

# QUESTIONS & ANSWERS

Kill your exam at first Attempt



E20-065 Dumps  
E20-065 Braindumps  
E20-065 Real Questions  
E20-065 Practice Test  
E20-065 dumps free



**EMC**

# E20-065

*Advanced Analytics Specialist Exam for Data Scientists*

<http://killexams.com/pass4sure/exam-detail/E20-065>



**QUESTION: 56**

Which representation is most suitable for a small and highly connected network?

- A. Edge list
- B. Adjacency matrix
- C. Eigenvector centrality
- D. Adjacency list

**Answer: B**

**QUESTION: 57**

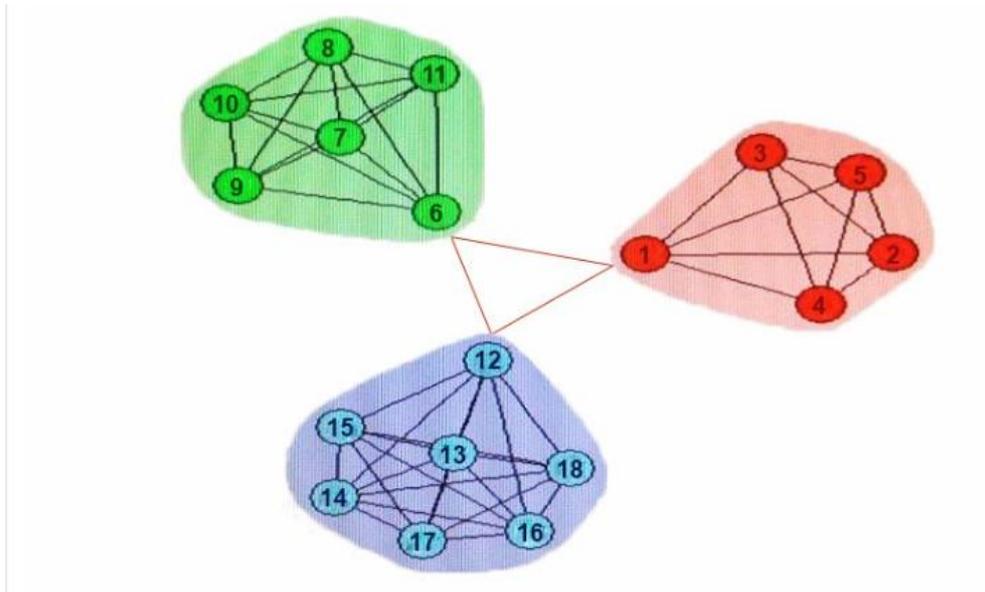
What is a characteristic of spark?

- A. Unable to run map -> reduce execution plans
- B. Supports applications written in Python, Java, and Scala
- C. Less efficient processing small files than Hadoop MapReduce
- D. Supports workflows that can return to previous work steps

**Answer: B**

**QUESTION: 58**

If two of the communities are re-designated to be one community, how does that change the network characteristics? Refer to the exhibit.



- A. Neighborhood overlap would increase
- B. Network diameter would decrease
- C. Modularity would increase
- D. Modularity would decrease

**Answer:** B

**QUESTION:** 59

What is the maximum degree of a node in an undirected graph with 50 nodes'?

- A. 49
- B. 50
- C. 1250
- D. 2500

**Answer:** A

**QUESTION:** 60

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:** 61

What do first-order and second-order Markov processes have in common concerning next word prediction?

- A. Both use WordNet to model the probability of the next word
- B. Both are unsupervised methods
- C. Both provide the foundation to build a trigram language model
- D. Neither makes assumptions about the probability of the next word

**Answer:** C

**QUESTION: 62**

A marketing team creates a graph using a square for each data point, where the length of each side is set to the data value. The data values are 10 and 20. What is the lie factor of the graph?

- A. 1
- B. 2
- C. 3
- D. 6

**Answer: B**

**QUESTION: 63**

You are analyzing written transcripts of focus groups conducted on product X. Your approach is to use TF-IDF for your analysis. What combination of TF-IDF scores should you examine to ensure you only report on the most important terms?

- A. High TF score and high DF score
- B. High TF score and high IDF score
- C. High TF score and low IDF score
- D. Low TF score and low DF score

**Answer: C**

**QUESTION: 64**

What does YARN provide over and above MapReduce?

- A. Separate cluster and resource management
- B. Parallelized processing
- C. Serialized processing
- D. Access to HDFS data

**Answer: A**

**QUESTION: 65**

What is an intended application of the MapReduce framework?

- A. Processing can be broken into smaller pieces
- B. Processing a large number of small files

- C. Processing in real time is required
- D. Processing a small subset of data

**Answer: A**

**QUESTION: 66**

What do lemmatization and stemming have in common?

- A. Use WordNet
- B. Remove common words in a natural language
- C. Reduce the high dimensionality in text
- D. Use a set of heuristics

**Answer: C**

For More exams visit <https://killexams.com/vendors-exam-list>



*Kill your exam at First Attempt....Guaranteed!*