



# Customer Shopping Behaviour Analysis

Uncovering insights from 3,900 purchases to guide 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 categories

**18**

## Data Columns

Comprehensive customer  
and purchase attributes

**50**

## Locations

Geographic diversity in  
customer base

**25**

## Products

Unique items across four  
categories

## Customer Demographics

- Age, Gender, Location
- Subscription Status
- Purchase History

## Shopping Behavior

- Discount & Promo Usage
- Review Ratings
- Shipping Preferences



# Data Preparation & Analysis

01

## Data Loading

Imported dataset using pandas and performed initial exploration with `df.info()` and `.describe()` for summary statistics.

02

## Data Cleaning

Handled 37 missing values in Review Rating column using median imputation by product category.

03

## Feature Engineering

Created `age_group` and `purchase_frequency_days` columns. Standardized column names to `snake_case` format.

04

## Database Integration

Connected to PostgreSQL and loaded cleaned data for structured SQL analysis.

# Key Statistical Insights

## Customer Profile

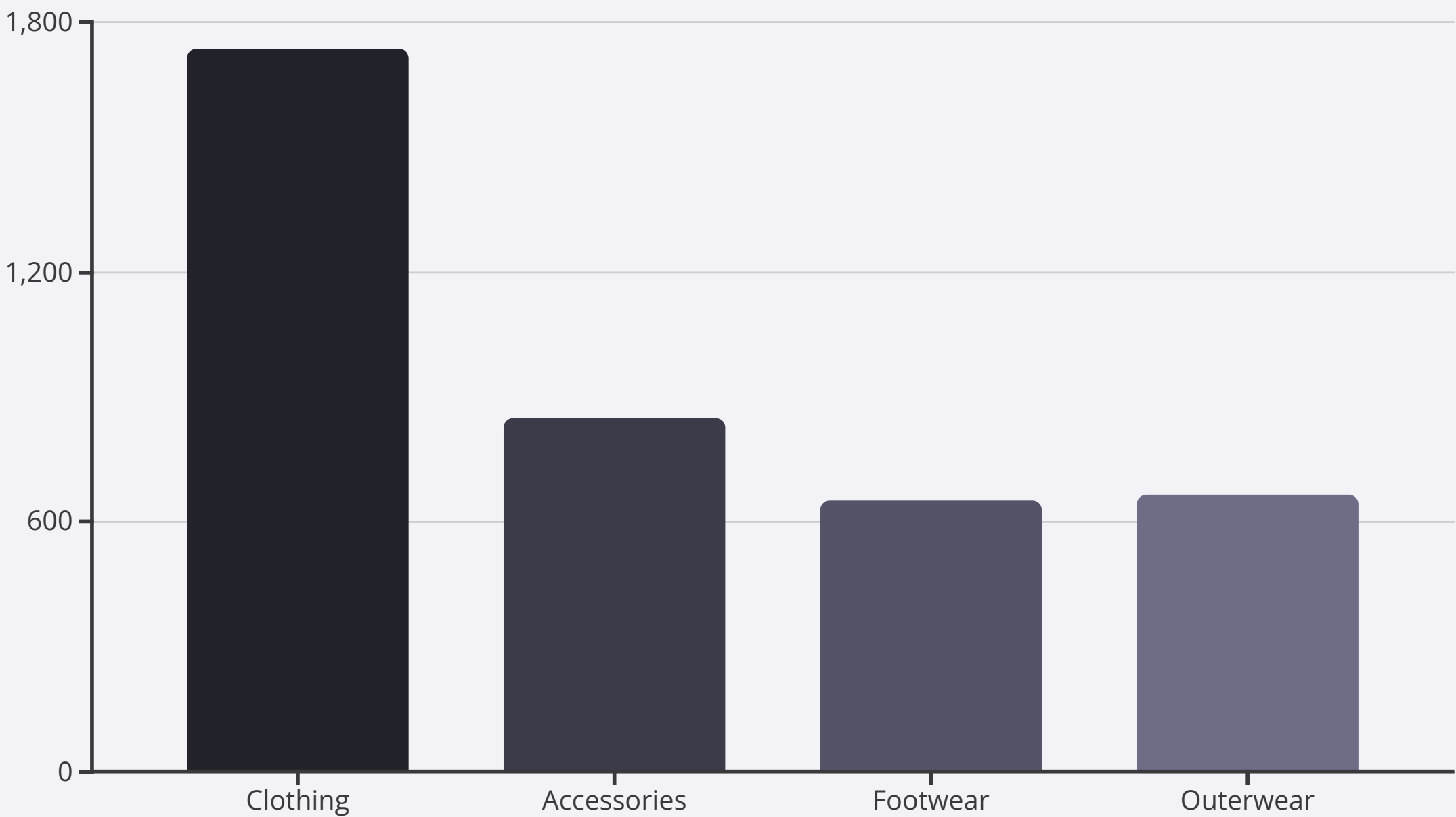
**Average Age:** 44 years  
**Age Range:** 18-70 years  
**Gender Split:** 68% Male, 32% Female

## Purchase Behavior

**Avg Amount:** \$59.76  
**Price Range:** \$20-\$100  
**Avg Rating:** 3.75/5.0

## Engagement

**Avg Previous Purchases:** 25  
**Subscribers:** 27%  
**Discount Users:** 43%



# Revenue Analysis by Gender

## Male Customers

**\$157,890** total revenue

68% of customer base driving 68% of total revenue

## Female Customers

**\$75,191** total revenue

32% of customer base contributing 32% of revenue

Revenue distribution aligns proportionally with customer demographics, indicating consistent spending patterns across genders.



# Discount Strategy Insights

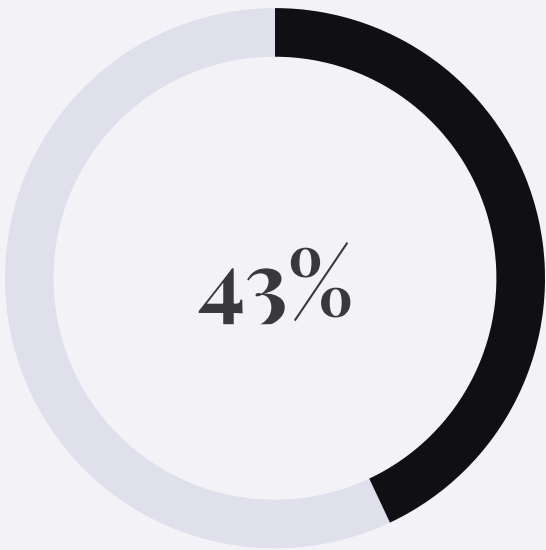
## High-Spending Discount Users

**839 customers** used discounts but still spent above the average purchase amount of \$59.76.

This segment represents strategic buyers who maximize value without compromising on quality or quantity.

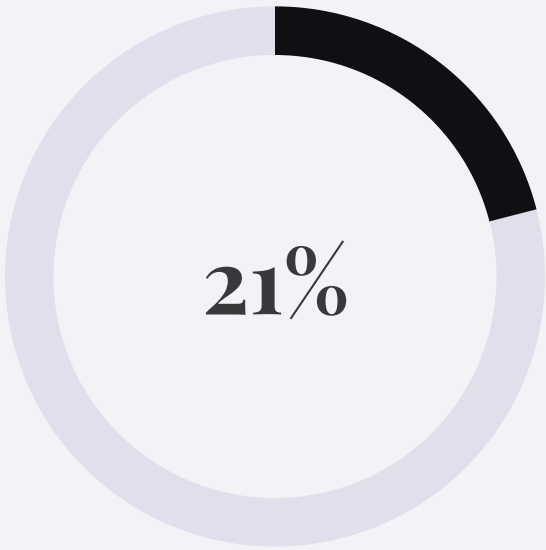
## Top Discount-Dependent Products

- 1. Hat (50% discount rate)
- 2. Sneakers (49.66%)
- 3. Coat (49.07%)
- 4. Sweater (48.17%)
- 5. Pants (47.37%)



Discount Usage

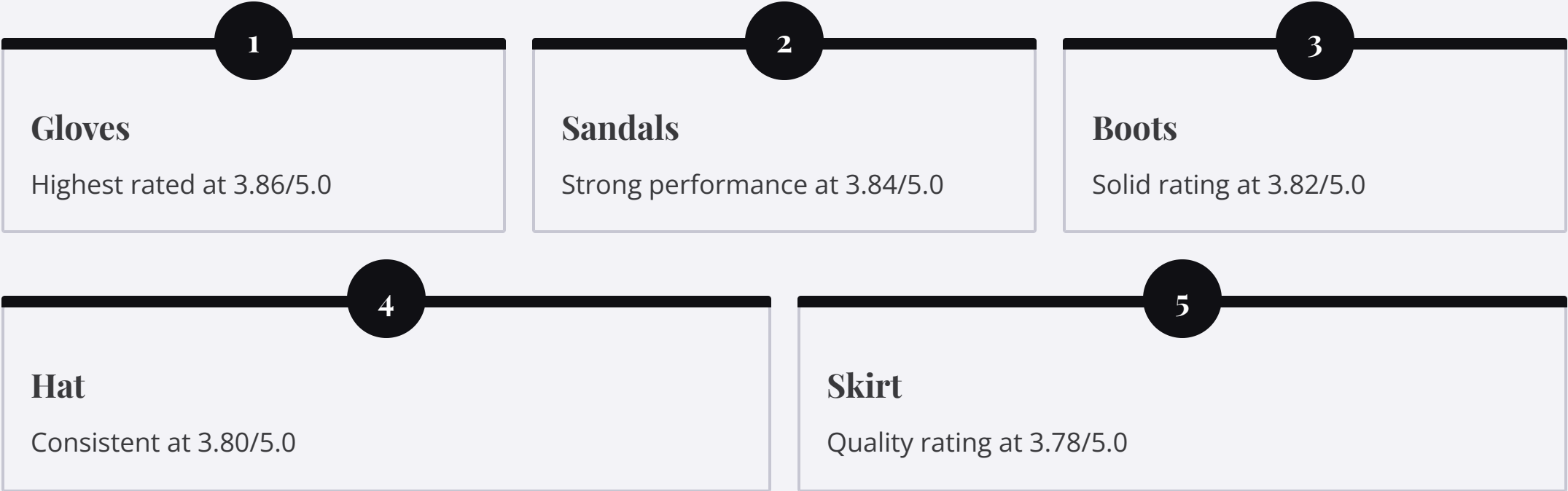
Of all purchases



Above Average

Discount users spending more

# Product Performance & Ratings



## Shipping Preferences

**Express:** \$60.48 avg spend

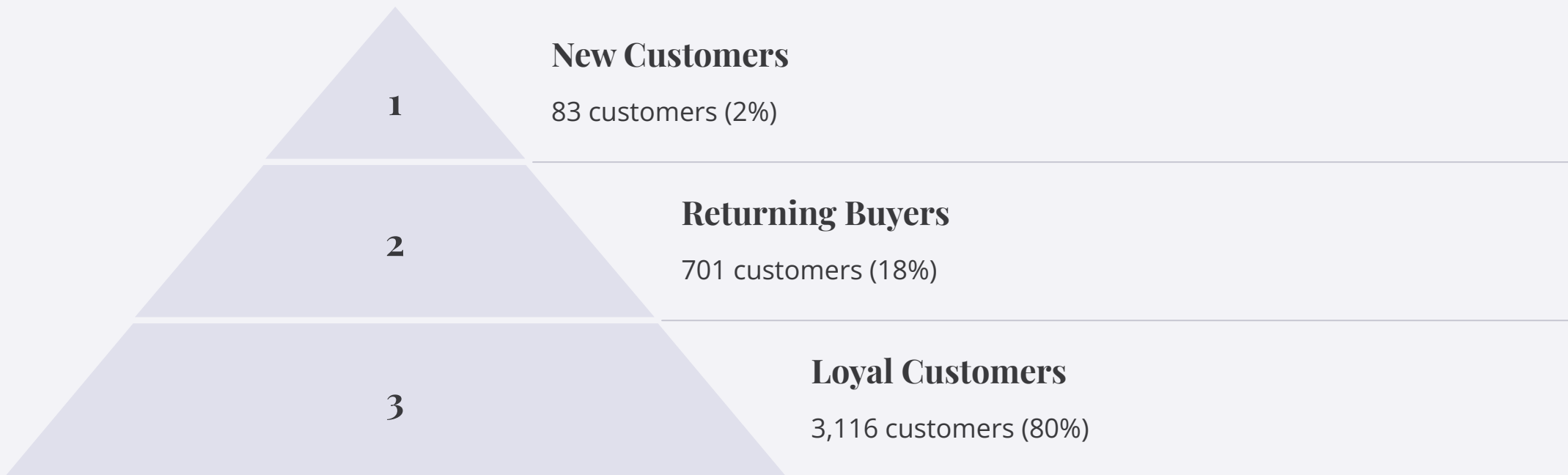
**Standard:** \$58.46 avg spend

Express shipping users spend 3.5% more per transaction.





# Customer Segmentation Analysis



## Subscription Impact

**Subscribers:** 1,053 customers

Average spend: \$59.49

Total revenue: \$62,645

## Non-Subscribers

**Non-subscribers:** 2,847 customers

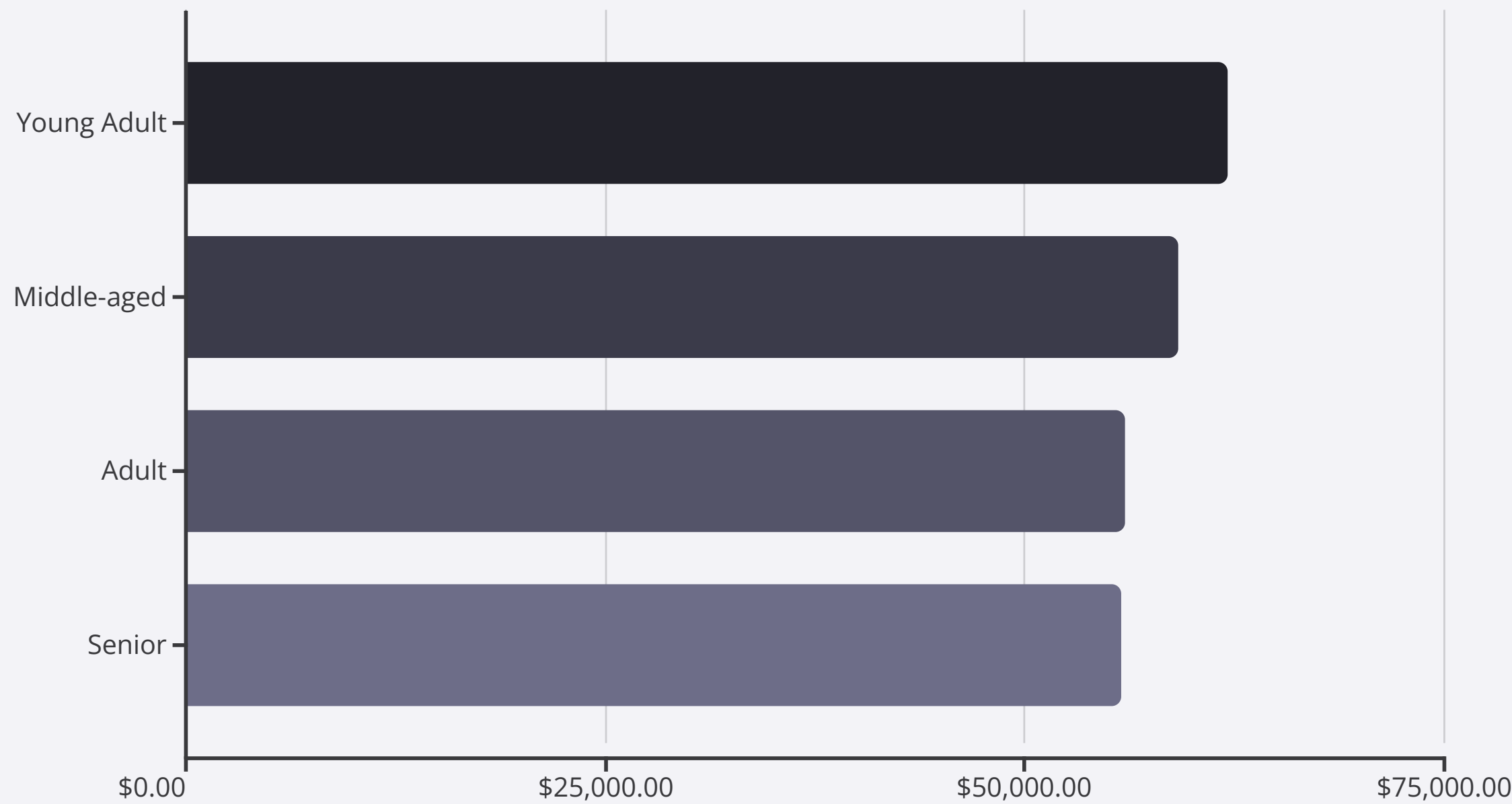
Average spend: \$59.87

Total revenue: \$170,436

Among repeat buyers (>5 purchases), 958 are subscribers while 2,518 are not—indicating significant opportunity for subscription conversion.



# Revenue by Age Demographics



Young adults lead revenue generation at \$62,143, followed closely by middle-aged customers at \$59,197. Revenue distribution is relatively balanced across age groups, suggesting broad market appeal.

# Strategic Recommendations



## Boost Subscriptions

Promote exclusive benefits to convert the 2,518 repeat buyers who aren't yet subscribers.



## Loyalty Programs

Reward repeat buyers to accelerate movement from Returning to Loyal segment.



## Optimize Discounts

Balance promotional sales with margin control, especially for high-discount products like hats and sneakers.



## Product Positioning

Highlight top-rated products (gloves, sandals, boots) in marketing campaigns.



## Targeted Marketing

Focus on high-revenue young adults and express-shipping users for premium offerings.