Team Name: Frog 1

Project Name: Understanding Earthquake Strength

Members: Soor Hansalia, Hugo Liu, David Bickram, Arossa Adhikary

Github URL: https://github.com/hugoliu-code/Group27-Final

Datasets: Earthquakes

Report: Earthquake Report

Responsibilities and Roles:

- Soor Hansalia: Earthquake magnitude/depth relationship, Github Organization, Abstract, References, Earthquake Report, What We Learned
- Arossa Adhikary: Earthquake magnitude/significance relationship, Trials & Tribulations sections, map of locations for earthquakes greater than 4.5 magnitude, report editing
- Hugo Liu: Earthquake/Longitude relation, Github Creation, data organization of Corgis .data/.py files (into accessible pickle files), Overall Experience Section, Changes we would Make Section
- David Bickram: Normal Distribution: Earthquake Magnitude, Finding Datasets

Overall Experience:

The overall experience was smooth and interesting. The team was able to schedule all necessary meetings relatively easily and group work sessions were easygoing and fruitful; we had no issue collaborating and helping each other. Additionally, we were able to use various resources related to the tools/libraries we used to streamline our learning process, including: youtube videos, documentation, and online tutorials. Since we decided to focus on learning our tools well first, we were able to seamlessly use them together to create our visuals and data manipulation. This was many of our members' first interaction with python's data libraries, so coming to meaningful conclusions was an eye-opening experience to the data capabilities of python.

Trials & Tribulations:

The largest challenge we faced surrounded one main idea: novelty. The newness of this project made it difficult to begin traversing its waters. The data visualizations required us to use skills that were not taught in class, such as matplotlib and panda libraries. In the beginning, we were

hesitant to choose the data analysis project as the guidelines were open-ended, too; however, we craved the uniqueness of the task at hand and decided to dive in. Looking through the thorough datasets, we were overwhelmed with the amount of data processing in our near future. The earthquakes.py file that was given to us from CORGIS was bombarded with information; the code we saw was a true application of the fundamental skills we have learned in class especially in relation to nested lists/dictionaries and exposure to a variety of data types. It took a group effort to be able to map through the data and truly understand how to relate the data points to one another. It was difficult to filter and clean the data, as well. We were unsure what data would be useful to our group, too. Resultantly, it required hours of discussion to determine what our hypotheses were and how they related to the data. To navigate through this, we made a correlation matrix to determine how all the variables were related to one another. The matrix paved a clear pathway for discovering in the future how these concepts are related. Another significant challenge we encountered was trying to convert longitude and latitude data into countries. We used the Geopy module.geopy to get the country that pertains to a certain latitude and longitude. The code was not difficult to figure out rather the part we were unable to loop around was the runtime. There are over 8000 rows of data for the program to run through and we could not find a way to decrease the runtime of the program. Instead we opted to decrease the size of the data we were looking at; for example, when discovering the relationship between Magnitude and Significance of an Earthquake, the dataset was filtered to find values only with a magnitude greater than 5.5, meaning there is at least “slight damage to buildings and other structures” (Michigan Technological University). Once we shortened the dataset, the processing time was quicker however a new problem arose with the module’s ability to identify the coordinates within the earthquake dataset. To overcome this issue, we manually input each coordinate pairing into another online coordinate to address the converter and mapped each figure accordingly.

Wisdom Gleaned:

If we were to approach this project again, we would first establish a clearer vision when it comes to a greater goal of the project. Much of our project started separately with each member learning individually before we came together to compile and deliberate. Because of this, however, we had members work with data that had little relation to each other and this eventually led to some

of our work being discarded. Additionally we would establish clearer git best practices that all members would strictly follow. We ended up pushing many miscellaneous files to our repository, so much so that we made a new one to reorganize the project. Such best practices may include: naming conventions for files, communication for what files we were working on, file organization, as well as making sure every member was confident working with git.

What We Learned:

- Hugo Liu: I learned how to clean and format data using Numpy and Pandas, as well as how to visualize it using Matplotlib. Specifically I modified dataframes and displayed them in scatterplots and polar scatterplots.
- Arossa Adhikary: I enjoyed discovering Python's capabilities in a new light with data analysis. Specifically, I was most enthusiastic about learning how to navigate around Pandas and understanding the importance of dataframes. I also absorbed new knowledge of coding in a team and thus how to use GitHub in various methods to ensure all the group members' code was saved. It was also great to learn how to code various types of graphs.
- David Bickram: I learned how to use pandas, scipy and numpy to navigate and manipulate raw data. I also learned how to use matplotlib to display the manipulated data.
- Soor Hansalia: I learned how to convert different types of data into usable dataframes in Pandas, use SciPy to perform some statistical analysis, and use MatPlotLib to display my results. I also learned how to properly pull, push and organize a GitHub with a group.

Resources:

- CORGIS Datasets Project. (n.d.). Think.cs.vt.edu. Retrieved December 6, 2023, from https://think.cs.vt.edu/corgis/python/
- for, S. (2022). Anderson-Darling Test for Normality. In YouTube. https://www.youtube.com/watch?v=VKLzYXpSzvY
- Holtz, Y. (n.d.). Map with markers with Python and Folium. The Python Graph Gallery. Retrieved December 6, 2023, from https://python-graph-gallery.com/312-add-markers-on-folium-map/

- Pandas. (2018). Python Data Analysis Library — pandas: Python Data Analysis Library. Pydata.org. https://pandas.pydata.org/
- Numpy. (2009). NumPy. Numpy.org. https://numpy.org/
- Matplotlib. (2012). Matplotlib: Python plotting — Matplotlib 3.1.1 documentation. Matplotlib.org. https://matplotlib.org/
- SciPy. (2020). SciPy.org — SciPy.org. Scipy.org. https://scipy.org/
- https://www.statsmodels.org/stable/index.html
- ektamaini ektamain Follow. (2023, January 10). *Python - coefficient of determination-R2 score*. GeeksforGeeks. https://www.geeksforgeeks.org/python-coefficient-of-determination-r2-score/
- Zach. (2021, October 5). *How to plot line of best fit in python (with examples)*. Statology. https://www.statology.org/line-of-best-fit-python/
- GeeksforGeeks. (2023, December 1). *How to sort pandas DataFrame?*. GeeksforGeeks. https://www.geeksforgeeks.org/how-to-sort-pandas-dataframe/
- *Machine learning - scatter plot*. Python Machine Learning Scatter Plot. (n.d.). https://www.w3schools.com/python/python_ml_scatterplot.asp
- *ANSS Comprehensive Earthquake Catalog (COMCAT) documentation*. U.S. Geological Survey. (n.d.). https://earthquake.usgs.gov/data/comcat/#tsunami
- Bressan, D. (2023, September 12). *New study shows updated map of Earth's tectonic plates*. Forbes. https://www.forbes.com/sites/davidbressan/2022/06/09/new-map-shows-earths-tectonic-plates-in-unprecedented-detail/?sh=5318a9433d72
- Michigan Technological University. (2021, October 4). *Earthquake magnitude scale*. https://www.mtu.edu/geo/community/seismology/learn/earthquake-measure/magnitude/