

Paper 173-2008

Creating Complex Reports

Cynthia L. Zender, SAS Institute, Inc., Cary, NC

ABSTRACT

Are you confused about whether you need a detail report or a summary report? Do you wonder whether you are using the right reporting procedure for your report? Have you ever spent a lot of time going down the road with one procedure, only to discover that you need to switch to a different procedure to get what you want or closer to what you want?

This paper is for intermediate SAS programmers who want to see a gallery of complex reports and review the code that produced them. It covers the pros and cons of PROC REPORT, PROC TABULATE, and DATA _NULL_ for each report. All report output is produced using ODS (rather than LISTING) output. Get out your wrenches and tool belts and fasten your seat belt for the ride, this is going to be a real nuts and bolts paper presentation.

INTRODUCTION

SAS users have a variety of reporting procedures available to them. PROC PRINT produces detail reports or listings, where one report row represents one observation in the data set or subset being processed. PROC REPORT produces both detail reports and summary reports. Like PROC PRINT, a detail report row represents one observation. However, in a summary report, each report row represents the summary of a group of observations. PROC REPORT offers more control over subtotals and grand totals than PROC PRINT. In addition, PROC REPORT allows the use of selected DATA step language statements to calculate new report items or change existing report items (such as the custom text on the summary line). PROC TABULATE, on the other hand, produces only summary reports, based on the class variables and the analysis variables used to calculate the statistics for the report cells. Although many users consider PROC SQL to be a query and reporting language, a discussion of using PROC SQL for reporting falls outside the scope of this paper.

BASIC REPORT TYPES

Reports serve many purposes. There is the “just get me a number” report, the “I’m checking the data” report, the “who/what/where/when” report, the production report, the “get the information for another product” report, the “send this to the VP’s BlackBerry” report, the “put this information on a Web site” report, and so on. What all reports have in common is that there is some kind of input (usually the data), some kind of process, selection criteria to limit the data, statistics to be produced, groups to be set, or recoding to be done. And, there is some kind of output (the report). The report has to go someplace. It has to be used, distributed, or routed to a printer, a Web server, an e-mail box, or a snail mail box. Frequently, what happens to a report after it is produced controls how the report is written or what the report looks like. If you are generating the “just get me a number” report, then the cosmetics of the report do not need much consideration. But, if you are generating a production report, it is possible that the report will be e-mailed to company directors, and cosmetics need more consideration. It is possible that the report will be submitted to a governing body where cosmetics are controlled by the governing body.

Data is not the only input for the report process. All of the considerations mentioned above are input. These considerations are related to the data, and are questions you should answer before you start coding.

- Do you need a detail report or a summary report?
- Is the data ready to use or are there transformations to be performed?
- How much weight should be given to cosmetic issues? What are the cosmetic issues or requirements?
- What is the ultimate distribution of the report?

The answers to all of these questions control the ODS destination that you use and the procedure and syntax you use to create and enhance the program output. The answers will help you make other decisions before you start coding. For example, your boss wants some summary numbers, but wants them in Microsoft Excel to use them as input to a budget process driven by Microsoft Excel macros. The data might need to be lightly summarized, perhaps with PROC MEANS or PROC TABULATE. To get the output to Microsoft Excel in the quickest way, you might use ODS CSV or PROC EXPORT. There are no cosmetic issues if all your boss wants are the data points.

But, what if your boss wants a report for distribution? What if the report needs to be composed of two parts—a summary report and a detailed list of all data rows. The report is going to be delivered via the Web, but cannot be in HTML format. It has to be in a format that is not easily edited. These requirements point to a PDF file. In one ODS “sandwich,” you can create a single PDF file that holds the summary report and the detailed list of all data rows (the detail report). When creating a PDF with ODS, you can automatically create PDF bookmarks and a Table of Contents for the two different parts of the report. The first, most important question to answer is whether your report needs to be a detail report or a summary report.

DETAIL OR LISTING REPORTS

Consider the data set SASHELP.CLASS. It has 19 observations, 9 female students, and 10 male students. The following reports are detail reports, produced with PROC PRINT and PROC REPORT for the subset of female students.

PROC PRINT Detail Report				PROC REPORT Detail Report			
Sex	Name	Age	Height	Sex	Name	Age	Height
F	Alice	13	56.5	F	Alice	13	56.5
	Barbara	13	65.3		Barbara	13	65.3
	Carol	14	62.8		Carol	14	62.8
	Jane	12	59.8		Jane	12	59.8
	Janet	15	62.5		Janet	15	62.5
	Joyce	11	51.3		Joyce	11	51.3
	Judy	14	64.3		Judy	14	64.3
	Louise	12	56.3		Louise	12	56.3
	Mary	15	66.5		Mary	15	66.5
N = 9				Female Students N = 9			

Figure 1: PROC PRINT Detail Report Compared to PROC REPORT Detail Report

Many users would call this a listing report because it lists a subset of observations, and each observation has its own row. I prefer not to mix two different meanings of the term “listing,” one being a listing report and the other being the ODS LISTING destination. My preference is to describe the above reports as detail reports. Except for minor cosmetic differences, you get the same kind of detail report from PROC PRINT and PROC REPORT. In fact, you could generate the report with a DATA _NULL_ program as well, but that seems to me like using a sledgehammer to pound in a thumbtack. DATA _NULL_ reporting capabilities have changed significantly since the days of FILE PRINT and PUT statements sending monospace text to the LISTING window. DATA _NULL_ can produce detail or summary reports. When used with ODS, these reports can be only tabular in structure (in SAS 9.1.3). Future developments for DATA _NULL_ reporting capabilities include a new object dot syntax for report creation, as described in the SUGI paper “Next Generation Data _NULL_ Report Writing Using ODS OO Features.” This new object dot syntax allows free-format reports to be created with DATA _NULL_ and ODS.

If you did use DATA _NULL_ to create a detail report on the female students from SASHELP.CLASS, the report would appear as shown in Figure 2.

Sex	Name	Age	Height
F	Alice	13	56.5
	Barbara	13	65.3
	Carol	14	62.8
	Jane	12	59.8
	Janet	15	62.5
	Joyce	11	51.3
	Judy	14	64.3
	Louise	12	56.3
	Mary	15	66.5
	Female Student N =		9

Figure 2: DATA _NULL_ Detail Report

You can customize the look of this report by changing the ODS table template associated with the DATA step, or by creating a custom table template to use with DATA _NULL_ (as described in my SUGI paper “The Power of Table Templates and DATA _NULL_”). However, both PROC PRINT and PROC REPORT can produce this kind of detail report, so no further discussion of DATA _NULL_ for detail reports is necessary.

SIMPLE SUMMARY REPORTS

Let’s look at a summary report so we can see the difference between summary reports produced with PROC REPORT versus PROC TABULATE. This time, we will use SASHELP.SHOES, which has 10 regions, 8 products, and 395 observations of sales information. Figure 3 shows the summary report from PROC REPORT and Figure 4 shows the summary report from PROC TABULATE.

The PROC REPORT table is a flat, more streamlined table. The PROC TABULATE table looks like a crosstab table. Regardless of which procedure produced the report, the report is a summary report because every report row represents the summary of a group of observations. For example, the report row for Canada represents the summary of 37 observations. More levels of granularity are available with both procedures. For example, to see the breakdown of products within Canada requires only minor syntax changes in both procedures. This output is shown in Figure 5 and Figure 6.

Region	Sales				
	Total	Count	Min	Avg	Max
Africa	\$2,342,588	56	801	41,832	360,209
Asia	\$460,231	14	937	32,874	149,013
Canada	\$4,255,712	37	1,190	115,019	757,798
Central America/Caribbean	\$3,657,753	32	936	114,305	576,112
Eastern Europe	\$2,394,940	31	712	77,256	304,093
Middle East	\$5,631,779	24	449	234,657	1,298,717
Pacific	\$2,296,794	45	325	51,040	373,908
South America	\$2,434,783	54	1,716	45,089	245,757
United States	\$5,503,986	40	554	137,600	456,985
Western Europe	\$4,873,000	62	736	78,597	502,636
Grand Total	\$33,851,566	395	325	85,700	1,298,717

Figure 3: PROC REPORT Summary Report

Region	Sales				
	Total	Count	Min	Mean	Max
Africa	\$2,342,588	56	801	41,832	360,209
Asia	\$460,231	14	937	32,874	149,013
Canada	\$4,255,712	37	1,190	115,019	757,798
Central America/Caribbean	\$3,657,753	32	936	114,305	576,112
Eastern Europe	\$2,394,940	31	712	77,256	304,093
Middle East	\$5,631,779	24	449	234,657	1,298,717
Pacific	\$2,296,794	45	325	51,040	373,908
South America	\$2,434,783	54	1,716	45,089	245,757
United States	\$5,503,986	40	554	137,600	456,985
Western Europe	\$4,873,000	62	736	78,597	502,636
Grand Total	\$33,851,566	395	325	85,700	1,298,717

Figure 4: PROC TABULATE Summary Report

		Sales				
Region	Product	Total	Count	Min	Avg	Max
Canada	Boot	\$385,613	5	7,892	77,123	286,497
	Men's Casual	\$441,903	4	15,403	110,476	353,361
	Men's Dress	\$920,101	4	12,775	230,025	757,798
	Sandal	\$14,798	5	1,190	2,960	5,120
	Slipper	\$952,751	5	5,676	190,550	700,513
	Sport Shoe	\$140,389	5	2,598	28,078	64,026
	Women's Casual	\$410,807	4	18,712	102,702	304,106
	Women's Dress	\$989,350	5	12,601	197,870	756,347
Canada		\$4,255,712	37	1,190	115,019	757,798

Figure 5: PROC REPORT Summary Report for Canada

		Sales				
		Total	Count	Min	Mean	Max
Region	Product					
Canada	Boot	\$385,613	5	7,892	77,123	286,497
	Men's Casual	\$441,903	4	15,403	110,476	353,361
	Men's Dress	\$920,101	4	12,775	230,025	757,798
	Sandal	\$14,798	5	1,190	2,960	5,120
	Slipper	\$952,751	5	5,676	190,550	700,513
	Sport Shoe	\$140,389	5	2,598	28,078	64,026
	Women's Casual	\$410,807	4	18,712	102,702	304,106
	Women's Dress	\$989,350	5	12,601	197,870	756,347
	Subtotal	\$4,255,712	37	1,190	115,019	757,798

Figure 6: PROC TABULATE Summary Report for Canada

SUMMARY REPORTS WITH MANY CROSSINGS

I consider PROC TABULATE to be the "Iron Chef" of SAS software, slicing and dicing its way to report victory. PROC TABULATE allows you to have crossings in the row and column dimensions of a table, and to have a third dimension—the page dimension—as shown in Figure 7.

Region Canada	Shoe Styles				N	All Products				
	Men's Casual	Men's Dress	Women's Casual	Women's Dress		Sum	Region %	% of All Regions		
Subsidiary										
Calgary	.	\$12,775	.	\$12,601	2	\$25,376	0.91%	0.56%		
Montreal	\$53,929	\$112,009	\$24,497	\$132,638	4	\$323,073	11.69%	7.22%		
Ottawa	\$19,210	.	\$18,712	\$33,824	3	\$71,746	2.59%	1.60%		
Toronto	\$15,403	\$37,519	\$63,492	\$53,940	4	\$170,354	6.16%	3.81%		
Vancouver	\$353,361	\$757,798	\$304,106	\$756,347	4	\$2,171,612	78.62%	48.58%		
SubTotal	\$									
	Region Pacific	Shoe Styles				N	All Products			
	Men's Casual	Men's Dress	Women's Casual	Women's Dress	Sum		Region %	% of All Regions		
	Subsidiary									
	Auckland	.	\$4,892	\$18,189	\$42,760	3	\$65,841	3.85%	1.47%	
	Canberra	\$24,733	\$4,231	\$15,032	\$21,486	4	\$65,482	3.83%	1.46%	
	Jakarta	\$373,908	\$186,631	\$18,761	\$4,659	4	\$583,959	34.19%	13.06%	
	Kuala Lumpur	\$106,657	\$55,667	\$36,110	\$61,108	4	\$259,542	15.19%	5.80%	
	Manila	\$128,309	\$144,151	\$131,794	\$229,372	4	\$633,626	37.10%	14.17%	
	Singapore	\$28,761	\$30,619	.	\$40,056	3	\$99,436	5.82%	2.22%	
	SubTotal	\$662,368	\$426,191	\$219,886	\$399,441	22	\$1,707,886	100.00%	38.20%	

Figure 7: PROC TABULATE Summary Report with Page, Row, and Column Dimensions

PROC TABULATE shines when multiple crossings of highly summarized data need to be presented. Although PROC REPORT and DATA _NULL_ could produce this report, a summary report of this nature falls within PROC TABULATE's capabilities. PROC TABULATE can have 10 variables or statistics nested in the row dimension, and 10 variables or statistics nested in the column dimension. The data for PROC TABULATE does not have to be sorted. And, in SAS 9, you can have up to 2,147,483,647 ($2^{31} - 1$) potential interactions for a crossing. That is a lot of slicing and dicing!

GRAPHICAL REPORTS

SAS/GRAPH procedures like PROC GPLOT produce the graphical equivalent of a detail report. Instead of listing every observation in a table, the observations might be graphically represented on a plot across time. PROC GCHART can produce the graphical equivalent of a summary report, where each bar or pie slice represents a group of observations. The height of the bar or size of the pie slice represents a statistic calculated for the group. Some bar chart reports are shown later in the complex examples.

In addition to standard SAS/GRAPH capabilities, the new ODS GRAPHICS facility allows SAS/STAT procedures to produce complex analytical graphical reports directly from the procedure, as long as ODS GRAPHICS is turned on. A discussion of ODS GRAPHICS falls outside the scope of this paper.

SUMMARY OF SAS REPORT TABLE PRODUCTION

Procedure	Report Type	Highlights/Features
PRINT	Detail Reports (one report row for every observation in the data, or for every observation in the subset)	Easy to use Statement-oriented syntax (VAR, SUM, FORMAT, etc.) Data must be sorted to perform BY-group processing Data must be summarized by another procedure (such as PROC MEANS) if summary reports with PROC PRINT are needed In 9.2, you can get reader breaks automatically with PROC PRINT Easy to use syntax for ODS for style changes and trafficlighting
REPORT	Detail Reports (one report row for every observation in the data, or for every observation in the subset) Summary Reports (one report row that collapses or summarizes data based on the group or category variables)	Can calculate new report items Report-item-oriented syntax (DEFINE statement for every report item) Can change report items (COMPUTE block syntax) Can use temporary variables in report (COMPUTE block syntax) Allows spanning headers More extensive group break control (can get subtotals only, end-of-report totals only, or both) Data does not have to be sorted to define variables as GROUP or ORDER variables Easy to use syntax for ODS for style changes
TABULATE	Summary Reports (one report row that collapses or summarizes data based on the group or category variables)	Statement-oriented syntax (CLASS, VAR, TABLE, etc.) Many crossings in row and column dimensions Page dimension or BY-group processing extends table capabilities Can use standard denominator for report percentages (row percentage, column percentage, overall percentage) or can specify your own denominator definition Extensive options for LISTING and other ODS destinations Succinct TABLE statement syntax Data must be sorted to perform BY-group processing Easy to use syntax for ODS for style changes
DATA_NULL_ with LISTING Destination Syntax	Detail Reports Summary Reports Free-Format Reports (FILE PRINT and PUT syntax)	The power of DATA step programming available Can write report lines based on internal program criteria Can calculate new report items in DATA step Can write free-format reports in monospace font in the OUTPUT or LISTING window Can perform data manipulation and reporting in one step
DATA_NULL_ with Default ODS Syntax	Detail Reports Summary Reports	The power of DATA step programming available Can write report lines based on internal program criteria Can calculate new report items in DATA step Can perform data manipulation and reporting in one step
DATA_NULL_ with Table Templates	Detail Reports Summary Reports Complex Tabular Reports	Can calculate new report items in table template or DATA step Can change report items Can use DATA step variables, macro variables, and dynamic table template variables in report Allows spanning headers New CELLSTYLE-AS syntax available for conditional style changes in template Can define and use generic columns

Although the preceding table outlines the major report procedures, it does not include PROC SQL, SAS/GRAPH procedures, or ODS GRAPHICS for SAS/STAT procedures. In the next table, the types, benefits, and capabilities of the major ODS destinations are presented.

SUMMARY OF ODS DESTINATION FEATURES

Destination	Report Type	Highlights/Features
LISTING	Verify report structure or data Quick and dirty ad hoc reports	Monospace font No style elements supported for tabular results Starting in SAS 9, style specification supported for SAS/GRAPH results
HTML-Based HTML3 HTML MSOFFICE2K HTML HTMLCSS	Production reports Reports for the Web	ODS creates all HTML tags according to HTML specifications Many software programs can open HTML Easy to use syntax for ODS for style changes HTML-based destinations use SAS style templates or CSS files
Spreadsheet-Based CSV CSVALL EXCELXP	Production reports where data or report needs to be in spreadsheet	ODS creates CSV file or SpreadsheetML file CSV files are in one worksheet with no style ExcelXP creates multiple worksheet structure automatically ExcelXP uses SAS style template
RTF-Based RTF TAGSETS.RTF OpenOffice RTF	Production reports where report needs to be opened in word processor	Most word processors open RTF files Files can be easily edited, text can be added, and files can be resaved RTF destinations use SAS style templates Output can be enhanced with RTF control strings and ODS ESCAPECHAR
Printer-Family PDF PS PCL	Production reports where editing is not required	Most PDF files can be opened with PDF readers, such as Adobe Acrobat Reader or Ghostview PostScript files can be changed and distilled to PDF PCL files are designed for HP printers Printer-family destinations use ODS style templates

Choosing a procedure or destination depends on each report's requirements. Let's look at some complex reports sent to a variety of ODS destinations.

GALLERY OF COMPLEX REPORTS

The examples in this paper span the gamut from a jazzed-up detail report to a graph with a tabular report. To help users navigate the programs that they download, the following gallery shows complex reports addressed in this paper. A brief discussion of each report and techniques used to produce the report are included.

As I set out to write this paper, I used examples collected from many years of programming with SAS and teaching about SAS reporting procedures. What I have discovered is that when designing reports, it's better for me to start with an idea of what the report should look like and then work backward to the data. Perhaps this is a holdover from my Fortran and COBOL days when I had to map out a final report byte by byte on an IBM report layout form.

Seeing the reports makes transformations clear and helps you identify any extra formatting you need or any styles you should override. The programs for these complex reports are lengthy so they are not included in the paper. Instead, all of the data and programs are downloadable from the SAS Global Forum Web site.

COMPLEX EXAMPLE 1: DEMOGRAPHIC REPORTS

There are two different demographic reports in the first example. They are basically two different versions of the same report. PROC TABULATE cannot produce this report because disparate variable values (such as **N** and **PCTN**) have to appear in the same cell in a specific format (nn, (%)) and PROC TABULATE cannot handle this requirement. Also, PROC TABULATE cannot combine other statistics in a cell with punctuation. This is the report most frequently requested in class.

Demographic Report in JOURNAL Style	Demographic Report in Default Style																																																																																																																
Using PROC REPORT With JOURNAL Style <hr/> Patient Characteristics (N= 1,049 Patients) <hr/> Patient Demographics Age 69.80 ± 11.409 (32 - 88) Gender Female 473 (45.09) Male 576 (54.91) LVEF 32.66 ± 17.762 (5 - 65) CAD(required) No 283 (26.98) Yes 766 (73.02) MI Status Without MI 367 (34.99) With MI 682 (65.01) Hypertension Status No Hypertension 549 (52.34) Hypertension 500 (47.66) NYHA Functional Class I 376 (35.84) II 234 (22.31) III 158 (15.06) IV 149 (14.20) UNK 132 (12.58)	Proc Report Alternate Demographic Example Table 1: Multivariate Predictors of Outpatient Alcohol Treatment (n=284) <table border="1"> <thead> <tr> <th>Variable</th> <th>N (%)</th> <th>OR (95% CI)</th> <th>p-value</th> </tr> </thead> <tbody> <tr> <td colspan="4">Age</td> </tr> <tr> <td>18-25</td> <td>20 (7.04)</td> <td>1</td> <td></td> </tr> <tr> <td>26-30</td> <td>19 (6.69)</td> <td>0.932 (0.435 1.996)</td> <td>0.8356</td> </tr> <tr> <td>31-35</td> <td>25 (8.80)</td> <td>2.042 (1.069 3.898)</td> <td>0.0305</td> </tr> <tr> <td>36-40</td> <td>95 (33.45)</td> <td>2.762 (1.511 5.048)</td> <td>0.001</td> </tr> <tr> <td>41-45</td> <td>65 (22.89)</td> <td>2.331 (1.248 4.356)</td> <td>0.008</td> </tr> <tr> <td>>45</td> <td>60 (21.13)</td> <td>2.483 (1.326 4.648)</td> <td>0.0045</td> </tr> <tr> <td colspan="4">Education level</td> </tr> <tr> <td>Less than HS</td> <td>28 (9.86)</td> <td>1</td> <td></td> </tr> <tr> <td>HS/GED</td> <td>93 (32.74)</td> <td>0.948 (0.544 1.650)</td> <td>0.8493</td> </tr> <tr> <td>Some College</td> <td>114 (40.14)</td> <td>0.931 (0.532 1.628)</td> <td>0.8018</td> </tr> <tr> <td>College</td> <td>28 (9.86)</td> <td>0.479 (0.243 0.946)</td> <td>0.034</td> </tr> <tr> <td>Graduate School</td> <td>21 (7.39)</td> <td>0.723 (0.347 1.503)</td> <td>0.3845</td> </tr> <tr> <td colspan="4">Annual Pre-tax Income Level</td> </tr> <tr> <td>None</td> <td>28 (9.86)</td> <td>1</td> <td></td> </tr> <tr> <td><\$20,000</td> <td>170 (59.86)</td> <td>1.121 (0.641 1.962)</td> <td>0.6879</td> </tr> <tr> <td>\$20,000 - 59,999</td> <td>61 (21.48)</td> <td>0.708 (0.380 1.319)</td> <td>0.2766</td> </tr> <tr> <td>\$60,000 - 100,000</td> <td>7 (2.46)</td> <td>0.447 (0.166 1.207)</td> <td>0.1123</td> </tr> <tr> <td>>\$100,000</td> <td>18 (6.34)</td> <td>0.542 (0.163 1.807)</td> <td>0.319</td> </tr> <tr> <td colspan="4">Race/Ethnicity</td> </tr> <tr> <td>White</td> <td>76 (26.76)</td> <td>1</td> <td></td> </tr> <tr> <td>African American</td> <td>16 (5.63)</td> <td>1.751 (1.183 2.590)</td> <td>0.0051</td> </tr> <tr> <td>Hispanic/Latino</td> <td>31 (10.92)</td> <td>0.900 (0.554 1.463)</td> <td>0.6728</td> </tr> <tr> <td>Native American</td> <td>34 (11.97)</td> <td>2.141 (0.412 11.124)</td> <td>0.365</td> </tr> <tr> <td>Asian/PI</td> <td>45 (15.83)</td> <td>0.283 (0.037 2.152)</td> <td>0.2225</td> </tr> <tr> <td>Mixed</td> <td>48 (16.90)</td> <td>1.338 (0.775 3.052)</td> <td>0.2184</td> </tr> <tr> <td>Other</td> <td>34 (11.97)</td> <td>1.486 (0.464 4.756)</td> <td>0.5043</td> </tr> </tbody> </table>	Variable	N (%)	OR (95% CI)	p-value	Age				18-25	20 (7.04)	1		26-30	19 (6.69)	0.932 (0.435 1.996)	0.8356	31-35	25 (8.80)	2.042 (1.069 3.898)	0.0305	36-40	95 (33.45)	2.762 (1.511 5.048)	0.001	41-45	65 (22.89)	2.331 (1.248 4.356)	0.008	>45	60 (21.13)	2.483 (1.326 4.648)	0.0045	Education level				Less than HS	28 (9.86)	1		HS/GED	93 (32.74)	0.948 (0.544 1.650)	0.8493	Some College	114 (40.14)	0.931 (0.532 1.628)	0.8018	College	28 (9.86)	0.479 (0.243 0.946)	0.034	Graduate School	21 (7.39)	0.723 (0.347 1.503)	0.3845	Annual Pre-tax Income Level				None	28 (9.86)	1		<\$20,000	170 (59.86)	1.121 (0.641 1.962)	0.6879	\$20,000 - 59,999	61 (21.48)	0.708 (0.380 1.319)	0.2766	\$60,000 - 100,000	7 (2.46)	0.447 (0.166 1.207)	0.1123	>\$100,000	18 (6.34)	0.542 (0.163 1.807)	0.319	Race/Ethnicity				White	76 (26.76)	1		African American	16 (5.63)	1.751 (1.183 2.590)	0.0051	Hispanic/Latino	31 (10.92)	0.900 (0.554 1.463)	0.6728	Native American	34 (11.97)	2.141 (0.412 11.124)	0.365	Asian/PI	45 (15.83)	0.283 (0.037 2.152)	0.2225	Mixed	48 (16.90)	1.338 (0.775 3.052)	0.2184	Other	34 (11.97)	1.486 (0.464 4.756)	0.5043
Variable	N (%)	OR (95% CI)	p-value																																																																																																														
Age																																																																																																																	
18-25	20 (7.04)	1																																																																																																															
26-30	19 (6.69)	0.932 (0.435 1.996)	0.8356																																																																																																														
31-35	25 (8.80)	2.042 (1.069 3.898)	0.0305																																																																																																														
36-40	95 (33.45)	2.762 (1.511 5.048)	0.001																																																																																																														
41-45	65 (22.89)	2.331 (1.248 4.356)	0.008																																																																																																														
>45	60 (21.13)	2.483 (1.326 4.648)	0.0045																																																																																																														
Education level																																																																																																																	
Less than HS	28 (9.86)	1																																																																																																															
HS/GED	93 (32.74)	0.948 (0.544 1.650)	0.8493																																																																																																														
Some College	114 (40.14)	0.931 (0.532 1.628)	0.8018																																																																																																														
College	28 (9.86)	0.479 (0.243 0.946)	0.034																																																																																																														
Graduate School	21 (7.39)	0.723 (0.347 1.503)	0.3845																																																																																																														
Annual Pre-tax Income Level																																																																																																																	
None	28 (9.86)	1																																																																																																															
<\$20,000	170 (59.86)	1.121 (0.641 1.962)	0.6879																																																																																																														
\$20,000 - 59,999	61 (21.48)	0.708 (0.380 1.319)	0.2766																																																																																																														
\$60,000 - 100,000	7 (2.46)	0.447 (0.166 1.207)	0.1123																																																																																																														
>\$100,000	18 (6.34)	0.542 (0.163 1.807)	0.319																																																																																																														
Race/Ethnicity																																																																																																																	
White	76 (26.76)	1																																																																																																															
African American	16 (5.63)	1.751 (1.183 2.590)	0.0051																																																																																																														
Hispanic/Latino	31 (10.92)	0.900 (0.554 1.463)	0.6728																																																																																																														
Native American	34 (11.97)	2.141 (0.412 11.124)	0.365																																																																																																														
Asian/PI	45 (15.83)	0.283 (0.037 2.152)	0.2225																																																																																																														
Mixed	48 (16.90)	1.338 (0.775 3.052)	0.2184																																																																																																														
Other	34 (11.97)	1.486 (0.464 4.756)	0.5043																																																																																																														

Figure 8: Demographic Summary Reports

Both versions of this report are delivered to the RTF destination. This type of report is rarely delivered in HTML and only sometimes delivered in PDF. Because the data set was not in the appropriate format, the major steps to create the report were:

- 1) Summarize the relevant variables.
- 2) Restructure the summarized data with a DATA step to be in the same structure as the report, making a character string from the summary statistics.
- 3) Use ODS and PROC REPORT to make each major header bold and to indent the category information.

COMPLEX EXAMPLE 2: COLOR-CODED DETAIL REPORT WITH VERTICAL HEADERS

This report is a true detail report. The data is a simple table of instructor names. Each variable in the data also represents a course in a particular curriculum. The report did not require data to be transformed, but two summary lines were needed. One line was needed for courses where instructor certification was completed. And, one line was needed for courses where instructor certification was in process.

Instructor	BI User Courses							BI Analyst Courses					BI AppDev			Other					
	SBIFAST	LWSSIOV1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	
Adams, Alfred				C	C	C	C	P		C	C										
Baffin, Barb		C	C		C	C	C			C	C					C			P	P	
Chevell, Carla			C		C	C	C			C	C	C			P				P		
DuPont, Dan			C		C	C	C				C	C							C		
Edgewater, Elizabeth		P								C	C										
Frankhart, Frederick							C				C	C									
Gawaine, Georgia	C		C		C		C		C	C	C	C	C	C	C						
Herman, Harry																P					
Ignatz, Irene			C	C	C	C	P				C	C	P					P			
Johanssen, Jerry			C	C	C									P					P	P	
Kimmet, Katherine		C	C	C																P	
LaFontaine, Larry								C					P			P					
Morrison, Marie			C					P		P						C	C				
Nordstrom, Norma			C																		
Olivier, Oscar					C	P	C				C	C				C	P				
Pendergast, Peter			C		C	C	C			C	C	C			C						
Quincy, Quentin			C		C	C	C			C	C	C	P	C							
Ristow, Rita													C						C	C	C
Steinmetz, Steve							P		P			C									
Thomas, Tony		P	C		C	C	C			C	C										
Now Certified	1	2	12	4	11	9	10	1	1	8	12	10	1	2	4	3	4	2	1	1	
In Process	0	2	0	0	0	2	1	3	0	1	0	1	2	1	4	3	4	2	2	2	
Instructor	SBIFAST	LWSSIOV1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	SBIOR1	

Figure 9: Certification Detail Report

Trafficlighting was needed for each instructor, based on whether he or she had a **C** or a **P** in all of the courses for a particular group. I am not a fan of vertical headers. But, in this case, vertical headers were required for the managers to make the report as narrow as possible and fit in one browser screen. To produce the report, the specifics are:

- 1) The basic report is a detail report, which usually points to PROC PRINT. But, the need for alternate row highlighting pointed to PROC REPORT or DATA _NULL_ with ODS.
- 2) The vertical headers at the top of the table could be done with PROC PRINT. But, the need for a table footer and the need to do conditional trafficlighting based on the values in more than one cell required PROC REPORT or DATA _NULL_ with ODS.
- 3) Although the first draft of the report just used PROC REPORT to produce the basic listing with alternate row highlighting, the need for trafficlighting was verbose when PROC REPORT was the only thing used. As a result, preprocessing of the data with a DATA step program was performed before running PROC REPORT.

COMPLEX EXAMPLE 3: SUMMARY REPORT ORDERED DIFFERENTLY FOR EACH REGION

The following report is also frequently requested—a summary report ordered by the summary value of a column (for example, by descending grouped sales within each region such as the report on the left in Figure 10). This includes a summary report that is in an order that is not a default way to order a report. The report on the right in Figure 10 is in the same order as the overall total sales report, where slipper is the product with the highest overall sales, followed by women's dress shoes.

Every Region Ordered by Region's Descending Sales			Every Region Ordered by Overall Total Descending Sales		
Region	Product	Total Sales	Region	Product	Total Sales
Asia	Slipper	\$152,032	Asia	Slipper	\$152,032
	Men's Dress	\$119,366		Women's Dress	\$78,234
	Women's Dress	\$78,234		Men's Dress	\$119,366
	Boot	\$62,708		Men's Casual	\$11,754
	Women's Casual	\$25,837		Women's Casual	\$25,837
	Men's Casual	\$11,754		Boot	\$62,708
	Sandal	\$8,208		Sport Shoe	\$2,092
	Sport Shoe	\$2,092		Sandal	\$8,208
Asia		\$460,231	Asia		\$460,231
Canada	Women's Dress	\$989,350	Canada	Slipper	\$952,751
	Slipper	\$952,751		Women's Dress	\$989,350
	Men's Dress	\$920,101		Men's Dress	\$920,101
	Men's Casual	\$441,903		Men's Casual	\$441,903
	Women's Casual	\$410,807		Women's Casual	\$410,807
	Boot	\$385,613		Boot	\$385,613
	Sport Shoe	\$140,389		Sport Shoe	\$140,389
	Sandal	\$14,798		Sandal	\$14,798
Canada		\$4,255,712	Canada		\$4,255,712

Figure 10: Summary Reports in Different Orders

To produce the report:

- 1) Make multiple passes through the file to create a regional order variable and an overall order variable on which ordering can be performed.
- 2) Perform break processing to create subtotals at the region level.

In the end, this summary report was produced using PROC REPORT.

COMPLEX EXAMPLE 4: TUFTE-INSPIRED SPARKLINE

This example was inspired by Edward Tufte’s ideas about data presentation. In one of his books, Tufte advocated the use of sparkline data representations to convey a lot of information in a condensed format.

		Record (24 Games)		Wins	Ties	Losses
Division	Team	Sparkline 1 color	Sparkline 2 colors			
Eastern	NESUGgers	██-███	███-███	15	3	6
	SAS Macros	███-███	███-███	15	2	7
Central	MWSUG Data Steppers	███-███	███-███	16	3	5
	Pharma Sluggers	███-███	███-███	14	2	8
Western	PNW Forecasters	███-███	███-███	15	3	6
	West Coast WUSSers	███-███	███-███	15	3	6

(Green=Home Game; Gray=Away Game)

Figure 11: Sparkline Summary Report of Wins, Losses, and Ties

Creating Figure 11 required team wins, losses, and ties to be transformed from W, L, and T values into keyboard characters +, -, and /, respectively. Then, the Bissantz SparkFonts 4 TrueType fonts were loaded and used for the text string of wins, losses, and ties. To get the sparkline in colors, the foreground color was changed with ODS ESCAPECHAR in a style override.

COMPLEX EXAMPLE 5: FOUR GRAPHS IN A TABLE

PROC GREPLAY is one way to get four graphs on one page. However, PROC GREPLAY has a few disadvantages. In SAS 9.1.3, the internal SAS catalog image is not stored with style template information. So, the replayed images require a lot of special formatting using SAS/GRAPH syntax to make them look like they were created with the ACTXIMG or JAVAXIMG driver (which uses style template information in SAS 9.1.3). To place four graphs in a table side by side, the HTMLPANEL destination works for HTML-based output. However, if you want PDF or RTF output, your only choice is to use the experimental ODS layout syntax.

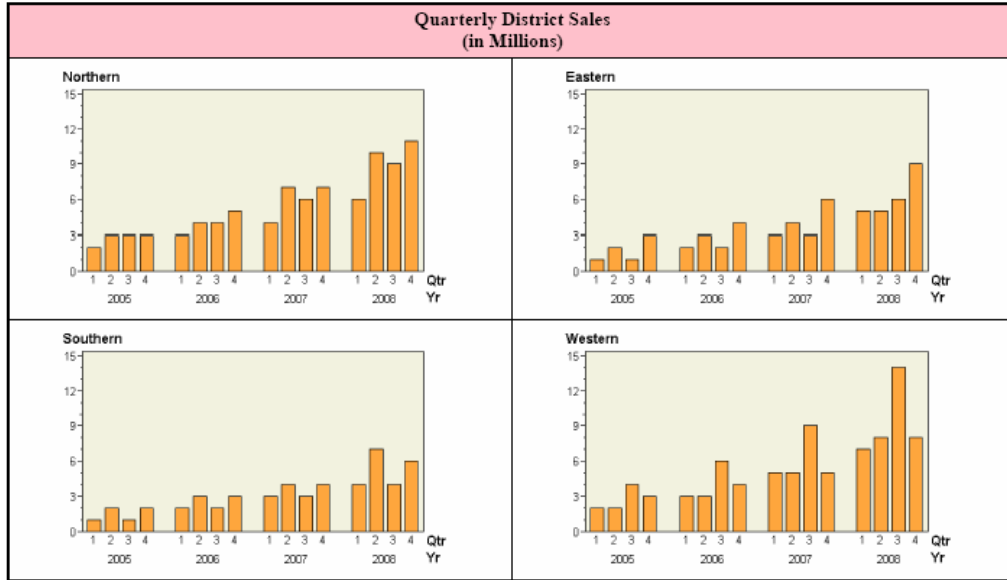


Figure 12: Four Graphs in a Table (Viewed in Web Browser)

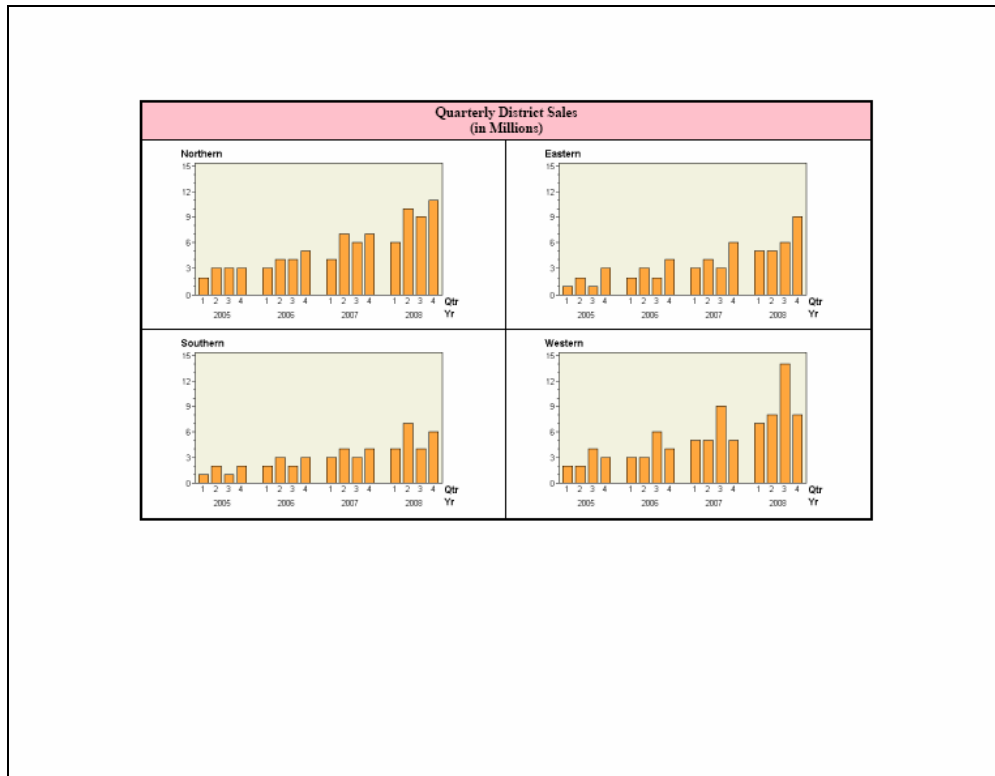


Figure 13: Four Graphs in a Table (Viewed in Adobe Acrobat Reader)

Here is another way to create this kind of report:

- 1) Create the graph files using ODS and the ACTXIMG driver. Use the HSIZE and VSIZE GOPTIONS to size the graph.
- 2) Create a data set that has the number of rows and columns that you want your table to have. In the previous output, there are two rows and two columns.
- 3) Use the PREIMAGE style attribute to put the correct graph file into the correct table cell.

COMPLEX EXAMPLE 6: GRAPH AND TABLE RESULTS DISPLAYED IN SAME REPORT

This example is a variation of the previous example. The quarterly graph for each district is displayed next to the yearly totals for that district.

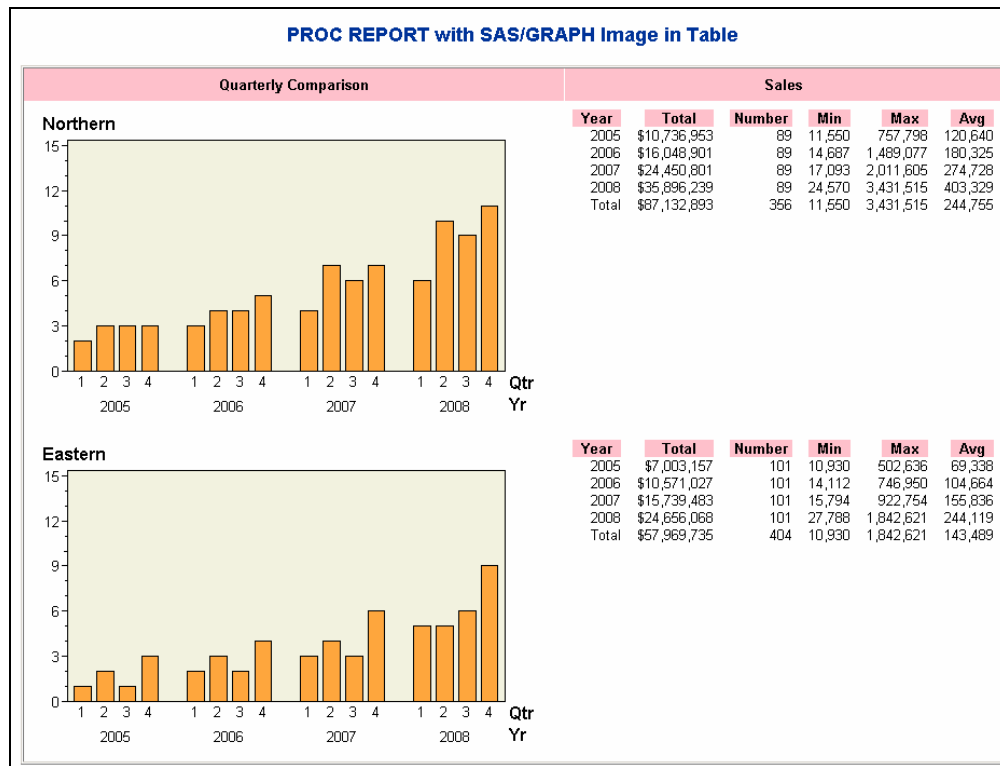


Figure 14: Table and Graph in Same Report Created with PROC REPORT

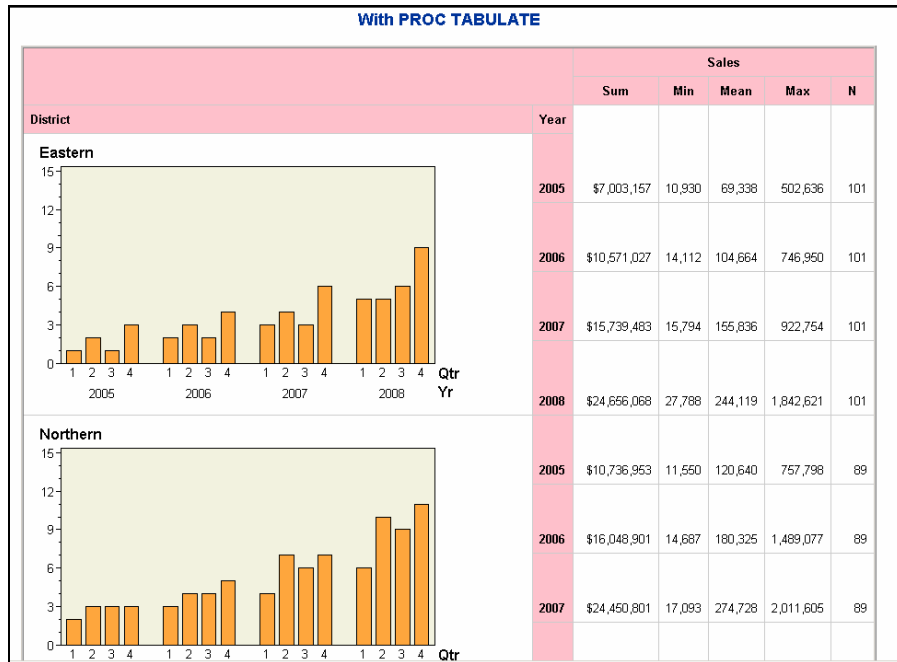


Figure 15: Table and Graph in Same Report Created with PROC TABULATE

To produce the report:

- 1) Use the PREIMAGE style attribute to place each district's image into the table.
- 2) Modify the cell with the graph so that it spans all of the other report rows. This can be achieved with PROC TABULATE and the HTML destination, as shown in Figure 15.
- 3) To make the District cell span all of the yearly rows in all destinations, using PROC REPORT, build a large character variable for YEAR, TOTAL, N, MIN, MAX, and AVG variable information, and insert a line break between each year's value.
- 4) Run PROC REPORT on the data from step 3.

COMPLEX EXAMPLE 7: HIGHLIGHTING TABLE ROWS WITH <PARENT> SYNTAX

It is easy to highlight a row in PROC REPORT, just use CALL DEFINE syntax. It is not as easy in PROC TABULATE. In the following example, highlighting is set by using the <PARENT> specification for a style element to indicate to PROC TABULATE that row highlighting should be inherited from the parent cells. Figure 16 shows the row-level banding.

		Retail		Wholesale		Total	
		Quantity	Amount	Quantity	Amount	Quantity	Amount
		Sum	Sum	Sum	Sum	Sum	Sum
NC	L	Missing	Missing	2272.00	45440.00	2272.00	45440.00
	M	1066.00	26600.00	1066.00	21320.00	2132.00	47920.00
	S	472.00	11800.00	472.00	9440.00	944.00	21240.00
NE	L	2421.00	60525.00	2421.00	48420.00	4842.00	108945.00
	M	1825.00	45625.00	1825.00	36500.00	3650.00	82125.00
	S	623.00	15575.00	623.00	12460.00	1246.00	28035.00
SO	L	2303.00	57575.00	2303.00	46060.00	4606.00	103635.00
	M	2149.00	54725.00	2149.00	42980.00	4298.00	97705.00
	S	1254.00	31150.00	Missing	Missing	1254.00	31150.00
WE	L	2655.00	66375.00	2655.00	53100.00	5310.00	119475.00
	M	2360.00	59000.00	2360.00	47200.00	4720.00	106200.00
	S	561.00	14025.00	561.00	11220.00	1122.00	25245.00

Figure 16: PROC TABULATE Report with Banded Highlighting (Viewed in Web Browser)

Furthermore, it was hard for me to imagine using this same <PARENT> specification in a style override until David Kelley explained it to me.

- 1) A user-defined format sets the background colors for all four regions. NC and SO are set to beige, and NE and WE are set to pink (the same as the column headings).
- 2) CITYSIZE style attributes are inherited from the parent REGION.
- 3) Because of parent-level style-attribute inheritance, all of the statistics calculated for the crossing use the REGION background color for each cell. This gives a banded or row-level highlighting effect.

COMPLEX EXAMPLE 8: DEMOGRAPHIC REPORT WITH DATA _NULL_ AND TABLE TEMPLATE

The last example brings us around full circle to a demographic report. Only this time, the demographic report is produced using DATA _NULL_ and FILE PRINT ODS. Figure 15 shows the report, which uses the JOURNAL style for ODS RTF.

<i>Demographic Report Using ODS and Table Templates</i>			
	NYHA I, II, UNK (N= 742)	NYHA III (N= 158)	NYHA IV (N= 149)
<i>Patient Age</i>			
N	742	158	149
Mean (SD)	70.1 (11.15)	69.0 (12.07)	69.4 (11.98)
Median	73.0	71.0	73.0
Min, Max	32, 88	32, 86	32, 86
<i>MI Categories</i>			
No	273 (26.0%)	42 (4.0%)	52 (4.9%)
Yes	469 (44.7%)	116 (11.0%)	97 (9.2%)

Figure 17: Demographic Report with DATA _NULL_

I consider this to be the hardest program in the set. That is because to achieve this report with DATA _NULL_, a custom table template had to be used.

- 1) The COMPLEX.PATIENT_INFO data had too many levels for the NYHA variable, so a DATA step was used to create only three levels for NYHA. PROC TABULATE was used to create a summary file based on the values of the CADMI and AGE variables.
- 2) A table template was designed with four report items: ROWLABEL, COL1, COL2, and COL3. The ROWLABEL report item serves the same purpose in this demographic report as the VARINFO column in the first two demographic reports.
- 3) Macro variables were created from the summary file for the count (N) for each NYHA group.
- 4) A DATA _NULL_ program wrote to the table template using N=PS syntax. For example, when _TYPE_=11, then the information from the summary file needed to be written to lines 9 and 10 of the report for the CADMI values in each column. When _TYPE_=10, then the information needed to be written to lines 2 through 6 of the report.

CONCLUSION

Complex reports are doable with SAS. Sometimes you have to think outside the box or outside of a single procedure. Consider the programs used in this paper:

Program Name	DATA Step Used	Procedures Used
Complex1_demog.sas	yes	MEANS, FREQ, SORT, FORMAT, REPORT
Complex2_certify.sas	yes	SORT, FORMAT, REPORT (with macro program for multiple COMPUTE blocks)
Complex3_diff_order_report.sas	no	REPORT (for summary), SQL (for join), REPORT (for final)
Complex4_win_loss.sas	yes	FORMAT, REPORT
Complex5a_smallmult_report.sas	yes	SORT, FORMAT, SQL, GCHART, REPORT
Complex6a_graph_report.sas	yes	SORT, REPORT (for summary), SQL, REPORT (uses graphs created in example 5 GCHART)
Complex7_tabulate_style.sas	no	FORMAT, TABULATE
Complex8_demog_data_null.sas	yes	TABULATE (for summary), TEMPLATE (for table template)

By the time I finished assembling all of the programs, the theme of my code was that the DATA step, whether it was manipulating data or building arrays or adjusting data set structure, was essential to the successful creation of complex reports. DATA step programs and PROC REPORT are unbeatable in the flexibility they bring to the creation of complex reports.

REFERENCES

O'Connor, Dan. 2003. "Next Generation Data _NULL_ Report Writing Using ODS OO Features." *Proceedings of the Twenty-eighth Annual SAS Users Group International Conference*. Cary, NC: SAS Institute Inc. Available at <http://www2.sas.com/proceedings/sugi28/022-28.pdf>.

Zender, Cynthia. 2005. "The Power of TABLE Templates and DATA _NULL_." *Proceedings of the Thirtieth Annual SAS Users Group International Conference*. Cary, NC: SAS Institute Inc. Available at <http://www2.sas.com/proceedings/sugi30/088-30.pdf>.

RESOURCES

You can download the programs and data in this paper from the SAS Web site for Technical Papers and Presentations for SAS Global Forum 2008 Conference at support.sas.com/rnd/papers/. Alternate versions of some of the programs are included in the downloadable zip file. These versions show more macro coding, alternate versions of the same program, or the downside of using some of the techniques in non-HTML destinations.

ACKNOWLEDGMENTS

Thanks to David Kelley and the rest of the ODS team for their help with ODS examples. In addition, this paper would not have been possible without the encouragement and support of Herb Kirk, Larry Stewart, Eric Rosslund, and Sue Rakes. Thanks to Michele Ensor, Jane Stroupe, Linda Jolley, Christine Riddiough, and Sue Rakes, who reviewed the programs in this paper and told me when it was time to stop fussing over them.

CONTACT INFORMATION

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

Cynthia L. Zender
 SAS Institute, Inc.
 Work Phone: 575-522-3803
 E-mail: Cynthia.Zender@sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.