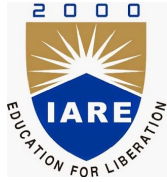


Hall Ticket No

|  |  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|--|

Question Paper Code: ACS012



# INSTITUTE OF AERONAUTICAL ENGINEERING (Autonomous)

Dundigal, Hyderabad - 500 043

MODEL QUESTION PAPER-I

B.Tech VII Semester End Examinations, November 2020

Regulations: IARE - R16

## **BIG DATA AND BUSINESS ANALYTICS** (COMPUTER SCIENCE AND ENGINEERING)

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**

### UNIT-I

- (a) What is Big Data? Explain how big data processing differs from distributed processing. [7m]

(b) List various application of big data. How it can be used to improve business for a super-store. [7m]
- (a) Explain characteristics of Big Data in detail and illustrate in which condition data is called by Big Data. [7m]

(b) How Big Data Analytics can be useful in the development of smart cities and explain the landscape of Big Data Technology? [7m]

### UNIT-II

- (a) What are the advantages of Hadoop? Explain Hadoop architecture and its components with proper diagram. [7m]

(b) Write short note on Hadoop Ecosystem also explain various use cases involved in Hadoop. [7m]
- (a) Why to choose Hadoop for processing Big Data in detail and explain the concept of Distributed and parallel computing challenges? [7m]

(b) Explain in detail the interacting process with Hadoop Ecosystem? List out various big data processing technologies? [7m]

### UNIT-III

- (a) 5. a) Define HDFS? Discuss the HDFS Architecture and HDFS Commands in brief. Write down the goals of HDFS. [7m]

(b) How does HDFS ensure data Integrity in a Hadoop Cluster? [7m]
- (a) Discuss racks in Hadoop Cluster? Explain how Hadoop Clusters are arranged in several racks with an real time example? [7m]

- (b) Create a file in HDFS, Explain the Anatomy of a File Read and Write? [7m]

**UNIT-IV**

7. (a) Explain Map-reduce framework in brief and Draw the architectural diagram for Physical Organization of Compute Nodes. [7m]  
(b) Explain working of following phases of Map Reduce with one common example. (i) Map Phase (ii) Combiner Phase (iii) Shuffle and Sort Phase (iv) Reducer Phase [7m]
8. (a) Determine the working of the map reduce algorithm? [7m]  
(b) Develop the map reduce code for counting occurrences of specific words in the input text file(s). Also write the commands to compile and run the code. [7m]

**UNIT-V**

9. (a) What is Apache Pig and why do we need it and draw the architecture of Apache Pig and explain in brief? [7m]  
(b) Elaborate Pig data Model in detail and Discuss how it will help for effective data flow. [7m]
10. (a) Explain architecture of APACHE HIVE. Explain various data insertion techniques in HIVE with example. [7m]  
(b) What do you mean by Hive SQL Data Definition Language? [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:**

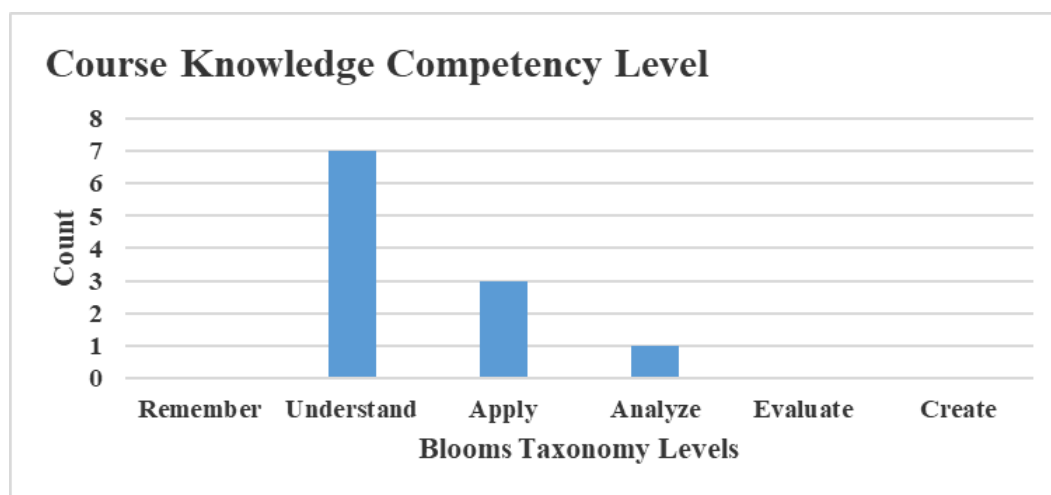
|       |   |
|-------|---|
| CO 1  | Explain the evolution of big data with its characteristics and challenges with traditional business intelligence.   |
| CO 2  | Compare big data analysis and analytics in optimizing the business decisions.   |
| CO 3  | Classify the key issues and applications in 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 | What is Big Data? Explain how big data processing differs from distributed processing.   | Remember   | CO 1  | PO 1     |
|      | b | List various application of big data. How it can be used to improve business for a superstore.   | Remember   | CO 3  | PO 1,2,3 |
| 2    | a | Explain characteristics of Big Data in detail and illustrate in which condition data is called by Big Data.  | Understand | CO 1  | PO 1     |
|      | b | How Big Data Analytics can be useful in the development of smart cities and explain the landscape of Big Data Technology?.                                   | Remember   | CO 2  | PO 1     |
| 3    | a | What are the advantages of Hadoop? Explain Hadoop architecture and its components with proper diagram.   | Remember   | CO 11 | PO 4     |
|      | b | Write short note on Hadoop Ecosystem also explain various use cases involved in Hadoop.  | Understand | CO 5  | PO 1,2,3 |
| 4    | a | Why to choose Hadoop for processing Big Data in detail and explain the concept of Distributed and parallel computing challenges?                             | Remember   | CO 5  | PO 1,2,3 |
|      | b | Explain in detail the interacting process with Hadoop Ecosystem? List out various big data processing technologies?  | Remember   | CO 4  | PO 4     |
| 5    | a | 5. Define HDFS? Discuss the HDFS Architecture and HDFS Commands in brief. Write down the goals of HDFS.  | Remember   | CO 6  | PO 1,2   |
|      | b | How does HDFS ensure data Integrity in a Hadoop Cluster?   | Understand | CO 6  | PO 1,2   |
| 6    | a | Discuss racks in Hadoop Cluster? Explain how Hadoop Clusters are arranged in several racks with an real time example?  | Create     | CO 6  | PO 1,2   |
|      | b | Create a file in HDFS, Explain the Anatomy of a File Read and Write?   | Create     | 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 1,3,4 |
|      | b | Explain working of following phases of Map Reduce with one common example. (i) Map Phase (ii) Combiner Phase (iii) Shuffle and Sort Phase (iv) Reducer Phase | Understand | CO 7  | PO 1,3,4 |

|    |   |  |            |       |          |
|----|---|--|------------|-------|----------|
| 8  | a | Determine the working of the map reduce algorithm?   | Evaluate   | CO 7  | PO 1,3,4 |
|    | b | Develop the map reduce code for counting occurrences of specific words in the input text file(s). Also write the commands to compile and run the code. | Apply      | CO 7  | PO 1,3,4 |
| 9  | a | What is Apache Pig and why do we need it and draw the architecture of Apache Pig and explain in brief?   | Remember   | CO 8  | PO 2,3,4 |
|    | b | Elaborate Pig data Model in detail and Discuss how it will help for effective data flow.   | Understand | CO 9  | PO 2,3,4 |
| 10 | a | Explain architecture of APACHE HIVE. Explain various data insertion techniques in HIVE with example.   | Understand | CO 9  | PO 1,2,3 |
|    | b | What do you mean by Hive SQL Data Definition Language?   | Remember   | CO 11 | PO 1,2   |

### KNOWLEDGE COMPETENCY LEVELS OF MODEL QUESTION PAPER



Signature of Course Coordinator  
Dr. M Madhubala, Professor

HOD,CSE