

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.



The Foundation: Our Data

3,900

Total Purchases

Transactions analyzed

18

Data Points

Features per customer

50

Locations

Geographic coverage

Key Features Captured

- Customer demographics (Age, Gender, Location, Subscription Status)
- Purchase details (Item, Category, Amount, Season, Size, Color)
- Shopping behavior (Discounts, Promo Codes, Previous Purchases, Review Ratings, Shipping Type)

```
4 # Python script file
5 # Create class for Data Cleaning
6
7 class DataCleaning:
8     def __init__(self):
9         self.data = None
10
11     def load_data(self, file_path):
12         # Load data from CSV file
13         self.data = pd.read_csv(file_path)
14
15     def handle_na(self):
16         # Handle missing values
17         self.data.fillna(0, inplace=True)
18
19     def remove_outliers(self):
20         # Remove outliers
21         self.data = self.data[(self.data['Age'] < 100) & (self.data['Age'] > 10)]
22
23     def encode_labels(self):
24         # Encode categorical labels
25         self.data['Gender'] = self.data['Gender'].map({'Male': 0, 'Female': 1})
26
27     def scale_features(self):
28         # Scale features
29         self.data['Age'] = self.data['Age'].apply(lambda x: (x - self.data['Age'].min()) / (self.data['Age'].max() - self.data['Age'].min()))
30
31     def save_data(self, file_path):
32         # Save cleaned data
33         self.data.to_csv(file_path, index=False)
```



Data Preparation in Python

01

Data Loading & Exploration

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

02

Missing Data Handling

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

03

Feature Engineering

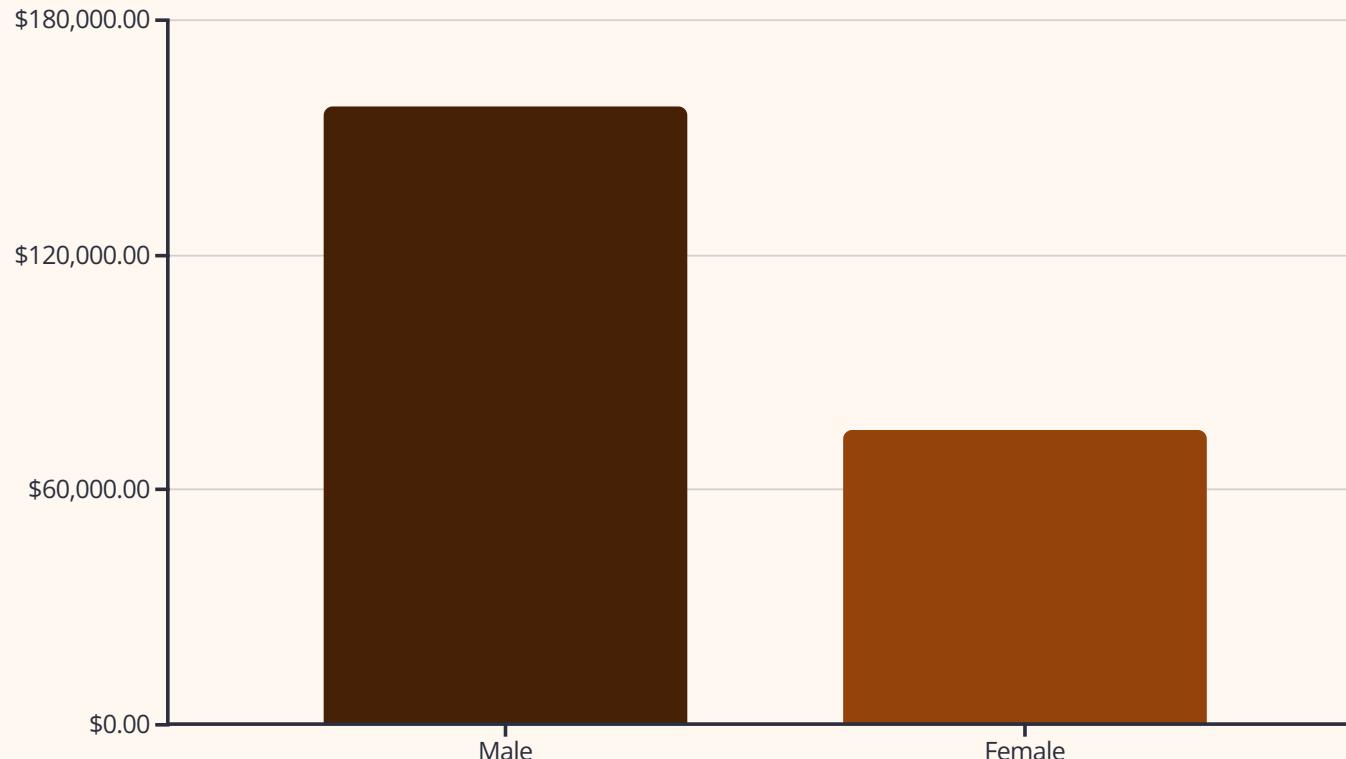
Created age_group bins and purchase_frequency_days columns for deeper analysis

04

Database Integration

Connected to PostgreSQL and loaded cleaned data for SQL analysis

Revenue Insights by Gender



Key Finding

Male customers generate 2.1x more revenue than female customers, representing 68% of total revenue. This significant disparity suggests opportunities for targeted marketing campaigns to better engage female shoppers.

Smart Discount Users



High-Value Discount Shoppers

839 customers used discounts while spending above the \$60 average purchase amount. These savvy shoppers represent a valuable segment that responds to promotions without sacrificing basket size.

Strategic Opportunity

Balance promotional offers with margin control

Customer Value

Discount users maintain high spending levels

Top-Rated Products



Gloves

3.86 average rating



Sandals

3.84 average rating



Boots

3.82 average rating



Hat

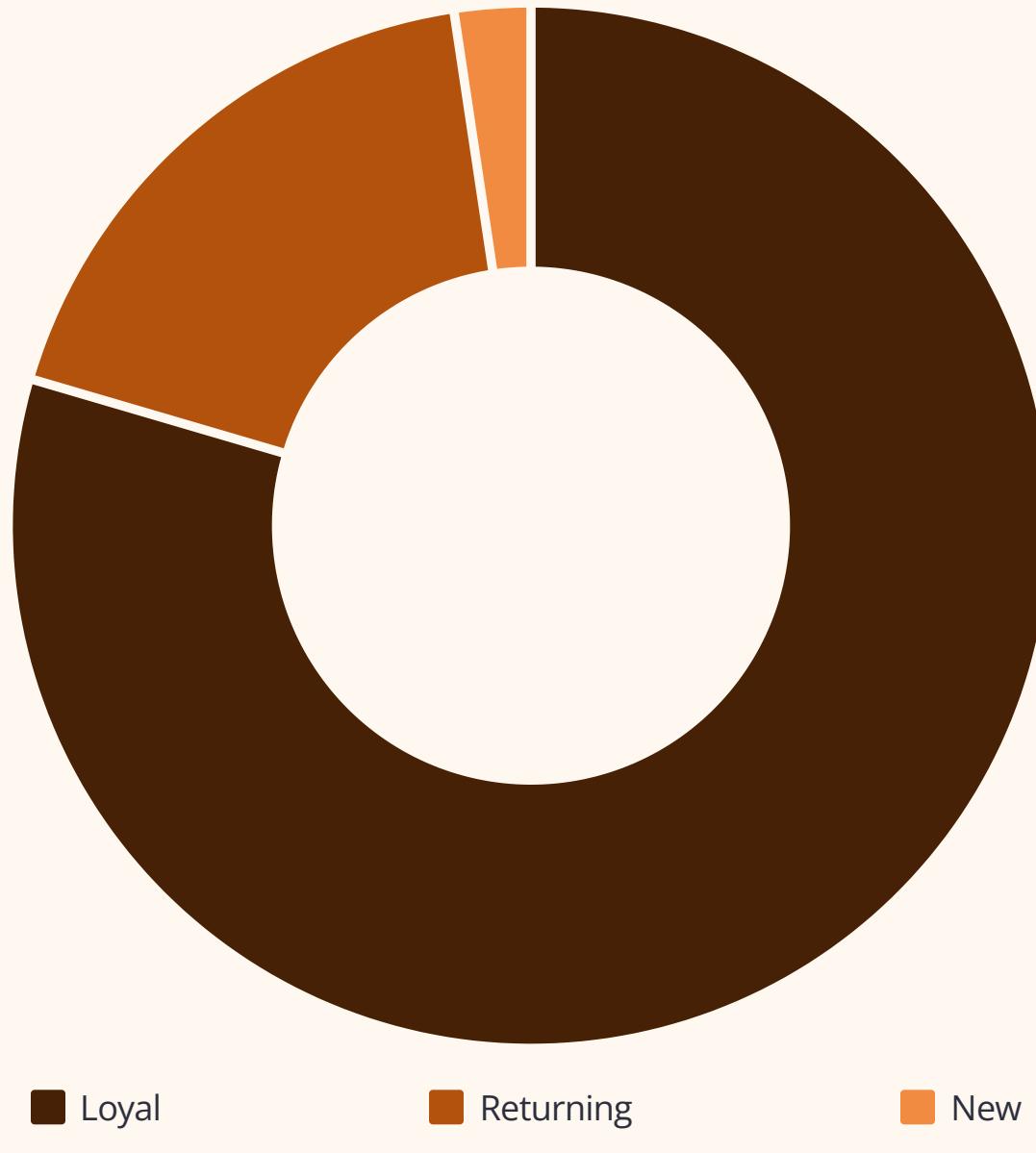
3.80 average rating



Skirt

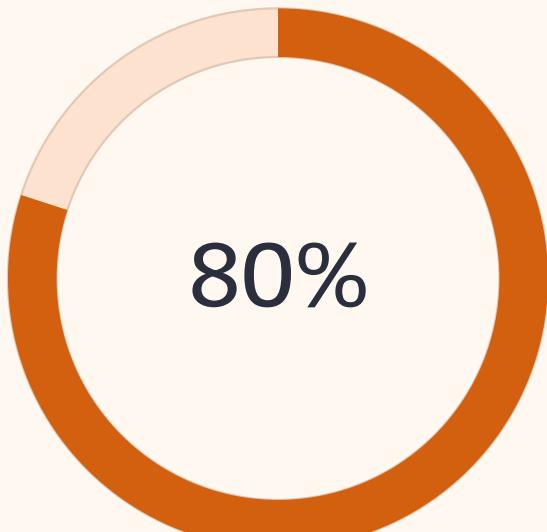
3.78 average rating

Customer Segmentation Analysis

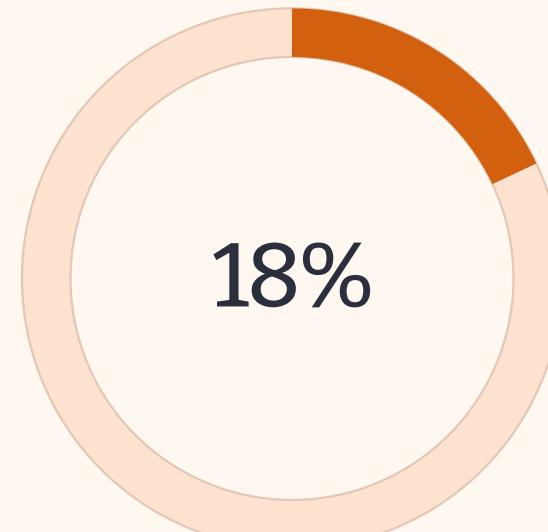


Segment Breakdown

80% of customers are classified as Loyal (3,116), with 18% Returning (701) and just 2% New (83). This strong loyalty base presents opportunities for retention programs and upselling.



Loyal customers



Returning shoppers



POWER BI DASHBOARD

Interactive Visual Insights

3.9K Customers

Total analyzed

\$59.76 Average

Purchase amount

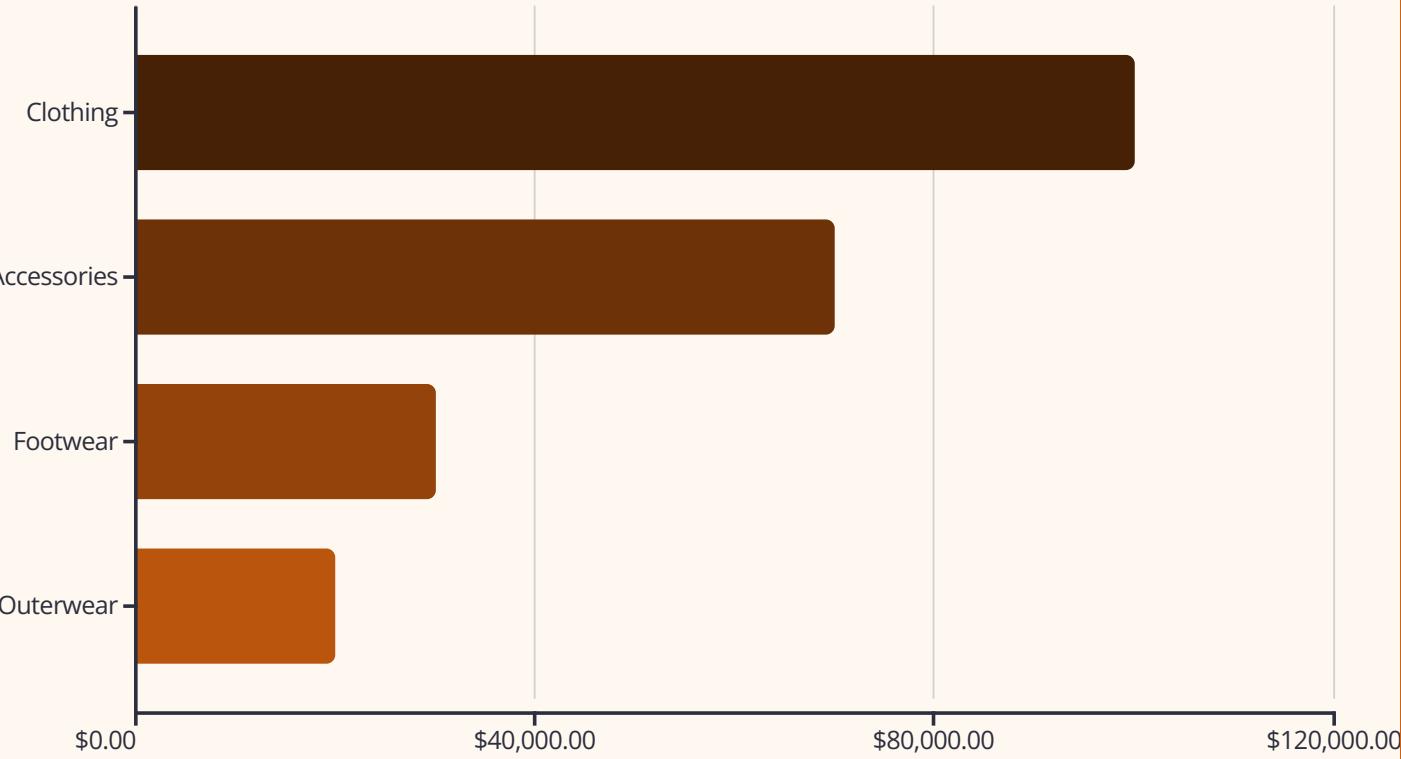
3.75 Rating

Average review score

27% Subscribe

Subscription rate

Revenue Distribution Insights



Category Performance

Clothing dominates with \$100K revenue (45% of total), followed by Accessories at \$70K. Young Adults lead age groups with \$62K revenue contribution, making them a prime target for marketing campaigns.

Strategic Recommendations

1

Boost Subscriptions

Promote exclusive benefits to increase the 27% subscription rate among loyal customers

2

Loyalty Programs

Reward repeat buyers to move Returning customers into the Loyal segment

3

Review Discount Policy

Balance promotional sales boosts with margin control for sustainable growth

4

Product Positioning

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

5

Targeted Marketing

Focus on high-revenue Young Adults and express-shipping users for maximum ROI