Worksheet 4

LG Grace C. Sabio

2022-11-22

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

```
shoe_size <- 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, 64.5, 65.0, 70.0, 64.0, 70.0, 71.0, 72.0,
                  64.0, 74.5, 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')

)
shoe_size
```

```
##    shoeSize height Gender
## 1       6.5   66.0      F
## 2       9.0   68.0      F
## 3       8.5   64.5      F
## 4       8.5   65.0      F
## 5      10.5   70.0      M
## 6       7.0   64.0      F
## 7       9.5   70.0      F
## 8       9.0   71.0      F
## 9      13.0   72.0      M
## 10      7.5   64.0      F
## 11     10.5   74.5      M
## 12      8.5   67.0      F
## 13     12.0   71.0      M
## 14     10.5   71.0      M
## 15     13.0   77.0      M
## 16     11.5   72.0      M
## 17      8.5   59.0      F
## 18      5.0   62.0      F
## 19     10.0   72.0      M
## 20      6.5   66.0      F
## 21      7.5   64.0      F
## 22      8.5   67.0      M
## 23     10.5   73.0      M
## 24      8.5   69.0      F
```

```
## 25     10.5   72.0      M
## 26     11.0   70.0      M
## 27      9.0   69.0      M
## 28     13.0   70.0      M
```

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

```
##    Shoe size Height Gender
## 1        6.5   66.0      F
## 2        9.0   68.0      F
## 3        8.5   64.5      F
## 4        8.5   65.0      F
## 5       10.5   70.0      M
## 6        7.0   64.0      F
## 7        9.5   70.0      F
## 8        9.0   71.0      F
## 9       13.0   72.0      M
## 10       7.5   64.0      F
## 11      10.5   74.5      M
## 12       8.5   67.0      F
## 13      12.0   71.0      M
## 14      10.5   71.0      M
## 15      13.0   77.0      M
## 16      11.5   72.0      M
## 17       8.5   59.0      F
## 18       5.0   62.0      F
## 19      10.0   72.0      M
## 20       6.5   66.0      F
## 21       7.5   64.0      F
## 22       8.5   67.0      M
## 23      10.5   73.0      M
## 24       8.5   69.0      F
## 25      10.5   72.0      M
## 26      11.0   70.0      M
## 27       9.0   69.0      M
## 28      13.0   70.0      M
```

a. Describe the data

```
# The table shows three columns such as "Shoe size", "Height", and "Gender".
#I noticed that the taller the person is, the bigger the shoe size is.
#Mostly males had bigger shoe size and taller compare to females.
```

b. Find the mean of shoe size and height of the respondents.

```
sizeNheight <- subset(shoe_size[ 1:2])
sizeNheight
```

```
##    Shoe size Height
## 1        6.5   66.0
```

```
## 2        9.0   68.0
## 3        8.5   64.5
## 4        8.5   65.0
## 5       10.5   70.0
## 6        7.0   64.0
## 7        9.5   70.0
## 8        9.0   71.0
## 9       13.0   72.0
## 10       7.5   64.0
## 11      10.5   74.5
## 12       8.5   67.0
## 13      12.0   71.0
## 14      10.5   71.0
## 15      13.0   77.0
## 16      11.5   72.0
## 17       8.5   59.0
## 18       5.0   62.0
## 19      10.0   72.0
## 20       6.5   66.0
## 21       7.5   64.0
## 22       8.5   67.0
## 23      10.5   73.0
## 24       8.5   69.0
## 25      10.5   72.0
## 26      11.0   70.0
## 27       9.0   69.0
## 28      13.0   70.0
```

```
meanSH <- colMeans(sizeNheight)
meanSH
```

```
## Shoe size    Height
##  9.410714 68.571429
```

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

```
sizeNheight <- subset(shoe_size[c( 1, 2)])
meanSH <- colMeans(sizeNheight)
mean(meanSH)
```

```
## [1] 38.99107
```

```
 # Yes. because I think the taller a person is, the bigger the shoe size he has.
```

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. |

```
summary_months <- summary(months_vector)
summary_months
```

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

```
summary_factor_months <- summary(factor_months_vector)
summary_factor_months
```

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

```
#
```

Interpret the results of both vectors. Are they both equally useful in this case?

```
# The summary() of the months_vector displays the length, class, and mode of the vector
#months_vector while the summary() of the factor_months_vector displays the different months
#of the year and their corresponding frequency or how many times it was mentioned in the vector.
# They are both equally useful in this case.
```

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

```
factor_data <- c("East", "West", "North")
factor_data
```

```
## [1] "East"  "West"  "North"
```

```
new_order_data <- factor(factor_data,levels = c("East","West","North"))
print(new_order_data)
```

```
## [1] East  West  North
## Levels: East West North
```

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.

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

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

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

```r
# View(import_march_csv)
```