**Exercise 1: Introduction to R**

**Name…………………………………………………..ID………………………..**

1. Read data “Exercise1P1.txt” from the folder. The file contains a grade sheet of one class. Fill in the following spaces and write the first 10 lines of the data frame in the following space.
   1. Command:………………………………………………………
   2. The first 10 lines are:
   3. The data frame has ……………rows and …………………….columns.
   4. The column names of the data frame are………………………………

……………………………………………………………………………

* 1. How many students who get an “A” from the class……………………….
  2. How many students who get an “F” from the class…………………….
  3. What are the minimum and maximum of total scores…………………
  4. How many girls in the class……………………………………………..
  5. How many girls in the class who get an “A”……………………………
  6. Create a column which is the log of “TOTAL” and name it as

“ LogT”. Print the first 5 rows below.

* 1. Create a data frame containing only “ID”, “GRADE” and name it as “EX1M”. Print the first 5 rows below.
  2. Create an Excel file named “Exercise1P1.csv” from the file “Exercise1P1.txt” .
  3. Create an SPSS file named “Exercise1P1.SAV” from the file “Exercise1P1.txt” .
  4. Create a text file named “Exercise1M.txt” from EX1M.

1. Read data file “Exercise1P2.SAV”, make it to be a data frame named “Exer2D”.
   1. What is the command you use to read the file?

………………………………………………………………………..

* 1. What is the dimension of the data frame?.......................................
  2. Find quantiles of each column of “Exer2D” and show results in the form of a matrix.

Command:

Output:

* 1. Find quantiles of each column of “Exer2D” and show results in the form of lists.

Command:

Output:

* 1. Create one new column, named MIN, containing the minimum values of each row and one new column, named MAX, containing the maximum values of each row. Write down the new data frame below.

Command:

Output:

* 1. Create a new data frame named “Exer2New” containing only the two new columns, MAX and MIN created in 2.3. Write down the new data frame below.

Command:

Output:

* 1. Create a new data frame named “Exer2New2” containing three columns: “MAX”, “DIF” =MAX-MIN, “SUM”= MAX+MIN.

Command:

Output:

* 1. Merge “Exer2New” and “Exer2New2”, and name it as Exer2Merge.

Command:

Output:

* 1. Create a text file named “Exercise2New.txt” from Exer2Merge.
  2. Read the first 5 lines of the file “Exercise2New.txt” into R and name it as EXer2N

………………………………………………………………………….

Write your output below:

Command:

Output:

* 1. Change the data frame EXer2N to be a matrix and create a text file named “EXer2Mat.txt” containing the matrix. Write down the command below.

Command:

Output:

* 1. Read the file “EXer2Mat.txt”.

Command:

Output:

1. Read data file “Exercise1P3.txt” into a data frame named “EX1P3”.
   1. Create a contingency table of SEX and EDU variables.

Command:

Output:

* 1. Create a contingency table of SEX and EDU variables with row margins.

Command:

Output:

* 1. Create a contingency table of SEX and EDU variables with both row margins and column margins.

Command:

Output:

* 1. Create a contingency table of SEX and EDU variables with proportions instead of counts(row sum =1).

Command:

Output:

* 1. Create a contingency table of SEX and EDU variables with proportions instead of counts(column sum =1).

Command:

Output:

* 1. Create a flat contingency table of SEX, EDU and STATUS. Write down your command and output below.

Command:

Output:

* 1. Find average income of those whose who has “SEX”=”Male” and “STATUS” =1.

Command:

Output:

* 1. Find averages of “GRADE” of Males and Females.

Command:

Output:

* 1. Create a list named “P3list” containing three sub-lists as the following:

P3list$ID = 1:5

P3list$SEX= c(“M”,”M”,”F”,”M”,”F”)

P3list$EDUC=c(25,32,44,18,38)

Command:

Output:

* 1. Create a list named P3L1 containing three sub-lists as the following:

P3L1$ID= EX1P3$ID

P3L1$INCOME= EX1P3$ INCOME

P3L1$GRADE= EX1P3$ GRADE

Command:

Output:

* 1. Create a data frame from P3L1 and name it “P3L1D1”.

Command:

Output:

* 1. Create a data frame from P3L1 containing columns “EDUC” and “STATUS” and a new column “NUMB”=1:lengh(P3L1D1). The name of new data frame is “P3L1D2”.

Command:

Output:

* 1. Merge the two data frames, “P3L1D1” and “P3L1D2” by matching P3L1D1$ID and P3L1D2$NUMB .

Command:

Output: