

Customer Shopping Behavior Analysis

Uncovering insights from 3,900 purchases to guide strategic business decisions

~By Vineet Patel



Project Overview

This comprehensive analysis examines customer shopping behavior using transactional data from 3,900 purchases across various product categories.

Our goal: uncover insights into spending patterns, customer segments, product preferences, and subscription behavior to drive strategic business decisions.



Dataset at a Glance

3,900

Total Purchases

Transactions analyzed

18

Data Columns

Feature dimensions

37

Missing Values

In Review Rating only

Demographics

Age, Gender, Location, Subscription Status

Purchase Details

Item, Category, Amount, Season, Size, Color

Shopping Behavior

Discounts, Promo Codes, Purchase History, Reviews, Shipping

Data Preparation with Python

01

Data Loading & Exploration

Imported dataset using pandas, analyzed structure with `df.info()` and summary statistics

03

Column Standardization

Renamed columns to snake_case for better readability and documentation

05

Data Consistency

Verified redundancy between `discount_applied` and `promo_code_used`, dropped duplicate

02

Missing Data Handling

Imputed missing Review Rating values using median rating per product category

04

Feature Engineering

Created `age_group` and `purchase_frequency_days` columns for deeper analysis

06

Database Integration

Connected to MySQL and loaded cleaned DataFrame for SQL analysis

Revenue Insights

Revenue by Gender

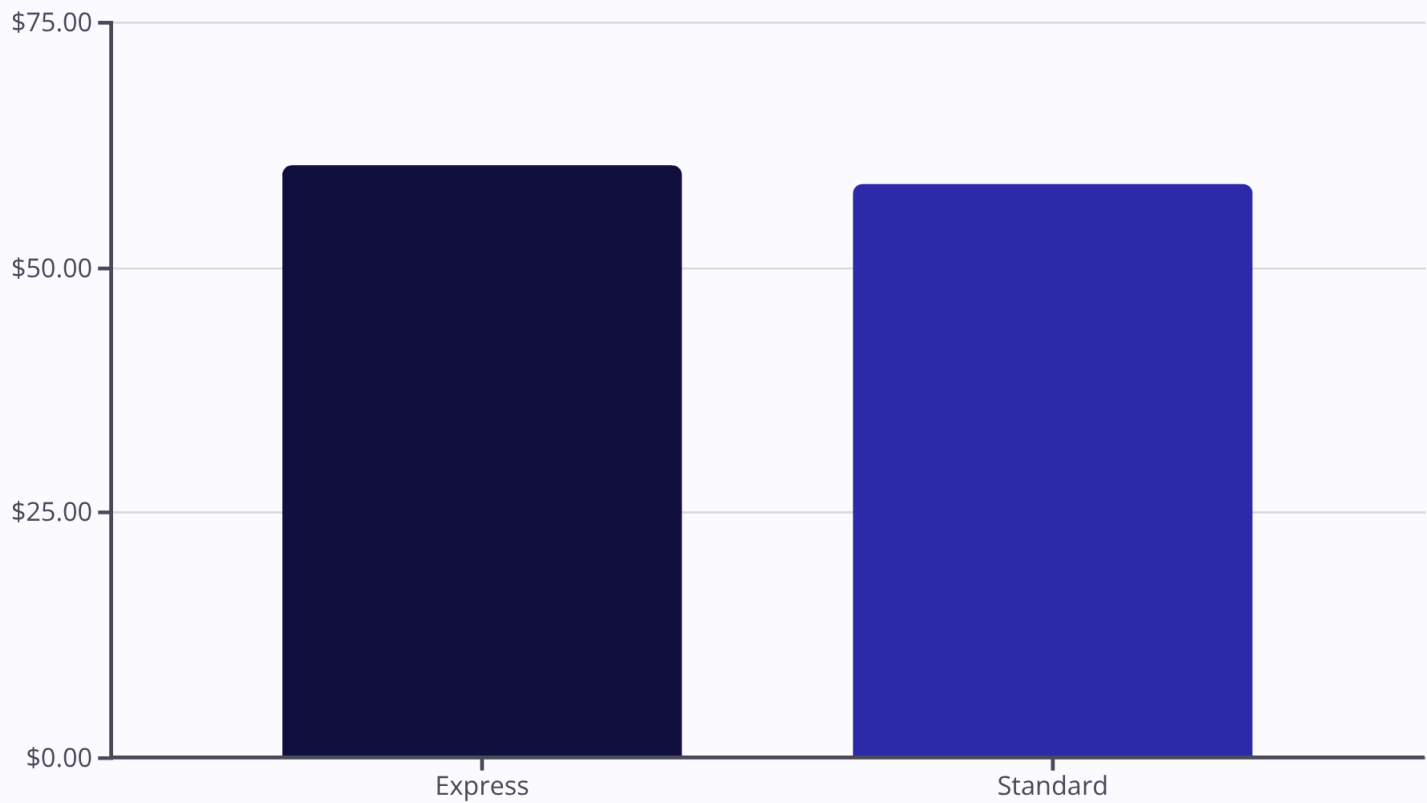
Male customers generated **\$157,890** in total revenue compared to **\$78,191** from female customers.



Shipping Type Impact

Express shipping customers spend an average of **\$60.48** versus **\$58.46** for Standard shipping.

Higher-value customers prefer faster delivery options.



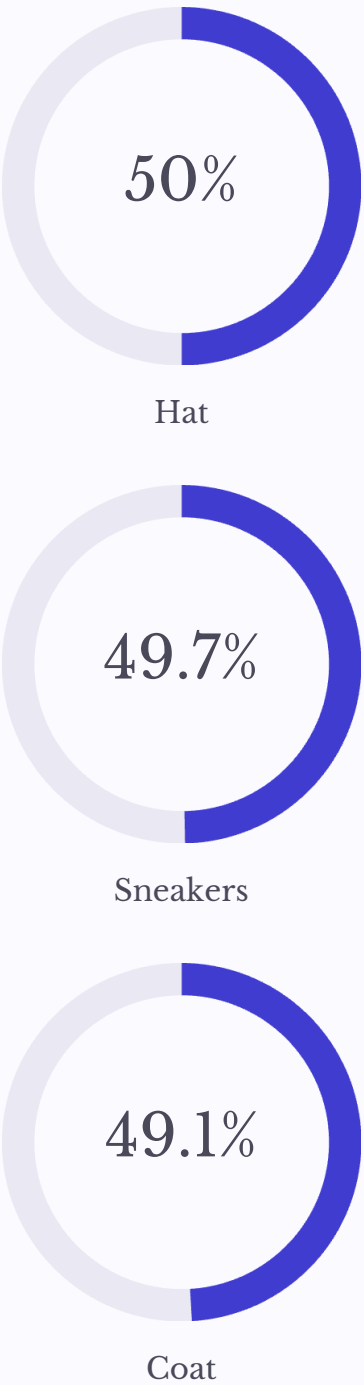
Product Performance Analysis

Top 5 Products by Rating



Discount-Dependent Products

Products with the highest percentage of discounted purchases:



Customer Segmentation



Customer classification based on purchase history reveals a strong loyal customer base representing **80% of total customers**.

Subscription Analysis

Subscribers: **1,053 customers** with \$59.49 average spend

Non-subscribers: **2,847 customers** with \$59.87 average spend

Repeat Buyer Behavior

Among customers with >5 purchases:

- 958 are subscribers
- 2,518 are non-subscribers

Top Products by Category

1

Accessories

Jewelry leads with 171 orders, followed by Sunglasses and Belt (161 each)

2

Clothing

Blouse and Pants tie at 171 orders, with Shirt close behind at 169 orders

3

Footwear

Sandals top the category with 160 orders, followed by Shoes (150) and Sneakers (145)

4

Outerwear


Jacket leads with 163 orders, followed closely by Coat at 161 orders





Interactive Power BI Dashboard

We built a comprehensive interactive dashboard in Power BI to visualize key insights and enable data-driven decision making.




Real-Time Insights

Dynamic visualizations of customer behavior and sales trends



Interactive Filters

Drill down by category, demographics, and time periods



KPI Tracking

Monitor key performance indicators and business metrics

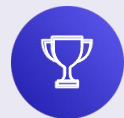


Business Recommendations



Boost Subscriptions

Promote exclusive benefits and perks for subscribers to increase conversion from non-subscribers



Customer Loyalty Programs

Reward repeat buyers to move them into the "Loyal" segment and increase lifetime value



Review Discount Policy

Balance sales boosts with margin control, especially for discount-dependent products



Product Positioning

Highlight top-rated and best-selling products in marketing campaigns



Targeted Marketing

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



Thank You!

We appreciate your time and attention today. We're happy to answer any questions you may have regarding our analysis and recommendations.