Paper 10920-2016

Using Animation to Make Statistical Graphics Come to Life

Jesse Pratt, Cincinnati Children's Hospital Medical Center

ABSTRACT

The Statistical Graphics (SG) procedures and the Graph Template Language (GTL) are capable of generating powerful individual data displays. What if one wanted to visualize how a distribution changes with different parameters, how data evolve over time, or to view multiple aspects of a three dimensional plot? By utilizing macros to generate each a graph for each frame, combined with the ODS PRINTER destination, it is possible to create .gif files to create effective animated data displays. This paper outlines the syntax and strategy necessary to generate these displays, as well as provides a handful of examples. Intermediate knowledge of PROC SGPLOT, PROC TEMPLATE, and the SAS® MACRO language is assumed.

INTRODUCTION

Using either the SG procedures or GTL in combination with the ODS PRINTER destination can generate animated displays of data. The steps, which will be expanded upon later, are as follows:

- Determine the general shape of the graph, which will be extended using a macro. The same principle applies to the first step in writing any macro, test one case first.
- Determine what the macro parameter(s) will be. The macro will be used to generate one graph per value of said parameter. This will be a value that changes, thereby changing each graph and creating each frame used in the animation.
- Write the macro. The individual macro written will depend on the type of graph that is desired and could require a separate graph per value of the macro parameter, or additionally require a separate data set per value of the macro parameter to generate each graph as well.
- Set the animation options.
- Invoke ODS PRINTER.
- Call the macro.
- Stop the animation.
- Close ODS PRINTER.

Writing the macro for the animation, to reiterate, is typically dependent on the type of graphics that are to be displayed, and thus general ideas will be expressed in the examples to follow. First, the general syntax will be discussed.

GENERAL SYNTAX

Below is the general syntax to create an animated .gif file from SAS® graphics:

```
options printerpath=gif animation=start animduration=<number>
animloop=<yes/no> noanimoverlay nodate nonumber;

ods printer file="<filename>";

<call macro to generate graphs>
```

```
options printerpath=gif animation=stop;

ods printer close;
```

The options used are explained as follows:

- `printerpath=gif`
  Tells SAS what type of file to put all of the graphs into. .gif is recommended.
- `animation=start`
  Tells SAS to start the animation.
- `animduration=`
  Sets frame speed in seconds. This can be changed to speed up or slow down the animation, accordingly.
- `animloop=`
  Loop the animation.
- `noanimoverlay`
  Replace each graph with the next instead of overlaying.
- `animoverlay`
  Overlapy each graph with the next.
- `nodate nonumber`
  Suppress the date and page number of each graph.
- `animation=stop`
  Tells SAS to stop the animation
- Note that the syntax for ODS PRINTER is the same for other ODS destinations.

EXAMPLE 1 – RELATIONSHIP BETWEEN TWO SERIES GRAPHS

This first example illustrates the basic process of creating an animated .gif with PROC SGPLOT. Consider the following SAS data set below (created by generating random numbers) containing one independent variable (XVAR) and two dependent variables (Y1VAR, Y2VAR):

	XVAR	Y1VAR	Y2VAR
1	1	0.1849625698	0.9700887157
2	2	0.3998243061	0.2593986454
3	3	0.9216025779	0.9692773498
4	4	0.5429791731	0.5316917228
5	5	0.0497940262	0.0665665516
6	6	0.8193185706	0.5238705215
7	7	0.8533943109	0.0671845768
8	8	0.9570238576	0.2971939642
9	9	0.2726117891	0.6899296309
10	10	0.9767648624	0.2265075185

Suppose we want to view the relationship between Y1VAR and Y2VAR as XVAR is increasing as an animation. This involves creating a data set starting with the first observation being complete and the rest of the observations having missing values of Y1VAR and Y2VAR, and then filling in a value one by one until the final data set is complete:

	XVAR	Y1VAR	Y2VAR
1	1	0.1849625698	0.9700887157
2	2	.	.
3	3	.	.
4	4	.	.
5	5	.	.
6	6	.	.
7	7	.	.
8	8	.	.
9	9	.	.
10	10	.	.

	XVAR	Y1VAR	Y2VAR
1	1	0.1849625698	0.9700887157
2	2	0.3998243061	.
3	3	.	.
4	4	.	.
5	5	.	.
6	6	.	.
7	7	.	.
8	8	.	.
9	9	.	.
10	10	.	.

	XVAR	Y1VAR	Y2VAR
1	1	0.1849625698	0.9700887157
2	2	0.3998243061	0.2593986454
3	3	.	.
4	4	.	.
5	5	.	.
6	6	.	.
7	7	.	.
8	8	.	.
9	9	.	.
10	10	.	.

	XVAR	Y1VAR	Y2VAR
1	1	0.1849625698	0.9700887157
2	2	0.3998243061	0.2593986454
3	3	0.9216025779	.
4	4	.	.
5	5	.	.
6	6	.	.
7	7	.	.
8	8	.	.
9	9	.	.
10	10	.	.

	XVAR	Y1VAR	Y2VAR
1	1	0.1849625698	0.9700887157
2	2	0.3998243061	0.2593986454
3	3	0.9216025779	0.9692773498
4	4	.	.
5	5	.	.
6	6	.	.
7	7	.	.
8	8	.	.
9	9	.	.
10	10	.	.

	XVAR	Y1VAR	Y2VAR
1	1	0.1849625698	0.9700887157
2	2	0.3998243061	0.2593986454
3	3	0.9216025779	0.9692773498
4	4	0.5429791731	0.5316917228
5	5	0.0497940262	0.0665665516
6	6	0.8193185706	0.5238705215
7	7	0.8533943109	0.0671845768
8	8	0.9570238576	0.2971939642
9	9	0.2726117891	0.6899296309
10	10	0.9767648624	0.2265075185

Figure 1: Sample frame data sets for series plots

Graphing each of these data sets produces the following plots used as animation frames:

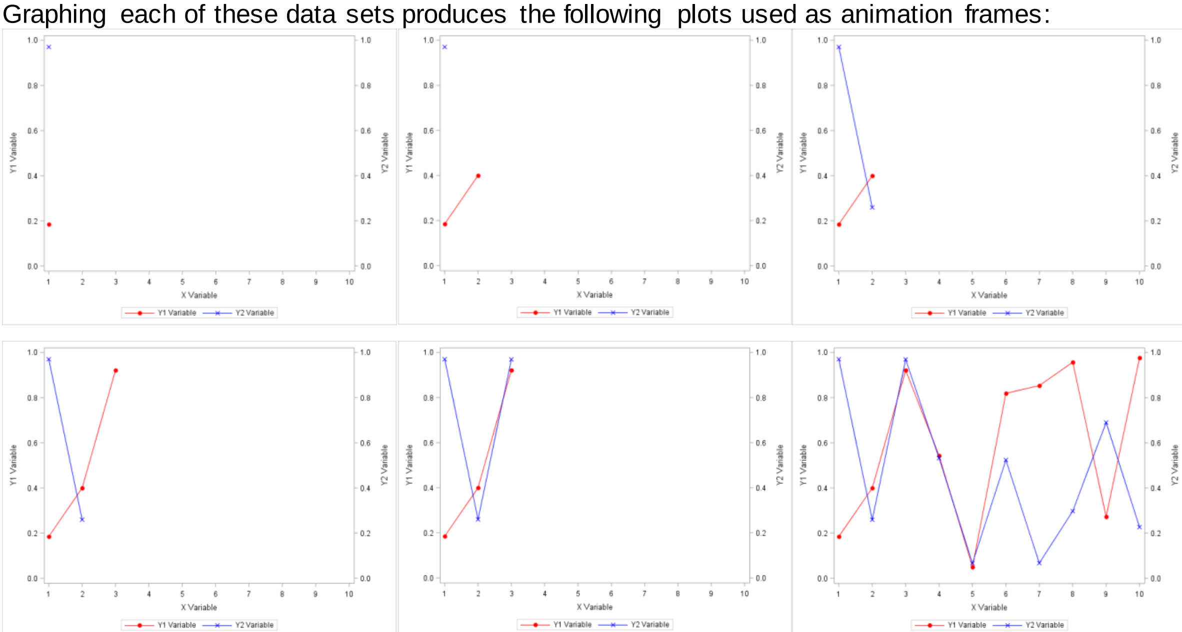

Figure 2: Individual animation frames for series plots

Running a macro that produces all of these graphs inside of the ODS PRINTER block will produce the animated .gif file that runs each frame. See the appendix for the code that generates this.

EXAMPLE 2 – THE BETA DISTRIBUTION

This example illustrates the changes in the shape of the beta distribution when one parameter is kept fixed and the other is allowed to vary. Consider the following data set where the independent variable (XVAR) is generated from from a random uniform distribution and the dependent variable (YVAR) is a beta distribution with alpha parameter equal to 2:

	i	j	XVAR	YVAR	j10
32	1	4.1	0.5452258906	0.9910538949	41
33	1	4.2	0.9039809983	0.0109384285	42
34	1	4.3	0.4140916483	1.6168836485	43
35	1	4.4	0.9805190768	0.0000356431	44
36	1	4.5	0.5501460403	0.831387731	45
37	1	4.6	0.5039294146	1.0405718722	46
38	1	4.7	0.548427094	0.7755052809	47
39	1	4.8	0.0090250815	0.2427495776	48
40	1	4.9	0.3326741314	1.9860270936	49
41	1	5	0.967108103	0.0000339588	50
42	2	1	0.8385434867	1.6770869734	10
43	2	1.1	0.9167508296	1.6515863428	11
44	2	1.2	0.7061726678	1.4592619873	12
45	2	1.3	0.7101744351	1.464459043	13
46	2	1.4	0.0953314887	0.3077311374	14
47	2	1.5	0.5936839355	1.419118174	15

In this example, the variable j10 is the variable that defines each frame and is based on the varying parameter of the beta distribution, j. To illustrate the data sets and subsequent graphs generated, consider some sample data sets where j = 1, 1.9, and 5:

	i	j	XVAR	YVAR	j10
1	1	1	0.2812413384	0.5624826767	10
2	2	1	0.8385434867	1.6770869734	10
3	3	1	0.9670992484	1.9341984968	10
4	4	1	0.6938883437	1.3877766875	10
5	5	1	0.9700242006	1.9400484012	10
6	6	1	0.9398488267	1.8796976534	10
7	7	1	0.6228938357	1.2457876714	10
8	8	1	0.0753918868	0.1507837736	10
9	9	1	0.676605082	1.3532101639	10
10	10	1	0.7304002641	1.4608005283	10
11	11	1	0.8820012384	1.7640024767	10
12	12	1	0.0565259196	0.1130518392	10
13	13	1	0.856203217	1.712406434	10
14	14	1	0.8999579013	1.7999158026	10
15	15	1	0.691173978	1.382347956	10
16	16	1	0.9901513769	1.9803027538	10
17	17	1	0.387305747	0.7746114939	10

	i	j	XVAR	YVAR	j10
1	1	1.9	0.5765730746	1.4659075451	19
2	2	1.9	0.2087524446	0.9316747209	19
3	3	1.9	0.5364595221	1.4796788118	19
4	4	1.9	0.1454995499	0.6959119832	19
5	5	1.9	0.1059575812	0.5278444348	19
6	6	1.9	0.3387808531	1.2864157096	19
7	7	1.9	0.5661076235	1.471276912	19
8	8	1.9	0.457288041	1.4536300597	19
9	9	1.9	0.8703074208	0.7628653687	19
10	10	1.9	0.5975252388	1.4513516106	19
11	11	1.9	0.2935444202	1.1830464647	19
12	12	1.9	0.552823222	1.4762815592	19
13	13	1.9	0.9623727049	0.2769794801	19
14	14	1.9	0.6063284304	1.4437107092	19
15	15	1.9	0.9782675672	0.1717935146	19
16	16	1.9	0.4456488674	1.4439456649	19
17	17	1.9	0.7499159653	1.1869749922	19

	i	j	XVAR	YVAR	j10
1	1	5	0.967108103	0.0000339588	50
2	2	5	0.1877369672	2.4516462512	50
3	3	5	0.8965986934	0.0030748467	50
4	4	5	0.3621742755	1.7982372051	50
5	5	5	0.0705171205	1.5789973585	50
6	6	5	0.331302337	1.9873070614	50
7	7	5	0.8505962732	0.0127142414	50
8	8	5	0.0683837505	1.5453341101	50
9	9	5	0.4730995993	1.0939260142	50
10	10	5	0.1889570375	2.4527866121	50
11	11	5	0.7750838491	0.0595048702	50
12	12	5	0.7769638443	0.0576797206	50
13	13	5	0.7850521825	0.0502749838	50
14	14	5	0.2541139863	2.3596070139	50
15	15	5	0.9513119394	0.0001603741	50
16	16	5	0.2220526669	2.4399337452	50
17	17	5	0.7557244235	0.0807244139	50

Figure 3: Sample frame data sets for the beta distribution

When fitted to a LOESS curve in PROC SGPLOT, these data sets generate the following plots:

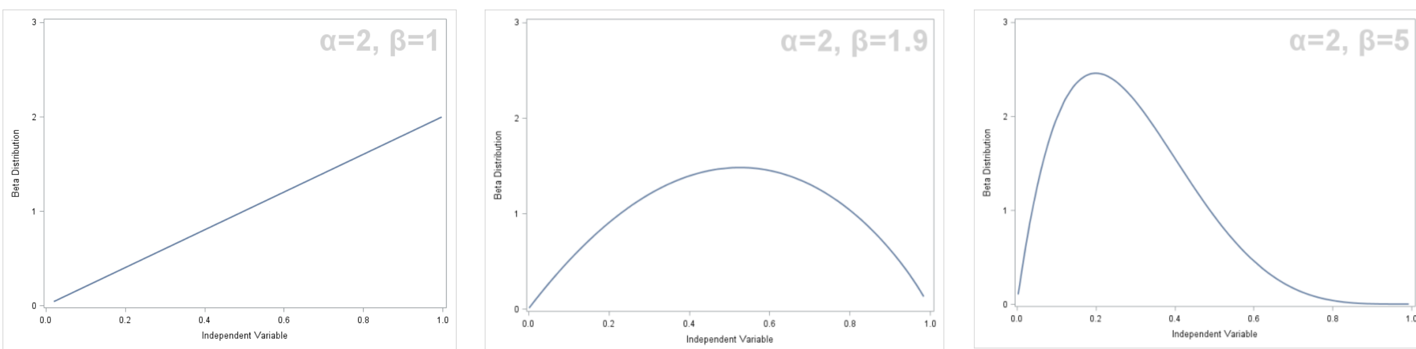

Figure 4: Individual animation frames for the beta distribution

Note that these are only a sample of 3 graphs; one graph for each value of beta from 1 to 5 by 0.1 increments are generated for animation. See the appendix for precise code.

EXAMPLE 3 – POWER VS. HORIZONTAL SHIFT OF A NORMAL DISTRIBUTION

This example illustrates what happens to the power of a standard normal distribution test as the mean differs from zero. Consider the following data set that starts the animation. The anomaly at zero is due to the floating point calculation due to looping with a decimal index.

	XVAR	YVAR	YVAR2	LOW	HI
391	-0.1	0.3969525475	0.194186055	0	0
392	-0.09	0.3973298316	0.1965204989	0	0
393	-0.08	0.3976677055	0.1988631194	0	0
394	-0.07	0.3979660682	0.2012135427	0	0
395	-0.06	0.3982248302	0.2035713883	0	0
396	-0.05	0.3984439141	0.2059362687	0	0
397	-0.04	0.3986232542	0.20830779	0	0
398	-0.03	0.3987627968	0.2106855517	0	0
399	-0.02	0.3988624999	0.2130691468	0	0
400	-0.01	0.3989223338	0.2154581618	0	0
401	-4.09915E-14	0.3989422804	0.217852177	0	0
402	0.01	0.3989223338	0.2202507667	0	0
403	0.02	0.3988624999	0.2226534988	0	0
404	0.03	0.3987627968	0.2250599353	0	0
405	0.04	0.3986232542	0.2274696325	0	0
406	0.05	0.3984439141	0.2298821407	0	0
407	0.06	0.3982248302	0.2322970047	0	0
408	0.07	0.3979660682	0.2347137639	0	0

The variable XVAR is the independent variable, YVAR is the dependent variable that represents the standard normal distribution, YVAR2 is the dependent variable that represents the shifted standard normal distribution, and LOW and HI are variables that will assist with graphing. The following data sets and subsequent graphs represent the shifted normal distribution moving down the x-axis:

	XVAR	YVAR	YVAR2	LOW	HI
510	1.09	0.2202507667	0.3989223338	0	0
511	1.1	0.217852177	0.3989422804	0	0
512	1.11	0.2154581618	0.3989223338	0	0
513	1.12	0.2130691468	0.3988624999	0	0
514	1.13	0.2106855517	0.3987627968	0	0
515	1.14	0.20830779	0.3986232542	0	0
516	1.15	0.2059362687	0.3984439141	0	0
517	1.16	0.2035713883	0.3982248302	0	0
518	1.17	0.2012135427	0.3979660682	0	0
519	1.18	0.1988631194	0.3976677055	0	0
520	1.19	0.1965204989	0.3973298316	0	0
521	1.2	0.194186055	0.3969525475	0	0

	XVAR	YVAR	YVAR2	LOW	HI
650	2.49	0.017971133	0.3989223338	0	0.3989223338
651	2.5	0.0175283005	0.3989422804	0	0.3989422804
652	2.51	0.0170946705	0.3989223338	0	0.3989223338
653	2.52	0.0166701008	0.3988624999	0	0.3988624999
654	2.53	0.0162544505	0.3987627968	0	0.3987627968
655	2.54	0.015847579	0.3986232542	0	0.3986232542
656	2.55	0.0154493471	0.3984439141	0	0.3984439141
657	2.56	0.0150596163	0.3982248302	0	0.3982248302
658	2.57	0.0146782491	0.3979660682	0	0.3979660682
659	2.58	0.014305109	0.3976677055	0	0.3976677055
660	2.59	0.0139400605	0.3973298316	0	0.3973298316
661	2.6	0.0135829692	0.3969525475	0	0.3969525475

	XVAR	YVAR	YVAR2	LOW	HI
800	3.99	0.000139285	0.3989223338	0	0.3989223338
801	4	0.0001338302	0.3989422804	0	0.3989422804
802	4.01	0.0001285762	0.3989223338	0	0.3989223338
803	4.02	0.0001235162	0.3988624999	0	0.3988624999
804	4.03	0.0001186434	0.3987627968	0	0.3987627968
805	4.04	0.0001139514	0.3986232542	0	0.3986232542
806	4.05	0.000109434	0.3984439141	0	0.3984439141
807	4.06	0.0001050853	0.3982248302	0	0.3982248302
808	4.07	0.0001008992	0.3979660682	0	0.3979660682
809	4.08	0.0000968702	0.3976677055	0	0.3976677055
810	4.09	0.0000929928	0.3973298316	0	0.3973298316
811	4.1	0.0000892617	0.3969525475	0	0.3969525475

Figure 5: Sample frame data sets for the shifting normal distribution

Note the shading that represents an increase in power as the mean shifts from zero:

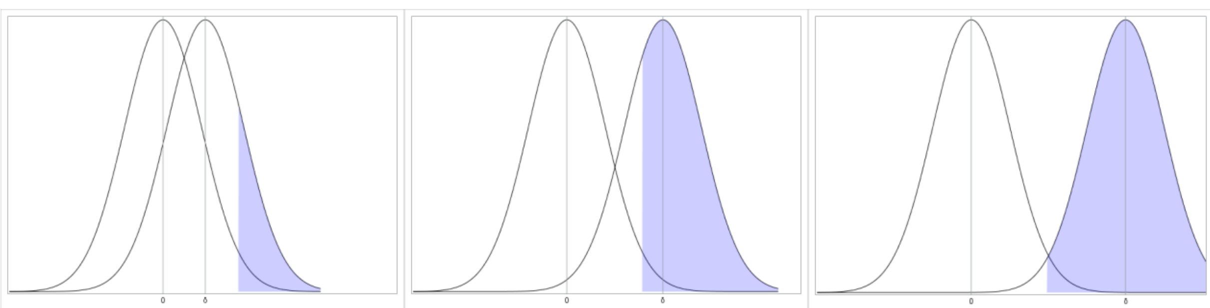

Figure 6: Individual animation frames for the shifting normal distribution

See the appendix for precise code.

EXAMPLE 4 – ROTATING A THREE DIMENSIONAL PLOT

When viewing a three dimensional plot, sometimes viewing from only one angle is insufficient. This example illustrates how to rotate a three dimensional plot in order to get a more complete viewing. It utilizes the Graph Template Language and will be different than the previous ones in that the templates create the frames instead of individual data sets. The details of the generation of the graph can be found in the SAS® 9.4 Graph Template Language User's Guide on pages 169-174. The key to animating this plot is to vary the value in the ROTATE= option of the LAYOUT OVERLAY3D statement as the macro below outlines:

```
ods escapechar="~";
%macro template;
%do i=1 %to 36;
 /* This creates a macro variable for the number of degrees of rotation */
  %let j=%sysevalf(10*&i);
  proc template;
```

```
    /* Create one template for each rotation */
    define statgraph surfaceplot&i;
      begingraph;
      entrytitle "SURFACECOLORGRADIENT=TEMPERATURE";
    /* Note the ROTATE option that rotates the graph by 10 degrees */
        layout overlay3d / cube=false rotate=&j
                           xaxisopts=(display=(line))
                           yaxisopts=(display=(line))
                           zaxisopts=(display=(line));
          surfaceplotparm x=length y=width z=depth / name="surf"
          surfacetype=fill
          surfacecolorgradient=temperature
          reversecolormodel=true
          colormodel=twocoloraltramp;
          continuouslegend "surf" /
                        title="Temperature (~{unicode  '00B0'x}F)";
        endlayout;
      endgraph;
    end;
    run;
  %end;
  %mend;
  %template;
```

After creating the templates, a data set is generated from the LAKE data set in the SASHELP library as outlined in the User's Guide referenced above. PROC SGRENDER is then run for each template to create the animation frames. Below are sample graphs:

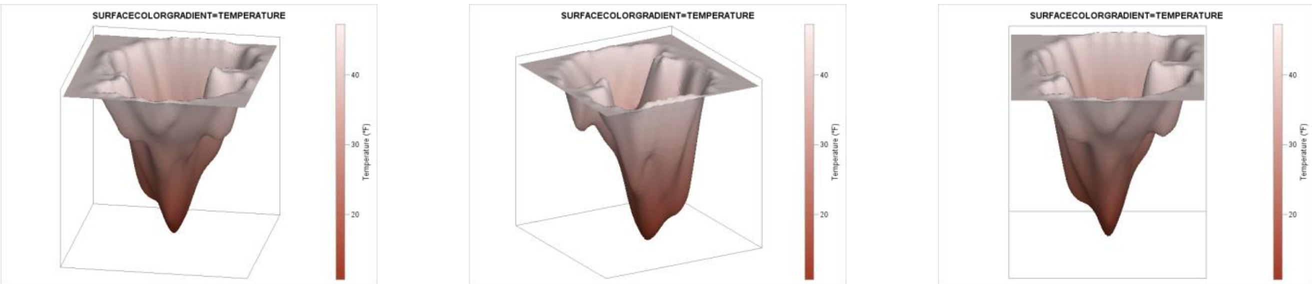

Figure 7: Individual animation frames for a rotated three dimensional plot

CONCLUSION

While the Statistical Graphics Procedures and the Graph Template Language are powerful tools for producing effective data displays, there are times they fall short when showing variation in parameters or multiple angles of a three dimensional plot. The ODS PRINTER destination in combination with these tools and a carefully crafted macro comes to the rescue in producing effective and informative animations.

APPENDIX – CODE USED TO GENERATE EACH EXAMPLE

EXAMPLE 1

```
/* Generate the starting data set */
data one;
do XVAR=1 to 10;
  Y1VAR=ranuni(XVAR);
  Y2VAR=ranuni(11*XVAR);
  output;
```

```
end;
run;

/* Transpose the data in order to easily make each value missing one at a
time */
proc transpose data=one out=two;
  by XVAR;
  var Y1VAR Y2VAR;
run;

/* Create an index variable to assist with missing values */
data three;
  set two;
  INDEX=_N_;
run;


%macro fill;
%do j=2 %to 20;
  /* Create data sets with sequentially missing values */
  data three&j;
  set three;
  if INDEX > &j then COL1=.;
  run;

  /* Get the data back into the original structure */
  proc transpose data=three&j out=five&j (drop=_NAME_);
  by XVAR;
  id  NAME ;
  var COL1;
  run;

  /* Create each frame using PROC SGPLOT */
  proc sgplot data=five&j noautolegend;
  series x=XVAR y=Y1VAR / markers markerattrs=(symbol=circlefilled color=red)
                      lineattrs=(pattern=solid color=red)
                      legendlabel="Y1 Variable" name="y1";
  series x=XVAR y=Y2VAR / y2axis markers markerattrs=(symbol=x color=blue)
                      lineattrs=(pattern=solid color=blue)
                      legendlabel="Y2 Variable" name="y2";
  xaxis label="X Variable" values=(1 to 10 by 1);
  yaxis label="Y1 Variable" values=(0 to 1 by 0.2);
  y2axis label="Y2 Variable" values=(0 to 1 by 0.2);
  keylegend "y1" "y2";
  run;
%end;
%mend;

/* Call the macro to generate the animation */
options printerpath=gif animation=start animduration=1 animloop=yes
noanimoverlay nodate nonumber;
ods printer file="H:\SAS Training\Animation Paper\seriesplot.gif";
%fill;
options printerpath=gif animation=stop;
ods printer close;
```

EXAMPLE 2

```
/* Create the starting data set */
data beta;
  do i=1 to 200;
    do j=1 to 5 by 0.1;
      XVAR=ranuni(234);
      YVAR=pdf("beta",XVAR,2,j);
      output;
    end;
  end;
run;
/* Create animation variable */
data beta2;
  set beta;
  j10=round(10*j,1);
run;

/* The next two blocks of code are used to generate macro variables used in
the INSET statement in each of the graphs */
data beta3;
  set beta2;
  if i=1;
  keep j;
run;
proc sql noprint;
  select j
  into :BETA10-:BETA50
  from beta3;
quit;

ods escapechar="~";
%macro betas;
%do i=10 %to 50;
  /* Subset on the animation variable */
  data beta2&i;
    set beta2;
    if j10=&i;
  run;
  /* Graph each frame */
  proc sgplot data=beta2&i noautolegend;
    loess x=XVAR y=YVAR/nomarkers;
    yaxis label="Beta Distribution" values=(0 to 3 by 1);
    xaxis label="Independent Variable";
    inset "~{unicode alpha}=2, ~{unicode beta}=&&BETA&i" / position=topright
           textattrs=(size=30 weight=bold color=lightgrey);
  run;
%end;
%mend;

/* Call the macro to generate the animation */
options printerpath=gif animation=start animduration=0.1 animloop=yes
noanimoverlay nodate nonumber;
ods printer file="H:\SAS Training\Animation Paper\beta.gif";
%betas;
options printerpath=gif animation=stop;
ods printer close;
```

EXAMPLE 3

```
ods escapechar="~";
%macro normals;
%do i=1 %to 30;
/* Want to index by 0.1 */
%let k=%sysevalf(&i/10);
/* Shifting the mean point */
%let m=%sysevalf(&k + 1);
/* Create a macro variable for the "scooting" of the shifted distribution */
%let j=%sysevalf(&k+4);
data normal&i;
  do XVAR=-4 to &j by 0.01;
    /* Generate values of a standard normal distribution */
     YVAR=pdf("normal",XVAR,0,1);
    /* Generate values of a shifted standard normal distribution */
     YVAR2=pdf("normal",XVAR,&m,1);
    /* LOW is always set to zero as a lower bound for the BAND statement in
PROC SGPLOT */
     LOW=0;
    /* Shades after the 97.5th percentile */
     if XVAR ge 1.96 then HI=YVAR2;
     else HI=0;
     output;
  end;
run;

/* Generate the plots for each frame */
proc sgplot data=normal&i noautolegend;
  series x=XVAR y=YVAR/lineattrs=(color=black);
  series x=XVAR y=YVAR2/lineattrs=(color=black);
  band x=XVAR lower=LOW upper=HI/fillattrs=(color=blue transparency=0.8);
  /* Create a reference line for the moving delta */
  refline &m/axis=x label="~{unicode delta}" labelpos=min;
  /* Create a reference line at zero */
  refline 0/axis=x label="0" labelpos=min;
  yaxis display=none;
  xaxis values=(-4 to 6) display=none;
run;
%end;
%mend;

/* Call the macro to generate the animation */
options printerpath=gif animation=start animduration=0.1 animloop=yes
noanimoverlay nodate nonumber;
ods printer file="H:\SAS Training\Animation Paper\normal slide.gif";
%normals;
options printerpath=gif animation=stop;
ods printer close;
```

CONTACT INFORMATION

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

Jesse Pratt
Cincinnati Children's Hospital Medical Center
513-803-3084
Jesse.Pratt_2@cchmc.org