Customer Shopping Behavior Analysis

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

Project Overview

Goal

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

Data Source

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

Purpose

Guide strategic business decisions and optimize marketing efforts.
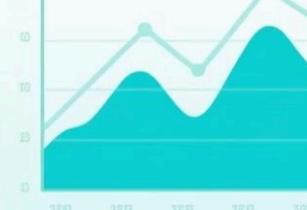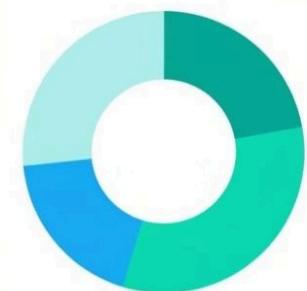
Dataset Summary

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

- **Customer Demographics:** Age, Gender, Location, Subscription Status.

- **Purchase Details:** Item Purchased, Category, Purchase Amount, Season, Size, Color.

- **Shopping Behavior:** Discount Applied, Promo Code Used, Previous Purchases, Frequency of Purchases, Review Rating, Shipping Type.

Missing Data: 37 values in the "Review Rating" column were handled during cleaning.

Exploratory Data Analysis (EDA) in Python

01

Data Loading & Exploration

Imported data with pandas, checked structure and summary statistics.

02

Missing Data Handling

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

03

Column Standardization

Renamed columns to snake_case for readability and consistency.

04
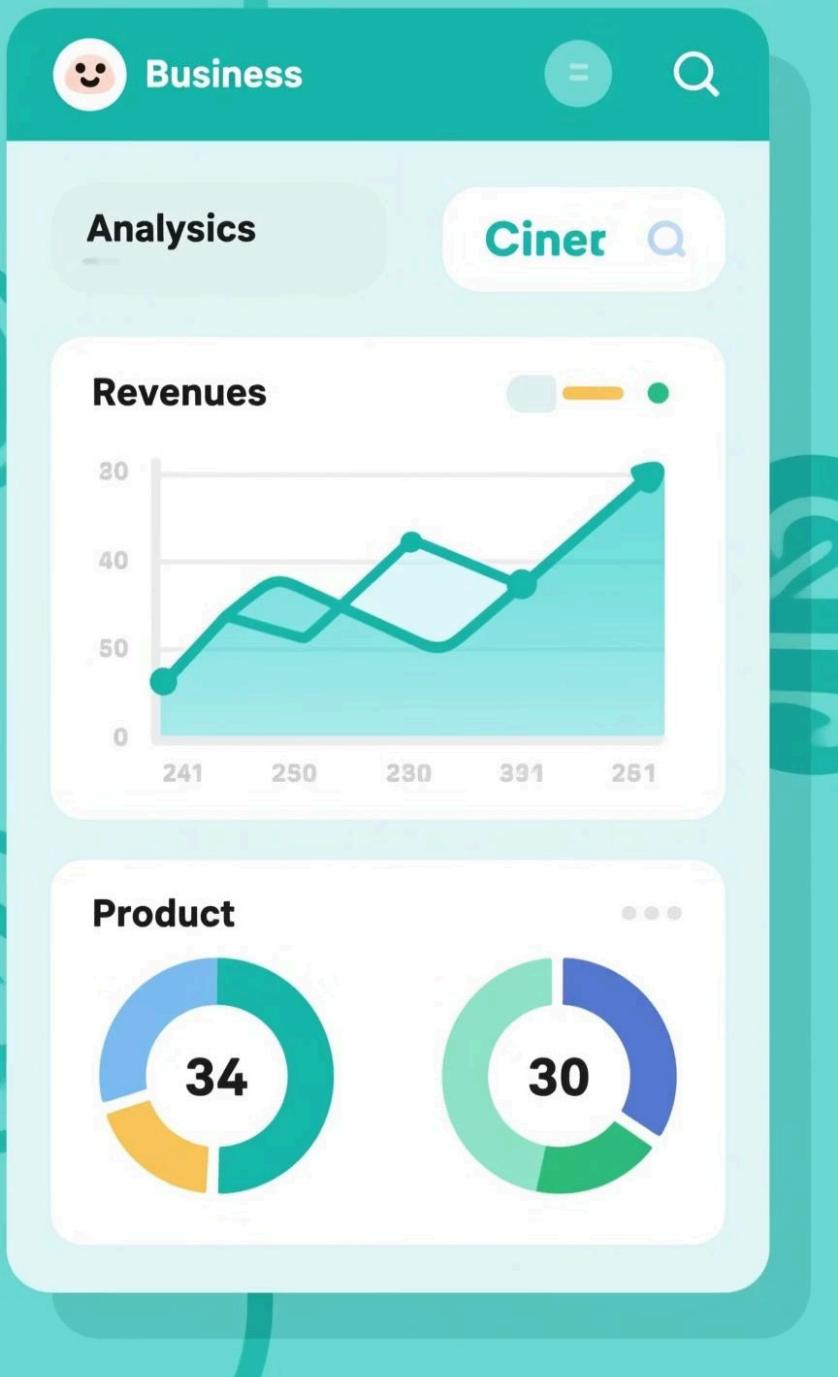
Feature Engineering

Created 'age_group' and 'purchase_frequency_days' columns.

05

Database Integration

Loaded cleaned data into PostgreSQL for SQL analysis.

Key Business Insights from SQL Analysis

Revenue by Gender

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

Top 5 Products by Rating

Gloves, Sandals, Boots, Hat, and Skirt received the highest average review ratings.

Shipping Type

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

Made with GAMMA

Customer Behavior & Product Insights

High-Spending Discount Users

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

Discount-Dependent Products

Hat (50%), Sneakers (49.66%), and Coat (49.07%) had the highest percentage of discounted purchases.

Subscribers vs. Non-Subscribers

Non-subscribers had a higher total revenue ($170,436) compared to subscribers ($62,645), despite similar average spend.

Customer Segmentation & Loyalty

New Customers

83 customers

Loyal Customers

3,116 customers

Returning Customers
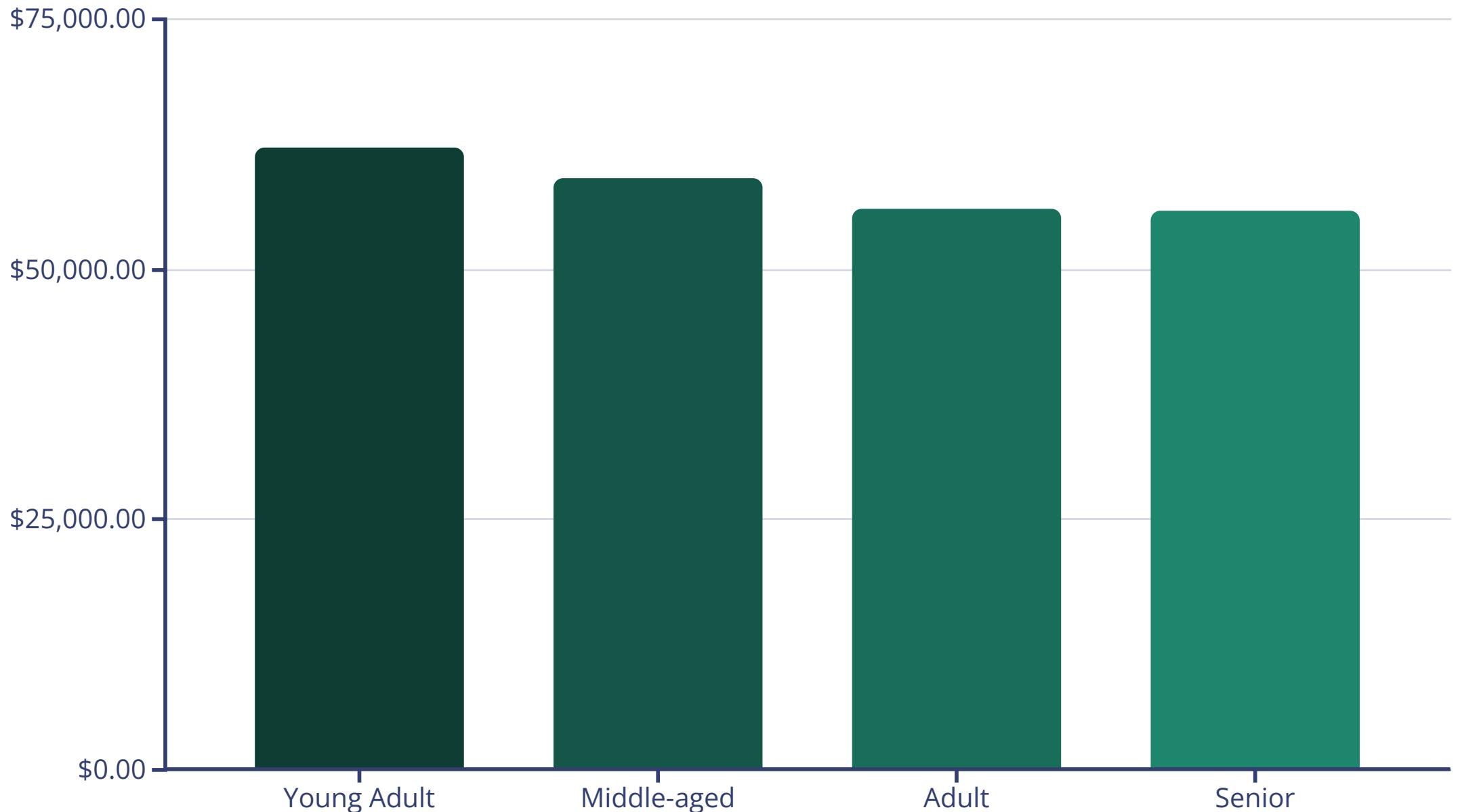
701 customers

Customers with more than 5 purchases are more likely to subscribe (958 Yes vs. 2518 No).

Top Products by Category

1	Accessories	Jewelry	171
1	Clothing	Blouse	171
1	Footwear	Sandals	160
1	Outerwear	Jacket	163

Revenue by Age Group

Young Adults contribute the highest revenue, followed closely by Middle-aged customers.

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**

Balance sales boosts with margin control for sustainable growth.

→ **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.