Chapter 2 Preparing Data

Mohammad Saqib Ansari

2023-11-29

Preparing data

```
# Creating vectors for different attributes
surname <- c("Tony", "James", "John")
height <- c(184, 175, 158)
weight <- c(80, 78, 72)
size <- c(9.5, 8.5, 8)
sex <- c("M", "M", "M")

# Creating a data frame 'my.data' with the vectors
my.data <- data.frame(surname, height, weight, size, sex)

# Writing 'my.data' to a CSV file named 'my.data.csv'
write.csv(my.data, file = "my.data.csv")
```

The above R code is used to prepare sample data and perform the following steps:

1. **Creating Data:**
   - Vectors for 'surname', 'height', 'weight', 'size', and 'sex' are created to represent different attributes.
   - These vectors are then combined into a data frame named 'my.data' using the `data.frame()` function in R.
2. **Writing Data to CSV:**
   - The `write.csv()` function is used to write the 'my.data' data frame to a CSV file named 'my.data.csv'. This CSV file will contain the structured data in a comma-separated format.

```
# Reading the CSV file 'my.data.csv' into R as a data frame 'tab'
tab <- read.table("my.data.csv", sep = ",", header = TRUE, dec = ".", row.names = 1)
```

The code snippet above demonstrates reading the previously saved CSV file 'my.data.csv' back into R as a data frame named 'tab'. The `read.table()` function is used to read the CSV file, specifying parameters such as the separator (`,`), header existence, decimal point representation (`"."`), and using the first column as row names (`row.names = 1`).

This R Markdown document provides code explanations and code chunks to execute the data preparation steps and file operations in R, aiding in understanding the process of creating a data frame, saving it to a CSV file, and subsequently reading it back into R for further analysis. "'