# Case OLIST

Foram criadas visualizações que auxiliam no monitoramento das categorias dos produtos vendidos pelo e-commerce, incluindo valores mínimos, máximos e o total de vendas por estado. A visualização permite acompanhar o total de vendas feitas por estado, as compras por localidade e as avaliações por categoria de produtos. Além disso, uma seção foi criada para comparar as avaliações com a quantidade de comentários dos produtos. Na última tela, é possível acompanhar a média de tempo gasto para realizar uma entrega, a estimativa de entrega, o total de itens entregues, o total de itens por status de entrega e as notas dos itens entregues.

**Página 1: Monitoramento de Categorias de Produtos**

* **KPIs Principais**:
  + **Total de Categorias de Produtos**: Mostra o total de categorias existentes na base de dados.
  + **Menor Valor de Produto**: Exibe o produto com o menor valor.
  + **Maior Valor de Produto**: Exibe o produto com o maior valor.
* **Funcionalidades**:
  + Filtro de **Categoria de Produto** para manipulação da visualização.
  + Tabela com visualização de categorias e quantidades de venda.
  + Mapa mostrando o total de compras por estado.

**Página 2: Vendas por Estado**

* **Visualizações**:
  + Gráfico de funil mostrando o total de vendas por estado.
  + Quatro mapas destacando:
    - Total de clientes por estado.
    - Total de vendedores por estado.
    - Compras por UF.
    - Vendas por UF.

**Página 3: Avaliações e Comentários**

* **Visualizações**:
  + Tabela com a quantidade de avaliações e comentários, detalhada por notas de 1 a 5.
  + Gráfico de área mostrando a proporção entre avaliações e comentários.
  + Gráfico de barras exibindo notas por categoria de produtos.

**Página 4: Prazos de Entrega**

* **Visualizações**:
  + Gráficos focados nos prazos de entrega:
    - Média de tempo desde a aprovação do pedido até a entrega.
    - Média real do tempo de entrega, considerando a data de aprovação e a data de entrega.
  + Tabela com status de entrega e total de cada status.
  + Gráfico de barras com notas dos itens entregues.

**Fonte de Dados e ETL**

* **Fonte de Dados**:
  + Dados baseados em arquivos CSV, carregados através do Pentaho Data Integration.
  + Dados transformados e carregados no banco de dados PostgreSQL.
* **Tabelas Identificadas**:
  + **Tabelas de Fato**: order\_items, order\_payments, order\_reviews, orders.
  + **Tabelas de Dimensão**: customers, geolocation, products, sellers, category\_name\_translation.
  + Dimensão de tempo criada pela stored procedure UP\_CARGA\_DIMENSAO\_TEMPO, populada com o período de 2000 a 2100 (dados disponíveis a partir de 2017).
* **Data Warehouses**:
  + **dw\_olist**: Armazena informações gerais.
  + **dw\_entrega**: Armazena informações específicas de entregas.
* **Processo ETL**:
  + Criada uma transformação para cada arquivo CSV, carregando os dados em tabelas no banco de dados.
  + Transformações divididas em jobs para dimensões e fatos.
  + Dimensão de tempo populada para o período de 2000 a 2100.

**Modelo de Predição**

* **Desenvolvimento**:
  + Modelo de predição criado em Python usando dados do **dw\_entrega**.
  + Objetivo: Prever a chance de um pedido ser entregue com sucesso ou ser cancelado.
* **Resultados**:
  + Dados insuficientes para um modelo de aprendizado 100% acurado.
  + Decisão de não utilizar o modelo como fonte no dashboard devido à baixa acurácia.

Esta documentação oferece uma visão abrangente das visualizações criadas, fontes de dados utilizadas, processos ETL implementados e a tentativa de integração de um modelo de predição.

# Arquitetura

Diagrama

Descrição gerada automaticamente