

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: 5011

B.E. / B.Tech. DEGREE END-SEMESTER EXAMINATIONS – MAY / JUNE 2024

Sixth Semester

Computer Science and Engineering

U19CSV32 – DATA SCIENCE AND ANALYTICS

(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.	What are the Benefits and uses of Data science?	2	K1	CO1
2.	Define Exploratory Data Analysis?	2	K1	CO1
3.	Brief the Conventional System?	2	K1	CO2
4.	Why the sampling distributions contain less variables when sample size is large?	2	K2	CO2
5.	State the Bayesian Modelling.	2	K1	CO3
6.	List out the regression models. Define linear regression.	2	K2	CO3
7.	What are the characteristics of Data streams?	2	K2	CO4
8.	What are the applications of Data Streams?	2	K2	CO4
9.	Define Data Visualization?	2	K1	CO5
10.	What is the use of Collective inferencing?	2	K2	CO5

PART – B

(5 x 13 = 65 Marks)

Q. No.	Questions	Marks	KL	CO
11. a)	i. Explain Cleaning, Integration and Transformation?	6	K2	CO1
	ii. Elaborate data science process in detail.	7	K2	CO1
(OR)				
b)	i. Explain the steps to build a model?	6	K2	CO1
	ii. Where can we use Facets? Explain any facet insights.	3+4	K2	CO1
12. a)	i. What is a sample? Why the samples are used and what are the specific types of samples?	6	K3	CO2
	ii. The sum of two numbers is S and their product is P . Find the 2 numbers using analytical method.	7	K4	CO2
(OR)				
b)	i. What is Analytic Process and Tools? Explain.	7	K2	CO2
	ii. The mean and standard deviation of the tax value of all vehicles registered in a certain state are $\mu = \text{Rs. } 13,525$ and $\sigma = \text{Rs. } 4,180$. Suppose random sample size 100 are drawn from the population of vehicles. What are the mean and standard deviation of the sample mean?	6	K4	CO2
13. a)	i. Explain the structure of neural networks.	6	K2	CO3
	ii. Illustrate the mathematical functions used in the data analysis process.	7	K3	CO3
(OR)				
b)	i. Describe Support vector machine.	6	K2	CO3
	ii. Sketch and elaborate the Architecture of Fuzzy logic control design (or) controller?	7	K2	CO3
14. a)	i. Determine the tail length for each stream element and the resulting estimate of the number of distinct elements if the hash function is: a. $h(x) = 2x + 1 \text{ mod } 32$. b. $h(x) = 3x + 7 \text{ mod } 32$. c. $h(x) = 4x \text{ mod } 32$.	2+2+2	K3	CO4
	ii. Explain Sampling method for a fixed propagation?	7	K4	CO4
(OR)				
b)	i. What do you mean by Real time analytics platform? Explain.	7	K2	CO4
	ii. Based on what you know, plan how would you partition the following bit stream into buckets 1001011011101? Find all of them?	6	K4	CO4

15. a) What are the common measures used in Egocentric network Analysis. Discuss Alter Attributes with examples? Analyse what are the visualization techniques used to visualizing data. 7 K2 CO5
6 K4 CO5

(OR)

- b) i. Elaborate the limitations and challenges of social network. 7 K3 CO5
ii. Develop a visualization technique to represent the following data. 2+2+2 K3 CO5
a. 2D data
b. multi-dimensional data
c. pyramid-type data.

PART – C

(1 x 15 = 15 Marks)

Q. No.	Questions	Marks	KL	CO
16. a)	State the definition of stock market prediction? Explain any case study of real time stock market prediction with use case diagrams.	15	K6	CO4
(OR)				
b)	Brief the importance of social network analysis? Explain the social network analysis with a use case diagram.	15	K6	CO5