The task for today's lab is to simply get you to fit a regularized multiple regression model. The package I recommend `glmnet` is not as streamlined as the function `lm`, so it takes some effort to set up a model – but you'll be glad we're not working in Python.

You may pick a dataset to work with if you're done looking at the `bike` dataset, otherwise it seems reasonable to me to keep playing around to try to predict `cnt`.

Below are some specific tasks/points about regularization that you should understand and think through, no matter what dataset you choose to work on.

1. Penalty term. Look at the parameter `alpha` under the Arguments section and the objective function under the Details section of the help file for `glmnet`.

   ```
   ?glmnet
   ```

   (a) What piece of the puzzle are they calling the penalty term?
   (b) What piece of the puzzle are they calling the objective function? Notice that I had one two many λs in my objective function in class yesterday.

2. Ridge versus Lasso.
   (a) What is the main difference between Ridge regression and Lasso?
   (b) When might you be interested in one versus the other? Hint: what is the standard error of a coefficient forced to zero?

3. `cv.glmnet`. Notice that the output of `cv.glmnet` provides two specific values of λ to choose. Explain the difference?

4. Prediction. Theoretically, there exists a value for λ that minmizes Mean Squared Error (MSE), better than does say un-regularized multiple linear regression. It's not uncommon that the value `lambda.1se` does not reduce MSE compared to un-regularized multiple regression. Does this happen for you?

   (a) Manipulate the following pseudo code to create a (singluar) training and a (singular) testing dataset.

   ```
   library(caret)
   train_idx <- createDataPartition(y, p=0.75, list=FALSE)
   training <- df[train_idx, ]
   testing <- df[-train_idx, ]
   ```

   (b) Calculate predictions $\hat{y}$ for three separate models: un-regularized multiple regression, regularized multiple regression with `lambda.min`, and regularized multiple regression with `lambda.1se`.

   (c) Which of these three models produces the smallest MSE on the testing dataset?

5. Plots describing λs.

   (a) Replicate the plot that has the magnitude of your coefficients on the y-axis and the λs on the log scale on the x-axis?

   (b) How many variables are included in the Lasso version of your model? How many did you start with?

6. Overfitting. Compare the magnitude of the coefficients under regularized and unregularized models. Does the amount of regularization recommended by `cv.glmnet` seem surprising?