



# 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

Comprehensive customer  
and purchase attributes

**50**

## Locations

Geographic diversity in  
customer base

**4**

## Categories

Clothing, Footwear,  
Accessories, Outerwear

## Customer Data

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

## Transaction Details

- Product, Category, Amount
- Season, Size, Color
- Discounts & Ratings

# Data Preparation Journey

01

## Data Loading & Exploration

Imported dataset using pandas, analyzed 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` for deeper analysis

04

## Data Standardization

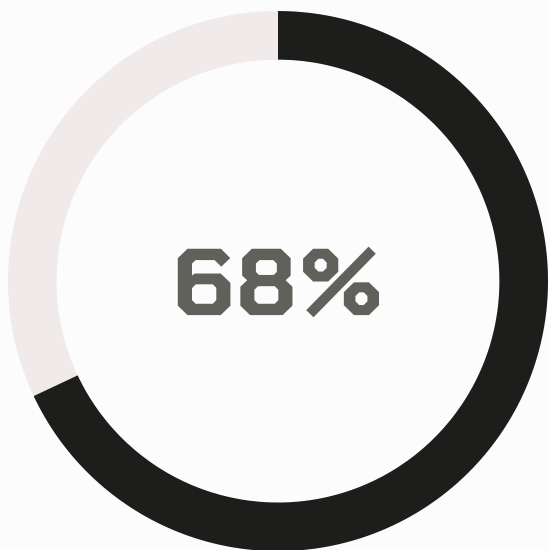
Renamed columns to snake\_case, removed redundant `promo_code_used` field

05

## Database Integration

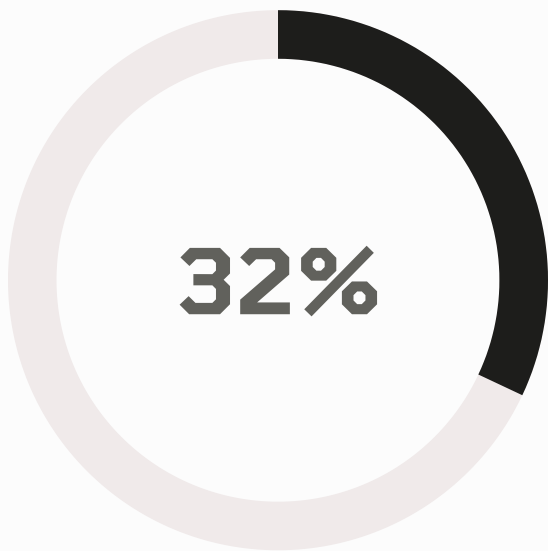
Connected to MySQL and loaded cleaned data for SQL analysis

# Customer Demographics



**Male Customers**

2,652 of total customer base



**Female Customers**

1,248 of total customer base



## Age Range

18-70 years

Average: 44 years

## Average Spend

\$59.76 per purchase

Range: \$20-\$100

	gender	Revenue
▶	Male	157890

## Revenue Analysis by Gender

### Male Customers

Higher total revenue due to larger customer base

### Female Customers

Comparable average spend per transaction

# Product Performance Insights

## Top Rated: Gloves

3.86 average rating

## Sandals

3.84 average rating

## Boots

3.82 average rating

## Most Purchased by Category

### Clothing

Blouse (171 orders)

Pants (171 orders)

### Accessories

Jewelry (171 orders)

Sunglasses (161 orders)

### Footwear

Sandals (160 orders)

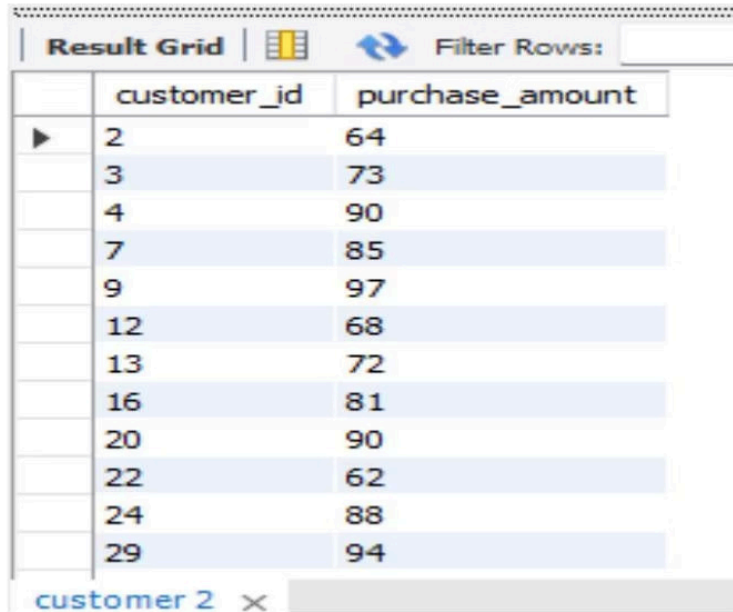
Shoes (150 orders)

### Outerwear

Jacket (163 orders)

Coat (161 orders)

# Discount & Promotion Impact



	customer_id	purchase_amount
▶	2	64
	3	73
	4	90
	7	85
	9	97
	12	68
	13	72
	16	81
	20	90
	22	62
	24	88
	29	94

customer 2 ×

## High-Value Discount Users

Customers using discounts while spending above average

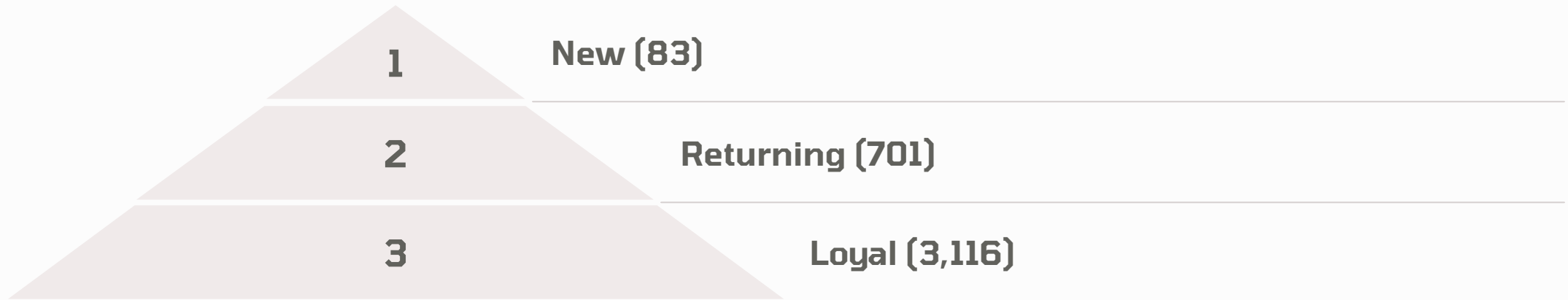
●●●●● 43%

## Discount Usage Rate

## Top Discounted Products

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

# Customer Segmentation



Customer classification based on purchase history reveals strong loyalty base

## Subscription Status

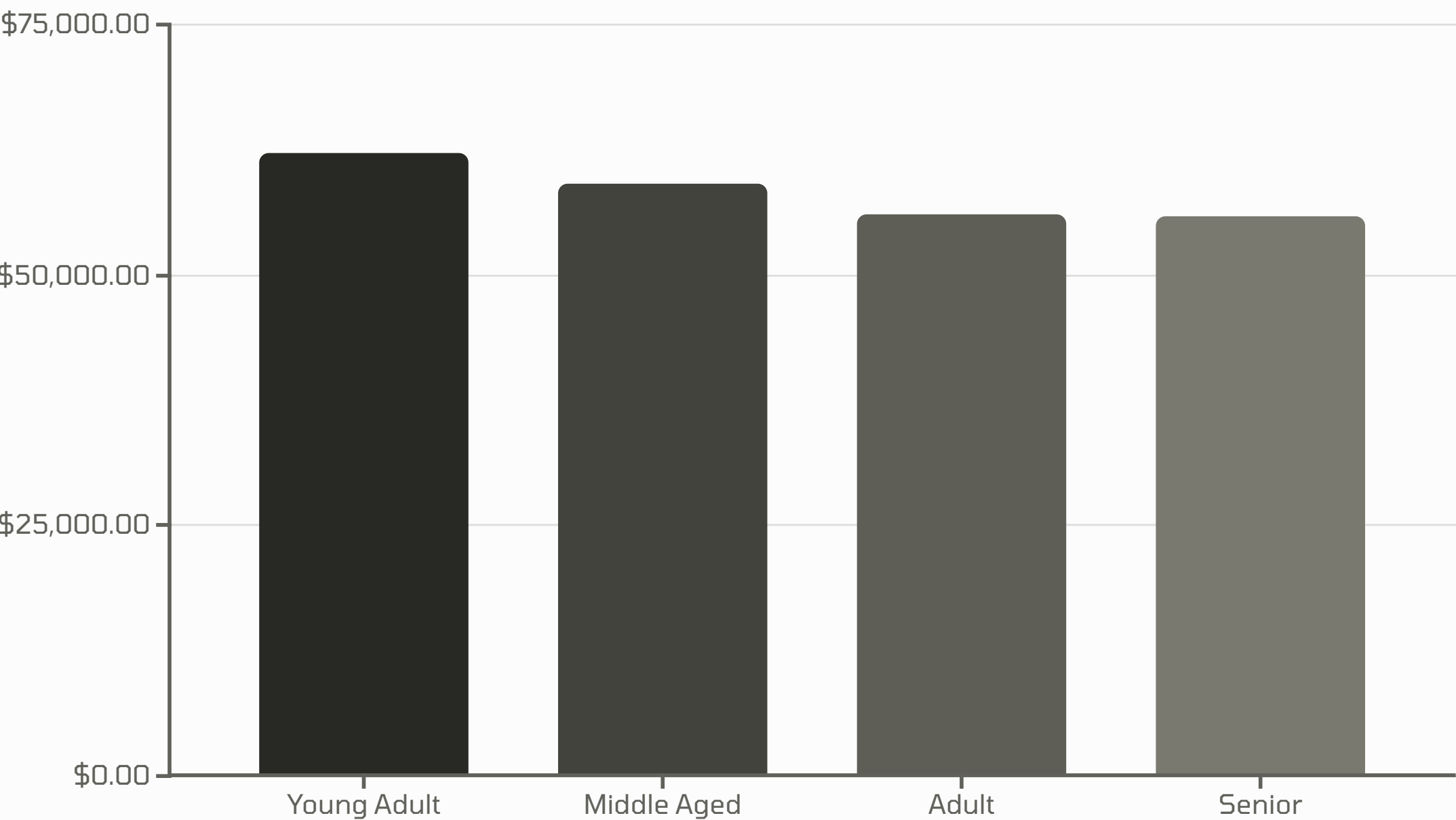
- Subscribers: 1,053 customers
- Non-subscribers: 2,847 customers
- Average spend similar across both groups

## Repeat Buyers

- 958 subscribers with 5+ purchases
- 2,518 non-subscribers with 5+ purchases
- Opportunity to convert repeat buyers



# Revenue by Age Group



## Shipping Preferences

Express: \$60.48 avg spend  
Standard: \$58.46 avg spend

## Key Insight

Young adults drive highest revenue; express shipping correlates with higher spend

# Strategic Recommendations

