Reading about Random Forests in *An Introduction to Statistics Learning*

I recommend to begin reading from the beginning of Chapter 8 to get the general tree building ideas in the context of both regression and classification trees. However I suggest skipping the Tree Pruning section in Section 8.1.1 because individual trees "generally do not have the same level of predictive accuracy as some of the other regression and classification approaches seen in this book." Do not skip section 8.1.2 on classification trees and pay particular attention to Gini index and the cross entropy alterative. Also read the rest of Section 8.1.

We focus quickly on Random Forests with description beginning in Section 8.2   The class Powerpoint on Random Forests Concepts in the course folder has more information. You may skip 8.2.3 on Boosting (although it is good to know this methodology) and jump to section 8.3.2 to run the first three lines of the R script that select a training set.   Then jump to Section 8.3.3 on Bagging and Random Forest and run that example. This is instructive but nothing is due. The graphics due use other data as indicated in assignment R scripts.   However the assignment posting have you skip some of these.