E20-026 $^{Q\&As}$

Enterprise Storage Networking Specialist Exam

Pass EMC E20-026 Exam with 100% Guarantee

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

https://www.leads4pass.com/e20-026.html

100% Passing Guarantee
100% Money Back Assurance

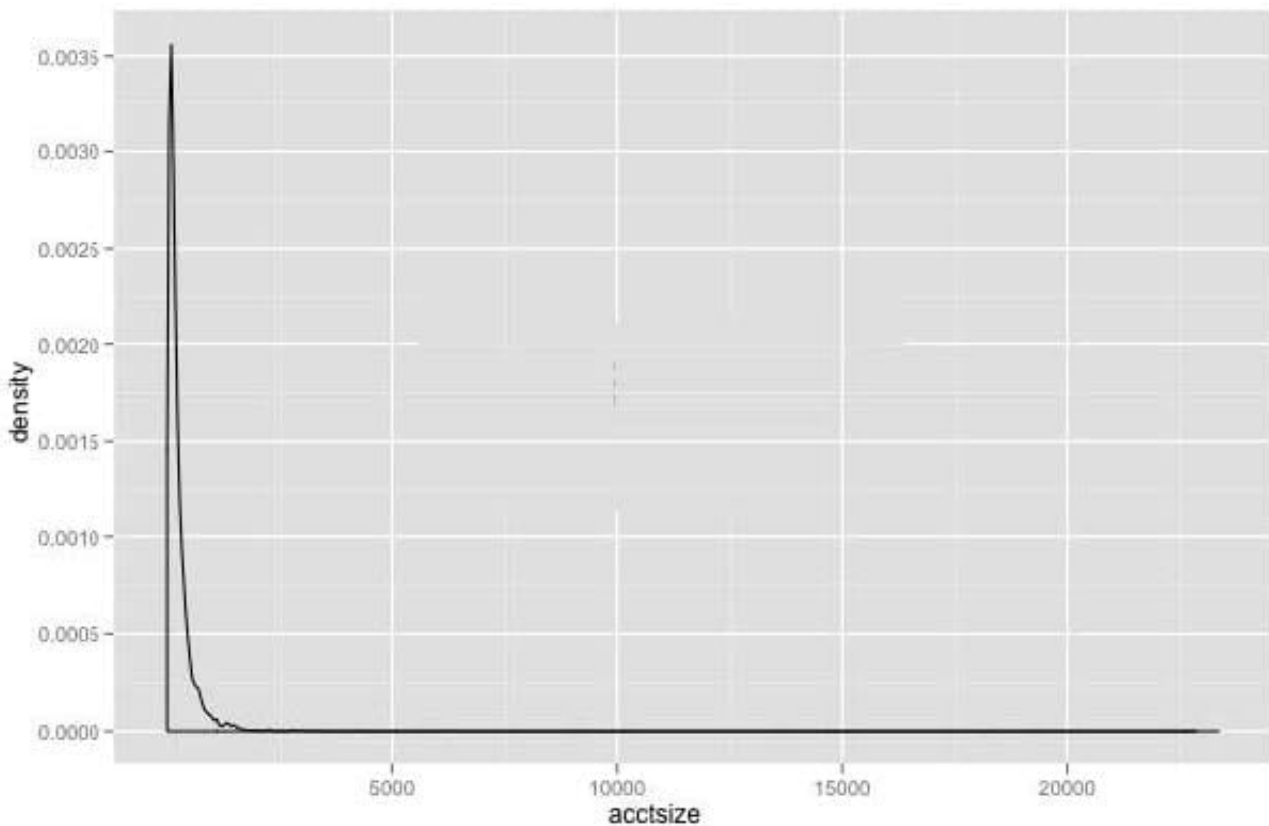Following Questions and Answers are all new published by EMC Official Exam Center

⚙ **Instant Download** After Purchase

⚙ **100% Money Back** Guarantee

⚙ **365 Days** Free Update

⚙ **800,000+** Satisfied Customers

QUESTION 1

Refer to the exhibit.

You have plotted the distribution of savings account sizes for your bank. How would you proceed, based

on this distribution?

A. The data is extremely skewed. Replot the data on a logarithmic scale to get a better sense of it.

B. The data is extremely skewed,but looks bimodal; replot the data in the range 2,500-10,000 to be sure.

C. The accounts of size greater than 2500 are rare,and probably outliers. Eliminate them from your future analysis.

D. The data is extremely skewed. Split your analysis into two cohorts: accounts less than 2500,and accounts greater than 2500

Correct Answer: A

QUESTION 2

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 3

Which word or phrase completes the statement?

Business Intelligence is to ad-hoc reporting and dashboards as Data Science is to _____ .

A. Optimization and Predictive Modeling

B. Alerts and Queries

C. Structured Data and Data Sources

D. Sales and profit reporting

Correct Answer: A

QUESTION 4

Which activity is performed in the Operationalize phase of the Data Analytics Lifecycle?

A. Define the process to maintain the model

B. Try different analytical techniques

C. Try different variables

D. Transform existing variables

Correct Answer: A

QUESTION 5

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 6

You have run the association rules algorithm on your data set, and the two rules {banana, apple} => {grape} and {apple, orange}=> {grape} have been found to be relevant. What else must be true?

A. {grape,apple,orange} must be a frequent itemset.

B. {banana,apple,grape,orange} must be a frequent itemset.

C. {grape} => {banana,apple} must be a relevant rule.

D. {banana,apple} => {orange} must be a relevant rule.

Correct Answer: A

QUESTION 7

Refer to the Exhibit.

You are working on creating an OLAP query that outputs several rows of with summary rows of subtotals

and grand totals in addition to regular rows that may contain NULL as shown in the exhibit. Which function

can you use in your query to distinguish the row from a regular row to a subtotal row?

```
store   | customer | product | sum
--------+----------+---------+----+
s1      | c2       | p2      | 70
s1      | c2       |         | 70
s1      | c3       | p1      | 40
s1      | c3       |         | 40
s1      |          |         | 110
s2      | c1       | p1      | 90
s2      | c1       | p2      | 50
s2      | c1       |         | 140
s2      |          | p1      | 44
s2      |          |         | 44
s2      |          |         | 184
        |          |         | 294
(12 rows)
```
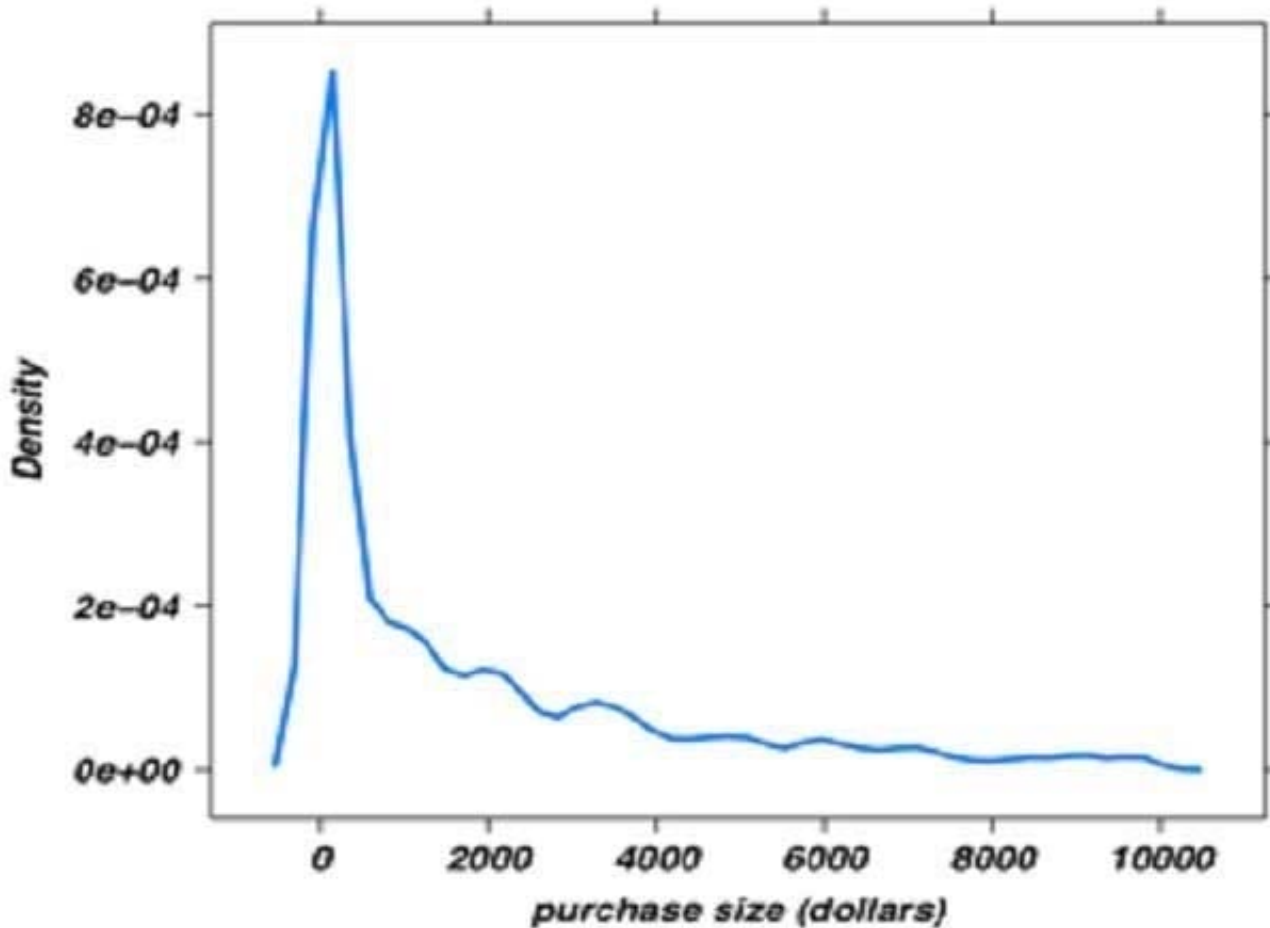
A. GROUPING

B. RANK

C. GROUP_ID

D. ROLLUP

Correct Answer: A

QUESTION 8

Refer to the exhibit.

You have created a density plot of purchase amounts from a retail website as shown. What should you do next?

A. Recreate the plot using the barplot() function

B. Use the rug() function to add elements to the plot

C. Recreate the density plot using a log normal distribution of the purchase amount data

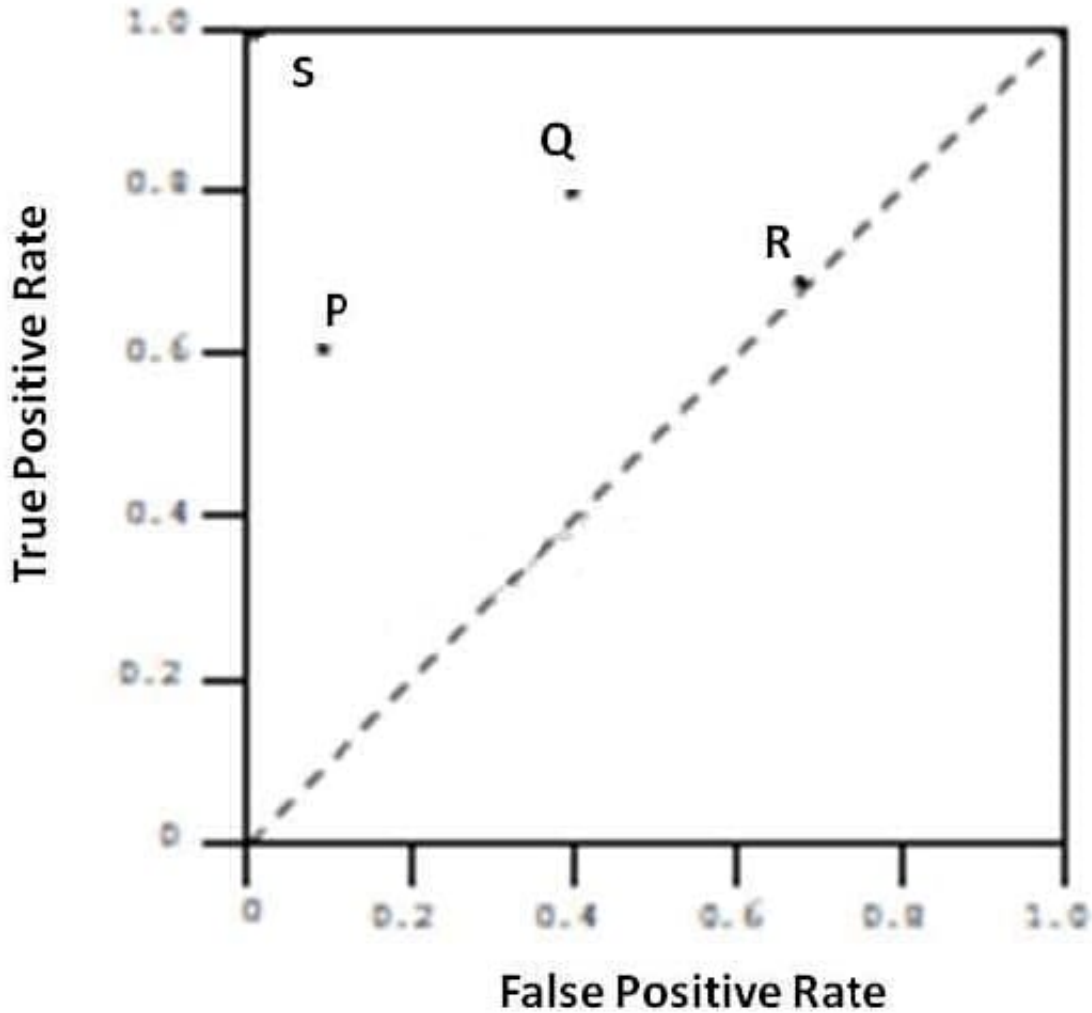D. Reduce the sample size of the purchase amount data used to create the plot

Correct Answer: C

QUESTION 9

Refer to the exhibit.

The graph represents an ROC space with four classifiers labelled A through D. Which point in the graph

represents a perfect classification?

A. S

B. P

C. Q

D. R

Correct Answer: A

QUESTION 10

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

A. Key-value pairs

B. Binary respresentation of keys concatenated with structured data

C. Compressed index

D. Unique key record and separate records of all possible values

Correct Answer: A

QUESTION 11

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 12

You are asked to write a report on how specific variables impact your client\\'s sales using a data set

provided to you by the client. The data includes 15 variables that the client views as directly related to

sales, and you are restricted to these variables only.

After a preliminary analysis of the data, the following findings were made:

1.

Multicollinearity is not an issue among the variables

2.

Only three variables--A, B, and C--have significant correlation with sales You build a linear regression model on the dependent variable of sales with the independent variables of A, B, and C. The results of the regression are seen in the exhibit. You cannot request additional datA. what is a way that you could try to increase the R2 of the model without artificially inflating it?

Independent Variable	Coefficient	P-Value
A	0.45	0.0000
B	3.67	0.0000
C	1.23	0.0000

$R^2 = 0.10$

A. Create clusters based on the data and use them as model inputs

B. Force all 15 variables into the model as independent variables

C. Create interaction variables based only on variables A,B,and C

D. Break variables A,B,and C into their own univariate models

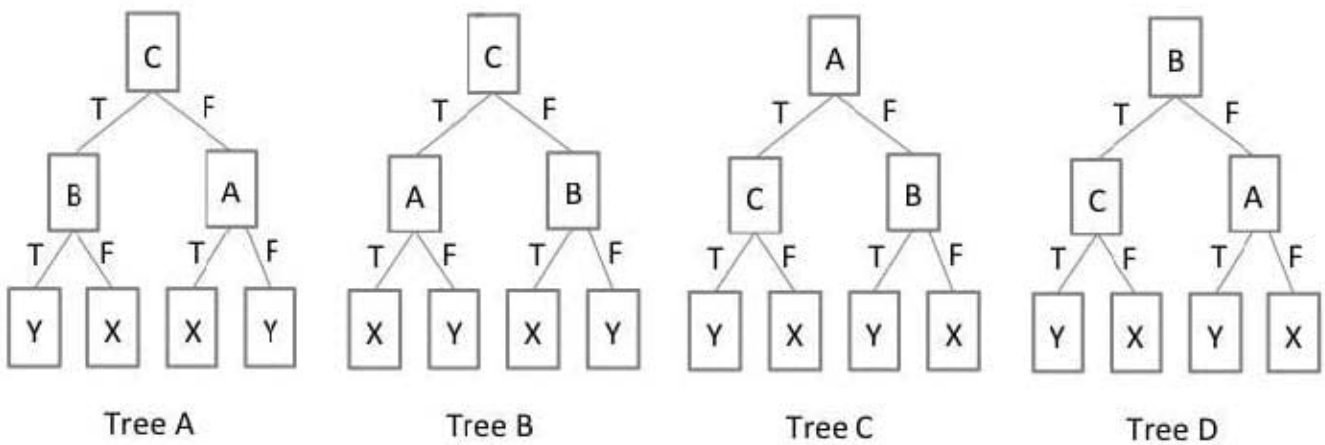Correct Answer: A

QUESTION 13

Refer to the Exhibit.

In the Exhibit, the table shows the values for the input Boolean attributes "A", "B", and "C". It also shows

the values for the output attribute "class". Which decision tree is valid for the data?

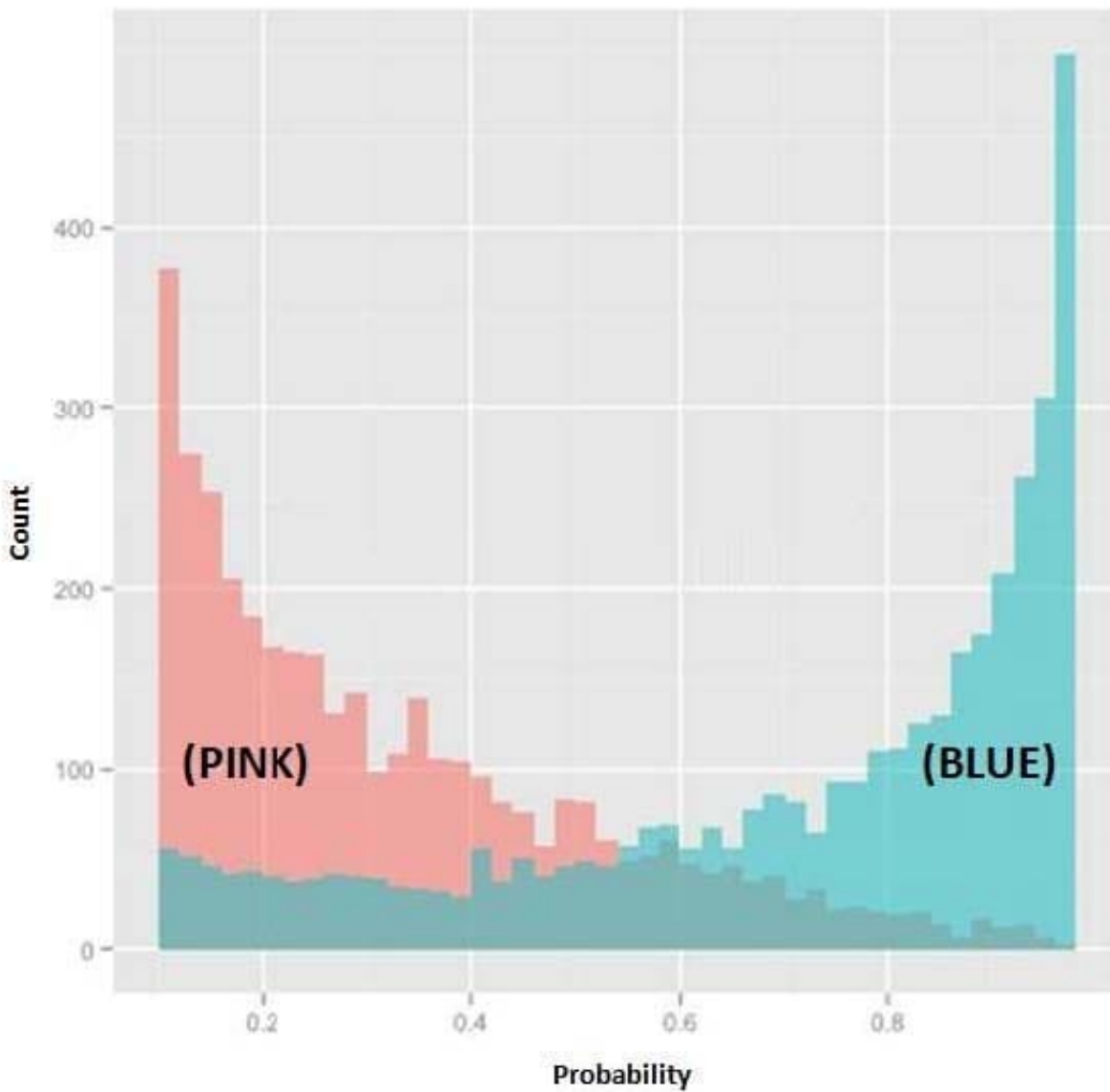| C	B	A	CLASS
T	T	T	X
T	T	F	Y
T	F	T	X
F	F	F	Y
F	T	T	X
F	F	T	Y

Tree A   Tree B   Tree C   Tree D

A. Tree B

B. Tree A

C. Tree C

D. Tree D

Correct Answer: A

QUESTION 14

Refer to the exhibit. In the exhibit, the x-axis represents the derived probability of a borrower defaulting on a loan. Also in the exhibit, the pink represents borrowers that are known to have not defaulted on their loan, and the blue represents borrowers that are known to have defaulted on their loan. Which analytical method could produce the probabilities needed to build this exhibit?

A. Logistic Regression

B. Linear Regression

C. Discriminant Analysis
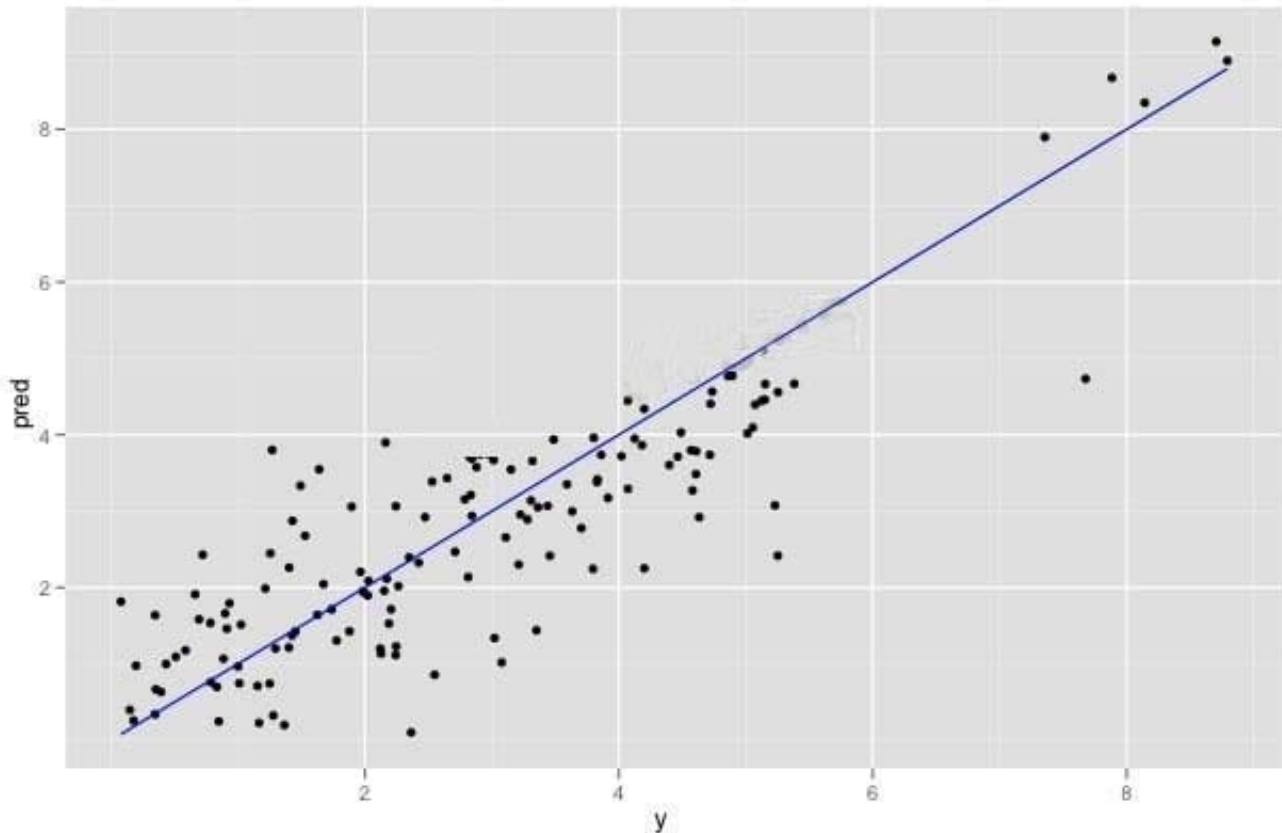
D. Association Rules

Correct Answer: A

QUESTION 15

Refer to the exhibit.

You have run a linear regression model against your data, and have plotted true outcome versus predicted

outcome. The R-squared of your model is 0.75. What is your assessment of the model?

A. The R-squared may be biased upwards by the extreme-valued outcomes. Remove them and refit to get a better idea of the model\\'s quality over typical data.

B. The R-squared is good. The model should perform well.

C. The extreme-valued outliers may negatively affect the model\\'s performance. Remove them to see if the R-squared improves over typical data.

D. The observations seem to come from two different populations,but this model fits them both equally well.

Correct Answer: A

E20-026 Practice Test   E20-026 Exam Questions   E20-026 Braindumps