E20-065^{Q&As}

Advanced Analytics Specialist Exam for Data Scientists

Pass EMC E20-065 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.passapply.com/e20-065.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by EMC
Official Exam Center

🌐 **Instant Download** After Purchase

🌐 **100% Money Back** Guarantee

🌐 **365 Days** Free Update

🌐 **800,000+** Satisfied Customers

QUESTION 1

A hotel chain runs a simul-ation on room pricing. They want to estimate revenue, per hotel, within +/- $10 with 95% confidence (Za/2=1.96). The estimated revenue standard deviation is $5000 based on previous booking data.

What is the optimal number of simulation trials to run?

A. A 32-bit operating system was used

B. The same number of trials was used

C. A linear congruential generator (LCG) was used (or pseudo-random number generation

D. Different seeds tor the random number generator were used.

Correct Answer: C

QUESTION 2

In a connected, undirected graph of 5 nodes with 10 edges, how many more edges need to be added to make the clustering coefficient of every node equal 1 ?

A. 0

B. 5

C. 10

D. 15

Correct Answer: A

QUESTION 3

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.
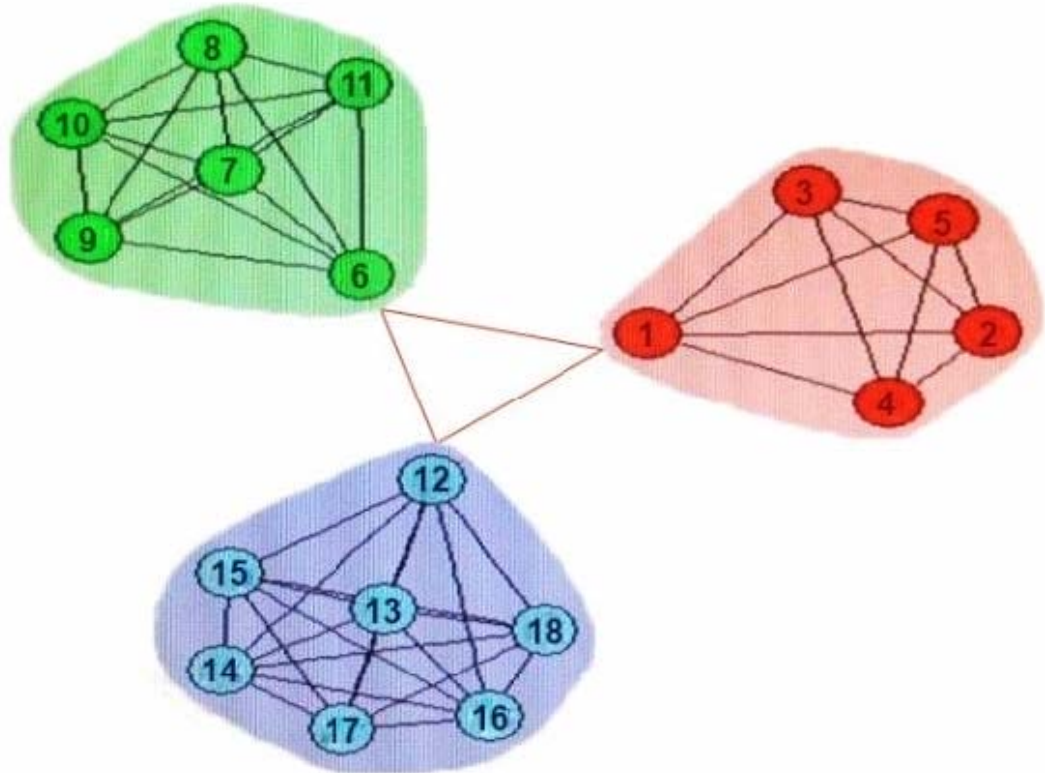
Why is this a problem?

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.

Correct Answer: D

QUESTION 4

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

Correct Answer: B

QUESTION 5

What is a random subspace of features, as used by Random Forests?

A. A random subset of features that are chosen at each split in the decision tree

B. Filtration of data that does not meet a pre-defined weighting thrsehold

C. The creation of out-of-bag (OOB) data that is used to select features

D. Removal of highly correlated variables to randomize the features

Correct Answer: A

E20-065 VCE Dumps   E20-065 Practice Test   E20-065 Braindumps