1. (40 pts.) Consider the dataframe `Cars93`, (from library `MASS`). It is of interest to predict `MPG.highway` using variables `Horsepower, Weight, EngineSize, RPM, Length, and Width`.

   a) Build a histogram of relative frequencies of `MPG.highway`.

   b) Fit the full regression model with no interactions. Report the coefficients table (do not simplify the model).

   c) What are the *standard regression assumptions*? Are these assumptions reasonable for the fitted model? State your conclusions but do not simplify the model.

2. (30 pts.) Consider predicting the price of a new car using observations from the csv file `GMcars` from folder `Exam 1` in Blackboard. Fit a multiple regression model with predictors `Mileage` (X_1) and factors, the number of doors (X_2) and of cylinders (X_3), and all interactions. Report

   a) the coefficients table

   b) the fitted equation for estimating the price of a car with two doors and six cylinders.

   Do not simplify the model.

3. (30 pts.) Consider the data frame `Hitters`, (from library `ISLR`) which contains 20 variables about major league players. It is of interest to predict a baseball player's `Salary` using as predictors various statistics associated with his performance during the year.

   a) Remove all rows that have missing values using function `na.omit()` with argument the name of the dataframe.

   b) Set the seed equal to 1.

   c) Use function `regsubsets(...,nvmax=19)` to select the best model using 10-fold cross validation. By default `regsubsets()` finds best models up to eight predictors unless the argument `nvmax` is used (here we have 19 predictors). Use `nvmax` at all times.

   d) Report the adjusted R-squared for the best model.

SUBMISSION - Convert your report to a pdf file. Then submit to the instructor's email.