RWorksheet_Elizalde#4a

Christian Alimace Elizalde

2024-10-14

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

```
new_data <- read.csv("/cloud/project/Rworksheet4a/SshG.csv")
new_data
```

```
##    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   74.5      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 data shows the shoe size, height, and gender.

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

```
male <- subset(new_data, Gender == "M" & Height & Shoe.size)
male
```

```
##    Shoe.size Height Gender
## 5       10.5   70.0      M
```

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

```
female <- subset(new_data, Gender == "F" & Height & Shoe.size)
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.

```
mSS <- mean(new_data$Shoe.size)
mSS
```

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

```
mH <- mean(new_data$Height)
mH
```

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

#d. Is there a relationship between shoe size and height? Why? #Yes because taller individuals have larger shoe size while small individuals have smaller shoe size.

#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","Septembe
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. | Inter-pret the results of both vectors. Are they both equally useful in this case? #The result of months vector prints the length, class, and mode while the factor_months_vector prints the number of how many times a month has been mentioned in the vector. They are both useful.

```
summ <- summary(months)
summ
```

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

```
fsumm <- summary(factor_months_vector)
fsumm
```

```
##     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 <- c("East", "West", "North")
Frequency <- c(1,4,3)
factor_data <- data.frame(Direction, Frequency)
factor_data
```

```
##   Direction Frequency
## 1      East         1
## 2      West         4
## 3     North         3
```

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

```
## Direction Frequency
##      <NA>      <NA>
## 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.

```
readT <- read.table("import_march.csv", header = TRUE, sep = ",")
readT
```

```
##   Students Strategy1 Strategy2 Strategy3
## 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
```

#b. View the dataset. Write the R scripts and its result.

```
print(readT)
```

```
##   Students Strategy1 Strategy2 Strategy3
## 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

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

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

```
if(!is.na(num) == 20){
  print("TRUE")
}else if (!is.na(num) >=1 && num <=50){
  num
}else {
  print("The number selected is beyond the range of 1 to 50.")
}
```

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

#7. Change

```
snack <- as.numeric(readline(prompt="Enter price of snack: "))
```

```
## Enter price of snack:
```

```
minBills <- function(snack) {
  bills <- 0

  if (!is.na(snack) >= 1000) {
    bills <- bills + snack %/% 1000
    snack <- snack %% 1000
  }
  if (!is.na(snack) >= 500) {
    bills <- bills + snack %/% 500
    snack <- snack %% 500
  }
  if (!is.na(snack) >= 200) {
    bills <- bills + snack %/% 200
    snack <- snack %% 200
  }
  if (!is.na(snack) >= 100) {
    bills <- bills + snack %/% 100
    snack <- snack %% 100
  }
  if (!is.na(snack) >= 50) {
    bills <- bills + snack %/% 50
    snack <- snack %% 50
```

```
  }
  return(bills)
}
minBills(snack)
```

```
## [1] NA
```

#8. #a. create a data frame frome the above table. Write the codes and its output

```
df <- data.frame(Name = 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))
df
```

```
##    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
```

#b.

```
avg_scores <- rowSums(df[,-1]) / (ncol(df) - 1)
if (any(avg_scores > 90)) {
high_avg_names <- df$Name[avg_scores > 90]
high_avg_scores <- avg_scores[avg_scores > 90]
print0(paste(high_avg_names, "'s average grade this semester is ", high_avg_scores))
} else {
cat("No student's average grade is over 90.")
}
```

```
## No student's average grade is over 90.
```

#c.

```
test_avg <- colSums(df[, -1]) / nrow(df)

if(any(test_avg < 80)){
  test_hard <- which(test_avg<80)

cat(paste("The test", test_hard, "was difficult."))
 }else {
cat("All tests had an average scores of 80 or above.")

}
```

```
## The test 2 was difficult.
```

#d.

```
high_score <- apply(df[,-1], 1, function(x) sort(x, decreasing = TRUE)[1])
if (any(high_score > 90)) {
hnames <- df$Name[high_score > 90]
hscores <- high_score[high_score > 90]
cat(paste(hnames, "'s highest grade this semester is ", hscores,".\n"))
} else {
```

```
cat("No student's highest grade exceeded 90.")
}
```

```
## Annie 's highest grade this semester is  100 .
##  Hanna 's highest grade this semester is  100 .
```