

JAMIA HAMDARD
Department of Computer Science & Engineering
School of Engineering Sciences & Technology
New Delhi-110062

Examination: Even Semester Examination 2023 (Supplementary)

Paper Name: Data Preparation and Analysis
Time: 3 Hours

Paper Code: MTCSE PE 231 (Suppl)
Maximum Marks: 75

*Note: Mobile Phones, or programmable calculators, or any equipment with memory are not allowed.
Follow proper numbering for answering the questions.*

SECTION A

7*5=35

Section A is compulsory and each question carries 7 marks

Q.No	Question	CO	BL
1	What are the various file formats used for data storage? Explain the significance of transformation techniques.	CO1	2
2	Explain data visualization. What are the methods to visualize the data?	CO5	2
3	Imagine you were building a product to help people buy mobile phone online. How to generate and validate the hypothesis for mobile purchase module?	CO4	6
4	What is clustering? Perform k means clustering for the following data (1,3), (2,1), (4,3), (3,6), (7,2) Just do up to two iterations and state the termination condition.	CO4	3
5	For a given data, explain how the following terms: a) Skewness, Left skewness, and Right skewness. b) Kurtosis c) What does skewness tell us about data? Use a Diagram to explain the above.	CO4	5

SECTION B

Attempt any **four** questions. Each question carries 10 marks

Q.No	Question	CO	BL
1	The real data in the world is not in the form to be directly processed. Why do we say so? Give valid explanations as per problems faced in real-life.	CO3	5
2	Explain the various normalization techniques.	CO3	6
3	What is meant by missing data? How to handle missing data in real-time.	CO2+ CO3	4
4	What do we understand when we say that the data is heterogeneous in nature. Explain any three data collection techniques.	CO2	5
5	Answer the following: i) Define correlation ii) Is correlation and covariance the same concept or different? Justify.	CO5	2,3

	iii) If A and B are positively correlated, what does it mean? iv) If A and B are negatively correlated, what does it mean? v) Does a correlation of zero mean that there exists no dependency? Justify		
6	Explain the applications of data preparation and analysis in healthcare, and marketing?	CO6	2