APJ ABDULKALAM TECHNOLOGICAL UNIVERSITY 08 PALAKKAD CLUSTER

SECOND SEMESTER M.TECH. DEGREE EXAMINATION JULY 2022

Branch: Computer Science and Engineering Specialization: Computer Science

08CS6042(B) Big Data Essentials

Time: 3 hours Max. Marks: 60

Answer all six questions.

Modules 1 to 6: Part 'a' of each question is compulsory and answer either part 'b' or part 'c' of each question.

Q. No.	Module 1	Marks
1. a	Describe the roles and responsibilities of Data Scientist.	3
	Answer b or c	
b	Explain the steps involved in data analytics life cycle with neat sketch.	6
c	Write the advantages and sources of bigdata.	6
Q. No.	Module 2	Marks
2. a	Describe hadoop input output formats.	3
Answer b or c		
b	Explain the design of HDFS.	6
c	Write the configuration files needed while installing Hadoop.	6
Q. No.	Module 3	Marks
3. a	Describe the steps involved in developing a Map reduce algorithm.	3
*	Answer b or c	k
b	Explain Map reduce word count algorithm design.	6
c	How can you analyse the importance of partitioners and combiners in Map reduce programming?	6
Q. No.	Module 4	Marks
4. a	Discuss about HBase concepts?	3

Answer b or c

b /	Write a short note on NoSQL databases and its types.	6
c	Compare MangoDB with Hive.	6
Q. No.	Module 5	Marks
5. a	Discuss the issues of Data mining with BigData.	4
	Answer b or c	
b	Describe the clustering on Big Data.	8
c	Discuss the implementation of Graph algorithms in Big Data.	8
Q. No.	Module 6	Marks
6. a	Write a short note on in-memory models.	4
	Answer b or c	
b	Discuss about various models used for Big Data analytics.	8
c	Explain the different phases in real time data analytics.	8