SAS® GRAPHIC MACROS FOR PRODUCING 2-DIMENSIONAL AND 3-DIMENSIONAL STEP-GRAPHS
Akos Felsovalyi, J.C. Penney Company, Inc.

INTRODUCTION

Let data set A have variables X and Y1--YN (N>1). We can request a line graph (Figure 1) with the following commands:

```sas
PROC GPLOT DATA=A;
  PLOT (Y1--YN) * X / OVERLAY;
  SYMBOL I=JOIN;
```

The HIST2D, HIST3D and HIST3DM SAS graphic macros presented here transform data set A such that the new data sets yield 2-dimensional (Figure 2) and 3-dimensional (Figure 3 and 4) step-graphs.

The number of observations in data set A is k and the number of Y variables is n (n=N).

**%MACRO HIST2D**

Data set A contains n functions between X and Y1--YN. The functions are:

\[ Y_l = F(X) \]

... \[ Y_n = F(X) \]

Macro HIST2D transforms those functions into step functions:

\[ y_l = s(x) \]

... \[ y_n = s(x) \]

The domains of functions \( F \) are

\[ x, x', x', x', ... , x \]

\[ 1 \]

\[ 2 \]

\[ k \]

The domains of functions \( S \) are

\[ x', x', x', ... , x' \]

\[ 0 \]

\[ 1 \]

\[ 2 \]

\[ k \]

HIST2D has two methods to define \( x' \).

**Method 1:**

\[ x' = x - \frac{m}{k} \]

\[ 0 \]

\[ 1 \]

\[ 2 \]

\[ k \]

(where \( \sum_{i=2}^{k} (x - x_{i-1}) \)

\[ m = \frac{k}{k - 1} \]

i.e. \( m \) is the average of the distances between every two consecutive points)

\[ x'_i = x_i (i=1,2,...,k) \]

**Method 2:**

\[ x' = x - \frac{x - x_{i-1}}{i - 1} \]

\[ 1 \]

\[ 2 \]

\[ k \]

\[ k \]

\[ k \]

\[ (x' is the midpoint of x and x \text{ where } i = 1, ..., k-1) \]

Parameters of HIST2D:

- `DATA = SASdataset` specifies the input data set. Default is `_LAST_`.
- `OUT = SASdataset` specifies the output data set. Default is OUT.
- `ST = n` specifies the output data set. Default is OUT.
- `ZRO = n` specifies the method to define the step functions. See above. Default is 1.
- `METHOD = n` specifies the method to define the domains of the step functions.
- `X = variable` specifies the variable. Default is X.
- `Y1 = variable` specifies the first y variable. Default is Y.
- `YN = variable` specifies the last y variable. Default is Y.

In order to obtain the step-graph (histogram), the user needs to run procedure GPLOT using SYMBOL statement with I=JOIN.

The number of observations in the output data set is \( 2k \) if the "no border" option is specified (ZRO=0) and \( 3k+1 \) if the "border" option is specified (ZRO=1).

The listing of HIST2D is in Appendix B.
%MACRO HIST3D

Data set A contains functions
\[ Y_1 = F(x), \ldots, Y_N = F(x) \]

Macro HIST3D transforms those functions into a two variable step function: \[ Z = S(x, y) \]

The domain of each \( F \) is \( x_1, \ldots, x_n \). The domain of function \( S \) is \( (x_i, j) \), where \( j \in \{1, 2, \ldots, n\} \). The value of \( S \) at \( (x_i, j) \) is
\[ Z = S(x_i, j) = F(x_j) \]

Parameters of HIST3D:
- DATA, OUT, ZRO, X, Y1, YN are the same as those in HIST2D.
- YNAME variable specifies the name of the new y variable.
- ZNAME variable specifies the name of the new z variable.

In order to obtain the 3-dimensional step-graph the user needs to run procedure G3D with the variables X, YNAME, ZNAME.

The listing of HIST3D is in Appendix C.

%MACRO HIST3DM

Macro HIST3DM yields the same 3-dimensional step-graph that HIST2D does, but allows the user to color the top of every step. The coloring can be done because the graph is drawn by procedure GMAP. The macro creates two data sets (map and response).

Parameters of HIST3DM:
- DATA, ST, X, Y1, YN are the same as those in HIST2D.
- MAPDSET = SASdataset specifies the map data set. Default is MAP.
- RESPDSET = SASdataset specifies the response data set. Default is RESPONSE.

The unit area identifying variable in the map data set is ID and the response variable in the response data set is RESP.

Each observation of each Y variable generates a unit area in the map data set. These unit areas are unit squares and defined by 4 points. Thus the number of the unit areas in the map data set is \( n^2 \) and the number of observations is \( 4*n^2 \). In order to obtain the 3-dimensional step-graph, the user needs to run procedure GMAP with the PRISM statement, using the DISCRETE option (see 3). To have the graph with the right rotation, the user has to choose an appropriate viewing position (XVIEW, YVIEW and ZVIEW options).

A natural coloring is to assign the same color to all steps of one Y variable. Sometimes this can be cumbersome, since the coloring is done by the response variable. Macro HIST3DM makes this type of coloring easy by printing out the necessary PATTERN statements. The colors used in the PATTERN statements are called color1, color2, ..., colorn. (The above coloring can be achieved only if the values of one Y variable differ from the values of another Y variable.)

The listing of HIST3DM is in Appendix D.

EXAMPLE

Appendix A is a SAS program that compares the monthly sales data from three subdivisions. The program displays the sales in four ways, using a simple line graph (Figure 1), a 2-dimensional step-graph (Figure 2), a 3-dimensional step-graph without coloring (Figure 3) and a 3-dimensional step-graph with coloring (Figure 4). Note that the line graph and the 2-dimensional step-graph require the same GPLOT commands; that HIST2D is called with the "no border" option specified (ZRO=0), but HIST3D is called with the "border" option specified (ZRO=1); and that PRISM statement has the XVIEW=0.4, YVIEW=0.4 and ZVIEW=0.01 options.

These examples were run under VM/CMS and the macros resided in separate files.

REFERENCE


MORE INFORMATION

For more information please contact the author:

Akos Felsovalyi
J.C. Penney, 39th Floor
1301 Avenue of the Americas
New York, NY 10019
(212) 957-6696

SAS, SAS/GRAPH and SAS Communications are registered trademarks of SAS Institute Inc., Cary, NC, USA
Monthly Sales of Subdivisions 1–3

Figure 1

Monthly Sales of Subdivisions 1–3

Figure 2
Monthly Sales of Subdivisions 1–3

Figure 3

Monthly Sales of Subdivisions 1–3

Figure 4