

Customer Shopping Behavior Analysis

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



Dataset Overview

3,900

Total Purchases

Transactions analyzed across
all categories

18

Data Points

Features tracked per
customer

50

Locations

Geographic coverage

25

Products

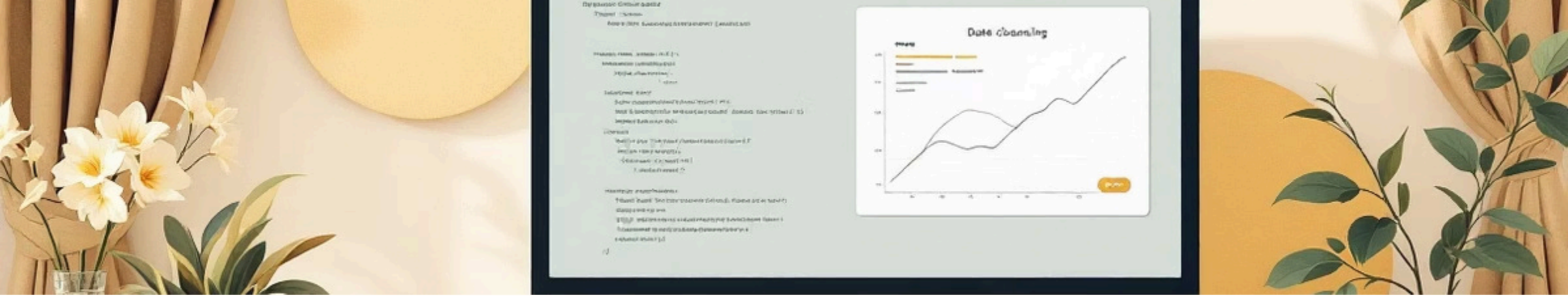
Unique items purchased

Customer Data

- Demographics & location
- Subscription status
- Purchase history

Transaction Details

- Product & category
- Pricing & discounts
- Shipping & ratings



Data Preparation Process

01

Data Loading

Imported dataset using pandas, explored structure with `df.info()` and `.describe()`

03

Feature Engineering

Created `age_group` and `purchase_frequency_days` columns

02

Missing Data Handling

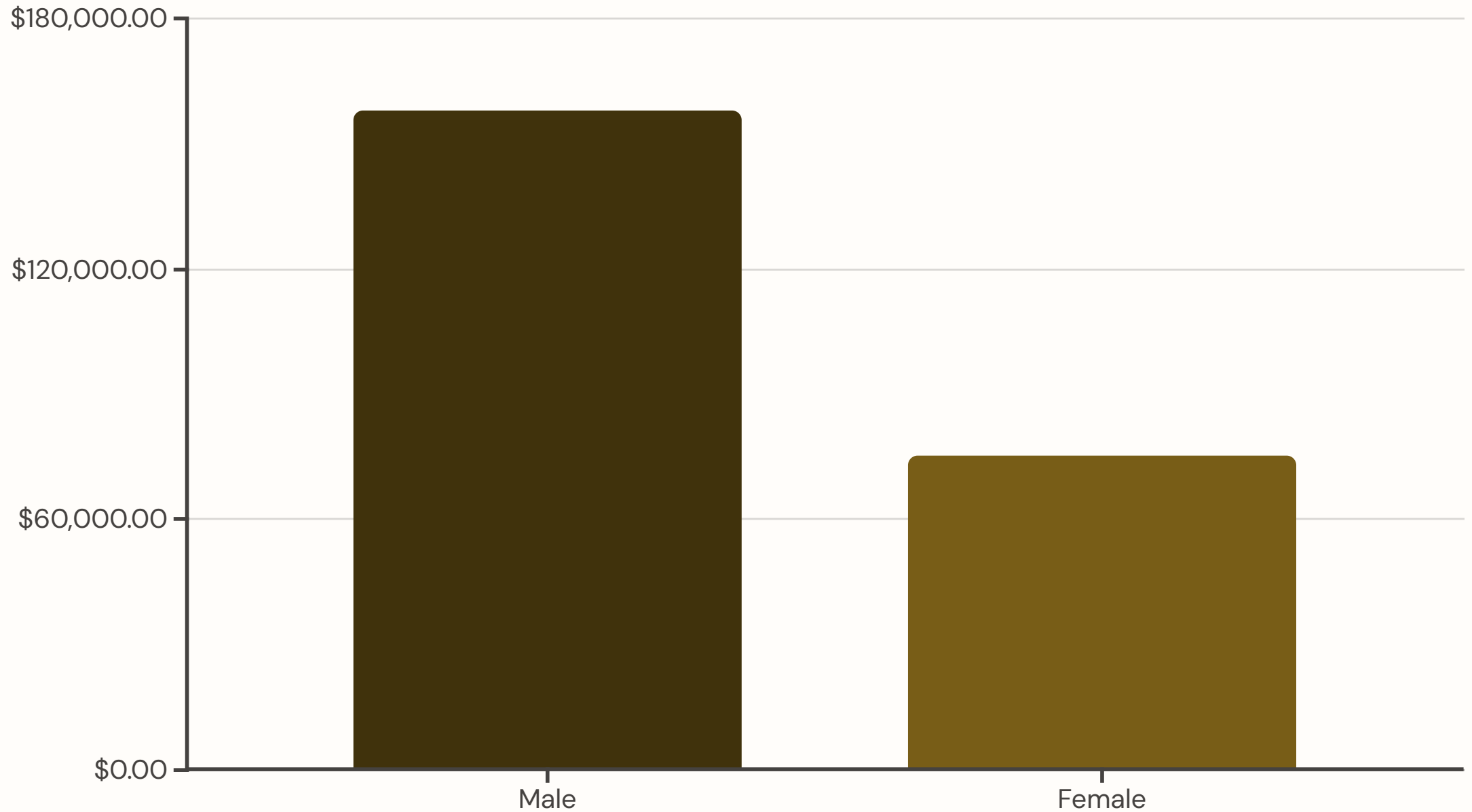
Imputed 37 missing Review Rating values using median by category

04

Database Integration

Connected to PostgreSQL for advanced SQL analysis

Revenue by Gender



Male customers generate **2.1x more revenue** than female customers, representing a significant opportunity for targeted marketing campaigns.

Customer Segmentation Insights

Loyal Customers

3,116 customers

Largest segment driving repeat business

Returning Buyers

701 customers

Growth opportunity segment

New Customers

83 customers

Smallest but critical for expansion

Subscription Analysis

Customers with 5+ purchases show strong subscription potential

- 958 repeat buyers are subscribers
- 2,518 repeat buyers not yet subscribed
- Massive conversion opportunity



Top Performing Products

Gloves

Highest rated: **3.86 stars**

Sandals

Second best: **3.84 stars**

Boots

Strong performer: **3.82 stars**

Top Products by Category

Accessories: Jewelry (171 orders)

Clothing: Blouse & Pants (171 orders each)

Footwear: Sandals (160 orders)

Outerwear: Jacket (163 orders)

Discount Strategy Analysis

High-Spending Discount Users

839 customers used discounts but spent above average

Average purchase: **\$59.76**

Discount-Dependent Products

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

📌 **Key Finding:** Discounts drive sales but need careful margin management. Nearly half of certain products sold with discounts.



Shipping & Subscription Comparison



Express Shipping

Average purchase amount



Standard Shipping

Average purchase amount

Subscribers

1,053 customers

Avg spend: **\$59.49**

Total revenue: **\$62,645**

Non-Subscribers

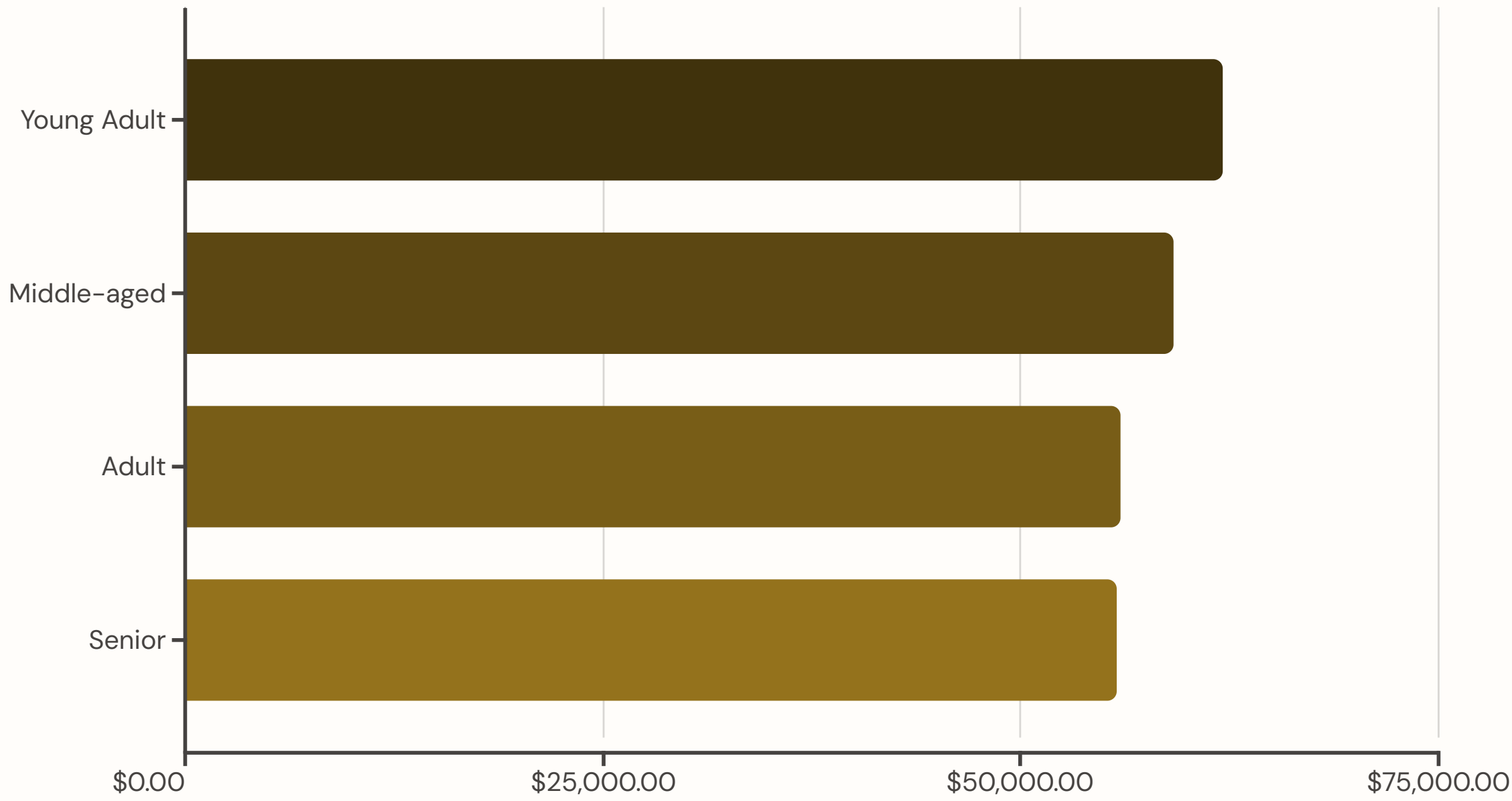
2,847 customers

Avg spend: **\$59.87**

Total revenue: **\$170,436**

Express shipping users spend slightly more. Non-subscribers represent **73%** of customer base with massive subscription conversion potential.

Revenue by Age Group



Young adults lead revenue generation, but all segments show balanced contribution. Target marketing by age group for maximum impact.

Strategic Recommendations



Boost Subscriptions

Promote exclusive benefits to convert 2,518 repeat non-subscribers



Loyalty Programs

Reward repeat buyers to grow the loyal customer segment



Optimize Discounts

Balance sales boosts with margin control on high-discount products



Targeted Marketing

Focus on high-revenue age groups and express shipping users



Product Positioning

Highlight top-rated items like Gloves and Sandals in campaigns