



Customer Shopping Behavior Analysis

Uncovering insights from 3,900 customer transactions to drive strategic business decisions through data-driven analysis of spending patterns, customer segments, and product preferences.



Dataset Overview

3,900

Total Purchases

Transactions analyzed across all product categories

18

Data Features

Comprehensive attributes tracked per transaction

\$59.76

Average Purchase

Mean transaction value across all customers

3.75

Average Rating

Customer satisfaction score out of 5.0

The dataset captures customer demographics, purchase details, shopping behaviors, and product preferences. Key features include age, gender, location, subscription status, item categories, purchase amounts, discount usage, and review ratings. Only 37 missing values were found in the Review Rating column.

Data Preparation & Cleaning

Python Analysis Steps

We began with comprehensive data preparation using Python to ensure data quality and consistency:

- Imported dataset using pandas for analysis
- Explored structure with `.info()` and `.describe()` methods
- Imputed missing Review Rating values using median by category
- Standardized column names to snake_case format
- Created age_group feature for segmentation

Key Statistical Insights

Distribution analysis revealed important patterns:

- **Age Distribution:** Shapiro-Wilk test showed non-normal distribution, indicating concentration in specific age ranges
- **Purchase Amounts:** Uneven spending patterns with small group contributing disproportionately higher amounts
- **Database Integration:** Connected to PostgreSQL for advanced SQL analysis



Revenue Analysis by Gender

Male Customers

Total Revenue: \$157,890

Average Spend: \$59.54

Males represent the larger customer base and contribute significantly more to total revenue.

Female Customers

Total Revenue: \$75,191

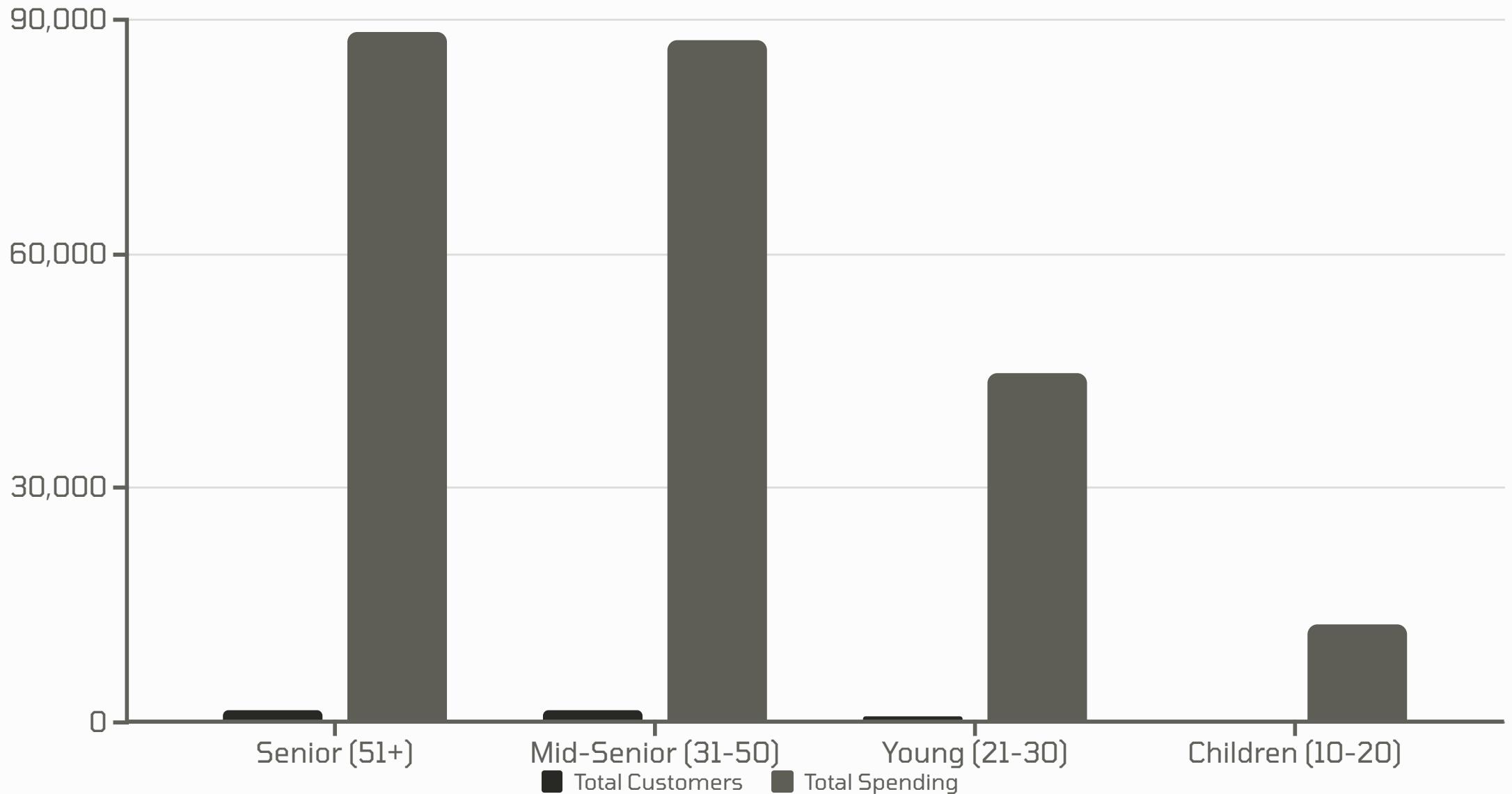
Average Spend: \$60.25

Females show slightly higher average spending per transaction despite lower total volume.

While male customers drive higher total revenue, female customers demonstrate marginally higher per-transaction spending. This suggests opportunities for targeted marketing to increase female customer acquisition while maintaining male customer engagement.



Customer Demographics & Spending



Senior and Mid-Senior age groups dominate both customer count and total spending, representing our core demographic. These segments contribute over 75% of total revenue, making them critical targets for retention and engagement strategies.



Category Performance Analysis

1

Clothing

Total Sales: \$104,264

Customers: 1,737

Avg Rating: 3.72

Largest category by revenue and customer base, representing 45% of total sales.

2

Accessories

Total Sales: \$74,200

Customers: 1,240

Avg Rating: 3.77

Second-highest performer with strong customer satisfaction scores.

3

Footwear

Total Sales: \$36,093

Customers: 599

Avg Rating: 3.79

Highest average rating indicates strong product quality and customer satisfaction.

4

Outerwear

Total Sales: \$18,524

Customers: 324

Avg Rating: 3.75

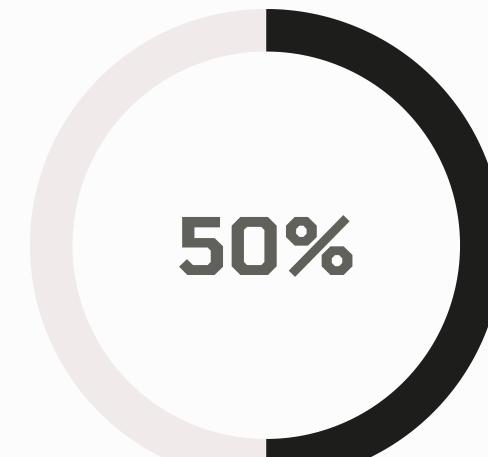
Smallest category but maintains solid customer satisfaction levels.

Top Performing Products

Highest Rated Items

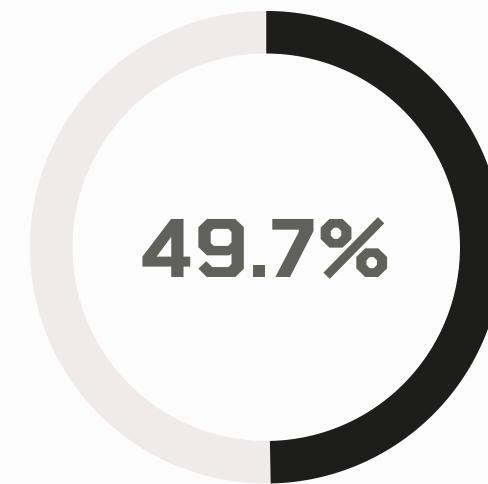
- Gloves - 3.86 rating
- Sandals - 3.84 rating
- Boots - 3.82 rating
- Hat - 3.80 rating
- Skirt - 3.79 rating

Discount Leaders



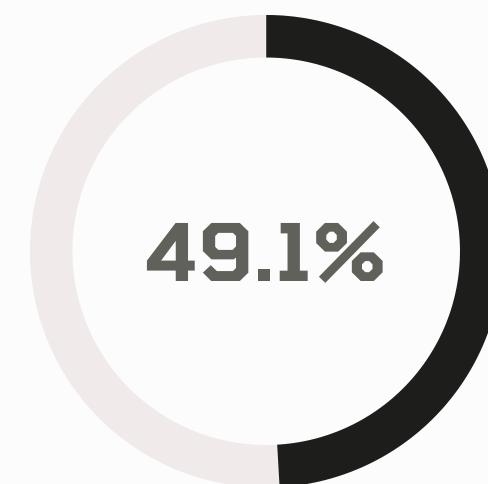
Hat

Highest discount usage rate



Sneakers

Second in promotional purchases



Coat

Strong discount attraction

These top-rated products should be prominently featured in marketing campaigns. The high discount usage rates suggest price sensitivity in certain categories, requiring careful balance between promotional strategies and margin protection.

Geographic & Shipping Insights

Top Markets

Montana leads with \$5,784 in sales, followed by Illinois (\$5,617) and California (\$5,605). These five states represent key geographic markets for targeted expansion.

Shipping Preferences

Standard shipping shows highest average rating (3.82), while Express and 2-Day options maintain strong performance (3.77). Express shipping users represent high-value customer segment.

High-Value Discount Users

A significant number of customers used discounts but still spent above the average purchase amount of \$59.76, demonstrating that promotional strategies can attract value-conscious yet high-spending customers without sacrificing revenue.

Customer Loyalty & Subscriptions



The customer base shows strong loyalty with 80% classified as loyal customers. However, subscription analysis reveals an opportunity:

Non-Subscribers

Average Spend: \$59.87

Total Revenue: \$170,436

Larger segment driving majority of revenue

Subscribers

Average Spend: \$59.49

Total Revenue: \$62,645

Opportunity to convert more customers

While spending patterns are similar, the lower subscriber base represents untapped potential for recurring revenue growth through enhanced subscription benefits and targeted conversion campaigns.

Strategic Recommendations

Boost Subscription Programs



Promote exclusive benefits and perks to convert non-subscribers. Focus on demonstrating value through early access, special discounts, and loyalty rewards.

Enhance Customer Loyalty



Implement reward programs for repeat purchases to move customers from returning to loyal segments, capitalizing on the strong existing loyalty base.

Optimize Discount Strategy



Balance promotional sales boosts with margin control. Target discounts strategically to high-value customers while protecting profitability.

Highlight Top Products



Feature highest-rated and best-selling items prominently in marketing campaigns, particularly Gloves, Sandals, and Boots.

Targeted Marketing Focus



Concentrate efforts on high-revenue Senior and Mid-Senior age groups, express-shipping users, and top geographic markets like Montana and Illinois.