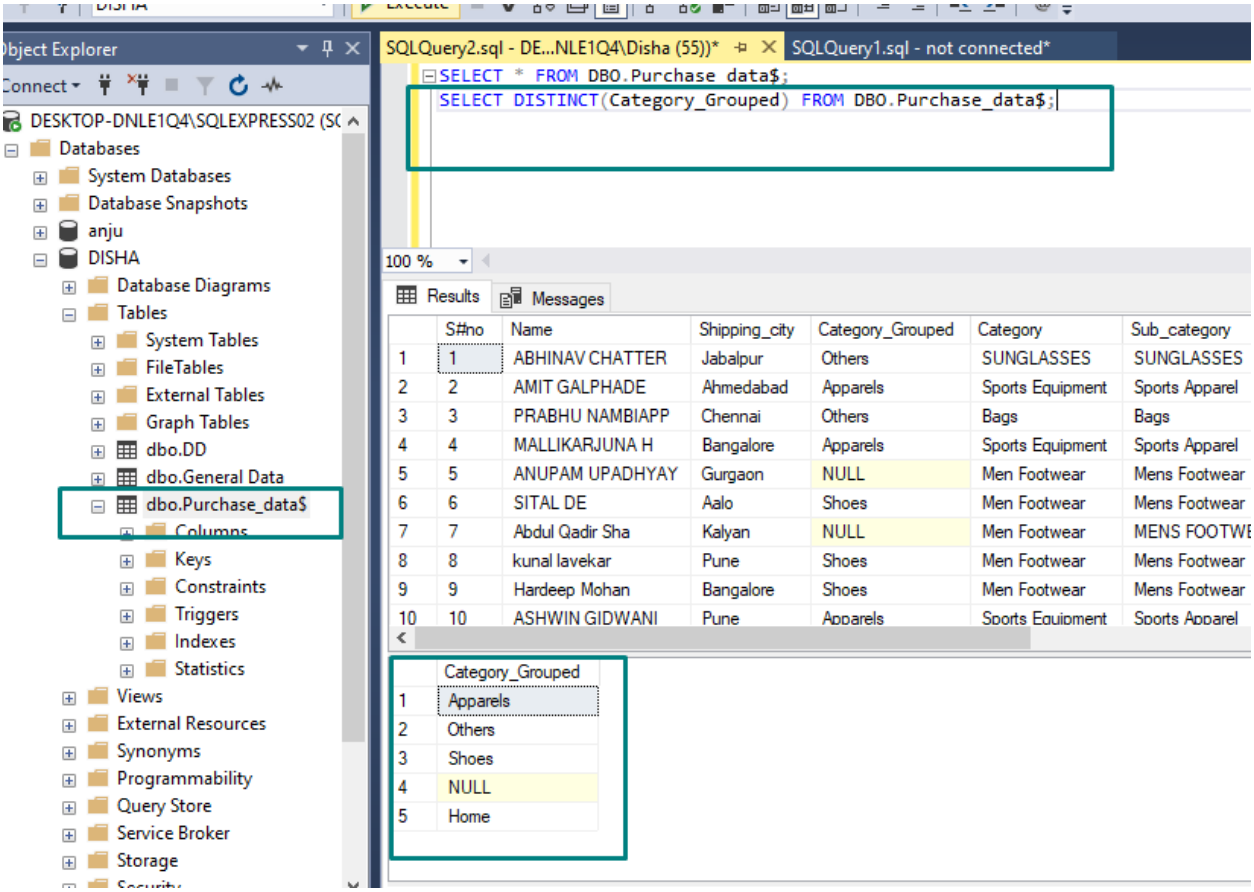
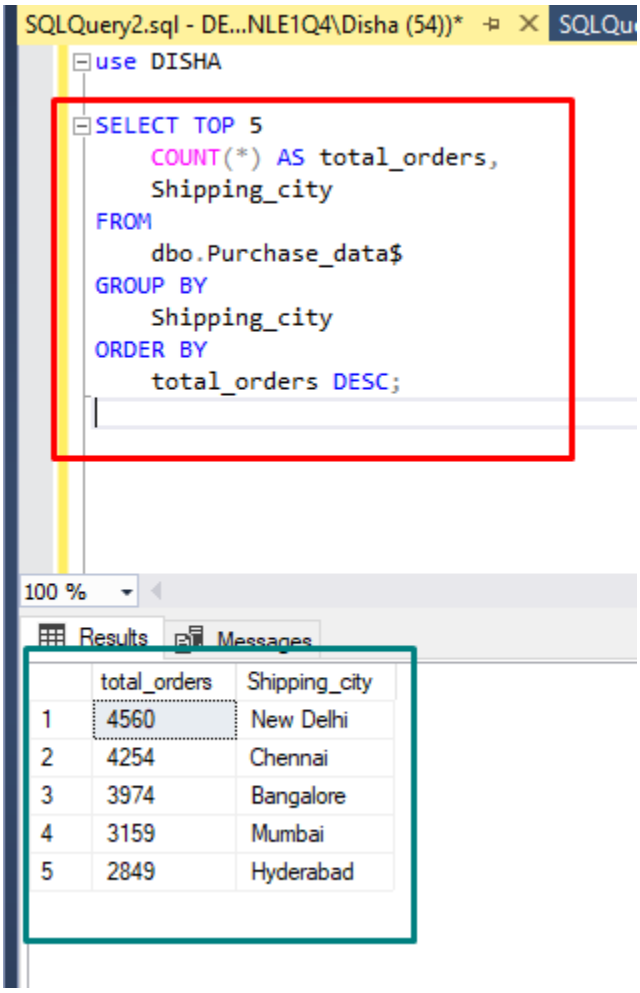
PAYTM EPURCHASE DATA

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



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

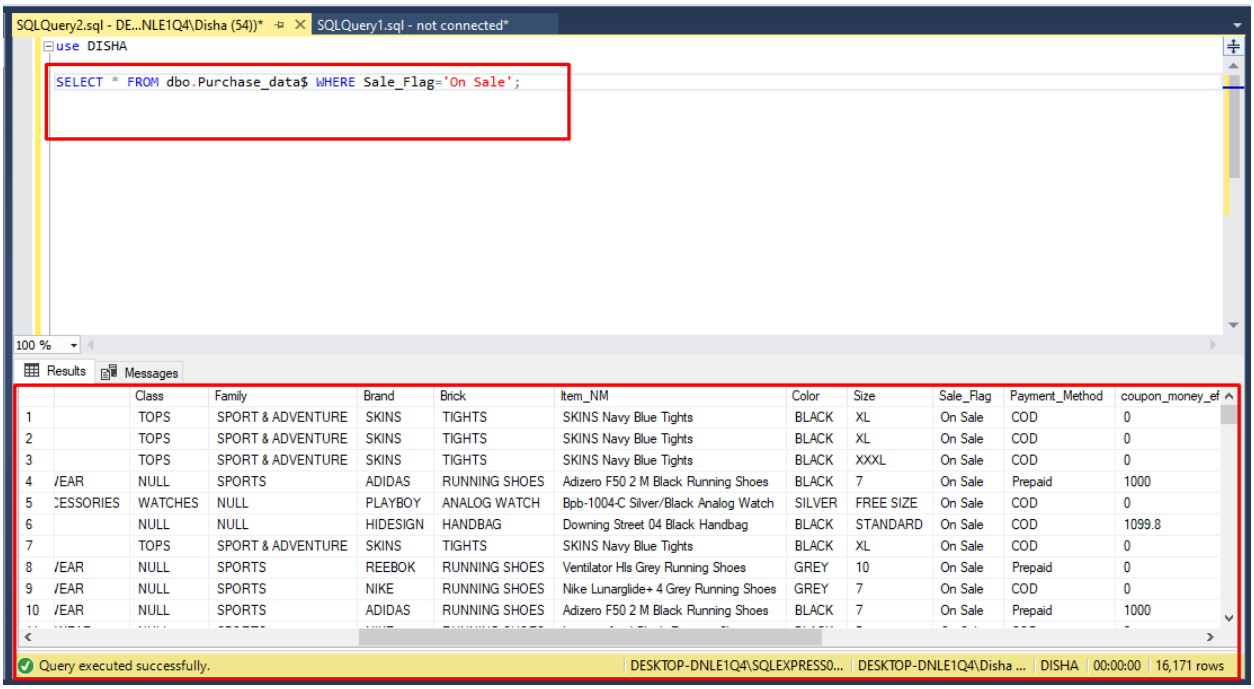


	total_orders	Shipping_city
1	4560	New Delhi
2	4254	Chennai
3	3974	Bangalore
4	3159	Mumbai
5	2849	Hyderabad

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

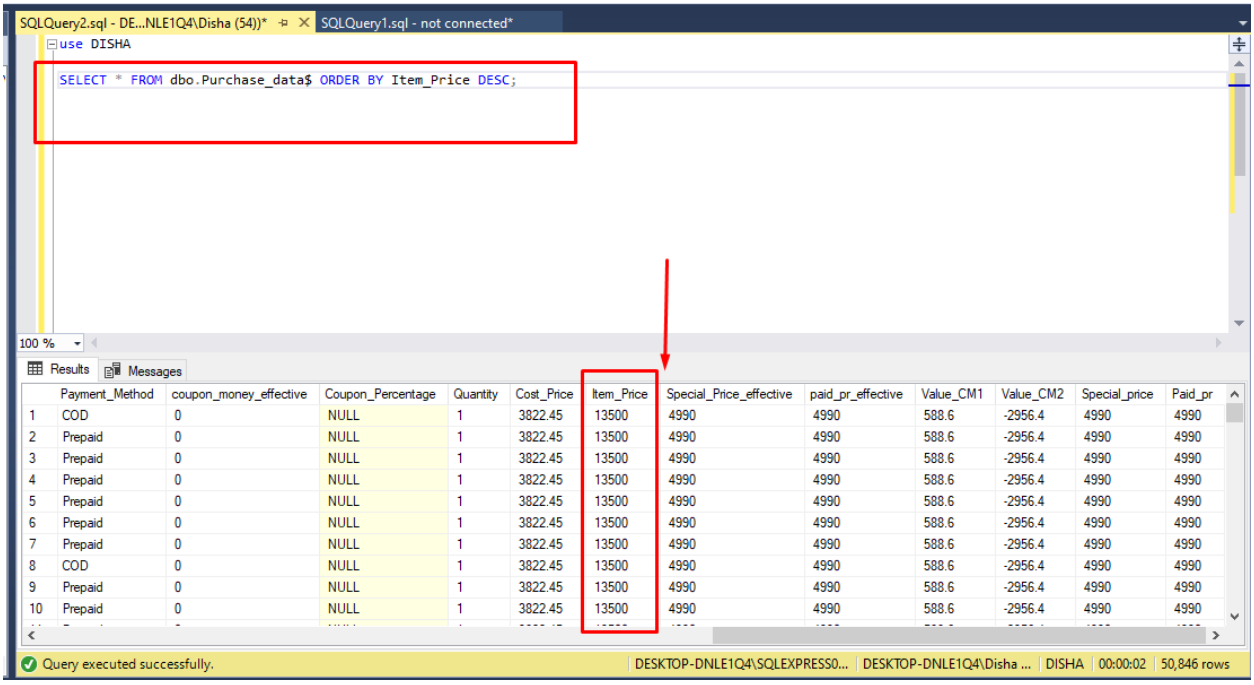
There’s no such category.

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

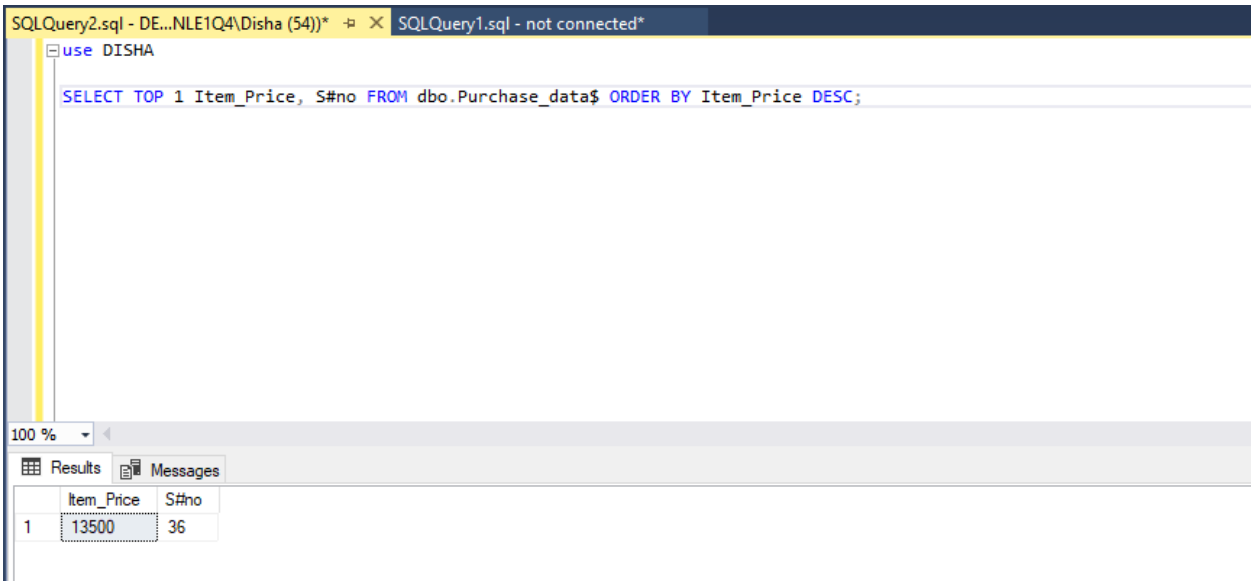


		Class	Family	Brand	Brick	Item_NM	Color	Size	Sale_Flag	Payment_Method	coupon_money_ef
1		TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKINS Navy Blue Tights	BLACK	XL	On Sale	COD	0
2		TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKINS Navy Blue Tights	BLACK	XL	On Sale	COD	0
3		TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKINS Navy Blue Tights	BLACK	XXXL	On Sale	COD	0
4	/EAR	NULL	SPORTS	ADIDAS	RUNNING SHOES	Adizero F50 2 M Black Running Shoes	BLACK	7	On Sale	Prepaid	1000
5	CESSORIES	WATCHES	NULL	PLAYBOY	ANALOG WATCH	Bpb-1004-C Silver/Black Analog Watch	SILVER	FREE SIZE	On Sale	COD	0
6		NULL	NULL	HIDESIGN	HANDBAG	Downing Street 04 Black Handbag	BLACK	STANDARD	On Sale	COD	1099.8
7		TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKINS Navy Blue Tights	BLACK	XL	On Sale	COD	0
8	/EAR	NULL	SPORTS	REEBOK	RUNNING SHOES	Ventilator Hls Grey Running Shoes	GREY	10	On Sale	Prepaid	0
9	/EAR	NULL	SPORTS	NIKE	RUNNING SHOES	Nike Lunarglide+ 4 Grey Running Shoes	GREY	7	On Sale	COD	0
10	/EAR	NULL	SPORTS	ADIDAS	RUNNING SHOES	Adizero F50 2 M Black Running Shoes	BLACK	7	On Sale	Prepaid	1000

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

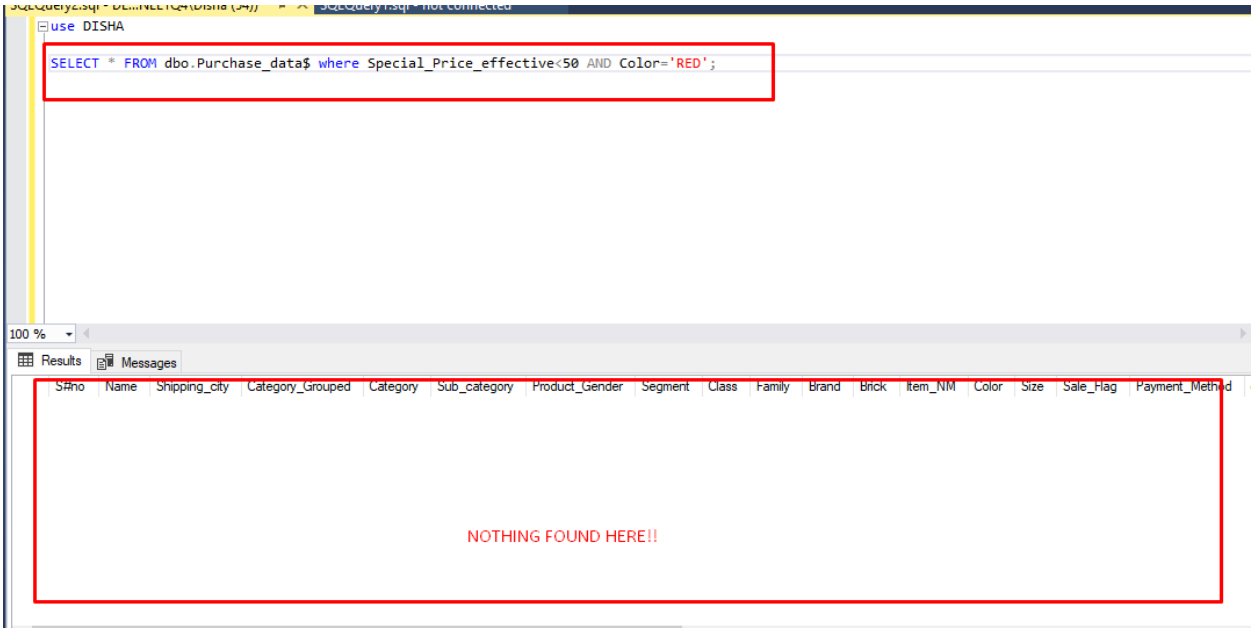


	Payment_Method	coupon_money_effective	Coupon_Percentage	Quantity	Cost_Price	Item_Price	Special_Price_effective	paid_pr_effective	Value_CM1	Value_CM2	Special_price	Paid_pr
1	COD	0	NULL	1	3822.45	13500	4990	4990	588.6	-2956.4	4990	4990
2	Prepaid	0	NULL	1	3822.45	13500	4990	4990	588.6	-2956.4	4990	4990
3	Prepaid	0	NULL	1	3822.45	13500	4990	4990	588.6	-2956.4	4990	4990
4	Prepaid	0	NULL	1	3822.45	13500	4990	4990	588.6	-2956.4	4990	4990
5	Prepaid	0	NULL	1	3822.45	13500	4990	4990	588.6	-2956.4	4990	4990
6	Prepaid	0	NULL	1	3822.45	13500	4990	4990	588.6	-2956.4	4990	4990
7	Prepaid	0	NULL	1	3822.45	13500	4990	4990	588.6	-2956.4	4990	4990
8	COD	0	NULL	1	3822.45	13500	4990	4990	588.6	-2956.4	4990	4990
9	Prepaid	0	NULL	1	3822.45	13500	4990	4990	588.6	-2956.4	4990	4990
10	Prepaid	0	NULL	1	3822.45	13500	4990	4990	588.6	-2956.4	4990	4990

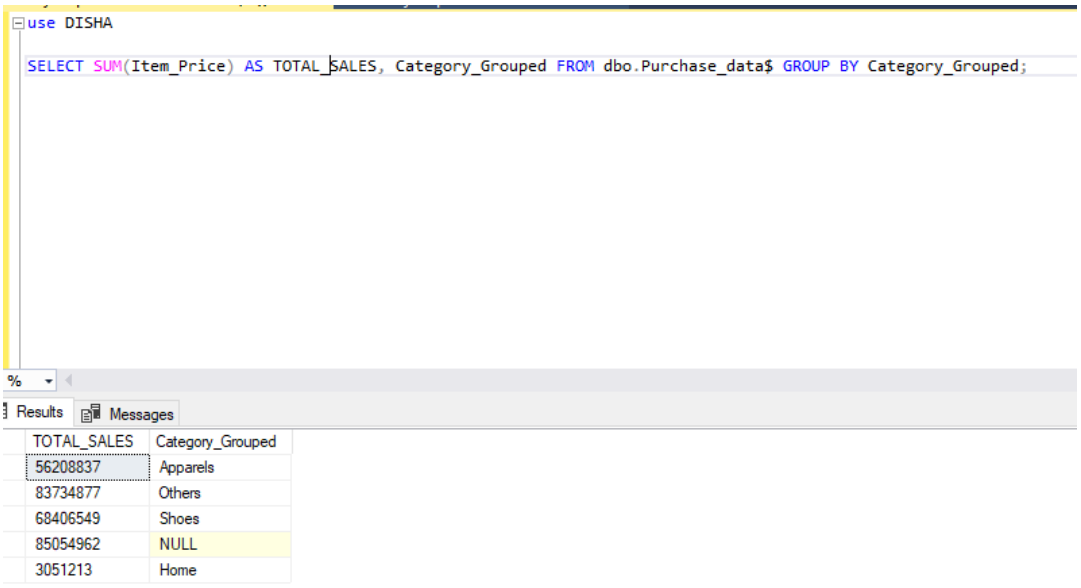


	Item_Price	S#no
1	13500	36

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.



```
use DISHA

SELECT SUM(Item_Price) AS TOTAL_SALES, Category_Grouped FROM dbo.Purchase_data$ GROUP BY Category_Grouped;
```

%

Results | Messages

TOTAL_SALES	Category_Grouped
56208837	Apparels
83734877	Others
68406549	Shoes
85054962	NULL
3051213	Home

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

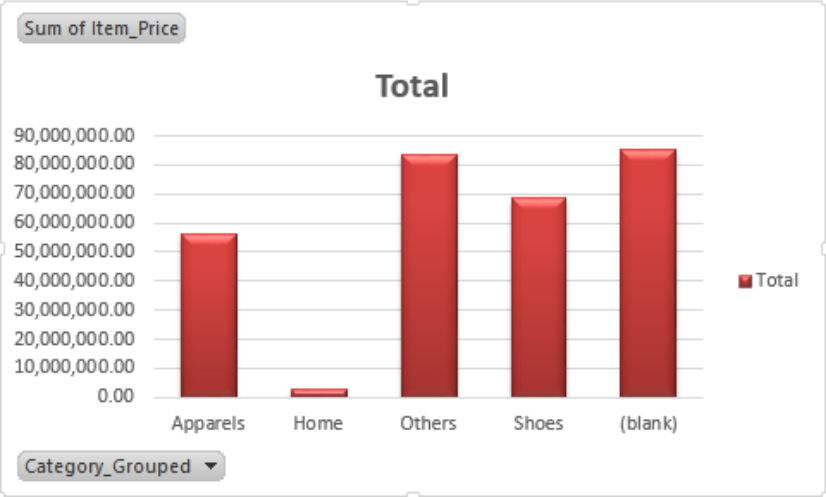
SQLQuery2.sql - DE...NLE1Q4\Disha (54))* | SQLQuery1.sql - not connected*

```
use DISHA

SELECT SUM(Item_Price) AS TOTAL_SALES,
Category_Grouped
FROM dbo.Purchase_data$
GROUP BY Category_Grouped
ORDER BY TOTAL_SALES DESC;
```

100 %

Results | Messages

	TOTAL_SALES	Category_Grouped
1	85054962	NULL
2	83734877	Others
3	68406549	Shoes
4	56208837	Apparels
5	3051213	Home



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

```
use DISHA

SELECT Family, Count(*) AS TOTAL_PRODUCTS
FROM dbo.Purchase_data$
GROUP BY Family;
```

100 %

Results | Messages

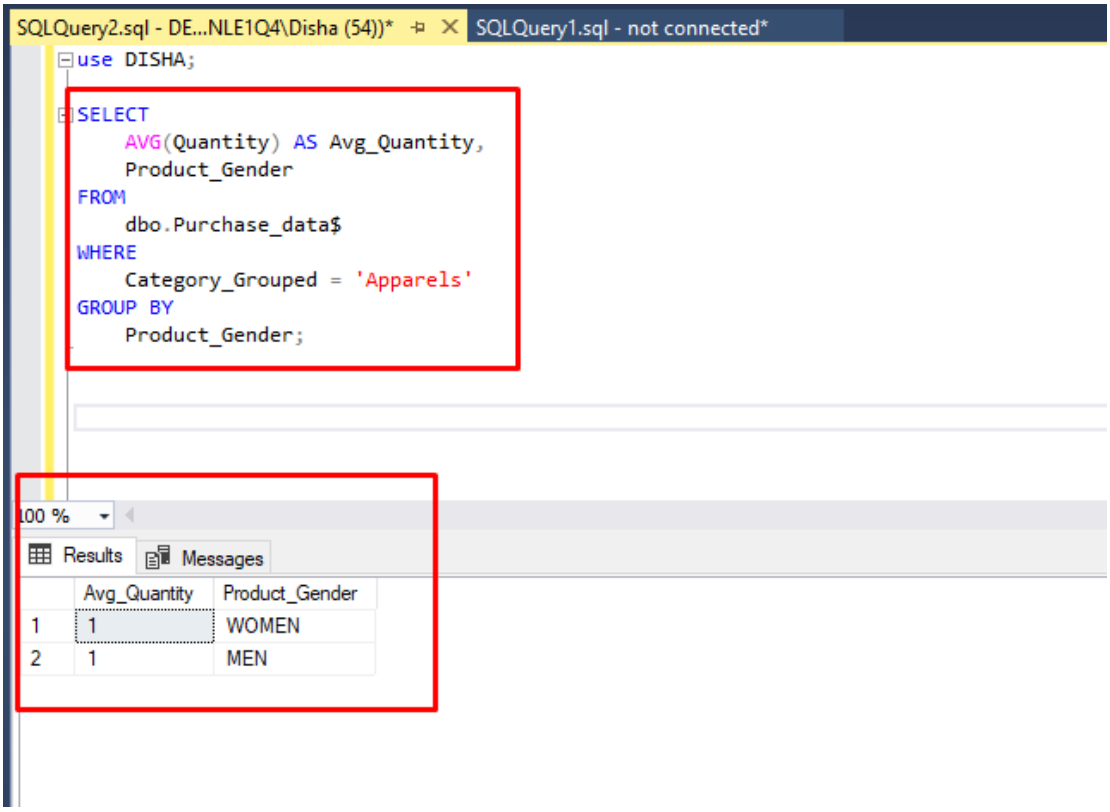
	Family	TOTAL_PRODUCTS
1	UNISEX	2912
2	NULL	15533
3	FLOORING	641
4	WOMEN	576
5	EQUIPMENTS	527
6	CASUAL WEAR	121
7	CASUAL	3670
8	ETHNIC	6426
9	ETHNIC WEAR	505
10	SPORT & ADVENTURE	3087
11	TABLES	524

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

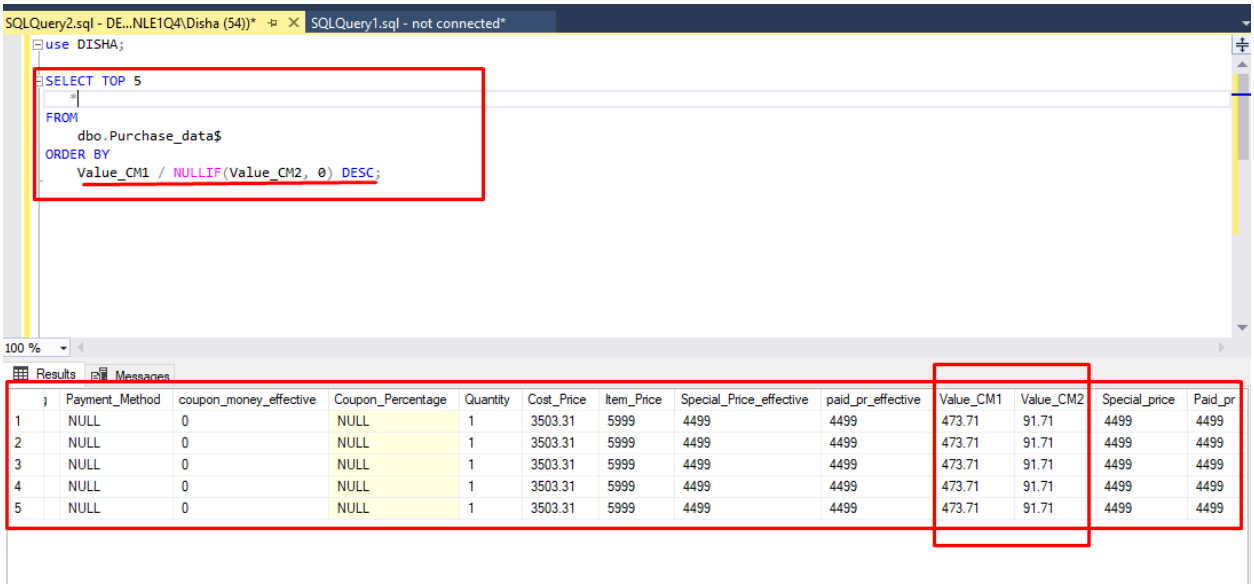
Visa, MasterCard).

```
use DISHA;

UPDATE dbo.Purchase_data$
SET Payment_Method =
    CASE
        WHEN Payment_Method NOT IN ('Visa', 'MasterCard') THEN 'Null'
        ELSE Payment_Method
    END;
```

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



	Payment_Method	coupon_money_effective	Coupon_Percentage	Quantity	Cost_Price	Item_Price	Special_Price_effective	paid_pr_effective	Value_CM1	Value_CM2	Special_price	Paid_pr
1	NULL	0	NULL	1	3503.31	5999	4499	4499	473.71	91.71	4499	4499
2	NULL	0	NULL	1	3503.31	5999	4499	4499	473.71	91.71	4499	4499
3	NULL	0	NULL	1	3503.31	5999	4499	4499	473.71	91.71	4499	4499
4	NULL	0	NULL	1	3503.31	5999	4499	4499	473.71	91.71	4499	4499
5	NULL	0	NULL	1	3503.31	5999	4499	4499	473.71	91.71	4499	4499

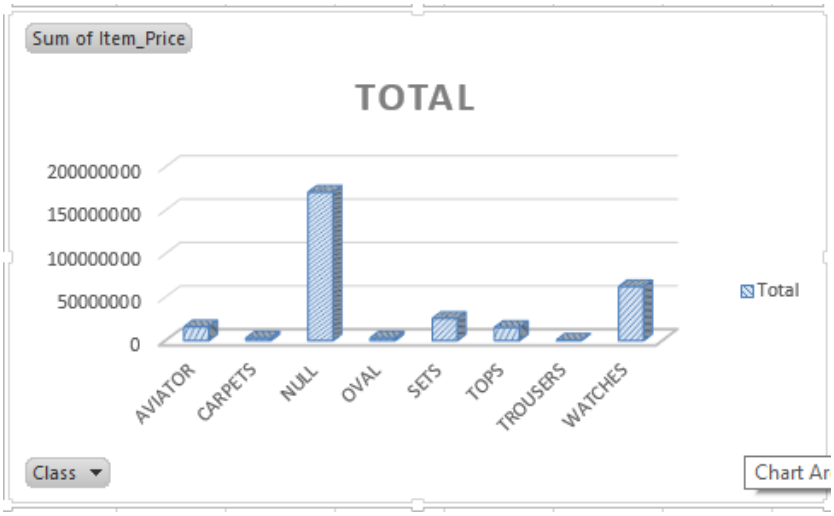
13. Identify the top 3 "Class" categories with the highest total sales. Create a stacked bar chart to represent this data.

```
use DISHA;

SELECT TOP 3
    Class, SUM(Item_Price) AS TotalSales
FROM
    dbo.Purchase_data$
GROUP BY
    Class
ORDER BY
    TotalSales DESC;
```

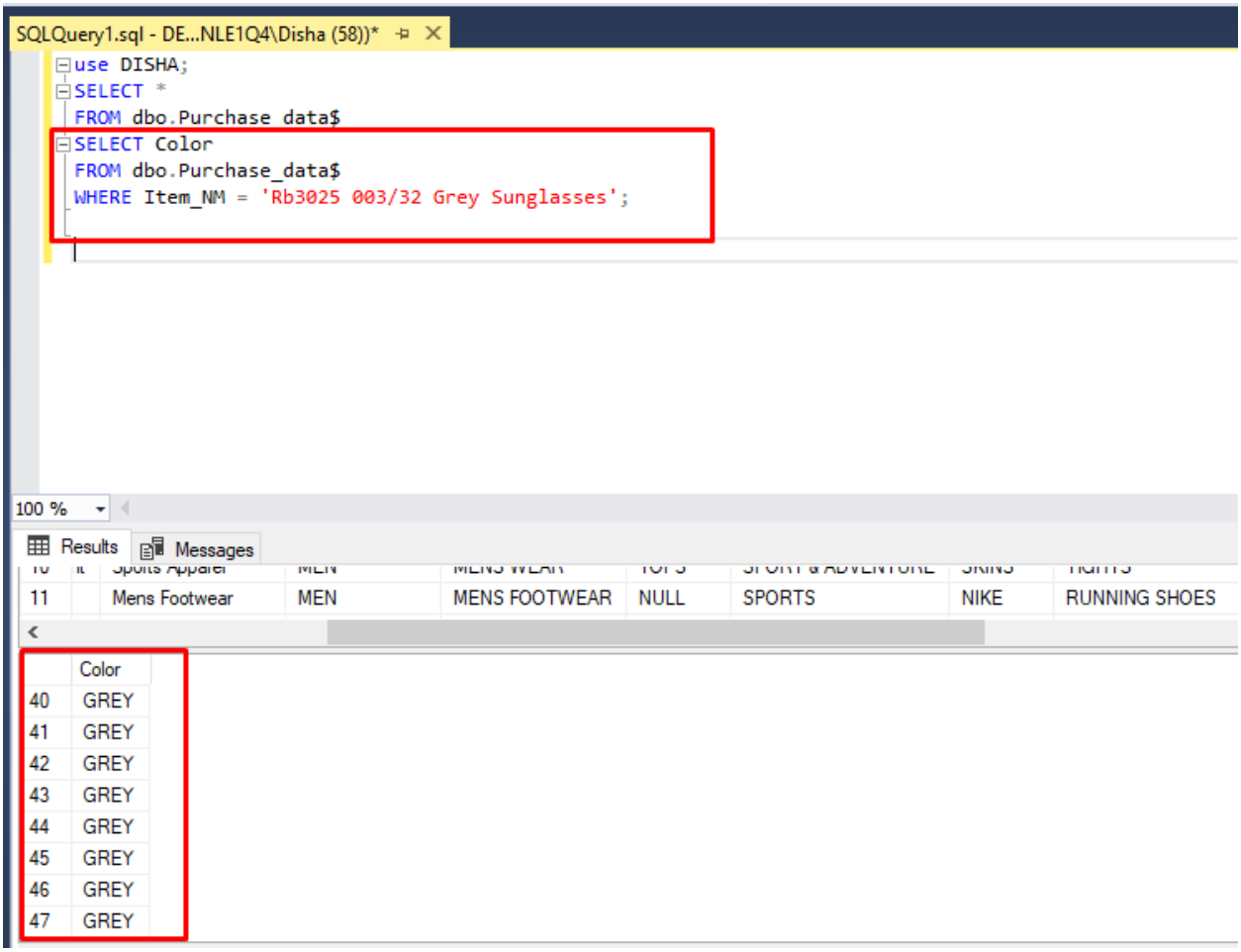
100 %

Results | Messages

	Class	TotalSales
1	NULL	170968090
2	WATCHES	62213793
3	SETS	25529167



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

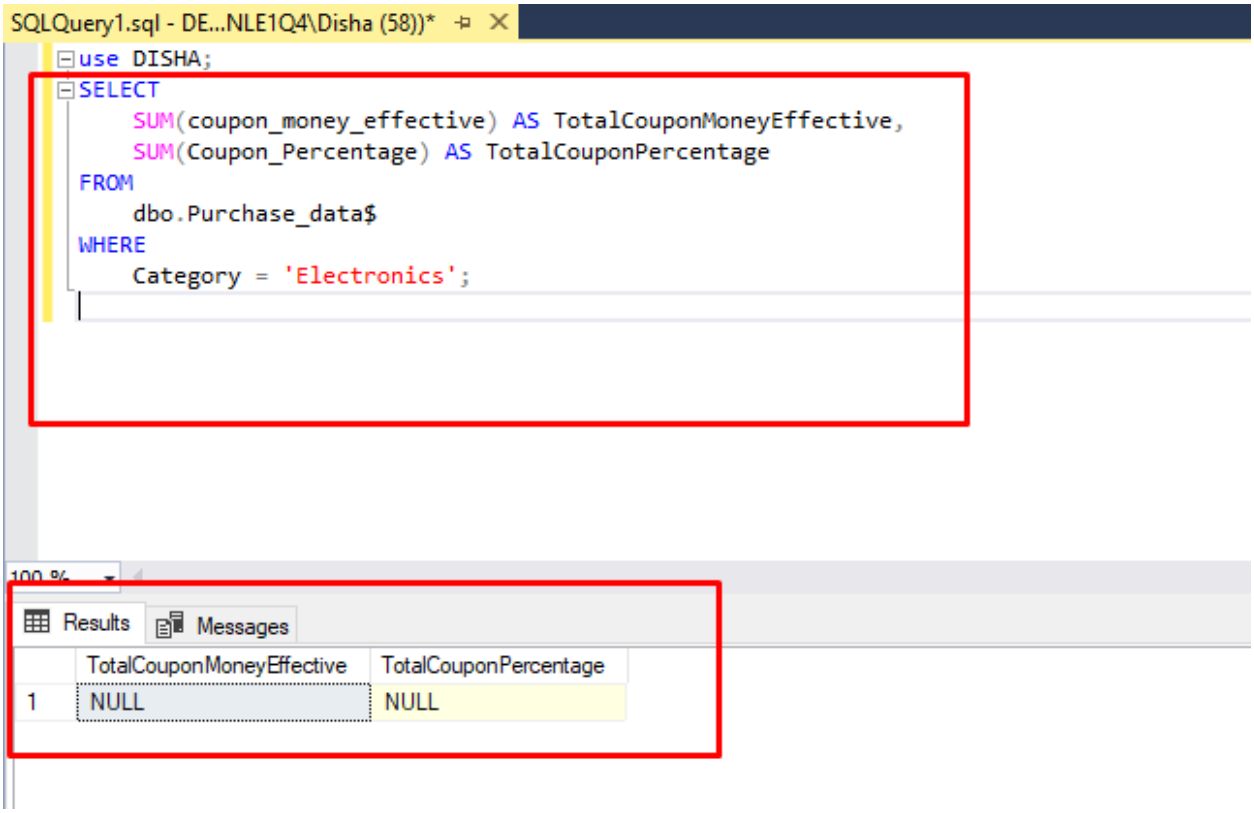


=INDEX(C1:C100, MATCH("specific_item_nm", A1:A100, 0))

0 in MATCH ensures an exact match

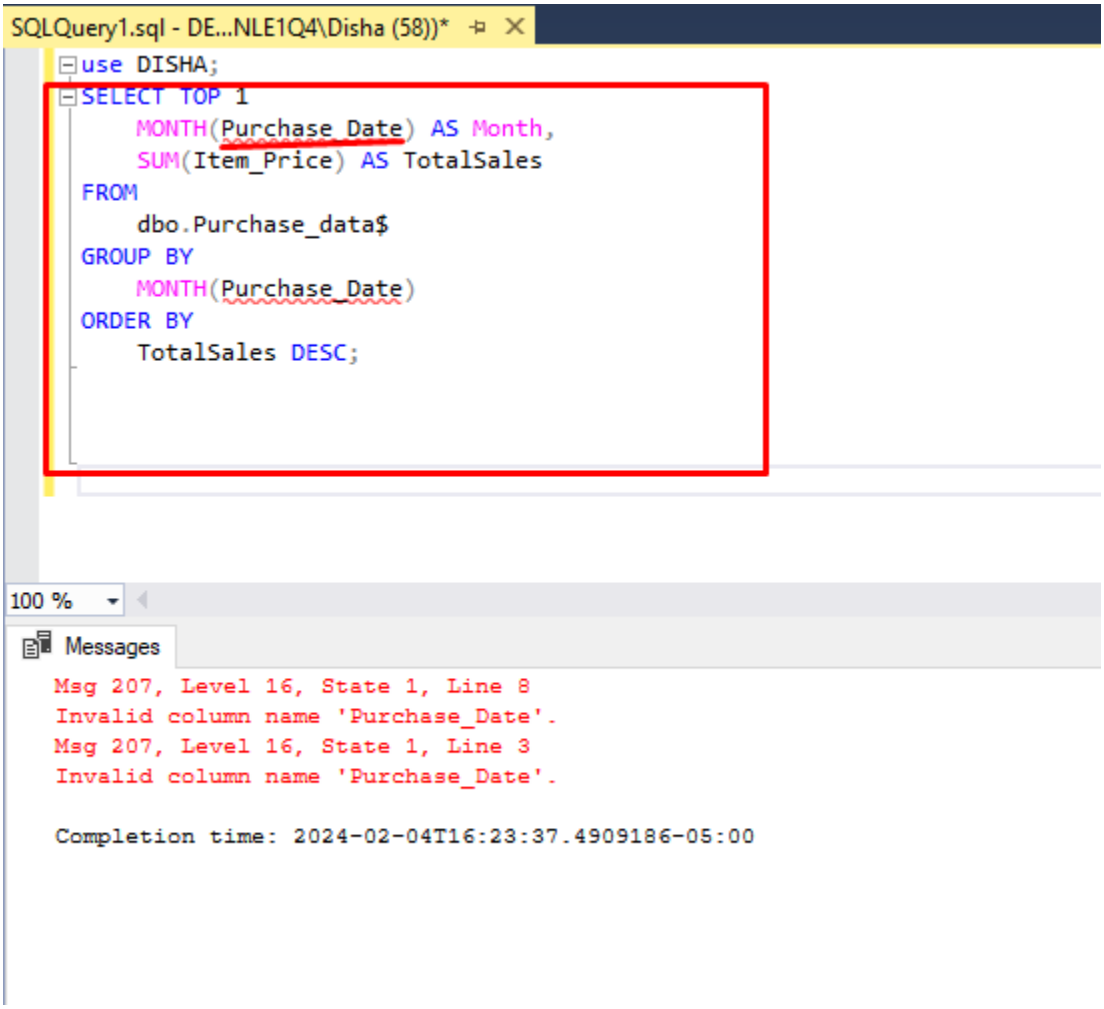
A1:A100 is the range of "Item_NM," and C1:C100 is the range of "Color.

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



2 of 2

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



Date hasn’t been provided to us.

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

SQLQuery2.sql - DE...NLE1Q4\Disha (54))* | SQLQuery1.sql - not connected*

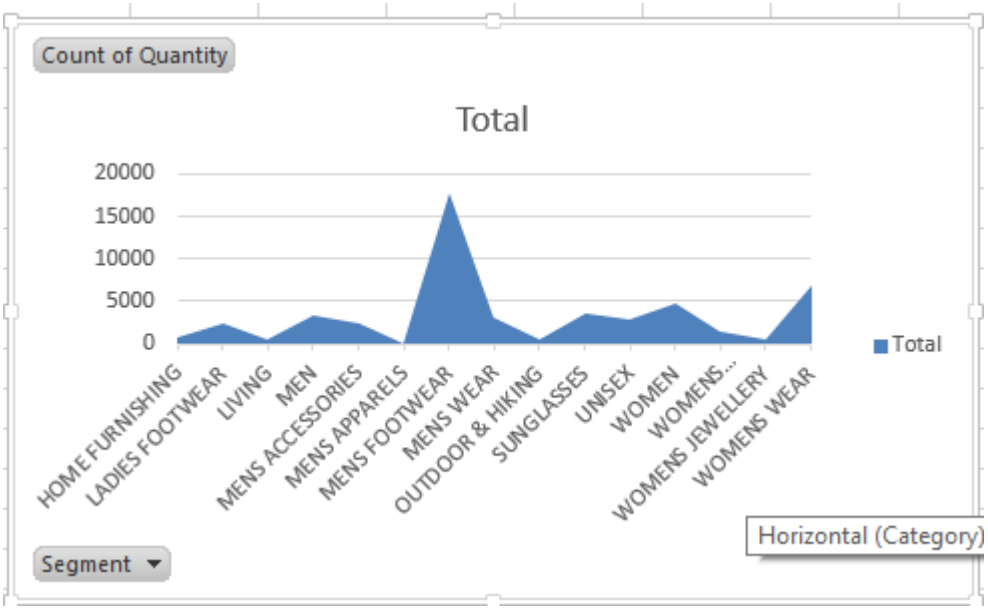
```
use DISHA;

SELECT
    Segment,
    SUM(Item_Price) AS TotalSales
FROM
    dbo.Purchase_data$
GROUP BY
    Segment;
```

100 %

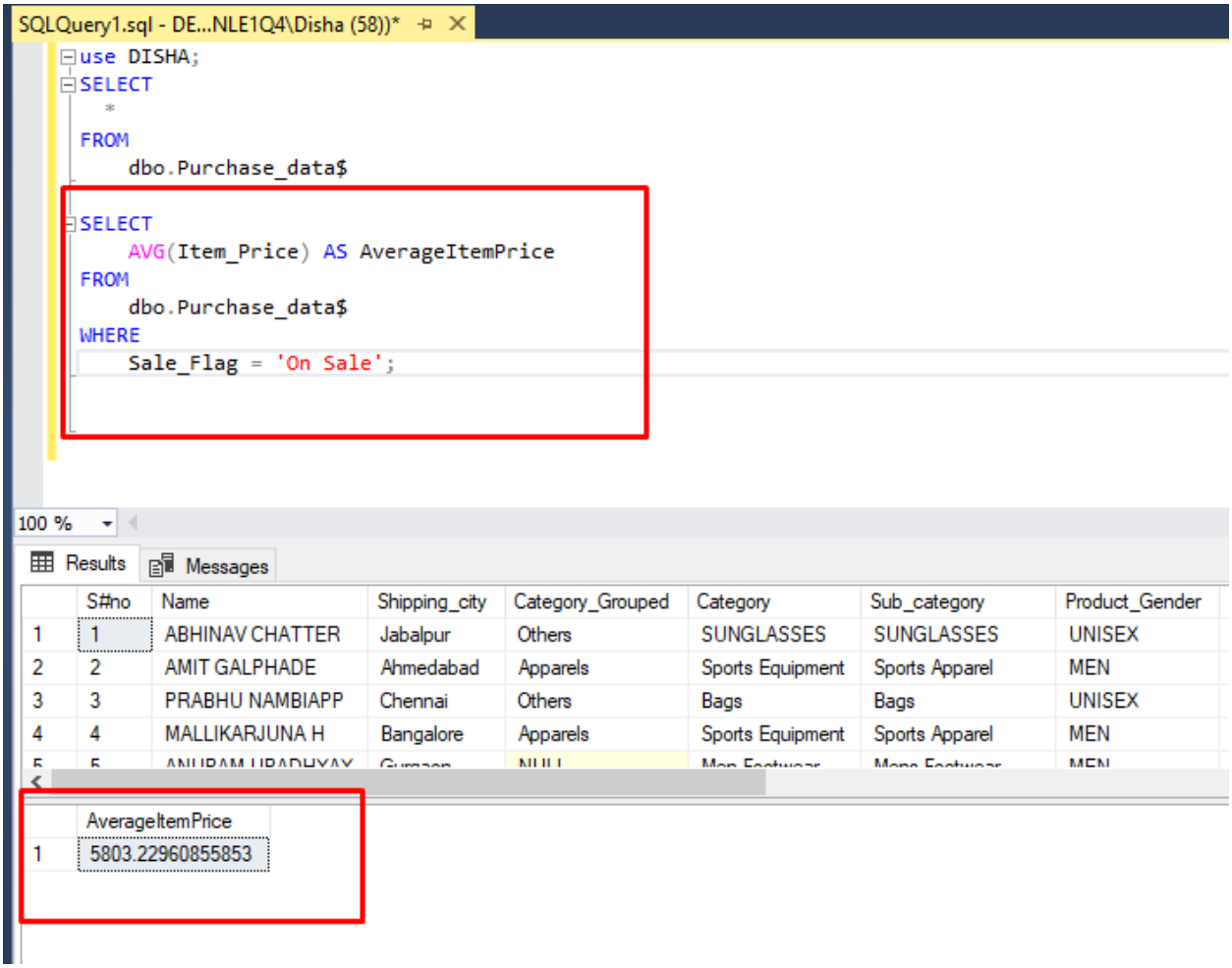
Results | Messages

	Segment	TotalSales
1	MENS ACCESSORIES	13279676
2	OUTDOOR & HIKING	3045195
3	MENS WEAR	15039825
4	WOMENS ACCESSORIES	7386940
5	UNISEX	15519256
6	LADIES FOOTWEAR	13408398
7	WOMEN	30013131
8	WOMENS JEWELLERY	4549307
9	MENS APPARELS	723305
10	HOME FURNISHING	3051213
11	LIVING	3961755

Query executed successfully. DESKTOP-DNLE1Q4\SQLEXPRESS0...



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



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

```
use DISHA;

SELECT
    S#no,
    Family,
    Brand,
    Paid_pr
FROM
    dbo.Purchase_data$ AS pd
WHERE
    Paid_pr > (
        SELECT AVG(Paid_pr)
        FROM dbo.Purchase_data$ AS sub_pd
        WHERE sub_pd.Family = pd.Family AND sub_pd.Brand = pd.Brand
    );
```

100 %

Results | Messages

	S#no	Family	Brand	Paid_pr
1	2	SPORT & ADVENTURE	SKINS	4999
2	4	SPORT & ADVENTURE	SKINS	4999
3	5	SPORTS	NIKE	5621
4	7	CASUAL	CLARKS	4560
5	8	SPORTS	NIKE	4995
6	13	CASUAL	WOODLAND	5621
7	17	SPORTS	REEBOK	5670
8	20	NULL	PLAYBOY	4799
9	26	SPORTS	NIKE	5299
10	30	ETHNIC	SANGRIA	6396
11	31	NULL	VIP	5983

Query executed successfully. DESKTOP-DNLE1Q4\SQLEXPRESS0...

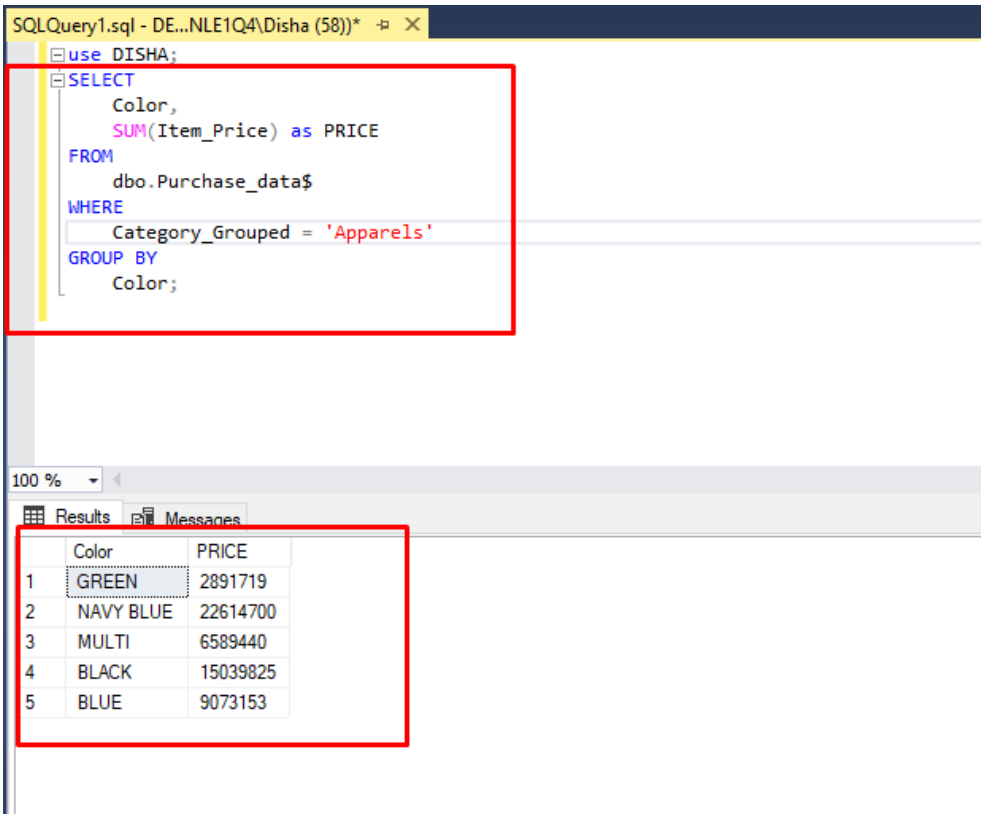
20. Create a pivot table to show the total sales for each "Color" within the "Clothing" category and use conditional formatting to highlight the highest sales.

There's no Clothing category but there is "Apparels" category.

Men Apparel	723305
BEIGE	723305

Women Apparel	44010575
BLACK	2841563
BLUE	9073153
GREEN	2891719
MULTI	6589440
NAVY BLUE	22614700

SQL



Item_Price	Special_Price_effective
4999	4999
4999	4999
4095	4095
4999	4999
7495	7495
6495	6495
4560	4560
5995	5995
5690	5690
4020	4020
4999	4999
5795	4636
7495	5621
5795	4636
5090	5090
5499	5499
6300	5670
4295	4295
4020	4020
4799	4799
4999	4999
5499	5499
4925	4186
4990	4990
5499	5499
5299	5299
5499	5499
4195	4195
4550	4095
7995	6396
11966	5983
4925	4186