STAT 428/528, HW2, Due March 5

Problems 1–3 use the data `salary.dat` that was analyzed in class and is available at the class website.

1. For the salary data used in class, we didn't check the residual plots for the final model. Read the data and fit the model used in class with sex, year, and rank as the predictors. Make sure that rank is treated as a factor variable. Make residual plots and comment on what you see. In particular, identify an observation that appears to be an outlier. What makes this observation unusual?

2. Reanalyze the salary data with the outlier identified in question 1 removed. You should now have 51 observations. Use stepwise automated model selection using the BIC criterion to select a model. What is the best model using this criterion? Does it differ (in terms of predictors retained) from the best model when the outlier is included in the model?

3. Find a 95% confidence interval for the effect of sex with the outlier in the model (this was done in class) and the outlier removed. How does removal of the outlier effect the confidence interval?

4. Here is some historical data relating temperature for boiling points (in Fahrenheit) and atmospheric pressure. Atmospheric pressure is related to altitudes, and the measurements were done at higher altitudes, where boiling points for water are lower than the usual 212 degrees F. Treat the boiling point (temperature) as a response variable, and the pressure as the predictor. Plot the data and fit a simple linear regression model and a quadratic model. Is the quadratic model signficantly better than the linear model? Justify your answer. Also plot residual plots and comment on the adequacies of the models. You'll have to copy and paste the data into a file. Note that there are only two variables and 31 observations.

Temp	Pressure	Temp	Pressure
210.8	29.211	210.2	28.559
208.4	27.972	202.5	24.697
200.6	23.726	200.1	23.369
199.5	23.030	197.0	21.892
196.4	21.928	196.3	21.654
195.6	21.605	193.4	20.480
193.6	20.212	191.4	19.758
191.1	19.490	190.6	19.386
189.5	18.869	188.8	18.356
188.5	18.507	185.7	17.267
186.0	17.221	185.6	17.062
184.1	16.959	184.6	16.881
184.1	16.817	183.2	16.385
182.4	16.235	181.9	16.106
181.9	15.928	181.0	15.919
180.6	15.376		