Customer Shopping Behavior Analysis

1. Project Overview

This project focuses on analyzing customer shopping behavior using transactional retail data. The goal is to identify patterns related to customer spending, product preferences, subscription impact, and discount usage to support data-driven business decisions.

2. Dataset Summary

The dataset consists of approximately 3,900 customer purchase records with multiple attributes including customer demographics, purchase details, and shopping behavior. The data was collected in CSV format and includes fields such as age, gender, category, purchase amount, discount usage, review ratings, and subscription status.

3. Data Cleaning and Preparation (Python)

Data cleaning and preparation were performed using Python and Pandas. Missing values in review ratings were handled using median imputation. Column names were standardized for better readability. Additional features such as age groups and customer segments were created to support deeper analysis.

4. Exploratory Data Analysis (EDA)

Exploratory data analysis was conducted to understand purchasing trends, category-wise sales distribution, subscription behavior, and customer segmentation. Visualizations were generated using Matplotlib to identify meaningful patterns and outliers.

5. SQL Analysis (PostgreSQL)

The cleaned dataset was loaded into a PostgreSQL database for structured analysis. SQL queries were written to answer key business questions such as revenue by gender, high-spending customers, top-rated products, shipping type comparison, subscriber behavior, discount dependency, and revenue contribution by age group.

6. Power BI Dashboard

An interactive Power BI dashboard was created to visualize important KPIs and insights. The dashboard includes charts for total customers, revenue distribution, category performance, subscription comparison, age-group analysis, and shipping type comparison.

7. Business Recommendations

Based on the analysis, it is recommended to strengthen subscription programs, optimize discount strategies, focus marketing efforts on high-revenue customer segments, and promote

top-performing products to maximize profitability.

8. Conclusion

This project demonstrates an end-to-end data analytics workflow including data cleaning, analysis, SQL-based querying, and dashboard visualization. It highlights practical skills in Python, SQL, and Power BI, making it suitable for academic and professional data analytics portfolios.