## Unit 2 - Week-1

### Assignment-1

The due date for submitting this assignment has passed. As per our records you have not submitted this assignment.

Due on 2019-03-13, 23:59 IST.

1) _____________ refers to the biases, noise and abnormality in data, trustworthiness of data.

- Value
- Veracity
- Velocity
- Volume

No, the answer is incorrect.

Score: 0

Accepted Answers:
- Veracity

2) _____________ refers to the connectedness of big data.

- Value
- Veracity
- Velocity
- Valence

No, the answer is incorrect.

Score: 0

Accepted Answers:
- Valence

3) Consider the following statements:

Statement 1: Volatility refers to the data velocity relative to timescale of event being studied

Statement 2: Viscosity refers to the rate of data loss and stable lifetime of data

© 2014 NPTEL - Privacy & Terms - Honor Code - FAQs -

A project of [NPTEL](https://onlinecourses.nptel.ac.in/) in association with [NASSCOM](https://www.nasscom.org/)

Funded by

Friday 21 June 2019 10:15 AM
4) What are the main components of Hadoop Ecosystem? 

- MapReduce, HDFS, YARN 
- MLlib, GraphX 
- Gelly, Table, CEP 
- None of the mentioned

No, the answer is incorrect.

Score: 0

Accepted Answers: 
MapReduce, HDFS, YARN

5) True or False?

NoSQL databases store unstructured data with no particular schema.

- True
- False

No, the answer is incorrect.

Score: 0

Accepted Answers:
True

6) Which of the following is guaranteed by Zookeeper?

- Interactivity
- Flexibility
- Scalability
- Reliability

No, the answer is incorrect.

Score: 0

Accepted Answers:
Reliability

7) Which of the following is not a NoSQL database?

- HBase
- SQL Server
- Cassandra
- None of the mentioned

No, the answer is incorrect.

Score: 0

Accepted Answers:
SQL Server

8) _____________ is a distributed machine learning framework on top of Spark.

- MLlib
- Spark Streaming
- GraphX
9) ____________ is a resource management platform responsible for managing compute resources in the cluster and using them in order to schedule users and applications. 

- Hadoop Common
- Hadoop Distributed File System (HDFS)
- Hadoop YARN
- Hadoop MapReduce

No, the answer is incorrect.
Score: 0
Accepted Answers:
MLlib

10) Which of the following tool is designed for efficiently transferring bulk data between Apache Hadoop and structured datastores such as relational databases.

- Apache Sqoop
- Pig
- Mahout
- Flume

No, the answer is incorrect.
Score: 0
Accepted Answers:
Apache Sqoop

11) ____________ is a distributed, reliable, and available service for efficiently collecting, aggregating, and moving large amounts of log data.

- Apache Sqoop
- Pig
- Mahout
- Flume

No, the answer is incorrect.
Score: 0
Accepted Answers:
Flume

12) ____________ brings scalable parallel database technology to Hadoop and allows users to submit low latencies queries to the data that's stored within the HDFS or the Hbase without acquiring a ton of data movement and manipulation.

- Apache Sqoop
- Impala
- Mahout
- Flume

No, the answer is incorrect.
Score: 0