

Customer Shopping Behavior Analysis

Uncovering insights from 3,900 purchases to guide strategic business decisions through data-driven analysis of spending patterns, customer segments, and product preferences.

Dataset Overview

3,900

Total Purchases

Transactions analyzed across
all categories

18

Data Columns

Comprehensive feature set

50

Locations

Geographic diversity

Key Features Captured

Our dataset encompasses customer demographics including age, gender, location, and subscription status. Purchase details cover item categories, amounts, seasonal trends, size, and color preferences. Shopping behavior metrics include discount usage, promo codes, previous purchases, frequency patterns, review ratings, and shipping preferences.



Data Preparation & Cleaning Process

01

Data Loading & Exploration

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

02

Missing Data Handling

Identified 37 missing values in Review Rating column, imputed using median rating by product category

03

Feature Engineering

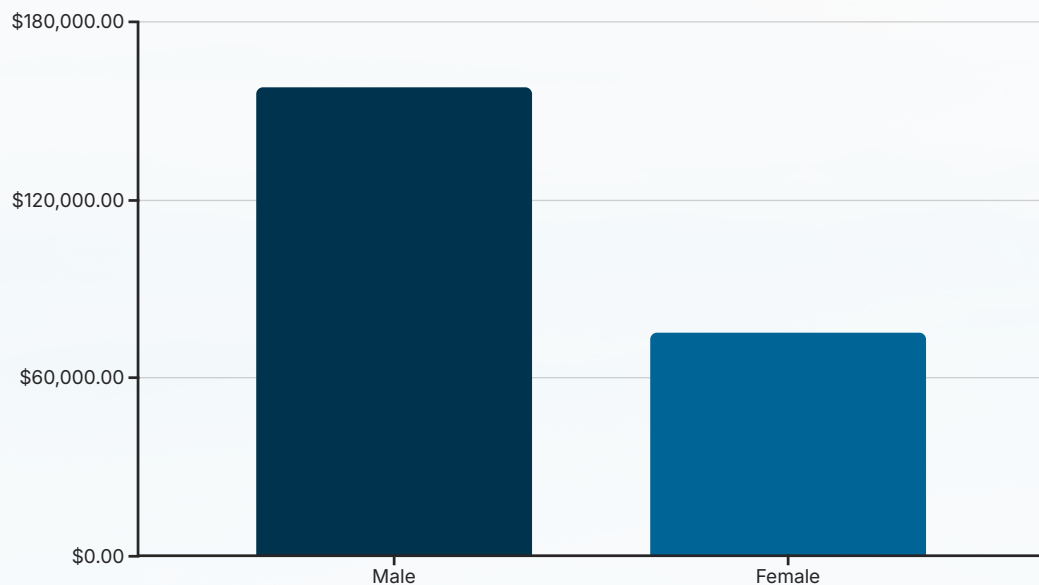
Created `age_group` column through binning, generated `purchase_frequency_days` from purchase data, standardized columns to snake_case

04

Database Integration

Connected Python to PostgreSQL, loaded cleaned DataFrame for SQL analysis

Gender Revenue Analysis



Male Demographic Drives Revenue

Our analysis reveals that male customers generate significantly higher revenue, contributing \$157,890 compared to \$75,191 from female customers. This represents a 2.1x difference in total revenue generation.

This insight suggests opportunities for targeted marketing strategies and product positioning to further capitalize on male customer preferences while exploring growth potential in the female demographic.

Product Performance & Customer Preferences

Top-Rated Products



Gloves

Highest customer satisfaction



Sandals

Strong summer performance



Boots

Consistent quality ratings



Hat

Popular accessory choice



Products with highest average review ratings demonstrate strong customer satisfaction. Gloves lead with 3.86 stars, followed closely by Sandals at 3.84. These top performers should be featured prominently in marketing campaigns.

Shipping & Discount Behavior Insights



Express Shipping Premium

Express shipping customers spend \$60.48 on average, compared to \$58.46 for standard shipping, indicating willingness to pay more for faster delivery.



High-Value Discount Users

839 customers used discounts but still spent above average purchase amount, demonstrating that discounts attract quality buyers, not just bargain hunters.



Discount-Dependent Products

Hats (50%), Sneakers (49.66%), and Coats (49.07%) show highest discount rates, suggesting these categories may need pricing strategy review.

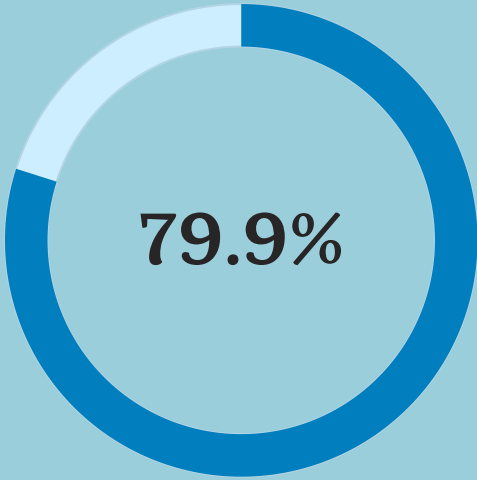
Customer Loyalty Breakdown

Three Distinct Segments

Our customer base divides into three clear segments based on purchase history. Loyal customers dominate at 3,116 individuals (79.9%), representing our most valuable segment with consistent repeat purchases.

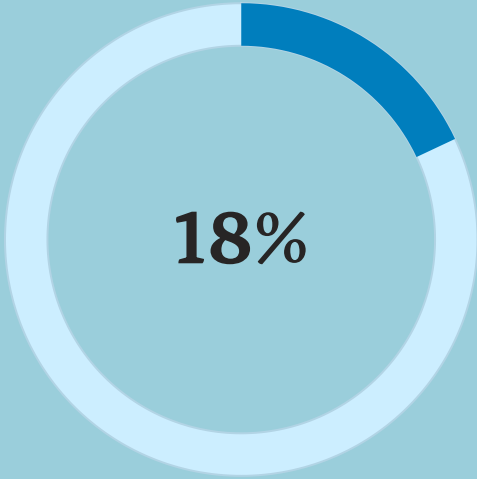
Returning customers number 701 (18%), showing positive engagement but not yet reaching loyal status. New customers comprise just 83 individuals (2.1%), indicating strong retention once customers make their first purchase.

This distribution suggests our business excels at converting new customers into loyal advocates, though new customer acquisition may need attention.



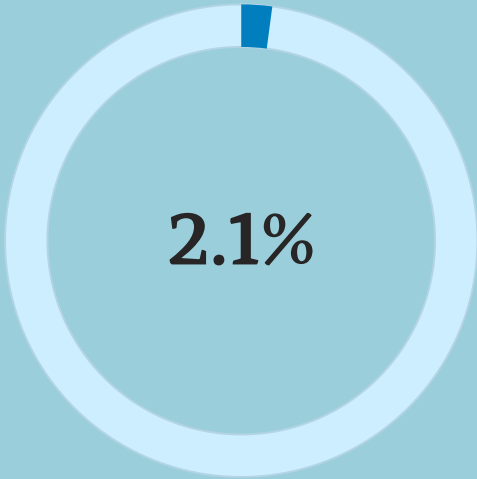
Loyal

3,116 customers



Returning

701 customers



New

83 customers

Subscription Status & Revenue Impact

Subscription Metrics

27% Subscribe

1,053 customers with active subscriptions

\$59.49 Avg

Average spend per subscriber

\$62,645

Total subscriber revenue

Non-Subscriber Metrics

73% No Sub

2,847 customers without subscriptions

\$59.87 Avg

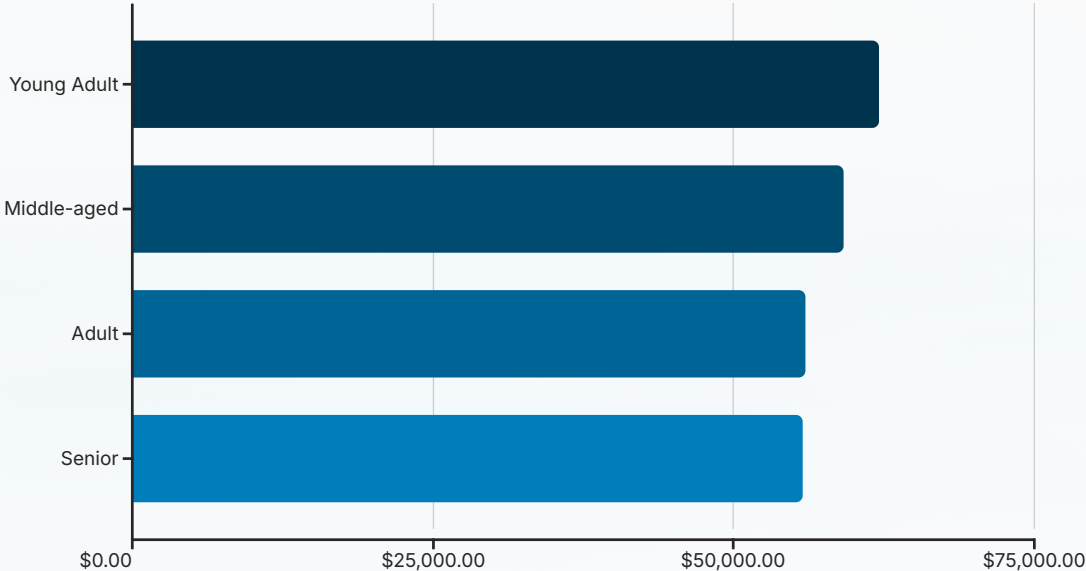
Average spend per non-subscriber

\$170,436

Total non-subscriber revenue

Repeat buyers with over 5 purchases show 958 subscribers versus 2,518 non-subscribers, indicating subscription opportunity among loyal customers.

Revenue Distribution Across Demographics



Age Group Performance

Young Adults lead revenue generation at \$62,143, followed closely by Middle-aged customers at \$59,197. The distribution is relatively balanced across age groups, with Adults contributing \$55,978 and Seniors \$55,763.

This even distribution suggests broad market appeal, though Young Adults represent the highest-value segment for targeted marketing initiatives and product development focus.

Business Recommendations

Boost Subscription Rates

Promote exclusive benefits and tempting offers to increase the 27% subscription rate. Target loyal customers with over 5 purchases who haven't yet subscribed.

Enhance Loyalty Programs

Reward repeat buyers to accelerate their transition into the Loyal segment. Focus on the 701 Returning customers to increase retention and lifetime value.

Optimize Discount Strategy

Review discount policy for products with 50% discount rates. Balance sales boosts with margin control, especially for Hats, Sneakers, and Coats.

Strategic Product Positioning

Highlight top-rated products (Gloves, Sandals, Boots) in marketing campaigns. Leverage high review ratings to drive conversion and build trust.

Targeted Marketing Focus

Concentrate efforts on high-revenue demographics: Young Adults and male customers. Develop campaigns for express-shipping users who demonstrate higher spending patterns.