

DASHBOARD COMERCIAL

USANDO PYTHON, SQL E POWER BI



Mario Henrique
Cosme Juvencio

DASHBOARD COMERCIAL USANDO POWER BI

Banco de dados em PostgreSQL / API em Python / *Dashboard* em Power BI

Contexto: Uma empresa opera em múltiplos locais, possuindo diversas lojas em diferentes cidades e estados. Ela comercializa uma variedade de produtos com o auxílio de uma equipe de vendedores. O negócio gera uma grande quantidade de dados a partir de suas atividades diárias. Esses dados têm um grande potencial para gerar insights valiosos para a empresa e resolver problemas de negócio. Para responder a essas perguntas, será construído um processo de análise de dados. Esse processo envolve várias etapas, desde a análise de requisitos, criação e manutenção do banco de dados até a construção dos dashboards interativos para visualização dos dados. Portanto, o objetivo desse projeto é transformar os dados brutos da empresa em informações estratégicas que possam ser usadas para tomada de decisões orientadas a dados e impulsionar o sucesso do negócio.

1 ANÁLISE DE REQUISITOS

A análise de requisitos envolve a compreensão das necessidades, problemas e objetivos do cliente ao desenvolver um sistema ou solução. No contexto desse projeto, envolve o entendimento das necessidades da empresa relacionadas ao gerenciamento e análise de seus dados de venda. O objetivo da análise de requisito é criar uma descrição clara e detalhada do que o sistema deve fazer, bem como qualquer restrição sobre como ele deve operar. Dessa forma, foi construída a análise a partir dos dados brutos e dos objetivos da solução a ser implementada.

a. Gerenciamento de Produtos

O sistema deve ser capaz de registrar produtos, juntamente com seus detalhes, como nome do produto, categoria e fabricante.

b. Gerenciamento de Vendedores

O sistema deve permitir o registro de vendedores com seus identificadores únicos e nomes.

c. Gerenciamento de Lojas

O sistema deve ser capaz de registrar lojas, juntamente com seus detalhes, como identificador da loja, cidade e estado.

d. Gerenciamento de Vendas

O sistema deve permitir o registro de vendas. Cada venda deve incluir um identificador do produto, identificador do vendedor, identificador da loja, data da venda, valor da venda, custo, lucro, segmento do cliente e comissão paga ao vendedor.

e. Análise de Vendas

O sistema deve permitir a análise de dados de vendas de várias maneiras, incluindo vendas por vendedor, vendas por loja, vendas por categoria, vendas por segmento de cliente e tendências de vendas ao longo do tempo.

f. Análise de Rentabilidade

O sistema deve permitir a análise da rentabilidade, calculando o lucro para cada venda e permitindo análises agregadas de rentabilidade por vendedor, loja, categoria, segmento de cliente e evolução ao longo do tempo da receita, custo e lucro.

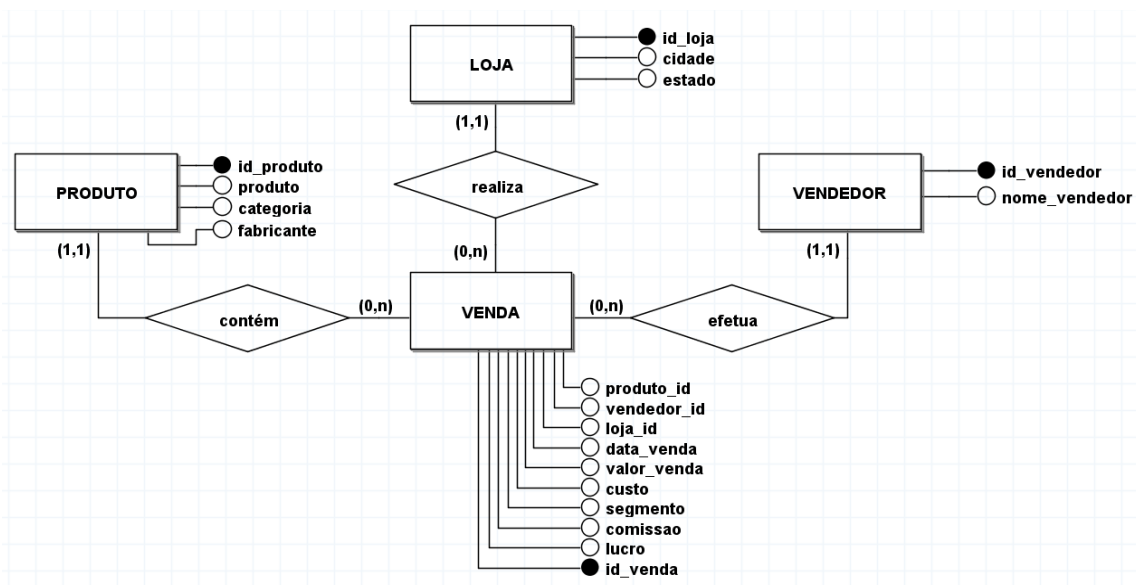
2 MODELAGEM DE DADOS

A modelagem de dados é um processo essencial na construção de um banco de dados. Ela envolve a definição e organização de dados de acordo com um modelo de dados para refletir com precisão os processos de negócios e requisitos de informação de uma organização. Em termos mais simples, é o processo de criar um modelo visual para representar os tipos de dados que serão usados, como eles estão interligados e como podem ser manipulados.

A modelagem de dados é vital no desenvolvimento de ferramentas de visualização de dados, tais como os dashboards em Power BI, pois promove eficiência e alta performance, especialmente ao lidar com grandes volumes de dados. Além disso, a modelagem garante a precisão e confiabilidade das análises, minimizando a ocorrência de erros, uma vez que os dados estão organizados de maneira lógica e consistente através das restrições de integridade definidas na construção do banco de dados. A modelagem de dados, através dos modelos conceituais, também permite a visualização da organização dos dados até mesmo para usuários não técnicos, proporcionando flexibilidade na análise dos dados e permitindo explorar as relações entre os dados e extrair insights.

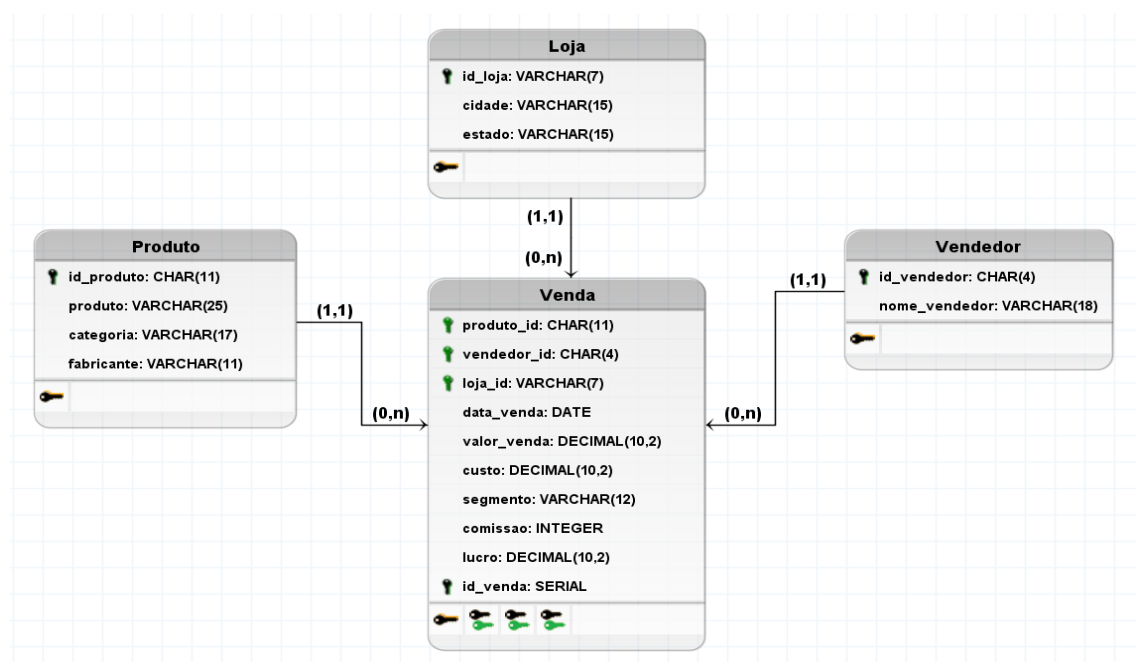
No presente projeto, o banco de dados foi construído a partir do SBGDR PostgreSQL. A partir dos dados brutos e da análise de requisitos, foram construídos os modelos conceitual e lógico do banco de dados.

Figura 1 – Modelo conceitual do banco de dados.



O modelo conceitual de dados é uma representação de alto nível do esquema do banco de dados e fornece uma visão geral dos dados sem se aprofundar nos detalhes técnicos. A partir da Figura 1 é possível visualizar as entidades de dados (produto, loja, vendedor, venda), os seus atributos e os relacionamentos entre elas. O modelo conceitual foi obtido na fase de análise dos dados brutos do negócio e levantamento dos requisitos do sistema, servindo como ponto de partida para o entendimento do negócio. Em seguida, o modelo conceitual foi traduzido em um modelo lógico, incorporando os detalhes técnicos, como tipos de dados específicos para os atributos. A Figura 2 apresenta o modelo lógico construído.

Figura 2 – Modelo lógico do banco de dados.



Após isso, o banco de dados relacional foi construído implementando as tabelas, atributos e restrições de integridade descritas na modelagem de dados.

3 TRATAMENTO DA BASE DE DADOS BRUTA

Os dados usados neste estudo foram obtidos por meio da plataforma Data Science Academy (<https://www.datascienceacademy.com.br/>). Originalmente, esses dados estavam contidos em uma única planilha, exigindo uma etapa preliminar de preparação para permitir análises adequadas. Este processo envolveu a limpeza e organização dos dados, seguido pela sua divisão em várias tabelas, conforme especificado durante a modelagem de dados. Estas tabelas foram então utilizadas para alimentar o banco de dados, construído no PostgreSQL, a partir de uma API em Python. Para visualizar o algoritmo usado para limpeza e organização dos dados, acesse meu repositório no GitHub através do link abaixo.

https://github.com/MarioHenriqueUFAL/ProjetosBI/blob/main/projeto_comercial.ipynb

O pré-tratamento dos dados foi realizado para eliminar as inconsistências e erros na base de dados. Além disso, para obtenção de novas métricas, foram inseridas novas colunas obtidas a partir de cálculos realizados em outras colunas com informações numéricas, tais como a coluna com dados sobre os lucros em cada venda realizada. O tratamento adequado dos dados nesse projeto tornou os dados mais fáceis de serem manipulados e entendidos para extração de insights mais precisos e relevantes para o negócio.

4 CRIAÇÃO DE API PARA ALIMENTAÇÃO DO BANCO DE DADOS

Para inserir os dados no banco de dados, foi necessária a construção de uma API em Python. O algoritmo construído realizou basicamente as seguintes tarefas:

- Conexão do Python com o banco de dados, utilizando a biblioteca “psycopg2”;
- Leitura da base de dados;
- Criação de um cursor, responsável por executar comandos SQL e interagir com o banco de dados;
- Execução de comandos SQL para inserção dos dados nas suas respectivas tabelas.

O algoritmo implementado pode ser visualizado abaixo.

```

import pandas as pd
import psycopg2
from psycopg2 import sql

df = pd.read_excel('Dados_Comerciais_modificado_v3.xlsx')

conn = psycopg2.connect(
    dbname='Comercial',
    user='postgres',
    password='*****',
    host='localhost',
    port='5432'
)
cur = conn.cursor()

for index, row in df.iterrows():

    query = sql.SQL("""INSERT INTO Venda (produto_id, vendedor_id,
                                   loja_id, data_venda, valor_venda, custo, segmento, comissao, lucro)
                                   VALUES (%s, %s, %s, %s, %s, %s, %s, %s, %s)""")

    # Executa a consulta
    cur.execute(query, (row['id_produto'], row['id_vendedor'],
                        row['loja'], row['data venda'], row['valor venda'],
                        row['custo'], row['segmento'], row['comissao percentual'],
                        row['lucro']))

conn.commit()

cur.close()
conn.close()

```

5 ELABORAÇÃO DO DASHBOARD NO POWER BI

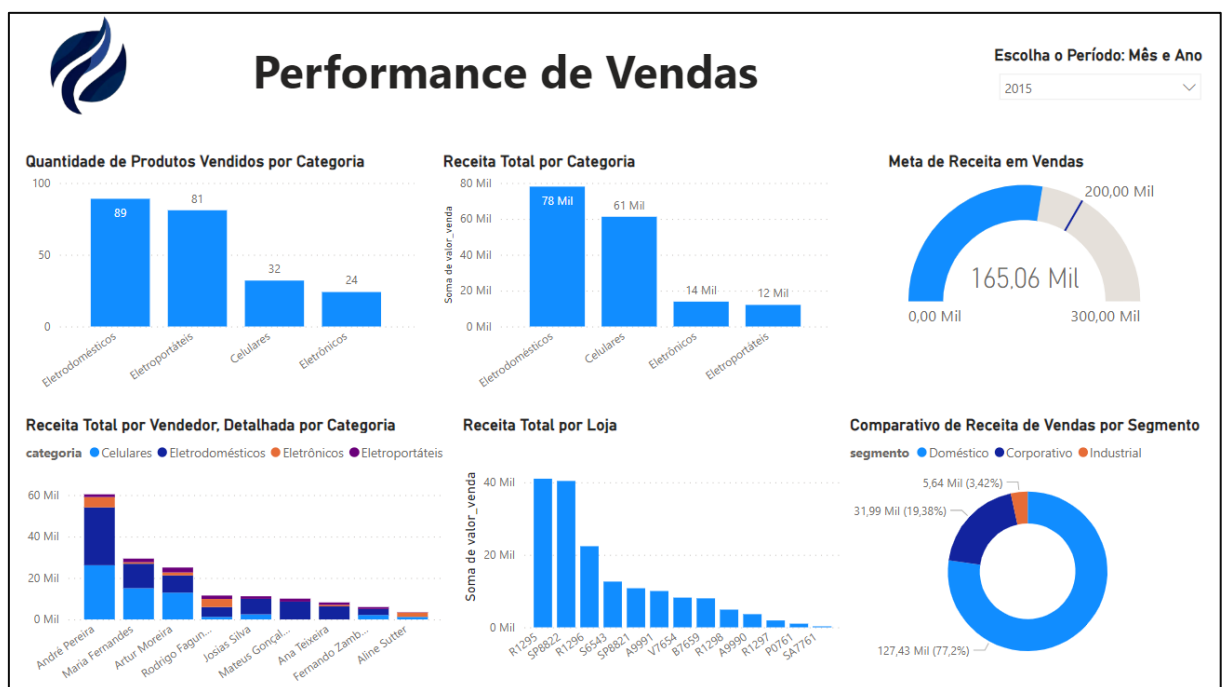
Com o banco de dados devidamente construído e os dados da empresa corretamente armazenados, procedeu-se à conexão do Power BI com o PostgreSQL para a importação das tabelas e informações correspondentes. As tabelas, seus respectivos relacionamentos e cardinalidades foram interpretados de maneira precisa pelo Power BI, tornando desnecessárias quaisquer manipulações adicionais na modelagem de dados dentro do *software*. A modelagem de dados reconhecida no Power BI pode ser vista na Figura 3.

Figura 3 – Modelagem de dados no Power BI.



No Power BI foram criados dois *dashboards*: o primeiro concentra-se na análise da performance de vendas por categoria de produto, por vendedor, por loja e por segmento de cliente, além de apresentar um indicador referente à meta de receita anual, conforme à Figura 4.

Figura 4 – *Dashboard* sobre performance de vendas.

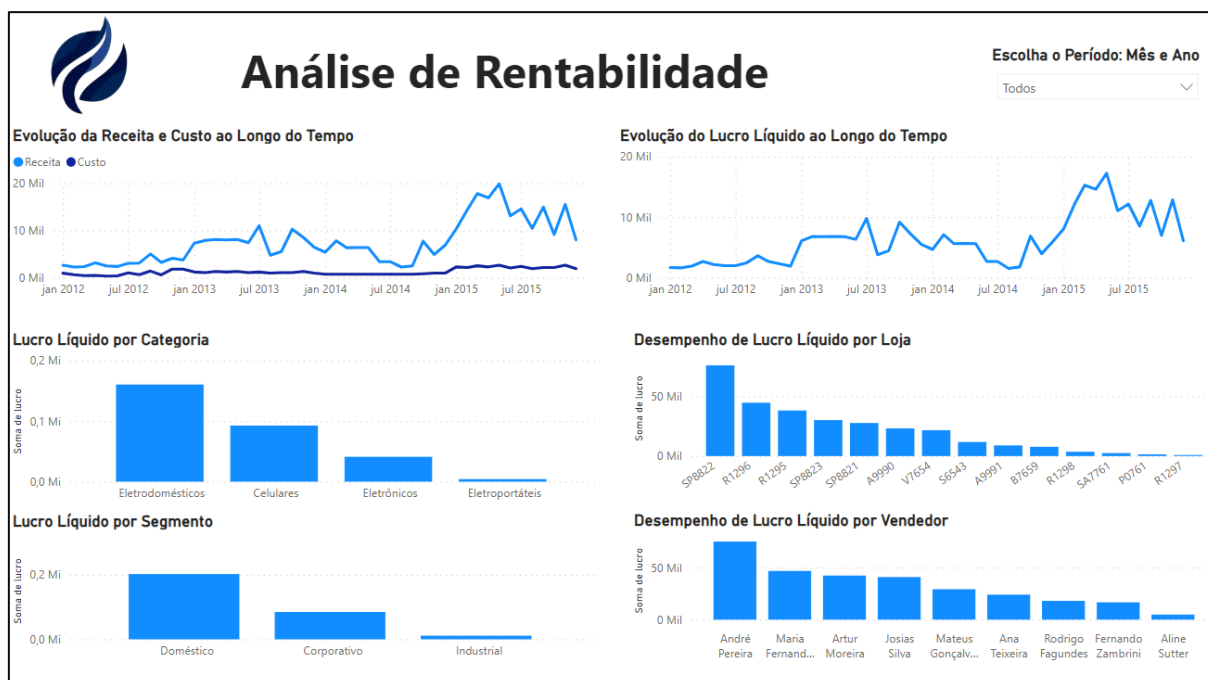


O *dashboard* da Figura 4 ajuda a responder as seguintes perguntas estratégicas:

- Quais categorias de produto estão performando melhor em vendas?
- Quais vendedores estão performando melhor? Quais estão abaixo da média?
- Quais lojas têm a maior quantidade de vendas? Existe alguma loja que está com desempenho ruim?
- Qual o segmento de cliente apresenta o maior volume de vendas?
- Como a empresa está performando em relação à meta de receita de vendas anual? Quanto falta para alcançar a meta?

O segundo *dashboard*, por sua vez, oferece uma visão abrangente da rentabilidade da empresa, realizando uma análise detalhada do lucro obtido por categoria de produto, por loja, por segmento de cliente e por vendedor. Além disso, o segundo relatório também fornece uma análise da evolução temporal da receita, custo e lucro ao longo dos anos para identificar possíveis padrões sazonais e tendências, como mostrado na Figura 5.

Figura 5 – *Dashboard* sobre análise de rentabilidade.



O *dashboard* de análise de rentabilidade ajuda a responder as seguintes questões?

- Quais categorias de produto são as mais rentáveis?
- Quais lojas estão trazendo mais lucros? Existe alguma loja que não está sendo rentável?

- Qual segmento de cliente mais rentável?
- Quais vendedores estão contribuindo mais para obtenção de lucros da empresa?
- Qual é a tendência ao longo do tempo em relação à receita, custo e lucro? Existe algum comportamento padrão ou sazonalidade perceptível?
- Qual é a previsão de lucro para o próximo período com base nas tendências atuais? Qual é a relação entre o custo e o lucro ao longo do tempo?

Dessa forma, os *dashboards* construídos no presente trabalho podem ajudar a empresa a tomar decisões estratégicas e operacionais baseadas em dados, melhorar a eficiência, aumentar a rentabilidade e definir metas e estratégias futuras.