

Exploratory Data Analysis (EDA) Report: Earthquake Data (1995-2023)

1. Introduction

This report presents an Exploratory Data Analysis (EDA) on global earthquake data spanning from 1995 to 2023. The analysis includes data cleaning, statistical exploration, and visualization to extract meaningful insights.

2. Dataset Overview

The dataset consists of 1000 records of earthquake events with features like magnitude, depth, location, continent, country, and various seismic measurements.

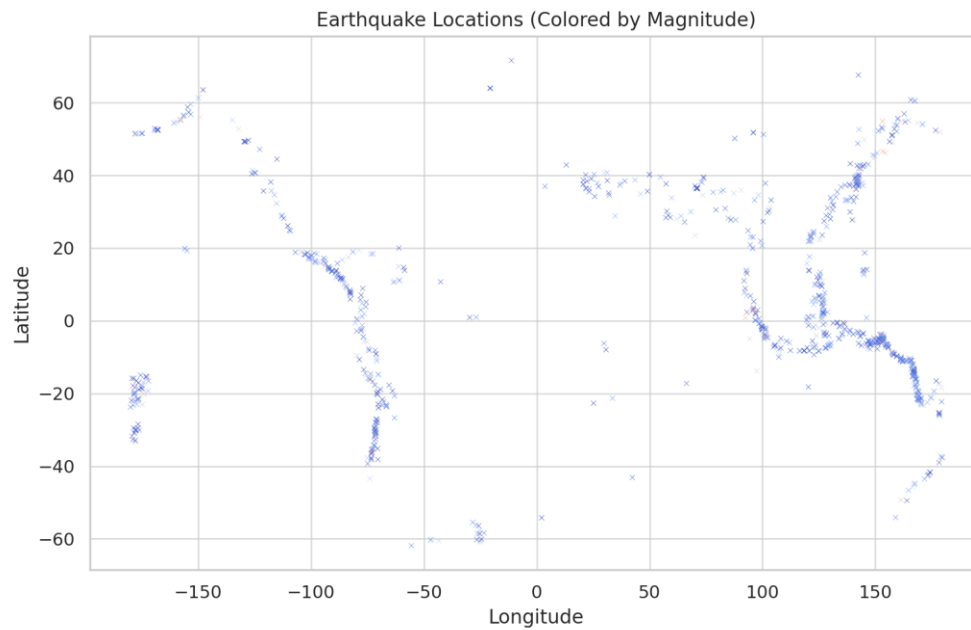
3. Data Cleaning

- Missing values in categorical columns were filled with the mode.
- Missing values in numerical columns were filled with the median.
- After cleaning, the dataset was complete without null values.

4. Visual Exploration

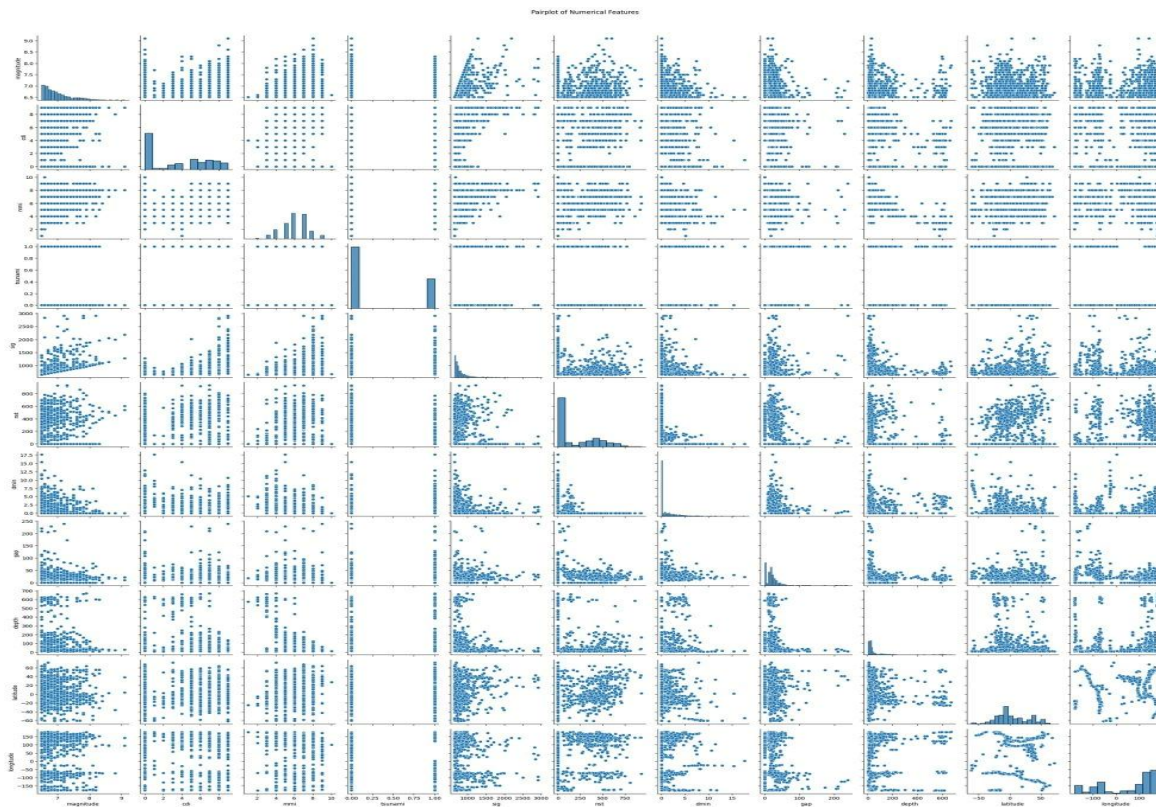
4.1 Earthquake Locations Scatter Plot

This plot shows the geographic distribution of earthquake epicenters around the globe.



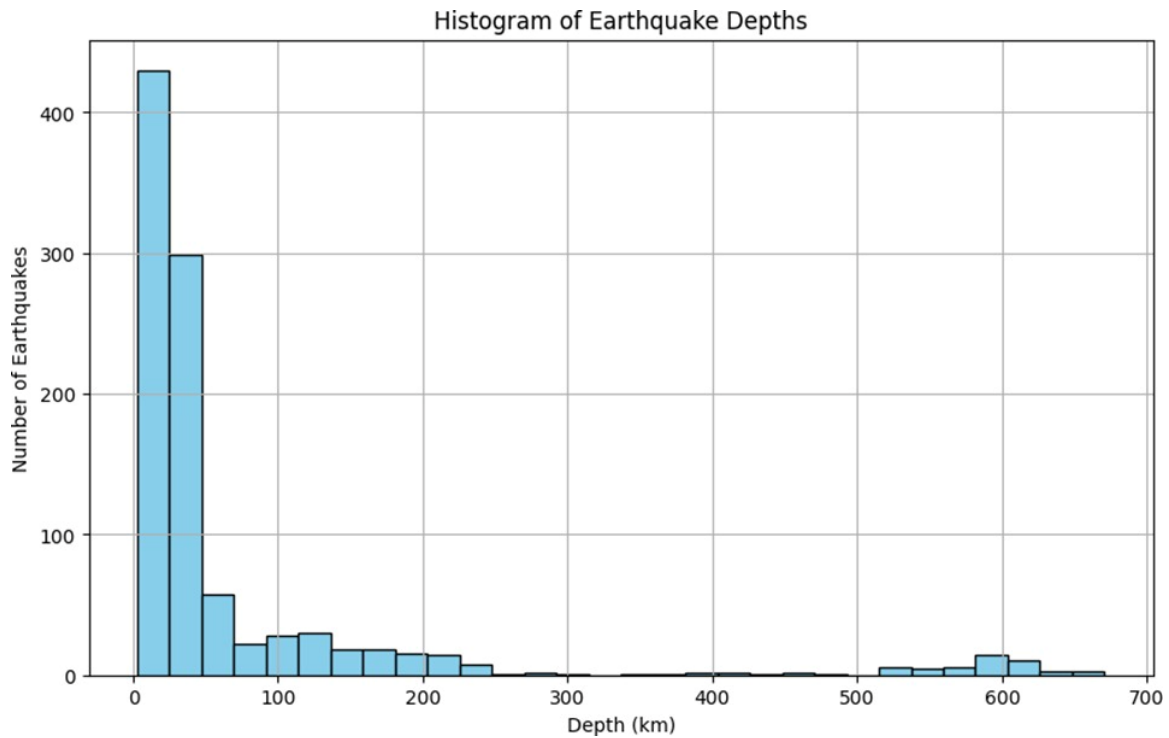
4.2 Pairplot of Numerical Features

Pairwise relationships among numerical features such as magnitude, depth, latitude, and longitude.



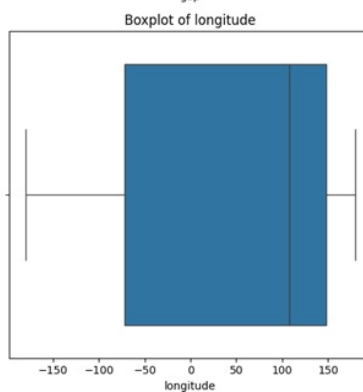
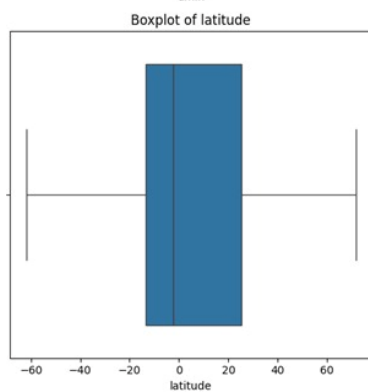
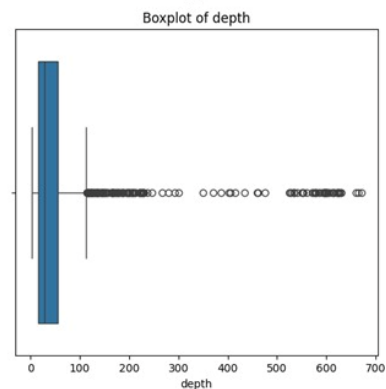
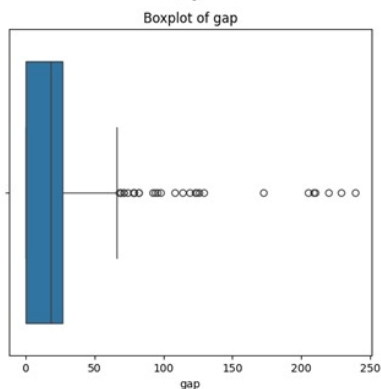
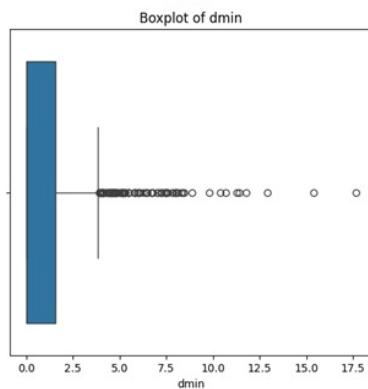
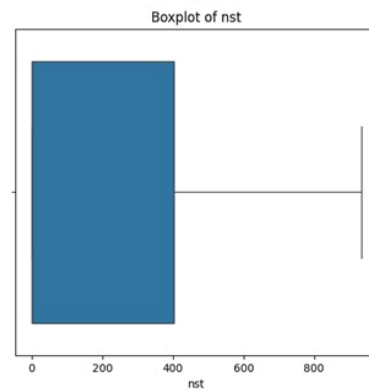
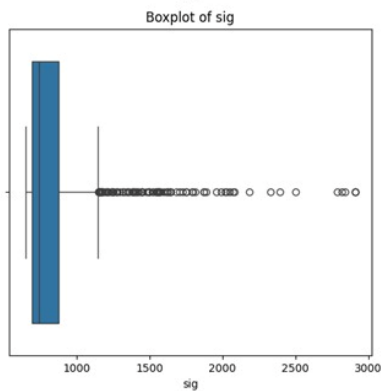
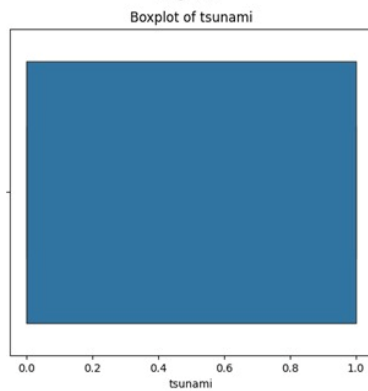
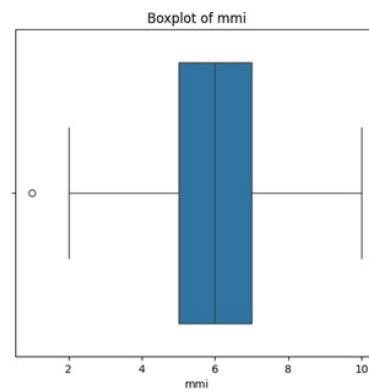
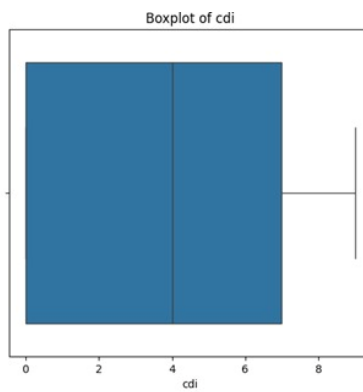
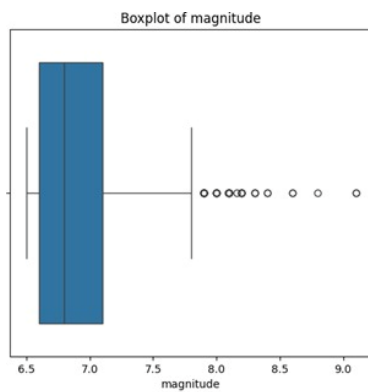
4.3 Histogram of Earthquake Depths

The histogram shows the distribution of earthquake depths, indicating most earthquakes are shallow.



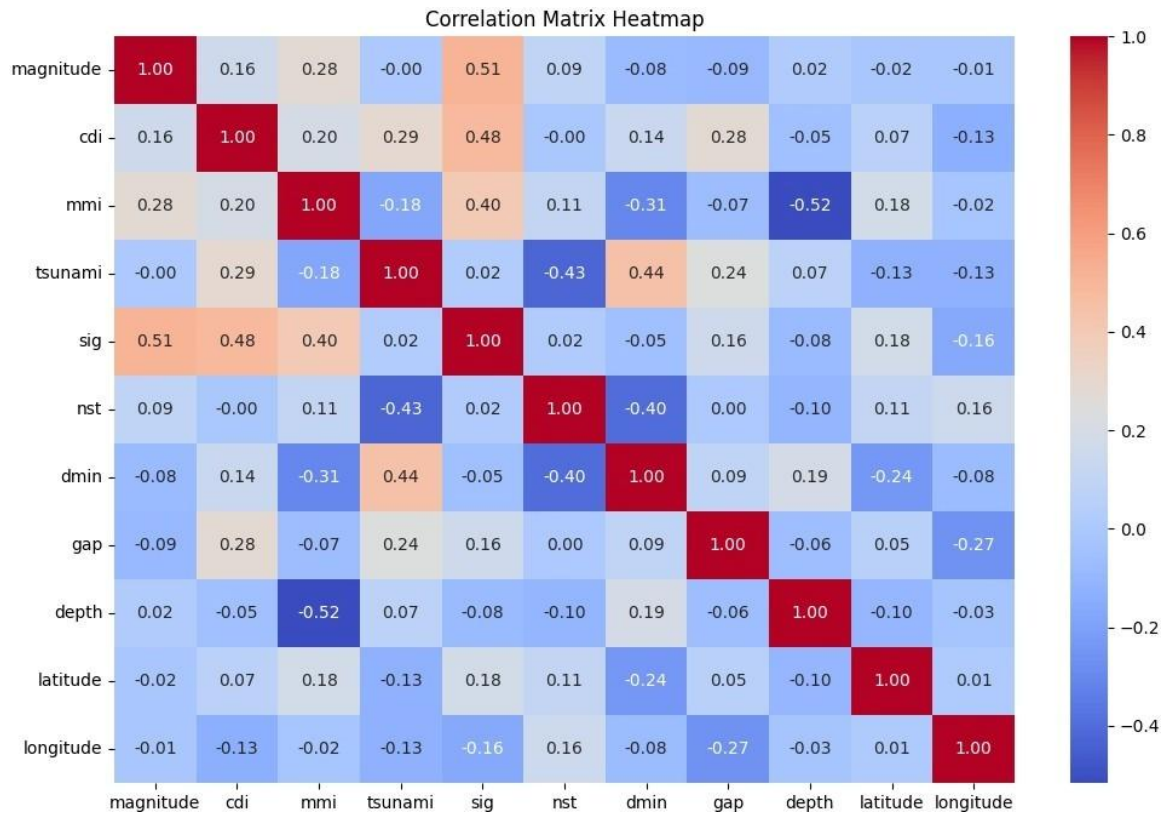
4.4 Boxplots of Numerical Columns

Boxplots reveal the spread and outliers of key numerical attributes.



4.5 Correlation Matrix Heatmap

Correlation between numerical features, showing weak to moderate correlations.



5. Conclusion

The EDA revealed that:

- Most earthquakes occur at shallow depths.
- The majority of events have moderate magnitudes (around 6.5-7.5).
- Certain geographic regions (e.g., Pacific Ring of Fire) are more prone to earthquakes.

Visualization helped uncover trends, detect outliers, and understand relationships among variables.