

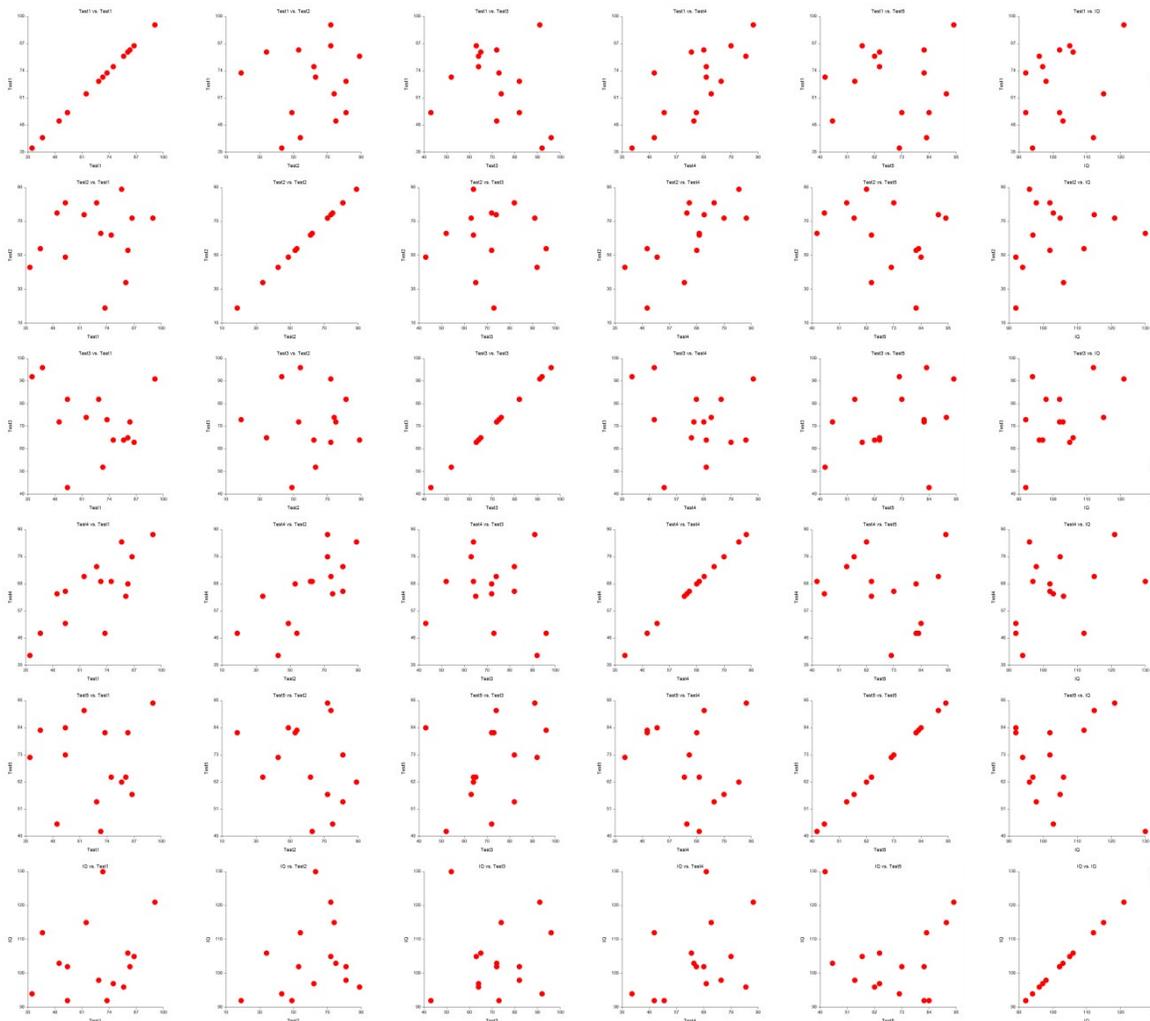
Chapter 162

Scatter Plot Matrix

Introduction

A *scatter plot matrix* is table of scatter plots. Each plot is small so that many plots can be fit on a page. When you need to look at several plots, such as at the beginning of a multiple regression analysis, a scatter plot matrix is a very useful tool.

Following is an example of a scatter plot matrix created during the initial phase of a multiple regression study. Five test scores form the pool of independent variables and the subjects IQ value is the dependent variable. Notice how quickly you can scan the plots for highly correlated variables and for outliers.



Data Structure

Each scatter plot is constructed from two numeric variables. A third, alphanumeric, variable may be used to control the plot symbol.

Procedure Options

These options are used to define the scatter plot matrix.

Variables Tab

This panel specifies which variables are in the scatter plot.

Variables

Vertical Variable(s)

Enter two or more variables to be displayed vertically down the page. If you leave this field blank, the horizontal variables will be used as the vertical variables as well.

Horizontal Variable(s)

Enter two or more variables to be displayed horizontally across the page. If you leave the Vertical Variable(s) field blank, these variables will also be used as the vertical variables.

Grouping (Symbol) Variable

This variable may be used to designate different plotting symbols for different rows of data. For example, you might want to use different plotting symbols to distinguish observations from different groups. The appearance of each symbol is designated on the Symbols Tab.

The symbol variable may also be used to “watch” certain rows throughout the dataset. For example, suppose that two or three rows appear as outliers in one of the plots. It might be useful to observe these rows in other plots. You would do this as follows. Add a new variable to your dataset that consists of zeros for all rows except the two or three of interest. Set the values of those rows to one and specify this new variable as the Symbol Variable. You will be able determine where those rows occur in all the plots in the scatter plot matrix.

Data Label Variable

A data label is text that is displayed beside each point. This field contains the name (or number) of the variable containing the labels. The values may be text or numeric. The size, style, and color of the text may be modified by pressing the second button to the right of the text box. This button brings up the text settings window.

Frequency Variable

Specify an optional frequency (count) variable. This variable contains integers that represent the number of observations (frequency) associated with each observation.

Weight Variable

A weight variable may be specified to set the (non-negative) weight given to each observation. By default, each observation receives an equal weight of $1/n$ (where n is the sample size). This variable allows you to specify different weights for different observations.

Scatter Plot Matrix

Format Options

Number of Plots Per Row

This option determines the upper limit for the number of plots on a single row.

Variable Names

This option selects whether to display only variable's name, label, or both.

Value Labels

This option selects whether to display only values, value labels, or both. Use this option if you want the group variable to automatically attach labels to the values (like 1=Yes, 2=No, etc.).

Symbol Size Options

Symbol Size Variable

The values of this optional column control the proportional size of the plot symbols.

Example 1 – Creating a Scatter Plot Matrix

This section presents a tutorial on generating a scatter plot matrix. To run this example, take the following steps using the Tests dataset.

You may follow along here by making the appropriate entries or load the completed template **Example 1** by clicking on Open Example Template from the File menu of the Scatter Plot Matrix window.

1 Open the Tests dataset.

- From the **File** menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Tests.NCSS**.
- Click **Open**.

2 Open the Scatter Plot Matrix window.

- Using the Graphics menu or the Procedure Navigator, find and select the **Scatter Plot Matrix** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

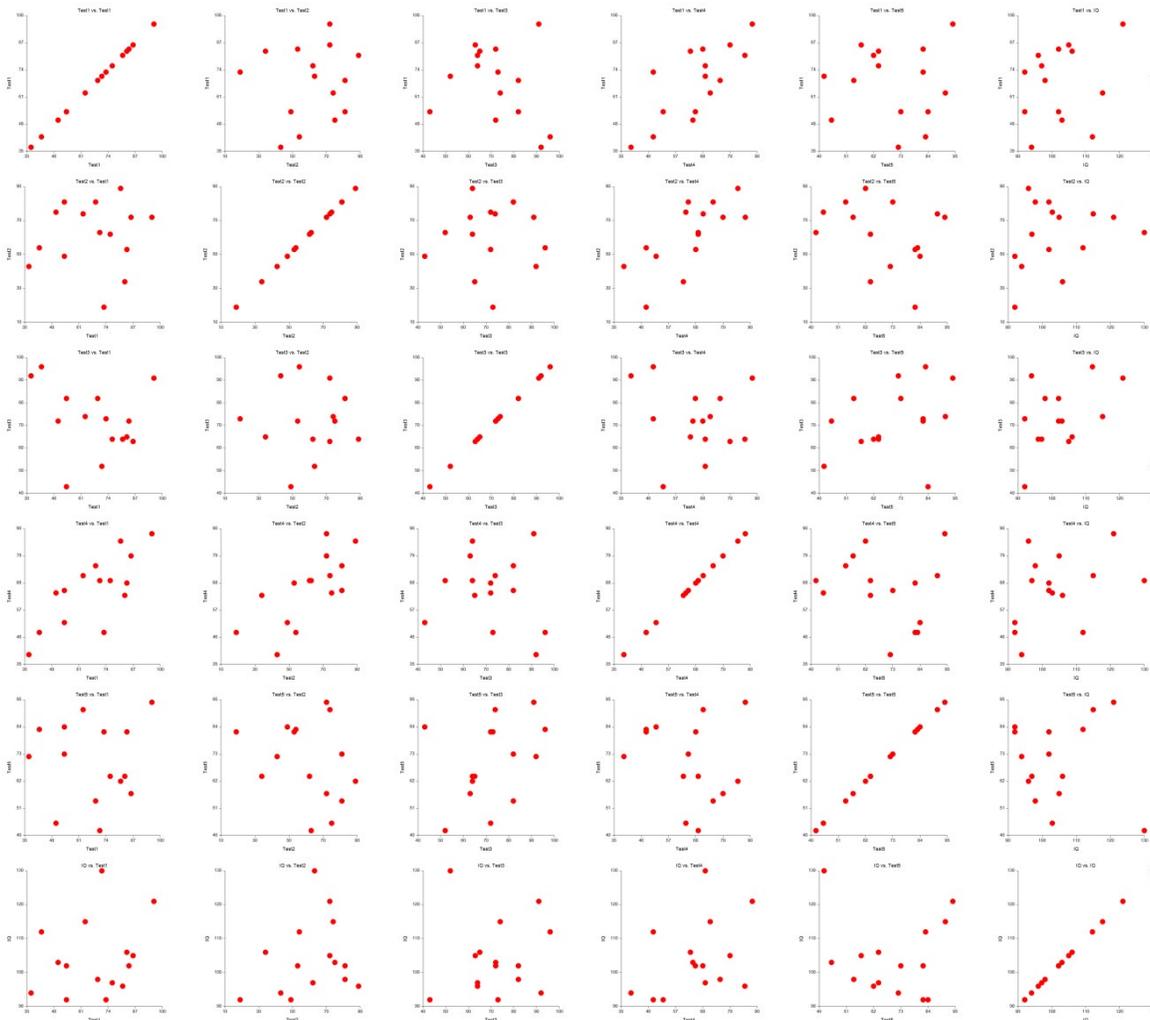
- On the Scatter Plot Matrix window, select the **Variables tab**.
- Double-click in the **Horizontal Variables** text box. This will bring up the variable selection window.
- Select variables **Test1,Test2,Test3,Test4,Test5,IQ** from the list of variables and then click **Ok**. "Test1-IQ" will appear in this box.
- Clear the **Vertical Variable(s)** box.

4 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

Scatter Plot Matrix

Scatter Plot Matrix Output



This report displays the scatter plot matrix. In the output screen you can double-click any of the plots to see it in full size.