Using PROC SGSCATTER to Explore Variable Relationships

To explore the relationships between variables, you can use the SGSCATTER procedure. PROC SGSCATTER creates a paneled graph of scatter plots for multiple combinations of variables. You can specify options to overlay fit plots and ellipses on your scatter plots. In PROC SGSCATTER, you must use one of the following three statements: COMPARE, MATRIX, or PLOT. Each statement produces a different type of scatter plot panel.

The COMPARE statement creates a comparative panel of scatter plots with shared axes. All cells in a row share the same row axis although the plots might have different ranges. All cells in a column share the same column axis although, again, the plots might have different ranges. You must include the X= argument to specify one or more variables for the X axis, and the Y= argument to specify one or more variables for the Y axis.

The MATRIX statement creates a scatter plot matrix. Each variable that you specify is graphed against each other variable. In the MATRIX statement, you must specify at least two numeric variables for the matrix.

The PLOT statement creates a paneled graph that contains multiple independent scatter plots. In other words, each scatter plot has its own independent set of axes. In the PLOT statement, you must specify one or more plot requests. Each plot request specifies the variables to plot and produces a separate cell.

PROC SGSCATTER offers many different options that you can specify to define the features of your graph. Two popular options that can be used in any of the three statements are ELLIPSE= and GROUP=. ELLIPSE= adds a confidence or prediction ellipse to the scatter plot. GROUP= specifies a classification variable to divide the values into groups. For information about increasing the size of graphs produced by PROC SGSCATTER, click the Information button.

Close