

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

Q.No	Questions	BLT	CO	Marks
1.	What is data science?	L1	1	2
2.	State the importance of exploratory data analysis in the data science process.	L2	1	2
3.	Define categorical and numerical variables.	L1	2	2
4.	How do you describe variability in a dataset?	L2	2	2
5.	What is a correlation coefficient?	L1	3	2
6.	Differentiate between regression and correlation.	L2	3	2
7.	What are NumPy arrays?	L1	4	2
8.	What is the use of hierarchical indexing in Pandas?	L2	4	2
9.	What is a scatter plot used for?	L1	5	2
10.	Explain the role of legends and annotations in data visualization.	L2	5	2

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

Q.No	Questions	BLT	CO	Marks
11.	a Illustrate the steps involved in data preparation and its role in model building.	L2	1	16
	Or			
	b Explain data mining and data warehousing. How do they support data science applications?	L2	1	16
12.	a Classify various methods for representing data using tables and graphs.	L2	2	16

Or

	b	Demonstrate the measures of central tendency and variability with appropriate examples.	L2	2	16
13.	a	Identify the process of calculating the least squares regression line with a dataset.	L3	3	16
		Or			
	b	Apply multiple regression to model relationships between variables. Provide examples.	L3	3	16
14.	a	Illustrate data manipulation techniques in Pandas. How do you handle missing data?	L2	4	16
		Or			
	b	How do you summarize and combine datasets? Perform group-wise aggregation in Pandas?	L2	4	16
15.	a	Describe the purpose of line plots, histograms and density plots with reference to visualization libraries such as Matplotlib and Seaborn.	L2	5	16
		Or			
	b	Explain the concepts and applications of 3D data visualization and geographic plotting in Python.	L2	5	16