Tips of Using Proc Report for Summary Tables and Listings

Jenny Zhang
ICON Clinical Research

ABSTRACT

The REPORT procedure is an easy-to-use, ad hoc report generator that simplifies the report development process. It combines features from the PRINT, MEANS, and TABULATE procedures with features of DATA step to provide a powerful report writing tool. It is frequently used in the pharmaceutical industry to create summary tables and listings. Many SAS® programmers have been very familiar with the basics of PROC REPORT, such as Define, Column, Break and Compute statements. However, there are many advanced futures and options available to generate customized reports. This paper presents some of these advanced options which are very useful in PROC REPORT. They can make programming job much easier. They include: (1). Using FLOW option. (2). Specify one or more ID variables to print on each page and (3). Using #BYVAL title option.

INTRODUCTION

This paper briefly describes FLOW, ID option and #BYVAL title option with several examples to demonstrate the usage of these options in PROC REPORT. These options make program tasks accomplished more efficiently.

TIP 1 - USING FLOW OPTION

The FLOW option of the PROC REPORT allows the user to include lengthy text in a report by wrapping the text within a specified column width. The FLOW option honors the split character.

In many situations, we have to use PROC REPORT to display some variables with long text within a fixed length of column, such as medical history terms in medical history summary tables.

In the following example, the length for variable C_MED is 200. The column width for C_MED is only 25. How can we arrange the proper display for the long string? The solution is option FLOW, which will wrap the string of C_MED to the next row automatically.

```
proc report data=final1 center missing headline headskip nowd split='|' spacing=0
formchar(2) = '_';
column order c_med ("Number of Subjects (%) |__" _19 _39 _49 );

define order /order  order=internal noprint;
define c_med  /display  width=25 left   "System Organ Class|    Preferred Term"
spacing=2 flow;
define _19    /display  width=15 left  "Treatment 1| (N=&bign_19)" spacing=2;
define _39    /display  width=15 left  "Treatment 2| (N=&bign_39)" spacing=2;
define _49    /display  width=15 left  "Treatemnt 3| (N=&bign_49)" spacing=2;
run;
```

Medical History by System Organ Class and Preferred Term

System Organ Class Preferred Term	Number of Subjects (%) Treatment 1 (N=182)	 Treatment 2 (N=205)	 Treatemnt 3 (N=175)
Subjects with Any Medical History	119 (65.4)	136 (66.3)	112 (64.0)
Blood and lymphatic system disorders	1 (0.5)	3 (1.5)	2 (1.1)
Anaemia	1 (0.5)	1 (0.5)	2 (1.1)
Iron deficiency anaemia	0 (0.0)	1 (0.5)	0 (0.0)
Leukopenia	0 (0.0)	1 (0.5)	0 (0.0)
Cardiac disorders	10 (5.5)	7 (3.4)	4 (2.3)
Myocardial infarction	5 (2.7)	2 (1.0)	2 (1.1)
Acute myocardial infarction	0 (0.0)	2 (1.0)	1 (0.6)
Angina unstable	1 (0.5)	1 (0.5)	1 (0.6)

If you don't use the FLOW option for a character variable that contains a large amount of text, PROC REPORT truncates the value to fit in one line of the width of the column for that variable.

TIP 2 - USING ID OPTION

When creating report, we sometimes have too many variables to print across one page and want to specify one or more ID variables to print on each page when wrapping occurs. We can use the ID option in the DEFINE statement for each ID column or subsequent pages. An ID variable and all columns to its left appear at the left of every page of ID ensures that we can identify each row of the report when report contains more columns than will fit on one page.

The following is the Lab listing generated by PROC REPORT:

Subject Profiles for Laboratory Data - Serum Chemistry

Site/ No. Subject No.	Sex/Age (years)	Collection Date	test1	test2	test3	test4
5001/001	Male/ 34	07SEP2006		57 H	27	46
		16NOV2006	1 L	57 H	31	46
		20OCT2006	1 L	72 H	37	43
		22SEP2006	1 L	56 H	26	48
		30AUG2006		58 H	30	46
5001/002	Male/ 32	04SEP2006		41	23	51 H
		11SEP2006		34	20	49
		24OCT2006	5	44	29	49
		25SEP2006	6	53 H	30	48
		27NOV2006	7	51 H	24	47

Subject Profiles for Laboratory Data - Serum Chemistry

Site/ No. Subject No.	Sex/Age (years)	Collection Date	test5	test6	test7	test8
5001/001	Male/ 34	07SEP2006	66		3.6	25
		16NOV2006	59		3.6	24
		20OCT2006	62		4.3	24
		22SEP2006	66		6.1	25
		30AUG2006	66	NEGATIVE	3.9	23
5001/002	Male/ 32	04SEP2006	97	NEGATIVE	4.6	27
		11SEP2006	95		4.3	25
		24OCT2006	86		5	23
		25SEP2006	94		4.6	26
		27NOV2006	98		6.1	25

Below is SAS code without using ID option:

```
proc report
data = final center missing headline headskip nowd split = '|' spacing = 0;
column sitesubj agesex datec _853 _107 _106 _110;
define sitesubj /order order=internal width=9 "Site/|No.|Subject|No." ;
define agesex   /order order=internal width=11 "Sex/Age|(years)" flow ;
define datec    /order order=internal width=12 left "Collection|Date" ;
define _853     /display width=11 "test1" flow ;
define _107     /display width=11 "test2" flow ;
define _106     /display width=11 "test3" flow ;
define _110     /display width=11 "test4" flow ;
break after sitesubj/ skip;
run;

proc report
data = final center missing headline headskip nowd split = '|' spacing = 0;
column sitesubj agesex datec _105 _928 _128 _116;
define sitesubj     /order order=internal width=9 "Site/|No.|Subject|No." ;
define agesex       /order order=internal width=11 "Sex/Age|(years)" flow ;
define datec        /order order=internal width=12 left "Collection|Date" ;
define _105         /display width=11 "test5|" flow ;
define _928         /display width=11 "test6" flow ;
define _128         /display width=11 "test7" flow ;
define _116         /display width=11 "test8" flow ;
break after sitesubj/ skip;
run;
```

Because of page size limitation, we can not display all 8 lab tests in one page. As a result, two PROC REPORT procedures must be used together. SAS will produce a table from the first PROC REPORT, then the result of the second PROC REPORT will be appended at the end of the first one. We need to add more PROC REPORT if we want to display more LAB tests, which is tedious.

Here is more efficient code using ID option:

```
proc report
data = final center missing headline headskip nowd split = '|' spacing = 0;
column sitesubj agesex datec 853 _107 _106 _110 _105 _928 _128 _116;
```

```
define sitesubj /order order=internal id width=9 "Site/|No.|Subject|No." ;
define agesex   /order order=internal id width=11 "Sex/Age|(years)" flow ;
define datec    /order order=internal id width=12 left "Collection|Date" ;
define _853     /display width=11 "test1" flow ;
define _107     /display width=11 "test2" flow ;
define _106     /display width=11 "test3" flow ;
define _110     /display width=11 "test4" flow ;
define _105     /display width=11 "test5|" flow ;
define _928     /display width=11 "test6" flow ;
define _128     /display width=11 "test7" flow ;
define _116     /display width=11 "test8" flow ;
break after sitesubj/ skip;
run;
```

When option ID is used, SAS will print the ID variables repeatedly in each page. As a result, 4 lab tests for each subject are displayed in the first page while the other 4 lab tests are displayed in the next page. By calculating the total width of the ID variables (11 + 12 = 23) and the first 4 test result variables (11 x 4 = 44), we know for sure that the fifth test will be displayed on next page since we already have width 67 while the page size is 70.

Tip 3 – Using #BYVAL title option

Sometimes we need to display certain titles at the beginning of a report, such as 'Treatment', 'Gender' and 'IBS Subtype' in the following listing. One approach to achieving it is to use the COMPUTE statement.

Appendix 3
Weekly IVRS Data

Treatment: X
Gender: Male
IBS Subtype: Diarrhea predominant

Patient No.	Date	Adequate Overall Relief	Complete Overall Relief	Adequate Relief of Most Prominent IBS Symptom	Complete Relief of Most Prominent IBS Symptom
03-002	05DEC2003	Yes	No	No	Not asked since answer to previous question was "no".
	06FEB2004	No	Not asked since answer to previous question was "no".	No	Not asked since answer to previous question was "no".
	07NOV2003	No	Not asked since answer to previous question was "no".	No	Not asked since answer to previous question was "no".

Here is the code using COMPUTE statement:

```
%let _titlen = 2;
%let _title1 = %nrstr(Appendix 3);
%let _title2 = %nrstr(Weekly IVRS Data);

proc report data =final missing headline headskip nowd split='\' spacing=0;
columns trtgrp gender dxdescrp usubjid entdate wq1a wq2 wq3a wq4;
define trtgrp     /order  order=internal noprint;
define gender     /order  order=internal noprint;
```

```
define dxdescrp   /order  order=internal noprint;
define usubjid    /order  width=8    center 'Patient\No.';
define entdate    /order  width=10   center 'Date';
define wq1a       /order  width=15   center 'Adequate\Overall\Relief' ;
define wq2        /order  width=15   left 'Complete\Overall\Relief' f=$wqa. flow;
define wq3a       /order  width=15   center 'Adequate\Relief\of Most\Prominent\IBS
Symptom' ;
define wq4        /order  width=15   left 'Complete\Relief\of Most\Prominent\IBS
Symptom' f=$wqa. flow;
compute before _page_;
   line @1 "Treatment: " trtgrp $12.;
   line @1 "Gender: " gender gender.;
   line @1 "IBS Subtype: " dxdescrp ibscurr.;
   line @1 "&_horzbar";
endcomp;
break after usubjid/skip;
run;
```

A more efficient alternative is to use the #BYVAL title option:

```
%let _titlen = 6;
%let _title1 = %nrstr(Appendix 3);
%let _title2 = %nrstr(Weekly IVRS Data);
%let _title3 = %nrstr(&_left.Treatment: #byval(trtgrp));
%let _title4 = %nrstr(&_left.Gender: #byval(gender));
%let _title5 = %nrstr(&_left.IBS Subtype: #byval(dxdescrp));
%let _title6 = %nrstr(&_horzbar.);

options nobyline ;

proc report data =final missing headline headskip nowd split='\' spacing=0;
columns usubjid entdate wq1a wq2 wq3a wq4;
by trtgrp gender dxdescrp;
format gender gender. dxdescrp ibscurr.;

define usubjid    / order    width=8    center 'Patient\No.';
define entdate    / order    width=10   center 'Date';
define wq1a       / order    width=15   center 'Adequate\Overall\Relief' ;
define wq2        / order    width=15   left 'Complete\Overall\Relief' f=$wqa. flow;
define wq3a       / order    width=15   center 'Adequate\Relief\of Most\Prominent\IBS
Symptom' ;
define wq4        / order    width=15   left 'Complete\Relief\of Most\Prominent\IBS
Symptom' f=$wqa. flow;
break after usubjid/skip;
run;

Three BY variable names (trtgrp gender dxdescrp) and their values are inserted into
titles using the #BYVAL option. They must also be used in the PROC REPORT BY
statement.
```

CONCLUSION

By using FLOW, ID option and #BYVAL title option in PROC REPORT, we can create summary listing and tables in a format we desire.

REFERENCE

SAS Guide to the REPORT Procedure, Usage and Reference, Version 6. Cary, NC: SAS Institute Inc, 1990.

SAS Report Writing: A Programming Approach: Course Notes, Cary, NC: SAS Institute Inc,2001.

CONTACT INFORMATION

Jenny Zhang
ICON Clinical Research, Inc.
1700 Pennbrook Parkway
North Wales, PA 19454
215-616-3142
Email: hui.zhang@iconplc.com