

THE KAVERY ENGINEERING COLLEGE*(An Autonomous Institution)*

UG/PG DEGREE END SEMESTER THEORY EXAMINATIONS, APRIL/MAY 2026

*Fourth Semester**Artificial Intelligence and Data Science*

AD3491 – FUNDAMENTALS OF DATA SCIENCE AND ANALYTICS

*Regulations - 2021**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.	List the characteristics of quality data.	L1	1	2
2.	Define Data Cleansing.	L1	1	2
3.	Represent the following data in stem- and-leaf display. 67, 74, 63, 88, 82, 97, 65, 79.	L3	2	2
4.	How the skewness of a data distribution can be identified?	L1	2	2
5.	Compare population and sample.	L2	3	2
6.	Define confidence interval (ci) for μ .	L1	3	2
7.	Define F-Test.	L1	4	2
8.	Compare dependent variables and explanatory variables.	L2	4	2
9.	Tell about forecasting.	L1	5	2
10.	List the use of auto correlation in time series analysis.	L1	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 Sketch and outline the step-by-step activities in data science process.	L2	1	8
	ii Explain the characteristics and benefits of data.	L2	1	8
	Or			
	b i Explain Data Science and Big Data.	L2	1	8
	ii Explain the different facts of data with examples.	L2	1	8
12.	a i Explain the steps to calculate IQR with an example.	L2	2	8
	ii Explain the 4 things to describe relationship between variables.	L2	2	8
	Or			
	b Explain about following measures and calculate them with given “residence changes” data. 1, 3, 4, 1, 0, 2, 5, 8, 0, 2, 3, 4, 7, 11, 0, 2, 3, 4	L2	2	16
	i. Variance			
	ii. Standard Deviation			
	iii. Inter Quartile Range (IQR)			
	iv. Z-Score			

13.	a	i	Explain sampling distribution and its types.	L2	3	8
		ii	Explain one tailed and two tailed test.	L2	3	8
			Or			
	b	i	Imagine a simple population consisting of only 5 observations: 2 4 6 8 10. List all possible sample of size two. Summarize relative frequency table showing the sample distribution of the mean.	L2	3	8
		ii	Calculate the value of the z test for each of the following situations. Also, given critical z score of +/- 1.96, calculate the critical confidence level.	L2	3	8
			i. X=12; σ=9; n=25; μhyp=15 ii. X=3600; σ=4000; n=100; μhyp=3500 iii. X=0.25; σ=010; n=36; μhyp=0.22			
14.	a	i	Estimate the calculation of Sum of Squares (Two-Factor ANOVA) with an example.	L5	4	8
		ii	10. Blood pressure of 8 patients are before and after are recorded: Before: 180,200,230, 240,170,190,200 and 165 After: 140,145, 150,155,120,130,140 and 130. Find, is there any significant difference between BP reading before and after by applying two-sample t-test.	L5	4	8
			Or			
	b	i	Evaluate about one factor ANOVA with example.	L5	4	8
		ii	Evaluate about the chi-square test with an example.	L5	4	8
	15.	a	i	Discover about time series analysis with example.	L4	5
		ii	Discover about multiple regression with an example.	L4	5	8
			Or			
b		i	Examine logistic regression model in predictive analysis with suitable example.	L4	5	8
		ii	Examine the role of weighted resample in linear model testing.	L4	5	8