

# 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.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**

You are analyzing written transcripts of focus groups conducted on product X. Your approach is to use TFIDF for your analysis.

What combination of TF-IDF scores should you examine to ensure you only report on the most important terms?

- A. High TF score and high DF score
- B. High TF score and high IDF score
- C. High TF score and low IDF score
- D. Low TF score and low DF score

Correct Answer: C

---

**QUESTION 2**

Which library is NOT part of the Apache Spark distribution?

- A. MLib
- B. NLTK
- C. GraphX
- D. Spark SQL

Correct Answer: B

---

**QUESTION 3**

What does YARN provide over and above MapReduce?

- A. Separate cluster and resource management
- B. Parallelized processing
- C. Serialized processing
- D. Access to HDFS data

Correct Answer: A

---

**QUESTION 4**

After a client submits a job request to the YARN ResourceManager, what happens next?

- A. The scheduler allocates a container to run an ApplicationMaster
- B. The ResourceManager allocates containers to run map and reduce tasks
- C. The Resource Manager requests load data from the NodeManagers
- D. The ApplicationManager starts an ApplicationMaster

Correct Answer: D

---

#### QUESTION 5

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

[Latest E20-065 Dumps](#)

[E20-065 PDF Dumps](#)

[E20-065 VCE Dumps](#)