



Customer Shopping Behavior Analysis

Insights from 3,900 Retail Transactions Across Product Categories

Project Overview

Our Mission

Analyze customer shopping behavior using transactional data from 3,900 purchases to uncover insights into spending patterns, customer segments, product preferences, and subscription behavior.

01

Data Exploration

Python analysis and cleaning

02

SQL Analysis

Business transactions queries

03

Visualization

Power BI dashboard creation

04

Recommendations

Strategic business insights



Dataset at a Glance

3,900

Total Transactions

Comprehensive purchase records analyzed

18

Data Columns

Rich feature set for analysis

4

Product Categories

Clothing, Accessories, Footwear, Outerwear

50

Locations

Geographic diversity across states

Data Preparation & Cleaning



Data Loading

Imported dataset using pandas and performed initial exploration with `df.info()` and `.describe()` for summary statistics.



Missing Data Handling

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



Feature Engineering

Created `age_group` column by binning customer ages and `purchase_frequency_days` from purchase data.

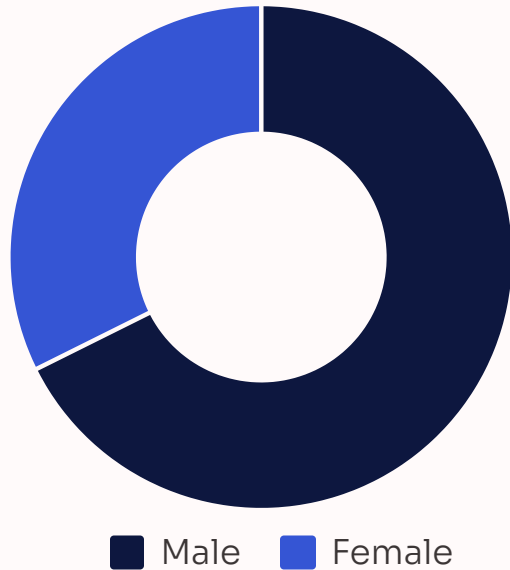


Database Integration

Connected Python to PostgreSQL and loaded cleaned DataFrame for SQL analysis.

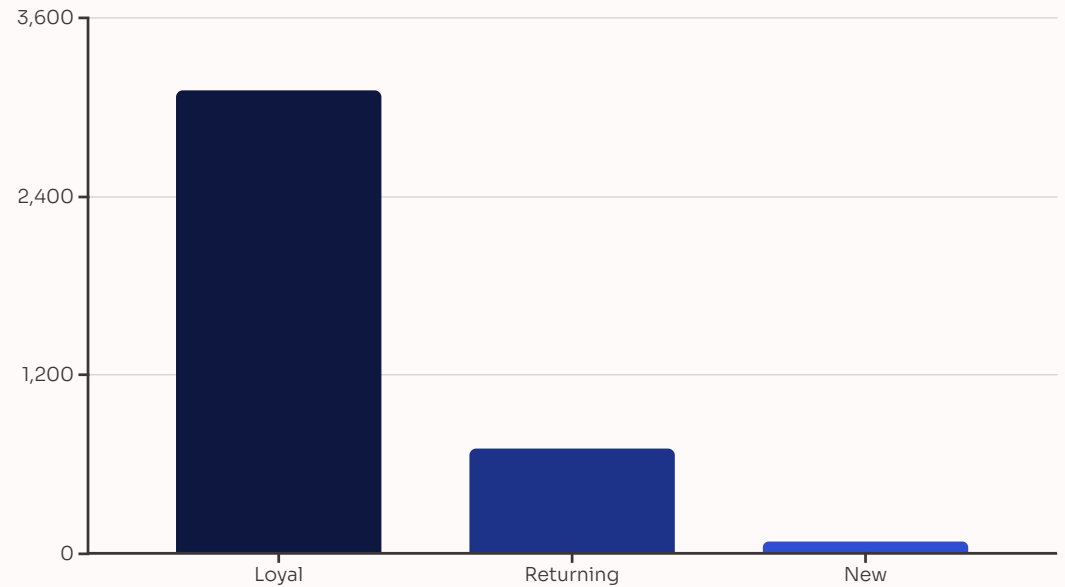
Revenue & Customer Insights

Revenue by Gender



Male customers generate 68% of total revenue, representing a significant opportunity for targeted marketing.

Customer Segmentation



80% of customers are loyal, indicating strong retention but limited new customer acquisition.

Product Performance Analysis



Top Rated Products

1. Gloves (3.86)
2. Sandals (3.84)
3. Boots (3.82)
4. Hat (3.80)
5. Skirt (3.78)



Best Sellers by Category

- Clothing: Blouse (171 orders)
- Accessories: Jewelry (171 orders)
- Footwear: Sandals (160 orders)
- Outerwear: Jacket (163 orders)

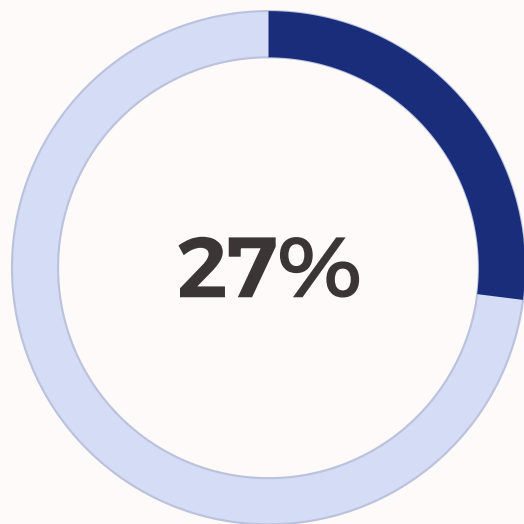


Discount-Dependent Items

Hat (50%), Sneakers (49.66%), Coat (49.07%), Sweater (48.17%), and Pants (47.37%) show highest discount rates.

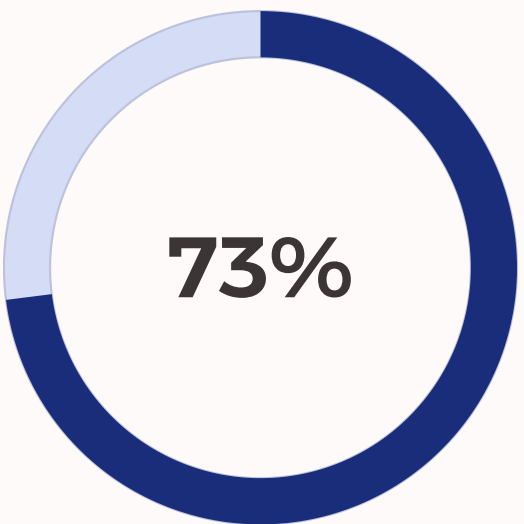


Subscription & Loyalty Patterns



Subscription Rate

1,053 customers with active subscriptions



Non-Subscribers

2,847 customers without subscriptions

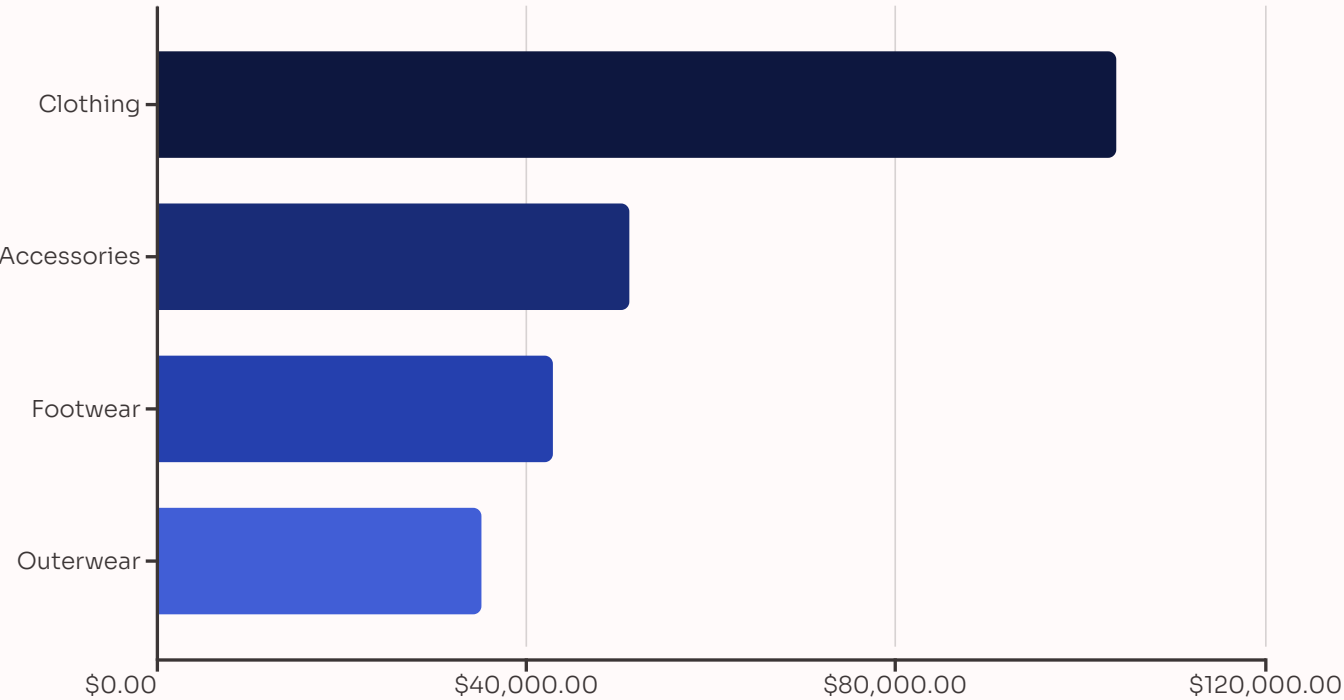
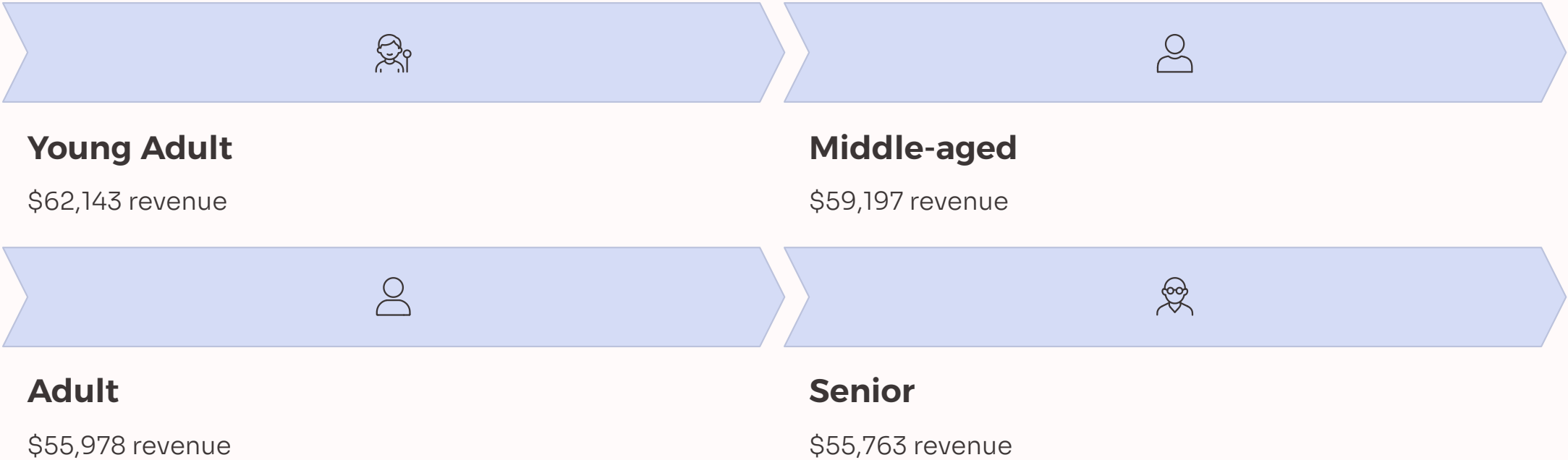
Spending Comparison

Subscribers average **\$59.49** per purchase while non-subscribers spend **\$59.87**. Despite similar spending, subscribers generate \$62,645 total revenue vs. \$170,436 from non-subscribers.

Repeat Buyer Insight

Among customers with 5+ purchases, 958 are subscribers while 2,518 are not—indicating significant untapped subscription potential.

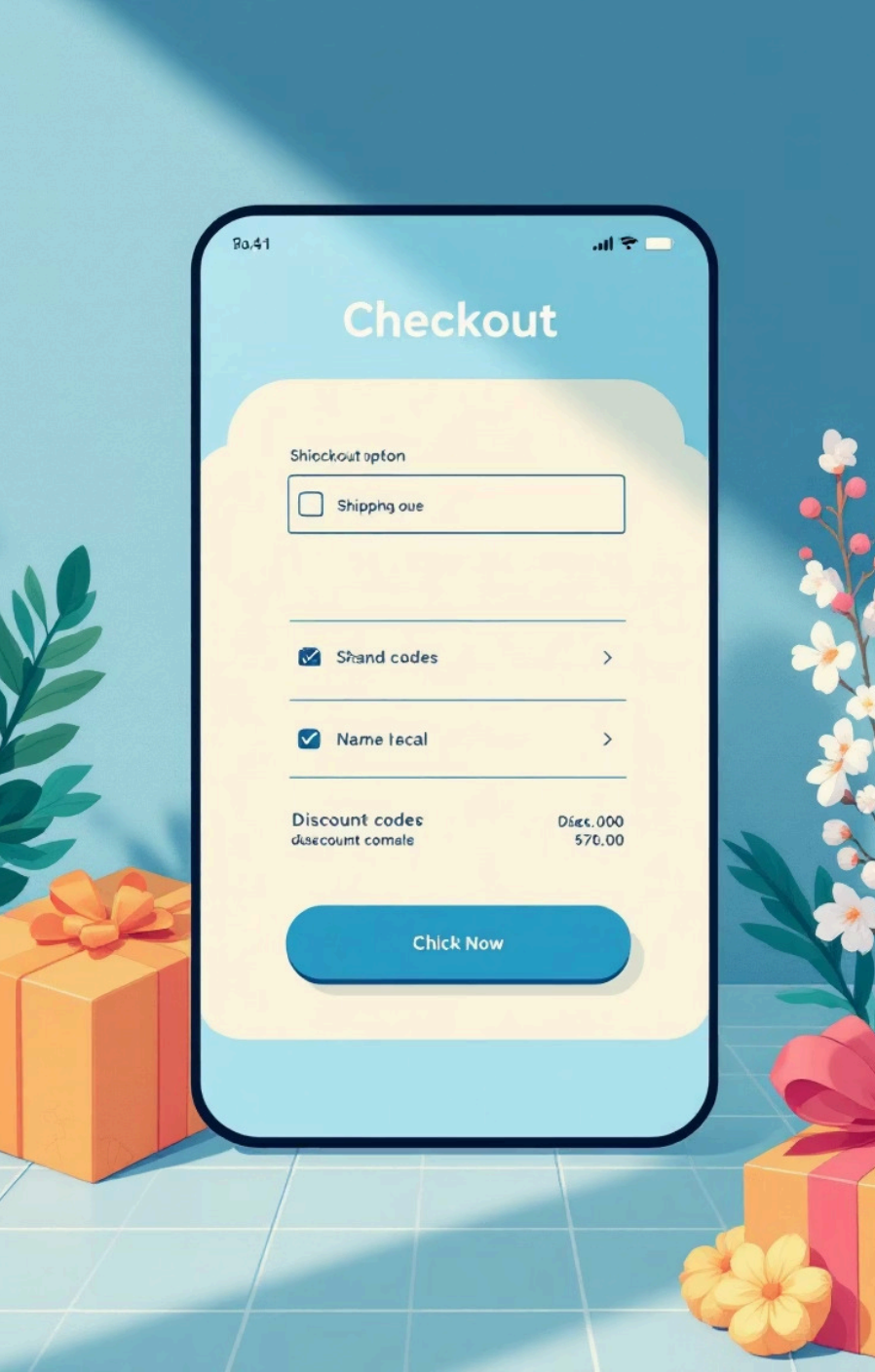
Revenue Distribution Analysis



Category Performance

Clothing dominates with 45% of total revenue. Young adults are the highest-spending demographic across all categories.

Shipping & Discount Behavior



Express Shipping

Average purchase: **\$60.48**

Premium shipping correlates with slightly higher spending, suggesting price-insensitive customer segment.



Standard Shipping

Average purchase: **\$58.46**

Most popular option with cost-conscious customers maintaining healthy purchase amounts.



High-Value Discount Users

Multiple customers used discounts but still spent above the \$59.76 average, indicating discounts don't necessarily reduce basket size.

Strategic Recommendations

Boost Subscriptions

With 2,518 repeat buyers not subscribed, promote exclusive benefits to convert loyal customers into subscribers.

Customer Loyalty Programs

Reward repeat buyers to accelerate movement from "Returning" to "Loyal" segment and increase lifetime value.

Review Discount Policy

Balance sales boosts with margin control—high-value customers spend above average even with discounts applied.

Product Positioning

Highlight top-rated products (Gloves, Sandals, Boots) and best-sellers (Blouse, Jewelry) in marketing campaigns.

Targeted Marketing

Focus efforts on young adults and middle-aged customers, plus express-shipping users who demonstrate higher spending patterns.