Customer Shopping Behavior Analysis

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

Project Overview

Data-Driven Insights
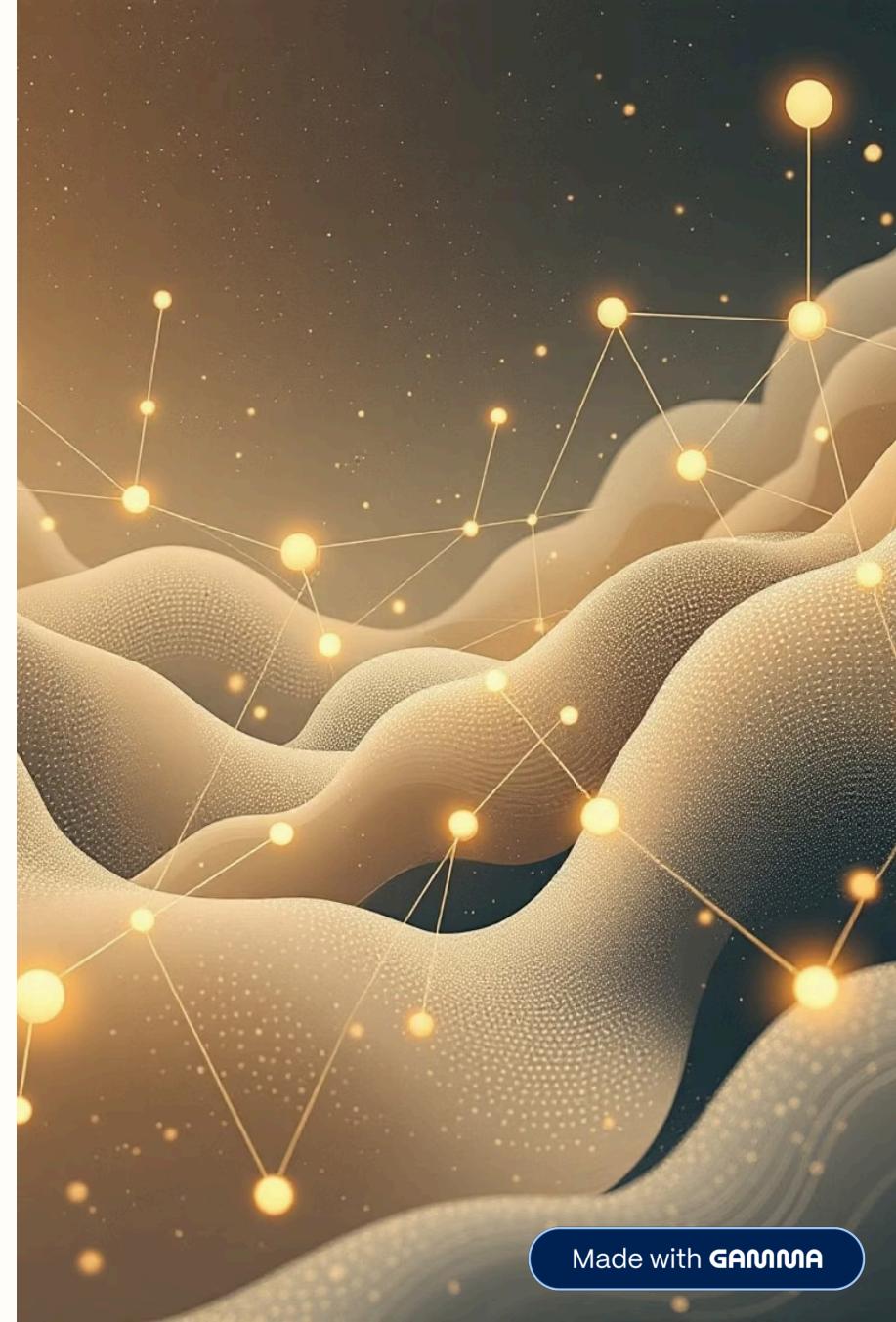
Analyzing 3,900 purchases to understand customer spending and product preferences.

Strategic Decisions

Guiding business strategies through insights into customer segments and subscription behavior.

Comprehensive Analysis

Covering demographics, purchase details, and shopping behavior metrics.

Dataset Summary

Data Volume
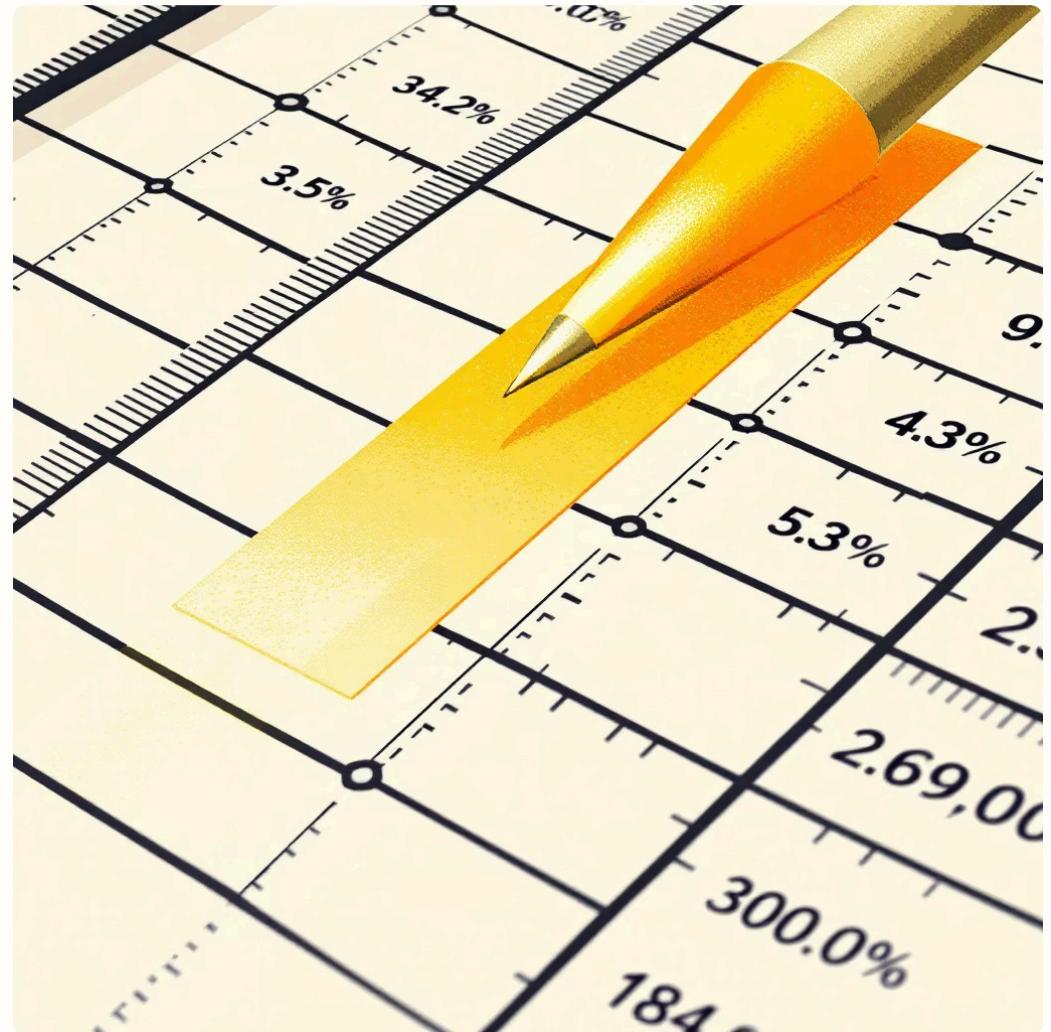
3,900 rows and 18 columns of transactional data.

Key Features

Demographics, purchase details, and shopping behavior (e.g., discount use, previous purchases).

Missing Data

37 values in the 'Review Rating' column.

Exploratory Data Analysis (EDA) with 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 category.

03

Column Standardization

Renamed columns to snake_case for readability.

04
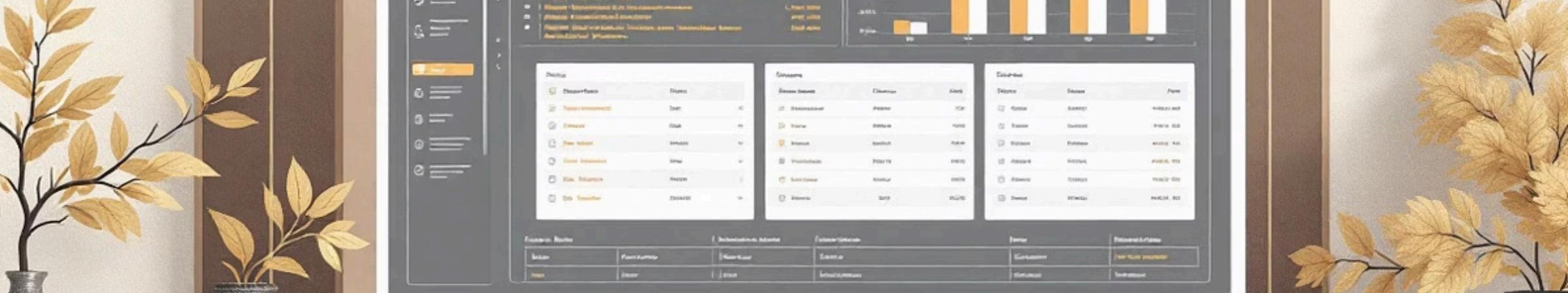
Feature Engineering

Created 'age_group' and 'purchase_frequency_days' columns.

05

Database Integration

Loaded cleaned data into PostgreSQL for SQL analysis.

Data Analysis with SQL: Key Business Insights

1

Revenue by Gender

Male customers generated significantly more revenue ($157,890) than female customers ($75,191).

2

High-Spending Discount Users

Identified 839 customers who used discounts but still spent above average.

3

Top 5 Products by Rating

Gloves (3.86), Sandals (3.84), Boots (3.82), Hat (3.80), Skirt (3.78) are top-rated.

4

Shipping Type Comparison

Express shipping users had a slightly higher average purchase amount ($60.48) than Standard ($58.46).

Made with GAMMA

SQL Insights: Subscriptions & Discounts

Subscribers vs. Non-Subscribers

Subscribers (1053) have an average spend of $59.49, non-subscribers (2847) average $59.87.

Discount-Dependent Products

Hat (50%), Sneakers (49.66%), Coat (49.07%), Sweater (48.17%), Pants (47.37%) are most discounted.
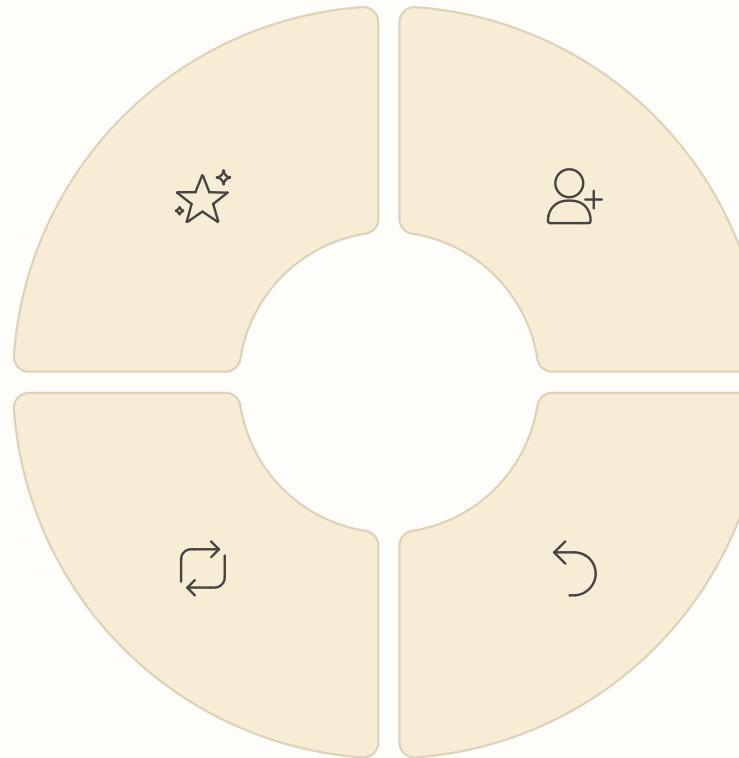
Customer Segmentation & Repeat Buyers

Loyal Customers

3,116 customers classified as loyal.

New Customers

83 customers identified as new.

Repeat Buyers & Subscriptions

958 repeat buyers are subscribers, 2518 are not.

Returning Customers

701 customers are returning.

Top Products by Category & Age Group Revenue

Top 3 Products per Category

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

Revenue by Age Group

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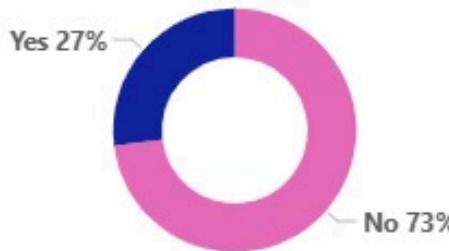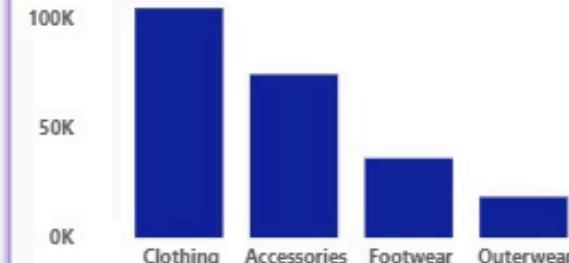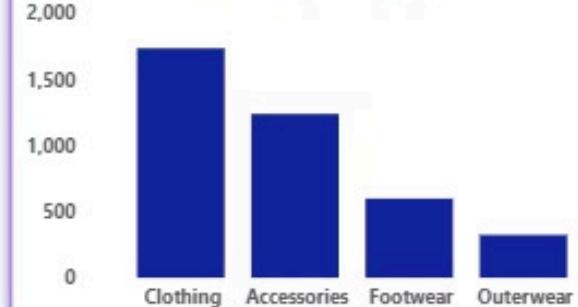
Interactive Dashboard in Power BI

An interactive dashboard was built in Power BI to visually present key insights from the customer shopping behavior analysis.

3.9K

Customers

$59.76

Avg. Purchase

3.75

Avg. Review

Business Recommendations

Boost Subscriptions
Promote exclusive benefits for subscribers to increase sign-ups.

Customer Loyalty Programs
Reward repeat buyers to foster loyalty and move them into the "Loyal" segment.

Review Discount Policy
Strategically balance sales boosts with margin control when applying discounts.

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