RWorksheet_Benedicto#4a.R

Gabriel R. Benedicto

2024-10-14

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

```
shoe_size <- c(6.5, 9.0, 8.5, 8.5, 7.0, 9.0, 9.5, 13.0, 7.5, 10.5, 10.5, 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, 65.0, 65.0, 64.0, 71.0, 72.0, 72.0, 74.5, 67.0, 74.5, 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", "F", "F", "F", "M", "F", "M", "M", "M", "M",
            "M", "M", "F", "F", "M", "F", "F", "F", "M", "F", "M", "M", "M", "M")

data <- data.frame(Shoe_size = shoe_size, Height = height, Gender = gender)

data
```

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

a. Describe the data. The table presents data about individuals' shoe size, height, and gender. The variables are:

-Shoe size: Numeric, represents the size of the shoe. -Height: Numeric, represents the height of the individual (presumably in inches). -Gender: Categorical, either "M" for male or "F" for female.

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

- The result is that it shows the Shoe-size and height of every Male and Female in the data.

```
male_data <- subset(data, Gender == "M", select = c(Shoe_size, Height))

male_data
```

```
##    Shoe_size Height
## 8       13.0   72.0
## 10      10.5   67.0
## 11      10.5   74.5
## 12      12.0   71.0
## 13      10.5   71.0
## 14      13.0   77.0
## 15      11.5   72.0
## 18      10.0   72.0
## 22      10.5   73.0
## 24      10.5   72.0
## 25      11.0   70.0
## 26       9.0   69.0
## 27      13.0   70.0
```

```
female_data <- subset(data, Gender == "F", select = c(Shoe_size, Height))

female_data
```

```
##    Shoe_size Height
## 1        6.5   66.0
## 2        9.0   68.0
## 3        8.5   65.0
## 4        8.5   65.0
## 5        7.0   64.0
## 6        9.0   71.0
## 7        9.5   72.0
## 9        7.5   74.5
## 16       8.5   59.0
## 17       5.0   62.0
## 19       6.5   66.0
## 20       7.5   64.0
## 21       8.5   67.0
## 23       8.5   69.0
```

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

```
mean_shoe_size <- mean(data$Shoe_size)

mean_shoe_size
```

```
## [1] 9.444444
```

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

mean_height
```

```
## [1] 69
```

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

- Yes. for me from personal experience, as the shoesize increase, the height also increases.

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)

print(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
```

```
levels(factor_months_vector)
```

```
##  [1] "April"     "August"    "December"  "February"  "January"   "July"
##  [7] "March"     "May"       "November"  "October"   "September"
```

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 result of months_vector is that is states the Length, Class and Mode. While the factor_months_vector states how many months in the data for example April, April has 2. -I think they are both useful because it is easy for me to understand and determine how many types of data from the raw data itself.

```
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
```

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

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

factor_data <- factor(direction_vector)

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

```
## [1] East  West  North West  North West  North West
## 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.
b. View the dataset. Write the R scripts and its result.

```
list.files()
```

```
## [1] "import_march.csv"            "RWorksheet_Benedicto-4a.pdf"
## [3] "RWorksheet_Benedicto-4a.Rmd" "RWorksheet_Benedicto#4a.Rmd"
```

```
list.files("/cloud/project/")
```

```
## [1] "project.Rproj"            "RWorksheet_Benedicto#2R."
## [3] "RWorksheet_Benedicto#3R." "RWorksheet_Benedicto4R."
## [5] "Worksheet1"
```

```
getwd()
```

```
## [1] "/cloud/project/RWorksheet_Benedicto4R."
```

```
setwd("/cloud/project/RWorksheet_Benedicto4R.")
data_excel <- read.table("import_march.csv")
data_excel
```

```
##                                     V1
## 1 Students,Strategy1,Strategy2,Strategy3
## 2                            Male,8,10,8
## 3                                 ,4,8,6
## 4                                 ,0,6,4
## 5                          Female,14,4,15
## 6                                ,10,2,12
## 7                                 ,6,0,9
```