A Technique for Producing Sorted Columns with a Hanging First Line Using PROC REPORT

David C. Izard, GlaxoSmithKline, Collegeville, PA
David C. Chen, GlaxoSmithKline, Collegeville, PA

Abstract

SAS® PROC REPORT does an excellent job allowing you to sort columns, order columns and even keep repeated values of columns from printing unnecessarily when organized in a vertical format. One challenge, however, is to “convince” PROC REPORT to work effectively with multiple variables appearing in a single column while still maintaining the identity of the original data. This paper discusses how you can influence SAS to produce a section of text with what appears to be a hanging first line, also known as an out-dented first line, and a business situation for using this technique effectively to present adverse event data in a coherent fashion.

Business Case

A common report requested in the Pharmaceutical arena is a listing of adverse events. There is a lot of information to fit into this report, including the verbatim term captured on the case report form, the preferred term as identified by processing the verbatim term through some type of coding dictionary such as COSTART or MedDRA, and perhaps a modified or enhanced term, usually produced by the Clinical Data Management or Clinical Monitoring function in conjunction with the Investigator in order to facilitate successful coding by the coding dictionary. All three of these terms are desired on the listing of adverse events, but there is not much real estate left on the report if you include all three of these terms as separate columns even if you allow the text to wrap.

Proposed Solution

The solution would be to produce a report that would deliver these three terms in a single column with easy to read text that clearly identifies the three different terms. Bonus features might be to present this information in a clear fashion but still having underlying control over the sort order of these terms within the combined column. A proposed solution could also include the concept of a hanging indent, keeping a notation of some kind out to the left on the first line, visually isolating the text that identifies the type of term from the actual term itself.

Technical Details

The solution involves tricking SAS into producing what visually appears to be a single column by actually displaying two columns with no space between them and taking advantage of PROC REPORT options to control the appearance. Along with these two display variables would be an underlying sort order variable that would not be displayed. The final detail is to add labels and printer control variables, specifically the definition and use of the split character within PROC REPORT, to make the titles of the two adjacent variables blend imperceptibly to produce a coherent combined title.
Specific Steps of the Solution

1. Create an analysis dataset of all variables for a single event on a single row in the SAS dataset.

-----Alphabetic List of Variables and Attributes-----

<table>
<thead>
<tr>
<th>#</th>
<th>Variable</th>
<th>Type</th>
<th>Len</th>
<th>Pos</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>BODY1</td>
<td>Char</td>
<td>120</td>
<td>75</td>
<td>MEDDRA System Organ Class</td>
</tr>
<tr>
<td>5</td>
<td>COMMTERM</td>
<td>Char</td>
<td>100</td>
<td>295</td>
<td>MEDDRA Lower Level Dictionary Term</td>
</tr>
<tr>
<td>6</td>
<td>INTENS</td>
<td>Char</td>
<td>21</td>
<td>395</td>
<td>AE: Intensity</td>
</tr>
<tr>
<td>4</td>
<td>PREF</td>
<td>Char</td>
<td>100</td>
<td>195</td>
<td>MEDDRA Preferred Term</td>
</tr>
<tr>
<td>2</td>
<td>RAND</td>
<td>Char</td>
<td>70</td>
<td>5</td>
<td>Randomization</td>
</tr>
<tr>
<td>7</td>
<td>SER</td>
<td>Char</td>
<td>14</td>
<td>416</td>
<td>AE: Serious</td>
</tr>
<tr>
<td>1</td>
<td>SUBJID</td>
<td>Char</td>
<td>5</td>
<td>0</td>
<td>Subject ID</td>
</tr>
<tr>
<td>8</td>
<td>VERBATIM</td>
<td>Char</td>
<td>200</td>
<td>430</td>
<td>AE: Verbatim Text</td>
</tr>
<tr>
<td>9</td>
<td>WITHAE</td>
<td>Char</td>
<td>14</td>
<td>630</td>
<td>AE: Withdrawal</td>
</tr>
</tbody>
</table>

The SAS System

2

14:59 Monday, February 21, 2005

Obs   SUBJID     RAND      BODY1
1    19732     PLACEBO    INFECTIONS AND INFESTATIONS
2    19732     PLACEBO    MUSCULOSKELETAL AND CONNECTIVE TISSUE DISORDERS
3    19732     PLACEBO    NERVOUS SYSTEM DISORDERS
4    20263     PLACEBO    CARDIAC DISORDERS
5    21257     ACTIVE     GASTROINTESTINAL DISORDERS

Obs   PREF                               COMMTERM                     INTENS
1    GASTROENTERITIS                    GASTROENTERITIS              Mild
2    MUSCULOSKELETAL CHEST PAIN         MUSCULOSKELETAL CHEST PAIN   Mild
3    HEADACHE                           HEADACHE                     Mild
4    MYOCARDIAL ISCHAEMIA               CARDIAC ISCHEMIA             Severe
5    GASTROESOPHAGEAL REFLUX DISEASE   GASTROESOPHAGEAL REFLUX       Moderate

Obs   SER    VERBATIM                                          WITHAE
1    No     DIARRHEA/ABDOMINAL PAIN                            No
2    No     PAIN RIGHT THORAX DUE TO MUSCLE PAIN               No
3    No     HEADACHES                                         No
4    Yes    CARDIAC ISCHEMIA REVEALED BY STRESS ECHOGRAPHY     Yes
5    No     REFLUX SYMPTOMS                                    No

2. Create a new dataset with four new variables: two display variables; one sort variable; and a sort/display variable. The first display variable is a prefix indicating the type of adverse event term – the verbatim, modified or preferred term. The second display variable is the actual adverse event term associated with the adverse event id. The sort order variable is also based on the adverse event id, controlling the order in which the terms are presented. The sort/display variable, cntr, will be used to add a blank line between adverse events in the final report.
data work.ae(keep=cntr subjid rand body1 term_o term_id term
intens ser withae);

set outmart.psug2005;

attrib term_o length= 8 label="AE: Term Order"
      term_id length= $ 2 label="AE: Term ID"
      term length= $ 200 label="AE: Term Variable"
      cntr length= 8 label="AE: Unique Adverse Event Counter";

cntr = _n_;             /* variable used to add a blank line between   */
      /* adverse events in the PROC REPORT output   */
term_o = 1;            /* assign variables, produce first observation */
term_id = "P:";         /* in output dataset                           */
term = pref;
output work.ae;

intens = '';           /* clear out assignments to display variables */
ser = '';             /* on other rows in order to produce a cleaner  */
withae = '';          /* report                                      */
term_o = 2;            /* assign variables, produce second observation*/
term_id = "M:";         /* in output dataset                           */
term = commterm;
output work.ae;

intens = '';           /* clear out assignments to display variables */
ser = '';             /* on other rows in order to produce a cleaner  */
withae = '';          /* report                                      */
term_o = 3;            /* assign variables, produce third observation */
term_id = "V:";         /* in output dataset                           */
term = verbatim;
output work.ae;

return;                 /* explicit return                              */
run;

During the creation of this dataset two key actions occur: (1) the individual adverse event
term variables are dropped from the dataset; and (2) for each observation in the input
dataset, three observations are written to the output dataset. The resulting dataset appears here:

Obs  SUBJID  RAND   BODY1                                           INTENS   SER
  1   19732  PLACEBO INFECTIONS AND INFESTATIONS                     Mild     No
  2   19732  PLACEBO INFECTIONS AND INFESTATIONS
  3   19732  PLACEBO INFECTIONS AND INFESTATIONS
  4   19732  PLACEBO MUSCULOSKELETAL AND CONNECTIVE TISSUE DISORDERS Mild     No
  5   19732  PLACEBO MUSCULOSKELETAL AND CONNECTIVE TISSUE DISORDERS
  6   19732  PLACEBO MUSCULOSKELETAL AND CONNECTIVE TISSUE DISORDERS
  7   19732  PLACEBO NERVOUS SYSTEM DISORDERS                         Mild     No
  8   19732  PLACEBO NERVOUS SYSTEM DISORDERS
  9   19732  PLACEBO NERVOUS SYSTEM DISORDERS
 10   20263  PLACEBO CARDIAC DISORDERS                                 Severe   Yes
 11   20263  PLACEBO CARDIAC DISORDERS
 12   20263  PLACEBO CARDIAC DISORDERS
 13   21257  ACTIVE  GASTROINTESTINAL DISORDERS                       Moderate No
 14   21257  ACTIVE  GASTROINTESTINAL DISORDERS
 15   21257  ACTIVE  GASTROINTESTINAL DISORDERS
<table>
<thead>
<tr>
<th>Obs</th>
<th>WITHAE</th>
<th>TERM_O</th>
<th>TERM_ID</th>
<th>TERM</th>
<th>CNTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
<td>1</td>
<td>P:</td>
<td>GASTROENTERITIS</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>2</td>
<td>M:</td>
<td>GASTROENTERITIS</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>3</td>
<td>V:</td>
<td>DIARRHEA/ABDOMINAL PAIN</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>1</td>
<td>P:</td>
<td>MUSCULOSKELETAL CHEST PAIN</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>No</td>
<td>2</td>
<td>M:</td>
<td>MUSCULOSKELETAL CHEST PAIN</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>No</td>
<td>3</td>
<td>V:</td>
<td>PAIN RIGHT THORAX DUE TO MUSCLE PAIN</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Yes</td>
<td>1</td>
<td>P:</td>
<td>HEADACHE</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>No</td>
<td>2</td>
<td>M:</td>
<td>HEADACHE</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>No</td>
<td>3</td>
<td>V:</td>
<td>HEADACHES</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>Yes</td>
<td>1</td>
<td>P:</td>
<td>MYOCARDIAL ISCHAEMIA</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>No</td>
<td>2</td>
<td>M:</td>
<td>CARDIAC ISCHEMIA</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>No</td>
<td>3</td>
<td>V:</td>
<td>CARDIAC ISCHEMIA REVEALED BY STRESS ECHOGRAPHY</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>No</td>
<td>1</td>
<td>P:</td>
<td>GASTROESOPHAGEAL REFLUX DISEASE</td>
<td>5</td>
</tr>
<tr>
<td>14</td>
<td>No</td>
<td>2</td>
<td>M:</td>
<td>GASTROESOPHAGEAL REFLUX</td>
<td>5</td>
</tr>
<tr>
<td>15</td>
<td>No</td>
<td>3</td>
<td>V:</td>
<td>REFLUX SYMPTOMS</td>
<td>5</td>
</tr>
</tbody>
</table>

3. Call PROC REPORT to produce the listing. Note specific text in **bold red** that allows for the effective use of this dataset organization to produce a table with the desired result.

```informal
proc report data=work.ae nowd missing headline headskip spacing=2
split="#";

column subjid rand body1 cntr term_o term_id term intens ser withae;

define subjid / order order=internal left width=7 "SUBJECT#ID";
define rand   / order order=internal left width=10 "TREATMENT#GROUP";
define body1   / order order=internal left width=50 flow "MedDRA System#Organ Class";
define cntr    / order order=internal noprint;
define term_o  / order order=internal noprint;
define term_id  / left width=3 "ADV#EXP";
define term    / left width=20 spacing=0 flow "ERSE#ERIENCE";
define intens  / left width=10 "INTENSITY";
define ser     / left width=10 "SERIOUS#AE?";
define withae  / left width=10 "WITHDRAWN#DUE TO AE?";

break after cntr / skip;
run;
```

Key statements in this call to PROC REPORT that influence the production of the desired hanging first line within the adverse event term variable are:

- “spacing=0” on the DEFINE statement for the variable `term` to allow the two variables to appear as a single column in the report
- The labels on the DEFINE statement for the variables `term_id` and `term` that appear to be incorrect when viewed individually but make sense when placed next to each other
- The `flow` option on the DEFINE statement for the variable `term` which ultimately produced the hanging first line
### Result

Using the dataset described in Step 1 of the Specific Steps above, the PROC REPORT call in Step 5 of the Specific Steps above would produce the following report

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>TREATMENT</th>
<th>MedDRA System</th>
<th>ADVERSE EXPERIENCE</th>
<th>SERIOUS</th>
<th>WITHDRAWN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>GROUP</td>
<td>Organ Class</td>
<td>INTENSITY</td>
<td>AE?</td>
<td>DUE TO AE?</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td>---------------</td>
<td>-----------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>19732</td>
<td>PLACEBO</td>
<td>INFECTIONS AND INFESTATIONS</td>
<td>P: GASTROENTERITIS</td>
<td>Mild</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>M: GASTROENTERITIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>V: DIARRHEA/ABDOMINAL PAIN</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MUSCULOSKELETAL AND CONNECTIVE TISSUE DISORDERS</td>
<td>P: MUSCULOSKELETAL CHEST PAIN</td>
<td>Mild</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>M: MUSCULOSKELETAL CHEST PAIN</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>V: PAIN RIGHT THORAX DUE TO MUSCLE PAIN</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NERVOUS SYSTEM DISORDERS</td>
<td>P: HEADACHE</td>
<td>Mild</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>M: HEADACHE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>V: HEADACHES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20263</td>
<td>PLACEBO</td>
<td>CARDIAC DISORDERS</td>
<td>P: MYOCARDIAL ISCHAEMIA</td>
<td>Severe</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>M: CARDIAC ISCHEMIA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>V: CARDIAC ISCHEMIA REVEALED BY STRESS ECHOGRAPHY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21257</td>
<td>ACTIVE</td>
<td>GASTROINTESTINAL DISORDERS</td>
<td>P: GASTROESOPHAGEAL REFLUX DISEASE</td>
<td>Moderate</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>M: GASTROESOPHAGEAL REFLUX</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>V: REFLUX SYMPTOMS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusion

There are many ways one can pursue reclaiming physical space on the printed page. This paper offers one solution on how to combine similar yet distinct fields into a single location on a report while still maintaining readability. This solution exploits features of PROC REPORT as well as intelligent dataset design to arrive at a simple yet elegant solution.

Acknowledgements

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.

Contact Information

David C. Izard
GlaxoSmithKline, Mail Stop UP4310
1250 South Collegeville Road
Collegeville, PA 19426
Phone: 610.917.6911
Fax: 610.917.4701
Email: david.c.izard@gsk.com

David C. Chen
GlaxoSmithKline, Mail Stop RN0420
2301 Renaissance Boulevard
King of Prussia, PA 19406
Phone: 610.787.3828
Fax: 610.787.7007
Email: david.c.chen@gsk.com