Metadata and Graphical Analysis

Author: Kush Paliwal

Input Data Source

https://earthquake.usgs.gov/earthquakes/feed/v1.0/summary/all_month.csv

The data has been downloaded from USGS Earthquake Hazard website as a csv file on 19th April 2020 at 12:36 PM (Eastern Time). The data file ("comma separated values" (CSV) ASCII text file) contains all the earthquakes which have occurred around the world in the past 30 days (from 03/20/2020 to 04/19/2020). The file contains 22 data sets such as time, latitude (degree), longitude (degree), depth (km), etc. The data types contain date, time, string, integer, float, etc.

In this case genfromtxt() does not work because the dataset contains different types of data.

Analysis Conducted with Python Script

- Read the data file "all_month.csv" by using read_table() function.
- Create a frequency histogram of earthquake magnitude (10 bins with width of 1 and a range of 0 to 10).
- Create a KDE plot of earthquake magnitude (Gaussian kernel type with kernel width 0.5).
- A scatter plot of latitude vs. longitude for all earthquakes.
- A normalized cumulative distribution plot of earthquake depths.
- A scatter plot of earthquake magnitude vs. depth.
- A Normal Quantile-Quantile plot of the earthquake magnitude.

Graphical Analysis

1. Frequency histogram of earthquake magnitude

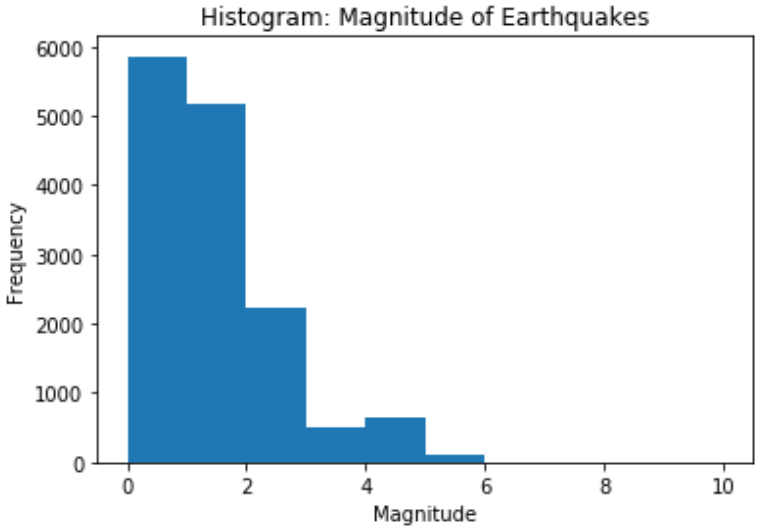

Figure 1 Histogram of Earthquake Magnitude

The width of the bins is 1 and the range is 1 to 10. When the bin width is decreased the plot will become smoother, but the negative aspect of a very small bin width might be gaps in the histogram. This histogram suggests that most of the earthquakes which have occurred in the past 30 days are of small magnitude.

2. KDE plot of earthquake magnitude

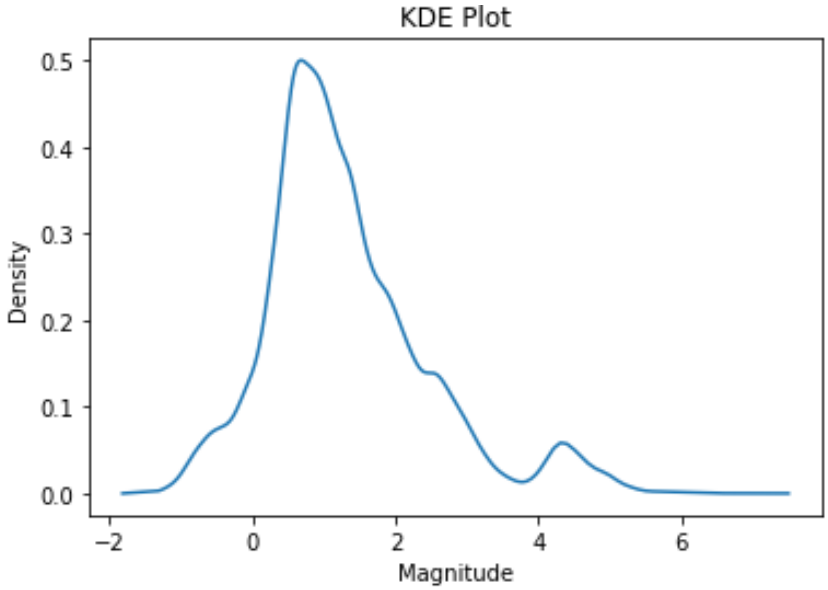

Figure 2 KDE Plot of Earthquake Magnitude

The kernel type is Gaussian, and width is 0.1. The Histogram and the KDE are showing the same distribution. The difference between the two is that the histogram is showing frequency of the earthquakes which have occurred in the past 30 days whereas the KDE is showing the probability density of these magnitudes. The KDE is a smoother version of the Histogram and might better match the true distribution function.

3. Scatter Plot of Latitude vs Longitude for all earthquakes

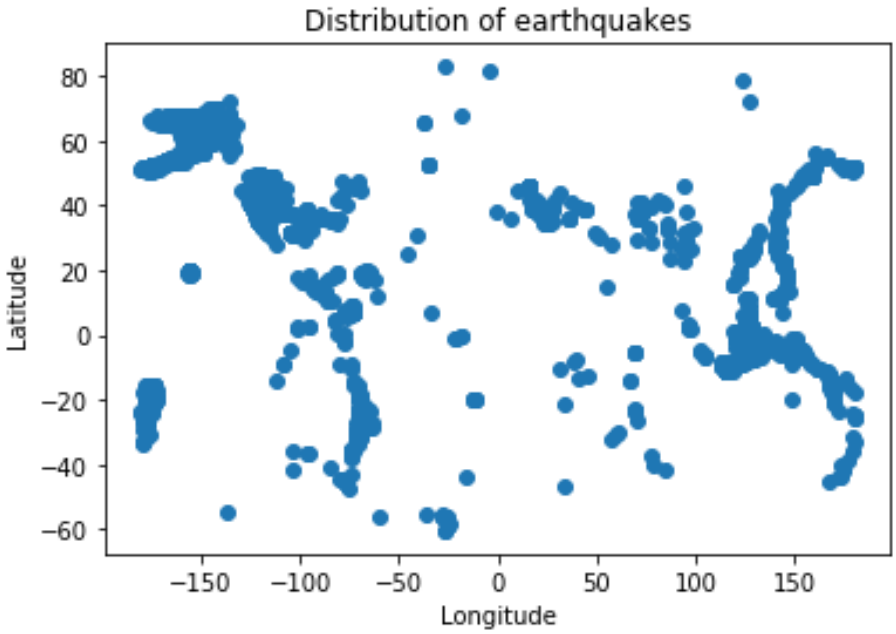

Figure 3 Scatter Plot of Latitude vs. Longitude for All Earthquakes.

From the plot it can be observed that the distribution of points follows the edges of the tectonic plates.

4. Normalized cumulative distribution plot of earthquake depths

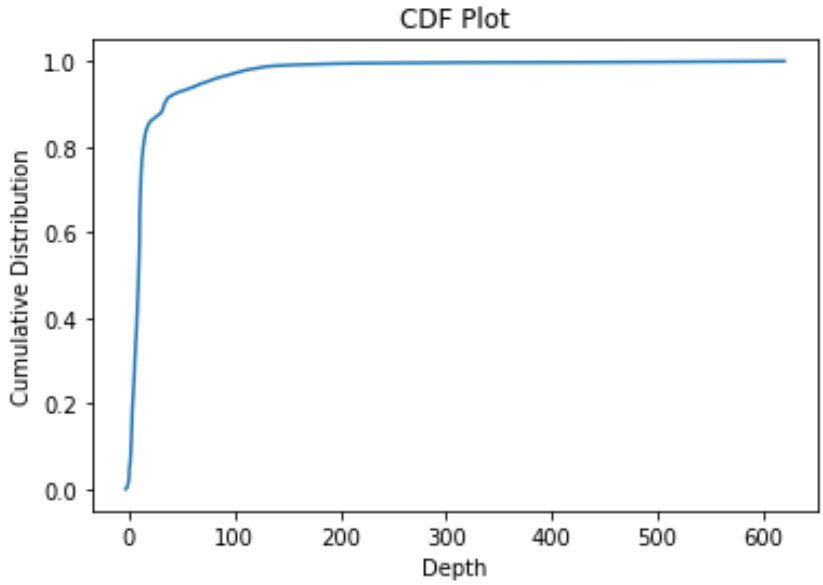

Figure 4 Normalized Cumulative Distribution Plot of Earthquake Depths

The cumulative distribution plot suggests that most of the earthquakes begin to rupture with a depth of less than 50-60 km. The slope before a depth of 100 km is very steep suggesting a dense distribution of earthquake occurrences in this region.

5. Scatter plot of earthquake magnitude vs. depth

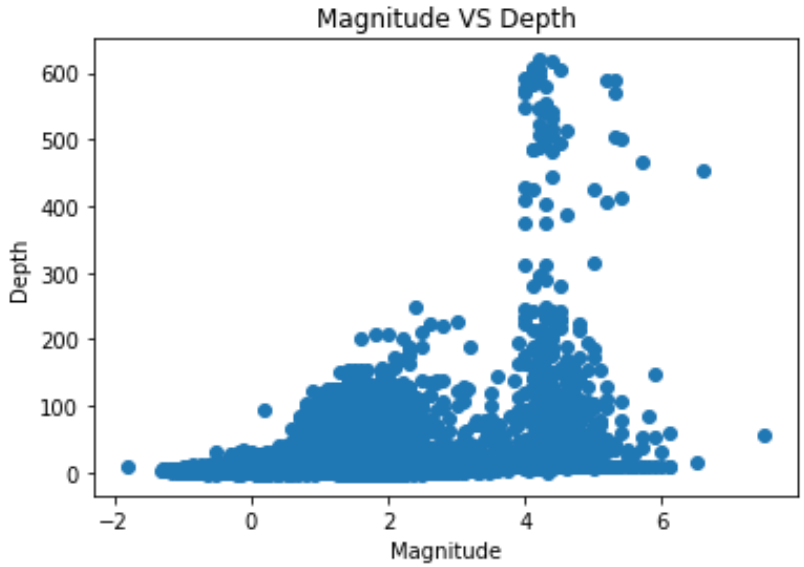

Figure 5 Scatter Plot of Earthquake Magnitude vs. Depth

The earthquakes with larger magnitudes seem to have bigger rupture depths which is also logical. The number of earthquakes having large rupture depths are very less than ones with smaller depths as the density in the plot suggests.

6. Quantile-Quantile plot of the earthquake magnitude

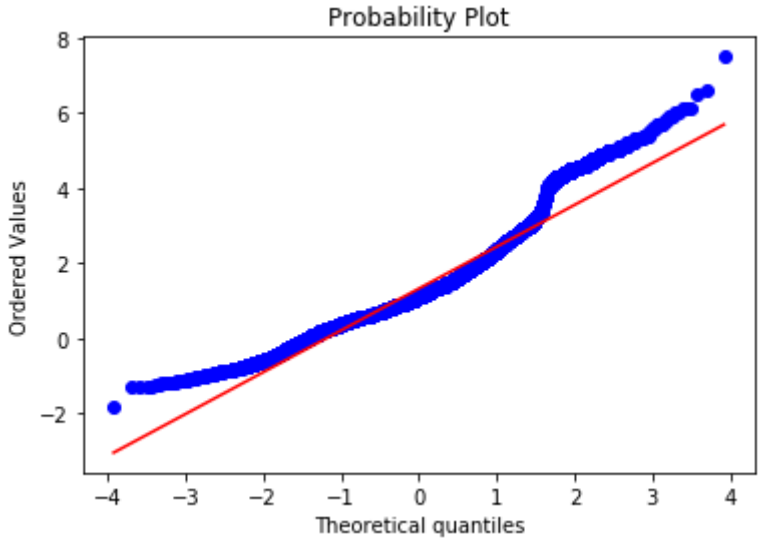

Figure 6 Normal Q-Q Plot of Earthquake Magnitude

The plot uses a normal distribution. The figure shows that there are some deviations at the extreme points of the data but most part of the data matches the theoretical normal distribution line. It can be inferred that the data might follow a normal distribution. Although the shape of the KDE suggests that it might also be gumbel distribution which can be confirmed by further statistical analysis.