

# Retail Sales Analytics – India

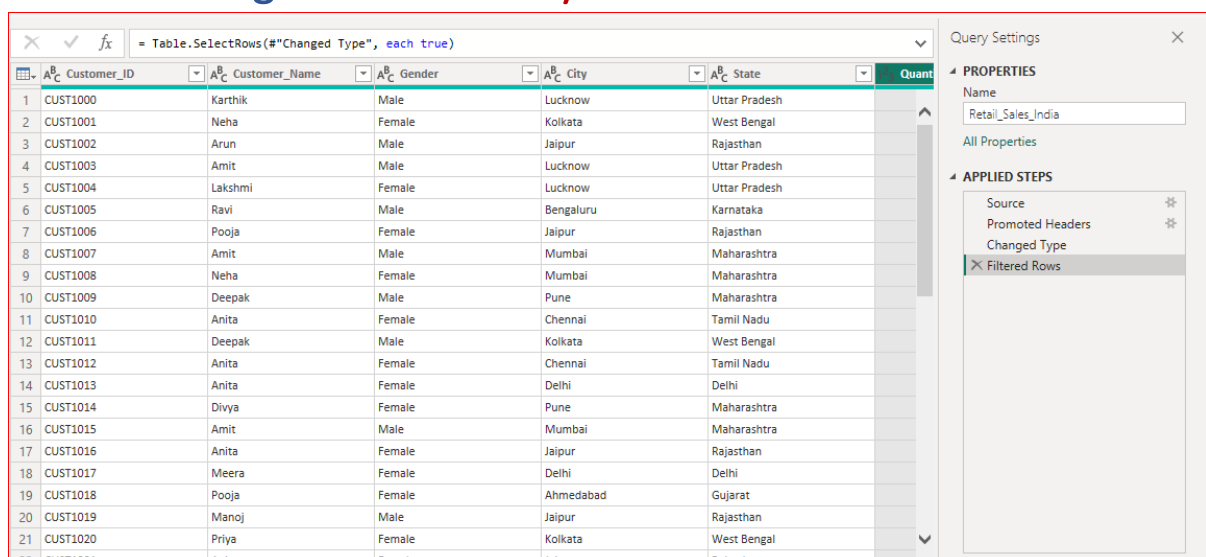
## About Dataset :

This dataset contains retail sales transactions from various cities and states across India, including customer details, product information, pricing, discounts, and total sales amounts.

It features multiple product categories and well-aligned brand associations, making it suitable for analyzing customer behavior, product performance, and regional sales trends.

A few missing values and duplicates are present, providing opportunities to apply practical data cleaning and transformation techniques in Power BI.

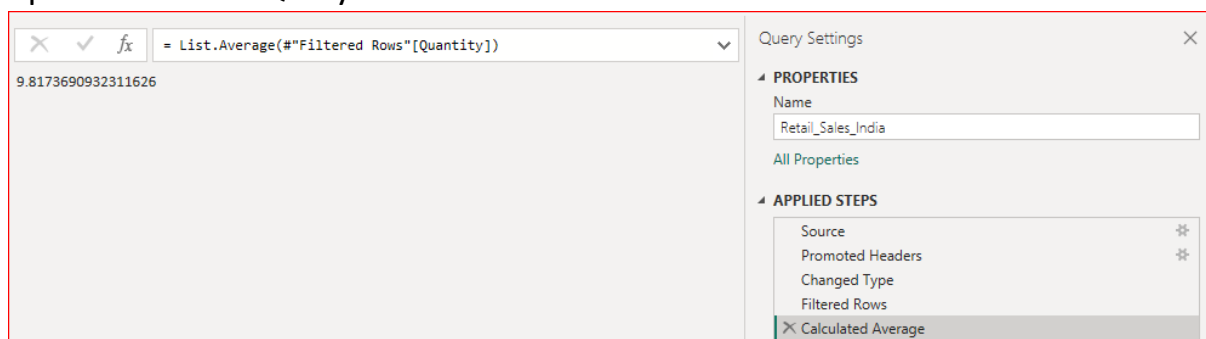
## Data Cleaning : Power Query



	Customer_ID	Customer_Name	Gender	City	State	Quantity
1	CUST1000	Karthik	Male	Lucknow	Uttar Pradesh	
2	CUST1001	Neha	Female	Kolkata	West Bengal	
3	CUST1002	Arun	Male	Jaipur	Rajasthan	
4	CUST1003	Amit	Male	Lucknow	Uttar Pradesh	
5	CUST1004	Lakshmi	Female	Lucknow	Uttar Pradesh	
6	CUST1005	Ravi	Male	Bengaluru	Karnataka	
7	CUST1006	Pooja	Female	Jaipur	Rajasthan	
8	CUST1007	Amit	Male	Mumbai	Maharashtra	
9	CUST1008	Neha	Female	Mumbai	Maharashtra	
10	CUST1009	Deepak	Male	Pune	Maharashtra	
11	CUST1010	Anita	Female	Chennai	Tamil Nadu	
12	CUST1011	Deepak	Male	Kolkata	West Bengal	
13	CUST1012	Anita	Female	Chennai	Tamil Nadu	
14	CUST1013	Anita	Female	Delhi	Delhi	
15	CUST1014	Divya	Female	Pune	Maharashtra	
16	CUST1015	Amit	Male	Mumbai	Maharashtra	
17	CUST1016	Anita	Female	Jaipur	Rajasthan	
18	CUST1017	Meera	Female	Delhi	Delhi	
19	CUST1018	Pooja	Female	Ahmedabad	Gujarat	
20	CUST1019	Manoj	Male	Jaipur	Rajasthan	
21	CUST1020	Priya	Female	Kolkata	West Bengal	
22	CUST1021	Anita	Female	Jaipur	Rajasthan	

## Quantity Column Cleaning :

The Quantity column contained null values, so I calculated the average quantity and used it to replace all missing entries through the Replace Values option in Power Query.



	Customer_ID	Customer_Name	Gender	City	State	Quantity
1	CUST1000	Karthik	Male	Lucknow	Uttar Pradesh	
2	CUST1001	Neha	Female	Kolkata	West Bengal	
3	CUST1002	Arun	Male	Jaipur	Rajasthan	
4	CUST1003	Amit	Male	Lucknow	Uttar Pradesh	
5	CUST1004	Lakshmi	Female	Lucknow	Uttar Pradesh	
6	CUST1005	Ravi	Male	Bengaluru	Karnataka	
7	CUST1006	Pooja	Female	Jaipur	Rajasthan	
8	CUST1007	Amit	Male	Mumbai	Maharashtra	
9	CUST1008	Neha	Female	Mumbai	Maharashtra	
10	CUST1009	Deepak	Male	Pune	Maharashtra	
11	CUST1010	Anita	Female	Chennai	Tamil Nadu	
12	CUST1011	Deepak	Male	Kolkata	West Bengal	
13	CUST1012	Anita	Female	Chennai	Tamil Nadu	
14	CUST1013	Anita	Female	Delhi	Delhi	
15	CUST1014	Divya	Female	Pune	Maharashtra	
16	CUST1015	Amit	Male	Mumbai	Maharashtra	
17	CUST1016	Anita	Female	Jaipur	Rajasthan	
18	CUST1017	Meera	Female	Delhi	Delhi	
19	CUST1018	Pooja	Female	Ahmedabad	Gujarat	
20	CUST1019	Manoj	Male	Jaipur	Rajasthan	
21	CUST1020	Priya	Female	Kolkata	West Bengal	
22	CUST1021	Anita	Female	Jaipur	Rajasthan	

×

## Replace Values

Replace one value with another in the selected columns.

Value To Find

null

Replace With

9

OK

Cancel

## Discount Coulmn Cleaning :

The Discount column contained null values, which were addressed by calculating the average unit price using statistical aggregation and replacing all missing entries with this value through the Replace Values option.

×

✓

fx

= List.Average("#Filtered Rows1"["Discount\_%"])

▼

15.154974489795919

Query Settings

PROPERTIES

Name

Retail\_Sales\_India

All Properties

APPLIED STEPS

Source

Promoted Headers

Changed Type

Filtered Rows

Quantity filled

Filtered Rows1

Calculated Average

×

## Replace Values

Replace one value with another in the selected columns.

Value To Find

null

Replace With

15

OK

Cancel

## Total Amount Column Cleaning :

The Total Amount column contained null values, so I created a Custom Column and applied the required calculation formula to accurately populate all missing values.

### Custom Column

Add a column that is computed from the other columns.

New column name

Custom column formula ⓘ

```
= [[Quantity]] * [Unit_Price] * (1-[#"Discount %"] /100)
```

Available columns

- Customer\_ID
- Customer\_Name
- Gender
- City
- State
- Quantity
- Unit\_Price

<< Insert

[Learn about Power Query formulas](#)

✓ No syntax errors have been detected.

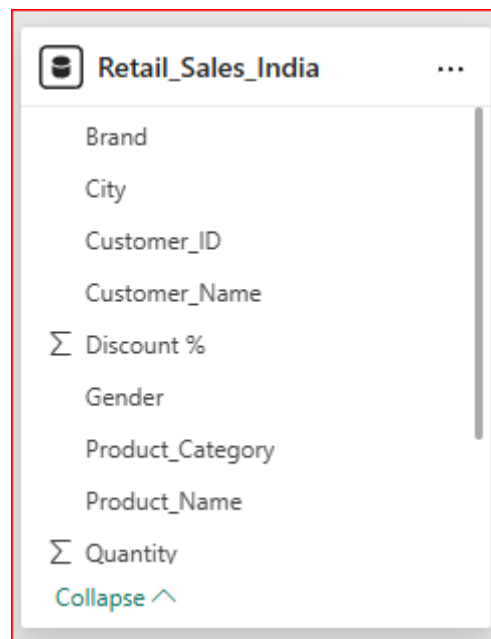
OK Cancel

## After Cleaning Dataset :

fx

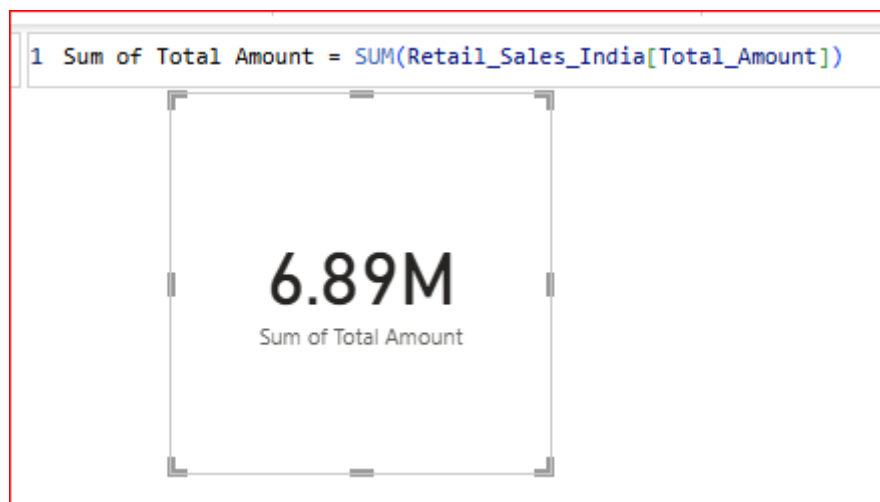
</

## Data Model :

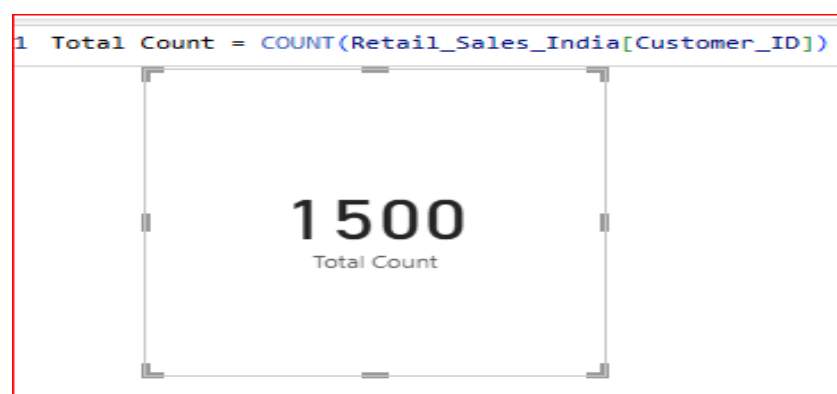


## Dax Measures :

### Sum of Total Amount :



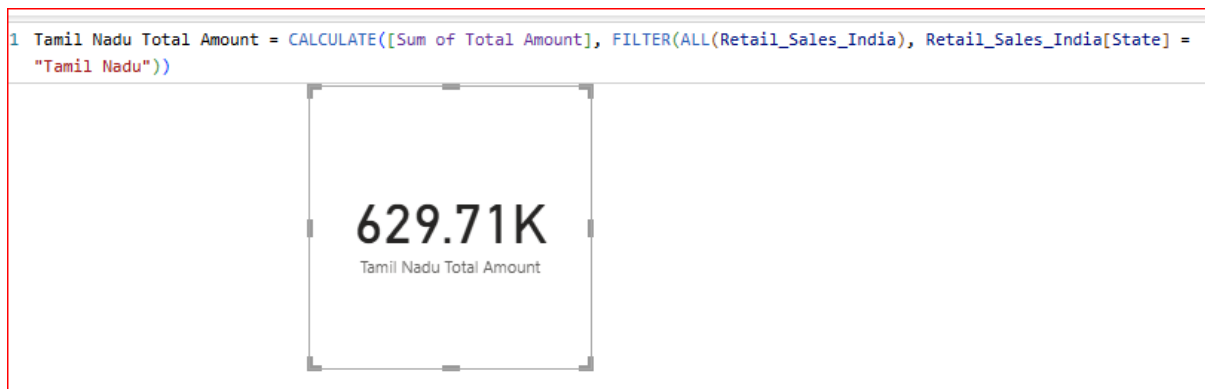
### Total Count :



## Average Price :

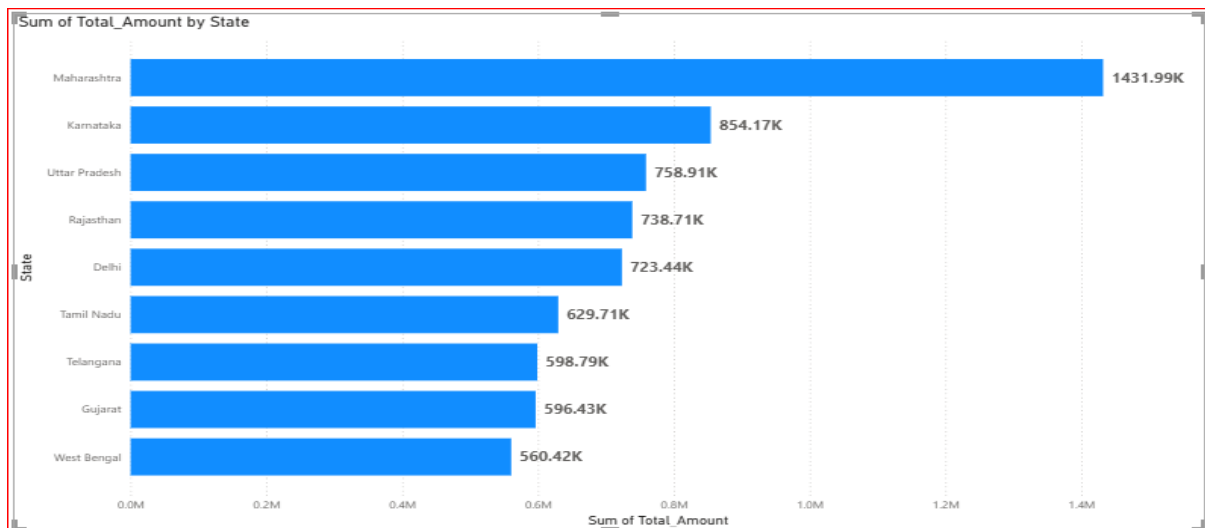


## Total Amount of Tamilnadu :

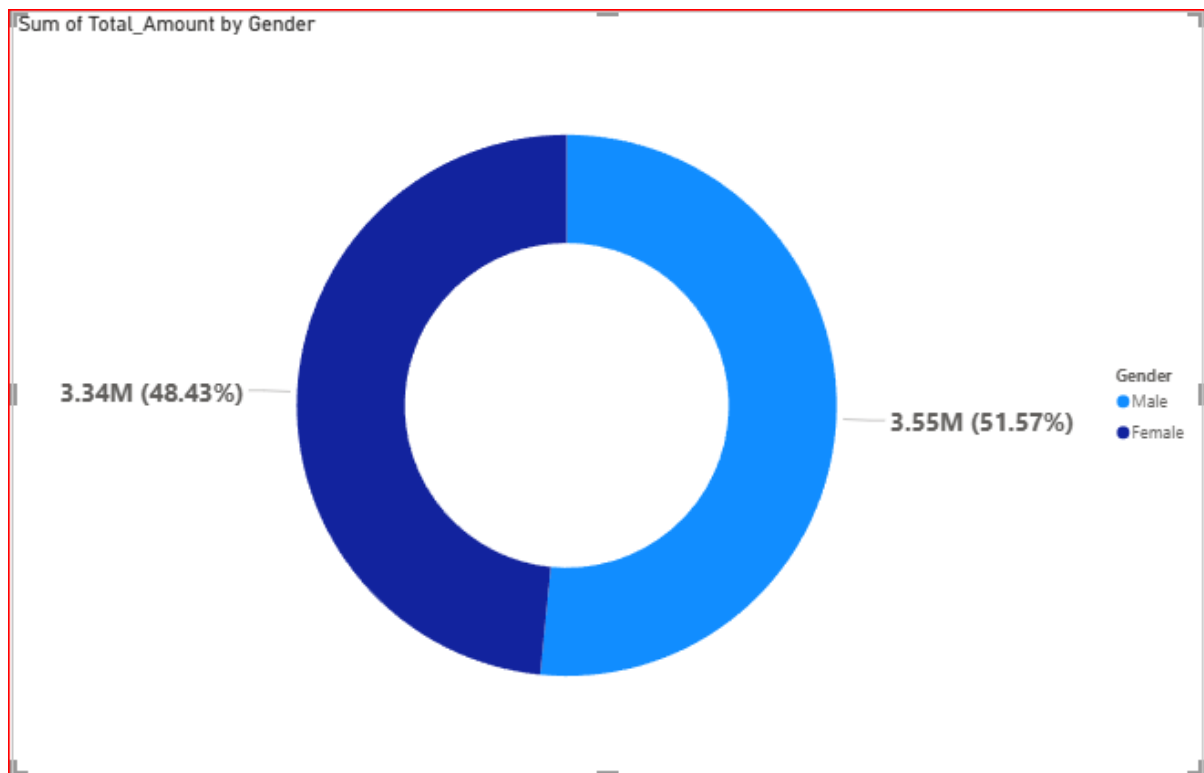


## Data Visualisation :

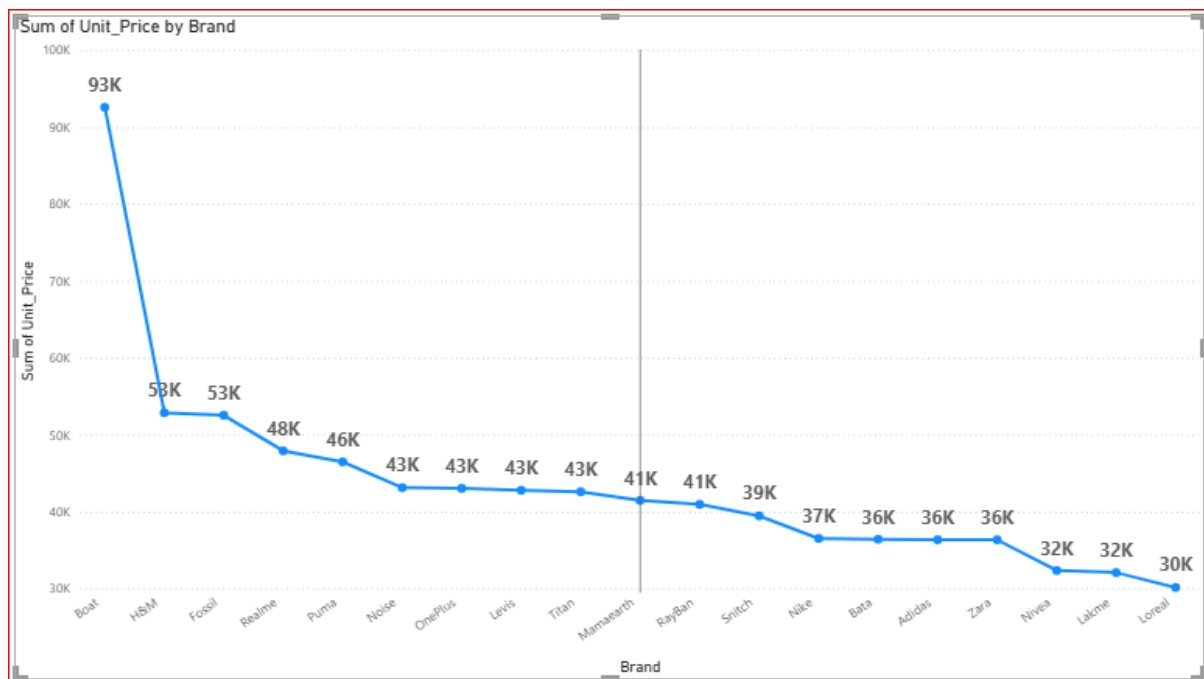
### Bar Chart : Sum of Total Amount by State



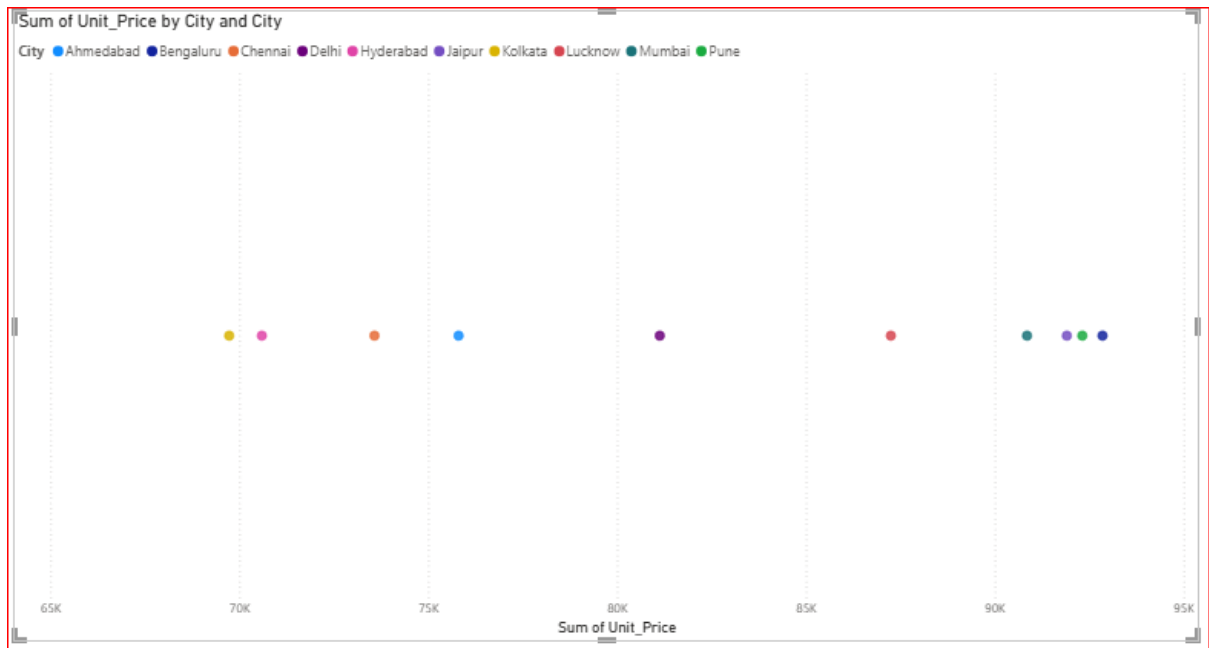
## Donut Chart : Sum of Total Amount by Gender



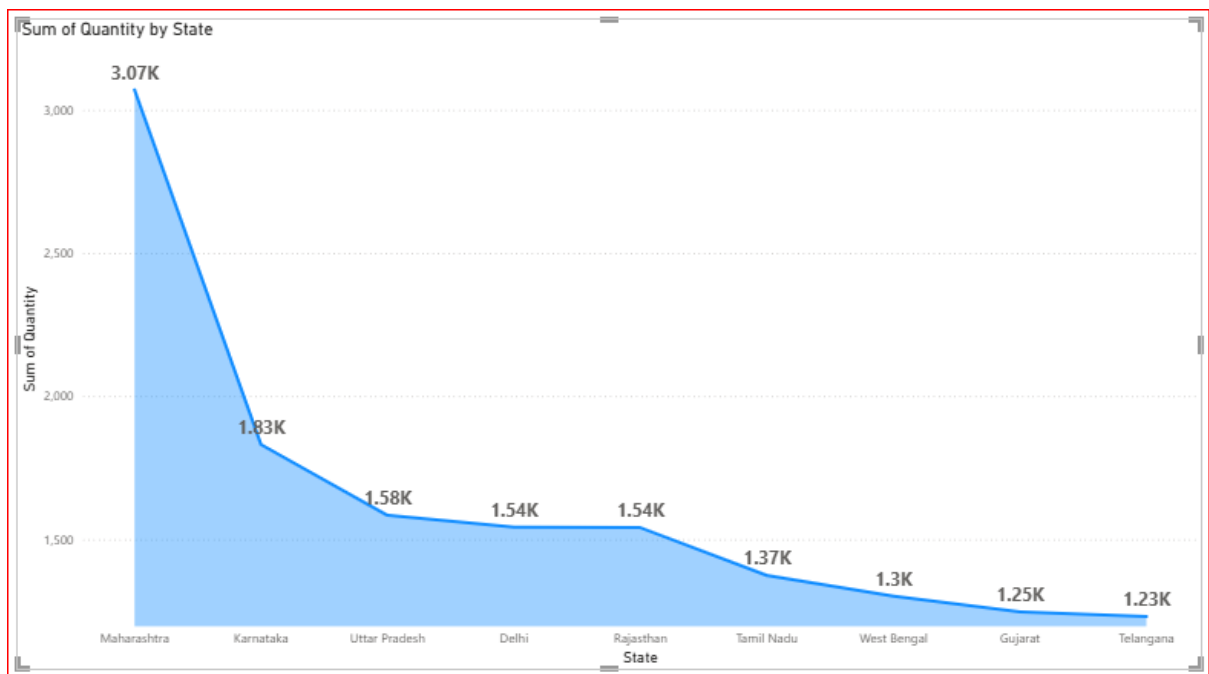
## Line Chart : Sum Of Unit Price by Brand



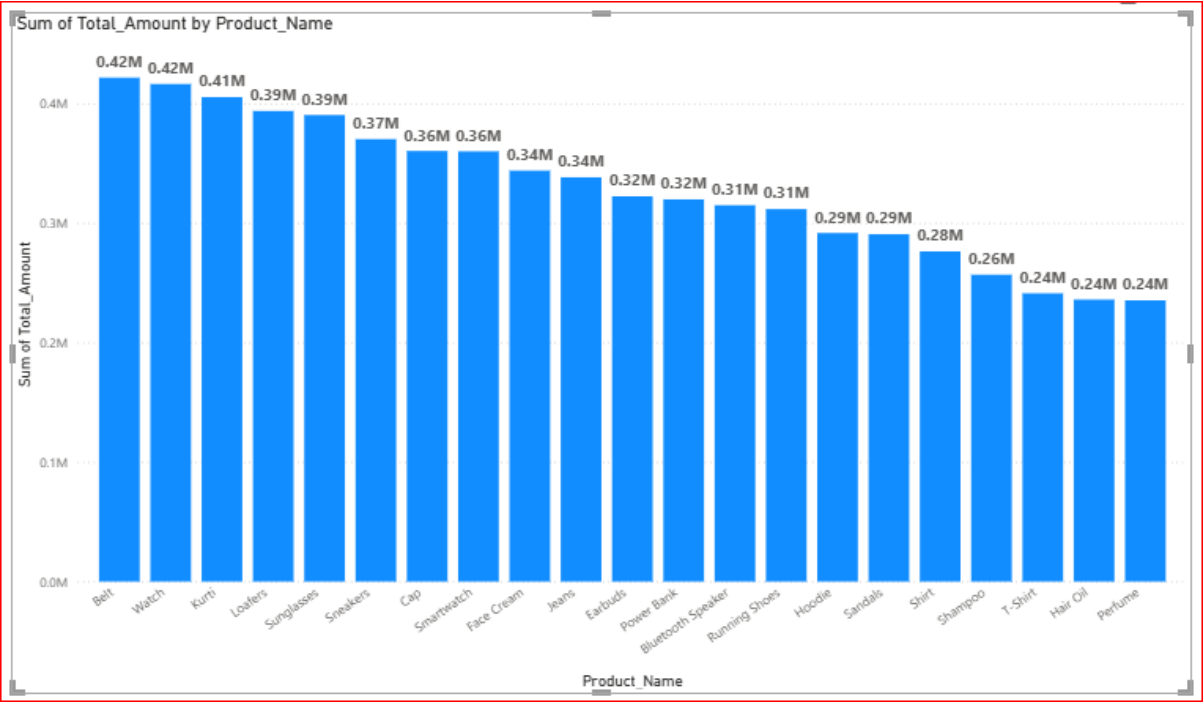
## Scatter Chart : Sum of Unit Price by City



## Area Chart : Sum of Quantity by State



Coulmn Chart : Sum of Total Amount by Product Name

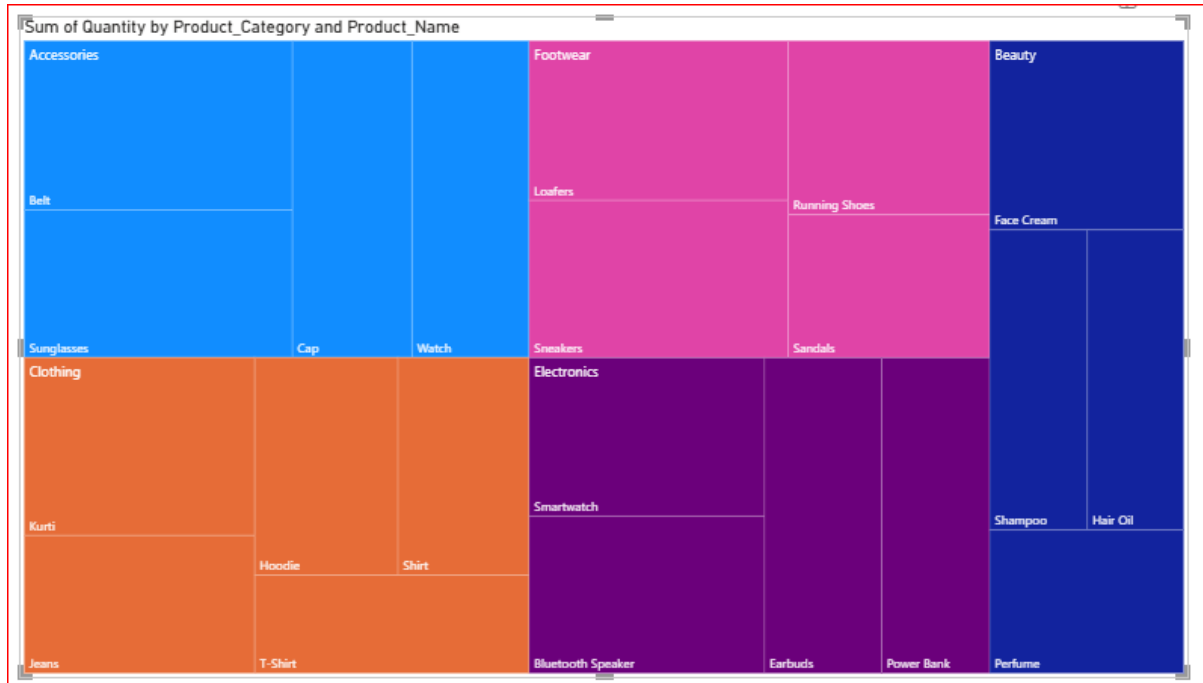


Matrix :

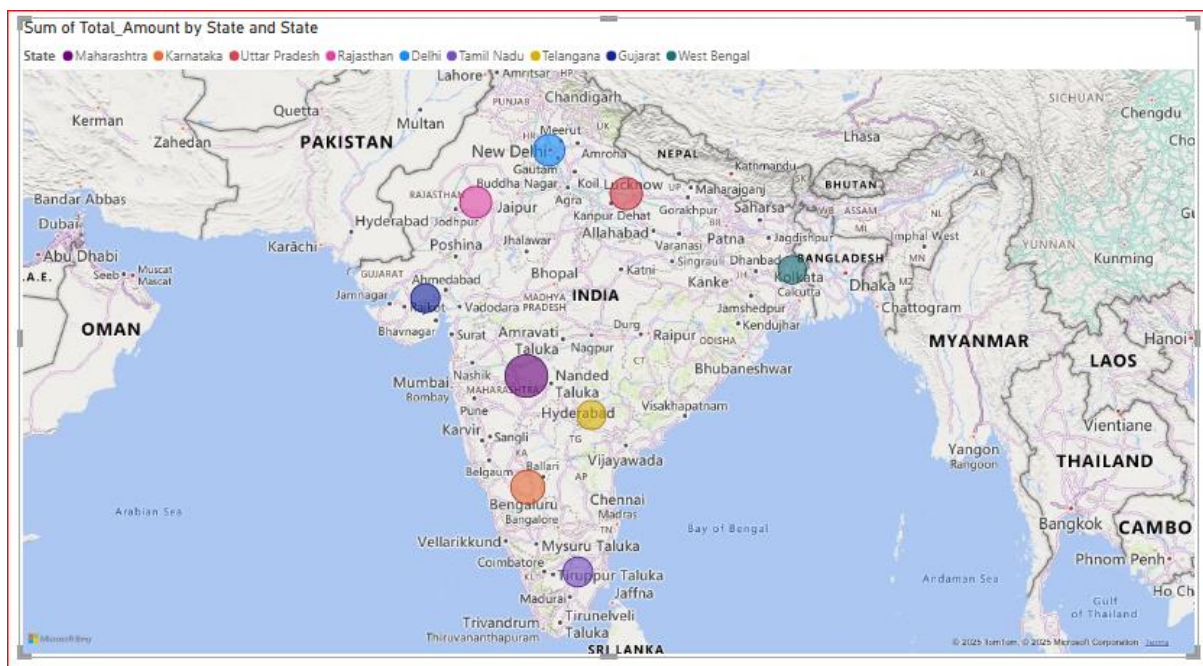
Product_Name	Delhi	Gujarat	Karnataka	Maharashtra	Rajasthan	Tamil Nadu	Telangana	Uttar Pradesh	West Bengal	Total
Belt	4854	5003	5158	11139	5465	3459	4628	5308	2620	47634
Accessories	4854	5003	5158	11139	5465	3459	4628	5308	2620	47634
Bluetooth Speaker	3078	4617	4314	11618	4931	4100	1146	5844	6270	45918
Electronics	3078	4617	4314	11618	4931	4100	1146	5844	6270	45918
Cap	4252	3496	5392	11641	4380	5127	2018	4084	5608	45998
Earbuds	4659	4025	3606	9394	4598	1211	4804	9334	2778	44409
Face Cream	5949	2540	4074	7973	5817	3162	3512	5091	1874	39992
Hair Oil	2631	2174	2321	6085	2027	3539	2666	1054	4398	26895
Hoodie	3352	2629	4415	5937	3199	1618	1730	3893	3107	29880
Jeans	4222	3524	4545	7602	3895	6036	4609	3907	1810	40150
Kurti	5346	2178	4292	9919	5247	5791	4241	3264	2971	43249
Loafers	694	1826	2817	13967	6838	4337	2180	5439	6253	44351
Perfume	4848	3000	3075	7828	1391	2866	2815	3963	3918	33704
Power Bank	1738	1770	3569	10309	5974	5622	4027	1187	5358	39554
Running Shoes	4125	4166	5106	5473	3131	3741	5069	2973	1390	35174
Sandals	3058	4273	3341	7873	3824	3354	3412	4285	334	33754
Shampoo	4408	5468	2776	6102	2019	2500	2957	4658	4579	35467
Shirt	4278	3197	5605	9075	3728	3411	365	1125	1048	31832
Smartwatch	4463	4916	6002	7552	6172	2496	5257	4577	3884	45319
Sneakers	3390	4023	4371	9616	4309	1816	5300	5396	4270	42491
Sunglasses	4082	6566	5075	9614	7921	3447	1891	6812	3335	48743
T-Shirt	2286	1786	4057	5964	4095	2617	4228	1044	265	26342
Watch	5416	4624	8942	8491	2947	3323	3736	4010	3656	45145
Total	81129	75801	92853	183172	91908	73573	70591	87248	69726	826001



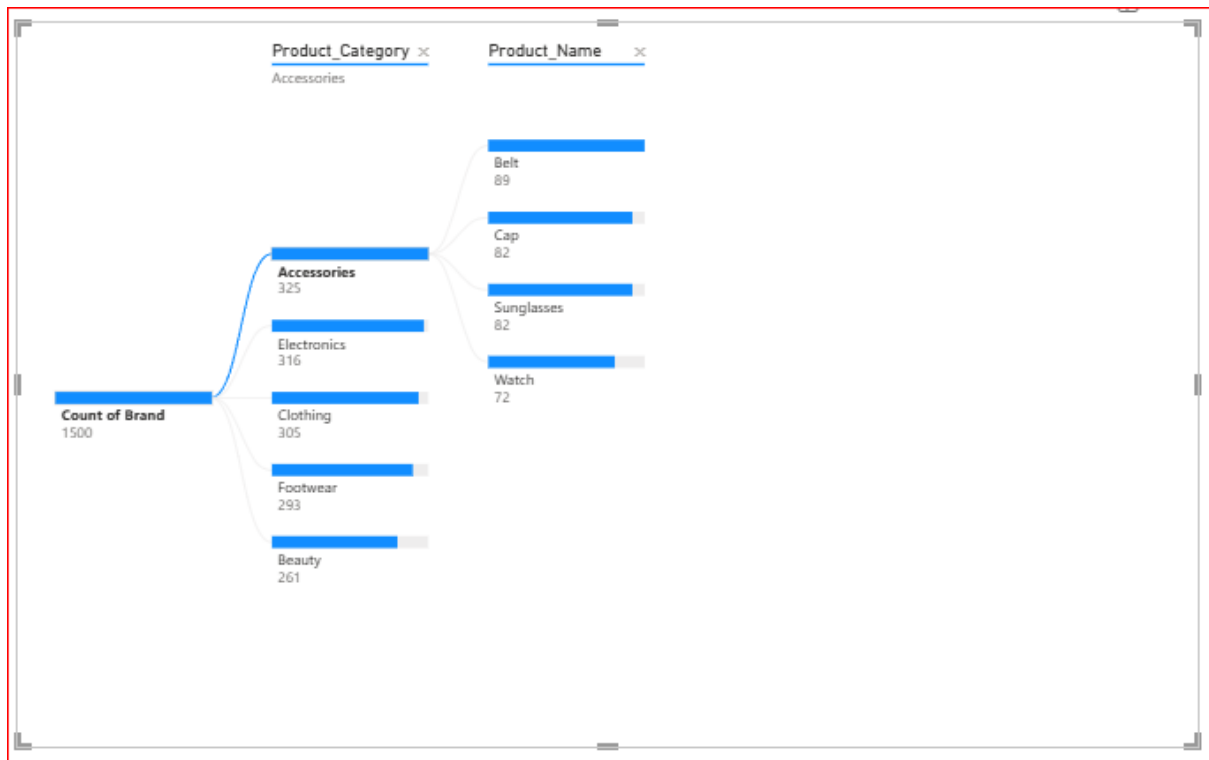
## Treemap Chart : Sum of Quantity By Product Category and Product Name



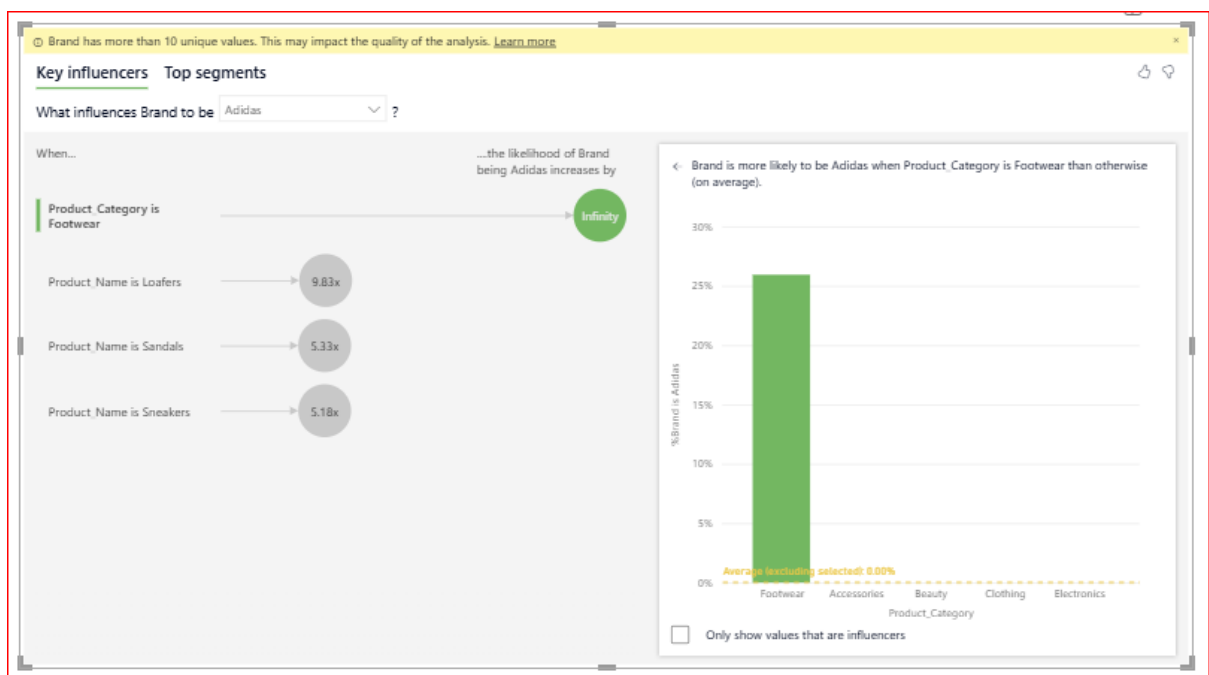
## Global (Map) Chart : Sum of Total Amount by state



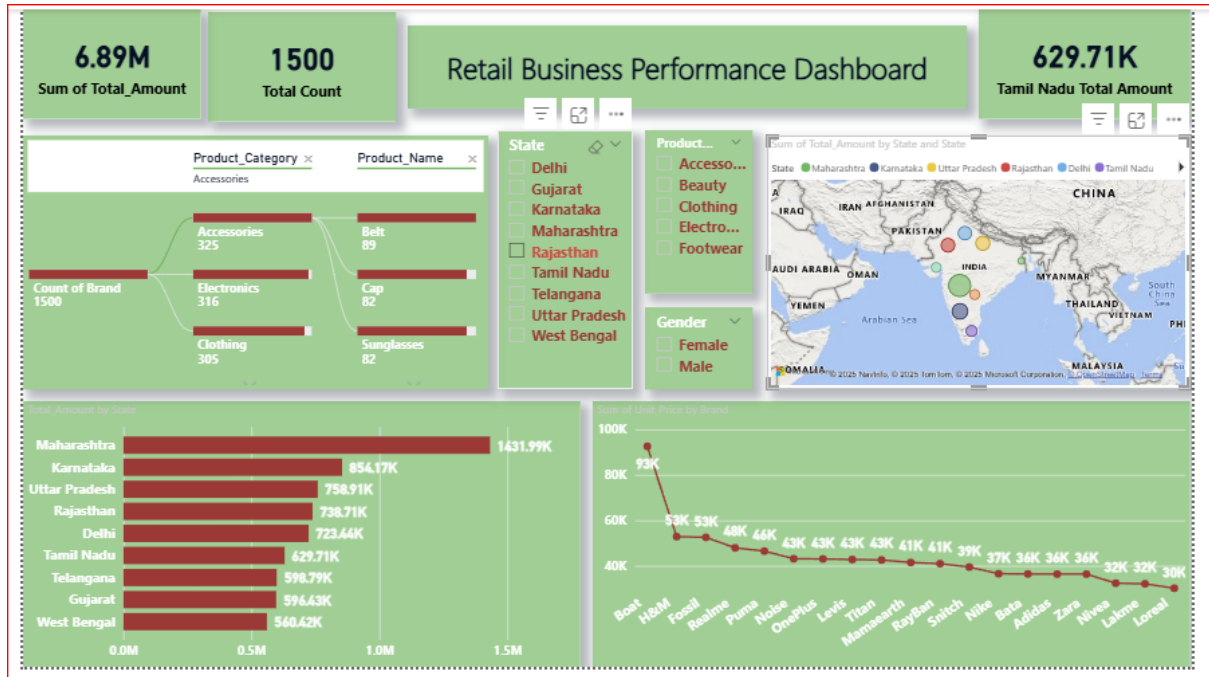
## Decomposition Tree (Ai) Chart :



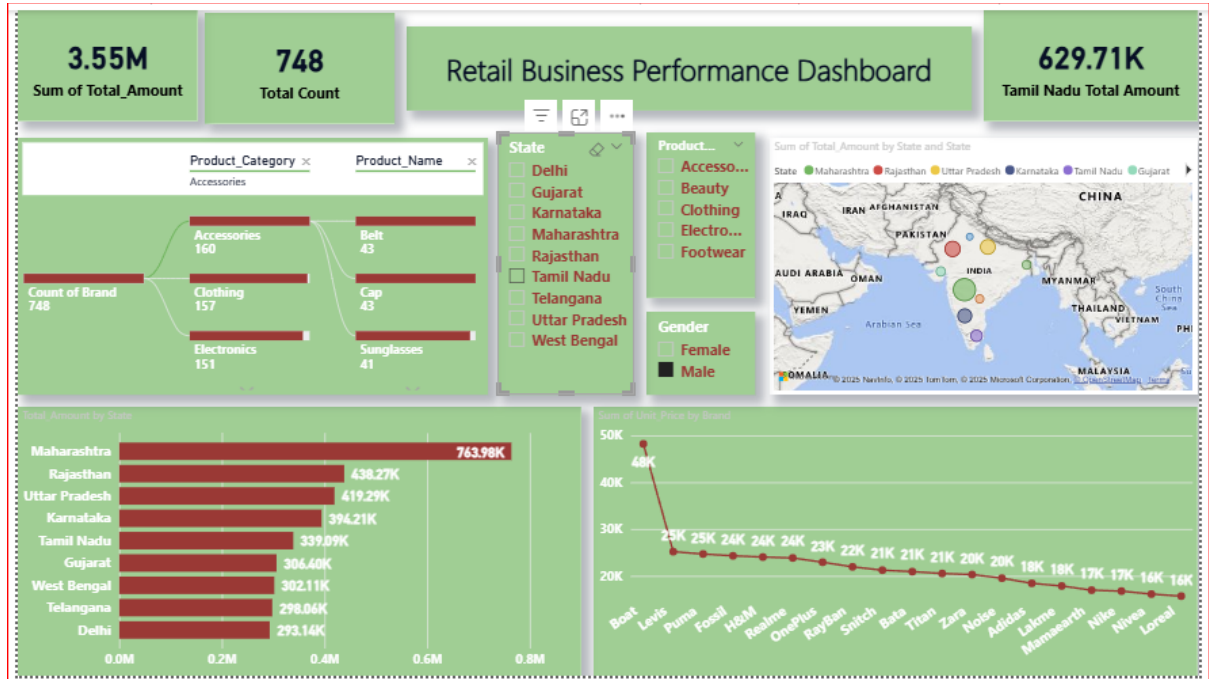
## Key Influencers (Ai) Chart :



## Final Dashboard



## Slicer Used Dashboard



## Insights Discovered :

- The business shows strong momentum across major Indian states, with **Maharashtra** emerging as the **top-performing market**, contributing the largest share of total revenue. Karnataka and Uttar Pradesh follow closely, reflecting consistent customer engagement and healthy sales distribution across regions.
- Customer interest is strongly centered around **Clothing and Accessories**, where brands like **Snitch, H&M, Levis, and RayBan** dominate the demand. These categories not only attract the highest number of orders but also drive a significant portion of the overall revenue.
- Tamil Nadu stands out as a stable contributor, generating over **₹629K in sales**, showcasing its importance as a consistent and reliable market even when not topping the chart.
- Brand performance reveals a clear pattern: a handful of popular brands deliver premium sales, while others show moderate traction. The price-trend visualization highlights that customers are willing to spend more on recognized brands, especially in Electronics and Fashion.
- When slicers are applied, the dashboard shifts dramatically — total sales drop from **6.89M to 3.55M**, and transaction count nearly halves — indicating that customer behavior changes significantly based on category, gender, and region filters. This suggests rich segmentation opportunities and highlights how different customer groups influence business outcomes.

## Conclusion :

The analysis clearly shows that sales performance is strongest in major urban markets, with Clothing and Accessories driving the highest demand across product categories. Brand preference plays a key role in customer purchasing behavior, and revenue distribution across states highlights valuable expansion opportunities. Overall, the dataset provides meaningful insights for strengthening product strategy, regional targeting, and customer engagement.

**THANK YOU**

~~~~~

