**InnoByte Services**

Data analyst Intern project

Name : Katari Pavan

Batch : 01-Nov-2024 to 30-Nov-2024

Mail : [pavankatari24@gmail.com](mailto:pavankatari24@gmail.com)

|  |  |
| --- | --- |
| Data set | Amazon sales data |
| Records | 128974 |
| Data variables | 21 |

Tools used : MS. Excel, MySQL, Tableau

About the Project :

This is the project given by innobyte services for data analyst interns as part of Internship. For completing this project, I used data analysis tools such as MS. Excel, MySQL , Tableau. This document will showcase my work that I done in completing this internship successfully.

What I done :

First, I cleaned the data using MS. Excel – removed data duplicates, Checked data types consistency, Deleted columns having missing values >30%, turned numeric values into structured format.

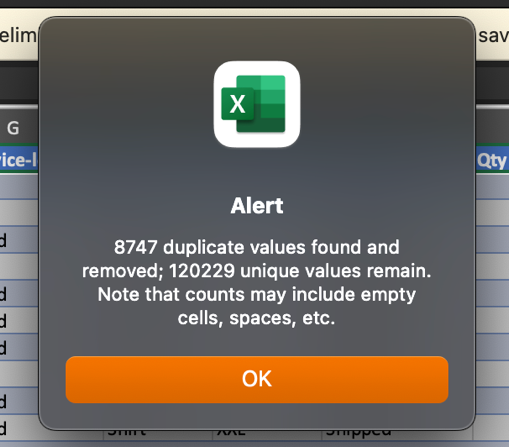
Imported all the data into Tableau and analyzed the data

For data visualization I used business intelligence tool Tableau, I developed relevant charts, graphs and Dashboards. Also mentioned all the key findings, insights, Suggestions.

Data cleaning :

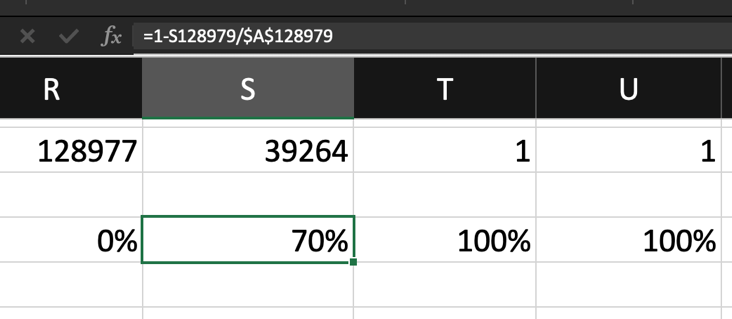
* Removed all the data duplicates present in dataset using MS. Excel, ( **Order ID** is Primary key ).

Total duplicate values : 8747



* Removed the data variables those having missing values percentage greater than 30%.

(fulfilled-by = 70%, New = 100%, Pendings = 100%)



* Converted the data types those are in wrong data type i.e., Date.
* Done all necessary changes in ship-state,

Ex. : NL 🡪 Nagaland, RJ 🡪 Rajasthan

**AFTER DATA CLEANING :**

|  |  |
| --- | --- |
| Data set | Amazon sales data |
| Records | 17 |
| Data variables | 120204 |

**Sales Overview :**

Total Sales :



Total Quantity(Units) :



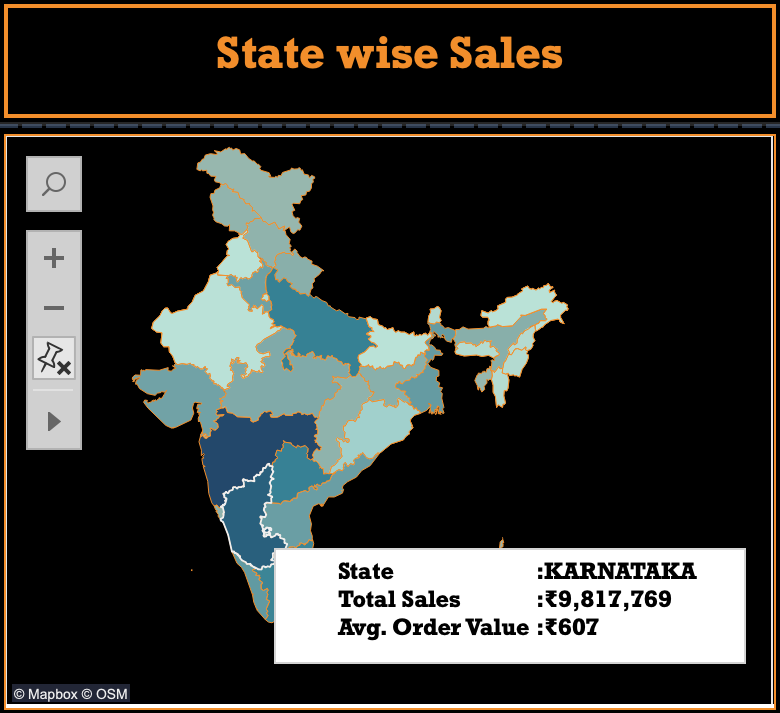
Average Order Value :



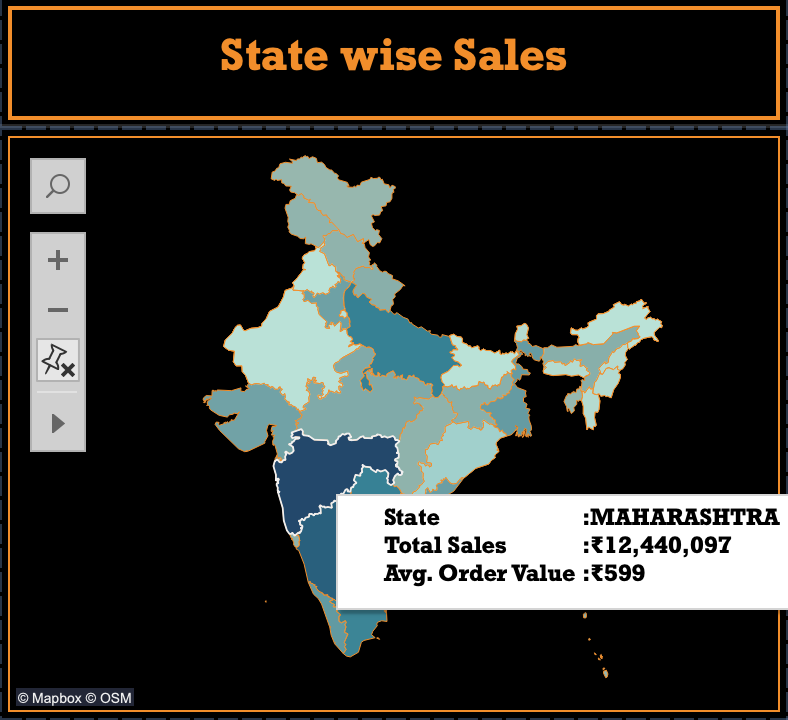
Total Orders :



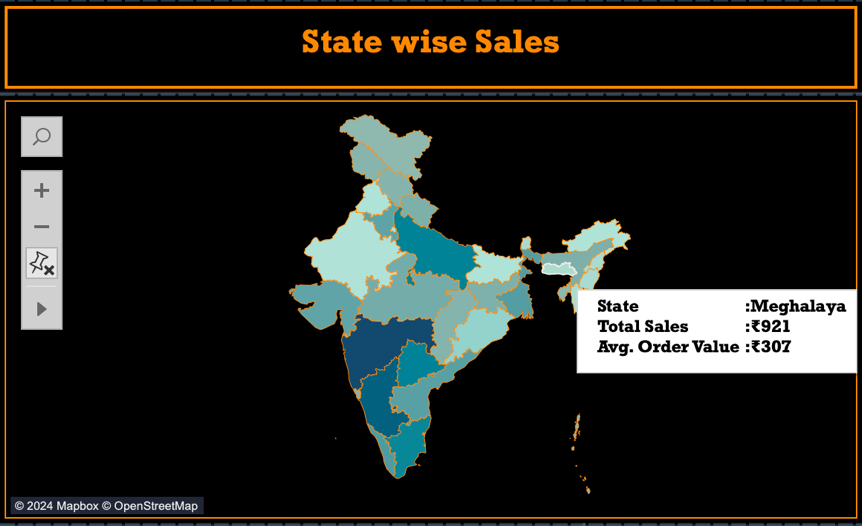
State wise Sales :



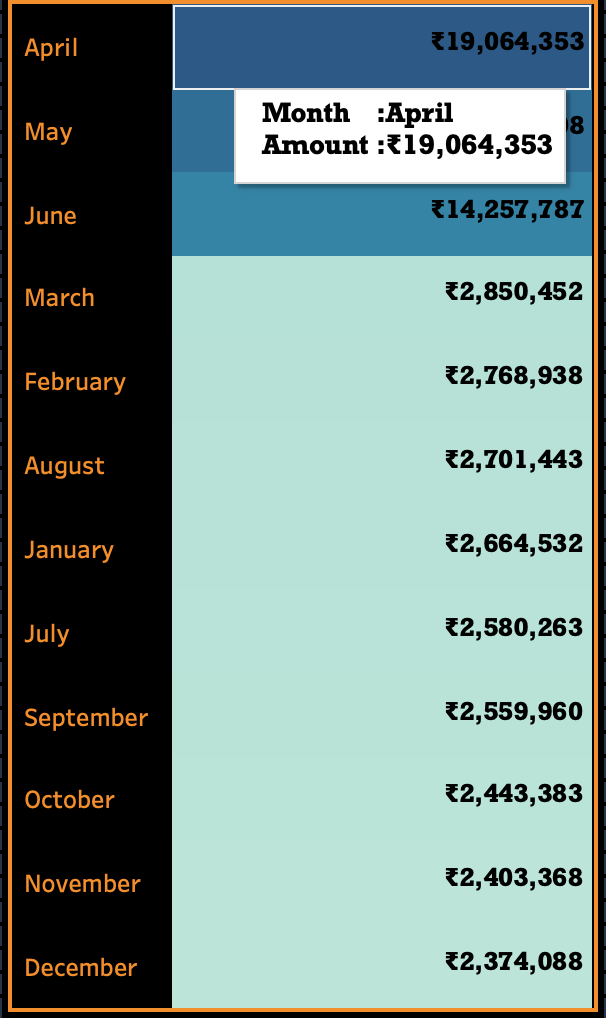
State with High Sales :



State with Low Sales :



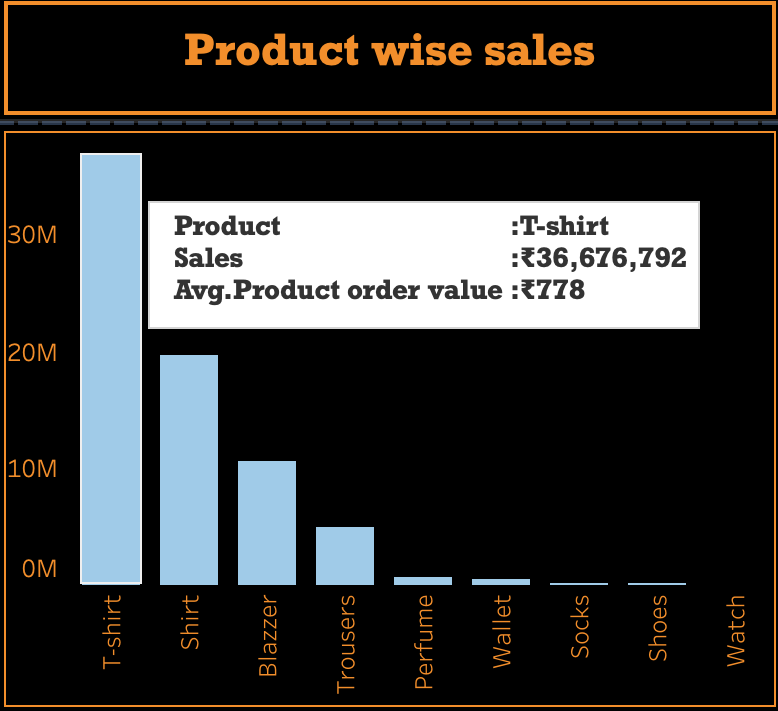
Month with high sales :



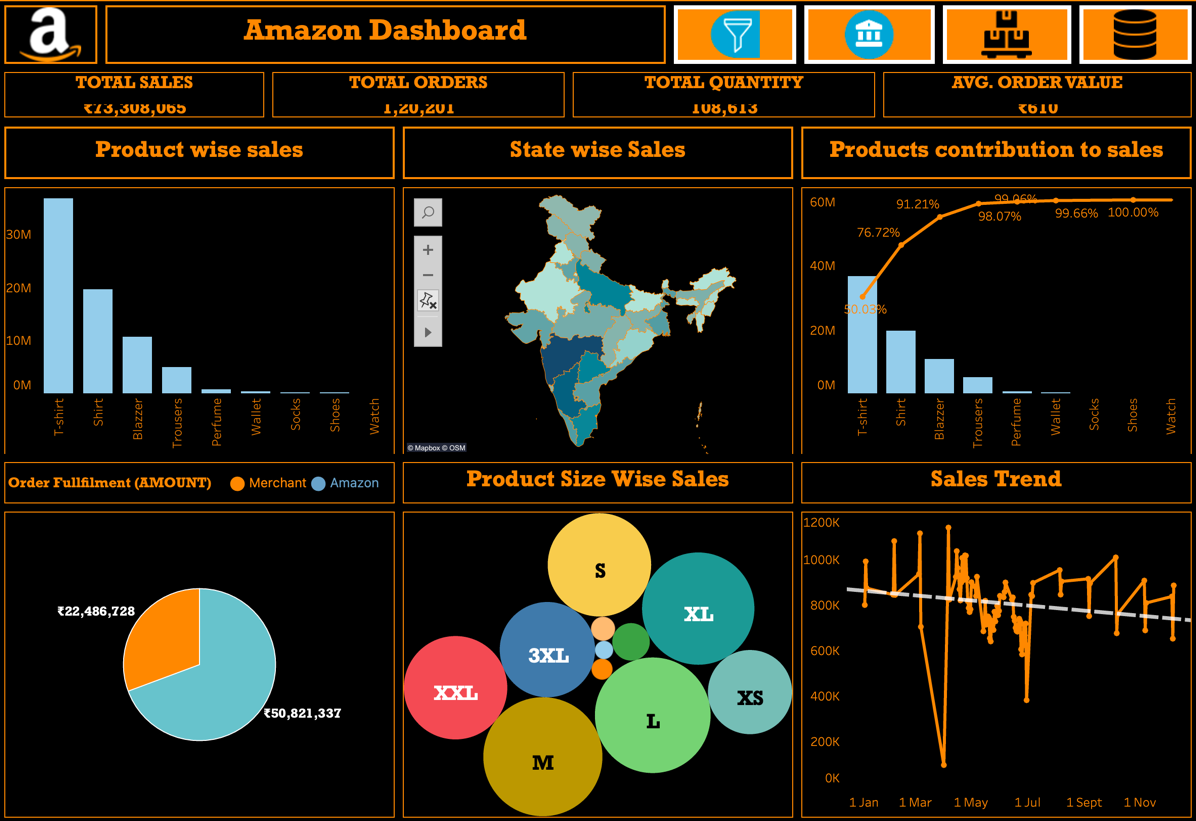
Month with low sales :



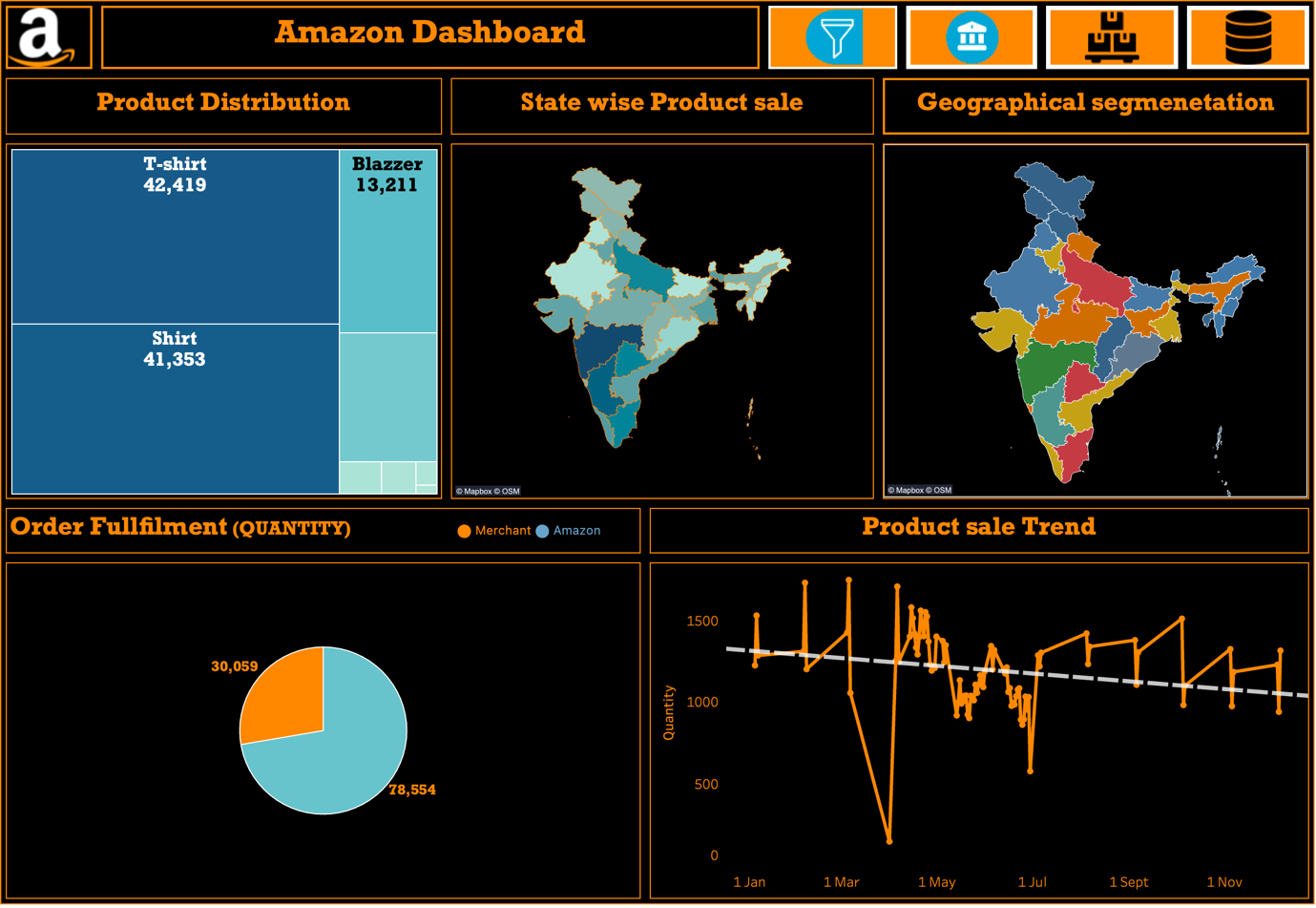
Product wise sales :



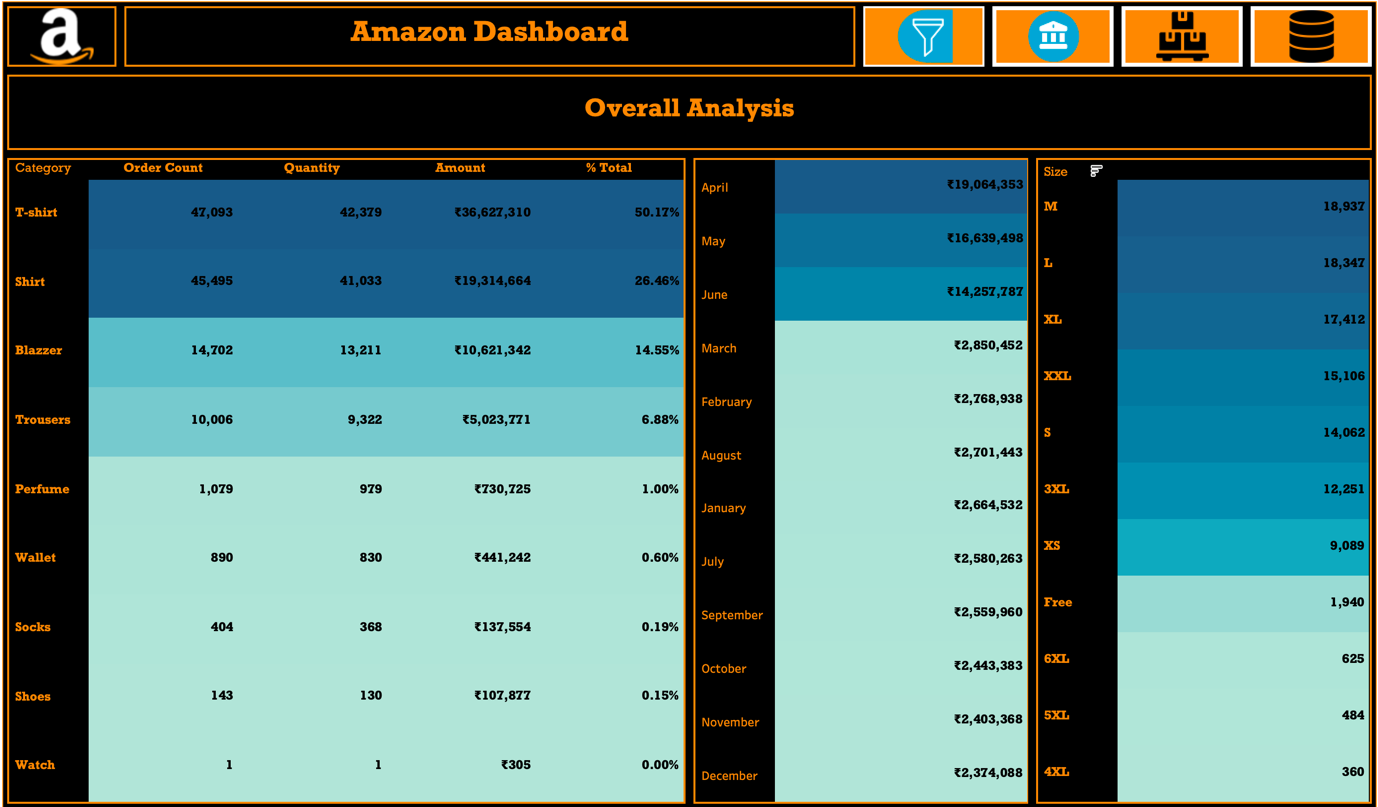
Sales Analysis Dashboard :

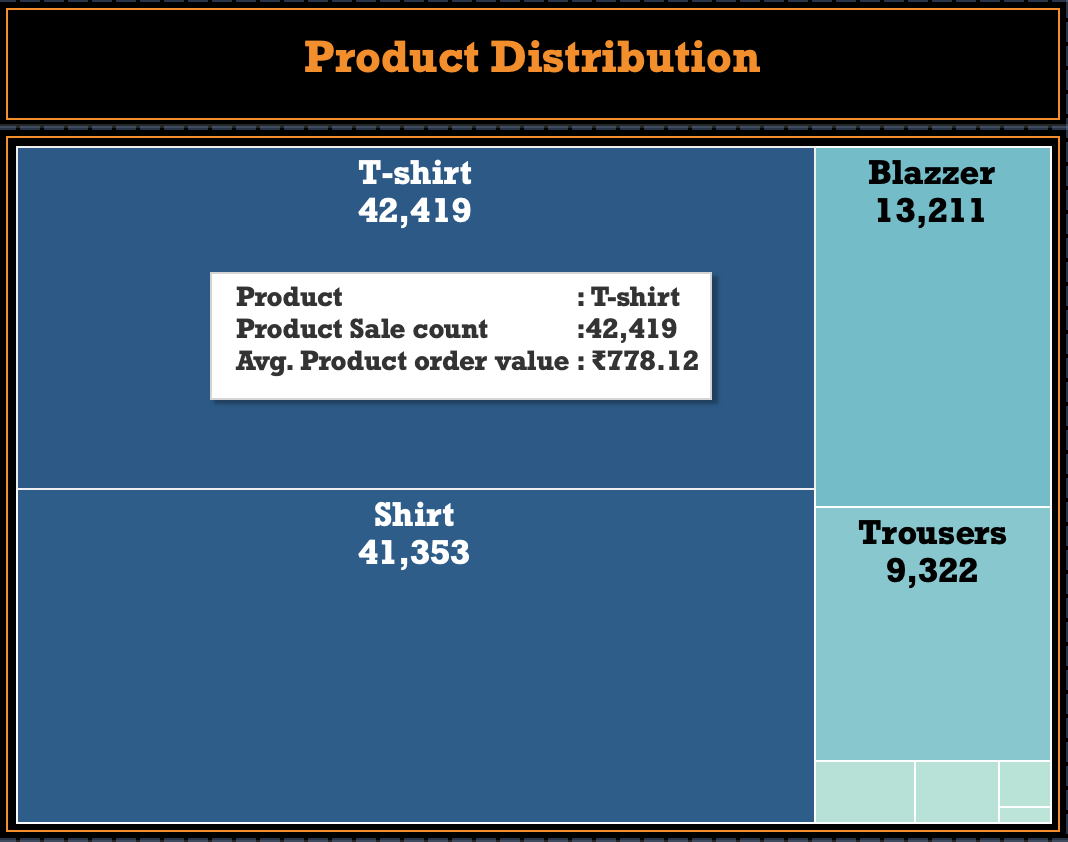


Category Analysis Dashboard :

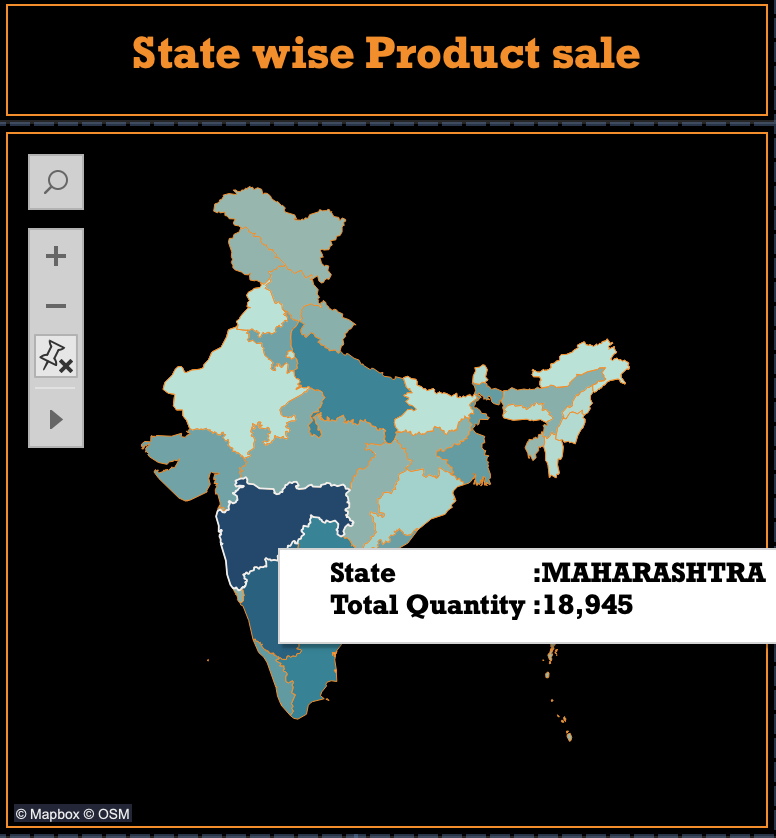


Overall Category Analysis Dashboard :

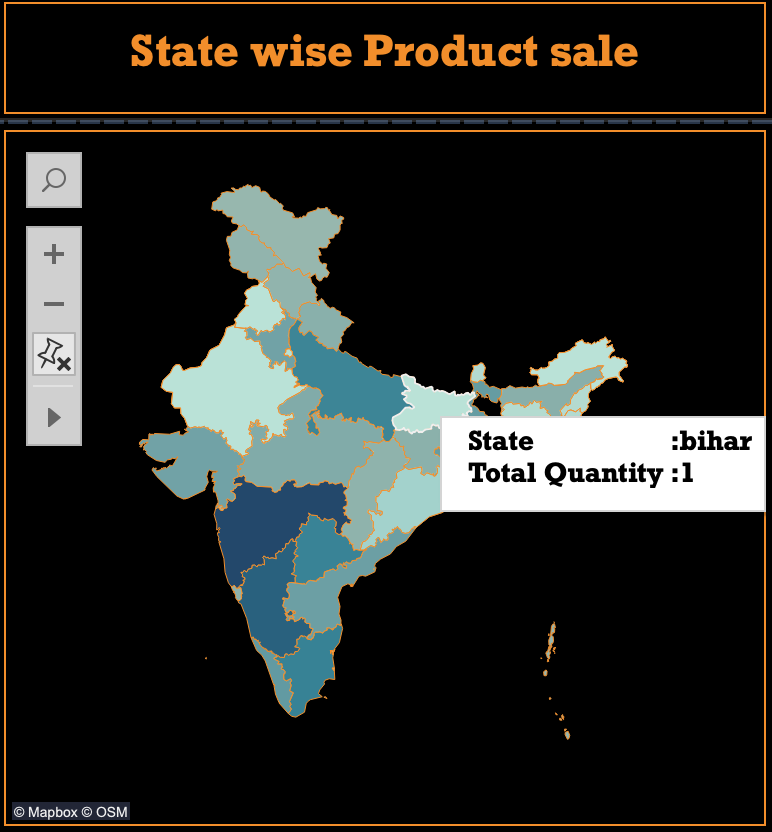


Product wise Orders: 

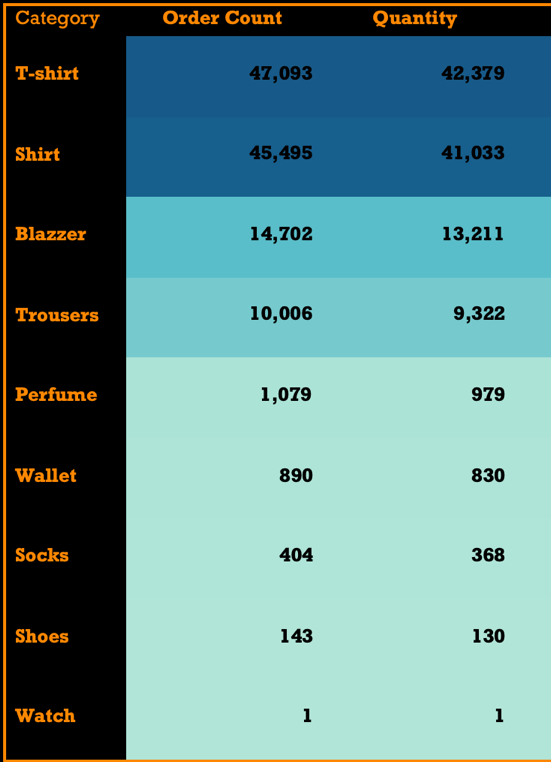
State with High Quantity Sold :



State with Low Quantity sold :



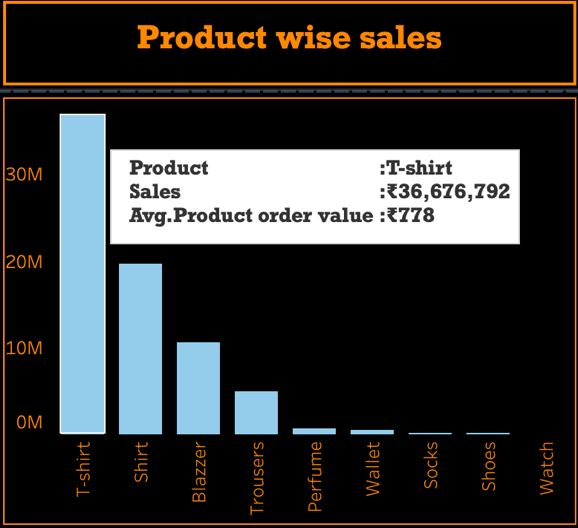
Product wise Orders count & Quantity :



Product size – high Orders:



Product with High sales Orders :



Product with low sales Orders: Watch

Top 2 Products in sales Orders: T-Shirts, Shirt

Month with high Orders: April – Rs.19,064,353

Month with low Orders: December – Rs.2,374,088

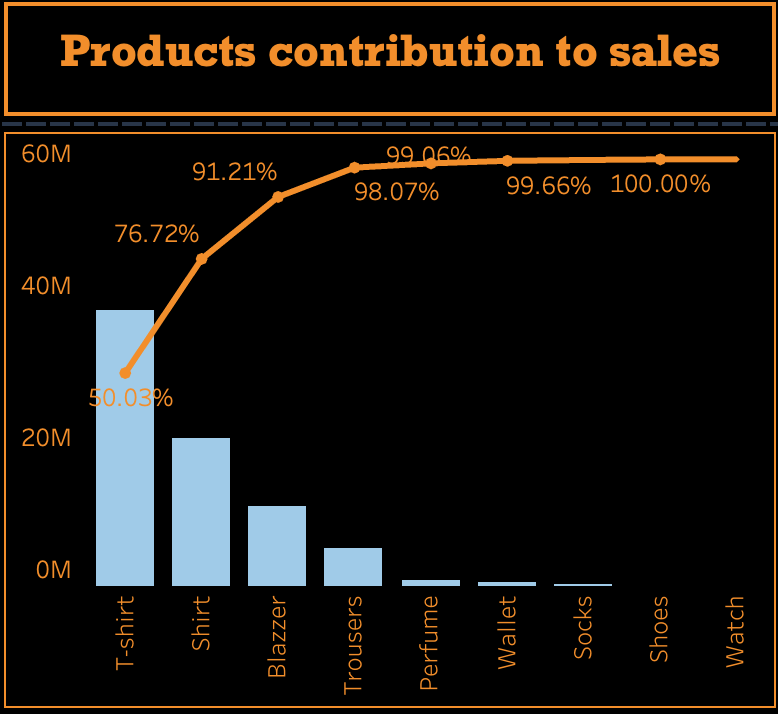
State with High Orders: Maharashtra

State with Low Orders: Meghalaya

Product size – high Orders & Quantity : T-Shirts

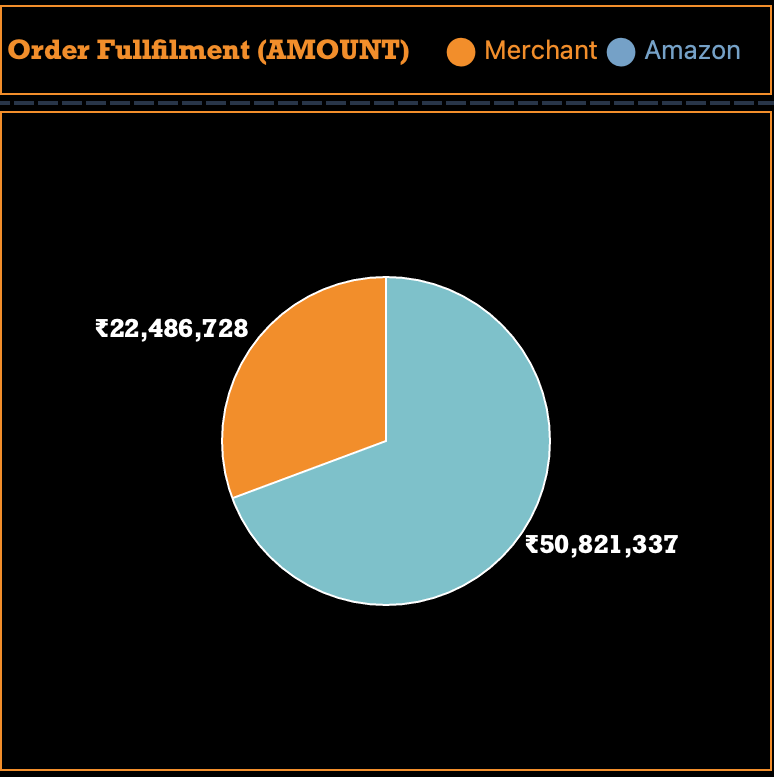
Top Product contribution to overall sales :

T-Shirts, Shirt (Both contributes More than 75% to overall sales)

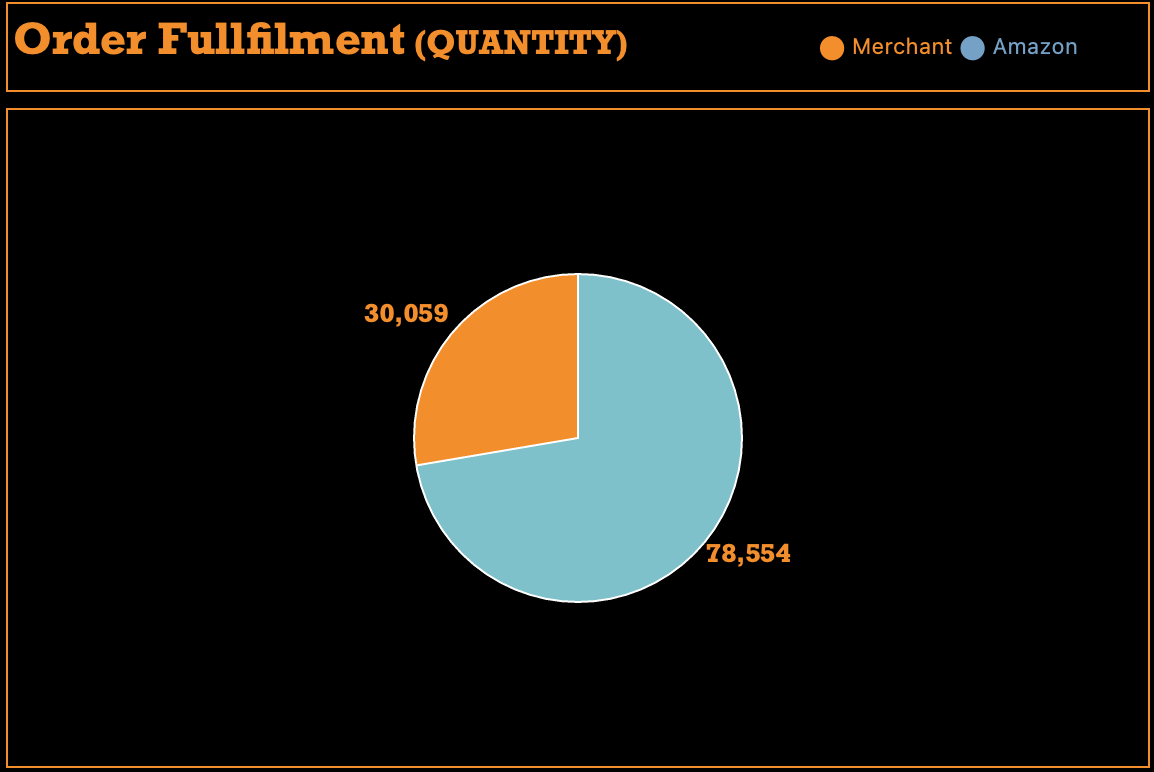


Fulfillment Analysis :

Amount generated from each Fulfilment Method :

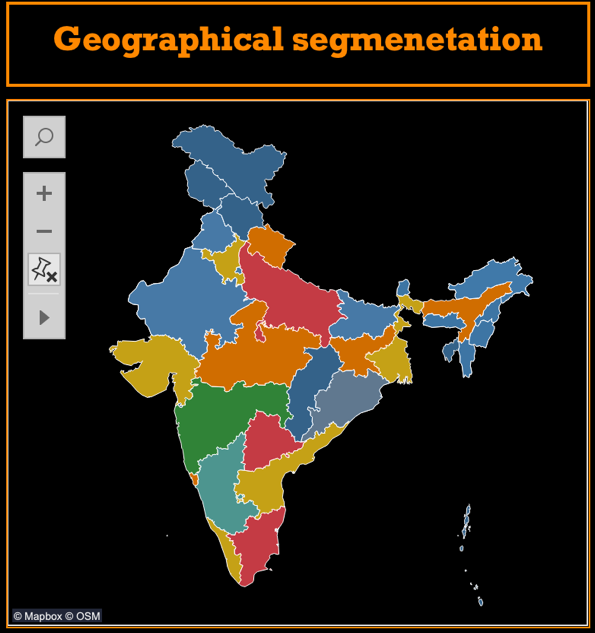


Orders generated from each Fulfilment Method :



Segmentation :

Geographic segmentation (Based on state) :



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.No | Performance | Cluster | Color | Amount |
| 01. | Very High Performing | 7 Star | Green | >1 CR |
| 02. | High Performing | 5 Star | Light Blue | >75lakhs and < 1CR |
| 03. | Good Performing | 4 Star | Red | >50 lakhs and < 75 lakhs |
| 04. | Average Performing | 3 Star | Yellow | >25 lakhs and < 50 lakhs |
| 05. | Low Performing | 2 Star | Orange | >6 lakhs and < 25lakhs |
| 06. | Very Low Performing | 1 Star | Blue | < 6 Lakhs |

Insights & Key Findings :

* Maharashtra performed very well in all areas.
* T-shirts most ordered(47,093) and most sold product (42,379).
* It is observed that North and North - East States performance is very low.
* Most of the orders are fulfilled by Amazon rather than Merchant
* Products size – M are sold most.
* It is identified that the sales and Quantity sold are showing town trend.
* April month is generated high sales.
* It is fined that more than 75% of sales are generated from

T-shirts and shirts.