E-Commerce Transactions SQL Project

Project Overview

The project showcases a comprehensive retail data analytics workflow using PostgreSQL and an open-source e-commerce dataset. The goal was to demonstrate advanced SQL skills by building a robust ETL pipeline, data modelling, and insightful business analyses typical of a professional data analyst role. Unlike typical analytics projects, this study avoids visualization tools and focuses entirely on SQL-driven analysis.

Challenges Faced

- Cleaning and normalizing raw transactional data required thoughtful filtering and structuring to ensure accuracy and usability.
- Balancing query complexity with performance to ensure efficient execution on larger datasets.
- Joining columns from different tables was challenging but it helped build meaningful insights which improved both the accuracy and efficiency of the analysis.

Key Learnings

- I strengthen my understanding of star schema design and how it boosts analytical efficiency.
- Advanced use of PostgreSQL window functions and CTEs significantly improved query power, performance and maintainability.
- Building customer segmentation and cohort analyses deepened my grasp of turning data into strategic business value.

Future Improvements

- Incorporate machine learning models for customer churn prediction or product recommendation within the pipeline.
- Automate the ETL process with scheduling tools and logging for production readiness.
- Build a BI dashboard to visualize KPIs dynamically.

Business Impacts

The analysis delivers key retail KPIs – from revenue trends to customer behavior segmentation – enabling data-driven decisions around marketing, retention, and product management. This project illustrates how SQL expertise can directly support business growth and operational improvements.