(2½ Hours) **[Total Marks: 75]**

N. B.: (1) **All** questions are **compulsory**.
(2) Make **suitable assumptions** wherever necessary and **state the assumptions** made.
(3) Answers to the **same question** must be **written together**.
(4) Numbers to the **right** indicate **marks**.
(5) Draw **neat labelled diagrams** wherever **necessary**.
(6) Use of **Non-programmable** calculators is **allowed**.

Q1. Attempt any the three of the following. **15 Marks**

a. Explain Apache Mesos and Apache Spark tools
b. Explain data structures in the functional layer of the ecosystem
c. Write a Python program to build directed acyclic graph.
d. Explain precision recall, precision recall curve, sensitivity, specificity and F1 measure.
e. Explain Hypothesis and CHI-SQUARE testing.

Q2. Attempt any the three of the following. **15 Marks**

a. Explain Hub in data science
b. List the super steps for processing the data lake.
c. Write a python program to Drop the Columns Where All Elements Are Missing Values.
d. What are the different typical reference satellites? Explain.
e. Explain importance of missing value treatment.

Q3. Attempt any three of the following. **15 marks**

a. Explain Schema on read ecosystem
b. Explain Data Lakes and Data Swamps.
c. Explain the different ways to deal with errors.
d. Explain the Event section of TPOLE.
e. Explain Horizontal Style and Vertical Style Slicing.

Q4. Attempt any three of the following. **15 marks**

a. Explain Scala, Akka and Kafka
b. Why is it necessary to train the data science team?
c. Explain the principles of data analysis.
d. What is a fishbone diagram? Explain with example.
e. Explain Association Rule Mining with market basket analysis.

Q5. Attempt any three of the following. **15 marks**

a. Explain Schema on write ecosystem
b. List and explain the different data stores used in data science.
c. Write a python program to Fill All Missing Values with the Mean of the particular column.
d. Explain the use of correlation and forecasting in data science.
e. Explain importance of Graphics and Picture in report.
