

1309-2014

## Make It Possible: Create Customized Graphs with GTL

Wen Song, ICF International, Calverton, MD

Ge Wu, Johns Hopkins University, MD

### 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 slope information in the legend
- Overlaying different plots together to create sophisticated analytical bottleneck test output

### UTILIZING BUBBLE PLOTS AS YOUR DISTRIBUTION DENSITY BARS

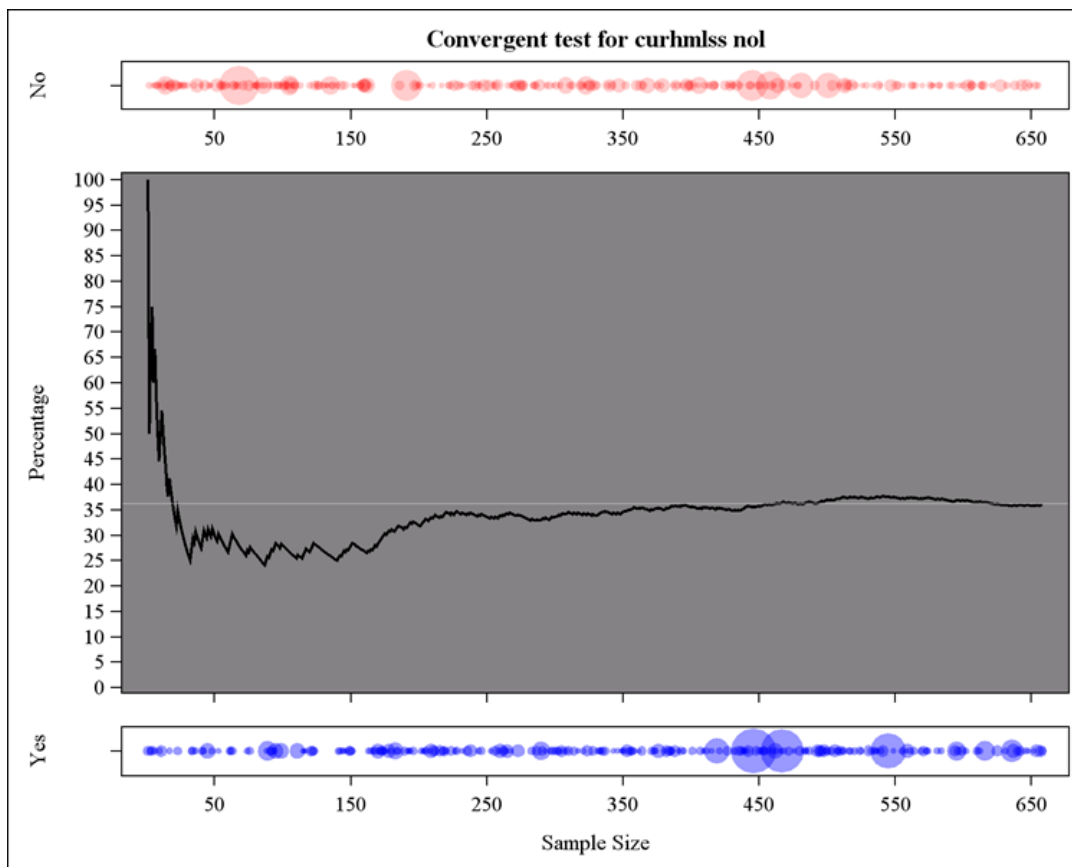


Figure1. Bubble density bars

As shown in Figure1, 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:

```

proc template;
define statgraph wave;
dynamic _N _P_HISPANIC _N2 _N3 _Y22 _Y12 _NSC_IDUS;
begingraph / designwidth=750 designheight=600;
  entrytitle halign=center "Convergent test for &var. &siteini.";
  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=(tickvaluesequence=(start=0 end=&n increment=50)))
yaxisopts=( display=(TICKS LINE LABEL )label=('No'));
  bubbleplot x=_N2 y=_Y22 size=_NSC_IDUS/ DISPLAY=(FILL ) BUBBLERADIUSMAX=16
BUBBLERADIUSMIN=2
  name='bubble' datatransparency=0.6 fillattrs=(color=CXFF0000 transparency=0.8 );
  endlayout;
  layout overlay / wallcolor=CX848284
yaxisopts=( display=(TICKS LINE TICKVALUES label)label='Percentage'
Linearopts=(viewmin=0 viewmax=100 tickvaluesequence=(start=0 end=100 increment=5)))
xaxisopts=( display=(TICKS LINE )
Linearopts=(tickvaluesequence=(start=0 end=&n increment=50)));
  seriesplot x=_N y=_P_HISPANIC / name='series' connectorder=xaxis lineattrs=(thickness=2
);
  referenceline y=&m / name='href' datatransparency=0.6 yaxis=Y curvelabelposition=max
lineattrs=(color=GreyDF thickness=0.3 );
  endlayout;
  layout overlay / xaxisopts=( display=(TICKS TICKVALUES LINE label )
Linearopts=(tickvaluesequence=(start=0 end=&n increment=50)) label=('Sample Size'))
yaxisopts=( display=(TICKS LINE LABEL ) label=('Yes'));
  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

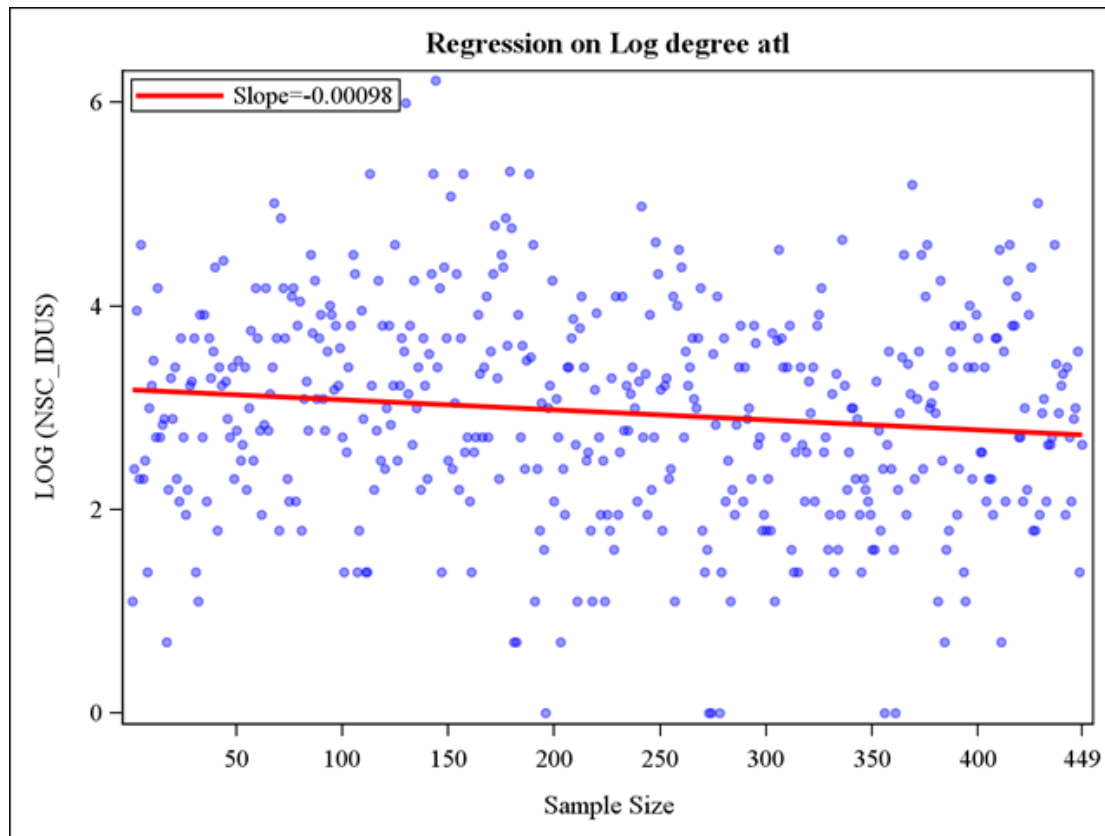


Figure2. Linear regression graphics

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 slop information to the output? GTL comes to rescue.

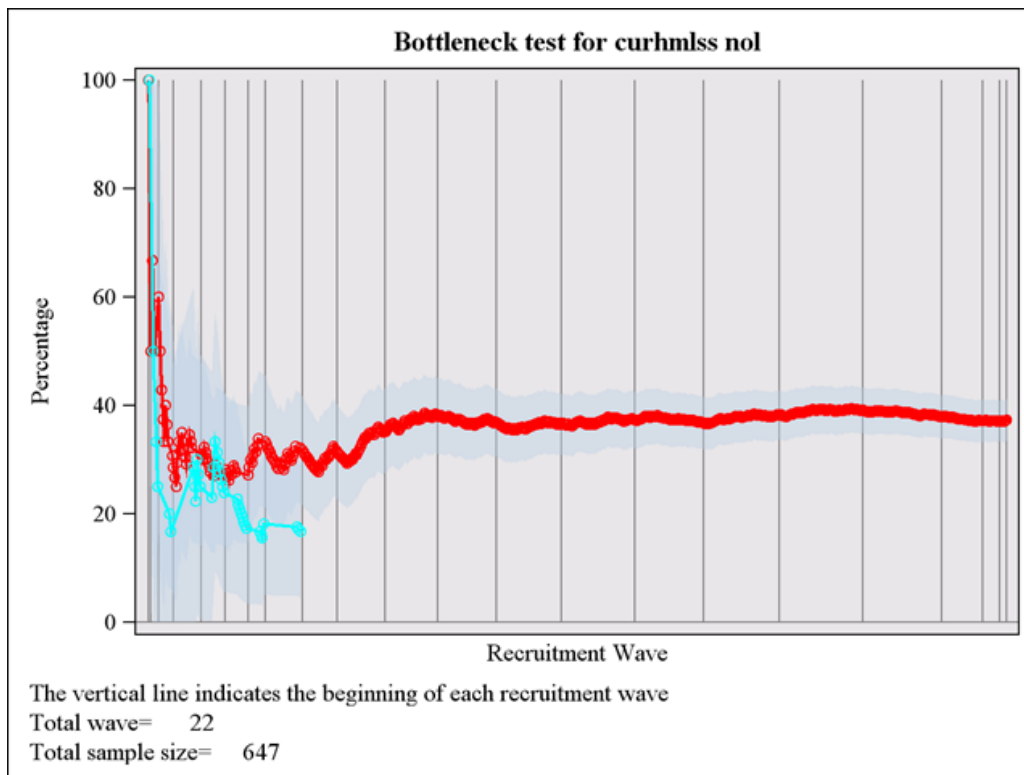
```
proc reg data=reg outest=t1 tableout alpha=0.1 noprint;
  model lncsc_idus = N;
run;
data t2;
set t1(obs=1);
run;
proc sql noprint;
select N into: slope from t2;
quit;
```

```

proc template;
define statgraph regression;
dynamic _N _LNSC_IDUS;
beginingraph;
  entrytitle halight=center "Regression on Log degree &siteini";
  layout lattice / rowdatarange=data columndatarange=data rowgutter=10 columngutter=10;
  layout overlay / xaxisopts=( label=('Sample Size')
Linearopts=(tickvaluesequence=(start=0 end=&n increment=50))
  yaxisopts=( label=('LOG (NSC_IDUS)'));
  scatterplot x=_N y=_LNSC_IDUS / name='scatter' legendlabel='Log(nsc_idu)'
datatransparency=0.6 markerattrs=(color=CX0000FF symbol=CIRCLEFILLED size=6 );
  regressionplot x=_N y=_LNSC_IDUS / name='regression' legendlabel="Slope=&slope"
lineattrs=(color=CXFF0000 thickness=3 );
  discretelegend 'regression' / opaque=false border=true halight=left valign=top
displayclipped=true across=1 order=rowmajor location=inside;
  endlayout;
  endlayout;
endgraph;
end;
run;

```

## OVERLAYING DIFFERENT PLOTS TOGETHER TO CREATE SOPHISTICATED ANALYTICAL BOTTLENECK TEST OUTPUT



**Figure3. Bottleneck test**

The figure3 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.

```

begingraph;
entryfootnote halign=Left "The vertical line indicates the beginning of each recruitment wave
";
entryfootnote halign=Left "Total wave=&nwave";
entryfootnote halign=Left "Total sample size=&total";

entrytitle halign=center "Bottleneck test for &var. &siteini.";
layout lattice / rowdatarange=data columndatarange=data rowgutter=10 columngutter=10;
layout overlay /wallcolor=CXE8E6E8 xaxisopts=( display=(LINE LABEL ) label=('Recruitment
Wave') griddisplay=off)
yaxisopts=( label=('Percentage'));
needleplot x=_WAVE y=_YWAVE / name='needle' lineattrs=(color=CX848284 );
bandplot x=_N2 limitupper=_UP1 limitlower=_LOW1 / name='band' datatransparency=0.6
connectororder=axis extend=true;
bandplot x=_N4 limitupper=_UP2 limitlower=_LOW2 / name='band2' datatransparency=0.6
connectororder=axis extend=true;
bandplot x=_N6 limitupper=_UP3 limitlower=_LOW3 / name='band3' datatransparency=0.6
connectororder=axis extend=true;
bandplot x=_N8 limitupper=_UP4 limitlower=_LOW4 / name='band4' datatransparency=0.6
connectororder=axis extend=true;

seriesplot x=_N y=_P1 / name='series' connectororder=xaxis lineattrs=(color=CXFF0000
thickness=2 );
scatterplot x=_N y=_P1 / name='scatter1' markerattrs=(symbol=circle color=CXFF0000);
seriesplot x=_N3 y=_P2 / name='series2' connectororder=xaxis lineattrs=(color=CX00FFFF
thickness=2 );
scatterplot x=_N3 y=_P2 / name='scatter2' markerattrs=(symbol=circle color=CX00FFFF);
seriesplot x=_N5 y=_P3 / name='series3' connectororder=xaxis lineattrs=(color=CX00FF00
thickness=2 );
scatterplot x=_N5 y=_P3 / name='scatter3' markerattrs=(symbol=circle color=CX00FF00);
seriesplot x=_N7 y=_P4 / name='series4' connectororder=xaxis lineattrs=(color=CXFF00FF
thickness=2 );
scatterplot x=_N7 y=_P4 / name='scatter4' markerattrs=(symbol=circle color=CXFF00FF);
endlayout;
endgraph;
end;
run;

```

## 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](mailto: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.