Worksheet 4a

Rica Marie Benliro

2023-11-01

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

```
shoe_size <- c(6.5,9.0,8.5,8.5,10.5,7.0,9.5,9.0,13.0,7.5,10.5,8.5,12.0,10.5,13.0,11.5,8.5,5.0,10.0,6.5,7
height <- c(66.0,68.0,64.5,65.0,70.0,64.0,70.0,71.0,72.0,64.0,74.0,67.0,71.0,71.0,77.0,72.0,59.0,62.0,72
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"

shoesizdf <- data.frame(shoe_size,height,gender)
shoesizdf
```

```
##    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   74.0      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 "Household Data" table includes three main variables: "Shoe size" and "Height" (numerical in inches
```

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

```r
male <- subset(shoesizdf, gender == "M")
male
```

```
##    shoe_size height gender
## 5       10.5     70      M
## 9       13.0     72      M
## 11      10.5     74      M
## 13      12.0     71      M
## 14      10.5     71      M
## 15      13.0     77      M
## 16      11.5     72      M
## 19      10.0     72      M
## 22       8.5     67      M
## 23      10.5     73      M
## 25      10.5     72      M
## 26      11.0     70      M
## 27       9.0     69      M
## 28      13.0     70      M
```

```r
female <- subset(shoesizdf, gender == "F")
female
```

```
##    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
## 6        7.0   64.0      F
## 7        9.5   70.0      F
## 8        9.0   71.0      F
## 10       7.5   64.0      F
## 12       8.5   67.0      F
## 17       8.5   59.0      F
## 18       5.0   62.0      F
## 20       6.5   66.0      F
## 21       7.5   64.0      F
## 24       8.5   69.0      F
```

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

```r
mean(shoesizdf$shoe_size)
```

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

```r
mean(shoesizdf$height)
```

```
## [1] 68.55357
```

d. Is there a relationship between shoe size and height? Why? '`Yes. The data shows that people with heights of 70 and above have bigger show sizes compared to heights 69 and below.`

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.

```r
months_vector <- c("March", "April","January", "November", "January", "September", "October", "Septembe
months_vector
```

```
##  [1] "March"     "April"     "January"   "November"  "January"   "September"
```

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

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

```r
summary(months_vector)
```

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

```r
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
```

```
The results of the summary of months_vector does not really show any contributtions. However, factor_mo
```

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

```r
direction <- c("East", "West","North")
frequency <- c(1,4,3)
```

```r
direction_factor <- factor(direction, levels = c("East", "West", "North"))
direction_factor
```

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

```r
freq <- factor(frequency, levels = c("1", "4", "3"))
freq
```

```
## [1] 1 4 3
## Levels: 1 4 3
```

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

```r
importmarch <- read.csv(file="import_march.csv")
importmarch
```

```
##   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
```

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

```r
read_importmarch <- read.table(file ='/cloud/project/import_march.csv', header = TRUE, sep = ',')
read_importmarch
```

```
##   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
```

6. Full Search

```r
input <-  as.numeric(readline(prompt = "Enter a number from 1-50:"))
```

```
## Enter a number from 1-50:
```

```r
input <- 49
if(input == 20){
  print("TRUE")
} else if (input > 50){
  print("The number selected is beyond the range of 1 to 50.")
} else{
  print(input)
}
```

```
## [1] 49
```

7. Change

```r
minBills <- function(price) {

#user input
price<- 100

  if (price %% 50 != 0) {
    return("The price must be a multiple of 50.")
  }

  bills <- c(500, 100, 50, 200, 1000)

  numBills <- 0

  for (i in 1:length(bills)) {
    while (price >= bills[i]) {
      price <- price - bills[i]
      numBills <- numBills + 1
    }
  }
  return(numBills)
}


price <- as.integer(readline(prompt="Enter the price of the snack (in pesos, divisible by 50): "))
```

```
## Enter the price of the snack (in pesos, divisible by 50):
```

```r
print(paste("The minimum number of bills needed to pay for a", price, "peso snack is", minBills(price),
```

```
## [1] "The minimum number of bills needed to pay for a NA peso snack is 1 ."
```

8. The following is each student's math score for one semester. Based on this, answer the following questions.

a. Create a dataframe from the above table. Write the R codes and its output.

```r
students_df <- data.frame(Name = c("Annie", "Thea", "Steve", "Hanna"),
                          Grade1 = c(85, 65, 65, 75),
                          Grade2 = c(65, 75, 95, 90),
                          Grade3 = c(65, 65, 100, 75),
                          Grade4 = c(75, 55, 90, 80))
students_df
```

```
##    Name Grade1 Grade2 Grade3 Grade4
## 1 Annie     85     65     65     75
## 2  Thea     65     75     65     55
## 3 Steve     65     95    100     90
## 4 Hanna     75     90     75     80
```

b. Without using the rowMean function, output the average score of students whose average math score over 90 points during the semester. write R code and its output.

```r
average_scores <- (students_df$Grade1 + students_df$Grade2 + students_df$Grade3 + students_df$Grade4) /
good_students <- which(average_scores > 90)

for (i in 1:nrow(students_df)) {
 if (i %in% good_students) {
    cat(paste0(students_df$Name[i], "'s score this semester is "), average_scores[i], "\n")
 } else {
    cat(paste0(students_df$Name[i], "'s does not have a score of 90 and above."), "\n")
 }
}
```

```
## Annie's does not have a score of 90 and above.
## Thea's does not have a score of 90 and above.
## Steve's does not have a score of 90 and above.
## Hanna's does not have a score of 90 and above.
```

c. Without using the mean function, output as follows for the tests in which the average score was less than 80 out of 4 tests. Example output: The nth test was difficult.

```r
count <- 0

for (i in 1:nrow(students_df)) {

  average_scores <- sum(students_df[i, -1]) / 4
  if (average_scores < 80) {
    count <- count + 1
    print(paste("The", count, "nth test was difficult for", students_df[i, 1], "."))
  }
}
```

```
## [1] "The 1 nth test was difficult for Annie ."
## [1] "The 2 nth test was difficult for Thea ."
```

d. Without using the max function, output as follows for students whose highest score for a semester exceeds 90 points. Example Output: Annie's highest grade this semester is 95.

```
highest_grade <- apply(students_df[, -1], 1, max)

students_above_90 <- students_df[highest_grade > 90, ]

for (i in 1:nrow(students_above_90)) {
  cat(students_above_90[i, 1], "'s highest grade this semester is", max(students_above_90[i, -1]), ".")
}
```

```
## Steve 's highest grade this semester is 100 .
```