ADA2:

Assignment 1: Due 01/31/2019 Thursday in class

The director of admissions of a small college selected 120 students at random from the new freshman class in a study to determine whether a student's grade point average (GPA) at the end of the freshman year (Y) can be predicted from students' ACT test score (test-score), IQ score (IQ), and high school average score (HS-Avg).

`ex.data<-read.table("http://www.math.unm.edu/~luyan/ADA219/gpa.txt",header=T)`

1. Plot scatter plot matrix, correlation matrix of Y and predictors, and briefly comment on them.

2. Fit a multiple regression model with the three predictors IQ, HS-Avg and test-score. State model assumptions and the fitted regression model.

3. Evaluate model assumptions using appropriate plots and tests.

4. Are there any outlying residuals in this data set? Are there any outliers in any of the predictor variables? If so which observation(s) and which predictor? Explain how this observation can be a fairly extreme outlier in terms of its x value but still not produce an outlying e_i value.

   Let's continue with the following problems even if the model above is inappropriate

5. Conduct statistical tests to determine the importance of each predictor variable. State hypotheses, p-values, and decisions. What do you conclude?

6. What is the multiple coefficient of determination? Interpret.

7. Obtain 95% CI for expected freshman GPA for a student with test score= 32, IQ = 82, HS-Avg= 72. Interpret.

8. Frank Buffay obtained a score of 32 on the test-score. He also has IQ = 82, HS-Avg= 72. Predict his freshman GPA using a 95% prediction interval. Interpret your prediction interval.

9. Consider the linear regression model with GPA as the response and test score (X_1), IQ (X_2), and HSAvg (X_3) as the predictors.

   Fit the model with X_1 as the only predictor. Then fit the model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$. Comment on the tests on β_1 in these two models. How do you explain the difference?

10. Now try a transformation on Y if needed, and reevaluate model assumptions.