Assignment 3: Volcanic Eruptions

MinJae Jo

2025-09-11

```
eruptions <- read_csv("eruptions_dataset.csv")
```

```
## Rows: 103 Columns: 16
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (4): name, location, country, type
## dbl (12): year, month, day, latitude, longitude, elevation, VEI, deaths, mis...
##
## 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.
```

Exercise 1

CSV (Comma-Separated Values) is a simple text format that stores data separated by commas. Each row represents one record (observation), and each column represents a variable (attribute). The file size is small and it is easy to read and write in various tools such as Excel, R, and Python, so it is highly compatible. However, if commas are included in the value, attention should be paid to quotation processing, encoding differences, and line modulation problems.

Exercise 2

i. 103 rows and 16 columns

ii. Each rows shown volcanic eruption event that records of one eruption on a particular volcano on a particular year, month, or day.

iii. Elevation is recorded in meters, so it is a metric system.

Exercise 3

i.

```
library(dplyr)

eruptions %>%
  select(name, elevation)
```

```
## # A tibble: 103 x 2
##    name             elevation
##    <chr>                <dbl>
##  1 Tungurahua            5023
##  2 Eyjafjallajokull      1651
##  3 Pacaya                2569
##  4 Zealandia Bank           0
##  5 Karangetang           1797
##  6 Sinabung              2460
##  7 Merapi                2910
##  8 Tungurahua            5023
##  9 Tengger Caldera       2329
## 10 Merapi                2910
## # i 93 more rows
```

ii. eruptions %>% select(name:elevation) 'name:elevation' selects all columns from 'name' to 'elevation' based on the order of the columns. Therefore, a total of six columns are output: 'name, location, country, latitude, longitude, elevation'.

iii. eruptions_stored <- eruptions %>% select(name, elevation)

glimpse(eruptions_stored)

Exercise 4

Exercise 5

Exercise 6

Exercise 7

Exercise 8

Exercise 9