

[Time:2.30 Hrs]

[Marks:60]

Please check whether you have got the right question paper.

N.B: (1) All questions are compulsory.

(2) Figures to the right indicate full marks.

(3) Assume additional data if necessary but state the same clearly.

(4) Symbols have their usual meanings and tables have their usual standard design unless stated otherwise.

(5) Use of calculators and statistical tables are allowed. / If required, keep it.

- Q.1** Attempt any two of the following **12**
- Explain the basics of Business layer.
 - Explain the basics of Utility layer.
 - Explain the Rapid Information Factory ecosystem.
 - List and explain different programming languages using in data science processing.
- Q.2** Attempt any two of the following **12**
- Explain Directed Acyclic Graph Scheduling.
 - List the super steps for processing the data lake.
 - Explain Data Lakes and Data Swamps.
 - Why is it necessary to train the data science team?
- Q.3** Attempt any two of the following **12**
- Write a python program to Drop the Columns Where All Elements Are Missing Values.
 - Write a python program to Fill All Missing Values with the Mode of the particular column.
 - Write Python program to create the network routing diagram from the given data.
 - Write a Python program to generate visitors data from the given csv file.
- Q.4** Attempt any two of the following **12**
- What are the different typical reference satellites? Explain.
 - What is a fishbone diagram? Explain with example.
 - Explain hypothesis testing, t-test and chi-square test with respect to data science.
 - Write a note on ANOVA.
- Q.5** Attempt any two of the following **12**
- Explain Tensorflow with example.
 - Write a note on Principal Component Analysis.
 - What is organize superstep? Explain in brief.
 - Why is report superstep important? Explain its importance.

*****END*****