

DUMPS ARENA

Advanced Analytics Specialist Exam for Data Scientists

EMC E20-065

Version Demo

Total Demo Questions: 10

Total Premium Questions: 66

Buy Premium PDF

<https://dumpsarena.co>

sales@dumpsarena.co

sales@dumpsarena.co
dumpsarena.co

QUESTION NO: 1

What advantage does replication provide while storing a file in HDFS?

- A. Data protection and scheduling flexibility
- B. Elimination of requirement for a combiner process
- C. Elimination of requirement for Shuffle and Sort process
- D. Memory optimization and minimizing tasks to run

ANSWER: A

QUESTION NO: 2

How can you improve processing performance in HIVE?

- A. Partition tables
- B. Run the SET hive.exec.parallel = false command
- C. Ensure highly normalized tables and use joins
- D. Minimize bucketing

ANSWER: A

QUESTION NO: 3

How is the relative value of a node visualized in a sunburst?

- A. Color
- B. Area
- C. Gradient
- D. Position

ANSWER: A

QUESTION NO: 4

A data engineer is asked to process several large datasets using MapReduce. Upon initial inspection the engineer realizes that there are complex interdependencies between the datasets.

Why is this a problem?

- A. MapReduce works best on unstructured data
- B. There is no problem; MapReduce accommodates all the data
- C. MapReduce can only parse one file at a time.
- D. MapReduce is not ideal when the processing of one dataset depends on another.

ANSWER: D**QUESTION NO: 5**

What is the maximum number of edges in an undirected graph of 10 nodes?

- A. 45
- B. 90
- C. 100
- D. 9

ANSWER: A**QUESTION NO: 6**

Which library is NOT part of the Apache Spark distribution?

- A. MLib
- B. NLTK
- C. GraphX
- D. Spark SQL

ANSWER: B

QUESTION NO: 7

Consider the two sentences below.

- I mailed my credit card application to the bank
- We walked along the river bank until we came to a waterwheel

What type of NLP ambiguity might occur when interpreting the word "bank"?

- A. Discourse
- B. Syntactic
- C. Semantic
- D. Acoustic

ANSWER: C**QUESTION NO: 8**

What best describes the meaning behind the phrase "Six Degrees of Separation"?

- A. Ability to use about six hops to reach any other node in an extremely large social network
- B. Erdos number of all scholars having written papers with Paul Erdos
- C. Maximum number of edges between nodes in a graph with a diameter of six
- D. Typical distance between nodes that are connected by triadic closure

ANSWER: A**QUESTION NO: 9**

In multinomial logistic regression, what is used to calculate the probability of outcome occurring?

- A. Logistic function applied to a linear combination of the input and outcome variables
- B. Linear regression applied to a combination of input variables
- C. Linear regression applied to a combination of input and outcome variables
- D. Logistic function applied to a linear combination of the input variables

ANSWER: B

QUESTION NO: 10

Which problem type is best suited for simulation?

- A. One with a few. non-random input variables
- B. One that has a closed-form solution
- C. One with numerous, non-random Input-variables
- D. One that compares "what-if scenarios

ANSWER: D