PROJECT REPORT: CUSTOMER SHOPPING BEHAVIOR ANALYSIS

Author: Nilesh Patil

Date: February 2025

Objective: Strategic Optimization of Retail Operations via Data Intelligence

1. EXECUTIVE SUMMARY

This project analyzes transactional data from **3,900 purchases** to uncover insights into spending patterns, customer segments, and product preferences. By integrating **Python**, **SQL**, and **Power BI**, the analysis provides actionable intelligence to solve the core business challenge: leveraging consumer data to identify trends and optimize marketing strategies.

2. BUSINESS PROBLEM STATEMENT

A leading retail company aims to improve sales and long-term loyalty by understanding shifting purchasing patterns across demographics and sales channels.

Core Question: *"How can the company leverage consumer shopping data to identify trends, improve customer engagement, and optimize marketing and product strategies?"*

3. TECHNICAL METHODOLOGY

3.1 Data Preparation & Modeling (Python)

The exploratory data analysis focused on ensuring a "SQL-friendly" and clean dataset:

- **Cleaning:** Addressed 37 missing values in the Review Rating column by applying category-wise median imputation.
- **Feature Engineering:**
  - **Age Segmentation:** Created four groups: young_adult, adult, middle_aged, and senior.
  - **Frequency Mapping:** Converted categorical frequency into numeric "purchase frequency days" (e.g., Weekly = 7, Annually = 365).
- **Optimization:** Removed redundant columns (e.g., promo_code_used) and standardized naming conventions for database compatibility.

3.2 Data Analysis & Business Transactions (SQL)

Structured queries were executed in **PostgreSQL** to extract high-value business insights:

- **Revenue Analysis:** Male customers generated **$157,890** in revenue, while Female customers contributed **$75,191**.
- **Category Performance:** Clothing was identified as the primary revenue driver, whereas Outerwear represents the lowest revenue contribution.
- **Discount Trends:** Hats (50%) and Sneakers (49.66%) lead the highest percentage of purchases with discounts applied.
- **Shipping Metrics:** Average spend for Express shipping ($60.48) is slightly higher than Standard shipping ($58.46).

3.3 Visualization & Insights (Power BI)

A dashboard was constructed to track critical KPIs:

- **Total Transactions:** 3.9K
- **Avg Purchase Amount:** $59.76
- **Avg Review Rating:** 3.75
- **Subscription Rate:** 27% (Yes) vs 73% (No).

4. STRATEGIC RECOMMENDATIONS

Based on the data trends, the following business strategies are proposed:

- **Subscription Growth:** With 73% of customers currently non-subscribers, there is a high opportunity to increase recurring revenue through targeted loyalty programs.
- **Segment Targeting:** Marketing campaigns should prioritize **Young Adult** and **Middle-Aged** segments, as they are the most active buyers.
- **Category Focus:** The **Outerwear** category requires specialized promotional strategies or targeted discounts to improve performance.
- **Quality Control:** Investigate products with lower review ratings to improve product quality and overall customer satisfaction.
- **Retention Strategy:** Implement a loyalty program specifically for "Repeat Buyers" (those with more than 5 previous purchases).

5. CONCLUSION

This multi-tool analysis provides a comprehensive roadmap for making data-driven business decisions. By focusing on the high-revenue Male demographic and converting the large pool of non-subscribers, the company can significantly enhance its market position and revenue growth.

Deliverables including Python scripts, SQL queries, and Power BI files are maintained in the project's GitHub Repository.