



Data Mining

LATEST SUBMISSION GRADE

80%

1. According to the reading, the output of a data mining exercise largely depends on the skills of the data scientist carrying out the exercise.

1 / 1 point

☐ True.

☒ False.

✓ **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

☒ False.

☐ True.

✓ **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.

0 / 1 point

☒ True.

☐ False.

! **Incorrect**
Incorrect. Data visualization is a good starting point for data mining.

5. "Formal evaluation could include testing the predictive capabilities of the models on observed data to see how effective and efficient the algorithms have been in reproducing data." This is known as:

- ☐ Prototyping.
- ☐ Overfitting.
- ☒ In-sample forecast.
- ☐ Reverse engineering.



Correct

Correct.
