

Worksheet#4a

#1. The table below shows the data about shoe size and height. Create a data frame. `DATA <- data.frame`
`(ShoeSize=c(6.5,9.0,8.5,8.5,10.5,7.0,9.5,9.0,13.0,7.5,10.5,8.5,12.0,10.5,13.0,11.5,8.5,5.0,10.0,6.5,7.5,8.5,10.5,8.5,10.5,11.0,9.0,13.0)`

`Height=c(66.0,68.0,64.5,65.0,70.0,64.0,70.0,71.0,72.0,64.0,74.5,67.0,71.0,71.0,77.0,72.0,59.0,62.0,72.0,66.0,64.0,67.0,73.0,69.0,72.0`

`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","F","M","M")`

`DATA`

#a. Describe the data. #The data has 28 entries with 3 columns #First Column shows Shoe_Size #Second Column displays Height #Third column shows Gender

#b. Create a subset by males and females with their corresponding shoe size and height. `male-sub <- DATA[DATA$Gender == "M", c("ShoeSize", "Height", "Gender")]` `malesub` `femalesub <- DATA[DATA$Gender == "F", c("ShoeSize", "Height", "Gender")]` `femalesub` #c. Find the mean of shoe size and height of the respondents. `Shoes <- mean(DATA$ShoeSize)` `ShoesHeights <- mean(DATA$Height)` `Heights` #d. Is there a relationship between shoe size and height? Why? #Yes #Shoe size generally is proportional to height. #We can expect people with large feet to be taller.

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

`months <- c("March", "April", "January", "November", "January", "September", "October", "September", "November", "August", "J`
`months` `factor_months_vector <- factor(months)` `factor_months_vector`

#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?

`summary(months)` `summary(factor_months_vector)` # In the summary of months, it shows the number of length, class, and mode of the vector. # In the summary of `factor_months_vector`, it shows the frequency of each months. # Both are useful in different cases where the no. of length, class, mode, or the frequency is needed.

#4. Create a vector and factor for the table below. `factor_freq <- c(1,4,3)` `factor_vector <- c("East", "West", "North")` `new_factor_vector <- factor(factor_vector, levels = c("East", "West", "North"))` `print(new_factor_vector)`

#5. Enter the data below in Excel with file name = `import_march.csv` `write.csv(import_march, file = "import_march.csv")` `file_path <- "/cloud/project/import_march.csv"` `imported_table <- read.csv(file = file_path, header = TRUE, sep = ",")` `View(imported_table)`

#6. Full Search `rannum <- readline(prompt = "Enter number from 1 to 50:")` `rannum <- as.numeric(rannum)` `paste("The number you have chosen is", rannum)` `if (rannum > 50) { paste("The number selected is beyond the range of 1 to 50") }` `else if (rannum == 20) { paste("TRUE") }` `else { paste(rannum) }`

#7. Change minibills `<- function(price) {`

`minibills <- price %/% 50` `paste("The minimum no. of bills:", minibills) }` `minibills(90)`

#8a. Create a dataframe from the above table. `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)` `mathgrades <- data.frame(Name = names, Grade1 = grade1, Grade2 = grade2, Grade3 = grade3, Grade4 = grade4)` `mathgrades`

#8b. Without using the `rowMean` function, output the average score of students whose average math score over 90 points during the semester. `mathgradesAverage <- (mathgrades$Grade1 + mathgrades$Grade2 + mathgrades$Grade3 + mathgrades$Grade4) / 4`

`highscorers <- mathgrades[mathgrades$Average > 90,]` `highscorers`

`if (nrow(highscorers) > 0) { paste(highscorers$Name, "saveragegradethissemesteris", highscorers$Average)`
`} else { paste("No students have an average math score over 90.") }`

```
#8c. Without using the mean function, output as follows for the tests in which the average score was less than 80 out of 4 tests. firstTest <- sum(mathgrades$Grade1) / nrow(mathgrades) firstTest
```

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

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

```
fourthTest <- sum(mathgrades$Grade4) / nrow(mathgrades) fourthTest
```

```
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.") }
```

```
#8d. Without using the max function, output as follows for students whose highest score for a semester exceeds 90 points. # annie scores if (mathgrades[1,2] > mathgrades[1,3] && mathgrades[1,2] > mathgrades[1,4] && mathgrades[1,2] > mathScore[1,5]) { annieHighest <- mathgrades[1,2] } else if (mathgrades[1,3] > mathgrades[1,4] && mathgrades[1,3] > mathgrades[1,5]) { annieHighest <- mathgrades[1,3] } else if (mathgrades[1,4] > mathgrades[1,5] && mathgrades[1,2] > mathgrades[1,5]) { annieHighest <- mathgrades[1,4] } else { annieHighest <- mathgrades[1,5] }
```

thea scores

```
if (mathgrades[2,2] > mathgrades[2,3] && mathgrades[2,2] > mathgrades[2,4] && mathgrades[2,2] > mathgrades[2,5]) { theaHighest <- mathgrades[2,2] } else if (mathgrades[2,3] > mathgrades[2,4] && mathgrades[2,3] > mathgrades[2,5]) { theaHighest <- mathgrades[2,3] } else if (mathgrades[2,4] > mathgrades[2,5] && mathgrades[2,2] > mathgrades[2,5]) { theaHighest <- mathgrades[2,4] } else { theaHighest <- mathgrades[2,5] }
```

steve scores

```
if (mathgrades[3,2] > mathgrades[3,3] && mathgrades[3,2] > mathgrades[3,4] && mathgrades[3,2] > mathgrades[3,5]) { steveHighest <- mathgrades[3,2] } else if (mathgrades[3,3] > mathgrades[3,4] && mathgrades[3,3] > mathgrades[3,5]) { steveHighest <- mathgrades[2,3] } else if (mathgrades[3,4] > mathgrades[3,5] && mathgrades[3,2] > mathgrades[3,5]) { steveHighest <- mathgrades[3,4] } else { steveHighest <- mathgrades[3,5] }
```

hanna scores

```
if (mathgrades[4,2] > mathgrades[4,3] && mathgrades[4,2] > mathgrades[4,4] && mathgrades[4,2] > mathgrades[4,5]) { hannaHighest <- mathgrades[4,2] } else if (mathgrades[4,3] > mathgrades[4,4] && mathgrades[4,3] > mathgrades[4,5]) { hannaHighest <- mathgrades[2,3] } else if (mathgrades[4,4] > mathgrades[4,5] && mathgrades[4,2] > mathgrades[4,5]) { hannaHighest <- mathgrades[4,4] } else { hannaHighest <- mathgrades[4,5] }
```

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

```
above90 <- mathgrades[mathgrades$HighestGrades > 90,] above90
```

```
if (nrow(above90) > 0) { paste(above90Name, "shighestgradethissemesteris", above90HighestGrade) } else { paste("No students have an average math score over 90.") }
```