

# Customer Shopping Behavior Analysis

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

# Dataset Overview

3,900

Total Purchases

Transactions analyzed

18

Data Points

Features per transaction

50

Locations

Geographic coverage

25

Products

Unique items tracked

## Customer Data

- Demographics & location
- Subscription status
- Purchase history

## Transaction Data

- Product details & pricing
- Discounts & promotions
- Review ratings

# Data Preparation Process

01

## Data Loading

Imported dataset using pandas, explored structure with df.info() and summary statistics

02

## Missing Data Handling

Imputed 37 missing Review Rating values using median rating per product category

03

## Feature Engineering

Created age\_group bins and purchase\_frequency\_days columns for deeper analysis

04

## Data Consistency

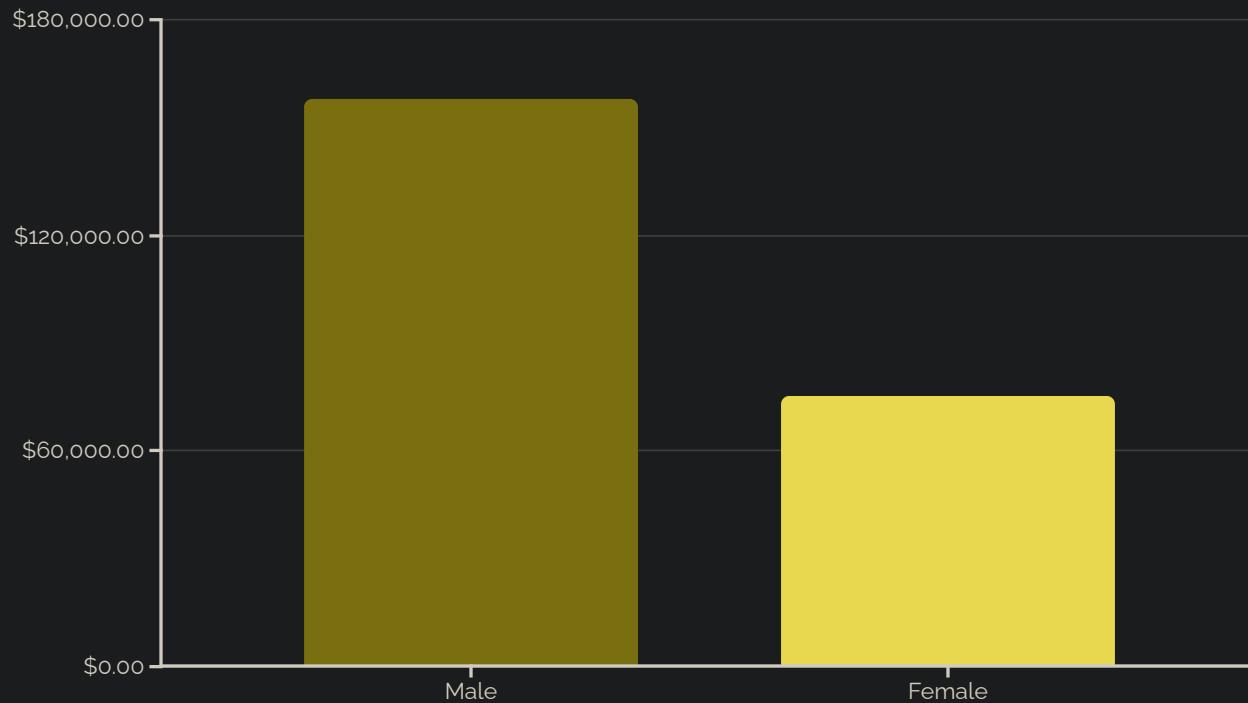
Standardized columns to snake\_case, removed redundant promo\_code\_used field

05

## Database Integration

Connected to PostgreSQL and loaded cleaned data for SQL analysis

# Revenue Insights



## Gender Revenue Gap

Male customers generate **2.1x more revenue** than female customers

Total revenue: \$233,081

# Customer Segmentation

## Loyal Customers

**3,116 customers**

Highest purchase frequency and lifetime value

## Returning Customers

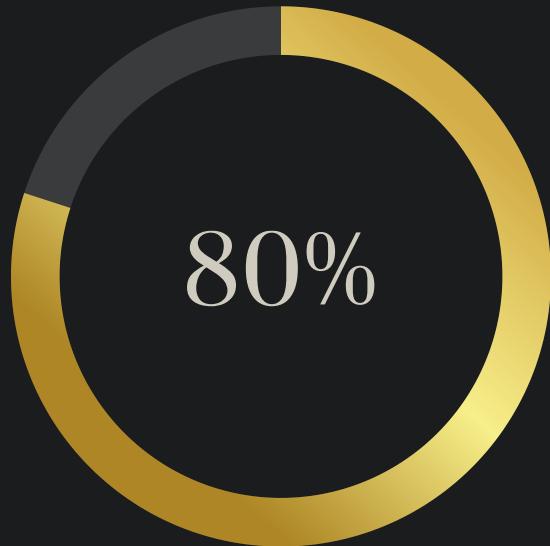
**701 customers**

Moderate engagement, growth potential

## New Customers

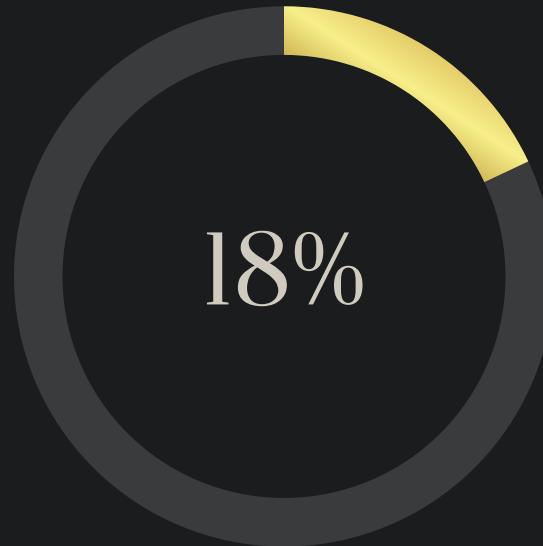
**83 customers**

First-time buyers, conversion opportunity



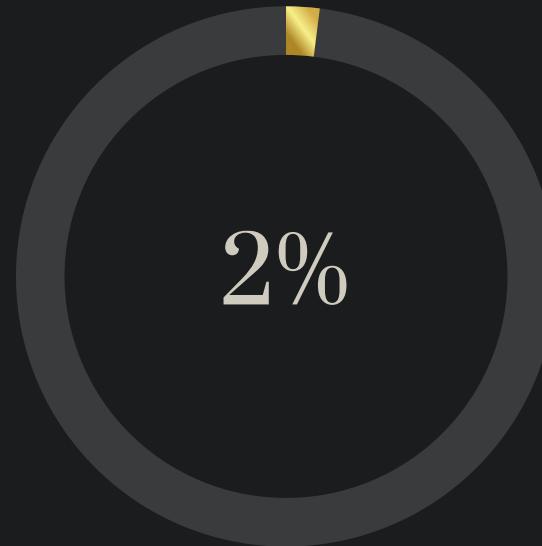
### Loyal Base

Strong customer retention



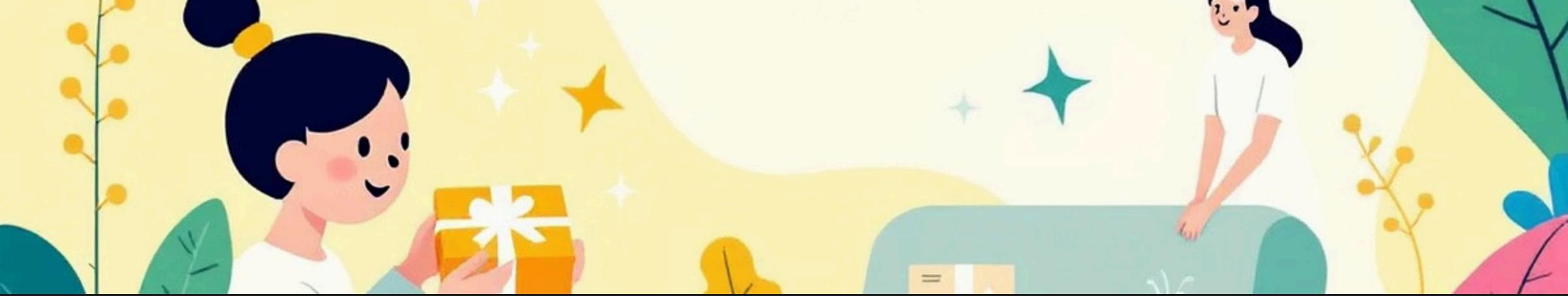
### Returning

Mid-tier engagement



### New

Acquisition focus needed



# Subscription Analysis

## Subscription Status

27%

Subscribers

1,053 customers

73%

Non-Subscribers

2,847 customers

## Spending Patterns

- Subscribers: **\$59.49** avg spend
- Non-subscribers: **\$59.87** avg spend
- Minimal spending difference

- ❑ Key Finding: Repeat buyers (>5 purchases) show **958 subscribers** vs **2,518 non-subscribers** - significant conversion opportunity exists

# Product Performance

1

## Top-Rated Products

Gloves (3.86), Sandals (3.84), Boots (3.82) lead customer satisfaction

2

## Best Sellers by Category

Clothing: Blouse (171), Accessories: Jewelry (171), Footwear: Sandals (160)

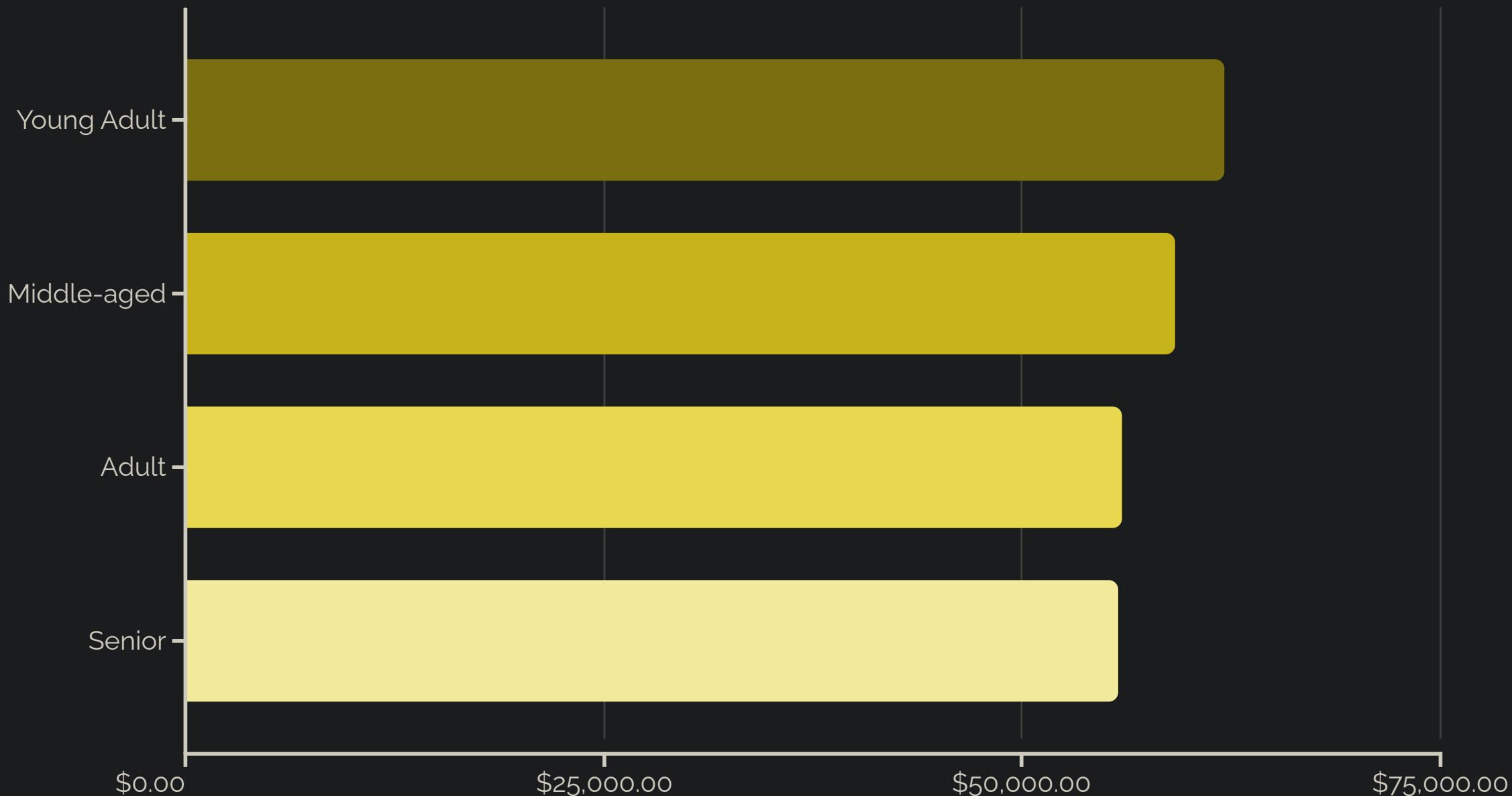
3

## Discount-Dependent Items

Hat (50%), Sneakers (49.66%), Coat (49.07%) rely heavily on discounts

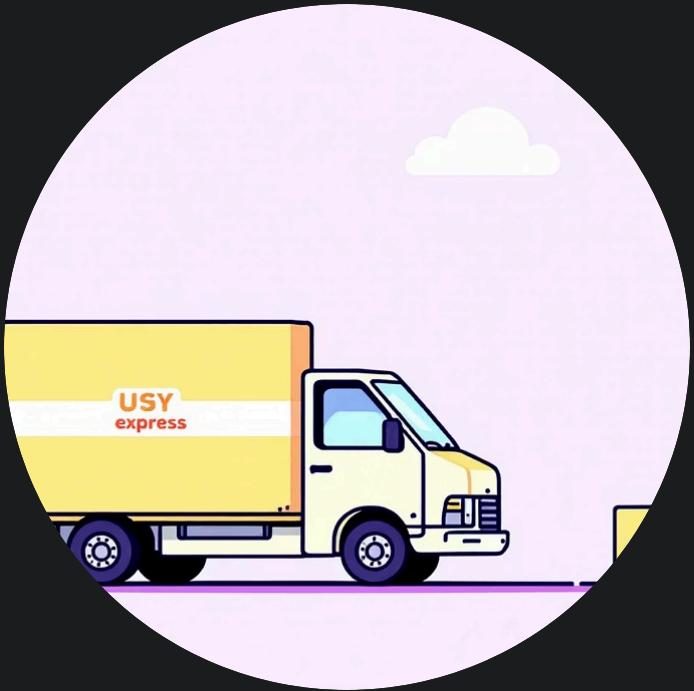


# Revenue by Age Group



Young adults drive highest revenue at **\$62,143**, followed closely by middle-aged customers. Revenue distribution relatively balanced across age groups.

# Shipping & Discount Insights



## Express Shipping

Average purchase: **\$60.48**

Premium service attracts higher  
spenders



## Standard Shipping

Average purchase: **\$58.46**

Most common shipping choice



## Smart Discounters

**839 customers** used discounts but spent  
above average

Discount-savvy high-value segment

# Strategic Recommendations



## Boost Subscriptions

Promote exclusive benefits to convert 73% non-subscribers, especially repeat buyers



## Loyalty Programs

Reward repeat buyers to strengthen 80% loyal customer base



## Review Discount Policy

Balance sales boosts with margin control for discount-dependent products



## Product Positioning

Highlight top-rated items (Gloves, Sandals, Boots) in marketing campaigns



## Targeted Marketing

Focus on young adults and express-shipping users for maximum ROI