Name:

Enrolment No:

UPES
UNIVERSITY WITH A PURPOSE

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, Dec 2021

Course: Big Data Ingestion **Semester: III**
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 number 09 and 10.

SECTION A

S. No.		Marks	CO
Q1	What is the difference between batch processing and stream processing? (4 differences)	**4**	**CO1**
Q2	Consider "--target-dir" and "--warehouse-dir" arguments, can we use both in the same import command? (Yes/No)	**4**	**CO4**
Q3	Import an RDBMS table order_details (in an order_db MySQL database) in HDFS using Sqoop considering table do not have a primary key column. (write command only)	**4**	**CO4**
Q4	Write a short note on "Big Data Ingestion". (maximum 60 words)	**4**	**CO1**
Q5	Name all arguments which is used to verify the Sqoop jobs. Also write complete command to verify scoop jobs? (1+3)	**4**	**CO2**
	SECTION B		
Q6	Design and illustrate the Big Data Ingestion Architecture. What are the different Big Data Ingestion challenges? (6+4)	**10**	**CO1**
Q7	Discuss all four Data Ingestion Parameters.	**10**	**CO3**
Q8	Describe all possible methods used to import table into HDFS when primary key is not defined.	**10**	**CO4**
Q9	Write a short note on topics and partitions in kafka with diagram. (maximum 150 words for each) (6+4) **OR** Define compaction in Kafka and how does it work? (6+4)	**10**	**CO4**
	SECTION-C		

Q10	a) Construct and discuss the Kafka Architecture in detail. What is the role of the Zookeeper in the Kafka cluster with the help of diagram? (10 + 10) **OR** b) Explain sqoop architecture in detail with diagram. Also discuss all sqoop import and export operation with command. (10 + 10)	**20**	**CO3, CO2**
Q11	Design and illustrate Kafka Producer components diagram. Identify the different steps involved in the sending the data to Kafka broker with diagram. (10 + 10)	**20**	**CO3**