ERSH 8320 Homework #3
Due Thursday, October 11th at 11:59:59 PM

This data appeared in the Wall Street Journal on March 1, 1984. The advertisements were selected by an annual survey conducted by Video Board Tests, Inc., a New York ad-testing company, based on interviews with 20,000 adults who were asked to name the most outstanding TV commercial they had seen, noticed, and liked. The retained impressions were based on a survey of 4,000 adults, in which regular product users were asked to cite a commercial they had seen for that product category in the past week.

Number of cases:
21
Variable Names:

1. FIRM: Firm name
2. SPEND: TV advertising budget, 1983 ($ millions)
3. MILIMP: Millions of retained impressions per week

a)   Fit a regression model for MILIMP using SPEND. State the estimated regression parameters and their standard errors.

b)   Following the regression, are any outliers present in the analysis? Describe your conclusion by showing any plots and/or statistics.

c)   Does the assumption of normally distributed error terms hold for this analysis? Provide any necessary plots and/or statistical tests (using a Type-I error rate of 0.05). For the statistical test, describe what is being tested, how it is being tested, and how you used this to reach your conclusions.

d)   Overall, does this model significantly account for variation in Y? Provide the statistical test used to answer this question (using a Type-I error rate of 0.05). For the statistical test, describe what is being tested, how it is being tested, and how you used this to reach your conclusions.

e)   What is the observation with the largest leverage? Delete this observation, rerun the regression analysis, and report the estimated regression parameters, their standard errors, and the overall model R^2.

f)   Return the original data (replace the deleted observation in e). What is the observation with the Cook's D? Delete this observation, rerun the regression analysis, and report the estimated regression parameters, their standard errors, and the overall model R^2.

g)   Which, if any, observations would you exclude from this analysis? Provide support for your answer in the form of evidence you find from statistical values you find in the analysis.