ABSTRACT

When generating tables from SAS® PROC TABULATE or PROC REPORT to summarize data, sometimes it is necessary to produce a cell index as part of the data display. A cell index is simply a data listing that reports the summarized count of each cell shown in the table.

This topic was presented and discussed at NESUG'01 Conference [1]. Several enhancements have since been made resulting in a major modification of the original code. It provides an automated SAS reporting macro to generate a cell index with several new features. The design goal for this program is to provide an efficient, flexible and easy-to-use macro.

The SAS product used in this paper is SAS BASE, with no limitation of operating systems.

INTRODUCTION

SAS System provides several useful tools for the production of tables and tabulations that summarize data. The procedures, PROC REPORT, and PROC TABULATE are very popular.

It is sometimes required to generate a cell index as part of the data display after producing a summarized table. The clinical study reviewer or SAS programmer uses a cell index for program validation purposes. The difference between a data listing and a cell index is that a cell index displays only categorical variables that appear in the table. It displays data observations to support the summary count shown in each cell of the table. The last column of a cell index usually displays the subject variable. It is also a preferred design to have the column before the last as a numeric count of the sub-categorical grouping of the subject variable.

The new and enhanced features are: num_var, freqsort, noprinv, and pageby.

This paper provides a step-by-step approach to developing a SAS reporting module for cell index production.

This paper is comprised of five parts. Part 1 includes the abstract and the introduction. Part 2 describes the sample data used throughout the paper. Part 3 is devoted to the program design and SAS code. Part 4 involves the SAS macro invocation and application. Part 5 concludes the paper.

SAS MACRO DESIGN SPECIFICATION FOR CELL INDEX

The program requirements for this macro are efficient, flexible and easy-to-use. This design goal can be achieved by minimizing the macro arguments. The framework of the program flow is described as follows.

The main code consists of six steps:

STEP 1:

This step accepts the user's input macro argument information and prepares the dataset for the cell index output.

There are seven macro arguments, namely, dsin, sortby, num_var, freqsort, noprinv, and idvar.

STEP 2:

This step performs three counts: one for number of variables in the sortby list, the other for number of subject observations in each sub-category grouping, the last one for the space width available for listing of all variables in the sortby list.

STEP 3:

This step performs specific features for variables freqsort, noprinv, pageby and num_var.

STEP 4:

This step creates all SAS statements related to the procedure PROC REPORT. Then the SAS macro statements create a temporary SAS macro file named ctemp.sas.
STEP 5:
SAS file `ctemp.sas` is run to create a cell index.

STEP 6:
All data files and temporary SAS macro files are deleted after the execution of the SAS macro `get_cell.sas`.

Note: The X statement used in this program is for the UNIX platform. You need to issue your operating system command to delete the temporary file `ctemp.sas`.

Note: The statement

```sas
%let dwidth=%eval(%eval(94/%eval(&j-1)));
```

is for reserved space width for all sortby variables. You can make further adjustments to fit your needs.

SAS CODE FOR CELL INDEX

The whole SAS macro code are shown as follows:

```sas
%macro get_cell(dsin=, sortby=, idvar=, num_var=, freqsort=N, noprintv=, pageby=);
  ***********************************************;
  * read in dataset ;
  ***********************************************;
  data dsin;
    set &dsin;
    count=1;
    proc sort;
      by &sortby;
      %if &pageby ne %then %do;
        proc sort data=dsin nodupkey;
          by &pageby &sortby &idvar;
      %end;
      run;
      ***********************************************;
  * using &sortby for grouping to count number of * objects in each cell
  ***********************************************;
  proc summary data= dsin;
    var count;
    output out=pxyz1 sum=n;
  data dsin;
    merge dsin pxyz1;
  by &pageby &sortby;
  run;
  ***********************************************;
  * covert &sortby list to a column that contains * each variable in &sortby list
  ***********************************************;
  data f1;
    length &sortby 4.;
    %do i=1 %to &sortbyn;
      %scan(&sortby,&i)=1;
    %end;
    run;
    proc transpose data=f1 out=f2(keep=_name_);
      var &sortby;
    data f2;
      set f2;
      if _name_ = %upcase(&noprintv) then delete;
    run;
  ***********************************************;
  * The following 3 datasets f1, f2, and f3, * construct a temp SAS macro to perform * proc report and produce a cell index by * constructing a dataset with define statement * for &sortby argument
  ***********************************************;
  data f2;
    set f2 end=eof;
    var1 = 'define ' || _name_  || "  / order left width=&dwidth flow      ; ";
    output;
    if eof then do;
      %if %upcase(&noprintv) ne %then %do;
        var1 = 'define ' || &noprintv  || "  / order noprint   ; ";
      %end;
      %end;
    %if %upcase(&freqsort)=Y %then %do;
      %let dwidth=%eval(%eval(94/%eval(&j-1)));
    %end;
    %end;
  %if %upcase(&pageby) ne %then %do;
    %let dwidth=%eval(%eval(81/%eval(&j-1)));
  %end;
  %else %do;
    %let dwidth=%eval(%eval(93/%eval(&j-1)));
  %end;
  %global sortbyn dwidth;
  %let sortbyn=%eval(&j-1);
  %if &num_var ne %then %do;
    %let dwidth=%eval(%eval(81/%eval(&j-1)));
  %end;
  %if &noprintv ne %then %do;
    %let dwidth=%eval(%eval(93/%eval(&j-1)));
  %end;
  %mend varn;
  %varn;
  ***********************************************;
  * The following 3 datasets f1, f2, and f3, * construct a temp SAS macro to perform * proc report and produce a cell index by * constructing a dataset with define statement * for &sortby argument
  ***********************************************;
  data f2;
    set f2 end=eof;
    var1 = 'define ' || _name_  || "  / order left width=&dwidth flow      ; ";
    output;
    if eof then do;
      %if %upcase(&pageby) ne %then %do;
        var1 = 'define ' || &pageby  || "  / order left width=&dwidth flow      ; ";
      %end;
      %end;
  %end;
  %if %upcase(&freqsort)=Y %then %do;
    %let dwidth=%eval(%eval(94/%eval(&j-1)));
  %end;
  %else %do;
    %let dwidth=%eval(%eval(93/%eval(&j-1)));
  %end;
  %global sortbyn dwidth;
  %let sortbyn=%eval(&j-1);
  %if &num_var ne %then %do;
    %let dwidth=%eval(%eval(81/%eval(&j-1)));
  %end;
  %if &noprintv ne %then %do;
    %let dwidth=%eval(%eval(93/%eval(&j-1)));
  %end;
  %mend varn;
  %varn;
```
run;
*****************************************************************************
* construct a dataset with define statement for
* variable n and &idvar;
*****************************************************************************
data f3;
  length var1 $90.;
  var1='define n / order width=5 "N";
  output;
  var1='define &idvar / display width=15 flow
  "";
  output;
  %if &num_var ne %then %do;
  var1='define &num_var / display right
  width=13 "";
  output;
  %end;
  var1='break after %scan(&sortby, &sortbyn )
  /skip;"
  output;
  var1='run;"
  output;
  var1='%mend ctemp;"
  output;
run;
*****************************************************************************
* construct a dataset with macro statement and
* proc report statement
*****************************************************************************
data f1;
  length var1 $90.;
  var1=' %macro ctemp; 
  output;
  var1='proc report data=dsin headline
  headskip split="~";
  output;
  %if %upcase(&pageby) ne %then %do;
  var1='by &pageby;"
  output;
  %end;
  %if %upcase(&freqsort) eq Y %then %do;
  var1='column ("==" orderv &sortby n &num_var &idvar);"
  output;
  %end;
  %else %if %upcase(&freqsort) ^= Y %then 
  %do;
  var1='column ("==" &sortby n &num_var &idvar);"
  output;
  %end;
run;
data f2(keep=var1);
  set f1 f2 f3; run;
*****************************************************************************
* macro call ctemp.sas to produce cell index
*****************************************************************************
data _null_;
  set f2;
  file 'ctemp.sas';
  put var1;
run;
%inc 'ctemp.sas'; run;
%ctemp; run;
*****************************************************************************
* remove temp SAS macro ctemp.sas and clean up
* datasets
*****************************************************************************
x 'rm ctemp.sas'; run;
proc datasets nolist;
  delete pxyz1 f1 f2 f3 dsin dsinn;
run;
%mend get_cell;

SAS MACRO INVOCATION AND APPLICATION

Before this macro invocation, it is important that you use
the DATA STEP to prepare the SAS data set and label for
variables to appear in the cell index. In this section,
individual an invocation example with code and a figure
showing the relationship between table and cell index are
provided. The sample macro invocation example 1 is as
follows:

\%get_cell(dsin = final,
  sortby = trt agec sex,
  idvar = patient);

Note: This macro accepts categorical variables in
the sortby argument. If you have numeric
variables in the sortby list, you need to convert
them to character variables before invoking the macro.

A summary table displays the treatment variable as the
across variable with categorical variables of age group and
sex. The output of summary table and cell index from this
example 1 is shown in Figure 1.

![Figure 1 Output of Summary Table and Cell Index from Example 1](image)

Example 2 uses the feature num_var, to add additional
variable to the cell index, either numeric or character
variable, and specifies it as age. The output from this
example is shown in Figure 2.

```sas
%get_cell(dsin = final,
          sortby = trt agec sex,
          num_var = age,
          idvar = patient);
```

**Figure 2** Cell Index with Additional Variable
num_var Specified

Sometimes the output table contains a certain sort order for
the key variable. It is desired that the cell index contains
the sort order for the same variable. You can create a 'no
print' variable that controls the desired sort order without
printing it on the output. Example 3 is a sample of this type
of output.

```sas
%get_cell(dsin = final,
          sortby = trt_o trt agec sex,
          num_var = age,
          idvar = patient);
          noprintv = trt_o,
```

**Figure 3** Cell Index with noprintv Specified

Sometimes the output table program contains a by
statement that output is breaking by this 'pageby' variable.
It is desired that the cell index contains the same feature to
break the output by 'pageby' variable. Example 4 is a
sample of this type of output.

```sas
%get_cell(dsin = lab,
          sortby = visit range,
          pageby = labtest trt,
          idvar = pid);
```

**Figure 4.** Cell Index with pageby Specified
Sometimes the output table contains a certain descending sort order for the key variable. It is desired that the cell index contains the descending sort order for the frequency count N. Example 5 is a sample of this type of output.

%get_cell(dsin = aesub, sortby = pref trt, freqsort = Y, idvar = pid);

CONCLUSION

This paper takes a comprehensive approach in designing a reporting module for creation of a cell index. SAS code and the examples of how to invoke the SAS macro are provided in this paper. This paper achieves the following goals:

* Provides an easy-to-use macro with only seven macro arguments
* Provides enhanced features to fit the different needs
* Provides a compact, efficient, and simple code for cell index production.

ACKNOWLEDGMENTS

The author would like to thank Ellen Forman, Shu-Mei Huang, Biostatistics & Data Management, GlaxoSmithKline, and Ju Zhang from EDP Contract Services, for their review, validation and enhancement suggestions of this macro.

REFERENCES


SAS is a registered trademark of SAS Institute Inc., Cary, NC, USA
© indicates USA registration.

Author
Shi-Tao Yeh, Ph. D.
(610)917-5883(W)
E-mail: shi-tao_yeh-1@gsk.com