Code No: **RT4105B** **R13** **Set No. 1**

IV B.Tech I Semester Supplementary Examinations, February - 2019

HADOOP AND BIG DATA

(Common to Computer Science and Engineering and Information Technology)

Time: 3 hours **Max. Marks: 70**

Question paper consists of Part-A and Part-B

Answer ALL sub questions from Part-A

Answer any THREE questions from Part-B

*****

PART–A *(22 Marks)*

1.	a)	Define Serialization Reference.	[3]
	b)	What is the function of Job tracker and Task Tracker in HDFS?	[4]
	c)	How can you perform combiner and partition in a Map Reduce?	[4]
	d)	Enumerate the writable wrapper classes for Java Primitives?	[4]
	e)	Write about the key design principles of Pig Latin.	[4]
	f)	List out and explain the data types in Hive.	[3]

PART–B *(3x16 = 48 Marks)*

2.	a)	What are the data structures in Java? Explain in brief about hash map and free map.	[8]
	b)	How can you implement Generic types in Java?	[8]
3.	a)	Define HDFS? Explain in brief about the basic building blocks of Hadoop?	[8]
	b)	How can you Configure XML files?	[8]
4.	a)	What is Map Reduce? Write a map reduce program for weather?	[8]
	b)	Discuss in brief about Mapper code and Reducer code.	[8]
5.	a)	How can you implement the custom writable?	[8]
	b)	Explain in brief about Writable class hierarchies.	[8]
6.	a)	How can you run the Pig scripts in Local and Distributed mode?	[8]
	b)	Write the syntax of a Pig program with suitable example.	[8]
7.	a)	How can you create and manage the data bases in Hive?	[8]
	b)	What are the steps to be followed Squirrel running on Apache Hive with a neat diagram?	[8]

||''|'''|''|'|''''||