DA Assignment – 1

~**Soumajit Pal**

(21BCE7214)

VIT-AP

The growth of supermarkets in most populated cities is increasing and market competitions are also high. The dataset is one of the historical sales of supermarket company which has recorded in 3 different branches for 3 months data. Predictive data analytics methods are easy to apply to this dataset.

Attribute information -

Invoice id: Computer-generated sales slip invoice identification number

Branch: Branch of supercenter (3 branches are available identified by A, B and C).

City: Location of supercenters

Customer type: Type of customers, recorded by Members for customers using member cards and Normal for those without member cards.

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: The price of each product in $

Quantity: Number of products purchased by the customer

Tax: 5% tax fee for customers buying

Total: Total price including tax

Date: Date of purchase (Record available from January 2019 to March 2019)

Time: Purchase time (10 am to 9 pm)

Payment: Payment used by the customer for the 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)

Challenge: Upload the dataset to Cognos Analytics, delete the unnecessary columns, create a data module, explore and visualize the dataset.

Data Modules-

Columns Removed:

1)**Invoice Id:** Because that attribute would not add any information to the visualizations.

2) **Branch:** Because that attribute and the City attribute would have given the same result.

3)**Gross Margin Percentage:** Because the value of all the records in the attribute is same and it would not make any difference to the visualizations.

Columns Split-

1)Date: The date was split into Months and Days of the Week as the data could be then divided into 3 months and the 7 days of the week respectively.

New Dataset-

↑↓	Row Id	City	Customer type	Gender	Product line	Unit price	Quantity
	1	Yangon	Member	Female	Health and beauty	74.69	7
	2	Naypyitaw	Normal	Female	Electronic accessories	15.28	5
	3	Yangon	Normal	Male	Home and lifestyle	46.33	7
	4	Yangon	Member	Male	Health and beauty	58.22	8
	5	Yangon	Normal	Male	Sports and travel	86.31	7
	6	Naypyitaw	Normal	Male	Electronic accessories	85.39	7
	7	Yangon	Member	Female	Electronic accessories	68.84	6
	8	Naypyitaw	Normal	Female	Home and lifestyle	73.56	10
	9	Yangon	Member	Female	Health and beauty	36.26	2
	10	Mandalay	Member	Female	Food and beverages	54.84	3
	11	Mandalay	Member	Female	Fashion accessories	14.48	4
	12	Mandalay	Member	Male	Electronic accessories	25.51	4
	13	Yangon	Normal	Female	Electronic accessories	46.95	5

Tax 5%	Total	Date	Month	Day of the week	Time	Payment
26.1415	548.9715	2019-01-05	1	Saturday	13:08:00	Ewallet
3.82	80.22	2019-03-08	3	Friday	10:29:00	Cash
16.2155	340.5255	2019-03-03	3	Sunday	13:23:00	Credit card
23.288	489.048	2019-01-27	1	Sunday	20:33:00	Ewallet
30.2085	634.3785	2019-02-08	2	Friday	10:37:00	Ewallet
29.8865	627.6165	2019-03-25	3	Monday	18:30:00	Ewallet
20.652	433.692	2019-02-25	2	Monday	14:36:00	Ewallet
36.78	772.38	2019-02-24	2	Sunday	11:38:00	Ewallet
3.626	76.146	2019-01-10	1	Thursday	17:15:00	Credit card
8.226	172.746	2019-02-20	2	Wednesday	13:27:00	Credit card
2.896	60.816	2019-02-06	2	Wednesday	18:07:00	Ewallet
5.102	107.142	2019-03-09	3	Saturday	17:03:00	Cash
11.7375	246.4875	2019-02-12	2	Tuesday	10:25:00	Ewallet

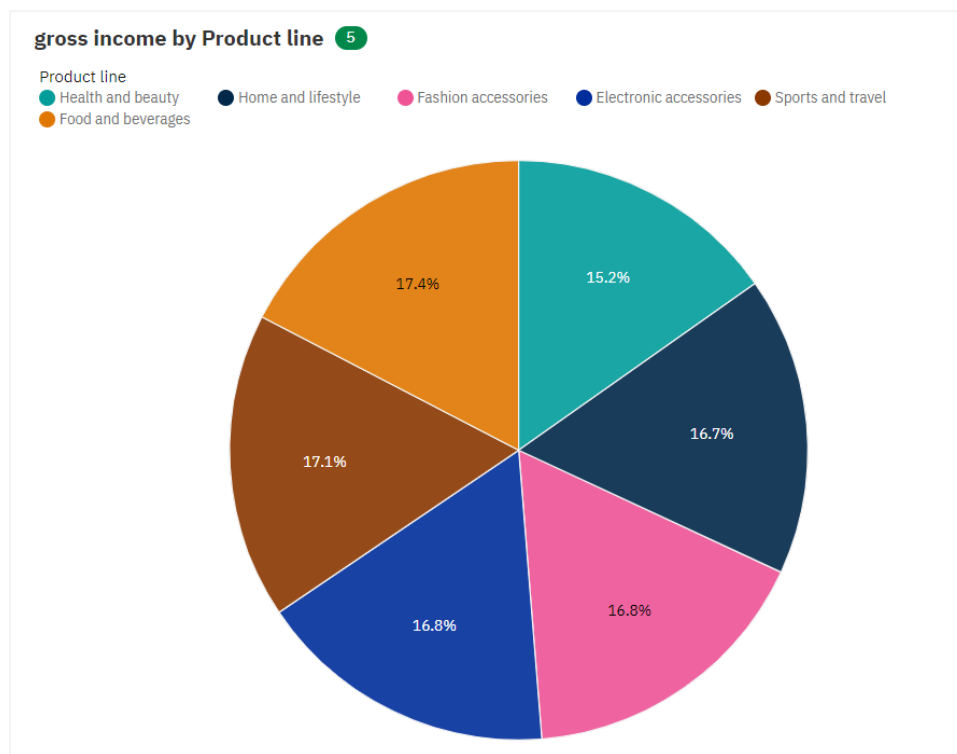| cogs	gross income	Rating
522.83	26.1415	9.1
76.4	3.82	9.6
324.31	16.2155	7.4
465.76	23.288	8.4
604.17	30.2085	5.3
597.73	29.8865	4.1
413.04	20.652	5.8
735.6	36.78	8
72.52	3.626	7.2
164.52	8.226	5.9
57.92	2.896	4.5
102.04	5.102	6.8
234.75	11.7375	7.1

Visualization-

1) Total, cogs and gross income by City –

Observation- The city of Naypyitaw had the maximum sales and gross income compared to the other cities.

2) Gross income by Product line-

gross income by Product line 5

Product line
● Health and beauty ● Home and lifestyle ● Fashion accessories ● Electronic accessories ● Sports and travel
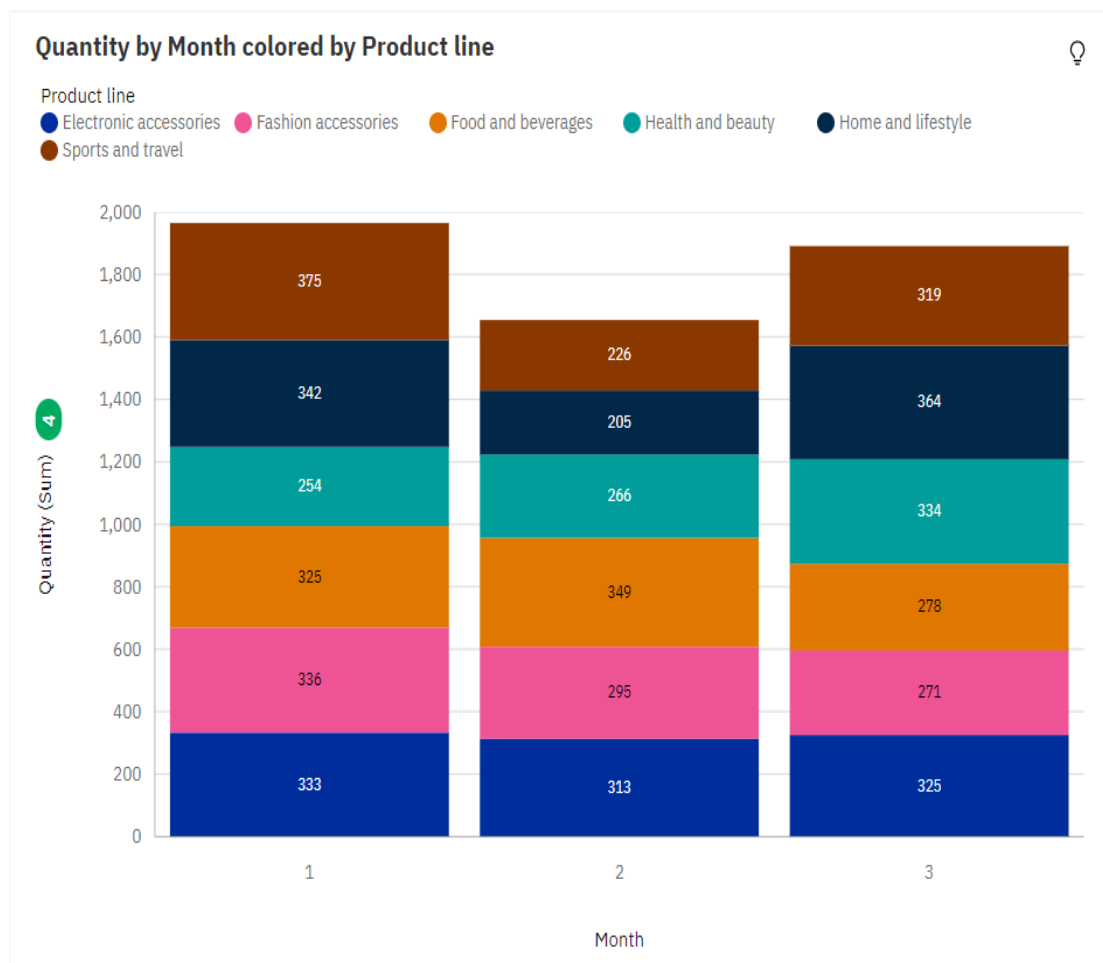● Food and beverages

15.2%
16.7%
16.8%
16.8%
17.1%
17.4%

Observation- Food and beverages had the maximum gross income while Health and Beauty earned the least.

3) Product line coloured by Product line sized by Total-

Product line colored by Product line sized by Total 5

Total (Sum)
49,193.74   56,144.84

Product line
● Electronic accessories ● Fashion accessories ● Food and beverages ● Health and beauty
● Home and lifestyle ● Sports and travel

Sports and travel

Home and lifestyle
Food and beverages
Electronic accessories

Fashion accessories

Observation- Food and beverages had the maximum total sales value while Health and Beauty earned the least.

4) Quantity by Month coloured by Product line-

Observations-

First Month-

- Maximum Quantity Sold- Sports and Travel(375)
- Minimum Quantity Sold- Health and Beauty(254)
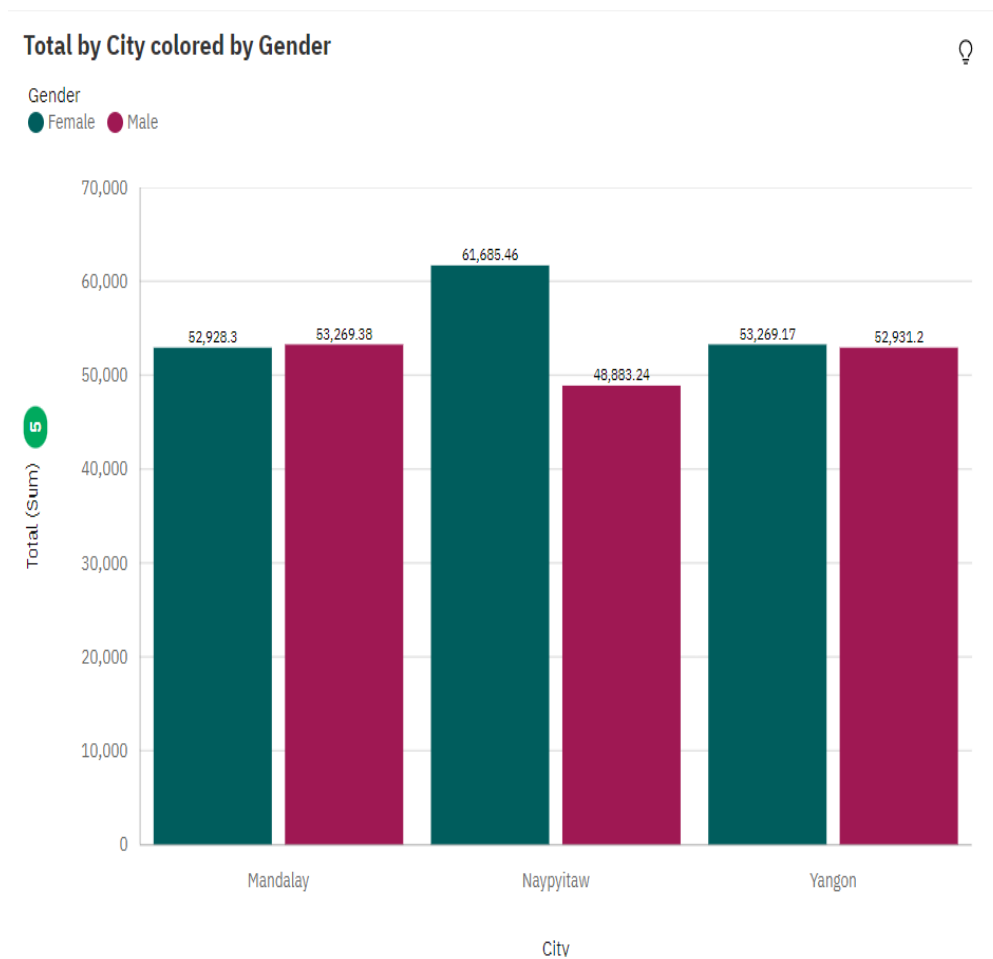
Second Month-

- Maximum Quantity Sold- Food and Beverages(349)
- Minimum Quantity Sold- Home and Lifestyle(205)

Third Month-

- Maximum Quantity Sold- Home and Lifestyle(364)
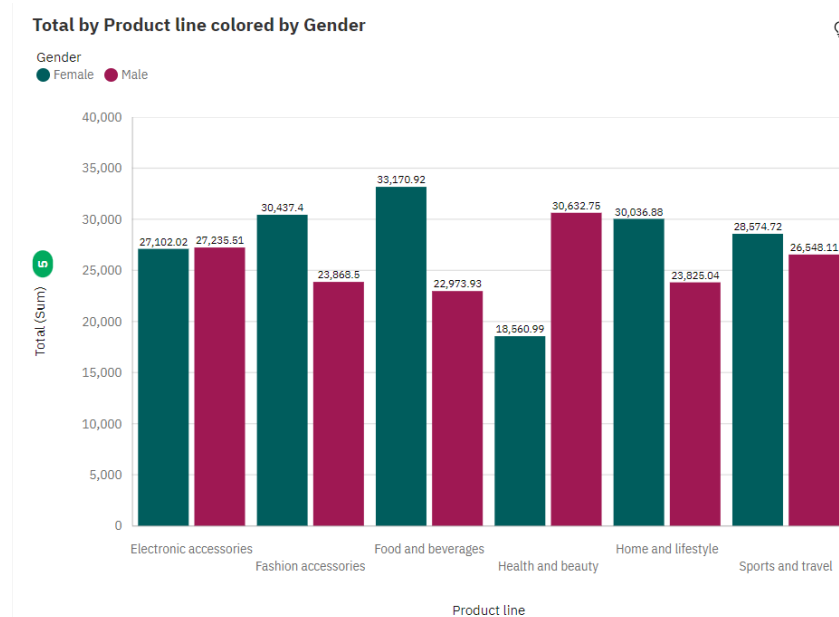- Minimum Quantity Sold- Fashion Accessories(271)

5) Total by City coloured by Gender-

Total by City colored by Gender

Gender
● Female ● Male

Observations-

- In the city of Naypyitaw, female customers purchased $12,802.22 more than the male customers.
- In other cities no major difference was observed in the gender-wise sales.

6) Total by Product line coloured by Gender-

Observation- In food and beverages & fashion accessories, there are more women customers compared to male customers. In health and beauty , there are more men consumers compared to women consumers.

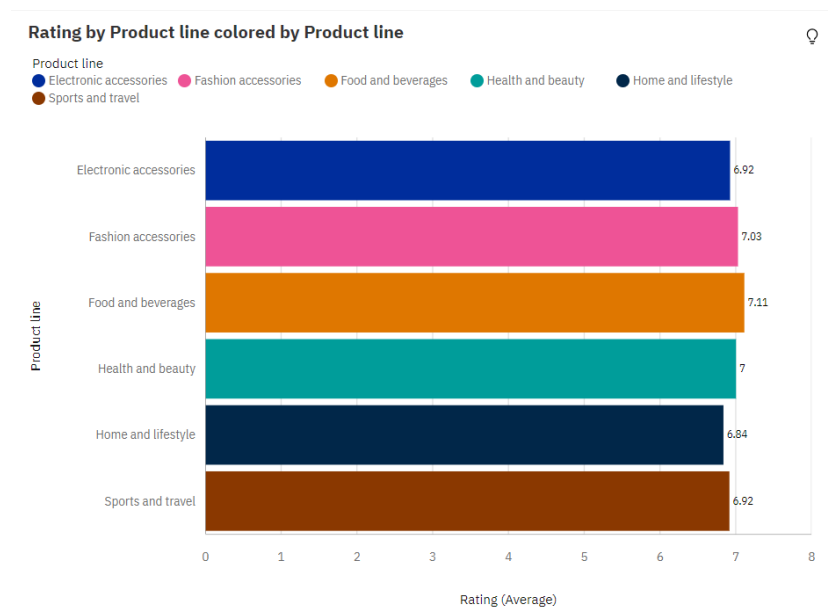## 7) Total by City and Day of the week-
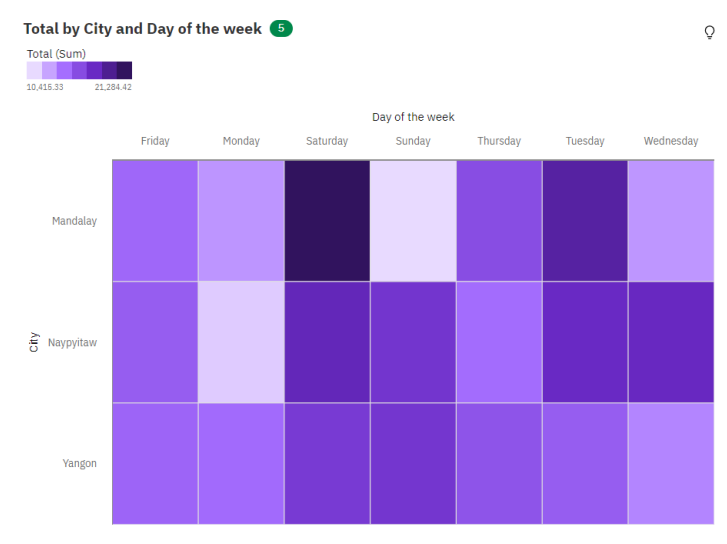
Observations-

- In the city of Mandalay, Saturday had the maximum sales while Sunday had the least.
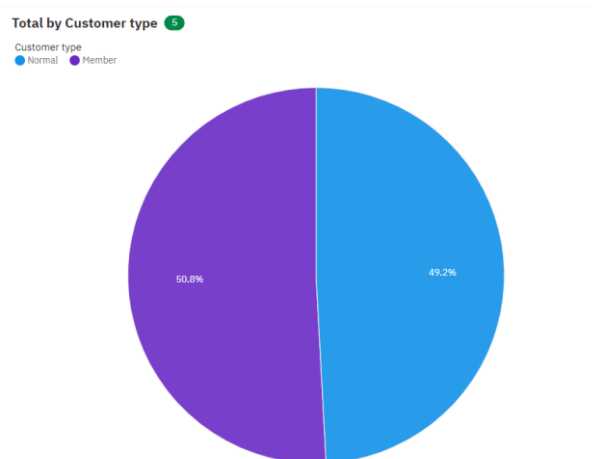- In the city of Naypyitaw, the lowest sales was on Monday.

## 8) Rating by Product line coloured by Product line-

Observation- The customers are most satisfied with food and beverages while they are least satisfied with home and lifestyles which needs improvement.

9) Total by Customer type-

Observation- There is no major difference between the purchasing amount between the normal customers and the member customers.

10) Total by City coloured by Payment-
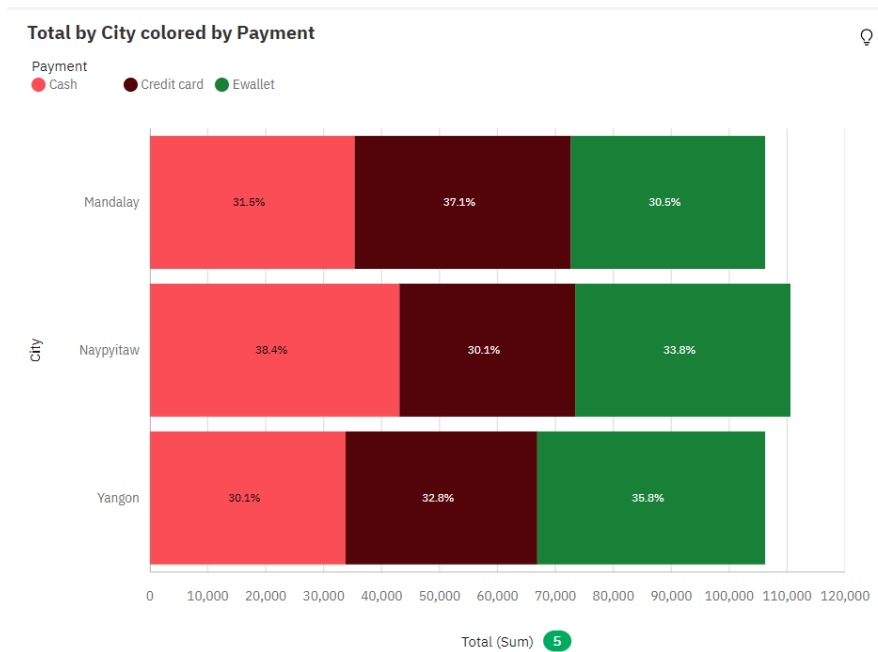
Observation-

- In Mandalay, people prefer to pay using credit cards(37.1%).
- In Naypyitaw, people prefer to pay using cash(38.4%).
- In Yangon, people prefer to pay using E-wallet(35.8%).