



Customer Shopping Behavior Analysis

Analyzing 3,900 purchases to uncover insights into spending patterns, customer segments, and product preferences that drive strategic business decisions.

Dataset Overview

3,900

Total Purchases

Transactions analyzed across all categories

18

Data Columns

Features covering demographics and behavior

50

Locations

Geographic distribution of customers

25

Product Types

Diverse items across four categories

Key features include customer demographics, purchase details, shopping behavior patterns, and subscription status. Only 37 missing values in Review Rating column.

Data Preparation Process

01

Data Loading & Exploration

Imported dataset using pandas and performed initial structure analysis with `df.info()` and summary statistics.

02

Missing Data Handling

Imputed missing Review Rating values using median rating of each product category.

03

Feature Engineering

Created `age_group` column by binning ages and `purchase_frequency_days` from purchase data.

04

Data Consistency

Verified redundancy between `discount_applied` and `promo_code_used`; dropped duplicate column.

05

Database Integration

Connected to PostgreSQL and loaded cleaned DataFrame for SQL analysis.

Revenue by Gender



Male Customers Drive Revenue

Male customers generated **\$157,890** in total revenue, more than double the **\$75,191** from female customers.

This significant gap suggests opportunities for targeted marketing to female segments.

High-Value Discount Users

839 Customers

Used discounts but spent above average purchase amount

Strategic Insight

Discounts attract high-value customers, not just bargain hunters

Opportunity

Balance promotional strategy with margin protection



Product Performance Insights

Top-Rated Products

1 Gloves

3.86 average rating

2 Sandals

3.84 average rating

3 Boots

3.82 average rating

4 Hat

3.80 average rating

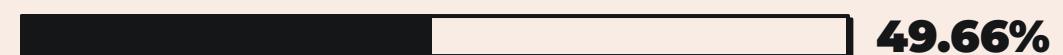
5 Skirt

3.78 average rating

Discount-Dependent Items

Hat  50%

Hat

Sneakers  49.66%

Sneakers

Coat  49.07%

Coat

Sweater  48.17%

Sweater

Pants  47.37%

Pants

Customer Segmentation



The majority of customers (80%) fall into the Loyal segment based on purchase history. Focus on moving Returning customers into Loyal status through targeted engagement.

Subscription Analysis

Subscription Status

Subscribers

1,053 customers

Avg spend: \$59.49

Total revenue: \$62,645

Non-Subscribers

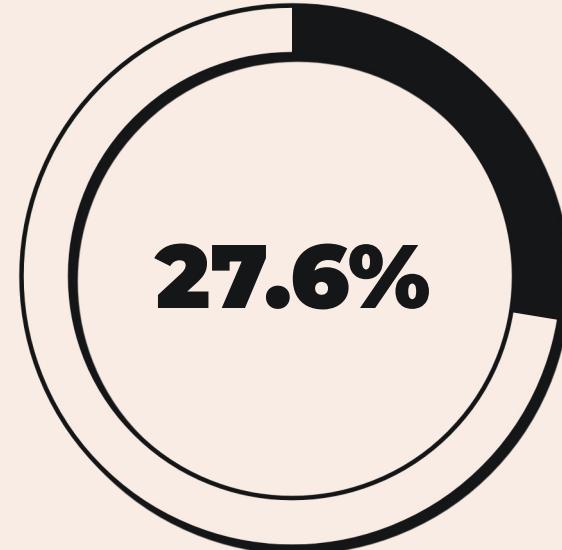
2,847 customers

Avg spend: \$59.87

Total revenue: \$170,436

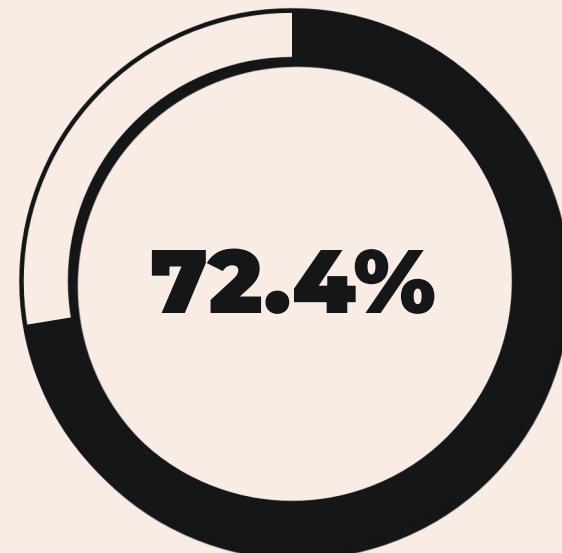
Repeat Buyers

Among customers with more than 5 purchases:



Subscribed

958 repeat buyers

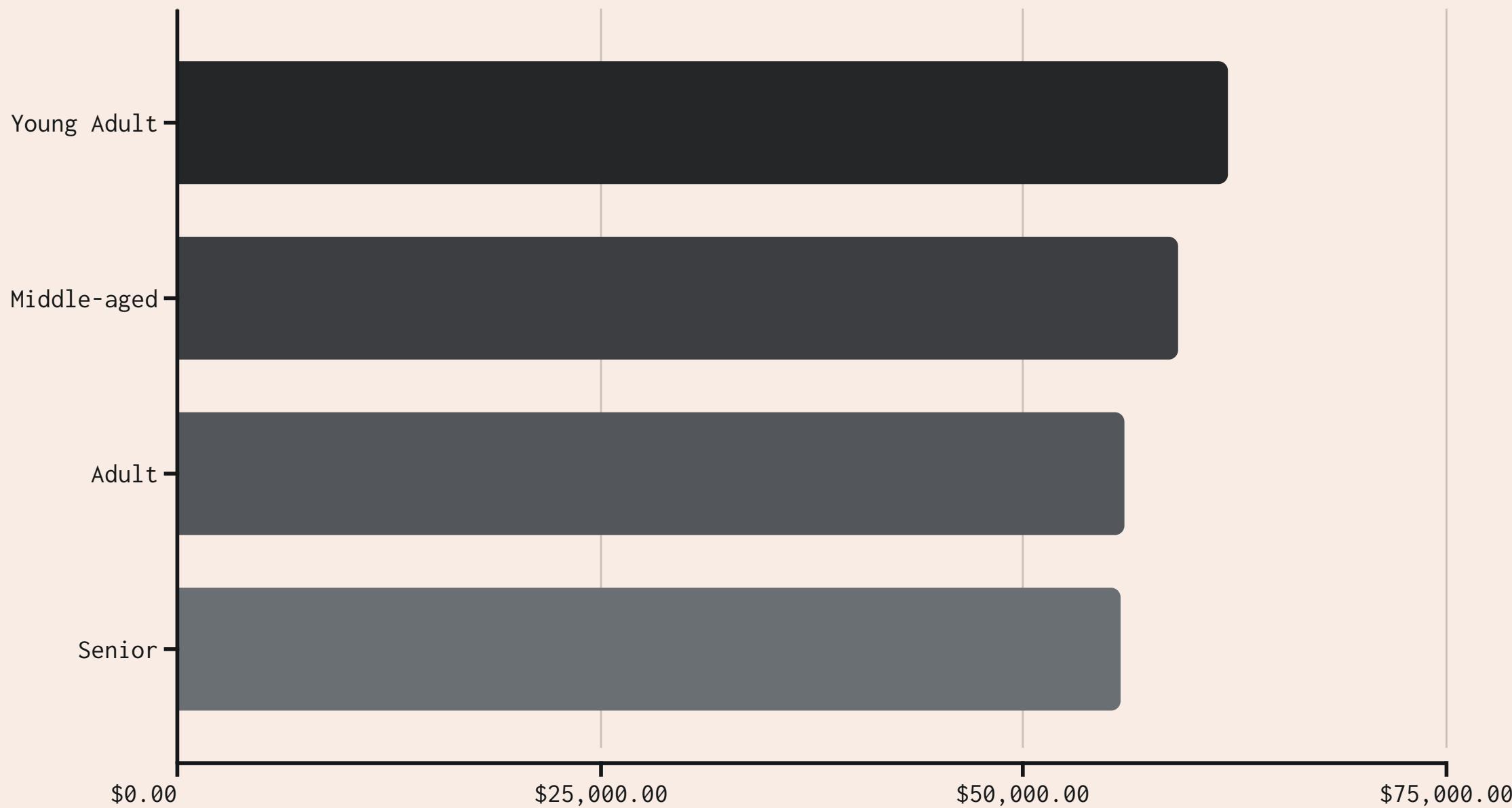


Not Subscribed

2,518 repeat buyers

Significant opportunity to convert loyal customers into subscribers.

Revenue by Age Group



Young Adults contribute the highest revenue at \$62,143, followed closely by Middle-aged customers. Revenue is relatively balanced across age groups, suggesting broad market appeal.

Strategic Recommendations



Boost Subscriptions

Promote exclusive benefits to convert 2,518 repeat buyers who haven't subscribed yet.



Loyalty Programs

Reward repeat buyers to accelerate movement into the Loyal segment.



Review Discount Policy

Balance promotional sales boosts with margin control for sustainable growth.



Product Positioning

Highlight top-rated products like Gloves and Sandals in marketing campaigns.



Targeted Marketing

Focus on high-revenue Young Adults and express-shipping users for maximum ROI.