ADVANCED DATA VISUALIZATION
(EXP 1)

Name: Darshit Bhagtani
UID: 2021700006
Branch:4th year CSE-DS

AIM:
Create basic charts using Tableau / Power BI / R / Python / D3.js to be performed on the dataset of Ecommerce field

- Complete all plots on practice dataset and reproduce on e-commerce dataset.
- Basic - Bar chart, Pie chart, Histogram, Timeline chart, Scatter plot, Bubble plot
- Calculate Product wise sales, region wise sales
- Write observations from each chart

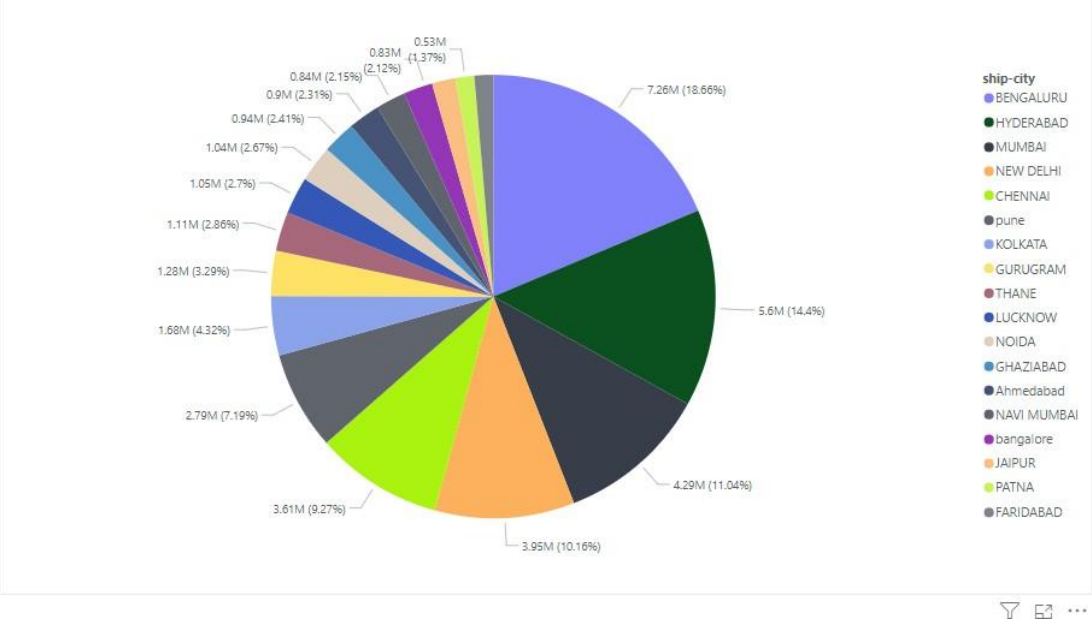
DATA:
In this assignment, I analyzed a dataset representing a single day's sales from an e-commerce platform, specifically Amazon.in. The dataset offers a comprehensive snapshot of the platform's operations and customer behavior. It encompasses a diverse range of products, primarily focused on Indian ethnic wear such as kurtas and traditional sets, alongside Western-style dresses and tops.

Geographically, the data reveals a wide distribution of customers across India, from major metropolitan areas to smaller towns, providing insights into the platform's market penetration. The varied order statuses - from shipped to canceled - along with different fulfillment methods (merchant-fulfilled, Amazon-fulfilled, Easy Ship) showcase the complexity of the platform's logistics operations. Pricing information and the presence of numerous promotion IDs hint at the platform's pricing strategies and marketing efforts.

This rich dataset not only illuminates daily sales patterns but also provides valuable insights into product popularity, customer preferences, geographical demand distribution, and the intricacies of e-commerce order processing and fulfillment in the Indian market.

DESCRIPTION WITH OUTPUT:

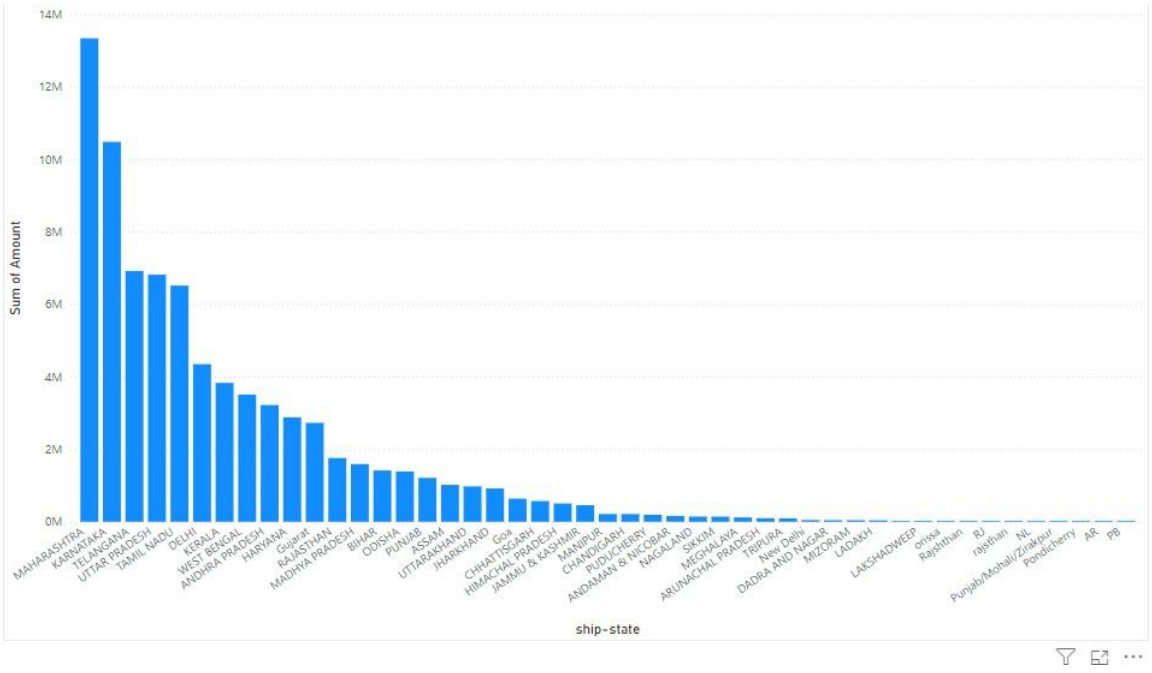
1. Pie chart of sum of amount by ship-city having amount more than 0.53M.



Observation:

1) Bengaluru city has the highest amount of sales i.e. 7.26M.
2) Followed by Hyderabad, Mumbai and New Delhi.
3) The gap between Bengaluru and Hyderabad is around 1.5M means it is highly dominating.

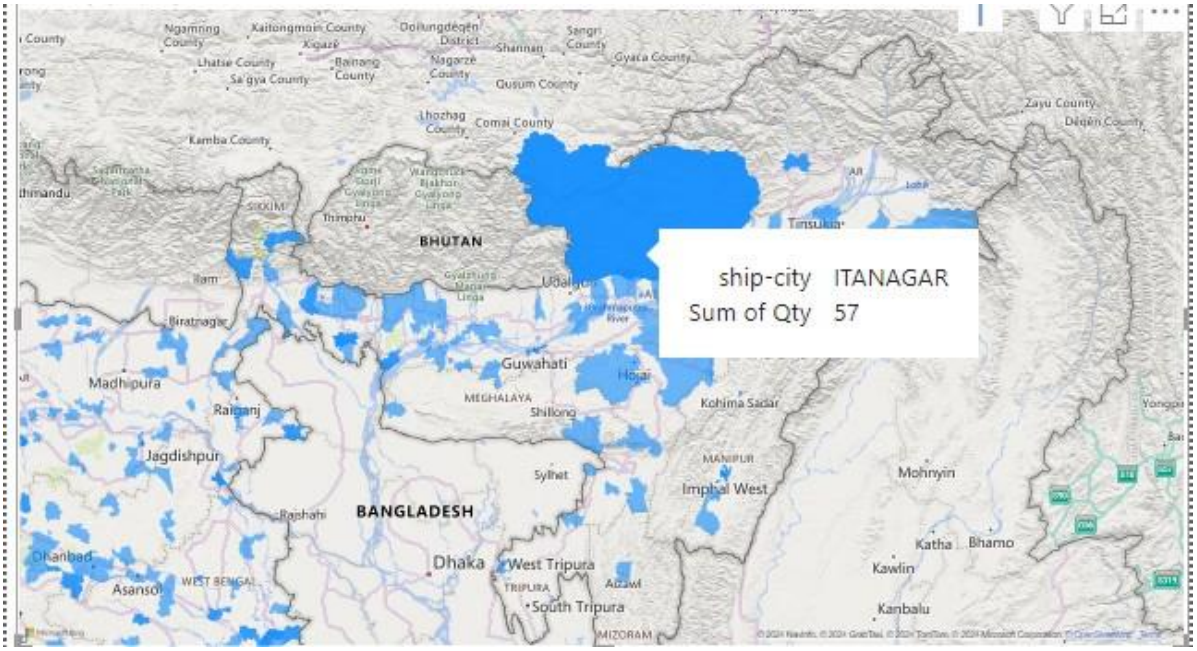
2. Sum of amount by ship state.



Observation:

1) Made a Bar chart and sorted it according to high sum amount.
2) Maharashtra, Karnataka, Telangana are the top three states.

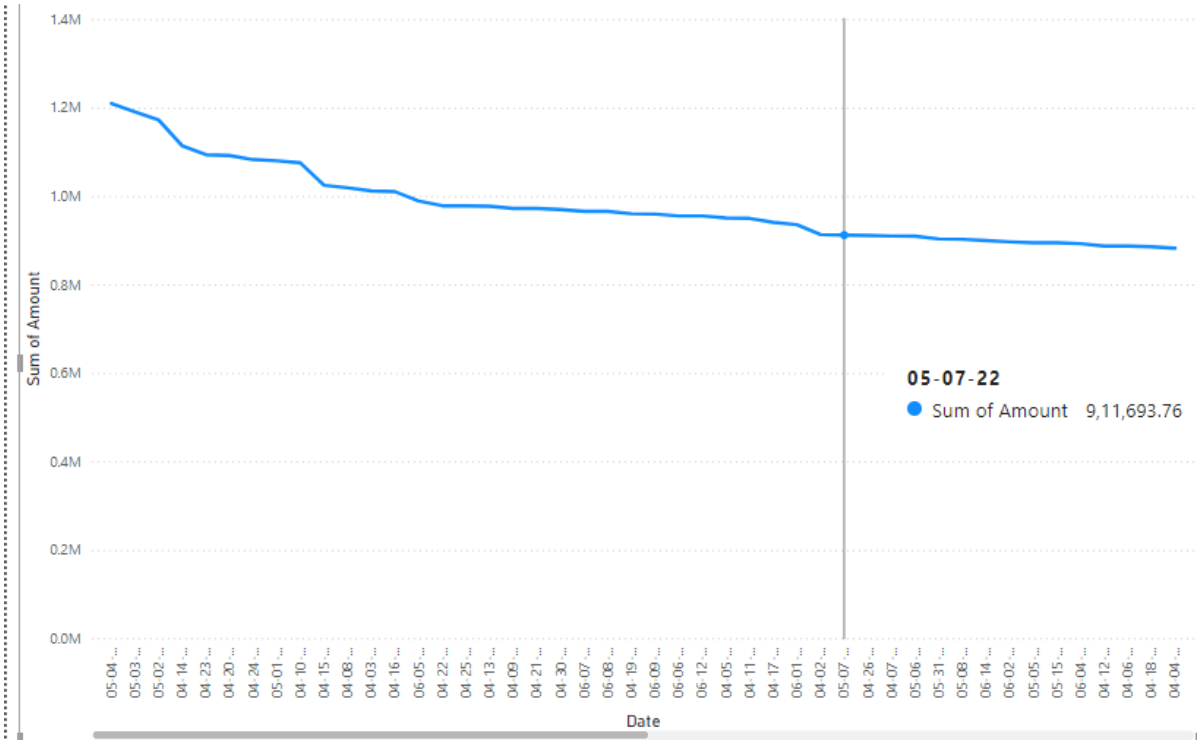
3. Geographical plot of Sum of quantity by ship-city.



Observation:

1) It is a geographical plot which fills the city with a particular and keeps it opacity according to the sum of quantity
2) Easy to visualize.
3) Can hover of each city to visualize

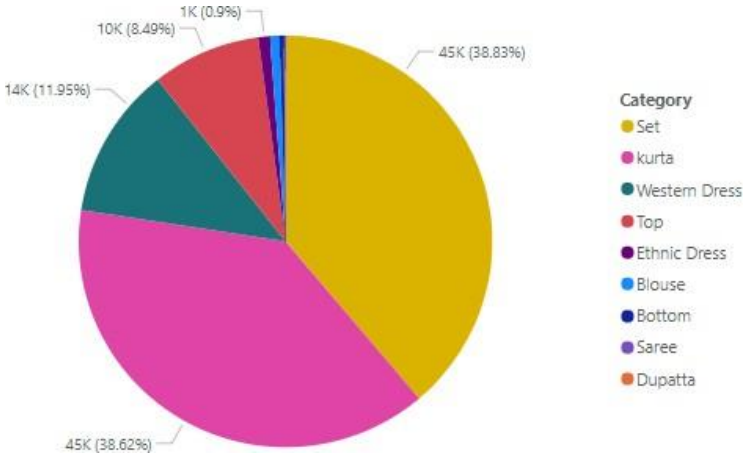
4. Time-Line chart of sum of amount by date.



Observations :

1. Time Range: The chart shows data across multiple dates, giving us a view of sales performance over time.
2. Highest Sales Day: There's a notable peak in the left of the chart, representing the day with the highest total sales amount. This could be due to a special promotion, a popular product launch, or other factors that drove increased purchases.
3. Sales Day: We can see the sales dropping with time. This could be because in the start there might be a festive season.

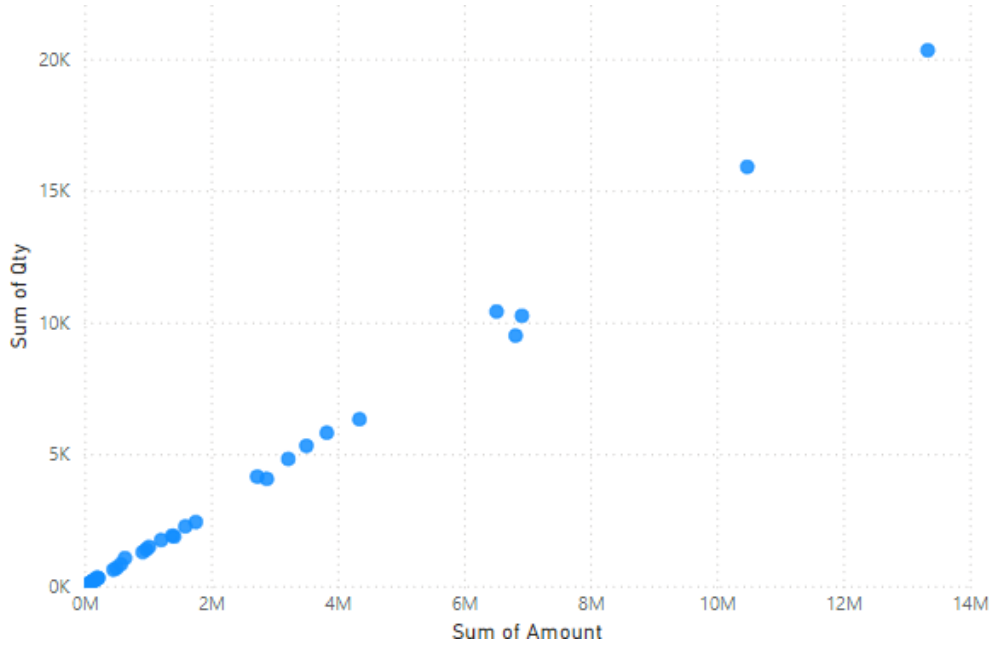
5. Pie chart of sum of quantity by Category.



Observation:

1. The pie chart shows the distribution of sales quantity across different product categories.
2. Kurta, Set appears to be the largest category, taking up a significant portion of the pie.
3. Other categories like Western Dress, Top, and Ethnic Dress make up smaller portions of the total quantity sold.
4. This visualization helps identify which product categories are most popular in terms of quantity sold.

6. Scatter Plot Sum of Amount and Sum of Quantity by ship-state



Observations:

1. Each point on this scatter plot represents a state, with its position determined by the sum of amount (x-axis) and sum of quantity (y-axis) for that state.
2. Some states stand out as having higher amounts and quantities than others, likely representing the more populous or economically active states.
3. There's a cluster of states with lower amounts and quantities, and a few outliers with significantly higher values in both dimensions.
4. This visualization helps identify which states are the top performers in terms of both sales volume and revenue.