DUMPSARENA

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.com

sales@dumpsarena.com

QUESTION NO: 1

What describes how nodes in a social network are similar to each other in characteristics?

A. Community clustering

B. Modularity

C. Homophily

D. Strongly tied network

ANSWER: C

QUESTION NO: 2

Which graph structure would best model the relationship between job seekers and employers?

A. Bipartite

B. Weighted

C. Directed acyclic

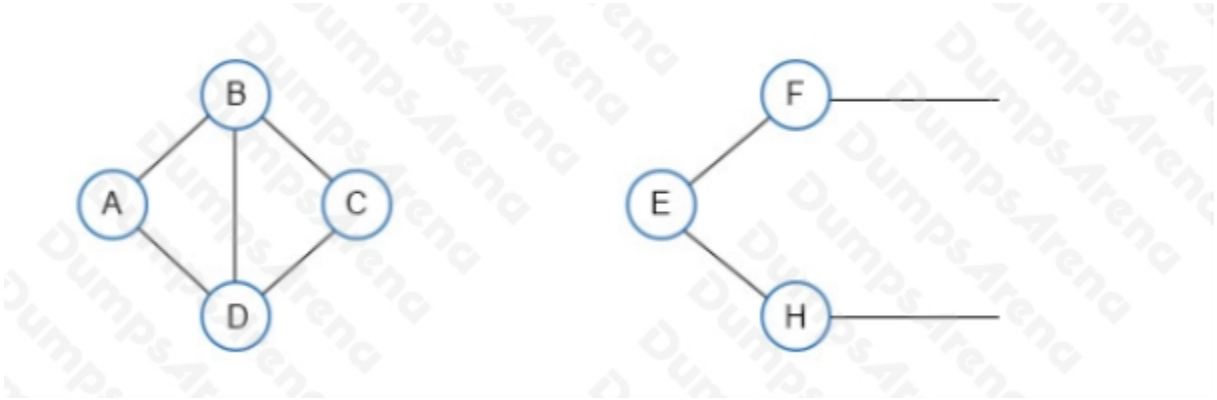
D. Ranked

ANSWER: A

QUESTION NO: 3

In the graph, which edge would be considered a weak lie?

Refer to the exhibit.



A. C-E

B. E-F

C. B-C

D. G-I

ANSWER: D

QUESTION NO: 4

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: 5

What is an ideal use case for HDFS?

A. Storing files that are updated frequently

B. Storing files that are written once and read many times

C. Storing results between Map steps and Reduce steps

D. Storing application files in memory

ANSWER: B

QUESTION NO: 6

Consider dataset that resides in HDFS. Which tool natively provides the capability to run a Random Forests model against this data?

A. Mahout

B. Pig

C. Hive

D. HBase

ANSWER: A

QUESTION NO: 7

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

QUESTION NO: 8

What are the major components of the YARN architecture?

A. ResourceManager and NodeManager

B. Task Tracker and NameNode

C. HDFS, Tez, and Spark

D. Avro, ZooKeeper, and HDFS

ANSWER: A

QUESTION NO: 9

What are two visualization tools used for trivariate data?

A. Scatter plot matrix

B. Hexbin plot and heatmap

C. Scatter plot matrix and density plot

D. Scatter plot matrix and heatmap

ANSWER: B

QUESTION NO: 10

What is a characteristic of stop words?

A. Used in term frequency analysis

B. Include words such as "a", "an", and "the"

C. Meaningful words requiring a parser to stop and examine them

D. Don't occur often in text

ANSWER: B

DUMPSARENA