

**Reg. No.:**

**Name :**



**VIT<sup>®</sup>**  
**Vellore Institute of Technology**  
(Deemed to be University under section 3 of UGC Act, 1956)

## **School of Computing Science and Engineering**

### **Digital Assignment – II**

**Course Code: ITA6008**

**Couse Name : Big Data Analytics**

**Instructor Name: Ramesh Ragala**

**Slot : B1+B2**

#### **1. Study of Hadoop Ecosystem**

Illustrate the differces among Hbase, Hadoop, Pig and Hive using examples.

#### **2. R practice for Exploratory Data Analysis**

A useful way to detect patterns and anomalies in the data is through the exploratory data analysis with visualization. Visualization gives a succinct, holistic view of the data that may be difficult to grasp from the numbers and summaries alone. Using visual representations of data is a hallmark of exploratory data analyses: letting the data speak to its audience rather than imposing an interpretation on the data a priori. Write a R-Code to examine ways of displaying data to help explain the underlying distributions of a single variable or the relationships of two or more variables.

#### **3. R practice for Exploratory Data Analysis**

Continuing with above question, write R-Code to examine ways of displaying data to help explain the underlaying distribution of multiple variables in data.