K. J. Somaiya Institute of Management Studies & Research

Program: MIM Batch 2017-20 (SEM V)   Subject: Big Data Analytics   Marks: 50

Exam Date: November 2019   Time: 3 hours   End Term Examination

Notes:

- All 5 questions are compulsory.

- Start each question on a fresh page.

- Draw diagrams, mention examples & answer pointwise.

Questions:

1. Explain Big Data Framework in context of Hadoop Ecosystem with diagrams. List / Map of each Apache Hadoop project with corresponding Big Data framework component.   [10 Marks]

2. Explain Apache Hive in context of Big Data Warehouse.   [10 Marks]

3. Define types of Big Data Analytics. List statistical algorithms used by Data Scientists to gain insight into data sets alongwith with example. List Data sources & Use Cases for Big Data Analytics in 3 industries (e.g. Retail / Telecom / Healthcare).   [10 Marks]

4. List the advantages of Hadoop Enterprise Data Lake over traditional Enterprise Data Warehouse. List the types / categories of big data with examples which Big Data Framework can handle.   [10 Marks]

5. Draw Data Lake Architecture for any one Industry (e.g. Retail / Telecom / Healthcare). Clearly Identify the Data Sources, Data Repository, Analytics Use Cases for that Industry.   [10 Marks]