RWorksheet_Calambro#4a

Christian Paul Calambro

2024-10-02

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

```
library(readxl)
Household <- read_excel("HouseHold_Data.xlsx")

Household
```

```
## # A tibble: 28 x 3
##    Shoe_size Height Gender
##        <dbl>  <dbl> <chr>
##  1       6.5   66   F
##  2       9     68   F
##  3       8.5   64.5 F
##  4       8.5   65   F
##  5      10.5   70   M
##  6       7     64   F
##  7       9.5   70   F
##  8       9     71   F
##  9      13     72   M
## 10       7.5   64   F
## # i 18 more rows
```

a. Describe the data

The data set includes 25 people with information on their shoe size, height, and gender. Shoe sizes range from 5.0 to 13.0, and heights range from 59.0 to 77.0 inches. Males generally have larger shoe sizes and are taller than females. Taller individuals tend to have bigger shoe sizes.

b. Create a subset by males and females with their corresponding shoe size and height. What its result? Show the R scripts.

```
males <- subset(Household, Gender == "M", select = c(Shoe_size, Height))
females <- subset(Household, Gender == "F", select = c(Shoe_size, Height))
males
```

```
## # A tibble: 14 x 2
##    Shoe_size Height
##        <dbl>  <dbl>
##  1      10.5   70
##  2      13     72
##  3      10.5   74.5
##  4      12     71
##  5      10.5   71
##  6      13     77
##  7      11.5   72
##  8      10     72
```

```
##  9       8.5   67
## 10      10.5   73
## 11      10.5   73
## 12      11     70
## 13       9     69
## 14      13     70
```

```
females
```

```
## # A tibble: 14 x 2
##    Shoe_size Height
##        <dbl>  <dbl>
##  1       6.5   66
##  2       9     68
##  3       8.5   64.5
##  4       8.5   65
##  5       7     64
##  6       9.5   70
##  7       9     71
##  8       7.5   64
##  9       8.5   67
## 10       8.5   59
## 11       5     62
## 12       6.5   66
## 13       7.5   64
## 14       8.5   69
```

c. Find the mean of shoe size and height of the respondents. Write the R scripts and its result.

```
meanofshoes <- mean(Household$Shoe_size)
meanofheight <- mean(Household$Height)

paste("The Mean of Shoes:", meanofshoes)
```

```
## [1] "The Mean of Shoes: 9.41071428571429"
```

```
paste("The Mean of Height:", meanofheight)
```

```
## [1] "The Mean of Height: 68.6071428571429"
```

d. Is there a relationship between shoe size and height? Why? Yes, there is a relationship between the shoe size and the height.The taller individuals tend to have larger shoe sizes.

FACTORS

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 <- 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 <- factor(Months)
Factor_Months
```

```
##  [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
```

```
levels(Months)
```

```
## NULL
```

3. Then check the summary() of the months_vector and factor_months_vector. | Inter- pret the results of both vectors. Are they both equally useful in this case?

```
("The Summary of Months:")
```

```
## [1] "The Summary of Months:"
```

```
summary(Months)
```

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

```
("The Summary of Factor_Months:")
```

```
## [1] "The Summary of Factor_Months:"
```

```
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
```

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

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

Frequency <- c(1, 4, 3)

new_order_data <- factor(Directions, levels = c("East", "West", "North"))

("Directions Factor with Specified Order:")
```

```
## [1] "Directions Factor with Specified Order:"
```

```
new_order_data
```

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

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

```
library(readxl)
data <- read_excel("import_march.csv")

data
```

```
## # A tibble: 6 x 4
##   Students Strategy1 Strategy2 Strategy3
##   <chr>        <dbl>     <dbl>     <dbl>
## 1 Male             8        10         8
## 2 Male             4         8         6
## 3 Male             0         6         4
```

```
## 4 Female          14         4        15
## 5 Female          10         2        12
## 6 Female           6         0         9
```