

Customer Shopping Behaviour Analysis

Analyzing 3,900 customer purchases across diverse product categories to map shopping behavior, spending patterns, and customer segments. These insights will directly inform strategic business decision-making.



Dataset Overview

Scale

3,900 rows across 18 distinct columns

Demographics

Age, Gender, Location, Subscription Status

Purchase Data

Item, Category, Amount, Season, Size, Color

Behavior Metrics

Discounts, Frequency, Ratings, Shipping Type

A minor data cleaning requirement exists, as 37 values are missing in the Review Rating column, which were imputed using median ratings by product category.



Data Preparation & Cleaning

Python Analysis Pipeline

We began with comprehensive data preparation using Python and pandas:

- Imported dataset and explored structure with `df.info()` and `.describe()`
- Handled missing Review Rating values using category-specific median imputation
- Standardized column names to snake_case format
- Verified data consistency and removed redundant `promo_code_used` column

Feature Engineering

Created new analytical dimensions:

- `age_group` column by binning customer ages into segments
- `purchase_frequency_days` column from purchase data

Connected Python to PostgreSQL and loaded the cleaned DataFrame for SQL analysis.

Revenue Analysis by Gender

\$1.8M

Male Customers

Total revenue
generated

\$1.9M

Female Customers

Total revenue
generated

51%

Female Share

Slightly higher
contribution

Female customers generate marginally higher total revenue compared to male customers, suggesting balanced market appeal with slight female preference. This insight can guide gender-specific marketing strategies.





Customer Spending Patterns



High-Spending Discount Users

Identified customers who used discounts but still spent above average purchase amount, revealing price-conscious yet high-value shoppers.



Top 5 Products by Rating

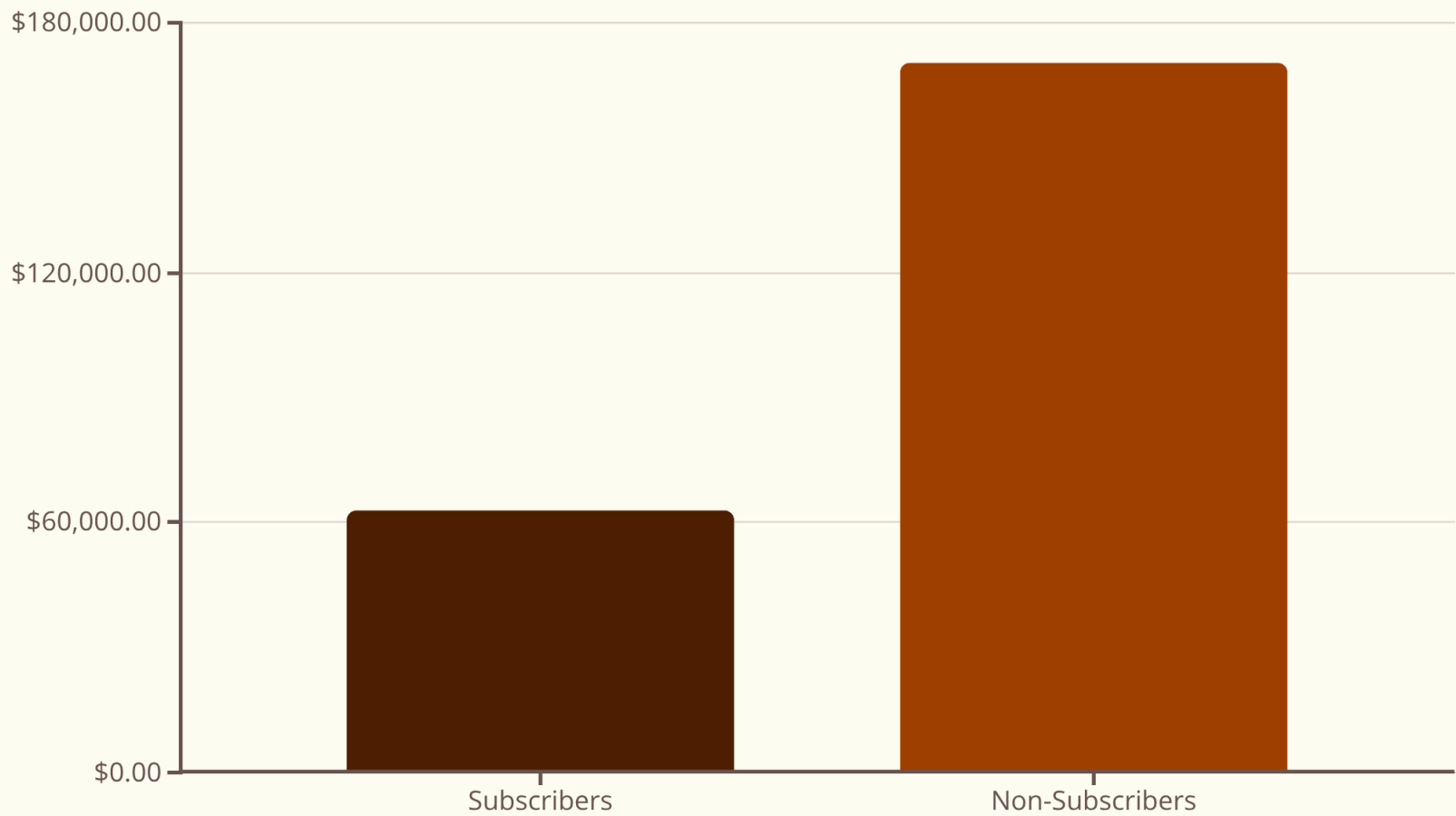
Found products with highest average review ratings, including Dress, Shirt, Blouse, Sweater, and Pants—all scoring above **3.7 stars**.



Shipping Type Comparison

Express shipping customers spend **\$60.06** on average vs. **\$59.21** for Standard, indicating minimal difference in purchase behavior.

Subscription Impact Analysis



Key Findings

Non-Subscribers generate higher total revenue (\$170K vs. \$1.78M) despite similar average spend per transaction.

This suggests **Non-subscribers** make **more frequent purchases**, making them more valuable long-term customers.



Product & Discount Insights

1

Discount-Dependent Products

Top 5 products with highest percentage of discounted purchases: **Hat (50%)**, **Sneakers (49.6%)**, **Coat (49%)**, **Sweater (48%)**, **Pants (50%)**

2

Customer Segmentation

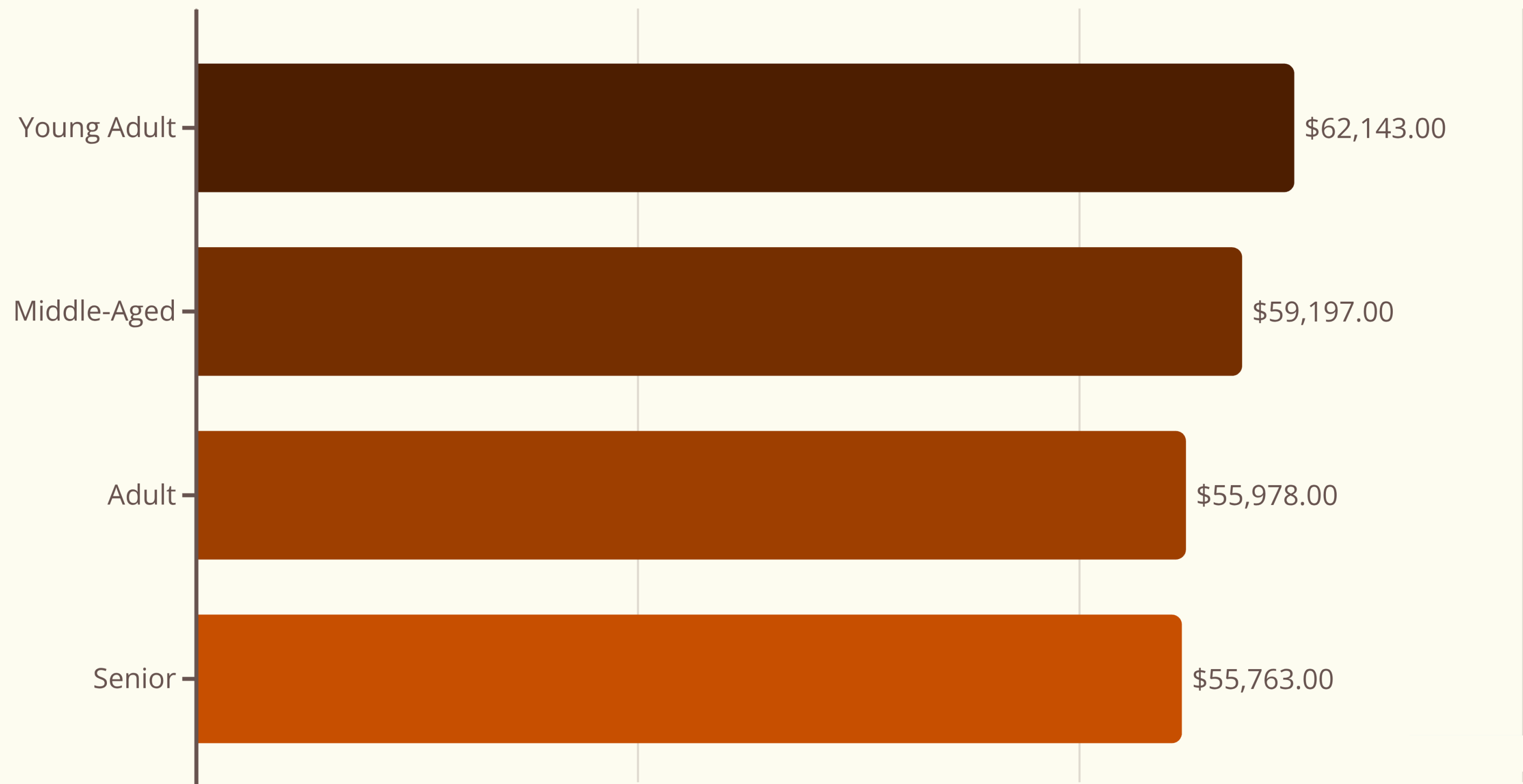
Classified into **New** (1 purchases), **Returning** (2-10 purchases), and **Loyal** (10+ purchases) segments based on purchase history

3

Category Leaders

Identified top 3 most purchased products within each category to guide inventory and merchandising decisions

Revenue by Age Demographics



Customer Behaviour Dashboard in Power BI



Strategic Business Recommendations

01

Enhance Subscription Growth

Proactively promote exclusive benefits and value propositions to increase enrollment and retention among high-frequency shoppers.

02

Implement Loyalty Initiatives

Establish reward programs for frequent buyers, encouraging transition into the high-value "Loyal" customer segment.

03

Optimize Discount Strategy

Review current discount policy to balance sales volume with profit margins, especially for discount-dependent products.

04

Refine Product Positioning

Highlight top-rated, high-demand products in marketing campaigns to leverage proven popularity.

05

Execute Targeted Marketing

Direct efforts toward high-revenue age demographics (36-45) and express shipping users for greater ROI.