SAS Publishing

SAS/SPECTRAVIEW® 9.1

User's Guide

The Power to Know®

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2004. *SAS/SPECTRAVIEW® 9.1 User's Guide*. Cary, NC: SAS Institute Inc.

SAS/SPECTRAVIEW® 9.1 User's Guide

ISBN 1-59047-213-6

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st printing, January 2004

Contents

PART 1

Techniques

CHAPTER

1

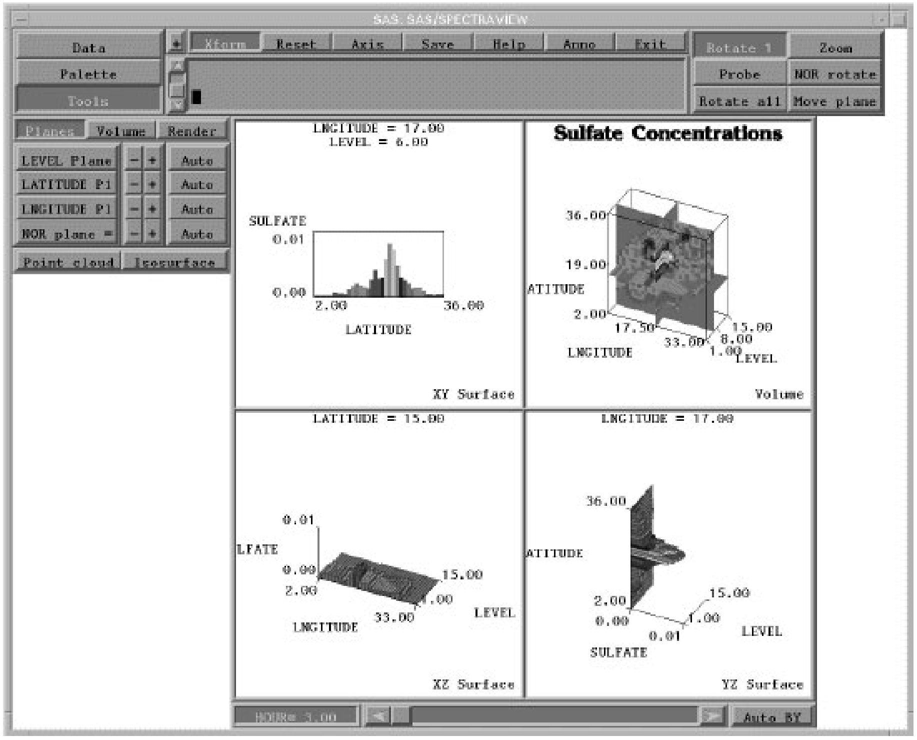
Getting Started

What is SAS/SPECTRAVIEW Software?

SAS/SPECTRAVIEW software is an interactive visualization product for viewing, exploring, and analyzing data as graphic images. The software reads a SAS data set, then based on your requests, creates two- and three-dimensional graphic images that represent the data. You can then visually explore both surface and interior views of the data to identify and analyze trends and relationships represented by the data.

Display 1.1 on page 4 shows four different images created by SAS/SPECTRAVIEW from a data set containing data about sulfate concentrations in the air. The image in the Volume window (top right window) displays intersecting cutting planes, which produce slices of data perpendicular to each axis. The other three windows show two- and three-dimensional surface views of the data at the position of the cutting planes.

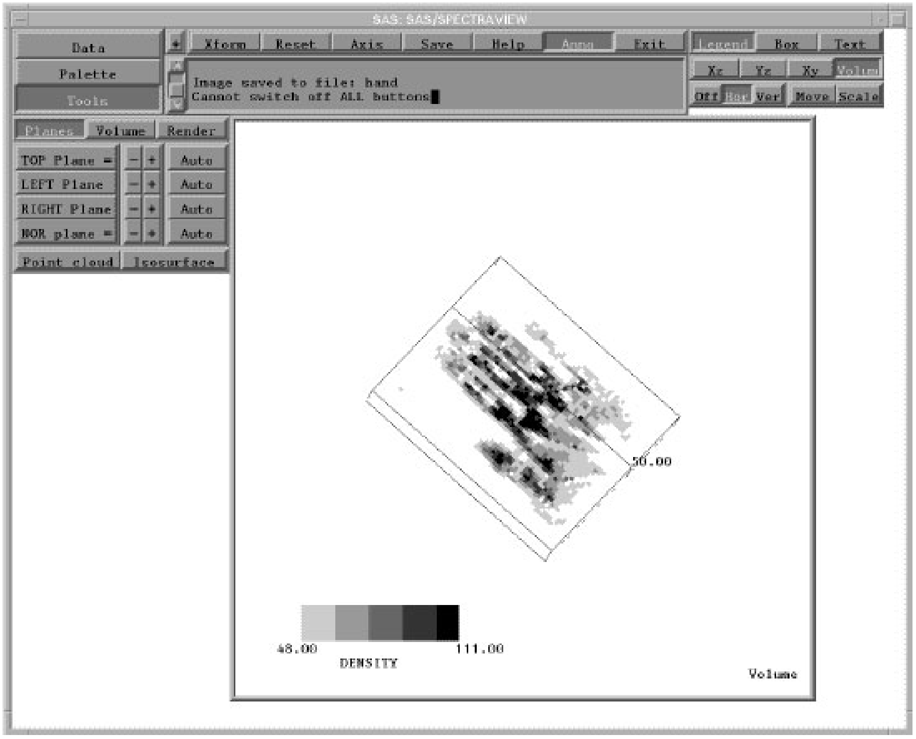
Display 1.1 SAS/SPECTRAVIEW Interface Displaying Images



The software's capabilities apply well to a variety of applications, including oil exploration, environmental sciences, chemical analysis, pharmaceutical studies, financial analysis, and medical imaging.

For example, medical professionals might perform a CT scan of a patient's hand to determine the severity of a fracture. Then, using various SAS/SPECTRAVIEW tools, data from the scan can be explored from several angles, resulting in more information than static two-dimensional images. Display 1.2 on page 5 shows a SAS/SPECTRAVIEW image of a hand created by requesting a cutting plane, which slices perpendicular along its axis. By probing the image, the software displays a density reading at a specific location.

Display 1.2 Multidimensional Medical Image



Using SAS/SPECTRAVIEW Software

Data Set Requirements

To use SAS/SPECTRAVIEW, your data must be stored in a SAS data set, which consists of variables and observations. A *variable* is a column in the data set, such as quantities or characteristics being measured, that has attributes such as a name and a type (character or numeric). An *observation* is the horizontal component of the data set, such as collections of values associated with a single entity; each observation contains one value for each variable in the data set.

The SAS data set that you use must have at least four variables:

- three variables, which can be character or numeric, to be the *axis variables*, that is, the X variable, Y variable, and Z variable. Each axis variable must contain at least two unique values. For best results, the axis variables should contain *discrete data*, which consists of distinct values (noncontinuous) containing natural gaps, like patient IDs and years.
- a fourth variable, which must be numeric, to be the *response variable*. The response variable must contain at least two unique values. For best results, the response variable should contain measured or modeled response values (like sales or population) that are related in some way to the axis values.

You can specify an optional fifth variable, which can be character or numeric, as a *BY variable*. A BY variable allows you to animate an image so that you can see how response values change according to some grouping, like over time.

How the Software Displays Data

When you load a SAS data set into SAS/SPECTRAVIEW, the software does the following:

1 The software reads each observation from the data set and creates a three-dimensional *volume grid* by plotting the values for the axis variables, which creates the x,y,z coordinates. That is, for each observation, the following occurs:
   - The value for the X variable is plotted on the X axis, which is the horizontal axis.
   - The value for the Y variable is plotted on the Y axis, which is the vertical axis.
   - The value for the Z variable is plotted on the Z axis, which is the depth axis.

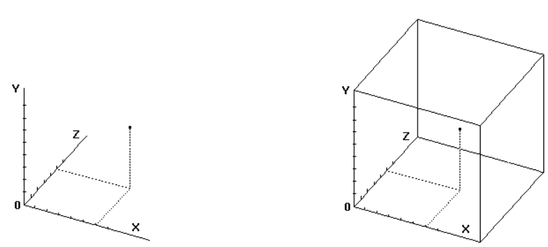
   The volume grid is actually an invisible network of lines that intersect in three-dimensional space. Each intersection of an x,y,z coordinate is a location in three-dimensional space, referred to as a *data point*. The shape and size of the volume grid is determined by the number of unique X, Y, and Z values and is displayed with a *bounding box*, which is a set of lines that outline the three-dimensional volume grid.

2 The software divides the values for the response variable (the response values) into ranges, then color codes them using default ranges and preset colors. According to the response value (and its associated response value range color), the software then maps an appropriate color to each data point that has an associated response value. *Note that you can customize the ranges and colors as explained in Chapter 4, "Setting Response Value Colors for Images," on page 45.*

Spatial Data

The variables that you specify for the axes frequently (but not always) represent dimensions of spatial data. For example, in a spatial diagram like the following cubic volume, the x,y,z coordinate 6,5,5 represents a location that is 6 ticks along the X axis, 5 ticks along the Y axis, and 5 ticks along the Z axis (counting from an origination point shown as 0 in this figure):

Figure 1.1 Coordinate 6,5,5 in Three-Dimensional Space



To illustrate the relationship among the axis values and the response values, consider the following spatial data example:

To determine the age of someone sitting in a specific seat in a stadium, you need to know the section, the row, and the seat number to locate that person. The three values are x,y,z coordinates that identify a specific location in space, which in this case is a stadium. Once located, the person can be asked his age; that number becomes the fourth value...the response value. You could collect the same information for everyone in the stadium. That is, you could attach an age response value to each location identified by section, row, and seat.

If you created a data set of the seating information and loaded it into SAS/SPECTRAVIEW, the following would occur:

- The resulting data points would form a volume grid displayed within a bounding box.
- The data points would represent the seat holders.
- The response value for each data point would be the age of the occupant.

You could then explore the data visually and determine, for example, whether age groupings occur in various locations in the stadium. You could also display any empty seats, which have no response value at that location.

Locations in three-dimensional space are similar to stadium seat locations. For example, if you want to test the amount of sulphur in the air at various locations, you would need three coordinates similar to section, row, and seat. These might be 20 km east, 10 km north, and 200 meters up. The coordinates describe a specific location in space where a sulphur sample can be taken and recorded. When you display the data, a color is mapped to each response value, representing ranges of values, for example, values between 0.0 and 0.5 could be red, values between 0.51 and 1.0 could be yellow, and so on.

Non-Spatial Data

Since the axis variables represent different dimensions of data, you can use SAS/SPECTRAVIEW to explore non-spatial data as well.

For example, the sample data set MORTGAGE (which contains mortgage payments for various numbers of years, interest rates, and loan amounts) can be represented several ways. The axis variables could be principal amount, percentage rate, and term of loan; it does not matter which variable you assign to X, which to Y, and which to Z. This presumes that the PAYMENT variable is the response you want to explore. Or you could assign the variables so that AMOUNT is the response to explore, with axes of term, rate, and affordable payment range.

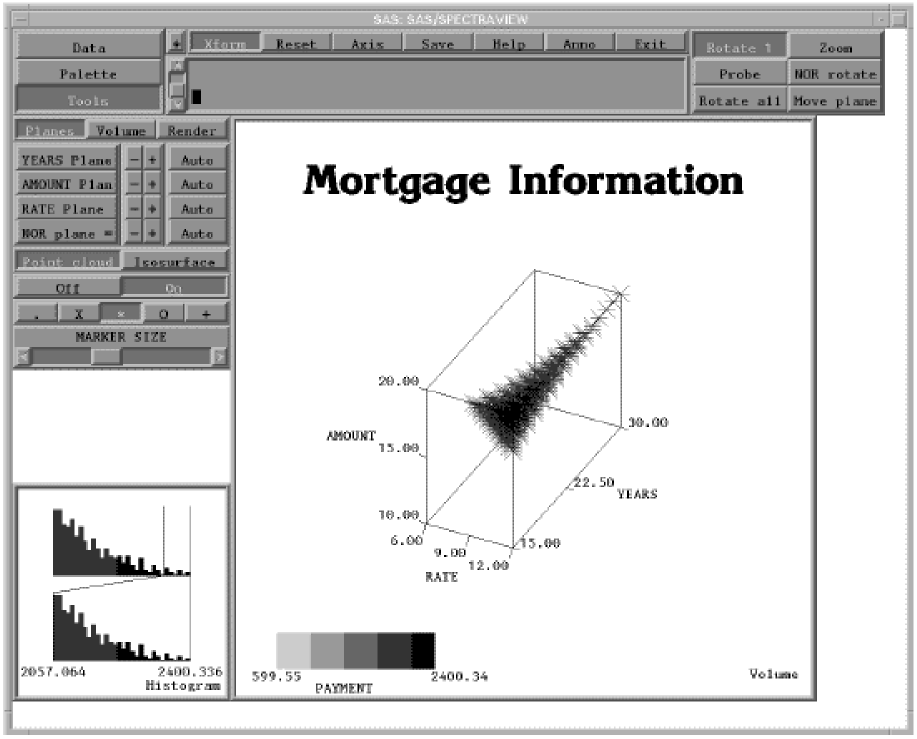
Assume the following variables for SAS/SPECTRAVIEW:

RATE	as the X variable, which is the loan interest percentage rate.
AMOUNT	as the Y variable, which is the loan amount.
YEARS	as the Z variable, which is the number of years for the loan.
PAYMENT	as the response variable, which is the monthly payment amount.

SAS/SPECTRAVIEW reads the data and generates the horizontal X axis (representing the loan interest rate), the vertical Y axis (representing the loan amounts), and the depth Z axis (representing the number of years for the loan). Each resulting data point is applied a color representing a response value, which is a monthly payment amount.

Display 1.3 on page 8 examines the data using a point cloud (which is one of the SAS/SPECTRAVIEW visualization techniques), showing the relationship of the monthly payments to percentage rate, loan amount, and length of the loan. This point cloud displays a subset of the data points, showing only the higher loan payments (response values). With this point cloud, you can determine that most of the higher loan payments are for a shorter number of years.

Display 1.3 MORTGAGE Data Set Displayed as a Point Cloud

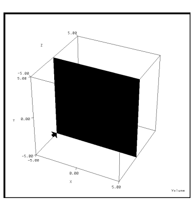


Summary of Software Tools

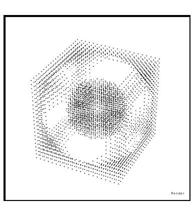
SAS/SPECTRAVIEW provides the following tools that you use to create and analyze images.

Visualization Techniques

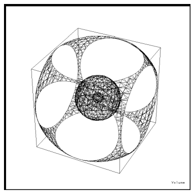
Visualization techniques create images representing your data.



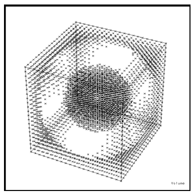
Cutting planes produce slices of data, either perpendicular to an axis or not perpendicular to any axis. In addition to the axis cutting planes, you can request two- and three-dimensional surface views (surfaces, charts, stacks, and plots) at a specific cutting plane's location.



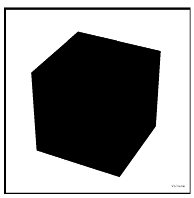
Direct volume rendering creates a two-dimensional image of the entire volume of data points with transparency.



Isosurface produces a three-dimensional surface by connecting all the data points with one response value.



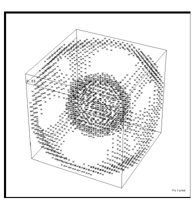
Point cloud displays response values with colored markers, showing individual data points.



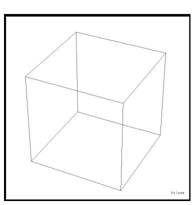
Solid-volume image produces a colored, three-dimensional solid-block image of the data points, providing a surface view of the data at the volume's borders.

Customization Options

SAS/SPECTRAVIEW provides a variety of options that you can use with visualization techniques to aid in data exploration and analysis:



Probe analysis displays the response value for a specific data point in an image.



Bounding box displays a set of lines that outline the three-dimensional volume representing the data in the Volume window. By default, the bounding box is on, but you can turn it off.

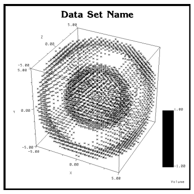


Image annotation allows you to add the following to two- and three-dimensional images: axis labels that name the variables associated with the axes, major and minor tick marks along the axes, text, and a response legend (a visual key to the colored data points).

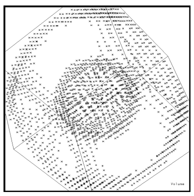
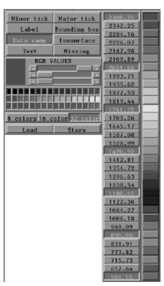


Image transformations let you view a three-dimensional image from any angle or in a different size. That is, you can rotate, move, and zoom an image.



Color palette lets you customize colors for the response value ranges, missing values, an isosurface, and image annotations (such as tick marks, axes labels, text, and the bounding box). You can also save and recall user-defined color palette files.

Processing Options

These options let you modify your view of the data:

BY variable
is a variable specification whose values define groups of observations, such as hour, month, or year. Specifying a BY variable allows you to animate an image so that you can see how response values change according to some grouping, like over time.

Categorizing data
groups numeric data to create discrete ranges for the X, Y, and Z axes, which results in reducing the number of data points created for the volume grid.

WHERE clause

allows you to specify a subset of data to be read into the software.

Data filtering
smooths or sharpens data to deemphasize or highlight variations in response values. The software provides several predefined filters, or you can create your own.

Data saving
lets you save all or part of the original data values to a SAS data set.

Image saving
lets you save a displayed image as a TIFF (Tagged Image File Format) file or a PostScript file.

Getting Acquainted with the Interface

The SAS/SPECTRAVIEW interface is a point-and-click, button-driven interface that provides the functionality to visually explore data.

Invoking SAS/SPECTRAVIEW Software

1 First, bring up the SAS System. *For information on bringing up the SAS System, see your SAS Installation Representative.*

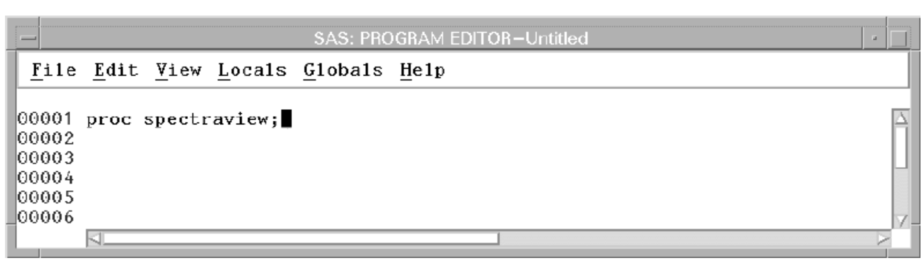
2 Issue one or more LIBNAME statements to assign a libref to each data library containing data sets that you want to use in the session. For example, **`libname mylib 'mypath';`**. Details are provided in "Assigning Librefs to Data" on page 22.

3 Then, invoke SAS/SPECTRAVIEW with a command or as a SAS procedure. On a command line, issue the command **`spectraview`**; acceptable abbreviations include **`spectrav`** and **`sview`**.

Display 1.4 Invoking SAS/SPECTRAVIEW with a Command

Or issue the SPECTRAVIEW procedure, which can be abbreviated to **`proc spectrav;`** or **`proc sview;`**. *Note that you do not need to include a RUN statement.*

Display 1.5 Invoking SAS/SPECTRAVIEW as a Procedure

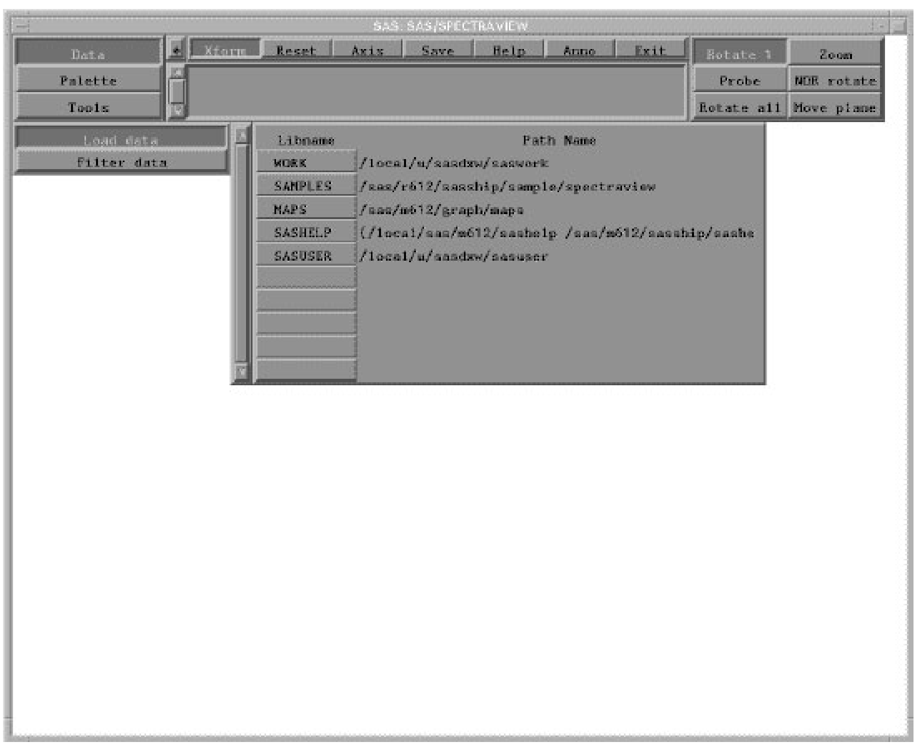


Looking at the Initial Display

Once you invoke the software, the initial interface display appears. By default, the interface is ready for you to load data; that is, Data and Load data are already selected. From the initial display, you can do the following:

- □ Load a SAS data set, which you must do before you can do anything else. That is, you cannot select any other buttons until you have loaded a data set. Details are provided in Chapter 2, "Loading Data," on page 17.
- □ Apply data filtering, which is explained in Chapter 3, "Adjusting Data with Filters," on page 37.
- □ Customize colors for your images with the color palette. From the color palette, you can define response value ranges and specify colors for those ranges. You can also specify colors for image customizations like text, axis tick marks, and so on. Of course, you can simply use the software defaults. Details are provided in Chapter 4, "Setting Response Value Colors for Images," on page 45.
- □ Request the various visualization techniques that produce the images such as cutting planes, point clouds, isosurfaces, a solid-volume display, and direct volume rendering. Details are provided in Chapter 5, "Exploring Data with Visualization Techniques," on page 53.

Display 1.6 SAS/SPECTRAVIEW Initial Display



Selecting Buttons

You interact with the software by selecting buttons with your graphics input device, which can be a mouse, a pick, or a stylus. SAS/SPECTRAVIEW treats all graphics input devices the same. For simplicity, this document refers to the graphics input device as a mouse.

To select a button, position the cursor on it, and click the left mouse button. Once selected, the button becomes shaded, and in many cases, additional buttons appear. *Note that in some situations, the software will select a button for you by default.*

Using the Windows

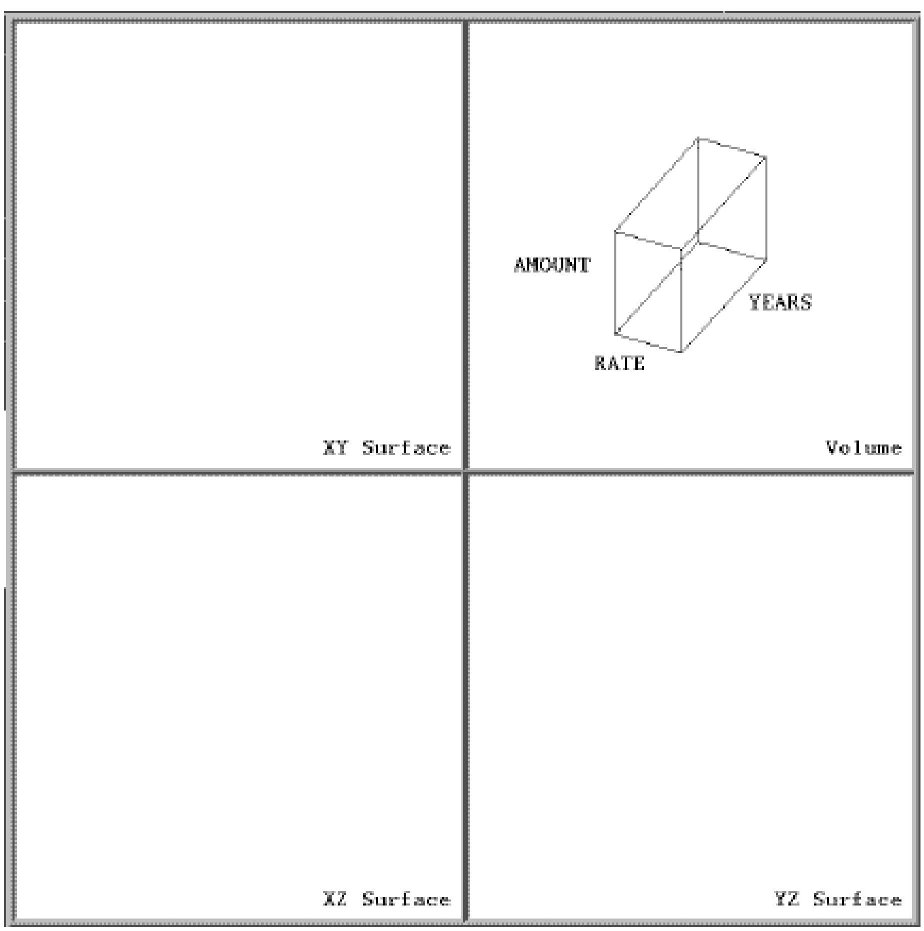
You view and manipulate images in display windows, and you communicate with the software in the text window.

Display Windows

There are four *display windows*, which provide different views of an image:

- The top right window, labeled Volume, displays the three-dimensional image or images, according to the technique(s) and options that you specify. For example, by requesting a point cloud, you can view all the data points in this window.
- The other three windows, labeled XY Surface, XZ Surface, and YZ Surface, display two- and three-dimensional images of your data. These are linked to the Volume window and show a single "slice" of your data.

Display 1.7 Display Windows



Instructions to resize the display windows are in “Resizing the Display Windows” on page 84.

Text Window

Using the *text window*, which is located to the right of the three basic function buttons (that is, Data, Palette, and Tools), you can enter text and receive messages from the software.

For example, when you position the cursor on a button, a one-line help message, which is usually a description or instruction, displays in the text window. In response, you usually select a button, enter a value, or use the cursor in a window. You also use the text window to enter specifications, such as for a WHERE clause or to enter text to be added to an image.

Display 1.8 Text Window

Instructions to resize the text window are in “Resizing the Text Window” on page 85.

Using Global Buttons

Global buttons, located at the top middle of the interface above the text window, display at all times. Global buttons provide the following capabilities:

Display 1.9 Global Buttons

Xform	To transform an image by rotating and zooming, to probe for a response value, or to manually move a cutting plane through the volume grid.
Reset	To reset the display window’s orientation, the color palette, or the entire software.
Axis	To define axis labels.
Save	To save an image as a PostScript file or a TIFF file or to save data values to a SAS data set.
Help	To display the online help.
Anno	To add annotation to an image by defining text, displaying the response legend, or displaying the bounding box.
Exit	To terminate SAS/SPECTRAVIEW software.

Accessing Online Help

SAS/SPECTRAVIEW provides online help for the interface buttons and sliders. By default, when you position the cursor on a button, a one-line help message displays in the text window. To display more detailed online help, you use the Help global button.

To access the detailed online help:

1 Select Help, then Detail. The software opens the Online Reference window.

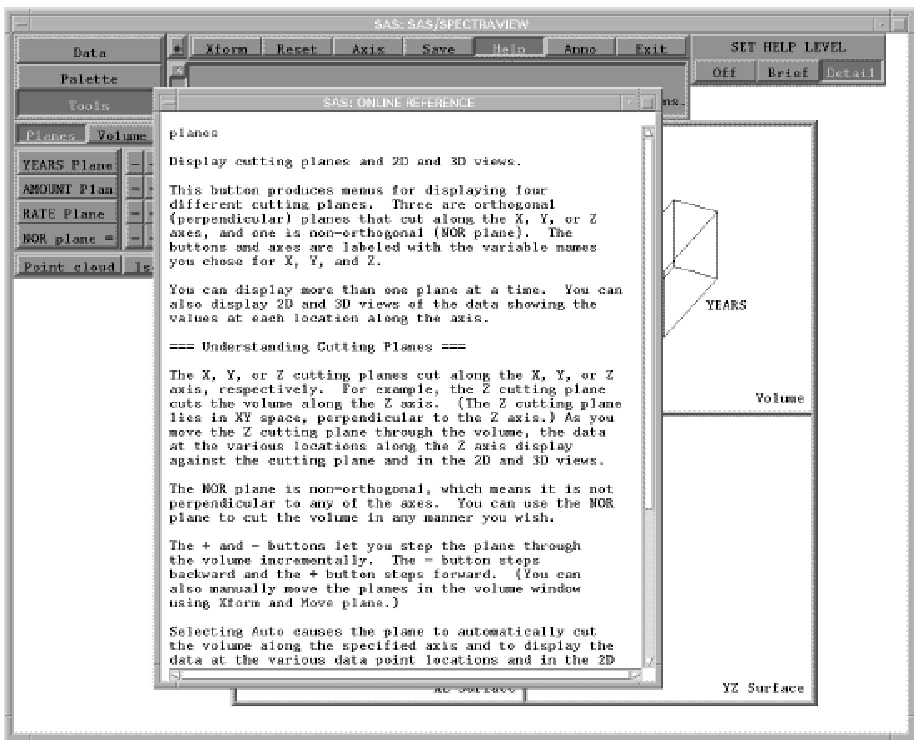
2 Move the window to a position so that you can view both it and the SAS/SPECTRAVIEW interface window. You can move the window by dragging it to any location. You can also pop it behind the interface.

In order for the Online Reference window to be visible, the SAS/SPECTRAVIEW interface should be set to pop-to-top only when you select its window border. If the Online Reference window does not stay on top, you need to modify your window manager files as explained in "Controlling the Online Reference Window" on page 86.

3 Click the SAS/SPECTRAVIEW interface window to make it the active window.

4 Position the cursor on a button or slider and wait a moment. *Note that to display online help, you only have to position the cursor on a button; you do not have to select it.*

Display 1.10 Online Reference Window



5 If the text is longer than the window, shift the text vertically with the scroll bar on the right. Clicking the up arrow shifts the text up; clicking the down arrow shifts the text down. You can also drag the scroll bar itself up or down.

To remove the Online Reference window and continue the one-line help message in the text window, select Help, then Brief. To turn off both, select Off. *Note that you cannot close the Online Reference window from the window itself.*

Using the Detailed Online Help

The text for each button follows a top-down structure. Information that you need frequently appears at the top. Concepts and other information that you may read only once appears lower down. Each button's information

- starts with a general description of that button's function
- describes subsequent button choices next (if applicable)
- ends with more detailed information, such as, information on understanding a tool.

Recognizing and Using Cursors

As you use the interface, you will notice several cursors presented as symbols. These symbols represent different actions to help you perform an activity. The default cursor is an arrow. The default cursor will change for some activities. For example:

- When the software is waiting for you to enter a response, such as for a filename when saving an image, the default cursor changes to an input cursor, which is a standard I-beam (eye-beam) symbol.
- When you request to rotate an image, the cursor changes to a directional symbol, indicating the directions for dragging the mouse.
- When you are deleting text, a hand symbol displays so that you can point to the text line to be deleted. The hand symbol also lets you indicate where you want axis labels to appear.

Exiting the Software

You exit SAS/SPECTRAVIEW with the Exit global button, which is located at the top right of the interface. Exiting SAS/SPECTRAVIEW returns you to the SAS System.

1. Select Exit. The button becomes highlighted, indicating that you need to confirm the request.
2. To confirm the request, select Exit again.

CHAPTER
2

Loading Data

Introduction to Loading Data

This section provides information that explains how SAS/SPECTRAVIEW loads data and how you can affect the results, and the section provides instructions on how to load data into SAS/SPECTRAVIEW .

Understanding the Volume Grid

When data is loaded into SAS/SPECTRAVIEW , the software creates a three-dimensional volume grid by plotting the values for the axis variables along the X, Y, and Z axes. Each intersection of an x,y,z coordinate is a data point in three-dimensional space. The shape and size of the volume grid is determined by the number of unique X, Y, and Z values.

The resulting total number of data points can be calculated by multiplying the number of unique X values * unique Y values * unique Z values. For example, if you have 10 X-axis values, 5 Y-axis values, and 2 Z-axis values, the result is 100 data points (10x5x2). If you have 10 values on each axis, the result is 1,000 data points (10x10x10).

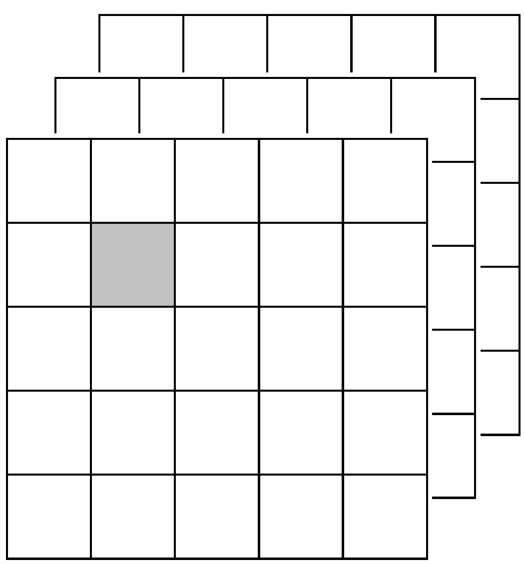
Loading Data Representing a Complete Grid

Data that represents a complete grid contains at least one set of x,y,z coordinates for each possible X, Y, Z variable combination. That is, when loading data, each time SAS/SPECTRAVIEW finds a unique axis value, the software creates a new grid intersection. For the grid to be complete, the data set must contain corresponding X, Y, and Z values for each possible intersection. The resulting number of data points would be the same as the number of observations in the data set, with the data points uniformly distributed in the volume grid, unless there are duplicate observations for a set of x,y,z coordinates. *Note that SAS/SPECTRAVIEW works best with data that represents a complete grid.*

Examples of data that would result in a complete grid is an air quality survey that includes a full grid of sample data from an entire area, scientific numerical models, medical images, or complete financial models like a mortgage table.

To have an idea of how much data is required for a complete grid, think of it like a three-dimensional spreadsheet where multiple sheets extend along the Z axis and where each cell on each sheet represents the values for one observation. Suppose the variables ROW represents X, COLUMN represents Y, and SHEET represents Z. The values ROW=2, COLUMN=2, and SHEET=1, which is one observation, would be located in the spreadsheet as shown in Figure 2.1 on page 18.

Figure 2.1 Three-Dimensional Spreadsheet



For a complete column 2, you would need these observations:

```
ROW   COLUMN   SHEET
1     2        1
2     2        1
```

3	2	1
4	2	1
5	2	1

For a complete sheet 1, you would need observations for all five columns:

ROW	COLUMN	SHEET
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
1	2	1
2	2	1
3	2	1
4	2	1
5	2	1
1	3	1
2	3	1
3	3	1
4	3	1
5	3	1
1	4	1
2	4	1
3	4	1
4	4	1
5	4	1
1	5	1
2	5	1
3	5	1
4	5	1
5	5	1

Finally, to complete the entire grid, you would need all those observations for sheet 2 and for sheet 3.

Loading Data Representing an Incomplete Grid

Data that represents less than a complete grid is data that does not have every possible combination but has at least one of the three values for X, Y, or Z. For example, data that represents an incomplete grid could be an air quality survey that consists of samples from random locations within a certain cubic area.

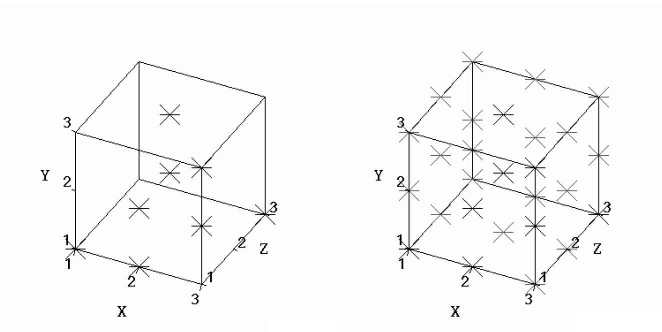
For an incomplete grid, when the software plots the actual axis values, any grid intersections without a data point are completed with software-generated *filler points* for the missing X, Y, or Z values to complete the grid.

For example, consider the following eight observations, which contain three unique values for each axis:

OBS	X	Y	Z	Response
1	1	1	1	111
2	2	1	1	211
3	3	1	3	313
4	3	2	1	321
5	3	3	1	331
6	2	2	1	221
7	2	2	2	222
8	2	3	2	232

The software would generate and plot 27 data points (3x3x3) — 8 actual data points representing the observations and 19 filler points as shown in Figure 2.2 on page 20. The first volume grid shows the actual data points; the second volume grid shows the actual data points and the filler points.

Figure 2.2 3x3x3 Volume Grid

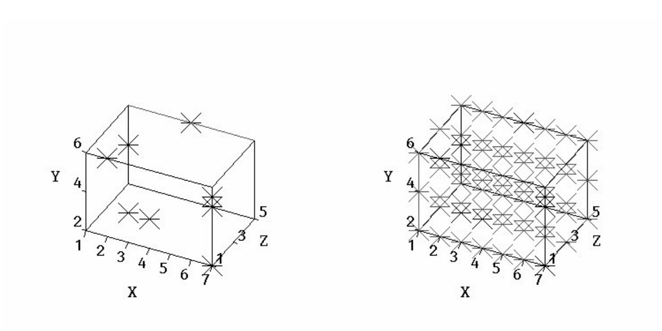


The larger the number of unique values for an axis, the larger the resulting number of data points. For example, consider the following eight observations, which contain 7 unique values for the X axis, and three unique values for the Y and Z axes.

OBS	X	Y	Z	Response
1	1	4	5	145
2	3	2	3	323
3	2	2	3	223
4	4	6	5	465
5	6	4	3	643
6	7	2	1	721
7	5	2	5	525
8	1	6	1	161

The software would generate and plot 63 data points (7x3x3) – .8 actual data points representing the observations and 55 filler points as shown in Figure 2.3 on page 20. The first volume grid shows the actual data points; the second volume grid shows the actual data points and the filler points.

Figure 2.3 7x3x3 Volume Grid



Loading Sparse Data

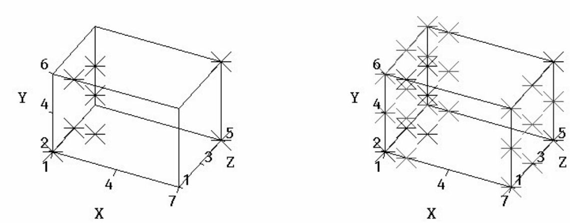
Data that does not contain at least one value for an x,y,z coordinate within the volume grid is referred to as *sparse* data. Generally, sparse data occurs when the unique values for an axis are widely distributed along the axis, for example, an air quality survey where an entire section of a test area was not sampled. And often, sparse data is not related spatially, for example, a data set where the X, Y, and Z values are height, weight, and age. *Note that sparse data can also result from subsetting.*

Unlike for locations having at least one value for x,y,z coordinate, the software does not replace non-existent x,y,z coordinates with filler points. Instead, the volume grid displays a visual gap indicating an area within the volume grid where no data is available. The actual data points appear to be non-uniformly distributed because of the gap in the data. Consider the following data, which contains three unique values for the axis variables:

OBS	X	Y	Z	Response
1	1	4	5	145
2	1	2	3	123
3	2	2	3	223
4	7	6	5	765
5	2	4	3	243
6	1	2	1	121
7	7	2	5	725
8	2	6	1	261

When the actual data values are plotted and the volume grid is completed, the actual data points are not uniformly distributed, resulting in a volume grid that appears to have gaps. The software would generate and plot 27 data points (3x3x3) – 8 actual data points representing the observations and 19 filler points as shown in Figure 2.4 on page 21. The first volume grid shows the actual data points; the second volume grid shows the actual data points, the filler points, and visual gaps:

Figure 2.4 Sparse Data Volume Grid



Note that when loading character data, gaps will not occur. The software assigns sequential numerical values to the character values, resulting in uniformly distributed data points.

Understanding Missing Values

A missing value is a value in the SAS System indicating that no data is stored for the variable in the current observation. In SAS/SPECTRAVIEW , any grid intersections with missing X, Y, or Z values or any x,y,z coordinate without an associated response value are completed with software-generated *filler points*. Filler points are handled as missing values.

Missing values, by default, have no color. If you want missing values to display in an image, you must use the color palette to assign a color as explained in "Assigning Color to Missing Values" on page 49.

If your data represents an incomplete grid or sparse data, the software may create many filler points. However, if your data represents a complete grid, displaying missing values lets you see holes, which may indicate a possible failure of the measuring equipment.

Assigning Librefs to Data

The SAS System requires that data stored in a SAS library or a specific SAS data set in a SAS library must have a *libref* (library reference) assigned to it. A libref is an arbitrary name that you make up to symbolically represent a SAS library. The SAS System uses a libref to associate a SAS library with its physical location.

The SAS System automatically provides the librefs SASUSER and WORK. SASUSER is for permanent SAS data sets; WORK is for temporary SAS data sets that are discarded at the end of a SAS session. You must assign librefs for any data library other than the libraries associated with SASUSER and WORK.

To assign a libref, issue the LIBNAME statement. You must assign a libref outside SAS/SPECTRAVIEW, for example, from the SAS PROGRAM EDITOR window. The most common form of the LIBNAME statement contains only the libref and the physical location for the SAS library, for example:

```
libname mylib 'mypath';
```

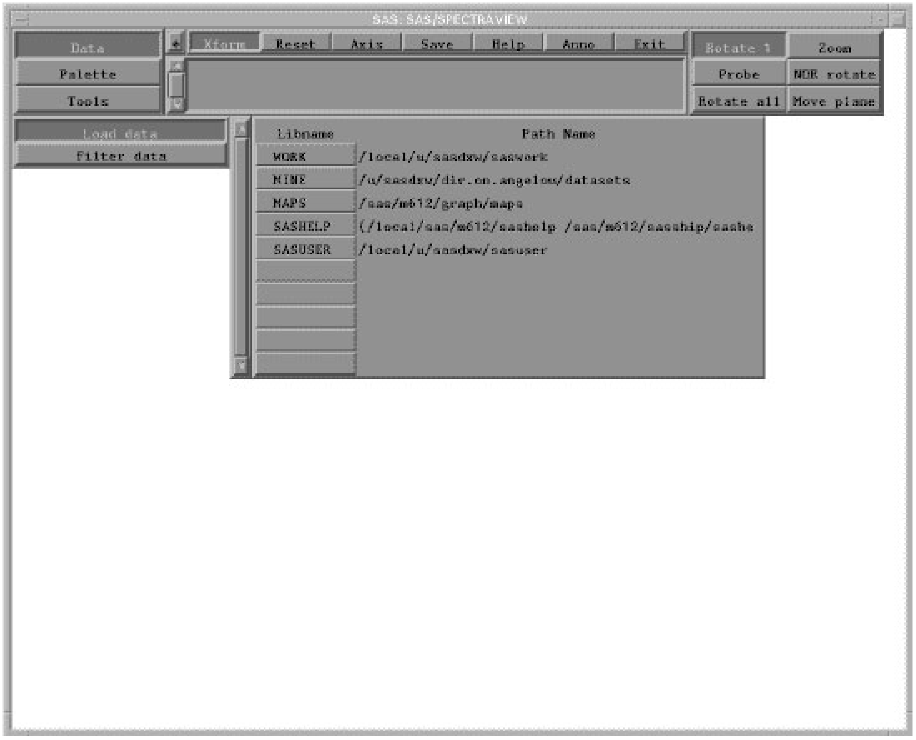
Note: You can include LIBNAME statements in an autoexec file, which is an external file containing SAS statements that are executed automatically when the SAS System is invoked. Then each time you bring up the SAS System, your librefs are automatically assigned. △

Loading the Data

The first step in the visualization process is selecting and reading your data into SAS/SPECTRAVIEW. The interface guides you through the process.

When you first invoke SAS/SPECTRAVIEW, Data is selected by default, ready for you to load data. Note that you can load data at any time during a SAS/SPECTRAVIEW session by reselecting Data.

Display 2.1 Loading Data



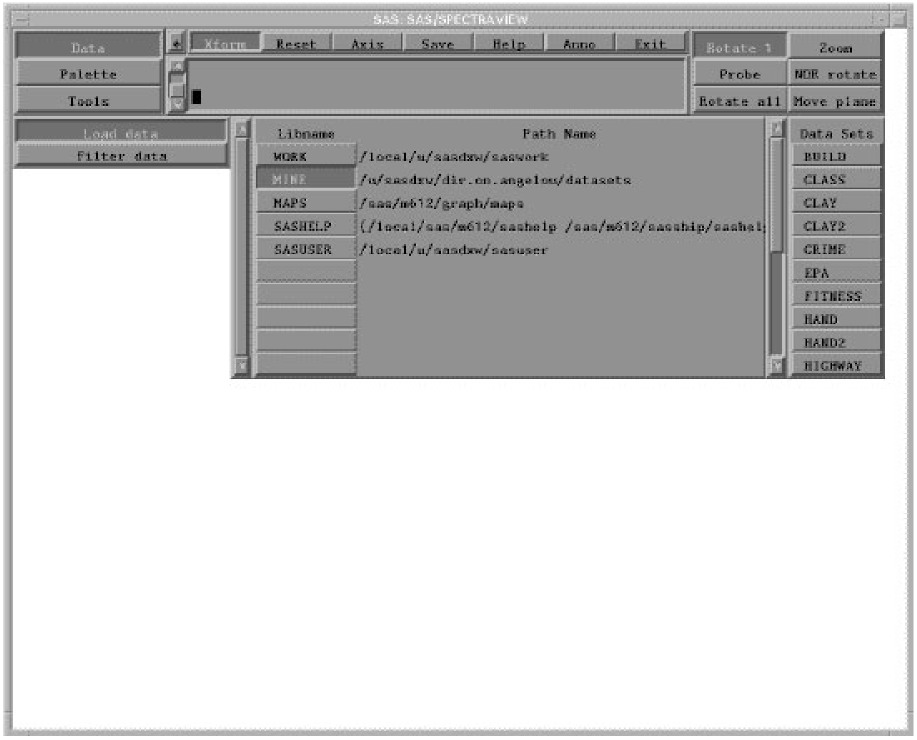
Selecting a Libref

To display the session's assigned librefs:

1 Select Load data. The software displays the assigned librefs under the label `Libname.` *To assign an additional libref for a session, you can do so from the SAS PROGRAM EDITOR window (if you invoked SAS/SPECTRAVIEW with a command), then refresh the session's librefs for SAS/SPECTRAVIEW by reselecting the Data and Load data buttons.*

2 Select the libref containing the data set that you want to load. Use the scroll bar if there are more than 10. Once you select the libref, the software displays the data sets associated with the libref.

Display 2.2 Selecting a Libref

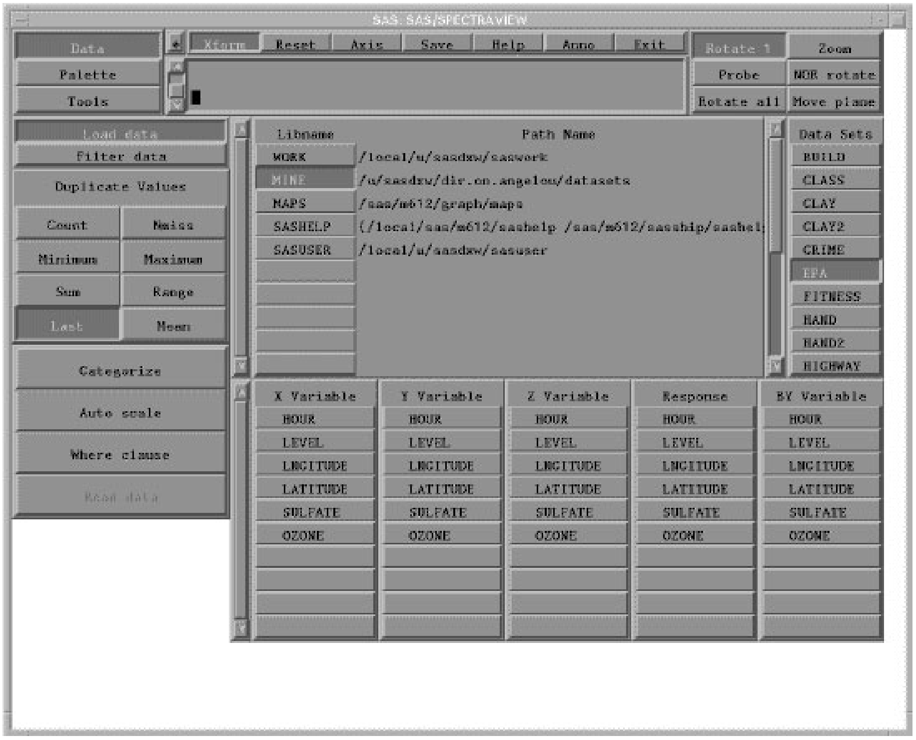


Selecting a Data Set

SAS/SPECTRAVIEW works as well with small data sets (such as 20 observations) as it does with large data sets (such as a quarter million observations). The SAS data set that you select must have at least four variables to be specified for the three axis variables and the response variable, the response variable must be numeric, and each variable specified for SAS/SPECTRAVIEW must contain at least two unique values. If you want to use a BY variable, the data set must have a fifth variable as well. *To load a data set that has only three variables, see "Loading a Data Set with Only Three Variables" on page 32.*

Select the input data set from the list of names. Use the scroll bar if there are more than 10. Once you select the input data set, the software lists the data set's variables in columns from which you can select SAS/SPECTRAVIEW variables.

Display 2.3 Selecting a Data Set



Specifying SAS/SPECTRAVIEW Variables

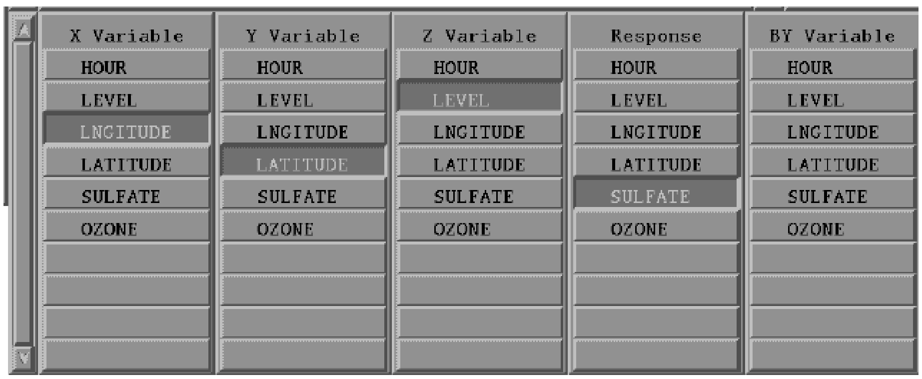
You must specify a different data set variable for each SAS/SPECTRAVIEW variable. That is, you must select a different variable from each of the **`X Variable`**, **`Y Variable`**, **`Z Variable`**, and **`Response variable`** columns. The axis variables can be either numeric or character, but the response variable must be numeric.

To help you select appropriate variables, you can place your cursor on a variable name, and the software will display a short description of it in the text window. For example, for the EPA data set, which contains the variables HOUR, LEVEL, LNGITUDE, LATITUDE, SULFATE, and OZONE, their descriptions provide the following information:

- All the variables are numeric. Specifically, the description for SULFATE is **`Type: Num, Label: Sulfate (ppm)`**.
- SULFATE and OZONE specify that their values are in ppm (parts per million). SULFATE and OZONE are good candidates for the Response variable, since you usually want a variable that is observed or generated in various quantities. A Response variable is one that contains the values that are of most interest.
- Variables LEVEL, LNGITUDE, and LATITUDE are described as RADM Model layer, RADM Cell X coordinate, and RADM Cell Y coordinate. Their values are most likely not sampled or generated but represent where SULFATE and OZONE values are located or from what types of data the response values were generated. Therefore, LEVEL, LNGITUDE, and LATITUDE are good candidates as axis variables, since they can be used to generate grid locations to display the response values.
- HOUR contains hour values. This type of variable is useful to generate groups of observations by assigning it as a BY variable as explained in "Grouping Observations with a BY Variable" on page 26.

Note that any variable that is appropriate as a Response variable is not a valid choice as an axis variable, and any variable that is appropriate as an axis variable is not a valid choice for a Response variable. Attempting to read a data set with inappropriate variables selected could result in the data set failing to load. You want to specify variables that are the best ones as the axis variables to build as complete a volume grid with actual data points as possible. And you want to avoid specifying axis variables that are sparsely valued or have continuous data.

Display 2.4 Specifying SAS/SPECTRAVIEW Variables



Once you select the four required variables, the software highlights Read data, but you still have the option of specifying BY variable processing, duplicate values handling, data categorizing, automatic axis scaling, and data subsetting with a WHERE clause, which are discussed in the following sections.

Grouping Observations with a BY Variable

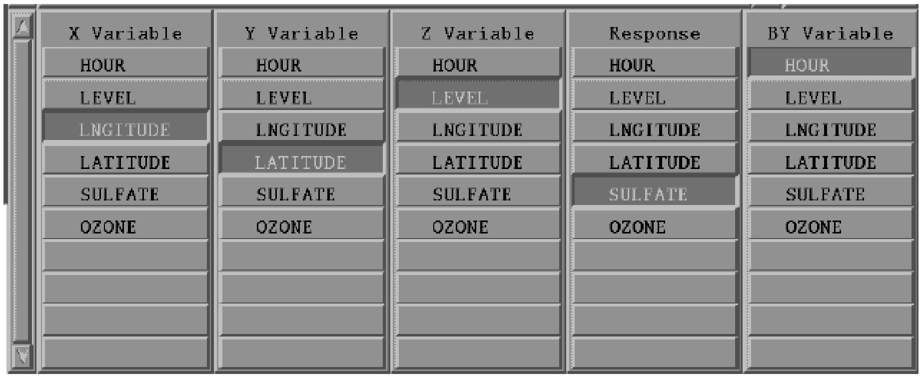
In addition to the four required variables, you have the option of specifying a fifth variable as a *BY variable*. The values of a BY variable define groups of observations, such as hour, month, or year. Specifying a BY variable allows you to animate an image so that you can see how response values change according to some grouping, like over time.

A BY variable can be either character or numeric. BY data usually includes multiple response values for a single data point.

For example, in the EPA data set, the variable HOUR contains hour values, which would be useful as a BY variable. If you imagine that the first four variables would generate a cube of data values, then specifying a BY variable would generate a sequence of cubes of data values that can be cycled through to determine how response values change over time (in this case).

If you select LNGITUDE, LATITUDE, and LEVEL as the axis variables, SULFATE as the Response variable, then HOUR as the BY variable, you will create a sequence of volumes of data to be displayed and analyzed.

Display 2.5 Specifying a BY Variable



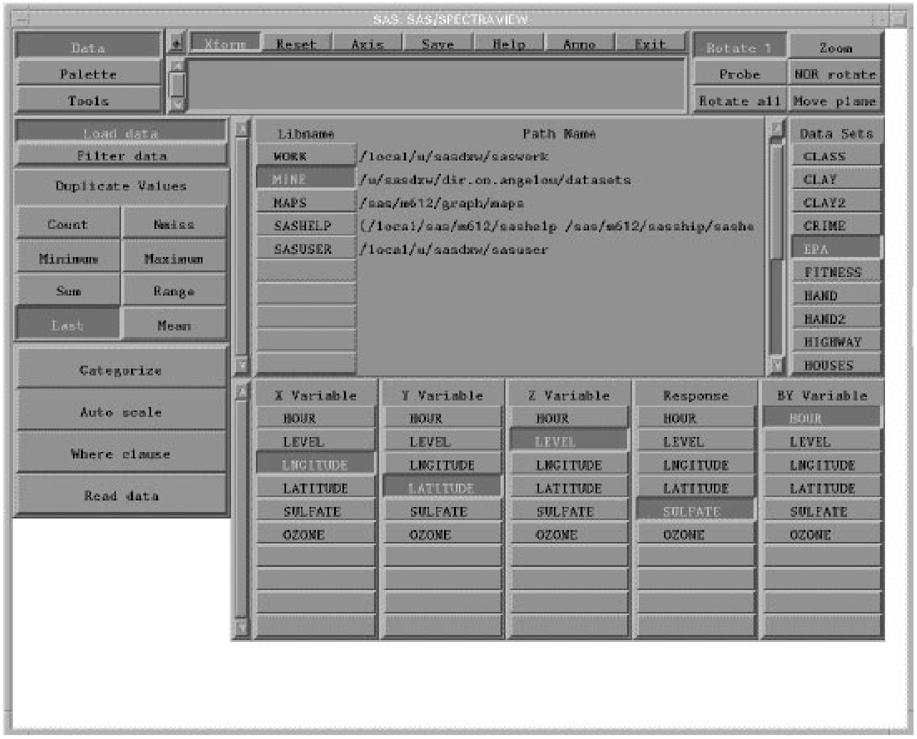
Note: If you do not specify a BY variable but your data contains BY data (like a time variable), you may receive a message in the text window after loading the data. The message warns that there is more than one response value for an x,y,z coordinate. When this occurs, the software handles the response values according to the setting on the **`Duplicate Values`** panel. △

Handling Duplicate Values

Duplicate values occur when the data has more than one observation for the same x,y,z coordinate, which could result in more than one response value for a data point. Note that if you also categorize the data or if you have specified a BY variable, the instances of duplicate values may increase.

You determine how the software handles duplicate values by selecting one of the choices under the label **`Duplicate Values.`** The default is Last, which means that the last response value encountered for a data point is used as that location's response value.

Display 2.6 Handling Duplicate Values



To specify how the software handles duplicate values, select one of the following options:

Count
: For each unique x,y,z location, the software counts the number of observations and uses that count as the response value. For example, if there are three observations that specify the x,y,z location 1,1,1, the response value is 3, regardless of the actual response values in the data.

When you load data, each response value for the resulting data points represents a count of the observations for that location. If there are no duplicate observations for a particular x,y,z location, the response value is 1, indicating that only one observation was found for that location. Similarly, if the data includes no observations for a particular x,y,z location, the response value would be 0, meaning that the data point is missing. Count allows you to find the number of response values that were used to calculate other values, for example, Mean or Sum. If you load data with Mean, you may want to know how many values were used to calculate the mean value shown at a particular x,y,z location. You can load again using Count, then probe the data to reveal the number used for the mean.

Nmiss
: For each unique x,y,z location, the software counts the number of observations with missing response values. For example, if an x,y,z location has two observations and both have a valid response value, the result is a response value of 0, meaning no observations with a missing response value were found for that location.

With Nmiss specified, every data point has a response value indicating how many missing response values were encountered for that location. If a valid data point has five observations and only three had response values, then that data point's response value is 2, meaning two observations were found missing a response value for that location. Nmiss only counts valid data points having no response value. It does not count filler points generated by the software. If the data does not contain an observation for an x,y,z location, the software inserts a data point that has a missing response value. This means that if you load a data set, display it as a point cloud, and discover there are several missing values in the volume grid, you can reload the data with Nmiss selected and determine which missing values are caused by missing response values as opposed to missing axis values.

Minimum
: If there are two or more response values for the same x,y,z location, the software uses the minimum value as the response value.

Maximum
: If there are two or more response values for the same x,y,z location, the software uses the maximum value as the response value.

Sum
: If there are two or more response values for the same x,y,z location, the software uses the sum as the response value.

Range
: If the data contains at least two response values for each x,y,z location, the software uses the range as the response value. The range is calculated by subtracting the minimum response value from the maximum response value. If there is only one value for a location, the response value is set to missing.

Last
: If the data contains two or more response values for the same x,y,z location, the software uses the last response value as the response value. This is the default.

Mean
: If the data contains two or more response values for the same x,y,z location, the software uses the mean as the response value.

Categorizing Data

Categorizing data is an option that groups numeric data to create distinct ranges (called categories) for each axis. *You cannot categorize character variables.* The result is a reduced number of data points in the volume grid. By categorizing all three axes, you can set exactly how many data points the software will create. Categorizing data is useful

- for data that has a large number of unique values for one or more axes, which would result in a large number of data points making the data difficult to analyze or possibly unable to be loaded.
- when it is simply not important that each axis value be distinct. In many cases, measurements are more precise than necessary for successful analysis.

Continuous data (containing few gaps that vary slightly over a large range like weight and height) are a good candidate for categorizing. For example, to analyze a group of people's heart rate based on their age, activity level, and weight, the weight values, which would be in pounds like 139.5, 143.6, would be considered continuous. That is, it is not likely that any two people (let alone several) would have the same weight but a different age and activity level. Categorizing the weight values by creating weight categories for ranges of weight with one value to represent each category would make the data clearer and easier to use.

Discrete data (containing natural gaps like patient IDs and years) would probably not be as useful to categorize. But discrete data such as hour could be categorized into groups if the degree of precision can be reduced without losing data integrity.

To categorize data:

1 Select Categorize. The software displays a group of sliders and buttons at the bottom of the interface.

Display 2.7 Categorizing Data

2 Under the label **CATEGORIZE AXIS**, specify which axis you want to categorize. By default, all three are turned on for categorizing. Use the on/off buttons to turn categorizing on or off for a particular axis. For example, selecting X on turns on categorizing for the X axis, and selecting Y off turns off categorizing for the Y axis.

3 Under **NUMBER GROUPS**, use the sliders to specify the number of categories you want for each axis. You can specify between two and 100 categories for each, with 10 being the default.

4 Under **GROUP AXIS VALUE**, for each categorized axis, specify the axis tick mark value:

Lower	Uses the lower bound value in each range.
Midpoint	Uses the midpoint value in each range. This is the default setting.
Upper	Uses the upper bound value in each range.
Bounds	Uses both the upper and lower bound values in each range. The values display as a range, for example, **125.1-225.1** for each major tick mark.

Effect on Duplicate Values Handling

Categorizing data makes it more likely that the software encounters more than one response value for a given x,y,z coordinate. (Uncategorized data usually contain only one response value for each x,y,z coordinate.) When one or more of the axes are categorized, some of the data points become duplicates within a group, which could result in more than one response value for a single data point.

For example, suppose values for the X variable are integers from 1 to 100. If you categorize the X values into groups of 10 values, 1-10 would be a single category. The data points 1,1,1 and 2,1,1 and 3,1,1 and so forth are viewed by the software as the same data point in the volume grid, because they would all have the same X, Y, and Z values.

The response values for the 10 data points would appear to be 10 different response values for the same data point. The response values for the duplicate locations are handled according to the method specified for duplicate values handling, with the default being to use the last response value found as the category's response value.

Automatically Scaling Axes

By selecting Auto scale, you can automatically scale the volume's three axes to the same length. The default is that the length of each axis is determined by the range of axis values. For example, an axis with values from 1 to 100 is ten times as long as an axis with values from 1 to 10.

Note: Once a data set is loaded, Auto scale is deselected. To load a subsequent data set with automatic scaling, you must select Auto scale again. △

Subsetting Data with a WHERE Clause

Optionally, you can specify a subset of data to be loaded into SAS/SPECTRAVIEW by specifying condition(s) that observations must meet. You can subset response values by specifying criteria for the response variable, and you can subset data points by specifying criteria for the axis variables.

Subsetting can change the size and shape of the volume grid. For example, subsetting data can create holes that are replaced with filler points, or subsetting can remove holes in data.

Prior to selecting Read data, you can specify subsetting conditions using a SAS WHERE clause:

1 Select Where clause.
2 In the text window, type a SAS WHERE clause, without the keyword WHERE and no ending semicolon. A condition consists of a variable name, an operator (such as EQ, NE, LT), and a value, such as **`sulfate > .00005060`**.

Display 2.8 Subsetting Data

3 Press Enter.

For details on specifying conditions, see the appropriate WHERE clause documentation. *Note that before you invoke SAS/SPECTRAVIEW, you can create a smaller SAS data set containing only the values that you want to use. For example, you could choose certain ranges of axis values or specific response values.*

Reading the Data Set

To have the software read the data, select Read data.

The software loads the input data, applying any optional specifications. For example, if a WHERE clause is specified, the software loads only those observations meeting the criteria, and if categorizing is specified, the software changes the number of data points accordingly. Once the data set is loaded, the variable list disappears, and the software is ready for you to

- apply optional data filtering, which is explained in Chapter 3, “Adjusting Data with Filters,” on page 37.
- customize colors, which is explained in Chapter 4, “Setting Response Value Colors for Images,” on page 45.
- request the visualization techniques that produce the images, which is explained in Chapter 5, “Exploring Data with Visualization Techniques,” on page 53.

If you have loading problems, see “Resolving Data Loading Problems” on page 31.

Resolving Data Loading Problems

The following topics provide suggestions on how to resolve possible data loading problems. Note that if a data set fails to load, the software displays an error message in the text window.

Loading a Data Set with Only Three Variables

SAS/SPECTRAVIEW requires four variables in order to load a SAS data set. However, with the following procedure, it is possible to load a data set that has only three variables.

1 Create a temporary SAS data set with the following DATA step code:

```
data temp;
   set yourdatasetname;
   dummy=1;
   output;
   dummy=2;
   output;
run;
```

2 Load the temporary data set TEMP into SAS/SPECTRAVIEW.
3 Select the X and Y axis variables that you are interested in, then select DUMMY as the Z variable.
4 Select the Response variable that you want.
5 Select Read data.
6 Use the data for your analysis.

Note that the Z plane will have two identical planes (z=1 and z=2). You can ignore the second one.

Changing Axis Variables

Sometimes data will load with certain axes and response variables specified but will not with different ones due to memory constraints. You want to specify variables that are the *best* ones as the axis variables to build as complete a volume grid with actual data points as possible. That is, you want to avoid specifying axis variables that are sparsely valued or have continuous data.

For example, the sample data set MORTGAGE loads without problems if YEARS, RATE, and AMOUNT are specified as the axis variables. However, if you specify PAYMENT for an axis and either YEARS, RATE, or AMOUNT as the response variable, the data may not load, because there are 16,400 unique values for PAYMENT. *Note that if a data set fails to load, the error message in the text window specifies the number of unique values found for each axis.*

See "Specifying SAS/SPECTRAVIEW Variables" on page 25 for details on specifying variables and determining which variables are best.

Categorizing Data

One of the main reasons that a data set will not load is that the data does not represent a complete grid, which most often occurs with random data or if the axis values are continuous rather than discrete. The data set may fail to load due to memory constraints, even when a larger data set loaded successfully. The problem is the number of resulting data points in the volume grid, not the number of observations.

Memory requirements for a data set depend on the number of unique X, Y, and Z values, which determines the number of data points that are created. If the number of data points becomes large, the data set may fail to load without additional memory. Of course, it takes thousands and thousands of data points to cause data loading problems.

To make the data clearer and easier to use in SAS/SPECTRAVIEW, you can *categorize* the data, which groups numeric data to create distinct ranges (called categories) for each axis. Instructions on how to categorize data are in “Categorizing Data” on page 29.

Changing Duplicate Values Handling

Specifying how the software handles duplicate values can cause data not to load. For example, if you select either Count or Nmiss under the label **Duplicate Values** and the data you want to load comprises a complete grid having no missing x,y,z locations and no duplicate observations for the same x,y,z location, the data would fail to load. That is,

- With Count specified, the response value for every data point would be 1. The data would fail to load because Count requires at least two different response values for an x,y,z location.
- With Nmiss specified, the response value for every data point would be 0. The data would fail to load because Nmiss requires at least two different response values for an x,y,z location.

Instructions for specifying how the software handles duplicate values are in “Handling Duplicate Values” on page 27.

Removing BY Variable Specification

Removing the BY variable specification will cut the amount of storage required by the number of BY groups in the data set.

To calculate storage requirements for a BY variable, multiply the number of unique values for each axis variable by the number of BY groups. For example, if you have five BY groups, you would need five times as much storage, because a grid is created for each value of the BY variable.

More information on BY variable processing is in “Grouping Observations with a BY Variable” on page 26.

Using G4GRID Procedure to Create a Complete Grid

You can run the G4GRID procedure on data to create a data set that represents a complete grid. For example, if your data is random in nature, PROC G4GRID may be a good choice. The procedure produces data that is derived from the original data. The amount of time it takes to produce the new data set is based on the number of observations in the data set and the size of the requested output grid.

PROC G4GRID enables the loading of a data set that could not otherwise be loaded due to memory constraints. By using PROC G4GRID, you can fill in missing values with interpolated values or resize the data set as required. PROC G4GRID is useful when

- the response values were sampled at discrete locations, for example, measurements of air pollution.
- the response data is functionally related to the axis variables. That is, the response is either analytically or physically a function of the axis variables. Air pollution measurements are a function of discrete locations identified by axis values, but a stock’s price is not a function of a stock’s name. That is, just because Granny’s Kitchen stock price is high does not mean Gerry’s Garage stock price is high even though they fall next to each other in the grid. Smoothing with PROC G4GRID would lower Granny’s stock and raise Gerry’s stock because they would be assumed to influence each other.

□ you want a complete grid of values and can accept some changes from your original values.

Complete documentation for PROC G4GRID is in Appendix 1, "The G4GRID Procedure."

Calculating Volume Grid Storage Requirements

To understand how to calculate storage requirements, compare the following two DATA step examples.

The first example produces 9,261 observations and would load with no problems. In fact, it is a relatively small data set by SAS/SPECTRAVIEW standards. There are 21 unique values for each axis, which results in a grid that has 9,261 data points (21x21x21). Each data point requires approximately four bytes of storage on most machines. Therefore, it requires 4x9,261=~36KB of storage for the grid.

```
data load;
  drop a b c;
  a=0.3;
  b=0.2;
  c=0.1;
  do x = -1 to 1 by 0.1;
    do y = -1 to 1 by 0.1;
       do z = -1 to 1 by 0.1;
       response = x**2/a**2 + y**2/b**2 + z**2/c**2;
       output;
       end;
    end;
  end;
run;
```

The second example, however, may not load, even though it has only 100 observations. The number of unique X, Y, and Z values is unknown, but by using the RANUNI function, it can be assumed that it will be close to 100 for each variable. The grid, therefore, requires 100x100x100=1,000,000 data points or about 108 times (~4MB) the storage requirement as compared to the first example.

```
data noload;
   drop seed I a b c;
   seed = -1;
   a = 0.3;
   b = 0.2;
   c = 0.1;
      do I = 1 to 100;
         x = 2.0*ranuni(seed) - 1.0;
         y = 2.0*ranuni(seed) - 1.0;
         z = 2.0*ranuni(seed) - 1.0;
         response = x**2/a**2 + y**2/b**2 + z**2/c**2;
         output;
      end;
run;
```

Specifying Larger Memory Size

To specify a larger memory size, invoke the SAS System and specify the system option MEMSIZE, which controls how much memory the SAS System uses, with a larger memory size. For example,

`-memsize 100m`.

Note that SAS/SPECTRAVIEW also requires additional memory for overhead, some of which is proportional to the size of the data set. It is possible that, while there is enough memory to build the grid, some other area may not succeed, which will prevent the SAS data set from loading.

CHAPTER

3

Adjusting Data with Filters

Introduction to Adjusting Data with Filters

After loading data, you can apply a data filter to smooth or sharpen data. Data filtering is optional, but applying one can remove drastic changes in neighboring response values or clean out false data. Applying a data filter can also highlight or exaggerate the differences between adjacent response values.

For example, suppose you have measurement values from an instrument that has some degree of noise associated with its measurements:

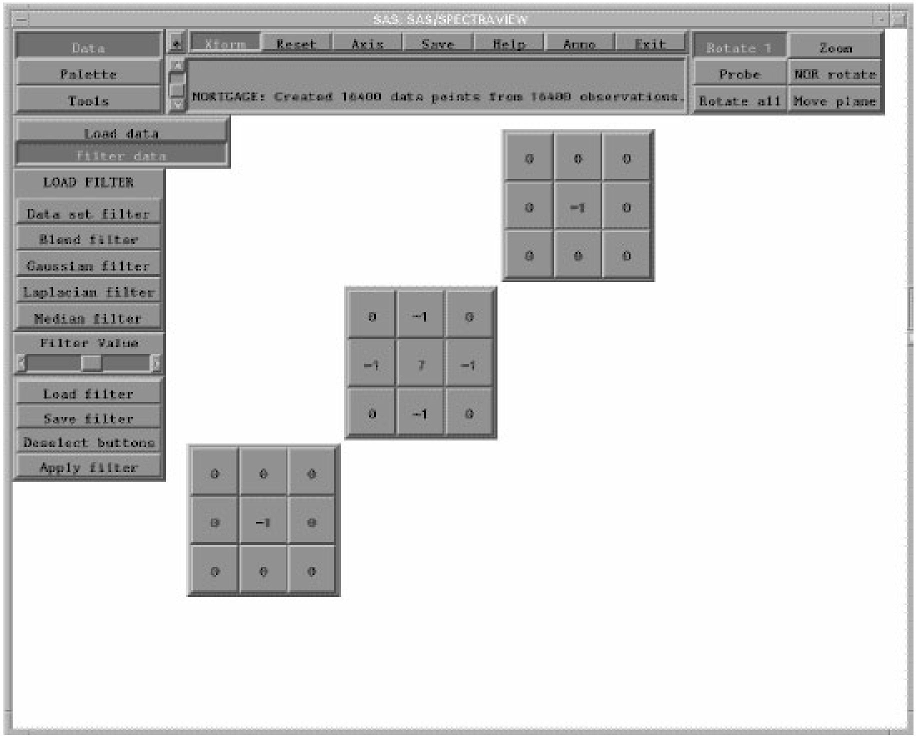
```
measure value = real value +/- instrument noise
```

To assist in analysis, you could smooth out the noise by applying a smoothing filter, like the Blend filter. Another example would be a data set representing a CT scan that you want smoothed or enhanced.

When you apply a data filter, the software adjusts the value for each response value in the data by performing a mathematical operation. In general, the operation replaces the response value being operated on by multiplying and averaging its value with the values of adjacent response values. Missing response values are ignored.

When you specify a filter, the software displays three pads of buttons that represent a 3x3x3 matrix. Each button (element) represents a response value location, with the center element representing the response value being operated on. For example, Display 3.1 on page 38 shows the matrix of preset values for the Laplacian filter (provided with the software), which sharpens data.

Display 3.1 Laplacian Filter Matrix of Preset Values



When you apply the Laplacian filter, the software does the following for each response value in the data:

1 The response value being operated on is multiplied by 7, which weights (increases) it.
2 Adjacent (surrounding) response values are multiplied by -1, which pulls down (decreases) their values.
3 The resulting values are then averaged, and the average replaces the value for the response value being operated on.

Using the Filters Provided with the Software

Understanding the Filters

SAS/SPECTRAVIEW provides four filters:

Blend filter
: smooths by averaging without weighting the response value.

Gaussian filter
: smooths by averaging and weights the response value.

Laplacian filter
: sharpens the data.

Median filter
: smooths by using the median value.

The filters are provided as templates and are preset with values that the software uses to multiply the response values. The templates differ only in their values, which you can use or modify to suit your needs. The Median filter does not provide values and cannot be modified.

Note that you can also create your own data filter, which is explained in “Creating and Applying Your Own Filter” on page 42.

Blend Filter

The Blend filter, which is the default, gently smooths data by averaging the response value with the values of adjacent response values. (In comparison, the Gaussian filter smooths by averaging and also weights the response value.) The Blend filter uses the following preset values, which you can modify:

```
1 1 1   1 1 1   1 1 1
1 1 1   1 1 1   1 1 1
1 1 1   1 1 1   1 1 1
```

Each element in the matrix has a value of 1. The multiplication does not change any values, so the response values are simply averaged to arrive at the replacement response value.

Gaussian Filter

The Gaussian filter smooths data by averaging and also weights the response value. (In comparison, the Blend filter smooths by averaging but does not raise the response value.) The Gaussian filter uses the following preset values, which you can modify:

```
0 0 0   0 1 0   0 0 0
0 1 0   1 4 1   0 1 0
0 0 0   0 1 0   0 0 0
```

The response value being operated on is multiplied by 4, which increases (weights) it. Adjacent response values are multiplied by 1, so they are not changed. The resulting values are then averaged, with the average replacing the response value being operated on. *Note that the Gaussian filter is useful to smooth out spikes in CAT scan data.*

Laplacian Filter

The Laplacian filter sharpens data by weighting the response value and pulling down the adjacent response values. The Laplacian filter uses the following preset values, which you can modify:

```
0 0 0    0-1 0   0 0 0
0-1 0   -1 7-1   0-1 0
0 0 0    0-1 0   0 0 0
```

The response value being operated on is multiplied by 7, which increases (weights) it. Adjacent response values are multiplied by -1, which decreases (pulls down) their values. The resulting values are then averaged, with the average replacing the response value being operated on. *Note that the Laplacian filter is helpful for sharpening contrast in data.*

Median Filter

The Median filter smooths data by using a median (middle) value. The Median filter is not a template. There are no values in the matrix; only asterisks, which cannot be changed.

The Median filter sorts the response value being operated on with surrounding response values in order, selects the median value, then replaces the response value being operated on with the median value. The median value is the value in the sorted list that has the same number of values below it as above it. All 27 values, except missing values, are included in each operation.

Combining Filters

You can apply multiple filters to the data to both smooth and sharpen data. For example:

- apply the Laplacian filter along with the Gaussian filter
- apply the Gaussian filter or Median filter along with the Laplacian filter
- apply the Laplacian filter along with the Median filter.

Note that the order in which you apply filters affects the result. That is, if you apply a Laplacian filter, then a Gaussian filter, the result will be different than if you apply the Gaussian filter first, then the Laplacian filter.

Applying a Filter Provided with the Software

The following instructions apply a data filter provided with SAS/SPECTRAVIEW :

1. Load the appropriate data set.
2. Select Data, then Filter data.
3. Specify the filter you want to load by selecting one of the following. The default is the Blend filter.

   Blend filter smooths by averaging without weighting the response value.

   Gaussian filter smooths by averaging and weights the response value.

   Laplacian filter sharpens the data.

   Median filter smooths by using the median value.
4. Select Apply filter to apply the filter to all response values in the data.

An Example Using the Blend Filter

The following steps illustrate the effects of applying the Blend filter:

1. Issue the following DATA step code to create a SAS data set named BUMPY, which produces a series of ellipses with additional random noise added to the response values:

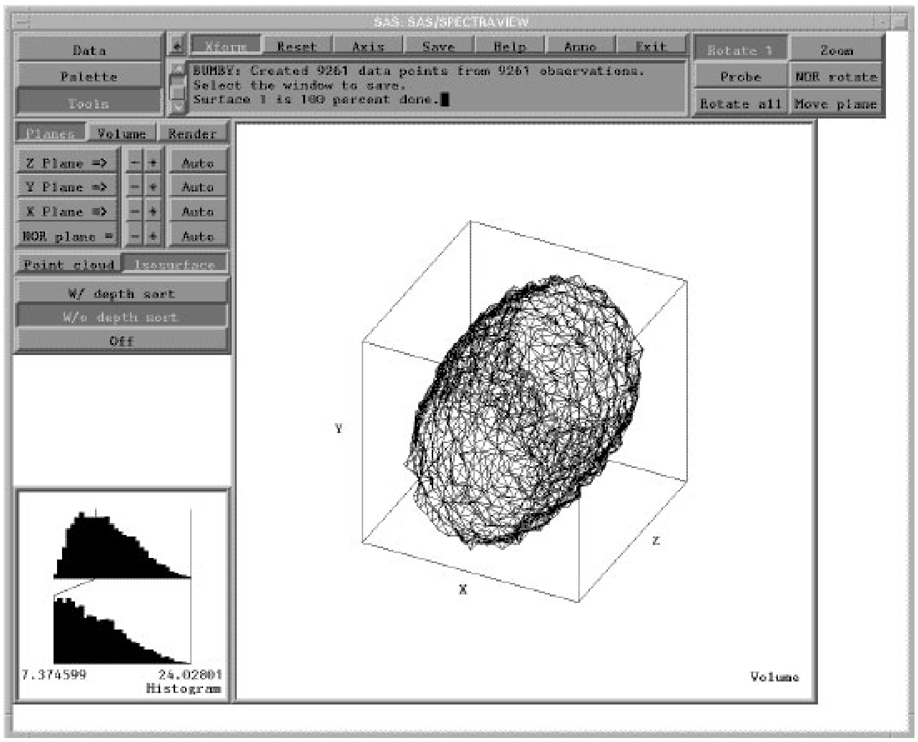
```
data bumpy;
  do x = -5 to 5 by 0.5;
  do y = -5 to 5 by 0.5;
  do z = -5 to 5 by 0.5;
     response = (x*x*0.5 + y*y*0.3 + z*z*0.1) + 2.0*ranuni(-1);
     output;
   end; end; end;
run;
```

2 Load the data set into SAS/SPECTRAVIEW .

3 Create an isosurface, which represents one response value:

- **a** Select Tools, then Isosurface.
- **b** To specify the response value, use the left mouse button and drag the histogram's left white line to a desired value. (The histogram appears at the bottom left of the interface.)
- **c** Then select W/ depth sort. The isosurface looks something like a football that is rather bumpy.

Display 3.2 Bumpy Isosurface



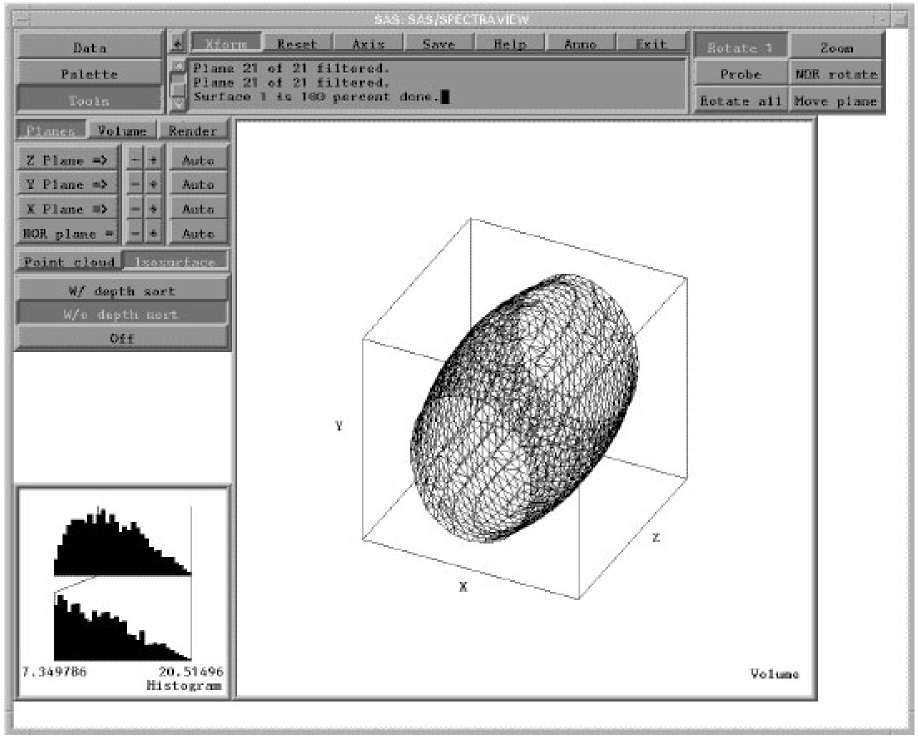
4 Next, smooth the data by applying a data filter:

- **a** Select Data, then Filter data.
- **b** Select Blend filter.
- **c** Select Apply filter.

5 Follow the same instructions as in Step 3 to re-create the isosurface.

The result is that removing the noise in the data displays a much smoother surface, which can help in viewing overall trends in data.

Display 3.3 Smoother Isosurface



Creating and Applying Your Own Filter

To create a filter to suit your own needs, you can customize a supplied filter by modifying its preset values, or you can define a filter using SAS DATA step code.

Customizing a Filter

The following steps create a filter by modifying the preset values of a filter provided with SAS/SPECTRAVIEW:

1. Load the appropriate data set.
2. Select [Data], then [Filter data].
3. Specify the filter template you want to customize by selecting one of the following. You cannot customize the Median filter.

   [Blend filter] smooths by averaging without weighting the response value.

   [Gaussian filter] smooths by averaging and weights the response value.

   [Laplacian filter] sharpens the data.
4. Select the filter element(s) that you want to have the same value, and adjust the slider under **`Filter Value`** to set the value.
5. Select [Deselect buttons] to release the selected elements.
6. If needed, adjust more element values.
7. Select [Apply filter] to apply the customized filter to all response values in the data.
8. To reuse the customized filter, store it in a SAS data set by selecting [Save filter] and specifying a one- or two-level SAS data set name in the text window.

Creating a Filter with SAS DATA Step Code

Using SAS DATA step code, you can create a SAS data set containing a filter, which can then be loaded into SAS/SPECTRAVIEW. The following restrictions apply:

- The data set must contain exactly four numeric variables, named as follows and in this order: FILTER_X, FILTER_Y, FILTER_Z, FILTER.
  - FILTER_X, FILTER_Y, and FILTER_Z are the indices of the filter for X, Y, and Z. The valid range for these values is 1, 2, and 3.
  - FILTER is the value for a given set of indices. Its range of values is -16 to 16, inclusive.
- All values must be integers or are converted to integers when read in. The data set must have at least 27 values, one for each location in the filter. More than 27 values produces a warning message. Incomplete filters are aborted. The data set cannot contain missing values.

The following DATA step code creates a SAS data set named BLEND, which contains a valid filter:

```
data blend;
  do filter_x = 1 to 3;
  do filter_y = 1 to 3;
  do filter_z = 1 to 3;
    filter = 1;  output;
  end; end; end;
run;
```

Applying a Filter Stored in a Data Set

The following instructions apply a filter stored in a SAS data set:

1. Load the appropriate data set.
2. Select Data, then Filter data.
3. Select Load filter and enter the data set name in the text window.
4. Select Data set filter to display the filter matrix.
5. Select Apply filter to apply the filter to all response values in the data.

CHAPTER

4

Setting Response Value Colors for Images

Introduction to Setting Response Value Colors for Images

When you load data, the software automatically

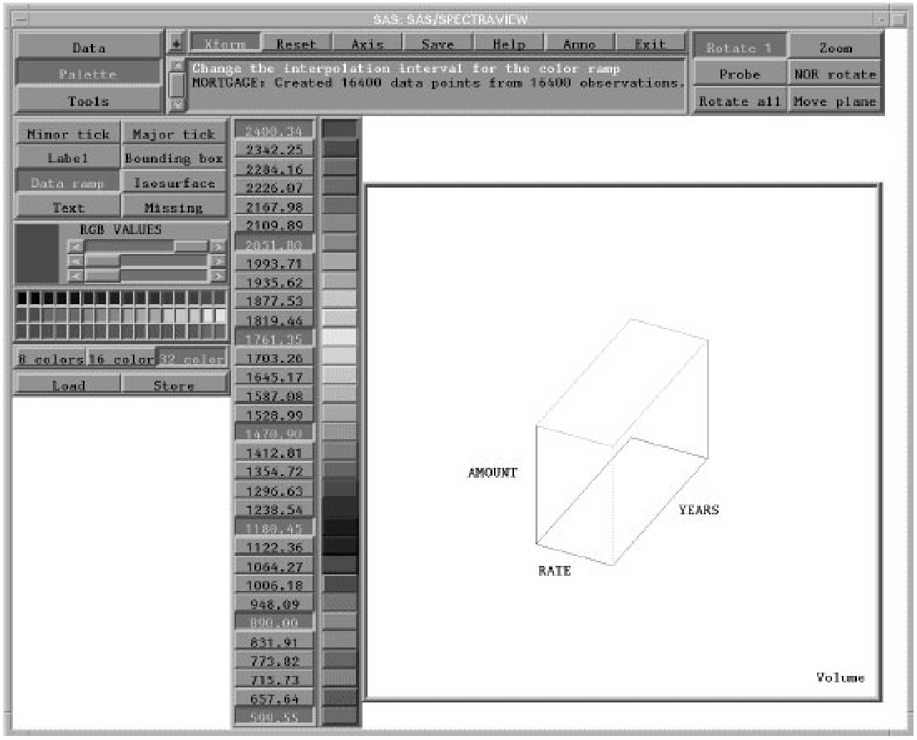
- sets default response value ranges
- color codes those ranges with preset colors.

These are the colors displayed when you create an image with a visualization technique. For example, for response values that are loan payments, the highest range of payments may be preset to red and the lowest range of payments may be preset to blue.

To meet individual needs, you can customize both the ranges and the mapped colors. You can also specify a color for missing values (which are by default transparent) so that you can see their locations in an image as well.

To set colors for an image, select Palette, which displays the software's color palette. By default, Data ramp is automatically selected, which displays the response value ranges and preset colors located below the text window as shown in Display 4.1 on page 46:

Display 4.1 Color Palette



When you are working with colors, any previously requested image remains displayed in the Volume window and will display your color specifications so that you can easily see the effects.

Note: Using the color palette, you can also specify colors for an isosurface (explained in “Viewing a Response Value as an Isosurface” on page 62) and various image customizations, such as axis labels, lines for axes, the bounding box, and text included in an image (explained in Chapter 6, “Customizing an Image,” on page 71). △

Specifying a Color with RGB Sliders and Color Chips

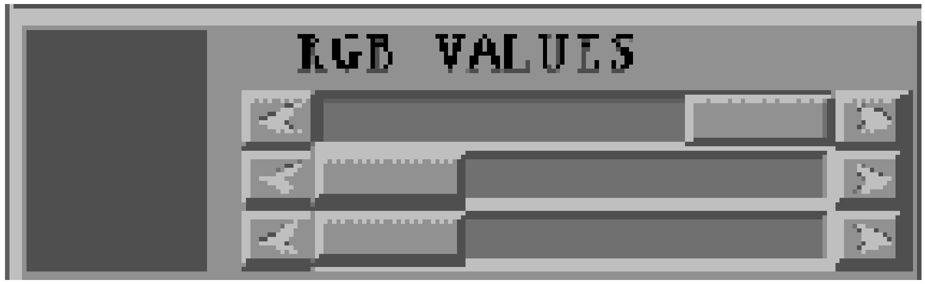
When you select Palette, the software automatically provides RGB sliders (located under the label **RGB VALUES**) and color chips (located below the RGB sliders). You use either the sliders, the chips, or both to select a color for response value ranges and for other SAS/SPECTRAVIEW items like tick marks, an isosurface, and so on.

Using the RGB Sliders

The sliders allow you to set the percentage of red, green, and blue to create a particular color. For example, the RGB values of 0, 0, 0 define black; 100, 100, 100 define white; and 0, 100, 0 define green.

The first slider sets the percentage of red, the second sets the percentage of green, and the third sets the percentage of blue. The resulting color appears in the square to the left of the RGB sliders.

Display 4.2 RGB Sliders

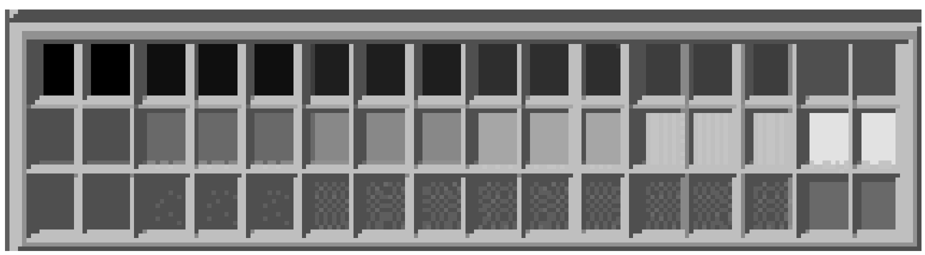


Using Color Chips

To select a color from the color chips, simply click one. The color appears in the square to the left of the RGB sliders and is reflected in the RGB slider values. Note that you can select a color chip, then refine the color with the RGB sliders.

The colors shown in the three rows of chips represent possible combinations that you can obtain by moving one of the three RGB sliders. The top row of color chips corresponds to the possible colors that can be defined by moving the top (red) slider. Similarly, the other two rows of color chips represent the effects of changing their corresponding RGB slider.

Display 4.3 Color Chips



Setting Response Value Ranges and Colors

To set response value ranges and assign specific colors to those ranges, you use the data ramp. The software then uses the specified colors to display response values in cutting planes, point clouds, the response legend, the response histogram, and so on.

For example, if the response values are sales data, you could color the low values in red, the medium values in yellow, and the high values in green, then visually analyze the data to spot trends and relationships not evident in standard reports and graphs.

Understanding the Response Value Ranges and Colors

Located below the text window, the software provides data buttons and color buttons:

- The *data buttons* display data values and allow you to interpolate between ranges of response values. You cannot change the values displayed on the data buttons. Turning a button on (selected) or off (deselected) changes how colors are interpolated between two selected buttons.
- The *color buttons* display colors associated with the data values and allow you to assign specific colors to defined ranges.

In the default data ramp, each selected data button represents the lower bound of a response value range. The range is bounded above by the value on the next selected data button (exclusive of that value itself). The software interpolates (ramps) colors between selected data buttons. For example,

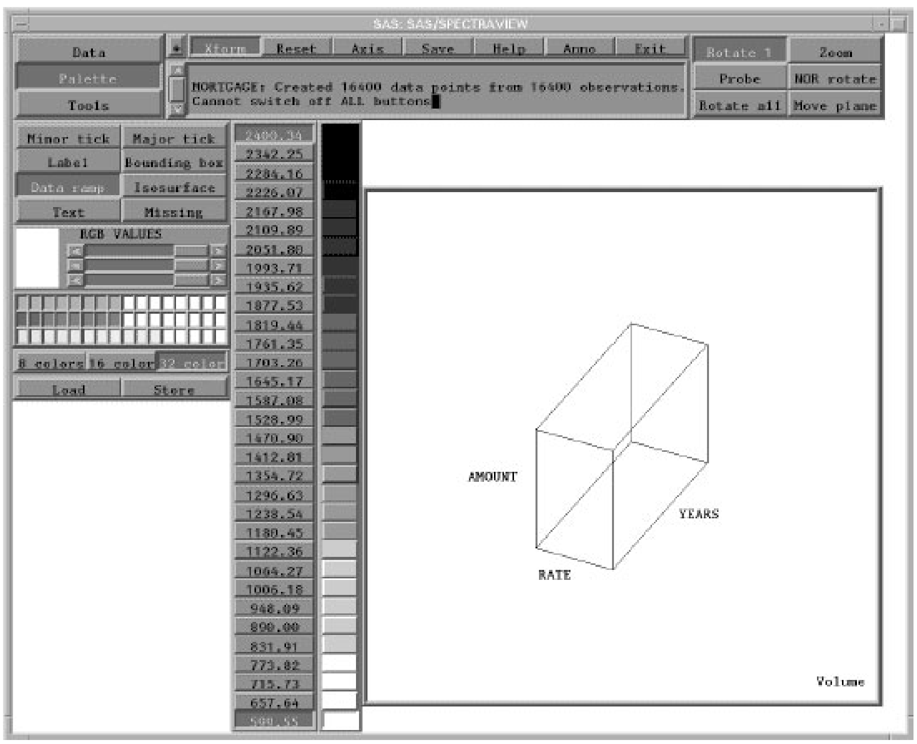
1 Deselect all data buttons except the top and bottom ones (which you cannot deselect).

2 Assign black to the top color button by clicking the top color button then using the color chips or sliders to assign 0, 0, 0 to the RGB values.

3 Assign white to the bottom color button by clicking the bottom color button then using the color chips or sliders to assign 100, 100, 100 to the RGB values.

The software uses the colors assigned to the top and bottom buttons to ramp intervening colors, as shown in Display 4.4 on page 48. All intervening values are assigned gradations between black and white, that is, a gray-scale ramp. Similarly, response values in a resulting image, such as a point cloud or a cutting plane, display in black to white gradations.

Display 4.4 Ramped Colors Between Top and Bottom Data Buttons



If you then select the color button in the middle of the top and bottom color buttons and assign it a third color (for example, yellow), the software ramps the colors between black (top) and yellow, and also between the yellow and white (bottom). All the response values falling above the middle button are displayed in colors ramped from black to yellow, and all response values falling below the middle button are displayed in colors ramped from yellow to white.

If you select multiple intervening color buttons and assign various colors, the software ramps the colors between each pair of selected buttons and displays response values within that range in the colors assigned.

Specifying Response Value Ranges and Colors

To define response value ranges and color:

1. Select [Palette].
2. Select [Data ramp], which is the default selection. The software displays RGB sliders, color chips, and the default data ramp, which has preset, evenly spaced ranges.
3. Determine the level of granularity by selecting:

[8 colors]	to divide the values in the data ramp into 8 segments and colors.
[16 colors]	to divide the values in the data ramp into 16 segments and colors.
[32 colors]	to divide the values in the data ramp into 32 segments and colors. This is the default.

   The software creates the color data ramp by determining the range of the response values in your data set, then dividing that range into 8, 16, or 32 equal parts. Even though you can modify the response value ranges so that the 32-color data ramp uses only eight colors, selecting the 8-color data ramp instead has the benefit of uncluttering line contours.
4. To define a response value range use the data ramp:
   a. Select the color button next to the data button that displays the lower bound value of the range. (When you select a color button, the data button is also selected.)
   b. Specify a color for the range by using the RGB sliders or by clicking a color chip.

   Note that to widen a range, you can deselect a data button. Once deselected, the associated color and those above and below it ramp as part of the surrounding range.
5. Repeat for each response value range.

Assigning Color to Missing Values

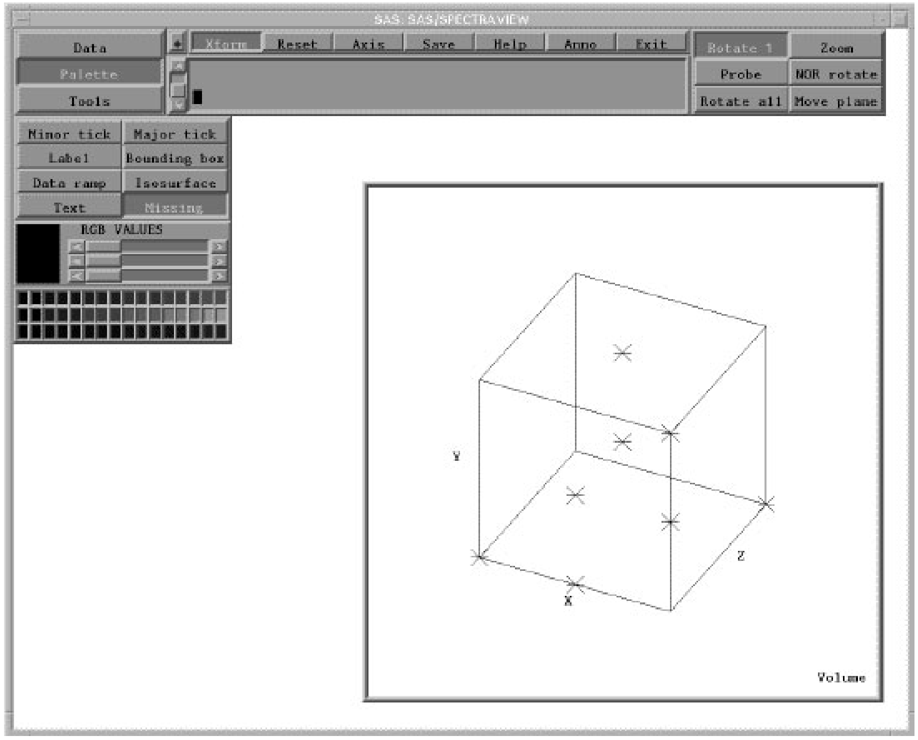
A missing value is a value in the SAS System indicating that no data is stored for the variable in the current observation. (Details about missing values in SAS/SPECTRAVIEW are explained in "Understanding Missing Values" on page 22.)

By default, missing values are transparent in an image; that is, the color assigned to missing values has RGB values set to zero. Changing the missing values to a color displays the data points in the grid without a response value.

To specify a color for missing values so that they can be displayed in an image:

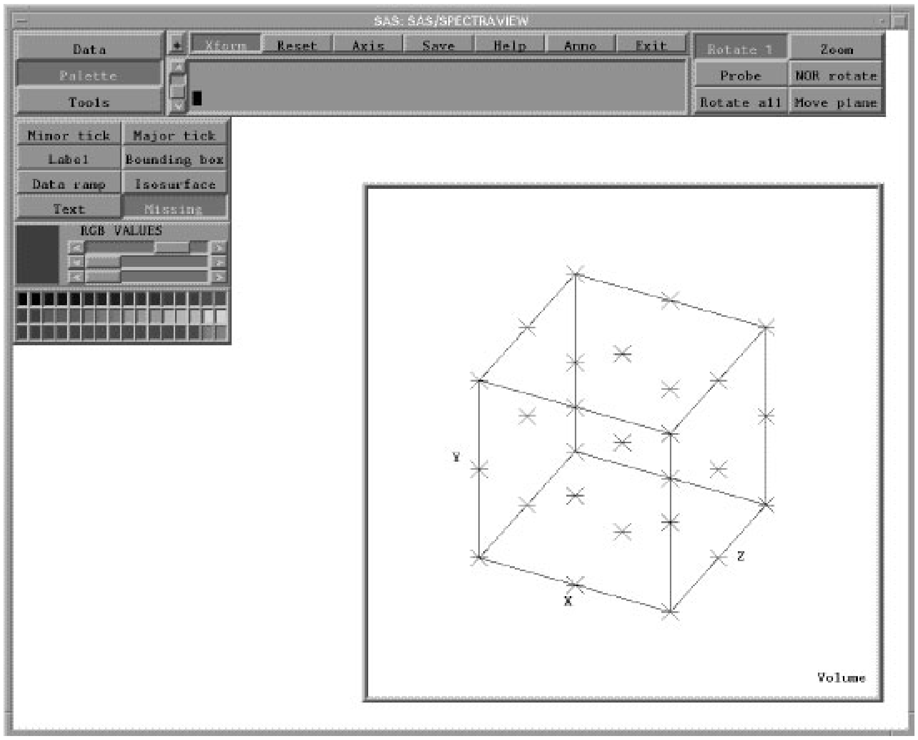
1. Select [Palette].
2. Select [Missing]. Display 4.5 on page 50 shows a point cloud with missing data points, due to the default transparent missing values.

Display 4.5 Transparent Missing Values



3 Adjust the RGB sliders. As you move the sliders, the color appears in a square at the left, and the missing values display in the Volume window. *You can also select a color chip, located below the RGB sliders.* Display 4.6 on page 51 shows the point cloud now displaying the data points that are missing.

Display 4.6 Colored Missing Values



Note that subsetting axes values with a WHERE clause when you load the data could cause otherwise valid data points to be discarded and regenerated as filler points, that is, missing.

CHAPTER

5

Exploring Data with Visualization Techniques

Introduction to Exploring Data with Visualization Techniques

Once you have loaded and perhaps color coded your data, it is time to explore and visualize the data. SAS/SPECTRAVIEW provides several interactive visualization techniques to search through your loaded data and display an image. The techniques include:

point cloud
: to display response values as colored markers, showing individual data points.

cutting planes
: to produce slices of data, either perpendicular to an axis or not perpendicular to any axis. In addition to the axis cutting planes, you can request two- and three-dimensional surface views (surfaces, charts, stacks, and plots) at a specific cutting plane's location.

isosurface
: to produce a three-dimensional surface by connecting all the data points with one response value.

solid-volume image
to produce a colored, three-dimensional solid-block image of the data points, providing an overall view of the data at the volume's border.

direct volume rendering
to create a two-dimensional image of the entire volume of data points with transparency.

BY variable processing
to animate an image so that you can see how response values change according to some grouping, like over time.

To decide which technique to use, consider the type of data. For example, to hold one axis value constant and examine the effect of the other two axis values on the response value, use a cutting plane. To view a subset of data based on response values, use a point cloud.

Note: You can combine multiple images, for example, you can display a point cloud, then a cutting plane. However, if the affect of multiple images is not what you want, be sure to turn off a technique before you request another. △

Color Coding Response Values

When you request a visualization technique, the software displays the response values using the colors from the current data ramp. You can customize the response value ranges and mapped colors as explained in Chapter 4, "Setting Response Value Colors for Images," on page 45.

Image Customizations

To enhance the visualization of data, the global buttons provide such options as image annotations, image transformations, and response value probing. For information, see Chapter 6, "Customizing an Image," on page 71.

Saving an Image

SAS/SPECTRAVIEW allows you to save a displayed image to either a TIFF (Tagged Image File Format) file or a PostScript file, for example, to use in presentations. You can also save current data values to a new SAS data set, for example, to use the new data set with other SAS System products. For information, see "Saving a Displayed Image" on page 83.

Displaying Response Values as a Point Cloud

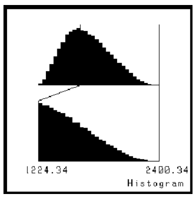
One of the first steps in analyzing data is to examine all the data values, for which the point cloud is ideal, for example, to show concentrations of data. A point cloud displays response values as colored markers, showing individual data points. The marker locations represent the plotted values of the X, Y, and Z variables, and the color of each marker represents the range in which that value resides as mapped in the data ramp.

You can display all response values in a point cloud, or you can specify a subset by using the response histogram.

Subsetting Response Values

Using the response histogram, you can specify a subset of response values for a point cloud by setting a range of responses. The response histogram, which appears at the bottom left of the interface when you request a point cloud or an isosurface, displays the frequency distribution of data points that share the same response value. The histogram may also include software-interpolated values between the actual response values.

Display 5.1 Response Histogram to Subset Point Cloud



There are actually two histograms:

- □ The upper one represents all the response values in the data set. The positions of the lines are the minimum and maximum response values.
- □ The lower one magnifies the portion of the data that lies between the two lines. The histogram stacks are larger or smaller relative to all the data between the lines. As you move the lines, you alter the span of values being compared to each other, so the stacks in the lower histogram change.

  Flat histogram stacks do not necessarily indicate no response data at that value. When you have a complete grid, the flat places may simply indicate much fewer values compared to the frequency of values represented by adjacent stacks. When you have only a few responses and they are widely separated, the values interpolated between them by the software may produce a flat stretch in the histogram.

Displaying a Point Cloud

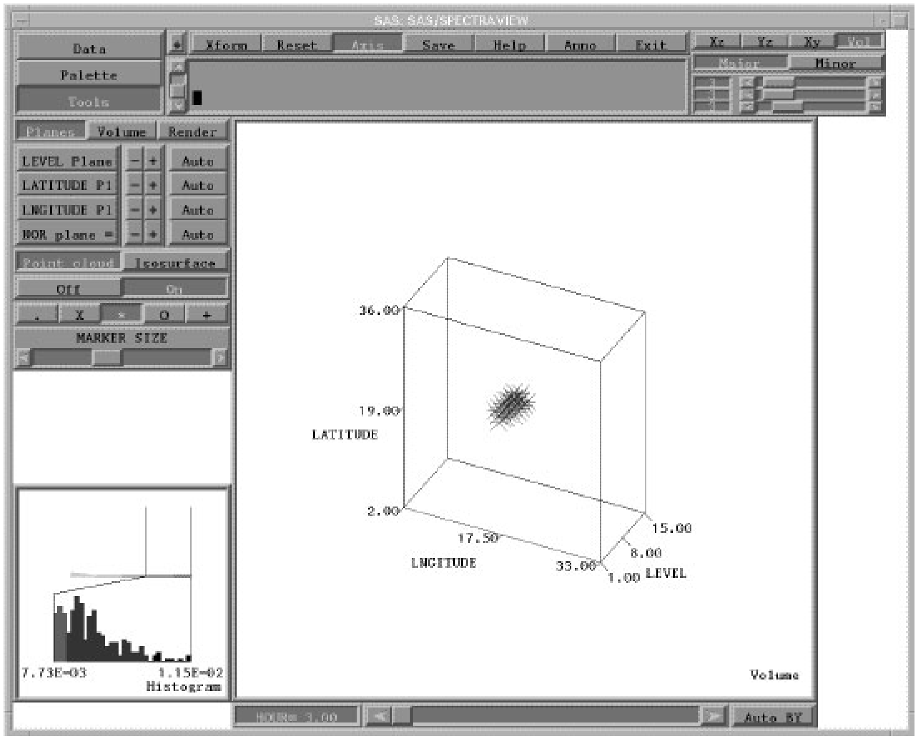
1. Select [Tools], [Planes], then [Point cloud].
2. To change the shape of the marker, select one from the panel of choices, which include a period, an X, an asterisk, a O, and a +. The default is the asterisk.
3. To change the size of the marker, move the slider below the label **MARKER SIZE**. The default is a moderate size marker. The larger the marker size, the more solid the point cloud.
4. To specify a subset of response values displayed in the point cloud, use the response histogram located in the lower left of the interface:
   - □ For a minimum response value, position the cursor on the left line, which represents the low-end values, then drag to the desired lower limit.
   - □ For a maximum response value, position the cursor on the right line, which represents the high-end values, then drag to the desired upper limit.
   - □ To pan across a subsection of the values, place the cursor inside the lines, then drag the cursor.

   Note that you can modify subsetting criteria with a point cloud displayed, but for a lot of data, you may want to turn off the point cloud to speed up the process.

5 Select [On].

The following point cloud displays a subset of response values from the EPA data set:

Display 5.2 Point Cloud Displaying Location of High Sulfate Concentrations



Creating Cut-Away Images with Cutting Planes

Cutting planes provide slices of your data...slicing the data as if with a knife. With cutting planes, you can investigate response values visually slice by slice, as you increment, decrement, or randomly select a position. The face of each slice is colored according to the response values. SAS/SPECTRAVIEW provides:

- three axis cutting planes that cut perpendicular to the X, Y, or Z axis, which display in the Volume window. In addition, you can display two- and three-dimensional surface views of the data (such as surfaces, charts, and plots), showing the values at each location along the specified axis, which display in the surface windows.
- one NOR cutting plane (non-orthogonal), meaning that the plane is not perpendicular to any axis, which displays in the Volume window.

Displaying an Axis Cutting Plane

The three axis cutting planes cut perpendicular to the X, Y, or Z axis. For example, the Y cutting plane lies parallel to the XZ plane and perpendicular to the Y axis. An axis cutting plane can be displayed in either block contour format (solid-block) or line contour format (continuous lines). To display an axis cutting plane in the Volume window:

1 Select [Tools], then [Planes].

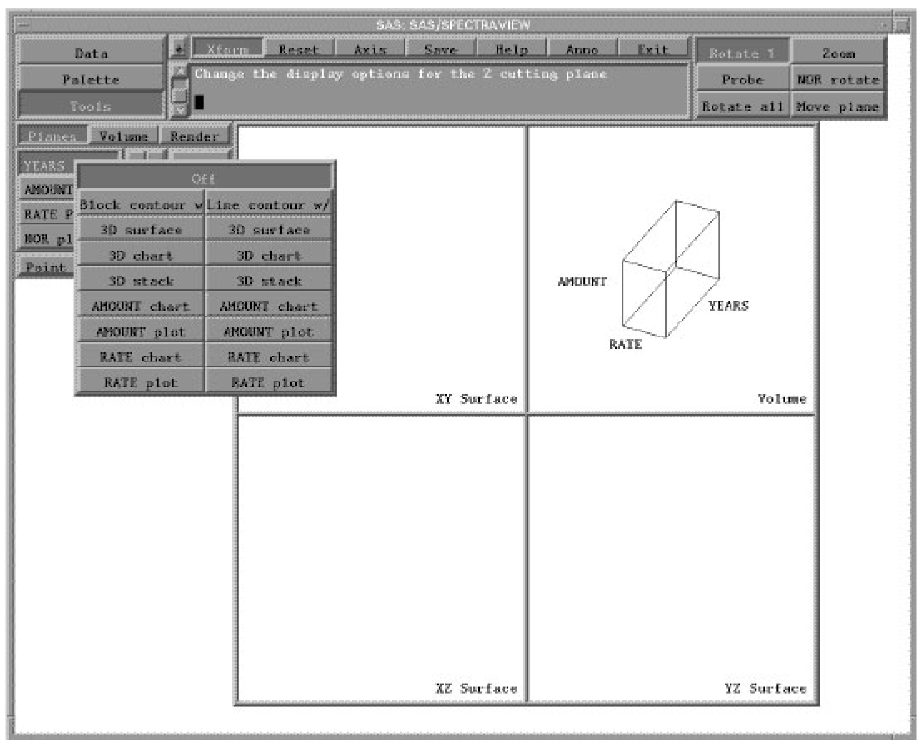
The three axis cutting planes are labeled with the variable names you specified for the X, Y, and Z axes. For example, using the MORTGAGE data set and specifying RATE as the X variable, AMOUNT as the Y variable, and YEARS as the Z variable, results in the axis cutting planes:

- YEARS Plane =>
- AMOUNT Plane =>
- RATE Plane =>

The labels for the axis cutting planes also contain a **=>** symbol indicating that the button is a *pull-down button*, which means that pressing the button will display a menu of choices.

2 Position the cursor on the pull-down button representing the axis cutting plane you want to display, then press and hold down the left mouse button to display a menu of choices. *The additional buttons are displayed only while you press the mouse button.*

Display 5.3 Pull-Down Buttons for Axis Cutting Planes



The top choices represent how the cutting plane will be displayed in the Volume window (as solid blocks or line contours), and the bottom choices control which type of image appears in the surface window.

3 Drag the mouse to one of the following choices, then release the mouse to select it:

Block contour w/
: to display the cutting plane in a solid-block format. Note that for sparse data, a block-contour image may appear as separate blocks, because each block represents one data point in the current plane.

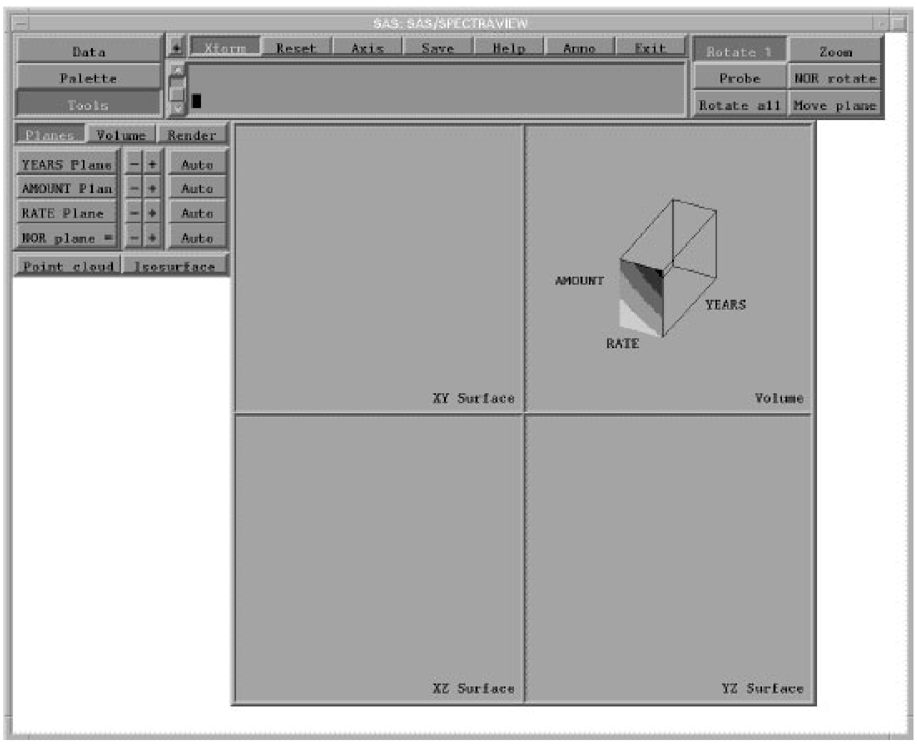
Line contour w/
: to display the cutting plane as continuous lines. Note that a line-contour image needs at least four data points at each grid intersection to be displayed, which may not exist for sparse data.

Off
  to turn off the displayed cutting plane.

You can display more than one cutting plane at a time in the Volume window, up to the four types. *If there is no data for the cutting plane's current location, the cutting plane will not display, unless you have assigned color to missing values.*

In the following example, which uses the MORTGAGE data set, the response values (for PAYMENT) are displayed against the Z cutting plane by selecting YEARS Plane=> then Block contour w/. The cutting plane shows the impact of varying interest rate and loan amounts for a given year. That is, the year is held constant and the impact of the other two variables on the payment amount is shown. You can move the cutting plane over the range of years to examine other combinations based on a given year value.

Display 5.4 Y-Axis Cutting Plane



Including Surface Views for an Axis Cutting Plane

You can display two- and three-dimensional surface views of data, such as surfaces, charts, stacks, and plots. The surface views represent the data at a selected cutting plane's current location, and they display in an appropriate surface window. To request a surface view:

1 Select Tools, then Planes.

The three axis cutting planes are labeled with the variable names you specified for the X, Y, and Z axes. For example, for the MORTGAGE data set, RATE is the X variable, AMOUNT is the Y variable, and YEARS is the Z variable, which results in the axis cutting planes

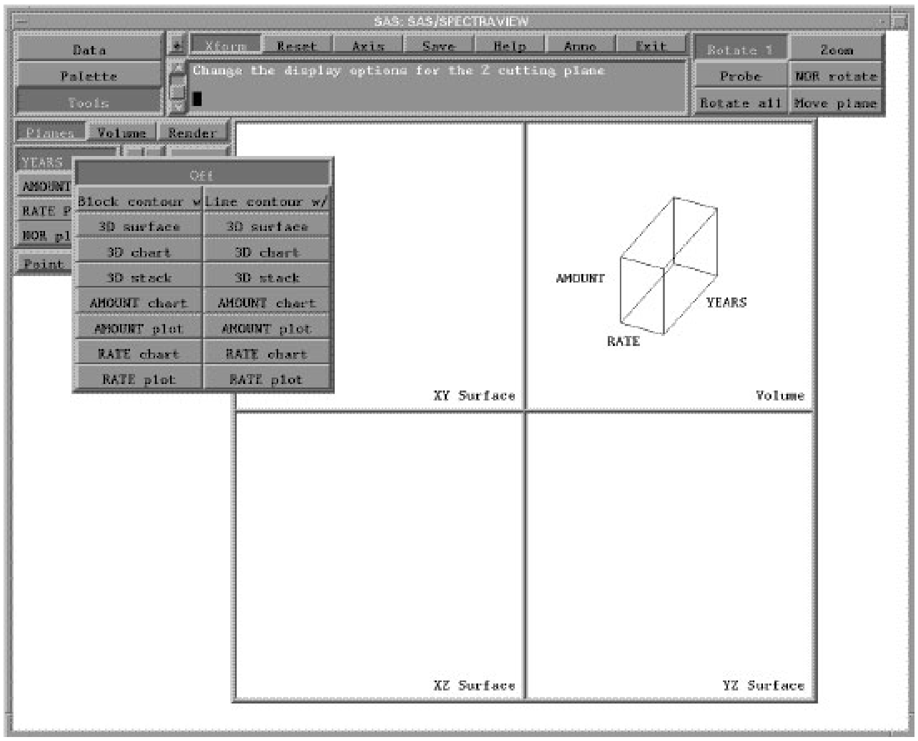
- YEARS Plane =>

- AMOUNT Plane =>
- RATE Plane =>

The labels for the axis cutting planes also contain a **=>** symbol indicating that the button is a *pull-down button*, which means that pressing it will display a menu of choices.

2 Position the cursor on the pull-down button representing the axis cutting plane you want to display, then press and hold down the left mouse button to display a menu of choices. *The additional buttons are displayed only while you press the mouse button.*

Display 5.5 Pull-Down Buttons for Surface Views



The top choices represent how the cutting plane will be displayed in the Volume window (as solid blocks or line contours), and the bottom choices control what type of image appears in the surface window.

3 Drag the mouse to one of the following choices, then release the mouse to select it:

3D surface	to create a three-dimensional surface of the axis cutting plane. The third dimension is response.
3D chart	to create a three-dimensional chart of the axis cutting plane. The third dimension is response.
3D stack	to create a three-dimensional stacked line contour of the axis cutting plane. The third dimension is response.

For the respective axes:

axis chart	to create a two-dimensional bar chart of the responses across all values of the axis. For example, for the X cutting plane, selecting the Z chart produces a color chart of the responses

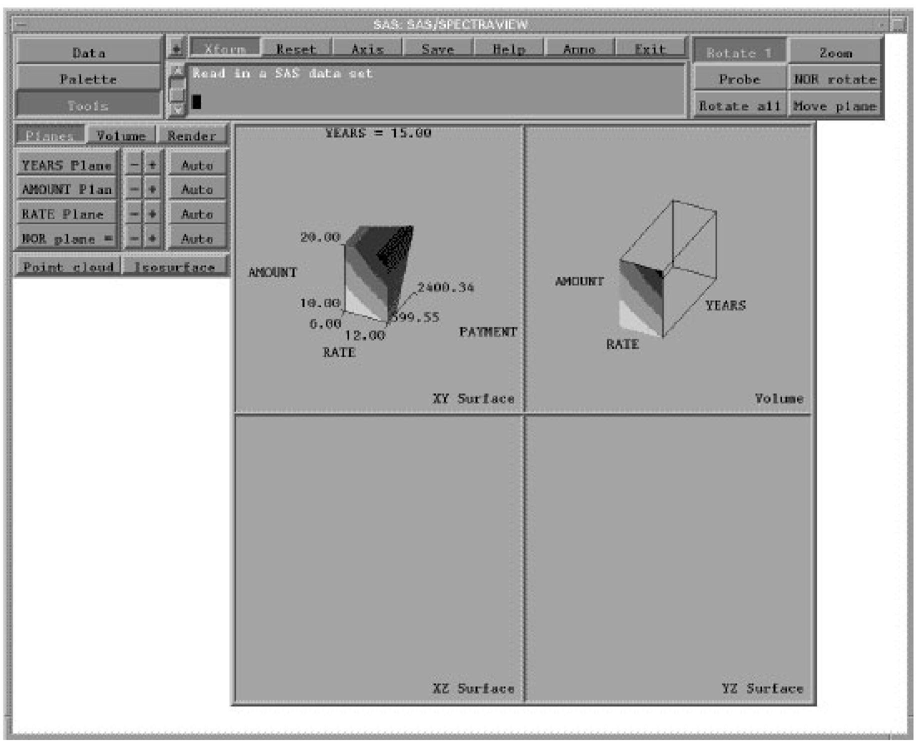
across all values of Z at the current X location of the cutting plane. The Y value remains constant at its current value.

axis plot — to create a two-dimensional plot of the response values across all values of the axis. For example, for the X cutting plane, selecting the Z plot produces a color plot of the responses across all values of Z at the current X location of the cutting plane. The Y value remains constant at its current value.

The image displays in the appropriate surface window in solid-block format. That is, a surface view of a cutting plane along the Z axis displays in the XY Surface window, a cutting plane along the Y axis displays in the XZ Surface window, and a cutting plane along the X axis displays in the YZ Surface window. *Note that when you request a surface view, if its associated axis cutting plane is not already displayed, it automatically displays in the Volume window.*

In the following example, a three-dimensional chart is displayed in the XY Surface window by selecting YEARS Plane=> (Z cutting plane), then 3D chart. The chart is linked to the Z cutting plane's current position.

Display 5.6 Three-Dimensional Chart in XY Surface Window



Displaying a NOR Cutting Plane

The NOR cutting plane (non-orthogonal) is not perpendicular to any axis. You can use the NOR cutting plane to cut the data in any manner you wish. It displays in block contour only. To display the NOR cutting plane in the Volume window:

1 Select Tools, then Planes.

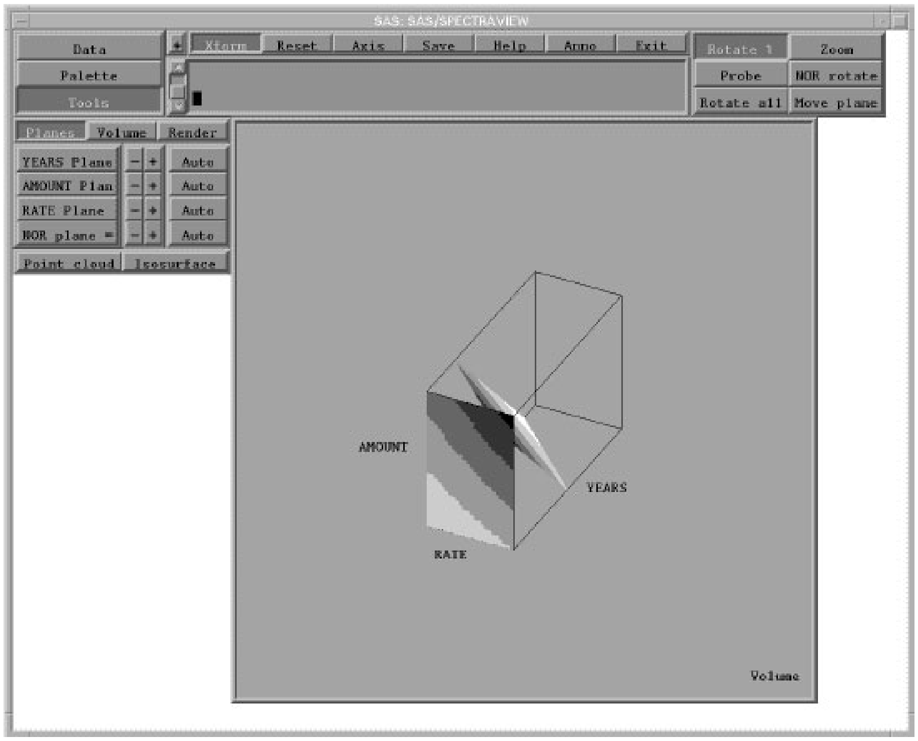
The labels for the cutting plane buttons contain a **=>** symbol indicating that the button is a *pull-down button*, which means that pressing it will display a menu of choices.

2 Position the cursor on the NOR plane => pull-down button, then press and hold down the left mouse button to display a menu of choices. *The additional buttons are displayed only while you press the mouse button.*

3 Drag the mouse and select Block contour to display the cutting plane in a solid-block format. (Selecting Off turns off the displayed cutting plane.)

In the following example, using the MORTGAGE data set, the NOR cutting plane is displayed along with the Y cutting plane.

Display 5.7 NOR Cutting Plane



Moving a Cutting Plane

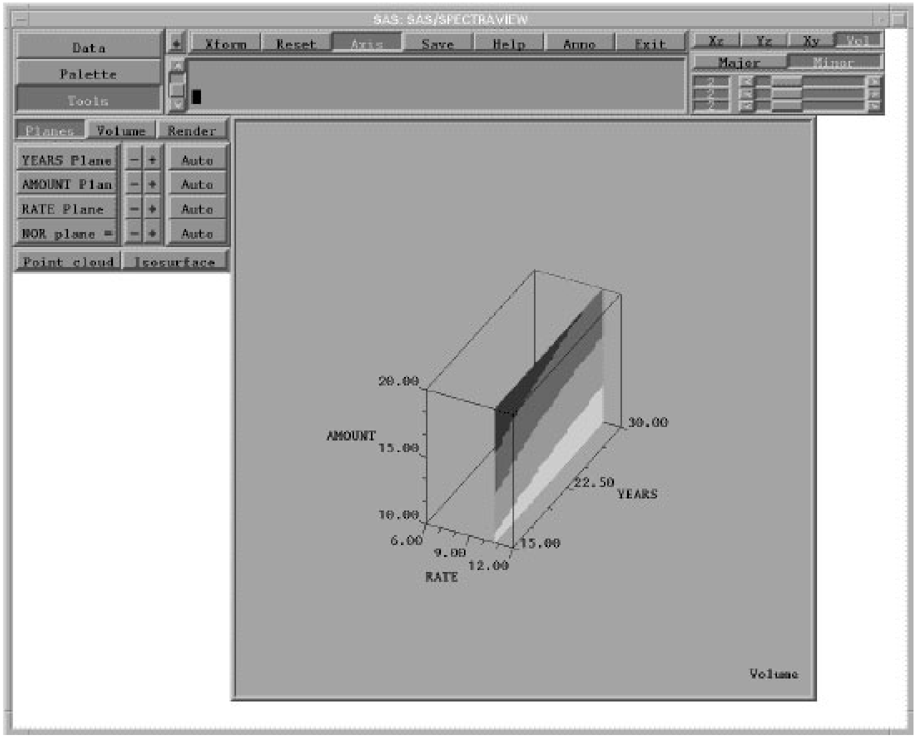
The initial position of a cutting plane shows a section at the minimum value for the axis. Moving a cutting plane lets you view the data slice by slice, allowing you to position a cutting plane at a specific location along the axis.

As you move a cutting plane, the data at the various locations along the axis displays against the cutting plane and in any two- or three-dimensional surface views. *If there is no data for the cutting plane's current location, the cutting plane will not display, unless you have assigned color to missing values.*

- To automatically move a displayed cutting plane continuously along its axis, select Auto . The response values display against the cutting plane as it moves through the volume grid, stopping at each unique value for the axis. Only one plane at a time can move continuously. (To make the plane stop moving, click the left mouse button.)
- To manually move a cutting plane to a specific point on the axis by stepping it through the data incrementally, use the − (backward) and + (forward) buttons.

  Or, select the Xform global button, select either Move plane (for the axis cutting planes) or NOR rotate (for the NOR cutting plane), drag the cursor along the axis to the appropriate point, and release the mouse button.

Display 5.8 Moved Cutting Plane



Viewing a Response Value as an Isosurface

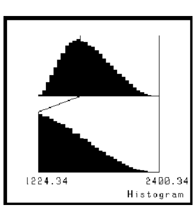
An isosurface produces a three-dimensional surface by connecting all the data points with one response value. That is, the isosurface lets you see all locations in the volume grid when the response values are equal. An isosurface is particularly useful in viewing scientific data and physical properties such as temperature, density, and stress.

You must specify the response value to be displayed by the isosurface using the response histogram. The value shown corresponds to the lower histogram level.

Isolating a Response Value

With the response histogram, you specify the response value for an isosurface. The response histogram, which appears at the bottom left of the interface when you request a point cloud or an isosurface, displays the frequency distribution of data points that share the same response value. The histogram may also include software-interpolated values between the actual response values, which can be used to generate an isosurface.

Display 5.9 Response Histogram to Specify Isosurface Value



There are actually two histograms:

- The upper one represents all the response values in the data set. The positions of the lines are the minimum and maximum response values.
- The lower one magnifies the portion of the data that lies between the two lines. The histogram stacks are larger or smaller relative to all the data between the lines. As you move the lines, you alter the span of values being compared to each other, so the stacks in the lower histogram change.

Flat histogram stacks do not necessarily indicate an absence of response data at that value. When you have a complete grid, the flat places may simply indicate much fewer values compared to the frequency of values represented by adjacent stacks. When you have only a few responses and they are widely separated, the values interpolated between them by the software may produce a flat stretch in the histogram.

Displaying an Isosurface

1 Select Tools, Planes, then Isosurface.

2 To specify the response value, use the left mouse button and drag the histogram's left line to the desired value. The default position of the left line is the minimum response value, which cannot be used to generate an isosurface. (The right line has no effect on an isosurface specification.)

3 Render the isosurface by selecting:

W/ depth sort
: displays the isosurface with depth sorting by sorting the planes front-to-back so that the image appears more realistically three-dimensional, for example, when moving planes through the volume grid. This method takes more time.

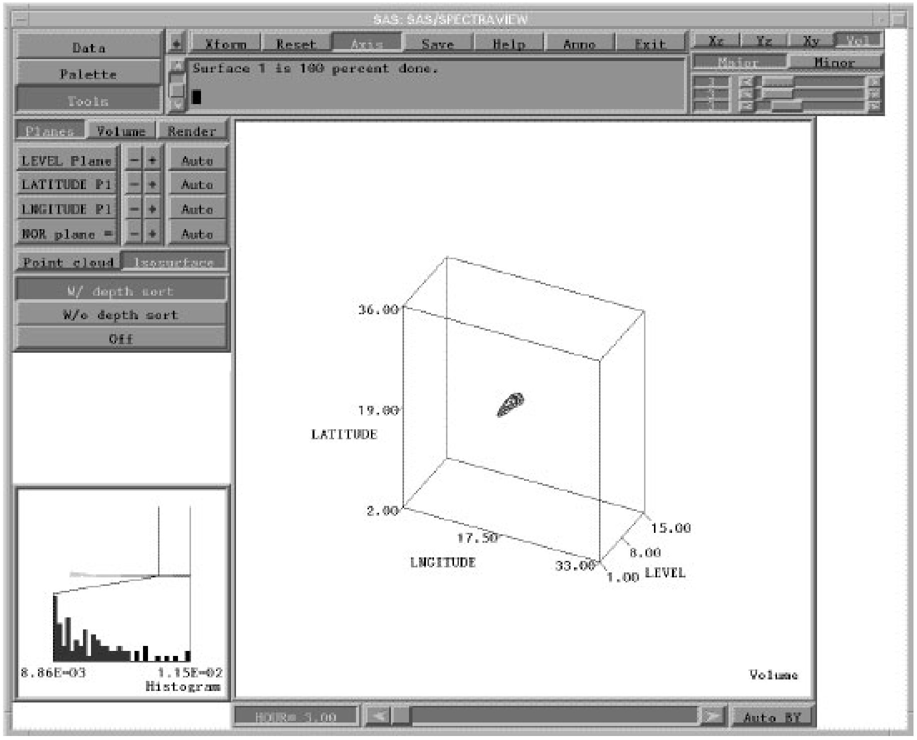
W/o depth sort
: displays the isosurface without depth sorting, which is a faster, but less realistic method. *It is recommended that you use this method when animating an isosurface with a BY variable to produce faster animation.*

Off
: turns off the displayed isosurface.

By default, the isosurface is drawn in the color shown in the histogram for the selected response value.

4 To assign a different color to the isosurface, select Palette, Isosurface, then User-defined. Adjust the RGB sliders. As you move the sliders, the current color appears in a square at the left, and the isosurface changes color in the Volume window. *You can also select a color chip, located below the RGB sliders.* (To return to the response value color, select Response level.)

Display 5.10 Isosurface



Note that you can specify a different response value when the isosurface is turned on or when it is turned off. However, if you have a large amount of response data, it may be faster to turn off the isosurface, then adjust the histogram.

Effect of a BY Variable on an Isosurface

For an isosurface with a BY variable specification, you can request a specific response value but not see the isosurface display. This is because the response value you selected is not contained in the data associated with the current BY value. (The histogram shows all responses, in ascending order, regardless of the value of the BY variable.)

To see the isosurface that was generated, use [Auto BY] to search for the appropriate BY value. Because the response value may be valid for more than one x,y,z location over time, an isosurface may be generated for more than one BY value. When the applicable BY value(s) is current, the isosurface is clearly visible.

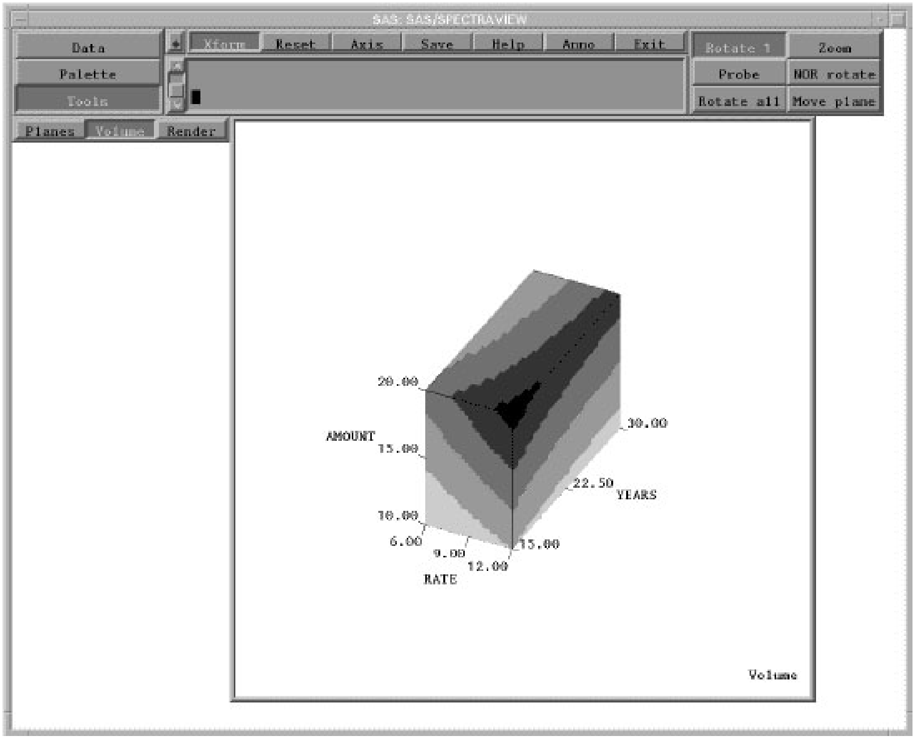
Displaying Data as a Solid-Volume Image

Displaying data as a solid-volume image produces a colored, three-dimensional solid-block image of the data points, providing an overall view of the data. The technique displays the response values at the bounds of the volume grid. *Note that to view interior data, you can specify a subset of the volume.* To display a solid-volume image:

1 Select [Tools].

2 Select [Volume]. The image displays in the Volume window.

Display 5.11 Solid-Volume Image



Note that from the Xform global button, you can select the Rotate 1 button to change the angle of the image by dragging the cursor along any of the three axes or the Move plane button to increase or decrease the area of volume by dragging the cursor along any of the three axes.

Rendering a Volume and Saving as a Two-Dimensional Image

Direct-volume rendering creates a two-dimensional image of the entire volume of data points with transparency displayed in the Volume window, which you can then save to a file.

You can render an image using either the Splat option or the Scanline option:

Splat — renders each response value using a technique that is much like splatting a snowball against the cutting plane as it moves through the volume grid.

Scanline — renders the image or images displayed in the Volume window as a high-quality, three-dimensional image.

Note that before rendering, consider changing the size of the Volume window so that the resulting image is larger or smaller. For example, for many data points, reducing the window size improves rendering time, whereas increasing the window size can provide a better view of dense data. For instructions, see "Resizing the Display Windows" on page 84.

Requesting Splat Rendering

Specifying the Splat option renders each response value using a technique that is much like splatting snowballs against the cutting plane as it moves through the volume grid...each data point is a snowball. You have the option of setting an opacity degree and a splat width.

To create a splat rendered image:

1 Select Tools, then Render.

2 In the Opacity window, which is located at the bottom of the interface, adjust the opacity line to specify a degree of opacity or transparency for response values with the various color ranges. That is, visibility of data points in the rendered image is determined by the location of the opacity line corresponding to the colors of the data points.

   To adjust the opacity degree, use the cursor to drag portions of the opacity line to the top or bottom of the window:

   - Colors represented on the line at the top of the window are rendered with a high opacity, making the colored data points more opaque (clearly visible in the rendered image). To make a specific range of responses more visible in the rendered image, drag those colors to the top of the window.
   - Colors represented on the line at the bottom of the window are rendered with a low opacity, making the colored data points more transparent (invisible in the rendered image). To exclude responses from the rendered image, drag those colors to the bottom of the window.

   Note that you can reset the opacity line to its default degrees by selecting the Reset global button, then the All button. However, the All button resets the entire software to its defaults, so use it cautiously.

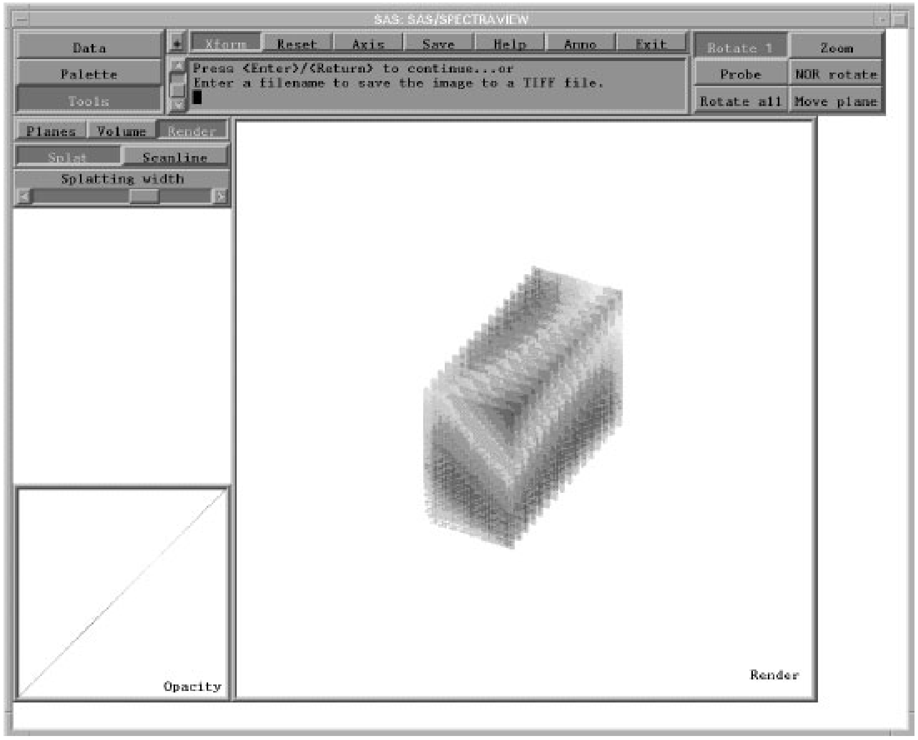
3 To set the size of the rendered data points, which determines how solid the rendered image is to be, adjust the splat width by dragging the slider to the right for a larger width or to the left for a smaller width. *The larger the splat width, the more solid the resulting image.*

4 Select Splat to initiate the rendering process.

5 Once the image is rendered, you are prompted to enter a filename in the text window. *The file type depends on the specification from the Save global button.* The default is a TIFF file.

6 To bypass the filename prompt or to clear the rendered image, press Enter.

Display 5.12 Splat Rendered Image



Note that when you request splat rendering, the software renders all response values even if you have a subset of response values displayed in a point cloud.

Requesting Scanline Rendering

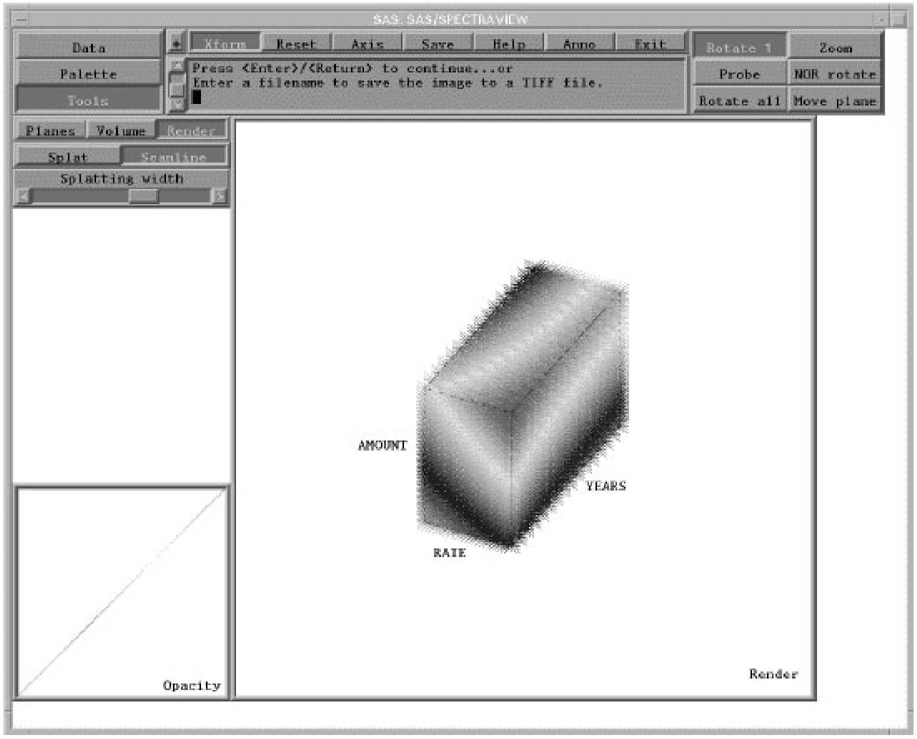
Specifying the scanline option renders the image displayed in the Volume window as a shaded, solid object, showing the effects of a light source. An isosurface and cutting planes are rendered as solid, Gouraud-shaded images.

For example, if you request an isosurface, rendering it using scanline produces a solid image from the wireframe one, which will show more clearly how the surface changes. In addition, a scanline rendering of a cutting plane is smoother.

To create a scanline rendered image:

1. Use a visualization technique to display an image or images in the Volume window, for example a point cloud.
2. Select Tools, then Render.
3. Select Scanline to initiate the rendering process.
4. Once the image is rendered, you are prompted to enter a filename in the text window. *The file type depends on the specification from the Save global button, with the default being a TIFF file.*
5. To bypass the filename prompt or to clear the rendered image, press Enter.

Display 5.13 Scanline Rendered Image



Animating an Image with a BY Variable

With a BY variable assigned, you can animate an image in the Volume window. The values of a BY variable define groups of observations, such as hour, month, or year. Specifying a BY variable lets you see how response values change according to some grouping, like over time.

For example, if the specified BY variable is the hour of the day and a point cloud is displayed showing the amount of sulphur in the air at various coordinate locations, the point cloud reflects the amount of sulphur for the locations at 9 o'clock, 10 o'clock, and so on.

To produce animation with a BY variable:

1 First, you must specify a BY variable when you load the data. For instructions, see "Grouping Observations with a BY Variable" on page 26.

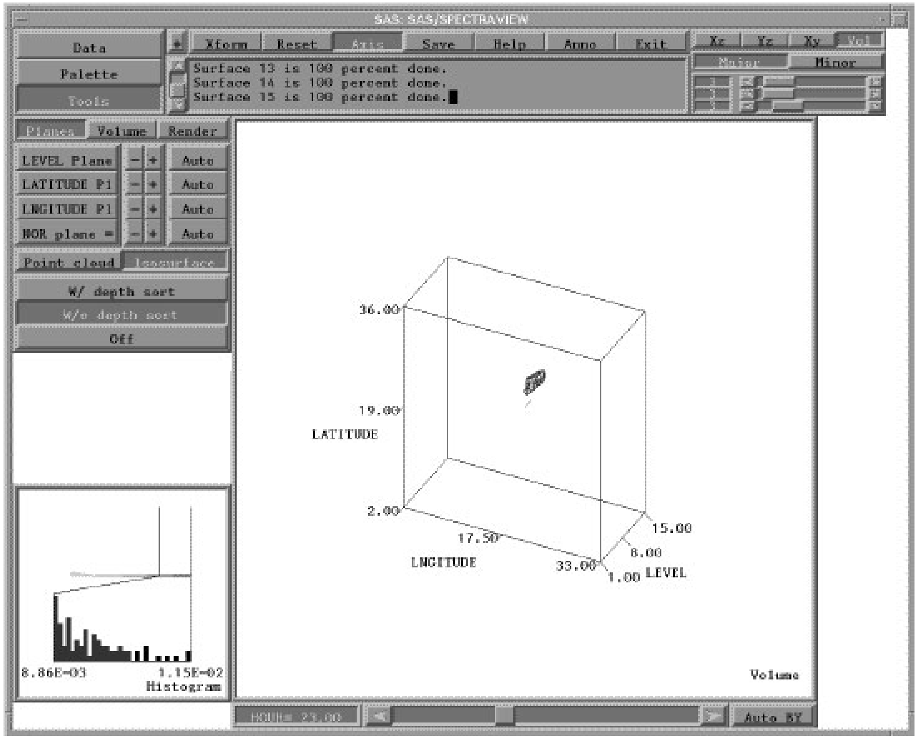
2 Use a visualization technique to display an image, such as a cutting plane, a point cloud, or an isosurface.

3 To automatically move through the BY variable values, which animates the current image, select Auto BY, which is located at the bottom right of the interface. The values for the BY variable appear at the left of the BY variable slider, changing as the software moves through the values. (To stop the animation, click the left mouse button.)

   To manually move through the BY variable values, for example, to see the image at a specific value, drag the BY variable slider, which appears below the display windows, to the left of Auto BY.

For example, using the EPA data set and specifying HOUR as the BY variable, the following isosurface shows high sulfate concentrations at hour 23.

Display 5.14 Specified BY Variable



CHAPTER

6

Customizing an Image

Introduction to Customizing an Image

In conjunction with visualization techniques, you can customize data exploration and analysis with the following options:

- annotating an image with axis labels, text, a response legend, and so on
- transforming an image by rotating and zooming it
- probing an image, which displays the response value for a specific data point in an image.

Saving an Image

SAS/SPECTRAVIEW allows you to save a displayed image to either a TIFF (Tagged Image File Format) file or a PostScript file, for example, to use in presentations. You can also save current data values to a new SAS data set, for example, to use the new data set with other SAS products. For information, see “Saving a Displayed Image” on page 83.

Annotating an Image

SAS/SPECTRAVIEW provides the following types of annotations that you can include for an image in the Volume window and the surface windows.

Setting Axis Labels and Tick Marks in the Volume Window

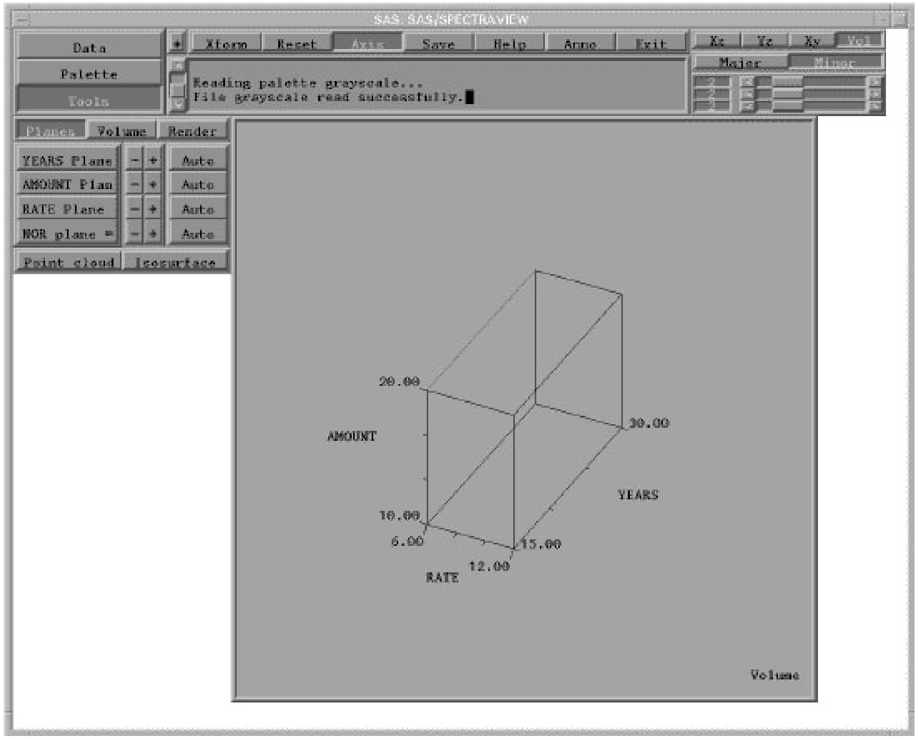
In the Volume window, you can display axis labels, which are the variable names associated with the X, Y, and Z axes, and major and minor tick marks. To display or hide these labels in the Volume window:

1. Select the [Axis] global button, then [Vol].
2. Select [Major] to set major tick marks or [Minor] to set minor tick marks. By default, major tick marks are set to 1, which displays variable names but no tick marks. *To display minor tick marks, major tick marks must be displayed.*
3. Set the tick marks by dragging the three sliders to specify the number of tick marks for the X, Y, and Z axes. (The value you specify appears to the left of the slider, and the tick marks appear on the appropriate axis in the Volume window.) *To hide variable names, set the slider to 0. To include variable names plus major tick marks, move the slider to a number greater than 1; for example, 2 displays variable names with two tick marks along the axis.*

   The maximum value for each major tick slider equals the number of unique data values for the respective axis. If you categorized the data, the maximum value equals the number of categories specified for that axis. The maximum value for the minor tick sliders is 10.
4. To specify where the variable name and tick marks for an axis are to appear, position the hand cursor in the Volume window so that it points to the axis line you want, and click the left mouse button. For example, if you have rotated an image and the variable name and tick marks for the X axis are at the back of the image, you can click the top front or bottom front X axis lines of the bounding box; the labels appear along the line you clicked.
5. The default color for both variable names and tick marks is black. To specify a different color, select [Palette], then either [Minor tick] or [Major tick] to assign a color to the tick marks or [Label] to assign a color to the variable names.

   Adjust the RGB sliders. As you move the sliders, the specified color appears in the square at the left, and the item changes color in the Volume window. *You can also select a color chip, located below the RGB sliders.*

Display 6.1 Volume Window with Axis Labels



Note that you can define axis labels and tick marks regardless of whether the bounding box is on or off.

Setting Axis Labels in the Surface Windows

In the surface windows, you can display axis labels, which are the variable names associated with the X, Y, and Z axes, and the minimum and maximum values. To display or hide labels in the surface windows:

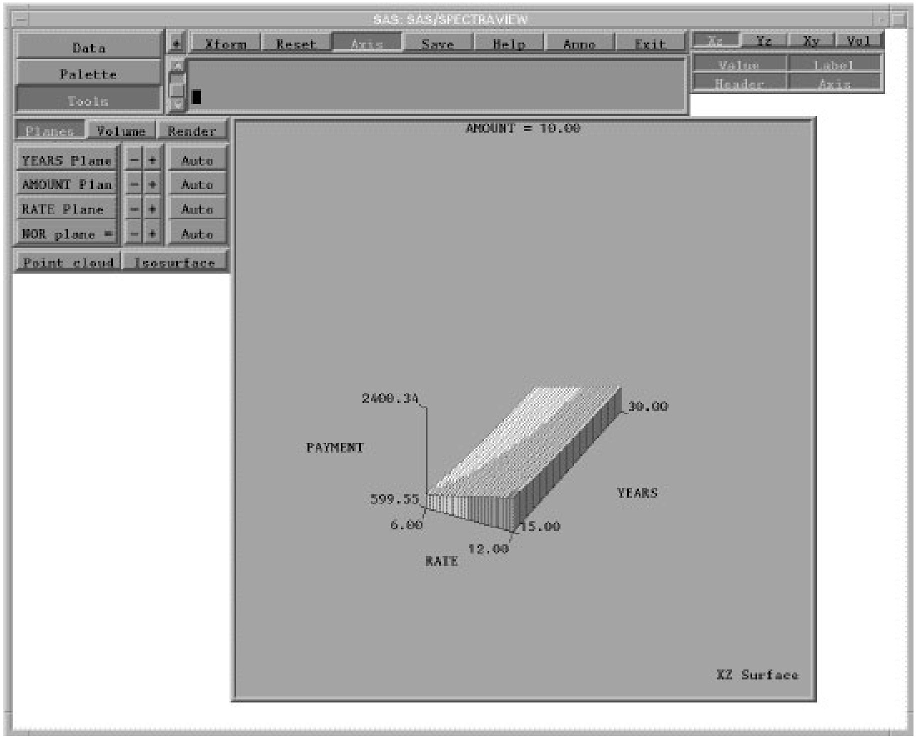
1 Select the Axis global button, then either Xz, Yz, or Xy to specify the surface window.

2 By default variable names and values are displayed. To hide variable names, deselect Label. To hide values, deselect Value.

3 The default color for both variable names and values is black. To specify a different color, select Palette, then Label.

   Adjust the RGB sliders. As you move the sliders, the specified color appears in the square at the left, and the item changes color in the Volume window. *You can also select a color chip, located below the RGB sliders.*

Display 6.2 Surface Window with Axis Labels



Note that you can define axis labels and values for surface window images regardless of whether the axes lines are on or off.

Setting the Bounding Box in the Volume Window

By default in the Volume window, the software displays a *bounding box*, which is a set of lines that outline the volume grid representing the loaded data. To include or hide the bounding box:

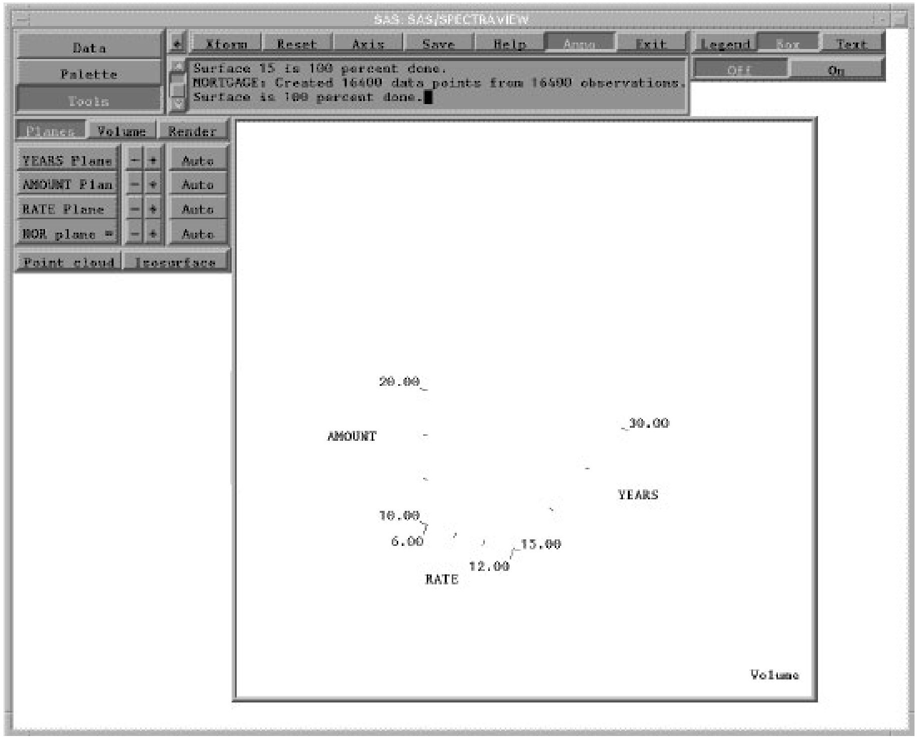
1 Select the Anno global button, then Box.

2 To turn off the bounding box, select Off. (To display the bounding box, select On.)

3 By default, the bounding box is multicolored. To specify one color, select Palette, then Bounding box.

4 Select User-defined to specify a color for the bounding box. Adjust the RGB sliders. As you move the sliders, the specified color appears in the square at the left, and the item changes color in the Volume window. *You can also select a color chip, located below the RGB sliders.* (To return to the default multiple colors, select Default.)

Display 6.3 Volume Window without Bounding Box



Setting Axes Lines in the Surface Windows

When you display a two- or three-dimensional image in a surface window, the software by default provides axes lines. To set axes lines in the surface windows:

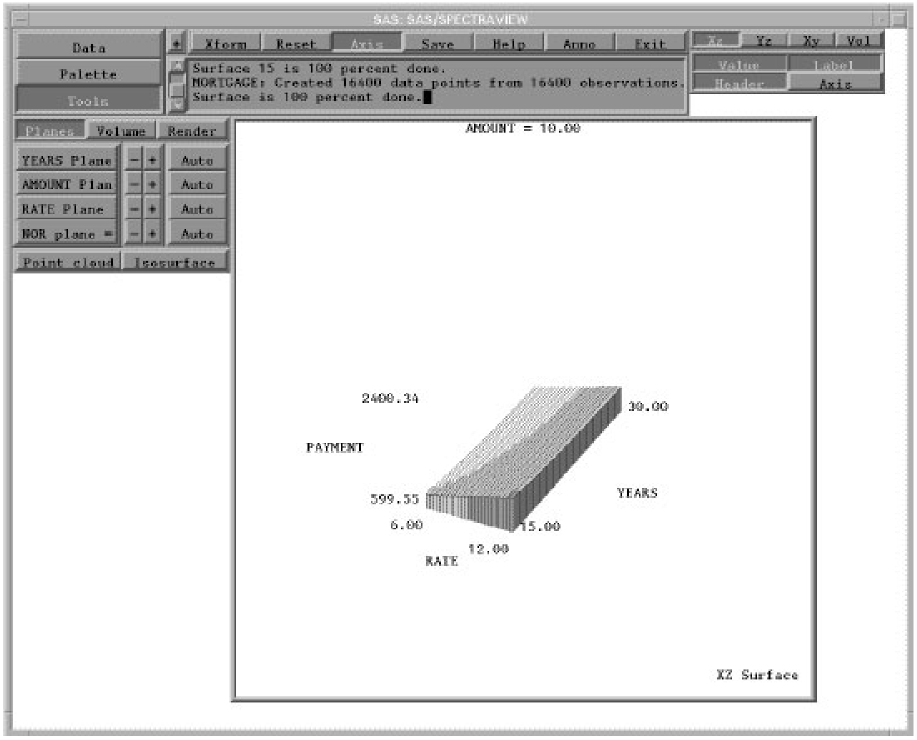
1 Select the [Axis] global button, then either [Xz], [Yz], or [Xy] to specify the surface window.

2 By default axes lines are displayed. To hide them, deselect [Axis] from the menu of buttons, which displayed when you selected the [Axis] global button.

3 The default color for axes lines is black. To specify a different color, select [Palette], then [Bounding box].

4 Select [User-defined] to specify a color for the lines. Adjust the RGB sliders. As you move the sliders, the specified color appears in the square at the left, and the item changes color in the Volume window. *You can also select a color chip, located below the RGB sliders.* (To return to the default, select [Default].)

Display 6.4 Surface Window without Axes Lines



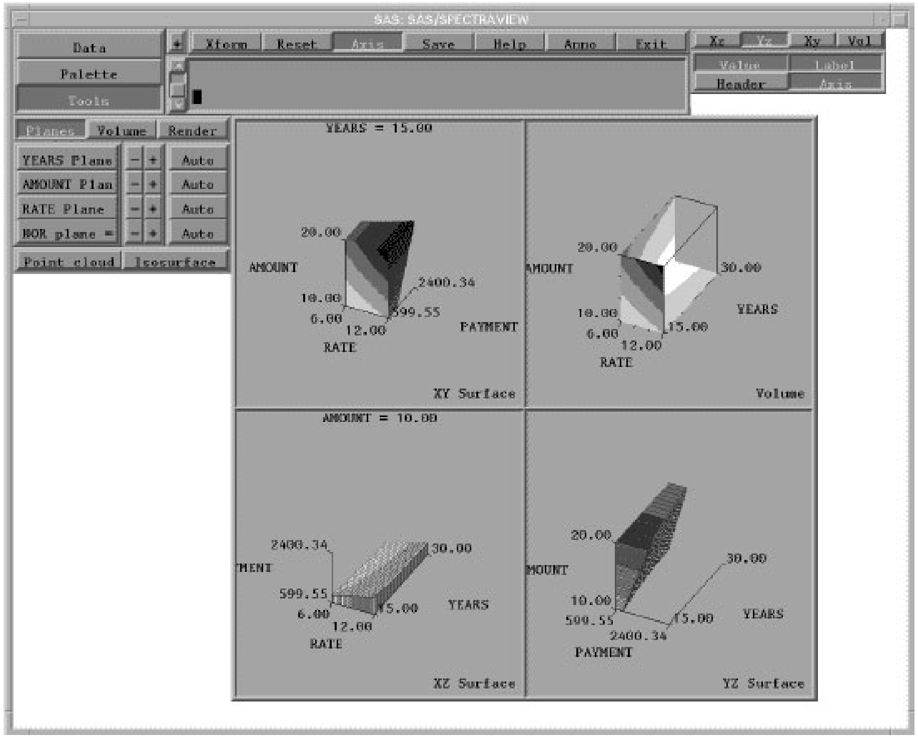
Setting Headers in the Surface Windows

When you display an axis cutting plane view in a surface window, the software by default provides a window header that specifies the variable name for the axis and the value for the cutting plane's location. To display or hide headers in the surface windows:

1. Select the Axis global button, then either Xz, Yz, or Xy to specify the surface window.
2. By default headers are displayed. To turn off headers, deselect Header.
3. The default color for headers is black. To specify a different color, select Palette, then Label.

   Adjust the RGB sliders. As you move the sliders, the specified color appears in the square at the left, and the item changes color in the Volume window. *You can also select a color chip, located below the RGB sliders.*

Display 6.5 YZ Surface Window with Header Turned Off



Including Text

You can add lines of text to a selected window, for example, to annotate it with a title or the current value of a variable. Each text line can contain one or more words or special characters, and you can include as many lines as you need. You can add simple text or a substitution string.

A *substitution string* can display the current value for a variable when you specify the variable name preceded by an ampersand (&). For example, for a variable named PAYMENT, entering **`Payment Amount=$&payment`** as the text line will display the current value for the variable PAYMENT, that is, **`Payment Amount=$1500.00`**. For the axis variables, the software displays the value for the axis cutting plane's current location. For the response variable, the software displays the response value at the intersection of the three cutting planes.

To add text:

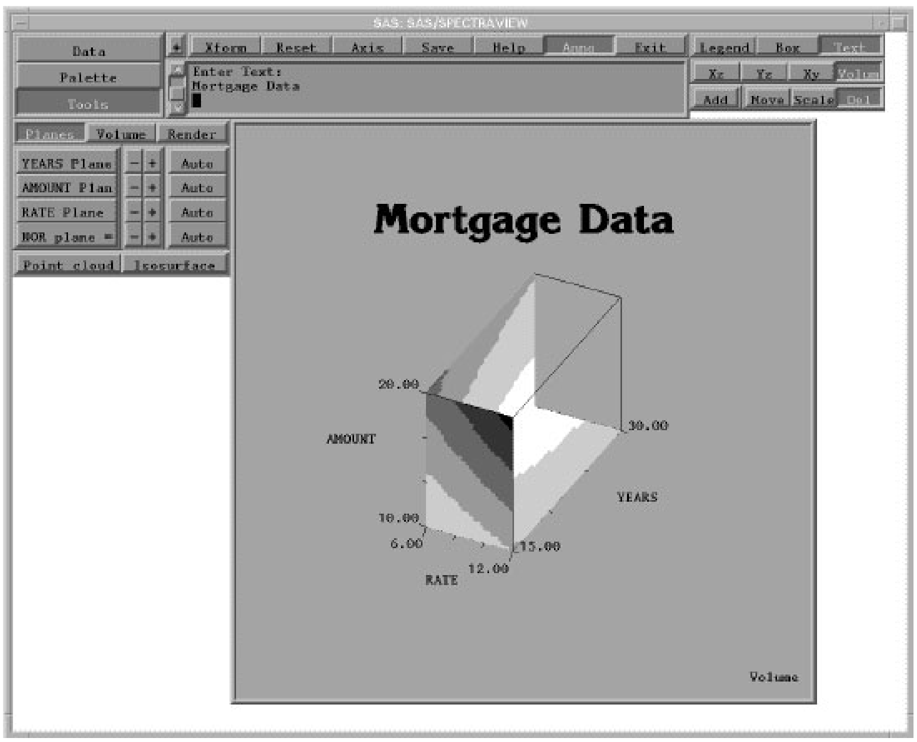
1. Select the Anno global button, then Text.
2. Specify the window by selecting Xz, Yz, Xy, or Volume. Note that the selected window must be visible.
3. Select Add. In the text window, the software displays a message asking you to enter text, and the cursor changes to the input cursor.
4. In the text window, type the text or text string and press Enter. For example, as shown in Display 6.6 on page 78, you might enter the title **`Mortgage Data`**. By default, the first text line is centered at the top of the selected window. The second text line is centered below the first, and so on.
5. To manipulate added text:
   - □ To move text, select Move, position the cursor on the text that you want to move, press the left mouse button, and drag to the appropriate position.

- □ To change the size of text, select [Scale], position the cursor on the text that you want to scale, press the left mouse button, and drag left (smaller) or right (larger).
- □ To delete text, select [Del], position the cursor on the text that you want to delete and click the left mouse button.

6 The default color for text is black. To specify a different color, select [Palette], then [Text].

Adjust the RGB sliders. As you move the sliders, the specified color appears in the square at the left, and the item changes color in the Volume window. *You can also select a color chip, located below the RGB sliders.*

Display 6.6 Volume Window with Text



Including a Response Legend

You can add a legend of response values and associated colors to a selected window. To include a response legend:

1 Select the [Anno] global button, then [Legend].

2 Specify the window by selecting [Xz], [Yz], [Xy], or [Volume].

3 Specify whether you want to display the legend horizontally or vertically by selecting [Hor] or [Ver], as shown in Display 6.7 on page 79.

4 To manipulate the legend:

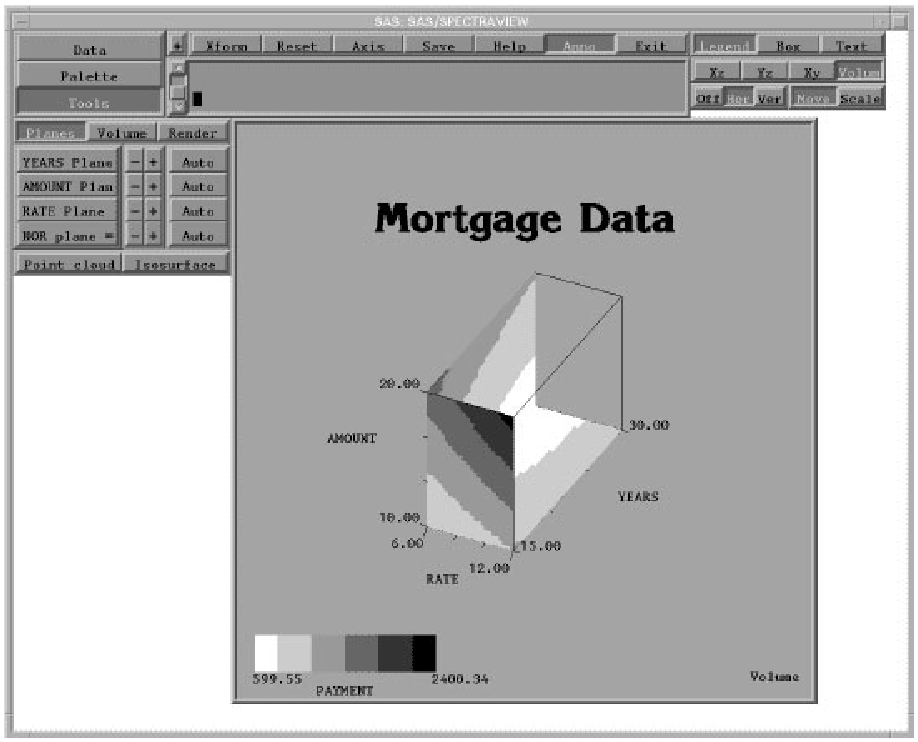
- □ To move the legend, select [Move], position the cursor on the legend, press the left mouse button, and drag to the appropriate position.
- □ To change the size of the legend, select [Scale], position the cursor on the legend, press the left mouse button, and drag left (smaller) or right (larger).

□ To remove a displayed legend, select Off.

5 The default color for the legend's labels, which are the response variable name and minimum and maximum values, is black. To specify a different color, select Palette, then Label.

Adjust the RGB sliders. As you move the sliders, the specified color appears in the square at the left, and the item changes color in the Volume window. *You can also select a color chip, located below the RGB sliders.*

Display 6.7 Volume Window with Response Legend



Transforming an Image

Transforming an image lets you view it from any direction and modify its size.

Rotating an Image

Rotating an image lets you view it from any angle, giving you better visualization to analyze the data. You can rotate any Volume window image and three-dimensional images in the surface windows; you cannot rotate two-dimensional images. To rotate an image:

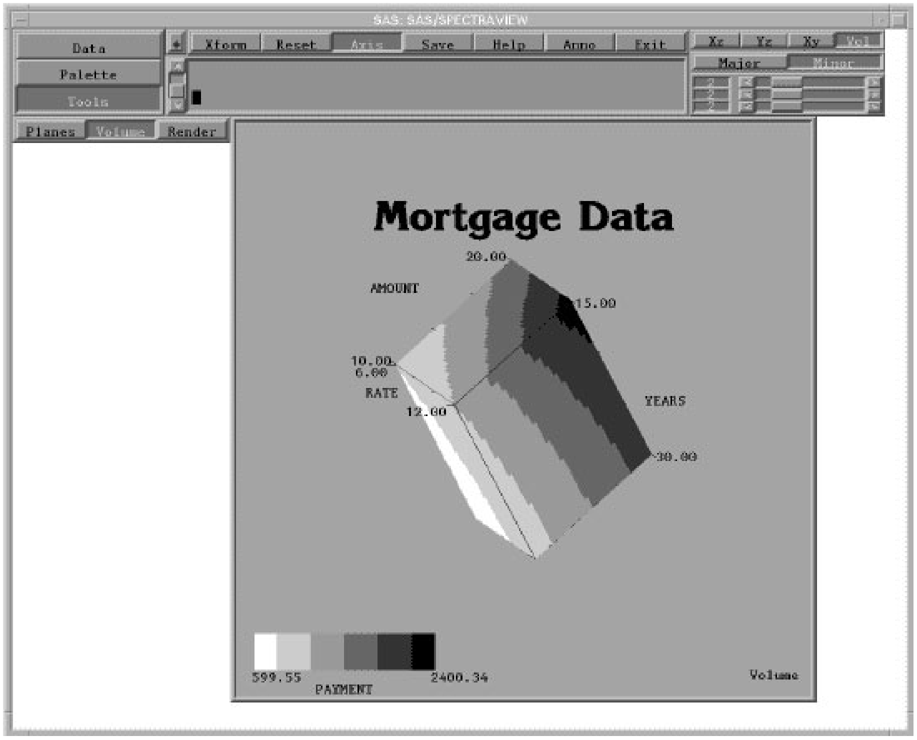
1 Select the Xform global button, then select

Rotate 1	to rotate the image in the Volume window or a three-dimensional image in a surface window.
NOR rotate	to rotate the NOR cutting plane.

Rotate all — to simultaneously rotate the images in all four windows, except for two-dimensional images.

2 Position the cursor within the Volume window or surface window and drag the cursor in any direction. The image becomes wire mesh during the rotation process. It returns to its original format when rotation is complete.

Display 6.8 Rotated Image



Note that if you have a point cloud with many data points, turning it off before rotating will speed the process. Once you reach the new angle (using the bounding box), turn the point cloud on again. Also, note that rotating an image may cause labels to become obscured. To make them visible again, select the Axis global button, specify which window, and click the axis where you want to relocate the labels.

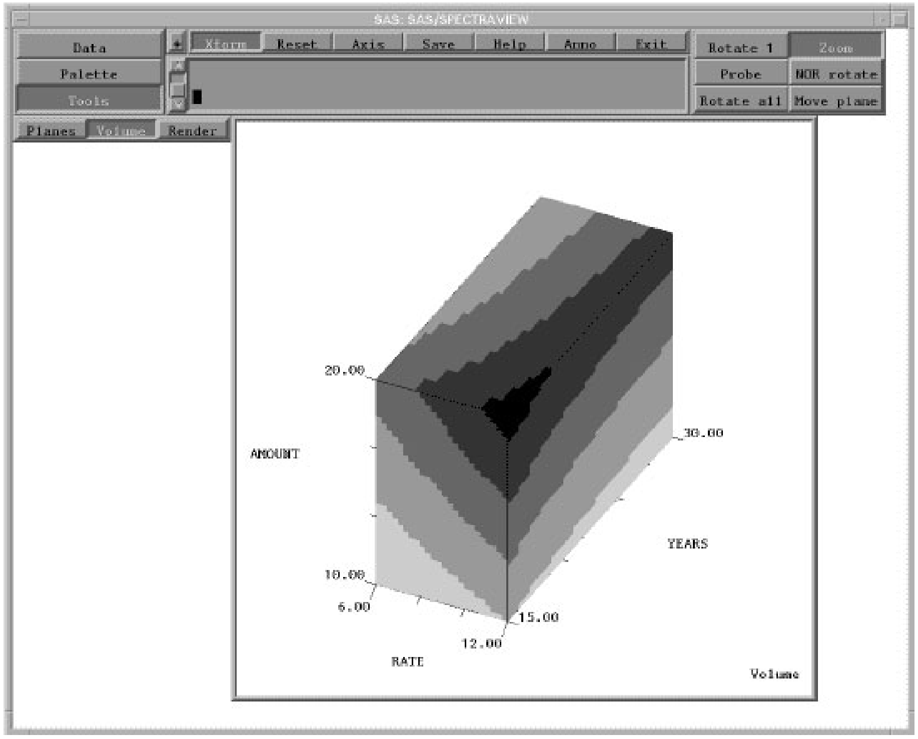
Zooming an Image

Zooming lets you move an image nearer or farther away, making it smaller or larger proportionally. To zoom an image:

1 Select the Xform global button, then select Zoom.

2 Position the cursor within the Volume window or surface window and drag the cursor left (to push the image farther away) or right (to bring the image closer to you).

Display 6.9 Zoomed Image



Probing an Image for a Specific Response Value

A useful customization option is to explore an image for specific response values, referred to as *probing*. Probing allows you to pinpoint the actual response value for a specific data point. To probe an image:

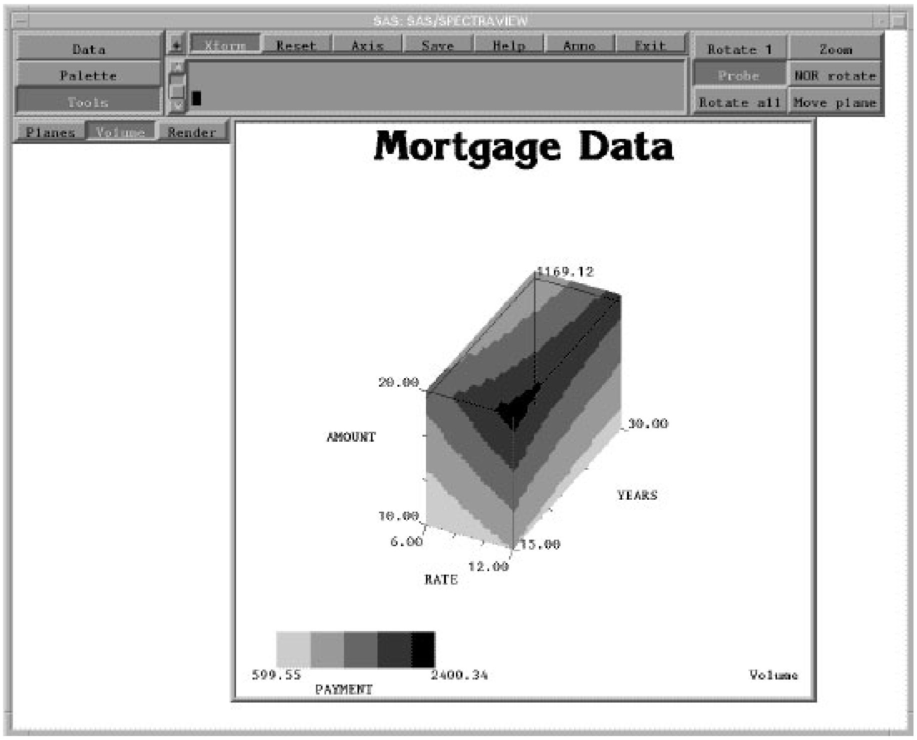
1. Select the Xform global button, then Probe.
2. Position the cursor in the appropriate window, press the left mouse button to display the probe lines, and do one of the following:
   - □ To probe a Volume window image, place the cursor on an axis and drag the probe line to an appropriate point along the axis. Then drag the other axis probe lines to display the response value at the intersection of all three probe lines.
   - □ To probe a three-dimensional image in a surface window, drag the probe lines along the variable axes to display the response value at the intersection of the two probe lines. (The response axis cannot be used to probe for response values.) The value for the horizontal axis is shown below the axis, and the response value is shown above it.
   - □ To probe a two-dimensional image in a surface window, drag the cursor along the horizontal axis. The value for the horizontal axis is shown below the axis, and the response value is shown above it.

   Note that if the data point being probed is a missing value, the software displays a ? for the value.
3. By default the probe lines and displayed response value are black.
   - □ To specify a different color for the displayed response value, select Palette, then Label.

- To specify a different color for the probe lines, select Palette, Bounding box, then User-defined.

Adjust the RGB sliders. As you move the sliders, the specified color appears in the square at the left, and the item changes color in the Volume window. *You can also select a color chip, located below the RGB sliders.*

Display 6.10 Probed Solid-Volume Image Displaying Response Value



CHAPTER 7

Managing the Software

Introduction to Managing the Software

The following topics provide information on how to manage SAS/SPECTRAVIEW, for example, how to save an image, how to configure the interface, and how to control the color palette.

Saving a Displayed Image

Saving to a File

SAS/SPECTRAVIEW allows you to save a displayed image to either a TIFF (Tagged Image File Format) file or a PostScript file, for example, to use in presentations. To save an image:

1 Use a visualization technique to display an image, and include customizations if desired.

2 Select the Save global button.

3 Specify the file format by selecting either PostScript file or TIFF file. *Note that TIFF files can be used as input to other SAS procedures and products.*

4 Use the left mouse button and click the display window you want to save. A message in the text window asks you to enter a filename, and the cursor changes to the input cursor.

5 In the text window, enter a one- or two-level filename for the file. *Pressing Enter without specifying a filename cancels the save request.*

Saving to a SAS Data Set

You can save current data values to a new SAS data set. Then, for example, you can use the new data set with other SAS System products. The values that are saved depend on what is displayed, which in essence is visual subsetting. That is, the data saved depends on what visualization technique(s) are currently active and how they are displayed:

- If a cutting plane is displayed, only the data points on that cutting plane are saved.
- If a point cloud is displayed, only the data points in the point cloud are saved.
- If two or more images are displayed, then the intersection of those images are saved. For example, if two cutting planes are displayed, then only the data points along the intersection of the two planes are saved. Likewise, if both a cutting plane and a point cloud are displayed, only the data points on the cutting plane and in the point cloud are saved.
- If no image is displayed or if all the data is displayed, then all values are saved. In addition, if you loaded a SAS data set for which you specified a WHERE clause to load a subset, you can save the subset of data as a new data set.

Note that missing values are saved, unless the image is a point cloud, then missing values are not saved.

To save data values:

1 Use a visualization technique to display an image, and include customizations if desired.

2 Select the Save global button.

3 Select SAS data set. A message in the text window asks you to enter a filename, and the cursor changes to the input cursor.

4 In the text window, enter a one- or two-level filename for the SAS data set. *Pressing Enter without specifying a filename cancels the save request.*

Configuring the Interface

Resizing the Display Windows

You can control the size of the display windows. That is, you can enlarge a window to fill the entire viewing area, or you can enlarge one window somewhat and maintain the other three on a smaller scale. For example, resizing the display windows to a single window could provide a better view of an image.

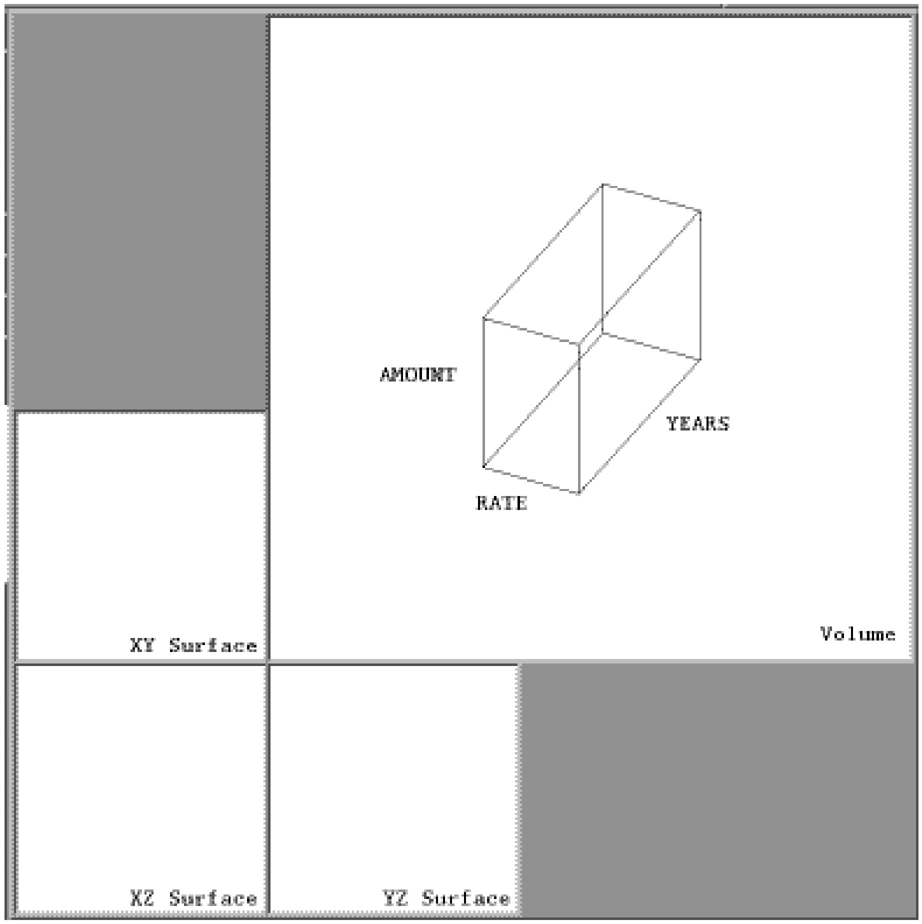
To resize the display windows:

1 Position the cursor in a window.

2 Press the middle (or right) mouse button, drag the cursor, then click the mouse button.

Or position the cursor within the display area where you want the intersection of all four windows to be, then double-click the middle (or right) mouse button.

Display 7.1 Resized Display Windows



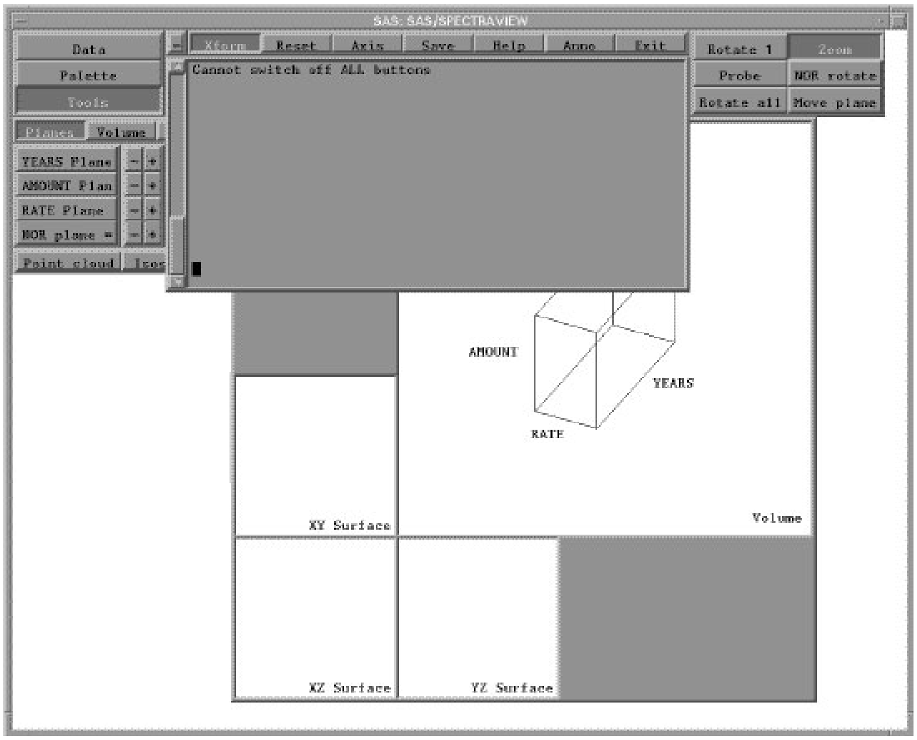
Note that double-clicking in the middle of the display area returns the windows to their default size.

Resizing the Text Window

You can increase the size of the text window with the ⊞ button, which is located in the window's upper left corner. Increasing the size of the text window provides a larger display of software messages.

To make the text window larger, select ⊞. To return the window to its default size, select ⊟.

Display 7.2 Resized Text Window



Controlling the Online Reference Window

In order for the Online Reference window to be visible, the SAS/SPECTRAVIEW interface should pop-to-top only when you select its window border. If the window does not stay on top, you need to modify your window manager files.

PC Systems

In order for the online help to work properly, you must make the SAS/SPECTRAVIEW interface window the active window. Depending on your screen resolution, this may require that the Online Reference window be partially obscured by the SAS/SPECTRAVIEW window while you are selecting the desired button.

X Window System

In order for the online help to work properly, the SAS/SPECTRAVIEW interface window must have focus. You can change your window manager resources to allow the Online Reference window to remain on top while the SAS/SPECTRAVIEW window retains focus. The values you change and the affected files depend on the window manager you are using. See your window manager documentation or your system administrator for assistance.

Resetting the Interface

From the Reset global button, you can reset the following for SAS/SPECTRAVIEW:

View — resets the four display windows' orientation and any zoomed or rotated image to the default view.

Palette	resets all color palette items to their defaults, which includes items such as axis tick marks and labels, bounding box, missing values, an isosurface, and the data ramp to 32 colors with the default ranges and colors. (Any stored color palette is not affected.)
All	resets all items in the software to their defaults, including windows, cutting planes, color palette, marker sizes, and so on.

Changing Background Colors

Prior to invoking SAS/SPECTRAVIEW, you can change the default background colors, which are the *primary background color* that defines the color behind the buttons and the *secondary background color* that defines the color of the display windows. You cannot set both background colors to the same color.

To change background colors:

1. From the SAS PROGRAM EDITOR window menu bar, select **`Globals`**, **`Options`**, then **`Color setup`**. The software opens the SASCOLOR window.
2. To change the primary background color, select **`Background`**, click the color from the color palette, and select Save.
3. To change the secondary background color, select **`Secondary Background`**, click the color from the color palette, and select Save.
4. To close the window, save the changes, and return to the Display Manager window, select OK.

The changes are saved to the catalog entry SASUSER.PROFILE.SAS.CPARMS.

Controlling the Color Palette

The following topics explain how to save your color customizations, how to load a saved color palette, and how to reset the color palette to its default specifications.

Saving Color Specifications to a File

You can save your color specifications for all color palette items, such as tick marks, labels, response value ranges, and so on. Then you can load the stored palette file to apply to a current image. To save color specifications:

1. Select Palette.
2. Select Data ramp.
3. Select Store. The software prompts you for a filename.
4. Enter a filename in the text window. The software saves the file to the directory from which you invoked SAS/SPECTRAVIEW. *You cannot specify a different directory in which to save the file.*

The software saves your color specifications for all color palette items, not just the data ramp.

Loading Color Specifications Stored in a File

Once you have saved color specifications in a file, you can load the stored palette file to apply to an image. Note that when you saved the file, the software saved it to the directory from which you invoked SAS/SPECTRAVIEW. To load that file, you must invoke SAS/SPECTRAVIEW from that location or copy the saved file to the appropriate location.

To load a stored palette file:

1. Select Palette.
2. Select Data ramp.
3. Select Load. The software prompts you for a filename.
4. Enter the filename in the text window.

The software loads the saved color specifications and applies them to the current session.

Resetting the Color Palette

To reset the current color palette to default colors, select the Reset global button, then Palette. The software highlights Palette, indicating that you need to confirm the request. To confirm, select Palette again.

When you reset the color palette, the software returns all currently assigned colors to their defaults. Stored palette files are not affected.

PART 2

Appendices

APPENDIX

1

The G4GRID Procedure

Introduction to the G4GRID Procedure

The G4GRID procedure creates a SAS data set to be used with the SAS/SPECTRAVIEW three-dimensional visualization software. The procedure enhances SAS/SPECTRAVIEW software's ability to create two- and three-dimensional graphic images from data.

PROC G4GRID allows you to interpolate three-dimensional data contained in a SAS data set. While PROC G4GRID is not required to load a data set into SAS/SPECTRAVIEW, the procedure does provide advantages:

- It enables the loading of data sets that could not otherwise be loaded due to memory constraints. SAS/SPECTRAVIEW creates a three-dimensional volume grid with values for each unique x,y,z coordinate contained in the data set. For sparse data sets, this can result in data points, many of which may be missing. By using PROC G4GRID prior to loading the data set, you can replace filler points with interpolated values or resize the data set as required.
- It creates larger or smaller resolutions of a data set to enhance or facilitate your view of the data. Using PROC G4GRID, you can create a data set of almost any desired resolution for use in SAS/SPECTRAVIEW.
- It creates more meaningful volume data from randomly sampled data points.

PROC G4GRID works by interpolating randomly distributed points in three dimensions and by creating a volume grid with data points at each grid location. The procedure uses a Gaussian interpolation function that varies as the inverse of the square of the distance from surrounding data points. While PROC G4GRID is similar in syntax and function to the G3GRID procedure, G4GRID uses a different interpolation method and does not render the same values as G3GRID.

You can run the procedure using either of these methods:

- PROC G4GRID syntax
- graphical interface .

Data Considerations

Each data set processed by the G4GRID procedure must contain at least four numeric variables:

- three axis variables
- one or more response variables.

If you specify more than one response variable, the procedure performs a separate analysis and produces interpolated values for each response variable. Unlike the G3GRID procedure, the G4GRID procedure does not require that the input data coordinate values be unique. It uses all coordinate values that fall within the grid volume.

Terminology

The following terms are specific to the G4GRID procedure:

coordinates
: the values representing the location of a data point along the X, Y, Z axes.

interpolate
: to estimate values between two or more known values.

response
: a value representing a measured response associated with each data point in the volume.

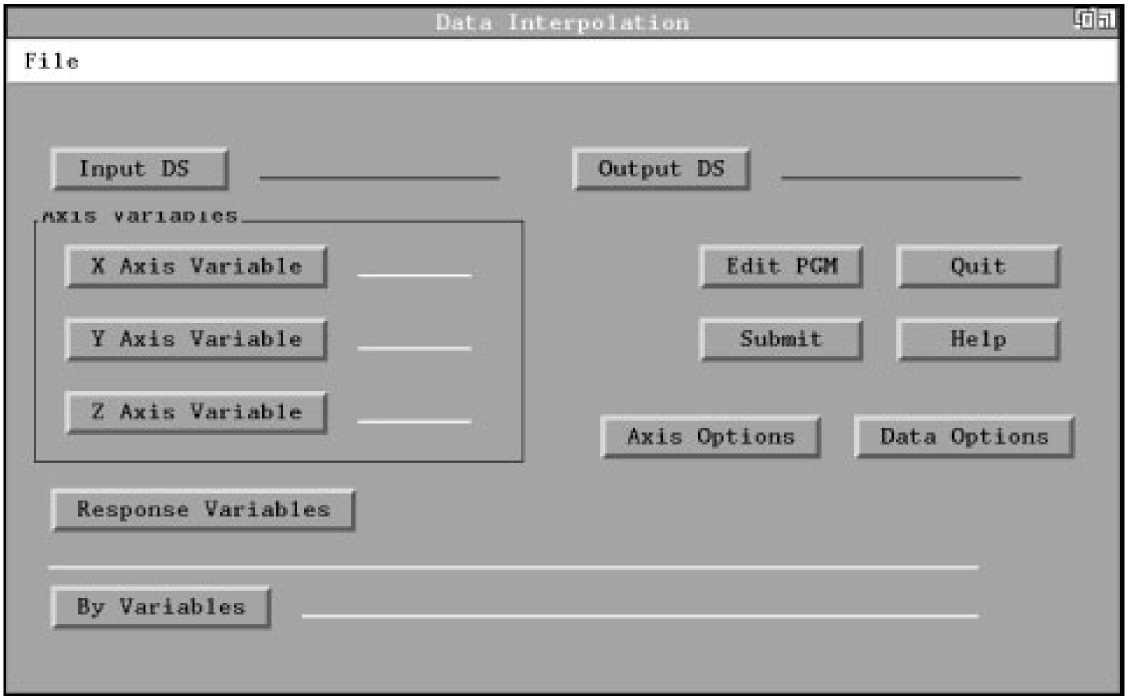
variables
: the SAS data set variables whose values are plotted on the X, Y, and Z axes.

Using the Graphical Interface

A FRAME entry application gives you a graphical interface for the G4GRID procedure so that you can process a G4GRID request without knowing the PROC G4GRID syntax. To use the G4GRID interface, issue the command:

```
af c=sashelp.g4grid.g4grid.frame
```

When the Data Interpolation window displays, see its online help for how to use it.



Proc G4GRID Syntax

PROC G4GRID<DATA=*SAS-dataset*>
<OUT=*SAS-dataset*>
<GAMMA=*gamma-factor*>
<D=*amplitude-factor*>
<RADIUS=*search-radius-value*>;
GRID *x*y*z=variable* <*...variable-n*>< / <*grid-options*>>;
BY <*options*> *variables*;

"The PROC G4GRID Statement" on page 93 initiates the procedure.
"The GRID Statement" on page 95 specifies the variables for interpolation.
"The BY Statement" on page 96 specifies the variable or variables by which data is grouped for processing.

The PROC G4GRID Statement

PROC G4GRID <DATA=*SAS-dataset*>
<OUT=*SAS-dataset*>
<GAMMA=*gamma-factor*>
<D=*amplitude-factor*>
<RADIUS=*search-radius-value*>;

The PROC G4GRID statement initiates the G4GRID procedure. If necessary, it can specify the input data set, the output data set, or both. This statement is required.

Options

DATA= ***SAS-dataset***

specifies the input data set. By default, the procedure uses the most recently created data set, or you can use the DATA= option to specify a data set. If no data set has been created in the current SAS session and the DATA= option is not supplied, an error occurs and the procedure stops.

The input data set must contain at least four numeric variables, and it must contain the numeric variables specified in the GRID statement used with the PROC G4GRID statement.

OUT=*SAS-dataset*

specifies the output data set. The data set will contain any BY variables specified, the values for the volume variables at the grid coordinates, and interpolated values for each response variable specified. By default, the output data set uses the *variable-n* naming convention as the data set name.

GAMMA=*gamma-factor*

is the first of three parameters that control the interpolation of data points. The GAMMA= value controls the "locality" of the data interpolation.

The routine makes two passes through the data. In each pass, a Gaussian weight is applied to each observation based on its distance from the grid point in question. On the second pass, the interpolation function weighting is narrowed so that points nearer the grid point location contribute more than points farther away. The numeric value of *gamma-factor* controls the narrowness of this second Gaussian pass.

The valid range for the GAMMA= value lies between zero (non-inclusive) and one (inclusive). A smaller value means closer points contribute more heavily during the second pass than when a larger value is specified. A normal range for gamma is 0.3 to 0.5, with 0.3 being the default.

D=*amplitude-factor*

is the second of three parameters that control the interpolation of data points. The D= value controls how the routine tries to resolve wavelengths embedded in the data.

The procedure tries to resolve some fraction of a target wavelength. The D= option indicates what fraction the procedure should attempt to resolve. Shorter wavelengths are resolved with greater amplitudes, while longer wavelengths are resolved with lesser amplitudes. The target wavelength that the procedure attempts to resolve is twice the average spacing of the input coordinates (that is, the smallest justifiable wavelength).

For example, assume that a data set has an average spacing of 3.5 units. The target wavelength for the data set is 7.0 units. The procedure attempts to resolve the D= fraction of the amplitude of this wave. If D=0.6 is specified, then the routine attempts to resolve 60 percent of amplitude of wavelengths in the data of 7.0 units. Longer waves have less than 60 percent of their amplitude resolved, while shorter waves have more than 60 percent resolved. The larger the value, the greater the detail resolved. However, too high a value results in spurious waves of smaller wavelengths being introduced into the data. That is, the larger the value specified, the more detailed the result, but also the more noise is introduced. The default value is 0.8.

RADIUS=*search-radius-value*

is the third of three parameters that control the interpolation of data points. The RADIUS= value sets the limit of the search radius during interpolation.

Points outside the search radius do not contribute to the interpolated value of the grid point in question. Data points that lie farther from a grid point contribute less to the interpolated value of that grid point. Some points are far enough away that their contribution can simply be ignored. Specifying a search radius that ignores them can significantly speed the data interpolation.

The valid range lies between zero (non-inclusive) and one (inclusive). For example, a search radius of 0.5 uses only data points that lie within half the grid volume distance from the grid location in question. To use all data points in the grid volume for each grid point, specify **`missing`**. The default value is 0.5.

Note: If an output grid contains missing values, increasing the search radius may eliminate them. △

The GRID Statement

GRID *x*y*z=variable< ...variable-n> </ <grid-options>>*;

The GRID statement specifies the variables for interpolation. These include the volume variables to be used as coordinates for the output grid, as well as one or more response variables to be interpolated. You can specify multiple response variables for the three volume variables that you specify. Options for the GRID statement control the observations produced in the output data set. You can specify only one GRID statement.

Options

Options affect all output produced by the GRID statement. If you use any of these options, separate them from the grid request with a slash (/). If you do not use any options, omit the slash.

XAXIS=*list-of-coordinates*
YAXIS=*list-of-coordinates*
ZAXIS=*list-of-coordinates*
specifies a list of numeric values to be assigned to the respective axis. The list does not have to be in order; it will be arranged in ascending numeric order. The list should not contain duplicates. The list overrides the number of values defined by the NXAXIS=, NYAXIS=, or NZAXIS= options. The list of coordinates may be specified in the following ways:

```
n n ...n
n,n, ...  ,n
n TO n <BY increment>
<n ... n> n,n ... ,n TO n <BY increment> <n, ... ,n>
```

NXAXIS=*n*
NYAXIS=*n*
NZAXIS=*n*
specifies the number of coordinates along the respective axis. The default is 11. The minimum and maximum values along each axis are used as its starting and ending points. Each axis is divided into *n*-1 equally spaced sections. These options are ignored if the corresponding XAXIS=, YAXIS=, or ZAXIS= options are specified.

XEXTRA=*n*
YEXTRA=*n*
ZEXTRA=*n*
specifies the number of temporary grid cells to be added to the interpolation along the respective axis. Normally, only points within the grid volume are included in the interpolation. However, this can result in some edge effect errors, since no data

outside this volume are used. To avoid such conditions, you can set the option to a positive non-zero value.

The distance of the new points from the end, which is the size of each "border," is the width of the first and last coordinate pair, respectively. All points that fall within the new coordinate boundaries are used in the interpolation. The size of the output grid is not affected.

UNIFORM
instructs the G4GRID procedure to assume that all coordinates along each axis are equally spaced. This allows the procedure to take some short cuts to slightly speed processing. The UNIFORM option is not necessary unless XAXIS=, YAXIS=, or ZAXIS= options are specified.

If XAXIS=, YAXIS=, or ZAXIS= is specified, PROC G4GRID must assume that the spacing between output grid points is not equal. As a result, special processing must occur for each correction pass.

If none of the above options are specified or the UNIFORM option is specified, the procedure assumes all points in the output grid are equally spaced and omits the special processing. If you specify one of the above options and the output grid points are known to be equally spaced, specifying the UNIFORM option may slightly speed processing. If, however, the output grid points are unevenly spaced to a significant extent, and the UNIFORM option is specified, erroneous results can occur. UNIFORM is the default unless XAXIS=, YAXIS=, or ZAXIS= is specified.

The BY Statement

BY <*options*> *variables*;

The BY statement specifies the variable or variables by which data is grouped for processing. A separate grid is produced for each value of the BY variable. The BY statement is optional.

Options

DESCENDING
indicates that the input data set is sorted in descending order. This is the default.

NOTSORTED
indicates that observations with the same BY variable values are to be grouped as they are encountered, without regard for whether the values are in numerical order.

Examples of Various G4GRID Invocations

The following examples illustrate various G4GRID invocations.

Example 1: Output Volume Grid Ranging from Minimum to Maximum

Assume that data set MYLIB.SAMPLE contains the variables MY_X, MY_Y, and MY_Z. At each x,y,z coordinate location specified by these variables, there is a value for the response variable RANDAT. Furthermore, there are multiple groups of these values specified by the BY variable BYIT.

To create a data set with an output volume grid that ranges from the minimum to the maximum along each coordinate for each BY group, you could use the following syntax:

```
proc g4grid data=mylib.sample;
    grid my_x*my_y*my_z=randat / nxaxis=5 nyaxis=5 nzaxis=5;
    by byit;
run;
```

Example 2: Output Volume Grid Ranging from 1 to 10

Using the same data set, assume there is also a variable RANDAT2 for which you would like output volume grids. This time, you would like the grid to run from 1 to 10 with values at each even unit distance. You could use the following syntax:

```
proc g4grid data=mylib.sample;
  grid my_x*my_y*my_z=randat randat2 /
  xaxis=1 to 10 yaxis=1 to 10 zaxis=1 to 10;
  by byit;
run;
```

References

Barnes, Stanley L., "Mesoscale Objective Map Analysis Using Weighted Time-Series Observations," United States Oceanic and Atmospheric Administration, Environmental Research Labs, Norman, OK, *Technical Memorandum (NOAA TMM ERL NSSL-62)*, March 1973.

APPENDIX 2

Sample SAS Data Sets

Introduction to Sample SAS Data Sets

This appendix gives information about the two SAS data sets that are used for this document's examples.

If your site has installed the sample data sets, they are found in the sample subdirectory for the product. For example, for a UNIX environment, the location may be **`!SASROOT/spectraview/sample`**. Please note that this location is the default and is not necessarily the location of the sample programs at your site. Contact your SAS Software Representative for information on how to access the sample library files.

EPA Data Set

The EPA data set describes the results of measuring sulfate concentrations.

The following table shows the variables in EPA and how they were assigned to SAS/SPECTRAVIEW variables. Note that the optional BY variable is included.

Table A2.1

Variable	Dimension	Description
HOURS	BY Variable	hour that the reading was taken
LEVEL	Z Variable	layer above ground
LNGITUDE	X Variable	longitude coordinate
LATITUDE	Y Variable	latitude coordinate

Variable	Dimension	Description
SULFATE	Response	sulfate concentration percentage
OZONE	not used	ozone concentration percentage

MORTGAGE Data Set

The MORTGAGE data set contains mortgage payment comparisons for various numbers of years, interest rates, and loan amounts. The data set has 16,400 observations.

The following table shows the variables in MORTGAGE and how they were assigned to the required SAS/SPECTRAVIEW variables.

Table A2.2

Variable	Dimension	Description
YEARS	Z Variable	number of years for loan
RATE	X Variable	interest percentage rate
AMOUNT	Y Variable	loan amount
PAYMENT	Response	monthly payment amount

APPENDIX

3

Recommended Reading

Recommended Reading

Here is the recommended reading list for this title:

- *SAS Language Reference: Concepts*
- *SAS Language Reference: Dictionary*
- *Base SAS Procedures Guide*

For a complete list of SAS publications, see the current *SAS Publishing Catalog*. To order the most current publications or to receive a free copy of the catalog, contact a SAS representative at

SAS Publishing Sales
SAS Campus Drive
Cary, NC 27513
Telephone: (800) 727-3228*
Fax: (919) 677-8166
E-mail: sasbook@sas.com
Web address: **support.sas.com/pubs**
* For other SAS Institute business, call (919) 677-8000.

Customers outside the United States should contact their local SAS office.

Glossary

autoexec file
: an external file containing SAS statements that are executed automatically when the SAS System is invoked. You can use the autoexec file to specify SAS system options, as well as librefs and filerefs.

axis
: a one-dimensional line representing the zero point on a scale used to plot values of the x,y,z coordinates. The X axis represents width, the Y axis represents height, and the Z axis represents depth.

axis variables
: the data set variables that you specify for the X axis, the Y axis, and the Z axis.

bounding box
: a set of lines that outline the three-dimensional volume grid representing the data.

BY variable
: an optional variable specification whose values define groups of observations, such as hour, month, or year. Specifying a BY variable allows you to animate an image so that you can see how response values change according to some grouping, like over time.

categorizing data
: to group numeric data creating distinct ranges for each axis.

click
: to press and quickly release a mouse button.

color palette
: lets you customize colors for response value ranges, missing values, an isosurface, and image annotations.

continuous data
: data containing few gaps that vary slightly over a large range, like weight and height.

coordinates
: the values representing the location of a data point along the X, Y, Z axes.

cutting plane
: a visualization technique that produces slices of data, either perpendicular to an axis or not perpendicular to any axis. In addition to the axis cutting planes, you can request two- and three-dimensional surface views (surfaces, charts, stacks, and plots) at a specific cutting plane's location.

data filtering
a data processing option that smooths or sharpens the data in order to deemphasize or highlight variations in response values. SAS/SPECTRAVIEW provides several predefined filters, or you can create your own.

data point
the location in three-dimensional space of an x,y,z coordinate.

data ramp
the column of data value ranges and associated colors on the color palette that you can use to customize response colors in images.

direct volume rendering
a visualization technique that creates a two-dimensional image of the entire volume of data points with transparency.

discrete data
data consisting of distinct values (non-continuous) containing natural gaps, like patient IDs and years.

display windows
the four interface windows that display the images for viewing and manipulation.

drag
to press and hold a mouse button while moving the mouse.

filler points
the data points generated by the software to fill the volume grid in locations not occupied by data points plotted from X, Y, and Z values.

grid
See *volume grid*.

interpolate
to estimate values between two or more known values.

isosurface
a visualization technique that produces a three-dimensional surface by connecting all the data points with one response value.

libref
an arbitrary name that you make up to symbolically represent a SAS library.

major tick marks
the points on an axis that mark the major divisions of the axis scale.

minor tick marks
the divisions of the axis scale that fall between major tick marks.

missing values
a value in the SAS System indicating that no data is stored for the variable in the current observation.

NOR plane
a cutting plane that is not perpendicular (non-orthogonal) to any axis.

observation
the horizontal component of a SAS data set. Each observation contains one data value for each variable in the data set. In SAS/SPECTRAVIEW software, an observation contains the x,y,z coordinates for a data point and the response value for that location.

opacity degree
the setting that determines the data points that are to be visible in an image produced by direct volume rendering.

point cloud
a visualization technique that displays response values with colored markers, showing individual data points.

probe analysis
a customization option that displays the response value for a specific data point in an image.

response histogram
a bar chart that displays both the frequency distribution of data points that share the same response value and the colors that represent those values. You use the histogram to specify a subset of response values for a point cloud and to specify the response value for an isosurface.

response values
the measured responses that are associated with each data point in the volume. The measured response could be, for example, density, porosity, or wind velocity.

response variable
the SAS data set variable whose values are measured responses that are associated with each data point in the volume.

RGB
a color-coding scheme that specifies a color in terms of percentages of red, green, and blue components.

SAS data set
the internal format used by the SAS System. A SAS data set contains descriptor information and its related data values organized as a table of observations (rows) and variables (columns) that can be processed by the SAS System.

solid-volume image
a visualization technique that produces a colored, three-dimensional, solid-block image of the data points, providing an overall view of the data at the volume's border.

sparse data
data that does not contain at least one value for an x,y,z coordinate within the volume grid, like random data.

splat width
the setting that determines how solid the rendered image is to be.

text window
a window in the interface where messages give you instructions, inform you of current conditions and errors, and so on. You also enter values and names in this window.

three-dimensional
having three dimensions, that is, width, height, and depth. A three-dimensional area can be defined by points having three coordinates relative to the X, Y, and Z axis.

TIFF file
a Tagged Image File Format, which is a format for scanned images.

two-dimensional
having height and width but not depth. A two-dimensional plane consists of points or lines defined by coordinates that are relative to two axes, for example, the X axis and the Y axis.

variable
a column in a SAS data set. Each SAS variable can have the following attributes: name, type (character or numeric), length, format, informat, and label.

volume grid
an invisible network of lines that intersect data points in three-dimensional space. The shape and size of the volume grid is determined by the number of distinct X, Y, and Z values.

volume visualization
the process of creating a three-dimensional image that represents data, providing both surface and interior views.

WHERE clause
allows you to specify a subset of data to be read into SAS/SPECTRAVIEW.

X axis
the axis that specifies width in the coordinate system. The first value in an x,y,z coordinate set represents a location in space relative to the X axis.

X variable
the SAS data set variable whose values are plotted on the X axis, which creates the horizontal axis.

Y axis
the axis that specifies height in the coordinate system. The second value in an x,y,z coordinate set represents a location in space relative to the Y axis.

Y variable
the SAS data set variable whose values are plotted on the Y axis, which creates the vertical axis.

Z axis
the axis that specifies depth in the coordinate system. The third value in an x,y,z coordinate set represents a location in space relative to the Z axis.

Z variable
the SAS data set variable whose values are plotted on the Z axis, which creates the depth axis.

Index

Your Turn

If you have comments or suggestions about *SAS/SPECTRAVIEW® 9.1 User's Guide*, please send them to us on a photocopy of this page, or send us electronic mail.

For comments about this book, please return the photocopy to

SAS Publishing
SAS Campus Drive
Cary, NC 27513
E-mail: **yourturn@sas.com**

For suggestions about the software, please return the photocopy to

SAS Institute Inc.
Technical Support Division
SAS Campus Drive
Cary, NC 27513
E-mail: **suggest@sas.com**