

# earthquakes

2024-01-25

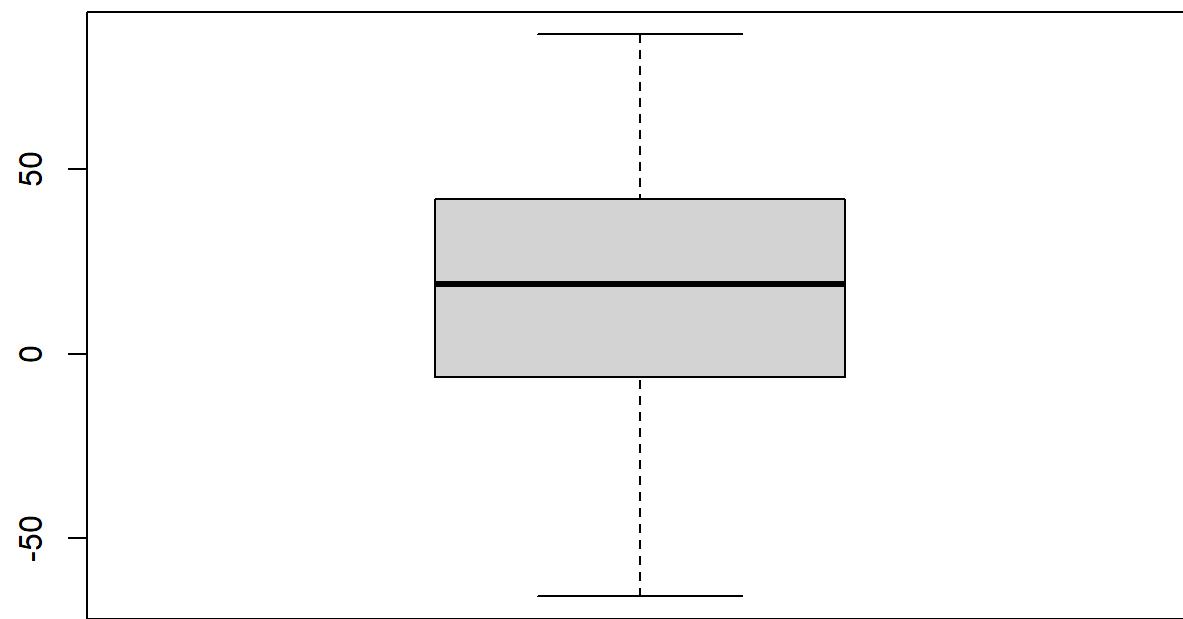
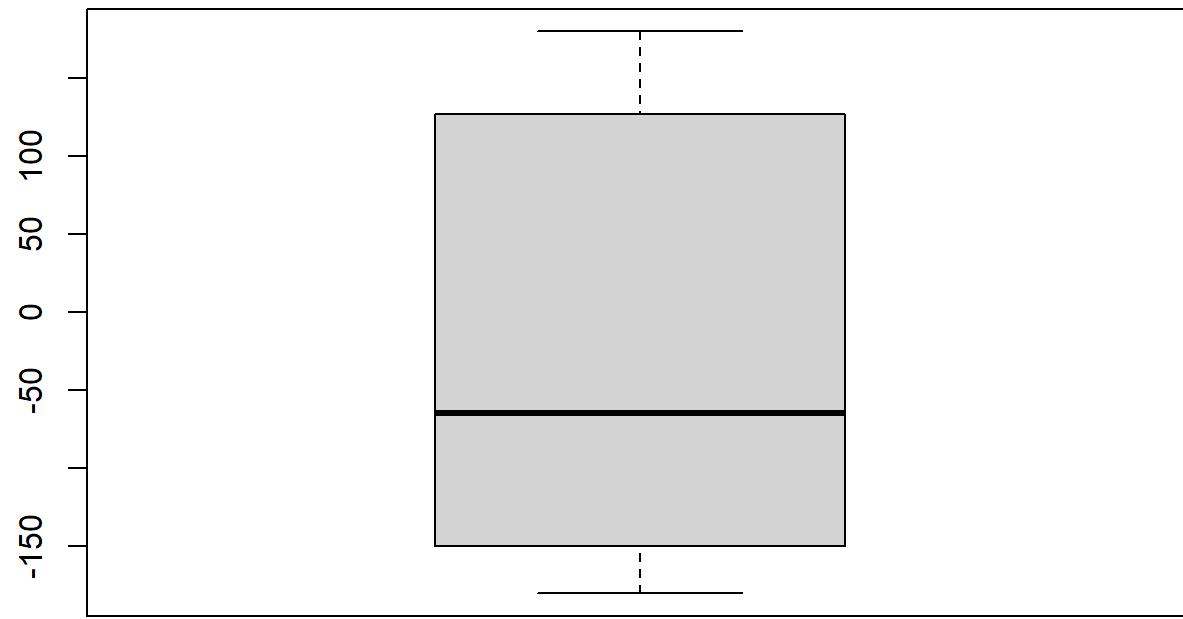
Install packages

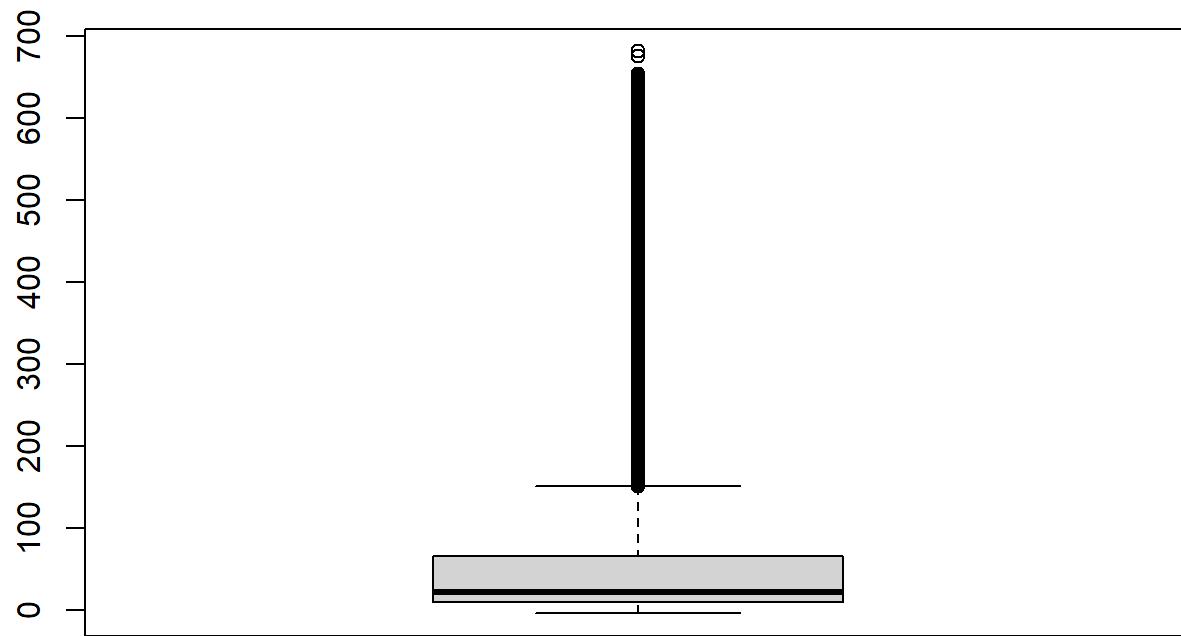
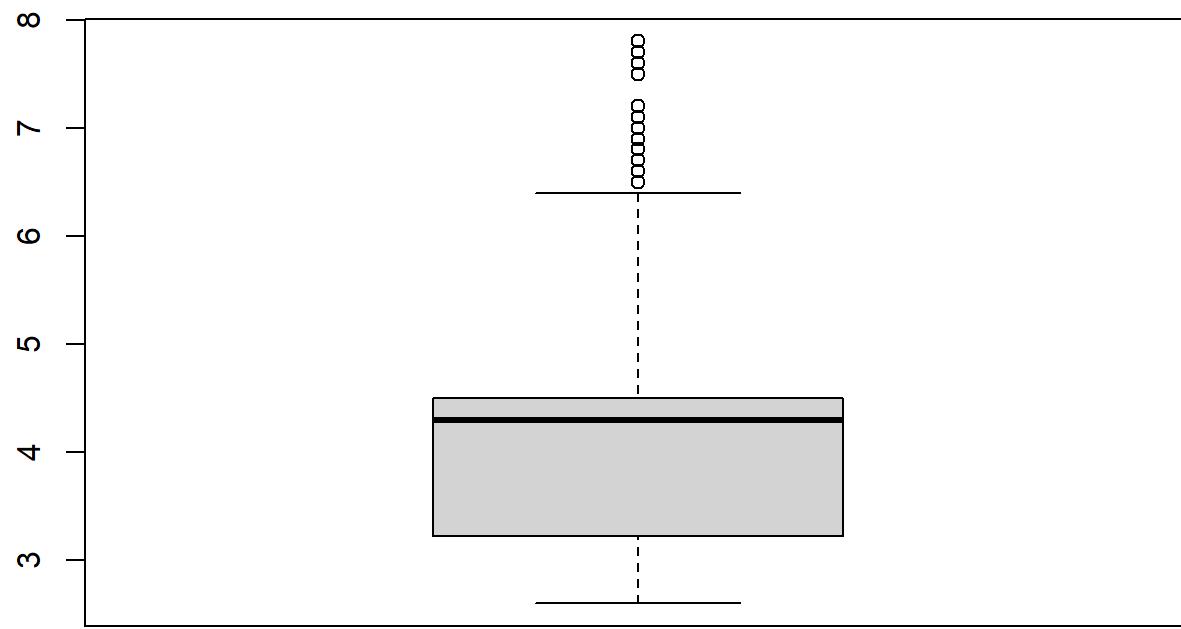
```
pacotes <- c("dplyr", "tidyverse", "RColorBrewer", "ggplot2", "gplots", "lattice", "fields", "viri
dis", "corrplot", "htmltools", "htmlwidgets")
if(sum(as.numeric(!pacotes %in% installed.packages())) != 0){
  instalador <- pacotes[!pacotes %in% installed.packages()]
  for(i in 1:length(instalador)) {
    install.packages(instalador, dependencies = T)
    break()
  }
  sapply(pacotes, require, character = T)
} else {
  sapply(pacotes, require, character = T)
}
```

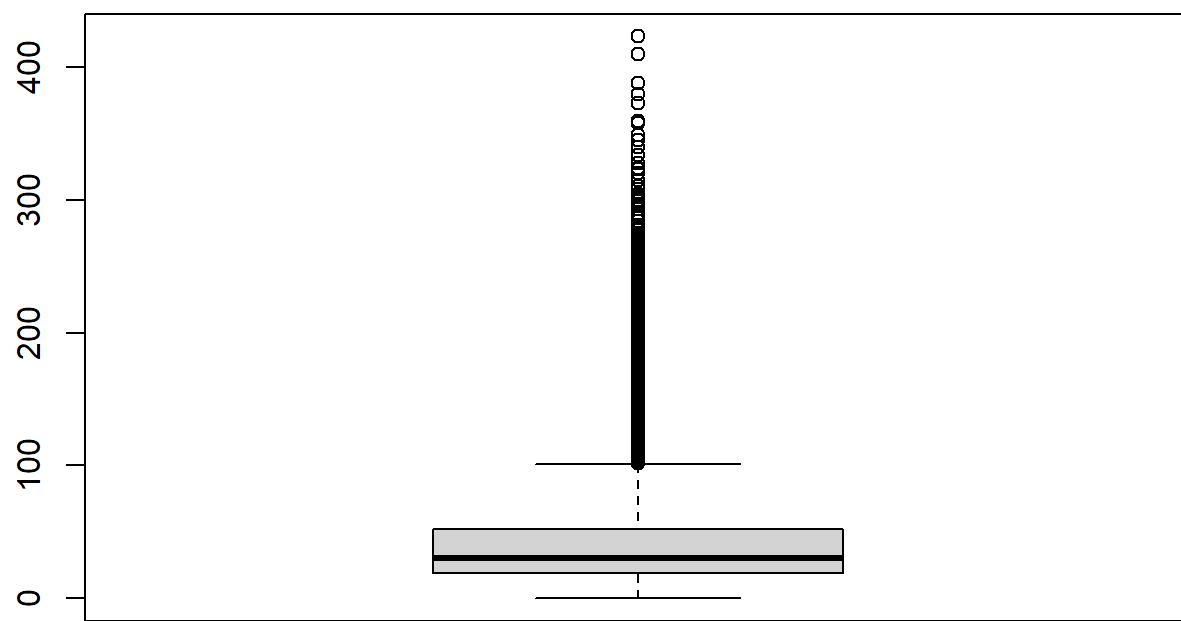
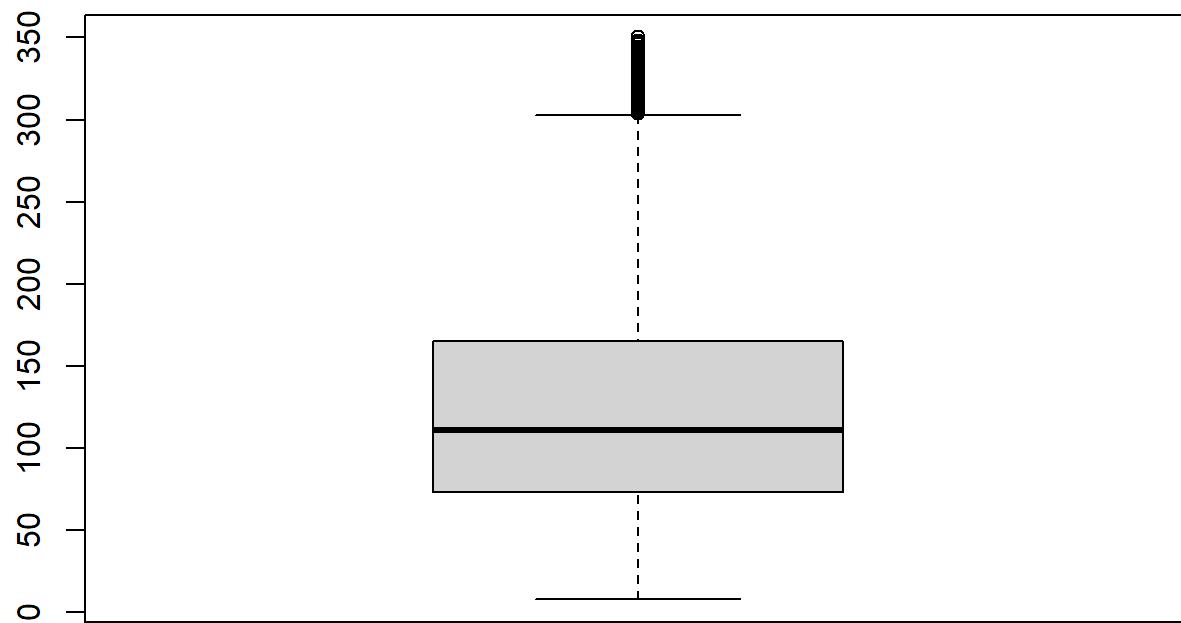
Data cleaning

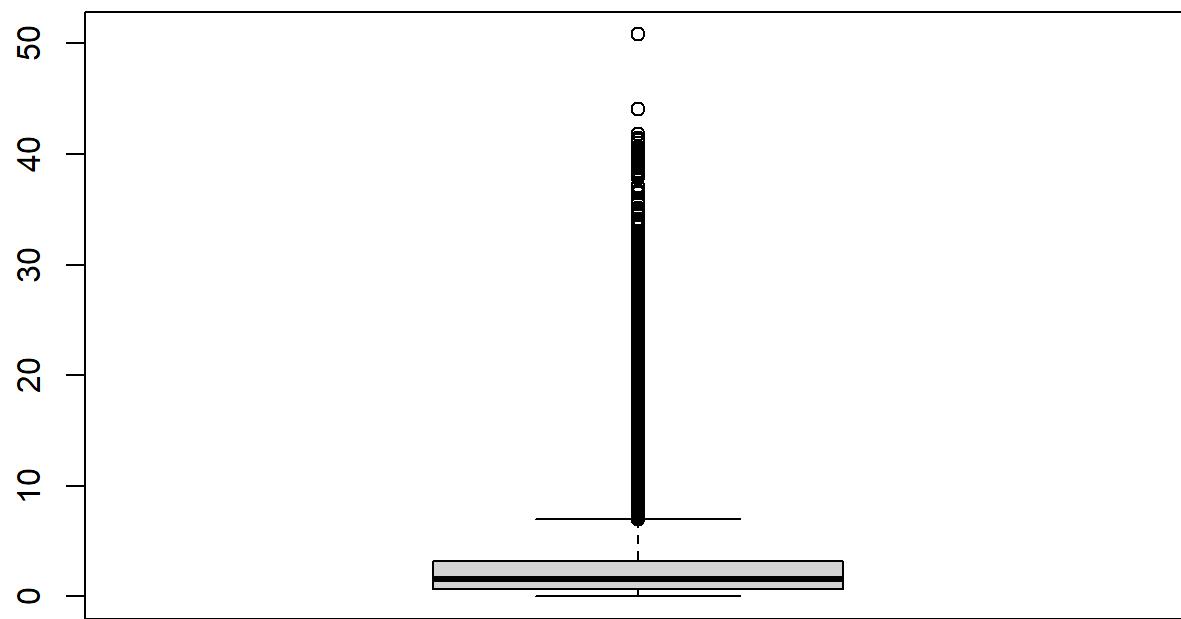
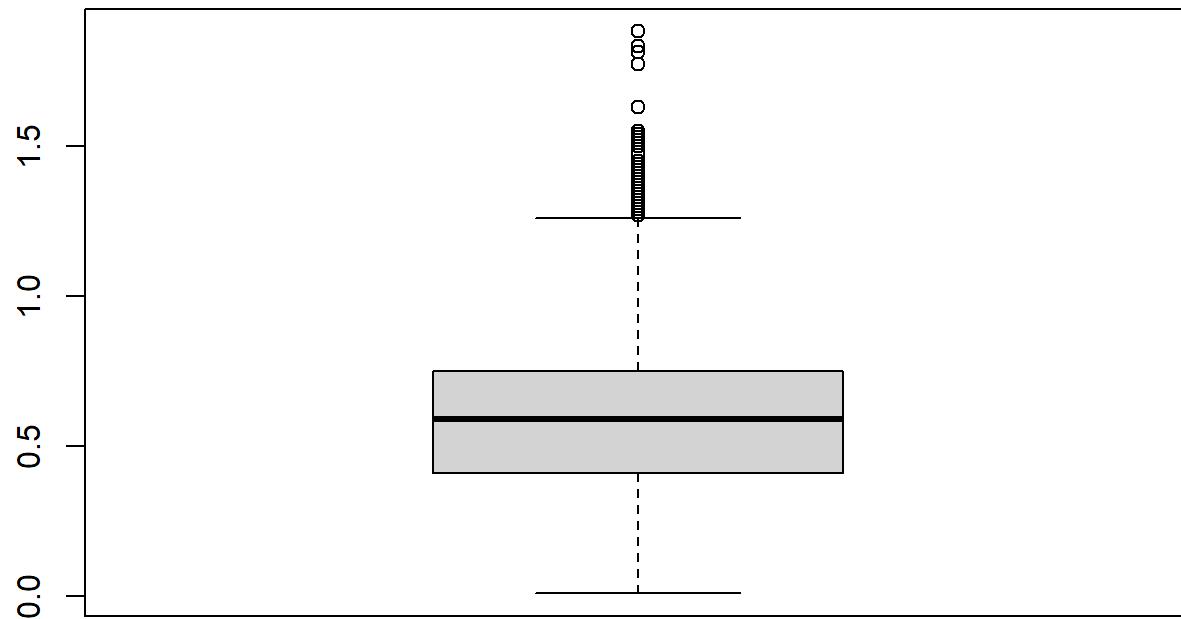
```
df <- read_csv("earthquakes_2023_global.csv")
head(df)
str(df)
dim(df) # 26642 obs 22 col
colSums(is.na(df)) # NAs = nst; gap; dmin; plac; horizontal error; magerror, magnst
sum(is.na(df)) # total NAs 11104
sum(rowSums(is.na(df)) > 0) # number of obs with NAs = 3771
#df$nst # Number of seismic stations that reported the earthquake.
#df$gap # The gap between different seismic stations' coverage.
#df$dmin # Minimum distance to the earthquake epicenter for the nearest station.
#df$horizontalError #Horizontal error in location determination.
#df$magError #Error in magnitude determination.
#df$magNst #Number of seismic stations used to calculate the magnitude.
#3771/26642 = 0.1415434 - 15% of obs has null values, we will try first EDA by removing them all
colnames(df)

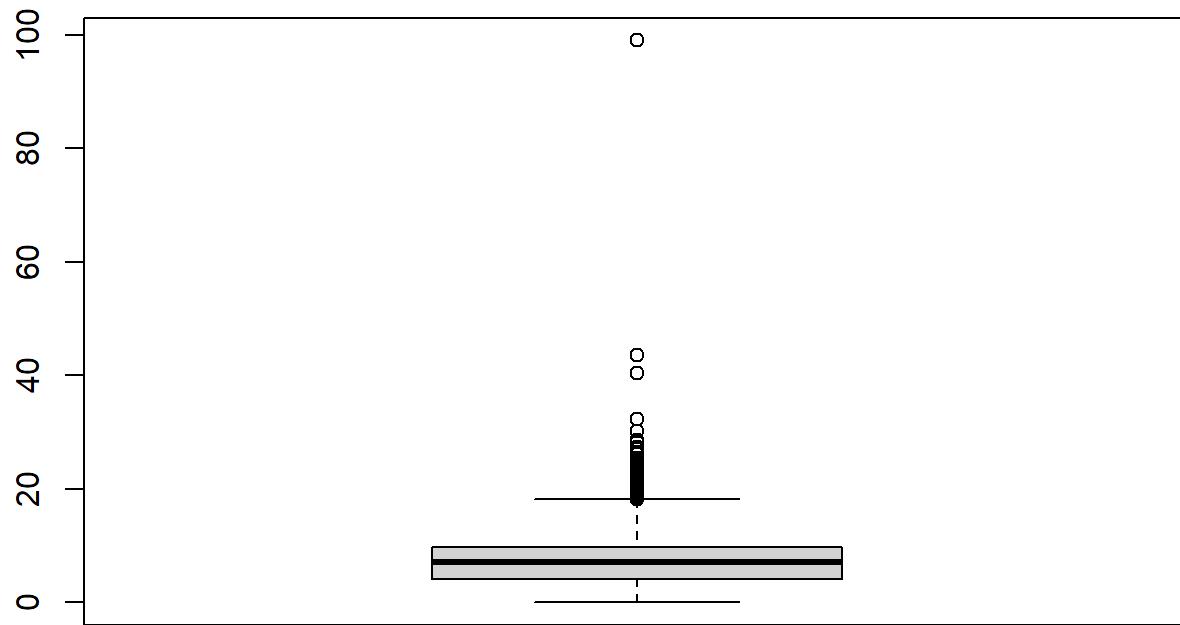
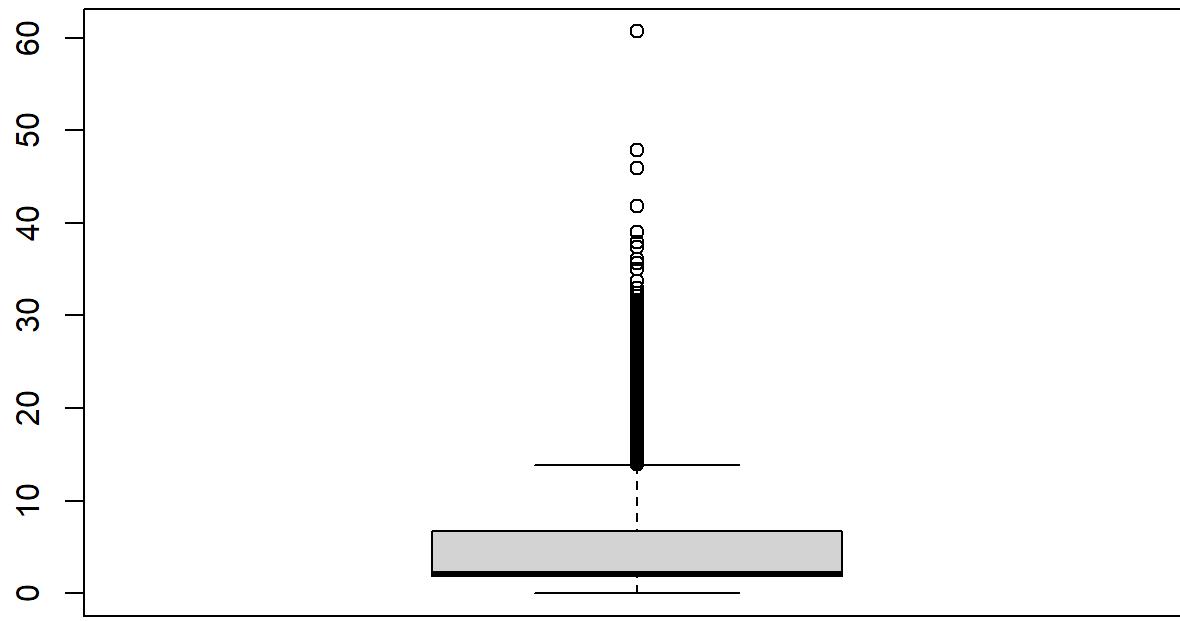
# Check outliers
lapply(names(df), function(col) {
  if (is.numeric(df[[col]])) {
    boxplot(df[[col]], main = col)
  }
})
```

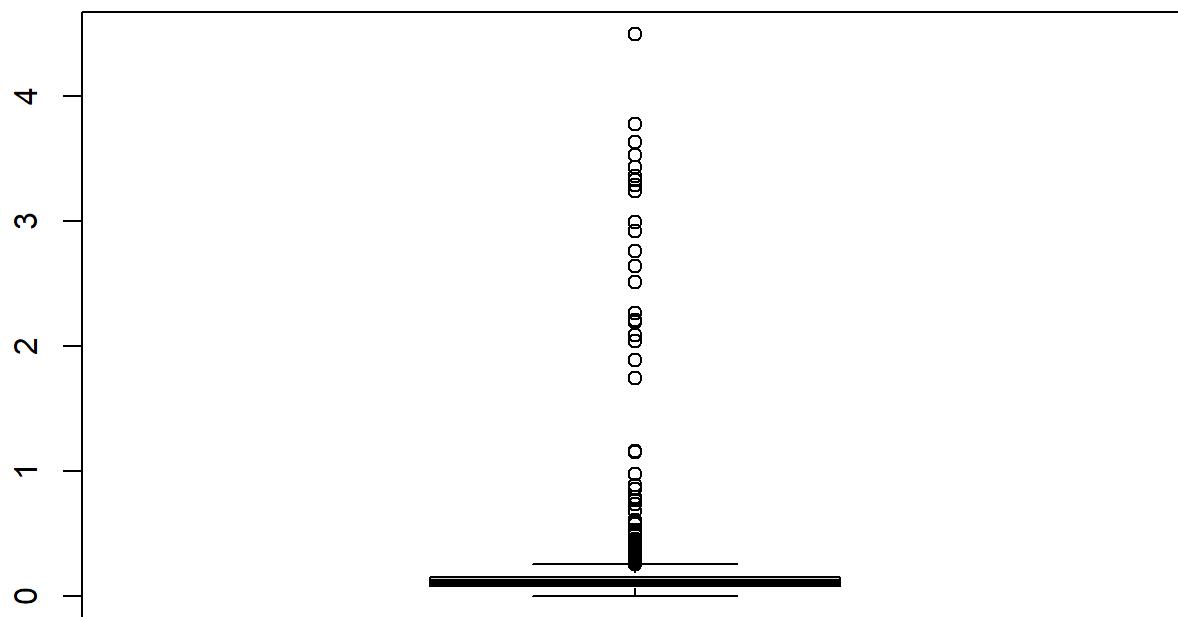
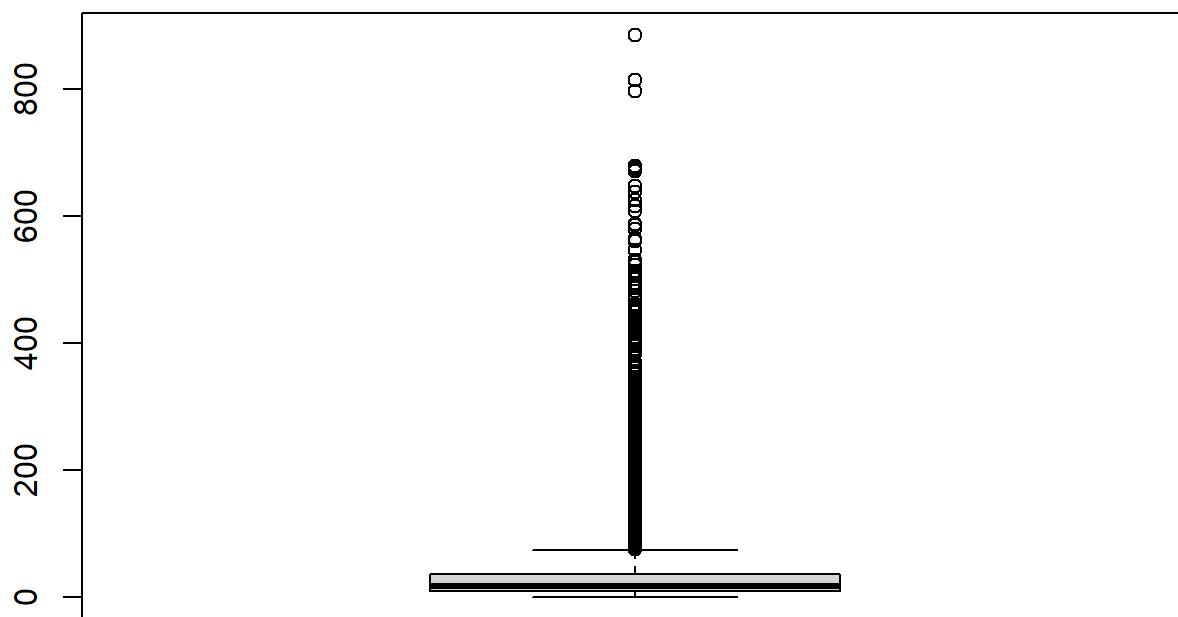
**latitude****longitude**

**depth****mag**

**nst****gap**

**dmin****rms**

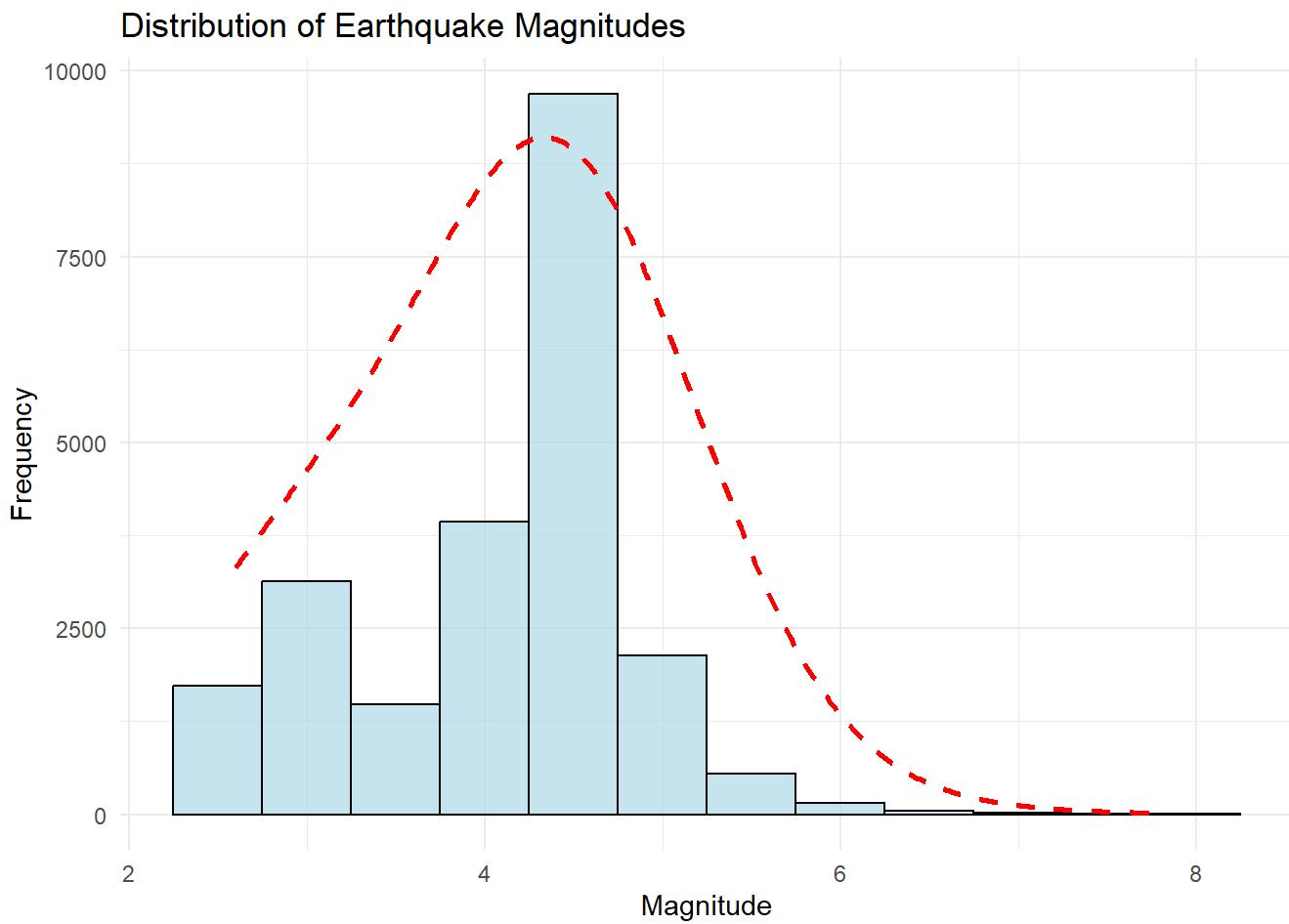
**horizontalError****depthError**

**magError****magNst**

```
df_cleaned <- na.omit(df) # remove obs with NAs
dim(df_cleaned) #22871 obs      22 col
head(df_cleaned)
```

Check magnitude distribution

```
ggplot(df_cleaned, aes(x = mag)) +
  geom_histogram(binwidth = 0.5, fill = "lightblue", color = "black", alpha = 0.7) +
  geom_density(aes(y = ..count..), color = "red", linetype = "dashed", size = 1, bw = 0.7) +
  labs(title = "Distribution of Earthquake Magnitudes",
       x = "Magnitude", y = "Frequency") +
  theme_minimal()
```



Calculate correlation matrix

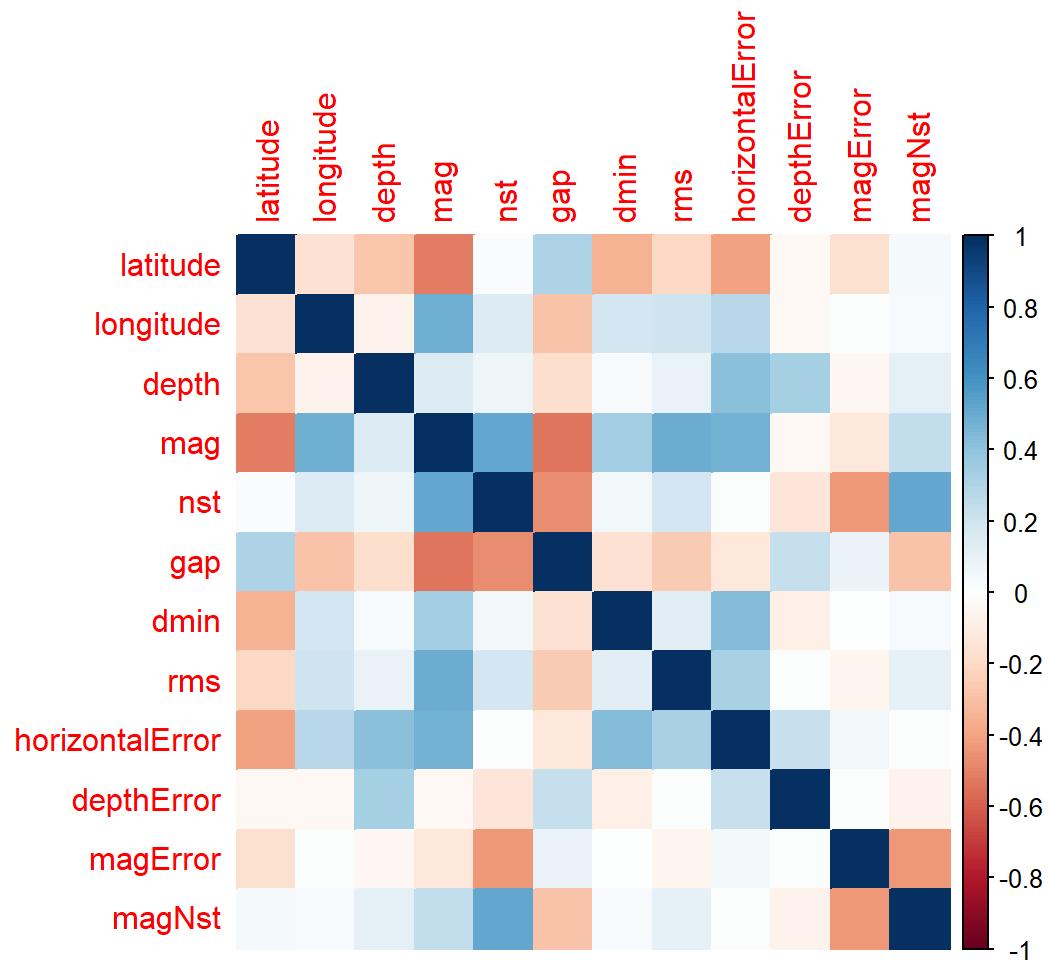
```
# Select numeric columns
numeric_columns <- df_cleaned[, sapply(df_cleaned, is.numeric)]

correlation_matrix <- cor(numeric_columns)
correlation_matrix <- round(cor(numeric_columns), 2)
correlation_matrix
```

```
##          latitude longitude depth   mag    nst   gap dmin   rms
## latitude      1.00     -0.16 -0.28 -0.52  0.02  0.30 -0.34 -0.21
## longitude     -0.16      1.00 -0.07  0.48  0.14 -0.29  0.18  0.20
## depth        -0.28     -0.07  1.00  0.14  0.07 -0.18  0.03  0.09
## mag          -0.52      0.48  0.14  1.00  0.51 -0.54  0.34  0.49
## nst           0.02      0.14  0.07  0.51  1.00 -0.47  0.05  0.18
## gap            0.30     -0.29 -0.18 -0.54 -0.47  1.00 -0.16 -0.26
## dmin         -0.34      0.18  0.03  0.34  0.05 -0.16  1.00  0.13
## rms           -0.21      0.20  0.09  0.49  0.18 -0.26  0.13  1.00
## horizontalError -0.41      0.27  0.41  0.47  0.01 -0.13  0.43  0.32
## depthError     -0.04     -0.04  0.33 -0.04 -0.14  0.23 -0.09  0.01
## magError       -0.17      0.01 -0.05 -0.13 -0.44  0.08  0.00 -0.06
## magNst          0.04      0.03  0.11  0.24  0.51 -0.29  0.03  0.10
##          horizontalError depthError magError magNst
## latitude        -0.41     -0.04   -0.17   0.04
## longitude        0.27     -0.04    0.01   0.03
## depth            0.41      0.33   -0.05   0.11
## mag              0.47     -0.04   -0.13   0.24
## nst              0.01     -0.14   -0.44   0.51
## gap             -0.13      0.23    0.08  -0.29
## dmin            0.43     -0.09    0.00   0.03
## rms              0.32      0.01   -0.06   0.10
## horizontalError  1.00      0.22    0.05   0.01
## depthError        0.22      1.00    0.01  -0.07
## magError          0.05      0.01    1.00  -0.44
## magNst            0.01     -0.07   -0.44   1.00
```

```
#correlation_matrix <- round(correlation_matrix, 2)
#correlation_matrix
```

```
corrplot(correlation_matrix, method="color")
```



Check magnitude of earthquakes around the world

```

color_palette <- leaflet:::colorBin(palette = "RdYlBu", domain = df_cleaned$mag, bins = 5, reverse = TRUE)

# Create the Leaflet map
mymap <- leaflet:::leaflet(df_cleaned) %>%
  leaflet:::addTiles() %>%
  leaflet:::addCircleMarkers(
    ~longitude, ~latitude,
    radius = ~sqrt(mag) * 3, # Adjust the size based on magnitude
    color = ~color_palette(mag),
    fillOpacity = 0.7, # Adjust the transparency of the circles
    popup = ~paste("Place: ", place, "<br>Magnitude: ", mag, "<br>Updated: ", updated)
  ) %>%
  leaflet:::addLegend(
    position = "bottomright",
    pal = color_palette,
    values = ~mag,
    title = "Magnitude"
  )

# Create an HTML title
title_html <- tags$h2("Earthquake Magnitudes Worldwide in 2023")

# Display the map with the title
htmltools:::tagList(title_html, mymap)

```

## Earthquake Magnitudes Worldwide in 2023



Leaflet (<https://leafletjs.com>) | © OpenStreetMap (<https://openstreetmap.org/copyright/>), ODbL (<https://opendatacommons.org/licenses/odbl/>)

```
#(mymap, file = "interactive_map.html")
```