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.pass2lead.com/e20-026.html
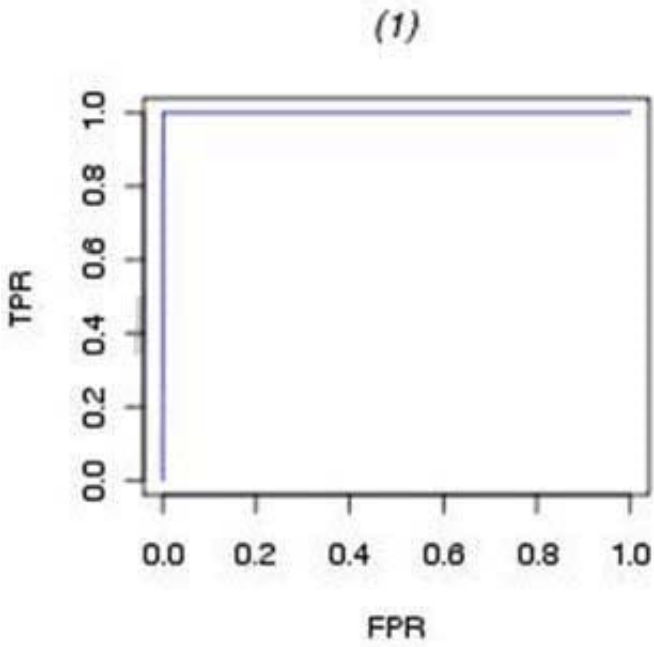
100% Passing Guarantee
100% Money Back Assurance

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

Instant Download After Purchase

100% Money Back Guarantee
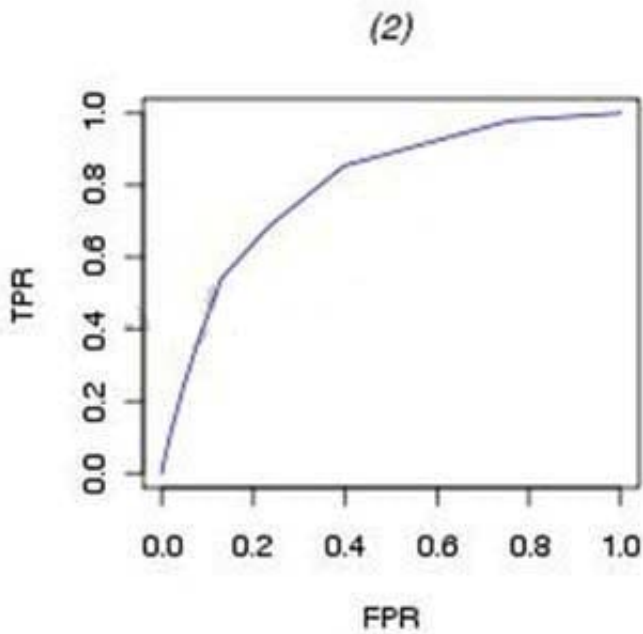
365 Days Free Update

800,000+ Satisfied Customers
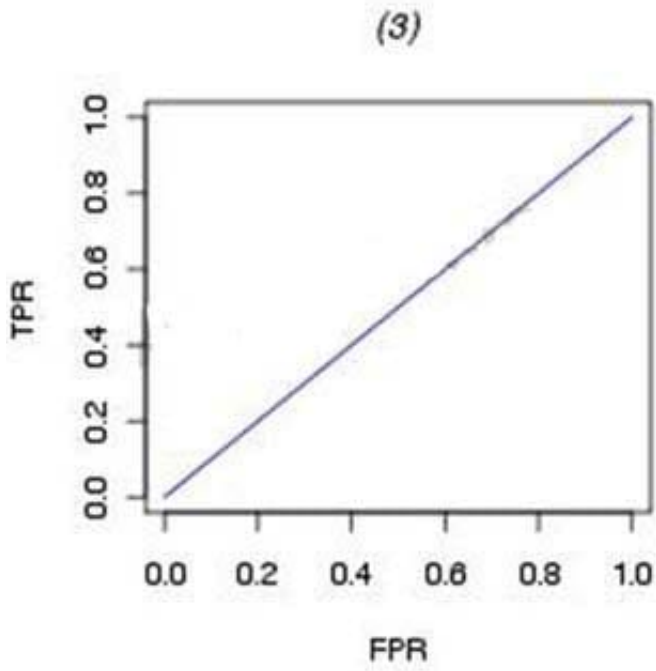
![Pass2Lead](https://Pass2Lead.com)
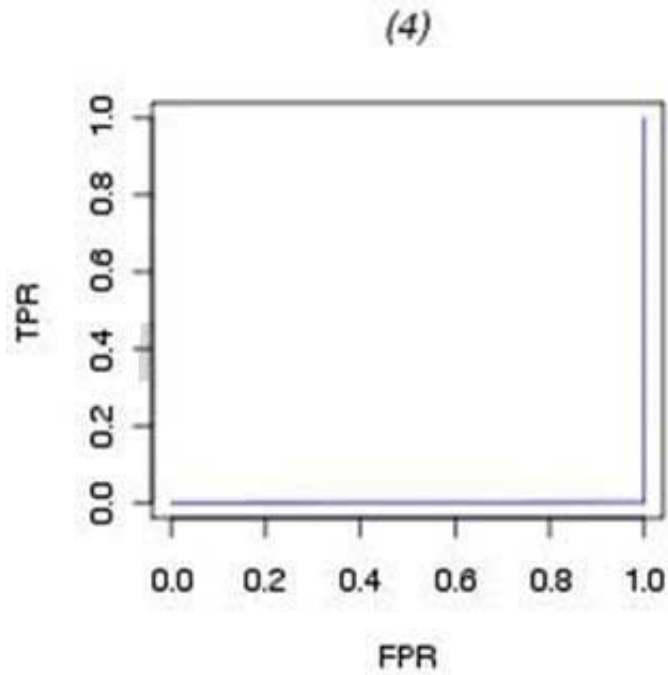**QUESTION 1**

Which ROC curve represents a perfect model fit? A)

B)

C) 52 / 55

(3)

D)

(4)

A. Exhibit A

B. Exhibit B

C. Exhibit C

D. Exhibit D

Correct Answer: A

QUESTION 2

What are the characteristics of Big Data?

A. Data volume,processing complexity,and data structure variety.

B. Data volume,business importance,and data structure variety.

C. Data type,processing complexity,and data structure variety.

D. Data volume,processing complexity,and business importance.

Correct Answer: A

QUESTION 3

Trend, seasonal, and cyclical are components of a time series. What is another component?

A. Irregular

B. Linear

C. Quadratic

D. Exponential

Correct Answer: A

QUESTION 4

Refer to the exhibit.

You are building a decision tree. In this exhibit, four variables are listed with their respective values of info-

gain.

Based on this information, on which attribute would you expect the next split to be in the decision tree?

Attribute	Info-Gain
Age	0.0310
Income	0.0100
Gender	0.0034
Credit Score	0.0456

A. Credit Score

B. Age

C. Income

D. Gender

Correct Answer: A

QUESTION 5

You submit a MapReduce job to a Hadoop cluster and notice that although the job was successfully submitted, it is not completing. What should you do?
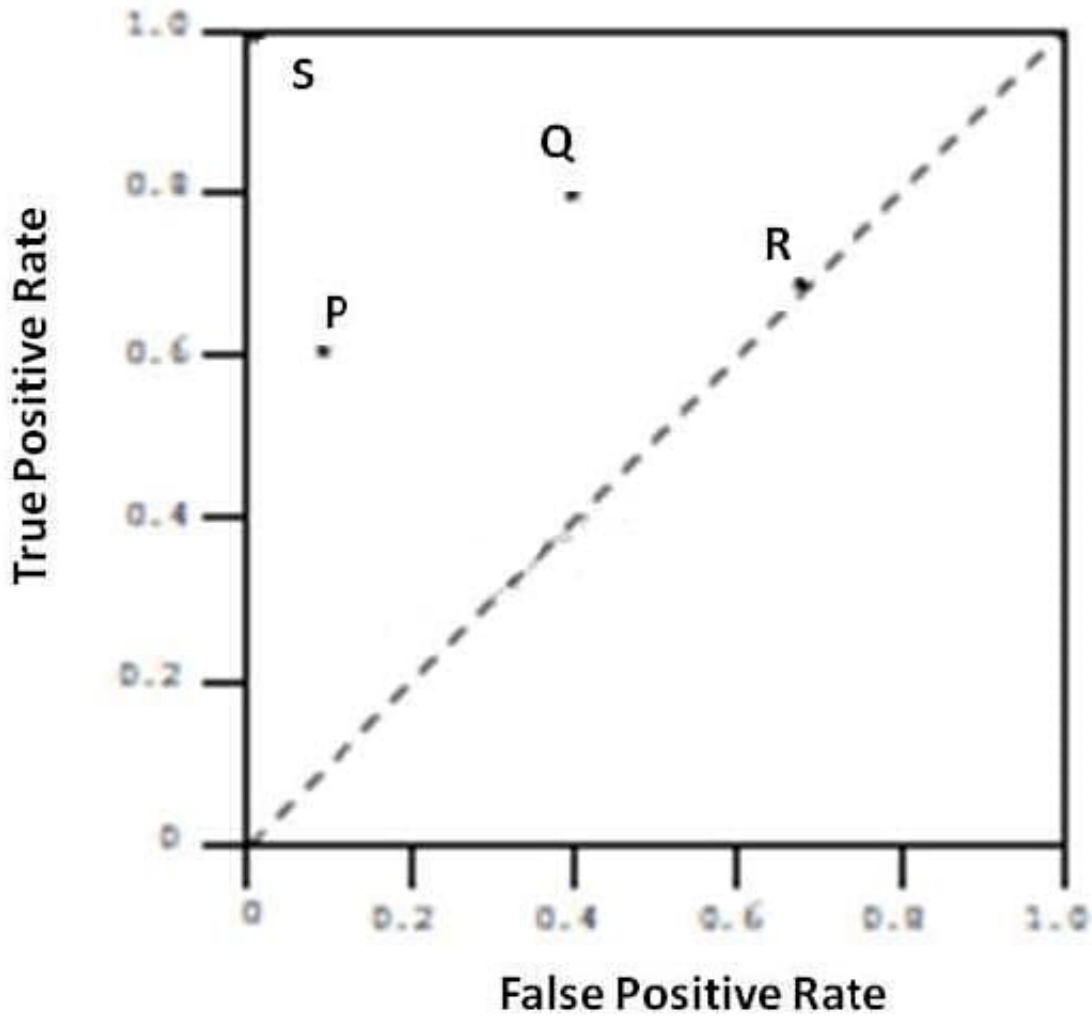
A. Ensure that the TaskTracker is running.

B. Ensure that the JobTracker is running

C. Ensure that the NameNode is running

D. Ensure that a DataNode is running

Correct Answer: A

QUESTION 6

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 7

You have been assigned to do a study of the daily revenue effect of a pricing model of online transactions. All the data

![Pass2Lead](https://Pass2Lead.com)
currently available to you has been loaded into your analytics database; revenue data, pricing data, and online transaction data. You find that all the data comes in different levels of granularity. The transaction data has timestamps (day, hour, minutes, seconds), pricing is stored at the daily level, and revenue data is only reported monthly. What is your next step?

A. Report back to the business owner that the current data model does not support the business question.

B. Interpolate a daily model for revenue from the monthly revenue data.

C. Aggregate all data to the monthly level in order to create a monthly revenue model.

D. Disregard revenue as a driver in the pricing model,and create a daily model based on pricing and transactions only.

Correct Answer: A

QUESTION 8

What is required in a presentation for business analysts?

A. Budgetary considerations and requests

B. Operational process changes

C. Detailed statistical explanation of the applicable modeling theory

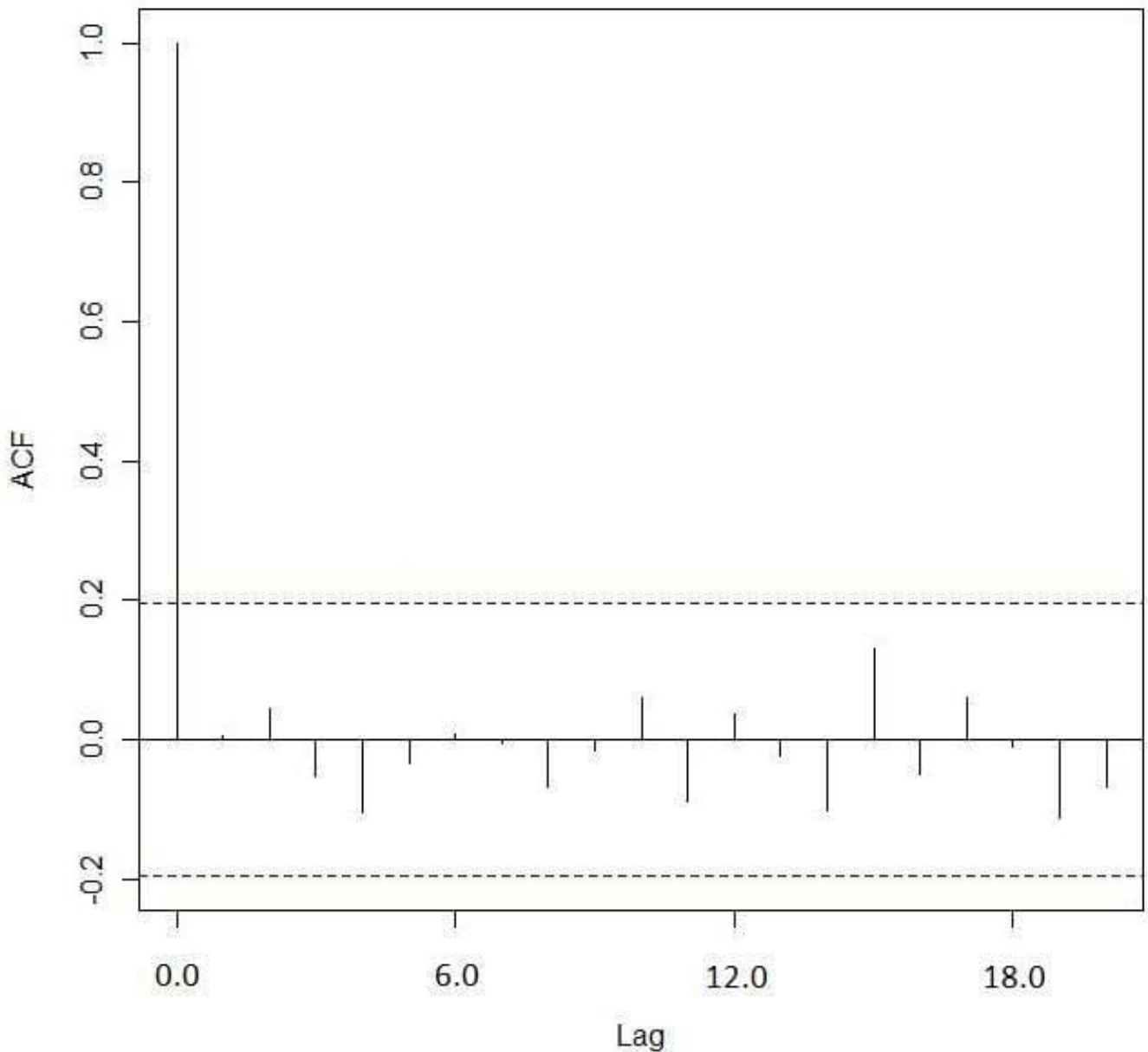D. The presentation author\\'s credentials

Correct Answer: B

QUESTION 9

Refer to the exhibit.

In the exhibit, a correlogram is provided based on an autocorrelation analysis of a sample dataset.

What can you conclude based only on this exhibit?

A. There appears to be no structure left to model in the data

B. There appears to be a seasonal component in the data

C. Lag 1 has a significant autocorrelation

D. There appears to be a cyclical component in the data

Correct Answer: A

QUESTION 10

Your colleague, who is new to Hadoop, approaches you with a question. They want to know how best to access their data. This colleague has previously worked extensively with SQL and databases. Which query interface would you recommend?

A. Hive

B. Pig

C. Howl

D. HBase

Correct Answer: A

QUESTION 11

What describes the use of UNION clause in a SQL statement?

A. Operates on queries and potentially increases the number of rows

B. Operates on queries and potentially decreases the number of rows

C. Operates on tables and potentially decreases the number of columns

D. Operates on both tables and queries and potentially increases both the number of rows and columns
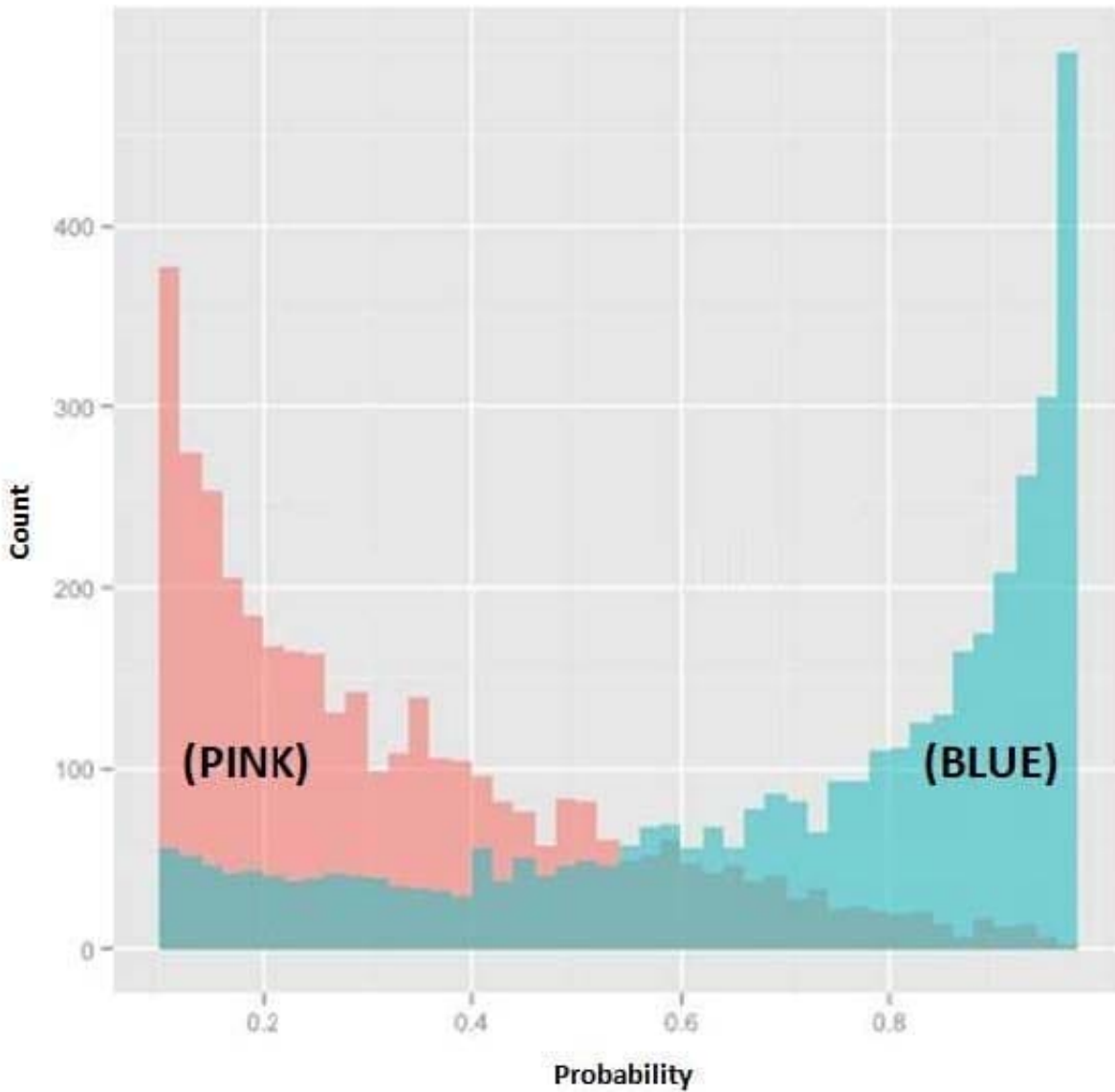
Correct Answer: A

QUESTION 12

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

QUESTION 13

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 14

In linear regression modeling, which action can be taken to improve the linearity of the relationship between the dependent and independent variables?
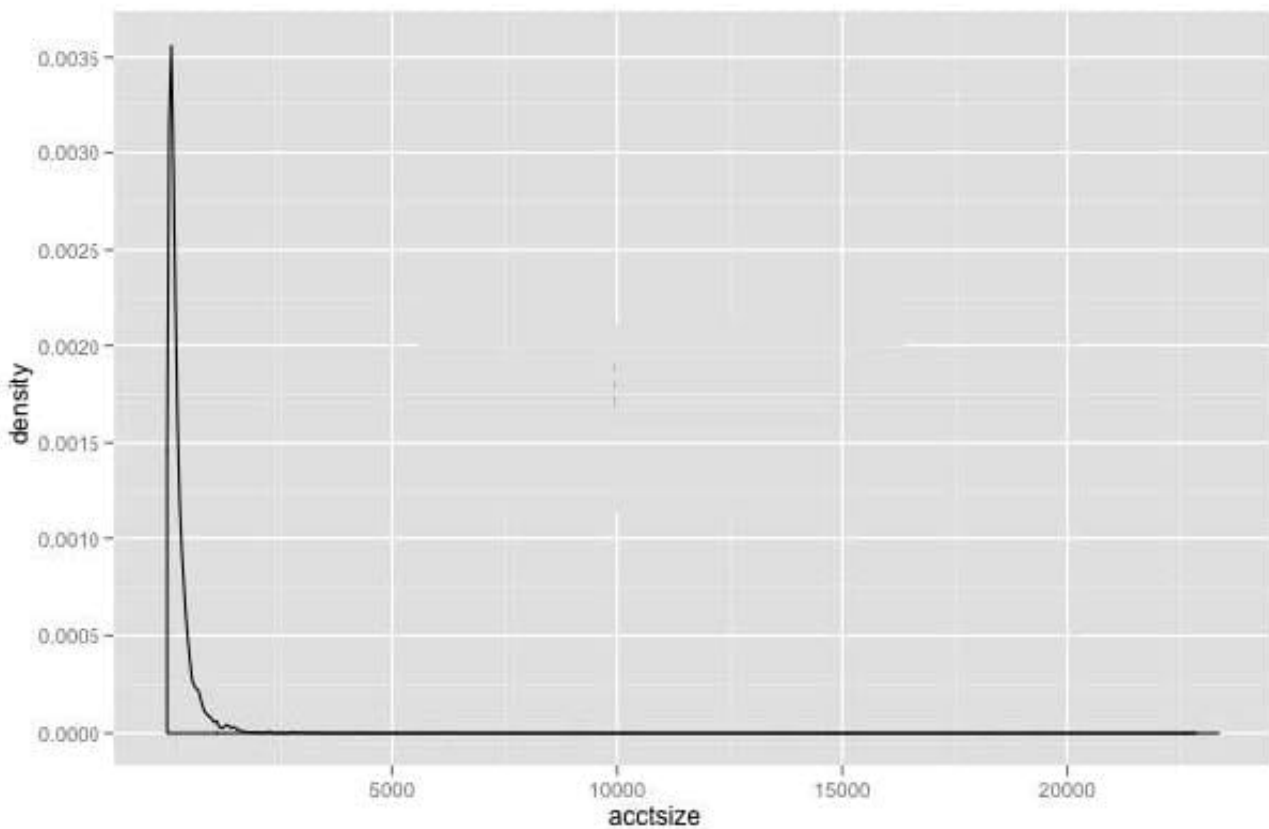
A. Apply a transformation to a variable

B. Use a different statistical package

C. Calculate the R-Squared value

D. Change the units of measurement on the independent variable

Correct Answer: A

QUESTION 15

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

![Pass2Lead](https://Pass2Lead.com)
E20-026 PDF Dumps   E20-026 VCE Dumps   E20-026 Practice Test