**THE SALES DATA OF AMAZON’S E-COMMERCE SITE**

**1. What is the count of distinct cities in the dataset?**

**select** count(distinct(city)) as count\_of\_cities **from** amazondb.amazon;

Insights: There are **3** Cities in this Dataset

**2. For each branch, what is the corresponding city?**

select distinct branch,city from amazondb.amazon

Insights: **A – Yangon, B – Naypyitaw, C - Mandalay**

**3. What is the count of distinct product lines in the dataset?**

select count(distinct `Product line`) as DistinctProductLines from amazondb.amazon;

Insights: There are **6** Product lines in this dataset

**4. Which payment method occurs most frequently?**

SELECT Payment, COUNT(payment) AS MethodCount FROM amazondb.amazon

GROUP BY Payment

LIMIT 1;

Insights: **Ewallet** have occurred most no of times

**5. Which product line has the highest sales?**

Select `product line`, sum(`unit price` \* quantity) as total\_sales from amazondb.amazon

Group by `product line`

Order by total\_sales asc

Limit 1;

Insights: **Health and Beauty** with **46851.1799** Sales

**6. How much revenue is generated each month?**

select Monthname, sum(total) as Revenue from amazondb.amazon

group by Monthname

order by Revenue desc;

Insights: **Jan - 116291.86800000005, Mar - 109455.50700000004, Feb - 97219.37399999997**

**7.In which month did the cost of goods sold reach its peak?**

select Monthname, SUM(cogs) as peak\_goods\_sold from amazondb.amazon

group by Monthname

order by peak\_goods\_sold desc;

Insights: **Jan – 110754.1600, Mar – 104243.339,Feb – 92589.88**

**8.Which product line generated the highest revenue?**

Select `product line`, sum(total) as Revenue from amazondb.amazon group by `Product line` order by Revenue desc;

Insights: **Food and Beverages – 56144.84400**

**9.In which city was the highest revenue recorded?**

select city, sum(total) as Revenue from amazondb.amazon group by City order by Revenue desc;

Insights: **Naypyitaw - 110568.70649999994**

**10. Which product line incurred the highest Value Added Tax?**

select `product line`,sum(`tax 5%`) as Highest\_VAT from amazondb.amazon group by `Product line` order by Highest\_VAT desc;

Insight: **Food and beverages - 2673.56399999**

**11. For each product line, add a column indicating "Good" if its sales are above average, otherwise "Bad."**

select `product line`, sum(`unit price` \* Quantity) as total\_sales,

case

when sum(`unit price` \* Quantity) > (

select avg(total\_sales)

from (

select `product line`, sum(`unit price` \* Quantity) as total\_sales

from amazondb.amazon

group by `product line`

) as subquery

) then 'Good'

else 'bad'

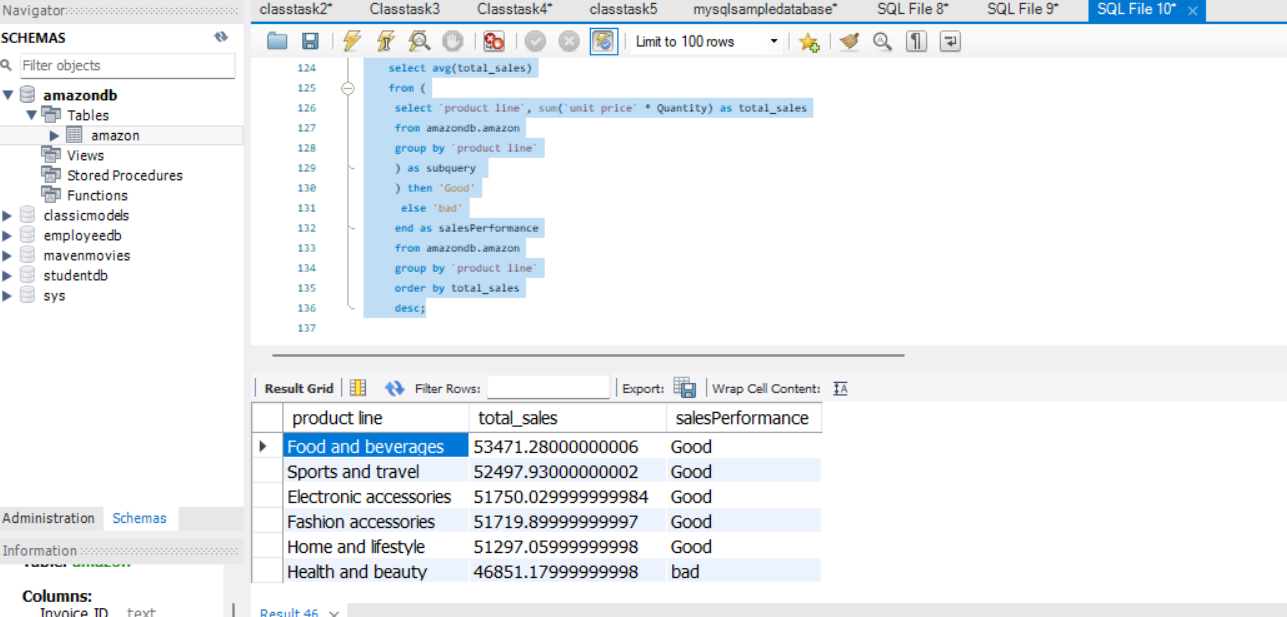
end as salesPerformance

from amazondb.amazon

group by `product line`

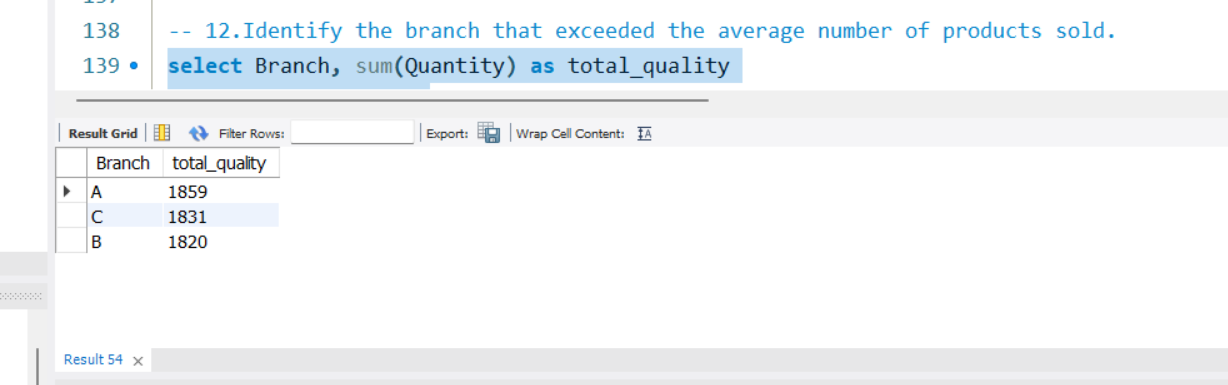
order by total\_sales

desc;

Insight: 

**12.Identify the branch that exceeded the average number of products sold.?**

select Branch, sum(Quantity) as total\_qualityfrom amazon\_data.amazon group by Branch having sum(Quantity) > (select avg(Quantity) from amazon\_data.amazon);



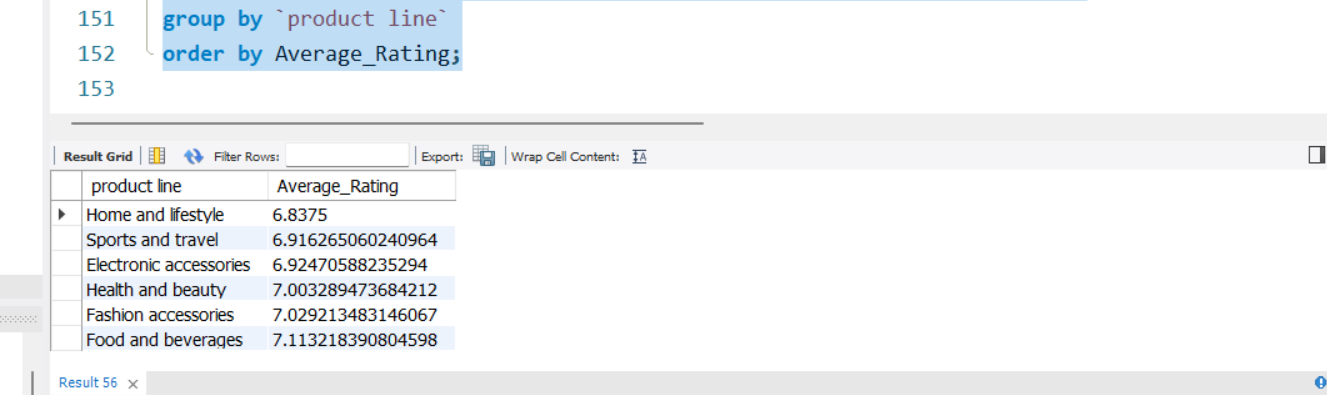
**13.Which product line is most frequently associated with each gender?**

select `product line`, Gender,count(`product line`) as no\_of\_Productline from amazondb.amazon group by gender,`product line` order by no\_of\_Productline;

Insights **Male – Sports and Travel , Female – Home and Lifestyle**

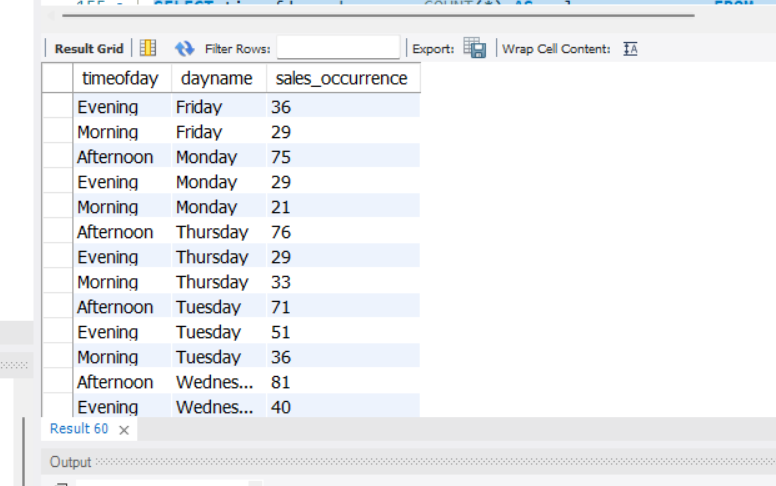
**14.Calculate the average rating for each product line?**

select `product line`, avg(rating) as Average\_Rating from amazondb.amazon group by `product line` order by Average\_Rating;



**15. Count the sales occurrences for each time of day on every weekday.?**

SELECT timeofday, dayname, COUNT(\*) AS sales\_occurrence FROM amazondb.amazon where dayname in ('Monday','Tuesday', 'Wednesday', 'Thursday', 'Friday') GROUP BY timeofday, dayname ORDER BY dayname, timeofday;



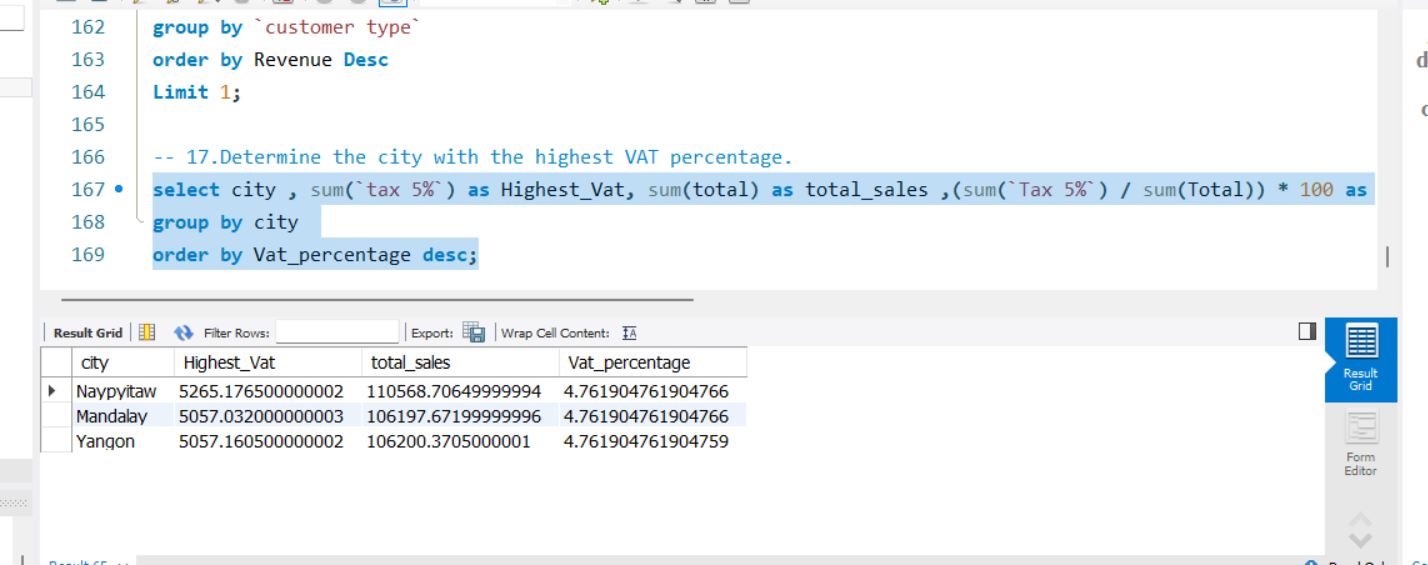
**16. Identify the customer type contributing the highest revenue.?**

select `Customer type` , sum(total) as Revenue from amazondb.amazon group by `customer type` order by Revenue Desc Limit 1;

Insight: **Member - 164223.444**

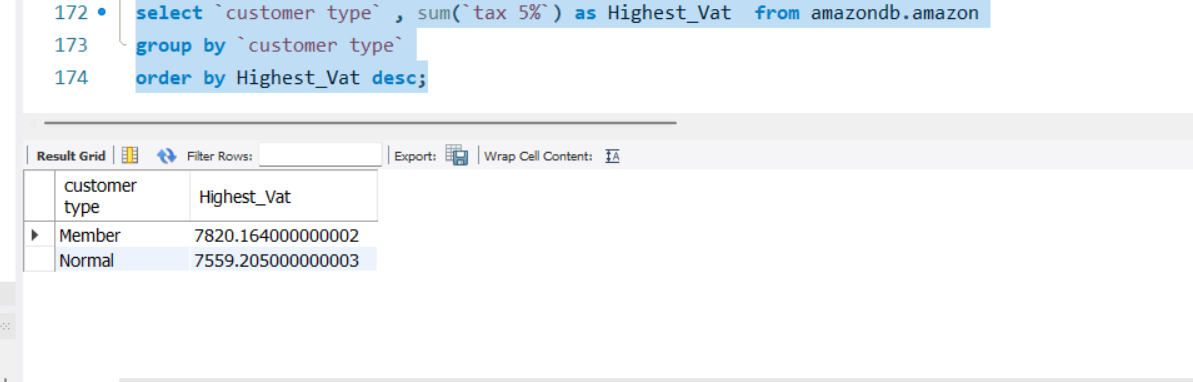
**17. Determine the city with the highest VAT percentage.?**

select city, sum(`tax 5%`) as Highest\_Vat, sum(total) as total sales ,(sum(`Tax 5%`) / sum(Total)) as Vat\_percentage from amazondb.amazon group by city order by Vat\_percentage desc;



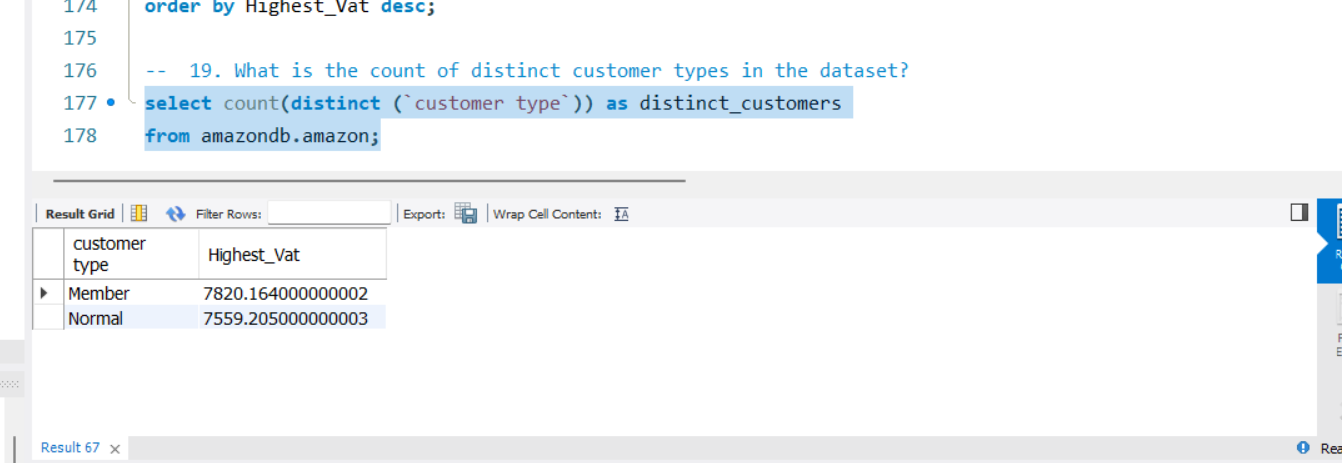
**18.Identify the customer type with the highest VAT payments?**

select `customer type` , sum(`tax 5%`) as Highest\_Vat from amazondb.amazon group by `customer type` order by Highest\_Vat desc;



**19. What is the count of distinct customer types in the dataset?**

select count(distinct (`customer type`)) as distinct\_customers from amazondb.amazon;



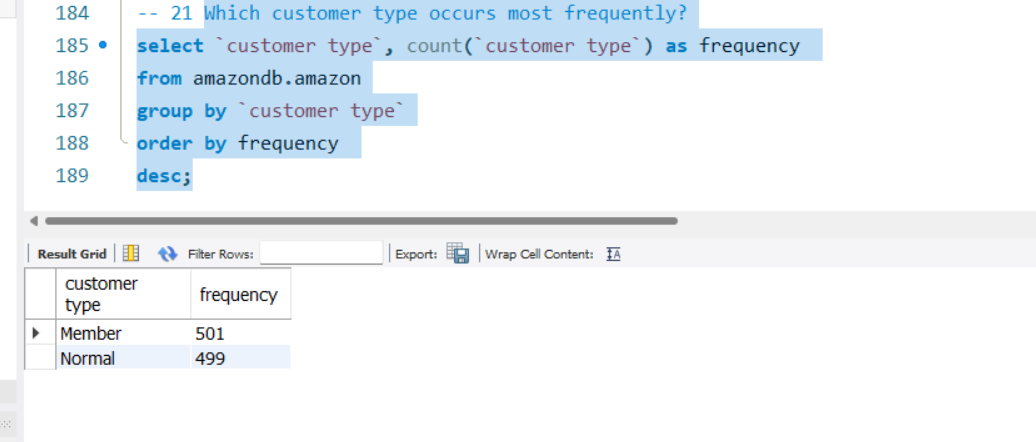
**20. What is the count of distinct payment methods in the dataset?**

select count(distinct(payment)) as distinct\_payment from amazondb.amazon;

Insights: There are **3** Payment Methods

**21. Which customer type occurs most frequently?**

select `customer type`, count(`customer type`) as frequency from amazondb.amazon group by `customer type` order by frequency desc;



**22.Identify the customer type with the highest purchase frequency.?**

select `customer type`, count(\*) as highest\_purchase from amazondb.amazon group by `customer type order by highest\_purchase desc limit 1;

Insights: **Member – 501**

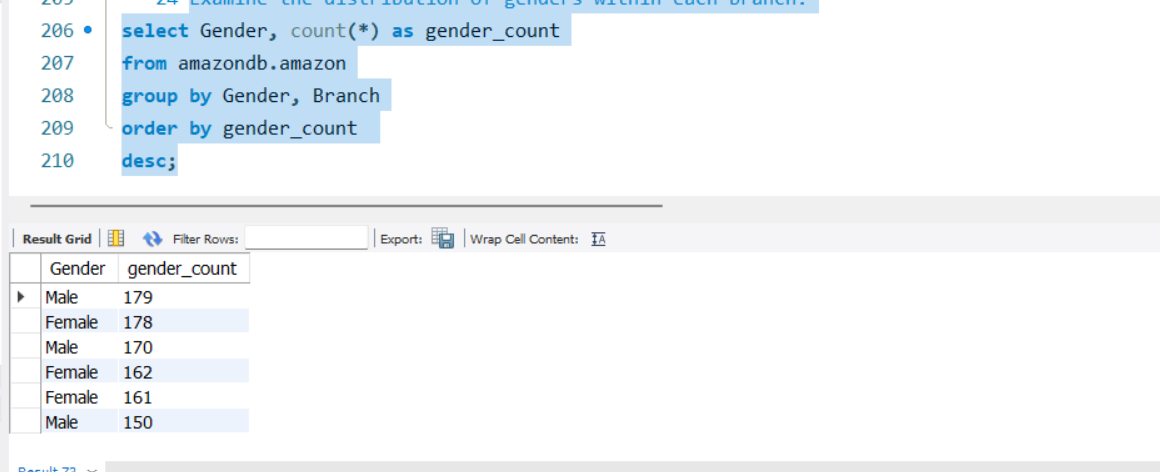
**23 Determine the predominant gender among customers.?**

select Gender, count(\*) as predominant\_genderfrom amazondb.amazongroup by Genderorder by predominant\_gender desc limit 1;

Insights: **Female – 501**

**24. Examine the distribution of genders within each branch.?**

select Gender, count(\*) as gender\_count from amazondb.amazon group by Gender, Branch order by gender\_count desc;



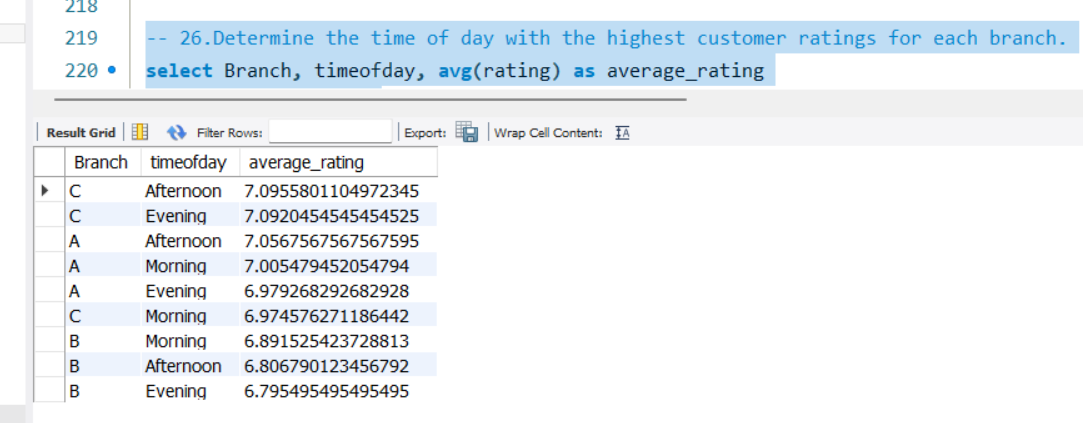
**25. Identify the time of day when customers provide the most ratings. ?**

select timeofday, count(Rating) as rating\_count from amazondb.amazon group by timeofday order by rating\_count desc limit 1;

Insight: **Afternoon – 528**

**26. Determine the time of day with the highest customer ratings for each branch. ?**

select Branch, timeofday, avg(rating) as average\_rating from amazondb.amazon group by Branch, timeofday order by average\_rating desc;



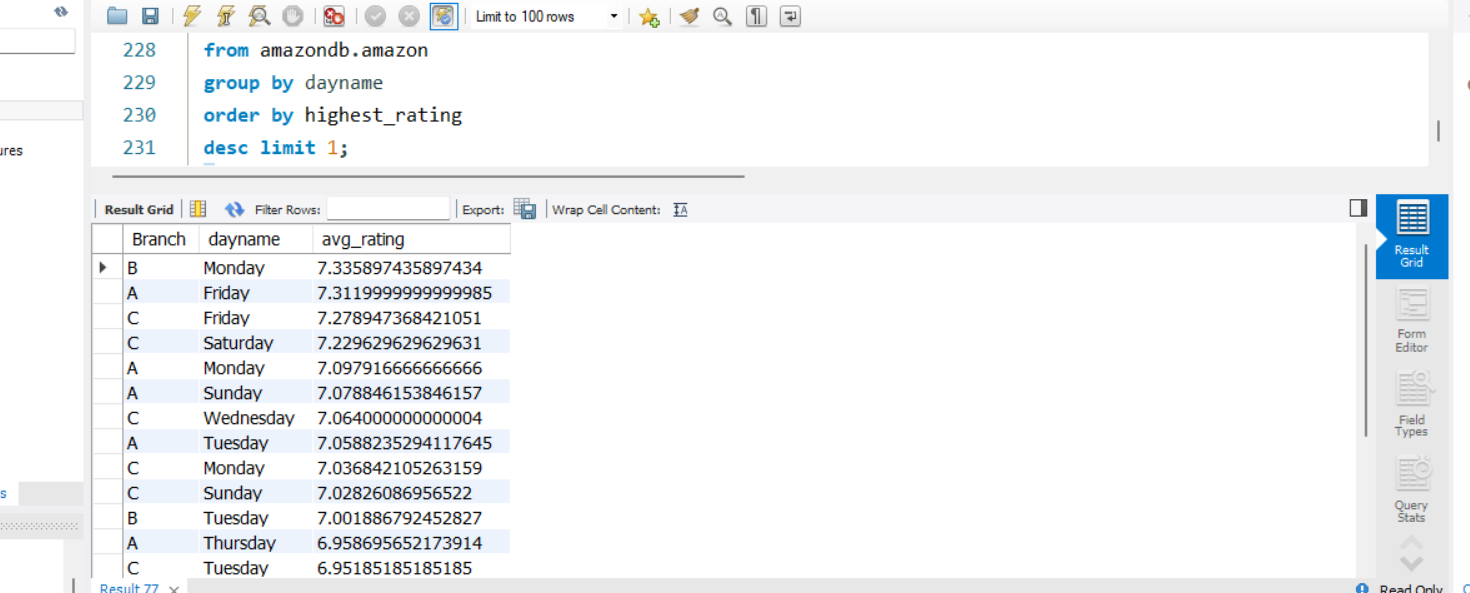
**27. Identify the day of the week with the highest average ratings.?**

select dayname, avg(Rating) as highest\_rating from amazondb.amazon group by dayname order by highest\_rating desc limit 1;

Insight: **Monday - 7.153599999999**

**28 Determine the day of the week with the highest average ratings for each branch.?**

select Branch, dayname, avg(Rating) as avg\_rating from amazondb.amazon group by Branch, dayname order by avg\_rating desc;



**INSIGHTS:**

**1.PRODUCT ANALYSIS:**

* **Food and Beverages** have generated highest revenue among all the product line.
* **Health and Beauty** have least sales in Product line
* City wise highest revenue **Mandalay – Sports and Travel, Naypyidaw – Food and Beverages, Yangon – Home and Lifestyle**
* City wise least revenue **Mandalay – Food and Beverages, Naypyidaw – Home and Lifestyle, Yangon – Health and Beauty**

**2.SALES ANALYSIS:**

* The Most of the sales occurred in **Afternoon** and **Evening**
* **Naypyidaw** have generated highest sales among three cities
* **January** has recorded highest sales
* Most of the customers used **E-wallet** has Payment method
* **Members** contribute **50.8%** of total sales, while **normal** customers account for **49.1%.**

**3.CUSTOMERS ANALYSIS:**

* **Members** customers were interested in **Food and Beverages** and **Normal** customers were interested in **Electronics accessories.**
* **Male** customers are interested in **Sports and Travel** and **Female** customers are interested in **Fashion accessories**
* **Female** have used **Cash** as payment method most of the time and **Male** have used **E-wallet** as payment method most of the time.