THE OPEN UNIVERSITY OF SRI LANKA
FACULTY OF ENGINEERING TECHNOLOGY
BACHELOR OF SOFTWARE ENGINEERING HONOURS
ACADEMIC YEAR – 2022/23
EEX4373– Data Science
Lab Test 01

Duration - 2 hours

Instructions

- Answer **All** the questions.
- You should **upload** the **ZIP File** to the given drop box within 2 hours.
- Write all your codes in a R script file and save the R script file using your index number.
- Paste your screenshots in a Word document. Save the Word document using your index number.
- Then **Word document** and **R script** file are combined into a zip file.
- Browsing the **web** is **prohibited**.
- Late submissions will **reduce 5marks** per every 10min.
- Use proper **conventions** and **indentation** when writing the code.

Question 01 (30 marks)

(a) Write an R code to create a vector with the elements 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 and 120 and name the vector as "*V1".*

(b) Write R codes to find mean, median and quartiles of the above numbers.

(c) Write an R code to create a matrix with 4 rows and 3 columns by using the data from above vector *V1* assigned row wise. Name the matrix as "*M1*".

(d) Write R codes to change the row names of *M1* as "*R1*", "*R2*", "*R3*" and "*R4*". And change the column names of *M1* as "*C1*", "*C2*" and "*C3*". Get the screen shot of your output and attach to the word document.

Question 02 (35 marks)

Table 1 shows the details of 10 employees working in an organization. Write R codes for each of the following.

Table 1: Employees' details

EmpID	EmpName	Gender	Department	Salary ($)
#01	A	M	IT	750
#02	B	F	HR	500
#03	C	F	FINANCE	550
#04	D	M	FINANCE	500
#05	E	M	HR	400
#06	F	F	IT	700
#07	G	M	IT	600
#08	H	M	HR	450
#09	I	M	FINANCE	500
#10	J	F	IT	650

(i) Create a data frame for Table 1 and name it as *EmpData*.

(ii) Add 100 dollars bonus to *Salary* of each employee.

(iii) Obtain the details of the employees who are attached to the FINANCE department.

(iv) Obtain the *EmpName* of the employers whose salary is above 600 dollars.

(v) Obtain the *EmpName* of the employers who are attached to the IT Department and whose salary is below 700 dollars. Get the screen shot of your output and attach to the word document.

Question 03 (35 marks)

Consider the "*mtcars*" R inbuild data set. Figure 1 shows the first six observations out of 32 observations of the *mtcars* data set.

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225	105	2.76	3.460	20.22	1	0	3	1

Figure 1: First six observations of *mtcars* data set

Table 2 shows the description of some variables in the *mtcars* data set.

Table 2: Variable details

Variable Name	**Description**
mpg	Miles/(US) gallon
cyl	Number of cylinders
disp	Displacement
hp	Gross horsepower
wt	Weight (1000 lbs)
vs	Engine (0 = V-shaped, 1 = straight)
am	Transmission (0 = automatic, 1 = manual)

(i) Create a suitable plot to visualize miles per gallon (*mpg*) of the vehicles.

(ii) Crete a suitable plot to visualize the displacement (*disp)* of automatic and manual vehicles separately in a single plot. Get the screen shot of your output and attach to the word document.

(iii) Obtain the descriptive statistics for the variable *wt*.

(iv) Obtain the descriptive statistics of the variable *wt* for each engine type. Get the screen shot of your output and attach to the word document.

(v) Obtain the deciles of the variable *hp*.

---- END OF THE PAPER ----