Customer Shopping Behavior Analysis

Uncovering insights from transactional data to drive strategic business decisions.

Project Overview

This project analyzes customer shopping behavior using transactional data from 3,900 purchases across various product categories. The goal is to uncover insights into spending patterns, customer segments, product preferences, and subscription behavior.

These insights will guide strategic business decisions, optimizing marketing efforts, product development, and customer retention strategies.

Dataset Summary

Our analysis is based on a comprehensive dataset with 3,900 rows and 18 columns, capturing a wide range of customer and purchase details.

Customer Demographics

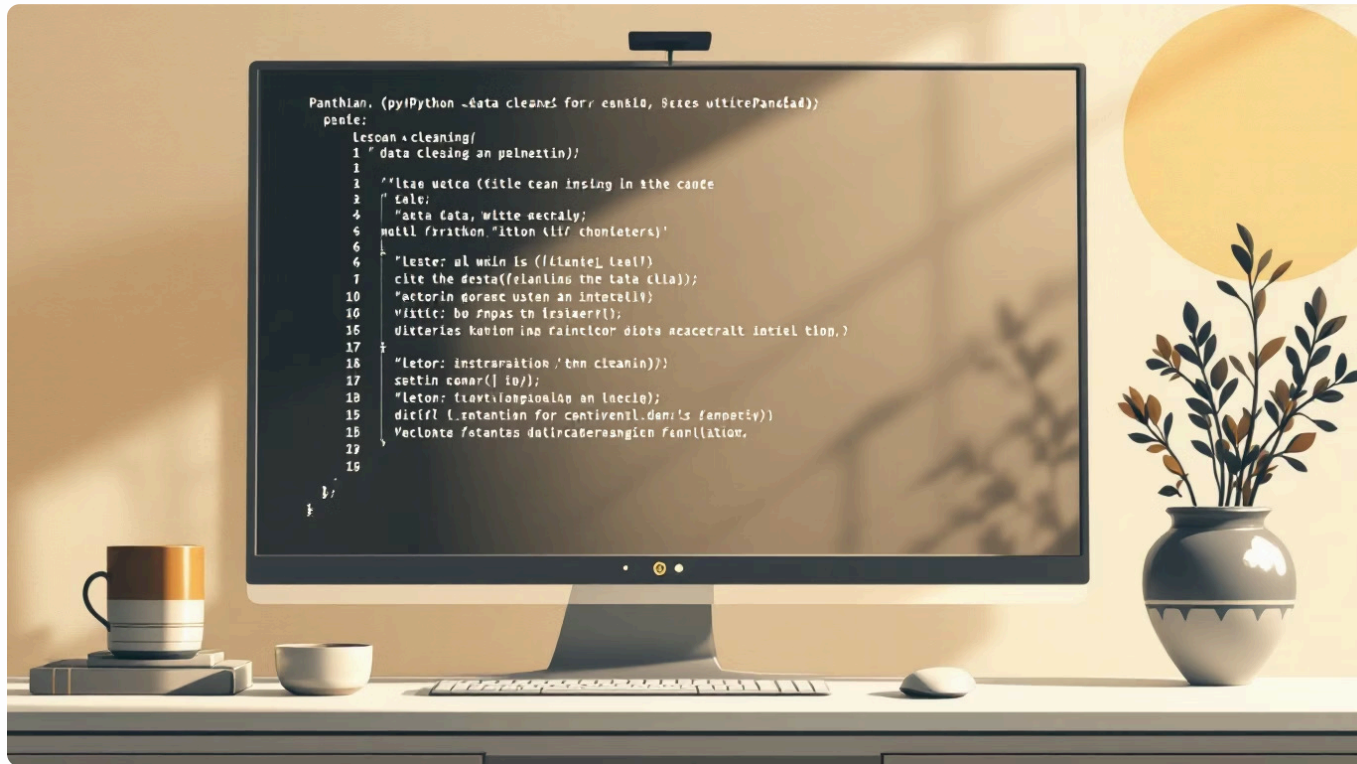Age, Gender, Location, Subscription Status

Purchase Details

Item Purchased, Category, Purchase Amount, Season, Size, Color

Shopping Behavior

Discount Applied, Promo Code Used, Previous Purchases, Frequency, Review Rating, Shipping Type

☐ Missing Data: 37 values in Review Rating column, imputed using median rating per category.

Made with GAMMA

Exploratory Data Analysis (Python)

We initiated our analysis with robust data preparation and cleaning using Python, ensuring data quality and consistency for deeper insights.

- **Data Loading:** Imported dataset with pandas.

- **Initial Exploration:** Used df.info() and .describe() for structural and statistical summaries.

- **Missing Data Handling:** Imputed missing Review Rating values using median per product category.

- **Feature Engineering:** Created age_group and purchase_frequency_days columns.

- **Database Integration:** Loaded cleaned data into PostgreSQL for SQL analysis.

Key Business Insights (SQL)

Structured SQL analysis revealed critical patterns in customer behavior and revenue generation.

Revenue by Gender

Male customers generated significantly more revenue ($157,890) than female customers ($75,191).

High-Spending Discount Users

Identified 839 customers who used discounts but still spent above the average purchase amount.

Top 5 Products by Rating

Gloves, Sandals, Boots, Hat, and Skirt received the highest average review ratings.

Shipping & Subscription Insights

Shipping Type Comparison

Express shipping users had a slightly higher average purchase amount ($60.48) compared to Standard shipping ($58.46).

Subscribers vs. Non-Subscribers

Non-subscribers contributed more total revenue ($170,436) than subscribers ($62,645), despite similar average spend.

Product & Customer Segmentation

1

Discount-Dependent Products

Hat, Sneakers, Coat, Sweater, and Pants had the highest percentage of discounted purchases.

2

Customer Segmentation

Customers classified into Loyal (3116), Returning (701), and New (83) segments based on purchase history.
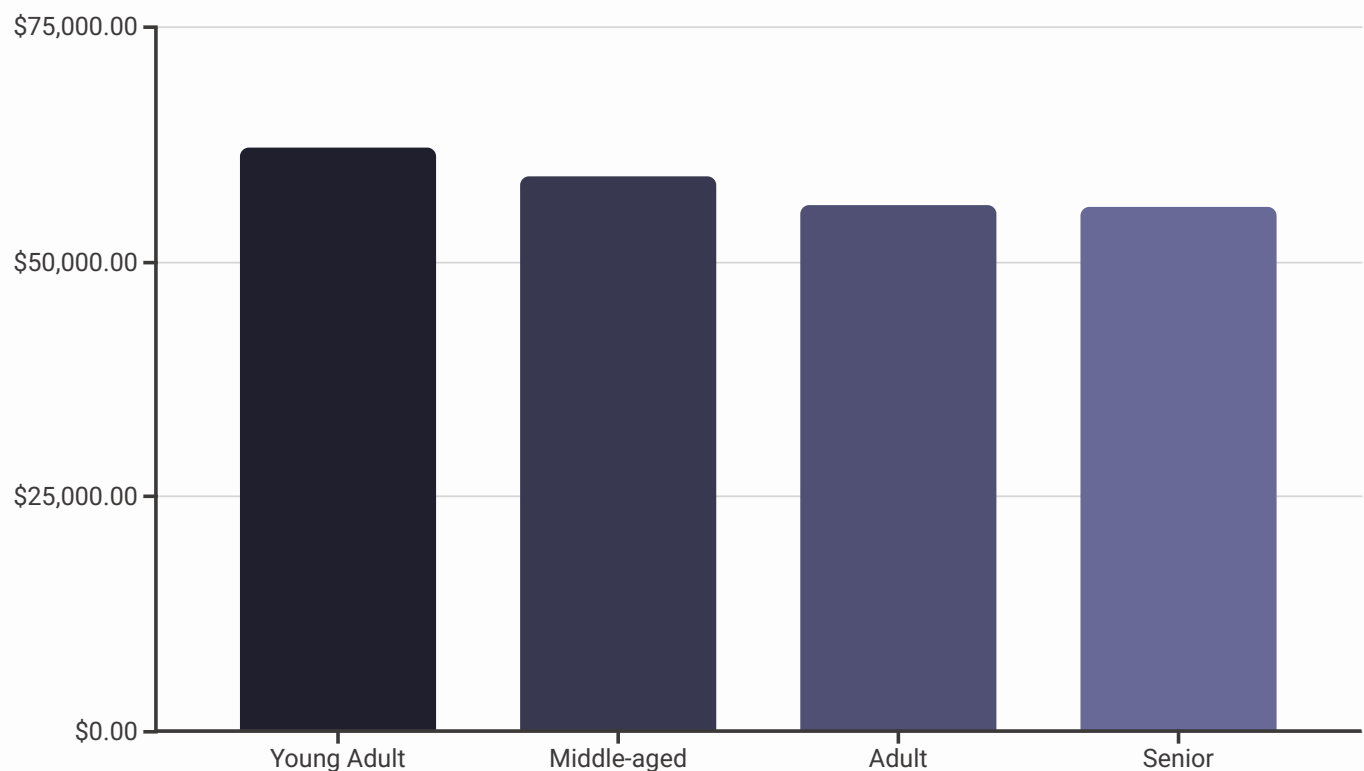
3

Repeat Buyers & Subscriptions

2518 non-subscribers were repeat buyers, compared to 958 subscribers, indicating an opportunity for conversion.

Top Products by Category

Accessories	Jewelry	171
Accessories	Sunglasses	161
Clothing	Blouse	171
Clothing	Pants	171
Footwear	Sandals	160
Footwear	Shoes	150
Outerwear	Jacket	163
Outerwear	Coat	161

Revenue by Age Group

Young Adults represent the highest revenue-generating segment, followed closely by Middle-aged customers. This highlights key demographics for targeted marketing efforts.

Understanding revenue contribution by age group allows for tailored product offerings and promotional strategies to maximize sales across all segments.

Made with GAMMA

Business Recommendations

Boost Subscriptions

Promote exclusive benefits to increase subscriber base.

Customer Loyalty Programs

Reward repeat buyers to foster loyalty and retention.

Targeted Marketing

Focus efforts on high-revenue age groups and express-shipping users.