



K. Sambal

Avinashilingam Institute for Home Science and Higher Education for Women
Deemed to be University Estd. u/s 3 of UGC Act 1956, Category A by MHRD (now MoE)
Re-accredited with A++ Grade by NAAC. CGPA 3.65/4, Category I by UGC
Coimbatore - 641 043, Tamil Nadu, India

Bachelor's Degree Examination – November 2025
V Semester

Class : III UG
Major : Computer Science

Time: 1^{1/2} Hours
Max. Marks: 50

23BCSDE2 Fundamentals of Data Science

Course Outcomes:

- CO1: Ability to apply the basic Data Science Knowledge on various problems.
CO2: Contribution of the knowledge gained on choosing and evaluation on different models.
CO3: Understanding the application and use of regression Techniques.
CO4: Attaining knowledge to identify the application areas of various unsupervised methods.
CO5: Analysing the documents for final deployment and effective presentation.

Part A

5 x 1 = 5

Choose the Correct Answer

1. Which of the following is not part of data science lifecycle? CO1 K1
a. Data Preparation b. Deployment
c. Marketing Strategy d. Modeling
2. A graphical representation of data set is referred as _____. CO2 K2
a. Sampling b. Modeling
c. Visualization d. Grouping
3. What is the command used in R to build the logistic regression model? CO3 K2
a. glm() b. gm() c. lm() d. lmn()
4. What technique can be used to improve the efficiency of apriori algorithm? CO4 K2
a. Sampling b. Cleaning
c. Hash based technique d. Transaction increase
5. Data Scientist are interested in _____ approach of a data science project. CO5 K2
a. Structural b. Modelling
c. Technique d. Pattern-based

Part B

5 x 5 = 25

Answer ALL questions

Each answer should not exceed 400 words or two pages

- 6.a. Write a note on the roles and Stages in Data Science process. CO1 K1
(or)
- 6.b. Elaborate the phases in managing and Exploring Data. CO1 K2
- 7.a. Brief the views on Modeling and Memorization methods. CO2 K2
(or)
- 7.b. Execute the steps in Loading Data into R. CO2 K3
- 8.a. Summarize the procedure on reading the model summary and characterizing the coefficient quality in Linear Regression. CO3 K1
(or)
- 8.b. Highlight the features on Linear Regression. CO3 K2
- 9.a. Express the views on hierarchical clustering with hclust(). CO4 K2
(or)
- 9.b. Explain the steps for K-means algorithm with suitable example. CO4 K3
- 10.a. Discuss about using knitr to produce milestone documentation. CO5 K2
(or)
- 10.b. Explain about presenting the result to the project sponsor. CO5 K2

Part C

2 x 10 = 20

Answer ALL questions

Each answer should not exceed 800 words or four pages

- 11.a. Illustrate on the significance and process of Deployment and Maintenance. CO1 K1
(or)
- 11.b. Explain on the important features for good data Visualization. CO2 K2