

**MADANAPALLE INSTITUTE OF TECHNOLOGY & SCIENCE, MADANAPALLE**  
(UGC-AUTONOMOUS INSTITUTION)**MCA II Year I Semester (R22) Supplementary End Semester Examinations –  
December 2025****FUNDAMENTALS OF DATA SCIENCE****Time: 3Hrs****Max Marks: 60**

Attempt all the questions. All parts of the question must be answered in one place only.

**In Q.no 1 to 5 answer either A or B only**

Q.No	Question	Marks	CO	BL
<b>Q.1(A)</b>	Briefly explain the steps in data science process with diagram	12M	1	2
<b>OR</b>				
<b>Q.1(B)</b>	Define outliers. How to Handle missing data and outliers using python? Explain with example.	12M	1	3
<b>Q.2(A)</b>	Explain data aggregation and summarization process using pandas and NumPy with a suitable example	12M	2	2
<b>OR</b>				
<b>Q.2(B)</b>	Explain hierarchical clustering algorithm with an example.	12M	2	2
<b>Q.3(A)</b>	How to handle missing data and duplicates in dataset using python libraries.	12M	3	3
<b>OR</b>				
<b>Q.3(B)</b>	Explain how to Deal with data inconsistencies and conflicts using python libraries.	12M	3	3
<b>Q.4(A)</b>	Explain different visualization tools in detail with an example.	12M	4	2
<b>OR</b>				
<b>Q.4(B)</b>	How to create geospatial visualization using GeoPandas library discuss with relevant python code.	12M	4	3
<b>Q.5(A)</b>	Describe Natural Language Processing. Explain the its preprocessing steps in details.	12M	5	3
<b>OR</b>				
<b>Q.5(B)</b>	Explain AR(p),MA(q) and ARIMA(p,d,q) models with example	12M	5	3

**\*\*\* END\*\*\***