

# EARTHQUAKE PREDICTION MODEL USING PYTHON

## PHASE 3

**NAME:** KINGSTON S

**REGISTER NUMBER:** 61772131021

### **DATA VISUALIZATION:**

Data visualization using machine learning in an earthquake prediction model refers to the integration of machine learning techniques with visualization tools to analyze and represent diverse data types, including seismic data, geospatial information, and geological characteristics. The outcomes of these analyses are then visually presented through graphs, maps, interactive dashboards, or other visual aids, enabling researchers and stakeholders to interpret and comprehend the data effectively. This fusion of machine learning and visualization enhances the understanding of earthquake-related data, aiding in informed decision-making and proactive earthquake prediction and mitigation efforts.

There are several common techniques used for data visualization: charts (bar, line, pie), plots (scatter, bubble, box), maps (heatmaps, dot distribution maps, cartograms), diagram and matrices.

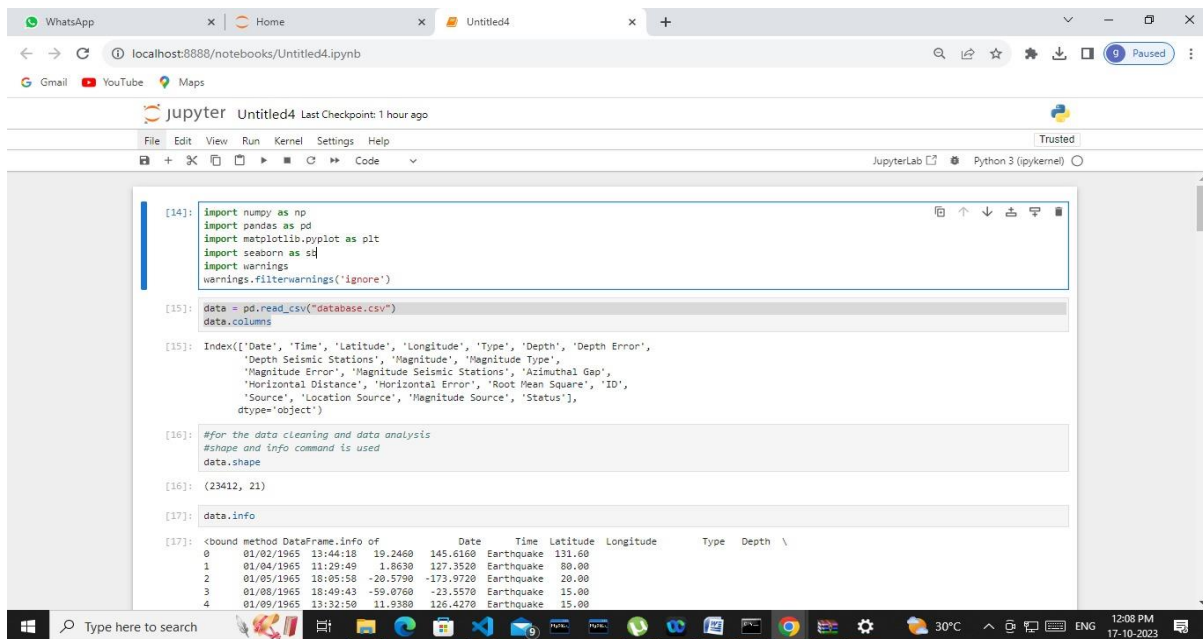
### **DATA VISUALIZATION IN EARTHQUAKE PREDICTION:**

The necessary libraries required for data analysis to data visualization of the earthquake.

#### **STEP 1:**

Import the pandas, numpy, matplotlib, seaborn and warnings. Unwanted warnings can be ignored and to display the data to index.

Read the data from csv and also columns which necessary for the model and the column which needs to be predicted. For the data can be cleaning and analysis the data.



```
[14]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sd
import warnings
warnings.filterwarnings('ignore')

[15]: data = pd.read_csv("database.csv")
data.columns

[15]: Index(['Date', 'Time', 'Latitude', 'Longitude', 'Type', 'Depth', 'Depth Error',
'Death Seismic Stations', 'Magnitude', 'Magnitude Type',
'Magnitude Error', 'Magnitude Seismic Stations', 'Azimuthal Gap',
'Horizontal Distance', 'Horizontal Error', 'Root Mean Square', 'ID',
'Source', 'Location Source', 'Magnitude Source', 'Status'],
dtype='object')

[16]: #For the data cleaning and data analysis
#shape and info command is used
data.shape

[16]: (23412, 21)

[17]: data.info

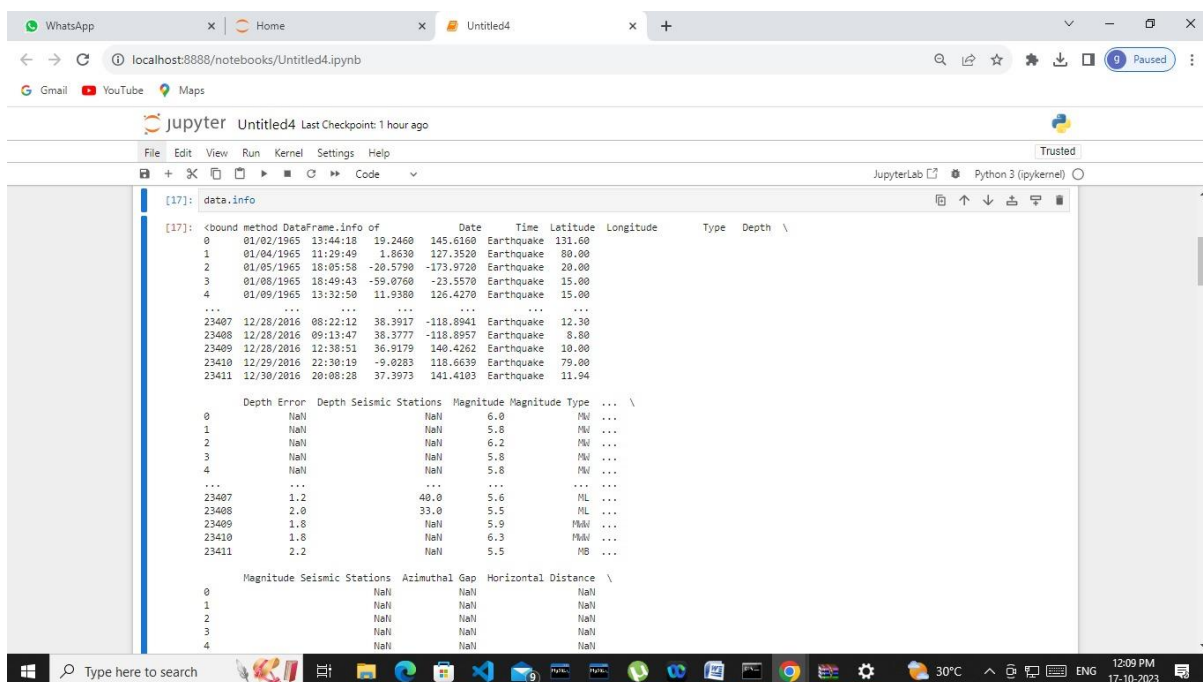
[17]: <bound method DataFrame.info of
0 01/02/1965 13:44:18 19.2460 145.6160 Earthquake 131.60
1 01/04/1965 11:29:49 1.8630 127.3520 Earthquake 80.00
2 01/05/1965 18:05:58 -20.5790 -173.9720 Earthquake 20.00
3 01/08/1965 18:49:43 -59.0760 -23.5570 Earthquake 15.00
4 01/09/1965 13:32:50 11.9380 126.4270 Earthquake 15.00
```

## STEP 2:

Shape and info command can be used in the given program.

Shape- shape can be represent qualitative data. Each category is represented by a specific shape.

Info- To explore relationships between data, to confirm ideas we hold about data or to explain data in easy to digest manner. Display the information of earthquake prediction.



```
[17]: data.info

[17]: <bound method DataFrame.info of
0 01/02/1965 13:44:18 19.2460 145.6160 Earthquake 131.60
1 01/04/1965 11:29:49 1.8630 127.3520 Earthquake 80.00
2 01/05/1965 18:05:58 -20.5790 -173.9720 Earthquake 20.00
3 01/08/1965 18:49:43 -59.0760 -23.5570 Earthquake 15.00
4 01/09/1965 13:32:50 11.9380 126.4270 Earthquake 15.00
...
23407 12/28/2016 08:22:12 38.3917 -118.8941 Earthquake 12.30
23408 12/28/2016 09:13:47 38.3777 -118.8957 Earthquake 8.80
23409 12/28/2016 12:38:51 36.9179 140.4262 Earthquake 10.00
23410 12/29/2016 22:30:19 -9.0283 118.6639 Earthquake 79.00
23411 12/30/2016 20:08:28 37.3973 141.4103 Earthquake 11.94

Depth Error Depth Seismic Stations Magnitude Magnitude Type ... \
0 NaN NaN NaN 6.0 Mw ...
1 NaN NaN NaN 5.8 Mw ...
2 NaN NaN NaN 6.2 Mw ...
3 NaN NaN NaN 5.8 Mw ...
4 NaN NaN NaN 5.8 Mw ...
...
23407 1.2 40.0 5.6 ML ...
23408 2.0 33.0 5.5 ML ...
23409 1.8 NaN 5.9 Mw ...
23410 1.8 NaN 6.3 Mw ...
23411 2.2 NaN 5.5 MB ...

Magnitude Seismic Stations Azimuthal Gap Horizontal Distance \
0 NaN NaN NaN NaN
1 NaN NaN NaN NaN
2 NaN NaN NaN NaN
3 NaN NaN NaN NaN
4 NaN NaN NaN NaN
```

```

23407      18.0      42.47      0.120
23408      18.0      48.58      0.129
23409      NaN      91.00      0.992
23410      NaN      26.00      3.553
23411      428.0      97.00      0.681

Horizontal Error Root Mean Square ID Source \
0      NaN      NaN      ISCGEN860706 ISCGEN
1      NaN      NaN      ISCGEN860737 ISCGEN
2      NaN      NaN      ISCGEN860762 ISCGEN
3      NaN      NaN      ISCGEN860856 ISCGEN
4      NaN      NaN      ISCGEN860890 ISCGEN
...      ...      ...      ...
23407      NaN      0.1898      MN00570710 NN
23408      NaN      0.2187      MN00570744 NN
23409      4.8      1.5200      US1000704F US
23410      6.0      1.4300      US1000704F US
23411      4.5      0.9100      US1000704F US

Location Source Magnitude Source Status
0      ISCGEN      ISCGEN      Automatic
1      ISCGEN      ISCGEN      Automatic
2      ISCGEN      ISCGEN      Automatic
3      ISCGEN      ISCGEN      Automatic
4      ISCGEN      ISCGEN      Automatic
...      ...      ...      ...
23407      NN      NN      Reviewed
23408      NN      NN      Reviewed
23409      US      US      Reviewed
23410      US      US      Reviewed
23411      US      US      Reviewed

[23412 rows x 21 columns]>

```

### STEP 3:

Read the describe data. The Descriptive statistical measures also gives us some idea regarding the distribution of the data.

The data are Longitude, Latitude, Depth, Depth Error, Depth Seismic Stations, Magnitude, Magnitude Seismic Stations, Azimuthal Gap, Horizontal Distance, Horizontal Error, Root Mean Square are describe the data can be read. The data can be described by the data visualization.

```

[18]: #descriptive statistical measures also gives us some idea regarding the distribution of the data.
data.describe()

[18]:

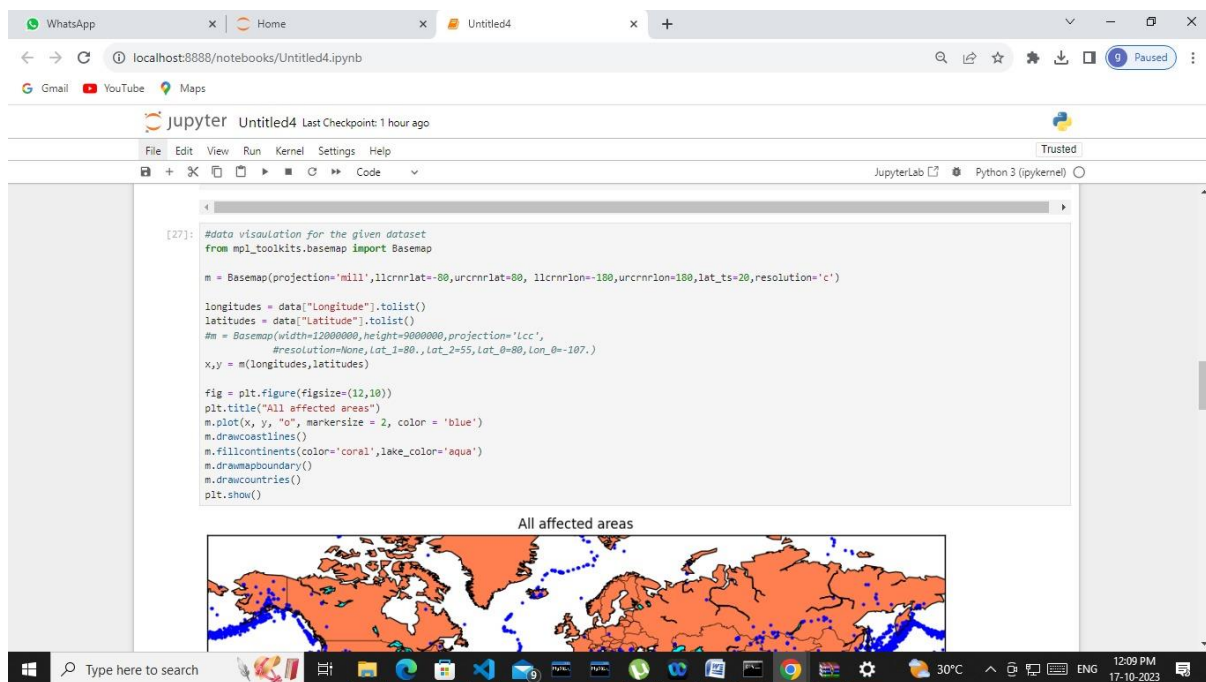
```

	Latitude	Longitude	Depth	Depth Error	Depth Seismic Stations	Magnitude	Magnitude Error	Magnitude Seismic Stations	Azimuthal Gap	Horizontal Distance	Horizontal Error	Root Mean Square
count	23412.000000	23412.000000	23412.000000	4461.000000	7097.000000	23412.000000	327.000000	2564.000000	7299.000000	1604.000000	1156.000000	17352.000000
mean	1.679033	39.639961	70.767911	4.993115	275.364098	5.882531	0.071820	48.944618	44.163532	3.992660	7.662759	1.022784
std	30.113183	125.511959	122.651898	4.875184	162.141631	0.423066	0.051466	62.943106	32.141486	5.377262	10.430396	0.188545
min	-77.080000	-179.997000	-1.100000	0.000000	0.000000	5.500000	0.000000	0.000000	0.004505	0.085000	0.085000	0.000000
25%	-18.653000	-76.349750	14.532500	1.800000	146.000000	5.600000	0.046000	10.000000	24.100000	0.968750	5.300000	0.900000
50%	-3.568500	103.982000	33.000000	3.500000	255.000000	5.700000	0.059000	28.000000	36.000000	2.319500	6.700000	1.000000
75%	26.190750	145.026250	54.000000	6.300000	384.000000	6.000000	0.075500	66.000000	54.000000	4.724500	8.100000	1.130000
max	86.005000	179.998000	700.000000	91.295000	934.000000	9.100000	0.410000	821.000000	360.000000	37.874000	99.000000	3.440000

## STEP 4:

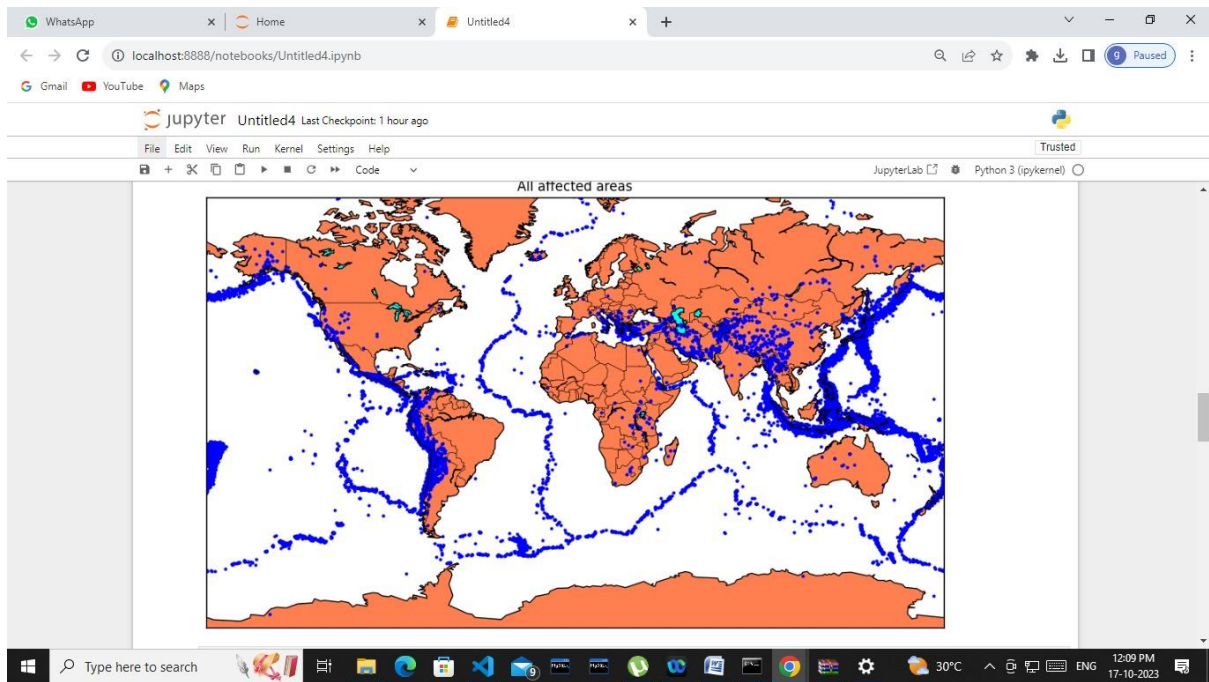
The data visualization for the given dataset. Import the Base map and declare the Longitude, Latitude.

Declare the figure size to be plotted. That can be plot the affected area in the given world ma and to draw coastlines, map boundary, countries. To fill the continents. After the plotted graph can be show in the given output. A visualization that shows data in map form using different shapes and colors to show the relationship between pieces of data and specific locations. A combination of visuals and words that represent data.



All affected area is shown in the output of the given world map.

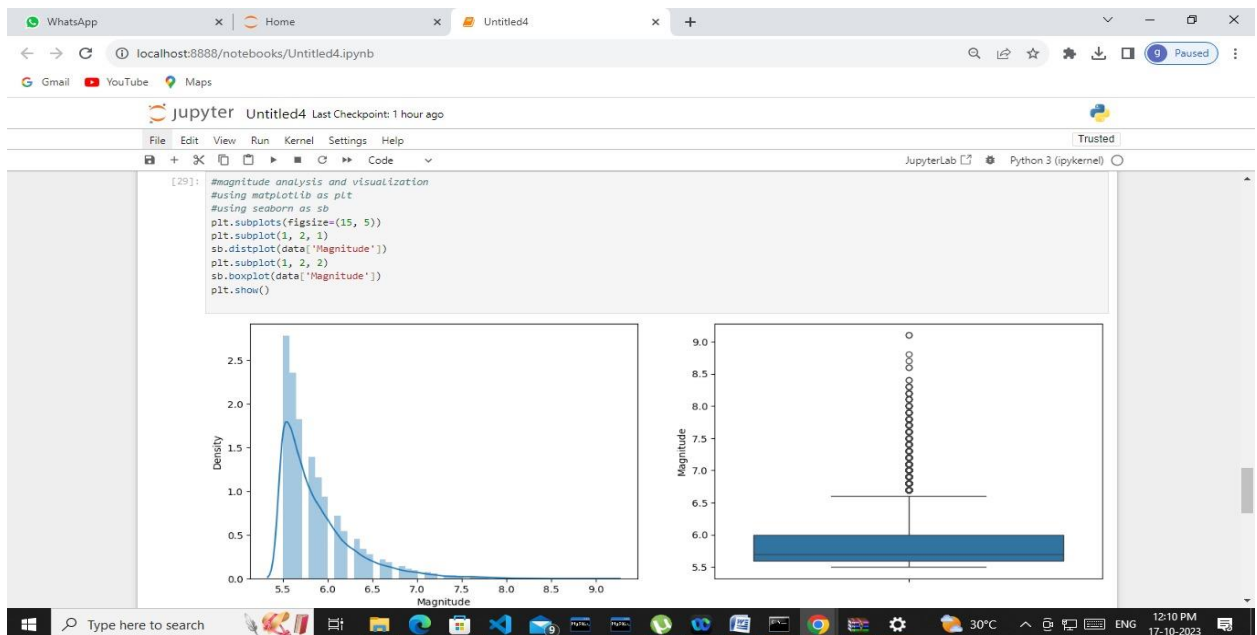
Map visualization also known as cartographic visualization. The geospatial data are usually visualized in the form of static or interactive maps.



**STEP 5:** Use the matplotlib, seaborn library in the magnitude analysis and visualization of the data.

The length of a bar in a bar chart can represent the magnitude of a specific data point.

The plots are subplots, displots, boxplots are shown in the output of given barchat. The bar chat are plot between density and magnitude and another chart are plot to the magnitude of the points to the data visualization of the dataset.



Thus, Data Visualization is completed for Earthquake failure detection by using Python.