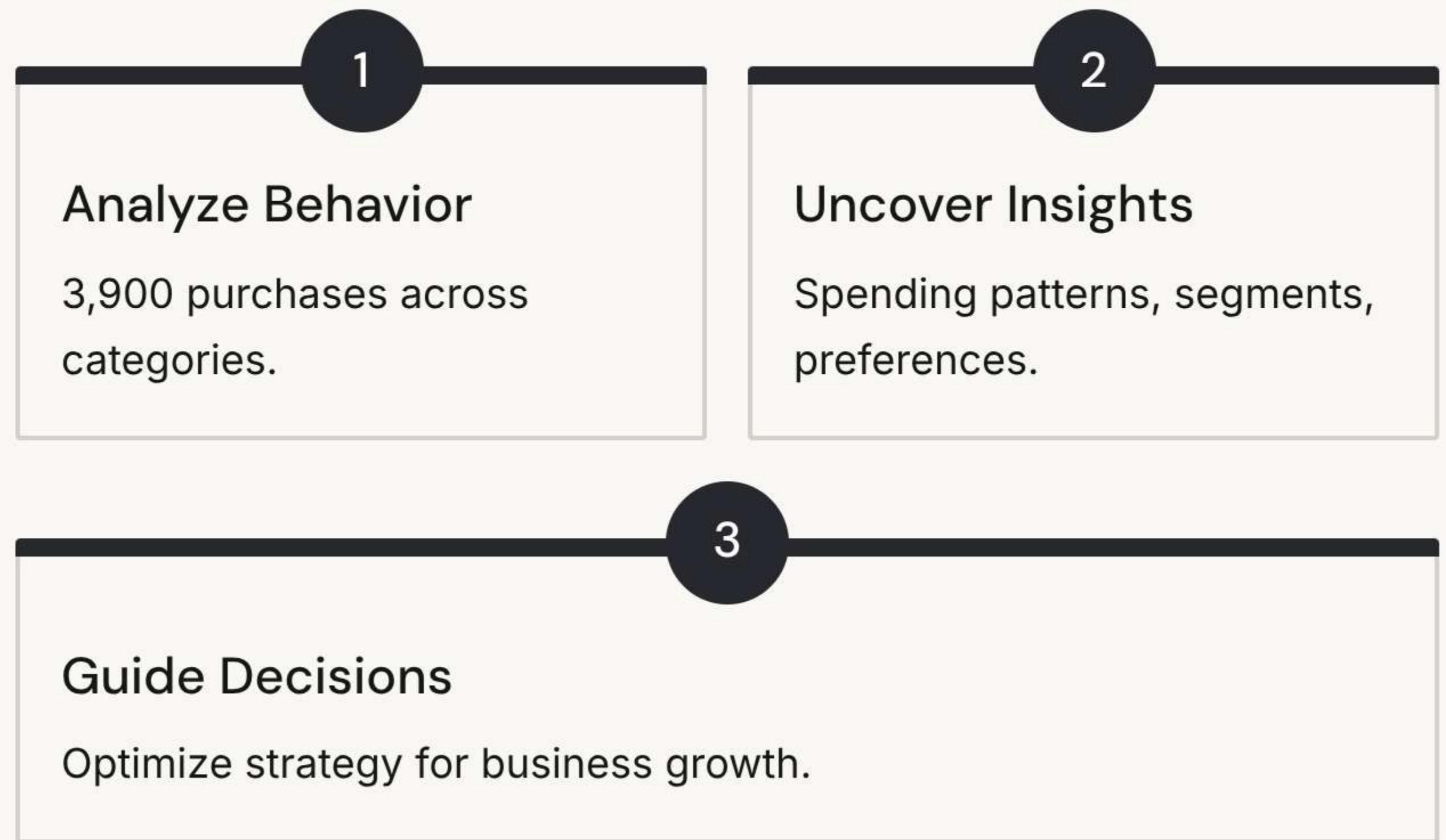




Customer Shopping Behavior Analysis

Uncovering insights from transactional data to guide strategic business decisions.

Project Overview





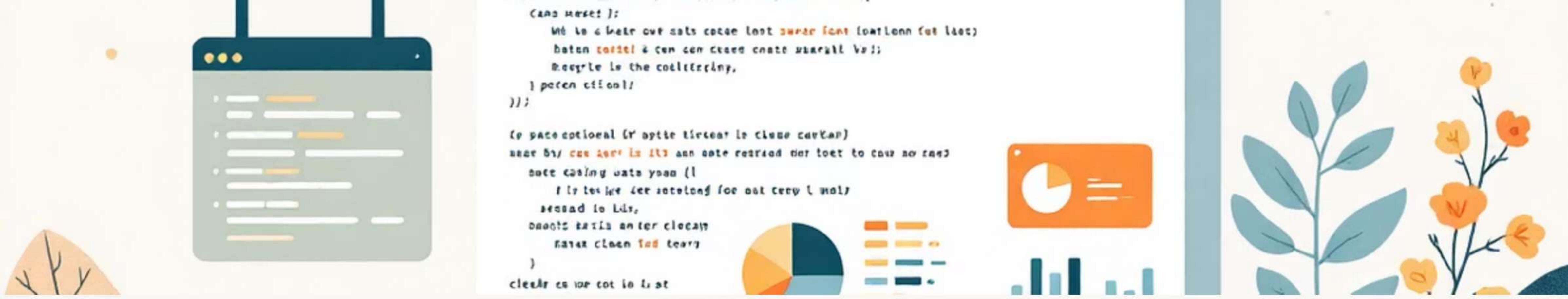
Dataset Summary

Key Features

- Customer demographics
- Purchase details
- Shopping behavior

Data Points

- Rows: 3,900
- Columns: 18
- Missing: 37 (Review Rating)



Exploratory Data Analysis (Python)

01

Data Loading & Exploration

Imported with pandas, checked structure and summary statistics.

03

Feature Engineering

Created age_group and purchase_frequency_days.

02

Missing Data Handling

Imputed Review Rating with median by category.

04

Database Integration

Loaded cleaned data to PostgreSQL for SQL analysis.

Data Analysis (SQL)

Revenue by Gender

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

High-Spending Discount Users

Identified 839 customers.

Top 5 Products by Rating

Gloves, Sandals, Boots, Hat, Skirt.

```
ROLLUP  
    customers FORMAT JSON;  
  
    SELECT  
        COUNT(*) AS count,  
        GROUP_CONCAT(DISTINCT gender) AS genders,  
        GROUP_CONCAT(DISTINCT age_group) AS age_groups,  
        GROUP_CONCAT(DISTINCT product_name) AS products  
    FROM customers  
    GROUP BY gender, age_group, product_name;  
  
    SELECT  
        SUM(amount) AS total_revenue  
    FROM sales  
    GROUP BY customer_id;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 1000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount < 1000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 1000 AND amount < 2000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 2000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 3000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 4000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 5000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 6000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 7000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 8000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 9000;  
  
    SELECT  
        COUNT(*) AS count  
    FROM sales  
    WHERE amount > 10000;
```

SQL Insights: Shipping & Subscriptions

Shipping Type Comparison

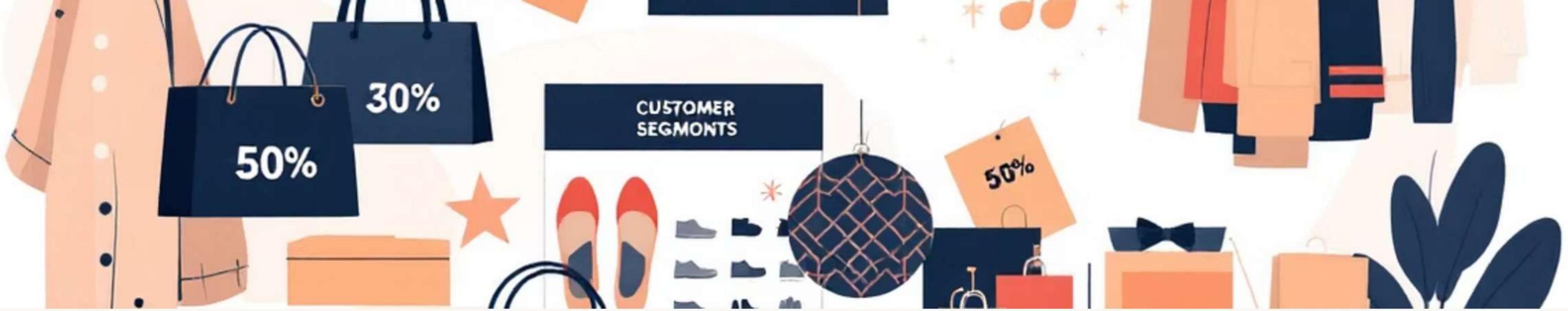
Standard: \$58.46, Express: \$60.48

Subscribers vs. Non-Subscribers

Subscribers: 1,053, Avg Spend: \$59.49

Non-Subscribers: 2,847, Avg Spend: \$59.87





SQL Insights: Products & Segments



Discount-Dependent Products

Hat, Sneakers, Coat, Sweater, Pants.



Customer Segmentation

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



Top 3 Products per Category

Jewelry, Blouse, Sandals, Jacket.



45%

SQL Insights: Repeat Buyers & Age

Repeat Buyers & Subscriptions

958 repeat buyers are subscribers.

2518 repeat buyers are not subscribers.

Revenue by Age Group

Young Adult: \$62,143

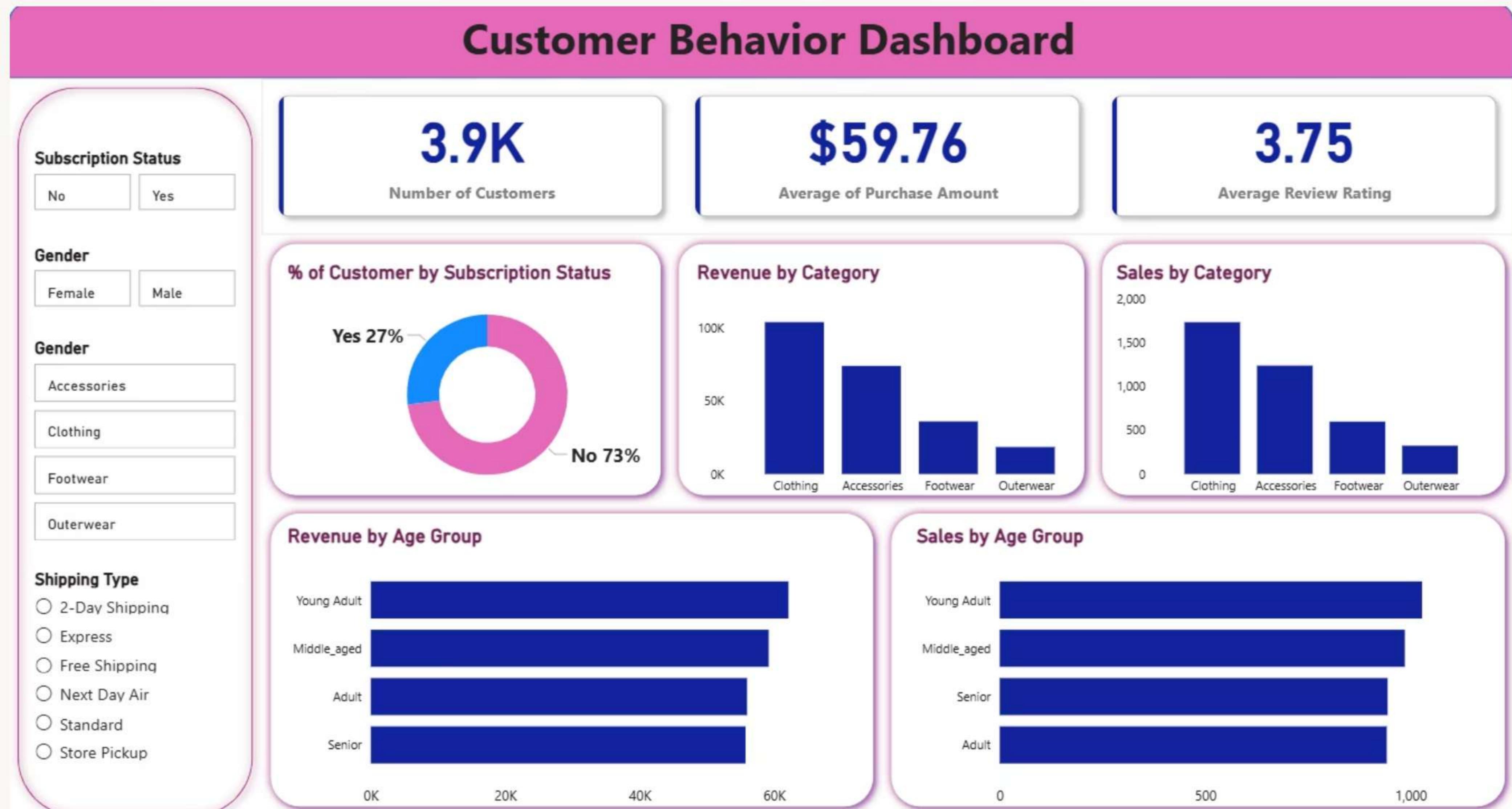
Middle-aged: \$59,197

Adult: \$55,978

Senior: \$55,763

Power BI Dashboard

Interactive visualization of key customer behavior metrics.





Business Recommendations

- **Boost Subscriptions**
Promote exclusive benefits.
- **Customer Loyalty Programs**
Reward repeat buyers.
- **Review Discount Policy**
Balance sales with margin control.
- **Targeted Marketing**
Focus on high-revenue age groups.