Customer Shopping Behavior Analysis

Analyzing 3,900 purchases to optimize business decisions.

Project Overview

Goal	Data	Impact
Uncover insights into spending patterns, segments, and preferences.	Transactional data from 3,900 purchases across categories.	Guide strategic business decisions for future growth.

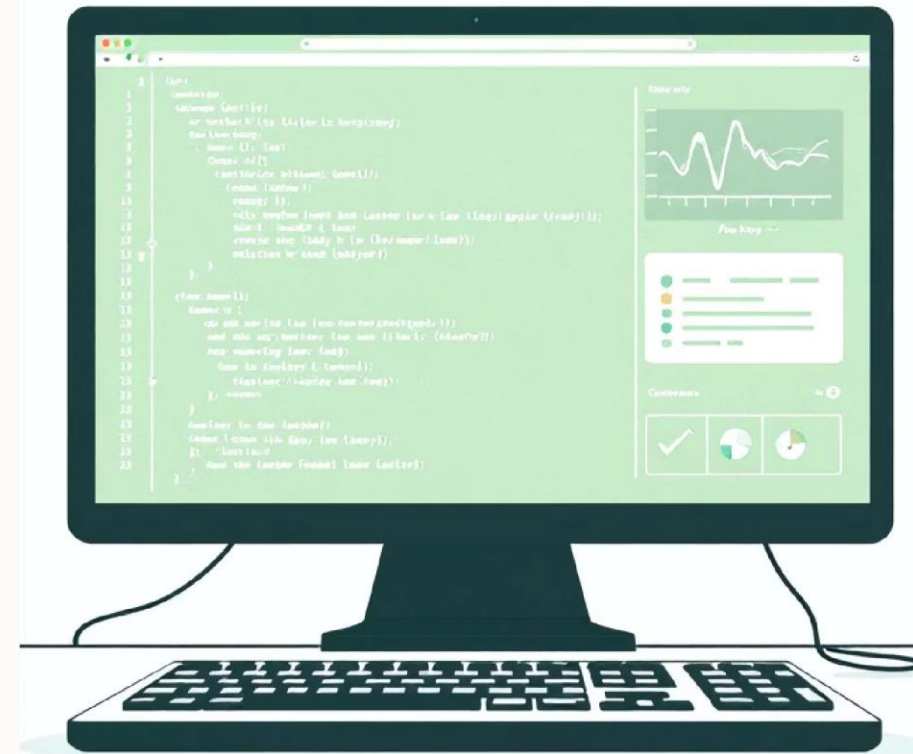# Dataset Summary

3,900 Rows

Comprehensive purchase records.

18 Columns

Key features: demographics, purchase details, behavior.

Missing Data

37 values in Review Rating column.

Exploratory Data Analysis (Python)

Data Preparation

- Loading with pandas

- Initial exploration (.info(), .describe())

- Missing data handling (median imputation)

- Column standardization (snake_case)

Feature Engineering

- Age groups created

- Purchase frequency days

- Dropped redundant promo_code_used

- Database integration (PostgreSQL)

Data Analysis (SQL)

1

Revenue by Gender

Male: $157,890, Female: $75,191

2

High-Spending Discount Users

839 customers identified.

3

Top 5 Products by Rating

Gloves (3.86), Sandals (3.84), Boots (3.82), Hat (3.80), Skirt (3.78).

4

Shipping Type Comparison

Standard: $58.46, Express: $60.48.
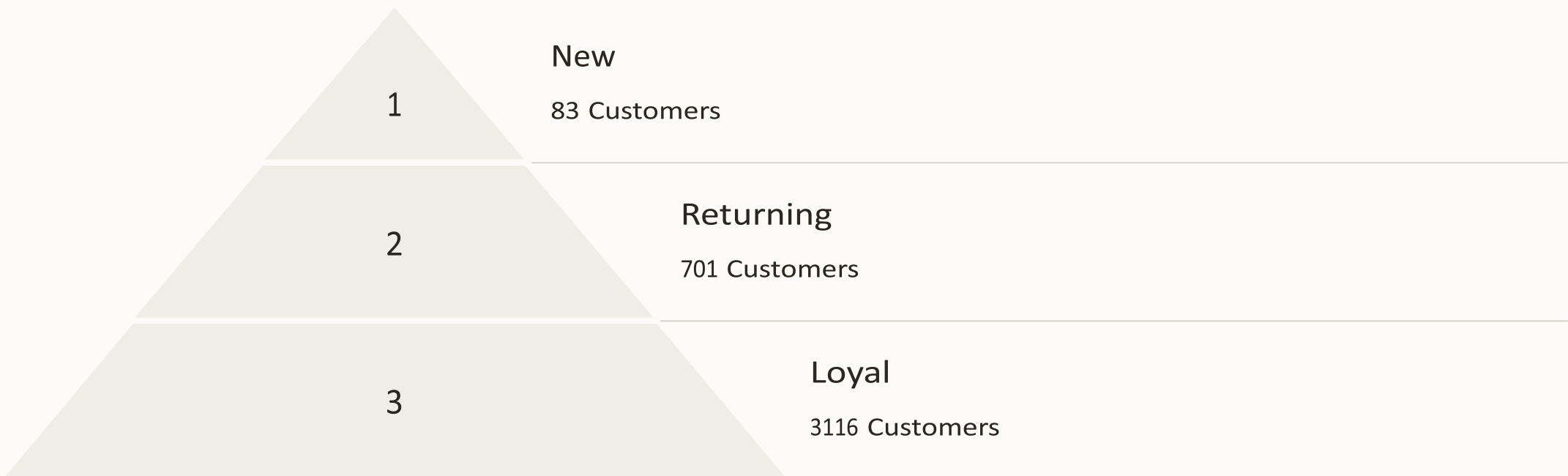
SQL Insights: Subscriptions & Discounts

Subscribers vs. Non-Subscribers

Yes	1053	$59.49	$62,645
No	2847	$59.87	$170,436

Discount-Dependent Products

- Hat: 50.00%
- Sneakers: 49.66%
- Coat: 49.07%
- Sweater: 48.17%
- Pants: 47.37%

Customer Segmentation

1 New

83 Customers

2 Returning

701 Customers

3 Loyal

3116 Customers

Top Products per Category

Accessories

Jewelry, Sunglasses, Belt

Clothing

Blouse, Pants, Shirt

Footwear

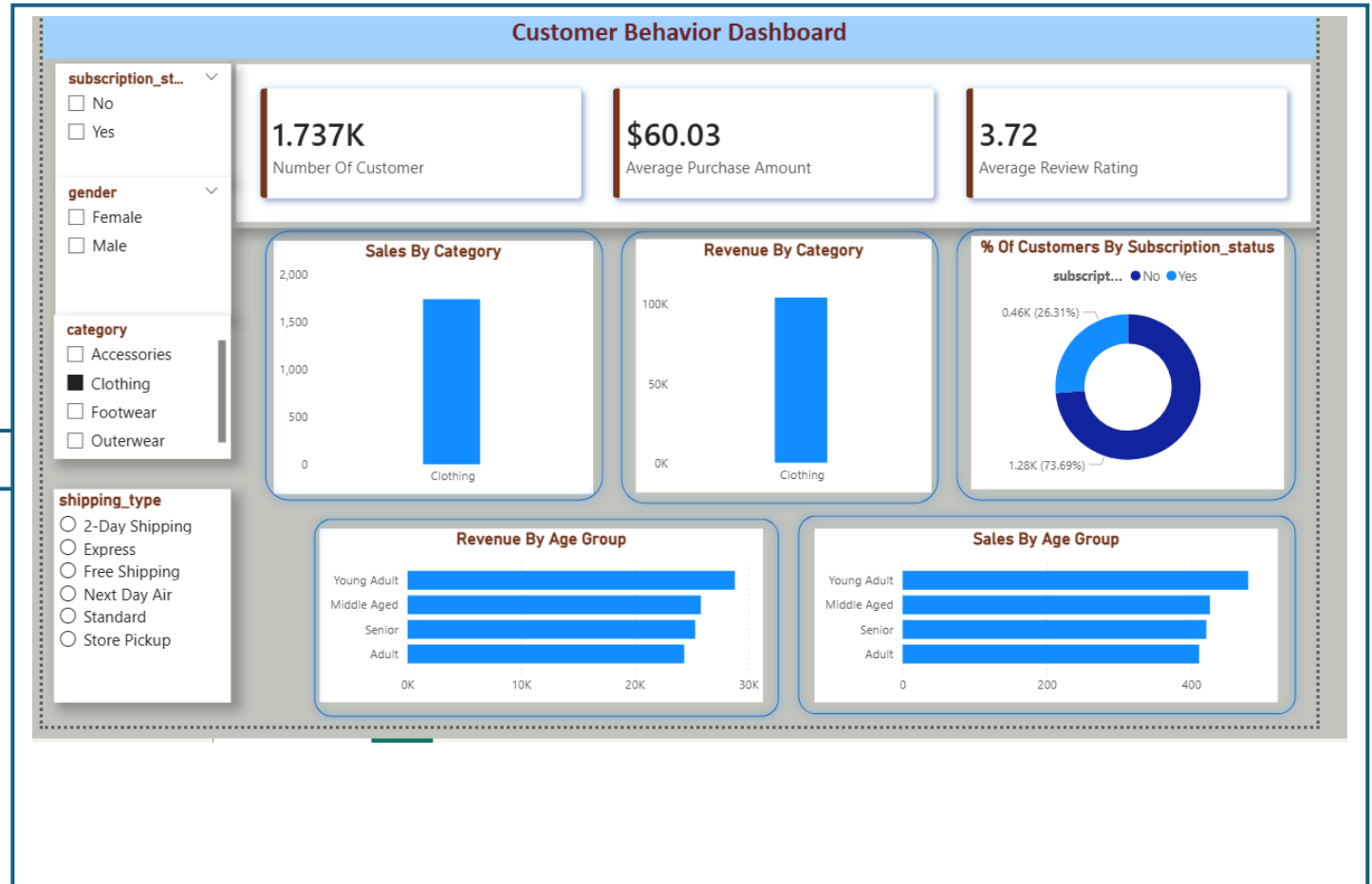Sandals, Shoes, Sneakers

Outerwear

Jacket, Coat

Power BI Dashboard

Interactive visualizations for deeper insights.

Business Recommendations

→ **Boost Subscriptions**

Promote exclusive benefits.

→ **Customer Loyalty Programs**

Reward repeat buyers.

→ **Review Discount Policy**

Balance sales with margin control.

→ **Product Positioning**

Highlight top-rated products.

→ **Targeted Marketing**

Focus on high-revenue groups.