**Assignment 1, part 4, SCMA 632 (15 points)**

**Submit by 9 AM Wednesday, November 23, 2016 in an Excel or possibly a Word document E-mail to scma.stat@gmail.com with PROJECT in the subject line**

R.L. Andrews

As was announced in the course syllabus, a term project is required for this class. The purpose of this quiz part is for you to select a problem and an associated data set for your proposed project. I will give you feedback so that you will know if your proposal and data set will be acceptable for this project. If you want to do something other than analyze a set of data, then outline in a couple of pages what you propose to do. Otherwise follow the guidelines below.

I encourage you to pair up with another class member for your project. If you do work in pairs then both members of the group will receive the same grade. Put both names on the submission and copy the partner with your submission e-mail.

1. Give an overview of the problem, phenomenon or situation you want to investigate using multiple regression methods. You do not need to give a review of literature on the subject, just a clear description of the phenomenon that may be a population or process. Target your description for someone who is essentially a novice on the subject.
2. Describe the data set you will be analyzing **and clearly specify the observational unit that will provide the individual rows/cases/records for the data set**. Describe your process for selecting and obtaining your data set. Will you be collecting the data or are they from an existing data set? For existing sets, describe how the data were collected.
3. List the variables and describe how each is measured. The variables are to meet these guidelines unless you get permission from me.
* The response/dependent variable must be a quantitative variable. I want a variable that potentially has **at least 20 different possible values** for the dependent variable.
* There must be **at least 2 categorical variables**. One must have **3 or 4 categories** and the other must have **2 or 3 categories**.
* There must be **at least 4 quantitative or numerical variables** that each potentially has **at least 8 different possible values**.
* There should be **at least 8 total independent or predictor variables**. (The dummy variables created for a categorical variable do not count toward the 8, only the one categorical variable used to create the dummies.) Additional categorical variables can have 2 or more categories each. Additional quantitative variables can have as few as 5 possible values (5-point Likert scale variables can be included).
* I strongly prefer that you have at least 100 valid cases (n ≥ 100) but as an absolute minimum you must have at least 50(n ≥ 50).

**I will give you a detailed list of things that I want you do with your data set for the project.**