Harman Socioeconomic Data -- MSDS 410 -- Logan Strouse

Bonus Assignment

What is the point of the analysis? What interpretations do you see?

This analysis was an exercise in looking at different factors and possible reductions through looking at the data. Two methods were used for the analysis and then compared. They were PROC FACTOR and PROC PRINCOMP. Through reading the pdf file and data set, I believe that the outcome of the analysis is dependent on the particular person doing the work. At first glance, it looks like they are analyzing correlations with the population variable as the response variable. The other variables are the predictors and the goal is to create a model to predict population growth or decline based on the predictor variables. I could also see this data being used to predict some of the current predictor variables and making population a predictor. Population could possibly be a good leading indicator for house value or employment. I see population and school and employment as having very high eigenvalue scores and because of this being able to explain 97 percent of the correlation. I would have need to remove population from the matrix if it truly is a response variable. By using the two factors, the researcher was able to calculate 4.67 of the total 5 of variance.