



**Free Questions for E20-065 by go4braindumps**

**Shared by Buckner on 15-04-2024**

**For More Free Questions and Preparation Resources**

**Check the Links on Last Page**

## Question 1

---

**Question Type:** MultipleChoice

---

What is the most likely reason for an HBase table to contain millions of columns?

### Options:

---

- A- Data is imported from a relational database table
- B- Data is stored in the column qualifier
- C- There are thousands of columns families
- D- The column names are randomly generated

### Answer:

---

B

## Question 2

---

**Question Type:** MultipleChoice

---

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

**Options:**

---

A- Mahout

B- Pig

C- Hive

D- HBase

**Answer:**

---

A

## Question 3

---

**Question Type:** MultipleChoice

---

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?

**Options:**

---

A- 1

B- 2

C- 3

D- 6

**Answer:**

---

B

## Question 4

---

**Question Type:** MultipleChoice

---

What are the major components of the YARN architecture?

**Options:**

---

- A- ResourceManager and NodeManager
- B- Task Tracker and NameNode
- C- HDFS, Tez, and Spark
- D- Avro, ZooKeeper, and HDFS

**Answer:**

---

A

## Question 5

---

**Question Type:** MultipleChoice

---

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

**Options:**

---

- 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 threshold
- 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

**Answer:**

---

A

## Question 6

---

**Question Type: MultipleChoice**

---

Which scenario would be ideal for processing Hadoop data with Hive?

**Options:**

---

**A-** Structured data, real-time processing

**B-** Unstructured data; batch processing

**C-** Unstructured data; real-time processing

**D-** Structured data; batch processing

**Answer:**

---

B

## Question 7

---

**Question Type:** MultipleChoice

---

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

### Options:

---

- 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

**To Get Premium Files for E20-065 Visit**

**<https://www.p2pexams.com/products/e20-065>**

**For More Free Questions Visit**

**<https://www.p2pexams.com/dell-emc/pdf/e20-065>**

