

Paper 110-25

How to Customize Axes in PROC GPLOT

Iza Peszek, Merck & Co., Inc., Rahway, NT

This paper is intended for SAS® users with at least intermediate experience using SAS/GRAPH®. It presents some simple applications of SAS ANNOTATE facility to achieve custom-tailoring of the axes for the plot produced via PROC GPLOT. While PROC GPLOT does a pretty good job with auto-selection of tick marks on the axes, sometimes the built-in functionality fails. This happens when the plotted data points are irregularly spaced and you want your tick marks to correspond to data points in the graph.

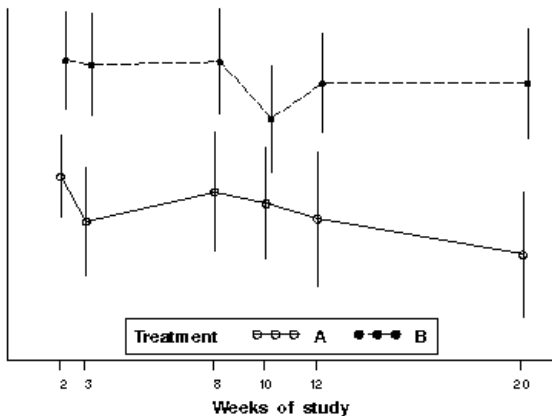
Another example is when PROC GPLOT is used with the ANNO option and the annotate data set has data points outside the natural data range used in PROC GPLOT.

In such cases, the common solution is to determine the new range of axes and to specify the ORDER option in the AXIS statement. Sometimes, however, it is not easy to determine the optimal value of the BY step in the ORDER option. Left to SAS, the selection of the BY step can be less than ideal. The notes below suggest an alternative way to label the tick marks when other solutions fail. We present two examples and suggest a method to format the axis in a desired manner.

INTRODUCTION

SAS System is the only graphic-capable software known to the author with virtually unlimited potential for custom-tailoring of the graphs. This is achieved either via a variety of available graphic statements or via SAS ANNOTATE facility. Some inexperienced users often find the ANNOTATE facility intimidating. We intend to demonstrate that it does not have to be so; and that even simple applications of the ANNOTATE facility can let you customize your graphs in a very impressive way.

While PROC GPLOT does a pretty good job with auto-selection of tick marks on the axes, sometimes the built-in functionality fails. This happens when the plotted data points are irregularly spaced and you want your tick marks to correspond to data points in the graph, as shown in Figure 1.



Another example is when PROC GPLOT is used with the ANNO option and the annotate data set has data points outside the natural data range used in PROC GPLOT. In such cases, the common solution is to determine the new range of axes and to specify the ORDER option in the AXIS statement. Sometimes, however, it is not easy to determine the optimal value of the BY step in the ORDER option. Left to SAS, the selection of the BY step can be less than ideal. The notes below suggest an alternative way to label the tick marks when other solutions fail. We present two examples and suggest how to format the axis to achieve the desirable result in each case.

EXAMPLE 1

You want to plot treatment means (TRTMEAN) over time (WEEK):

```
PROC GPLOT DATA=mydata;
PLOT TRTMEAN*WEEK=TRT;
...
```

The measurements are collected only during clinic visits which occur at 2, 3, 8, 10, 12 and 20 weeks, and you want your tick marks to correspond exactly to the clinic visits.

SOLUTION:

Our goal is to use the SAS ANNOTATE facility to supply the tick marks for the visits and to suppress the tick marks in the AXIS statement. The latter is relatively easy to accomplish:

```
AXIS1 order=&v_min to &v_max by &v_step
      minor=none major=none
      v=none
...
;
```

In this snippet, &v\_min and &v\_max denote the first and last visit (which can be easily determined using PROC MEANS) and &v\_step is the difference between &v\_max and &v\_min. Make sure that you store these values in the macro variables, as shown in the following snippet:

```
PROC MEANS means DATA=mydata NOPRINT;
VAR WEEK;
OUTPUT out=temp1 min=v_min max=v_max;
run;

DATA temp1; SET temp1;
v_step=v_max-v_min;
call symput ("v_min", put(v_min, best12.));
call symput ("v_max", put(v_max, best12.));
call symput ("v_step", put(v_step), best12.);
run;
```

Next, you create a data set with the tick marks (and their labels) to be used in the ANNO option of PROC GPLOT. SAS® provides a nifty macro, %label, which you can use to position the tick marks where you want them:

```
PROC SORT DATA=mydata NODUPKEY; BY WEEK;
%annomac;
DATA annodata; SET mydata (keep=visit);
length text $ 8;
xsys="2"; ysys="2";
%label(week, &v_min, "I", black, 0, 0, 1,
swiss, 5);
* tick marks;
xsys="2"; ysys="5";
%label(week, 5, put(week,2.), black, 0, 0, 1,
swiss, 5);
* tick mark labels;
run;
```

You may want to define the axis label so that it has two lines, the first line being blank :

```

AXIS1 ...
label=(h=4 j=c c=white ". "
      j=c c=black "Weeks of study");

```

This will leave room for the tick mark labels.

## EXAMPLE 2

Suppose that you want to plot the confidence intervals for treatment means over time. The confidence intervals are derived using PROC GLM, so you cannot plot them using the INTERPOLATION option inside the SYMBOL statement. You decided to use SAS ANNOTATE facility to supply the confidence intervals.

You create a dataset annodata to use with the ANNO option of PROC GPLOT; this dataset contains the variables WEEK (study week), TRT (treatment), LOWER (lower bound of the confidence interval) and UPPER (upper bound of the confidence interval). Your PROC GPLOT statements are shown in the snippet below:

```

PROC GPLOT DATA=mydata ANNO=annodata;
PLOT TRTMEAN*WEEK=TRT;
...
;

```

You do not know a-priori what should be the range of the vertical axis: it depends on your data. If you let SAS® determine the range, some of your confidence intervals may fail to display.

## SOLUTION

Once again, we need to determine the range of the horizontal axis and suppress the tick labels in the AXIS definition:

```

PROC MEANS means DATA=annodata NOPRINT;
VAR upper lower;
OUTPUT out=templ min=u_min l_min max=u_max
l_max;
run;

DATA templ; SET templ;
v_min=min(u_min, l_min);
v_max=max(u_max, l_max);
v_step=v_max-v_min;
call symput ("v_min", put(v_min, best12.));
call symput ("v_max", put(v_max, best12.));
call symput ("v_step", put(v_step, best12.));
run;

AXIS1 order=&v_min to &v_max by &v_step
      minor=none major=none
      v=none
...;

```

The macro variables &v\_min, &v\_max and &v\_step are assigned as in Example 1. To enhance the appearance of your custom axis, you may want to do some rounding of v\_min and v\_max before you calculate v\_step and store these values in macro variables. Ingenious programmers can write a program which does the rounding automatically. Since it is not the focus of this paper, we will not dwell on the details.

Now you need to add the tick labels to the annodata data set. But first you need to decide on the format of the tick mark

labels, and then to figure out how many tick mark you can have so that they do not overlap. Suppose that you rounded v\_min and v\_max to the nearest integer and you decided that you want at most 11 tick mark labels (the tick marks themselves can be more densely spaced if you wish).

The simplest way is to calculate the tick marks using the formula

$$\text{tick} = v_{\min} + i * (v_{\text{step}}) / 10; \quad i = 0, 1, 2, \dots, 10$$

This is all we need to create the tick marks and their labels in the annodata data set. Once again, we can use %label macro:

```

DATA annodata; set annodata;
xsys="2"; ysys="2";
if _N_=1 then do;
do i=0 to 9;
tick = &v_min+i*(&v_step)/9;
%label(...);
end;
%line(..);

```

When you specify this data set to be used in ANNO option of you PROC GPLOT, you will get the confidence intervals, the tick marks and their labels.

Sometimes, instead of setting the maximum number of tick mark labels, it may be more desirable to have the tick marks correspond to the points occurring on the graph (e.g., if you plot dates). Then your solution will be slightly different. You need to determine the minimum distance (in original data units) between the tick mark labels to assure no overlap. Suppose that you decided that the minimum distance is 30. The snippet below determines when the tick mark labels should be displayed:

```

Data annodata; set annodata;
...
retain tick_on 1;
prevweek = lag(week);
if _n_>1 then do;
if week-prevweek<30 and tick_on =1 then do;
%label();
tick_on=1;
end;
else do; tick_on=0; end;
end;
...;

```

In yet another scenario, you may want to display the tick marks labels for all treatment means, but to avoid overlap, you will "shift" their placement if necessary. For simplicity, assume that every other tick mark is to be shifted to the left:

```

DATA annodata; SET annodata;
...
if mod(_n_,2)=0 then do;
%label();
end;
else do;
%label();
end;
...;

```

Depending on your needs, you can define the position of the tick mark labels to suit your presentation best.

## CONCLUSION

The presented method shows how, with a little effort, you can achieve in SAS what is virtually impossible in any other graphic

package: create completely custom axes with the tick marks arbitrarily spaced. We hope that our hints will inspire you to solve more complicated graphical challenges.

#### **TRADEMARK INFORMATION**

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

Other brand and product names are registered trademarks or trademarks of their respective companies.

#### **CONTACT INFORMATION**

Your comments and questions are valued and encouraged.

Contact the author at:

Iza Peszek

Merck & Co., Inc.

Clinical Biostatistics, P.O. Box 2000

Rahway NJ 07065

Work Phone: (732) 594 3623

Fax: (732) 594 6075

**EMAIL: [IZABELLA\\_PESZEK@MERCK.COM](mailto:IZABELLA_PESZEK@MERCK.COM)**