

Total No. of Questions : 5]

SEAT No. :

PA-1026

[Total No. of Pages : 2

[5902]-54

T.Y. BSc.

COMPUTER SCIENCE

CS - 354 : Foundation of Datascience

(2019 Pattern) (CBCS) (Semester - V)

Time : 2 Hours]

[Max. Marks : 35

Instructions to the candidates:

- 1) *Figures to the right indicate full marks.*
- 2) *All questions are compulsory.*
- 3) *Neat labelled diagrams must be drawn wherever necessary.*

Q1) Attempt any EIGHT of the following :

[8×1=8]

- a) What is Data science?
- b) Define Data source?
- c) What is missing values?
- d) List the visualization libraries in python.
- e) List applications of data science.
- f) What is data transformation?
- g) Define Hypothesis Testing?
- h) What is use of Bubble plot?
- i) Define Data cleaning?
- j) Define standard deviation?

Q2) Attempt any FOUR of the following.

[4×2=8]

- a) List the tools for data scientist.
- b) Define statistical data analysis?

P.T.O.

- c) What is data cube?
- d) Give the purpose of data preprocessing?
- e) What is the purpose of data visualization?

Q3) Attempt any two of the following. **[2×4=8]**

- a) What are the measures of central tendency? Explain any two of them in brief.
- b) What are the various types of data available? Give example of each?
- c) What is venn diagram? How to create it? Explain with example.

Q4) Attempt any two of the following. **[2×4=8]**

- a) Explain different data formats in brief.
- b) What is data quality? Which factors are affected data qualities?
- c) Write details notes on basic data visualization tools?

Q5) Attempt any ONE of the following. **[1×3=3]**

- a) What is outlier? State types of outliers.
- b) State and explain any three data transformation techniques.

