

Total No. of Questions : 12]

SEAT No. :

P-3500

[Total No. of Pages : 2

[6005]-662

F.Y. M.E. (Computer/A.I & D.S)

**BASICS OF DATA SCIENCE**

**(2017 Pattern) (Semester - I) (510302)**

*Time : 3 Hours]*

*[Max. Marks : 50*

*Instructions to the candidates :*

- 1) Solve Q.1 or Q.2, Q.3 or Q.4, Q.5 or Q.6, Q.7 or Q.8, Q.9 or Q.10, Q.11 or Q.12.
- 2) Neat diagrams must be drawn wherever necessary.
- 3) Figures to the right indicate full marks.
- 4) Assume Suitable data, if necessary.

**Q1)** a) What is Data Science? Elaborate importance of data science. [5]

b) Explain : [4]

- i) Structured vs. Unstructured Data
- ii) Quantitative vs. Categorical Data

OR

**Q2)** a) Explain in detail the Data science process. [5]

b) Explain the Role of Data Scientist in Machine Learning. [4]

**Q3)** How does Exploratory Data Analysis help in analyzing the data? [8]

OR

**Q4)** What is Pearson's Correlation? Find the Pearson's correlation coefficient for the following data : [8]

No.	X	Y
1	43	99
2	21	65
3	25	79
4	42	75
5	57	87
6	59	81

P.T.O.

- Q5)** a) What is Linear Regression? Find a linear regression equation for the following sets of data : [5]

x	2	4	6	8
y	6	14	10	20

- b) Explain K-nearest Neighbors(k-NN) algorithm with an example. [4]

OR

- Q6)** a) Explain K-means clustering algorithm with an example. [5]  
b) Differentiate between Naive Bayes and k-NN algorithm. [4]

- Q7)** a) Explain types of data visualization. [4]  
b) What is data visualization and why it is important? [4]

OR

- Q8)** a) Explain various methods of encoding. [4]  
b) Explain following bivariate data visualization techniques. [4]  
i) Line plot  
ii) Bar plot

- Q9)** What is Collaborative Filtering? Explain how to compute the missing value of the utility matrix. [8]

OR

- Q10)** What is a Content-based Recommendation System? Explain the advantages and disadvantages of content-based recommendation system. [8]

- Q11)** Explain Social Network Graph. Also explain the varieties of Social Networks. [8]

OR

- Q12)** What are various methods to achieve the Clustering of Social-Network Graphs? [8]

