**Assignment 1, 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 a factor analysis and consider using 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. You are to perform an exploratory factor analysis to identify common factors that are being manifested in the rows and variables you decide are appropriate to include.

Individually each person will screen the data to determine if there are any rows in the original data set you recommend removing before beginning factor analysis. Clearly list the rows to be removed and the reason for removing each one. You may choose to give a list of reasons and then give all of the rows that are deleted because of each reason. Then you are to indicate if there are any variables that you think should be excluded from the factor analysis giving a reason for each variable that you recommend excluding and you may choose to create one or more new variables using values from one or more original variables. Each person submits this on one or more sheets before beginning group work. **Individual work** must be submitted to all of the team by **5 PM Monday, March 12**.

Each team will:

1. put each team member’s name on a tab or tabs with her/his individual work that lists recommended rows/cases and variables for exclusion as well as reasons for exclusion and lists any created variables.
2. include a text box on a tab describing the process the team used to arrive at the final factor analysis solution you are submitting as your final answer.
3. report the team’s list of rows and variables that were excluded for creating the final factor analysis solution, as well as reasons for each exclusion and any created variables. (It may be that reasons for excluding a row or column will arise while doing your factor analysis using a principal components solution.)

4. on a separate tab or tabs report (Do not put too much information on a single tab.)

* whether the covariance or correlation matrix was analyzed and why,
* the number of components retained for the solution and why you chose this number,
* the portion of the total variance explained (total variance of the retained variables),
* the communalities for each included variable,
* whether you chose an unrotated or rotated solution for your final solution and why,
* for a rotated solution indicate the rotation you used and why,
* the correlation between each variable and factor for your final solution, and
* name each factor in your final solution.

Each group will send their final submission to me in an Excel file via e-mail to **scma.stat@gmail.com**. Group members 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, March 29.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **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 |