

Data Mining

LATEST SUBMISSION GRADE

100%

1. According to the reading, the output of a data mining exercise largely depends on:

1 / 1 point

- ☐ The programming language used.
- ☐ The data scientist.
- ☒ The quality of the data.
- ☐ The scope of the project.

✓ **Correct**
Correct.

2. When data are missing in a systematic way, you can simply extrapolate the data or impute the missing data by filling in the average of the values around the missing data.

1 / 1 point

- ☐ True.
- ☒ False.

✓ **Correct**
Correct. When data are missing in a systematic way, you should determine the impact of missing data on the results and whether missing data can be excluded from the analysis.

3. Prior Variable Analysis and Principal Component Analysis are both examples of a data reduction algorithm.

1 / 1 point

- ☐ True.
- ☒ False.

✓ **Correct**
Correct. Prior Variable Analysis is not a data reduction algorithm.

4. After the data are appropriately processed, transformed, and stored, machine learning and non-parametric methods are a good starting point for data mining.

1 / 1 point

- ☒ False.
- ☐ True.

✓ **Correct**
Correct.

5. **In-sample forecast** is the process of formally evaluating the predictive capabilities of the models developed using observed data to see how effective the algorithms are in reproducing data.

1 / 1 point

- ☒ True.
- ☐ False.

✓ **Correct**
Correct.