**Assignment 2, MGMT643, Spring 2018**

*R.L. Andrews*

The assignment will involve individual and group work and each group will submit an Excel file electronically with both the individual and group work for the entire group. You are to use the data found on my homepage http://www.people.vcu.edu/~randrews/ under the title **U.S. Census Bureau Data (Excel 2007)**. In this file use the tab labeled DataSet with the variable named "county" as the row/case identifier. The data set has 82 variables using the numbers 1 through 82 for the columns with numerical data. You are to perform your analyses for this assignment using the rows and columns your decided to retain and/or create in Assignment 1, when you started with 42 numerical variables in the columns numbered: 2, 4-6, 8, 11, 13-20, 29-31, 37-40, 45, 50, 52, 53, 55-58, 60, 61, 63-65, 72-74, 76-79, 81. (Note that you may create appropriate new variables that use one or more of these variables.) Hence, data for columns numbered 1, 3, 7, 9, 10, 12, 21-28, 32-36, 41-44, 46-49, 51, 54, 59, 62, 66-71, 75, 80 & 82 will not be analyzed.

Each team will use the data set with the rows and columns you used for your Assignment 1 factor analysis and,

1. use the factors in your final solution for Assignment 1 and calculate numerical values for each factor using the names you gave to each factor. Rather than having computational software calculate factor scores, calculate what the Hair text calls a “summated scale” which is a simple linear combination of the individual variables that are primary contributors to the factor for the summated scale values represent. **For each factor give the linear combination used to calculate the summated scale values.** **Individually each person will submit their suggestion for this simple linear function for each factor. Report these on a tab with the individual’s name on it for the final submission.** **The functions that the group selects to use will be placed on a separate tab.**
2. **report the mean and standard deviation for the group’s summated scale values for each factor along with the covariance and correlation matrices using all of your factors.**
3. using your summated scale values as variable values, **perform a cluster analysis** of the cases in these data. **Report how many clusters you decided to use for your final solution and tell why you chose this number. For each cluster provide a general description of characteristics of the type of localities that make up the cluster.**
4. **perform a cluster analysis of all of the original variables** you used for the factor analysis in Assignment 1. **Create the same number of clusters as the number of factors you retained in your Assignment 1 final solution**. **List the variables in each cluster. Compare the variables in each cluster with the variables in the group’s simple linear functions for each factor, reporting the similarities and differences between the two.**

Copy each group member with an e-mail to **scma.stat@gmail.com** attaching an Excel file. Each group member will have a tab(s) with their name on it and their responses to the individual part. There will also be tabs with the group results. **Deadline for Group submission is 9 AM Thursday, April 5.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Group 1** | **Group 2** | **Group 3** | **Group 4** | **Group 5** |
| Huang | Miller | Meenakshi-Sundaram | Howard | Dalal |
| Logesh | Shivaprasad | Mohammed Elnagar | Sandhu | Menon |
| Sharma | Uppalaguptam | Onwujekwe | Srikumar | Stanley-Brown |