

Guia do Projeto - Análise de Dados de Vendas (E-commerce)

O objetivo deste projeto é atuar como um Analista de Dados que transforma dados brutos de transações de vendas em insights acionáveis para a tomada de decisão de negócio. O projeto utiliza dados reais de e-commerce e abrange desde a preparação dos dados até a criação de dashboards interativos.

1. ■■ Ferramentas e Tecnologias

- **Python:** Linguagem principal para manipulação, limpeza e análise de dados.
- **Pandas e NumPy:** Para agrupamento, fusão e tratamento de valores.
- **Matplotlib e Seaborn (ou Plotly/Altair):** Visualização de dados (EDA).
- **Power BI / Tableau / Streamlit:** Dashboarding interativo.
- **Git e GitHub:** Versionamento e documentação do projeto.

2. ■ Base de Dados Sugerida

Será utilizado o ***Brazilian E-commerce Public Dataset by Olist***, disponível no Kaggle. Este dataset contém múltiplas tabelas (pedidos, clientes, itens, produtos, pagamentos, avaliações, geolocalização), permitindo a aplicação de joins e análises realistas.

3. ■ Entregáveis do Projeto

- A. Notebook de Análise Exploratória (EDA):** Limpeza, pré-processamento, análise descritiva e criação de métricas.
- B. Análises de Negócio e Insights:** Responder perguntas-chave sobre performance, rentabilidade e comportamento de clientes.
- C. Dashboard de Visualização:** Exibir KPIs e insights em formato visual e interativo.

4. ■■ Roadmap em 4 Fases

Fase 1: Aquisição e Preparação dos Dados (ETL com Python)

- Obtenção e carregamento dos dados CSV.
- Inspeção e entendimento das tabelas.
- Fusão de dados usando joins.
- Limpeza e tratamento de dados.

Fase 2: Análise Exploratória de Dados (EDA)

- Criação de features (datas, lucro, prazo de entrega).
- Análise de distribuição (histogramas, boxplots).
- Agrupamentos e cálculo de KPIs.
- Análise de correlação entre variáveis.

Fase 3: Geração de Insights e Documentação

- Implementar Análise RFM para segmentar clientes.
- Interpretar gráficos e resultados com clareza.
- Documentar todo o processo no GitHub.

****Fase 4: Visualização e Dashboarding****

- Escolher visualizações mais relevantes.
- Criar dashboard no Power BI, Tableau ou Streamlit.
- Certificar-se de que o painel responde às perguntas de negócios.

5. ■ Melhorias e Extensões do Projeto

- Aplicar modelos de Machine Learning para prever demanda futura.
- Criar alertas automáticos de performance (via Python + Email/API).
- Publicar o dashboard em nuvem (Streamlit Cloud ou Power BI Service).
- Adicionar storytelling no README, com visualizações explicativas.
- Implementar testes automatizados para as funções de ETL.

■ Entregue o projeto completo em um repositório GitHub contendo o código, o notebook EDA, o dashboard e o README documentando o processo e os insights obtidos.