

Customer Shopping Behavior Analysis

Analyzing 3,900 transactions across product categories to uncover spending patterns, customer segments, and strategic insights for data-driven business decisions.





Dataset Overview

3,900

Total
Transactions

Purchase records
analyzed

18

Data Features

Columns tracking
behavior

37

Missing Values

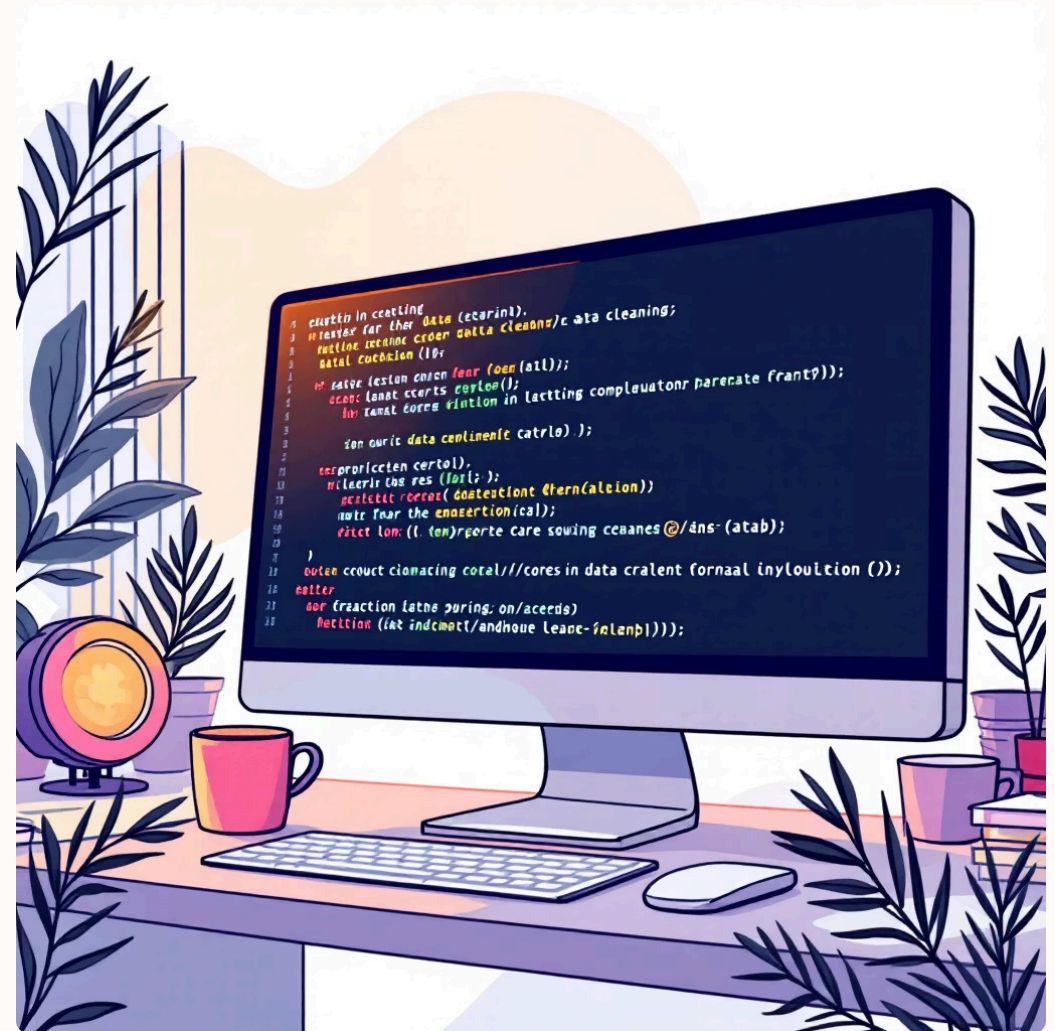
In Review Rating
column

Key features include customer demographics (age, gender, location, subscription status), purchase details (items, category, amount, season), and shopping behavior (discounts, promo codes, purchase frequency, ratings, shipping type).

Data Preparation & Cleaning

Python Analysis Steps

- Loaded dataset using pandas
- Explored structure with `df.info()` and `df.describe()`
- Imputed missing Review Ratings using median by category
- Standardized columns to snake_case
- Created `age_group` and `purchase_frequency_days` features
- Dropped redundant `promo_code_used` column



Connected Python to MySQL and loaded cleaned data for structured business analysis.

Revenue Insights

Gender Revenue Comparison

Male and female customers show comparable spending patterns across all product categories.

High-Spending Discount Users

43 customers used discounts but still spent above average, ranging from \$60-\$100 per purchase.



Product Performance

1

Gloves

3.86 average rating

2

Sandals

3.84 average rating

3

Boots

3.82 average rating

4

Hat

3.80 average rating

5

Skirt

3.78 average rating

Top-rated products show strong customer satisfaction, with accessories and footwear leading in reviews.

Shipping & Subscription Analysis

Shipping Type Comparison

Express: \$60.48 average purchase

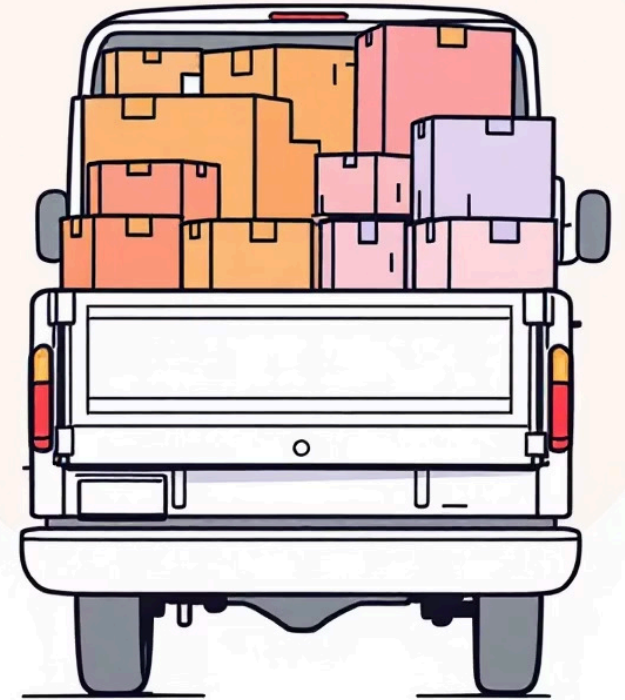
Standard: \$58.46 average purchase

Express shipping customers spend slightly more per transaction.

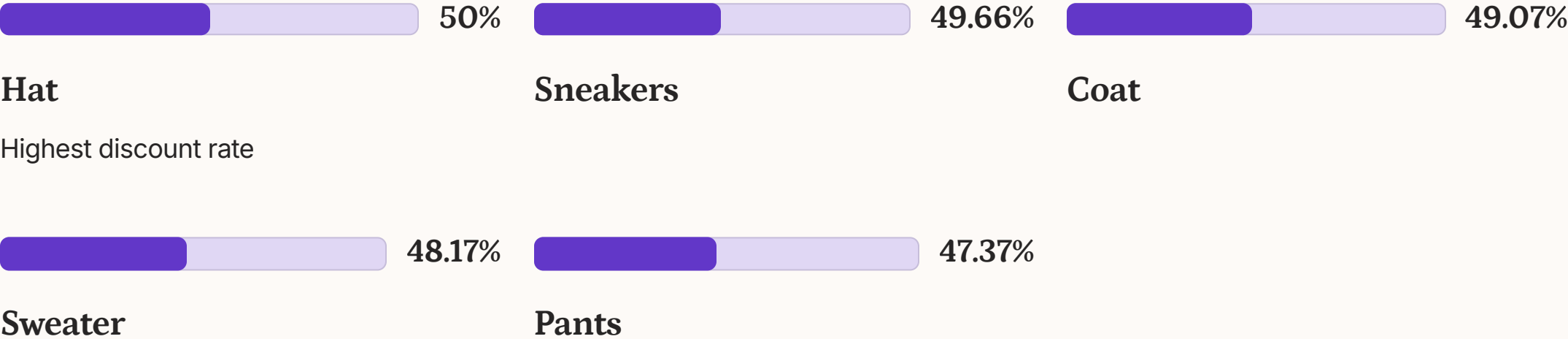
Subscriber vs. Non-Subscribers

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



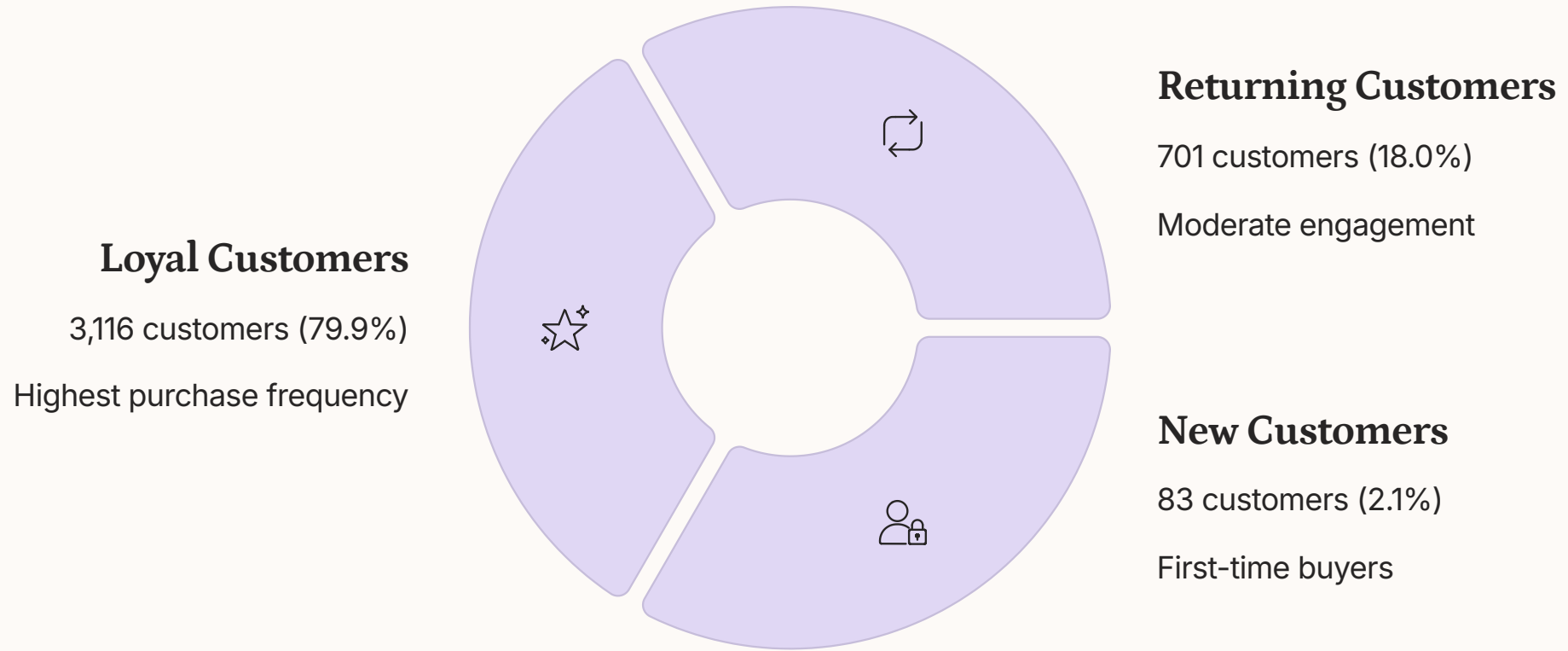
Discount Strategy Insights



Highest discount rate

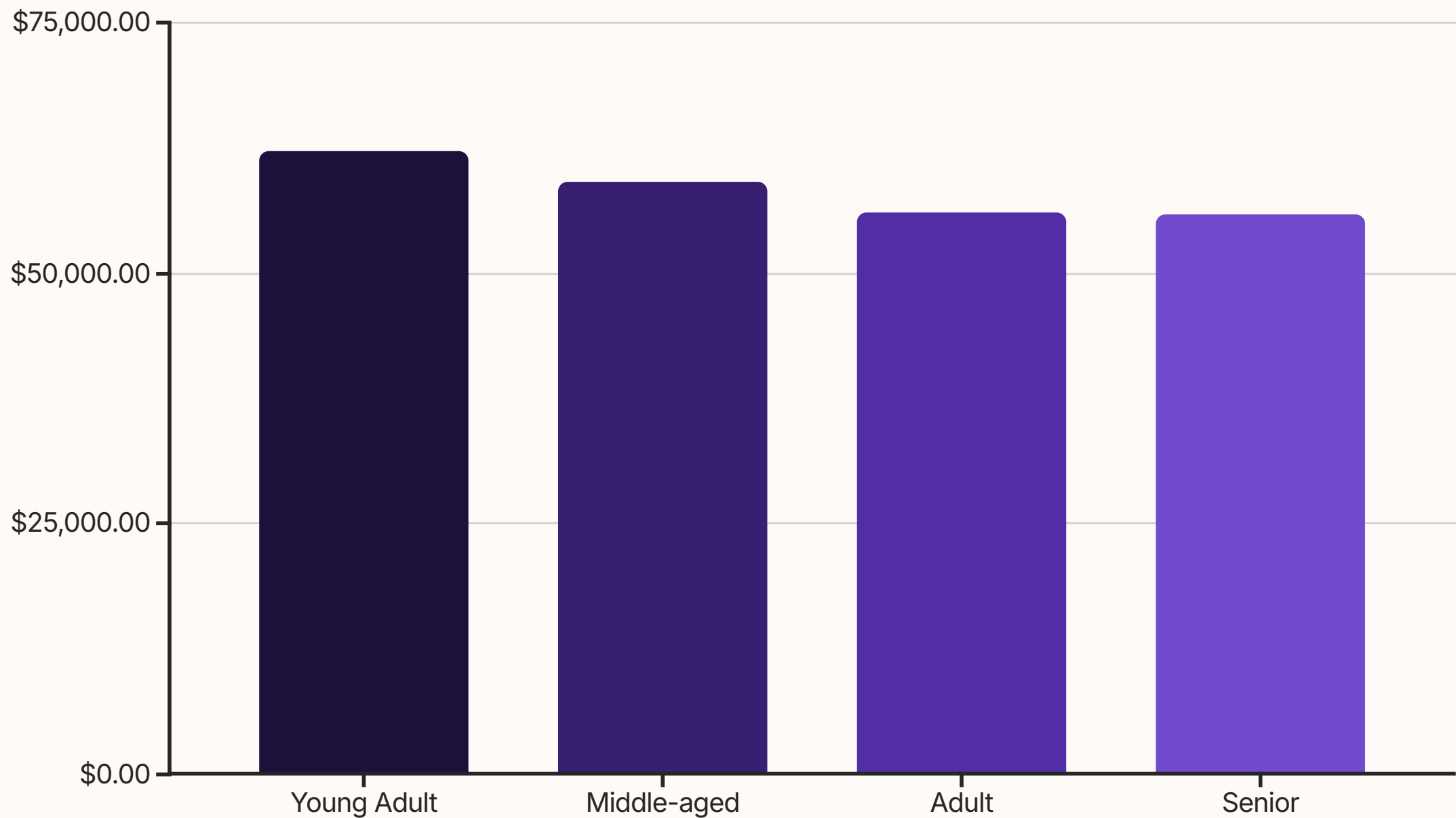
These products show highest percentage of discounted purchases, indicating discount-dependent demand.

Customer Segmentation



Strong loyal customer base presents opportunity to convert returning customers into loyal segment. Only 958 of 3,476 repeat buyers (>5 purchases) are subscribers.

Revenue by Age Group



Young adults generate highest revenue (\$62,143), followed closely by middle-aged customers (\$59,197). Revenue distribution is relatively balanced across age groups.

Strategic Recommendations



Boost Subscriptions

Promote exclusive benefits to convert repeat buyers into subscribers



Loyalty Programs

Reward returning customers to move them into loyal segment



Review Discount Policy

Balance sales boosts with margin control for discount-dependent products



Product Positioning

Highlight top-rated and best-selling items in marketing campaigns



Targeted Marketing

Focus on high-revenue age groups and express-shipping users