

# Customer Shopping Behavior Analysis

Understanding spending habits, product preferences, and customer types across 3,900 purchase records



# Dataset Overview

3,900

## Purchase Records

Comprehensive customer transactions analyzed

18

## Data Columns

Rich attributes covering customer and purchase details

37

## Missing Values

Only in Review Rating column

### Customer Details

Age, Gender, Location, Subscription Status

### Purchase Info

Item, Category, Amount, Season, Size, Color

### Shopping Behavior

Discounts, Promo Codes, Ratings, Shipping, Previous Purchases

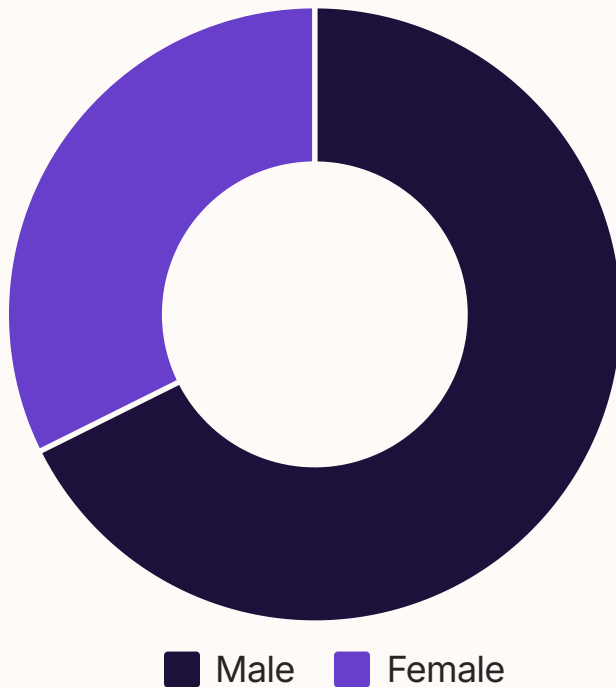
# Data Cleaning & Preparation

All data preparation performed in Python using Pandas and NumPy

|   | Customer ID | Age | Gender | Item Purchased | Category | Purchase Amount (USD) | Location      | Size | Color     | Season | Review Rating | Subscription Status | Shipping Type | Discount Applied | Promo Code Used | Previous Purchases | Payment Method |
|---|-------------|-----|--------|----------------|----------|-----------------------|---------------|------|-----------|--------|---------------|---------------------|---------------|------------------|-----------------|--------------------|----------------|
| 0 | 1           | 55  | Male   | Blouse         | Clothing | 53                    | Kentucky      | L    | Gray      | Winter | 3.1           | Yes                 | Express       | Yes              | Yes             | 14                 | Venmo          |
| 1 | 2           | 19  | Male   | Sweater        | Clothing | 64                    | Maine         | L    | Maroon    | Winter | 3.1           | Yes                 | Express       | Yes              | Yes             | 2                  | Cash           |
| 2 | 3           | 50  | Male   | Jeans          | Clothing | 73                    | Massachusetts | S    | Maroon    | Spring | 3.1           | Yes                 | Free Shipping | Yes              | Yes             | 23                 | Credit Card    |
| 3 | 4           | 21  | Male   | Sandals        | Footwear | 90                    | Rhode Island  | M    | Maroon    | Spring | 3.5           | Yes                 | Next Day Air  | Yes              | Yes             | 49                 | PayPal         |
| 4 | 5           | 45  | Male   | Blouse         | Clothing | 49                    | Oregon        | M    | Turquoise | Spring | 2.7           | Yes                 | Free Shipping | Yes              | Yes             | 31                 | PayPal         |

|   |                                      |                                 |
|---|--------------------------------------|---------------------------------|
| 01  | 02                                   | 03                              |
| <b>Data Import</b>                                | <b>Missing Values</b>                | <b>Standardization</b>          |
| Loaded CSV and validated structure                | Filled ratings with category medians | Renamed columns for consistency |
| 04  | 05                                   |                                 |
| <b>Feature Engineering</b>                        | <b>Database Upload</b>               |                                 |
| Created age groups and purchase frequency metrics | Connected to MySQL for SQL analysis  |                                 |

# Revenue by Gender



## Male Customers Drive Revenue

Male customers contribute 68% of total revenue, more than double female customer spending.

Key opportunity to increase female customer engagement and spending.

# Top-Rated Products



**Gloves**

Highest average rating: 3.86



**Sandals**

Strong performance: 3.84



**Boots**

Customer favorite: 3.82



**Hat**

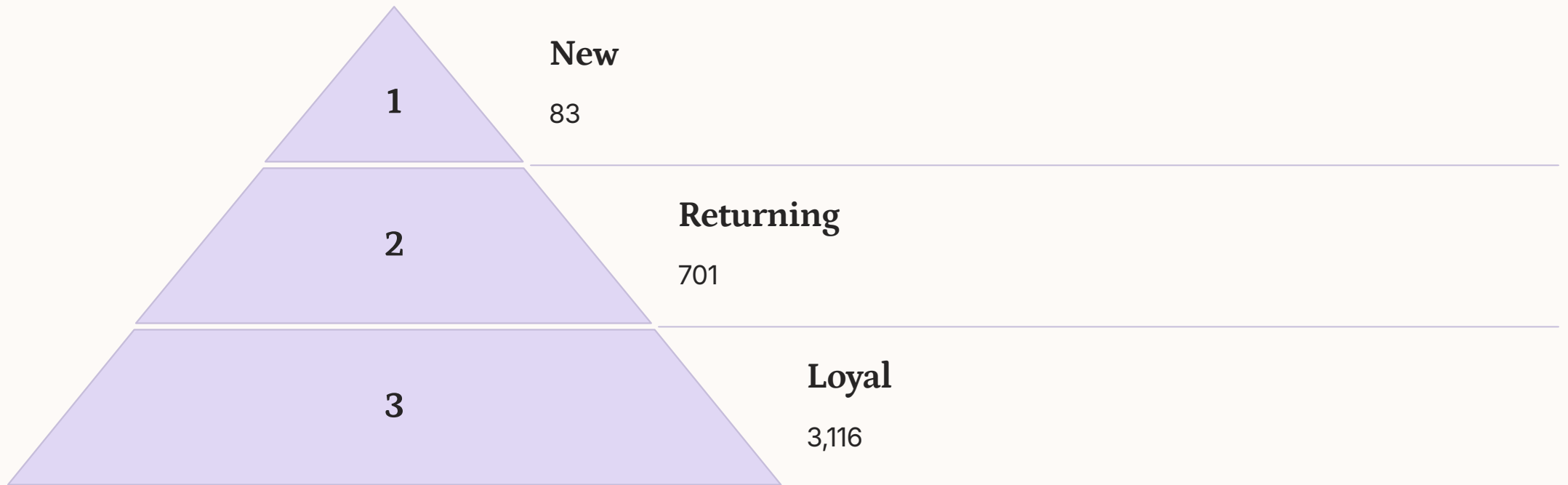
Consistent quality: 3.80



**Skirt**

Top clothing item: 3.78

# Customer Segmentation Insights



80% of customers are loyal repeat buyers, demonstrating strong retention. Focus on converting new customers to returning status.

# Subscription Analysis

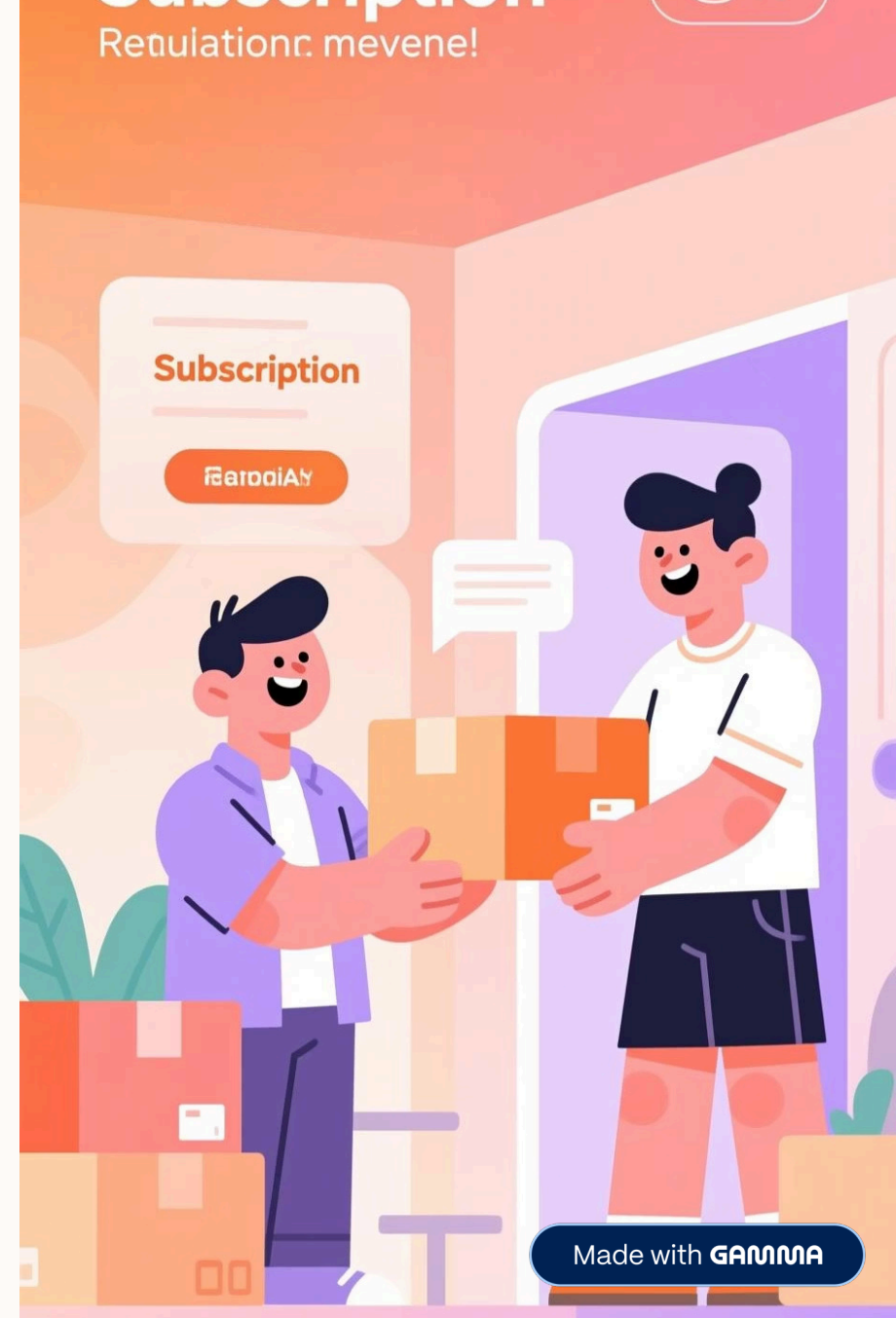
## Subscribers vs Non-Subscribers

- Subscribers: 1,053 customers
- Average spend: \$59.49
- Total revenue: \$62,645
- Non-subscribers: 2,847 customers
- Average spend: \$59.87
- Total revenue: \$170,436

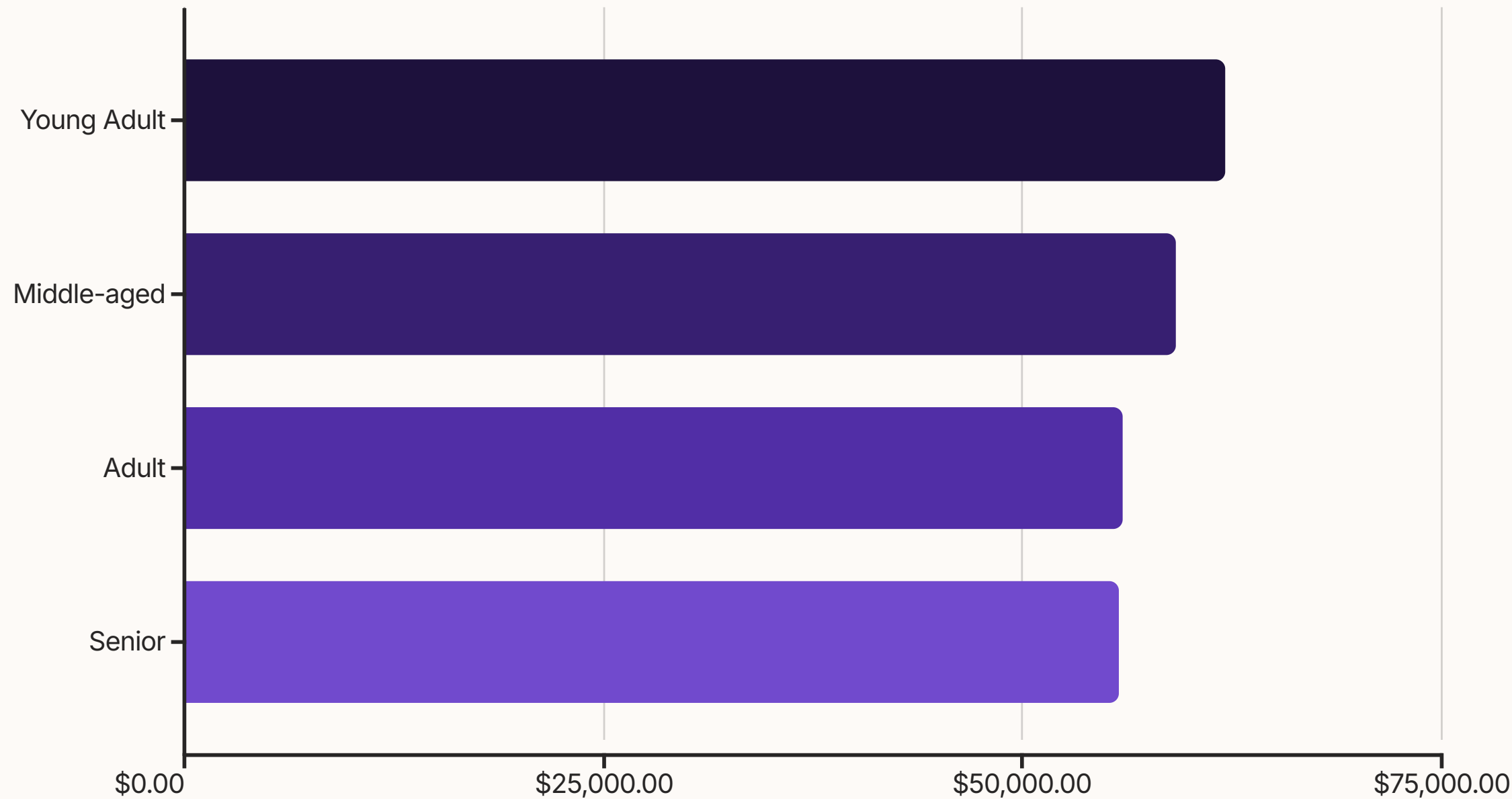
## Key Finding

Similar spending patterns between groups, but only 27% are subscribers.

**Opportunity:** Convert more customers to subscription model for predictable revenue.



# Revenue by Age Group



Young adults lead revenue generation, but all age groups contribute relatively evenly. Balanced demographic appeal across customer base.



# Power BI Dashboard



## Total Sales & Revenue

Real-time performance tracking



## Customer Insights

Top customers and segments



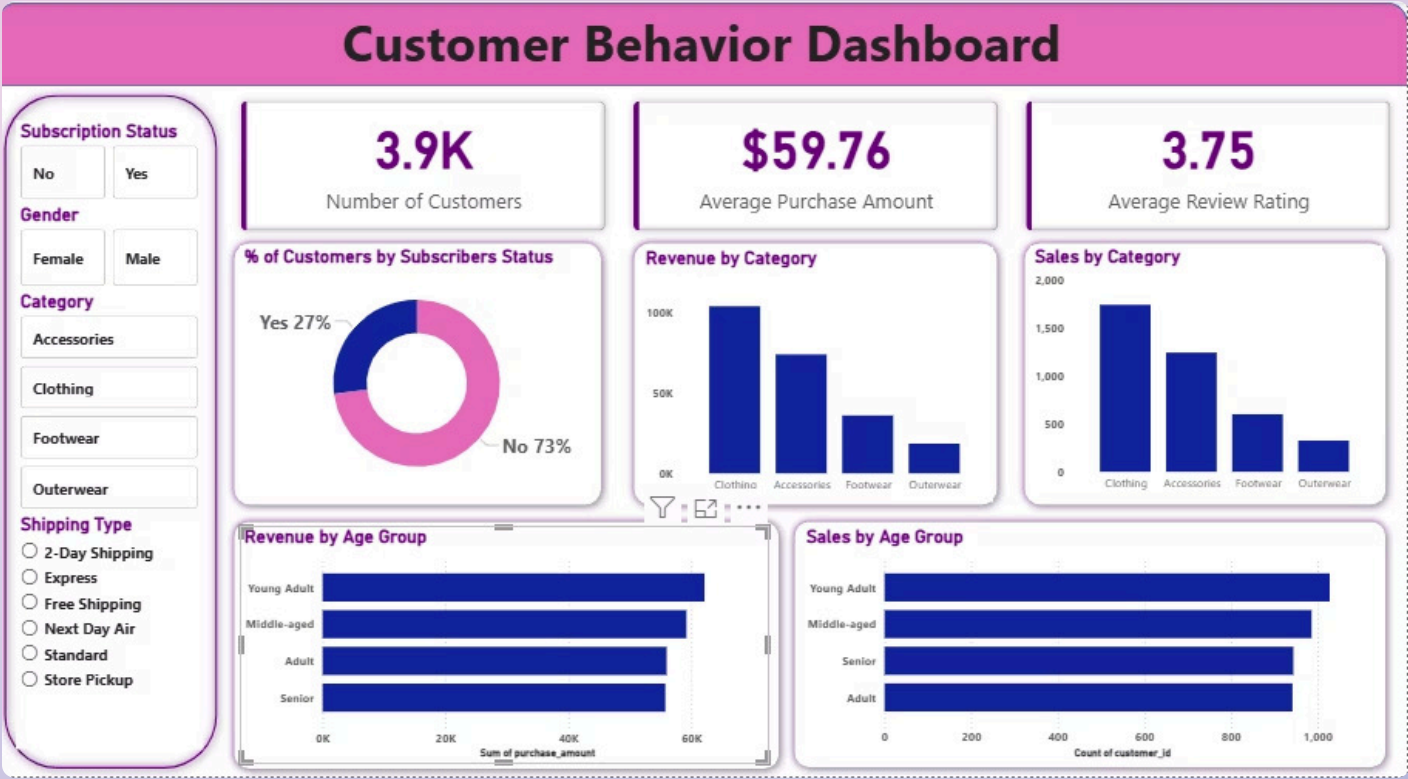
## Geographic Analysis

Revenue by age, gender, region



## Discount Impact

Promotion effectiveness tracking



# Business Recommendations



## Boost Subscriptions

Offer exclusive discounts and early access to convert non-subscribers



## Loyalty Program

Reward repeat buyers to strengthen retention and increase lifetime value



## Smart Discounting

Balance promotions to drive sales without eroding profit margins



## Promote Top Products

Highlight high-rated items like gloves and sandals in marketing



## Targeted Marketing

Focus campaigns on young adults and express shipping users

---

**Tools Used:** Python (Pandas, NumPy), MySQL, Power BI, VS Code