

Supermarket Sales Analysis

Data Analysis Portfolio by Fara Rizkhi Karunia



Tools Used :

- Python for data cleaning process
- SQL for analysis and querying the data
- Microsoft Power BI for visualizing the data

Objective Statement

The objective of this supermarket data analysis portfolio is to extract valuable insights from each branch. These insights will subsequently inform strategic decision-making processes, such as optimizing stock replenishment quantities and optimizing employee work shifts. The portfolio includes a comprehensive dashboard presenting monthly sales reports from the three supermarket branches over the past three months.

- Retrieve data on top-performing and predominant gender associated with specific product categories to optimize stock replenishment quantities.
- Collect peak hour and busiest days data to optimize cashier staffing levels during specific shifts.
- Identify the best performing branches by considering several aspects of assessment to reward employees.
- Generate a comprehensive dashboard featuring sales reports for the past three months.

Datasets

Supermarket Sales (Source : [Kaggle](#))

This dataset contains the historical sales information of a supermarket company recorded across three different branches for the period of three months (January-March 2019). It comprises 1,000 rows representing individual orders and includes 17 columns of relevant information.



Data Dictionary

Invoice ID	Computer generated sales slip invoice identification number
Branch	Branch of supercenter (3 branches identified by A, B and C)
City	Location of supercenter
Customer type	Type of Customers identified by member and normal.
Gender	Gender type of customer
Product Line	General item categorization groups
Unit price	Price of each product in \$
Quantity	Number of products purchased by customer
Tax	5% tax fee for customer buying

Data Dictionary

Total	Total price including tax
Date	Date of purchase (January - March 2019)
Time	Purchase time (10 am to 9 pm)
Payment	Payment used by customer for purchase (Cash, CC, and Ewallet)
COGS	Cost of goods sold
GMP	Gross margin percentage
Gross income	Gross income
Rating	Customer stratification rating on their overall shopping experience (On a scale of 1 to 10)

Business Questions

- What are the top-performing product categories in terms of sales and customer demand?
- Are there gender-based preferences for certain product categories?
- What are the peak hours of customer activity?



Assesing Data

Tools : Python



Data types checking

```
data.info()
```

Result :

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 17 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Invoice ID      1000 non-null    object  
 1   Branch          1000 non-null    object  
 2   City            1000 non-null    object  
 3   Customer type   1000 non-null    object  
 4   Gender          1000 non-null    object  
 5   Product line    1000 non-null    object  
 6   Unit price     1000 non-null    float64 
 7   Quantity        1000 non-null    int64   
 8   Tax 5%          1000 non-null    float64 
 9   Total           1000 non-null    float64 
 10  Date            1000 non-null    object  
 11  Time            1000 non-null    object  
 12  Payment         1000 non-null    object  
 13  cogs            1000 non-null    float64 
 14  gross margin percentage 1000 non-null    float64 
 15  gross income    1000 non-null    float64 
 16  Rating          1000 non-null    float64 
dtypes: float64(7), int64(1), object(9)
memory usage: 132.9+ KB
```

Missing values & duplicated data

```
data.isnull().sum()
```

```
data.duplicated().sum()
```

Result :

```
Invoice ID              0
Branch                 0
City                   0
Customer type          0
Gender                 0
Product line            0
Unit price              0
Quantity                0
Tax 5%                  0
Total                   0
Date                   0
Time                   0
Payment                 0
cogs                    0
gross margin percentage 0
gross income             0
Rating                  0
dtype: int64
```

Out[6]: 0

Data Cleaning

Tools : Python



In the process of assessing the data, we found that the dataset is quite clean. There are no missing values, duplicate entries, or statistical parameter errors in the data. The only issue we encountered was related to the data type in the date and time columns, which we promptly adjusted. Additionally, we refined the column names in the dataset to make them more representative, organized, and suitable for easy querying in SQL.

Data types for the date and time columns

```
data['Date']=pd.to_datetime(data['Date'])  
  
data['Time']=data['Time']+':00'
```

Result :

```
0    13:08:00  
1    10:29:00  
2    13:23:00  
3    20:33:00  
4    10:37:00
```

Name: Time, dtype: object

Rename some columns name

```
data = data.rename(columns={  
    'Invoice ID':'Invoice_ID',  
    'Customer type':'Customer_Type',  
    'Product line':'Product_Line',  
    'Unit price':'Unit_Price',  
    'Tax 5%':'Tax_5', 'Total':'Revenue',  
    'cogs':'COGS', 'gross margin percentage':'GMP',  
    'gross income':'Gross_Income'})
```

Result :

	Invoice_ID	Branch	City	Customer_Type	Gender	Product_Line
0	750-67-8428	A	Yangon	Member	Female	Health and beauty

Retrieving Data insights

Tools : SQL



Optimizing Stock Replenishment Quantities

In this section, I will extract key data insights crucial for addressing the optimization of restocking quantities. Among these insights, I will focus on identifying the best-selling product categories to determine the optimal restocking levels. Additionally, a gender preference analysis will be conducted across specific product categories. This analysis is essential in guiding decisions on the quantity of items to be restocked. For instance, in the health & beauty category, it has been observed that items with higher marketability are predominantly preferred by the female demographic. Consequently, the branch will prioritize restocking more female-oriented products than male-oriented ones.



Optimizing Stock Replenishment Quantities

- Retrieving the best-selling product lines

```
--- BRANCH A
-- Best Selling Product Line
SELECT branch, product_line, SUM(quantity) as sales
FROM supermarket_sales_clear_csv ssc
where branch = 'A'
group by 1, 2
order by 3 DESC;
```

	branch	product_line	sales
1	A	Home and lifestyle	371
2	A	Sports and travel	333
3	A	Electronic accessories	322
4	A	Food and beverages	313
5	A	Fashion accessories	263
6	A	Health and beauty	257

Branch : A

- Retrieving gender preferences for each product line

--- Gender Preference

```
SELECT branch, product_line, gender, count(gender) as jumlah
FROM supermarket_sales_clear_csv ssc
where branch = 'A'
GROUP BY 1,2,3
order by 1,2,4 DESC;
```

	branch	product_line	gender	jumlah
1	A	Electronic accessories	Male	32
2	A	Electronic accessories	Female	28
3	A	Fashion accessories	Female	28
4	A	Fashion accessories	Male	23
5	A	Food and beverages	Male	35
6	A	Food and beverages	Female	23
7	A	Health and beauty	Male	26
8	A	Health and beauty	Female	21
9	A	Home and lifestyle	Male	33
10	A	Home and lifestyle	Female	32
11	A	Sports and travel	Male	30
12	A	Sports and travel	Female	29

Optimizing Stock Replenishment Quantities

- Retrieving the best-selling product lines

```
--- BRANCH B
--- Best Selling Product Line
SELECT branch, product_line, SUM(quantity) as sales
FROM supermarket_sales_clear_csv ssc
where branch = 'B'
group by 1, 2
order by 3 DESC;
```

	branch	product_line	sales
1	B	Sports and travel	322
2	B	Health and beauty	320
3	B	Electronic accessori	316
4	B	Fashion accessorie	297
5	B	Home and lifestyle	295
6	B	Food and beverag	270

Branch : B

- Retrieving gender preferences for each product line

--- Gender Preference

```
SELECT branch, product_line, gender, count(gender) as jumlah
FROM supermarket_sales_clear_csv ssc
where branch = 'B'
GROUP BY 1,2,3
order by 1,2,4 DESC;
```

	branch	product_line	gender	jumlah
1	B	Electronic accessories	Female	28
2	B	Electronic accessories	Male	27
3	B	Fashion accessories	Female	33
4	B	Fashion accessories	Male	29
5	B	Food and beverages	Female	29
6	B	Food and beverages	Male	21
7	B	Health and beauty	Male	33
8	B	Health and beauty	Female	20
9	B	Home and lifestyle	Male	28
10	B	Home and lifestyle	Female	22
11	B	Sports and travel	Male	32
12	B	Sports and travel	Female	30

Optimizing Stock Replenishment Quantities

- Retrieving the best-selling product lines

```
--- BRANCH C  
-- Best Selling Product Line  
SELECT branch, product_line, SUM(quantity) as sales  
FROM supermarket_sales_clear_csv sscs  
where branch = 'C'  
group by 1, 2  
order by 3 DESC;
```

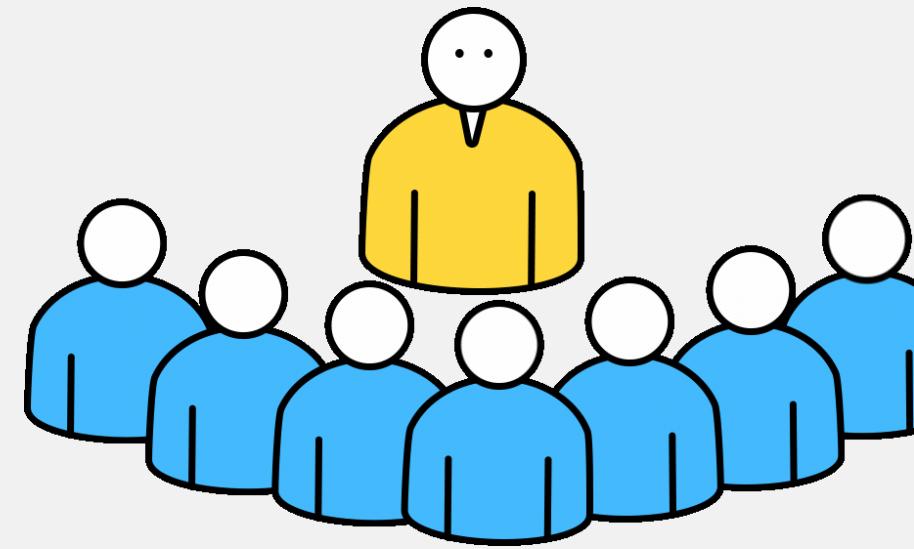
	branch	product_line	sales
1	C	Food and beverages	369
2	C	Fashion accessories	342
3	C	Electronic accessories	333
4	C	Health and beauty	277
5	C	Sports and travel	265
6	C	Home and lifestyle	245

- Branch : C
- Retrieving gender preferences for each product line

```
-- Gender Preference
```

```
SELECT branch, product_line, gender, count(gender) as jumlah  
FROM supermarket_sales_clear_csv sscs  
where branch = 'C'  
GROUP BY 1,2,3  
order by 1,2,4 DESC;
```

	branch	product_line	gender	jumlah
1	C	Electronic accessories	Female	28
2	C	Electronic accessories	Male	27
3	C	Fashion accessories	Female	35
4	C	Fashion accessories	Male	30
5	C	Food and beverages	Female	38
6	C	Food and beverages	Male	28
7	C	Health and beauty	Male	29
8	C	Health and beauty	Female	23
9	C	Home and lifestyle	Female	25
10	C	Home and lifestyle	Male	20
11	C	Sports and travel	Female	29
12	C	Sports and travel	Male	16



Optimizing Cashier Staffing Levels during Specific Shifts

In this section I will identify the busy hours and days in each supermarket branch. This is useful for the decision-making process in optimising the number of cashier staff on a particular shift based on the peak hours and days that have been obtained. The identification is done at each branch because usually each region will have quite different preferences, depending on the demographics of the surrounding population.

Peak Hours

Branch : A

- Collect peak hour data

```
--busy hour
SELECT pelabelan, COUNT(*) AS busy_hour
FROM (
    SELECT
        "Time",
        case
            WHEN CAST("Time" as TIME) BETWEEN '10:00:00' AND '11:00:00' THEN '10:00-11:00'
            WHEN CAST("Time" as TIME) BETWEEN '11:01:00' AND '12:00:00' THEN '11:01-12:00'
            WHEN CAST("Time" as TIME) BETWEEN '12:01:00' AND '13:00:00' THEN '12:01-13:00'
            WHEN CAST("Time" as TIME) BETWEEN '13:01:00' AND '14:00:00' THEN '13:01-14:00'
            WHEN CAST("Time" as TIME) BETWEEN '14:01:00' AND '15:00:00' THEN '14:01-15:00'
            WHEN CAST("Time" as TIME) BETWEEN '15:01:00' AND '16:00:00' THEN '15:01-16:00'
            WHEN CAST("Time" as TIME) BETWEEN '16:01:00' AND '17:00:00' THEN '16:01-17:00'
            WHEN CAST("Time" as TIME) BETWEEN '17:01:00' AND '18:00:00' THEN '17:01-18:00'
            WHEN CAST("Time" as TIME) BETWEEN '18:01:00' AND '19:00:00' THEN '18:01-19:00'
            WHEN CAST("Time" as TIME) BETWEEN '19:01:00' AND '20:00:00' THEN '19:01-20:00'
            WHEN CAST("Time" as TIME) BETWEEN '20:01:00' AND '21:00:00' THEN '20:01-21:00'
            WHEN CAST("Time" as TIME) BETWEEN '21:01:00' AND '22:00:00' THEN '21:01-22:00'
            ELSE 'Other'
        END AS pelabelan
    FROM supermarket_sales_clear_csv ssc
    where branch='A'
) AS labeled_data
GROUP BY 1
ORDER BY 2 DESC;
```

	pelabelan	busy_hour
1	10:00-11:00	40
2	15:01-16:00	37
3	12:01-13:00	33
4	11:01-12:00	33
5	13:01-14:00	32
6	18:01-19:00	32
7	16:01-17:00	32
8	17:01-18:00	28
9	19:01-20:00	27
10	14:01-15:00	24
11	20:01-21:00	22

Busiest Days

Branch : A

- Collecting data on crowded days

```
--BRANCH A  
-- crowded days  
SELECT TO_CHAR(CAST("Date" AS DATE), 'Day') AS Day, COUNT(TO_CHAR(CAST("Date" AS DATE), 'Day')) AS Transaction  
FROM supermarket_sales_clear_csv ssc  
where branch='A'  
GROUP BY 1  
order by 2 DESC;
```

	day	transaction
1	Sunday	52
2	Tuesday	51
3	Friday	50
4	Saturday	50
5	Monday	48
6	Thursday	46
7	Wednesday	43

Peak Hours

Branch : B

- Collect peak hour data

```
--busy hour
SELECT pelabelan, COUNT(*) AS busy_hour
FROM (
    SELECT
        "Time",
        case
            WHEN CAST("Time" as TIME) BETWEEN '10:00:00' AND '11:00:00' THEN '10:00-11:00'
            WHEN CAST("Time" as TIME) BETWEEN '11:01:00' AND '12:00:00' THEN '11:01-12:00'
            WHEN CAST("Time" as TIME) BETWEEN '12:01:00' AND '13:00:00' THEN '12:01-13:00'
            WHEN CAST("Time" as TIME) BETWEEN '13:01:00' AND '14:00:00' THEN '13:01-14:00'
            WHEN CAST("Time" as TIME) BETWEEN '14:01:00' AND '15:00:00' THEN '14:01-15:00'
            WHEN CAST("Time" as TIME) BETWEEN '15:01:00' AND '16:00:00' THEN '15:01-16:00'
            WHEN CAST("Time" as TIME) BETWEEN '16:01:00' AND '17:00:00' THEN '16:01-17:00'
            WHEN CAST("Time" as TIME) BETWEEN '17:01:00' AND '18:00:00' THEN '17:01-18:00'
            WHEN CAST("Time" as TIME) BETWEEN '18:01:00' AND '19:00:00' THEN '18:01-19:00'
            WHEN CAST("Time" as TIME) BETWEEN '19:01:00' AND '20:00:00' THEN '19:01-20:00'
            WHEN CAST("Time" as TIME) BETWEEN '20:01:00' AND '21:00:00' THEN '20:01-21:00'
            WHEN CAST("Time" as TIME) BETWEEN '21:01:00' AND '22:00:00' THEN '21:01-22:00'
            ELSE 'Other'
        END AS pelabelan
    FROM supermarket_sales_clear_csv ssc
    where branch='B'
) AS labeled_data
GROUP BY 1
ORDER BY 2 DESC;
```

	pelabelan	busy_hour
1	19:01-20:00	50
2	13:01-14:00	37
3	18:01-19:00	35
4	11:01-12:00	33
5	15:01-16:00	32
6	14:01-15:00	30
7	12:01-13:00	26
8	10:00-11:00	26
9	20:01-21:00	25
10	17:01-18:00	21
11	16:01-17:00	17

Busiest Days

Branch : B

- Collecting data on crowded days

```
--BRANCH B  
-- crowded days  
SELECT TO_CHAR(CAST("Date" AS DATE), 'Day') AS Day, COUNT(TO_CHAR(CAST("Date" AS DATE), 'Day')) AS Transaction  
FROM supermarket_sales_clear_csv ssc  
where branch='A'  
GROUP BY 1  
order by 2 DESC;
```

	day	transaction
1	Saturday	60
2	Tuesday	53
3	Friday	51
4	Wednesday	50
5	Thursday	44
6	Monday	39
7	Sunday	35

Peak Hours

Branch : C

- Collect peak hour data

```
--busy hour
SELECT pelabelan, COUNT(*) AS busy_hour
FROM (
    SELECT
        "Time",
        case
            WHEN CAST("Time" as TIME) BETWEEN '10:00:00' AND '11:00:00' THEN '10:00-11:00'
            WHEN CAST("Time" as TIME) BETWEEN '11:01:00' AND '12:00:00' THEN '11:01-12:00'
            WHEN CAST("Time" as TIME) BETWEEN '12:01:00' AND '13:00:00' THEN '12:01-13:00'
            WHEN CAST("Time" as TIME) BETWEEN '13:01:00' AND '14:00:00' THEN '13:01-14:00'
            WHEN CAST("Time" as TIME) BETWEEN '14:01:00' AND '15:00:00' THEN '14:01-15:00'
            WHEN CAST("Time" as TIME) BETWEEN '15:01:00' AND '16:00:00' THEN '15:01-16:00'
            WHEN CAST("Time" as TIME) BETWEEN '16:01:00' AND '17:00:00' THEN '16:01-17:00'
            WHEN CAST("Time" as TIME) BETWEEN '17:01:00' AND '18:00:00' THEN '17:01-18:00'
            WHEN CAST("Time" as TIME) BETWEEN '18:01:00' AND '19:00:00' THEN '18:01-19:00'
            WHEN CAST("Time" as TIME) BETWEEN '19:01:00' AND '20:00:00' THEN '19:01-20:00'
            WHEN CAST("Time" as TIME) BETWEEN '20:01:00' AND '21:00:00' THEN '20:01-21:00'
            WHEN CAST("Time" as TIME) BETWEEN '21:01:00' AND '22:00:00' THEN '21:01-22:00'
            ELSE 'Other'
        END AS pelabelan
    FROM supermarket_sales_clear_csv ssc
    where branch='C'
) AS labeled_data
GROUP BY 1
ORDER BY 2 DESC;
```

	pelabelan	busy_hour
1	10:00-11:00	37
2	19:01-20:00	36
3	12:01-13:00	33
4	13:01-14:00	33
5	15:01-16:00	32
6	16:01-17:00	28
7	14:01-15:00	28
8	17:01-18:00	27
9	20:01-21:00	27
10	18:01-19:00	25
11	11:01-12:00	22

Busiest Days

Branch : A

- Collecting data on crowded days

```
--BRANCH C  
-- crowded days  
SELECT TO_CHAR(CAST("Date" AS DATE), 'Day') AS Day, COUNT(TO_CHAR(CAST("Date" AS DATE), 'Day')) AS Transaction  
FROM supermarket_sales_clear_csv ssc  
where branch='C'  
GROUP BY 1  
order by 2 DESC;
```

	day	transaction
1	Tuesday	54
2	Saturday	54
3	Wednesday	50
4	Thursday	48
5	Sunday	46
6	Friday	38
7	Monday	38



Best Performing Branch

In this section, we will look for the best performing branches by considering several assessment aspects including number of transactions, total sales, total revenue, customer ratings, and number of members in each months. The best performing branch in each month will be rewarded. The reward can be in the form of performance bonuses, shopping vouchers for employees, or certificates of appreciation. Giving rewards to the best performing branches is expected to increase employee motivation, stimulate employee innovation and creativity in increasing sales, and reward employee performance to retain employees with superior performance in the company.

Best Performing Branch on January

- We consider five aspects in determining the best supermarket branch on January: the number of transactions, total sales, overall revenue, customer ratings, and the number of members.

```
--Best Branch
SELECT
    branch,
    TO_CHAR(CAST("Date" AS DATE), 'Month') AS Month,
    count(invoice_id) as number_of_transaction,
    sum(quantity) as total_sales,
    sum(revenue) as total_revenue,
    avg(rating) as Avg_rating,
    count(customer_type='Member') as member
FROM
    supermarket_sales_clear_csv sscs
where
    EXTRACT(MONTH FROM CAST("Date" AS DATE)) = 1
GROUP BY 1,2
ORDER by 3 DESC;
```

	branch	month	number_of_transaction	total_sales	total_revenue	avg_rating	member
1	C	January	122	680	40,434.67578125	7.154918057	122
2	A	January	119	685	38,681.140625	7.0781512461	119
3	B	January	111	600	37,176.0546875	6.8018017674	111

Best Performing Branch on February

- We consider five aspects in determining the best supermarket branch on February: the number of transactions, total sales, overall revenue, customer ratings, and the number of members.

```
--Best Branch February
SELECT
    branch,
    TO_CHAR(CAST("Date" AS DATE), 'Month') AS Month,
    count(invoice_id) as number_of_transaction,
    sum(quantity) as total_sales,
    sum(revenue) as total_revenue,
    avg(rating) as Avg_rating,
    count(customer_type='Member') as member
FROM
    supermarket_sales_clear_csv sscc
where
    EXTRACT(MONTH FROM CAST("Date" AS DATE)) = 2
GROUP BY 1,2
ORDER by 3 DESC;
```

	branch	month	number_of_transaction	total_sales	total_revenue	avg_rating	member
1	B	February	109	624	34,424.26953125	7.0082568597	109
2	C	February	100	537	32,934.9921875	7.1999999666	100
3	A	February	94	493	29,860.12695312	7.0074468014	94

Best Performing Branch on March

- We consider five aspects in determining the best supermarket branch on March: the number of transactions, total sales, overall revenue, customer ratings, and the number of members.

```
--Best Branch March
SELECT
    branch,
    TO_CHAR(CAST("Date" AS DATE), 'Month') AS Month,
    count(invoice_id) as number_of_transaction,
    sum(quantity) as total_sales,
    sum(revenue) as total_revenue,
    avg(rating) as Avg_rating,
    count(customer_type='Member') as member
FROM
    supermarket_sales_clear_csv ssc
where
    EXTRACT(MONTH FROM CAST("Date" AS DATE)) = 3
GROUP BY 1,2
ORDER by 3 DESC;
```

	branch	month	number_of_transaction	total_sales	total_revenue	avg_rating	member
1	A	March	127	681	37,659.11328125	6.9937008009	127
2	B	March	112	596	34,597.35546875	6.6491071539	112
3	C	March	106	614	37,199.02734375	6.8584905705	106

Dashboard

Tools : Microsoft Power BI

In this segment, I have created several dashboards to facilitate the analysis process. Firstly, I designed a report dashboard for each branch in the first quarter of 2019, containing visual to aid the decision-making process for optimizing restocking of goods and determining the optimal number of cashier employees. Secondly, I developed an overall report dashboard for the first quarter of 2019, comprising visual to compare the performance of each branch and identify the best-performing one.



Branch A First Quarter Report

Total Gross Income

5.06K

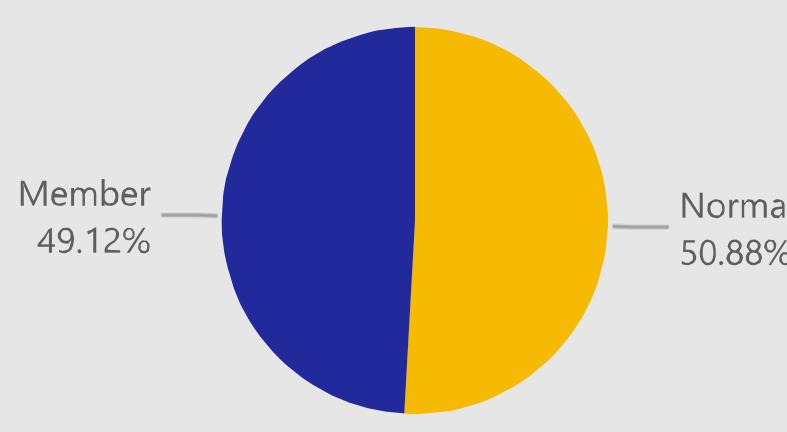
Total Quantity

1859

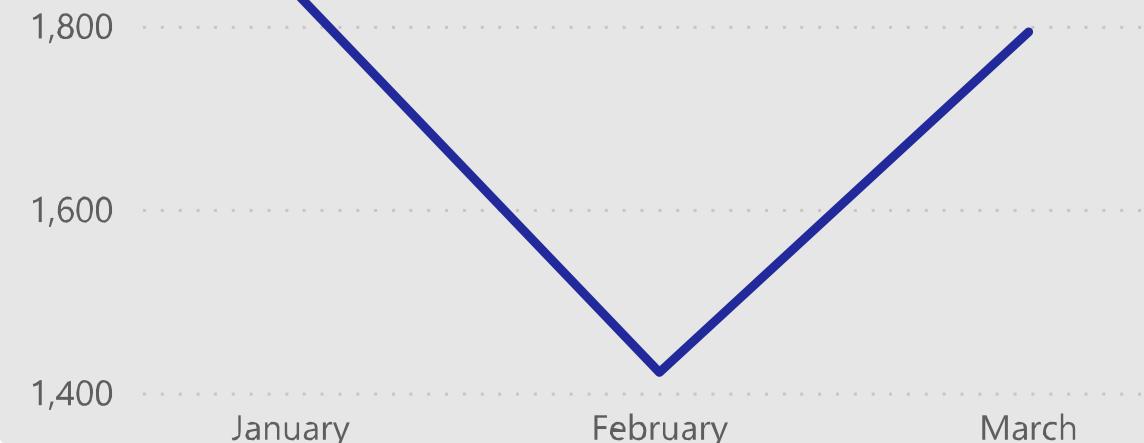
Average Rating

7.03

Customer Type



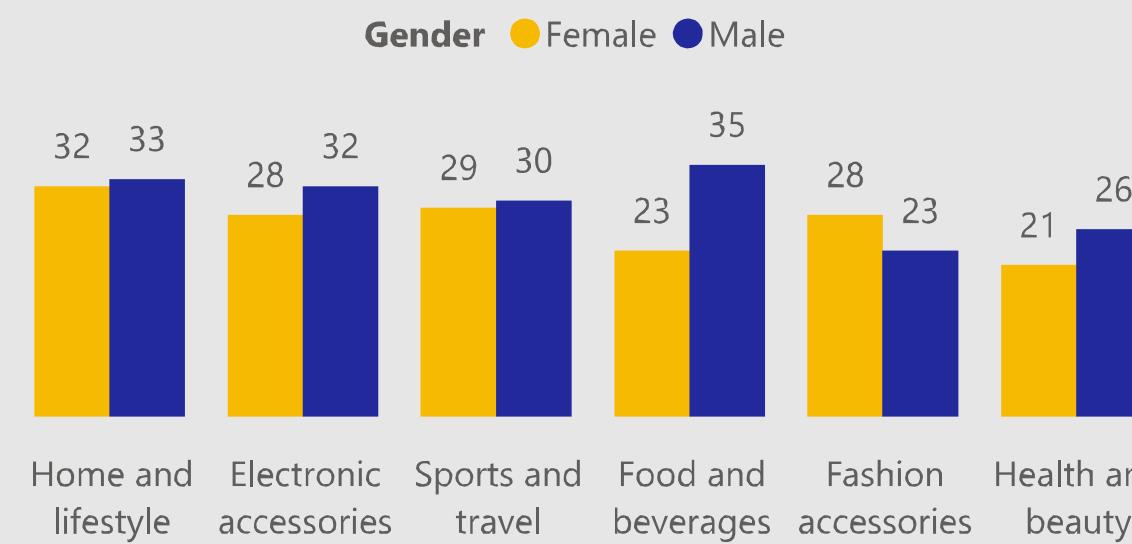
Total Gross Income



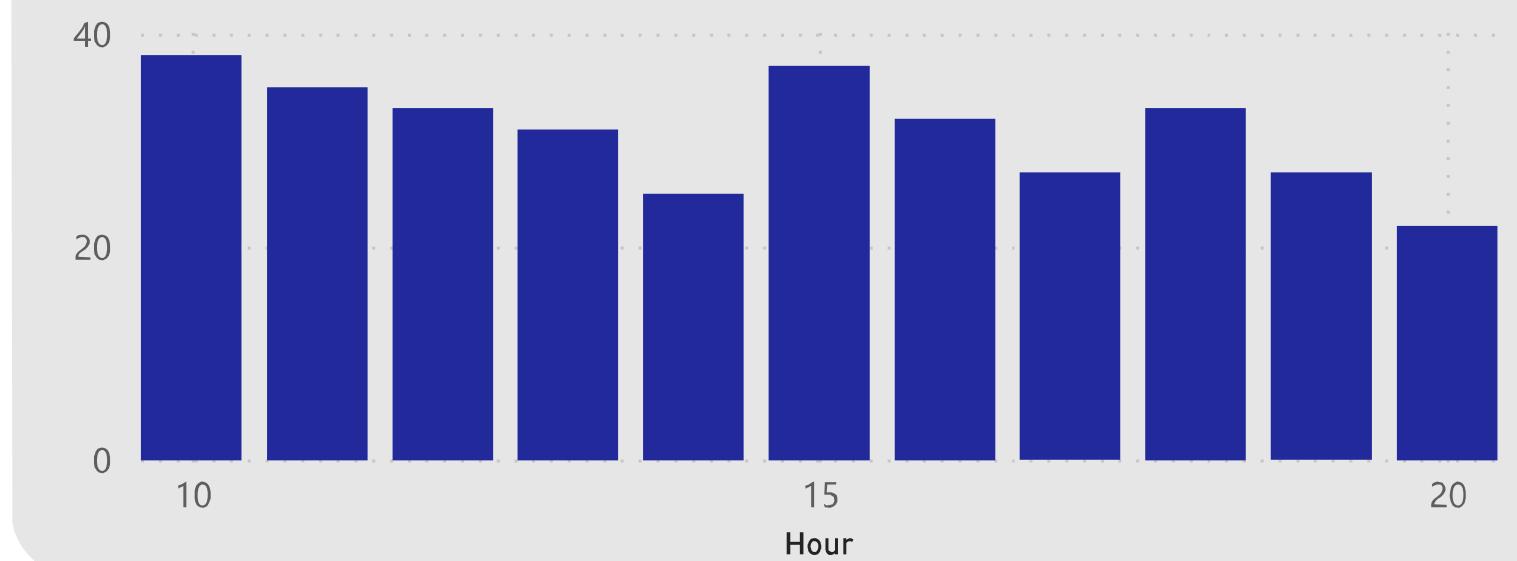
Total Revenue



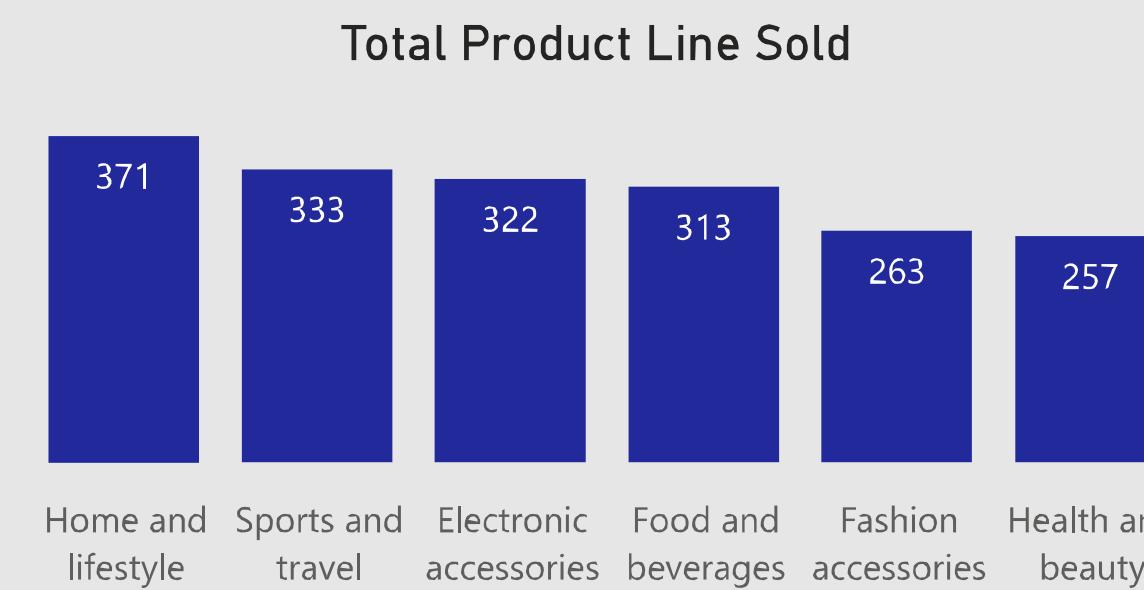
Gender Preference by Product Line



Peak Hours



Busiest Day



Branch B First Quarter Report

Total Gross Income

5.06K

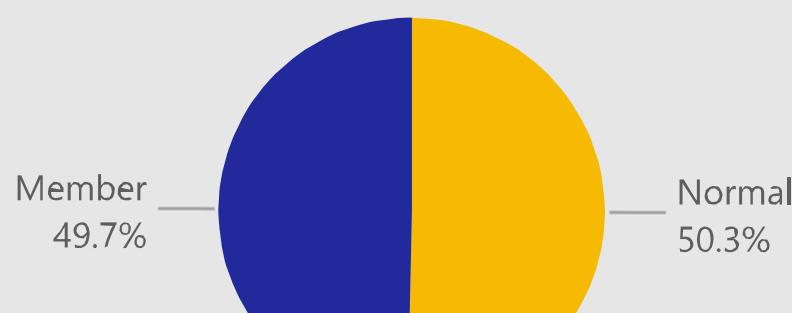
Total Quantity

1820

Average Rating

6.82

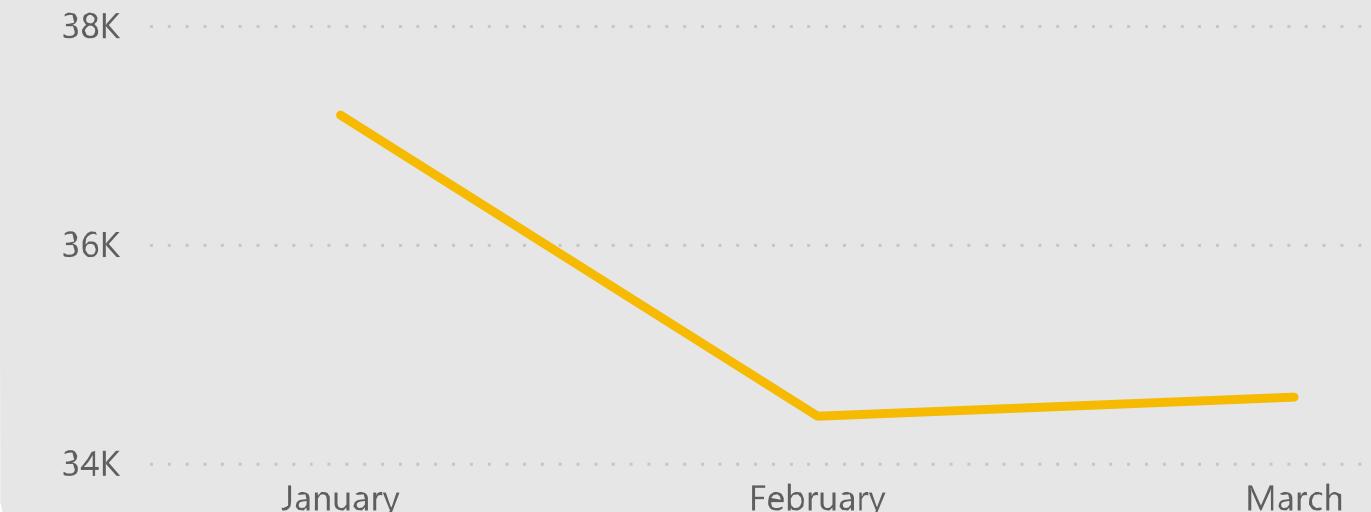
Customer Type



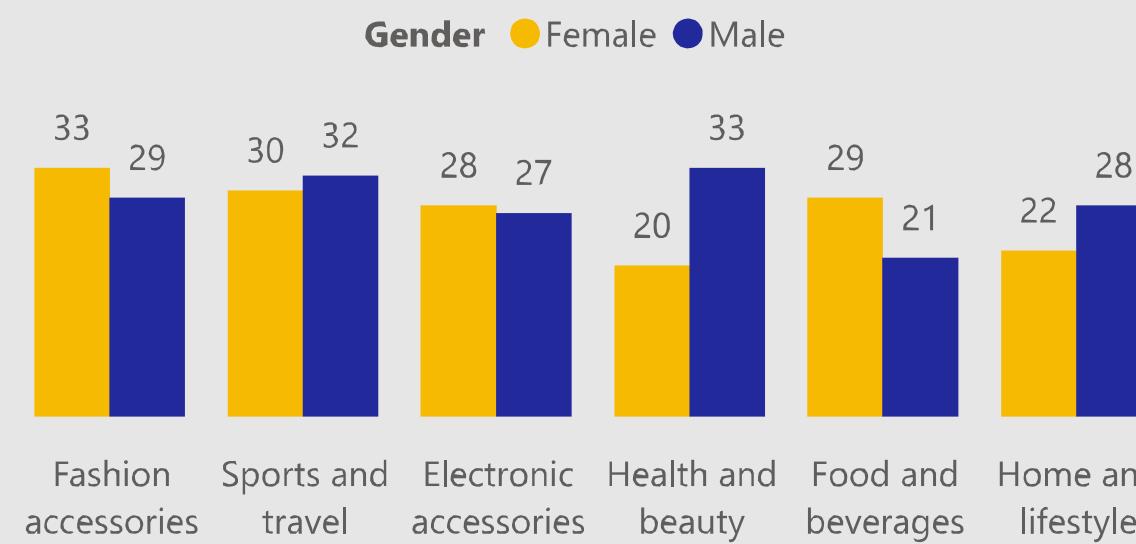
Total Gross Income



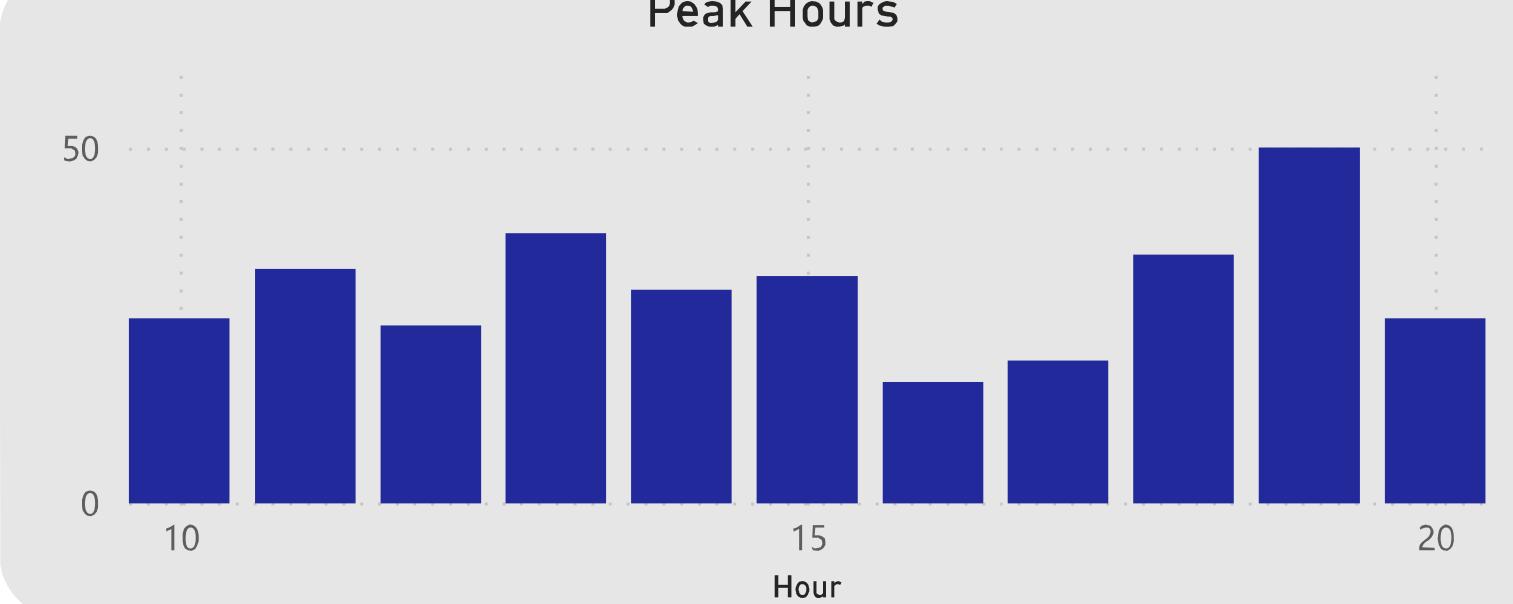
Total Revenue



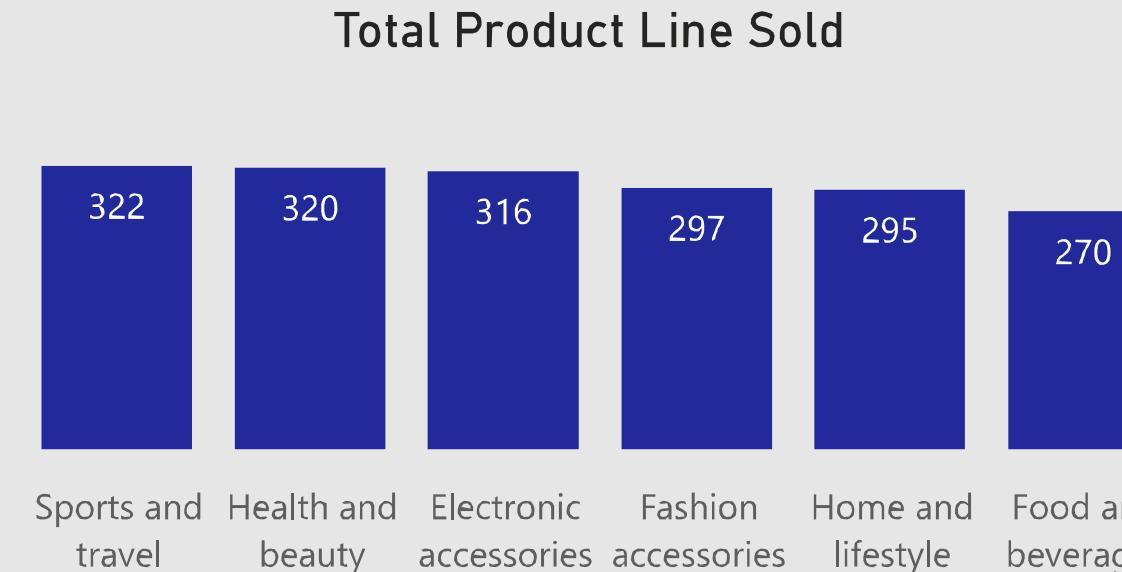
Gender Preference by Product Line



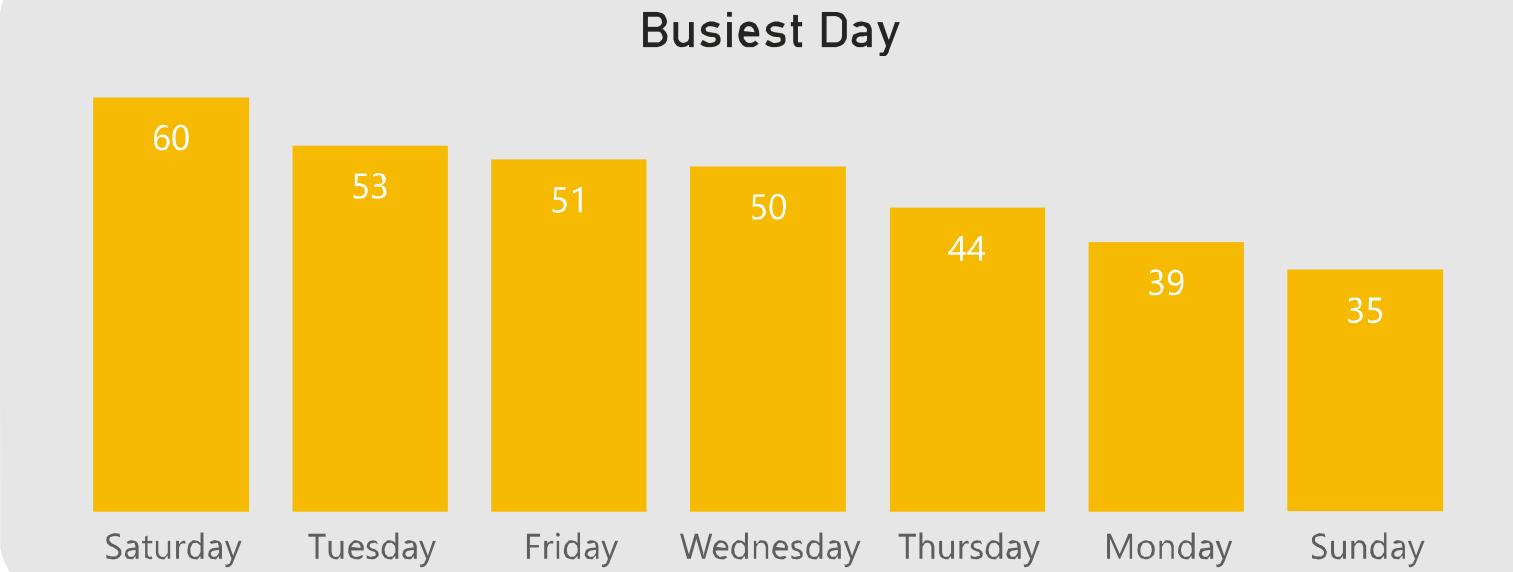
Peak Hours



Total Product Line Sold



Busiest Day



Branch C First Quarter Report

Total Gross Income

5.27K

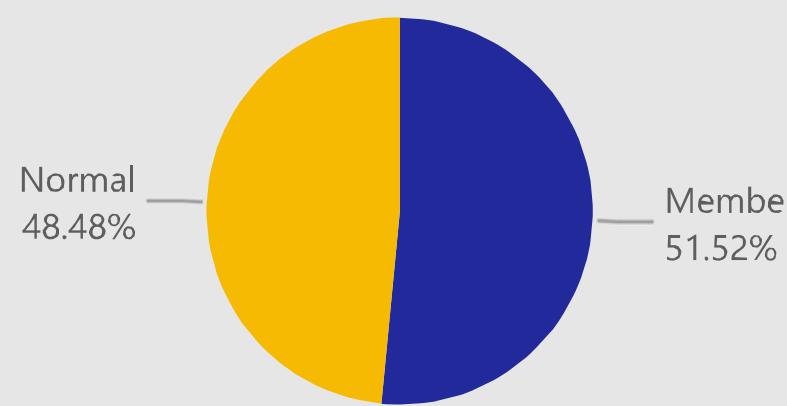
Total Quantity

1831

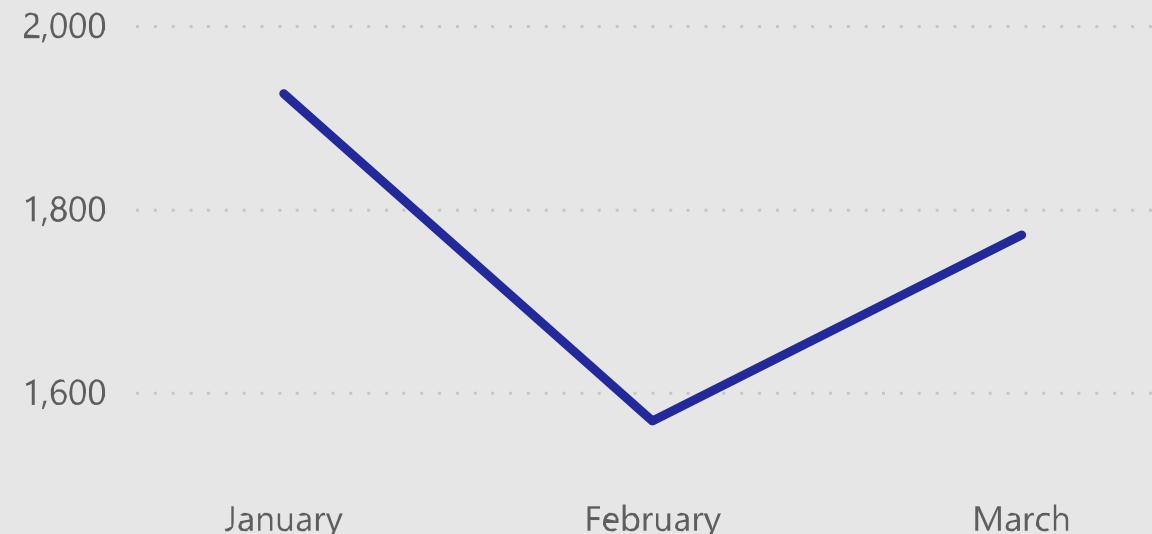
Average Rating

7.07

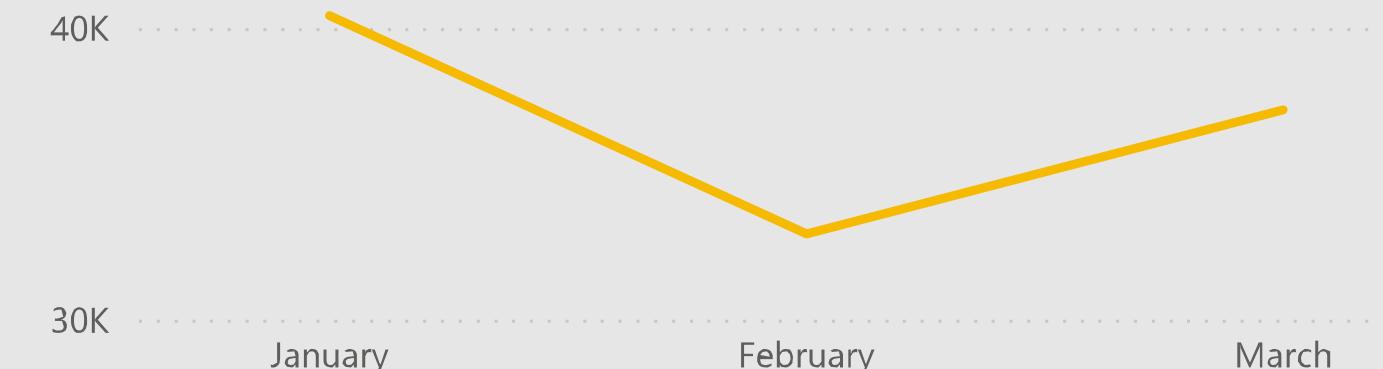
Customer Type



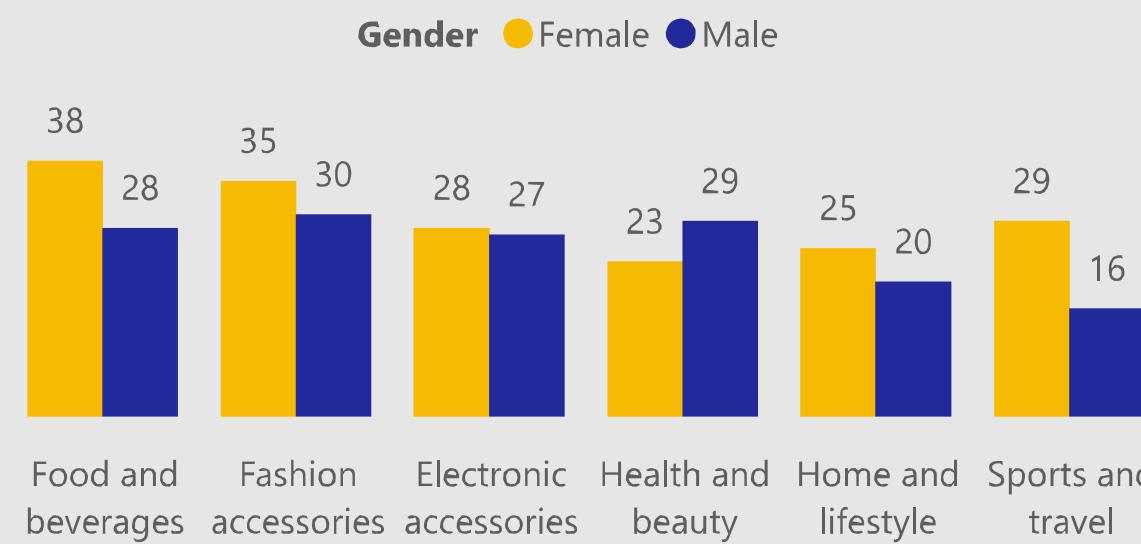
Total Gross Income



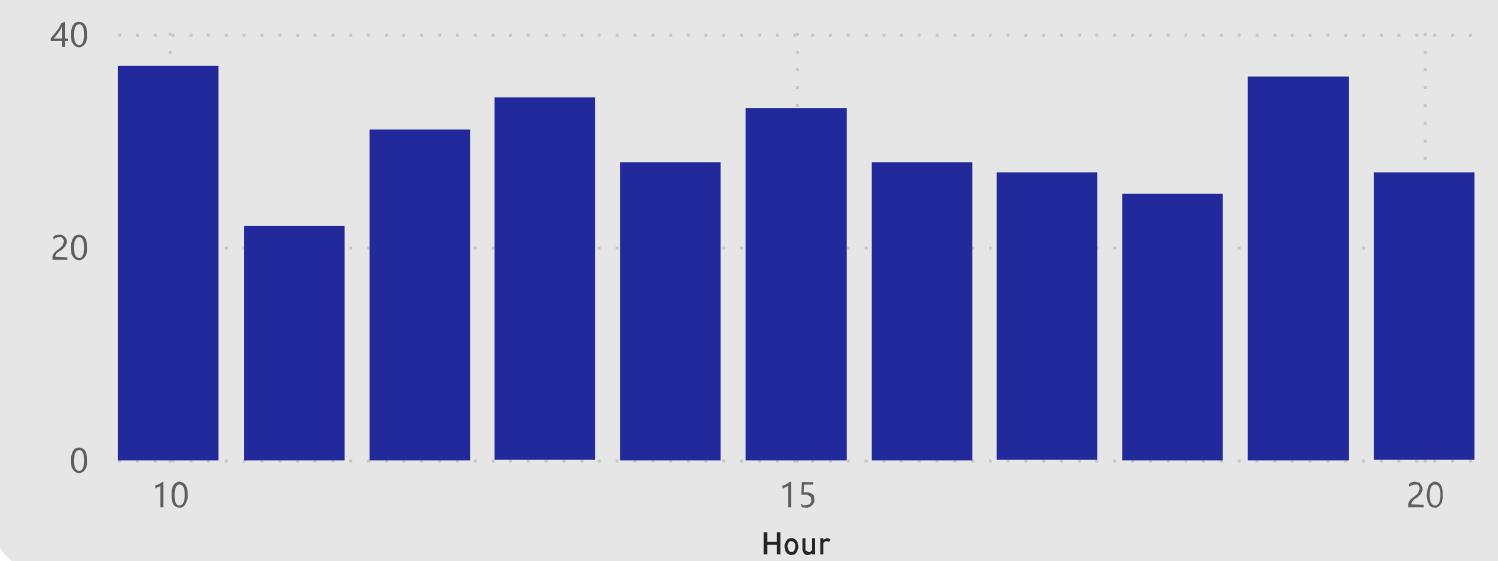
Total Revenue



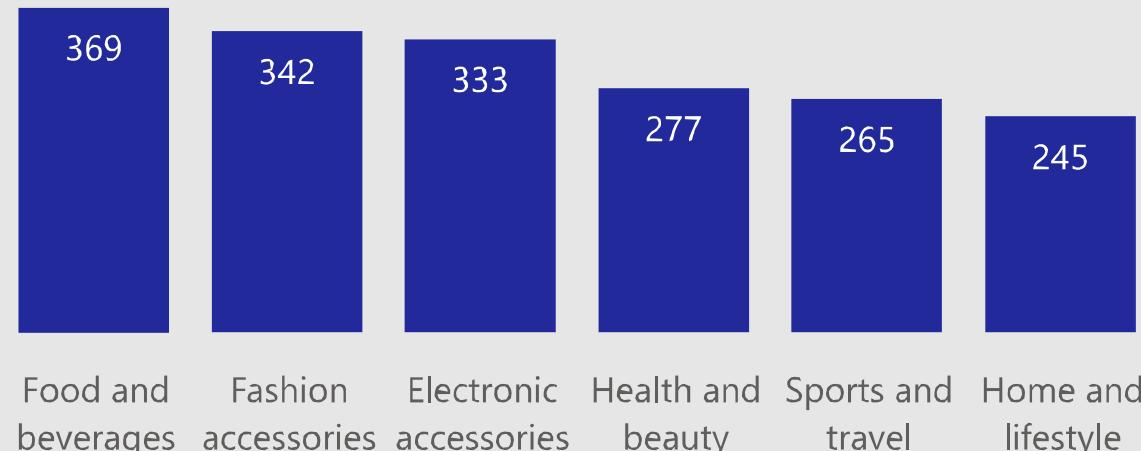
Gender Preference by Product Line



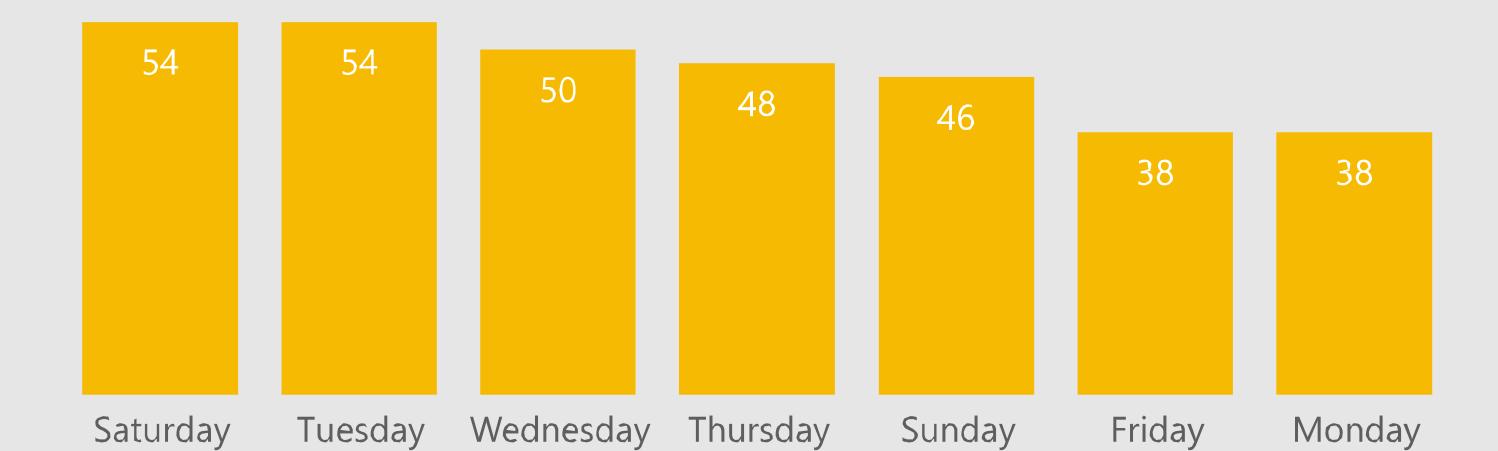
Peak Hours



Total Product Line Sold



Busiest Day



CONCLUSION

Optimizing Stock Replenishment Quantities

- **Branch A**

1. In Branch A, the best-selling product lines include home & lifestyle, sports & travel, and electronic accessories. This information can be taken into consideration when deciding to restock more of these particular product lines.
2. In Branch A, a subtle gender preference was observed in certain product lines. Home & Lifestyle, Electronic Accessories, Sports & Travel, and Health & Beauty lines showed a slightly higher proportion of male shoppers, although the difference was not significant. On the other hand, in the Fashion Accessories line, there was a slight inclination towards female buyers, though the difference was not too pronounced. This gender preference variation could be considered when restocking, taking into account factors such as product colors, usability, and other relevant aspects. However, in the Food & Beverages line, while there were more male buyers, it is my opinion that there is no distinct gender preference in this category.



CONCLUSION

Optimizing Stock Replenishment Quantities

- **Branch B**

1. In Branch B, the best-selling product lines include sports & travel, health & beauty, and electronic accessories. This information can be taken into consideration when deciding to restock more of these particular product lines.
2. In Branch B, a subtle gender preference was observed in certain product lines. Sports & Travel, Health & Beauty, and Home & Lifestyle lines showed a slightly higher proportion of male shoppers, although the difference was not significant. On the other hand, in the Fashion Accessories and Electronic Accessories lines, there was a slight inclination towards female buyers, though the difference was not too pronounced. This gender preference variation could be considered when restocking, taking into account factors such as product colors, usability, and other relevant aspects. However, in the Food & Beverages line, while there were more female buyers, it is my opinion that there is no distinct gender preference in this category.



CONCLUSION

Optimizing Stock Replenishment Quantities

- **Branch C**

1. In Branch C, the best-selling product lines include Food & Beverages, Fashion Accessories, and Electronic Accessories. This information can be taken into consideration when deciding to restock more of these particular product lines.
2. In Branch C, a subtle gender preference was observed in certain product lines. Health & Beauty line showed a slightly higher proportion of male shoppers, although the difference was not significant. On the other hand, in the Fashion Accessories, Electronic Accessories, Sports & Travel, and Home & Lifestyle lines, there was a slight inclination towards female buyers, though the difference was not too pronounced. This gender preference variation could be considered when restocking, taking into account factors such as product colors, usability, and other relevant aspects. However, in the Food & Beverages line, while there were more female buyers, it is my opinion that there is no distinct gender preference in this category.



CONCLUSION

Optimizing Cashier Staffing Levels during Specific Shifts

- **Branch A**

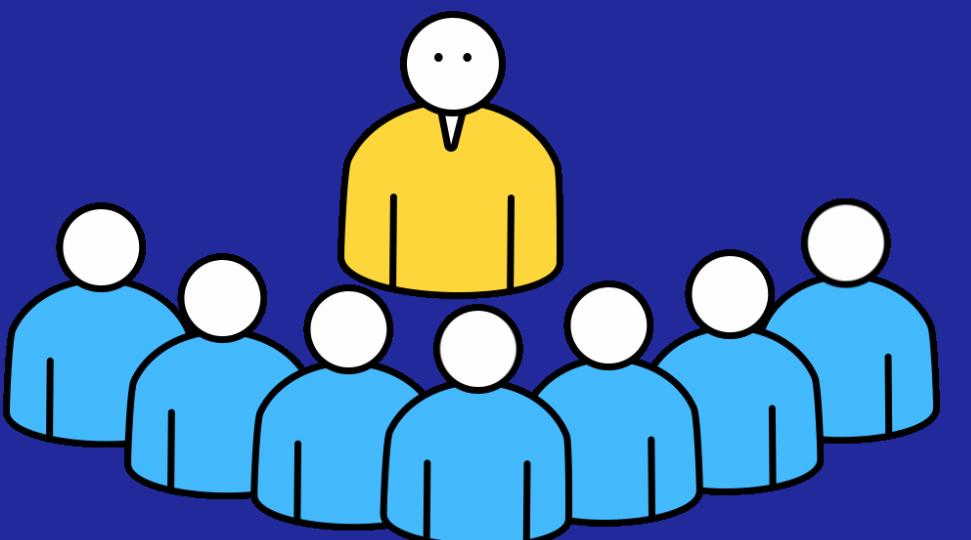
Based on the analysis results, it was found that the busiest days at Branch A are Sunday, Tuesday, Friday, and Saturday, with peak hours occurring around 10 am and 3 pm. This information can be taken into consideration when determining the optimal number of workers or cashier staff during these busy days and hours.

- **Branch B**

Based on the analysis results, it was found that the busiest days at Branch A are Saturday, Tuesday, and Friday with peak hours occurring around 7 pm. This information can be taken into consideration when determining the optimal number of workers or cashier staff during these busy days and hours.

- **Branch C**

Based on the analysis results, it was found that the busiest days at Branch A are Saturday, Tuesday, and Wednesday with peak hours occurring around 10 am and 7 pm. This information can be taken into consideration when determining the optimal number of workers or cashier staff during these busy days and hours.



First Quarter Report (All Branch)

Total Revenue

322.97K

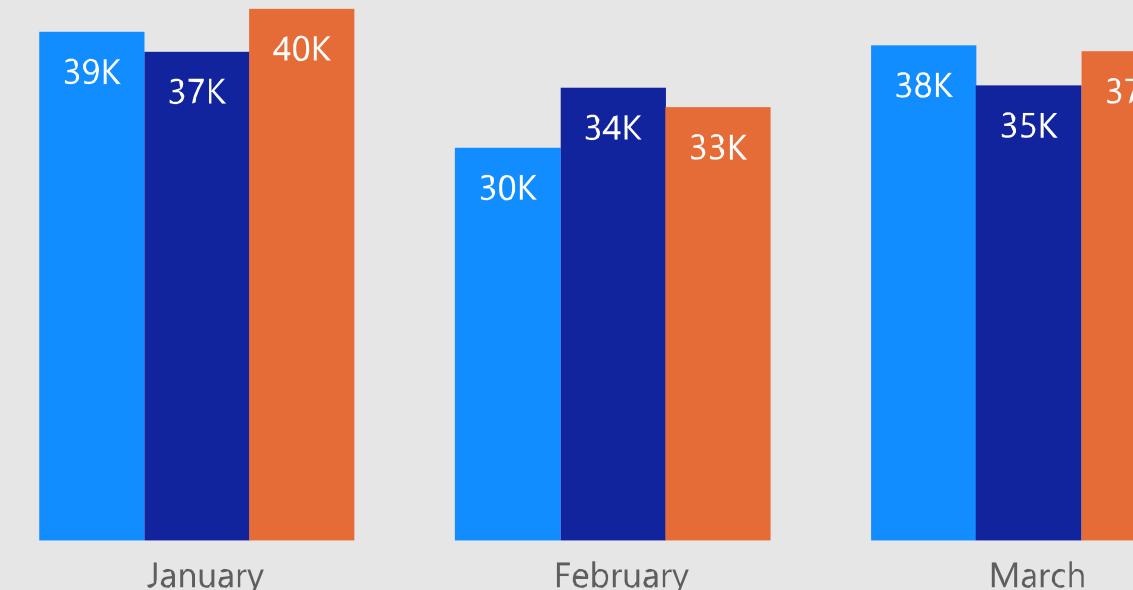
COGS: 307.59K (+5%)

Total Gross Income

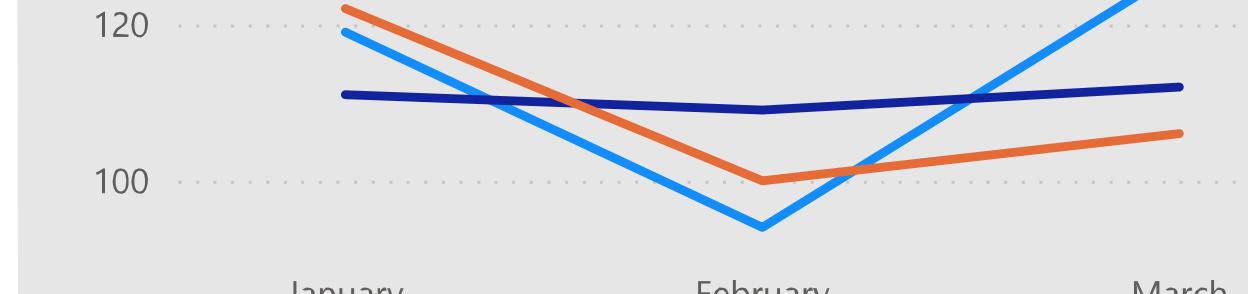
15.38K

Total Revenue by Months

Branch A Branch B Branch C

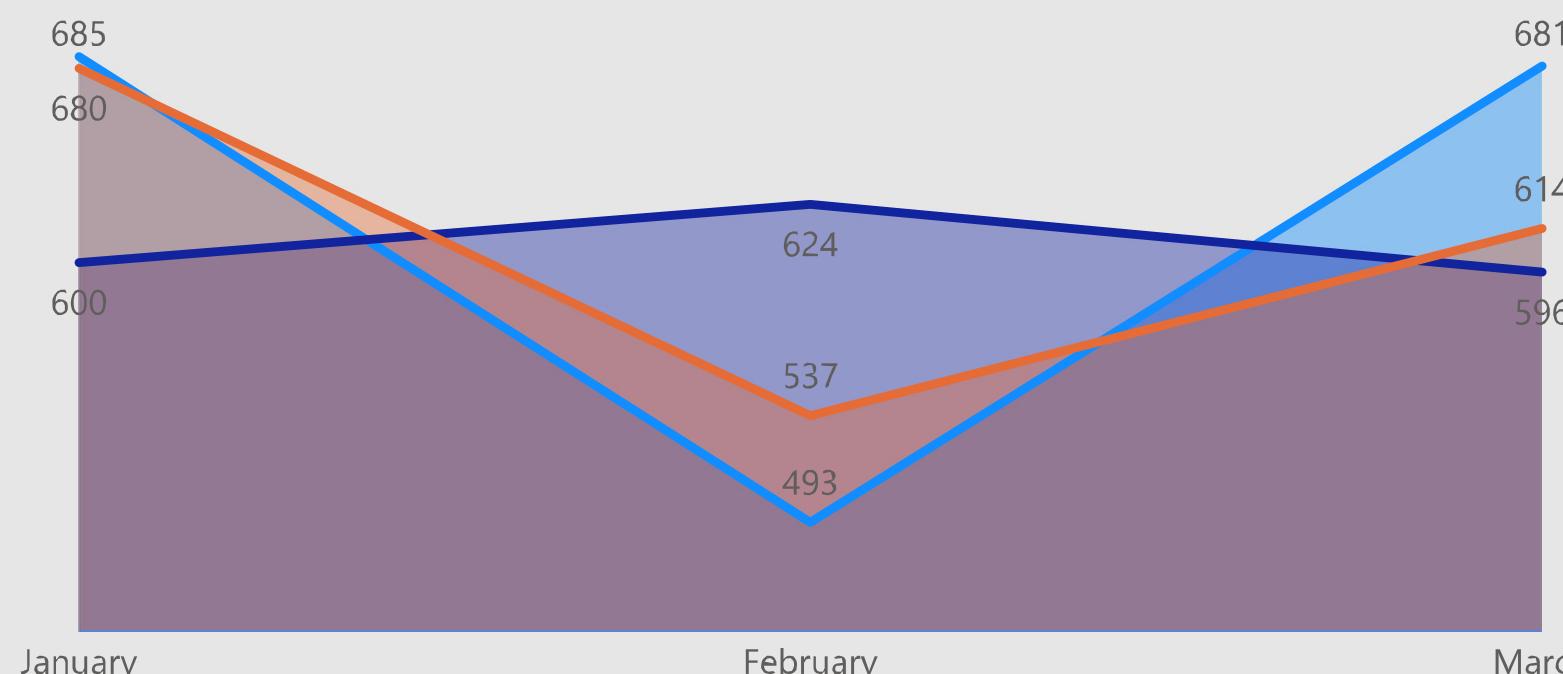


Total Transaction



Total Products Sold

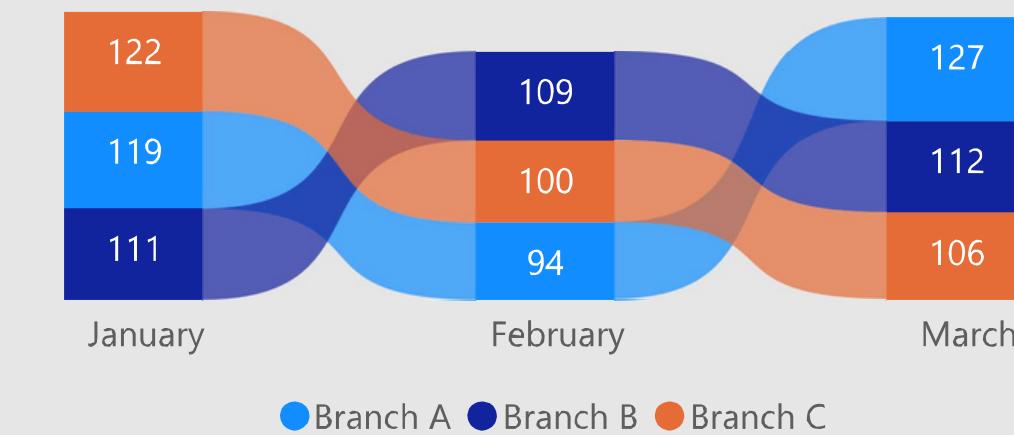
Branch A Branch B Branch C



Average Rating per Branch



Total Member by Months



CONCLUSION

Best Performing Branch

In this section, we will look for the best performing branches by considering several assessment aspects including number of transactions, total sales, total revenue, customer ratings, and number of members in each months.

- **January**

In January, Branch C was entitled to the reward as it achieved the highest scores in four key aspects: revenue, transactions, total members, and average rating. Although Branch C secured the second position in total products sold, it lagged behind Branch A by just 5 items.

- **February**

In February, Branch B was eligible for the reward as it secured the highest scores in four key aspects: revenue, total products sold, transactions, and total members. However, in terms of average rating, Branch B secured the second position, with a difference of 0.2 points from Branch C.

- **March**

In March, Branch A was deserving of the reward as it attained the highest score in all aspects: revenue, total products sold, transactions, average rating, and total members.



Thank You!



: furarizkhi@gmail.com



: linkedin.com/in/faraniaa



: github.com/faraniaa