UNIVERSITY OF COLOMBO SCHOOL OF COMPUTING
09 FEB 2022
EXAMINATION & REGISTRATION
U. C. S. C. LIBRARY UNIVERSITY OF COLOMBO

UNIVERSITY OF COLOMBO, SRI LANKA

UNIVERSITY OF COLOMBO SCHOOL OF COMPUTING

BACHELOR OF SCIENCE IN COMPUTER SCIENCE

*Academic Year 2021/2022 – Second Year Examination – Semester II*

# *SCS2211 – Laboratory II*
# *Part B*

*TWO (2) HOURS (For Part A and B)*



*To be completed by the candidate*

Examination Index No: .....

**Important Instructions to candidates:**

1. The medium of instruction and question is **English**.
2. **Write your answers** in **English.**
3. If a page or a part of this question paper is not printed, please inform the supervisor immediately.
4. Note that questions appear on both sides of the paper. If a page is not printed, please inform the supervisor immediately.
5. Write your index number on each and every page of the Question paper.
6. This paper has **02** questions on **07** pages.
7. Answer **ALL** questions. All questions carry equal marks (25 marks).
8. **This paper consists of two parts, Part A (Question No 1 and Question No 2) and Part B (Question No 3 and Question No 4) which have submit separately.**
9. Any electronic device capable of storing and retrieving text including electronic dictionaries and mobile phones are **not allowed**.
10. **Non-Programmable** calculators are **allowed**.

| For Examiner's use only | |
|---|---|
| | |
| **Question No** | **Marks** |
| | |
| 3 | |
| 4 | |
| **Total** | |

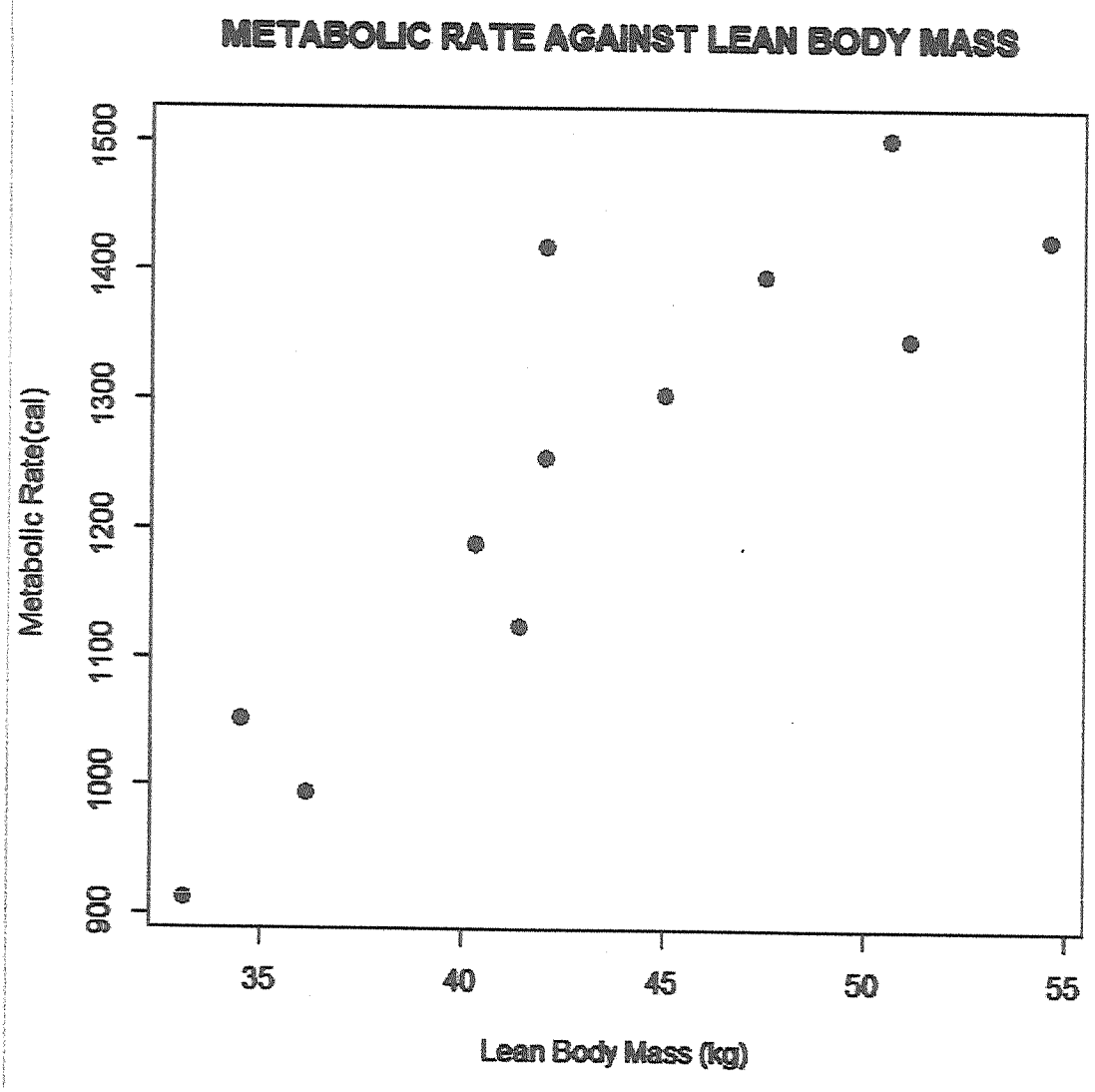
Index No: .....

## Part B

**Question 3**

Answer the questions (a) and (b) with respect to the given scatterplot. Assume that the following two vectors are already created for you in your R workspace.

Mass = c(36.1, 54.6, 47.5, 42.0, 50.6, 42.0, 40.3, 33.1, 41.4, 34.5, 51.1, 45)
Rate = c(995, 1425, 1396, 1418, 1502, 1256, 1189, 913, 1124, 1052, 1347, 1304)



**Index No: .....**

(a)

(i) Write the R code to produce the given scatter plot.

[5 Marks]

| |
|---|
| |
| |
| |
| |
| |

(ii) Identify the explanatory variable and response variable.

[2 Mark]

| |
|---|
| Explanatory: |
| Response: |

(iii) Interpret the overall pattern of the scatterplot given above.

[3 Marks]

| |
|---|
| |
| |

(iv) What is the difficulty you faced when interpreting the above relationship? And which statistical method can be applied to overcome that difficulty?

[2 Marks]

| |
|---|
| |
| |

**Index No: .....**

(b) For the above Mass and Rate data **correlation** and $r^2$ are calculated as **0.88** and **0.78** respectively. Using your statistical knowledge, interpret the relationship between Mass and Rate using correlation and $r^2$.

[6 Marks]

| **Correlation:** |
|---|
| |
| |
| $r^2$: |
| |
| |
| |

(c) Following snapshot gives the output of a least squares regression line fitted for Mass and Rate data.

```
Coefficients:
(Intercept)        Mass
      189.7        24.4
```

(i) Write the R code to find least squares regression line.

[2 Mark]

| |
|---|
| |

(ii) Find the metabolic rate when lean body mass is 38.

[5 Marks]

| |
|---|
| |
| |
| |

**Index No: .....**

**Question 4**

(a) Write the output of the following R commands with respect to following given vector x:

[10 Marks]

**> x = c(1, 3, 5, 7, 8, 9)**

| |
|---|
| > x>3 |
| |
| |
| > x = = 3 |
| |
| |
| > x= = 3 \| x != 3 |
| |
| |
| > x[x>3] |
| |
| |
| > x[x!=3] |
| |
| |

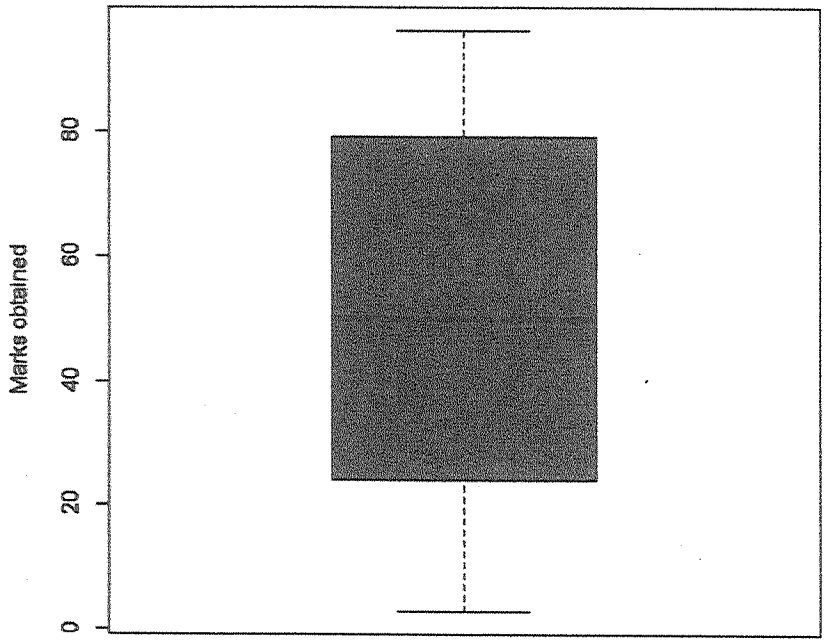
**Index No: .....**

**(b)** Following R code generates a graphical representation of marks obtained by 50 students in a programming course. This student sample is taken from the recent 03 batches from ABC college. Answer the questions (i) to (iii) based on these details.

```
> boxplot(marks, ylab="Marks obtained")
> summary(marks)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   3.00   24.75   50.00   48.70   78.25   96.00
> sd(marks)
[1] 29.90615
```

(i) Identify and markdown the five number summary on the given boxplot diagram.

[5 marks]




(ii) Calculate the range of the middle 50% of the data (IQR).

[2 mark]

**Index No: .....**

(iii) State 03 numerical measures that can be used to measure the central tendency of the marks.

[3 marks]

(iv) Variance is used as a numerical measure of a spread. What is the variance of the marks?

[2marks]

(v) If the SEM (Standard Error of the Mean) for the all the students in last 03 batches is 3.13, calculate the interval estimation of the population mean with 95% confidence. Please note that $Z^*$ for 95% confidence level is 1.96.

[3 marks]

**********