Walmart sales data analysis using SQL

Objective: The primary objective of this project is to analyze Walmart's sales data to identify and understand the various factors influencing sales across its different branches.

Dataset description:

Column	Description	Data Type
invoice_id	Invoice of the sales made	VARCHAR(30)
branch	Branch at which the sales were made	VARCHAR(5)
city	Location of branch	VARCHAR(30)
customer_type	Type of customer	VARCHAR(30)
gender	Gender of the customer	VARCHAR(10)
product_line	Product line of the product sold	VARCHAR(100)
unit_price	Price of each product	DECIMAL(10,2)
quantity	Amount of the product sold	INT
VAT	Amount of tax on the purchase	FLOAT(6,4)
total	Total cost of purchase	DECIMAL(10,2)
date	Purchase date	DATE
time	Purchase time	TIMESTAMP
payment_method	Mode of payment	VARCHAR(15)
cogs	Cost of goods sold	DECIMAL(10,2)
gross_margin_percentage	Gross margin percentage	FLOAT(11,9)
gross_income	Gross income	DECIMAL(10,2)
rating	Rating	FLOAT(2,1)

The dataset was obtained from Kaggle.

TABLE CREATION:

```
CREATE TABLE IF NOT EXISTS sales(
            invoice_id VARCHAR(30) NOT NULL PRIMARY KEY,
    branch VARCHAR(5) NOT NULL,
    city VARCHAR(30) NOT NULL,
    customer_type VARCHAR(30) NOT NULL,
    gender VARCHAR(30) NOT NULL,
    product_line VARCHAR(100) NOT NULL,
    unit_price DECIMAL(10,2) NOT NULL,
    quantity INT NOT NULL,
    VAT FLOAT(6,4) NOT NULL,
    total DECIMAL(12, 4) NOT NULL,
    date DATETIME NOT NULL,
    time TIME NOT NULL,
    payment VARCHAR(15) NOT NULL,
    cogs DECIMAL(10,2) NOT NULL,
    gross_margin_pct FLOAT(11,9),
    gross_income DECIMAL(12, 4),
    rating FLOAT(2, 1)
);
```

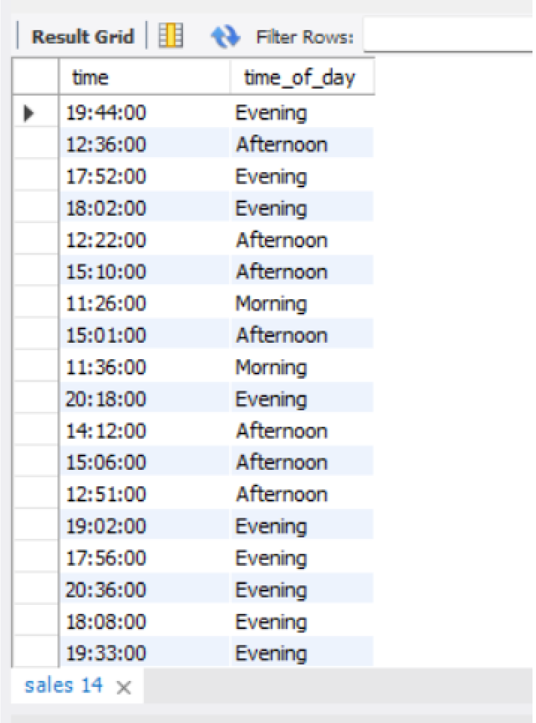
Feature Engineering

1. Add a new column named `time_of_day` to give insight of sales in the Morning, Afternoon and Evening.

CODE:

```
SELECT time, (CASE
              WHEN `time` between "00:00:00" AND "12:00:00" THEN "Morning"
              WHEN `time` between "12:01:00" AND "16:00:00" THEN "Afternoon"
              ELSE "Evening"
              END) AS time_of_day from sales;
ALTER TABLE sales ADD COLUMN time_of_day VARCHAR(20);

UPDATE sales
SET time_of_day = (
          CASE
          WHEN `time` between "00:00:00" AND "12:00:00" THEN "Morning"
          WHEN `time` between "12:01:00" AND "16:00:00" THEN "Afternoon"
          ELSE "Evening"
          END);
SELECT time, time_of_day FROM sales;
```

OUTPUT:



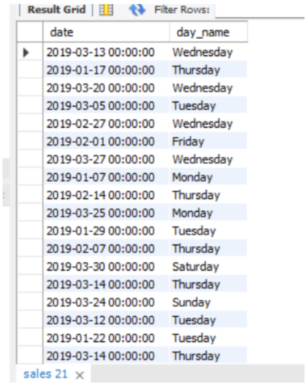
time	time_of_day
19:44:00	Evening
12:36:00	Afternoon
17:52:00	Evening
18:02:00	Evening
12:22:00	Afternoon
15:10:00	Afternoon
11:26:00	Morning
15:01:00	Afternoon
11:36:00	Morning
20:18:00	Evening
14:12:00	Afternoon
15:06:00	Afternoon
12:51:00	Afternoon
19:02:00	Evening
17:56:00	Evening
20:36:00	Evening
18:08:00	Evening
19:33:00	Evening

2. Add a new column named `day_name` that contains the extracted days of the week on which the given transaction took place (Mon, Tue, Wed, Thur, Fri). This will help answer the question on which week of the day each branch is busiest.

CODE:

```
SELECT date, DAYNAME(date)  FROM sales;
ALTER TABLE sales ADD COLUMN day_name VARCHAR(10);
UPDATE sales SET day_name = DAYNAME(date);
SELECT date, day_name FROM sales;
```

OUTPUT:



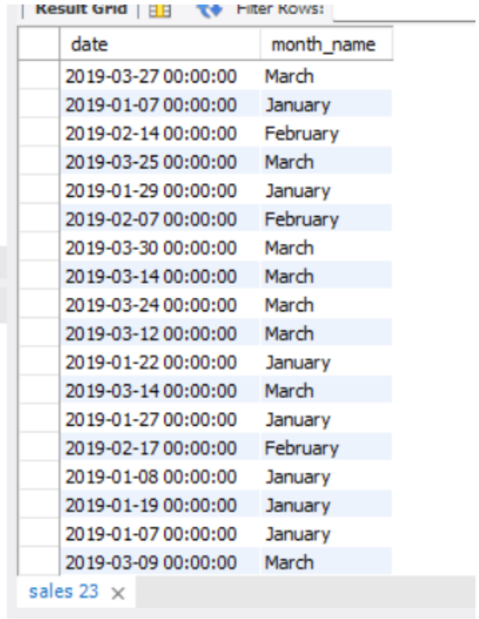
date	day_name
2019-03-13 00:00:00	Wednesday
2019-01-17 00:00:00	Thursday
2019-03-20 00:00:00	Wednesday
2019-03-05 00:00:00	Tuesday
2019-02-27 00:00:00	Wednesday
2019-02-01 00:00:00	Friday
2019-03-27 00:00:00	Wednesday
2019-01-07 00:00:00	Monday
2019-02-14 00:00:00	Thursday
2019-03-25 00:00:00	Monday
2019-01-29 00:00:00	Tuesday
2019-02-07 00:00:00	Thursday
2019-03-30 00:00:00	Saturday
2019-03-14 00:00:00	Thursday
2019-03-24 00:00:00	Sunday
2019-03-12 00:00:00	Tuesday
2019-01-22 00:00:00	Tuesday
2019-03-14 00:00:00	Thursday

3. Add a new column named `month_name` that contains the extracted months of the year on which the given transaction took place (Jan, Feb, Mar). Help determine which month of the year has the most sales and profit.

CODE:

```
SELECT date, MONTHNAME(date)  FROM sales;
ALTER TABLE sales ADD COLUMN month_name VARCHAR(10);
UPDATE sales SET month_name = MONTHNAME(date);
SELECT date, month_name FROM sales;
```

OUTPUT:



date	month_name
2019-03-27 00:00:00	March
2019-01-07 00:00:00	January
2019-02-14 00:00:00	February
2019-03-25 00:00:00	March
2019-01-29 00:00:00	January
2019-02-07 00:00:00	February
2019-03-30 00:00:00	March
2019-03-14 00:00:00	March
2019-03-24 00:00:00	March
2019-03-12 00:00:00	March
2019-01-22 00:00:00	January
2019-03-14 00:00:00	March
2019-01-27 00:00:00	January
2019-02-17 00:00:00	February
2019-01-08 00:00:00	January
2019-01-19 00:00:00	January
2019-01-07 00:00:00	January
2019-03-09 00:00:00	March

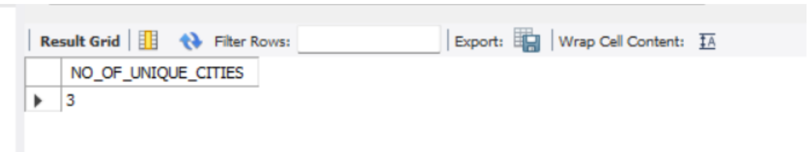
Generic Questions

1] How many unique cities does the data have?

CODE

```
SELECT COUNT(distinct(city)) AS NO_OF_UNIQUE_CITIES FROM sales;
```

OUTPUT:



NO_OF_UNIQUE_CITIES
3

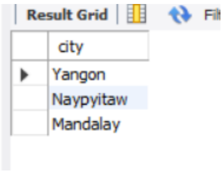
There are a total of 3 unique cities.

2] In which city is each branch?

CODE:

```
SELECT distinct city FROM sales;
```

OUTPUT:



	city
▶	Yangon
	Naypyitaw
	Mandalay

There are 3 cities in which the branches are located - Yangon, Naypyitaw, Mandalay.

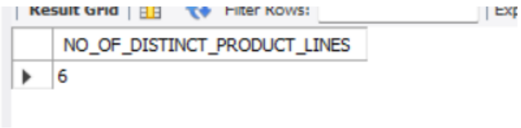
Product Analysis

1] How many unique product lines does the data have?

CODE:

SELECT COUNT(distinct product_line) AS NO_OF_DISTINCT_PRODUCT_LINES FROM sales;

OUTPUT:



	NO_OF_DISTINCT_PRODUCT_LINES
▶	6

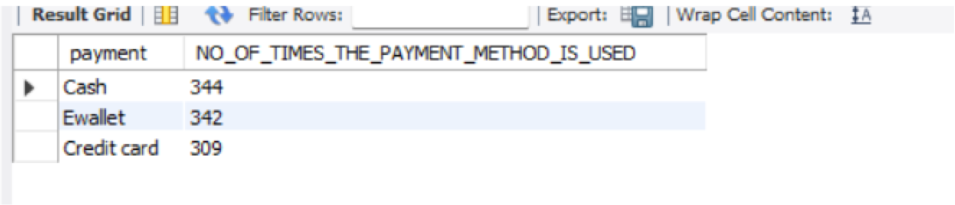
There are a total of 6 unique product lines in the data

2] What is the most common payment method?

CODE:

SELECT payment, COUNT(payment) AS NO_OF_TIMES_THE_PAYMENT_METHOD_IS_USED FROM sales GROUP BY payment ORDER BY COUNT(payment) DESC;

OUTPUT:



	payment	NO_OF_TIMES_THE_PAYMENT_METHOD_IS_USED
▶	Cash	344
	Ewallet	342
	Credit card	309

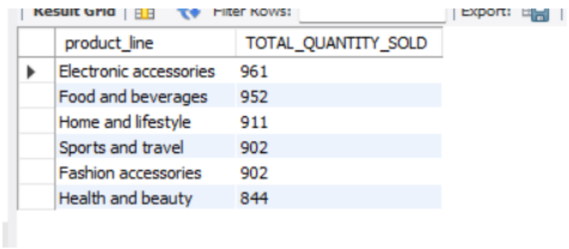
Cash mode of payment is the most common method.

3] What is the most selling product line?

CODE:

SELECT product_line, SUM(quantity) AS TOTAL_QUANTITY_SOLD FROM sales GROUP BY product_line ORDER BY SUM(quantity) DESC;

OUTPUT:



product_line	TOTAL_QUANTITY_SOLD
Electronic accessories	961
Food and beverages	952
Home and lifestyle	911
Sports and travel	902
Fashion accessories	902
Health and beauty	844

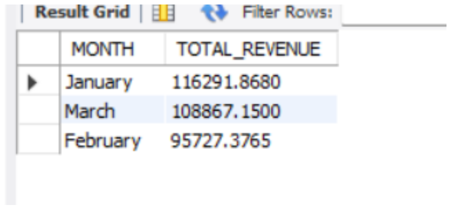
Electronic accessories is the most sold product line.

4] What is the total revenue by month?

CODE:

SELECT month_name AS MONTH, SUM(total) AS TOTAL_REVENUE FROM sales GROUP BY month_name ORDER BY SUM(total) DESC;

OUTPUT:



MONTH	TOTAL_REVENUE
January	116291.8680
March	108867.1500
February	95727.3765

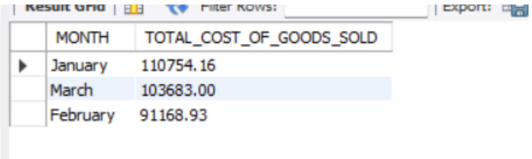
January boasts the highest total revenue

5] What month had the largest COGS?

CODE:

SELECT month_name AS MONTH, SUM(cogs) AS TOTAL_COST_OF_GOODS_SOLD FROM sales GROUP BY month_name ORDER BY SUM(cogs) DESC;

OUTPUT:



MONTH	TOTAL_COST_OF_GOODS_SOLD
January	110754.16
March	103683.00
February	91168.93

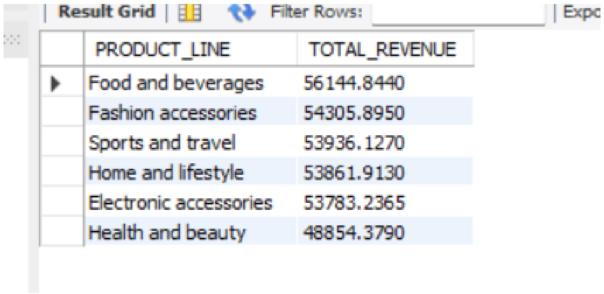
The month of January had the largest COGS.

6] What product line had the largest revenue?

CODE:

```
SELECT product_line AS PRODUCT_LINE, SUM(total) AS TOTAL_REVENUE FROM sales
GROUP BY product_line ORDER BY SUM(total) DESC;
```

OUTPUT:



PRODUCT_LINE	TOTAL_REVENUE
Food and beverages	56144.8440
Fashion accessories	54305.8950
Sports and travel	53936.1270
Home and lifestyle	53861.9130
Electronic accessories	53783.2365
Health and beauty	48854.3790

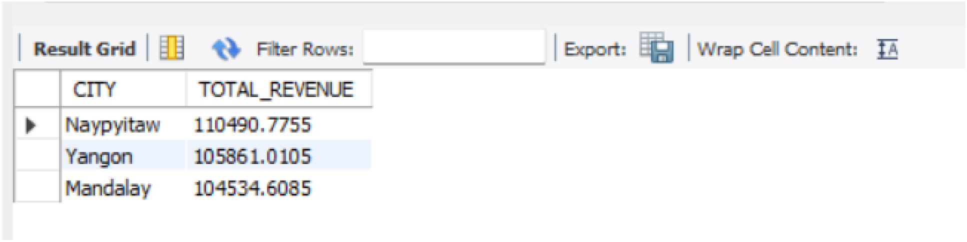
Food and Beverages have the largest revenue.

7] What is the city with the largest revenue?

CODE:

```
SELECT city AS CITY, SUM(total) AS TOTAL_REVENUE FROM sales GROUP BY city ORDER
BY SUM(total) DESC;
```

OUTPUT:



CITY	TOTAL_REVENUE
Naypyitaw	110490.7755
Yangon	105861.0105
Mandalay	104534.6085

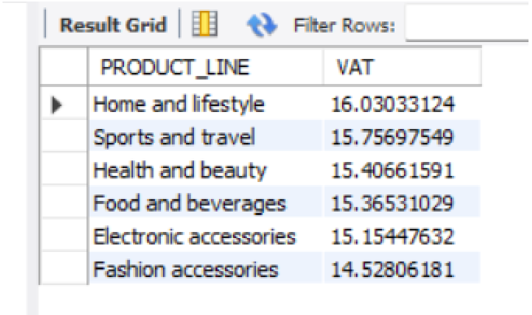
The city of Naypyitaw has the largest revenue.

8] What product line had the largest VAT?

CODE:

```
SELECT product_line AS PRODUCT_LINE, AVG(VAT) AS VAT FROM sales GROUP BY
product_line ORDER BY AVG(VAT) DESC;
```

OUTPUT:



PRODUCT_LINE	VAT
Home and lifestyle	16.03033124
Sports and travel	15.75697549
Health and beauty	15.40661591
Food and beverages	15.36531029
Electronic accessories	15.15447632
Fashion accessories	14.52806181

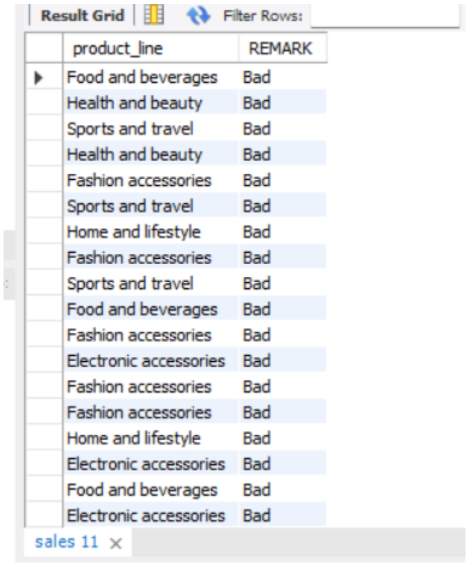
The product line Home and lifestyle has the largest VAT.

9] Fetch each product line and add a column to those product line showing "Good", "Bad". Good if its greater than average sales

CODE:

```
SELECT product_line AS PRODUCT_LINE, (CASE
            WHEN AVG(quantity)>6 THEN "Good"
            ELSE "Bad"
            END) AS REMARK FROM sales GROUP BY product_line;

ALTER TABLE sales ADD COLUMN REMARK VARCHAR(10);

UPDATE sales s
JOIN (
   SELECT product_line, AVG(quantity) OVER () as avg_quantity
   FROM sales
) sub ON s.product_line = sub.product_line
SET s.REMARK = CASE
   WHEN sub.avg_quantity > 6 THEN 'Good'
   ELSE 'Bad'
END;

SELECT product_line, REMARK from sales;
```

OUTPUT:



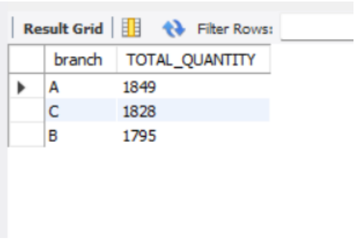
product_line	REMARK
Food and beverages	Bad
Health and beauty	Bad
Sports and travel	Bad
Health and beauty	Bad
Fashion accessories	Bad
Sports and travel	Bad
Home and lifestyle	Bad
Fashion accessories	Bad
Sports and travel	Bad
Food and beverages	Bad
Fashion accessories	Bad
Electronic accessories	Bad
Fashion accessories	Bad
Fashion accessories	Bad
Home and lifestyle	Bad
Electronic accessories	Bad
Food and beverages	Bad
Electronic accessories	Bad

10] Which branch sold more products than average product sold?

CODE:

```
SELECT branch, SUM(quantity) AS TOTAL_QUANTITY FROM sales GROUP BY branch
HAVING SUM(quantity) > (SELECT AVG(quantity) from sales) ORDER BY SUM(quantity)
DESC;
```

OUTPUT:



branch	TOTAL_QUANTITY
A	1849
C	1828
B	1795

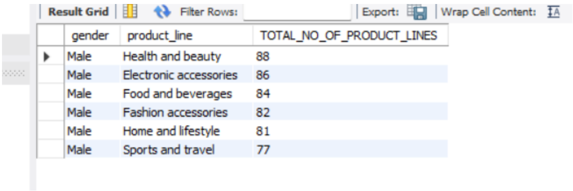
The branch A sold more products than the average products sold.

11] What is the most common product line by gender?

CODE:

```
SELECT gender, product_line, COUNT(product_line) AS TOTAL_NO_OF_PRODUCT_LINES
FROM sales WHERE gender='Male' GROUP BY gender, product_line
ORDER BY COUNT(product_line) DESC;
```

OUTPUT:



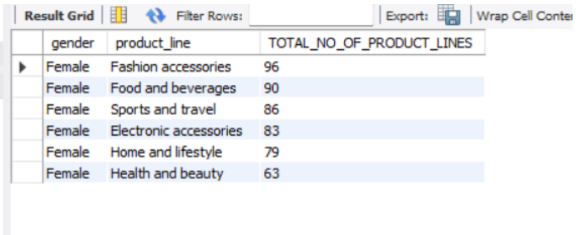
gender	product_line	TOTAL_NO_OF_PRODUCT_LINES
Male	Health and beauty	88
Male	Electronic accessories	86
Male	Food and beverages	84
Male	Fashion accessories	82
Male	Home and lifestyle	81
Male	Sports and travel	77

For the Male gender, the most common product line is Health and Beauty.

CODE:

```
SELECT gender, product_line, COUNT(product_line) AS TOTAL_NO_OF_PRODUCT_LINES
FROM sales WHERE gender='Female' GROUP BY gender, product_line
ORDER BY COUNT(product_line) DESC;
```

OUTPUT:



gender	product_line	TOTAL_NO_OF_PRODUCT_LINES
Female	Fashion accessories	96
Female	Food and beverages	90
Female	Sports and travel	86
Female	Electronic accessories	83
Female	Home and lifestyle	79
Female	Health and beauty	63

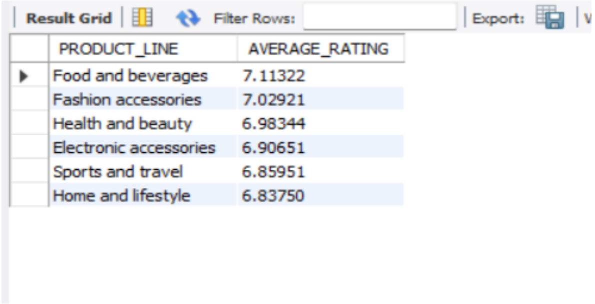
For the Female gender, the most common product line is Fashion accessories.

12] What is the average rating of each product line?

CODE:

```
SELECT product_line AS PRODUCT_LINE, AVG(rating) AS AVERAGE_RATING from sales
GROUP BY product_line ORDER BY AVG(rating) DESC;
```

OUTPUT:



PRODUCT_LINE	AVERAGE_RATING
Food and beverages	7.11322
Fashion accessories	7.02921
Health and beauty	6.98344
Electronic accessories	6.90651
Sports and travel	6.85951
Home and lifestyle	6.83750

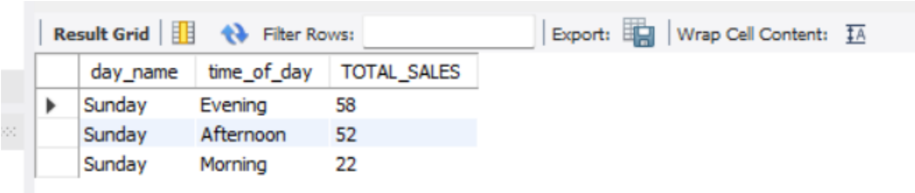
Sales Analysis

1] Number of sales made in each time of the day per weekday

CODE:

```
SELECT day_name, time_of_day, COUNT(*) AS TOTAL_SALES FROM sales WHERE
day_name='Sunday' GROUP BY day_name, time_of_day ORDER BY COUNT(*) DESC;
```

OUTPUT:



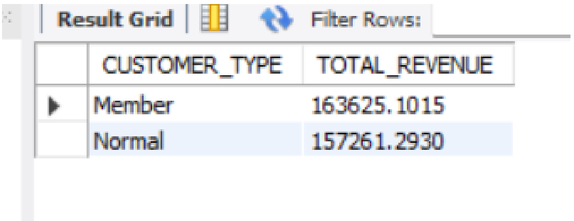
day_name	time_of_day	TOTAL_SALES
Sunday	Evening	58
Sunday	Afternoon	52
Sunday	Morning	22

2] Which of the customer types brings the most revenue?

CODE:

```
SELECT customer_type AS CUSTOMER_TYPE, SUM(total) AS TOTAL_REVENUE FROM
sales GROUP BY customer_type ORDER BY SUM(total) DESC;
```

OUTPUT:



CUSTOMER_TYPE	TOTAL_REVENUE
Member	163625.1015
Normal	157261.2930

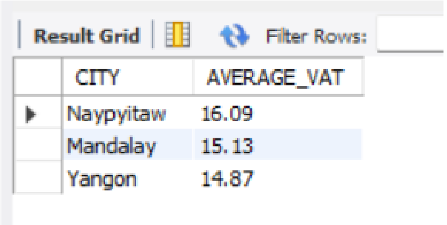
The member customer type brings the most revenue

3] Which city has the largest tax percent/ VAT (Value Added Tax)?

CODE:

```
SELECT city AS CITY, ROUND(AVG(VAT),2) AS AVERAGE_VAT from sales GROUP BY city
ORDER BY AVG(VAT) DESC;
```

OUTPUT:



CITY	AVERAGE_VAT
Naypyitaw	16.09
Mandalay	15.13
Yangon	14.87

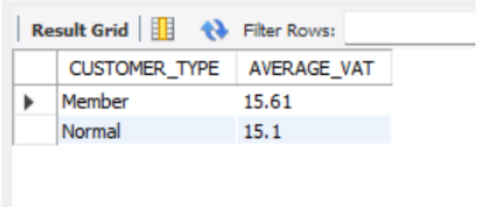
The city of Naypyitaw has the largest VAT.

4] Which customer type pays the most in VAT?

CODE:

```
SELECT customer_type AS CUSTOMER_TYPE, ROUND(AVG(VAT),2) AS AVERAGE_VAT
from sales GROUP BY customer_type ORDER BY AVG(VAT) DESC;
```

OUTPUT:



CUSTOMER_TYPE	AVERAGE_VAT
Member	15.61
Normal	15.1

The customer type - Member pays the most in VAT.

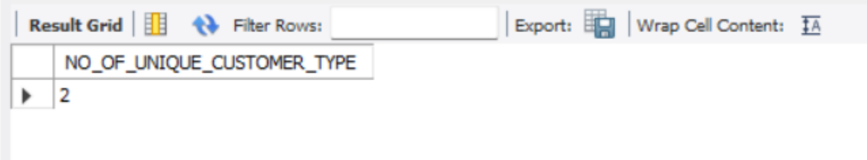
Customer Analysis

1] How many unique customer types does the data have?

CODE:

```
SELECT COUNT(distinct customer_type) AS NO_OF_UNIQUE_CUSTOMER_TYPE from sales;
```

OUTPUT:



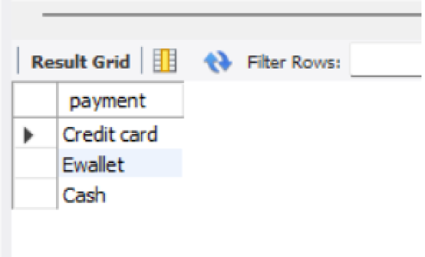
NO_OF_UNIQUE_CUSTOMER_TYPE
2

2] How many unique payment methods does the data have?

CODE:

```
SELECT distinct payment from sales;
```

OUTPUT:



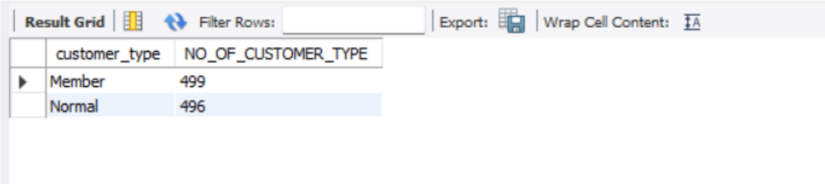
payment
Credit card
Ewallet
Cash

3] What is the most common customer type?

CODE:

```
SELECT customer_type, COUNT(customer_type) AS NO_OF_CUSTOMER_TYPE FROM sales GROUP BY customer_type ORDER BY COUNT(customer_type) DESC;
```

OUTPUT:



customer_type	NO_OF_CUSTOMER_TYPE
Member	499
Normal	496

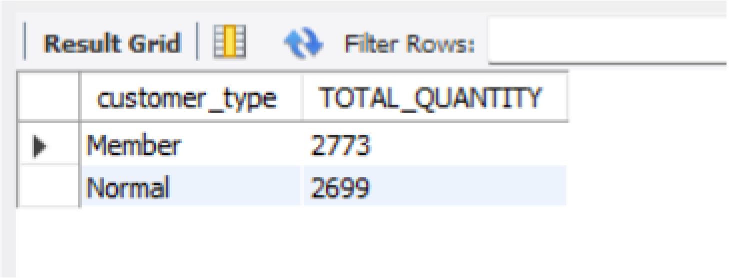
In total, there are 499 members and 496 normal customers.

4] Which customer type buys the most?

CODE:

SELECT customer_type, SUM(quantity) AS TOTAL_QUANTITY FROM sales GROUP BY customer_type ORDER BY SUM(quantity) DESC;

OUTPUT:



customer_type	TOTAL_QUANTITY
Member	2773
Normal	2699

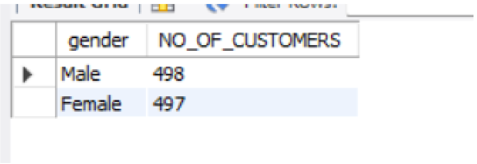
The total quantity of goods purchased by members is 2773, whereas normal customers bought 2699.

5] What is the gender of most of the customers?

CODE:

SELECT gender, COUNT(*) AS NO_OF_CUSTOMERS from sales GROUP BY gender ORDER BY COUNT(*) DESC;

OUTPUT:



gender	NO_OF_CUSTOMERS
Male	498
Female	497

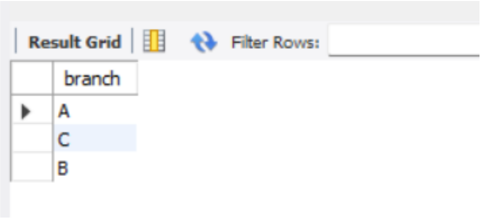
The total number of male customers is 498, while the total number of female customers is 497.

6] What is the gender distribution per branch?

CODE:

SELECT distinct branch from sales;

OUTPUT:

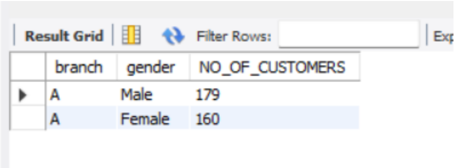


branch
A
C
B

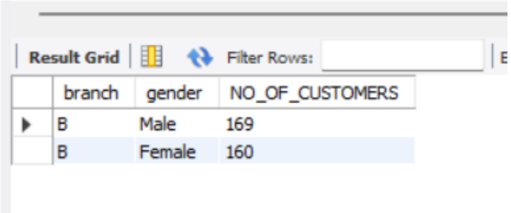
CODE:

SELECT branch, gender, COUNT(*) AS NO_OF_CUSTOMERS FROM sales WHERE branch='A' GROUP BY branch, gender ORDER BY COUNT(*) DESC;
SELECT branch, gender, COUNT(*) AS NO_OF_CUSTOMERS FROM sales WHERE branch='B' GROUP BY branch, gender ORDER BY COUNT(*) DESC;
SELECT branch, gender, COUNT(*) AS NO_OF_CUSTOMERS FROM sales WHERE branch='C' GROUP BY branch, gender ORDER BY COUNT(*) DESC;

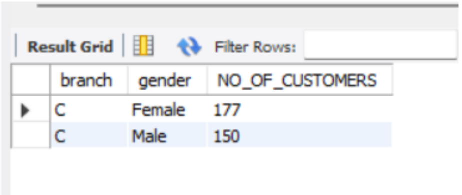
OUTPUT:



branch	gender	NO_OF_CUSTOMERS
A	Male	179
A	Female	160



branch	gender	NO_OF_CUSTOMERS
B	Male	169
B	Female	160



branch	gender	NO_OF_CUSTOMERS
C	Female	177
C	Male	150

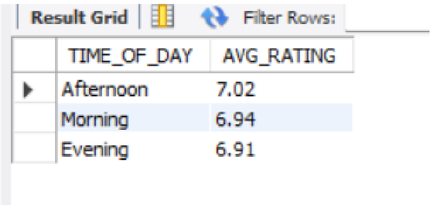
In Branch A, there are 179 male customers compared to 160 female customers. In Branch B, there are 169 male customers and 160 female customers. In Branch C, there are 177 female customers versus 150 male customers.

7] Which time of the day do customers give most ratings?

CODE:

SELECT time_of_day AS TIME_OF_DAY, ROUND(AVG(rating),2) AS AVG_RATING FROM sales GROUP BY time_of_day ORDER BY AVG(rating) DESC;

OUTPUT:



TIME_OF_DAY	AVG_RATING
Afternoon	7.02
Morning	6.94
Evening	6.91

Overall, the afternoon has the highest average rating.

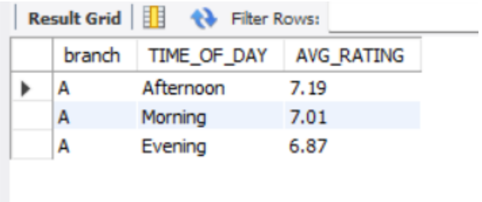
8] Which time of the day do customers give the most ratings per branch?

CODE:

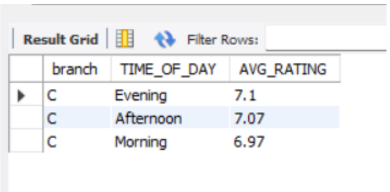
SELECT branch, time_of_day AS TIME_OF_DAY, ROUND(AVG(rating),2) AS AVG_RATING FROM sales WHERE branch='A' GROUP BY branch, time_of_day ORDER BY AVG(rating) DESC;

```
SELECT branch, time_of_day AS TIME_OF_DAY, ROUND(AVG(rating),2) AS AVG_RATING
FROM sales WHERE branch='B' GROUP BY branch, time_of_day ORDER BY AVG(rating)
DESC;

SELECT branch, time_of_day AS TIME_OF_DAY, ROUND(AVG(rating),2) AS AVG_RATING
FROM sales WHERE branch='C' GROUP BY branch, time_of_day ORDER BY AVG(rating)
DESC;
```

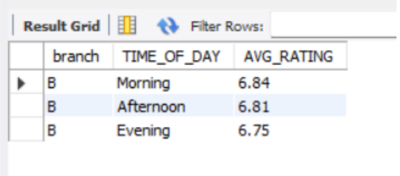
OUTPUT:



branch	TIME_OF_DAY	AVG_RATING
A	Afternoon	7.19
A	Morning	7.01
A	Evening	6.87



branch	TIME_OF_DAY	AVG_RATING
C	Evening	7.1
C	Afternoon	7.07
C	Morning	6.97



branch	TIME_OF_DAY	AVG_RATING
B	Morning	6.84
B	Afternoon	6.81
B	Evening	6.75

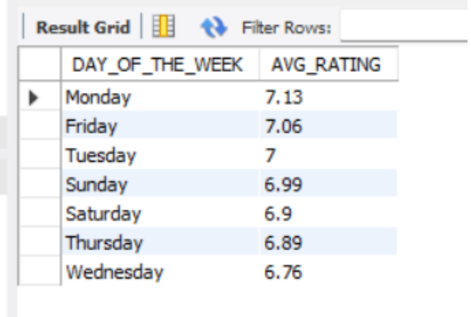
For Branch A, the afternoon holds the highest average rating. For Branch B, the evening has the highest average rating. For Branch C, the morning achieves the highest average rating.

9] Which day of the week has the best avg ratings?

CODE:

```
SELECT day_name AS DAY_OF_THE_WEEK, ROUND(AVG(rating),2) AS AVG_RATING
FROM sales GROUP BY day_name ORDER BY AVG(rating) DESC;
```

OUTPUT:



DAY_OF_THE_WEEK	AVG_RATING
Monday	7.13
Friday	7.06
Tuesday	7
Sunday	6.99
Saturday	6.9
Thursday	6.89
Wednesday	6.76

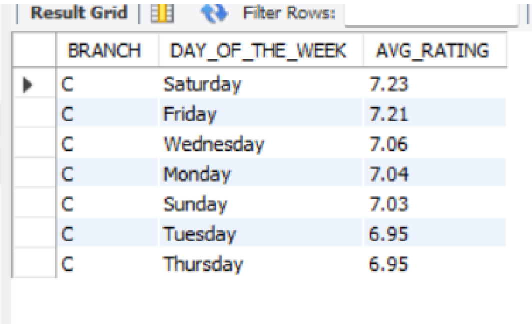
Monday boasts the highest average rating of the week.

10] Which day of the week has the best average ratings per branch?

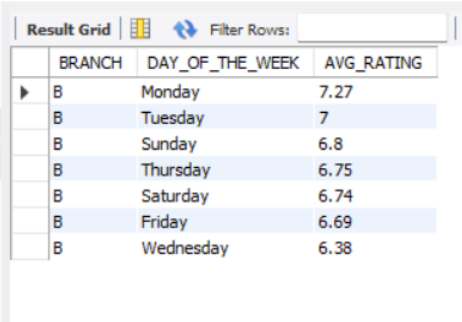
CODE:

```
SELECT branch AS BRANCH, day_name AS DAY_OF_THE_WEEK, ROUND(AVG(rating),2)
AS AVG_RATING FROM sales WHERE branch='A' GROUP BY branch, day_name ORDER BY
AVG(rating) DESC;

SELECT branch AS BRANCH, day_name AS DAY_OF_THE_WEEK, ROUND(AVG(rating),2)
AS AVG_RATING FROM sales WHERE branch='B' GROUP BY branch, day_name ORDER BY
AVG(rating) DESC;

SELECT branch AS BRANCH, day_name AS DAY_OF_THE_WEEK, ROUND(AVG(rating),2)
AS AVG_RATING FROM sales WHERE branch='C' GROUP BY branch, day_name ORDER BY
AVG(rating) DESC;
```

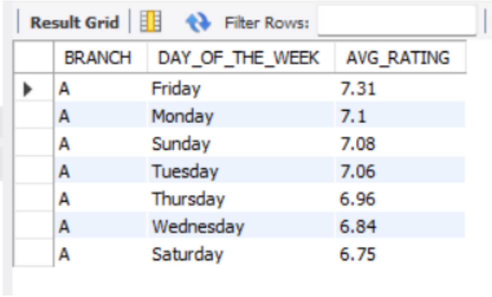
OUTPUT:



BRANCH	DAY_OF_THE_WEEK	AVG_RATING
C	Saturday	7.23
C	Friday	7.21
C	Wednesday	7.06
C	Monday	7.04
C	Sunday	7.03
C	Tuesday	6.95
C	Thursday	6.95



BRANCH	DAY_OF_THE_WEEK	AVG_RATING
B	Monday	7.27
B	Tuesday	7
B	Sunday	6.8
B	Thursday	6.75
B	Saturday	6.74
B	Friday	6.69
B	Wednesday	6.38



BRANCH	DAY_OF_THE_WEEK	AVG_RATING
A	Friday	7.31
A	Monday	7.1
A	Sunday	7.08
A	Tuesday	7.06
A	Thursday	6.96
A	Wednesday	6.84
A	Saturday	6.75

For Branch A, the average ratings are highest on Friday. For Branch B, Monday has the best average ratings. For Branch C, Saturday is the top day for average ratings.

Insights and Recommendations:

1] Most Common Payment Method: The predominant mode of payment is cash.

Recommendation: Consider offering more payment options to cater to diverse customer preferences and enhance convenience.

2] Top Selling Product Line: Electronic accessories emerge as the best-selling product category.

Recommendation: Capitalize on the popularity of electronic accessories by expanding the product range or launching targeted marketing campaigns.

3] Highest Revenue Month: January records the highest total revenue.

Recommendation: Analyze the factors contributing to January's success and implement strategies to replicate or sustain this revenue peak in other months.

4] Largest COGS Month: January had the largest Cost of Goods Sold (COGS).

Recommendation: Evaluate cost-efficiency measures and supply chain management practices to optimize COGS throughout the year.

5] Top Revenue Category: Food and beverages generate the highest revenue.

Recommendation: Explore opportunities for product diversification or promotion within the food and beverage segment to further boost revenue.

6] Top Revenue City: Naypyitaw stands out with the highest revenue.

Recommendation: Consider investing more resources in Naypyitaw, such as targeted marketing campaigns or special promotions, to maintain or enhance its revenue contribution.

7] Product Line with Highest VAT: Home and lifestyle products have the largest average Value-Added Tax (VAT).

Recommendation: Review pricing strategies and VAT implications to ensure competitiveness and profitability within the home and lifestyle category.

8] Branch Performance: Branch A surpasses the average product sales volume.

Recommendation: Identify the factors contributing to Branch A's success and implement best practices across other branches to improve overall sales performance.

9] Gender-based Product Preferences: Health and beauty products are most popular among males, while females prefer fashion accessories.

Recommendation: Tailor marketing efforts and product offerings to align with gender-specific preferences and enhance customer satisfaction.

10] Top Revenue Customer Type: Members contribute the highest revenue.

Recommendation: Implement loyalty programs or exclusive offers to incentivize membership and encourage repeat purchases.

11] City with Highest VAT: Naypyitaw registers the highest VAT.

Recommendation: Monitor VAT trends and adjust pricing or promotional strategies accordingly to optimize revenue and VAT collection.

12] Customer Type with Highest VAT: Members pay the most in VAT.

Recommendation: Communicate the value proposition of membership, such as exclusive discounts or rewards, to justify the higher VAT expenditure for members.

13] Customer Demographics: There are 499 members and 496 normal customers in total.

Recommendation: Develop targeted marketing strategies to engage and retain both member and non-member customer segments effectively.

14] Total Quantity Purchased: Members purchase 2773 goods, while normal customers buy 2699.

Recommendation: Analyze purchase patterns to tailor inventory management and product assortment strategies to meet customer demand more efficiently.

15] Gender Distribution: The total number of male customers is 498, and female customers total 497.

Recommendation: Ensure gender-inclusive marketing and product offerings to cater to diverse customer demographics effectively.

16] Gender Distribution by Branch: Branch A has 179 male customers and 160 female customers, Branch B has 169 male customers and 160 female customers, and Branch C has 177 female customers and 150 male customers.

Recommendation: Investigate branch-specific factors influencing gender distribution and tailor marketing strategies accordingly to achieve better gender balance.

17] Peak Rating Time: Afternoons receive the highest average ratings overall.

Recommendation: Schedule customer service or promotional activities during peak rating times to maximize customer satisfaction and engagement.

18] Branch-specific Rating Peaks: Afternoons are best rated in Branch A, evenings in Branch B, and mornings in Branch C.

Recommendation: Align staffing and service offerings with peak rating times at each branch to enhance customer experience and satisfaction.

19] Weekly Rating Peak: Mondays receive the highest average ratings of the week.

Recommendation: Plan promotions or special events on Mondays to capitalize on positive customer sentiment and drive sales.

20] Branch-specific Weekly Rating Peaks: Fridays are best rated in Branch A, Mondays in Branch B, and Saturdays in Branch C.

Recommendation: Customize marketing strategies and promotions to coincide with each branch's peak rating day, optimizing customer engagement and satisfaction.