Customer Shopping Behavior Analysis

Uncovering purchasing patterns, customer segments, product preferences, and subscription behavior to drive smarter business decisions.

Project Overview

What We Set Out to Do

This project analyzes customer shopping behavior using transactional data to uncover insights across four key dimensions:

- Purchasing patterns
- Customer segments
- Product preferences
- Subscription behavior

Dataset at a Glance

3,900

Total Rows

Individual customer transactions analyzed

18

Columns

Distinct features per transaction record

37

Missing Values

Found only in the Review Rating column

Customer Demographics

Key demographic features captured in the dataset to segment and understand the customer base.

Age

Spans a wide range of customer age groups

Gender

Captures male, female, and other identities

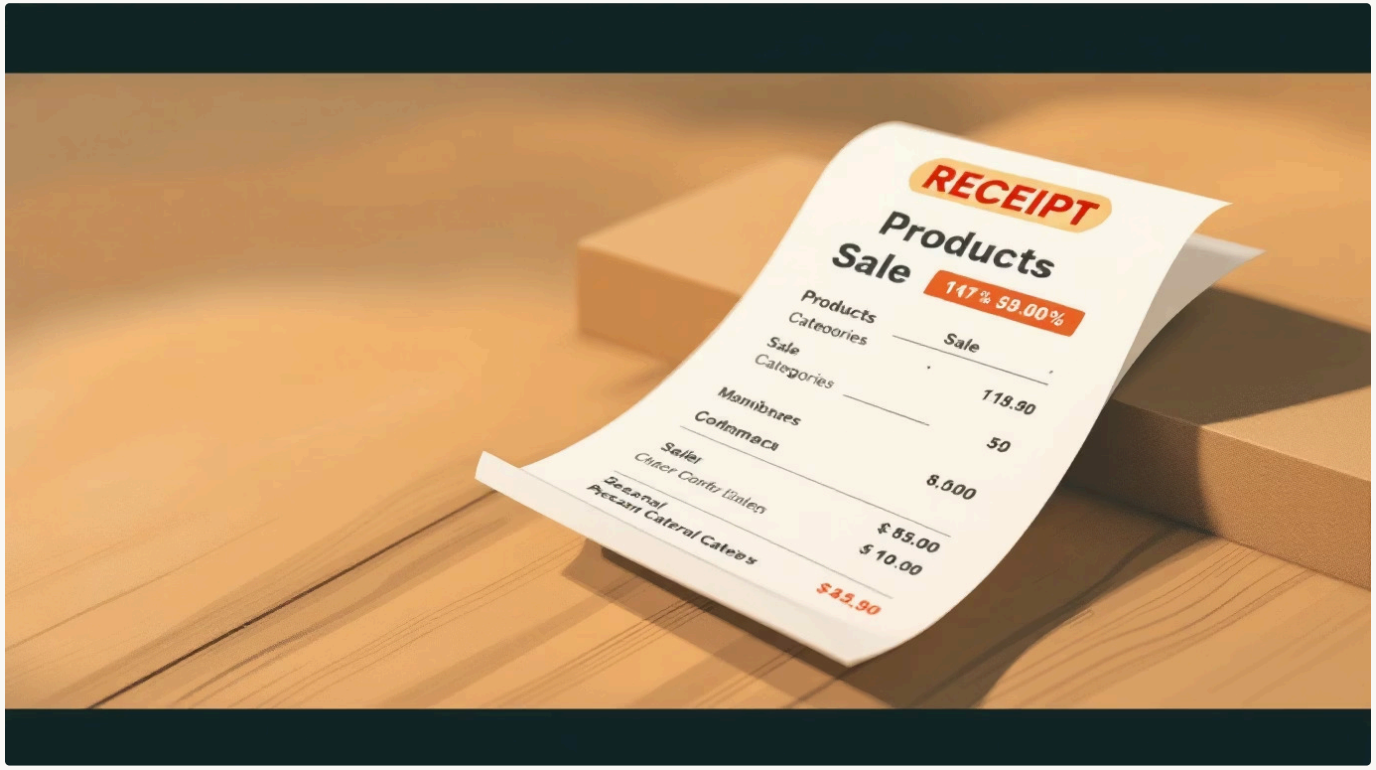
Location

Geographic distribution of customers

Subscription Status

Whether the customer holds an active subscription

Purchase Details



What Was Bought & When

Each transaction captures rich purchase-level detail:

- Item Purchased & Category
- Purchase Amount
- Season of purchase

These fields enable trend analysis across product lines and time periods.

Shopping Behavior Features

Discounts & Promos

Tracks whether a discount or promo code was applied at checkout

Previous Purchases

Number of prior transactions per customer

Review Rating

Customer satisfaction score (37 missing values noted)

Shipping Type

Preferred delivery method selected by the customer

Data Quality Note

Identify Missing Data

Find 37 nulls in Review Rating column.

Handle Gaps

Impute values or exclude rows for clean analysis.

Assess Impact

Confirm issue is minimal and localized.

Missing Data

Only **37 missing values** were identified across the entire dataset — all concentrated in the **Review Rating** column.

This represents a minimal data quality issue and was addressed prior to analysis.

Tools: Python

Pandas

Data wrangling, cleaning, and transformation

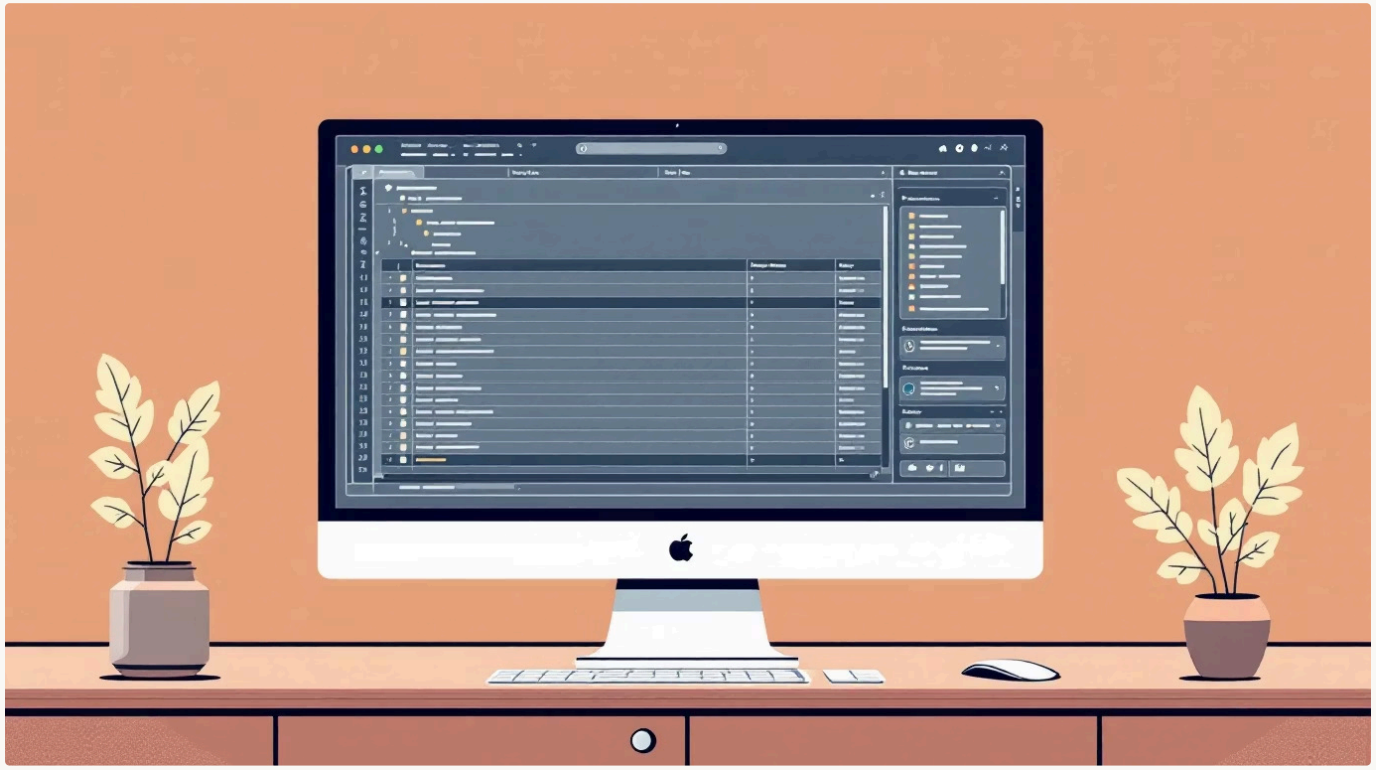
Matplotlib

Custom static visualizations and plots

Seaborn

Statistical graphics and heatmaps

Tools: SQL & Power BI



SQL

Used for structured querying, aggregation, and filtering of transactional records directly from the database.

Power BI

Interactive dashboards built to visualize customer segments, purchase trends, and subscription insights for stakeholders.

Key Takeaways

Rich Dataset

3,900 rows × 18 features covering demographics, purchases & behavior

Minimal Data Issues

Only 37 missing values — high-quality foundation for analysis

Full Toolchain

Python, SQL & Power BI deliver end-to-end analytical capability