Data Analyst Portfolio Project

By: Ilham Prawira Hasbi

About

- This dataset is a snapshot of a fictional retail landscape, capturing essential attributes that drive retail operations and customer interactions. It includes key details such as Transaction ID, Date, Customer ID, Gender, Age, Product Category, Quantity, Price per Unit, and Total Amount. These attributes enable a multifaceted exploration of sales trends, demographic influences, and purchasing behaviors.

Business Objective

- How does customer age and gender influence their purchasing behavior?

- Are there discernible patterns in sales across different time periods?

- Which product categories hold the highest appeal among customers?

- What are the relationships between age, spending, and product preferences?

- How do customers adapt their shopping habits during seasonal trends?

- Are there distinct purchasing behaviors based on the number of items bought per transaction?

- What insights can be gleaned from the distribution of product prices within each category?

TOOLS

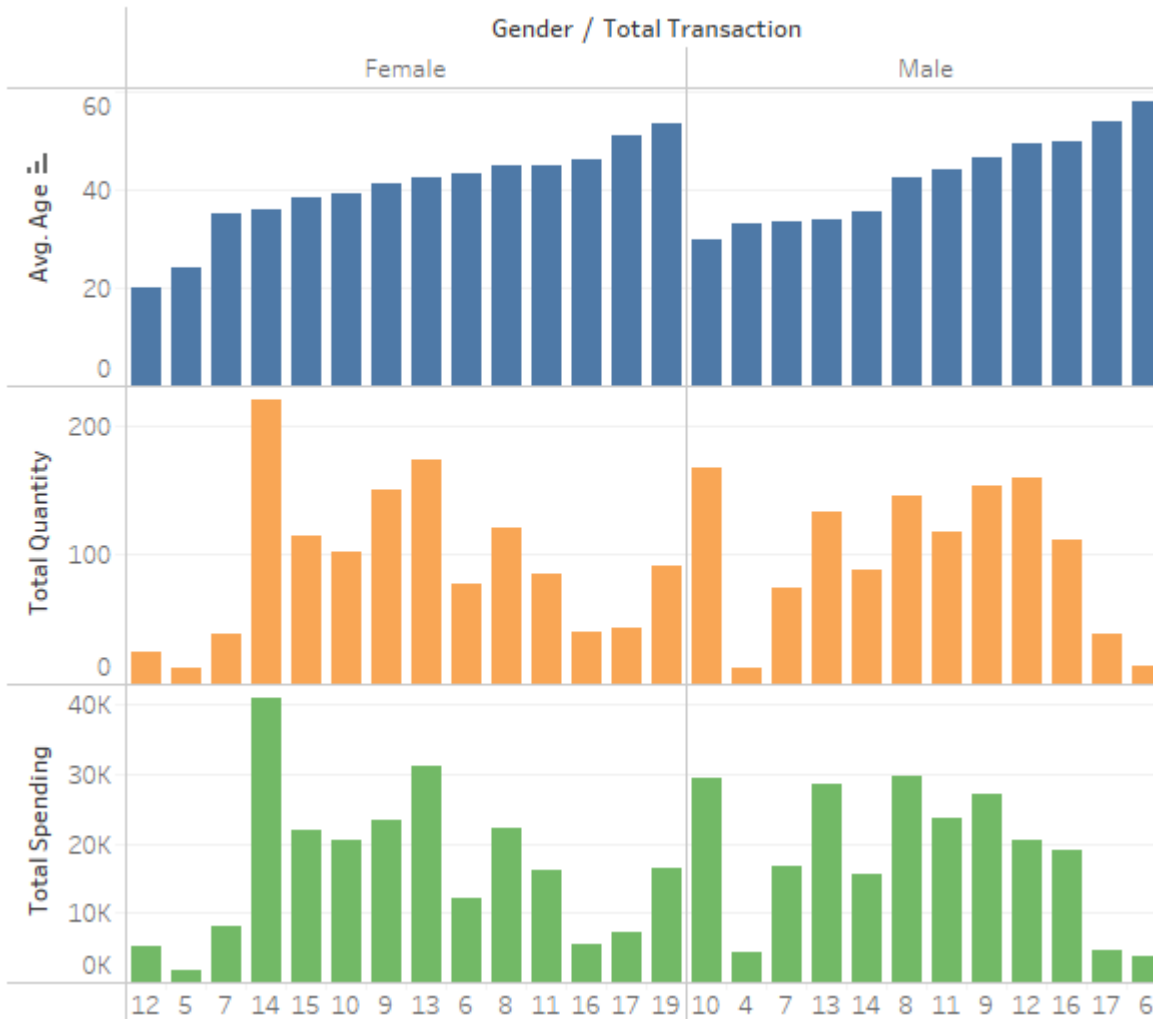Data Preparation

DBMS, Analysis

Dashboard visualization

Exploratory Data Analysis

Influence of customer age and gender on purchasing behavior

Influence of customer age and gender on purchasing behavior

- Query:

```sql
select DISTINCT Gender, Age, count(Transaction_ID) as Total_Transaction,
sum(Quantity) AS Total_Quantity,
sum(Total_Amount) AS Total_Spending
from sp.retail_sales_dataset
group by Gender, Age;
```

According to the graph, purchasing behavior is not influenced by age and gender. Female aged 36 have the highest total quantity and total spend compared to other female age groups. Male aged 29 have the highest total quantity and male aged 42 have the highest total spend.

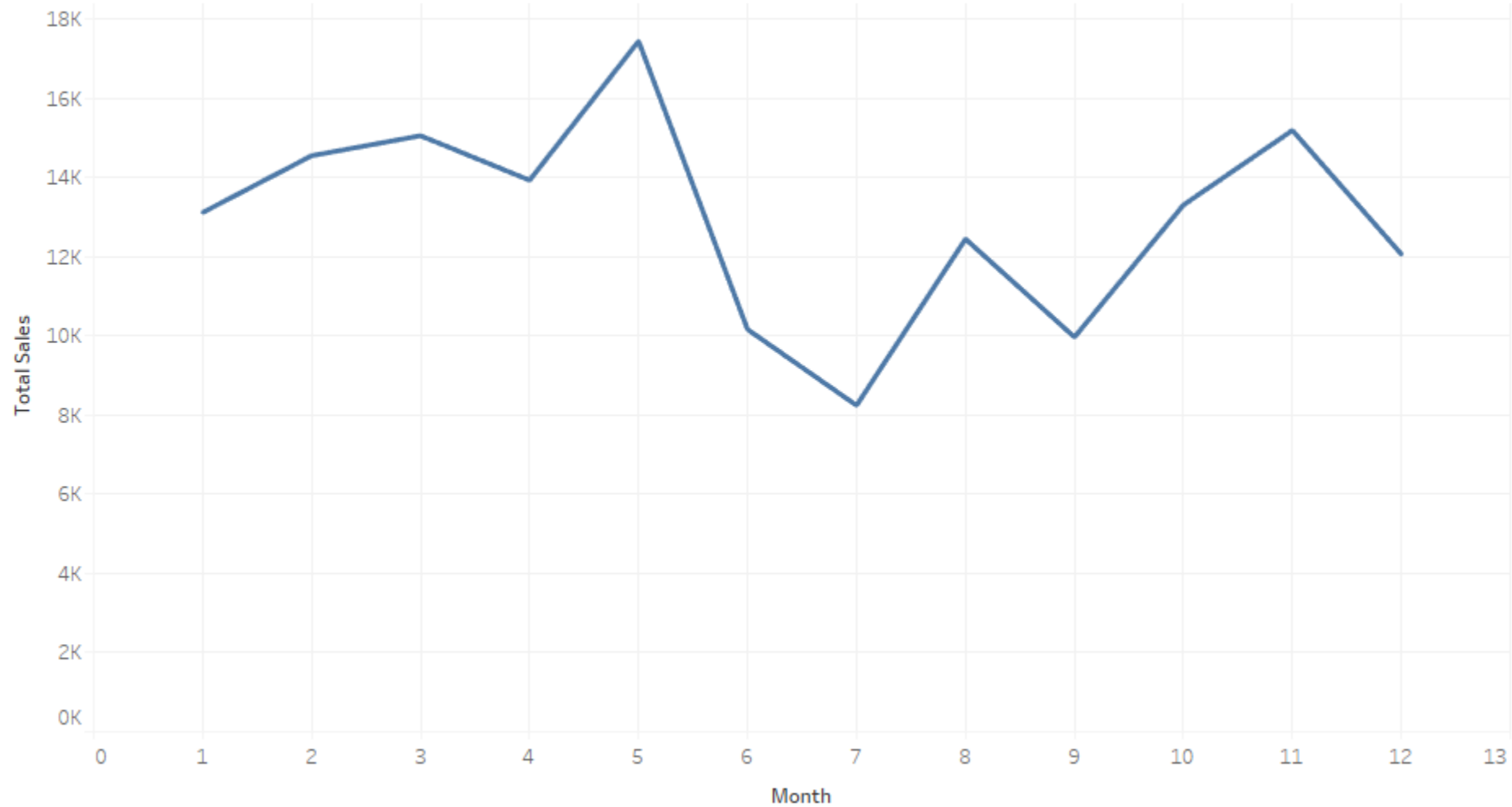# Sales patterns from 3 product category in different time period

- Query:

```sql
select month(Date) AS Month, SUM(Total_Amount) Total_Sales
from sp.retail_sales_dataset
WHERE Product_Category = 'Electronics'
GROUP BY month(Date)
ORDER BY month(Date);
```

```sql
select month(Date) AS Month, SUM(Total_Amount) Total_Sales
from sp.retail_sales_dataset
WHERE Product_Category = 'Beauty'
GROUP BY month(Date)
ORDER BY month(Date);
```

```sql
select month(Date) AS Month, SUM(Total_Amount) Total_Sales
from sp.retail_sales_dataset
WHERE Product_Category = 'Clothing'
GROUP BY month(Date)
ORDER BY month(Date);
```

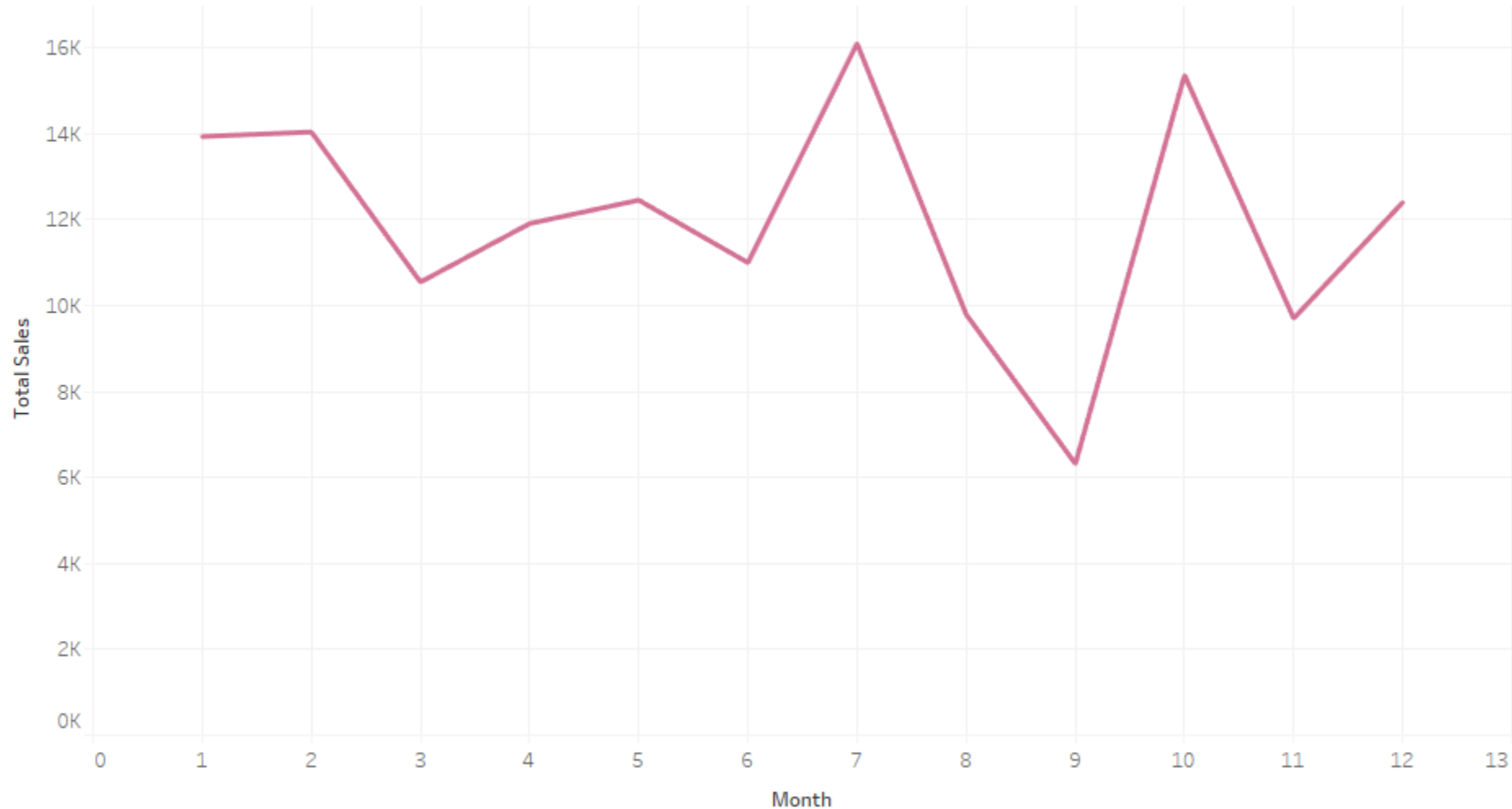# Sales patterns from 3 product category in different time period
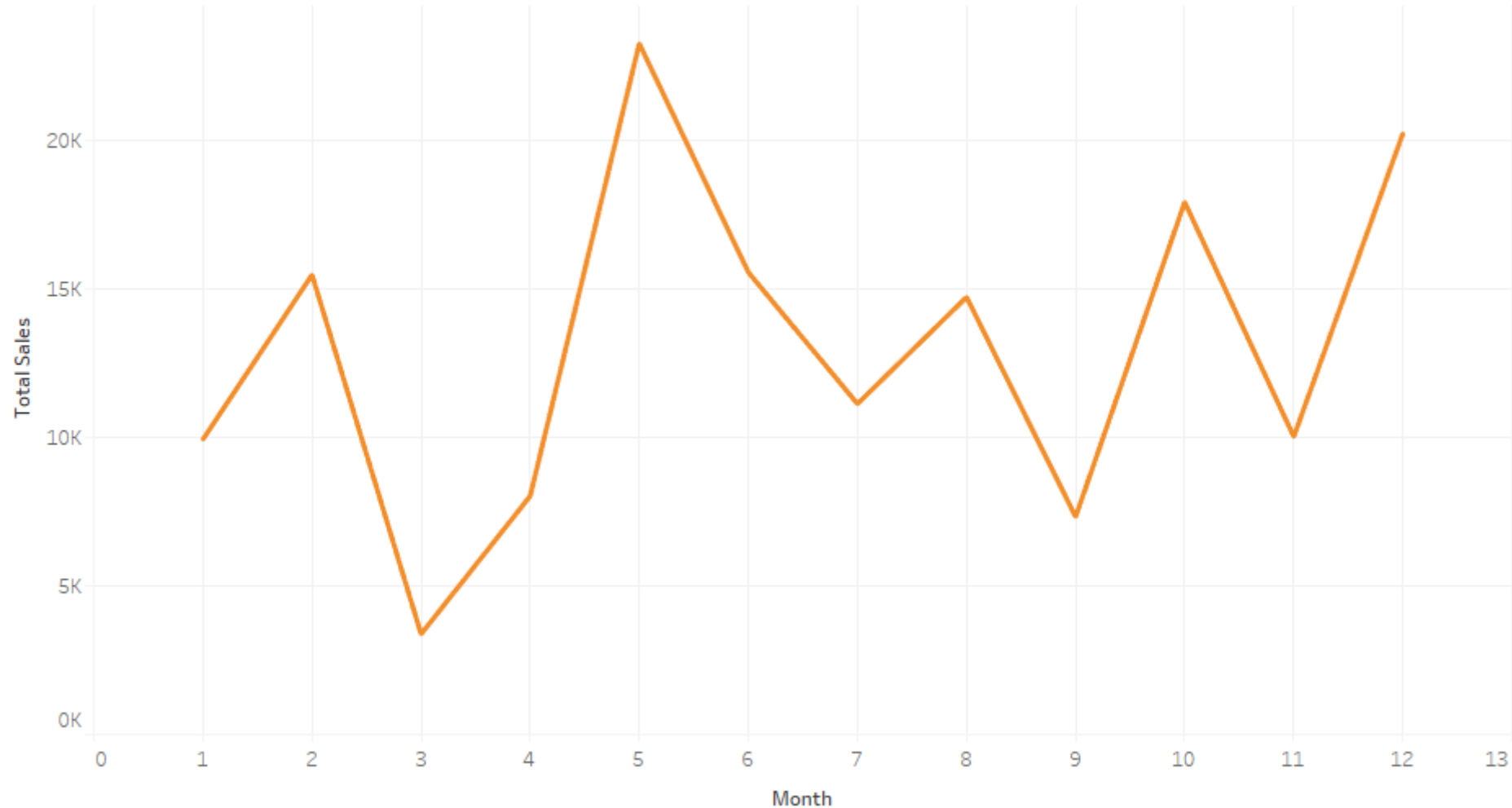

Clothing sales patterns in different time periods

Sales patterns from 3 product category in different time period

Beauty product sales patterns in different time periods

Sales patterns from 3 product category in different time period

Electronic product sales patterns in different time periods

Sales patterns from 3 product category in different time period

The highest sales of the electronics and clothing product categories were in the fifth month, while in the beauty product category in the seventh month. All three product categories experienced a decline in month 9.

# Product categories hold the highest appeal among customers

• Query:

```sql
select Count(Customer_ID) AS Customer_count, Product_Category,
SUM(Total_Amount) AS Total_Sales
from sp.retail_sales_dataset
group by Product_Category
order by Total_Sales DESC;
```

Customer_count	Product_Category	Total_Sales
342	Electronics	156905
351	Clothing	155580
307	Beauty	143515

Electronic products are the ones that customers are most interested in.

Product categories hold the highest appeal among customers

Product categories hold the highest appeal among customers
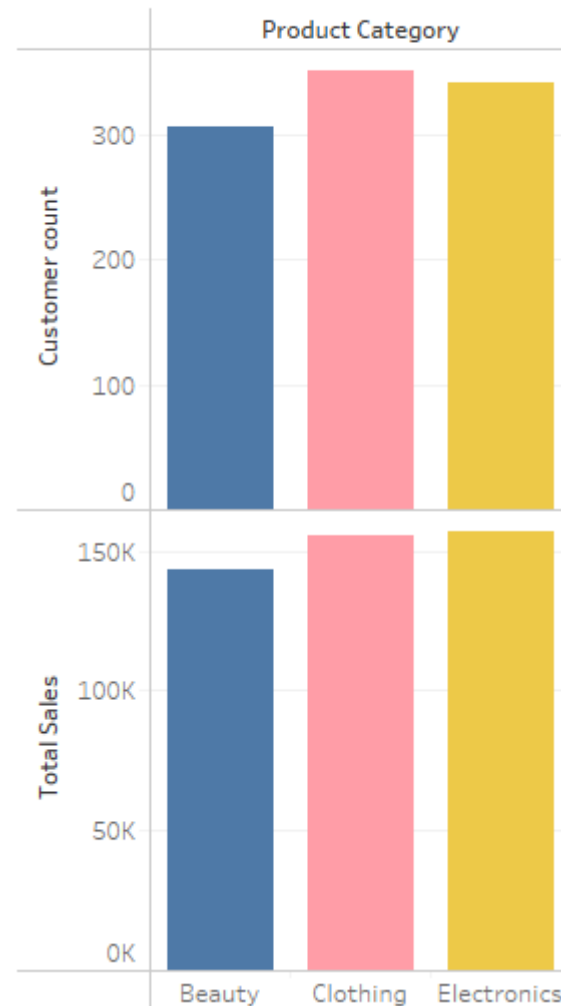
# Relationships between age, spending, and product preferences

- Query:

```sql
select DISTINCT SUM(Total_Amount) as Total_Sales,
Product_Category, ROUND(AVG(Age),0) Age, COUNT(Transaction_ID) Total_transactions
from sp.retail_sales_dataset
group by Product_Category
order by Total_Sales DESC;
```

Total_Sales	Product_Category	Age	Total_transactions
156905	Electronics	42	342
155580	Clothing	42	351
143515	Beauty	40	307

Age, spending, and product preference have no relationship with each other. Total sales are highest in the electronic product category but total transactions are highest in the clothing product category.

Relationships between age, spending, and product preferences
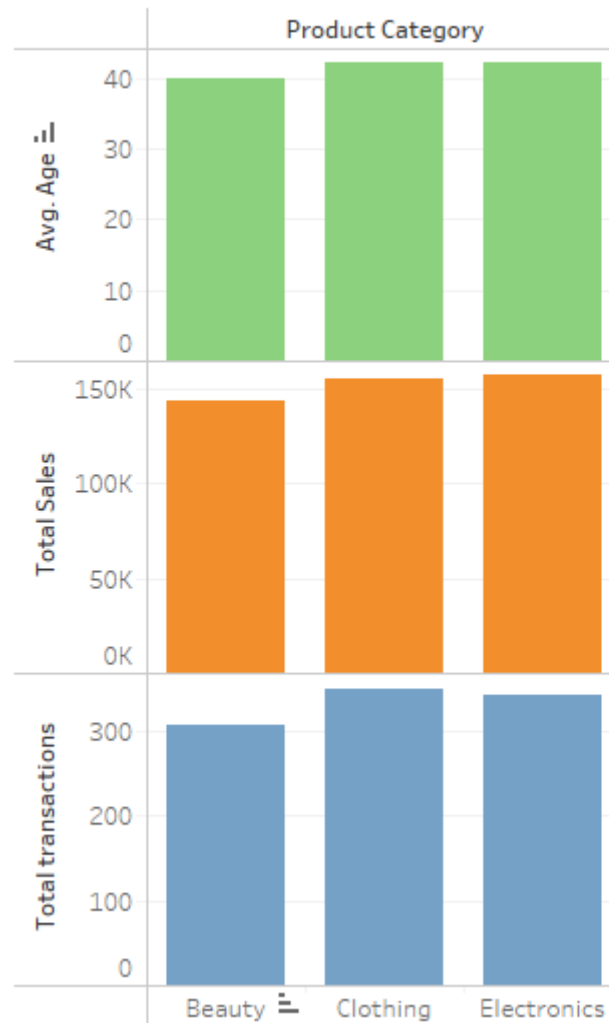
Customers shopping habits during seasonal trends

- Query:

```sql
select month(Date) as transaction_month, Product_Category,
Sum(Total_Amount) as total_sales, COUNT(Transaction_ID) Total_Transaction
from sp.retail_sales_dataset
group by transaction_month, Product_Category
order by transaction_month;
```

- Consumer spending in the clothing and electronics product categories increased in months 2, 5, 8, 10, with the highest sales in month 5.

- Consumer spending in the three categories (clothing, electronics, beauty) all increased in months 2, 5, 10.

- The highest sales in the beauty category were in month 9.

transaction_month	Product_Category	total_sales	Total_Transaction
1	Beauty	13930	26
1	Clothing	13125	26
1	Electronics	9925	26
2	Beauty	14035	26
2	Clothing	14560	33
2	Electronics	15465	26
3	Beauty	10545	21
3	Clothing	15065	38
3	Electronics	3380	14
4	Beauty	11905	29
4	Clothing	13940	36
4	Electronics	8025	21
5	Beauty	12450	28
5	Clothing	17455	37
5	Electronics	23245	40
6	Beauty	10995	25
6	Clothing	10170	28
6	Electronics	15550	24
7	Beauty	16090	27
7	Clothing	8250	19
7	Electronics	11125	26
8	Beauty	9790	24

8	Clothing	12455	32
8	Electronics	14715	38
9	Beauty	6320	20
9	Clothing	9975	20
9	Electronics	7325	25
10	Beauty	15355	31
10	Clothing	13315	30
10	Electronics	17910	35
11	Beauty	9700	25
11	Clothing	15200	26
11	Electronics	10020	27
12	Beauty	12400	25
12	Clothing	12070	26
12	Electronics	20220	40

Customers shopping habits during seasonal trends
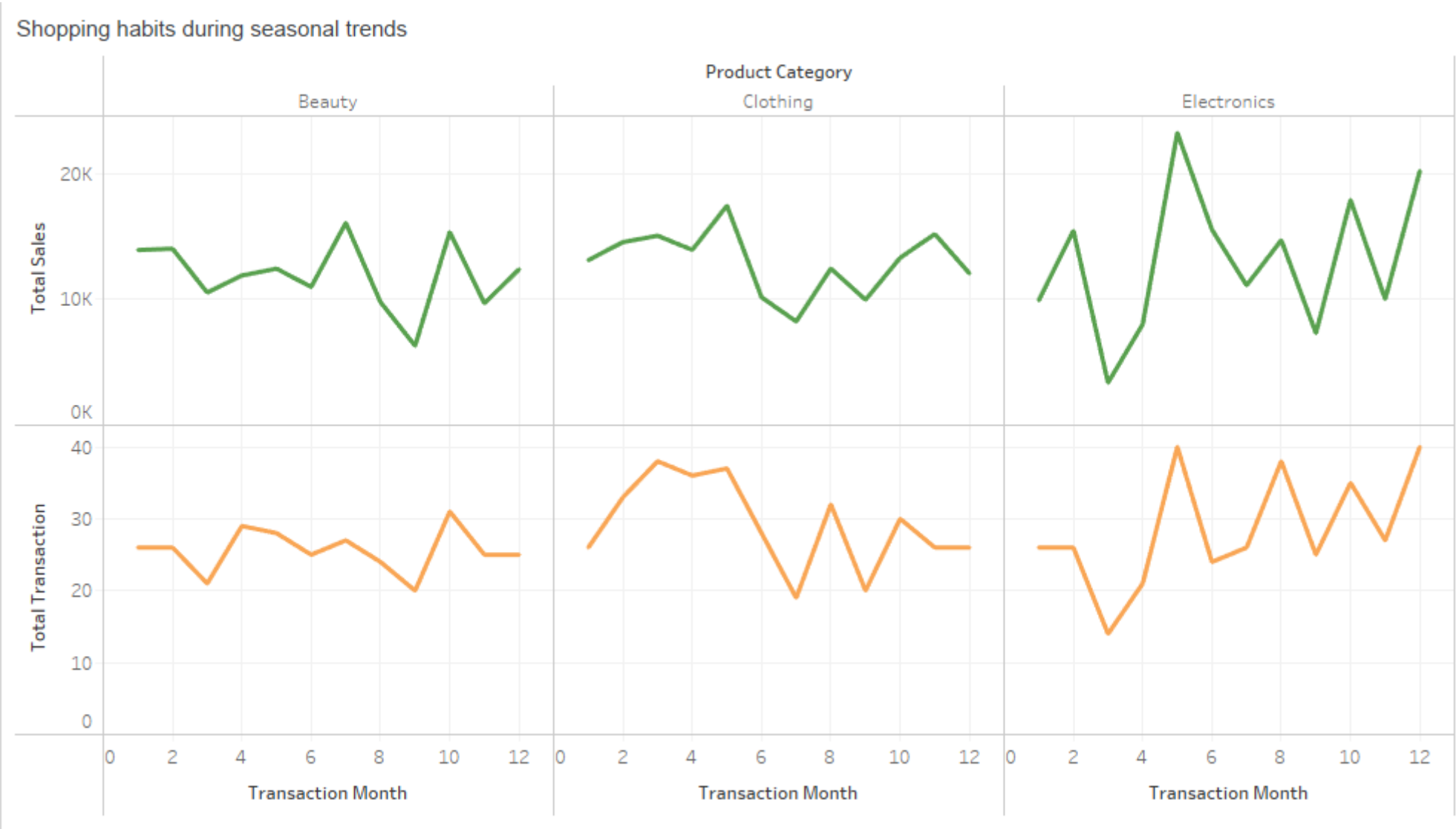
# Purchasing behaviors based on the number of items bought per transaction

- Query:

```sql
select Product_Category, Quantity, COUNT(Transaction_ID) Total_transactions,
ROUND(AVG(Total_Amount),2) Total_Sales
from sp.retail_sales_dataset
group by Product_Category, Quantity
```

Product_Category	Quantity	Total_transactions	Total_Sales
Beauty	3	85	613.59
Clothing	2	80	308.38
Electronics	1	91	186.37
Clothing	1	88	183.13
Beauty	2	75	399.07
Beauty	1	74	158.51
Electronics	4	93	749.25
Electronics	2	88	300.57
Clothing	4	97	685.36
Electronics	3	70	625.93
Clothing	3	86	561.80
Beauty	4	73	680.82

Purchasing preferences based on number of items purchased are mostly electronics and clothing based on quantity, total sales and total transactions compared to the beauty category.

Purchasing behaviors based on the number of items bought per transaction
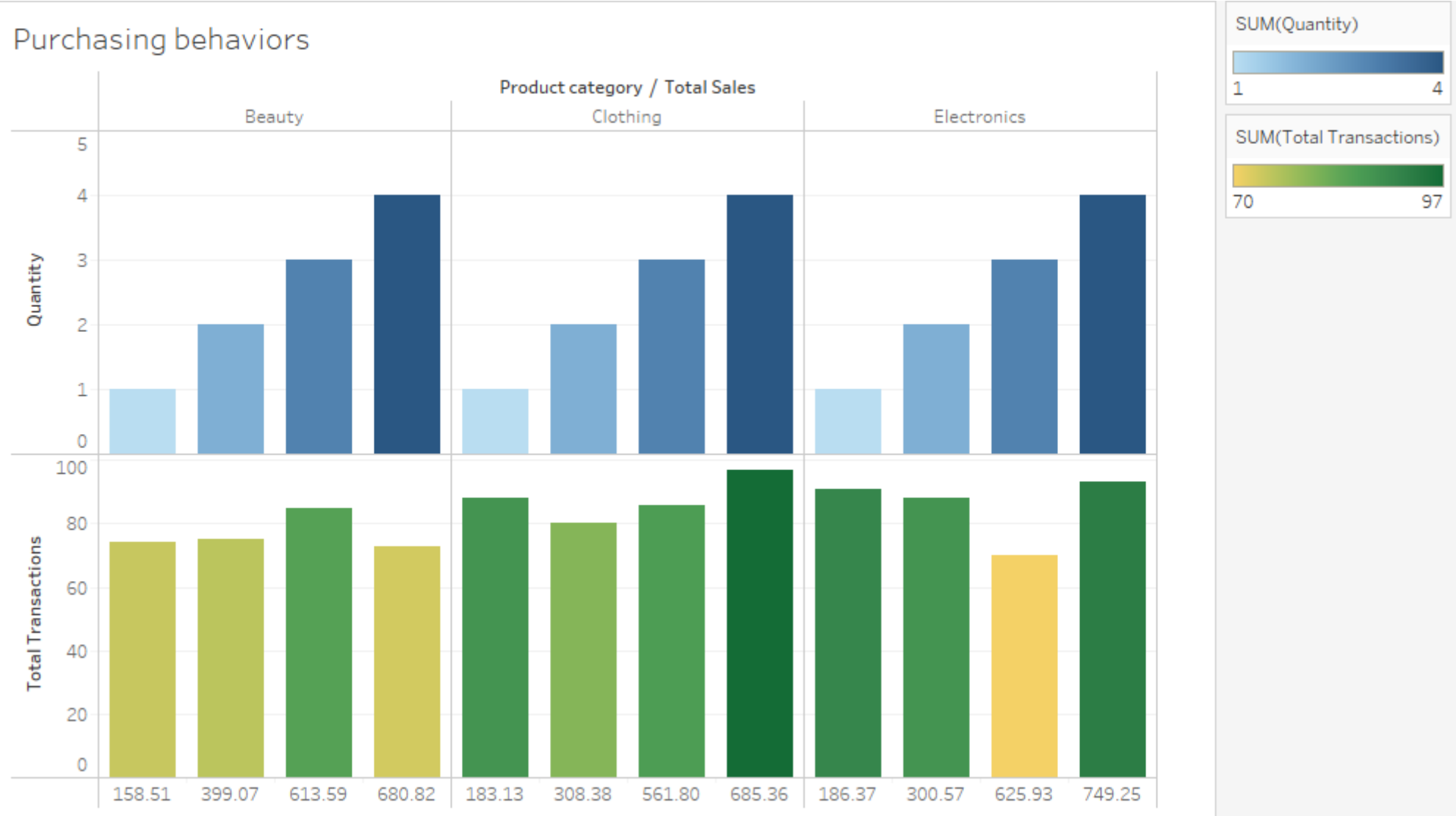
Distribution of product prices within each category

- Query:

```sql
SELECT Product_Category,
MIN(Price_per_Unit) AS Min_Price,
MAX(Price_per_Unit) AS Max_Price,
ROUND(AVG(Price_per_Unit),2) AS Average_Price
FROM sp.retail_sales_dataset
GROUP BY Product_Category;
```

Distribution of product prices

Average Price	Product Category	Maximum Price	Minimum Price
174.29	Clothing	500.0	25.0
181.90	Electronics	500.0	25.0
184.06	Beauty	500.0	25.0

The highest average price is in the beauty product category, while the lowest average price in the clothing product category.

Conclusions

- Age and gender do not influence shopping behavior.

- There is no relationship between age, spending, and product preference.

- Sales of the three product categories increase in months 2, 5, and 10, so it is necessary to increase inventory and promotions based on these time periods. Sales of the three product categories decrease in month 9, so promotions or discounts need to be based on that time period.

- The product that consumers are most interested in is electronics, so it is necessary to increase inventory and promotions.