Dashboard / My courses / BIG DATA ANALYTICS / UNIT 3 BIG DATA TOOLS II

/ QUIZ 1 - BIG DATA ANALYTICS (25.02.2022 - 6.00 PM to 6.45 PM)

Started on	Friday, 25 February 2022, 6:00 PM
State	Finished
Completed on	Friday, 25 February 2022, 6:26 PM
Time taken	26 mins 25 secs
Grade	**30.00** out of 30.00 (**100**%)

Question **1**

Correct

Mark 1.00 out of 1.00

What is the name of the programming framework originally developed by Google that supports the development of applications for processing large data sets in a distributed computing environment?

○ Google Cloud Dataproc

○ HIVE

◉ MAP REDUCE   ✔

○ ZOOKEEPER

Your answer is correct.

The correct answer is:
MAP REDUCE

Question **2**

Correct

Mark 1.00 out of 1.00

What are the main components of Big Data?

- ⦿ ALL THE OPTIONS GIVEN ✔
- ◯ MAP REDUCE
- ◯ YARN
- ◯ HDFS

Your answer is correct.

The correct answer is:
ALL THE OPTIONS GIVEN

Question **3**

Correct

Mark 1.00 out of 1.00

_____ has the world's largest Hadoop cluster.

- ◯ APPLE
- ◯ NONE OF THE OPTIONS GIVEN
- ⦿ FACEBOOK ✔
- ◯ DATAMATICS

Your answer is correct.

The correct answer is:
FACEBOOK

Question **4**

Correct

Mark 1.00 out of 1.00

HDFS stands for

○ Hadoop Digital Filesystem

○ Hadoop Distinguishing Filesystem

◉ Hadoop Distributed Filesystem   ✔

○ Hadoop Distributing Filesystem

Your answer is correct.

The correct answer is:
Hadoop Distributed Filesystem

Question **5**

Correct

Mark 1.00 out of 1.00

The .... phase uses one or more mappers to process the input data.

○ REDUCE

◉ MAP   ✔

○ BOTH MAP AND REDUCE

○ NONE OF THE OPTIONS GIVEN

Your answer is correct.

The correct answer is:
MAP

Question **6**

Correct

Mark 1.00 out of 1.00

The .... phase uses zero or more reducers to process the data output during the map phase.

○ MAP

○ BOTH MAP AND REDUCE

◉ REDUCE   ✔

○ NONE OF THE OPTIONS GIVEN

Your answer is correct.

The correct answer is:
REDUCE

Question **7**

Correct

Mark 1.00 out of 1.00

Which type of data Hadoop can deal with

○ UNSTRUCTURED

○ STRUCTURED

◉ ALL OF THE OPTIONS GIVEN   ✔

○ SEMI - STRUCTURED

Your answer is correct.

The correct answer is:
ALL OF THE OPTIONS GIVEN

Question **8**

Correct

Mark 1.00 out of 1.00

Which of the following is true about Hadoop?

- ⦿ Hadoop is schema-less – we don't need to define a schema before writing the data ✔
- ◯ Hadoop MapReduce processes the data at very low latency
- ◯ Hadoop can't handle structured data
- ◯ None of the Options given

Your answer is correct.

The correct answer is:
Hadoop is schema-less – we don't need to define a schema before writing the data

Question **9**

Correct

Mark 1.00 out of 1.00

Which of the following platforms does Apache Hadoop run on?

- ◯ DEBIAN
- ◯ UNIX - LIKE
- ⦿ CROSS - PLATFORM ✔
- ◯ BARE METAL

Your answer is correct.

The correct answer is:
CROSS - PLATFORM

Question **10**

Correct

Mark 1.00 out of 1.00

Hadoop Framework is written in

- ⊙ JAVA   ✔
- ○ PYTHON
- ○ C++
- ○ SCALA

Your answer is correct.

The correct answer is:
JAVA

Question **11**

Correct

Mark 1.00 out of 1.00

Which of the following is the daemon of Hadoop?

- ○ NODE MANAGER
- ⊙ ALL THE OPTIONS GIVEN   ✔
- ○ DATA NODE
- ○ NAME NODE

Your answer is correct.

The correct answer is:
ALL THE OPTIONS GIVEN

Question **12**

Correct

Mark 1.00 out of 1.00

Apache Hadoop achieves reliability by replicating the data across multiple hosts and hence does not require _____ storage on hosts.

○ OPERATING SYSTEM

○ ZFS

◉ RAID   ✔

○ Standard RAID levels

Your answer is correct.

The correct answer is:
RAID

Question **13**

Correct

Mark 1.00 out of 1.00

Hadoop works in

○ ALL OF THE OPTIONS GIVEN

◉ MASTER - SLAVE FASHION   ✔

○ WORKER / SLAVE FASHION

○ MASTER - WORKER FASHION

Your answer is correct.

The correct answer is:
MASTER - SLAVE FASHION

Question **14**

Correct

Mark 1.00 out of 1.00

Which data tool is regarded as being very necessary before stepping into big data industry?

○ HPCC

◉ HADOOP ✔

○ CASSANDRA

○ NEO4J

Your answer is correct.

The correct answer is:
HADOOP

Question **15**

Correct

Mark 1.00 out of 1.00

Which data tool does one need to deal with large volume of network data or graph related issue like social networking or demographic pattern?

◉ Neo4j ✔

○ R Programming Tool

○ MongoDB

○ RapidMiner

Your answer is correct.

The correct answer is:
Neo4j

Question **16**

Correct

Mark 1.00 out of 1.00

Which of the following is ideal for the business that needs fast and real-time data for instant decisions?

○ R PROGRAMMING TOOL

○ RAPIDMINER

○ NEO4J

⦿ MONGODB   ✔

Your answer is correct.

The correct answer is:
MONGODB

Question **17**

Correct

Mark 1.00 out of 1.00

Which of the following mainly processes data sets?

○ HADOOP

○ APACHE STORM

⦿ CASSANDRA   ✔

○ APACHE SPARK

Your answer is correct.

The correct answer is:
CASSANDRA

Question **18**

Correct

Mark 1.00 out of 1.00

Which of the following is flexible to work with HDFS as well as with other data stores?

- ⦿ APACHE SPARK　　　　　　　　　　　　　　　　　　　　　　　　　　✔
- ○ HADOOP
- ○ CASSANDRA
- ○ APACHE STORM

Your answer is correct.

The correct answer is: APACHE SPARK

Question **19**

Correct

Mark 1.00 out of 1.00

Which is the abstraction of Apache Spark?

- ⦿ BOTH SHARED VARIABLE AND RDD　　　　　　　　　　　　　　　　　✔
- ○ NONE OF THE OPTIONS GIVEN
- ○ RDD
- ○ SHARED VARIABLE

Your answer is correct.

The correct answer is:
BOTH SHARED VARIABLE AND RDD

Question **20**

Correct

Mark 1.00 out of 1.00

Which of the following is not output operation on DStream?

○ SaveAsTextFiles

○ ForeachRDD

◉ ReduceByKeyAndWindow   ✔

○ SaveAsHadoopFiles

Your answer is correct.

The correct answer is:
ReduceByKeyAndWindow

Question **21**

Correct

Mark 1.00 out of 1.00

Which Cluster Manager do Spark Support?

◉ ALL THE OPTIONS GIVEN   ✔

○ Standalone Cluster Manager

○ YARN

○ MESOS

Your answer is correct.

The correct answer is:
ALL THE OPTIONS GIVEN

Question **22**

Correct

Mark 1.00 out of 1.00

Which is not a component on the top of Spark Core?

○ MLlib

○ NONE OF THE OPTIONS GIVEN

◉ Spark RDD   ✔

○ Spark Streaming

Your answer is correct.

The correct answer is:
Spark RDD

Question **23**

Correct

Mark 1.00 out of 1.00

Which of the following provide the Spark Core's fast scheduling capability to perform streaming analytics.

○ SPARK RDD

○ GRAPHX

○ SPARK R

◉ SPARK STREAMING   ✔

Your answer is correct.

The correct answer is:
SPARK STREAMING

Question **24**

Correct

Mark 1.00 out of 1.00

CQL stands for?

○ Component Query Language

○ Cluster Query Language

○ Commit Query Language

◉ Cassandra Query Language ✔

Your answer is correct.

The correct answer is:
Cassandra Query Language

Question **25**

Correct

Mark 1.00 out of 1.00

A _____ is a special column, therefore, it is also a key-value pair, but it stores a map of sub-columns.

○ a. COLUMN

◉ b. SUPER COLUMN ✔

○ c. KEY COLUMN

○ d. SUB COLUMN

Your answer is correct.

The correct answer is:
SUPER COLUMN

Question **26**

Correct

Mark 1.00 out of 1.00

What is the maximum size of a MongoDB document?

○ 12 MB

○ There is no maximum size. It depends on the RAM.

◉ 16 MB ✔

○ 2 MB

Your answer is correct.

The correct answer is:
16 MB

Question **27**

Correct

Mark 1.00 out of 1.00

Which of the following commands create an unique index on author field of the posts collection?

○ db.posts.createIndex({"author": {"$unique":1} });

◉ db.posts.createIndex({"author":1 }, {"unique": true}); ✔

○ db.posts.createIndex({"author": unique });

○ db.posts.createIndexUnique({"author":1 });

Your answer is correct.

The correct answer is:
db.posts.createIndex({"author":1 }, {"unique": true});

Question **28**

Correct

Mark 1.00 out of 1.00

What does the following $slice query return using the following command?

```
db.posts.find( {}, { comments: { $slice: [ -10, 5 ] } } )
```

○ Returns 10 comments, beginning with the last

○ Returns 5 comments, beginning with the first 10 items

◉ Returns 5 comments, beginning with the last 10 items   ✔

○ Returns 10 comments, beginning with the first

Your answer is correct.

The correct answer is:
Returns 5 comments, beginning with the last 10 items

Question **29**

Correct

Mark 1.00 out of 1.00

Which of the following is the most important consideration while designing the schema for MongoDB?

○ The schema should contain maximum indexes

◉ The schema should match the data access and query patterns   ✔

○ The schema should focus on creating possible embedded documents

○ The schema should be kept in 3NF similar to SQL schemas

Your answer is correct.

The correct answer is:
The schema should match the data access and query patterns

Question **30**

Correct

Mark 1.00 out of 1.00

Consider that you are using { upsert : true } option in your update command. Which of the following parameters will be used to determine if any new documents were inserted:

○   nMatched

○   nModified

○   nInserted

◉   nUpserted   ✔

Your answer is correct.

The correct answer is:
nUpserted

◀ **UNIT 2 COURSE MATERIAL**

Jump to...   ⇕

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

Sathyabama Learning Management System

📖📚 **e-Resources**

INFO

Facebook

Twitter

Instagram

YouTube

Sathyabama Staff Forum

Course Material

GET SOCIAL

Copyright © 2022