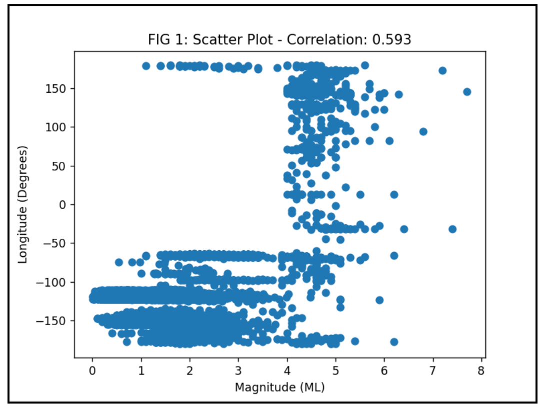
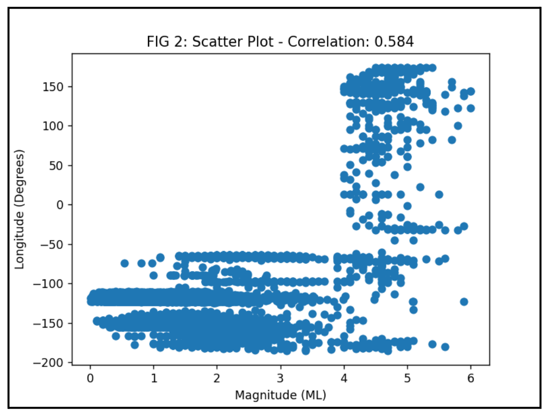
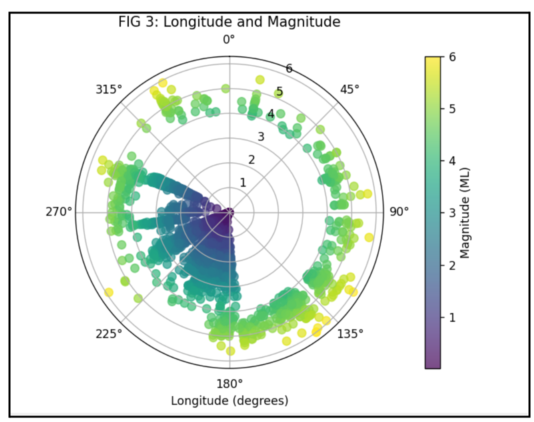
Abstract/Datasets: Our project examined different earthquake metrics; we aim to predict/realize the consequences of earthquake impact soon after occurrence. After selecting the dataset from CORGI's website, we curated hypotheses to give us a better understanding of the earthquake dataset. Our statistical analyses are described throughout this report, underscoring the various correlations found to enhance our understanding of earthquake behavior and the CORGIS dataset. CORGI's dictionaries of earthquakes were flattened into a 2D dictionary and converted into a .pkl file to simplify the data filtering process.

Earthquake Significance in Relation to Longitude





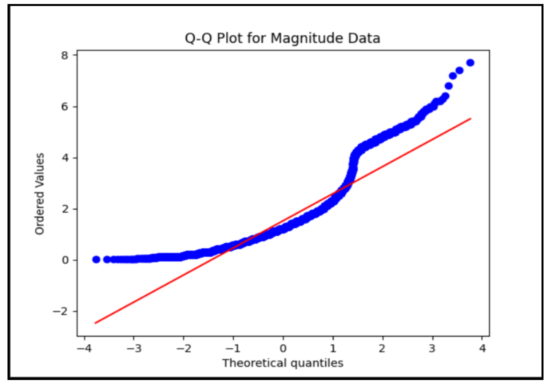


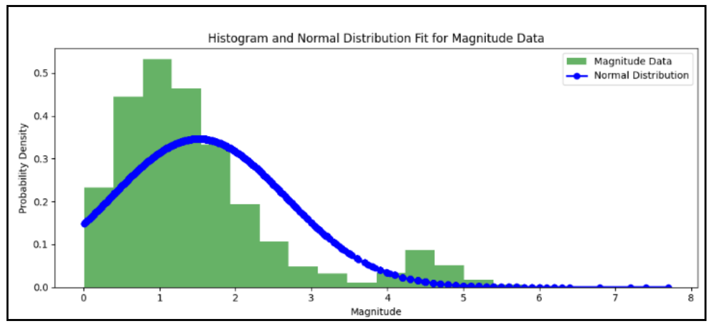
Dataset: Longitude and Magnitude pairs from earthquake dataset, describing the longitudinal location and Richter scale measurement of earthquakes.
Null Hypothesis: Longitude has no relation to magnitude.
Alternative Hypothesis: Longitude has a relation to magnitude.
Analysis: The data (FIG 1) immediately shows a clear pattern. There are two "blocks" of data determined by longitude. Two things stand out: some outliers in the upper magnitudes as well as a "slice" of data in the top of the data that looks like it belongs to the bottom "block".
I cleaned the data (FIG 2) by deleting outliers and shifting the top "slice" to the bottom (longitude is a periodic value; 170 longitude can also be represented as -190). Although the correlation dropped, it is not relevant for this kind of "group" relationship. It is clear that the longitudes from ~ negative 45 to ~ positive 175 are seemingly guaranteed to have a magnitude > 4, **(Conclusion) therefore we reject the Null Hypothesis that there is no relation between magnitude and longitude.** This is shown better in a polar scatterplot (FIG 3), which accounts for the periodic nature of longitude.

Normal Distribution: Earthquake Magnitude

Dataset: Magnitude column from the earthqauke dataset.
Null Hypothesis: Magnitudes of earthquakes follow a normal distribution.
Alternative Hypothesis: Magnitudes of earthquakes do not follow a normal distribution.

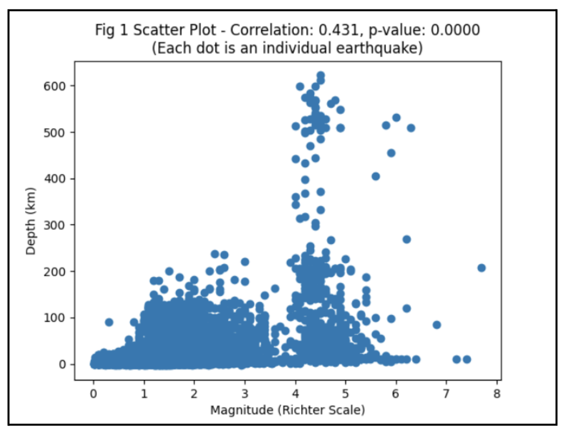
Analysis*:* The earthquake data is loaded from a pickle file into a pandas DataFrame which is subsetted to include only the 'magnitude' column. The Anderson-Darling test for normality is performed on the earthquake magnitudes at a 5% significance level using scipy and the test statistic and critical values are extracted from the result. A Q-Q plot is created using scipy to visually assess whether the magnitudes follow a normal distribution by comparing sample quantiles against theoretical quantiles expected under normality. The Q-Q plot is displayed using matplotlib. Using matplotlib, a histogram was created using the magnitude values and bins of size 20. Also, a normal distribution, which is fitted to the magnitude data, was overlaid on the same plot. This shows that the magnitude data (histogram) does not follow the normal distribution (blue line).





Conclusion: Since the Anderson-Darling Statistic, 383.34, is greater than the critical value, 0.787, the null hypothesis was rejected and the magnitudes were found to not follow a normal distribution.

Relationship Between Magnitude and Depth



Dataset: Magnitude (Richter Scale) and Depth (km) from dataset

Null Hypothesis: Magnitude and Depth possess a strong positive correlation

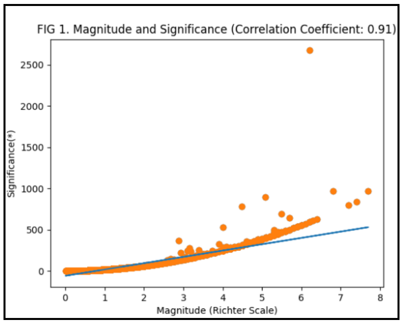
Alternative Hypothesis: The correlation between Magnitude and Depth is moderate or weak

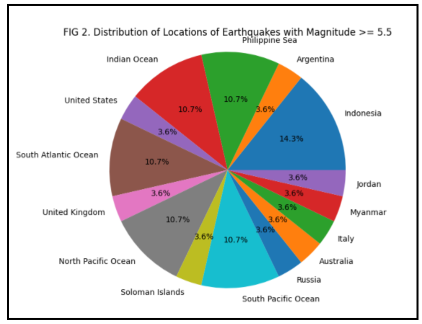
Analysis: An initial look at the graph on the right shows there is an insignificant relationship between magnitude and depth. The earthquakes are relatively evenly distributed between magnitudes of 0-6 and depths of 0-300km. There are some outliers of earthquakes around magnitudes of 4 and 5; such extraneous solutions could be explained by the small number of locations on earth that can accurately measure earthquakes at these extensive depths. Additionally, these results are surprising because the original belief was a strong positive correlation between magnitude and depth because deep earthquakes would have more energy from the earth's core. **(Conclusion) Since the p-value is less than 0.01 we**

reject the null hypothesis that r is greater than or equal to 0.75. This is reflected in the correlation calculation and the graphical representation above.

Relationship Between Magnitude and Earthquake Significance(*)

** Significance is measured on a multitude of factors including magnitude, maximum MMI, felt reports, and estimated impact (USGS)*





[MAP DEMO (LOCALLY HOSTED THUS, 3-SECOND VIDEO REQUIRED TO DEMONSTRATE INTERACTIVITY)]

Dataset: Magnitude (Richter Scale) and Significance(*) from CORGIS dataset

Null Hypothesis: Significance of an earthquake has a significant difference/correlation to magnitude of earthquake; most earthquakes are distributed close to areas with tectonic plates.

Alternative Hypothesis: Significance of an earthquake has a significant difference/correlation to magnitude of earthquake; most earthquakes are not distributed close to areas with tectonic plates.

Analysis: In relation to FIG 1. **(Conclusion) the null hypothesis was accepted as there is a strong positive correlation (0.91) between magnitude and significance**. The results make sense when applying physical science applications: the stronger an earthquake, there becomes a higher probability the event will cause damage within a community. The line of best fit further asserts such a comparison is valid as the data points closely reside along the line. The R^2 value of 0.8355 again proves such a dataset was apt to be compared together; it also signifies that 83.5% of the variability is explained by the model. Further investigation of such a relationship to determine causation would thus be viable and encouraged. From the analyses of FIG 1., another question was brought about regarding the locations of the earthquakes with the highest magnitudes and identifying whether their positions are of notable discussion. FIG 2. and the map explains how the majority of high impact earthquakes are found along a similar region of the world; such discoveries were made through comparing the Python-curated map to a map with plate boundaries (Bressan). **(Conclusion) Again, the null hypothesis was accepted that most earthquakes are distributed along tectonic plate boundaries.**