earthquakes_visualization_(1) (1)

December 3, 2023

1 Earthquakes visualization

2 Project Overview:

The goal of this project is to develop a visualization tool that effectively represents and communicates earthquake data from the National Earthquake Information Center (NEIC). The tool should enable users to explore and understand earthquake patterns, trends, and potential impacts.

3 A bout the dataset

The National Earthquake Information Center (NEIC) determines the location and size of all significant earthquakes that occur worldwide and disseminates this information immediately to national and international agencies, scientists, critical facilities, and the general public. The NEIC compiles and provides to scientists and to the public an extensive seismic database that serves as a foundation for scientific research through the operation of modern digital national and global seismograph networks and cooperative international agreements. The NEIC is the national data center and archive for earthquake information.

4 Part 1 : Data Importation

1. Import pandas and the different packages of Plotly

[3]:
```
import pandas as pd
import plotly.express as px
import plotly.graph_objects as go
import matplotlib.pyplot as plt
from plotly.subplots import make_subplots
import plotly.io as pio

pd.options.display.max_columns = None


dataset = pd.read_csv("earthquake_1995-2023.csv")
dataset.head()
```

[3]:
```
                                    title  magnitude         date_time  cdi  \
0          M 6.5 - 42 km W of Sola, Vanuatu        6.5  16-08-2023 12:47    7
```

```
1  M 6.5 - 43 km S of Intipucá, El Salvador        6.5  19-07-2023 00:22    8
2  M 6.6 - 25 km ESE of Loncopué, Argentina        6.6  17-07-2023 03:05    7
3     M 7.2 - 98 km S of Sand Point, Alaska        7.2  16-07-2023 06:48    6
4                  M 7.3 - Alaska Peninsula        7.3  16-07-2023 06:48    0

   mmi   alert  tsunami  sig net  nst      dmin    gap magType    depth  \
0    4   green        0  657  us  114  7.177000   25.0     mww  192.955
1    6  yellow        0  775  us   92  0.679000   40.0     mww   69.727
2    5   green        0  899  us   70  1.634000   28.0     mww  171.371
3    6   green        1  860  us  173  0.907000   36.0     mww   32.571
4    5     NaN        1  820  at   79  0.879451  172.8      Mi   21.000

   latitude  longitude              location      continent    country
0  -13.8814   167.1580          Sola, Vanuatu            NaN    Vanuatu
1   12.8140   -88.1265  Intipucá, El Salvador            NaN        NaN
2  -38.1911   -70.3731    Loncopué, Argentina  South America  Argentina
3   54.3844  -160.6990     Sand Point, Alaska            NaN        NaN
4   54.4900  -160.7960       Alaska Peninsula            NaN        NaN
```

2. Show the first rows as well as basic statistics about the dataset

[6]:
```
dataset.describe(include="all")
```

[6]:
```
                                        title    magnitude         date_time  \
count                                    1000  1000.000000              1000
unique                                    984          NaN               990
top     M 6.5 - Kermadec Islands, New Zealand          NaN  11-01-2022 12:39
freq                                        3          NaN                 3
mean                                      NaN     6.940150               NaN
std                                       NaN     0.438148               NaN
min                                       NaN     6.500000               NaN
25%                                       NaN     6.600000               NaN
50%                                       NaN     6.800000               NaN
75%                                       NaN     7.100000               NaN
max                                       NaN     9.100000               NaN

                cdi         mmi  alert      tsunami          sig   net  \
count   1000.000000  1000.00000    449  1000.000000  1000.000000  1000
unique          NaN         NaN      4          NaN          NaN    11
top             NaN         NaN  green          NaN          NaN    us
freq            NaN         NaN    353          NaN          NaN   960
mean       3.605000     6.02700    NaN     0.325000   847.915000   NaN
std        3.328972     1.43399    NaN     0.468609   301.802632   NaN
min        0.000000     1.00000    NaN     0.000000   650.000000   NaN
25%        0.000000     5.00000    NaN     0.000000   691.000000   NaN
50%        4.000000     6.00000    NaN     0.000000   744.000000   NaN
75%        7.000000     7.00000    NaN     1.000000   874.250000   NaN
```

```
max        9.000000    10.00000    NaN     1.000000  2910.000000   NaN

                nst         dmin          gap magType        depth  \
count   1000.000000  1000.000000  1000.000000    1000  1000.000000
unique          NaN          NaN          NaN       9          NaN
top             NaN          NaN          NaN     mww          NaN
freq            NaN          NaN          NaN     502          NaN
mean     193.918000     1.125174    20.926290     NaN    74.612541
std      239.045858     2.073164    24.415895     NaN   130.812590
min        0.000000     0.000000     0.000000     NaN     2.700000
25%        0.000000     0.000000     0.000000     NaN    16.000000
50%        0.000000     0.000000    18.000000     NaN    29.000000
75%      403.000000     1.549250    27.000000     NaN    55.000000
max      934.000000    17.654000   239.000000     NaN   670.810000

           latitude    longitude                 location continent  \
count   1000.000000  1000.000000                      994       284
unique          NaN          NaN                      502         6
top             NaN          NaN  Kokopo, Papua New Guinea      Asia
freq            NaN          NaN                       29       137
mean       4.315554    51.486576                      NaN       NaN
std       26.633320   117.478302                      NaN       NaN
min      -61.848400  -179.968000                      NaN       NaN
25%      -13.518500   -71.694450                      NaN       NaN
50%       -2.443500   107.791000                      NaN       NaN
75%       25.167250   148.364750                      NaN       NaN
max       71.631200   179.662000                      NaN       NaN

          country
count         651
unique         56
top     Indonesia
freq          140
mean          NaN
std           NaN
min           NaN
25%           NaN
50%           NaN
75%           NaN
max           NaN
```

3. To avoid problems with your visualizations, use pandas to convert this column into `Datetime` type

[9]:
```
# convert all observations of Date to datetime
dataset.loc[:,'date_time'] = pd.to_datetime(dataset['date_time'], utc=True)
```

```
# sort values by date
dataset = dataset.sort_values(by = 'date_time')
dataset.head()
```

[9]:

```
                                      title  magnitude  \
975       M 6.7 - 73 km WSW of Vallenar, Chile        6.7
982         M 7.0 - 45 km SE of Sucúa, Ecuador        7.0
981         M 6.5 - 50 km SE of Sucúa, Ecuador        6.5
973     M 7.9 - 128 km ESE of Kuril'sk, Russia        7.9
999  M 7.1 - 14 km NE of Cabatuan, Philippines        7.1

                    date_time  cdi  mmi alert  tsunami  sig net  nst  dmin  \
975 1995-01-11 00:35:00+00:00    0    7   NaN        0  691  us    0   0.0
982 1995-03-10 01:51:00+00:00    0    7   NaN        0  754  us    0   0.0
981 1995-03-10 12:44:00+00:00    0    7   NaN        0  650  us    0   0.0
973 1995-03-12 18:01:00+00:00    0    6   NaN        0  960  us    0   0.0
999 1995-05-05 03:53:00+00:00    0    7   NaN        0  776  us    0   0.0

     gap magType  depth  latitude  longitude              location  \
975  0.0      mw   19.9   -28.906    -71.417       Vallenar, Chile
982  0.0      mw   24.4    -2.750    -77.881         Sucúa, Ecuador
981  0.0      mw   16.7    -2.778    -77.851         Sucúa, Ecuador
973  0.0      mw   33.0    44.663    149.300      Kuril'sk, Russia
999  0.0      mw   16.0    12.626    125.297  Cabatuan, Philippines

         continent      country
975            NaN        Chile
982  South America      Ecuador
981  South America      Ecuador
973            NaN          NaN
999            NaN  Philippines
```

[11]:

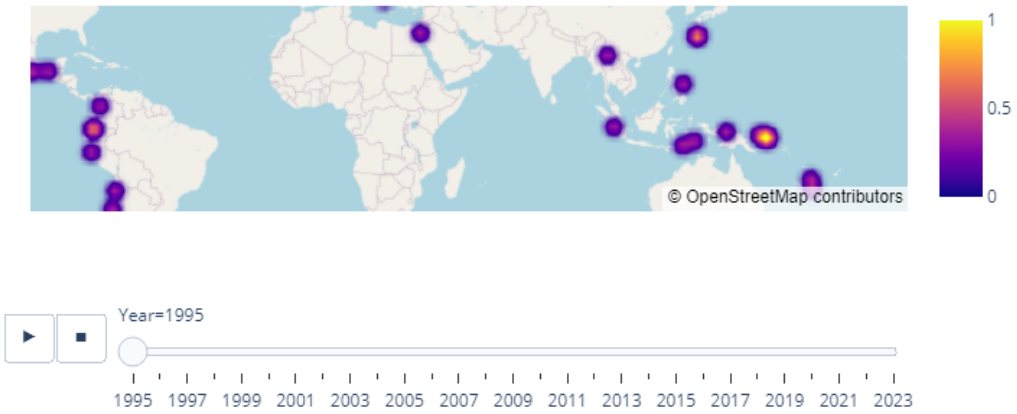
```
# Convert 'Date_Time' to datetime format
dataset['date_time'] = pd.to_datetime(dataset['date_time'])

# Extract the year and create a new column 'Year'
dataset['Year'] = dataset['date_time'].dt.year
```

5 the distribution of earthquake locations

[14]:

```
fig = px.density_mapbox(dataset , lat="latitude", lon="longitude",␣
 ↪mapbox_style="open-street-map",
                        animation_frame = 'Year', zoom = 0.5, radius = 10)
fig.show()
```

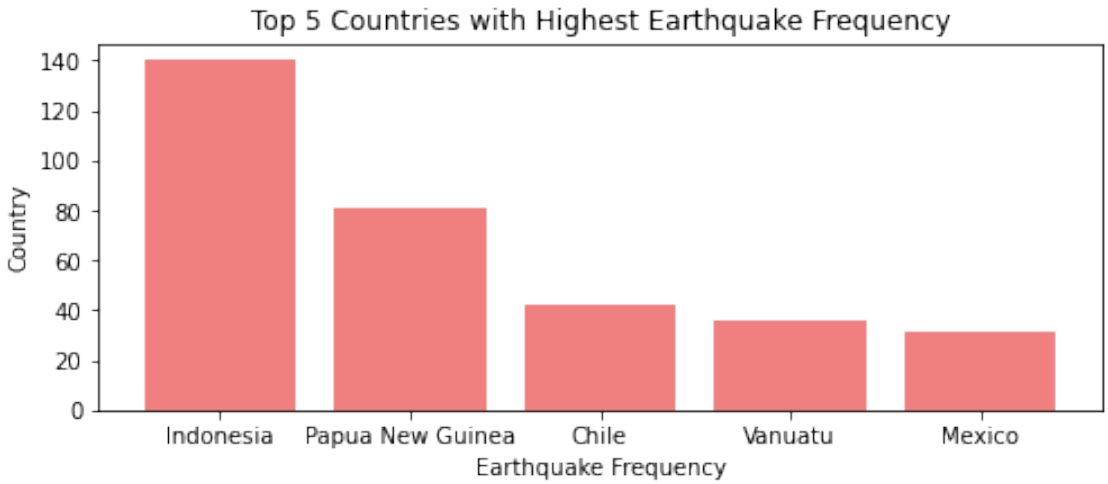


[15]:
```
top_countries = dataset.groupby('country').size().nlargest(5).index
filtered_top_countries = dataset[dataset['country'].isin(top_countries)]

# Create a table showing earthquake frequency for the top countries
table_data = filtered_top_countries['country'].value_counts().reset_index()
table_data.columns = ['Country', 'Earthquake Frequency']

# Plot a horizontal bar chart
plt.subplots(figsize=(8, 3))
plt.bar(table_data['Country'], table_data['Earthquake Frequency'],␣
 ↪color='lightcoral')
plt.title('Top 5 Countries with Highest Earthquake Frequency')
plt.xlabel('Earthquake Frequency')
plt.ylabel('Country')

plt.show()
```



[16]:
```
table_data
```

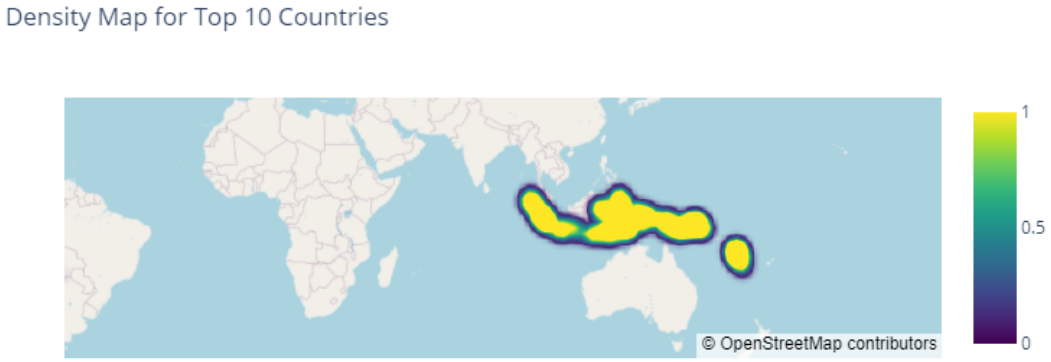
[16]:

	Country	Earthquake Frequency
0	Indonesia	140
1	Papua New Guinea	81
2	Chile	42
3	Vanuatu	36
4	Mexico	31

6 The distribution of earthquake ten top country

[18]:
```
# Create density maps for both top and least countries
fig_top_countries = px.density_mapbox(filtered_top_countries, lat="latitude",␣
 ↪lon="longitude", color_continuous_scale="Viridis",
                                      mapbox_style="open-street-map", zoom=0.5,␣
 ↪radius=10,
                                      title='Density Map for Top 10 Countries')


# Display the plots
fig_top_countries.show()
```



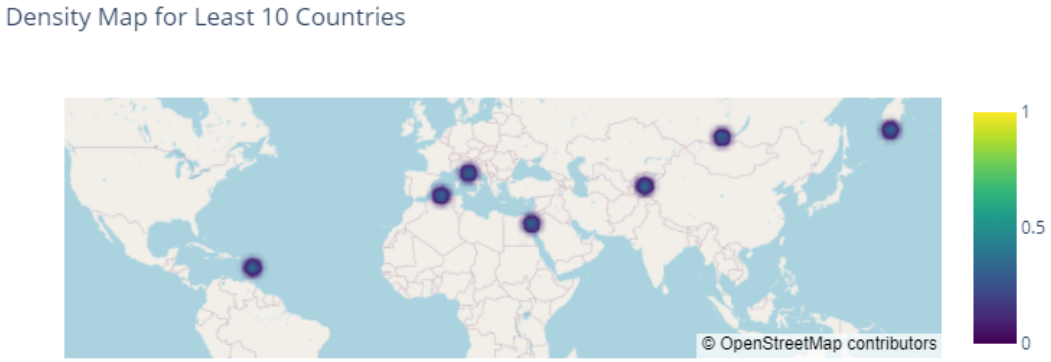
[41]:
```
least_countries = dataset.groupby('country').size().nsmallest(10).index
filtered_least_countries = dataset[dataset['country'].isin(least_countries)]

# Create a table showing earthquake frequency for the top countries
table_data = filtered_top_countries['country'].value_counts().reset_index()
table_data.columns = ['Country', 'Earthquake Frequency']
```

Plot the country have leaset

[44]:
```
fig_least_countries = px.density_mapbox(filtered_least_countries,␣
 ↪lat="latitude", lon="longitude", color_continuous_scale="Viridis",
                                        mapbox_style="open-street-map", zoom=0.
 ↪5, radius=10,
                                        title='Density Map for Least 10␣
 ↪Countries')


fig_least_countries.show()
print(least_countries)
```

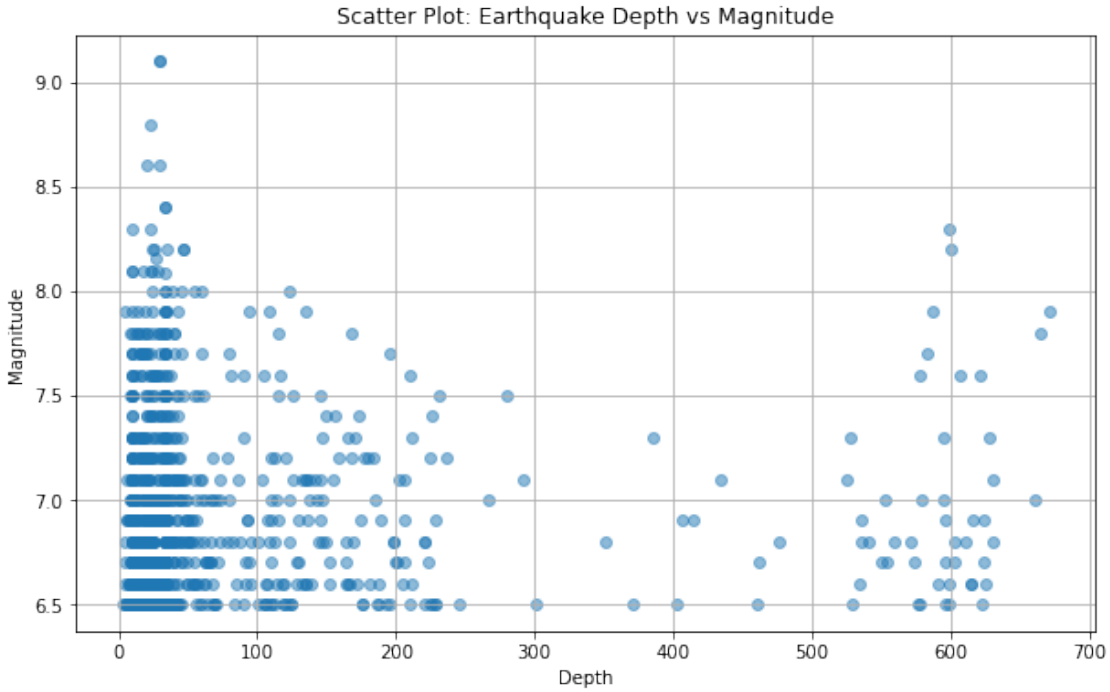



```
Index(['Algeria', 'Botswana', 'Canada', 'Italy', 'Kyrgyzstan', 'Martinique',
       'Mongolia', 'Mozambique', 'Russian Federation (the)', 'Saudi Arabia'],
      dtype='object', name='country')
```

Is there a relationship between earthquake depth and magnitude?

[47]:
```
import matplotlib.pyplot as plt

plt.figure(figsize=(10, 6))
plt.scatter(dataset['depth'], dataset['magnitude'], alpha=0.5)
plt.title('Scatter Plot: Earthquake Depth vs Magnitude')
plt.xlabel('Depth')
plt.ylabel('Magnitude')
plt.grid(True)
plt.show()
```



7 Reference

- https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.scatter.html
- https://lightupthesky1111.medium.com/python-geographical-plotting-using-plotly-f18e3f590f7f

[ ]: