

#### Question 1

What is the main purpose of data modeling in the data science methodology? 1 / 1 point

- To refine and adjust the problem statement
- To develop models for descriptive or predictive purposes
- To select an appropriate analytical approach
- To collect and compile raw data

Correct

To develop models for descriptive or predictive purposes. Data modeling involves creating models that provide insights or predictions based on the data.

#### Question 2

How does a training set contribute to predictive modeling? 1 / 1 point

- A training set provides a set of unknown outcomes
- A training set serves as a calibration gauge for the model
- A training set contains variables that are not required for modeling
- It helps select appropriate algorithms

Correct

A training set serves as a calibration gauge for the model. Data scientists use training sets consisting of historical data with known outcomes to assess and adjust the model's performance.

#### Question 3

What is the primary purpose of model evaluation in the data science methodology? 1 / 1 point

- To deploy the model and monitor its performance
- To refine the data collection process
- To determine the optimal parameter values for model building
- To assess the quality of the model and ensure it meets the initial request

Correct

To assess the quality of the model and ensure it meets the initial request. Model evaluation helps determine if the model effectively addresses the problem and provides accurate results.

#### Question 4

What is the purpose of diagnostic measures during model evaluation? 1 / 1 point

- To assess the quality of relationships in a descriptive model
- To test the model's statistical significance
- To ensure the model is working as intended
- To refine the model's design for predictive outcomes

Correct

To ensure the model is working as intended. Diagnostic measures help identify areas where the model may need adjustments or refinements.

Question 5

What does the ROC curve help determine in model evaluation? 1 / 1 point

The relative misclassification cost of the model

The model's statistical significance

The optimal model based on diagnostic measures

The true-positive rate and false-positive rate for different criteria

Correct

The true-positive rate and false-positive rate for different criteria. The ROC curve helps visualize the performance of a binary classification model by showing how well it classifies positive and negative outcomes.