**19CSE304**

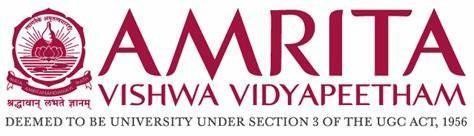
**Foundations of Data Science**

**Assignment**

**Visual Analysis of Global Climate Resilience and Natural Events**

**Submitted by: Ojas**

**Ojas, Bl.En.U4CSE22281, section ‘C’ and semester ‘V’**



AMRITA SCHOOL OF ENGINEERING, BANGALORE

AMRITA VISHWA VIDYAPEETHAM

**BENGALURU-560035**

October -2024

**Problem Statement:**

This dataset provides insights into **global climate resilience** by recording natural events like earthquakes and droughts across North America in 2024. Each row represents an event with details on **geographic location** (latitude, longitude, and region) and **weather conditions** (temperature, humidity, wind speed, precipitation, UV index). This data helps analyze patterns of natural events, weather variations, and temporal trends, offering a basis for understanding climate impacts across different regions.

**Features in Dataset:**

* **Number of Rows:** The dataset contains a series of rows, each representing an event or observation with specific climate and geographic attributes. The total number of rows indicates the number of recorded events across different regions and dates.
* **Number of Columns:** The dataset includes several columns, each capturing a different aspect of the event or weather conditions.

**Data Columns Explanation:**

1. **Date**: Represents the date when the event was recorded, helping in analyzing time-based trends such as seasonal changes or specific event timings.
2. **Event Type**: Specifies the type of natural event, such as "earthquake" or "drought." This column is essential for categorizing and comparing different event impacts.
3. **Location**: Describes the area or region (e.g., state or country) where the event occurred, allowing for geographic analysis and identification of event-prone areas.
4. **Latitude and Longitude**: These columns provide the geographical coordinates of the event, making it possible to plot the events on a map and analyze spatial patterns.
5. **Temperature (Average, Max, Min)**: Records the average, maximum, and minimum temperatures for each event, giving insight into temperature fluctuations and extreme weather conditions associated with natural events.
6. **Humidity**: Indicates the humidity level at the time of the event, useful for assessing dryness in drought events or moisture levels in other regions.
7. **Wind Speed**: Measures the wind speed recorded during the event. High wind speeds can have significant effects, particularly in dry conditions or during storms.
8. **Precipitation**: This shows the amount of precipitation (rainfall) associated with each event, helping to understand rainfall patterns, especially in regions facing droughts.
9. **UV Index**: Indicates the UV radiation level at the time and place of the event. High UV levels may correlate with sunny, dry weather, whereas lower levels could suggest cloud cover.
10. **Weather Conditions**: Describes the overall weather during the event (e.g., sunny, rainy, snowy). This provides additional context to other climate factors like temperature and precipitation.

**Introduction:**

Dataset offers climate resilience information containing the events such as EQ, and DR in the USA, Canada, and Mexico in 2024. Each record captures Event details: Earthquake or drought, geographical area, and year of occurrence. Geographical data: Details of longitude and latitude. Weather conditions: Such as average, maximum, and minimum temperature, relative humidity, wind velocity, amount of precipitation, the UV radiation index, and the state of the atmosphere such as sunny, light snow, and the like. I like that this dataset explains how weather variables develop in areas prone to natural disasters. Auxiliary features associated with each event type or location can be analyzed and can provide knowledge of how temperature or humidity changes to the event.

**Objective:**

* Explain objectives in bullet points for visualization tasks

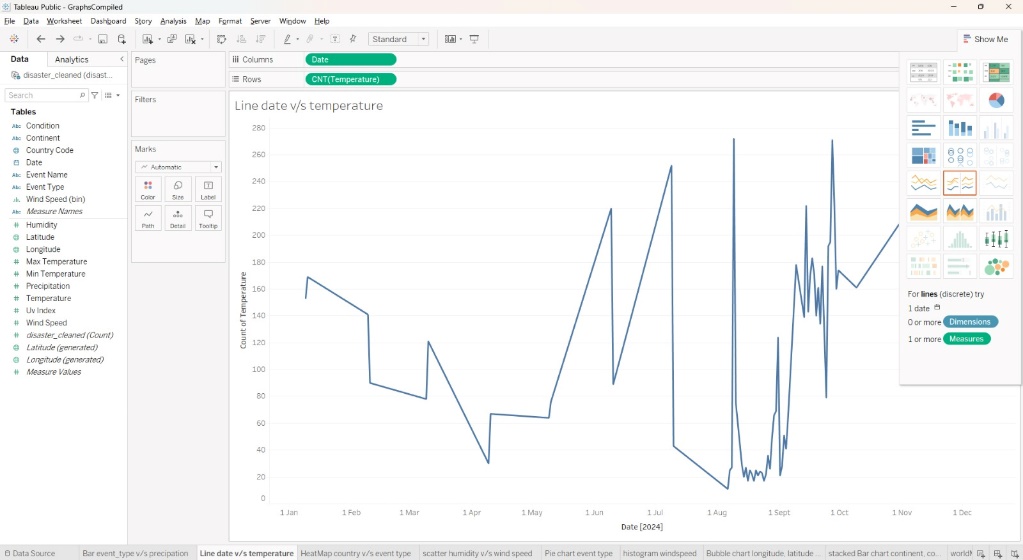
**Advantages of Data Visualization:**

1. Enhanced Understanding: Simplifies complex data, making it easier to interpret and draw insights.
2. Pattern Identification: Helps in spotting trends, outliers, and relationships within the data.
3. Quick Decision-Making: Facilitates faster analysis, supporting timely and informed decisions.
4. Engagement and Clarity: Makes data more accessible and engaging, improving communication with diverse audiences.

**Research Questions:**

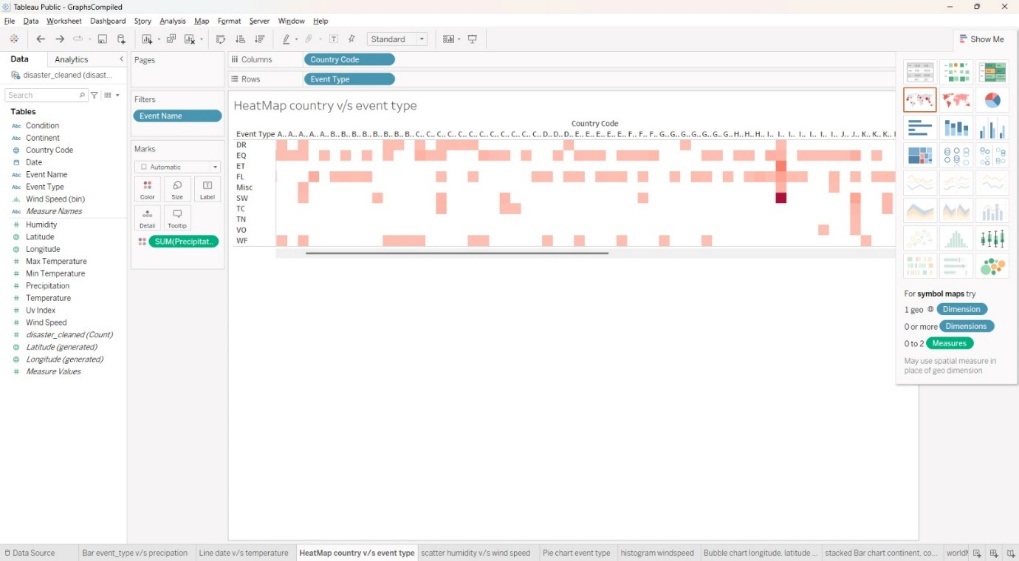
1. **LINE CHART**

* **Purpose:** The line chart shows the average temperature which has been constantly increasing throughout the year.
* **Observations:** If the format used dates, then the appearance of any fluctuations in temperature. For instance, fluctuation in temperature change may be in harmony with the timescale or, it may also reflect the effects relating to a particular event.



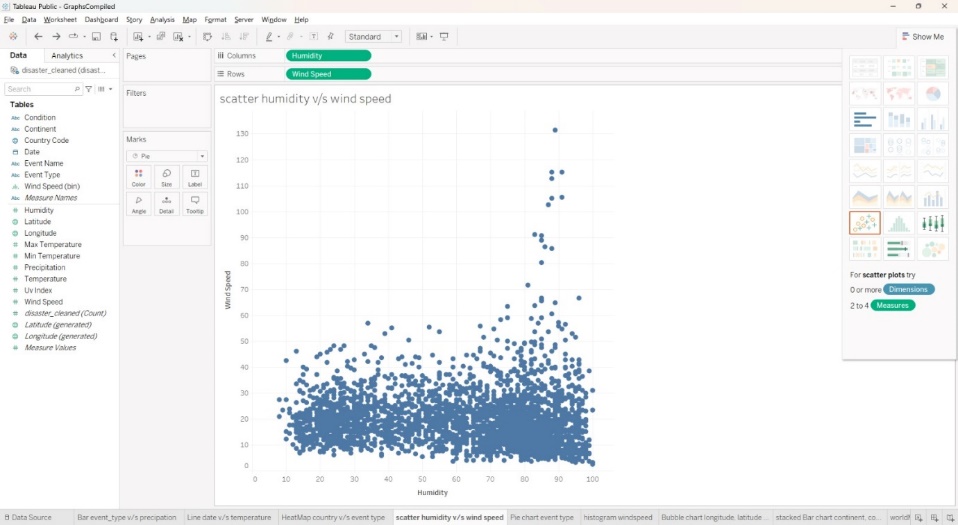
1. **HEATMAP**

* **Purpose:** This would show the distribution of earthquakes and droughts across different regions.
* **Observations:** Identify if certain locations experience more earthquakes or droughts. For example, California-Nevada and Central California may show more earthquakes, reflecting their seismically active nature.



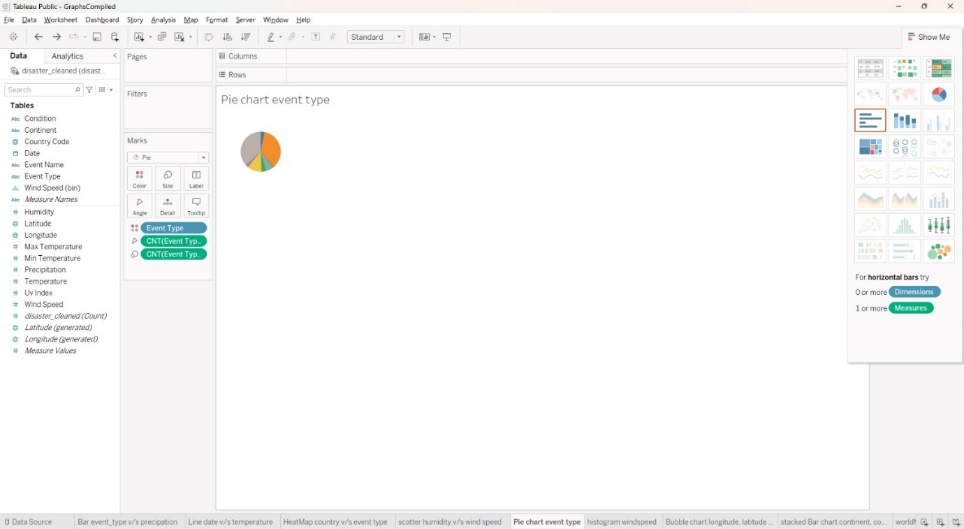
1. **SCATTER PLOT**

* **Purpose:** It is seen in the scatter plot given above how much humidity and wind speed are there for different events in the dataset. This assists in determining whether or not two of these weather variables have a relation.
* **Observations:** They show trends, thereby pointing out which ranges of the humidity are related to what ranges of wind speed. For instance, low rainfall zones with low humidity and different wind velocities may be interpreted as dry conditions sometimes associated with irregular winds. That is high humidity accompanied by high velocity could represent storm-inclined conditions probably in areas prone to earthquakes.



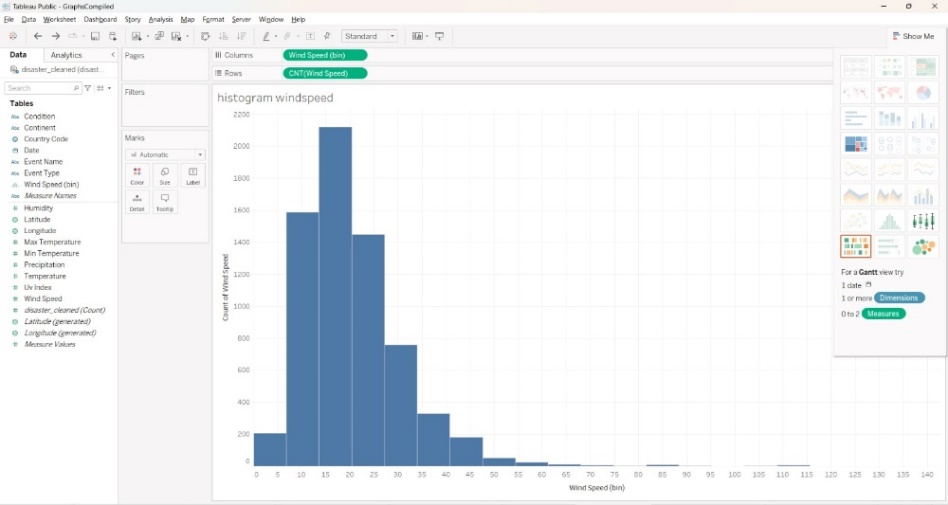
1. **PIE CHART**

* **Purpose:** The pie chart depicts the proportion or percentage of various kinds of events including earthquakes, droughts, etc in the given input data set. This visualization provides an idea of the relative occurrence of each type of event.
* **Observations:** The pie chart depicts the difference between one type of event and another type. This proved useful for quickly getting an overview of what types of occurrences dominate the natural realm, and therefore what planning for preparedness and recovery should primarily entail in each of these categories.



1. **HISTOGRAM**

* **Purpose:** The histogram visualizes the distribution of **wind speeds** across different events in the dataset.
* **Observations:** The histogram may show certain wind speed ranges as more common, with peaks around specific values. For instance, if lower wind speeds are more frequent, it could indicate that most events occur in relatively calm conditions. On the other hand, if higher wind speeds are common, it might suggest that certain regions are prone to more intense weather activity.



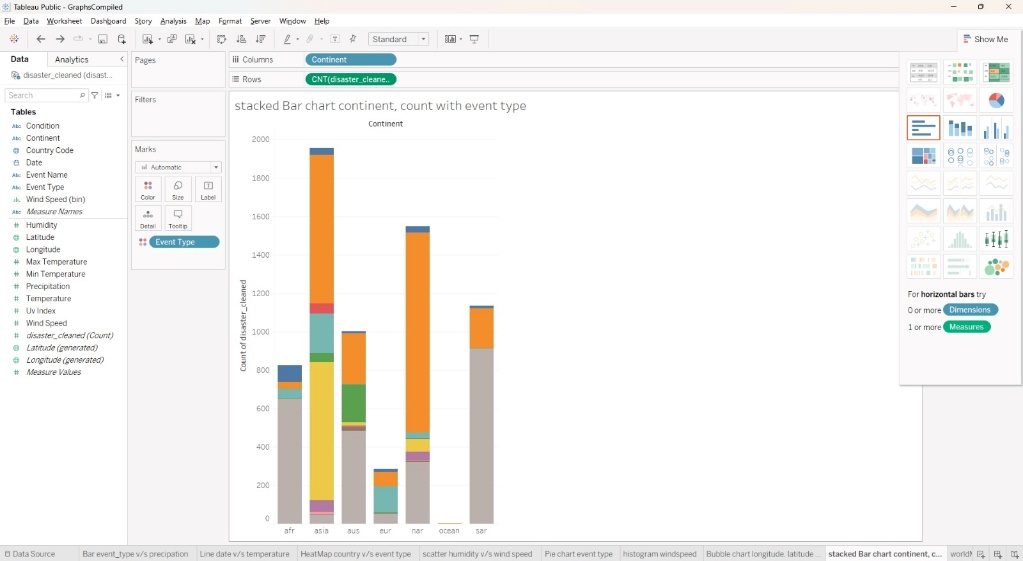
1. **BUBBLE CHART**

* **Purpose:** The bubble chart represents **longitude** and **latitude** as the position coordinates for different events, with the **size of each bubble corresponding to the UV index**.
* **Observations:** The chart shows clusters of events in certain geographical regions, which can indicate hotspots of natural events like earthquakes or droughts. Larger bubbles in specific areas suggest higher UV indexes, which could correlate with regions experiencing more intense sunlight or specific climate conditions.



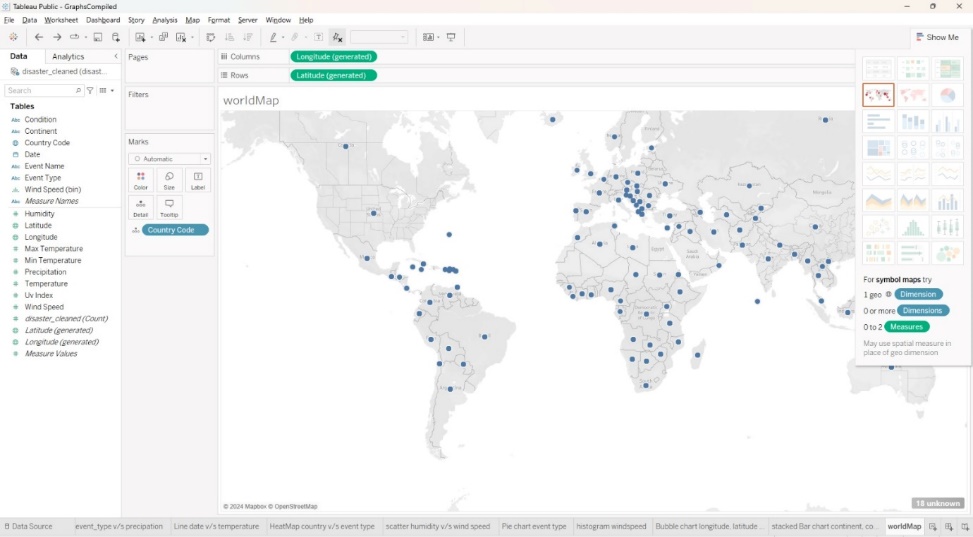
1. **STACKED BAR CHART**

* **Purpose:** The stacked bar chart shows the **count of events** for each **continent**, with different colors representing **event types** (e.g., earthquakes and droughts).
* **Observations:** The chart reveals that certain continents experience more of one event type than another. For instance, North America might have a higher frequency of earthquakes, indicated by a larger section of the corresponding color in the stacked bars. Conversely, other continents might show a more balanced or different distribution, highlighting regional differences in natural events.



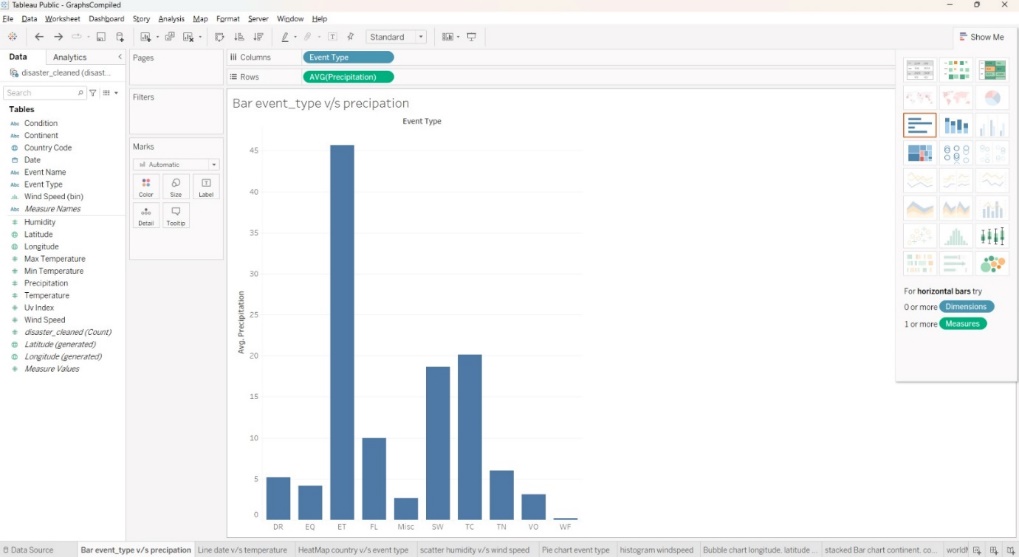
1. **SYMBOL MAP**

* **Purpose:** The symbol map plots **longitude** and **latitude** to display the **geographic locations** of recorded events on a map.
* **Observations:** The map may show concentrations of events in certain areas, identifying regional hotspots. For instance, clusters along specific longitude and latitude coordinates could indicate areas more prone to seismic or drought activity. Patterns in event distribution can reveal geographic trends, such as coastal or inland areas experiencing more frequent events.



1. **HORIZONTAL BAR CHART**

* **Purpose:** The horizontal bar chart displays **event types** (e.g., earthquakes and droughts) on the y-axis with corresponding **precipitation levels** on the x-axis.
* **Observations:** The chart shows that certain event types are associated with higher or lower levels of precipitation. For instance, drought events might have little to no precipitation, reinforcing dry conditions, while earthquakes could display a broader range of precipitation levels.



1. **PACKED BUBBLE CHART**

* **Purpose:** The packed bubble chart displays the **maximum temperature** for each **continent**, with bubble size corresponding to the temperature level.
* **Observations:** The chart shows larger bubbles for certain continents, indicating higher maximum temperatures in those areas. For instance, if Africa or Asia has the largest bubbles, it suggests that these regions experience more intense heat events. Smaller bubbles on other continents could reflect milder maximum temperatures.

