Worksheet-4 in R

Worksheet for R Programming

Instructions:

- Use RStudio or the RStudio Cloud accomplish this worksheet.

- Save the R script as *RWorksheet_lastname#4.R.*
- On your own *GitHub repository*, push the R script, the Rmd file, as well as this pdf worksheet to the repo you have created before.
- Do not forget to comment your Git repo on our VLE
- Accomplish this worksheet by answering the questions being asked and writing the code manually.

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

Shoe size	Height	Gender	Shoe size	Height	Gender
6.5	66.0	F	13.0	77.0	M
9.0	68.0	F	11.5	72.0	M
8.5	64.5	F	8.5	59.0	F
8.5	65.0	F	5.0	62.0	F
10.5	70.0	M	10.0	72.0	M
7.0	64.0	F	6.5	66.0	F
9.5	70.0	F	7.5	64.0	F
9.0	71.0	F	8.5	67.0	M
13.0	72.0	M	10.5	73.0	M
7.5	64.0	F	8.5	69.0	F
10.5	74.5	M	10.5	72.0	M
8.5	67.0	F	11.0	70.0	M
12.0	71.0	M	9.0	69.0	M
10.5	71.0	M	13.0	70.0	M

a. Describe the data.

```
df <- 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),

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

df
```

```
> df
   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      M
8       9.0   71.0      F
9      13.0   72.0      M
10      7.5   64.0      M
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      M
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
```

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

summary(df)

```
> summary(df)
   Shoesize         Height         Gender
 Min.   : 5.000   Min.   :59.00   Length:28
 1st Qu.: 8.500   1st Qu.:65.75   Class :character
 Median : 9.000   Median :69.50   Mode  :character
 Mean   : 9.411   Mean   :68.57
 3rd Qu.:10.500   3rd Qu.:71.25
 Max.   :13.000   Max.   :77.00
```

#SHOESIZE: Mean : 9.411
HEIGHT: Mean :68.57

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

Yes, if the height is heigher then the shoesize is greater.

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")

```
#2
Months <- 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")
Months
```

```
> Months
 [1] "March"     "April"     "January"   "November"  "January"   "September" "October"
"September" "November"
[10] "August"    "January"   "November"  "November"  "February"  "May"       "August"
"July"      "December"
```

```
[19] "August"    "August"    "September" "November"  "February"  "April"


factor_Months <- factor(Months)
factor_Months

 factor_Months
 [1] March     April     January   November  January   September October   September
November  August    January
[12] November  November  February  May       August    July      December  August
August    September November
[23] February  April
Levels: April August December February January July March May November October
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)

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

summary(factor_Months)

```
> summary(factor_Months)
    April    August  December  February   January      July     March       May  November   October September
        2         4         1         2         3         1         1         1         5         1         3
>
```

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

Direction	Frequency
East	1
West	4
North	3

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(1,4,3)
> factor_data
[1] 1 4 3
>
> n_order_data <- factor(factor_data,levels = c("East","West","North"))
> print(n_order_data)
[1] <NA> <NA> <NA>
Levels: East West North
```

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

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

Figure 1: Excel File

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

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