

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.pass2lead.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 best describes tokenization?

- A. Adding lexical relations to the raw text
- B. Converting text into the list of terms
- C. Converting text into a list of unique terms
- D. Reducing variant forms of tokens to their base forms

Correct Answer: B

QUESTION 2

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

Correct Answer: A

QUESTION 3

Which scenario is a proper use case for multinomial logistic regression?

- A. A marketing firm wants to estimate the personal income of a group of potential customers. Using inputs such as age, education, marital status, and credit card expenditures, a data scientist is building a model that will estimate a person's income
- B. A logistic distribution company wants to minimize the distance traveled by its delivery trucks. A data scientist is building a model to determine the optimal route for each of its trucks
- C. To improve the initial routing of a loan application, a financial institution plans to classify a loan application as Approve, Reject, or Possibly_Approve. Based on the company's historical loan application data, a data scientist is building a model to assign one of these three outcomes to each submitted application.
- D. A manufacturer plans to determine the optimal number of workers to employ in an assembly line process. Utilizing the observed distributions of the task durations of each process step, a data scientist is building a model to mimic the interactions and dependencies between each stage in the manufacturing process.

Correct Answer: C

QUESTION 4

What is an effective use of color in visualization?

- A. Use self-explanatory colors so a legend is unnecessary
- B. Maximize use of color to make a more lasting impression
- C. Use high contrast colors such as red and blue
- D. Minimize use of color except for emphasis

Correct Answer: C

QUESTION 5

In a social network, what does it mean for a node to have a high degree but low betweenness?

- A. The node is adjacent to a few nodes, each of which has high Page Ranks.
- B. The node has the only edge connecting its community to the rest of the graph.
- C. The node can be easily bypassed by communications taking other shorter paths.
- D. The node acts as the hub of the graph.

Correct Answer: D

[E20-065 Study Guide](#)

[E20-065 Exam Questions](#)

[E20-065 Braindumps](#)