**Modern Seismology and Geophysics Analysis**

**Elijah Taber, Michael Speer, Caleb Trimble**

**April 12, 2025**

**Abstract**

Earthquakes remain among the most significant natural hazards worldwide, posing an ever-present threat to human life, infrastructure, and economies (Federal Emergency Management Agency [FEMA], 2020). Monitoring seismic activities in real time and extracting meaningful insights from the data has become crucial in enhancing earthquake preparedness and risk mitigation efforts. This project leverages a publicly available dataset from the U.S. Geological Survey (USGS) (2023) to examine earthquakes recorded over the span of one month, seeking to identify key patterns, trends, and correlations that can inform both governmental and private organizations about seismic hazards and potential impacts.

Beyond immediate risk management, this project addresses important research questions about the underlying tectonic processes that generate earthquakes. By comparing our month-long dataset to historically reported patterns (Stein & Wysession, 2009), we can better understand the consistency or deviation of current seismicity from long-term norms. Preliminary findings suggest that while most seismic events occur along well-known tectonic plate boundaries, deviations occasionally emerge, emphasizing the dynamic nature of the Earth’s crust. Such anomalies can stimulate further scientific inquiries, prompting collaborations among seismologists, geologists, and data scientists to refine existing earthquake prediction models and develop novel approaches to detect pre-seismic signals.

This project not only provides a snapshot of global seismic activity within a one-month window but also shows the important role that data-driven methods play in contemporary hazard management, such as clearly showing how concentrated earthquakes can be within a specific cluster zone, giving insight into present-day unstable tectonic boundaries. From informing policy decisions at federal agencies to supporting risk assessment in the private sector, the insights derived from such an analysis carry profound implications for public safety, economic stability, and scientific advancement.

**Introduction**

Earthquakes are natural phenomena characterized by the sudden release of accumulated strain energy in the Earth’s lithosphere, resulting in ground shaking and potential surface rupture (Stein & Wysession, 2009). They typically occur along faults, fractures in the Earth’s crust where large blocks of rock move relative to each other, and are strongly associated with tectonic plate boundaries. Within these fault zones, stress gradually builds until it exceeds the strength of the rocks, triggering a seismic event. This accumulation and release of stress serve as fundamental processes that shape Earth’s surface, giving rise to mountain ranges, rift valleys, and other notable geological features. When an earthquake occurs, it often releases energy in the form of seismic waves: primary (P) waves, secondary (S) waves, and surface waves, that propagate outward, causing vibrations felt at varying distances from the source (Freed, 2005).

A key element in understanding earthquakes lies in properly interpreting their measurements, which predominantly include magnitude, depth, and geographical coordinates (Allen, Gasparini, Kamigaichi, & Böse, 2009). Magnitude provides a numerical estimate of the energy released at the quake’s source. One common scale is the Moment Magnitude Scale (M\_w), which estimates the total energy based on factors such as fault rupture area, the slip on the fault, and the rock rigidity. The Moment Magnitude Scale superseded the older and less comprehensive Richter Scale, offering more reliable measurements for large-magnitude quakes. An important point to note is that each increment of one step in magnitude corresponds to roughly 32 times more energy release (Stein & Wysession, 2009). Thus, a relatively small numerical increase in magnitude can represent substantially larger seismic energy, underscoring the need for precise measurements in assessing the potential destructive impact of earthquakes.

Geographicalcoordinates pinpoint an earthquake’s epicenter, the point on the Earth’s surface directly above the quake’s hypocenter (focus). Mapping these epicenters allows researchers to visualize clusters of seismic activity, revealing patterns such as alignment along plate boundaries, hotspots, and fault systems (Freed, 2005). Earthquake epicenters often accumulate along convergent boundaries, where plates collide, as well as divergent boundaries and transform faults. By studying these spatial patterns, scientists and policymakers can highlight high-risk zones, a critical step for effective disaster preparedness and urban planning.

Despite centuries of scientific inquiry, earthquake prediction remains a formidable challenge. Because of the complexity of fault systems, varying local geological conditions, and the chaotic nature of stress accumulation, predicting the exact time, location, and magnitude of an earthquake with precision is not feasible (Stein & Wysession, 2009). Instead, contemporary approaches focus on hazard assessment and early warning systems. Earthquake Early Warning (EEW) systems detect initial earthquake waves, particularly the faster but less damaging P waves, and rapidly transmit warnings to nearby populations before the arrival of the slower, more destructive S waves and surface waves (Allen et al., 2009). Although these warnings typically provide only seconds to tens of seconds of lead time, this short interval can be critical for individuals to seek cover, for trains to slow down, or for medical professionals to pause delicate procedures.

# **Business Problem/Hypothesis**

Given the central role of seismic data in both scientific analysis and disaster mitigation planning, this project aims to clarify how various features of earthquake events interact and whether patterns exist that can inform improved forecasting and risk assessment efforts. Specifically, we hypothesize that significant clustering of earthquakes near tectonic boundaries correlates with elevated magnitude and shallower depths. Furthermore, we seek to determine whether errors in magnitude and depth measurements are spatially non-uniform and whether human-reviewed data consistently demonstrates lower uncertainty. By answering these questions, the project contributes to the broader field of geoinformatics, where machine learning, spatial modeling, and geophysical principles intersect to improve our understanding of Earth’s dynamic processes (Huang et al., 2023).

The hypothesis also extends into temporal and categorical patterns in seismic activity. For example, it is anticipated that automatic classifications (i.e., machine-reviewed events) are more common for lower-magnitude quakes or those occurring in less densely populated regions where human oversight is limited. Similarly, we posit that temporal clusters may reflect aftershock sequences or episodic tectonic events. These insights have direct business implications: insurance firms, civil engineering consultants, and municipal planners can utilize these findings to refine risk maps, allocate resources, and reassess infrastructure resilience, particularly in zones showing high-frequency, shallow, and high-magnitude events (Worden et al., 2010).

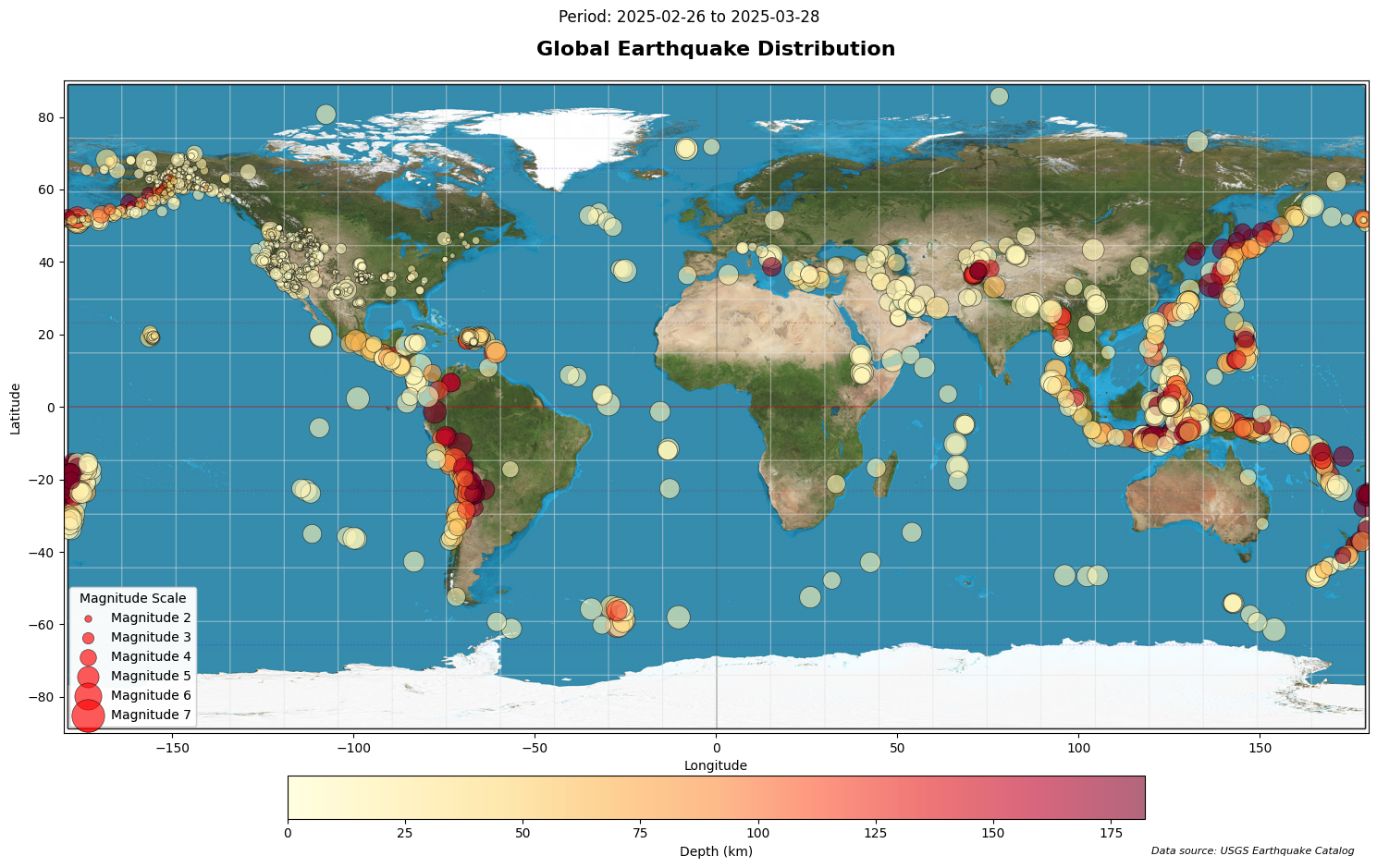
# **Methods/Analysis**

The data analysis pipeline began with exploratory data analysis (EDA), including summary statistics on magnitude, depth, and error values such as magError , depthError, and horizontalError. Visualizations provided insight into variable distributions and outliers, allowing for the assessment of overall data integrity. Next, the project employed time series techniques to examine how earthquake frequency varied over the month-long study period, highlighting daily seismic fluctuations and short-term surges that may indicate clustered geophysical activity or aftershock patterns. Spatial visualizations used geospatial mapping libraries to render the epicenter locations against tectonic plate boundaries, allowing for a visual inspection of spatial density and pattern alignment with major fault zones.

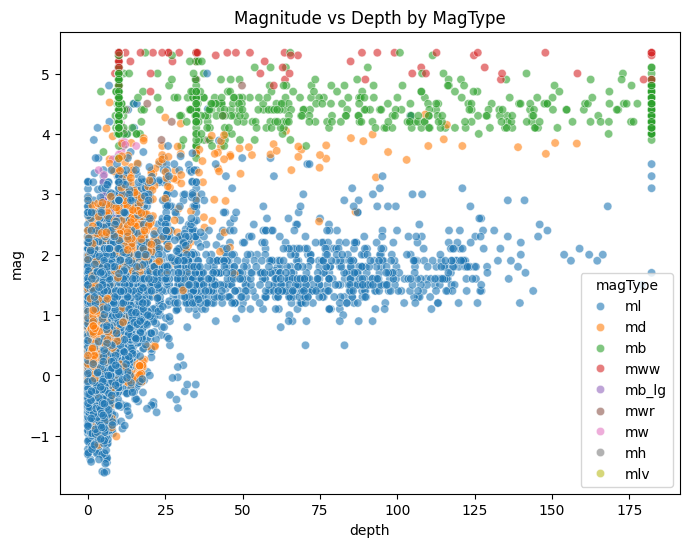
For deeper insights, the project utilized unsupervised clustering algorithms, particularly DBSCAN, to identify spatial clusters of earthquakes. This method, which groups points based on density and proximity, proved effective at revealing concentrated zones of activity not necessarily aligned with tectonic boundaries, suggesting the presence of minor faults or localized stress accumulation. Correlation matrices further assessed the relationship between variables, identifying moderate negative correlations between magnitude and depth, and strong associations between magError and automated status classification. This multifaceted approach allowed for robust pattern recognition and error evaluation, ensuring that findings were both statistically valid and geologically meaningful (Mignan & Broccardo, 2019).

# **Results**

The findings reveal several notable patterns. Earthquakes in the dataset predominantly occurred near major tectonic boundaries, with the Pacific Ring of Fire accounting for a large portion of seismic activity. The spatial clustering analysis corroborated this, showing that the most intense and numerous earthquake clusters corresponded to subduction zones and transform faults. Interestingly, shallow-focus earthquakes (depth < 70 km) composed a substantial majority of the dataset, aligning with expectations for destructive surface events. These events also tended to carry higher magnitude readings, particularly along active convergent boundaries, reinforcing the importance of localized stress conditions in seismic hazard evaluation.



Temporal analysis uncovered multiple peaks in daily earthquake occurrences, some of which could be linked to aftershock sequences. Notably, these periods also displayed a spike in measurement uncertainty—specifically higher depthError values, which may suggest either lower signal quality or rapid succession events that challenge sensor resolution. Additionally, a comparison between human-reviewed and automatic event classifications showed that automated classifications, while more frequent, generally had greater associated error metrics. This disparity underlines the continued necessity of expert review for high-risk events, even as machine-based classifications grow more prevalent.



The scatter plot of magnitude versus depth by magType shows key trends in seismic behavior. Most low-magnitude earthquakes (mag < 2) cluster at shallow depths (< 50 km), consistent with typical near-surface stress releases. In contrast, larger quakes (mag > 4) appear across a broader depth range, likely reflecting deeper fault activity. The distribution of magType also varies: local magnitude (ml) dominates shallow, weaker events, while moment magnitude (mww) and body wave magnitude (mb) are more common in deeper or stronger quakes, suggesting that different magnitude scales are suited to different tectonic settings.

# **Recommendations/Ethical Considerations**

* **Enhance Automated Classification Systems**: Invest in improving machine-learning algorithms for automatic earthquake classification, especially to reduce measurement error in lower-magnitude and shallow-depth events, which showed higher uncertainty in this analysis.
* **Prioritize Sensor Deployment in Underreported Regions**: Address global reporting skew by increasing the density and coverage of seismic sensors in regions with limited seismic data, enabling a more accurate global understanding of seismic risks.
* **Promote Public Awareness on Coastal Seismic Risks**: Encourage public awareness campaigns and pre-travel advisories about seismic hazards in coastal or tectonically active regions to support safer decision-making for tourists and residents alike.
* **Expand Temporal Data Analysis for Trend Detection**: Support long-term seismic monitoring to detect patterns and anomalies over extended periods, which could refine risk models and potentially identify precursory seismic behavior more reliably.

# **Conclusion**

This project demonstrates how data-driven analysis of seismic activity can offer meaningful insights into earthquake behavior and associated risks. By examining a one-month dataset from the USGS, we identified clear spatial clustering along tectonic boundaries, a dominance of shallow earthquakes, and relationships between magnitude, depth, and classification accuracy. The use of geospatial visualization, clustering algorithms, and error analysis allowed for a multi-faceted view of seismic events, stressing the importance of combining statistical methods with geophysical understanding.

Key findings include higher measurement uncertainty in automatically classified and low-magnitude events, as well as distinct usage patterns of magnitude types based on depth and intensity. These insights suggest targeted improvements in sensor deployment and classification workflows. Although this study focuses on a brief temporal window, the methodology is scalable and could support long-term monitoring, infrastructure planning, and public safety initiatives as part of a broader seismic risk management strategy that relies on up-to-date earthquake analysis.

# **References**

Allen, R. M., Gasparini, P., Kamigaichi, O., & Böse, M. (2009). The status of earthquake early warning around the world: An introductory overview. *Seismological Research Letters, 80*(5), 682–693.

Federal Emergency Management Agency (FEMA). (2020). *Earthquake risk in the U.S.* FEMA.<https://www.fema.gov>

Freed, A. M. (2005). Earthquake triggering by static, dynamic, and postseismic stress transfer. *Annual Review of Earth and Planetary Sciences, 33*, 335–367.

Stein, S., & Wysession, M. (2009). *An introduction to seismology, earthquakes, and earth structure.* Wiley-Blackwell.

U.S. Geological Survey (USGS). (2023). *Earthquake Catalog.* Retrieved from<https://earthquake.usgs.gov/earthquakes/feed/v1.0/summary/all_month.csv>

Huang, Y., Meng, L., & Ampuero, J. P. (2023). Machine learning in earthquake source studies. *Nature Reviews Earth & Environment*, 4(2), 115–129. https://doi.org/10.1038/s43017-022-00357-2

Mignan, A., & Broccardo, M. (2019). One neuron versus deep learning in aftershock prediction. *Nature Communications*, 10, 1–7. https://doi.org/10.1038/s41467-019-11739-4

Worden, C. B., Gerstenberger, M. C., Rhoades, D. A., & Wald, D. J. (2010). Probabilistic relationships between ground-motion parameters and modified Mercalli intensity in California. *Bulletin of the Seismological Society of America*, 100(6), 3083–3101.

**Appendix**

**Data Source**

* **Dataset**: USGS Earthquake Catalog (all earthquakes, past 30 days)
* **Access Date**: March 2025
* **URL**: https://earthquake.usgs.gov/earthquakes/feed/v1.0/summary/all\_month.csv

**Key Variables Used in Analysis**

* mag: Earthquake magnitude
* depth: Hypocenter depth (km)
* latitude, longitude: Epicenter coordinates
* magType: Method of magnitude estimation (e.g., ml, mb, mww)
* magError, depthError, horizontalError: Measurement uncertainty metrics
* type: Earthquake event classification
* status: Reviewed status (automatic or human-reviewed)

**Tools & Libraries**

* Python: pandas, matplotlib, seaborn, scikit-learn, numpy, urllib
* Jupyter Notebook environment
* DBSCAN clustering for spatial analysis
* Basemap and contextily for global geographic visualizations