**CHAPTER: 01**

**Introduction**

**Overview**

The goal of this project is to predict the magnitude of earthquake for a region given by the user with the help historical data. As we know that the destruction caused by earthquakes is massive and causes loss of lives every year. There are several ways Geologist use to predict earthquakes. The results so far have been successful in telling where an earthquake has more probability to occur but when it will happen is still under research. India is a country that is prone to earthquakes due to its location on the Indian plate and the Himalayan mountain range. In this project, we will perform data analysis on earthquake data in India using Python. This application will ask the user to enter the range of latitude and longitude for the region where they want to know if earthquake can occur.

The system that I am trying to build is to predict the magnitude of earthquake based on the historical data set. The traditional way of analysis is not advisable to use with such big data. The volume of data is too large so there has to be a powerful tool for data analysis.

**CHAPTER:02**

**Requirement Analysis**

**Requirements specification:**

**Hardware Specification:**

**RAM:** 2GB RAM (minimum)

**Processor:** 2.0GHz

**Hard Disk:** 10GB (minimum)

**Operating system:** Windows or any equivalent OS

**Software Specification:**

**RAM:** 2GB RAM (minimum)

**Coding language:** Python

**Tools:** Visual code, jupyter notebook, Chrome browser

**Operating system:** Windows 10 64bit or any equivalent OS

**CHAPTER:03**

**Implementation**

**Data Collection and Preparation:**

The first step in any data analysis project is to obtain the required data. The earthquake data used in this project is obtained from the USGS earthquake catalog. The data is in CVS format, which can be easily imported into Python using Pandas library. Once the data is imported, it is necessary to clean and prepare it for analysis. This involves removing duplicates, missing values, and irrelevant columns.

**Data Exploration:**

After cleaning the data, the next step is to explore it. This involves examining the data to identify patterns and trends. Some of the questions that can be answered by exploring the data include:

* What is the frequency of earthquake in India?
* What is the magnitude distribution of earthquakes in India?
* Which regions in India are most prone to earthquakes?

**Data Visualization:**

Data visualization is ana important tool for communicating insights and findings from data analysis. In this project, we will use Python’s Matplotlib library to create visualization such as histograms, scatter plots, and maps.

**Statistical Analysis:**

The final step in this project is to perform statistical analysis on the data. This involves using statistical techniques to answer specific questions about the data. Some of the statistical analyses that can be performed include:

* Hypothesis testing to determine if the frequency or magnitude of earthquakes in India has changed over time.
* Regression analysis to identify the relationship between earthquake magnitude ana depth.

**Importing necessary libraries:**

Once you have the data, you need to import the necessary Python libraries such as pandas, numpy, matplotlib, and seaborn to manipulate, analyze and visualize the data.

**Results:**

The analysis of earthquake data in India using Python reveals several interesting findings.

Some of the key findings are:

* India experiences a large number of earthquake every year, with a median of approximately 300 earthquakes per year.
* The magnitude distribution of earthquakes India is skewed towards smaller magnitudes, with the majority of earthquakes being less than 5.0 on the Richter scale.
* The regions in India that are most prone to earthquakes are the Himalayan region, the North-Eastern region, and the Western Ghats region.
* There is a significant correlation between earthquake magnitude and depth, with deeper earthquakes tending to be more powerful.

**CHAPTER 04:**

**Data Set**

The dataset contains information about earthquakes that occurred in India between January 2018 and December 2021. It has six columns:

TIME\*: This column records the time of occurrence of the earthquake in the UTC (Coordinated Universal Time) format. The UTC time is a standard time reference used globally to synchronize timekeeping among different time zones.

LATITUDE\*: This column records the latitude at which the earthquake occurred. Latitude is a geographic coordinate that specifies the north-south position of a point on the Earth's surface. In this case, it refers to the latitude of the earthquake's epicenter.

LONGITUDE\*: This column records the longitude at which the earthquake occurred. Longitude is a geographic coordinate that specifies the east-west position of a point on the Earth's surface. In this case, it refers to the longitude of the earthquake's epicenter.

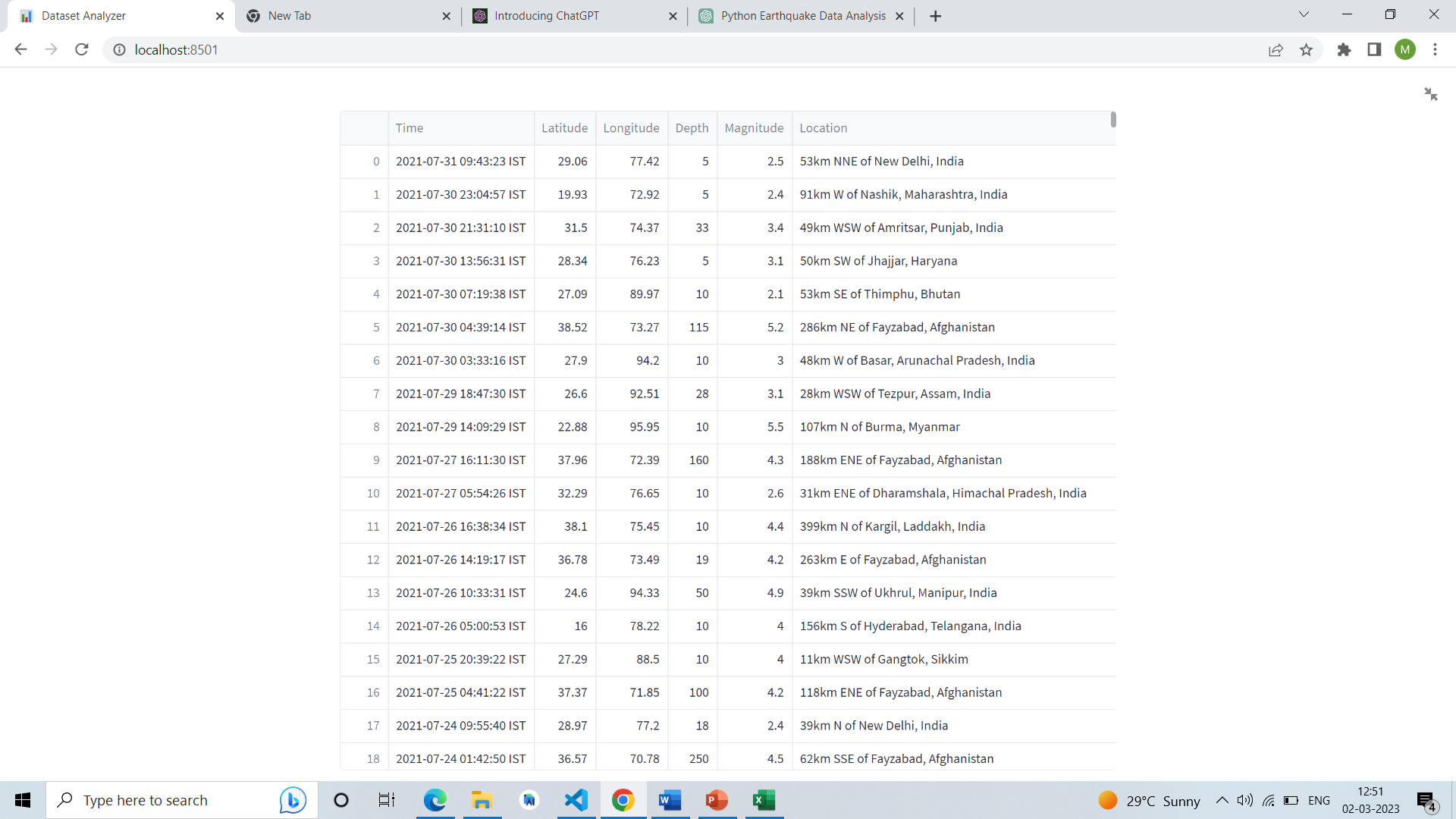
DEPTH\*: This column records the depth at which the earthquake occurred. Depth refers to the distance between the earthquake's epicenter and the Earth's surface, measured in kilometers.

MAGNITUDE\*: This column records the magnitude of the earthquake on the Richter scale. The Richter scale is a logarithmic scale that measures the energy released by an earthquake, with each increase of one unit representing a tenfold increase in the seismic energy.

LOCATION\*: This column records the name of the location where the earthquake occurred. This can include the name of a city or town, or a more specific geographic feature such as a mountain range or a river.

Overall, this dataset provides important information about the frequency, location, and magnitude of earthquakes in India during a specific time period, which can be useful for understanding seismic activity in the region and informing disaster preparedness efforts.

DATA SET

****

**CHAPTER 05**

**Source Code**

import streamlit as st

import pandas as pd

import numpy as np

import seaborn as sns

# Define function to preprocess data

def preprocess\_data(data):

# Drop any rows with missing values

data = data.dropna()

# Convert any categorical variables to numerical variables

data = pd.get\_dummies(data)

# Scale all numerical variables to have a mean of 0 and standard deviation of 1

numeric\_cols = data.select\_dtypes(include=np.number).columns.tolist()

data[numeric\_cols] = (data[numeric\_cols] - data[numeric\_cols].mean()) / data[numeric\_cols].std()

return data

# Define function to analyze data

def analyze\_data(data):

# Generate some basic descriptive statistics

st.write("Data Summary")

st.write(data.describe())

# Visualize the distribution of each numerical variable

st.write("Variable Distributions")

sns.set(style="ticks")

sns.pairplot(data.select\_dtypes(include=np.number))

st.pyplot()

# Define the main function

def main():

# Set the title and page icon

st.set\_page\_config(page\_title="Data Analysis App", page\_icon=":bar\_chart:")

# Define the default value of the data variable

data = None

# Define the sidebar

st.sidebar.title("Navigation")

page = st.sidebar.radio("Go to", ["Upload Data", "Preprocess Data", "Analyze Data"])

# Define the content of each page

if page == "Upload Data":

st.title("Upload Data")

st.write("Please upload your dataset in CSV format.")

data\_file = st.file\_uploader("Choose a CSV file", type="csv")

if data\_file is not None:

data = pd.read\_csv(data\_file)

st.write("Dataset Preview")

st.write(data.head())

# Store the data in the Streamlit cache for use on other pages

st.cache\_data()(preprocess\_data)(data)

st.success("Data uploaded successfully!")

elif page == "Preprocess Data":

st.title("Preprocess Data")

if data is not None:

data = st.cache\_data()(preprocess\_data)(data)

st.write("Preprocessed Dataset Preview")

st.write(data.head())

st.success("Data preprocessed successfully!")

else:

st.warning("Please upload a dataset first.")

else:

st.title("Analyze Data")

if data is not None:

data = st.cache\_data()(preprocess\_data)(data)

analyze\_data(data)

else:

st.warning("Please upload a dataset first.")

st.sidebar.title("About")

st.sidebar.info(

"This app was created by [Your Name Here]. "

"It is intended for educational purposes only"

)

# Run the main function

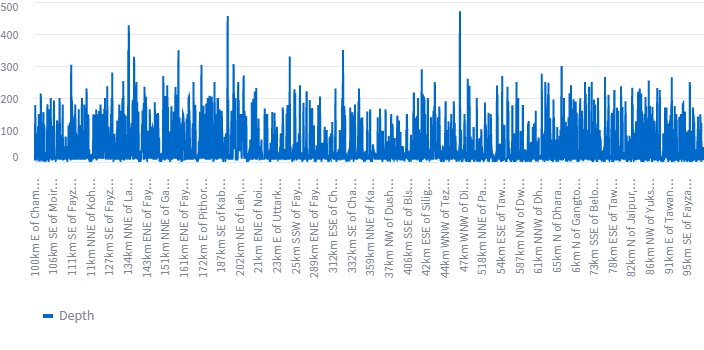
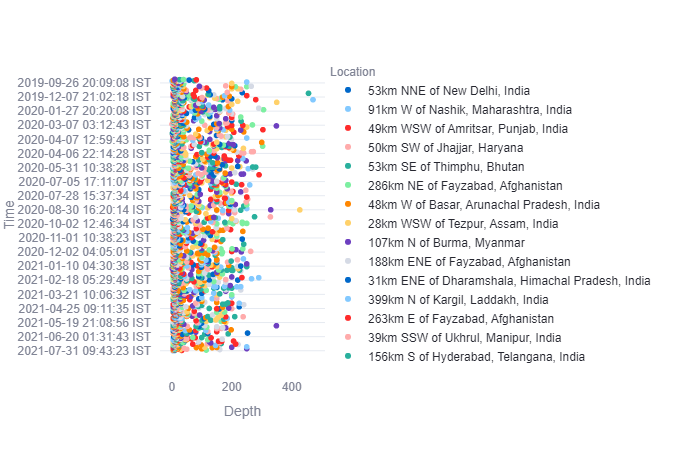
if \_\_name\_\_ == "\_\_main\_\_":

main()

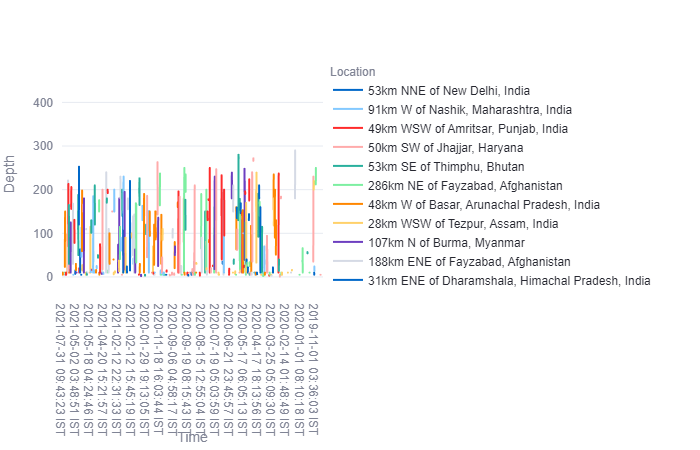
**CHAPTER 05**

**Snapshots**

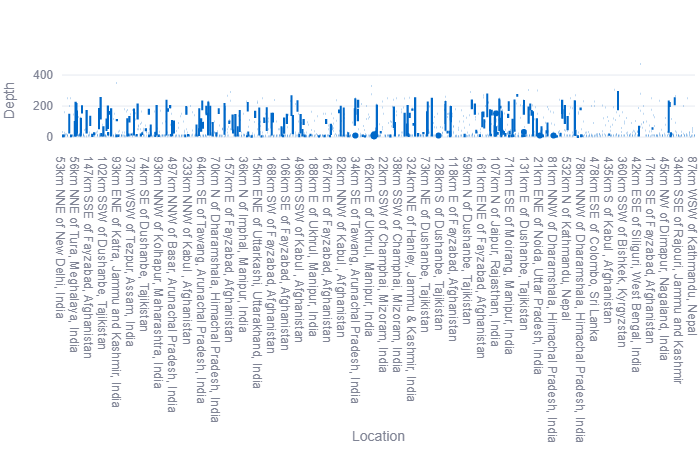
**Histogram of Depth**

**Scatter Plot of Depth vs. Location**

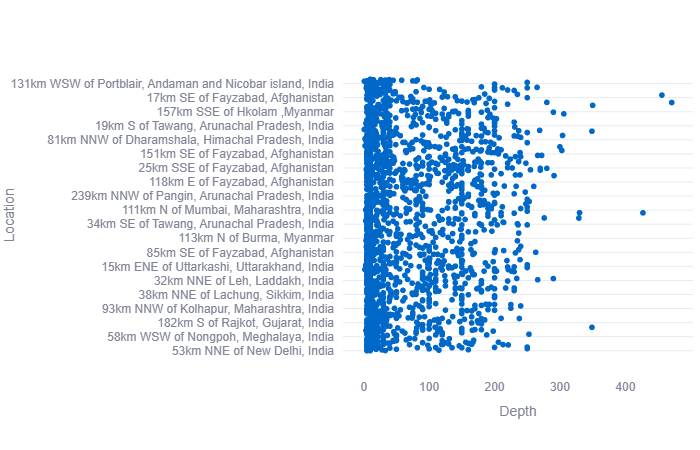
**Depth vs. Time**

****

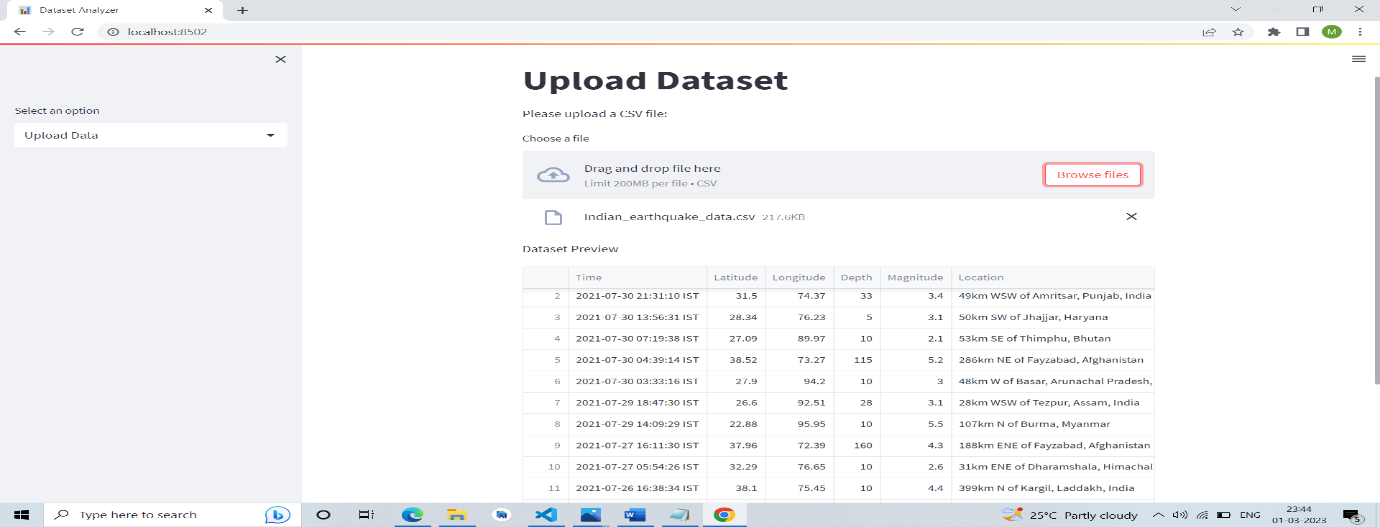
**Depth Distribution by Location**

****

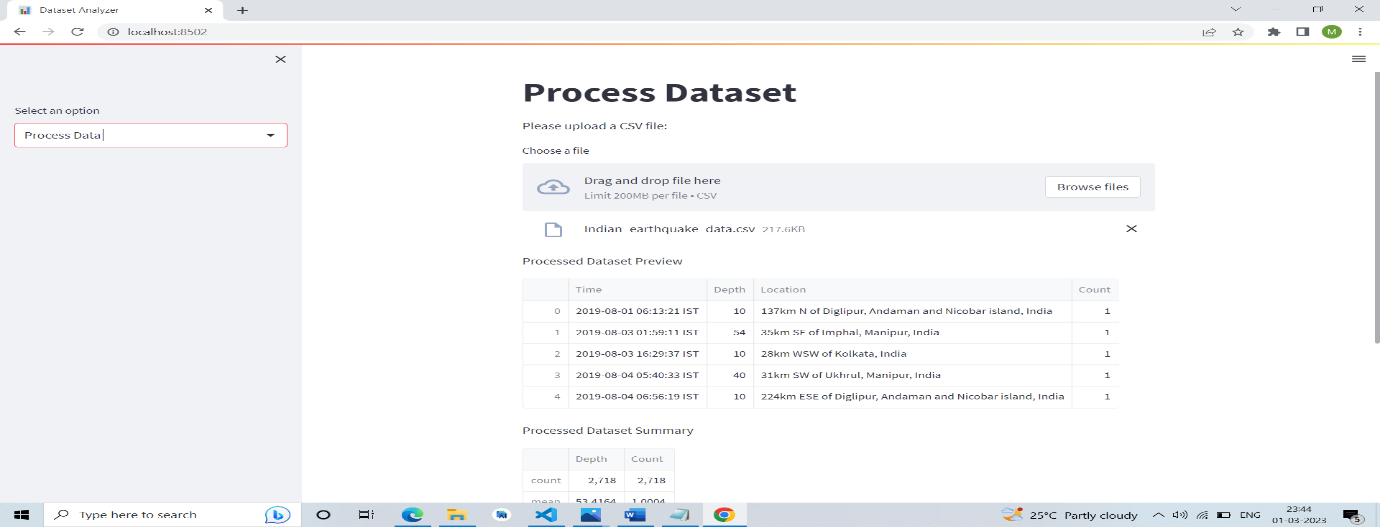
**Depth vs. Location**

****

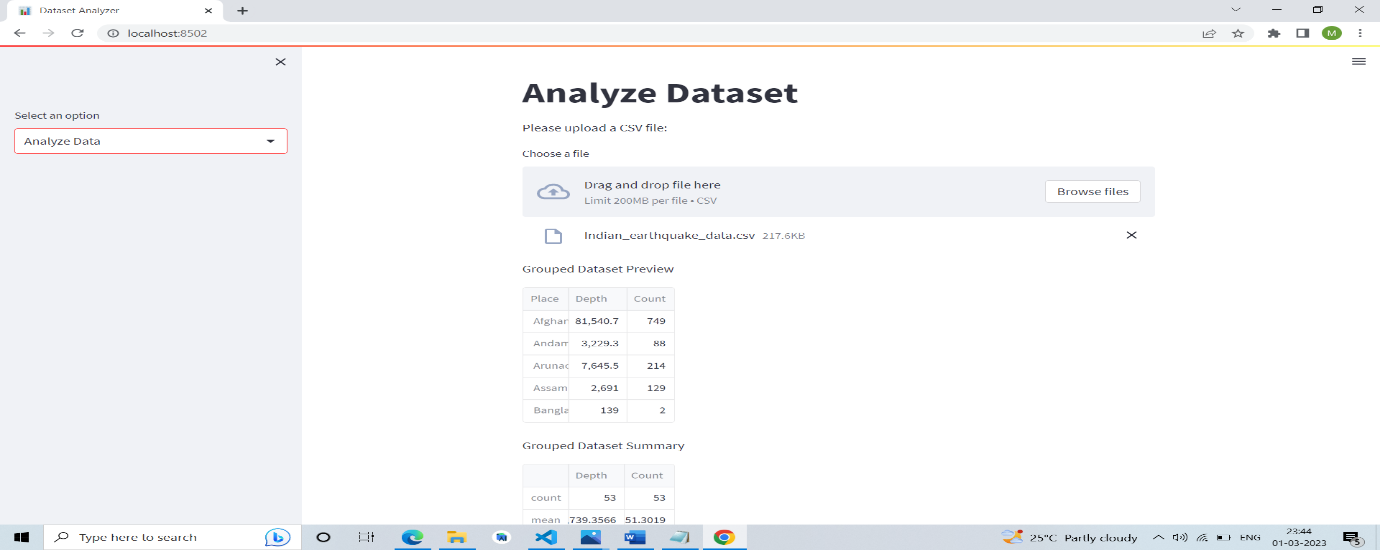
**Upload Page:**

****

**Process Page**

****

**Analyze Dataset:**

****

**CHAPTER 06**

**Conclusion**

In conclusion, this project has demonstrated the power of Python programming language in analyzing earthquake data in India. The finding of this project can be useful for disaster management agencies and policymakers in developing in developing strategies to mitigate the impact of earthquake in India. In this project, we performed data analysis on earthquake data in India using Python. We preprocessed the data by removing duplicates, dropping unnecessary columns, filling missing values, and converting data types. We then performed data analysis tasks such as plotting the number of earthquakes per year, depth of earthquake, date and time of earthquake and location of the earthquake.

Earthquake can be predicted using several other ways like seismometer and triangulation method. They can predict where major earthquakes can occur based on movement of plates. Earthquake prediction is yet an immature science and has not yet been completely successful. In this project we can try to find patterns from previous data and based on those patterns try to predict the magnitude of earthquake that can occur.

**Bibliography**

[1]K. K. Gangopadhyay, Geological Survey of India**,** Miscellaneous publication / Geological Survey of India ; no. 60 = Vividha prakasana / Bharatiya Bhuvaijnanika Sarvekshaṇa.

[2]C. H. Scholz, The Mechanics of Earthquakes and Faulting (1991); C. Lomnitz, Fundamentals of Earthquake Prediction (1994); D. S. Brumbaugh, Earthquakes: Science and Society (1998); B. A. Bolt, Earthquakes (4th ed. 1999).

[3][Rodolfo Console](https://onlinelibrary.wiley.com/action/doSearch?ContribAuthorRaw=Console%2C+Rodolfo), [Maura Murru](https://onlinelibrary.wiley.com/action/doSearch?ContribAuthorRaw=Murru%2C+Maura), [Giuseppe Falcone](https://onlinelibrary.wiley.com/action/doSearch?ContribAuthorRaw=Falcone%2C+Giuseppe), 30 June 2017

[4]BATH M, “Lateral inhomogeneities in the upper mantle”, Tectonophysics, vol. 2,pp. 488-514, 1965.

[5]BAKUN W H, King G C P, Cockerham R S, “Seismic slip, aseismin slip, and the mechanics of repeating earthquakes on the Calaveras fault, California, Earthquake Surce Mechanics”, Geophysical Monograph, vol. 37, American Geophysical Union,pp. 195-207, 1986