*Big Data Developer*



Session 1: How to Solve Big Data



Assignment



**A C A D G I L D Page 1**



*Big Data Developer*

*Assignment 2 – Try the given quiz questions and send the*

*answers for them in a word document.*

**Table of Contents**

1. Introduction 3

2. Objective 3

3. Prerequisites: 3

4. Associated Data Files 3

5. Problem Statement: 3

6. Expected Output 7

7. Approximate Time to Complete Task 7

**A C A D G I L D Page 2**



*Big Data Developer*

**1. Introduction**

In this assignment you need to select one right choice for the questions given on the topics discussed in the first session.

**2. Objective**

This assignment will help you to consolidate the concepts learnt in the session 1.

**3. Prerequisites:**

None

**4. Associated Data Files**

None

**5. Problem Statement:**

**1. What constitutes to Big Data?**

a.volume

b.velocity

c.variety

d.All the above

**Answer D**

**2. Solution Identified to handle Big Data is?**

a.Hadoop

b.storm

c.spark

d.All of these

**Answer D**

**3. Doug cutting added DFS &MapReduce support to:**

a)Nutch

b)spark

c)GFS

d)facebook

**Answer A**

**4.Which file system is followed by Hadoop?**

a.DFS

b.GFS

c.HDFS

d.NTFS

**Answer C**

**5.How much data can be taken as a BigData in regards to volume?**

a.Mega Bytes

b.Terabytes

c.Petabytes

d. b & c

**Answer D**

**6 .Hadoop can process data in Real time?**

a.Yes

b.No

c.None of these

**Answer A**

**7 .Hadoop is developed using?**

1. C
2. C++
3. Java
4. Ruby

**Answer C**

**8 .Hadoop reduces cost of operations by?**

1. using commodity Hardware
2. Using less number of nodes
3. Limited cluster size
4. None of these

**Answer A**

**9 .Search engine adopted in Hadoop is?**

a.Google

b.yahoo

c.Solr

c.AOL

**Answer A**

**10. Stream processing can be achieved in Hadoop using?**

a.Storm

b.Spark

c.Both of these

d.None of these

**Answer C**

**11 What is the highest data unit we have at present?**

1. petabyte
2. exabyte
3. terabyte
4. Brontobyte

**Answer D**

**12.Hadoop is well suitable for clusters where**

1. Number of nodes keep changing
2. Number of nodes are constant
3. Most number of free nodes
4. None of these

**Answer A**

**13 .what is batch processing?**

a.processing of previously collected jobs in a single batch. b.processing of previously collected jobs in different batches. c.processing of jobs at that moment they are collcted

d.None of these

**Answer C**

**14 .what is stream processing?**

1. Processing data as streams
2. Processing data in batches
3. Both of these
4. None of these

**Answer A**

**15 .Hadoop uses?**

1. Batch processing
2. Near Real time processing
3. Real time processing
4. All the above

**Answer D**

**6. Expected Output**

None.

**7. Approximate Time to Complete Task**

10mins

**A C A D G I L D Page 7**

