<table>
<tr><td>Name:

Enrolment No:</td><td>UPES
UNIVERSITY OF TOMORROW</td></tr>
</table>

Course: Big Data Analytics   **Semester: III**
Program: MBA BA - KPMG   **Time   : 03 hrs.**
Course Code: DSBA8002   **Max. Marks: 100**

Instructions:

SECTION A
10Qx2M=20Marks

S. No.		Marks	CO
Q1			
i.	What is a key characteristic of Big Data? A. Small Volume B. Structured Format C. Low Velocity D. Variety of Data	**2**	**CO1**
ii.	Which of the following is a challenge in handling Big Data? A. Data Security B. Limited Variety C. Low Volume D. Slow Velocity	**2**	**CO2**
iii.	What is a benefit of adopting Big Data analytics? A. Reduced Data Variety B. Improved Data Security C. Enhanced Decision-Making D. Slower Data Processing	**2**	**CO3**
iv.	What is a common adoption issue for Big Data implementation? A. Data Security Concerns B. Lack of Variety in Data C. Slow Data Velocity D. Limited Data Volume	**2**	**CO4**
v.	Hadoop 2.0 introduced which resource manager? A. YARN B. HDFS C. MapReduce D. Hive	**2**	**CO1**
vi.	What does HDFS stand for in the context of Big Data? A. High-Density File System B. Hadoop Data Format System C. Hadoop Distributed File System D. High-Velocity Data Storage	**2**	**CO2**

vii.	What is the primary function of MapReduce in Hadoop? A. Data Storage B. Data Analysis C. Data Security D. Data Velocity Control	2	CO3
viii.	Which of the following is a type of NoSQL database? A. MySQL B. Oracle C. MongoDB D. SQL Server	2	CO2
ix.	In the context of databases, what does RDBMS stand for? A. Relational Database Management System B. Rapid Data Backup and Management System C. NoSQL Database System D. Remote Data Backup and Monitoring System	2	CO4
x.	What is a key concept that distinguishes RDBMS from NoSQL databases? A. Data Security B. Data Velocity C. Data Structure Flexibility D. Data Volume	2	CO1
SECTION B **4Qx5M= 20 Marks**			
Q 2	Explain how volume, velocity, and variety contribute to the complexity of Big Data.	5	CO2
Q 3	List and explain three key benefits organizations can gain from adopting Big Data analytics.	5	CO1
Q 4	Briefly describe the key features of Hadoop 2.0.	5	CO3
Q 5	How does HBase contribute to the processing of Big Data?	5	CO4
SECTION-C **3Qx10M=30 Marks**			
Q 6	Compare and contrast Big Data analytics with traditional Business Intelligence.	10	CO3
Q 7	Define MapReduce and explain its fundamental principles. Discuss the advantages of using MapReduce for processing large datasets.	10	CO2
Q 8	How does the process of analytics involve transforming and making sense of raw data to generate valuable information? Provide explanations and associated statistical or machine learning models for the Descriptive Analytics, Diagnostic Analytics, Predictive Analytics, and Prescriptive Analytics.	10	CO1
SECTION-D **2Qx15M= 30 Marks**			
Q 9	Explain the architecture of Hadoop Distributed File System (HDFS). Discuss the key features that make HDFS suitable for storing large-scale data. How does HDFS ensure fault tolerance in distributed data storage? **OR**	15	CO3

	Examine the essential characteristics of NoSQL databases. Elaborate on the four primary classifications of NoSQL databases, citing prominent examples within each type. Additionally, analyze the pros and cons associated with employing NoSQL databases, taking into account their appropriateness for various applications and the challenges they present to developers.		

Q 10	**Write a short note on the following:**
	   A. Define YARN and explain its role in the Hadoop ecosystem.
	   B. Discuss the advantages of YARN in terms of resource management.
	   C. How does YARN enable multi-tenancy in Hadoop clusters?
		15	**CO4**