**Homework Formatting Longitudinal Data**

The data set fram40 is in the fram directory.

1. Use a PROC CONTENTS step to find the names of the variables on the dataset fram40 that contain determinations of serum cholesterol. All variables have a character prefix followed by a number. The number represents the sequence of the determinations: 1 means the first determination, 2 means the second, and so on.
2. Use a PROC SURVEYSELECT step to select a simple random sample of twelve observations that contains only the values of the first six determinations of serum cholesterol.
   1. Use the seed, 54321 for the random number generator within PROC SURVEYSELECT .
   2. Output the random sample to a data set work.ctmp.
3. Use a DATA step with work.ctmp to create a new version of work.ctmp
   1. Assign a sequential ID variable (i.e. numbered 1 to 12)
   2. Rename the variables containing the cholesterol determinations as chol1, chol2, chol3, chol4, chol5, and chol6.
4. Do a random sort of the data.
   1. Use a data step to create a new version of ctmp by setting ctmp.
      1. Create a variable ranx, that contains a random uniform (0,1) observation.
      2. Use the seed, 54321 in a call to the streaminit function to initialize the random number generator
      3. Sort the new version of the dataset ctmp by sorting the existing version on the variable ranx.
5. Use a PROC CORR step to examine the correlation structure for the repeated measurements of cholesterol.
6. Use a DATA step to create six new data sets named work.chol1, work.chol2, work.chol3, work.chol4, work.chol5, and work.chol6. Each observation on each dataset should contain only the variables id and chol***i*** where the variable chol***i*** is the value of cholesterol at that determination (1, 2,…,6).
7. Use a DATA step to create a data set longchol that is in long format. I.e. each observation on the data set contains three variables, id, exam, and chol.
   1. Exam is the determination number for the value (1, 2, …, 6)
   2. Chol is the determination of serum cholesterol for that id and examination combination.
8. Use a PROC SGPLOT step with the dataset longchol to create a spaghetti plot that displays the longitudinal data for the 12 individuals.
   * 1. Exam is on the x axis
     2. Cholesterol is on the y axis
     3. A separate series plot for each individual (id) is plotted
     4. Plot symbols are included on the plot for each individual determination.