

# Customer Shopping Behavior Analysis

Uncovering insights from transactional data to guide strategic business decisions.





# Project Overview

## Goal

Uncover insights into spending patterns, customer segments, product preferences, and subscription behavior.

## Data Source

Transactional data from 3,900 purchases across various product categories.

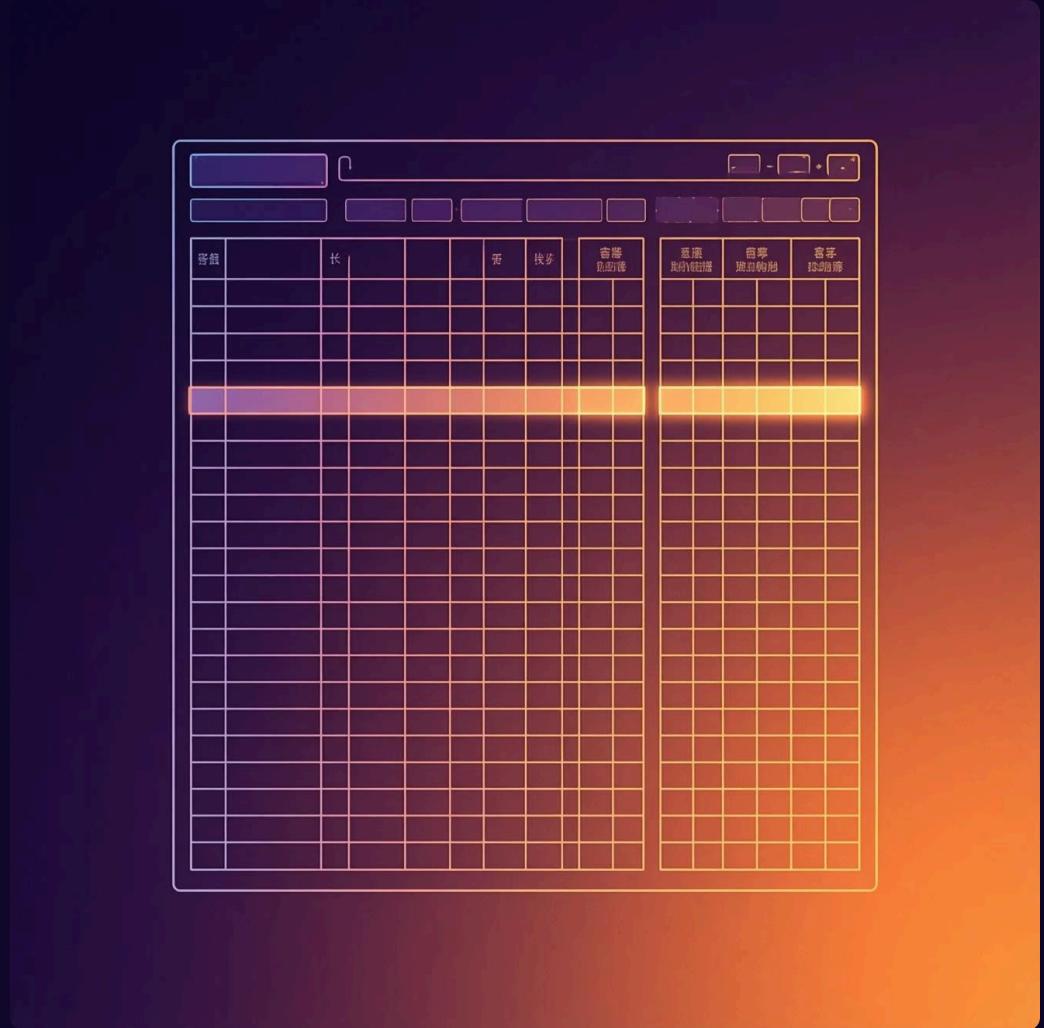
## Impact

Guide strategic business decisions for optimization and growth.

# Dataset Summary

## Key Features

- Customer demographics (Age, Gender, Location, Subscription Status)
- Purchase details (Item, Category, Amount, Season, Size, Color)
- Shopping behavior (Discount, Promo Code, Previous Purchases, Frequency, Review, Shipping)



A screenshot of a data visualization interface, likely Gamma, showing a heatmap of a dataset. The interface has a dark theme with orange highlights. At the top, there are several input fields and buttons. Below them is a grid of colored cells representing the data. The columns are labeled with Chinese characters: 长 (Chang), 短 (Duan), 喜 (Xi), 惠 (Wei), 营 (Ying), 蓝 (Lan), 绿 (Lue), and 白 (Bai). The rows are numbered from 1 to 3900. A prominent orange bar highlights the first two rows of the grid.

1	长	短	喜	惠	营	蓝	绿	白
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								
36								
37								
38								
39								
40								
41								
42								
43								
44								
45								
46								
47								
48								
49								
50								
51								
52								
53								
54								
55								
56								
57								
58								
59								
60								
61								
62								
63								
64								
65								
66								
67								
68								
69								
70								
71								
72								
73								
74								
75								
76								
77								
78								
79								
80								
81								
82								
83								
84								
85								
86								
87								
88								
89								
90								
91								
92								
93								
94								
95								
96								
97								
98								
99								
100								

## Dataset Snapshot

- Rows: 3,900
- Columns: 18
- Missing Data: 37 values in Review Rating

# Exploratory Data Analysis (Python)

01

## Data Loading & Exploration

Imported dataset with `pandas`, checked structure and summary statistics.

02

## Missing Data Handling

Imputed missing 'Review Rating' using median per product category.

03

## Column Standardization

Renamed columns to snake case for readability.

04

## Feature Engineering

Created 'age\_group' and 'purchase\_frequency\_days'.

05

## Database Integration

Loaded cleaned data into PostgreSQL for SQL analysis.



# Data Analysis using SQL: Key Findings

## Revenue by Gender

Male: \$157,890

Female: \$75,191

## Top 5 Products by Rating

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

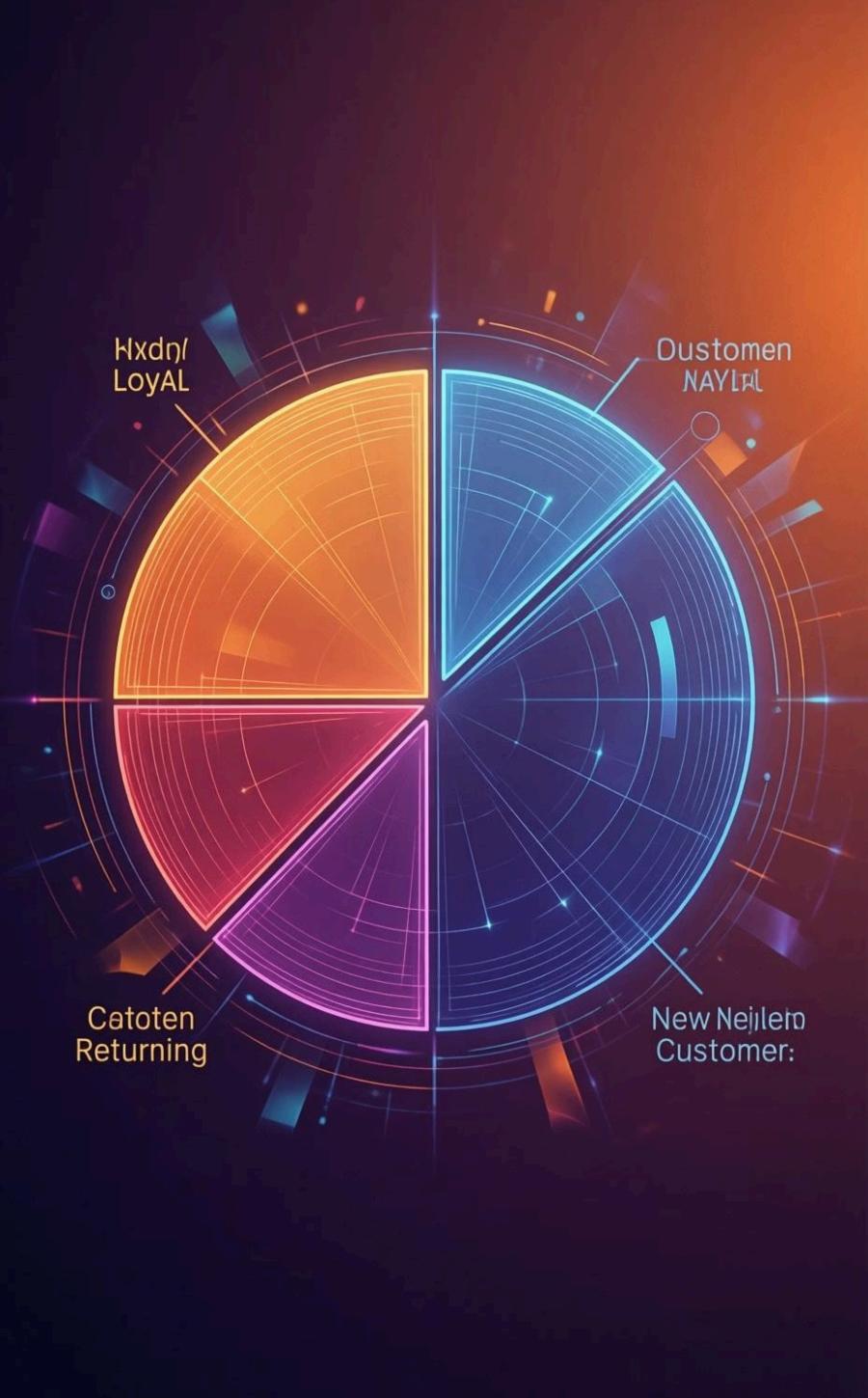
## Shipping Type Comparison

Standard: \$58.46 avg.

Express: \$60.48 avg.

## Discount-Dependent Products

Hat (50%), Sneakers (49.66%),  
Coat (49.07%), Sweater (48.17%),  
Pants (47.37%)



# SQL Insights: Customer Segmentation



Loyal

3,116 Customers



Returning

701 Customers



New

83 Customers

## Subscribers vs. Non-Subscribers

- Subscribers (1,053): \$59.49 avg. spend, \$62,645 total revenue
- Non-Subscribers (2,847): \$59.87 avg. spend, \$170,436 total revenue

## Repeat Buyers & Subscriptions

- 958 repeat buyers are subscribers.
- 2,518 repeat buyers are not subscribers.



# Top Products by Category & Age Group Revenue

## Top 3 Products per Category

- Accessories:** Jewelry, Sunglasses, Belt
- Clothing:** Blouse, Pants, Shirt
- Footwear:** Sandals, Shoes, Sneakers
- Outerwear:** Jacket, Coat

## Revenue by Age Group

- Young Adult:** \$62,143
- Middle-aged:** \$59,197
- Adult:** \$55,978
- Senior:** \$55,763

Gender

Female

Male

Category

Accessories

Clothing

Footwear

Outerwear

Number of Customers

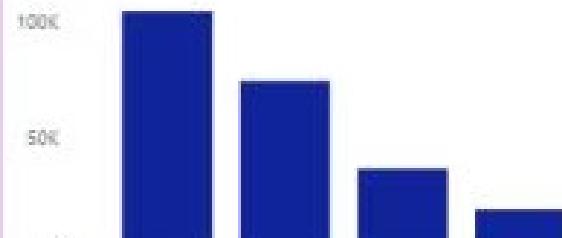
Average Purchase Amount

Average Review Rating

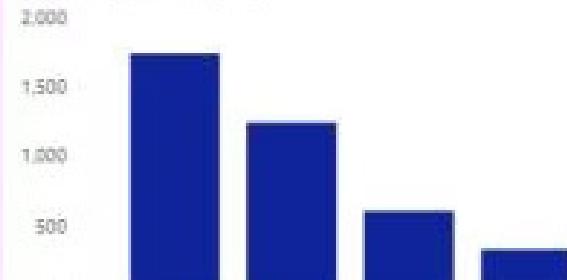
% of Customers by Subscription Status



Revenue by Category



Sales by Category



# Interactive Dashboard in Power BI

Visualizing insights for actionable business strategies.

3.9K

Customers

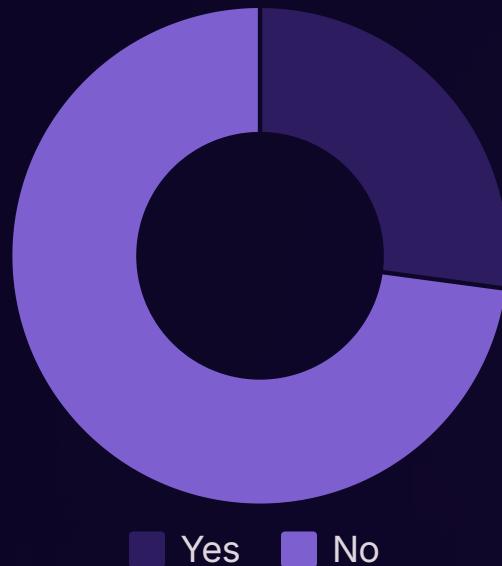
\$59.76

Avg. Purchase

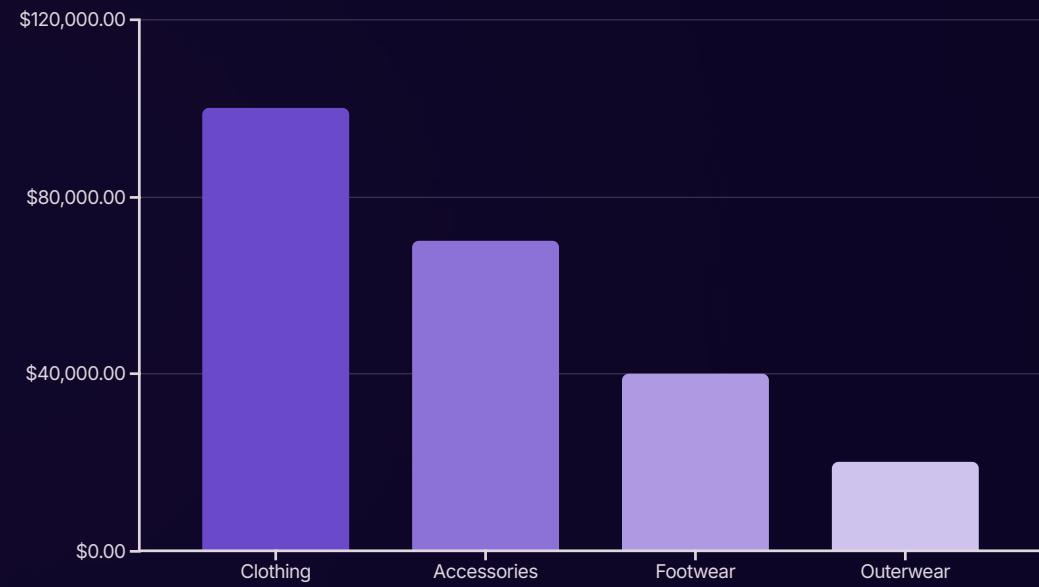
3.75

Avg. Review

# Dashboard Highlights



27% of customers are subscribers.



Clothing generates the highest revenue.

# Business Recommendations



## Boost Subscriptions

Promote exclusive benefits.



## Customer Loyalty

Reward repeat buyers.



## Review Discount Policy

Balance sales with margins.



## Product Positioning

Highlight top-rated items.



## Targeted Marketing

Focus on high-revenue groups.