CSE574 Introduction to Machine Learning
Programming Assignment 2 (in R-Studio)
Lab: Decision Trees
Due Date: **April 24, 2018 before midnight**

For programming assignment 2, we will read *An Introduction to Statistical Learning with Applications in R* (ISLR) chapter 8 *Tree-Based Methods*; and, individually replicating 8.3 *Lab: Decision Trees*. Chapter 8 of ISLR is available on Piazza.

TASK 0 Preparation 10 points (state in your report the platform you use)
Setup (or find on a campus system) your R infrastructure. You will want to use R-Studio and R-markdown (an executable document environment for producing your report). If not previously installed, for your platform:

1. Install R https://cran.rstudio.com/
2. Install R-Studio https://www.rstudio.com/products/rstudio/download/#download
3. Install R packages required for this lab
   `install.packages(c("tree","ISLR","MASS","randomForest","gbm"))`
   (You may need to have a compiler and the packages "devtools" installed as well, as MASS seems to compile source code during install. Windows probably has binary available, if not see https://cran.rstudio.com/bin/windows/Rtools/ for a compiler and related tools. Then, `install.packages("devtools") )`

TASK 1 Fitting Classification Trees 20 points
Replicate section 8.3.1, turn in output (R-markdown output is probably easiest; pdf, docx, or html is fine)

TASK 2 Fitting Regression Trees 20 points
Replicate section 8.3.2, turn in output (pdf, docx, or html is fine)

TASK 3 Bagging and Random Forests 20 points
Replicate section 8.3.3, turn in output (pdf, docx, or html is fine)

TASK 4 Boosting 20 points
Replicate section 8.3.4, turn in output (pdf, docx, or html is fine)

TASK 5 Summary 10 points
Report time required for homework, difficulties encountered, and collaborators.

Submission requirements
Each student should submit a zip file name named *<ubperson#>**-project2.zip***
The zip file should contain a single file (pdf, docx, html) containing your output for Task 1-4, plus your comments for Task 5.
Remember: submit your own work; collaboration is ok; accessing libraries / other resources on web is ok. The report must be your own, meaning: if we ask you about your work, you must be able to explain it in detail to us.