

com. S

6th Sem./ CSE / 2022(S)
Th-4 Data Science & Analytics

Full Marks: 80

Time- 3 Hrs

Answer any five Questions including Q No.1& 2
Figures in the right hand margin indicates marks

1. Answer **All** questions 2 x 10
 - a. Differentiate between data science & data analysis.
 - b. Name the popular toolkits used in data science.
 - c. What do you mean by Bokeh?
 - d. What are the different data types used in python programming?
 - e. Define data visualization.
 - f. Define metadata.
 - g. What is R programming ?
 - h. What is data lake?
 - i. Why SPSS is important?
 - j. How to measure the central tendency?
2. Answer **Any Six** Questions 6 x 5
 - a. Explain the different methods of data repository.
 - b. Describe the different types of data used in data science process.
 - c. Explain the salient features of data visualization.
 - d. State and explain the types of statistical data.
 - e. Discuss about the data collection method.
 - f. Explain about the normal distribution curve.
 - g. Illustrate the different analysis techniques used in data science.
3. How data science process works? Explain in detail. 10
4. Discuss the different methods used in data visualization? 10
5. Explain the application of data science technologies for visualization. 10
6. Describe the challenges of Big data visualization. 10
7. How data can be explore & import using IBM SPSS? Explain. 10