# Earthquake Data Analysis: Last 15 Years (2010-2025)

*A Comprehensive Look at Global Seismic Activity and Impact*

***By Suzan Wheed***

***July,2025***

# Introduction & Project Goal

**Project Goal**: To analyze earthquake data from 2010 to 2025 to identify trends, patterns, and key insights related to earthquake occurrences, magnitudes, and their impact.

**Why this analysis?** Understanding seismic activity is crucial for disaster preparedness, risk assessment, and scientific research.

# Data Sources

## USGS Earthquake Hazards Program:

Primary source for earthquake event data: time, place, magnitude, depth, coordinates.

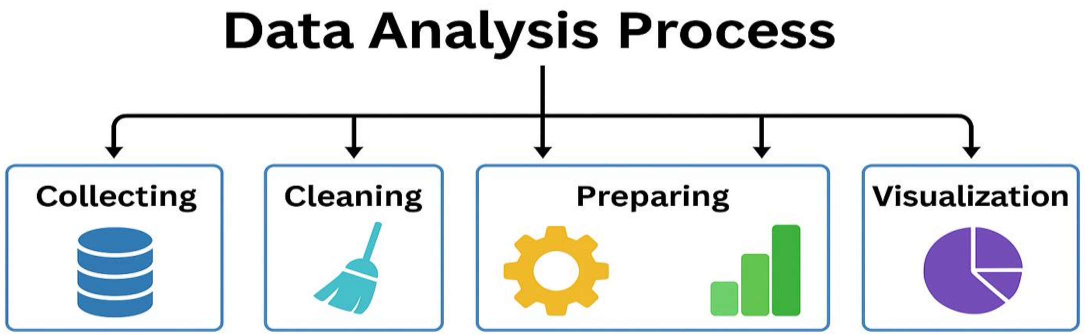
Provides real-time and historical earthquake information.

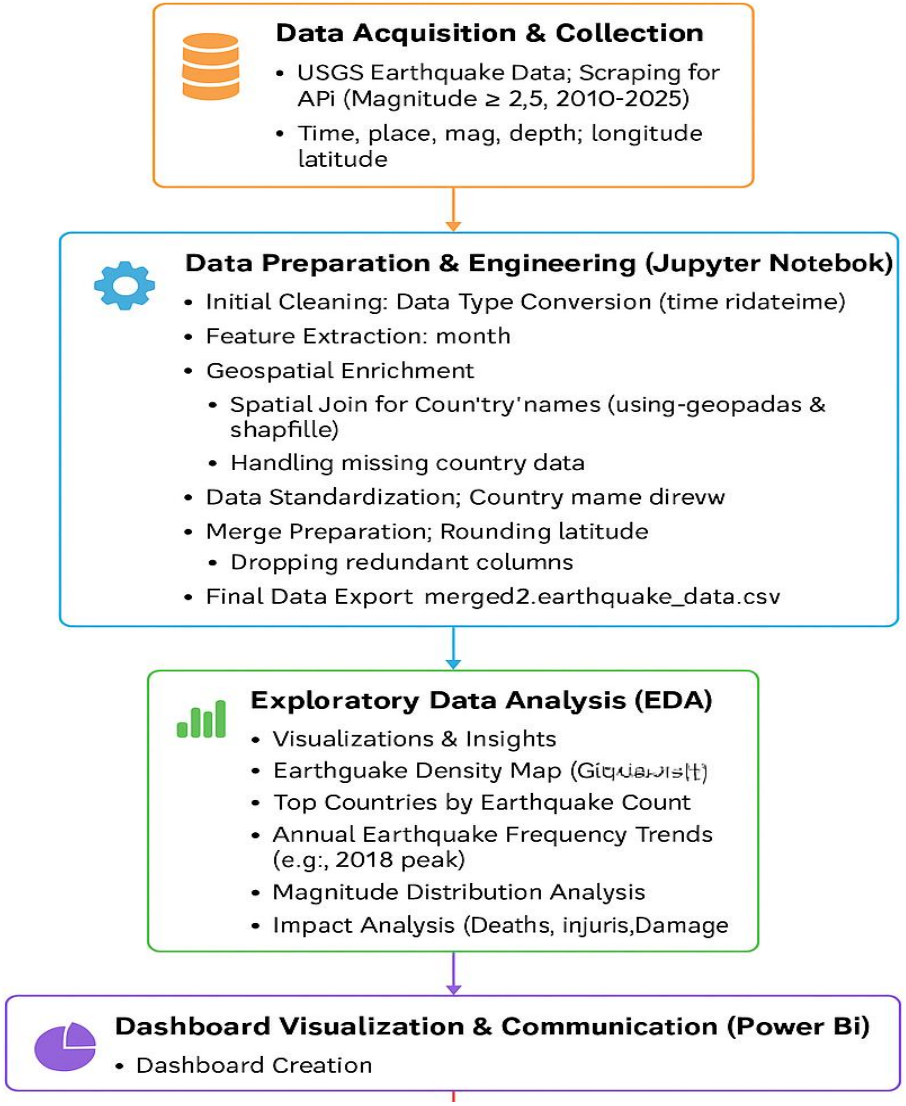
## NOAA National Centers for Environmental Information (NCEI):

Supplements USGS data with additional impact details: deaths, injuries, damage, etc.

Offers a historical database of significant earthquake events.

# Analysis Process



# Tools used in Analysis

- Python  for scrapping earthquake data from USGS API and cleaning these data and preparing it for the analysis
- Python for EDA
- Power Bi for dashboard
- Power Point for presenting data analysis

# Data Collection 1- USGS Data Scraping

**Methodology:** Utilized the USGS API to fetch earthquake data (magnitude ≥2.5).

**Timeframe:** Monthly data collection from January 2010 to the current year (2025).

**Key Data Points Collected:** time, place, mag, depth, longitude, latitude.

**Efficiency:** Implemented a check to avoid re-downloading if CSV files already existed, optimizing the process.

# Data Collection 2- NOAA Data Download

**Methodology:** Obtained a TSV/Excel file directly from the NOAA website.

**URL:**
https://www.ngdc.noaa.gov/hazel/view/hazards/earthquake/search

**Content:** Contains crucial additional earthquake impact details, allowing for a more comprehensive analysis of consequences.

# Data Preprocessing & Cleaning - Initial Steps

## Initial Data Overview (USGS Data):

Thorough checks of data types and non-null counts to ensure data integrity.

Conversion of time column to proper datetime objects for temporal analysis.

## Feature Engineering:

Extraction of month names and year from the time column to facilitate time-based grouping and analysis.

# Data Preprocessing & Cleaning - Geospatial & Standardization

## Geospatial Join for Country Names:

- Used geopandas for a spatial join operation.
- Mapped longitude and latitude coordinates to country names using a shapefile (ne_10m_admin_0_countries.shp).
- Handled missing country values by extracting names from the place column.

## Country Name Standardization:

- Implemented a two-step mapping process using external Excel/CSV files (Country_Standardization6_Map.xlsx, region_country_mapping2.csv).
- Resolved regional names and variations to consistent country names, with careful review of mismatches.

# Data Merging

### Integrating USGS and NOAA Data:

Prepared common columns (lat, lon, year) by rounding coordinates and ensuring consistent year data types.

Performed a left join to merge detailed NOAA impact data (data DataFrame) with comprehensive USGS earthquake data (quakes DataFrame).

**Join Keys:** latitude, longitude, and year.

### Post-Merge Cleanup:

Dropped redundant or unnecessary columns (e.g., duplicate coordinates, raw date/time components) to streamline the dataset.

# Data Merging - Final Touches & Export

## Label Encoding for Descriptions:

- Converted numeric description codes (e.g., 1, 2, 3, 4) in impact columns (Death Description, Damage Description) to more descriptive categorical labels ('few', 'some', 'many', 'very many') based on source documentation.

## Final Data Export:

- Saved the cleaned, merged, and enhanced DataFrame to a new CSV file (merged2_earthquake_data.csv).
- This ensures the processed data is readily available for further analysis or sharing.

# Exploratory Data Analysis (EDA) - Overview

**Purpose:** To uncover patterns, detect anomalies, test hypotheses, and extract valuable insights from the processed data.

**Key Areas Explored:**

•Geographical distribution of earthquakes.

•Temporal trends in earthquake frequency.

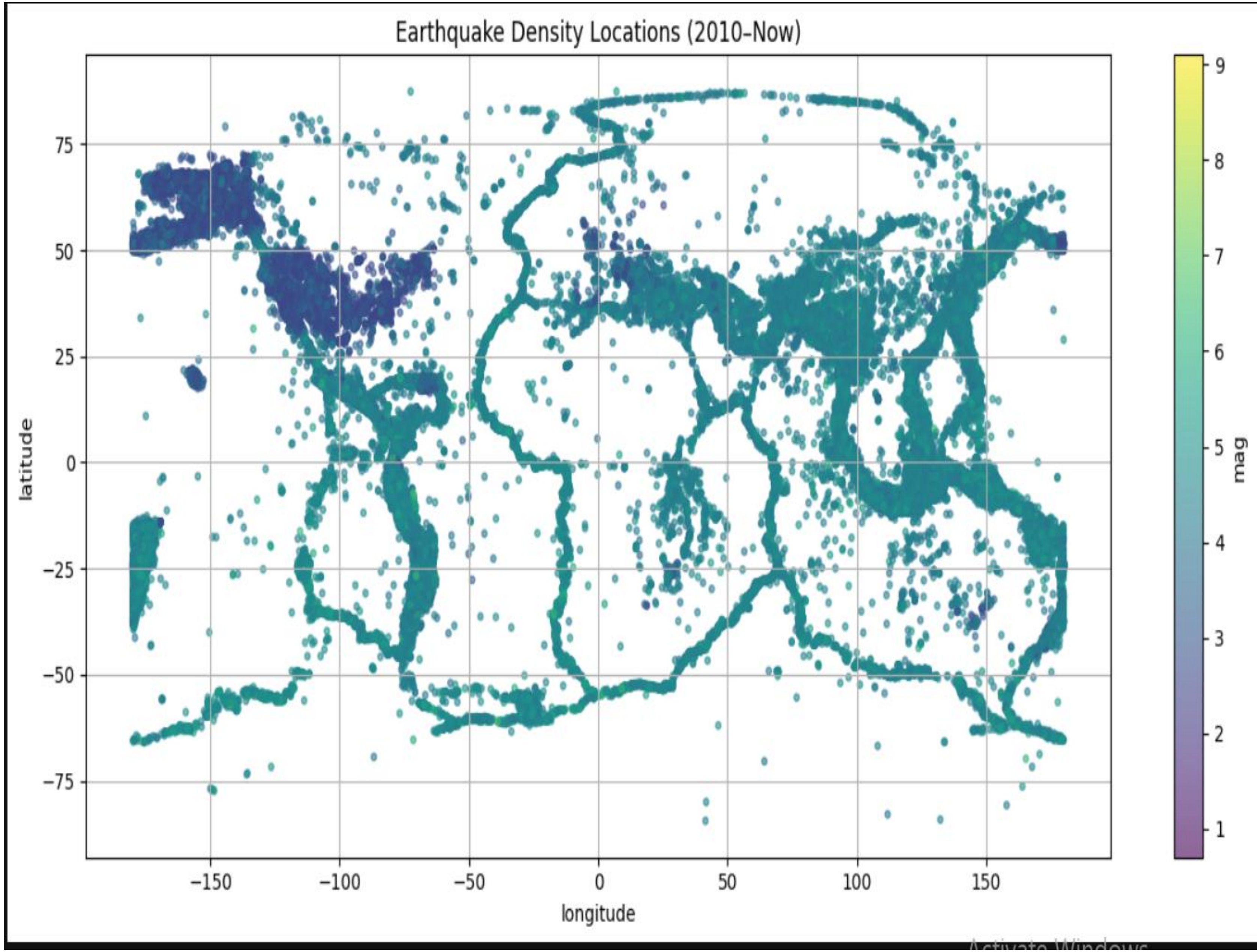
- •Magnitude characteristics.
- •Impact analysis (deaths, injuries, damage).

# EDA - Earthquake Density Visualization

•**Visualization:** Scatter plot using matplotlib.pyplot.

•**Focus:** Global distribution and density of earthquakes based on longitude and latitude.

•**Insight:** The plot clearly shows earthquake epicenters concentrated along tectonic plate boundaries, particularly around the Pacific Ring of Fire, indicating areas of high seismic activity.

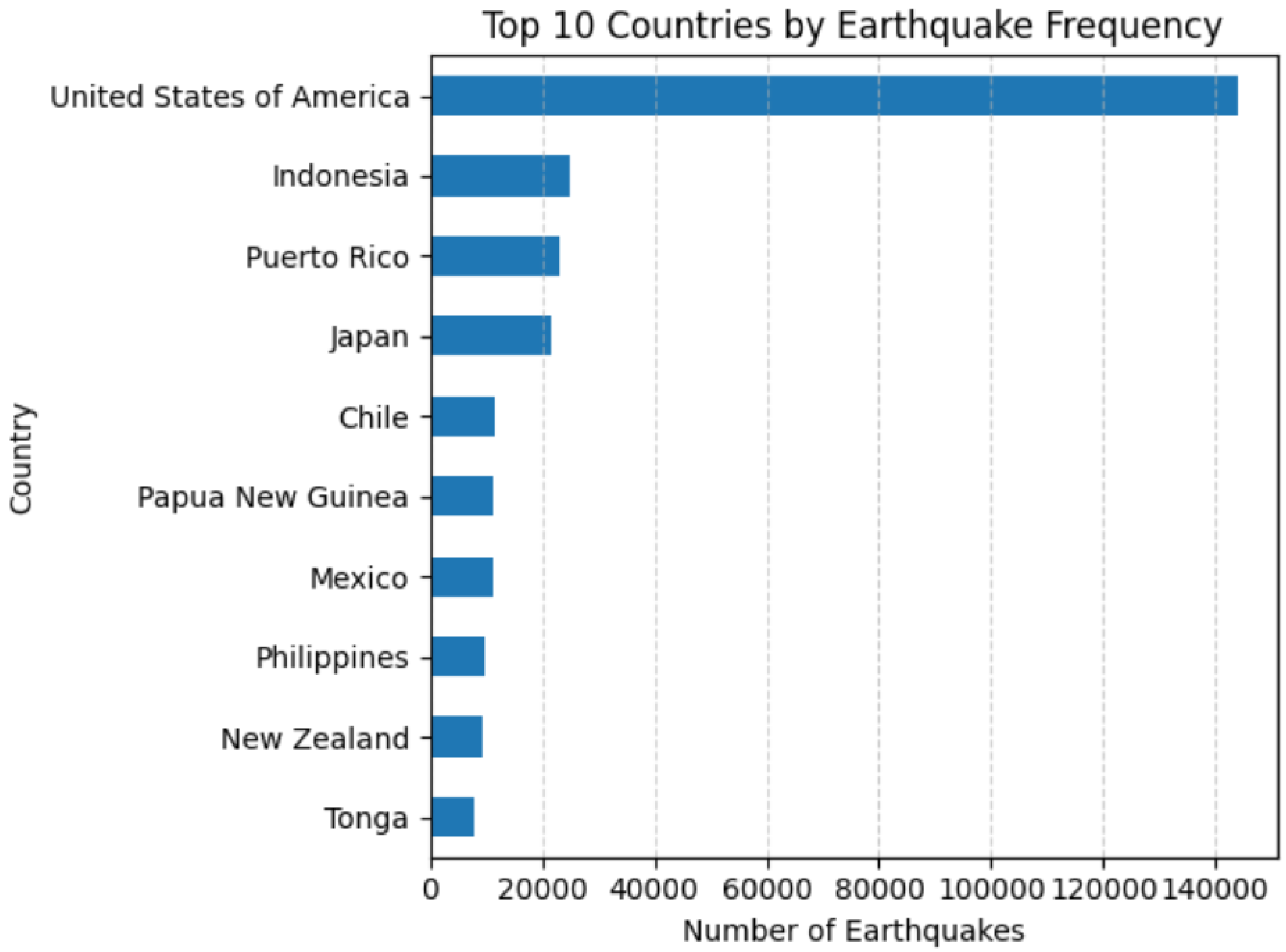


# EDA - Top 10 Countries by Earthquake Count

•**Visualization:** Scatter plot using matplotlib.pyplot.

•**Focus:** Global distribution and density of earthquakes based on longitude and latitude.

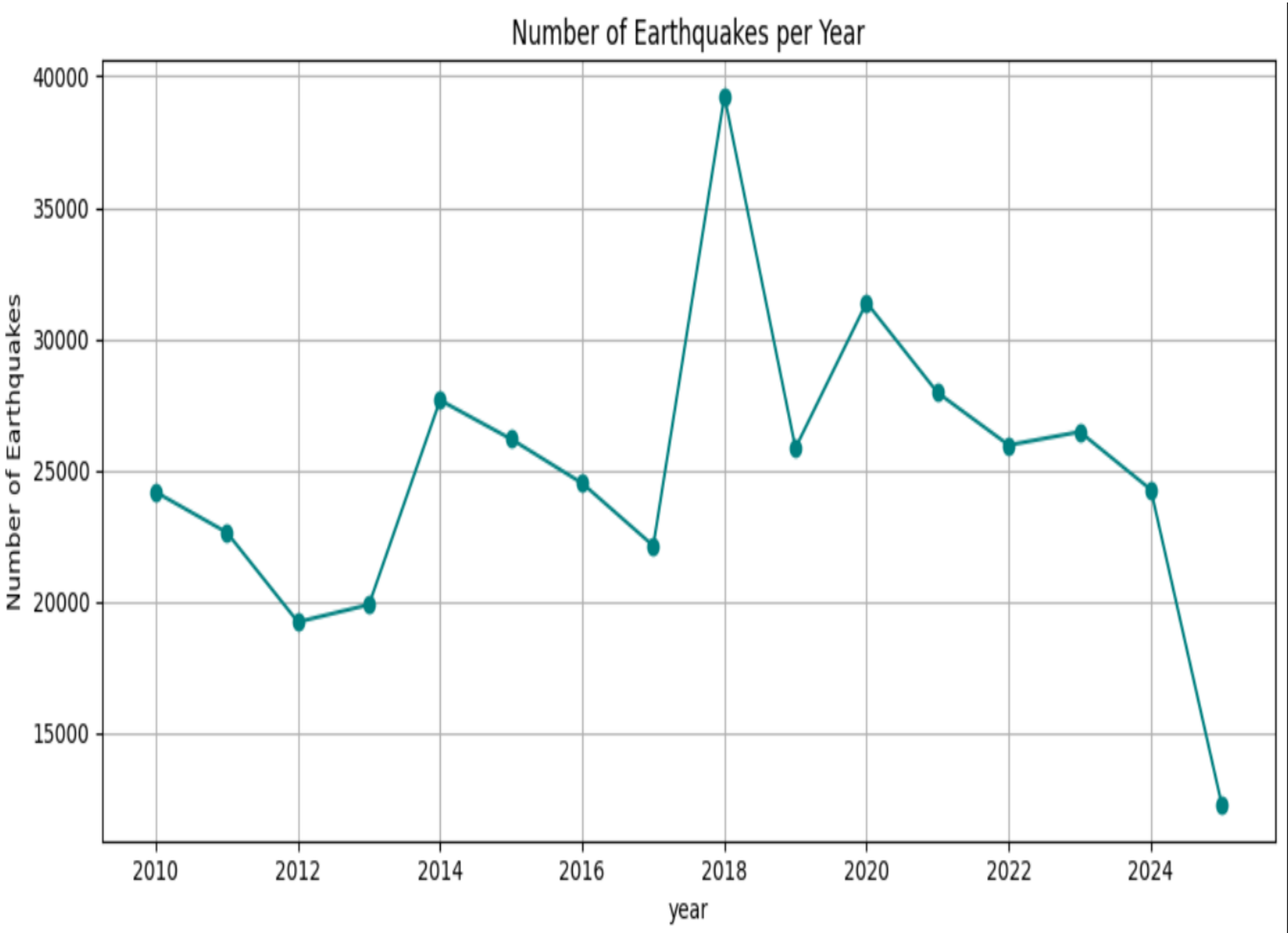
•**Insight:** The plot clearly shows earthquake epicenters concentrated along tectonic plate boundaries, particularly around the Pacific Ring of Fire, indicating areas of high seismic activity.



# Earthquakes Over Time (Annual Trends)

•**Analysis:** Grouped data by year to observe the annual frequency of earthquakes.

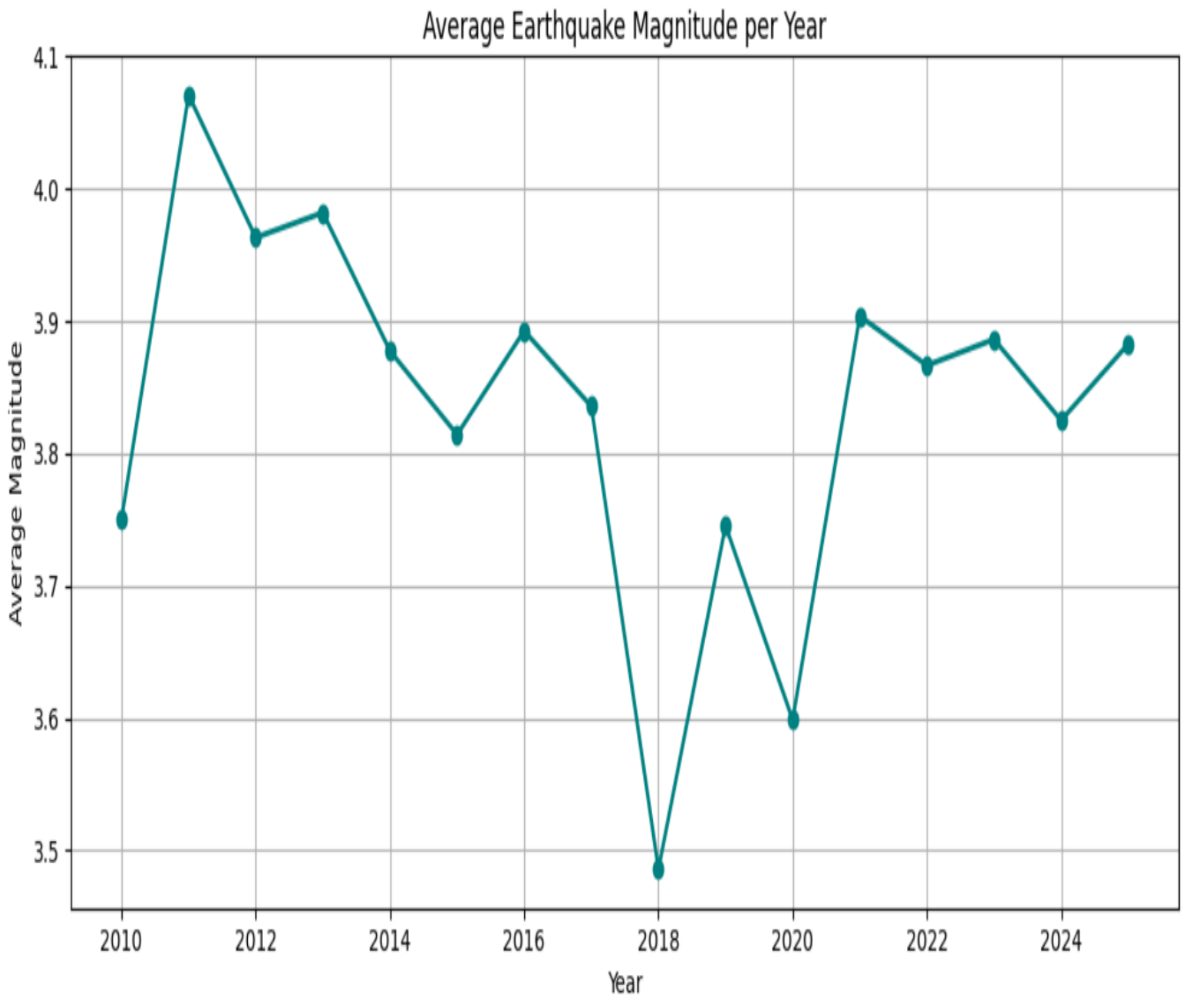
•**Insight:** The year 2018 recorded the highest number of earthquake events (39244 events), indicating a peak in seismic activity during that period.



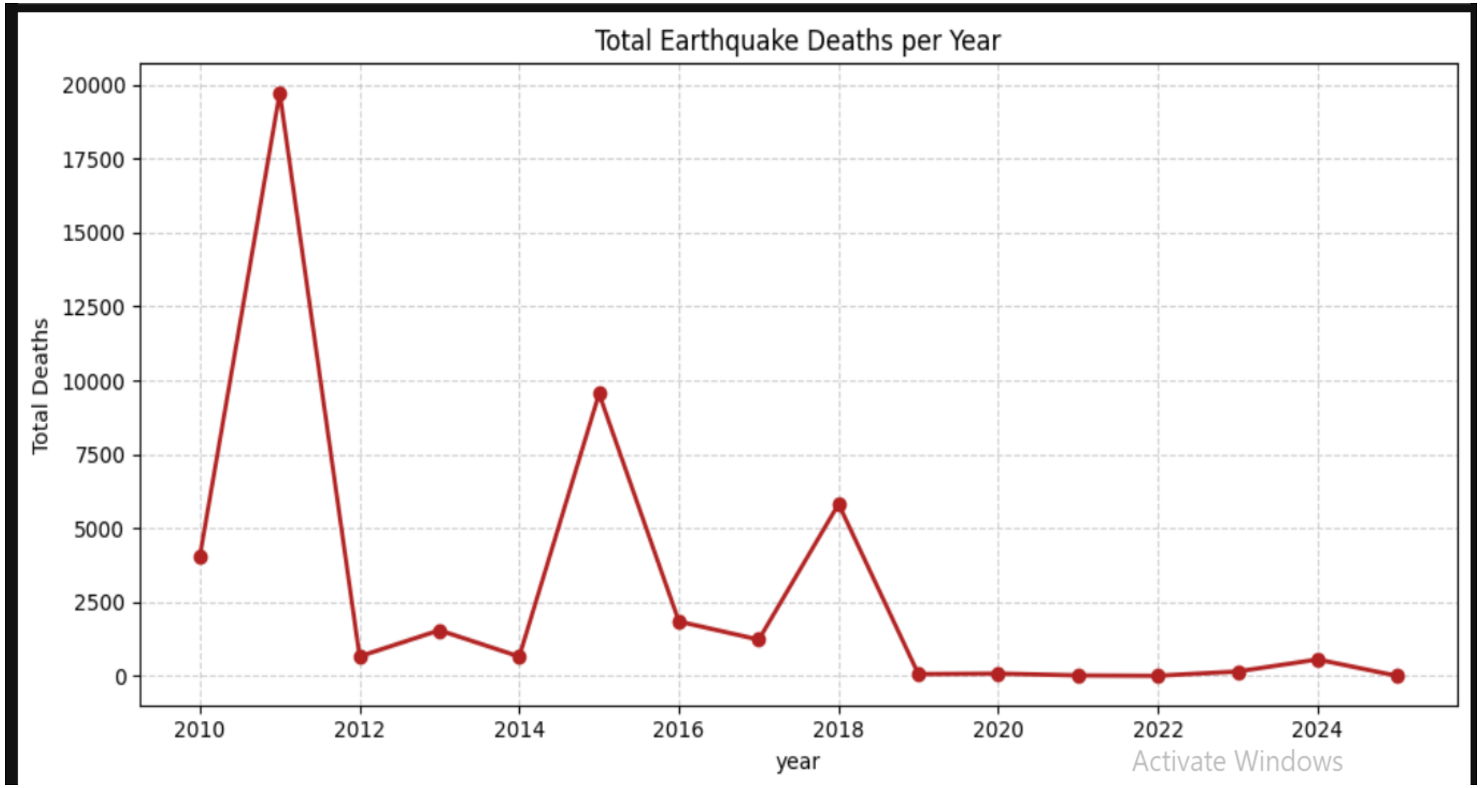
# EDA - Magnitude Distribution

**Analysis:** Investigated the distribution of earthquake magnitudes across the dataset.

**Insight:** The majority of earthquakes recorded are of lower magnitudes (e.g., 2.5-4.5), with a significantly smaller number of high-magnitude events. This aligns with the inverse relationship between earthquake frequency and magnitude.



# EDA - Impact Analysis: Total Deaths




- **Analysis:** Identified the earthquake event with the highest number of total deaths.
- **Insight:** The Great East Japan Earthquake in 2011 (magnitude 9.0) resulted in significant loss of life and widespread destruction, highlighting the devastating human impact of high-magnitude events, particularly when accompanied by tsunamis.

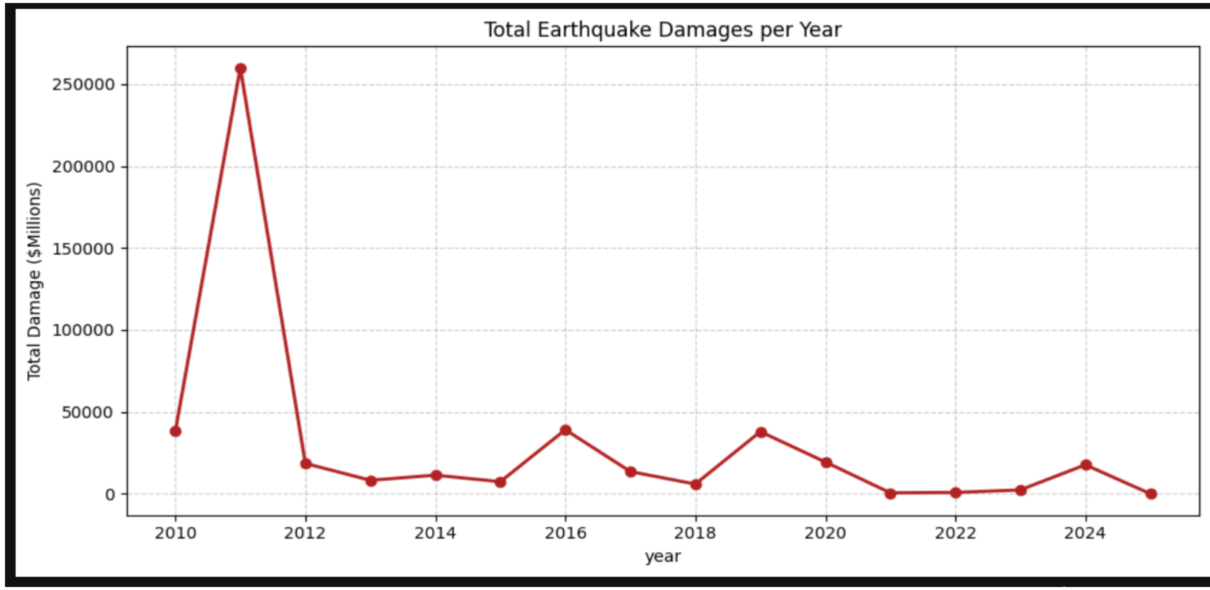
# EDA - Impact Analysis: Injuries & Damage

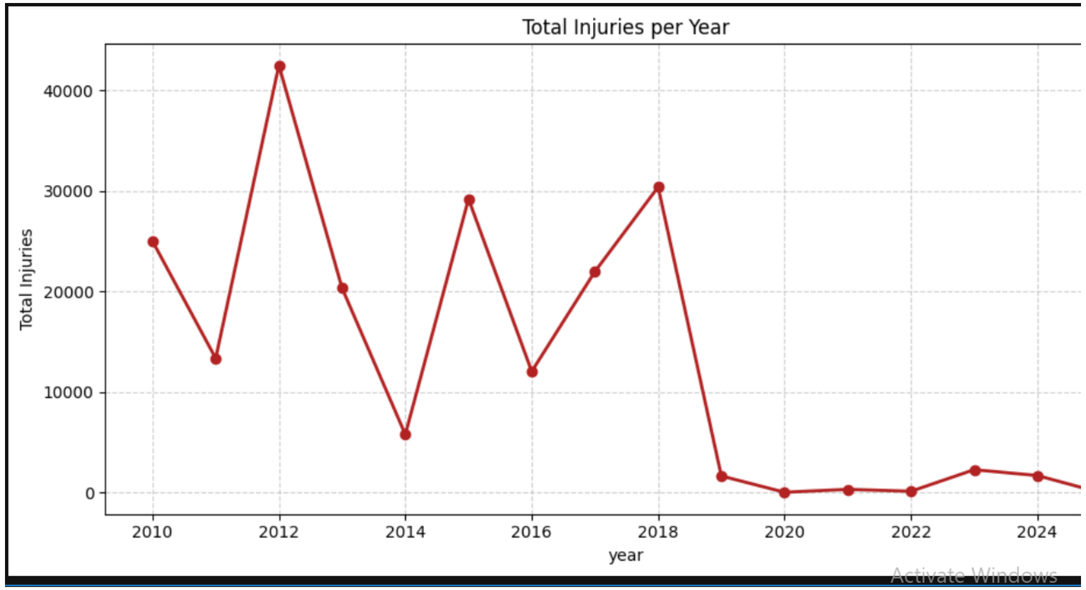
**Total Injuries per Year:**

**Insight:** This chart shows the annual trend of total injuries caused by earthquakes. We can observe fluctuations, with some years experiencing significantly higher injury counts.

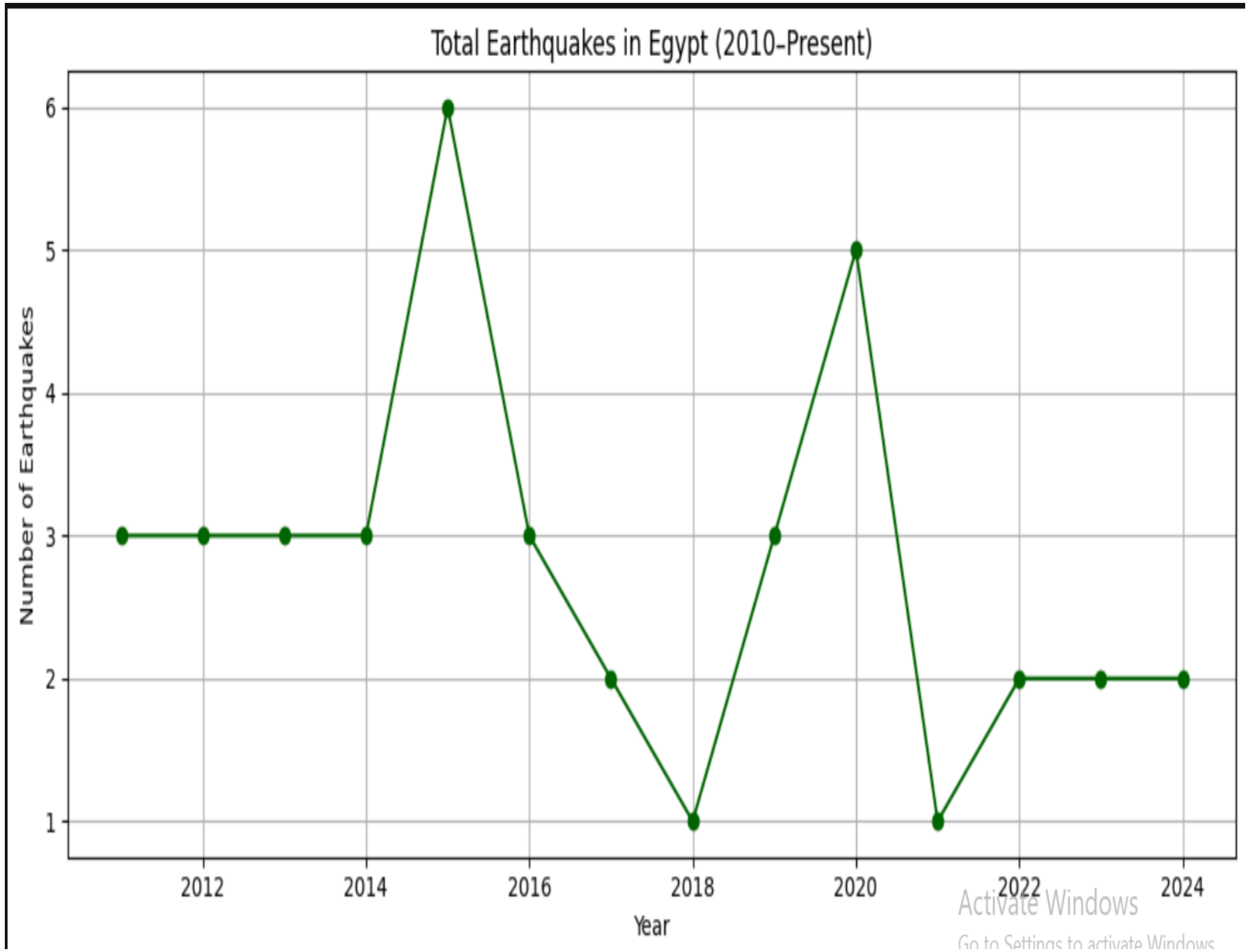
**Total Earthquake Damages per Year:**

**Insight:** This chart illustrates the annual trend of total monetary damages (in millions of dollars) caused by earthquakes. It highlights years with particularly costly events.



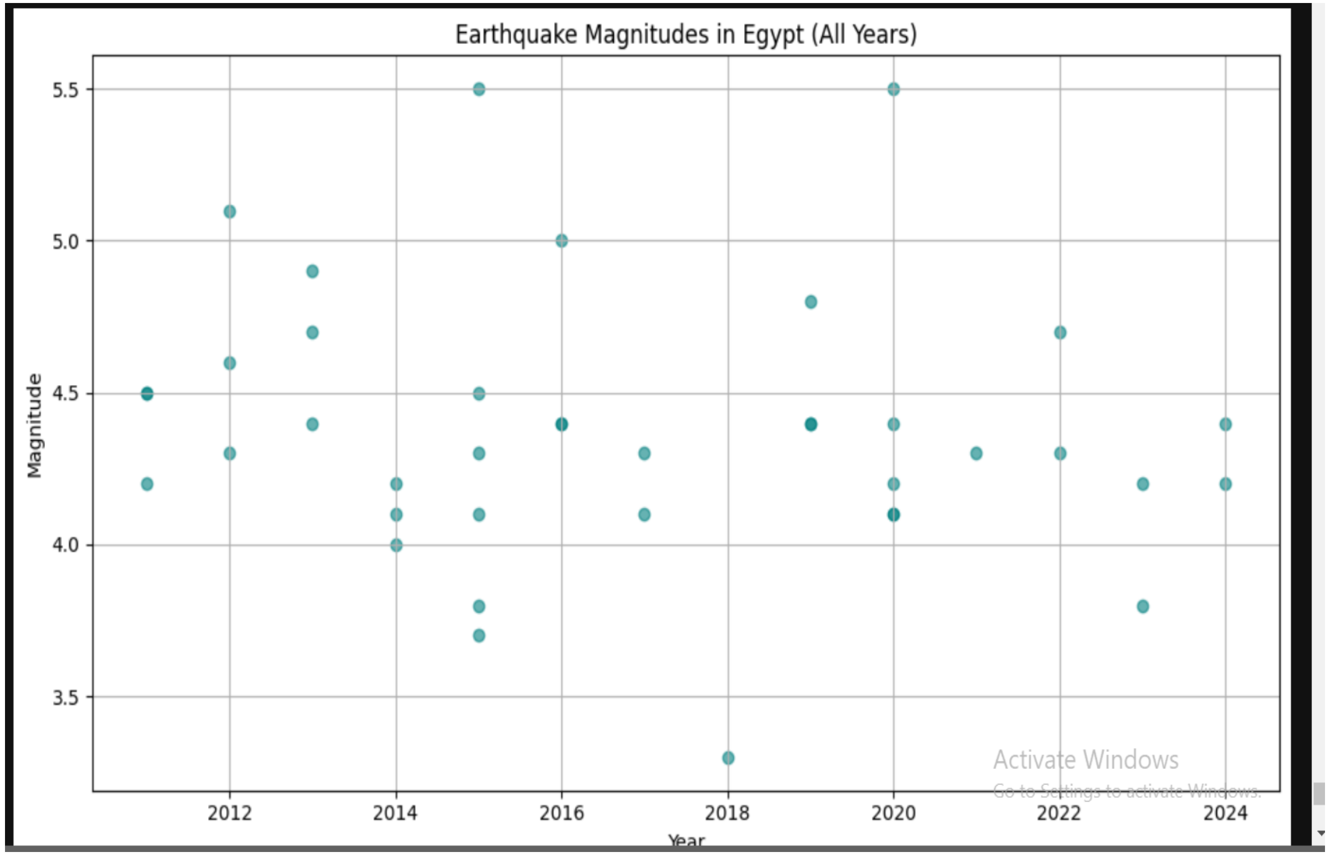
# EDA - Impact Analysis: Total Earthquakes in Egypt (2010-Present)





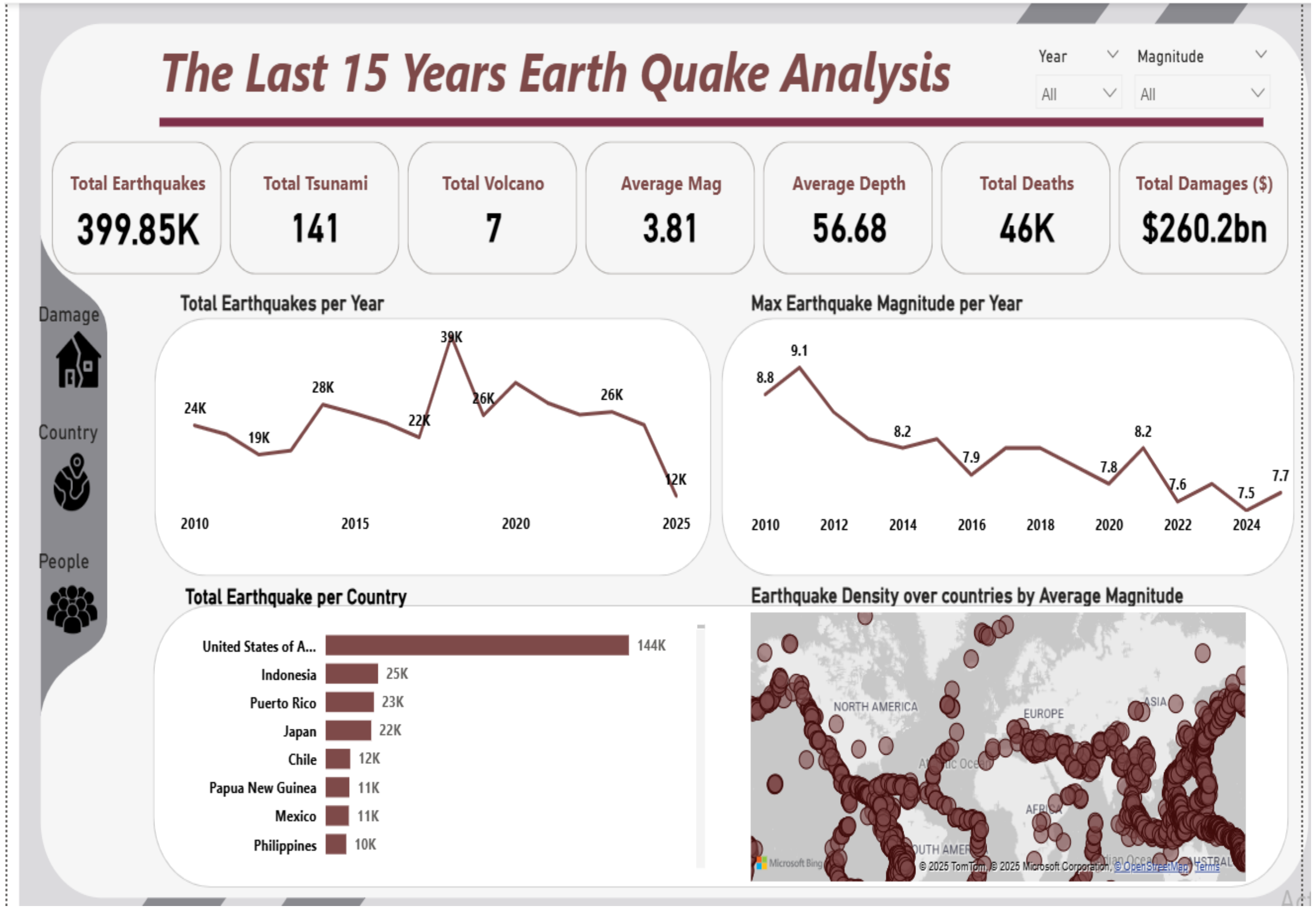
- **Insight:** This chart illustrates the annual frequency of earthquakes specifically within Egypt from 2010 to the present.
- **Observation:** We can observe fluctuations in the number of recorded seismic events in Egypt over the years, with peaks and troughs indicating periods of higher or lower activity.
- **Total Earthquake** 39

# EDA - Impact Analysis: Earthquake Magnitudes in Egypt (All Years)




- **Insight:** This scatter plot visualizes the distribution of earthquake magnitudes in Egypt across all analyzed years.
- **Observation:** It allows us to see the range of magnitudes recorded and how they are distributed over time, helping to identify any patterns or significant events.

# Dashboard Overview: Visualizing Key Insights



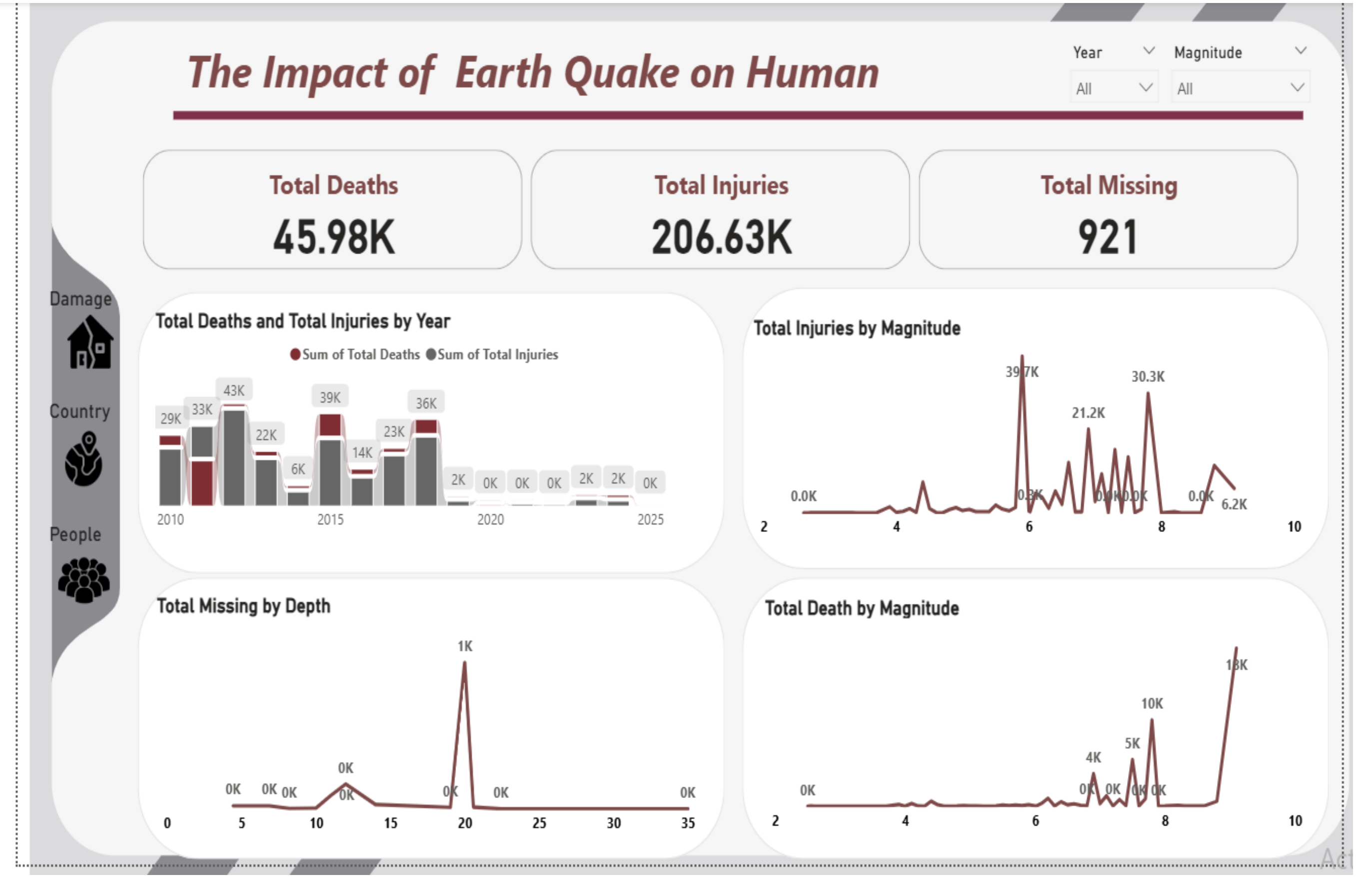

**Purpose:** This dashboard provides a dynamic and comprehensive visual summary of our earthquake data analysis.

**Key Visualizations:**

**Magnitude Trends:** Observe the frequency of different magnitude earthquakes over time.

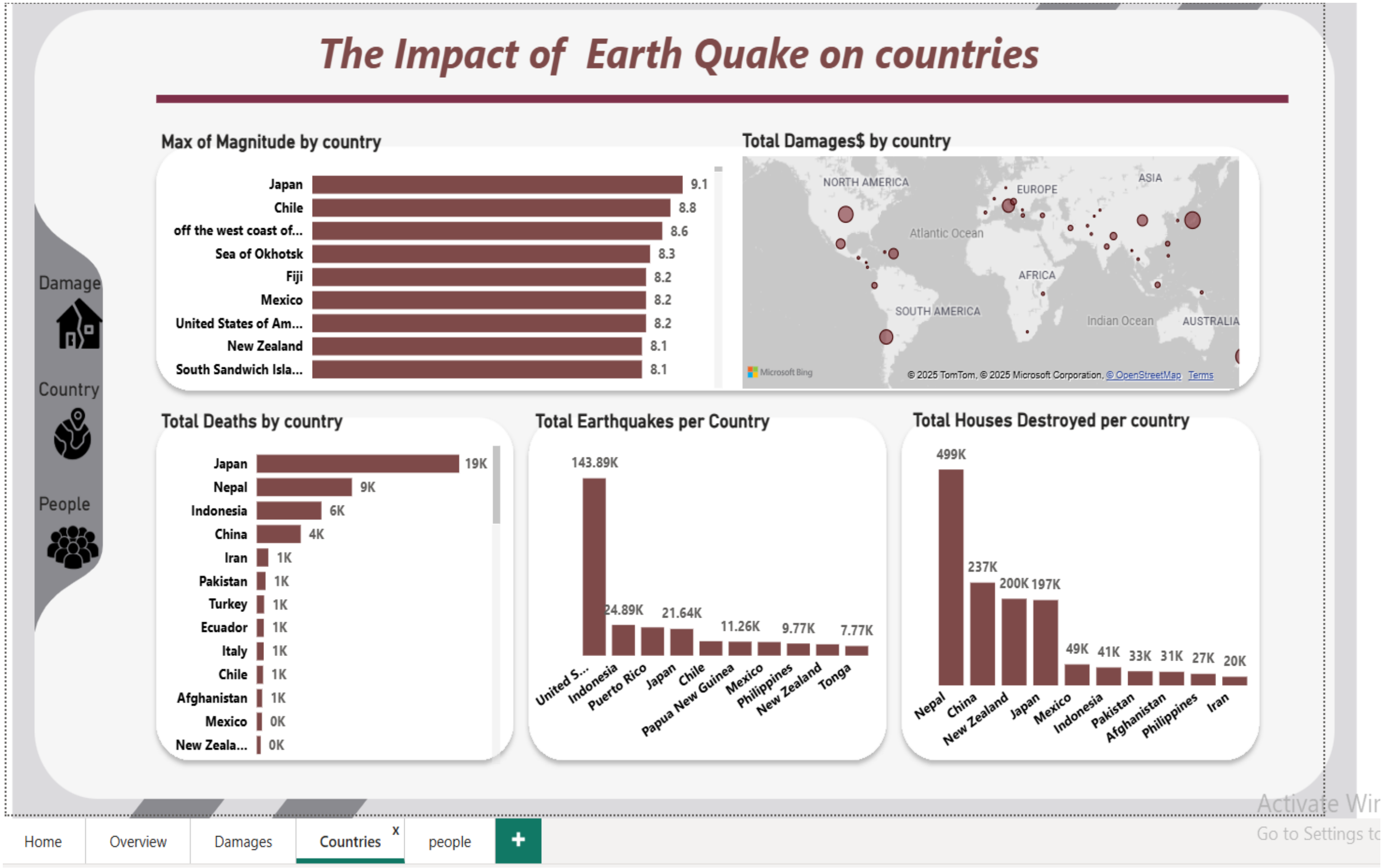
**Value:** Offers a quick, intuitive way to grasp the main findings and explore the data at a glance.

# Dashboard Overview: Visualizing Key Insights



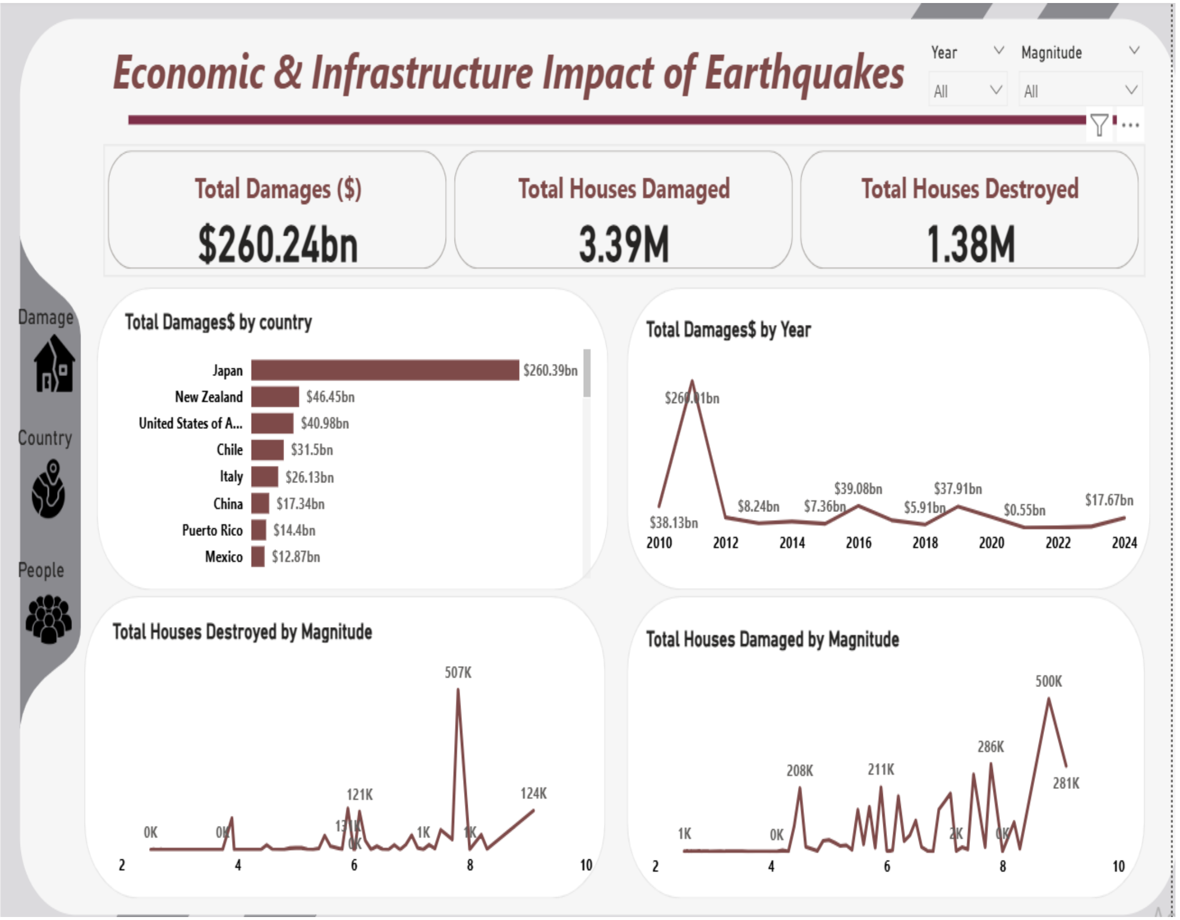

**Impact Summary:** Understand the scale of deaths, injuries, and damage associated with seismic events.

# Dashboard Overview: Visualizing Key Insights




**Geographical Distribution:** Quickly identify earthquake hotspots around the globe.

# Dashboard Overview: Visualizing Key Insights




**Significant Financial and Structural Toll:** The dashboard highlights the massive scale of damage, with total damages reaching over $260 billion, and millions of houses either damaged or destroyed.

**Japan as a Major Impact Zone:** It clearly identifies Japan as the country with the highest economic impact from earthquakes, incurring the vast majority of the total damages.

**High Magnitude = High Impact:** The visuals demonstrate a direct correlation between higher earthquake magnitudes (especially around magnitude 8) and a significant increase in the destruction and damage to houses.

# Key Finding: Geographical Concentration

**Insight:** That earthquakes are not randomly distributed across the globe but are heavily concentrated along well-known tectonic plate boundaries, such as the "Pacific Ring of Fire" (which includes Japan, Indonesia, and the west coasts of North and South America), the Alpine-Himalayan belt, and mid-ocean ridges. This highlights the inherent geological instability of these regions.

**Observation:** There are clear concentrations in seismically active

**Implication:** This highlights regions with inherently higher seismic risk due to tectonic plate interactions.

zones, such as the Pacific Ring of Fire.

# Key Finding: Magnitude vs. Impact

**Insight:** While many earthquakes are low-magnitude, their impact can vary significantly.

**Observation:** High-magnitude events, especially in densely populated or vulnerable areas, can lead to catastrophic human and structural damage.

**Implication:** Preparedness and building codes are crucial in high-risk zones, regardless of overall frequency.

# Key Finding: Temporal Fluctuations

**Insight:** Earthquake frequency shows variations over time.

**Observation:** Our analysis reveals fluctuations in earthquake frequency over the last 15 years, with 2018 being a particularly active year (39,244 events).

**Implication:** While long-term trends exist, short-term variations in seismic activity are also observable.

# Thank you