## In this document, type only in the boxes provided (which you may enlarge if necessary) and only use the courier font

(this is what the courier font looks like)

## Dataset Code and Spreadsheet Used

|  |  |
| --- | --- |
| Code (beginning with ds..) | Spreadsheet (csv/xlsx/xls) |
| Ds152 | Earthquakes\_v3.csv |

## Question 1

What is the dataset about?

(Response limit: Max 30 words but ideally less)

|  |
| --- |
| This dataset contains records of earthquakes that occurred from 1965 to 1966, including details such as latitude, longitude, depth and magnitude for each event. |

## Question 2

Produce the first 5 rows of the dataset. Only include the maximum 5 most interesting columns and rename them if necessary. Do not add any formatting (spaces etc) unless produced by R.   
(Only use the courier font for this and make sure the columns align. If you rename the columns to make them shorter, underneath the 5 rows list the renamings you have done e.g  
\*acc renamed from “Acceration of the cars”)

|  |
| --- |
| selected\_cols <- c("DATETIME",  "LAT",  "LONG",  "DEPTH",  "MAGNITUDE")  df\_subset <- df[1:5, selected\_cols]  print(as\_tibble(df\_subset))  # A tibble: 5 × 5  DATETIME LAT LONG DEPTH MAGNITUDE  *<chr>* *<dbl>* *<dbl>* *<dbl>* *<dbl>*  1 1/7/1965 10:22 36.5 26.5 10 5.3  2 1/10/1965 8:02 39.2 22.2 10 4.9  3 1/12/1965 17:26 37 22 10 4  4 1/15/1965 14:56 36.8 21.8 10 4.5  5 3/9/1965 19:16 39 24 10 4.2 |

## Question 3

What does each line of the dataset represent?

(Response limit: Max 30 words)

|  |
| --- |
| Each line in the dataset represents the date and time the earthquake had occurred, and it also includes latitude, longitude, depth and magnitude. |

## Question 4

Using basic descriptive statistics, summarize the dataset and highlight some salient features about it on initial viewing.

(Response limit: Max 150 words)

|  |
| --- |
| The dataset gives us insights into earthquakes that occurred between 1965 and 1966. It shows that earthquake magnitudes varied widely, ranging from minor tremors below 4.0 to powerful quakes over 7.0. This highlights the different levels of intensity these events can have. When it comes to depth, the earthquakes also show a lot of differences, with some happening close to the surface and others occurring deep underground. On average, most of these earthquakes had moderate magnitudes around 5.0 and were located at depths of less than 100 km. Looking at the geographic coordinates—latitude and longitude—we can see that these quakes are spread across the globe, with a notable concentration near tectonic plate boundaries like the Pacific Ring of Fire. |

**Question 5**

What is the most surprising thing you can see when first looking at the data?

(Response limit: Max 50 words)

|  |
| --- |
| The most surprising feature is the recorded minimum depth of 0 km. Many people assume that earthquakes occur several kilometers underground, but this data includes events happening right at the surface. This could suggest extremely shallow tremors or reflect a reporting convention for very minor events. |

## Question 6

Draw a graph using R base graphics showing something interesting about your dataset, and explain it in a single sentence afterwards.

(Put image here with explanation below – max text 30 words)

|  |
| --- |
| A graph of a earthquake  AI-generated content may be incorrect. |

|  |
| --- |
| Most earthquakes that are recorded tend to be quite minor, with the majority registering below a 3.0 on the Richter scale. These smaller quakes are so common that we often don’t notice them, but they provide valuable information about the earth's activity beneath our feet. |

## Question 7

Reproduce the R Code you used to create the graph. Only use graphical techniques taught in the course. Do not use any external libraries. The code you provide should be the same as in the R file you supplied and capable of being copied and pasted to an R prompt to generate a graph

|  |
| --- |
| mean\_magnitude <- mean(Earthquakes\_data$MAGNITUDE, na.rm = TRUE)  mag\_bins <- seq(0, 7, by = 0.5)  hist(Earthquakes\_data$MAGNITUDE,  breaks = mag\_bins,  main = "Distribution of Earthquake Magnitudes",  xlab = "Magnitude (Richter Scale)",  ylab = "Frequency",  col = "lightblue",  border = "black") |

## Question 8

What interesting conclusions do you think it might be possible to make from a more in depth view of this data, and what real-world implications might it have?

(Response limit: Max 100 words)

|  |
| --- |
| A more in-depth analysis could uncover significant patterns related to earthquake frequency, clusters of magnitude, and correlations with depth. This exploration could help pinpoint areas at higher risk, providing essential insights for disaster preparedness, urban development, and the implementation of early-warning systems. Such proactive measures could greatly reduce the potential for loss of life and damage to property. Moreover, recognizing these trends could enhance the ability of geologists to forecast seismic activity with greater accuracy. |

## Question 9

Provide a short anecdote ideally from your own experience or if not that, relating to something in your country of origin, which has commonalities with some of the issues and concepts visible in the dataset (max 100 words)

|  |
| --- |
| In my hometown in India, we occasionally experience minor tremors, which often remind us of the devastating 2001 Gujarat earthquake. In response, local communities engage in safety drills and make efforts to reinforce buildings, emphasizing the importance of understanding seismic patterns. This community resilience reflects the dataset's insights into variations in earthquake magnitude and frequency, highlighting the need for continued awareness and preparedness. |

## Question 10

Include maximum 2 citations (for things cited earlier)

|  |
| --- |
| [1]. U.S. Geological Survey. (n.d.). Earthquake Data and Statistics. Retrieved from https://www.usgs.gov Kanamori, H. (1977).  [2]. The energy release in major earthquakes. \*Journal of Geophysical Research\*, 82(20), 2981–2987. |