Question 1

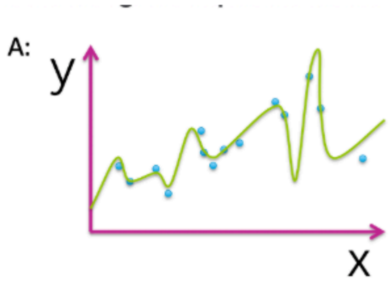
The model that best minimizes training error is the one that will perform best for the task of prediction on new data.

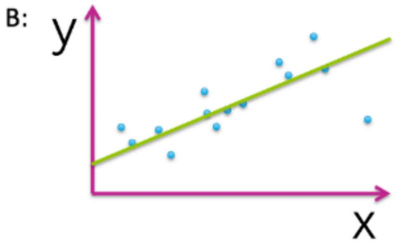
a. True

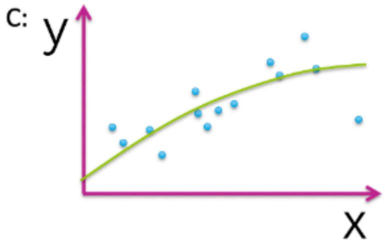
[✓ b. False]

Question 2

Which figure represents an overfitted model?







a. B

[✓ b. A]

c. C

Question 3

If the features of Model 1 are a strict subset of those in Model 2, which model will **USUALLY** have lowest **TRAINING** error?

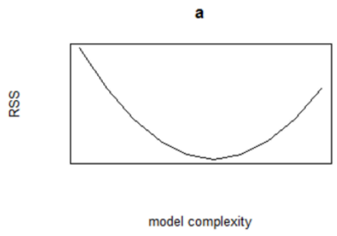
a. It’s impossible to tell with only this information

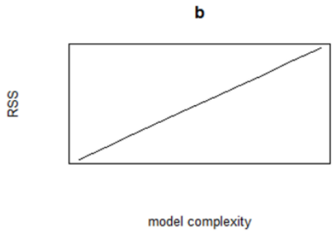
[✓ b. Model 2]

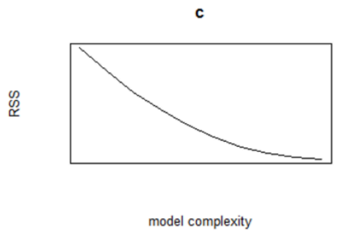
c. Model 1

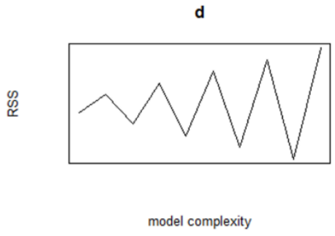
Question 4

Which of the following plots of model complexity vs. Residual Sum of Squares (RSS) is most likely from **TRAINING** data (for a fixed data set)?









a. C
b. A
c. D
d. B

Question 5
It is **always** optimal to add more features to a regression model.
a. True
b. False

Question 6
A simple model with few parameters is most likely to suffer from:
a. High Variance
b. High Bias

Question 7
A common process for selecting a parameter like the optimal polynomial degree is:
a. Bootstrapping
b. Minimizing test error
c. Minimizing validation error
d. Model estimation
e. Multiple regression

Question 8
Selecting model complexity on test data (choose all that apply)
a. Is computationally inefficient
b. Provides an overly optimistic assessment of performance of the resulting model
c. Should never be done
d. Allows you to avoid issues of overfitting to training data

Question 9
Which of these could be an acceptable sequence of operations using scikit-learn to apply the k-nearest neighbors classification model?
a. read_table, train_test_split, fit, KNeighborsClassifier, score
b. read_table, train_test_split, KNeighborsClassifier, fit, score
c. KNeighborsClassifier, train_test_split, fit, score, read_table
d. read_table, fit, train_test_split, KNeighborsClassifier, score

Question 10
Given a dataset with 10,000 observations and 50 features plus one label, what would be the dimensions of X_train, y_train, X_test, and y_test? Assume a train/test split of 75%/25%.
a. X_train: (2500,) y_train: (2500, 50) X_test: (7500,) y_test: (7500, 50)
b. X_train: (10000, 28) y_train: (10000,) X_test: (10000, 12) y_test: (10000,)
c. X_train: (2500, 50) y_train: (2500,) X_test: (7500, 50) y_test: (7500,)

d. X_train: (7500, 50) y_train: (7500,) X_test: (2500, 50) y_test: (2500,)
e. X_train: (10000, 50) y_train: (10000,) X_test: (10000, 50) y_test: (10000)