Customer Shopping Behavior Analysis

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

Project Overview & Dataset Summary

Project Goals

- Analyze spending patterns
- Identify customer segments
- Understand product preferences
- Examine subscription behavior
- Guide strategic business decisions

Dataset Snapshot

Our dataset comprises 3,900 rows and 18 columns, detailing customer demographics, purchase specifics, and shopping habits. Key features include Age, Gender, Location, Item Purchased, Category, Purchase Amount, and Subscription Status.

> Note: 37 missing values in the Review Rating column were addressed during data preparation.

Exploratory Data Analysis with Python

We initiated our analysis using Python for data preparation and cleaning, ensuring a robust foundation for insights.

01

Data Loading & Initial Exploration

Imported the dataset with `pandas` and used `df.info()` and `.describe()` for structural and statistical overviews.

02

Missing Data Handling

Null values were identified, and missing `Review Rating` values were imputed using the median rating per product category.

03

Column Standardization

Columns were renamed to `snake_case` for improved readability and consistency.

04

Feature Engineering
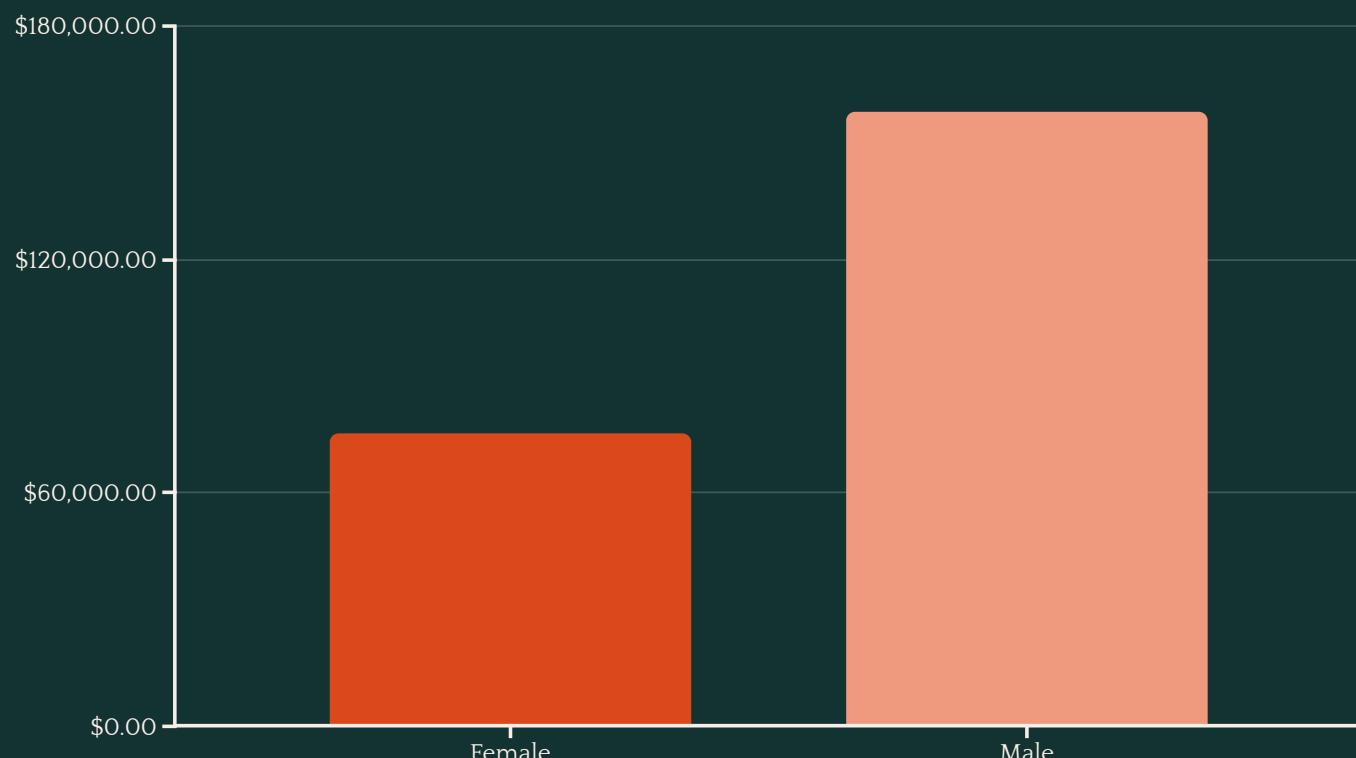
New columns like `age_group` and `purchase_frequency_days` were created to enrich the dataset.

05

Database Integration

The cleaned DataFrame was loaded into PostgreSQL for subsequent SQL-based analysis.

Key Insights: Revenue & Discounts

Revenue by Gender

Our analysis revealed a significant difference in revenue contribution, with male customers generating nearly double the revenue compared to female customers.

High-Spending Discount Users

We identified 839 customers who utilized discounts yet maintained purchase amounts above the average. This segment represents an opportunity for targeted promotions.

Bar chart:

	Revenue
Female	~$75,000.00
Male	~$150,000.00

Y-axis: $0.00, $60,000.00, $120,000.00, $180,000.00

Product Performance & Shipping Preferences
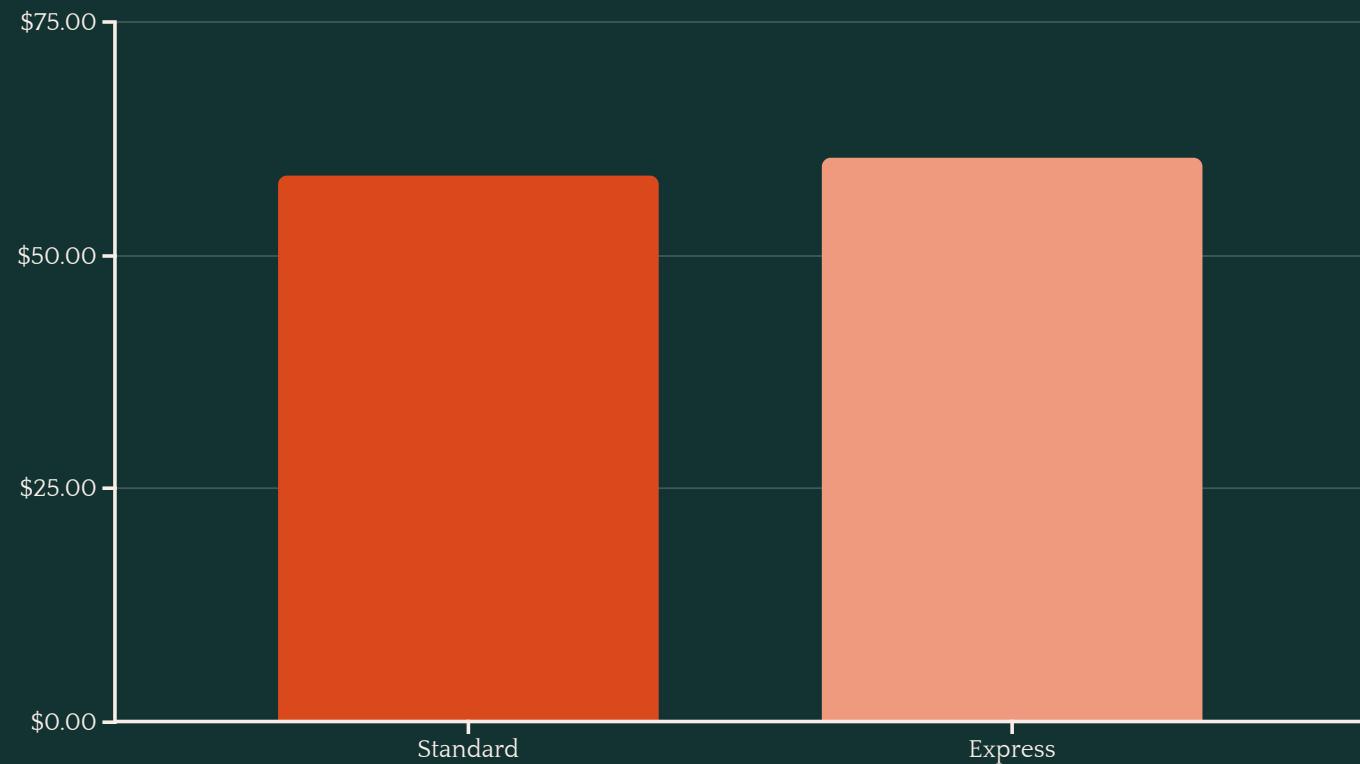
Top 5 Products by Rating

Gloves, Sandals, Boots, Hat, and Skirt emerged as the top-rated products, indicating strong customer satisfaction in these categories.

Gloves	3.86
Sandals	3.84
Boots	3.82
Hat	3.80
Skirt	3.78

Shipping Type Comparison

Customers opting for Express shipping tend to have slightly higher average purchase amounts ($60.48) compared to Standard shipping ($58.46).

Subscription & Discount Dynamics

Subscribers vs. Non-Subscribers

Non-subscribers contribute significantly more to total revenue ($170,436) than subscribers ($62,645), despite similar average spend per customer.

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

Discount-Dependent Products

Products like Hat, Sneakers, Coat, Sweater, and Pants show the highest percentage of discounted purchases, suggesting these items are often bought with promotions.

Hat	50.00
Sneakers	49.66
Coat	49.07
Sweater	48.17
Pants	47.37

Customer Segmentation & Loyalty

Customer Segments

Customers were classified into three segments based on their purchase history: Loyal (3116), Returning (701), and New (83). The majority are loyal customers.

Loyal	3116
Returning	701
New	83

Repeat Buyers & Subscriptions

Customers with more than 5 purchases are more likely to be subscribers (958) compared to non-subscribers (2518), indicating a correlation between purchase frequency and subscription status.
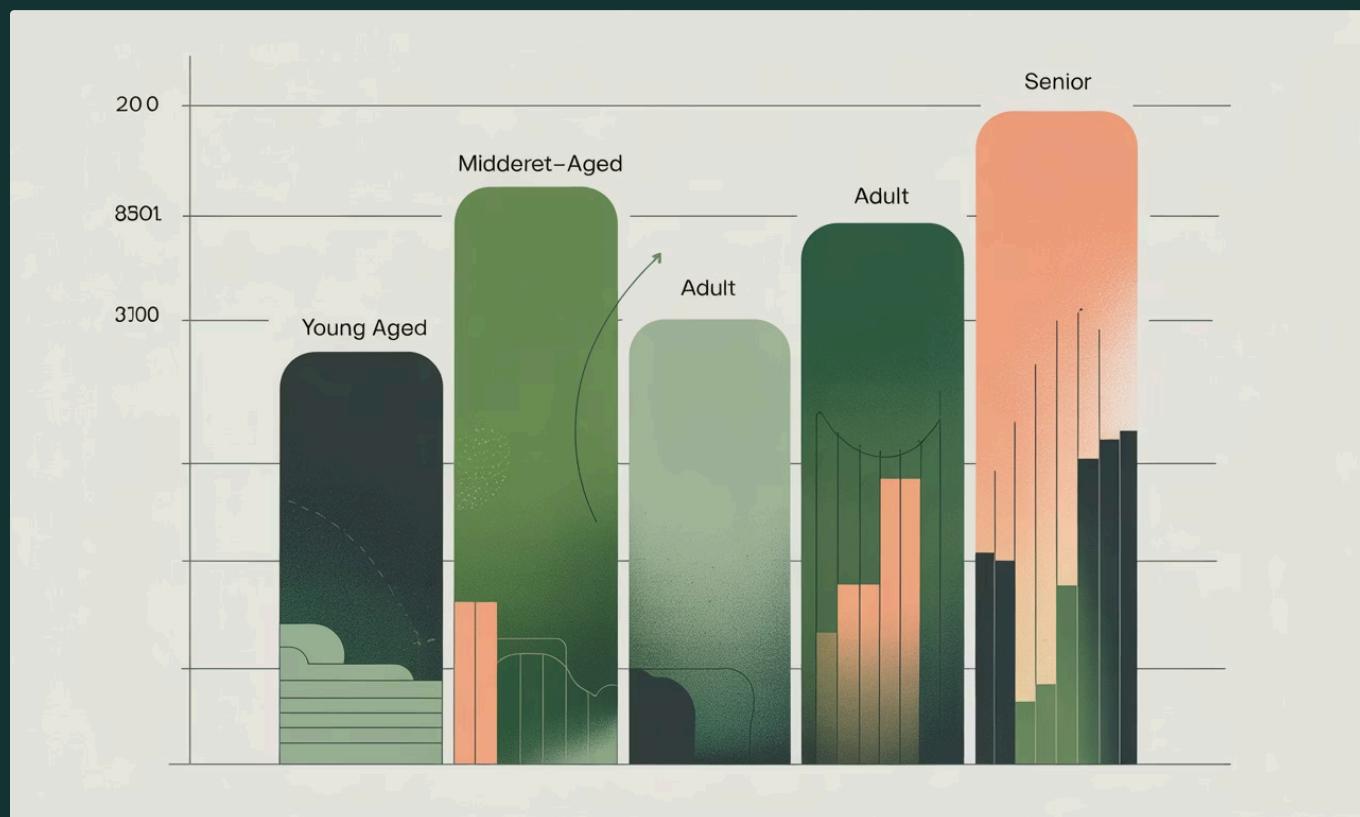
No	2518
Yes	958

Top Products by Category & Age Group Revenue

Top 3 Products per Category

Jewelry, Blouse, Sandals, and Jacket are consistently top sellers within their respective categories, highlighting strong product appeal.

1	Accessories	Jewelry	171
1	Clothing	Blouse	171
1	Footwear	Sandals	160
1	Outerwear	Jacket	163

Revenue by Age Group

Young Adults contribute the highest total revenue ($62,143), followed closely by Middle-aged customers ($59,197), indicating key demographic targets.

Young Adult	$62,143
Middle-aged	$59,197
Adult	$55,978
Senior	$55,763
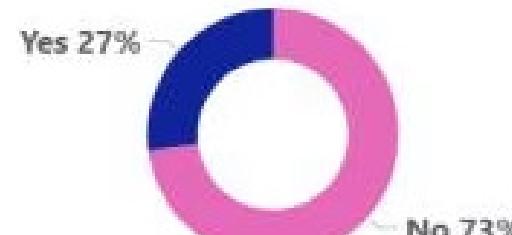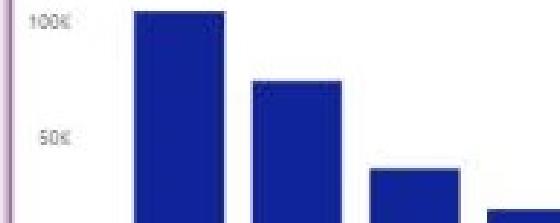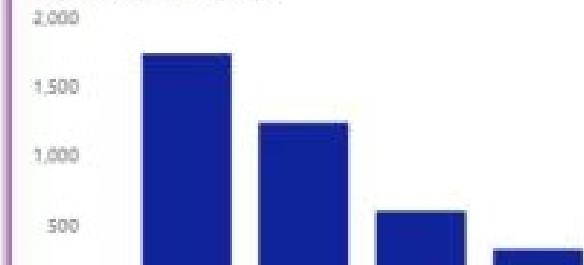
Interactive Dashboard in Power BI

An interactive Power BI dashboard was developed to visually present these insights, offering a dynamic way to explore customer behavior data.

3.9K
Customers

$59.76
Avg. Purchase

3.75
Avg. Review

Made with GAMMA

Strategic Business Recommendations

Based on our comprehensive analysis, we propose the following actionable recommendations to optimize business strategies and enhance customer engagement:

Boost Subscriptions

Promote exclusive benefits to increase subscriber base and retention.

Customer Loyalty Programs

Implement programs to reward repeat buyers and foster loyalty.

Review Discount Policy

Strategically balance sales boosts with margin control to maximize 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 optimal impact.