

10 Things Experienced SAS® Programmers Don't Know – But Should

Doug Zirbel, Pinnacle Solutions Inc., Indianapolis, IN

ABSTRACT

Many SAS programmers have years of experience, but sometimes within only a narrow set of procedures and techniques. Numerous SAS professionals were asked for their views on an essential skill set, an industry-specific skill set, and the best means to learn it. These are their recommendations. The list covers SAS Version 6 as well as Version 8. Base SAS products listed as a part of this are ODS, Proc SQL dictionary tables, data step, macro language, and Proc Report.

INTRODUCTION

What makes a well-rounded, experienced SAS programmer?

What are 10 things that experienced SAS programmers don't know, but should?

What are the industry-specific SAS skill sets? And finally,

What's the best way to learn them?

The first two questions were posed to the SAS-L discussion group beginning in 2000, then reintroduced again in 2001, and updated further with the additional questions in 2002. Some of the answers were technically specific, for example, "the Scan function", or "Proc Report". Many others suggested foundational theory, such as System Life Cycle Development methodology. Some suggested that an expert programming skill set is specific to the industry you happen work in. Still others waxed philosophical:

"Programming and musical composition have much in common", or "the 'well-rounded SAS programmers' that I know got that way either by eating too much or by first learning how to do a few things well and then studying alternative methods". – Sig Hemansen

SAS has expanded so much, e.g. WebAf, JavaServer Pages, Data Mining, Data Warehousing, that I have little fear of falling into a specialization hole. Seventeen years ago, I knew what I didn't know. Now I don't know, what I don't know. – Jim Loughlin

METHOD AND FINDINGS

The respondents were asked to rate "What are 10 things that experienced SAS programmers don't know, but should?" items from a list of categories on a scale of 10, with 10 being the highest score and 1 the lowest. These can be found in Appendix A.

The question, *What's the best way to learn these skills?*, was sent by email to those who had already responded. Later, the question was posted to the SAS-L discussion group. The responses often provided detailed, thoughtful suggestions, and generally fell into one of five classes: Specific techniques, Books – Manuals, Mentors and Other Programmers' Code, Classes and SUGI, and Experience – Experiment – Practice. This unranked collection is found in Appendix B.

SAS-L participants were later asked to rank several "best ways" categories, and these rankings are shown in Appendix C. Interestingly, the top two items were "curiosity/motivation" and "Applying code examples presented in Sugi/SUG papers or on SAS-L to your own work".

The industry-specific skill sets listed in Appendix D summarize suggestions offered by SAS-L discussion group as of August 2002.

THE ORIGINAL LIST WHICH RESPONDENTS RANKED

- Macros - basic knowledge of macro language
- Macros – knowing how to create and use macro var "arrays"
- Macros - knowing how to write macros that write SAS programs
- SCL - basic knowledge of SAS-AF SCL
- SCL - SCL Functions usable in e.g., Base SAS data steps
- SQL Dictionary tables - basics of retrieving SAS metadata
- SAS IntrNet - basic knowledge of
- SAS Help - How to get it
- Data Step Vector - how it works
- Data step - basic knowledge of
- Data step reports - "Data _null_" report writing
- Display Manager - basic knowledge of
- Debugging Techniques - DEBUG facility and other techniques
- Data file manipulation - INFILE, UPDATE, SET intricacies etc
- ODS - capabilities and more broadly, SAS8 improvements
- Your Operating System native (and other) languages - REXX, unix scripts,
- VBScript, Syncsort
- Proc Format
- Proc SQL
- Proc Summary/Means
- Proc Freq
- Proc Tabulate
- Proc Report
- Proc Transpose
- Scan()
- Index()
- Substr()
- Efficiency
- Arrays / matrices
- Web technology - HTML, JavaScript, SAS IntrNet, etc
- Configuration Management
- Relational Database Design
- Structured programming concepts
- System Development Life Cycle processes
- Understanding how programs translate into machine operations

CONCLUSION

This paper lists the results of surveys of experienced SAS programmers regarding a) the most important skills needed by experienced SAS programmers, b) views on the best way to learn these skills, and c) a list of some of these skills categorized by industry.

REFERENCES

Carpenter, Art (1998), *Carpenter's Complete Guide To the SAS Macro Language*, Cary, NC: SAS Institute Inc.

Davis, Michael L., "SCL for the Rest of Us: Nonvisual Uses of Screen Control Language," *Observations: The Technical Journal for SAS Software Users*, www.sas.com/service/library/periodicals/obs/obswww05/

Delwiche, Lora D., and Slaughter, Susan J. (1998), *The Little SAS Book: A Primer, Second Edition*, Cary, NC: SAS Institute Inc.

Dilorio, Frank and Michal, Nancy (1996), "Data About Data: An Introduction To Dictionary Tables", *Proceedings of the Twenty-first Annual SAS Users Group International*, 21, 213-219.

McConnell, Steve, *Code Complete* (1993), Microsoft Press.

Miron, Thomas (1996), "The Secret Life of the Data Step", *Proceedings of the Twenty-first Annual SAS Users Group International*, 21, 171-177.

SAS Institute Inc. (2000), *SAS Macro Language Reference, Version 8*, Cary NC: SAS Institute Inc.

SAS Institute Inc. (1989), *SAS Language and Procedures: Usage, First Edition*, Cary NC: SAS Institute Inc.

ACKNOWLEDGMENTS

Thanks to:

April Sanders
Amy Swinford
Allan Page
Andrea Wainwright
Andy Kuligowski
Anthony Gargon
Arthur Tabachneck
Bill LeBlanc
Brad Goldman
Casey Pierce
Charles S. Patridge
Christina Gullion
Curt Seeliger
Curtis Smith
David Cassell
Diana Kornbrot
Dianne Rhodes
DJ Penix
Douglas Dame
Dwight Eggers
Ed Heaton
F.J. Kelley
Frank Dilorio
Frank Liu
Frank Schiffel
Gary McQuown
Greg Woolridge
Ian Whitlock
Jack Hamilton
James Green
Jim Loughlin
John Gerstle
John LaBore
John Stimpson

Karsten M. Self
Kim LeBouton
Lee Medoff
Mark Moran
Maurice Sebastian
Michael L. Davis
Michael Raithe
Michael S Hines
Michael Thomas
Mike Harris
Mike Rhoads
Nancy Brucken
Nancy Michal
Paul Dorfman
Pedro Gomez
Peter Flom
Quentin McMullen
Richard DeVenezia
Richard Palmer
Ronald J. Fehd
Scott Miller
Sig Hermansen
Steve Albert
Steve First
Susan Fehrer
Thomas Hauge
Thomas Miron
Tom Winn
Wei Cheng
William W. Viergever

CONTACT INFORMATION

Your comments and questions are valued and encouraged.
Contact the author at:

Doug Zirbel
Pinnacle Solutions, Inc.
One North Meridian Street, Suite 400
Indianapolis, IN 46204
Work Phone: (317) 423-0128
Fax: (317) 423-9143
Email: doug.zirbel@psiconsultants.com
Web: <http://www.psiconsultants.com>

This paper can also be found at
<http://www.psiconsultants.com/about/publications.htm>.

APPENDIX A

10 Things Experienced SAS Programmers Don't Know – But Should

Topic	Score	Comments
THE TOP 10		
Macros - basic knowledge of macro language	63	Macros are crucial. Next in priority would be Screen Control Language. – Jim Loughlin
Data step - basic knowledge of	60	
Data file manipulation - INFILE, UPDATE, SET intricacies etc	59	I'd ask/want to see that they REALLY know to control a merge so that they know that exactly what's supposed to happen does happen. – Douglas Dame
ODS - capabilities and more broadly, SAS8 improvements	53	
Efficiency	49	
Data Step Vector - how it works	47	
Macros- knowing how to create and use macro var "arrays"	46	
System Development Life Cycle processes	44	...many SAS programmers are clueless about the application of a documented software development life cycle to SAS programming. Without that, it is impossible to build valid SAS systems. – Mike Harris
SAS Help - How to get it	40	
Proc Report	34	
THE NEXT 12 (RECEIVED AT LEAST 4 VOTES)		
SQL Dictionary tables - basics of retrieving SAS metadata	31	
Proc Summary/Means	28	
Debugging Techniques - DEBUG facility and other techniques	27	
Proc SQL	27	Not merely the SAS proc, but fundamentals of database design and the query language. Joe Celko is the canonical reference. – Karsten Self
Structured programming concepts	24	Programming and musical composition have much in common. The programmer and the composer use symbols that humans can write and understand, yet also correspond to signals that precisely control devices... good programmers check syntax carefully and test building blocks of larger programs. – Sig Hermansen
Proc Format	18	
Relational Database Design	18	I think that data modeling is one of the most basic skills for a good SAS programmer. This is one of the mature aspects of computer science, going back to Date's work at IBM on

		relation databases. It is important to have good intuition about such concepts as primary keys, foreign keys, and data normalization to be able to use SAS data structures wisely. – Dwight Eggers
Web technology - HTML, JavaScript, SAS IntrNet, etc	17	
Your Operating System native (and other) languages - REXX, unix scripts, VBScript, Syncsort	17	Know the features (and limitations) of your OS. My own experience is that I wasn't a SAS programmer, I was a SAS/Unix programmer. – Karsten Self
Proc Transpose	16	I did eventually (and haphazardly) come to discover that PROC TRANSPOSE would be ideally suited to [...a certain] task. – Lee Medoff
Data step reports - "Data _null_" report writing	14	
SCL - SCL Functions usable in e.g., Base SAS data steps	14	

APPENDIX B

What's the best way to learn them (unranked)

Specific techniques and "other"

In one sense, each of these different 'subject areas' is the proof of that rule of "know many languages". which might be reframed as: know many -perspectives-. certainly I would recommend at least two of each of: database spreadsheet word-processing number-crunching other	Ron Fehd
So you want to become a SAS Programmer (http://groups.google.com/groups?hl=en&selm=88f2gg%24r60%241%40nnrp1.deja.com)	Charles Patridge
own your own niche	Ron Fehd
SAS Observations was a wonderful tool for moving my skills forward. I don't think that we've had anything of that caliber from SAS since its demise.	Kim LeBouton
A grounding in the CS basics	Karsten Self
...keeping an eye on www.sas.com (especially the "communities" and the Service and Support home page)	Mike Rhoads
Heavily utilize the tech support section on http://www.sas.com	Jim Loughlin
I think the first most important thing is to learn the philosophy with which a language is constructed. Without that it is very hard to locate answers to anything. That kind of knowledge comes by asking lots of questions and searching for answers.	Ian Whitlock
The [OnlineDoc] CD for 8.2.	Frank Schiffel
I have become a strong believer in software engineering principles (http://www.sei.cmu.edu/cmm/cmms/cmms.html)	Dwight Eggers
Most SAS programmers really do need to know their data.	Dianne Rhodes
I use post-it tabs to mark important examples that I can refer back to often.	Andrea Wainwright

Books, Manuals

When I first got my own set of SAS manuals, I sat down and read about all the string-processing functions.	Ron Fehd
Books and manuals help... sometimes half the battle is just finding out that a particular method or technique is available for doing something.	Nancy Brucken
My recommended text is Steve McConnell, _Code Complete_, Microsoft Press, 1993. Everything in the literature which isn't in the text itself is referenced in the bibliography. SAS specific: Rick Aster's books, which are the only ones I am aware of which treat SAS as a programming language, rather than simply an analytic toolkit.	Karsten Self
SAS Language Reference...Aster and Seidman, Professional SAS Programming Secrets. A good beginners book I've used to teach intro. SAS is DiIorio, SAS Applications Programming.	
I'm a big fan of hardcopy documentation and BBU. I often will look at them while watching TV--exciting gal I know!!! Many times I've had the situation that I have a problem that needs a solution, and I recall that I saw something earlier from my perusal of documentation.	Kim LeBouton

One of the best techniques for me has been picking things up through the manuals. I don't try to memorize syntax, but rather to familiarize myself with capabilities ("Aha -- you can wrap long text fields with PROC REPORT!"). That way, when a relevant problem comes up later, I have a clue about where to look. I always try to do this with the Changes and Enhancements documentation when a new version comes out	Mike Rhoads
The next-easiest way would be to purchase one or more of the excellent "Books by Users"-books (depending on the subject, some of the authors that come to mind include: Delwiche & Slaughter, Schlotzhauer & Littell, Cody & Smith, Cody & Pass, Aster, DiIorio, Carpenter Raithel, Khattree & Naik. etc.), and to work through the examples they include.	Tom Winn
Sounds trite, but read the manuals. Review every 6 months. Concepts that made no sense 6 months ago, now become clear.	Jim Loughlin
Constantly review the chapter on SAS Functions. Also check out the SAS macros that come with software. I hate to tell you how many times I was patting myself on the back about some clever code I wrote for which there was already a SAS function or SAS supplied macro which did the same thing.	Jim Loughlin
At the beginning one has to do some reading...	Ian Whitlock
The first time I exercise a new skill I have learned, I make sure that I have both the users manual and the class notes [available from SAS publications] within easy reach. I feel that both books compliment each other. The users manual gives you the basic description of the statement or PROC or whatever you are trying to use including the options available. The class notes tend to clarify the descriptions and provide coding examples, usually for more than one type of problem.	Greg Woolridge
Mastering the SAS system, SAS programming secrets [by Aster?] are necessary tools.	Frank Schiffel
I took the SAS manuals home with me and studied and underlined. I still have my 1982 edition, which is held together with duct tape.	Dianne Rhodes
I found that the best way to learn a new computing language is [for me, anyway] to begin with a good book.	David Cassell
Manuals.	Karsten Self
Mentors, Other Programmers' SAS Code, SAS-L	
I've been reading SAS-L since early 90s. Take a look at SAS-L these days. Are we a Guild? Are there Mentors, coaching Mentees(sp)?... I'm considering this as the title of my book on macros: "A Journeyman's Reference ..." Remember the Guild System, which were organizations of craftsmen at the beginning of the Industrial Revolution?	Ron Fehd
I've learned more from SAS-L than any other one source.	Karsten Self
A mentor. Jack Hamilton was mine.	Karsten Self
SAS-L	Mike Rhoads
Having a well-experienced, patient, and helpful SAS programmer with good communication skills, as a mentor should be mentioned as a wonderful learning resource for many learners.	Tom Winn
To me the outstanding leaning tool is SAS-L. Perhaps, it will not often take you to the extremes of the SAS world, but it will provide many problems with many answers worthy of study.	Ian Whitlock
Next I would say read your colleagues code and try to get them to read yours. Also try to reads your code a year after it is written. You learn to appreciate your coding problems more with enough time.	Ian Whitlock

Work on as many different projects from as many different people as you can.	Frank Schiffel
Seeing how other people solve problems. You can do that on SAS-L	Dianne Rhodes
I've found that the best way for me to get started on each of these was through a simple introduction that piqued my curiosity and provided a set of working examples with which I could begin. In most cases these days, that intro is either: (1) a full discussion on SAS-L,	David Cassell
Number 6, in particular ["efficiency"], could be presented as a series of SAS-L messages by Paul Dorfman and Ian Whitlock.	David Cassell
SAS-L is a great resource. I use that after I have exhausted my class notes and SAS manuals.	Andrea Wainwright
Experience – Experiment – Practice	
One of the things I've done over the years is cobble code. Read a paper, find an idea, test it, or more aptly, fit it into, streamline, a program that you have working.	Ron Fehd
The best way I've found is just by writing code. Reading and especially modifying code written by other people has taught me quite a bit, both in terms of learning great ways to solve problems, and in seeing how not to approach and code something.	Nancy Brucken
Solving problems both at your own workplace, and seeing the perspectives of others, is the best teacher.	Karsten Self
Most of us learn **best** by actually using the ideas and techniques ourselves, usually (though not necessarily) with good documentation (either printed or online) nearby as a guide. Ideally, this step would follow active participation in a training course. The lessons we teach ourselves seem to stay with us the longest. However, without comprehensive training on a particular subject, our self-taught learnings tend to focus only on the problems that are at hand, and there will be gaps in our knowledge.	Tom Winn
The least efficient way to acquire SAS knowledge is by "unguided" trial and error. This situation can be improved somewhat (but often only a little bit!) if the reference manual is used as the only source of information. In the long run, the trail-and-error approach may also be the most expensive -- in terms of wasted time and effort.	Tom Winn
Practice, practice, practice.	Jim Loughlin
...but reading without programming will not take you very far. Work problems.	Ian Whitlock
I don't feel that I really "understand" what is taught until I have a chance to use it in my job.	Greg Woolridge
Experience, Experience, Experience. Books are a great starting point for anybody to learn new skills, but in all my years of using SAS I've not yet seen a substitute for hands-on getting the job done kind of learning.	Allan Page

Classes, SUGI and other SUGs

Online tutorials from SAS Institute! Check out -- http://www.sas.com/service/edu/courses/tutorials.html	Tom Winn
Self Study? Review Ian's (Whitlock) papers. Then all the SUGI and regional papers you can get your hands on.	Ron Fehd
The Writing for Reading Style Sheet (SUGI paper 38-25) has some hints about considering what you're doing, in the overview of your career.	Ron Fehd
SAS Classes	Nancy Brucken
Take several SAS classes a year.	Kim LeBouton
Proceedings (SUGI and WUSS) have also been very helpful too.	Kim LeBouton
Users group meetings, conferences, and paper	Mike Rhoads
The easiest, and maybe the most expensive, way for would-be SAS programmers to acquire those skills is by taking a well-designed training course (presented either by SAS Institute trainers, or by reputable non-SAS training vendors).	Tom Winn
The best way is take the appropriate SAS class and then "immediately" start applying whatever was taught. And then follow-up with lots of practice, on your own time if necessary.	Jim Loughlin
SAS courses, especially those taught by the SAS Institute, are a good starting point.	Greg Woolridge
I first learned SAS by taking a 40 hour 5 day course given by SAS.	Dianne Rhodes
Attend...your local user groups, regional user groups, and SUGI.	Dianne Rhodes
I also have found it useful to write papers for conferences and do the research by searching on line on Google, and going through past proceedings from SUGI and NESUG.	Dianne Rhodes
(2) a SUGI presentation, or ...a presentation at a regional SUG like WUSS or PNWSUG	David Cassell
For me, the classes are always a best first step.	Andrea Wainwright

APPENDIX C

"Best Ways To Learn SAS" Categories Ranked In Descending Vote order

Score	Rank	Way to learn SAS
53	13	"deep, compelling intellectual curiosity", or "personal motivation" ...("hungry enough")
52	12	Read Sugi/SUG Applying code examples presented in papers or on SAS-L to your own work
39	11	Take a class SAS Institute and SAS Alliance Partner Courses, immediately followed by practice, practice, practice
38	10	Read code other people's code and have them read yours. Comparing the (at least) 5 different ways to do something in SAS (this includes "The Wrong Way") helps you figure out the best way for you (see SAS-L message http://www.listserv.uga.edu/cgi-bin/wa?A2=ind0207D&L=sas-l&P=R3684 and those that follow in the thread for the perfect example of this!!!)
34	9	The Online Documentation and SAS System Help
26	8	Hard copy of the Manuals
25	7	From a Mentor
25	7	Subscribe to SAS-L
22	6	Spending time working through examples
21	5	SAS Books By Users
21	5	Use SAS Tech Support
20	4	Practice(trial and error messages)
15	3	SAS Class Notes (available by taking a SAS institute course, or from SAS Institute Publishing)
12	2	Present a paper at a SUGI or a SUG/LUG
9	1	SAS Computer Based Tutorial
9	1	Set up a user group within your workgroup, help other people with coding problems

APPENDIX D

Industry-Specific SAS Skill Sets

Industry	Skill Set
Pharmaceutical (clinical trials)	Ability to follow written instructions well, Basic Statistics , Documentation , data step , Arrays, data _null_ reports, Date Functions (intck, etc.), Input multiple file formats (CSV, XLS, TXT) , Merging, Text Functions (substr, index, etc.) , AF, Formats, Macros, Proc FREQ , Proc MEANS, Proc REPORT, PROC SQL, PROC TABULATE. , PROC TEMPLATE, SAS/ACCESS to Oracle, DB2, SAS/GRAPH , SAS/STAT
Social Science	Ability to think for yourself and problem solve, Arrays, Input multiple file formats (CSV, XLS, TXT), Text Functions (substr, index, etc.), Basic Statistics, Data cleaning, Documentation, Macros , Proc MEANS / Proc FREQ, PROC SQL
Financial	Ability to follow instructions, ACCESS to Oracle/DB2, Basic Statistics, including Regression Techniques, data step, arrays, Data _NULL_ reports, Data Automation, Excel OLE , VB/Excel Macros, Forecasting - ETS, Macros, Proc COMPUTAB - (if you have ETS), Proc FREQ, Proc REPORT, PROC SQL , Proc SUMMARY/MEANS, Proc TABULATE, PROC TRANSPOSE, SAS/AF , SAS/Graph: GPLOT – ANNOTATE – CHART – DSGI – or equivalent in Excel Spreadsheets, SAS/OR
Manufacturing	Base - DATA Input, SAS/GRAPH, SAS/QC – Shewhart, SAS/STAT
Forestry	reading data from different sources: plain text - delimited text - Excel - dbf - access - oracle, processing: merge - sort – PROC SQL - FORMAT - date and variable functions - (format - input/put - scan - substr - etc) - save file out reporting: PROC SQL - PROC TABULATE - PROC REPORT - PROC PLOT and SAS/GARPH analyzing: PROC FREQ - PROC RANK - PROC MEANS - PROC SUMMARY - PROC UNIVARIATE statistical analysis: PROC PROC REG - PROC NLIN - PROC MIXED - PROC LOGISTIC - PROC PRINCOMP Macros
Marketing	data step, macro, AF