ABSTRACT

The graphics annotate facility allows the user to create customised modifications to the graphic output. The coordinate system values (e.g. XSYS=3, YSYS=3) in combination with the coordinates (e.g. X=50, YSYS=50) plays vital role in annotating the graphics. When the coordinate system assumes graphics/procedure output area (e.g. XSYS=3, YSYS=5 etc.), determining the coordinates is often judgemental and is done by trial and error. Hence it is a time consuming task. There is no functionality in SAS/Graph which will provide the required coordinates to user instantly.

This paper details the creation and use of dummy frames which can be superimposed on the graph (to be produced) in order to determine the coordinates quickly.

INTRODUCTION

A coordinate system determines how coordinates are interpreted. User specify a coordinate system to use for each dimension, using the XSYS, YSYS, and ZSYS variables (for X, Y, and Z, respectively).

Each coordinate system refers to one of three drawing areas: data area, procedure output area, and graphics output area. Coordinates are measured from a different origin for each area; they also have different limits. Areas and their Coordinate Systems show the areas on the graphics output and the coordinate systems that use them.
Let's discuss the creation and use of the frame using the following graph template. The graph below contains the horizontal bars representing the odds ratio and corresponding upper and lower confidence interval for different countries. The basic graph can be produced using PROC GPLOT.

Figure 1. Basic Forest Plot using PROC GPLOT

The interpretation of odds ratio i.e. the labels below x-axis (Treatment A better & Treatment B better) needs to be annotated. Since these labels are outside data area, the coordinate system value can't assume data area (i.e. XSYS=2, YSYS=2), so we need to use graphics/procedure output area (i.e. XSYS=3, YSYS=3). Once the values for coordinate system are fixed, next step is to determine the corresponding coordinates (X, Y).

Since we are using graphics output area, the permissible values for the coordinates ranges from 0 to 100. Since we need the labels below x-axis, the coordinates (X, Y) corresponding to (XSYS, YSYS) need to be determining by mere guess. A typical approach is to start with some seed coordinate say (X=20, Y=20). If the seed coordinates doesn't work then try with some other values until the required coordinates are obtained.

The use of following frames when overlaid with graph to be produced helps to identify the coordinates instantly. We will see how to create the frames and determine the coordinates as follows.
Frame 1:

This frame contains horizontal and vertical lines at the interval specified by the user. In the above frame horizontal and vertical lines are drawn at the interval of 5% of the total graphics/procedure output area. The horizontal and vertical lines have been labelled from 5 to 100 which will help the user to identify the coordinates. For example, the coordinate (50, 50) is indicated by the arrow in the above frame.

Code to generate the frame:

```plaintext
%macro frame1 (cosys=, /*Required co-ordinate systems i.e. 3 (Graphics output area) or 5 (Procedure output area) */
    coord= /* Intervals on X and Y axis where the horizontal and vertical lines would be drawn. Should be greater than or equal to 2. Preferably an integer value in order to read the labels from x-axis */);

data frame1;
length function $8. text $10.;
******************************************************************************;
** Draw the horizontal and vertical at the specified interval **;
******************************************************************************;

do i=coord to 100 by coord;
do j=0 to 100 by coord;
    function='move';xsys="&cosys";ysys="&cosys";x=i;y=j;output;
    function='draw';xsys="&cosys";ysys="&cosys";x=i;
```
Determination of the coordinates:

Figure 3. Forest Plot with superimposed Frame 1

After superimposing the frame decide the position where the labels are needed. Looking at the shells for the graphs, we can fix the position as indicated by the arrows. So it is very much clear that the label "Treatment A better" can be displayed on the coordinate (X=20, Y=5). Similarly, the label "Treatment B better" can be displayed on the coordinate (X=70, Y=5).
Figure 4. Frame 2

This frame is divided into 100 squares of 10% of height and 10% breadth of the total graphics/procedure output area each. Inside each square digits 0 to 9 are mentioned horizontally which will serve as dummy x-axis for that square. Similarly, the vertical axis with digit 0 to 9 mentioned will serve as dummy y-axis for that square. Each square is labeled as (X, Y), (1X, Y), (2X, Y) etc. Suppose if the required position of the coordinates falls in (1X, 2Y) square then determine that coordinate by tracing (X, Y) coordinate from dummy x-axis and y-axis of that square. For example, the coordinate (15, 16) is indicated by the arrow in the above frame.
Code to generate the frame:

```sourcecode
%macro frame2(cosys= /* Required co-ordinate systems i.e. 3 (Graphics output area) or 5 (Procedure output area) */);

data frame2;
length function $8. text color $5.;
******************************************************************;
** Get the horizontal and vertical lines at the interval of 10% **;
******************************************************************;
do i=10 to 100 by 10;
  do j=0 to 100 by 10;
    function='move';xsys="&cosys";ysys="&cosys";x=i;y=j;output;
    function='draw';xsys="&cosys";ysys="&cosys";x=i;
    y=100-j;line=2;color='red';output;
  end;
end;

*****************************************;
** Label the X-axis with digits 0 to 9 **;
*****************************************;
do m=0 to 99;
  do n=1 to 91 by 10;
    do l=0 to 90 by 10;
      if l<=m<=l+9 then do;
        function='label';xsys="&cosys";ysys="&cosys";x=m;y=n;
        text=trim(left(m-l));color='green';when='A';output;
      end;
    end;
  end;
end;

*******************************;
** Label each square **;
*******************************;
do a=5 to 95 by 10;
  do b=5 to 95 by 10;
    if a ne 5 and b ne 5 then do;
      function='label';xsys="&cosys";ysys="&cosys";x=a;y=b;
      text=compress(((a-5)/10)||'X,'||(b-5)/10)||'Y');
      color='green';when='A';output;
    end;
    else if a=5 and b ne 5 then do;
      function='label';xsys="&cosys";ysys="&cosys";x=a;y=b;
      text=compress('X,'||(b-5)/10)||'Y');
      color='green';when='A';output;
    end;
    else if a ne 5 and b=5 then do;
      function='label';xsys="&cosys";ysys="&cosys";x=a;y=b;
      text=compress(((a-5)/10)||'X','Y');
      color='green';when='A';output;
    end;
    else if a=5 and b=5 then do;
      function='label';xsys="&cosys";ysys="&cosys";x=a;y=b;
      text='X,Y';color='green';when='A';output;
    end;
  end;
end;
end;
```

**Label the Y-axis with digits 0 to 9**;

do q=1 to 99;
do p=0 to 99 by 10;
do r=0 to 99 by 10;
if r<q then do;
    function='label';xsys='&cosys';ysys='&cosys';x=p;y=q;
text=trim(left(qr));color='green';when='A';angle=90;output;
end;
end;
end;
end;
run;
%mend;

**Determination of the coordinates:**

![Figure 5. Forest Plot with superimposed Frame 2](image-url)
After superimposing the frame decide the position where the labels are needed. Looking at the shells for the graphs, we can fix the position as indicated by the arrows. So it is very much clear that the label “Treatment A better” can be displayed on the coordinate (X=20, Y=5). Similarly, the label “Treatment B better” can be displayed on the coordinate (X=70, Y=5).

**How to use the frames:**

**Step1:** Produce the basic graph without annotation.

**Step 2:** Use ANNO=FRAME1 or ANNO=FRAME2 in the PROC GPLOT which is used to create the above graph. This will superimpose the frame on the above graph.

**Step 3:** Once the coordinates are fixed, we can create annotated datasets to display the labels as follows.

```sas
data anno;
  length function $8. text $40.;
  function='label';xsys='3';ysys='3';x=20;y=5;
    text='Treatment A better';output;
  function='label';xsys='3';ysys='3';x=70;y=5;
    text='Treatment B better';output;
run;
```

**Step 4:** Replace the FRAME1 or FRAME2 data with ANNO and run the PROC GPLOT which will provide the required graph.

**CONCLUSION**

Use of the above frames simplifies the task of the user while using annotate facility of SAS/Graph. It is very useful when the graph has many texts/symbols to be annotated. With the help of the frames above, position of the co-ordinates can be obtained within few seconds and hence saves lots of time.

**ACKNOWLEDGMENTS**

I thank all my colleagues for guiding this work and carefully reviewing the paper with comments and suggestions.

**CONTACT INFORMATION**

Your comments and questions are valued and encouraged. Contact the author at:
Sandeep Sawant
i3 Statprobe,
7th Floor, Corporate Center,
Opp. To VITS Hotel,
Andheri-Kurla Road,
Andheri (E)- 400059
Mumbai, India
Work Phone: +91-22-30554032
E- mail: sandeep.sawant@i3global.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.