[6237]-504

T.Y. B.Sc.

COMPUTER SCIENCE

CS-354 : Foundation of Data Science

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

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) What is word cloud?

b) Why data cleaning is important operation of data preprocessing?

c) Define Statistical data analysis?

d) What is the purpose of data Visualization?

e) List any two applications of data Science.

f) What do you mean by data transformation?

g) What is Visual encoding?

h) Define Bubble plot?

i) What is CSV format?

j) Define Standard derivation?

Q2) Attempt any four of the following. [4×2=8]

a) What is 3V's?

b) What are the uses of XML files?

c) Calculate mean, mode, median and range:

22, 24, 35, 47, 23, 23, 31, 25

d) Why data preprocessing is important before data analysis?

e) List different types of data attributes with example.

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

a) What do you mean by Data Reduction? Explain any two Data Reduction strategies.

b) What do you mean by hypothesis testing? Explain null and alternate hypothesis.

c) How to visualize geospatial data? Explain in detail.

Q4) Attempt any Two of the following: [2×4=8]

a) What are the components of data scientist tool box? Explain two of them in detail.

b) What is Outlier? Explain types of outlieas.

c) What are the various types of data? Explain in detail.

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

a) What do you mean by data discretization? Explain discretization by Histogram analysis.

b) Write short note on different methods fer measuring data simillarity and Dissimilarity.

◀▶ ◀▶ ◀▶