|  |
| --- |
| **True / False** |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. Lack of specificity is what leads to ambiguity in defining Big Data.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1 Big Data | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2. For a data set to be considered Big Data, it must display only one of the 3 Vs (volume, velocity and variety).   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | False | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1 Big Data | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3. ​Scaling out is keeping the same number of systems, but migrating each system to a larger one.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | False | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1a Volume | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 4. In many ways, the issues associated with volume and velocity are the same.​   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-1b Velocity | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 5. The analysis of data to produce actionable results is feedback loop processing.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1b Velocity | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 6. ​​Relational databases rely on unstructured data.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | False | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-1c Variety | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 7. Big Data processing imposes a structure on the data as needed for applications as a part of retrieval and processing.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-1c Variety | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 8. The ability to graphically present data in a way that makes it understandable is the concept of value.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | False | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1d Other Characteristics | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 9. Characteristics that are important in working with data in the relational database model also apply to Big Data.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-1d Other Characteristics | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 10. Hadoop is a database that has become the de facto standard for most Big Data storage and processing.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2 Hadoop | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 11. ​​​Under the HDFS system, using a write-one, ready-many model simplifies concurrency issues.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2a HDFS | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 12. ​​A block report is used to let the name node know that the data mode is still available.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | False | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2a HDFS | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 13. A reduce function takes a collection of key-value pairs with the same key value and summarizes them into a single result.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2b MapReduce | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 14. Hive is a good choice for jobs that require a small subset of data to be returned very quickly.​   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | False | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2c Hadoop Ecosystem | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 15. ​​​Flume is a tool for converting data back and forth between a relational database and the HDFS.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | False | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2c Hadoop Ecosystem | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 16. ​Most NoSQL products run only in a Linux or Unix environment.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3 NoSQL | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 17. Key-value and document databases are structurally similar.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3b Document Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 18. A column family database is a NoSQL database model that organizes data in key-value pairs with keys mapped to a set of columns in the value component.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3c Column-Oriented Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 19. Interest in graph databases can be tied to the area of social networks.​   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3d Graph Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 20. The name, MongoDB, comes from the word humongous as its developers intended their new product to support extremely large data sets.   |  |  |  | | --- | --- | --- | |  | a. | True | |  | b. | False |  |  |  | | --- | --- | | *ANSWER:* | True | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-5 Working with Document Databases Using MongoDB | | *LEARNING OBJECTIVES:* | 14.07 - Understand how to work with document databases using MongoDB | |

|  |
| --- |
| **Multiple Choice** |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 21.  \_\_\_\_\_\_ is NOT one of the "3 Vs" of Big Data.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | Volume | b. | Velocity | |  | c. | Validation | d. | Variety |  |  |  | | --- | --- | | *ANSWER:* | c | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1 Big Data | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 22. \_\_\_\_\_\_ is keeping the same number of systems, but migrating each system to a larger system.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | Clustering | b. | Scaling up | |  | c. | Streaming | d. | Scaling out |  |  |  | | --- | --- | | *ANSWER:* | b | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1a Volume | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 23. \_\_\_\_\_\_ focuses on filtering data as it enters the system to determine which data to keep and which to discard.​   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | ​Scaling up | b. | ​Feedback loop processing | |  | c. | ​Stream processing | d. | ​Scaling out |  |  |  | | --- | --- | | *ANSWER:* | c | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1b Velocity | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 24. ​A(n) \_\_\_\_\_\_ is a process or set of operations in a calculation.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | ​algorithm | b. | ​feedback loop | |  | c. | ​stream | d. | ​structure |  |  |  | | --- | --- | | *ANSWER:* | a | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1b Velocity | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 25. Big Data \_\_\_\_\_\_.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | relies on the use of structured data | b. | captures data in whatever format it naturally exists | |  | c. | relies on the use of unstructured data | d. | imposes a structure on data when it is captured |  |  |  | | --- | --- | | *ANSWER:* | b | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-1c Variety | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 26. In the context of Big Data, \_\_\_\_\_\_ relates to changes in meaning.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | variety | b. | variability | |  | c. | veracity | d. | viability |  |  |  | | --- | --- | | *ANSWER:* | b | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1d Other Characteristics | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 27. In the context of Big Data, \_\_\_\_\_\_ refers to the trustworthiness of a set of data.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | value | b. | variability | |  | c. | veracity | d. | viability |  |  |  | | --- | --- | | *ANSWER:* | c | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1d Other Characteristics | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 28. By default, Hadoop uses a replication factor of \_\_\_\_\_\_.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | one | b. | two | |  | c. | three | d. | four |  |  |  | | --- | --- | | *ANSWER:* | c | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2a HDFS | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 29. Which of the following is NOT a key assumption of the Hadoop Distributed File System?   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | High volume | b. | Write many, read-once | |  | c. | Streaming access | d. | Fault-tolerance |  |  |  | | --- | --- | | *ANSWER:* | b | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2a HDFS | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 30. When using a HDFS, the \_\_\_\_\_\_ node creates new files by communicating with the \_\_\_\_\_\_ node.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | client; name | b. | data; name | |  | c. | data; client | d. | host; client |  |  |  | | --- | --- | | *ANSWER:* | a | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-2a HDFS | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31. When using a HDFS, a heartbeat is sent every \_\_\_\_\_\_ to notify the name node that the data mode is still available.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | 3 hours | b. | 3 seconds | |  | c. | 6 hours | d. | 6 seconds |  |  |  | | --- | --- | | *ANSWER:* | b | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2a HDFS | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 32. ​When using MapReduce, a \_\_\_\_\_\_\_ function takes a collection and data and sorts and filters it into a set of key-value pairs.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | ​reduce | b. | ​map | |  | c. | ​data | d. | ​block |  |  |  | | --- | --- | | *ANSWER:* | b | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2b MapReduce | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 33. When using MapReduce, best practices suggest that the number of mappers on a given node should be \_\_\_\_\_\_.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | 50 or less | b. | over 100 but less than 300 | |  | c. | 100 or less | d. | at least 300 |  |  |  | | --- | --- | | *ANSWER:* | c | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2b MapReduce | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 34. \_\_\_\_\_\_ processing occurs when a program runs from beginning to end without any user interaction.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | Hadoop | b. | Block | |  | c. | Hive | d. | Batch |  |  |  | | --- | --- | | *ANSWER:* | d | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2b MapReduce | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 35. Two of the most popular applications to simplify the process of creating MapReduce jobs are Hive and \_\_\_\_\_\_.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | Flume | b. | Pig | |  | c. | Sqoop | d. | Impala |  |  |  | | --- | --- | | *ANSWER:* | b | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2c Hadoop Ecosystem | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 36. \_\_\_\_\_\_ is a tool for converting data back and forth between a relational database and the HDFS.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | ​Flume | b. | ​Pig | |  | c. | ​Sqoop | d. | ​Impala |  |  |  | | --- | --- | | *ANSWER:* | c | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2c Hadoop Ecosystem | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 37. \_\_\_\_\_\_ was the first SQL on Hadoop application.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | Flume | b. | Pig | |  | c. | Sqoop | d. | Impala |  |  |  | | --- | --- | | *ANSWER:* | d | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2c Hadoop Ecosystem | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 38. Which of the following is NOT one of the standard NoSQL categories?   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | Document databases | b. | Column-oriented databases | |  | c. | Graph databases | d. | Chart databases |  |  |  | | --- | --- | | *ANSWER:* | d | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3 NoSQL | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 39. To query the value component of the pair when using a key-value database, use get or \_\_\_\_\_\_.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | store | b. | fetch | |  | c. | retrieve | d. | gather |  |  |  | | --- | --- | | *ANSWER:* | b | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3a Key-Value Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 40. Document databases group documents into logical groups called \_\_\_\_\_\_.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | buckets | b. | sets | |  | c. | collections | d. | blocks |  |  |  | | --- | --- | | *ANSWER:* | c | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3b Document Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 41. ​\_\_\_\_\_\_ minimizes the number of disk reads necessary to retrieve a row of data.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | C​olumn-oriented database | b. | ​Row-centric storage | |  | c. | Column-family database | d. | ​Column-centric storage |  |  |  | | --- | --- | | *ANSWER:* | b | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3c Column-Oriented Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 42. Modeling and storing data about relationships is the focus of \_\_\_\_\_\_ databases.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | key-value | b. | column-oriented | |  | c. | document d | d. | graph |  |  |  | | --- | --- | | *ANSWER:* | d | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3d Graph Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 43. Graph theory is a mathematical and computer science field that models relationships, or edges, between objects called \_\_\_\_\_\_.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | maps | b. | scales | |  | c. | buckets | d. | nodes |  |  |  | | --- | --- | | *ANSWER:* | d | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3d Graph Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 44. A query in a graph database is called a \_\_\_\_\_\_.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | schema | b. | hierarchy | |  | c. | traversal | d. | script |  |  |  | | --- | --- | | *ANSWER:* | c | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3d Graph Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 45. Data collected or aggregated around a central topic or entity is said to be \_\_\_\_\_\_ aware.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | aggregate | b. | transversally | |  | c. | feedback | d. | visually |  |  |  | | --- | --- | | *ANSWER:* | a | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3e Aggregate Awareness | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 46. A \_\_\_\_\_\_ is a programmed function within an object used to manipulate the data in that same object.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | batch | b. | method | |  | c. | block | d. | node |  |  |  | | --- | --- | | *ANSWER:* | b | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-5b Example of a MongoDB Query Using find() | | *LEARNING OBJECTIVES:* | 14.07 - Understand how to work with document databases using MongoDB | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 47. In MongoDB, \_\_\_\_\_\_ method retrieves objects from a collection that match the restrictions provided.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | count\* | b. | read\* | |  | c. | review[] | d. | find() |  |  |  | | --- | --- | | *ANSWER:* | d | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-5b Example of a MongoDB Query Using find() | | *LEARNING OBJECTIVES:* | 14.07 - Understand how to work with document databases using MongoDB | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 48. In MongoDB, the \_\_\_\_\_\_ method  is used to improve the readability of retrieved documents through the use of line breaks and indention.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | pretty() | b. | clean\* | |  | c. | break[] | d. | filter+ |  |  |  | | --- | --- | | *ANSWER:* | a | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-5b Example of a MongoDB Query Using find() | | *LEARNING OBJECTIVES:* | 14.07 - Understand how to work with document databases using MongoDB | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 49. Neo4j is a \_\_\_\_\_\_ database.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | graph | b. | column family | |  | c. | key-value | d. | row-centric |  |  |  | | --- | --- | | *ANSWER:* | a | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-6 Working with Graph Databases Using Neo4j | | *LEARNING OBJECTIVES:* | 14.08 - Understand how to work with graph databases using Neo4j | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 50. A(n) \_\_\_\_\_\_ is a tag that is used to associate a collection of nodes as being of the same type or belonging to the same group.   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | a. | edge | b. | key | |  | c. | label | d. | bucket |  |  |  | | --- | --- | | *ANSWER:* | c | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-6a Creating Nodes in Neo4j | | *LEARNING OBJECTIVES:* | 14.08 - Understand how to work with graph databases using Neo4j | |

|  |
| --- |
| **Completion** |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 51. \_\_\_\_\_\_ is the Big Data 3 V that relates to the speed at which data is entering the system.   |  |  | | --- | --- | | *ANSWER:* | Velocity | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1 Big Data | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 52. Scaling out is also referred to as \_\_\_\_\_\_\_.   |  |  | | --- | --- | | *ANSWER:* | clustering | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-1a Volume | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 53. \_\_\_\_\_\_ refers to the analysis of the data to produce actionable results.   |  |  | | --- | --- | | *ANSWER:* | Feedback loop processing​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1b Velocity | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 54. A method of text analysis that attempts to determine if a statement conveys a positive, negative, or neutral attitude is referred to as \_\_\_\_\_\_ analysis.   |  |  | | --- | --- | | *ANSWER:* | sentimental​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1d Other Characteristics | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 55. ​\_\_\_\_\_\_ is the coexistence of a variety of data storage and data management technologies within an organization’s infrastructure.   |  |  | | --- | --- | | *ANSWER:* | ​Polyglot persistence | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-1d Other Characteristics | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 56. Within MapReduce, a(n) \_\_\_\_\_\_ runs maps and reduces tasks on nodes.   |  |  | | --- | --- | | *ANSWER:* | task tracker​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2b MapReduce | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 57. Most organizations that use Hadoop also use a set of other related products that interact and complement each other to produce an entire \_\_\_\_\_\_ of applications and tools.   |  |  | | --- | --- | | *ANSWER:* | ecosystem​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2c Hadoop Ecosystem | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 58. \_\_\_\_\_\_\_ ​languages allow the user to specify what they want, not how to get it which is very useful for query processing.   |  |  | | --- | --- | | *ANSWER:* | Declarative​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2c Hadoop Ecosystem | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 59. ​Within Hadoop,               is used for producing data pipeline tasks that transform data in a series of steps.   |  |  | | --- | --- | | *ANSWER:* | Pig​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2c Hadoop Ecosystem | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 60. ​Within Hadoop,               can transfer data in both directions - into and out of HDFS.   |  |  | | --- | --- | | *ANSWER:* | Sqoop | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-2c Hadoop Ecosystem | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 61. \_\_\_\_\_\_ databases simply store data with no attempt to understand the contents of the value component or its meaning.   |  |  | | --- | --- | | *ANSWER:* | Key-value  KV​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3a Key-Value Databases | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 62. \_\_\_\_\_\_ is a human-readable text format for data interchange that defines attributes and values in a document.   |  |  | | --- | --- | | *ANSWER:* | JavaScript Object Notation  JSON | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3b Document Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 63. \_\_\_\_\_\_ do not store relationships as perceived in the relational model and generally have no support for join operations.   |  |  | | --- | --- | | *ANSWER:* | Document databases​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3b Document Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 64. \_\_\_\_\_\_ ​refers to traditional, relational database technologies that use column-centric, not  row-centric storage.   |  |  | | --- | --- | | *ANSWER:* | Column-oriented database  Columnar database​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3c Column-Oriented Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 65. In a column family database, a column that is composed of a group of other related columns is called a(n) \_\_\_\_\_\_.   |  |  | | --- | --- | | *ANSWER:* | super column​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3c Column-Oriented Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 66. In a graph database, the representation of a relationship between nodes is called a(n) \_\_\_\_\_\_.   |  |  | | --- | --- | | *ANSWER:* | ​edge | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3d Graph Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 67. \_\_\_\_\_ are like attributes; they are the data that we need to store about the node.   |  |  | | --- | --- | | *ANSWER:* | Properties | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-3d Graph Databases | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 68. A database model that attempts to provide ACID-compliant transactions across a highly distributed infrastructure is \_\_\_\_\_\_.   |  |  | | --- | --- | | *ANSWER:* | NewSQL | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-4 NewSQL Databases | | *LEARNING OBJECTIVES:* | 14.06 - Describe the characteristics of NewSQL databases | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 69. \_\_\_\_\_\_  is used to extract knowledge from sources of data—NoSQL databases, Hadoop data stores, and data warehouses—to provide decision support to all organizational users.   |  |  | | --- | --- | | *ANSWER:* | Data analytics​ | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-6c Retrieving Relationship Data with MATCH | | *LEARNING OBJECTIVES:* | 14.08 - Understand how to work with graph databases using Neo4j | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 70. The interactive, declarative query language in Neo4j is called \_\_\_\_\_\_.   |  |  | | --- | --- | | *ANSWER:* | Cypher | | *DIFFICULTY:* | Easy | | *REFERENCES:* | 14-6a Creating Nodes in Neo4j | | *LEARNING OBJECTIVES:* | 14.08 - Understand how to work with graph databases using Neo4j | |

|  |
| --- |
| **Essay** |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 71. Discuss the 3 Vs of Big Data.  How has the definition of Big Data regarding these items changed over time?   |  |  | | --- | --- | | *ANSWER:* | The three V’s are Volume, Velocity and Variety:  ​  Volume is the quantity of data to be stored and a key characteristic of Big Data.  The storage capacities associated with Big Data are very large.  As storage needs increase, they can be handled by scaling up or scaling out.  Scaling up is keeping the same number of systems but migrating each to a larger system.  Scaling out involves distributing data storage structures across a cluster of commodity servers.  ​  Velocity is the speed at which data enters the system and is another key characteristic. In many ways, the issues of velocity mirror those of volume. The velocity of processing can be broken down into two categories: stream and feedback loop.  ​  Variety refers to the vast array of formats and structures in which the data may be captured.  Big Data requires that the data be captured in whatever format it naturally exists.  ​ | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-1 Big Data | | *LEARNING OBJECTIVES:* | 14.01 - Explain the role of Big Data in modern business | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 72. Define the four key assumptions of the Hadoop Distributed File System (HDFS).   |  |  | | --- | --- | | *ANSWER:* | *High volume*: The volume of data in Big Data applications is expected to be in terabytes, petabytes or larger.  Hadoop assumes HDFS files will be extremely large  ​  *Write-once, ready-many*: This model simplifies concurrent issues and improves overall data throughput.  Using this model, a file is created, written to the file system and then closed.  Once the file is closed, changes cannot be made to its contents which improves overall system performance and works well for the types of tasks performed by many Big Data applications.  ​  *Streaming access*: Unlike transaction processing systems, Big Data applications typically process entire files. Hadoop is optimized for batch processing of entire files as continuous streams of data.  ​  *Fault tolerance*: Hadoop is designed to be distributed across thousands of low-cost, commodity computers.  The HDFS is designed to replicate data across many devices so that, when one fails, the data is still available from another device. By default, Hadoop uses a replication factor of three, meaning that each block of data is stored on three devices. | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-2a HDFS | | *LEARNING OBJECTIVES:* | 14.03 - Explain how the core components of the Hadoop framework operate | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 73. Discuss the need for a Hadoop ecosystem and identify the key components.   |  |  | | --- | --- | | *ANSWER:* | Because Hadoop is a very low-level tool requiring considerable effort to create, manage, and use, it presents quite a few obstacles.  This has resulted in a host of related applications that attempt to make Hadoop easier to use and more accessible to users who are not skilled at complex Java programming. Most organizations that use Hadoop also use a set of other related products that interact and complement each other to produce an entire ecosystem of applications and tools.  ​  MapReduce simplification applications have been developed to simplify the process of creating MapReduce jobs.  Two of the most popular are Hive and Pig.  ​  Data ingestion applications help to “ingest” or gather data into Hadoop from existing systems and include Flume.  Sqoop is a tool for converting data back and forth between a relational database and HDFS.  ​  Direct query applications attempt to provide faster query access than is possible through MapReduce and include HBase and Impala. | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-2c Hadoop Ecosystem | | *LEARNING OBJECTIVES:* | 14.04 - Identify the major components of the Hadoop ecosystem | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 74. What is NoSQL and what are the major NoSQL approaches (categories)?   |  |  | | --- | --- | | *ANSWER:* | ​NoSQL is the unfortunate name given to a broad array of nonrelational database technologies that have developed to address Big Data challenges.  The name is unfortunate because it does not describe what the NoSQL technologies are, but rather what they are not.  Even that explanation is poor.  Literally hundreds of products can be considered as NoSQL. Most of them fit into one of four categories: key-value data stores, document databases, column-oriented databases and graph databases. | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-3 NoSQL | | *LEARNING OBJECTIVES:* | 14.05 - Summarize the four major approaches of the NoSQL data model and how they differ from the relational | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 75. Discuss NewSQL and what it attempts to do.​   |  |  | | --- | --- | | *ANSWER:* | ​NewSQL is a database model that attempts to provide ACID-compliant transactions across a highly distributed infrastructure and are the latest technologies to appear to appear in the data management arena to address Big Data problems.  As a new category of data management products, NewSQL databases have not yet developed a track record of success and have been adopted by relatively few organizations.  ​  ​Because no technology can perfectly provide the advantages of both RDBMS and NoSQL, NewSQL has disadvantages, principally centered around its heavy use of in-memory storage. | | *DIFFICULTY:* | Moderate | | *REFERENCES:* | 14-4 NewSQL Databases | | *LEARNING OBJECTIVES:* | 14.06 - Describe the characteristics of NewSQL databases | |