olist

Project

Road Map

Project Overview

Olist Insight Hub is an advanced data analytics project that simulates a professional Business Intelligence (BI) ecosystem for a major e-commerce marketplace. The project analyzes the "Olist" dataset, integrating logistics, customer behavior, and product performance. It aims to transform raw, fragmented data into a unified "Source of Truth" to drive strategic decision-making and operational excellence.

Project Objectives

- **Data Integration:** To collect, clean, and integrate 8 relational tables into a robust SQL data warehouse.
- **Logistics Optimization:** To measure delivery performance by analyzing the gap between estimated and actual delivery dates.
- **Customer Segmentation:** To implement a dynamic **RFM (Recency, Frequency, Monetary) Model** to categorize customers and guide marketing strategies.
- **Quality Analysis:** To investigate the correlation between shipping speed, product categories, and customer review scores.
- **Performance Monitoring:** To develop interactive dashboards for tracking Key Performance Indicators (KPIs) and Year-over-Year (YoY) growth.

Scope of Work

The project covers four primary analytical domains:

- **Executive Analysis:** High-level overview of revenue, order volume, and sales growth patterns (2016–2018).
- **Logistics & Fulfillment:** Analysis of carrier performance, delivery lead times, and geographic delay patterns.
- **Customer Insights (RFM):** Detailed segmentation of the customer base to identify "Champions" vs. "At-Risk" users.
- **Product & Review Analysis:** Evaluation of product category success and its impact on customer satisfaction levels.

Tools & Technologies

- **Python (Pandas/NumPy):** Used for data profiling, datetime conversion, and quality assessment (preserved nulls for SQL-stage imputation).
- **SQL Server (T-SQL):** Used for database schema design, enforcing referential integrity (PK/FK), and denormalizing data into analytical "Flat Tables."
- **Power BI:** Used for advanced DAX modeling, Time Intelligence calculations, and interactive dashboard design.
- **DBDiagram.io:** Used for architecting the Entity Relationship Diagram (ERD).

Expected Deliverables

- **Cleaned Relational Database:** A well-structured SQL schema ready for analytical querying.
- **Advanced DAX Library:** A collection of measures for YoY growth, RFM scoring, and logistics KPIs.
- **Interactive Dashboard:** Comprising Executive Summary, Logistics, Customer RFM, and Product Quality pages.
- **Actionable Insight Report:** A summary of findings and strategic recommendations for Olist stakeholders

Timeline

Phase	Task	Duration
1	Project Planning & Data Profiling	Week 1
2	Data Cleaning & Preprocessing (Python)	Week 2
3	Relational Modeling & SQL Engineering	Week 3
4	DAX Development & RFM Implementation	Week 4
5	UI/UX Dashboard Design (Power BI)	Week 5
6	Documentation & Final Presentation	Week 6

Project Goal

To create a high-performance, data-driven analytics report that delivers actionable insights, optimizes delivery operations, and enhances customer retention strategies across the Olist marketplace.