



# E20-065<sup>Q&As</sup>

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

How can you improve processing performance in HIVE?

- A. Partition tables
- B. Run the SET hive.exec.parallel = false command
- C. Ensure highly normalized tables and use joins
- D. Minimize bucketing

Correct Answer: A

---

### QUESTION 2

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 3

What is a typical use of a UDF in Pig?

- A. Creating functionality outside of what is provided by the built-in functions
- B. Providing Functional access to user-defined data in HDFS
- C. Providing advanced analytics to Hadoop
- D. Providing an interface from Pig to Microsoft Excel for easier data manipulation

Correct Answer: A

---



#### QUESTION 4

What is a characteristic of the trigram language model?

- A. Based on the second-order Markov process
- B. Equivalent to trigram hidden Markov models
- C. Uses smoothing to reduce the high dimensionality in text
- D. Can be used for part-of -speech tagging

Correct Answer: A

---

#### QUESTION 5

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

Correct Answer: B

[E20-065 PDF Dumps](#)

[E20-065 VCE Dumps](#)

[E20-065 Study Guide](#)