Assignment #2: Due September 26, 2019, 5pm

INSTRUCTIONS: Please upload **ONLY** your SAS program (saved as LASTNAME_FIRSTNAME_HW2) to Canvas and a format file (saved as LASTNAME_FIRSTNAME_FORMAT) if you choose to create any formats (they should not be in the main program). The only program that will be run is your main program, so make sure to use %include to call the formats from the main program. DO NOT create any permanent formats or data sets. Make sure to use the %let statement to create a path as was done in class and use that path to locate your data sets. This will be the only modification that I will make to your code (replacing your path with mine). Please make sure to comment your code and title any output (even though you won't submit the output, I will be looking at it to make sure that your code ran correctly).

One hundred college students were recruited to join a quality of life (QOL) study. They were given five assessments using different subscales. If an individual missed the original measurement time for a subscale, he/she was offered a makeup assessment.

Several authors have devised ways in which to combine the assessments to come up with an overall QOL score:

Author A: The average of the first four scores is worth 70% of the score, while the last score is worth 30% of the score. In order to get a score, you must have completed all 5 assessments (i.e. if someone misses an assessment, they get a missing value for the summary score).

Author B: The average of assessment 1, 2 and 3 counts for 50% of the score, while the average of assessments 4 and 5 count for 50% of the score. If only one assessment is missing, the assessment is ignored and it does not impact the overall assessment. If more than one assessment is missing, then the individual gets a missing value for the summary score.

Author C: If a student is from out-of-state, the score is the half of the average of the completed assessments (regardless of how many are missing) plus 25. If the student is from in-state, the student gets a 50 for any missing assessments and the total score is the average of all of the assessments.

Please create a Data Set (called FINALDATA) that contains the following variables: PATNO, Year, Residency, Major, Author A score, Author B score, Author C score, and an indicator of which author's score was the highest (A, B, or C). All scores in this data set should be rounded to 2 decimal places.

Deliverables:

A sorted FINALDATA (by PATNO) and the first 15 observations from FINALDATA set printed

There are three files for this assignment and they contain the following information:

1. *Demographics.csv*:
    a. PATNO (participant number)
    b. Year:1 = Freshman, 2= Sophomore, 3=Junior, 4=Senior, 5=Senior Plus
    c. Gender: 1=Male, 2=Female
    d. Residency: 1=In State, 2=Out of State
    e. Major: 1=Chemistry, 2=Biology, 3=Mathematics, 4=Physics, 5=Psychology, 6=Other
2. *Assessment_original.csv*
    a. PATNO
    b. QOL
    c. Scale
3. *Assessment_makeup.csv*
    a. PATNO
    b. QOL

c. Scale

The *Assessment_original.csv* contains the QOL score from the original assessment (one line per test), while *Assessment_makeup.csv* contains the QOL score for individuals who missed the first offering of the assessment and took the make-up assessment.