



# Customer Shopping Behavior Analysis

Uncovering insights from transactional data to drive strategic business decisions.

# Project Overview

1

## Analyze Customer Behavior

Using 3,900 purchases across various product categories.

2

## Uncover Key Insights

Spending patterns, customer segments, product preferences, and subscription behavior.

3

## Guide Decisions

Inform strategic business decisions with data-driven recommendations.

# Dataset Summary

## Data Volume

3,900 rows of transactional data.

## Data Points

18 columns including demographics, purchase details, and shopping behavior.

## 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, Shipping)

## Missing Data

37 values in Review Rating column.

# Exploratory Data Analysis (Python)

01

## Data Loading & Initial Exploration

Imported dataset with pandas; checked structure and summary statistics.

02

## Missing Data Handling

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

03

## Column Standardization

Renamed columns to snake\_case for readability.

04

## Feature Engineering

Created 'age\_group' and 'purchase\_frequency\_days'.

05

## Data Consistency & Integration

Dropped redundant 'promo\_code\_used'; loaded data into PostgreSQL for SQL analysis.

# Data Analysis (SQL)

## Revenue by Gender

Female: \$75,191, Male: \$157,890.

## High-Spending Discount Users

Identified 839 customers who used discounts but spent above average.

## 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.

## Subscribers vs. Non-Subscribers

Subscribers (1,053 customers) avg. spend \$59.49, Non-subscribers (2,847 customers) avg. spend \$59.87.

## Discount-Dependent Products

Hat (50%), Sneakers (49.66%), Coat (49.07%), Sweater (48.17%), Pants (47.37%).

# Customer Segmentation & Product Insights

## Customer Segmentation

- Loyal: 3,116 customers
- Returning: 701 customers
- New: 83 customers

## Top 3 Products per Category

- Accessories: Jewelry, Sunglasses, Belt
- Clothing: Blouse, Pants, Shirt
- Footwear: Sandals, Shoes, Sneakers
- Outerwear: Jacket, Coat

# Subscription & Age Group Analysis

## Repeat Buyers & Subscriptions

Customers with >5 purchases: 958 subscribers vs. 2,518 non-subscribers.

## Revenue by Age Group

- Young Adult: \$62,143
- Middle-aged: \$59,197
- Adult: \$55,978
- Senior: \$55,763

# Customer Behavior Dashboard



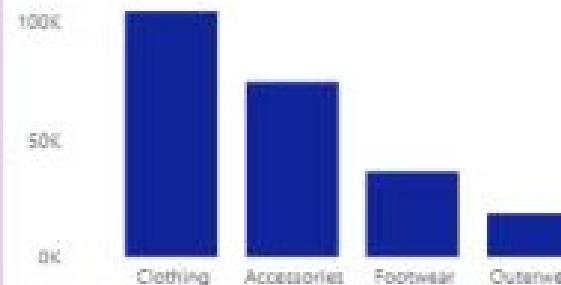
Customers

Subscription Status



\$59.76  
Average Purchase Amount

Revenue by Category



## Dashboard in Power BI

Interactive dashboard visualizing key metrics and trends for customer behavior.

Sales by Age



# Key Dashboard Metrics

**3.9K**

**Customers**

Total number of unique customers analyzed.



**\$59.76**

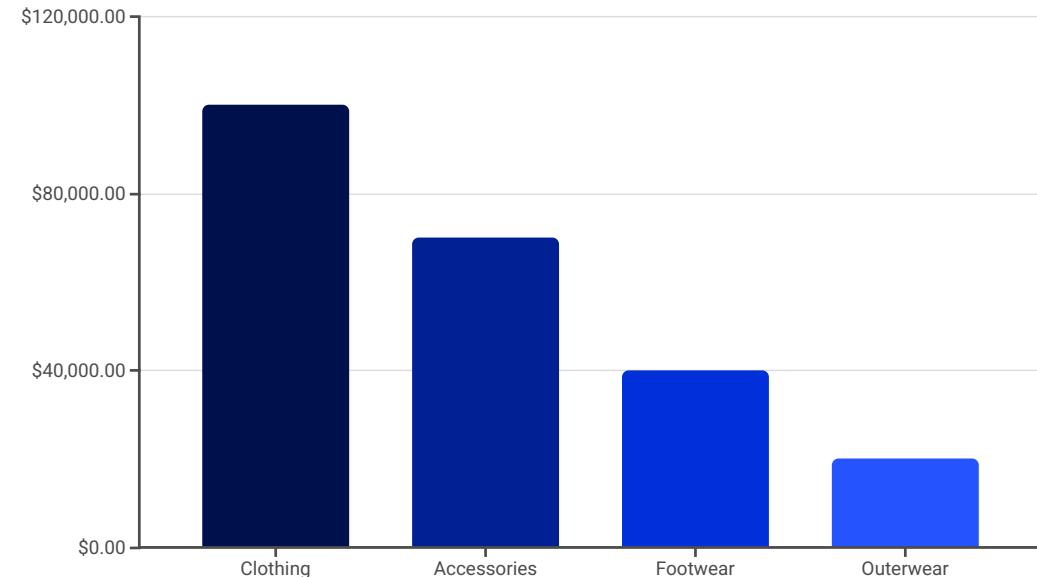
**Avg. Purchase Amount**

Average spend per transaction.

**3.75**

**Avg. Review Rating**

Overall average product review score.



Subscription Status: 27% Yes, 73% No.

Revenue by Category: Clothing leads with \$100K.

# Business Recommendations

## → **Boost Subscriptions**

Promote exclusive benefits to increase subscriber base.

## → **Customer Loyalty Programs**

Reward repeat buyers to foster loyalty and retention.

## → **Review Discount Policy**

Optimize discount strategies to balance sales and profit margins.

## → **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.