Paper 192–31

Creating Statistical Graphics in SAS® 9.2: What Every Statistical User Should Know

Robert N. Rodriguez and Tonya E. Balan
SAS Institute Inc.
Cary, North Carolina, USA

ABSTRACT

Effective graphics are essential to modern statistical analysis. SAS 9.1 introduced an experimental extension to the Output Delivery System (ODS), which is used by over two dozen SAS/STAT and SAS/ETS procedures to create statistical graphics as automatically as they create tables. This extension, referred to as ODS Statistical Graphics (or ODS Graphics for short), requires minimal additional syntax, and it provides displays commonly needed for data analysis and statistical modeling, including scatter plots, histograms, and box-and-whisker plots. General ODS features, such as styles and destination statements, apply equally to tables and graphs.

With the production release of ODS Graphics in SAS 9.2, many more statistical procedures have been modified to use this new functionality; see Appendix A for a list. For example, PROC LOGISTIC produces effects plots, and PROC RSREG produces contour and ridge plots. New SAS/GRAPH procedures, as well as existing SAS/QC procedures, also take advantage of ODS Graphics functionality. SAS 9.2 introduces an interactive graphics editor, with which you can make immediate, data-specific changes

to your graphs, such as customizing titles and annotating points. SAS 9.2 also adds ODS styles that are designed specifically for statistical work. In addition to describing this new work, this paper presents the basics of creating, modifying, and managing graphics.

Note: At the time of this writing, some of the features described in this paper were being completed. When SAS 9.2 is released, an updated version of this paper will be provided in the Papers section at http://www.sas.com/statistics.

INTRODUCTION

Effective graphics are indispensable for modern data analysis, statistical modeling, and data mining. They reveal patterns, identify differences, and express uncertainty that is not readily apparent in tabular output. Effective graphics add visual clarity and rich content to analytical presentations, provoking questions and stimulating deeper investigation.

In the past, creating graphics with statistical procedures in SAS has generally required additional programming. For example, plotting survival curves for several risk groups with the LIFETEST procedure in SAS Version 8 involves three steps: creating output data sets with curve values and metadata such as p-values; modifying the data sets programmatically; and finally, using SAS/GRAPH procedures with the annotate facility to render the plot.

Recall that the Output Delivery System (ODS), added in Version 8, manages all tabular output created by procedures and enables you to display it in a variety of destinations, such as HTML and RTF. SAS 9.1 introduced an extension to ODS, referred to as ODS Graphics, which—together with corresponding modifications to statistical procedures—equips these procedures to create graphics as automatically as tables. This eliminates the need for additional programming. ODS Graphics was introduced as experimental software in SAS 9.1 (see Rodriguez 2004) and will be production software in SAS 9.2.

Note: In SAS 9.2, SAS/GRAPH is required for ODS Graphics functionality in Base SAS, SAS/STAT, SAS/ETS, SAS High Performance Forecasting, and SAS/QC.

ODS Graphics is invoked when you provide the statement

```
ods graphics on;
```

prior to your procedure statements. Any procedures that have been modified to use ODS Graphics then create graphics as part of their output, either by default or when you specify procedure options for requesting specific graphs.

The following statements illustrate how you can create a survival plot with ODS Graphics in SAS 9.2.

```
ods html;
ods graphics on;

proc lifetest data=BMT plots=survival(hwb test=logrank);
    time T * Status(0);
    strata Group;
run;

ods html close;
ods graphics off;
```

The ODS HTML statement specifies an HTML destination for the output. The ODS GRAPHICS statement requests ODS Graphics in addition to the usual tabular output, and the PLOTS= option requests the survival plot, which is shown in Figure 1. Note that the complete HTML output, which is not shown here, integrates graphs and tables. The ODS GRAPHICS OFF statement disables ODS Graphics, and the ODS HTML CLOSE statement closes the HTML destination.
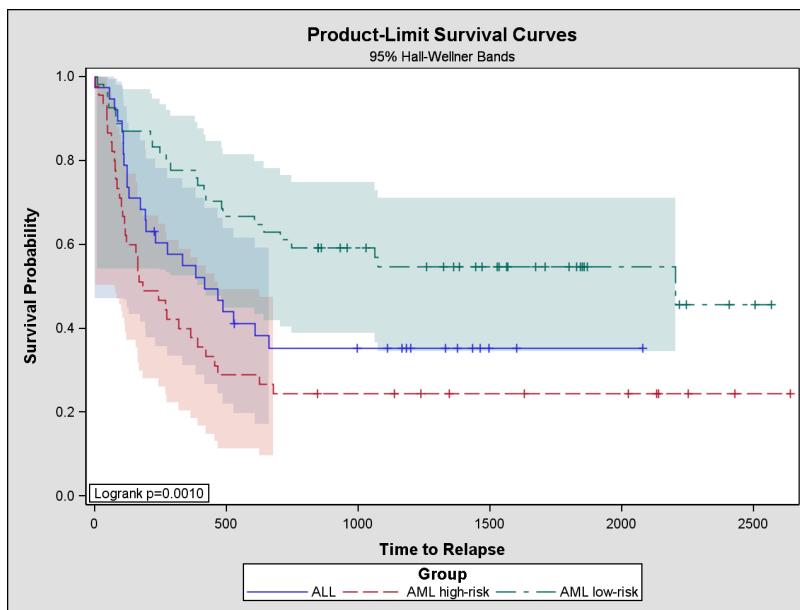

Figure 1. Survival Plot Produced by LIFETEST Procedure

With ODS Graphics, procedures create graphs that are commonly needed for a particular analysis—either by default or when you specify options, as in the preceding example. ODS Graphics enables a procedure to take full advantage of its computational results, and consequently procedures can enrich graphs with useful statistical information and metadata, such as sample sizes and p-values.

The next section provides a primer on ODS Graphics, which explains the basics of creating and managing graphics. Subsequent sections describe and illustrate new functionality in SAS 9.2.

THE ABC'S OF ODS GRAPHICS

To invoke ODS Graphics with the procedures that support it, you specify the statement

```
ods graphics on;
```

before the procedure statement. To turn off this facility, you specify

```
ods graphics off;
```

Once you have turned on ODS Graphics, creating graphics with procedures is as simple as requesting tabular output. You use

- ODS destination statements (such as ODS HTML) to specify *where you want your graphs to be displayed*. Figure 2 illustrates HTML output. See Appendix C for a list of destinations supported in SAS 9.2.
- procedure options and defaults to specify *which graphs you want*. For each procedure, these options are described in the Syntax section of the procedure chapter in the forthcoming *SAS/STAT 9.2 User's Guide*, and the graphs available are listed in the Details section titled "ODS Graph Names".
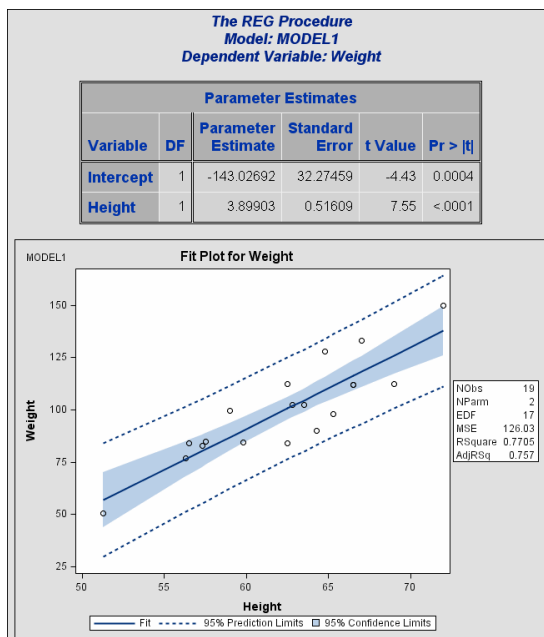

Figure 2.   HTML Output from PROC REG with Default Style

As with ODS tables, you can use

- graph names in ODS SELECT and ODS EXCLUDE statements to select and exclude graphs from your output. Appendix B explains how to look up the name for a graph.
- ODS styles to control the general appearance and consistency of all graphs and tables.

3

Viewing Your Graphs with Different ODS Destinations

For almost all ODS destinations, including HTML, RTF, and PDF, graphs and tables are integrated in the output, as illustrated for HTML in Figure 2. The exception is the default LISTING destination, which is now supported by ODS Graphics in SAS 9.2. If you are using the LISTING destination and running SAS 9.2 in the SAS windowing environment (Display Manager), you view your graphs individually by clicking on the graph icons in the Results output tree, as illustrated in Figure 3. This action invokes a host-dependent graph viewer (for instance, Microsoft Photo Editor on Windows). Note that graphs are *not* displayed with traditional GRSEG graphs in the Graph window.
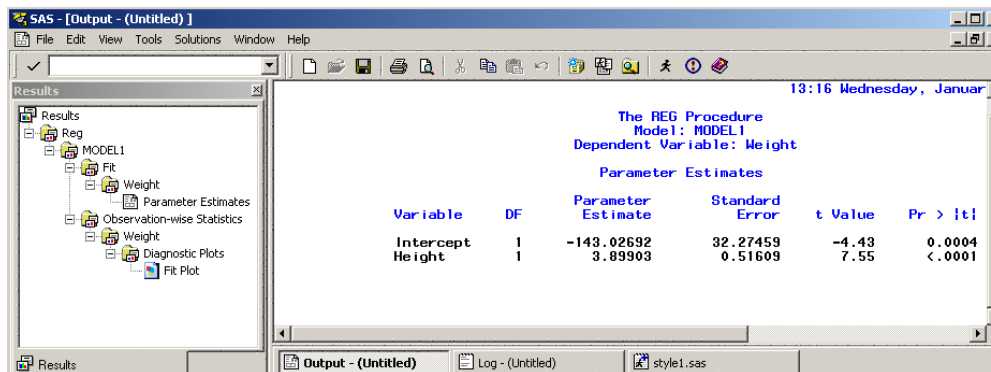

Figure 3.   Results Output Tree in SAS Display Manager

If you are using Display Manager on Windows and prefer to view integrated output, you should specify a destination such as HTML or RTF. You can prevent the Output window from appearing by closing the LISTING destination with the statement

```
ods listing close;
```

Accessing Individual Graphs

If you are writing a paper or creating a presentation, you will need to access your graphs individually. There are various ways to do this, depending on the ODS destination. Two particularly useful methods are the following:

- If you are viewing RTF output on Windows, you can simply copy and paste your graphs from the viewer into a Microsoft Word document or a Microsoft PowerPoint slide.
- You can save your graphs in files and then include them in a document. For instance, using the LATEX destination, you can save your graphs as PNG or PostScript files to include in a paper that you are writing with LATEX. Similarly, using the HTML destination, you can save your graphs in GIF files to include in an HTML document.

   You specify the graphics image format and the file name in the ODS GRAPHICS statement. For example, the statements

   ```
   ods latex;
   ods graphics on / imagefmt=ps imagename='myname';
   ```

   create graphs in PostScript files named myname.ps, myname1.ps, and so on. Appendices C, D, and E provide details about the file types available with various destinations, how they are named, and how they are saved.

4

Specifying the Resolution and Size of Graphs

Two factors to consider when you are creating graphs for a document are the resolution of the graph and its dimensions. The default resolution of graphs created with HTML and LISTING is 100 dpi (dots per inch), whereas the default with RTF is 200 dpi. Because of its higher resolution, RTF is recommended if you are cutting and pasting graphs into a Microsoft PowerPoint presentation or a Microsoft Word document.

You can increase the resolution with the IMAGE_DPI= option in the ODS destination statement, as illustrated by the following statement.

```
ods html image_dpi=300;
```

The increase in resolution between 100 dpi and 200 dpi is noticeable, but beyond 200 dpi it is offset by the sharply increased size of the image file. If you increase the resolution, you may need to compensate by reducing the graph in the container document.

The default dimensions of graphs produced with ODS Graphics are a width of 640 pixels and a height of 480 pixels, and these values determine the default aspect ratio. You can specify the dimensions with the WIDTH= and HEIGHT= options in the ODS GRAPHICS statement, as in the following example.

```
ods graphics on / width=4.5 in  height=3.5 in;
```

If you specify only one dimension, the other is determined by the default aspect ratio. For best results, you should create your graphs in the dimensions displayed in your document.

By default in SAS 9.2, fonts and symbol markers are automatically scaled with the size of the graph. You can suppress this scaling with the SCALE= option, as in the following statement.

```
ods graphics on / scale=off;
```

Appendix F provides a more detailed example of graph resizing and font scaling.

Creating Graphs with Style

ODS styles control the overall visual appearance of graphs and tables in your output, and they specify colors, fonts, line styles, and other attributes of graph elements. SAS 9.1 introduced several styles for statistical output, and additional styles are planned for SAS 9.2. Designed by user interface specialists at SAS, these styles achieve consistency and professional appearance of statistical graphics by providing appropriate default attributes for common graphics elements, such as fitted lines and confidence bands.

In SAS 9.2, the following styles are recommended for statistical work:

- the DEFAULT style for each destination, which is a color style intended for general-purpose work; see Figure 2 for an example. This style is a compromise in the sense that some graph elements are intentionally overdistinguished to facilitate black-and-white printing. For instance, fit lines corresponding to different classification levels are distinguished by both colors and line patterns.

- the STATISTICAL family of styles (named STATISTICAL, STATISTICAL2, . . .) are color styles recommended for output in Web pages or color print media. STATISTICAL styles may not necessarily print well on black-and-white devices. See Figure 4 for an example.

- the JOURNAL family of styles (named JOURNAL, JOURNAL2, . . .) are black-and-white or gray-scale styles recommended for journal articles and other publications that are printed in black and white. See Figure 5 for an example.
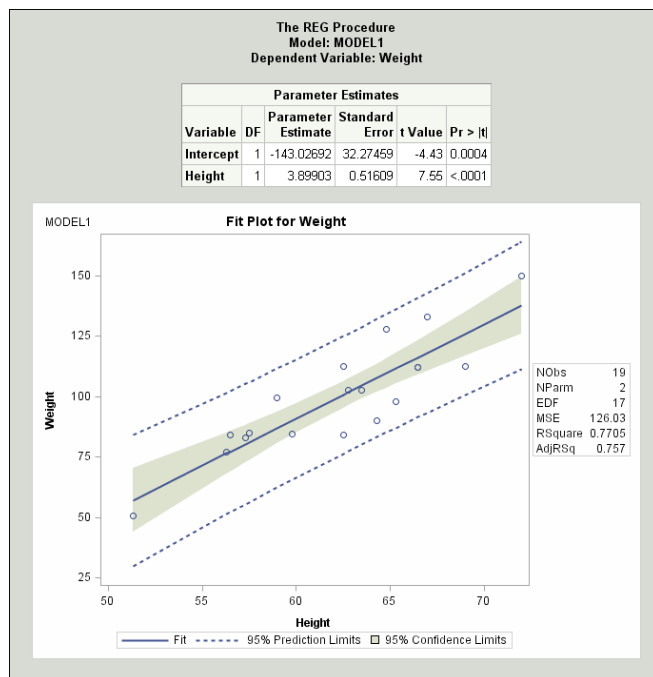
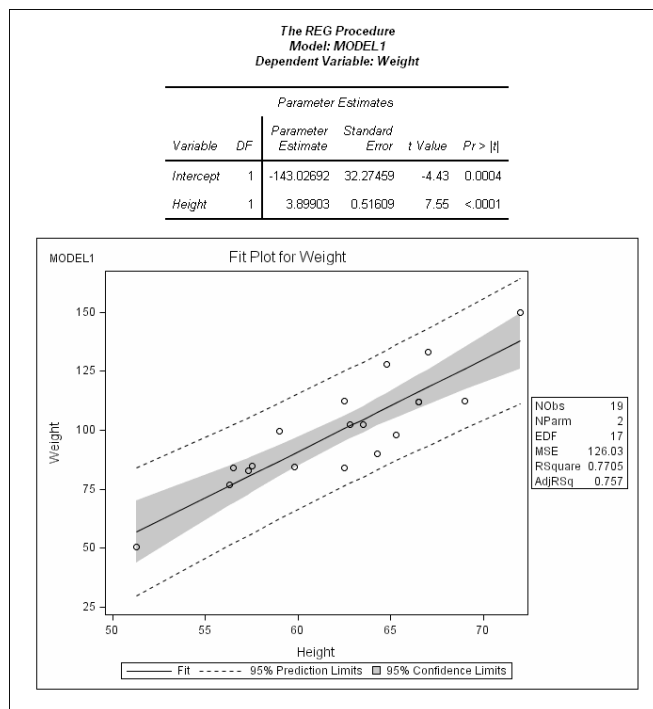**Figure 4.** HTML Output from PROC REG with Statistical Style

Figure 5. HTML Output from PROC REG with Journal Style

You specify a style with the STYLE= option in the ODS destination statement. For example, the following statement requests HTML output using the Journal style.

```
ods html style = Journal;
```

Note that the STYLE= option in the ODS LISTING statement applies only to graphs; the legacy SAS monospace format is used for output tables.

You can modify the colors, fonts, and other attributes of graph elements in a SAS-supplied style by editing the style definition file. This is explained in the Base SAS documentation for the TEMPLATE procedure.

Modifying Your Graphs

Although ODS Graphics is designed to automate the creation of high-quality statistical graphics, modifications are inevitable. SAS 9.2 provides you with two ways to modify your graphs, depending on whether the changes you want to make are data-dependent and *immediate* (for a specific graph you are preparing for a paper or presentation), or whether they are *persistent* (to be applied to a graph each time you run the procedure).

For making immediate changes, SAS 9.2 introduces the ODS Graph Editor, which you can use to modify your graphs with a point-and-click interface. You can invoke the Graph Editor from Display Manager by selecting the Edit option for a graph icon in the Results window.

Figure 6 illustrates the Graph Editor invoked for a graph created by the ROBUSTREG procedure. Figure 7 illustrates modifications made with tools in the Graph Editor. The title has been changed to incorporate the name of the data, and an outlier has been highlighted with a custom label.

After you have modified a graph with the Graph Editor, you can save it as a graphics image file. Alternatively, you can save the graph as an XML file, which allows you to resume editing it at a later time. You can edit an image file with the Graph Editor, but it does not preserve the context of the XML representation.

To make persistent, data-independent changes, you can modify the ODS graph template for a particular plot. A graph template is a program, written in the Graph Template Language (GTL) and pre-compiled with PROC TEMPLATE, which specifies the layout and details of a graph. Because SAS provides a default template for each graph created by a procedure, *it is completely unnecessary for statistical users to have knowledge of templates in order to create statistical graphics.* However, a SAS programmer who is familiar with the GTL can access, customize, and save templates to make changes that remain in effect each time the procedure is run.

Note that the experimental version of the GTL that was introduced in SAS 9.1 has been substantially modified and extended in SAS 9.2. The topic of writing templates is outside the scope of this paper, but more information can be found in a SUGI 31 tutorial paper by Cartier (2006).

For More Information

If you are trying out ODS Graphics for the first time, begin with the graphics examples in the chapters of the documentation for procedures that use ODS Graphics. To take full advantage of ODS Graphics, you will want to learn more about graph management, which is explained in the chapter on "Statistical Graphics Using ODS" in the forthcoming *SAS/STAT 9.2 User's Guide*.
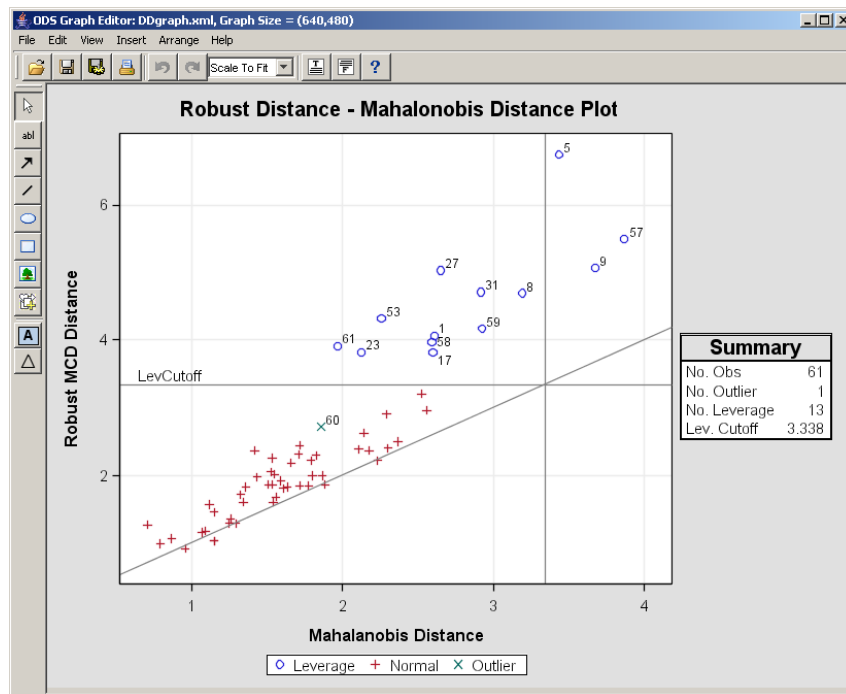
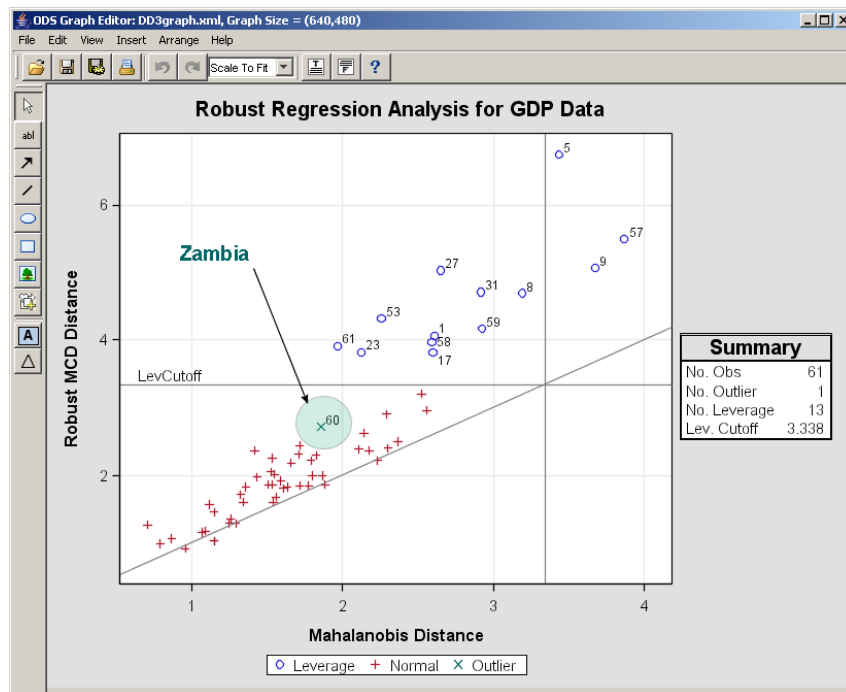**Figure 6.** Graph Editor Invoked with Graph from PROC ROBUSTREG

Figure 7. Point-and-Click Modifications Made with Graph Editor

WHAT'S NEW IN SAS 9.2?

ODS Graphics was introduced as experimental functionality in SAS 9.1, where it was used by approximately 30 SAS/STAT and SAS/ETS procedures. In SAS 9.2, ODS Graphics is production software, and coverage has grown to over 50 procedures; see Appendix A for a list. Many plots have been added or enhanced in response to user requests, including

- plots of odds ratios over strata in the FREQ procedure
- residual, predicted value, and regression diagnostic plots in the GENMOD procedure
- means comparison and analysis of means plots for LS-means in the GLM and GLIMMIX procedures (see Example 3 on page 17)
- plots of Kaplan-Meier curves displaying the number of subjects at risk in the LIFETEST procedure (see Example 2 on page 14)
- plots of smoothed hazard functions for ungrouped data in the LIFETEST procedure
- effects plots for binary and polytomous responses in the LOGISTIC procedure
- ROC plots for comparing models in the LOGISTIC procedure
- adjusted p-value and uniform probability plots in the MULTTEST procedure
- plots for predicted values, residuals, and actual values in the PANEL and TSCSREG procedures
- correlation loadings plots in the PLS procedure
- score plots with confidence and prediction ellipses in the PRINCOMP procedure
- score plots with ID labeling in the PRINCOMP procedure
- residual, diagnostic, ridge, contour, and surface plots in the RSREG procedure (see Figure 8)
- vector time series plots and thumbnail plots in the TIMESERIES procedure
- graphics for Box-Cox smoothing, preference and ideal point mapping, and penalized B-spline smoothing (new in SAS 9.2) in the TRANSREG procedure
- histograms, box plots, plots for paired and crossover data in the TTEST procedure
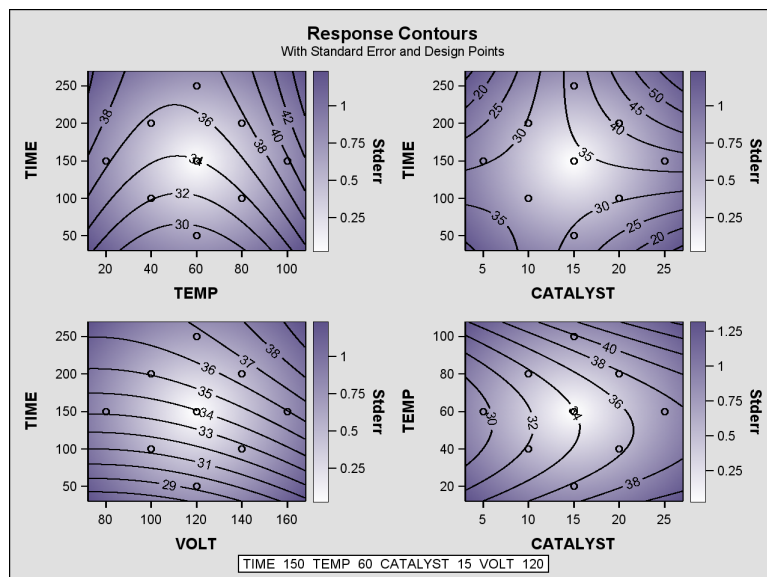

Figure 8. Contour Plots of Surface Slices Created with RSREG Procedure

In SAS 9.2, a consistent PLOTS= procedure syntax has been introduced for requesting optional graphs; this syntax is explained in Example 1 below. Support for ODS Graphics is being added to two Base SAS statistical procedures, FREQ and UNIVARIATE, and is planned for seven SAS/QC procedures: ANOM, CAPABILITY, CUSUM, MACONTROL, PARETO, RELIABILITY, and SHEWHART (see Example 4 on page 20.

In addition to these procedure-level enhancements, the general functionality of ODS Graphics has grown significantly in SAS 9.2 New features include the following:

- The LISTING destination is now supported (see page 4).
- An interactive graph editor has been added for making immediate changes to graphs (see page 7).
- New statistical styles are being introduced (see page 5).
- Fonts are automatically scaled when graphs are resized (see Appendix F).
- Graph sizing options have been added to the ODS GRAPHICS statement (see Appendix F).
- An option has been added to the ODS destination statements for specifying graph resolution; see page 5. You can also specify this in the SAS registry.
- A family of SAS/GRAPH procedures has been introduced that provides general-purpose tools for creating scatter plots (see Example 5 on page 22).
- The SGRENDER procedure has been provided in SAS/GRAPH for creating graphs with the Graph Template Language (see Cartier 2006).

The next five sections of this paper provide a series of examples that illustrate new ODS Graphics features in SAS 9.2.

EXAMPLE 1: GRAPHICS FOR REGRESSION MODELING USING THE GLMSELECT AND REG PROCEDURES

This example illustrates the use of ODS Graphics with two procedures for regression modeling, and it also illustrates the PLOTS= syntax for requesting graphics, which has been introduced in SAS 9.2.

The following data analyzed by Efron et al. (2004) are measurements of ten baseline variables: age, sex, body mass index (BMI), blood pressure (BP), and six blood serum measurements for 442 diabetes patients. The response variable is a measure of disease progression one year after baseline.

```
data Diabetes;
   input Patient Age Sex BMI BP TC LDL HDL TCH LTG GLU Progression;
datalines;
    1   59   2   32.1  101  157   93.2   38   4   4.9   87  151
    2   48   1   21.6   87  183  103.2   70   3   3.9   69   75
    3   72   2   30.5   93  156   93.6   41   4   4.7   85  141
    .
    .       more data lines
    .
  441   36   1   30.0   95  201  125.2   42   5   5.2   85  220
  442   36   1   19.6   71  250  133.2   97   3   4.6   92   57
  ;
```

LASSO selection (Tibshirani 1996) is a recent variable selection algorithm that is implemented in the new GLMSELECT procedure in SAS/STAT; see Cohen (2006) for an introduction to this procedure. The following statements perform LASSO selection to find baseline variables that are important predictors of disease progression.

```
ods html;
ods graphics on;
proc glmselect data=diabetes
               plots=CoefficientPanel(StepAxis=NormBeta);
   model Progression = Age Sex BMI BP TC LDL HDL TCH LTG GLU
               / selection=lasso(stop=none choose=cp);
run;
ods graphics off;
ods html close;
```

Note: All of the graphics examples in this paper open and close an ODS destination, and they all turn ODS Graphics on and off. This is done to make the examples self-contained, but in practice you can place these statements at the top and bottom of your SAS program.

The "Coefficient Panel" in Figure 9, requested with the PLOTS=CoefficientPanel option, helps you visualize the selection process. (For a list of PLOTS= options available with PROC REG, refer to the Syntax section and the "ODS Graph Names" section in the "REG Procedure" chapter of the forthcoming *SAS/STAT 9.2 User's Guide*.) The upper plot in Figure 9 shows how standardized coefficients of selected variables evolve as the selection method proceeds, so that you can assess the relative importance of each of the predictors. The lower plot shows the progression fit criterion, Mallow's C(p), that is minimized by the selected model.
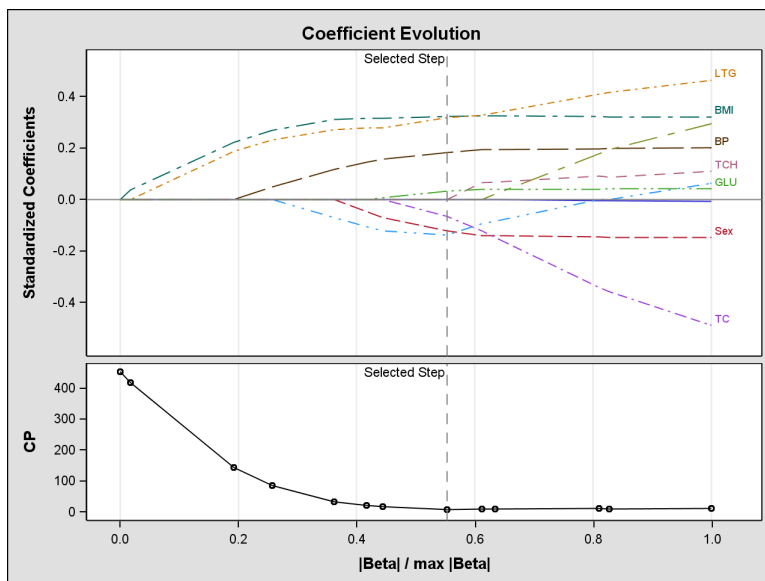

Figure 9.   Coefficient Evolution Plot Created with PROC GLMSELECT

You can use diagnostics plots available with PROC REG to further investigate the model containing the variables selected using LASSO selection. The following statements produce the panel shown in Figure 10.

```
ods html;
ods graphics on;
proc reg data=diabetes plots(only)=DiagnosticsPanel;
   model Progression = Sex BMI BP TC HDL LTG GLU;
run;
ods graphics off;
ods html close;
```

The PLOTS option

```
plots(only)=DiagnosticsPanel
```

specifies that this is the only graphic to be produced. If you omit the ONLY suboption, or if you omit the PLOTS= option entirely, then you obtain all the diagnostic plots that PROC REG produces by default.

In some cases, you may prefer to obtain paneled plots as a series of individual plots. You request this by specifying the UNPACK suboption for each such panel. The following statements show how you use the UNPACK suboption to obtain all the plots shown in Figure 10 as individual plots.

```
ods html;
ods graphics on;
proc reg data=diabetes plots(only)=DiagnosticsPanel(unpack);
    model Progression = Sex BMI BP TC HDL LTG GLU;
run;
ods graphics off;
ods html close;
```
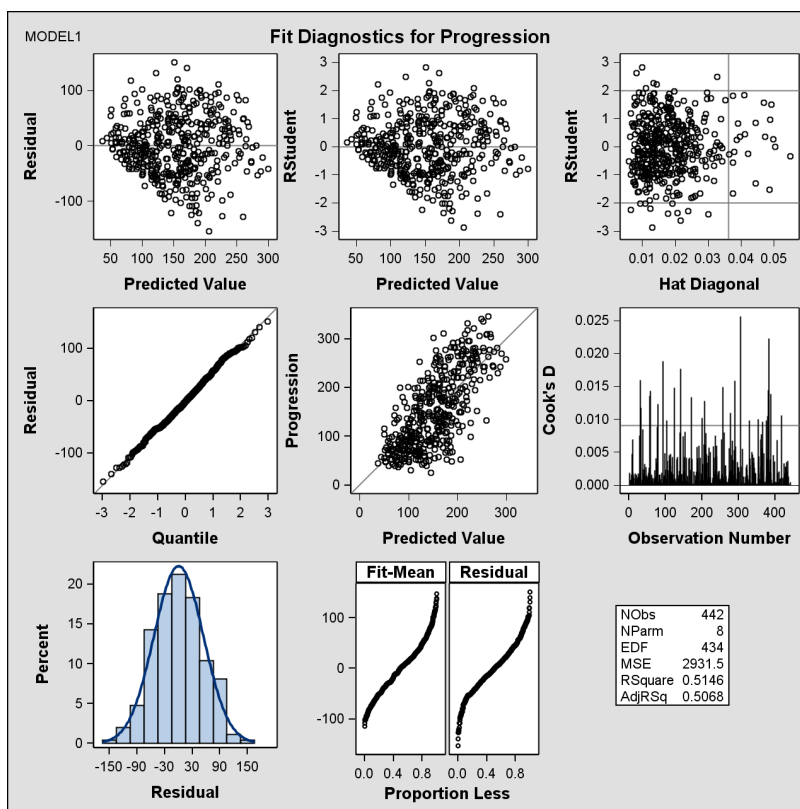

Figure 10. Fit Diagnostics Panel Created with PROC REG

In this case, the UNPACK suboption unpacks the "Diagnostics Panel" into eight individual plots, one of which is shown in Figure 11. Note that if you want to unpack *all* of the paneled displays produced by a procedure, you can simply specify UNPACK as a global PLOTS suboption by providing the keyword UNPACK in parentheses immediately following the keyword PLOTS.
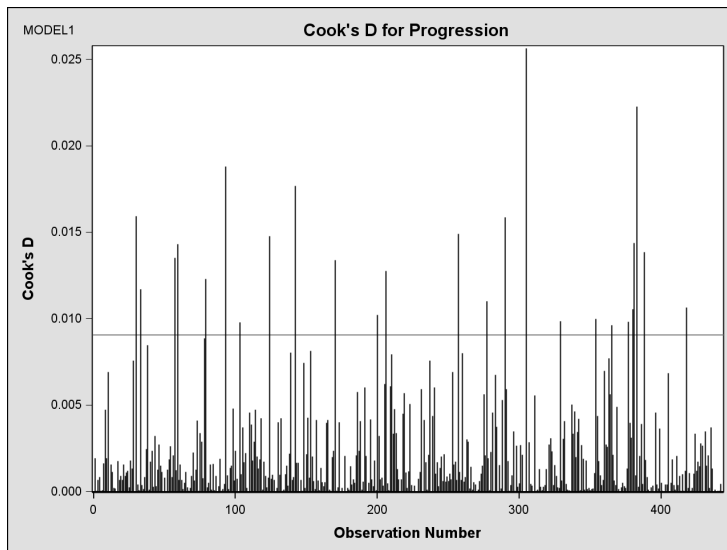
Figure 11.   Cook's D Plot Created with PROC REG

You can also directly request one or more of the plots appearing in the "Diagnostics Panel". For example, the following statements request the residual histogram shown in Figure 12.

```
ods html;
ods graphics on;
proc reg data=diabetes plots(only)=ResidualHistogram;
    model Progression = Sex BMI BP TC HDL LTG GLU;
run;
ods graphics off;
ods html close;
```
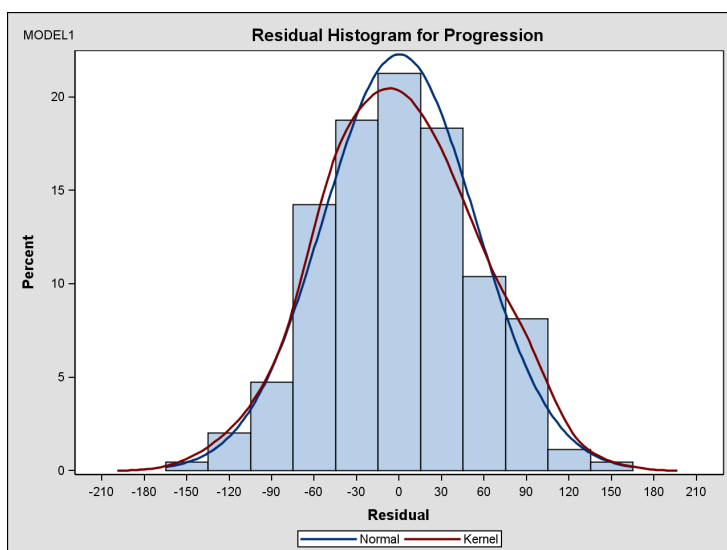

Figure 12.   Residual Histogram with Overlaid Densities Created with PROC REG

EXAMPLE 2: GRAPHICS FOR SURVIVAL ANALYSIS USING THE LIFETEST PROCEDURE

This example is based on the study of Bone Marrow Transplantation for Leukemia described by Klein and Moeschberger (1997). The data contain disease-free survival times of 137 bone marrow transplant patients in three risk groups: ALL (Acute Lymphoblastic Leukemia), AML (Acute Myeloctic Leukemia) low-risk, and AML high-risk. The endpoint of interest is the disease-free survival, which is the time to relapse or death or end of the study in days. The variable Group represents the Patient's risk category, the variable T represents the disease-free survival time, and the variable Status is the censoring variable with value 1 for an uncensored time and 0 for a censored time.

```
proc format;
    value risk 1='ALL' 2='AML low-risk' 3='AML high-risk';
run;

data BMT;
    input Group T Status @@;
    format Group risk.;
    label T='Time to Relapse';
datalines;
1 2081 0 1 1602 0 1 1496 0 1 1462 0 1 1433 0
1 1377 0 1 1330 0 1  996 0 1  226 0 1 1199 0
1 1111 0 1  530 0 1 1182 0 1 1167 0 1  418 1
.
.  more data lines
.
3  625 1 3   48 1 3  273 1 3   63 1 3   76 1
3  113 1 3  363 1
;
```

The following statements illustrate the default plot of estimated survival curves that the LIFETEST procedure creates when ODS Graphics is on.

```
ods html;
ods graphics on;
proc lifetest data=BMT;
    time T * Status(0);
    strata Group;
    run;
ods graphics off;
ods html close;
```

The plot is shown in Figure 13. AML low-risk patients have the best survival (that is, the longest time to relapse), and AML high-risk patients have the worst survival.

You can also display the number of patients at risk at specific time points in the survival plot by using the ATRISK= suboption, as shown by the following statements.

```
ods html;
ods graphics on;
proc lifetest data=BMT plot=survival(atrisk=0 to 2000 by 250);
    time T * Status(0);
    strata Group;
    run;
ods graphics off;
ods html close;
```
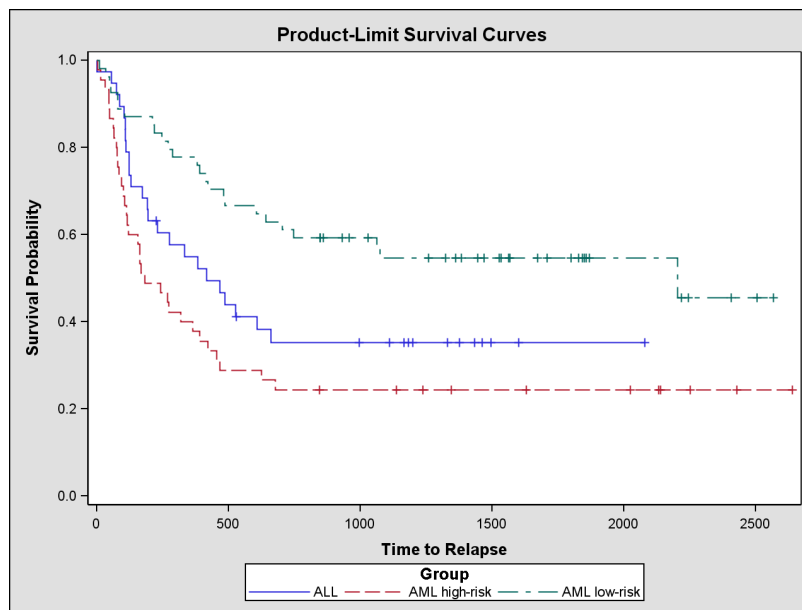
Figure 13.   Survival Plot Created with PROC LIFETEST

The PLOTS=SURVIVAL option requests the plot of the estimated survival curves, and the suboption ATRISK= requests that the numbers of patients at risk at 0, 250, 500, ..., 2000 days are to be displayed. The survival plot is shown in Figure 14.
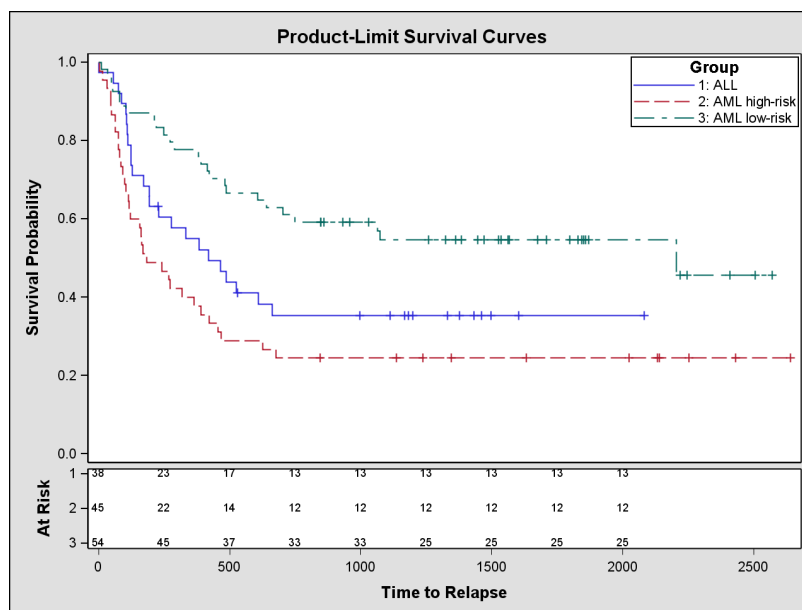

Figure 14.   Survival Plot with Numbers of At-Risk Patients

At 500 days, the numbers of patients at risk are 17, 37, and 14 for ALL, AML low-risk, and AML high-risk, respectively.

You can add pointwise and simultaneous confidence intervals for the survivor functions to the plot. You specify the suboption CL to request pointwise confidence intervals for the survivor functions, as illustrated by the following statements.

```
ods html;
ods graphics on;
proc lifetest data=BMT plot=survival(cl);
    time T * Status(0);
    strata Group;
    run;
ods graphics off;
ods html close;
```

You can request Hall-Wellner confidence bands and equal-precision confidence bands with the suboptions HW and EP, respectively. See Figure 1 for an illustration of the suboption HW.

With stratified data, PROC LIFETEST computes a test of the homogeneity of the survival curves across strata. You can use the TEST=LOGRANK suboption to inset the p-value inset in the survival plot, as shown by the following statements.

```
ods html;
ods graphics on;
proc lifetest data=BMT plot=survival(atrisk=0 to 2000 by 250 test=logrank);
    time T * Status(0);
    strata Group;
    run;
ods graphics off;
ods html close;
```

The suboption TEST=LOGRANK requests the p-value of the log-rank test, which is a default test. The survival plot is shown in Figure 15, which shows a p-value of 0.001.
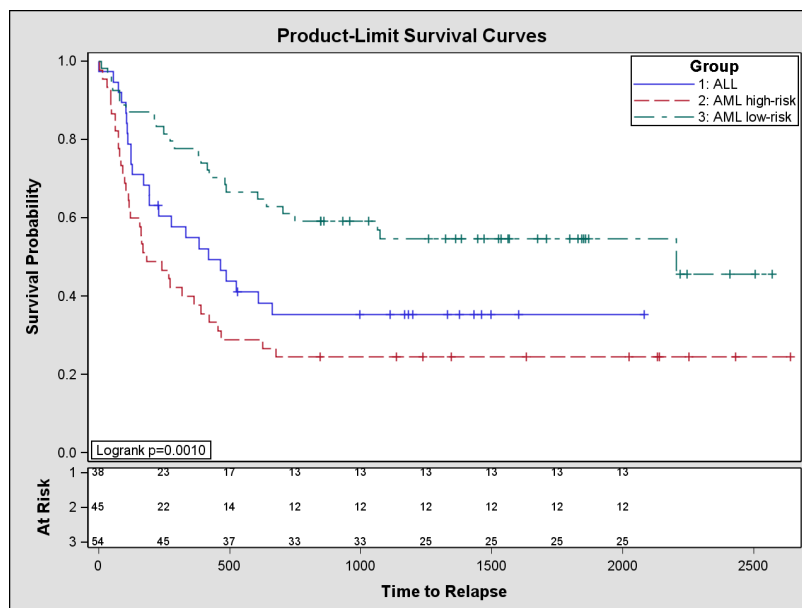

Figure 15.   Survival Plot with At-Risk Numbers and Homogeneity Test p-Value

EXAMPLE 3: MEANS COMPARISON PLOTS USING THE GLIMMIX PROCEDURE

This example analyzes a hypothetical data set of yields in an industrial process. Material was available from five vendors to perform a chemical reaction whose yield depends on two factors: pressure (three levels) and temperature (two levels). The effects of all three factors are assumed to be fixed. Because vendor-to-vendor variation in the material affects only the baseline process yield, a statistical model was fit that accounted for the main effect of vendor and the two-way factorial effects of pressure and temperature.

```
data Yields;
   input Vendor Pressure Temp Yield @@;
   datalines;
1  1  1  10.20   1  1  2   9.48   1  2  1   9.74
1  2  2   8.92   1  3  1  11.79   1  3  2   8.85
2  1  1  10.43   2  1  2  10.59   2  2  1  10.29
2  2  2  10.15   2  3  1  11.12   2  3  2   9.30
3  1  1   6.46   3  1  2   7.34   3  2  1   9.44
3  2  2   8.11   3  3  1   9.38   3  3  2   8.37
4  1  1   7.36   4  1  2   9.92   4  2  1  10.99
4  2  2  10.34   4  3  1  10.24   4  3  2   9.96
5  1  1  11.72   5  1  2  10.60   5  2  1  11.28
5  2  2   9.03   5  3  1  14.09   5  3  2   8.92
;
```

The following statements use the GLIMMIX procedure (Schabenberger 2005) to fit this model and request the "Type III Tests of Fixed Effects" table to examine the pattern of significant effects.

```
ods select Tests3;
proc glimmix data=Yields;
   class Vendor Pressure Temp;
   model Yield = Vendor Pressure Temp Pressure*Temp;
run;
```

The ODS SELECT statement specifies that only the table named Tests3 is to be displayed; see Figure 16. Note that you can use the ODS SELECT statement to select any combination of tables or graphs. Appendix B provides details concerning the use of the ODS SELECT and ODS EXCLUDE statements.

```
                    The GLIMMIX Procedure

              Type III Tests of Fixed Effects

                        Num      Den
         Effect          DF       DF    F Value   Pr > F

         Vendor           4       20      5.68    0.0032
         Pressure         2       20      1.42    0.2646
         Temp             1       20      6.48    0.0193
         Pressure*Temp    2       20      3.82    0.0393
```

Figure 16.   Tests of Model Effects

This initial step of the analysis reveals significant differences among the vendors as well as a significant interaction between the process factors. Note that the main pressure effect is not significant; it could be masked by the interaction. To gain visual insight into the interaction pattern and the vender-to-vendor

differences, the following PROC GLIMMIX statements request several plots that are related to least-squares means.

```
ods html;
ods graphics on;
ods select DiffPlot MeanPlot;

proc glimmix data=Yields;
   class Vendor Pressure Temp;
   model Yield = Vendor Pressure Temp Pressure*Temp;
   lsmeans Vendor          / plot=diff;
   lsmeans Pressure*Temp   / plot=mean(sliceby=Temp join);
run;

ods graphics off;
ods html close;
```

The CLASS and MODEL statements remain the same. The first LSMEANS statement requests a plot of the vendor least-squares means differences, shown in Figure 17, which is known as a Diffogram. When based on arithmetic means, this display is also known as a mean-mean scatter plot (Hsu 1996; Hsu and Peruggia 1994). The Diffogram displays significant and nonsignificant differences by rotating the confidence interval for least-squares mean differences. Whenever these lines cross the 45 degree reference line, the least-squares means associated with the center point of the line are not significantly different.
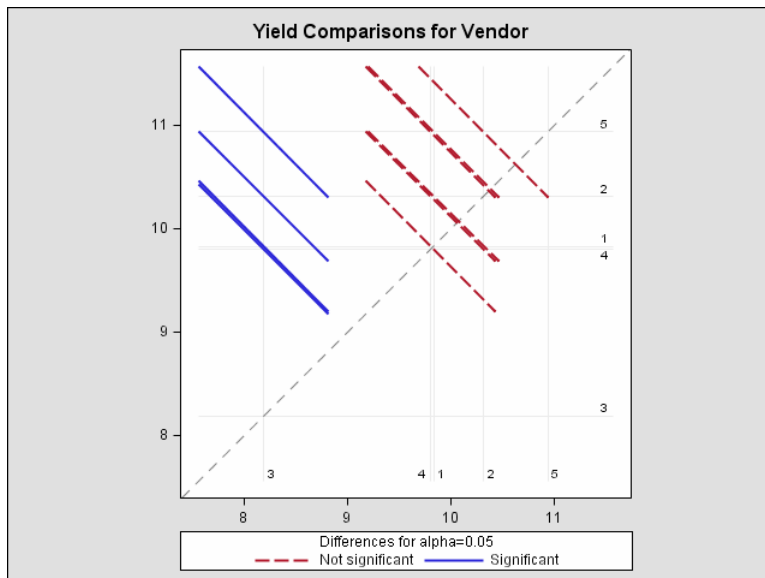

Figure 17.   Diffogram for Vendor Effect Created with GLIMMIX Procedure

Here, the Diffogram shows that yields from Vendor 3 are significantly different from the yields of the other four vendors. The yield differences among those four are not statistically significant.

The second LSMEANS statement requests a graph of the pressure $\times$ temperature interaction means, which is displayed in Figure 18. These cell means are sliced by temperature, that is, a separate relationship between yield and pressure is shown for each temperature.
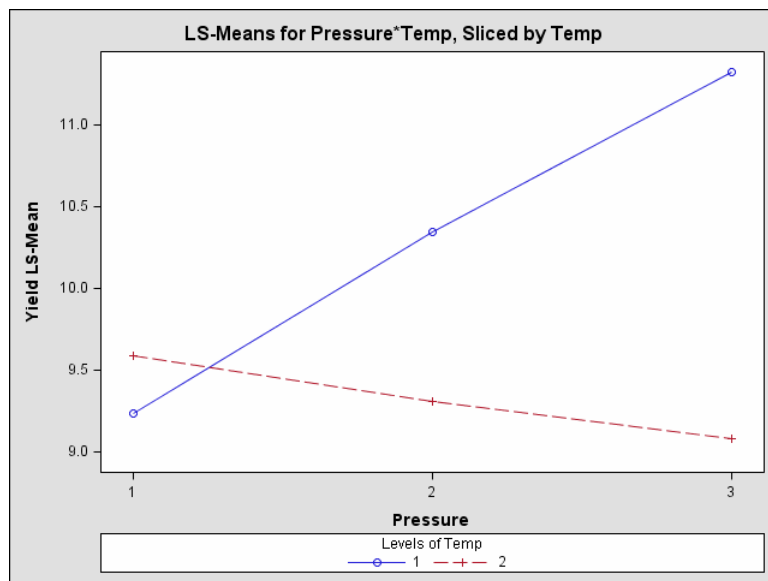
Figure 18. Plot of Pressure $\times$ Temperature Interaction Created with PROC GLIMMIX

The means plot in Figure 18 reveals the nature of the interaction (lack of parallelism) between the two factors that influence the reaction. At temperature level 2, the yield depends only weakly on pressure and declines with increasing pressure. At temperature level 1, the yield increases strongly with pressure.

There are occasions when you want to modify the default Diffogram. For example, you may want to separate the means depending on the sign of differences. When the number of levels is large, suppressing the lines or marking the center point of the lines can be a helpful visual device.

The following statements construct a Diffogram without lines (NOLINES option) in which the intersection of the least-squares means is marked (CENTER option) and in which the comparisons are separated by sign (NOABS option, SAS 9.1). The CENTER and NOLINES option have been added for SAS 9.2 as a result of user feedback.

```
ods html;
ods graphics on;

ods select DiffPlot;
proc glimmix data=Yields;
   class Vendor Pressure Temp;
   model Yield = Vendor Pressure Temp Pressure*Temp;
   lsmeans Vendor / plot=diff(center nolines noabs);
run;

ods graphics off;
ods html close;
```

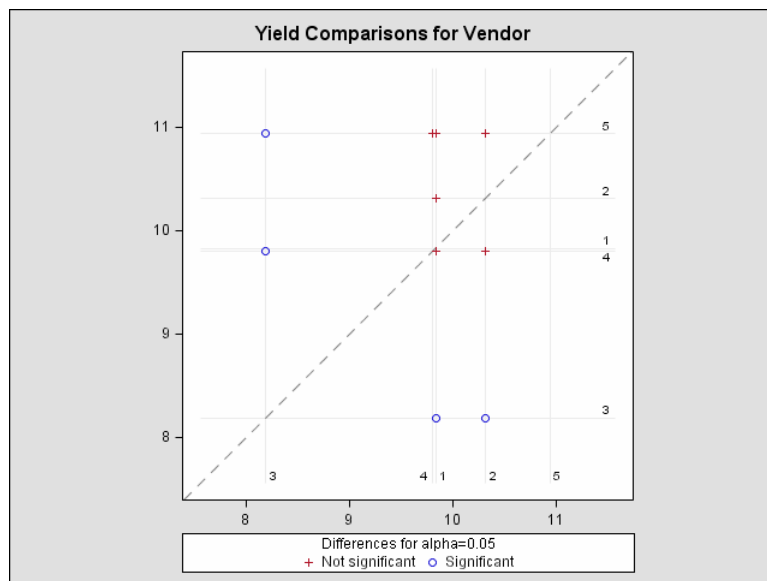The modified Diffogram is shown in Figure 19.

Figure 19. Diffogram with Additional SAS 9.2 Options (NOLINES, CENTER)

EXAMPLE 4: CONTROL CHARTS USING THE SHEWHART PROCEDURE

In SAS 9.2, a number of SAS/QC procedures, including the SHEWHART procedure, will make use of ODS Graphics; see Appendix A for a complete list. These procedures already provide extensive options and features for creating high-resolution GRSEG graphics (the same type of graphics created by SAS/GRAPH procedures), and they will continue to support GRSEG graphics. Adding ODS Graphics functionality to SAS/QC procedures such as SHEWHART and CAPABILITY gives SAS/QC users the option of creating graphics that are consistent with displays produced by the statistical procedures.

To request ODS Graphics with SAS/QC procedures, you specify the ODS GRAPHICS statement, as usual. Note, however, that neither GOPTIONS statement options nor procedure options for specifying colors, fonts, line types, and other attributes will affect displays requested with ODS Graphics. Instead, these attributes will be determined by the ODS style that is in effect for your ODS destination.

This behavior is illustrated by the following statements, which use a data set containing delay time measurements that are subgrouped by day. The data are provided in Example 39.2 of the *SAS/QC 9.1 User's Guide* (SAS Institute 2004b). First, the SHEWHART procedure is used *without* the ODS GRAPHICS statement to create a control chart with box-and-whisker plots for the subgroup samples.

```
proc shewhart data=times limits=timelim ;
   boxchart delay * day /
      boxstyle = schematicid
      cboxes   = dagr
      cboxfill = ywh
      cframe   = vligb
      idcolor  = black
      nohlabel
      nolegend
      notches;
   id reason;
   label delay = 'Delay in Minutes';
run;
```

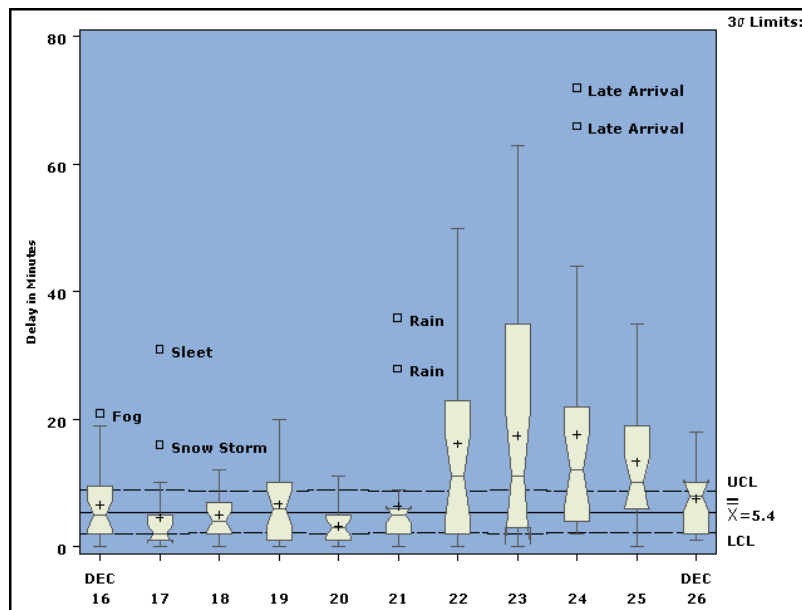The control chart, which is created with GRSEG graphics, is shown in Figure 20.

Figure 20. Control Chart Created with PROC SHEWHART Using GRSEG Graphics

Now, consider the same statements but with the addition of the ODS GRAPHICS statement.

```
ods html style=default;
ods graphics on;

proc shewhart data=times limits=timelim ;
   boxchart delay * day /
      boxstyle = schematicid
      cboxes   = dagr
      cboxfill = ywh
      cframe   = vligb
      idcolor  = black
      nohlabel
      nolegend
      notches;
   id reason;
   label delay = 'Delay in Minutes';
run;

ods graphics off;
ods html close;
```

The control chart produced with ODS Graphics is shown in Figure 21. The color options specified in the BOXCHART statement are not used, and instead the colors are determined by the Default style for HTML.
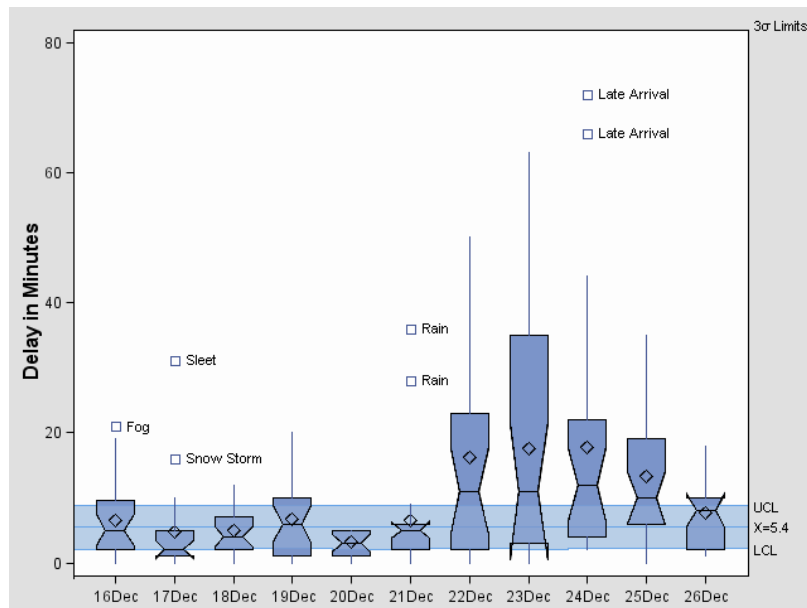
Figure 21.   Control Chart Created with PROC SHEWHART Using ODS Graphics

In addition to the SAS/QC procedures listed in Appendix A, there are other procedures, including BOXPLOT, REG, and UNIVARIATE, that are being modified to support ODS Graphics in addition to legacy GRSEG graphics. These procedures will continue to support GRSEG graphics, but when the ODS GRAPHICS statement is used their appearance will be determined by the style that is in effect for the ODS destination.

EXAMPLE 5: SCATTER PLOTS USING NEW STATISTICAL GRAPHICS PROCEDURES IN SAS/GRAPH

Making a plot of the data is often the first step in a statistical analysis, and general-purpose tools for creating displays such as scatter plots are indispensable to the data analyst. SAS 9.2 introduces a family of SAS/GRAPH procedures that use the functionality of ODS Graphics and are specifically designed for creating displays commonly needed for exploratory data analysis and statistical work.

The initial set of procedures in this family is

- the SGPLOT procedure, which you can use to create *overlaid* plots. These plots consist of a *principal* plot (such as a scatter plot, line plot, step plot, or needle plot) that is superimposed with data smooths (such as regression lines, loess smooths, spline smooths) and other features.
- the SGCATTER procedure, which you can use to create a variety of scatter plots, including scatterplot matrices and comparative scatter plots.
- the SGRENDER procedure, which you can use to create a custom plot by rendering a graph template with the Graph Template Language. The topic of graph templates lies outside the scope of this paper and is discussed by Cartier (2006).

The names of these procedures begin with the prefix SG to distinguish them from traditional SAS/GRAPH procedures, which create GRSEG graphics and are controlled by the GOPTIONS statement. Graphs created by these new procedures are consistent with graphs created by the statistical procedures that use ODS Graphics because they are produced with the same technology and the same ODS styles. In addition to this consistency, the SGPLOT and SGSCATTER procedures offer the advantage of a syntax that is convenient for statistical graphics.

To illustrate the SGPLOT procedure, consider the data set CARS, which provides descriptive variables for a number of car models.

```
data cars;
   length country $9 type $7;
   label horsepower = "Horsepower"
         displace   = "Displacement"
         weight     = "Weight (kg)"
         mileage    = "Mileage (MPG)";
   input country type price reliability mileage weight displace horsepower;
datalines;
USA        Small    8895      4          33    2560        97          113
USA        Small    7402      2          33    2345       114           90
Korea      Small    6319      4          37    1845        81           63
 .
 .         more data lines
 .
run;
```

The following statements use the SGPLOT procedure to create a scatter plot of the variables horsepower and mileage overlaid with a loess smooth.

```
title  "Comparison of Mileage and Horsepower";
title2 "Points Labeled With MPG";

proc sgplot data=cars;
   loess y=horsepower x=mileage / clm nomarkers;
   scatter y=horsepower x=mileage / group=type datalabel=mileage name="scatter";
   discretelegend "scatter";
run;
```

The display is shown in Figure 22. Note that the loess fit is computed directly by the SGPLOT procedure.
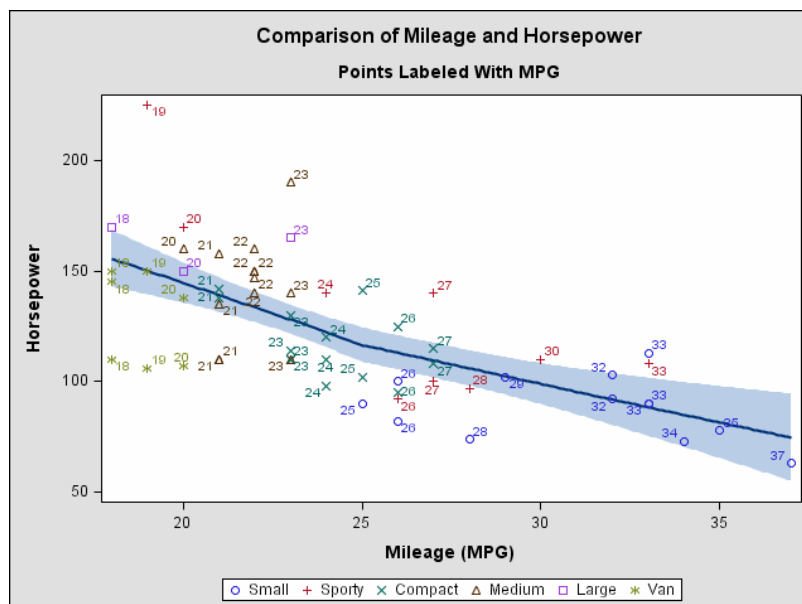

Figure 22.   Scatter Plot and Loess Smooth Created with PROC SGPLOT

The following statements use the SGSCATTER procedure to create a matrix of scatter plots of mileage and weight versus displace and horsepower, using type as a group (classification) variable.

```
proc sgscatter data=cars tmplout="temp.sas";
    compare y=(mileage weight) x=(displace horsepower) /
    group=type loess=(nogroup) spacing=5;
run;
```

The LOESS= option requests loess smooths for the plots. The NOGROUP suboption requests a single smooth for all the data in each plot (ignoring the GROUP= variable). The display is shown in Figure 23.
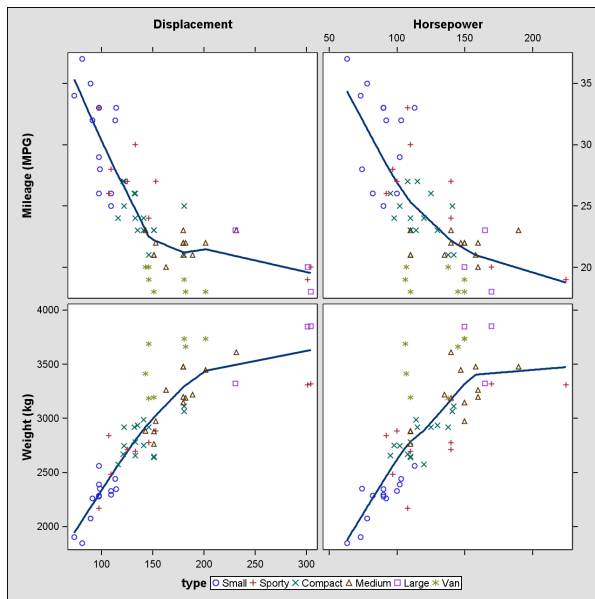

Figure 23. Comparative Scatter Plot Created with PROC SGCATTER

SUMMARY: WHAT DO YOU REALLY NEED TO KNOW?

The production release of ODS Graphics in SAS 9.2 delivers the second major installment of new functionality for providing graphics as an integrated part of the output of statistical procedures. Over 50 procedures have been modified to take advantage of this functionality, with more to come in future releases.

Creating graphics is as simple as requesting tables. You turn on ODS Graphics with the statement

```
ods graphics on;
```

Then you use procedure options and defaults to specify the graphs you want, and you use standard ODS destination statements to specify where you want your graphs to be displayed.

If you are viewing RTF output on Windows, you can cut and paste your graphs into a Word document or a PowerPoint presentation. You can also save your graphs in files for inclusion in LaTeX and other documents. ODS Graphics provides a variety of features for managing these files and for specifying the size and resolution of graphs.

ODS styles control the visual appearance of both graphs and tables, and SAS 9.2 provides styles that are designed for professional statistical work. The new ODS Graph Editor provides a point-and-click environment for modifying graphs that you are preparing for a paper or presentation.

APPENDIX A: PROCEDURES SUPPORTING ODS GRAPHICS

The following procedures have been modified to support ODS Graphics functionality in SAS 9.2. Note that in SAS 9.2, SAS/GRAPH is required for ODS Graphics functionality. ODS Graphics coverage for SAS/QC is anticipated to be experimental.

Base SAS

- CORR
- FREQ
- UNIVARIATE

SAS/ETS

- ARIMA
- AUTOREG
- ENTROPY
- EXPAND
- MODEL
- SYSLIN
- TIMESERIES
- TSCSREG
- UCM
- VARMAX
- X12

SAS High-Performance Forecasting

- HPF
- HPFENGINE

SAS/QC (planned)

- ANOM
- CAPABILITY
- CUSUM
- MACONTROL
- PARETO
- RELIABILITY
- SHEWHART

SAS/STAT

- ANOVA
- BOXPLOT
- CALIS
- CLUSTER
- CORRESP
- FACTOR
- GAM
- GENMOD
- GLM
- GLIMMIX
- GLMSELECT
- KDE
- LIFEREG
- LOESS
- LOGISTIC
- MDS
- MI
- MIXED
- MULTTEST
- NPAR1WAY
- PHREG
- PLS
- PRINQUAL
- QUANTREG
- REG
- ROBUSTREG
- RSREG
- TRANSREG
- TTEST

APPENDIX B: SELECTING AND EXCLUDING GRAPHS

You can specify the name of a graph in the ODS SELECT and ODS EXCLUDE statements to select or exclude it from your output, just as you can select or exclude a table from your output. There are three ways to determine a graph name:

- Names for all the graphs produced by a procedure are listed in the Details section titled "ODS Graph Names" of the procedure chapter in the forthcoming *SAS/STAT 9.2 User's Guide.*
- If you are running in the Display Manager, you can use the Results window to view the names of ODS graphs created in your SAS session.
- You can run the procedure with the ODS TRACE ON statement, which requests a record of the output objects created by ODS.

Note: The ODS graph name is used to derive the name of the image file that captures the graph; see Appendices C and D.

APPENDIX C: GRAPHICS IMAGE FILES

Accessing your graphs as individual image files is useful when you want to include them in various types of documents. The default image file type depends on the ODS destination, but there are other supported image file types that you can specify. You can also specify the names for your graphics image files and the directory in which you want to save them.

If you are using an HTML or a LATEX destination, your graphs are individually produced in a specific image file type, such as PNG or PostScript.

If you are using a destination in the PRINTER family or the RTF destination, the graphs are contained in the ODS output file and cannot be accessed as individual image files. However, you can open an RTF output file in Microsoft Word and then copy and paste the graphs into another document, such as a Microsoft PowerPoint presentation.

The following table shows the various ODS destinations supported by ODS Graphics in SAS 9.2, the viewer for displaying graphs in each destination (if applicable), and the image file types supported for each destination.

Destination	Destination Family	Viewer	Image File Types
DOCUMENT		Not Applicable	Not Applicable
HTML	MARKUP	Browser	PNG (new 9.2 default), GIF, JPEG
LATEX	MARKUP	Ghostview	PostScript (default), EPSI, GIF, JPEG, PNG
PCL	PRINTER	Ghostview	Contained in PostScript file
PDF	PRINTER	Acrobat	Contained in PDF file
PS	PRINTER	Ghostview	Contained in PostScript file
RTF		Microsoft Word	Contained in RTF file
LISTING		Host-specific	GIF (default)

Note that the default image file type of PNG is planned for HTML in SAS 9.2 (the default is GIF in SAS 9.1).

You can specify the image file type for the HTML and LATEX destinations with the IMAGEFMT= option in the ODS GRAPHICS statement as follows.

```
ods graphics on / imagefmt = png;
```

Currently, the default image resolution is 100 dpi for the HTML and LISTING destinations, and it is 200 dpi for other destinations.

APPENDIX D: NAMING GRAPHICS IMAGE FILES

The names of your graphics image files are composed of a *base file name*, an *index counter*, and an *extension*. By default, the base file name is the ODS graph name. The counter is set to one when you begin a SAS session; it is applied each time you create the same graph in a session, starting with the *second* time, and then it is increased by one each time. The extension indicates the image file type.

For instance, suppose you run the following statements at the beginning of a SAS session.

```
ods html;
ods graphics on;

proc reg data=one plots(only)=(qq cooksd);
model y=x;
quit;

proc reg data=one plots(only)=(qq cooksd);
model y=x;
quit;

ods graphics off;
ods html close;
```

The two graphics image files created after the first procedure step are CooksD.png and QQPlot.png. The two graphics image files created after the second step are CooksD1.png and QQPlot1.png.

You can specify the RESET option in the ODS GRAPHICS statement to reset the index counter. This avoids duplication of graphics image files and saves disk space if you are rerunning a SAS program in the same session. For instance, suppose you run the following statements at the beginning of a SAS session.

```
ods html;
ods graphics on;

proc reg data=one plots(only)=(qq cooksd);
model y=x;
quit;
ods graphics off;

ods graphics on / reset;
proc reg data=one plots(only)=(qq cooksd);
model y=x;
quit;

ods graphics off;
ods html close;
```

The only two graphics image files created after the second procedure step are CooksD.png and QQPlot.png.

You can specify a base file name for all your graphics image files with the IMAGENAME= option in the ODS GRAPHICS statement. For example, suppose you run the following statements.

```
ods html;
ods graphics on / imagename = "MyPlots";

proc reg data=one plots(only)=(qq cooksd);
model y=x;
quit;

proc reg data=one plots(only)=(qq cooksd);
model y=x;
quit;

ods graphics off;
ods html close;
```

Then the two graphics image files created after the first procedure step are MyPlots.png and MyPlots1.png rather than CooksD.png and QQPlot.png. The two graphics image files created after the second step are MyPlots2.png and MyPlots3.png rather than CooksD1.png and QQPlot1.png.

Finally, suppose you run the following statements.

```
ods html;
ods graphics on / imagename="MyPlots";

proc reg data=one plots(only)=(qq cooksd);
model y=x;
quit;
ods graphics off;

ods graphics on / imagename="MyPLots" reset;
proc reg data=one plots(only)=(qq cooksd);
model y=x;
quit;

ods graphics off;
ods html close;
```

The only two graphics image files created after the second procedure step are MyPlots.png and MyPlots1.png.

APPENDIX E: SAVING GRAPHICS IMAGE FILES

Knowing where your graphics image files are saved and how they are named is particularly important if you are running in batch mode or if you plan to access the files for inclusion in a document. The following discussion assumes you are running SAS under the Windows operating system.

Your graphics image files are saved by default in the SAS current folder. If you are using the SAS windowing environment, the current folder is displayed in the status line at the bottom of the main SAS window. If you are running your SAS programs in batch mode, the graphs are saved by default in the same directory where you started your SAS session.

With the HTML, LATEX, and LISTING destinations, you can specify a directory for saving your graphics image files. With the PRINTER and RTF destinations, you can only specify a directory for your output file.

If you are using the HTML destination, the individual graphs are created as PNG files by default in SAS 9.2. You can use the PATH= and GPATH= options in the ODS HTML statement to specify the directory where your HTML and graphics files are saved, respectively. This also gives you more control over your graphs. For example, if you want to save your HTML file named test.htm in the C:\myfiles directory, but you want to save your graphics image files in C:\myfiles\images, then you specify

```
ods html path  = "C:\myfiles"
         gpath = "C:\myfiles\images"
         file  = "test.htm";
```

You can add the URL= suboption to create a *relative* path rather than a complete path in the HTML source file test.htm for the graphs it references. This is illustrated by the following statement, which creates a relative path to the images subdirectory.

```
ods html path  = "C:\myfiles"
         gpath = "C:\myfiles\images" (url = "images\")
         file  = "test.htm";
```

You can specify URL=NONE to create a relative path to the same directory in which the source file is saved. Note that without the URL= suboption, a deployed HTML file might not display the graph content.

If you are using the LATEX destination, the individual graphs are saved as PostScript files by default. If you specify the PATH= and GPATH= options in the ODS LATEX statement, your tabular output is saved as a LATEX source file in the PATH= directory, and your graphs are saved as PostScript files in the GPATH= directory. If you are writing a paper in LATEX, you can reference the graph PostScript files in your paper. In this situation, you might not need to use the source file.

APPENDIX F: RESIZING GRAPHS AND RESCALING FONTS

This example illustrates how you can use the HEIGHT= and WIDTH= options in the ODS GRAPHICS statement to change the size of the graphics produced.

The following statements generate data where the dependent variable is a nonlinear function of a regressor.

```
data ScaleExample;
   do x=-2 to 4 by 0.1;
     y=x*x*x-2*x-x+2+3*rannor(1);
     output;
   end;
run;
```

You can use the LOESS procedure to fit a nonparametric model for y as a function of x. This fit is built from many local linear regressions, each using a small range of the regressor data. The smoothness of the fit you obtain is controlled by a smoothing parameter that specifies the fraction of the data used in each local regression. While PROC LOESS has facilities to choose an appropriate smoothing parameter automatically, you may want to visualize how the fit changes as the smoothing parameter varies. The following statements produce four local regression fits corresponding to four specified smoothing parameter values. The PLOTS=FITPANEL option in the PROC LOESS statement requests the panel of LOESS fits shown in Figure 24.

```
ods html;
ods graphics on;

proc loess data=ScaleExample plots=fitPanel;
   model y=x/smooth=0.1 0.25 0.5 0.9;
run;

ods graphics off;
ods html close;
```
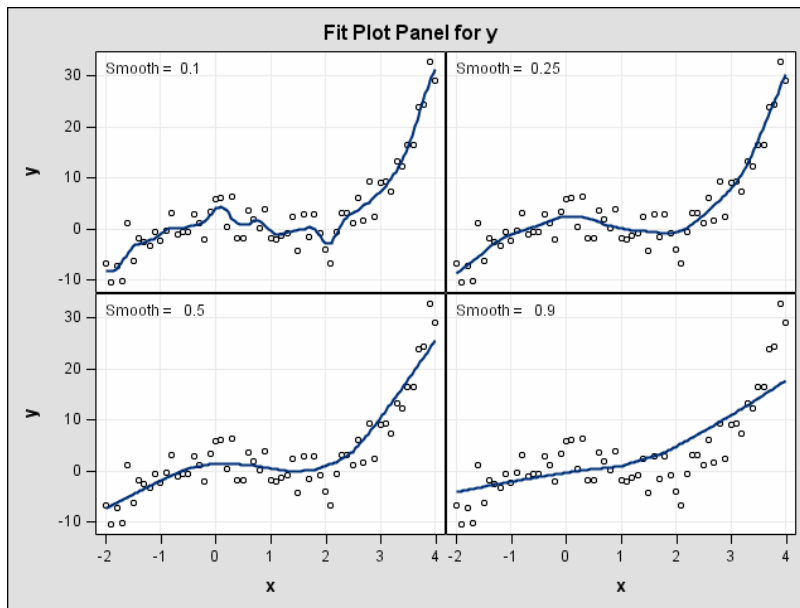
Figure 24.   Fit Panel with Default Size

Suppose you want to include the plot in Figure 24 in an article for a journal that requires a two-column format for both text and figures. The statement

```
ods graphics on / width=320px height=240px;
```

halves the width and height of this plot from their default values of 640 pixels and 480 pixels, respectively, producing a plot appropriately sized for a two-column format. This plot is shown in Figure 25.
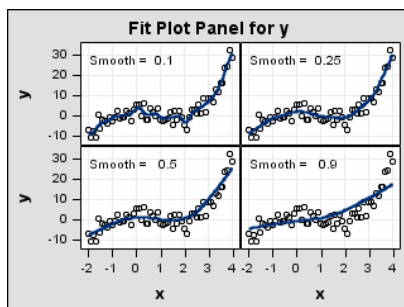

Figure 25.   Fit Panel with Reduced Size

Observe that the fonts in this smaller plot are automatically scaled with the size of the graph. In some cases, you may prefer to reduce the size of the graph, but still retain the larger default font sizes. You can suppress the font scaling with the SCALE= option in the ODS GRAPHICS statement. In the preceding statements, if you use the modified ODS GRAPHICS statement

```
ods graphics on / width=320px height=240px scale=off;
```

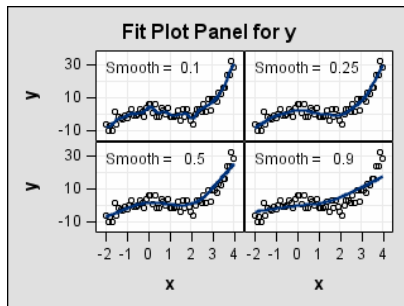then you obtain the plot in Figure 26.

Figure 26.   Fit Panel with Reduced Size and No Font Scaling

REFERENCES

Cartier, J. (2006), "A Programmer's Introduction to the Graphics Template Language," *Proceedings of the Thirty-first Annual SAS Users Group International Conference.* Cary, NC: SAS Institute Inc.

Cohen, R. A. (2006), "Introducing the GLMSELECT PROCEDURE for Model Selection," *Proceedings of the Thirty-first Annual SAS Users Group International Conference.* Cary, NC: SAS Institute Inc.

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004), "Least Angle Regression (with discussion)," *Annals of Statistics*, 32, 407–499.

Hsu, J. C. (1996), *Multiple Comparisons: Theory and Methods*, London: Chapman & Hall.

Hsu, J. C. and Peruggia, M. (1994), "Graphical Representation of Tukey's Multiple Comparison Method," *Journal of Computational and Graphical Statistics*, 3: 143–161.

Klein, J. P. and Moeschberger, M. L. (1997), *Survival Analysis: Techniques for Censored and Truncated Data*, New York: Springer-Verlag.

Rodriguez, R. N. (2004), "An Introduction to ODS for Statistical Graphics in SAS 9.1," *Proceedings of the Twenty-ninth Annual SAS Users Group International Conference.* Cary, NC: SAS Institute Inc.

SAS Institute Inc. (2004a), *Getting Started with the ADX Interface for Design of Experiments.* Cary, NC: SAS Institute Inc.

SAS Institute Inc. (2004b), *SAS/QC 9.1 User's Guide.* Cary, NC: SAS Institute Inc.

Schabenberger, O. (2005), "Introducing the GLIMMIX Procedure for Generalized Linear Mixed Models," *Proceedings of the Thirtieth Annual SAS Users Group International Conference.* Cary, NC: SAS Institute Inc.

Tibshirani, R. (1996), "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society Series B*, 58, 267–288.

CONTACT INFORMATION

Robert N. Rodriguez　　　Tonya E. Balan
SAS Institute Inc.　　　　　SAS Institute Inc.
SAS Campus Drive　　　　 SAS Campus Drive
Cary, NC 27513　　　　　　Cary, NC 27513
(919) 531-7650　　　　　　 (919) 531-9487
Bob.Rodriguez@sas.com　 Tonya.Balan@sas.com