SQL Project PPT

Task 1:

Walmart wants to identify which branch has exhibited the highest sales growth over time. Analyze the total sales for each branch and compare the growth rate across months to find the top performer.

- SELECT Branch, MONTH(date) AS month_num, gross_margin_percentage, LAG(gross_margin_percentage) OVER (PARTITION BY Branch ORDER BY date) AS previous_margin, ((gross_margin_percentage - LAG(gross_margin_percentage) OVER (PARTITION BY Branch ORDER BY date))) * 100 AS growth_rate FROM walmartsales_dataset ORDER BY Branch, date;

Output

Branch	month_num	gross_margin_percentage	previous_margin	growth_rate
A	1	4.761904762	Null	Null
A	1	4.761904762	4.761904762	0
A	1	4.761904762	4.761904762	0
A	1	4.761904762	4.761904762	0
A	1	4.761904762	4.761904762	0
A	1	4.761904762	4.761904762	0
A	1	4.761904762	4.761904762	0
A	1	4.761904762	4.761904762	0
A	1	4.761904762		

Task 2:

Walmart needs to determine which product line contributes the highest profit to each branch.The profit margin should be calculated based on the difference between the gross income and cost of goods sold.

- select Product_line,round(sum(cogs-gross_income),2) AS Profit from walmartsales_datasetgroup by Product_line order by Profit desc;

Product_line	Profit
Food and beverages	50797.72
Sports and travel	49873.03
Electronic accessories	49162.53
Fashion accessories	49133.91
Home and lifestyle	48732.21
Health and beauty	44508.62

- According to Output we can say that the Top 1st Profit Making product is **Food and beverages** and second One is **Sports and Travel**.

Task 2: Chart


Profit
52000
51000
50000
49000
48000
47000
46000
45000
44000
43000
42000
41000
Profit
50797.72
49873.03
49162.53
49133.91
48732.21
44508.62
Food and beverages
Sports and travel
Electronic accessories
Fashion accessories
Home and lifestyle
Health and beauty
Product

Task 3:

Walmart wants to segment customers based on their average spending behavior. Classify customers into three tiers: High, Medium, and Low spenders based on their total purchase amounts.

- According to Question I give **Low Category to 19700** and **between 19701 and 22700** I Give Medium and for other I give **High**.
- select Customer_ID,round(sum(Total),2) AS Total_Spend, case when round(sum(Total),2) <= 19700 then 'Low' when round(sum(Total),2) between 19701 and 22700 then 'Medium' else 'High'end AS Category from walmartsales_dataset group by Customer_ID;

Customer_ID	Total_Spend	Category
2	23392.28	High
3	23402.26	High
11	21398.82	Medium
9	19661.6	Low
13	21063.66	Medium
1	22634.55	Medium
8	26634.34	High
12	21720.65	Medium
14	21049.4	Medium


Customer Segmentation
30000
25000
20000
15000
10000
5000
0
Total Spend
1
2
3
8
9
11
12
13
14
Customer Id

Task 4:

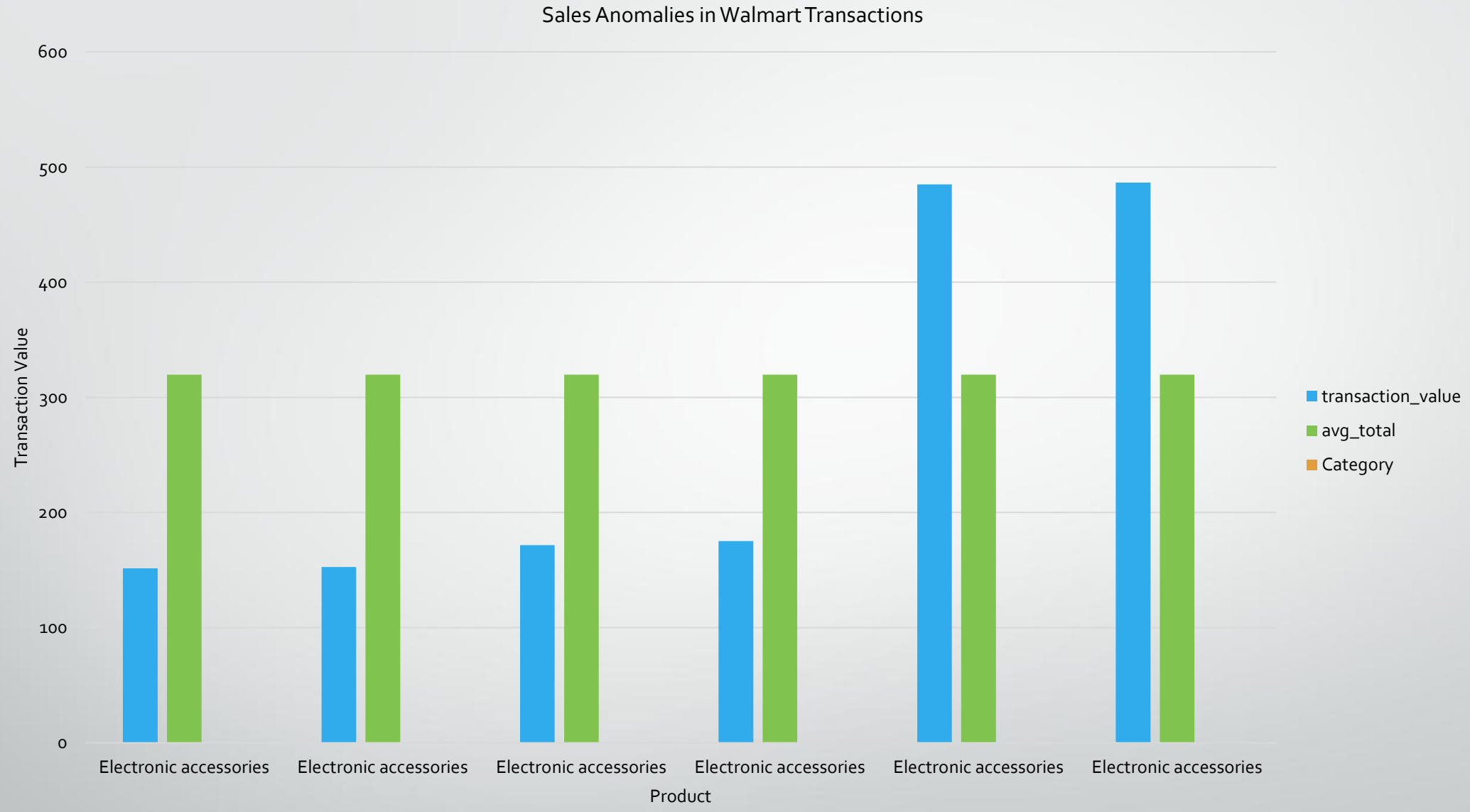
Walmart suspects that some transactions have unusually high or low sales compared to the average for the product line. Identify these anomalies.

- WITH ProductLineStats AS (SELECT Product_line,AVG(Total) AS avg_total FROM walmartsales_dataset GROUP BY Product_line) SELECT w.Product_line,w.Total AS transaction_value,pl.avg_total,CASE WHEN w.Total > pl.avg_total * 1.5 THEN 'High' WHEN w.Total < pl.avg_total * 0.5 THEN 'Low' ELSE 'Normal'END AS CategoryFROM walmartsales_dataset wJOIN ProductLineStats plON w.Product_line = pl.Product_lineORDER BY w.Product_line, w.Total;

- This query calculates the average sales for each product line and categorizes transactions as 'High,' 'Low,' or 'Normal' based on their value compared to the average. It then sorts the results by product line and transaction amount.

Task 4: output

Product_line	transaction_value	avg_total	Category
Electronic accessories	151.4835	319.63253823529413	Low
Electronic accessories	152.712	319.63253823529413	Low
Electronic accessories	171.7275	319.63253823529413	Normal
Electronic accessories	175.14	319.63253823529413	Normal
Electronic accessories	484.974	319.63253823529413	High
Electronic accessories	486.444	319.63253823529413	High

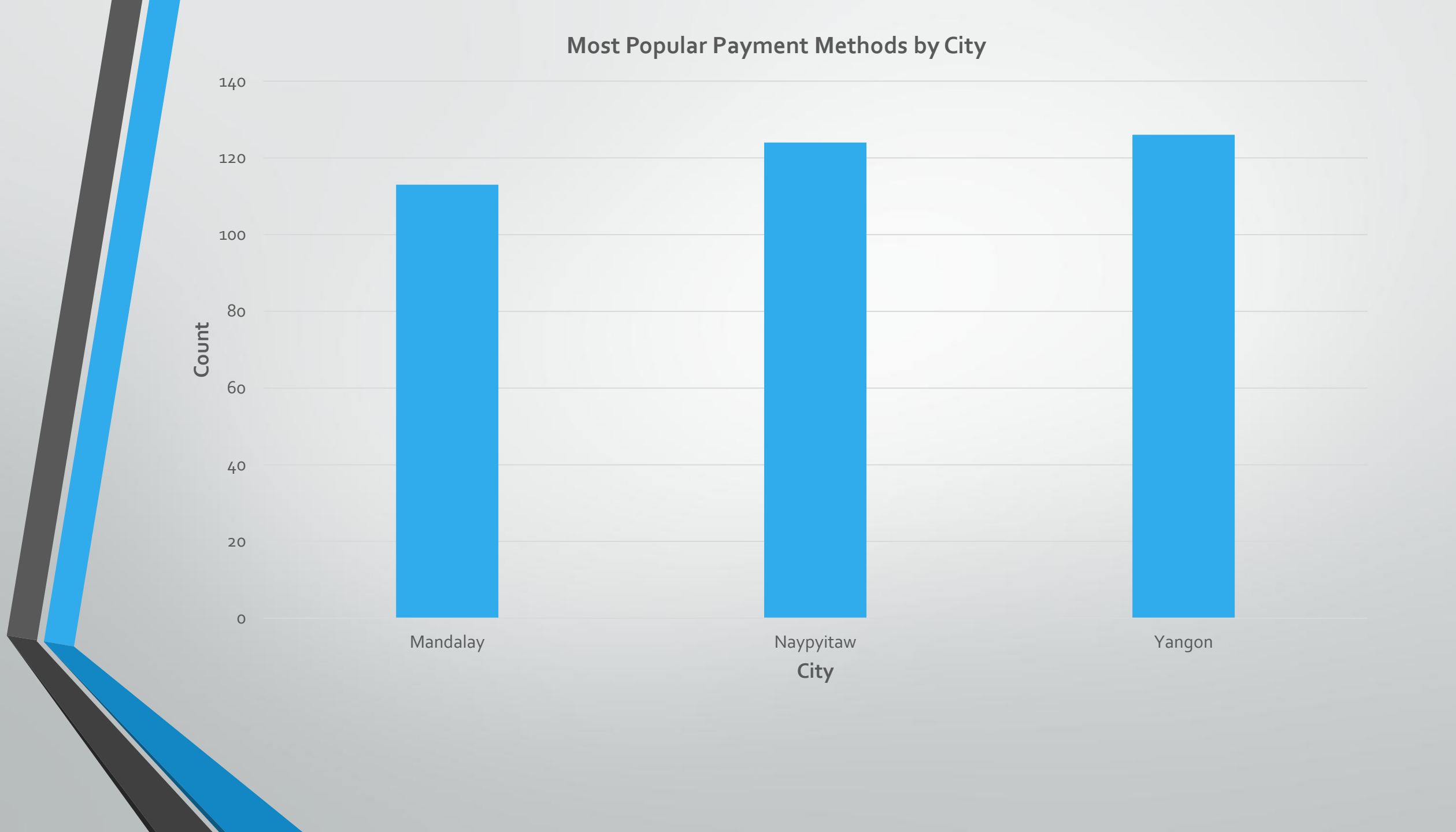

Sales Anomalies in Walmart Transactions
600
500
400
300
200
100
0
Transaction Value
Electronic accessories
Electronic accessories
Electronic accessories
Electronic accessories
Electronic accessories
Electronic accessories
Product
transaction_value
avg_total
Category

Task 5:

Walmart needs to determine the most popular payment method in each city to tailor marketing strategies.

- select city,Payment,count(*) AS Count from walmartsales_datasetgroup by City,Paymentorder by Count desc,city;

city	Payment	Count
Mandalay	Ewallet	113
Naypyitaw	Cash	124
Yangon	Ewallet	126

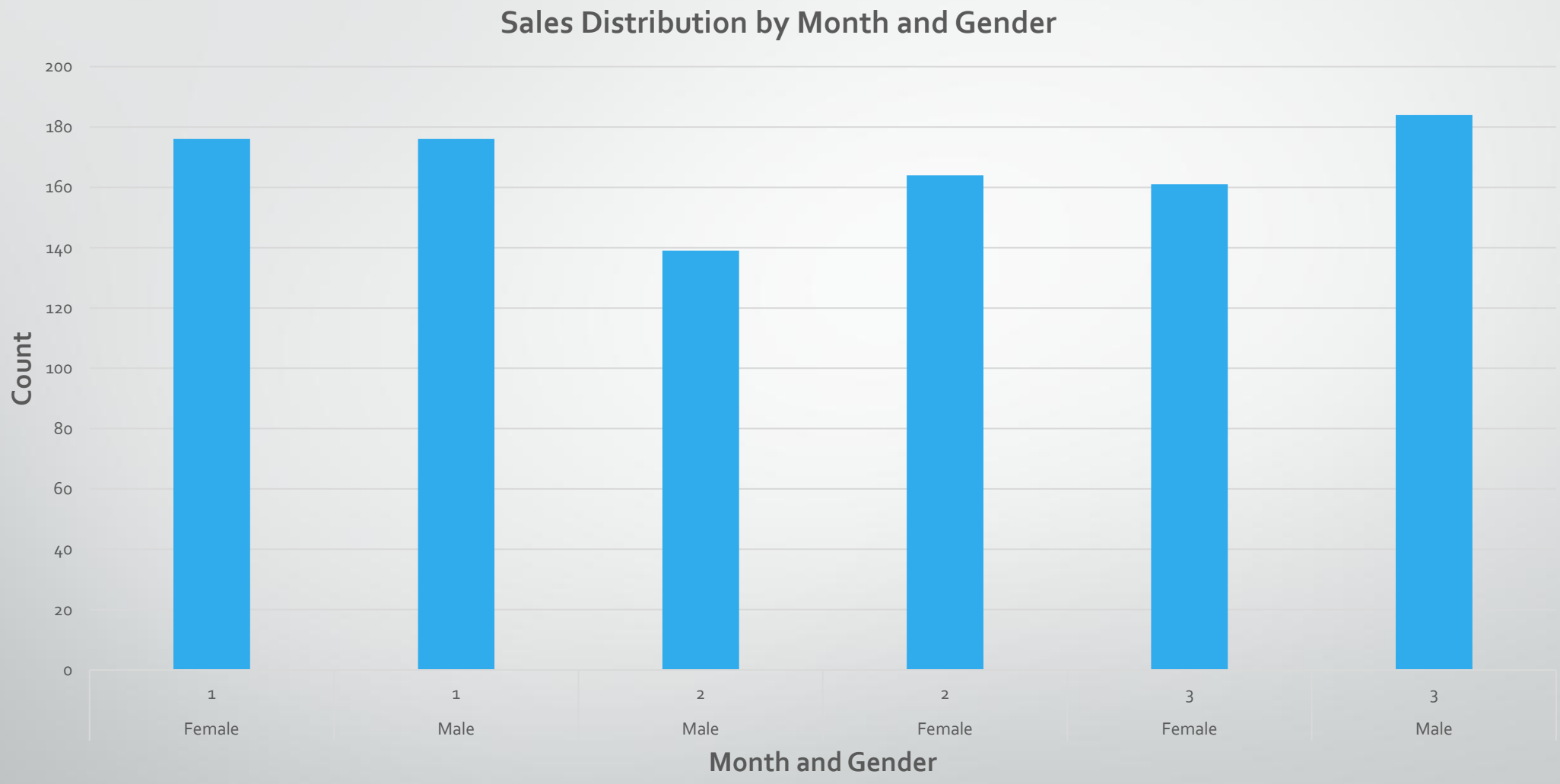

Most Popular Payment Methods by City
140
120
100
80
60
40
20
0
Count
Mandalay
Naypyitaw
Yangon
City

Task 6:

Walmart wants to understand the sales distribution between male and female customers on a monthly basis.

- select Gender,month(date) as month,count(*) AS count from walmartsales_datasetgroup by Gender,month(date)order by month;

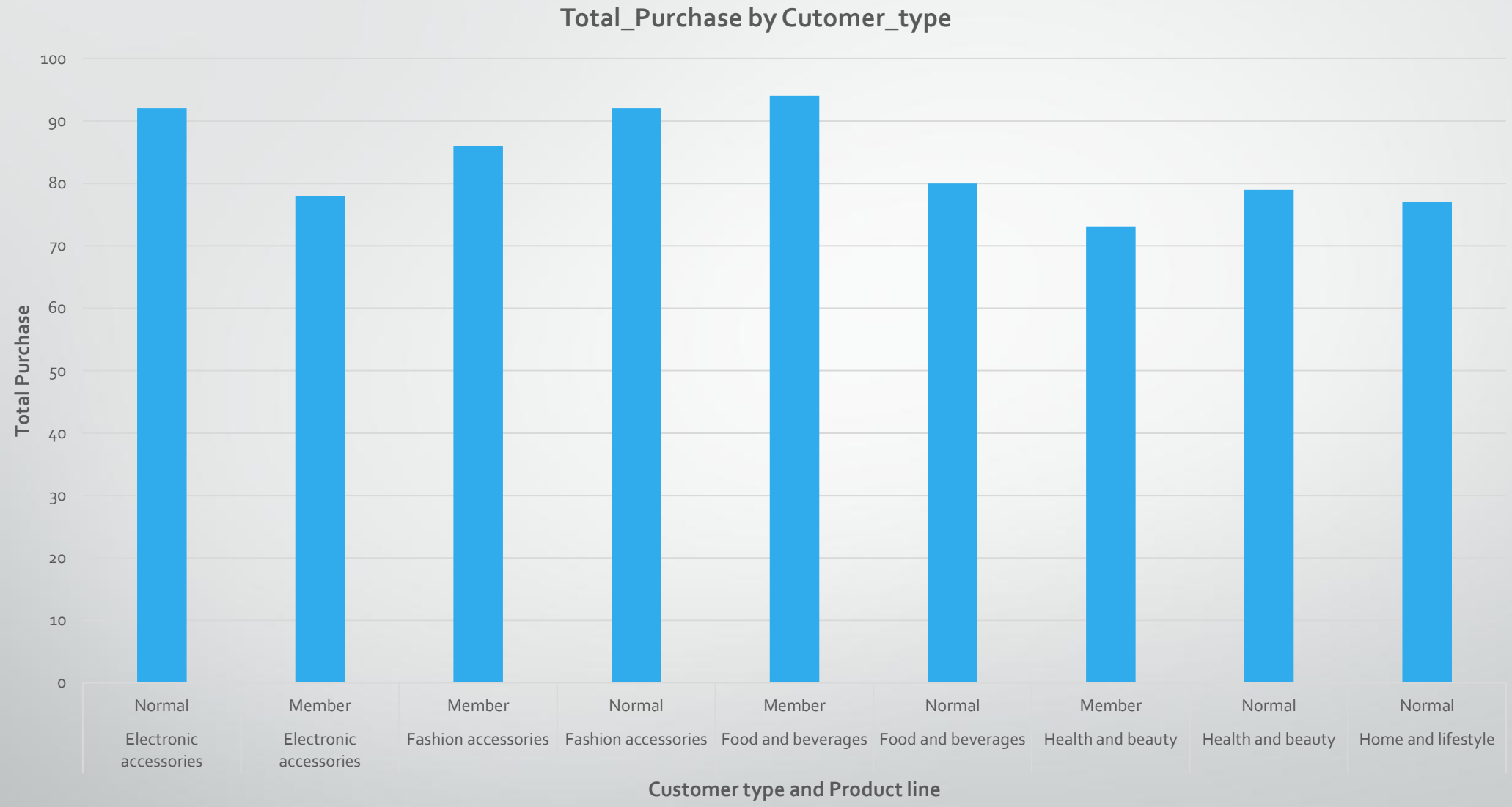
Gender	month	count
Female	1	176
Male	1	176
Male	2	139
Female	2	164
Female	3	161
Male	3	184


Sales Distribution by Month and Gender
Count
200
180
160
140
120
100
80
60
40
20
0
1
Female
1
Male
2
Male
2
Female
3
Female
3
Male
Month and Gender

Task 7:
Walmart wants to know which product lines are preferred by different customer types(Member vs. Normal).

- select Product_line,Customer_type,count(*) AS Total_Purchase from walmartsales_datasetgroup by Customer_type,Product_lineorder by Product_line;

Product_line	Customer_type	Total_Purchase
Electronic accessories	Normal	92
Electronic accessories	Member	78
Fashion accessories	Member	86
Fashion accessories	Normal	92
Food and beverages	Member	94
Food and beverages	Normal	80
Health and beauty	Member	73
Health and beauty	Normal	79
Home and lifestyle	Normal	77


Total_Purchase by Cutomer_type
Total Purchase
100
90
80
70
60
50
40
30
20
10
0
Normal
Electronic accessories
Member
Electronic accessories
Member
Fashion accessories
Normal
Fashion accessories
Member
Food and beverages
Normal
Food and beverages
Member
Health and beauty
Normal
Health and beauty
Normal
Home and lifestyle
Customer type and Product line

Task 8:
Walmart needs to identify customers who made repeat purchases within a specific time frame (e.g., within 30 days).

- SELECT customer_id,date,next_purchase_date FROM (SELECT customer_id,date,LEAD(date) OVER (PARTITION BY customer_id ORDER BY date) AS next_purchase_dateFROM walmartsales_dataset) AS abcWHERE DATEDIFF(next_purchase_date, date) <= 30;

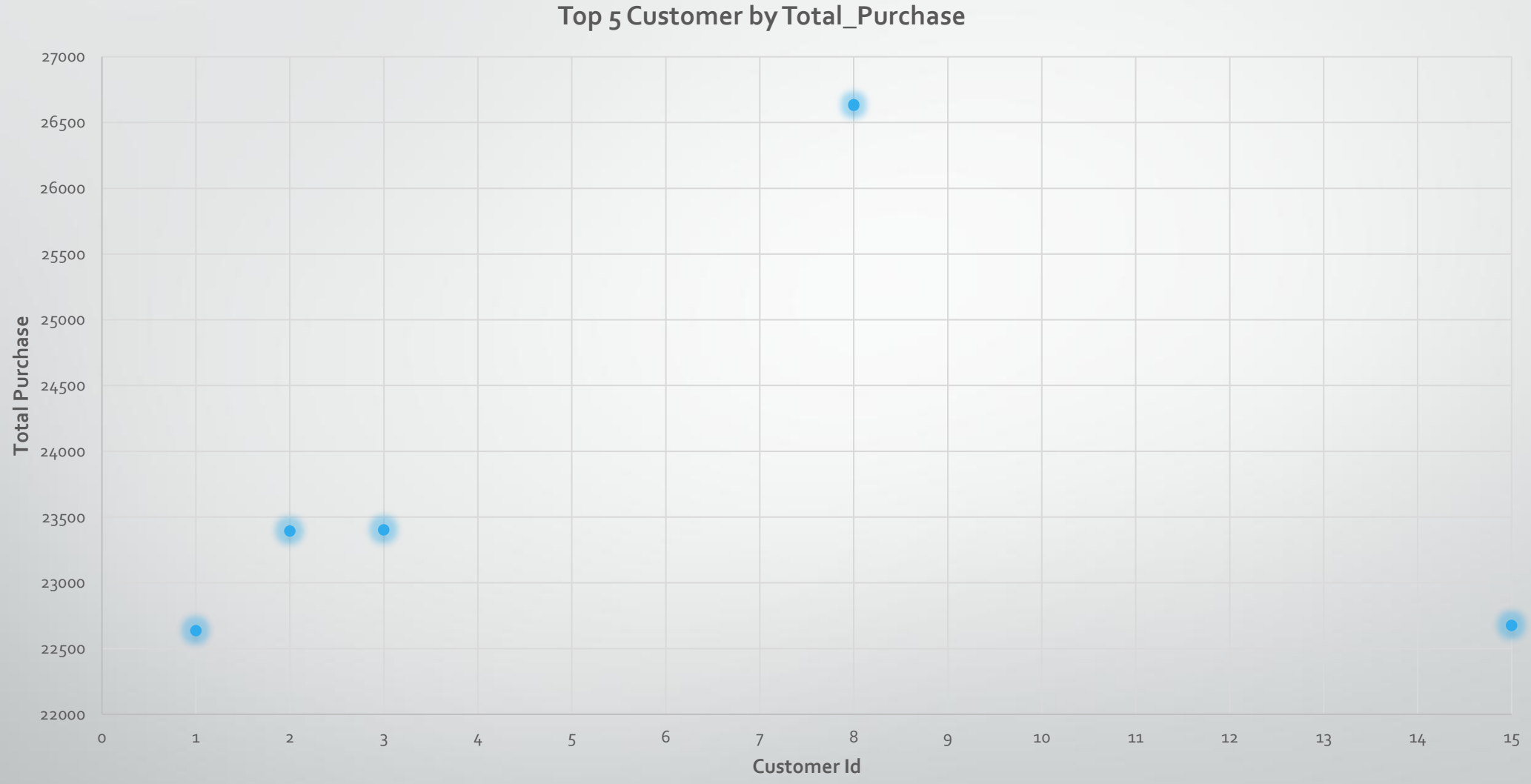
customer_id	date	next_purchase_date
1	2019-03-29	2019-03-30
1	2019-03-30	2019-03-30
2	2019-01-01	2019-01-01
2	2019-01-01	2019-01-01
3	2019-03-27	2019-03-27
3	2019-03-27	2019-03-30
4	2019-01-02	2019-01-03
4	2019-01-03	2019-01-05
5	2019-02-19	2019-02-20
5	2019-02-20	2019-02-21

Task 9:

Walmart wants to reward its top 5 customers who have generated the most sales Revenue.

- select Customer_ID,round(sum(Total),2) AS Total_Purchase from walmartsales_dataset group by Customer_ID order by total_purchase desc limit 5;

Customer_ID	Total_Purchase
8	26634.34
3	23402.26
2	23392.28
15	22674.46
1	22634.55


Top 5 Customer by Total_Purchase
Total Purchase
27000
26500
26000
25500
25000
24500
24000
23500
23000
22500
22000
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Customer Id

Task 10:

Walmart wants to analyze the sales patterns to determine which day of the week brings the highest sales.

- select dayofweek(date) AS Day_Of_Week,dayname(date) AS Day_Name,round(sum(Total),2) AS Total_Sale from walmartsales_datasetgroup by Day_Of_Week,day_nameorder by total_sale desc;

Day_Of_Week	Day_Name	Total_Sale
7	Saturday	56120.81
3	Tuesday	51482.25
5	Thursday	45349.25
1	Sunday	44457.89
6	Friday	43926.34
4	Wednesday	43731.14
2	Monday	37899.08


Total_Sale by Day of Week
Total Sale
100
90
80
70
60
50
40
30
20
10
0
Normal
Electronic accessories
Member
Electronic accessories
Member
Fashion accessories
Normal
Fashion accessories
Member
Food and beverages
Normal
Food and beverages
Member
Health and beauty
Day of Week