BUAN 6356: Problem Set Sample

This is a sample problem set. I will complete this as if it were a problem set. You can find my results stored in the ps1 folder. YOU SHOULD **NOT** COMPLETE THIS ASSIGNMENT.

Deliverable: (what you turn in) an R source-code file named ps-sample.r

Question 1

Data: (A description of the data set)

> The data for this question comes from the file Meap93.csv. This data is a summary of various schools in Michigan, their enrollments, information on their teachers (salaries, totals, etc), and their pass rates for the MEAP (a standardized test used in Michigan).

Analysis: (Steps for you to complete in your code)

- Read the data Meap93.csv into a new variable using either the data.table package (use a data.frame if you still can't install data.table): context1
- Use summary statistics to familiarize yourself with the data.
- Generate a new variable that is the sum of salary and benefits. Name it: totcomp
- Generate a new variable representing the natural logarithm of enrollment. Name it: lenroll
- Run the following linear model using the 'lm' function. Store the result in: model1

$$math10_i = \beta_0 + \beta_1 totcomp_i + \beta_2 lenroll_i + e_i \quad (1)$$

- Run the following linear model using the 'lm' function. Store the result in: model2

$$sci11_i = \beta_0 + \beta_1 totcomp_i + \beta_2 lenroll_i + e_i \quad (2)$$

Interpretations: (Add these as comments in your code)

a. Interpret the estimated coefficient on totcomp from model1 (eq 1).
b. Interpret the estimated coefficient on lenroll from model1 (eq 1).
c. Interpret the estimated coefficient on totcomp from model2 (eq 2).
d. Interpret the estimated coefficient on lenroll from model2 (eq 2).
e. Interpret the estimated intercept from model2 (eq 2).