Customer Shopping Behavior Analysis

Analyzing 3,900 purchases to uncover insights into spending patterns, customer segments, and product preferences that drive strategic business decisions.

Dataset Overview

3,900
Total Purchases

Transactions analyzed across all categories

18
Data Columns

Features covering demographics and behavior

50
Locations

Geographic distribution of customers

25
Products

Unique items purchased

Key features include customer demographics, purchase details, shopping behavior patterns, and review ratings. Only 37 missing values found in Review Rating column.

Data Preparation Process

01

Data Loading & Exploration

Imported dataset using pandas and analyzed structure with df.info() and summary statistics

02

Missing Data Handling

Imputed 37 missing Review Rating values using median rating by product category

03

Feature Engineering

Created age_group bins and purchase_frequency_days columns for deeper analysis

04

Database Integration

Connected to PostgreSQL and loaded cleaned data for SQL business analysis

Revenue by Gender

Male   Female

Male Customers Drive Revenue

Male customers generated $157,890 in total revenue compared to $75,191 from female customers—representing 68% of total sales.

This significant gap suggests opportunities for targeted marketing to female demographics.

Customer Spending Insights

High-Value Discount Users

839 customers used discounts but still spent above average ($59.76), indicating price-conscious yet high-spending segments.

Shipping Preferences

Express shipping customers spend $60.48 on average vs. $58.46 for Standard—a 3.5% premium for faster delivery.

Subscription Impact

Subscribers and non-subscribers show similar average spend ($59.49 vs. $59.87), but non-subscribers drive 73% of total revenue.

Made with GAMMA

Top-Rated Products

77%

Gloves

Highest average rating at 3.86 out of 5.0

77%

Sandals

Strong performance with 3.84 rating

76%

Boots

Solid customer satisfaction at 3.82

76%
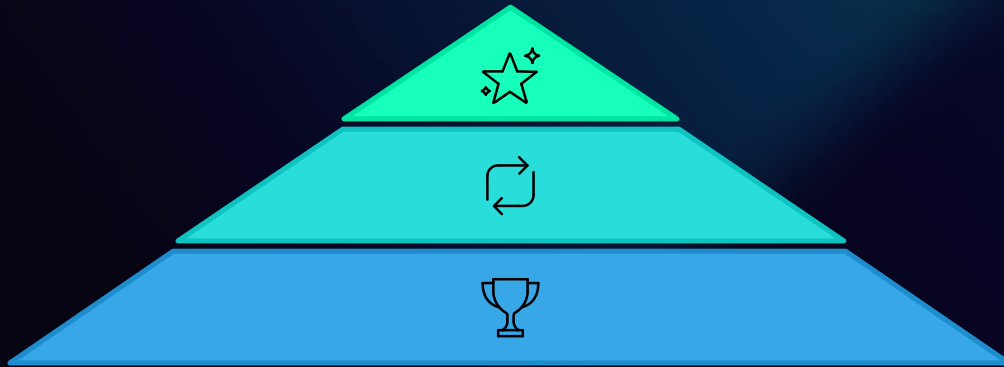
Hat

Consistent quality with 3.80 rating

76%

Skirt

Rounds out top 5 at 3.78 rating

These products demonstrate consistent quality and customer satisfaction, making them ideal for promotional campaigns and inventory prioritization.

Made with GAMMA

Customer Segmentation Analysis

Loyalty Dominates

80% of customers are classified as Loyal based on purchase history, indicating strong retention.

Among repeat buyers (5+ purchases), 958 are subscribers while 2,518 are not—suggesting untapped subscription potential.

New

83 customers

Returning

701 customers

Loyal

3,116 customers

Discount-Dependent Products

1

Hat

50% of purchases used discounts

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 rely heavily on discounts to drive sales. Review pricing strategy to balance volume with profit margins.

Revenue by Age Group

Young Adults lead revenue generation at $62,143, followed closely by Middle-aged customers. Revenue distribution is relatively balanced across age groups, suggesting broad market appeal.

Strategic Recommendations

Boost Subscriptions

Promote exclusive benefits to convert the 2,518 repeat buyers who aren't yet subscribers.

Loyalty Programs

Reward repeat buyers to strengthen the 80% loyal customer base and increase lifetime value.

Review Discount Policy

Balance promotional discounts with margin-control, especially for the 5 discount-dependent products.

Targeted Marketing

Focus campaigns on Young Adults, express-shipping users, and top-rated products like Gloves and Sandals.