



E20-007^{Q&As}

Data Science and Big Data Analytics

Pass EMC E20-007 Exam with 100% Guarantee

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

<https://www.pass4itsure.com/e20-007.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**

A data scientist is given an R data frame, "empdata", with the columns Age, Salary, Occupation, Education, and Gender. The data scientist would like 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 2

When creating a presentation for a technical audience, what is the main objective?

- A. Show that you met the project goals
- B. Show how you met the project goals
- C. Show if the model will meet the SLA
- D. Show the technique to be used in the production environment

Correct Answer: B

QUESTION 3

Refer to the exhibit.

	FREE HOUSING	HOME OWNER	RENTER	TOTAL
BAD CREDIT	54	476	270	800
GOOD CREDIT	75	1245	460	1780
TOTAL	129	1721	730	2580

Click on the calculator icon in the upper left corner. You are given a list of pre-defined association rules:

- A. `RENTER => BAD CREDIT`
- B. `RENTER => GOOD CREDIT`
- C. `HOME OWNER => BAD CREDIT`
- D. `HOME OWNER => GOOD CREDIT`



E. FREE HOUSING => BAD CREDIT

F. FREE HOUSING => GOOD CREDIT For your next analysis, you must limit your dataset based on rules with confidence greater than 60%. Which of the rules will be kept in the analysis?

A. Rules B and D

B. Rules A and F

C. Rules C and E

D. Rules D and E

Correct Answer: A

QUESTION 4

Which method is used to solve for coefficients b_0, b_1, \dots, b_n in your linear regression model : $Y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$

A. Ordinary Least squares

B. Apriori Algorithm

C. Ridge and Lasso

D. Integer programming

Correct Answer: A

QUESTION 5

Refer to the exhibit.



Table A: Term inverse document frequencies

	Solid	State	Disk
IDF	2	1.5	0.5

Table B: Term per-document frequencies

Document	Tf		
	Solid	State	Disk
A	0.1	0.1	0.1
B	0.1	0.1	0.5
C	0.2	0.2	0.1
D	0.1	0.3	0.1

Click on the calculator icon in the upper left corner. An analyst is searching a corpus of documents for the topic "solid state disk". In the Exhibit, Table A provides the inverse document frequency for each term across the corpus. Table B provides each term's frequency in four documents selected from corpus. Which of the four documents is most relevant to the analyst's search?

- A. Document C
- B. Document A
- C. Document B
- D. Document D

Correct Answer: A

[Latest E20-007 Dumps](#)

[E20-007 PDF Dumps](#)

[E20-007 VCE Dumps](#)