Statistics 6430 (SAS)
Final project
Due: Thursday, July 20

Background:
A university runs a consulting center called the Analytical Consulting Lab for small- to moderate-sized consulting projects within the university. There are three consultants who meet with clients, offer advice, and possibly help analyze data, help plan and implement a research study, or help develop resources. The last names of these three consultants are Smith, Jones, and Brown. There is also a Lab Coordinator (LC) who assigns incoming projects to the consultants and tracks the time the consultants spend on each project.

Each work day, each consultant tracks the amount of time spent working on each project and submits this information to the LC on a form (one per project). On this form, the consultants specify the date, the total number of hours (rounded to 15 minutes) spent on the project during the day, the project number, the project stage, and whether the project has been completed. The project stages are:

1. Initial consultation with the client
2. Planning stage, which may involve meetings with client and other relevant parties
3. Implementation and analysis stage
4. Interpretation and reporting stage, which may be the writing of a final report or a results meeting with the client
5. Optional follow-up meeting with the client (at most one)

A project is considered complete at the end of the fourth stage.

If the consultants work on multiple stages of a single project during the day, when filling out the form for the LC, the consultants are asked to specify the stage on which they spent the majority of time that day. Thus, a consultant will only record one stage per day per project. Additionally, the stages must be completed consecutively. Although, not all five stages are required for every project, so some may be skipped.

The LC has recently been experimenting with project classifications, which designate an entire project as one of four types:

1. Study coordination: the most extensive type of project where the consultant essentially coordinates an entire experiment, survey, or research study
2. Information technology and web development: where the consultant directs the development of analytic software
3. Data analysis: where the consultant performs basic data analysis of results of a completed research study

4. Advising: the smallest-scale projects where the consultant may only provide their expert advice on a topic

In terms of record-keeping, the LC began a master file of the forms starting at the beginning of the year, but it was sparsely updated after September 1st. Many of the forms after September 1st are recorded in a different file. Additionally, when consultants alert the LC to corrections to the hours and stage, they are kept in a separate file until they have been corrected in the master file. The forms the consultants submit each day do not request their name, thus for the projects initiated after September 1st and not listed in the master file, the LC needs the file of project assignments made. Lastly, since the project classifications are a developing system, these types have been recorded in a separate file.

The LC from this consulting center has hired your team to streamline the project reporting from the year and to produce specific reports that summarize various characteristics of the center's work over the year. To this end the LC has provided the following five datasets:

1. *Master.csv*: The LC's master file, which contains the assigned consultant, the project number, the date, the number of hours worked, the stage, and whether the project is completed (1 designates completed)

2. *NewForms.csv*: The LC's file of forms that have not been added to the master file, which contains the project number, the date, the hours worked, the stage, and whether the project is completed (1 designates completed)

3. *Corrections.csv*: The LC's file that tracks corrections to be made to the master file, which contains the project number, the date, the corrected hours and the corrected stage (only nonmissing values are corrections)

4. *Assignments.csv*: The LC's list for which consultant is assigned to which project, which contains the consultant and the project number

5. *ProjClass.csv*: The LC's list of project classifications, which contains the type of project and the project number

Objective 1:
Create a new master file that is completely up to date, error-free, and includes the project classification types.

- Incorrect values should not be kept in the new master file.

- Whenever a correction is made, it should be noted by a 'Yes' in an appropriately named additional variable.

- The new master file should be written to a .csv file named *NewMaster.csv* and should contain appropriate headers. (You do not need to use SAS to supply the headers.)

- Any formatting should be done in SAS.

- The new master file should also be saved in a permanent SAS dataset named *NewMaster*.

Objective 2:
Starting with the new master file, generate a report of ongoing projects as of the last entry date (November 4th). Ongoing projects are those that have not yet been completed. This report should show only project numbers.

Objective 3:
Starting with the new master file, generate a report of the consulting activity of each consultant on each project as of the last entry date (November 4th). There should be three separate reports, each showing the project numbers on which the consultant has worked. For each project the following information should be given: the total number of hours worked, the project type, whether the project has been completed, the start date of the project, and the end date of the project (determined by the last form submitted for the project).

Objective 4:
Starting with the new master file, generate two or more additional reports that you think offer useful information to the LC. These reports should summarize the data in some meaningful way. At least one report should make use of a PROC procedure other than PROC PRINT. Graphical reports are acceptable, but not required.

Notes for the reports:

- Each report should be given an appropriate title.
- Each report should look professional (ie. more complete and descriptive headers are used for columns instead of the variable names).