Duration: 3hrs   **[Max Marks:80]**

N.B. : (1) Question No 1 is Compulsory.
      (2) Attempt any three questions out of the remaining five.
      (3) All questions carry equal marks.
      (4) Assume suitable data, if required and state it clearly.

1	Attempt any **FOUR**	[20]
a	What are different facets of data science?	5
b	Explain different data cleaning and data transformation techniques.	5
c	What is the difference between business intelligence and data science?	5
d	Explain the various Data science tools.	5
e	Explain numerical and categorical variables.	5
2 a	Elaborate and explain all the steps of the Data Science Process.	10
b	Explain Distributed File system and Hadoop.	10
3 a	Explain applications of machine learning in data science.	10
b	Explain SVM and decision tree Algorithm.	10
4 a	There are 2 stocks X and Y. Their share prices on particular days are as follows.	10

Sr No	Stock X	Stock Y
1	58	7
2	50	8
3	53	8
4	45	9
5	60	5

Find out the (R^2) correlation coefficient from the given data.

b	Compare NoSQL database and traditional RDBMS. Explain the architecture of Graph based NoSQL databases.	10
5 a	Explain CAP Theorem, BASE principles and applications of NoSQL databases.	10
b	Explain fraud detection and stock price prediction in detail.	10
6 a	Explain Recommendation system in detail.	10
b	Write short note on any two :	10

      1. Naive Bayes Algorithm
      2. Sentiment analysis
      3. Information Gain and Entropy

****

3564192C60C540B03A2807A9D2A37A6A