

Block 1: From questionnaire to SPSS saved file**1.3.3.9 Some general advice on file building in SPSS**

[23 November 2010]

Previous sessions: 1.3.3.1 to 1.3.3.8

Very small surveys or data sets can be generated in a single run, but this is not advisable for beginners, or even for experienced researchers, especially if you want to keep track of your work. On very large data sets, several runs may be needed before the final version is achieved.

Although it is possible to work directly in the SPSS Data Editor or using the drop-down menus, it can be quite fiddly. You are best advised to use SPSS in syntax mode to assemble your SPSS saved files in stages, starting with **DATA LIST**, then follow with **MISSING VALUES**, **VARIABLE LABELS** and **VALUE LABELS** in that order. Permanent data transformations and generation of derived variables can be done later. This may sound cumbersome and tiresome, but years of experience have shown it to be more effective and much less frustrating in the long run. Many experienced researchers also copy and paste both to and from SPSS using Word and Excel to write, store or manipulate their data and syntax outside SPSS.

Files need to have the correct extension if SPSS is to read or write them properly (**.sps** for syntax files, **.sav** for data editors, **.spo** for output files and **.por** for portable files) This is usually done automatically by SPSS using point-and-click, one of the few occasions on which this method triumphs over syntax mode. Initial raw data files can use **.txt** (or **.dat**, which may appear as WordPerfect or Windows Movie icons on screen, but SPSS can still read them). You can also input raw data directly into the **Data Editor**, although I personally wouldn't recommend this except for lone researchers entering their own data and provided there weren't a large number of cases and/or variables (for which Excel may be preferable). There's no harm in beginners playing around with a few cases and variables, but only as an exercise in editing.

Although this is only a preliminary exercise, you should **get into the habit of saving your work** every so often. This applies to data sets, syntax files and your working SPSS **Data Editor**. However, there is no point in saving output files resulting from runs with lots of errors in them. SPSS is so fast on modern PC's that, when errors occur (and there will be some, even if you're an experienced user like me!) as soon as you have found and corrected the error(s) **it is better to close the current output file** and then click on **No** when asked if you want to save it. Otherwise you'll finish up with dozens of error messages in your output file. In the old days, this used to be reams of wide printout, so at least we're saving a few trees!

Once you have an error-free syntax file, run it again to get the correct results on a new output file. SPSS generates its own names in sequence for syntax, output and saved files, but you would be well advised to change these when you save them with **File > Save As ...** since you won't necessarily remember what you did in **syntax109.sps**!

Recommended practice would be to create a new folder for each survey and keep all the related files for each survey together in the same folder. If the number of files gets too large, you can always create sub-folders.

Since these were our first exercises we called the files **myclass1** and saved them in a folder called **myclass**. First we created a new folder **myclass** in our working area and then saved the SPSS data editor as **myclass1.sav** and the SPSS syntax file as **myclass1.sps**. The filename displayed at top left of the Data Editor changed from **Untitled** to ***myclass1** and the filename in the syntax file changed from ***syntax1.sps** to ***myclass1.sps**.

Some of you will want to save the (correct) output file as well. Just save it as **myclass1.spo**

From now on, if you attempt to close a syntax or output file SPSS will ask you if you want to save the changes to a *.sps or *.spo file. If you try to close the data editor SPSS will warn you that this will exit SPSS and ask if you really mean it. If you say Yes it will ask you if you want to save the data editor as a *.sav file.

If you only write the first part of the filename, SPSS will add the extensions automatically. Once SPSS has closed down, all you need to do to carry on working next time is to navigate to folder **myclass** and double-click the saved file **myclass1.sav**. The next set of SPSS runs can then be called **myclass2.*** and so on. This way all files relating to a particular run will have the same file name part and you will be able to keep track of your work, or at least know which files go together.

I also use file names such as **freq.*** and **tab.*** etc. (to indicate frequency and tabulation runs) or names indicating the substantive content (eg **protest.*** or **anomy.***).

Until you are absolutely certain of the accuracy and completeness of any subsequent files, and in case of computer crashes, power failures or other incidents such as accidental deletion, you should **always keep the earlier and current editions** and also **copy them** to a safe external medium such as a rewriteable CD, an external hard drive or a remote server.

You have been warned! Don't find out the hard way!

End of tutorial

Next tutorial: 1.4.1 Labelling your variables in SPSS

[\[Back to Block 1 menu\]](#)