# 36-402/608        Homework #8      due 10:30AM 3/18
(Optional)

1. Pig fat (50 points)

   (a) Look at Sleuth problem 17.08 and the answer on page 528. Load the data from ex1708.csv. Verify that you cannot run `step(lm(fat ., data=fat), direction="backward")` (even after correcting for whatever you called your data.frame and the upper/lower case of your variables). Although R could have sensible carried out this analysis, why do you think it did not?

   (b) Starting with my code, add enough code to carry out the forward selection to the point where the AIC starts to rise. Turn in the `summary(lm())` of the best model using forward selection with AIC.

   ```
   aic1 = rep(NA,13)
   for (i in 1:13) aic1[i] = AIC(lm(fat$fat~fat[,i+1]))
   one = which.min(aic1)
   cat("AIC with m", one, " = ", aic1[one], "\n", sep="")
   # AIC with m8=56.91094
   aic2 = rep(NA,13)
   for (i in (1:13)[-one]) aic2[i] = AIC(lm(fat$fat~fat[,one+1]+fat[,i+1]))
   two = which.min(aic2)
   cat("+m", two, " = ", aic2[two], "\n", sep="")
   # +m4 = 48.16103
   ```

   (c) Carry out PCA on the pig fat dataset. Turn in the cumulative variance percents and the loadings (rotation) for the first 4 PCs.

   (d) Let $k$ be the number of variables chosen in part $b$. Create $k$ new variables that are suggested by the first $k$ PCs, but are simpler and more interpretable. Turn in the R code to create the variables and the `summary(lm())` that regresses the pig fat on the $k$ new variables.

   (e) Write a paragraph explaining what you think the pig farmers should do in practice to predict pig fat from MRI.

2. Insurance (50 points)

   Do Sleuth problem 17.14.

   For part d, you can use the code in zipmap.R to make a (crude) map.