Tectonic Plate Visualization via Earthquake, Tsunami, and Volcano Data

Alex DeGroat, Clayton Kramp, Jeffrey Stenerson, Allee Zarrini

November 21, 2018

1 Introduction and Motivation

All around the world, natural disasters occur very frequently and have caused a lot of destruction and death among the human population. Majority of the natural disasters occurs in certain locations around the world with a connection to the tectonic plates, but majority of the population does not really know where the tectonic plates are located and how often the natural disaster occurs around them. The team wanted to analyze the results and display the location of the natural disasters that have occurred. The user will be able to identify the location of the tectonic plates and visualize the impact these plates have on natural disasters.

There are variety of data sets that have been recording natural disasters of all types around the world. The team decided to focus on three specific natural disasters: earthquakes, volcanoes, and tsunamis. The team have found three separate data sets from Kaggle, each for a single natural disaster, which identifies the date, location and the size of the disaster. We will use this data along with a mapping visualizer to create geographical projections to show where each of these disasters occur.

The data will allow insight onto particular disasters that have occurred around the world. For example, the team decided to look more closely in the Fukushima Nuclear Disaster that occurred in Japan. With the data, we will be able to see where the earthquake started and the impacts the earthquakes had. The earthquake created a massive tsunami that was the main cause of destruction, which the team will be able to pinpoint the locations the tsunami hit around the area.

The team will also analyze the ring of fire, which is known to have the most earthquakes and volcanoes than any other place in the world. The ring of fire is a single tectonic plate that is moving under the surface of the earth and when it moves, it collides with other tectonic plates that results in earthquakes and eruptions of volcanoes. The team will be able to identify that the ring of fire, is in fact the area that earthquakes are most abundant.

2 Data

2.1 Earthquake

The "Significant Earthquake" data set identifies the date, time, geographical coordinates, and the magnitude of the earthquake. The data set also contains more information that is not necessarily needed for our purposes, therefore the team created a select query that obtained the data that is needed. There were several events that the data set contained that didn't have a specific geographical coordinates and magnitudes, therefore, the team had to remove these events from the data set. This data set was created by the US Geological Survey and was published to Kaggle for educational purposes. The data set looked at earthquakes from 1965 to 2016.

2.2 Tsunami

The "Tsunami" data set identifies the date, time, country, and geographical coordinates of the event, and the magnitude and intensity of a tsunami. The data set also contains more information that is not necessarily needed for our purposes such as the damage that resulted from the tsunami. This data would be very useful, unfortunately majority of the tsunami has no information for this particular data, therefore the team had to remove it from the select query. There were several events that the data set contained that didn't have a specific geographical coordinates and magnitudes, therefore, the team had to remove these events from the data set. The data set has been created by NOAA (National Oceanic and Atmospheric Administration) and was published to Kaggle for educational purposes. NOAA has been collecting data since 2000 BC.

2.3 Volcano

The volcano eruption data set identifies the date, time, name of volcano, location of volcano with both geographical coordinates and the country the volcano is located in, and the VEI. The VEI is the volcanic explosively index that the team will use in order to determine the magnitude of the eruption. The data set also contains information about the destruction of the volcano, but since volcanoes was the only data set that had the data and had the least amount of information, the team decided to remove the data from the data set. The data set has been keeping tracks of all eruptions starting in 2010 up until today in 2018. This data set has been published to Kaggle for educational purposes.

2.4 Data Availability

The data is loaded onto `ckramp`'s schema for use.

3 Technologies Used

3.1 PostgreSQL

After we downloaded the `csv` files containing the relevant data, we used `PostgreSQL` to pre-process the data. We first constructed raw tables which contained all the columns originally provided by the data sets. Then, we constructed smaller tables from the raw data that only contains relevant data. We added an ID column to be the primary key as no good candidates for primary keys existed in our refined tables. We also deleted rows missing important data (latitude and longitude) as well as updated rows that had `NULL` values for the magnitude column (default to 1). Finally, the data from the processed tables were exported to a `csv` file.

3.2 Python

The post-processing was done using `python3`. We first opened and read in the necessary information from the `csv` files, and constructed a large dictionary that creates a dictionary for each row in the data. This way, we can easily access columns as latitude, longitude, and magnitude.

3.3 Folium

The visualization is completed on a `python` package called `folium`. `Folium` is a package that allows us to visualize data using the `leaflet.js` mapping library. Thus, we initially create a map using `folium` and after reading in each line of the data, we add a bubble to our map at the corresponding latitude, longitude, and radius dependent on the magnitude value. We then finaly save the map as a `.html`.

4 Visualization

4.1 Earthquake

For visualization purposes, the radius of each circle is proportional to the magnitude of the earthquake. Thus the wider the circle, the stronger the earthquake. Magnitude values ranged from 5.5 to 9.1.

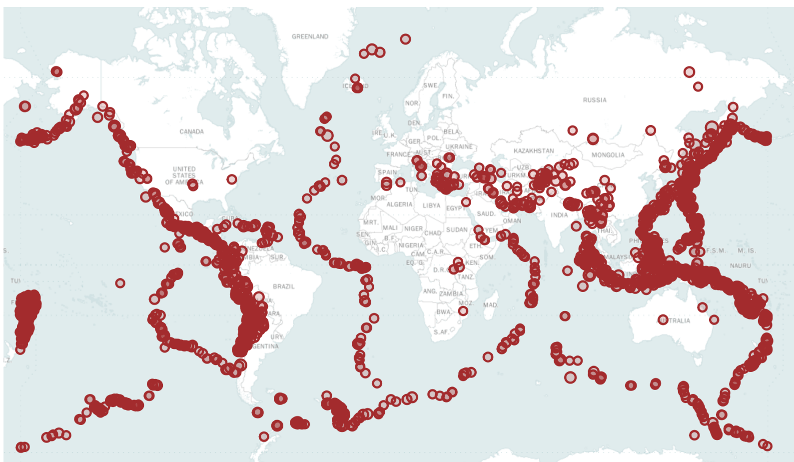

Figure 1: Earthquake data visualized from 2010

4.2 Tsunami

For visualization purposes, the radius of each circle is proportional to the magnitude of the tsunami. Thus the wider the circle, the stronger the tsunami. Magnitude values ranged from 1 to 9.5.

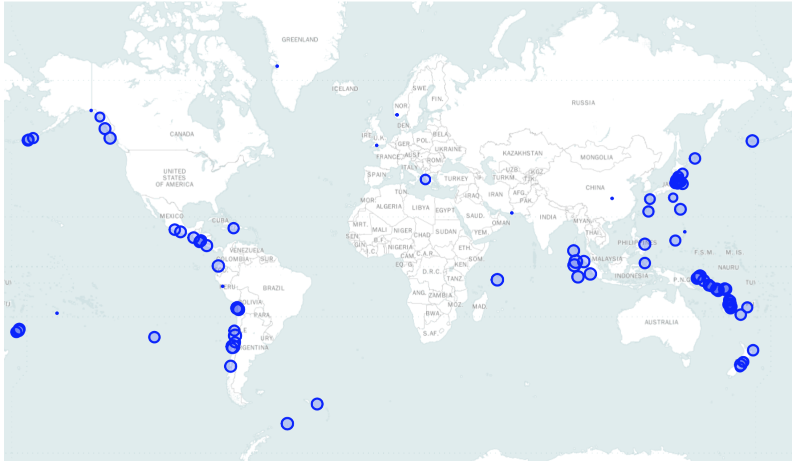

Figure 2: Tsunami data visualized from 2010

4.3 Volcano

For visualization purposes, the radius of each circle is proportional to the volcano explosive index (VEI) of the volcano. Thus the wider the circle, the more destructive the volcano. VEI values ranged from 1 to 4. As volcano are more sparse and do not occur as frequently as earthquakes, we added a multiplier to the volcano radius so that the data is more visible when overlayed with the other datasets.

Figure 3: Volcano data visualized from 2010

The volcano dataset was the least significant of the sets with have, because of the duration of data collected. Looking at Figure 3, we can see that majority of the volcanoes lies in Indonesia, which makes sense because Indonesia has more active volcanoes than any other countries. Indonesia contains 76 active volcanoes and had around 1,171 dated eruptions. This is due to the Asian tectonic plate that resides in the Indian Ocean (Simkin and Siebert).

5 Case Study/ Analysis

5.1 Fukushima Nuclear Disaster

For our first case study, we examine the Fukushima Daiichi Nuclear Disaster. On March 11, 2011 a 9.1 magnitude earthquake struck 130 km offshore of of the city Sendai in the Miyagi prefecture. This large earthquake was quickly followed by 15-meter tsunami that struck the Fukushima Daiichi Nuclear Power Plant. This disabled the cooling and power supply of the reactors and caused a nuclear accident. The following figures depict the earthquake and subsequent tsunami that occurred:

Figure 4: 3/11/2011
Fukushima Daiichi Nuclear Disaster

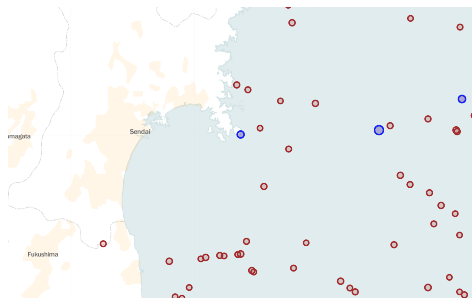

Figure 5:
Tsunami and Earthquake Data near Fukushima

We can obtain the information regarding this tsunami with the following query:

```
select magnitude, latitude, longitude from tsunami where
year = 2011 and month = 3 and date = 11 and country = 'JAPAN';
```

This retrieves the following response:

magnitude	latitude	longitude
9.1	38.297	142.372

The earthquakes that occurred in the similar area on March 11, 2018 can also be retrieved with the following query:

```
select magnitude from earthquake where
date = '3-11-2011'::date and latitude between 38 and 39 and longitude between 142
    and 143;
```

This retrieves the following response:

magnitude
9.1
6.2
6.2
5.8
6
5.9

We can clearly see that many powerful earthquakes occurred as a consequence of the initial strong hit. This visualization technique can help avoid future disasters as we can project and predict where major earthquakes, tsunamis, and volcanoes hit. In this way, we can examine locations that are low in risk of such natural disasters based off of past data. If Japan were to construct a new nuclear power plant, with our application we could suggest building one along the coast of the Yamagata prefecture, as the risk for earthquakes and tsunamis are substantially lower, and it still has the same access to water for cooling purposes. Yamagata prefecture is the prefecture neighboring the Miyagi prefecture (with its capital city of Sendai) to the left.

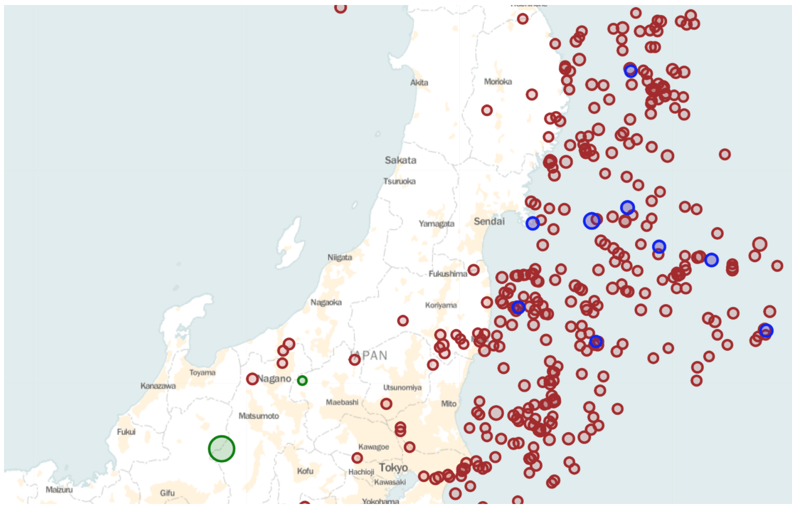

Figure 6: Tsunami and Earthquake Threats

5.2 Ring of Fire

For our second case study, the team examined the Ring of Fire and how it relates to earthquakes, more specifically the higher magnitude earthquakes. The Ring of Fire is known to have the most number of volcanoes and earthquakes, but as well as the strongest earthquakes. From the figures shown below, we can see that there is a massive tectonic plate the lies in the Pacific Ocean.

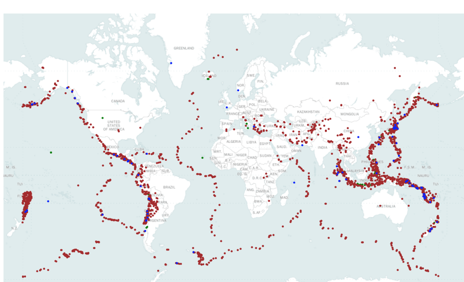

Figure 7: All data displaying the lining of tectonic plates

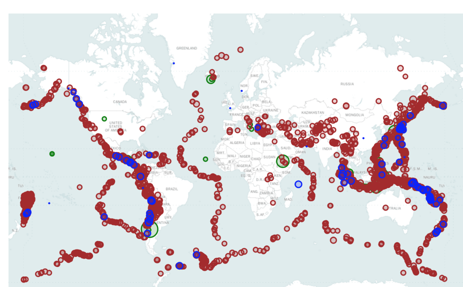

Figure 8: All data put together with magnitude

This image shows all of the volcanoes, tsunamis, and earthquakes from the data set. This clearly shows the location and size of each of the tectonic plates around the world. This can be proven by the following figure, that shows the actual outlines of the tectonic plates.

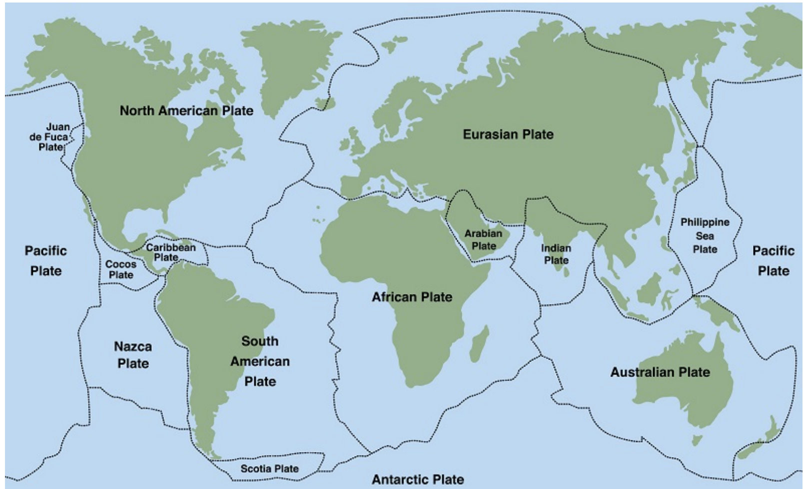

Figure 9: Real tectonic plates

Further into our analysis, the team wanted to determine where the higher magnitude of earthquakes occur and it relates to a specific region. In order to determine this, the team had to do several query in order to obtain the data that the team wanted. The team obtained the earthquakes that had magnitudes greater than 8 with the following query:

```
select magnitude, latitude, longitude from earthquake where magnitude >= 8;
```

The team was received with data of around 50 different earthquakes that occurred with a magnitude greater than 8. This data then was displayed on Folium to visualize the data on a map, which is shown below:

Figure 10: Earthquakes with magnitude greater than 8

The figure doesn't really show a clear trend, but the team noticed something happening odd around the Pacific Ocean. Moving the map to focus on this particular region, the following figure is produced:

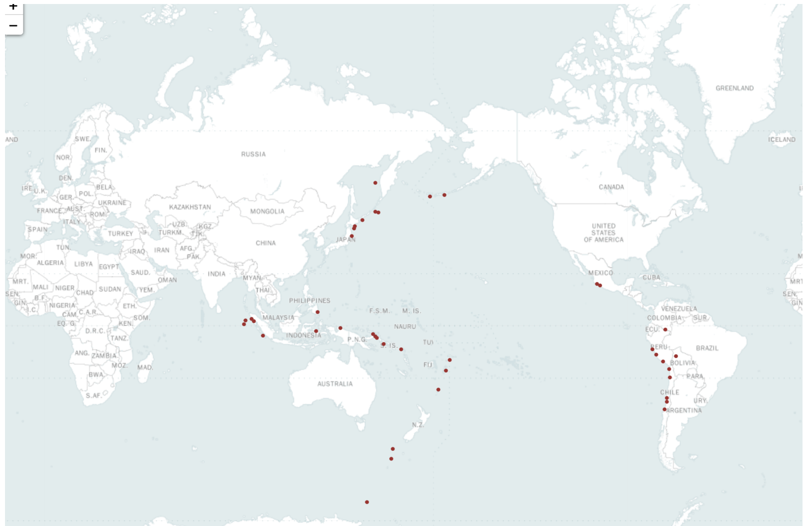

Figure 11: Earthquakes with magnitude greater than 8

From the figure above, it clearly shows a trend around the tectonic plate located in the Pacific Ocean. This region is also known as the Ring of Fire, which the team predicted was going to have the most earthquakes in general as well as the earthquakes with the highest magnitude. The ring of fire is shown in the figure below to specifically identify the region, thus approving our hypothesis.

Figure 12: Ring of Fire

6 Technical Challenges

There were several technical challenges that occurred throughout our analysis. These challenges included modifying the data set to fit the project needs, researching map visualization libraries for python, and integrating everything together to produce proper visualizations. Research was done for several methods of visualizing the data. Because of its simplicity and ease of use, the team decided to use a Python library for visualizations. Folium was chosen because it was easy to learn and did not require a large amount of dependencies.

7 Conclusion

With the abundance of data available on natural disasters such as earthquakes, tsunamis, and volcanic eruptions, a proper analysis can seem daunting. Using modern technologies such as PostgreSQL and Python3, we were able to analyze a large amount of data and glean some interesting information about natural disasters. Using our analysis of the Fukushima Nuclear Disaster, we can predict an improved location for future Nuclear Power Plants. Using these predictions we can reduce the chances of a similar disaster happening again. Our Ring of Fire analysis can be used to visualize the most powerful natural disasters as well as the tectonic plates all round the world. Using modern database management techniques and Python visualizations, we were able to produce a useful analysis of a large amount of data.

8 References

Tectonic plate image:
`https://www.worldatlas.com/articles/major-tectonic-plates-on-earth.html`

Japan Earthquake image:
`https://www.power-technology.com/projects/fukushima-daiichi/attachment/fukushima-daiichi3/`

Ring of Fire image:
`https://www.nationalgeographic.org/encyclopedia/ring-fire/`

Data:
`https://www.kaggle.com/noaa/seismic-waves`
`https://www.kaggle.com/usgs/earthquake-database`
`https://www.kaggle.com/texasdave/volcano-eruptions`