Name: Muhammad Isa Dadi Hasibuan
NIM: 207056013

Data set

The data set I'm using comes from kaggle.com. The data set is in the form of sales data in supermarkets. The data set contains historical sales of supermarket company which has recorded in 3 different branches for 3 months data. The following are the attributes in the data set:

- Invoice id: Computer generated sales slip invoice identification number
- Branch: Branch of supermarket (3 branches are available identified by A, B and C).
- City: Location of supermarket
- Customer type: Type of customers, recorded by Members for customers using member card and Normal for without member card.
- Gender: Gender type of customer
- Product line: General item categorization groups - Electronic accessories, Fashion accessories, Food and beverages, Health and beauty, Home and lifestyle, Sports and travel
- Unit price: Price of each product in $
- Quantity: Number of products purchased by customer
- Tax: 5% tax fee for customer buying
- Total: Total price including tax
- Date: Date of purchase (Record available from January 2019 to March 2019)
- Time: Purchase time (10am to 9pm)
- Payment: Payment used by customer for purchase (3 methods are available – Cash, Credit card and Ewallet)
- COGS: Cost of goods sold
- Gross margin percentage: Gross margin percentage
- Gross income: Gross income
- Rating: Customer stratification rating on their overall shopping experience (On a scale of 1 to 10)

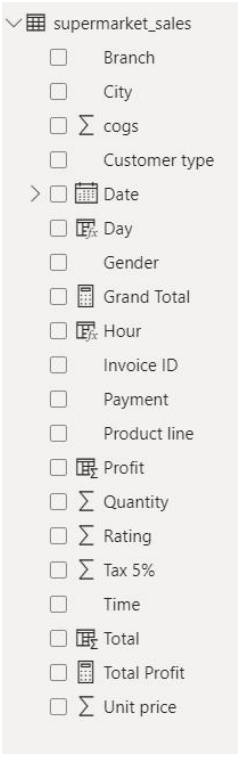
Expected Insights

Expected insight are to get the day of the most sales and get product category with the most product sales per branch.

Name: Muhammad Isa Dadi Hasibuan
NIM: 207056013

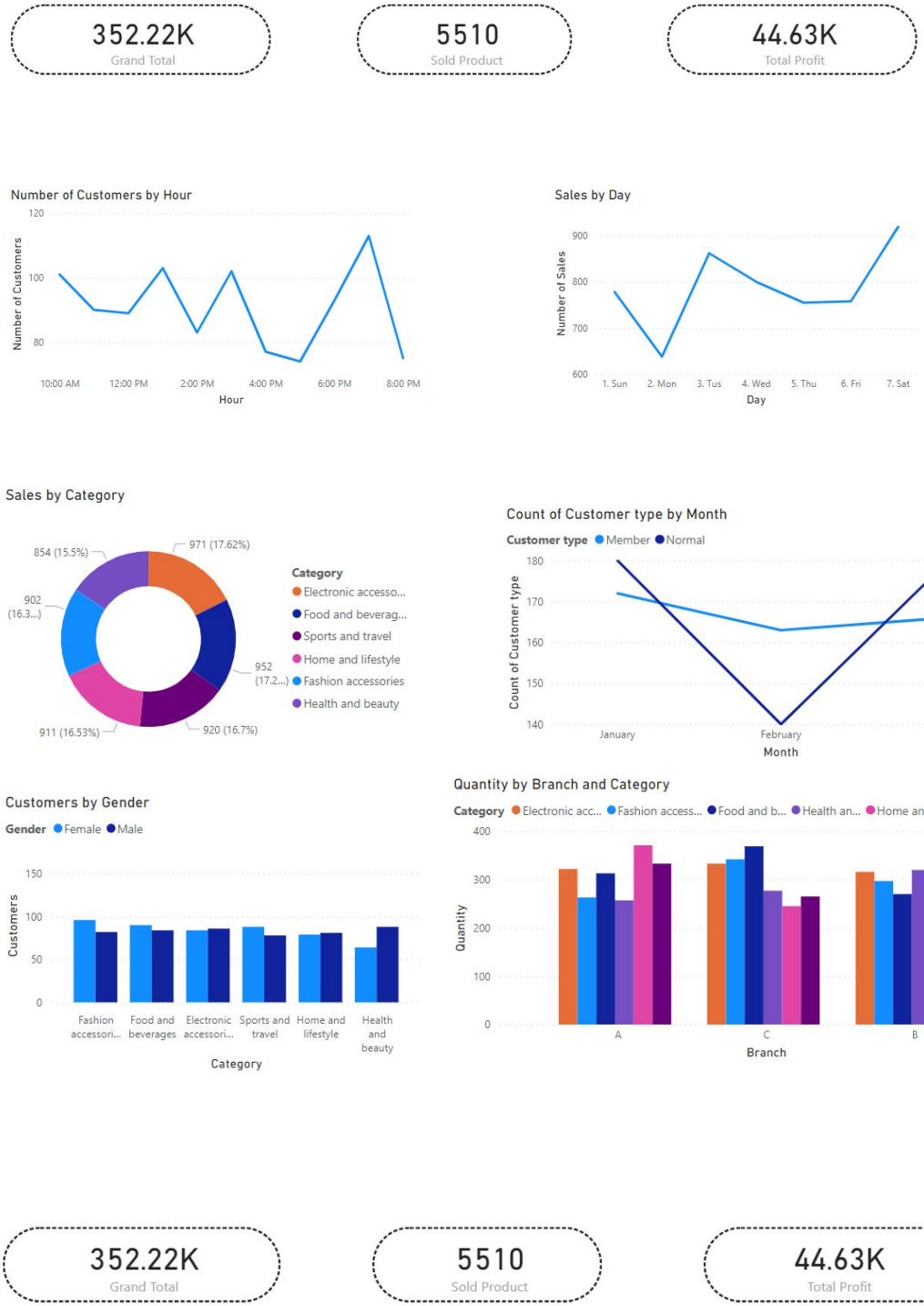
Model

The model only consists of one table name supermarket_sales. In the model there is a total column, but when I importing the data, comma error occurs. So I removed the column and replaced it using DEX. I use the quantity, unit price, and tax attributes to create the total column. Then I made measures to calculate the grand total of sales and profit. Because I wanted to get insight about the most sales days, so I created a column to convert the dates in the date column into days. In the power BI, the DAX function to get days is available, but the output is numbers 1-7 which represent the days of the week. so I created a column to convert those numbers into text like sun, mon, etc. I also create hour column to group the sales by hour. I also use the branch column to group sales by branch. as well as gender and product line (category).



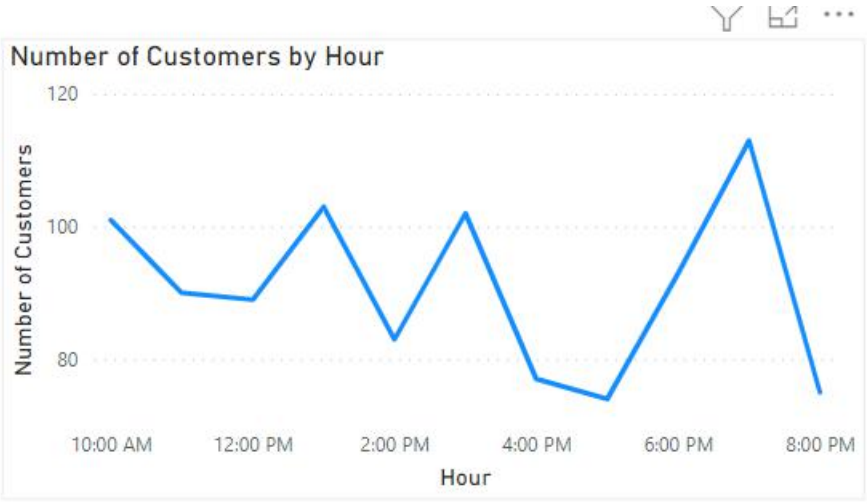
Name: Muhammad Isa Dadi Hasibuan
NIM: 207056013

Visualization and Obtained Insights



From the figure above, we can see the grand total that created using DAX, count of sold products, and total profits.

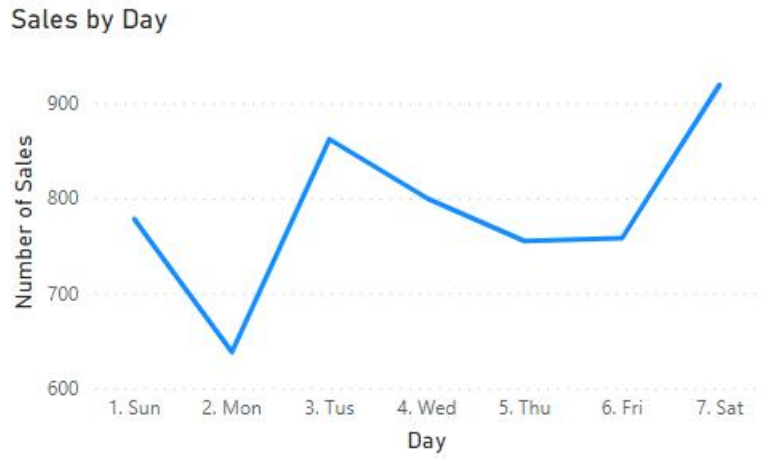
Name: Muhammad Isa Dadi Hasibuan
NIM: 207056013



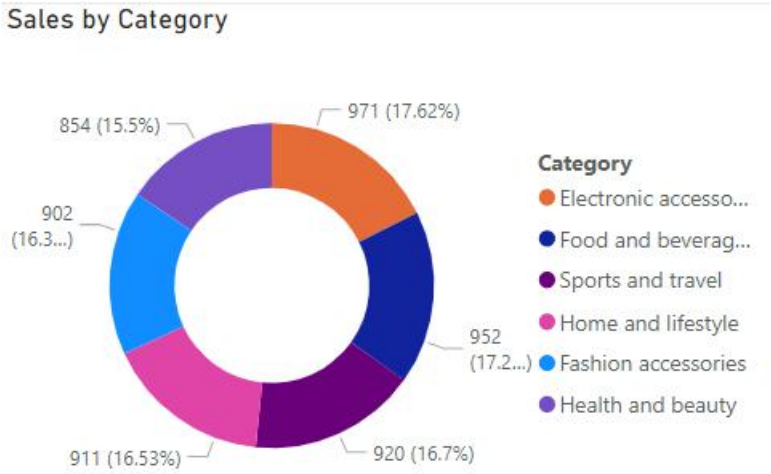
From the figure above, we can see that most customers come at 7:00 PM. At 7:30 PM Number of customers had the highest number at 113 and was 52.70% higher than 5:00 PM, which had the lowest Number of Customers at 74. 7:00:00 PM accounted for 11.30% of Number of Customers across all 11 Hour ranged from 74 to 113. Almost everyday most customers came at 7:00 PM except on Thursday.

On Thursday the highest number of customers happened at 10:00 AM. From 6:00 PM to 8:00 PM had the lowest customers.

Name: Muhammad Isa Dadi Hasibuan
NIM: 207056013

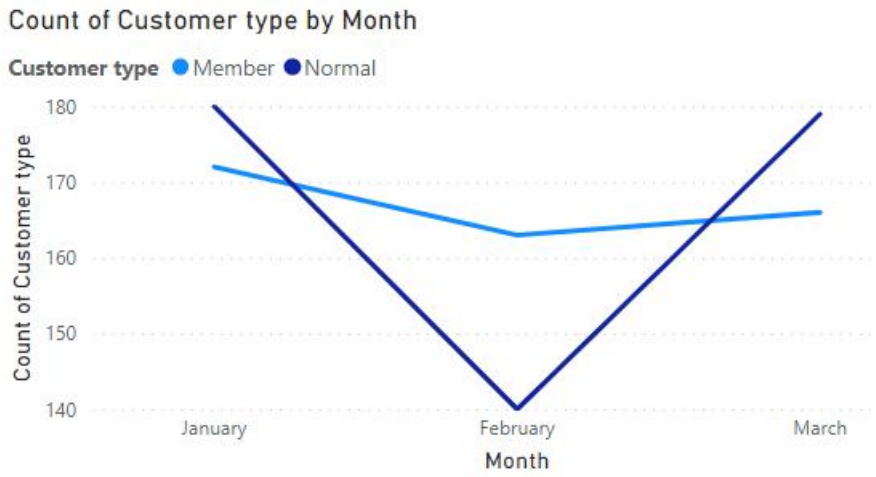


From the figure above, We can see that the highest Number of Sales happened on Saturday at 919 and was 44.04% higher than Monday, which had the lowest Number of Sales at 638. Saturday accounted for 16.68% of Number of Sales across all 7 Day ranged from 638 to 919.

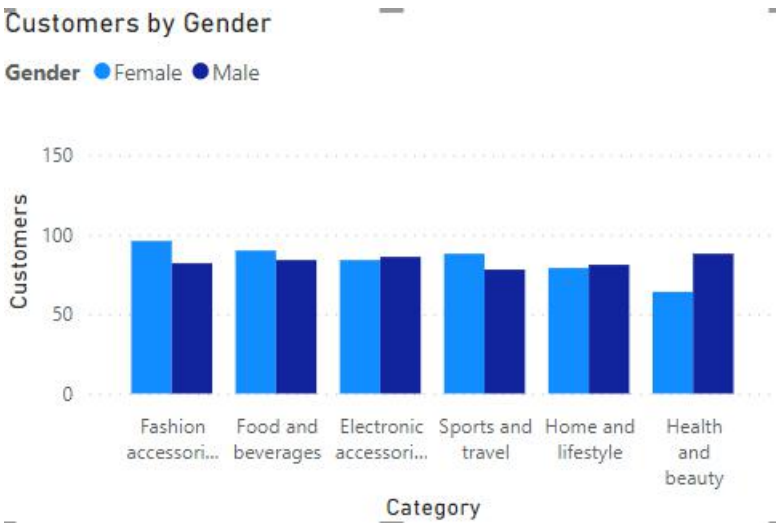


From the figure above, we can see that the sales differ only slightly. Electronic accessories had the highest percentage of 17.62%

Name: Muhammad Isa Dadi Hasibuan
NIM: 207056013

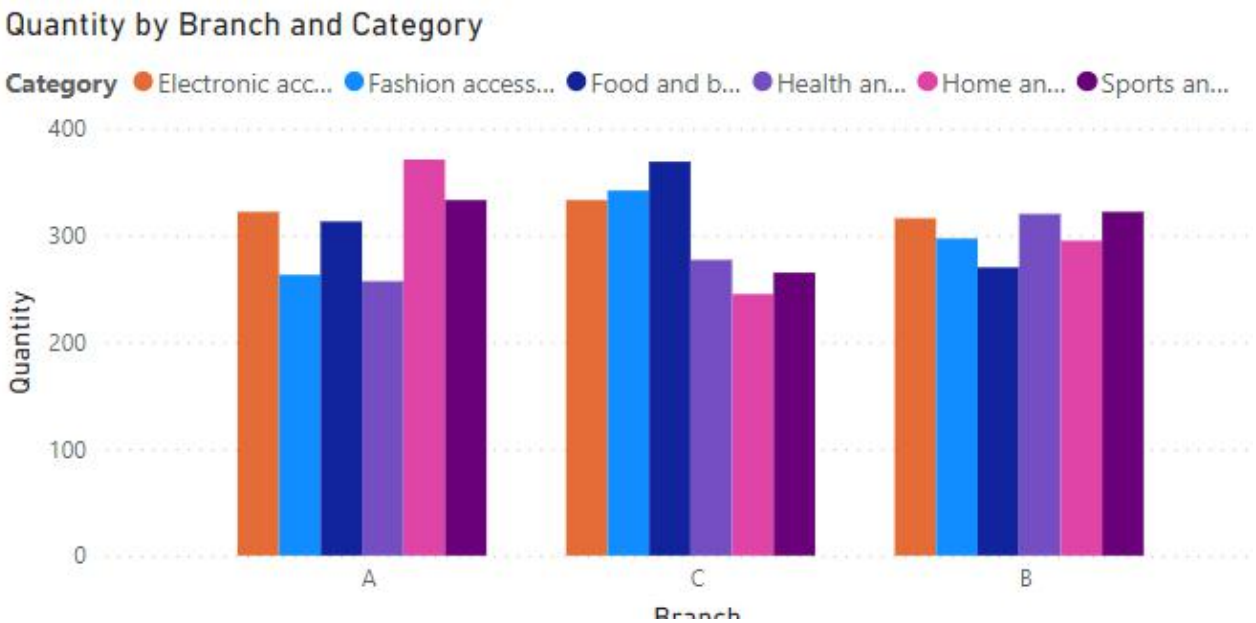


Total Count of Customer type was slightly higher for Member (501) than Normal (499). In January, Customer type made up 18.00% of total customer. Customer type for Member and Normal diverged the most when the Month was February, when Member were 23 higher than Normal.



Total Customers was higher for Female (501) than Male (499). Fashion accessories was the highest, made up 9.60% of Customers. Average Customers was higher for Female (83.50) than Male (83.17). Customers for Male and Female diverged the most when the Category was Health and beauty, when Male were 24 higher than Female.

Name: Muhammad Isa Dadi Hasibuan
NIM: 207056013



From the picture above we can see that home and lifestyle had the highest sales in branch A (371). While in branch B the most sales was food and beverages (369) and sport and travel was the highest in branch C (322).