

RWorkSheet#4a

2023-10-25

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

```
df <- data.frame(
  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, 10.5),
  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, 62.0),
  Gender = c('F','F','F','F','M','F','F','F','M','F','M','F','M','M', 'M', 'M', 'F','F', 'M', 'F', 'F'),
)
df
```

##	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

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

```
df_males <- subset(df, Gender == 'M')
df_females <- subset(df, Gender == 'F')
df_males
```

```
##      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
```

```
df_females
```

```
##      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.
```

```
mean_shoe_size <- mean(df$Shoe_Size)
mean_height <- mean(df$Height)
paste("The mean shoe size of the respondents is:", mean_shoe_size)
```

```
## [1] "The mean shoe size of the respondents is: 9.41071428571429"
```

```
paste("The mean height of the respondents is:", mean_height)
```

```
## [1] "The mean height of the respondents is: 68.5714285714286"
```

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

```
# 2.
```

```
# Constructing the character vector months
```

```
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")
```

```
# Converting the character vector months to a factor
```

```
factor_months_vector <- factor(months_vector)
```

```
# Printing out the result
```

```
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.
```

```
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
```

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

```
direction <- c("East", "West", "North")
```

```
frequency <- c(1,4,3)
```

```
factor_data <- factor(c(direction,frequency))
```

```
factor_data
```

```
## [1] East  West  North 1      4      3
## Levels: 1 3 4 East North West
```

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

```
print(new_order_data)
```

```
## [1] East  West  North <NA> <NA> <NA>
```

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

```
# 5
```

```
read.table(file = "/cloud/project/RWorksheet#4a/import_march.csv", header = TRUE, sep = ",")
```

```
##      X      X.1      X.2      X.3
## 1 Students Strategy 1 Strategy 2 Strategy 3
## 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
```

```
reading <- read.csv("import_march.csv")
```

```
reading
```

```
##      X      X.1      X.2      X.3
```

```
## 1 Students Strategy 1 Strategy 2 Strategy 3
## 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
```

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

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

```
#error cannot 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
```

```
minimumBills <- function(price) {

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

```
minimumBills(900)
```

```
## [1] "The minimum no. of bills: 18"
```

```
# 8.a
```

```
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)
```

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

```
# 8.b

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

highScorers <- grade[grade$Average > 90,]
highScorers

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

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

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

```
# 8.c
```

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

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

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

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

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

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

```
fourthTest <- sum(grade$Grade4) / nrow(grade)
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."
```

```
# 8.d
```

```
# Annie scores
```

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

```

    annieHighest <- grade[1,3]
  } else if (grade[1,4] > grade[1,5] && grade[1,2] > grade[1,5]) {
    annieHighest <- grade[1,4]
  } else {
    annieHighest <- grade[1,5]
  }

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

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

  # Hanna scores
  if (grade[4,2] > grade[4,3] && grade[4,2] > grade[4,4] && grade[4,2] > grade[4,5]) {
    hannaHighest <- grade[4,2]
  } else if (grade[4,3] > grade[4,4] && grade[4,3] > grade[4,5]) {
    hannaHighest <- grade[4,3]
  } else if (grade[4,4] > grade[4,5] && grade[4,2] > grade[4,5]) {
    hannaHighest <- grade[4,4]
  } else {
    hannaHighest <- grade[4,5]
  }

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

  above90 <- grade[grade$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"
```