

# Customer Shopping Behavior Analysis

Uncovering insights from 3,900 purchases to guide strategic business decisions.





# Project Overview

## Goal

Analyze spending patterns, customer segments, product preferences, and subscription behavior.

## Data Source

Transactional data from 3,900 customer purchases across various product categories.

## Impact

Provide actionable insights for informed business strategies.

# Dataset Summary

**Rows:** 3,900

**Columns:** 18

## Key Features:

- Customer demographics (Age, Gender, Location, Subscription Status)
- Purchase details (Item, Category, Amount, Season, Size, Color)
- Shopping behavior (Discount, Promo Code, Previous Purchases, Frequency, Review, Shipping)

**Missing Data:** 37 values in "Review Rating"



# Exploratory Data Analysis (Python)

01

## Data Loading & Exploration

Imported dataset with pandas, checked structure and summary statistics.

02

## Missing Data Handling

Imputed "Review Rating" using median per product category.

03

## Column Standardization

Renamed columns to `snake_case` for readability.

04

## Feature Engineering

Created `age_group` and `purchase_frequency_days`.

05

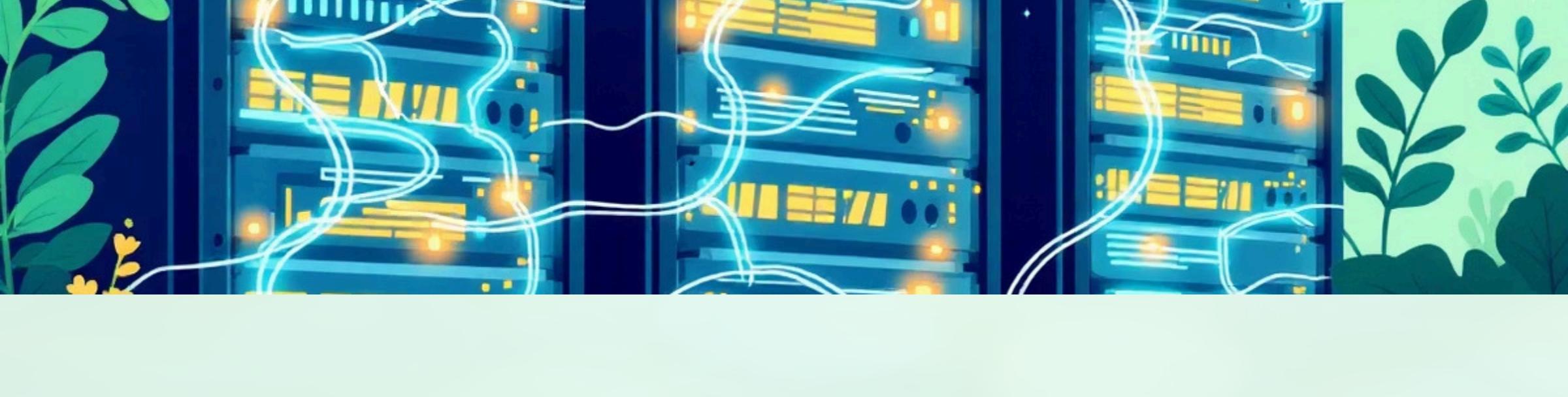
## Data Consistency

Dropped redundant `promo_code_used` column.

06

## Database Integration

Loaded cleaned data into PostgreSQL for SQL analysis.



# Data Analysis (SQL) - Key Findings

## Revenue by Gender

Male: \$157,890 vs. Female: \$75,191

## High-Spending Discount Users

839 customers used discounts but spent above average.

## Top 5 Products by Rating

Gloves (3.86), Sandals (3.84), Boots (3.82), Hat (3.80), Skirt (3.78).

# SQL Insights: Shipping & Subscriptions

## Shipping Type Comparison

Average purchase amounts:

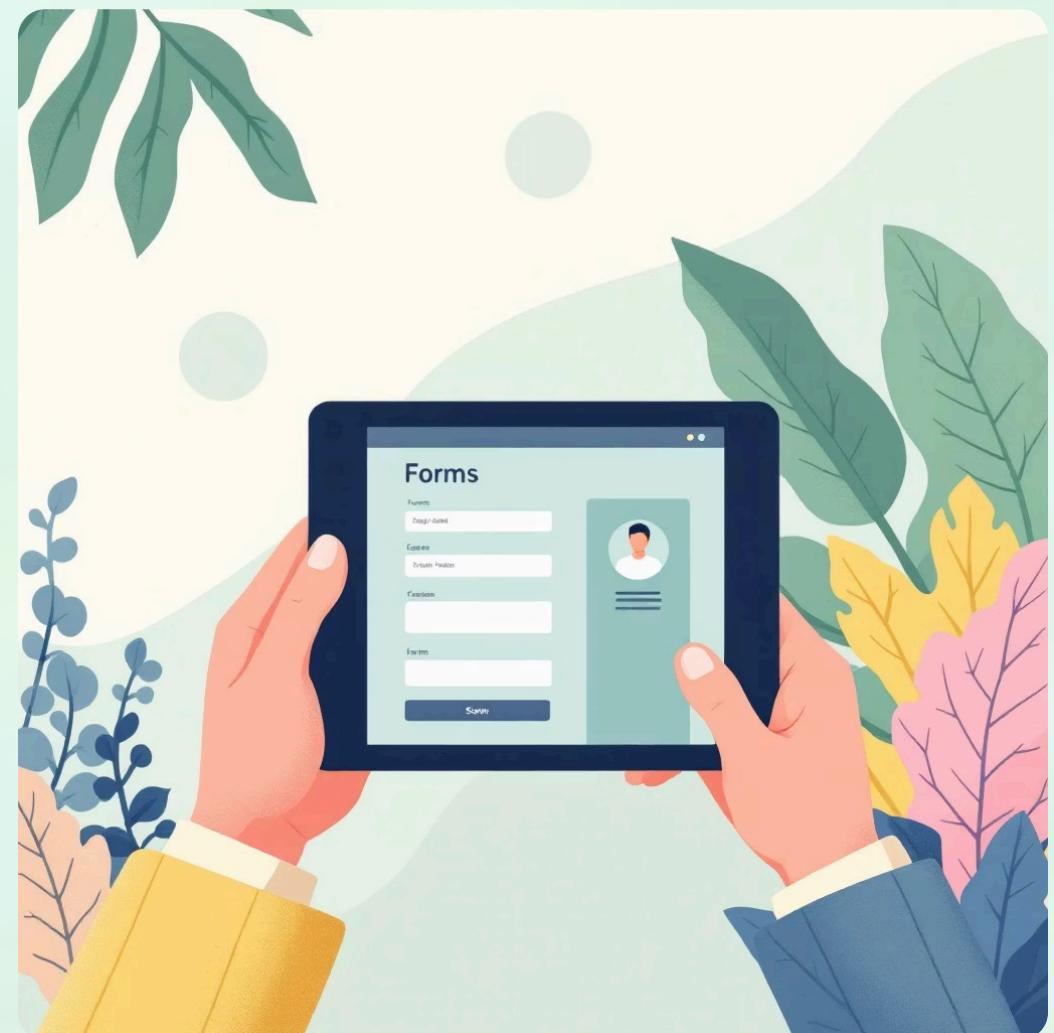
- Standard: \$58.46
- Express: \$60.48



## Subscribers vs. Non-Subscribers

**Subscribers (1,053):** Avg. Spend \$59.49, Total Revenue \$62,645

**Non-Subscribers (2,847):** Avg. Spend \$59.87, Total Revenue \$170,436



# SQL Insights: Product & Customer Behavior

1

## Discount-Dependent Products

Hat (50%), Sneakers (49.66%),  
Coat (49.07%), Sweater (48.17%),  
Pants (47.37%)

2

## Customer Segmentation

Loyal (3116), Returning (701),  
New (83)

3

## Repeat Buyers & Subscriptions

958 repeat buyers (5+ purchases) are subscribers, 2518 are not.



# Top Products by Category & Age Group Revenue

## Top 3 Products per Category

- **Accessories:** Jewelry, Sunglasses, Belt
- **Clothing:** Blouse, Pants, Shirt
- **Footwear:** Sandals, Shoes, Sneakers
- **Outerwear:** Jacket, Coat



## Revenue by Age Group

- **Young Adult:** \$62,143
- **Middle-aged:** \$59,197
- **Adult:** \$55,978
- **Senior:** \$55,763



# Power BI Dashboard: Visualizing Insights

An interactive dashboard presents key metrics and customer behavior visually.

3.9K

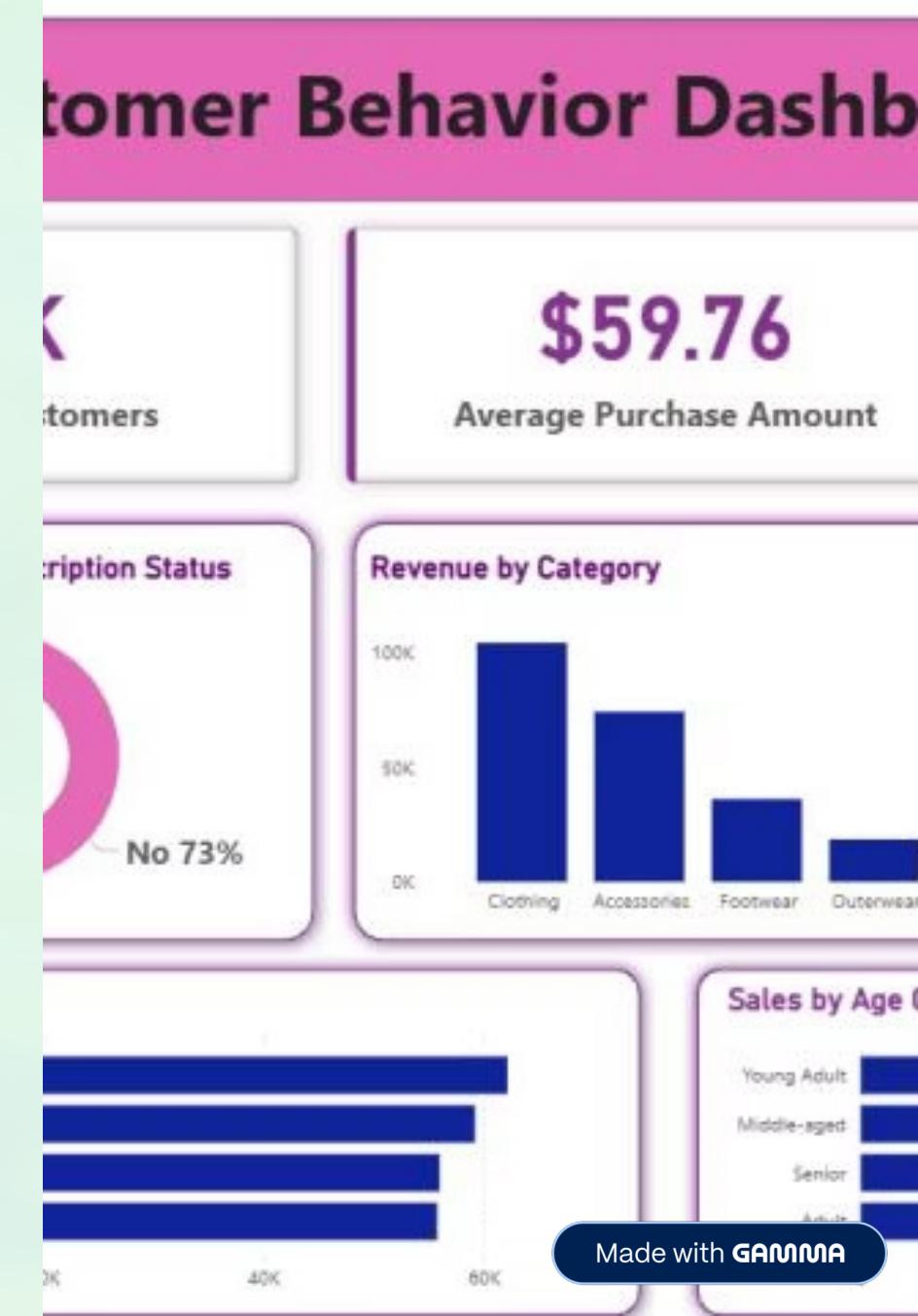
Customers

\$59.76

Avg. Purchase

3.75

Avg. Review



# Business Recommendations



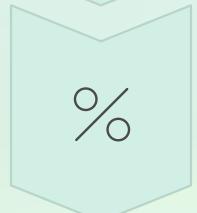
## Boost Subscriptions

Promote exclusive benefits to increase subscriber base.



## Customer Loyalty Programs

Reward repeat buyers to foster loyalty and retention.



## Review Discount Policy

Balance sales boosts with careful margin control.



## Product Positioning

Highlight top-rated and best-selling products in campaigns.



## Targeted Marketing

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