

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 various product categories

18

Data Columns

Features covering demographics, purchases, and behavior

37

Missing Values

Only in Review Rating column, handled via imputation

Customer Data

- Age, Gender, Location
- Subscription Status
- Previous Purchases

Purchase Details

- Item, Category, Amount
- Season, Size, Color
- Discounts & Shipping

Data Preparation in Python

01

Data Loading & Exploration

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

02

Missing Data Handling

Imputed Review Rating nulls using median rating per product category

03

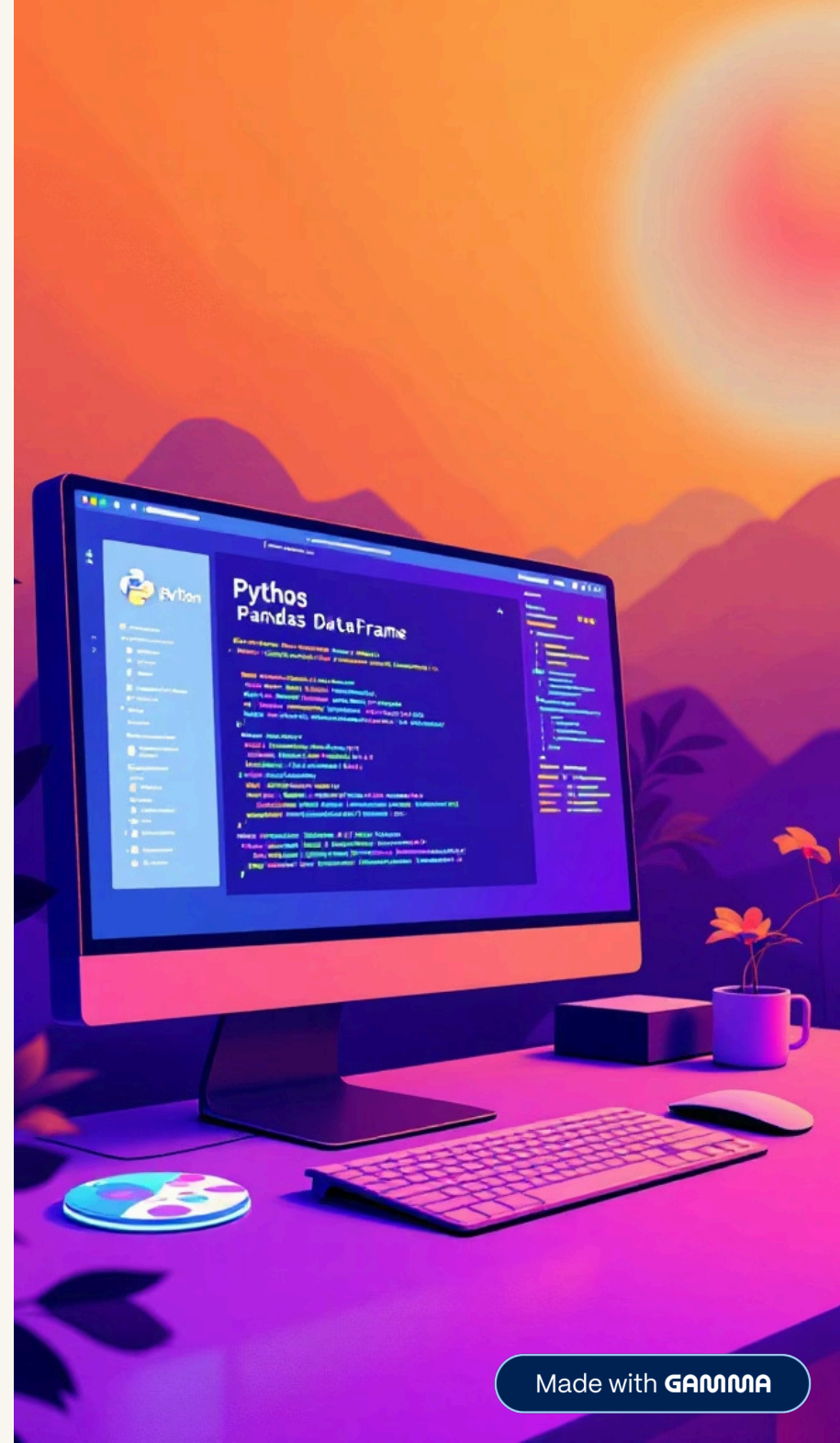
Feature Engineering

Created `age_group` bins and `purchase_frequency_days` columns for deeper analysis

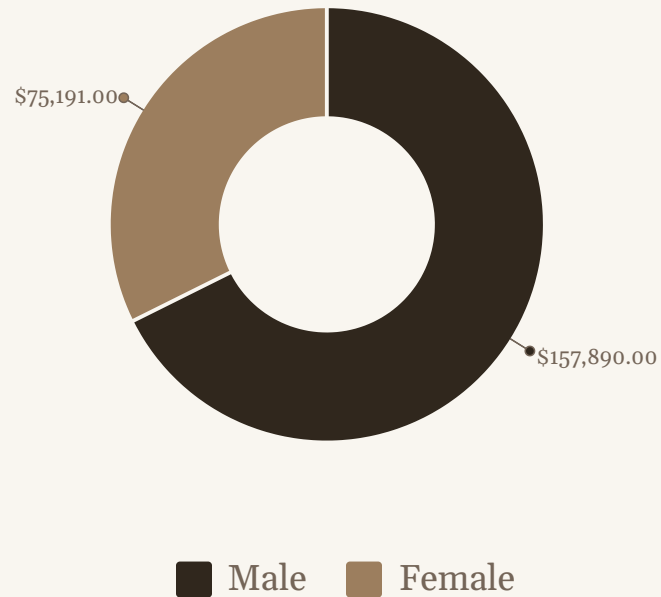
04

Database Integration

Connected to PostgreSQL and loaded cleaned data for SQL analysis



Revenue by Gender



Male customers generate **\$157,890** in revenue compared to **\$75,191** from female customers.

This 68% male revenue share suggests opportunities for targeted marketing to balance customer acquisition.

Key Business Insights



Top-Rated Products

Gloves (3.86), Sandals (3.84), and Boots (3.82) lead in customer satisfaction



Shipping Preferences

Express shipping users spend \$60.48 vs. \$58.46 for Standard—a 3.5% premium

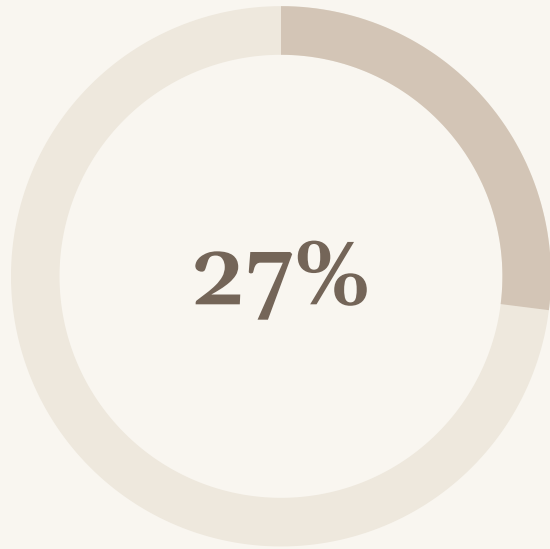


Smart Spenders

839 customers used discounts but still spent above average purchase amount

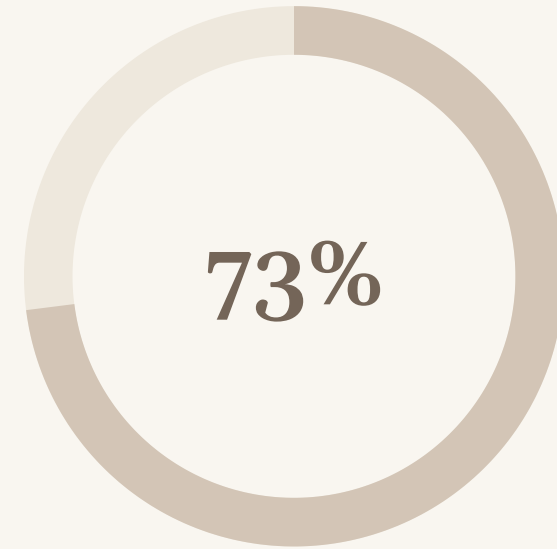


Subscription Analysis



Subscription Rate

1,053 of 3,900 customers are subscribers



Non-Subscribers

2,847 customers without subscriptions

Average Spend

Subscribers: **\$59.49**

Non-subscribers: **\$59.87**

Minimal difference suggests subscription benefits need enhancement

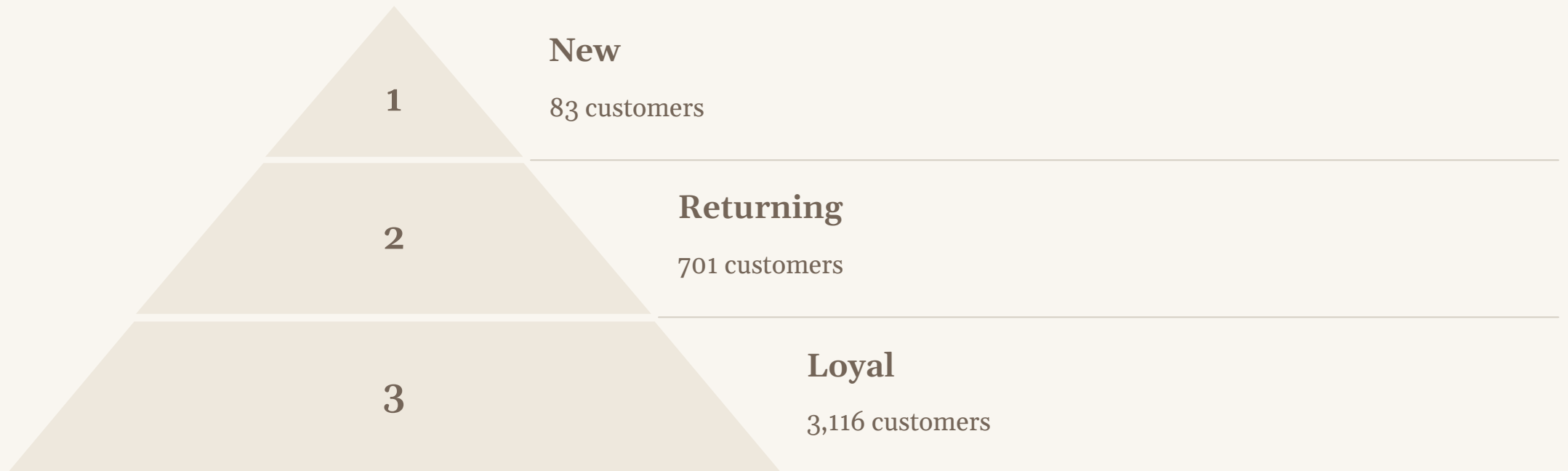
Repeat Buyers

Of customers with 5+ purchases:

- 958 are subscribers
- 2,518 are not subscribed

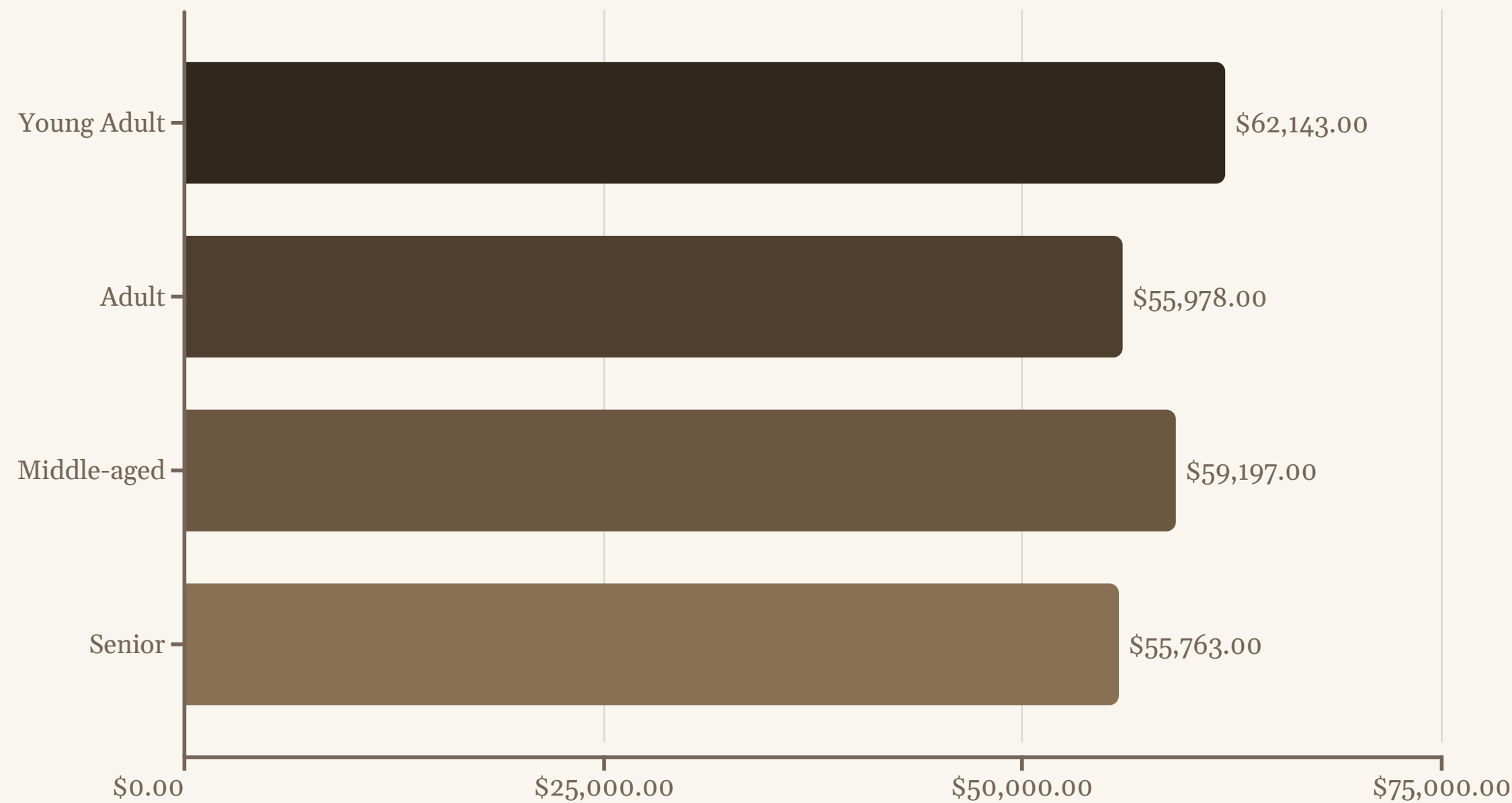
Huge opportunity for conversion

Customer Segmentation



Our customer base is **80% loyal customers**, demonstrating strong retention. Focus should shift to converting the 18% returning segment into loyal advocates through targeted loyalty programs.

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.



Product Performance Insights

Discount-Dependent Products

1. Hat (50% discount rate)
2. Sneakers (49.66%)
3. Coat (49.07%)
4. Sweater (48.17%)
5. Pants (47.37%)

These items rely heavily on promotions to drive sales

Top Products by Category

Accessories: Jewelry, Sunglasses, Belt

Clothing: Blouse, Pants, Shirt

Footwear: Sandals, Shoes, Sneakers

Outerwear: Jacket, Coat

Strategic Recommendations



Boost Subscriptions

Promote exclusive benefits to convert 2,518 repeat buyers who aren't subscribed



Loyalty Programs

Reward returning customers to accelerate their journey into the loyal segment



Review Discount Policy

Balance promotional sales with margin control, especially for discount-dependent products



Targeted Marketing

Focus on high-revenue age groups and express-shipping users who spend 3.5% more