RWorksheet_gagante#4a.Rmd

Liza Claire Gagante

2024-10-17

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

```r
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 shows data regarding the shoe size, height, and gender of each individuals.

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

```r
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
```

```r
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.

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

mean_shoe_size
```

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

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

mean_height
```

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

d. Is there a relationship between shoe size and height? Why? -Yes, beacuse as the height increases, shoe size tends to increase as well.

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. Consider data consisting of the names of months: “March”,“April”,“January”,“November”,“January”, “September”,“October”,“September”,“November”,“August”, “January”,“November”,“November”,“February”,“May”,“August”, “July”,“December”,“August”,“August”,“September”,“N April”)

```r
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
```

```r
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. | 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 result of months_vector it states the Length, Class and Mode. While the factor_months_vector states how many months in the data for example December, December has 1. As what I saw they are both useful because it is easy for me to understand and determine how many types of data from the raw data itself.

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

```r
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.

```
data_excel <- read.table("import_march.csv")
```

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

```
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
```

6. Full Search

a. Create an R Program that allows the User to randomly select numbers from 1 to 50.Then display the chosen number. If the number is beyond the range of the selected choice,it will have to display a string "The number selected is beyond the range of 1 to 50". If number 20 is inputted by the User, it will have to display "TRUE", otherwise display the input number.

```
exhaustive_search <- function(selected_number) {

  if(selected_number < 1 || selected_number > 50) {
    print("The selected number is beyond the range of 1 to 50")
  } else if(selected_number == 20) {
    print("TRUE")
  } else {
    print(selected_number)
  }
}
selected_number <- readline(prompt = "Select a number from 1 to 50: ")
```

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

```
  exhaustive_search(selected_number)
```

```
## [1] "The selected number is beyond the range of 1 to 50"
```

7. Change

a. Write a function that prints the minimum number of bills that must be paid, given the price of the snack. Input: Price of snack (a random number divisible by 50) Output: Minimum number of bills needed to purchase a snack.

```
min_bills <- function(price) {
  bills <- c(1000, 500, 200, 100, 50)

  num_bills <- 0
  for(bill in bills) {
    count <- price %/% bill
    num_bills <- num_bills + count

    price <- price %% bill
  }

```

```r
  print(paste("Minimum number of bills needed to purchase a snack: ", num_bills))
}

min_bills(1650)
```

```
## [1] "Minimum number of bills needed to purchase a snack:  4"
```

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 <- 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)
)

students
```

```
##    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. 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.

```r
for (i in 1:nrow(students)) {
  total_score <- students$Grade1[i] + students$Grade2[i] + students$Grade3[i] + students$Grade4[i]
  avg_score <- total_score / 4

  if (avg_score > 90) {
     formatted_output <- sprintf("%s's average grade this semester is %.2f.", students$Name[i], avg_sco
    print(formatted_output)
  }
}
```

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
for (j in 2:5) {
  total_test_score <- sum(students[, j])
  avg_test_score <- total_test_score / nrow(students)

  if (avg_test_score < 80) {
    print(paste("The", j-1, "th test was difficult."))
  }
}
```

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

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.

```r
for (i in 1:nrow(students)) {

  grades <- c(students$Grade1[i], students$Grade2[i], students$Grade3[i], students$Grade4[i])

  highest_grade <- grades[1]
  for (grade in grades) {
    if (grade > highest_grade) {
      highest_grade <- grade
    }
  }

  if (highest_grade > 90) {
    print(paste(students$Name[i], "'s highest grade this semester is", highest_grade))
  }
}
```

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