

## Project 2 Presentation Outline

**Group 3:** Melodie Vines, Tiffany Teasley, Stefanee Richardson, and Mabel Duran-Sanchez

### Project Theme:

Motivated by the [recent magnitude 5.1 earthquake](#) which struck North Carolina on August 9th, 2020, we decided to take a look back in history and explore the frequency, magnitudes, and locations of these natural events.

After reviewing several datasets, we came across a dataset collecting data on significant earthquakes around the world since 1600 from the [U.S. Department of Commerce, National Oceanic and Atmospheric Administration](#) available through [Kaggle](#). To view the dataset we utilized for this project, click [here](#).

### Coding Approach:

Python Flask powered RESTful API, HTML/CSS, Javascript, and MongoDB were utilized for the visualization. A dashboard was created with a heatmap and charts to best show the earthquake data. The user is able to apply filters to the data and the charts are updated upon their selections.

There were 2 Python files. In the first python file (datacleanup.py), preparation of data occurred including: cleansing of the data, renaming a column and exporting the CSV file into a JSON file to be utilized by MongoDB.

Within the second Python file (app.py), there are 2 different routes. One route pulls all of the data from MongoDB to send to Javascript upon the call to the route. The second route applies filters to the data based on user selections. When a user selects one or more filters on the HTML page, Javascript pulls those values and sends them in a request to Python. In Python, using MongoDB the filters are applied and the filtered data is then returned to Javascript. Regardless if the user selects one or multiple filters, the logic can pull the data and send the filtered data set to Javascript.

In Javascript (logic.js), upon reading the data from the call the dropdowns are filled in based on the data received. The data elements needed for the visualizations are stored in arrays and then rendered to the charts on the HTML page.

### Data Munging Techniques:

1. **Data Exploration:** We decided to keep the dataset almost intact given that significant amounts of information were missing, particularly prior to the 1900s. For the "Name" column which provides earthquake location information, we decided to only keep the location name prior to the ":" which most often is the country, although in some instances it refers to the state.
2. **Data Transformation:** Transforming the entire CSV file into a JSON. We discovered that the only columns of interest for our analysis included: "Year", "Month", "Name"

(which later for the purposes of our visualization in html we renamed “Location”), “Magnitude”, and “Deaths” (which later for the purposes of our visualization in html we renamed “Earthquake Related Deaths (Yes/No)”).

3. **Data Enrichment:** The team researched earthquakes to better understand the Richter Scale and discovered that the scale consists of six different levels ranging from “Minor” to “Great” according to ascending numerical values on the Richter Scale. For instance, “Minor” earthquakes range from 3-3.9 and those considered “Great” range from 8 and above.

**New JS library utilized:** anime.js for titles and headers of each of our pages.

- Anime.js is a lightweight JavaScript animation library with a simple, yet powerful API.
- It works with CSS properties, SVG, DOM attributes and JavaScript Objects.
- Anime.js was used to animate the header of each page. The class letter was used to animate the letters in the word EARTHQUAKE. The animation.css file was used to create the class with the margin attributes. The javascript file included the rotation, translation and speed of the text animation.

#### **Final Visualizations:**

1. **Heatmap:** Our heatmap titled *Earthquake Magnitude* helps visualize earthquake magnitudes around the world. The legend on the right-hand side of the heatmap showcases the magnitude types according to the Richter Scale in ascending order (from dark purple to bright yellow). From the data provided in our dataset, it is evident that the majority of earthquakes over the time period in consideration (1600 - 2020) are above 8 and towards the higher end of the Richter Scale, hence the prevalence of bright yellow tones throughout the map.
2. **Scatter Plot:** Our scatter plot titled *Earthquake Count by Year* shows the number of earthquakes recorded per each year within our dataset. The more we approach the 21st century, the more earthquake occurrences which are also observed. This is most likely due to the vast advancements in technology which have allowed for more precise monitoring and recording of earthquake events.
3. **Histogram:** Our histogram titled *Earthquake Magnitude Count* showcases the total number of earthquakes per Richter Scale magnitude class during the time period under study (1600 - 2020). According to the dataset, earthquakes at 7.5 magnitude seem to be the most frequent throughout this time period with 250 counts. The second most frequent Richter Magnitude class is 6 with 198 counts of occurrences.
4. **Stacked Bar Chart:** Our stacked bar chart titled *Earthquake Count by Year and Magnitude* combines the analysis from the scatter plot and the histogram and brings it together into a single visualization. This chart showcases the number of earthquakes throughout the time period under study (1600 - 2020) according to each of the Richter Scale magnitude classes while also counting each of those occurrences.