Paper 181-30

Creating bar charts and scatter plots using the SAS® GPLOT procedure. Both on the same page? It is easier than you think!

Rita Tsang, Averion Inc., Framingham, Massachusetts

ABSTRACT

The SAS GPLOT and GCHART procedures are powerful tools for generating graphical displays of clinical data. These two procedures, GPLOT and GCHART, are usually employed to create scatter plots and bar charts respectively. However, we often need to display both a scatterplot and a bar chart within the same graph. The GCHART procedure cannot generate scatter plots and it does not have the ability to superimpose graphical displays. How can we get around this limitation?

This paper will delineate a program that generates a graphical display of a scatter plot superimposed on a bar chart by using the GPLOT procedure, which proves generating complicated graphical displays in SAS can be done by using simple Data steps!

INTRODUCTION

Using the GPLOT procedure to create bar charts is not unprecedented. The past approach, however, requires the presence of pre-defined groups of data on the x-axis before the four corners of the y-value can be created for plotting (Carpenter and Shipp, 1995). This approach is inadequate to handle safety monitoring data over time. In our clinical data, the vertical bars on the x-axis are not represented by distinct groups. They denote safety lab data for an individual patient collected at different assessments. An alternative approach to create bar charts in GPLOT is to employ the annotate functions in the annotate data set (SAS Institute, 1999). This approach however becomes cumbersome when the values on the x-axis are clustered together.

Constrained by our data, we have come up with a solution to create the vertical bars by a simple data manipulation and by employing the GPLOT SYMBOL statement. With the GPLOT procedure, we can superimpose a scatter plot on the bar chart with a second plot statement.

THE DATA

Our task at hand is to display the urine output data (Urine__) in a vertical bar chart and plasma creatinine data (Labval) in a scatter plot by assessment date (Asdt) for each patient. Here is a sample display of the data.

```
Site   Randno__        Asdt          Labval        Urine__
3      1101   14MAR2004:20:00:00                   27
3      1101   15MAR2004:04:00:00                   595
3      1101   15MAR2004:04:20:00  1.4
3      1101   15MAR2004:12:00:00                   329
3      1101   15MAR2004:14:45:00                   29
3      1101   15MAR2004:16:45:00                   24
3      1101   15MAR2004:18:45:00                   6
3      1101   15MAR2004:20:45:00  1.1              48
3      1101   15MAR2004:22:45:00                   53
3      1101   16MAR2004:00:45:00                   53
3      1101   19MAR2004:03:00:00  1.2
```

CREATING A BAR-CHART IN GPLOT: A SIMPLE DATA MANIPULATION

The following DATA step is a simple manipulation of the clinical data to create additional records of zero value for urine output (Urine1) and missing value for plasma creatinine (Labval) at each assessment time point. By creating the extra records, we can create vertical bars for the urine output data at each assessment time point by linking the extreme values, zero and the original values, using the INTERPOL=BOX00F option in the GPLOT SYMBOL statement. With the plasma creatinine values set to missing, SAS GPLOT simply will not display those observations.

```
data urine__;
set urine__;
by site randno__ asdt;
urine1=urine__;
output;
if urine__>.then do;
urine1=0;
labval=.;
output;
end;
run;
```

Here is a sample display of the data after the DATA step:

```
Site   Randno__        Asdt           Labval        Urine1
3       1101    14MAR2004:20:00:00                  27
3       1101    14MAR2004:20:00:00                  0
3       1101    15MAR2004:04:00:00                  595
3       1101    15MAR2004:04:00:00                  0
3       1101    15MAR2004:04:20:00   1.4
3       1101    15MAR2004:12:00:00                  329
3       1101    15MAR2004:12:00:00                  0
```

CREATING A BAR-CHART IN GPLOT: THE SYMBOL STATEMENT
In PROC GPLOT, we can produce box and whisker plots by specifying INTERPOL=BOX in the SYMBOL statement
(SAS Institute, 1999). By varying the options in the INTERPOL=BOX <options>, the bottom and top edges of the box
can be redrawn at a desired location. For example, INTERPOL=BOX00 will display a box ranges from the high and
low extremes without displaying the whiskers.

Borrowing the idea of box plots, we can employ the following SYMBOL statement for the urine output, to link the
original urine output value (high extreme) with the newly created records of zero value (low extreme). This SYMBOL
statement will connect the high and low extreme values on the x-axis (I=BOX00F) with no plot symbol (V=NONE).
The I=BOX00F option will fill the box with the color black (CV=BLACK). A black outline of the box will also be drawn
(CO=BLACK). The BIWIDTH option specifies the width of the box. The width is about 0.4 cm in the current graphical
display.

```
symbol1 v=none i=box00f co=black cv=black bwidth=2;
```

CREATING A SCATTER-PLOT SUPERIMPOSED ON THE BAR-CHART: A SECOND PLOT STATEMENT
The SYMBOL statement for plasma creatinine is defined with the letter P as the plot symbol. The symbols will be
joined by a black line in a scatter plot.

```
symbol2 h=0.17 in v='P' cv=black i=join l=1;
```

The following axis statements are defined for the assessment date (horizontal axis), urine output (left vertical axis),
and plasma creatinine (right vertical axis) respectively.

```
axis1 label=(h=0.17 in f=triplex "Assessment date and time") value=(h=0.135 in
f=triplex) offset=(1,1);

axis2 label=(h=0.17 in a=90 f=triplex "Urine Output (CC)") minor=none order=(0
to 2100 by 100) offset=(5 pct) value=(h=0.135 in f=triplex);

axis3 label=(h=0.17 in a=270 f=triplex "Plasma Creatinine (mg/dl)") minor=none
order=(0 to 2 by 1) offset=(5 pct) value=(h=0.135 in f=triplex);
```

By employing two PLOT statements in GPLOT, we can superimpose a scatter plot of plasma creatinine on the bar
chart of urine output (SAS Institute, 1999). The first PLOT statement (PLOT URINE1*ASDT) displays the bar chart of
urine output by assessment date. The OVERLAY option allows the second PLOT statement to be superimposed on
the bar chart. The second PLOT statement (PLOT2 LABVAL*ASDT) displays a scatter plot of plasma creatinine by
assessment date. Figure 12.1 shows the final graphical display.

```
proc gplot data=urine__;
plot urine1*asdt/ overlay noframe nolegend haxis=axis1 vaxis=axis2;
plot2 labval*asdt / noframe nolegend haxis=axis1 vaxis=axis3;
title1 h=0.17 in f=triplex "THE GRAPHICAL DISPLAY:";
title3 h=0.17 in  f=triplex  "Figure 12.1 - Urine Output, and Creatinine Levels
for Patient 1101";
title4 j=c h=0.17 in f=triplex  "Treatment Assignment: XXX";
footnote1 j=c h=0.15 in f=triplex "P = Plasma Creatinine";
run;
quit;
```

CONCLUSION
This paper has introduced a solution to display a bar chart and a scatter plot on one page using the SAS GPLOT and a simple DATA step in SAS. This solution can accommodate data with no pre-defined groups on the x-axis which proves that generating complicated graphical displays in SAS *is* easier than you think.

REFERENCES
Carpenter, Arthur L. and Shipp, Charles E. (1995), Quick Results with SAS/GRAPH Software, Cary, NC: SAS Institute Inc.

SAS Institute (1999), SAS/GRAPH Software: Reference Volume 1, Version 8, Cary, NC: SAS Institute Inc.

SAS Institute (1999), SAS/GRAPH Software: Reference Volume 2 Version 8, Cary, NC: SAS Institute Inc.

ACKNOWLEDGEMENT
The author would like to express her appreciation to the following people for their invaluable comments and suggestions in this paper:

Pat Wesolowski, Averion Inc.

Randy Childs, Averion Inc.

CONTACT INFORMATION
The author may be contacted at:
    Rita Tsang
    Averion Inc.
    4 California Avenue
    Framingham, MA 01701
    (508)416-2694
    Fax: (508)416-2797
    E-mail: rtsang@averioninc.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.

THE GRAPHICAL DISPLAY:

Figure 12.1 – Urine Output, and Creatinine Levels for Patient 1101
Treatment Assignment: XXX

P = Plasma Creatinine