Worksheet-4 in R

Drake Francis M. Jaculina

2022-11-23

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

```
shoeDF <- 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),
  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),
  Gender= c("F", "F", "F", "F", "M", "F", "F", "F", "M", "F", "M", "F", "M", "M"),
  Shoesize2= c(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),
  Height2= c(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),
  Gender2= c("M", "M", "F", "F", "M", "F", "F", "M", "M", "F", "M", "M", "M", "M")
  )
shoeDF
```

```
##    Shoesize Height Gender Shoesize2 Height2 Gender2
## 1       6.5   66.0      F      13.0      77       M
## 2       9.0   68.0      F      11.5      72       M
## 3       8.5   64.5      F       8.5      59       F
## 4       8.5   65.0      F       5.0      62       F
## 5      10.5   70.0      M      10.0      72       M
## 6       7.0   64.0      F       6.5      66       F
## 7       9.5   70.0      F       7.5      64       F
## 8       9.0   71.0      F       8.5      67       M
## 9      13.0   72.0      M      10.5      73       M
## 10      7.5   64.0      F       8.5      69       F
## 11     10.5   74.5      M      10.5      72       M
## 12      8.5   67.0      F      11.0      70       M
## 13     12.0   71.0      M       9.0      69       M
## 14     10.5   71.0      M      13.0      70       M
```

a. Describe the data.

The data shows the different shoe size and height of male and female. The data is likely have a linear correlation.

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

```
mean1 <- shoeDF$Shoesize
mean(mean1)
```

```
## [1] 9.321429
```

```
mean2 <-shoeDF$Shoesize2
mean(mean2)
```

```
## [1] 9.5
```

```
both1 <- mean(c(mean1,mean2))
both1
```

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

```
mean3 <-shoeDF$Height
mean(mean3)
```

```
## [1] 68.42857
```

```
mean4 <-shoeDF$Height2
mean(mean4)
```

```
## [1] 68.71429
```

```
both2 <- mean(c(mean3,mean4))
both2
```

```
## [1] 68.57143
```

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

Yes, there is a relationship or correlation between shoe size and height. It is because as you can see in the given data, taller male or female tend to have a larger shoe size while shorter male or female tend to have a smaller shoe size.

Factors

A nominal variable is a categorical variable without an implied order. This means that it is impossible to say that 'one is worth more than the other'. In contrast, ordinal variables do have a natural ordering.

Example

```
Gender <- c("M","F","F","M")
factor_Gender <- factor(Gender)
factor_Gender
```

```
## [1] M F F M
## Levels: F M
```

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.

Consider data consisting of the names of months: "March","April","January","November","January", "September","October","September","November","August", "January","November","November","February","May","August", "July","December","August","August","September","November","February","April")

```
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_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
```

Interpret the results of both vectors.

For the summary of months_vector, it shows the total Length, Class, and the Mode while for the summary of factor_months_vector, it specify directly the months, the sequence or arrangement was still the same, and the number from which it's used repeatedly.

Are they both equally useful in this case?

Yes, their are both equally useful in this case.

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

Note: Apply the factor function with required order of the level. new_order_data <- factor(factor_data,levels = c("East","West","North")) print(new_order_data)

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

```
##  East  West North
##  <NA>  <NA>  <NA>
## 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.

```
import <- read.table("/cloud/project/RWorksheet_JACULINA#4/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.

```
view_dataset <- read.csv("/cloud/project/RWorksheet_JACULINA#4/import_march.csv")
view_dataset
```

```
##   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
```