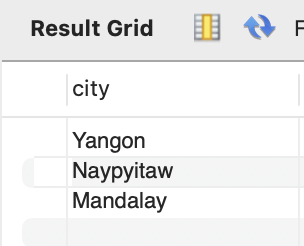
Nguyễn Văn Thống

### **Generic Question**

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

SELECT DISTINCT (city) AS city

FROM walmart



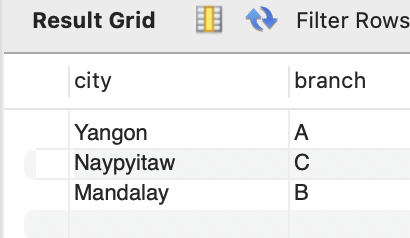
**2. In which city is each branch?**

SELECT

DISTINCT (city),

branch

FROM walmart

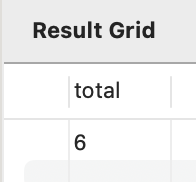


### **Product**

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

SELECT COUNT(DISTINCT `product line`) AS total

FROM walmart



**4. What is the most common payment method?**

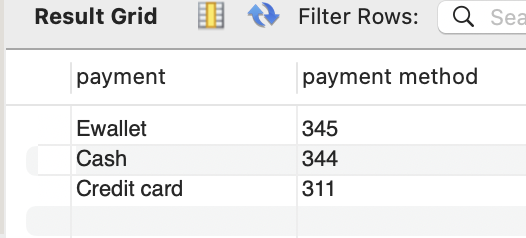
SELECT `payment`,

count(\*) AS `payment method`

FROM walmart

GROUP BY 1

ORDER BY 2 DESC



**5. What is the most selling product line?**

SELECT

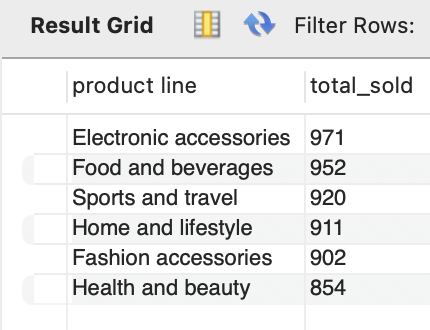
`product line`,

SUM(`quantity`) AS total\_sold

FROM walmart

GROUP BY 1

ORDER BY 2 DESC



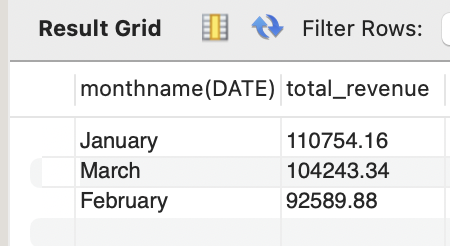
**6. What is the total revenue by month?**

SELECT monthname(DATE),

round(sum(`unit price` \* `quantity`),2) AS total\_revenue

FROM walmart

GROUP BY 1



**7. What month had the largest COGS?**

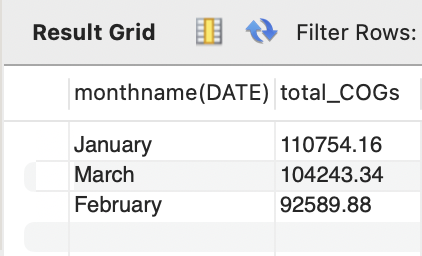
SELECT monthname(DATE),

round(sum(cogs),2) AS total\_COGs

FROM walmart

GROUP BY 1

ORDER BY 2 DESC



**8. What product line had the largest revenue?**

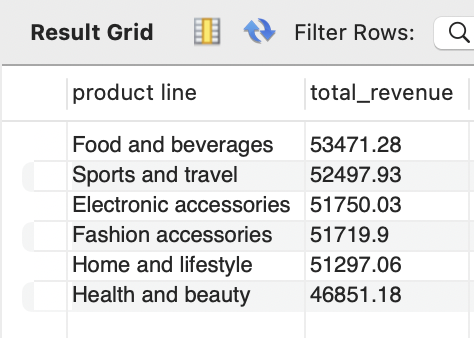
SELECT `product line`,

round(sum(`unit price` \* `quantity`),2) AS total\_revenue

FROM walmart

GROUP BY 1

ORDER BY 2 DESC



**9. What is the city with the largest revenue?**

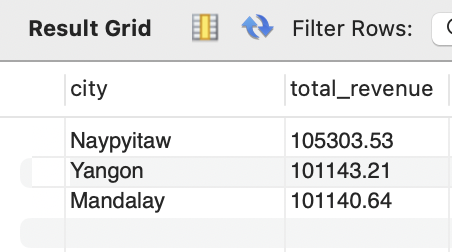
SELECT `city`,

round(sum(`unit price` \* `quantity`),2) AS total\_revenue

FROM walmart

GROUP BY 1

ORDER BY 2 DESC



**10. What product line had the largest VAT?**

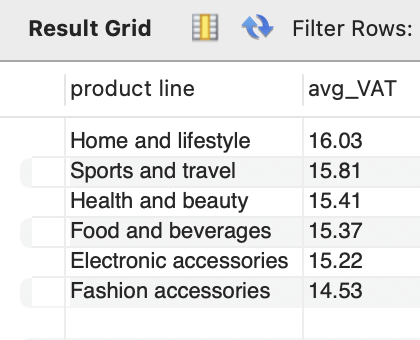
SELECT `product line`,

ROUND(AVG(`tax 5%`),2) AS avg\_VAT

FROM walmart

GROUP BY `product line`

ORDER BY avg\_VAT DESC;



**11. 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`,

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 walmart

GROUP BY `product line`) AS subquery)

THEN 'Good'

ELSE 'Bad'

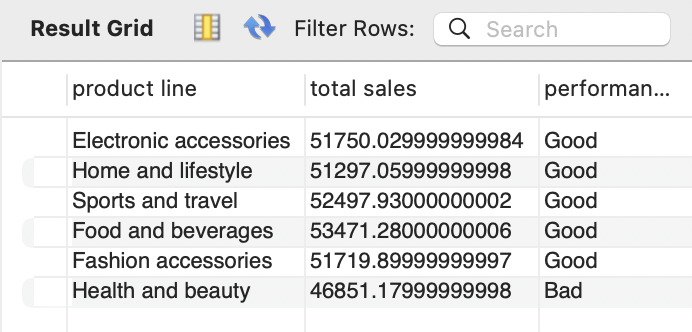
END AS performance

FROM walmart

GROUP BY `product line`

ORDER BY

FIELD(performance, 'Good', 'Bad');



**12. Which branch sold more products than average product sold?**

SELECT AVG(total\_quantity) AS avg\_quantity\_per\_branch

FROM (

SELECT `branch`, SUM(`quantity`) AS total\_quantity

FROM walmart

GROUP BY `branch`

) AS branch\_totals;

SELECT `branch`,

SUM(`quantity`) AS total\_quantity

FROM walmart

GROUP BY `branch`

HAVING total\_quantity > (

SELECT AVG(total\_quantity)

FROM (

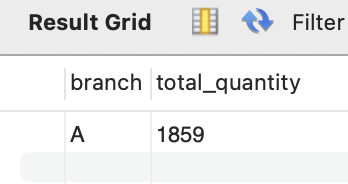
SELECT SUM(`quantity`) AS total\_quantity

FROM walmart

GROUP BY `branch`

) AS subquery

);



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

SELECT `gender`,

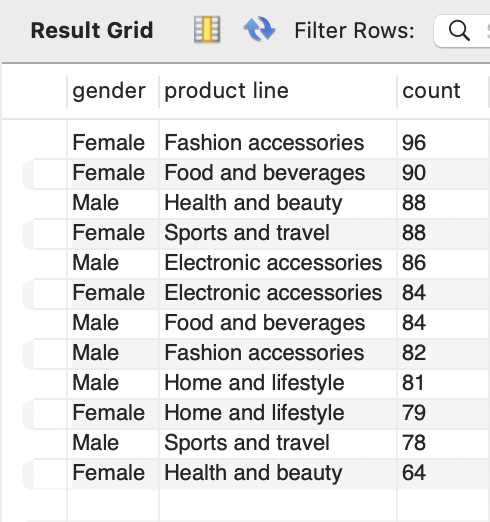
`product line`,

COUNT(gender) AS count

FROM walmart

GROUP BY `gender`, `product line`

ORDER BY count DESC;



**14. What is the average rating of each product line?**

SELECT

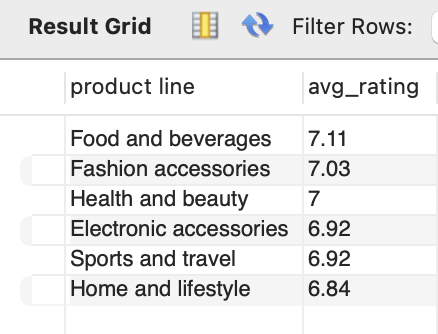
`product line`,

round(AVG(rating),2) AS avg\_rating

FROM walmart

GROUP BY 1

ORDER BY 2 DESC



### **Sales**

SET sql\_safe\_updates = 0

ALTER TABLE walmart ADD COLUMN `time day` varchar(20);

UPDATE walmart

SET `time day` = (

CASE

WHEN `time` BETWEEN "00:00:00" AND "12:00:00" THEN "Morning"

WHEN `time` BETWEEN "12:00:01" AND "18:00:00" THEN "Afternoon"

ELSE "Evening"

END

);

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

SELECT

DAYNAME(`date`) AS weekday,

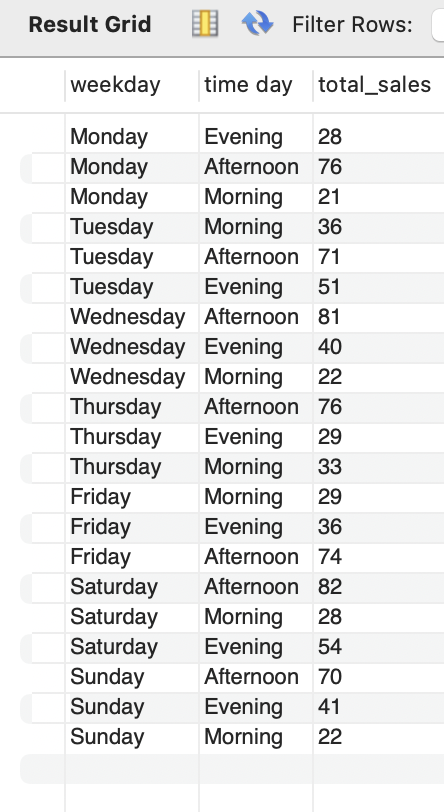
`time day`,

COUNT(\*) AS total\_sales

FROM walmart

GROUP BY 1,2

ORDER BY FIELD(weekday, 'Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday', 'Sunday');



**16. Which of the customer types brings the most revenue?**

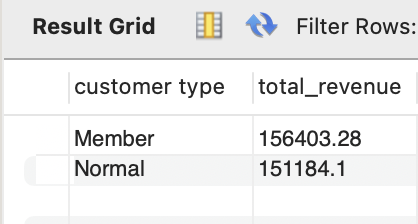
SELECT `customer type`,

ROUND(SUM(`unit price` \* `quantity`),2) AS total\_revenue

FROM walmart

GROUP BY 1

ORDER BY 2 DESC;



**17. Which city has the largest tax percent/ VAT (Value Added Tax)?**

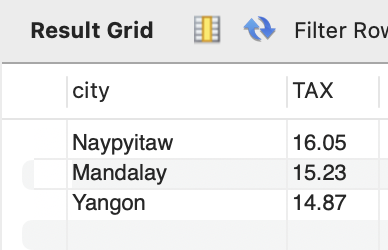
SELECT `city`,

ROUND(AVG(`Tax 5%`),2) AS TAX

FROM walmart

GROUP BY 1

ORDER BY 2 DESC;



**18. Which customer type pays the most in VAT?**

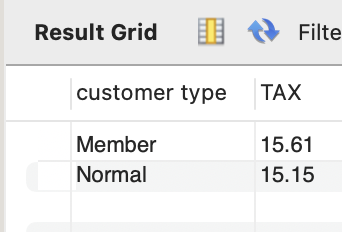
SELECT `customer type`,

ROUND(AVG(`Tax 5%`),2) AS TAX

FROM walmart

GROUP BY 1

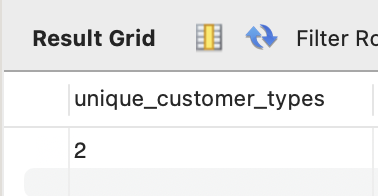
ORDER BY 2 DESC;



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

SELECT COUNT(DISTINCT `customer type`) AS unique\_customer\_types

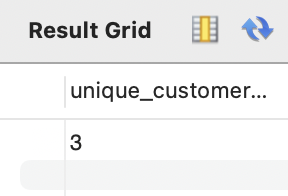
FROM walmart;



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

SELECT COUNT(DISTINCT `payment`) AS unique\_customer\_types

FROM walmart;



**21. What is the most common customer type?**

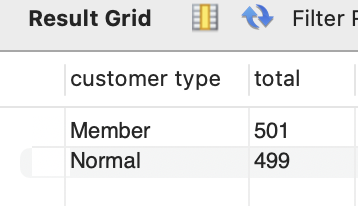
SELECT `customer type`,

COUNT(\*) AS total

FROM walmart

GROUP BY 1

ORDER BY 2 DESC;



**22. Which customer type buys the most?**

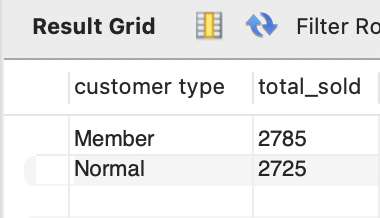
SELECT `customer type`,

SUM(quantity) AS total\_sold

FROM walmart

GROUP BY 1

ORDER BY 2 DESC;



**23. What is the gender of most of the customers?**

SELECT

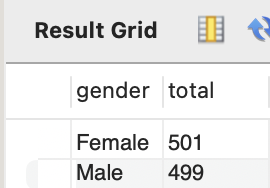
gender,

COUNT(\*) AS total

FROM walmart

GROUP BY 1

ORDER BY 2 DESC



**24. What is the gender distribution per branch?**

SELECT `branch`,

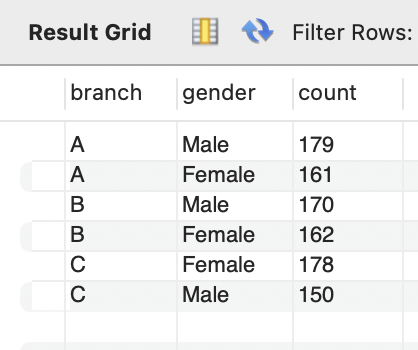
`gender`,

COUNT(\*) AS count

FROM walmart

GROUP BY 1,2

ORDER BY 1,3 DESC;



**25. Which time of the day do customers give most ratings?**

SELECT

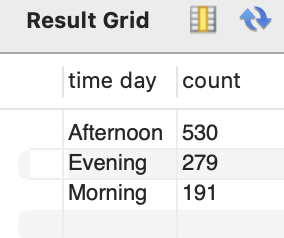
`time day`,

COUNT(rating) AS count

FROM walmart

GROUP BY 1

ORDER BY 2 DESC;



**26. Which time of the day do customers give most ratings per branch?**

SELECT `branch`,

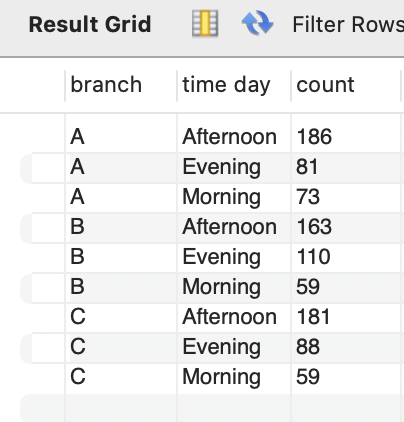
`time day`,

COUNT(\*) AS count

FROM walmart

GROUP BY 1,2

ORDER BY 1, 3 DESC;



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

SELECT DAYNAME(`date`) AS `Day of Week`,

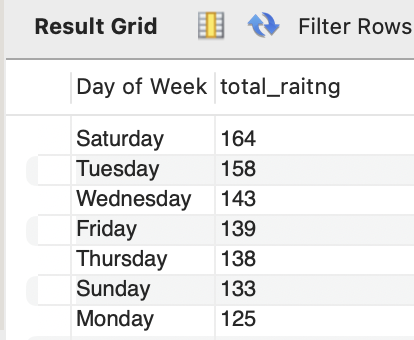
COUNT(`rating`) AS total\_raitng

FROM walmart

WHERE `rating` IS NOT NULL

GROUP BY 1

ORDER BY 2 DESC;

ß

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

SELECT

`branch`,

DAYNAME(`date`) AS `Day of Week`,

ROUND(AVG(`rating`),2) AS `AVG Rating`

FROM walmart

WHERE `rating` IS NOT NULL

GROUP BY 1, 2

ORDER BY 1 , 3 DESC;

