

Customer Shopping Behavior Analysis

This project analyzes customer shopping behavior using transactional data from 3,900 purchases across various product categories. The goal is to uncover insights into spending patterns, customer segments, product preferences, and subscription behavior to guide strategic business decisions.



Dataset Summary

Key Features

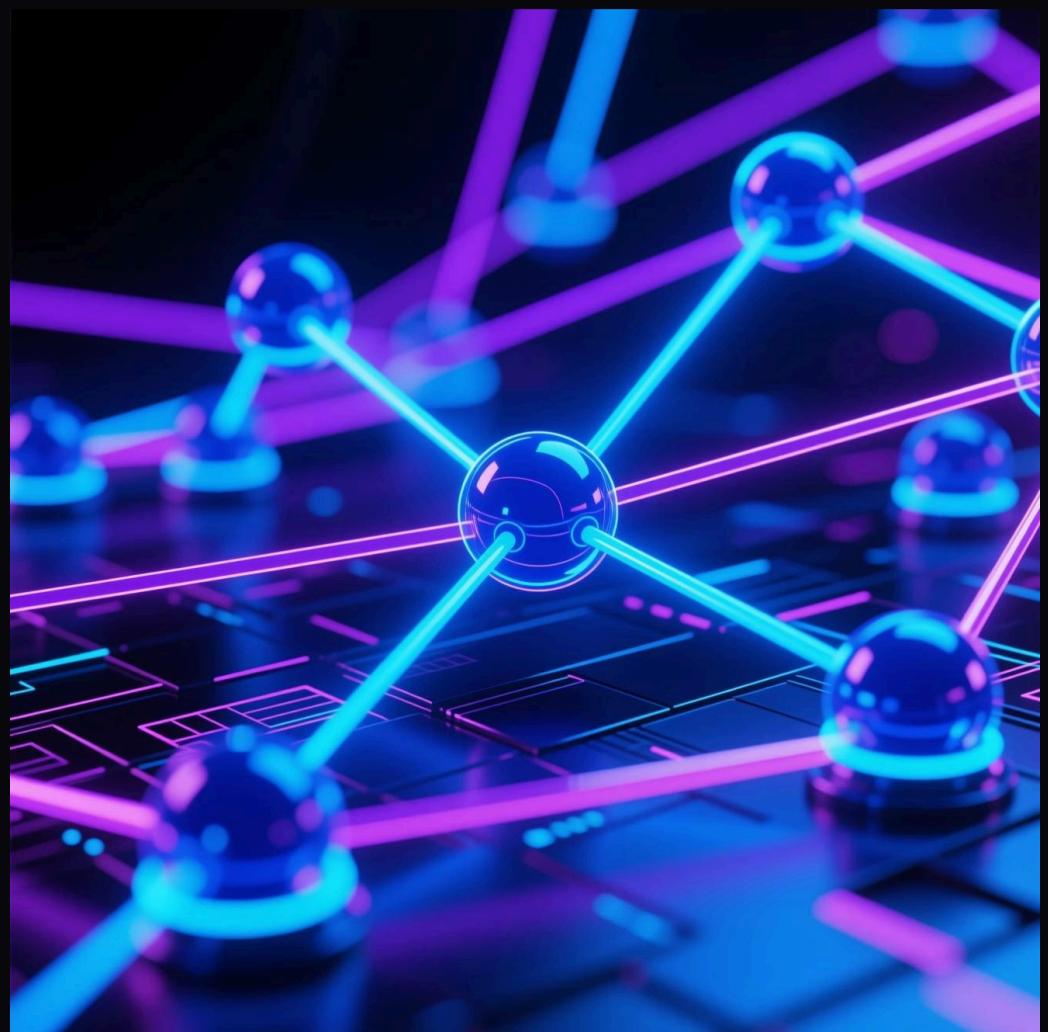
- Customer demographics (Age, Gender, Location, Subscription Status)
- Purchase details (Item, Category, Amount, Season, Size, Color)
- Shopping behavior (Discount, Promo Code, Previous Purchases, Frequency, Review Rating, Shipping Type)

Dataset Overview

Rows: 3,900

Columns: 18

Missing Data: 37 values in Review Rating



Exploratory Data Analysis with Python

Data Loading & Exploration

Imported dataset using pandas. Used `df.info()` and `.describe()` for initial checks.

Missing Data Handling

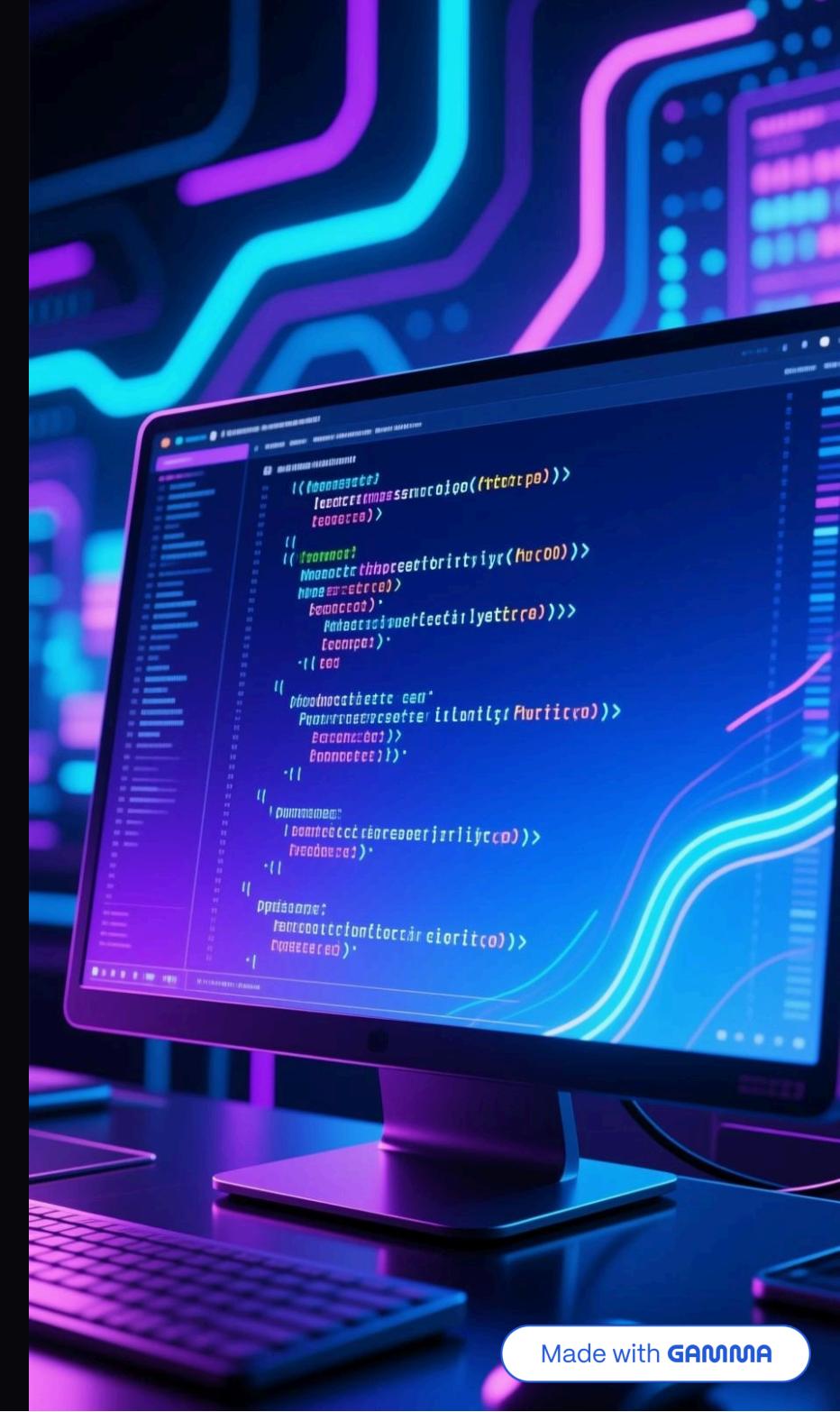
Imputed missing values in 'Review Rating' using the median rating per product category.

Column Standardization

Renamed columns to snake_case for readability.

Feature Engineering

Created 'age_group' and 'purchase_frequency_days' columns.





Data Analysis using SQL

Structured analysis in PostgreSQL answered key business questions.

1

Revenue by Gender

Compared total revenue from male vs. female customers.

2

High-Spending Discount Users

Identified customers using discounts but spending above average.

3

Top 5 Products by Rating

Found products with the highest average review ratings.

4

Shipping Type Comparison

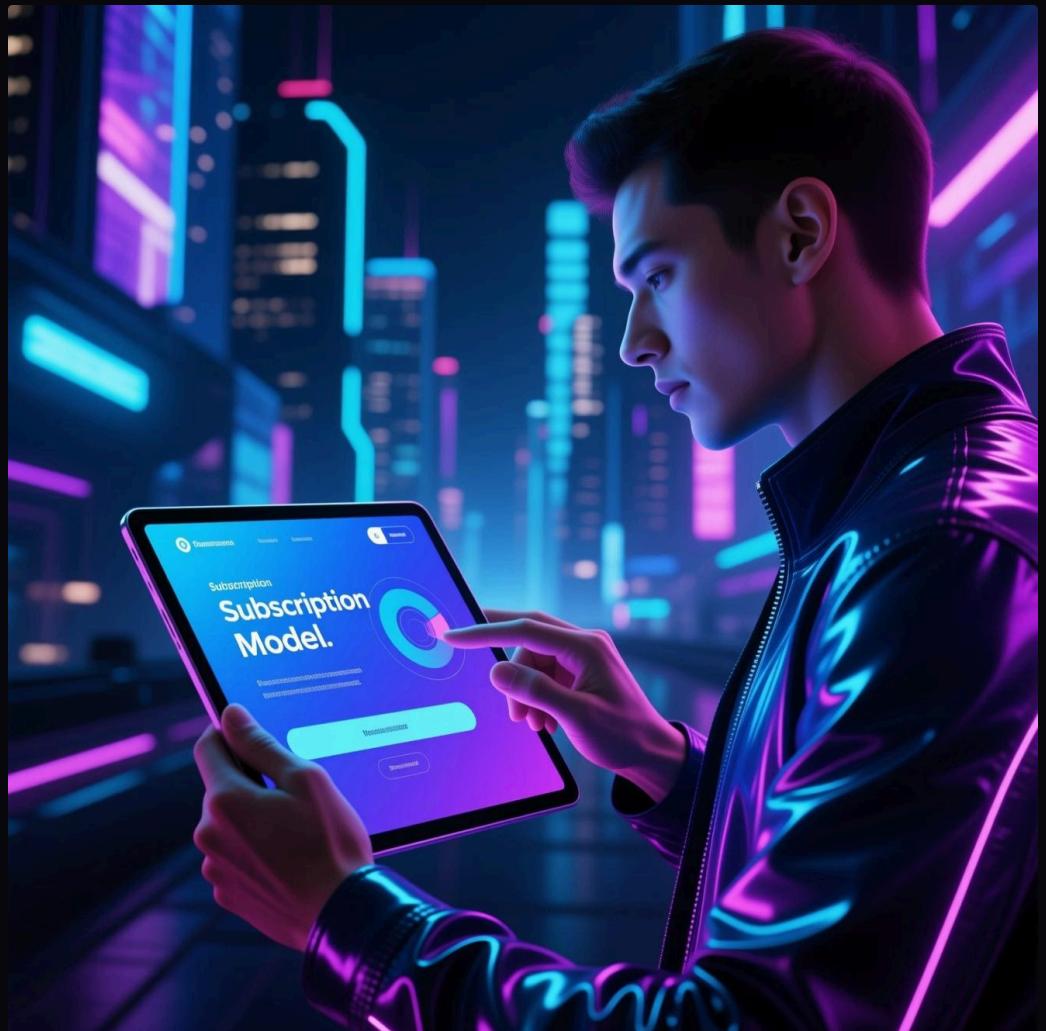
Compared average purchase amounts for Standard vs. Express shipping.

Subscription Insights

Subscribers vs. Non-Subscribers

Compared average spend and total revenue across subscription statuses.

Yes	1053	59.49	62645.00
No	2847	59.87	170436.00



Repeat Buyers & Subscriptions

Customers with >5 purchases are more likely to subscribe.

No	2518
Yes	958

Product & Discount Analysis



Discount-Dependent Products

Identified 5 products with the highest percentage of discounted purchases.



Top 3 Products per Category

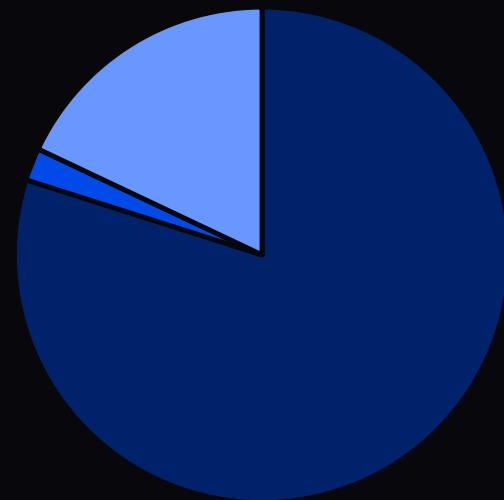
Listed the most purchased products within each category.

1	Accessories	Jewelry
1	Clothing	Blouse
1	Footwear	Sandals

Customer Segmentation & Revenue by Age

Customer Segmentation

Customers classified into New, Returning, and Loyal segments based on purchase history.



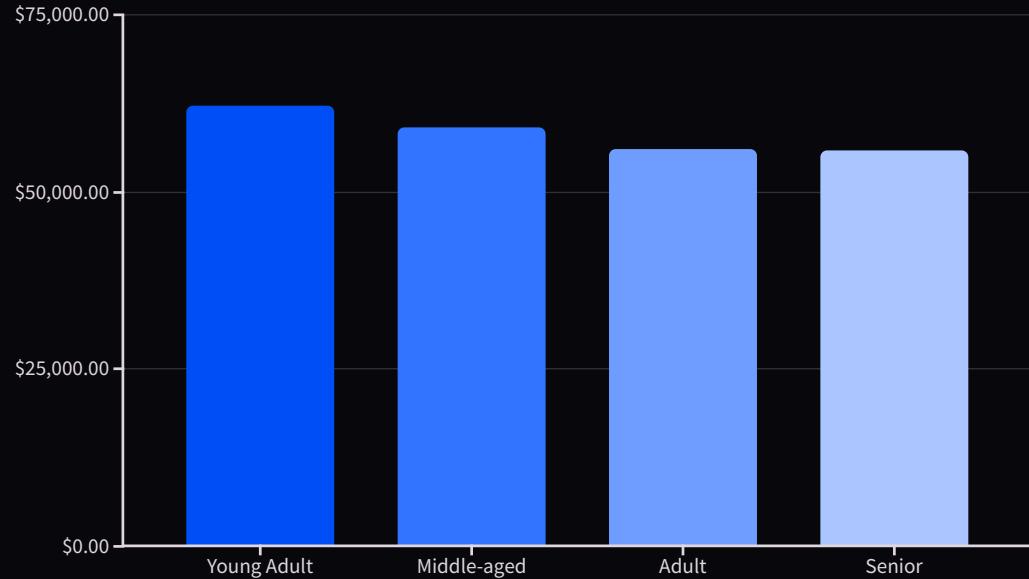
■ Loyal

■ New

■ Returning

Revenue by Age Group

Total revenue contribution of each age group.



Gender

Female

Male

Category

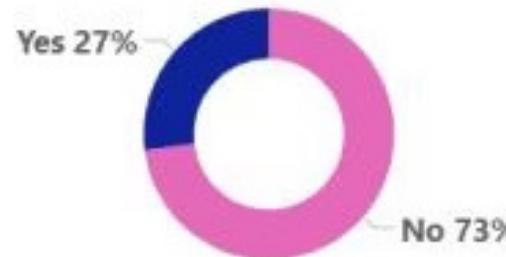
Accessories

Clothing

Footwear

Outerwear

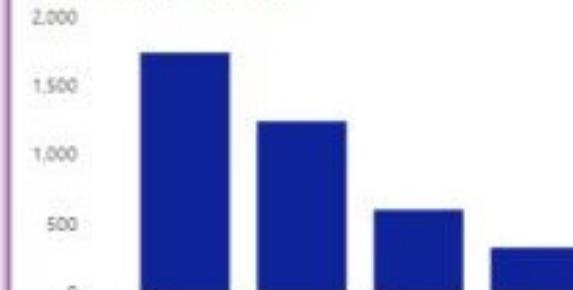
% of Customers by Subscription Status



Revenue by Category



Sales by Category



Power BI Dashboard: Key Metrics

An interactive dashboard in Power BI presents insights visually.

3.9K

Customers

\$59.76

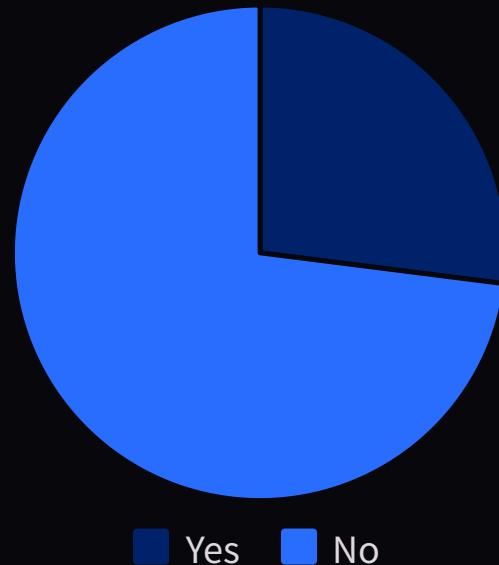
Avg. Purchase Amount

3.75

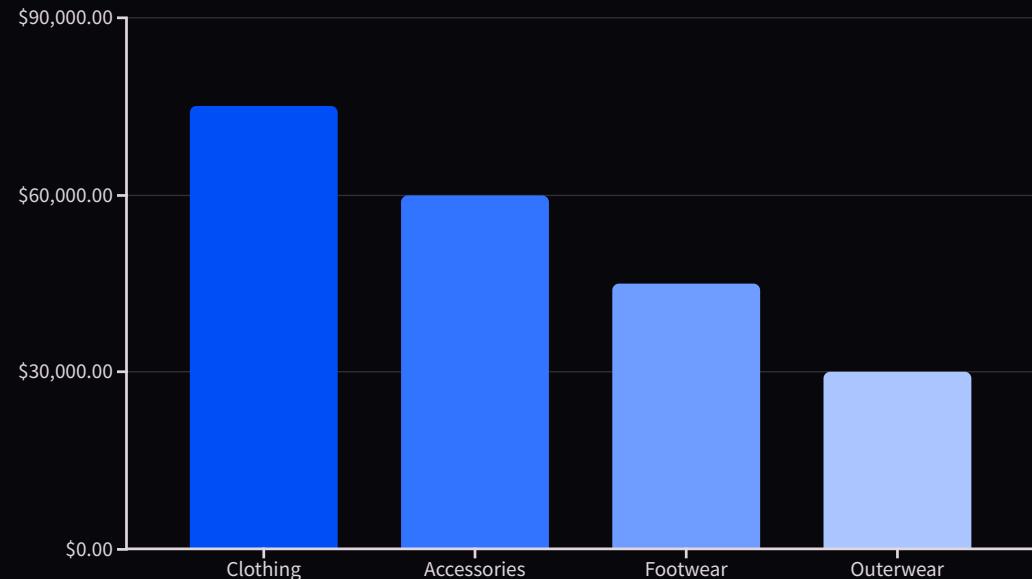
Avg. Review Rating

Dashboard Visualizations

% of Customers by Subscription Status



Revenue by Category



Business Recommendations

→ Boost Subscriptions

Promote exclusive benefits for subscribers to increase sign-ups.

→ Customer Loyalty Programs

Reward repeat buyers to foster loyalty and move them into the “Loyal” segment.

→ Review Discount Policy

Balance sales boosts with margin control to optimize profitability.

→ Product Positioning

Highlight top-rated and best-selling products in marketing campaigns.

→ Targeted Marketing

Focus efforts on high-revenue age groups and express-shipping users for maximum impact.