Customer Shopping Behavior Analysis

Uncovering insights from 3,900 purchases to guide strategic business decisions

Understanding Our Data

This analysis examines transactional data from 3,900 purchases across multiple product categories. Our goal: uncover spending patterns, customer segments, product preferences, and subscription behavior.
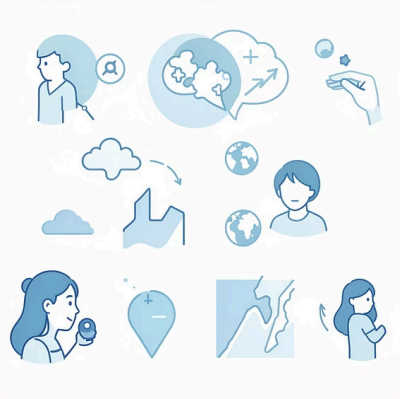
3,900

Total purchases analyzed

18

Data features tracked

4

Product categories

Dataset Overview

Customer Demographics

Age, gender, location, and subscription status

Purchase Details

Item, category, amount, season, size, and color

Shopping Behavior

Discounts, promo codes, purchase history, ratings, and shipping

Missing Data: 37 values in Review Rating column were imputed using median ratings by category

Data Preparation & Cleaning

01
Data Loading & Exploration

Imported dataset with pandas, checked structure and summary statistics

02
Missing Data Handling

Imputed Review Rating nulls using median by product category

03
Feature Engineering

Created age_group bins and purchase_frequency_days columns

04
Data Consistency

Verified redundancy, dropped promo_code_used column

05
Database Integration

Connected to PostgreSQL and loaded cleaned data for SQL analysis

Key Business Questions Answered

1

Revenue by Gender

Male customers generated $157,890 vs. Female $75,191

2

High-Spending Discount Users

839 customers used discounts but spent above average

3

Top Products by Rating

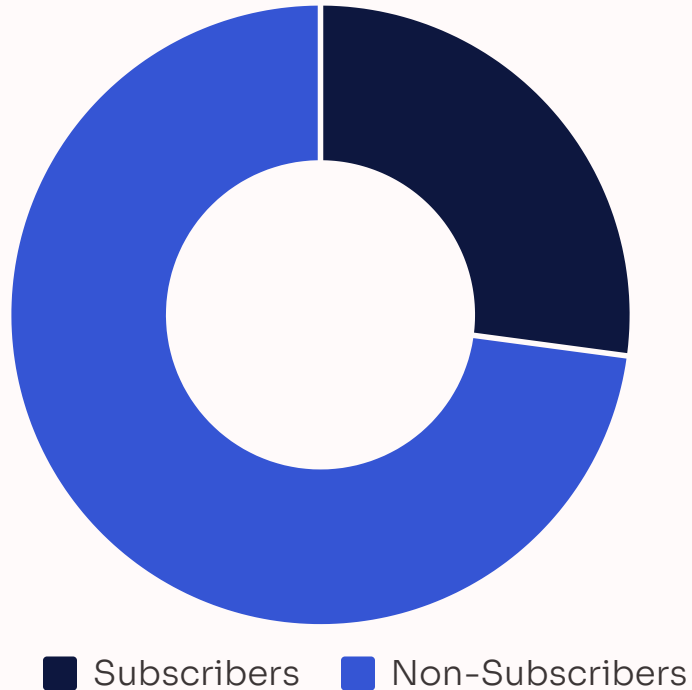Gloves (3.86), Sandals (3.84), Boots (3.82) led ratings

4

Shipping Preferences

Express shipping averaged $60.48 vs. Standard $58.46

Subscription Analysis

- Subscribers
- Non-Subscribers

Subscription Insights

Only 27% of customers are subscribers, yet they contribute significantly to revenue.
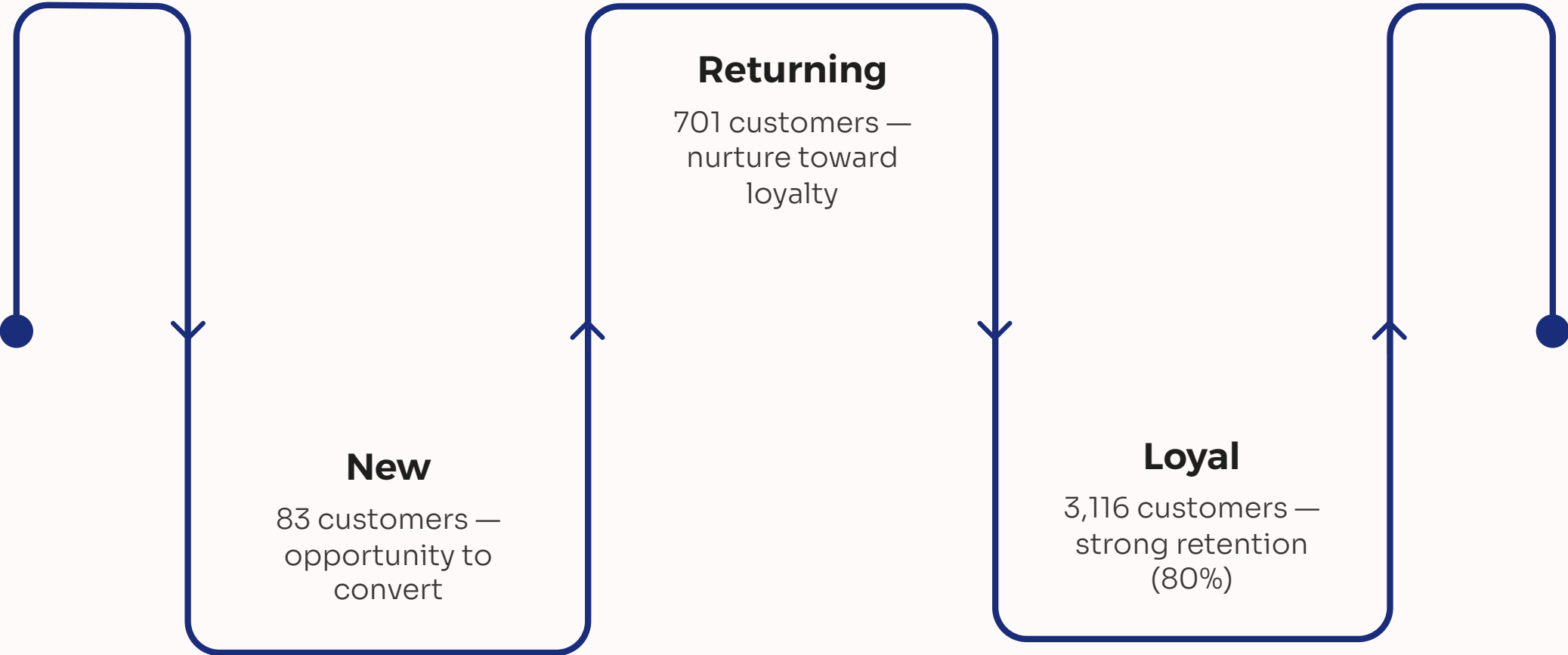
$59.49

Avg Subscriber Spend

$59.87

Avg Non-Subscriber Spend

Repeat buyers (958 subscribers vs. 2,518 non-subscribers) show higher loyalty potential.

Customer Segmentation

Returning

701 customers — nurture toward loyalty

New

83 customers — opportunity to convert

Loyal

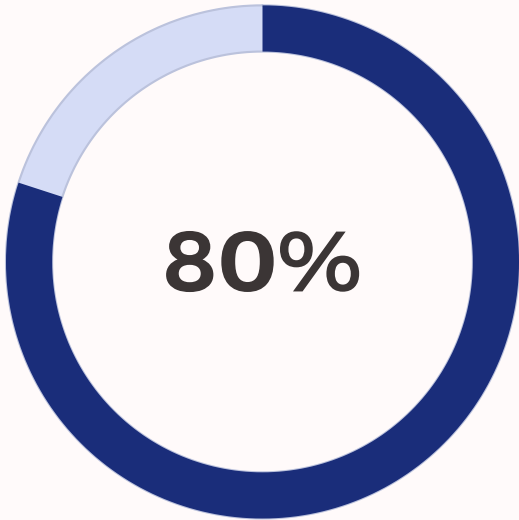3,116 customers — strong retention (80%)

The majority of customers (80%) fall into the Loyal segment, indicating strong retention. Focus on converting New and Returning customers into Loyal advocates.

2%

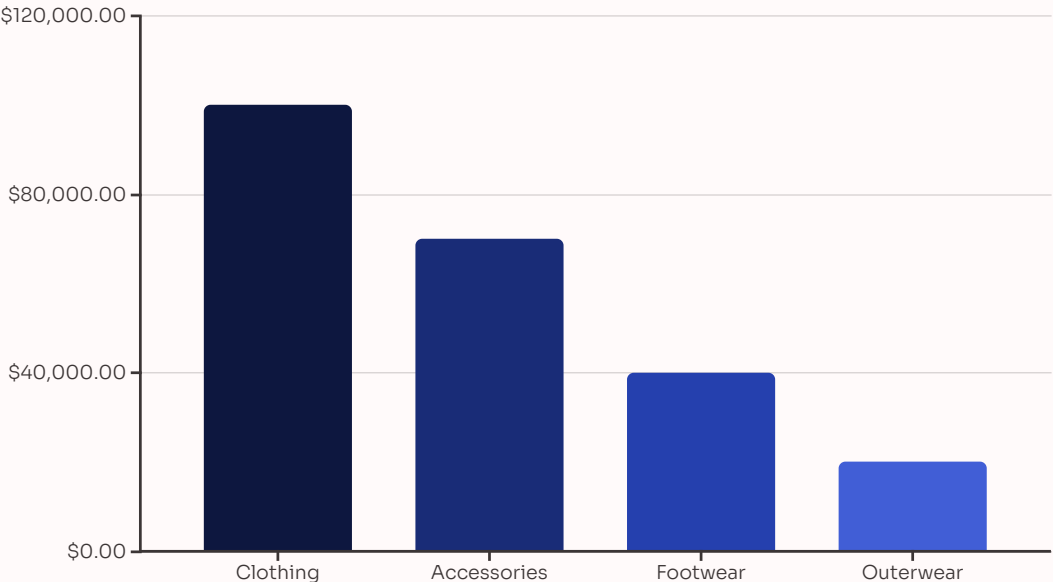**New Customers**

18%

Returning Customers

80%
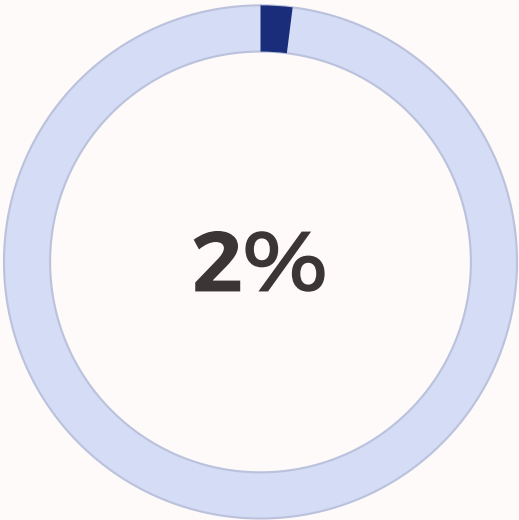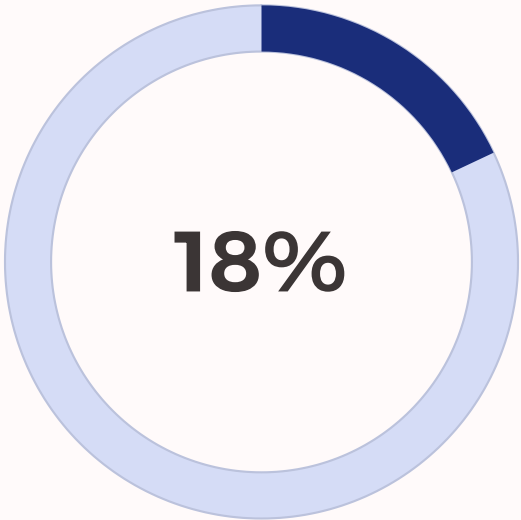
Loyal Customers

Revenue & Sales Performance

Category Performance

Clothing dominates with 1,737 sales generating the highest revenue. Accessories follow as a strong secondary category.

Top Item: Blouse

171 purchases in Clothing

Discount Impact

Hat: 50% discount rate

Age Group Revenue Distribution

Young Adult — 27%

$62,143 total revenue

Middle-aged — 25%

$59,197 total revenue

Adult — 24%

$55,978 total revenue

Senior — 24%

$55,763 total revenue

Key Takeaways & Next Steps

Boost Subscriptions

Promote exclusive benefits to convert the 73% non-subscriber base

Loyalty Programs

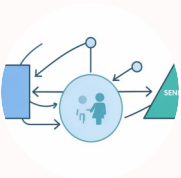Reward repeat buyers to accelerate movement into Loyal segment

Review Discount Policy

Balance sales boosts with margin control—839 high spenders used discounts

Product Positioning

Highlight top-rated items (Gloves 3.86, Sandals 3.84) in marketing campaigns

Targeted Marketing

Focus on Young Adults ($62K revenue) and express-shipping users