

Paytm Mall Epurchase Data Analysis Solutions

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

Solution:

It represents the categorical data of the items bought by the customers.

```
select distinct Category_Grouped from paytmdata;
```

	Category_Grouped
▶	Others
	Apparels
	Shoes
	Home

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

Solution:

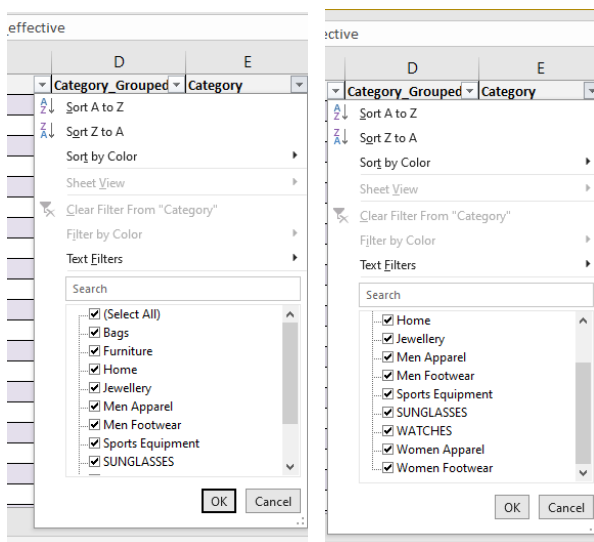
```
select max(result), Shipping_city  
from (SELECT count(*) as result, Shipping_city from  
paytmdata group by Shipping_city) as counts  
group by Shipping_city  
ORDER BY max(result) DESC  
LIMIT 5;
```

	max(result)	Shipping_city
▶	4560	New Delhi
	4254	Chennai
	3974	Bangalore
	3159	Mumbai
	2849	Hyderabad

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

Solution:

There isn't any category as "Electronics" in the given dataset



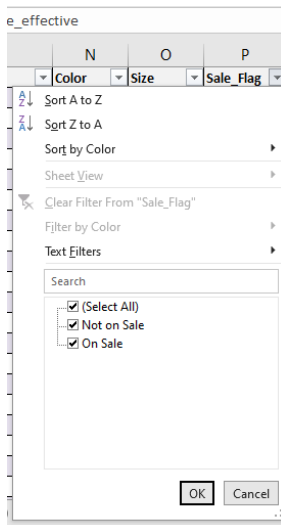
SQL Query If there is a "Electronics" category

```
select item_nm from paytmdata where category = 'Electronics';
```

4. Filter the data to show only rows with a "Sale_Flag" of 'Yes'.

Solution:

"**SALE_FLAG**" column contains only two unique values.



Assuming that **On Sale** value as a 'Yes'.

SELECT * from paytmdata where Sale_Flag = 'On Sale';

Brick	Item_NM	Color	Size	Sale_Flag	Payment_Method	coupon_mone
IGHTS	SKINS Navy Blue Tights	BLACK	XL	On Sale	COD	0
IGHTS	SKINS Navy Blue Tights	BLACK	XL	On Sale	COD	0
IGHTS	SKINS Navy Blue Tights	BLACK	XXXL	On Sale	COD	0
:UNNING SHOES	Adizero F50 2 M Black Running Shoes	BLACK	7	On Sale	Prepaid	1000
NALOG WATCH	Bpb-1004-C Silver/Black Analog Watch	SILVER	FREE SIZE	On Sale	COD	0
IANDBAG	Downing Street 04 Black Handbag	BLACK	STANDARD	On Sale	COD	1099.8
IGHTS	SKINS Navy Blue Tights	BLACK	XL	On Sale	COD	0
:UNNING SHOES	Ventilator Hls Grey Running Shoes	GREY	10	On Sale	Prepaid	0
:UNNING SHOES	Nike Lunarglide+ 4 Grey Running Sho...	GREY	7	On Sale	COD	0
:UNNING SHOES	Adizero F50 2 M Black Running Shoes	BLACK	7	On Sale	Prepaid	1000
:UNNING SHOES	Lunarswift+ 4 Black Running Shoes	BLACK	5	On Sale	COD	0
IGHTS	SKINS Navy Blue Tights	BLACK	XL	On Sale	COD	0
'RICKET	Tridium Ti Full Spike White Cricket Shoes	WHITE	10	On Sale	COD	1000

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

Solution:

```
SELECT item_price, item_NM from paytmdata
order by item_price DESC
LIMIT 1;
```

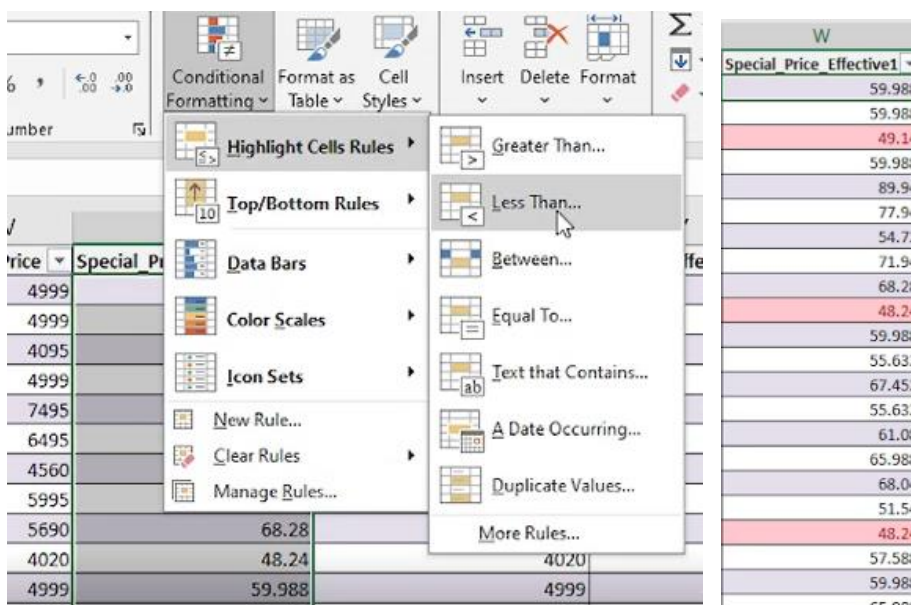
Family	Brand	Brick	itemcost	Item_NM	Color	Size	Sale_Flag	Payment_Method
SPORTS	PUMA	SNEAKERS	13500	Street Tuneo Mid Black Sneakers	BLACK		9 Not on Sale	COD
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	PUMA	SNEAKERS	13500	Street Tuneo Mid Black Sneakers	BLACK		9 Not on Sale	COD
SPORTS	ADIDAS	RUNNING SHOES	13500	Element Soul M Black Running Shoes	BLACK		9 Not on Sale	Prepaid
SPORTS	PUMA	SNEAKERS	13500	Street Tuneo Mid Black Sneakers	BLACK		9 Not on Sale	COD

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

Solution:

Select the entire Special_Price_effective column

Home tab → Conditional Formatting → Highlight Cells Rules → Less Than → Write \$50



Special_Price_Effective1
59.988
59.988
49.14
59.988
89.94
77.94
54.72
71.94
68.28
48.24
59.988
55.632
67.452
55.632
61.08
65.988
68.04
51.54
48.24
57.588
59.988
65.988

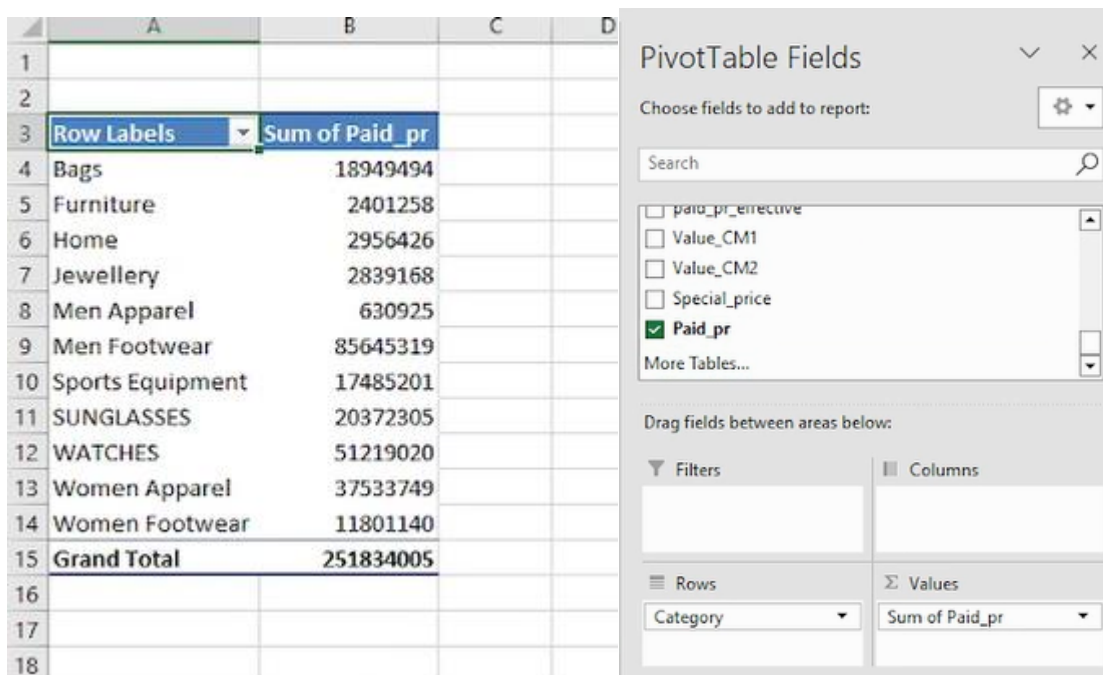
7. Create a pivot table to find the total sales value for each category.

Solution:

Go to INSERT tab → click on pivot table → From Table/Range →

Click on OK

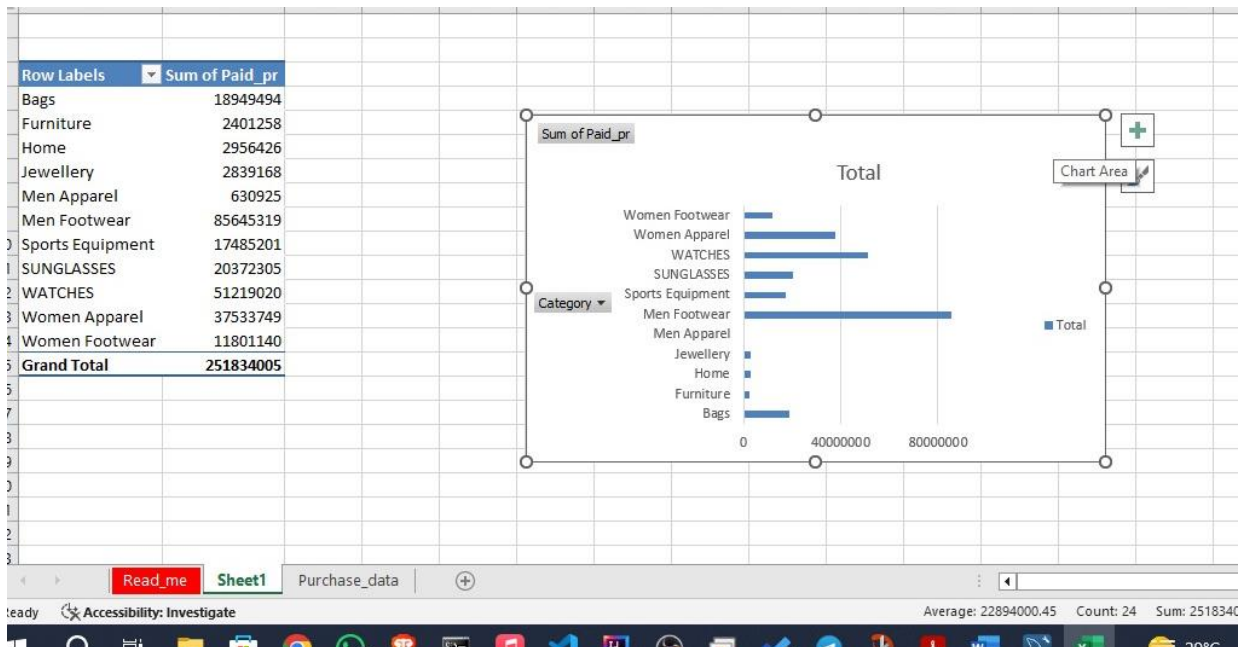
A pivot table will be create and you have to add the desired columns in fields



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

Solution:

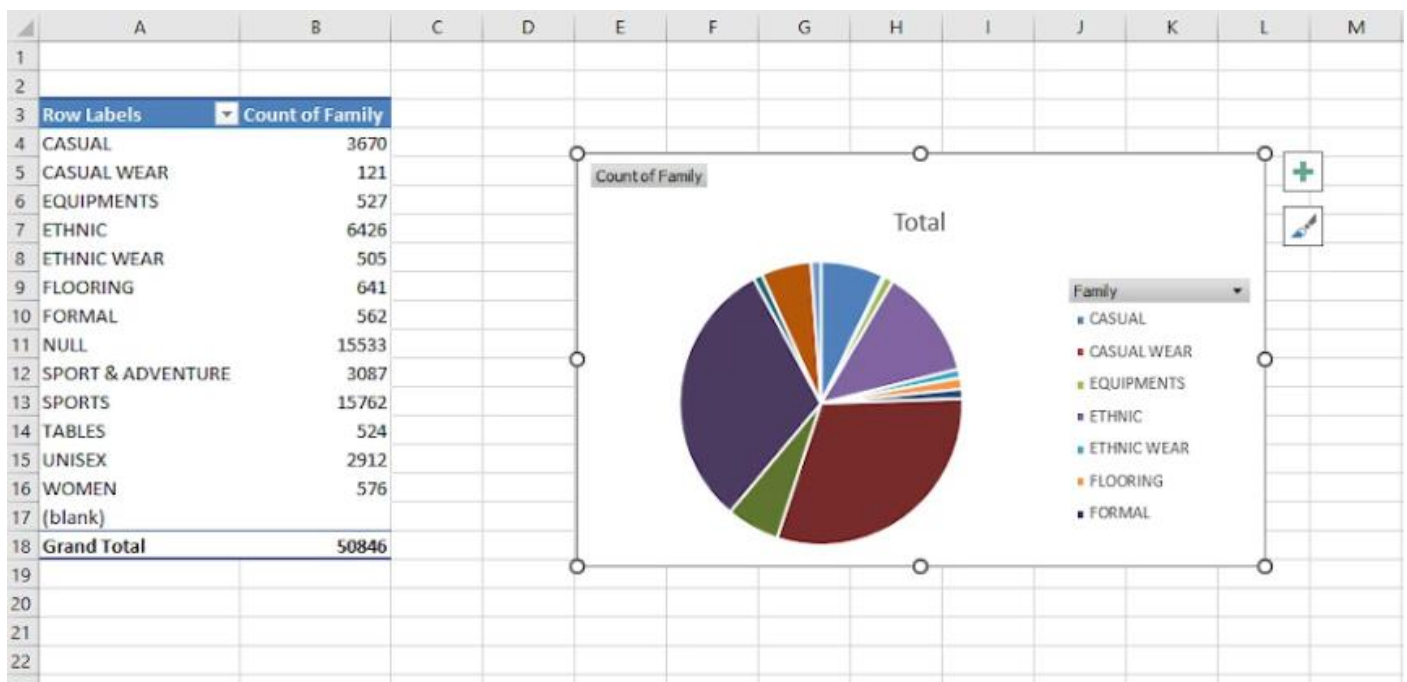
Go to INSERT tab → Recommended Charts → Bar



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

Solution:

First we have to create a pivot table and select that table then we have to click on the pie chart present in INSERT tab.



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

Solution:

```
=NOT (ISERROR (MATCH (Q2, {"Visa", "MasterCard"}, 0)))
```

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

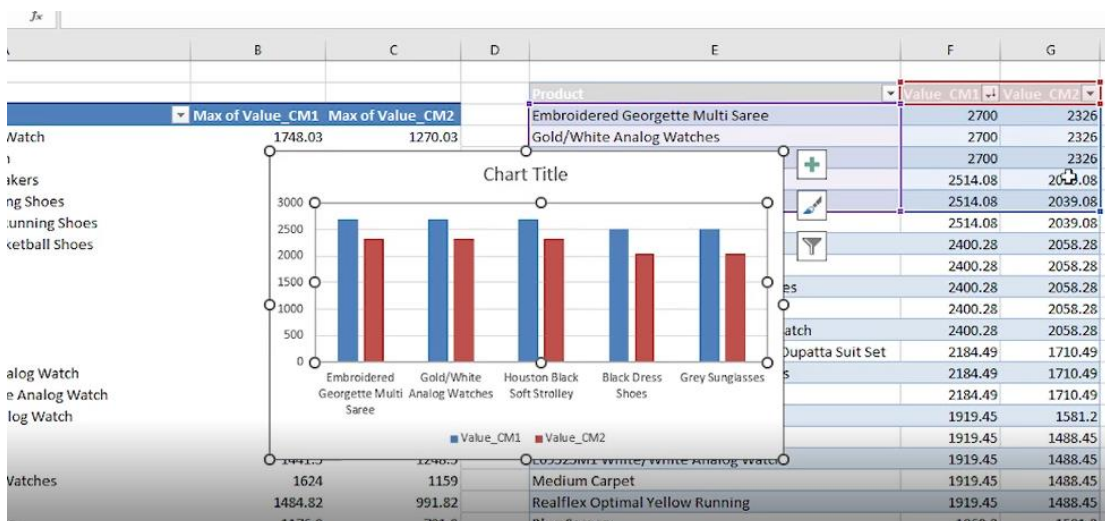
Solution:

```
select Product_Gender, AVG(Quantity) from paytmdata
where Category In ('Men Footwear', 'SUNGLASSES', 'Women Footwear')
group by Product_Gender;
```

	Product_Gender	AVG(Quantity)
	UNISEX	1.0000
	MEN	1.0003
	WOMEN	1.0000

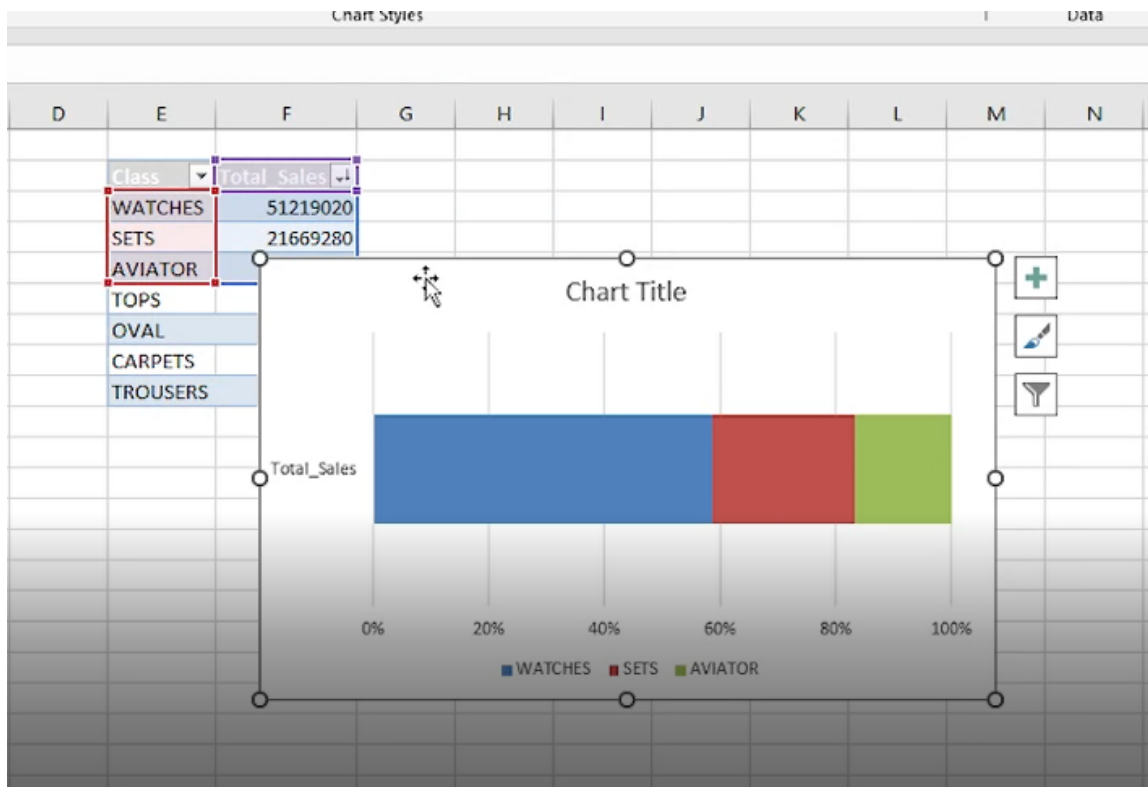
12. Find the top 5 products with the highest "Value_CM1" and "Value_CM2" ratios. Create a chart to visualize this data.

Solution:



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

Solution:



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

Solution:

```
=VLOOKUP([@Color], Table1[[#All],[item_nm]:[Color]], 2, 1)
```

15. Calculate the total "coupon_money_effective" and "Coupon_Percentage" for products in the "Electronics" category.

Solution:

SQL Query If there is a "Electronics" category

```
select item_nm, SUM(coupon_money_effective),
SUM(Coupon_Percentage)
from paytmdata
```


where Category = 'Electronics';

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

Solution:

There isn't any month column.

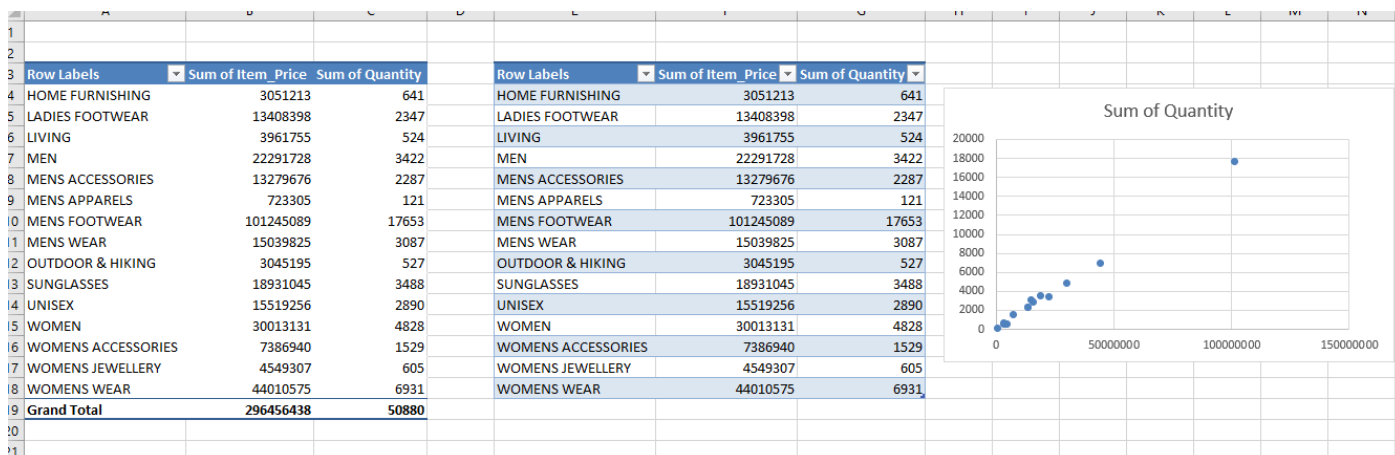
If there is a month column then as follows:

=MONTH(A2)

=SUMIF(\$C\$2:\$C\$1000, E2, \$B\$2:\$B\$1000)

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.

Solution:



18. Use the AVERAGEIFS function to find the average "Item_Price" for products that have a "Sale_Flag" of 'Yes.'

Solution:

=AVERAGEIFS([Item_Price], [Sale_Flag], "Yes")

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

Solution:

```
=AVERAGEIFS($AD$2:$AD$1000,$J$2:$J$1000,J3,$K$2:$K$1000,K3)
```

```
=IF(AD1>L1, "Above Avg Family", "Below Avg Family")
```

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.

Solution:

Row Labels	Sum of Paid_pr
Men Footwear	85645319
BLACK	32936218
BLUE	5927983
BROWN	2538884
COFFEE	4288977
GREY	20042856
RED	3229678
TAN	4263143
WHITE	12417580
SUNGLASSES	20372305
BROWN	5572768
GREEN	8095953
GREY	6703584
Women Footwear	11801140
BLACK	3475650
BROWN	756450
PINK	3203170
YELLOW	4365870