

Register No :

--	--	--	--	--	--	--	--	--	--	--	--

Question Paper Code :

St. JOSEPH'S COLLEGE OF ENGINEERING, CHENNAI - 119
(An Autonomous Institution)

B.E./B.Tech. DEGREE END SEMESTER EXAMINATIONS, APRIL/MAY 2024

Third Semester
Artificial Intelligence and Machine Learning
ML1301 – DATA FOUNDATION
(Regulations 2021)

Time : Three Hours

Maximum : 100 Marks

Answer ALL questions

CO1	Explore the fundamental concepts of Data science
CO2	Understand Data Science Process and Tools of EDA
CO3	Address how Organizational structure's influence efficiency and effectiveness.
CO4	Analyze and Validate data using Spreadsheets and Tableau.
CO5	Think through the ethics incorporating privacy, data sharing and decision-making and build interactive dashboards for Business

BL - Bloom's Taxonomy Levels

(BL1- Remembering, BL2- Understanding, BL3 - Applying, BL4 - Analyzing, BL5 - Evaluating, BL6 - Creating)

PART - A (10 X 2 = 20 Marks)

Q.No	Questions	Marks	CO	BL
1.	Give the importance of Big Data.	2	CO1	BL1
2.	What is called Mining of Frequent Patterns?	2	CO1	BL1
3.	Discuss the issues that come with lot of anomalies in data, in connection with EDA.	2	CO2	BL2
4.	Outline about Imbalanced dataset?	2	CO2	BL2
5.	What are the benefits of Data Science Centers of Excellence?	2	CO3	BL2
6.	Identify the types of Tree data structure.	2	CO3	BL2
7.	What are the steps in filtering the data in Excel?	2	CO4	BL1
8.	Give the generic formula to VLOOKUP between two sheets.	2	CO4	BL1
9.	What are the 5 C's of Data Science?	2	CO5	BL1
10.	Define Data insights	2	CO5	BL1
PART - B (5 X 13 = 65 Marks)				
Q.No	Questions	Marks	CO	BL
11.a	How can we categorize data based on their types?	(13)	CO1	BL3
	OR			

11.b	What is a Distributed File System? Explain the various characteristic features of DFS in detail.	(13)	CO1	BL3
12.a	Why is Data Science important and what are the steps involved in the Data Science process?	(13)	CO2	BL4
	OR			
12.b	Why is data preprocessing an important step in data analysis? Explain with an example.	(13)	CO2	BL4
13.a	Show linear and non-linear data structures with an example of each?	(13)	CO3	BL3
	OR			
13.b	Explain Data Organizational Models in detail.	(13)	CO3	BL3
14.a	Explain the needs for Data Visualization and Compare the various Data Visualization Techniques in terms of Quantitative and Qualitative Data in Excel	(13)	CO4	BL4
	OR			
14.b	Outline the process of creating a dashboard in Tableau along with the steps involved in each stage?	(13)	CO4	BL4
15.a	Give the procedure to build a Machine Learning Model.	(13)	CO5	BL3
	OR			
15.b	Present about natural language processing (NLP) and its significance in the field of artificial intelligence?	(13)	CO5	BL3
	PART - C (1 X 15 = 15 Marks)			
Q.No	Questions	Marks	CO	BL
16.a	Give a Plan to create a dashboard for a superstore showing the sales and profits for different segments and sub-category of products across all states.	(15)	CO4	BL5 BL6
	OR			
16.b	Analyze the various steps involved in exploratory data analysis (EDA) by applying with a sample dataset, and explain the significance of each step?	(15)	CO2	BL5 BL6