WALMART SALES DATA ANALYSIS

**# FEATURE ENGINEERING**

**Calculate the time of day from time column**

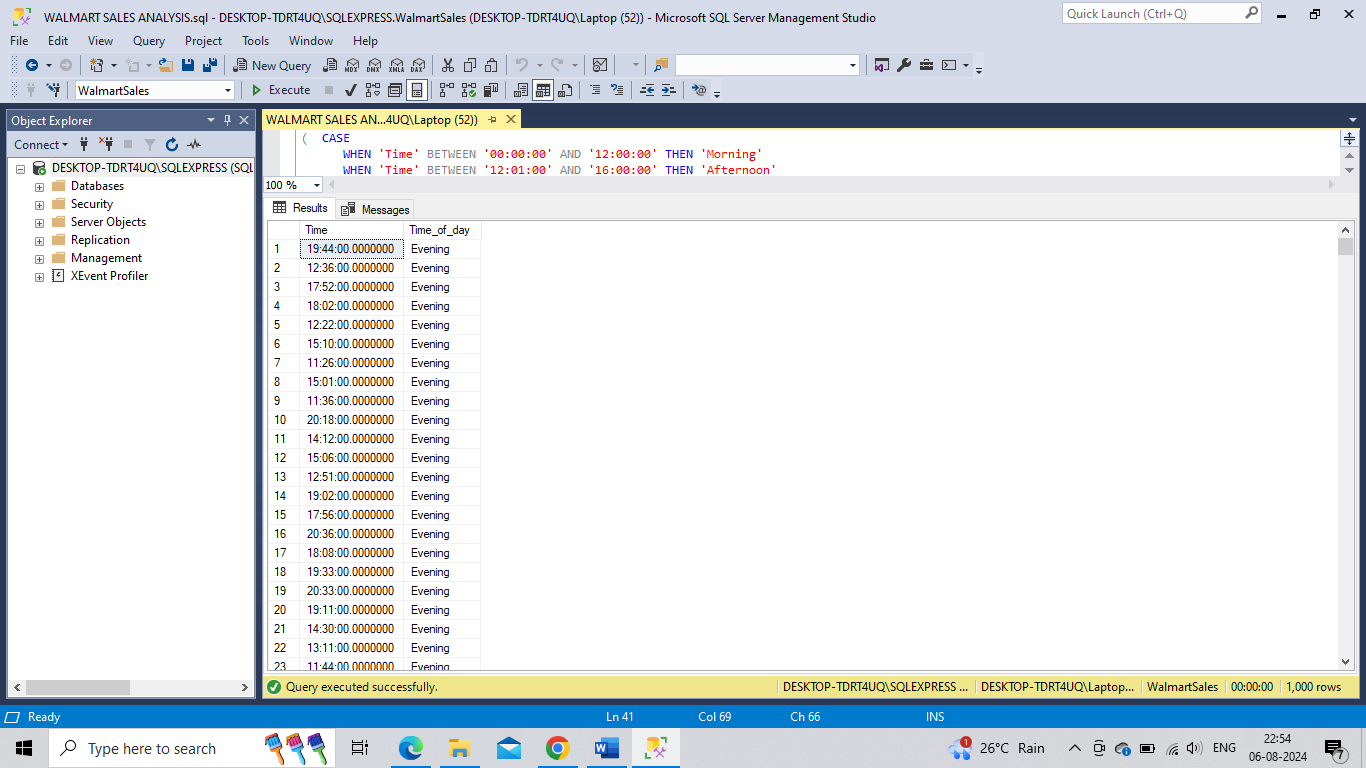
SELECT Time,( CASE

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 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 );

**ADD DAY NAME COLUMN**

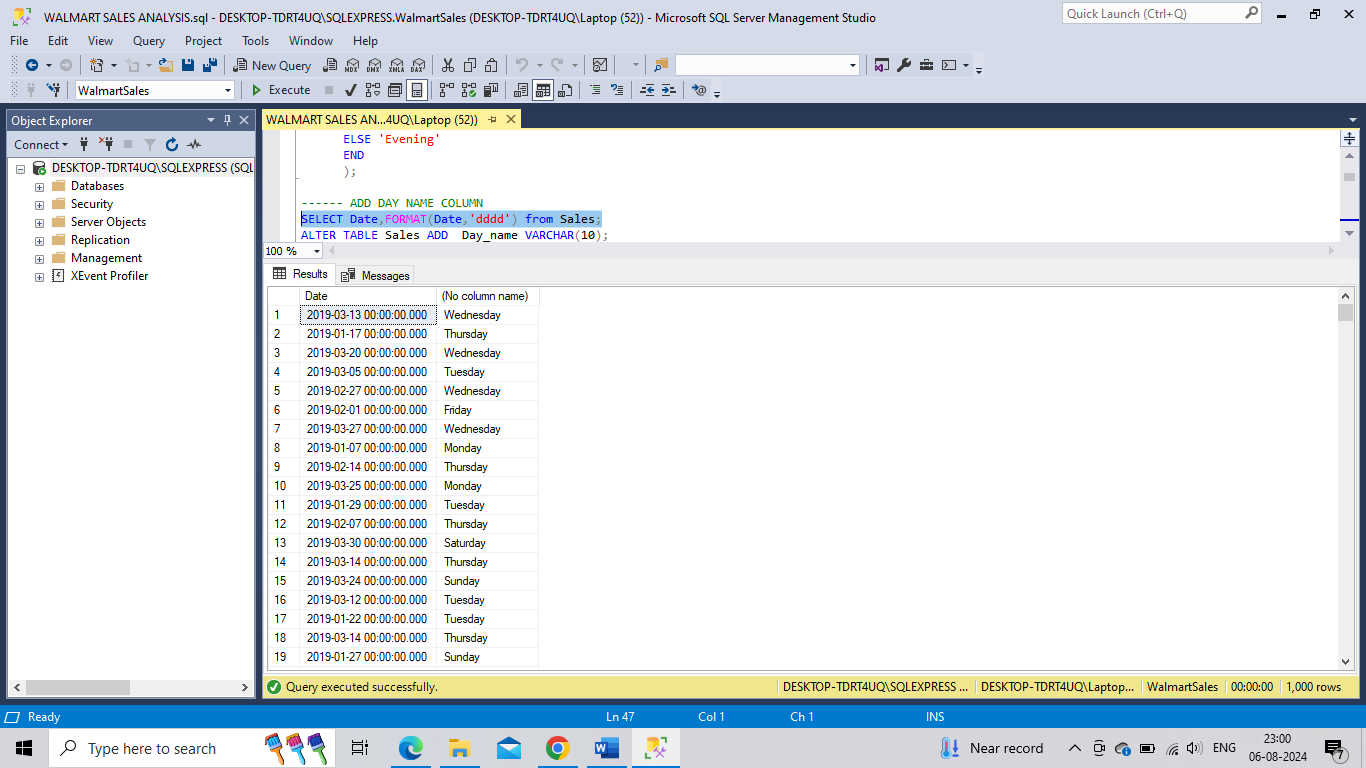
SELECT Date,FORMAT(Date,'dddd') from Sales;

ALTER TABLE Sales ADD Day\_name VARCHAR(10);

UPDATE Sales

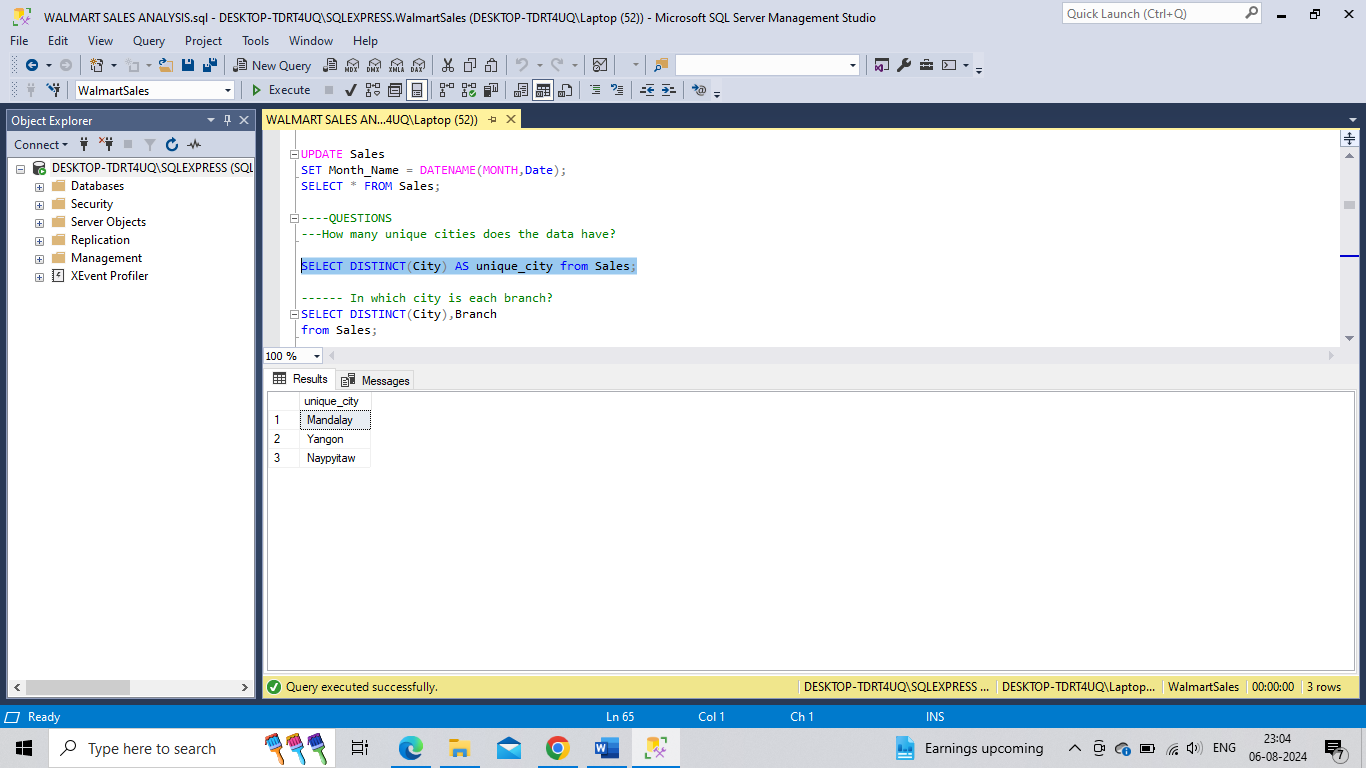
SET Day\_name = FORMAT(Date,'dddd');

SELECT \* FROM Sales;



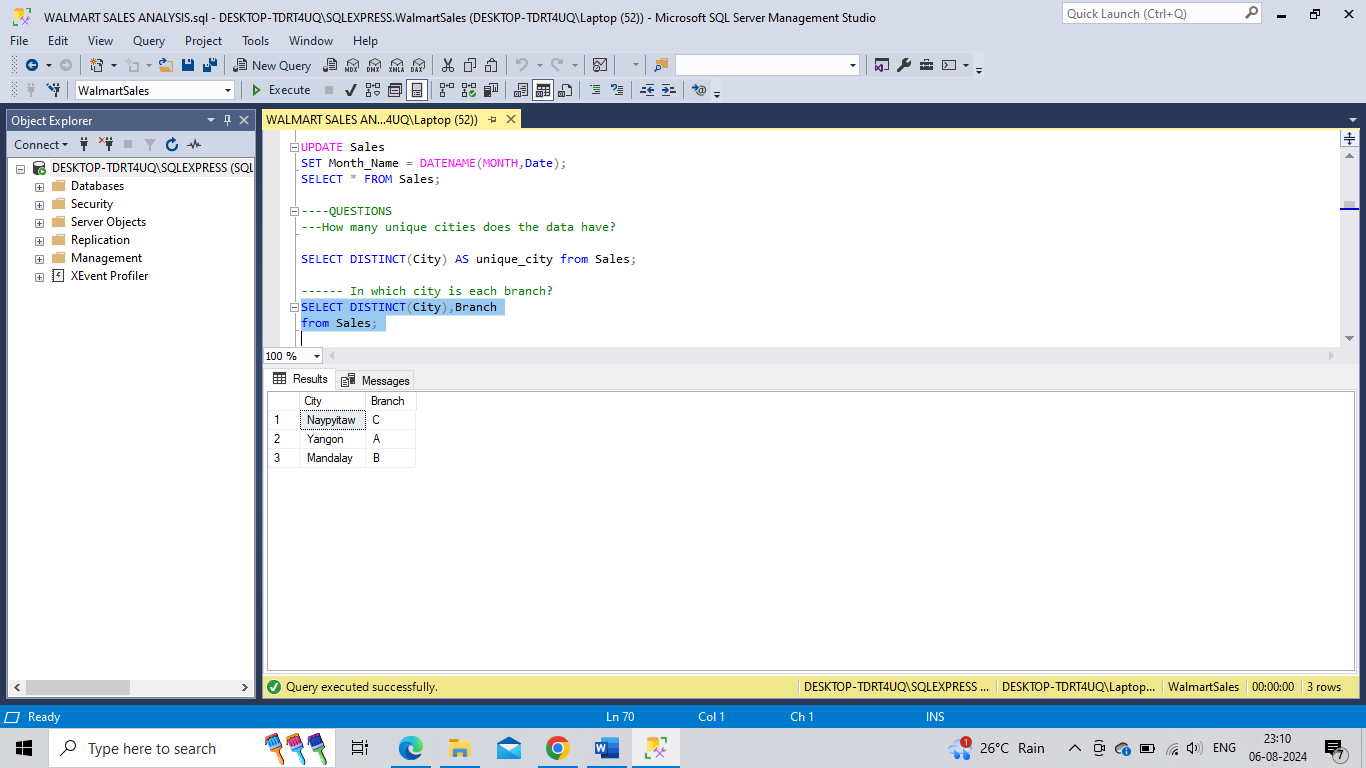
**Q.1. How many unique cities does the data have?**

SELECT DISTINCT(City) AS unique\_city from Sales;



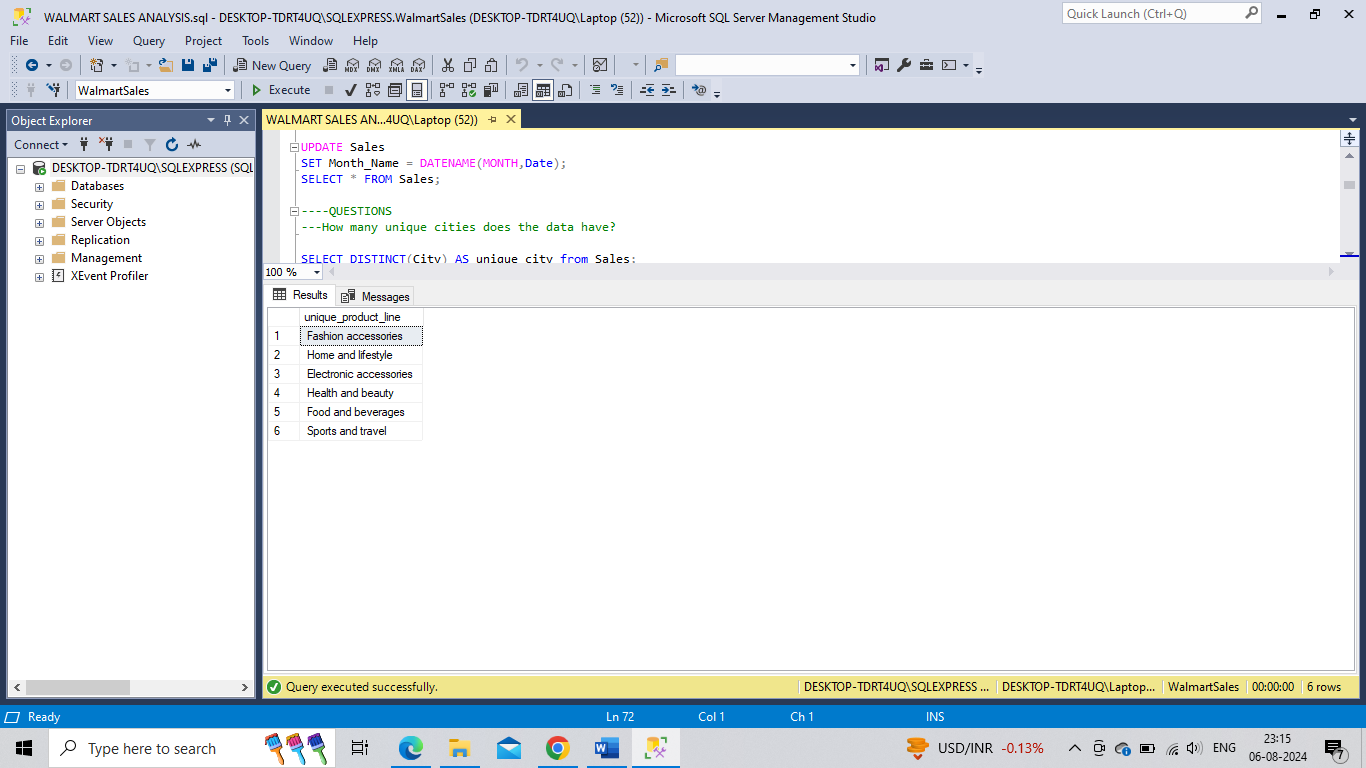
**Q.2. In which city is each branch?**

SELECT DISTINCT(City),Branch from Sales;



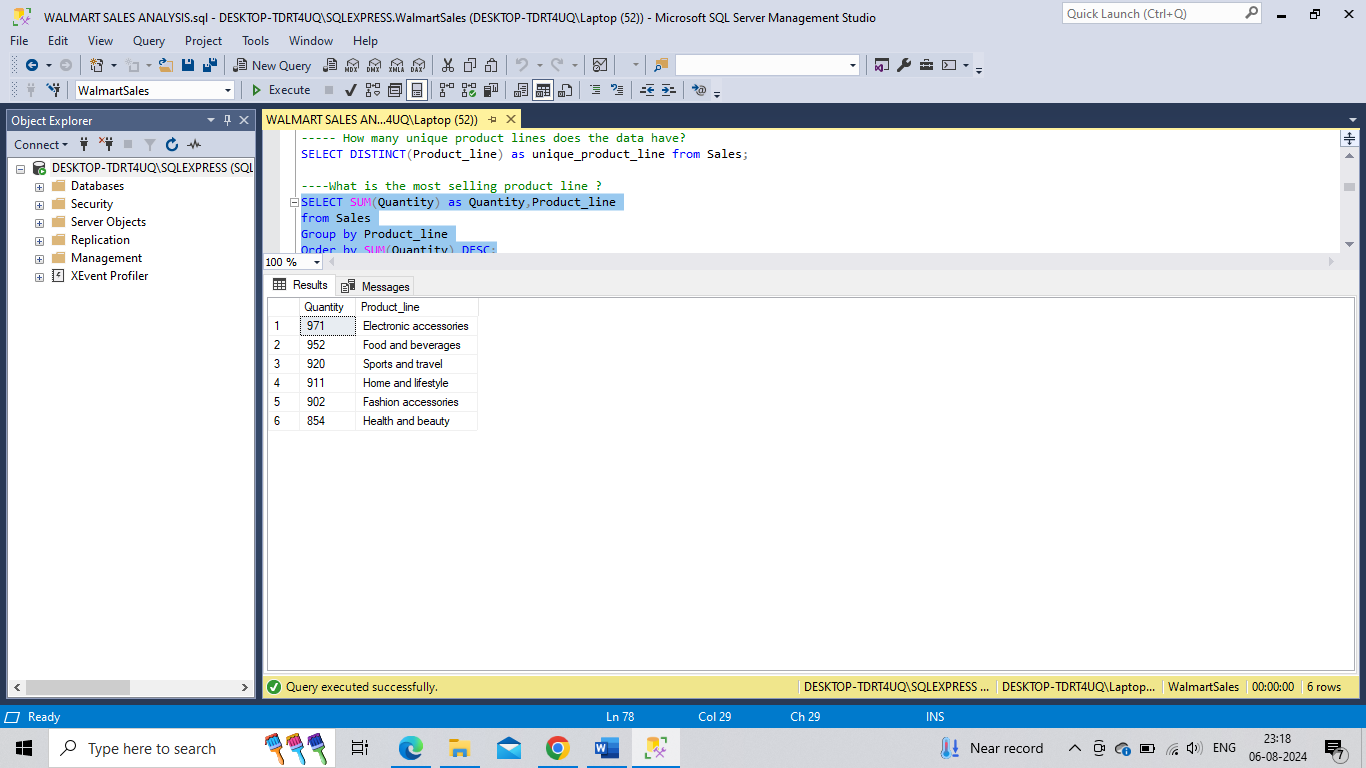
**Q.3. How many unique product lines does the data have?**

SELECT DISTINCT(Product\_line) as unique\_product\_line from Sales;



**Q.4. What is the most selling product line ?**

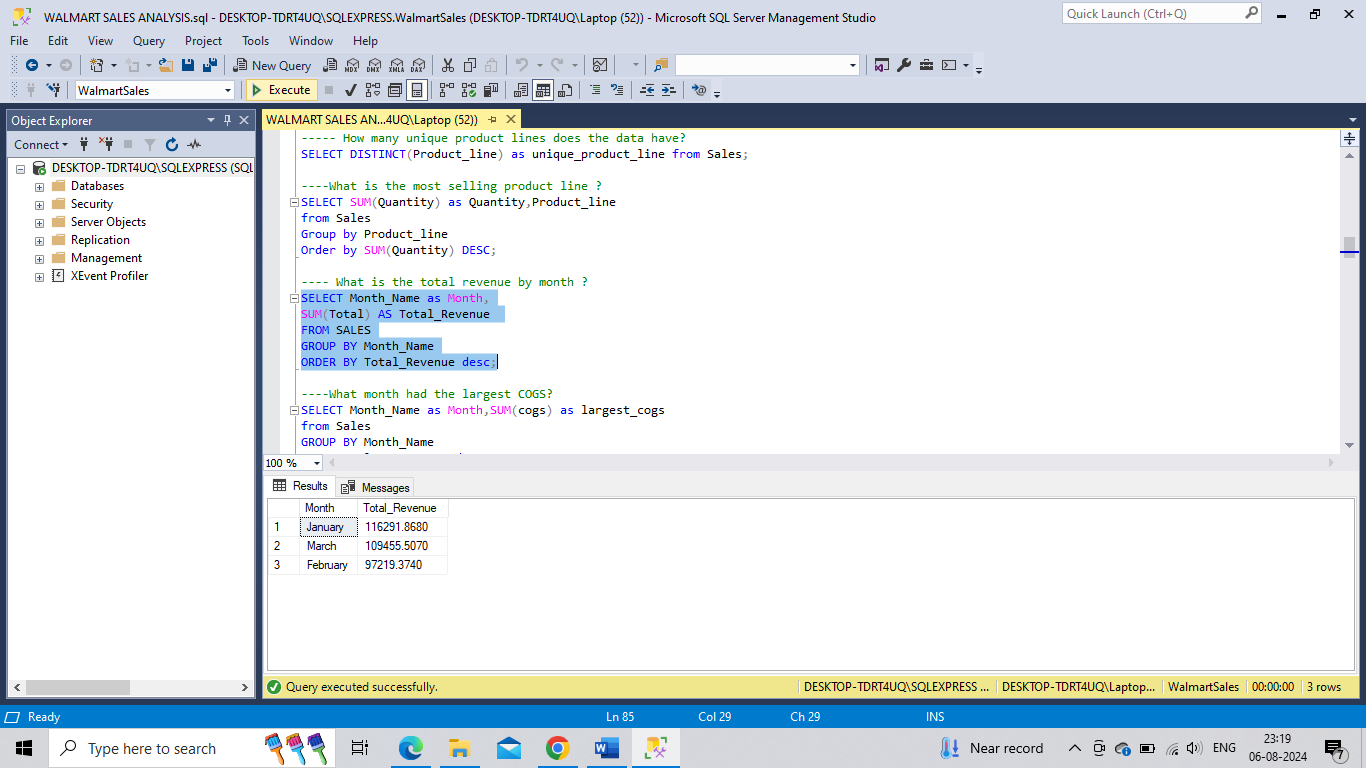
SELECT SUM(Quantity) as Quantity,Product\_line from Sales Group by Product\_line Order by SUM(Quantity) DESC;



**Q.5. What is the total revenue by month ?**

SELECT Month\_Name as Month, SUM(Total) AS Total\_Revenue FROM SALES

GROUP BY Month\_Name ORDER BY Total\_Revenue desc;



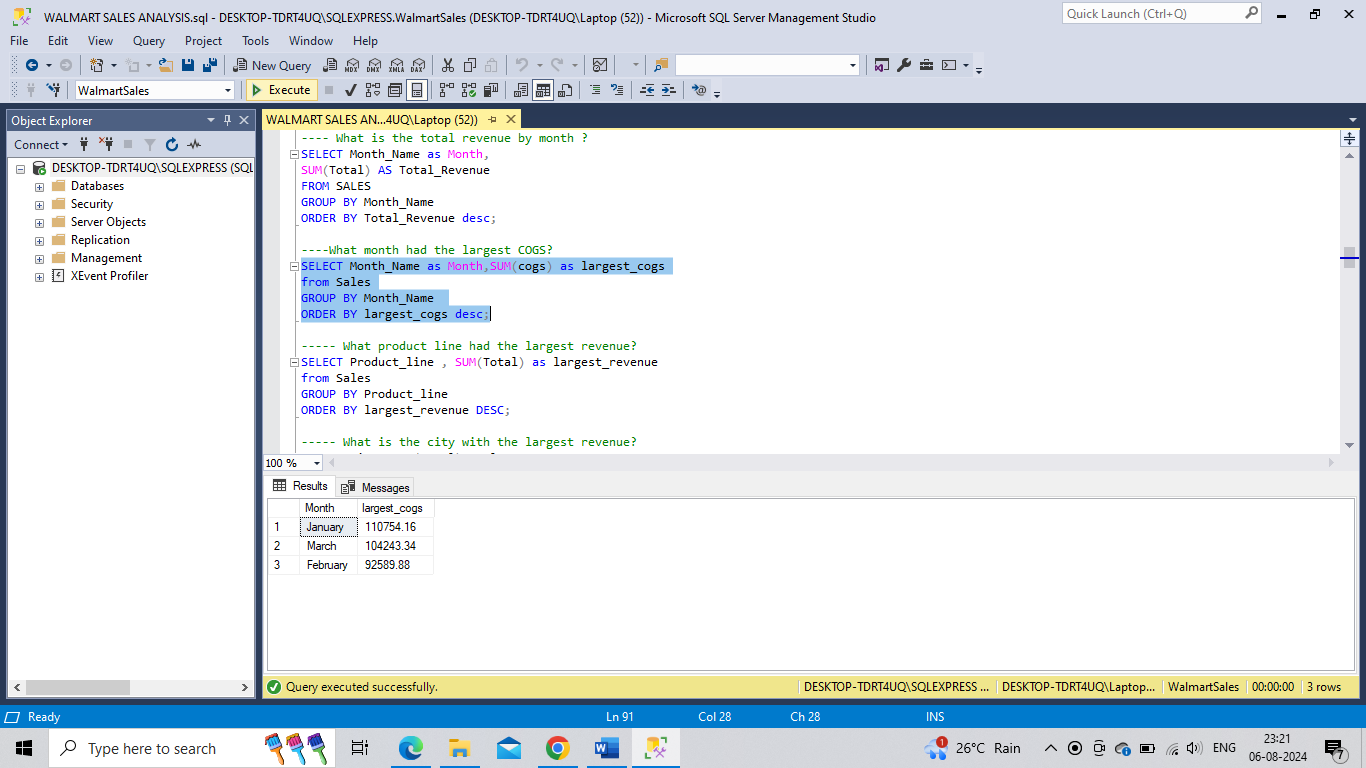
Q.6. **What month had the largest COGS?**

SELECT Month\_Name as Month,SUM(cogs) as largest\_cogs

from Sales

GROUP BY Month\_Name

ORDER BY largest\_cogs desc;



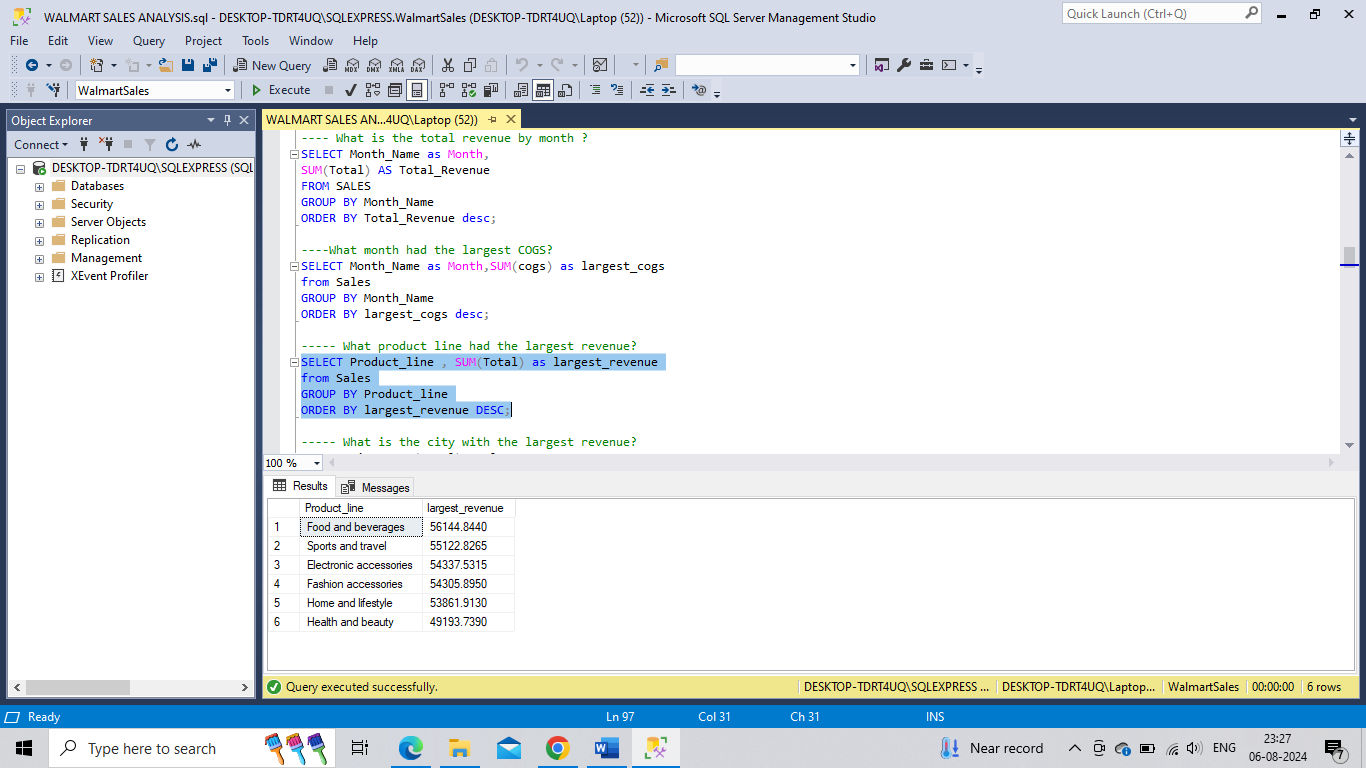
**Q.7. What product line had the largest revenue?**

SELECT Product\_line , SUM(Total) as largest\_revenue

from Sales

GROUP BY Product\_line

ORDER BY largest\_revenue DESC;



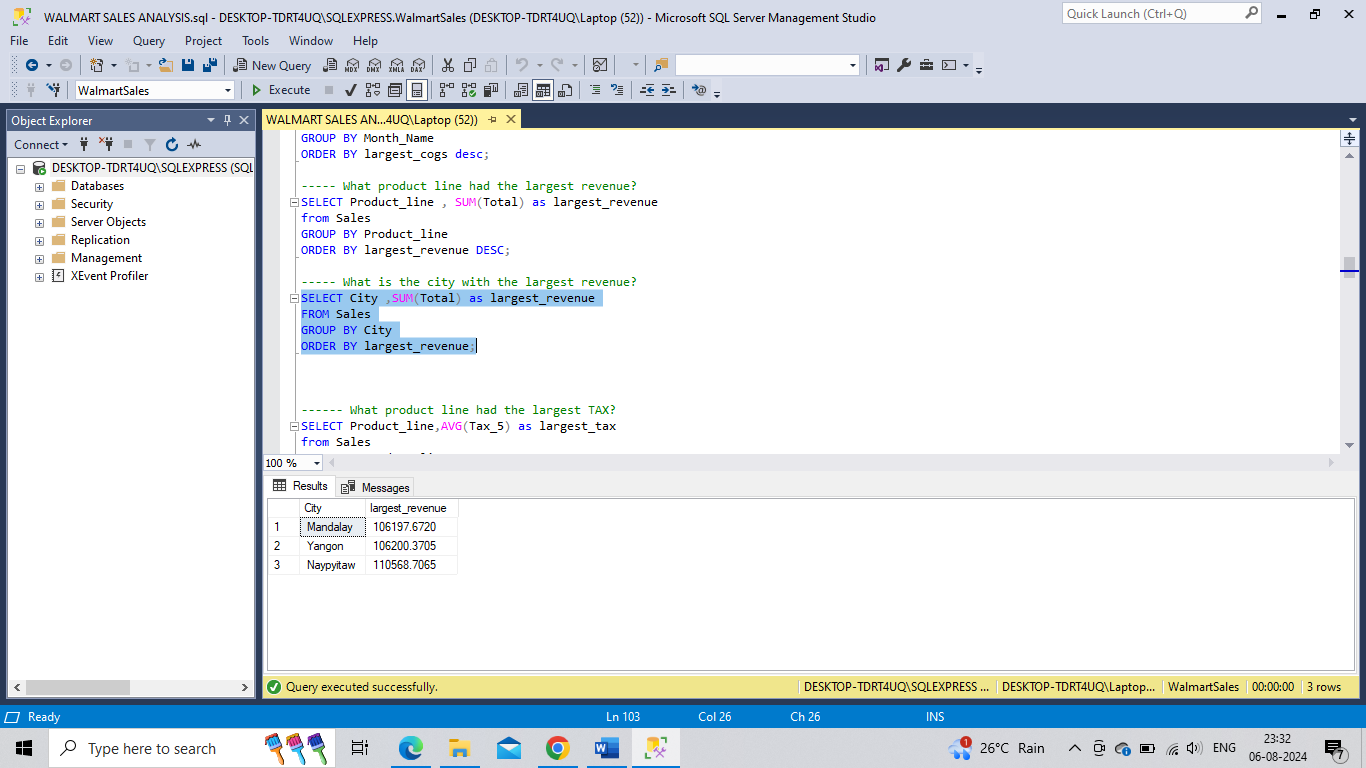
**Q.8. What is the city with the largest revenue?**

SELECT City ,SUM(Total) as largest\_revenue

FROM Sales

GROUP BY City

ORDER BY largest\_revenue;



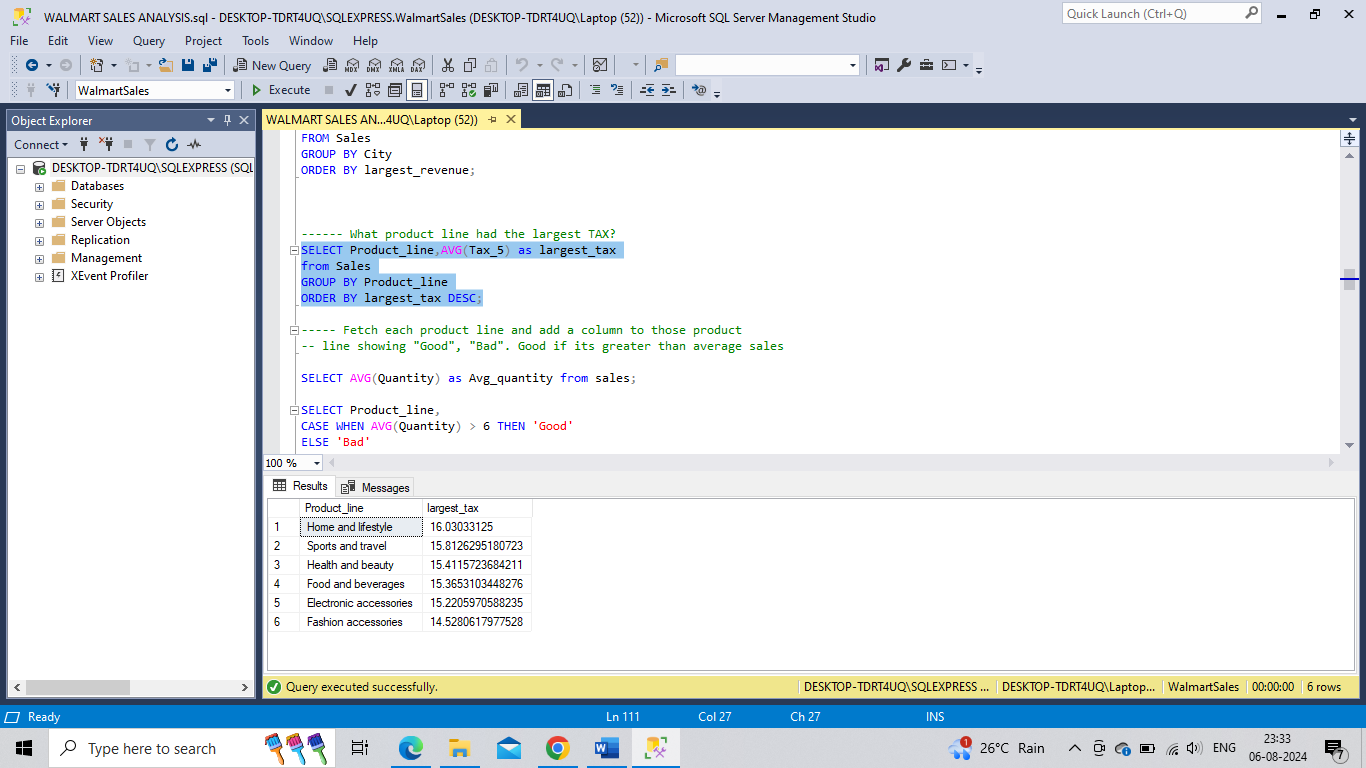
**Q.9. What product line had the largest TAX?**

SELECT Product\_line,AVG(Tax\_5) as largest\_tax

from Sales

GROUP BY Product\_line

ORDER BY largest\_tax DESC;

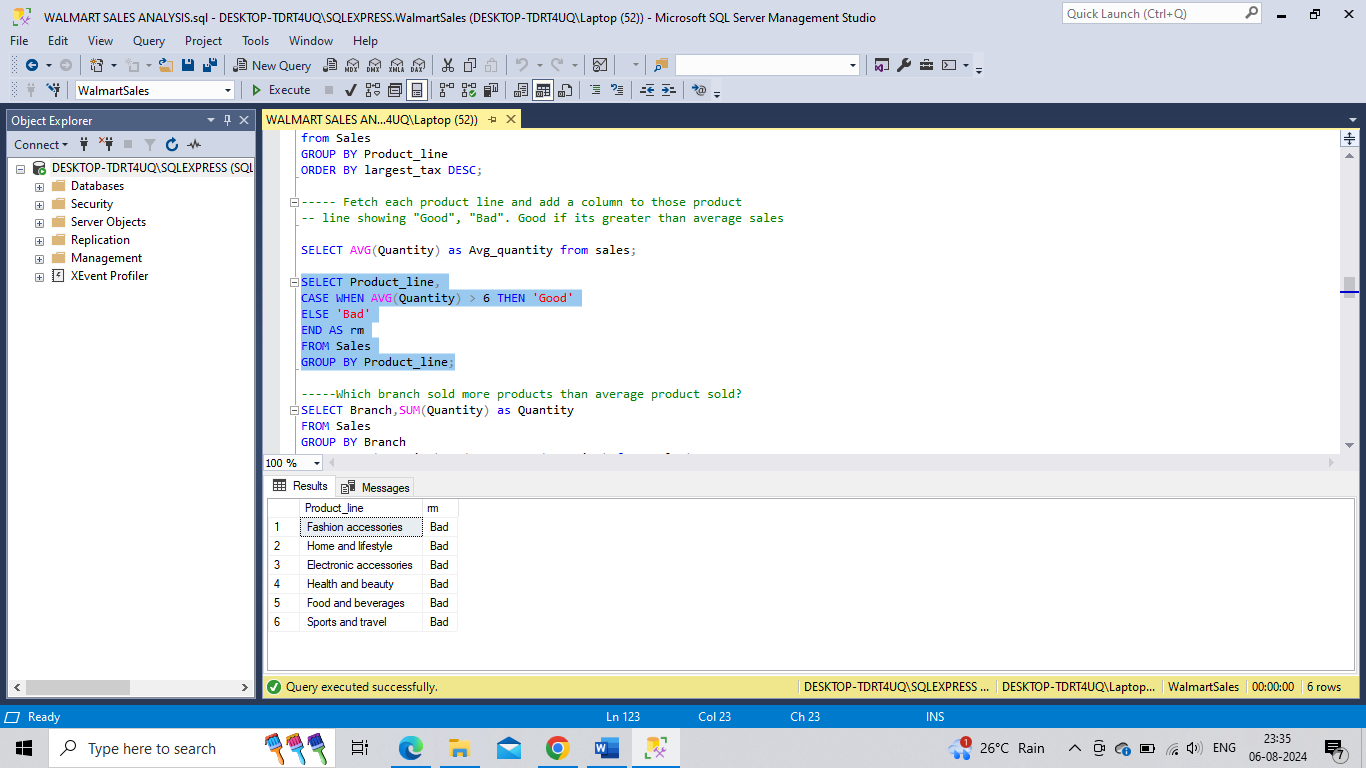


**Q.10. Fetch each product line and add a column to those product line showing "Good", "Bad". Good if its greater than average sales**

SELECT Product\_line,

CASE WHEN AVG(Quantity) > 6 THEN 'Good' ELSE 'Bad'

END AS rm FROM Sales GROUP BY Product\_line;



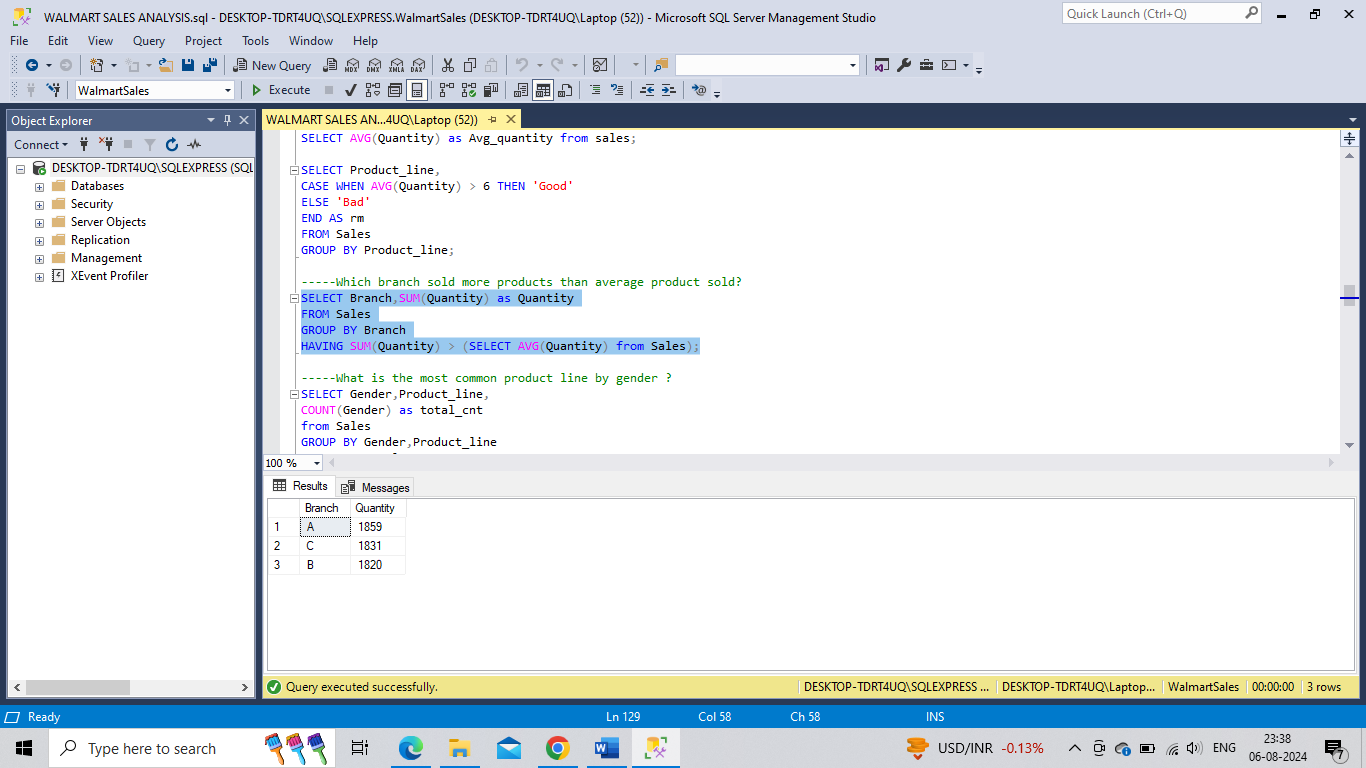
**Q.11. Which branch sold more products than average product sold?**

SELECT Branch,SUM(Quantity) as Quantity

FROM Sales

GROUP BY Branch

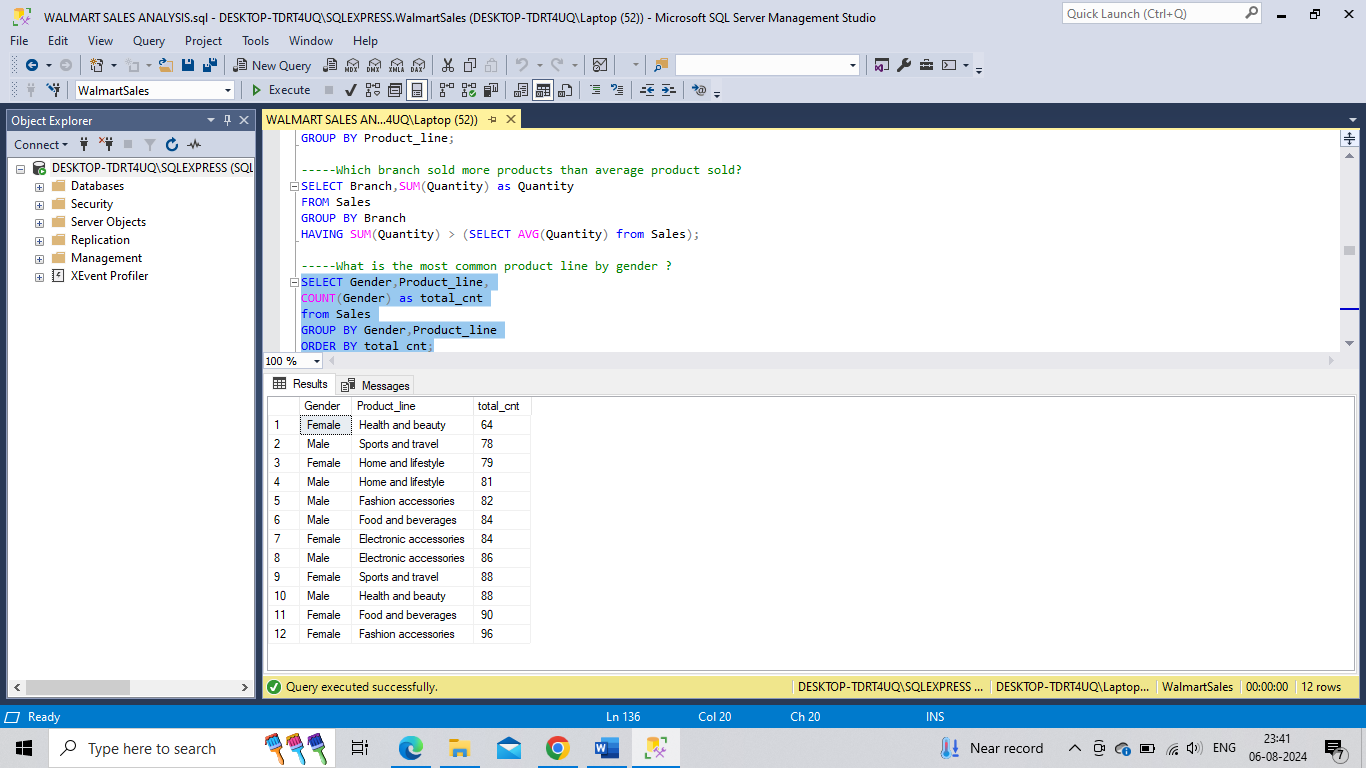
HAVING SUM(Quantity) > (SELECT AVG(Quantity) from Sales);



**Q.12. What is the most common product line by gender ?**

SELECT Gender,Product\_line, COUNT(Gender) as total\_cnt from Sales

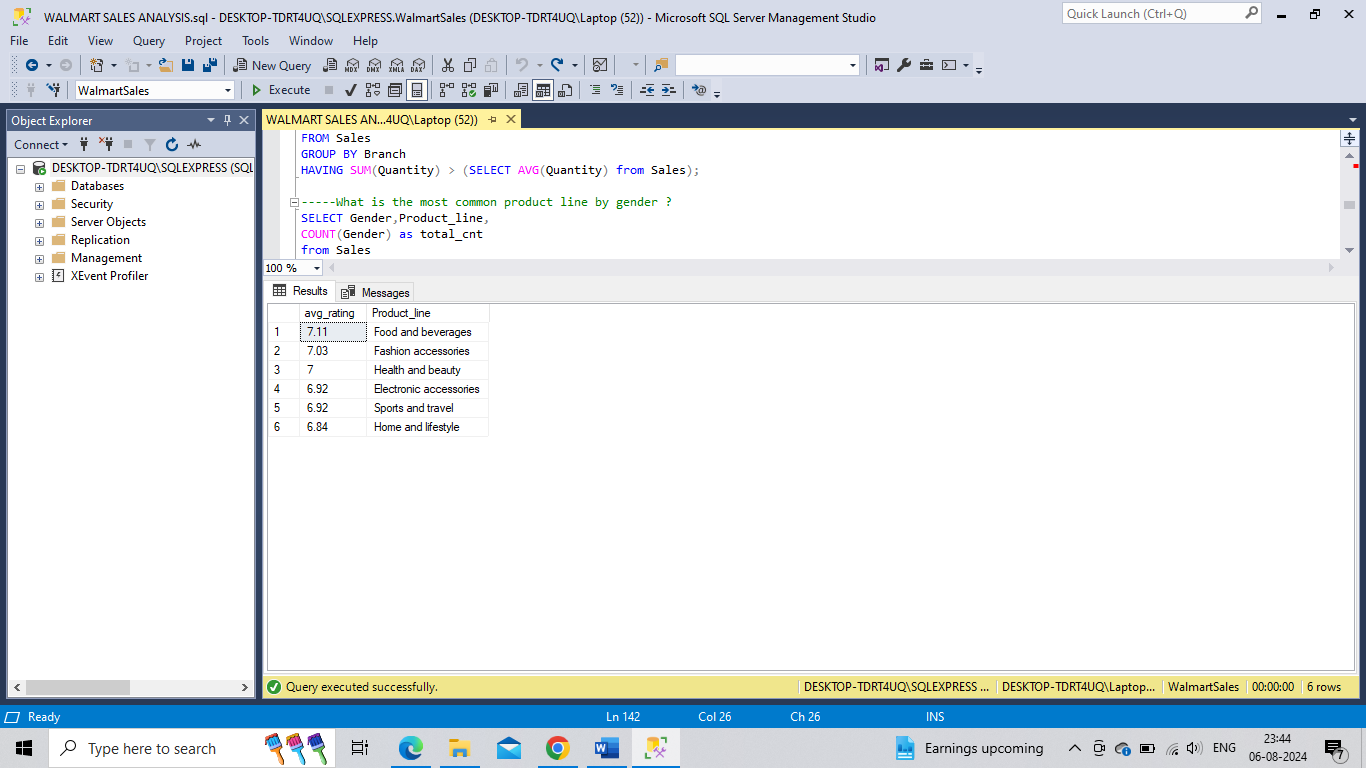
GROUP BY Gender,Product\_line ORDER BY total\_cnt;



**Q.13. What is the average rating of each product line**

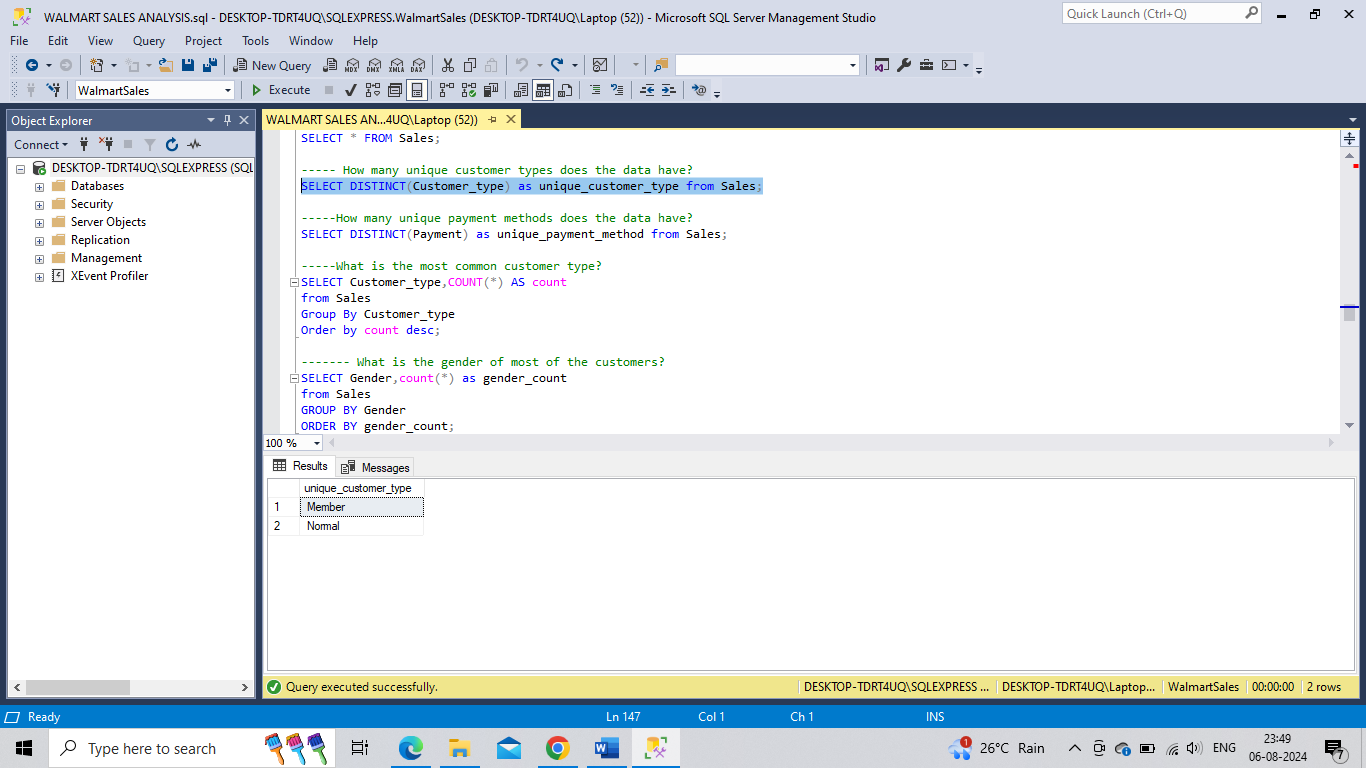
SELECT ROUND(AVG(Rating), 2)as avg\_rating ,Product\_line FROM

Sales GROUP BY Product\_line ORDER BY avg\_rating DESC;



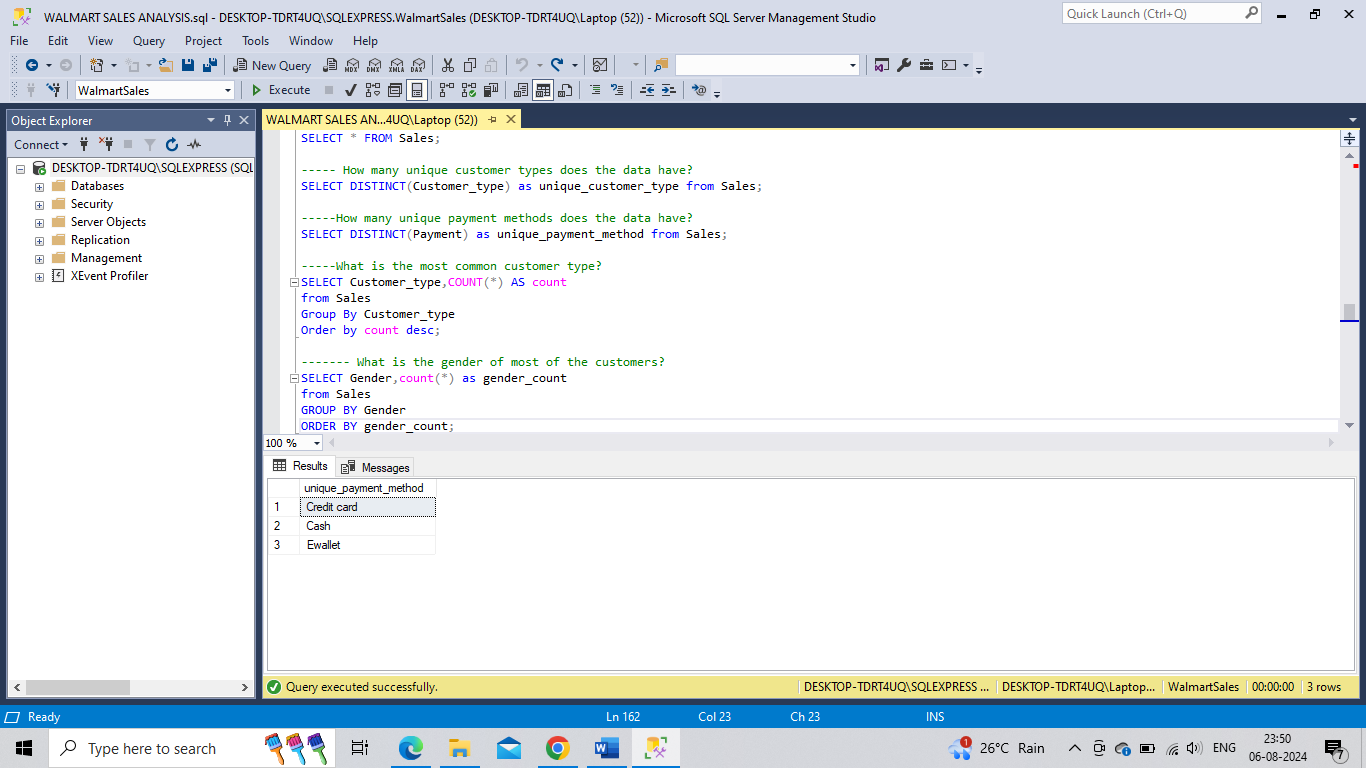
**Q.14. How many unique customer types does the data have?**

SELECT DISTINCT(Customer\_type) as unique\_customer\_type from Sales;



**Q.15. How many unique payment methods does the data have?**

SELECT DISTINCT(Payment) as unique\_payment\_method from Sales;

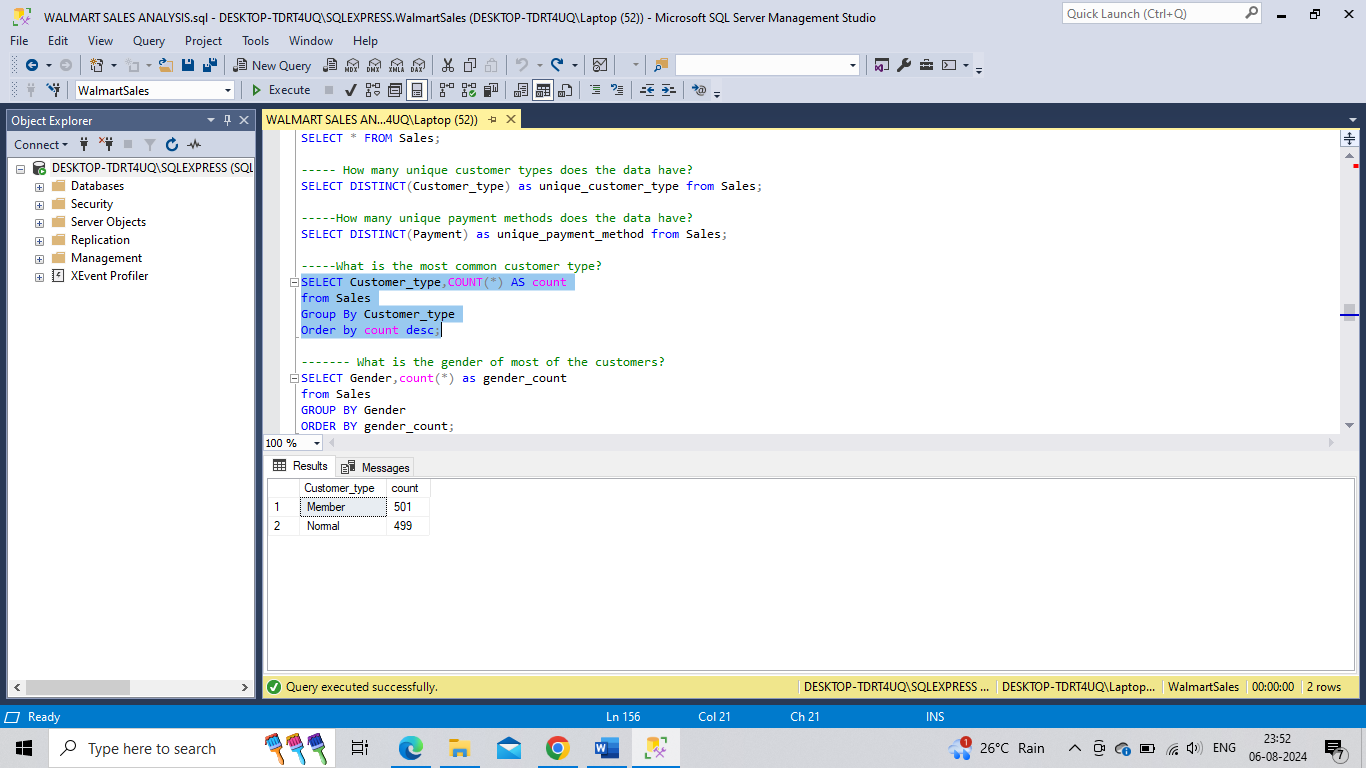


**Q.16. What is the most common customer type?**

SELECT Customer\_type,COUNT(\*) AS count

from Sales

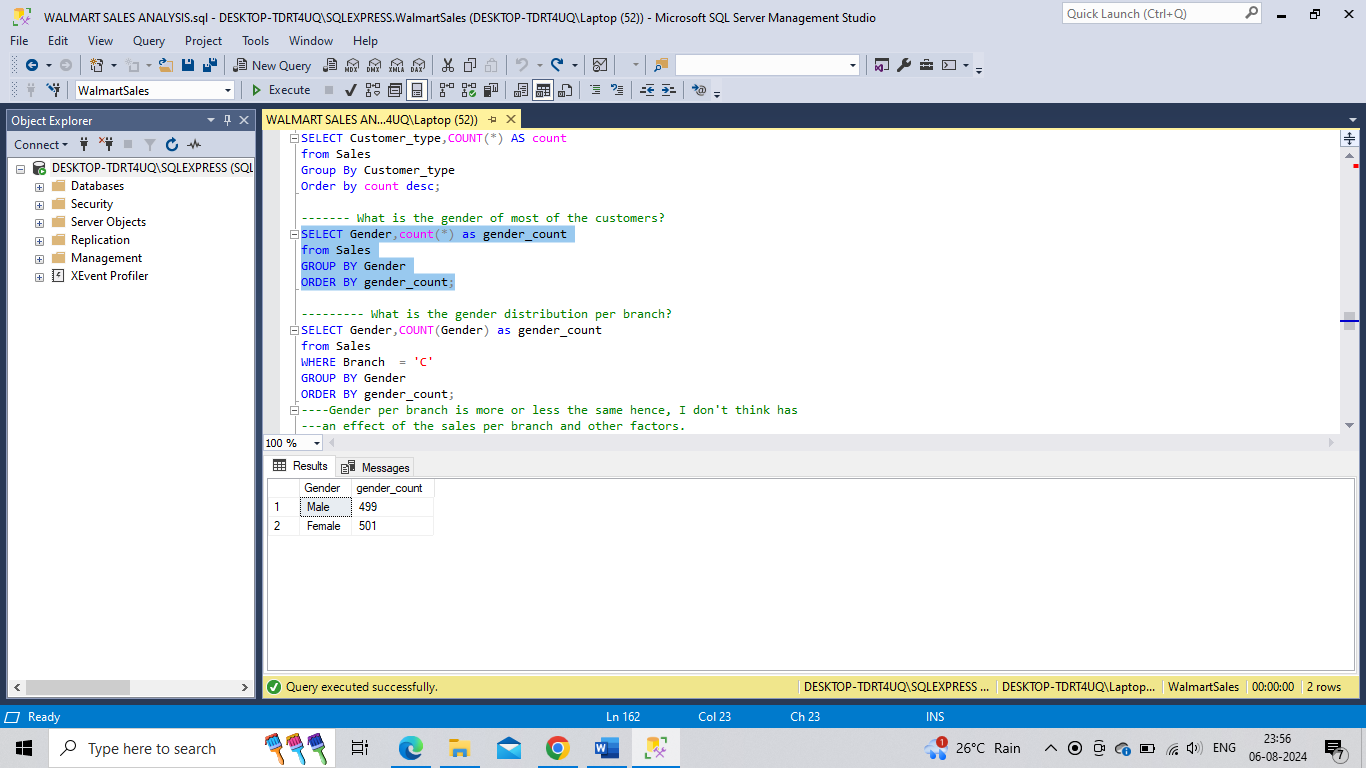
Group By Customer\_type Order by count desc;



Q.17. **What is the gender of most of the customers?**

SELECT Gender,count(\*) as gender\_count

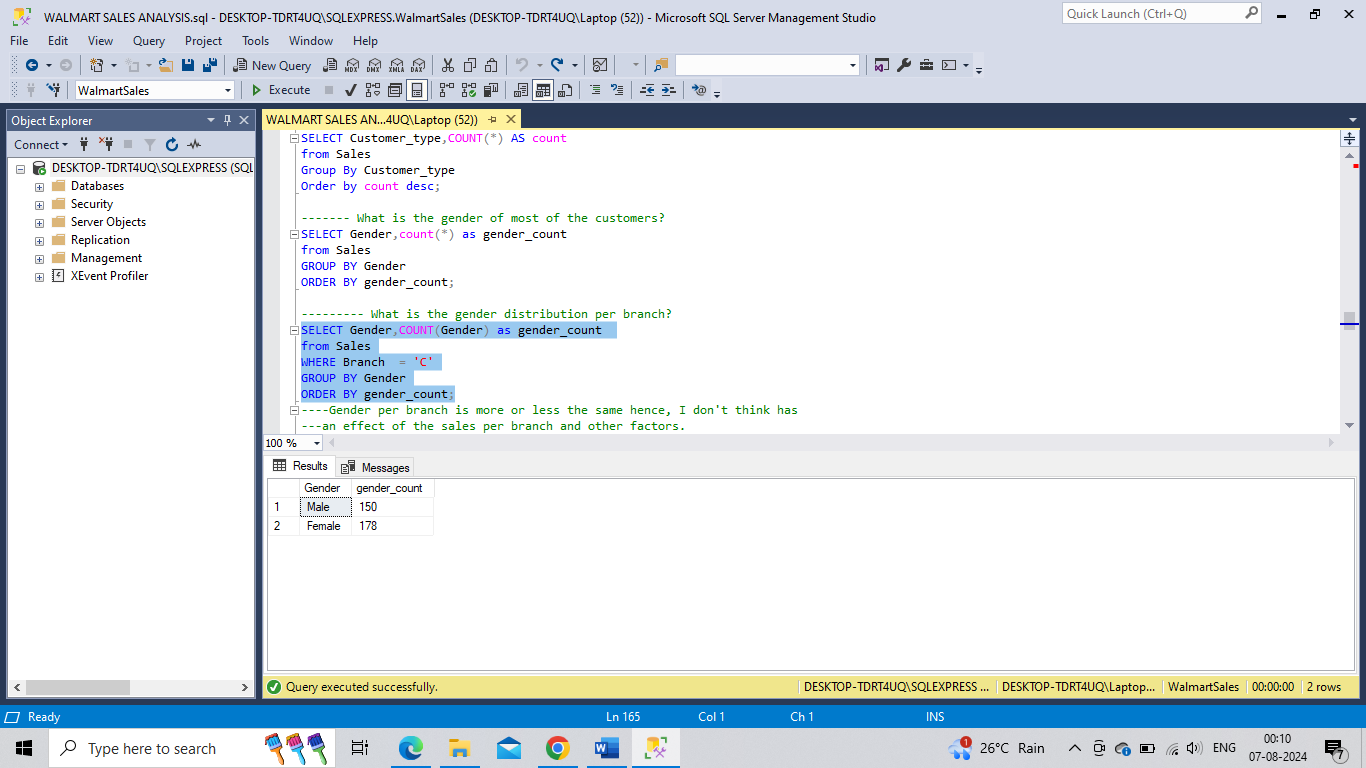
from Sales GROUP BY Gender ORDER BY gender\_count;



**Q.18.** **What is the gender distribution per branch?**

SELECT Gender,COUNT(Gender) as gender\_count

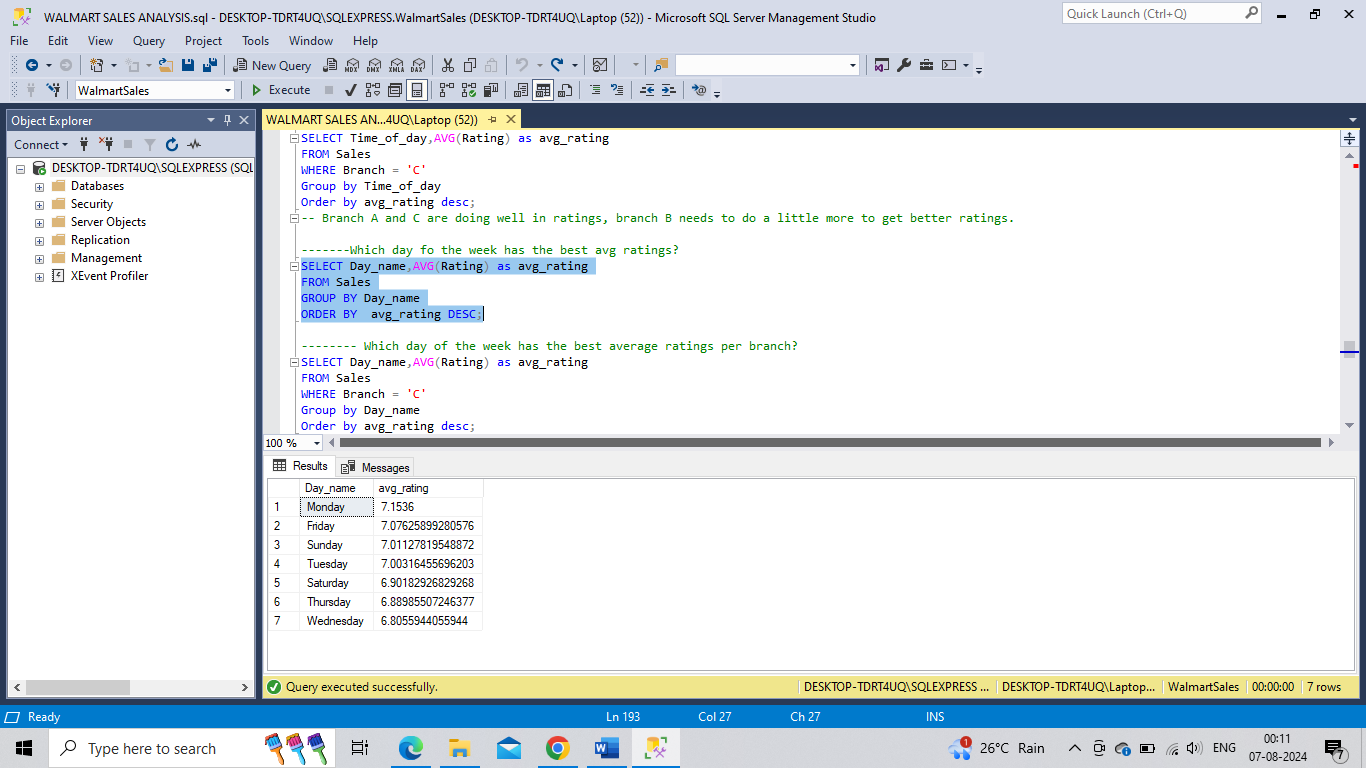
from Sales WHERE Branch = 'C' GROUP BY Gender ORDER BY gender\_count;



**Q.19. Which day of the week has the best avg ratings?**

SELECT Day\_name,AVG(Rating) as avg\_rating FROM Sales

GROUP BY Day\_name ORDER BY avg\_rating DESC;



**Q.20. Which day of the week has the best average ratings per branch?**

SELECT Day\_name,AVG(Rating) as avg\_rating FROM Sales

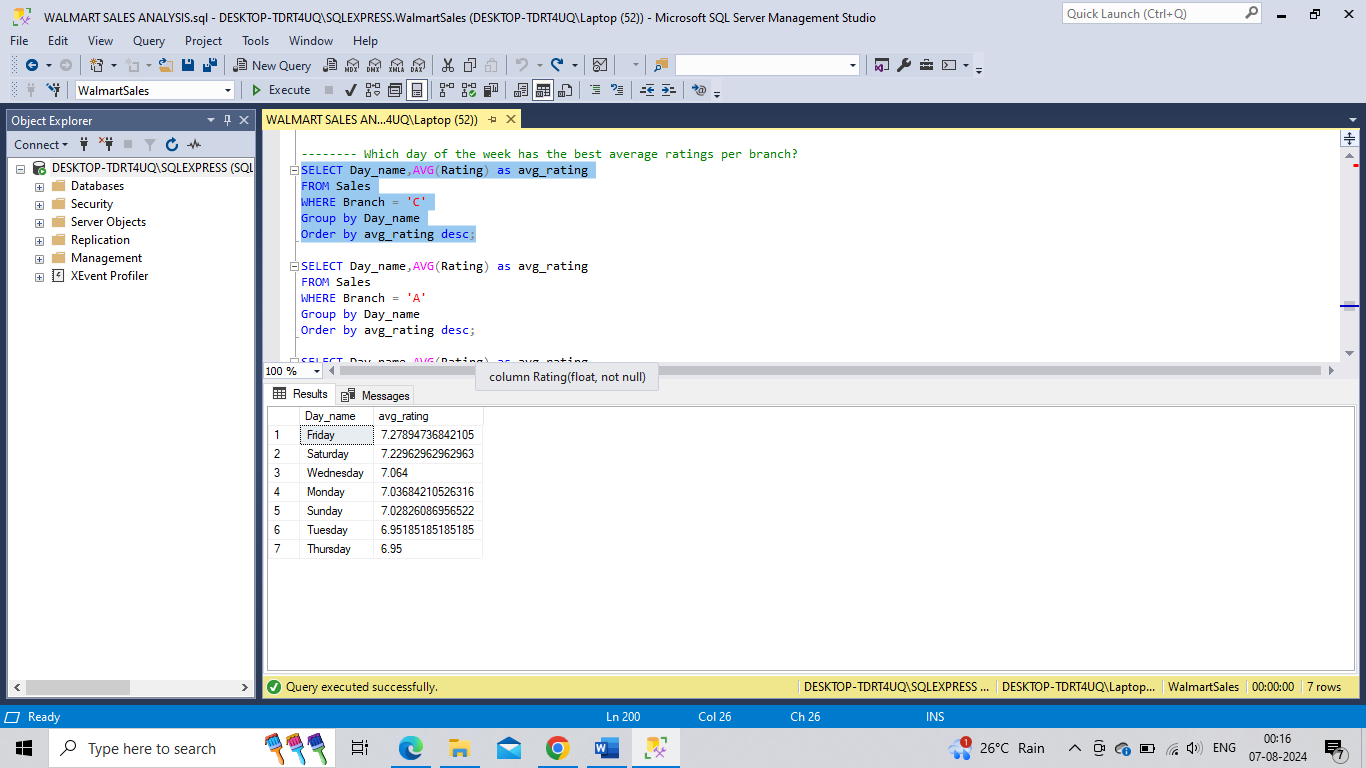
WHERE Branch = 'C' Group by Day\_name Order by avg\_rating desc;

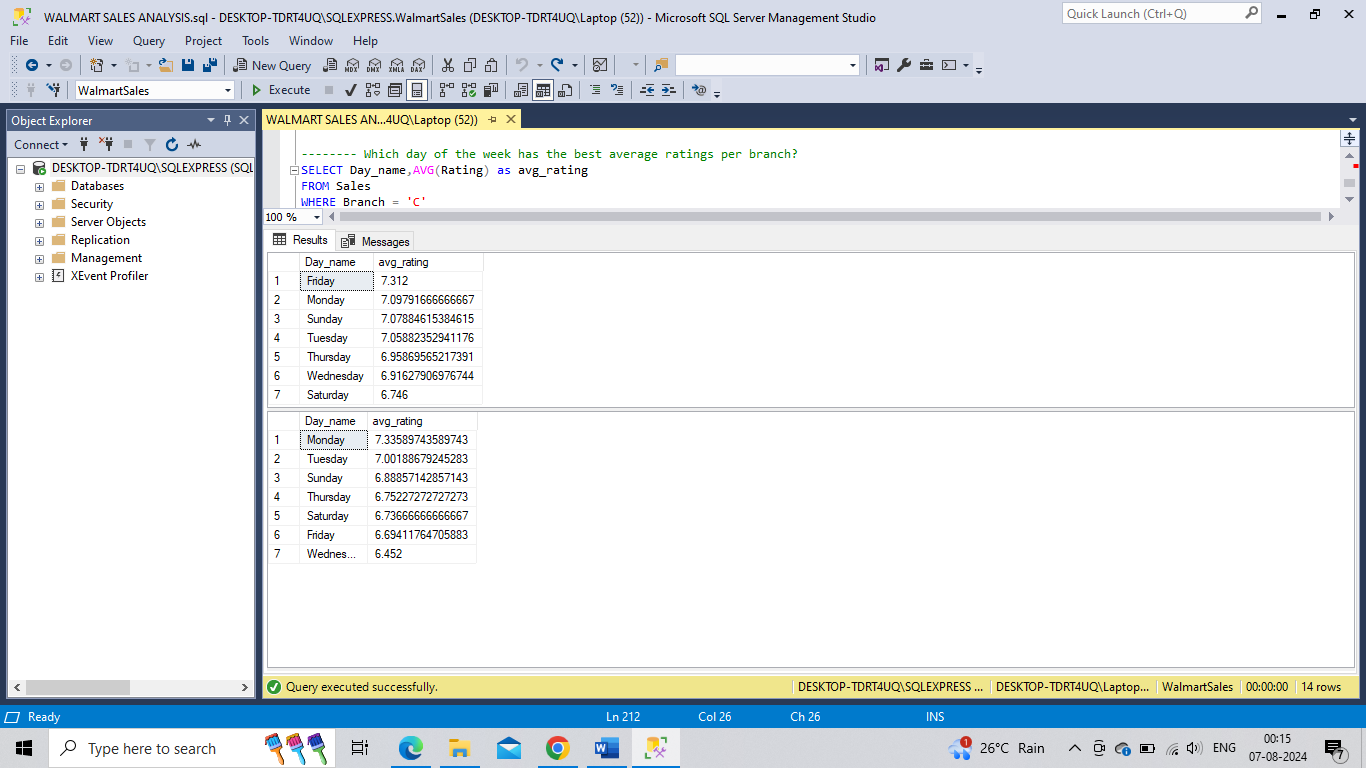
SELECT Day\_name,AVG(Rating) as avg\_rating FROM Sales WHERE Branch = 'A'

Group by Day\_name Order by avg\_rating desc;

SELECT Day\_name,AVG(Rating) as avg\_rating FROM Sales

WHERE Branch = 'B' Group by Day\_name Order by avg\_rating desc;

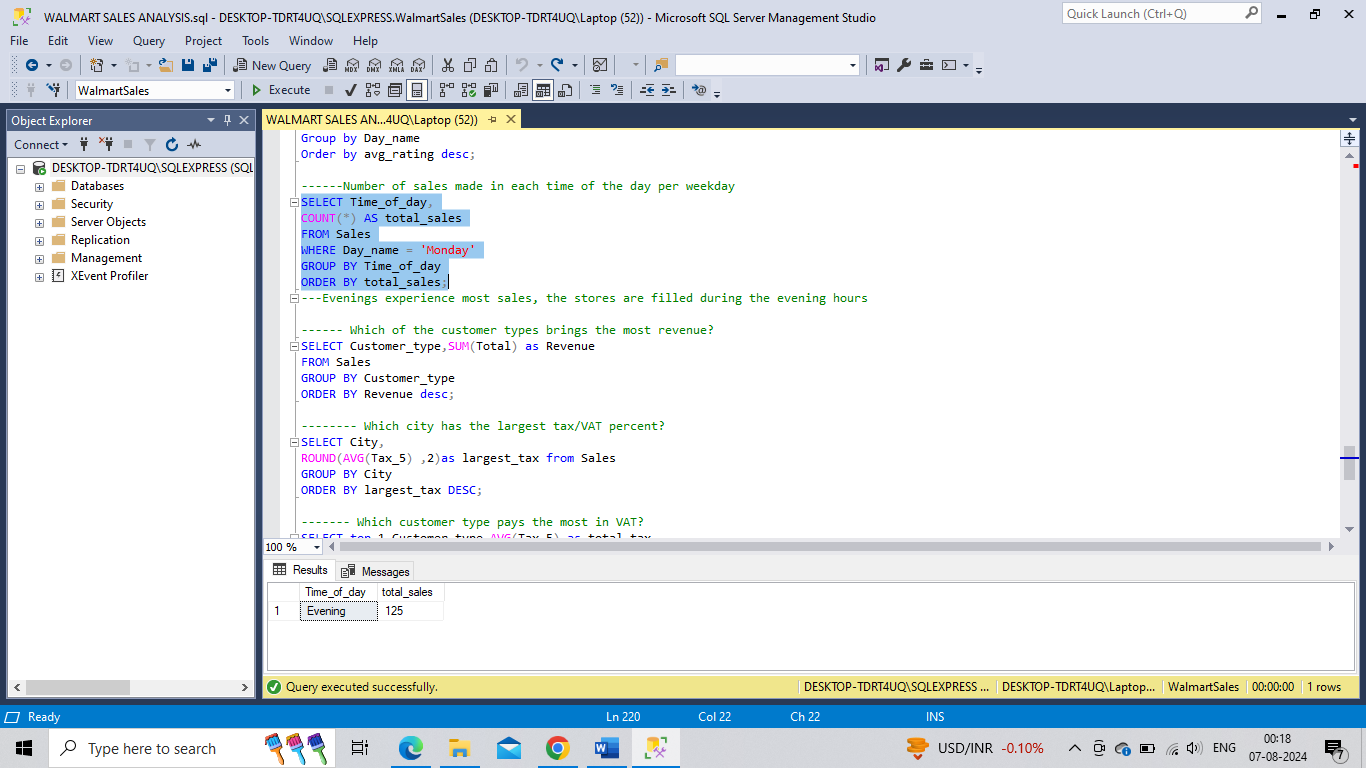




**Q.21. Number of sales made in each time of the day per weekday**

SELECT Time\_of\_day, COUNT(\*) AS total\_sales FROM Sales

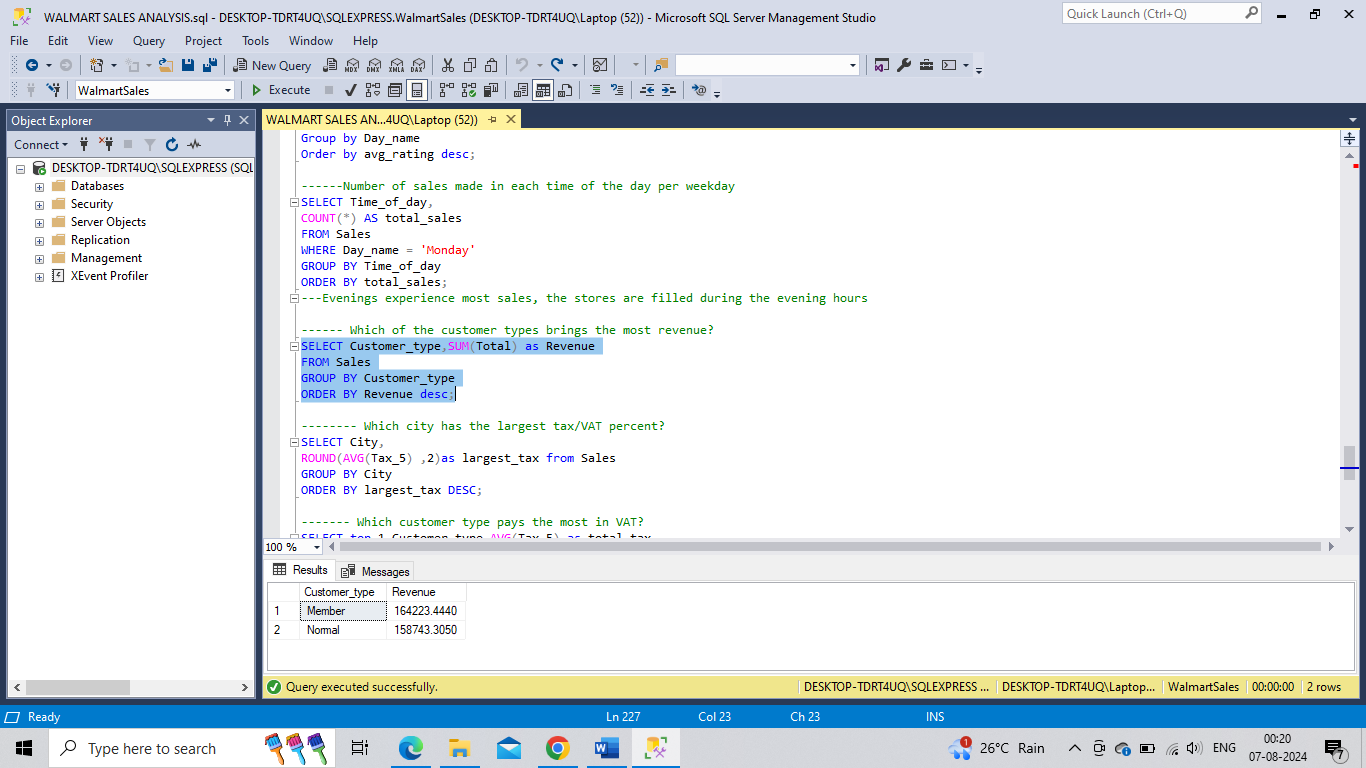
WHERE Day\_name = 'Monday' GROUP BY Time\_of\_day ORDER BY total\_sales;



**Q.22. Which of the customer types brings the most revenue?**

SELECT Customer\_type,SUM(Total) as Revenue FROM Sales

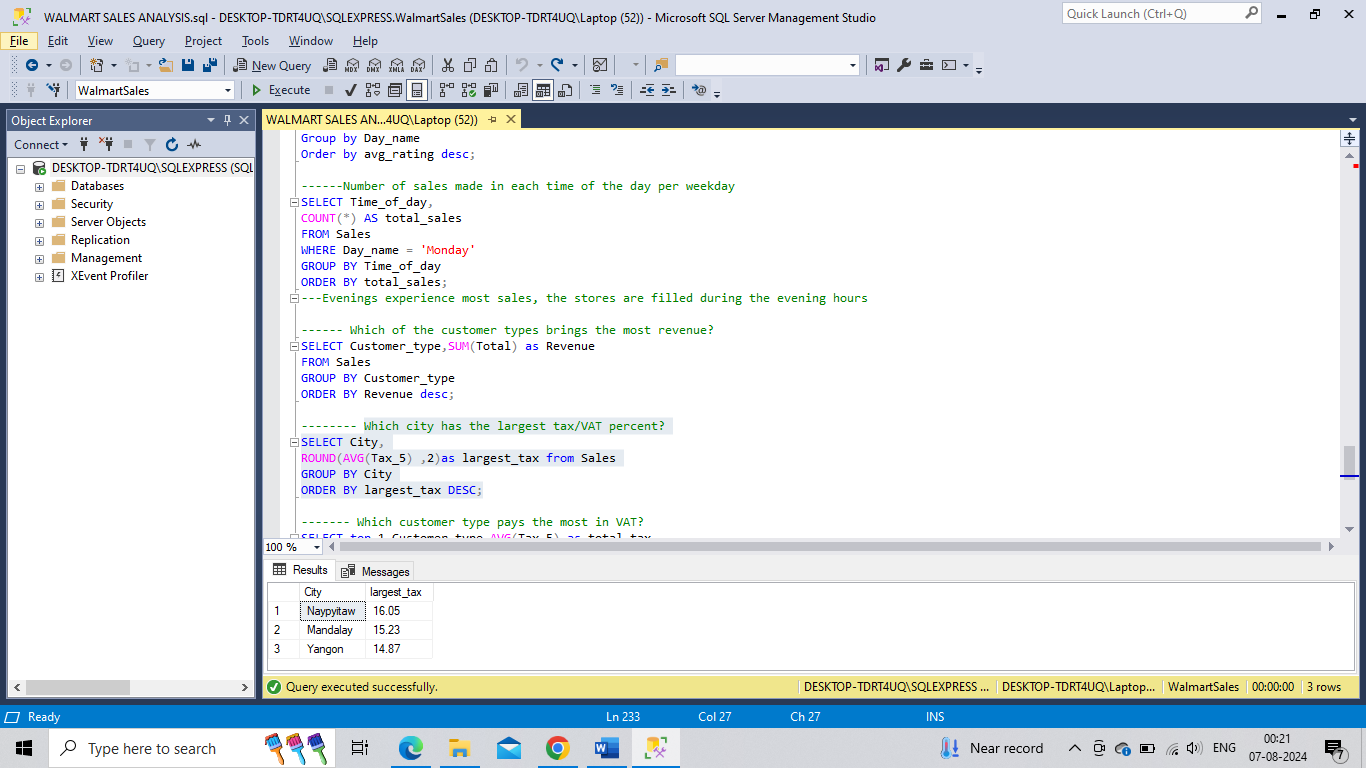
GROUP BY Customer\_type ORDER BY Revenue desc;



**Q.23.** **Which city has the largest tax/VAT percent?**

SELECT City, ROUND(AVG(Tax\_5) ,2)as largest\_tax from Sales

GROUP BY City ORDER BY largest\_tax DESC;



**Q.24. Which customer type pays the most in VAT?**

SELECT top 1 Customer\_type,AVG(Tax\_5) as total\_tax

from Sales

GROUP BY Customer\_type

ORDER BY total\_tax desc;

