Liegata_3361911_Assignment_3_RMD

Julius Viktor Liegata

2023-10-15

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com (http://rmarkdown.rstudio.com).

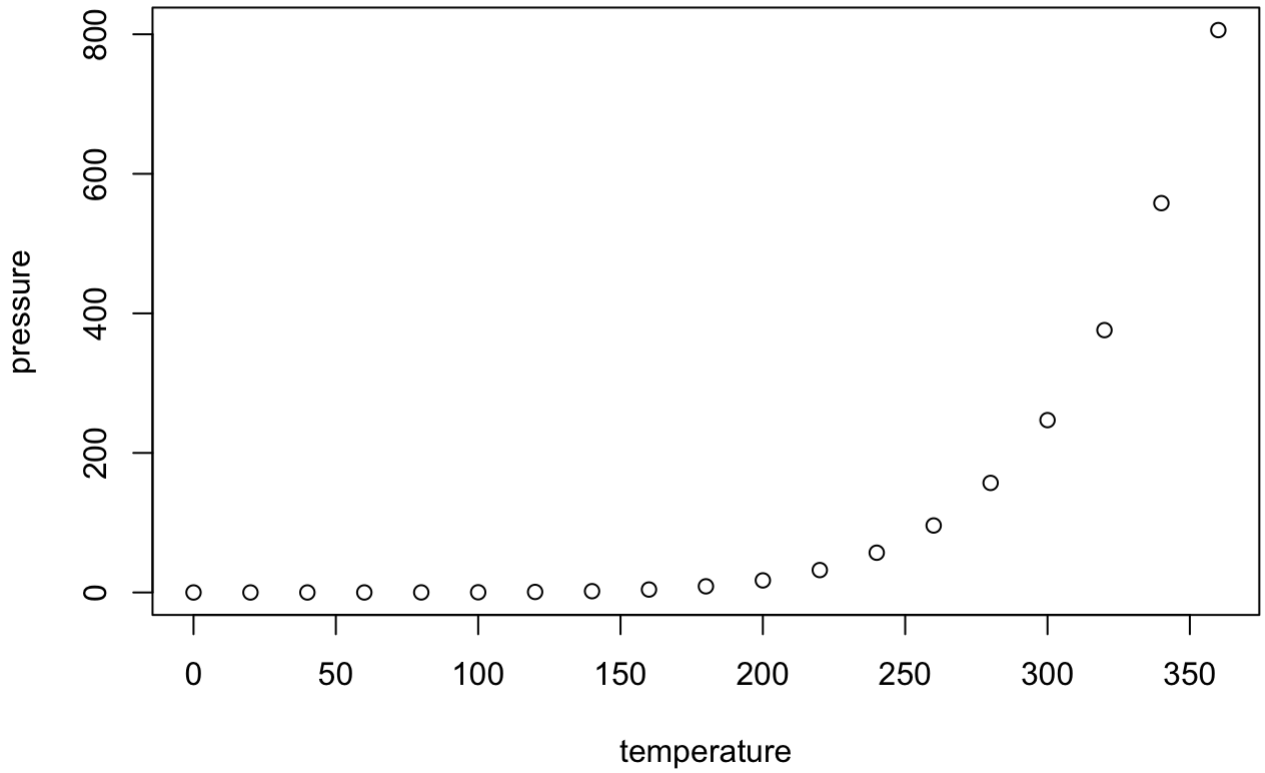
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist       
##  Min.   : 4.0   Min.   :  2.00  
##  1st Qu.:12.0   1st Qu.: 26.00  
##  Median :15.0   Median : 36.00  
##  Mean   :15.4   Mean   : 42.98  
##  3rd Qu.:19.0   3rd Qu.: 56.00  
##  Max.   :25.0   Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

```
library(grid)
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✔ dplyr     1.1.2          ✔ readr     2.1.4
## ✔ forcats   1.0.0          ✔ stringr   1.5.0
## ✔ ggplot2   3.4.4.9000     ✔ tibble    3.2.1
## ✔ lubridate 1.9.2          ✔ tidyr     1.3.0
## ✔ purrr     1.0.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflic
ts to become errors
```

```
library(sf)
```

```
## Linking to GEOS 3.11.0, GDAL 3.5.3, PROJ 9.1.0; sf_use_s2() is TRUE
```

```
library(rnaturalearth)
```

```
## The legacy packages maptools, rgdal, and rgeos, underpinning the sp package,
## which was just loaded, were retired in October 2023.
## Please refer to R-spatial evolution reports for details, especially
## https://r-spatial.org/r/2023/05/15/evolution4.html.
## It may be desirable to make the sf package available;
## package maintainers should consider adding sf to Suggests:.
## Support for Spatial objects (`sp`) will be deprecated in {rnaturalearth} and will
be removed in a future release of the package. Please use `sf` objects with {rnatural
earth}. For example: `ne_download(returnclass = 'sf')`
```

```
library(scales)
```

```
##
## Attaching package: 'scales'
##
## The following object is masked from 'package:purrr':
##
##     discard
##
## The following object is masked from 'package:readr':
##
##     col_factor
```

```
library(naniar)
library(ggplot2)
library(RColorBrewer)
library(rvest)
```

```
##
## Attaching package: 'rvest'
##
## The following object is masked from 'package:readr':
##
##     guess_encoding
```

```
earthquake_list <- read_tsv("//Users/julius/Monash/FIT5145/Assignment_3/Earthquake_da
taset/katalog_gempa_v2.tsv")
```

```
## New names:
## Rows: 102515 Columns: 38
## ── Column specification
## ──────────────────────────────────────────────────────── Delimiter: "\t" chr
## (5): eventID, mag_type, location, agency, magTypeFM dbl (31): ...1, latitude,
## longitude, magnitude, depth, phasecount, azimuth_... dttm (2): datetime,
## datetimeFM
## ℹ Use `spec()` to retrieve the full column specification for this data. ℹ
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## • `` -> `...1`
```

```
earthquake_list <- earthquake_list[, 2:11]
```

```
#seperate one of the column
location_list <- strsplit(earthquake_list$location,",")

location_list_df <- do.call(rbind.data.frame,location_list)

colnames(location_list_df) <- c("province","area")

earthquake_list <- cbind(earthquake_list,location_list_df)
```

```
#altering the datasets

#specifying the data that will be dropped
column_drop <- c("eventID","azimuth_gap","phasecount")

earthquake_list <- earthquake_list[, !names(earthquake_list) %in% column_drop]
```

```
#seperating datetime into two different column
datetime <- ymd_hms(earthquake_list$datetime)

earthquake_list$Date <- date(earthquake_list$datetime)
earthquake_list$Time <- format(earthquake_list$datetime, format = "%H:%M:%S")
```

```
#removing datetime column
earthquake_list <- earthquake_list[, -1]
```

```
earthquake_list <- earthquake_list%>%
  select(Date,Time,everything())
```

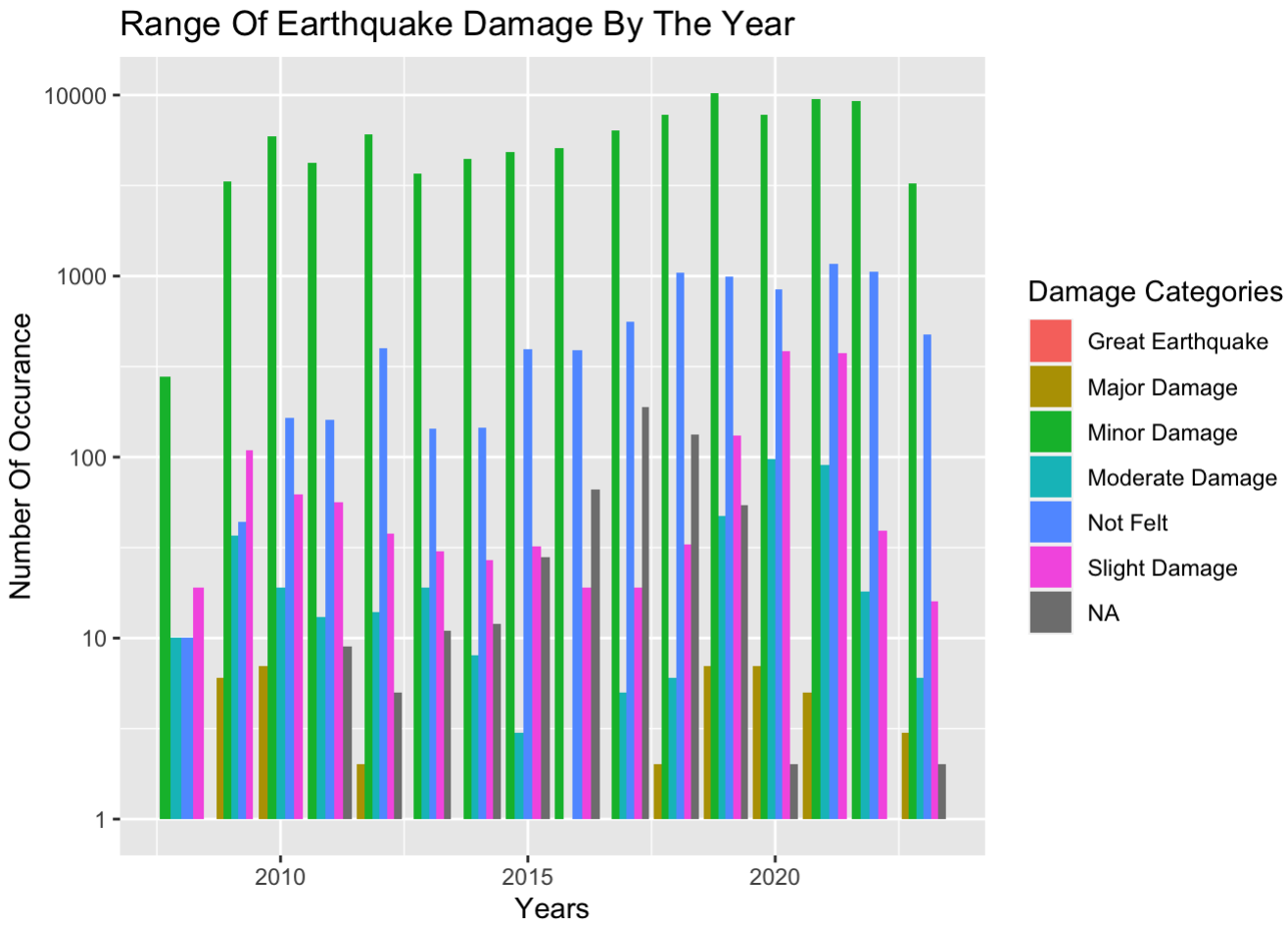
```
#removing white space
earthquake_list$area <- trimws(earthquake_list$area)
```

```
#categorizing the strength of the earthquake

earthquake_list$magn_category <-
earthquake_list$magn_category <- ifelse(
  earthquake_list$magnitude <= 2.5, "Not Felt",
  ifelse(
    earthquake_list$magnitude > 2.5 & earthquake_list$magnitude <= 5.4, "Minor Damag
e",
    ifelse(
      earthquake_list$magnitude > 5.4 & earthquake_list$magnitude <= 6.0, "Slight Dam
age",
      ifelse(
        earthquake_list$magnitude > 6.0 & earthquake_list$magnitude <= 6.9, "Moderate
Damage",
        ifelse(
          earthquake_list$magnitude > 6.9 & earthquake_list$magnitude <= 7.9, "Major
Damage",
          ifelse(
            earthquake_list$magnitude > 7.9, "Great Earthquake",
            "Unknown"
          )
        )
      )
    )
  )
)
```

```
#analysis 1
#showing the number of earthquake by the year
earthquake_list <- earthquake_list %>%
  mutate(year = as.integer(format(Date, "%Y")))

#merging the datasets
occurance_plot <- earthquake_list %>%
  group_by(year,magn_category) %>%
  summarise(occurance = n()) %>%
  ggplot(aes(x = year, y = occurance, fill = magn_category))+
  scale_y_log10()+
  geom_bar(stat = "identity" ,position = "dodge")+
  labs(title = "Range Of Earthquake Damage By The Year", x = "Years", y = "Number Of
Occurance",fill = "Damage Categories")
```

```
## `summarise()` has grouped output by 'year'. You can override using the
## `.groups` argument.
```

```
occurance_plot
```




```
library(jsonlite)
```

```
##
## Attaching package: 'jsonlite'
```

```
## The following object is masked from 'package:purrr':
##
##     flatten
```

```
#creating a indonesian map
#read json file
indonesia_regions <- fromJSON("/Users/julius/Monash/FIT5145/Assignment_3/indonesia-geojson-master/indonesia-province.json")

ID <- indonesia_regions$features
```

```
library(geojsonio)
```

```
## Registered S3 method overwritten by 'geojsonsf':
##   method        from
##   print.geojson geojson
```

```
##
## Attaching package: 'geojsonio'
```

```
## The following object is masked from 'package:base':
##
##     pretty
```

```
library(broom)
```

```
indonesian_map <- geojson_read("https://raw.githubusercontent.com/superpikar/indonesia-geojson/master/indonesia-en.geojson", what = "sp")

indonesia_map_fortified <- tidy(indonesian_map)
```

```
## Warning: `tidy.SpatialPolygonsDataFrame()` was deprecated in broom 1.0.4.
## ℹ Please use functions from the sf package, namely `sf::st_as_sf()`, in favor
##   of sp tidiers.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

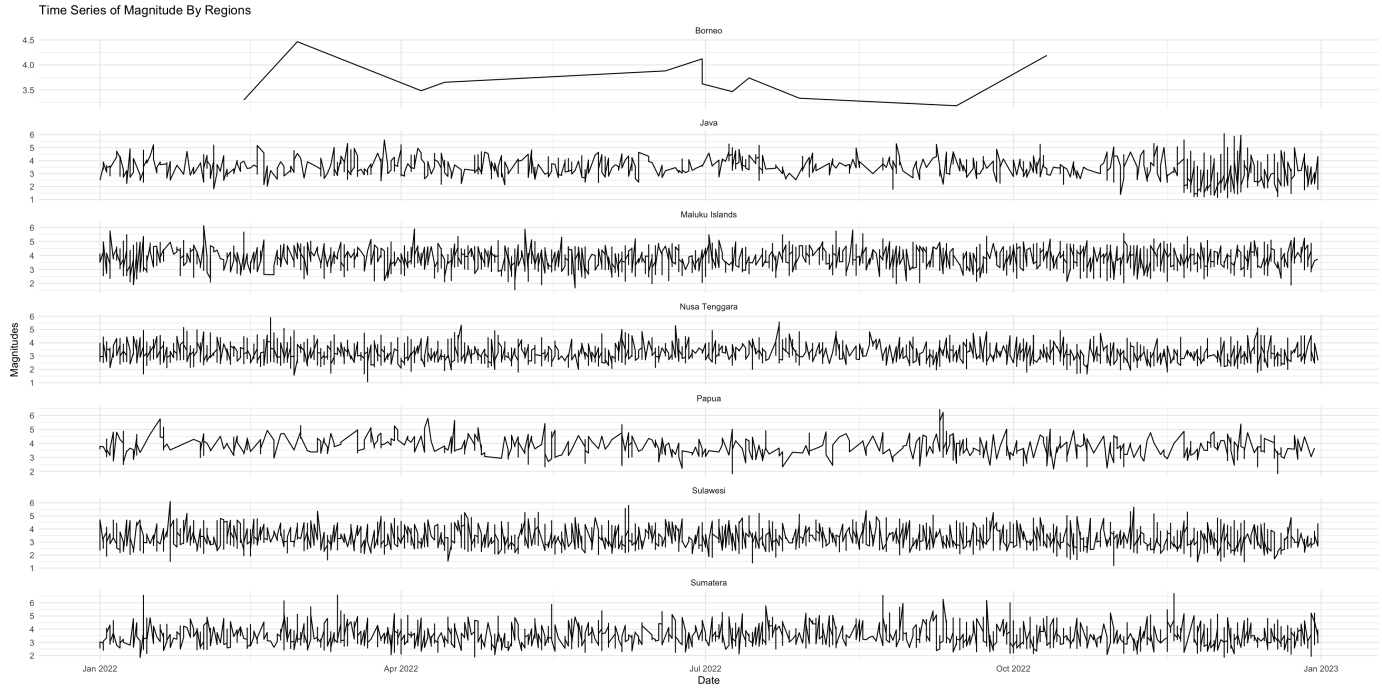
```
## Regions defined for each Polygons
```

```
#plotting the map
indonesia_map_plot <- indonesia_map_fortified %>%
  ggplot() +
  geom_polygon(
    data = indonesia_map_fortified,
    aes(x = long, y = lat, group = group, fill =long),
  ) +
  scale_fill_gradient(low = "#99aa33", high = "#dd9900") +  # Specify the gradient colors
  theme_void() +
  guides(fill = "none")+
  coord_map()

indonesia_map_plot
```

```
#altering the main file more
area_to_drop <- c("East of Philippine Islands","Mindanao","Northwest of Australia","P
hilippine Islands Region","South China Sea","South Indian Ocean","Sulu Archipelag
o","Timor Region","Timor Sea","Western Australia")

earthquake_list_2 <- earthquake_list %>%
  filter(!(province %in% area_to_drop))
```

```
#altering the name to fit the json files

maluku_islands <- c("Aru Islands Region","Banda Sea","Buru","Ceram Sea","Halmaher
a","North of Halmahera","Northern Molucca Sea","Seram","Tanimbar Islands Reg.","Tanim
bar Islands Region")

nusa_tenggara <- c("Bali Region","Bali Sea","Flores Region","Flores Sea","Savu Se
a","South of Sumba","South of Sumbawa","Sumba Region","Sumbawa Region")

borneo <- c("Borneo","Malay Peninsula")

sulawesi <- c("Celebes Sea","Minahassa Peninsula","Sulawesi","Talaud Islands")

papua <- c("Irian Jaya","Irian Jaya Region","Near North Coast of Irian Jaya","Near No
rth Coast of West Papua","Neara South Coast of Irian Jaya","nNew Guinea","W. Caroline
Islands","West Papua","West Papua Region")

java <- c("Java","Java Sea","South of Java")

sumatera <- c("Nicobar Islands","Northern Sumatra","Off West Coast of Northern Sumatr
a","Southern Sumatra","Southwest of Sumatra","Sunda Strait")

earthquake_list_3 <- earthquake_list_2 %>%
  mutate(regions = case_when(province %in% maluku_islands ~ "Maluku Islands",
                             province %in% nusa_tenggara ~ "Nusa Tenggara",
                             province %in% borneo ~ "Borneo",
                             province %in% sulawesi ~ "Sulawesi",
                             province %in% papua ~ "Papua",
                             province %in% java ~ "Java",
                             province %in% sumatera ~ "Sumatera",
                             TRUE ~ "Other"
  ))
```

```
#analysis 2
#filter the earthquakes

earthquake_2022 <- earthquake_list_3 %>%
  filter(year == 2022)

earthquake_2022 <- earthquake_2022 %>%
  filter(!(regions == "Other"))

earthquake_time_series <- ggplot(earthquake_2022, aes(x = Date, y = magnitude))+
  geom_line() +
  facet_wrap(~regions, scales = "free_y", ncol =1) +
  labs(title = "Time Series of Magnitude By Regions",x = "Date",y = "Magnitudes") +
  theme_minimal() +
  theme(legend.position = "none")

earthquake_time_series
```

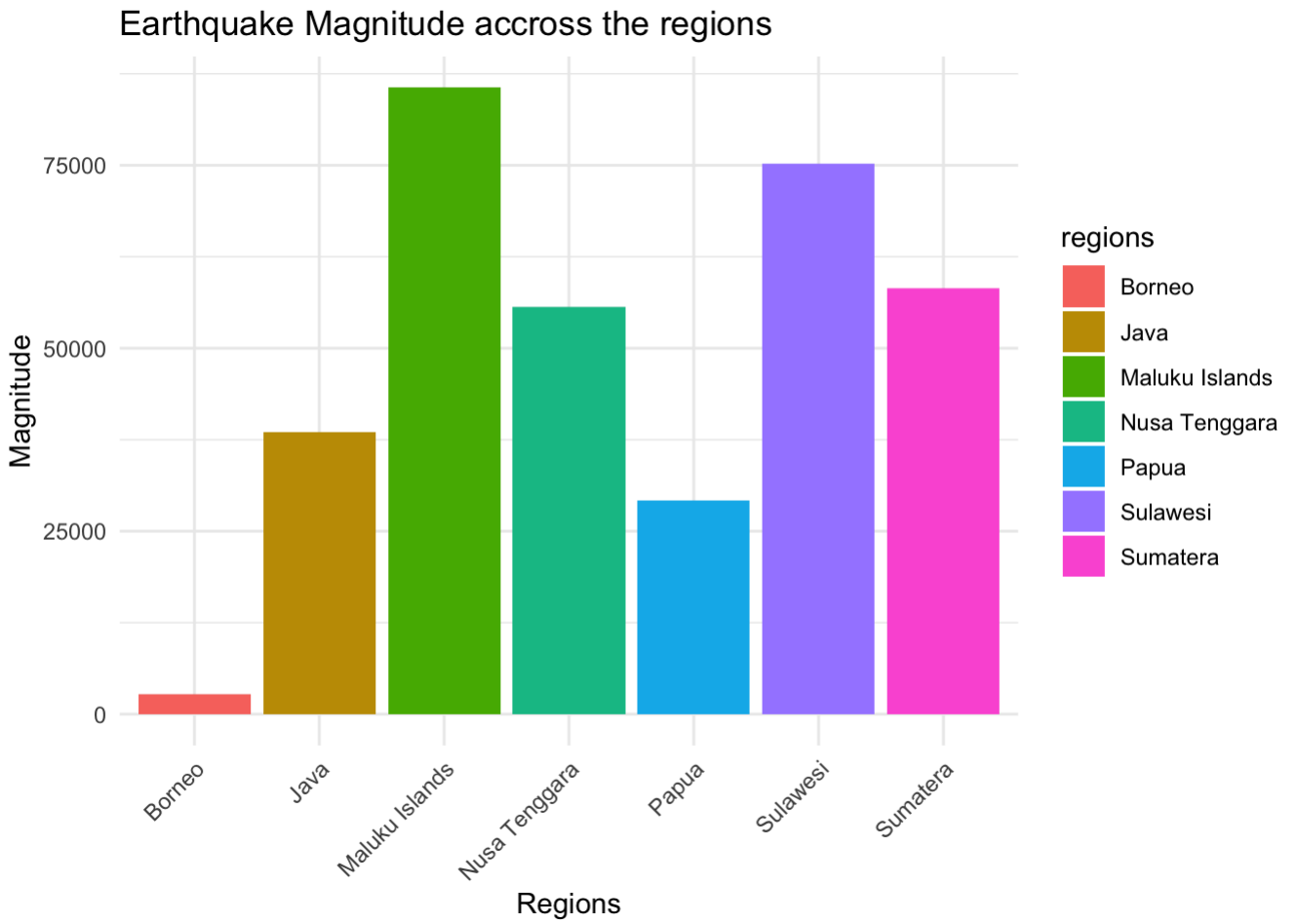



```
#pie chart base on the year
```

```
earthquake_list_3 <- earthquake_list_3 %>%
  filter(!(regions == "Other"))
```

```
#analysis 3
magnitude_region_plot <- ggplot(earthquake_list_3, aes(x = regions, y = magnitude, fi
ll = regions)) +
  geom_bar(stat = "identity") +
  labs(title = "Earthquake Magnitude accross the regions",
       x = "Regions",y = "Magnitude") +
  theme_minimal()+
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
magnitude_region_plot
```

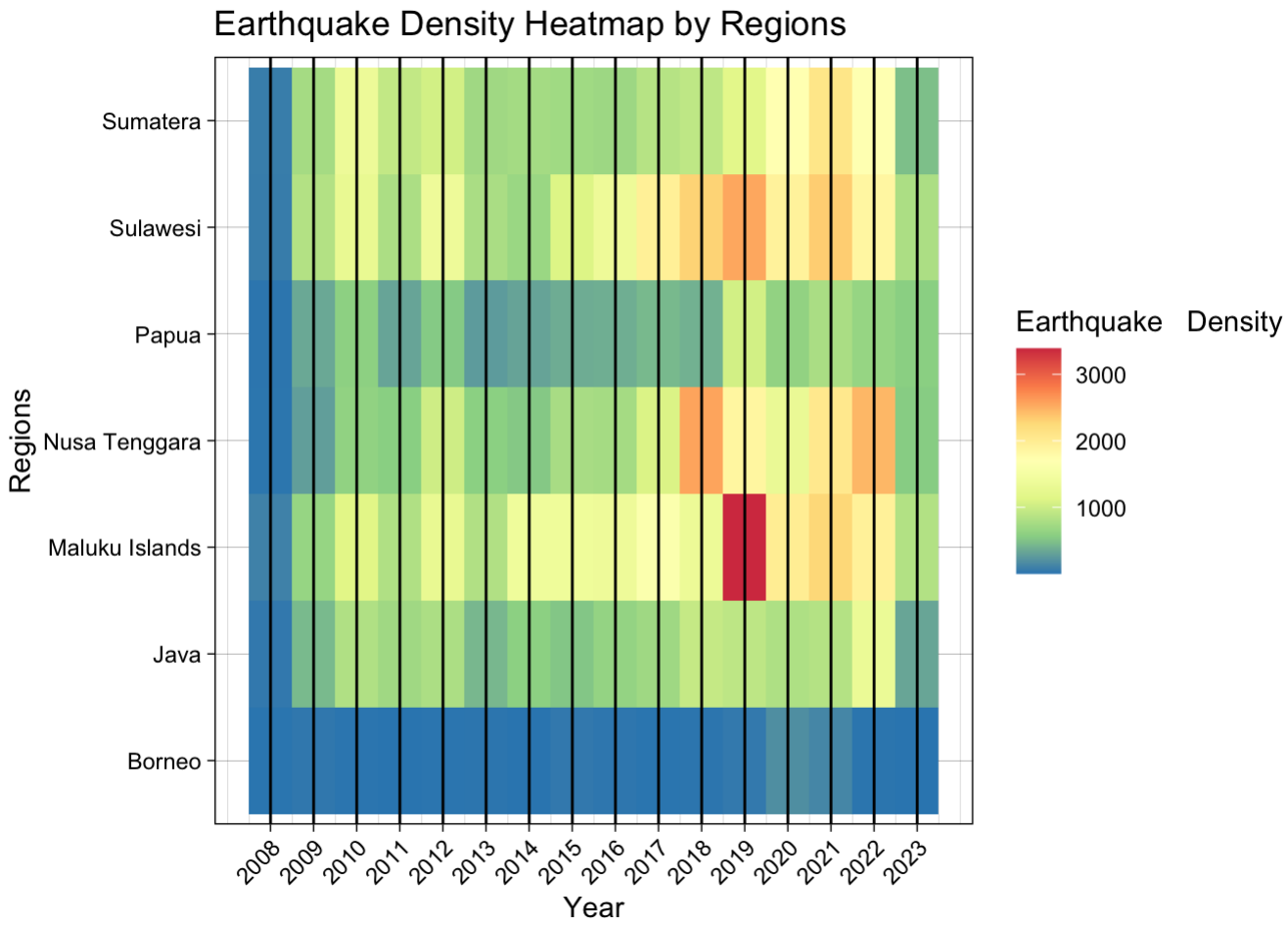
```
## Warning: Removed 484 rows containing missing values or values outside the scale ra
nge
## (`position_stack()`).
```




```
earthquake_heatmap <- earthquake_list_3 %>%
  group_by(regions,year) %>%
  summarise(density = n()) %>%
  ggplot() +
  aes(x = year,y = regions, fill = density) +
  geom_tile()+
  scale_fill_distiller(palette = "Spectral") +  # Example palette
  labs(title = "Earthquake Density Heatmap by Regions", x = "Year", y = "Regions", fi
ll = "Earthquake   Density") +
  theme_linedraw()+
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  scale_x_continuous(breaks = unique(earthquake_list_3$year)) +
  geom_vline(xintercept = unique(earthquake_list_3$year), color = "black", linetype =
"solid")
```

```
## `summarise()` has grouped output by 'regions'. You can override using the
## `.groups` argument.
```

```
earthquake_heatmap
```



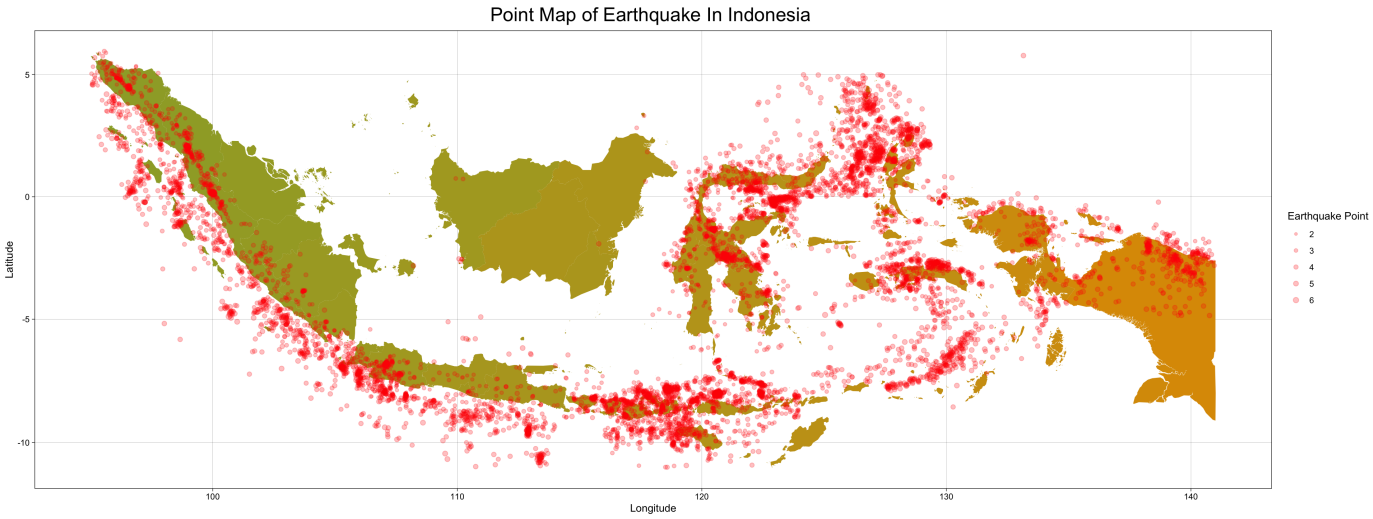

```
earthquake_2022_months <- earthquake_2022 %>%
  mutate(month = month(Date), days = day(Date))
```

```
#volcanoes

volcanoes_indonesia <- read_csv("/Users/julius/Monash/FIT5145/Assignment_3/Volcanoes_
data_Set/volcanoes around the world in 2021.csv")
```

```
## Rows: 1571 Columns: 11
## ── Column specification ────────────────────────────────────────────────────────
## Delimiter: ","
## chr (8): Region, Number, Volcano Name, Country, Location, Type, Status, Last...
## dbl (3): Latitude, Longitude, Elevation (m)
##
## ℹ Use `spec()` to retrieve the full column specification for this data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
#analysis

earthquake_last_5_year <- earthquake_list_3 %>%
    filter(year >= 2017 & year <= 2022)

#adding point map
indonesia_map_plot_2<- indonesia_map_plot +
  geom_point(
    data = earthquake_2022,
    aes(x = longitude, y = latitude, size = magnitude),
    color = "red",
    alpha = 0.25,
  ) +
  scale_size_continuous(range = c(0.1, 2.5))+
  labs(title = "Point Map of Earthquake In Indonesia ", x = "Longitude", y = "Latitud
e", size = "Earthquake Point") +
   theme_linedraw() +
  theme(
    plot.title = element_text(hjust = 0.5, size = 20))


indonesia_map_plot_2
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_point()`).
```



```
volcanoes_indonesia <- volcanoes_indonesia %>%
  filter(Country == "Indonesia")

volcanoes_map <- indonesia_map_plot_2 +
  geom_point(data = volcanoes_indonesia, aes(x = Longitude, y = Latitude, color = "Vo
lcanoes"), size = 3, shape = 17)+
  scale_color_manual(values = "#330000") +
  guides(color = guide_legend(title = "Volcanoes"))

volcanoes_map
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_point()`).
```