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: 6008

B.E. / B.Tech. DEGREE END-SEMESTER EXAMINATIONS – MAY / JUNE 2024
Seventh Semester
Information Technology
U19IT724 – BIG DATA 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.	Distinguish between Big Data and Big data Analytics.	2	K2	CO1
2.	Tell the significance of Data Science.	2	K2	CO1
3.	List out the characteristics of NoSQL.	2	K1	CO2
4.	Write the applications of Hadoop YARN.	2	K1	CO2
5.	Give the meaning of the following command: Select first_name, salary, date_of_join from employees where designation="Manager";	2	K2	CO3
6.	What is significance of MongoDB in data analysis?	2	K1	CO3
7.	Identify the steps involved in process of Hive QL process engine.	2	K2	CO4
8.	List out the benefits of Apache Kafka.	2	K1	CO4
9.	Write a FILTER operator act in pig.	2	K2	CO5
10.	List out the execution modes of pig.	2	K1	CO5

PART – B

(5 x 13 = 65 Marks)

Q.No.		Questions	Marks	KL	CO
11.	a)	Explain various approaches to analysis of Data.	13	K2	CO1
		(OR)			
	b)	Summarize the top challenges faced in big data.	13	K2	CO1
12.	a)	Differentiate SQL and NoSQL with examples.	13	K2	CO2
		(OR)			
	b)	Illustrate the function Hadoop Distributed File System with a neat sketch.	13	K2	CO2
13.	a)	Using collection, design a table/column family to support the following requirements. 1. Store the basic information about student such as Student Roll. no, Student Name, Student Date of Birth, and Student Address. 2. Store the Subject Preferences of each student. There should be a minimum of two subject preferences and a maximum of four. The order of preferences as given by the student should be preserved 3. Store the hobbies of each student. There should be a minimum of two hobbies and a maximum of four. The hobbies as given by the student should be arranged in alphabetical order.	13	K3	CO3
		(OR)			
	b)	Illustrate the process of MongoDB Query Language with suitable example.	13	K3	CO3
14.	a)	Write a Map Reduce program to count the occurrence of similar words in file. Use partitioner to partition key based on alphabets Input data: Welcome to Hadoop Session Introduction to Hadoop Introduction Hive Hive Session Pig Session	13	K4	CO4
		(OR)			
	b)	Analyze and Explain how the Netflix uses Apache Kafka.	13	K4	CO4
15.	a)	Demonstrate the function of various HDFS commands.	13	K3	CO5
		(OR)			

	b)	Explain the use case of Retail log processing using Pig.	13	K3	CO5

PART – C

(1 x 15 = 15 Marks)

Q.No.		Questions	Marks	KL	CO
16.	a)	Write the insert method to store the following document in MongoDB. Name: "Isai" Address: {"City" :"Banglore","Street" : "HSR layout","Affilication": "ABC Ltd"} Hobbies: Chess,Base Ball.	15	K5	CO2
		(OR)			
	b)	Design a table/column family to support the following requirements using TTL. Store the login details of the user such as User ID and Password. The information stored should expire in a day's time.	15	K5	CO5