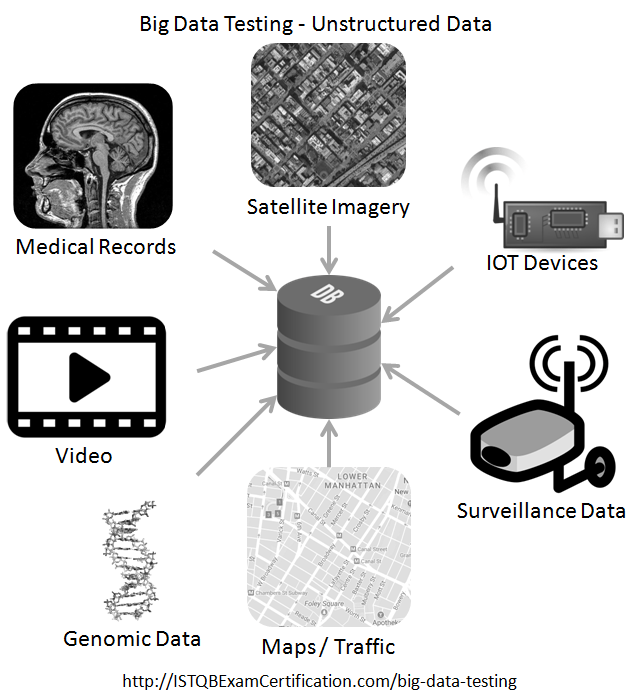
**EXPERIMENT-4**

**TESTING DOCUMENT**

* **TOPIC:** Big Data Analytics
* **INTRODUCTION:** To store large volume of unstructured data with the help of frameworks like hadoop, processing the data in a useful format with the help of framework such as map reduce and draw some insights out of it to convert the values into action.
* **TESTING PHASE:**

1. Testing is to be performed on the complete unstructured data that we’ll be processing for drawing out some insights out of it.
2. **Information about unstructured data**
   1. Data which is not organized into a predefined structure
   2. 80-90% of world’s data
   3. Examples: audios, videos, images, pdf’s emails, etc.
3. **Big Data Testing- The Unstructured View**



1. **Test, Strategy and steps for testing Big Data Application**
   1. DATABASE TESTING OF BIG DATA

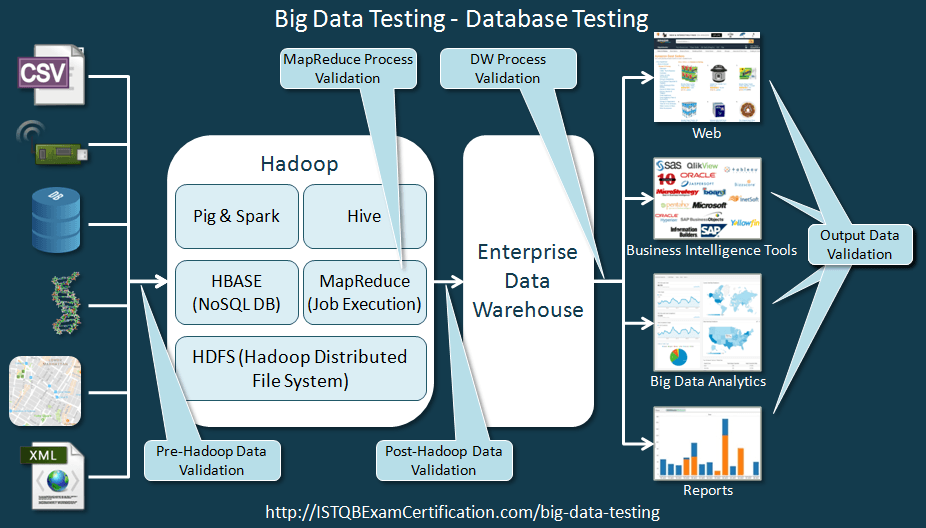
= Before going further in our project phase we’ll be checking that which kind of database we are going to use for storage of our data depending on the functional requirement.

The functional Requirements could be :

1.Storage and querring of data

2.For cloud storage

3.For Business Intelligence, etc.



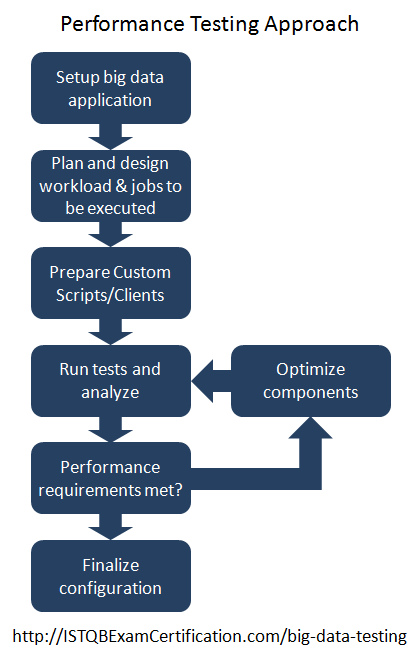
* 1. PERFORMANCE TESTING OF BIG DATA APPLICATION

= The Big Data project involves Huge amount of data processing which requires heavy computing resources and smooth data flow in the network.

So performance testing of a system measures metrics like throughput, memory utilization, CPU Utilization, time taken to complete a task, etc.

It is also recommended that some failover cases should also be run to test the fault tolerance of the system so thet if one node fails then also the processing should not stop.

The Performance tesing approach is shown in the image below:



* 1. FUNCTIONAL TESTING OF BIG DATA APPLICATION:

=Functional testing of big data application is performed by testing the front end application based on user requirement. The front end can be a web based application which interfaces with HADOOP.

* 1. ROLES AND RESPONSIBIITIES OF A TESTER IN BIG DATA APPLICATION

1. The tester should be able to work with unstructured and semi-structured data.
2. The software tester should be able to work with changing schema.
3. Tester should know how to work with frameworks like HADOOP, HDFS, etc.
4. Tester of big data application requires huge amount of technical skills and and there is huge demand for those who passes these skills.