



Customer Shopping Behavior Analysis

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

Dataset Summary



Data Volume

3,900 rows and 18 columns of transactional data.



Customer Demographics

Includes Age, Gender, Location, and Subscription Status.



Purchase Details

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

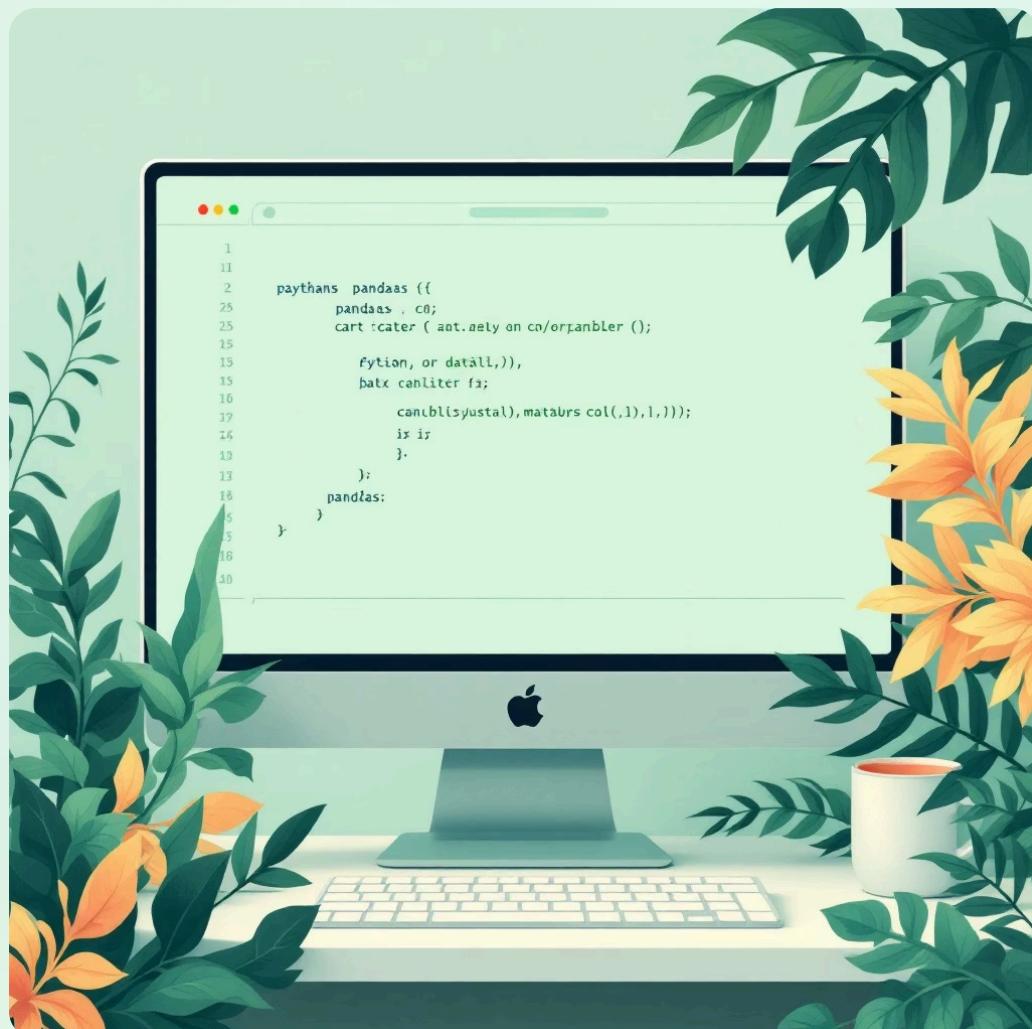


Shopping Behavior

Discount, Promo Code, Previous Purchases, Frequency, Review Rating, Shipping Type.

Missing data: 37 values in the Review Rating column were handled.

Exploratory Data Analysis in Python



01

Data Loading & Exploration

Imported dataset with pandas, checked structure and summary statistics.

02

Missing Data Handling

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

03

Column Standardization

Renamed columns to snake_case for readability.

04

Feature Engineering

Created 'age_group' and 'purchase_frequency_days' columns.

05

Database Integration

Loaded cleaned data into PostgreSQL for SQL analysis.

Revenue by Gender & High-Spending Discount Users

Revenue by Gender

Female	75191
Male	157890

Male customers generated significantly higher revenue.

High-Spending Discount Users

Identified 839 customers who used discounts but still spent above average.



Product Performance & Shipping Insights

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
Express	60.48

Express shipping users have slightly higher average purchase amounts.

Subscription & Discount Dependency

Subscribers vs. Non-Subscribers

Yes	1053	59.49	62645.00
No	2847	59.87	170436.00

Non-subscribers contribute more to total revenue due to higher numbers.

Discount-Dependent Products

Hat	50.00
Sneakers	49.66
Coat	49.07
Sweater	48.17
Pants	47.37

These products show the highest percentage of purchases with discounts.

Customer Segmentation & Loyalty



Loyal Customers

3,116 customers with consistent purchase history.



New Customers

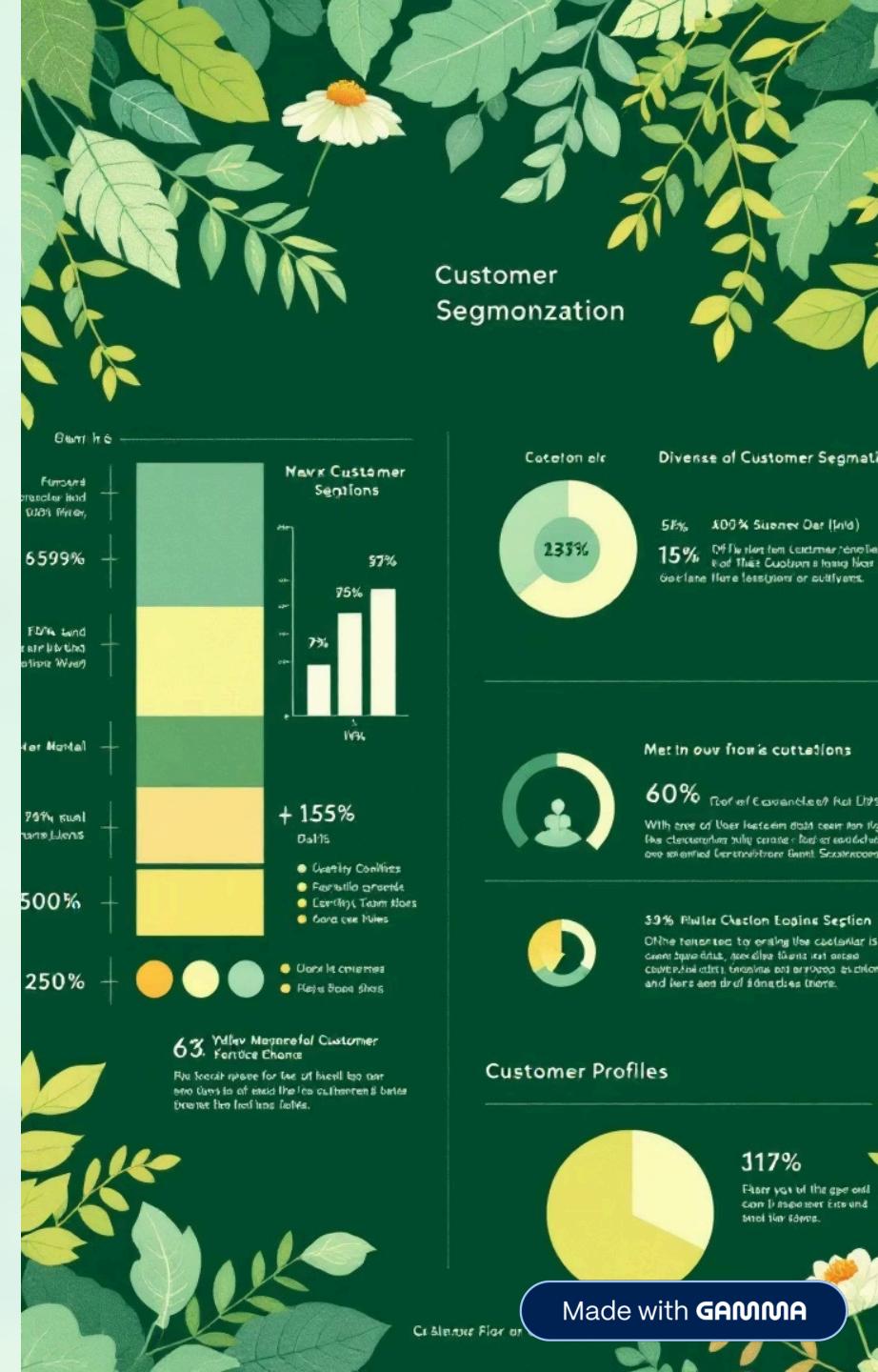
83 customers, indicating recent engagement.



Returning Customers

701 customers who have made repeat purchases.

Customers with more than 5 purchases are more likely to subscribe (958 vs 2518 non-subscribers).



Top Products per Category

1	Accessories	Jewelry	171
2	Accessories	Sunglasses	161
3	Accessories	Belt	161
1	Clothing	Blouse	171
2	Clothing	Pants	171
3	Clothing	Shirt	169
1	Footwear	Sandals	160
2	Footwear	Shoes	150
3	Footwear	Sneakers	145
1	Outerwear	Jacket	163
2	Outerwear	Coat	161

Blouse and Jewelry are top sellers in their respective categories.

Revenue by Age Group



Young Adults contribute the highest total revenue, followed closely by Middle-aged customers.

Business Recommendations

→ Boost Subscriptions

Promote exclusive benefits to increase subscriber base.

→ Customer Loyalty Programs

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

→ Review Discount Policy

Strategically balance sales boosts with margin control for discount-dependent products.

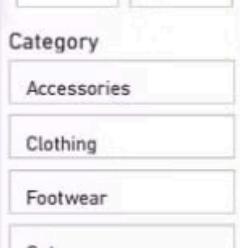
→ Product Positioning

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.

Customer Behavior Dashboard



3.9K

Number of Customers

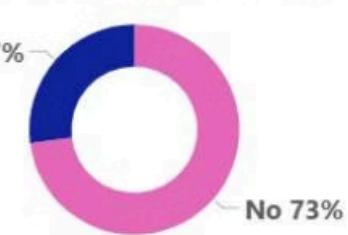
\$59.76

Average Purchase Amount

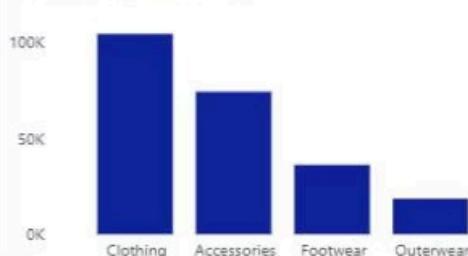
3.75

Average Review Rating

% of Customers by Subscription Status



Revenue by Category



Sales by Category



Revenue by Age Group



Sales by Age Group

