Q1-1: How many distinct (binary classification) decision trees are possible with *4* Boolean attributes? Here distinct means representing different functions.

1. 2^4
2. 2^8
3. 2^{16}
4. 2^{32}

Q1-1: How many distinct (binary classification) decision trees are possible with *4* Boolean attributes? Here distinct means representing different functions.

1. 2^4
2. 2^8
3. 2^{16} ←
4. 2^{32}

#distinct decision trees
= #distinct Boolean functions
= #functions of 2^4 = 16 inputs, binary label for each input
= **2^{16}**

Q1-2: Which of the following statements is TRUE?

1. If there is no noise, then there is no overfitting.
2. Overfitting may improve the generalization ability of a model.
3. Generalization error is monotone with respect to the capacity/complexity of a model.
4. More training data may help preventing overfitting.

Q1-2: Which of the following statements is TRUE?

1. If there is no noise, then there is no overfitting.
2. Overfitting may improve the generalization ability of a model.
3. Generalization error is monotone with respect to the capacity/complexity of a model.
4. More training data may help preventing overfitting. ⬅

1. We can still have false correlation that leads to overfitting.
2. Overfitting would undermine the generalization ability.
3. Generalization error would first decrease and then increase as the model capacity increases.
4. Increasing training data size would help better approximate the true distribution.

Q2-1: Are these statements true or not?
(A) The accuracy of a model is the training set accuracy, and its estimator is the test set accuracy.
(B) An unbiased estimator $\hat{\theta}$ always equals to its corresponding true parameter θ.

1. True, True
2. True, False
3. False, True
4. False, False

Q2-1: Are these statements true or not?
(A) The accuracy of a model is the training set accuracy, and its estimator is the test set accuracy.
(B) An unbiased estimator $\hat{\theta}$ always equals to its corresponding true parameter θ.

1. True, True
2. True, False
3. False, True
4. False, False ⬅

(A) The accuracy of a model should be based on the true distribution. The training set and test set only approximate the true distribution.
(B) An unbiased estimator equals to the true parameter in expectation, which means that they won't always be the same for single estimate but the average of a large number of estimates would well approximate the true parameter. An unbiased estimator just makes sure that there's no systematic error.

Q2-2: Are these statements true or not?
(A) The sample size on the learning curve is the size of test set.
(B) A larger training set would provide a lower variance estimate of the accuracy of a learned model.

1. True, True
2. True, False
3. False, True
4. False, False

Q2-2: Are these statements true or not?
(A) The sample size on the learning curve is the size of test set.
(B) A larger training set would provide a lower variance estimate of the accuracy of a learned model.

1. True, True
2. True, False
3. False, True
4. False, False ←

(A) The sample size on the learning curve is for training set.
(B) A larger test set rather than a larger training set does so.

Q2-3: Which of the following is NOT true?

1. Random resampling can tell us how sensitive accuracy of a learning method is.
2. Class proportions are maintained same in the stratified sampling.
3. In leave-one-out cross validation, the number of partition equals to the number of instances.
4. In cross validation, we are evaluating the performance of an individual learned hypothesis.

Q2-3: Which of the following is NOT true?

1. Random resampling can tell us how sensitive accuracy of a learning method is.
2. Class proportions are maintained same in the stratified sampling.
3. In leave-one-out cross validation, the number of partition equals to the number of instances.
4. In cross validation, we are evaluating the performance of an individual learned hypothesis.

In cross validation, we are evaluating a learning method as opposed to a specific individual learned hypothesis.