



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.geekcert.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

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

Correct Answer: C

QUESTION 2

You conduct a TFIDF analysis on 3 documents containing raw text and derive TFIDF ("data", document y) = 1.908. You know that the term "data" only appears in document 2.

What is the TF of "data" in document 2?

- A. 2 based on the following reasoning: $TFIDF = TF \cdot IDF = 1.908$ You then know that IDF will equal $\log(3/2) = 0.954$ Therefore, $TFIDF = TF \cdot 0.954 = 1.908$ TF will then round to 2
- B. 4 based on the following reasoning: $TFIDF = TF \cdot IDF = 1.908$ You then know that IDF will equal $\log(3/1) = 0.477$ Therefore, $TFIDF = TF \cdot 0.477 = 1.908$ TF will then round to 4
- C. 6 based on the following reasoning: $TFIDF = TF \cdot IDF = 1.908$ You then know that IDF will equal $3/1 = 3$ Therefore, $TFIDF = TF/3 = 1.908$ TF will then round to 6
- D. 11 based on the following reasoning: $TFIDF = TF \cdot IDF = 1.908$ You then know that IDF will equal $\log(3/2) = 0.176$ Therefore, $TFIDF = TF \cdot 0.176 = 1.908$ TF will then round to 11

Correct Answer: B

QUESTION 3

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

Correct Answer: A



QUESTION 4

Which HDFS feature protects against user errors causing accidental loss of data?

- A. Encryption
- B. Replication
- C. Namenode federation
- D. Snapshots

Correct Answer: B

QUESTION 5

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

Correct Answer: B

[E20-065 PDF Dumps](#)

[E20-065 Study Guide](#)

[E20-065 Braindumps](#)