Stakeholder Deck Data-Driven Insights

Customer Shopping

Behavior Analysis

2026-02-16 | Customer Shopping Behavior Analysis | 2026

CONTENTS

Table of Contents

01 Executive Summary & Key Findings

02 Project Overview & Dataset

03 Methodology (Python EDA & Preparation)

04 SQL Findings (1-3): Gender, Discounts, Ratings

05 SQL Findings (4-6): Shipping, Subscription, Discount-Dependent Products

06 SQL Findings (7-10): Segments, Category Leaders, Repeat Buyers, Age

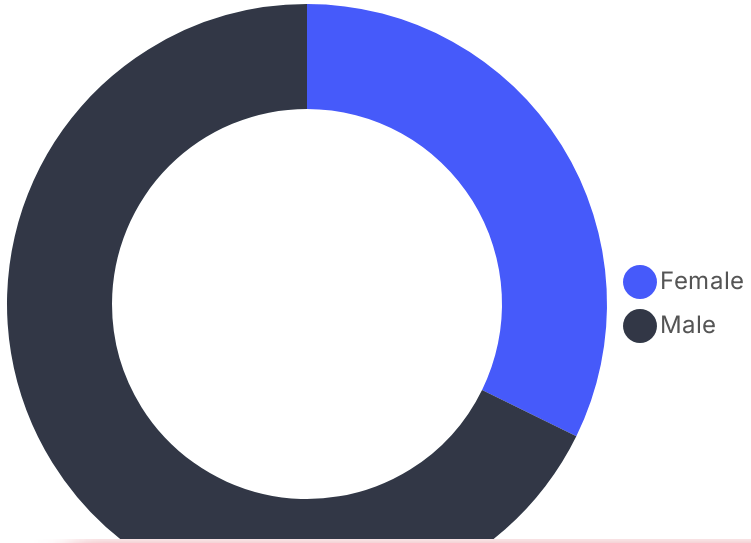
07 Power BI Dashboard

08 Business Recommendations

Executive Summary & Key Findings

Transactional analysis of 3,900 purchases surfaces clear patterns in revenue contribution, product performance, and customer segments.

Revenue by Gender



Avg Purchase Amount

Top 5 Products (Rating)

Gloves	3.86
Sandals	3.84
Boots	3.82
Hat	3.80
Skirt	3.78

Revenue by Age Group

Customer Segments

Loyal	3,116
Returning	701
New	83

Project Overview

Project Goal

Uncover insights on spending patterns, customer segments, product preferences, and subscription behavior to guide strategic decisions.

Data Scope

Analysis based on a comprehensive dataset of 3,900 transactions across various product categories.

Key Outputs

Delivering actionable insights, an interactive Power BI dashboard, and data-driven business recommendations.

Dataset Summary

Demographics

Age, Gender, Location, Subscription Status

Purchase Details

Item Purchased, Category, Purchase Amount, Season, Size, Color

Behavior & Experience

Discount Applied, Promo Code Used, Previous Purchases, Frequency of Purchases, Review Rating, Shipping Type

DATASET SHAPE

3,900
Total Rows

18
Columns

Data Quality Note

37 missing values identified in the *Review Rating* column.

Python EDA Workflow (Preparation)

Data Loading

Imported dataset with pandas

Initial Exploration

Checked structure with `df.info()`
Summary statistics with `.describe()`

Missing Data Handling

Imputed missing Review Rating using median rating within each product category

Column Standardization

Renamed columns to snake case

Workflow Summary

The preparation phase established a clean and consistent dataset foundation. By addressing missing values strategically and standardizing formats early, we ensured reliable downstream analysis in SQL and Power BI.

- Load → Explore → Clean missing → Standardize columns

Feature Engineering & Data Integration

Feature Engineering

- Created **age_group** by binning customer ages
- Created **purchase_frequency_days** from purchase data

Consistency Check

- Verified redundancy between *discount_applied* and *promo_code_used*
- Dropped *promo_code_used*

Database Integration

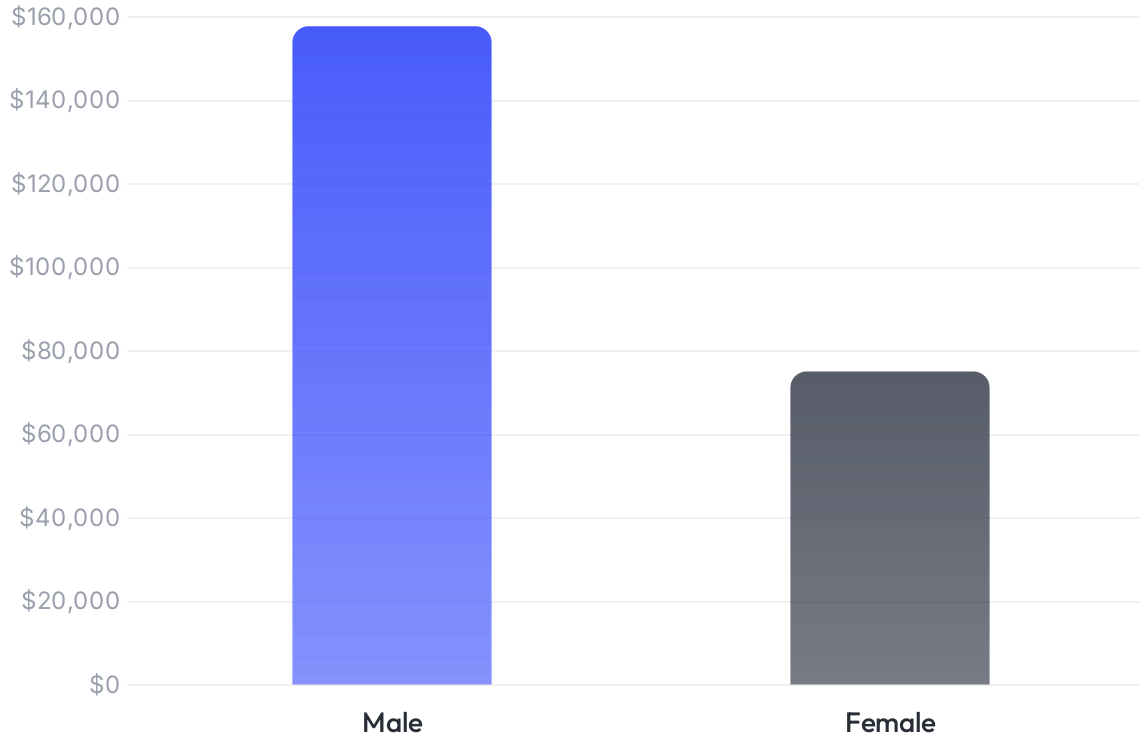
- Connected Python to PostgreSQL
- Loaded cleaned DataFrame for SQL analysis

Data Pipeline Flow

SQL Findings: Gender Revenue & Discount Behavior

Revenue by Gender

Comparison of total revenue generated by male vs. female customers.



High-Spending Discount Users

839 customers

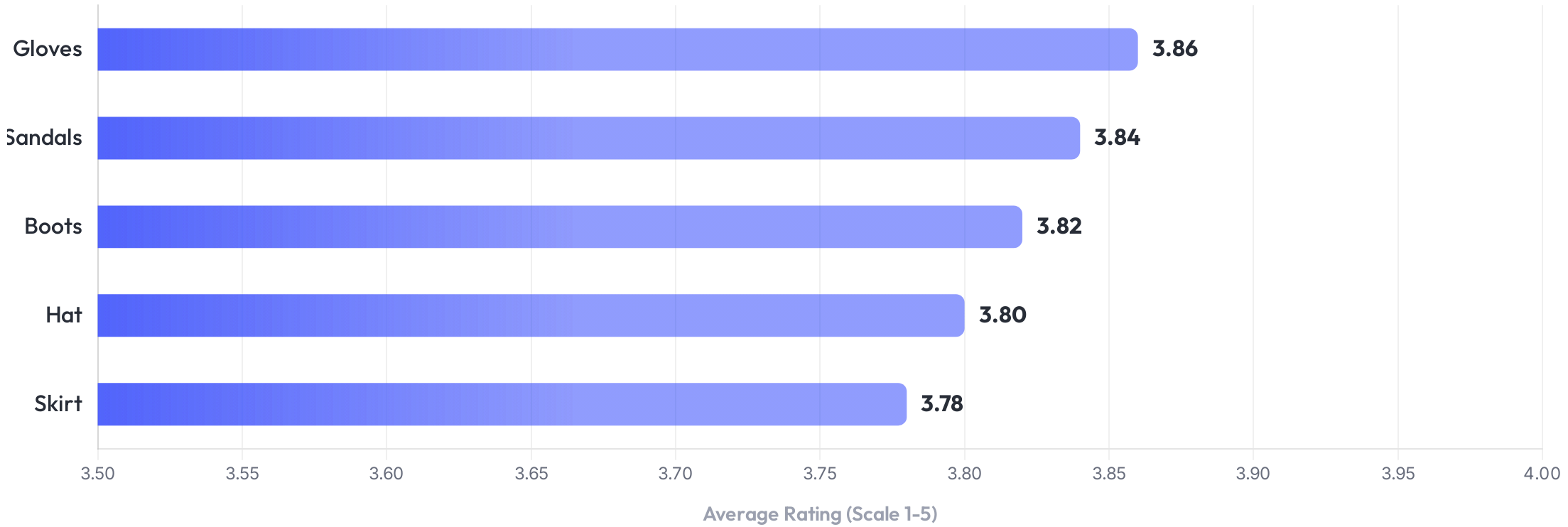
Identified customers who used discounts but still spent above the average purchase amount.

Sample Purchase Amounts ($)

64	73	90	85	97
68	72	81	90	62

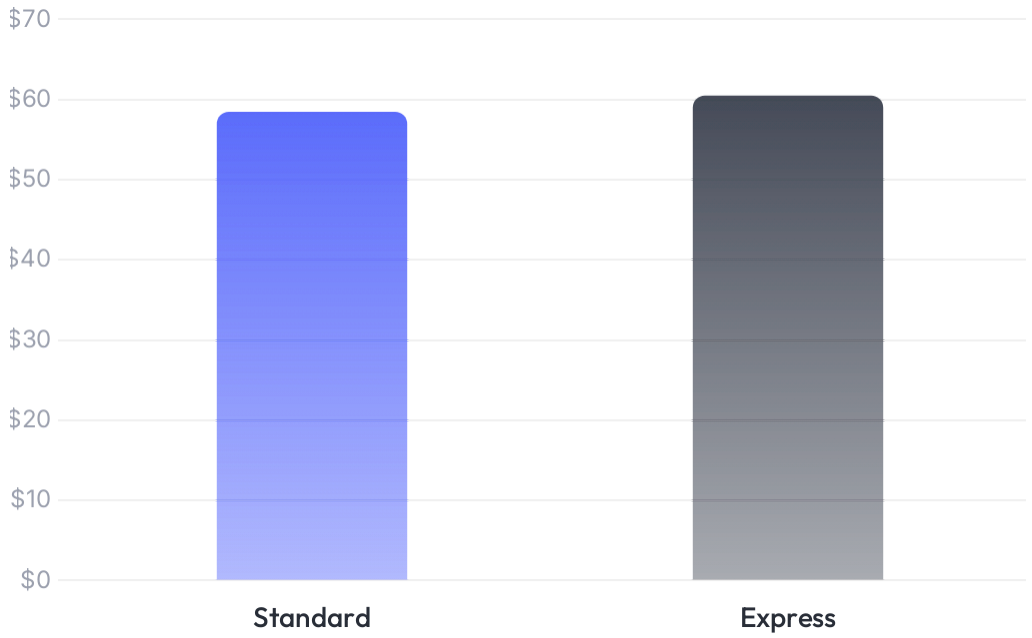
∗Sample of top 10 transactions

SQL Findings (3): Top 5 Products by Rating



SQL Findings (4–5): Shipping & Subscription

Avg Purchase Amount by Shipping Type



Subscribers vs Non-Subscribers

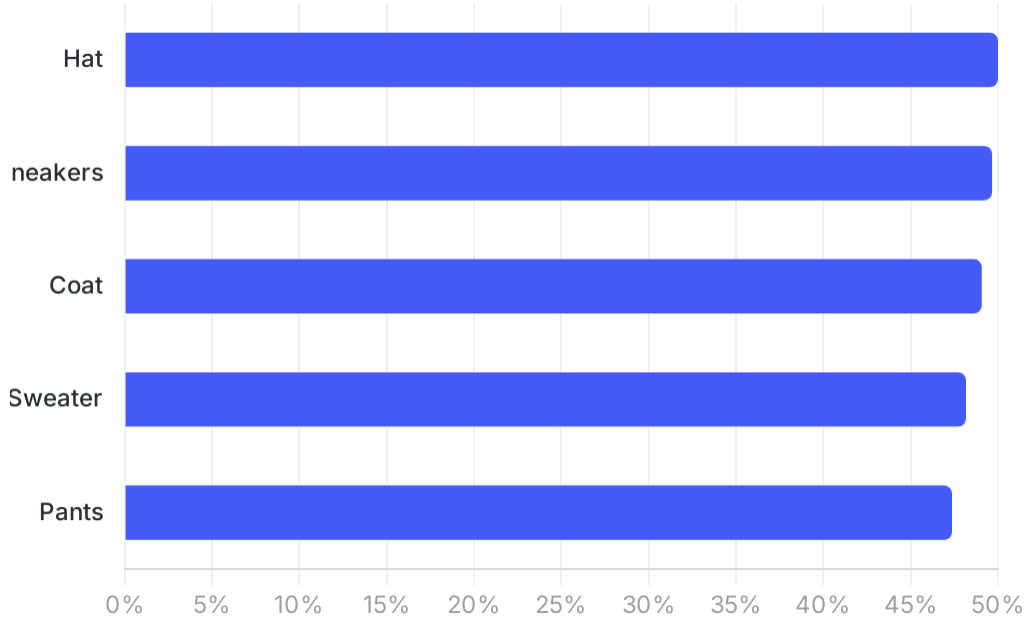
Status	Customers	Avg Spend	Revenue
Yes	1,053	$59.49	**$62,645**
No	2,847	$59.87	**$170,436**

Revenue Distribution

27% (Subscribers) 73% (Non-Subscribers)

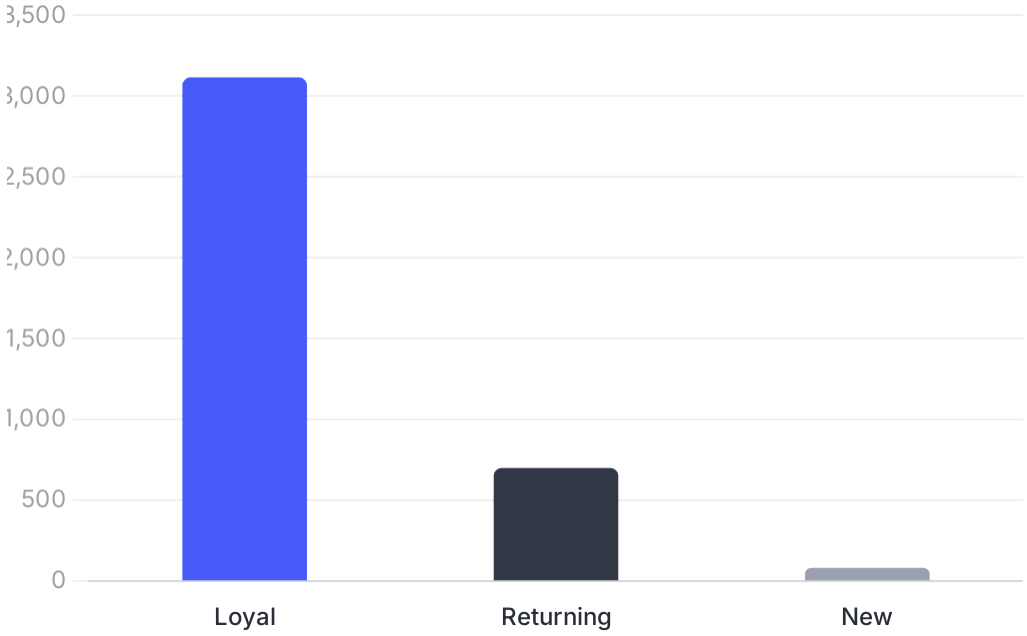
SQL Findings: Discount-Dependent Products & Segmentation

Top 5 Discount-Dependent Products



Products with highest percentage of discounted purchases.

Customer Segmentation by Count



Distribution of customers across Loyal, Returning, and New segments.

Additional SQL Insights

Top Products by Category

Accessories

Jewelry	171 orders
Sunglasses	161 orders
Belt	161 orders

Clothing

Blouse	171 orders
Pants	171 orders
Shirt	169 orders

Footwear

Sandals	160 orders
Shoes	150 orders
Sneakers	145 orders

Outerwear

Jacket	163 orders
Coat	161 orders

Repeat Buyer Insight

Customers with > 5 Purchases

WITHOUT SUBSCRIPTION

2,518 users

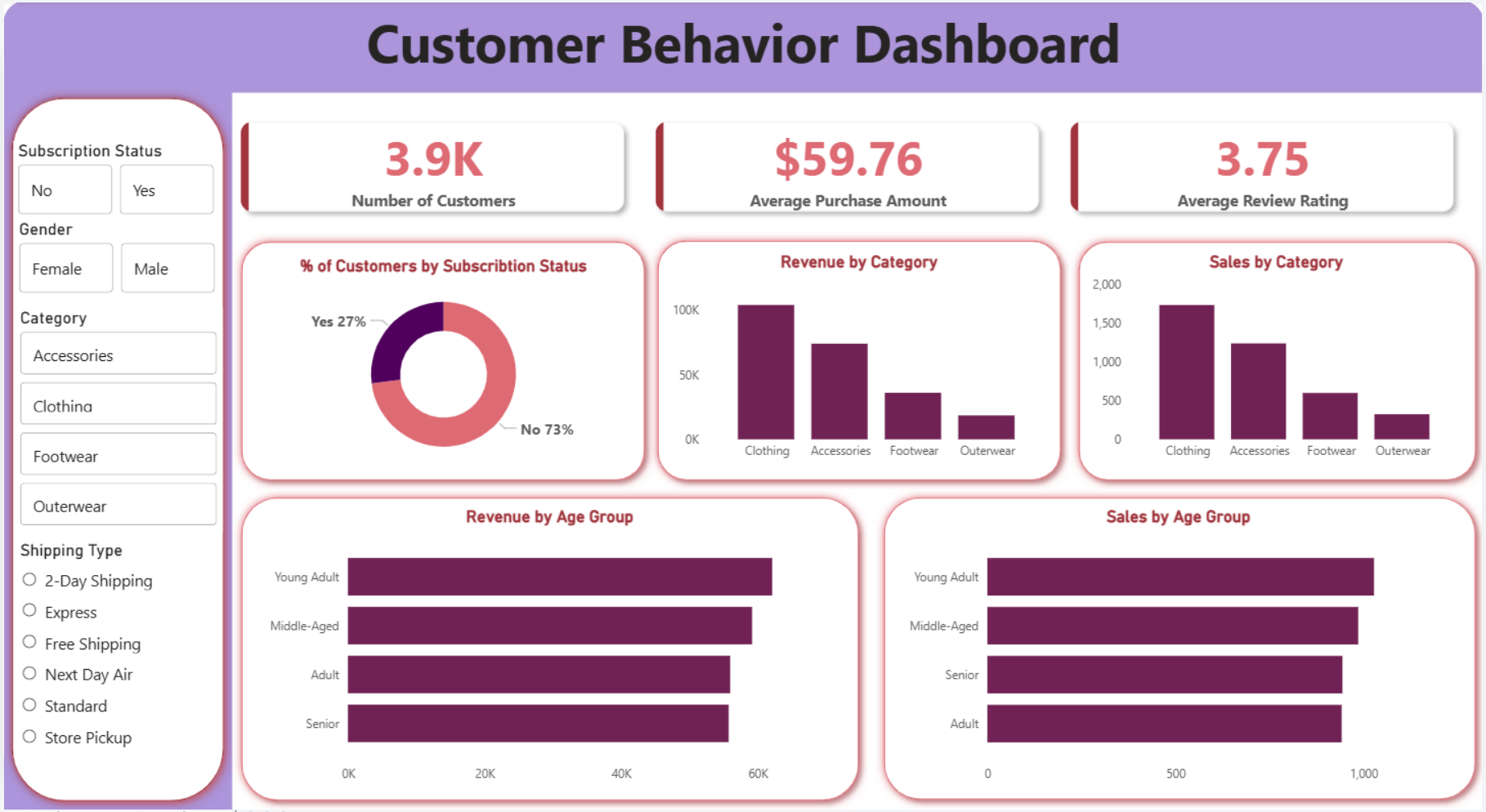
WITH SUBSCRIPTION

958 users

High volume of repeat buyers without subscription suggests opportunity for conversion.

Power BI Dashboard

Interactive dashboard built to present insights visually



Business Recommendations

Boost Subscriptions

Promote exclusive benefits for subscribers.

Customer Loyalty Programs

Reward repeat buyers to move them into the "Loyal" segment.

Review Discount Policy

Balance sales boosts with 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.

Strategic Growth Focus

Thank you

Data-driven customer insights for stakeholders