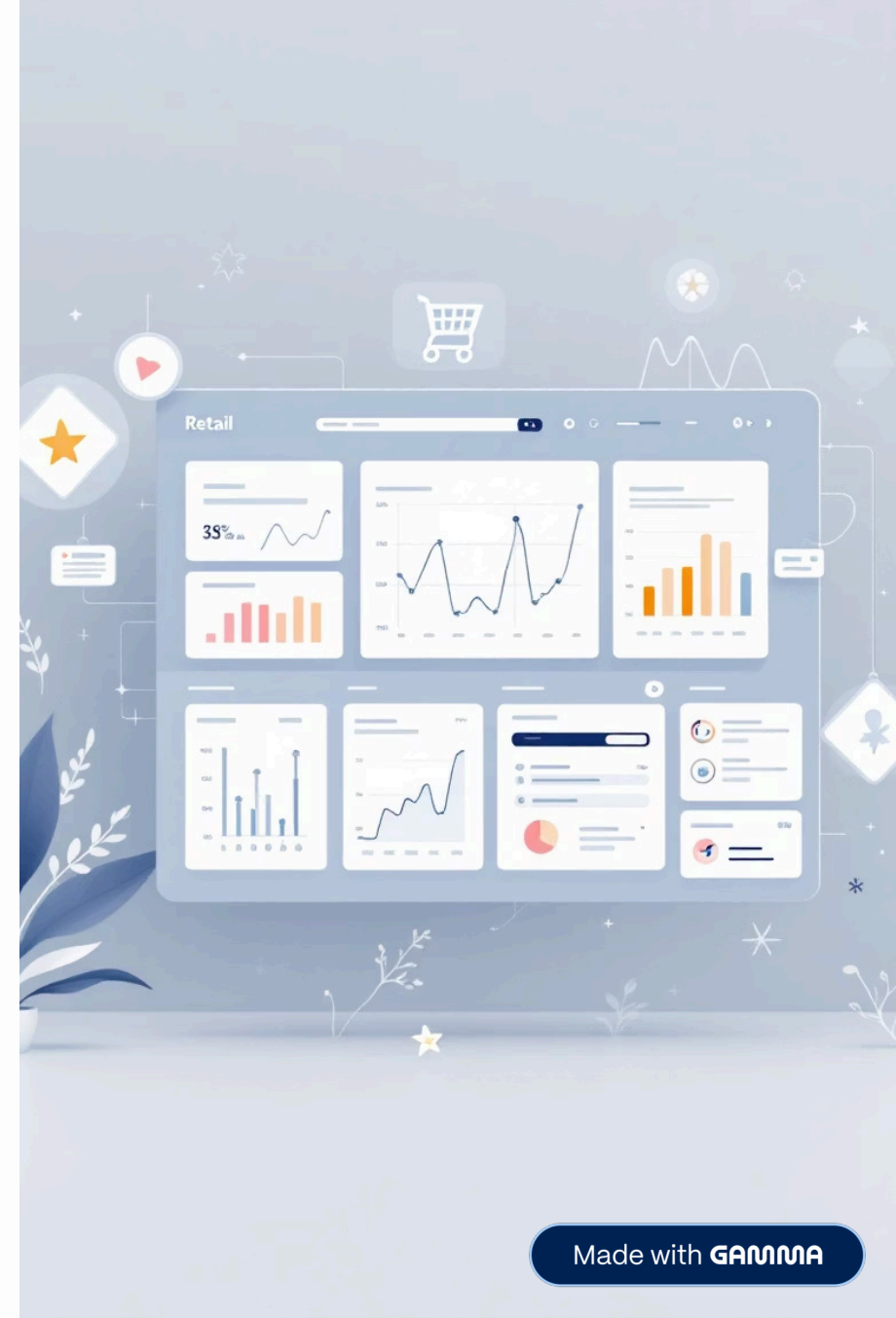


Customer Shopping Behavior Analysis

Analyzing 3,900 purchases to uncover insights into spending patterns, customer segments, and product preferences that drive strategic business decisions.



Dataset Overview

Scale & Scope

3,900 transactions analyzed across 18 key features including customer demographics, purchase details, and shopping behavior patterns.

- Customer demographics (Age, Gender, Location)
- Purchase details (Category, Amount, Season)
- Shopping behavior (Discounts, Reviews, Shipping)

Data Quality

Comprehensive dataset with minimal missing data—only 37 missing values in Review Rating column, handled through median imputation by product category.



Data Preparation & Engineering

01

Data Loading & Exploration

Imported dataset using pandas, performed initial exploration with summary statistics and structure analysis.

03

Feature Engineering

Created age_group bins and purchase_frequency_days columns to enable deeper segmentation analysis.

02

Missing Data Handling

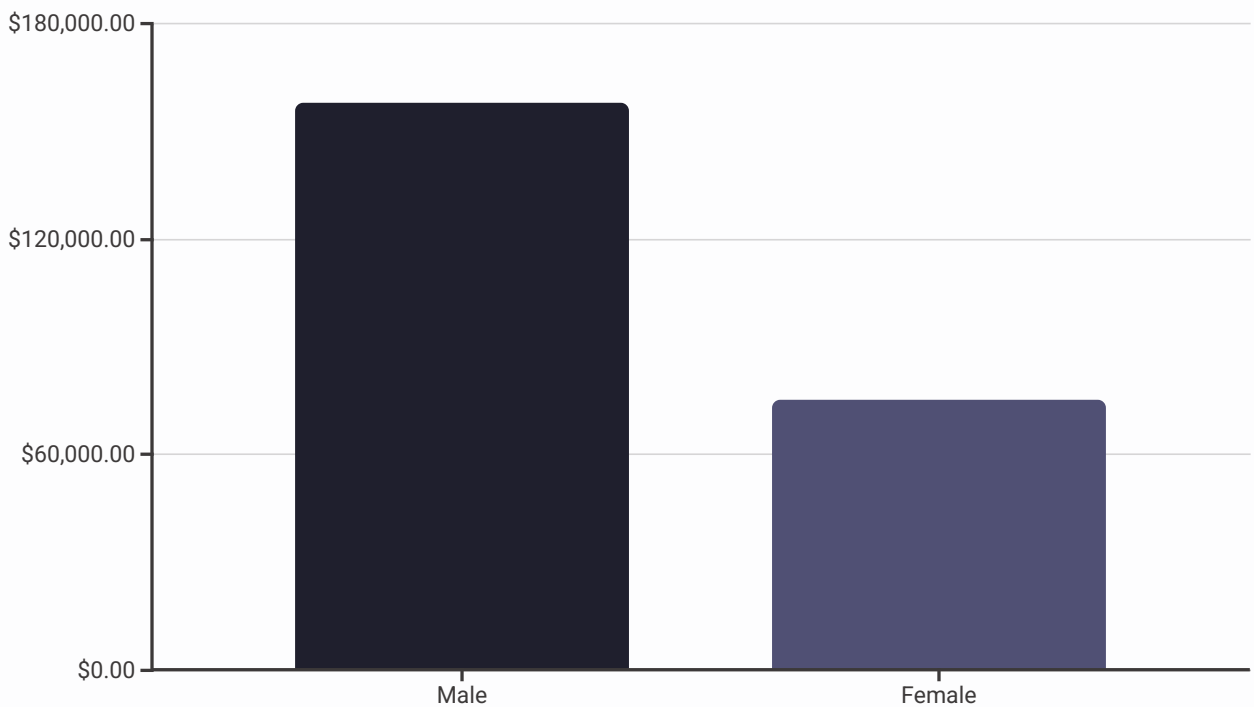
Imputed 37 missing Review Rating values using median rating per product category for accuracy.

04

Database Integration

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

Revenue Insights by Gender & Shipping



Male customers generate 2.1x more revenue than female customers, representing a significant opportunity for targeted marketing.

Shipping Preferences

Express shipping users spend **\$60.48** on average vs. **\$58.46** for Standard—a small but meaningful premium.



Top-Rated Products Drive Quality Perception



Gloves

Highest rated at 3.86 stars

Sandals

Strong rating of 3.84 stars

Boots

Rated 3.82 stars



Hat

Solid 3.80 star rating

Skirt

Rated 3.78 stars

These top-rated products should be prominently featured in marketing campaigns to leverage positive customer sentiment and drive conversions.



Discount Strategy Analysis

High-Value Discount Users

Multiple customers used discounts yet spent **above average**—purchases ranging from \$60 to \$100.

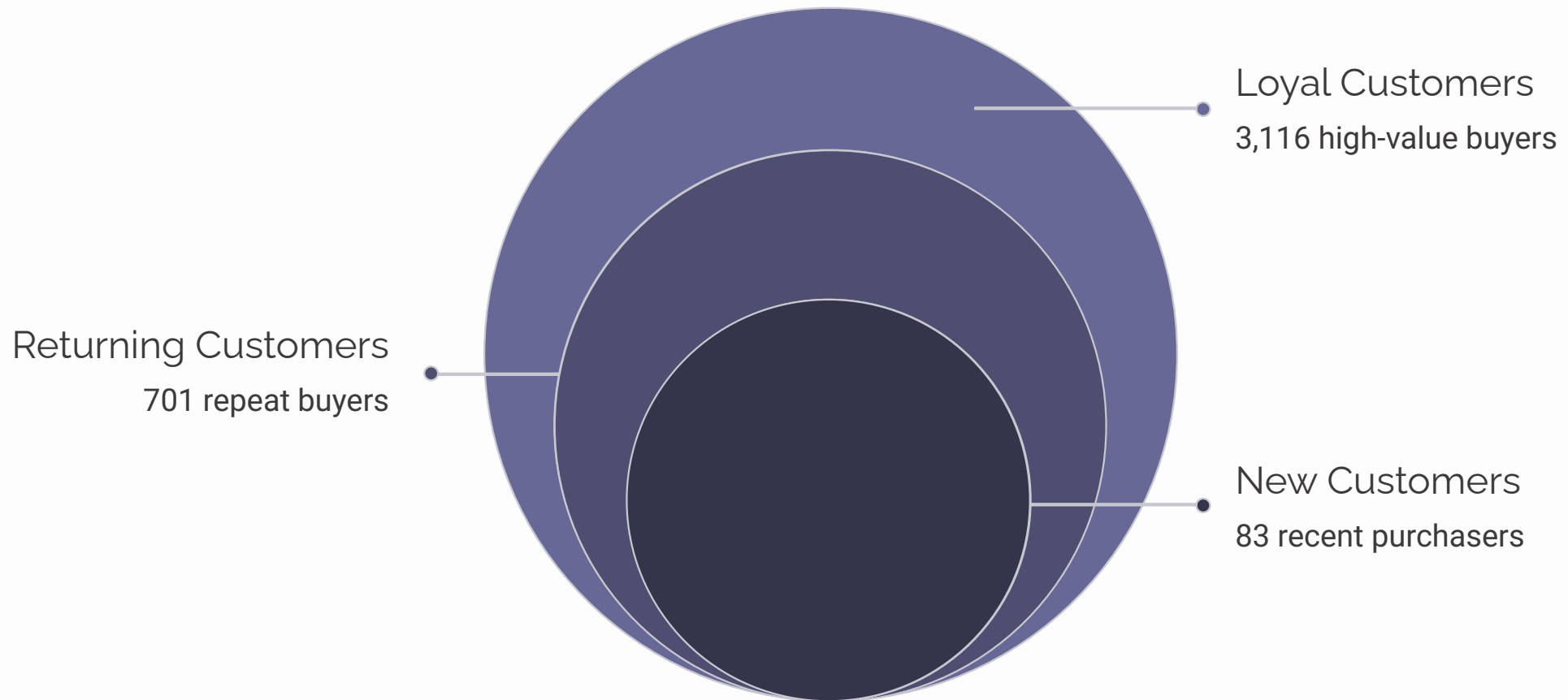
Top Discounted Products

Hats lead at 50% discount rate, followed by Sneakers (49.7%), Coats (49.1%), and Sweaters (48.2%).

Strategic Opportunity

Balance promotional activity with margin control—discounts drive volume without sacrificing high-value purchases.

Customer Segmentation Reveals Loyalty



Subscription Impact

Subscribers: **1,053 customers** averaging \$59.49 spend (\$62,645 total revenue)

Non-subscribers: **2,847 customers** averaging \$59.87 spend (\$170,436 total revenue)

Repeat Buyer Insight

Among customers with 5+ purchases: **958 are subscribers** vs. 2,518 non-subscribers—opportunity to convert loyal buyers.

Category Performance Breakdown

Clothing

\$104K revenue from 1,737 sales. Top items: Blouse, Pants, Shirt.

Accessories

\$74K revenue from 1,240 sales. Top items: Jewelry, Sunglasses, Belt.

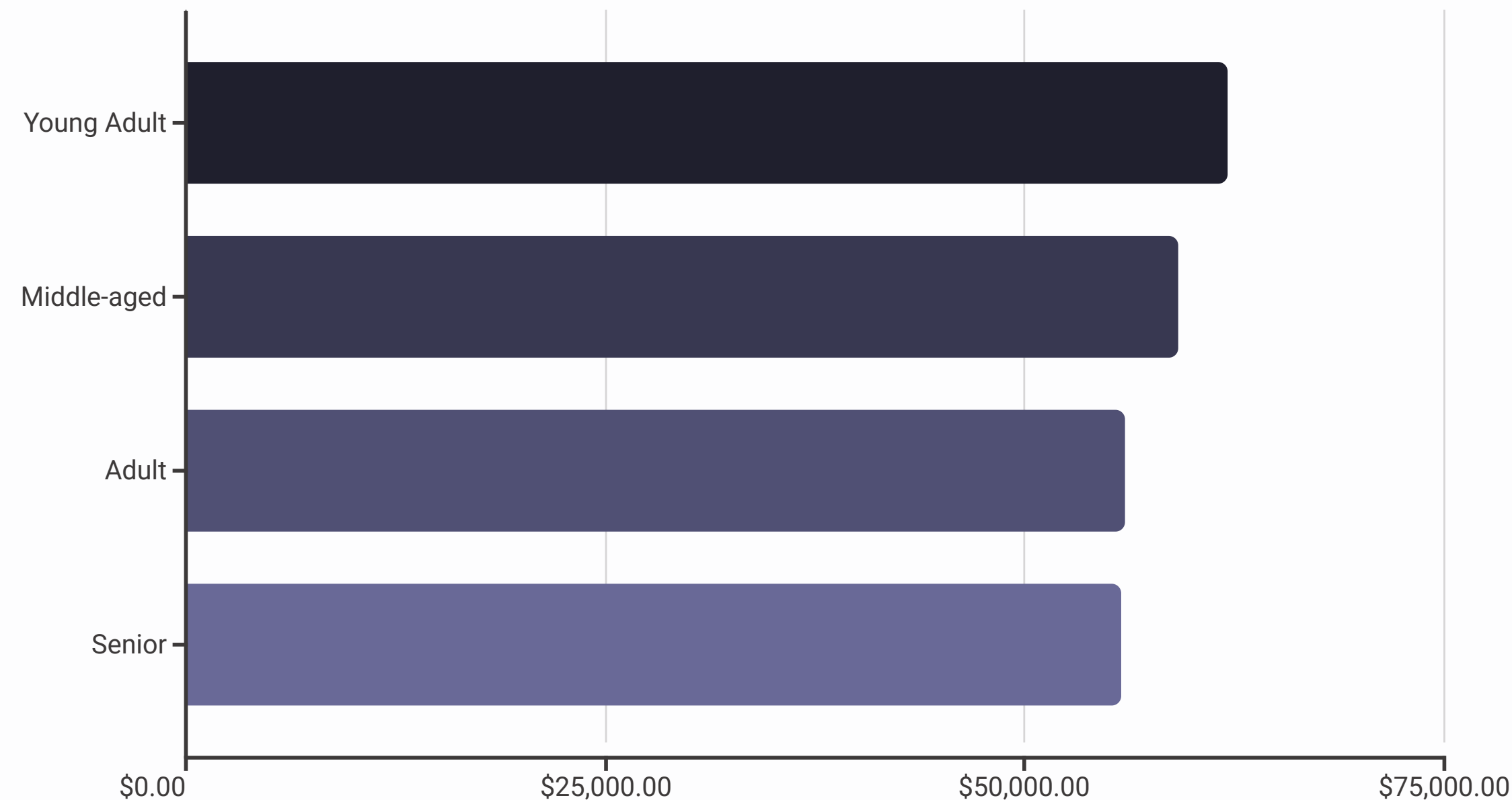
Footwear

\$36K revenue from 599 sales. Top items: Sandals, Shoes, Sneakers.

Outerwear

\$19K revenue from 324 sales. Top items: Jacket, Coat.

Age Group Revenue Distribution



Young Adults lead revenue generation at **\$62K**, followed closely by Middle-aged customers at **\$59K**. Focus marketing efforts on these high-value demographics.

Strategic Recommendations

Boost Subscriptions

Promote exclusive benefits to convert the 2,518 repeat buyers who aren't yet subscribers.

Customer Loyalty Programs

Reward repeat buyers to accelerate movement from Returning (701) to Loyal (3,116) segment.

Review Discount Policy

Balance promotional activity with margin control—high-value customers still respond to discounts.

Product Positioning

Highlight top-rated products (Gloves, Sandals, Boots) and best-sellers (Clothing category) in campaigns.

Targeted Marketing

Focus on Young Adults and Middle-aged customers, plus Express shipping users who demonstrate higher spend.