

B.TECH.**THEORY EXAMINATION (SEM–VI) 2016-17****BIG DATA****Time : 3 Hours****Max. Marks : 100****Note : Be precise in your answer. In case of numerical problem assume data wherever not provided.****SECTION – A****1. Explain the following:****10 x 2 = 20**

- (a) List the characteristics of big data.
- (b) How to calculate risk in marketing?
- (c) Why would you use inferential statistics in big data?
- (d) What do you mean by sharding?
- (e) State the usage of Hadoop pipes.
- (f) Compare Master-Slave and peer to peer architecture in NoSql.
- (g) What is the purpose of bloom filter?
- (h) Compare the classic Map Reduce with YARN.
- (i) Mention the usage of Grunt.
- (j) How Date and Time data types are used in Hive?
- (k) Why Hive is preferred instead of PigLatin?

SECTION – B**2. Attempt any five of the following questions:****5 x 10 = 50**

- (a) Relate crowd sourcing and big data. Justify the relationship with an example.
- (b) Write down the aggregate data model in detail with an example.
- (c) Differentiate “Scale up and Scale out” Explain with an example How Hadoop uses Scale out feature to improve the Performance.
- (d) Discuss in detail about the basic building blocks of Hadoop with a neat sketch.
- (e) Explain in detail about Map-reduce Workflows.
- (f) Provide overview of HBase data model.
- (g) Enumerate the rules followed while data modeling in Cassandra. How the relationships are handled in Cassandra?
- (h) Write down the Hive queries for natural join and outer join. Give examples.

SECTION – C**Attempt any two of the following questions:****2 x 15 = 30**

- 3
 - (i) Explain with a neat sketch about the processing of a job in hadoop.
 - (ii) List the various operational modes of hadoop cluster configuration and explain in detail about configuring/installing the hadoop in local/standalone mode.
- 4
 - (i) Consider the student data File (st.txt), Data in the following format Name, District, age, gender.
 - Write a PIG script to Display Names of all female students
 - Write a PIG script to find the number of Students form XXXX District
 - Write a PIG script to Display District wise count of all male students.
 - (ii) Explain the operators supported by pig w.r.to. data access, transformations and debugging operations.
- 5
 - (i) Discuss the different ways of constructing version stamps. What are their pros and cons?
 - (ii) Write in detail about the three dimensions of big data.