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

Shoesize <- data.frame (Gender=c("F","F","F","F","M","F","F","F","M","F","M","F","M","M","M","M","F","F","M","F",
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
) Shoesize

#a. Describe the data. #The shoe size and height of M is much more higher than F

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

malesubset <- Shoesize[Shoesize$Gender == "M", c("Gender", "ShoeSize", "Height")] femalesubset < -Shoesize[Shoesize$Gender == "F", c("Gender", "ShoeSize", "Height")] malesubset femalesubset

#c. Find the mean of shoe size and height of the respondents. Write the R scripts and its result.

meanofShoeSize <- mean(Shoesize$ShoeSize) meanofHeight < -mean(Shoesize$Height) meanofShoeSize meanofHeight

#d. Is there a relationship between shoe size and height? Why? correlationCoefficient <- cor(Shoesize$ShoeSize, Shoesize$Height
correlationCoefficient

#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","September","November","February", "April") factorMonthsVec <- factor(months) factorMonthsVec

3Then 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(factorMonthsVec)

#The number of observations, class, and mode of the months_vector are displayed in the summary.

The frequency of each month is shown in the factor_months_vector summary.

Both are beneficial in many situations when the quantity of observations, class, mode, or frequency are required.

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

factor_data <- c(1,4,3) new_order_data <- factor(factor_data,levels = c("East","West","North")) print(new_order_data)

5

imported_table <- read.table(file = "/cloud/project/worksheet4/import_march.csv", header = TRUE, sep = ",")

View(imported_table)

6

randomNumber <- readline(prompt = “Enter number from 1 to 50:”)

randomNumber <- as.numeric(randomNumber) paste(“The number you have chosen is”, randomNumber)

if (randomNumber > 50) { paste(“The number selected is beyond the range of 1 to 50”) } else if (randomNumber == 20) { paste(“TRUE”) } else { paste(randomNumber) }

7

minimumBills <- function(price) {

minBills <- price %/% 50 paste(“The minimum no. of bills:”, minBills) } minimumBills(90)

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)

Math <- data.frame(Name = Names,
Grade1 = Grade1,
Grade2 = Grade2,
Grade3 = Grade3,
Grade4 = Grade4
) Math

8.b

Math$Average < -(Math$Grade1 + Math$Grade2 + Math$Grade3 + Math$Grade4)/4high_scorers < -Math[Math$Average > 90,] average_score <- sum(high_scorers$Average) / nrow(high_scorers) print(average_score)

#8 c.

Math$Average < -(Math$Grade1 + Math$Grade2 + Math$Grade3 + Math$Grade4)/4below_80_scorers < -Math[Math$Average < 80,] print(below_80_scorers)

#8 d.

Math$HighestScore < -apply(Math[, 2 : 5], 1, function(x)max(x))high_scorers < -Math[Math$HighestScore > 90,] high_scorers <- high_scorers[, -ncol(high_scorers)] print(high_scorers)

0