

[Time: Three Hours]**[Max Marks:80]****"Please check whether you have got the right question paper."****N.B**

- i) Q.No.1 from section A and Q.No.6 from section B are compulsory.
- ii) Attempt any two questions from the remaining questions in each section.

Section A

Q.1	Differentiate between <u>any two</u> of the following	10
a)	Structured data Vs unstructured data.	
b)	Document stores Vs graph databases.	
c)	Key value stores Vs object oriented data stores.	
Q.2	a) Explain Big Data Management Architecture in detail. b) Explain the impact when an organization can handle big data that is streaming in real time.	08 07
Q.3	a) A differentiate between virtualization resources and services. Explain how to create different virtual systems within a single physical system. b) Explain following layers of the big data stack in detail: redundant physical infrastructure, security infrastructure, interfaces and feeds to and from application.	08 07
Q.4	a) Explain following kinds of structured analysis for big data: Basic analytic and advanced analytics. b) Explain in detail the text analytics using NLP method.	08 07
Q.5	a) Explain the role of virtualization in big data. b) Explain different sources of unstructured data with example.	08 07

Section B

Q.6	Write short notes on <u>(any two)</u>	10
a)	Hive services	
b)	File permission in HDFS	
c)	Combiner function in Map Reduce	
Q.7	a) Explain how hadoop provides a reliable shared storage and analysis system in detail. b) Explain the working of Map Reduce data flow with a single reduce task. Give an example.	08 07
Q.8	a) Justify the following statement: HDFS is a file system designed for storing very large files with streaming data. access patterns, running for a client to write data to HDFS. b) Explain different steps required for a client to write data to HDFS.	08 07
Q.9	a) Write short notes on the following: Pig latin Editors and Pig Data processing operators. b) Write a program using Hive to extract at least three columns from a given database file and insert it into a new table.	08 07
Q.10	a) Explain how the data is analyzed using Hadoop? Give an example. b) Discuss the case study: Hadoop and Hive at Facebook.	08 07