Customer Purchasing Behavior Analysis Report

1. Project Overview

This project presents an end-to-end data analytics workflow analyzing customer purchasing behavior using Python for preprocessing, PostgreSQL for structured data storage and querying, and Power BI for interactive dashboard visualization. The objective of this project is to identify purchasing trends, customer segmentation patterns, category-wise revenue performance, and subscription behavior insights to support data-driven business decisions.

2. Dataset Description

- Total Records: ~4,000 customers
- Attributes: Demographics, Purchase Details, Payment Methods, Subscription Status
- Categories: Clothing, Footwear, Accessories, Outerwear
- Metrics: Purchase Amount, Review Rating, Previous Purchases, Frequency of Purchases

3. Data Processing & Engineering

Python was used to clean column names, handle missing values (median imputation for review ratings), create age groups using quantile-based segmentation, and map purchase frequency to numeric values. The cleaned dataset was then loaded into PostgreSQL using SQLAlchemy for further structured analysis. SQL queries were used to aggregate revenue by category, calculate average purchase amount, and evaluate subscription impact on sales performance.

4. Dashboard Insights (Power BI)

The Power BI dashboard provides interactive filtering by Subscription Status, Gender, Category, and Shipping Type. Key KPIs: • Total Customers: 4K • Average Review Rating: 3.75 • Average Purchase Amount: 59.76 Visualizations include: • Revenue by Category • Sales Distribution by Category • Revenue by Age Group • Subscription Percentage Distribution

5. Key Business Insights

- Clothing category generated the highest revenue among all categories.
- Young and middle-aged customers contributed the largest share of revenue.
- Subscription customers represent a significant portion of total sales.
- Higher review ratings correlate with stronger repeat purchasing behavior.

6. Conclusion

The Customer Purchasing Behavior Analysis project demonstrates strong analytical capability by combining data cleaning, SQL querying, and interactive dashboard development. The insights generated can assist businesses in optimizing product strategy, targeting high-value customer segments, and improving operational decisions through data-driven intelligence.