Customer Shopping Behavior Analysis

Analyzing 3,900 customer transactions 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 product categories

18
Data Columns
Comprehensive customer and purchase attributes

50
Locations
Geographic diversity in customer base

25
Products
Unique items across four major categories

The dataset captures customer demographics including age, gender, and location, alongside detailed purchase information such as item category, amount, season, size, and color. Shopping behavior metrics include discount usage, promo codes, previous purchases, frequency, review ratings, and shipping preferences.

Data Preparation & Python Analysis

01

Data Loading & Exploration

Imported dataset using pandas, performed initial exploration with df.info() and .describe() for summary statistics

02

Missing Data Handling

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

03

Feature Engineering

Created age_group column by binning customer ages and purchase_frequency_days from purchase data
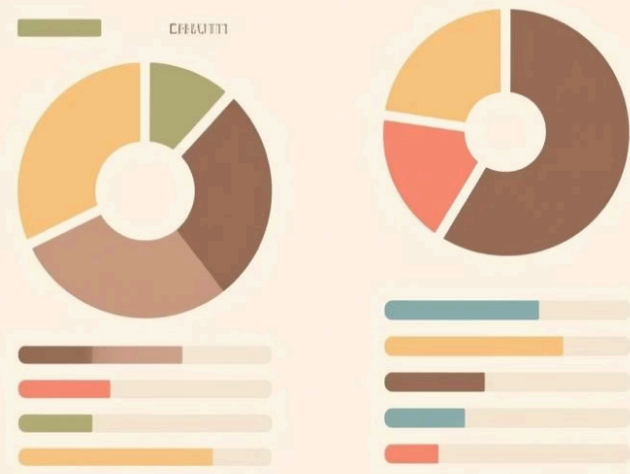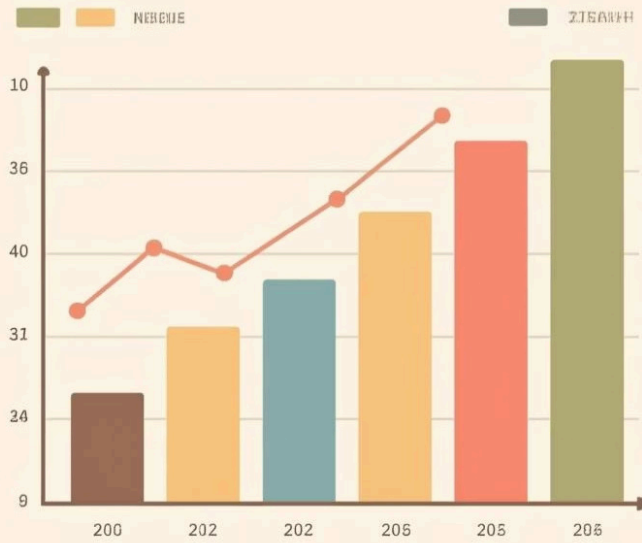
04

Data Standardization

Renamed columns to snake_case, verified discount_applied and promo_code_used redundancy, dropped duplicates

05

Database Integration

Connected Python to PostgreSQL and loaded cleaned DataFrame for SQL analysis

Key Statistical Insights

Customer Demographics

- Average age: **44 years** (range: 18-70)
- Gender split: **68% Male**, 32% Female
- Most common size: **Medium** (45%)
- Top location: Montana (96 customers)

Purchase Behavior

- Average purchase: **$59.76**
- Price range: $20 - $100
- Average previous purchases: **25.4**
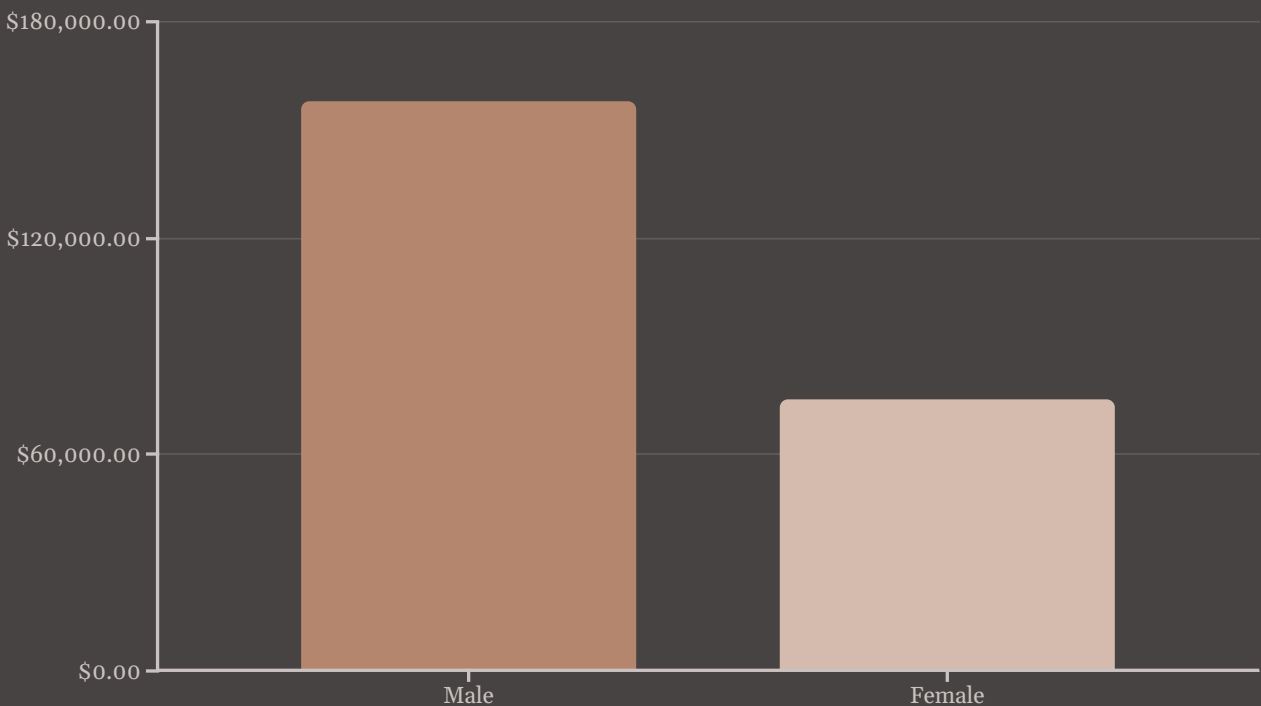- Average review rating: **3.75/5.0**

Shopping Patterns

- Subscription rate: **27%** (1,053 subscribers)
- Discount usage: **43%** of transactions
- Most popular season: **Spring** (999 purchases)
- Top category: **Clothing** (1,737 items)

Shipping Preferences

- Free shipping: **17.3%**
- Most common: Every 3 Months frequency
- Top payment: **PayPal** (677 transactions)

Revenue Analysis by Gender

Gender Revenue Gap

Male customers generate **$157,890** in total revenue, more than double the **$75,191** from female customers. This significant disparity represents a 68% vs 32% split in total revenue contribution.

Despite representing 68% of the customer base, male shoppers demonstrate proportional spending patterns, suggesting consistent engagement across both segments.

Product Performance Insights

Top Rated: Gloves

Highest average rating at **3.86/5.0**, indicating strong customer satisfaction

Runner-Up: Sandals

Second highest rating at **3.84/5.0**, popular seasonal choice

Strong Performer: Boots

Rated **3.82/5.0**, consistent quality across seasons

Most Purchased: Blouse

171 orders in Clothing category

Accessories Leader: Jewelry

171 orders, tied for highest volume

## Footwear Favorite: Sandals

160 orders lead the category

Made with GAMMA

Customer Segmentation Analysis

Returning Customers

701 customers (18.0%)

Growing engagement, potential for loyalty program conversion

Loyal Customers

3,116 customers (79.9%)

High-value segment with consistent purchase history and strong brand affinity

New Customers

83 customers (2.1%)

Fresh acquisition opportunities requiring onboarding focus

The overwhelming majority of customers fall into the Loyal segment, demonstrating strong retention. However, the small percentage of New customers (2.1%) suggests opportunities to enhance acquisition strategies and convert Returning customers into the Loyal tier.
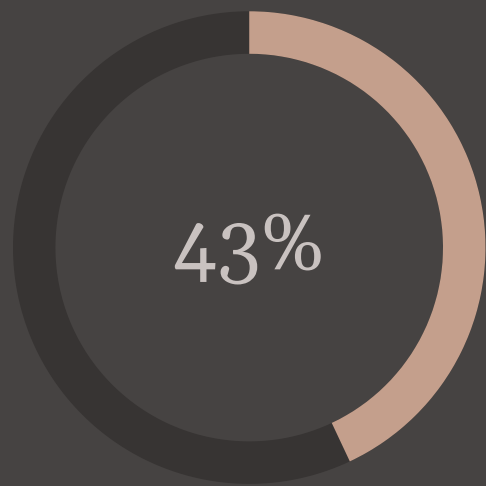
Subscription & Discount Behavior

Subscription Impact

27%
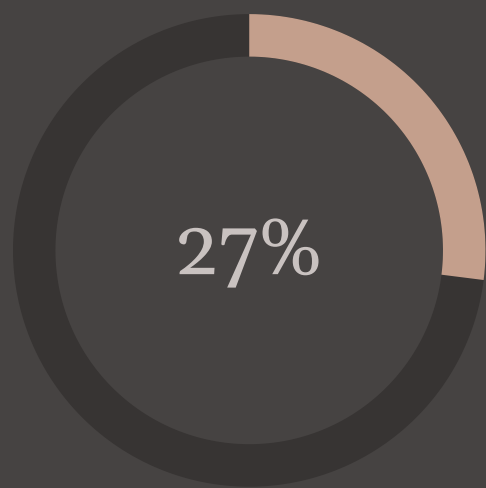
Subscription Rate

1,053 of 3,900 customers

43%

Discount Usage

Applied across purchases

Subscribers spend an average of **$59.49** compared to **$59.87** for non-subscribers, showing minimal difference. Total revenue: subscribers contribute $62,645 while non-subscribers generate $170,436.

Among repeat buyers (>5 purchases), **958 are subscribers** vs 2,518 non-subscribers, suggesting subscription benefits may not be compelling enough.

High-Value Discount Users

839 customers used discounts but still spent above the $59.76 average, indicating price-conscious shoppers who maintain high purchase values.

Top Discount-Dependent Products

1. **Hat** - 50.0% discount rate
2. **Sneakers** - 49.66% discount rate
3. **Coat** - 49.07% discount rate
4. **Sweater** - 48.17% discount rate
5. **Pants** - 47.37% discount rate

Shipping & Age Demographics

Express Shipping Premium

Average purchase: **$60.48**

Customers choosing Express shipping spend $2.02 more on average than Standard shipping users ($58.46), indicating willingness to pay for faster delivery correlates with higher order values.

Young Adult Revenue Leaders

Total revenue: **$62,143**

The Young Adult segment generates the highest revenue, followed by Middle-aged ($59,197), Adult ($55,978), and Senior ($55,763) groups, showing relatively balanced contribution across age demographics.
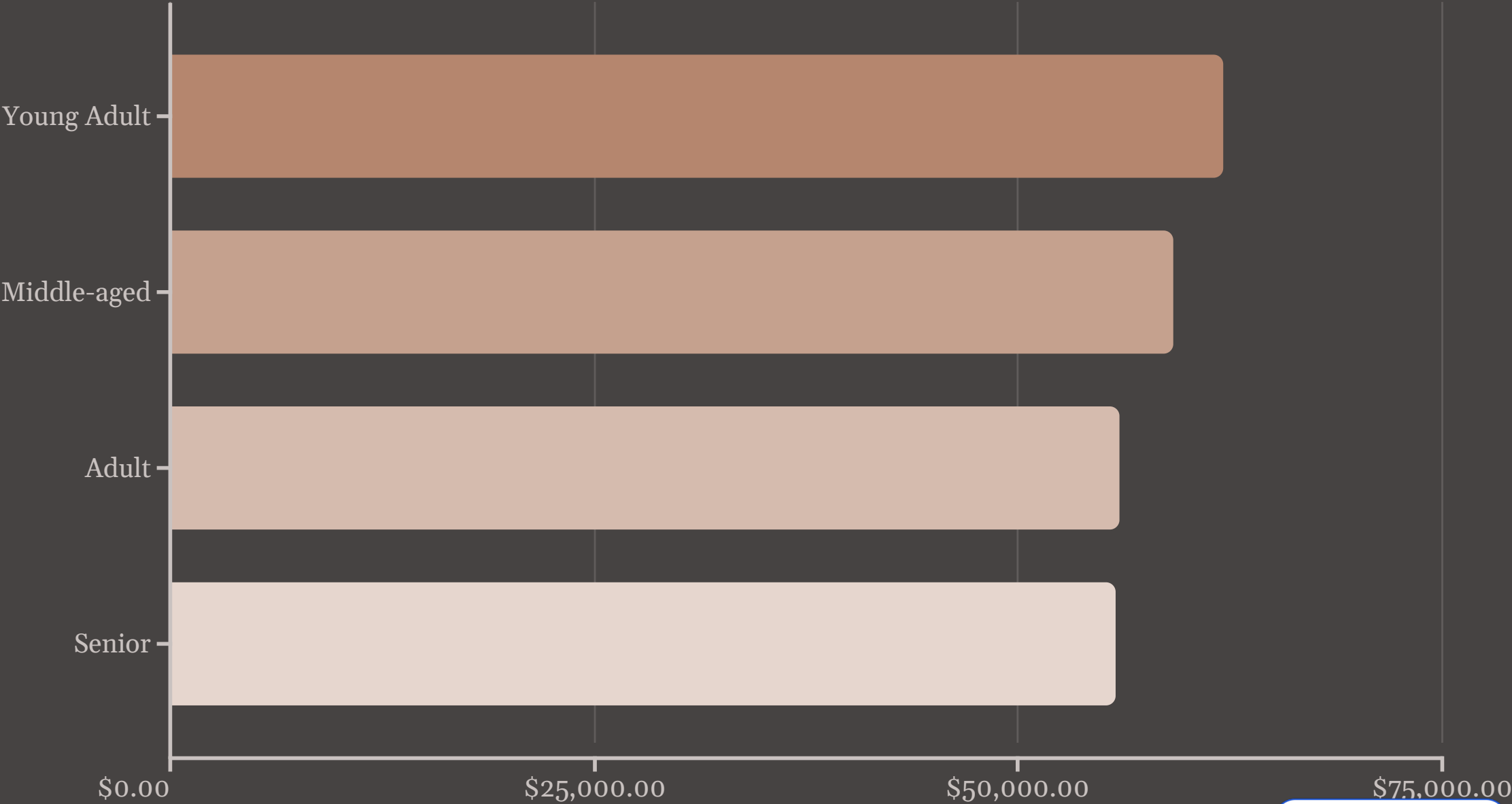
Strategic Business Recommendations

Boost Subscriptions

Promote exclusive benefits and perks for subscribers to increase the 27% subscription rate and convert repeat buyers

Customer Loyalty Programs

Reward repeat buyers to accelerate movement from Returning (18%) to Loyal (80%) segment

Review Discount Policy

Balance sales boosts with margin control, especially for products with 47-50% discount rates

Product Positioning

Highlight top-rated products (Gloves, Sandals, Boots) and best-sellers (Blouse, Jewelry) in marketing campaigns

Targeted Marketing

Focus efforts on high-revenue Young Adult segment and Express shipping users who demonstrate higher purchase values