HW 7

Due: April 9th at 3:00 pm

1. Use the college applications dataset from James, Witten, Hastie and Tibshirani's book, available as `College` in the `ISLR` package, or here: http://www-bcf.usc.edu/~gareth/ISL/College.csv

   a) Use `set.seed(25)` and split the dataset randomly into 50% training set and 50% test set. Report the average applications received (`Apps` variable) in each of the two datasets.

   b) Fit a OLS linear regression model on the *training* set to predict applications based on the other numerical variables. Report the *test* set mean prediction error.

   c) Fit a ridge regression model on the *training* set, with λ chosen by cross-validation (use the `glmnet` package and examples from class). Report the *test* set prediction error.

   d) Repeat part c) but this time use lasso model.

   e) Compute and compare the R^2 values of the three models computed on both the training and test datasets.

   f) Compare and comment on the three models.