A Productive Word Wrap Macro

Monte Jarvis, Genzyme Inc
Frank Liu, Genzyme Inc

INTRODUCTION:

If a data set contains a large number of variables, and one needs to print full observations with some wrapped variables on each page, PROC SQL and PROC REPORT cannot handle the job although they both have flow options. This paper presents a productive SAS macro using a Split-Merge approach to wrap as many long text variables as required. Generally the wrap macro can be implement before using PROC PRINT. The output is neat and readable.

THE SPLIT-MERGE APPROACH

When a long text variable needs to be wrapped, normally, this variable can be split into lines based on a given output length, and then the split text can be written out to a SAS data set. However, the SAS data set will include the duplicated value of non-wrapped variables. To avoid this duplication, you can split an input data set into two output data sets. The first data set contains the first split line of the long text variables. The second data set contains the remaining split lines of the long text variables and other variables. Because both data sets also include a runtime variable about the order of observation, therefore, merging these two data sets by the variable will avoid the duplicated value of non-wrapped variable.

THE WRAP MACRO: WORDWRAP

WORDWRAP is invoked with four keyword parameters. These keywords, used throughout the macro in the form of macro variables, enable the user to achieve a great deal of flexibility from using this productive macro: one can choose any text variables for wrapping, and one can wrap as many long text variables as required.

Here is the text of the macro with a detailed explanation:

```
* **********************************************;
* * Parameters Specifications               *
* * --------------------------------------- *
* * data Input data set name.              *
* * id  ID variable name. ID               *
* * variable can be one or more,           *
* * such as, id=id1, id2, id3...          *
* * var Long text variable name.           *
* * This var can be one or more,           *
* * such as, var=var1 var2 var3...         *
* * len Width of each line of              *
* * wrapped variables. Len                 *
* * contains different lengths             *
* * corresponding to each long             *
* * text variable, such as,                *
* * len=len1 len2 len3...                  *
* * **********************************************;

%MACRO WORDWRAP
(data=, id=, var=, len=);

* **********************************************;
* * Description of Macro Variables           *
* * ---------------------------------------- *
* * _n  Number of the ID variable.          *
* * _m  Number of the long text             *
* *   variables.                            *
* * _com Hold _COM1, _COM2, _COM3...        *
* * _blk Hold _BLK1, _BLK2, _BLK3...         *
* * **********************************************;

%let _n=1;
%let _m=1;
%let _com=
%let _blk=
%do %until (%scan(&id,&_n)=);
  %let id&_n=%scan(&id,&_n);
  %let _n=%eval(&_n+1);
%end;
%let _n=%eval(&_n-1);
  /* Get the number of the id variables */
```
Posters

%do %until (%scan(&var,&_m)=);
%let var&_m=%scan(&var, &_m);
/* Get var1, var2, var3 ... */
%let len&_m=%scan(&length,&_m);
/* Get len1, len2, len3 ... */
%let _m=%eval(_m+1);
end;

%let _m=%eval(_m-1);
/* Get the number of the long text variables */
%do _i=_m %to 1 %by -1;
%let _com=COM&_i &_com;
%let _blk=BLK&_i &_blk;
end;

******************************************************************************;
* * SUB1 contains the first split line *;
* of long text variables. *;
* * SUB2 contains the remaining split *;
* lines of long text variables. *;
* * Description of Variables *;
* ------------------------- *
* * _COM1, _COM2, ... *;
* Store the remaining text of *;
* the long text variables after *;
* split line. *
* _BLK1, _BLK2 *;
* Word break point at end line. *
* CTR Counter for observation to be *
* Written to a SAS data set. *
* _L split line number of the long *
* text varables. *
******************************************************************************;

DATA SUB1(DROP=_com &_blk _J _L)
SUB2(KEEP=ctl &var CTR);
LENGTH &_com $200;
RETAIN CTR 0;
SET &data;
_L=1;
RUN;

******************************************************************************;
* * PREWRAP: *;
* ----- *
* 1. Determines whether the long *;
* text variable need to be *;
* wrapped or not. *
* 2. _L =1 indicates that the all *;
* long text variables of current *;
* observation do not need to be *;
* wrapped.
******************************************************************************;

* 3. _L > 1 indicates some or all *;
* long text variables have been *;
* successfully wrapped.Last line *;
* of long text variables remains.*;
******************************************************************************;

PREWRAP:

%do _i=1 %to &_m:
 _COM&_i=&var&_i;
end;
%do _i=1 %to &_m:
 IF length(&&var&_i)>&len&_i THEN
 GOTO WRAP;
 ELSE
 IF &_i=&_m AND _L=1 THEN
 DO;
 CTR+1;
 OUTPUT SUB1;
 END;
 ELSE
 IF &_i=&_m AND _L>1 THEN
 DO;
 CTR+1;
 OUTPUT SUB2;
 END;
 %end;
 RETURN;

******************************************************************************;
* * WRAP: *;
* ----- *
* A wrap cycle starts with scanning *;
* the long text variable letter by *;
* letter from right to left until *
* word break point is found at the *
* end of line. *
* * *
******************************************************************************;

WRAP:

%do _i=1 %to &_m:
 _BLK&_i=0;
 IF LENGTH(&&COM&_i)>&len&_i THEN
 DO;
 _J=&len&_i TO 1 BY -1;
 IF SUBSTR(_COM&_i, _J, 1)=" ">
/* Look backwards for a blank until a blank is found */
 THEN
 DO;
 _BLK&_i=_J;
 GOTO LOOP&_i;
 END;
 END;
 %end;
LOOP&_i:

%end;
%do _i=1 %to &m;
   IF _BLK&_i ne 0 THEN
   &var&_i=SUBSTR(_COM&_i,1,_BLK&_i);
%end;
IF _L=1 then
   DO;
      CTR+1;
      OUTPUT SUB1;
   END;
ELSE
   DO;
      CTR+1;
      OUTPUT SUB2;
      /* if _blk ne 0, then you are in the middle of the word */
   END;
   _L+1;
%do _i=1 %to &m;
   IF _BLK&_i NE 0 THEN
   &var&_i=SUBSTR(_COM&_i, _BLK&_i+1);
   ELSE
      &var&_i= ',
%end;
GOTO PREWRAP;

DATA &data;
   SET SUB1 SUB2;
   BY CTR &id;
RUN;
%MEND WORDWRAP;

REFERENCES:

CONTACT INFORMATION:
Monte Jarvis
Biostatistics & Medical Data Processing
Genzyme, Inc.
One Kendall Square
Cambridge, MA 02139

Tel: (617)252-7693
E-mail: mejarvis@world.std.com

Frank Liu
Biostatistics & Medical Data Processing
Genzyme, Inc.
One Kendall Square
Cambridge, MA 02139

Tel: (617)374-7262
E-mail: fliu@world.std.com

SAS is a registered trademark of SAS Institute Inc., in the USA and other countries.

LIBNAME SSD '[PROJECT.STUDY.SSD]';

DATA A;
   SET SSD.ADVERSE;
   BY PATID;
   IF FIRST.PATID THEN AENUM=1;
   ELSE AENUM+1;
RUN;
### Appendix: An Example of Word-Wrapped Data Listing

#### Genzyme Corporation
CAMBRIDGE, MASS.

**STUDY NUMBER**
PRODUCT DESCRIPTION
AE 14:56/18DEC95

**ADVERSE EVENTS**

<table>
<thead>
<tr>
<th>PATID</th>
<th>AE ID</th>
<th>PRESENT AT BASELINE</th>
<th>AE DESCRIPTION</th>
<th>HISTORIO OF EVENT</th>
<th>SEVERITY</th>
<th>START DATE</th>
<th>STOP DATE</th>
<th>START TIME</th>
<th>STOP TIME</th>
<th>TYPE OF EPISODE</th>
<th>DUR (MINS)</th>
<th>DUR (HRS)</th>
<th>DUR (DAYS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>901</td>
<td>1</td>
<td>Y</td>
<td>WITHDRAWAL SEX</td>
<td>N</td>
<td>MILD</td>
<td>14MAY94</td>
<td>25MAY94</td>
<td>0:00:00</td>
<td>12:00:00</td>
<td>INTERMITTENT</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>902</td>
<td>1</td>
<td>N</td>
<td>NAUSEA</td>
<td>N</td>
<td>MILD</td>
<td>16MAY94</td>
<td>17MAY94</td>
<td>0:00:00</td>
<td>12:00:00</td>
<td>INTERMITTENT</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>903</td>
<td>2</td>
<td>Y</td>
<td>FEVER</td>
<td>Y</td>
<td>MILD</td>
<td>24MAY94</td>
<td>25MAY94</td>
<td>0:00:00</td>
<td>12:00:00</td>
<td>INTERMITTENT</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>904</td>
<td>3</td>
<td>N</td>
<td>SHAKING CHILLS</td>
<td>N</td>
<td>MILD</td>
<td>25MAY94</td>
<td>25MAY94</td>
<td>0:00:00</td>
<td>12:00:00</td>
<td>INTERMITTENT</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>904</td>
<td>1</td>
<td>N</td>
<td>BACK ITCHING RASH</td>
<td>N</td>
<td>MILD</td>
<td>11JUN94</td>
<td>11JUN94</td>
<td>0:00:00</td>
<td>12:00:00</td>
<td>SINGLE</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PATID</th>
<th>ACTION TAKEN</th>
<th>OTHER ACTION TO PRODUCT</th>
<th>RELATIONSHIP TO PRODUCT</th>
<th>OUTCOME</th>
<th>AE COMMENT</th>
<th>-COSTART-</th>
<th>-COSTART-</th>
<th>-COSTART-</th>
</tr>
</thead>
<tbody>
<tr>
<td>901</td>
<td>NONE</td>
<td>Y</td>
<td>NO</td>
<td>Res</td>
<td>PATIENT RECEIVED</td>
<td>WRITING</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NONE</td>
<td>Y</td>
<td>NO</td>
<td>Rec</td>
<td>TREATED WITH TYLENOL</td>
<td>FEVER</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NONE</td>
<td>Y</td>
<td>NO</td>
<td>Rec</td>
<td>PATIENT MONITORED,</td>
<td>NAUSEA</td>
<td>GEN BODY</td>
<td></td>
</tr>
<tr>
<td>902</td>
<td>NONE</td>
<td>Y</td>
<td>NO</td>
<td>Rec</td>
<td>LAXATIVE CONSID</td>
<td>CV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>903</td>
<td>NONE</td>
<td>Y</td>
<td>NO</td>
<td>Rec</td>
<td>ANTIABOTIC CONSID</td>
<td>CV</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NONE</td>
<td>Y</td>
<td>NO</td>
<td>Rec</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NONE</td>
<td>Y</td>
<td>NO</td>
<td>Rec</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>904</td>
<td>NONE</td>
<td>Y</td>
<td>NO</td>
<td>Rec</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**AE = ADVERSE EVENT**

**NA = NOT AVAILABLE**

**ND = NOT DONE**

**-99 = MISSING NUMERIC VALUE**

**N/A = NOT APPLICABLE**

**UNK = UNKNOWN**

**01JAN01 = MISSING/PARTIAL DATE**