

Session 1: How to Solve Big

Data

Assignment 2

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

*Assignment 2 – Try the given quiz questions and forward the answers 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** |  |
|  |  | | |

**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 Big Data?**

a. volume b. velocity c. variety

d. All the above

**Solution : d**

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

a. Hadoop b. storm

c. spark

d. All of these



**Solution : d**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **A C A D G I L D** | **Page 3** |  |
|  |  | | |

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

a) Nutch b) spark c) GFS

d) facebook

**Solution : c**

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

a. DFS

b. GFS

c. HDFS

d. NTFS

**Solution: c**

**5. How much data can be considered as a BigData with regards to volume?**

a. Mega Bytes b. Terabytes

c. Petabytes

d. b & c

**Solution: d**

**6. Can Hadoop process data in real time?**

a. Yes b. No

c. None of these

**Solution: d**

**7. Hadoop is developed using?**

a. C

b. C++

c. Java



|  |  |  |  |
| --- | --- | --- | --- |
|  | **A C A D G I L D** | **Page 4** |  |
|  |  | | |

d. Ruby

**Solution: c**

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

a. Using commodity Hardware b. Using less number of nodes c. Limited cluster size

d. None of these

**Solution: a**

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

a. Google b. Yahoo c. Solr

c. AOL

**Solution: c**

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

a. Storm b. Spark

c. Both of these

d. None of these

**Solution: c**

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

a. Petabyte b. Exabyte c. Terabyte

d. Brontobyte



**Solution: d**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **A C A D G I L D** | **Page 5** |  |
|  |  | | |

**12. Hadoop is well suited for clusters where:**

a. Number of nodes keep changing b. Number of nodes are constant

c. Most number of free nodes d. None of these

**Solution: 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 collected

d. None of these

**Solution: a**

**14. What is stream processing?**

a. Processing data as streams b. Processing data in batches c. Both of these

d. None of these

**Solution: a**

**15. Hadoop uses?**

a. Batch processing

b. Near Real time processing c. Real time processing

d. All the above



**Solution: a**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **A C A D G I L D** | **Page 6** |  |
|  |  | | |

**6. Expected Output**

None.

**7. Approximate Time to Complete Task**

10 mins



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