RWorksheet_Octaviano#4

Jirraine Octaviao

2023-10-25

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

```
Gender <- c("F", "F", "F", "F", "M", "F", "F", "F", "M", "F", "M" , "F", "M", "M", "M", "M", "F", "F",

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

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



householdData <- data.frame(
  Gender = Gender,
  ShoeSize = shoeSize,
  Height = Height
)
householdData
```

```
##    Gender ShoeSize Height
## 1       F      6.5   66.0
## 2       F      9.0   68.0
## 3       F      8.5   64.5
## 4       F      8.5   65.0
## 5       M     10.5   70.0
## 6       F      7.0   64.0
## 7       F      9.5   70.0
## 8       F      9.0   71.0
## 9       M     13.0   72.0
## 10      F      7.5   64.0
## 11      M     10.5   74.5
## 12      F      8.5   67.0
## 13      M     12.0   71.0
## 14      M     10.5   71.0
## 15      M     13.0   77.0
## 16      M     11.5   72.0
## 17      F      8.5   59.0
## 18      F      5.0   62.0
## 19      M     10.0   72.0
## 20      F      6.5   66.0
## 21      F      7.5   64.0
## 22      M      8.5   67.0
```

```
## 23      M     10.5   73.0
## 24      F      8.5   69.0
## 25      M     10.5   72.0
## 26      M     11.0   70.0
## 27      M      9.0   69.0
## 28      M     13.0   70.0
```

#a # The data contains 28 unique data entries on individuals, including gender, shoe size, and height.

#b

```
males <- householdData[householdData$Gender == "M",]
males
```

```
##    Gender ShoeSize Height
## 5       M     10.5   70.0
## 9       M     13.0   72.0
## 11      M     10.5   74.5
## 13      M     12.0   71.0
## 14      M     10.5   71.0
## 15      M     13.0   77.0
## 16      M     11.5   72.0
## 19      M     10.0   72.0
## 22      M      8.5   67.0
## 23      M     10.5   73.0
## 25      M     10.5   72.0
## 26      M     11.0   70.0
## 27      M      9.0   69.0
## 28      M     13.0   70.0
```

```
females <- householdData[householdData$Gender == "F",]
females
```

```
##    Gender ShoeSize Height
## 1       F      6.5   66.0
## 2       F      9.0   68.0
## 3       F      8.5   64.5
## 4       F      8.5   65.0
## 6       F      7.0   64.0
## 7       F      9.5   70.0
## 8       F      9.0   71.0
## 10      F      7.5   64.0
## 12      F      8.5   67.0
## 17      F      8.5   59.0
## 18      F      5.0   62.0
## 20      F      6.5   66.0
## 21      F      7.5   64.0
## 24      F      8.5   69.0
```

```
#c
meanOfShoeSize <- mean(householdData$ShoeSize)
meanOfShoeSize
```

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

```
meanOfHeight <- mean(householdData$Height)
meanOfHeight
```

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

#d. # The relationship between the two is that the shoe size is directly proportional to the height. If the height is small, the shoe size is also small.

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","No

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"
```

```
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. | Inter-pret the results of both vectors. Are they both equally useful in this case?

```
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
```

```
#In the summary of months_vector, it shows the number of observations, class, and mode of the vector.
#In the summary of factor_months_vector, it shows the frequency of each month.
#Both are useful in different cases where the number of observations, class, mode, or frequency are nee
```

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

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

```
## [1] <NA> <NA> <NA>
```

```
## Levels: East West North
```

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

```
imported_table <- read.table(file = "/cloud/project/RWorksheet#4/RWorksheet#4a/import_march.csv", heade
imported_table
```

```
##   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
```

Using Conditional Statements (IF-ELSE)

6. Full Search

```
randomNum <- readline(prompt = "Enter number from 1 to 50: ")
```

```
## Enter number from 1 to 50:
```

```
#can't knit if there is as.numeric
#randomNum <- as.numeric(randomNum)

paste("The number you have chosen is", randomNum)
```

```
## [1] "The number you have chosen is "
```

```
if (randomNum > 50) {
  paste("The number selected is beyond the range of 1 to 50")
} else if (randomNum == 20) {
  paste("TRUE")
} else {
  paste(randomNum)
}
```

```
## [1] ""
```

7. Change

```
minimumBills <- function(price) {

  minBills <- price %/% 50
  paste("The minimum no. of bills:", minBills)
}
minimumBills(90)
```

```
## [1] "The minimum no. of bills: 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.

```
names <- c("Annie", "Thea", "Steve", "Hanna")
grade1 <- c(85,65,75,95)
grade2 <- c(65,75,55,75)
grade3 <- c(85,90,80,100)
grade4 <- c(100,90,85,90)

mathScore <- data.frame(
  Name = names,
  Grade1 = grade1,
  Grade2 = grade2,
  Grade3 = grade3,
  Grade4 = grade4
)
```

#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. Example Output: Annie's average grade this semester is 88.75.

```
mathScore$Average <- (mathScore$Grade1 + mathScore$Grade2 + mathScore$Grade3 + mathScore$Grade4) / 4

highscorers <- mathScore[mathScore$Average > 90,]
highscorers
```

```
## [1] Name    Grade1  Grade2  Grade3  Grade4  Average
## <0 rows> (or 0-length row.names)
```

```
if (nrow(highscorers) > 0) {
  paste(highscorers$Name, "'s average grade this semester is", high_scorers$Average)
} else {
  paste("No students have an average math score over 90.")
}
```

```
## [1] "No students have an average math score over 90."
```

#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.

```
firstTest <- sum(mathScore$Grade1) / nrow(mathScore)
firstTest
```

```
## [1] 80
```

```
secondTest <- sum(mathScore$Grade2) / nrow(mathScore)
secondTest
```

```
## [1] 67.5
```

```
thirdTest <- sum(mathScore$Grade3) / nrow(mathScore)
thirdTest
```

```
## [1] 88.75
```

```
fourthTest <- sum(mathScore$Grade4) / nrow(mathScore)
fourthTest
```

```
## [1] 91.25
```

```
if (firstTest < 80) {
  paste("The 1st test was difficult.")
} else if(secondTest < 80) {
  paste("The 2nd test was difficult.")
} else if(thirdTest < 80) {
  paste("The 3rd test was difficult.")
} else if(fourthTest < 80) {
  paste("The 4th test was difficult.")
} else {
  paste("No test had an average score less than 80.")
}
```

```
## [1] "The 2nd test was difficult."
```

#d. 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.

```
# -annie scores-
if (mathScore[1,2] > mathScore[1,3] && mathScore[1,2] > mathScore[1,4] && mathScore[1,2] > mathScore[1,5
  annieHighest <- mathScore[1,2]
} else if (mathScore[1,3] > mathScore[1,4] && mathScore[1,3] > mathScore[1,5]) {
  annieHighest <- mathScore[1,3]
} else if (mathScore[1,4] > mathScore[1,5] && mathScore[1,2] > mathScore[1,5]) {
  annieHighest <- mathScore[1,4]
} else {
  annieHighest <- mathScore[1,5]
}


# -thea scores-
if (mathScore[2,2] > mathScore[2,3] && mathScore[2,2] > mathScore[2,4] && mathScore[2,2] > mathScore[2,5
  theaHighest <- mathScore[2,2]
} else if (mathScore[2,3] > mathScore[2,4] && mathScore[2,3] > mathScore[2,5]) {
  theaHighest <- mathScore[2,3]
} else if (mathScore[2,4] > mathScore[2,5] && mathScore[2,2] > mathScore[2,5]) {
  theaHighest <- mathScore[2,4]
} else {
  theaHighest <- mathScore[2,5]
}

# -steve scores-
if (mathScore[3,2] > mathScore[3,3] && mathScore[3,2] > mathScore[3,4] && mathScore[3,2] > mathScore[3,5
  steveHighest <- mathScore[3,2]
} else if (mathScore[3,3] > mathScore[3,4] && mathScore[3,3] > mathScore[3,5]) {
  steveHighest <- mathScore[2,3]
} else if (mathScore[3,4] > mathScore[3,5] && mathScore[3,2] > mathScore[3,5]) {
  steveHighest <- mathScore[3,4]
} else {
  steveHighest <- mathScore[3,5]
}

# -hanna scores-
if (mathScore[4,2] > mathScore[4,3] && mathScore[4,2] > mathScore[4,4] && mathScore[4,2] > mathScore[4,5
  hannaHighest <- mathScore[4,2]
```

```
} else if (mathScore[4,3] > mathScore[4,4] && mathScore[4,3] > mathScore[4,5]) {
  hannaHighest <- mathScore[2,3]
} else if (mathScore[4,4] > mathScore[4,5] && mathScore[4,2] > mathScore[4,5]) {
  hannaHighest <- mathScore[4,4]
} else {
  hannaHighest <- mathScore[4,5]
}

mathScore$HighestGrades <- c(annieHighest, theaHighest, steveHighest, hannaHighest)


above90 <- mathScore[mathScore$HighestGrades > 90,]
above90
```

```
##    Name Grade1 Grade2 Grade3 Grade4 Average HighestGrades
## 1 Annie     85     65     85    100   83.75           100
## 4 Hanna     95     75    100     90   90.00           100
```

```
if (nrow(above90) > 0) {
  paste(above90$Name, "'s highest grade this semester is", above90$HighestGrade)
} else {
  paste("No students have an average math score over 90.")
}
```

```
## [1] "Annie 's highest grade this semester is 100"
## [2] "Hanna 's highest grade this semester is 100"
```