RWorksheet_TAN#4

John Kenneth D. Tan

2022-11-23

1. The table below shows the data about shoe size and height. Create a data frame.

```
table_data <- data.frame(
  Shoesize = c(6.5, 9.0, 8.5, 8.5, 10.5, 7.0, 9.5, 9.0,
               13.0, 7.5, 10.5, 8.5, 12.0, 10.5, 13.0,
               11.5, 8.5, 5.0, 10.0, 6.5 ,7.5, 8.5, 10.5,
               8.5, 10.5, 11.0, 9.0, 13.0),

  Height = c(66.0, 68.0, 64.0, 65.0, 70.0, 64.0, 70.0, 71.0,
             72.0, 64.0, 74.0, 67.0, 71.0, 71.0, 77.0, 72.0,
             59.0, 62.0, 72.0, 66.0, 64.0, 67.0, 73.0, 69.0,
             72.0, 70.0, 69.0, 70.0),

  Gender = c("F","F","F","F","M","F","F","F","M","F","M","F",
             "M","M","M","M","F","F","M","F","F", "M","M","F",
             "M","M","M","M"))

names(table_data)<-list("Shoe size", "Height", "Gender")

table_data
```

```
##    Shoe size Height Gender
## 1        6.5     66      F
## 2        9.0     68      F
## 3        8.5     64      F
## 4        8.5     65      F
## 5       10.5     70      M
## 6        7.0     64      F
## 7        9.5     70      F
## 8        9.0     71      F
## 9       13.0     72      M
## 10       7.5     64      F
## 11      10.5     74      M
## 12       8.5     67      F
## 13      12.0     71      M
## 14      10.5     71      M
## 15      13.0     77      M
## 16      11.5     72      M
## 17       8.5     59      F
## 18       5.0     62      F
## 19      10.0     72      M
## 20       6.5     66      F
```

```
## 21       7.5     64      F
## 22       8.5     67      M
## 23      10.5     73      M
## 24       8.5     69      F
## 25      10.5     72      M
## 26      11.0     70      M
## 27       9.0     69      M
## 28      13.0     70      M
```

a. Describe the data.

Answer: The data is composed of three columns and 28 rows that contains the Shoe size, Height and Gender.

b. Find the mean of shoe size and height of the respondents. Copy the codes and results

```
Shoesize <- c(6.5, 9.0, 8.5, 8.5, 10.5, 7.0, 9.5, 9.0, 13.0, 7.5, 10.5, 8.5, 12.0, 10.5, 13.0,
              11.5, 8.5, 5.0, 10.0, 6.5 ,7.5, 8.5, 10.5 ,8.5, 10.5, 11.0, 9.0, 13.0)
Shoesize
```

```
##  [1]  6.5  9.0  8.5  8.5 10.5  7.0  9.5  9.0 13.0  7.5 10.5  8.5 12.0 10.5 13.0
## [16] 11.5  8.5  5.0 10.0  6.5  7.5  8.5 10.5  8.5 10.5 11.0  9.0 13.0
```

```
mean_Shoesize <- mean(Shoesize)
mean_Shoesize
```

```
## [1] 9.410714
```

```
Height <- c(66.0, 68.0, 64.0, 65.0, 70.0, 64.0, 70.0, 71.0, 72.0, 64.0, 74.0, 67.0, 71.0, 71.0,
            77.0, 72.0, 59.0, 62.0, 72.0, 66.0, 64.0, 67.0, 73.0, 69.0, 72.0, 70.0, 69.0, 70.0)
Height
```

```
##  [1] 66 68 64 65 70 64 70 71 72 64 74 67 71 71 77 72 59 62 72 66 64 67 73 69 72
## [26] 70 69 70
```

```
mean_height <- mean(Height)
mean_height
```

```
## [1] 68.53571
```

c. Is there a relationship between shoe size and height? Why?

Yes, when the respondents are tall, their shoe sizes are also big. Whereas, if their height is below 70.0, their shoe size are smaller.

2. Construct character vector months to a factor with factor() and assign the result to factor_months_vector. Print out factor_months_vector and assert that R prints out the factor levels below the actual values.

```
months_vector <- c("March","April","January","November","January",
"September","October","September","November","August",
"January","November","November","February","May","August",
"July","December","August","August","September","November","February","April")

factor_months_vector <- factor(months_vector)
factor_months_vector
```

```
##  [1] March     April     January   November  January   September October
##  [8] September November  August    January   November  November  February
## [15] May       August    July      December  August    August    September
## [22] November  February  April
## 11 Levels: April August December February January July March May ... September
```

3. Then check the summary() of the months_vector and factor_months_vector. #Interpret the results of both vectors. Are they both equally useful in this case?

```
  summary(months_vector)
```

```
##    Length     Class      Mode
##        24 character character
```

```
  summary(factor_months_vector)
```

```
##     April    August  December  February   January      July     March       May
##         2         4         1         2         3         1         1         1
##  November   October September
##         5         1         3
```

Answer: Yes, it is beneficial because it aids in the efficient understanding of data. In months_vector, you do not need to count the number of words because of the function length. Its mode and class are also displayed. You don't have also to manually count the number of times a specific month appears in factor_months_vector because it's already computed.

4. Create a vector and factor for the table below.

```
  factor_data <- c("East" = '1', "West" = '4', "North" = '3')
  factor_data
```

```
##  East  West North
##   "1"   "4"   "3"
```

5. Enter the data below in Excel with file name = import_march.csv

   a. Import the excel file into the Environment Pane using read.table() function. Write the code.

```
setwd("C:/Users/Kenneth/Desktop/RProg_Worksheets/Worksheet  4/")

import_table<-read.table("import_march.csv", header=TRUE, sep=",")
import_table
```

```
##   Students Strategy.1 Strategy.2 Strategy.3
## 1                    8         10          8
## 2     Male           4          8          6
## 3                    0          6          4
## 4                   14          4         15
## 5   Female          10          2         12
## 6                    6          0          9
```

b. View the dataset. Write the code and its result.

```
view_data<- read.csv("import_march.csv")
view_data
```

```
##   Students Strategy.1 Strategy.2 Strategy.3
## 1                    8         10          8
## 2     Male           4          8          6
## 3                    0          6          4
## 4                   14          4         15
## 5   Female          10          2         12
## 6                    6          0          9
```