizando o problema

A escolha do banco de dados utilizado neste projeto se deu pelo fato de conter dados com valores simples de fácil uso para manipulação e análise de dados, fundamental para que todos os integrantes do grupo consigam participar independente do nível de conhecimento em python e também porque dados de compras de cliente são um dos principais exemplos encontrados no mercado, mesmo com outras ferramentas, como excel ou power bi.

Em muitos negócios a falta de governança de dados e tomada de decisão com base em “achismos” é um problema frequente que pode acabar impactando negativamente, com isso o nosso projeto, apesar de fictício, vem para resolver esses problemas. Para isso passamos pela etapa de pré-processamento para garantir que os dados estejam no formato correto, como dizia o renomado físico teórico Richard Feynman: “Nós temos muita informação e muito menos compreensão”, ou seja, dados sem qualidade resultam em análises sem valor. E após os dados estarem adequadamente processados, passamos para a etapa de análise de dados com a construção de uma painel interativo, com visualizações simples para facilitar a tomada de decisão, vale refletir, que, em dados atualmente falasse muito de inteligência artificial como uma forma revolucionária de se fazer as coisas, mas um simples painel ou dashboard como o deste projeto pode ajudar muito pequenos negócios que não tem estrutura e orçamento suficiente para construção de modelos de machine learning, como diz a seguinte metáfora: “matar uma formiga com um canhão”, ou seja às vezes o simples é a melhor opção em vez de demandar esforço excessivo ou desproporcional para resolver um problema.

5. Etapas do Projeto:

5.1 Pré-Processamento:

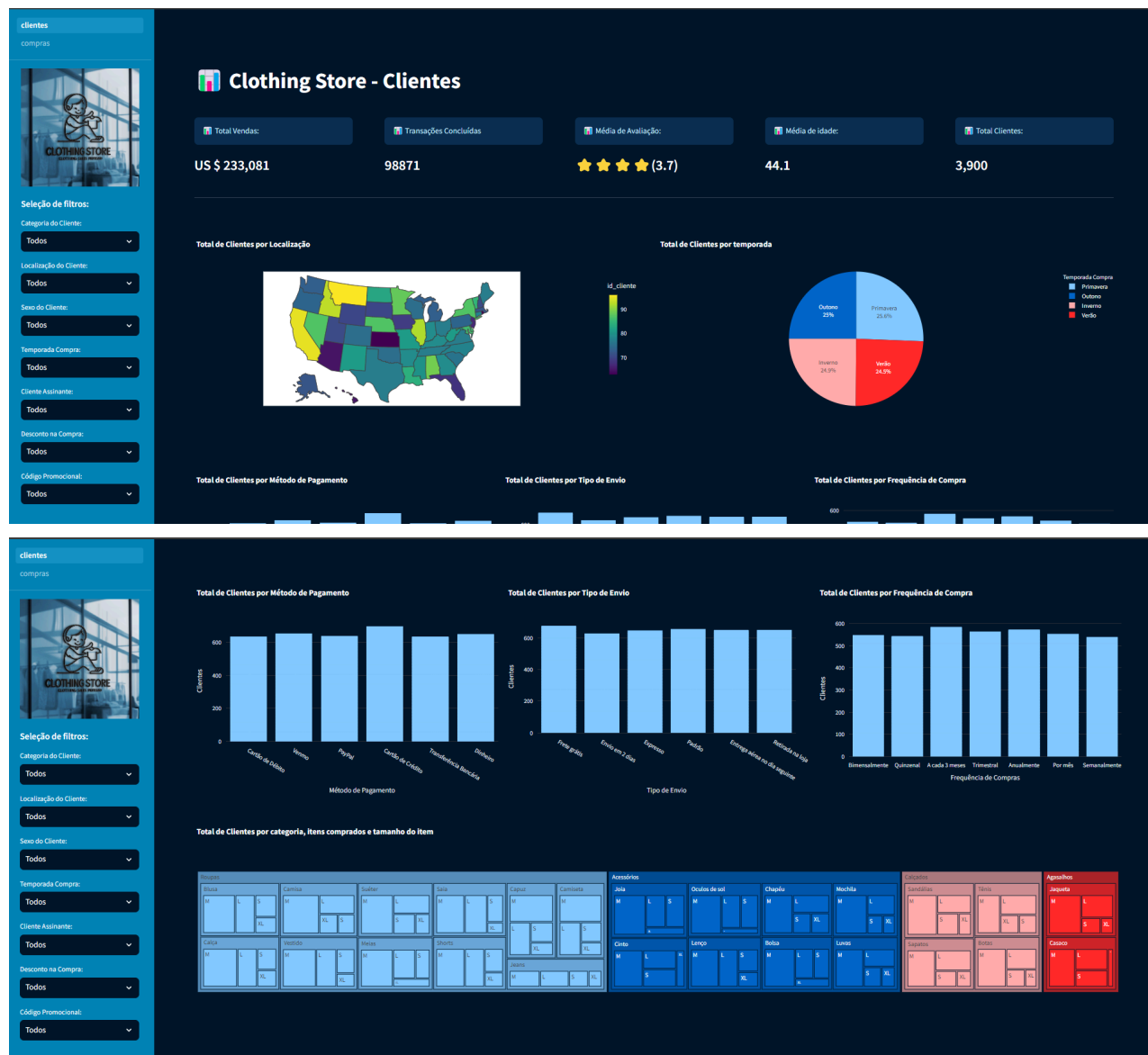
Na etapa de pré-processamento começou verificando a estrutura do projeto, valores nulos e duplicatas e foi notado que os dados estavam todos em um ótimo formato e não era necessário fazer nenhuma modificação neles, no entanto, por motivos didáticos e pela fácil interpretação dos dados pelo público e pelos integrantes do grupo, foi decidido substituir os valores dos dados de língua inglesa por sua devida tradução em língua portuguesa, com a renomeação de colunas e valores categóricos. E também foi criada uma coluna a mais com base na coluna de localização dos clientes, com o código dos estados americanos para poderem ser reconhecidos pelo gráfico do mapa e a etapa de pré-processamento chega ao fim ao salvar o arquivo agora denominado como df com as alterações feitas e que será o caminho para nossas análises na etapa da construção do painel interativo.

5.2 Painel Interativo:

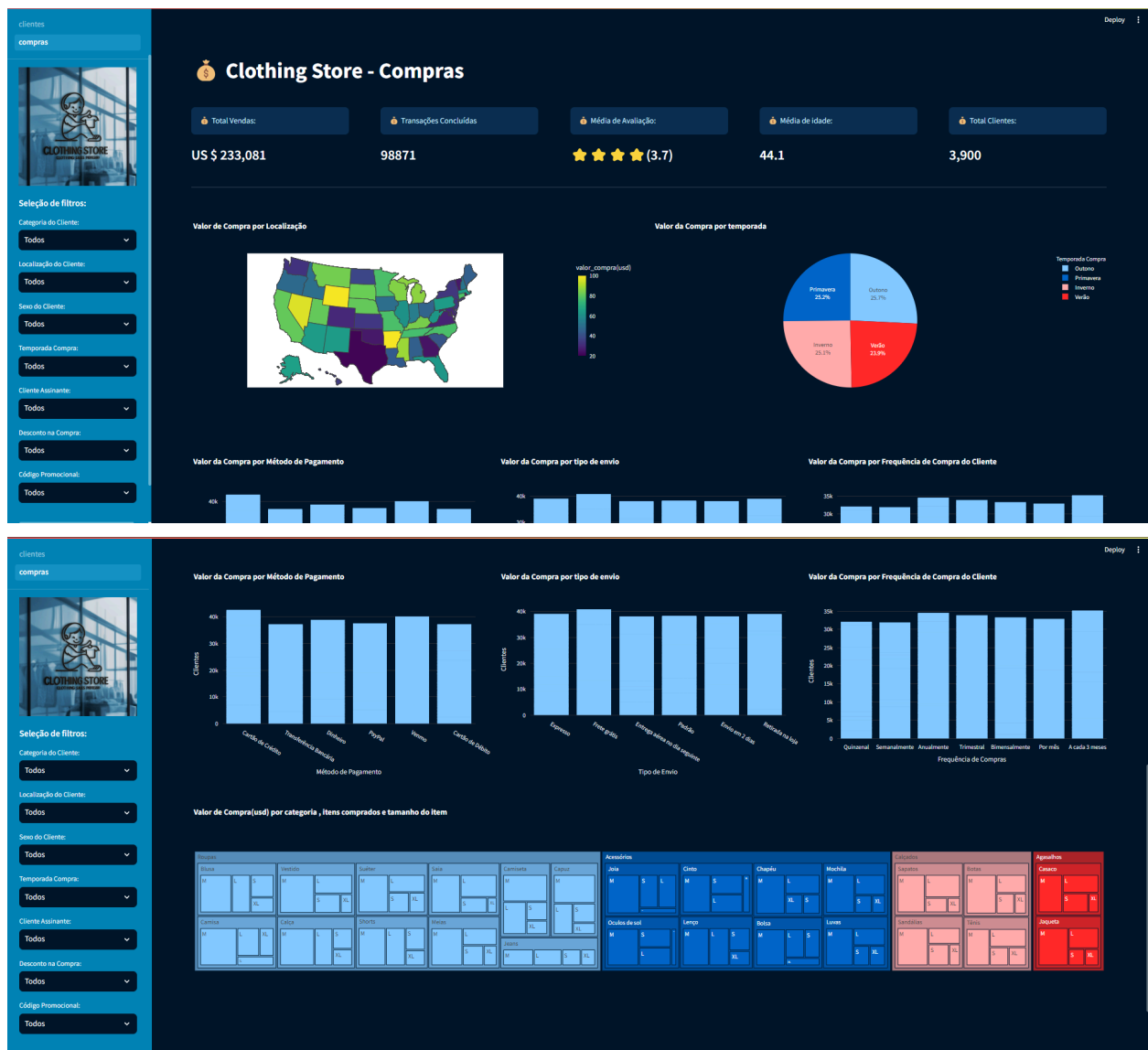
Ao longo da construção do painel interativo foi lidado com obstáculos sobre como seria construído, começou primeiramente com a biblioteca streamlit, mas por sugestão do professor Kaio Monteiro o projeto foi mudado para biblioteca dash, mas ocorreram problemas técnicos ao longo do caminho e consideramos o power bi uma opção mais viável, mas como a regra do projeto é visualizações em linguagem python, optamos por voltar a escolha inicial e foi construído duas páginas com a biblioteca streamlit, a primeira página com visualizações focada na distribuição de clientes e a segunda página focado nos valores de compra dos clientes, ambas aplicando os mesmo atributos categóricos, mas para valores numéricos diferentes. Ambas as páginas são divididas em sidebar e conteúdo principal, o sidebar contém botões de navegação, logo da loja e os filtros, enquanto o conteúdo principal possui os cartões (KPIs relevantes) e os gráficos. Os gráficos presentes no projeto são: gráfico de barras, setores, treemap e mapa com o intuito de deixar o projeto variado e facilitar a geração de insights.

Na 1º página foi adicionado seis gráficos, ambos os gráficos referentes ao valor de contagem de clientes, o gráfico de mapa traz dados a respeito da quantidade de clientes por localização, o gráfico de pizza traz os clientes por temporada, o que foi uma escolha ideal para esse tipo de gráfico por ter poucas categorias, também foi adicionado três gráficos de barra sobre métodos de pagamento, tipo de envio e frequência de compra dos clientes e a último gráfico é o treemap que traz dados sobre a categoria do produto, item comprado e tamanho do item. Todos esses visuais com o

objetivo de dar uma visão geral de como está a situação da quantidade de clientes na loja de acordo com diferentes atributos.



Na 2º página também foi adicionado seis gráficos, ambos os gráficos referentes ao valor de compra dos clientes, o gráfico de mapa traz dados sobre o valor total em cada localização, o gráfico de pizza traz o valor comprado por temporada, também foi adicionado três gráficos de barra sobre os valores de compra por métodos de pagamento, tipo de envio e frequência de compra dos clientes e por fim, o treemap traz dados sobre o total comprado por categoria do produto, item comprado e tamanho do item. Todos esses visuais com o objetivo de dar uma visão geral de como está a situação de compra dos clientes na loja de acordo com diferentes colunas do dataset.



6. Conclusão

Nesse projeto foi entregue como produto final um painel interativo feito em python com a utilização das bibliotecas pandas, numpy, plotly e streamlit através das etapas de pré-processamento e análise de dados, em que foi aplicado conteúdos visto dentro e fora de sala de aula para entregar um projeto que tem como objetivo análises que ajudem na tomada de decisão de maneira mais assertiva, trazendo consigo informações claras que o público possa entender, além de um visual elegante que entrega análises variadas com diferentes atributos, KPIs e filtros.

7. Referências e links úteis

- Repositório no Github: https://github.com/micaellimaj/Projeto_Dev_Python
- Link do deploy: <https://clothingstore.streamlit.app/>
- Link do dataset:
<https://www.kaggle.com/datasets/iamsouravbanerjee/customer-shopping-trends-dataset>
- Conteúdos de referência:
 - <https://youtu.be/nPz88occPYw?si=OeO1va0WgSn-uBAg>
 - <https://youtu.be/pWxDxhWXJos?si=GWELoHAVm6PRR8Cs>
 - <https://youtu.be/Sb0A9i6d320?si=i3XI849cJuHd1XdK>
 - <https://streamlit.io/>
 - <https://plotly.com/python/>