Reg.No.:												

VIVEKANANDHA COLLEGE OF ENGINEERING FOR WOMEN
[AUTONOMOUS INSTITUTION AFFILIATED TO ANNA UNIVERSITY, CHENNAI]
Elayampalayam – 637 205, Tiruchengode, Namakkal Dt., Tamil Nadu.

Question Paper Code: 130015

B.E. / B.Tech. DEGREE END-SEMESTER EXAMINATIONS – NOV. / DEC. 2024
Sixth Semester
Computer Science and Technology
U19CT611 – FOUNDATIONS OF DATA SCIENCE
(Regulation 2019)

Time: Three Hours Maximum: 100 Marks

Answer ALL the questions

Knowledge Levels (KL)	K1 – Remembering	K3 – Applying	K5 - Evaluating
	K2 – Understanding	K4 – Analyzing	K6 - Creating

PART – A

(10 x 2 = 20 Marks)

Q.No.	Questions	Marks	KL	CO
1.	Define data science with an example.	2	K2	CO1
2.	What is datafication?	2	K1	CO1
3.	What are the different types of data collection strategies?	2	K1	CO2
4.	What is are exploring the data?	2	K2	CO2
5.	Find the Central Tendency of the given data set (2,4,5,7,8,3,4,5,6,7,9).	2	K2	CO3
6.	Write about ANOVA.	2	K1	CO3
7.	Define K-FOLD cross validation.	2	K1	CO4
8.	List out the properties of Clustering.	2	K1	CO4
9.	How to produce an effective presentation?	2	K2	CO5
10.	How will you plot multiple plots in one window?	2	K2	CO5

PART – B

(5 x 13 = 65 Marks)

Q.No.		Questions	Marks	KL	CO
11.	a)	Describe in detail about different Stages in a Data Science Project.	13	K2	CO1

(OR)

	b)	What are the applications of data science in various fields? Explain with real world examples.	13	K2	CO1
12.	a)	Why data preprocessing technique is important in Data Science and explain the working of various data preprocessing techniques.	13	K3	CO2
		(OR)			
	b)	Explain the concept of data integration transformation with examples.	13	K2	CO2
13.	a)	Explain about Basic Statistical Description of Data with examples.	13	K2	CO3
		(OR)			
	b)	What are the five parts involved in Whisker plot and Compute Whisker Plot for the given data set. (100,120,110,150,110,140,130,170,120,220,140,110)	13	K3	CO3
14.	a)	Illustrate sampling techniques for Modeling and Validation.	13	K2	CO4
		(OR)			
	b)	Write short notes on		K1	CO4
		i. Evaluating clustering models.	8		
		ii. Validating models.	5		
15.	a)	How will you Export the graph for the dataset in R-Programming?	13	K3	CO5
		(OR)			
	b)	What is multivariate data? How do you visualite display multivariate data? Explain.	13	K3	CO5

PART – C

(1 x 15 = 15Marks)

Q.No.		Questions	Marks	KL	CO
16.	a)	Calculate the central Tendency and dispersion of data for the given data set.	15	K3	CO3

x	66	68	68	70	71	72	72
y	68	70	69	72	72	72	74

(OR)

	b)	How will you produce presentations to the end users and clients in an effective way? Explain with examples.	15	K2	CO5