Customer Shopping Behavior Analysis

Uncovering insights from transactional data to drive strategic business decisions.

Project Overview

This project analyzes customer shopping behavior using transactional data from 3,900 purchases across various product categories.

Goal

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

Impact

Guide strategic business decisions through data-driven recommendations.

Dataset Summary

Our dataset comprises 3,900 rows and 18 columns, capturing a comprehensive view of customer interactions.

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

Key Features

Customer demographics, purchase details, and shopping behavior.

Demographics

Age, Gender, Location, Subscription Status.

Purchase Details
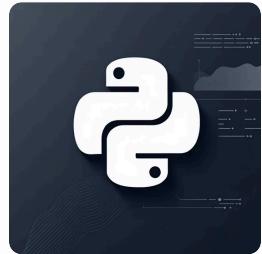
Item Purchased, Category, Amount, Season, Size, Color.

Behavior

Discount, Promo Code, Previous Purchases, Frequency, Review Rating, Shipping Type.
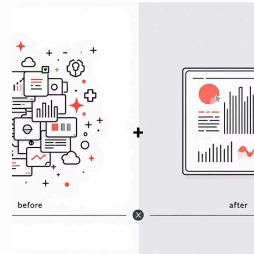
Exploratory Data Analysis with Python

Data preparation and cleaning were performed using Python to ensure data quality and readiness for analysis.

Data Loading & Exploration

Imported dataset using `pandas` and explored structure with `df.info()` and `describe()`.

Missing Data Handling

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

Feature Engineering
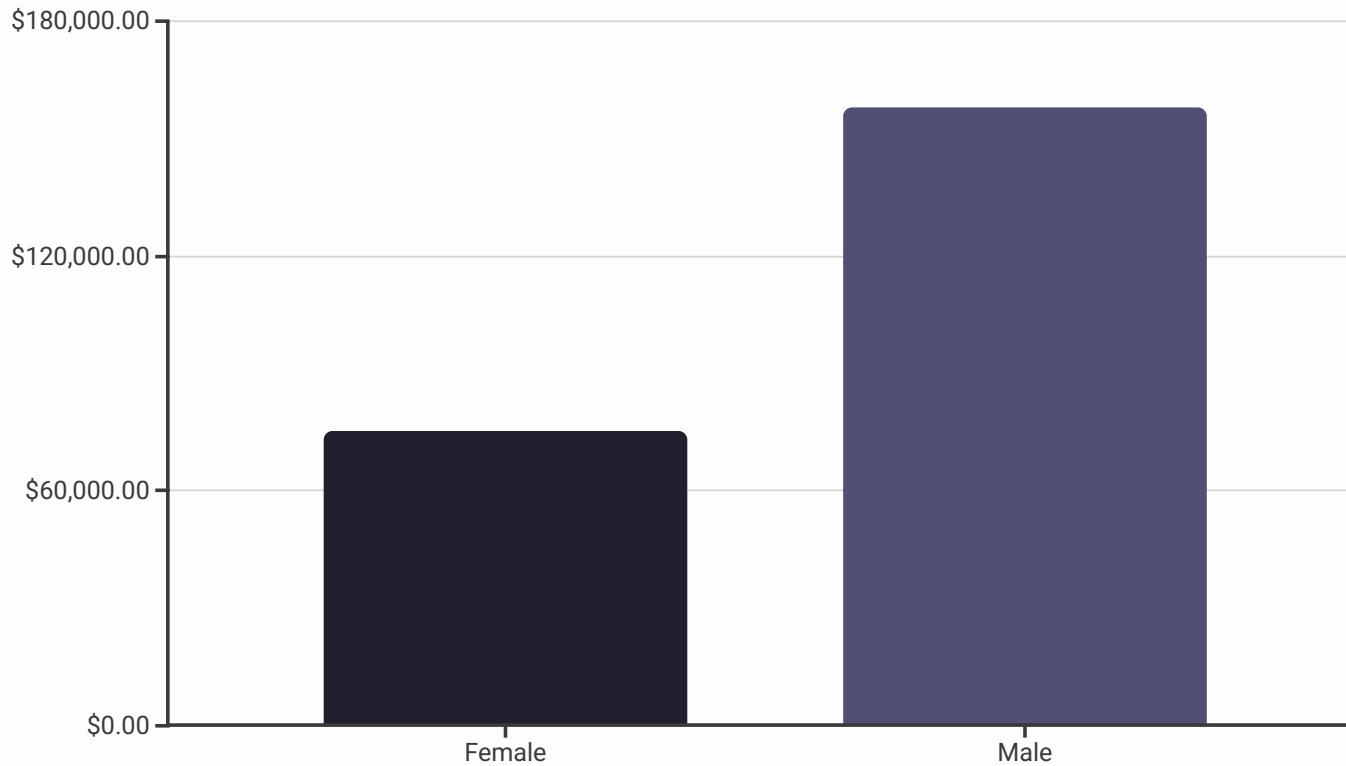
Created `age_group` and `purchase_frequency_days` columns for deeper insights.

Database Integration

Loaded the cleaned DataFrame into PostgreSQL for SQL analysis.

Revenue by Gender

Analysis revealed a significant difference in revenue contribution between genders.

- **Male Customers:** Generated $157,890 in revenue.
- **Female Customers:** Generated $75,191 in revenue.

This insight can inform targeted marketing strategies.

Top Products & Shipping Insights

Shipping Type Comparison

Average purchase amounts vary by shipping preference, suggesting different customer segments.

Shipping Type	Avg. Purchase
Standard	$58.46
Express	$60.48

Top 5 Products by Rating

Products with the highest average review ratings indicate strong customer satisfaction.

Item Purchased	Average Rating
Gloves	3.86
Sandals	3.84
Boots	3.82
Hat	3.80
Skirt	3.78

Subscription & Discount Behavior

Subscribers vs. Non-Subscribers

Comparing average spend and total revenue across subscription statuses.

Status	Customers	Avg. Spend	Total Revenue
Yes	1053	$59.49	$62,645
No	2847	$59.87	$170,436

Discount-Dependent Products

Identified products with the highest percentage of discounted purchases.

Item Purchased	Discount Rate (%)
Hat	50.00
Sneakers	49.66
Coat	49.07
Sweater	48.17
Pants	47.37

Customer Segmentation & Loyalty

Customer Segments

Customers were classified into segments based on their purchase history.

Repeat Buyers & Subscriptions

Customers with more than 5 purchases are more likely to subscribe.

Subscription Status	Repeat Buyers
No	2518
Yes	958

Loyal

3,116 Customers

Returning

701 Customers

New

83 Customers

Dashboard in Power BI

An interactive Power BI dashboard was created to visualize key insights and facilitate data exploration.

Business Recommendations

1 Boost Subscriptions

Promote exclusive benefits to increase subscriber base.

2 Customer Loyalty Programs

Reward repeat buyers to foster loyalty and retention.

3 Review Discount Policy

Optimize discount strategies to balance sales and profit margins.

4 Product Positioning

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

5 Targeted Marketing

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