

Congratulations! You passed!
Grade received 81.25%
Latest Submission Grade 81.25%
To pass 80% or higher
Go to next item

1. Which of the following statements accurately describe the general categories of feature engineering? Select all that apply. 0.5 / 1 point

- ☒ Feature selection involves taking multiple features to create a new one that will improve the accuracy of the algorithm.
- ☐ This should not be selected
Review [the video that introduces feature engineering.](#)
- ☒ Feature transformation involves modifying existing features in a way that improves accuracy when training a model.
- ☐ Correct
- ☒ The three general categories of feature engineering are selection, extraction, and transformation.
- ☐ Correct
- ☒ Feature extraction involves choosing the features in the data that contribute the most to predicting the response variable.
- ☐ This should not be selected
Review [the video that introduces feature engineering.](#)

2. A data professional resolves a class imbalance in a very large dataset. They alter the majority class by using fewer of the original data points in order to produce a split that is more even. What does this scenario describe? 1 / 1 point

- ☐ Smoothing
- ☒ Downsampling
- ☐ Upsampling
- ☐ Merging
- ☐ Correct

3. Fill in the blank: Customer _____ is the business term that describes how many customers stop using a product or service, or stop doing business with a company altogether, and at what rate this occurs. 1 / 1 point

- ☐ transfer
- ☐ retention
- ☐ exchange

☒ churn

☒ Correct

4. Naive Bayes's theorem enables data professionals to calculate posterior probability for a data project. What does posterior probability describe? 1 / 1 point

- ☐ The likelihood of an event occurring after taking into consideration only the most suitable observations and information
- ☐ The likelihood of an event occurring based upon only observations and information that align with current hypotheses
- ☒ The likelihood of an event occurring after taking into consideration all new, relevant observations and information
- ☐ The likelihood of an event occurring based upon the observations and information that were available at the start of the data project

☒ Correct

5. Fill in the blank: When using a scaler to _____ the columns in a dataset using MinMaxScaler, a data professional must fit the scaler to the training data and transform both the training data and the test data using that same scaler. 1 / 1 point

- ☐ customize
- ☒ normalize
- ☐ sort
- ☐ filter

☒ Correct

6. A data professional examines a dataset to reveal key details about the data that will help inform the plans for building a model. Which PACE stage does this scenario describe? 0 / 1 point

- ☒ Plan
- ☐ Analyze
- ☐ Construct
- ☐ Execute

☒ Incorrect

Review [the video about modeling and the PACE workflow](#).^[↗]

7. In the model-development process, which type of feature is useful by itself because it contains information that will be useful when forecasting the target?

1 / 1 point

- ☐ Irrelevant
- ☐ Redundant
- ☒ Predictive
- ☐ Interactive

☒ Correct

8. Fill in the blank: Log normalization involves reducing _____ in order to make and making the data more effective for modeling.

1 / 1 point

- ☐ probability
- ☐ continuity
- ☒ skew
- ☐ normality

☒ Correct