Customer Shopping

Exploratory Data Analysis
using Python

by Chyntia Novrianti Kudadiri

01 Understanding Data

02 Bussiness Problem

03 Data Preparation

04 Data Visualization

05 Recommendation for Business

Understanding Data

Costumer Shopping Dataset is a dataset containing shopping information in Istanbul.

This dataset comes from 10 different shopping malls starting from Jan 1, 2021 to March 8, 2023.

The dataset can be downloaded on Kaggle

Attribute Information:

- **invoice_no**: Invoice number.
- **customer_id**: Customer number.
- **gender**: customer's gender.
- **age**: customer's age.
- **category**: category of the purchased product.
- **quantity**: The quantities per transaction.
- **price**: Product price per unit in Turkish Liras (TL).
- **payment_method**: payment method used for the transaction.
- **invoice_date**:The day when a transaction was generated.
- **shopping_mall**: shopping mall where the transaction was made.

Business Problem

Shopping is one of the factors that contribute to economic growth. Shopping can increase a country's income. Therefore, by analyzing customer behavior using Customer Shopping Dataset, companies can determine business strategies to attract customers to shop.

In this data set, I'll explore some of the following through visualization :

- What month is the most and the least total spending?
- Which gender spend more money?
- What payment method is the most used by customers?
- Which category the customer spend most money?

- Which shopping mall is most visited by customers?
- Distribution by customer's age
- Distribution by price
- In 2023, Which shopping mall is most visited by customers?
- In 2023, Which Shopping mall has the highest total revenue?
- In 2023, What day in a week have the most transaction?
- In 2023, What day in a week have the most total spending?
- In 2023, Which category the customer spend most money?

Data Preparation

This project use Google Colab. If you want to see the whole project, you can find it **here**.

I did data preparation before explore the data :
1. Add column total_price = quantity * price
2. Add column invoice_month (month of invoice_date)
3. Change type of invoice_date and invoice_month into datetime
4. Add column invoice_date (day of week of invoice_date)
5. Add column age_group (range of age)
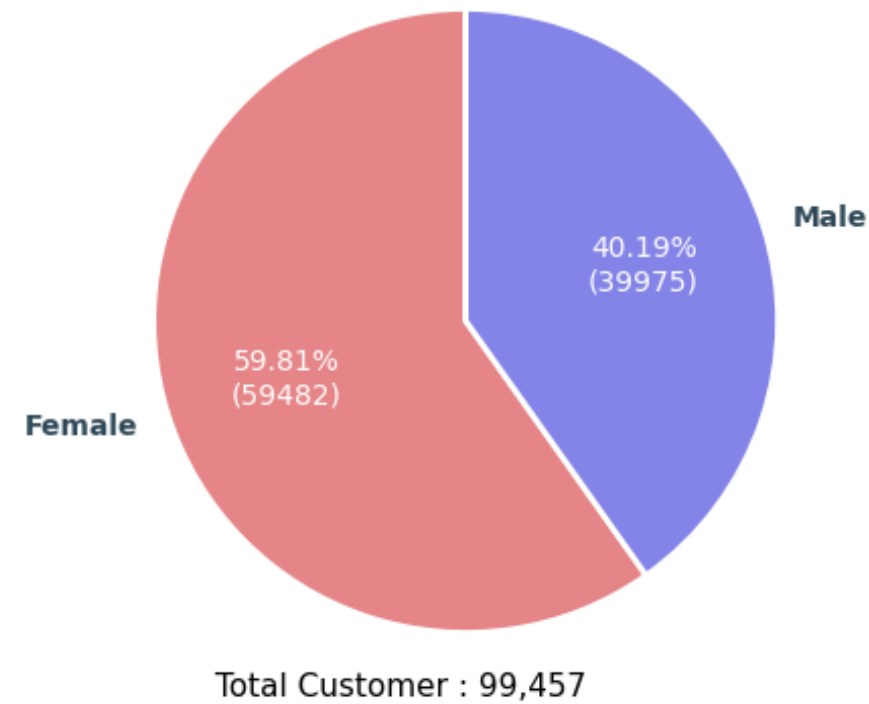6. Add column price_group (range of price)

Data Visualization

Total Spending per Month
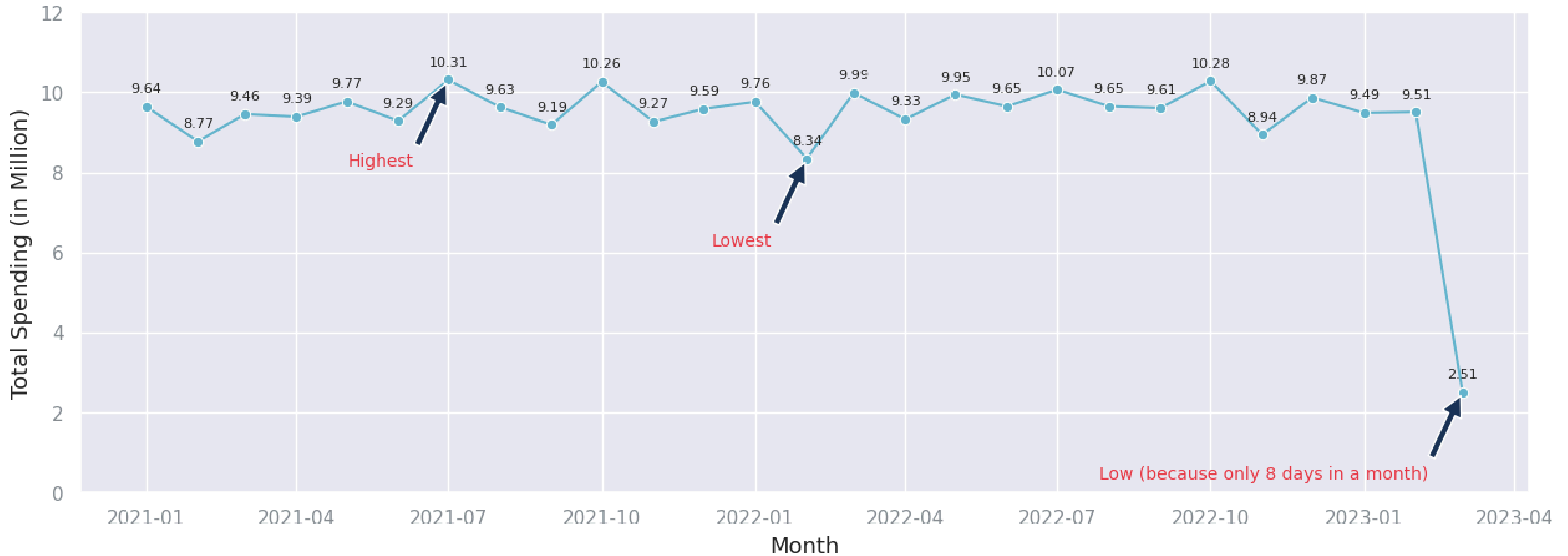
Customers spent around **10.31 million in July 2021** which is the **highest** total spending and the **lowest** total spending **in February 2022** around **8.34 million**.

Percentage Total Customer per Gender

Male 40.19% (39975)

Female 59.81% (59482)

Total Customer : 99,457

Percentage Total Spending per Gender

Male 40.28% (101.30 M)

Female 59.72% (150.21 M)

Total Spending : 251,505,794

Number of Customers and Total Spending by Gender
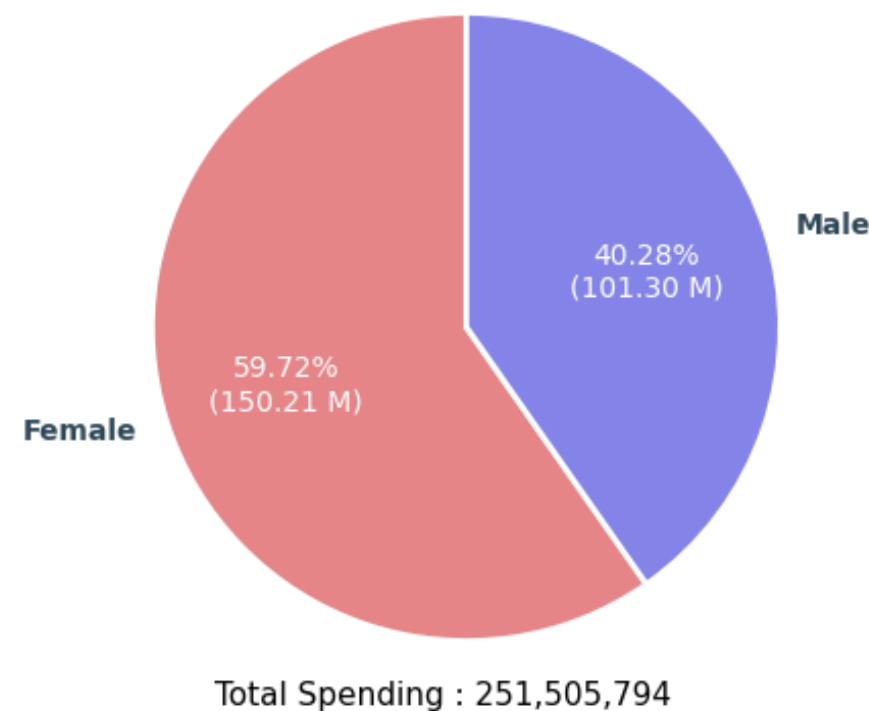
From January 2021 to March 2023, **female customers dominated** the customer base, with total spending **19.44%** higher than male customers

The company is recommended to focus on developing relevant marketing strategies for female customers to increase loyalty and purchase value, while also implementing targeted strategies to boost male customers' purchasing interest

Total Customer per Payment Method
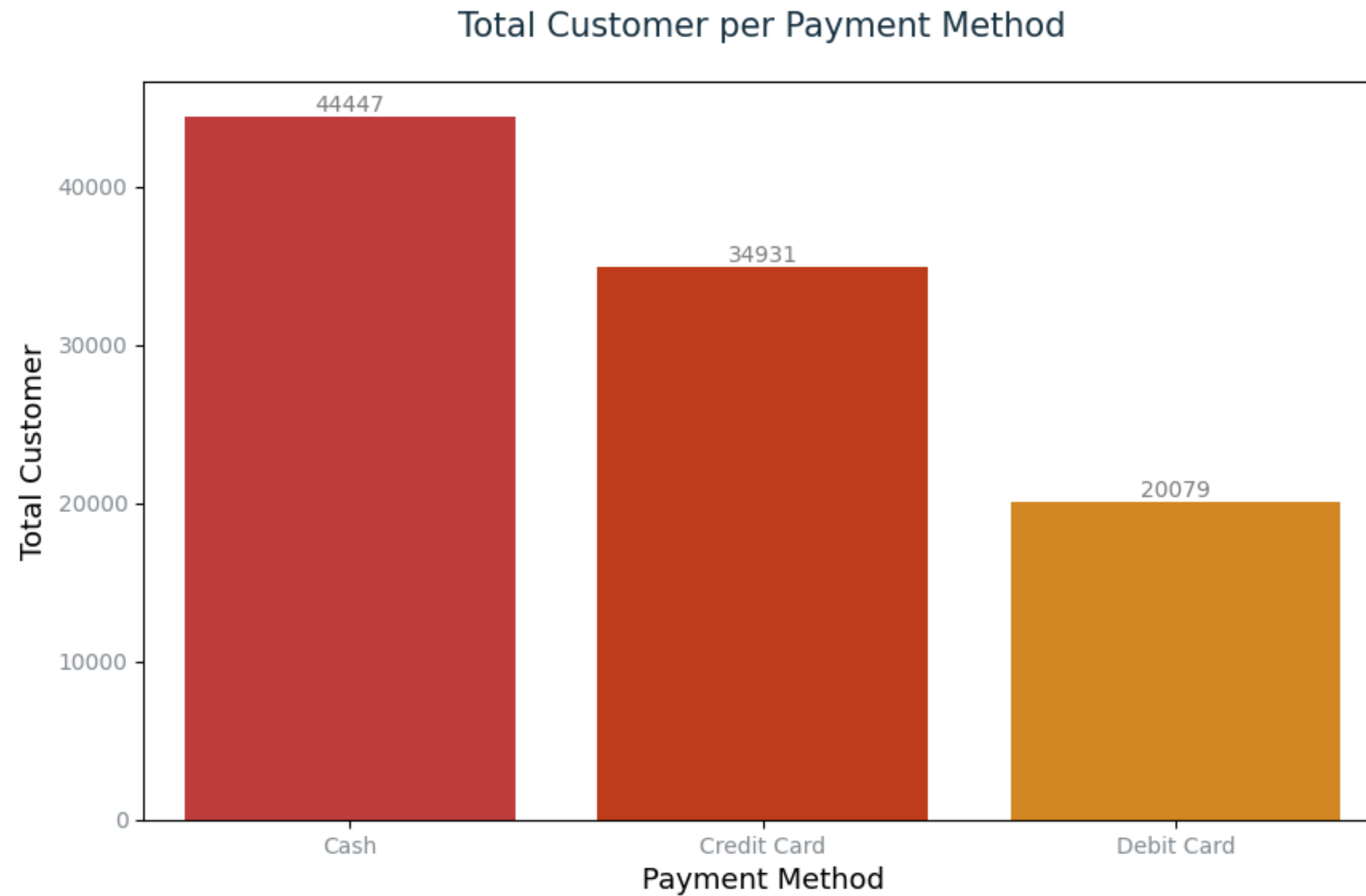
Since cash payments dominate over debit and credit cards, the company should maintain cash payment options while encouraging non-cash transactions through targeted promotions or incentives to improve payment efficiency.
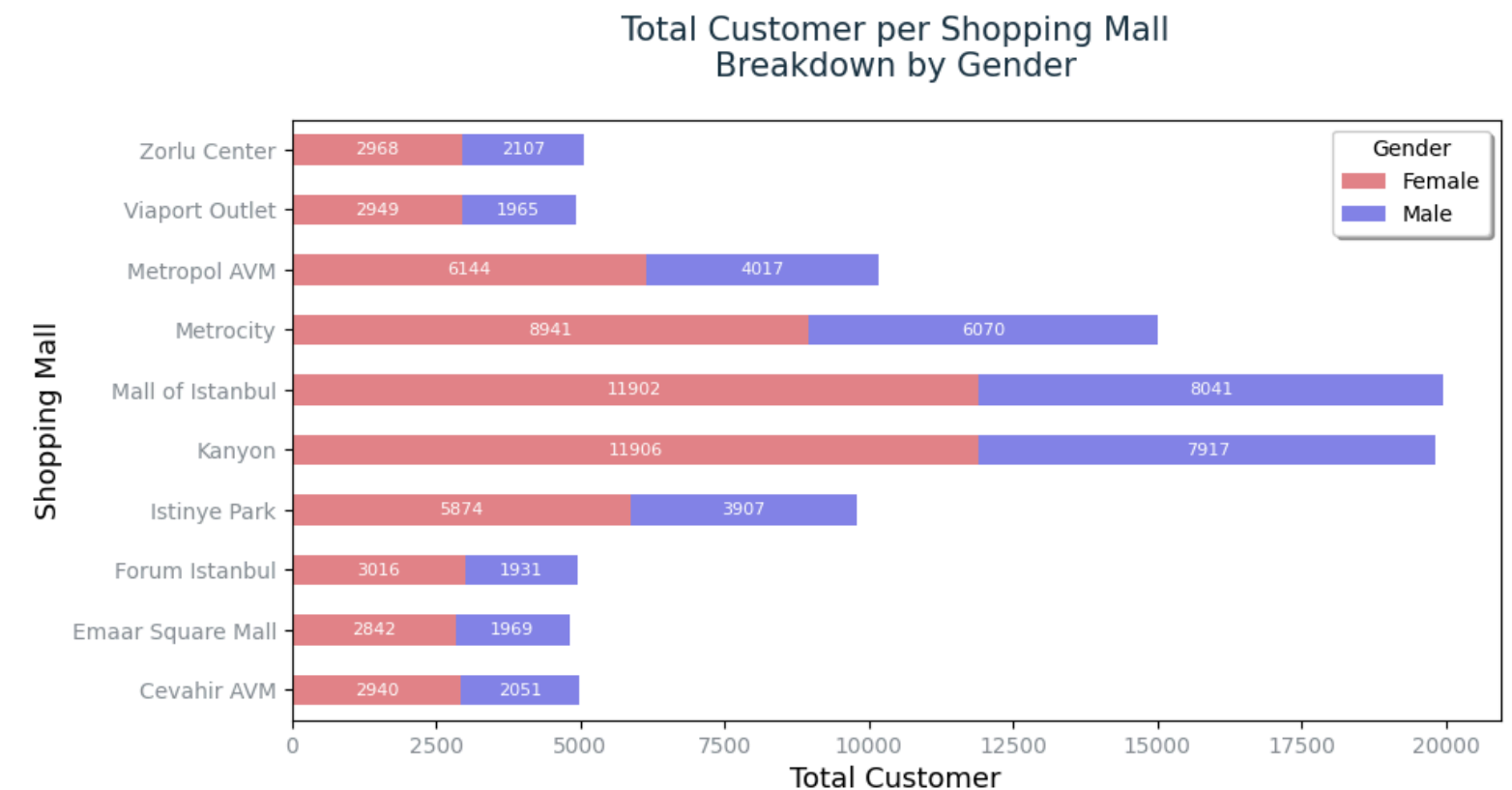
Total Sales per Category
Breakdown by Gender
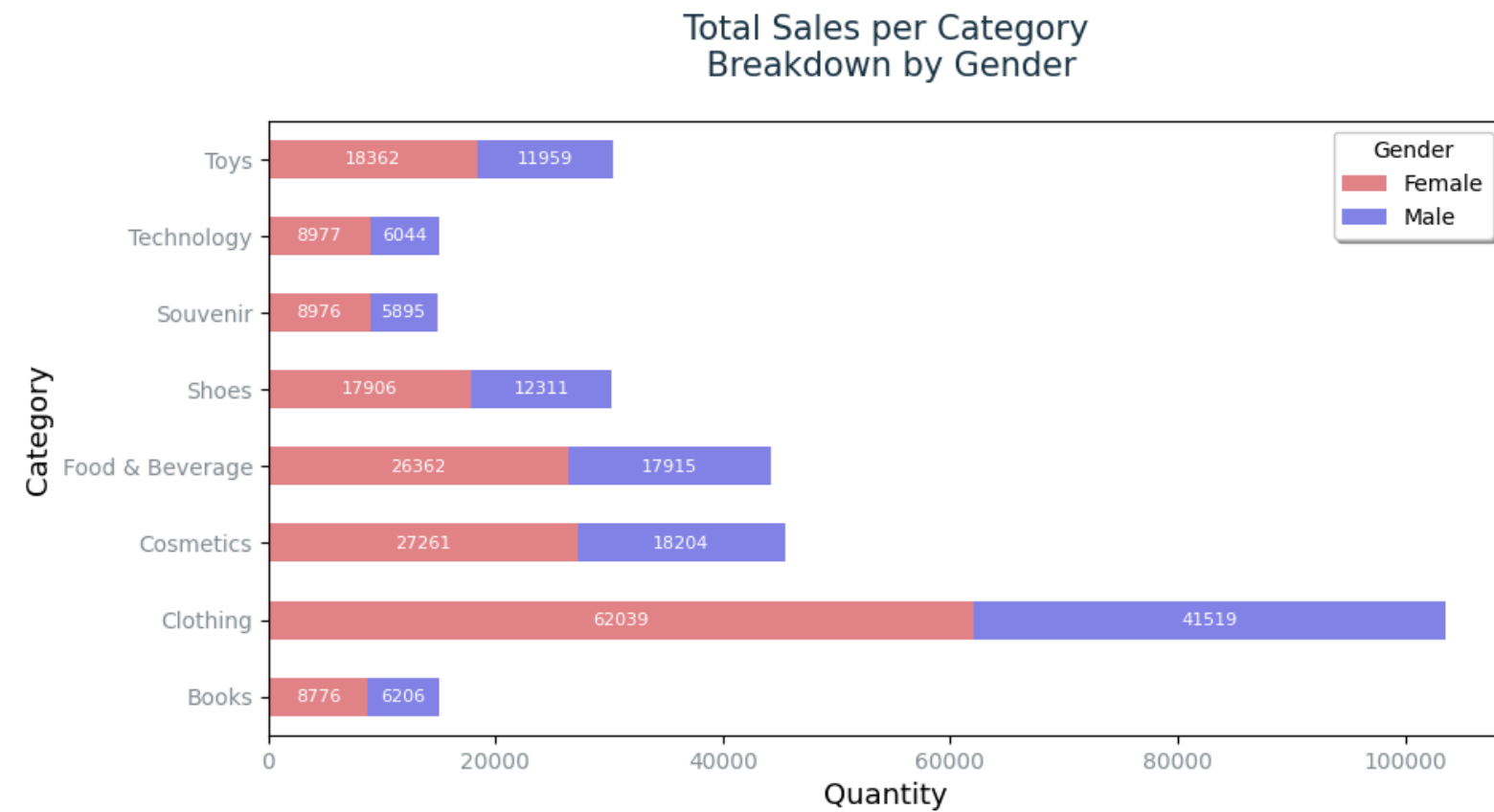
Top Category

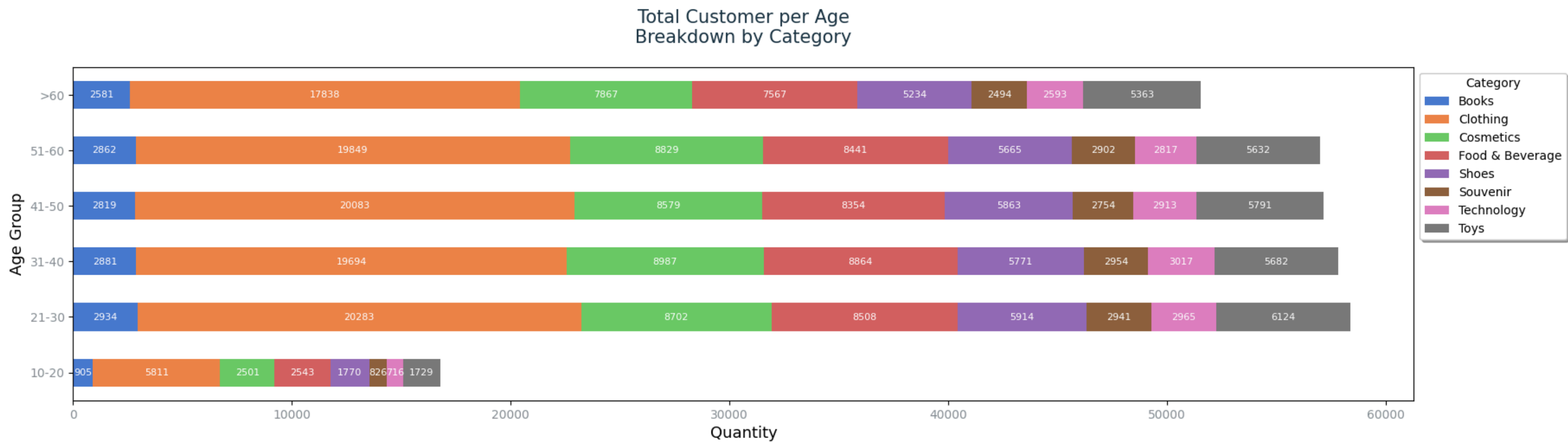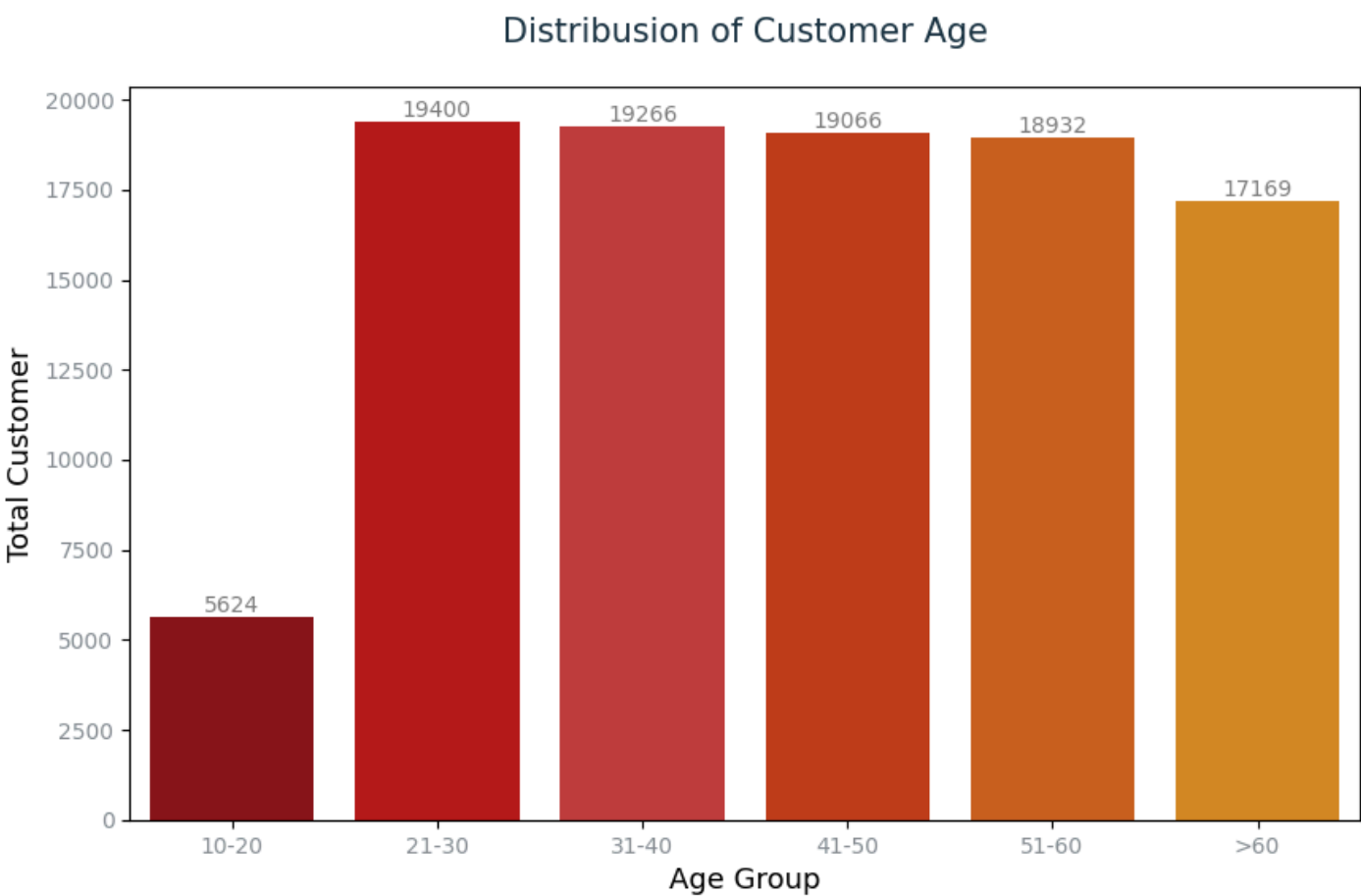The most purchased by customers is Clothing and the least purchased is Souvenir.

Top Shopping Mall

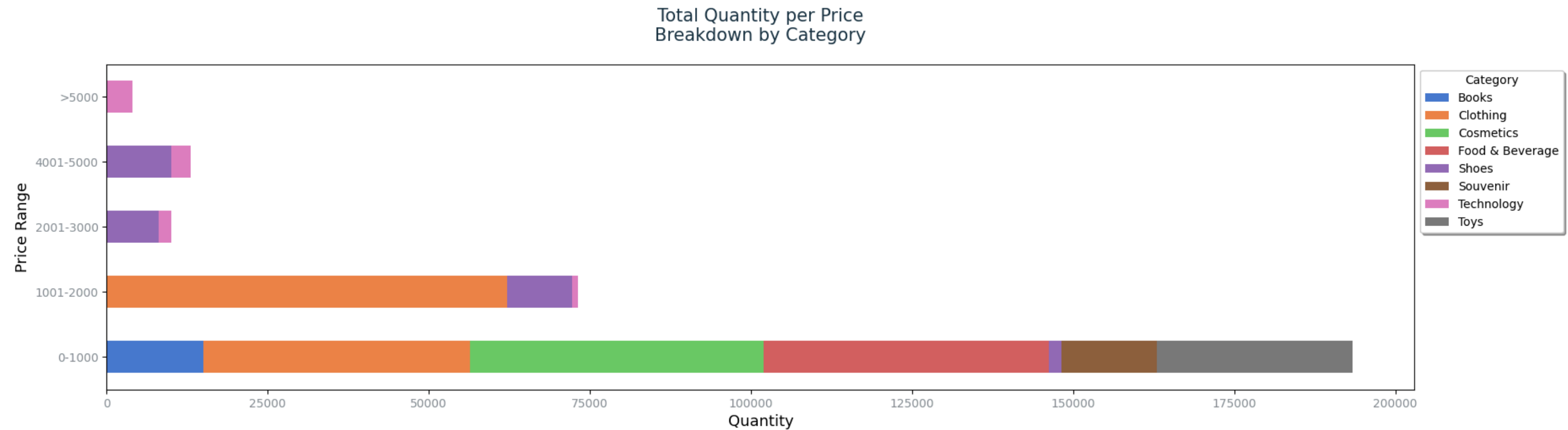Most visited by customer is Mall of Istanbul and the least visited is Emaar Square Mall.

Total Customer per Shopping Mall
Breakdown by Gender

Customer shopping is dominated by customers aged 21-30 years or teenagers. Customers aged 31-40 are in second place after teenagers which is only about 200 fewer.

The category most purchased by every age range is clothing

Distribusion of Customer Age

Total Customer per Age
Breakdown by Category

Total Quantity per Price Breakdown by Category

Total Purchased by Price
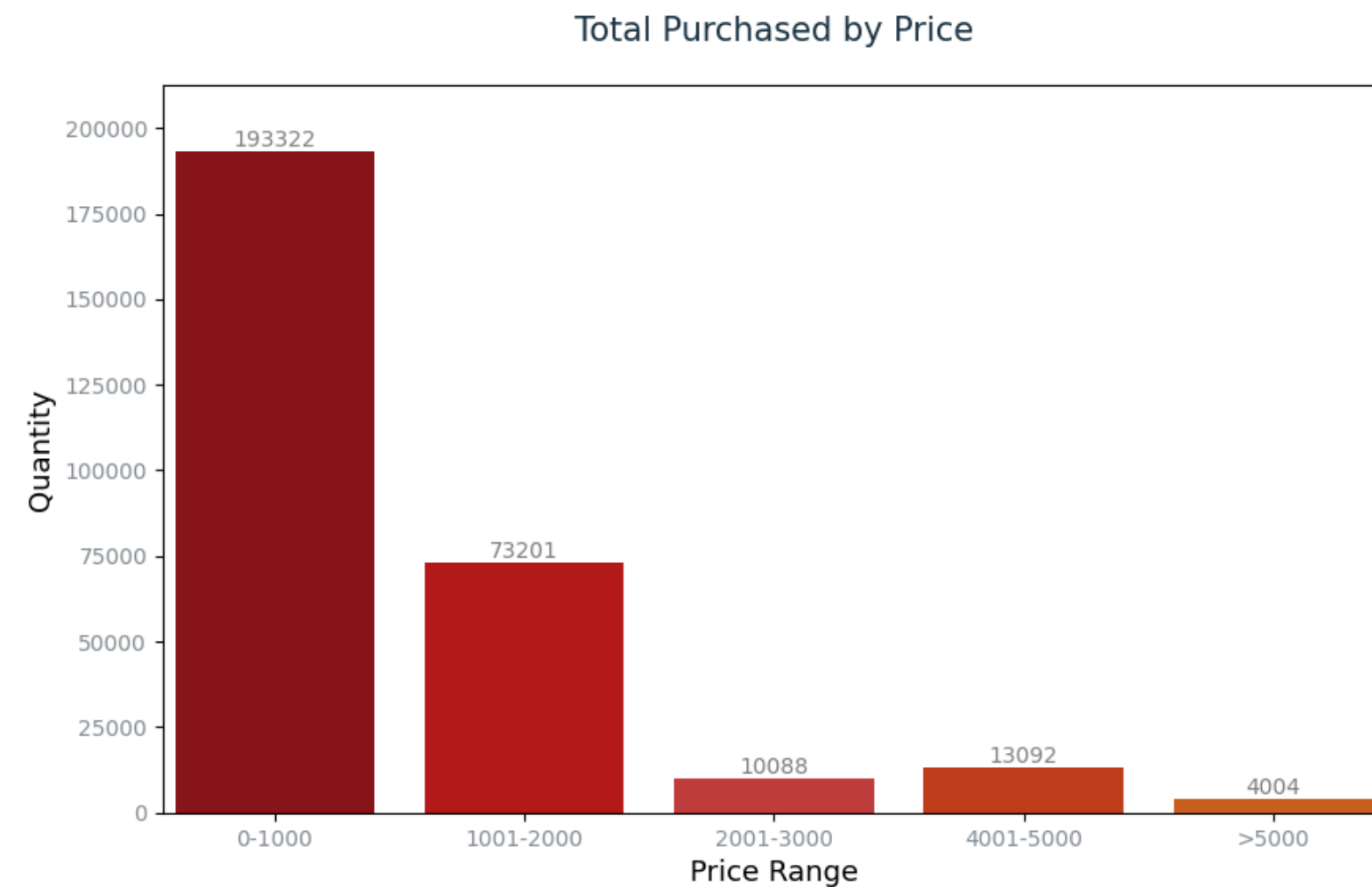
Customers buy more categories with a price range of 0-1000 TL and the most purchased category is clothing

# Customer Shopping Analysis in 2023

Top 5 Shopping Mall with the Most Customers

Top 5 Shopping Mall with the Highest Revenue
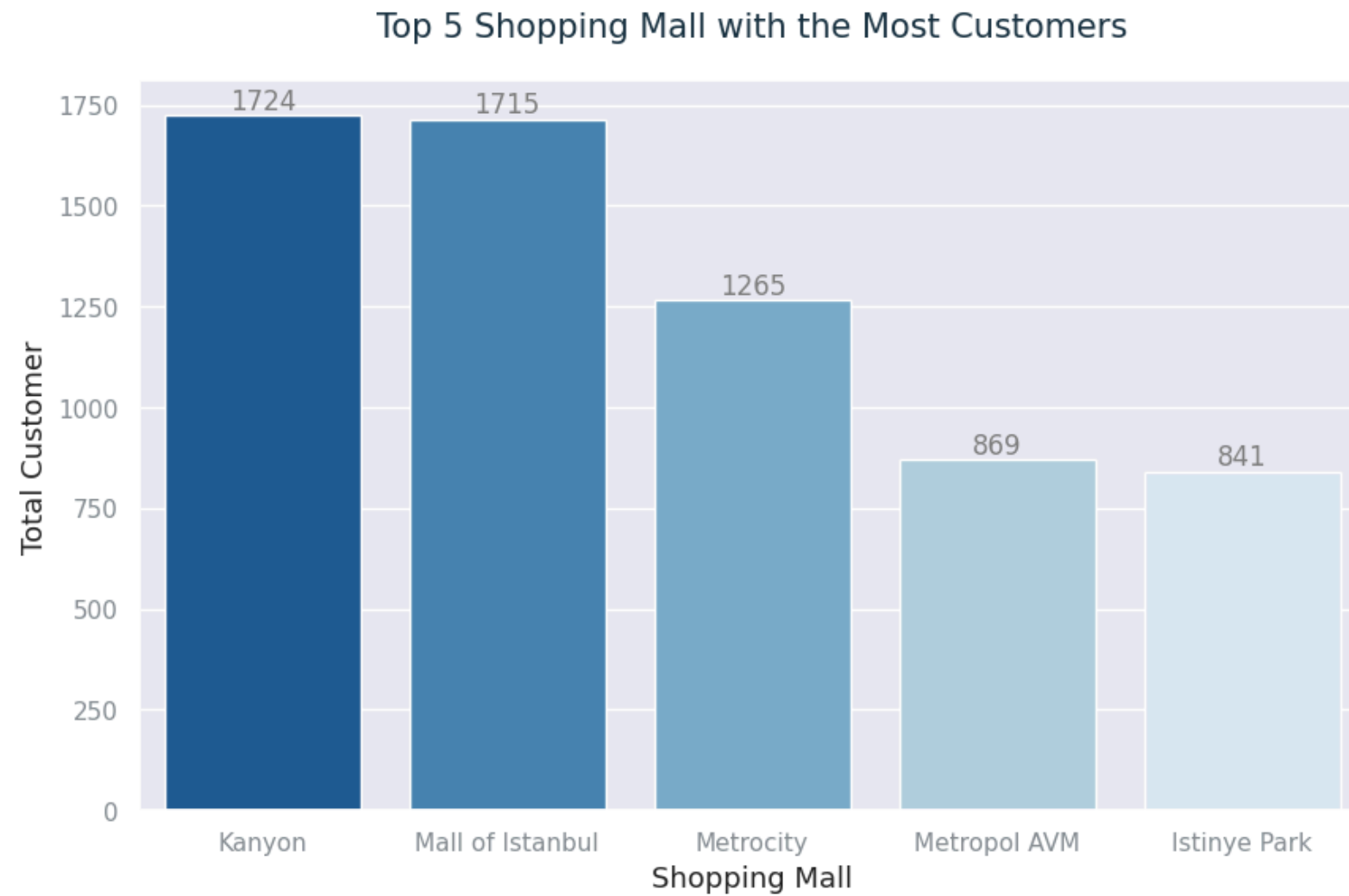
Charts show 5 shopping malls with the most of customers and the highest total revenue. In 2023, **Kanyon is the mall with the most number of customers** with a total revenue of more than 4 million TL.
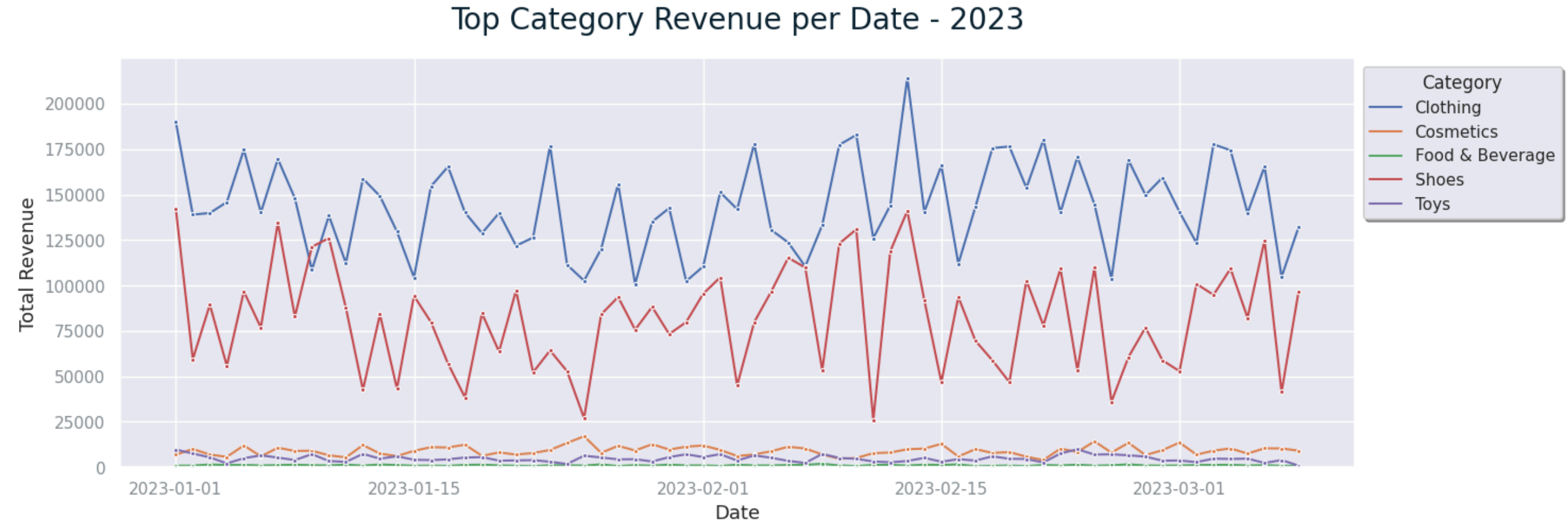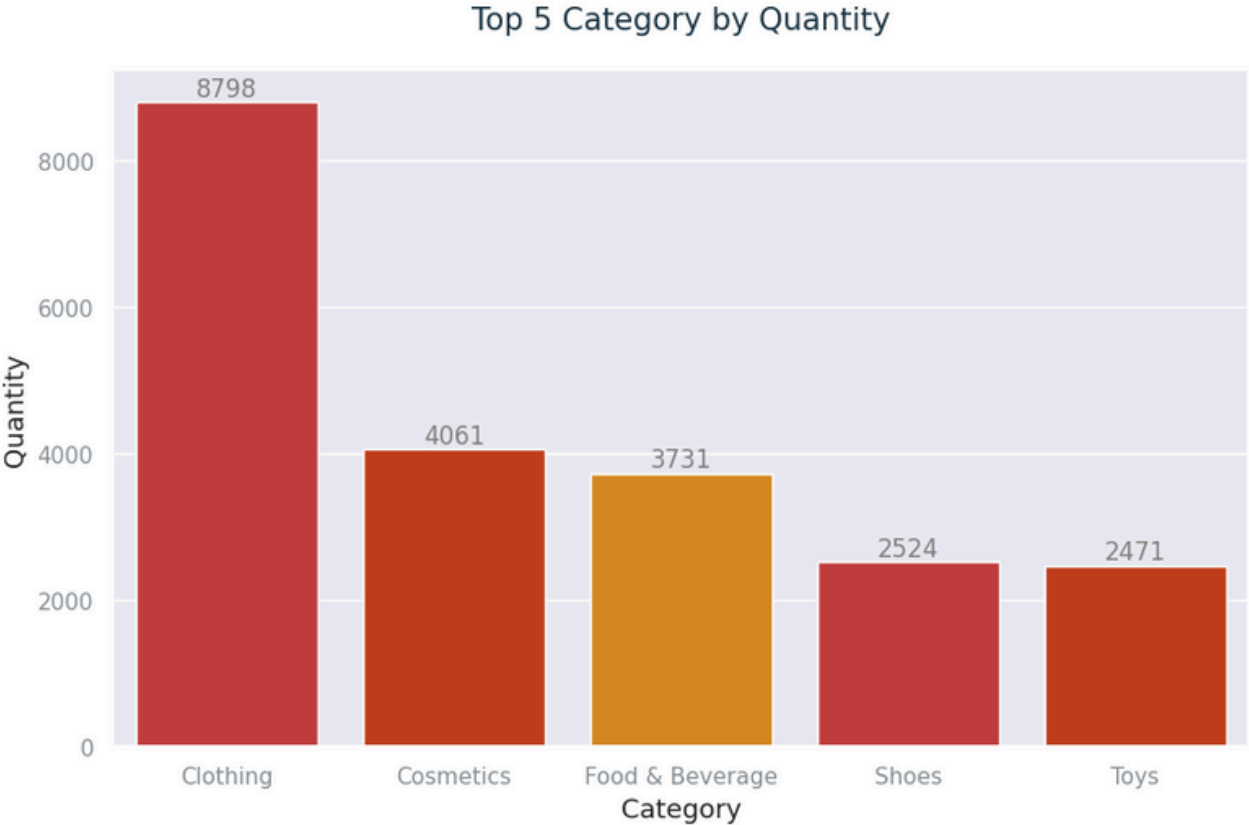
Total Customer by Weekday - 2023

Total Spending by Weekday - 2023
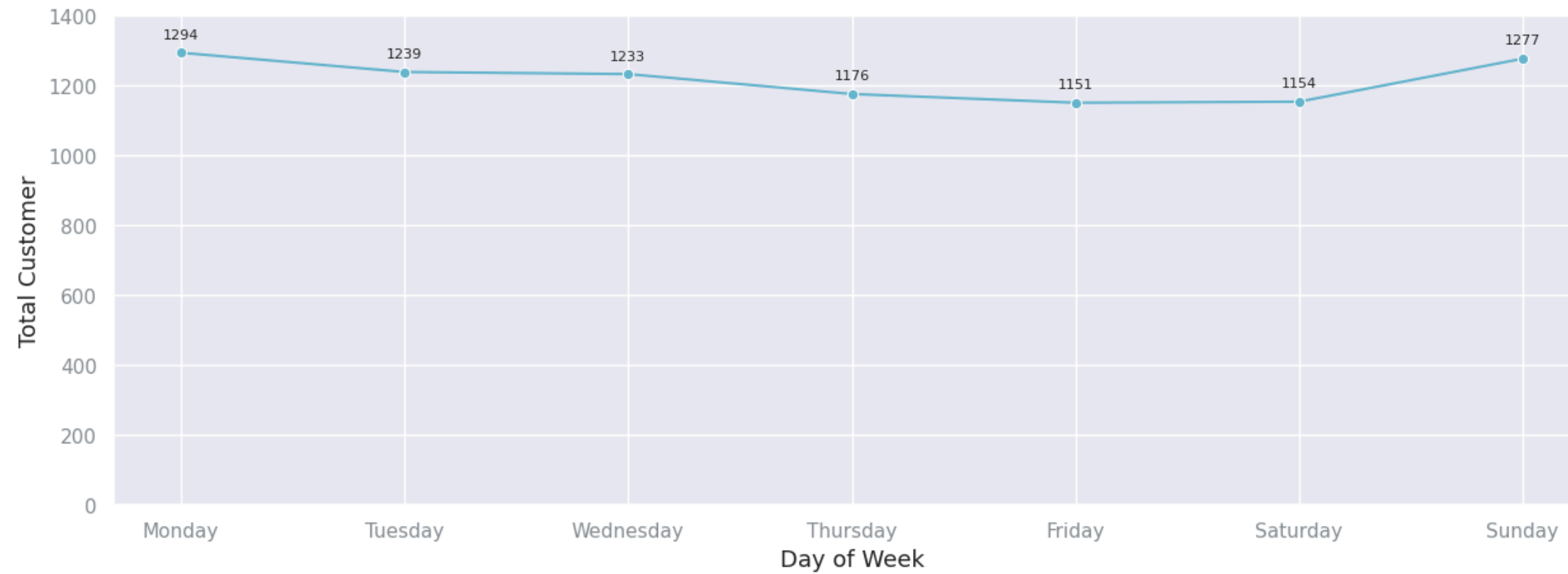
In 2023, Customers do more transactions and spend money to shop on **Monday.**
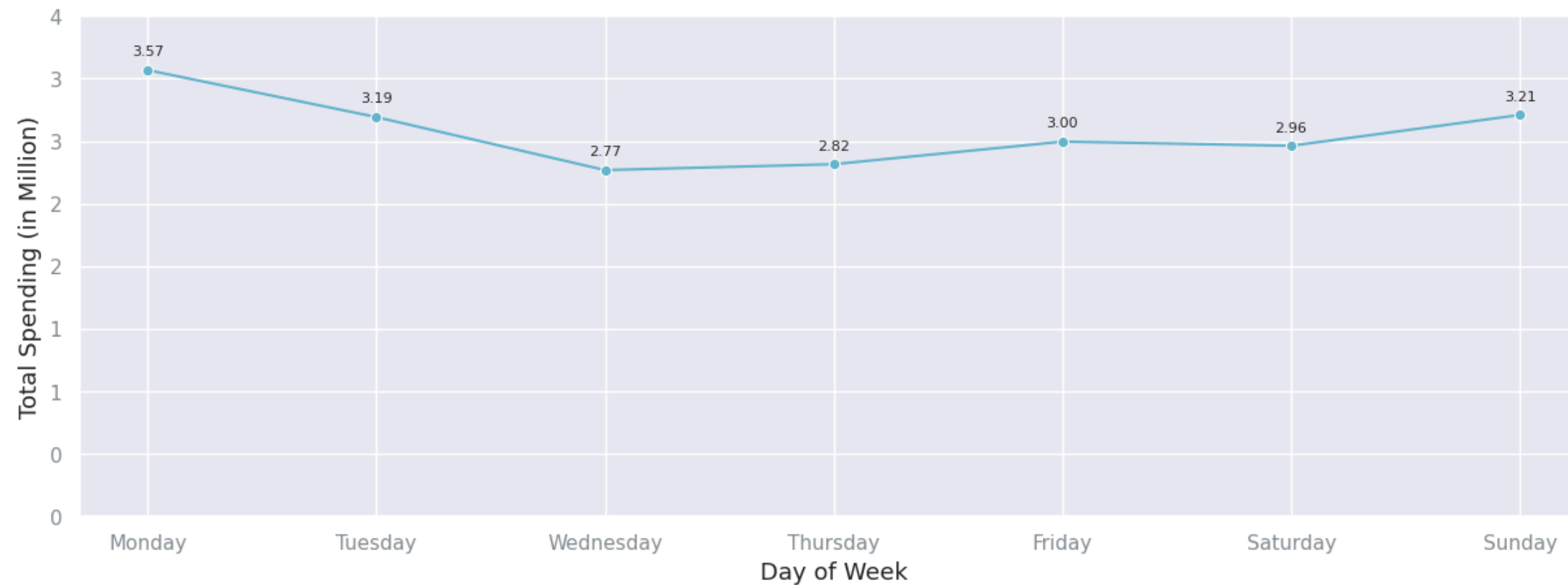
On Monday, there were 1294 transactions by customers with a total spending of 3.57 million TL.

Category with the most total purchased and the highest revenue in 2023.

Top 5 Category by Quantity

Top Category Revenue per Date - 2023

Recommendation for Business

Based on the visualization that has been done, there are several business strategy recommendations that can be used by companies/malls to increase sales.

Customers are dominated by female customers with an age range of 21-30 years old. Companies can attract the attention of these customers by adding the types of products that they usually buy.

Category with the highest price that is bought the most is technology. Companies can provide discounts or guarantees for every purchase of technology products to increase sales and maintain customer trust.

Thank you

Let's Connect

in **Chyntia Novrianti Kudadiri**   **Chyntia Kudadiri**   M **chyntiakudadiri14@gmail.com**