| | | | |
|---|---|---|---|
| **Course/Branch** : B.E / CSE | **Year / Semester** : IVth YR / VII Sem | Format No. | NAC/TLP-07a.13 |
| **Subject Code** : CS8091 | **Subject Name** : Big Data Analytics | Rev. No. | 02 |
| **Unit No** : 01 | **Unit Name** : Introduction to Big Data | Date | 30.09.2020 |

**OBJECTIVE TYPE QUESTION BANK**

| S. No. | Objective Questions (MCQ /True or False / Fill up with Choices ) | BTL |
|---|---|---|
| 1. | Which of the following is not an example of Social Media?<br>a. Twitter<br>**b. Google**<br>c. Insta<br>d. Youtube | L3 |
| 2. | By 2025, the volume of digital data will increase to<br>a. TB<br>b. YB<br>**c. ZB**<br>d. EB | L1 |
| 3. | For Drawing insights for Business what are need?<br>a. Collecting the data<br>b. Storing the data<br>c. Analysing the data<br>d. **All the above** | L5 |
| 4. | Does Facebook uses "Big Data " to perform the concept of Flashback? Is this True or False.<br>**a. TRUE**<br>b. FALSE | L3 |
| 5. | The Process of describing the data that is huge and complex to store and process is known as<br>a. Analytics<br>b. Data mining<br>**c. Big Data**<br>d. Data Warehouse | L1 |
| 6. | Data generated from online transactions is one of the example for volume of big data. Is this true or False.<br>**a. TRUE**<br>b. FALSE | L3 |
| 7. | Velocity is the speed at which the data is processed<br>a. TRUE<br>b. **FALSE** | L4 |
| 8. | _____ have a structure but cannot be stored in a database.<br>a. Structured<br>**b. Semi-Structured**<br>c. Unstructured<br>d. None of these | L2 |
| 9. | _____refers to the ability to turn your data useful for business.<br>a. Velocity<br>b. Variety<br>**c. Value**<br>d. Volume | L1 |

| | | | |
|---|---|---|---|
| **Course/Branch** : B.E / CSE | **Year / Semester** : IVth YR / VII Sem | Format No. | NAC/TLP-07a.13 |
| **Subject Code** : CS8091 | **Subject Name** : Big Data Analytics | Rev. No. | 02 |
| **Unit No** : 01 | **Unit Name** : Introduction to Big Data | Date | 30.09.2020 |

**OBJECTIVE TYPE QUESTION BANK**

| | | |
|---|---|---|
| 10. | Value tells the trustworthiness of data in terms of quality and accuracy.<br>a. TRUE<br>b. **FALSE** | L3 |
| 11. | GFS consists of a _____ Master and _____ Chunk Servers<br>a. Single, Single<br>b. Multiple, Single<br>**c. Single, Multiple**<br>d. Multiple, Multiple | L1 |
| 12. | Files are divided into _____ sized Chunks.<br>a. Static<br>b. Dynamic<br>**c. Fixed**<br>d. Variable | L2 |
| 13. | _____is an open source framework for storing data and running application on clusters of commodity hardware.<br>a. HDFS<br>**b. Hadoop**<br>c. MapReduce<br>d. Cloud | L1 |
| 14. | HDFS Stores how much data in each clusters that can be scaled at any time?<br>a. 32<br>b. 64<br>**c. 128**<br>d. 256 | L2 |
| 15. | Hadoop MapReduce allows you to perform distributed parallel processing on large volumes of data quickly and efficiently… is this MapReduce or Hadoop… i.e statement is True or False<br>**a. TRUE**<br>b. FALSE | L4 |
| 16. | Hortonworks was introduced by Cloudera and owned by Yahoo.<br>a. TRUE<br>b. **FALSE** | L1 |
| 17. | Hadoop YARN is used for Cluster Resource Management in Hadoop Ecosystem.<br>**a. TRUE**<br>b. FALSE | L4 |
| 18. | Google Introduced MapReduce Programming model in 2004.<br>**a. TRUE**<br>b. FALSE | L4 |
| 19. | _____ phase sorts the data & _____creates logical clusters.<br>a. Reduce, YARN<br>b. MAP, YARN<br>c. REDUCE, MAP<br>d. **MAP, REDUCE** | L2 |

| | | | |
|---|---|---|---|
| **Course/Branch** : B.E / CSE | **Year / Semester** : IVth YR / VII Sem | Format No. | NAC/TLP-07a.13 |
| **Subject Code** : CS8091 | **Subject Name** : Big Data Analytics | Rev. No. | 02 |
| **Unit No** : 01 | **Unit Name** : Introduction to Big Data | Date | 30.09.2020 |

**OBJECTIVE TYPE QUESTION BANK**

| | | |
|---|---|---|
| 20. | There is only one operation between Mapping and Reducing is it True or False…<br>**a. TRUE**<br>b. FALSE | L4 |
| 21. | _____ is factors considered before Adopting Big Data Technology.<br>**a. Validation**<br>b. Verification<br>c. Data<br>d. Design | L3 |
| 22. | _____ for improving supply chain management to optimize stock management, replenishment, and forecasting;<br>a. Descriptive<br>b. Diagnostic<br>c. **Predictive**<br>d. Prescriptive | L3 |
| 23. | which among the following is not a Data mining and analytical applications?<br>a. profile matching<br>b. social network analysis<br>c. facial recognition<br>d. **Filtering** | L2 |
| 24. | _____ as a result of data accessibility, data latency, data availability, or limits on bandwidth in relation to the size of inputs.<br>a. Computation-restricted throttling<br>b. Large data volumes<br>**c. Data throttling**<br>d. Benefits from data parallelization | L1 |
| 25. | As an example, an expectation of using a recommendation engine would be to increase same-customer sales by adding more items into the market basket.<br>a. Lowering costs<br>**b. Increasing revenues**<br>c. Increasing productivity<br>d. Reducing risk | L2 |
| 26. | Which storage subsystem can support massive data volumes of increasing size.<br>a. Extensibility<br>b. Fault tolerance<br>**c. Scalability**<br>d. High-speed I/O capacity | L5 |
| 27. | _____provides performance through distribution of data and fault tolerance through replication<br>**a. HDFS**<br>b. PIG<br>c. HIVE<br>d. HADOOP | L3 |

| | | | |
|---|---|---|---|
| **Course/Branch** : B.E / CSE | **Year / Semester** : IVth YR / VII Sem | Format No. | NAC/TLP-07a.13 |
| **Subject Code** : CS8091 | **Subject Name** : Big Data Analytics | Rev. No. | 02 |
| **Unit No** : 01 | **Unit Name :** Introduction to Big Data | Date | 30.09.2020 |

**OBJECTIVE TYPE QUESTION BANK**

| | | |
|---|---|---|
| 28. | _____ is a programming model for writing applications that can process Big Data in parallel on multiple nodes.<br>a. HDFS<br>**b. MAP REDUCE**<br>c. HADOOP<br>d. HIVE | L1 |
| 29. | _____ takes the grouped key-value paired data as input and runs a Reducer function on each one of them.<br>a. MAPPER<br>**b. REDUCER**<br>c. COMBINER<br>d. PARTITIONER | L2 |
| 30. | _____ is a type of local Reducer that groups similar data from the map phase into identifiable sets.<br>a. MAPPER<br>b. REDUCER<br>**c. COMBINER**<br>d. PARTITIONER | L3 |
| 31. | While Installing Hadoop how many xml files are edited and list them ?<br>i. **core-site.xml**<br>ii. hdfs-site.xml<br>iii. mapred.xml<br>iv. yarn.xml | L4 |
| 32. | Write the code for core-site.xml ?<br><?xml version="1.0" encoding="UTF-8"?><br><?xml-stylesheet type="text/xsl" href="configuration.xsl"?><br><br><configuration><br><property><br><name>hadoop.tmp.dir</name><br><value>D:\hadoop\temp</value><br></property><br><property><br><name>fs.default.name</name><br><value>hdfs://localhost:50071</value><br></property><br></configuration><br><br></?xml > | L6 |
| 33. | Write the code for hdfs-site.xml ? | L3 |

| | | | |
|---|---|---|---|
| **Course/Branch** : B.E / CSE | **Year / Semester** : IVth YR / VII Sem | Format No. | NAC/TLP-07a.13 |
| **Subject Code** : CS8091 | **Subject Name** : Big Data Analytics | Rev. No. | 02 |
| **Unit No** : 01 | **Unit Name :** Introduction to Big Data | Date | 30.09.2020 |

**OBJECTIVE TYPE QUESTION BANK**

| | | |
|---|---|---|
| | <?xml version="1.0" encoding="UTF-8"?><br><?xml-stylesheet type="text/xsl" href="configuration.xsl"?><br><configuration><br><property><name>dfs.replication</name><value>1</value></property><br><property><br><name>dfs.namenode.name.dir</name><value>/hadoop2.6.0/data/name</value><final>true</final></property><br><property><name>dfs.datanode.data.dir</name><value>/hadoop2.6.0/data/data</value><final>true</final> </property><br></configuration><br></xml> | |
| 34. | Write the code for mapred.xml?<br><?xml version="1.0"?><br><configuration><br><property><br><name>mapreduce.framework.name</name><br><value>yarn</value><br></property><br><property><br><name>mapred.job.tracker</name><br><value>localhost:9001</value><br></property><br><property><br><name>mapreduce.application.classpath</name><br><value>/hadoop-2.6.0/share/hadoop/mapreduce/*,<br>/hadoop-2.6.0/share/hadoop/mapreduce/lib/*,<br>/hadoop-2.6.0/share/hadoop/common/*,<br>/hadoop-2.6.0/share/hadoop/common/lib/*,<br>/hadoop-2.6.0/share/hadoop/yarn/*,<br>/hadoop-2.6.0/share/hadoop/yarn/lib/*,<br>/hadoop-2.6.0/share/hadoop/hdfs/*,<br>/hadoop-2.6.0/share/hadoop/hdfs/lib/*,<br></value><br></property><br></configuration> | L3 |
| 35. | Write the code for yarn-site.xml ?<br><?xml version="1.0"?><br><configuration><br><property><br><name>yarn.nodemanager.aux-services</name><br><value>mapreduce_shuffle</value><br></property> | L3 |

| | | | |
|---|---|---|---|
| **Course/Branch** : B.E / CSE | **Year / Semester** : IVth YR / VII Sem | Format No. | NAC/TLP-07a.13 |
| **Subject Code** : CS8091 | **Subject Name** : Big Data Analytics | Rev. No. | 02 |
| **Unit No** : 01 | **Unit Name :** Introduction to Big Data | Date | 30.09.2020 |

**OBJECTIVE TYPE QUESTION BANK**

| | | |
|---|---|---|
| | <property><br><name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name><br><value>org.apache.hadoop.mapred.ShuffleHandler</value><br></property><br><property><br><name>yarn.nodemanager.log-dirs</name><br><value>D:\hadoop\userlog</value><final>true</final><br></property><br><property><name>yarn.nodemanager.local-dirs</name><value>D:\hadoop\temp\nm-localdir</value></property><br><property><br><name>yarn.nodemanager.delete.debug-delay-sec</name><br><value>600</value><br></property><br><property><name>yarn.application.classpath</name><br><value>/hadoop-2.6.0/,/hadoop-2.6.0/share/hadoop/common/*,/hadoop2.6.0/share/hadoop/common/lib/*,/hadoop-2.6.0/share/hadoop/hdfs/*,/hadoop2.6.0/share/hadoop/hdfs/lib/*,/hadoop-2.6.0/share/hadoop/mapreduce/*,/hadoop2.6.0/share/hadoop/mapreduce/lib/*,/hadoop-2.6.0/share/hadoop/yarn/*,/hadoop2.6.0/share/hadoop/yarn/lib/*</value><br></property><br></configuration> | |
| 36. | what are the environmental variable set for Hadoop ?<br>**i. User variables:**<br>• Variable: HADOOP_HOME<br>• Value: D:\hadoop-2.6.0<br>**ii. System variable:**<br>• Variable: Path<br>• Value: D:\hadoop-2.6.0\bin<br>D:\hadoop-2.6.0\sbin<br>D:\hadoop-2.6.0\share\hadoop\common*<br>D:\hadoop-2.6.0\share\hadoop\hdfs<br>D:\hadoop-2.6.0\share\hadoop\hdfs\lib*<br>D:\hadoop-2.6.0\share\hadoop\hdfs*<br>D:\hadoop-2.6.0\share\hadoop\yarn\lib*<br>D:\hadoop-2.6.0\share\hadoop\yarn*<br>D:\hadoop-2.6.0\share\hadoop\mapreduce\lib*<br>D:\hadoop-2.6.0\share\hadoop\mapreduce*<br>D:\hadoop-2.6.0\share\hadoop\common\lib* | L1 |