Please submit your annotated R code in a file titled `hw05.R`,
along with the text files requested below.

For the first problem, once you have your predictions, put them into a vector such as `predvect`. (Only the predictions, not the x values.) Executing the code

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

will generate a CSV file `predfile.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 `hw05.R`.

1. For this problem you are to predict the mileage (mpg = miles per gallon) for a set of automobiles. The variables in the data are

   ```
   #     mpg = miles per gallon (response)
   #     cyl = number of cylinders (4, 6, or 8)
   #     disp = engine displacement (cubic inches)
   #     hp = engine horse power
   #     accel = 0-to-60 time (seconds)
   #     year = year car produced
   ```

   The file `hw05p01train.csv` contains data to use to train your model(s), and the file `hw05p01predict.csv` gives explanatory variables to be used for predicting mpg. As usual, your goal is to minimize the MSE for the prediction data. Don't use any transformations, just adjust your model through variable selection.

2. For this problem you will need to import the text file `comscore.csv`. This file contains internet site visits logged by ComScore, an internet analytics company providing marketing data and analytics to many of the worlds largest enterprises, agencies, and publishers. (Wikipedia) The format of the file is Comma Separated Values (CSV), so it can be opened in either a text editor or Excel. Review the file, and then answer the following questions. Put your answers in the R file.

   (a) Determine the total of the average daily visits (column J) for the 100 sites with a digit in the first column.

   (b) Determine the total number of unique visits for the sites with a "[G]" in the second column.

   (c) Determine the total number of unique visits for the sites with a reference to Canada in the third column.

   (d) Find the sum of the total pages viewed (column L) for all entries with an "[S]" in the second column that have a figure reported for total pages viewed.

When you are finished, upload the files `hw05.R` and `hw05p01mypredictions.csv` into the Assignments in Collab.