STAT 6430 Homework 4

## Please submit your annotated R code in a file titled TAO3.R, along with the text file requested below.

This Team Assignment requires predictions for the value of a response variable based on the value of a seven explanatory variables. For this assignment, you should use a linear model to make your predictions. It is fine to use transformations and variable selection, provided your predictions come from a linear model.

For this problem, once you have your predictions, put them into the vector predvect. (Only the predictions, not the x values.) Execute the code

```
write.table(predvect, file = "TAO3preds.csv", row.names=F, col.names=F, sep=",")
```

This should generate a CSV file TAO3preds.csv that will correctly format your predictions for review. (Open in Excel – you should see the first column (no header) filled with your predictions.)

For this assignment, you will submit your R code. Your code should be neatly organized with comments to provide documentation. You should not only include the code used to make your predictions, but also include a brief statement describing how you arrived at your choice. The code should be included in a single R file, labeled TAO3.R.

1. This problem requires the data contained in two files, TAO3train.csv and TAO3predict.csv. The objective is to predict the y-value based on the given values of  $x_1, \ldots, x_7$ . Use the data in TAO3train.csv to develop a linear model, then predict the y-values corresponding to the x-values contained in TAO3predict.csv. Export your predictions to the file named TAO3preds.csv.

When you are finished, upload the files TAO3.R and TAO3preds.csv into the Assignments in Collab. (As usual, one set of submissions per team.)