



Progressive Education Society's
Modern College of Arts, Science & Commerce Ganeshkhind, Pune – 16
(Autonomous)
End Semester Examination: OCT / NOV 2024
Faculty: Science and Technology

Program: BSc(Computer Science) Semester:
Program (Specific): (Computer science)
Class: TYBSc (Computer Science)
Name of the Course: Foundation of Data Science
Course Code: 24CS354
Paper: I

SET: A
Course Type: DSEC-II
Max.Marks: 35

Time: 2Hr

Instructions to the candidate:

- 1) *There are 4 sections in the question paper. Write each section on separate page.*
- 2) *All Sections are compulsory.*
- 3) *Figures to the right indicate full marks.*
- 4) *Draw a well labeled diagram wherever necessary.*

SECTION: A

Q1) Multiple choice question

[5 x 1 = 5]

1. Factors comprising the data quality including.
 - a) Accuracy c) completeness
 - b) Consistency d) All of mentioned
2. Which file is text file in which the values in the columns are sperated by a comma?
 - a) csv c) json
 - b) html d) xlsx
3. The dispersion of data values can be measured.
 - a) Range c) standard deviation
 - b) Variance d) All of the mentioned
4. Which charts display data as a cluster of circles?
 - a) Bubble c) bar
 - b) Scatter d) line
5. Which is property of an object?
 - a) Attribute value c) parameter
 - b) Argument d) value

Q2) Answer any 4 of the following in one sentence.

[4 x 1 = 4]

1. What do you mean by Data Quality?
2. Define visual encoding.
3. Define outlier.
4. State one hot coding.
5. Describe the methods of feature selection.
6. Give examples of semi structured data.



Progressive Education Society's
Modern College of Arts, Science & Commerce Ganeshkhind, Pune – 16
(Autonomous)
End Semester Examination: OCT / NOV 2024
Faculty: Science and Technology

SECTION: B

Q3) Answer any 4 of the following

[4 x 2 = 8]

1. Explain any two tools in data scientist tool box.
2. Explain null and alternate hypothesis.
3. Explain any two data visualization tools.
4. What is data cleaning? Explain any two data cleaning methods.
5. Compare between structured & unstructured data.
6. State the different methods of feature extraction.

SECTION: C

Q4) Answer any 4 of the following.

[4 x 2 = 8]

1. What is Data Discretization?
2. Write advantages and disadvantages of tag cloud.
3. List all applications of Data Science.
4. What is binary attribute and ordinal attribute?
5. Explain donut chart with diagram.
6. Explain data cubes.

SECTION: D

Q5) Answer any 2 of the following

[2 x 5 = 10]

1. Explain all Data Visualization Libraries.
2. Illustrate the different types of data. Explain in detail with appropriate example.
3. Explain following data transformation technique: Rescaling, Normalizing.
4. Find mean, median, mode & range for following list of values:

13, 18, 13, 14, 13, 16, 14, 21, 13