Total No. of Questions : 5]

SEAT No. : ____

P-1303

[Total No. of Pages : 2

[6055]-304

T.Y. B.Sc.

COMPUTER SCIENCE

CS - 354 : Foundations of Data Science

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

Time : 2 Hours] ***[Max. Marks : 35***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***

Q1) **Attempt any Eight of the following :** **[8 × 1 = 8]**

a) List any two application of Data Science.

b) What is outlier?

c) What is missing values?

d) Define variance.

e) What is nominal attribute?

f) What is data transformation?

g) What is one hot coding?

h) What is the use of Bubble plot?

i) Define Data visualisation.

j) Define Standard deviation?

Q2) **Attempt <u>any four</u> of the following :** **[4 × 2 = 8]**

a) Differentiate structured and Unstructured Data.

b) What is inferential statistics?

c) What do you mean by data preprocessing?

d) Define data discretization.

e) What is visual encoding?

P.T.O.

Q3) **Attempt <u>any two</u> of the following :** **[2 × 4 = 8]**

a) Explain outlier detection methods in brief.

b) Write different data visualization libraries in python.

c) What is data cleaning? Explain any two data cleaning methods.

Q4) **Attempt <u>any two</u> of the following :** **[2 × 4 = 8]**

a) Explain 3V's of Data Science.

b) Explain data cube aggregation method in detail.

c) Explain any two data transformation technique in detail.

Q5) **Attempt <u>any one</u> of the following :** **[1 × 3 = 3]**

a) Write a short note on feature extraction.

b) Explain Exploratory Data Analysis (EDA) in detail.