



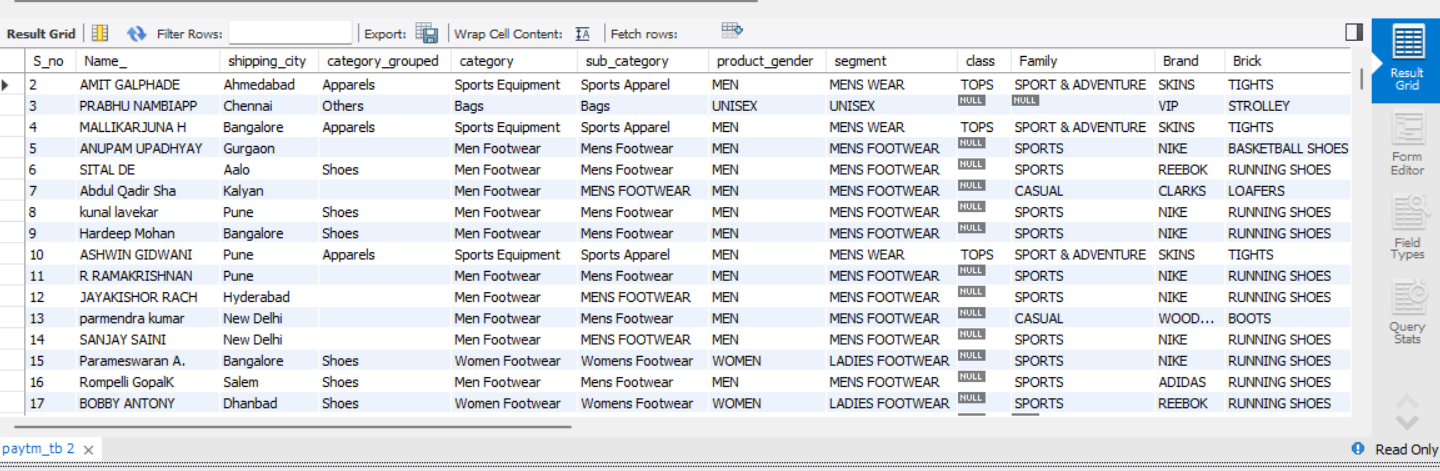
INTERNSHIP PROJECT

**Thenmozhi R**

**BY**

PAYTM EPURCHASE DATA ANALYSIS USING SQL

LOADING THE DATA



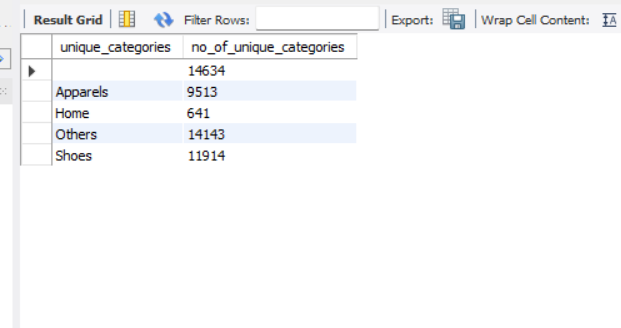
1. What does the "Category\_Grouped" column represent, and how many unique categories are there?

select distinct(category\_grouped)as unique\_categories,count(\*)as no\_of\_unique\_categories

from paytm\_tb

group by category\_grouped

order by unique\_categories;



2. Can you list the top 5 shipping cities in terms of the number of orders?

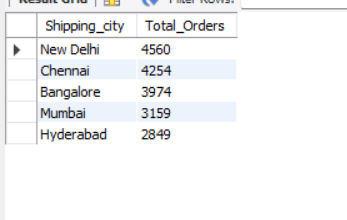
SELECT Shipping\_city ,COUNT(Shipping\_city) as Total\_Orders

FROM paytm\_tb

GROUP BY Shipping\_city

ORDER BY Total\_Orders desc

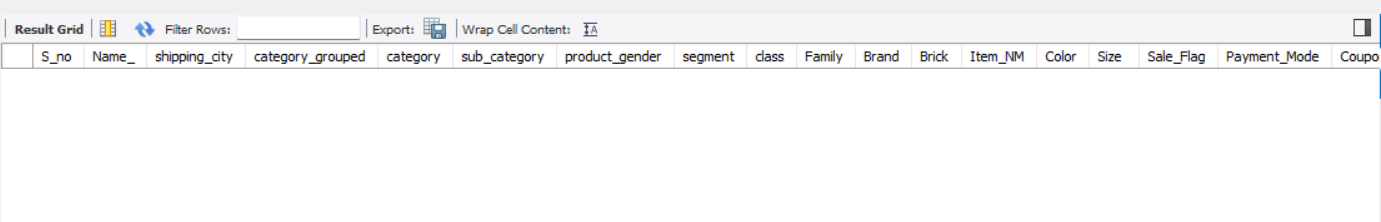
LIMIT 5 ;



3. Show me a table with all the data for products that belong to the "Electronics" category.

SELECT \* FROM paytm\_tb

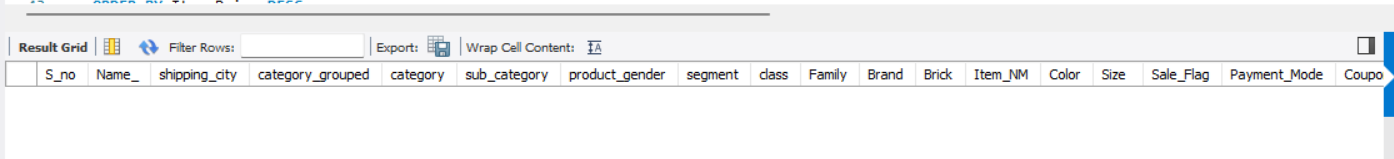
WHERE Category = "Electronics";



4. Filter the data to show only rows with a "Sale\_Flag" of 'Yes'.

SELECT \* FROM paytm\_tb

WHERE Sale\_Flag LIKE "Sale" ;



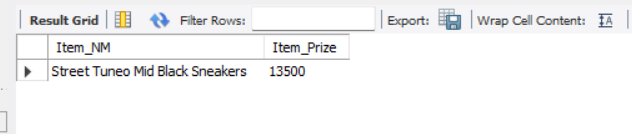
5. Sort the data by "Item\_Price" in descending order. What is the most expensive item?

SELECT Item\_NM , Item\_Prize

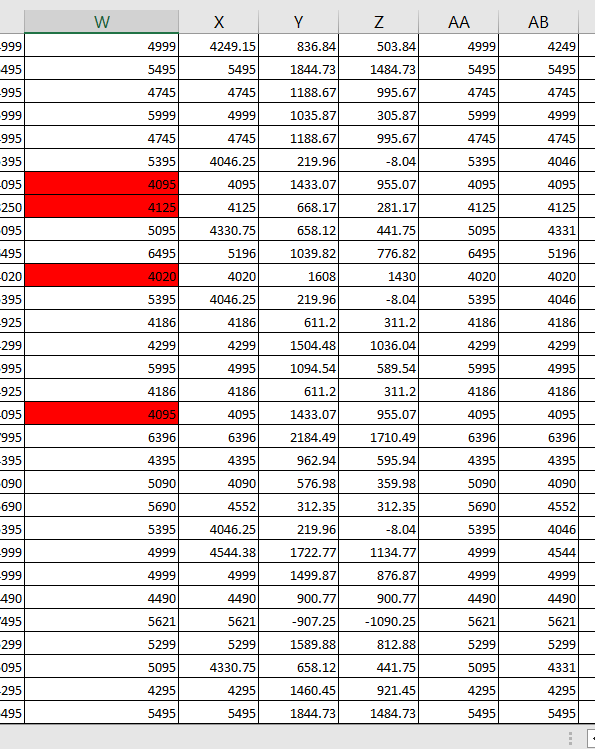
FROM paytm\_tb

ORDER BY Item\_Prize DESC

LIMIT 1 ;



6. Apply conditional formatting to highlight all products with a "Special\_Price\_effective" value below $50 in red.

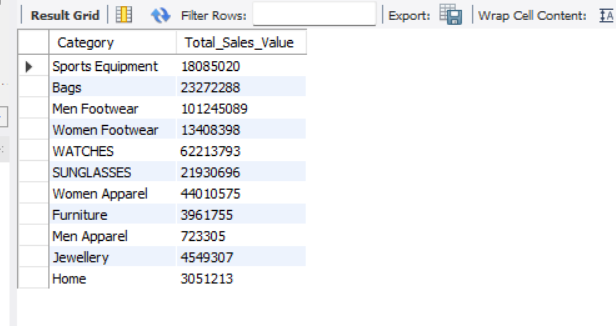


7. Create a pivot table to find the total sales value for each category.

SELECT Category , SUM(Item\_Prize) as Total\_Sales\_Value

FROM paytm\_tb

Group by Category ;



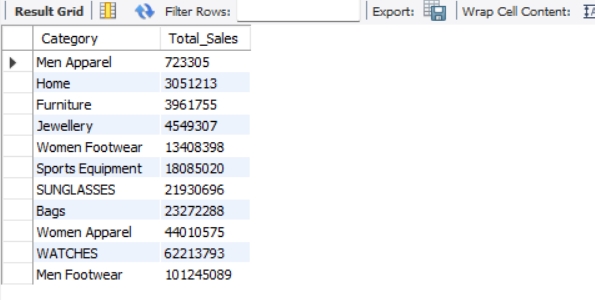
8. Create a bar chart to visualize the total sales for each category

SELECT Category , SUM(Item\_Prize) as Total\_Sales

FROM paytm\_tb

Group by Category

ORDER BY Total\_Sales;

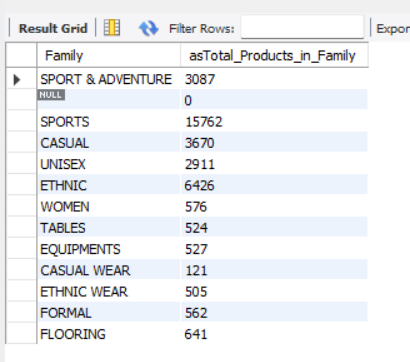


9. Create a pie chart to show the distribution of products in the "Family" category.

SELECT Family ,COUNT(Family)asTotal\_Products\_in\_Family

FROM paytm\_tb

Group by Family ;

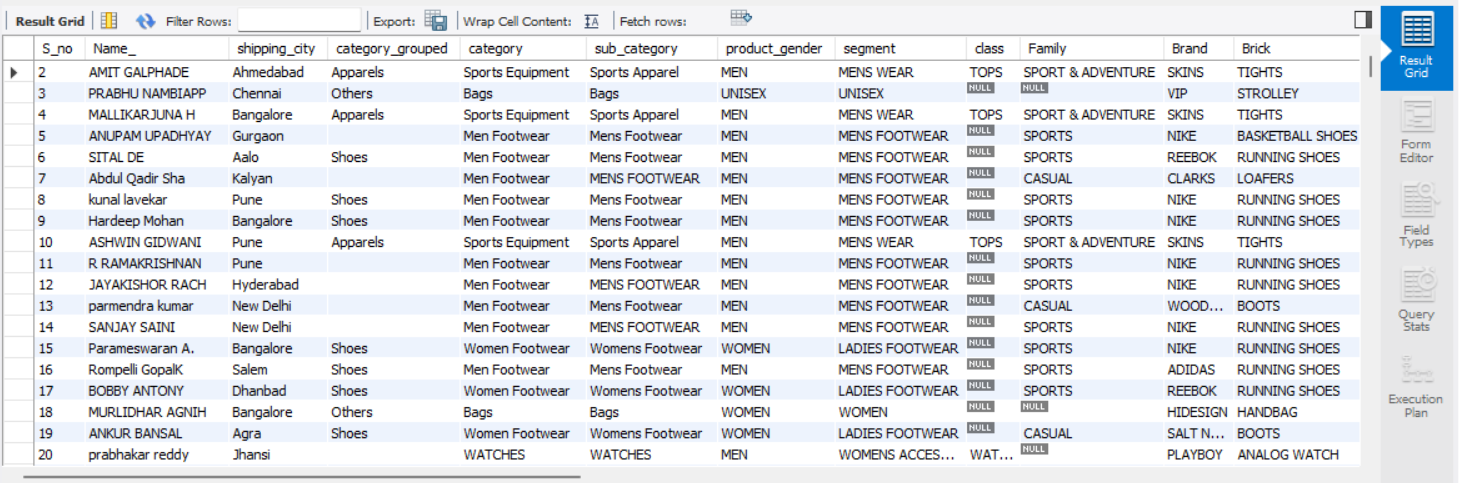


10. Ensure that the "Payment\_Method" column only contains valid payment methods (e.g., Visa, MasterCard)

select \*

from paytm\_tb

where Payment\_Mode in ("COD","Prepaid");



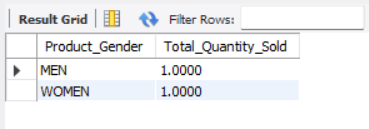
11. Calculate the average "Quantity" sold for products in the "Clothing" category, grouped by "Product\_Gender."

SELECT Product\_Gender , AVG(Quantity) as Total\_Quantity\_Sold

FROM paytm\_tb

WHERE Category\_Grouped ="Apparels"

GROUP BY Product\_Gender;



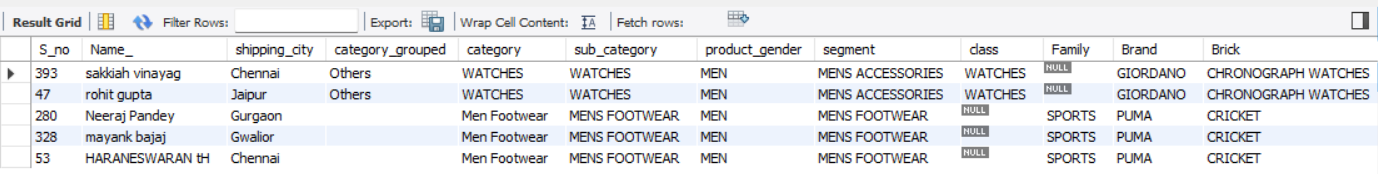
12. Find the top 5 products with the highest "Value\_CM1" and "Value\_CM2" ratios. Create a chart to visualize this data.

SELECT \*, ROUND((Value\_CM1/Value\_CM2),2) AS Ration

FROM paytm\_tb

ORDER BY Ration DESC

LIMIT 5 ;



13. Identify the top 3 "Class" categories with the highest total sales. Create a stacked bar chart to represent this data.

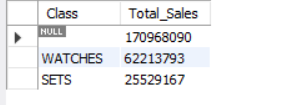
SELECT Class, SUM(Item\_Prize) AS Total\_Sales

FROM paytm\_tb

Group by Class

Order by Total\_Sales DESC

LIMIT 3 ;



14. Use VLOOKUP or INDEX-MATCH to retrieve the "Color" of a product with a specific "Item\_NM."

select Color

from paytm\_tb

where Item\_NM ="your specific Item\_NM";



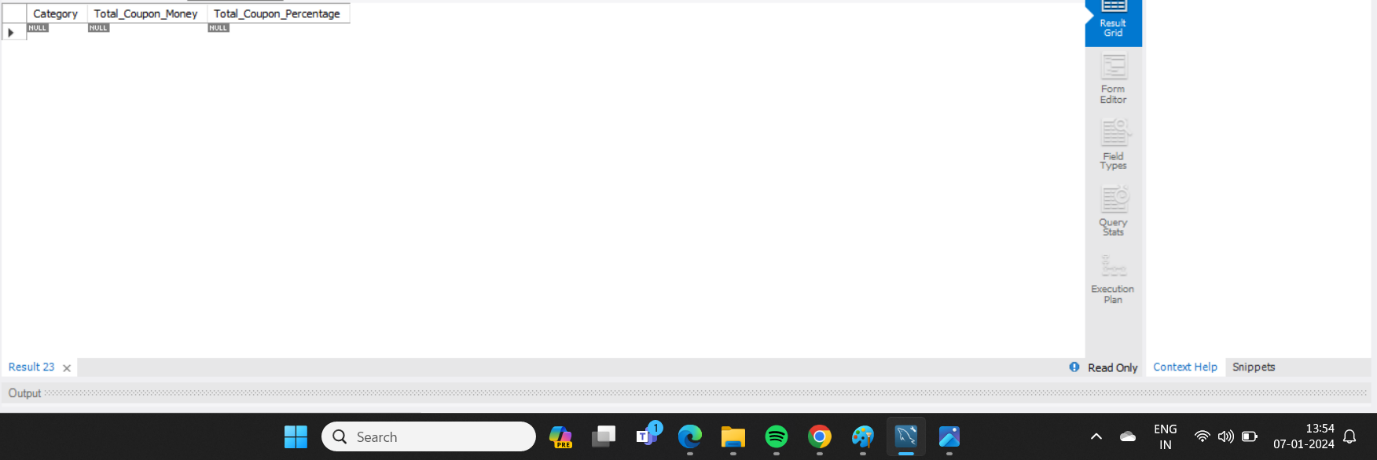
15. Calculate the total "coupon\_money\_effective" and "Coupon\_Percentage" for products in the "Electronics" category.

SELECT Category , SUM(coupon\_money\_effective) AS Total\_Coupon\_Money ,SUM(Coupon\_Percentage) AS

Total\_Coupon\_Percentage

FROM paytm\_tb

WHERE Category = "Electronics" ;



16. Perform a time series analysis to identify the month with the highest total sales.

select extract(Month from Sale\_Flag) as Month,

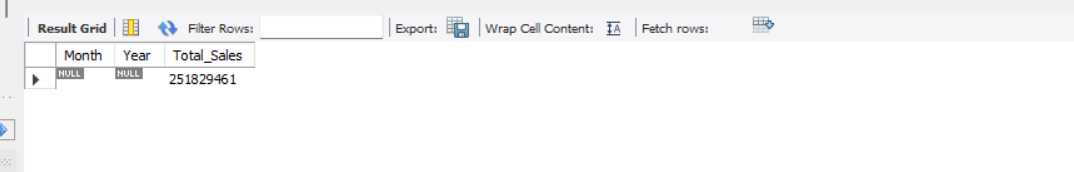
extract(Year from Sale\_Flag) as Year , Sum(Paid\_pr) as Total\_Sales

from paytm\_tb

group by Month,Year

order by Total\_Sales desc

limit 1;

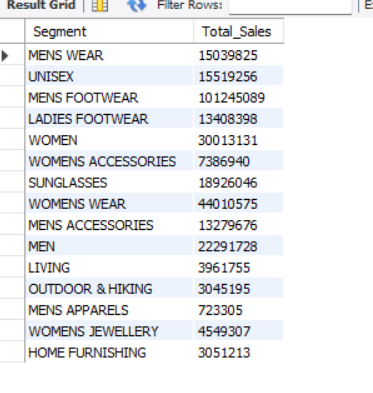


17. Calculate the total sales for each "Segment" and create a scatter plot to visualize the relationship between "Item\_Price" and "Quantity" in this data.

SELECT Segment , SUM(Item\_Prize) AS Total\_Sales

FROM paytm\_tb

GROUP BY Segment;



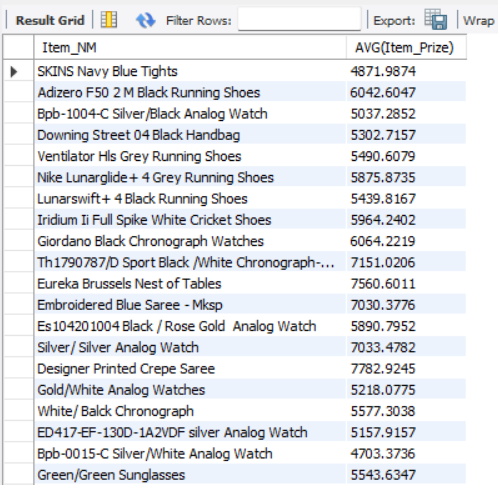
18. Use the AVERAGEIFS function to find the average "Item\_Price" for products that have a "Sale\_Flag" of 'Yes.'

SELECT Item\_NM,AVG(Item\_Prize)

FROM paytm\_tb

WHERE Sale\_Flag="ON Sale"

Group by Item\_NM ;



19. Identify products with a "Paid\_pr" higher than the average in their respective "Family" and "Brand" groups.

select Family , Brand , Item\_NM,Paid\_pr

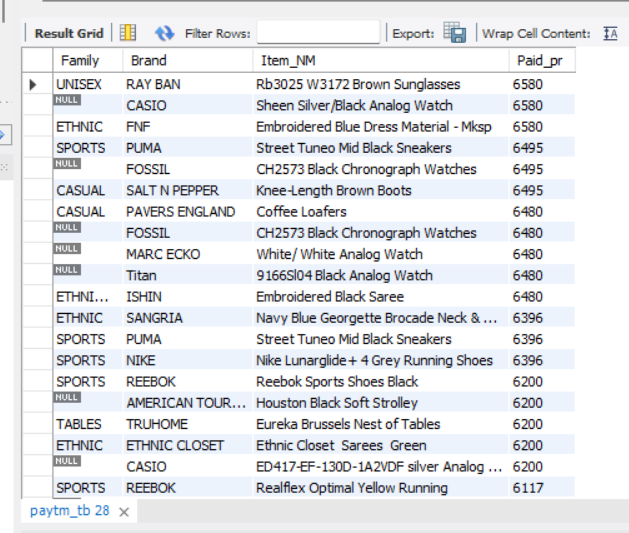
from paytm\_tb

where paid\_pr >( select avg(paid\_pr)

from paytm\_tb)

group by Family,Brand , Item\_NM,Paid\_pr

order by Paid\_pr desc ;



20. Create a pivot table to show the total sales for each "Color" within the "Clothing" category and use conditional formatting to highlight the highest sales.

SELECT DISTINCT Color ,SUM(Item\_Prize) AS Total\_Sales

FROM paytm\_tb

WHERE Category\_Grouped = "Apparels"

GROUP BY Color ;

