Name: Enrolment No:	UPES

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, December 2018

Course: MBA (BA) **Semester: III**
Programme: Big Data Analytics **Course Code: DSBA 8002**
Time: 03 hrs. **Max. Marks: 100**

SECTION A

S. No.		Marks	CO
Q 1	**Attempt all questions.** 1. What does commodity Hardware in Hadoop world mean? a) Very cheap hardware b) Industry standard hardware c) Discarded hardware d) Low specifications Industry grade hardware 2. Which of the following are NOT big data problem(s)? a) Parsing 5 MB XML file every 5 minutes b) Processing IPL tweet sentiments c) Processing online bank transactions d) both (a) and (c) 3. What does "Velocity" in Big Data mean? a) Speed of input data generation b) Speed of individual machine processors c) Speed of ONLY storing data d) Speed of storing and processing data	**10X2= 20**	**CO1**

	4. The term Big Data first originated from: a) Stock Markets Domain b) Banking and Finance Domain c) Genomics and Astronomy Domain d) Social Media Domain 5. Which of the following are NOT true for Hadoop? a) It's a tool for Big Data analysis b) It supports structured and unstructured data analysis c) It aims for vertical scaling out/in scenarios d) Both (a) and (c) 6. Which of the following are the core components of Hadoop? a) HDFS b) Map Reduce c) HBase d) Both (a) and (b) 7. Hadoop is open source. a) ALWAYS True b) True only for Apache Hadoop c) True only for Apache and Cloudera Hadoop d) ALWAYS False 8. What is the default HDFS block size? a) 32 MB b) 64 KB		

	c) 128 KB d) 64 MB 9. What is the default HDFS replication factor? a) 4 b) 1 c) 3 d) 2 10. Which of the following is NOT a type of metadata in NameNode? a) List of files b) Block locations of files c) No. of file records d) File access control information		
SECTION B			
	Attempt all questions		
Q1.	What is Big Data? Explain the different characteristics of Big Data.	**5**	**CO2**
Q2.	Explain MongoDB CRUD operations with examples.	**5**	**CO2**
Q3.	What is a HIVE? Specify its Role in Hadoop?	**5**	**CO2**
Q4.	Explain in brief about Name node, Data Node and Secondary Name node in HDFS.	**5**	**CO2**
Q5.	Differentiate between structured, semi structured and un-structured data with examples?	**5**	**CO2**

Q6.	Differentiate between DFS and HDFS.	**5**	**CO2**
	SECTION-C		
	Instruction: Before leaving the examination hall, kindly save your work in folder as your SAP ID at the instructed location.		

Q1.	**A) Insert following records in collection name restaurants:**
	
	{
	"address": {
	"building": "1007",
	"coord": [-73.856077, 40.848447],
	"street": "Morris Park Ave",
	"zipcode": "10462"
	},
	"borough": "Bronx",
	"cuisine": "Bakery",
	"grades": [
	{ "date": { "$date": 1393804800000 }, "grade": "A", "score": 2 },
	{ "date": { "$date": 1378857600000 }, "grade": "A", "score": 6 },
	{ "date": { "$date": 1358985600000 }, "grade": "A", "score": 10 },
	{ "date": { "$date": 1322006400000 }, "grade": "A", "score": 9 },
	{ "date": { "$date": 1299715200000 }, "grade": "B", "score": 14 }
	],
	"name": "Morris Park Bake Shop",
	"restaurant_id": "30075445"
	}
		20	**CO3**

	B) Write the following queries using MongoDB: i) Write a MongoDB query to display all the documents in the collection restaurants. ii) Write a MongoDB query to display the fields restaurant_id, name, borough and cuisine for all the documents in the collection restaurant. iii) Write a MongoDB query to display the fields restaurant_id, name, borough and cuisine, but exclude the field _id for all the documents in the collection restaurant. iv) Write a MongoDB query to display all the restaurant which is in the borough Bronx v) Write a MongoDB query to display the first 5 restaurant which is in the borough	**10X3= 30**	**CO3**

	Bronx. vi)Write a MongoDB query to find the restaurants who achieved a score more than 90. vii)Write a MongoDB query to display the next 5 restaurants after skipping first 5 which are in the borough Bronx. viii)Write a MongoDB query to find the restaurants that achieved a score is more than 80 but less than 100. ix)Write a MongoDB query to find the restaurant Id, name, borough and cuisine for those restaurants which contain 'Wil' as first three letters for its name. x)Write a MongoDB query to find the restaurant Id, name, borough and cuisine for those restaurants which contain 'ces' as last three letters for its name.		