ABSTRACT
SAS® statistical graphics procedures (SG procedures) that were introduced in SAS 9.2 help you to create high quality statistical graphics. These procedures include SGPLOT, SGPANEL, SGSCATTER, and SGRENDER. The procedures SGPLOT, SGPANEL, SGSCATTER need only little coding. These procedures can be used to create most of the commonly used statistical graphics. Each procedure is created for different purpose. To create customized statistical graphics and layouts, you can use the procedure SGRENDER. To use SGRENDER procedure, you need to create templates using Graph Template Language (GTL). The coding for these templates may be little complicated. Here we illustrate some use of these procedures through some examples. This paper also shows how to enhance the appearance of the graphs created by these procedures using ODS styles.

INTRODUCTION
The statistical graphic procedures include SGPLOT, SGPANEL, SGSCATTER, and SGRENDER. The procedures SGPLOT, SGPANEL, SGSCATTER can be used easily to create many commonly used statistical graphics. These procedures require very little coding and can produce very high quality graphics. The SGRENDER procedure can be used to create customized statistical graphics. To use SGRENDER procedure, you need to create templates using Graph Template Language (GTL). The coding for these templates may be little complicated.

The SGPLOT procedure can create one or more plots overlaid on a single set of axes. Statistical graphics SGPLOT can produce include scatter plots, line plots, histograms, bar charts, and regression plots.

The SGPANEL procedure can create a panel of graph cells for the values of one or more classification variables. Like SGPLOT, SGPANEL can create a variety of plot types, and overlay these plots on each graph cell of the panel. It can produce several types of layout.

The SGSCATTER procedure produces a paneled graph of scatter plots for different combination of variables. You can overlay fit plots and ellipses on your scatter plots. It can create many different types of paneled graphs.

The SGRENDER procedure creates graphics using templates created with GTL language. The GTL and SGRENDER procedure can be used to create customized layouts and graphs that cannot be created using SGPLOT, SGPANEL, or SGSCATTER procedures.

There are statements and options in all these procedures to control the appearance of your graph and to add features like legends and reference lines.

The graphics produced by these procedures can be sent to ODS destinations, and use ODS styles. The graph sizes can be changed using ODS GRAPHICS statement.

SGPLOT
SGPLOT procedure can produce the following:
- Basic Plots: scatter series, step, band, and needle.
- Fits and Confidence: Loess, regression, penalized B-spline, and computed ellipse
- Distribution: horizontal and vertical box plots, histograms, normal curves, and kernel density estimates.
- Categorization: dot plots, horizontal and vertical bar charts, horizontal, and vertical line charts.

We will show how to create some of these plots through some examples.
SCATTER PLOT

The following example creates a scatter plot using SGPLOT. This plot can be sent to any ODS destination.

```plaintext
PROC SGPLOT DATA = Sashelp.Class;
SCATTER X = Height Y = Weight;
RUN;
```

The following example fits a regression line to the above scatter plot. The option DATALABEL displays the values of the variable 'name'.

```plaintext
PROC SGPLOT DATA = Sashelp.Class;
TITLE 'Scatter Plot and Linear Regression';
SCATTER X = Height Y = Weight;
REG X = Height Y = Weight / DATALABEL = Name;
RUN;
```
BOX PLOT

The following example creates a box plot for the quantitative variable 'weight' for different values of the categorical variable 'origin'.

```plaintext
PROC SGPLOT DATA = Sashelp.Cars;
    HBOX Weight / CATEGORY = Origin;
RUN;
```

![Box Plot Example]

The following example adds a vertical reference line at the x-value 3500 to the above plot.

```plaintext
PROC SGPLOT DATA = Sashelp.Cars;
    RERLINES 3500 / AXIS = X LABEL = '3500';
    HBOX Weight / CATEGORY = Origin;
RUN;
```

![Box Plot with Reference Line Example]
SERIES PLOT

The following code produces a series plot for the variable “coal.”

```
PROC SGPLOT DATA = Sashelp.Electric (WHERE = (Customer = "Residential"));
  SERIES X = Year Y = Coal;
RUN;
```

![Series plot for coal](image)

The following code is for a series plot for a different variable “naturalgas.” In the above example, the x-axis label values are the default values which are decimals. The option `discrete` will use the actual data values.

```
PROC SGPLOT DATA = Sashelp.Electric (WHERE = (Customer = "Residential"));
  XAXIS TYPE = DISCRETE;
  SERIES X = Year Y = Naturalgas;
RUN;
```

![Series plot for natural gas](image)
The following code overlays the above two graphs. The option MARKERS mark the plotted points.

```sas
PROC SGPLOT DATA = Sashelp.Electric (WHERE = (Customer = "Residential"));
   XAXIS TYPE = DISCRETE;
   SERIES X = Year Y = Coal / MARKERS;
   SERIES X = Year Y = Naturalgas / MARKERS Y2AXIS;
RUN;
```

In the following example, using the keylegend option the legend in the above example is moved inside and to the top left position.

```sas
PROC SGPLOT DATA = Sashelp.Electric (WHERE = (Customer="Residential"));
   XAXIS TYPE=DISCRETE;
   SERIES X = Year Y = Coal/ MARKERS;
   SERIES X = Year Y = Naturalgas / MARKERS Y2AXIS;
   KEYLEGEND / LOCATION = INSIDE POSITION = TOPLEFT;
RUN;
```
HISTOGRAM AND DENSITY CURVES

This example creates a histogram for the variable ‘height’ in the data set ‘class’.

```
PROC SGPLOT DATA = Sashelp.Class;
   HISTOGRAM Height;
RUN;
```

If you need count information also for the variable height, you can use the following code.

```
PROC SGPLOT DATA = Sashelp.Class;
   HISTOGRAM Height;
   HISTOGRAM Height/ SCALE = COUNT Y2AXIS;
RUN;
```
In the above example, the legend is confusing and unnecessary. The following example gets rid of the legend using the option `noautolegend`.

```SAS
PROC SGPLOT DATA = Sashelp.Class NOAUTOLEGEND;
  HISTOGRAM Height;
  HISTOGRAM Height / SCALE = COUNT Y2AXIS;
RUN;
```

In the following example, a **Normal density Curve and a Kernel Density Curve** are overlaid on a Histogram.

```SAS
PROC SGPLOT DATA = Sashelp.Class NOAUTOLEGEND;
  HISTOGRAM Height;
  HISTOGRAM Height / SCALE = COUNT Y2AXIS;
RUN;
```
In the following example options `fillattrs` and `lineattrs` are included to the above code.

```sas
PROC SGPLOT DATA = Sashelp.Class;
  HISTOGRAM Height /FILLATTRS = (COLOR=GREEN);
  DENSITY Height / LINEATTRS = (COLOR=RED PATTERN = DASH THICKNESS = 5);
  DENSITY Height / TYPE=KERNEL  LINEATTRS = (COLOR = YELLOW PATTERN = SOLID THICKNESS = 5);
RUN;
```

**BAR CHART**

The following example is a bar chart for the quantitative variable ‘predict’ at different levels of the categorical variable ‘country’.

```sas
PROC SGPLOT DATA = Sashelp.Prdsale;
  VBAR Country / RESPONSE = Predict;
RUN;
```
The example below overlays bar chart for the variable actual on the bar chart of the variable predict. We can hardly see the bar chart of the variable predict. The next example shows how to modify this code so we can see the bar charts for both variables clearly.

```sas
PROC SGPLOT DATA = Sashelp.Prdsale;
   VBAR Country / RESPONSE = Predict;
   VBAR Country / RESPONSE = Actual;
RUN;
```

In the following example, options **barwidth** and **transparency** are used to make the above graphic clearer.

```sas
PROC SGPLOT DATA = Sashelp.Prdsale;
   VBAR Country / RESPONSE = Predict;
   VBAR Country / RESPONSE = Actual
      BARWIDTH = 0.5 TRANSPARENCY = 0.2;
RUN;
```
SGPANEL

The SGPANEL procedure creates a panel of graph cells based on classification variables. The SGPANEL procedure has a required PANELBY statement that is used to define the classifier variables for the panel. You can use options in the PANELBY statement to control the attributes of the panel. For example, you can use the ROWS= option to specify the number of rows in the panel.

SGPANEL can use different layouts, which are specified by the LAYOUT= option in the PANELBY statement. The layout determines how your classifier variables are used to create the panel. The default layout is PANEL. With this layout, you can specify any number of classifier variables.

The following example shows a panel of graph cells with histograms and plots with default layout PANEL.

```
PROC SGPANEL DATA = Sashelp.Heart NOAUTOLEGEND;
  TITLE "Cholesterol Distribution in Heart Study";
  PANELBY Sex;
  HISTOGRAM Cholesterol;
  DENSITY Cholesterol;
RUN;
```

In the following example rows and columns are used.

```
PROC SGPANEL DATA = Sashelp.Cars;
  TITLE "Cars by Type";
  PANELBY Type / ROWS = 2 COLUMNS = 3;
  SCATTER X = Mpg_City Y = Mpg_Highway;
RUN;
```
The following example shows a panel of bar charts with COLUMNLATTICE layout. In the COLUMNLATTICE layout, the graph cells are arranged in a single row. The option NOBORDER removes the borders around the cells. NOVARNAME option removes the variable name.

PROC SGPANEL DATA = Sashelp.Pordsale;
  TITLE "Yearly Sales by Product";
  PANELBY Year / NOVARNAME LAYOUT = COLUMNLATTICE NOBORDE;
  VBAR Product / RESPONSE = Actual;
RUN;
SGSCATTER

The SGSCATTER procedure contains three statements that you can use to create a paneled graph of scatter plots:

- MATRIX
- PLOT
- COMPARE

It supports overlaying of fits and ellipses. To create a single scatter plot it is better to use SGPLOT procedure.

SCATTER PLOT MATRIX

The following example shows a scatter plot matrix.

PROC SGSCATTER DATA = Sashelp.Iris;
  TITLE "Scatterplot Matrix for Iris Data";
  MATRIX Sepallength Petallength Sepalwidth Petalwidth;
RUN;

The following example creates a scatter plot matrix with group data. It also creates a histograms and normal density curves for the variables sepal length, petal length, sepal width, and petal width.

PROC SGSCATTER DATA = Sashelp.Iris;
  TITLE "Scatterplot Matrix for Iris Data";
  MATRIX Sepallength Petallength Sepalwidth Petalwidth / GROUP = Species DIAGONAL = (HISTOGRAM NORMAL);
RUN;
COMPARATIVE PANEL USING PLOT STATEMENT

The following example creates a comparative scatter plot.

```sas
PROC SGSCATTER DATA = Sashelp.Iris (WHERE = (Species = "Virginica"));
  TITLE "Multi-Celled plot for Species Virginica";
  PLOT (Sepallength Sepalwidth)*(Petallength Petalwidth);
RUN;
```

![Scatterplot Matrix for Iris Data](image)

![Multi-Celled Spline Curve for Species Virginica](image)
COMPARATIVE PANEL USING COMAPRE STATEMENT

The COMPARE statement is used to create a shared axis panel. The option ELLIPSE creates a prediction ellipse. A prediction ellipse is a region for predicting a new observation.

```
PROC SGSCATTER DATA = Sashelp.Iris (WHERE = (Species = "VERSICOLOR"));
  TITLE "Versicolor Length and Width";
  COMPARE Y = (Sepalwidth Petalwidth)
              X = (Sepallength Petallength) / REG ELLIPSE SPACING = 4;
RUN;
```

SGRENDER

This procedure identifies the data set that contains the plot variables and the template that is used to graph. The templates should be written using Graphic Template Language (GTL). The SGRENDER procedure also gives you the option to specify the name of the output object and the label for the output object.

The following few examples shows how templates are produced using GTL and how they are used with SGRENDER procedure. Except the last graphic all the other graphics in the following examples can be produced using one of the other graphic procedures mentioned above.

GRAPH FROM A SIMPLE GTL TEMPLATE

In the following example, SGRENDER procedure uses the following template and the data set ‘class’ to create the scatter plot below. There are three blocks in a GTL template: Define block, Graph block, Layout block.

```
PROC TEMPLATE;
  DEFINE STATGRAPH Scatter;
    BEGINGRAPH;
      ENTRYTITLE "Weight vs Height";
      LAYOUT OVERLAY;
      SCATTERPLOT Y = Weight X = Height;
      ENDLAYOUT;
    ENDGRAPH;
END;
RUN;

PROC SSGRENDER DATA = Sashelp.Class TEMPLATE = Scatter;
RUN;
```
In the following example, option `datalabel` is used.

```
PROC TEMPLATE;
  DEFINE STATGRAPH Scatter;
  BEGINGRAPH;
    ENTRYTITLE "Weight vs Height";
    LAYOUT OVERLAY;
      SCATTERPLOT Y = Weight X = Height / DATALABEL = Name;
    ENDLAYOUT;
  ENDGRAPH;
END;
RUN;

PROC SGRENDER DATA = Sashelp.Class TEMPLATE = Scatter;
RUN;
```
In the following, a **Cubic Regression** is added. The **DEGREE** option can take integer values between 0 and 174.

```plaintext
PROC TEMPLATE;
   DEFINE STATGRAPH Scatter;
      BEGINGRAPH;
         ENTRYTITLE "Scatter Plot of the Class Data Set with Fit Functions";
         LAYOUT OVERLAY;
            SCATTERPLOT Y = Weight X = Height;
            REGRESSIONPLOT Y = Weight X = Height / DEGREE = 3;
         ENDLAYOUT;
      ENDGRAPH;
   END;
RUN;

PROC SGRENDER DATA = Sashelp.Class TEMPLATE = Scatter;
RUN;
```
The following example creates a four-panel display of a data set. The **ROWGUTTER** adds horizontal gap between cells, and **COLUMNGUTTER** adds vertical gap between cells.

PROC TEMPLATE;
  DEFINE STATGRAPH Panel;
  BEGINGRAPH;
    ENTRYTITLE "Paneled Display ";
    LAYOUT LATTICE / ROWS = 2 COLUMNS = 2 ROWGUTTER = 10 COLUMNGUTTER = 10;
      LAYOUT OVERLAY;
        SCATTERPLOT Y = Weight X = Height;
        REGRESSIONPLOT Y = Weight X = Height;
      ENDLAYOUT;
      LAYOUT OVERLAY / XAXISOPTS = (LABEL = 'Weight');
        HISTOGRAM Weight;
      ENDLAYOUT;
      LAYOUT OVERLAY / YAXISOPTS = (LABEL = 'Height');
        BOXPLOT Y = Height;
      ENDLAYOUT;
      LAYOUT OVERLAY;
        SCATTERPLOT Y = weight X = height / GROUP = sex NAME = "Scat";
        DISCRETELEGEND "Scat" / TITLE = 'Sex';
      ENDLAYOUT;
    ENDLAYOUT;
  ENDGRAFPH;
END;

PROC SGRRENDER DATA = Sashelp.Class TEMPLATE = Panel;
RUN;

**Graphics and Data Visualization**

**LANED DISPLAY**

![Graphical representation of paneled display](image-url)
ODS DESTINATIONS for GRAPHS

The graphs produced by these procedures can sent to any ODS destination. The following example sends the plot to a PDF destination.

```
ODS PDF FILE = 'Scatter.pdf';
PROC SGPLOT DATA = Sashelp.Class;
  SCATTER X = Height Y = Weight;
RUN;

ODS PDF CLOSE;
```

ODS STYLES for GRAPHS

ODS STYLES control the overall appearance of the graphics produced by ODS Graphics. Some of the styles used are DEFAULT, STATISTICAL, ANALYSIS, JOURNAL, SCIENCE, and RTF. The following example creates SCIENCE style pdf output. CLI creates prediction limits for the individual predicted values.

```
ODS PDF FILE = 'style.pdf'  STYLE=SCIENCE;
PROC SGPLOT DATA = Sashelp.Class;
  TITLE 'Science Style Output for PDF';
  SCATTER X = Height Y = Weight;
  REG X= Height Y= weight / CLI;
RUN;

ODS PDF CLOSE;
```
SPECIFYING SIZE AND IMAGE FORMAT FOR GRAPHS

The graphs size and image format can be changed as in the following example. For RTF destinations PNG is the default format. JPEG, JPG, and GIF image formats are also supported for RTF. Different ODS destinations are supported by different sets of image formats.

```
ODS GRAPHICS / RESET IMAGEFMT = JPEG
HEIGHT = 2in WIDTH = 2in;
ODS RTF FILE = 'a.rtf';

PROC SGPLOT DATA=SASHELP.CLASS;
  SCATTER X=HEIGHT Y=WEIGHT;
  REG X=HEIGHT Y=WEIGHT/ CLI;
RUN;

ODS RTF CLOSE;
```

CONCLUSIONS

The SG procedures SGPLOT, SGSCATTER, SGPANEL can be used quickly to graph most of the statistical graphics. To create customized graphics or layouts, one needs to first create an appropriate template using GTL language and use SGRENDER procedure with this template. All these procedures create very high quality graphics. Many other graphs can be created using SG procedures. You could find those graphs in the references below.

REFERENCES

http://support.sas.com/documentation/onlinedoc/graph/index.html


ACKNOWLEDGMENTS
SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

DISCLAIMER
This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review by the Census Bureau than its official publications. This report is released to inform interested parties and to encourage discussion.

CONTACT INFORMATION
Your comments and questions are valued and encouraged. Contact the author at:
Selvaratnam Sridharma
Economic Planning and Coordination Division
U.S. Bureau of the Census
Address
Washington, DC 20233-6100
301-763-6774
Email: selvaratnam.sridharma@census.gov