ABSTRACT

Effective graphs are indispensable for modern statistical analysis. They reveal tendency that are not readily apparent in simple tables and add visual clarity to reports. My client is a big graph fan; he always shows me a lot of high quality and complex sample graphs that were created by other software and asks me “Can SAS® duplicate these outputs?” Often, by leveraging the capabilities of the ODS Graphics Template Language (GTL) and the SGRENDER procedure, the answer is “Yes”. GTL offers SAS users more direct approach to customize the output and to overlay graphs in different level. This paper uses cases drawn from a real work situation to demonstrate how to get the seemingly unattainable results with the power of GTL:

- Utilizing bubble plots as your distribution density bars
- Creating refreshing looking linear regression graphics with the slop information in the legend
- Overlaying different plots together to create sophisticated analytical bottleneck test output

UTILIZING BUBBLE PLOTS AS YOUR DISTRIBUTION DENSITY BARS

Figure 1. Bubble density bars

As shown in Figure 1, the series plot does tell a clear story about whether the percentage of participants who answer ‘Yes’ to the question indicates a convergent tendency as the sample size increasing. The bubbles plots here work as distribution density bars that indicate how many individuals choose the “Yes” or “No” as the answers to the question and the area
of each bubble demonstrates the network size of each individual. If we ignore the matter of program efficiency for a moment, there is barely a way to create this combination by using SGPLOT. Fortunately, with knowledge of SAS® GTL, you can make this dream a reality. Here is the code:

```sas
proc template;
define statgraph wave;
dynamic _N_ _P_HISPANIC N2 N3 Y22 Y12 NSC_IDUS;
begingroup / designwidth=750 designheight=600;
  entrytitle halign=center "Convergent test for &var. &siteinl."
  layout lattice / rowdatarange=data columndatarange=data rows=3 rowgutter=10
columngutter=10 rowweights=(0.14 0.72 0.14);
  layout overlay / xaxisopts=(
    display=(TICKS TICKVALUES LINE )
    Linearopts=(tickvaluelsequence=(start=0 end=&n increment=50))
  )
  yaxisopts=(
    display=(TICKS LINE LABEL )label=('No')
  )
  bubbleplot x=_N y=Y22 size=NSC_IDUS/ DISPLAY=(FILL ) BUBBLERADIUSMAX=16
  BUBBLERADIUSMIN=2
  name='bubble' datatransparency=0.6 fillattrs=(color=CXFFFF00 transparency=0.8 );
  endlayout;
  layout overlay / wallcolor=CX348284
  yaxisopts=(
    display=(TICKS LINE TICKVALUES label)label='Percentage'
    Linearopts=(viewmin=0 viewmax=100 tickvaluelsequence=(start=0 end=100 increment=5))
  )
  xaxisopts=(
    display=(TICKS LINE )
    Linearopts=(tickvaluelsequence=(start=0 end=&n increment=50))
  );
  seriesplot x=_N y=_P_HISPANIC / name='series' connectorder=xaxis lineattrs=(thickness=2)
  xaxisopts=(
  )
  reference line y=&m / name='href' datatransparency=0.6 yaxis=Y curvelabelposition=max
  lineattrs=(color=CX0000FF thickness=0.3 );
  endlayout;
  layout overlay / xaxisopts=(
    display=(TICKS TICKVALUES LINE label )
    Linearopts=(tickvaluelsequence=(start=0 end=&n increment=50))
  )
  yaxisopts=(
    display=(TICKS LINE LABEL ) label=('Sample Size')
  )
  bubbleplot x=N3 y=Y12 size=NSC_IDUS/ DISPLAY=(FILL ) BUBBLERADIUSMAX=16
  BUBBLERADIUSMIN=2
  name='bubble2' datatransparency=0.6 fillattrs=(color=CX0000FF transparency=0.6 );
  endlayout;
endlayout;
endgraph;
end;
run;
```
CREATING REFRESHING LOOKING LINEAR REGRESSION GRAPHICS

Figure 2. Linear regression graphics

When it comes to linear regression graphics, most people would think of PROC REG. Admittedly, PROC REG meets normal needs for simple linear regression information. But what if we want the output looks prettier? What if we want to get rid of the confidence interval band? And what if we want to add the slope information to the output? GTL comes to rescue.

```plaintext
proc reg data=reg outest=t1 tableout alpha=0.1 noprint;
   model lnsce_idus = N;
   run;
data t2;
set t1(obs=1);
run;
proc sql noprint;
select N into: slope from t2;
quit;
```
OVERLAYING DIFFERENT PLOTS TOGETHER TO CREATE SOPHISTICATED ANALYTICAL BOTTLENECK TEST OUTPUT

Figure 3. Bottleneck test
The figure contains so much information! The needle plot here indicates the starts of different waves, the scatter plot and series plot together show whether the percentage of participants who answer ‘Yes’ to the question possesses a convergent tendency when passes each waves and sample size increases. And the band plots demonstrate the 95% confidence interval of the percent.

CONCLUSION
This paper provides a gentle and example-driven introduction to the graph template language. With the GTL, we can easily create professional-looking graphs that cannot be produced by the SG procedures. Although my methods are basic, and the codes could be improved a lot if we introduce the FCMP procedure to retrieve certain metrics, they could be quite useful when creating similar outputs.

ACKNOWLEDGMENTS
I would like to thank my boss Tonja Kyle and our client (CDC/OID/NCHHSTP) for providing me with great challenges over the years. I also want to thank my colleague Ge Wu, who is the co-author for this paper, for her helpful discussions.

CONTACT INFORMATION
Your comments and questions are valued and encouraged. Contact the author at:
Name: Wen Song
Company: ICF International
Address: 11785 Beltsville Dr, suite 300
City, State ZIP: Beltsville, MD, 20705
Work Phone: 301-572-0963
E-mail: wen.song@icfi.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.