

Duration: 3hrs**[Max Marks:80]**

- N.B.: (1) Question No 1 is Compulsory.
 (2) Attempt any three questions out of the remaining five.
 (3) All questions carry equal marks.
 (4) Assume suitable data, if required and state it clearly.

- 1 a By using matrices, solve the following system of linear equation [5]
 $x + y + z = 3, x + 2y + 3z = 4, x + 4y + 9z = 6$
 c Enumerate the different techniques used in data cleaning. [5]
 d Describe feature engineering [5]
 e Explain the need of Dimension Reduction Algorithm [5]

- 2 a Find the singular value decomposition of the matrix $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$ [10]
 b In a manufacturing industry, item supplied on a critical raw material have been [10]
 received in ten lots against a monthly order of 200 units. Test at 5% level of
 significance whether the supplies are uniform (Chi Square method)

Lot No	1	2	3	4	5	6	7	8	9	10
No of Units	25	12	21	18	14	20	30	15	22	23

- 3 a Describe stem and leaf plot. Display data in stem and leaf plot the following [10]
 20 student's right data (cm).
 143, 163, 154, 159, 172, 165, 162, 171, 146, 165, 176, 145, 165, 182, 175, 186,
 160, 158, 167, 172. Find mode.
 b i) Explain scatter plot with example [5]
 ii) Differentiate between simple, random sampling, stratified random sampling. [5]

- 4 a What is a Graph? Explain any four types of Graphs along with its uses. [10]
 b Explain types of data. Compare and contrast the quantitative and qualitative data. [10]

- 5 a Partition the given data into 4 bins using Equi-depth binning method and [10]
 perform smoothing according to the following method

- i) Smoothing by bin mean
 ii) Smoothing by bin median
 iii) Smoothing by bin boundaries

Data : 11, 13, 13, 15, 15, 16, 19, 20, 20, 20, 21, 21, 22, 23, 24, 30, 40, 45, 45, 45,
 71, 72, 73, 75

- b Minimize $f(x_1, x_2) = 4x_1 - 2x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$. Starting from point [10]
 $x_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
 6 a What is feature mapping in dimensionality reduction? Explain any 5 techniques [10]
 of feature mapping.
 b Compute the standard deviation for the following sets [10]
 Set 1 = {0, 8, 12, 20} Set 2 = {8, 9, 11, 12}
