

Supermarket Sales

DEPI Data Analysis First Project Using Python



By: Ahmed Essam Abdallah

❖ Introduction:

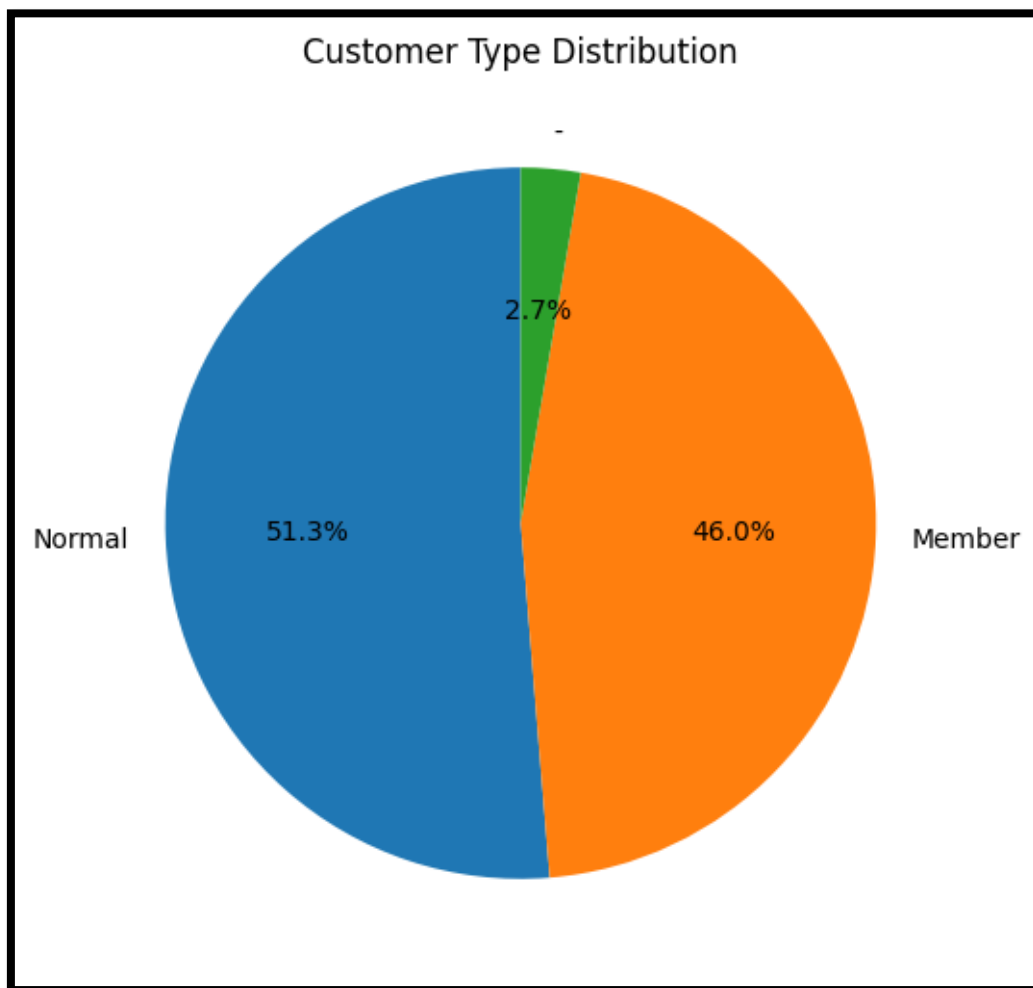
This project aims to analyze and visualize supermarket sales data to gain insights into customer behavior, sales trends, and product performance. The project involves data cleaning, transformation, and visualization using Python (Colab Notebook) libraries such as Pandas, NumPy and Matplotlib.

	Invoice ID	Branch	Yangon	Naypyitaw	Mandalay	Customer type	Gender	Product line	Unit price	Quantity	Tax 5%	Total	Date	Time	Payment	Rating
0	750-67-8428	A	1	0	0	Normal	Male	Health and beauty	74.69	7	26.1415	NaN	1/5/2019	13:08	Ewallet	9.1
1	226-31-3081	C	0	1	0	Normal	Male	Electronic accessories	15.28	5	3.8200	80.2200	3/8/2019	10:29	Cash	9.6
2	631-41-3108	A	1	0	0	Normal	Male	Home and lifestyle	46.33	7	16.2155	340.5255	3/3/2019	13:23	Credit card	7.4
3	123-19-1176	A	1	0	0	Normal	Male	Health and beauty	58.22	8	NaN	489.0480	1/27/2019	8 - 30 PM	Ewallet	8.4
4	373-73-7910	A	1	0	0	Normal	Male	Sports and travel	86.31	7	30.2085	634.3785	2/8/2019	10:37	Ewallet	5.3
...
1001	861-77-0145	C	0	1	0	Member	Male	Electronic accessories	81.97	10	40.9850	860.6850	3/3/2019	14:30	Cash	9.2
1002	479-26-8945	B	0	0	1	Member	Female	Sports and travel	16.49	2	1.6490	34.6290	2/5/2019	11:32	Ewallet	4.6
1003	210-67-5886	C	0	1	0	Member	Female	Health and beauty	98.21	3	14.7315	309.3615	2/5/2019	10:41	Credit card	7.8
1004	227-78-1148	B	0	0	1	Normal	Female	Fashion accessories	72.84	7	25.4940	535.3740	2/15/2019	12:44	Cash	8.4
1005	645-44-1170	A	1	0	0	Member	Male	Home and lifestyle	58.07	9	26.1315	548.7615	1/19/2019	20:07	Ewallet	4.3

1006 rows x 16 columns

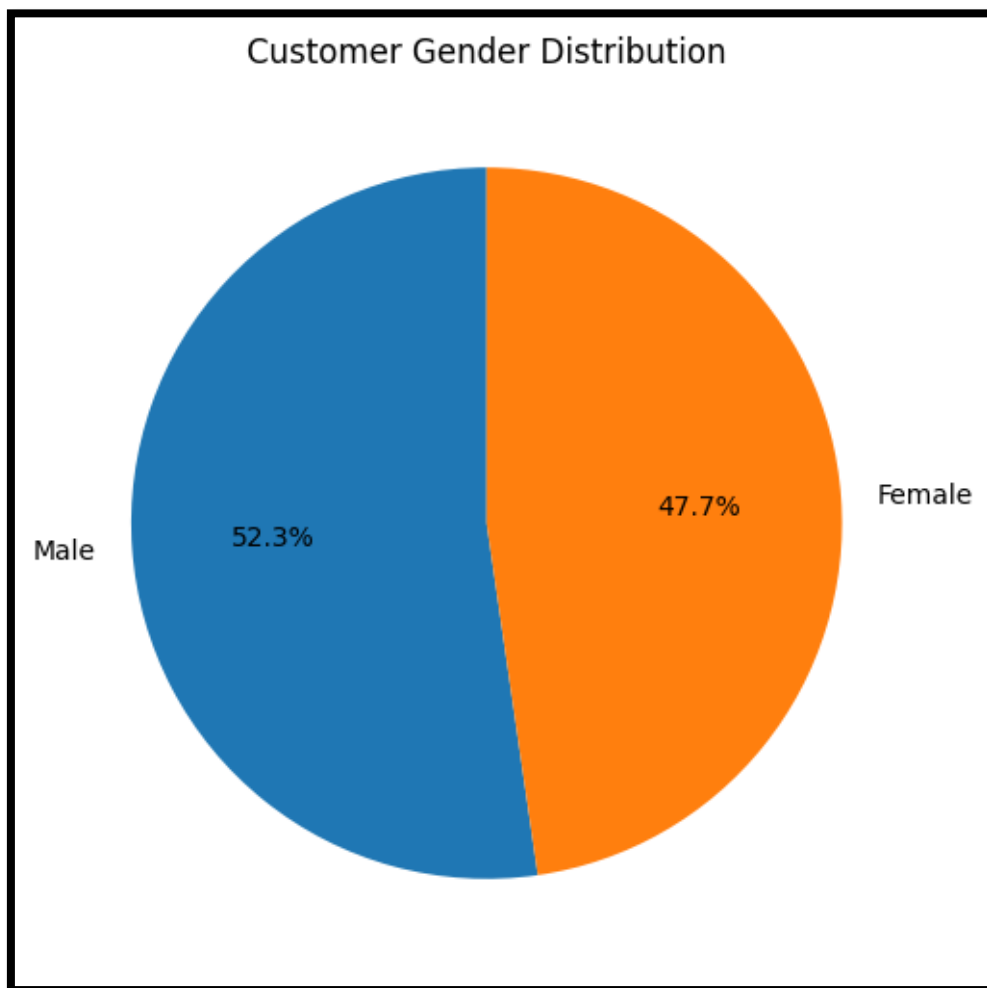
❖ Data Visualization:

1. Customer Type Distribution:



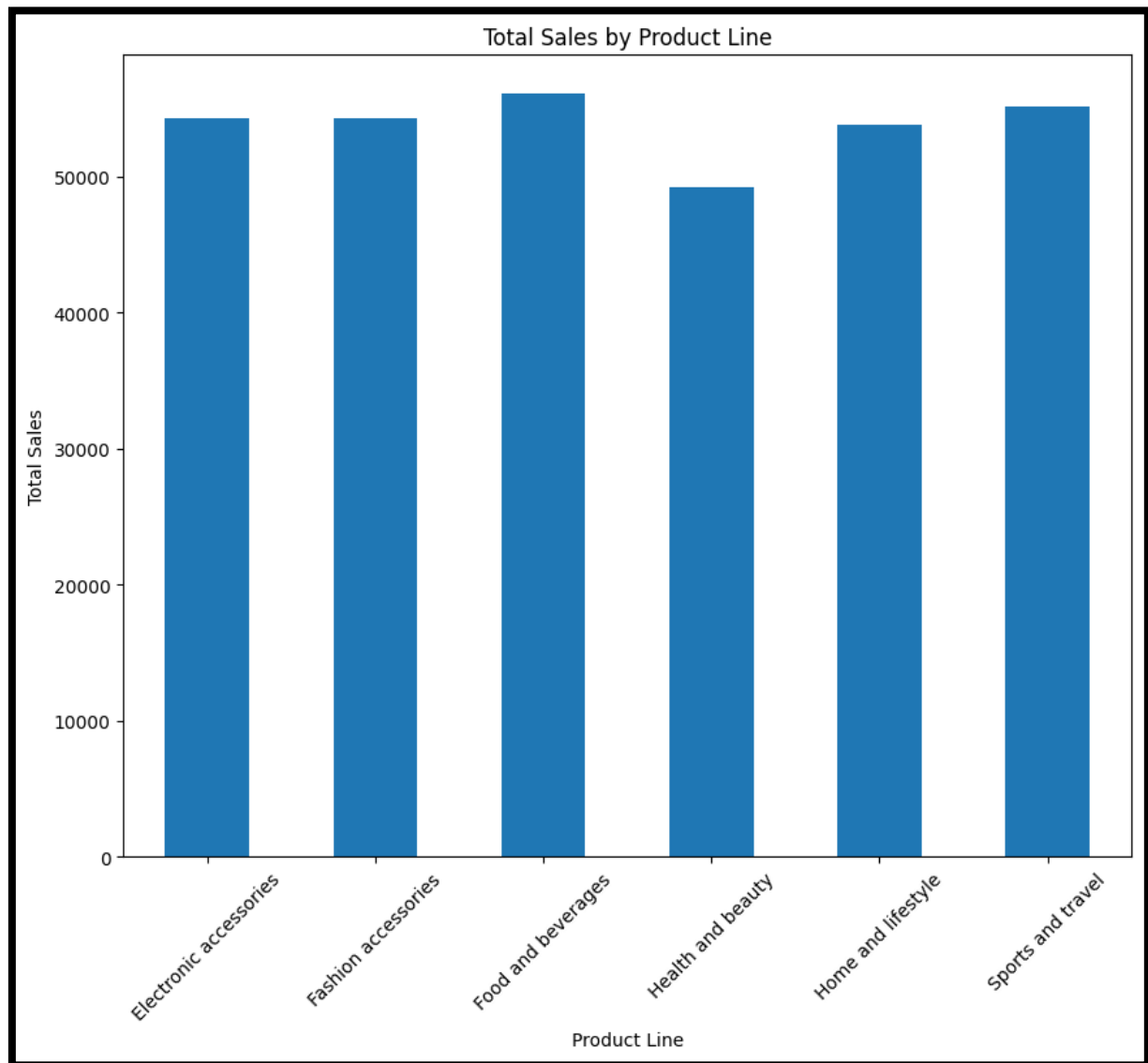
As we see, Normal customers are more than Member customers, but I was unable to know what the meaning of the dash (-).

2. Customer Gender Distribution:



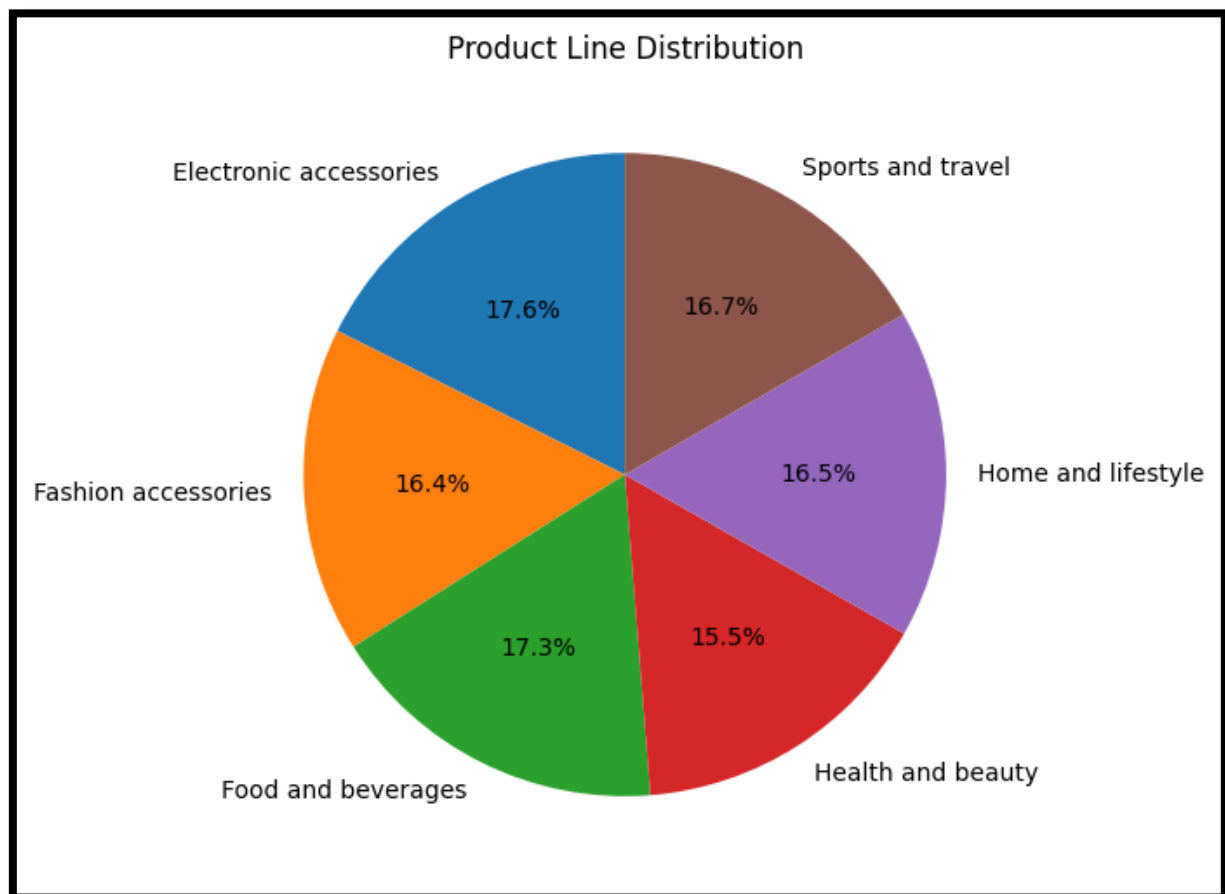
The Males customers is more than the Females Customers.

3. Total Sales by Product Line:



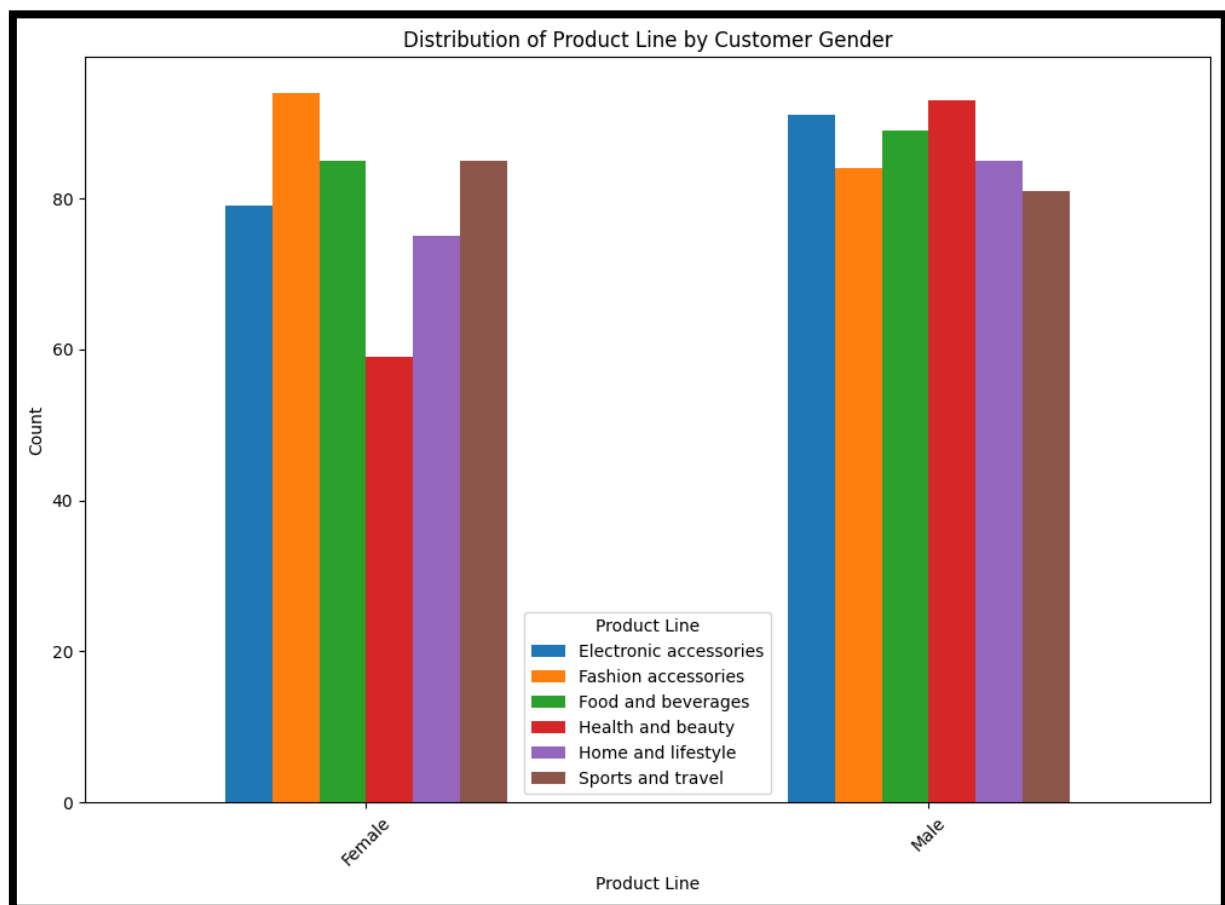
The Food and beverages line has the most total sales (Unit price * Quantity), but there are slight differences between the Product lines total sales.

4. Product Line Distribution:



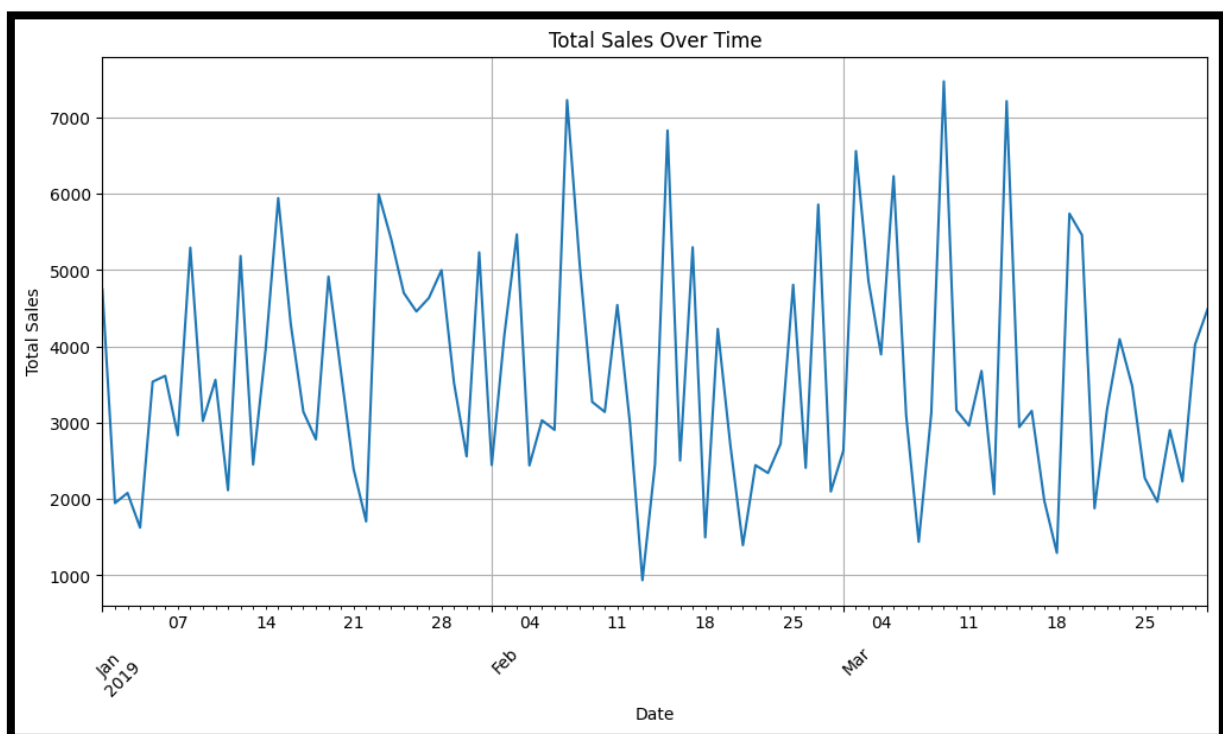
Although the Food and beverages line have the most total sales, but Electronic accessories have the most Quantity sold.

5. Total Product Line by Customer Gender:



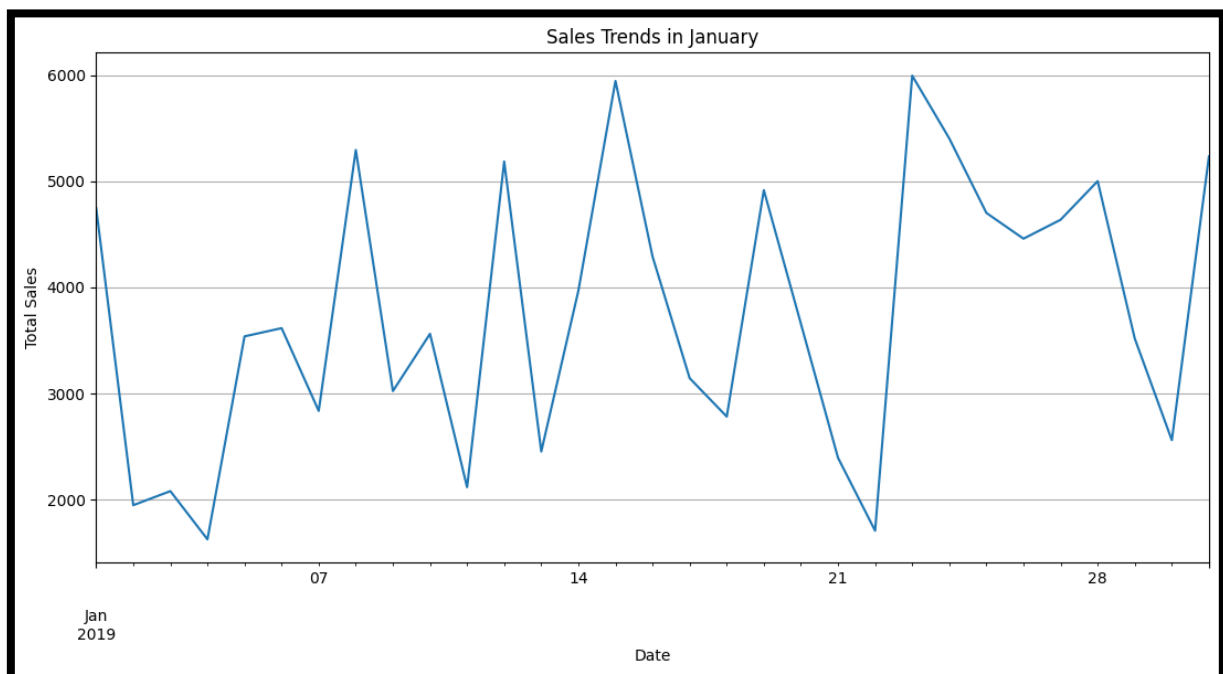
We can see that Females make most purchases from Fashion accessories line (What a Surprise). When Males make most purchases from Health and beauty.

6. Total Sales Over Time:



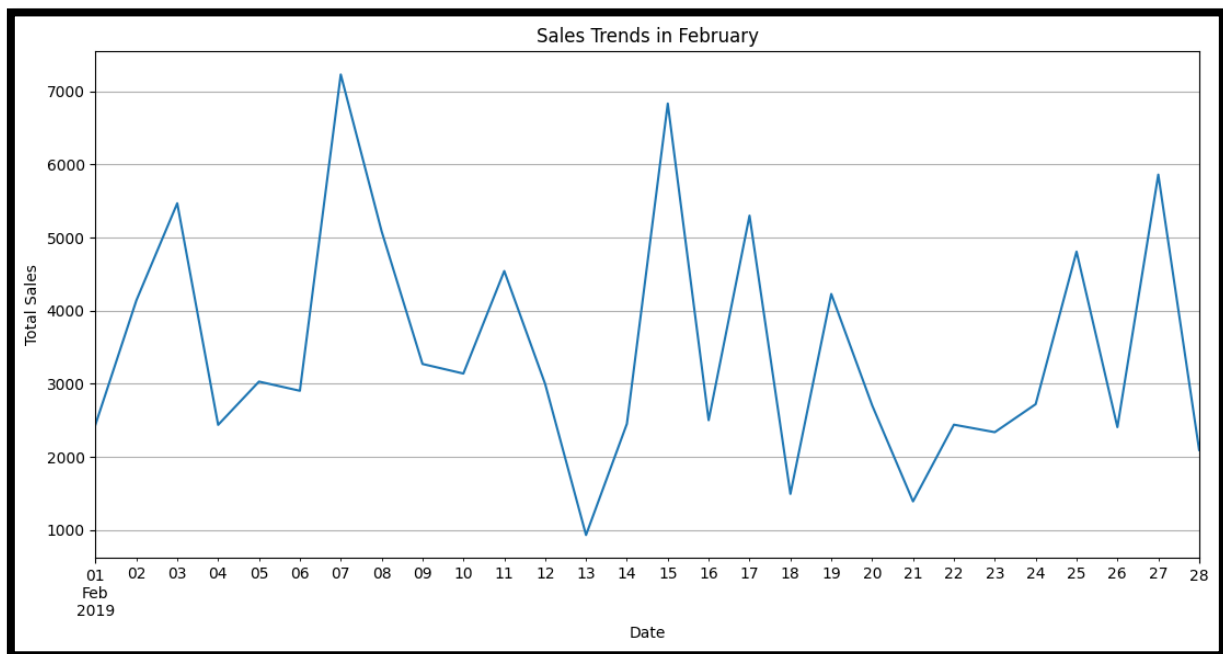
The Highest Sales made in 09/03/2019.

7. Sales Trends in January:



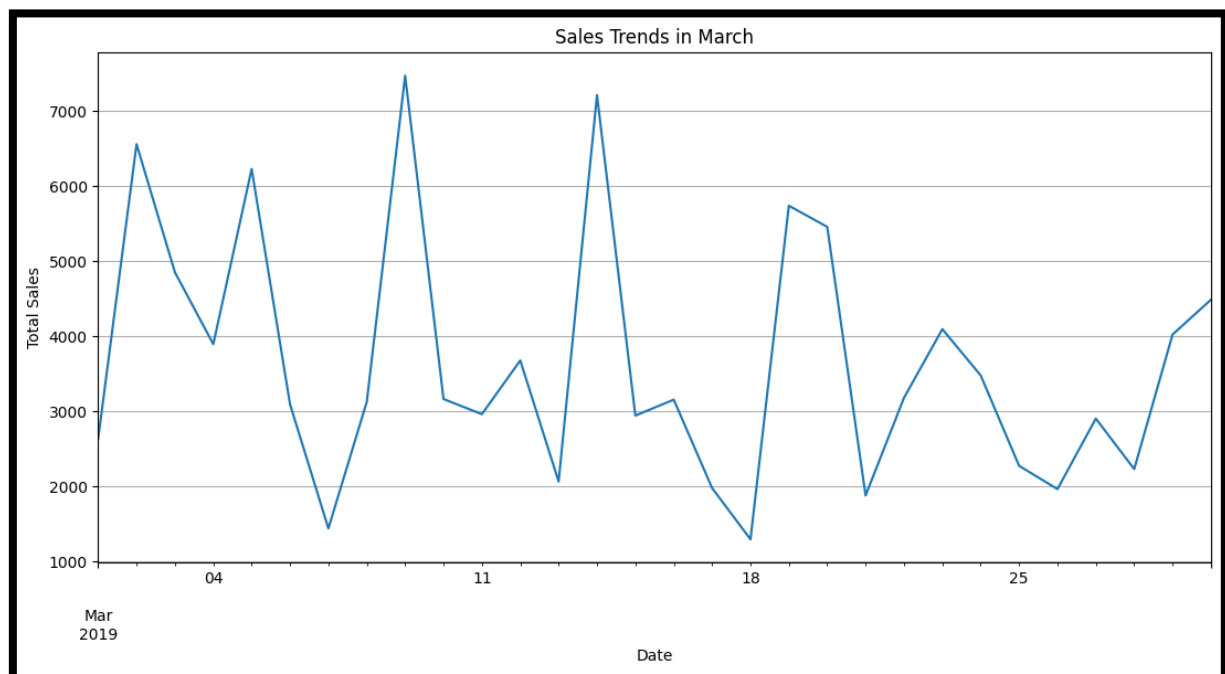
The Highest Sales in January made in 23/01/2019.

8. Sales Trends in February:



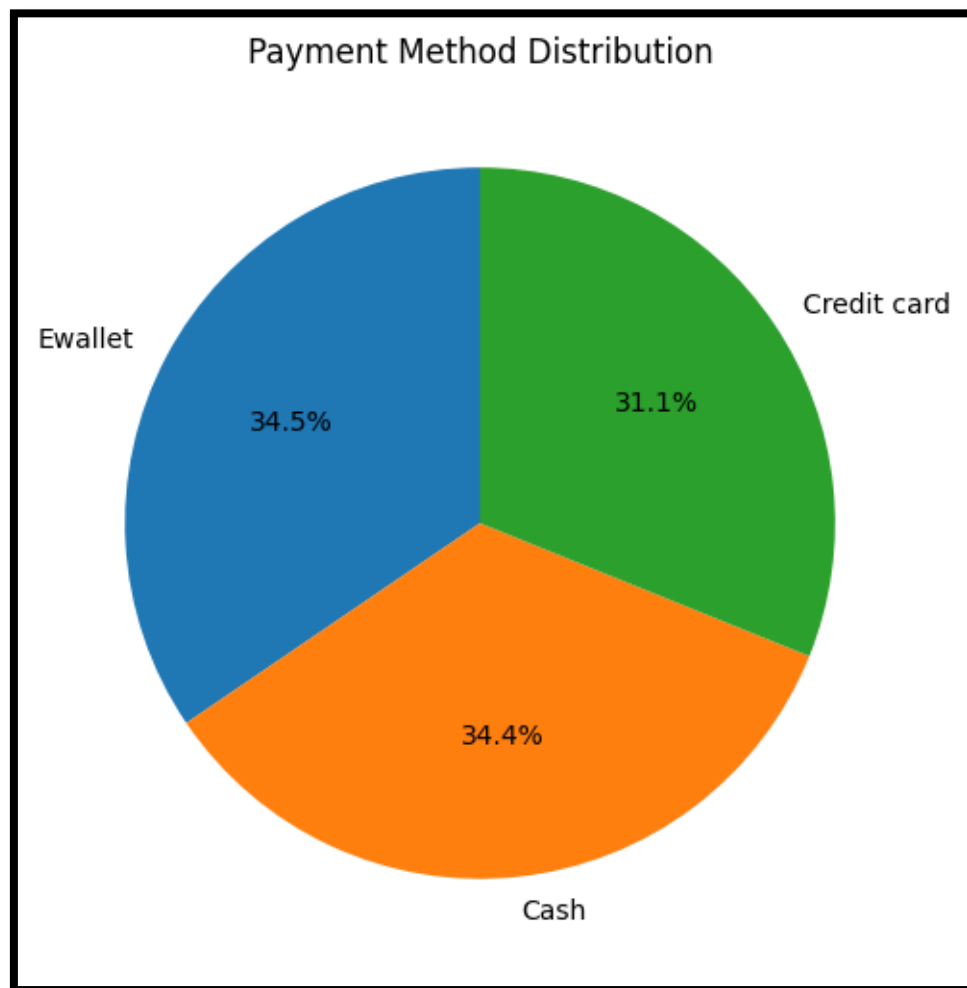
The Highest sales in February made in 07/02/2019.

9. Sales Trends in March:



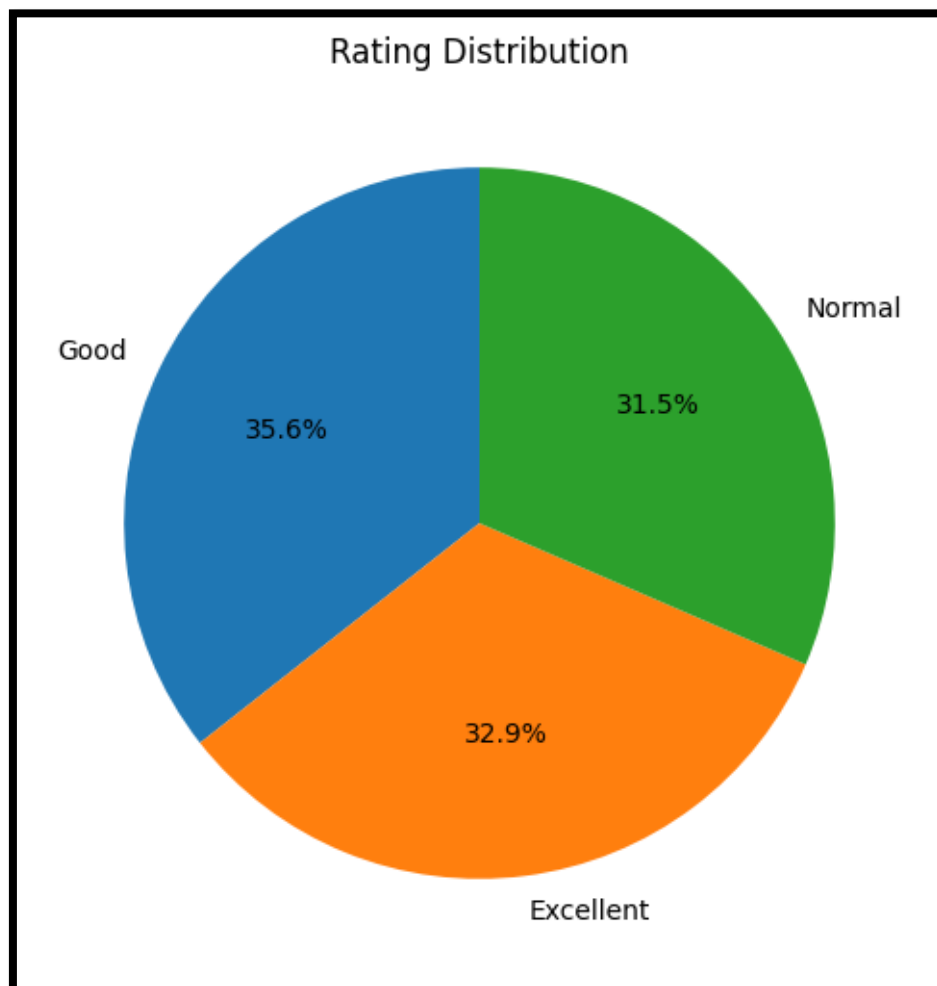
The Highest sales in March made in 09/03/2019.

10. Payment Method Distributions:



Most payment made by Ewallet.

11. Rating Distribution:

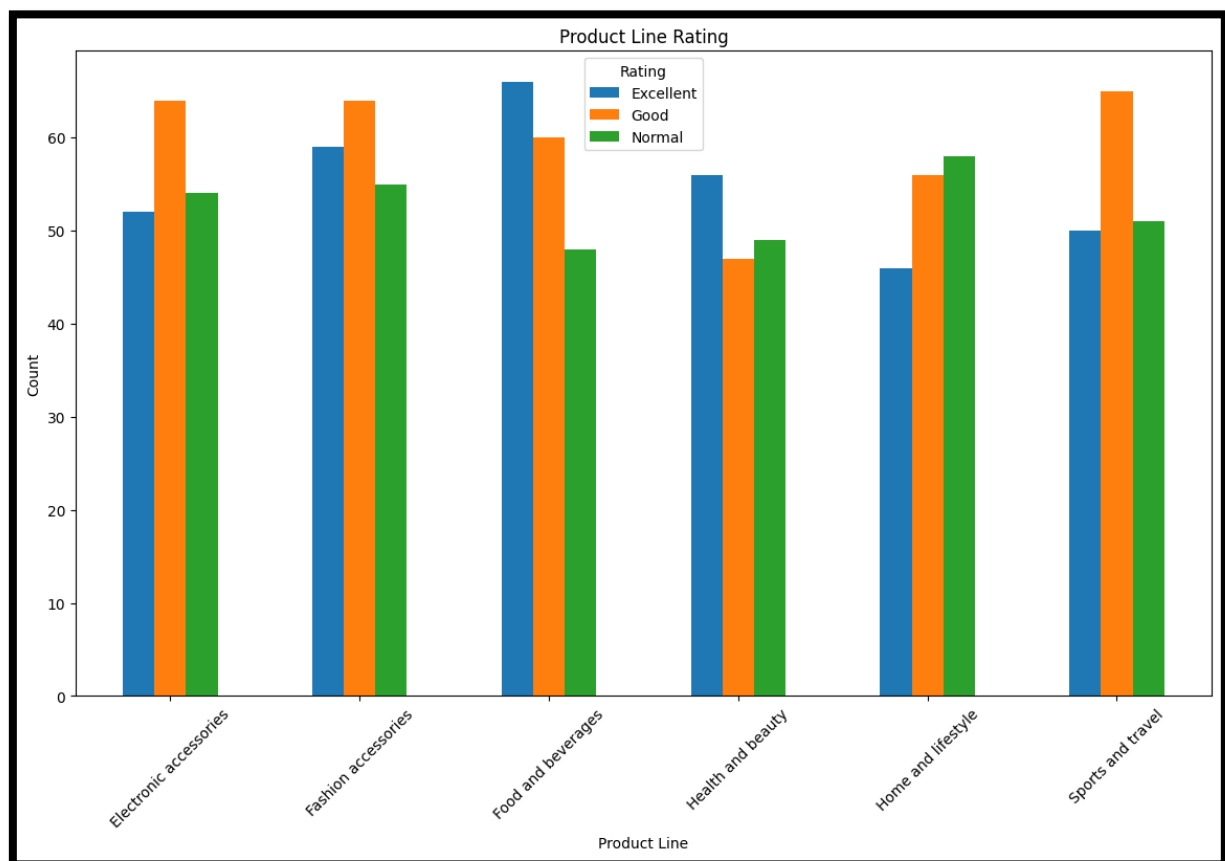


I have divided the Rating into categories:

Bad: (0-4), Normal: (4-6), Good: (6-8), Excellent: (8-10)

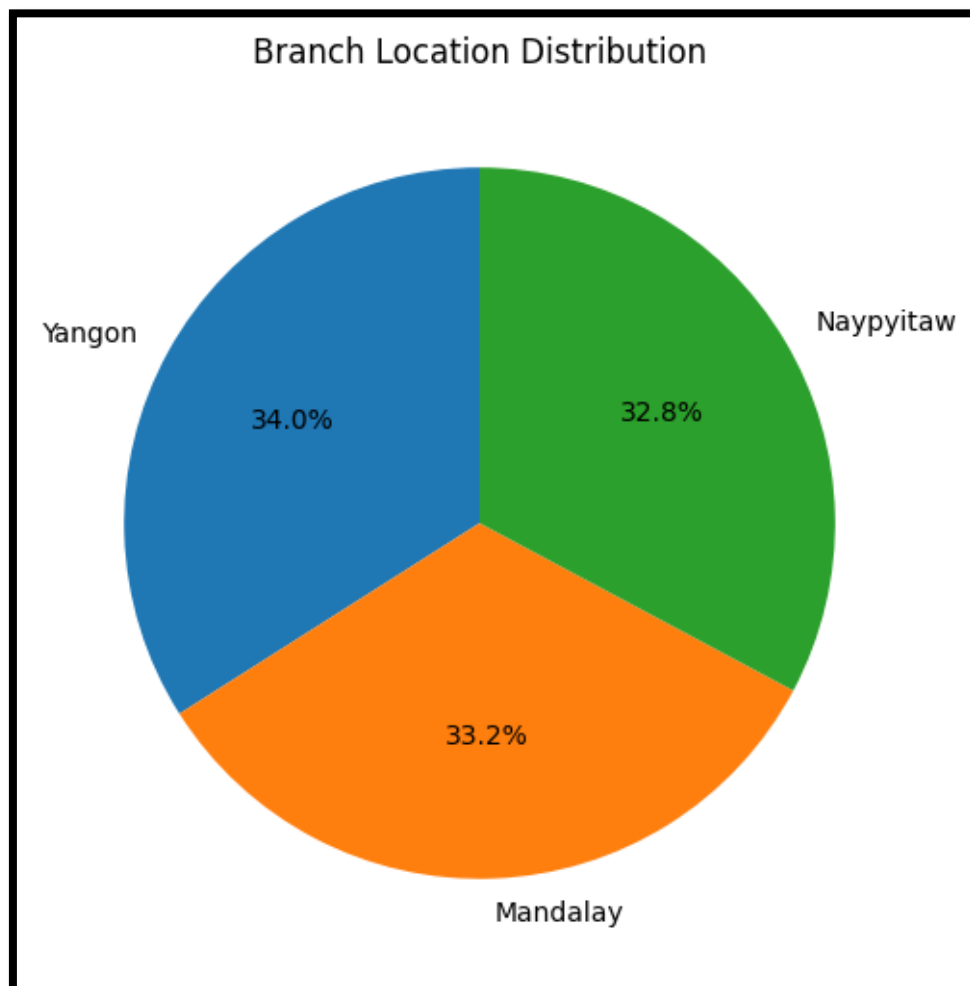
Good is the most Rating and there is no Rating under 4 (Bad).

12. Product Line Rating:



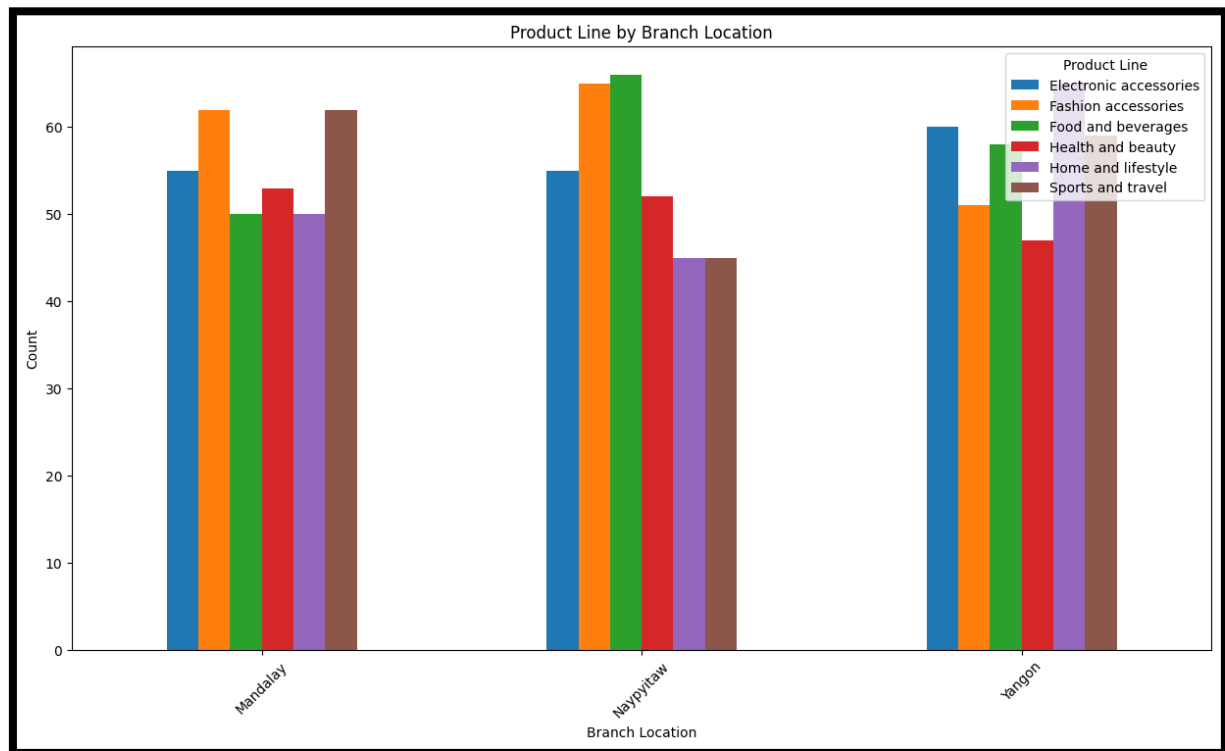
The customers are happy with Food and beverages line and Health and beauty line.

13. Branch Location Distribution:



The Yangon branch has the highest sales.

14. Product Line by Branch Location:



Mandalay Branch: Fashion accessories and Sports and travel have the highest sales.

Naypyitaw Branch: Food and Beverages has the highest sales.

Yangon Branch: Home and lifestyle has the highest sales.

Thank You