1. **What is the estimated depth of a Decision Tree trained (unrestricted) on a one million instance training set?**

The depth of a well-balanced binary tree containing m leaves is equal to log2(m), rounded up Thus, if the training set contains one million instances, the Decision Tree will have a depth of log2  
(106) ≈ 20.

2**. Is the Gini impurity of a node usually lower or higher than that of its parent? Is it always lower/greater, or is it usually lower/greater?**

A node's Gini impurity is generally lower than its parent's. This is ensured by the CART training algorithm's cost function, which splits each node in a way that minimizes the weighted sum of its children's Gini impurities. However, if one child is smaller than the other, it is possible for it to have a higher Gini impurity than its parent, as long as this increase is more than compensated for by a  
decrease of the other child's impurity.

3**. Explain if its a good idea to reduce max depth if a Decision Tree is overfitting the training set?**

If a Decision Tree is overfitting the training set, it may be a good idea to decrease max\_depth, since this will constrain the model, regularizing it.

**4. Explain if its a good idea to try scaling the input features if a Decision Tree underfits the training set?**

Decision Trees don't care whether or not the training data is scaled or centered, **scaling the input features will just be a waste of time.**

5**. How much time will it take to train another Decision Tree on a training set of 10 million instances if it takes an hour to train a Decision Tree on a training set with 1 million instances?**

The computational complexity of training a Decision Tree is O(n × m log(m)). So if you multiply the training set size by 10, the training time will be multiplied by K = (n × 10m × log(10m)) / (n × m × log(m)) = 10 × log(10m) / log(m). If m = 10^6, then K ≈ 11.7, so you can expect the training time to be roughly 11.7 hours.

**6. Will setting presort=True speed up training if your training set has 100,000 instances?**

Presorting the training set speeds up raining only if the dataset is smaller than a few thousand instances. If it contains 100,000 instances, setting presort=True will considerably slow down training.

7 –

8-