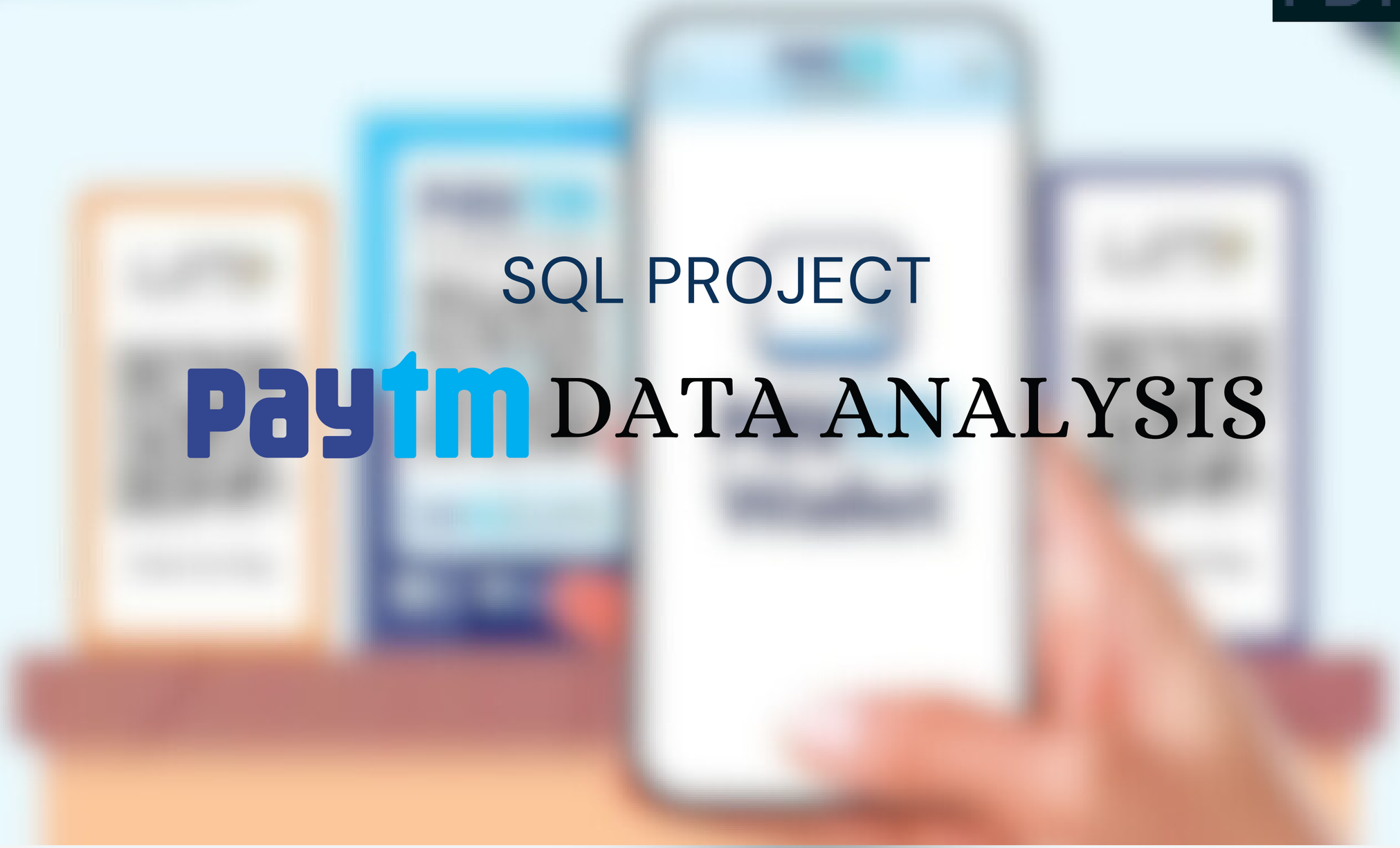
Presented by : Boloram Mitra



SQL PROJECT

Paytm DATA ANALYSIS

ABOUT DATASET

Table Name : paytm_data

Columns:	
S.no	int
Name	text
Shipping_city	text
Category_Grouped	text
Category	text
Sub_category	text
Product_Gender	text
Segment	text
Class	text
Family	text
Brand	text
Brick	text
Item_NM	text
Color	text
Size	varchar(255)

Sale_Flag	text
Payment_Method	text
coupon_money_effective	double
Coupon_Percentage	double
Quantity	double
Cost_Price	double
Item_Price	double
Special_Price_effective	double
paid_pr_effective	double
Value_CM1	double
Value_CM2	double
special_price	double
Paid_pr	double

Our database has 1 table and 28 columns as shown above

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

```
SELECT DISTINCT Category_Grouped AS Unique_Categories
FROM paytm_data;
```

OUTPUT :

	Unique_Categories
▶	Others
	Apparels
	shoes
	Home

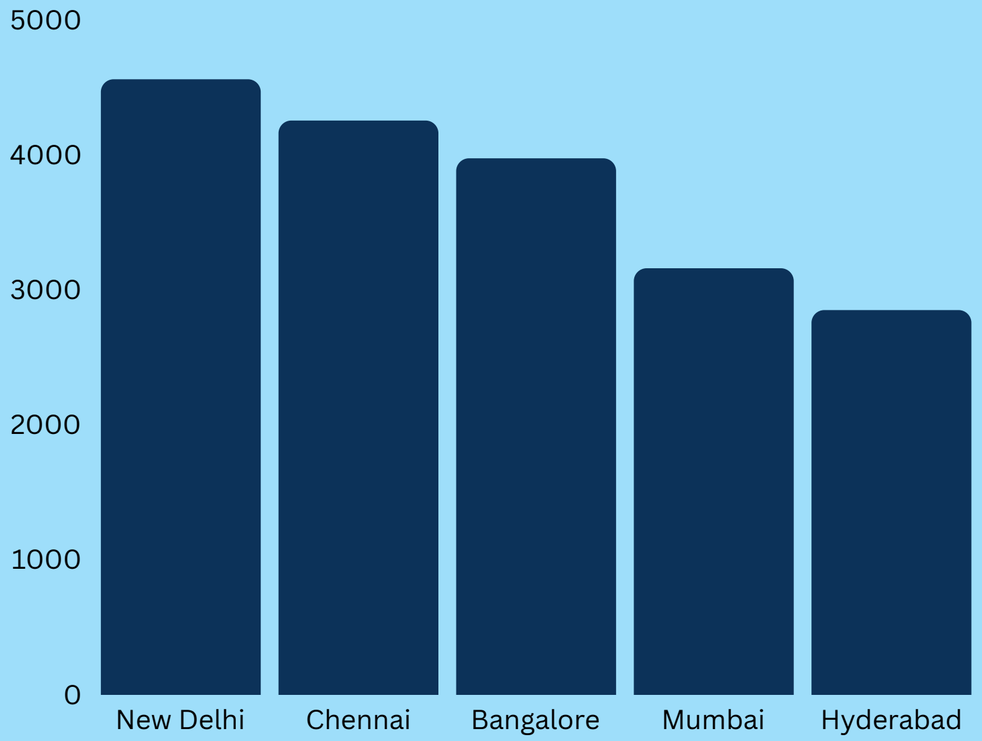
From the Dataset we can found that there are **4 unique categories** which are **Others, Apparels, Shoes,** & **Home**

2. List the top 5 shipping cities in terms of the number of orders.

```sql
SELECT Shipping_city, COUNT(*) AS Order_Count
FROM paytm_data
GROUP BY Shipping_city
ORDER BY Order_Count DESC
LIMIT 5;
```

OUTPUT :

Shipping_city	Order_Count
New Delhi	4560
Chennai	4254
Bangalore	3974
Mumbai	3159
Hyderabad	2849



The data shows **top 5 cities** according to **Order Count**

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

```
SELECT *
FROM paytm_data
WHERE Category_grouped = "Apparels" ;
```

OUTPUT :

S.no	Name	Shipping_city	Category_Grouped	Category	Sub_category	Product_Gender	Segment	Class	Family	Brand	Brick	Item_NM
2	AMIT GALPHADE	Ahmedabad	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKINS Navy Blue Tights
4	MALLIKARJUNA H	Bangalore	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKINS Navy Blue Tights
10	ASHWIN GIDWANI	Pune	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKINS Navy Blue Tights
23	NAGA KISHORE	Bangalore	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKINS Navy Blue Tights
30	DEEPAK THAMBI	Bangalore	Apparels	Women Apparel	Ethnic	WOMEN	WOMENS WEAR	SETS	ETHNIC	SANGRIA	SUIT SET	Navy Blue Georgette Bro
32	SATHIYA NARAYAN	Chennai	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKINS Navy Blue Tights
38	roto robo	Kottayam	Apparels	Women Apparel	Ethnic	WOMEN	WOMENS WEAR	SETS	ETHNIC	SANGRIA	SUIT SET	Navy Blue Georgette Bro
46	sagar barman	Kolkata	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKINS Navy Blue Tights

The query **retrieves all information** from the database where the **category** is **Apparels.**

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

```
SELECT *
FROM paytm_data
WHERE Sale_Flag = 'On Sale';
```

OUTPUT :

S.no	Name	Shipping_city	Category_Grouped	Category	Sub_category	Product_Gender	Segment	Class	Family	Brand	Brick	Ite
2	AMIT GALPHADE	Ahmedabad	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKI
4	MALLIKARJUNA H	Bangalore	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKI
10	ASHWIN GIDWANI	Pune	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKI
16	Rompelli GopalK	Salem	Shoes	Men Footwear	Mens Footwear	MEN	MENS FOOTWEAR	Other	SPORTS	ADIDAS	RUNNING SHOES	Adi
20	prabhakar reddy	Jhansi	Others	WATCHES	WATCHES	MEN	WOMENS ACCESSORIES	WATCHES	NULL	PLAYBOY	ANALOG WATCH	Bpb
22	RAHUL SINGH PAT	Jabalpur	Others	Bags	Bags	WOMEN	WOMEN	Other	NULL	HIDESIGN	HANDBAG	Dov
23	NAGA KISHORE	Bangalore	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKI
24	kamla singh	Lucknow	Shoes	Men Footwear	Mens Footwear	MEN	MENS FOOTWEAR	Other	SPORTS	REEBOK	RUNNING SHOES	Ver

The query **retrieves all information** from the database where the **Sale Flag** is **On Sale.**

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

```
SELECT *
FROM paytm_data
WHERE Sale_Flag = 'On Sale';
```

OUTPUT :

S.no	Name	Shipping_city	Category_Grouped	Category	Sub_category	Product_Gender	Segment	Class	Family	Brand	Brick	Ite
2	AMIT GALPHADE	Ahmedabad	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKI
4	MALLIKARJUNA H	Bangalore	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKI
10	ASHWIN GIDWANI	Pune	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKI
16	Rompelli GopalK	Salem	Shoes	Men Footwear	Mens Footwear	MEN	MENS FOOTWEAR	Other	SPORTS	ADIDAS	RUNNING SHOES	Adi
20	prabhakar reddy	Jhansi	Others	WATCHES	WATCHES	MEN	WOMENS ACCESSORIES	WATCHES	NULL	PLAYBOY	ANALOG WATCH	Bpb
22	RAHUL SINGH PAT	Jabalpur	Others	Bags	Bags	WOMEN	WOMEN	Other	NULL	HIDESIGN	HANDBAG	Dov
23	NAGA KISHORE	Bangalore	Apparels	Sports Equipment	Sports Apparel	MEN	MENS WEAR	TOPS	SPORT & ADVENTURE	SKINS	TIGHTS	SKI
24	kamla singh	Lucknow	Shoes	Men Footwear	Mens Footwear	MEN	MENS FOOTWEAR	Other	SPORTS	REEBOK	RUNNING SHOES	Ver

The query **retrieves all information** from the database where the **Sale Flag** is **On Sale.**

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

```sql
SELECT *
FROM (
    SELECT *, RANK() OVER (ORDER BY Item_Price DESC) AS price_rank
    FROM paytm_data
) AS ranked_items
WHERE price_rank = 1;
```

OUTPUT :

coupon_money_effective	Coupon_Percentage	Quantity	Cost_Price	Item_Price	Special_Price_effective	paid_pr_effective	Value_CM1	Value_CM2	special_price
	0	1	3822.45	13500	4990	4990	588.6	-2956.4	4990
	0	1	3822.45	13500	4990	4990	588.6	-2956.4	4990
	0	1	3822.45	13500	4990	4990	588.6	-2956.4	4990
	0	1	3822.45	13500	4990	4990	588.6	-2956.4	4990
	0	1	3822.45	13500	4990	4990	588.6	-2956.4	4990
	0	1	3822.45	13500	4990	4990	588.6	-2956.4	4990
	0	1	3822.45	13500	4990	4990	588.6	-2956.4	4990
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

The query **retrieves all information** from the database where the **Item Price** is **maximum** in the database, which is **13,500 INR**

6. Apply conditional formatting to highlight all products with a "Special_Price_effective" value below $50 (INR 4,179.69) in red.

```
SELECT *
FROM paytm_data
WHERE Special_Price_effective < 4179.69;
```

OUTPUT :

coupon_money_effective	Coupon_Percentage	Quantity	Cost_Price	Item_Price	Special_Price_effective	paid_pr_effective	Value_CM1	Value_CM2	special_price	Paid_pr
0	0	1	2186.66	4095	4095	4095	1433.07	955.07	4095	4095
0	0	1	2412	4020	4020	4020	1608	1430	4020	4020
0	0	1	2412	4020	4020	4020	1608	1430	4020	4020
0	0	1	2429.55	4550	4095	4095	1190.18	1041.03	4095	4095
0	0	1	3945.85	5900	4130	4130	-294.86	-490.86	4130	4130
0	0	1	2186.66	4095	4095	4095	1433.07	955.07	4095	4095
0	0	1	2429.55	4550	4095	4095	1190.18	1041.03	4095	4095
0	0	1	2186.66	4095	4095	4095	1433.07	955.07	4095	4095

The query **retrieves all information** from the database where the **Special price effective** is **below $50**

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

```
SELECT Category_grouped, SUM(Item_Price) AS Total_Sales_Value
FROM paytm_data
GROUP BY Category_grouped;
```

OUTPUT :

Category_grouped	Total_Sales_Value
Others	115016278
Apparels	59773705
shoes	114653487
Home	7012968

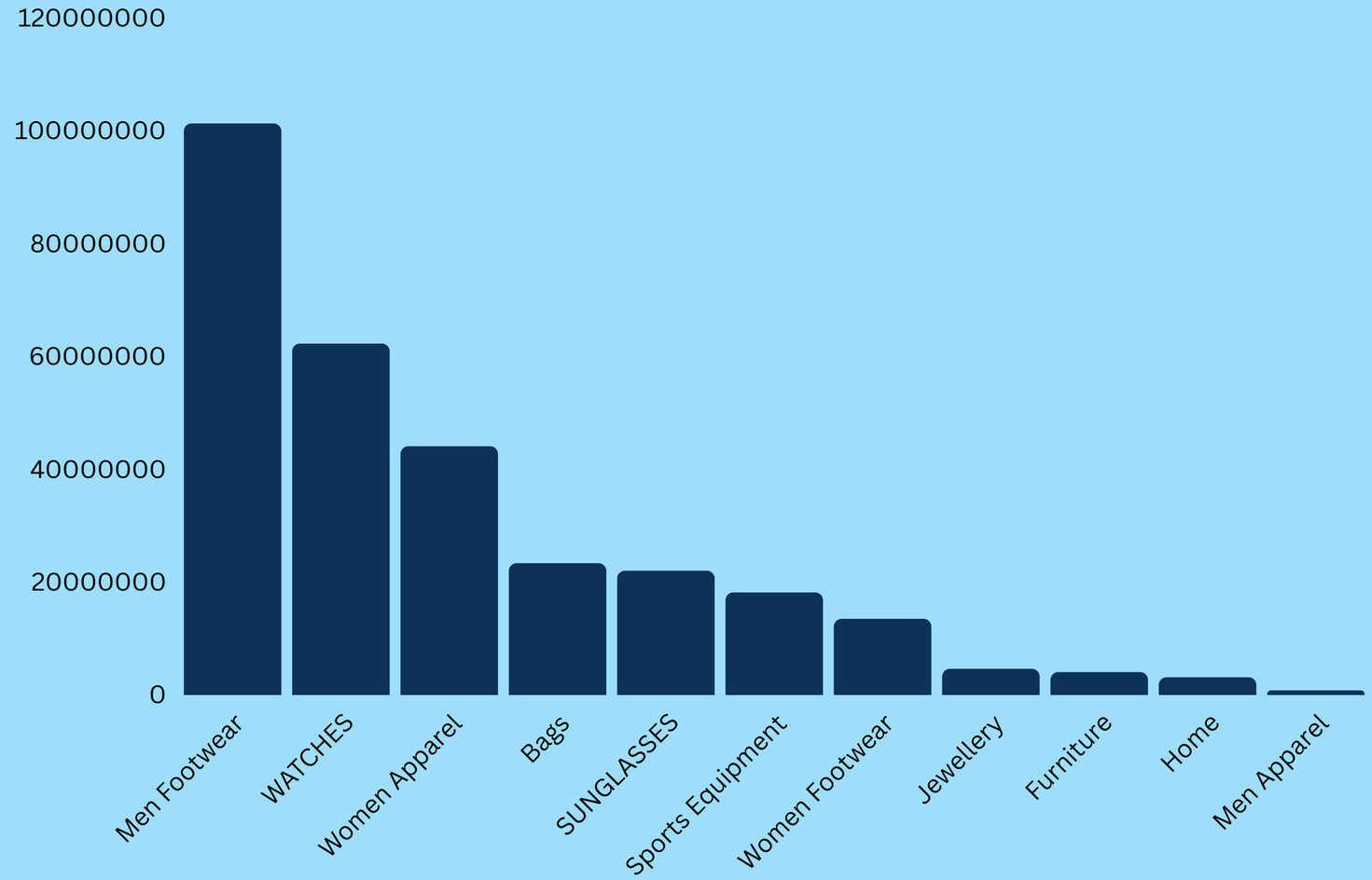
From this query we can find the **Total sales** for each **Category Group**.

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

```
SELECT Category, SUM(Item_Price) AS Total_Sales
FROM paytm_data
GROUP BY Category;
```

OUTPUT :

Category	Total_Sales
SUNGLASSES	21935695
Sports Equipment	18085020
Bags	23272288
Men Footwear	101245089
Women Footwear	13408398
WATCHES	62213793
Women Apparel	44010575
Furniture	3961755
Men Apparel	723305
Jewellery	4549307
Home	3051213



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

```
SELECT Product_Gender, AVG(Quantity) AS Average_Quantity
FROM paytm_data
WHERE category like "%Apparel%"
GROUP BY Product_Gender;
```

OUTPUT :

	Product_Gender	Average_Quantity
▶	WOMEN	1
	MEN	1

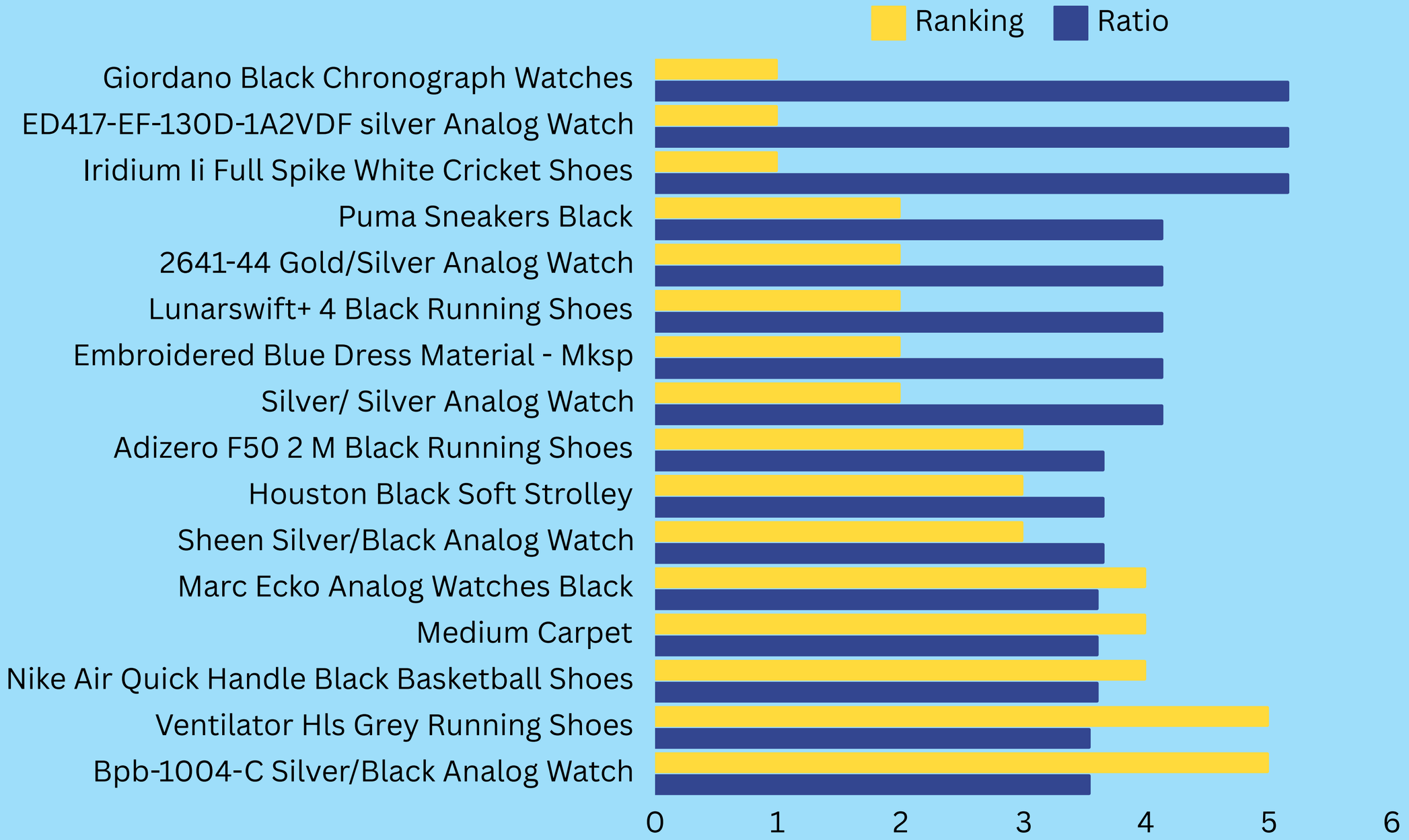
From this query we can find that **average quantity** for **any apparel category** is **1** for **Men** as well as **Women**

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

```sql
WITH ranked_data AS (
    SELECT Item_NM,
           DENSE_RANK() OVER (ORDER BY Ratio DESC) AS Ranking,
           Ratio FROM (
        SELECT *,
               Value_CM1 / NULLIF(Value_CM2, 0) AS Ratio
        FROM paytm_data
    ) AS ranked_ratios)
SELECT DISTINCT Ranking, Item_NM, Ratio
FROM ranked_data
WHERE Ranking <= 5;
```

OUTPUT :

Ranking	Item_NM	Ratio
1	Giordano Black Chronograph Watches	5.16530367462654
1	ED417-EF-130D-1A2VDF silver Analog Watch	5.16530367462654
1	Iridium Ii Full Spike White Cricket Shoes	5.16530367462654
2	Puma Sneakers Black	4.139717425431711
2	2641-44 Gold/Silver Analog Watch	4.139717425431711

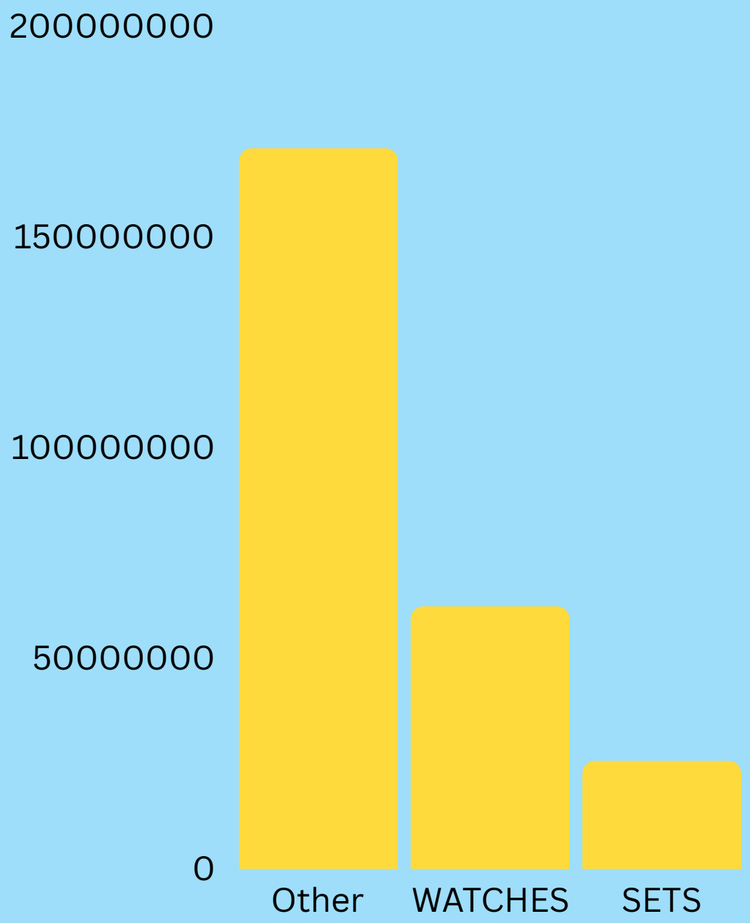

Ranking
Ratio
Giordano Black Chronograph Watches
ED417-EF-130D-1A2VDF silver Analog Watch
Iridium Ii Full Spike White Cricket Shoes
Puma Sneakers Black
2641-44 Gold/Silver Analog Watch
Lunarswift+ 4 Black Running Shoes
Embroidered Blue Dress Material - Mksp
Silver/ Silver Analog Watch
Adizero F50 2 M Black Running Shoes
Houston Black Soft Strolley
Sheen Silver/Black Analog Watch
Marc Ecko Analog Watches Black
Medium Carpet
Nike Air Quick Handle Black Basketball Shoes
Ventilator Hls Grey Running Shoes
Bpb-1004-C Silver/Black Analog Watch
0
1
2
3
4
5
6

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

```sql
SELECT Class, SUM(Item_Price) AS Total_Sales
FROM paytm_data
GROUP BY Class
ORDER BY Total_Sales DESC
LIMIT 3;
```

OUTPUT :

Class	Total_Sales
Other	170968090
WATCHES	62213793
SETS	25529167

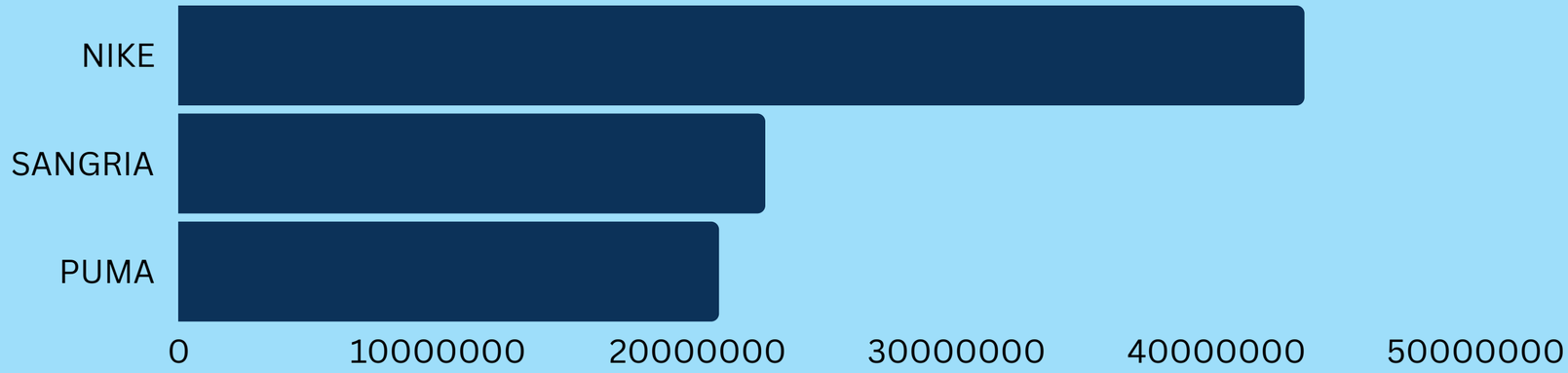


12. Find the total sales for each "Brand" and display the top 3 brands in terms of sales.

```sql
SELECT Brand, SUM(Item_Price) AS Total_Sales
FROM paytm_data
GROUP BY Brand
ORDER BY Total_Sales DESC
LIMIT 3;
```

OUTPUT :

Brand	Total_Sales
NIKE	43406505
SANGRIA	22614700
PUMA	20832027



13. Calculate the total revenue generated from "Apparels" category products with a "Sale_Flag" of 'On Sale'.

```
SELECT Category_grouped,Sale_Flag, SUM(Item_Price) AS Total_Revenue
FROM paytm_data
WHERE Category_grouped = 'Apparels' AND Sale_Flag = 'On Sale';
```

OUTPUT :

	Category_grouped	Sale_Flag	Total_Revenue
▶	Apparels	On Sale	22813541

From this query we get to know the **Total Revenue** for the **Apparels category** where the items are **On Sale**.

14. Identify the top 5 shipping cities based on the average order value (total sales amount divided by the number of orders) and display their average order values.

```sql
WITH ranked_cities AS (
    SELECT Shipping_city,
           SUM(Item_Price) / COUNT(*) AS Average_Order_Value,
           DENSE_RANK() OVER (ORDER BY SUM(Item_Price) / COUNT(*) DESC) AS City_Rank
    FROM paytm_data
    GROUP BY Shipping_city
)
SELECT Shipping_city, Average_Order_Value,City_Rank
FROM ranked_cities
WHERE City_Rank <= 5;
```

OUTPUT :

Shipping_city	Average_Order_Value	City_Rank
Panagudi	13500	1
Nathdwara	13500	1
kavaratti	13500	1
Karauli	13500	1
Gopeshwar	13500	1
Vallabh Vidyanagar	13500	1
Dalhousie	13500	1
laxminagar	13500	1
pulivendla	13500	1
akole	13500	1
chhourai	13500	1
Seoni Malwa	13500	1
Rameswaram	12375	2

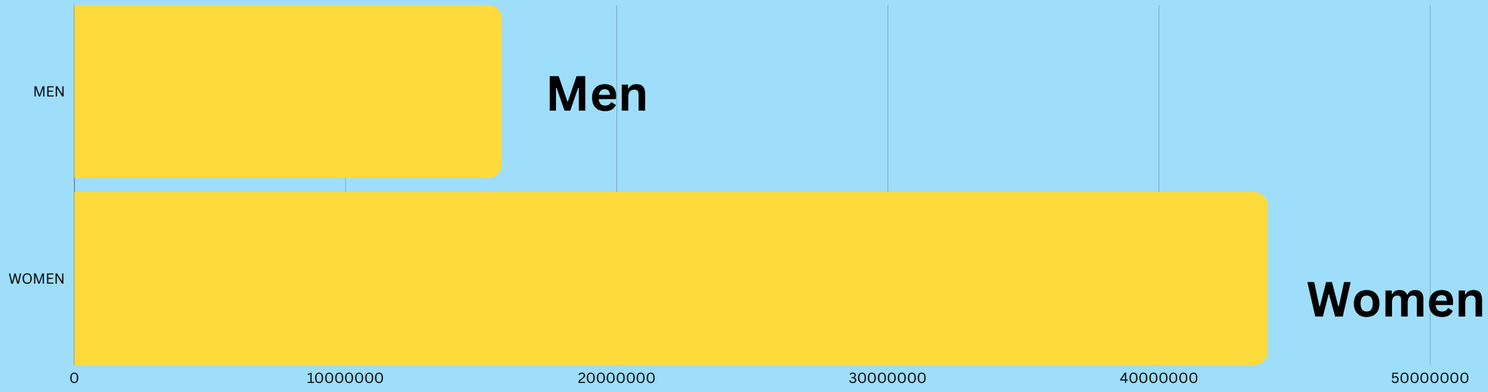
This output shows the **top 5 shipping cities** , **ranked** according to their **average order values**. As multiple cities has same average order value we will get **34 results** in **top 5.**

15. Determine the total number of orders and the total sales amount for each "Product_Gender" within the "Apparels" category_grouped.

```sql
SELECT Product_Gender,
       COUNT(*) AS Total_Orders,
       SUM(Item_Price) AS Total_Sales_Amount
FROM paytm_data
WHERE Category_Grouped = 'Apparels'
GROUP BY Product_Gender;
```

OUTPUT :

	Product_Gender	Total_Orders	Total_Sales_Amount
▶	MEN	3208	15763130
	WOMEN	6931	44010575



Men

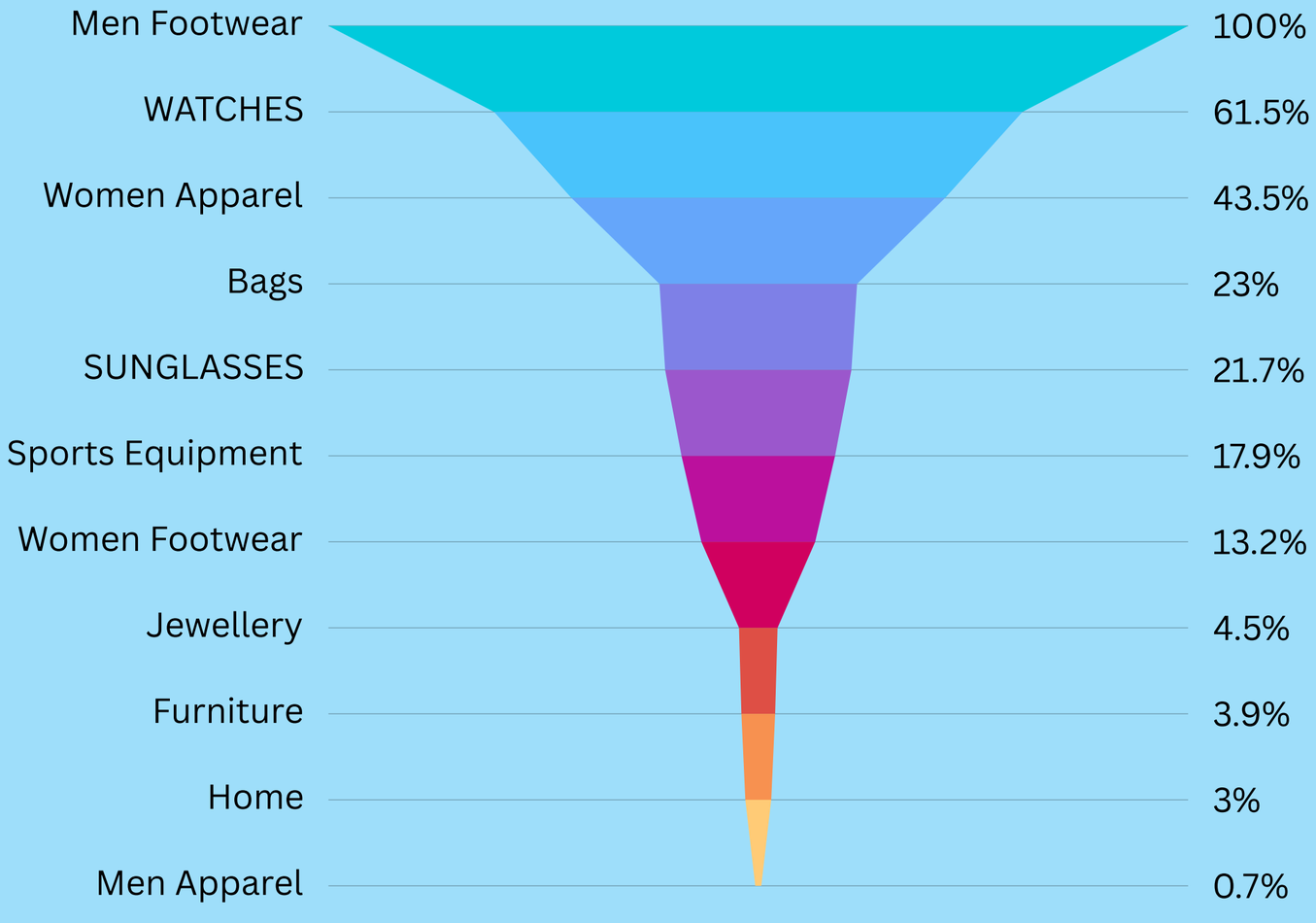
Women

16. Calculate the percentage contribution of each "Category" to the overall total sales.

```sql
SELECT Category,
Round(((SUM(Item_Price) / (SELECT SUM(Item_Price) FROM paytm_data)) * 100),2) AS Percentage_Contribution
FROM paytm_data
GROUP BY Category
ORDER BY Percentage_Contribution DESC;
```

OUTPUT :

	Category	Percentage_Contribution
▶	Men Footwear	34.15
	WATCHES	20.99
	Women Apparel	14.85
	Bags	7.85
	SUNGLASSES	7.4
	Sports Equipment	6.1
	Women Footwear	4.52
	Jewellery	1.53
	Furniture	1.34
	Home	1.03
	Men Apparel	0.24



17. Identify the "Category" with the highest average "Item_Price" and its corresponding average price.

```
SELECT Category, round(AVG(Item_Price),2) AS Average_Item_Price
FROM paytm_data
GROUP BY Category
ORDER BY Average_Item_Price DESC
LIMIT 1;
```

OUTPUT :

	Category	Average_Item_Price
▶	Furniture	7560.6

From the query we can find out that the **Furniture** is the category which has the **highest average item price**, which is **7560 INR.**

18. Calculate the total sales for each "Segment" and the average quantity sold per order for each segment.

```
SELECT
    Segment,
    SUM(Item_Price) AS Total_Sales,
    AVG(Quantity) AS Average_Quantity_Per_Order
FROM
    paytm_data
GROUP BY
    Segment;
```

OUTPUT :

Segment	Total_Sales	Average_Quantity_Per_Order
SUNGLASSES	18931045	1
MENS WEAR	15039825	1
UNISEX	15519256	1
MENS FOOTWEAR	101245089	1.0003400011333372
LADIES FOOTWEAR	13408398	1
WOMEN	30013131	1.0018676073874249
WOMENS ACCESSORIES	7386940	1.0105750165234633
WOMENS WEAR	44010575	1
MENS ACCESSORIES	13279676	1
MEN	22291728	1
LIVING	3961755	1
OUTDOOR & HIKING	3045195	1
MENS APPARELS	723305	1

The output shows the **average quantity ordered** for **each segment,** and most of the segment has average order **quantity 1.**

THANK'S FOR WATCHING

PRESENTED BY : BOLORAM MITRA