Bigdata :

* Big data is nothing but collection of large amount of data from different sources .

The examples of big data is emails, mobile apps and social media

* five characteristics or 5 v’s of big data is

**volume :** The size of dada consider as the volume

**value :**

**variety :** The type of data like unstructured, semi-structured, structured, and raw

data etc.

**velocity :** velocity is at what speed data is coming and storing the data

**veracity** : veracity is quality of the data

Disadvantages of big data :

1.storage issue

2.transfer

3.process

**HADOOP**:

Hadoop is a frame work that helps to process the large amount of data.

The major elements of hadoop is

**HDFS :**Hadoop Distributed File System

It will store and distribute the data

**YARN :** Yet Another Resource Negotiator

Yarn helps to mange the resource at cluster manger.  it performs scheduling and resource allocation for the Hadoop System.

Yarn consist of

Resource Manager

Nodes Manager

Application Manager

**MapReduce:** Programming based Data Processing

In map reduce it uses two functions.one is map and second one is reduce. In this reduce function takes the output from map function.