INSTITUTE OF AERONAUTICAL ENGINEERING
(Autonomous)
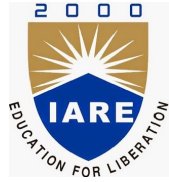Dundigal, Hyderabad - 500 043

MODEL QUESTION PAPER-II

B.Tech III Semester End Examinations, November 2020

Regulations: IARE - R18

DATA STRUCTURES

(COMMON TO ME/CSE/IT/ECE/CE)

Time: 3 hour   Maximum Marks: 70

Answer ONE Question from each MODULE
All Questions Carry Equal Marks
All parts of the question must be answered in one place only

MODULE-I

1. (a) Illustrate the Evolution of Big Data in detail? In perspective of Doug Laney and a Gartner analyst coined the term Big Data.

   (b) Explain the various applications of Big Data analytics and why this sudden hype around Big Data Analytics?   [**7m**]

2. (a) Summarize in detail the challenges of Big Data with a neat diagram?   [**7m**]

   (b) Recall the CAP theorem and how it is different from ACID properties in a distributed computing environment?   [**7m**]

MODULE-II

3. (a) Recall the introduction of Hadoop over Big Data and show what are the basic difference between traditional RDBMS and Hadoop?   [**7m**]

   (b) Write short note on Hadoop Ecosystem also explain various use cases involved in Hadoop. [**7m**]

4. (a) Explain the key distinctions of Hadoop which are very flexible to handle the huge volume of data?   [**7m**]

   (b) Interpret the integrated Hadoop systems offered by leading market vendors with Cloud-based Hadoop solutions.   [**7m**]

MODULE-III

5. (a) Explain in brief bout the Hadoops rack topology with the following terms: Rack Awareness and Fault Tolerance?   [**7m**] .

   (b) Discuss racks in Hadoop Cluster? Explain how Hadoop Clusters are arranged in several racks with a real time example?   [**7m**]

6. (a) Define HDFS? Discuss the HDFS Architecture and HDFS Commands in brief. Write down the goals of HDFS.   [**7m**]

(b) What is a heartbeat in HDFS and how Name Node handles data nodes failures? Explain [**7m**]

MODULE-IV

7. (a) Explain Map-reduce framework in brief and Draw the architectural diagram for Physical Organization of Compute Nodes. [**7m**]

   (b) Illustrate a Map Reduce programming model provide a way for reducers to communicate with each other? [**7m**]

8. (a) Demonstrate in detail about the Driver class, map and reducer phases with a real time example? [**7m**]

   (b) Demonstrate workflow in a basic word count MapReduce program to understand MapReduce Paradigm. [**7m**]

MODULE-V

9. (a) What is Apache Pig and Demonstrate in detail about the scalar and complex data types in PIG? [**7m**]

   (b) Express DDL concepts in detail? In Hive, explain the term aggregation and its uses? [**7m**]

10. (a) Illustrate PIG Latin language is case-sensitive or not? What does FOREACH do? [**7m**]

    (b) Describe the importance of partitions in Hive with an example? [**7m**]

END OF EXAMINATION

COURSE OBJECTIVES:
The course should enable the students to:

1	The scope and essentiality of Big Data and Business Analytics.
2	The technologies used to store, manage, and analyze big data in a Hadoop ecosystem.
3	The techniques and principles in big data analytics with scalability and streaming capability.
4	The hypothesis on the optimized business decisions in solving complex real-world problems.

COURSE OUTCOMES:
**After successful completion of the course, students should be able to:**

CO 1	Explain the evolution of big data with its characteristics and challenges with traditional business intelligence.
CO 2	Compare big data analysis and analytics in optimizing the business decisions.
CO 3	Classify the key issues and applications in intelligent business and scientific computing.
CO 4	Explain the big data technologies used to process and querying the big data in Hadoop, Map Reduce, Pig and Hive.
CO 5	Make use of appropriate components for processing, scheduling and knowledge extraction from large volumes in distributed Hadoop Ecosystem.
CO 6	Translate the data from traditional file system to HDFS for analyzing big data in Hadoop ecosystem.
CO 7	Develop a Map Reduce application for optimizing the jobs.
CO 8	Develop applications for handling huge volume of data using Pig Latin.
CO 9	Explain the importance of big data framework HIVE and its built-in functions, data types and services like DDL.
CO 10	Demonstrate business models and scientific computing paradigms, and tools for big data analytics.
CO 11	Categorize Hadoop components for developing real time big data analytics in various applications like recommender systems, social media applications etc.

MAPPING OF SEMESTER END EXAMINATION QUESTIONS TO COURSE OUTCOMES

Q.No		All Questions carry equal marks	Taxonomy	CO's	PO's
1	a	Illustrate the Evolution of Big Data in detail? In perspective of Doug Laney and a Gartner analyst coined the term "Big Data".1. Bubble sort 2. Selection sort	Understand	CO 1	PO 1
	b	Explain the various applications of Big Data analytics and why this sudden hype around Big Data Analytics?	Understand	CO 2	PO 1
2	a	Summarize in detail the challenges of Big Data with a neat diagram?	Understand	CO 1	PO 1
	b	Recall the CAP theorem and how it is different from ACID properties in a distributed computing environment?	Remember	CO 3	PO 1,2,3
3	a	Recall the introduction of Hadoop over Big Data and show what are the basic difference between traditional RDBMS and Hadoop?	Understand	CO 5	PO 1,2,3
	b	Write short note on Hadoop Ecosystem also explain various use cases involved in Hadoop.	Understand	CO 5	PO 1,2,3
4	a	Explain the key distinctions of Hadoop which are very flexible to handle the huge volume of data?	Understand	CO 5	PO 1,2,3
	b	Interpret the integrated Hadoop systems offered by leading market vendors with Cloud-based Hadoop solutions.	Understand	CO 4	PO 1,2,3
5	a	Explain in brief bout the Hadoop's rack topology with the following terms: Rack Awareness and Fault Tolerance?	Understand	CO 6	PO 1,2
	b	Discuss racks in Hadoop Cluster? Explain how Hadoop Clusters are arranged in several racks with a real time example?	Understand	CO 6	PO 1,2
6	a	Define HDFS? Discuss the HDFS Architecture and HDFS Commands in brief. Write down the goals of HDFS.	Apply	CO 6	PO 1,2
	b	What is a heartbeat in HDFS and how Name Node handles data nodes failures?	Apply	CO 6	PO 1,2
7	a	Explain Map-reduce framework in brief and Draw the architectural diagram for Physical Organization of Compute Nodes.	Understand	CO 7	PO 1-5

	b	Illustrate a MapReduce programming model provide a way for reducers to communicate with each other?	Understand	CO 7	PO 1-5
8	a	Demonstrate in detail about the Driver class, map and reducer phases with a real time example?	Understand	CO 7	PO 1-5
	b	Demonstrate workflow in a basic word count MapReduce program to understand MapReduce Paradigm.	Understand	CO 7	PO 1-5
9	a	What is Apache Pig and Demonstrate in detail about the scalar and complex data types in PIG?	Remember	CO 8	PO 1-5
	b	Express DDL concepts in detail? In Hive, explain the term aggregation and its uses?	Understand	CO 9	PO 1,2,3
10	a	Illustrate PIG Latin language is case-sensitive or not? What does FOREACH do?	Understand	CO 12	PO 1-5
	b	Describe the importance of partitions in Hive with an example?	Remember	CO 10	PO 1,2,3

KNOWLEDGE COMPETENCY LEVELS OF MODEL QUESTION PAPER
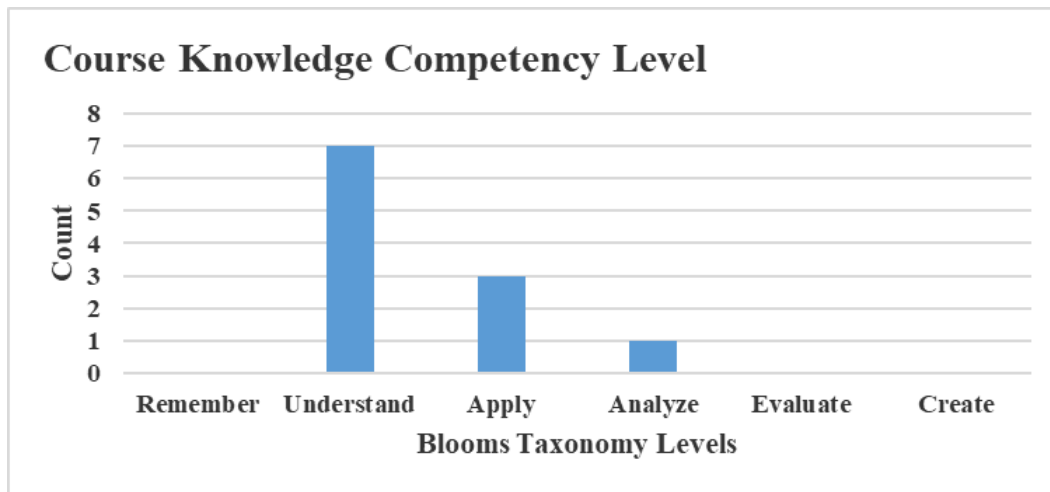
Signature of Course Coordinator

HOD, CSE

Dr. M Madhubala, Professor CSE