



E20-007^{Q&As}

Data Science and Big Data Analytics

Pass EC-COUNCIL E20-007 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.geekcert.com/e20-007.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by EC-COUNCIL Official Exam Center

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers





QUESTION 1

You have been assigned to run a linear regression model for each of 5,000 distinct districts, and all the data is currently stored in a PostgreSQL database. Which tool/library would you use to produce these models with the least effort?

- A. MADlib
- B. Mahout
- C. R
- D. HBase

Correct Answer: A

QUESTION 2

A data scientist is given an R data frame (i.e., `empdata`) with the following columns: Age Salary Occupation Education Gender. The scientist wants to examine only the Salary and Occupation columns for ages greater than 40. Which

command extracts the appropriate rows and columns from the data frame?

- A. `empdata[empdata$Age > 40, c("Salary", "Occupation")]`
- B. `empdata[c("Salary", "Occupation"), empdata$Age > 40]`
- C. `empdata[Age > 40, ("Salary", "Occupation")]`
- D. `empdata[, c("Salary", "Occupation")]$Age > 40`

Correct Answer: A

QUESTION 3

Which data asset is an example of semi-structured data?

- A. XML data file
- B. Database table
- C. Webserver log
- D. News article

Correct Answer: A

QUESTION 4

Which word or phrase completes the statement? Emphasis color is to standard color as _____ .



- A. Main message is to context
- B. Main message is to key findings
- C. Frequent item set is to item
- D. Pie chart is to proportions

Correct Answer: A

QUESTION 5

You are provided four different datasets. Initial analysis on these datasets show that they have identical mean, variance and correlation values. What should your next step in the analysis be?

- A. Visualize the data to further explore the characteristics of each data set
- B. Select one of the four datasets and begin planning and building a model
- C. Combine the data from all four of the datasets and begin planning and bulding a model
- D. Recalculate the descriptive statistics since they are unlikely to be identical for each dataset

Correct Answer: A

[E20-007 PDF Dumps](#)

[E20-007 Study Guide](#)

[E20-007 Braindumps](#)