

# Project Name:

## DMart Sales

---

### Description:

This Power BI project analyzes DMart's sales, customer demographics, product performance, operational efficiency, and payment trends. It brings raw sales and customer data into Power BI, cleans and shapes it using Power Query, builds a star schema data model, and creates interactive visuals to answer key business questions.

---

### Purpose / Problem Statement:

DMart needs an easy-to-use dashboard to:

- Track monthly and yearly sales trends
- Understand customer behavior by age, gender, and engagement (time on website, clicks)
- Identify best-selling products and impact of discounts
- Monitor shipping performance and cancellations
- Analyze payment method preferences and order status patterns

Without this, analysis is slow and scattered across multiple files, making decision-making harder.

---

### Scope:

- Importing and cleaning DMart CSV dataset in Power BI
- Handling missing and incorrect values (e.g., replacing "Cancelled" text in numeric columns with nulls)
- Creating calculated columns (Discount %, Delivery Days, Month-Year, Status Flags, Time Spent Buckets)
- Building a star schema with fact and dimension tables
- Creating DAX measures for KPIs (Total Sales, Orders, AOV, Cancellation Rate, etc.)
- Building visuals to answer business questions
- Applying filters (e.g., Completed orders only for certain charts)

- Formatting visuals for clear, interactive dashboards
- Integration with external APIs

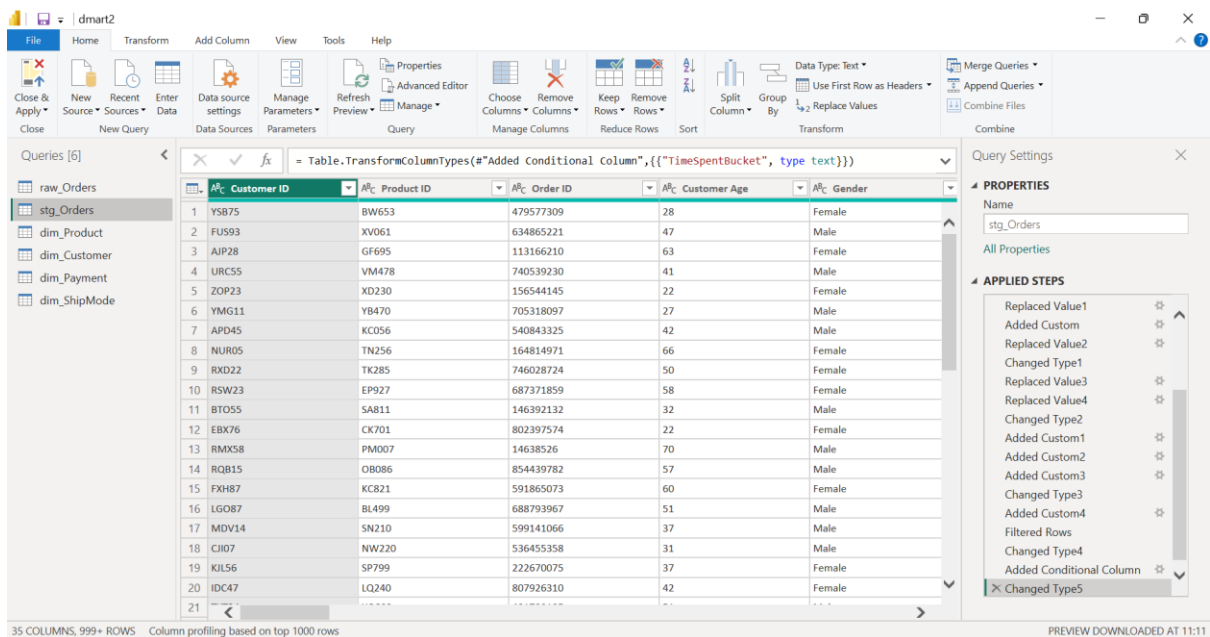
## Steps Performed:

### 1. Data Import

- Loaded Dmart+Dataset\_Tableau.csv into Power BI.

### 2. Data Cleaning in Power Query

- Removed extra spaces, fixed text case.
- Replaced text values ("Cancelled", "Returned") in numeric columns with nulls.
- Converted data types (currency to decimal, dates to date format).



### 3. Data Transformation

- Added calculated columns:
  - DiscountPercent
 
$$\text{if } [\text{MRP}] \neq \text{null and } [\text{MRP}] \neq 0$$

$$\text{then } ([\text{MRP}] - [\text{Discount Price}]) / [\text{MRP}]$$

$$\text{else null}$$
  - DiscountAmount
 
$$\text{if } [\text{MRP}] \neq \text{null and } [\text{Discount Price}] \neq \text{null}$$

then [MRP] - [Discount Price]

else null

- DeliveryDays (Delivery Date – Order Date)

if [Order Date] <> null and [Delivery Date] <> null

then Duration.Days([Delivery Date] - [Order Date])

else null

- Order\_MonthYear

Text.From([Year]) & "-" &

Text.PadStart(Text.From([Month]), 2, "0")

- OrderStatusFlag (Completed / Cancelled / Returned)

if [Total Order Value] = null then "Check Status"

else if Text.Upper([Order Status]) = "CANCELLED" then  
"Cancelled"

else if Text.Upper([Order Status]) = "RETURNED" then  
"Returned"

else "Completed"

- TimeSpentBucket (grouped time on website into ranges)

- Created dimension tables for products, customers, payment, and shipping.

#### 4. Data Modeling

- Connect dimensions to fact table

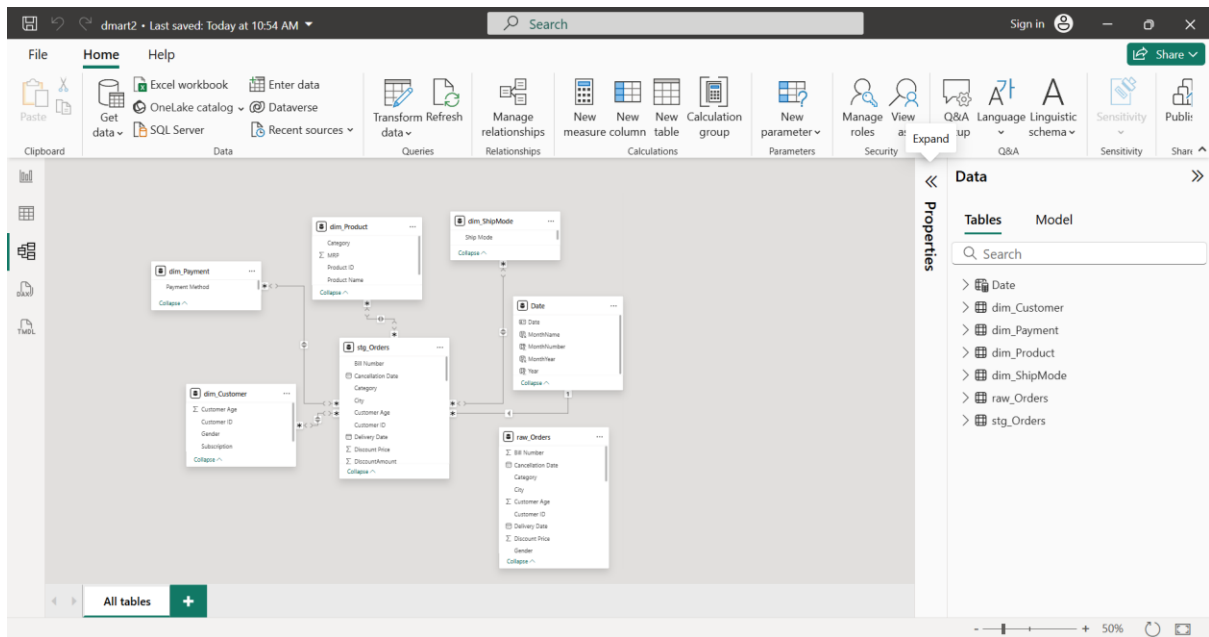
- dim\_Product[Product ID] → stg\_Orders[Product ID]

- dim\_Customer[Customer ID] → stg\_Orders[Customer ID]

- dim\_Payment[Payment Method] → stg\_Orders[Payment Method]

- dim\_ShipMode[Ship Mode] → stg\_Orders[Ship Mode]

- Built a **star schema** linking fact table to dimension tables.



- Added a **Date Table** for time-based analysis.

Date = CALENDAR(DATE(2021,1,1), DATE(2023,12,31))

Add columns in the Date table:

Year = YEAR('Date'[Date])

MonthNumber = MONTH('Date'[Date])

MonthName = FORMAT('Date'[Date], "MMMM")

MonthYear = FORMAT('Date'[Date], "YYYY-MM")

Date	Year	MonthNumber	MonthName	MonthYear
01-01-2021 00:00:00	2021	1	January	2021-01
02-01-2021 00:00:00	2021	1	January	2021-01
03-01-2021 00:00:00	2021	1	January	2021-01
04-01-2021 00:00:00	2021	1	January	2021-01
05-01-2021 00:00:00	2021	1	January	2021-01
06-01-2021 00:00:00	2021	1	January	2021-01
07-01-2021 00:00:00	2021	1	January	2021-01
08-01-2021 00:00:00	2021	1	January	2021-01
09-01-2021 00:00:00	2021	1	January	2021-01
10-01-2021 00:00:00	2021	1	January	2021-01
11-01-2021 00:00:00	2021	1	January	2021-01
12-01-2021 00:00:00	2021	1	January	2021-01
13-01-2021 00:00:00	2021	1	January	2021-01
14-01-2021 00:00:00	2021	1	January	2021-01
15-01-2021 00:00:00	2021	1	January	2021-01
16-01-2021 00:00:00	2021	1	January	2021-01
17-01-2021 00:00:00	2021	1	January	2021-01
18-01-2021 00:00:00	2021	1	January	2021-01
19-01-2021 00:00:00	2021	1	January	2021-01
20-01-2021 00:00:00	2021	1	January	2021-01
21-01-2021 00:00:00	2021	1	January	2021-01

Table: Date (1,095 rows)

## 5. DAX Measures

- Total Sales, Total Orders, Average Order Value, Total Discount Amount

- Cancellation Rate, Repeat Customer Rate, Repeat Customer Count
- 1. Total Sales = SUM(stg\_Orders[Total Order Value])
- 2. Total Orders = DISTINCTCOUNT(stg\_Orders[Order ID])
- 3. Average Order Value = DIVIDE([Total Sales], [Total Orders], 0)
- 4. Total Discount Amount = SUM(stg\_Orders[DiscountAmount])
- 5. Cancellation Rate  
=DIVIDE(CALCULATE(DISTINCTCOUNT(stg\_Orders[Order ID]),  
stg\_Orders[Order Status] = "Cancelled"), [Total Orders], 0)
- 6. Repeat Customer Count =  
COUNTROWS(FILTER(VALUES(stg\_Orders[Customer  
ID]),CALCULATE(COUNTROWS(stg\_Orders)) > 1))
- 7. Repeat Customer Rate = DIVIDE([Repeat Customer Count],  
DISTINCTCOUNT(stg\_Orders[Customer ID]), 0)

## 6. Visualization

The dashboard contains multiple interactive pages and visuals answering the key business questions.

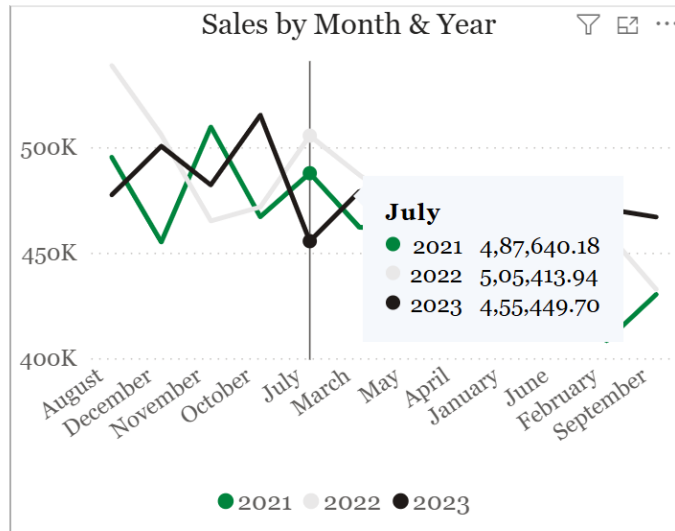
Each visual is designed for quick interpretation and filtering by year, category, state, and payment mode.

---

## 1. Sales Performance

- **Sales by Month & Year** (Line/Column chart)

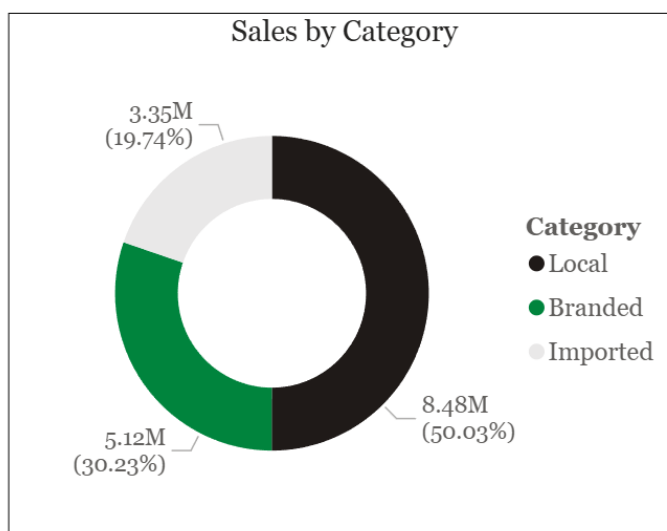
Displays monthly sales trends for 2021, 2022, and 2023, allowing year-over-year and seasonal comparisons.



- **Sales by Category** (Donut chart)

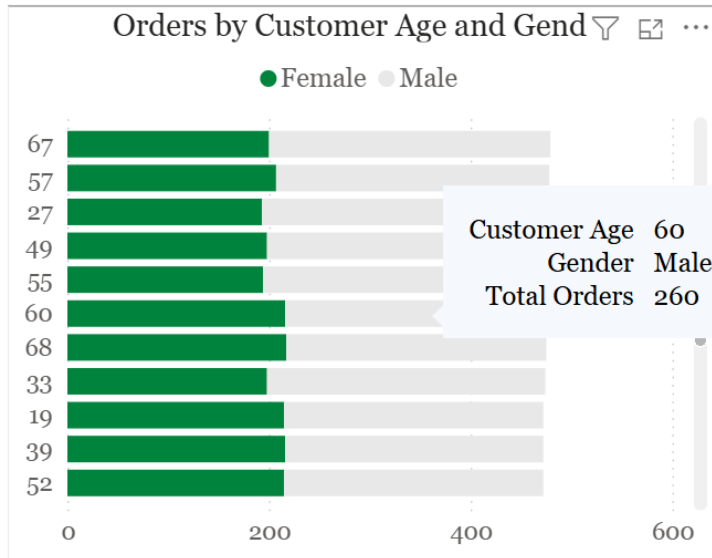
Shows revenue contribution by category:

- Local (50.03%)
- Branded (30.23%)
- Imported (19.74%)

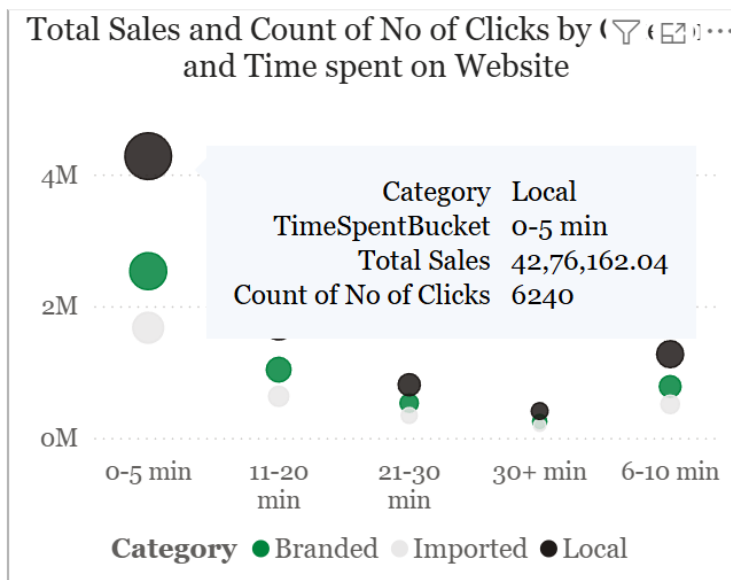


## 2. Customer Demographics & Behavior

- **Orders by Customer Age and Gender** (Clustered bar chart)  
Breaks down order counts by customer age groups for male and female customers.

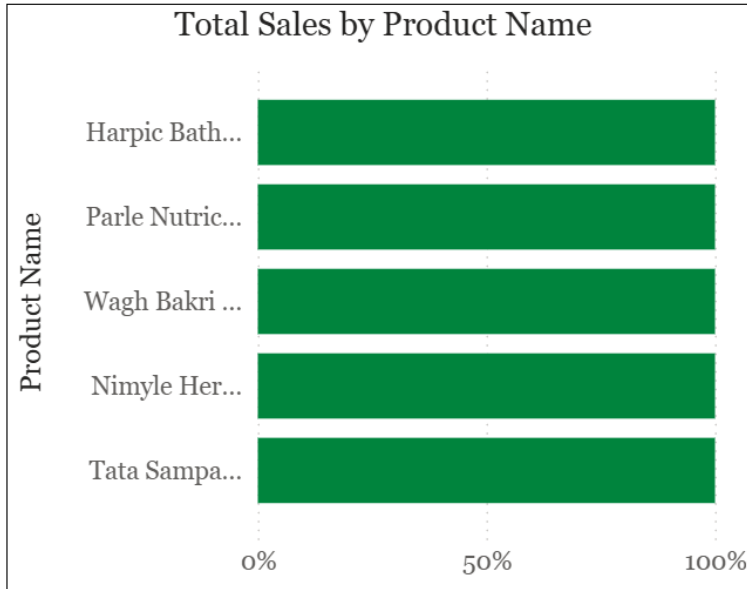


- **Total Sales & No. of Clicks by Category and Time Spent on Website** (Clustered column with legend)  
Analyzes how engagement (time spent online and number of clicks) correlates with sales across Branded, Imported, and Local products.

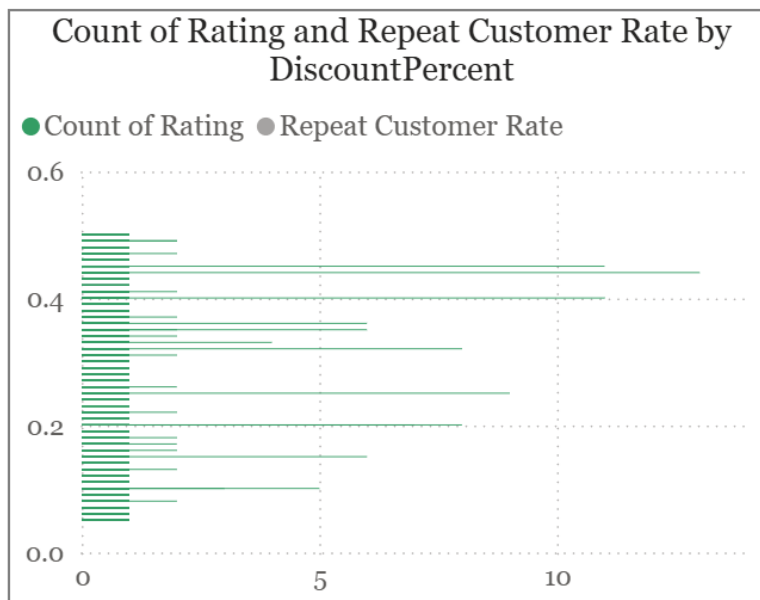


### 3. Product Insights

- **Total Sales by Product Name** (Stacked bar chart)  
Ranks products such as Harpic Bathroom Cleaner, Parle Nutricrunch, Wagh Bakri Tea, Nimyle Herbal, and Tata Sampann based on sales share.



- **Count of Rating & Repeat Customer Rate by Discount Percent** (Dual-axis combo chart)  
Compares the number of ratings and repeat purchase rate across varying discount levels.



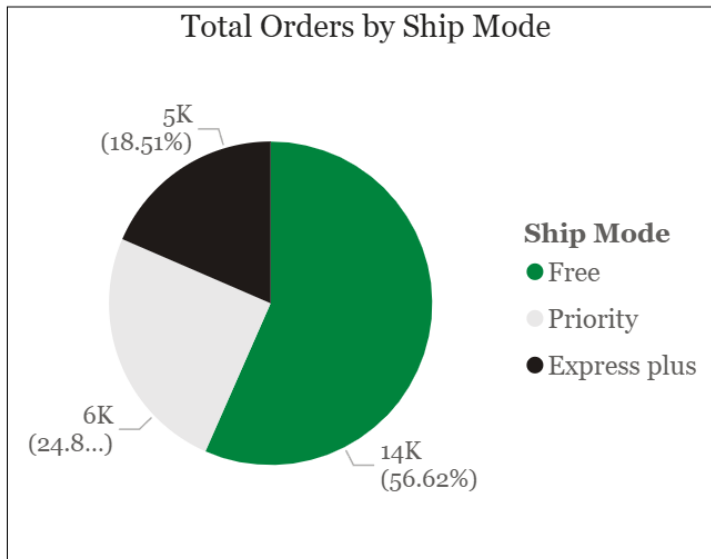


#### 4. Operational Performance

- **Total Orders by Ship Mode** (Donut chart)

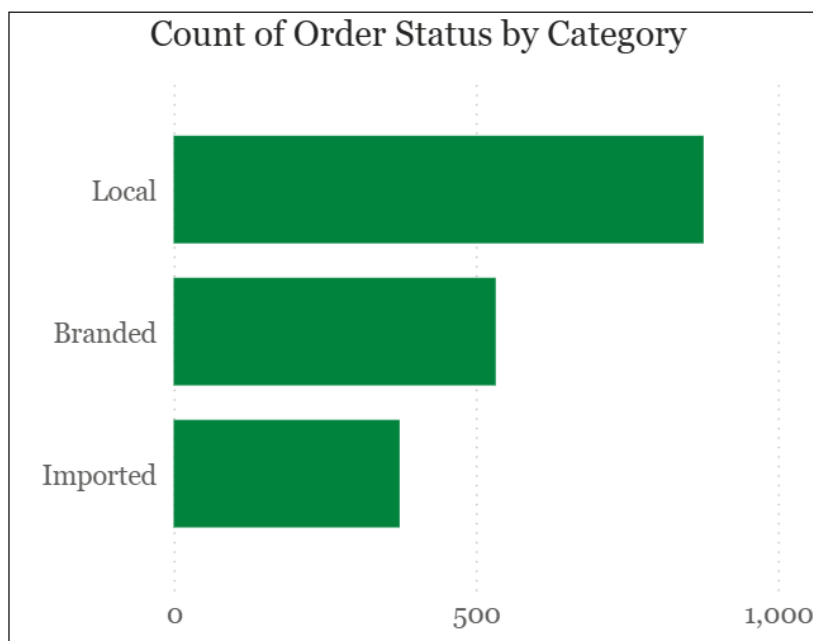
Distribution of orders by shipping type:

- Free (56.62%)
- Priority (24.8%)
- Express Plus (18.51%)



- **Count of Order Status by Category** (Clustered column chart)

Shows completed, shipped, and cancelled order counts for Local, Branded, and Imported categories.

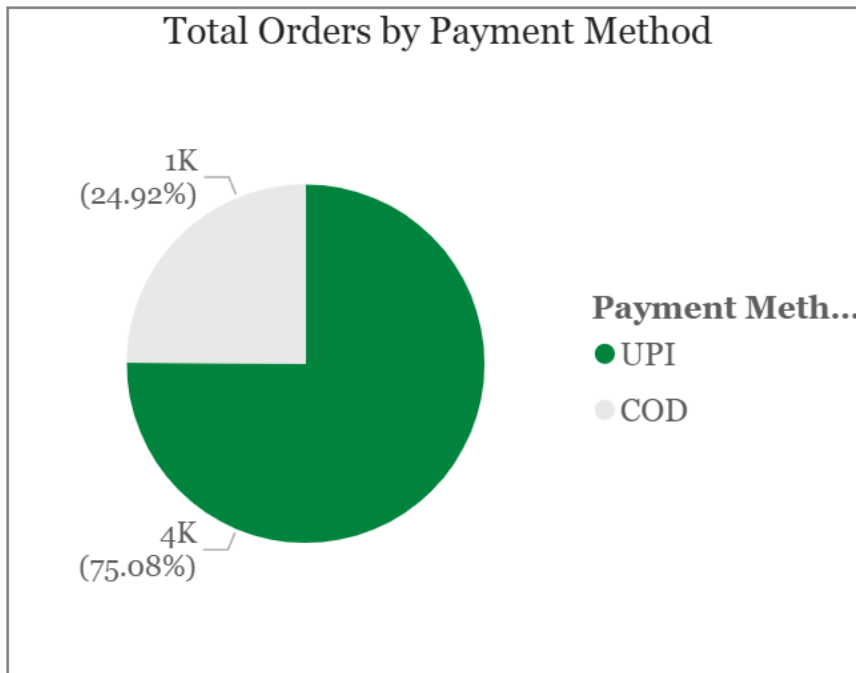


## 5. Payments & Order Status

- **Total Orders by Payment Method** (Donut chart)

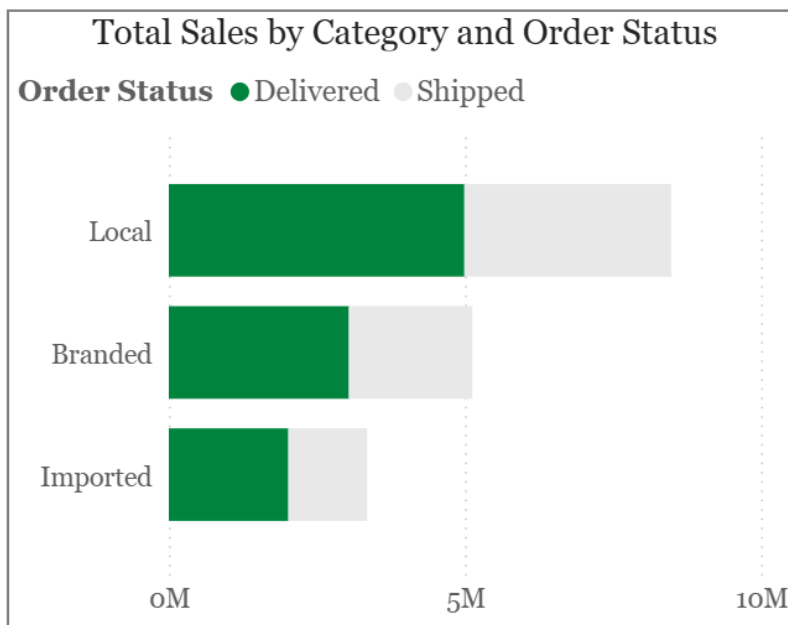
Payment method breakdown:

- UPI (75.08%)
- COD (24.92%)



- **Total Sales by Category and Order Status** (Stacked bar chart)

Compares sales amounts across product categories, segmented by order status (Delivered, Shipped).



## 6. Interactive Filters

Slicers allow real-time filtering of all visuals by:

- **Year:** 2021, 2022, 2023



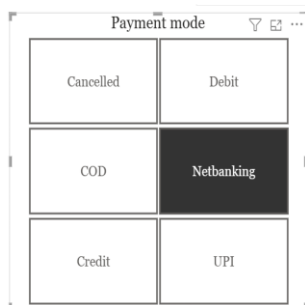
- **Category:** Branded, Imported, Local



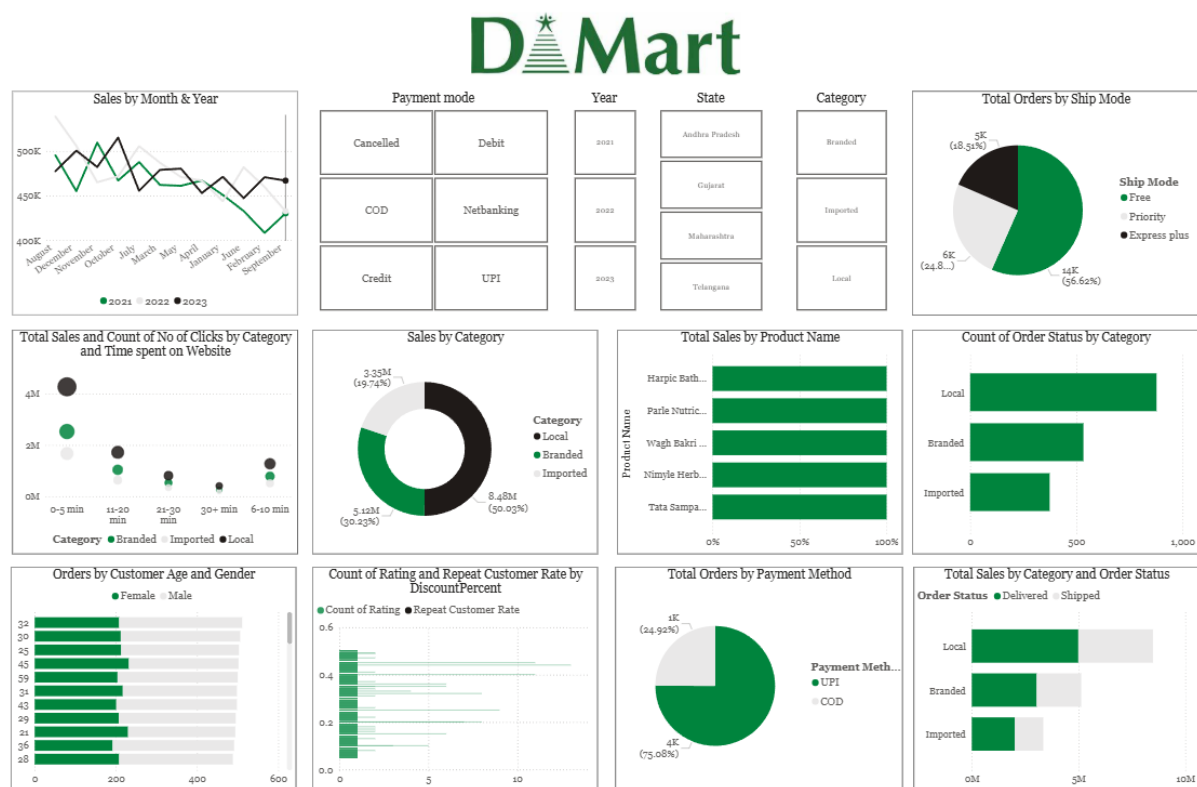
- **State:** Andhra Pradesh, Gujarat, Maharashtra, Telangana



- **Payment Mode:** Cancelled, COD, Credit, Debit, Netbanking, UPI



## 7. Dashboard



## Conclusion

This project successfully transformed DMart's raw sales dataset into a clear, interactive Power BI dashboard.

It provides management with quick insights into sales performance, customer behaviour, product trends, operational metrics, and payment patterns.

The dashboard is easy to filter, supports data-driven decision-making, and is structured for future expansion, such as live data updates and advanced analytics.