### Paper 001-2011

# Point-and-Click Programming Using SAS® Enterprise Guide®

Mira Shapiro, Analytic Designers LLC, Bethesda, Maryland Kirk Paul Lafler, Software Intelligence Corporation, Spring Valley, California

### **Abstract**

SAS® Enterprise Guide® empowers organizations exploiting the power of SAS by offering programmers, business analysts, statisticians and end-users with powerful built-in wizards to perform a multitude of reporting and analytical tasks, access multi-platform enterprise data sources, deliver data and results to a variety of mediums and outlets, perform important data manipulations without the need to learn complex coding constructs, and support data management and documentation requirements quickly and easily. Attendees learn how to use the graphical user interface (GUI) to access tab-delimited and Excel input files; subset, group, and summarize data; join two or more tables together; flexibly export results to HTML, PDF and Excel; and visually manage projects using flowcharts and diagrams.

### Introduction

SAS® Enterprise Guide® (EG) provides a powerful programming platform to accomplish many tasks previously only possible using more traditional techniques found in the DATA and PROC steps. EG provides access to multi-platform enterprise data sources including SAS data sets, tab-delimited data, and Microsoft Excel files; satisfies "custom" reporting as well as complex analytical tasks; delivers data and results to a variety of mediums and outlets including HTML and Microsoft Excel; performs data manipulations without the need to learn complex coding constructs; and supports data management and documentation requirements including flowcharts and diagrams quickly and easily using the power of the built-in wizards.

# **Data Used In Examples**

The data used in all the examples in this paper consist of a selection of movie classics, along with an actors table. The Movies tab-delimited file, SAS data set, and Microsoft Excel file consists of six columns: title, length, category, year, studio, and rating. Title, category, studio, and rating are defined as character columns with length and year being defined as numeric columns. The Movies data is illustrated below.

### **Tab-delimited MOVIES File**

| Brave Heart                 | 177 Action Adventure                    | 1995 Paramount Pictures | R      |
|-----------------------------|---|-------------------------|--------|
| casablanca                  | 103 Drama                               | 1942 MGM / UA           | PG     |
| Christmas Vacation          | 97 Comedy                               | 1989 Warner Brothers    | PG-13  |
| Coming to America           | 116 Comedy                              | 1988 Paramount Pictures | R      |
| Dracula                     | 130 Horror                              | 1993 Columbia Tristar   | R<br>R |
| Dressed to Kill             | 105 Drama Mysteries                     | 1980 Filmways Pictures  | R      |
| Forrest Gump                | 142 Drama                               | 1994 Paramount Pictures | PG-13  |
| Ghost                       | 127 Drama Romance                       | 1990 Paramount Pictures | PG-13  |
| Jaws                        | 125 Action Adventure                    | 1975 Universal Studios  | PG     |
| Jurassic Park               | 127 Action                              | 1993 Universal Pictures | PG-13  |
| Lethal Weapon               | 110 Action Cops & Robber                | 1987 Warner Brothers    | R      |
| Michael                     | 106 Drama                               | 1997 Warner Brothers    | PG-13  |
| National Lampoon's Vacation | 98 Comedy                               | 1983 Warner Brothers    | PG-13  |
| Poltergeist                 | 115 Horror                              | 1982 MGM / UA           | PG     |
| Rocky                       | 120 Action Adventure                    | 1976 MGM / UA           | PG     |
| Scarface                    | 170 Action Cops & Robber                | 1983 Universal Studios  | R<br>R |
| Silence of the Lambs        | 118 Drama Suspense                      | 1991 Orion              |        |
| Star Wars                   | 118 Drama Suspense<br>124 Action Sci-Fi | 1977 Lucas Film Ltd     | PG     |
| The Hunt for Red October    | 135 Action Adventure                    | 1989 Paramount Pictures | PG     |
| The Terminator              | 108 Action Sci-Fi                       | 1984 Live Entertainment | R<br>G |
| The Wizard of Oz            | 101 Adventure                           | 1939 MGM / UA           |        |
| Titanic                     | 194 Drama Romance                       | 1997 Paramount Pictures | PG-13  |

#### **MOVIES Data Set**

|    | Title                          | Length | Category             | Year | Studio             | Rating |
|----|--------------------------------|--------|----------------------|------|--------------------|--------|
| 1  | Brave Heart                    | 177    | Action Adventure     | 1995 | Paramount Pictures | R      |
| 2  | Casablanca                     | 103    | Drama                | 1942 | MGM / UA           | PG     |
| 3  | Christmas Vacation             | 97     | Comedy               | 1989 | Warner Brothers    | PG-13  |
| 4  | Coming to America              | 116    | Comedy               | 1988 | Paramount Pictures | R      |
| 5  | Dracula                        | 130    | Horror               | 1993 | Columbia TriStar   | R      |
| 6  | Dressed to Kill                | 105    | Drama Mysteries      | 1980 | Filmways Pictures  | R      |
| 7  | Forrest Gump                   | 142    | Drama                | 1994 | Paramount Pictures | PG-13  |
| 8  | Ghost                          | 127    | Drama Romance        | 1990 | Paramount Pictures | PG-13  |
| 9  | Jaws                           | 125    | Action Adventure     | 1975 | Universal Studios  | PG     |
| 10 | Jurassic Park                  | 127    | Action               | 1993 | Universal Pictures | PG-13  |
| 11 | Lethal Weapon                  | 110    | Action Cops & Robber | 1987 | Warner Brothers    | R      |
| 12 | Michael                        | 106    | Drama                | 1997 | Warner Brothers    | PG-13  |
| 13 | National Lampoon's<br>Vacation | 98     | Comedy               | 1983 | Warner Brothers    | PG-13  |
| 14 | Poltergeist                    | 115    | Horror               | 1982 | MGM / UA           | PG     |
| 15 | Rocky                          | 120    | Action Adventure     | 1976 | MGM / UA           | PG     |
| 16 | Scarface                       | 170    | Action Cops & Robber | 1983 | Universal Studios  | В      |
| 17 | Silence of the Lambs           | 118    | Drama Suspense       | 1991 | Orion              | B      |
| 18 | Star Wars                      | 124    | Action Sci-Fi        | 1977 | Lucas Film Ltd     | PG     |
| 19 | The Hunt for Red October       | 135    | Action Adventure     | 1989 | Paramount Pictures | PG     |
| 20 | The Terminator                 | 108    | Action Sci-Fi        | 1984 | Live Entertainment | R      |
| 21 | The Wizard of Oz               | 101    | Adventure            | 1939 | MGM / UA           | G      |
| 22 | Titanic                        | 194    | Drama Romance        | 1997 | Paramount Pictures | PG-13  |

### **MOVIES Microsoft Excel File**



The data stored in the ACTORS table is illustrated below.

### **ACTORS Data Set**

|    | Title                       | Actor_Leading        | Actor_Supporting |
|----|-----------------------------|----------------------|------------------|
| 1  | Brave Heart                 | Mel Gibson           | Sophie Marceau   |
| 2  | Christmas Vacation          | Chevy Chase          | Beverly D'Angelo |
| 3  | Coming to America           | Eddie Murphy         | Arsenio Hall     |
| 4  | Forrest Gump                | Tom Hanks            | Sally Field      |
| 5  | Ghost                       | Patrick Swayze       | Demi Moore       |
| 6  | Lethal Weapon               | Mel Gibson           | Danny Glover     |
| 7  | Michael                     | John Travolta        | Andie MacDowell  |
| 8  | National Lampoon's Vacation | Chevy Chase          | Beverly D'Angelo |
| 9  | Rocky                       | Sylvester Stallone   | Talia Shire      |
| 10 | Silence of the Lambs        | Anthony Hopkins      | Jodie Foster     |
| 11 | The Hunt for Red October    | Sean Connery         | Alec Baldwin     |
| 12 | The Terminator              | Arnold Schwarzenegge | Michael Biehn    |
| 13 | Titanic                     | Leonardo DiCaprio    | Kate Winslet     |

# **Exploring Enterprise Guide**

Enterprise Guide (EG) provides users with a graphical user interface (GUI) to make programming tasks easier. Once EG is started you'll see the 'Welcome to SAS Enterprise Guide' dialog. Users can select an existing project from the list of available projects displayed under the 'Open a project' heading; New Project, New SAS Program and New Data under the 'New' heading; or request assistance under the 'Assistance' heading, as illustrated in Figure 1.



Figure 1. Welcome to SAS Enterprise Guide dialog

We'll begin exploring EG's many capabilities by selecting 'New Project'. Once a new project is initiated, EG's three main windows appear: Project Explorer, Project Designer, and Task Status, as illustrated in Figure 2.

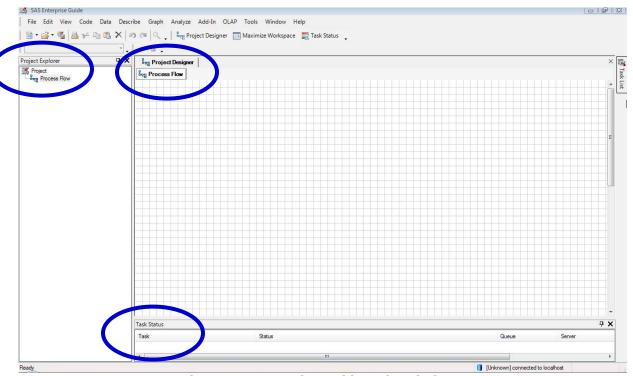


Figure 2. Enterprise Guide Main Windows

Additional windows can be opened using the point-and-click capabilities found in EG. Once open, a tab displays at the top of the screen to enable navigation to other windows. For example, a list of available tasks can be displayed by clicking the "Task List" button located at the right of the EG main windows, as Figure 3 illustrates.

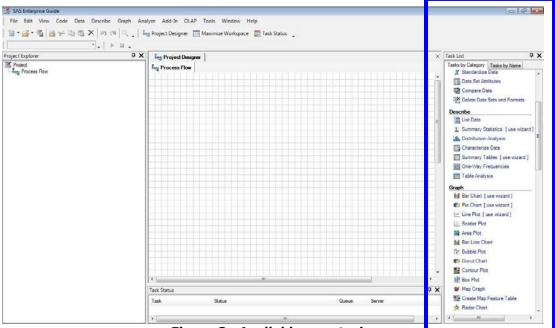


Figure 3. Available user tasks

Tasks under the 'Tasks by Category' tab are displayed within the following functional categories: Data, Describe, Graph, ANOVA, Regression, Multivariate, Survival Analysis, Capability, Control Charts, Pareto, Time Series, Model Scoring, and Tools, as illustrated in Figure 4. Tasks under the 'Tasks by Name' tab are displayed in alphabetical task name order along with each task associated SAS Procedure, as illustrated in Figure 5.

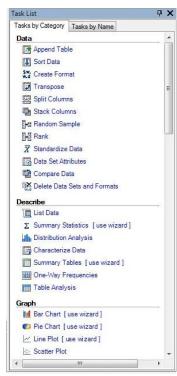


Figure 4. Task List by Category

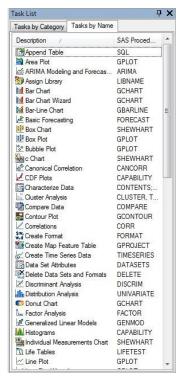


Figure 5. Task List by Name

### **Accessing Multiple Data Sources**

SAS EG has the ability to access a variety of remote servers, including Windows, Unix, and IBM mainframe operating systems, containing data from many types of input data sources. From text files to SAS data sets; Windows data sources including Microsoft Excel, Microsoft Access, Lotus, Paradox, and HTML; relational database tables including Oracle, DB2, SQL-Server, MySQL, among others; and ODBC, Microsoft Exchange folders, and OLE DB, EG is capable of adding data files to a project using **View ... Server List** and/or **File** and **Import Data...** .

### Importing SAS Data

To illustrate the process of importing a SAS data set located on the authors' local computer, the 'Local Computer' icon is clicked on the **Open Data** dialog as illustrated in Figure 6.



Figure 6. Open Data dialog

The data importation process illustrated in Figure 7 demonstrates the selection of the Movies data set for import purposes, the entire data set (all rows and columns) imported and made available to EG as a SAS data set in 'read-only' mode, and finally after the successful completion of the requested task the data set is created and opened in 'read-only' mode.

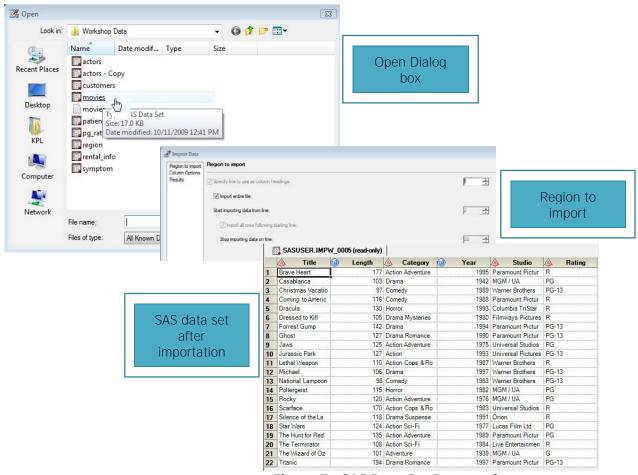


Figure 7. SAS Data Set Importation process

As an added bonus, EG provides users with a convenient way to view any, and all, SAS Log messages and task-specific EG-generated SAS code following the completion of the requested importation task. Figure 8 and Figure 9 illustrate the available log messages and task-generated SAS code from the specific data set importation task respectively.

```
Log (Import Data (movies (Process Flow)))
read-only) Last Submitted Code (Import Data (movies (Process Flow)))
 ⊟1
                                                                       The SAS System
                ;*';*";*/;quit;run;
                OPTIONS PAGENO=MIN;
                LIBNAME ECLIB000 "E:\Workshops\Workshop Data";
   NOTE: Libname ECLIB000 refers to the same physical library as EC100017.
   NOTE: Libref ECLIB000 was successfully assigned as follows:
          Physical Name: E:\Workshops\Workshop Data
                %LET CLIENTTASKLABEL=%NRBQUOTE(Import Data);
                %LET _EGTASKLABEL=%NRBQUOTE(Import Data);
%LET _CLIENTPROJECTNAME=%NRBQUOTE();
               ODS ALL CLOSE;
   10
                ODS PROCTITLE;
   12
               OPTIONS DEV=ACTIVEX:
   NOTE: Procedures may not support all options or statements for all devices. For detail
   13
               FILENAME EGHTML TEMP:
                ODS HTML(ID=EGHTML) FILE=EGHTML ENCODING='utf-8' STYLE=EGDefault
             ! STYLESHEET=(URL="file:///C:/Program%20Files/SAS/Shared%20Files/BIClientStyl ! ATTRIBUTES=("CODEBASE"="http://www2.sas.com/codebase/graph/v91/sasgraph.exe
   NOTE: Writing HTML (EGHTML) Body file: EGHTML
   WARNING: Argument 1 to function TSLVL referenced by the %SYSFUNC or %QSYSFUNC macro fu
   17
```

Figure 8. SAS Log results

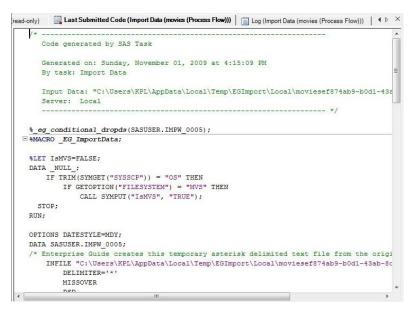


Figure 9. SAS generated code

### Importing Tab-delimited Files

To further illustrate the data importation process we'll look at the process of importing a tab-delimited file. As before, the specific text file is located on the authors' local computer, so the 'Local Computer' icon is clicked on the **Open Data** dialog, the Movies (with tabs) file is selected, with the entire file (all rows and columns) selected for import, and converted and opened as a SAS data set in 'read-only' mode, as illustrated in Figure 10.

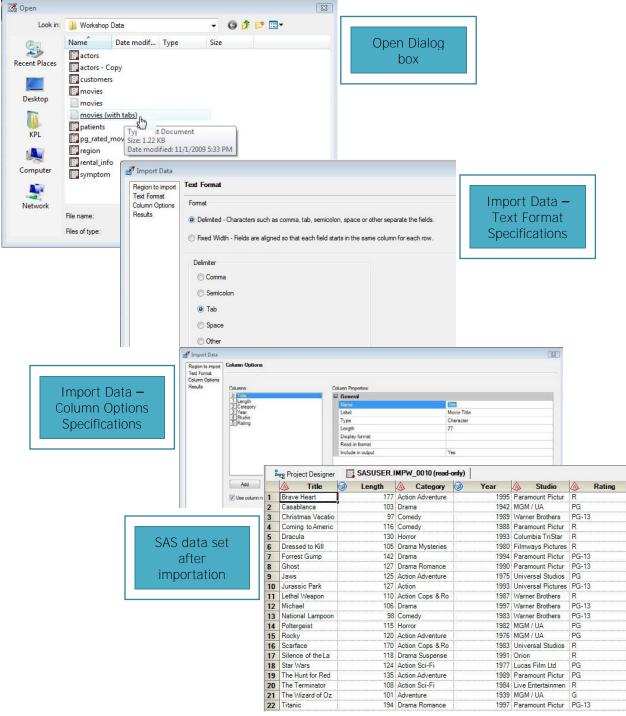


Figure 10. Tab-delimited File Importation process

### Importing Microsoft Excel Files

Finally, to illustrate the flexibility and power of the data importation process, we'll look at the process of importing a Microsoft Excel file. As with the previous data importation examples, the specific Excel file is located on the authors' local computer. The Excel file, Movies, is selected; the entire file (all rows and columns) selected for import; and converted and opened as a data set in 'read-only' mode, as illustrated in Figure 11.

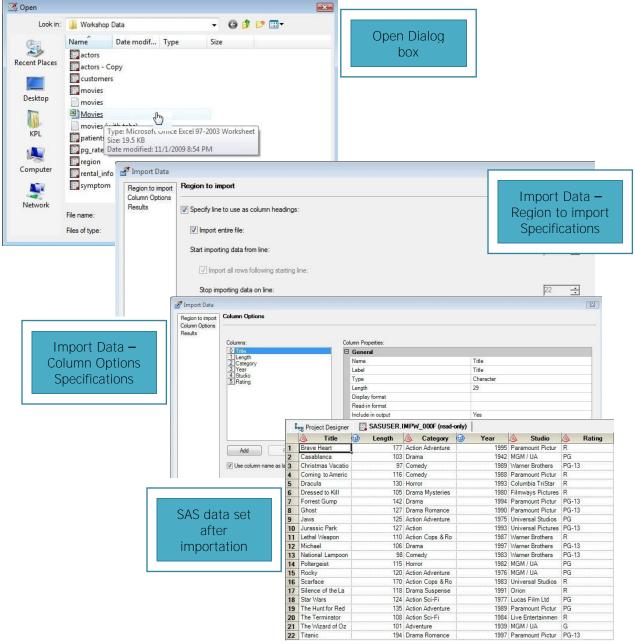


Figure 11. Excel File Importation process

### Manipulating Data - No Programming Required

EG provides users with powerful point-and-click data summarization and manipulation capabilities without the need to learn formal programming language techniques. Supported features include recoding data values, sorting or rearranging the data order, producing descriptive statistics, merging (or joining) tables of data, transposing data, data concatenation, and comparing data. Due to size restrictions of this paper we'll confine our attention to illustrating the production of descriptive statistics and a match merge (or join) operation on the Movies and Actors data sets.

### Producing "Quick and Dirty" Descriptive Statistics

In order to perform most types of analyses, it is necessary to fully understand your data. In addition to cleaning and organizing the data, the first stage should always include using descriptive statistics to obtain some basic measures for each variable. All of these tasks can be performed within EG by using Wizards without the use of complex programming. SAS EG Wizards provide the ability to produce "quick and dirty" descriptive statistics. The Wizard allows the user to select a task by category or name, select/verify the data source, assign variables to roles, select the desired statistics and result types, customize the output and create the report.

In the first set of steps, the Wizard guides the user through the task selection process, verifying the data source and assigning variables to roles. The Wizard guides users by providing a list of variables to assign as categorical and continuous, as illustrated in Figure 12.

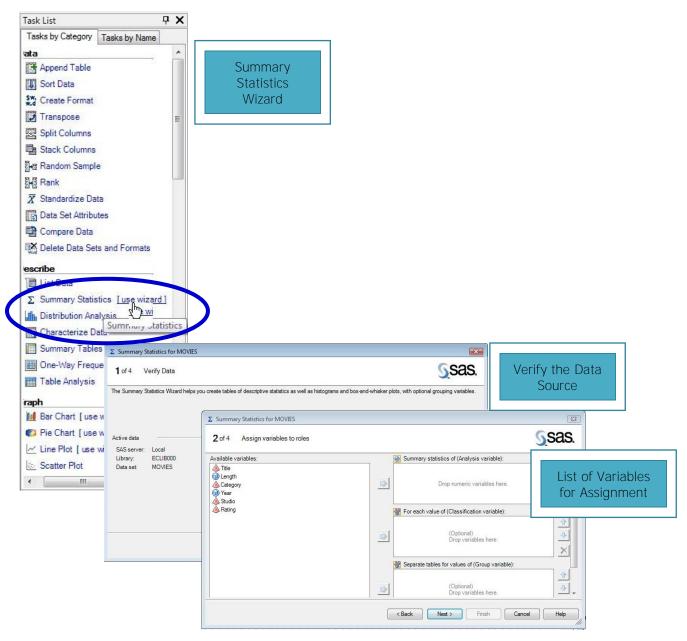


Figure 12. Process of producing "Quick and Dirty" Statistics (Part 1)

After specifying the variables and their roles, the Wizard guides the user through the selection of statistics and results. Titles and footnotes for the final report can be easily specified using the Wizard, as illustrated in Figure 13.

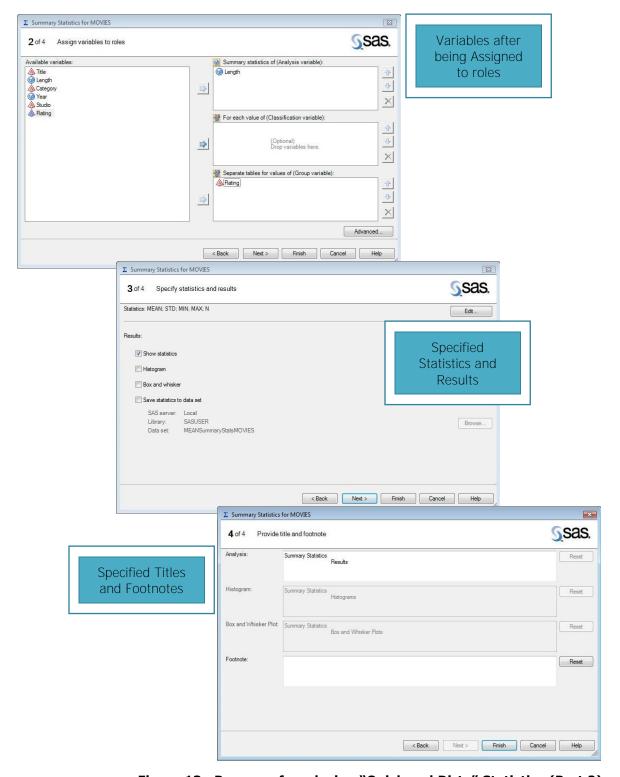


Figure 13. Process of producing "Quick and Dirty" Statistics (Part 2)

Finally, the Wizard runs the report with the selected options and displays the results. As with other reports and summaries created in EG, numerous options are available for embellishing and exporting results. Additional options for descriptive measures and options for complex statistical analysis are available through EG. In this example the mean, minimum, maximum and standard deviation were calculated, as illustrated in Figure 14.

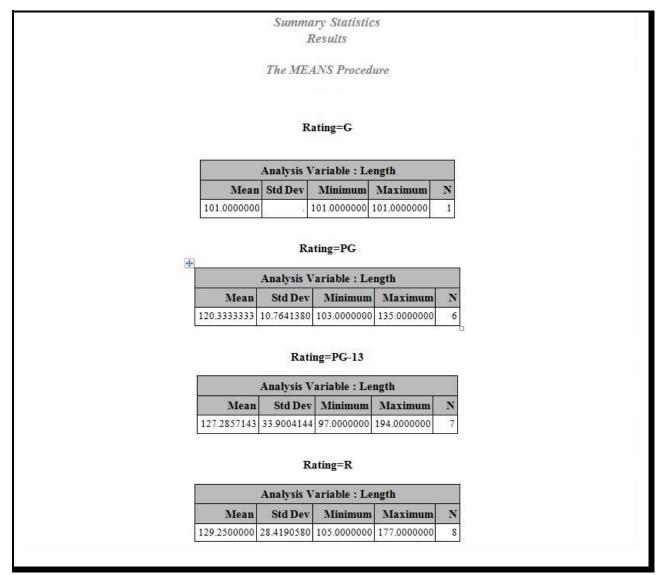


Figure 14. "Quick and Dirty" Statistics Output

### Manipulating Data with Merges (or Joins)

A merge (or join) of two or more tables provides a way to bring data together horizontally. The process requires a minimum of two tables, where a column from each table is used for the purpose of connecting the tables. Connecting columns should have "like" values and is most successful when the joining columns have the same datatype attributes. The following task applies a match-merge process using the TITLE value in both tables as the matching column, as illustrated in Figure 15.

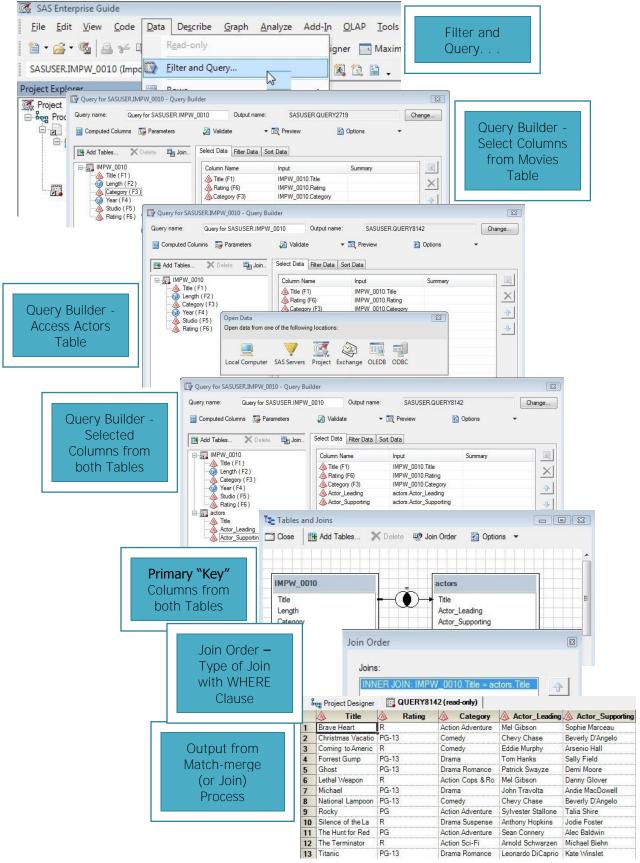


Figure 15. Match-merge process

# Show Me the Results - Creating "Custom" Reports

SAS EG provides numerous point-and-click features designed for reporting and presentation. The GUI front-end is designed to be simple to use, and is what differentiates SAS from other software products. EG and its built-in capabilities offer users a unique ability to generate quick results – requiring little, if any, programming skills. In the following examples we'll see how EG can be used to export results to HTML and Microsoft Excel.

# **Exporting Results to HTML**

With the widespread use of the Internet, EG and Output Delivery System (ODS) combine to turn tired-looking monospace output into great looking information using Hyper-text Markup Language (HTML). EG and ODS take the pain out of creating and deploying selected pieces of SAS output in HTML format by providing a level of control without the need to learn complicated coding techniques, illustrated in Figure 16. The HTML-generated output can be deployed to a server (e.g., the Web, Intranet, and Extranet), or a stand-alone workstation for easy access using a Web browser such as Internet Explorer, Firefox, or Netscape Navigator. As you explore the power of EG and ODS, you'll begin to appreciate the relative ease in delivering SAS output and data to HTML.

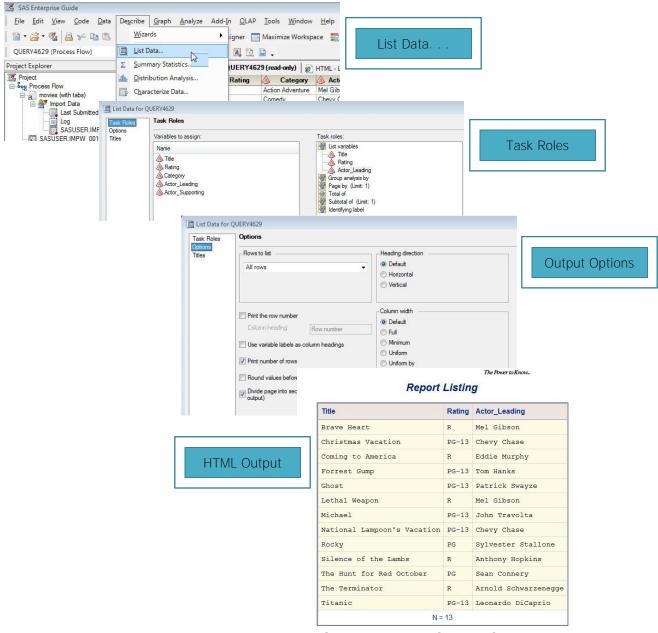


Figure 16. Exporting results to HTML

### Exporting Results to Microsoft Excel

Microsoft Excel is not only one of the most widely used software products in the world; it is without a doubt an essential component in an organization's inventory of mission-critical software tools. Figure 17 illustrates the process of using EG to deliver data and results to Microsoft Excel. EG makes creating Microsoft Excel output from data and/or selected pieces of SAS output as easy as 1-2-3.

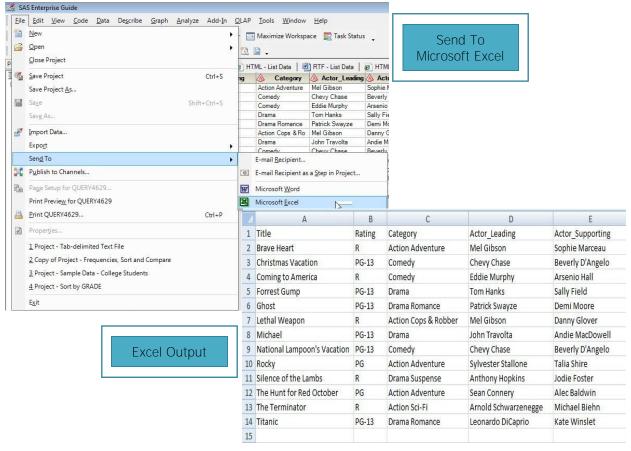


Figure 17. Exporting results to Microsoft Excel

# **Accessing Flow Diagrams and Generated Code**

EG provides users with application-generated flow diagrams for visually organizing, viewing, and managing projects. These process and flow diagrams are important system and application documentation components. As illustrated in Figure 18 and 19, input and output data sources, along with "key" processes are readily available with a saved project.

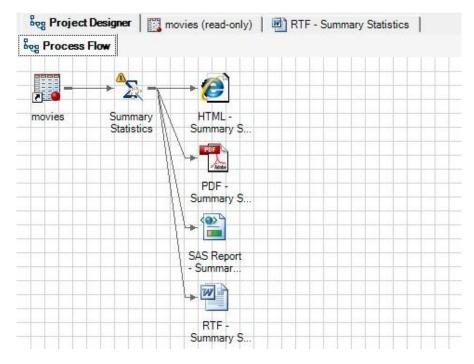


Figure 18. Project Designer - "Quick and Dirty" Statistics Process Flow diagram

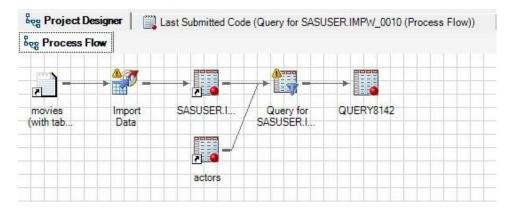


Figure 19. Project Designer - Merge (or Join) Process Flow diagram

Another wonderful feature built into EG is its ability to provide users with application-generated syntactically-correct SAS code. As Figures 20 and 21 illustrate, EG's point-and-click steps along with all user selected options for producing "quick and dirty" descriptive statistics and the match-merging (or joining) process presented earlier generated an assortment of SAS programming code including SORT, SQL, and MEANS procedure code. EG provides users with working code to help learn the many programming techniques available in the SAS System, the ability to execute the generated code without having to revisit the numerous steps provided through the GUI, as well as the actual source code for system documentation purposes.

```
%macro _SASTASK_DROPDS(dsname);
    %IF %SYSFUNC(EXIST(&dsname)) %THEN %DO;
            DROP TABLE &dsname;
    %END;
    %IF %SYSFUNC(EXIST(&dsname, VIEW)) %THEN %DO;
             DROP VIEW &dsname;
    %END;
%mend _SASTASK_DROPDS;
%LET _EGCHARTWIDTH=0;
%LET _EGCHARTHEIGHT=0;
 Code generated by SAS Task
 Generated on: Sunday, November 01, 2009 at 6:34:59 PM
 By task: Summary Statistics
 Input Data: ECLIB000.MOVIES
 Server: Local
PROC SQL:
%_SASTASK_DROPDS(WORK.SORTTempTableSorted);
QUIT;
Sort data set ECLIB000.MOVIES
   DATA=ECLIB000.MOVIES(KEEP=Length Rating)
    OUT=WORK.SORTTempTableSorted
    BY Rating;
RUN;
Run the Means Procedure
TITLE;
TITLE1 "Summary Statistics";
TITLE2 "Results";
FOOTNOTE;
PROC MEANS DATA=WORK.SORTTempTableSorted
    FW=12
    PRINTALLTYPES
    CHARTYPE
    VARDEF=DF
             MEAN
             STD
             MIN
             MAX
    VAR Length;
    BY Rating;
RUN;
End of task code.
RUN; QUIT;
PROC SOL:
%_SASTASK_DROPDS(WORK.SORTTempTableSorted);
QUIT;
TITLE; FOOTNOTE;
```

Figure 20. Project Explorer - Generated Code for "Quick and Dirty" Statistics Process

Figure 21. Project Explorer - Generated SQL Code for Merge (or Join) Process

### Conclusion

SAS® Enterprise Guide® (EG) empowers an organization's end-users with a powerful graphical user interface (GUI) environment for exploiting a multitude of data, analytical, and reporting tasks. EG provides access to multi-platform enterprise data sources including SAS data sets, tab-delimited data, and Microsoft Excel files; create "custom" report generation; deliver data and results to a variety of mediums and outlets including HTML and Microsoft Excel; produce "quick and dirty" descriptive statistics; perform data manipulations without the need to learn complex coding constructs; while supporting data management and documentation requirements by producing system flowcharts and diagrams quickly and easily using the built-in wizards.

#### References

- Delwiche, Lora D. and Susan J. Slaughter (2006), "Producing Summary Tables in SAS® Enterprise Guide®," Proceedings of the 2006 NorthEast SAS Users Group (NESUG) Conference.
- Hemedinger, Chris (2005), "Boost Your Programming Productivity with SAS® Enterprise Guide®," Proceedings of the Thirtieth SAS Users Group International (SUGI) Conference, SAS Institute Inc., Cary, NC, USA.
- Hettinger, Patricia (2009), "Tips for Moving SAS® Enterprise Guide® on Unix," Proceedings of the 2009 SouthEast SAS Users Group (SESUG) Conference.
- Lafler, Kirk Paul (2004), "Creating HTML Output with Output Delivery System," Proceedings of the 2004 Western Users of SAS Software (WUSS) Conference, Software Intelligence Corporation, Spring Valley, CA, USA.
- SAS® Software Essentials Using SAS Enterprise Guide® Course Notes, First Edition (2009). Software Intelligence Corporation, Spring Valley, CA, USA.
- Delwiche, Lora D. and Susan J. Slaughter (2006), "Producing Summary Tables in SAS® Enterprise Guide®," Proceedings of the 2006 North East SAS Users Group (NESUG) Conference.
- Shapiro, Mira and Kirk Paul Lafler (2010), "Point-and-Click Programming Using  $SAS^{\otimes}$  Enterprise  $Guide^{\otimes}$ ," Proceedings of the 2010 MidWest SAS Users Group (MWSUG) Conference.
- Shapiro, Mira and Kirk Paul Lafler (2010), "Point-and-Click Programming Using SAS® Enterprise Guide®," Proceedings of the 2010 Western Users of SAS Software (WUSS) Conference.
- Shapiro, Mira and Kirk Paul Lafler (2010), "Point-and-Click Programming Using  $SAS^{\otimes}$  Enterprise  $Guide^{\otimes}$ ," Proceedings of the 2010 South East SAS Users Group (SESUG) Conference.
- Todd, Michael (2008), "Transitioning to SAS® Enterprise Guide®," Proceedings of the 2008 NorthEast SAS Users Group (NESUG) Conference.

### **Acknowledgments**

The authors would like to thank George Fernandez, SGF 2011 Applications Development Section Chair for accepting our abstract and paper, as well as Debbie Buck, SGF 2011 Conference Chair for a great conference!

### **Trademark Citations**

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.

# **Author Bios**

Mira Shapiro is Principal Consultant of Analytic Designers LLC and has been a SAS user since 1979 and is currently serving as SAS User Liaison for DC-SUG, a Washington-DC SAS Users Group. She has used SAS throughout her career as a Capacity Planner, Consultant and Biostatistician. She holds a BA in Statistics / Computer Science and an MS in Public Health / Biostatistics and works on analytics and pre-sales projects across multiple industries.

Kirk Paul Lafler is consultant and founder of Software Intelligence Corporation and has been using SAS since 1979. He is a SAS Certified Professional, provider of IT consulting services, trainer to SAS users around the world, and sasCommunity.org Advisory Board member. As the author of four books including PROC SQL: Beyond the Basics Using SAS, Kirk has written more than four hundred peer-reviewed papers, been an Invited speaker and trainer at more than three hundred SAS International, regional, local, and special-interest user group conferences and meetings throughout North America, and is the recipient of 17 "Best" contributed paper awards. His popular SAS Tips column, "Kirk's Korner of Quick and Simple Tips", appears regularly in several SAS User Group newsletters and Web sites, and his fun-filled SASword Puzzles is featured in SAScommunity.org.

#### **Contact Information**

Mira Shapiro E-mail: Mira.Shapiro@gmail.com

~ ~ ~

Kirk Paul Lafler E-mail: <u>KirkLafler@cs.com</u>