Register No : _____

THE KAVERY ENGINEERING COLLEGE

(An Autonomous Institution)

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

Sixth Semester

Agricultural Engineering

OCS353 - DATA SCIENCE FUNDAMENTALS

Regulations – 2021

Duration : 3 Hrs ***Maximum : 100 Marks***

PART – A (ANSWER ALL QUESTIONS) (Marks 10*2=20)

Q.No	*Questions*	*BLT*	*CO*	*Marks*
1.	Interpret the role of data science in technology.	L2	1	2
2.	How are organizations using the information from data warehouses?	L2	1	2
3.	List the aggregate functions used for data manipulation with example.	L1	2	2
4.	Compare the use of loc and iloc with an example.	L2	2	2
5.	Differentiate Supervised and Unsupervised Machine learning techniques.	L2	3	2
6.	Why is outlier detection important in data analysis?	L1	3	2
7.	Write a python code to draw a simple line plot using matplotlib.	L2	4	2
8.	Write the source code to draw a simple histogram.	L2	4	2
9.	Identify the tools and techniques commonly used for processing large datasets.	L2	5	2
10.	What are the basic steps involved in preparing data for building a recommender system?	L1	5	2

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

Q.No			*Questions*	*BLT*	*CO*	*Marks*
11.	a		Briefly explain the steps in data science process with diagram.	L2	1	16
			Or			
	b		Elaborate the architecture of a typical Data Mining System with a neat diagram.	L2	1	16
12.	a	i	Demonstrate how pandas dataframe can be constructed? Explain with an example.	L2	2	8
		ii	Write a numpy program to construct an array by repeating k times.	L2	2	8
			Or			
	b	i	How broadcasting is applied when performing the following operation and explain the rules with the below code. np.arange(3)+5 np.ones(3,3)+ np.arange(3)	L2	2	8
		ii	Discuss the steps to handle the missing data in Pandas with example.	L2	2	8
13.	a		Apply the concept of supervised learning in a case where a hospital wants to predict patient readmission rates based on historical data.	L3	3	16
			Or			
	b		Make use of an example explain K means clustering algorithm.	L3	3	16

14.	a	i	Write a simple python program that draws scatter plot where x = [1,2,3,4] and y = [1,4,9,16] and give labels, annotations, and legends to the graph.	L3	4	8
		ii	Write a python code to visualize the three-dimension function using contour plots.	L3	4	8
			Or			
	b	i	How graphical data can be projected using matplotlib? Explain with example.	L3	4	8
		ii	Explain how error bars can be used to represent uncertainty in data visualization.	L3	4	8
15.	a		Design a pipeline to automate the entire process of data preparation, model building, and deployment for a malicious URL prediction system.	L3	5	16
			Or			
	b		Design a recommender system for an e-commerce platform that integrates user behaviour data, product features, and demographic information.	L3	5	16

****