Lab_07

Fantastic Four

Who is Hiring us?

Fantastic Four is being hired by the American Red Cross in Honolulu, Hawaii to look at tsunami data. There are two reasons that this organization is hiring the team. First, they would like to map out all the tsunamis in the Pacific in a creative way using the data we have. This could be economic impact of more recent tsunamis (since that part of the data is available) or the relationship between magnitude and source of the waves. Second, they would like to know if there is a better way to predict, educate, and minimize impact to society for islands in the Pacific.

Questions and Importance

There are two questions we have set out to answer. The first is:

What are the patterns in tsunamis that come to Hawaii?

The reason we are looking at this question is to find the patterns of tsunamis that come to Hawaii. This is important because it can help us better prepare the community for impact (via insurance, warning, education) and furthermore minimize the damage to our beloved Hawaiian Islands.
The second question to look at is:

What are the patterns in tsunamis in the Pacific and how to they relate to ones recorded in Hawaii?

This question is important because by being able to see relationships between waves in Hawaii and other places in the Pacific, we can put together our resources in order to better prepare and educate society in order to minimize impact.

Dataset to Analyze

The dataset Fantastic Four will be using will be a tsunami dataset from

`https://www.kaggle.com/noaa/seismic-waves/data`

This is an NOAA dataset of tsunami data for recordable history. A preview of the column headers of the raw data is the following:

```
## # A tibble: 2,582 x 45
## # ... with 2,582 more rows, and 45
## #   variables: SOURCE_ID <chr>,
## #   YEAR <chr>, MONTH <chr>, DAY <chr>,
## #   HOUR <chr>, MINUTE <chr>,
## #   CAUSE <chr>, VALIDITY <chr>,
## #   FOCAL_DEPTH <chr>,
```

```
## #   PRIMARY_MAGNITUDE <chr>,
## #   REGION_CODE <chr>, COUNTRY <chr>,
## #   `STATE/PROVINCE` <chr>,
## #   LOCATION <chr>, LATITUDE <chr>,
## #   LONGITUDE <chr>,
## #   MAXIMUM_HEIGHT <chr>,
## #   MAGNITUDE_ABE <chr>,
## #   MAGNITUDE_IIDA <chr>,
## #   INTENSITY_SOLOVIEV <chr>,
## #   WARNING_STATUS <chr>,
## #   MISSING <chr>,
## #   MISSING_ESTIMATE <chr>,
## #   INJURIES <chr>,
## #   INJURY_ESTIMATE <chr>,
## #   FATALITIES <chr>,
## #   FATALITY_ESTIMATE <chr>,
## #   DAMAGE_MILLIONS_DOLLARS <chr>,
## #   DAMAGE_ESTIMATE <chr>,
## #   HOUSES_DAMAGED <chr>,
## #   HOUSE_DAMAGE_ESTIMATE <chr>,
## #   HOUSES_DESTROYED <chr>,
## #   HOUSE_DESTRUCTION_ESTIMATE <chr>,
## #   ALL_MISSING <chr>,
## #   MISSING_TOTAL <chr>,
## #   ALL_INJURIES <chr>,
## #   INJURY_TOTAL <chr>,
## #   ALL_FATALITIES <chr>,
## #   FATALITY_TOTAL <chr>,
## #   ALL_DAMAGE_MILLIONS <chr>,
## #   DAMAGE_TOTAL <chr>,
## #   ALL_HOUSES_DAMAGED <chr>,
## #   HOUSE_DAMAGE_TOTAL <chr>,
## #   ALL_HOUSES_DESTROYED <chr>,
## #   HOUSE_DESTRUCTION_TOTAL <chr>
```

This data contains date, magnitude, location, depth, time, and damage attributes for tsunamis dating back to 2000 BC. This data set needs tidying for the following reasons:
* The region we are looking at is the Pacific
* There are a LOT of values are N/A and we need to evaluate/get rid of tsunamis that do not give significant data.
* For economic impacts, it would be important to look at more recent data, which would be more realistic
* Simplify columns (total damage, date/time, etc.)

Subsequent Questions and Importance:

- How has Tsunami affected an island? Fatalities, house damages, etc.
- Where is the most common occurrence of Tsunami's? Does maximum height have a correlation with how much damage occurs?
- Is there a relationship between time of year a tsunami occurs and how much damage is inflicted?
- What is the relationship between the tsunami source and the wave, and how does the distance affect the tsunami severity and amount of damage? This is significant in analyzing where tsunamis cause the

most damage, which could indicate areas that are at greater risk of being impacted by a tsunami. It could possibly lead to formulating predictions of where future tsunamis may occur and how intense they will be. This could help with pre-storm evacuations and safety preparation which would limit injuries, possible deaths, and destruction. Which in turn would limit resource and economical strain following tsunamis.

- How the warning status relate to the damage to people and environment? This is important because predicting an catastrophe such as tsunami, earthquake and so on will be an cutting-edge technique for people to minimize their losses.

Contributions

Contributions will be added in the Lab 8 writeup since this is a two-part lab.