Customer Shopping Behavior Analysis – Project Report

1. Project Overview

This project analyzes 3,900 customer transactions across multiple product categories to uncover spending patterns, product preferences, subscription behavior, and customer segmentation. The insights aim to guide strategic business decisions and improve revenue, loyalty, and marketing effectiveness.

2. Dataset Summary

- **Rows:** 3,900

- **Columns:** 18

- **Key Features:** Customer demographics (Age, Gender, Location, Subscription Status), purchase details (Item, Category, Amount, Season, Size, Color), shopping behavior (Discount Applied, Promo Code, Previous Purchases, Frequency, Review Rating, Shipping Type)

- **Missing Data:** 37 values in Review Rating column

3. Methodology

Data Preparation & Python Analysis

- Imported data with **Pandas** and explored structure using `.info()` and `.describe()`.

- Imputed missing Review Ratings using **median by product category**.

- Standardized column names and engineered features: `age_group` and `purchase_frequency_days`.

- Dropped redundant columns (`promo_code_used`).

SQL Analysis (PostgreSQL)

- Revenue by gender, age group, and subscription status.

- Identified high-spending discount users.

- Top 5 products by average rating and top 3 per category.

- Customer segmentation into **New, Returning, Loyal**.

- Shipping type comparisons and repeat buyer analysis.

Visualization

- Built interactive dashboards in **Power BI** to highlight revenue trends, customer segments, and top-performing products.

4. Key Insights

- Male and female customers contribute similarly to revenue, with certain age groups driving higher sales.

- High-spending customers often leverage discounts, indicating effective promotions.

- Subscribers spend more consistently than non-subscribers.

- Repeat buyers are more likely to subscribe, suggesting loyalty potential.

- Top-rated products align closely with high-revenue items, indicating successful offerings.

5. Recommendations

1. **Boost Subscriptions:** Offer exclusive benefits to subscribers.

2. **Customer Loyalty Programs:** Reward repeat buyers to increase retention.

3. **Optimize Discount Strategy:** Balance promotions with profit margins.

4. **Product Positioning:** Highlight top-rated and high-revenue products.

5. **Targeted Marketing:** Focus campaigns on high-value age groups and frequent buyers.

6. Tools & Technologies

- **Python:** Pandas, NumPy

- **SQL:** PostgreSQL

- **Visualization:** Power BI