Instructions: As always, evaluate assumptions, apply transforms if needed, and answer the questions (!). Feel free to include graphics and tables if it will help your answer, but only if it will help your answer. And a hint: if you use SMA regression and want to use AICctab, include this inside the parentheses: nobs=length(X) where X is one of the variables, and length(X) just gives a count of the number of observations (nobs) for AICctab to use (and not needed for lm).

1. It's World Series Season! So here's a baseball data set, from 2010. The MLB2020.txt lists batting statistics for every baseball player in the 2010 season. Variables we will use for each batter are:
   - G = games played
   - AB = at bats
   - R = runs scored
   - H = hits
   - HR = home runs
   - OBP = on base percentage
   - AVG = batting average

   [2 pts] Since scoring runs is the reason for batting, which variable most plausibly predicts runs scored (R) by a batter?

2. The cod fishery in the North Atlantic is heavily fished, and thus divided into mapped zones for licensed harvest (Fishery; estimated tons harvested per zone), with each of those zones evaluated for fish density (Density; estimated number of adult fish per $100m^2$) and recruitment (Recruits; estimated number of juvenile fish per $100m^2$).
   [2 pts] Which of the two estimates (Density or Fishery) in the data (fishery.txt) best predicts Recruits, which can then be used to organize next year's harvest?

3. [2 pts] Provide general recommendations to the National Marine Fisheries Service for cod fishing in the following year. Use graphs and statistical results from the fishery data above to justify your recommendations.