

PharmaSUG 2013 –CC26
Automating the Labeling of X- Axis
Sanjiv Ramalingam, Vertex Pharmaceuticals, Inc. , Cambridge, MA

ABSTRACT

Labeling of the X-axis usually involves a tedious axis statement specifying tick marks and typing labels for each of the time points. A methodology to implement and automate the generation of the axis statement for X-axis labeling with real time representation of the X-axis is discussed.

INTRODUCTION

Creating customized labels for the X-axis usually involves creating a value format or specifying the labels within the axis statement. The paper proposes another methodology where in the visit numbers (numeric variable), their respective labels (character variable) are already present in the input dataset. This information is then used to automate the creation of the labels.

AVISITN	AVISIT
0	D1
1	
2	
3	D2
4	D3

Figure 1. Illustration of structure of a typical input dataset used to automate labeling. Note that no label exists when AVISITN equals 1 or 2.

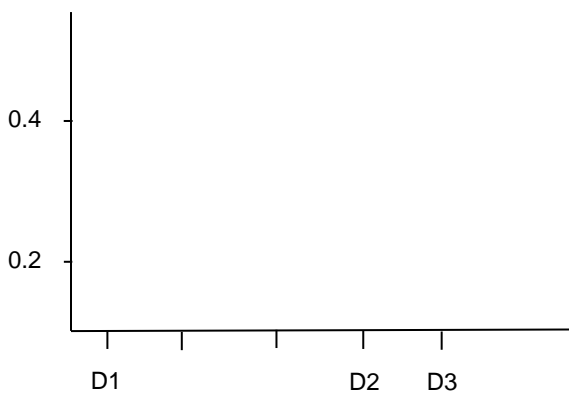


Figure 2. Illustration of how labeling appears in the X-axis when the algorithm is implemented. Note that though tick marks are created when AVISITN equals 1 or 2 no labels are created.

METHODOLOGY

Automating the labeling of X-axis involves dynamically creating two different formats.

The first format is a value format that is created for a unique combination of visits (numeric) and their labels. i.e.

```
Proc Format;
  VALUE val_fmt
    0='D1'
    2='D2'
;
```

The second format is a picture format created to insert blank labels for visit labels that are not unique. i.e.

```
Proc Format;
  PICTURE pic_fmt
    1=" "
    3=" "
;
```

The picture format is recommended because the value format cannot be used to create blank labels. Use of value format results in an apostrophe (') instead of blanks.

Dynamic creation of the formats involves creating a string/text of the required statement and the use of PROC SQL to combine all the string/text and convert them into a single macro variable. Two major steps are involved. In the first step, a text/string of the required code is created. This can be implemented using the concatenation operator.

Assuming that for each unique combination of visit numbers and their labels that is contained in a dataset, a text string that is basically an assignment statement for each of the visits needs to be created.

E.g. To create a statement 0='D1' where 0 is the visit number (avisitn) and D1 is the corresponding label (avisit) and to store it in a variable by name val_fmt,

```
val_fmt=strip(put(avisitn,best.))||strip("=")||strip(avisit)||strip("'");
```

It is then required to make all the records of the variable of interest(val_fmt) into one macro variable using PROC SQL.

```
proc sql;
  select val_fmt
  into :val_fmt separated by ' '
  from val_fmt;
quit;
/* Variable of interest */
/* Macro variable to be created */
/* Source dataset */
```

Automating the Labeling of X- Axis, continued

The step by step procedure to create such a graph of interest is as follows:

STEP 1

The macro variables required are first initialized.

```
/* Input/Source dataset */
%let inds=srcds;

/* Numeric Visit Variable */
%let visitn =avisitn;

/* Character Visit Variable that contains label name for Visit*/
%let visit =avisit;

/* Order/Space between visits; Typical in behavior compared to the by option within
the axis statement */
%let visitn =avisitn;

/* Maximum value of visit */
%let max_dmy=20;

/* Maximum value of visit */
%let min_dmy=0;
```

STEP 2

The input dataset is used to get the unique combination of visits and their respective labels. Two datasets are created; one for "value-format" and another for "picture-format". Records that do not have a unique combination of visit labels are output to the picture-format dataset. The required string/text statements are also created.

```
proc sort data=&inds nodupkey out=temp1(keep=&visitn &visit);
  by &visit &visitn;
run;

data val_fmt pic_fmt1;
  length pic_fmt $15 val_fmt $50;
  set temp1;
  by &visit;
  if not first.&visit then do;
    &visit=" ";
    pic_fmt=strip(put(&visitn,best.))||strip("=")||strip(' ' ' ');
    output pic_fmt1;
  end;
  else do;
    val_fmt=strip(put(&visitn,best.))||strip("=")||strip(&visit)||strip("");
    output val_fmt;
  end;
run;
```

STEP 3

If the ordering between the visits is not even then it is required to generate a dummy sequence to create blank labels. This also allows for real time representation of the X-Axis.

```
data temp2;
  do i=&min_dmy to &max_dmy by &spvisit;
    &visitn=i;
    output;
  end;
run;
```

Automating the Labeling of X- Axis, continued

STEP 4

A picture format needs to be assigned to all values that are not there in parent dataset.

```
proc sort data=temp1;
  by &visitn &visit;
run;

data pic_fmt2;
  length pic_fmt $15;
  merge temp2(in=a) temp1(in=b);
  by &visitn;
  if a and not b then do;
    pic_fmt=strip(put(&visitn,best.))||strip("=")||strip(' ' ' ');
    output;
  end;
run;

/* final picture format dataset */
data pic_fmt;
  set pic_fmt1 pic_fmt2;
run;
```

STEP 5

Macro Variables for each of the formats are then created. In certain cases a picture format may not be required at all. In such cases the macro variable for picture format is created only if the observations in the picture format dataset pic_fmt is greater than zero.

```
proc sql;
  select val_fmt into:val_fmt separated by ' '
  from val_fmt;
quit;

proc sql noprint;
  select count(*)
  into :OBSCOUNT
  from pic_fmt;
quit;

%macro chkobs();

%if %eval(&obscount gt 0) %then %do;

  proc sql;
    select pic_fmt into:pic_fmt separated by ' '
    from pic_fmt;
  quit;

  proc format ;
    value val_fmt
    &val_fmt
    other=[pic_fmt.]
    ;
    picture pic_fmt
    &pic_fmt
    ;
  run;

%end;
```

Automating the Labeling of X- Axis, continued

```
%if %eval(&obscount eq 0) %then %do;

    proc format ;
        value val_fmt
            &val_fmt
        ;
    run;

%end;

%mend;

%chkobs;
```

The dynamically created format can then be invoked in the graphical procedure statements such as a GPLOT as a simple format statement.

e.g.

```
format avisitn val_fmt. ;
```

STEP 6

To dynamically create customized tick marks the annotate facility in combination with the GANNOMAC feature can be used to dynamically place tick marks by creating the required annotate dataset.

```
/* V_min - Minimum Y-value in the Source Dataset */
%let v_min=0;

data anno_tick;
    length when $1;
    retain when 'a';
    set temp1;
    xsys='2'; ysys='2';
    %label(&visitn, &v_min, "I", black, 0, 0, 0.5,arial,8);
run;
```

CONCLUSION

A methodology to automate the creation of X –axis labels has been explained. To further automate, the source dataset can be used to determine the maximum (max_dmy) and minimum (min_dmy) values using the PROC MEANS procedure.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Please contact the author at:

Name: Sanjiv Ramalingam
Enterprise: Vertex Pharmaceuticals, Inc.
Address: 130 Waverly Street
City, State, ZIP: Cambridge, MA 02139
Email : sanjiv_ramalingam@vrtx.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.