



Total No. of Questions: 4/18

Total No. of Pages: 2

SECOND YEAR (SYBCA (Science))
BCA23204: Data Science
(Semester III)

Program: BCA Code:BCASc08
Program Specific: BCA (Science)
Course Type: Minor
Paper: -

Credits: 2
Time: 2 Hours
Max. Marks: 30
SET: A

Instructions to the candidate:

- 1) All questions are compulsory.
- 2) Figures to the right indicate full marks.
- 3) Draw a well labelled diagram wherever necessary.

SECTION: A

Q1) Answer the following

[5 X 1= 5]

1. Define statistical data analysis.
2. What is compressed data?
3. State Data cube Aggregation.
4. Define noisy data.
5. List tools for data scientist.

SECTION: B

Q2) Answer the following (Attempt any 5)

[5 X 2 =10]

1. Define Binarizing and also write python code for same.
2. Write formula of dissimilarity for symmetric and asymmetric binary attributes.
3. What are different sources of data in data science?
4. Define scatter plot and write python code for same.
5. Explain standard dataset with one example.
6. List application of outlier.
7. Short note on Dimensionality reduction.

SECTION: C

Q3) Answer the following (Attempt any 2)

[2 X 5 = 10]

1. What is outlier? Also explain methods for outlier detection
2. Write the python code to create a simple Venn diagram for the following two sets:
 $X = \{1, 3, 5, 7, 9,\}$ $Y = \{2, 3, 4, 5, 6,\}$
3. Differentiate between equal – width binning and equal - frequency binning.
4. Explain any five data formats in brief.

SECTION: D

Q4) Answer the following (Attempt any 1)

[5 X 1 = 5]

1. What are measures of central tendency? Explain each in detail
2. Explain the following data transformation techniques:
 - a) Rescaling data
 - b) Normalizing data
 - c) Binarizing data
 - d) Standardizing data
 - e) Label encoding
