

Metadata

The assignment 7 was carried out using most updated data for 30 days World Earthquake from USGS [1]. My previous work was based on the data downloaded in April 1st. However, due to the COVID-19 circumstance, I had to cease my work and redid again. Hence, the data used here is based on May 7th. The downloaded file was in the csv format and the view is as found in Figure 1. As seen in the figure, the file is using comma as the delimiter. In order to make sure that the data is correct, the printed head is shown in Figure 2.

```
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated,place
gNst,status,locationSource,magSource
2020-05-07T15:05:40.340Z,33.861,-116.802,20.24,1.08,ml,14,140,0.1603,0.35,ci,ci
of Cabazon, CA",earthquake,1.04,2.16,0.205,26,automatic,ci,ci
2020-05-07T15:04:44.886Z,61.419,-152.2194,0,1.6,ml,,,1,ak,ak0205vxqvji,2020-0
Alaska",earthquake,0.3,,,automatic,ak,ak
2020-05-07T14:45:02.290Z,32.8991667,-116.2121667,9.92,1.11,ml,22,85,0.09456,0.1
07T14:48:29.710Z,"27km NW of Ocotillo, CA",earthquake,0.33,0.84,0.154,14,automa
2020-05-07T14:43:18.520Z,33.5016667,-116.5076667,9.23,1.38,ml,52,57,0.05900,0.18
ESE of Anza, CA",earthquake,0.18,0.41,0.18,26,automatic,ci,ci
2020-05-07T14:17:14.183Z,30.0584,141.4828,54.34,4.6,mb,,125,3.015,1.13,us,us700
Japan region",earthquake,10.5,8.6,0.087,39,reviewed,us,us
2020-05-07T14:09:28.880Z,19.2224998,-155.3820038,31.17,1.9,md,46,154,0.02091,0.
07T14:12:40.700Z,"10km E of Pahala, Hawaii",earthquake,0.57,0.86,0.1,6,automati
2020-05-07T13:58:35.160Z,40.5108333,-112.165,-0.41,1.53,md,12,95,0.06867,0.2,uu
of Tooele, Utah",earthquake,0.53,1.42,0.412,8,reviewed,uu,uu
2020-05-07T13:54:00.200Z,33.342,-115.7998333,1.06,1.34,ml,8,181,0.2172,0.18,ci,
of Bombay Beach, CA",earthquake,1.76,31.61,0.216,30,automatic,ci,ci
2020-05-07T13:49:28.230Z,33.6363333,-116.7361667,15.5,0.53,ml,22,84,0.0485,0.16
07T13:53:03.086Z,"11km NNW of Anza, CA",earthquake,0.32,0.72,0.171,16,automatic
2020-05-07T13:34:20.400Z,19.4015007,-155.2663269,1.54,1.82,ml,27,33,0.003254,0.
07T13:40:58.380Z,"4km SW of Volcano, Hawaii",earthquake,0.19,0.14,0.27,14,auton
2020-05-07T13:17:49.620Z,42.0008333,-112.4563333,5.82,0.72,md,6,115,0.1535,0.12
07T14:08:16.770Z,"27km SW of Malad City, Idaho",earthquake,0.59,31.61,0.185,5,r
2020-05-07T13:11:57.240Z,38.8318329,-122.8184967,1.71,0.56,md,15,53,0.01069,0.0
07T13:45:04.148Z,"8km NW of The Geysers, CA",earthquake,0.26,0.63,,1,automatic,
2020-05-07T12:52:21.670Z,61.469,-150.0475,37.5,1.3,ml,,,0.72,ak,ak0205vwtbdj,2
Alaska",earthquake,,1.3,,,automatic,ak,ak
```

Figure 1 The 30 days earthquake csv file view

time	latitude	longitude	depth	mag	magType	nst	gap	dmin	rms
2020-05-07T14:45:02.290Z	32.899167	-116.212167	9.92	1.11	ml	22.0	85.0	0.09456	0.18
2020-05-07T14:43:18.520Z	33.501667	-116.507667	9.23	1.38	ml	52.0	57.0	0.05900	0.18
2020-05-07T14:09:28.880Z	19.222500	-155.382004	31.17	1.90	md	46.0	154.0	0.02091	0.16
2020-05-07T13:58:35.160Z	40.510833	-112.165000	-0.41	1.53	md	12.0	95.0	0.06867	0.20
2020-05-07T13:54:00.200Z	33.342000	-115.799833	1.06	1.34	ml	8.0	181.0	0.21720	0.18

Figure 2 The head view of 30 days Earthquake data in Jupyter Notebook

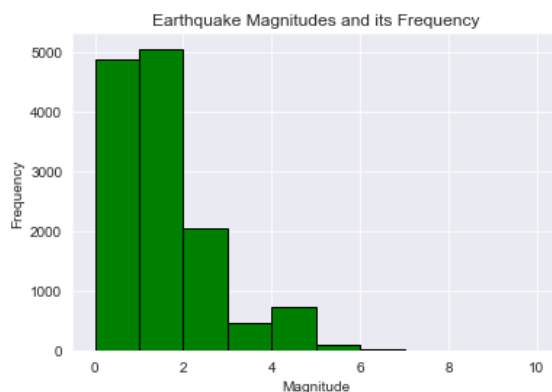


Figure 3 Histogram of Earthquake Magnitude (2.2)

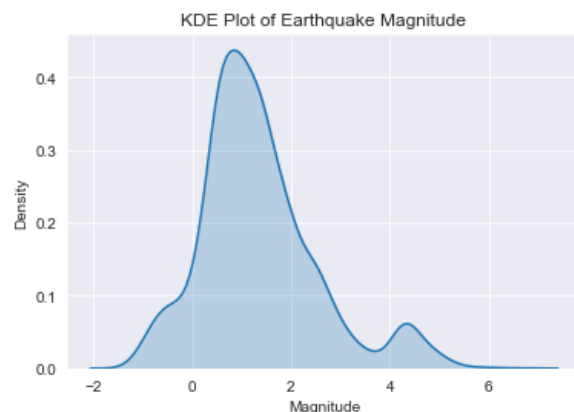


Figure 4 KDE plot of earthquake magnitude (2.3)

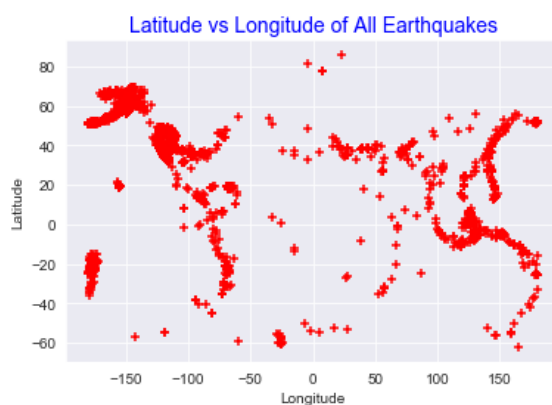


Figure 5 Latitude vs longitude for all earthquakes (2.4)

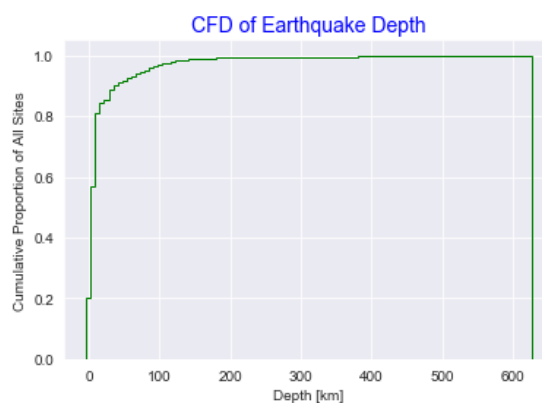


Figure 6 Normalized CDF of earthquake depths (2.5)

- 2.1 The numpy function `genfromtxt()` does not work because the source file contains values in many different types. The `genfromtxt()` simply cannot handle, hence needs the `pd.read_csv()`.
- 2.2 The histogram of magnitude as a function of frequency is depicted in Figure 3. It shows skewed-right, 7 magnitude and 1.5-2 mag as the highest and most frequencies occurred, respectively. The range is for the range of data. For instance, in this particular case, the magnitude is from 0 to mostly 9 for the earthquake. Hence, showing 0-10 is logical.
- 2.3 KDE plot and its relation to the histogram are shown in Figure 4 and Figure 9, respectively. The difference between the KDE and the histogram is the KDE starts at around -2, where the histogram is at 0. The highest peak and the biggest magnitude are 4.5 density and around 7 magnitude, respectively.
- 2.4 The latitude versus longitude for all earthquakes in the form of scattered plot is shown in Figure 5. As shown in the figure, it resembles the form of the ring of fire around the globe. Setting the longitude and latitude as x and y-axis, respectively, for the best fit in plotting to be confirmed with global map based on the global consensus, i.e. north is at the upper side.

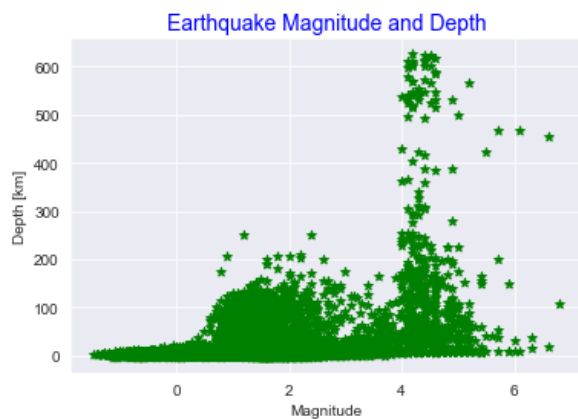


Figure 7 Scatterplot of earthquake (2.6)

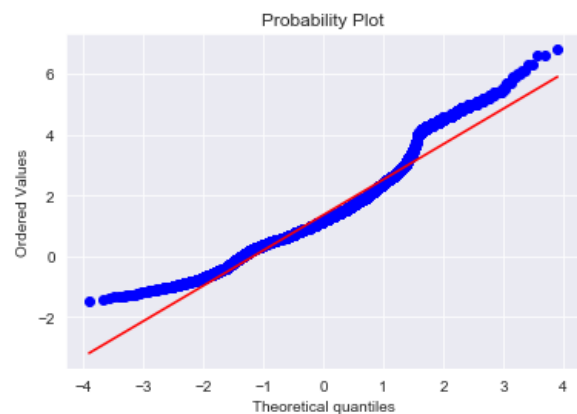


Figure 8 Q-Q plot of earthquake magnitudes

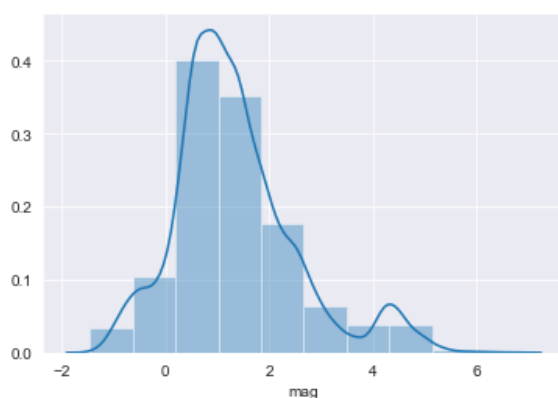


Figure 9 relation of KDE and histogram plots

- 2.5 The normalized cumulative distribution plot of earthquake depths is depicted in Figure 6. To my understanding, this graph shows that most of the earthquakes occurred in the depth range of 0-600 km. From those range, the biggest cumulative proportion of all sites are found in around 50-100 km.

- 2.6 The scatter plot of magnitude as a function of the depth is shown in Figure 7. The lowest depth is found in near 0, whereas the highest depth reach around 630 km from the ground. The magnitude of those earthquakes in the deepest ground are within the range of 4-4.5 magnitude.
- 2.7 The Q-Q plot was made based on the stackoverflow guidance [2] and the result is shown in Figure 8. The figure shows that it has normal distribution since the dots are close to the red diagonal line. The normal distributions here is the comparison between the theoretical and the ordered value.

Reference

- [1] "Earthquake_Data." <https://earthquake.usgs.gov/earthquakes/feed/v1.0/csv.php> (accessed May 07, 2020).
- [2] "python - Quantile-Quantile Plot using SciPy," *Stack Overflow*. <https://stackoverflow.com/questions/13865596/quantile-quantile-plot-using-scipy> (accessed May 07, 2020).