EPCET LIBRARY BANGALORE-49

CBCS SCHEME

USN | | | | | | | | | | |

21CS/IS644

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2026
Data Science and Visualization

Time: 3 hrs. Max. Marks: 100

Note: ***Answer any FIVE full questions, choosing ONE full question from each module.***

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.
2. Any revealing of identification, appeal to evaluator and /or equations written eg, 42+8 = 50, will be treated as malpractice.

Module-1

1 a. Interpret the hype about Data Science, Datafication and current landscape. **(08 Marks)**

b. Explain the Skill Set needed for Data Scientist profile. **(08 Marks)**

c. Explain the Data Scientist in terms of academia and industry. **(04 Marks)**

OR

2 a. Summarize the definition of Data Science and the components of the Conway Venn diagram. **(10 Marks)**

b. Explain with example the following with respect to Data Science.

i) Statistical Inference
ii) Population
iii) Sample **(10 Marks)**

Module-2

3 a. Illustrate Data Science process with a neat diagram. **(10 Marks)**

b. Explain the important steps carried out during Exploratory Data Analysis. **(10 Marks)**

OR

4 a. Comprehend the real direct case study and highlight the major issues addressed. **(06 Marks)**

b. Explain the three primary ways to extend the linear regression beyond the least squares. **(06 Marks)**

c. Explain the evaluation metric process in KNN. Given the dataset with income represented in Thousand.

Age	Income	Credit
69	3	Low
66	57	Low
49	79	Low
49	17	Low
58	26	High
44	71	High

Predict the credit rating whose age is 57 year and Income is $37,000? **(08 Marks)**

Module-3

5 a. Define feature generation. Explain Filter and Wrapper method feature generation algorithm and also selecting criteria in feature selection. **(10 Marks)**

b. Write Decision Tree algorithm in detail. Explain Decision Tree for Chasing Dragon's problem. **(10 Marks)**

OR

6 a. Explain Bagging and Boosting in Random Forest and also importance of Hyper parameterization with example. **(10 Marks)**

b. What is Dimensionality problem? Explain Singular Value Decomposition (SVD) in detail. **(10 Marks)**

Module-4

7 a. Explain the steps involved in the Data wrangling process with neat diagram. **(10 Marks)**

b. Compare Line Bar and Radar Charts along with their uses and design practices with example. **(10 Marks)**

OR

8 a. What are Scatter Plots? Explain in detail Characteristics of Scatter Plots. **(06 Marks)**

b. Demonstrate with a Python Code to customize the bubble plot with the following:

i) Shape
ii) Edges **(06 Marks)**

c. Discuss various Geoplots with their uses and design practices with example. **(08 Marks)**

Module-5

9 a. Describe the purpose of Matplotlib in data visualization and also create a basic line plot using Matplotlib with example. **(06 Marks)**

b. Explain Legend functions with syntax and example. **(06 Marks)**

c. Write Syntax for the following with respect to pyplot:

i) Adding X Label & Y label
ii) Creating Figures and Format strings
iii) Pandas Data Frames
iv) Subplots **(08 Marks)**

OR

10 a. Develop a histogram with Specific color, edge color and line width. **(06 Marks)**

b. Apply functions of stacked area chart to write a python program and demonstrate stacked area chart with suitable lable in X axis, Y-axis, title gridline
(Assume data Values) **(06 Marks)**

c. Explain the following with syntax and example with respect to Matplotlib :

i) Basic Image Operations
ii) How to write Mathematical Expressions. **(08 Marks)**

* * * * *