UPES

UNIVERSITY WITH A PURPOSE

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, May 2019

Course: Big Data Ingestion **Semester: IV**
Program: BTech CSE + Big Data **Time 03 hrs.**
Course Code: CSBD 2002 **Max. Marks: 100**

Instructions: Attempt all questions from section A. There is an internal choice in question 11.

	SECTION A		
S. No.		**Marks**	**CO**
Q1	In which of the following components of the big data architecture, data is prepared in a structured format and stored for further analysis? A)- Batch Processing System B)- Data Storage C)- Analytical Data Store D)- Message Ingestion	**5**	**CO1**
Q2	Consider commands --target-dir and --warehouse-dir arguments? Can we use both in the same import command? (Yes/No)	**5**	**CO4**
Q3	Import an RDBMS table order_details (in an order_db MySQL database) in HDFS using Sqoop when we do not have a primary key column. (write command only)	**5**	**CO4**
Q4	Which of the following Statement is true ? A)- Data is analysed at source and loaded to data lake. B)- Big Data is predominantly structured C)- Batch data is analysed real time D)- Big Data collection involves both batch and streaming data	**5**	**CO3**
Q5	Which of the following arguments is used to verify the Sqoop jobs ? A) --show B) –list C) –verify D) –check	**5**	**CO2**

Q6	Which of the following tool is used for batch processing ?
	   A) Kafka
	   B) Storm
	   C) Heron
	   D) Sqoop
		5	**CO1**

SECTION B			
Q7	Define compaction in Kafka and how does it work?	**10**	**CO4**
Q8	Design and illustrate the Big Data Ingestion Architecture. What are the different Big Data Ingestion challenges?	**10**	**CO1**
Q9	Discuss all four Data Ingestion Parameters.	**10**	**CO3**
Q10	Write all possible methods used to import table into HDFS when primary key is not defined.	**10**	**CO4**
Q11	Write a note about topics and partitions in kafka with diagram.	**10**	**CO3**
SECTION-C			
Q12	a) Construct the Kafka Architecture in detail. What is the role of the Zookeeper in the Kafka cluster with the help of diagram? **OR** b) Sketch the Kafka Producer components diagram and illustrate the different steps involved in the sending the data to Kafka broker.	**20**	**CO2, CO3**