CS/INFO 3300; INFO 5100
Project 2
Due 11:59pm Thursday April 20

*You are encouraged to get data and inspiration from other sites. Make sure you acknowledge these in comments and in your written description. Any code that you did not write yourself (such as d3) should go in a separate .js file.

Your final submission has two parts, a d3-based interactive data visualization (85 pts) and a written description of your visualization (10 pts).

A. A description of the work done by each team member. (Consider this your final status update.)
Jean Maldonado Caizapanta
- Finalized the datasets
- Cleared and formatted the datasets
- Integrated 2 different datasets
- Finalized the visualization and interactive features
- Implemented visualization through code (map)
- Integrate with other's code

Yang Qin
- Finalized the datasets
- Cleared and formatted the datasets
- Integrated 2 different datasets
- Finalized the visualization and interactive features
- Implement visualization through code (chart)
- Integrate with other's code

Po-Yuan Chang
- Finalized the datasets
- Cleared and formatted the datasets
- Integrated 2 different datasets
- Finalized the visualization and interactive features
- Implement visualization through code (chart)
- Integrate with other's code

B. A description of the data. Report where you got the data. Describe the variables. If you had to reformat the data or filter it in any way, provide enough details that someone could repeat your results. If you combined multiple datasets, specify how you integrated them. Mention any additional data that you used, such as shape files for maps. Editing is important! You are not required to use every part of the dataset. Selectively choosing a subset can improve usability. Describe any

criteria you used for data selection.

- Data was obtained through NOAA National Centers for Environmental Information (https://www.ngdc.noaa.gov/hazard/tsu_db.shtml)
- The website provides information on tsunamis, earthquakes, and volcanoes. Data was downloaded for years 1850-2017. Files for each natural disaster provides information on time, causes, location (latitude & longitude), parameters (magnitude), and effects (deaths, damages, etc.). Data was downloaded for the locations of tsunami origins and its runups showing impacted regions.
- Originally, we planned to visualize 4 different natural disasters (including hurricanes/typhoons). Tsunamis, earthquakes, and volcanoes were selected due to how interrelated these natural disaster were. However, after initial visualizations we found that earthquakes were the main cause and showing locations did not show anything interesting since the origin of tsunamis and it's corresponding cause were pretty much on the same point in a map. Therefore, we chose to investigate tsunami activities and their runups (impacted locations). We still considered tsunami causes but not on a detailed level. The tsunami runup data was also obtained from the aforementioned website. The tsunami origin data was integrated with the runup data using a tsunami ID which was available in both files. Given that the tsunami files had many fields, only the following were retained:
    - **Tsunami origins:** ID, YEAR, CAUSE_CODE, FOCAL_DEPTH, PRIMARY_MAGNITUDE, COUNTRY, LATITUDE, LONGITUDE, MAXIMUM_WATER_HEIGHT, DEATHS, DAMAGE_MILLIONS_DOLLARS
    - **Tsunami runups:** I_D, TSEVENT_ID, YEAR, LATITUDE, LONGITUDE
- We only wanted to visualize tsunamis that resulted in considerable damages, so tsunamis that caused less than 50 deaths were filtered out. We also filtered out tsunamis prior to 1850.

C. A description of the mapping from data to visual elements. Describe the scales you used, such as position, color, or shape. Mention any transformations you performed, such as log scales.

Timeline histogram

- The total number of major tsunamis was plotted onto a horizontal histogram divided into 25 year segments from 1850-2017. Hovering over each of these bars would provide a summary within that timeframe and reveal a change in color indicating interactivity.

Bubble chart

- Once a time frame is selected, the bubble chart will update showing tsunamis within that period. Data for number of deaths is shown on the bubble chart based on the size of each bubble as indicated in the legend (plotted on a log scale). In addition to visualizing deaths data by size, additional information on date, damage ($), and location are tagged to each bubble, appearing when users hover over bubbles. Data for causes of tsunamis (earthquakes, volcanoes, or others) are represented by the color of each bubble and labelled accordingly in the legend.

Map

- The locations of each tsunami is provided in both datasets with longitude and latitude numbers. These locations are plotted on the map in two different colors. One color represents the origin locations of major tsunamis; the other color shows locations impacted by these tsunamis. An animation shows the traveling of tsunami impact locations from the origin.

D. The story. What does your visualization tell us? What was surprising about it?

- Our visualization allows users to explore some of the deadliest tsunamis in the past two centuries through interacting with 3 visual elements: world map, bubble chart, and timeline. The timeline provides an overview of the total number of major tsunamis segmented every 25 years from 1850-2017. The timeline is divided into 7 segments. Hovering over each segment will display details of that time frame. Once a segment is clicked, the bubble chart and map will update with further details on each of these major tsunamis. (Note: major tsunamis are defined by ones that have caused 50 or more deaths) Next, the bubble chart compares different major tsunamis within the selected time frame. Hovering over each bubble provides details of that specific tsunami, including year, deaths, economic damages, and location of origin. Clicking on a bubble will update the map to specifically visualize that tsunami. The map will display the origin location of the selected tsunamis and the areas it impacted.
- It is interesting to see how most locations of these major tsunamis center around the Ring of Fire, which correlates directly with a majority of them being caused by earthquakes. Additionally, the death numbers and economic losses of some of the largest tsunamis are astonishing.

Libraries used:

- Tooltip.js - used to show tooltips next to histogram bars and bubble chart bubbles
- https://github.com/vlandham/gates_bubbles - baseline code for bubble chart