



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**

What is the output format from the Map function of MapReduce?

- A. Key-value pairs
- B. Binary representation of keys concatenated with structured data
- C. Compressed index
- D. Unique key record and separate records of all possible values

Correct Answer: A

QUESTION 2

Refer to the exhibit.

| Independent Variable | Coefficient | P-Value |
|----------------------|-------------|---------|
| A | 1.23 | 0.001 |
| B | -6.72 | 0.035 |
| C | 4.57 | 0.020 |
| D | 10.23 | 0.123 |

After analyzing a dataset, you report findings to your team:

1.

Variables A and C are significantly and positively impacting the dependent variable.

2.

Variable B is significantly and negatively impacting the dependent variable.

3.

Variable D is not significantly impacting the dependent variable.

After seeing your findings, the majority of your team agreed that variable B should be positively impacting the dependent variable.

What is a possible reason the coefficient for variable B was negative and not positive?

- A. Variable B is interacting with another variable due to correlated inputs
- B. Variable B needs a quadratic transformation due to its relationship to the dependent variable
- C. The information gain from variable B is already provided by another variable



D. Variable B needs a logarithmic transformation due to its relationship to the dependent variable

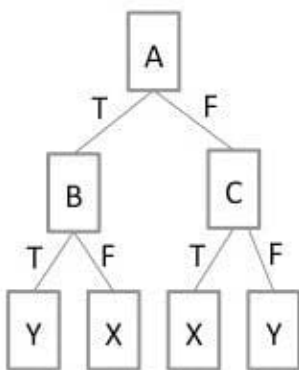
Correct Answer: A

QUESTION 3

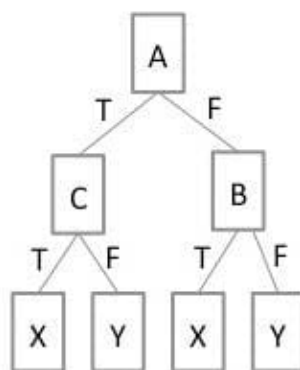
Based on the exhibit, the table shows the values for the input Boolean attributes A, B, and

C. In addition, the exhibit shows the values for the output attribute "class".

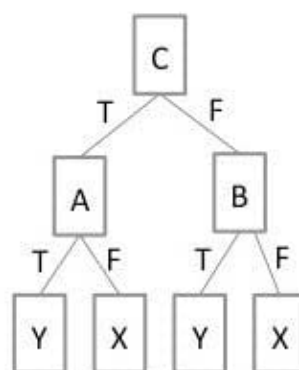
| A | B | C | CLASS |
|---|---|---|-------|
| T | T | T | X |
| T | T | F | Y |
| T | F | T | X |
| F | F | F | Y |
| F | T | T | X |
| F | F | T | Y |



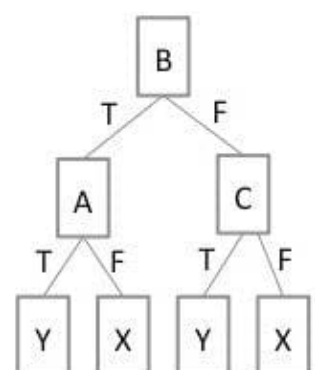
Tree A



Tree B



Tree C



Tree D

Which decision tree is valid for the data?

A. Tree A

B. Tree B

C. Tree C

D. Tree D

Correct Answer: B



QUESTION 4

If R factors are categorical variables, which data classification level are they most closely related?

- A. Nominal
- B. Ordinal
- C. Interval
- D. Ratio

Correct Answer: A

QUESTION 5

You are building a logistic regression model to predict whether a tax filer will be audited within the next two years. Your training set population is 1000 filers. The audit rate in your training data is 4.2%. What is the sum of the probabilities that the model assigns to all the filers in your training set that have been audited?

- A. 42.0
- B. 4.2
- C. 0.42
- D. 0.042

Correct Answer: A

[E20-007 PDF Dumps](#)

[E20-007 VCE Dumps](#)

[E20-007 Practice Test](#)