International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 6, pp 101-105, June-2025

https://doi.org/10.47001/IRJIET/2025.906011

Earthquake Analysis

[1]Siluveru Sanhith Kumar, [2]Gaikwad Pranay, [3]Dr. T Praveen Kumar, [4]Dr. Diana Moses

[1,2]UG Scholar, Department of AI & DS, Methodist College of Engineering and Technology, Hyderabad, India
[3,4]Professor, Department of CSE, Methodist College of Engineering and Technology, Hyderabad, India

***Abstract* - This study provides an in-depth analysis of a global earthquake dataset, offering valuable insights into seismic activity patterns and trends worldwide. The dataset includes key parameters such as magnitude, depth, geographical coordinates, and the date and time of occurrence, collected from reliable seismic monitoring networks over several decades. The aim is to understand the spatial and temporal distribution of earthquakes, identify high-risk zones, and explore correlations between earthquake characteristics. Using statistical analysis, geospatial visualization, and data mining techniques, the research reveals that regions along tectonic plate boundaries, such as the Pacific Ring of Fire, experience significantly higher seismic activity. The distribution of magnitude and depth highlights trends that are crucial for earthquake preparedness and risk mitigation strategies. Additionally, temporal analysis suggests clustering of events in some regions, potentially linked to aftershock sequences or foreshocks. The dataset also serves as a foundation for machine learning applications, including earthquake prediction models and anomaly detection. While precise prediction remains a complex challenge, data-driven approaches offer promising avenues for future research. The study emphasizes the importance of high-quality, well-structured seismic data for scientific research, public awareness, and policymaking. Continued efforts in refining data collection and analysis are essential to enhance resilience against future earthquakes and guide infrastructure development, urban planning, and emergency response strategies worldwide.**

Keywords: Earthquake Analysis, Earthquake dataset, Seismic monitoring.

I. LITERATURE SURVEY

Nwankwo et al. analyzed global earthquake data using tools such as Excel, Tableau, and Python to identify temporal and spatial patterns in seismic activity. Their work highlights that large-magnitude earthquakes are more frequent in tectonically active zones, with clustering seen in regions like the Pacific Ring of Fire. The findings suggest that understanding temporal spikes (e.g., increased activity during specific decades) can help inform international disaster readiness [1].

Garg et al. compared R and Tableau for handling large-scale earthquake datasets, revealing Tableau's superior performance in terms of speed, interactivity, and usability. While R offered more flexibility in modeling, Tableau's visual dashboards proved effective for quick pattern recognition and communication with non-technical stakeholders [2].

Gaur conducted a public perception survey comparing earthquake risk awareness with actual seismic activity statistics. Interestingly, in regions with minimal seismic activity, public fear was often high due to media amplification. This discrepancy emphasizes the need for better science communication and education on real vs. perceived risks [3].

Silva Atencio et al. used linear regression and time series models to analyze seismic risk in Central America, particularly Costa Rica. Their findings pointed to a rising earthquake frequency in urban regions such as the Gran Área Metropolitana. They advocated for transparent data sharing and collaborative planning between governments and communities [4].

Ajagbe et al. employed Folium and other mapping tools to visualize earthquake epicenters and intensity on choropleth maps. Their analysis of African and Asian tectonic zones demonstrated that interactive visualizations offer more granular insights into spatial risk distribution than traditional static maps [5].

Kane et al. examined the correlation between weather patterns and earthquake occurrences in North America from 1908 to 2014. Using Holt-Winters exponential smoothing, they found weak seasonal correlations but noted that temperature and barometric shifts could influence post-earthquake landslides and secondary disasters [6].

Nair et al. compared D3.js and Tableau in visualizing massive earthquake datasets, particularly for trend analysis across decades. They concluded that while Tableau is better for fast prototyping and interactive dashboards, D3.js provides more customizable, web-based visualization options suited to academic research [7].

Sharma et al. conducted a comparative study of seismic activity in regions like Japan, California, Chile, and Turkey, correlating earthquake frequencies with population density,

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 6, pp 101-105, June-2025

https://doi.org/10.47001/IRJIET/2025.906011

economic development, and fault proximity. Their results support using seismic zoning maps to guide urban development and emergency planning [8].

Liu et al. designed a prototype for a geospatial earthquake dashboard that integrates live seismic feeds with historical data, enabling dynamic risk assessment. Their review of existing systems revealed a lack of advanced GIS features, highlighting the potential for dashboards that support multi-layered data analysis and prediction [9].

Akhtar et al. evaluated Tableau's capability in visualizing time-series disaster data, particularly for COVID-19 and historical earthquakes. They demonstrated how real-time dashboards with predictive features could be repurposed for early warning systems in seismically active regions [10].

Feng et al. applied deep learning models such as LSTM and Prophet to earthquake time-series data from San Francisco, Japan, and Chile. They found that a 3-year training window offered optimal forecasting accuracy. Their future work aims to integrate graph mining and spatial analysis for multi-regional earthquake prediction [11].

Agarwal et al. used Rapid Miner and k-means clustering to segment historical earthquake data based on magnitude and location. Their analysis, covering the period from 1904 to 2011, revealed distinct seismic zones and highlighted declining trends in specific regions, potentially due to improved building codes and mitigation efforts [12].

Wang et al. proposed a model for earthquake impact prediction using neighborhood-level infrastructure and transportation data. Incorporating features such as proximity to critical facilities and mobility flow (analogous to taxi flow), their negative binomial regression model improved accuracy by over 17%, showcasing the value of integrating urban data with seismic analytics [13].

Feng et al. and Nwankwo et al. have also applied modern big data tools and machine learning techniques to earthquake forecasting. Their works show that hybrid approaches combining deep learning, historical pattern analysis, and socioeconomic data are increasingly effective in modeling seismic risks and supporting global readiness [14].

II. MATERIALS & METHODS

Datasets

The dataset comprises earthquake statistics from various regions around the world, detailing seismic events, their magnitude, location, and time. The dataset includes temporal information (date and time of the earthquake), geographic identifiers (regions, cities, and countries), and metadata such as magnitude, depth, and impact. The primary focus is on earthquake occurrences and their intensity, which can be critical for seismic risk assessments, disaster preparedness, and resource allocation for emergency response.

This dataset provides critical data for analyzing global seismic activity and helps in identifying patterns, predicting potential aftershocks, and assessing the long-term impacts on local populations and infrastructure. It can be used by seismologists, disaster management agencies, governments, and researchers in developing earthquake preparedness strategies.

Dataset Structure

The dataset has 16 columns and 1130 rows, though only about 305 rows contain valid data, with the remaining rows being either empty or filled with placeholder values. The dataset includes columns for earthquake data such as magnitude, depth, affected regions, and temporal data.

- **States:** Country or region where the earthquake occurred (e.g., California, Japan) (String)
- **Cities:** Name of the city closest to the earthquake's epicenter (e.g., Los Angeles, Tokyo) (String)
- **Magnitude:** The magnitude of the earthquake on the Richter scale (e.g., 5.6) (Float/String)
- **Depth:** The depth of the earthquake in kilometers (e.g., 10.5 km) (Float)
- **Latitude:** Latitude of the earthquake's epicenter (e.g., 37.7749) (Float)
- **Longitude:** Longitude of the earthquake's epicenter (e.g., -122.4194) (Float)
- **Duration:** Duration of shaking in seconds (e.g., 30) (Float)
- **Intensity:** Earthquake intensity based on the Modified Mercalli Intensity (MMI) scale (e.g., VII) (String)
- **Affected_population:** Estimated population impacted by the earthquake (e.g., 150,000) (Integer/String)
- **Fatalities:** Number of reported fatalities (e.g., 23) (Integer)
- **Injuries:** Number of reported injuries (e.g., 300) (Integer)
- **Damage_cost:** Estimated economic cost of damage caused by the earthquake (e.g., $1.5 billion) (String)
- **Unnamed:** 12: Empty/unused column (NaN) (Null/Unknown)
- **Date:** Date when the earthquake occurred (e.g., 04-Jan-23) (String - Date format)
- **Time:** Time when the earthquake occurred (e.g., 21:30) (String - Time format)
- **Area:** Region or area code where the earthquake was recorded (e.g., 7.0) (Float, could be Int)

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 6, pp 101-105, June-2025

https://doi.org/10.47001/IRJIET/2025.906011

III. SOFTWARE

Tableau

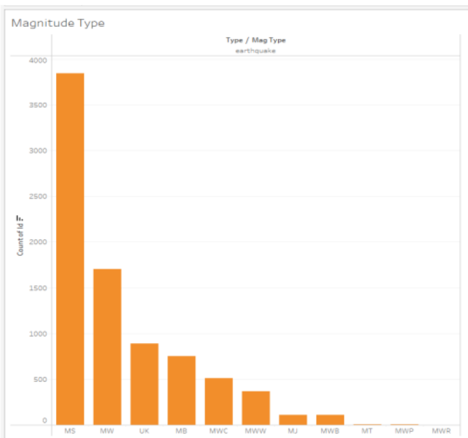
Tableau is a data visualization and business intelligence (BI) software that allows users to connect, visualize, and share data insights in an interactive, easy-to-understand way. It transforms raw data into visually appealing and insightful dashboards and reports without requiring any coding.

Key Features of Tableau includes Data Connectivity, Interactive Dashboards, Ease of Use (No Coding Required), Advanced Analytics, Sharing and Collaboration and Data Preparation.

Tableau is powerful data visualization and business intelligence tool designed to help users analyze, interpret, and visualize data. Its primary purpose is to transform raw data into interactive dashboards, charts, and graphs that are easy to understand and interpret. Tableau's key features include a drag-and-drop interface, real-time data analysis, forecasting, and advanced analytics such as trend identification and clustering. It also supports data blending, allowing users to combine data from various sources for more comprehensive analysis. With its ability to map geographic data, Tableau is highly effective for geospatial analysis, providing location-based insights.

The application of Tableau spans across a wide range of industries, aiding businesses in decision-making by providing insights into past and present performance, as well as customer behavior. It enables automated reporting, ensuring that stakeholders have timely access to critical data. Tableau also facilitates collaboration through Tableau Server and Tableau Online, allowing teams to share and discuss insights efficiently. Additionally, Tableau Prep helps clean and structure data before analysis, and calculated fields allow users to create customized metrics to suit specific analytical needs, enhancing the overall data analysis process.

IV. DATA VISUALIZATIONT

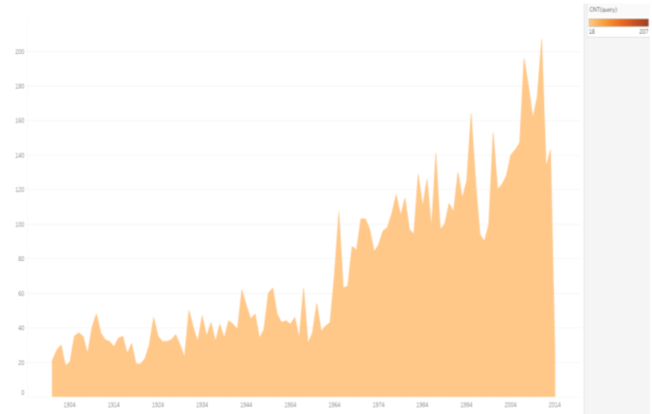


This graph will show how the distribution of earthquake magnitudes changes over time. It could highlight whether larger magnitude earthquakes have become more frequent in the 20th century.

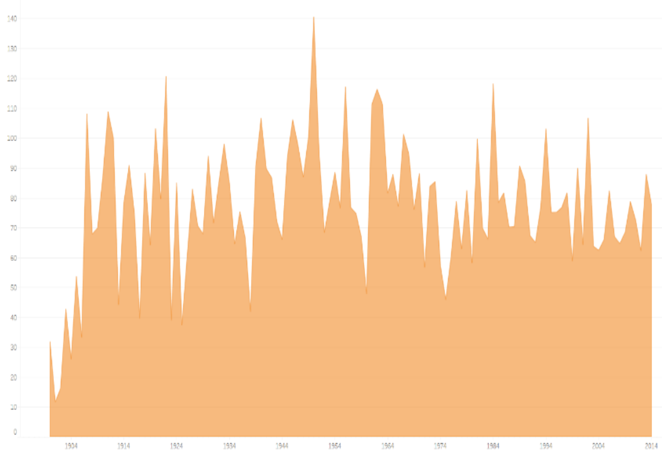
Magnitude type count

Mag T..	
MS	3,847
MW	1,706
UK	887
MB	754
MWC	510
MWW	367
MJ	111
MWB	108
MT	4
MWP	3
MWR	1

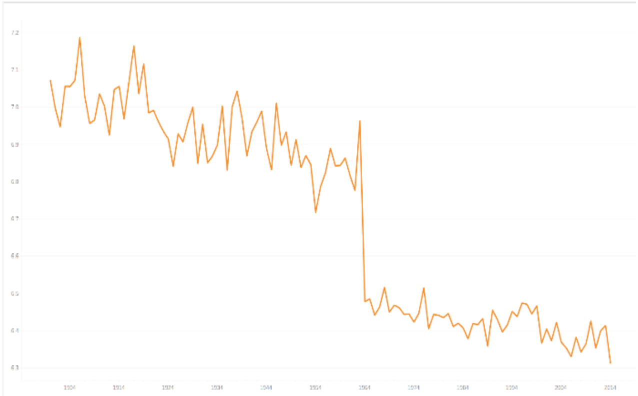
Identifying patterns of particularly high-magnitude earthquakes could help scientists understand whether there are any increasing trends or anomalies.

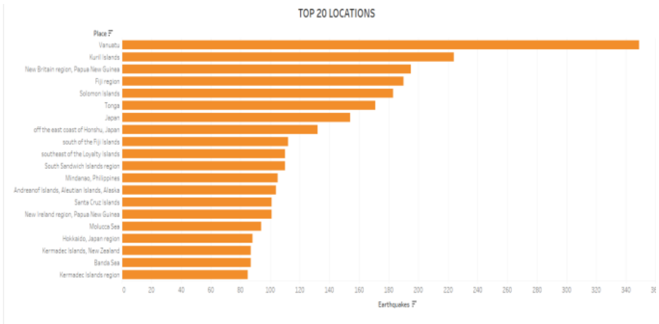


This graph will show the frequency of earthquakes occurring each year. It will help identify trends, such as whether earthquake occurrences have increased over time or if there were significant spikes in particular years.



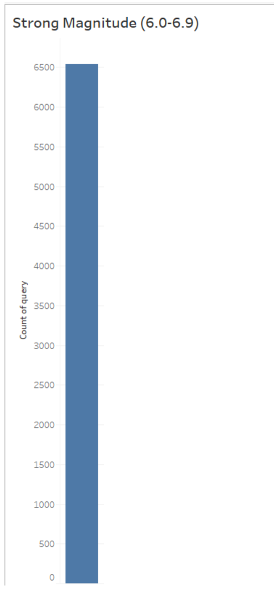
International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Issue 6, pp 101-105, June-2025
https://doi.org/10.47001/IRJIET/2025.906011



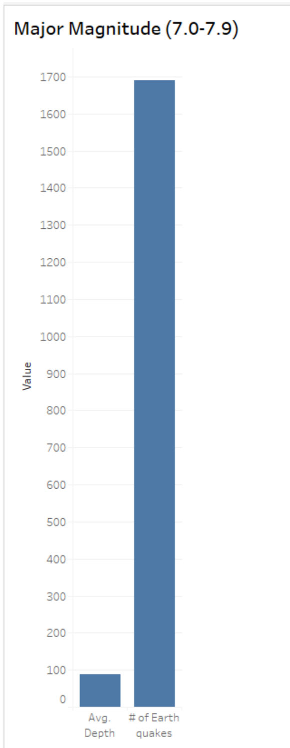


The chart highlights the top 20 earthquake-prone locations, with Vanuatu experiencing the highest number (~350), followed by the Kuril Islands and New Britain region of Papua New Guinea. Most of these regions lie along the Pacific Ring of Fire, known for intense seismic activity. Areas in the South Pacific, East Asia, and parts of Alaska and the South Atlantic are prominently affected. The data underscores the need for strong disaster preparedness, seismic monitoring, and resilient infrastructure in these vulnerable regions.

The map shows the top 5 locations of the most significant earthquakes globally, with events concentrated in high-risk seismic zones. These include Japan, the Aleutian Islands (Alaska), the Kuril Islands (Russia), the region near Indonesia, and the southern coast of Chile. All these locations are positioned along major tectonic boundaries, primarily within the Pacific Ring of Fire, indicating zones of intense geologic activity. This highlights the global distribution of major seismic events and the ongoing need for robust disaster mitigation efforts in these vulnerable areas.



The chart highlights a high frequency of strong magnitude earthquakes (ranging from 6.0 to 6.9), with **over 6,500 recorded events**. This magnitude range is considered capable of causing significant damage, especially in densely populated or poorly constructed areas. The large count suggests that strong earthquakes are relatively common globally and emphasizes the need for continuous seismic monitoring, infrastructure resilience, and emergency preparedness in vulnerable regions.



ISSN (online): 2581-3048
Volume 9, Issue 6, pp 101-105, June-2025
https://doi.org/10.47001/IRJIET/2025.906011

V. RESULTS & DISCUSSION

Based on the visual data provided across the screenshots, it is evident that global earthquake activity has shown a marked increase over the years, particularly from the mid-20th century onward. The line charts indicate a clear upward trend in the number of recorded earthquakes from 1904 to 2014, with a noticeable spike after the 1950s. This rise can be attributed both to improved seismic monitoring technology and a possible actual increase in tectonic activity. A decade-wise breakdown further supports this trend, showing higher counts of earthquakes in each successive decade, especially during the 2000s and 2010s. When examining earthquake magnitudes, the data reveals that the majority fall within the 6.0–6.9 range, totaling over 6,500 events, suggesting that while these are considered strong, they occur relatively frequently. In contrast, more powerful quakes in the 7.0–7.9 range are significantly fewer, numbering between 500 and 1,700 events, highlighting their comparative rarity. Overall, the data underscores an increasing frequency of seismic events over time, with moderate earthquakes being the most prevalent.

VI. CONCLUSION

- The dataset spans **over a century** (1904–2014), showing clear growth in detected seismic events, potentially due to both natural cycles and improved detection technologies.
- **Strong earthquakes (6.0–6.9)** are the most prevalent among large-magnitude quakes.
- **Major earthquakes (7.0–7.9),** while fewer in number, still show consistent activity globally.
- The **Pacific region** dominates in frequency of significant earthquakes, highlighting its seismic vulnerability.

REFERENCES

[1] Havskov, Jens, and Lars Ottemoller. "SEISAN earthquake analysis software." *Seismological Research Letters* 70.5 (1999): 532-534.

[2] Anagnos, Thalia, and Anne S. Kiremidjian. "A review of earthquake occurrence models for seismic hazard analysis." *Probabilistic Engineering Mechanics* 3.1 (1988): 3-11.

[3] Kan, Christopher L., and Anil K. Chopra. "Elastic earthquake analysis of torsionally coupled multistorey buildings." *Earthquake Engineering & Structural Dynamics* 5.4 (1977): 395-412.

[4] Fenves, Gregory L., and Anil K. Chopra. Earthquake analysis and response of concrete gravity dams. Berkeley, CA: University of California, *Earthquake Engineering Research Center*, 1984.

[5] Sever, Mehmet Şükrü, et al. "The Marmara earthquake: epidemiological analysis of the victims with nephrological problems." *Kidney international* 60.3 (2001): 1114-1123.

[6] Battul, Vaishnavi, M. Helen Santhi, and G. Malathi. "Earthquake data analysis and data visualization of Maharashtra state, India from 1912 to 2009 using R programming." *IOP Conference Series: Materials Science and Engineering*. Vol. 989. No. 1. IOP Publishing, 2020.

[7] Pun, W. K., and N. N. Ambraseys. "Earthquake data review and seismic hazard analysis for the Hong Kong region." *Earthquake engineering & structural dynamics* 21.5 (1992): 433-443.

[8] Pulinets, Sergey A., et al. "Irpinia earthquake 23 November 1980 Lesson from Nature reviled by joint data analysis." *Annals of Geophysics* 50.1 (2007).

[9] Mousavi, S. Mostafa, et al. "Stanford Earthquake Dataset (STEAD): A global data set of seismic signals for AI." *IEEE Access* 7 (2019): 179464-179476.

[10] Campbell, Kenneth W. "Bayesian analysis of extreme earthquake occurrences. Part I. Probabilistic hazard model." *Bulletin of the Seismological Society of America* 72.5 (1982): 1689-1705.

Citation of this Article:

Siluveru Sanhith Kumar, Gaikwad Pranay, Dr. T Praveen Kumar, & Dr. Diana Moses. (2025). Earthquake Analysis. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 9(6), 101-105. Article DOI https://doi.org/10.47001/IRJIET/2025.906011

****