Visualization of global earthquakes from 1965 to 2016 based on R

(Dai Yanhui 12031280)

Background

Earthquake, also known as ground motion or ground vibration, is a natural phenomenon in which seismic waves are generated during the rapid release of energy from the earth's crust. The plates and plates on the earth squeeze and collide with each other, causing dislocations and ruptures on the edges and inside of the plates, which is the main cause of earthquakes. Earthquakes often cause serious casualties, can cause fires, floods, toxic gas leakage, the spread of bacteria and radioactive materials, and may also cause secondary disasters such as tsunamis, landslides, collapses, and ground fissures. Although earthquakes are unpredictable, we can visually analyze the historical data of earthquakes to find the laws and provide a basis for scientific prevention of earthquakes and minimize the loss of life and property caused by earthquakes.

Data and preprocessing

Data: Significant Earthquakes, 1965-2016

This dataset includes a record of the date, time, location, depth, magnitude, and source of every earthquake with a reported magnitude 5.5 or higher since 1965 (please refer to Fig. 1). The National Earthquake Information Center (NEIC) determines the location and size of all significant earthquakes that occur worldwide and disseminates this information immediately to national and international agencies, scientists, critical facilities, and the general public. The NEIC compiles and provides to scientists and to the public an extensive seismic database that serves as a foundation for scientific research through the operation of modern digital national and global seismograph networks and cooperative international agreements. The NEIC is the national data center and archive for earthquake information.

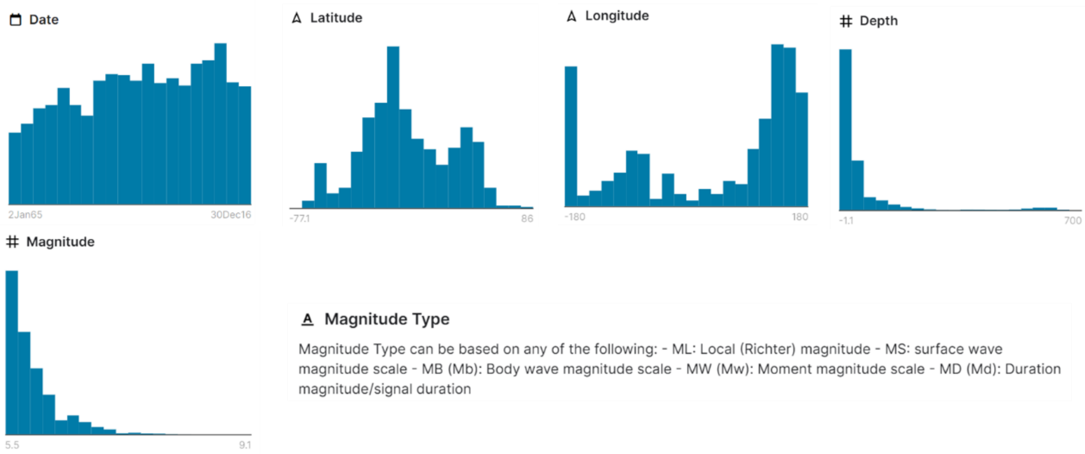

Fig 1. Earthquakes with magnitude 5.5 or higher

Data preprocessing

The first thing we need to do before data analysis is to clean the data and remove the missing data. See the appendix at the end of the text for the specific code (the codes of the other figures are also in this appendix).

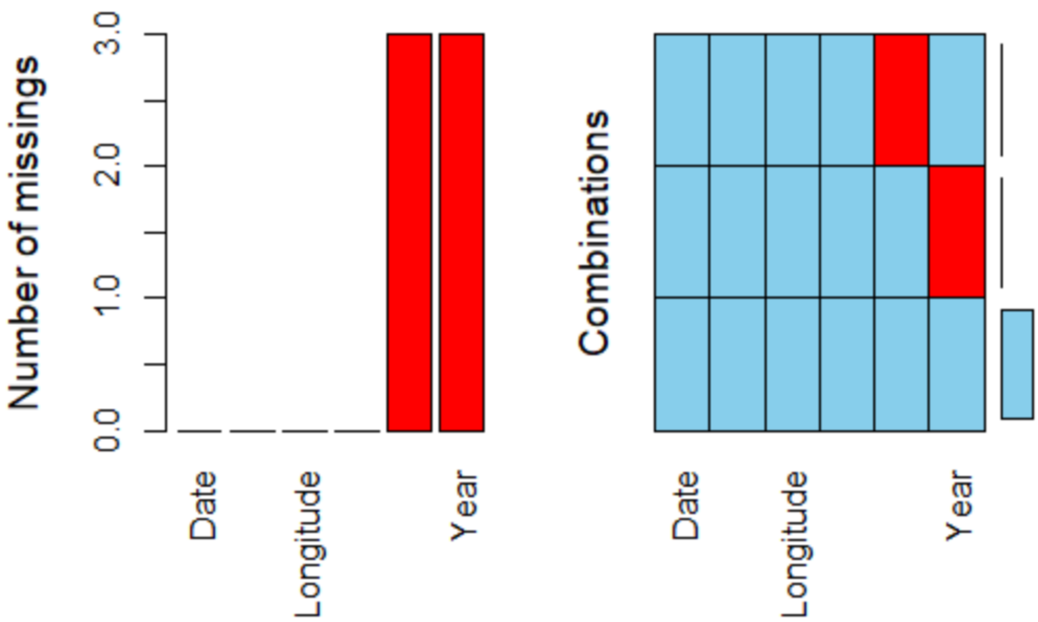

Fig 2. Eliminate missing data

Data analysis

Long-term trend of the total number of global earthquakes from 1965 to 2016

First, we conducted a time series analysis on the data, and obtained the long-term trend of the total number of global earthquakes from 1965 to 2016, as shown in Fig. 3. The occurrence of earthquakes is related to the degree of activity of the earth's plates, and the degree of activity of the earth's plates may change periodically. Therefore, through time-series analysis of the number of earthquakes in the world, the changing law of the number of earthquakes can be discovered, which can provide guidance for scientific earthquake prevention. It can be seen that from 1965 to 2016, the total number of global earthquakes showed an increasing trend over time (a spiral increase).

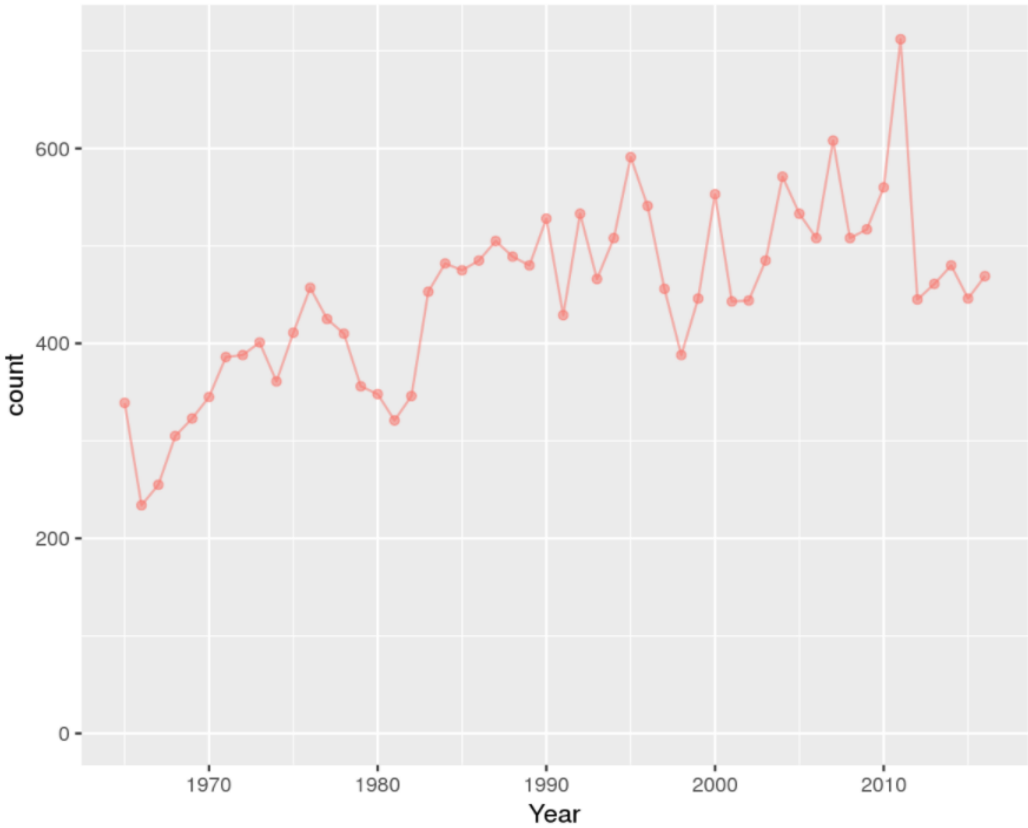

Fig 3. Long-term trend of the total number of global earthquakes from 1965 to 2016

Intuitively, the global trend of the total number of earthquakes has four rises and four falls. In order to clarify the specific details of the upward spiral trend, we have drawn an interactive bar histogram (Fig. 4).

In the interactive histogram, you can use the mouse to frame selection in the graph, and zoom in on the data of the time period of interest (the partial enlarged view in Figure 4 is an example of the first ascending stage). Through the partial zoom operation on the interactive histogram, we found the following rules:

(1) From 1965 to 1966, from 1976 to 1981, from 1995 to 1998, and from 2011 to 2012, the total number of global earthquakes showed a decreasing trend;

(2) From 1966 to 1976, from 1981 to 1995, from 1998 to 2011, and from 2012 to 2016, the total number of global earthquakes showed an increasing trend.

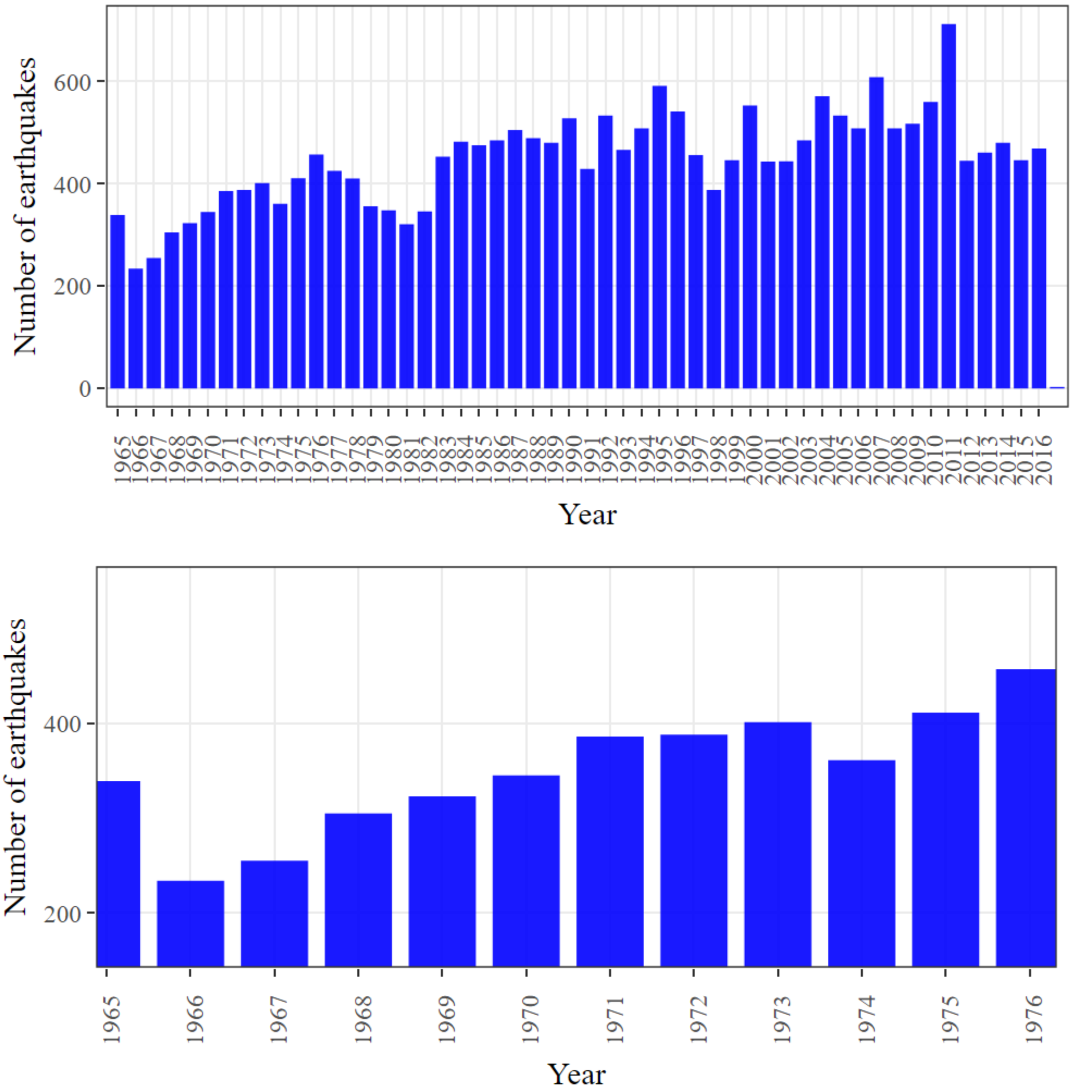

Fig 4. Interactive histogram of the total number of global earthquakes from 1965 to 2016

Earthquake level distribution

When analyzing the law of earthquake occurrence, in addition to the number of earthquakes, the distribution of magnitude is also a very important data. Therefore, we plot the density distribution of earthquake amplitude (see Fig. 5).

It can be seen from Figure 5 that most of the earthquakes have magnitudes between 5.5

and 6.5.

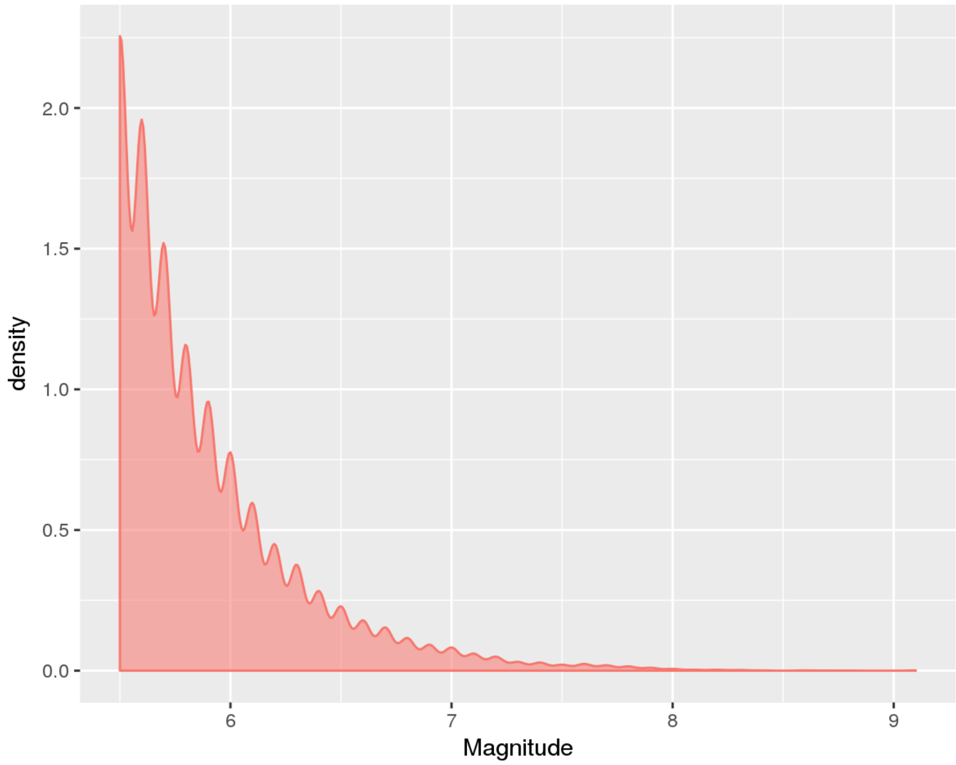

Fig 5. The density distribution of earthquake amplitude

Magnitude Type can be based on any of the following: - ML: Local (Richter) magnitude - MS: surface wave magnitude scale - MB (Mb): Body wave magnitude scale - MW (Mw): Moment magnitude scale - MD (Md): Duration magnitude/signal duration.

Fig. 6 plots interactive histograms of different magnitude types. It can be seen that the magnitude calculation methods in recent years have mostly adopted MWC and MWW.

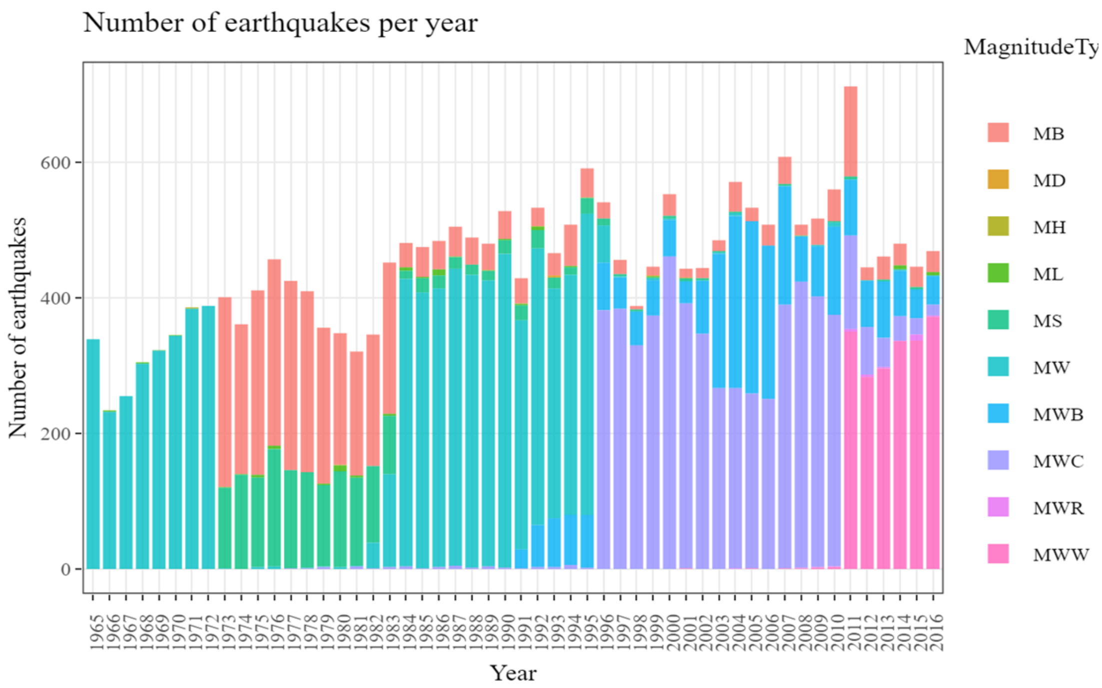

Fig 6. Time distribution histogram of different types of Magnitude

The depth of earthquake

The depth of earthquake refers to the vertical distance from the source to the ground (epicenter). The study of focal depth is of great significance for exploring the deep environment of earthquake gestation and occurrence, the active tectonic background of seismic energy accumulation and release, as well as the internal structural deformation and mechanical properties of the crust. Therefore, we mapped the distribution of the depth of earthquakes that occurred globally from 1965 to 2016.

Therefore, we have drawn the distribution of the depth of earthquakes that occurred globally from 1965 to 2016, as shown in Fig. 7, where the size of the circle represents the depth, the larger the circle, the deeper the depth, and the smaller the circle, the shallower the depth.

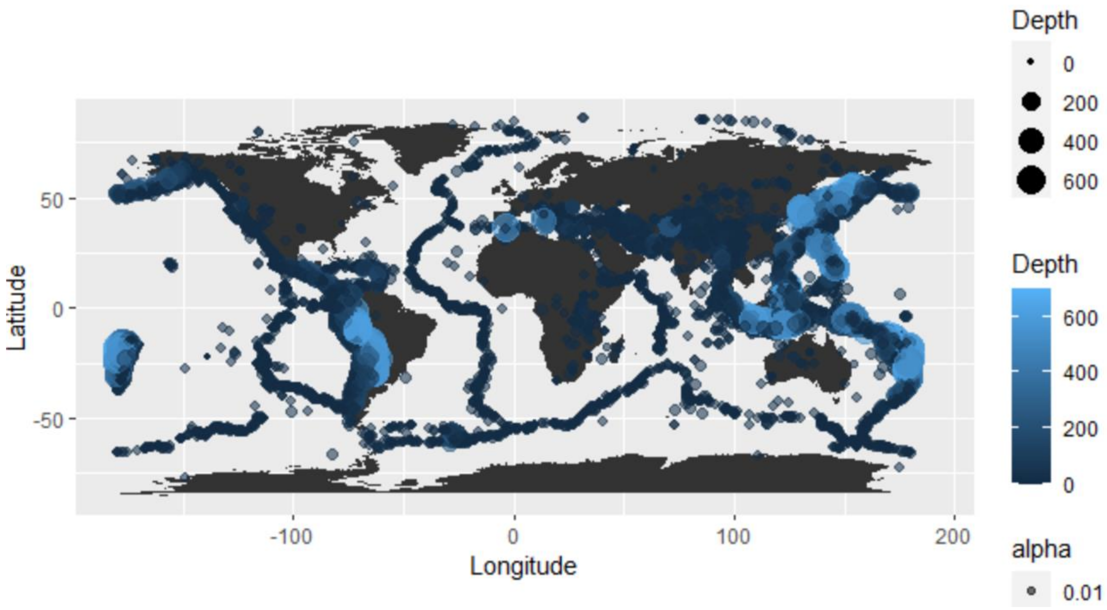

Fig 7. The distribution of the depth of earthquakes that occurred globally from 1965 to 2016

We found that most of the deeper earthquakes are distributed in the Pacific Rim seismic belt, while the shallower earthquakes mostly occur on the Eurasian continent and part of the sea.

It can be seen from Fig. 8 that the MWW calculation method is used to assess the magnitude of earthquakes on the Pacific Rim seismic belt.

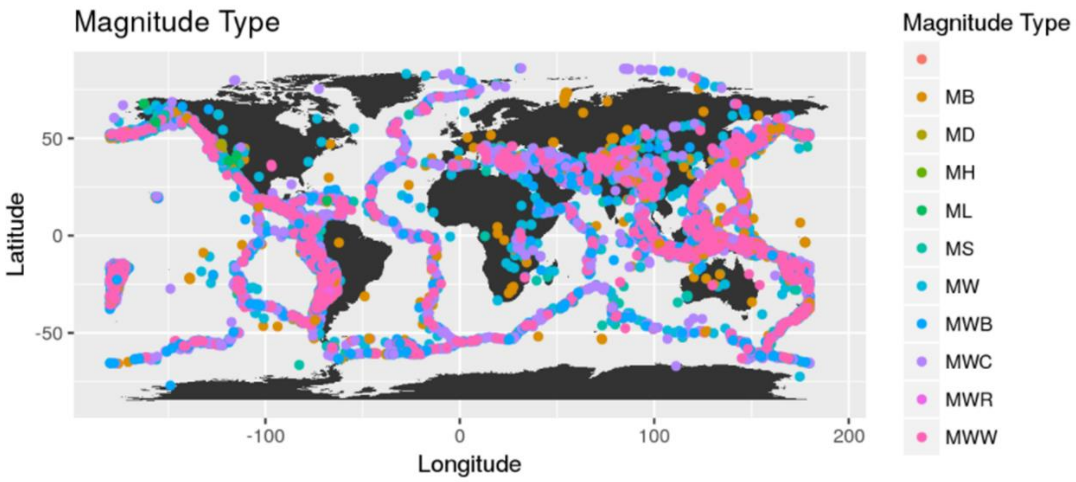

Fig 8. The distribution of the Magnitude type of earthquakes that occurred globally from 1965 to 2016

The spatial distribution of earthquakes on a global scale

As shown in Fig. 9, we mapped each earthquake that occurred from 1965 to 2016 in this data set on a world map (the map is built in R language). In addition, in order to more intuitively see the frequency of earthquakes in different regions of the world, we clustered the results in Fig. 9 to obtain Fig. 10.

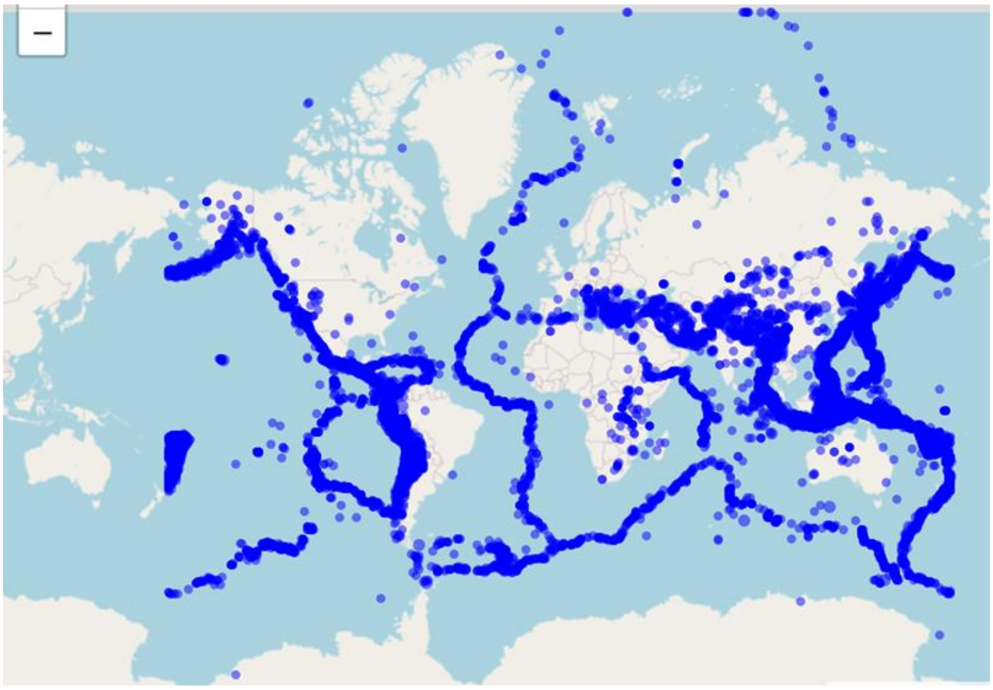

Fig 9. The spatial distribution of earthquakes on a global scale

Fig 10. Clustering results of the number of earthquakes

Most of the world's earthquakes occur in the Pacific Rim seismic belt, and Indonesia, the Philippines, and Japan in Asia are in the hardest hit areas. The second is the Eurasian seismic zone starting from Indonesia, passing through the western Indochina Peninsula and my country's Yunnan-Guizhou-Sichuan-Tibet region to the northern coast of the Mediterranean and extending to the Atlantic Ocean.

Earthquake distribution in China:

Fig. 9 and Fig. 10 are interactive maps. Therefore, we can get the distribution of earthquakes in China through partial zooming operations (Fig. 11 and Fig. 12).

(Since this map is a built-in map of the R language system, the border of China is wrong.)

Fig 11. Distribution of earthquakes in China

Fig 12. Clustering results of Fig. 11

China is located between the two major seismic belts in the world-the Pacific Rim seismic belt and the Eurasian seismic belt. The earthquakes are mainly distributed in five regions: (1) Taiwan Province and its adjacent seas; (2) Yunnan, Guizhou, Sichuan and Tibet in the southwest; (3) Gansu, Qinghai, and Ningxia in the northwest; (4) Taihang Mountains in North China, Fenwei River Valley, Yinshan-Yanshan area, central Shandong and Bohai Bay; (5) Southeastern coastal areas such as Guangdong Province and Fujian Province.

The frequency of earthquakes in the west is high, and the impact of earthquakes in the east is great. For example, Tangshan is located in my country's Bohai Rim seismic belt. The South Pacific plate and the Asian plate on this seismic belt frequently squeeze, causing the internal rupture of the Pacific plate.

Conclusion

- From 1965 to 2016, the number of global earthquakes with magnitude >5.5 showed an overall upward trend.

- From 1965 to 2016, earthquakes with a magnitude of 5.5-6 occurred with the highest frequency.
- Most of the earthquakes occur in the Pacific Rim seismic belt.
- There are five earthquake-prone areas in China.

The code for Fig. 2

```
library(dplyr)
library(readr)
library(ggplot2)
library(plotly)
library("RColorBrewer")
library(VIM)
Eqdata <- read_csv("database.csv")
usedataname <- c("Date","Latitude","Longitude", "Magnitude","Magnitude Type")
usedata = Eqdata[usedataname]
colnames(usedata) <- c("Date","Latitude","Longitude", "Magnitude","MagnitudeType")
usedata["Year"] <- sapply(usedata$Date,function(x) {strsplit(x,"/")[[1]][3]})
## 剔除缺失数据
VIM::aggr(usedata,prop = FALSE)
usedata = usedata[!is.na(usedata$Year),]
usedata = usedata[!is.na(usedata$MagnitudeType),]
summary(usedata)
```

The code for Fig. 3 and Fig. 4

```
yeardata <- usedata%>%
  group_by(Year)%>%
  summarise(Eqnum = n())

ggplot(yeardata, aes(x =   Year, y = Eqnum, color = "blue" ) ) +
```

```
  geom_point()  + geom_line()
## 可交互图像
p1 <- ggplot(yeardata) +
  theme_bw(base_family = "STKaiti") +
  geom_bar(aes(Year,Eqnum),width = 0.8,fill = "blue",stat = "identity",alpha = 0.9) +
  theme(axis.text.x = element_text(angle = 90)) +
  labs(x = "Year",y = "Number of earthquakes",title = "Number of earthquakes per year")
ggplotly(p1)
```

The code for Fig. 5

```
ggplot(usedata, aes(x = Magnitude, colour = 'Magnitude Type',
                    fill = 'Magnitude Type', alpha = 0.01)) +
  geom_density()
```

The code for Fig. 6

```
## 地震每年每种类型发生次数
YTdata <- usedata%>%
  group_by(Year,MagnitudeType)%>%
  summarise(Eqnum = n())

## 可交互图像
p2 <- ggplot(YTdata,aes(fill = MagnitudeType)) +
  theme_bw(base_family = "STKaiti") +
  geom_bar(aes(Year,Eqnum),width = 0.8,stat = "identity",alpha = 0.8) +
  theme(axis.text.x = element_text(angle = 90)) +
  labs(x = "Year",y = "Number of earthquakes",title = "Number of earthquakes per year")
ggplotly(p2)
```

The code for Fig. 7

```
earthquake <- read_csv("database.csv")
dim(earthquake)
```

```
str(earthquake)

earthquake <- earthquake[, Date := as.Date(Date, format = "%m/%d/%Y")]

earthquake <- earthquake[, Year := year(Date)]

world <- map_data("world")

ggplot() + geom_polygon(data = world, aes(x =    long,    y= lat, group = group)) +

  geom_point(data = earthquake,

             aes(x = Longitude, y = Latitude , alpha = 0.01, color = Depth, size = Depth)) +

  stat_density2d(show.legend = FALSE) + xlab("Longitude") +ylab("Latitude") + coord_quickmap()
```

The code for Fig. 8

```
ggplot() + geom_polygon(data = world, aes(x =    long,    y= lat, group = group)) +

  geom_point(data = earthquake,

             aes(x = Longitude, y = Latitude, colour = 'Magnitude Type' )) +

  stat_density2d(show.legend = FALSE) + xlab("Longitude") +ylab("Latitude") + ggtitle("Magnitude Type")+ coord_quickmap()
```

The code for Fig. 9-12

```
###地震发生的地图可视化

library(leaflet)

usedata["infor"] = paste(usedata$Year," 年 :",usedata$Magnitude," Type:",usedata$MagnitudeType,sep = "")

map <- leaflet(data = usedata,width = 800, height = 500) %>%

  addTiles() %>%

  addCircleMarkers(lng = usedata$Longitude, lat = usedata$Latitude,

                   stroke = FALSE,color = "blue",

                   fillOpacity = 0.5,radius = ~(usedata$Magnitude / 2),

                   popup = ~(usedata$infor))


map
```

```
ggplot() + geom_polygon(data = world, aes(x =    long,    y= lat, group = group)) +
  geom_point(data = earthquake,
                    aes(x = Longitude, y = Latitude,    alpha = 0.01, size = Magnitude)) +
  stat_density2d(show.legend = FALSE) + xlab("Longitude") +ylab("Latitude") + ggtitle("Reviewed Only    ") + coord_quickmap()

#聚类
map <- leaflet(data = usedata) %>%
  addTiles() %>%
  addCircleMarkers(lng = usedata$Longitude, lat = usedata$Latitude,
                           clusterOptions = markerClusterOptions(),
                           popup = ~(usedata$infor))

map
```