Retail Store Sales Dashboard

🔍 Objective:

The objective of this project is to **analyze sales performance and customer behavior** using key metrics such as **order volume, profit, payment modes, and geographic distribution**. Through interactive dashboards and visualizations in Power BI, the aim is to:

- Identify **high-performing product categories and sub-categories**
- Understand **customer preferences** in terms of **payment modes and locations**
- Gain insights into **profitability trends** across different states
- Provide **data-driven recommendations** for improving sales strategies and operational efficiency

---

🗃️ Dataset Overview:

The dataset used in this analysis is the *"Retail Store Sales"* dataset, which contains two CSV files – Details and Orders records of customer orders, including:

- **Order ID, Amount, Profit, Quantity category, Sub-Category, Payment Mode**
- **Order Date, Customer Name, State and City** respectively.

⚙️ Tools Used:

- **Python (Pandas):** Initial data cleaning and preprocessing
- **Power BI:** Interactive dashboard creation and visual analytics

---

📈 Key Areas of Analysis:

1. **Sales and Profit Trends:** Understanding how sales and profit vary across different regions, segments, and time periods.
2. **Profitability by Category/Sub-Category:** Identifying which products contribute most to overall profit or loss.
3. **Discount Impact:** Analyzing how varying discount levels influence profitability.
4. **Customer Insights:** Recognizing top-performing customers and segments.
5. **Shipping & Logistics:** Examining how shipping modes affect cost and delivery efficiency.

🧠 Conclusion:

This dashboard acts as a strategic tool to monitor sales health, optimize pricing strategies, and streamline inventory decisions. By visualizing these patterns, business stakeholders can make informed decisions to drive growth and profitability.

Loading the Dataset into Power BI

First, we check for any missing values or null values and duplicate values in the datasets using Python.

```
#Step 2:loading the dataset

df = pd.read_csv("Details.csv", encoding='ISO-8859-1')
df
```

	Order ID	Amount	Profit	Quantity	Category	Sub-Category	PaymentMode
0	B-25681	1096	658	7	Electronics	Electronic Games	COD
1	B-26055	5729	64	14	Furniture	Chairs	EMI
2	B-25955	2927	146	8	Furniture	Bookcases	EMI
3	B-26093	2847	712	8	Electronics	Printers	Credit Card
4	B-25602	2617	1151	4	Electronics	Phones	Credit Card
...	...	...	...	...	...	...	...
1495	B-25700	7	-3	2	Clothing	Hankerchief	COD
1496	B-25757	3151	-35	7	Clothing	Trousers	EMI
1497	B-25973	4141	1698	13	Electronics	Printers	COD
1498	B-25698	7	-2	1	Clothing	Hankerchief	COD
1499	B-25993	4363	305	5	Furniture	Tables	EMI

1500 rows × 7 columns

```
df1 = pd.read_csv("Orders.csv", encoding='ISO-8859-1')
df1
```

	Order ID	Order Date	CustomerName	State	City
0	B-26055	10-03-2018	Harivansh	Uttar Pradesh	Mathura
1	B-25993	03-02-2018	Madhav	Delhi	Delhi
2	B-25973	24-01-2018	Madan Mohan	Uttar Pradesh	Mathura
3	B-25923	27-12-2018	Gopal	Maharashtra	Mumbai
4	B-25757	21-08-2018	Vishakha	Madhya Pradesh	Indore
...	...	...	...	...	...
495	B-25742	03-08-2018	Ashwin	Goa	Goa
496	B-26088	26-03-2018	Bhavna	Sikkim	Gangtok
497	B-25707	01-07-2018	Shivani	Maharashtra	Mumbai
498	B-25758	22-08-2018	Shubham	Himachal Pradesh	Simla
499	B-26095	28-03-2018	Monisha	Rajasthan	Jaipur

500 rows × 5 columns

```
# Step 3: View basic info

print("Initial dataset shape:", df.shape)
print("\nColumn-wise missing values:\n", df.isnull().sum())
```

```
Initial dataset shape: (1500, 7)

Column-wise missing values:
 Order ID        0
Amount          0
Profit          0
Quantity        0
Category        0
Sub-Category    0
PaymentMode     0
dtype: int64
```

```
# Step 3: View basic info

print("Initial dataset shape:", df1.shape)
print("\nColumn-wise missing values:\n", df1.isnull().sum())
```

```
Initial dataset shape: (500, 5)

Column-wise missing values:
 Order ID        0
Order Date      0
CustomerName    0
State           0
City            0
dtype: int64
```

```
#Step 4:removing duplicates

df = df.drop_duplicates()
print("\nShape after removing duplicates:", df.shape)
```

```
Shape after removing duplicates: (1500, 7)
```

```
#Step 5:removing null values

df = df.dropna()
df.shape
```

```
(1500, 7)
```

```
#Step 4:removing duplicates

df1 = df1.drop_duplicates()
print("\nShape after removing duplicates:", df1.shape)
```

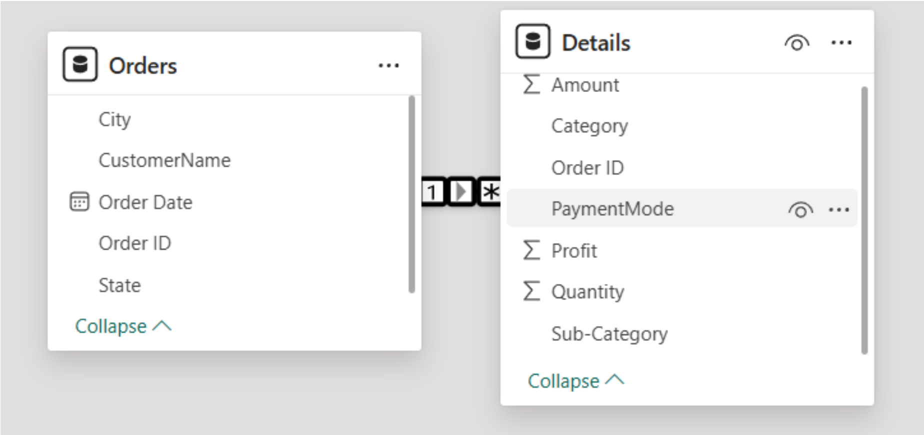
```
Shape after removing duplicates: (500, 5)
```

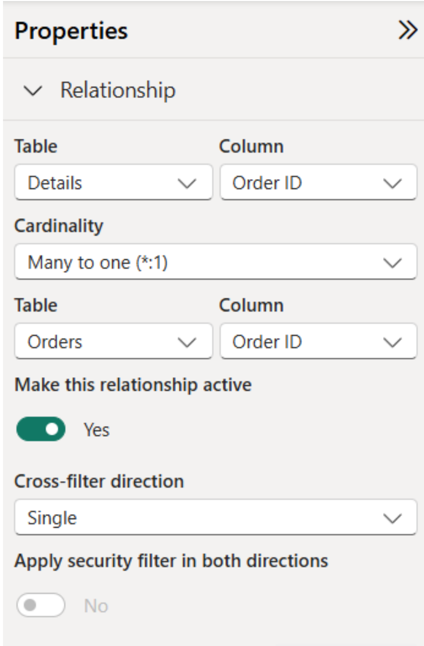
```
#Step 5:removing null values

df1 = df1.dropna()
df1.shape
```

```
(500, 5)
```

```
#THERE ARE NO DUPLICATES, NULL VALUES OR MISSING VALUES IN THE DATASETS
```

Since our dataset is split into two CSV files, using Power BI for generating dashboards will be insightful. I will also enable us to see the relationship between the two tables (Details and Orders) since both have a common column – Order ID. We load our datasets in Power BI after cleaning the data.



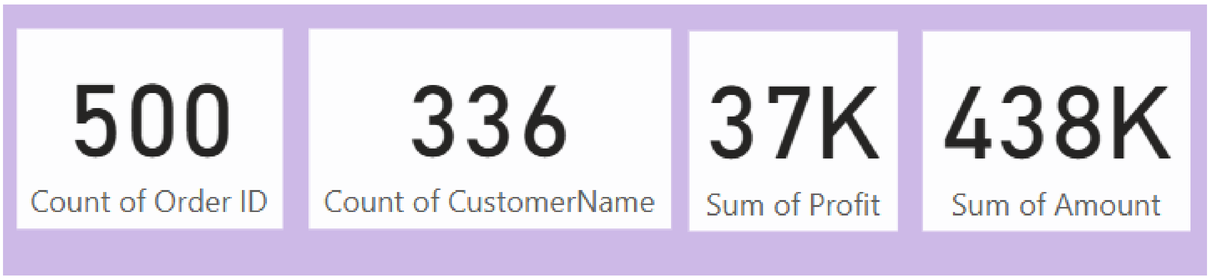


There is a Many-to-One (*:1) relationship from
Details[Order ID] → Orders[Order ID]
This is because each order (in Orders) can appear many times in the Details table (e.g., multiple products in one order) but each Order ID in Orders is unique.

It is a standard star schema design, where:

- Orders = Dimension Table (one row per order)
- Details = Fact Table (multiple rows per order for each product/item)

1. KPI Cards



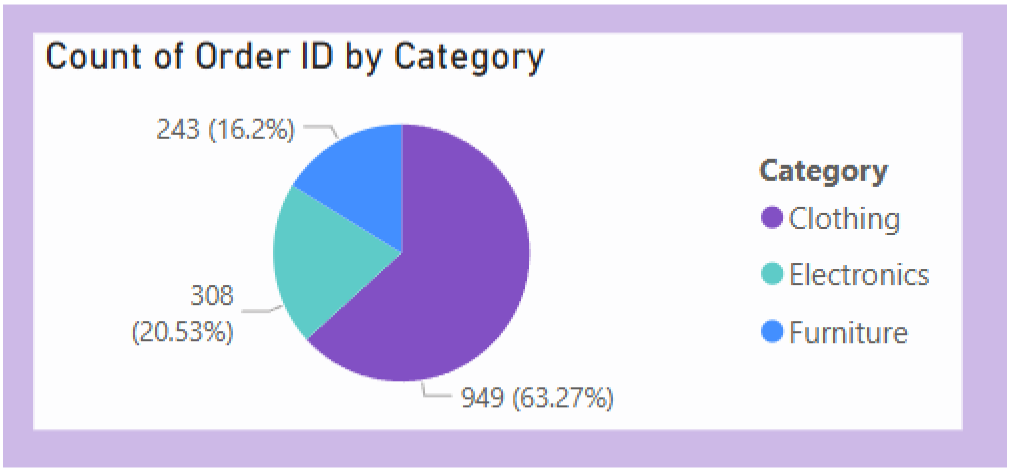
The top section of the dashboard features **4 KPI cards** providing high-level metrics:

KPI Metric	Value	Field Used
Total Orders	500	Count of Order ID
Total Customers	336	Count of CustomerName
Total Profit	37K	Sum of Profit
Total Sales Amount	438K	Sum of Amount

These cards provide a quick snapshot of the business performance.

- 500 orders were placed by 336 unique customers, indicating a customer repeat rate.
- The business generated ₹438,000 in revenue, yielding ₹37,000 in profit — about an 8.4% profit margin.

2. **Pie Chart – Order Distribution by Category**



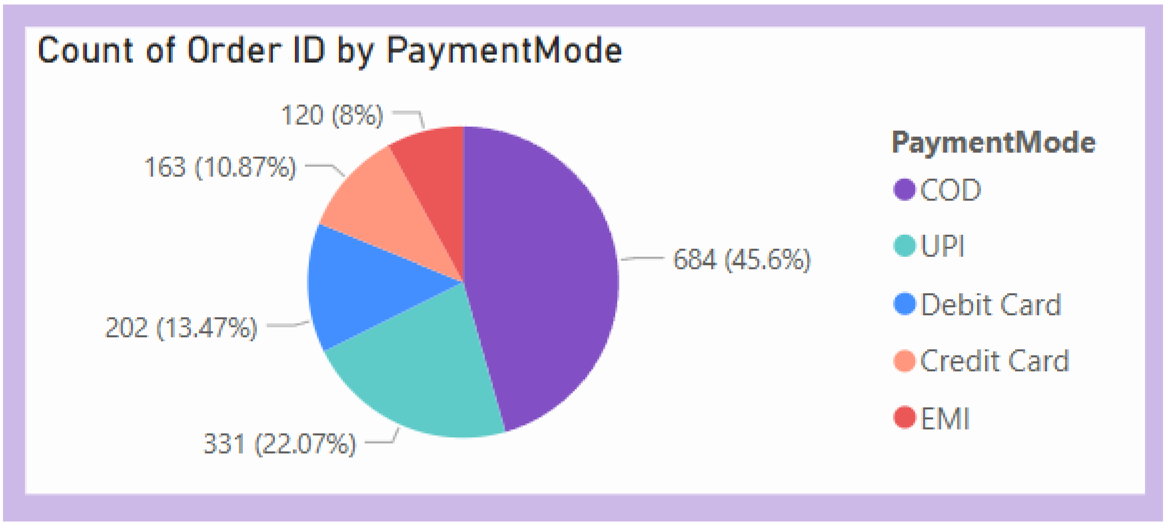
This **Pie Chart** breaks down the total number of orders based on **product categories**.

Category	**Count of Orders**	**% of Total Orders**
Clothing	949	63.27%
Electronics	308	20.53%
Furniture	243	16.2%

Purpose & Insights:

- Clothing dominates the order volume, contributing to over 63% of total orders, indicating it's the most popular product category.
- Electronics and Furniture are far behind, with 20% and 16% respectively — this insight can help guide stock/inventory or promotional focus.
- This chart helps stakeholders visualize product performance by category quickly and clearly.

3. Pie Chart – Order Distribution by Payment Mode



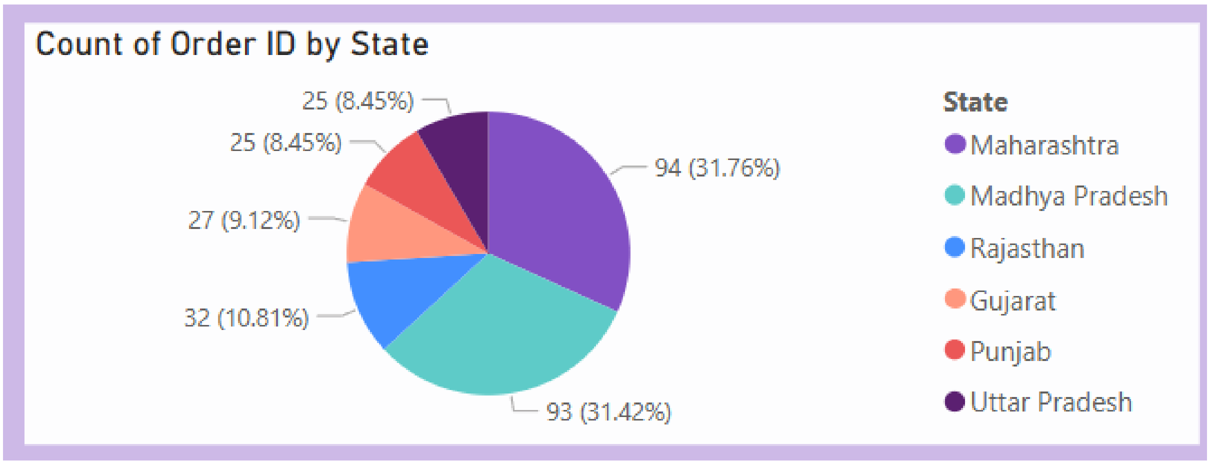
This pie chart shows the distribution of orders based on the payment method used by customers.

Payment Mode	Count of Orders	% of Total Orders
COD	684	45.6%
UPI	331	22.07%
Debit Card	202	13.47%
Credit Card	163	10.87%
EMI	120	8%

Purpose & Insights:

- Cash on Delivery (COD) is the most preferred payment method, used in nearly half of all orders.
- Digital payment methods (UPI, Debit Card, Credit Card) collectively make up a significant portion (~46.4%), indicating strong adoption of online transactions.
- EMI is the least used option, possibly due to product pricing or customer segment targeting.
- This chart can help in financial planning, policy decisions, or tailoring marketing strategies based on payment trends.

4. **Pie Chart – Count of Order ID by State**



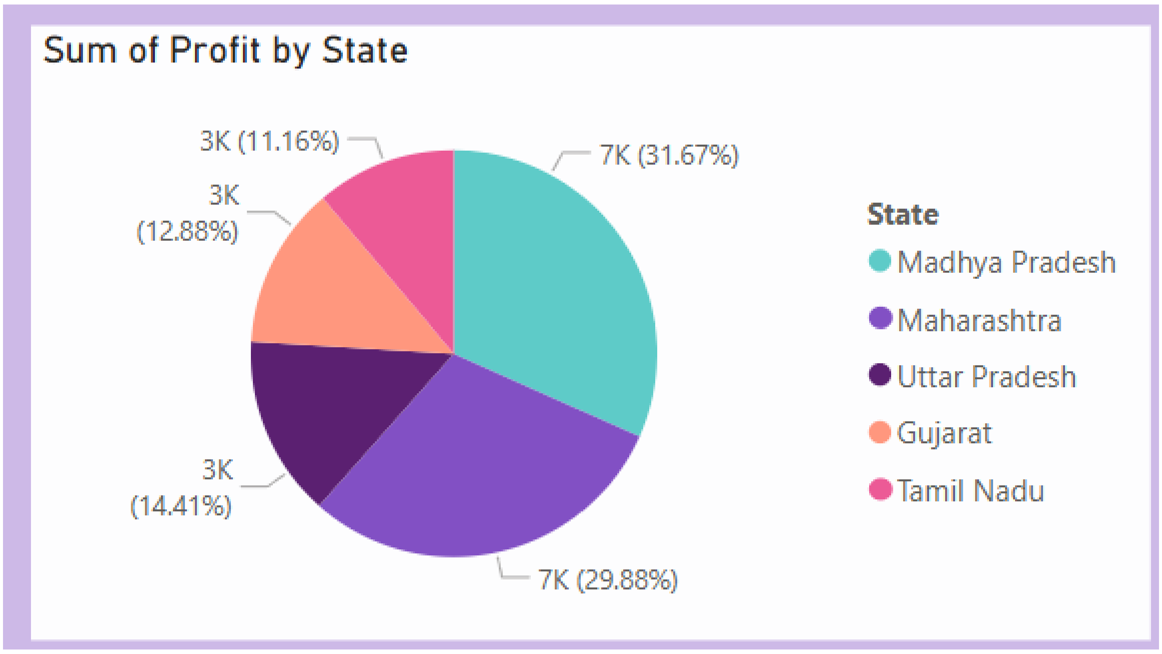
This pie chart displays the distribution of orders across different Indian states, showing the volume of orders placed per state.

State	**Count of Orders**	**% of Total Orders**
Maharashtra	94	31.76%
Madhya Pradesh	93	31.42%
Rajasthan	32	10.81%
Gujarat	27	9.12%
Punjab	25	8.45%
Uttar Pradesh	25	8.45%

Purpose & Insights:

- Maharashtra and Madhya Pradesh dominate the chart with over 63% of total orders, indicating these as major market zones.
- The rest of the states have moderate to low order activity, potentially signaling a need for increased marketing or regional strategy shifts.
- Businesses can use this insight for targeted sales efforts, inventory distribution, or logistics planning.

5. **Pie Chart – Sum of Profit by State**



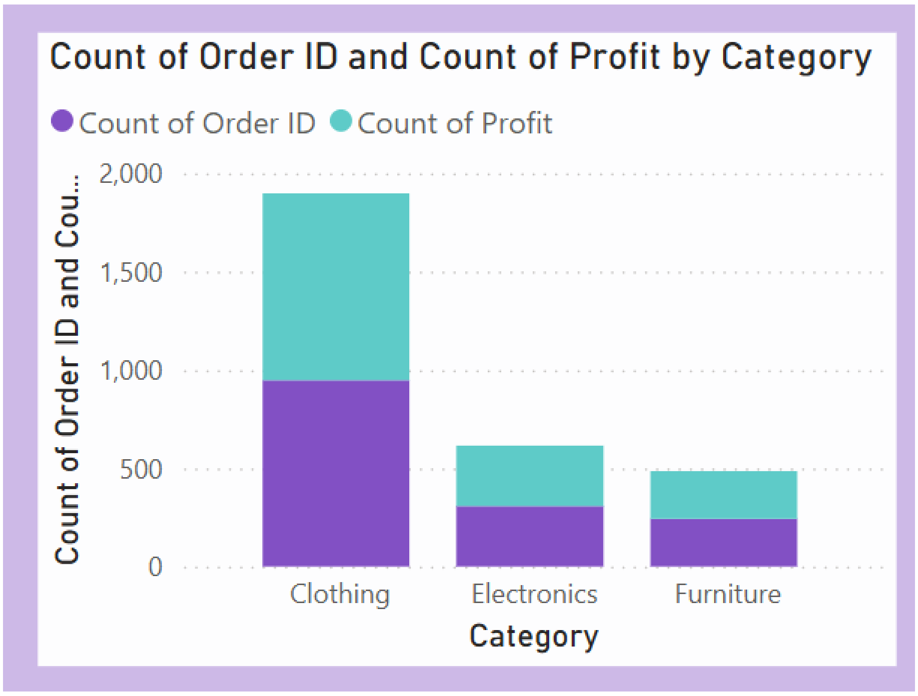
This chart presents the total profit generated from sales in various states, offering a clear view of which regions contribute most to the company's profitability.

State	**Sum of Profit**	**% of Total Profit**
Madhya Pradesh	₹7K	31.67%
Maharashtra	₹7K	29.88%
Uttar Pradesh	₹3K	14.41%
Gujarat	₹3K	12.88%
Tamil Nadu	₹3K	11.16%

Purpose & Insights:

- Madhya Pradesh and Maharashtra again take the lead, contributing to over 61% of total profits, aligning with their high order volumes.
- Uttar Pradesh, Gujarat, and Tamil Nadu generate moderate profits, potentially from fewer but higher-margin orders.
- Businesses can leverage this to focus on high-profit states for new product launches or optimize pricing and promotions in less profitable regions.

6. **Stacked Column Chart – Count of Order ID and Count of Profit by Category**



This chart compares the volume of orders and the number of profit instances across three product categories: Clothing, Electronics, and Furniture. The visual uses stacked columns where:

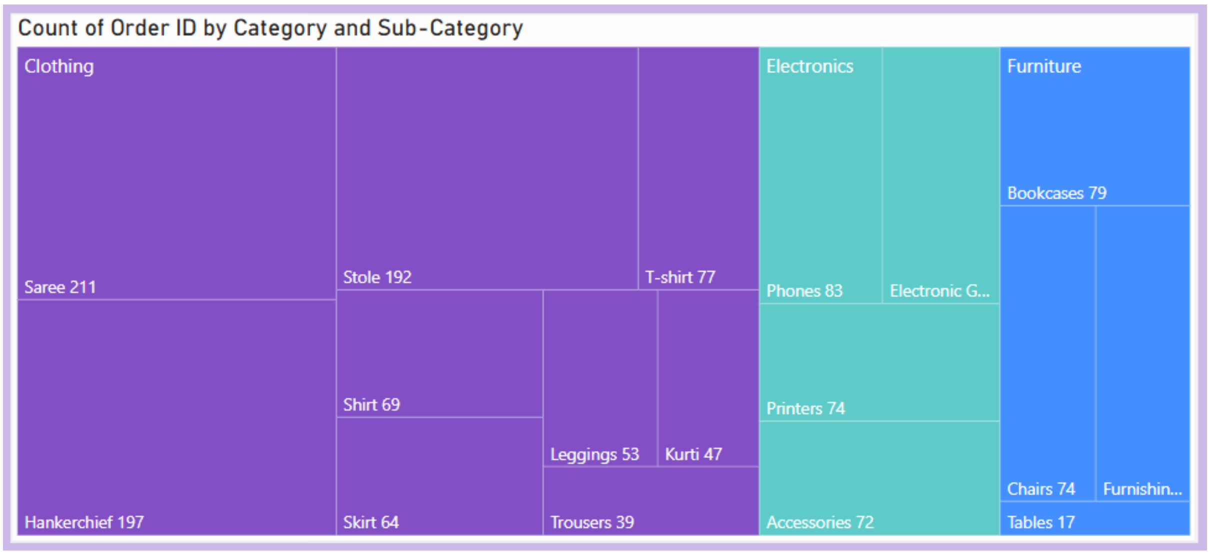
- Purple represents the Count of Order IDs
- Teal represents the Count of Profit records

Category	**Order Count**	**Profit Count**
Clothing	~950	~950
Electronics	~300	~300
Furniture	~240	~240

Purpose & Insights:

- Clothing dominates in both order volume and profit records, indicating it is the most popular and profitable category.
- Electronics and Furniture have significantly lower values, suggesting less engagement or possibly higher costs affecting profitability.
- This helps stakeholders prioritize inventory, marketing, and promotional strategies around the most responsive category.

7. Treemap – Count of Order ID by Category and Sub-Category



This treemap visualizes the distribution of order counts across various sub-categories within the main categories: Clothing, Electronics, and Furniture. Each rectangle represents a sub-category, sized proportionally based on the number of orders.

Category	**Top Sub-Categories by Order Count**
Clothing	Saree (211), Handkerchief (197), Stole (192)
Electronics	Phones (83), Printers (74), Accessories (72)
Furniture	Bookcases (79), Chairs (74), Furnishings (53)

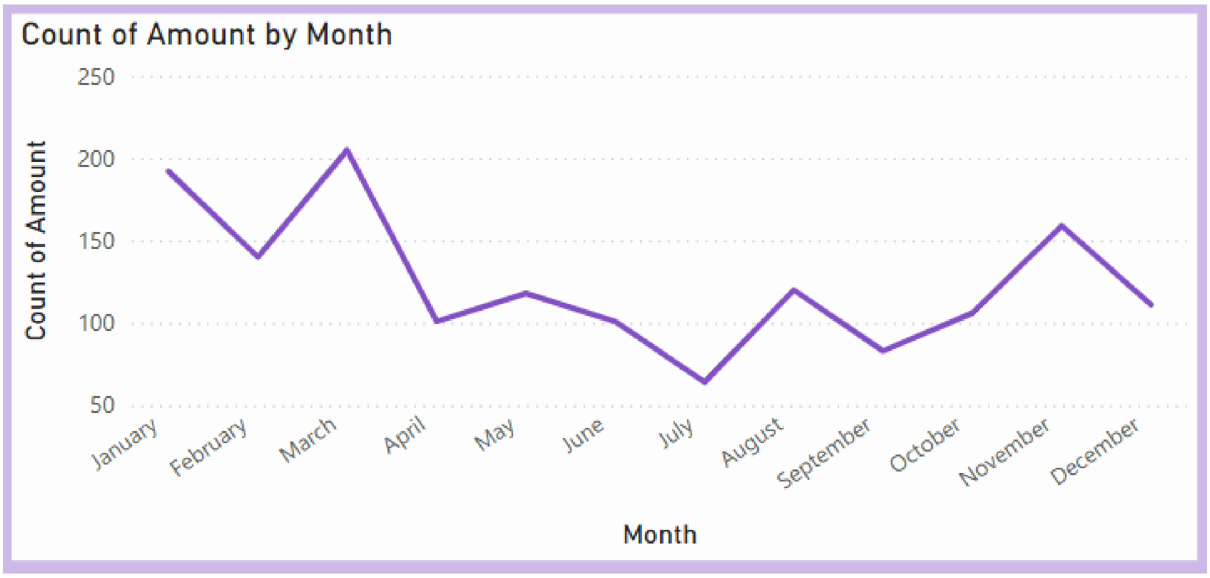
Purpose & Insights:

- Clothing dominates this treemap with the highest number of sub-categories and order volume.
- Within clothing, Saree, Handkerchief, and Stole are the leading products, indicating high customer interest and demand.
- In Electronics, Phones lead the count, followed closely by Printers and Accessories.
- Furniture has fewer orders overall, with Bookcases and Chairs being the most requested.

This chart helps decision-makers identify:

- Top-selling sub-categories worth investing in.
- Low-performing items where marketing or discounts could boost sales.
- Areas where product diversity might need to expand or be consolidated.

8. **Treemap–Count of Order ID by Category and Sub-Category**



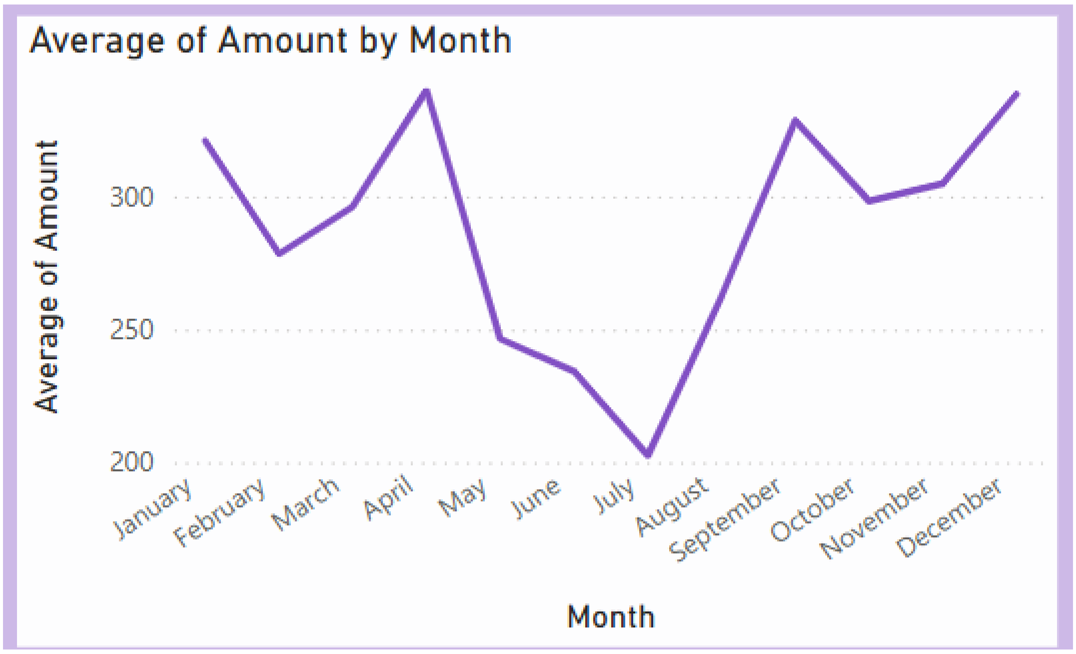
This chart presents the total count of amount by month, showcasing the fluctuation in transaction volume across the year. It helps identify which months experience higher or lower activity, guiding decision-making in terms of demand cycles and resource allocation.

Month	**Count of Amount (approx.)**
January	190
February	140
March	210
April	100
May	120
June	105
July	65
August	120
September	85
October	105
November	160
December	115

Purpose & Insights:

- March and January see the highest activity, with March peaking at over 210 transactions. This could indicate seasonal demand or marketing effectiveness during Q1.
- July records the lowest count, suggesting a potential dip in business activity during mid-year. This might be due to seasonal slowdowns or market behavior shifts.
- November shows a significant rise post-October, possibly due to festive season campaigns or end-of-year sales pushes.
- Fluctuations across the year suggest the need for dynamic inventory and workforce planning, especially during high-activity months.
- Businesses can utilize these insights to align marketing and sales efforts with demand cycles, ensuring optimized resource use and maximized return during peak periods.

9. Treemap–Count of Order ID by Category and Sub-Category



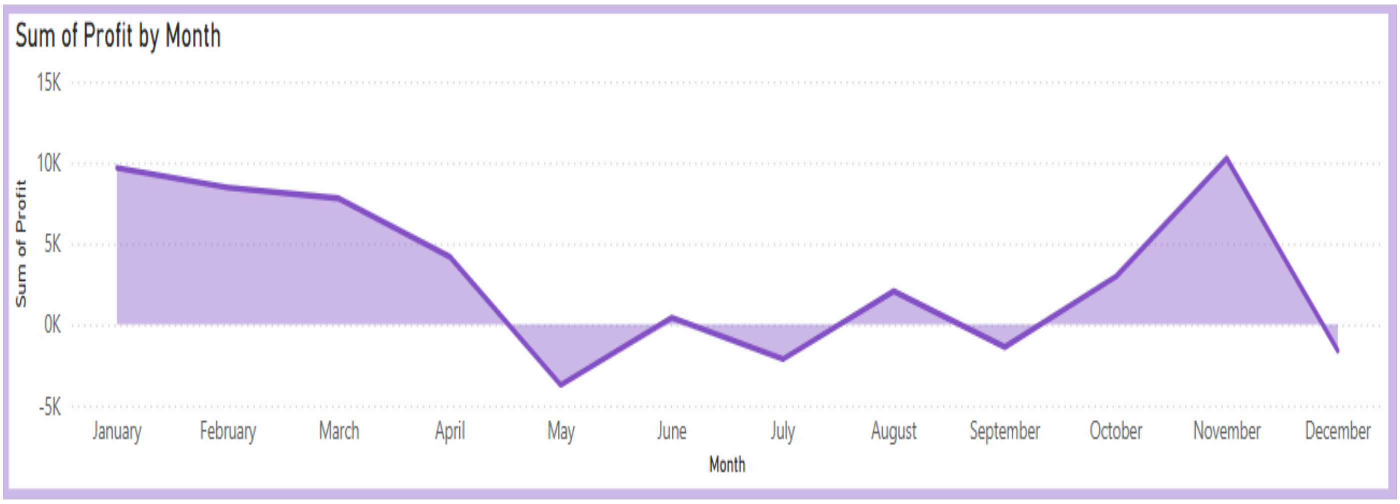
This chart displays the average transaction amount for each month, offering a clear view of monthly value trends and indicating periods of higher or lower-value transactions.

Month	**Average Amount (approx.)**
January	320
February	280
March	295
April	330
May	340
June	245
July	230
August	200
September	280
October	330
November	300
December	340

Purpose & Insights:

- May and December record the highest average amounts (around ₹340), suggesting these months bring in high-value transactions—possibly due to seasonal sales or bulk purchases.
- August has the lowest average amount (~₹200), pointing to either smaller transactions or discount-driven behavior during this period.
- October also shows strong performance, rebounding sharply from August, likely aligning with festive or promotional periods.
- A clear mid-year dip from June to August indicates a phase of lower-value activity, potentially signaling the need for promotional strategies or campaign boosts during that time.
- Businesses can use this insight to plan for premium product launches or targeted upselling in high average-amount months like May, October, and December, and explore ways to lift transaction values during low-avg months.

10. **Treemap–Count of Order ID by Category and Sub-Category**



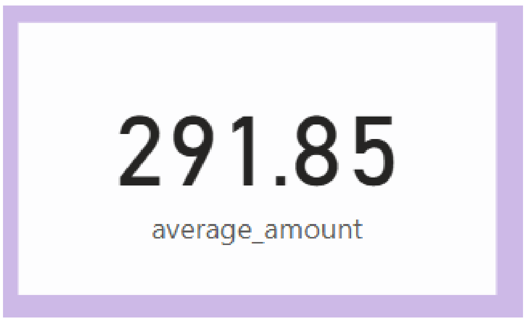
This area chart illustrates the total profit generated per month, highlighting fluctuations in profitability across the year. It visually indicates not only high and low-profit periods but also months with negative profit (losses).

Month	Sum of Profit (approx.)
January	₹9K
February	₹8K
March	₹7K
April	₹4K
May	-₹4K
June	₹1K
July	-₹2K
August	₹2K
September	-₹1K
October	₹3K
November	₹9K
December	-₹3K

Purpose & Insights:

- Strongest profit months: January and November lead with profits close to ₹9K, indicating peak revenue periods that may align with seasonal or promotional spikes.
- Significant losses: May and December show sharp dips into negative territory, with losses nearing ₹4K and ₹3K respectively—suggesting poor performance or high expenditure during these months.
- Mid-year volatility: From May to September, profit fluctuates between minor gains and losses, indicating an unstable phase that may benefit from cost control or operational optimization.
- Notably, despite higher average amounts in May and December (as per the previous chart), the profit dips, possibly due to high discounts, operational costs, or returns during these periods.
- Businesses should explore reasons for losses in high-average-value months and capitalize on high-profit periods like January and November by replicating successful strategies.

11. **KPI Cards**



These cards highlight the **key statistics** related to transaction amounts, helping to quickly understand the scale and variability in the dataset.

Metric	Value
Average Amount	₹291.85
Highest Amount	₹5,729
Lowest Amount	₹4

Purpose & Insights:

- The average transaction amount stands at ₹291.85, indicating the typical order size across all months.
- The highest transaction recorded is ₹5,729, which is significantly above average—pointing to occasional large-value purchases that could be from bulk or premium buyers.
- The lowest amount is just ₹4, suggesting that some transactions are very small, potentially due to micro-purchases, test transactions, or heavy discounting.
- The wide range (₹4 – ₹5,729) reveals a high variability in transaction values. This emphasizes the need to analyze customer segmentation and pricing strategies to balance volume with profitability.
- When combined with earlier insights (e.g., average amount trend by month), this further supports the idea of targeting high-value transaction periods and managing low-profit months effectively.