1. B
2. A
3. A
4. C
5. D
6. B
7. C
8. A and D
9. B
10. A
11. One-hot encoding creates d-dimensional vectors for each instance where d is the unique number of feature values in the dataset. For a feature having a large number of unique feature values or categories, one-hot encoding is not a great choice. Binary encoding might be a good alternative to one-hot encoding because it creates fewer columns when encoding categorical variables. Ordinal encoding is a good choice if the order of the categorical variables matters.
12. When we are using an imbalanced dataset, we can oversample the minority class using replacement. This technique is called oversampling. Similarly, we can randomly delete rows from the majority class to match them with the minority class which is called undersampling.
13. The key difference between ADASYN and SMOTE is that the former uses a density distribution, as a criterion to automatically decide the number of synthetic samples that must be generated for each minority sample by adaptively changing the weights of the different minority samples to compensate for the skewed distributions. The latter generates the same number of synthetic samples for each original minority sample.
14. GridSearchCV is the process of performing hyperparameter tuning in order to determine the optimal values for a given model. if you do grid search on all of your data, the error on your test set will be biased low, and when you go to apply your model to new data, the error could be much higher (and likely will, except for the effects of randomness).
    In summary, you should only use gridsearch on the training data after doing the train/test split, if you want to use the performance of the model on the test set as a metric for how your model will perform when it really does see new data.
15. There are three error metrics that are commonly used for evaluating and reporting the performance of a regression model; they are:
    Mean Squared Error (MSE)- It is also an important loss function for algorithms fit or optimized using the least squares framing of a regression problem. Here “least squares” refers to minimizing the mean squared error between predictions and expected values.
    Root Mean Squared Error (RMSE)- the square root of the error is calculated, which means that the units of the RMSE are the same as the original units of the target value that is being predicted.
    Mean Absolute Error (MAE)- MAE, is a popular metric because, like RMSE, the units of the error score match the units of the target value that is being predicted.