

Paper 148-2009

**SAS® Enterprise Guide® for the Business Analyst**

Lina Clover, SAS Institute Inc., Cary, NC

**ABSTRACT**

If you are a business analyst and have a hard time figuring out how to use SAS to answer your business questions, you are not alone. SAS Enterprise Guide makes SAS analytics available to a wide range of users with a variety of skill levels, including programmers who are unfamiliar with SAS. In this paper, you will learn how to answer the “What” and “Why” questions about your business via the easy-to-use interface of SAS Enterprise Guide. You will learn how to prepare your data, perform statistical analyses, and interpret the results. You will also learn how to use the basic methods of data exploration with descriptive statistics such as min (minimum value), max (maximum value), mean, standard deviation, and correlation, and how to create graphical representations with bar and pie charts, histograms, and box plots. You’ll learn how to draw conclusions and make simple inferences with statistics such as linear regression, and how to share your results via e-mail, Microsoft Word and PowerPoint, PDF, and the Web.

**INTRODUCTION**

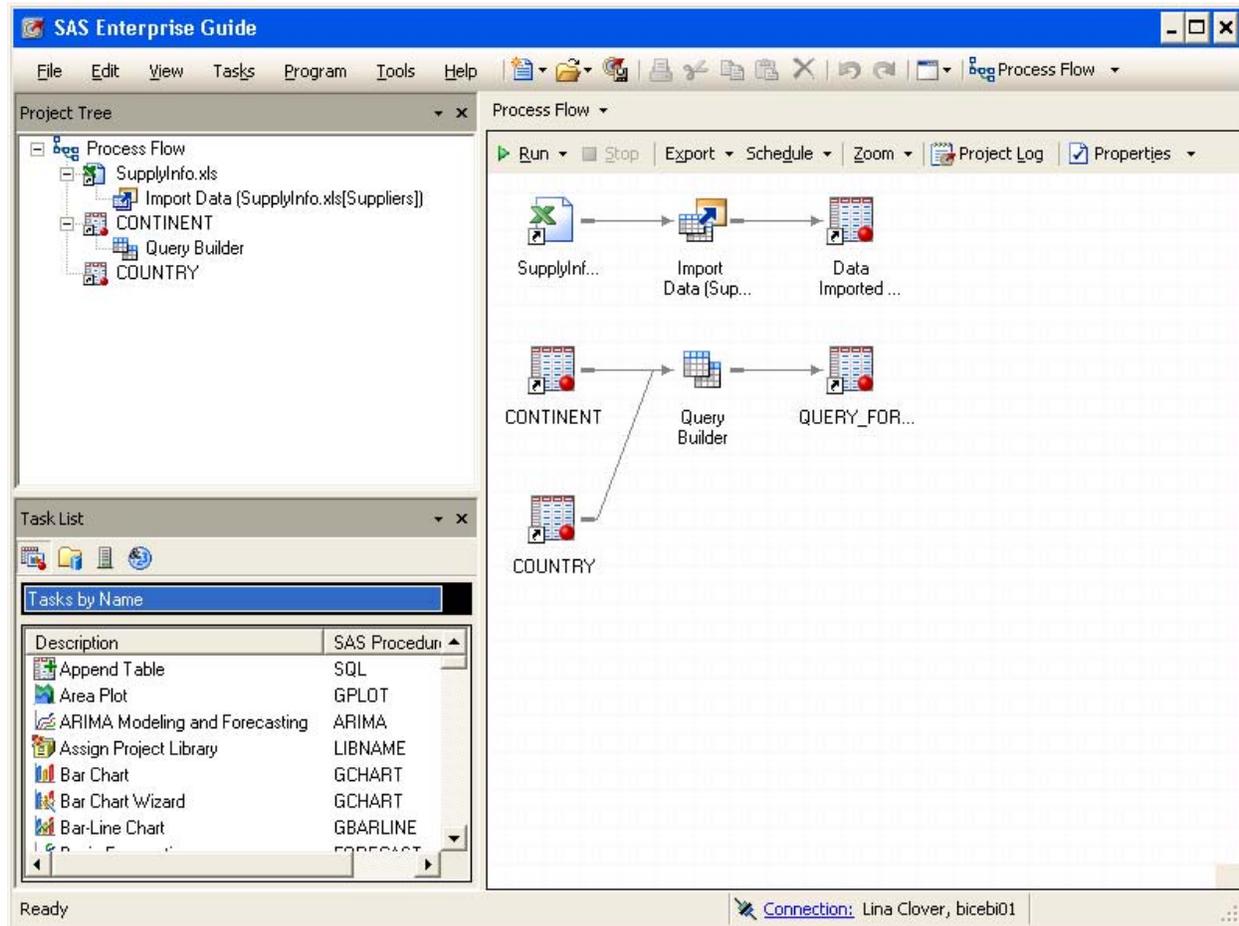
SAS Enterprise Guide delivers the power of SAS via an intuitive point-and-click interface; you don’t have to be a programmer, and you don’t have to learn SAS code to perform both simple and advanced statistical analyses. This paper focuses on the most common functions needed by business analysts in order to perform a given analysis with their business data. First, we’ll demonstrate how SAS Enterprise Guide enables you to manage your data effectively and efficiently to prepare it for analysis. Second, we’ll show you how to master the statistical tasks that you need to perform frequently. Finally, we’ll show you how to easily create charts and reports and distribute them throughout the enterprise.

**OVERVIEW OF SAS ENTERPRISE GUIDE WORK ENVIRONMENT**

The SAS Enterprise Guide work environment organizes data, analyses, results, and reports into project files. Project files have names and can be saved to your local machine or on the network. Project files can be retrieved later to continue the analytical work. The project file saves all the data references, the statistical work, and the results obtained from the data analysis. The actual SAS functionality is performed by work segments called tasks. Each task in SAS Enterprise Guide enables you to run a SAS procedure that is used to analyze your data. Tasks are the means by which most of the statistical analyses are performed. Tasks generate the necessary SAS PROC code and submit that code to the SAS server to execute on your behalf. The results are then displayed in one of the many reporting formats available to you via the software. You can view the code that is generated from each task, and you can even view the execution logs that are generated from the SAS server.

Within a project, the Process Flow diagram shows a visual representation of your work. The representation includes the data used, the tasks performed, and the results from each of the tasks. The Process Flow is a handy tool for organizing your work and ordering it in the sequence in which you want the various analyses to be executed. You can execute (run) the entire flow or selectively execute portions of it, and you can repeat the process as often as you want. If you are performing several related yet independent analyses, you can choose to create several process flows within a single project file.

Figure 1. SAS Enterprise Guide Work Environment



## IT'S ALL IN THE DATA: BRING IT ON

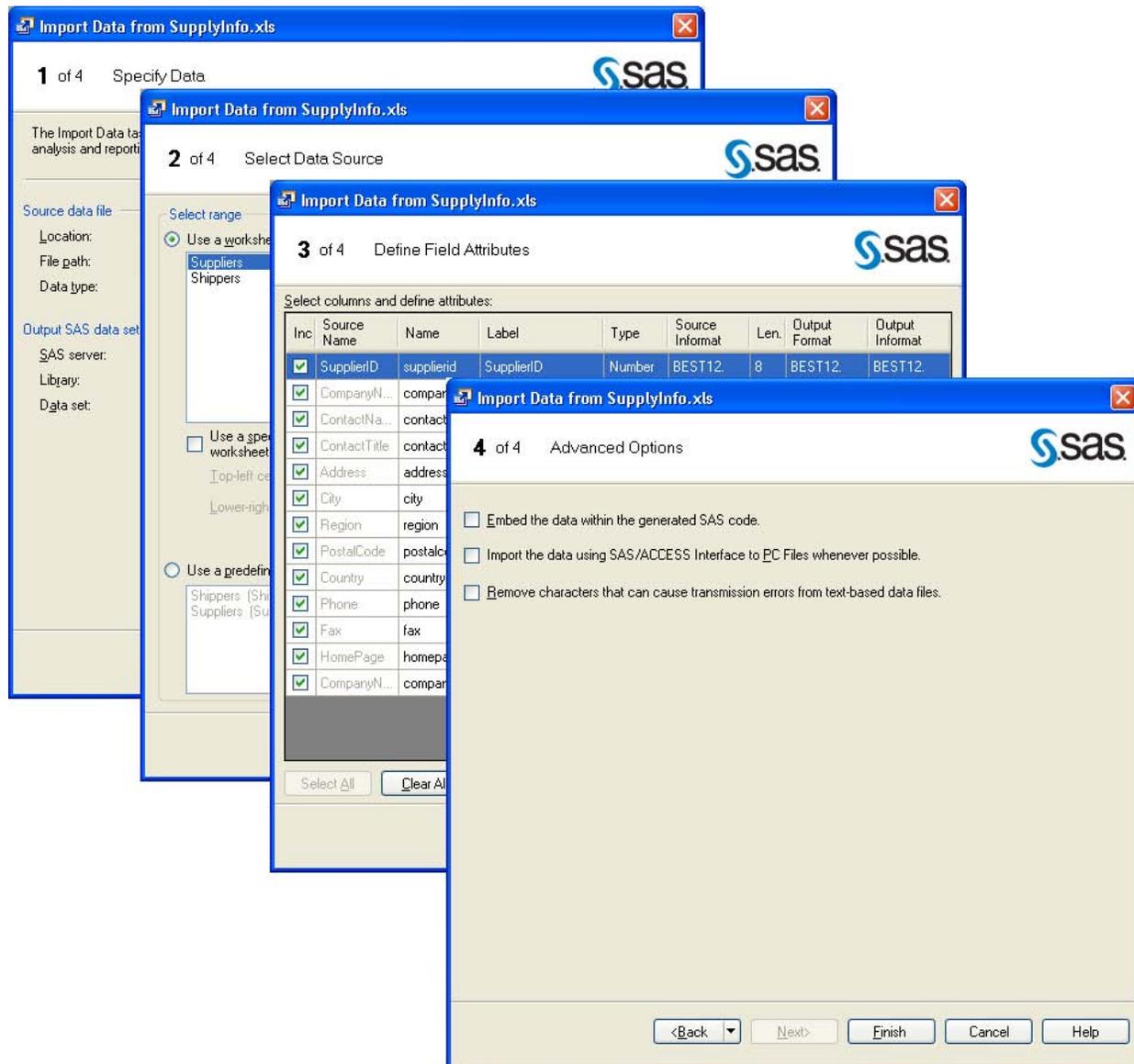
The first area that is covered in this paper is data preparation: bringing the data into the SAS System and using the Query Builder to prepare the data for analysis. One of the greatest strengths of SAS Enterprise Guide is its ability to work with data regardless of its type and source. Most of the SAS procedures require that you have data in a SAS data set (SAS data table) or imbedded in the SAS code in a format understood by SAS. SAS Enterprise Guide offers a comprehensive set of data tools that provide methods for opening existing data tables, creating new ones, and importing content from various file representations and storage systems into SAS.

### Data Files and Tables

You can create new SAS data sets and enter data directly from the data grid. The New Data wizard enables you to define columns of data with the two basic data types understood by SAS: character and numeric. Numeric data columns can be further formatted to indicate the type of numeric data they contain, such as date, time, or currency. Existing SAS tables can easily be accessed by SAS Enterprise Guide directly from the SAS server or from your local hard drive. The File Open dialog box enables you to access a variety of file types and open them in the work environment.

In most cases your data exists outside of SAS, in a dBase file, a relational database, a spreadsheet, or even a text file. Your data can be local or remote; the Import Data task can bring data into SAS and shape it into a form that is understood by the SAS server. The Import Data task is one of the most useful data access tasks in SAS Enterprise Guide. You can specify whether the data is comma delimited, tab delimited, fixed length, or variable length, before it is imported into a SAS table or imbedded in the SAS code.

Figure 2. Import Data Task



You can also open data from exchange servers, OLE DB, OLAP, and ODBC data sources. When necessary, SAS Enterprise Guide generates all data registration needed by the SAS server to read the data (that is, SAS Library definitions).

### The Query Builder Task

Bringing the data into SAS is only half of the story. In most cases, data is several separate data files and is not in the shape or format that is needed for business analysis. The Query Builder task in SAS Enterprise Guide is a powerful data management tool for manipulating data, joining tables, splitting columns, filtering rows, sorting rows and columns, and many other data operations. The advanced expression builder enables you to create new computed columns based on expressions or ranges of recoded values. The Query Builder modules generate standard SQL syntax that is needed for manipulating data. Like other tasks in SAS Enterprise Guide, the generated SQL Query is submitted to the SAS server to execute. Queries that result in new tables are automatically inserted into the project.

Figure 3. Example: Table Joins in the Query Builder Task

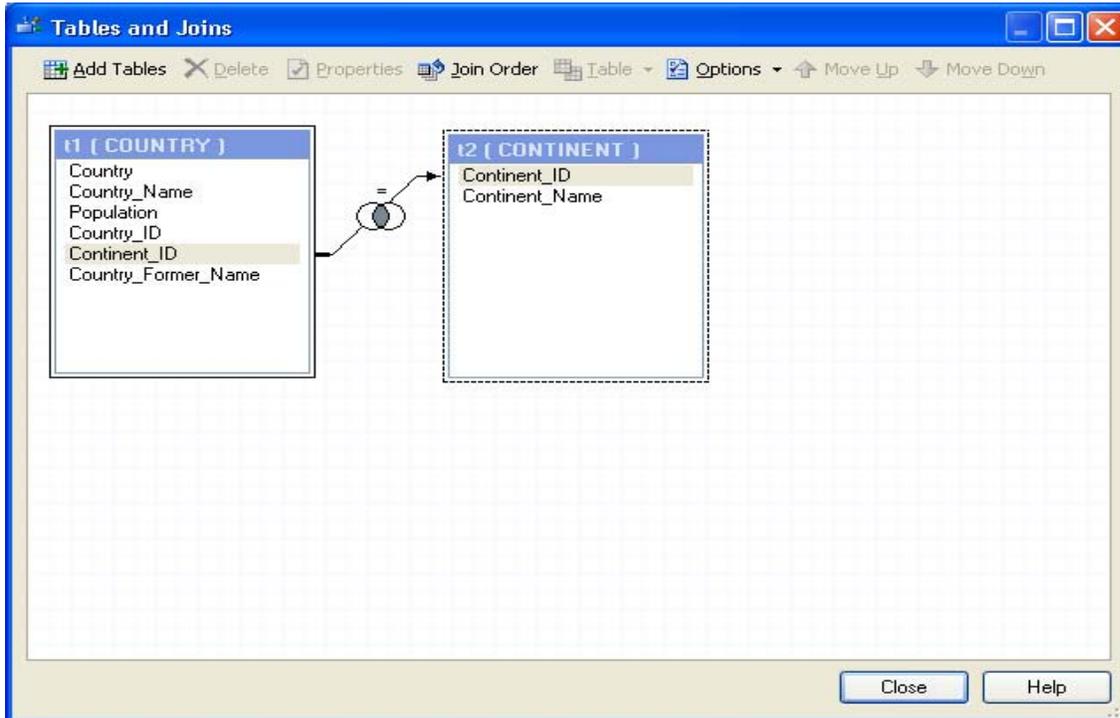


Figure 4. Example: Filtering Data in the Query Builder Task

The screenshot shows the 'Query Builder for SASMain - IWA:ORION\_RE.COUNTRY' window. The 'Filter Data' tab is active. The 'Where' clause is defined as follows:

Filter the raw data	Operator
Where	
t1.Population > 2800000000	OR
( GROUP )	
t1.Country_Name = 'United States'	OR
t1.Country_Name = 'Canada'	OR
t1.Country_Name = 'Mexico'	

The left pane shows the table structure for t1 (COUNTRY) and t2 (CONTINENT). The bottom of the window has buttons for 'Run', 'Save and Close', 'Cancel', and 'Help'.

## SAS STATISTICS, THE WAY TO GO

The power of SAS statistical and analytical capabilities is based on widely accepted statistical principles and theorems. The statistical procedures in the SAS programming language (often referred to as PROCs) provide the ability to apply statistical principles to business data. SAS is the most powerful statistical programming language and is used by statisticians and programmers worldwide.

SAS Enterprise Guide is the Windows desktop interface to SAS. It offers a comprehensive set of tasks to SAS procedures, enabling analysts who are not familiar with the SAS language the ability to run statistical analyses without having to learn to code. Many of the statistical tasks in SAS Enterprise Guide offer the option of producing both data and graphs that will help you turn raw data into usable information. Although SAS Enterprise Guide offers tasks to a large number of SAS procedures with advanced statistical capabilities, this paper covers only a few of the basic descriptive statistics most commonly used in business analysis.

### Summary Statistics Task

The basic statistics measurements of a non-discrete numeric variable are Mean-X (also Average, or value at the center), Minimum-MIN (the smallest value), Maximum-MAX (the largest value), Standard Deviation-Std Dev (the measure of how far the observations are spread out from the center), and N (the Number of observations). The Summary Statistics task can produce all of these statistics and is one of the most often-used statistical procedures in SAS.

Figure 5. Example: Summary Statistics Task

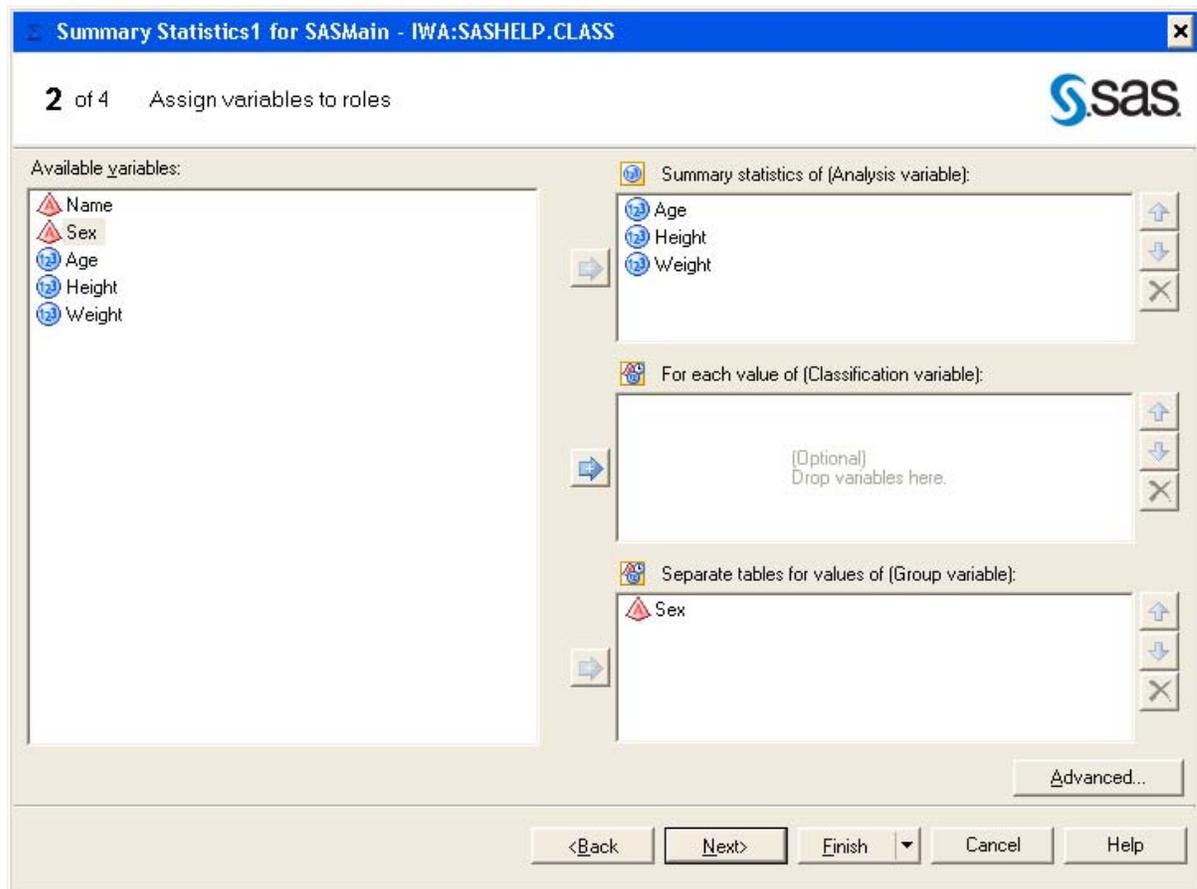


Figure 6. Example: Results of Summary Statistics Task

Sex=F					
Variable	Mean	Std Dev	Minimum	Maximum	N
Age	13.2222222	1.3944334	11.0000000	15.0000000	9
Height	60.5888889	5.0183275	51.3000000	66.5000000	9
Weight	90.1111111	19.3839137	50.5000000	112.5000000	9

Sex=M					
Variable	Mean	Std Dev	Minimum	Maximum	N
Age	13.4000000	1.6465452	11.0000000	16.0000000	10
Height	63.9100000	4.9379370	57.3000000	72.0000000	10
Weight	108.9500000	22.7271864	83.0000000	150.0000000	10

A large number of other statistics, such as standard error, variance, mode, range, sum, and the number of missing values, can also be produced by this task. Other statistics for percentiles, quartiles, confidence limits of the mean, and  $t$  statistics can also be obtained. Summary Statistics will quickly become your favorite task.

### Frequency Distribution of Discrete Data

You often need to know the frequency of occurrences of distinct values in one or more columns in your data table. The One-Way Frequency task is another simple task that enables you to conduct such analysis, producing two statistics—one statistic that contains the distinct data values, and another statistic that contains the frequency that the value occurs. There are other statistics, such as percent distribution, cumulative frequency distribution, and cumulative percent distribution, that can also be produced by this task that are useful in various types of analysis.

Figure 7. Example: One-Way Frequency Task

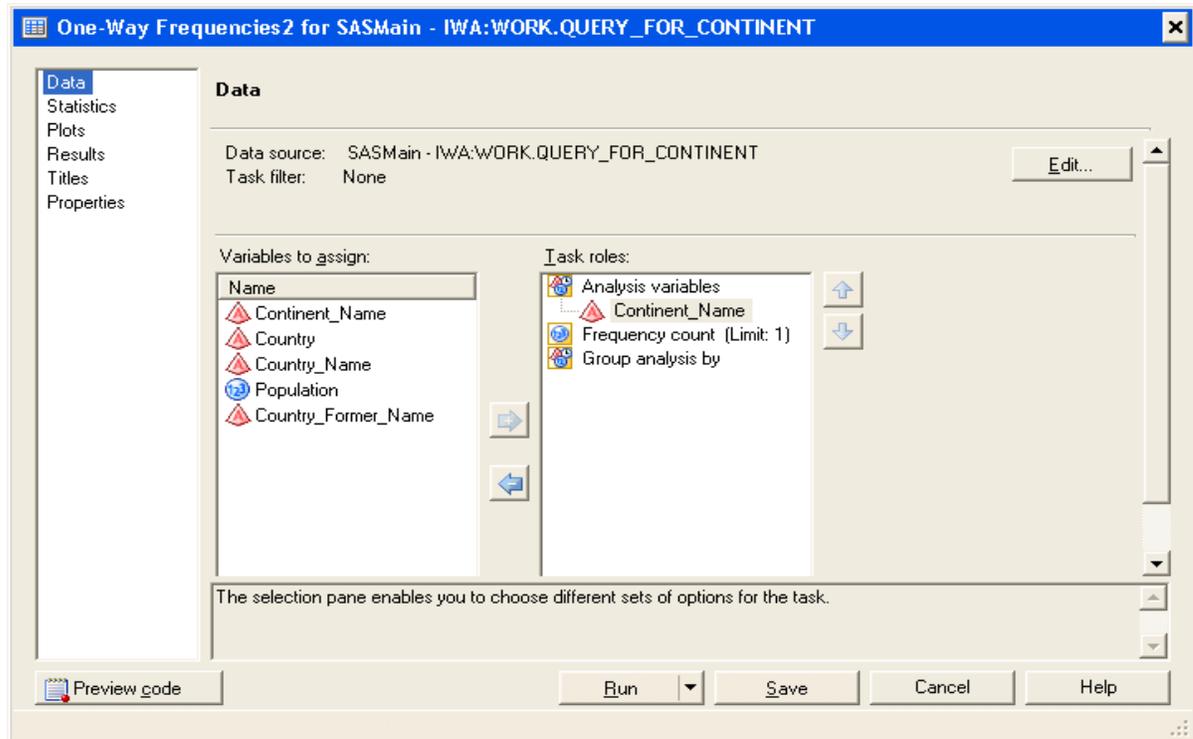


Figure 8. Example: Results of One-Way Frequency Task

Continent Name				
Continent_Name	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Africa	56	23.73	56	23.73
Antarctica	1	0.42	57	24.15
Asia	46	19.49	103	43.64
Australia/Pacific	34	14.41	137	58.05
Europe	48	20.34	185	78.39
North America	26	11.02	211	89.41
South America	25	10.59	236	100.00

Other useful tasks are Summary Table, and Characterize Data. The Summary Table task produces similar results to Summary Statistics and One-Way Frequency; it differs in that it enables you to produce a table that contains statistical data. The resulting table can then be used for further analysis or as an input to another statistical task. The Characterize Data task helps you create a summary report with graphs, frequency counts, and tables that describe the main characteristics of the data.

### Relationship between Two Numeric Variables

The Correlation task is used to determine the type of relationship between two numerical variables, X and Y, that belong to the same person or object. The goal of this analysis is to determine whether a positive or a negative relationship exists between X and Y. A positive relationship indicates that the smaller values of X tend to be paired with the smaller values of Y, and the larger values of X tend to be paired with the larger values of Y. A negative relationship indicates that the smaller values of X tend to be paired with the larger values of Y, and the larger values of X tend to be paired with the smaller values of Y.

Figure 9. Correlation Task

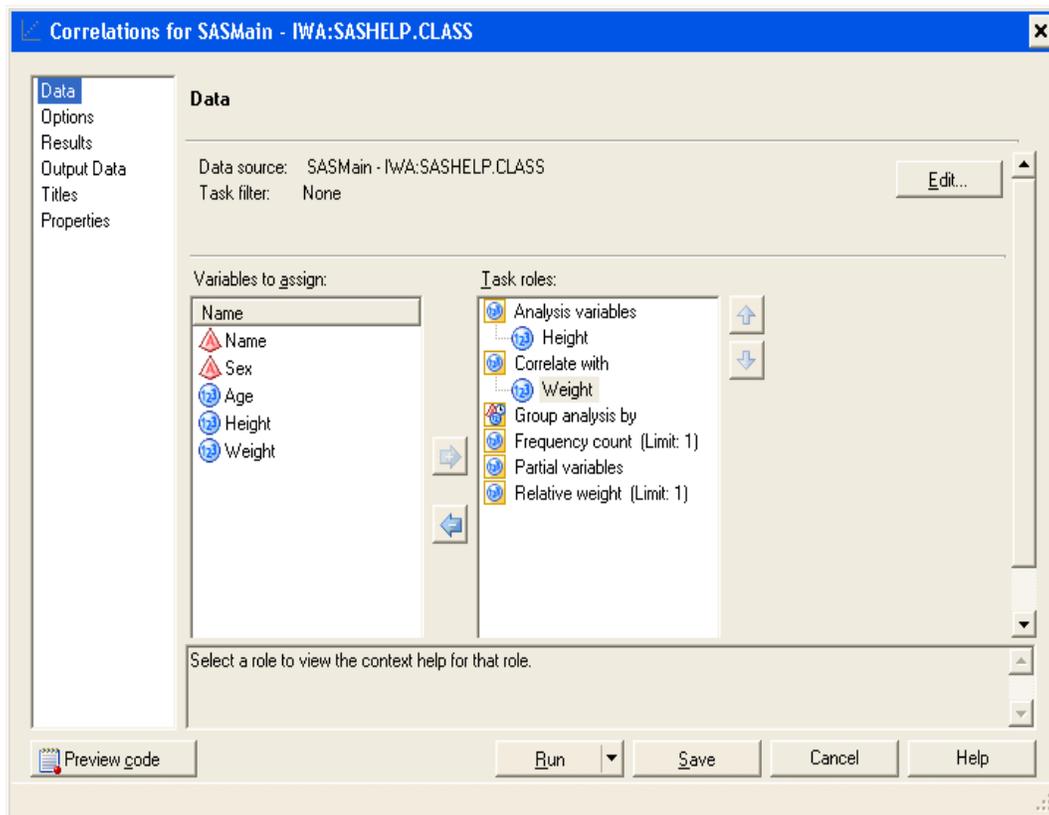
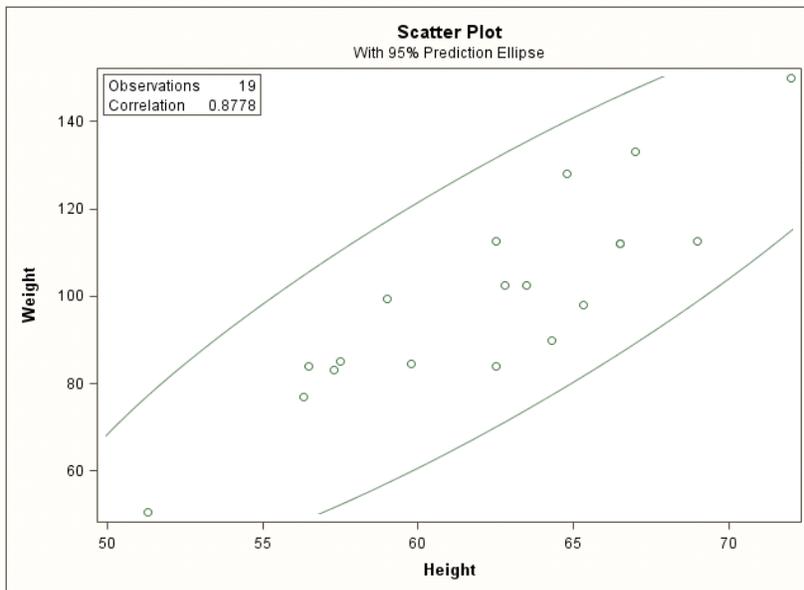


Figure 10. Results of Correlation Task

1 With Variables:		Weight				
1 Variables:		Height				
Simple Statistics						
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
Weight	19	100.02632	22.77393	1901	50.50000	150.00000
Height	19	62.33684	5.12708	1184	51.30000	72.00000

Pearson Correlation Coefficients, N = 19	
Prob >  r  under H0: Rho=0	
	Height
	0.87779
Weight	<.0001



Simple linear regression is another task that studies the linear relationship between two or more independent (also called explanatory and predictor) variables with a dependent (also called response) variable. A linear regression equation with one independent variable represents a straight line.

### The Plot Thickens with Graph Tasks

One of the most effective ways to analyze your data is with charts. SAS Enterprise Guide has a large gallery of business charts and graphics. They are most useful for comparing two or more values. The bars can be horizontal or vertical; stacked and grouped; two dimensional or three dimensional.

Figure 11. Gallery of Bar Chart Types

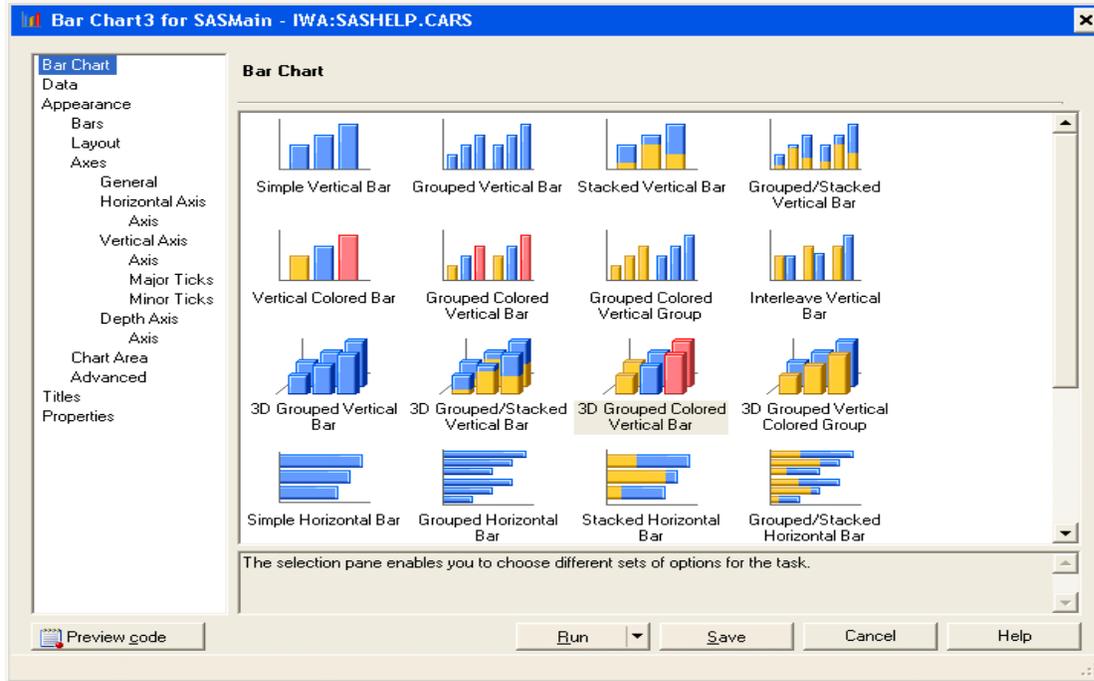
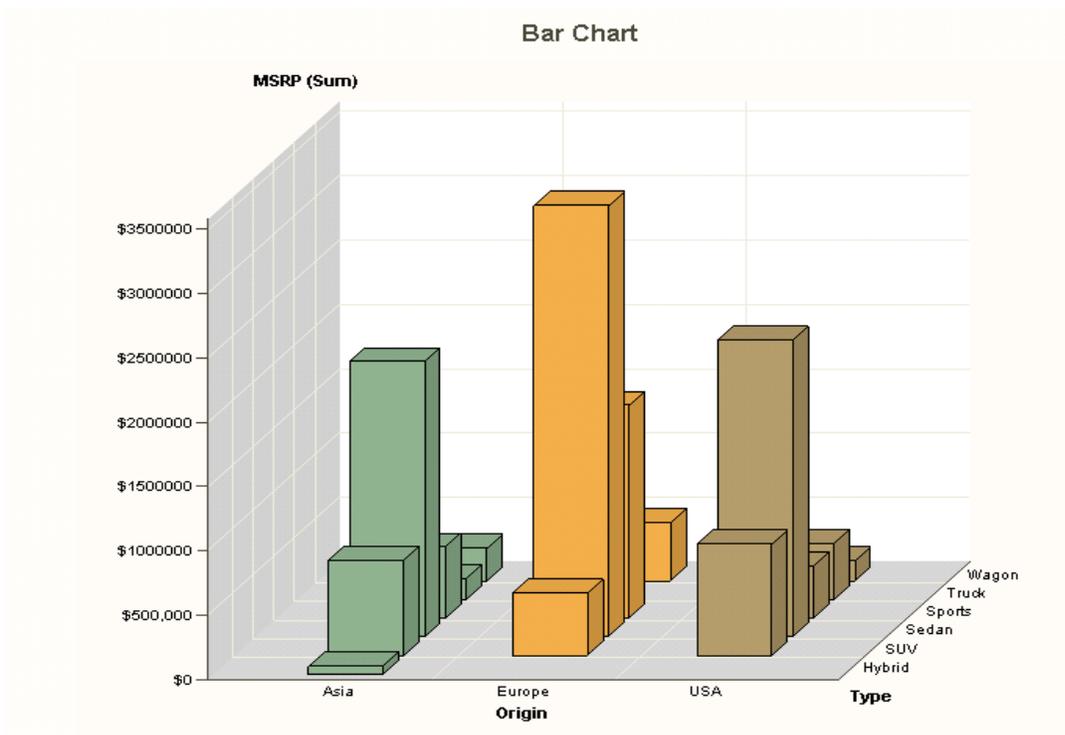


Figure 12. Results of Bar Chart Task

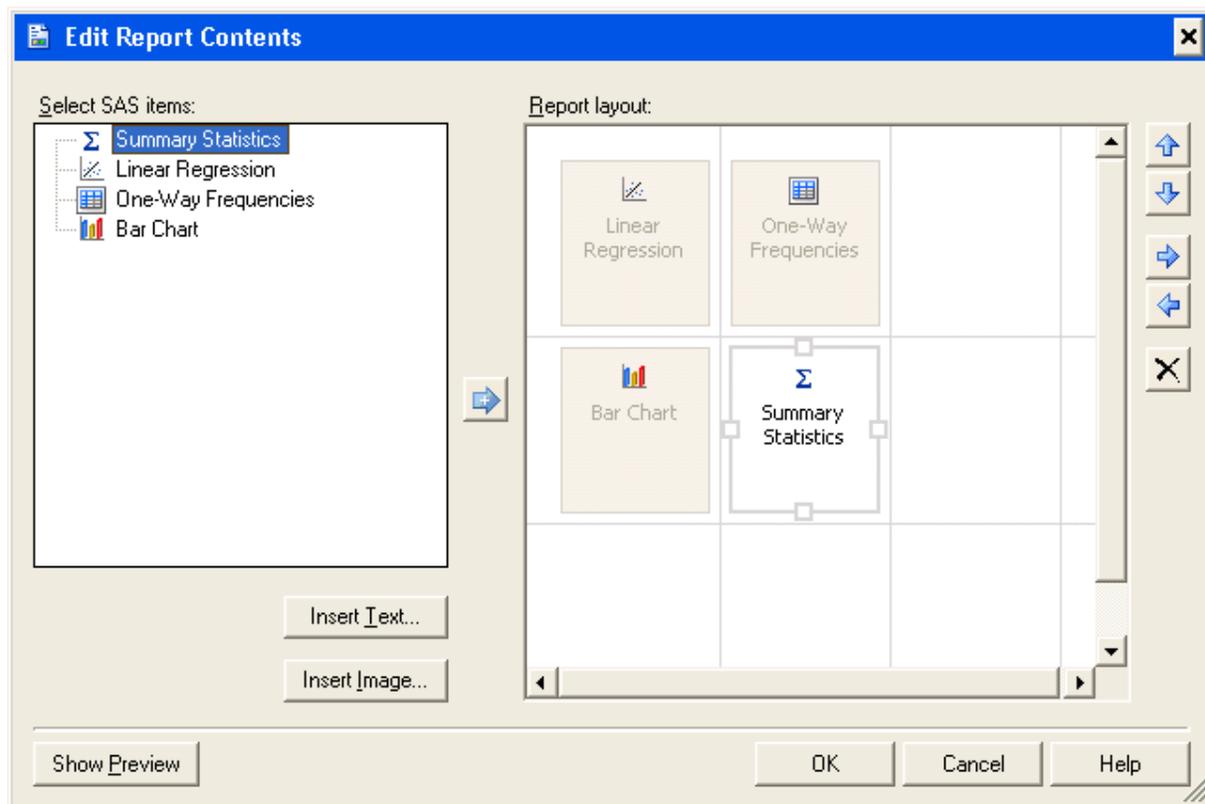


Some other types of business charts are pie, donut, tile, contour, line plot, scatter plot, area plot, bubble plot, and box plot.

## REPORT AND DISTRIBUTE, LET'S TELL EVERYONE

Reporting your statistical analysis results is one of the easiest functions in SAS Enterprise Guide 4.2. As a matter of fact, it is so easy you don't have to do anything to make reports from your statistical analysis. By default, SAS Enterprise Guide returns all task results in a SAS Report format. You can combine results from several tasks into a single report. You can also apply different styles for your report via the Style Gallery, which is available from the Tools menu.

**Figure 13. Example: Editing Report Content**



Depending on your business needs, you might choose other types of results. SAS Enterprise Guide and the SAS server support several formats for task output, such as HTML, RTF (Rich Text Format), PDF, or plain text. You can also specify several graph representations such as ActiveX, Java, JPEG, and GIF, to name a few. The HTML Document Builder is another tool that can be used to create an HTML document from several tasks and other documents. The Document Builder enables you to create a Web document with a table of contents and choose only the results that you want to include.

### Sharing the Results

There are several ways to share your analytical reports from SAS Enterprise Guide. SAS Report format is common to all SAS applications and solutions. SAS Reports created in SAS Enterprise Guide can be viewed in the SAS® Add-In for Microsoft Office in Microsoft Excel, Word, and PowerPoint. On the Web, reports can be viewed in SAS® Web Report Studio and the SAS® Information Delivery Portal.

Several export capabilities are available to transfer your data and reports to other applications such as Excel, Word, PowerPoint, Internet Explorer, and Adobe PDF. Another useful feature is the Send To E-mail Recipient; the results

are packaged and zipped as an e-mail attachment and sent to an Exchange account. You can also send the entire project file to e-mail recipients.

## CONCLUSION

The goal of this paper is to demonstrate how easy it is to perform most statistical analyses by using the intuitive interface of SAS Enterprise Guide. It is an ideal tool for business analysts who need to run analyses, yet don't have the time or inclination to learn how to program in SAS. For experienced professionals, we hope that you've learned about data management and reporting tools that will enable you to quickly distribute your analytical reports to others in your organization and on the Web.

## REFERENCES

Davis, James B. 2007. *Statistics Using SAS Enterprise Guide*. Cary, NC: SAS Institute Inc.

Slaughter, Susan J., and Lora D. Delwiche. 2006. *The Little SAS Book for Enterprise Guide 4.1*. Cary, NC: SAS Institute Inc.

## RECOMMENDED READING

Davis, James B. 2007. *Statistics Using SAS Enterprise Guide*. Cary, NC: SAS Institute Inc.

Der, Geoff, and Brian S. Everitt. 2007. *Basic Statistics Using SAS Enterprise Guide: A Primer*. Cary, NC: SAS Institute Inc.

McDaniel, Stephen, and Chris Hemedinger. 2007. *SAS for Dummies*. New York: Wiley.

Slaughter, Susan J., and Lora D. Delwiche. 2006. *The Little SAS Book for Enterprise Guide 4.1*. Cary, NC: SAS Institute Inc.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author:

Lina Clover  
SAS Institute Inc.  
Work Phone: 919-531-4616  
E-mail: [Lina.Clover@sas.com](mailto:Lina.Clover@sas.com)

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

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