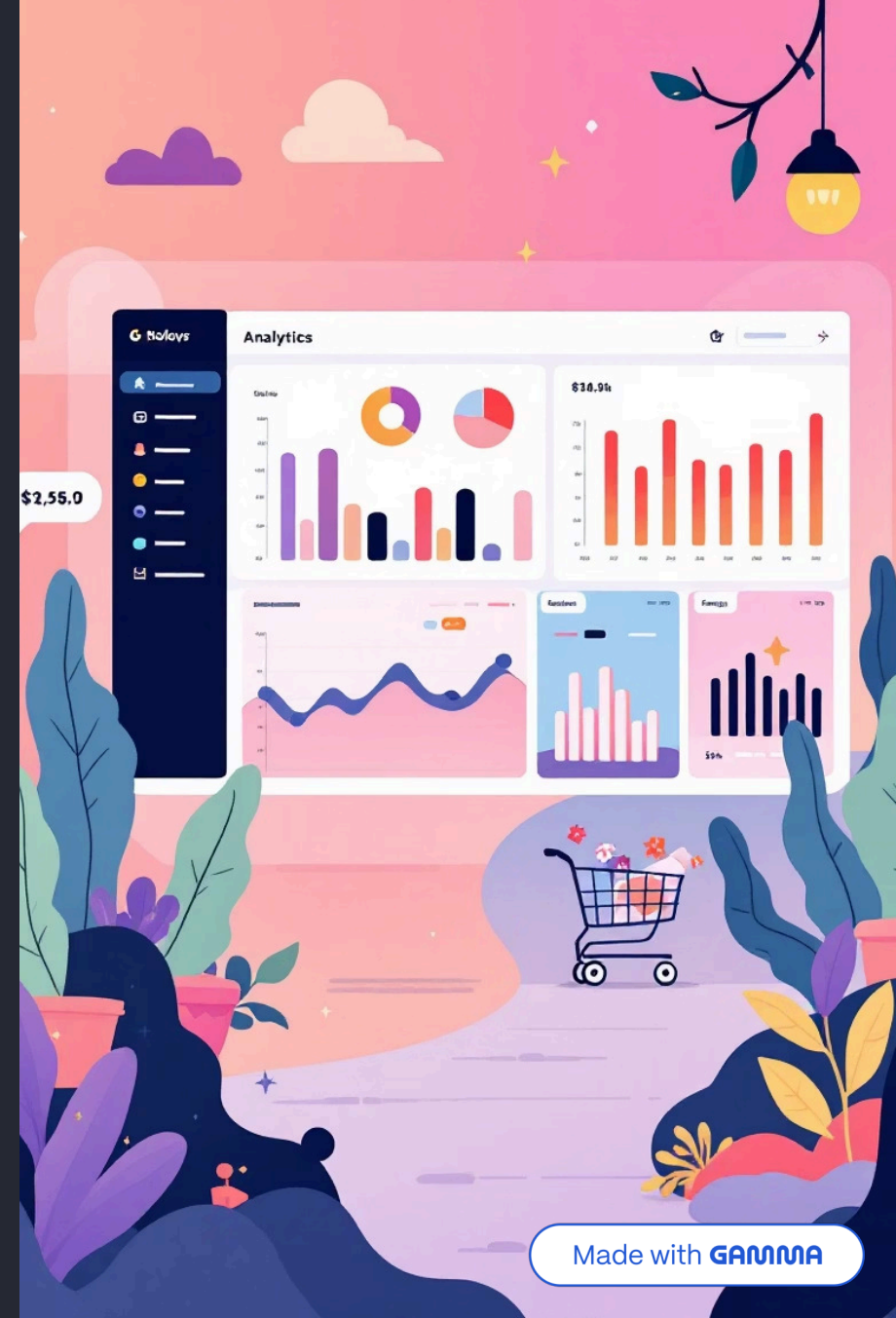


Customer Shopping Behavior Analysis

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



Dataset Overview

3,900

Total Purchases

Transactions analyzed across all
categories

18

Data Features

Columns capturing customer and
purchase details

50

Locations

Geographic distribution of
customers

\$59.76

Average Purchase

Mean transaction value in USD

The dataset includes comprehensive customer demographics (age, gender, location, subscription status), purchase details (item, category, amount, season, size, color), and shopping behavior metrics (discounts, promo codes, previous purchases, frequency, review ratings, shipping type). Only 37 missing values were found in the Review Rating column.

Data Preparation & Cleaning Process

01

Data Loading & Exploration

Imported dataset using pandas, examined structure with `df.info()` and summary statistics with `.describe()`

02

Missing Data Handling

Imputed 37 missing Review Rating values using median rating of each product category

03

Standardization & Engineering

Renamed columns to snake case, created `age_group` bins and `purchase_frequency_days` features

04

Data Consistency

Verified redundancy between `discount_applied` and `promo_code_used`, dropped duplicate column

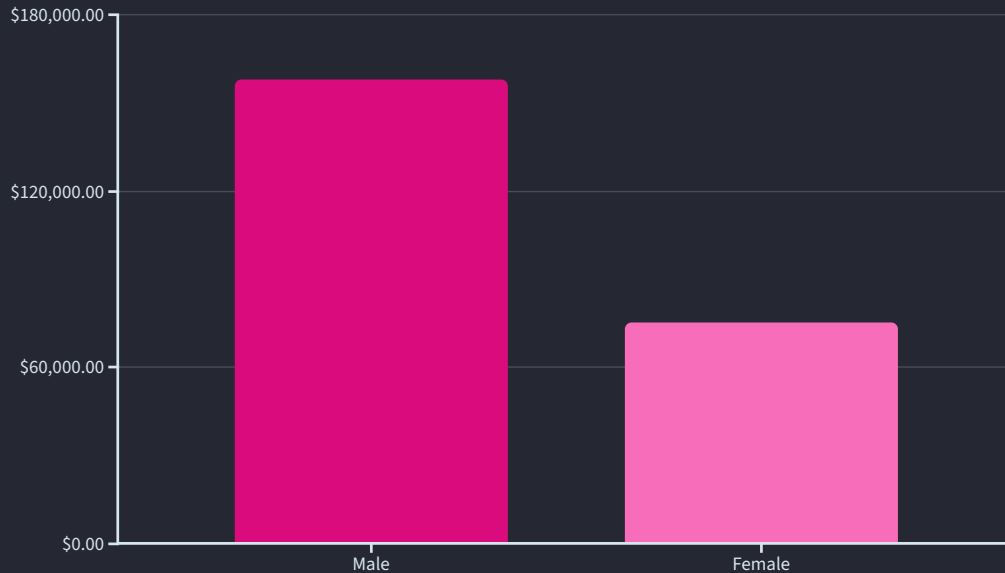
05

Database Integration

Connected Python to PostgreSQL and loaded cleaned DataFrame for SQL analysis

Revenue Analysis by Gender & Subscription

Gender Revenue Comparison



Male customers generate **2.1x more revenue** than female customers, representing 68% of total revenue.

Subscription Impact

Subscribers

1,053 customers

\$59.49 avg spend

\$62,645 total revenue

Non-Subscribers

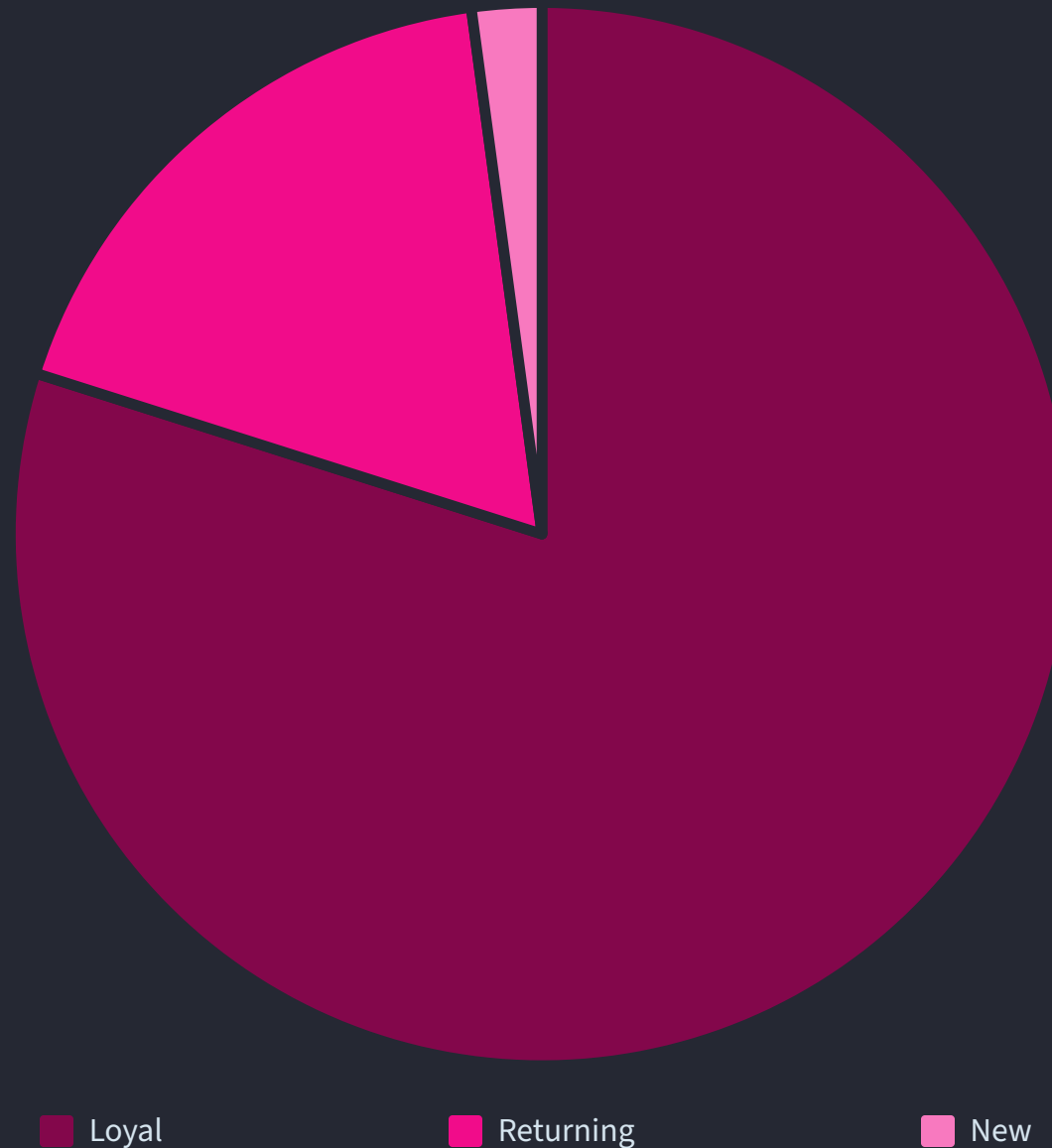
2,847 customers

\$59.87 avg spend

\$170,436 total revenue

Despite similar average spending, non-subscribers drive 73% of total revenue due to larger customer base.

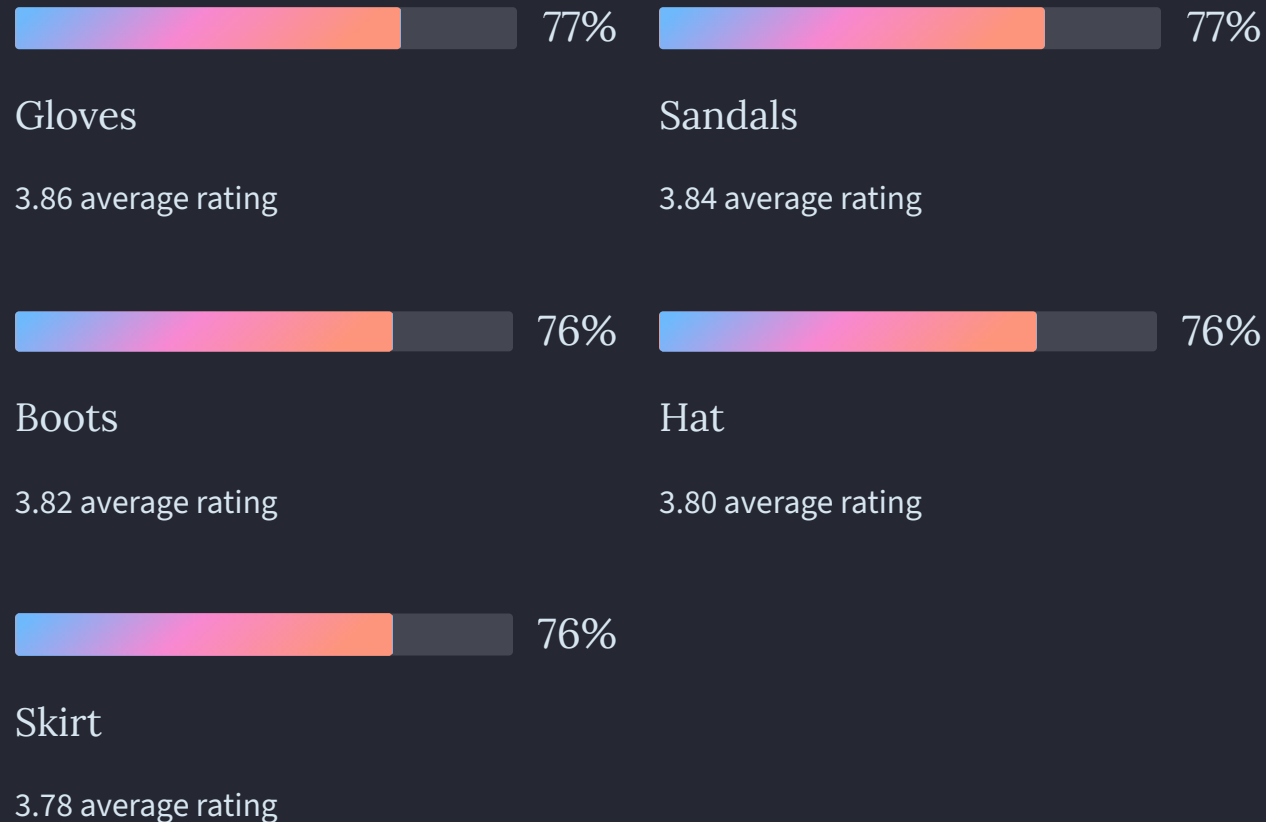
Customer Segmentation Insights



Customer segmentation reveals a strong loyal base comprising **80% of all customers**. Returning customers represent 18%, while new customers account for just 2%. Among repeat buyers with more than 5 purchases, 958 are subscribers while 2,518 are not, indicating significant opportunity for subscription conversion.

Product Performance & Ratings

Top 5 Products by Rating



Most Discount-Dependent

1. **Hat** - 50% discount rate
2. **Sneakers** - 49.66% discount rate
3. **Coat** - 49.07% discount rate
4. **Sweater** - 48.17% discount rate
5. **Pants** - 47.37% discount rate

These products show high dependency on discounts, suggesting potential margin pressure.

Category Performance Analysis

Clothing

Top sellers: Blouse (171), Pants (171), Shirt (169)

Largest category with highest transaction volume

Accessories

Top sellers: Jewelry (171), Sunglasses (161), Belt (161)

Strong performance across jewelry and fashion accessories

Footwear

Top sellers: Sandals (160), Shoes (150), Sneakers (145)

Consistent demand across casual and formal options

Outerwear

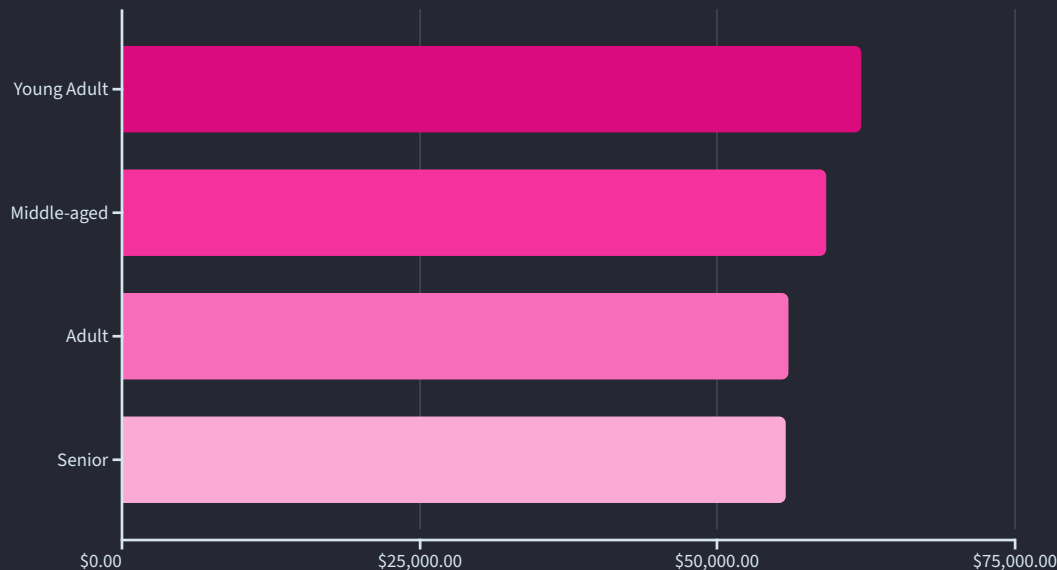
Top sellers: Jacket (163), Coat (161)

Seasonal category with concentrated product preferences



Age Group & Shipping Analysis

Revenue by Age Group



Shipping Type Preferences



Express

\$60.48 avg purchase



Standard

\$58.46 avg purchase

Express shipping customers spend **\$2.02 more** on average, indicating higher-value customer segment.

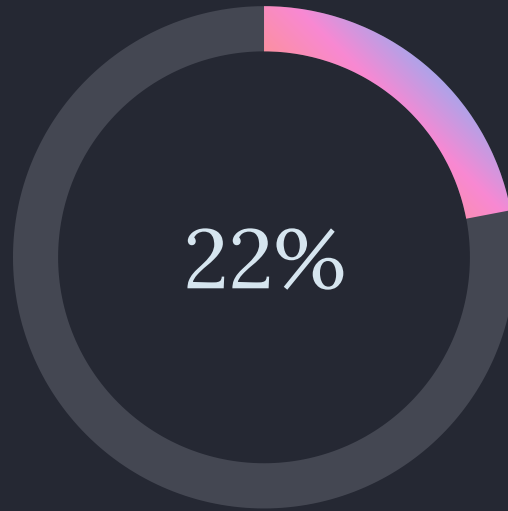
Young adults lead revenue generation at \$62,143, with relatively balanced distribution across all age groups. This suggests broad market appeal and opportunity for age-specific targeting strategies.

High-Value Discount Users



Smart Shoppers

Customers who used discounts but spent above average (\$59.76)



Of All Customers

Significant portion combining value-seeking with high spending



Average Spend

These customers spend 25% more than overall average

This segment represents a valuable opportunity: customers who respond to promotions while maintaining high purchase values. They demonstrate that discounts can drive revenue without necessarily reducing transaction size, suggesting strategic discount targeting could boost both volume and value.

Strategic Recommendations

1

Boost Subscription Conversion

With 73% of customers unsubscribed, promote exclusive benefits and rewards to convert high-frequency buyers into subscribers

2

Implement Loyalty Programs

Reward repeat buyers to accelerate movement from "Returning" to "Loyal" segment, focusing on the 701 returning customers

3

Optimize Discount Strategy

Review products with 45%+ discount rates to balance sales volume with margin protection, especially for Hats and Sneakers

4

Highlight Top Performers

Feature high-rated products (Gloves, Sandals, Boots) and best-sellers (Blouse, Jewelry, Jacket) in marketing campaigns

5

Target High-Value Segments

Focus marketing on Young Adults, express-shipping users, and male customers who drive disproportionate revenue