



- 3.) Transform/derive new variables
  - a. Transform your raw data into variable types useful for modeling (e.g. bin numeric numbers, create binary dummy variables, standardize giving totals, etc.)
  
- 4.) Explore the data/data validation
  - a. Go through each variable and compare your expectations with the actual data
  - b. Make initial observations of each variable and spot potential problems
  - c. Think about what variables are still missing
  - d. Identify the things that need to be corrected, but resist temptation to make corrections as you go
  - e. Document every step you take in this process
  
- 5.) Fix problems with the data
  - a. Make changes to your data based on your observations, make notes of each change and the methodology behind it
  
- 6.) Analysis/modeling