

Metadata

Creator: Linji Wang (wang1278@purdue.edu)

Date of Creation: 03/06/20

File Name: program-07.py

File Type: Python 3 script

Purpose:

Graphical analysis of USGS “all earthquakes” data for the past “30 days”, including:

- 1) histogram of earthquake magnitude;
- 2) KDE plot of earthquake magnitude;
- 3) scatter plot of earthquake locations;
- 4) normalized CDF of earthquake depths;
- 5) scatter plot of earthquake magnitude with depth;
- 6) Q-Q plot of earthquake magnitudes.

Input Data:

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

Time of Download: 03/06/2020 1:04 PM EST

Format: .csv comma separated values ASCII text file

Descriptions:

This file contains earthquake related data including time, latitude, longitude, depth, magnitude, magnitude type, and other key information. Time period covered by this file is 30 days from 02/05/2020 to 03/06/2020. For more information about this data, please visit:

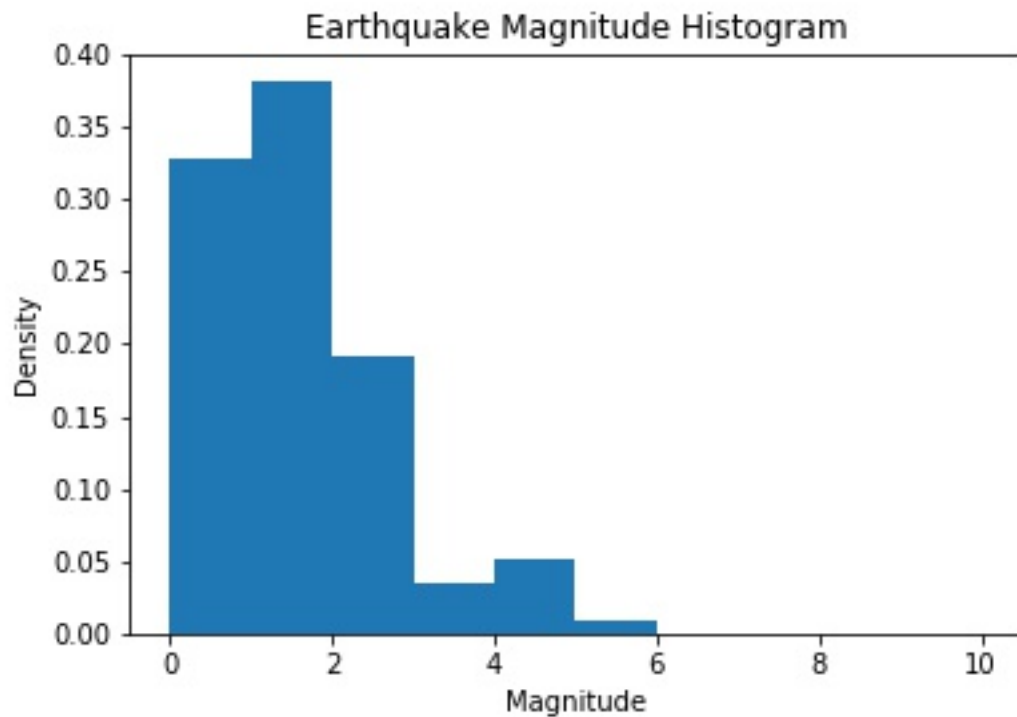
<https://earthquake.usgs.gov/earthquakes/feed/v1.0/csv.php>

Script Structure Note:

In order to read the earthquake data file, pandas read_table() is used instead of genfromtxt() since the data file contains multiple types of data (string, float, integer). Also, it is better to store the data as pandas dataframe rather than array.

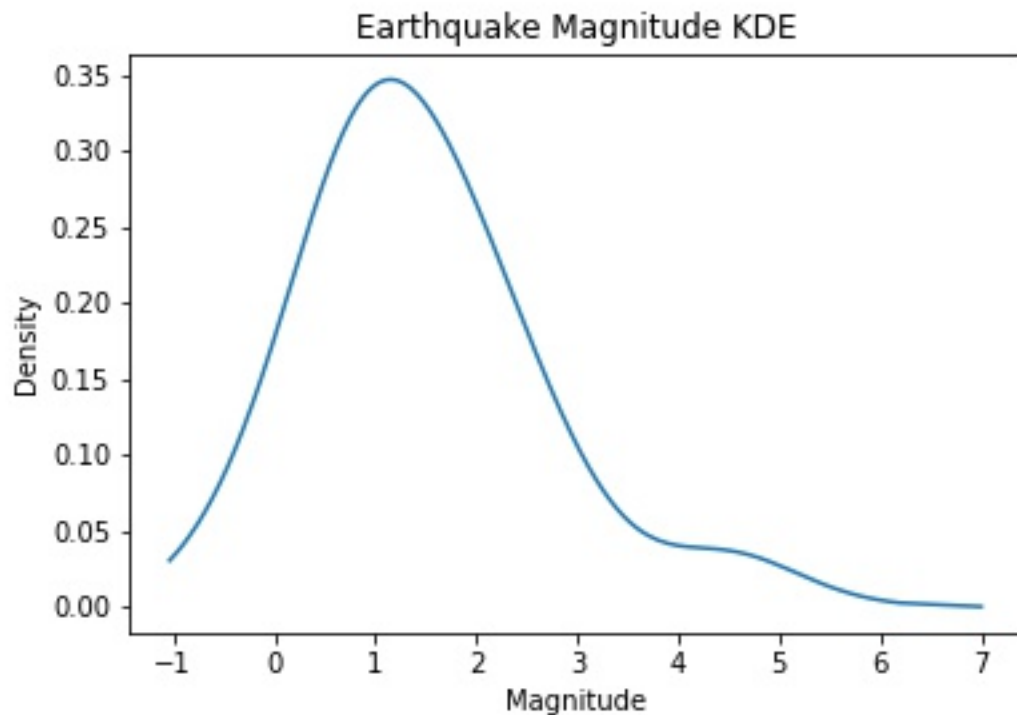
Output Graphs Analysis:

1) Histogram of earthquake magnitude



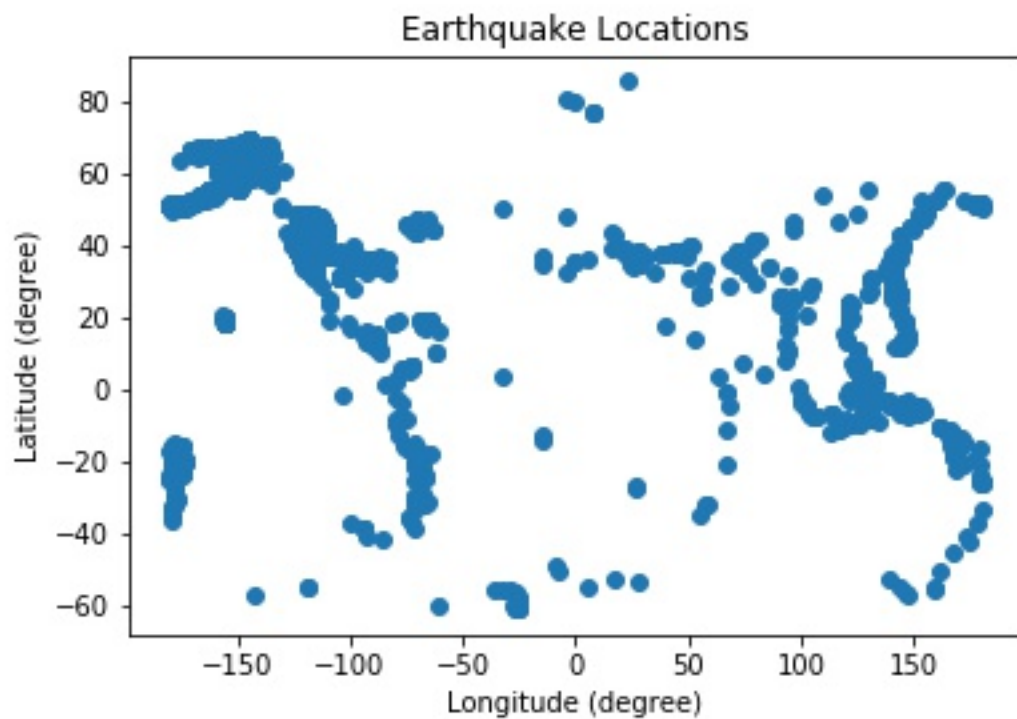
This histogram of earthquake magnitude, uses a bin width of 1 and a range of 0 to 10. The selection of bin size and range affect the density in each bin. For example, a bin width of 2 would result in a density of sum of the densities of magnitude 1 and 2. It can be selected according to the goal of analysis. This histogram shows the density distribution of different magnitude of earthquakes. Earthquakes with magnitude of 1-3 have higher probability of occurrence than others and magnitude-2 earthquake has the highest probability of occurrence.

2) KDE plot of earthquake magnitude



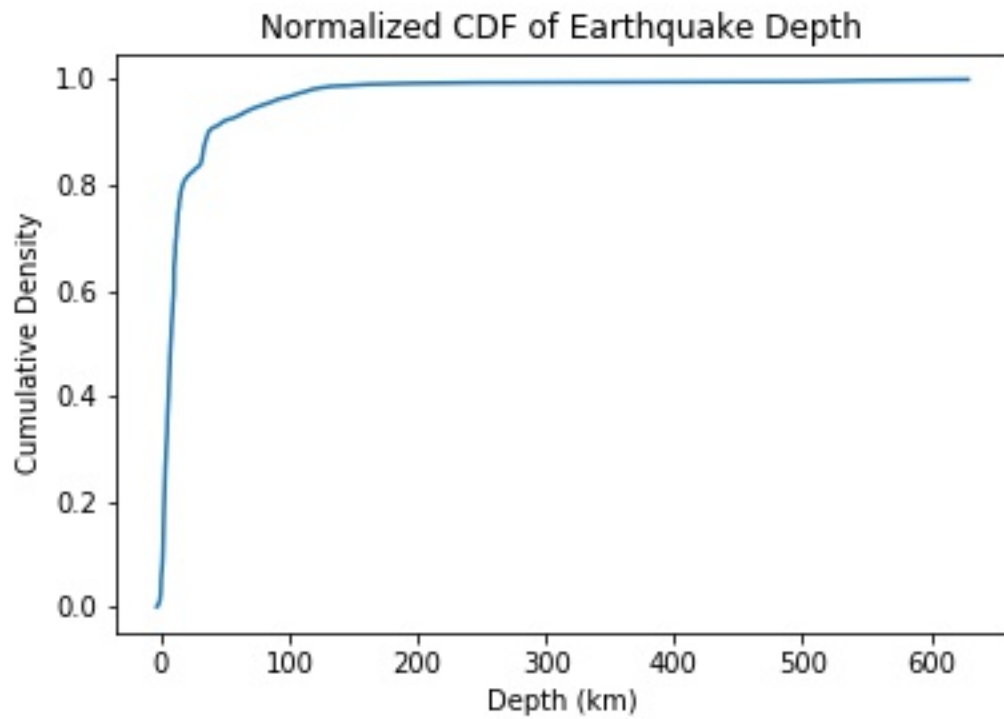
The KDE of earthquake magnitude uses a Gaussian kernel type with a bandwidth of 0.5. Comparing to the probability density histogram, it shows similar information (the probability density of each magnitude of earthquake). In comparison, the KDE uses line shape rather than bar shape to better represent the trend or movement of density with respect of magnitude. Additionally, the histogram accounts the sum probability density of each integer interval while the KDE represents the actual (or “calculated”) probability density of each data points.

3) scatter plot of earthquake locations



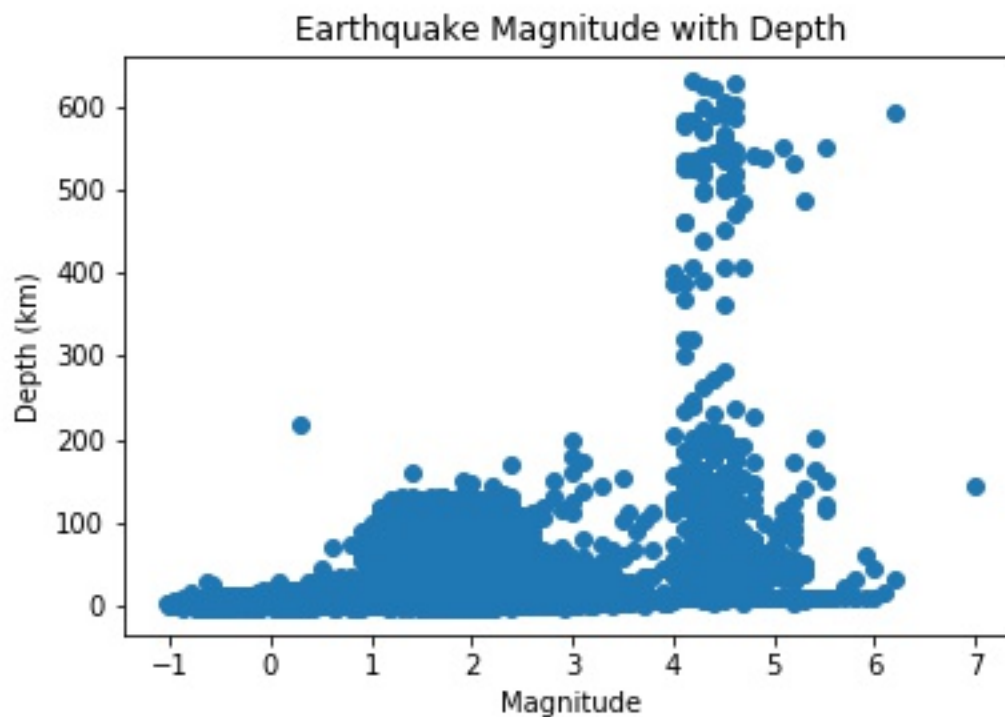
This earthquake location uses longitude as x and latitude as y because it is a common sense to refer North and South as “up” and “down”, West and East as “left” and “right”, thus longitude as x-axis and latitude as y-axis to make the plot “upright”. The plot shows that more earthquakes occurred at plate boundaries in the past 30 days.

4) Normalized CDF of earthquake depths



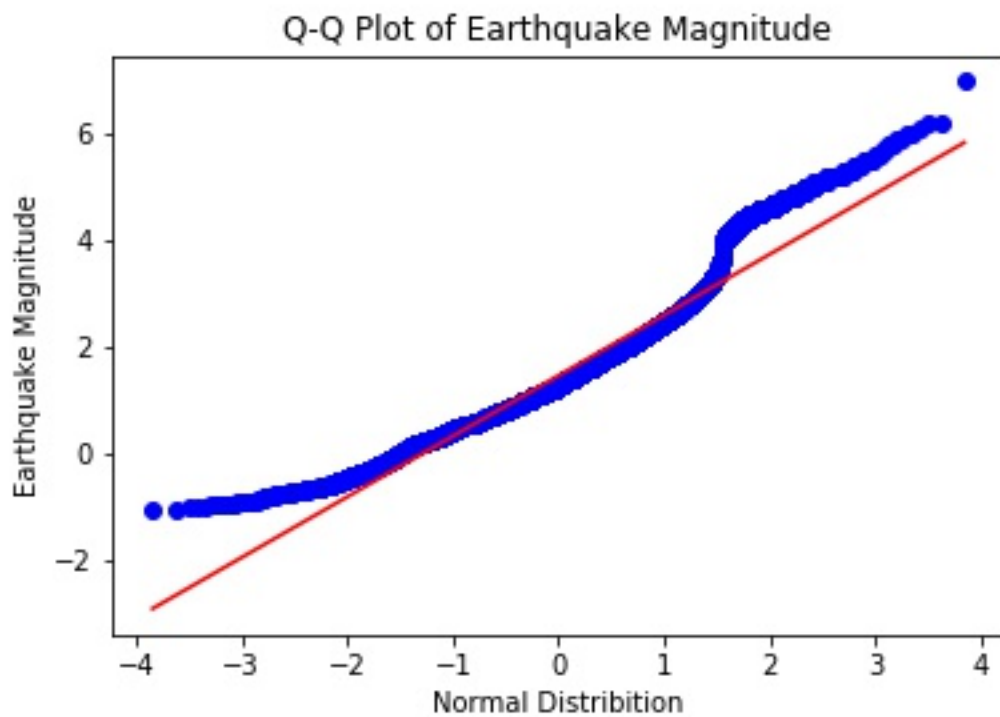
Based on the distribution trend in this normalized CDF plot, the very rapid slope change in the domain of 0-100 km depth shows that almost all earthquakes occurred at the depth of 0-100 km.

5) Scatter plot of earthquake magnitude with depth



This scatter plot shows a non-linear relationship between earthquake magnitude and earthquake depth. Earthquake magnitude increases as depth increases and the depth peaked over 600 km between magnitude of 4 and 5, and then mildly dropped to lower than 600 km after magnitude of 5.

6) Q-Q plot of earthquake magnitudes



The Q-Q plot assumes a normal distribution. The earthquake magnitude data comply with the distribution except for some deviations at the extreme value from -2 to 0.