Group 3

Team Members:

Name	CID	Email
Fardifa Fathmiul Alam	C17219280	fardifa@g.clemson.edu
Taoli Wuyun	C60799916	twuyun@g.clemson.edu
Ranga Baminiwatte	C16768228	abamini@g.clemson.edu

Project Repository: https://github.com/ranga519/CPSC-6030-Project

Project Website: https://ranga519.github.io/CPSC-6030-Project/

Visualization Video: https://youtu.be/zU00vFGjqlQ

Natural Disasters in the Last 50 Years: Where and from which disasters do people get most affected?

Background and Motivation:

The motivation behind exploring natural disasters in the last 50 years lies in understanding and mitigating their impact on human populations. Without a comprehensive overview of natural disasters' geographical and typological distribution, communities and governments may struggle to implement effective disaster preparedness and response measures. The absence of a visualized dataset on natural disasters could lead to uninformed decision-making, inadequate resource allocation, and heightened vulnerability to the devastating effects of such events. Data visualization is crucial in simplifying the complexity of disaster-related data, providing actionable insights into patterns, vulnerable regions, and types of disasters that have historically caused the most harm.

Visualizing the occurrence and impact of natural disasters over the past 50 years helps identify high-risk areas, enabling proactive measures to enhance resilience and reduce the human and economic toll. Furthermore, understanding the prevalence of specific types of disasters in different regions allows for tailored preparedness strategies. This information is pivotal for governments, NGOs, and communities to allocate resources effectively, plan evacuation routes, and implement sustainable infrastructure to withstand recurring threats.

Project Goal:

To depict a comprehensive narrative of natural disasters over the last 50 years, highlighting affected regions, prevalent disaster types, and their impact on human populations, to inform targeted disaster management and mitigation strategies.

Project Objectives:

1. Geographical Distribution - Identify regions most affected by natural disasters and visualize the frequency and severity of incidents in each area over the past 50 years.

2. Typological Analysis - Analyze the types of natural disasters that have historically caused the most significant impact and visualize their distribution globally.

3. Humanitarian Impact - Examine natural disasters' demographic and socioeconomic effects, including the number of casualties and economic losses, focusing on vulnerable countries.

4. Correlation between different measures of impact - Visualize and compare how the countries are affected in terms of various measures of impact such as the frequency of occurrence, death rate and financial damages

5. This project aims to provide a comprehensive understanding of the patterns and impact of natural disasters, enabling stakeholders to make informed decisions, allocate resources effectively, and implement proactive measures to enhance resilience and reduce the overall impact on affected populations.

Data: The data has been collected from Kaggle | Natural Disasters 1900-2021

Data Processing: The primary dataset contained natural disaster data over 121 years. However, as we are interested in events over the last 50 years, we reduced entries to 14,450. We aim to perform some of the following data processing tasks to make the visualization more effective:

1. Remove the events that contained null values for the number of disasters, Financial damage and number of deaths,
2. Change the given date format and convert it into year and month separately to establish a proper time series.
3. Convert the binary values into categorical values
4. Extract the latest name of a country from the “Location” attribute to replace older names of a country that isn’t being used in modern context anymore

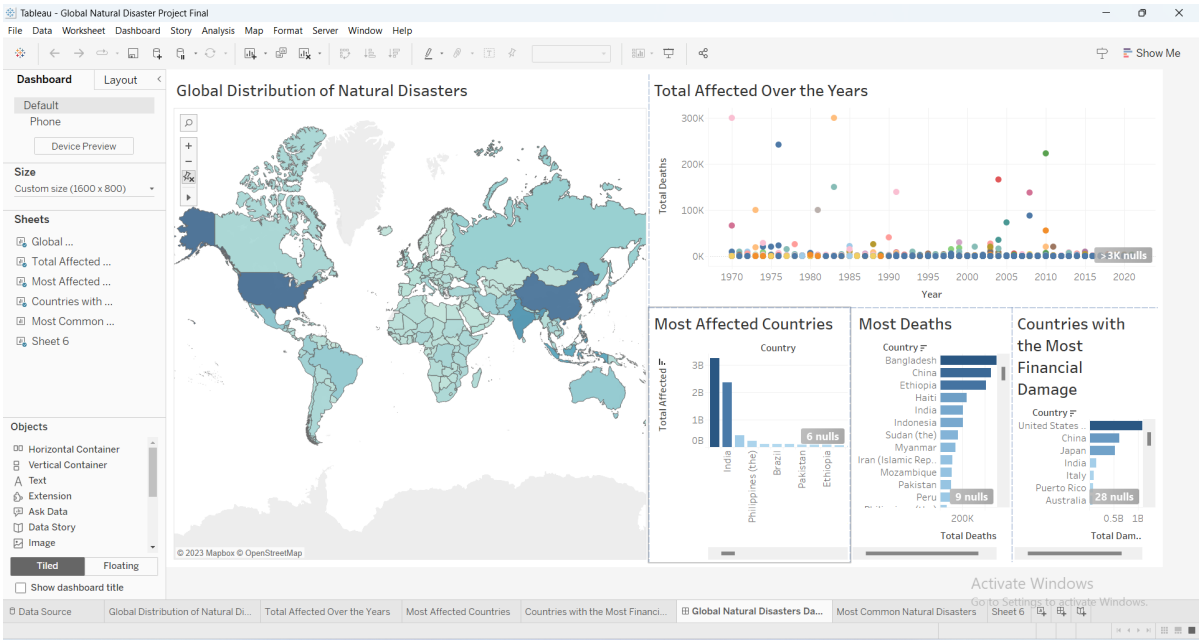
DESIGN EVOLUTION

Initial Questions We Wanted to Answer:

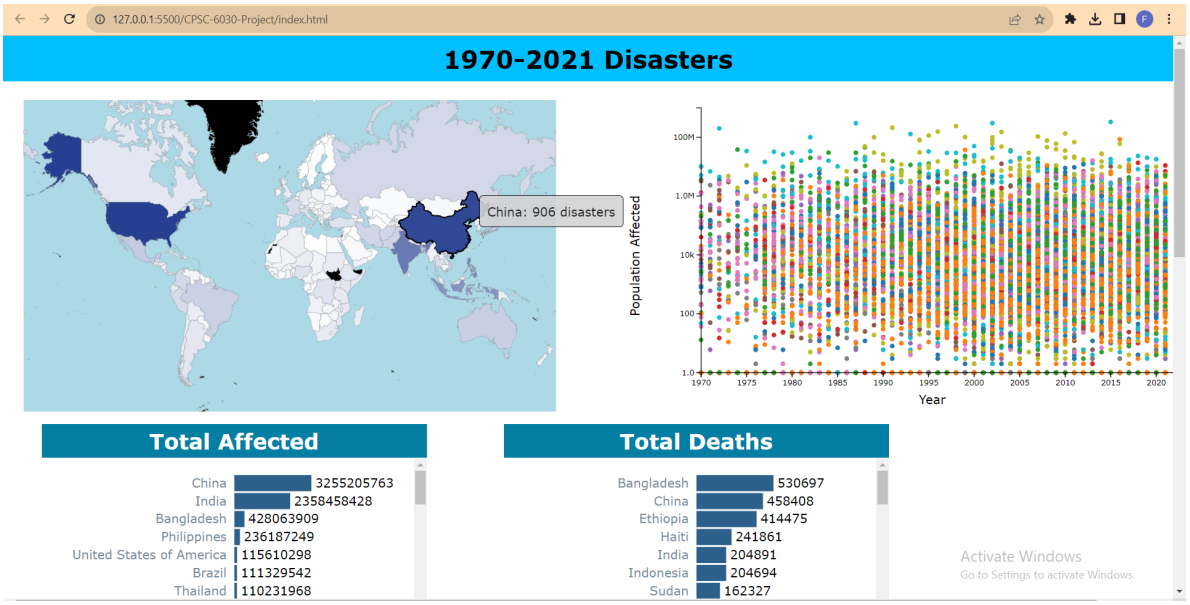
1. What is the geographical distribution of the frequency of natural disasters across the world?
2. What are the top 10 Countries affected the most by number of affected, number of deaths and financial damages?
3. How have the total number of people affected by various natural disasters changed over the years?

We created the initial prototype and dashboard based on these questions

Initial Prototype:



Initial Dashboard using D3JS:



For our initial prototype and dashboard, we showed a global distribution of natural disasters across countries of the world by frequency of a disaster. We showed a scatter plot of disasters by singular events by total number of deaths over the years which we later replaced by total number of affected people as it shows a bigger picture of how a natural disaster impacts a country. The scatter plot events were extracted from a yearly data hence they were seen vertically linear at first. One column of points meant all disasters occuring on that specific year. We also had 2 more bar charts showing countries ranked by total affected and deaths.

After working on our initial dashboard, we found some critical insights about the data that encouraged us to change our frame of questions and subsequent visualizations:

1. **Nonlinear Correlation between Frequency of Occurrence and Impact resulting in changing from map view to multiple bar charts side by side to show comparison:**

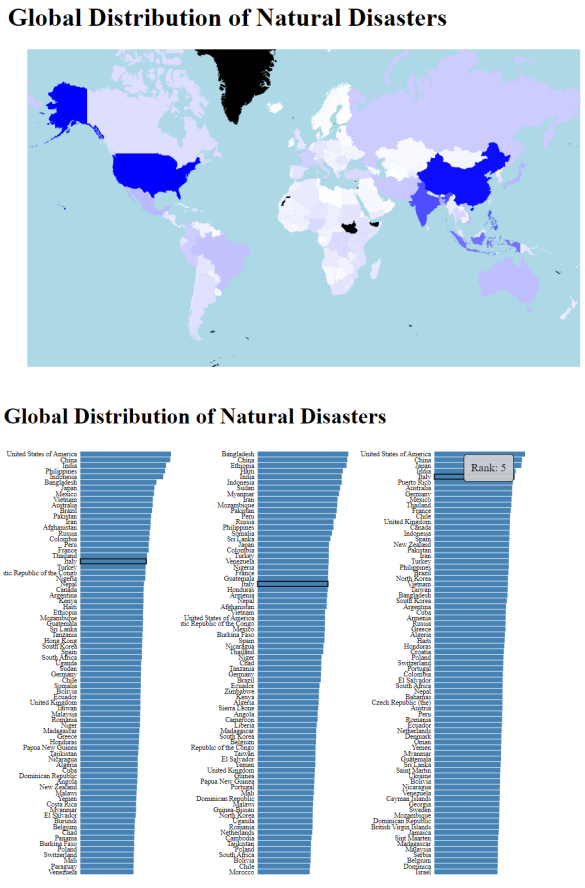


Figure: Before and after

Initially, we wanted that if we click on a country, we will see the rank of that country within the bar chart of most affected, deaths and financial damages. But the correlation between

them is not linear - a critical insight we found from the data. For example, Japan doesn't have a lot of occurrences of natural disasters but it does rank one of the highest in terms of financial damages. On the other hand, a country might have a lot of financial damages, but it has much better control in terms of death rate, such as Canada. Showing the comparison of the various metrics of impact through connecting maps and three bar charts was not easy for the users to visualize. So, to solve this, at first we dropped the bar chart for total affected as we already had a scatterplot showing its changes. Secondly, during one of our classes on abstraction of data, we saw a side by side comparison of three bar charts that compare data between its various attributes. There was a line connecting three bars of the same item. And that inspired us to change the focus of our work to visualize the COMPARISON of various metrics. The geographical distribution didn't add much value to our analysis. So, we made 3 bar charts showing 3 measures of impact of a natural disaster - frequency, death rate and financial damages. Initially we showed the top 10 only before we added comparison of measures but later we added all the countries for the users to choose for what country they want to see the comparison.

2. **Varying Impact of Various Disaster Categories across the countries and the years**

We wanted to show through the scatterplot how natural disasters occur throughout the years and how their impact on the number of affected people changes as well. Working with the data, we found that different natural disasters impacted different countries at different rates throughout history. Some disasters are becoming more widespread such as the increase of floods due to the increasing global warming (probable hypothesis), some disasters are more infrequent such as volcanic activity. There was a need to introduce a visualization with disaster types as categorical values to show these changes. That's why we created a bubble chart for various disaster types where if a user interacts and clicks on it, can see patterns for a specific disaster.

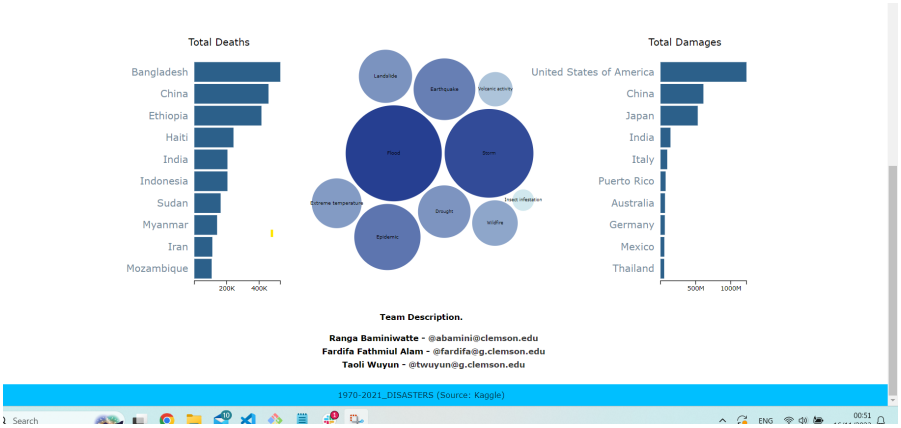


Figure: Before and after

We added the bubble charts while still playing around with top 10 disasters. The initial one took different shades of blue to show disasters by the number of occurrences but later changed to color code it by giving unique values to each categorical value. The size of the radius already depicts the frequency of the disaster type, having color as a second channel to show the same thing seemed redundant.

3. **Importance of Natural Disasters Occurrence by month and not only year and by disaster type**

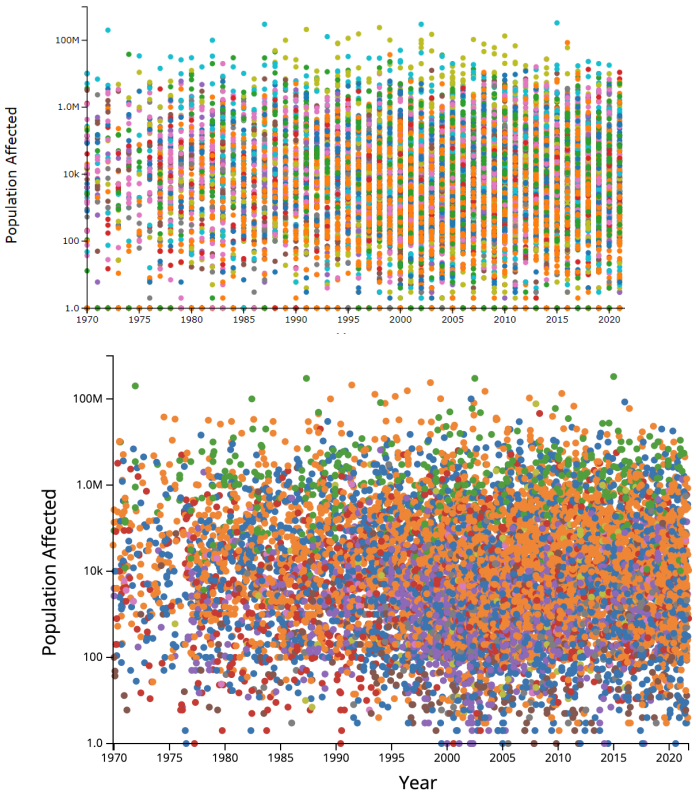


Figure: Before and after

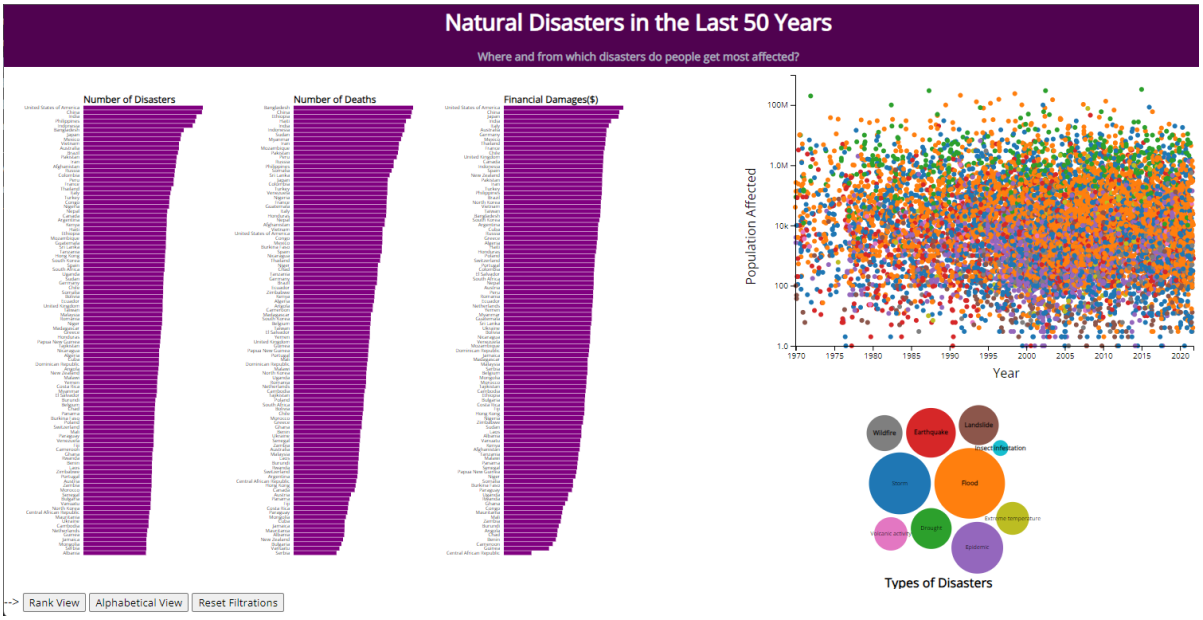
Previously, the events were sorted only by year value, one column meant all the disasters happening in that year. Later, through data processing we extracted months and mapped the disaster events in the scatterplot using both months and years to give a more in-depth view of the occurrence throughout the years. On the other hand, previously, the points were color coded as per the name of the country - but having 198 shades of colors for 198 different countries meant that there was no visual difference between colors - lightest blue showing USA and slightly darker blue showing Argentina wouldn't be visible to the naked eyes of the users. A distinct scale of color values was necessary to understand the plot, so the colors were coded later as per the disaster category. It was coherent with the bubble chart as well. If someone clicks on a specific disaster, all the disasters of that type will be shown in the scatterplot with the exact same color scale between these two charts.

With this gradual evolution of our designs based on the data we were working with, our final visualization focused on answering the following questions that generated crucial insights:

1. Correlation between frequency of a natural disaster and various measures of impact - death rate and financial damages
2. Top countries affected by various types of natural disasters
3. How have the total number of people affected by various natural disasters changed over the years and how does it vary for various natural disaster types and for each country?

Based on these questions, we have the following final dashboard

The Final Dashboard:

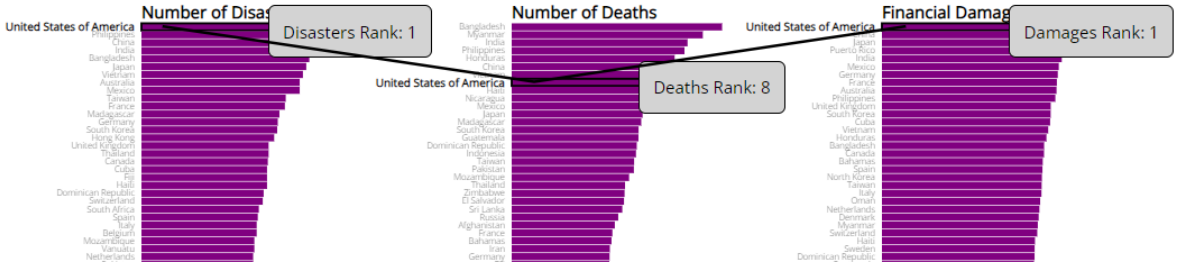


FINAL IMPLEMENTATION

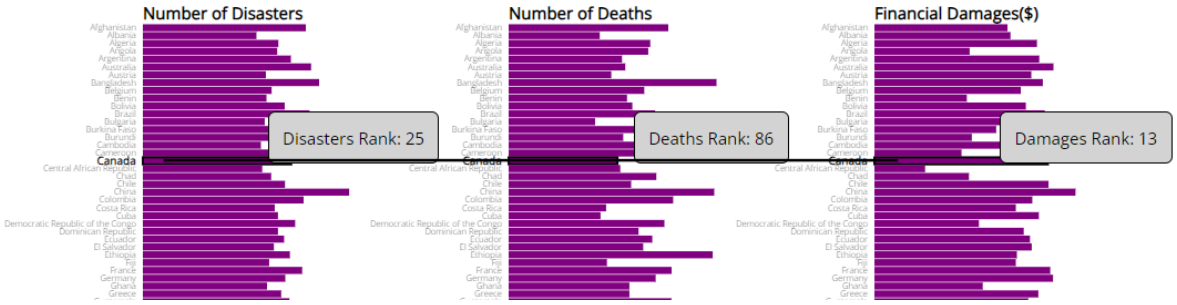
The main dashboard contains 3 visualizations,

1. Triple Bar chart
2. Bubble chart
3. Scatter Plot

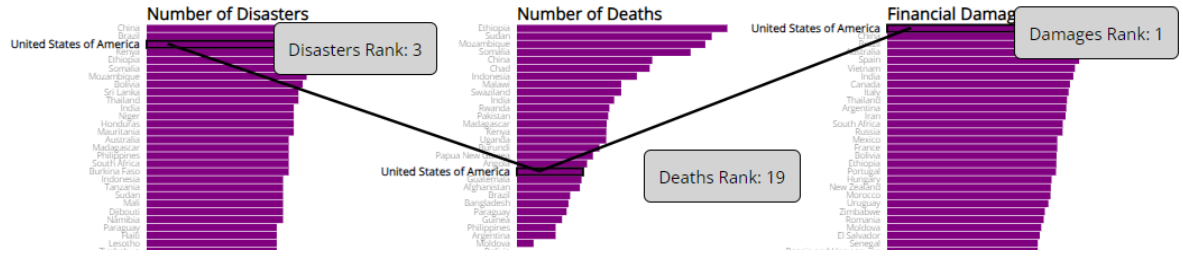
1. **Barchart**: The barchart is presenting a comprehensive overview of global disasters by utilising three bar charts that highlight the total number of disasters, deaths, and economic damages for each country. The dataset is filtered to showcase the top 100 countries based on the total number of disasters, providing users with valuable insights into the most affected regions.



Two distinct views are offered to enhance user experience: one displaying countries ranked from largest to smallest based on the total number of disasters, and the other presenting countries in alphabetical order. Users can seamlessly switch between these views by clicking on the "ranking" and "alphabet" buttons, tailoring the visualization to their preferences. Here is the alphabet view:

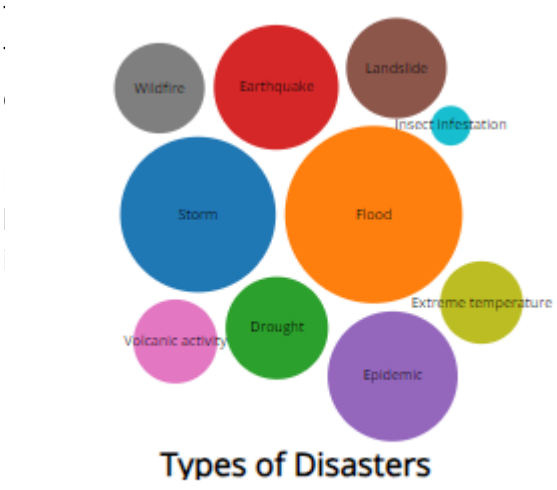


The project incorporates interactivity by allowing users to explore specific types of disasters. Upon selecting a particular disaster type, the bar chart visualization dynamically adjusts to showcase the number of disasters, deaths, and economic damages caused by the chosen disaster category. This interactive feature is further refined to display data only for the top 100 countries with the highest occurrence of the selected disaster type. Following figure shows the ranking after choosing the Drought as a disaster category. For the Drought disaster, the USA ranked 3rd in terms of numbers. But for overall disaster ranking, the USA ranks 1st. In summary, this bar chart provides a user-friendly and dynamic exploration of global disasters, offering valuable insights into the distribution of incidents, their impact on human lives, and the economic ramifications across the top 100 affected countries for each category of disasters.



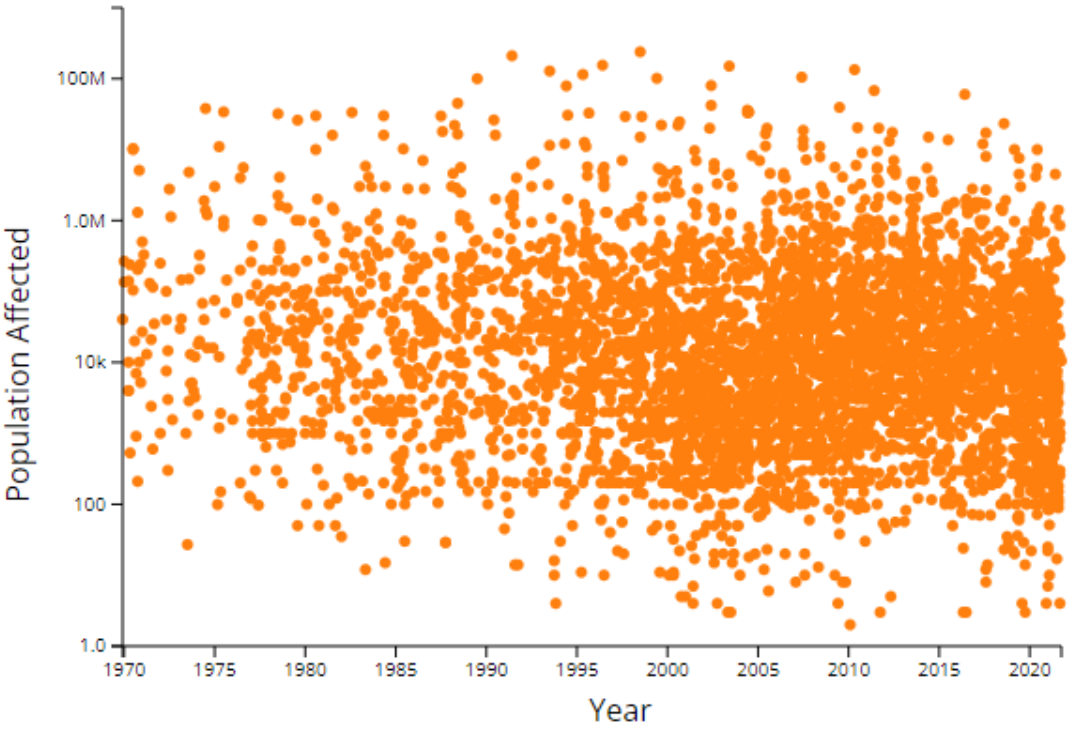
2. **Bubble Chart:**

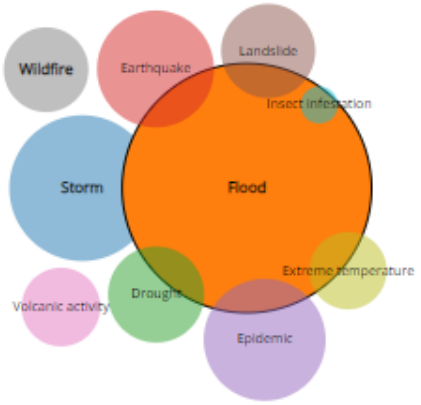
The primary purpose behind this visualization is to give the user a quantitative description of the aggregated number of types of disasters that have occurred in the last 50 years.



eir categorical values of “Disaster
ɔt dots corresponding to a single
gated count numerically.

· plot and triple bar charts. The idea
ic disaster type and conduct further

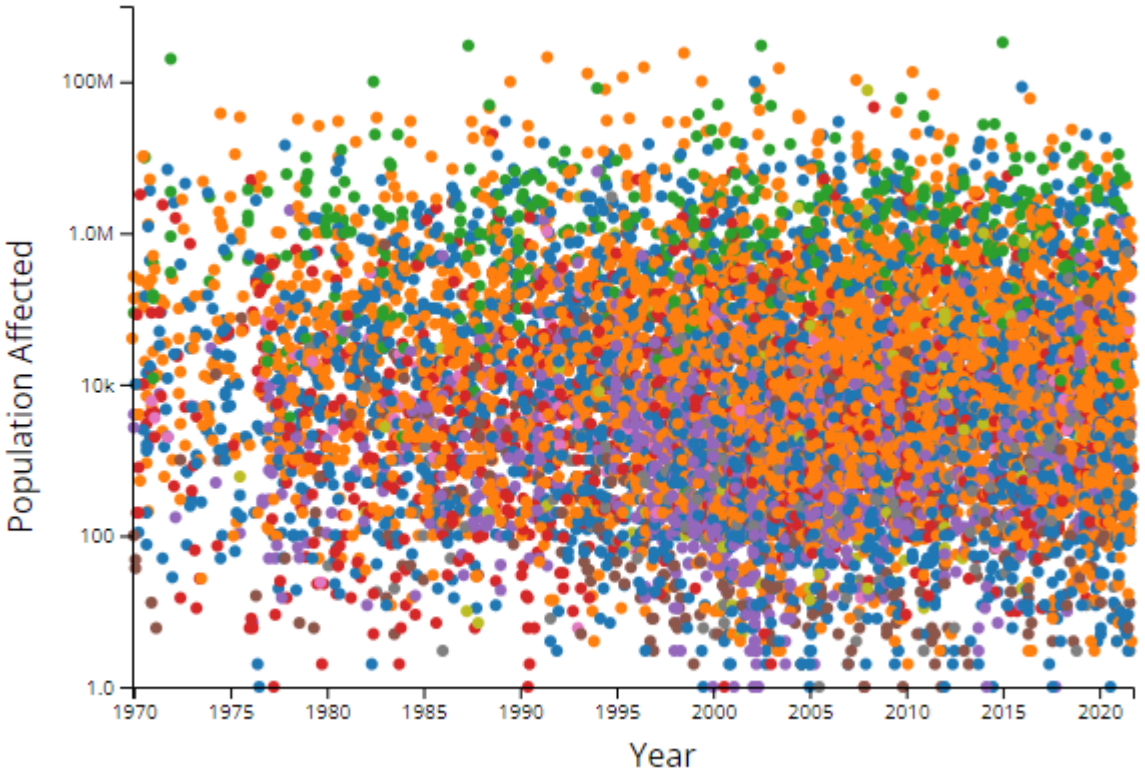




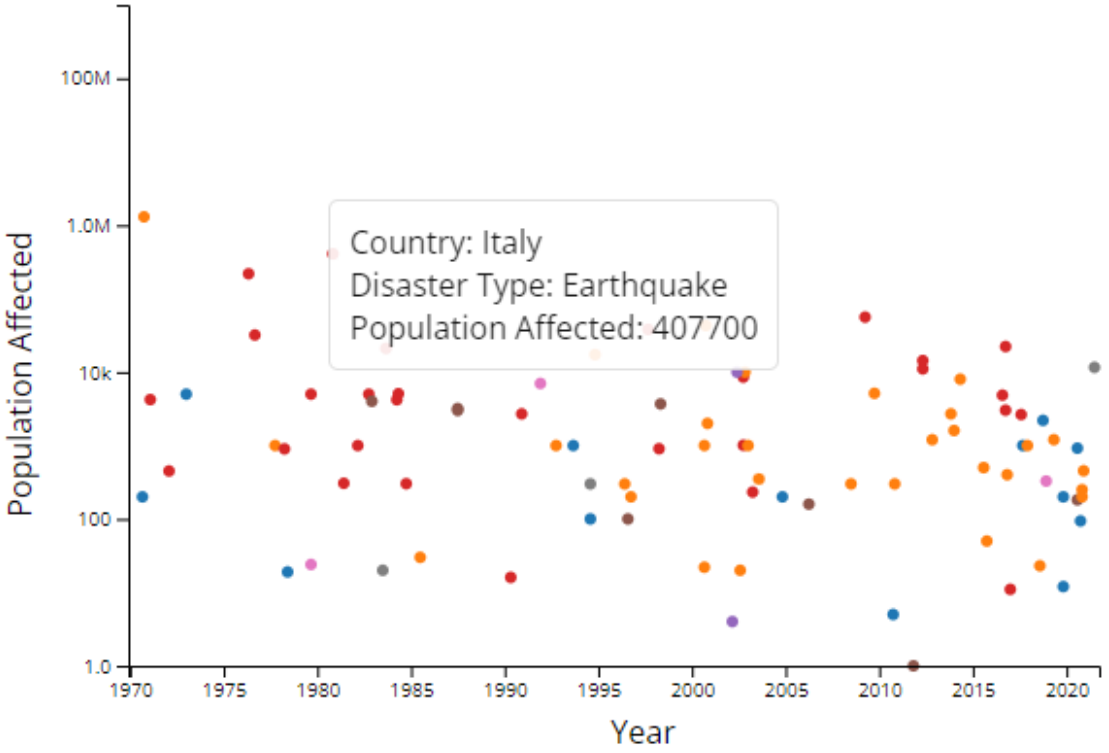
Types of Disasters

3. **Scatter Plot**:

The primary purpose behind this scatterplot is to visualize every single event that happened over the 50 year period. The scatter plot is plotted on year vs population affected. The population affected y axis was log scaled to provide more clarity. The dots in the scatter plot are color coded according to categorical disaster type



A tooltip is provided for users to get information about a singular even by hovering over a dot in scatter plot. Furthermore, filtrations are included, to filter out events based on country and disaster type.



Evaluation:

The correlation between the frequency of natural disasters and their impact, measured in terms of death rates and financial damages, does not exhibit clear patterns. Surprisingly, the number of disasters does not necessarily correlate with the rates of death or financial losses. A notable observation is that relatively less developed countries often experience higher death tolls, even if they encounter fewer natural disasters. This suggests that factors beyond the sheer occurrence of disasters, such as infrastructure, preparedness, and healthcare systems, play a crucial role in determining the actual impact.

When examining the countries most affected by various types of natural disasters, China and the USA emerge as frequent targets. The geographical location, population density, and other influential factors contribute to their susceptibility. India, for instance, faces recurrent droughts that have affected an astonishing 300 million people. Similarly, China contends with a history of devastating floods that have impacted millions, highlighting the disasters and their global distribution.

Analyzing the trends in the total number of people affected by natural disasters over the years reveals interesting insights. China, for instance, consistently experiences a similar trend in the occurrence of floods. This suggests that the country might be prone to certain types of disasters, possibly due to geographical and climatic factors.