

**ASSIGNMENT 6**  
**SPECIAL TOPICS IN EDUCATIONAL DATA MINING**  
**PROFESSOR RYAN S.J.d. BAKER**  
**SOCIAL NETWORK ANALYSIS**  
**DUE NOON, MONDAY MARCH 25**

The goal of this assignment is to build a regression model predicting student course grades, using the data in Asgn6-data-v4.xlsx.

The data set has two sheets with data. The first sheet has each student's grade. The second sheet has data from student use of an online class discussion forum. Each row represents a post on the discussion forum. Each row has the student who posted, the thread ID, the length of the post in terms of the number of characters, and the student being responded to.

This data set can be represented as a social network, of students responding to each other. In fact, only by doing so can you obtain an acceptable model predicting student grades (hint, hint).

In this assignment, you must compute the social network and use it to help predict student course grades.

You can build the social network using any software package you like (I like UCINet). You can do it by hand too, but trust me, that's just crazy. You can also build the regression algorithm in any fashion you like; cross-validation is not required for this assignment.

Please turn in:

- Evidence of model goodness for your model predicting student grades
- All data features created from the original data
- A document explaining how you completed the assignment

You will be graded on completeness and comprehensibility of your hand-in, whether you correctly and validly apply the method you choose to this data, and whether the methods you chose fit the requirements of this assignment.

**BONUS:** The student who succeeds in producing the regression model with the best correlation, gets the bonus. Note that you cannot receive the bonus unless you create a social network.