



# Customer Shopping Behavior Analysis

A comprehensive data-driven exploration of 3,900 transactions to uncover spending patterns, product preferences, and strategic growth opportunities.



# Project Overview & Dataset

Our analysis leverages a robust dataset of 3,900 purchases across 18 key features, providing a 360-degree view of the customer journey.



## Demographics

Age, Gender, Location, and Subscription Status.



## Purchase Details

Item, Category, Amount, Season, Size, and Color.



## Behavioral Data

Review Ratings, Shipping Type, and Purchase Frequency.

# Data Preparation in Python

Before analysis, we performed rigorous data cleaning and feature engineering to ensure high-quality insights.

01

## Imputation

Filled 37 missing Review Ratings using the median rating per category.

02

## Standardization

Renamed columns to snake\_case and dropped redundant promo code data.

03

## Engineering

Created age\_group bins and purchase\_frequency\_days for segmentation.

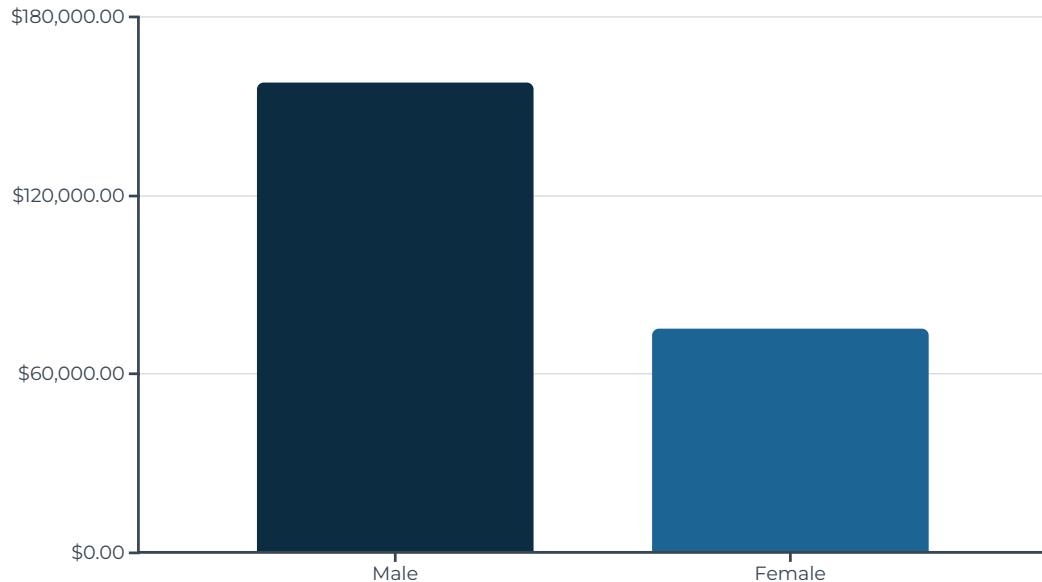
04

## Integration

Loaded the cleaned DataFrame into MySQL Server for structured SQL querying.

# Revenue & Spending Insights

SQL analysis revealed significant disparities in revenue generation across gender and shipping preferences.



## Key Findings

- **Gender Gap:** Male customers contribute over 2x the revenue of female customers.
- **Shipping Impact:** Express shipping users spend slightly more (\$60.48) than Standard users (\$58.46).
- **Discount Efficiency:** 839 high-spenders still exceed average purchase amounts despite using discounts.

# Top Performing Products

Quality and popularity metrics highlight our most successful inventory items.



## Gloves

Highest Average Rating: 3.86



## Sandals

Average Rating: 3.84



## Jewelry

Top Accessory: 171 Orders

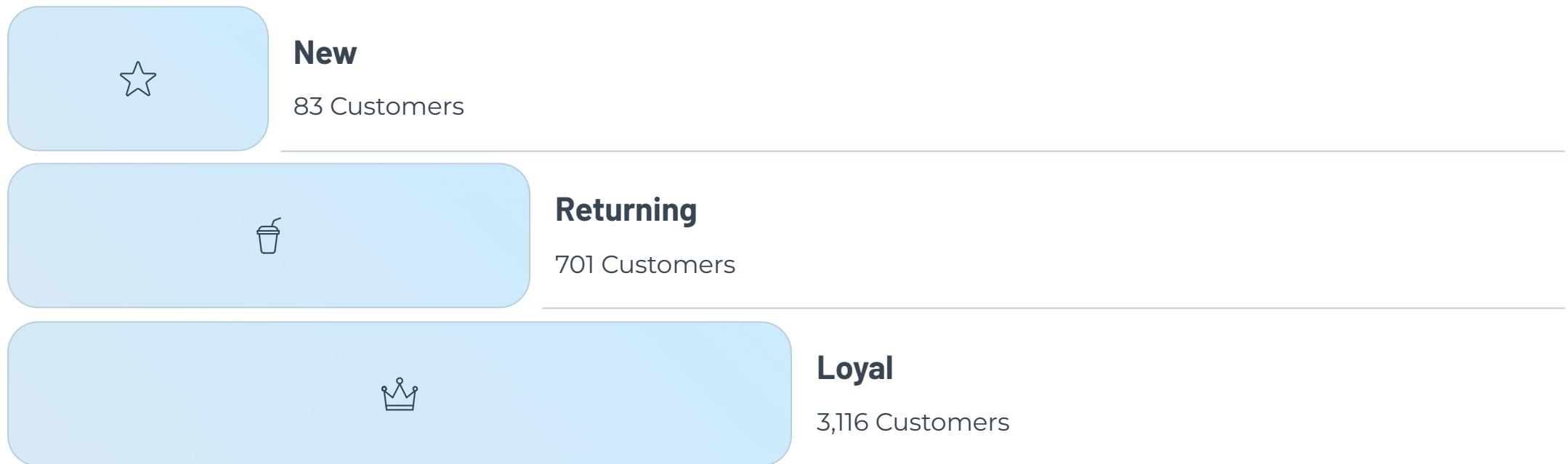


## Blouse

Top Clothing: 171 Orders

# Customer Segmentation

We classified our 3,900 customers into three distinct segments based on their purchase history to better target marketing efforts.



The vast majority of our base is "Loyal," indicating high retention but a potential need for more aggressive new customer acquisition.

# Subscription & Loyalty Analysis

Comparing the behavior of subscribers versus non-subscribers reveals opportunities for conversion.

## Subscribers

1,053 Customers

Total Revenue: \$62,645

## Non-Subscribers

2,847 Customers

Total Revenue: \$170,436



- Insight:** Average spend is nearly identical between groups (\$59.49 vs \$59.87), suggesting that subscription status currently doesn't drive higher individual transaction value.

# Demographic Revenue Breakdown

Revenue is remarkably well-distributed across age groups, with Young Adults leading the contribution.

**\$62K**

**Young Adult**

Highest revenue contribution group.

**\$59K**

**Middle-aged**

Strong secondary market segment.

**\$56K**

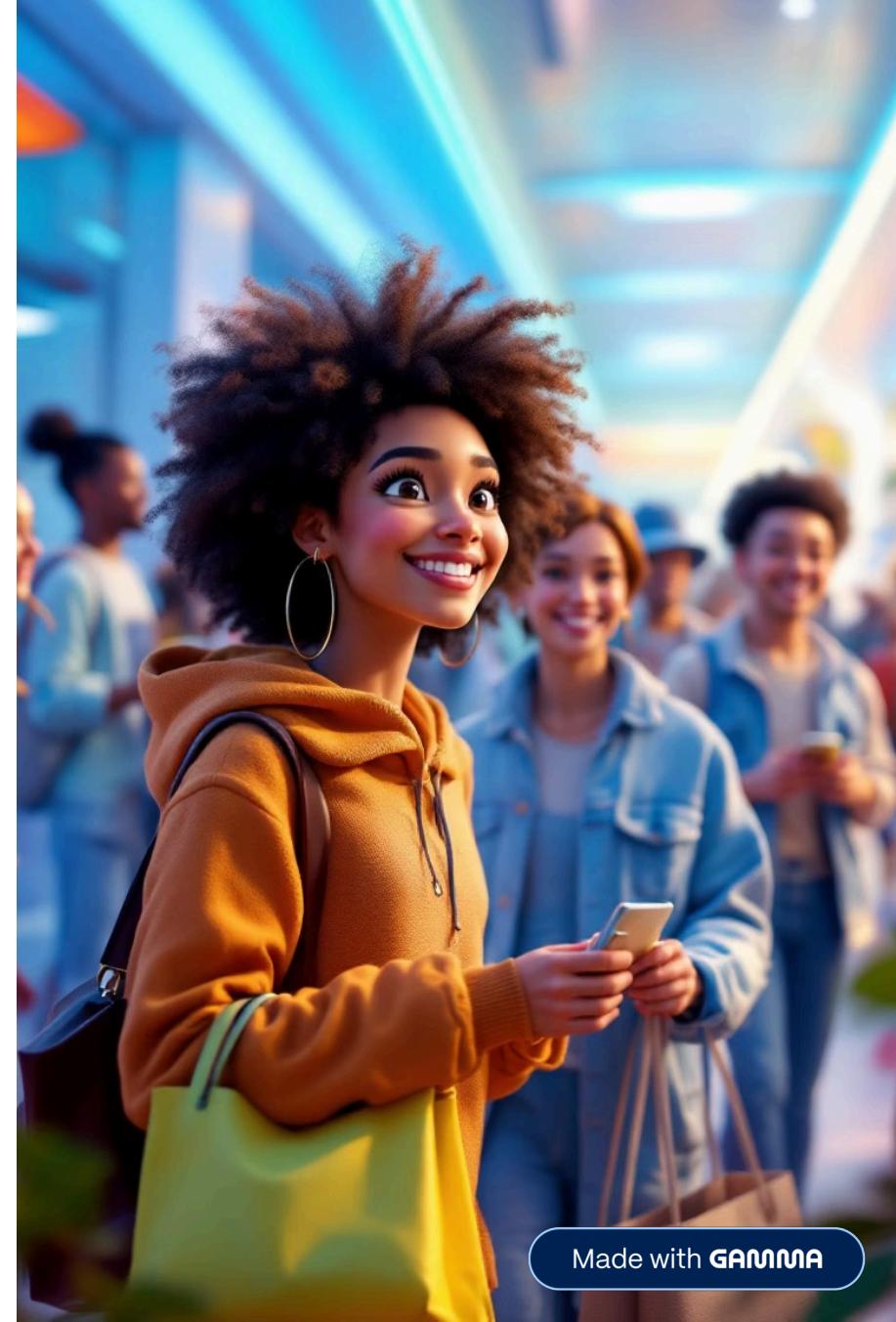
**Adult**

Consistent spending patterns.

**\$55K**

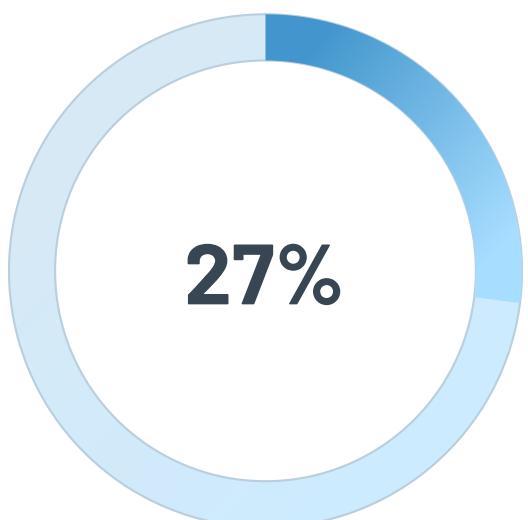
**Senior**

Stable revenue source.

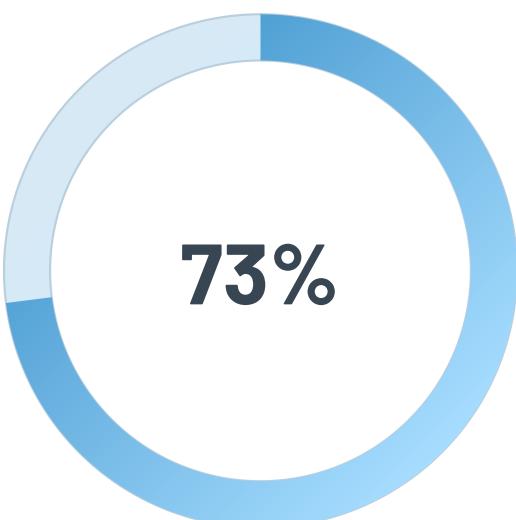


# Interactive Dashboard

Our Power BI dashboard provides real-time filtering by category, gender, and shipping type.



Subscribed Customers



Non-Subscribed Customers



Average Review Rating

# Strategic Recommendations

Based on the data, we recommend the following actions to optimize revenue and customer lifetime value.



## Boost Subscriptions

Promote exclusive benefits to convert the 73% non-subscriber base.



## Targeted Marketing

Focus on high-revenue age groups and express-shipping users.



## Product Positioning

Highlight top-rated items like Gloves and Sandals in campaigns.



## Discount Policy

Refine discount strategies for "Discount-Dependent" items like Hats.

