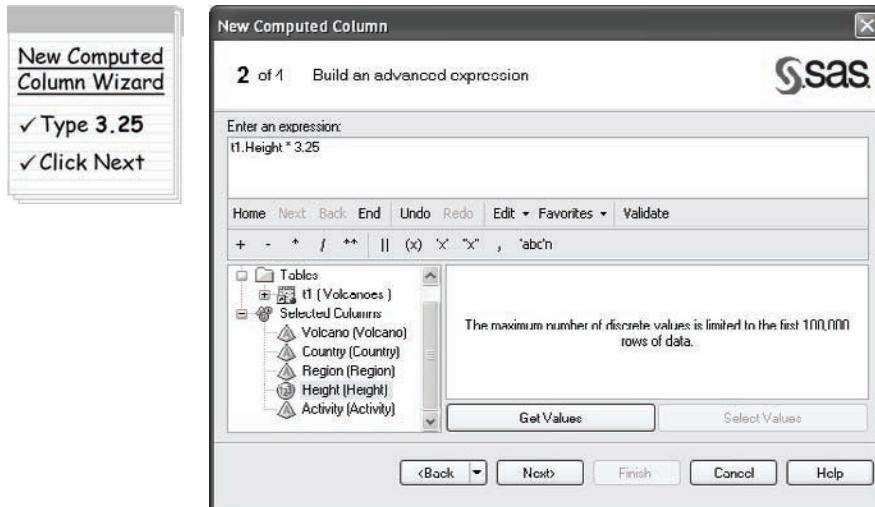
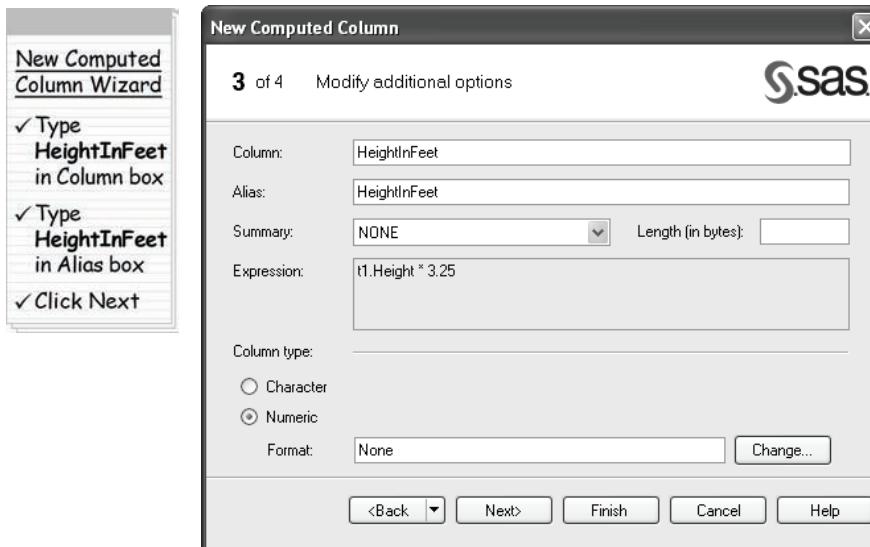


To complete the expression, type the number 3.25 in the Expression text box after the asterisk.

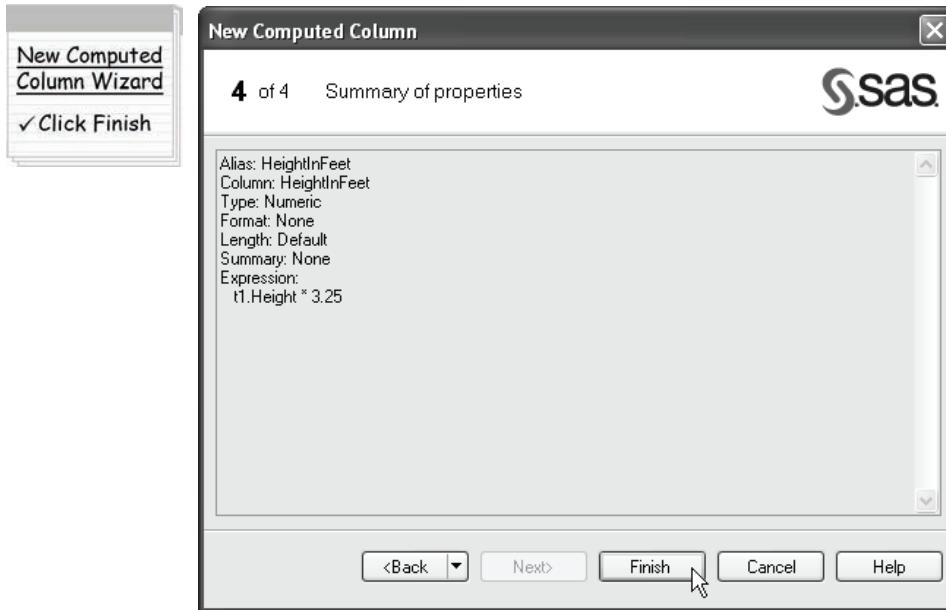


Click Next. Notice that the column has been given the name _Calculation and the alias Calculation. The alias will become the column name in the resulting data table. You could leave these as they are, but it is better to give the column a meaningful name. In both the Column text box and Alias text box, enter **HeightInFeet**.

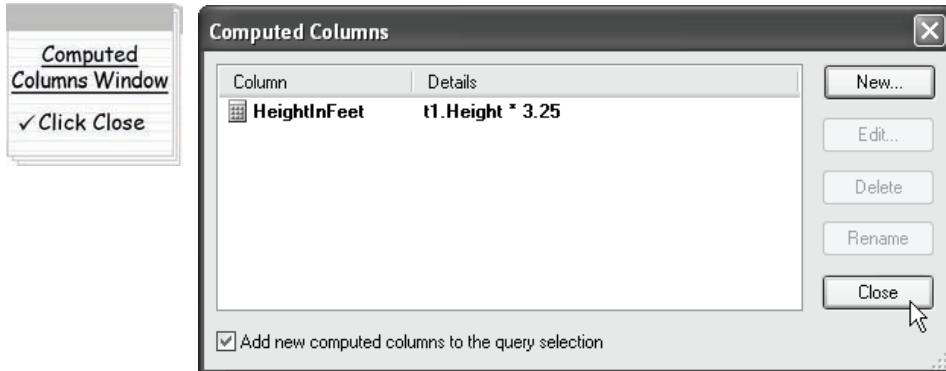


Click Next.

The final window of the New Computed Column wizard simply gives a summary of the new computed column.

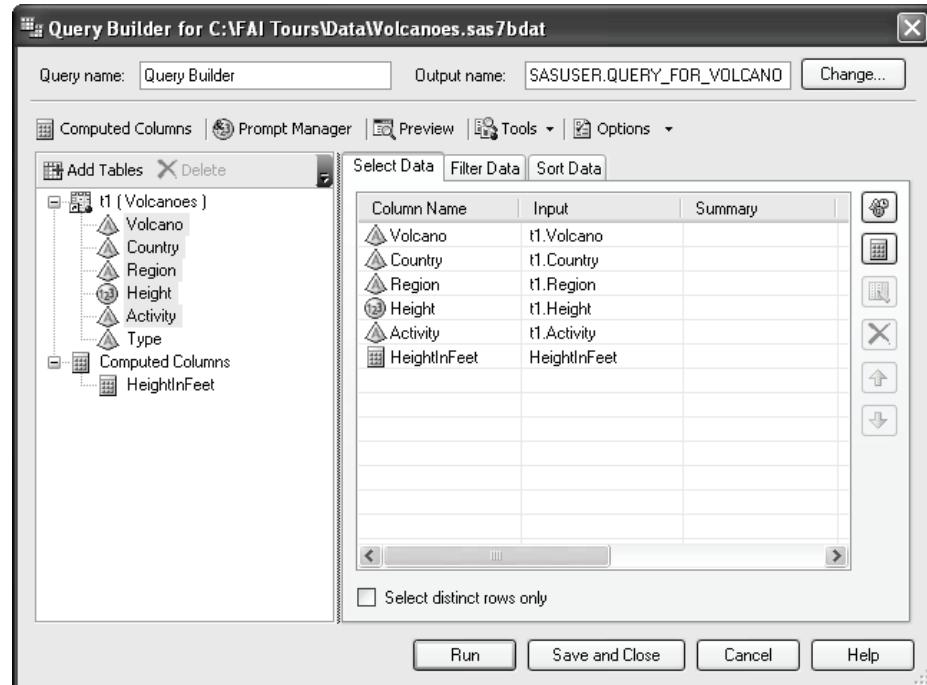


Click **Finish** to return to the Computed Columns window.



Click **Close**.

Notice that the new column, HeightInFeet, appears in the list of columns on the Select Data tab as well as under Computed Columns in the list on the left.

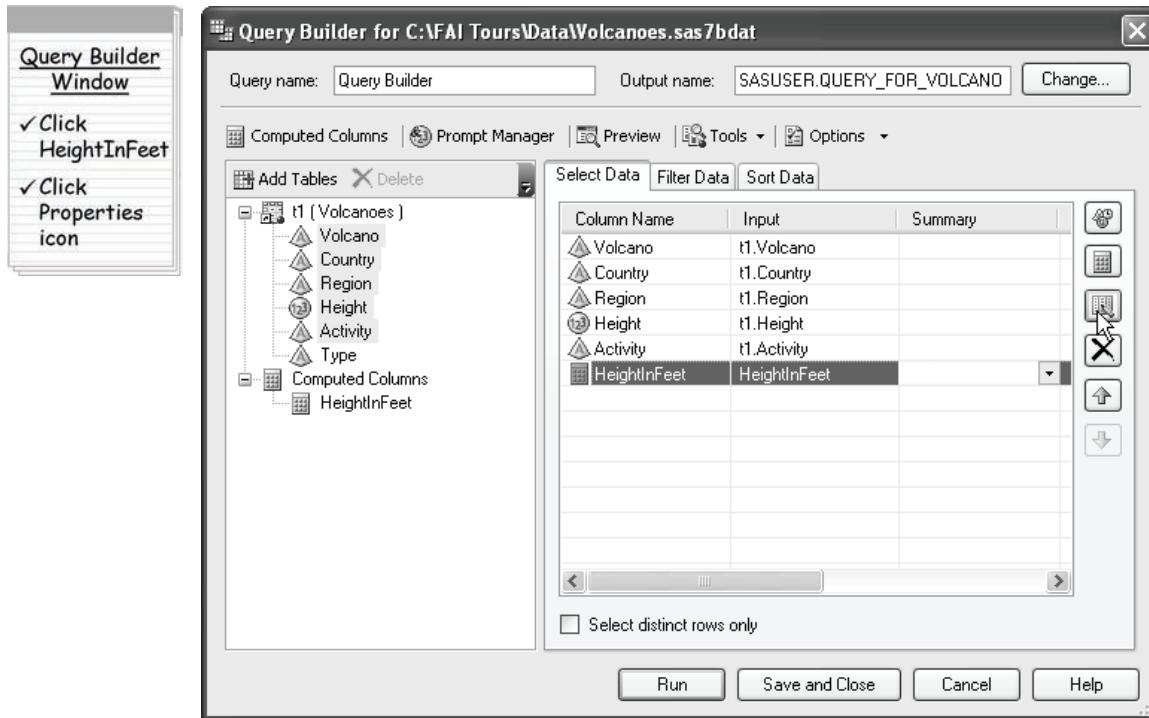


Tutorial C

Click the Refresh icon  in the Preview window to preview the result of the query with the newly computed column.

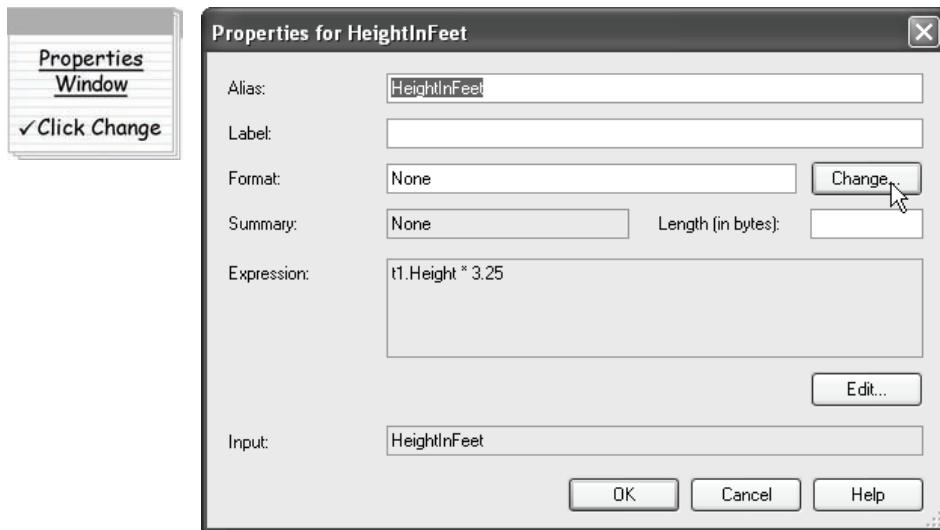
Volcano	Country	Region	Height	Activity	HeightInFeet
Altar	Ecuador	SA	5321	Extinct	17293.25
Arthur's Seat	UK	Eu	251	Extinct	815.75
Barren Island	India	As	354	Active	1150.5
Elbrus	Russia	Eu	5633	Extinct	18307.25
Erebus		An	3794	Active	12330.5
Pt. no	T+olay	Pi	2250	Active	10887.5

The values for HeightInFeet are correct, but they could use some formatting. It is not necessary to show fractions of feet, and it would be nice to have commas in the numbers to make them easier to read. To change the display format of the HeightInFeet column, open the Properties window for the column. Click **HeightInFeet** in the list of column names on the Select Data tab.

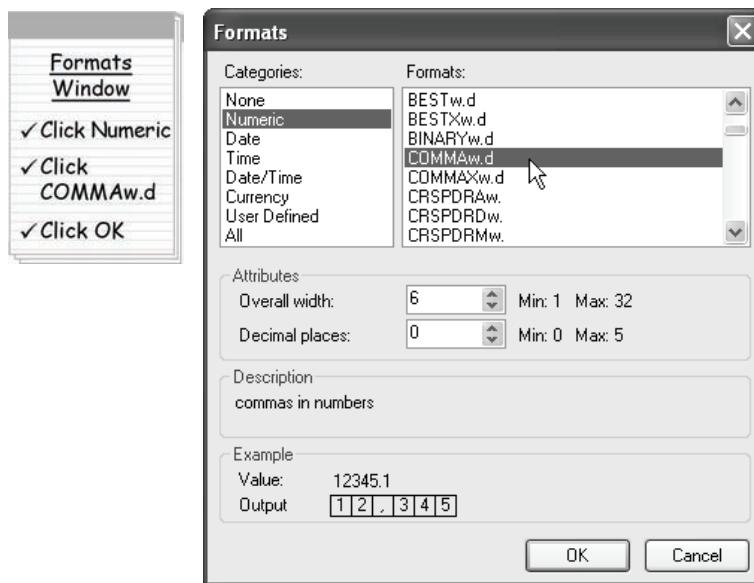


Then click the Properties icon located on the right side of the Query Builder window.

The HeightInFeet column currently has no format associated with it.

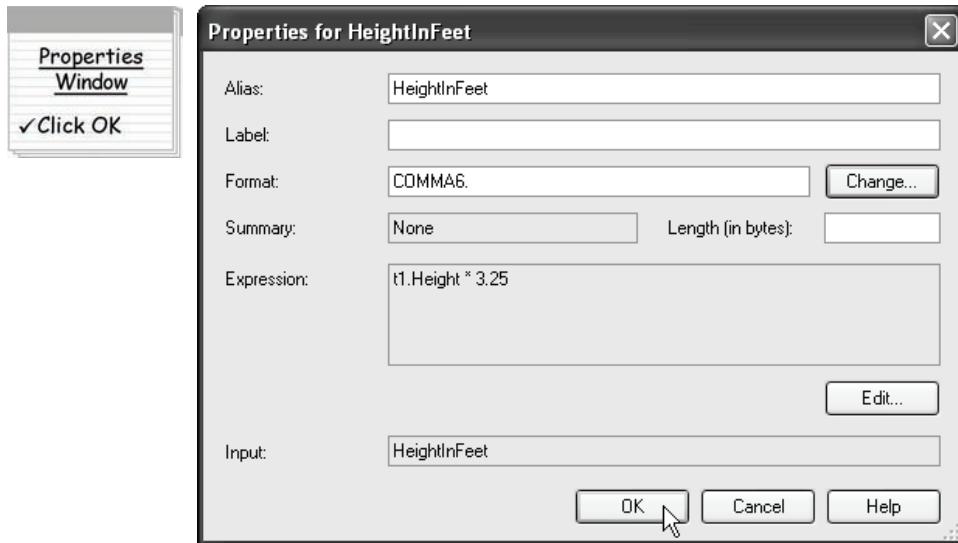


Click Change. This opens a window where you can select a format for the column. Select the **COMMAdw** format from the **Numeric** group. The default width of 6 is fine for this column because the heights of all the volcanoes are at most 6 digits including the comma. Make sure that the number of decimal places is set to 0.

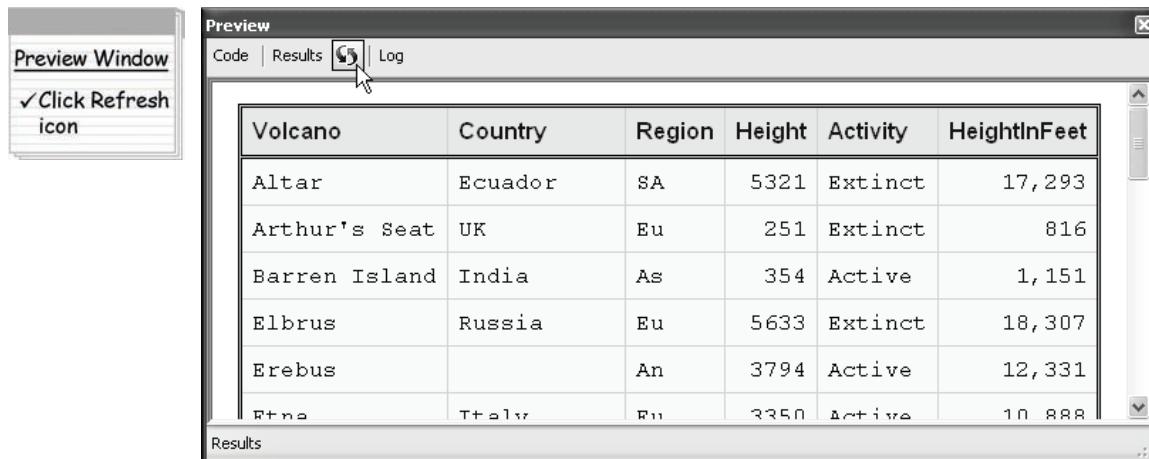


Click **OK** to return to the Properties window.

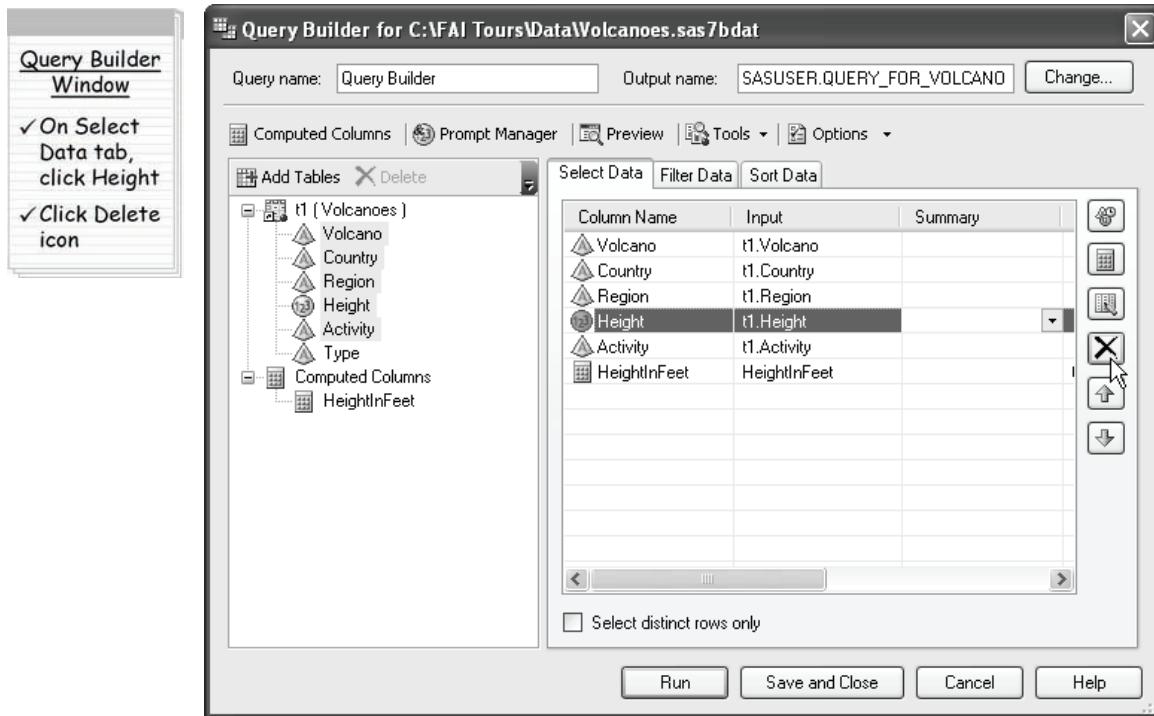
Now the format for the column, COMMA6., appears in the Format area of the Properties window.



Click **OK**, then click the Refresh icon in the Preview window to see the result of setting the format for the HeightInFeet column.

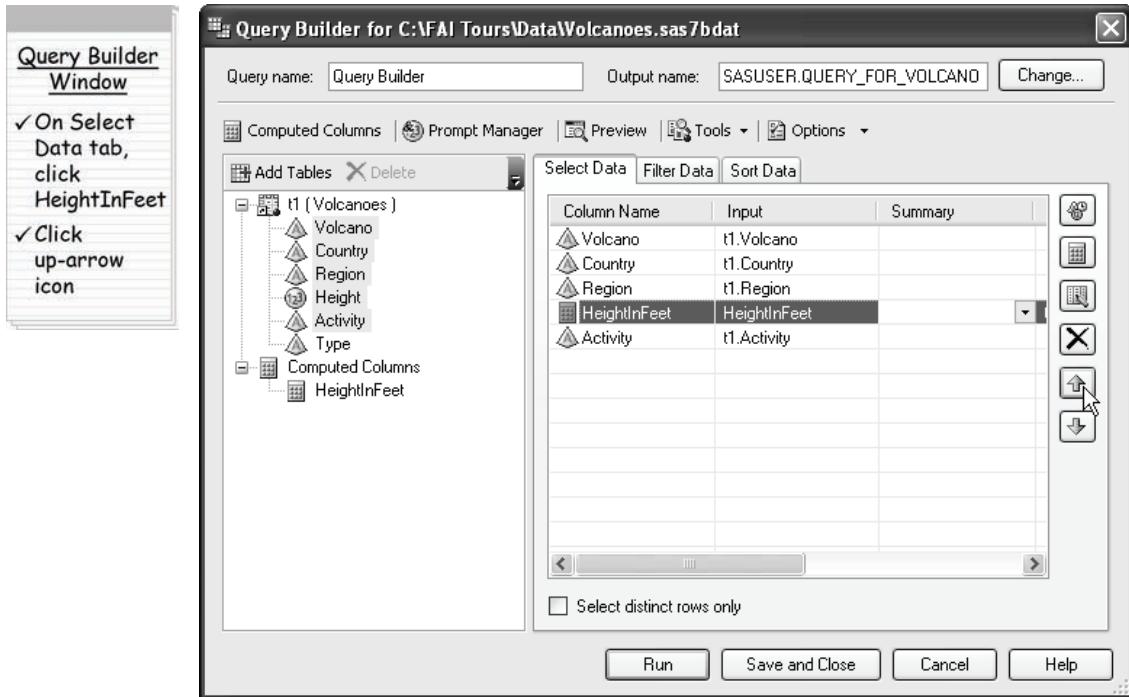


Ordering and removing columns Now that you have the HeightInFeet column, you no longer need the Height column. To remove it from the query, first select it by clicking **Height** in the list of columns on the Select Data tab.

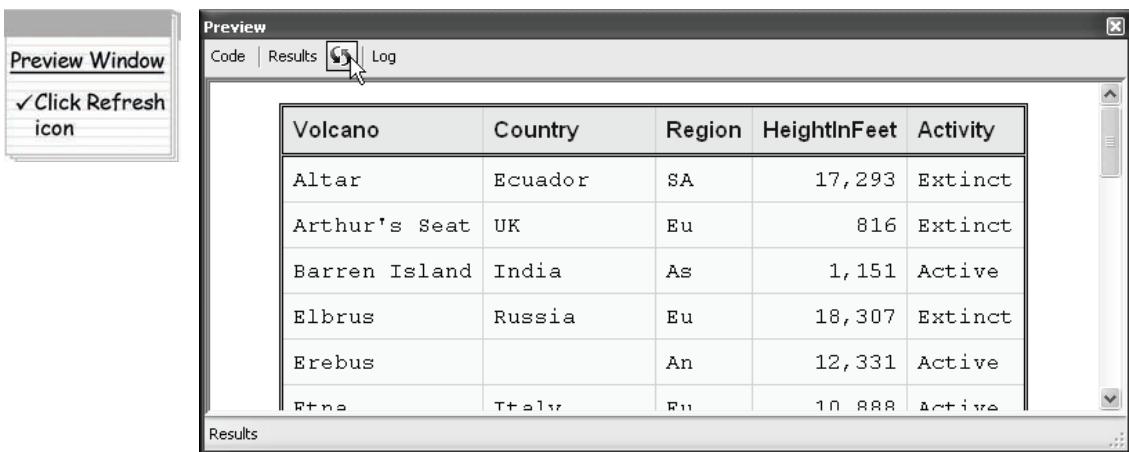


Then click the Delete icon located on the right side of the Query Builder window. It is important to note that deleting the column from the query result does not delete the column from the original data table.

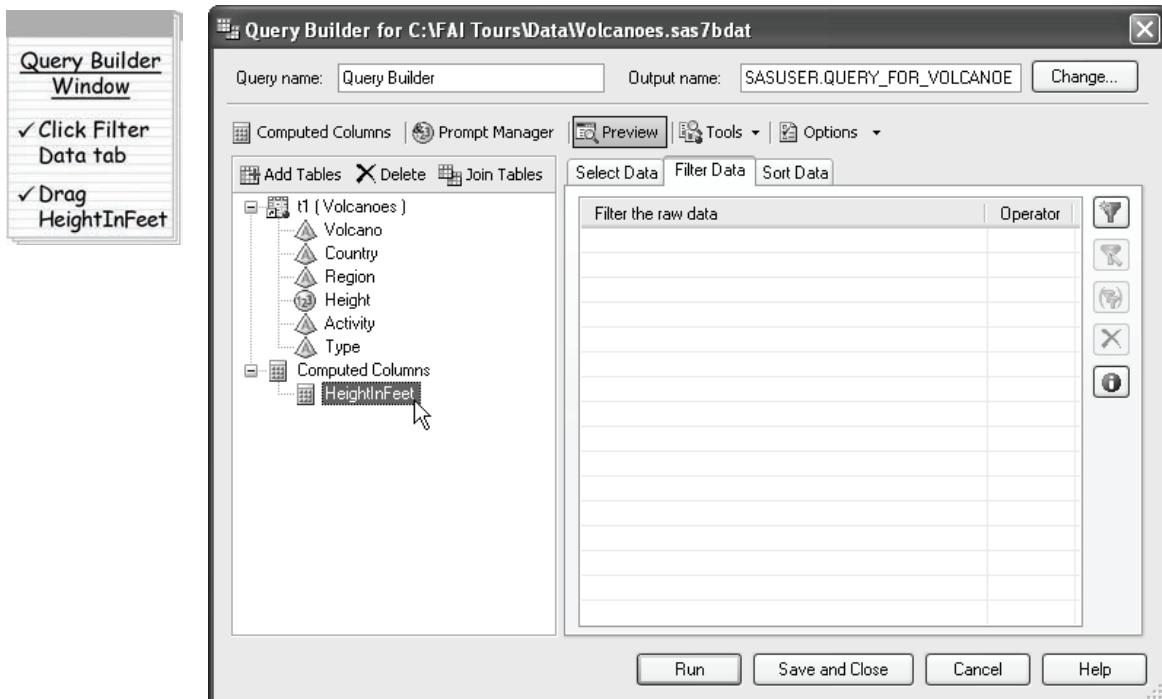
In the results, the columns will be listed in the order that they appear in the Query Builder. You can change the order on the Select Data tab. Click the new column HeightInFeet and then click the up-arrow icon located on the right side of the Query Builder window until the column is listed just above the column Activity.



Click the Refresh icon in the Preview window. Notice that HeightInFeet now appears before Activity and the Height column is no longer in the result.



Filtering data To filter or create subsets of your data, use the Filter Data tab of the Query Builder. Click the **Filter Data** tab to bring it forward. To create a filter, drag the column that you want to use as the basis of your filter over to the Filter Data tab. For this example, use the HeightInFeet column to select only the volcanoes with heights over 12,000 feet.



Click the **HeightInFeet** column on the left side of the window and drag it over to the **Filter Data** tab.

Other Ways to Filter Data

In addition to the Filter Data tab in the Query Builder, you can filter data using the Filter and Sort task accessible by clicking **Filter and Sort** on the workspace toolbar for Data Grids, or by selecting **Tasks ▶ Data ▶ Filter and Sort** from the menu bar. You can also filter data used for tasks by clicking the **Edit** button located near the top of the Data page for tasks. This opens the Edit Data and Filter window where you can define your filter.

As soon as you release the mouse button, the New Filter wizard will open. The column for the filter is automatically set to HeightInFeet, and the Operator is initially set to Equal to.

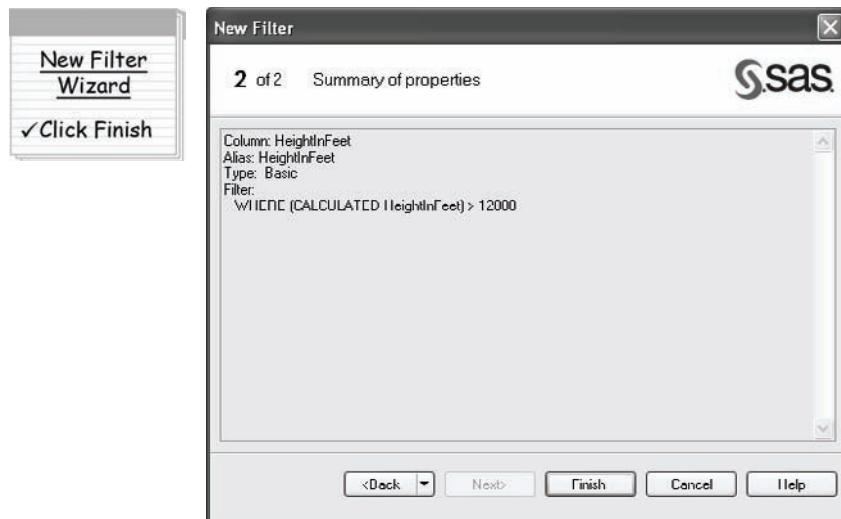


Because you want all volcanoes with a height over 12,000 feet, you need to select a different operator. Click the down-arrow to the right of **Operator** to display the drop-down list of operators. Select the **Greater than** operator. Next, type the value **12000** in the box labeled **Value**. When you enter numeric values, do not enter any commas or dollar signs.

New Filter Wizard

- ✓ From Operator list, select Greater than
- ✓ In Value box, type 12000
- ✓ Click Next

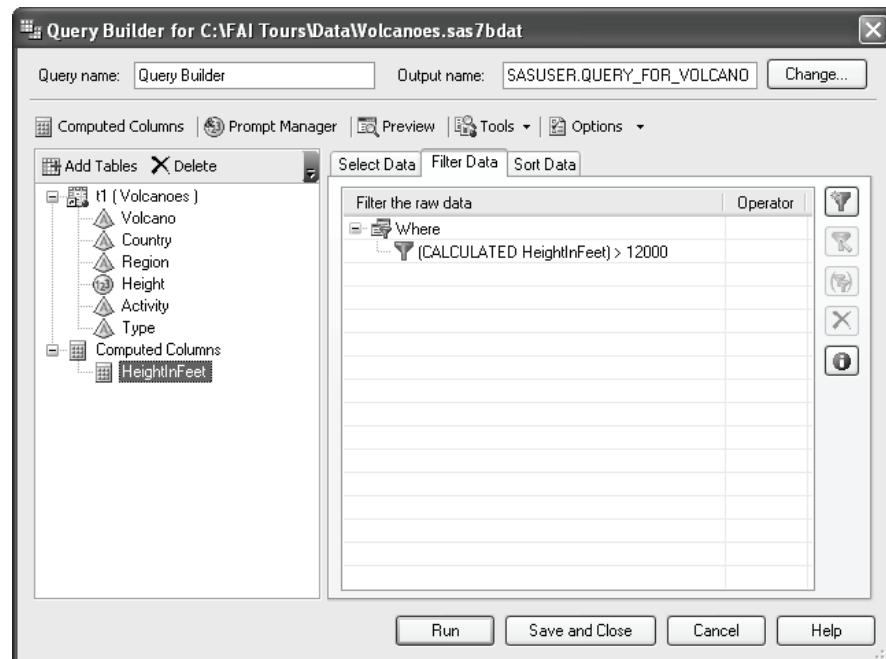
Click **Next** to display a summary of the filter.



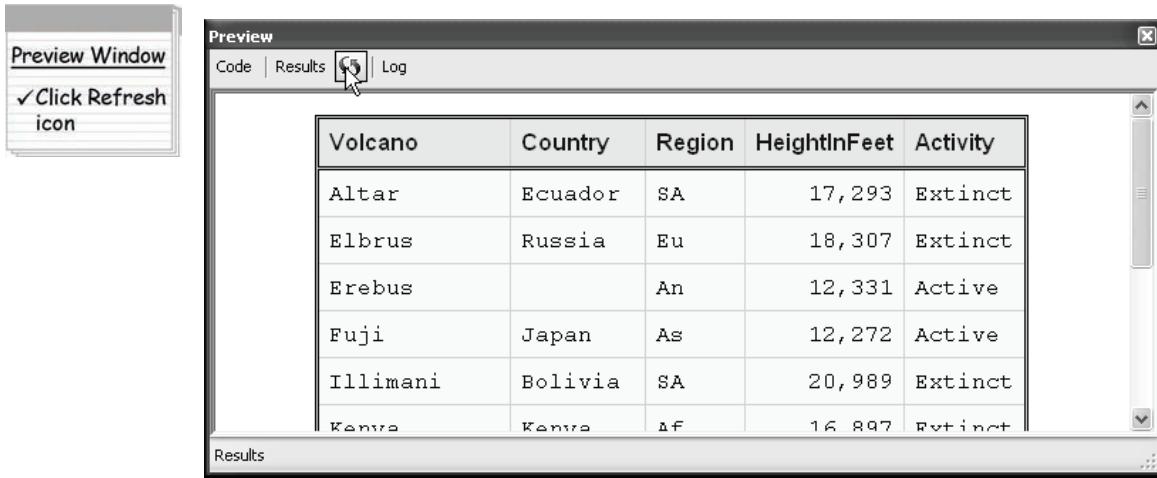
Tutorial C

Click **Finish**.

Now the Filter Data tab shows the filter you created.

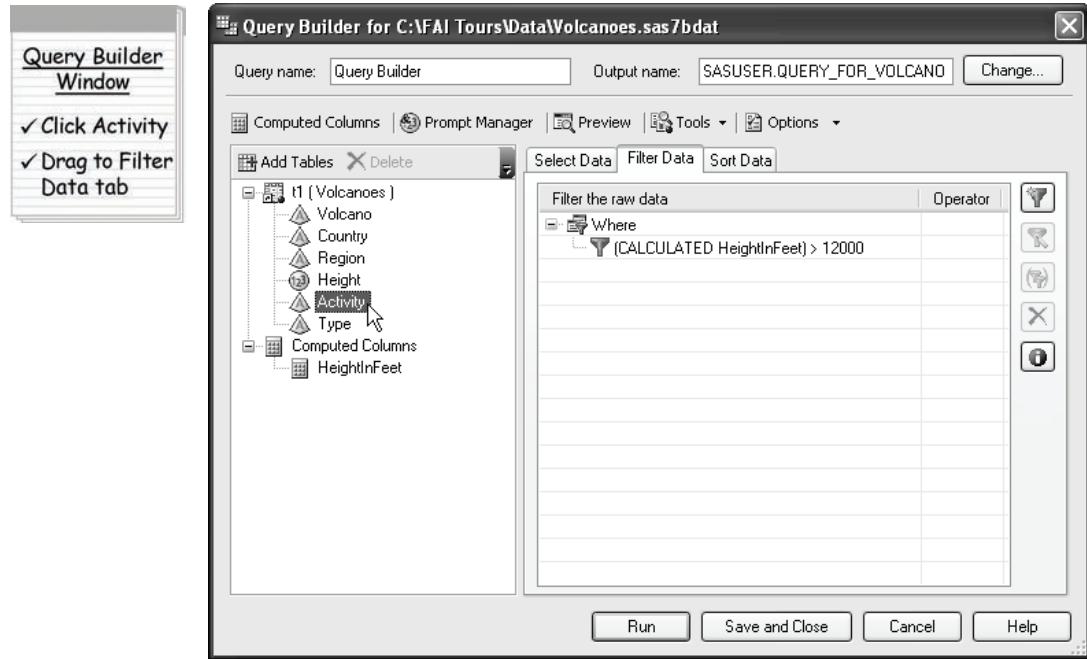


Click the Refresh icon  in the Preview window and see that now all the volcanoes listed are over 12,000 feet.



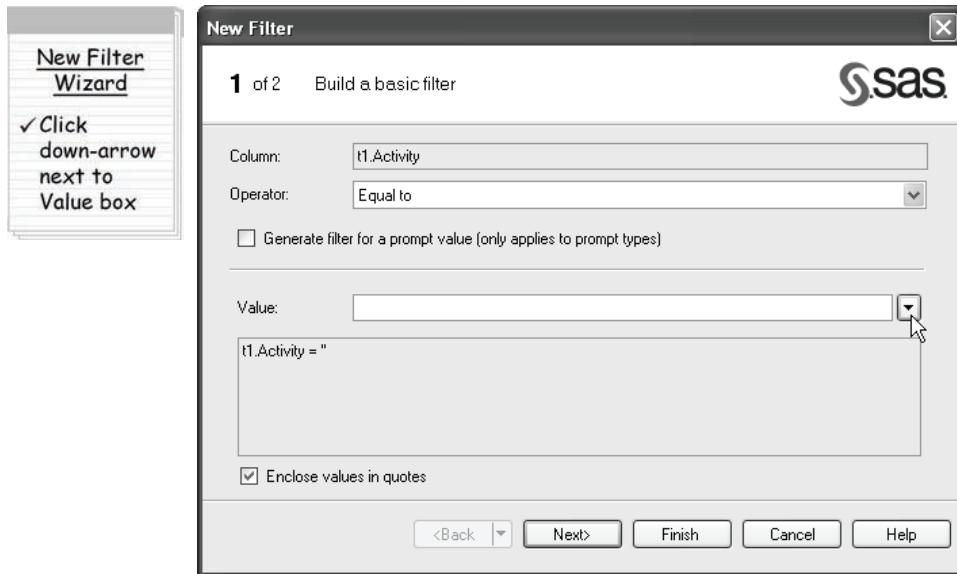
Volcano	Country	Region	HeightInFeet	Activity
Altar	Ecuador	SA	17,293	Extinct
Elbrus	Russia	Eu	18,307	Extinct
Erebus		An	12,331	Active
Fuji	Japan	As	12,272	Active
Illimani	Bolivia	SA	20,989	Extinct
Kaimiro	Kenya	Af	16,897	Extinct

You can create more complicated filters by adding more conditions to your filter. Add the Activity column to the filter to create a data table having only volcanoes that are over 12,000 feet and are active.



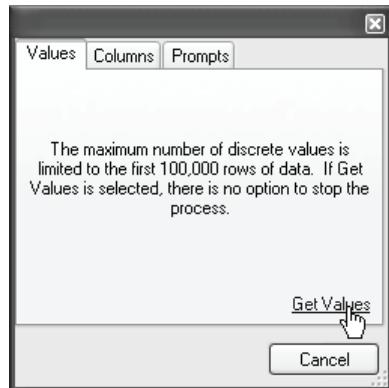
Click the column **Activity** and drag it over to the **Filter Data** tab.

When you release the mouse button, the New Filter wizard will open. This time, there is no need to change the operator because you want all volcanoes where the column Activity equals Active. At this point, you could type the value **Active** (paying attention to the casing of the letters) in the **Value** box to complete the filter. However, SAS Enterprise Guide gives you the option of choosing from a list of values for the column. Choosing the value from a list has the advantage that you can't accidentally misspell the value or use lowercase where it should be uppercase. But use caution if your data tables are very large, as it may take a long time to generate the list of values.



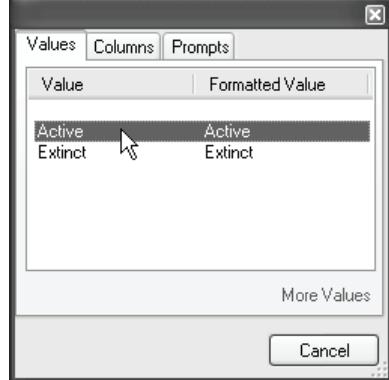
Click the down-arrow next to the **Value** box to open a new window where you can view the values for the Activity column.

Values Tab
✓ Click Get Values



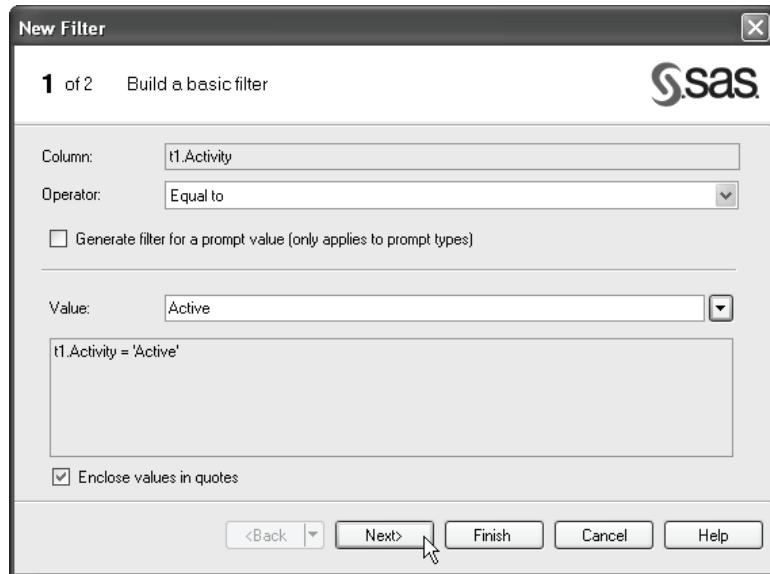
Click **Get Values** to load all possible values for Activity.

Values Tab
✓ Click Active

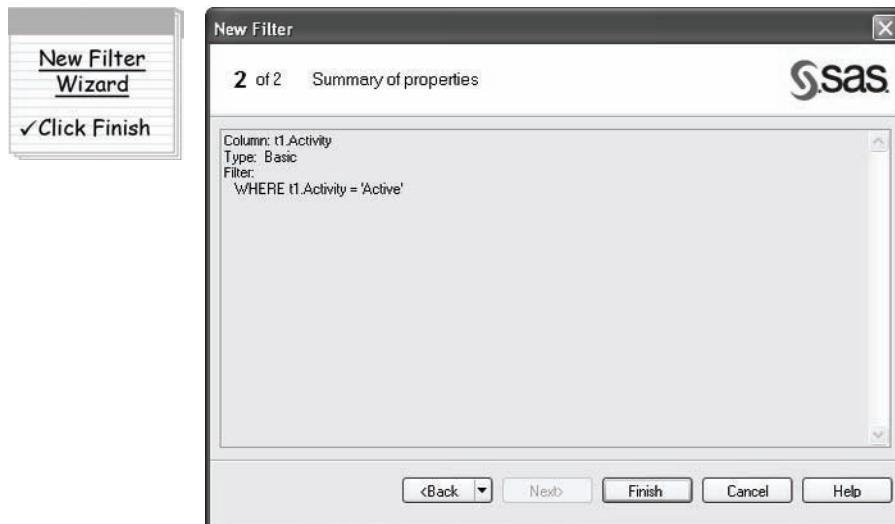


The Activity column has three values: Active, Extinct, and a null or missing value. Click **Active** in the list of values. The window will close, and the value will appear in the Value box of the New Filter window.

New Filter Wizard
✓ Click Next

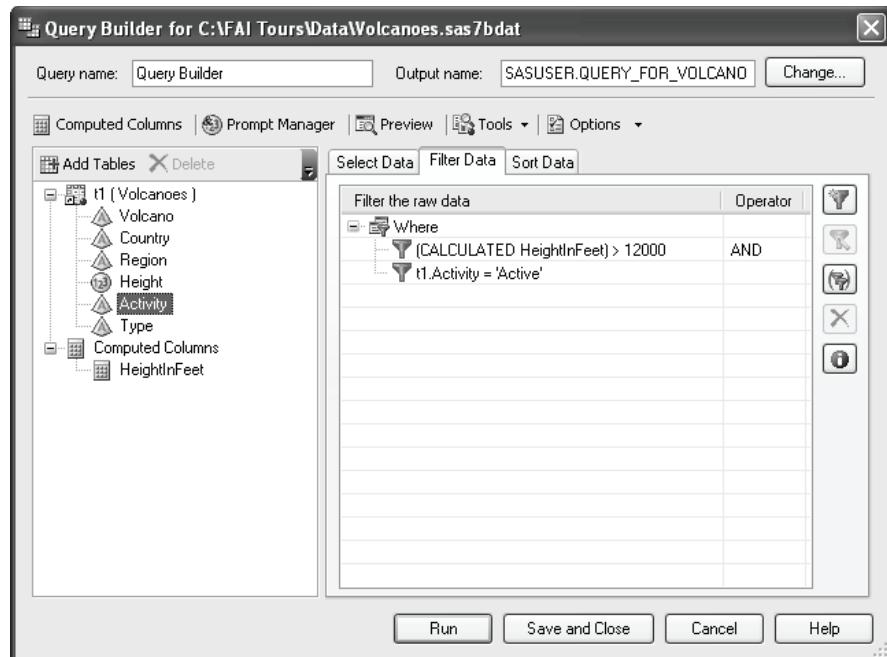


Click **Next** to display a summary of the filter.



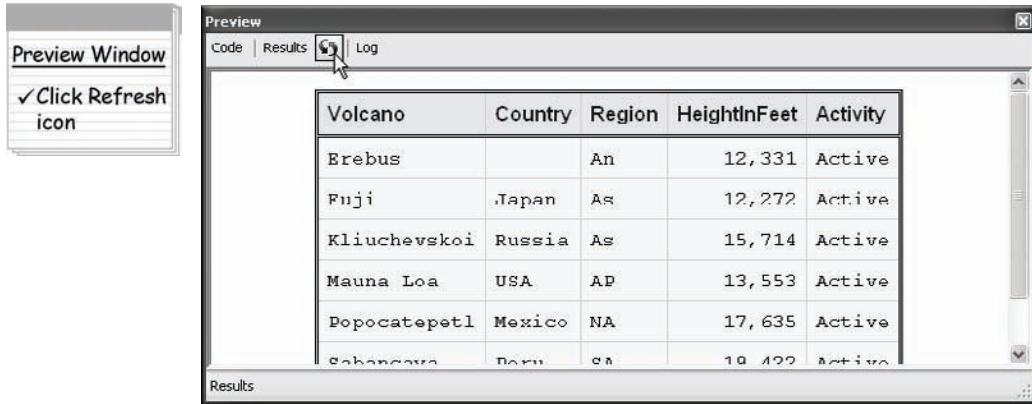
Click **Finish**.

Notice that a new condition has been added to the filter on the Filter Data tab.



If you wanted to change either filter condition, you could simply double-click the condition to reopen the Edit Filter window, or highlight the filter on the Filter Data tab and click the Edit Filter icon  located on the right side of the window.

Now click the Refresh icon  in the Preview window and notice that all the volcanoes listed are over 12,000 feet and are active.



The screenshot shows the 'Preview' window with a toolbar at the top labeled 'Code', 'Results', and 'Log'. Below the toolbar is a table with the following data:

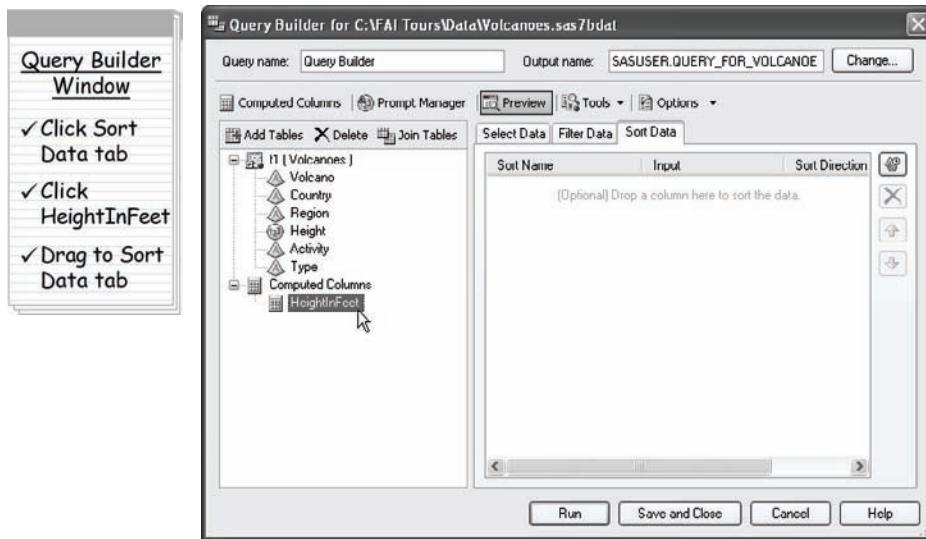
Volcano	Country	Region	HeightInFeet	Activity
Erebus		An	12,331	Active
Fuji	Japan	As	12,272	Active
Kliuchevskoi	Russia	As	15,714	Active
Mauna Loa	USA	AP	13,553	Active
Popocatepetl	Mexico	NA	17,635	Active
Sabancaya	Boru	an	10,499	Active

A note on the left side of the window says 'Click Refresh icon' with a checkmark.

AND or OR?

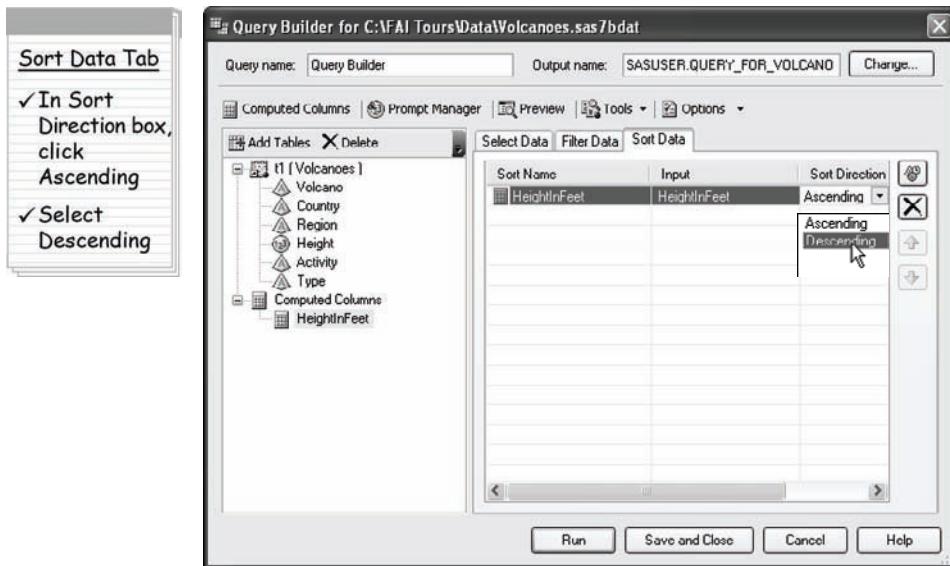
In this example, you want all rows that meet two conditions: volcanoes over 12,000 feet and Active. Because the volcano must pass both conditions, you use the AND operator (the default). But suppose you are not such a thrill seeker, and you would rather just look at volcanoes that are either extinct or less than 8,000 feet. The volcano has to meet only one of the conditions to be included. For this type of filter, use the OR operator. To change the operator from AND to OR, click the AND operator and choose OR from the drop-down list that appears after you click AND.

Sorting the data rows There is one last change to make to this query. The data came sorted alphabetically by the name of the volcano. For this list, it would be better to sort the volcanoes by height, showing the tallest volcano at the top of the list. To sort the data, click the **Sort Data** tab.



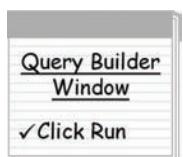
Click and drag the **HeightInFeet** column over to the **Sort Data** tab.

Initially, the sort direction for the column is set to Ascending. To change the sort direction so that the tallest volcano will be first, click **Ascending** and select **Descending** from the drop-down list.



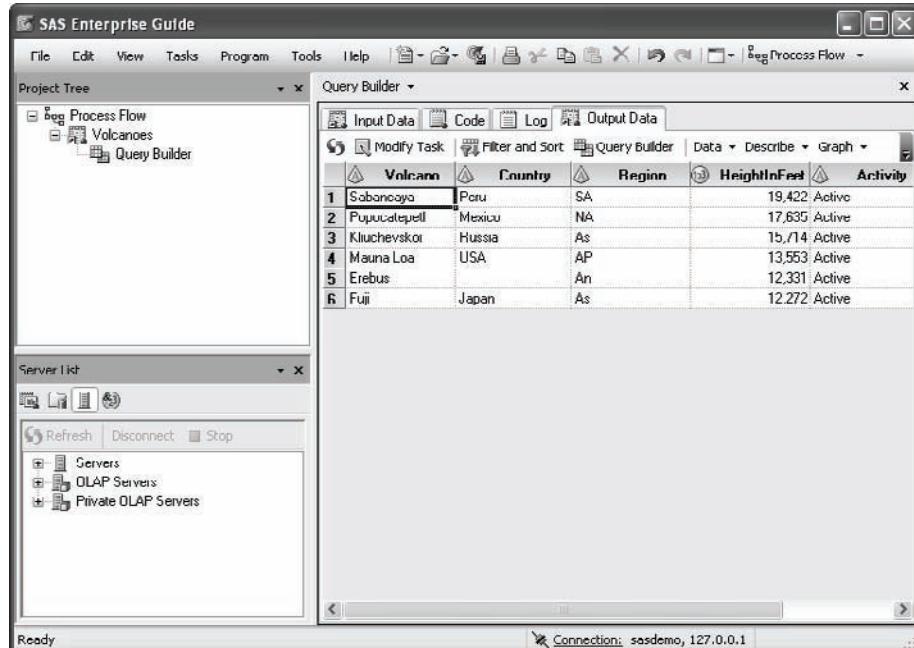
Click the Refresh icon in the Preview window and notice how the volcanoes are now sorted according to height, with the tallest volcano listed first.

Volcano	Country	Region	HeightInFeet	Activity
Sabancaya	Peru	SA	19,422	Active
Popocatepetl	Mexico	NA	17,635	Active
Kliuchevskoi	Russia	As	15,714	Active
Mauna Loa	USA	AP	13,553	Active
Erebus	An		12,331	Active
Fuji	Japan	As	12,272	Active



You have no more changes to make to this query, so click **Run** in the Query Builder window to run the query and close the window.

When you run the query, SAS Enterprise Guide will create a SAS data table by default. The new data table is given a name starting with **Query**, and is stored in a default location. The data table is displayed in the workspace. This data table is now ready for any other tasks that you may want to perform.

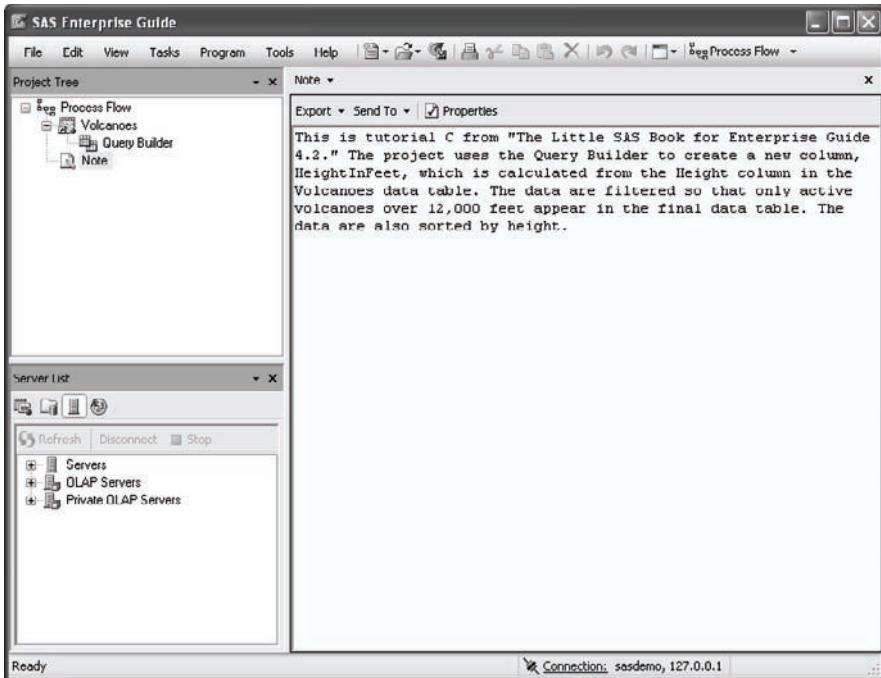


If you want to make any changes to the query, click **Modify Task** on the workspace toolbar for the query result to reopen the Query Builder.

Specifying the Format, Name, and Location for the Results of a Query

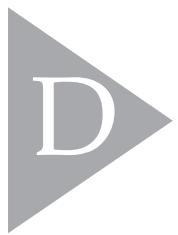
In this example, you are creating a SAS data table as a result of the query. If you want, you can create a report instead. Also, you may want to specify a meaningful name for your results, or store the results in a different location. To change the name and/or storage location for the resulting data table, click the Change button located in the upper right corner of the Query Builder window next to the Output Name. To change the result type of the query from data table to report, select **Options for this query** from the Options drop-down menu in the Query Builder window. On the Results page of the Query Options window that opens, check **Override the corresponding default settings in Tools -> Options**. Then check **Report**.

Completing the tutorial To complete the tutorial, add a note documenting the project. Click the words **Process Flow** in the Project Tree, and then select **File ▶ New ▶ Note** from the menu bar. Type comments about the project into the Note window in the workspace.



Now save the project and exit SAS Enterprise Guide. Select **File ▶ Save Project as** from the menu bar. Navigate to the location where you want to save the project, give the project the name **TutorialC**, and click **Save**. Then select **File ▶ Exit** from the menu bar to close SAS Enterprise Guide.





“One must travel, to learn.”

MARK TWAIN

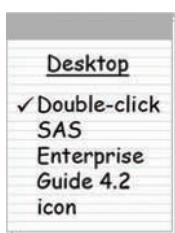
D ► Joining Two Data Tables Together

Often the data you need for a particular analysis are in more than one table. To perform the analysis, you need to join tables together. In this tutorial, you will be joining together two data tables, and then manipulating the data after the join. Here are the topics covered in this tutorial:

- Joining two tables together
- Filtering data after the join
- Selecting which data rows to keep

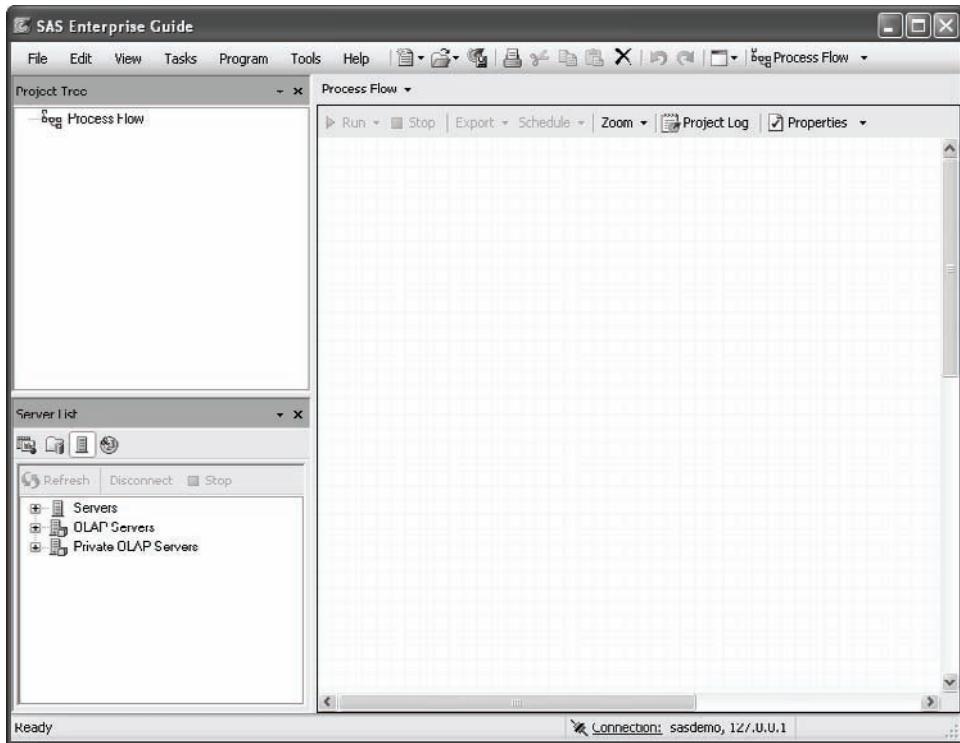
Before beginning this tutorial This tutorial uses the Volcanoes SAS data table, which contains information about volcanoes around the world. This tutorial also uses the Tours data table, which contains information about the volcano tours offered by the Fire and Ice Tours company. The data and instructions for downloading the data tables can be found in Appendix A.

The Fire and Ice Tours company wants to produce a list of tours for all volcanoes in Europe. The problem is that the Tours data set does not contain information about the region of the volcano. The region of the volcano is contained in the Volcanoes file. So, for the company to produce the desired list, the Volcanoes data and the Tours data must be joined together.



Starting SAS Enterprise Guide Start SAS Enterprise Guide by either double-clicking the **SAS Enterprise Guide 4.2** icon on your desktop, or selecting **SAS Enterprise Guide 4.2** from the Windows **Start** menu. Starting SAS Enterprise Guide brings up the SAS Enterprise Guide windows in the background, with the Welcome window in the foreground. The Welcome window allows you to choose between opening an existing project or starting a new project. Click **New Project**.

This opens an empty SAS Enterprise Guide window.



Tutorial D



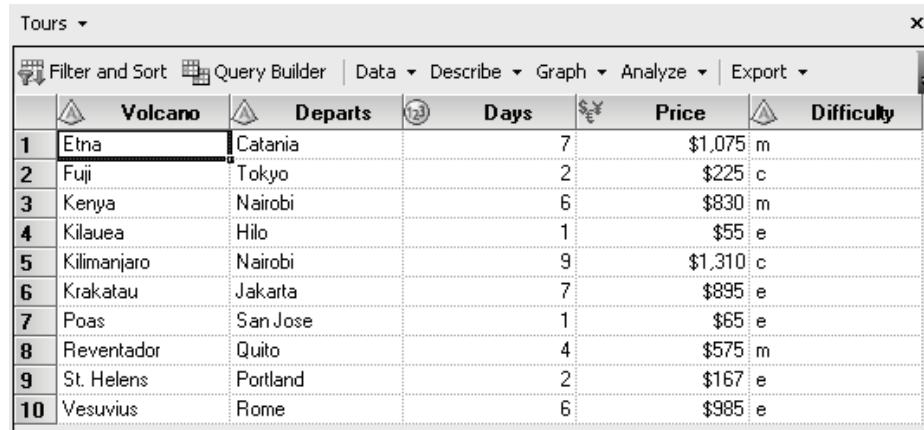
Opening the two data files to be joined Open the Volcanoes and Tours data tables by selecting **File ▶ Open ▶ Data** from the menu bar, clicking **Local Computer**, and navigating to the location where you saved the data for this book. You can select both tables at once by clicking one table, holding the control (CTRL) key down and then clicking the other table. Click **Open**. (If you created the Tours data table in Tutorial A and saved it in the SASUSER library, you may prefer to open your version instead by selecting **File ▶ Open ▶ Data** from the menu bar, clicking **Servers**, and navigating to the SASUSER library.)

After opening both files, your SAS Enterprise Guide window should look like the following.

The screenshot shows the SAS Enterprise Guide application window. On the left, there is a 'Project Tree' pane containing a 'Process Flow' node, a 'Volcanoes' node, and a 'Tours' node. Below it is a 'Server List' pane showing a connection to 'sasdemo, 127.0.0.1'. The main area displays a data table titled 'Tours' with the following columns: Volcano, Departs, Days, Price, and Difficulty. The data consists of 10 rows, numbered 1 to 10, listing various火山 (Volcanoes) and their tour details:

	Volcano	Departs	Days	Price	Difficulty
1	Etna	Catania	7	\$1,075 m	m
2	Fuji	Tokyo	2	\$225 c	c
3	Kenya	Nairobi	6	\$830 m	m
4	Kilauea	Hilo	1	\$55 e	e
5	Kilimanjaro	Nairobi	9	\$1,310 c	c
6	Krakatau	Jakarta	7	\$895 e	e
7	Poas	San Jose	1	\$65 e	e
8	Reventador	Quito	4	\$575 m	m
9	St. Helens	Portland	2	\$167 e	e
10	Vesuvius	Rome	6	\$965 e	e

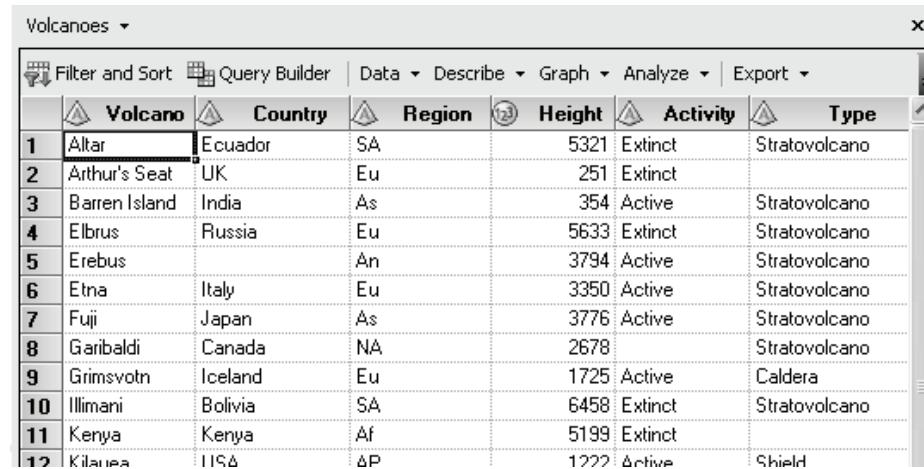
Here is what the Tours data table looks like. The Tours data contain information about the tour: the name of the volcano, the city where the tour departs, the number of days the tour lasts, the price, and a difficulty rating for the tour. The region of the volcano is not part of the Tours data table, so it would not be possible, using this data table alone, to produce a list of volcano tours in Europe.



A screenshot of the SAS Enterprise Guide interface showing the Tours data table. The table has columns: Volcano, Departs, Days, Price, and Difficulty. The rows are numbered 1 through 10. The first row, Etna, is selected. The data is as follows:

	Volcano	Departs	Days	Price	Difficulty
1	Etna	Catania	7	\$1,075	m
2	Fuji	Tokyo	2	\$225	c
3	Kenya	Nairobi	6	\$830	m
4	Kilauea	Hilo	1	\$55	e
5	Kilimanjaro	Nairobi	9	\$1,310	c
6	Krakatau	Jakarta	7	\$895	e
7	Poas	San Jose	1	\$65	e
8	Reventador	Quito	4	\$575	m
9	St. Helens	Portland	2	\$167	e
10	Vesuvius	Rome	6	\$985	e

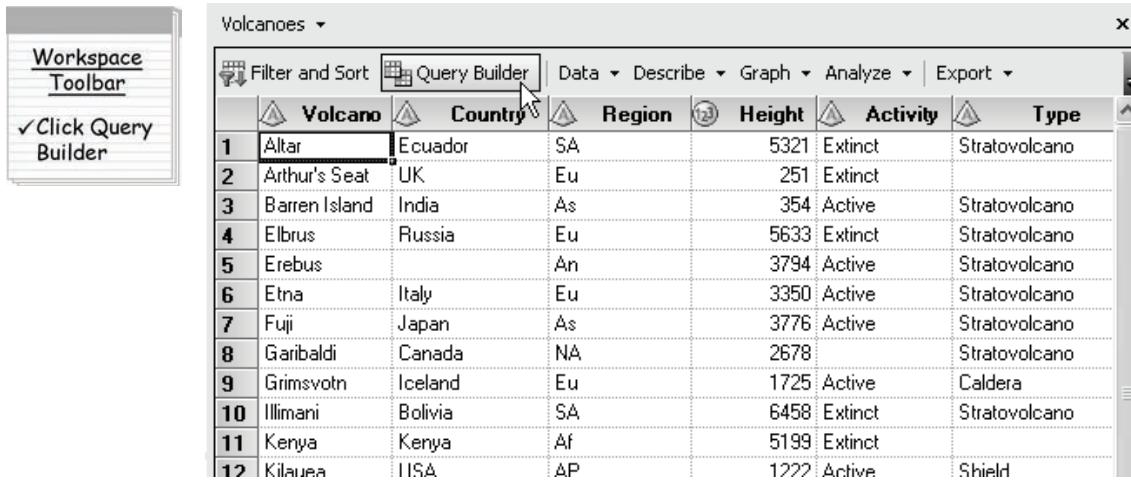
Here is a partial listing of the Volcanoes data table. This table includes the country and region of the volcano, as well as the height, activity, and type of volcano. While the two data tables contain different information, they do have one column in common, the name of the volcano. To join data tables together in a meaningful way, the tables must have at least one column that appears in both data tables. The common column does not have to have the same name in both data tables, but it must contain the same information and have the same possible values.



A screenshot of the SAS Enterprise Guide interface showing the Volcanoes data table. The table has columns: Volcano, Country, Region, Height, Activity, and Type. The rows are numbered 1 through 12. The first row, Altar, is selected. The data is as follows:

	Volcano	Country	Region	Height	Activity	Type
1	Altar	Ecuador	SA	5321	Extinct	Stratovolcano
2	Arthur's Seat	UK	Eu	251	Extinct	
3	Barren Island	India	As	354	Active	Stratovolcano
4	Elbrus	Russia	Eu	5633	Extinct	Stratovolcano
5	Erebus		An	3794	Active	Stratovolcano
6	Etna	Italy	Eu	3350	Active	Stratovolcano
7	Fuji	Japan	As	3776	Active	Stratovolcano
8	Garibaldi	Canada	NA	2678		Stratovolcano
9	Grimsvotn	Iceland	Eu	1725	Active	Caldera
10	Illimani	Bolivia	SA	6458	Extinct	Stratovolcano
11	Kenya	Kenya	Af	5199	Extinct	
12	Kilanea	Hawa	AP	1222	Active	Shield

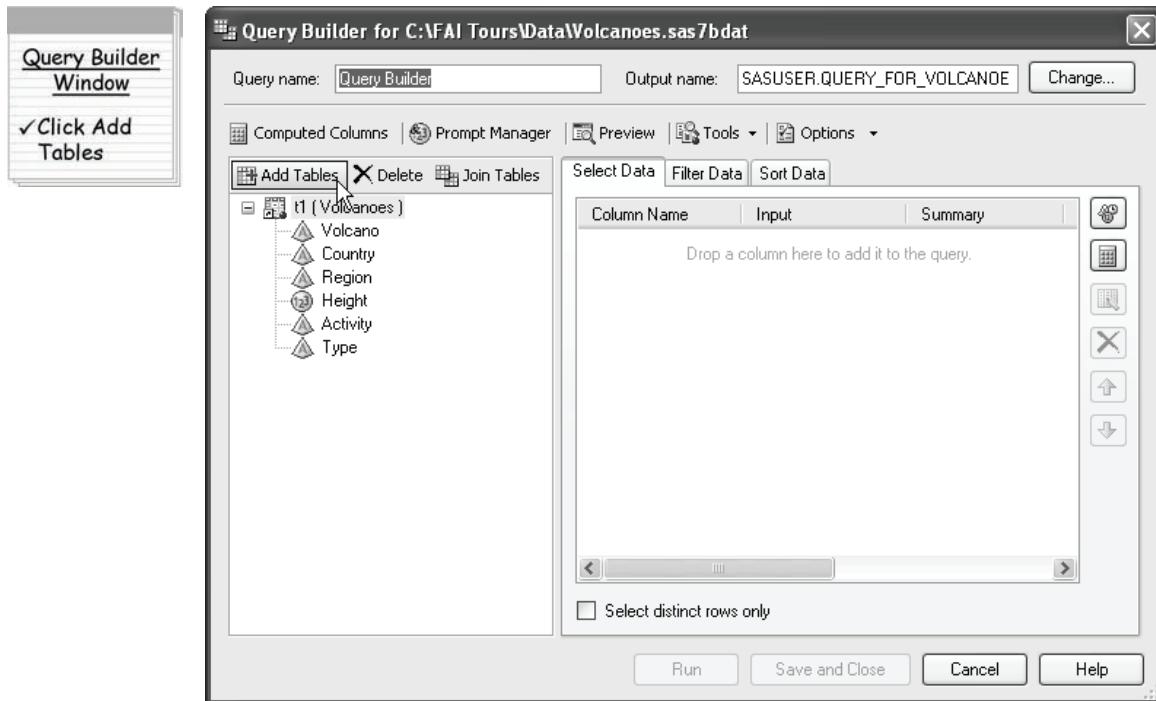
Joining tables You join tables using the Query Builder. In the Query Builder, in addition to joining tables, you can filter and sort data, and create new columns. To open the Query Builder, click **Query Builder** on the workspace toolbar for the Volcanoes data table. Or, you can right-click the **Volcanoes** data icon in the Project Tree or Process Flow and select **Query Builder**.



The screenshot shows the Oracle SQL Developer interface. On the left, there's a 'Workspace Toolbar' with a 'Query Builder' button. A tooltip 'Click Query Builder' is displayed over this button. The main area is titled 'Volcanoes'. It contains a table with the following data:

	Volcano	Country	Region	Height	Activity	Type
1	Altar	Ecuador	SA	5321	Extinct	Stratovolcano
2	Arthur's Seat	UK	Eu	251	Extinct	
3	Baren Island	India	As	354	Active	Stratovolcano
4	Elbrus	Russia	Eu	5633	Extinct	Stratovolcano
5	Erebus		An	3794	Active	Stratovolcano
6	Etna	Italy	Eu	3350	Active	Stratovolcano
7	Fuji	Japan	As	3776	Active	Stratovolcano
8	Garibaldi	Canada	NA	2678		Stratovolcano
9	Grimsvotn	Iceland	Eu	1725	Active	Caldera
10	Illimani	Bolivia	SA	6458	Extinct	Stratovolcano
11	Kenya	Kenya	Af	5199	Extinct	
12	Kilauea	USA	AP	1222	Active	Shield

The Query Builder window has three tabs: Select Data, Filter Data, and Sort Data. In addition, in the box on the left side of the Query Builder, you can choose: Add Tables, Delete, and Join Tables. The name of the data table appears in the list on the left, along with all the columns in the data table.

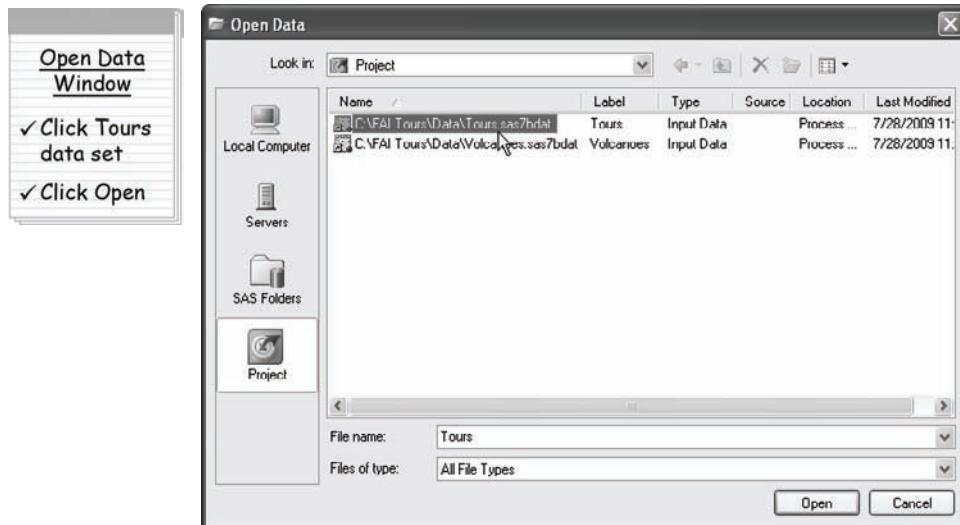


Even though both the Tours and the Volcanoes data tables are in the project, only the Volcanoes table is in the current query. In order to join the two tables together, you must add the Tours table to the query. Click **Add Tables**.

In the Open Data window, you can choose to add data from your Local Computer, Servers, SAS Folders, or the Project. Both data tables are open in your project, so click Project in the selection pane on the left.

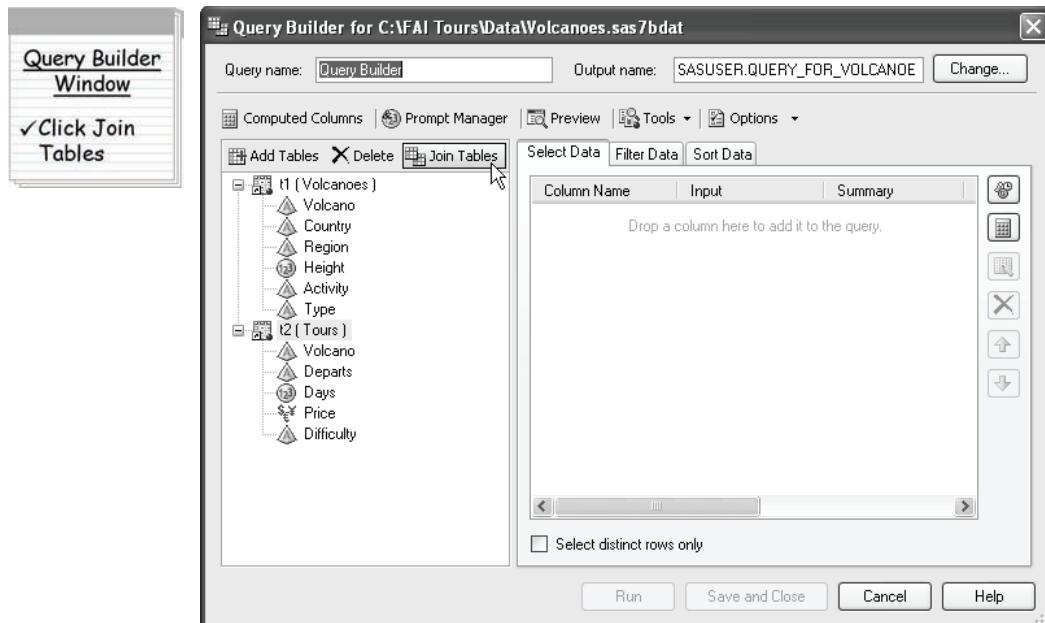


All the tables open in the project are listed in the Open Data window. The name includes the full path for the data and the extension for SAS data tables (sas7bdat).



Add the Tours table to the query by clicking **Tours**, and then clicking **Open**.

When you add data tables to the query, SAS Enterprise Guide checks to see if the data tables have a column in common. Common columns must have the same name and be the same type (numeric or character). If there is a common column, SAS Enterprise Guide will automatically use that column for the join. Because the Volcanoes data table and the Tours data table both have a character column named Volcano, SAS Enterprise Guide will use it to find matching rows.

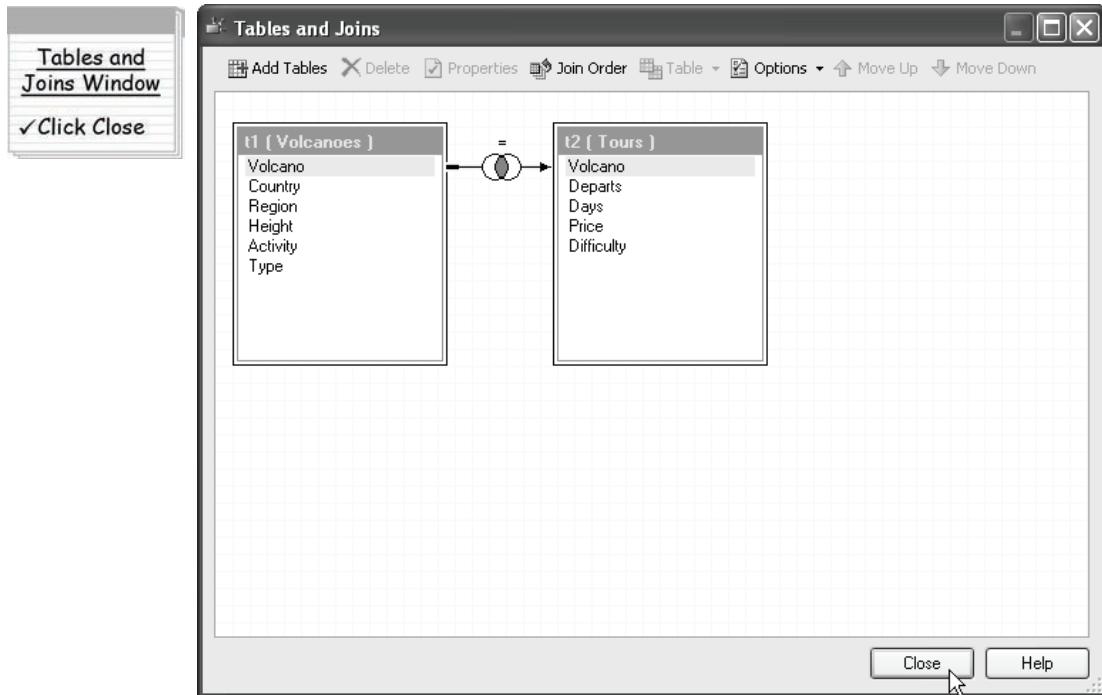


It's not obvious from the Query Builder window that SAS Enterprise Guide has found common columns for the join. To see what the join looks like, click **Join Tables** to open the Tables and Joins window.

What If the Common Columns Have Different Names?

No problem! When you try to join the two tables, SAS Enterprise Guide will let you know that it cannot determine how to join them and that you will need to do it manually. Click **OK**, and the Tables and Joins window will open automatically. Click the first table, then right-click the common column, select **Join With** from the pop-up menu, and select the column from the second table. Next, choose a join type from the Join Properties window, and click **OK**. SAS Enterprise Guide will draw a line between the two columns, and the columns will be linked.

In the Tables and Joins window, both tables are visible along with the columns in the tables. Notice the line drawn between the Volcano column in the Volcanoes table and the Volcano column in the Tours table. This shows how the tables will be joined. Also, there is a diagram on the line that shows the type of join. In this case, only rows found in both tables will be included in the resulting table.

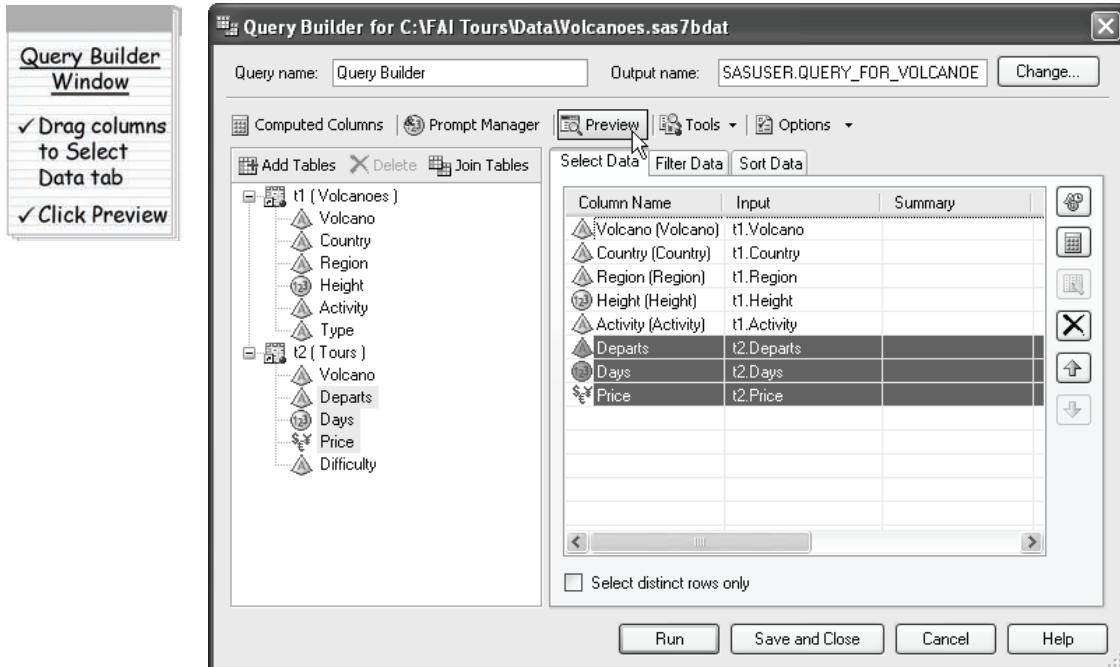


Click Close to close the Tables and Joins window and return to the Query Builder.

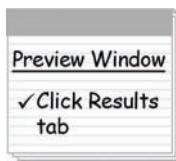
Tables with More Than One Common Column

It is possible to join data tables that have more than one common column. For example, you may have year and month columns in both data tables, and you want to match the tables based on the values of both columns. SAS Enterprise Guide does not handle this type of join automatically, but it is easy to do it yourself. SAS Enterprise Guide will link the first pair of columns for you. To link the second pair, open the Tables and Joins window by clicking **Join Tables** in the Query Builder window. Create another link by clicking the first table, then right-clicking the column name, choosing **Join With** in the pop-up menu, and selecting the column from the second table.

Before you can run the query, you must select the columns you want in the result. To select columns, click the column in the box on the left and drag it over to the Select Data tab. For this query, select the **Volcano**, **Country**, **Region**, **Height**, and **Activity** columns from the Volcanoes table. Then select the **Departs**, **Days**, and **Price** columns from the Tours table.



To get an idea of what the results of the join will look like, click the **Preview** button near the top of the Query Builder window. Then click the **Results** tab in the Preview window.

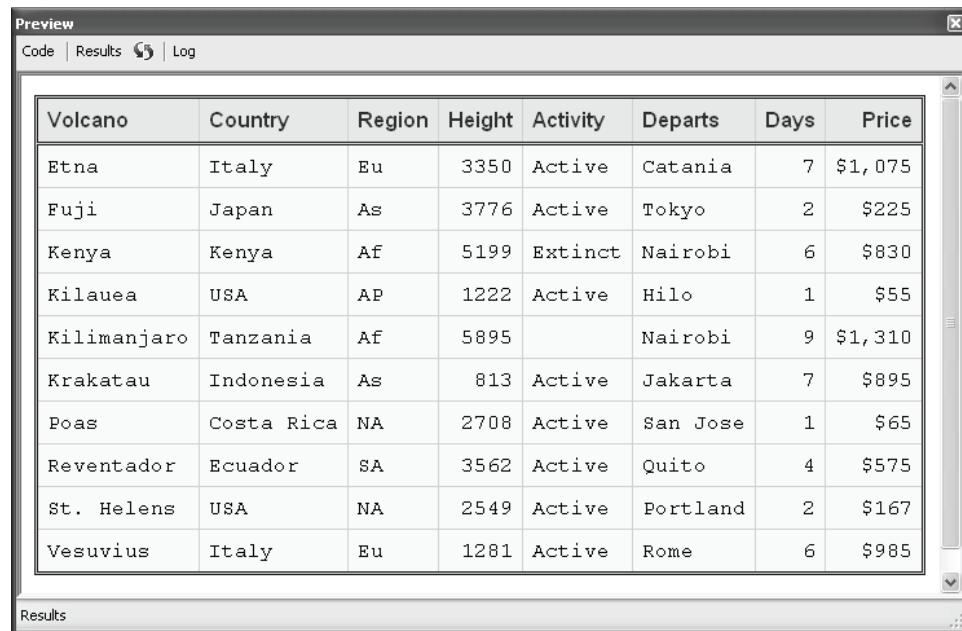


What If One Table Has Repeated Values of the Common Column?

In this example, each table has only one entry for each volcano, so there is a one-to-one match between the tables. But suppose the Tours table had two tours for one volcano. Then the values for the columns in the Volcanoes data table will be repeated for these two tours for the same volcano.

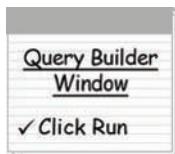
When you preview the results of a query, usually you will not see the complete result. The Preview window is designed to show you what to expect from your join without having to complete the entire join. With these small data tables, it wouldn't take long to perform the join, so the Preview window does not save you much time. But if you have very large data tables, it can take a long time for SAS Enterprise Guide to complete the join, so the Preview window allows you to make sure the query is correct before starting the join.

Notice that the volcanoes from the Volcanoes data table have been matched with the corresponding row from the Tours data table.



The screenshot shows the SAS Enterprise Guide Preview window. The title bar says "Preview". Below it is a menu bar with "Code" and "Results" selected. A toolbar icon is next to the menu. The main area contains a table with the following data:

Volcano	Country	Region	Height	Activity	Departs	Days	Price
Etna	Italy	Eu	3350	Active	Catania	7	\$1,075
Fuji	Japan	As	3776	Active	Tokyo	2	\$225
Kenya	Kenya	Af	5199	Extinct	Nairobi	6	\$830
Kilauea	USA	AP	1222	Active	Hilo	1	\$55
Kilimanjaro	Tanzania	Af	5895		Nairobi	9	\$1,310
Krakatau	Indonesia	As	813	Active	Jakarta	7	\$895
Poas	Costa Rica	NA	2708	Active	San Jose	1	\$65
Reventador	Ecuador	SA	3562	Active	Quito	4	\$575
St. Helens	USA	NA	2549	Active	Portland	2	\$167
Vesuvius	Italy	Eu	1281	Active	Rome	6	\$985



Click **Run** in the Query Builder window to run the query. The data table created by the query will display in the workspace. Notice that only the rows that appear in both data tables are part of the result. This is the default type of join for SAS Enterprise Guide. Because all the volcanoes in the Tours data table also appear in the Volcanoes data table, all the tours are represented here. However, some volcanoes do not have matching tours, and so they are not included in the result.

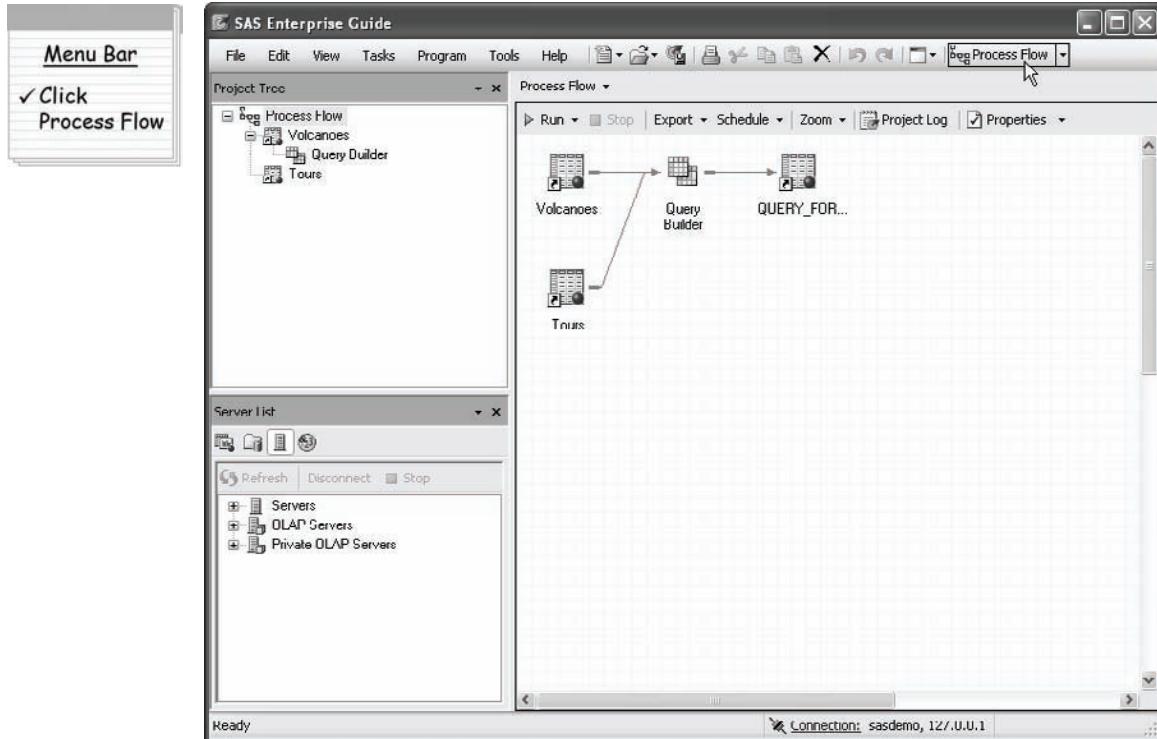
Query Builder ▾ x

Input Data (2) Code Log Output Data

Modify Task Filter and Sort Query Builder Data ▾ Describe ▾ Graph ▾

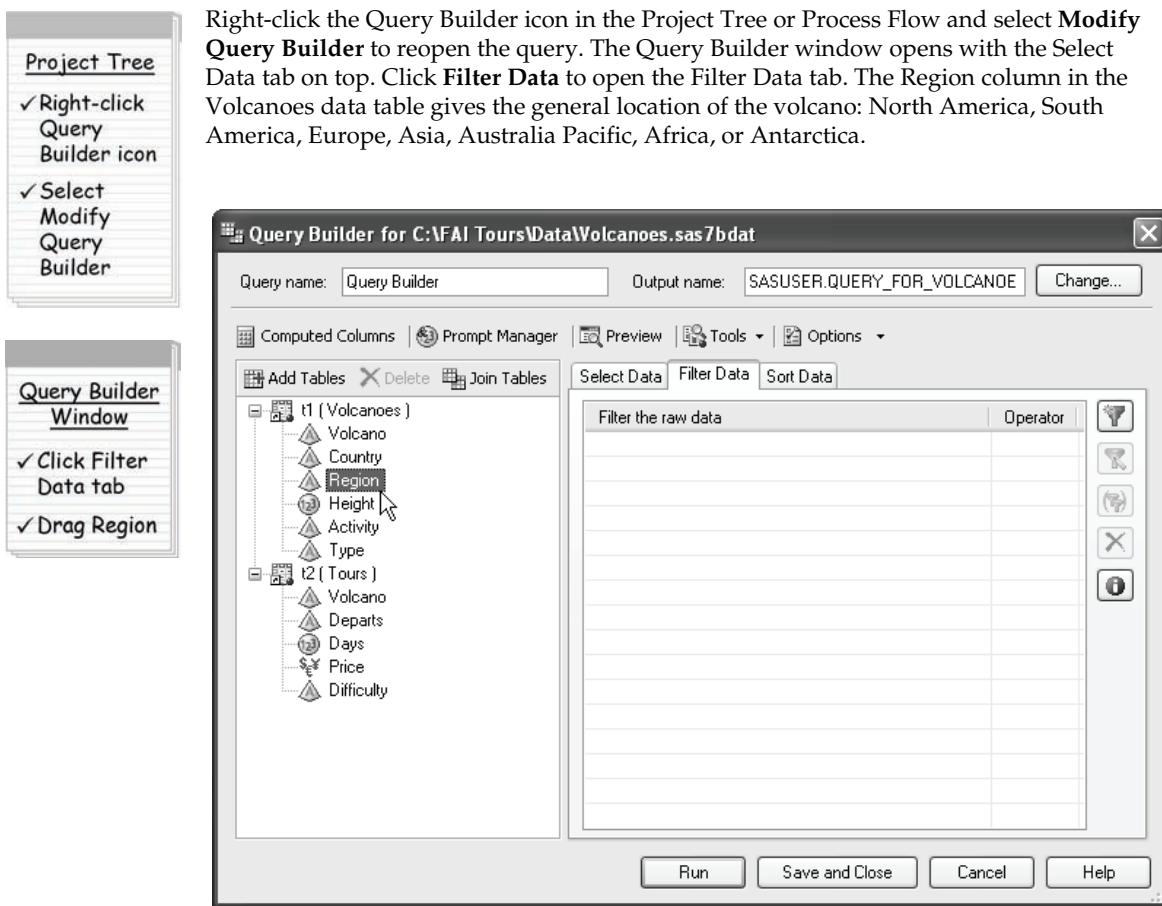
	Volcano	Country	Region	Height	Activity	Departs	Days	Price
1	Etna	Italy	Eu	3350	Active	Catania	7	\$1,075
2	Fuji	Japan	As	3776	Active	Tokyo	2	\$225
3	Kenya	Kenya	Af	5199	Extinct	Nairobi	6	\$830
4	Kilauea	USA	AP	1222	Active	Hilo	1	\$55
5	Kilimanjaro	Tanzania	Af	5895		Nairobi	9	\$1,310
6	Krakatau	Indonesia	As	813	Active	Jakarta	7	\$895
7	Poas	Costa Rica	NA	2708	Active	San Jose	1	\$65
8	Reventador	Ecuador	SA	3562	Active	Quito	4	\$575
9	St. Helens	USA	NA	2549	Active	Portland	2	\$167
10	Vesuvius	Italy	Eu	1281	Active	Rome	6	\$985

Click the **Process Flow** button to view the Process Flow for the project. It is not obvious from the Project Tree that the query uses both the Tours and the Volcanoes tables, but in the Process Flow it is easy to see that both tables contribute to the query. By default, results of queries are given an arbitrary name starting with the letters QUERY and are stored in a default location.



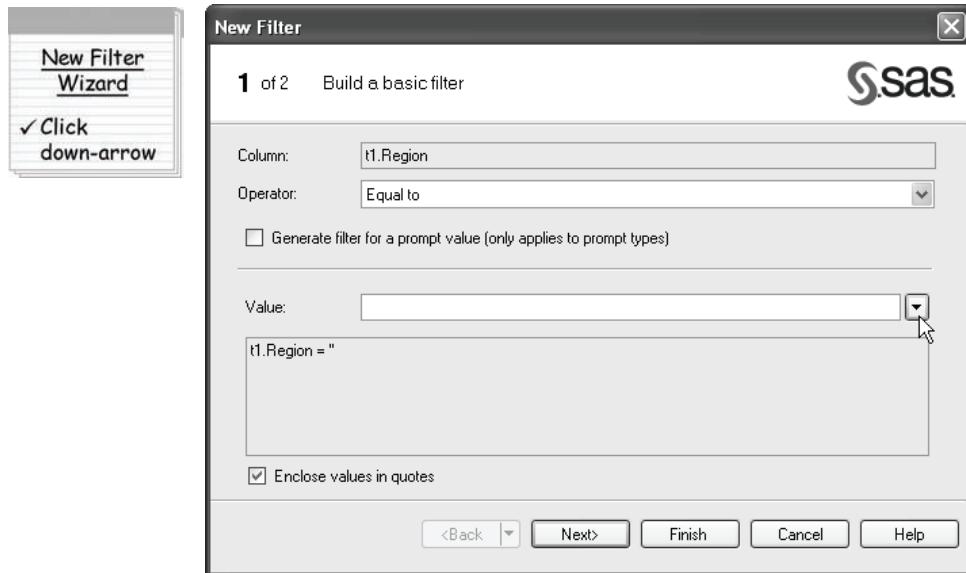
Filtering the data Now that the Tours data table and the Volcanoes data table have been joined together, it is possible to create a data table of tours for volcanoes in Europe. To do this, you create a filter as part of the same query that you used to join the data tables.

Right-click the Query Builder icon in the Project Tree or Process Flow and select **Modify Query Builder** to reopen the query. The Query Builder window opens with the Select Data tab on top. Click **Filter Data** to open the Filter Data tab. The Region column in the Volcanoes data table gives the general location of the volcano: North America, South America, Europe, Asia, Australia Pacific, Africa, or Antarctica.

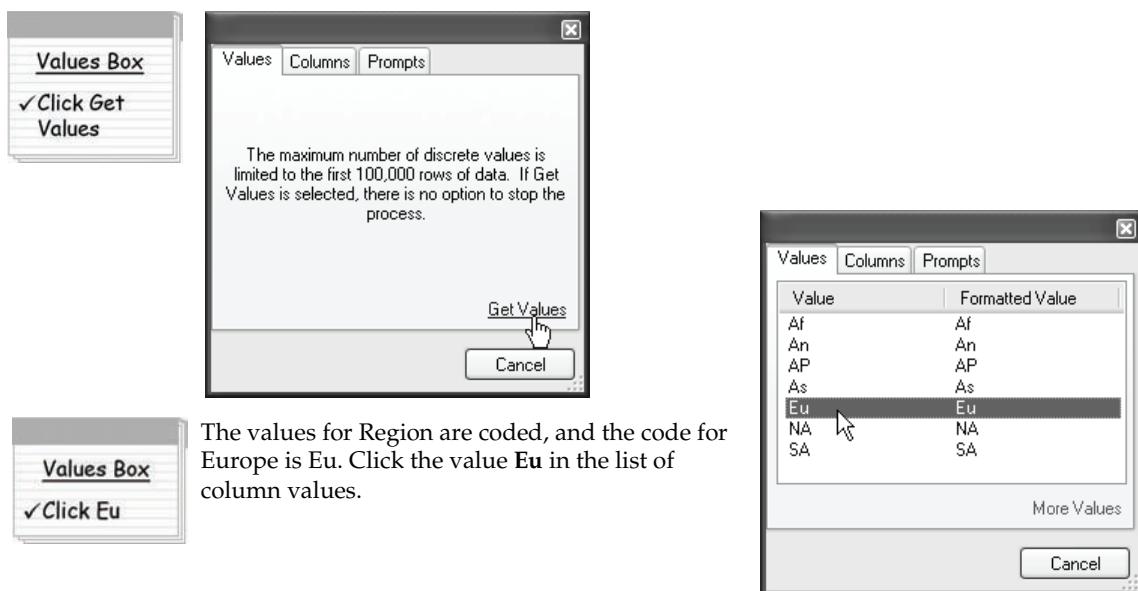


To create a filter based on the Region column, click **Region** in the list on the left, and drag it to the **Filter Data** tab.

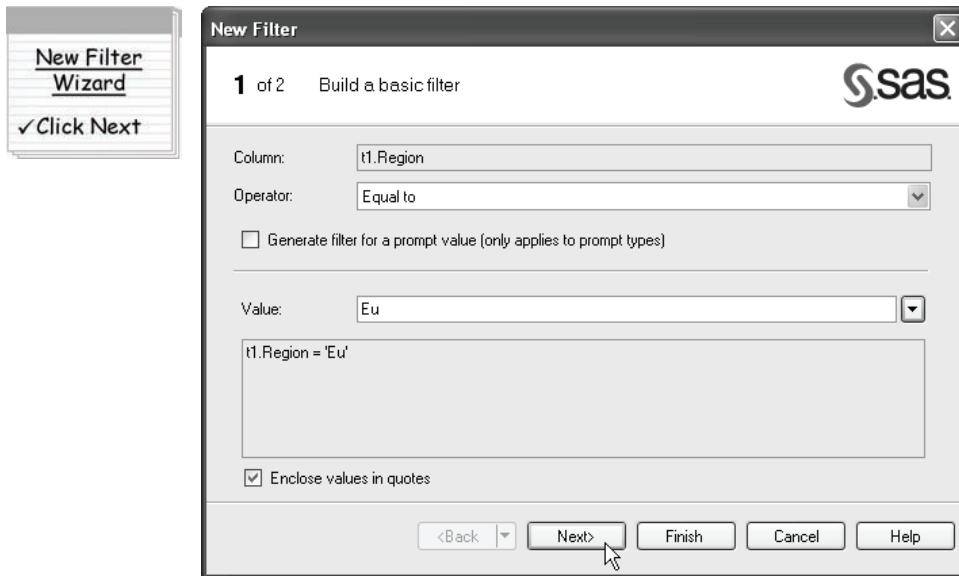
This automatically opens the New Filter wizard with Region from the Volcanoes data table (table 1) listed as the Column.



Click the down-arrow next to the **Value** box to open a window where you can display all possible values for Region. Click **Get Values**.



After you select Eu from the list of values, the window will close and Eu will appear in the box next to Value in the New Filter wizard.

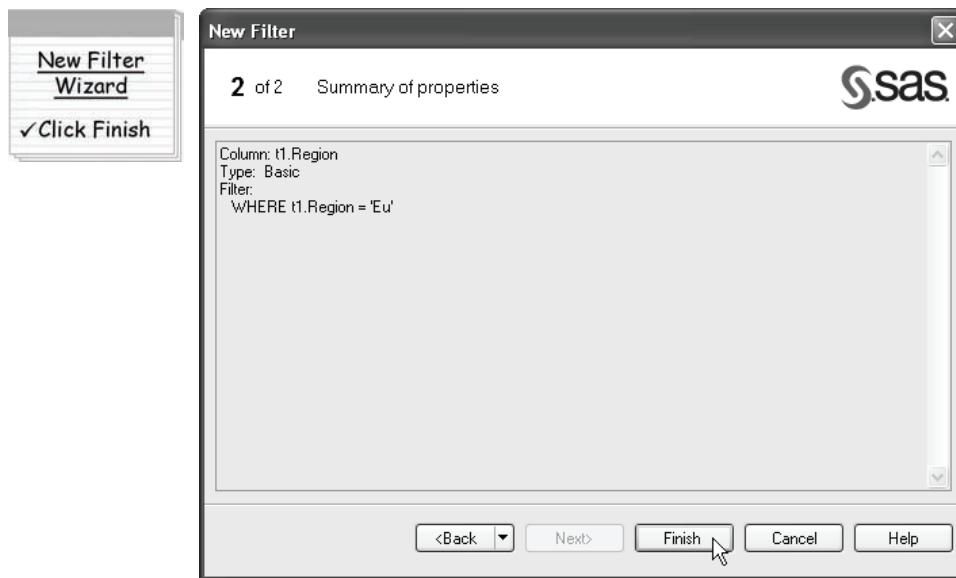


Click **Next**.

Choosing an Operator

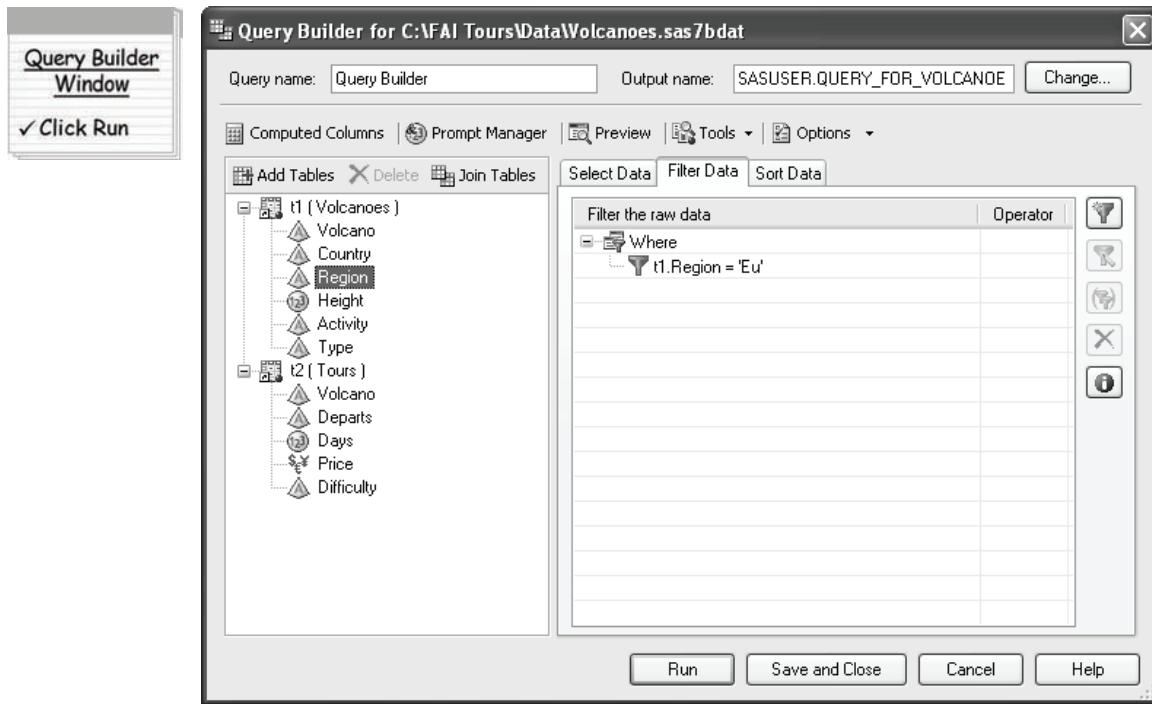
In this example, you want all the volcanoes whose region equals Eu, so you use the **Equal to** operator. The **Equal to** operator is the default operator, so you didn't need to change it. But you can change it if you want. To choose a different operator, click the down-arrow in the **Operator** box, and choose from the list. For example, if you wanted all volcanoes in North and South America, you would use the **In a list** operator. The **In a list** operator selects all rows in a table whose values for the column are contained in the specified list.

The final window of the New Filter wizard shows a summary of the filter you just created.



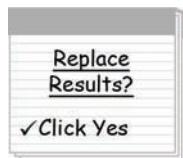
Click **Finish** to complete the filter condition and return to the Query Builder.

Now the filter appears on the Filter Data tab of the Query Builder window.

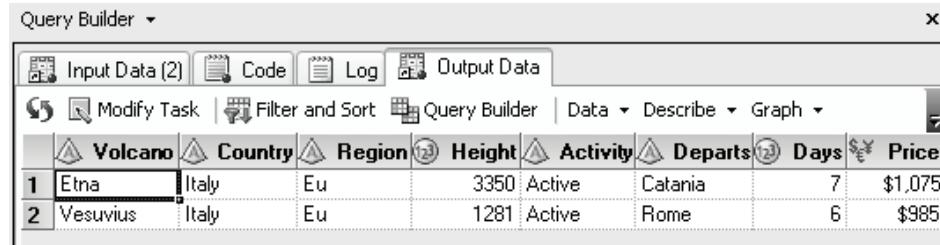


Click **Run** to see the results of this change to the query.

When SAS Enterprise Guide asks if you would like to replace the results from the previous run, click **Yes**.



Look at the resulting data table and note that only the tours of volcanoes in Europe appear in the data table.

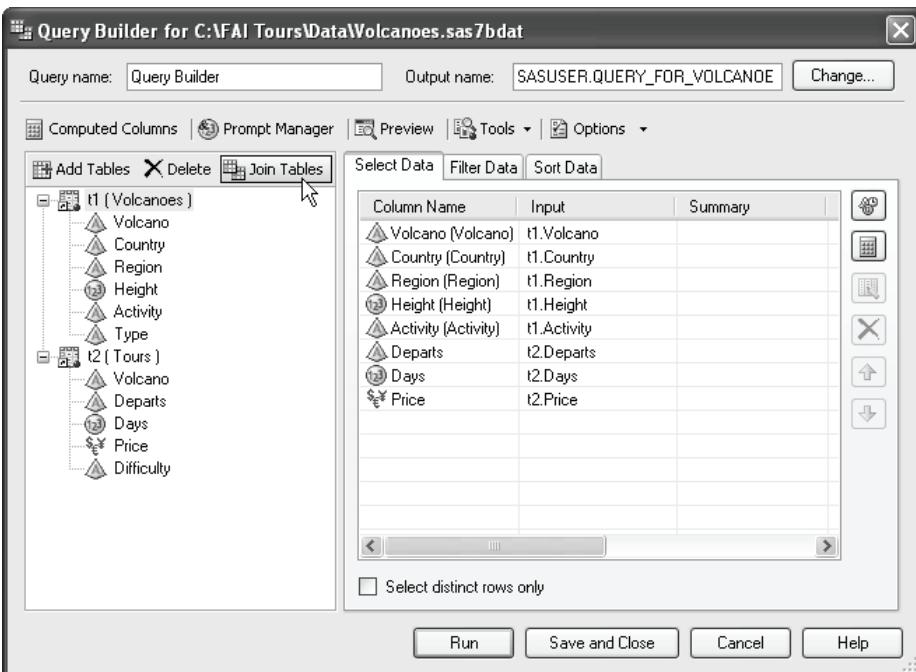


The screenshot shows the SAS Query Builder interface. The toolbar includes Input Data (2), Code, Log, Output Data, Modify Task, Filter and Sort, Query Builder, Data, Describe, and Graph. The main area displays a data table with columns: Volcano, Country, Region, Height, Activity, Departs, Days, and Price. The data shows two rows: Etna (Italy, Eu, 3350, Active, Catania, 7, \$1,075) and Vesuvius (Italy, Eu, 1281, Active, Rome, 6, \$985).

	Volcano	Country	Region	Height	Activity	Departs	Days	Price
1	Etna	Italy	Eu	3350	Active	Catania	7	\$1,075
2	Vesuvius	Italy	Eu	1281	Active	Rome	6	\$985

Selecting which rows to keep Suppose the Fire and Ice Tours company wants to expand the number of tours that it offers in Europe. The company wants to include in the list all the volcanoes in Europe, not just volcanoes that currently have tours. In SAS Enterprise Guide, the default action of a join is to include only the rows that appear in both tables. To change this default action, you need to change the type of join.

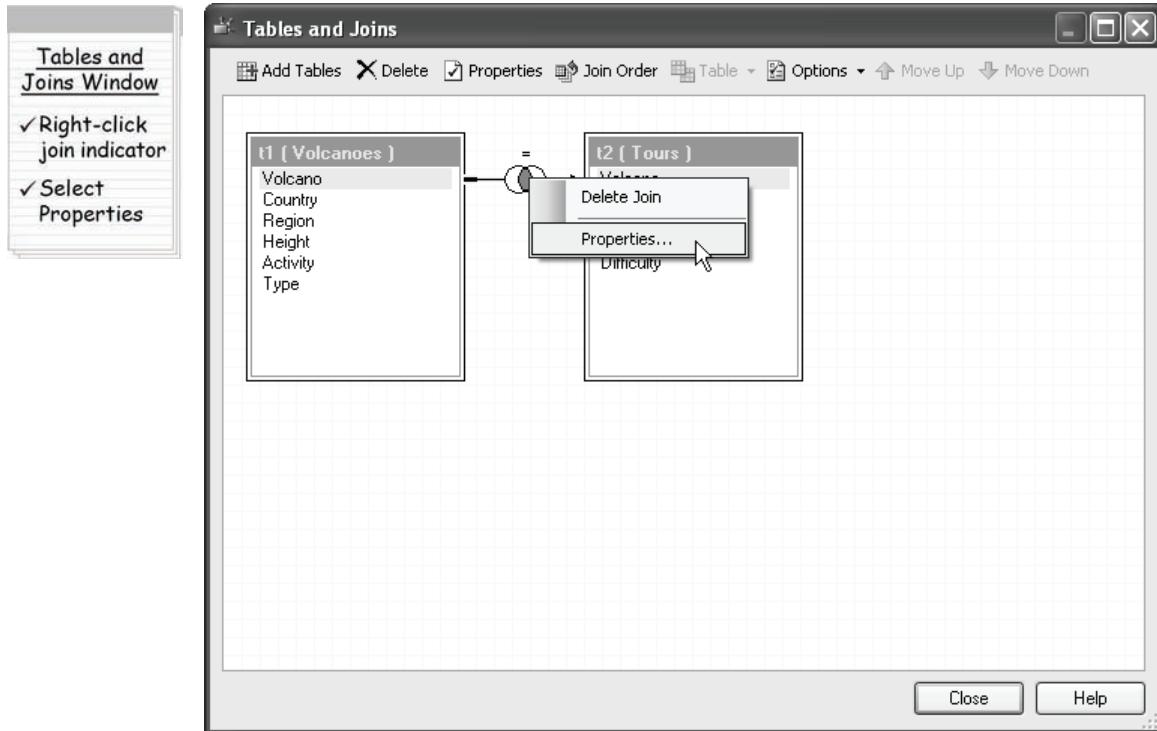
To reopen the query, right-click the Query Builder icon in the Project Tree or Process Flow and select **Modify Query Builder** or simply click **Modify Task** on the workspace toolbar for the query result.



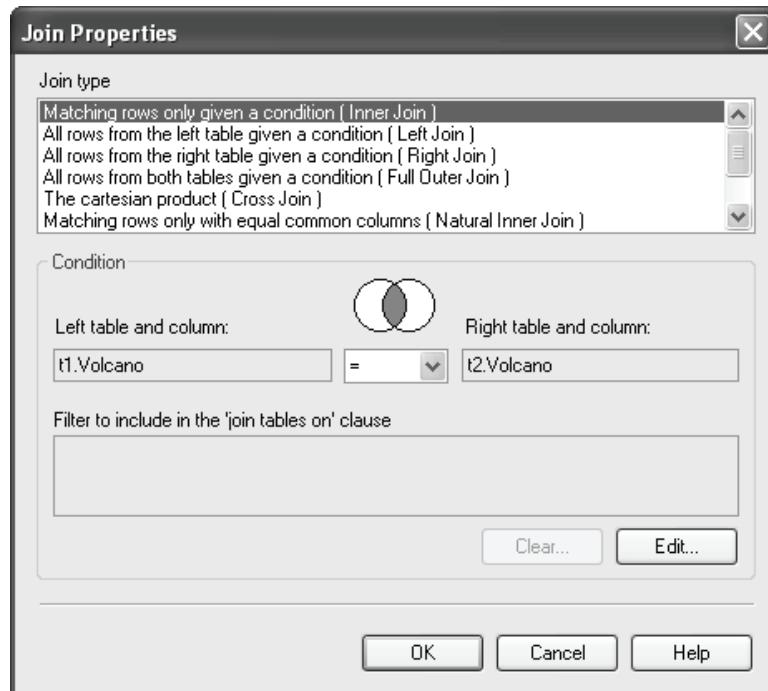
The screenshot shows the SAS Query Builder for the file C:\FAI Tours\Data\Volcanoes.sas7bdat. The workspace toolbar has a note: "Workspace Toolbar" and "Click Modify Task". The Query Builder window has a note: "Query Builder Window" and "Click Join Tables". The main window shows the "Join Tables" dialog. It lists two tables: t1 (Volcanoes) and t2 (Tours). The t1 table contains columns: Volcano, Country, Region, Height, Activity, Type. The t2 table contains columns: Volcano, Departs, Days, Price, Difficulty. A grid on the right shows the join conditions: Volcano (Volcano) is mapped to t1.Volcano, Country (Country) is mapped to t1.Country, Region (Region) is mapped to t1.Region, Height (Height) is mapped to t1.Height, Activity (Activity) is mapped to t1.Activity, Departs is mapped to t2.Departs, Days is mapped to t2.Days, and Price is mapped to t2.Price. Buttons at the bottom include Run, Save and Close, Cancel, and Help.

Click **Join Tables** to open the Tables and Joins window.

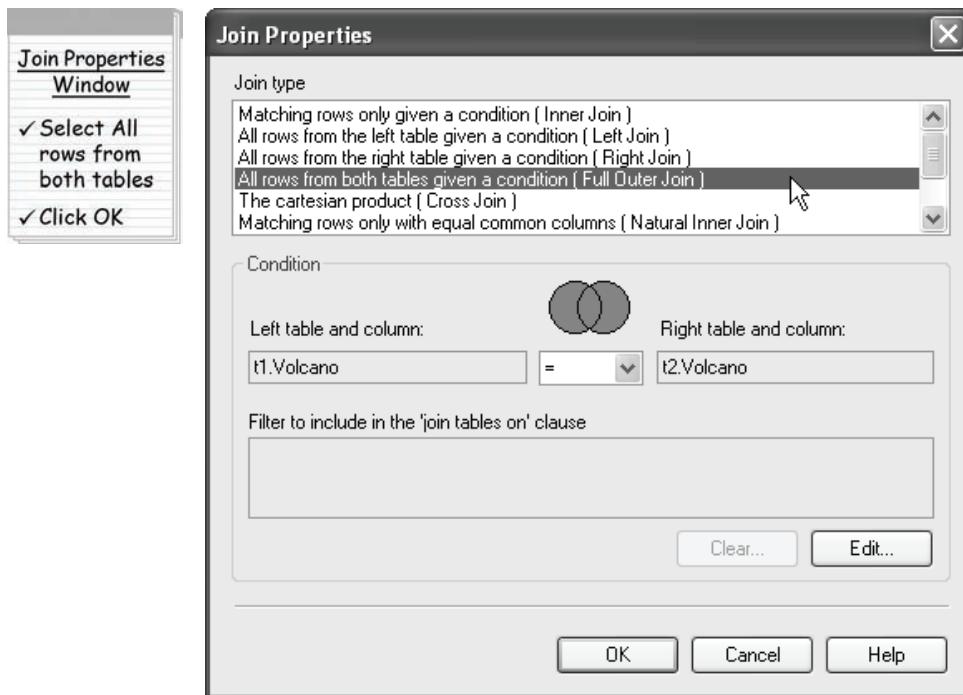
In the Tables and Joins window, notice the diagram  with the equal sign (=) above it on the line connecting the two tables. Right-click this join indicator and select **Properties** from the pop-up list.



This opens the Join Properties window. There are several types of joins listed under Join type. Each type of join has its own diagram using overlapping circles. For the **Matching rows only given a condition (Inner Join)** type of join, only the intersection of the two circles is filled with black.

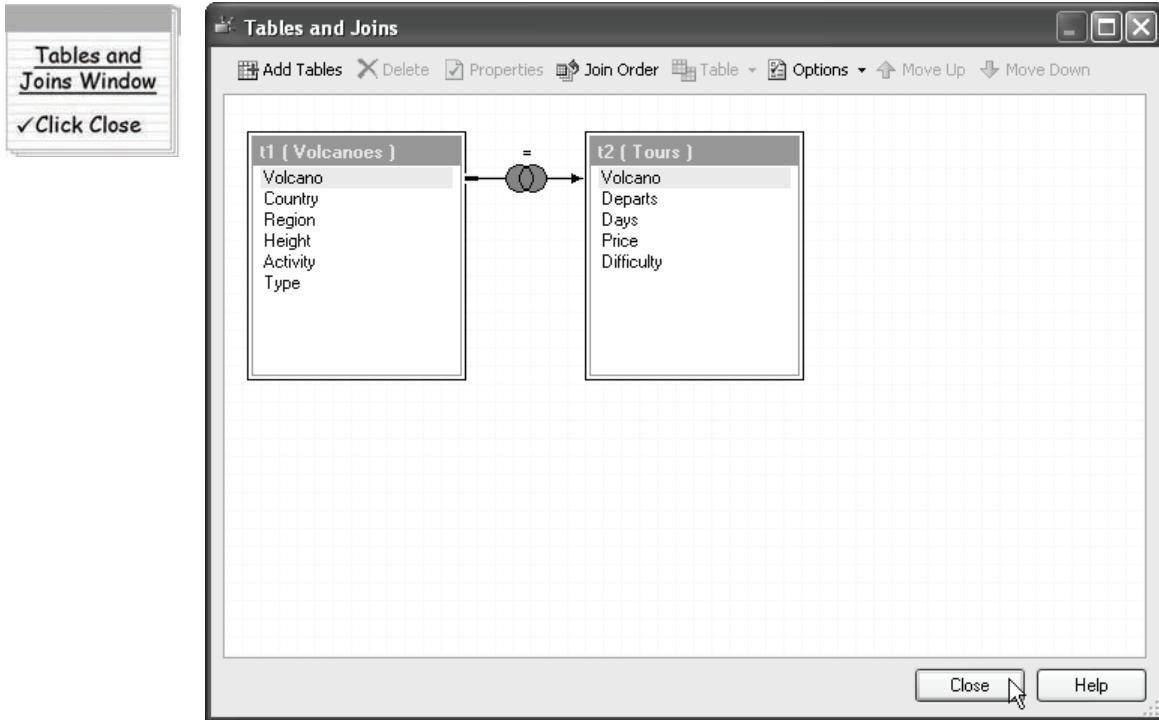


For this join, you want all the rows from each data table, even if there is no match. So, select **All rows from both tables given a condition (Full Outer Join)**.



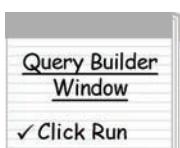
Notice that the join symbol has changed so that now both circles are completely filled with black. Click **OK**.

The join indicator in the Tables and Joins window has changed to reflect the type of join you just selected.

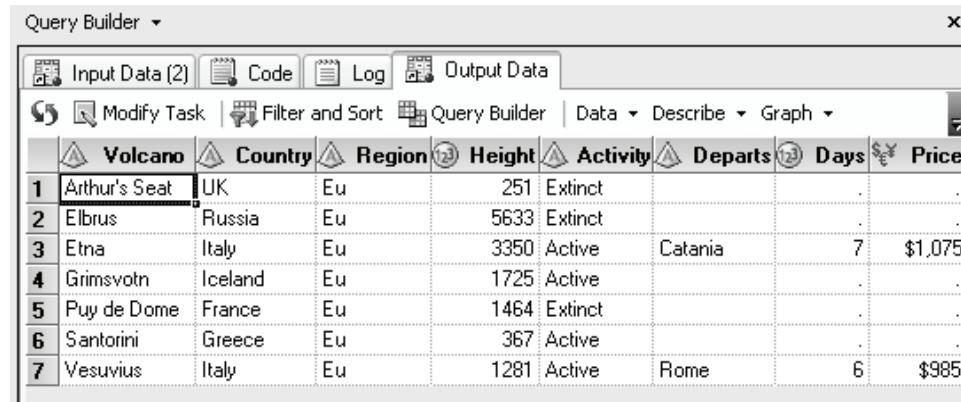


Click **Close** to return to the Query Builder window.

In the Query Builder window, click **Run** to see the results.



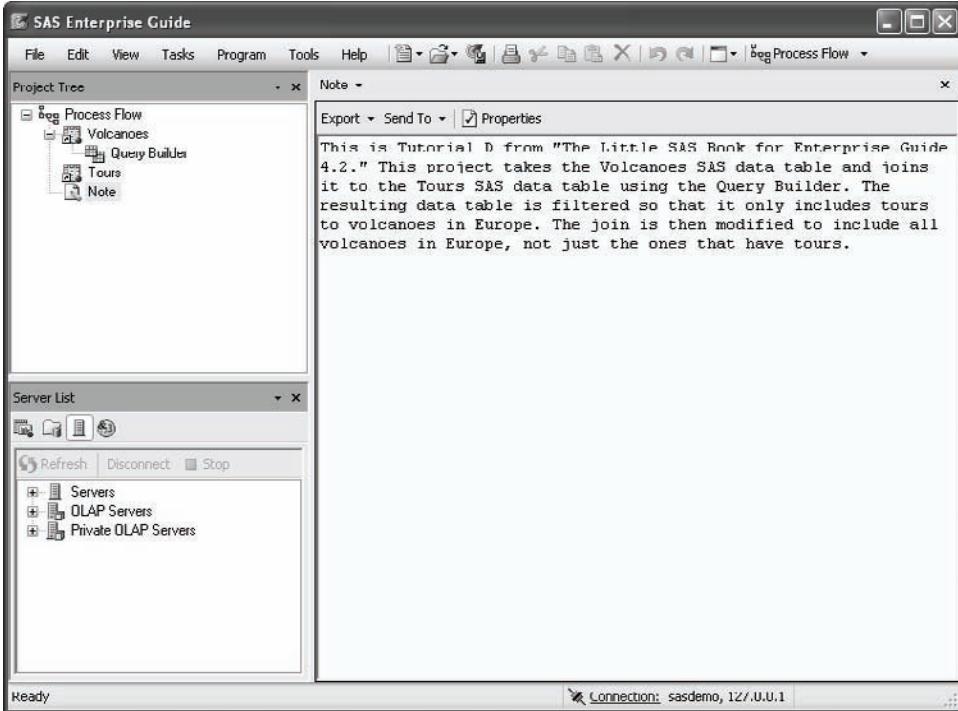
Look at the resulting data table in the workspace. Now the data include all the volcanoes in Europe, even if they don't have tours. Notice that all the columns from the Tours data table (Departs, Days, and Price) have missing values for the volcanoes for which there are no tours. Because all the volcanoes in the Tours data table also appear in the Volcanoes data table, all the columns from the Volcanoes data table have values.



The screenshot shows the SAS Query Builder interface with a data table titled "Volcano". The table has 7 rows and 9 columns. The columns are labeled: Volcano, Country, Region, Height, Activity, Departs, Days, and Price. The data includes information about various European volcanoes, such as Arthur's Seat, Elbrus, Etna, Grimsvotn, Puy de Dome, Santorini, and Vesuvius, along with their respective details like height, activity status, and price.

	Volcano	Country	Region	Height	Activity	Departs	Days	Price
1	Arthur's Seat	UK	Eu	251	Extinct		.	.
2	Elbrus	Russia	Eu	5633	Extinct		.	.
3	Etna	Italy	Eu	3350	Active	Catania	7	\$1,075
4	Grimsvotn	Iceland	Eu	1725	Active		.	.
5	Puy de Dome	France	Eu	1464	Extinct		.	.
6	Santorini	Greece	Eu	367	Active		.	.
7	Vesuvius	Italy	Eu	1281	Active	Rome	6	\$985

Completing the tutorial To complete the tutorial, add a note describing the project. Click the words **Process Flow** in the Project Tree. Then select **File ▶ New ▶ Note** from the menu bar. Enter a brief description of the project in the Note window in the workspace.



The screenshot shows the SAS Enterprise Guide interface. The Project Tree on the left contains a node for 'Process Flow' which is expanded to show 'Volcanoes', 'Query Builder', 'Tours', and 'Note'. The Note window on the right contains the following text:

```
This is Tutorial D from "The Little SAS Book for Enterprise Guide 4.2." This project takes the Volcanoes SAS data table and joins it to the Tours SAS data table using the Query Builder. The resulting data table is filtered so that it only includes tours to volcanoes in Europe. The join is then modified to include all volcanoes in Europe, not just the ones that have tours.
```

Below the Note window is the Server List, which shows 'Servers', 'OLAP Servers', and 'Private OLAP Servers' under the 'Refresh' tab. The status bar at the bottom indicates a connection to 'sasdemo, 127.0.0.1'.

Project Tree

- ✓ Click Process Flow

Menu Bar

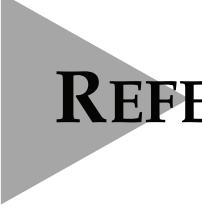
- ✓ Select File ▶ New ▶ Note

Note

- ✓ Type descriptive text

Menu Bar

- ✓ Save project
- ✓ Exit



REFERENCE SECTION

- Chapter 1** SAS Enterprise Guide Basics 133
- Chapter 2** Bringing Data into a Project 167
- Chapter 3** Changing the Way Data Values Are Displayed 189
- Chapter 4** Modifying Data in the Query Builder 201
- Chapter 5** Sorting and Filtering Data 223
- Chapter 6** Combining Data Tables 239
- Chapter 7** Producing Simple List and Summary Reports 249
- Chapter 8** Producing Complex Reports in Summary Tables 269
- Chapter 9** Basic Statistical Analysis 283
- Chapter 10** Producing Graphs 297
- Chapter 11** Changing Result Styles and Formats 311
- Chapter 12** Adding Flexibility with Prompts and Conditions 325

1

“Every tradition begins as an innovation, and every innovation is built on the traditions before it.”

JOE CRAVEN

From Joe Craven, award-winning multi-instrumentalist, musical archaeologist, and educator,
www.joecraven.com. Reprinted by permission of the author.



CHAPTER 1

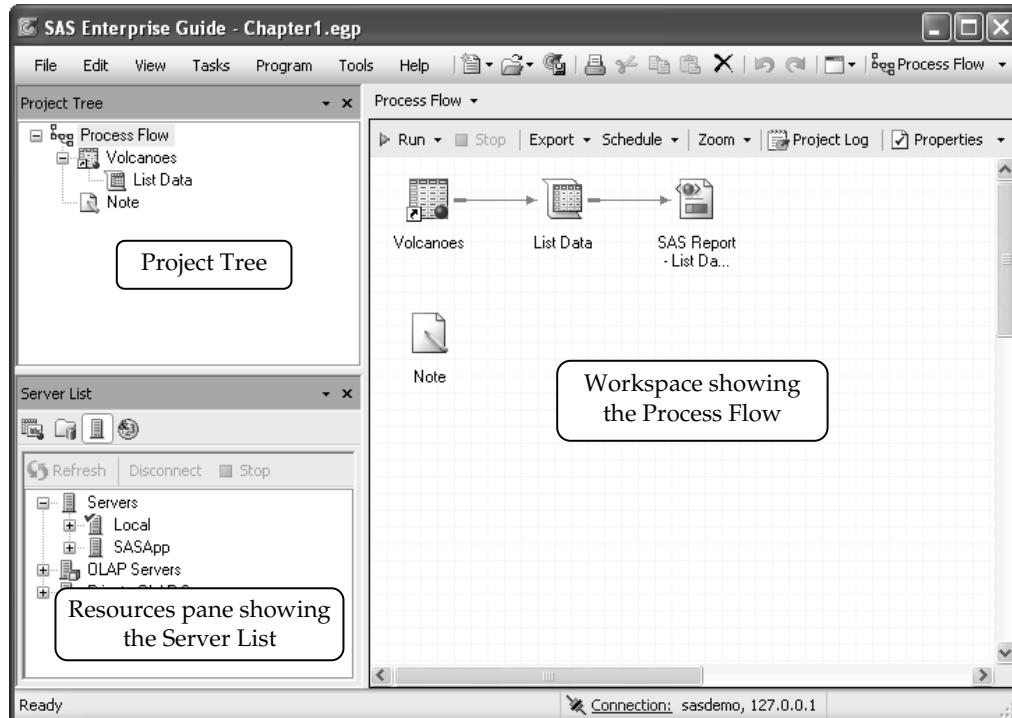
SAS Enterprise Guide Basics

- 1.1 SAS Enterprise Guide Windows **134**
- 1.2 Splitting the Resources Pane and Workspace **136**
- 1.3 Projects **138**
- 1.4 Managing Process Flows **140**
- 1.5 Running and Rerunning Tasks **142**
- 1.6 Creating and Exporting Task Templates **144**
- 1.7 SAS Data Tables **146**
- 1.8 Properties of Data Tables **148**
- 1.9 Properties of Columns **150**
- 1.10 Selected Informats **152**
- 1.11 Selected Standard Formats **154**
- 1.12 Scheduling Projects to Run at Specific Times **156**
- 1.13 Editing SAS Programs Generated by a Task **158**
- 1.14 Writing and Running Custom SAS Programs **160**
- 1.15 Viewing Program and Project Logs **162**
- 1.16 Using the Options Window **164**

1.1 ► SAS Enterprise Guide Windows

SAS Enterprise Guide has many windows. You can customize the appearance of SAS Enterprise Guide—closing some windows, opening others, and resizing them all—until it looks just the way you want. Then SAS Enterprise Guide will remember those settings so the next time you open it, everything will be just where you left it.

Here is SAS Enterprise Guide with its windows in their default positions.



Some windows are open by default while some are closed or hidden behind other windows. You can open or unhide the major windows using the **View** menu.

Docked windows Some of the windows in SAS Enterprise Guide are docked. Most of the docked windows can appear on the left or right side of the application. To change a window from one side to the other, click the down arrow (▼) in the upper-right corner of the window and select **Dock Left** or **Dock Right** from the pop-up menu. From this menu, you can also select **Auto Hide**. If you hide a window, it will be reduced to a tab along the side. To view a hidden window, position your cursor over the window's tab. When you move the cursor out of the window, it will be reduced to a tab again. To unhide a window, click its tab or select it from the **View** menu. These windows are docked:



Project Tree The Project Tree window displays the items in a project in a hierarchical tree diagram. This window is open by default.



Server List The Server List window lists available SAS servers, and the files and SAS data libraries on those servers. A SAS server is any computer on which SAS software is installed. The computer on which you run SAS Enterprise Guide may or may not be a SAS server. This window appears in the Resources pane, and is open by default.



Task List The Task List window lists all available tasks and task templates. Using the drop-down list at the top of this window, you can choose to display tasks by category, tasks by name, or task templates. You can open a task by double-clicking its name in this window. This window appears in the Resources pane, and is closed by default. To open this window, click its icon in the Resources pane.



SAS Folders The SAS Folders window lists any folders that have been defined in metadata. This window appears in the Resources pane, and is closed by default. To open this window, click its icon in the Resources pane.



Prompt Manager The Prompt Manager window lists any prompts defined for the current project. This window appears in the Resources pane, and is closed by default. To open this window, click its icon in the Resources pane.



Task Status The Task Status window displays notes about tasks that are currently running. This window is different from other docked windows because it is docked to the bottom of the application, and you cannot move it or reduce it to a tab. This window is closed by default. To open the Task Status window, select it from the **View** menu.

Workspace The workspace is not itself a window, but it is very important. This is where the Process Flow and document windows appear. The workspace is always there and cannot be closed. However, you can open and close individual items inside the workspace.



Process Flow The Process Flow window displays the items in a project and their relationship using a process flow diagram. You can open only one project at a time, but you can create as many process flows as you wish inside a single project. You can open the Process Flow by selecting it from the **View** menu, by double-clicking its name in the Project Tree, by selecting it from the drop-down list at the top of the workspace, by selecting the **Open Process Flow ▾** drop-down list on the menu bar, or by pressing **F4**.



Document windows The document windows display your data, results, programs, logs, and notes. There is a different type of icon for every kind of document. This icon represents a SAS data table.

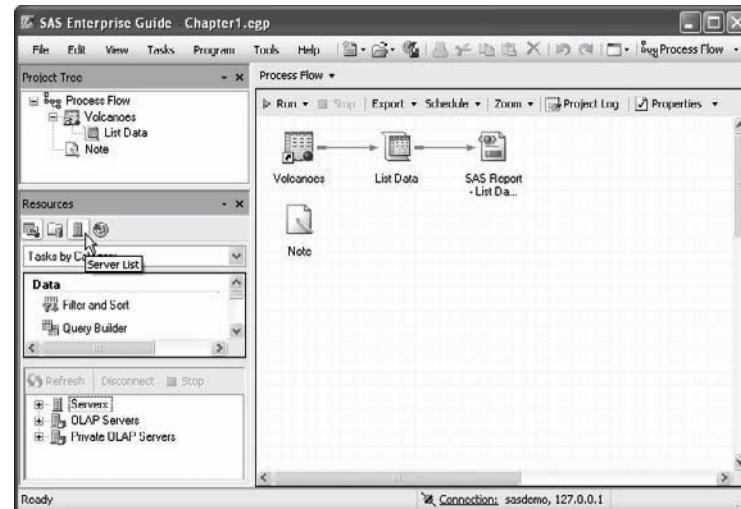
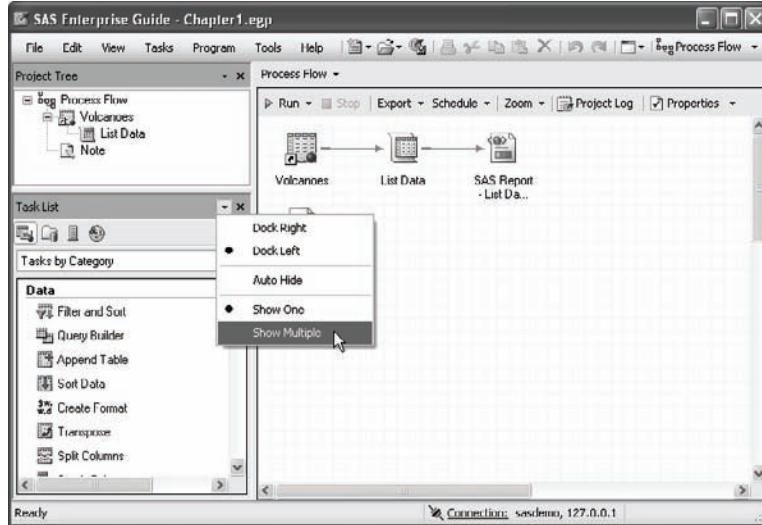
Menus and tools The menus and tools across the top of SAS Enterprise Guide (also called the menu bar) are always the same. However, the menus and tools inside the workspace (also called the workspace toolbar) change. For example, the options above a Process Flow are different from the options above a data table. You can also right-click many objects to open a pop-up menu for that object. So you can see that there are often several ways to do the same thing. This book cannot list all the ways to do every action, but with a little exploration you can find them.

Restoring windows Once you have rearranged your windows, you may decide you want them back where they started. To restore them to their original locations, select **Tools ▶ Options** from the menu bar. Then in the General page of the Options window, click **Restore Window Layout**.

1.2 Splitting the Resources Pane and Workspace

The Resources pane and the workspace are busy places. The Resources pane is home to four windows, while the workspace accommodates even more. By default, you can see only one item at a time, but you can see more if you split the Resources pane or workspace.

Splitting the Resources pane To split the Resources pane, click the down-arrow (\blacktriangledown) at the top of the pane, and select **Show Multiple** from the pull-down list. At first, you will only see one window because only one window is open.



To open other windows, click their icons: Task

List Folders

Server List

or Prompt Manager .

You can open all four windows at once in the Resources pane if you wish. In this example, two windows are open, the Task List and the Server List.

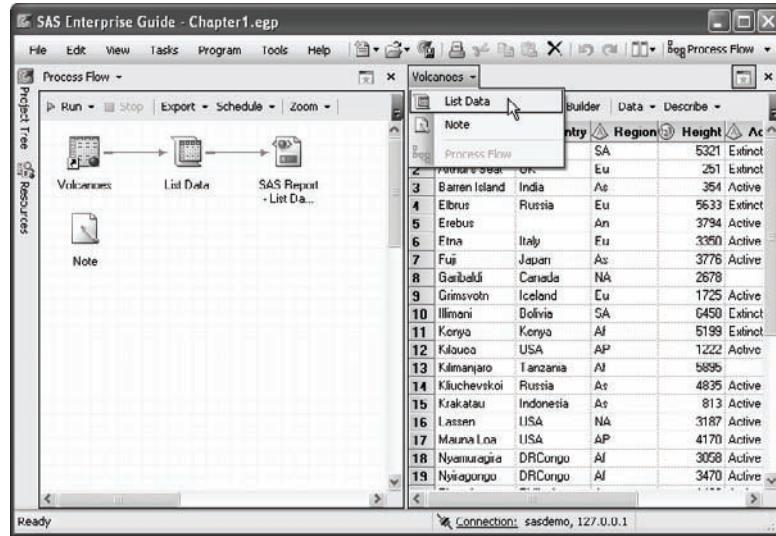
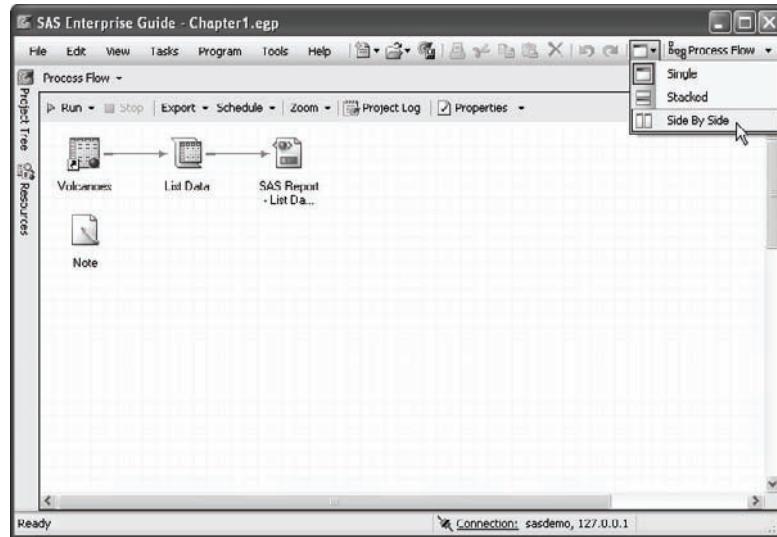
To return the Resources pane to normal, click the down-arrow and select **Show One**.

Maximizing the workspace It may be helpful to make the workspace as large as possible before you split it. To do this, select **View ▶ Maximize Workspace** from the menu bar. When you maximize the workspace, the Project Tree and Resources pane become tabs pinned to the edge of SAS Enterprise Guide. You can temporarily expand those windows by moving the cursor over a tab. When you move the cursor away, the window will be reduced to a tab again. To return the workspace to its normal size, select **View ▶ Maximize Workspace** again.

Splitting the workspace You can split the workspace into two pieces. First, open any items you wish to view. Then click the Workspace Layout icon on the menu bar



, and select either **Stacked** or **Side By Side** from the pull-down list. You can also do this by selecting **View ▶ Workspace Layout** from the menu bar.



You can click the down-arrow at the top of the workspace to view a drop-down list of all items that are currently open. To display an item, select it from the list.

To unsplit the workspace, click the Workspace Layout icon again, and select **Single** from the pull-down list. You can also click one of the Xs in the upper-right corners to close that section of the workspace.

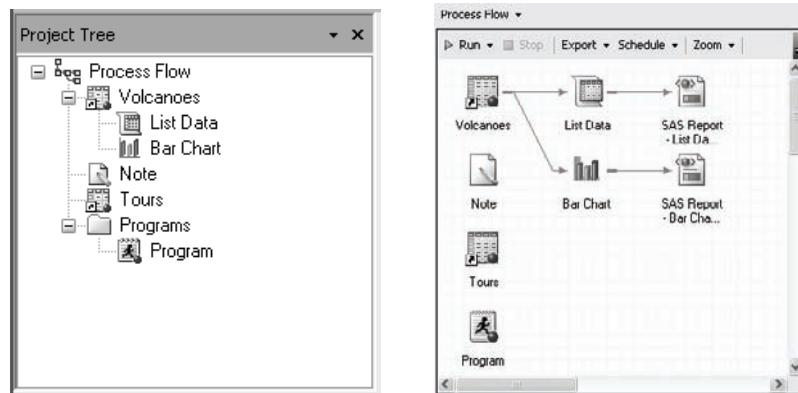
1.3 Projects

In SAS Enterprise Guide, all the work you do is organized into projects. A project is a collection of related data, tasks, results, programs, and notes. Projects help you by keeping track of everything, even if your data are scattered in many directories or on more than one computer. That way, when you come back to an old project six months or a year later, you won't be left wondering which data sets you used or what reports you ran.

You can have as many projects as you like, and you can use a data set over and over again in different projects, so there is a lot of flexibility. However, you can have only one project open at a time. Also, if you share a project file with someone else, that person must have access to your data files and any other items you reference.

To create a new project, select **File ▶ New ▶ Project** from the menu bar. To open an existing project, select **File ▶ Open ▶ Project** and navigate to your project.

Project Tree and Process Flow The Project Tree window displays projects in a hierarchical tree diagram, while the Process Flow window displays projects using a process flow diagram. In either window, the items in your project are represented by icons, and connected to show the relationship between items. Here are examples of a Project Tree and a Process Flow showing the same project. This project contains several types of items: data, tasks, results, a program, and a note.



Data Data files in a project may be SAS data tables, raw data files, or files from other databases or applications, such as Microsoft Excel spreadsheets. Projects contain shortcuts to data files, not the actual data. If you delete a project, your data files will still exist. This icon represents a SAS data table.



Tasks Tasks are specific analyses or reports that you run, such as List Data or Bar Chart. Every time you run a task, SAS Enterprise Guide adds an icon representing that task. This icon represents the Bar Chart task.



Results Results are the reports or graphs produced by tasks you run. Results are represented by icons labeled with the type of output (SAS Report, HTML, PDF, RTF, or text) and the name of the task. This icon represents output in SAS Report format.

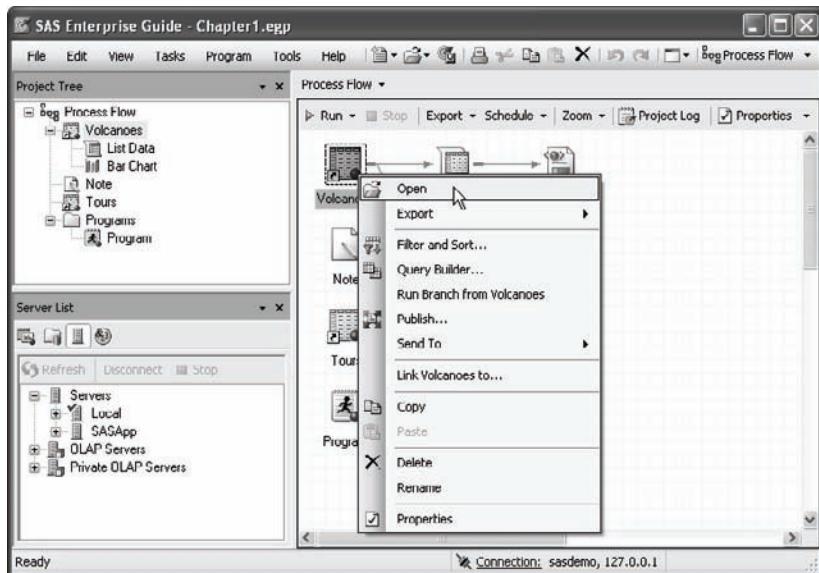


Notes Notes are optional text files you can use to document your work, or record comments or instructions for later use. To create a note, select **File ▶ New ▶ Note** from the menu bar. A text window will open, allowing you to type whatever you wish.



Programs Programs are files that contain SAS code. You can open existing programs in SAS Enterprise Guide, or you can write new programs.

Showing properties and opening items You can display the properties for any item by right-clicking its icon in the Project Tree or Process Flow and selecting **Properties** from the pop-up menu. You can open any item by double-clicking its icon, or by right-clicking its icon and selecting **Open** from the pop-up menu.



Renaming and deleting items You can rename most items by right-clicking the item and selecting **Rename** from the pop-up menu. You can delete an item in a project by right-clicking and selecting **Delete**. Note that if you delete data from a project, only the shortcut to that data is deleted, not the actual data file.

Saving a project To save a project, select **File ▶ Save project-name** or **File ▶ Save project-name As** from the menu bar. Each project is saved as a single file and has a file extension of .egp. You can save data, programs, and results in separate files by right-clicking the icon for that item and selecting **Export** from the pop-up menu.

1.4 Managing Process Flows

In SAS Enterprise Guide, you can have only one project open at a time. However, you can have an unlimited number of process flows within a single project. So, if you have a complex project, you may want to divide it into several process flows.

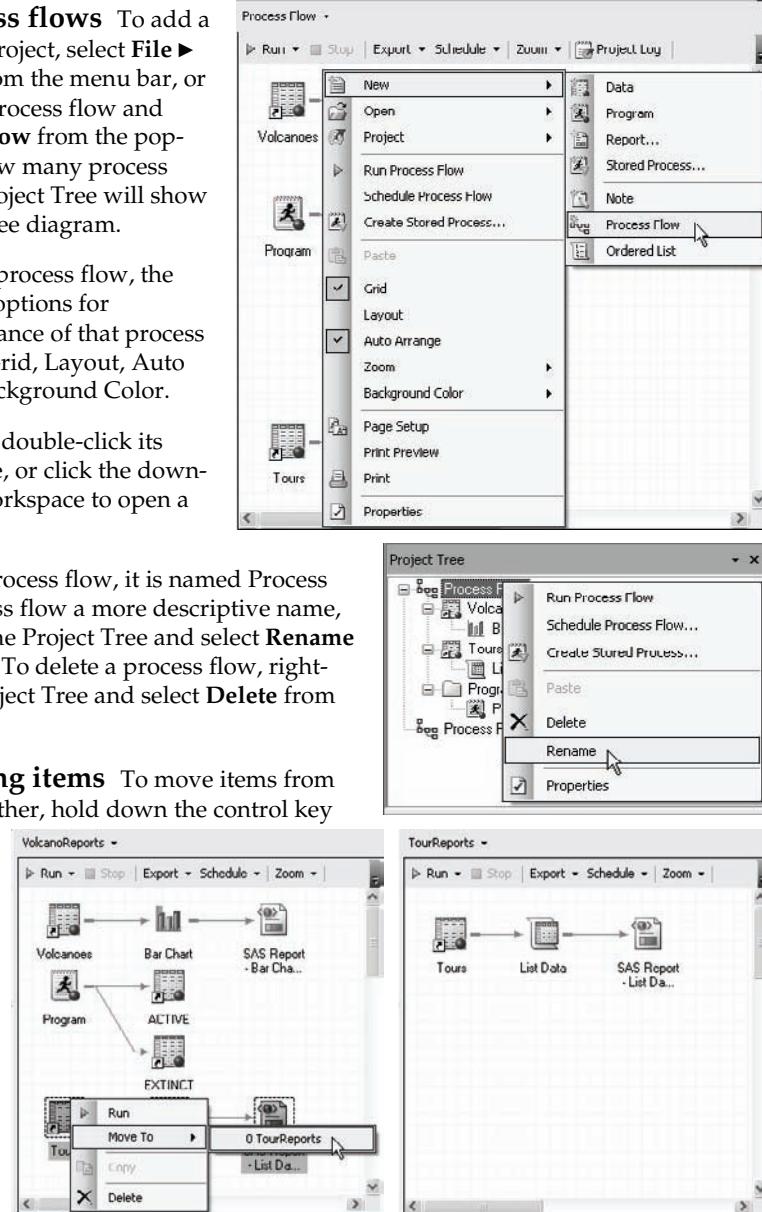
Adding new process flows To add a new process flow to a project, select **File ▶ New ▶ Process Flow** from the menu bar, or right-click the current process flow and select **New ▶ Process Flow** from the pop-up menu. No matter how many process flows you create, the Project Tree will show all of them in a single tree diagram.

When you right-click a process flow, the pop-up menu displays options for customizing the appearance of that process flow. Options include Grid, Layout, Auto Arrange, Zoom, Background Color, Paste, Grid, Layout, Auto Arrange, Zoom, Background Color, Page Setup, Print Preview, Print, Properties, and Process Flow.

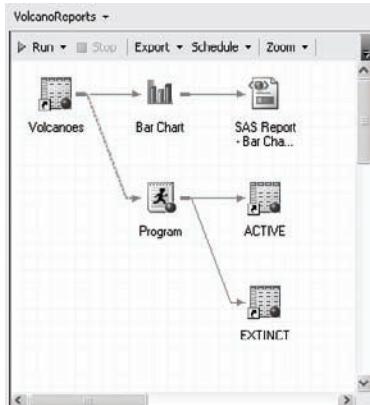
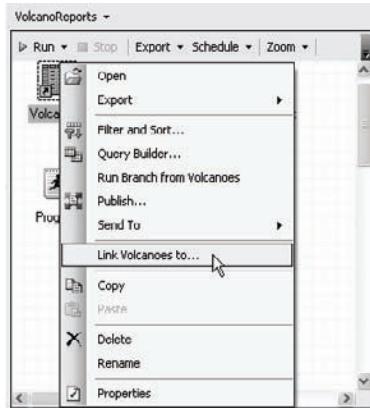
To view a process flow, double-click its name in the Project Tree, or click the down-arrow (▼) above the workspace to open a pull-down list.

When you add a new process flow, it is named Process Flow *n*. To give a process flow a more descriptive name, right-click its name in the Project Tree and select **Rename** from the pop-up menu. To delete a process flow, right-click its name in the Project Tree and select **Delete** from the pop-up menu.

Moving and copying items To move items from one process flow to another, hold down the control key (CTRL), and click all the items you want to move. Then right-click, and select **Move to ▶ process-flow-name** from the pop-up menu. In this example, three items are being moved to the process flow named TourReports.



Copying items is similar to moving items except that you cannot copy results. Select the items to be copied using control-click. Then right-click the items and select **Copy** from the pop-up menu, and right-click the target process flow and select **Paste**.

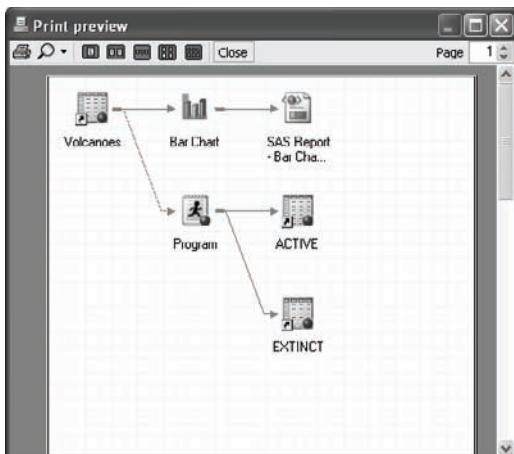


Linking items When you run a process flow, items are executed from top left to bottom right, following the branches created by links between items. You can add links between items to show relationships that may not be clear, or to force items to run in a particular order. For example, if you create a format that is used by a task, you might want to add a link indicating that the task follows the format. To add a link, right-click the initial item and select **Link item-name to** from the pop-up menu. A Link window will open

showing all the other items to which you can link. Select the item to which you want to link, and click **OK**.

In this process flow, the Volcanoes data icon has been linked to a program icon to show that this program uses the Volcanoes data table. Notice that when you add links they use a dashed line instead of a solid line.

To delete a link that you previously added, right-click the icon for that link ➔ in the Project Tree and select **Delete** from the pop-up menu.



Printing process flows You can print a copy of your process flow. To control page size and orientation, click the process flow and select **File ▶ Page setup for Process Flow** from the menu bar. To preview a printout, select **File ▶ Print preview for Process Flow**. To print the process flow, select **File ▶ Print Process Flow**. Here is the Print preview window for the VolcanoReports process flow.

1.5 Running and Rerunning Tasks

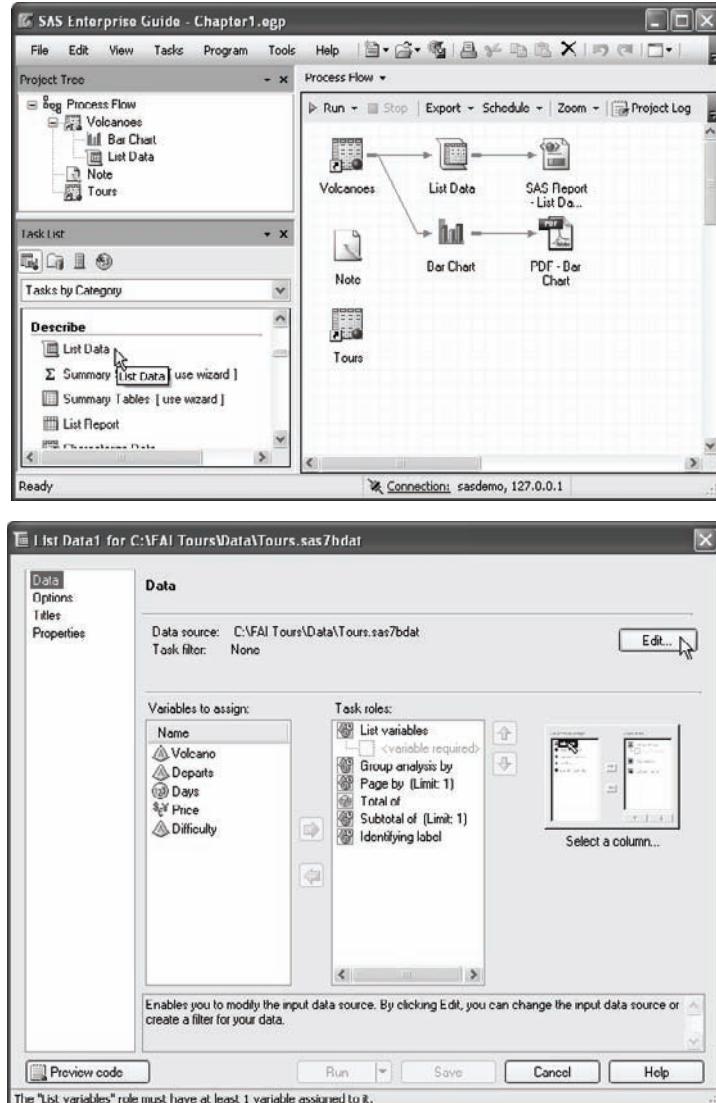
Running tasks is, of course, what SAS Enterprise Guide is all about. Regardless of which task you choose to run, the basic steps are the same: open the task, select the data, and then run the task.

Opening a task To open a task, select it from the **Tasks** menu, or click its name in the Task List window, or open a Data Grid and then select the task from the workspace toolbar. The window for that task will open. In this example, the List Data task is being selected in the Task List window.

Selecting the data

table When you open a task, it will use the data table that is currently active. If a Data Grid is open, then that data table will be active. You can also make a data table active by simply clicking its icon in the Project Tree or Process Flow before you open a task.

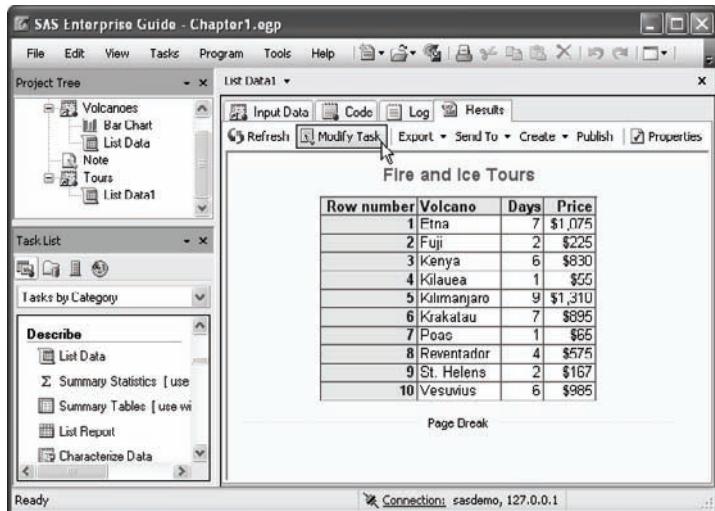
After you open a task, you can change the data table by clicking the **Edit** button in the Data page of the task window. The Edit Data and Filter window (not shown) will open where you can choose an alternate data table for the task. See section 5.1 for details about the Edit Data and Filter window.



Running a task Every task includes a Data page where you assign variables to task roles. Using the selection pane on the left, you can open other pages. The preceding image shows the List Data task, which has four pages: Data, Options, Titles, and Properties. When you are satisfied with all the settings, click the **Run** button. If you have more than one SAS server, your task will run on the same server where the data table is stored. If you decide you want to stop a task while it's running, select **Program ▶ Stop** from the menu bar, or click the **Stop** button ■ on the workspace toolbar above the Process Flow. When the task has finished running, the results will be displayed in the workspace.

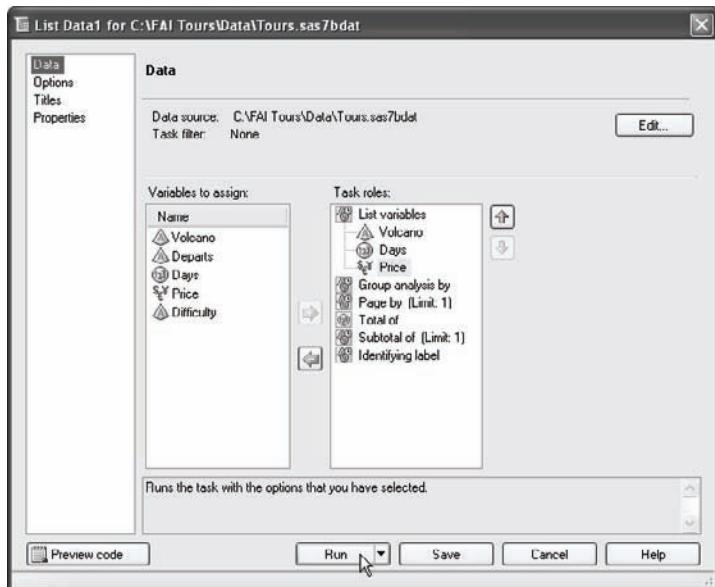
Rerunning a task

To make changes to a task and run it again, first reopen the task window. You can do this by clicking **Modify Task** on the workspace toolbar for the Results tab. You can also reopen a task by right-clicking the task icon in the Project Tree or Process Flow, and selecting **Modify task-name** from the pop-up menu.



Once the task window is open, you can make changes. Then click the **Run** button to rerun the task.

If you just want to rerun a task without reopening the task window, click **Refresh** on the workspace toolbar for the Results tab. You can also right-click the task icon in the Project Tree or Process Flow, and select **Run** from the pop-up menu.

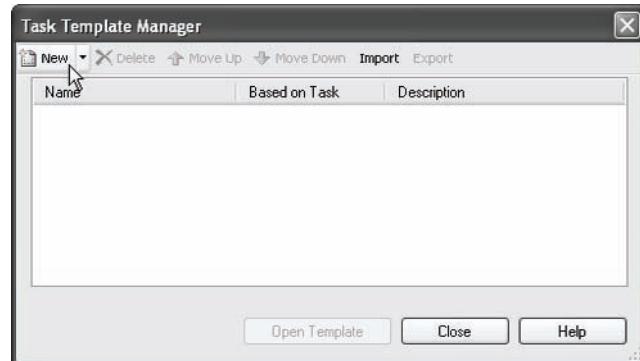
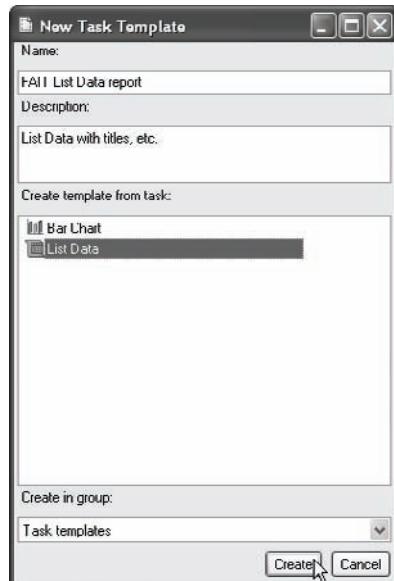


1.6 Creating and Exporting Task Templates

Even with the simplest tasks, there are many ways to customize your results. Once you have spent a lot of time changing titles, choosing options, and specifying a style; you might wish you could save all those settings and use them to create new results. With task templates, you can.

Task templates allow you to save tasks in a form that is independent of data. In other words, task templates save all your settings except the assignment of variables to task roles and certain data-dependent options. Most tasks can create templates, but a few of the more data-driven tasks (including Summary Tables and Append Table) cannot. SAS programmers will be interested to know that task templates are unrelated to the various kinds of templates created by the Output Delivery System.

Creating a task template To create a task template from a task that you have already run, select **Tasks ▶ Task Templates ▶ Task Template Manager** from the menu bar. This opens the Task Template Manager. Click **New** to open the New Task Template window.



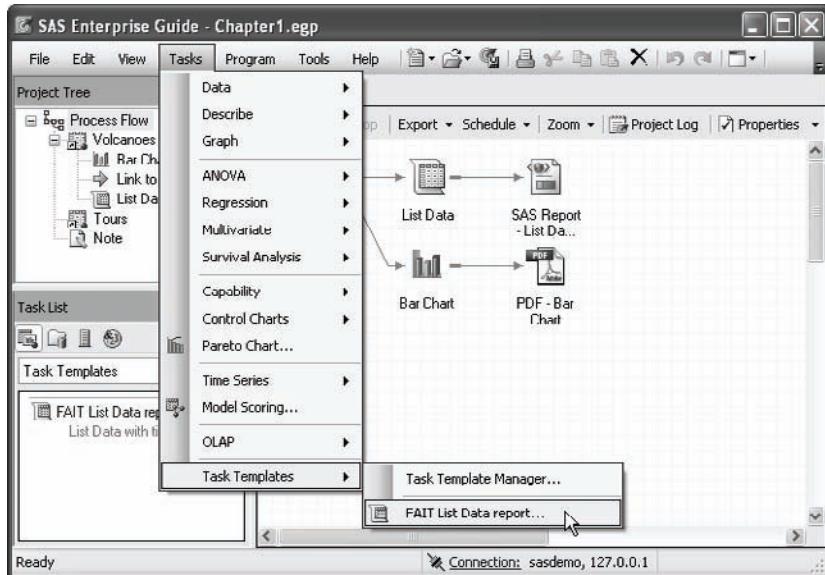
In the New Task Template window, type a name for the new template in the **Name** box. You can type an optional description in the **Description** box if you wish. The area labeled **Create template from task** lists all the tasks currently in the project. Choose a task by clicking its name. Then click **Create**. The new task template will be listed in the Task Template Manager. Click **Close**.

You can also create task templates directly from tasks. If the task window is open, then you can click the down-arrow (▼) on the **Run** button, and select **Create Template** from the pull-down list.

To delete a task template, simply open the Task Template Manager, click the task template name, and click **Delete**.

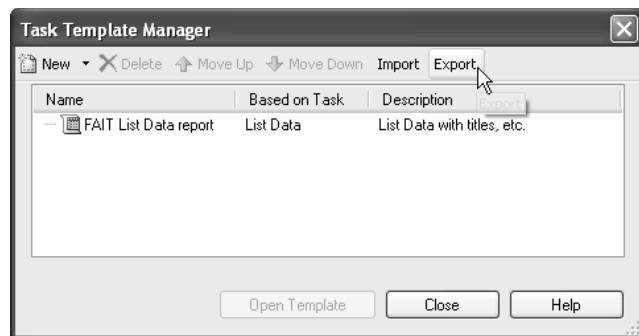
Using a task template Once you create a task template, unless you delete it, it will be available to you every time you open SAS Enterprise Guide. There are two ways to open a task template. You can select **Tasks ▶ Task Templates ▶ task-template-name** from the menu bar, or you can open the Task List window, select **Task Templates** from the drop-down list, and click the name of your task template.

By default, when you open a task template, it uses the active data table. After you open a task template, you can choose a different data table by clicking the **Edit** button in the Data page.



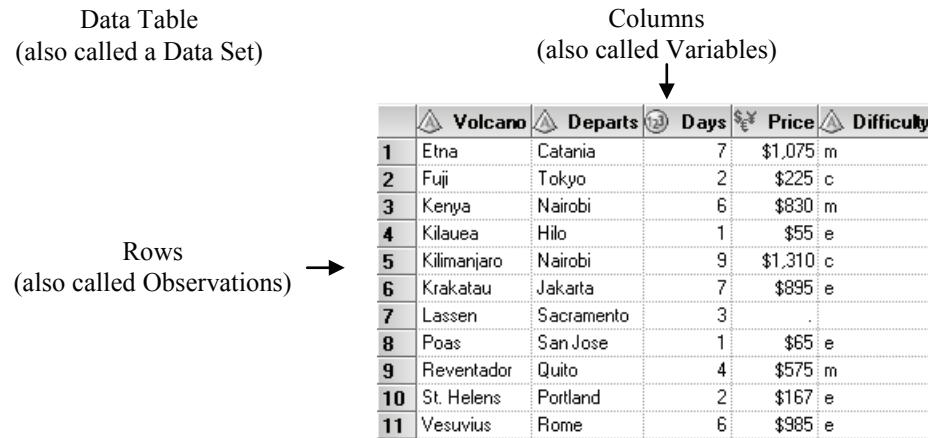
Exporting a task template When you create a task template, it will be saved in a default location. This location is associated with your Windows user account. To share task templates with other people (including anyone who uses the same computer, but a different Windows account), you must export the templates. To do this, open the Task Template Manager, and click **Export**. SAS Enterprise Guide will prompt you to select the templates you wish to export, and to specify a location for saving them.

To import task templates, open the Task Template Manager and click **Import**. Then navigate to the location of the task templates you wish to import.



1.7 SAS Data Tables

SAS Enterprise Guide can read and write many kinds of data files (see Chapter 2 for more on this topic), but for most purposes, you will want to have your data in a special form called a SAS data table. When you open a SAS data table, it is displayed in the workspace in a Data Grid. The following Data Grid shows the Tours data table that was created in Tutorial A. A new tour has been added for the volcano Lassen.



Terminology In SAS Enterprise Guide, rows are also called observations, columns are also called variables, and data tables are also called data sets. SAS Enterprise Guide uses all these terms. Some tasks use the term columns and others refer to variables, depending on the context.

Data types and data groups In SAS Enterprise Guide, there are two basic types of data: numeric and character. Numeric data are divided into four data groups: numeric, currency, time, and date. For each of these, SAS Enterprise Guide has special tools: informats for reading that type of data, functions for manipulating that type of data, and formats for displaying that type of data. SAS Enterprise Guide uses a different icon to identify each kind of data.



Character data may contain numerals, letters, or special characters (such as \$ and !) and can be up to 32,767 characters long. Character data are represented by a red pyramid with the letter A on it.



Currency data are numeric values for money and are represented by a picture of the dollar, euro, and yen symbols.



Date data are numeric values equal to the number of days since January 1, 1960. The table below lists four dates, and their corresponding SAS date and formatted values:

Date	SAS date value	MMDDYY10. formatted value
January 1, 1959	-365	01/01/1959
January 1, 1960	0	01/01/1960
January 1, 1961	366	01/01/1961
January 1, 2010	18263	01/01/2010

You will rarely see unformatted SAS date values in SAS Enterprise Guide. However, because dates are numeric, you can use them in arithmetic expressions to find, for example, the number of days between two dates. Datetime values are included in this data group, and are the number of seconds since January 1, 1960. Date data are represented by a picture of a calendar.



Time data are numeric values equal to the number of seconds since midnight. Time data are represented by a picture of a clock.



Other numeric data, that are not dates, times, or currency, are simply called numeric. They may contain numerals, decimal places (.), plus signs (+), minus signs (-), and E for scientific notation. Numeric data are represented by a blue ball with the numbers 1, 2, and 3 on it.

Numeric versus character If the values of a column contain letters or special characters, they must be character data. However, if the values contain only numerals, then they may be either numeric or character. You should base your decision on how you will use the data. Sometimes data that consist solely of numerals make more sense as character data than as numeric. Zip codes, for example, are made up of numerals, but it just doesn't make sense to add or subtract zip codes. Such values work better as character data.

Column names Column names in SAS Enterprise Guide may be up to 32 characters in length, and can begin with or contain any character, including blanks.

Moving data between SAS Enterprise Guide and Base SAS Any data created in SAS Enterprise Guide can be used in Base SAS, but the default rules for naming variables are different. Base SAS uses the VALIDVARNAME=V7 SAS system option, while SAS Enterprise Guide uses VALIDVARNAME=ANY. For the sake of compatibility, you may want to follow these rules when naming columns: choose column names that are 32 characters or fewer in length, start with a letter or underscore, and contain only letters, numerals, and underscores.

Missing data Sometimes, despite your best efforts, your data may be incomplete. The value of a particular column may be missing for some rows. In those cases, missing character data are represented by blanks, and missing numeric data are represented by a single period (.). In the preceding Data Grid, the value of Price is missing for the tour of Lassen, and its place is marked by a period. The value of Difficulty is missing for the same tour and is left blank.

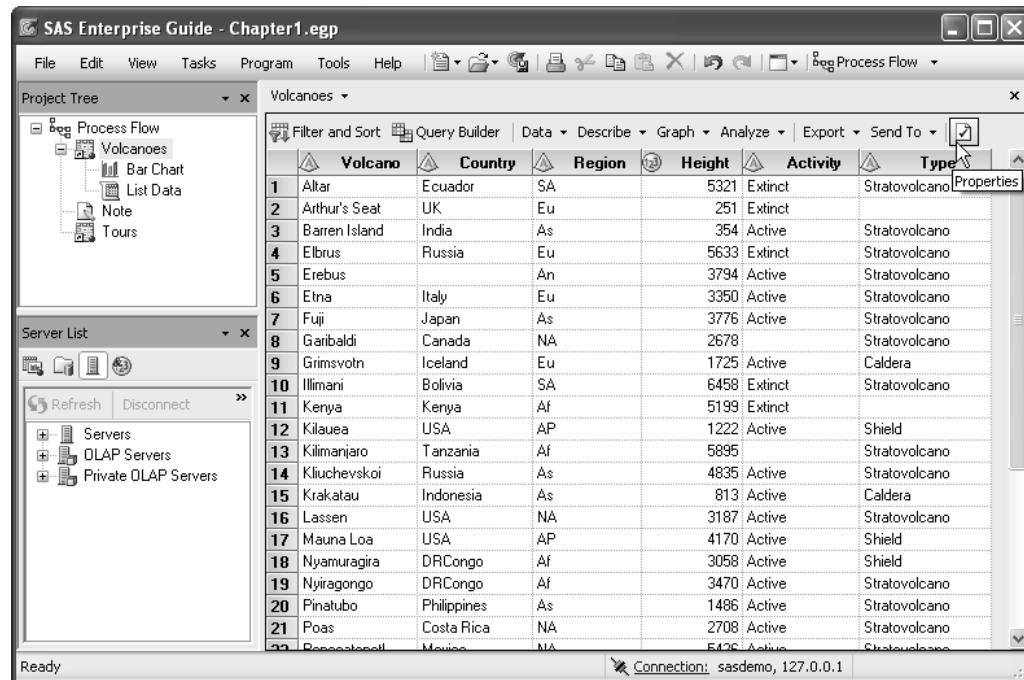
Documentation stored in SAS data tables In addition to your actual data, SAS data tables contain information about the data table, such as its name, the date that you created it, and the version of SAS you used to create it. SAS also stores information about each column in the data table, including its name, type, and length. This information is sometimes called the descriptor portion of the data table, and it makes SAS data tables self-documenting. This information is what you see in the Properties windows for data tables and columns. These Properties windows are described in more detail in the next two sections.

1.8 Properties of Data Tables

Someday you may be given a SAS Enterprise Guide project that was created by someone else. If you are unsure what the project does, then it would be a good idea to start by checking the properties of the data tables.

Opening the Properties window To display information about a data table, first open it in a Data Grid by double-clicking the data icon in the Project Tree or Process Flow. Then click the

Properties icon  on the workspace toolbar to open the table Properties window.

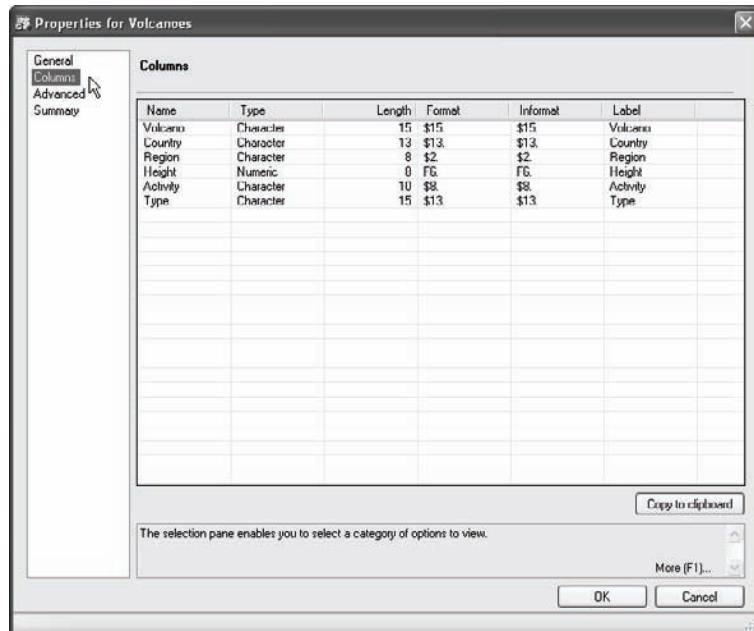
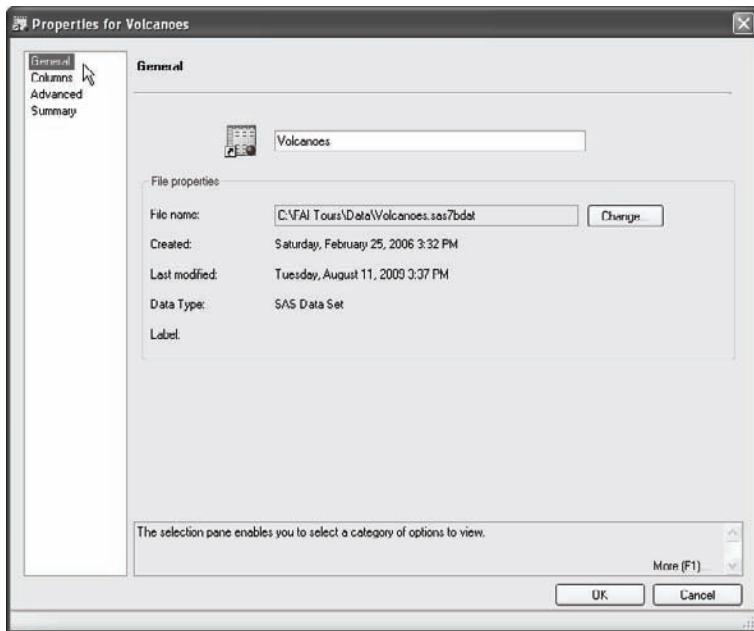


The screenshot shows the SAS Enterprise Guide interface with the title bar "SAS Enterprise Guide - Chapter1.egp". The menu bar includes File, Edit, View, Tasks, Program, Tools, Help, and a toolbar with various icons. The "Project Tree" pane on the left shows a "Process Flow" node expanded, with a "Volcanoes" data icon selected. The "Server List" pane on the right shows connections to "Servers", "OLAP Servers", and "Private OLAP Servers". The main workspace displays a data grid titled "Volcanoes" with the following columns: Volcano, Country, Region, Height, Activity, and Type. The "Type" column header has a dropdown arrow pointing down, and the "Properties" icon (a document with a checkmark) is highlighted with a red box. The data grid contains 21 rows of volcano information, such as Altar (Ecuador, SA, 5321, Extinct, Stratovolcano) and Mauna Loa (USA, AP, 4170, Active, Shield). The bottom status bar shows "Ready" and "Connection: sasdemo, 127.0.0.1".

You can also right-click a data icon in the Project Tree or Process Flow, and select Properties from the pop-up menu.

General page

When the table Properties window opens, it displays the General page. The General page lists basic information about the table: its name, when it was created and last modified, and whether it is a SAS data table or some other type of file.



Columns page If you click **Columns** in the selection pane on the left, the Columns page will open. Here, SAS Enterprise Guide displays information about each column: its name, type, length, format, informat, and label. You cannot change the properties of columns in the Properties window for a data table. To make changes, use the Properties window for an individual column as described in the next section.

1.9 Properties of Columns

The column Properties window displays properties for an individual column. You can use this window inside a task to change labels and display formats, but those changes will apply only to the results of that task rather than the original data table. However, if you open the column Properties window inside a Data Grid, then any changes you make will be saved with the data table.

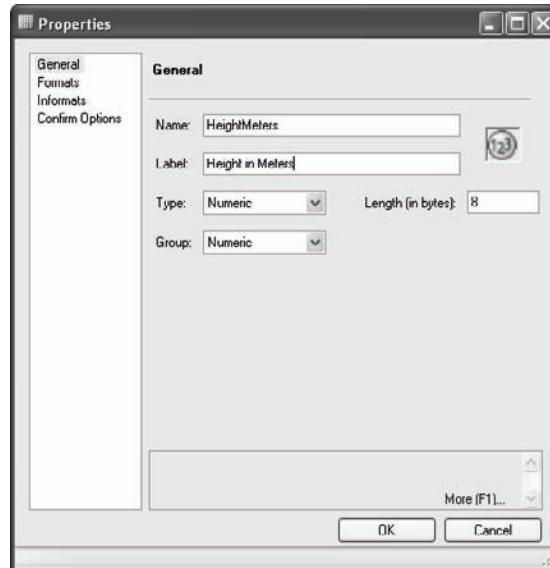
Setting the update mode The Data Grid opens in read-only mode. In this mode you cannot edit the data, and you cannot change column properties. To switch to update mode, select **Edit ▶ Protect Data** from the menu bar. This toggles the data table from read-only to update mode. To return to read-only mode, select **Edit ▶ Protect Data** again.

Opening the Properties window To open the column Properties window, right-click the header of a column and select **Properties** from the pop-up menu. In this Data Grid, **Properties** is being selected for the column Height.

Volcano	Country	Region	Height	Type
Altar	Ecuador	SA	532	Cut
Arthur's Seat	UK	Eu	25	Copy
Barren Island	India	As	35	Paste
Elbrus	Russia	Eu	5033	Hide
Erebus		An	379	Show
Etna	Italy	Eu	3356	Hold
Fuji	Japan	As	3776	Free
Gauthier	Canada	NA	2671	Delete
Grimsvoth	Iceland	Eu	1721	Insert Column...
Ilmorri	Bolivia	SA	6456	Width...
Kenya	Kenya	Af	5105	Properties
Kilaeed	USA	AP	1222	
Kilimanjaro	Tanzania	Af	5895	Stratovolcano
Kluchevskoi	Russia	As	4835	Active
Krakatau	Indonesia	As	813	Caldera
Lassen	USA	NA	3187	Stratovolcano
Mauna Loa	USA	AP	4170	Shield
Nyamuragira	DRCongo	Af	3050	Shield

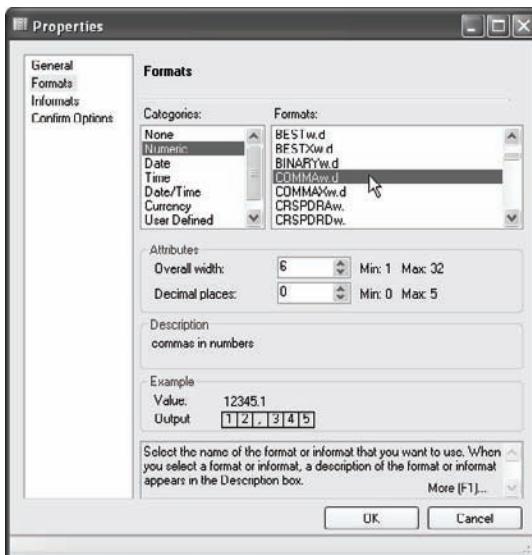
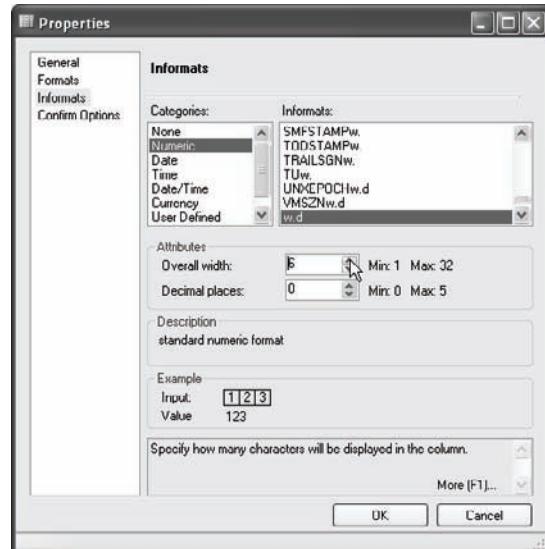
General page The Properties window has several pages. If there is no selection pane on the left, then the data table is in read-only mode and you need to switch to update mode.

The General page displays basic information for the column: its name, label, type, group, and length. You can change any of these properties. In this example, the column name has been changed to **HeightMeters**, and the label to **Height in Meters**. This column is **numeric** and has a length of 8.



Informats page Click **Informats** in the selection pane on the left to open the **Informats** page. **Informats** (also called input formats or read-in formats) tell SAS Enterprise Guide how to interpret input data. There are different informats for character, numeric, date, time, and currency data. In this example, the column uses the default numeric informat, *w.d*, with a width of 6 and no decimal places. This informat can be written as 6.0. See the next section for a table of commonly used informats.

In SAS Enterprise Guide 4.2, you can use informats when you import data files, and when you write SAS programs. However, informats are not used when you type data values into a Data Grid. Instead, the Data Grid uses the data type and data group that you specify to determine how to interpret any data values you enter.



Formats page Click **Formats** in the selection pane on the left to open the **Formats** page. **Formats** (also called display formats) tell SAS Enterprise Guide how data should look in Data Grids or reports. There are different formats for character, numeric, date, time, and currency data. In this example, the format **COMMAw.d** with a width of 6 and no decimal places has been selected. This format can be written as **COMMA6.0**. See section 1.11 for a table of commonly used formats.

	Volcano	Country	Region	HeightMeters	Activity	Type
1	Altar	Ecuador	SA	5,321	Extinct	Stratovolcano
2	Arthur's Seat	UK	Eu	251	Extinct	
3	Barren Island	India	As	354	Active	Stratovolcano
4	Elbrus	Russia	Eu	5,633	Extinct	Stratovolcano
5	Erebus		An	3,794	Active	Stratovolcano
6	Ftna	Italy	Fri	3,350	Active	Stratovolcano

Results Here is the Data Grid showing the new name, **HeightMeters**, and the format with commas.

1.10 Selected Informats

SAS informats (also called input formats or read-in formats) tell SAS Enterprise Guide how to interpret input data. You can specify informats when you import data, in a SAS program, or in a Data Grid. However, in SAS Enterprise Guide 4.2, informats are not used to interpret data that you type into a Data Grid. Here are a few of the many informats available in SAS Enterprise Guide.

Informat	Definition	Width range	Default width
Character			
\$w.	Reads character data—trims leading blanks	1–32,767	none
\$UPCASEw.	Converts character data to uppercase	1–32,767	8
Date, Time, and Datetime¹			
ANYDTDTEw.	Reads dates in any form—when dates are ambiguous, uses the DATESTYLE system option to determine	5–32	9
DATEw.	Reads dates in the form: <i>ddmonyy</i> or <i>ddmonyyyy</i>	7–32	7
DATETIMEw.	Reads datetime values in the form: <i>ddmonyy hh:mm:ss.ss</i>	13–40	18
DDMMYYw.	Reads dates in the form: <i>ddmmyy</i> or <i>ddmmmyyyy</i>	6–32	6
JULIANw.	Reads Julian dates in the form: <i>yyddd</i> or <i>yyyyddd</i>	5–32	5
MMDDYYw.	Reads dates in the form: <i>mmddyy</i> or <i>mmdyyyyy</i>	6–32	6
TIMEw.	Reads time in the form: <i>hh:mm:ss.ss</i> (hours:minutes:seconds—24-hour clock)	5–32	8
Numeric			
w.d	Reads standard numeric data	1–32	none
COMMAw.d	Removes embedded commas and \$, converts left parentheses to minus sign	1–32	1
PERCENTw.	Converts percentages to proportions	1–32	6

¹ SAS date values are the number of days since January 1, 1960. Time values are the number of seconds past midnight, and datetime values are the number of seconds past midnight on January 1, 1960.

The examples below show input data and resulting data values for each informat. The results shown are unformatted data values. See sections 3.1 and 3.2 for information about assigning display formats.

Informat	Input data	Results	Input data	Results
Character				
\$10.	Lassen	Lassen	St. Helens	St. Helens
\$UPCASE10.	Lassen	LASSEN	St. Helens	ST. HELENS
Date, Time, and Datetime				
ANYDTDTE10.	01jan1961 1961001	366 366	31.01.1961 01/31/61	396 396
DATE9.	1jan1961	366	31 jan 61	396
DATETIME14.	1jan1960 10:30	37800	1jan1961 10:30	31660200
DDMMYY10.	01.01.1961	366	31/01/61	396
JULIAN7.	1961001	366	61031	396
MMDDYY10.	01-01-1961	366	01/31/61	396
TIME8.	10:30	37800	10:30:15	37815
Numeric				
5.1	1234	123.4	-12.3	-12.3
COMMA10.0	\$1,000,001	1000001	(1,234)	-1234
PERCENT5.	5%	0.05	(20%)	-0.2

1.11 Selected Standard Formats

SAS formats (also called display formats) tell SAS Enterprise Guide how to display or print data. You can apply formats in a column Properties window in a Data Grid, a task, or a query. Here are a few of the many formats available in SAS Enterprise Guide.

Format	Definition	Width range	Default width
Character			
\$UPCASE <i>w.</i> \$ <i>w.</i>	Converts character data to uppercase Writes standard character data—default for character data	1–32767 1–32767	Length of variable or 8 Length of variable or 1
Date, Time, andDatetime¹			
DATE <i>w.</i>	Writes SAS date values in form <i>ddmonyy</i> or <i>ddmonyyyy</i>	5–9	7
DATETIME <i>w.d</i>	Writes SAS datetime values in form <i>ddmmmyy:hh:mm:ss.ss</i>	7–40	16
DTDATE <i>w.</i>	Writes SAS datetime values in form <i>ddmonyy</i> or <i>ddmonyyyy</i>	5–9	7
EURDFDD <i>w.</i>	Writes SAS date values in form <i>dd.mm.yy</i> or <i>dd.mm.yyyy</i>	2–10	8
JULIAN <i>w.</i>	Writes SAS date values in Julian date form <i>yyddd</i> or <i>yyyyddd</i>	5–7	5
MMDDYY <i>w.</i>	Writes SAS date values in form <i>mm/dd/yy</i> or <i>mm/dd/yyyy</i> —default for dates	2–10	8
TIME <i>w.d</i>	Writes SAS time values in form <i>hh:mm:ss.ss</i> —default for times	2–20	8
WEEKDATE <i>w.</i>	Writes SAS date values in form <i>day-of-week, month-name dd, yy</i> or <i>yyyy</i>	3–37	29
WORDDATE <i>w.</i>	Writes SAS date values in form <i>month-name dd, yyyy</i>	3–32	18
Numeric			
BEST <i>w.</i>	SAS System chooses best format—default format for numeric data	1–32	12
COMMA <i>w.d</i>	Writes numbers with commas	2–32	6
DOLLAR <i>w.d</i>	Writes numbers with a leading \$ and commas separating every three digits—default for currency	2–32	6
E <i>w.</i>	Writes numbers in scientific notation	7–32	12
EUROX <i>w.d</i>	Writes numbers with a leading € and periods separating every three digits	2–32	6
PERCENT <i>w.d</i>	Writes numeric data as percentages	4–32	6
	Writes standard numeric data	1–32	none

¹SAS date values are the number of days since January 1, 1960. Time values are the number of seconds past midnight, and datetime values are the number of seconds past midnight on January 1, 1960.

The examples below show unformatted data values and formatted results for each display format.

Format	Data value	Results	Data value	Results
Character				
\$UPCASE10. \$6.	Lassen Lassen	LASSEN Lassen	St. Helens St. Helens	ST. HELENS St. He
Date, Time, and Datetime				
DATE9.	366	01JAN1961	396	31JAN1961
DATETIME16.	37800	01JAN60:10:30	2629800	31JAN60:10:30
DTDATE9.	37800	01JAN1960	2629800	31JAN1960
EURDFDD10.	366	01.01.1961	396	31.01.1961
JULIAN7.	366	1961001	396	1961031
MMDDYY10.	366	01/01/1961	396	01/31/1961
TIME8.	37800	10:30:00	37815	10:30:15
WEEKDATE15.	366	Sun, Jan 1, 61	396	Tue, Jan 31, 61
WORDDATE12.	366	Jan 1, 1961	396	Jan 31, 1961
Numeric				
BEST10.	1000001	1000001	-12.34	-12.34
BEST6.	1000001	1E6	100001	100001
COMMA12.2	1000001	1,000,001.00	-12.34	-12.34
DOLLAR13.2	1000001	\$1,000,001.00	-12.34	\$-12.34
E10.	1000001	1.000E+06	-12.34	-1.234E+01
EUROX13.2	1000001	€1.000.001,00	-12.34	€-12,34
PERCENT9.2	0.05	5.00%	-1.20	(120.00%)
10.2	1000001	1000001.00	-12.34	-12.34

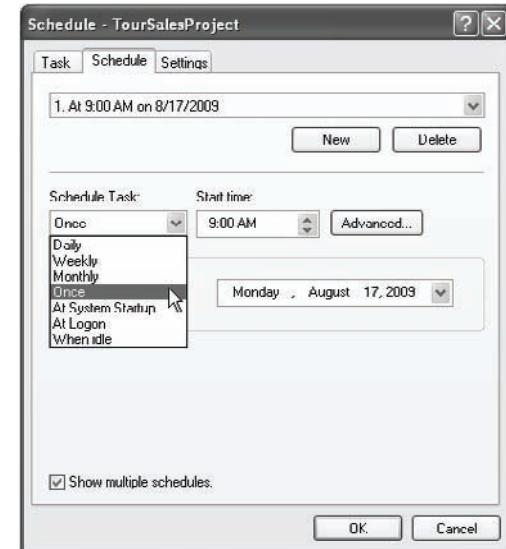
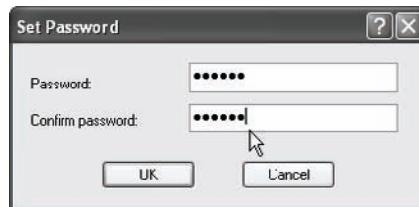
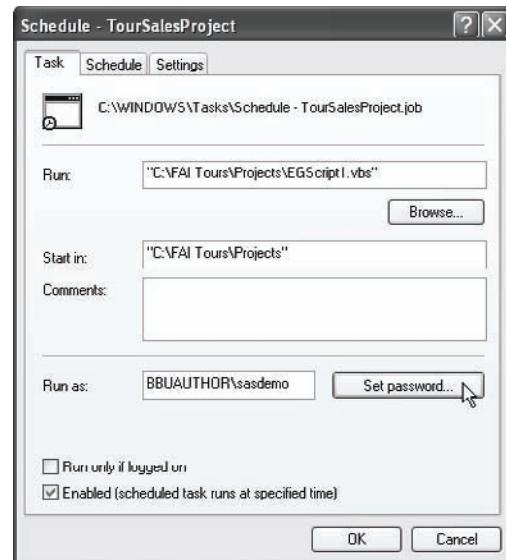
1.12 Scheduling Projects to Run at Specific Times

Sometimes you may want to create a project now, but run it later. For example, if you have data files that are updated on a regular basis, you might want to automatically rerun the project once a week using the new data. Or, if your data files are very large, you might want to run your projects at night so that SAS Enterprise Guide is not using valuable resources during work hours.

Opening the Schedule window You can schedule a complete project or just a process flow. To schedule a project, select **File ► Schedule project-name** from the menu bar.

To schedule a process flow, right-click the name of the process flow in the Project Tree and select **Schedule process-flow-name** from the pop-up menu. This opens the Microsoft Windows Task Scheduler with the Task tab on top. When you schedule a project, SAS Enterprise Guide creates a script that is saved in a file on your computer. The name and path of this script is displayed in the **Run** box. The **Start in** box displays the folder in which the script will run. Your computer and user name are displayed in the **Run as** box.

If you will not be logged on at the time the project runs, then make sure the box next to **Run only if logged on** is unchecked, and click the **Set password** button to open the Set Password window. Enter the password for your user name (the same password you use when you log on to your computer), and click **OK**.



Setting the run frequency To tell SAS Enterprise Guide when to run the project, click the **Schedule** tab. Click the **New** button. Then select the frequency to run the project from the drop-down list under **Schedule Task**. You can schedule your project to run just once at a specified time as shown here, or you can schedule your project to run on a regular basis.

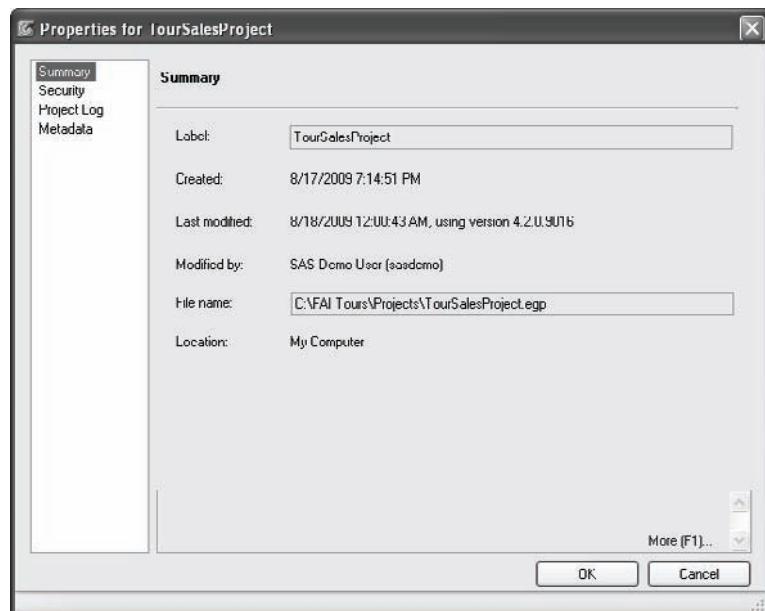
Setting the date and time To set the time the project will start running, click the up and down arrows on the **Start time** box, or simply click the time and type a new value. To choose a date other than today, click the down-arrow in the **Run on** box and select a date from the pull-down calendar.

Other settings If you click the **Settings** tab, you will see other options, including the maximum length of time a project will be allowed to run, and whether it will run if your computer has gone into sleep mode.

When you are satisfied with all the settings, click **OK** to schedule the project.

Running the project The project will not run if it is open or if the computer is turned off at the time the project is scheduled to run. However, if you have a different project open, the scheduled project will still run.

Viewing the results To see the results of your scheduled run, open the project after it has completed running. If you are not sure whether a project ran, you can confirm this by opening the Properties window for that project. To open the Properties window for a project, select **File ▶ Project Properties** from the menu bar. The **Last modified** field shows the date and time that the project last ran.

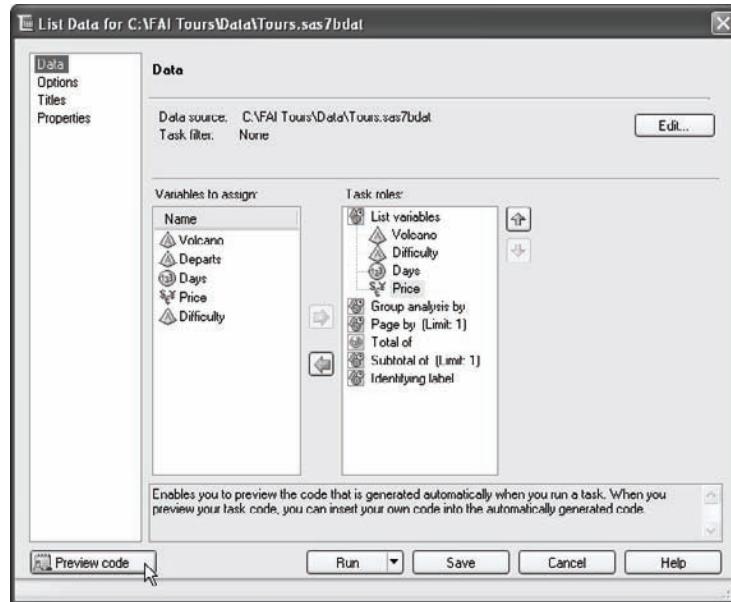


1.13 Editing SAS Programs Generated by a Task

If you are a SAS programmer, you may want to make a few changes to the programs generated by SAS Enterprise Guide. There is more than one way to do this. You can insert your own SAS code into the program associated with a task, or you can save the code generated by a task in a separate file which you can then edit and run.

Previewing code generated by a task

Many task windows have a **Preview code** button in the lower-left corner. If you click this button, SAS Enterprise Guide will open a Code Preview window displaying the code that SAS Enterprise Guide has written for that task.

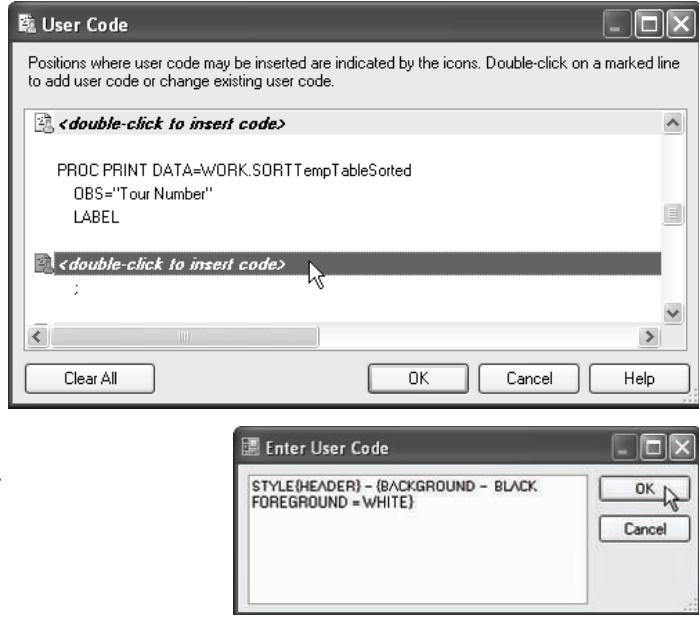


```
Code Preview for Task
Insert Code... 
TITLE1 "Fire and Ice Tours";
FOOTNOTE;

PROC PRINT DATA=WORK.SORTTempTableSorted
OBS="Tour Number"
LABEL
;
VAR Volcano Difficulty Days Price;
RUN;
/* -----
   End of task code.
-----
RUN; QUIT;
```

Inserting code in a task Here is a Code Preview window for a List Data task. You can see that it uses PROC PRINT. If you want to add code to the task, click the **Insert Code** button. This opens a User Code window. You cannot edit the existing code generated by a task, but the User Code window allows you to add code at specific points in the program.

In the User Code window, double-click **<double-click to insert code>** at the point where you wish to add your own custom code. An Enter User Code window will open. Type the custom code you wish to add. When you are done, click **OK**. Your new code will appear in the User Code window. Click **OK** in the User Code window. When you run the task, SAS Enterprise Guide will run the code you inserted along with the code generated by the task.



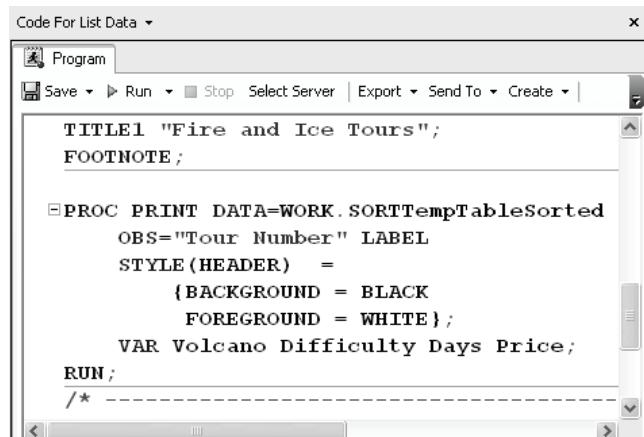
Editing code generated by a task If you want to be able to edit the entire program generated by a task, or code from tasks that do not have a Preview code button, you can make a copy of the program, and then edit it. To do this, run the task, and then right-click the task icon in the Project Tree or Process Flow, and select **Add As Code Template**. SAS Enterprise Guide will open a Program window containing the code generated by the task.

You can edit this code in any way you wish. Because this code is a copy of the code generated by the task, any changes you make here will not affect the task, nor will any changes you make to the task be reflected in this code.

When you have made all the changes you wish and are ready to run the program, click **Run** on the workspace toolbar for the Program window. Your program will run on the server that has been set as your default. To choose a different server, click

Select Server. If you decide you want to stop a program while it's running, click the Stop button on the workspace toolbar for the Program window. You can also use the Program menu on the menu bar to control execution of your program.

Programs created in this way are embedded in your project, and are not saved as separate files. For more information on embedding programs, see the next section.



1.14 Writing and Running Custom SAS Programs

You can accomplish a lot using tasks in SAS Enterprise Guide, but sometimes you may need to do something for which there is no predefined task. At those times, you can run a SAS program that was written outside SAS Enterprise Guide, or you can write a new one.

Writing a new SAS program

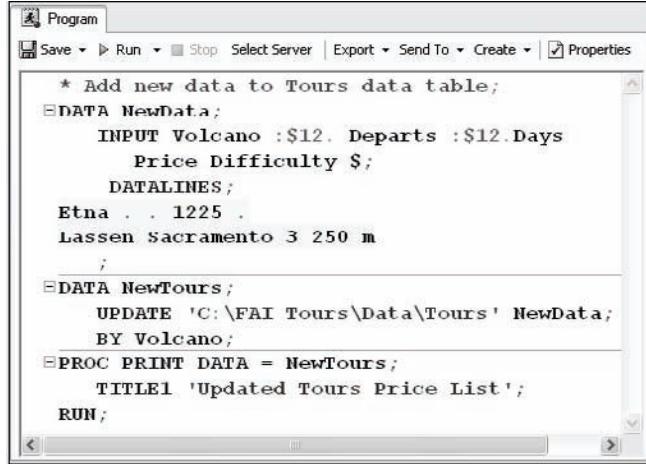
To create a new SAS program, open an empty Program window by selecting **File ▶ New ▶ Program** from the menu bar. A Program window will open in the workspace. The program editor in SAS Enterprise Guide is syntax-sensitive, which means that SAS keywords are displayed in blue, comments are green, quoted strings are magenta, and so forth.

Opening an existing SAS program

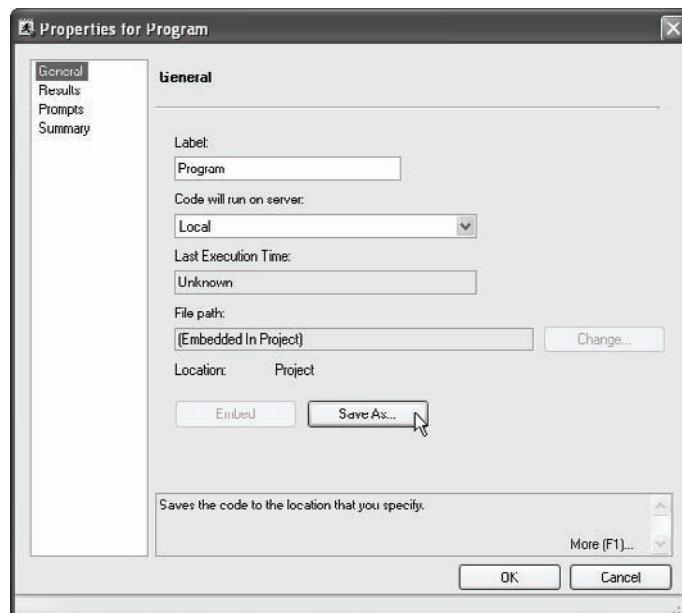
If you have existing SAS programs that you want to include in your project, you can open them by selecting **File ▶ Open ▶ Program** from the menu bar. Navigate to the existing SAS program and click **Open**. This opens a Program window in the workspace, where you can edit the program.

Saving a program in a file

Any new programs you write are automatically embedded in your project. This means that the program's code does not exist in a file outside of the project. To save a SAS program outside its project, click **Save** on the workspace toolbar for the Program window, or right-click the program icon in the Project Tree or Process Flow and select **Save program-name As** from the pop-up menu. You can also save a program from the Properties for *program-name* window. To view the properties of a program, click **Properties** on the workspace toolbar for the Program window, or right-click the program icon in the Project Tree or Process Flow and select **Properties** from the pop-up menu. Then in the General page, click **Save As**.



```
* Add new data to Tours data table;
DATA NewData;
  INPUT Volcano :$12. Departs :$12. Days
        Price Difficulty $;
  DATALINES;
Etna . . 1225 .
Lassen Sacramento 3 250 m
;
DATA NewTours;
  UPDATE 'C:\FAI Tours\Data\Tours' NewData;
  BY Volcano;
PROC PRINT DATA = NewTours;
  TITLE1 'Updated Tours Price List';
RUN;
```



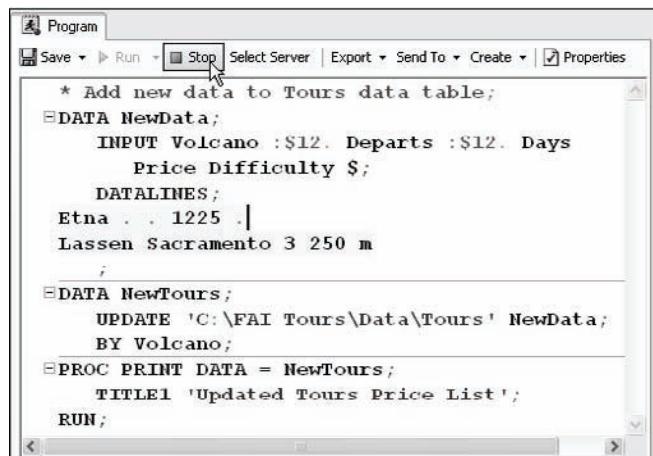
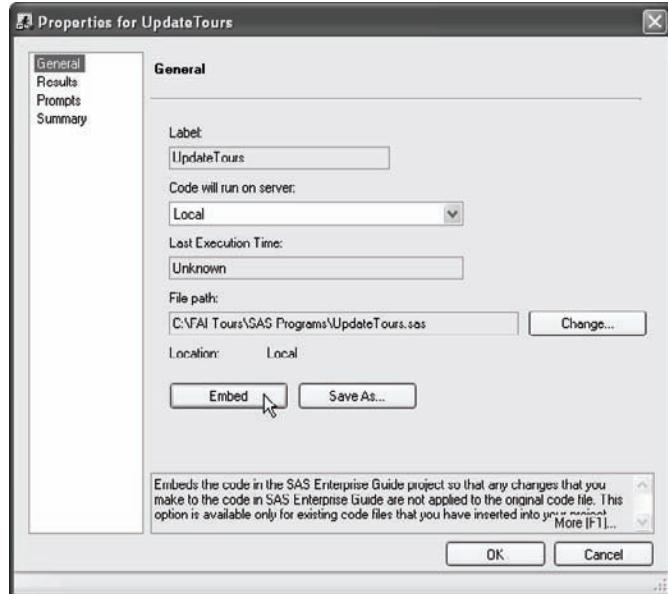
If you save the

program in a file, then it is not embedded, and any changes you make to it in SAS Enterprise Guide will be saved in the file rather than as part of your project. The icon for a program saved in a file includes a little arrow indicating that the project contains a shortcut to the program rather than

the actual program .

Embedding a program in a project When you open a SAS program that has been saved in a separate file, it is not automatically embedded in your project. If you want to embed the program in your project, then open the Properties window for the program and click **Embed**. After you embed the program, any changes you make to it in SAS Enterprise Guide will be saved as part of your project rather than in the separate file. The icon for embedded code looks like

this .



Running your program

When you are ready to run your program, click **Run** on the workspace toolbar for the Program window. Your program will run on the server that has been set as your default. To choose a different server, click **Select Server**. If you decide you want to stop a program while it's running, click the **Stop** button  on the workspace toolbar for the Program window. You can also use the Program menu on the main menu bar to control your program.

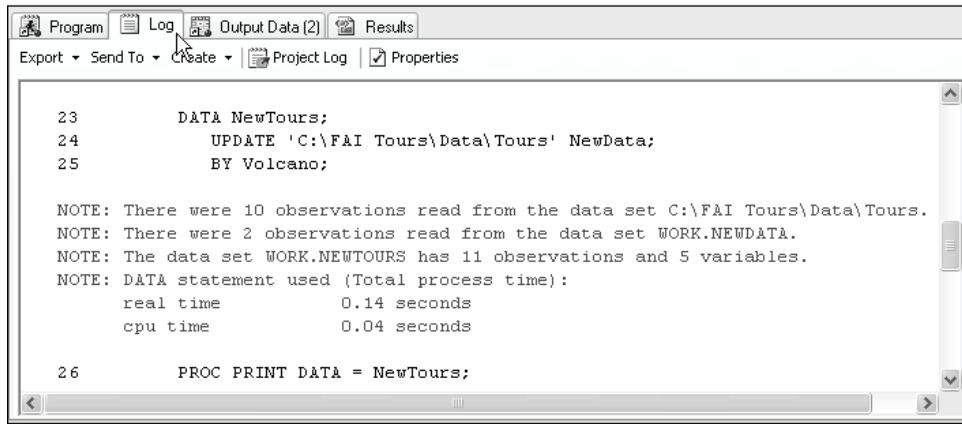
1.15 Viewing Program and Project Logs

A SAS log is a record of what SAS did. Just about everything you do in SAS Enterprise Guide generates a SAS log. Logs contain the actual code that SAS ran, plus any error messages, warnings, or notes.

Different types of logs A program log is the log that is generated when you run a SAS program. Tasks generate logs too, but when you run tasks, you have little need to view the task log. That is because tasks rarely produce errors or warnings. Every time you rerun a program or task, the old log is replaced with a new one.

The Project Log, on the other hand, is a single cumulative record of everything that has been run in a particular project. By default, the Project Log is turned off. Once you turn the Project Log on, nothing disappears from it unless you clear the log.

Viewing a program log After a program runs, the results are displayed in the workspace. To open the program log, click the tab labeled **Log**. Here is a portion of the program log generated by the SAS program in the preceding section.



The screenshot shows the SAS Enterprise Guide interface with the 'Log' tab selected in the top navigation bar. The main window displays a portion of a SAS program log. The log output includes:

```
23      DATA NewTours;
24      UPDATE 'C:\FAI Tours\Data\Tours' NewData;
25      BY Volcano;

NOTE: There were 10 observations read from the data set C:\FAI Tours\Data\Tours.
NOTE: There were 2 observations read from the data set WORK.NEWDATA.
NOTE: The data set WORK.NEWTOURS has 11 observations and 5 variables.
NOTE: DATA statement used (Total process time):
      real time          0.14 seconds
      cpu time          0.04 seconds

26      PROC PRINT DATA = NewTours;
```

One of the first things you will notice when you look at a log is that it contains more lines of SAS code than were in your original program. That is because SAS Enterprise Guide adds housekeeping statements to the beginning and end of your program to make sure that it runs properly when it is passed to your SAS server.

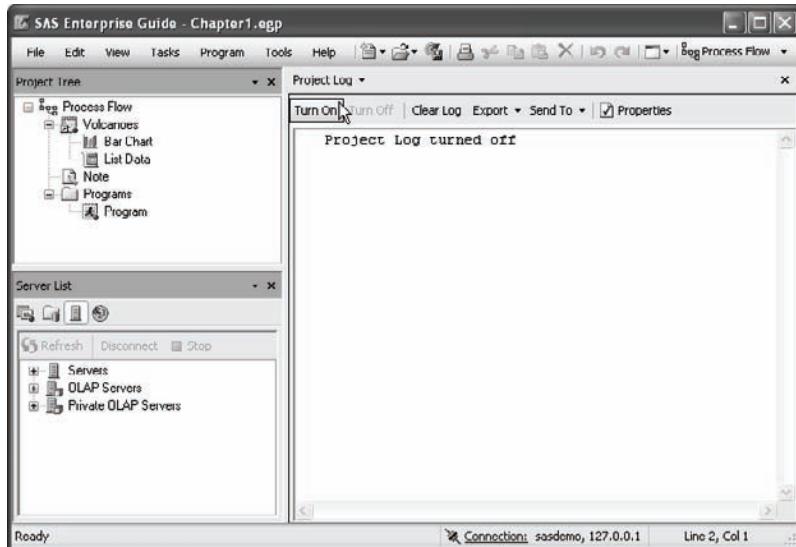


If your program contains any errors, its icon will include a red X. Programs that contain

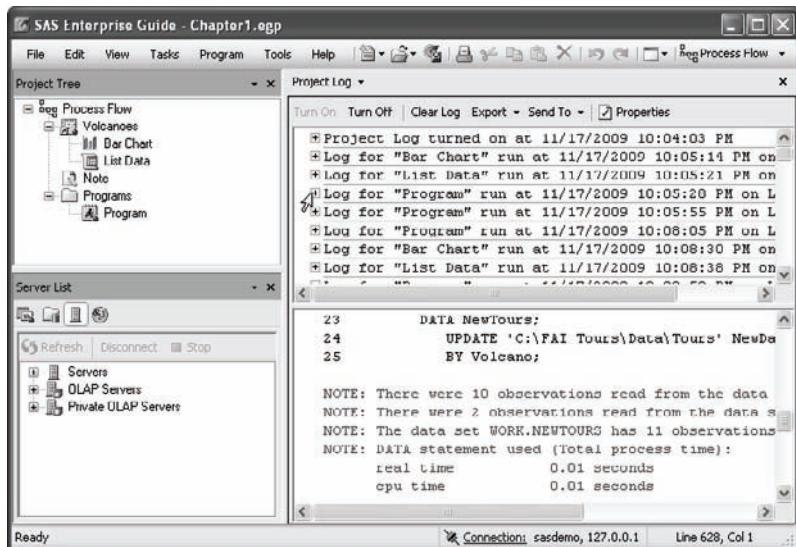


warnings (but no errors) have icons with yellow triangles. Even if there are no errors or warnings, it is a good habit to check the program log when you write your own SAS programs. Just because a program runs without errors or warnings does not mean that it produced the correct results.

Viewing the Project Log To turn on the Project Log, first open it by clicking **Project Log** on the workspace toolbar for the Process Flow or selecting **View ▶ Project Log** from the menu bar. Then on the workspace toolbar for the Project Log, click **Turn On**. Once the Project Log is turned on, it will keep a continuous history of everything that runs in that project.



The Project Log includes the date and time when each action occurred. Click the plus sign (+) to expand a section, or the minus sign (-) to collapse it. You can also split the Project Log into two pieces by clicking and dragging the top border (the line just below the workspace toolbar) of the Project Log window.

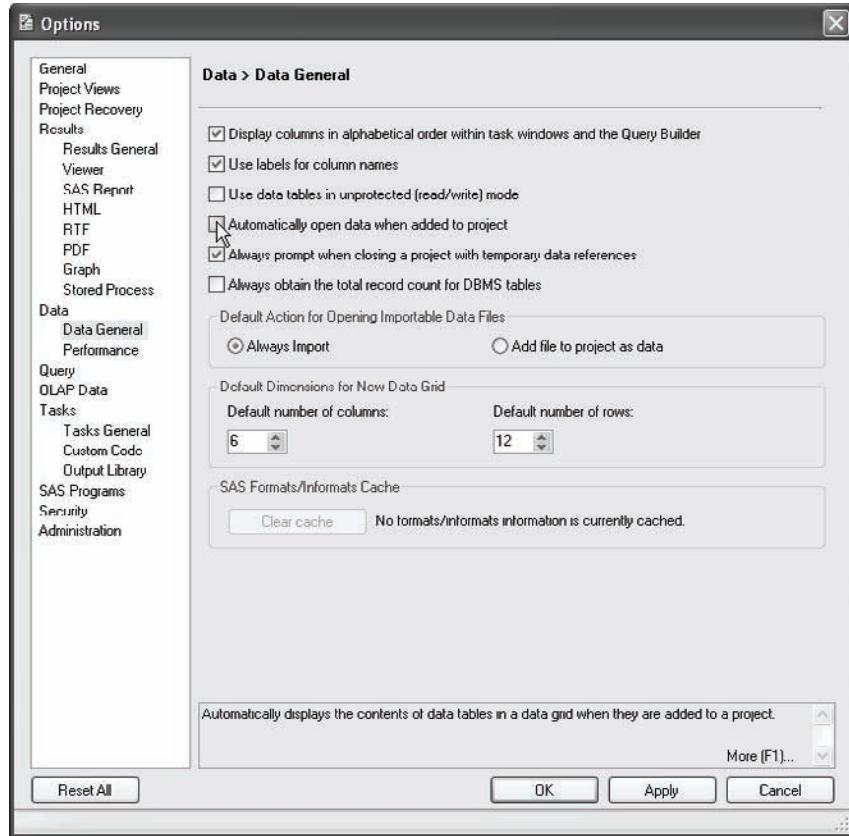


To clear the Project Log, click **Clear Log** on the workspace toolbar for the Project Log. To turn it off, click **Turn Off**.

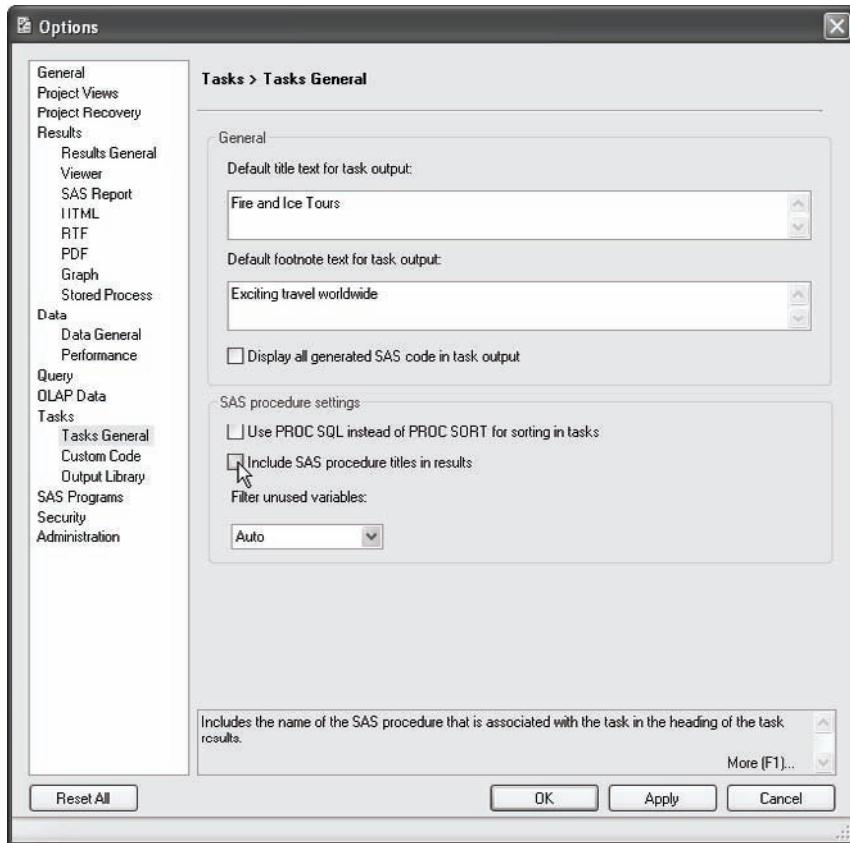
1.16 Using the Options Window

The Options window allows you to change many default behaviors in SAS Enterprise Guide. To open the Options window, select **Tools ▶ Options** from the menu bar.

Changing the way data are handled To see options for data, click **Data General** in the selection pane on the left. If you have large data tables, columns might be easier to find if they are arranged in alphabetical order. To list columns alphabetically in task windows, check the box in front of **Display columns in alphabetical order**. By default, the Data Grid uses column names, not labels, for column headers. To change this, check the box in front of **Use labels for column names**. If you have large data tables on remote servers, you may be able to improve performance by unchecking **Automatically open data when added to project**.



Changing the default titles and footnotes To change the default titles and footnotes, click **Tasks General** in the selection pane on the left. In this page, you can specify new default titles and footnotes, or set them to blank. Some tasks include in the results the name of the SAS procedure used by that task. In these results you will see titles like "The FREQ Procedure" or "The ANOVA Procedure." You can eliminate these titles by unchecking the box labeled **Include SAS procedure titles in results**.



Changing the default result format and style To change the default format for results, click **Results General** in the selection pane on the left to open the Results General page (not shown). Then check all the formats you want to use: **SAS Report**, **HTML**, **PDF**, **RTF**, or **text output**. To change the default style for results, click the name of the format (such as **PDF**) in the selection pane on the left to open a page for that format (not shown), and then select a style. See Tutorial B or Chapter 11 for more about changing result formats and styles.

Running code automatically If you have SAS code that you would like to run automatically, click **SAS Programs** in the selection pane on the left to open the SAS Programs page (not shown). Select an option under the heading **Additional SAS code**. Then click **Edit**, type your code in the Edit window, and click **Save**. The option **Submit SAS code when server is connected** is particularly useful for submitting LIBNAME statements. You can also click **Custom Code** in the selection pane on the left, and specify code to be run before or after tasks.

Restoring the window layout To restore windows to their default layout, click **General** in the selection pane on the left to open the General page (not shown). Then click the **Restore Window Layout** button.

Saving and resetting options To close the Options window and save the changes you have made, click **OK**. Once you set options, they stay in effect for future SAS Enterprise Guide sessions. If at a later time, you decide you want to restore everything in the Options window to the default settings, simply click the **Reset All** button in the lower left corner.

2

“ There is no knowledge that
is not power. ”

RALPH WALDO EMERSON

From the essay “Old Age” in *The Atlantic*, 1862.



CHAPTER 2

Chapter 2

Bringing Data into a Project

- 2.1 Sources of Data **168**
- 2.2 Locations for Data **170**
- 2.3 Assigning Libraries with the Assign Project Library Task **172**
- 2.4 Creating New SAS Data Tables in a Data Grid **174**
- 2.5 Editing SAS Data Tables in a Data Grid **176**
- 2.6 Inserting Computed Columns in a Data Grid **178**
- 2.7 Importing Microsoft Excel Spreadsheets **180**
- 2.8 Importing Delimited Raw Data **182**
- 2.9 Importing Fixed-Column Raw Data **184**
- 2.10 Exporting Data **186**

2.1 Sources of Data

Before you can analyze your data, before you can run a report, before you can do anything with your data, SAS Enterprise Guide must be able to read your data. Your data might be in a data warehouse on a mainframe, or on a piece of paper sitting on your desk. Whatever form your data take, there is a way to get your data into SAS Enterprise Guide.



New SAS data tables If you have a small amount of data, or if you've collected the data yourself, then you may find the easiest way to get your data into SAS Enterprise Guide is to type them directly into a Data Grid. To do this, just open an empty Data Grid like the one below, set the properties for the columns, and enter your data. SAS Enterprise Guide makes a SAS data table (also called a SAS data set) from the data you type into a Data Grid. Sections 2.4 to 2.6 describe Data Grids in more detail.

	A	B	C	D	E	F
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						



Existing SAS data tables You may have SAS data tables that were created in SAS or in another project in SAS Enterprise Guide. To open a SAS data table in SAS Enterprise Guide, select **File ▶ Open ▶ Data** from the menu bar. In the Open Data window, navigate to the location of the SAS data table. Once you have found the data table you want, click **Open**. SAS Enterprise Guide will immediately add the SAS data table to your project and display it in a Data Grid.

	Volcano	Departs	Days	Price	Difficulty
1	Etna	Catania	7	\$1,075	m
2	Fuji	Tokyo	2	\$225	c
3	Kenya	Nairobi	6	\$830	m
4	Kilauea	Hilo	1	\$55	e
5	Kilimanjaro	Nairobi	9	\$1,310	c
6	Krakatau	Jakarta	7	\$895	e
7	Poas	San Jose	1	\$65	e
8	Reventador	Quito	4	\$575	m
9	St. Helens	Portland	2	\$167	e
10	Vesuvius	Rome	6	\$985	e



Raw data files Raw data files are files that contain no special formatting. They are sometimes called text, ASCII, sequential, or flat files and can be viewed using a simple text editor, such as Microsoft Notepad. SAS data tables and Microsoft Excel files are not raw data files. If you open a spreadsheet or SAS data table in Microsoft Notepad, you'll see lots of strange characters that Microsoft Notepad simply can't interpret.

SAS Enterprise Guide can read just about any type of raw data file including delimited data files and fixed-column data files. In delimited data files, a delimiter separates the data values. CSV (comma-separated values) files use commas as delimiters. Other files may use a different delimiter such as a tab, semicolon, or space. Fixed-column data files are similar to delimited data files, but instead of having a delimiter separating the data values, the data values are lined up in tidy vertical columns. Reading raw data files is described in more detail in sections 2.8 and 2.9.



Other software files SAS Enterprise Guide can read files produced by many other types of software. When you install SAS Enterprise Guide, you get everything you need to read most PC data files. You do not need to install any additional software to read data in these formats:

- dBASE files
- HTML files
- IBM Lotus 1-2-3 files
- Microsoft Access files
- Microsoft Excel files
- Microsoft Exchange files
- Paradox files

However, if you have large PC data files, you may be able to improve performance by using SAS/ACCESS software. To read files this way, you must have SAS/ACCESS (either SAS/ACCESS Interface to PC Files or SAS/ACCESS Interface to ODBC, depending on the type of data files you are reading), and it must be installed on the same computer where SAS is installed. Then you select the option **Import the data using SAS/ACCESS Interface to PC Files whenever possible** in the Import wizard. Section 2.7 shows an example of reading a Microsoft Excel file.



If you have SAS/ACCESS Interface to PC Files software, you can also read data in these formats:

- JMP files
- SPSS save files in Microsoft Windows format (with a .sav extension)
- Stata files in Microsoft Windows format (with a .dta extension)

To read these files, select **Tasks ▶ Data ▶ Import *data-type* file** from the menu bar. A task icon will be added to your project, and the imported table will be saved in your WORK library. The WORK library is erased when you exit SAS Enterprise Guide. To recreate the table at a later time, just rerun the task.

SAS Enterprise Guide can read many other kinds of database files, including Oracle and DB2. To read these other database files, you must have the corresponding SAS/ACCESS product (such as SAS/ACCESS Interface to Oracle or SAS/ACCESS Interface to DB2) installed on the same computer where SAS is installed. Then you define a SAS data library on that computer using the corresponding SAS/ACCESS engine. The SAS data library tells SAS Enterprise Guide where to find the data file and how to read it.

2.2 Locations for Data

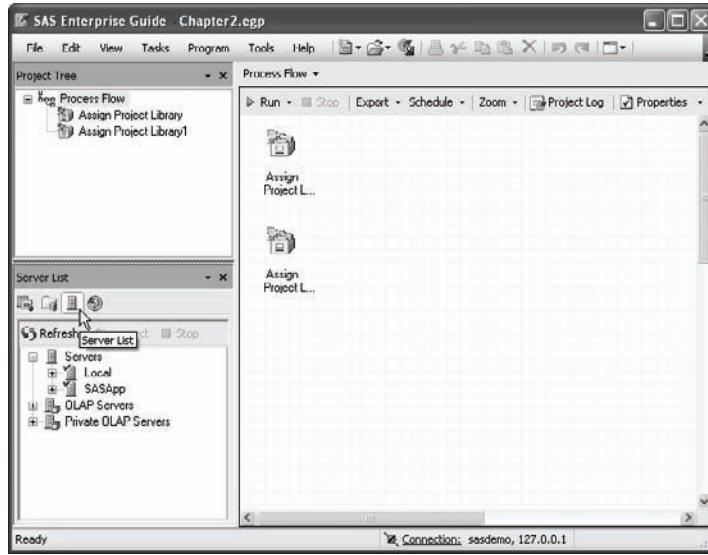
Before you can read a data file, you must tell SAS Enterprise Guide where to find it. For any particular file, there may be several ways to do this.

Server List window

Most types of files can be brought into a project using the Server List window. This window appears in the Resources pane. To view the Server List window, select **View ▶ Server List** from the menu bar, or click the

Server List icon  in the Resources pane.

A SAS server is any computer running SAS software. The Server List window lists all the SAS servers that are available to you, and the files and SAS data libraries on those servers. To see more detail for any part of the list (such as **Libraries** or **Files**) click the plus signs (+).

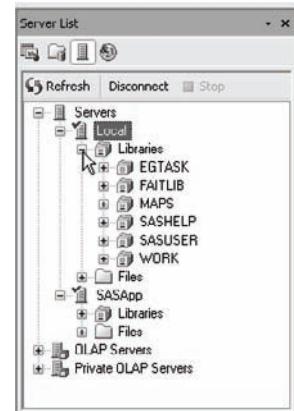


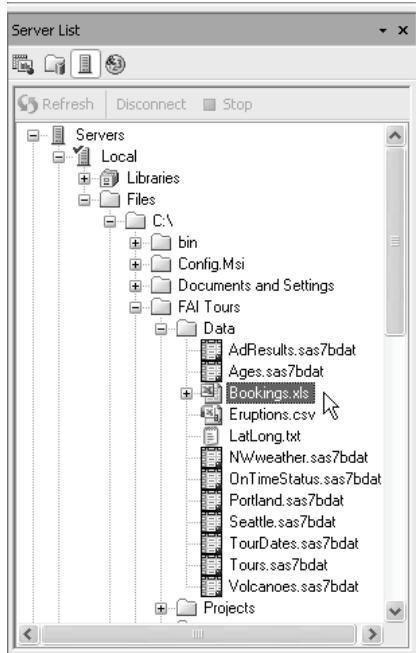
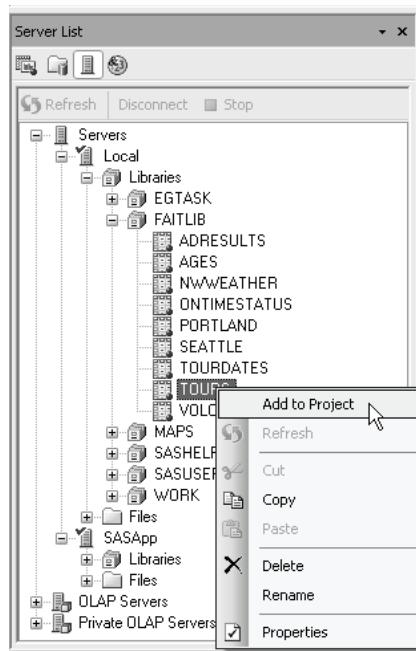
Libraries A SAS data library is a set of SAS files residing in a particular location. Instead of referring to the location by its full path, you identify it by a short name, called a libref. Here are some libraries you are likely to see:

WORK is a special library for temporary data tables. The WORK library is erased when you exit SAS Enterprise Guide.

SASUSER is the default library for output data tables and is permanent on most systems, so your data tables will not be erased by SAS Enterprise Guide. However, on z/OS and some UNIX systems, SASUSER may be temporary. In some environments, SASUSER may be read-only. If you have more than one SAS server, then you may have a SASUSER library on each server.

EGTASK is another special library. If a library with this name has been defined on your system, then SAS Enterprise Guide will use it as the default for output data sets instead of SASUSER.





You may have other libraries that have been defined for your use. This Server List window shows that a library named FAITLIB has been defined.

If you want to define your own libraries, there are several ways to do that. These include using the Assign Project Library task (see the next section), using the SAS Enterprise Guide Explorer, and submitting a LIBNAME statement from a Program window or via the OPTIONS window (see section 1.16).

To view the data sets in a SAS data library, click the plus sign (+) next to the library's name. You can add a data set to your project by clicking its name and dragging it into the Process Flow, or by right-clicking its name and selecting **Add to Project** from the pop-up menu. In this window, the Tours data set is being added to the project.

Files If you click **Files** in the Server List window, SAS Enterprise Guide will list the folders or directories on that SAS server. From there, you can navigate through the server's directory system to find the file you want. These files may be SAS data tables or other types of files such as Excel spreadsheets. The file extension for a standard SAS data table in Windows is .sas7bdat.

You can add a file to your project by clicking its name and dragging it into the Process Flow, or by right-clicking its name and selecting **Add to Project** from the pop-up menu.

Notice that the SAS data tables in the library FAITLIB are also listed under **Files**. That is because the FAITLIB data library points to that Windows subdirectory. If you add a SAS data table from **Files**, SAS Enterprise Guide will automatically define a library for you and give it a name like ECLIB000.

Other ways to bring data into a project
Most types of files can be brought into a project by selecting either **File ▶ Open ▶ Data**, or **File ▶ Import Data** from the menu bar. For some types of files, opening the data produces the same result as importing it. For other types of files, importing produces a different result. See sections 2.7 to 2.9 for more information about importing Microsoft Excel and raw data files.

To bring JMP, SPSS, or Stata data files into a project, you must import them. See the preceding section for a description of how to do this.

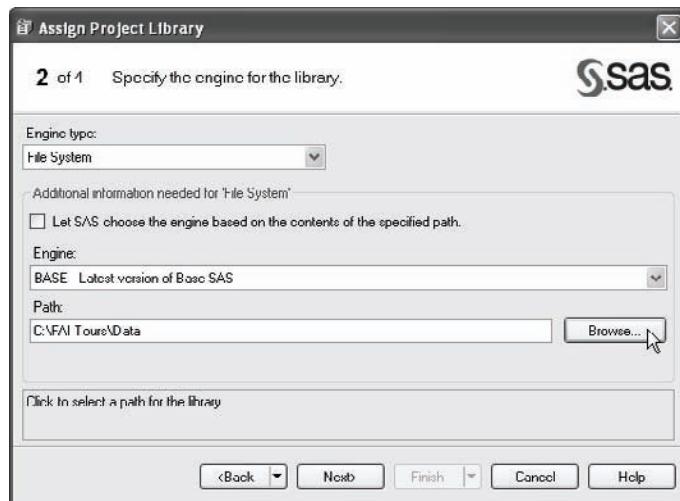
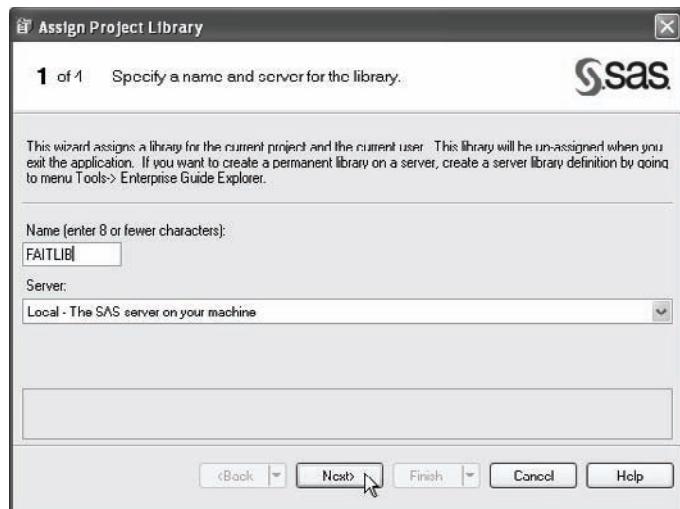
2.3 Assigning Libraries with the Assign Project Library Task

A SAS data library is a set of SAS files stored in a particular location. You use SAS data libraries to tell SAS Enterprise Guide where to find existing data tables, and where to save new ones. This section describes how to define SAS data libraries using the Assign Project Library task.

Assign Project Library wizard To define a project library, select **Tools ▶ Assign Project Library** from the menu bar. The Assign Project Library wizard will open. This wizard has four windows.

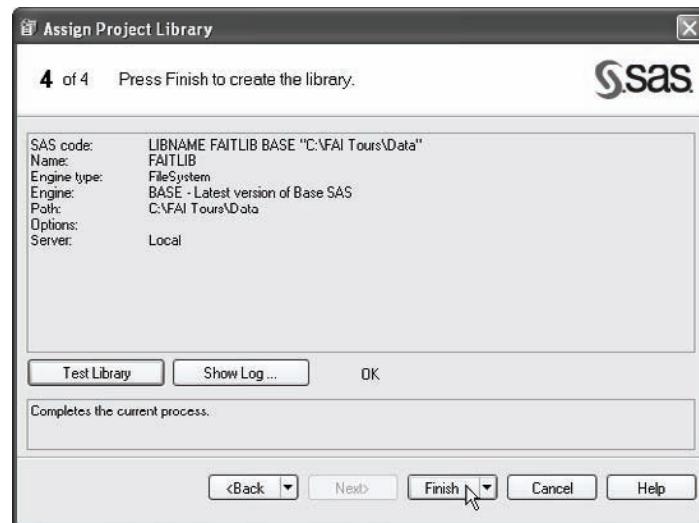
In the first window, type a name for your new library. The library name cannot be longer than eight characters; must start with a letter or underscore; and can contain only letters, numerals, and underscores. This name is called a libref, and it is like a nickname for your library. Next select the server where your data are stored. In this example, a library named FAITLIB is being created, and it will reside on the Local server. Click **Next** when you are done.

In the second window, you specify the engine. The SAS data engine determines the type of data that will reside in this library. Click the down-arrow on the box labeled **Engine type** to select the general type of engine, and then click the down-arrow on the box labeled **Engine** to select the engine. Depending on the engine type you choose, other options may appear in the lower portion of the window. For ordinary SAS data tables, use an engine type of **File System**, and the **BASE** engine. In the box labeled **Path**, type the path for your library, or click the **Browse** button to navigate to the location of your data. Click **Next** when you are done.

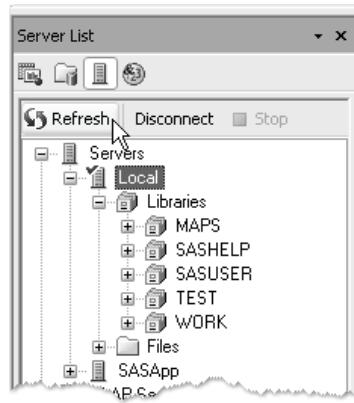


In the third window (not shown), you can specify options for your library.

In the fourth window, you can test your new library by clicking **Test Library**. Click **Finish** to close the wizard and create your library.



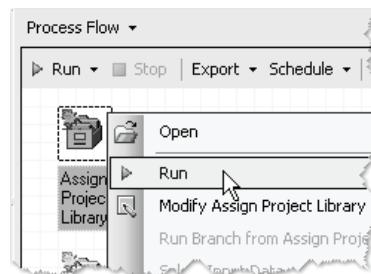
An Assign Project Library icon  will be added to your Project Tree and Process Flow windows. However, the new project library may not automatically appear in the Server List window. If not, then click the server (such as Local), and click **Refresh**.



Rerunning the Assign Project Library task

When you open SAS Enterprise Guide, project libraries are not automatically assigned. Before you can use project libraries (and any data accessed via them), you must rerun any Assign Project Library tasks.

You can rerun an Assign Project Library task by right-clicking its icon in the Project Tree or Process Flow and selecting **Run** from the pop-up menu, or by clicking its icon and selecting **Run ▶ Run Assign Project Library** on the workspace toolbar above the Process Flow.



Alternatively, you can rerun your entire process flow. Click **Run** on the workspace toolbar above the Process Flow, or right-click the name of the process flow in the Project Tree and select **Run process-flow-name** from the pop-up menu. Just be sure that your Assign Project Library icons appear before any tasks that use those libraries. Since projects run from top left to bottom right, you may want to move any Assign Project Library icons to the upper-left corner of the Process Flow.

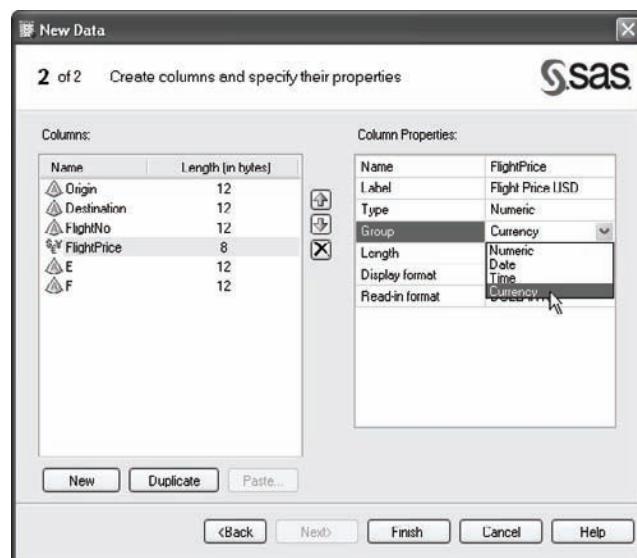
2.4 Creating New SAS Data Tables in a Data Grid

Using the Data Grid, you can browse and edit existing SAS data tables, or create new ones. To open an empty Data Grid, select **File ▶ New ▶ Data** from the menu bar. This opens the New Data wizard, which has two windows.

Name and location In the first window of the New Data wizard, type a name for your new data table in the **Name** box. This is a name that you make up. The name must follow the rules for standard SAS names (32 characters or fewer; start with a letter or underscore; and contain only letters, numerals, and underscores). Next choose a location where the new data table will be stored. This location must be a SAS data library. If a suitable library is not listed, then cancel the New Data wizard and define a new library before starting over (see section 2.3). In this example, the name Seattle has been typed in the Name box and the FAITLIB library has been selected. When you are satisfied, click **Next**.



Column properties In the second window of the New Data wizard, specify the properties for each column in your new data table. At first the columns are named A, B, C, and so on. To display the properties of a particular column, click its name in the list on the left under the heading **Columns**. To change a column property, click its value in the list on the right under the heading **Column Properties**. You can type a name for the column and an optional label. Then specify the data type and data group by clicking the word **Type** or **Group** and selecting from the pull-down list. In the Seattle data table, the fourth column has the name



FlightPrice, a label of Flight Price USD, a data type of Numeric, and a data group of Currency.

To change a read-in format (also called an informat) or a display format (also called a format), click its value under Column Properties.

The ellipsis button will appear . Click the ellipsis button to open a Formats window. The Formats window that opens looks the same regardless of whether you are modifying an informat or a format. See sections 1.10 and 1.11 for tables of commonly used informats and formats.

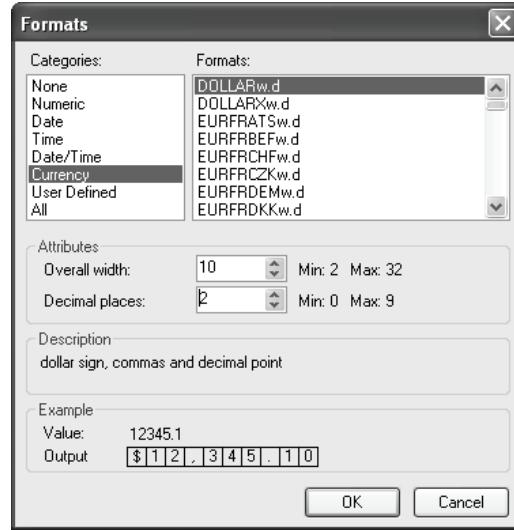
The Formats window shown here is for the format for FlightPrice. The format is set to DOLLARw.d with a width of 10 and 2 decimal places. This format can be written as DOLLAR10.2. When you are satisfied, click OK to return to the second window of the New Data wizard.

You can also open a window for the informat, and any informats you specify will be saved with the column properties. However, in SAS Enterprise Guide 4.2, Data Grids do not use informats to interpret input data. Instead, Data Grids use the data type and data group that you specify to determine how to interpret any data values you enter.

In the second window of the New Data wizard, you can add more columns by clicking **New**. You can delete columns by clicking the name of that column and then clicking the delete button . When you are satisfied with your columns and their properties, click **Finish** to close the wizard and create an empty data table.

Entering data Once you have created the data table, you can begin typing data into it. Any value you enter must fit the data type and data group you specified. You can copy and paste values. To move the cursor, click a cell, or use the tab and arrow keys.

Here is the Seattle data table with column attributes defined, and three rows of data entered.

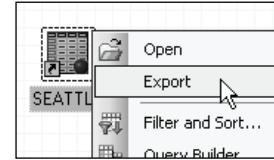


	Origin	Destination	FlightNo	FlightPrice
1	Seattle	Catania	BA48	\$802.00
2	Seattle	Hilo	HA21	\$677.00
3	Seattle	Jakarta	AA119	1815
4				.
5				.
6				.
7				.
8				.
9				.
10				.
11				.
12				.

2.5 Editing SAS Data Tables in a Data Grid

Editing SAS data tables is easy—whether you need to add new rows or columns, fix errors, or update values—but there are a few points to keep in mind.

Copying a SAS data table When you edit a SAS data table in a Data Grid, any changes you make are permanent, even if you don't save the project. Therefore, unless you are absolutely sure about the changes you make, you should make a copy of the data table before editing. To make a copy, right-click the data icon in the Project Tree or Process Flow window and select **Export ▶ Export *table-name*** from the pop-up menu. Choose a location for the new data table and give it a name. Or, open the table in a Data Grid, and select **Export ▶ Export *table-name*** from the workspace toolbar. When you export a data table, your new copy does not appear in your project, so after you export the data table, you must open it in your project.

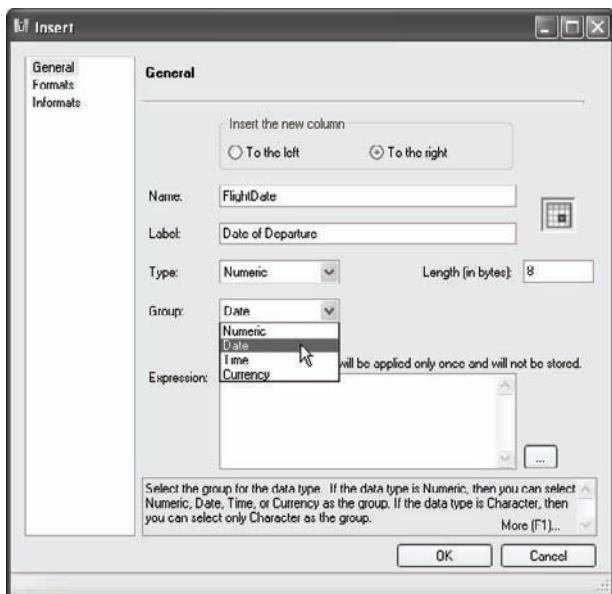


Setting the update mode Unless you are creating a new SAS data table, any data tables you open in SAS Enterprise Guide will initially be set to read-only mode. This prevents you from accidentally changing the data. To edit SAS data tables, you must change to update mode. Click the data icon in the Project Tree or Process Flow to make it active, and select **Edit ▶ Protect Data** from the menu bar. To switch back to read-only mode, select **Edit ▶ Protect Data** again.

Editing data values To change a value in a Data Grid, simply click the cell and start typing. You can also copy and paste values. In SAS Enterprise Guide 4.2, the Data Grid does not use informats to interpret input data. Instead, data values you enter will be interpreted based on the data type and data group for that column. After you enter a value, it will be displayed using the format for that column. If you are not sure which data type, data group, and format a column uses, you can find out by displaying the column properties. To do this, right-click the column header in the Data Grid, and select **Properties** from the pop-up menu.

Adding or removing a column To add a column to a data table, right-click a column header next to the place where you want to add a column, and select **Insert Column** from the pop-up menu. The Insert window will open.

Origin	Dest	FlightPrice
1 Seattle	Catania	\$1,045.00
2 Seattle	Hilo	\$677.00
3 Seattle	Jakarta	\$1,815.00
4 Seattle	Nairobi	\$1,761.00
5 Seattle	Quito	\$833.00
6 Seattle	Rome	\$596.00
7 Seattle	Sacrame	\$352.00
8 Seattle	San Jose	\$480.00
9 Seattle	Tokyo	\$721.00



In the General page of the Insert window, indicate whether you want the new column to be inserted to the left or the right. Then type a name and optional label for the column, and select the length, data type, and group. If you click **Formats** or **Informats** in the selection pane on the left, then you can specify display or read-in formats for the new column. In this case, the column will be named FlightDate, and will have a label of Date of Departure. Its type will be Numeric. It will be in the Date group, and will use the default format for dates, MMDDYY10.

To delete a column, right-click the column you want to delete and select **Delete** from the pop-up menu.

Adding or removing rows To add rows to a data table, right-click the row number at the point where you want to add rows, and select **Insert rows** from the pop-up menu. The Insert Rows window will open. Specify the number of rows you want to insert, and whether you want them to be inserted above or below. Then click **OK**.



You can also right-click a row and select **Append a row** to add a row at the bottom of the table. To delete rows, click the row you want to delete (or use shift-click to select more than one row).

Then right-click, and select **Delete rows**.

In this data table, a new row has been added for a flight to Sacramento, and a new column has been added for FlightDate.

	Origin	Destination	FlightNo	FlightPrice
1	Seattle	Catania	BA48	\$802.00
2	Seattle	Hilo	HA21	\$677.00
3	Seattle	Jakarta	AA119	\$1,815.00
4	Seattle	Nairobi	KLM6034	\$1,761.00
5	Seattle	Quito	CA1086	\$833.00
6	Seattle	Cut	A6	\$596.00
7	Seattle	Copy	1100	\$480.00
8	Seattle	Paste	875	\$721.00

	Origin	Destination	FlightDate	FlightNo	FlightPrice
1	Seattle	Catania	08/14/2011	BA48	\$802.00
2	Seattle	Hilo	07/14/11	HA21	\$677.00
3	Seattle	Jakarta		AA119	\$1,815.00
4	Seattle	Nairobi		KLM6034	\$1,761.00
5	Seattle	Quito		CA1086	\$833.00
6	Seattle	Rome		USA6	\$596.00
7	Seattle	Sacramento		SW345	\$352.00
8	Seattle	San Jose		CA1100	\$480.00
9	Seattle	Tokyo		UA875	\$721.00

2.6 Inserting Computed Columns in a Data Grid

In addition to entering and editing data, the Data Grid also allows you to compute new columns based on the value of existing columns. For example, if a teacher had scores for five exams, she could add a new column that would equal the mean of all the scores.

When you add a computed column to a data table, all the same caveats apply as when editing data tables. You'll probably want to save a copy of your table before making changes, and you must switch to update mode. For a discussion of these topics, see the previous section.

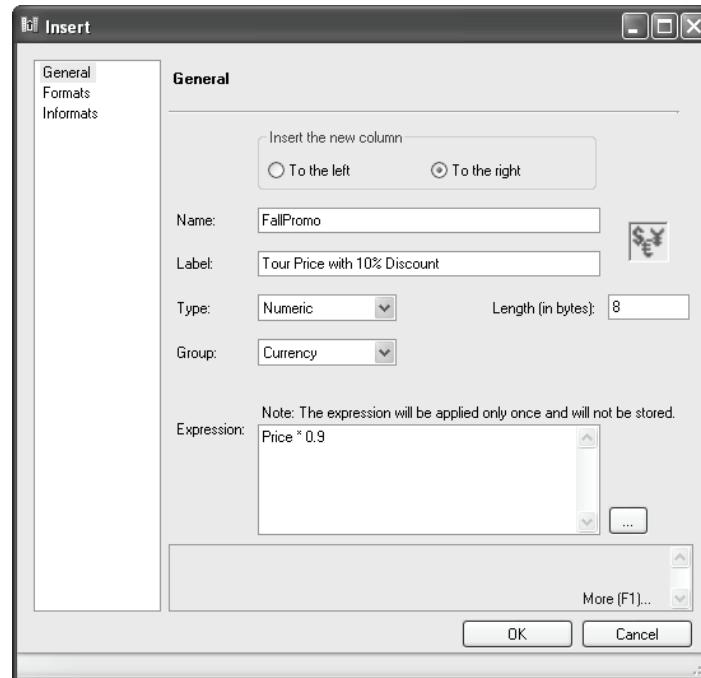
The Fire and Ice Tours company hopes to increase sales by offering a ten percent discount off the price of tours. To compute the new prices, insert a new column.

Inserting the column To add a computed column to a data table, you start the same way you would to insert an empty column. First right-click a column header next to the place where you want to add a column, and select **Insert Column** from the pop-up menu. The Insert window will open.



In the General page of the Insert window, indicate whether you want the new column to be inserted to the left or the right. Then type a name and optional label for the column, and select the length, data type, and group. In this example, the new column will have the name FallPromo, a label of Tour Price with 10% Discount, a type of Numeric, a group of Currency, and a length of 8.

If you want to specify the read-in or display format, click **Formats** or **Formats** in the selection pane on the left. This example uses the default format and informat.



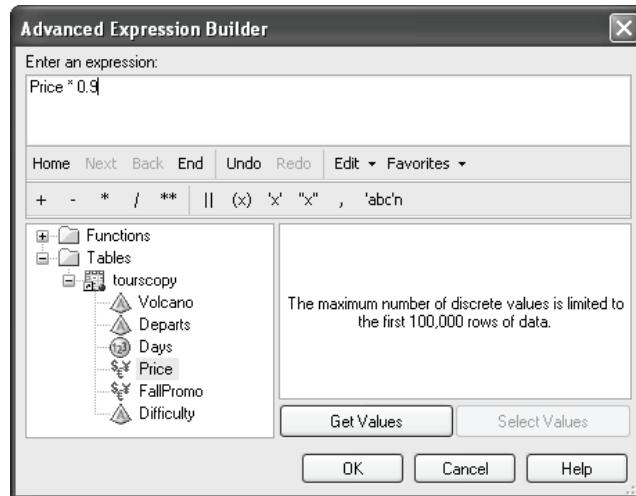
At the bottom of the Insert window is a box labeled **Expression**. You can type an expression in this box and SAS Enterprise Guide will use that expression to compute the values of the new column. In this example, you can see that the expression **Price * 0.9** has been entered so the values of the new column will equal the value of the Price column multiplied by 0.9. When you are satisfied, click **OK**, and the new column will be added to the Data Grid.

Here is the Data Grid showing the new column FallPromo.

	Volcano	Departs	Days	Price	FallPromo	Difficulty
1	Etna	Catania	7	\$1,075	\$968	m
2	Fuji	Tokyo	2	\$225	\$203	c
3	Kenya	Nairobi	6	\$830	\$747	m
4	Kilauea	Hilo	1	\$55	\$50	e
5	Kilimanjaro	Nairobi	9	\$1,310	\$1,179	c
6	Krakatau	Jakarta	7	\$895	\$806	e
7	Poas	San Jose	1	\$65	\$59	e
8	Reventador	Quito	4	\$575	\$518	m
9	St. Helens	Portland	2	\$167	\$150	e
10	Vesuvius	Rome	6	\$985	\$887	e

Building the expression In this example, the expression was simple so it was easy to type it in the Expression box. However, if the syntax of your expression were more complicated, you might want some help. To get help building your expression, click the ellipsis button  next to the box labeled **Expression** in the Insert window. The Advanced Expression Builder will open.

At the top of the Advanced Expression Builder is a box labeled **Enter an expression**. You can type in this box, or you can use the buttons and lists below to construct an expression.



The Advanced Expression Builder is similar to the expression builder available in the Query Builder. See Tutorial C or sections 4.3 and 4.4 for examples of building expressions.

Computing columns in a Data Grid versus in a query You can add a computed column to a data table in a Data Grid or in the Query Builder, but there is an important difference. When you compute a column in a Data Grid, the expression is applied only once, and is not saved. You cannot change an expression after it has been applied. So, if you wanted to see tour prices with a 20 percent discount, you would have to insert a completely new column. On the other hand, if you compute a column in a query, then you can re-open the query, make changes, and rerun it as many times as you wish.

2.7 Importing Microsoft Excel Spreadsheets

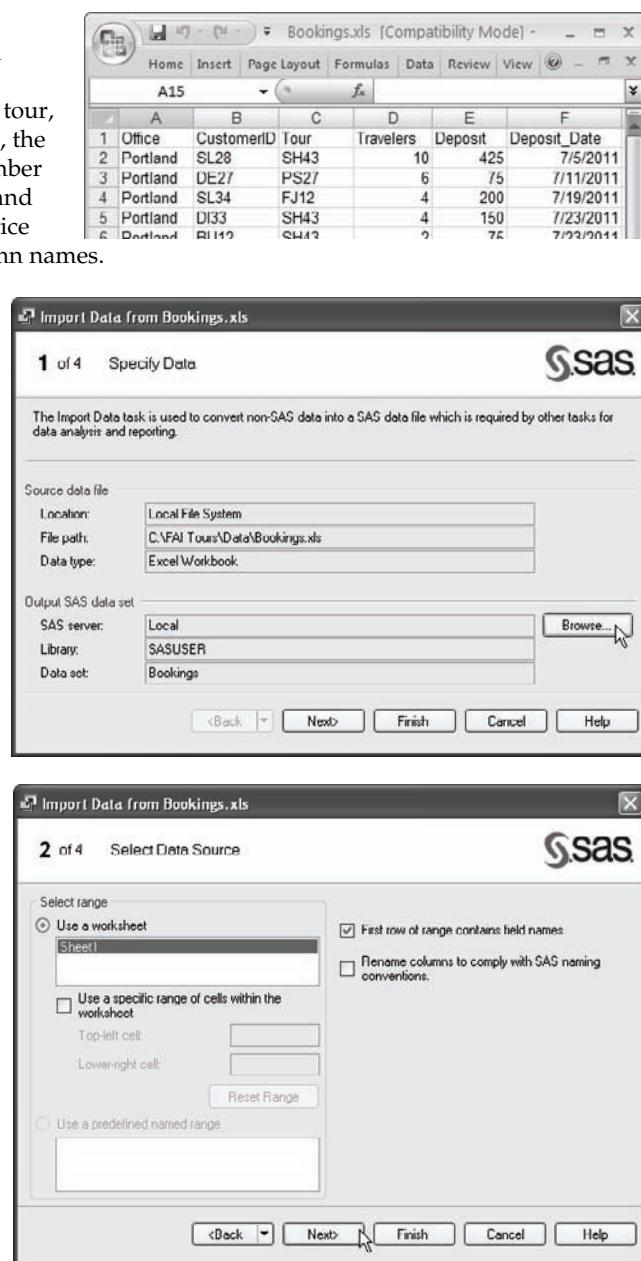
When you open or import a Microsoft Excel spreadsheet, SAS Enterprise Guide converts it to a SAS data set.

Input data This example uses a Microsoft Excel spreadsheet named Bookings.xls, which contains six columns: the office that booked the tour, the customer identification number, the tour identification number, the number of travelers, the money deposited, and the date the deposit was made. Notice that the first row contains the column names.

Import Data wizard There are several ways to open a Microsoft Excel file. You can select **File ▶ Open ▶ Data** or **File ▶ Import Data** from the menu bar, or drag the file from the Server List window to the Process Flow. The Import Data wizard will open.

The Import Data wizard has four windows. In the first window under the heading **Output SAS data set**, appears the SAS server, SAS data library, and data set name for the data set you are creating. To change any of these, click **Browse**. When you are satisfied, click **Next**.

In the second window under the heading **Select range**, you specify either a sheet, a specific range of cells, or a named range. You can also check the option **First row of range contains field names** to use values from the first row of the spreadsheet for column names. You can check the option **Rename columns to comply with SAS naming conventions** to tell SAS Enterprise Guide to automatically rename columns according to traditional rules for SAS names.



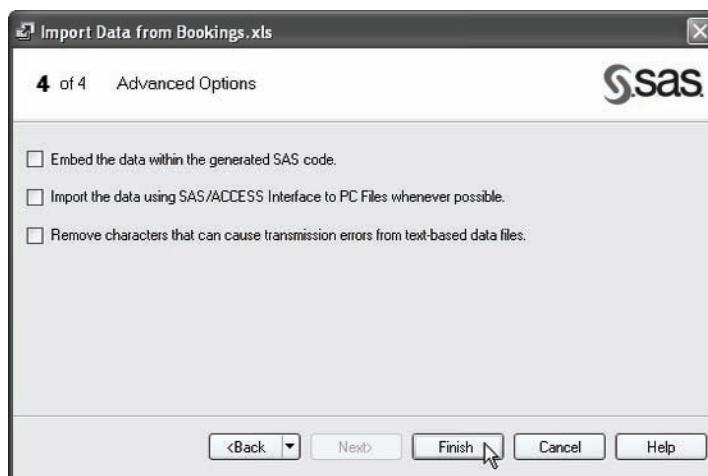
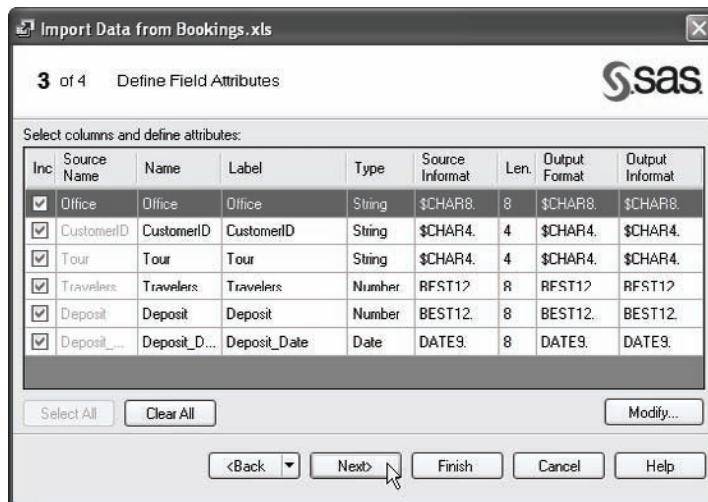
(See section 1.7 for a discussion of column names, and moving data between SAS Enterprise Guide and Base SAS.) For the Bookings data, import Sheet1 and use the first row as column names, and click **Next**.

In the third window, you see the column properties that SAS Enterprise Guide suggests for your data. To make changes, highlight the column you wish to change, and then click **Modify**. The Field Attributes window will open allowing you to change any column attribute. No changes are needed for this example. Click **Next**.

In the fourth window, you can choose options including **Import the data using SAS/ACCESS Interface to PC Files whenever possible**. If you have SAS/ACCESS Interface to PC Files software and you have large spreadsheets, using this option may be faster than the default.

When you are satisfied with the settings, click **Finish**.

Results The new data set will appear in a Data Grid. The data set opens in read-only mode, but because the data are now in a SAS data set, you can change to update mode to edit the data. Any changes you make will not be applied to the original Microsoft Excel spreadsheet.



	Office	CustomerID	Tour	Travelers	Deposit	Deposit_Date
1	Portland	SL28	SH43	10	425	05JUL2011
2	Portland	DE27	PS27	6	75	11JUL2011
3	Portland	SL34	FJ12	4	200	19JUL2011
4	Portland	DI33	SH43	4	150	23JUL2011
5	Portland	BU12	SH43	2	75	23JUL2011
6	Portland	DF31	FI12	2	175	25JUL2011

2.8 Importing Delimited Raw Data

Delimited raw data files have a special character separating the data values. That character is often a comma (as in CSV or comma-separated values files), but it can also be a tab, semicolon, space, or some other character.

Input data This example uses data from a file named Eruptions.csv. There are four variables: the volcano name, followed by the date an eruption started, the date that eruption ended, and the Volcanic Explosivity Index (VEI). Notice that this file has commas between the data values, and the first line contains the column names.

```
Volcano, StartDate, EndDate, VEI
Barren Island, 12/20/1795, 12/21/1795, 2
Barren Island, 12/20/1994, 06/05/1995, 2
Erebus, 12/12/1912, ., 2
Erebus, 01/03/1972, ., 1
Etna, 02/06/1610, 08/15/1610, 2
Etna, 06/04/1787, 08/11/1787, 4
Etna, 01/30/1865, 06/28/1865, 2
```

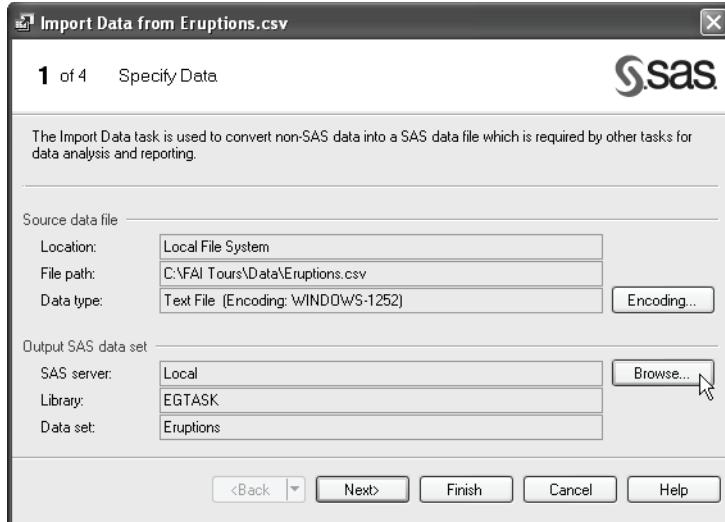
Opening a delimited data file You can open a raw data file by selecting **File ▶ Open ▶ Data** from the menu bar. When you do this, SAS Enterprise Guide opens the file in a simple text editor. Using this editor, you can view the data, and you can edit the data, but you cannot use the data in any task when it is opened this way. In order to use raw data in a task, you must import it.

Import Data wizard

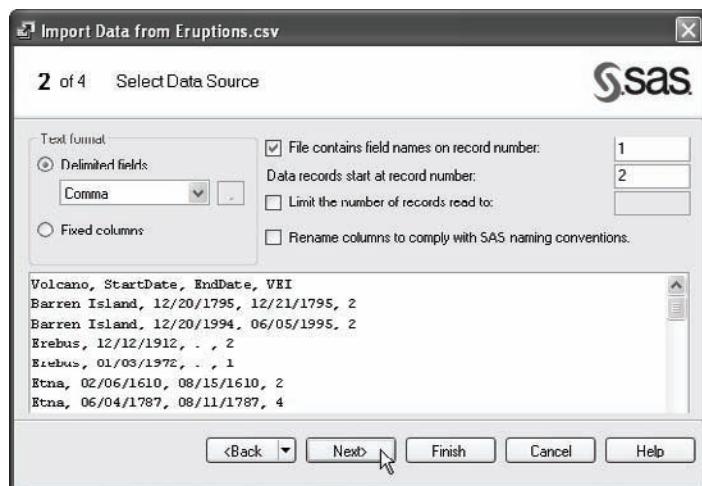
To import a raw data file, select **File ▶ Import Data** from the menu bar. The **Open** window will appear. Navigate to the file you wish to import, and click **Open**. The Import Data wizard will open.

The Import Data wizard has four windows. In the first window under the heading **Output SAS data set**, appears the SAS server, SAS data library, and data set name for the data set you are creating.

To change any of these, click **Browse**. When you are satisfied, click **Next**.



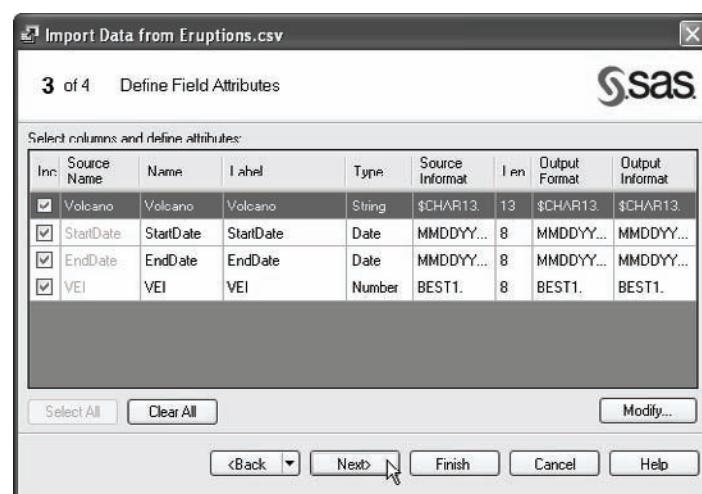
In the second window, select **Delimited fields**. To specify the delimiter for your file, click the down-arrow on the box and either select a delimiter (such as **Comma** or **Tab**) from the pull-down list, or select **Other** and type the delimiter into the box on the right. You can also check the option **File contains field names on record number**, and type a number in the corresponding box to use values in that row as column names. You can check the option **Rename columns to comply with SAS naming conventions** to tell SAS Enterprise Guide to automatically rename columns according to traditional rules for SAS names. (See section 1.7 for a discussion of column names, and moving data between SAS Enterprise Guide and Base SAS.) For the Eruptions data, select **Comma** as the delimiter, and use the first record for column names. Click **Next**.



In the third window, you see the column properties that SAS Enterprise Guide suggests for your data. To make changes, highlight the column you wish to change, and then click **Modify**. The Field Attributes window will open allowing you to change any column attribute. No changes are needed for this example. Click **Next**.

In the third window, you see the column properties that SAS Enterprise Guide suggests for your data. To make changes, highlight the column you wish to change, and then click **Modify**. The Field Attributes window will open allowing you to change any column attribute. No changes are needed for this example. Click **Next**.

In the fourth window (not shown), you can specify options for embedding data and removing characters that may produce errors. When you are satisfied with the settings, click **Finish**.



Results The data set will appear in a Data Grid. The data set opens in read-only mode, but because the data are now in a SAS data set, you can change to update mode to edit the data. Any changes you make will not be applied to the original data file.

	Volcano	StartDate	EndDate	VEI
1	Barren Island	12/20/1795	12/21/1795	2
2	Barren Island	12/20/1994	06/05/1995	2
3	Erebus	12/12/1912	.	2
4	Erebus	01/03/1972	.	1
5	Etna	02/06/1610	08/15/1610	2
6	Etna	06/04/1787	08/11/1787	4
7	Fra	01/30/1985	06/29/1985	2

2.9 Importing Fixed-Column Raw Data

Fixed-column raw data files are similar to delimited raw data files, but instead of having a delimiter separating the data values, the data values are lined up in tidy vertical columns.

Input data This example uses data from a file named LatLong.txt, which contains three variables: the name of the volcano, followed by its latitude and longitude. Notice that the column names appear in the first row, and the data values are vertically aligned.

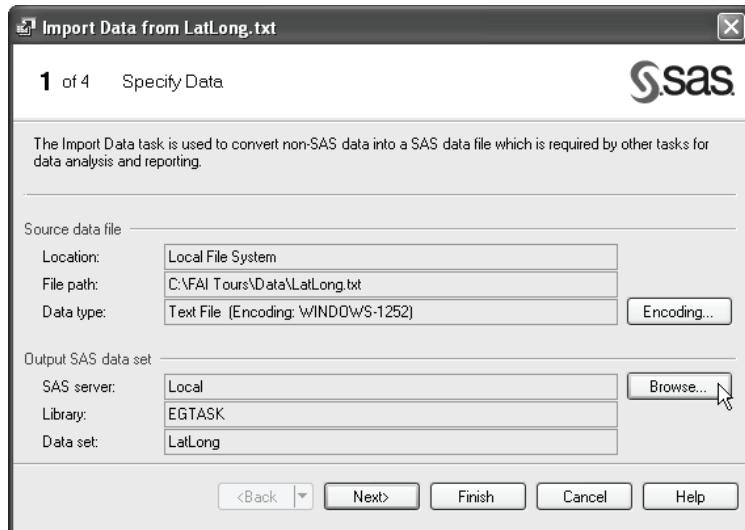
Volcano	Latitude	Longitude
Altar	-1.67	-78.42
Barren Island	12.28	93.52
Elbrus	43.33	42.45
Erebus	-77.53	167.17
Etna	37.73	15.00
Fuji	35.35	138.73
Garielaldi	40.95	-122.00

Opening a fixed-column data file You can open a raw data file by selecting **File ▶ Open ▶ Data** from the menu bar. When you do this, SAS Enterprise Guide opens the file in a simple text editor. Using this editor, you can view the data, and you can edit the data, but you cannot use the data in any task when it is opened this way. In order to use raw data in a task, you must import it.

Import Data wizard

To import a raw data file, select **File ▶ Import Data** from the menu bar. The Open window will appear. Navigate to the file you wish to import, and click **Open**. The Import Data wizard will open.

The Import Data wizard has four windows. In the first window under the heading **Output SAS data set**, appears the SAS server, SAS data library, and data set name for the data set you are creating. To change any of these, click **Browse**. When you are satisfied, click **Next**.



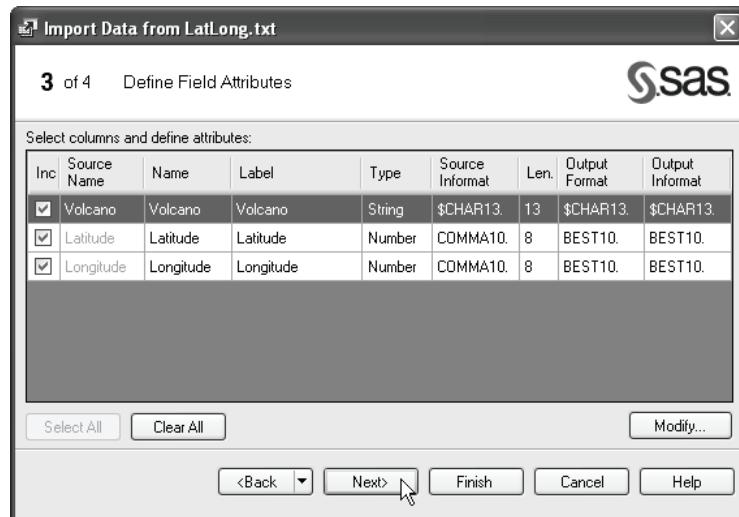
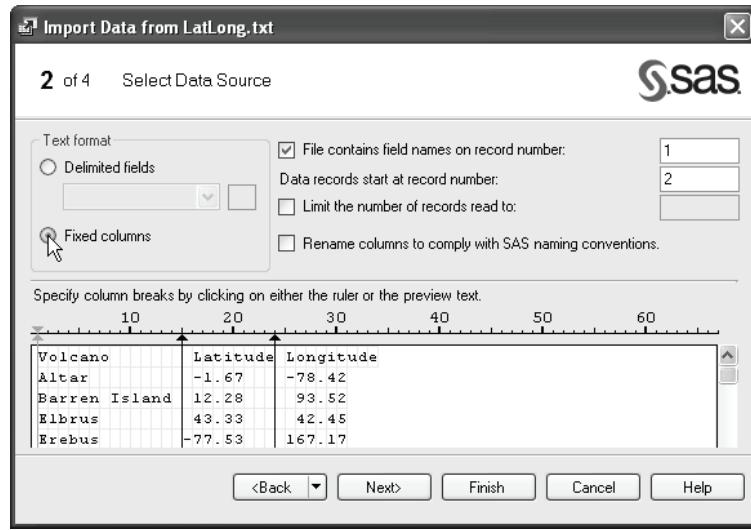
In the second window, select **Fixed columns**. At the bottom of this window is a box displaying the data file. Click the ruler above the data to tell SAS Enterprise Guide where each variable

begins. For the LatLong data, the volcano name starts at 1, latitude starts at 15, and longitude starts at 24. You can also check the option **File contains field names on record number**, and type a number in the corresponding box to use values in that row as column names. You can check the option **Rename columns to comply with SAS naming conventions** to tell SAS Enterprise Guide to automatically rename columns according to traditional rules for SAS names. (See section 1.7 for a discussion of column names, and moving data between SAS Enterprise Guide and Base SAS.) For the LatLong data, use the first record as column names, and click **Next**.

In the third window, you see the column properties that SAS Enterprise Guide suggests for your data. To make changes, highlight the column you wish to change, and then click **Modify**. The **Field Attributes** window will open allowing you to change any column attribute. No changes are needed for this example. Click **Next**.

In the fourth window (not shown), you can specify options for embedding data and removing characters that may produce errors. When you are satisfied with the settings, click **Finish**.

Results The data set will appear in a Data Grid. The data set opens in read-only mode, but because the data are now in a SAS data set, you can change to update mode to edit the data. Any changes you make will not be applied to the original data file.



	Volcano	Latitude	Longitude
1	Altar	-1.67	-78.42
2	Barren Island	12.28	93.52
3	Elbrus	43.33	42.45
4	Erebus	-77.53	167.17
5	Etna	37.73	15
6	Fiji	35.35	138.73
7	Garihaldi	49.85	-123

2.10 Exporting Data

After you have read your data into SAS Enterprise Guide and worked with it, you may want to access it in some other form. SAS Enterprise Guide can write data in these formats:

- Comma-separated values (CSV) files
- dBASE files
- HTML files
- IBM Lotus 1-2-3 files
- Microsoft Access 2002-2003 files
- Microsoft Excel 97-2003 files
- Paradox files
- SAS data tables
- Space-delimited text files
- Tab-delimited text files

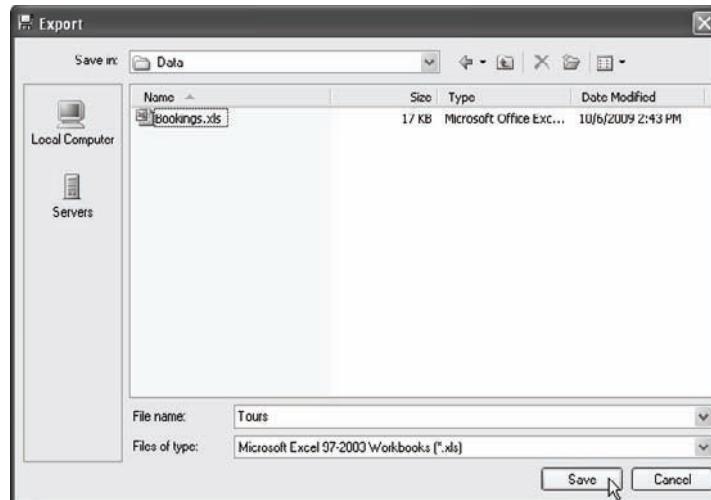
Here is a Data Grid showing the Tours data table, which will be exported as a Microsoft Excel file.

To export data from SAS Enterprise Guide, click **Export** on the workspace toolbar for the Data Grid and select **Export data-table-name** or **Export data-table-name As A Step In Project**.

You can also export data by right-clicking the data icon in the Project Tree or Process Flow, and selecting **File ▶ Export ▶ Export data-table-name** or **File ▶ Export ▶ Export data-table-name As A Step In Project** from the pop-up menu.

Exporting If you select **Export data-table-name**, then an Export window will open. Navigate to the location where you want to save the new file, specify a name for the file, select the type of file you want to create, and click **Save**. When you export data, no icon will appear in the Project Tree or Process Flow.

	Volcano	Departs	Days	Price	Difficulty
1	Etna	Catania	7	\$1,075	m
2	Fuji	Tokyo	2	\$225	c
3	Kenya	Nairobi	6	\$830	m
4	Kilauea	Hilo	1	\$55	e
5	Kilimanjaro	Nairobi	9	\$1,310	c
6	Krakatau	Jakarta	7	\$895	e
7	Pnae	San Jose	1	\$65	e



Exporting as a step in a project If you select **Export As A Step In Project**, the Export wizard will open. The number of windows in the wizard depends on the type of file you are creating. For an Excel file, there are five windows.

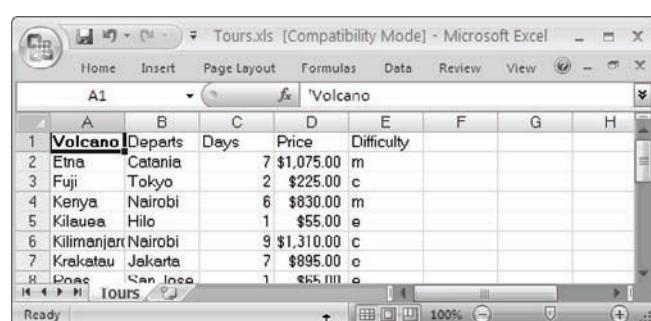
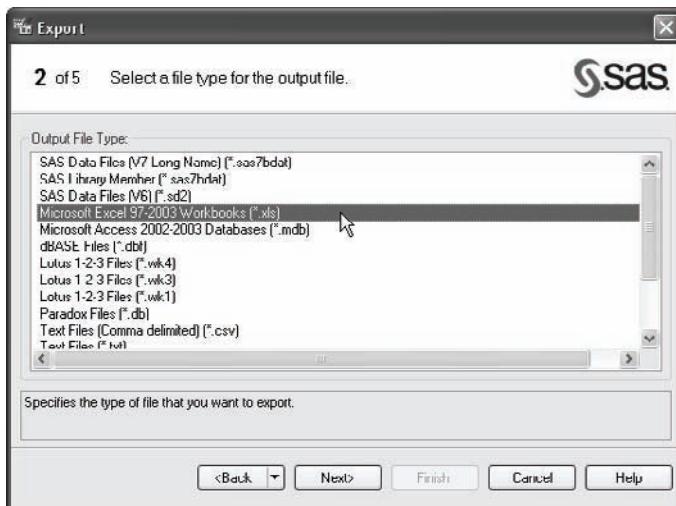
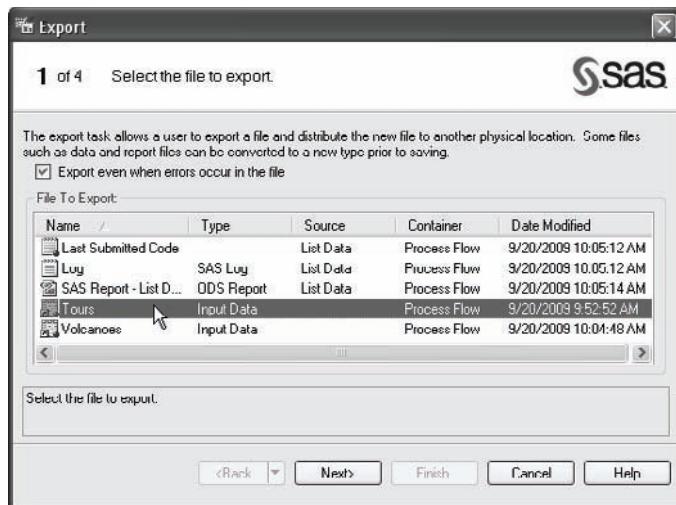
In the first window, select the data table you wish to export, and click **Next**.

In the second window, select the type of file you wish to create, and click **Next**.

In the third window (not shown), indicate whether you want to use labels for column names, and click **Next**.

In the fourth window (not shown), choose either **Local Computer** or **SAS Servers**, and then click the **Browse** button to navigate to the location where you want the table to be saved. You can only export a SAS data table to a computer that has SAS installed on it. In this window, you can also choose whether to **Overwrite existing output**. When you are satisfied, click **Next**.

In the fifth window (not shown),
confirm your settings by clicking **Finish**.



An Export File task icon  will be added to your project, along with an icon for the newly exported data. Every time you run your project, your data file will be automatically re-exported. Here is the Tours data table after being exported as a spreadsheet and opened in Microsoft Excel.

3

“ Knowledge is of two kinds.
We know a subject ourselves or
we know where we can find
information upon it. ”

SAMUEL JOHNSON

From Boswell's *Life of Johnson*, 1775. As quoted in *The Cyclopaedia of Practical Quotations: English, Latin, and Modern Foreign Languages* by Jehiel Keeler Hoyt, 1896.



CHAPTER 3

Changing the Way Data Values Are Displayed

Chapter 3

- 3.1 Applying Standard Formats in a Data Grid **190**
- 3.2 Applying Standard Formats in a Task **192**
- 3.3 Defining Your Own Character Formats **194**
- 3.4 Defining Your Own Numeric Formats **196**
- 3.5 Applying User-Defined Formats **198**

3.1 Applying Standard Formats in a Data Grid

When you open a Data Grid, SAS Enterprise Guide decides how the data should be displayed—how many decimal places to show, for example, and whether to use a dollar or percent sign. Most of the time this is fine, but sometimes you may want to change the way a column looks. You can do this by telling SAS Enterprise Guide to use a different display format.

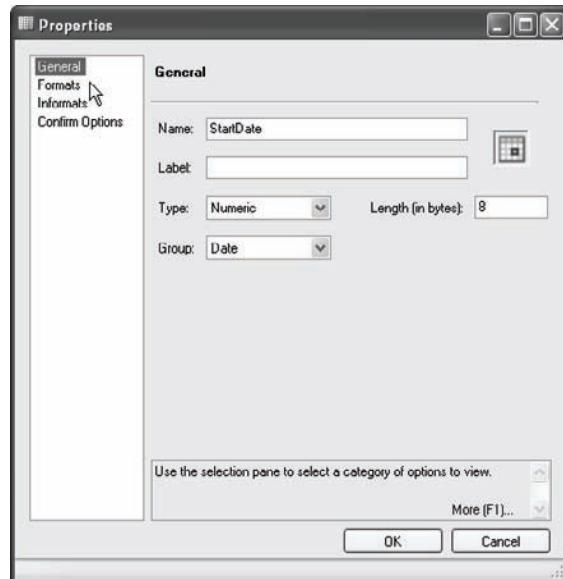
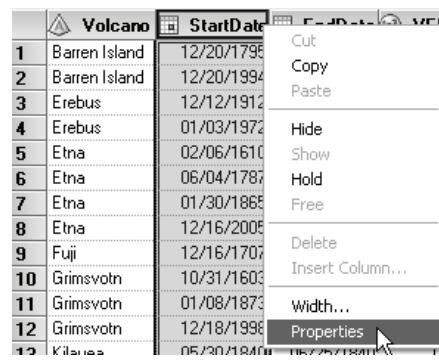
	Volcano	StartDate	EndDate	VEI
1	Barren Island	12/20/1795	12/21/1795	2
2	Barren Island	12/20/1994	06/05/1995	2
3	Erebus	12/12/1912		2
4	Erebus	01/03/1972		1
5	Etna	02/06/1610	08/15/1610	2
6	Fuji	06/14/1787	08/11/1787	4

Here is a Data Grid showing the Eruptions data table. The columns StartDate and EndDate use the default format for dates, MMDDYY10.0 (month followed by day and year), but SAS Enterprise Guide offers many other date formats.

Setting the update mode If the Data Grid is not already open, then open it by double-clicking the data icon in the Project Tree or Process Flow. Before you can change a display format, your data table must be in update mode. To switch from read-only mode to update mode, select **Edit ▶ Protect Data** from the menu bar. To switch back, select **Edit ▶ Protect Data** again.

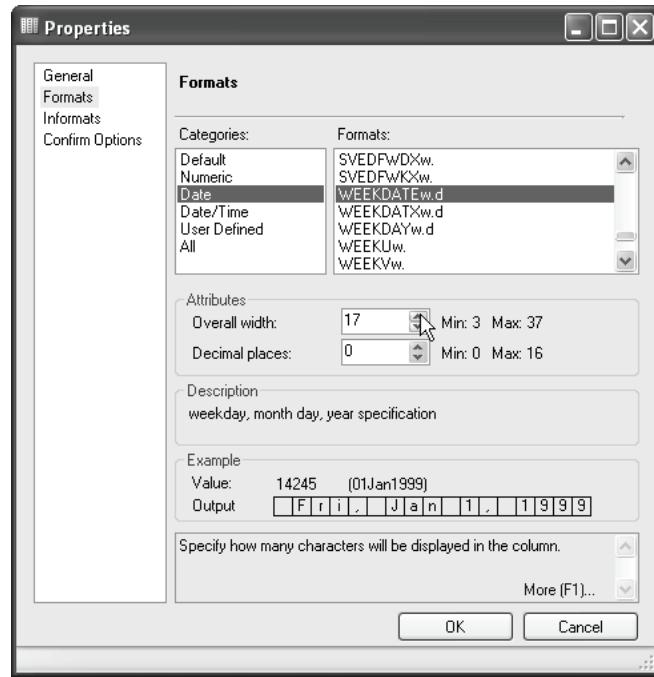
Opening the Properties window In the Data Grid, right-click the name of the column you want to modify, and select **Properties** from the pop-up menu. This opens the Properties window for that column.

The Properties window has several pages. The General page displays the data type and group. The data type and group determine which formats you can use for the column. If the type or group is wrong for any reason, you can change it here. Then click **Formats** in the selection pane on the left to open the Formats page.



Selecting formats In the Formats page, choose the category of formats you want to see, and then click the name of the format you want to use. In the area labeled **Attributes**, specify the overall width (the longest number of characters or digits that will be displayed for this column). For numeric columns, you may also specify the number of decimal places. The area labeled **Example** shows a sample of how this format will look. See section 1.11 for a table of commonly used formats.

In this example, the category Date has been selected, along with the format WEEKDATEw.d, an overall width of 17, and no decimal places. This format can be written as WEEKDATE17.0.



When you are satisfied with the format, click **OK** to apply the new format to the data.

Results Here is the Eruptions data table. Notice how different the values of StartDate look. When you make changes in a Data Grid, they are immediately saved with the data table, and will be applied in any tasks you run using that data table. However, formats affect only the way data are displayed—not the actual data values. That means you can easily select a different format or change the format back to what it was.

You can also apply formats in tasks and queries. Applying formats in tasks is discussed in the next section, and applying formats in queries is discussed in section 4.2.

	Volcano	StartDate	EndDate	VEI
1	Barren Island	Sun, Dec 20, 1795	12/21/1795	2
2	Barren Island	Tue, Dec 20, 1994	06/05/1995	2
3	Erebus	Thu, Dec 12, 1912	.	2
4	Erebus	Mon, Jan 3, 1972	.	1
5	Etna	Sat, Feb 6, 1610	08/15/1610	2
6	Etna	Mon, Jun 4, 1787	08/11/1787	4
7	Etna	Mon, Jan 30, 1865	06/28/1865	2
8	Etna	Fri, Dec 16, 2005	12/22/2005	1
9	Fuji	Fri, Dec 16, 1707	02/24/1708	5
10	Grimsvotn	Fri, Oct 31, 1603	11/01/1603	2
11	Grimsvotn	Wed, Jan 8, 1873	08/01/1873	4
12	Grimsvotn	Fri, Dec 18, 1998	12/28/1998	3
13	Kilauea	Sat, May 29, 1840	06/05/1840	0

3.2 Applying Standard Formats in a Task

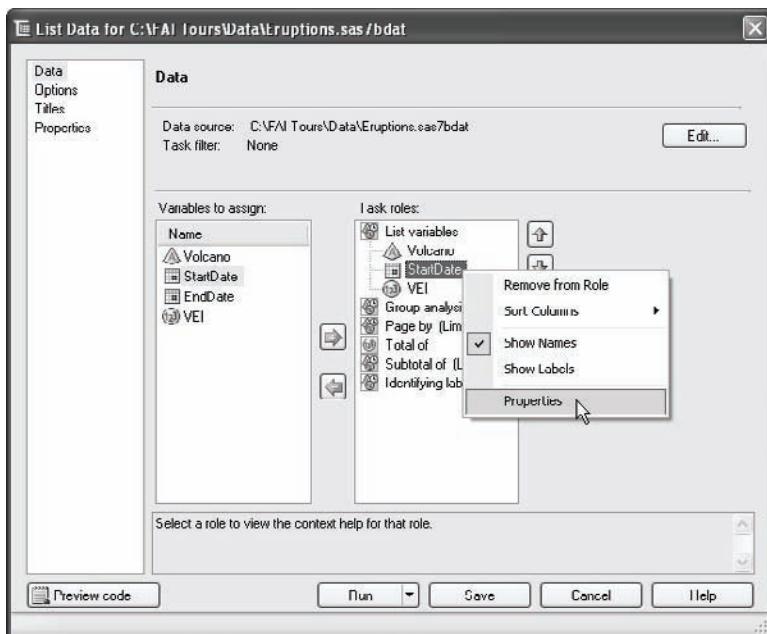
Every time you run a task that produces a report, SAS Enterprise Guide decides how the data should be displayed. That's good, but sometimes the way that SAS Enterprise Guide displays data may not be exactly what you want. You can change the way data are displayed by applying a format in a Data Grid or query, but then the format will be saved with the data set. If you don't want the format to be saved with the data, then you can apply the format directly in a task.

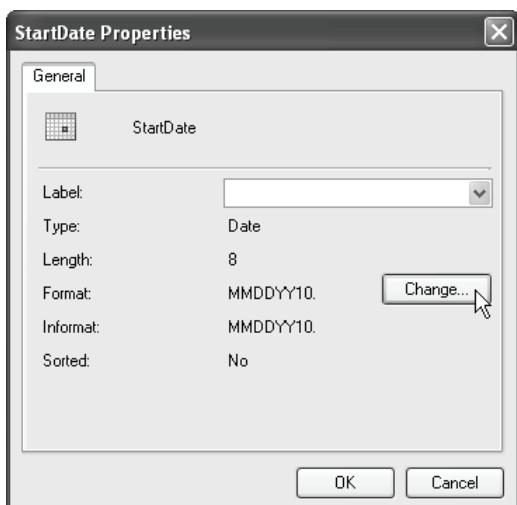
	Volcano	StartDate	EndDate	VEI
1	Barren Island	12/20/1795	12/21/1795	2
2	Barren Island	12/20/1994	06/05/1995	2
3	Erebus	12/12/1912	.	2
4	Erebus	01/03/1972	.	1
5	Etna	02/06/1610	08/15/1610	2
6	Etna	06/04/1787	08/11/1787	4
7	Etna	01/30/1865	06/28/1865	2
8	Etna	12/16/2005	12/22/2005	1
9	Fuji	12/16/1707	02/24/1708	5
10	Grimsvötn	10/31/1603	11/01/1603	2

Here is the Eruptions data set. The previous section showed how to apply the format WEEKDATEw.d to the variable StartDate in a Data Grid. This example uses the List Data task to show how you can apply the same format in a task.

To open the task, click the data icon in the Project Tree or Process Flow and select **Tasks ▶ Describe ▶ List Data** from the menu bar. The List Data window will open, displaying the Data page.

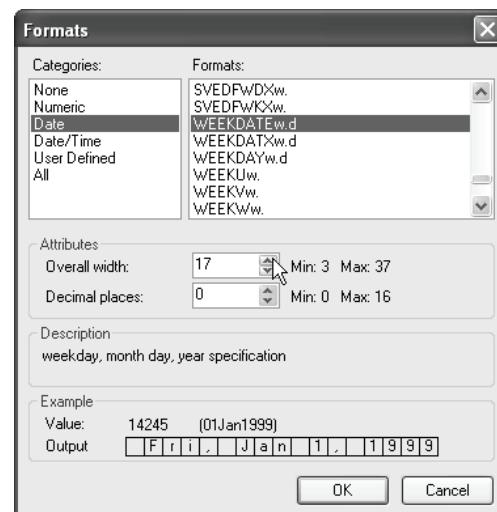
Opening the Properties window To open a Properties window for a variable, right-click the name of the variable you want to modify (in either the **Variables to assign** area or the **Task roles** area), and select **Properties** from the pop-up menu. In this example, the variables Volcano, StartDate, and VEI have been assigned to the **List variables** role, and **Properties** is being selected for StartDate.





Here is the Properties window for the variable StartDate. You cannot change a variable's data type or group using the Properties window inside a task. To make that change, use the Properties window inside a Data Grid, as described in the preceding section.

Click **Change** to open the Formats window.



Selecting formats In the Formats window, choose the category of formats you want to see, and then click the name of the format you want to use. In the area labeled **Attributes**, specify the overall width (the longest number of characters or digits that will be allowed for this variable). For numeric variables, you may also specify the number of decimal places. The area labeled **Example** shows a sample of how this format will look. See section 1.11 for a list of commonly used formats.

In this Formats window, the category Date has been selected, along with the format WEEKDATEw.d, an overall width of 17, and no decimal places. This format can be written as WEEKDATE17.0.

When you are satisfied with the format, click **OK**. Then click **OK** in the Properties window, and click **Run** in the task window.

Report Listing				
Row number	Volcano	StartDate	VEI	
1	Barren Island	Sun, Dec 20, 1795	2	
2	Barren Island	Tue, Dec 20, 1994	2	
3	Erebus	Thu, Dec 12, 1912	2	
4	Erebus	Mon, Jan 3, 1972	1	
5	Etna	Sat, Feb 6, 1610	2	
6	Etna	Mon, Jun 4, 1787	4	
7	Etna	Mon, Jan 30, 1865	2	
8	Etna	Fri, Dec 16, 2005	1	
9	Fuji	Fri, Dec 16, 1707	5	
10	Grimsvtn	Fri, Oct 31, 1603	2	

Results Here is the beginning of the report using the new format for StartDate. Any formats you apply in a task are not saved in the original data set and will not be used in other tasks.

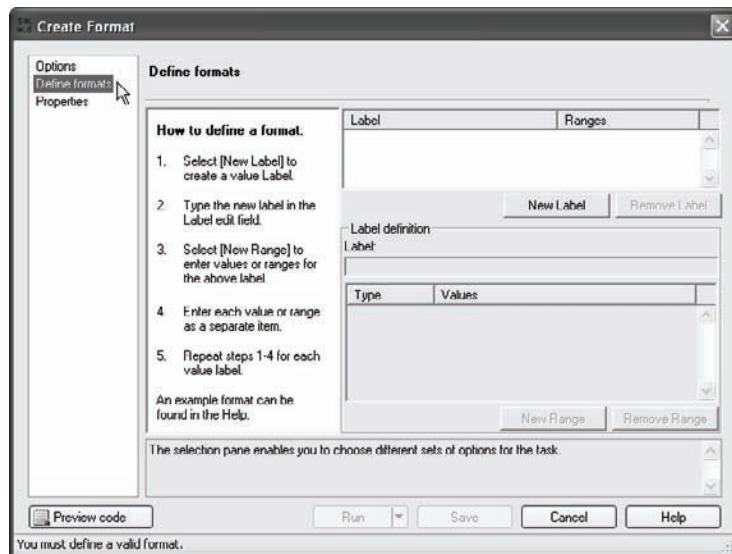
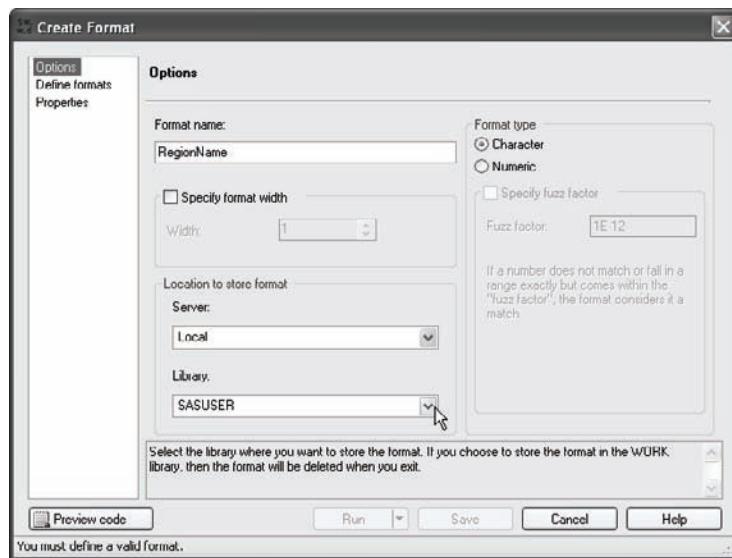
3.3 Defining Your Own Character Formats

Even with all the standard formats provided by SAS Enterprise Guide, there are times when you need something different. In those cases, you can create a user-defined format. Basically, user-defined formats allow you to specify a set of labels that will be substituted for specific values or ranges of values in your data. To do this, open the Create Format window by selecting **Tasks ▶ Data ▶ Create Format** from the menu bar.

Create Format

options To create a format for a character variable, select a **Format type** of **Character**. Then type a name for the new format in the **Format name** box. This name must be 31 characters or fewer in length; cannot start or end with a numeral; and can contain only letters, numerals, or underscores. This example shows a character format named RegionName being created.

Any formats stored in the WORK library (the default) will be deleted when you exit SAS Enterprise Guide. To save your format, choose a different library. If you have more than one SAS server, be sure to save your format on the same server where you run tasks. If you choose to leave your format in the WORK library, you can always regenerate it later by rerunning the Create Format task. When you are satisfied, click **Define formats** in the selection pane on the left.



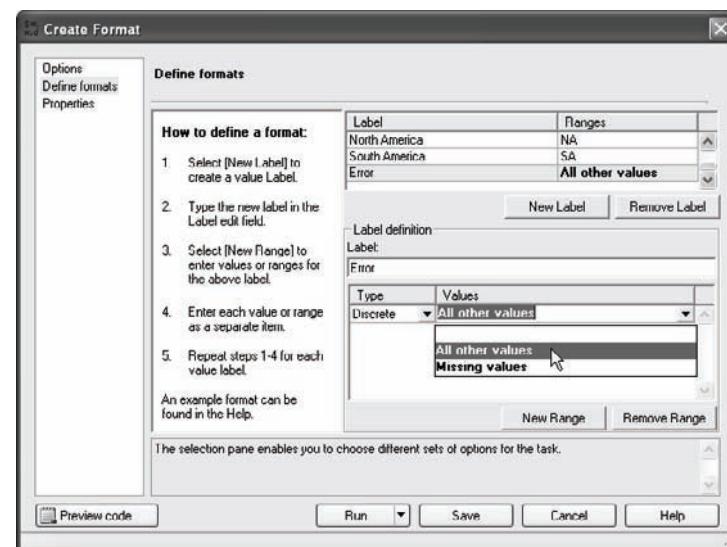
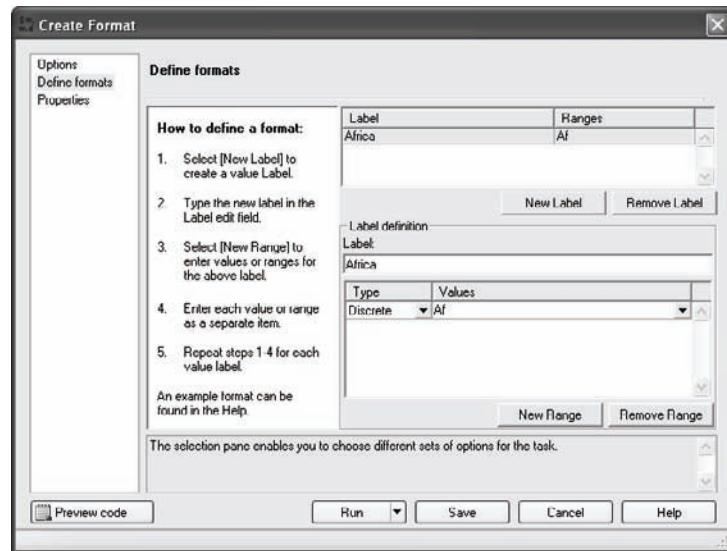
Defining formats You define formats in a stepwise fashion. First, click **New Label** and type a label in the **Label** box. Then click **New Range**, and under **Values** type the data value that corresponds to that label. Data values are case-sensitive, so “yes” is not the same as “Yes.” Repeat these steps until you have created all the labels you wish. In this example, you can see that the label Africa is being applied to the data value Af. Repeat the process to apply Antarctica to An, Australia/Pacific to AP, Asia to As, Europe to Eu, North America to NA, and South America to SA.

To specify a range of data values (such as A–D) rather than a discrete value, click the down-arrow (▼) under **Type** and select **Range** from the pull-down list. When you do that, a second box will appear under **Values** so that you can type in the two end points for your range.

You can specify a label to be used for missing values or for all other values by clicking the down-arrow in the box labeled **Values**. In this window, you can see that the label Error will be applied to all other data values. When you are satisfied with the format labels and ranges, click **Run** to create the format.

All character format names begin with a dollar sign, and end with a period, so this format will be named \$RegionName. A more detailed example of creating a character format appears in Tutorial B.

Using custom formats You can apply a user-defined format to a variable in the same ways you apply standard formats: in a Data Grid, a task, or a query. Section 3.5 shows the \$RegionName. format being used in a List Data task.



3.4 Defining Your Own Numeric Formats

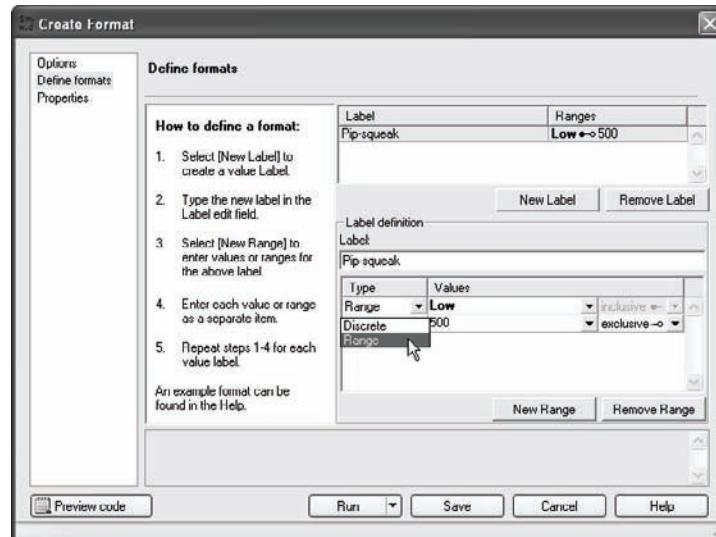
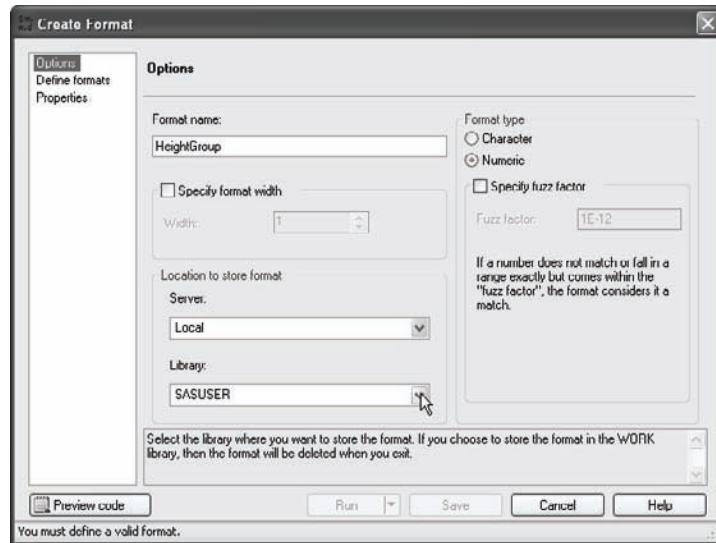
The previous section showed how to create a user-defined format for a character variable. Creating a user-defined format for a numeric variable is similar, but you have a few more options. Start by selecting **Tasks ▶ Data ▶ Create Format** from the menu bar to open the Create Format window.

Create Format

options Select a **Format type** of **Numeric**. Then type a name for the new format in the **Format name** box. This name must be 32 characters or fewer in length; cannot start or end with a numeral; and can contain only letters, numerals, or underscores. This example shows a numeric format named HeightGroup being created.

Any formats stored in the WORK library (the default) will be deleted when you exit SAS Enterprise Guide. To save your format, choose a different library. If you have more than one SAS server, be sure to save your format on the same server where you run tasks. If you choose to leave your format in the WORK library, you can always regenerate it later by rerunning the Create Format task. When you are satisfied, click **Define formats** in the selection pane on the left.

Defining formats You define formats in a stepwise fashion. First, click **New Label** and type a label in the **Label** box. Then click **New Range** and enter the data values corresponding to that label. Repeat these steps for the second label, and so on, until you have created all the labels you wish.



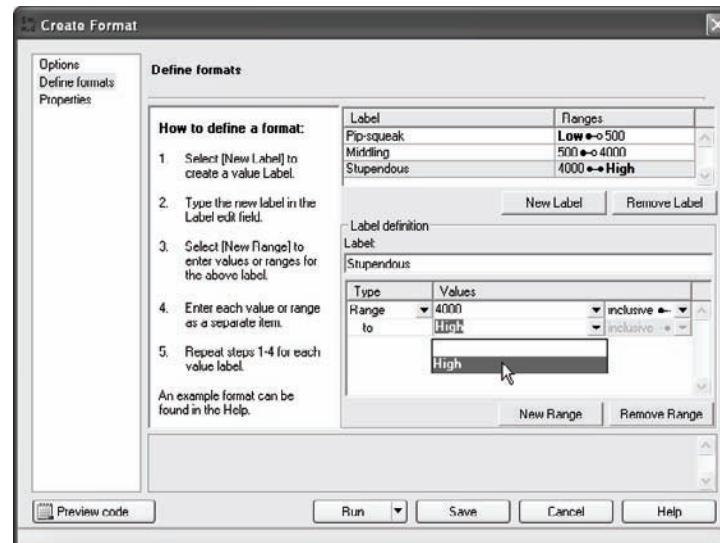
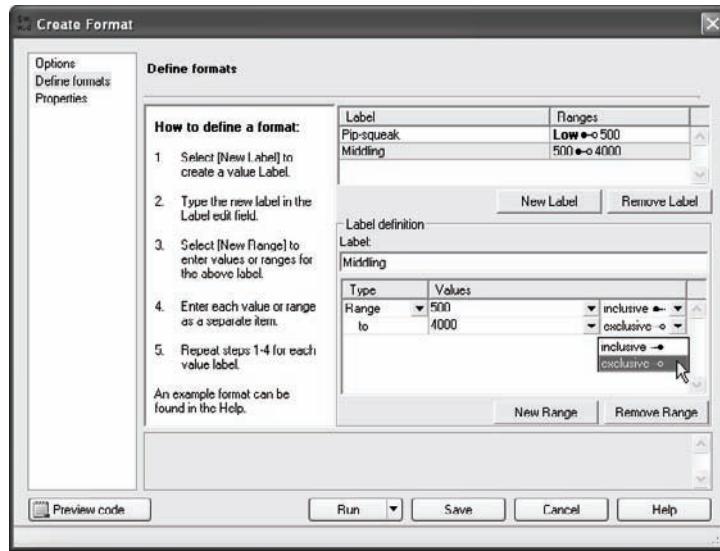
When you specify the data values, you have some choices. Under **Type**, click the down-arrow (**▼**) to open the pull-down list and select either **Discrete** (if you have a single value) or **Range**. In this example, you can see the label Pip-squeak will be substituted for a range of values up to 500.

You can make ranges inclusive or exclusive. In this example, the label Middling maps to values from 500 up to (but excluding) 4000. If you see a red X over ranges at the top of the window, it means that your ranges are overlapping and you will probably want to make one of them exclusive.

You can specify a label to be used for special values by clicking the arrow in the box labeled **Values**. For discrete values, you can select **All Other Values** or **Missing Values**. For ranges, you can select **Low** (the lowest possible value) and **High** (the highest). In this window, the label Stupendous has been mapped to data values from 4000 to High. When you are satisfied with the format labels and ranges, click **Run** to create the format.

Unlike character formats, numeric formats do not begin with a dollar sign. However, they do end with a period, so the name of this format will be HeightGroup.

Using custom formats You can apply a user-defined format to a variable in the same ways you apply standard formats: in a Data Grid, a task, or a query. The next section shows the HeightGroup. format being applied in a List Data task.



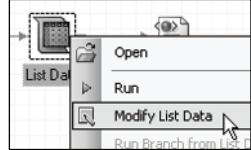
3.5 Applying User-Defined Formats

You apply user-defined formats in exactly the same ways you apply standard formats: in a Data Grid, a task, or a query. The example in this section applies two user-defined formats, \$RegionName. and HeightGroup. (from sections 3.3 and 3.4), in a List Data task.

Here is a simple report from a List Data task using the Volcanoes data set. The variables Volcano, Region, and Height have been assigned to serve in the List variables role. Notice that the data values are unformatted.

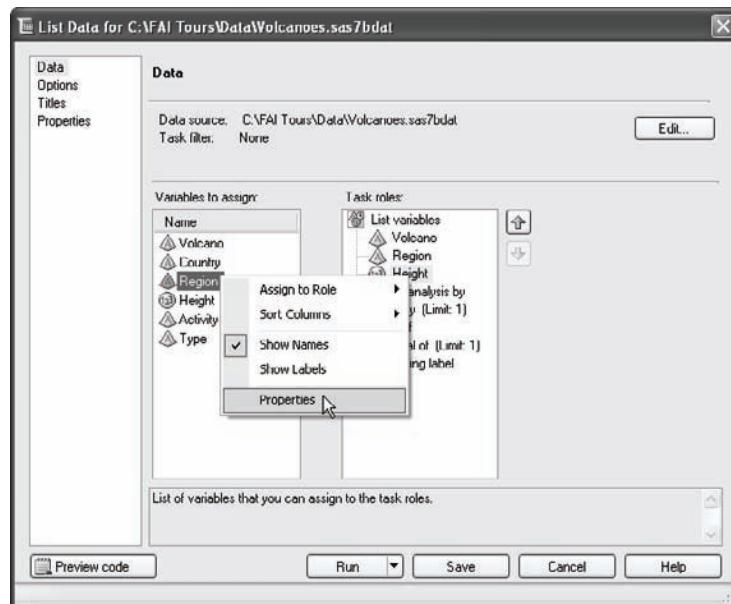
You can apply a format when you first run a task or you can add it later. To change an existing report, re-open the task window by right-clicking the task icon in the Project Tree or Process Flow and selecting **Modify task-name** from the pop-up menu.

Report Listing			
Row number	Volcano	Region	Height
1	Altar	SA	5321
2	Arthur's Seat	Eu	251
3	Barren Island	As	354
4	Elbrus	Eu	5633
5	Erebus	An	3794
6	Etna	Eu	3360
7	Fuji	As	3776
8	Garibaldi	NA	2678
9	Grimsvotn	Eu	1725
10	Illimani	SA	6458



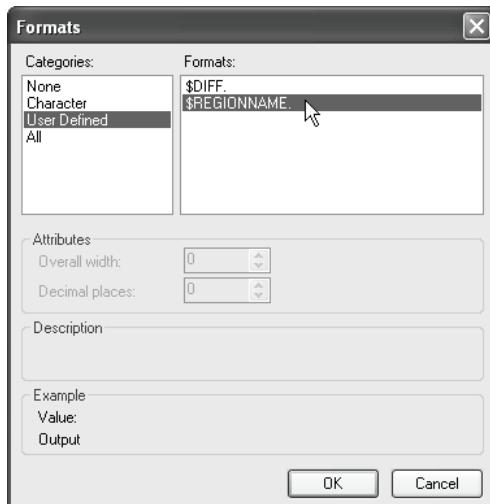
Opening the Properties window

To apply a format in a task, right-click the name of the variable you want to change in the Data page (in either the **Variables to assign** area or the **Task roles** area), and select **Properties** from the pop-up menu. In this example, **Properties** is being selected for the variable Region.



Using custom formats In the Properties window, click **Change** to open the Formats window for that variable.

Then in the Formats window, select the category **User Defined**. All the formats you have created will be listed. Here is the Formats window for the variable Region. Because Region is character, only character formats are listed. In this case, \$REGIONNAME. is being selected.



Results Here is the new report with user-defined formats applied. You can see that the values of Region are now displayed using the \$RegionName. format that was created in section 3.3. In addition, the values of Height are now displayed using the HeightGroup. format created in section 3.4.

In this example, the user-defined formats were applied to list variables. The result was that the formats simply replaced one value with another. However, if you apply a user-defined format to a variable assigned to a task role that groups the data, then it changes the structure of the report. See section 7.9 for an example of creating a grouped report with a user-defined format.



Once you have selected the correct format, click **OK** in the Formats window and click **OK** in the Properties window.

From the Data page of the task window, you can open the Properties window for other variables. Once you have applied all the formats you want, click **Run** in the task window.

Report Listing			
Row number	Volcano	Region	Height
1	Altar	South America	Stupendous
2	Arthur's Seat	Europe	Pip-squeak
3	Barren Island	Asia	Pip-squeak
4	Elbrus	Europe	Stupendous
5	Erebus	Antarctica	Middling
6	Etna	Europe	Middling
7	Fuji	Asia	Middling
8	Garibaldi	North America	Middling
9	Grimsvotn	Europe	Middling
10	Illimani	South America	Stupendous

4

“ The power to question is
the basis of all human
progress. ”

INDIRA GANDHI

Attributed to Indira Gandhi (1917-1984), prime minister of India.