RWorksheet_Sison#4.

Kathrina Casandra Sison

2024-10-14

```
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
Height <- c(66.0, 68.0, 64.5, 65.0, 70.0, 64.0, 70.0, 71.0, 72.0, 64.0, 74.75, 67.0, 71.0, 71.0, 77.0, 
Gender <- c("F", "F", "F", "F", "M", "F", "F", "F", "M", "F", "M", "F", "M", "M", "M", "M", "F", "F", "
shoe_data1 <- data.frame(Shoe_Size,Height, Gender)
```

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

#a. Describe the data. The data is consist of different shoe sizes along with its height and gender of the user.

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

```
str(shoe_data1)
```

```
## 'data.frame':    28 obs. of  3 variables:
##  $ Shoe_Size: num  6.5 9 8.5 8.5 10.5 7 9.5 9 13 7.5 ...
##  $ Height   : num  66 68 64.5 65 70 64 70 71 72 64 ...
##  $ Gender   : chr  "F" "F" "F" "F" ...
```

```
male_subset <- shoe_data1[shoe_data1$Gender == "M",]
female <- shoe_data1[shoe_data1$Gender == "F",]
print("Male_subset")
```

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

```
print(male_subset)
```

```
##    Shoe_Size Height Gender
## 5       10.5  70.00      M
## 9       13.0  72.00      M
## 11      10.5  74.75      M
## 13      12.0  71.00      M
## 14      10.5  71.00      M
## 15      13.0  77.00      M
## 16      11.5  72.00      M
## 19      10.0  72.00      M
## 22       8.5  67.00      M
## 23      10.5  73.00      M
## 25      10.5  72.00      M
## 26      11.0  70.00      M
## 27       9.0  69.00      M
## 28      13.0  70.00      M
```

```
print("Female Susbet")
```

```
## [1] "Female Susbet"
```

```
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.

```
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
Height <- c(66.0, 68.0, 64.5, 65.0, 70.0, 64.0, 70.0, 71.0, 72.0, 64.0, 74.75, 67.0, 71.0, 71.0, 77.0, 
mean_shoe_size <- mean(Shoe_Size)
mean_height <- mean(Height)
mean_shoe_size
```

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

```
mean_height
```

```
## [1] 68.58036
```

d. Is there a relationship between shoe size and height? Why?

```
correlation <- cor(shoe_data1$Shoe_Size, shoe_data1$Height, use = "complete.obs")
print(paste("Correlation between Shoe Size and Height:", correlation))
```

```
## [1] "Correlation between Shoe Size and Height: 0.775187654599107"
```

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")

```
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")
factor_months_vector <- factor(months)

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
```

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

```
factor_months <- factor(months)
summary(months)
```

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

```
summary(factor_months)
```

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

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

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

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

5.

```
write.csv("import_march.csv", row.names = FALSE)
```

```
## "x"
## "import_march.csv"
```

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

```
##   Students Strategy1 Strategy2 Strategy3
## 1     Male         8        10         8
## 2     Male         4         8         6
## 3     Male         0         6         4
## 4   Female        14         4        15
## 5   Female        10         2        12
## 6   Female         6         0         9
```

6

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
number <- as.integer(readline(prompt = "Enter a number: "))
```

```
## Enter a number:
```

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

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

7.

```
mBills <- function(price) {
  bills <- c(50, 100, 200, 500, 1000)
  count <- 0
  for ( bill in bills) {
  count <- count + price %/% bill
  price <- price %%bill
  }
  return(count)
}
snack_price <- as.integer(readline(prompt = "Enter the price of the snack:"))
```

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

```
print(paste("Minimum number of bills needed: ", mBills(snack_price)))
```

```
## [1] "Minimum number of bills needed:  NA"
```

8.

```
#A
math_score <- 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)
)
print(math_score)
```

```
##    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
math_score$Average <- (math_score$Grade1 + math_score$Grade2 + math_score$Grade3 + math_score$Grade4) /
for (i in 1:nrow(math_score)) {
  if (math_score$Average[i] > 90) {
    cat(math_score$Name[i], "'s average grade this semester is", math_score$Average[i], "\n")
  }
}
"There is no student who got over 90 points during the semester."
```

```
## [1] "There is no student who got over 90 points during the semester."
```

```
#C
test_averages <- c(
  (sum(math_score$Grade1) / nrow(math_score)),
  (sum(math_score$Grade2) / nrow(math_score)),
  (sum(math_score$Grade3) / nrow(math_score)),
  (sum(math_score$Grade4) / nrow(math_score))
)
for (i in 1:length(test_averages)) {
  if (test_averages[i] < 80) {
    cat("The", i, "nd test was difficult.\n")
  }
}
```

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

```
#D
for (i in 1:nrow(math_score)) {
  highest_score <- 0
  for (j in 2:5) {
    if (math_score[i, j] > highest_score) {
      highest_score <- math_score[i, j]
    }
  }
  if (highest_score > 90) {
    cat(math_score$Name[i], "'s highest grade this semester is", highest_score, ".\n")
  }
}
```

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