

**100%** Money Back  
**Guarantee**

**Vendor:**EMC

**Exam Code:**E20-026

**Exam Name:**Enterprise Storage Networking Specialist  
Exam

**Version:**Demo

### QUESTION 1

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

- A. Run a pilot
- B. Try different analytical techniques
- C. Try different variables
- D. Transform existing variables

Correct Answer: A

---

### QUESTION 2

A disk drive manufacturer has a defect rate of less than 1.5% with 98% confidence. A quality assurance team samples 1000 disk drives and finds 14 defective units. Which action should the team recommend?

- A. The manufacturing process is functioning properly and no further action is required
- B. A larger sample size should be taken to determine if the plant is operating correctly
- C. A smaller sample size should be taken to determine if the plant is operating correctly
- D. There is a flaw in the quality assurance process and the sample should be repeated

Correct Answer: A

---

### QUESTION 3

Which word or phrase completes the statement? Business Intelligence is to monitoring trends as Data Science is to \_\_\_\_\_ trends.

- A. Predicting
- B. Discarding
- C. Driving
- D. Optimizing

Correct Answer: A

---

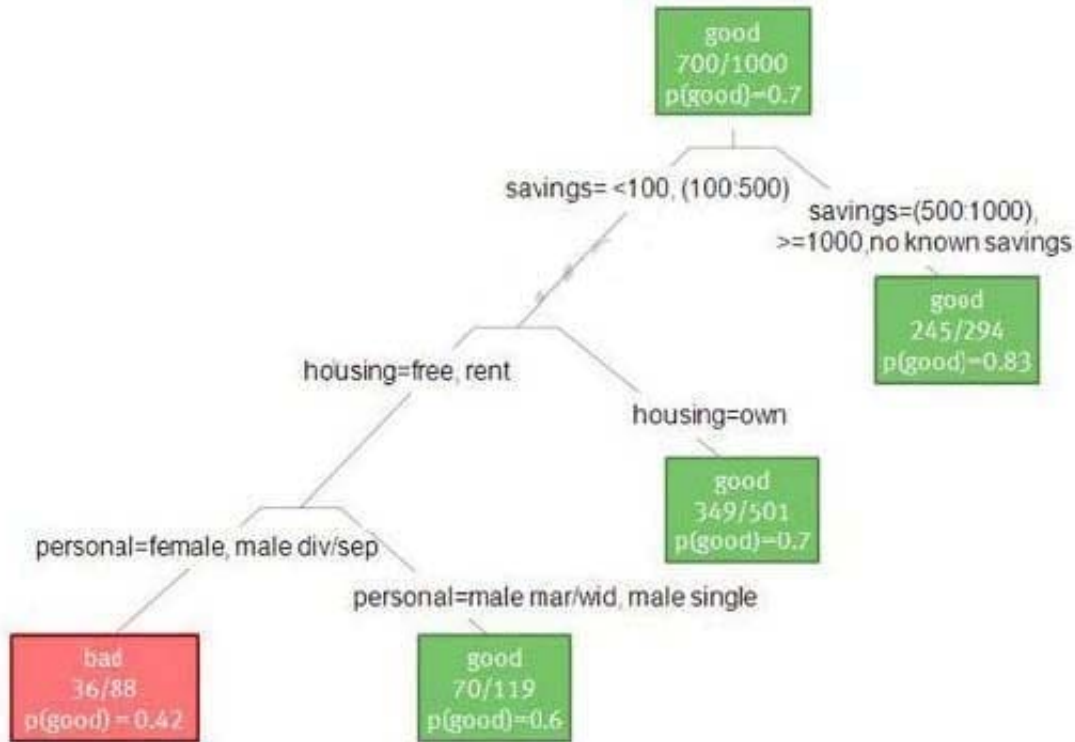
### QUESTION 4

Refer to the exhibit.

What provides the decision tree for predicting whether or not someone is a good or bad credit risk. What

would be the assigned probability,  $p(\text{good})$ , of a single male with no known savings?

## Example: The Credit Prediction Problem



- A. 0.83
- B. 0
- C. 0.498
- D. 0.6

Correct Answer: A

### QUESTION 5

You do a Student's t-test to compare the average test scores of sample groups from populations A and B. Group A averaged 10 points higher than group B. You find that this difference is significant, with a p-value of 0.03. What does that mean?

- A. There is a 3% chance that you have identified a difference between the populations when in reality there is none.
- B. The difference in scores between a sample from population A and a sample from population B will tend to be within 3% of 10 points.
- C. There is a 3% chance that a sample group from population A will score 10 points higher than a sample group from population B.
- D. There is a 97% chance that a sample group from population A will score 10 points higher than a sample group from

population B.

Correct Answer: A

---

### QUESTION 6

The Marketing department of your company wishes to track opinion on a new product that was recently introduced. Marketing would like to know how many positive and negative reviews are appearing over a given period and potentially retrieve each review for more in-depth insight. They have identified several popular product review blogs that historically have published thousands of user reviews of your company's products. You have been asked to provide the desired analysis. You examine the RSS feeds for each blog and determine which fields are relevant. You then craft a regular expression to match your new product's name and extract the relevant text from each matching review. What is the next step you should take?

- A. Convert the extracted text into a suitable document representation and index into a review corpus
- B. Use the extracted text and your regular expression to perform a sentiment analysis based on mentions of the new product
- C. Read the extracted text for each review and manually tabulate the results
- D. Group the reviews using Na.ve Bayesian classification

Correct Answer: A

---

### QUESTION 7

In R, functions like `plot()` and `hist()` are known as what?

- A. generic functions
- B. virtual methods
- C. virtual functions
- D. generic methods

Correct Answer: B

---

### QUESTION 8

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

- A.  $\{\text{grape,apple,orange}\}$  must be a frequent itemset.
- B.  $\{\text{banana,apple,grape,orange}\}$  must be a frequent itemset.
- C.  $\{\text{grape}\} \Rightarrow \{\text{banana,apple}\}$  must be a relevant rule.
- D.  $\{\text{banana,apple}\} \Rightarrow \{\text{orange}\}$  must be a relevant rule.

Correct Answer: A

---

**QUESTION 9**

Which data type value is used for the observed response variable in a logistic regression model?

- A. Any positive real number
- B. Any integer
- C. A binary value
- D. Any real number

Correct Answer: C

---

**QUESTION 10**

In which phase of the analytic lifecycle would you expect to spend most of the project time?

- A. Discovery
- B. Data preparation
- C. Communicate Results
- D. Operationalize

Correct Answer: C

---

**QUESTION 11**

In linear regression, what indicates that an estimated coefficient is significantly different than zero?

- A. A small p-value
- B. R-squared near 1
- C. R-squared near 0
- D. The estimated coefficient is greater than 3

Correct Answer: A

---

**QUESTION 12**

When would you use GROUP BY ROLLUP clause in your OLAP query?

- A. where all subtotals and grand totals are to be included in the output

B. where only the subtotals are to be included in the output

C. where only the grand totals are to be included in the output

D. where only specific subtotals and grand totals for a combination of variables are to be included in the output

Correct Answer: A