

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

1 / 1 point

- ☒ To develop models for descriptive or predictive purposes
- ☐ To refine and adjust the problem statement
- ☐ 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.

2. How does a training set contribute to predictive modeling?

1 / 1 point

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

☒ **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.

3. What is the primary purpose of model evaluation in the data science methodology?

1 / 1 point

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

☒ **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.

4. What is the purpose of diagnostic measures during model evaluation?

1 / 1 point

- ☐ To refine the model's design for predictive outcomes
- ☐ 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

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

5. What does the ROC curve help determine in model evaluation?

1 / 1 point

- ☒ The true-positive rate and false-positive rate for different criteria
- ☐ The optimal model based on diagnostic measures
- ☐ The relative misclassification cost of the model
- ☐ The model's statistical significance

☒ **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.