

THE KAVERY ENGINEERING COLLEGE*(An Autonomous Institution)***B.E/B.TECH DEGREE END SEMESTER THEORY EXAMINATIONS, NOV/DEC 2025***Third Semester**Computer Science and Engineering***CS3352 / UCST24303 - Foundations of Data Science****(Common to Information Technology)***Regulations – 2021 & 2024**Duration: 3 Hrs**Maximum: 100 Marks***PART – A (ANSWER ALL QUESTIONS) (Marks 10*2=20)**

<i>Q.No</i>	<i>Questions</i>	<i>BLT</i>	<i>CO</i>	<i>Marks</i>
1.	Define Data Science.	L1	1	2
2.	What is the purpose of Exploratory Data Analysis (EDA)?	L1	1	2
3.	What is the difference between quantitative and qualitative data?	L1	2	2
4.	Define mean, median, and mode.	L1	2	2
5.	Define correlation.	L1	3	2
6.	What is multiple regression, and when is it used?	L2	3	2
7.	What is fancy indexing in NumPy?	L2	4	2
8.	What is the purpose of pivot tables in Pandas?	L2	4	2
9.	Define density and contour plots.	L2	5	2
10.	How do you add text and annotations to a plot?	L2	5	2

PART – B (ANSWER ALL QUESTIONS) (Marks 5*16 = 80)

<i>Q.No</i>	<i>Questions</i>	<i>BLT</i>	<i>CO</i>	<i>Marks</i>
11.	a i Explain the Data Science Process and its key stages.	L1	1	8
	ii Describe the importance of data preparation and retrieval in Data Science.	L1	1	8
	Or			
	b i Write short notes on Data Mining and Data Warehousing.	L1	1	8
	ii Explain different Basic Statistical Descriptions of Data with examples.	L1	1	8
12.	a i Explain the different types of data with examples.	L1	2	8
	ii Discuss the different measures of central tendency (Averages) in data.	L1	2	8
	Or			
	b i What is Normal Distribution? Explain its significance in statistics.	L1	2	8
	ii Explain the importance of data visualization using tables and graphs.	L1	2	8

13.	a	i	Explain the least squares regression line with an example.	L2	3	10
		ii	Describe the computational formula for correlation coefficient and its application.	L2	3	6
Or						
	b	i	Explain standard error of estimate and its role in regression analysis.	L2	3	10
		ii	What is the importance of R^2 (coefficient of determination) in regression analysis?	L2	3	6
14.	a		Implement aggregations and computations on NumPy arrays with examples.	L3	4	16
	Or					
	b		Implement Hierarchical indexing in Pandas and explain its significance.	L3	4	16
15.	a		Compare and analyze the use of line plots, scatter plots, and histograms for data visualization.	L4	5	16
	Or					
	b		Demonstrate three-dimensional plotting and analyze how it enhances data representation.	L4	5	16