RWorksheet#4a

Charles Bjarne Huervana

2023-10-25

```
#1. The table below shows the data about shoe size and height. Create a data frame.
dataShoeHeight <- data.frame(

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",

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(66.0,68.0,64.5,65.0,70.0,64.0,70.0,71.0,72.0,64.0,74.5,67.0,71.0,71.0,77.0,72.0,59.0,62.0,72
)

dataShoeHeight
```

```
##    Gender Shoe_Size 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
```

```r
#A Describe the data.

##This data frame contains information about individuals' gender, shoe size, and height.
#With the corresponding values for each attribute. It appears to be a small dataset with 28 data points

#B Create a subset by males and females with their corresponding shoe size and height.

females_subset <- dataShoeHeight[dataShoeHeight$Gender == "F", c("Gender", "Shoe_Size", "Height")]
females_subset
```

```
##    Gender Shoe_Size 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
```

```r
males_subset <- dataShoeHeight[dataShoeHeight$Gender == "M", c("Gender", "Shoe_Size", "Height")]
males_subset
```

```
##    Gender Shoe_Size 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
```

```r
#C Find the mean of shoe size and height of the respondents. Write the R scripts and its result.

mean_shoe_size <- mean(dataShoeHeight$Shoe_Size)
mean_height <- mean(dataShoeHeight$Height)

mean_shoe_size
```

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

```r
mean_height
```

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

```
#D Is there a relationship between shoe size and height? Why?
#The connection between the two is that shoe size increases in direct proportion to height; when height
```

```
#FACTORS
```

```
#2 Construct character vector months to a factor with factor() and assign the result to factor_months_v

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
```

```
##  [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)

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 b

summaryMonths <- summary(months)
summaryMonths
```

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

```
summaryFactorMonths <-summary(factor_months_vector)
summaryFactorMonths
```

```
##     April    August  December  February   January      July     March       May
##         2         4         1         2         3         1         1         1
##  November   October September
##         5         1         3
```

```
#"months_vector" provides information about the count of data points, data type, and the most frequentl
#"factor_months_vector" displays the distribution of how often each month appears.
#two types of summaries serve distinct purposes and are valuable in different situations

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

factorDir <- c("East", "West", "North")
factorFreq <- c(1,4,3)
```

```r
new_order_data <- factor(factorDir,levels = c("East","West","North"))
new_order_data
```

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

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

print(getwd())
```

```
## [1] "/cloud/project/RWorksheet4"
```

```r
file.exists("import_march.csv")
```

```
## [1] TRUE
```

```r
imported_table <- read.table(file =  "/cloud/project/RWorksheet4/import_march.csv", header = TRUE, sep =

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
```

```r
## Using Conditional Statements (IF-ELSE)


# 6 Full Search

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

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

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

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

```r
if (randomNum > 50) {
  paste("Must be 1 to 50 numbers only")
} else if (randomNum == 20) {
  paste("TRUE")
} else {
  paste(randomNum)
}
```

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

```r
# ---------

# 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 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
)
mathScore
```

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

```
# 8 B Without using the rowMean function, output the average score of students whose average math score

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

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

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

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

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

```
# 8 C Without using the mean function, output as follows for the tests in which the average score was l

Test1 <- sum(mathScore$Grade1) / nrow(mathScore)
Test1
```

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

```
Test2  <- sum(mathScore$Grade2) / nrow(mathScore)
Test2
```

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

```
Test3  <- sum(mathScore$Grade3) / nrow(mathScore)
Test3
```

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

```
Test4 <- sum(mathScore$Grade4) / nrow(mathScore)
Test4
```

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

```
if (Test1 < 80) {
  paste("The 1st test was difficult.")
} else if(Test2 < 80) {
  paste("The 2nd test was difficult.")
} else if(Test3 < 80) {
  paste("The 3rd test was difficult.")
} else if(Test4 < 80) {
  paste("The 4th test was difficult.")
} else {
  paste("No test had an average score less than 80.")
}
```

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

```
# 8 D Without using the max function, output as follows for students whose highest score for a semester

# ANNIE
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]
}
cat("Annie's highest score is:", annieHighest, "\n")
```

```
## Annie's highest score is: 100
```

```
# THEA
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]
}
cat("Thea's highest score is:", theaHighest, "\n")
```

```
## Thea's highest score is: 90
```

```
# STEVE
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]
}
cat("Steve's highest score is:", steveHighest, "\n")
```

```
## Steve's highest score is: 85
```

```
# HANNA
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]
}
cat("Hannah's highest score is:", hannaHighest, "\n")
```

```
## Hannah's highest score is: 100
```

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


check_above90 <- mathScore[mathScore$HighestGrades > 90,]
check_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(check_above90) > 0) {
  paste(check_above90$Name, "'s highest grade this semester is", check_above90$HighestGrade)
} else {
  paste("No one got the average of 90")
}
```

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