Data Management in STATA (Lecture Notes for Tues, October 2\textsuperscript{nd} 2012)
(All course programs and datasets can be found in the Resources section of the website)

**General Information about data structure and variables:**

i. Basics of flatfile database structure in STATA
   a. Difference between STATA and EXCEL, etc

ii. Basic Data Setup (examine the GSS96\_small.dta file).
   a. Unique ID for each case
   b. Numeric responses whenever possible
      i. Includes categorical, scale, count, etc
   c. Variable names: the short, usually descriptive, name for each variable. Descriptions are always better: e.g., ‘age’ instead of ‘variable4’.
   d. Variable label: labels that you apply to the shorter variable name. Such as, “rincome” is a variable, but the label may be “Respondent’s average annual income”.
   e. Value labels: labels that you apply to individual values such as 1=male, 2=female.
   f. Missing Codes: how we deal with no response, as well as “don’t know”, “refused”, etc.

**Working with STATA:**

1) Working with the Stata windows
   a. Command
   b. Output
   c. Variables
   d. Review

2) Opening STATA program and data files
   a. Do files (.do)
   b. Data files (.dta)
      i. View your datafile (browse)
   c. Relationship between .do and .dta files
      i. Writing programs on the fly
         1. Logging your work ‘log using filename.log’ (without quotes)
         2. Stop logging by using ‘log close’.
      ii. Writing programs with comments: *, /* and */

3) Working with a simple program (simple.do and GSS96\_small.dta)
   a. In-class walkthrough of simple.do
      i. basics of simple variable creation and recoding.