

Helpful Hints on Record Keeping for People Working on Projects¹

A challenge in working on long-term projects is keeping track of programs, data files, variables, etc. Even in small projects, such as a single paper, many files can be produced. In this document, I hope to provide suggestions that will make it easier to keep track of your work, and to make it easier for others working on the project to find what you have done. It takes a little extra time to document your work, but in the long run, it is worth it.

1. Decide on a short name (one word) for the project or paper, and use this in the naming scheme of files and directories. For example, if the paper is going to be about immigrant women, you might choose the handle of "immw". Store all of the files pertaining to that paper in that directory and subdirectories. You may want to create subdirectories for systems files, for revisions, or for different aspects of the paper. For example, if you are using 1980 and 1990 data, you may want to have separate subdirectories for those years.
2. Create a file in the directory called "anote" that has information about the work that will be found in the directory. Briefly describe the goals of the project and the data that will be used. Update this note when necessary, putting the most recent information at the top, with dates. When analysis is finished for specific tables, put that information in the file called "anote". For example, "Output for table 1 was created by immwt1.sas. Output for table 2 was created by immwt2.sas. Output for footnote 1 was created by immwf1.sas. ..." Also, include information about where the text and tables for the paper are located - on whose PC and in what directory.
3. Create a directory where you store your programs for the project or paper. Use the short word you decided. If you will also do PC work relating to that project, create a directory on the PC for tables, figures, text, etc., for the paper. Create a file in the directory called "anote" that contains the same information described above. Also include information about the location of the program and listing files that were used to create the tables and other information for the project.
4. **VERY IMPORTANT.** It is very important to be able to determine how systems files and the variables in them were created. I believe that the best way to assure that you (and those that follow you) will be able to do this is by using the following strategy. Every systems file should have a program file with the same first name that corresponds to it. For example, if you create a SAS systems file called "immw90.ssd01", it should be created by a program named "immw90.sas". Similarly, an SPSS systems file called "immw90.spssfile" should be created by "immw90.spss". It is also a good idea put information in the file "anote" (described above) about systems files you create, the dates, etc.
5. To help you and others easily identify a file type, always name SAS programs with the extension of ".sas" and SPSS program files with the extension of ".spss" (or choose some other naming convention - just be consistent). Always name SPSS systems files with the extension of ".spssfile". With SAS systems files, there is not a choice - on the UNIX, the extension will be ".ssd01" and ".ssd" on a PC.
6. Document your program files with comments and titles. The first line of the program should always be a document line that tells the name of the program. For example, the first line of the program file "immw1.sas" would be "* immw1.sas, your initials or name, the date.". The second line should be a title that includes the full path name of the program (for example "/disk1/immw/y90/immw1.sas"). Document the date you created the program and your name (if you work or will work with others). If the program reads data, document the sources of the data carefully. Be generous with documentation throughout the program, especially with programming that may otherwise be confusing. And, create variable and value labels.
7. When printing tables and figures, include information that will make it easy to revise the table when necessary - the full path name of the file (for example, c:/immw/table1, from program /disk1/immw/table1.sas, and the date. Keep a copy of the tables with this documentation with the paper or project. You will probably want to print a copy with this information omitted to the journal.
8. Create an index of your directories and projects. For example, you would have an entry for the project on immigrant women. You would include information about the subject and purpose of the paper/project, the data being used, the directories where files are stored (PC and UNIX), the title of the paper, status of the project (for example, paper under review at AJS, or revises and resubmitted for ASR in 4/99), and any other information that is important.

S:\CSDA-SHR\HELP\records.wpd. July 8, 1999. jpg

¹Prepared by Patty Glynn, Center for Social and Demographic Analysis, University at Albany. July 16, 1999.