



Shafi Consultancy Ltd.

Quick Guide to Good Programming Practice

Getting started

1. Create an initialisation program such as **autoexec.sas** where **all libnames** and standard options, including **NOFMTERR** and **MSGLEVEL=I** are defined. If the initialisation program is not **autoexec.sas**, then include this at the top of every program. No libnames should be defined in individual programs.
2. Create a format program such as **Studyfmt.sas** where all study formats are created. This will be called from the initialisation program.
3. Create **Abbreviations** for header, comment and section boxes for consistency.

Programming Style

4. Use a **standard program header**.
5. Comment programs in the form ***** Comments *****; and not **/* comments */**. Insert comments in boxes, and have a different type of box to separate the program into sections.
6. **Avoid inserting comments inside DATA steps** or PROC SQL steps, put them in a box above the data step. It makes the program easier to read.
7. Set up program into **INPUT, PROCESS** and **OUTPUT** sections.
8. Read all external datasets **only once** in at the top of the program.
9. Use **"KEEP="** option and specify variable names both when reading in external datasets and when creating permanent datasets.
10. Only write **ONE** SAS statement per line.
11. Use **different meaningful names** for each temporary dataset.
12. Use indentation. **Indent 2 characters** (or one tab), but remain consistent throughout the program.

Checking

13. If the LOG window contains **Cartesian product** after a PROC SQL statement, check the dataset produced very carefully, as it will most likely have more observations than what is expected, and therefore be incorrect.
14. Check the LOG for **ERROR, WARNING, uninitialized** and **repeats of BY values**. The "Repeats of BY values" only appears as a NOTE, but it should be treated as an ERROR and the code must be changed.

www.shaficonsultancy.com