Decorative InfoGraphs using SAS
Sanjay Matange, SAS Institute Inc.
Presenter: Prashant Hebbar
Sanjay Matange, Director, Data Visualization Division
SAS Institute Inc.

Sanjay is responsible for the development and support of ODS and ODS Graphics, including Graph Template Language (GTL) and Statistical Graphics (SG) procedures.

He has been with SAS for over 25 years, is the co-author of three patents, author of four SAS Press books and author of Graphically Speaking, a blog on data visualization.

Prashant Hebbar, Principal Software Developer, Data Visualization, SAS.

Areas of expertise are the Graph Template Language (GTL), the ODS Graphics Designer and image file formats.

Effective Graphs vs Decorative InfoGraph

The results of statistical analyses are best viewed graphically.

Effective graphs allow easy and accurate comparisons with minimal distractions or “Chart Junk” (Tufte).

• SG Procedures and GTL output follows these principles by default.

• But, what if you need a “cool” graph for a handout or brochure? Do you go learn a new software?

• Here, we will show you how to use our trusted SGPLOT procedure to create “Decorative InfoGraphics”.

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.
What is a “Decorative InfoGraph”? 

Every day we are flooded with graphs and text. 
• How does one stand out in this flood of data?  
• Use “Decorative Infographs” to catch the eye of the reader.
Revenue by Year: A Comparison

Traditional bar chart created with VBAR statement of the SGPLOT Procedure

InfoGraph of the same data also created using SGPLOT Procedure...
A Definition of Decorative InfoGraph

An aesthetic and eye catching visual.

- Invites the reader to spend a bit more time on its aesthetic details and information content.
- Delivers the relevant information in a different way.
- Some accuracy may be traded off for aesthetic.
- Not intended for technical journals; more likely in glossy brochures or magazines.
Creating Decorative InfoGraphs with PROC SGLOT

We use:

• SAS data sets
• Some data step code to transform the data
• SGPLOT procedure to render the visual
• The process is mostly scalable to different data.
• Key feature: SYMBOLIMAGE statement
  • Allows you to use any image as a marker symbol in plots, attribute maps and STYLEATTR statements
SYMBOLIMAGE NAME=S symbolic NAME image=“FILE-PATH” / OPTIONS;
quotes significant – no filerefs

Options:

- HOFFSET=offset  The horizontal offset from the data value to the image center.
- VOFFSET=offset  The vertical offset from the data value.
- ROTATE= number  Angle of rotation in degrees.
- SCALE=number  Scale factor for the image.

SYMBOLCHAR lets you use a character as a marker symbol.
Using Image Markers in Graphs

• Define a new symbol from an image using the SYMBOLIMAGE statement and give it a name.
• Next, use this new symbol name like any other predefined symbol in SGPLOT, such as CIRCLE or TRIANGLEFILLED.

```sas
%let coinOpaque=C:\OpaqueCoin.png;
proc sgplot data=class_age13_15;
    symbolimage name=coinOpaque image="&coinOpaque";
    scatter x=age y=height / markerattrs=(symbol=coinOpaque size=100);
    xaxis offsetmin=0.15 offsetmax=0.15;
    yaxis offsetmin=0.15 offsetmax=0.15;
run;
```
Using Standard Images

• Most images on the web or pictures taken on your phone are rectangular in shape, and all pixels are opaque.

• When such opaque images are used for marker symbols, the results are not ideal: background can be distracting.
Using Transparent Images

• To get markers with the intended shape of the object in the image, we make the outer background pixels transparent.

• Using transparent background images works better!
Creating Transparent Images

1. Using standard software like **Photoshop** or **Gimp**, make the outer “red” pixels transparent.

2. Save this image as a transparent PNG file (GIF works for 256 colors or less)
Coin Stack InfoGraph

Revenues (Millions) by Year

- 2011: 8
- 2012: 14
- 2013: 3
- 2014: 12

Indian 2 Annas Coin from 1942

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.
Coin Stack: 1. Display Markers Instead of Bars

For the original bar chart, we used the VBAR statement.

```sas
vbar year / response=resp;
```

Now lets start with a scatter plot with filled markers.

```sas
scatter x=year y=resp; text x=year y=resp text=resp;
```
Coin Stack: 2. Display a Stack of Markers

Generate additional observations for the response (Y) to display multiple markers **ordered from bottom to top**: `DO VAL=1 TO RESP BY 1;`

Creates a stack of markers for each YEAR value (X)

```
title 'Revenues (Millions) by Year';
proc sgplot data=bar noborder noautolegend;
scatter x=year y=val /
    markerattrs=(symbol=circlefilled size=80)
    filledoutlinedmarkers dataskin=gloss;

text x=year y=yValue text=resp /
    textattrs=(weight=bold color=white);
    xaxis display=(noticks noline nolabel) integer <options>;
    yaxis display=none <options>;
run;
```
title 'Revenues (Millions) by Year';
proc sgplot data=Coins noborder noautolegend;
symbolimage name=Indian image="&Indian";
scatter x=year y=val / markerattrs=(symbol=Indian size=110);
text x=year y=YVal text=resp / textattrs=(size=20 weight=bold color=white)
strip position=top backlight=0.75;
xaxis display=(noticks noline nolabel) integer;
yaxis display=none;
run;
Coin Stack: Variations

- Finally, use the JITTER option to display a slightly staggered stack.
- Numeric X and Y axes: random noise jitter.
- You can use different coin markers!
  NOTE: “Thicker coin” marker needs more vertical spacing (lesser iterations)
College Readiness Graph

Next: college readiness of a school by grade, compared to state average.

• The best visualization is a bar chart (left)
• But a SAS user wanted a graph like the one on the right.
The idea is similar to the Coin Stack.

1. Convert the bar lengths into multiple observations at every 10% - full group
   If response is not a full multiple, use **partial** group
   Pad to 100% with **empty** group

2. Use a scatter plot to draw the graph.

   Key difference: usage of 5 different markers (groups):

   one marker each for full 10% of School and State
   one marker each for partial value
   one for empty.
• Define the 5 symbols, and insert them into list of group symbols.

/*--Infographics Icons--*/
\$let file1=C:\A1.png;
\$let file2=C:\B1.png;
\$let file3=C:\C1.png;
\$let file4=C:\A2.png;
\$let file5=C:\B2.png;

/*--Create Graph with Headers--*/
title 'College Readiness';
proc sgplot data=Readiness4;
    symbolimage name=A1 image="&file1";
    symbolimage name=B1 image="&file2";
    symbolimage name=C1 image="&file3";
    symbolimage name=A2 image="&file4";
    symbolimage name=B2 image="&file5";
    styleattrs datasymbols=(A1 A2 C1 B1 B2);
    scatter y=level x=val / group=group markerattrs=(size=40);
    <Other statements to draw rest of the graph>
run;
1. Draw stacked bars of the ingredients for each recipe with the names of each ingredient.

2. Overlay a coffee cup symbol on top of each bar.

The inside pixels of the cup are transparent, so the ingredients show through. The outer opaque pixels draw the cup shape.
Using Transparent Image as a Mask

In the coin stack example, we used an image with outer transparent outer pixels to see the shape of the content.

- In the image on the left, the white pixels inside the cup are transparent.
- Let us place this image over the figure on the right (as a mask).
- The figure behind shows through only where there are transparent pixels.
We use a HIGHLOW statement to draw the ingredient levels and TEXT plot to display the ingredient names.

We overlay SCATTER markers with image symbol of coffee cup to mask the highlow bars into coffee cups.

title j=1 h=1 'Coffee Recipes';
proc sgplot data=Coffee noautolegend dattrmap=attrmap;
  pad=(bottom=20pct);
  symbolimage name=Cup image="&file1" / voffset=0.15;
  symbolimage name=Steam image="&file2" / voffset=0.0;
  highlow x=name low=low high=high / group=group
    type=bar nooutline attrid=Coffee;
  text x=name y=mid text=group / backlight=0.4
    textattrs=(size=5 color=white);
  scatter x=name y=y / markerattrs=(symbol=Cup size=100);
  scatter x=name y=ys / markerattrs=(symbol=Steam size=20);
run;
Shapes, Masks and Skins

Image markers can be used as:

- **Shapes** when outer pixels are transparent.
- **Masks** when inner pixels are transparent.
- **Skins** when overlaid with a transparency.
Soda Sales by Region

- We draw the basic bar chart to display the Sales by Region.
- We carve out the bottle shape by using markers as mask.
- We overlay the skin to get the bottle sculpting and highlights.
%let Mask=C:\Coke_Mask.png;
%let Skin=C:\Coke_Skin.png;
%let Bubbles=C:\Bubbles.png;

title j=1 h=1 'Soda Sales by Region (in Millions $)';
proc sgplot data=coke noborder noautolegend nocycleattrs;
    symbolimage name=Mask image="&Mask";
    symbolimage name=Skin image="&Skin";
    symbolimage name=Bubbles image="&Bubbles";
    highlow x=Region low=low high=high / type=bar nooutline
                       barwidth=0.7 fillattrs=(color=cx3f0f00);
    scatter x=Region y=y / markerattrs=(symbol=Mask size=200);
    scatter x=Region y=y / markerattrs=(symbol=Skin size=200)
                       transparency=0.4;
    scatter x=Region y=bubble / transparency=0
                       markerattrs=(symbol=bubbles size=20);
    run;
Conclusion

• The SGPLOT procedure is ideal for effective analytical graphs.
• But sometimes you need to stand out in the crowd...
• You can use the SYMBOLIMAGE feature to create Decorative InfoGraphs.

Let’s have some fun with ODS Graphics!!
Online Resources

• Graph Focus Area Page at: support.sas.com/rnd/datavisualization

• Graphically Speaking Blog at: blogs.sas.com/content/graphicallyspeaking
SAS Press Books

  - Sanjay Matange, Jeanette Bottittis

- Clinical Graphs Using SAS®
  - Sanjay Matange

- Getting Started with the Graph Template Language in SAS®: Examples, Tips, and Techniques for Creating Custom Graphs
  - Sanjay Matange

- Statistical Graphics Procedures by Example: Effective Graphs Using SAS®
  - Sanjay Matange, Dan Health