**DATA VISUALISATION ASSIGNMENT**

**B9DA106 CA 2**

|  |  |
| --- | --- |
| **Student ID** | **Name** |
| 10389713 | Abhinav Singh |
| 10509732 | Pushpa Palavesam |

**Interactive Visualisations of Earthquakes from 1965 to 2016**



Dataset: World Earthquakes Survey from 1965 to 2016

A dataset of Earthquakes around the World from the year 1965 to 2016 has been taken from the [United States Geological Survey](https://earthquake.usgs.gov/earthquakes/search/) website.

<https://earthquake.usgs.gov/earthquakes/search/>

The file Earthquake dataset.csv contains 23412 records and 21 attributes. The file includes data of earthquakes recorded across the globe from the year 1965 to 2016 and it contains significant attributes for visualising the earthquakes like:

* Geographic Latitude and Longitude
* Date of the event occurrence
* Depth in Km
* Magnitude in Richter Scale
* Location

The Data processing and visualisations have been done in Google Colaboratory using Python 3.

Data Processing

The Source dataset CSV has been uploaded to Github repository for Visualisation CA2 Group.

<https://raw.githubusercontent.com/AbhinavSingh16/10389713-10509732-CA2-EARTHQUAKES/master/Earthquake%20dataset.csv>

The data is read using colab from github which can be found in Appendix [Read CSV](#_30j0zll) .

Following pre-processing steps are performed on the data, the code can be found in [Appendix Data Pre-processing](#_1fob9te)

1. The data was processed for missing values using imputation using Python.
2. The Date attribute was converted to Date Type for Timeseries analysis
3. New attribute called ‘Year’ was added which was the year of event extracted from Date

Github Link

Our Group github repository can be found at below location with the following files:

1. Earthquake dataset.csv
2. Visualisation\_CA2\_Earthquakes.doc
3. Visualisation\_CA2\_Earthquakes.ipynb

<https://github.com/AbhinavSingh16/10389713-10509732-CA2-EARTHQUAKES>

**Overview of Development Process:**

The work was split evenly as we are only 2 members in the group. We referred our course’s Moodle page and Classroom Labs and searched different visualization libraries for python online and introduced interactive plotting techniques with plot.ly and altair.

|  |  |  |
| --- | --- | --- |
| Student | Task | Man days |
| 10389713 Abhinav Singh | Data preprocessing, Visualisation, Summary, Github | 13 |
| 10509732 Pushpa Palavesam | Data sourcing, Visualisation, Summary | 13 |

Initial Analysis Questions

Earthquakes are unavoidable natural disasters that have a huge impact on our lives as well as properties. With advancement in engineering and technology, numerous earthquakes have been recorded over recent years and these recordings have helped scientist understand the activity of earth much better.

This visualisation aims at visualising the earthquakes recorded from 1965 to 2016 and enable the users to understand the primary questions as follows:

1. **Overarching Question:** Which place had the most frequent earthquake activity?
2. What was the maximum Magnitude of earthquake ever recorded and where?
3. Does Magnitude affect the Depth of earthquakes?
4. Where are the fault lines of major earthquakes in world map?
5. What has been the major cause of earthquakes recorded?
6. How many nuclear explosions have caused earthquakes in recent years and where?
7. Which Seismic movement activity has greater depth and impact?
8. What are the different types of Earthquake Magnitudes and How often have each Occurred?
9. Where are the most number of earthquakes greater than a Magnitude of 5 been recorded and from what sources?
10. Which Attributes correlates to each other the Most?

Discoveries & Insights

Visualisation has been done using **Python 3 in Google Colaboratory environment**. All the interactive visualisations can be viewed from below link:

<https://colab.research.google.com/drive/1UxfuUECbj0lXVQ7V1TczU7JnAmnSmUw3>

Visualisation 1: Timeseries plot of Occurrence of Earthquakes

**Question**: Render an exploratory analysis of earthquake events from 1965 to 2016. Which year observed the most occurring events?

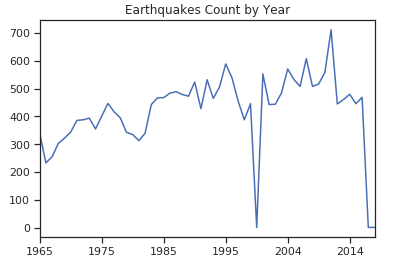


Figure 1:Timeseries plot of earthquakes

The code used for rendering the plot can be found in [Appendix Visualisation 1.](#_Visualisation_1:)

**Insight:** We can see that there hasn't been much activity in 2016 but the peak year of earthquakes has been in the year 2010.

Visualisation 2: Interactive Timeseries plot of Earthquake types

**Question**: What have been the causes of earthquakes over the given period of time and when?

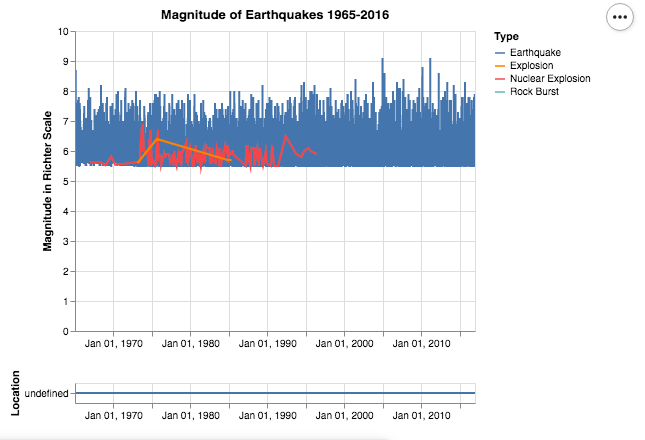


Figure 2: Interactive Timeseries plot of earthquakes causes

The code used for rendering the plot can be found in [Appendix Visualisation 2.](#_Visualisation_2:)

**Insight:** There has been no Nuclear Explosions big enough to cause earthquakes since 1996. And no Explosion since 1985. This concludes that the recent years have been peaceful to avoid natural disasters triggered by mankind.

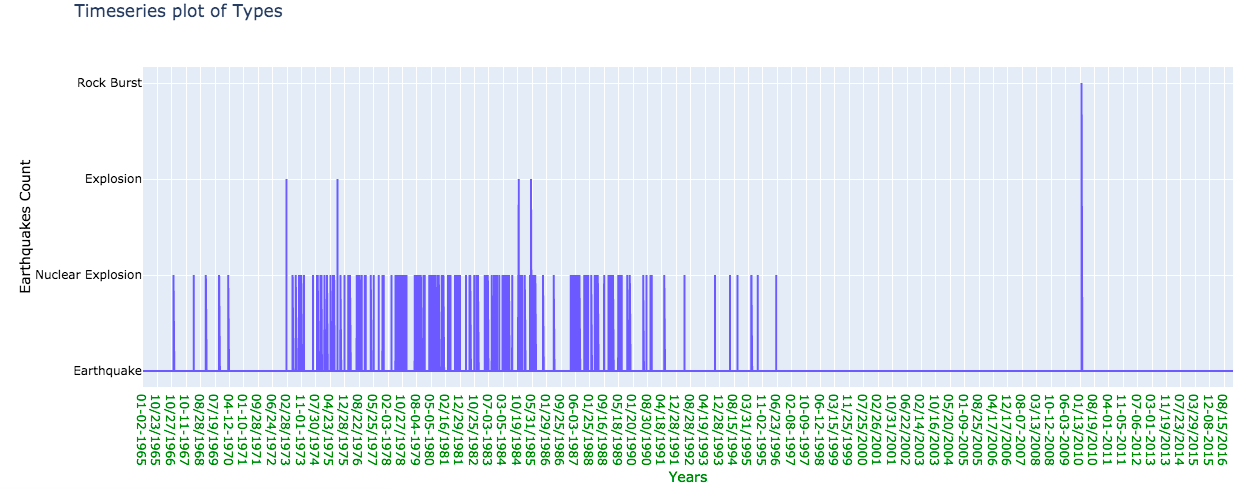


Figure 3: Interactive Timeseries of Types of earthquake

**Insight:** We can see that there has been a recent Rock burst that has caused earthquake in 2010. There have been nuclear explosions in 1973, 1975, 1984 and 1985 that have been the causes of earthquakes.

Let us see where the nuclear explosions have happened on Map visualisation 4.

Visualisation 3: Plot of Causes of Earthquakes

**Question**: What are the causes of Earthquakes? What has been the major cause of Earthquakes?

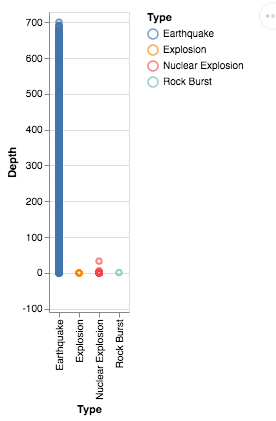


Figure 4: Bar plot of Causes of Earthquakes

The code used for rendering the plot can be found in [Appendix Visualisation 3](#_Visualisation_3:).

**Insight:** There are 4 nominal values. Earthquakes as natural disasters have been the major cause of earthquakes. Though there have been man-made cause like Nuclear Explosion and Explosion too.

Visualisation 4: World Map of Earthquakes and fault lines

**Question**: Visualise the earthquakes and its fault lines in world map. Which country have had the highest magnitude of earthquake recorded in world?

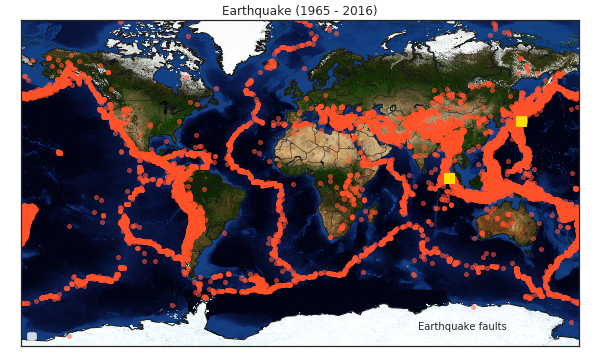


Figure 5: World Map of Earthquakes and Intensity

The code used for rendering the plot can be found in [Appendix Visualisation 4.](#_Visualisation_4:)

**Insight:** The Red points denote the faults lines of the earthquakes on world map. The Yellow squares are the location of earthquakes with highest magnitude recorded- Indonesia and Japan!

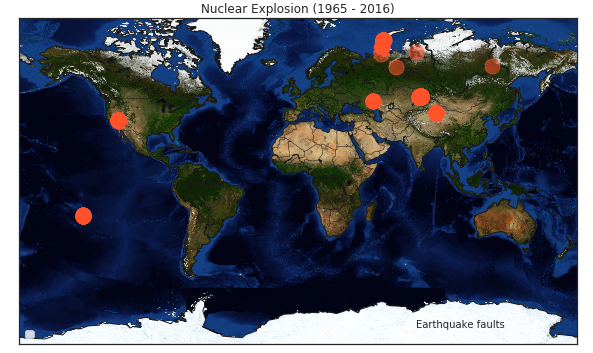


Figure 6: World Map of Nuclear Explosion induced Earthquakes

**Insight:** The Red points denote the areas in world recorded earthquakes due to Nuclear explosion. There has been Nuclear Explosion induced Quakes in US, Russia, Middle East.

Visualisation 5: Magnitude of Earthquakes Vs Frequency

**Question**: .What magnitude of intensity of earthquakes has been most recorded?

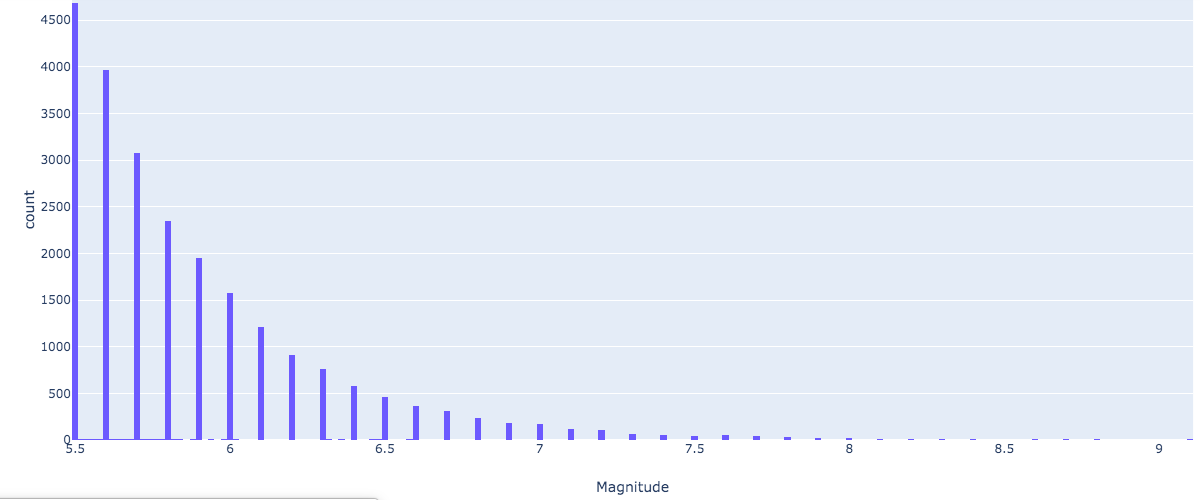


Figure 7: Magnitude Vs Frequency

The code used for rendering the plot can be found in [Appendix Visualisation 5.](#_Visualisation_5:)

**Insight:** 5.5 Richter scale has been the most frequently observed earthquakes, which is strong to move furniture. The Highest intensity recorded is 9.1 which is extremely violent and throws objects into air.

Visualisation 6: Magnitude of Earthquakes Vs Frequency

**Question**: Given the Magnitude Classes as below ordered factors:

Great: M > =8

Major: 7 < =M < 7.9

Strong: 6 < = M < 6.9

Moderate: 5.5 < =M < 5.9

Render a Magnitude Class distribution and which intensity of earthquake was most frequent?

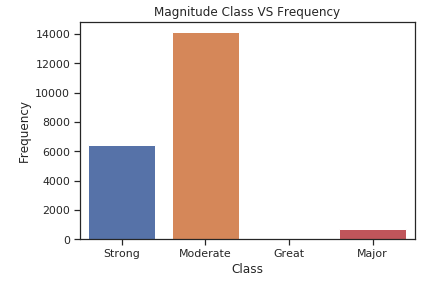


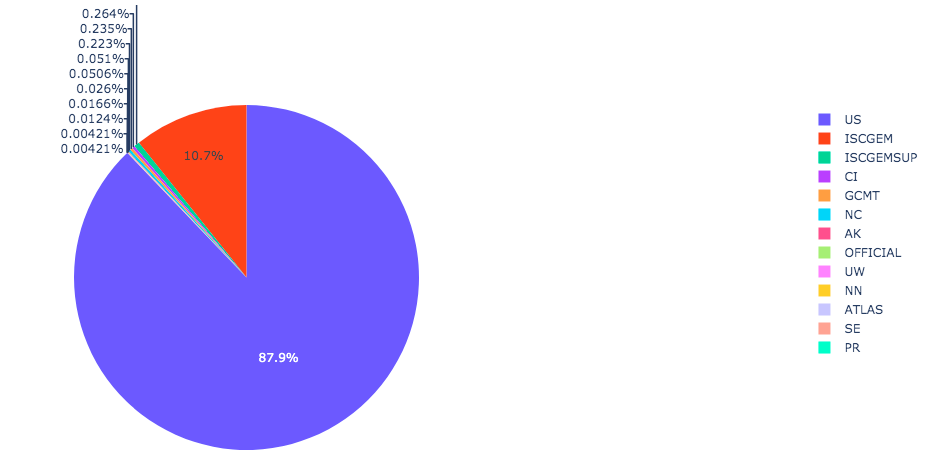
Figure 8: Bar plot of Frequency of Magnitude classes

The code used for rendering the plot can be found in [Appendix Visualisation 6.](#_Visualisation_6:)

**Insight:** Moderate Earthquakes have occurred the most and Great scale have been the least, which is >=8 on Richter scale and most devastating.

Visualisation 7: Interactive Pie Chart of Location by Magnitude/Intensity

**Question**: .What location recorded the most earthquakes?



**Figure 9:** Pie Chart of Location split by Magnitude of earthquake observed

The code used for rendering the plot can be found in [Appendix Visualisation 7.](#_Visualisation_7:)

**Insight:** US has recorded the highest number of earthquakes.

Visualisation 8: Depth of Earthquake in Km on 3D Axes

**Question**: How Deep does the earthquakes affect the crust of earth?

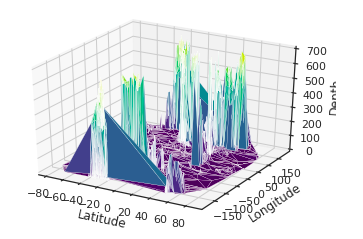


Figure 10:3D Axes plot of Depth on 2D canvas of Latitude-Longitude

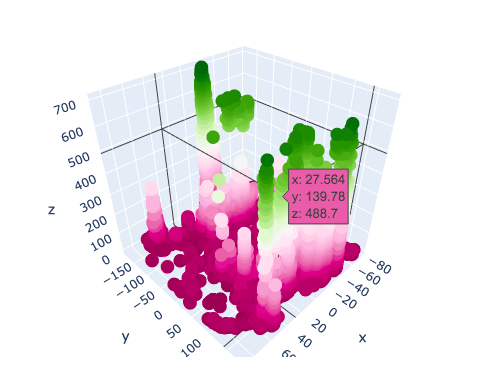


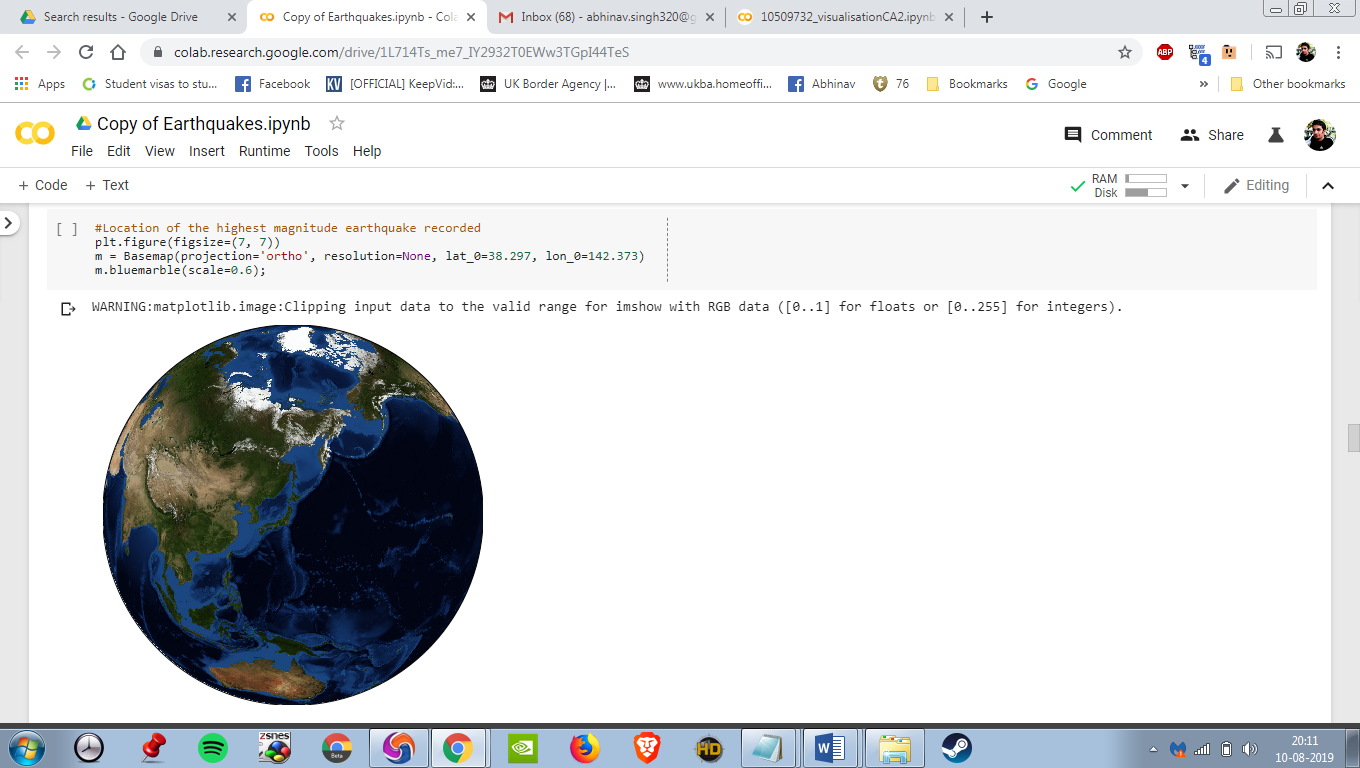
Figure 11: 3D Interactive plot of Depth on a 2D Canvas of Latitude and Longitude

The code used for rendering the plot can be found in [Appendix Visualisation 8.](#_Visualisation_8:)

**Insight:** There seems to be pretty Deep Impact on earth’s crust due to the Tectonic plate shifts. We can see how the Depth on z-axis mapped on a 2D canvas of Latitude and Longitude goes as deep as 600Km at some points on the earth’s crust

Visualisation 9: Location of last highest magnitude of earthquake by its coordinates

**Question: Where was the last maximum Magnitude of earthquake ever recorded?**



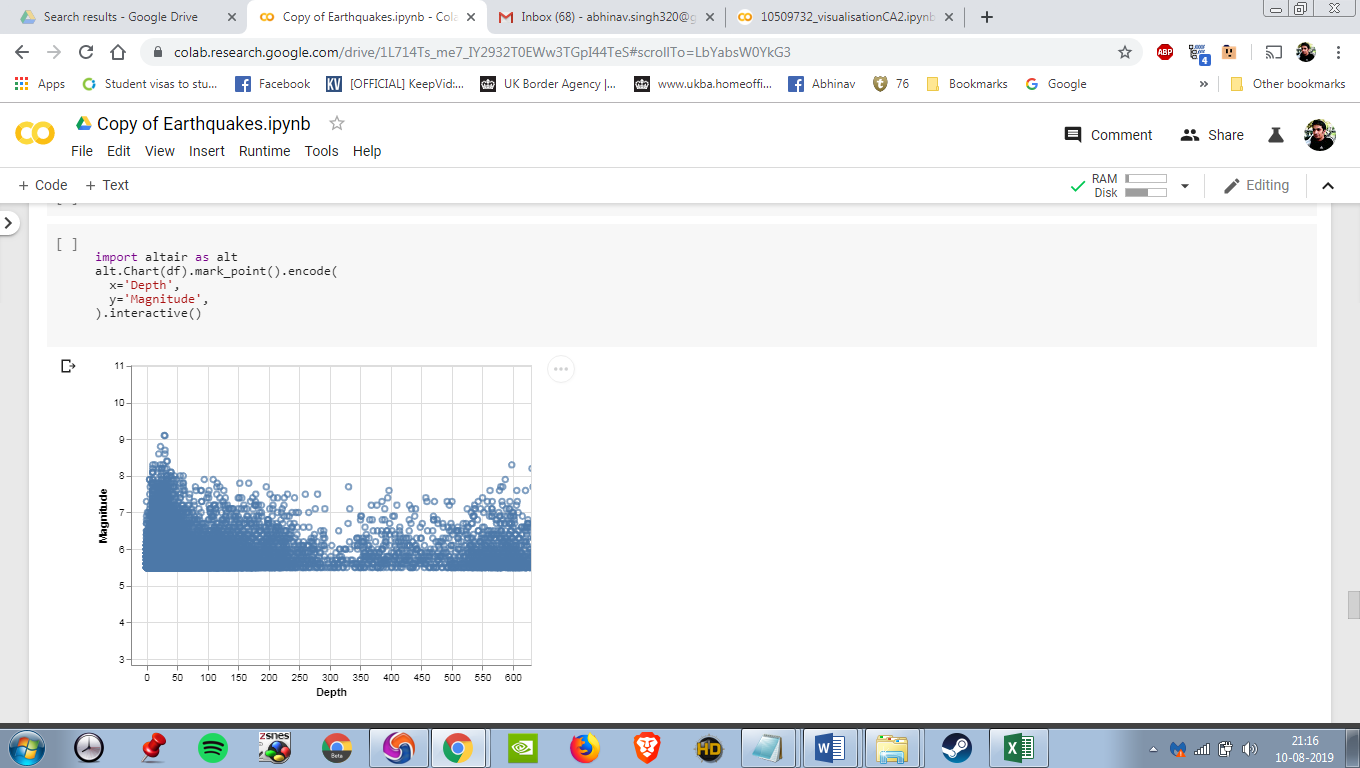
**Figure 12: geolocation of highest magnitude by its coordinates**

The code used for rendering the plot can be found in [Appendix Visualisation 9.](#_Visualisation_9:)

**Insight:** Location of the last highest magnitude earthquake recorded i.e. 9.1 on Richter scale in our dataset was on 3-11-2011 is shown as per its coordinates i.e. Latitude and Longitude which was near Japan in the Pacific Ocean.

Visualisation 10: Depth versus magnitude Scatter Chart

**Question: Does Depth affect the Magnitude of earthquakes?**



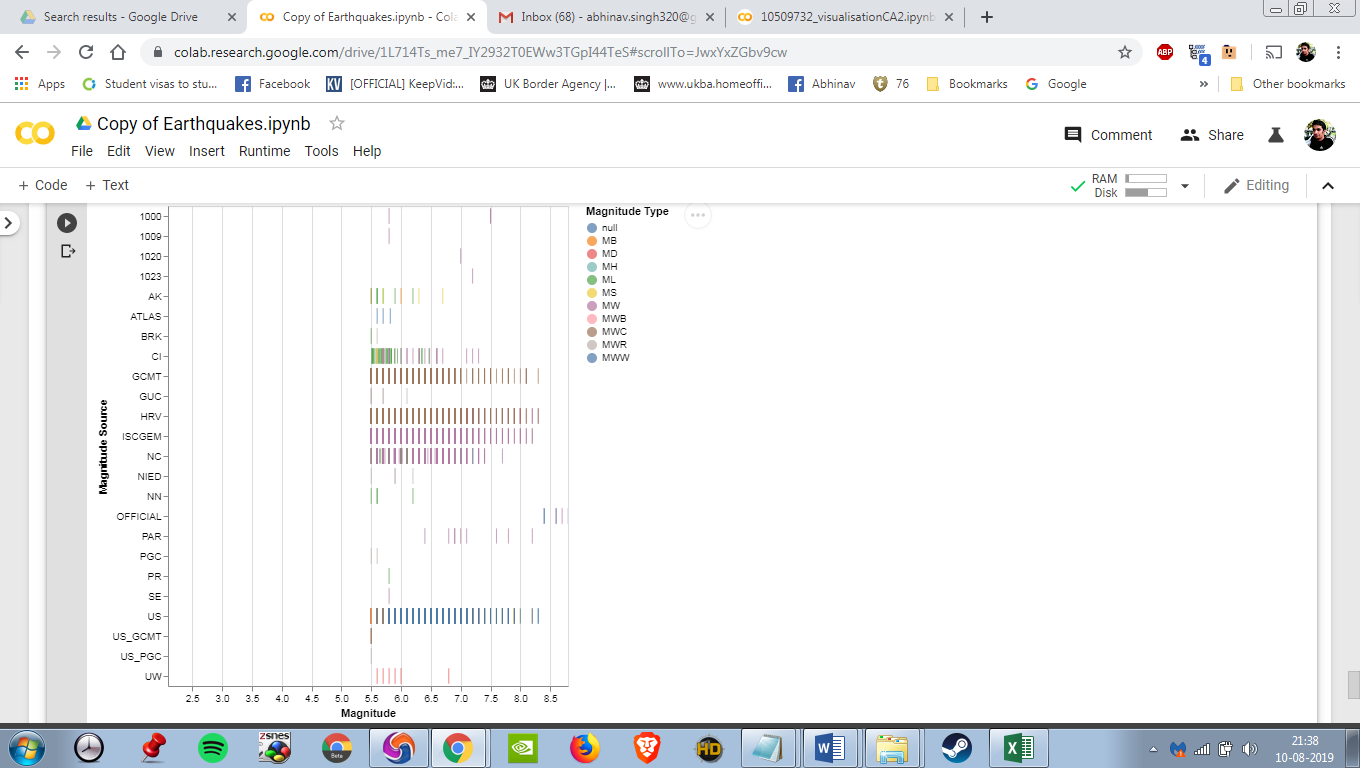
**Figure 13: Depth versus magnitude Scatter Chart**

The code used for rendering the plot can be found in Appendix [Visualisation10:](#_Visualisation10:)

**Insight**: From the above scatter plot we notice that the earth quakes can be very deep over 600 km but it doesn’t necessarily translates to high severity in magnitude as an earthquakes severity depends on other factors besides its strength such as its location and its distance from the epicentre. They can occur near the surface or deep within the Earth.

Visualisation 11: Magnitude Source and Type Ticker Plot

**Question: Where are the most number of earthquakes greater than a Magnitude of 5 been recorded and from what sources?**



**Figure 14**: Magnitude Source and Type Ticker Plot

The code used for rendering the plot can be found in Appendix [Visualisation](#_Visualisation_11:) 11:

**Insight:** The most number of Earthquakes above the Magnitude of 5 have been recorded in US which have the most number of earthquakes with MWW magnitude ,following US are the source HRV and GCMT with mostly MW and MWC magnitude Types.

Visualisation 12: Correlation Heat Map

**Question: Which Attributes correlates to each other the Most?**



**Figure 15**: Correlation Heatmap between Attributes

The code used for rendering the plot can be found in Appendix [Visualisation](#_Visualisation_11:) [12](#_Visualisation_12:)

We notice in the heat map that that there is a correlation between magnitude Seismic Stations and Depth Seismic Stations which is surprising as seen from above visualization that Magnitude and Depth are not entirely inter related.

Summary

There has been a drop in earthquakes towards 2016. 2010 has been the peak of earthquake activities in recent years. There has been no Nuclear Explosion induced Earthquakes since 1996 indicating peace among nations in recent years and avoiding mankind induced disasters. Natural Earthquakes due to tectonic plate shift has been the major cause of earthquakes recorded. Nuclear Explosion, Explosion and Rock Burst contribute to a very meagre amount of earthquakes. Japan and Indonesia have recorded the highest ever earthquakes of Magnitude 9 in Richter scale which was quite devastating. Russia, Middle East, US etc. have recorded Nuclear Explosions induced earthquakes.5.5 Richter Scale has been the most frequently observed earthquakes, which is strong enough to move furniture. The Highest intensity recorded is 9 which is extremely violent and throws objects into air

Moderate Earthquakes have occurred most. Extremely violent earthquakes classified on Scale >=8 have been least. There seems to be pretty Deep Impact on earth’s crust due to the Tectonic plate shifts as deep as 600Km.

Appendix

This section contains all the Python code used in Google Colaboratory:

# Read CSV:

import pandas as pd

import numpy as np

import altair as alt

import matplotlib.pyplot as plt

import seaborn as sns

sns.set(style="ticks", color\_codes=True)

%matplotlib inline

from mpl\_toolkits.basemap import Basemap

import warnings

warnings.filterwarnings('ignore')

from datetime import datetime

!pip install https://github.com/matplotlib/basemap/archive/master.zip

!pip install pyproj==1.9.6

**Read the CSV from github** link using following command

url = 'https://raw.githubusercontent.com/AbhinavSingh16/10389713-10509732-CA2-EARTHQUAKES/master/Earthquake%20dataset.csv'

earthQuake = pd.read\_csv(url)

# Data Pre-processing:

#Since colab allows only maximum of 5000 rows, use below command to disable it.

alt.data\_transformers.disable\_max\_rows()

earthQuake.head()

earthQuake.shape

earthQuake.info()

earthQuake.isnull().sum()

#Impute the missing values. Here all the missing values (NaN) are numeric values.

#Extract the Year from Date column and create a new attribute called Year

earthquakes = earthQuake.dropna(axis = 1)

earthQuakes['Year'] = earthQuakes['Date'].str[6:]

# Visualisation 1:

earthQuakes["Year"][earthQuakes.Type =="Earthquake"].value\_counts().sort\_index().plot(kind = 'line')

plt.title('{} Count by Year'.format('Earthquakes'))

# Visualisation 2:

!pip install plotly

!pip install chart-studio

import plotly.graph\_objs as go

from plotly.offline import init\_notebook\_mode,iplot

init\_notebook\_mode(connected=True)

def configure\_plotly\_browser\_state():

import IPython

display(IPython.core.display.HTML('''

<script src="/static/components/requirejs/require.js"></script>

<script>

requirejs.config({

paths: {

base: '/static/base',

plotly: 'https://cdn.plot.ly/plotly-latest.min.js?noext',

},

});

</script>

'''))

top = alt.Chart(earthQuakes).mark\_line().encode(

x = alt.X('Date:T', timeUnit = 'yearmonthdate', title = None),

y = alt.Y('Magnitude:Q', title = 'Magnitude in Richter Scale'),

color = alt.Color('Type:N'),

tooltip = ['Magnitude:Q', 'Location:O', 'Depth:O','Type:O', alt.Tooltip('Date:T', format = '%d-%m-%Y')]

).properties(

title = 'Magnitude of Earthquakes 1965-2016'

).interactive(bind\_y = False)

bottom = alt.Chart(earthQuakes).mark\_line().encode(

x = alt.X('Date:T', timeUnit = 'yearmonthdate', title = None),

y = alt.Y('Location:O'),

order = alt.Order('Year:T'),

tooltip = ['Magnitude:Q', 'Location:O', 'Depth:O','Type:O', alt.Tooltip('Date:T', format = '%d-%m-%Y')]

).interactive(bind\_y = False)

chart = (top & bottom).configure\_legend(labelLimit = 0).configure\_axisY(grid=True)

chart

earthquakeTs = go.Line(x=earthQuakes.Date, y=earthQuakes.Type)

quakeslayout = go.Layout(title='Timeseries plot of Types', xaxis=dict(title='Years',color='green'),

yaxis=dict(title='Earthquakes Count',color='black'))

quakesFig = go.Figure(data=[earthquakeTs], layout=quakeslayout)

configure\_plotly\_browser\_state()

iplot(quakesFig)

# Visualisation 3:

alt.Chart(earthQuakes).mark\_point().encode(

x='Type',

y='Depth',

color='Type'

).interactive()

# Visualisation 4:

def creat\_map(Type, size = 4):

fig = plt.figure(figsize=(10, 10))

fig.text(.8, .3, 'Earthquake faults', ha='right')

m = Basemap(projection='mill',llcrnrlat=-80,urcrnrlat=80, llcrnrlon=-180,urcrnrlon=180,lat\_ts=20,resolution='c')

m.drawcoastlines()

m.drawcountries()

m.bluemarble()

x, y = m(list(earthQuakes[earthQuakes.Type == Type].Longitude), list(earthQuakes[earthQuakes.Type == Type].Latitude))

points = m.plot(x, y, "o", markersize = size, color = 'tomato', alpha = .5)

if Type is 'Earthquake':

a,b = m(list(earthQuakes[(earthQuakes.Type == Type) & (earthQuakes.Magnitude == earthQuakes.Magnitude.max())].Longitude),

list(earthQuakes[(earthQuakes.Type == Type) & (earthQuakes.Magnitude == earthQuakes.Magnitude.max())].Latitude))

points = m.plot(a, b, "s", markersize = 10, color = 'gold', alpha = 1)

plt.title("{} (1965 - 2016)".format(Type))

plt.legend(loc ='lower left', prop= {'size':11})

plt.show()

creat\_map('Earthquake')

creat\_map('Nuclear Explosion', size = 15)

# Visualisation 5:

!pip install plotly.express

import plotly.express as px

tips = px.data.tips()

fig = px.histogram(earthQuakes, x="Magnitude")

configure\_plotly\_browser\_state()

fig.show()

# Visualisation 6:

earthQuakes.loc[earthQuakes['Magnitude'] > 8, 'Class'] = 'Great'

earthQuakes.loc[ (earthQuakes['Magnitude'] >= 7) & (earthQuakes['Magnitude'] < 7.9), 'Class'] = 'Major'

earthQuakes.loc[ (earthQuakes['Magnitude'] >= 6) & (earthQuakes['Magnitude'] < 6.9), 'Class'] = 'Strong'

earthQuakes.loc[ (earthQuakes['Magnitude'] >= 5.5) & (earthQuakes['Magnitude'] < 5.9), 'Class'] = 'Moderate'

sns.countplot(x="Class", data=earthQuakes)

plt.ylabel('Frequency')

plt.title('Magnitude Class VS Frequency')

# Visualisation 7:

labels = earthQuakes['Source'].values

values = earthQuakes['Magnitude'].values

trace = go.Pie(labels=labels, values=values)

layout = go.Layout(title='Earthquake Cause with Magnitude')

fig = go.Figure(data=[trace], layout=layout)

configure\_plotly\_browser\_state()

iplot(fig)

# Visualisation 8:

from mpl\_toolkits.mplot3d import Axes3D

!pip install ipympl

%matplotlib widget

import ipympl

import matplotlib.pyplot as plt

fig = plt.figure()

ax = fig.gca(projection='3d')

ax.plot\_trisurf(earthQuakes['Latitude'], earthQuakes['Longitude'], earthQuakes['Depth'],cmap=plt.cm.viridis, linewidth=0.2)

ax.set\_xlabel('Latitude')

ax.set\_ylabel('Longitude')

ax.set\_zlabel('Depth')

plt.show()

data = [go.Scatter3d(x = earthQuakes["Latitude"],y = earthQuakes["Longitude"],z = earthQuakes["Depth"],mode = 'markers',

marker=dict(color = earthQuakes["Depth"]))]

fig = go.Figure(data=data)

configure\_plotly\_browser\_state()

# Render the plot.

iplot(fig)

# [Visualisation](#_Visualisation_9:) 9:

plt.figure(figsize=(7, 7))

m = Basemap(projection='ortho', resolution=None, lat\_0=38.297, lon\_0=142.373)

m.bluemarble(scale=0.6);

# Visualisation 10:

import altair as alt

alt.Chart(df).mark\_point().encode(

x='Depth',

y='Magnitude',

).interactive()

# Visualisation 11:

alt.Chart(earthQuakes).mark\_tick().encode(

x='Magnitude',

y='Magnitude Source',

color='Magnitude Type'

).interactive()

# Visualisation 12:

plt.figure(figsize = (16,5))

sns.heatmap(df2.corr(), cmap="coolwarm", annot=True)

plt.show()