

Customer Shopping Behavior Analysis

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



Project Overview & Dataset

This analysis examines transactional data from 3,900 purchases across multiple product categories to uncover spending patterns, customer segments, product preferences, and subscription behavior.

Dataset Size

3,900 transactions with 18 key features

Data Coverage

Demographics, purchase details, shopping behavior, and customer preferences

Data Quality

37 missing values in Review Rating; cleaned and imputed using category medians



Data Preparation & Exploration

We began with comprehensive data cleaning and feature engineering in Python to prepare the dataset for analysis.

01

Data Loading

Imported 3,900 transactions using pandas and checked structure with df.info()

03

Feature Engineering

Created age_group bins and purchase_frequency_days columns for deeper segmentation

02

Missing Data Handling

Imputed 37 missing Review Rating values using median ratings by product category

04

Database Integration

Connected Python to PostgreSQL and loaded cleaned data for SQL analysis



Revenue & Customer Spending Insights

Analysis reveals significant differences in spending patterns across customer demographics and subscription status.

Revenue by Gender

Male customers generated **\$157,890** in revenue, while female customers contributed **\$75,191**—a 2:1 ratio indicating higher average spending among male shoppers.

Subscription Impact

Non-subscribers generated **\$170,436** total revenue (2,847 customers) vs. subscribers' **\$62,645** (1,053 customers), though average spend is similar at ~\$60 per transaction.

Customer Segmentation & Loyalty

Customers naturally segment into three distinct groups based on purchase history, with the majority classified as loyal repeat buyers.

3,116

Loyal Customers

Repeat buyers with established purchase patterns

701

Returning Customers

Customers with multiple purchases showing engagement

83

New Customers

First-time buyers representing growth opportunity

Notably, 958 repeat buyers (customers with >5 purchases) have subscriptions, while 2,518 do not—indicating subscription adoption remains a key growth lever.

Product Performance & Ratings

Top-performing products show consistent quality ratings, while certain items drive disproportionate discount dependency.

Highest-Rated Products

- Gloves: 3.86 average rating
- Sandals: 3.84 average rating
- Boots: 3.82 average rating
- Hat: 3.80 average rating
- Skirt: 3.78 average rating

Discount-Dependent Items

- Hat: 50.00% discounted
- Sneakers: 49.66% discounted
- Coat: 49.07% discounted
- Sweater: 48.17% discounted
- Pants: 47.37% discounted



Category Performance & Top Products

Analysis across four product categories reveals consistent top performers and category-specific purchasing patterns.



Accessories

Jewelry (171), Sunglasses (161), Belt (161)



Clothing

Blouse (171), Pants (171), Shirt (169)



Footwear

Sandals (160), Shoes (150), Sneakers (145)



Outerwear

Jacket (163), Coat (161), Vest (157)

Shipping & Discount Strategy Analysis

Shipping method and discount application significantly influence customer behavior and purchasing decisions across segments.

Shipping Type Impact

Express shipping customers spend slightly more at **\$60.48** average vs. Standard shipping at **\$58.46**. This suggests willingness to pay premium for faster delivery.

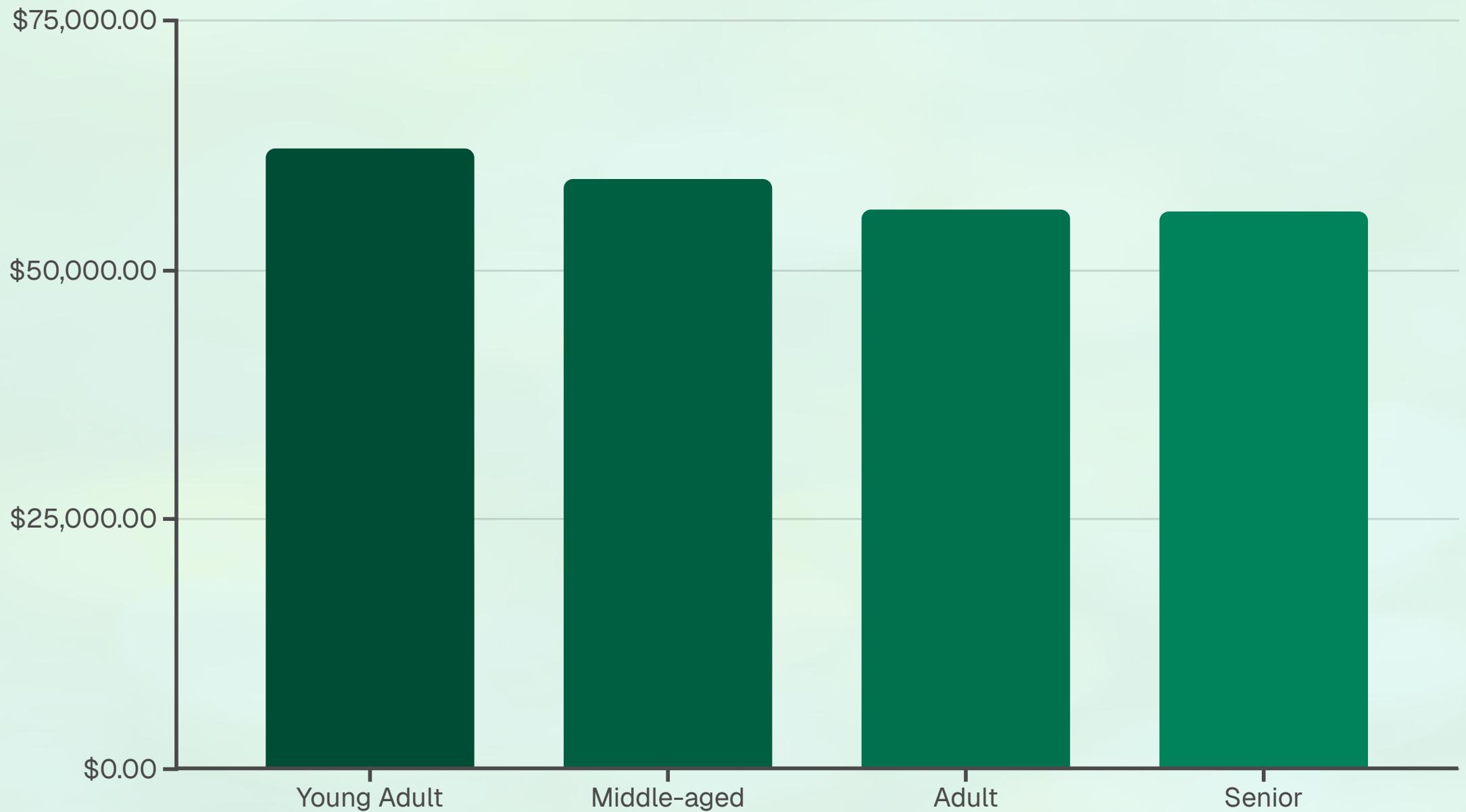
Discount Effectiveness

839 customers used discounts while spending above average purchase amount, demonstrating that strategic discounting drives volume without sacrificing margin on high-value transactions.



Age Group Revenue Contribution

Revenue distribution across age groups shows relatively balanced contribution, with Young Adults leading slightly, indicating diverse customer base across demographics.



Young Adults generate the highest revenue at \$62,143, followed closely by Middle-aged customers at \$59,197. This balanced distribution suggests effective market penetration across age demographics.

Strategic Recommendations

Based on comprehensive analysis, we recommend focused initiatives to drive growth, improve margins, and enhance customer lifetime value.

1 Boost Subscriptions

Promote exclusive benefits and loyalty rewards for subscribers to increase the 27% subscription rate and improve customer lifetime value.

2 Optimize Discount Strategy

Review discount policy on Hat, Sneakers, and Coat (50%+ discounted). Balance sales volume with margin protection through tiered promotions.

3 Targeted Marketing

Focus campaigns on high-revenue segments: male customers, Young Adults, and Express shipping users. Highlight top-rated products (Gloves, Sandals, Boots).

4 Loyalty Programs

Reward the 3,116 loyal customers and convert 701 returning customers into loyal segment through personalized incentives and exclusive offers.

