

# **USING SAS® PROC GREPLAY WITH ANNOTATE DATA SETS FOR EFFECTIVE MULTI-PANEL GRAPHICS**

**Walter T. Morgan, R. J. Reynolds Tobacco Company**

## **ABSTRACT**

This presentation introduces SAS users to PROC GREPLAY and the ANNOTATE facility and demonstrates the use of user defined templates to produce a large number of similar graphs (up to 100 or more) on a single page. Of course this would preclude reading values off the axis, but with uniform scaling, these graphs provide very useful visualizations of large data sets. These graphs include scatterplot matrices, but beyond that are great for comparing individual subjects in a study, or any other grouping with a large number of levels (i.e., trellis graphs). They provide quick identification of outliers and influential individuals or groups, allowing for easy data manipulation to examine the robustness of results. The flexibility provided by the ANNOTATE facility combined with the power of GREPLAY make these kinds of graphs surprisingly easy to create. Attendees will be provided with a set of templates developed by the speaker.

## **INTRODUCTION**

With the proliferation of data collected and generated today, intelligent interpretation of data is of greater importance than ever. Data sets commonly feature large numbers of measures at a range of test conditions, for multiple test subjects taken at numerous time points. It is easy to become overwhelmed by the sheer mass of data. In the absence of tools for closely examining data, analysis may consist of comparison of means through analysis of variance or a simple regression analysis. This approach can result in complacency about the strength of conclusions, in that the experimenter sub-consciously expects that an effect found significant occurs more or less uniformly across test subjects / conditions. Multi-panel graphics can provide a way to visualize large amounts of data, resulting in a better understanding of the data, as well as help to identify outliers and unanticipated structural patterns in the data.

While more specialized software products may be able to produce multi-panel graphic displays, SAS already has these capabilities through the use of the ANNOTATE facility combined with PROC GREPLAY. Of course, as the number of graphs presented on a page increases, each graph becomes smaller. However, with uniform scaling, even very small graphs can provide insight into the complexities of a large data set. PROC GREPLAY provides the capability to present multiple graphs on a single page through the use of templates, while the ANNOTATE facility provides the capability of adding identifying labels to each graph. In this paper I will present an introduction to the ANNOTATE facility and to PROC GREPLAY, describing those aspects most relevant to creating effective multi-panel graphics.

## THE SAS/GRAPH ANNOTATE FACILITY

The SAS/GRAPH ANNOTATE facility allows users to customize graphs by adding text, lines or other graphic elements to a SAS graph created by PROC GPLOT, GCHART, GCONTOUR, G3D and other SAS/GRAPH procedures. These elements are stored in a SAS data set that the user creates for this purpose. An annotate data set is composed of variables with pre-defined names which the SAS/GRAPH procedure is expecting to find. These variables specify *what* to do, *where* to do it, and *how* to do it. These three basic functions are addressed by *action* variables, *positioning* variables, and *attribute* variables.

The only action variable is FUNCTION. This is a text variable which identifies the type of action to be taken. Values of FUNCTION include MOVE, DRAW, LABEL, POLY, BAR, and PIE. The value of FUNCTION in each observation in the annotate data set determines the action that the observation will perform.

Positioning variables include coordinate *system* positioning variables and coordinate *value* positioning values. Coordinate system positioning variables include XSYS, YSYS, and ZSYS. They are all character variables of length 1, with possible values given in Table 1. Different values alter the choices of drawing area, relevant units, and absolute vs. relative placement of coordinates.

Table 1 Coordinate System Values

Placement	Data Area		Graphics Output Area		Procedure Output Area	
	% Units	Value Units	% Units	Cell Units	% units	Cell Units
Absolute	'1'	'2'	'3'	'4'	'5'	'6'
Relative	'7'	'8'	'9'	'A'	'B'	'C'

The Data area refers to the graphics area within the axes of the graph. The Graphics Output area includes the graph and all the area around it. The Procedure Output area includes the graph and area around it, but only to include chart titles, axis labels, and legends. The choice of units indicates whether the value specified indicates the percent of the relevant area, or actual values or cell references. Absolute coordinates specify the exact location for a graphics element, while relative coordinates specify the location relative to another graphics element. For the purposes discussed in this paper, I generally use either '1' or '2' for these values.

Coordinate value positioning variables are X, Y, and Z for numeric coordinates, XC and YC for character coordinates, and GROUP, MIDPOINT, and SUBGROUP for GCHART output. The value assigned must be meaningful relative to the coordinate system value.

Attribute variables specify attributes of the graphic element to be drawn. These include COLOR, SIZE, LINE, STYLE, and POSITION (right aligned, left aligned, etc.).

ANNOTATE data sets may include other variables, but they will be ignored. Values for variables not meaningful for the FUNCTION specified will also be ignored. If a graphics procedure includes a BY statement, then all BY variables must also be present in the ANNOTATE data set.

### Example 1

A series of measurements from a manufacturing process are collected. Add lines to the graph indicating the target and control limits, and label the lines and a point out of control.

```
data Example1 ;
do Week=1 to 30 ;
  input PctDefect @@ ;
  output ;
end ;
datalines;
5.2 5.1 4.3 6.2 5.3 6.2 7.3 5.8 6.4 2.1
5.3 4.6 3.2 6.3 4.9 5.2 9.2 3.1 4.7 6.3
4.4 4.8 5.9 6.1 4.4 5.2 4.7 5.5 5.2 4.9
;
data AnnoEx1 ;
length text $ 20 function $ 8 color $ 5 ;
/* Draw lines for lower limit, upper limit, and target */
xsys='1' ; /* % of Data Area */
ysys='2' ; /* Data Values */
function='move' ; x=0 ; y=1.9 ; output ;
function='draw' ; x=100 ; y=1.9 ; line=2 ; color='red' ; output ;
function='move' ; x=0 ; y=7.9 ; output ;
function='draw' ; x=100 ; y=7.9 ; line=2 ; output ;
function='move' ; x=0 ; y=4.9 ; output ;
function='draw' ; x=100 ; y=4.9 ; line=1 ; color='black' ; output ;

/* Label lines outside graph */
xsys='3' ; /* % of Graphics Output Area */
ysys='2' ; /* Data Values */
function='label' ; size=1 ; position='C' ;
x=85 ; text='Target' ; y=4.9 ; output ;
text='Upper Control' ; y=7.9 ; output ;
text='Lower Control' ; y=1.9 ; output ;

/* Label outlier point */
xsys='2' ; ysys='2' ; /* Data Values */
function='label' ; size=1.5 ; color='red' ; position='4' ;
x=16.5 ; y=9.2 ; text='Substitute Operator' ; output ;
run ;

proc print data=AnnoEx1 ;
run ;
```

In the data step above, the values for the variables that the ANNOTATE facility will require are assigned. Note that variables XSYS and YSYS are character variables and are thus entered with quotes. Also note that only the variables required for a FUNCTION need be specified; thus the variables LINE and COLOR are not specified for the first observation for which FUNCTION='MOVE'. The user should be sure that the length of character variables is sufficient for all values that will be used. Table 2 below presents the contents of data set AnnoEx1.

Table 2 Example 1 Annotate Data Set AnnoEx1

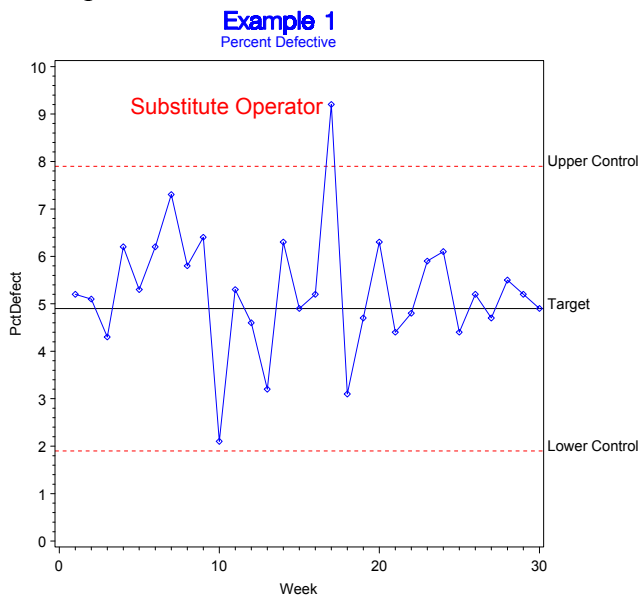
Obs	Text	Function	XSys	YSys	X	Y	Line	Color	Position	Size
1		Move	1	2	0	1.9	.			.
2		Draw	1	2	100	1.9	2	Red		.
3		Move	1	2	0	7.9	2	Red		.
4		Draw	1	2	100	7.9	2	Red		.
5		Move	1	2	0	4.9	2	Red		.
6		Draw	1	2	100	4.9	1	Black		.
7	Target	Label	3	2	85	4.9	1	Black	C	1
8	Upper Control	Label	3	2	85	7.9	1	Black	C	1
9	Lower Control	Label	3	2	85	1.9	1	Black	C	1
10	Substitute Operator	label	2	2	16.5	9.2	1	Red	4	1.5

```
goptions dev=win target=winprtc rotate=landscape ;
```

```
proc gplot data=Example1 ;
  plot pctdefect*week=1 / haxis=axis1 vaxis=axis2 anno=AnnoEx1 ;
  title1 c=blue 'Example 1' ;
  title2 c=blue 'Percent Defective' ;
  axis1 length=7.0 in ;
  axis2 label=(a=90) order=0 to 10 ;
run ;
```

Figure 1 presents the graph produced by GPLOT, using the ANNOTATE data set AnnoEx1. In this example, the ANNOTATE data set is accessed by adding it as an option to the PLOT statement. It can also appear in the PROC GPLOT statement, in which case it would apply to all subsequent PLOT statements.

Figure 1 Example 1 GPLOT with ANNOTATE Data Set



## PROC GREPLAY

PROC GREPLAY combines multiple graphs into a single graph. Given a catalog of graphs, it allows the user to select from the available graphs and position them on a single page according to a template. The format of the graph output will be as specified by the current GOPTIONS settings. Parameters that can be specified in a GREPLAY statement are the graphics catalog containing the available graphs (IGOUT), the graphics catalog to receive output graphs (GOUT) if different, the catalog of templates (TC), and the template (TEMPLATE) to be used. The template defines positions on the page for the #1 graph, the #2 graph, etc., for as many graphs as are defined. When GREPLAY is run in an interactive mode, the user assigns graphs from the displayed list to a position on the template by entering the number of the template graph beside the graph on the list.

A set of templates is provided in SASHELP and can be accessed by the following PROC GREPLAY statement –

```
PROC GREPLAY igout=work.gseg tc=sashelp.templt template=L2R2S ;  
run ;
```

In this example, graphs from work.gseg (the default graphics catalog) will be presented in a list. The template L2R2S (Left-2-Right-2 with Space) is for presenting four graphs in a 2 by 2 layout, with space between them. The user would enter the numbers 1, 2, 3, and 4 beside the graphs in the list to be placed in the first, second, third, and fourth template positions. Note that graphs are named according to the procedure that produced them, followed by a numerical index. Thus GPLOT, GPLOT1, GPLOT2, and GPLOT3 would be the names of the first four graphs produced by PROC GPLOT. In the interactive mode, PROC GREPLAY also presents a brief description of each graph.

PROC GREPLAY can also be operated in a non-interactive mode (specify NOFS for no full-screen), using the REPLAY (replay a single graph) or TREPLAY (replay a template) statements as follows –

```
PROC GREPLAY igout=work.gseg tc=sashelp.templt template=L2R2S nofs ;  
TREPLAY 1=GPLOT 2=GPLOT1 3=GPLOT2 4=GPLOT3 ;  
run ;
```

A handy statement to use with GREPLAY is DELETE \_ALL\_, especially for non-interactive use. Running GREPLAY as shown below before producing graphs ensures that your graphs will have the names you expect.

```
PROC GREPLAY IGOUT=WORK.GSEG nofs ;  
delete _all_ ;  
run ;
```

## USER DEFINED TEMPLATES

While SAS does provide some templates (and the user can select replayed graphs to place in templates as well), the capability of effectively displaying a large number of graphs

requires creating templates. They are created in PROC GREPLAY using the TDEF statement to specify the x and y coordinates (in percent units) of the four corners (ll=lower left, lr=lower right, ul=upper left, and ur=upper right) of the graph. The following GREPLAY call shows how to create a template and use it to combine 6 graphs.

```
PROC GREPLAY igout=work.gseg nofs ;
tc work.templt ;
tdef Box2by3 des='2 By 3 Box'
  1/llx=0    ly=50  lrx=33.33 lry=50  ulx=0    uly=100  urx=33.33 ury=100
  2/llx=33.33 lly=50  lrx=66.67 lry=50  ulx=33.33 uly=100  urx=66.67 ury=100
  3/llx=66.67 lly=50  lrx=100  lry=50  ulx=66.67 uly=100  urx=100  ury=100
  4/llx=0    lly=0   lrx=33.33 lry=0   ulx=0    uly=50   urx=33.33 ury=50
  5/llx=33.33 lly=0   lrx=66.67 lry=0   ulx=33.33 uly=50   urx=66.67 ury=50
  6/llx=66.67 lly=0   lrx=100  lry=0   ulx=66.67 uly=50   urx=100  ury=50 ;
template Box2by3 ;
treplay 1=GPLOT 2=GPLOT1 3=GPLOT2 4=GPLOT3 5=GPLOT4 6=GPLOT5 ;
run ;
```

There are other statements and options for replaying graphs, but these are all that is needed to create effective multi-panel displays. It is convenient to save defined templates in a catalog for general use, or at least to %include a SAS job that creates them as needed.

## CREATING EFFECTIVE MULTI-PANEL DISPLAYS

It is clear from the above that graphs can be easily combined on a single page using GREPLAY. A template of as many as 12 graphs in a 3-by-4 layout might be still be legible. But as the number of graphs increases, axis labels, legends and titles become more difficult to read. However, by 1) maximizing the portion of each graph that actually presents data, 2) adding text using the ANNOTATE facility that will be large enough to read, and 3) using uniform scaling on graphs, graphs can be squeezed down even smaller to present more graphs on a single page in a meaningful way.

To maximize the portion of each graph that presents data, all individual titles can be deleted. Use the NOBYLINE option to exclude By-variable titles. Also, use the NOLEGEND option to suppress legends, since they will be too small to read. But think about how necessary information could be included in an overall title, or by using meaningful colors. Be creative!

As we have seen, it is fairly straightforward to add text to graphs using the ANNOTATE facility. Simply adding a subject number (or machine ID or city or other identifier) in the corner of each graph is the simplest way to preserve sample IDs. Such an ANNOTATE data set can be created fairly easily from the data set itself by using FIRST. variables and assigning the TEXT variable to the desired identifier or combination of identifiers. Finally, uniform scaling is easy (at least for GPLOT) either by using the UNIFORM option on the GPLOT statement, or specifying the range of values for both axes.

## Example 2

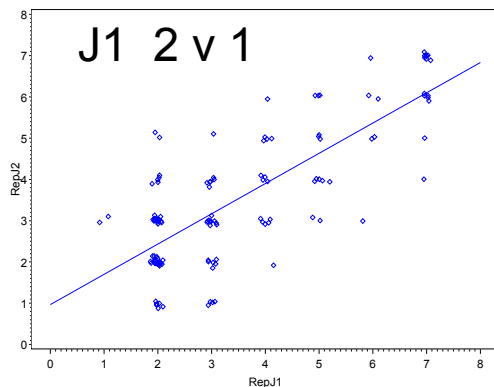
Each of 16 judges at the county fair has rated 12 apple pies 3 times for 11 different characteristics, each using a 7 point scale. I want to create a matrix of scatterplots comparing ratings for the same judge for his three replicates. The data set has variables Judge, RepJ1, RepJ2, and RepJ3 (as well as pie number and characteristic) which represent the ratings for a particular pie and characteristic for the three replicates. If a judge were consistent, his three sets of ratings would be identical. I added “jitter” to the values since they are on a 7 point scale and I want to minimize the effects of overprinting. Since I need to use three different labels and am actually graphing different pairs of variables, I created three similar annotate data sets. Note that the data set is used to create the ANNOTATE data sets.

```
data AnnoInd1 AnnoInd2 AnnoInd3 ; set pies ; by judge ;
retain jnum 0 ;
if first.judge ;
jnum=jnum+1 ;
xsys='1' ; ysys='1' ; function='label' ; color='black' ;
size=4 ; position='c' ; x=10 ; y=85 ;
text='J'||left(put(jnum,2.))||' 2 v 1' ; output AnnoInd1 ;
text='J'||left(put(jnum,2.))||' 3 v 1' ; output AnnoInd2 ;
text='J'||left(put(jnum,2.))||' 3 v 2' ; output AnnoInd3 ;
run ;
```

```
proc gplot data=pies ;
plot RepJ2*RepJ1 / vaxis=axis4 haxis=axis5 anno=annoind1 ;
plot RepJ3*RepJ1 / vaxis=axis4 haxis=axis5 anno=annoind2 ;
plot RepJ3*RepJ2 / vaxis=axis4 haxis=axis5 anno=annoind3 ;
symbol1 c=blue v=diamond i=r1 ;
axis4 label=(a=90) order=0 to 8 ;
axis5 offset=(5,5) order=0 to 8 ;
options nobyline ;
by judge ;
run ;
```

```
proc gslide ;
title1 c=blue 'Judge Replicate Scatterplots' ;
run ;
```

Figure 2. The First Graph – First Judge, Rep 2 vs. Rep 1

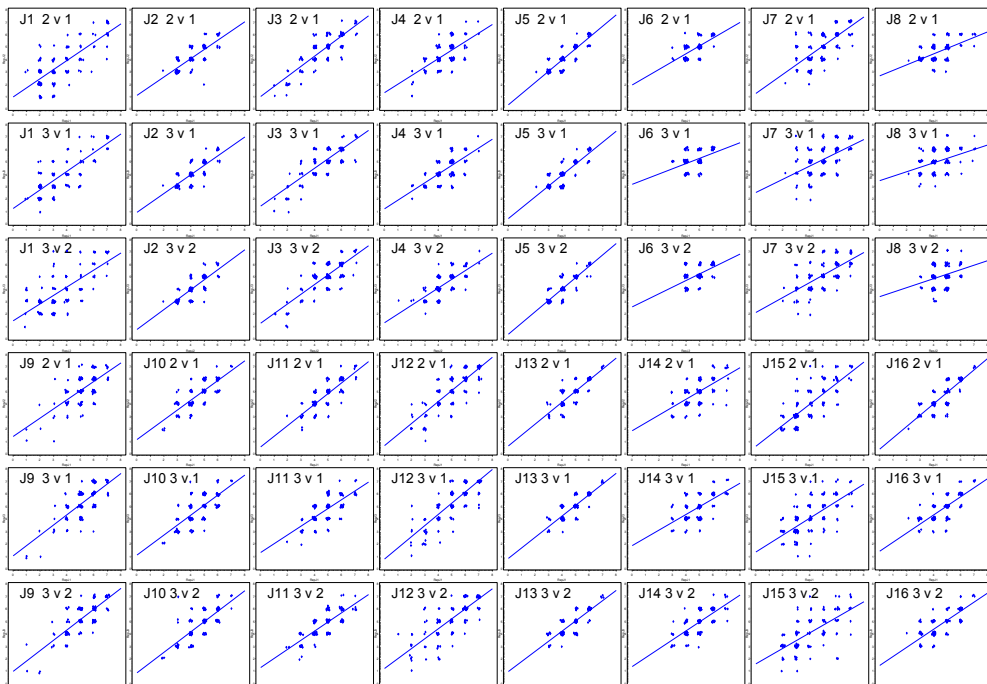


PROC GPGLOT has created three graphs like the one in Figure 2 for each of the 16 judges (=48 graphs), and PROC GSLIDE has created an overall title. Now to create a multi-panel display of these graphs, use a 6-by-8 template plus a panel for the overall title. In this template, the panels are defined as equal size blocks for 8 blocks across the row (0 to 100%) and 6 blocks high (0 to 90%). The panel for the overall title is simply 0 to 100% in both directions (overlapping the area for the other graphs) to maintain its size and aspect ratio. Note that the panels in the template are ordered across the rows, while the order of the graphs in the catalog are by judge – this makes setting up the TREPLAY statement a little bit tricky. Figure 3 presents the resulting multi-panel display.

```
proc greplay igout=work.gseg tc=work.templt template=box6by8t nofs ;
treplay 1=gplot 9=gplot1 17=gplot2 2=gplot3 10=gplot4 18=gplot5
      3=gplot6 11=gplot7 19=gplot8 4=gplot9 12=gplot10 20=gplot11
      5=gplot12 13=gplot13 21=gplot14 6=gplot15 14=gplot16 22=gplot17
      7=gplot18 15=gplot19 23=gplot20 8=gplot21 16=gplot22 24=gplot23
      25=gplot24 33=gplot25 41=gplot26 26=gplot27 34=gplot28 42=gplot29
      27=gplot30 35=gplot31 43=gplot32 28=gplot33 36=gplot34 44=gplot35
      29=gplot36 37=gplot37 45=gplot38 30=gplot39 38=gplot40 46=gplot41
      31=gplot42 39=gplot43 47=gplot44 32=gplot45 40=gplot46 48=gplot47
      49=gslide ;
run ;
quit ;
```

Figure 3. Multi-Panel Display of Judges' Replicate Ratings

### Judge Replicates Scatterplots



Despite the small scale, there is much information to be gleaned from this display. The ANNOTATE labels identify each graph. The three graphs for each judge are arranged vertically, with graphs for the first eight judges comprising the top three rows of the

display, and graphs for the other eight judges filling the bottom three rows. Adding a regression line to each graph provides a visual test of consistency. The lines for judges 6 and 8 don't seem to have as great a slope as the other judges. Also judge 6 seems to use only a small part of the scale no matter what he is rating, while judges 3 and 11 use more of the scale and are more consistent with their ratings. These are the kinds of things that you just don't know without looking at graphs. But most of us would be overwhelmed after looking at 48 graphs (one per page), and come away with little insight.

### Example 3

Now compare judges by plotting each judge's average ratings across three replicates against the average ratings of each other judge. This scatterplot matrix will include 120 graphs. The data set of averages originally contained one variable for each judge's averages, plus variables identifying the pie and characteristic rated. To produce graphs for all possible pairs of judges, the data were reformatted to define variables that contain the rating for the judge on the x-axis (x) and the rating for the judge on the y-axis (y), plus a variable to keep track of the variable pair. Also, variable PairDes is created, containing the label to be used on the graph.

```
data pie3a ; set pie3 ;
length pairdes $ 7 ;
array a[16] judgej1-judgej16 ;
pair=0 ;
do i=1 to 15 ;
  x=a[i] ;
  do j=i+1 to 16 ;
    pair=pair+1 ;
    y=a[j] ;
    PairDes=trim(left(put(j,2.)))||' v '||left(put(i,2.)) ;
  output ;
end ; end ;
run ;
```

The annotate data set to be used with these graphs is obtained from the data set as before.

```
data annoavg ; set pie3a ; by pair ;
if first.pair ;
xsys='1' ; ysys='1' ; function='label' ; color='black' ;
size=6 ; position='c' ; x=5 ; y=85 ; text=pairdes ; output ;
run ;
```

The individual graphs are created by pair.

```
options nobyline ;
proc gplot data=pie3a ;
plot y*x=1 / vaxis=axis4 haxis=axis5 anno=annoavg ;
symbol1 c=blue v=diamond i=r1 ;
axis4 label=(a=90) order=0 to 8 ;
axis5 offset=(5,5) order=0 to 8 ;
by pair ;
title1 ;
run ;
```

```

proc gslide ;
title1 c=blue "Scatterplot Matrix of Judge Means" ;
run ;

```

Finally, the multi-panel display is created using GREPLAY with a template called LOW15 (lower triangular matrix with 15 rows), which has graphs ordered down the first column, then down the second column, etc. This makes for a simpler assignment of graphs to template positions, especially with 121 template positions. Figure 4 presents the resulting multi-panel display.

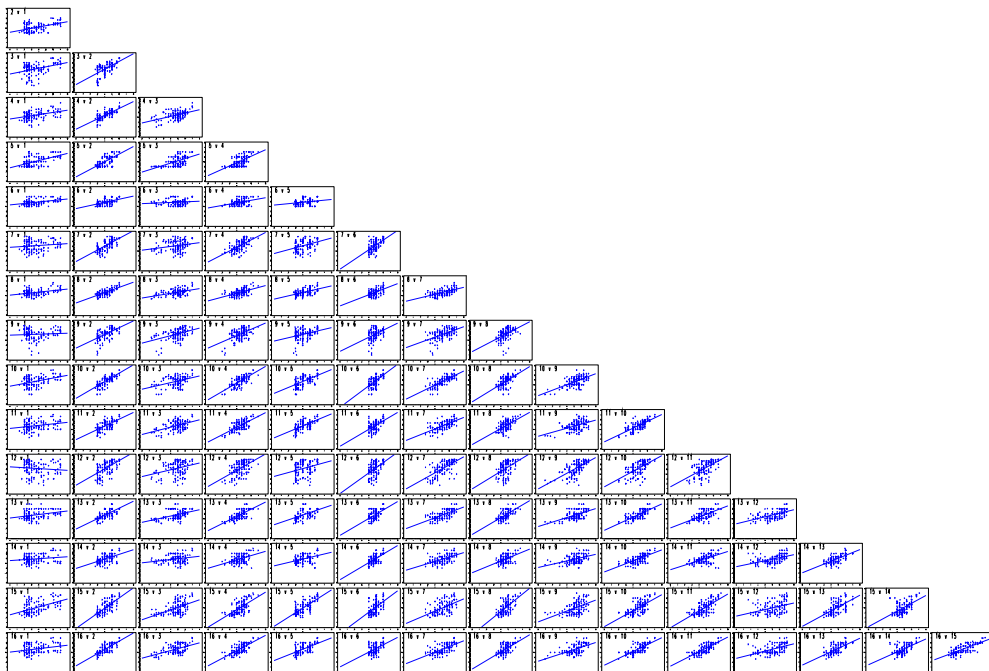
```

proc greplay igout=work.gseg tc=work.templt template=Low15 nofs ;
treplay 1=gplot 2=gplot1 3=gplot2 4=gplot3 5=gplot4 6=gplot5 7=gplot6 8=gplot7
9=gplot8 10=gplot9 11=gplot10 12=gplot11 13=gplot12 14=gplot13 15=gplot14 16=gplot15 17=gplot16
18=gplot17 19=gplot18 20=gplot19 21=gplot20 22=gplot21 23=gplot22 24=gplot23 25=gplot24
26=gplot25 27=gplot26 28=gplot27 29=gplot28 30=gplot29 31=gplot30 32=gplot31 33=gplot32
34=gplot33 35=gplot34 36=gplot35 37=gplot36 38=gplot37 39=gplot38 40=gplot39 41=gplot40
42=gplot41 43=gplot42 44=gplot43 45=gplot44 46=gplot45 47=gplot46 48=gplot47 49=gplot48
50=gplot49 51=gplot50 52=gplot51 53=gplot52 54=gplot53 55=gplot54 56=gplot55 57=gplot56
58=gplot57 59=gplot58 60=gplot59 61=gplot60 62=gplot61 63=gplot62 64=gplot63 65=gplot64
66=gplot65 67=gplot66 68=gplot67 69=gplot68 70=gplot69 71=gplot70 72=gplot71 73=gplot72
74=gplot73 75=gplot74 76=gplot75 77=gplot76 78=gplot77 79=gplot78 80=gplot79 81=gplot80
82=gplot81 83=gplot82 84=gplot83 85=gplot84 86=gplot85 87=gplot86 88=gplot87 89=gplot88
90=gplot89 91=gplot90 92=gplot91 93=gplot92 94=gplot93 95=gplot94 96=gplot95 97=gplot96
98=gplot97 99=gplot98 100=gplot99 101=gplot100 102=gplot101 103=gplot102 104=gplot103
105=gplot104 106=gplot105 107=gplot106 108=gplot107 109=gplot108 110=gplot109 111=gplot110
112=gplot111 113=gplot112 114=gplot113 115=gplot114 116=gplot115 117=gplot116 118=gplot117
119=gplot118 120=gplot119 121=gslide ;

```

Figure 4. Multi-Panel Display ScatterPlot Matrix

## Scatterplot Matrix of Judge Means



Again, despite the small size of each graph, some judges can be identified as different from the others. We can see again that Judge 6 does not use much of the scale. Judge 1 has very noisy graphs, indicating that his ratings are not consistent with the other judges. The simple label on each graph makes clear which judges are being compared on each graph.

#### Example 4

This example shows results of fabricated measurements taken at a series of time points among 75 subjects while testing two products. There are four different groups of subjects, but I want to see the individual data to determine if any subjects display especially unusual behavior. The data set includes columns for subject number (GroupID), time, and response. The ANNOTATE data set is created from the data set.

```
data anno1 ; set response ; by groupid ;
if first.groupid ;
xsys='1' ; ysys='1' ; x=15 ; y=90 ; function='label' ; position='6' ; size=7 ;
text=groupid ;
keep groupid xsys ysys function x y text position size ;

proc gplot data=response uniform ;
plot response*time=product / vaxis=axis2 haxis=axis5 nolegend anno=anno1 ;
by groupid ;
title1 ;
axis2 label=(a=90) ;
axis5 offset=(5,5) order=-5 to 10 by 5 label=(h=1 'Minutes') ;
symbol1 c=blue v=dot r=1 h=2 i=join l=2 w=3 ;
symbol2 c=red f= v=dot r=1 h=2 i=join l=2 w=3 ;
options nobyline ;
run ;

proc gslide ;
title1 c=blue 'Individual Data - Blue = A,' c=red 'Red = B' ;
run ;
```

Thus I have created a  $y \times x = z$  graph for each of the 75 subjects, and labeled it with the subject number (which includes group identification). I use a 7-by-12 template with an overall title for our multi-panel display as indicated in the PROC GREPLAY code.

```
proc greplay igout=work.gseg tc=work.templt template=box7by12 nofs ;
treplay 1=gplot 2=gplot1 3=gplot2 4=gplot3 5=gplot4 6=gplot5 7=gplot6
8=gplot7 9=gplot8 10=gplot9 11=gplot10 12=gplot11 13=gplot12
14=gplot13 17=gplot14 18=gplot15 19=gplot16 20=gplot17
21=gplot18 22=gplot19 23=gplot20 24=gplot21 25=gplot22
26=gplot23 27=gplot24 28=gplot25 29=gplot26 30=gplot27
31=gplot28 32=gplot29 33=gplot30 34=gplot31 37=gplot32
38=gplot33 39=gplot34 40=gplot35 41=gplot36 42=gplot37
43=gplot38 44=gplot39 45=gplot40 46=gplot41 47=gplot42
48=gplot43 49=gplot44 50=gplot45 51=gplot46 52=gplot47
53=gplot48 54=gplot49 55=gplot50 56=gplot51 59=gplot52
60=gplot53 61=gplot54 62=gplot55 63=gplot56 64=gplot57
65=gplot58 66=gplot59 67=gplot60 68=gplot61 69=gplot62
70=gplot63 71=gplot64 72=gplot65 73=gplot66 74=gplot67
```

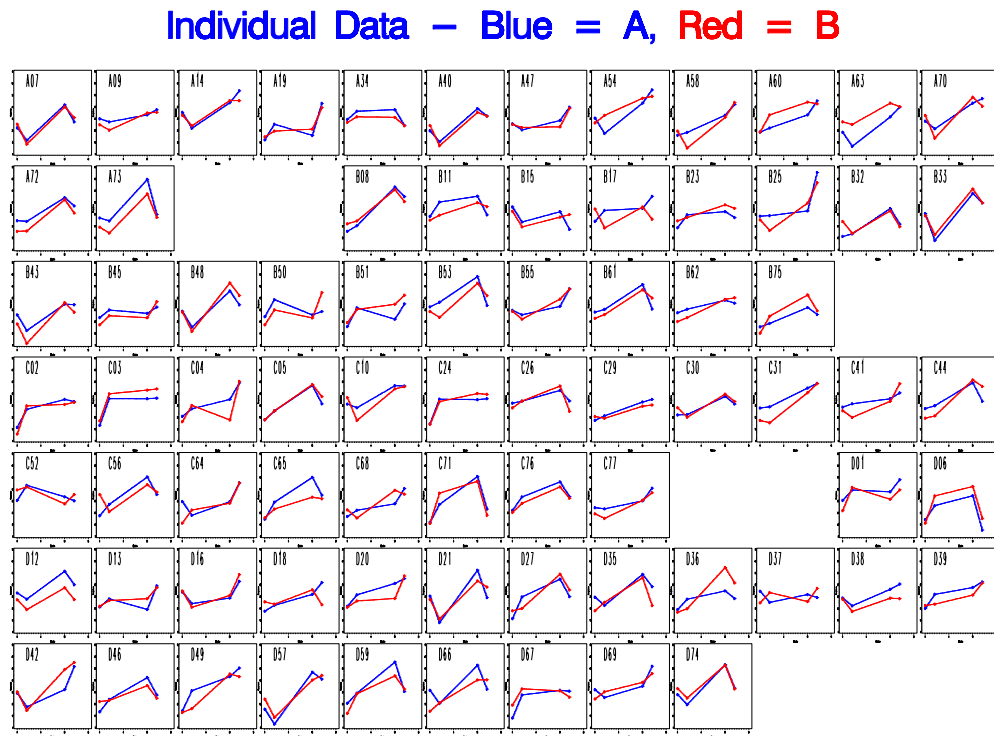
```

75=gplot68 76=gplot69 77=gplot70 78=gplot71 79=gplot72
80=gplot73 81=gplot74
85=gslide ;
run ;

```

However, as Figure 5 below shows, I have done something further to identify the four groups – there are two blank panels after the last subject in each group. In addition, I included a “legend” in the title, making it clear which product corresponds to which color line.

Figure 5. Multi-Panel Display of Product Comparisons by Subject



Again, despite the small size of each graph, this presentation allows us to identify apparent bad data and outlier subjects, and look for consistency of differences across subjects.

## CONCLUSIONS

With data sets becoming larger all the time, it is more important than ever to be confident that there are no surprises lurking in them. As we have seen, combination of the ANNOTATE facility with user defined templates in PROC GREPLAY allows the user to create useful multi-panel displays for effective visual review of large subsets of data.

SAS® is a registered trademark of SAS Institute, Inc.