**Assignment 3 SCMA643** R.L. Andrews

**Submit by 9 AM Tuesday, April 11, 2017**

**E-mail and Excel file to scma.stat@gmail.com with “Assignment 3” in the subject line and copy your partner if you are doing this with a classmate.**

**This assignment is the first step for a class project.**

**You may do the project alone or you may partner with one other person and submit one report and make one presentation for the two of you, with the understanding that the credit for the project will be the same for both people.**

**1.** **Give an overview of the problem**, phenomenon or situation you want to investigate using multivariate methods. **You do not need to give a review of literature on the subject**, just a clear description for someone who is not an expert in the field.

**2.** D**escribe the data set you will be analyzing**. Tell how you are going to obtain the data. Will you be collecting the data yourself or is this an existing data set? Describe how the data are to be or have been collected. **List the variables (Must have at least 6 quantitative variables!) and describe how each is measured.**  The minimum number of observations/cases/rows is 100 with a total of at least 1,000 quantitative measurements. **To use anything less than n=100 or less than 1,000 quantitative measurements, you need to get prior permission from me.**

**3. Give an overview of things you want to learn from the data and the analyses you intend to perform.** For the project you are to use at least two multivariate methods, Factor Analysis is required and you must use at least one more. For each analysis procedure indicate the variables to be analyzed by that procedure. **One of these procedures is to be a factor analysis** to determine the number of dimensions being measured in a set of quantitative variables. **In addition** to factor analysis**, one or more of the methods below must be used**.

Acceptable Multivariate Methods & data set requirements:

* Factor Analysis (Required analysis for the project)
* MANOVA (This is like ANOVA except there are multiple measurements for each unit and one tests for a difference between groups for the group centroids based on the multiple measurements.)
* Discriminant Analysis (A dependent variable with 2 or more categories/groups & a set of predictors.) (Predictors are assumed to be continuous with equal covariance matrices for each dependent variable category/group.)
* Bivariate Logistic Regression (The regression method uses a set of predictors that can include dummy variables and a bivariate dependent variable that is categorical with exactly 2 groups or categories. I will not cover Multinomial Logistic that uses more than 2 groups.)
* Cluster Analysis (Elements are clustered into groups based on a measure of similarity. These groups are not previously defined.)
* Canonical Correlation (Analyzes the relationships between two sets of quantitative variables.)