

Code No: **R20DS501****MALLA REDDY COLLEGE OF ENGINEERING & TECHNOLOGY****(Autonomous Institution – UGC, Govt. of India)****II B.Tech I Semester Supplementary Examinations, June 2025****Introduction to Data Science****(CSE-DS & B.Tech-AIDS)**

Roll No									
----------------	--	--	--	--	--	--	--	--	--

Time: 3 hours**Max. Marks: 70**

Note: This question paper Consists of 5 Sections. Answer **FIVE** Questions, Choosing **ONE** Question from each **SECTION** and each Question carries 14 marks.

SECTION-I

- | | | | BCLL | CO(s) | Marks |
|----------|----------|--|-------------|--------------|--------------|
| 1 | A | Explain the role of Python in data science, focusing on essential libraries used by data scientists. | L2 | CO-I | [7M] |
| | B | Describe various data operations in Python, such as filtering, grouping, and sorting, with examples. | L1 | CO-I | [7M] |

OR

- | | | | | | |
|----------|----------|---|-----------|-------------|-------------|
| 2 | A | Discuss the significance of an Integrated Development Environment (IDE) for data science workflows. | L1 | CO-I | [7M] |
| | B | Explain the data manipulation process and its importance in data science. Include an example of data rearrangement. | L2 | CO-I | [7M] |

SECTION-II

- | | | | | | |
|----------|----------|--|-----------|--------------|-------------|
| 3 | A | What is hypothesis testing? Explain the role of confidence intervals and p-values in hypothesis testing. | L4 | CO-II | [7M] |
| | B | Differentiate between sample mean and estimated mean. When is each used in data analysis? | L4 | CO-II | [7M] |

OR

- | | | | | | |
|----------|----------|--|-----------|--------------|-------------|
| 4 | A | Describe the frequency approach in statistical inference and its applications in data science. | L3 | CO-II | [7M] |
| | B | Explain how data preparation aids in enhancing data quality for analysis and modeling. | L2 | CO-II | [7M] |

SECTION-III

- | | | | | | |
|----------|----------|--|-----------|---------------|-------------|
| 5 | A | Define supervised learning and discuss its significance in machine learning. | L1 | CO-III | [7M] |
| | B | Explain the concept of training, validation, and test datasets in supervised learning. | L1 | CO-III | [7M] |

OR

- | | | | | | |
|----------|----------|--|-----------|---------------|-------------|
| 6 | A | Describe learning curves and discuss their role in evaluating model performance. | L3 | CO-III | [7M] |
| | B | Explain the support vector machine (SVM) algorithm and its applications in classification. | L2 | CO-III | [7M] |

SECTION-IV

- | | | | | | |
|----------|----------|--|-----------|--------------|-------------|
| 7 | A | Describe a case study where regression analysis is used to predict outcomes. | L3 | CO-IV | [7M] |
| | B | Compare and contrast different types of regression techniques | L2 | CO-IV | [7M] |

used in data science.

OR

- | | | | | | |
|----------|-----------------|---|-----------|--------------|-------------|
| 8 | <i>A</i> | Explain the process and importance of using clustering in identifying patterns within data. | L2 | CO-IV | [7M] |
| | <i>B</i> | Discuss the key steps in applying unsupervised learning to a real-world dataset. | L1 | CO-IV | [7M] |

SECTION-V

- | | | | | | |
|----------|-----------------|--|-----------|-------------|-------------|
| 9 | <i>A</i> | Define network analysis and its applications in data science, particularly in social networks. | L3 | CO-V | [7M] |
| | <i>B</i> | Explain centrality in network analysis and discuss its different types. | L1 | CO-V | [7M] |

OR

- | | | | | | |
|-----------|-----------------|---|-----------|-------------|-------------|
| 10 | <i>A</i> | Describe PageRank and its importance in ranking web pages within a network. | L2 | CO-V | [7M] |
| | <i>B</i> | Explain ego-networks and their role in analyzing social relationships. | L4 | CO-V | [7M] |
