

# Analisi dei Terremoti e delle faglie nel mondo (1970-2025)

## Introduzione

In questo lavoro di gruppo, abbiamo deciso di analizzare un dataset preso USGS.gov, che contiene varie informazioni sui terremoti avvenuti in tutti il mondo dal 1970 al 2025. Il dataset è stato selezionato in modo da avere solo terremoti di magnitudo superiore a 5.0 della scala Richter. Ci siamo limitati a questa magnitudo in quanto il sito consente di scaricare fino a 20.000 eventi la volta, quindi per coprire tutta la serie storica è stato necessario unificare tutte le richieste creando un unico dataset con 87798. Inoltre, dal sito globalquakemodel.org il quale contiene la mappa di tutte le faglie attive nel mondo. In questo modo abbiamo potuto confrontare le posizioni dei terremoti rispetto alle faglie. I dati sono stati analizzati tenendo conto sia della loro posizione geografica e di altre caratteristiche presenti nel dataset e sia della loro frequenza nel corso del tempo. Queste analisi sono state effettuate sia attraverso l'utilizzo di mappe, sia attraverso l'utilizzo di grafici al fine di descrivere al meglio tutte le variabili quantitative e qualitative presenti.

## Dataset Faglie

Il dataset contiene la mappa di tutte le faglie attive presenti nel mondo, ed è possibile importarlo tramite il seguente codice:

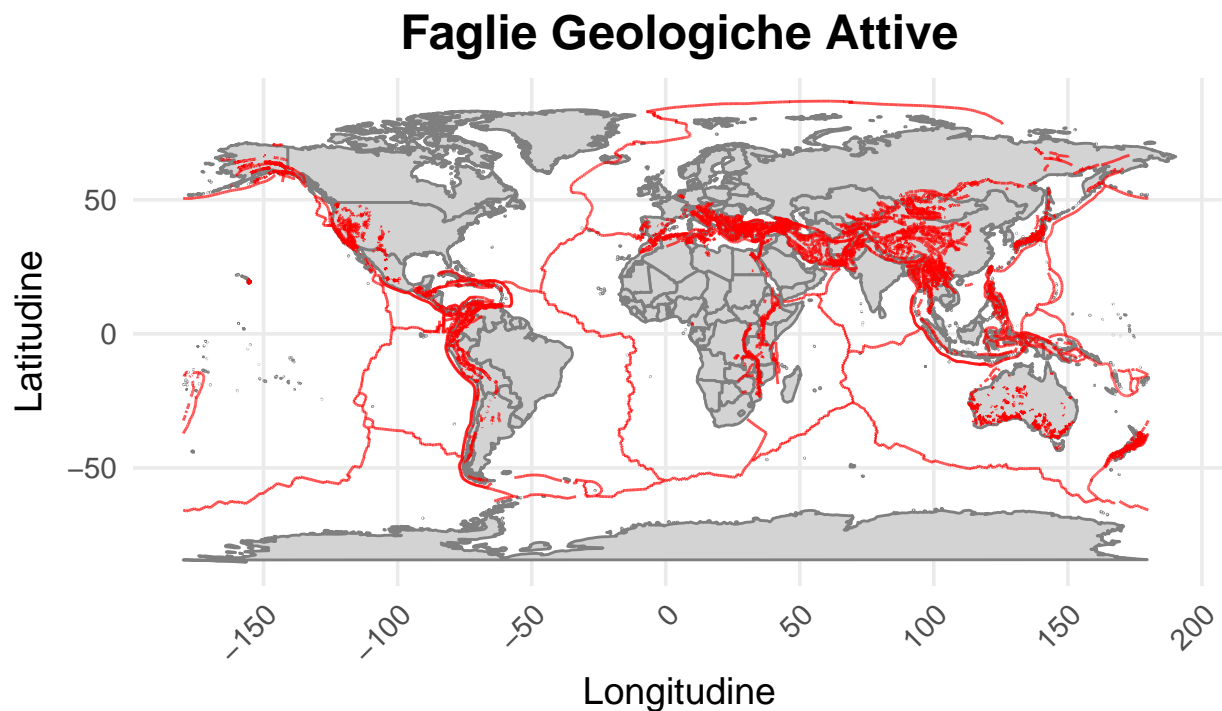
```
library(readr)
library(sf)
library(ggplot2)
library(dplyr)
library(maps)
library(tidyr)
library(RColorBrewer)
library(stringr)
library(rnaturalearth)
library(rnaturalearthdata)
library(ggrepel)
library(lubridate)
```

Per semplicità si è deciso di presentare tutte le librerie utilizzate nel lavoro nel precedente blocco di codice.

```
faglie <- st_read("gem_active_faults.shp")

## Reading layer 'gem_active_faults' from data source
##   'C:\Users\Roberto\Desktop\Progetto in R. L-D-R\gem_active_faults.shp'
##   using driver 'ESRI Shapefile'
## Simple feature collection with 16195 features and 26 fields
## Geometry type: LINESTRING
## Dimension:      XY
## Bounding box:   xmin: -180 ymin: -66.163 xmax: 180 ymax: 86.805
## CRS:           NA
```

```
ggplot() +
  # Aggiungi la mappa del mondo
  borders("world", colour = "gray50", fill = "lightgray") +
  # Aggiungi le faglie
  geom_sf(data = faglie, color = "red", size = 1, alpha = 0.7) +
  labs(title = "Faglie Geologiche Attive", x = "Longitudine", y = "Latitudine",
        color = "Magnitudo") +
  theme_minimal(base_size = 14) +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    plot.title = element_text(hjust = 0.5, size = 18, face = "bold"),
    legend.position = "right")
```



Nel grafico, viene mostrata la mappa nel mondo in grigio, e in rosso, le faglie geologiche attive. Come si può vedere le faglie si distribuiscono principalmente lungo i confini delle placche tettoniche con una forte concentrazione nella parte sud-ovest dell'Asia a formare la catena dell'Himalaya. La presenza di faglie geologiche è di particolare interesse in quanto nei pressi di queste zone si osserva una maggiore probabilità che vi siano terremoti. Questo fenomeno verrà visualizzato in seguito, attraverso il dataset "Earthquake" in cui è presente la lista dei terremoti registrati dal 1970-2025.

### Dataset Earthquake:

I dati sono stati presi dal sito USGS.gov filtrati con una magnitudo superiore a 5.0. Essendo la frequenza dei terremoti correlata alla magnitudo il sito consentiva solo di scaricare 20000 eventi per volta, producendo troppe richieste di dati all'ente fornitore. Inoltre, il dataset sull'intero fenomeno era piuttosto frammentato si è deciso quindi di unire le informazioni distribuite su diversi anni in un unico blocco per avere una visione globale del fenomeno.

In R, i dati sono stati importati tramite il seguente codice:

```
terremoti <- read.csv("Earthquake_1970-2025.csv")
```

## Analisi Premilinare.

Il nostro dataset è composto da 22 variabili di varia natura per un totale di quasi 88000 osservazioni. Per semplicità, mostriamo un breve estratto delle variabili più significative usate nell'analisi del fenomeno.

```
library(pander)
pander(head(terremoti %>% select(latitude, longitude, depth, mag, nst, type, time)))
```

latitude	longitude	depth	mag	nst	type	time
-20.18	-70.62	35	5	37	earthquake	2024-12-31T23:13:20.048Z
-6.668	150.6	10	5.1	61	earthquake	2024-12-31T20:09:41.043Z
-4.05	151.6	14.52	5	59	earthquake	2024-12-31T17:09:39.374Z
-17.66	168.2	66.61	5.1	121	earthquake	2024-12-30T05:49:02.808Z
-29.93	-72	10	5.5	127	earthquake	2024-12-30T05:41:06.678Z
-29.92	-72.06	10	5.5	178	earthquake	2024-12-30T05:40:49.261Z

Nella seguente tabella, le variabili indicano:

- **latitudine e longitudine:** rappresentano le coordinate geografiche di dove è avvenuto il terremoto.
- **depth:** La profondità, misurata in km rispetto alla superficie terrestre.
- **mag:** la magnitudo del terremoto in scala Richter.
- **nst:** numero di stazioni che hanno registrato il terremoto.
- **type:** l'origine del terremoto inteso come causa scatenante. (Terremoto terrestre, esplosione nucleare, collasso di miniere, frane ed esplosioni.)
- **time:** data e ora della rilevazione del fenomeno.