

Getting ANNOTATE Data Sets to Work For You A Beginning Tutorial

Cynthia J. Musser, Marshfield Medical Research Foundation, Marshfield, WI

Abstract

A beginning tutorial will be presented for SAS/GRAPH® software users who would like to enhance graphics output using ANNOTATE data sets but have been frustrated with or have not had the time to tackle the SAS manual.

Building an ANNOTATE data set to enhance existing graphics output containing:

- function variables - what to do
- position variables - where to do it
- attribute variables - how to do it

will be discussed and illustrated with examples of code and resulting output. Most examples will be from a biomedical environment but will be generalizable enough to aid anyone in enhancing existing graphics output.

Introduction

If you have time to read and experiment with SAS/GRAPH software you will find you can do amazing things. Unfortunately, when many people find they need to modify existing graphs they are under tight time deadlines. The purpose of this paper is to demonstrate how to enhance existing SAS/GRAPH output using ANNOTATE data sets. Examples will illustrate writing text on a graph, customizing bar graphs, drawing on a graph, and creating a legend.

Creating an ANNOTATE Data Set

An annotate data set contains three types of variables. Function variables tell what to do. Position variables tell where to do it. Attribute variables tell how to do it. These variables form observations in the ANNOTATE data set and each observation performs an action.

	<u>Variable</u>	<u>Description</u>
FUNCTION (WHAT TO DO)	FUNCTION	LABEL Writes text
	MOVE	Moves to a point
	DRAW	Draws a line
POSITION (WHERE TO DO IT)	XSYS	Coordinate system for x
	YSYS	Coordinate system for y
	X	Horizontal coordinate
	Y	Vertical coordinate
	GROUP	Used with GCHART instead of X or Y
	MIDPOINT	Used with GCHART instead of X or Y
ATTRIBUTE (HOW TO DO IT)	SIZE	Height of text
	STYLE	Text font
	TEXT	Message to be written
	COLOR	Color
	POSITION	Placement and alignment of text
	WHEN	Draw before/after graph
	LINE	Line type to use

All of the following examples assume you have an existing graph and you need to enhance the graph with text or by drawing on the graph.

Writing Text on a Graph (Figures 1 and 2)

To modify an existing graph you need to create an ANNOTATE data set that contains observations which assign values to the function, position and attribute variables. The variable FUNCTION is assigned the value LABEL to indicate you are writing text on the graph.

```
DATA ANNOT1;
FUNCTION = 'LABEL';
```

You must decide on the coordinate system to use (TABLE 1). Probably the easiest to use is XSYS = '2' and YSYS '2'. This tells SAS you want to write the text in the actual graph area and not in the surrounding border area. It also allows you to use axis values for the X and Y coordinates.

```
XSYS = '2'; YSYS = '2';
X = 1990; Y = 58.5;
```

Specifying values for the attribute variables determine text size, style, and content.

```
SIZE = 1.2; STYLE = 'SWISSB';
TEXT = 'Regression eq: Age at implant =
      .382 * Implant year - 698.388';
```

Last, in the code for the existing graph tell SAS to use the ANNOTATE data set as an option in this example, on the PLOT statement.

```
PROC GPLOT DATA=PRED;
PLOT AGE * IM_YR P_AGE * IM_YR/
FRAME HMINOR = 0 VMINOR = 4 OVERLAY
VAXIS = AXIS1 HAXIS = AXIS2
ANNOTATE = ANNOT1;
```

You can write several text messages on a single graph by creating an observation for each message and outputting that observation to the ANNOTATE data set (Figure 2).

To write text anywhere on the graphics output area (entire page) specify XSYS = '4' and YSYS = '4'. The unit of measurement for X and Y is then a cell (TABLE 1). The number of cells range from zero to HPOS or VPOS (number of columns or rows in the graphics output area).

```
XSYS = '4'; YSYS='4';
X = 80; Y = 2;
```

The only attribute to be changed is the text. Finally, output the observation to the ANNOTATE data set.

```
TEXT = 'Regression Analysis';
OUTPUT;
```

To continue to write text on an existing graph, for each observation change position and attribute variables and then output the observation.

Customizing a Bar Graph (Figure 3)

When you have created a graph using GCHART and want to write or draw on the graph, GROUP, MIDPOINT, and either X or Y (for HBAR or VBAR respectively) are used for positioning. For a vertical bar chart with a group axis specify the horizontal positioning variables of GROUP and MIDPOINT and the vertical position variable Y. In this example we want each vertical bar labeled with the average difference in angle measurement. The vertical coordinate is determined by the value of the variable DIFF. When the data used in the ANNOTATE data set belong to an existing SAS data set, use the SET statement to read the observations into the ANNOTATE data set.

```
DATA ANNOGRP;
SET DIFFER;
*FUNCTION WHAT TO DO;
FUNCTION = 'LABEL';
*POSITION WHERE TO DO IT;
XSYS='2'; YSYS='2';
GROUP = FACTOR;
MIDPOINT = TIME;
Y = DIFF;
```

When writing the value of a variable on the graph, you can use a formatted PUT statement specifying the variable name and an optional format.

Values for the variable POSITION affect placement of text. POSITION='8' indicates the text is centered one cell below the point specified. A complete listing of positions can be found on page 522 in SAS/GRAPH software.

```
TEXT = PUT (DIFF, 5.2);
POSITION = '8';
WHEN = 'A';
```

When you want to write on an existing chart produced by PROC GCHART specify WHEN = 'A'. This writes the annotate graphics after the GCHART output thus overlaying the procedure output.

By specifying WHEN = 'B' (default) annotate text or graphics are drawn before the GCHART procedure output.

Drawing on a Graph (Figure 4)

If you want to draw on an existing graph MOVE allows you to move to a specific point within the graph and DRAW is used to draw a line.

In this example to show variability, we want to draw bars indicating standard deviation on a vertical bar chart. To position the pointer at the top of the bar, we specify the horizontal coordinate using MIDPOINT, the vertical coordinate using Y and the FUNCTION 'MOVE'.

```
MIDPOINT = METHOD;
Y = MEAN + SD;
FUNCTION = 'MOVE';
OUTPUT;
```

To output the top of the bar:

```
FUNCTION = 'LABEL';
TEXT = '-';
OUTPUT;
```

Change the vertical coordinate and output the bottom of the bar.

```
Y = MEAN - SD;  
FUNCTION = 'LABEL';  
TEXT = '-';  
OUTPUT;
```

Since two points have already been specified in the ANNOTATE data set we will draw the line to connect the points.

```
FUNCTION = 'DRAW';  
OUTPUT;
```

Creating a Legend (Figure 5)

Creating or customizing your own legend can clarify an existing graph. It is easiest to draw the box for the legend and then write the text of the legend inside the box.

In general, decide where to place the legend box on the graph. Start in the lower left corner of the box. Draw the bottom of the box, the right side, the top, and then the left side of the box. The attribute LINE allows you to select line type.

```
LINE = 1;  
FUNCTION = 'MOVE'; X=295; Y=23;  
OUTPUT;/*LOWER LEFT*/  
FUNCTION = 'DRAW'; X=415; Y=23;  
OUTPUT;/*BOTTOM*/  
FUNCTION = 'DRAW'; X=415; Y=77;  
OUTPUT;/*RIGHT*/  
FUNCTION = 'DRAW'; X=295; Y=77;  
OUTPUT;/*TOP*/  
FUNCTION = 'DRAW'; X=295; Y=23;  
OUTPUT;/*LEFT*/
```

Once you have the legend box drawn decide on the legend content. By using the LABEL function along with positioning variables X and Y, and attribute variables describing the text, the legend can be written in the box. Using STYLE = 'SPECIAL'; and TEXT = "H" specifies the special character circle be drawn in the legend.

```
FUNCTION= 'LABEL'; X=300; Y=70; SIZE=1;  
POSITION='C'; TEXT='Data Values:';  
OUTPUT;  
FUNCTION= 'MOVE'; X=380; Y=74; OUTPUT;  
FUNCTION= 'LABEL'; STYLE='SPECIAL';  
TEXT='H'; OUTPUT;
```

To create the line for the regression equation in the legend use the functions of MOVE and DRAW. To no longer print special characters remember to change style to none.

```
FUNCTION= 'LABEL'; X=300; Y=55; SIZE=1;  
POSITION='C'; TEXT='Regression Eq:';  
STYLE='NONE'; OUTPUT;  
FUNCTION= 'MOVE'; X=380; Y=59; OUTPUT;  
FUNCTION= 'DRAW'; X=410; Y=59; OUTPUT;
```

Conclusion

In conclusion, ANNOTATE data sets can be a powerful tool enabling you to enhance existing graphics output. Just remember to specify FUNCTION, POSITION, and ATTRIBUTE variables telling what, where, and how. Experiment when and if you have a free minute. Look for an example similar to what you want to do. Keep a notebook of completed graphs for future reference.

Acknowledgements

SAS and SAS/GRAPH are registered trademarks of SAS Institute, Inc. in the USA and other countries.

References

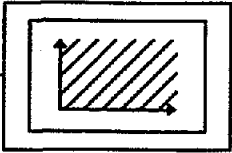
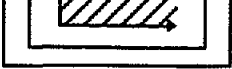
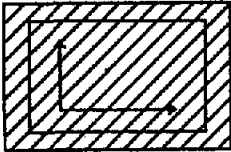
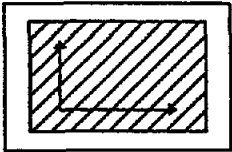
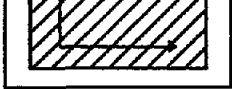
- SAS Institute Inc., SAS/GRAPH® Software: Reference, Version 6, First Edition, Volume 1, Cary, NC: SAS Institute Inc., 1990, pp. 794.
- SAS Institute Inc., SAS® System for Statistical Graphs, First Edition, Cary, NC: SAS Institute Inc., 1991, pp. 697.

Author

Comments are welcome, contact:

Cynthia J. Musser, MS
Marshfield Medical Research Foundation
1000 N. Oak Ave.
Marshfield, WI 54449
Internet: musserc@mfldclin.edu

TABLE 1
Selecting a Coordinate System

Value for XSYS, YSYS	Area	Unit of Measurement	Range
	DATA		
1		%	0-100%
2		axis values	min to max of axis
	Graphics Output		
3		%	0-100%
4		cell	0 to HPOS* 0 to VPOS*
	Procedure Output		
5		%	0-100%
6		cell	0 to HPOS* 0 to VPOS*

Sized after titles, footnotes
and legends

*Device Dependent
To determine use PROC GOPTIONS;RUN;

Age VS Year of Primary ICD Transplant Mean Age +/- SE

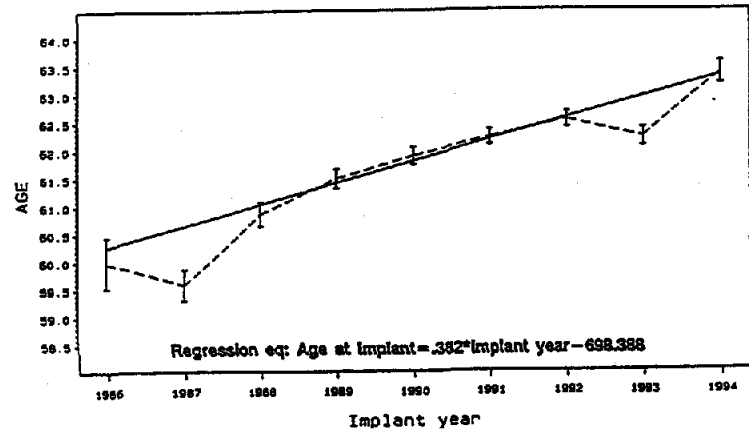


FIGURE 1

```
*****Create Annotate Data Set (Figure 1)*****;
data annot1;
  length function $8 text $60;
  *FUNCTION WHAT TO DO;
  function='label';
  *POSITION WHERE TO DO IT;
  xsys='2'; yeys='2';
  x=1990; y=58.5;
  *ATTRIBUTE HOW TO DO IT;
  size=1.2; style='swissb';
  text='Regression eq: Age at Implant=.382*Implant year-
        698.388';

*****Main Program (Figure 1)*****;
title f=swissb h=2.5 'Age VS Year of Primary ICD Transplant';
title2 h=2 f=swissb 'Mean Age +/- SE';

proc gplot data=pred;
  plot age*im_yr p_age*im_yr
    /frame hminor=0 vminor=4 overlay
    vaxis=axis1 haxis=axis2 annotate=annot1;

  symbol1 i=hiloctj cv=black ci=black l=3;
  symbol2 i=join v=none c=black;

  axis1 label=(a=90 h=1.5) value=(h=1) offset=(2)
    order=(58.5 to 64 by .5);
  axis2 label=(h=1.5) value=(h=1) offset=(4);
  format im_yr im_yr.;
run;
quit;
```

Age VS Year of Primary ICD Transplant Mean Age $\pm$ SE

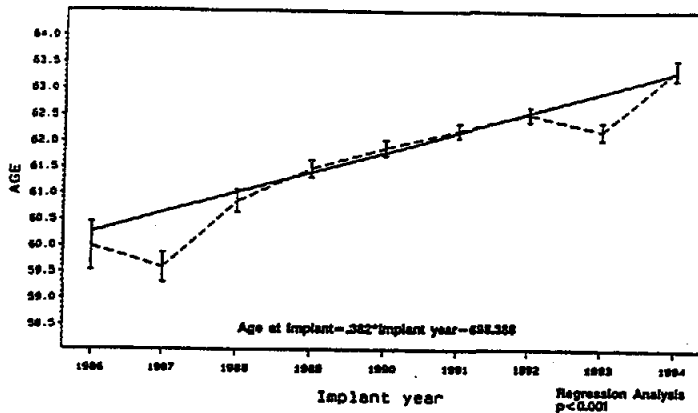


FIGURE 2

```
*****Create Annotate Data Set (Figure 2)*****;
*****Containing Several Observations*****;
data annot1;
  length function $8 text $40;
  *FUNCTION WHAT TO DO;
  function = 'label';
  *POSITION WHERE TO DO IT;
  xsys = '2'; ysys = '2';
  x = 1990; y = 58.5;
  *ATTRIBUTE HOW TO DO IT;
  size = 1.2; style = 'swissb';
  text = 'Age at Implant = .382 * Implant year - 698.388';
  output;

  *FUNCTION DOES NOT CHANGE;
  *POSITION WHERE TO DO IT;
  xsys = '4'; ysys = '4';
  x = 80; y = 2;
  *ATTRIBUTE ONLY TEXT CHANGES;
  text = 'Regression Analysis';
  output;

  *FUNCTION DOES NOT CHANGE;
  *POSITION ONLY COORDINATES CHANGE;
  x = 80; y = 1;
  *ATTRIBUTE CHANGE THE TEXT;
  text = 'p < 0.001';
  output;

*****MAIN PROGRAM SAME AS FIGURE 1*****;
```

Average Difference in Angle Measurements

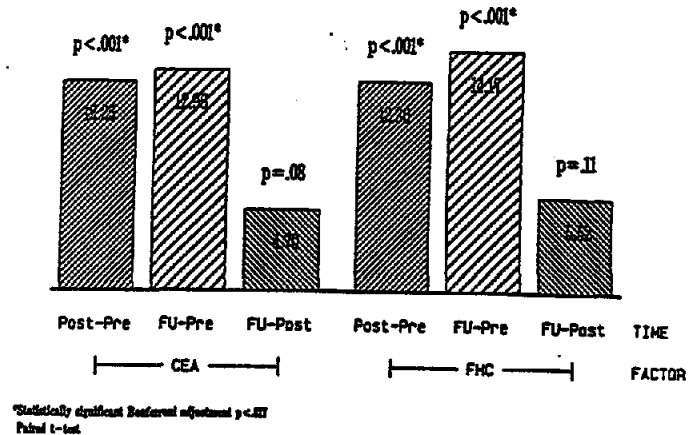


FIGURE 3

```
*****Create Annotate Data Set (Figure 3)*****;
data annogr;
  set differ;
  length text $ 7;
  *****Write data values on chart*****;
  *FUNCTION WHAT TO DO;
  function = 'label';
  *POSITION WHERE TO DO IT;
  xsys = '2'; ysys = '2';
  group = factor;
  midpoint = time;
  y = diff;
  *ATTRIBUTE HOW TO DO IT;
  size = 1; style = 'centx';
  text = put(diff,5.2);
  position = '8';
  when = 'a';
  output;

  *****Write p-values on chart*****;
  *FUNCTION HAS NOT CHANGED;
  *POSITION ONLY Y NEEDS TO BE CHANGED;
  y = diff + 1;
  *ATTRIBUTE TEXT AND POSITION HAVE CHANGED;
  text = pvalue;
  position = '2';
  output;

  *****Main Program (Figure 3)*****;
  title "Average Difference in Angle Measurements" font=centx;
  footnote j=1 h=.6 font=centb
    'Statistically significant Bonferroni adjustment
    p < .017';
  footnote2 j=1 h=.6 font=centb
    'Paired t-test';
  axis1 width=3 order=('Post-Pre' 'FU-Pre' 'FU-Post');
  axis2 label=none major=none minor=none style=0
    value=none;

  pattern1 value=l2 color=red;
  pattern2 value=r3 color=blue;
  pattern3 value=r2 color=green;

  proc gchart data=differ;
    vbar time/
      sumvar=diff
      subgroup=time
      group=factor
      maxis=axis1
      raxis=axis2
      nolegend
      annotate=annogr;
  format time $time.;
  run;
  quit;
```

Average Glucose Levels By Method

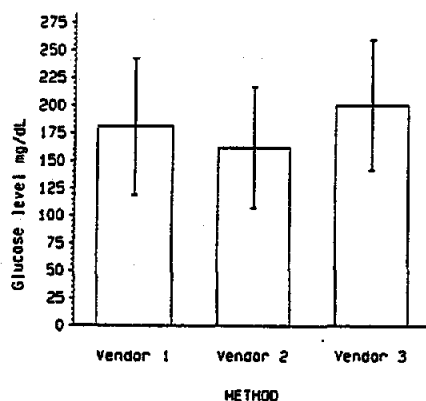


FIGURE 4

```
*****Create Annotate Data Set (Figure 4)*****;
data annoerr;
  set error;
  length function $8 text $4;
  xsys='2'; ysys='2';
  *****POSITION POINTER AND DRAW SD BARS*****;
  midpoint=method;
  y=mean+sd;
  function='move'; output;
  function='label'; text='-'; output;
  y=mean-sd;
  function='label'; text='-'; output;
  function='draw'; output;
  when='a';

  *****Main Program (Figure 4)*****;
  title h=2 font=swissb 'Average Glucose Levels By Method';
  axis1 order=(0 to 275 by 25)
    label=(a=90 'Glucose level mg/dL');

  pattern1 color=red value=e;
  pattern2 color=blue value=e;
  pattern3 color=green value=e;

  proc gchart data=error;
    vbar method/
      space=5
      sumvar=mean
      patternid=midpoint
      raxis=axis1
      annotate=annoerr;
    format method $method.;
  run;
quit;
```

Regression Analysis Glucose Results

$$\text{Lab Value} = 1.11 * (\text{Vendor A Value}) + 26.8$$

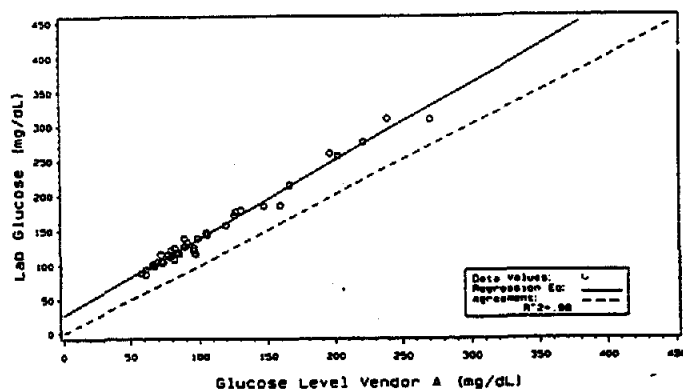


FIGURE 5

```
*****Create Annotate Data Set (Figure 5)*****;
data annot1;
  length function $8 text $25;
  xsys='2'; ysys='2';
  *****Draw box*****;
  line=1;
  function='move'; x=295; y=23; output; /*Lower Left*/
  function='draw'; x=415; y=23; output; /*Bottom*/
  function='draw'; x=415; y=77; output; /*Right*/
  function='draw'; x=295; y=77; output; /*Top*/
  function='draw'; x=295; y=23; output; /*Left*/
  *****Data Legend*****;
  function='label'; x=300; y=70; size=1; position='C';
  text='Data Values:'; output;
  function='move'; x=380; y=74; output;
  function='label'; style='special'; text='H'; output;
  *****Regression*****;
  function='label'; x=300; y=55; size=1; position='C';
  text='Regression Eq:'; style='none'; output;
  function='move'; x=380; y=59; output;
  function='draw'; x=410; y=59; output;
  *****Agreement*****;
  function='label'; x=300; y=40; size=1; position='C';
  text='Agreement:'; output;
  line=3;
  function='move'; x=380; y=44; output;
  function='draw'; x=410; y=44; output;
  *****R squared*****;
  function='label';
  y=25; x=355;
  position='B';
  size=1;
  text='R^2=.98';
  output;

  *****Main Program (Figure 5)*****;
  title f=swissb h=2.5 'Regression Analysis Glucose Results';
  title2 f=swissb h=2 'Lab Value = 1.11 * (Vendor A Value) + 26.8';
  proc gplot data=graph;
    plot lab * onetouch labref*accref
      /frame hminor=4 vminor=4
      vaxis=axis1 haxis=axis2
      overlay annotate=annot1;
    symbol1 i=r c=red value=circle;
    symbol2 i=join c=blue value=none l=3;
    axis1 label=(h=1.5 a=90) value=(h=1)
      order=(0 to 450 by 50);
    axis2 label=(h=1.5) value=(h=1) order=(0 to 450 by 50);
  run;
quit;
```