Relatório de Documentação Interna - Projeto Final: Summer Products

# Índice

# Introdução ………………………………………………………………………………………….2

# Objetivo do Projeto …………………………………………………………………………….3

Fontes de Dados …………………………………………………………………………4

Ferramentas e Tecnologias Utilizadas………………………………………… 4

Tratamento e Limpeza de Dados …………………………………………………5

Análises Realizadas……………………………………………………………………..8

Resultados e Insights…………………………………………………………………..18

Conclusões e Recomendações……………………………………………………..18

Introdução

* O nosso grupo de trabalho recebeu em mãos para o desenvolvimento do projeto final, a base de dados da plataforma Wish, com o respetivo guião do trabalho. O nosso grupo tinha a função de, de acordo com o guião, trabalhar e limpar a base de dados de forma a procedermos às nossas respetivas análises.
* Sobre a empresa em si e à sua atividade, antes de iniciarmos o nosso planeamento do projeto , debruçámo-nos sobre a mesma e fizemos uma pequena pesquisa, na qual, do que investigámos, verificámos que esta plataforma além de ser mundialmente conhecida tem como principal atividade a venda dos mais variados tipos de produtos e a baixo custo , em que, a maioria deles, são quase todos provenientes da China .
* O business model da empresa baseia-se sobretudo em preços fortemente competitivos, produtos de alta rotatividade e experiência de compra ramificada e por recomendação

Objetivo do Projeto

* O objetivo inicial do nosso projeto foi criar e desenvolver o business goal da empresa, de acordo com os dados que dispúnhamos. Definimos então que era identificar quais os principais fatores que influenciam o sucesso de vendas dos produtos de Verão da plataforma Wish, através da atratividade e visibilidade dos produtos, ajustar os seus preços e promoções e identificar da melhor forma possível quais os principais vendedores e produtos que mais contribuem para o sucesso de vendas.
* Após o business goal definido, o nosso objetivo passará por absorver e interpretar ao máximo a base de dados desta plataforma, de forma a podermos simplificar e clarificar a nossa análise, chegar aos resultados desejados, desenvolver o nosso espírito crítico e elaborar ações estratégicas para os resultados gerados pelos gráficos obtidos

Fontes de Dados

* Em relação aos dados fornecidos para desenvolvermos o nosso projeto final, além de já termos previamente á nossa disposição o Work Cycle onde explica todo os passos detalhados para trabalharmos este projeto enquanto analistas, recebemos a base de dados da Wish em formato Excel e tivemos em nossa mão este link “Sales of summer clothes in E-commerce Wish | Kaggle” bastante importante a nosso ver e que deu nos a conhecer como era o Dataset da plataforma Wish e em que dados se baseava .

Tecnologias Utilizadas

* Em relação às tecnologias que utilizámos ao longo deste trabalho, nós planeámos em primeira mão trabalhar a base de dados em Excel ( limpeza , colunas calculadas , gráficos , etc) . No entanto chegámos á conclusão que seria mais eficaz utilizar o Python, pois tínhamos acabado de trabalhar esta tecnologia no nosso último trabalho. Acabámos por dar uso na limpeza de dados, calcular colunas e criar gráficos.
* Por último em relação á parte estética e visual para dispormos para apresentação resolvermos utilizar o Power Bi pois considerámos uma ferramenta fundamental na criação de gráficos criativos que resultam numa melhor interpretação de dados

Tratamento e Limpeza de Dados

* Em relação ao tratamento e limpeza de dados realizado, como mencionámos anteriormente a nossa ideia inicial era executar em Excel mas depois decidimos fazer a limpeza de dados em Python porque seria mais eficaz.
* Começando então pelo tratamento começámos por verificar quantas linhas tinha a base de dados através da linha de código :

A screenshot of a computer

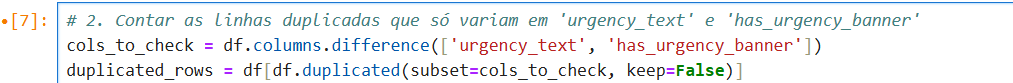
AI-generated content may be incorrect.

* Prosseguindo no tratamento, resolvemos efetuar a remoção de duplicados :

A white rectangular sign with blue text

AI-generated content may be incorrect.

* Após essa remoção reparámos que havia linhas que praticamente continham a mesma informação e só havia diferenças nas colunas ‘urgency\_text’ e ‘has\_urgency\_banner’. Como pretendemos preservar a informação de quais os produtos que foram marcados como urgentes, optámos por remover as linhas duplicadas que cuja informação nas colunas ‘urgency\_text’ e ‘has\_urgency\_banner’ estavam por preencher .

 Para tal fizemos uma contagem em primeiro lugar dos duplicados que só variam em ‘urgency\_text’ e ‘has\_urgency\_banner’:

Jjjjjkk

Para de seguida então eliminar as linhas duplicadas:

Por último em relação a este passo atualizámos o Data Frame já removendo as linhas duplicadas com urgência vazias:



* De acordo com a análise que fizemos das colunas em questão , resolvemos excluir por opção própria as colunas

“title” , “product\_url” , “ product\_picture”, “merchant\_profile\_picture” pois entendemos que não tinham relevância para a nossa análise em questão:

kkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkk

Feito esta parte inicial da limpeza quisemos apenas agora verificar quantas colunas e linhas tínhamos na nossa base de dados:

A screenshot of a computer code

AI-generated content may be incorrect.

* Partindo agora para um outro tratamento de dados quisemos preencher algumas colunas, das quais, tinham valores vazios em relação á origem dos países em falta, como também preencher registos vazios na coluna “urgency\_text” através de um texto que seria mais compreensível para quem está a fazer a análise:
* A screenshot of a computer code

  AI-generated content may be incorrect.gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg
* Em resumo , foi este o tipo de tratamento e limpeza que fizemos antes de procedermos à nossa análise e criação de colunas e gráficos que iremos explicar de seguida de que maneira o fizemos e se os resultados obtidos iam de encontro às nossas expectativas.

Análises Realizadas

* Em primeiro lugar quisemos atribuir diferentes níveis de rating a cada produto , estabelecendo e criando uma nova coluna “rating\_group”, no sentido de percebermos se esta base de dados seguia a política da empresa: venda de produtos a baixo custo e com boas avaliações.

A screenshot of a computer program

AI-generated content may be incorrect.

Sss

Após isso, verificámos então o total de classificações que havia por grupo através do rating\_group:

A screenshot of a computer code

AI-generated content may be incorrect.

* Para criar algumas colunas que achámos relevantes calcular para a nossa base de dados resolvemos então calcular qual era o Revenue e quanto era a percentagem de desconto por produto. Nesta medida , nós decidimos que para calcular o desconto seria pegar no valor de mercado menos o valor que era praticado na wish através da seguinte fórmula:

A screenshot of a computer code

AI-generated content may be incorrect.

* A screenshot of a computer program

  AI-generated content may be incorrect.Um dos focos importantes que demos à nossa análise foi categorizarmos os nossos produtos em diferentes categorias porque como o nome de cada produto era muito extenso , resolvemos simplificar cada nome e marcá-lo numa determinada categoria , para isso criámos a coluna “producto\_category” e depois o respetivo gráfico :

A graph with numbers and text

AI-generated content may be incorrect.

* A partir deste ponto resolvemos focar a nossa análise na criação de diferentes gráficos com especial atenção às unidades vendidas em relação às variáveis em estudo como por exemplo o preço, a avaliação e os países :

A screen shot of a computer code

AI-generated content may be incorrect.

A graph with blue dots

AI-generated content may be incorrect.

A screenshot of a computer code

AI-generated content may be incorrect.

A graph with blue dots

AI-generated content may be incorrect.

A chart of different colored dots

AI-generated content may be incorrect.A white background with text

AI-generated content may be incorrect.

* A screenshot of a computer program

  AI-generated content may be incorrect.Em relação a outros pontos estratégicos da análise que considerámos relevantes, ponderámos analisar então as categorias do produto com algumas variáveis, como as unidades vendidas, a receita e a avaliação média . Para tal tivemos de categorizar os produtos primeiro e definir condições na linha de código em relação aos nomes dessas mesmas categorias. De seguida, aplicámos o rating médio de cada categoria e calculámos o seu revenue. A linha de código em baixo bem como a tabela a seguir mostra como foi feito o nosso raciocínio.

A screenshot of a computer

AI-generated content may be incorrect.

* Um dos factos curiosos que ficámos a conhecer foi em primeiro lugar saber como era a variação das unidades vendidas nos produtos que tinham levado ad boost da parte dos vendedores. O resultado foi interessante pois os produtos que não tinham levado ad boost foram os que venderam mais. De seguida iremos mostrar como fizemos o nosso raciocínio em código:

A screenshot of a computer program

AI-generated content may be incorrect.

A blue pie chart with text

AI-generated content may be incorrect.

* Relativamente ainda sobre os produtos , nós enquanto grupo decidimos analisar quais eram os produtos que apresentaram um maior risco de devolução , para tal definimos que os produtos que tinham um rating inferior a 3.5 e com uma elevada quantidade de unidades vendidas , considerámos então que estavam expostos a que fossem mais tardes devolvidos. Isto de facto impacta negativamente a imagem da empresa. O código que utilizámos para efetuar esta análise foi o seguinte:

A screenshot of a computer program

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

A graph of a number of products

AI-generated content may be incorrect.

* Ainda sobre a análise dos produtos , resolvemos então explorar por tamanhos e cores quais eram então, as mais comuns . No entanto em relação aos resultados que obtivemos, não conseguimos estabelecer razões plausíveis para os mesmos porque não tínhamos informação sobre nomes de clientes, datas de nascimento , entre outros dados pessoais e, devido a isso só podemos analisar dados, nos resultados diretos que obtivemos :

A screenshot of a computer program

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.

A graph with different colored bars

AI-generated content may be incorrect.

* A nossa última análise incide-se sobre os vendedores porque nós achamos que os vendedores têm um papel fundamental no desempenho desta plataforma , principalmente nas avaliações médias que lhes são atribuídas, é com base nas avaliações e unidades vendidas que os vendedores poderão sempre melhorar o seu desempenho em relação ao sucesso das vendas de produtos de verão desta mesma plataforma .

De seguida o que pretendemos mostrar foi a análise dos top 10 vendedores com base na avaliação média e unidades vendidas

A screenshot of a computer program

AI-generated content may be incorrect.

A graph of different colored bars

AI-generated content may be incorrect.

* De seguida apresentamos então o top 10 vendedores por nº de avaliações e receitas estimadas. Acreditamos que, com este gráfico conseguimos perceber quais os vendedores que estão mais bem posicionados no mercado e se realmente o número de avaliações é diretamente proporcional à receita estimada

A screenshot of a computer program

AI-generated content may be incorrect.

A graph with a bar chart

AI-generated content may be incorrect.

Resultados e Insights

* A maioria dos produtos têm boas avaliações e preços baixos — o que favorece as vendas.
* Produtos com avaliações de 3 e 5 estrelas apresentam maior impacto positivo nas vendas.
* As unidades que não tiveram o impulsionamento de anúncios (Boost) têm melhores resultados que as unidades que têm Boost.
* As cores preto, branco e cinza são as mais vendidas. O tamanho "S" (pequeno) destaca-se como o mais popular entre os consumidores, o que pode ajudar na definição de **stock e marketing**.
* Algumas categorias têm descontos mais agressivos (ex.: bikinis), o que não correlacionou diretamente com um aumento de vendas.
* Produtos com baixa avaliação devem ser monitorizados para **prevenir devoluções**.

Conclusões e Recomendações

Identificar os principais fatores que influenciam o sucesso de vendas de produtos de verão na plataforma Wish, de forma a otimizar a estratégia de listagem e promoção de produtos:

* Melhorar a visibilidade e atratividade dos produtos
* Otimizar preços e promoções
* **Ações estratégicas** para aumentar as vendas e diminuir devoluções incluem: Cupões personalizados, gamificação (jogos), provas nas redes sociais, melhoria nas descrições, imagens realistas, controlo de qualidade dos vendedores.
* Aumentar confiança e fidelização dos clientes: Garantias visíveis, reputação do vendedor, política clara, programa de pontos, cashback, recompensas por engajamento