SMALL SEISMIC EVENT P.F.

1. TEAM MEMBERS

1. Thomas, Prouty – A02302709 (Grad Student)

2. DATASET SELECTION

- **Dataset Name**: USGS Earthquake Catalog
- **Source of Data**:
  - https://earthquake.usgs.gov/earthquakes/search/
  - This dataset is public domain and can be accessed via python in the following method:
  -

```python
from obspy.clients.fdsn import Client

client = Client("USGS")

catalog = client.get_events(starttime="2024-10-01", endtime="2024-10-11", minmagnitude=3)
```

- **Dataset Description**:
  - Summarize the content of the dataset, including the type of data, number of records, and number of variables/features.
  - The USGS Earthquake Catalog contains seismic event data from around the world and is available for public use. To date, the catalog contains documentation for over 100,000 seismic events. Each event is uniquely identified and described by a variety of features. Generally, each events documentation contains 7 or more features which typically include:
    - Event ID
    - Event Type
    - Time
    - Latitude and Longitude
    - Magnitude
    - Depth
    - Uncertainty
  - The goal of this project is to analyze and produce probabilistic forecasts for small-scale seismic events, specifically those with magnitudes between 2.5 and 5. The USGS Earthquake Catalog is invaluable for that purpose as it contains much of the information what I wish to analyze as a part of this project. The ability to filter and sort the data by magnitude, location, and other key attributes will greatly help in preprocessing the data analysis and use.
- **Data Sample**:
  - The following image is a portion of the data contained in each events documentation:

```
Event: Event:   2024-10-10T23:11:24.670000Z | -17.942, -178.649 | 4.4  mb | manual

                    resource_id: ResourceIdentifier(id="quakeml:earthquake.usgs.gov/fdsnws/event/1/query?eventid=us6000nxzd&form
                     event_type: 'earthquake'
                  creation_info: CreationInfo(agency_id='us', creation_time=UTCDateTime(2024, 10, 11, 0, 19, 8, 262000))
             preferred_origin_id: ResourceIdentifier(id="quakeml:earthquake.usgs.gov/product/origin/us6000nxzd/us/1728605903040/p
        preferred_magnitude_id: ResourceIdentifier(id="quakeml:earthquake.usgs.gov/product/origin/us6000nxzd/us/1728605903040/p
                           ---------
            event_descriptions: 1 Elements
                       origins: 1 Elements
                    magnitudes: 1 Elements, Magnitude: 4.4
Origin
                  resource_id: ResourceIdentifier(id="quakeml:earthquake.usgs.gov/product/origin/us6000nxzd/us/1728605903040/produ
                         time: UTCDateTime(2024, 10, 10, 23, 11, 24, 670000)
                    longitude: -178.6485
                     latitude: -17.9425
                        depth: 547409.0 [uncertainty=7255.0]
                      quality: OriginQuality(used_phase_count=66, used_station_count=66, standard_error=0.68, azimuthal_gap=49.0,
           origin_uncertainty: OriginUncertainty(horizontal_uncertainty=12490.0, preferred_description='horizontal uncertainty')
              evaluation_mode: 'manual'
                creation_info: CreationInfo(agency_id='us', creation_time=UTCDateTime(2024, 10, 11, 0, 18, 23, 40000))
```

3. PROJECT OBJECTIVES

- **Research Questions**:
  - The primary question being examined in this project is whether seismic events between magnitudes 2.5 and 5 can be automatically probabilistic forecasted for a region given historical precedent, and if so, what methods work well.
  - A secondary question for the project is where magnitude 2.5 and 5 earthquakes occur most often and what landscape features are commonly found in those locations.
- **Goals and Outcomes**:
  - **Goal 1:** Can seismic events between magnitudes 2.5 and 5 be automatically probabilistic forecasted within a region using past seismic data? Specifically, the aim is to determine if it is possible to "predict" (via probabilistic forecasting) an earthquake between magnitudes 2.5 and 5 occurring near a specific location within a following year.
  - **Goal 2:** Determine key location features that could lead to the common occurrence of earthquakes between magnitudes 2.5 and 5.
  - **Impact:** If it is possible to automatically probabilistic forecast small-scale seismic events, then it may be possible to use these mappings to determine when larger-scale events are likely to occur. Alternatively, the information obtained from this project could be used to attempt to estimate the occurrence of past small-scale seismic events.

4. TIMELINE AND MILESTONES

- Week 1: Project Draft Proposal and Submission (Due 10/12/2024):
  - Develop and research a hypothesis/research question.
  - Find publicly available datasets that could be used to investigate the question
  - Create draft of proposal
  - Edit Project Proposal and submit on 10/12/2024
- Week 2-3: Initial Research and Planning (Due 11/05/2024)
  - Perform further investigation of dataset being used of the project
  - Determine what types of preprocessing will be needed in order to perform analysis
  - Make and save a file containing the pre-process data from the dataset
  - Submit data file to the professor by 11/05/2024
- Week 4-5: Perform Analysis on Data File (Due 11/25/2024)
  - Perform analysis of data contained in the data file
  - Determine if the analysis supports or contradicts the hypothesis
  - If contradiction perform extra research to see if this result matches other research in the field
  - Submit results of analysis by 11/25/2024
- Week 6: Create Drafts of Final Report and Presentation (Due 12/02/2024)
  - Create Draft of Final Report
  - Create Draft of Final Presentation
  - Have the drafts reviewed by a peer
  - Submit drafts to the professor by 12/02/2024
- Week 7: Final Report and Final Presentation (Due 12/09/2024)
  - Submit the Final Report
  - Perform Final Presentation
  - Complete by 12/09/2024

5. QUESTIONS FOR INSTRUCTOR

- What actions should the student take if there seems to be a discrepancy between the student's analysis and published work in the field of the student's research question?
- How long should the final presentations be?

- Should the dataset being used by the student be saved to their GitHub repository or does the code to access/create the dataset suffice?