# Earthquakes analysis-final report

The code begins with the exploration and cleaning of a worldwide earthquake dataset. The dataset is loaded from a CSV file into a Pandas DataFrame, and initial exploration is performed to understand its structure and contents.

## Dataset Overview

The dataset contains information about earthquakes worldwide, including various parameters such as earthquake magnitude, location, date, and impact.

## Column Selection

To focus on relevant features for analysis, a new DataFrame, worldwide\_earthquake\_df, is created by selecting specific columns, including earthquake intensity, location (latitude and longitude), country, and other relevant information.

## Handling Null Values

The code addresses missing values by narrowing down the dataset to columns with the most significant information. The strategy is to drop columns with a substantial number of null values, considering the dataset's nature and characteristics.

## Data Types Conversion

The latitude and longitude columns are converted to numeric values, and non-numeric entries are replaced with NaN.

# Visualization of the Dataset

## Top 20 Countries with the Highest Earthquake Counts

A bar chart is generated to visualize the top 20 countries with the highest earthquake counts. The data is based on the selected columns, providing insights into the distribution of seismic activity globally.

## Geospatial Distribution of Earthquakes

An interactive map is created to visualize the distribution of earthquakes in the top 20 countries. The map uses hvplot and geoviews to display earthquake locations, including additional information like tsunami flags, year, focal depth, earthquake magnitude, and intensity.

## Top 20 Countries with High-Intensity Earthquakes

Another bar chart is generated to show the top 20 countries with earthquakes of intensity equal to or greater than 7.5. This provides insights into regions experiencing higher-magnitude earthquakes.

## Global Geospatial Distribution of All Earthquakes

An interactive map is created to visualize the geospatial distribution of all earthquakes in the dataset. The map includes information such as tsunami flags, year, focal depth, earthquake magnitude, and intensity.

# Analysing Patterns in Earthquake Occurrences

## Heatmap Based on Earthquake Intensity

A heatmap is generated to visually represent the distribution of earthquakes based on intensity. This provides insights into areas with higher earthquake magnitudes.

## Linear Regression: Earthquake Magnitude vs. Intensity

A scatter plot is created to analyze the linear relationship between earthquake magnitude (EQ\_PRIMARY) and intensity. Linear regression is performed, and the results indicate a positive moderate correlation between earthquake magnitude and intensity.

## Time Series Analysis: Earthquake Occurrences Over Time

Line plots are generated to analyze earthquake occurrences over time. The initial plot covers the entire dataset, while the second plot zooms in from the year 1900 onward. Linear regression is applied to the zoomed-in data, revealing an increasing trend in earthquake occurrences over the years.

## Trends in Earthquake Magnitudes Over Time

Line plots are generated to visualize trends in earthquake magnitudes over time. The plots suggest a decreasing trend in earthquake magnitude over time, with periodic fluctuations.

## Heatmap Comparing Earthquake Locations with the Ring of Fire

A heatmap is created to compare earthquake locations with the Ring of Fire. The map uses Folium to display a global map with a heatmap layer indicating earthquake concentrations.

## Earthquakes Triggering Tsunamis

A pie chart is generated to show the percentage of earthquakes triggering tsunamis. This provides an overview of the proportion of tsunamigenic earthquakes in the dataset.

# Conclusion and Insights

The analysis provides a comprehensive exploration of the worldwide earthquake dataset, offering insights into the distribution of earthquakes, their intensity, and trends over time. The visualizations help in understanding the geographical and temporal patterns of seismic activity, contributing to a broader understanding of earthquake occurrences globally.