

Customer Shopping Behavior Analysis

This project analyzes 3,900 customer purchases to uncover spending patterns, product preferences, and subscription behavior. Our goal is to guide strategic business decisions with data-driven insights.





Project Overview

Objective

Analyze customer shopping behavior from transactional data to inform strategic business decisions.

Data Source

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

Key Insights

Focus on spending patterns, customer segments, product preferences, and subscription impact.

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

Demographics, purchase details, shopping behavior, and subscription status.

Purchase Details

Item, category, amount, season, size, and color of purchased items.

Behavioral Data

Discounts, promo codes, previous purchases, frequency, and shipping type.

Exploratory Data Analysis (EDA) in Python

01

Data Loading & Initial Exploration

Imported data with `pandas`; used `df.info()` and `.describe()` for structural and statistical overview.

02

Missing Data Handling

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

03

Column Standardization

Renamed columns to `snake_case` for improved readability and consistency.

04

Feature Engineering

Created 'age_group' and 'purchase_frequency_days' for deeper analysis.

05

Data Consistency & Integration

Dropped redundant 'promo_code_used' and loaded cleaned data into PostgreSQL for SQL analysis.

Data Analysis using SQL: Key Business Insights

Revenue by Gender

Female	75191
Male	157890

Male customers generated significantly higher revenue.

Shipping Type Comparison

Standard	58.46
Express	60.48

Express shipping users have slightly higher average purchases.

Product Performance & Customer Segmentation

Top 5 Products by Rating

Gloves	3.86
Sandals	3.84
Boots	3.82
Hat	3.80
Skirt	3.78

Gloves, Sandals, and Boots lead in average customer satisfaction.

Customer Segmentation

Loyal	3116
New	83
Returning	701

A large loyal customer base, with opportunities to nurture new and returning buyers.

Subscription & Discount Dynamics



1

Subscribers vs. Non-Subscribers

Subscribers: 1053 customers, avg spend \$59.49, total revenue \$62,645. Non-subscribers: 2847 customers, avg spend \$59.87, total revenue \$170,436.

2

Discount-Dependent Products

Hat (50%), Sneakers (49.66%), Coat (49.07%), Sweater (48.17%), Pants (47.37%) show highest discount rates.

3

Repeat Buyers & Subscriptions

Out of repeat buyers (>5 purchases), 958 are subscribers, 2518 are not. Opportunity to convert more repeat buyers.

Customer Behavior Dashboard

K

Customers

Subscription Status

No 73%

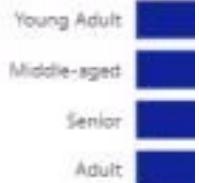
\$59.76

Average Purchase Amount

Revenue by Category



Sales by Age (K)



Interactive Power BI Dashboard

Our Power BI dashboard provides an interactive visual summary of key metrics and insights.

3.9K

Customers

\$59.76

Avg. Purchase

3.75

Avg. Rating

Strategic Business Recommendations

→ Boost Subscriptions

Promote exclusive benefits to increase subscriber base and retention.

→ Customer Loyalty Programs

Implement rewards for repeat buyers to foster loyalty and move them into the "Loyal" segment.

→ Review Discount Policy

Optimize discount strategies to balance sales growth with profit margins.

→ Product Positioning

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

→ Targeted Marketing

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