STAT-S631 – Final Exam – Take-home

Due December 14, 2017

Intructions:

- Please start your document with the following statement: "On my honor, I have not had any form of communication about this exam with any other individual (including other students, teaching assistants, instructors, etc.)" and sign by hand (if work was done partially or completely by hand and scanned) or typing your name (if work was done all typed).
- Your answers should be presented in the **first 2 pages** of your submission paper. If your arguments are appropriate and complete, you will receive full credit based on those answers entirely. If on the other hand, your arguments are incomplete or incorrect, I'll inspect the additional syntax/output after those two pages to determine if you can receive partial credit. In any case, you are strongly encouraged to have succinct answers. Lengthy answers are neither required nor desired (and they are prone to contain incorrect statements that would be penalized).
- Please note that you may or may not arrive to same model than your classmates and/or instructor, but if your explanations are appropriate and the used of the methods learned in class is correct (alongside the syntax/output your provide), you may still receive full credit. By contrast, incorrect and/or incomplete explanations will be penalized, even if the chosen models are appropriate. Moreover, include appropriate interpretations, if required.
- Supporting material (R syntax used, relevant output, graphs, etc.) is also required to get full credit and should be added/attached after your answers pages (unless a question explicitly requires otherwise). There is no limit of space for supporting material.
- Submit your solutions to Canvas as a single PDF file by Thursday, December 7th. Do not wait until the very last minute since late submissions will not be graded.

Questions

Use the data `Angell.txt` and description file `Angell.pdf` used previously in Assignment 6. Use `moralIntegration` as the response and all the other variables as predictors.

1. Use methods and techniques learned in class to determine
   - which predictors to include,
   - if you should include polynomials and/or interactions,
   - if weights or other adjustemnts should be used for the variance,
   - if transformations to predictors and/or response should be considered.
2. Based on the model you have obtained based on your work in part 1,
   - interpret at least two coefficient estimates,
   - perform and analysis of residuals, determine if the assumptions made about the mean and variances functions are appropriate,
   - make any appropriate changes to your model.
3. With the resulting model from part 3, perform an influence analysis, in particular
   - using the Cook's distance, determine which observations are the most influential (select up to four),
   - determine which observations, if any, could be considered outliers,
   - compare the coefficient estimates obtained with and without the selected influential observations and comment your results.