

# SAS® GLOBAL FORUM 2017

April 2 - 5 | Orlando, FL

## **Quick Results with SAS® Enterprise Guide®**

*A Hands-on Workshop by*

USERS PROGRAM

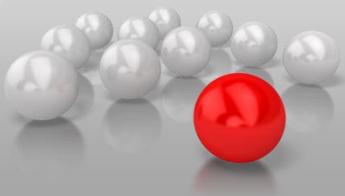


Kirk Paul Lafler, Mira Shapiro and Ryan Paul Lafler

Copyright © 2010-2017 by  
Kirk Paul Lafler and Software Intelligence Corporation.  
All rights reserved.

SAS is the registered trademark of SAS Institute Inc., Cary, NC,  
USA.

All other company and product names mentioned are used for  
identification purposes only and may be trademarks of their  
respective owners.



# Workshop Objectives



**Introducing  
and  
Launching  
SAS  
Enterprise  
Guide (EG)**



**Accessing  
Multi-  
platform Data  
Sources**



**Summarizing  
Statistics /  
Data**



**Manipulating  
Data**



**Managing  
Projects and  
Flowcharts**

# Tables Used in Exercises

	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 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	R
17	Silence of the Lambs	118	Drama Suspense	1991	Orion	R
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**

**Actors**

	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 Schwarzenegger	Michael Biehn
13	Titanic	Leonardo DiCaprio	Kate Winslet

# Excel File Used in Exercises

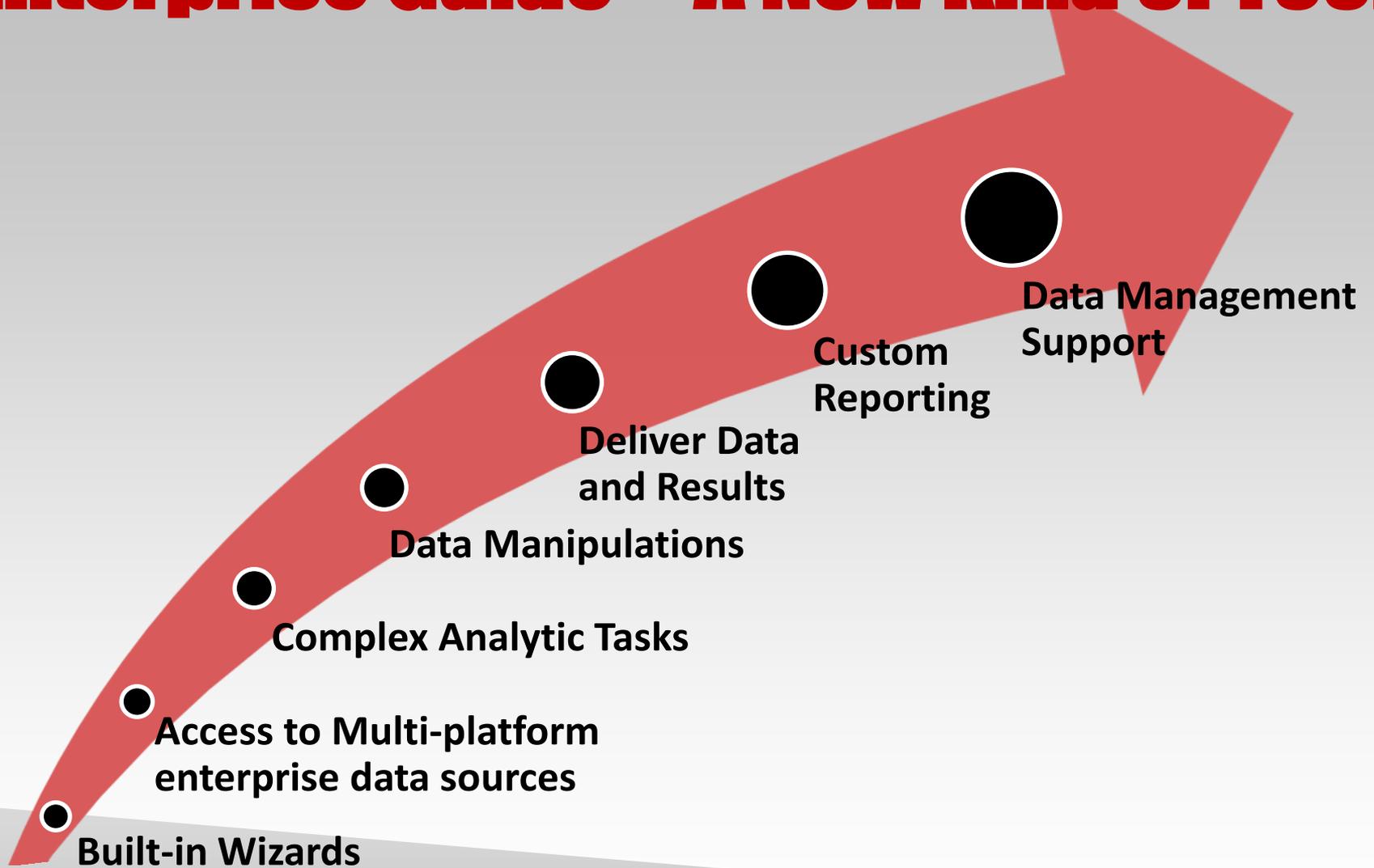
	A	B	C	D	E	F
1	Title	Length	Category	Year	Studio	Rating
2	Brave Heart	177	Action Adventure	1995	Paramount Pictures	R
3	Casablanca	103	Drama	1942	MGM / UA	PG
4	Christmas Vacation	97	Comedy	1989	Warner Brothers	PG-13
5	Coming to America	116	Comedy	1988	Paramount Pictures	R
6	Dracula	130	Horror	1993	Columbia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

**Movies.xls**

# **Introducing and Launching Enterprise Guide (EG)**



# Enterprise Guide – A New Kind of Tool



# Enterprise Guide (EG) Features

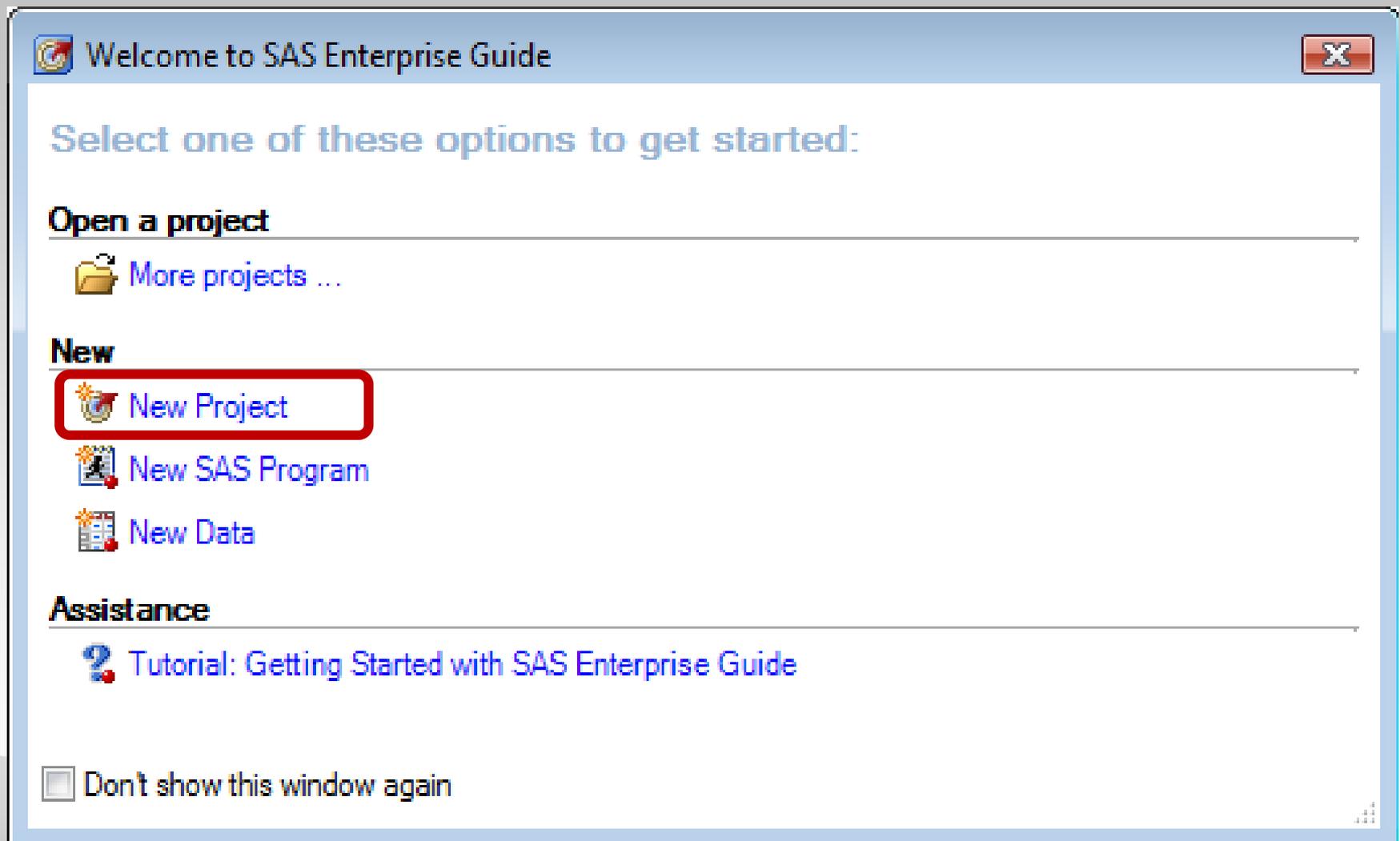
- Graphical User Interface (GUI)
- Point-and-click interaction
- Makes programming tasks easier
- Enterprise Guide (EG) generates SAS code:
  - ✓ PROC PRINT
  - ✓ PROC SQL
  - ✓ PROC MEANS
  - ✓ PROC TABULATE
  - ✓ and other PROCs
- Submits code to the SAS software

# Exercise #1 – Launching EG

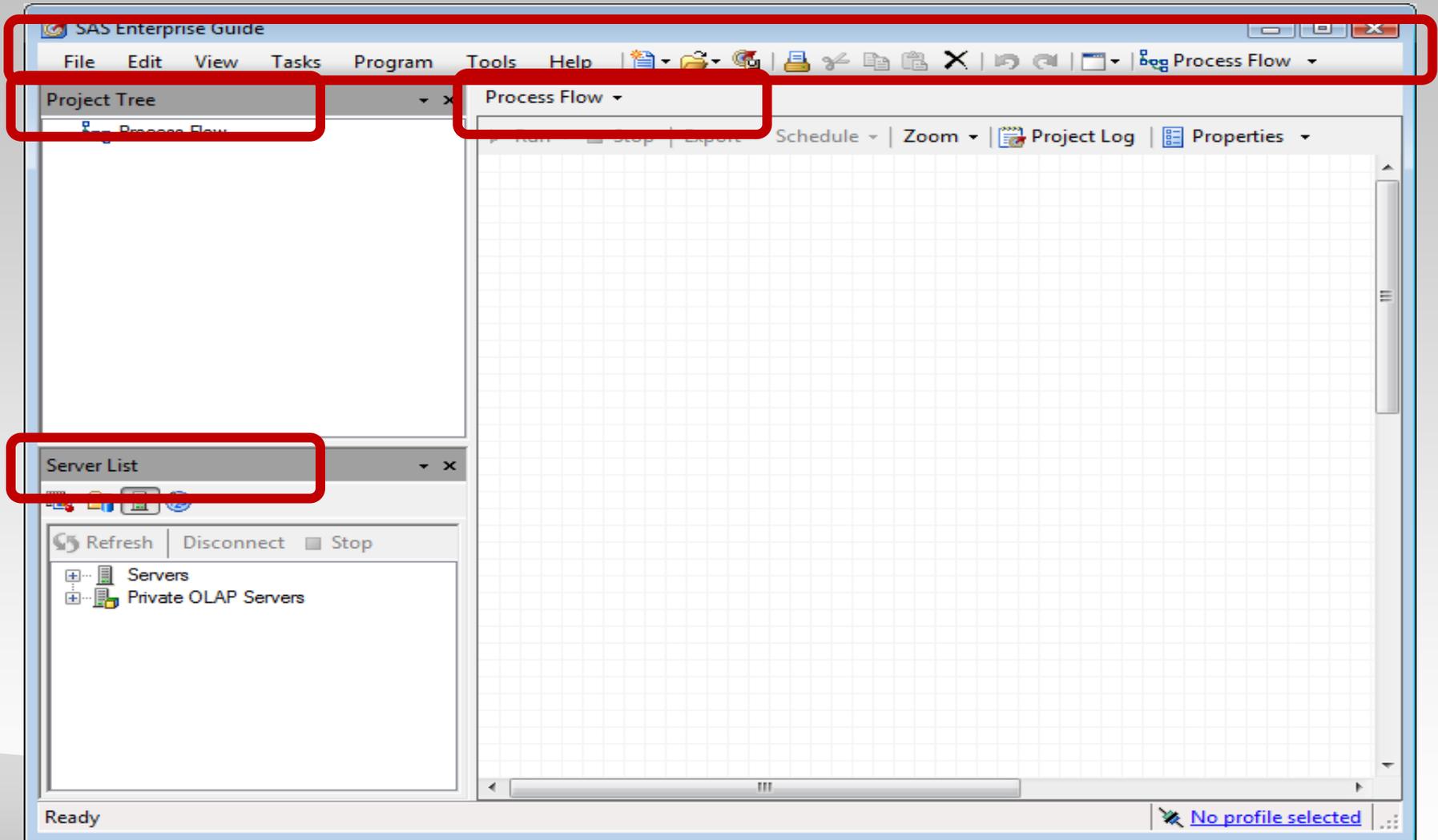
1. Launch Enterprise Guide by clicking the  icon located on the desktop.
2. The Enterprise Guide application should be active.
3. Let's discuss the exercise and corresponding display.



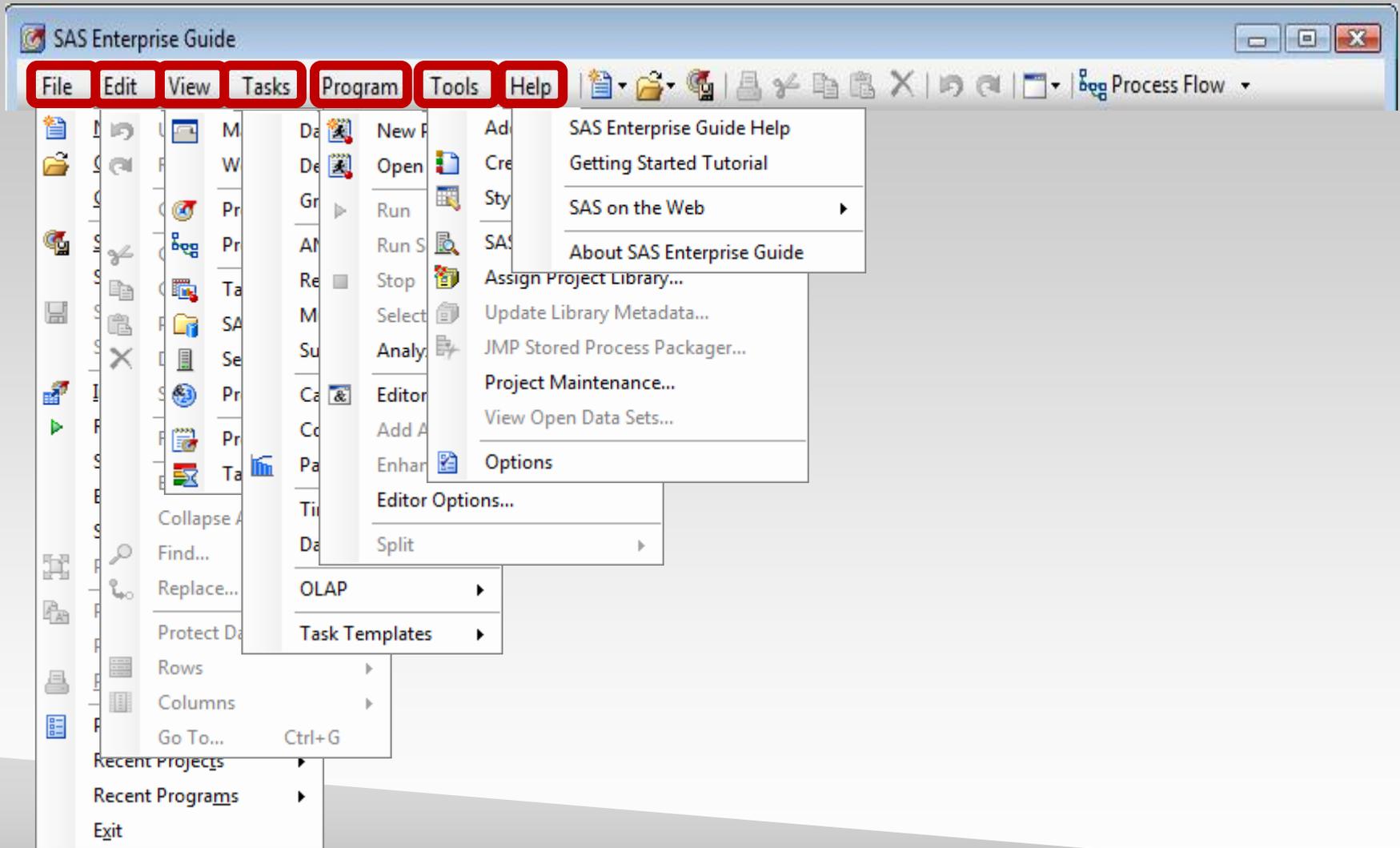
# Exercise #1 – EG's Welcome Screen



# Exercise #1 – EG's Main Window



# Exercise #1 – EG's Menu Bar



# Accessing Multi-platform Data Sources



# Accessing Multi-platform Data

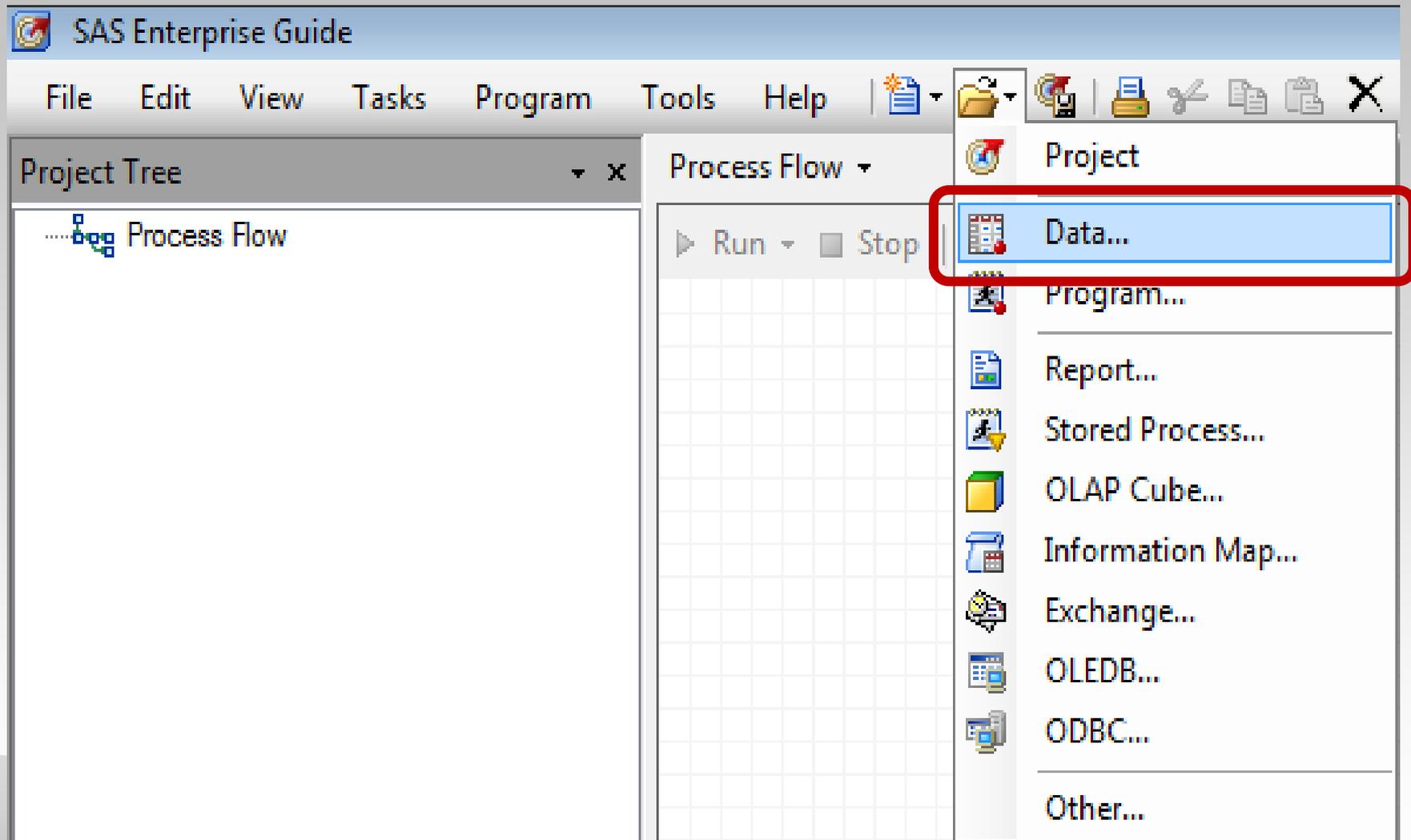
- SAS data sets
- Text files
- Microsoft Excel
- Microsoft Access
- SPSS
- Stata
- JMP
- Relational database tables:
  - ✓ Oracle
  - ✓ MySQL
  - ✓ DB2
  - ✓ Teradata
  - ✓ SQL-Server
  - ✓ Other databases

# Exercise #2 – Reading MOVIES Table

1. Click EG's "Open"  icon on the main menu bar.
2. Select "Data" from the drop-down list.
3. Select the MOVIES dataset in the Open Data window and click the Open button to read the data into EG.
4. Let's discuss the exercise and the resulting dataset.



# Exercise #2 (continued)



# Exercise #2 (continued)

Open Data Window

The screenshot shows the SAS Open Data window with the following details:

- Window title: Open Data
- Buttons: Browse, Search SAS Folders
- Look in: Workshop Data
- Left sidebar: Desktop, My Documents, My Computer, My Network Places, Servers, SAS Folders
- File list table:

Name	Date modif...	Type	Size
SAS Tables from PROC SQL Book			
actors			
actors_dups			
comedy			
customers			
<b>movies</b>			
Movies		SAS Data Set	17.0 KB
movies (wi			
movies_du			
my_extremeobs			
mymovies			
patients			
pg_rated_movies			
region			
rental_info			
symptom			

File name: movies.sas7bdat  
Files of type: All Known Data Files (\*.sas7bdat;\*.sas7bview;\*.sd2;\*.mdb;\*.acccdb;\*.dbf;\*.xls;\*.xlsx;\*.xslm)

Buttons: Open, Cancel



# Exercise #2 (continued)

MOVIES dataset  
(table) after being  
read into EG

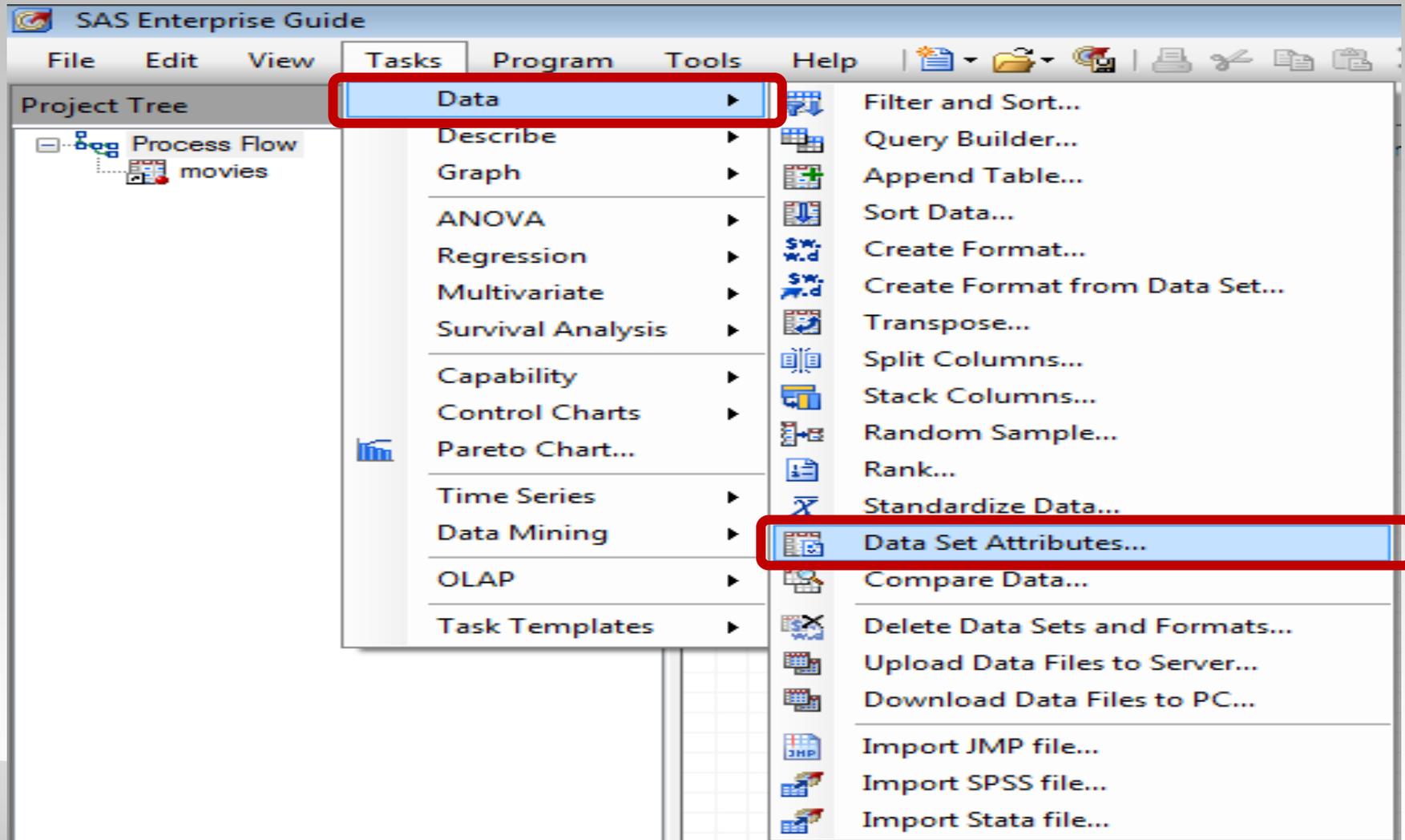
	⚠ Title	12 Length	⚠ Category	123 Year	⚠
1	Brave Heart	177	Action Adventure	1995	Param
2	Casablanca	103	Drama	1942	MGM / UA
3	Christmas Vacati...	97	Comedy	1989	Warner Brothers PG-13
4	Coming to Ameri...	116	Comedy	1988	Paramount Pictu... R
5	Dracula	130	Horror	1993	Columbia TriStar R
6	Dressed to Kill	105	Drama Mysteries	1980	Filmways Pictures R
7	Forrest Gump	142	Drama	1994	Paramount Pictu... PG-13
8	Ghost	127	Drama Romance	1990	Paramount Pictu... 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 & R...	1987	Warner Brothers R
12	Michael	106	Drama	1997	Warner Brothers PG-13
13	National Lampoo...	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 & R...	1983	Universal Studios R
17	Silence of the La...	118	Drama Suspense	1991	Orion R
18	Star Wars	124	Action Sci-Fi	1977	Lucas Film Ltd PG
19	The Hunt for Red...	135	Action Adventure	1989	Paramount Pictu... PG
20	The Terminator	108	Action Sci-Fi	1984	Live Entertainme... R
21	The Wizard of Oz	101	Adventure	1939	MGM / UA G
22	Titanic	194	Drama Romance	1997	Paramount Pictu... PG-13

# Exercise #3 – Dataset Attributes

1. Click EG's "Tasks" tab on the main menu bar.
2. Select "Data" from the Tasks drop-down list.
3. Select "Data Set Attributes..." from the Data drop-down list.
4. Select the desired options from the "Data Set Attributes" Wizard.
5. Let's discuss the exercise and the resulting output.



# Exercise #3 (continued)



# Exercise #3 (continued)

Dataset  
Attributes  
Window  
(Part 1 of 2)

Data Set Attributes for E:\Workshops\Workshop Data\movies.sas7bdat

1 of 2 Select the data you want to report on.

The Data Set Attributes Wizard helps you create a report with the data set's creation date, location, and number of observations as well as the variable names, labels, types, and formats.

SAS server: Local

Data:

Add...  
Remove

<Back Next> Finish Cancel Help



# Exercise #3 (continued)

Dataset  
Attributes  
Window  
(Part 2 of 2)

Data Set Attributes for E:\Workshops\Workshop Data\movies.sas7bdat

2 of 2 Select the output options.

Output data set  
SAS server: Local  
Library: WORK  
Data set: CONTCONTENTSFORMOVIES\_0000 Browse...

Other outputs

Default Report

Enhanced Report

In the data variable table, sort the rows by:

Variable name ▼

Order sequence:

Ascending

Descending

<Back Next> Finish ▼ Cancel Help



# Exercise #3 (continued)

The DATASETS Procedure

Data Set Name	ECLIB000.MOVIES	Observations	22
Member Type	DATA	Variables	6
Engine	V9	Indexes	3
Created	Monday, August 13, 2007 02:49:43 AM	Observation Length	88
Last Modified	Sunday, October 11, 2009 12:41:53 PM	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	WINDOWS_32		
Encoding	wlatin1 Western (Windows)		

Engine/Host Dependent Information	
Data Set Page Size	8192
Number of Data Set Pages	2
First Data Page	1
Max Obs per Page	92
Obs in First Data Page	22
Index File Page Size	4096
Number of Index File Pages	4
Number of Data Set Repairs	0
Filename	E:\Workshops\Workshop Data\movies.sas7bdat
Release Created	9.0101M0
Host Created	XP_PRO

Alphabetic List of Variables and Attributes			
#	Variable	Type	Len
3	Category	Char	20
2	Length	Num	3
6	Rating	Char	5
5	Studio	Char	25
1	Title	Char	



# Exercise #3 (continued)

Alphabetic List of Indexes and Attributes			
#	Index	# of Unique Values	Variables
1	CATRATE	16	Category Rating
2	Category	11	
3	Rating	4	

Generated by the SAS System ('Local', W32\_VSPRO) on October 12, 2011 at 12:50:55 AM

Page Break

## Tables on 'Local'

Table Name	Label	Type	Date Created	Date Modified	Number of Obs.	Char. Set	Password Protected	Data Set Type
MOVIES		DATA	13AUG07:02:49:44	11OCT09:12:41:53	22		---	

Generated by the SAS System ('Local', W32\_VSPRO) on October 12, 2011 at 12:50:55 AM

Page Break

## Variables in Table: MOVIES

Name	Variable Number	Type	Format	Label	Length
Category	3	Character			20
Length	2	Numeric			3
Rating	6	Character			5
Studio	5	Character			25
Title	1	Character			30
Year	4	Numeric			4

Generated by the SAS System ('Local', W32\_VSPRO) on October 12, 2011 at 12:50:55 AM

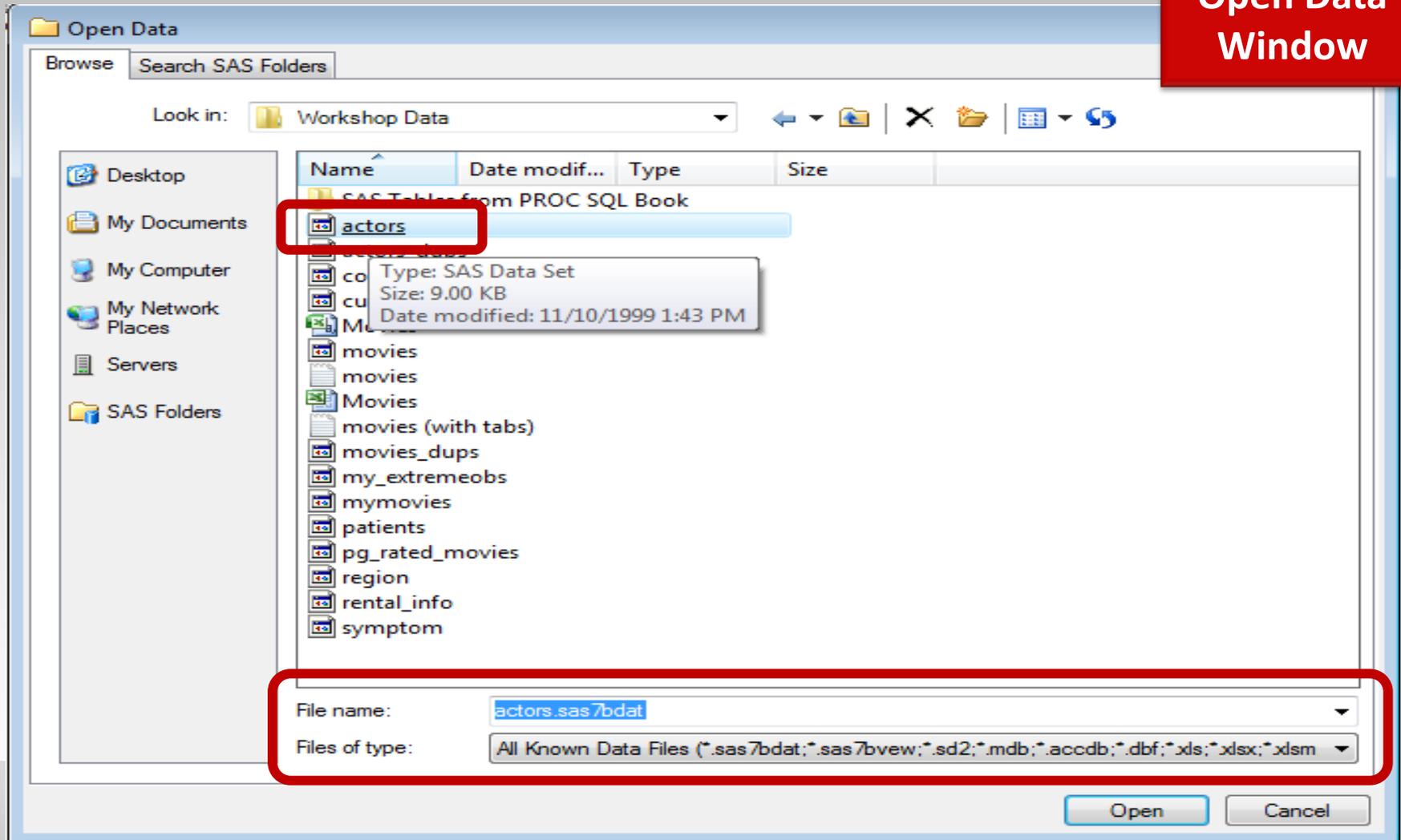
# Exercise #4 – Reading ACTORS Table

1. Following the steps presented in Exercise #2, click EG's "Open"  icon on the main menu bar.
1. Select "Data" from the drop-down list.
2. Select the ACTORS dataset in the Open Data window and click the Open button to read the data into EG.
3. Let's discuss the exercise and the resulting dataset.



# Exercise #4 (continued)

Open Data Window



# Exercise #4 (continued)

ACTORS dataset  
(table) after being  
read into EG

	Title	Actor_Leading	Acto
1	Brave Heart	Mel Gibson	Sophie M
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

# Tab-delimited Data Files

## Tab-delimited Data Files:

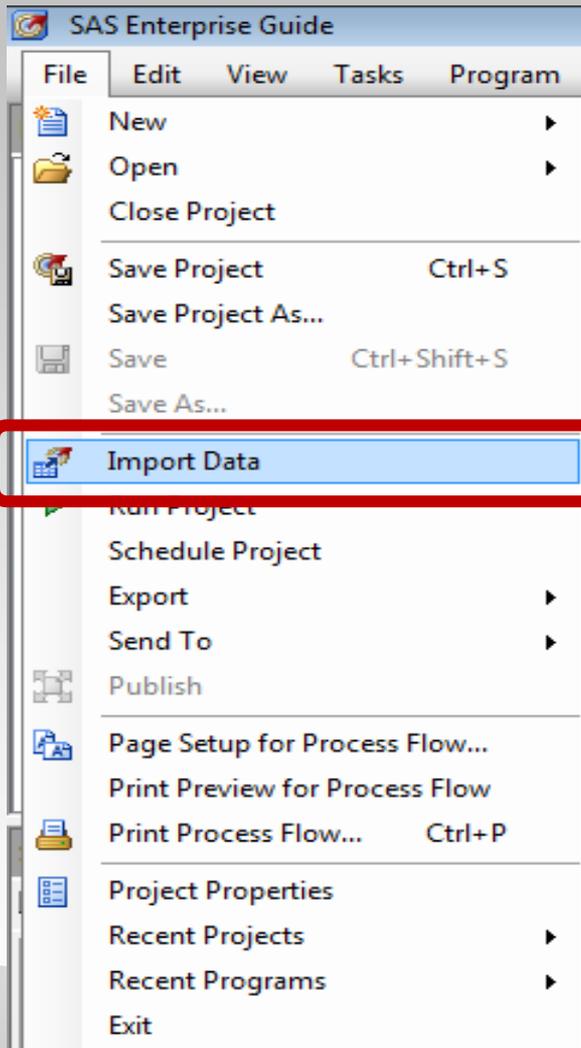
- Store data in a tabular structure
- Use a tab stop character between each value of a record
- Are very common and exported by many database and spreadsheet applications
- Can easily be accessed and processed using SAS Enterprise Guide (EG)

# Exercise #5 – Tab-delimited Data

1. Click the “File” tab on the main menu bar and select “Import Data” from the drop-down list.
2. Select “Importable Data Files - Tab” file type.
3. Select the “Movies (with Tabs)” file from the list.
4. Click the “Open” button.
5. Select the desired options from the Import Data Wizard.
6. Let’s discuss the exercise and the resulting dataset.

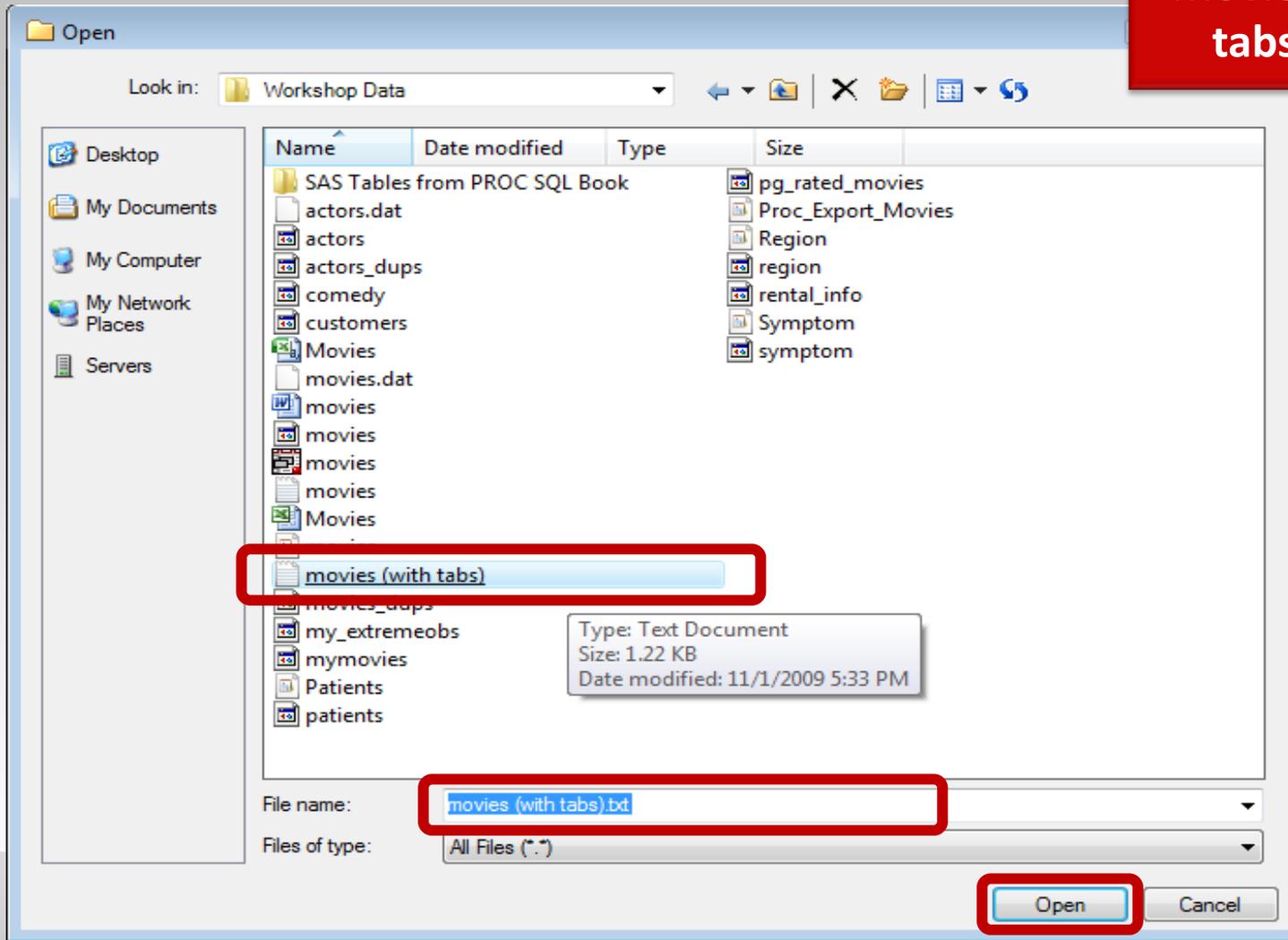


# Exercise #5 (continued)



# Exercise #5 (continued)

Movies (with tabs) File



# Exercise #5 (continued)

Specify Data  
to Import

Import Data from movies (with tabs).txt

1 of 4 Specify Data

The Import Data task is used to convert non-SAS data into a SAS data file which is required by other tasks for data analysis and reporting.

Source data file

Location: Local File System

File path: E:\Workshops\Workshop Data\movies (with tabs).bt

Data type: Text File (Encoding: WINDOWS-1252)

Encoding...

Performance...

Output SAS data set

SAS server: Local

Library: WORK

Data set: movies\_\_with\_tabs\_

Browse...

<Back | Next> | Finish | Cancel | Help



# Exercise #5 (continued)

Select Data Source

Import Data from movies (with tabs).txt

2 of 4 Select Data Source

Text format

Delimited fields

Tab

Text qualifier: "

Fixed columns

File contains field names on record number:

Data records start at record number: 1

Limit the number of records read to:

Rename columns to comply with SAS naming conventions.

Brave Heart	177	Action Adventure	1995	Paramount Pictures		
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	
Dressed to Kill	105	Drama	Mysterie	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	
Silence of the Lambs	118	Drama	Suspense	1991	Orion	R

<Back Next> Finish Cancel Help



# Exercise #5 (continued)

Column (Field)  
Attribute  
Definitions

Import Data from movies (with tabs).txt

3 of 4 Define Field Attributes

Select columns and define attributes:

Inc	Source Name	Name	Label	Type	Source Informat	Len.	Output Format	Output Informat
<input checked="" type="checkbox"/>	F1	F1	F1	String	\$CHAR27.	27	\$CHAR27.	\$CHAR27.
<input checked="" type="checkbox"/>	F2	F2	F2	Number	BEST3.	8	BEST3.	BEST3.
<input checked="" type="checkbox"/>	F3	F3	F3	String	\$CHAR20.	20	\$CHAR20.	\$CHAR20.
<input checked="" type="checkbox"/>	F4	F4	F4	Number	BEST4.	8	BEST4.	BEST4.
<input checked="" type="checkbox"/>	F5	F5	F5	String	\$CHAR18.	18	\$CHAR18.	\$CHAR18.
<input checked="" type="checkbox"/>	F6	F6	F6	String	\$CHAR5.	5	\$CHAR5.	\$CHAR5.

Select All Clear All Modify...

<Back Next> Finish Cancel Help



# Exercise #5 (continued)

Advanced  
Options

Import Data from movies (with tabs).txt

4 of 4 Advanced Options

Embed the data within the generated SAS code.

Import the data using SAS/ACCESS Interface to PC Files whenever possible.

Remove characters that can cause transmission errors from text-based data files.

Generalize import step to run outside SAS Enterprise Guide.

Maximum record length for input text file in bytes:

<Back ▼ Next> **Finish** Cancel Help



# Exercise #5 (continued)

SAS data set  
after  
importation

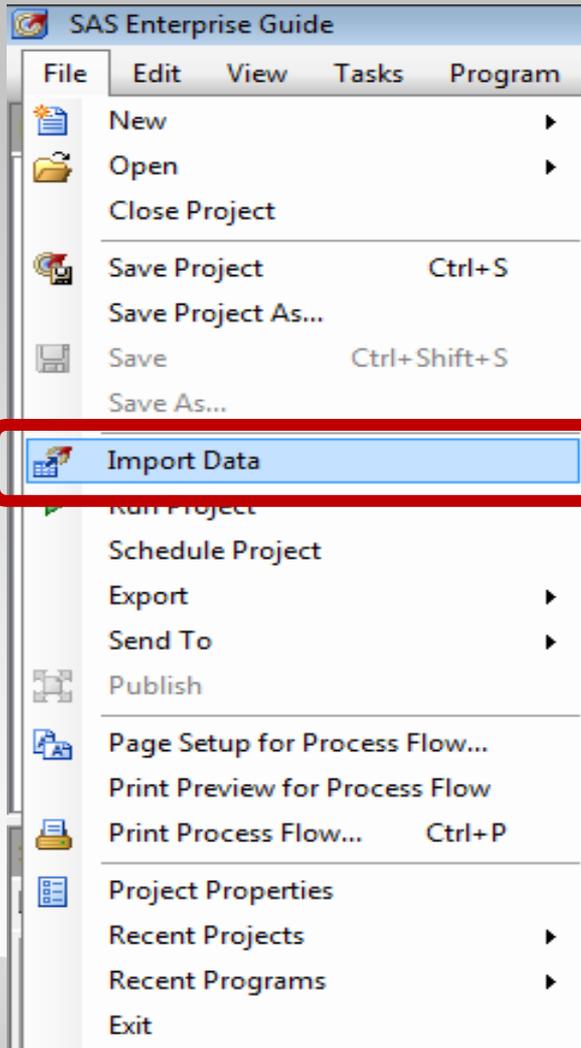
	 F1	 F2	 F3	 F4	 F5	 F6
1	Brave Heart	177	Action Adventure	1995	Paramount Pictu...	R
2	Casablanca	103	Drama	1942	MGM / UA	PG
3	Christmas Vacati...	97	Comedy	1989	Warner Brothers	PG-13
4	Coming to Ameri...	116	Comedy	1988	Paramount Pictu...	R
5	Dracula	130	Horror	1993	Columbia TriStar	R
6	Dressed to Kill	105	Drama Mysteries	1980	Filmways Pictures	R
7	Forrest Gump	142	Drama	1994	Paramount Pictu...	PG-13
8	Ghost	127	Drama Romance	1990	Paramount Pictu...	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 & R...	1987	Warner Brothers	R
12	Michael	106	Drama	1997	Warner Brothers	PG-13
13	National Lampoo...	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 & R...	1983	Universal Studios	R
17	Silence of the La...	118	Drama Suspense	1991	Orion	R
18	Star Wars	124	Action Sci-Fi	1977	Lucas Film Ltd	PG
19	The Hunt for Red...	135	Action Adventure	1989	Paramount Pictu...	PG
20	The Terminator	108	Action Sci-Fi	1984	Live Entertainme...	R
21	The Wizard of Oz	101	Adventure	1939	MGM / UA	G
22	Titanic	194	Drama Romance	1997	Paramount Pictu...	PG-13

# Exercise #6 – Import MS-Excel

1. Click the “File” tab on the main menu bar and select “Import Data” from the drop-down list.
2. Select “Microsoft Excel Workbooks” file type.
3. Select the “Movies” Excel file from the list.
4. Select the desired options from the Import Data Wizard.
5. Let’s discuss the exercise and the resulting dataset.

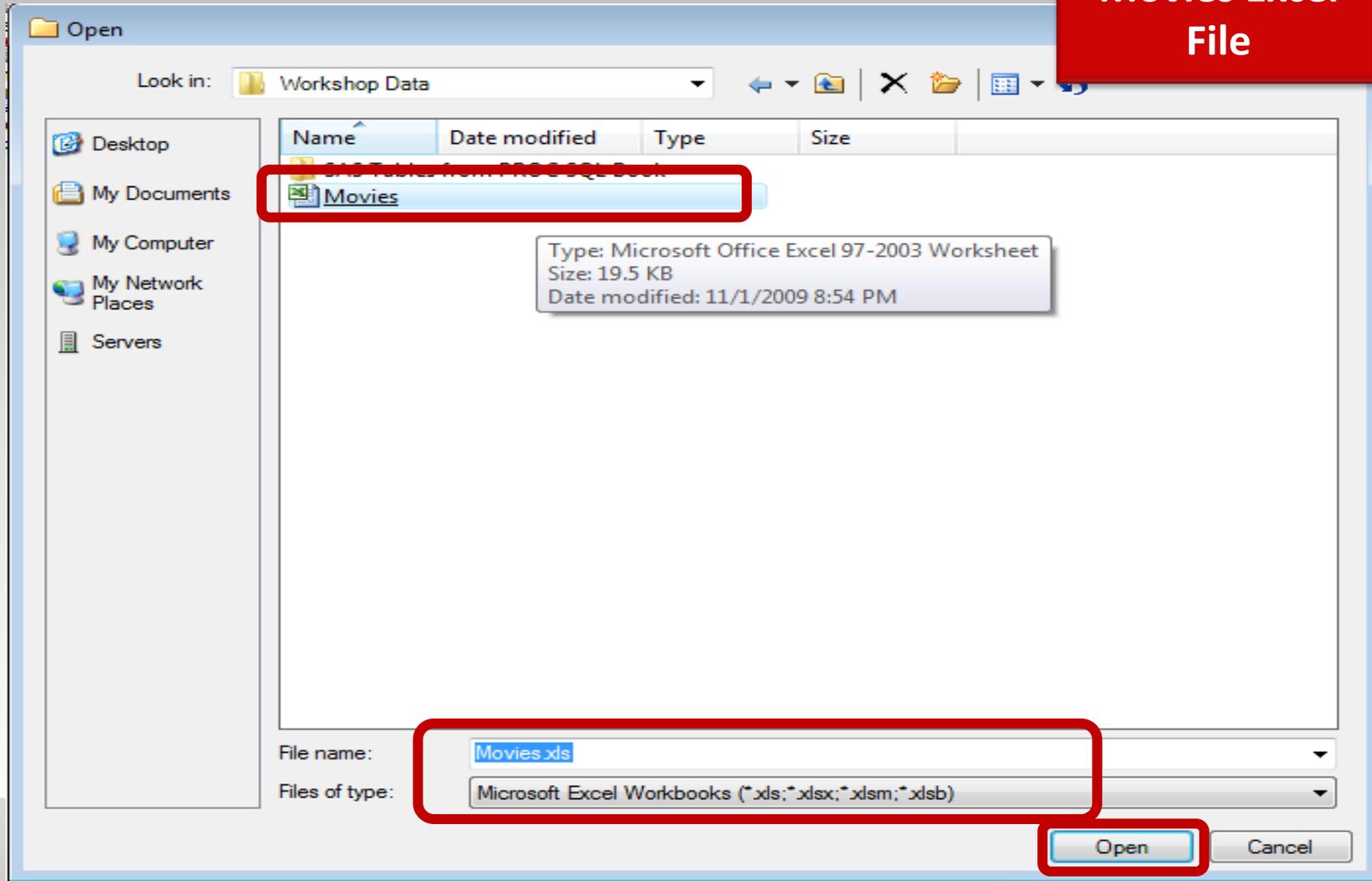


# Exercise #6 (continued)



# Exercise #6 (continued)

Movies Excel  
File



# Exercise #6 (continued)

Specify Data  
to Import

Import Data from Movies.xls

1 of 4 Specify Data 

The Import Data task is used to convert non-SAS data into a SAS data file which is required by other tasks for data analysis and reporting.

Source data file

Location: Local File System

File path: E:\Workshops\Workshop Data\Movies.xls

Data type: Excel workbook

Output SAS data set

SAS server: Local

Library: WORK

Data set: Movies1

<Back



# Exercise #6 (continued)

Select Data Source

Import Data from Movies.xls

2 of 4 Select Data Source

**Select range**

Use a worksheet

Sheet 1

First row of range contains field names

Rename columns to comply with SAS naming conventions.

Use a specific range of cells within the worksheet

Top-left cell:

Lower-right cell:

Expand row range as needed

Use a predefined named range



# Exercise #6 (continued)

Column (Field)  
Attribute  
Definitions

Import Data from Movies.xls

3 of 4 Define Field Attributes

Select columns and define attributes:

Inc	Source Name	Name	Label	Type	Source Informat	Len.	Output Format	Output Informat
<input checked="" type="checkbox"/>	Title	Title	Title	String	\$CHAR27.	27	\$CHAR27.	\$CHAR27.
<input checked="" type="checkbox"/>	Length	Length	Length	Number	BEST12.	8	BEST12.	BEST12.
<input checked="" type="checkbox"/>	Category	Category	Category	String	\$CHAR20.	20	\$CHAR20.	\$CHAR20.
<input checked="" type="checkbox"/>	Year	Year	Year	Number	BEST12.	8	BEST12.	BEST12.
<input checked="" type="checkbox"/>	Studio	Studio	Studio	String	\$CHAR18.	18	\$CHAR18.	\$CHAR18.
<input checked="" type="checkbox"/>	Rating	Rating	Rating	String	\$CHAR5.	5	\$CHAR5.	\$CHAR5.

Select All Clear All Modify...

<Back Next> Finish Cancel Help



# Exercise #6 (continued)

Advanced  
Options

Import Data from movies (with tabs).txt

4 of 4 Advanced Options

Embed the data within the generated SAS code.

Import the data using SAS/ACCESS Interface to PC Files whenever possible.

Remove characters that can cause transmission errors from text-based data files.

Generalize import step to run outside SAS Enterprise Guide.

Maximum record length for input text file in bytes:

<Back ▾ Next> **Finish** Cancel Help



# Exercise #6 (continued)

SAS data set  
after  
importation

	 F1	 F2	 F3	 F4	 F5	 F6
1	Brave Heart	177	Action Adventure	1995	Paramount Pictu...	R
2	Casablanca	103	Drama	1942	MGM / UA	PG
3	Christmas Vacati...	97	Comedy	1989	Warner Brothers	PG-13
4	Coming to Ameri...	116	Comedy	1988	Paramount Pictu...	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 Pictu...	PG-13
8	Ghost	127	Drama Romance	1990	Paramount Pictu...	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 & R...	1987	Warner Brothers	R
12	Michael	106	Drama	1997	Warner Brothers	PG-13
13	National Lampoo...	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 & R...	1983	Universal Studios	R
17	Silence of the La...	118	Drama Suspense	1991	Orion	R
18	Star Wars	124	Action Sci-Fi	1977	Lucas Film Ltd	PG
19	The Hunt for Red...	135	Action Adventure	1989	Paramount Pictu...	PG
20	The Terminator	108	Action Sci-Fi	1984	Live Entertainme...	R
21	The Wizard of Oz	101	Adventure	1939	MGM / UA	G
22	Titanic	194	Drama Romance	1997	Paramount Pictu...	PG-13

# Summarizing Statistics and Data



# Summarizing Statistics and Data

- Little or no programming required
- Supported features:
  - ✓ Producing “Quick and dirty” summary statistics
  - ✓ Recoding data values
  - ✓ Rearranging data order
  - ✓ Merging (or joining) tables of data
  - ✓ Transposing data
  - ✓ Concatenating data
  - ✓ Comparing data

# Exercise #7 – Characterize Data

1. Select the “Movies” dataset from the Process Flow work surface.
2. Click the “Tasks” tab on the main menu bar and select “Describe” from the drop-down list.
3. Select “Characterize Data” from the Describe drop-down list.
4. Select the desired options shown by the Characterize Data Wizard.
5. Let’s discuss the exercise and the resulting dataset.



# Exercise #7 (continued)

The image shows a software menu with three tabs: 'Tasks', 'Program', and 'Tools'. The 'Tasks' tab is active, displaying a list of analysis tasks. The 'Describe' task is highlighted with a red box. A sub-menu is open for 'Describe', listing various options. The 'Characterize Data...' option in this sub-menu is also highlighted with a red box. Other options in the sub-menu include 'List Data...', 'Summary Statistics Wizard...', 'Summary Statistics...', 'Summary Tables Wizard...', 'Summary Tables...', 'List Report Wizard...', 'Distribution Analysis...', 'One-Way Frequencies...', and 'Table Analysis...'. The 'Data' task is also visible at the top of the 'Tasks' list.

Task	Sub-menu Item
Data	List Data...
Describe	Characterize Data...
Describe	Summary Statistics Wizard...
Describe	Summary Statistics...
Describe	Summary Tables Wizard...
Describe	Summary Tables...
Describe	List Report Wizard...
Describe	Distribution Analysis...
Describe	One-Way Frequencies...
Describe	Table Analysis...
Graph	
ANOVA	
Regression	
Multivariate	
Survival Analysis	
Capability	
Control Charts	
Pareto Chart...	
Time Series	
Data Mining	
OLAP	
Task Templates	



# Exercise #7 (continued)

Select the  
Data for  
Reporting

Characterize Data for E:\Workshops\Workshop Data\movies.sas7bdat

1 of 3 Select the data you want to report on.

The Characterize Data Wizard helps you create a summary report, graphs, and frequency and univariate SAS data sets that describe the main characteristics of the data.

SAS server: Local

Data:

Add...  
Remove

<Back **Next>** Finish Cancel Help



# Exercise #7 (continued)

Select Report Options

Characterize Data for E:\Workshops\Workshop Data\movies.sas7bd

2 of 3 Select the report options.

Output Types

- Summary Report
- Graphs
- SAS Data Sets

Frequency Data:

SAS server: Local  
Library: WORK  
Data set: FREQCHARFREQUENCYFORMOVIES\_0000

Univariate Data:

SAS server: Local  
Library: WORK  
Data set: UNIVCHARUNIVARIATEFORMOVIES\_0000



# Exercise #7 (continued)

**Specify  
Output Limits**

Characterize Data for E:\Workshops\Workshop Data\movies.sas7bdat

3 of 3 Specify output limits.

Limit categorical values

Maximum number of unique categorical values to be reported per variable:

30

<Back Next> **Finish** Cancel Help



# Exercise #7 (continued)

Summary  
Output and  
Statistics  
Listing

Summary of Categorical Variables for ECLIB000.MOVIES  
Limited to the 30 Most Frequent Distinct Values per Variable

Variable	Label	Value	Frequency Count	Percent of Total Frequency
Category		Action Adventure	4	18.1818
		Comedy	3	13.6364
		Drama	3	13.6364
		Action Cops & Robber	2	9.0909
		Action Sci-Fi	2	9.0909
		Drama Romance	2	9.0909
		Horror	2	9.0909
		Action	1	4.5455
		Adventure	1	4.5455
		Drama Mysteries	1	4.5455
		Drama Suspense	1	4.5455

Variable	Label	Value	Frequency Count	Percent of Total Frequency
Rating		R	8	36.3636
		PG-13	7	31.8182
		PG	6	27.2727
		G	1	4.5455

Variable	Label	Value	Frequency Count	Percent of Total Frequency
Studio		Paramount Pictures	6	27.2727
		MGM / UA	4	18.1818
		Warner Brothers	4	18.1818
		Universal Studios	2	9.0909
		Columbia TriStar	1	4.5455
		Filmways Pictures	1	4.5455
		Live Entertainment	1	4.5455
		Lucas Film Ltd	1	4.5455
		Orion	1	4.5455
		Universal Pictures	1	4.5455



# Exercise #7 (continued)

## Summary Output and Statistics Listing

Variable	Label	Value	Frequency Count	Percent of Total Frequency
Title		Brave Heart	1	4.5455
		Casablanca	1	4.5455
		Christmas Vacation	1	4.5455
		Coming to America	1	4.5455
		Dracula	1	4.5455
		Dressed to Kill	1	4.5455
		Forrest Gump	1	4.5455
		Ghost	1	4.5455
		Jaws	1	4.5455
		Jurassic Park	1	4.5455
		Lethal Weapon	1	4.5455
		Michael	1	4.5455
		National Lampoon's Vacation	1	4.5455
		Poltergeist	1	4.5455
		Rocky	1	4.5455
		Scarface	1	4.5455
		Silence of the Lambs	1	4.5455
		Star Wars	1	4.5455
		The Hunt for Red October	1	4.5455
		The Terminator	1	4.5455
		The Wizard of Oz	1	4.5455
		Titanic	1	4.5455

Generated by the SAS System ('Local', W32\_VSPRO) on October 12, 2011 at 3:47:37 AM

Page Break

### Summary of Numeric Variables for ECLIB000.MOVIES

Variable	Label	N	NMiss	Total	Min	Mean	Median	Max	StdMean
Length		22	0	2748	97	124.91	119.0	194	5.50793

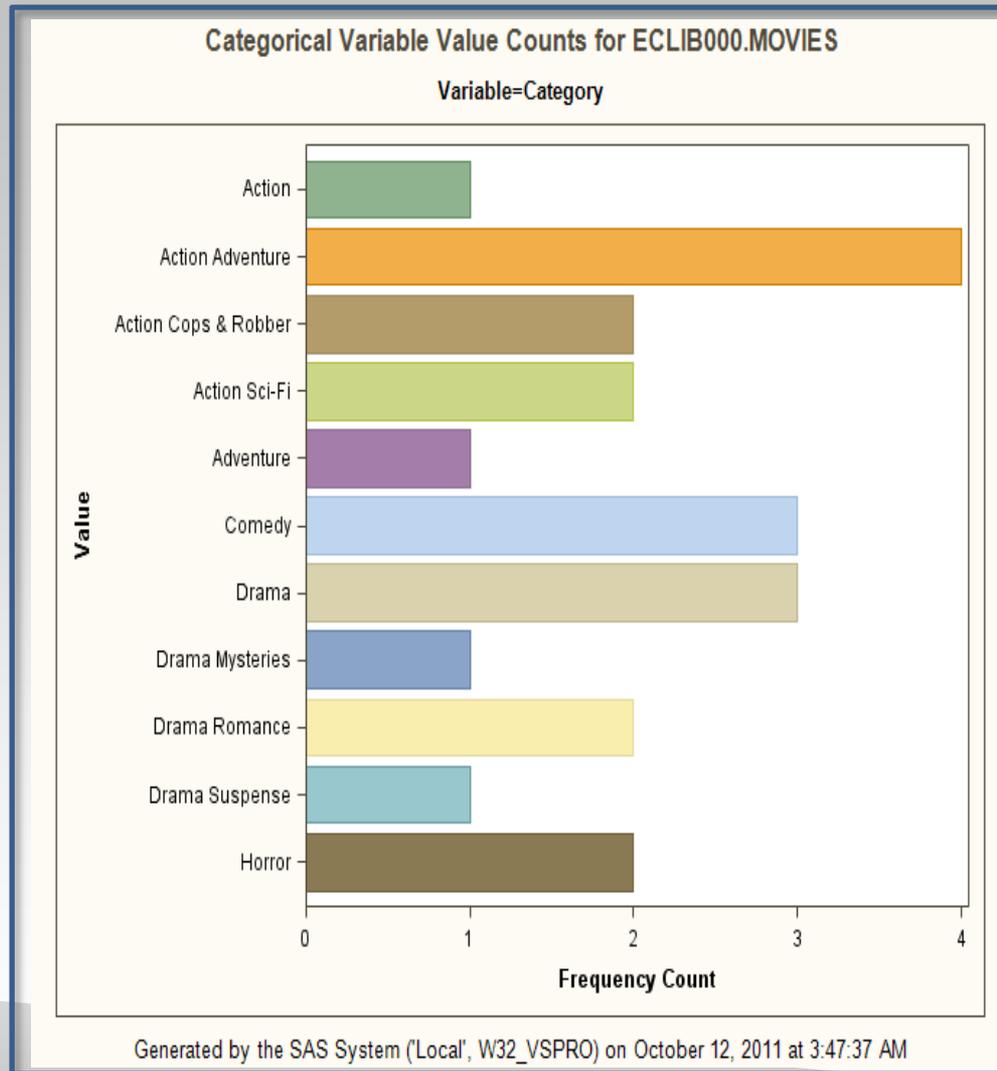
Variable	Label	N	NMiss	Total	Min	Mean	Median	Max	StdMean
Year		22	0	43624	1939	1982.91	1987.5	1997	3.24331

Generated by the SAS System ('Local', W32\_VSPRO) on October 12, 2011 at 3:47:37 AM



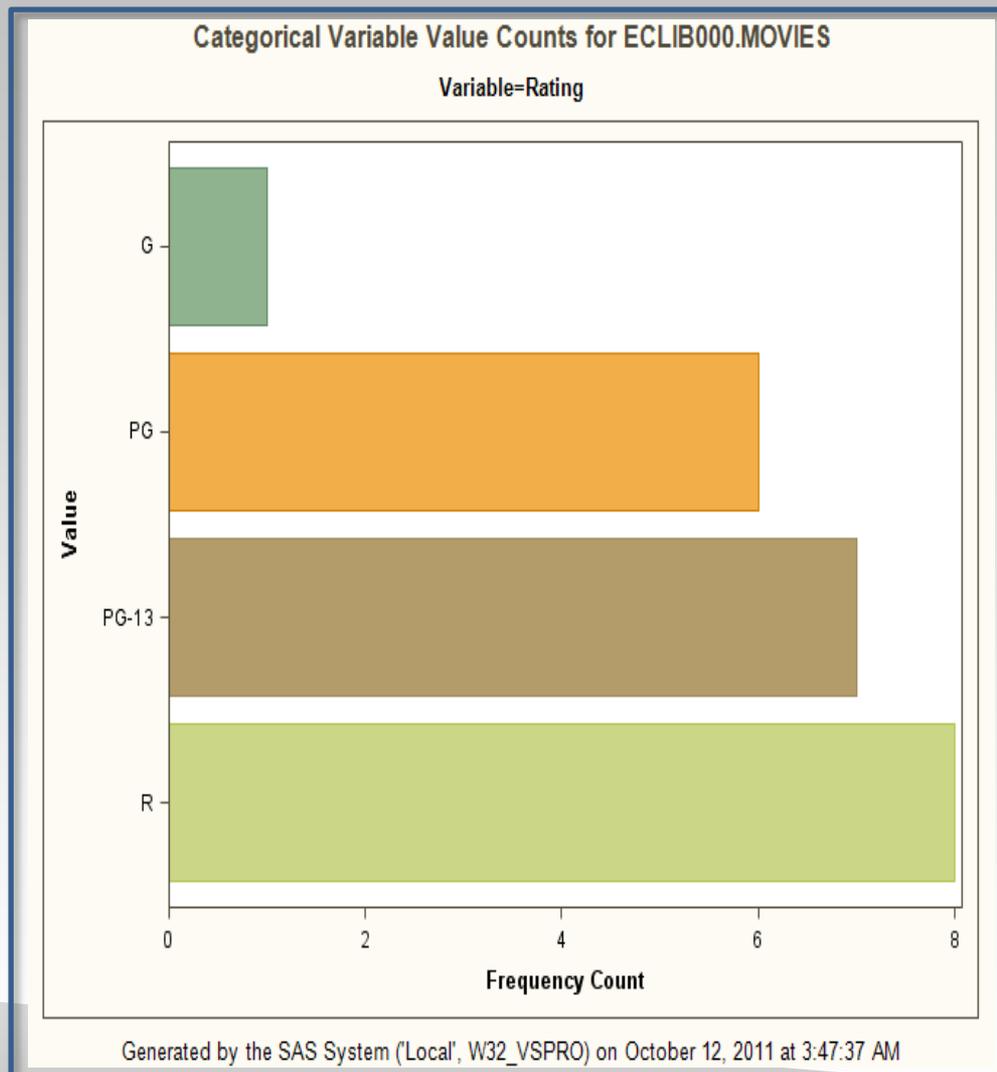
# Exercise #7 (continued)

Summary  
Output and  
Statistics  
Listing



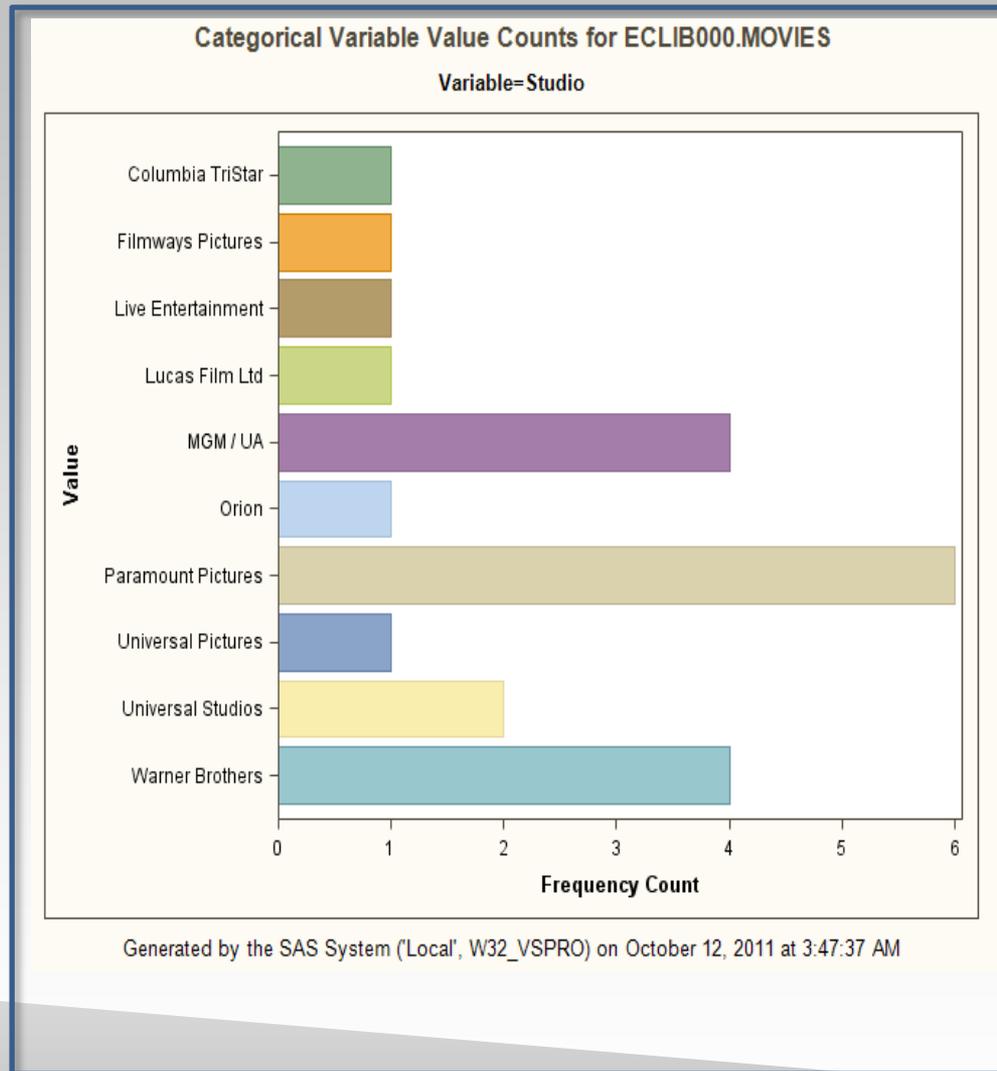
# Exercise #7 (continued)

Summary  
Output and  
Statistics  
Listing



# Exercise #7 (continued)

Summary  
Output and  
Statistics  
Listing



# **Manipulating Data Structures and Data**



# What is a Match-Merge (Equi-Join)?

- Process of combining tables side-by-side
- Performs a matching process between rows in tables
- Some or all of the tables' contents are brought together
- Gather and manipulate data from across tables

Visually, it looks something like this:



# Why Perform an Equi-Join?

- Data is often stored in separate tables
- Joins allow data to be combined as if it were stored in one huge file
- Provide exciting insights into data relationships
- Types of joins:
  - ✓ Inner joins – a maximum of 256 tables can be joined
  - ✓ Outer joins – a maximum of 2 tables can be joined

# What Happens During an Equi-Join?

**When joining two tables:**

- An intermediate Cartesian product is constructed from the two tables
- Rows are selected that match the WHERE clause

**When joining more than two table:**

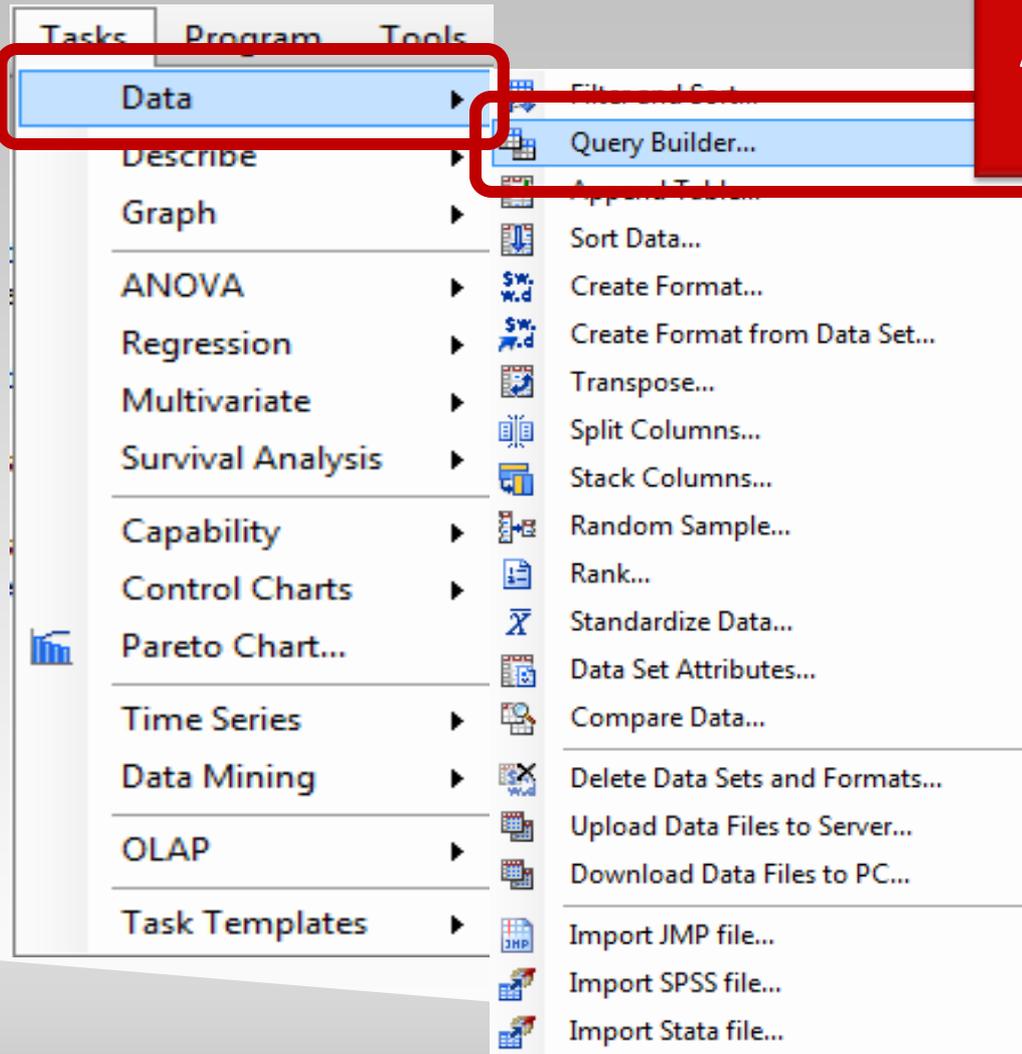
- Determines the order of processing to reduce the size of the intermediate Cartesian product
- Reconstructs the join into several two-way joins
- Unwanted rows and columns from the intermediate tables are removed

# Exercise #8 – Equi-Join

1. Click the “Tasks” tab on the main menu bar and select “Data” from the drop-down list.
2. Select “Query Builder” from the Data drop-down list.
3. Select the “Movies” and “Actors” datasets from the list.
4. Select the desired options displayed on the Query Builder Wizard.
5. Let’s discuss the exercise and the resulting merged dataset.



# Exercise #8 (continued)



**Access Query  
Builder**



# Exercise #8 (continued)

Select  
MOVIES Table  
for Joining

The screenshot shows the SAS Query Builder interface. The window title is "Query Builder for E:\Workshops\Workshop Data\movies.sas7bdat". The "Query name" is "Query Builder" and the "Output name" is "WORK.QUERY\_FOR\_MOVIES\_SAS7E". The interface includes tabs for "Computed Columns", "Prompt Manager", "Preview", "Tools", and "Options". The "Add Tables" pane on the left is highlighted with a red box and contains a tree view for "t1 (movies)" with the following fields: Title, Length, Category, Year, Studio, and Rating. The "Select Data" pane on the right is empty, with a table header showing "Column Name", "Identifier", and "Summary". A red arrow at the bottom of the slide points to the right.

Column Name	Identifier	Summary
Drop a column here to add it to the query.		

Select distinct rows only

Run Save and Close Cancel Help

# Exercise #8 (continued)

Click  
"Add Tables"  
to Add  
2<sup>nd</sup> Table

The screenshot shows the SAS Query Builder window for a query named 'Query Builder'. The interface includes a menu bar with 'Computed Columns', 'Prompt Manager', 'Preview', 'Tools', and 'Options'. Below the menu bar, there are three buttons: 'Add Tables' (highlighted with a red box), 'Delete', and 'Join Tables'. The main area is divided into two panes. The left pane shows a tree view of a table named 't1 (movies)' with columns: Title, Length, Category, Year, Studio, and Rating. The right pane is titled 'Select Data' and contains a table with columns 'Column Name', 'Identifier', and 'Summary'. The table is currently empty, with the text 'Drop a column here to add it to the query.' displayed. Below the table is a checkbox labeled 'Select distinct rows only'. At the bottom of the window are buttons for 'Run', 'Save and Close', 'Cancel', and 'Help'.



# Exercise #8 (continued)

Select  
**ACTORS** Table  
for Joining

Query Builder for E:\Workshops\Workshop Data\movies.sas7bdat

Query name: Query Builder      Output name: WORK.QUERY\_FOR\_UNIVCHARUNI

Computed Columns    Prompt Manager    Preview    Tools    Options

Add Tables    Delete    Join Tables

t1 (movies)

- Title
- Length
- Category
- Year
- Studio
- Rating

t2 (actors)

- Title
- Actor\_Leading
- Actor\_Supporting

Select Data    Filter Data    Sort Data

Column Name	Identifier	Summary
Drop a column here to add it to the query.		

Select distinct rows only

Run    Save and Close    Cancel    Help



# Exercise #8 (continued)

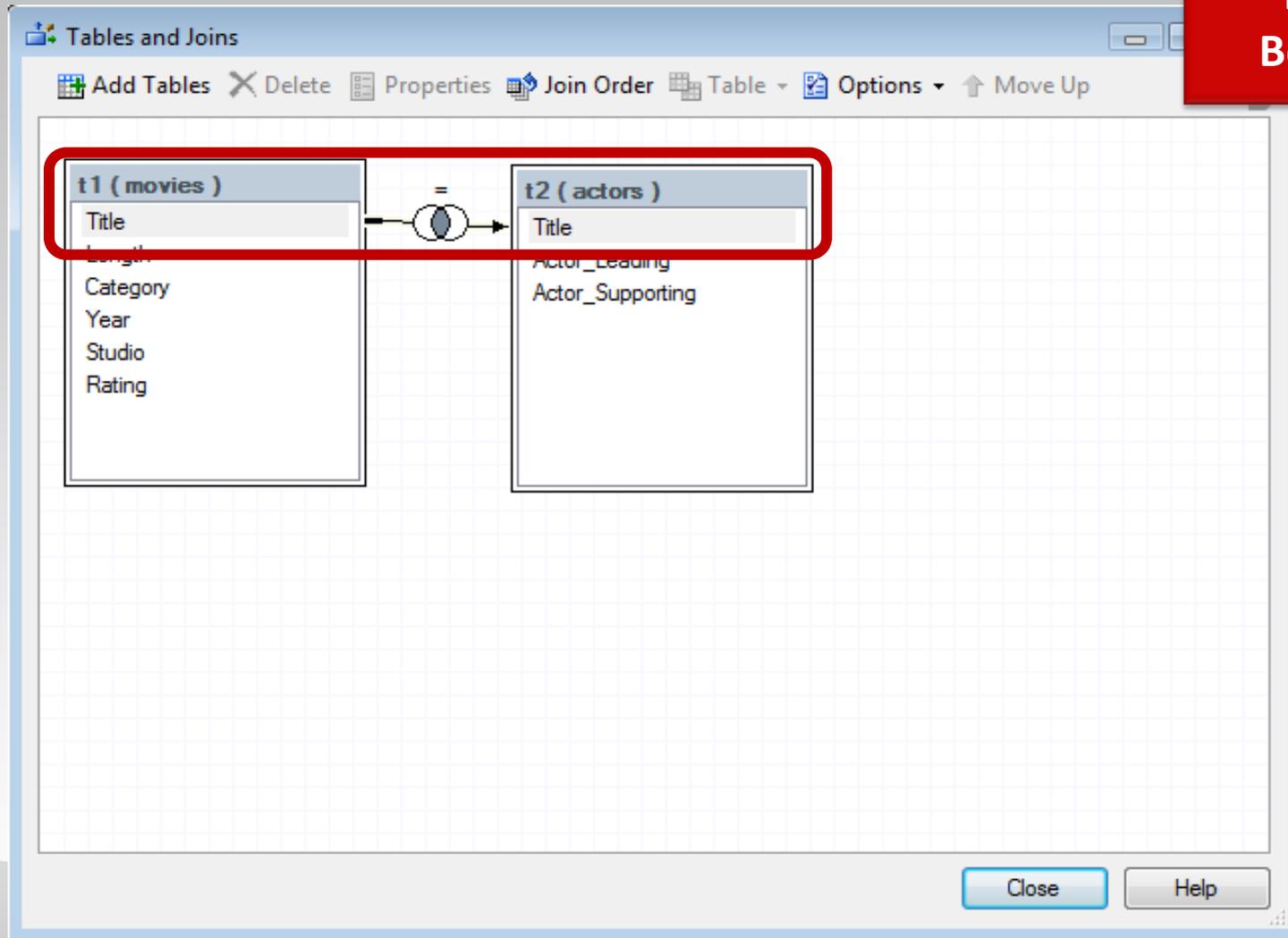
Click  
"Join Tables" to  
Describe Type  
of Join Desired

The screenshot shows the SAS Query Builder interface. The window title is "Query Builder for E:\Workshops\Workshop Data\movies.sas7bdat". The "Query name" is "Query Builder" and the "Output name" is "WORK.QUERY\_FOR\_UNIVCHARUNI". The "Join Tables" button is highlighted with a red box. The left pane shows a tree view of tables: "t1 (movies)" with columns "Title", "Length", "Category", "Year", "Studio", and "Rating"; and "t2 (actors)" with columns "Title", "Actor\_Leading", and "Actor\_Supporting". The right pane is titled "Select Data" and contains a table with columns "Column Name", "Identifier", and "Summary". The table is empty, with the text "Drop a column here to add it to the query." in the center. Below the table is a checkbox labeled "Select distinct rows only". At the bottom of the window are buttons for "Run", "Save and Close", "Cancel", and "Help".



# Exercise #8 (continued)

Select Primary  
"Key" from  
Both Tables



# Exercise #8 (continued)

Drag one or more Columns to Add it to Query

Query Builder for E:\Workshops\Workshop Data\movies.sas7bdat

Query name: Query Builder      Output name: WORK.QUERY\_FOR\_UNIVCHARUN1

Computed Columns    Prompt Manager    Preview    Tools    Options

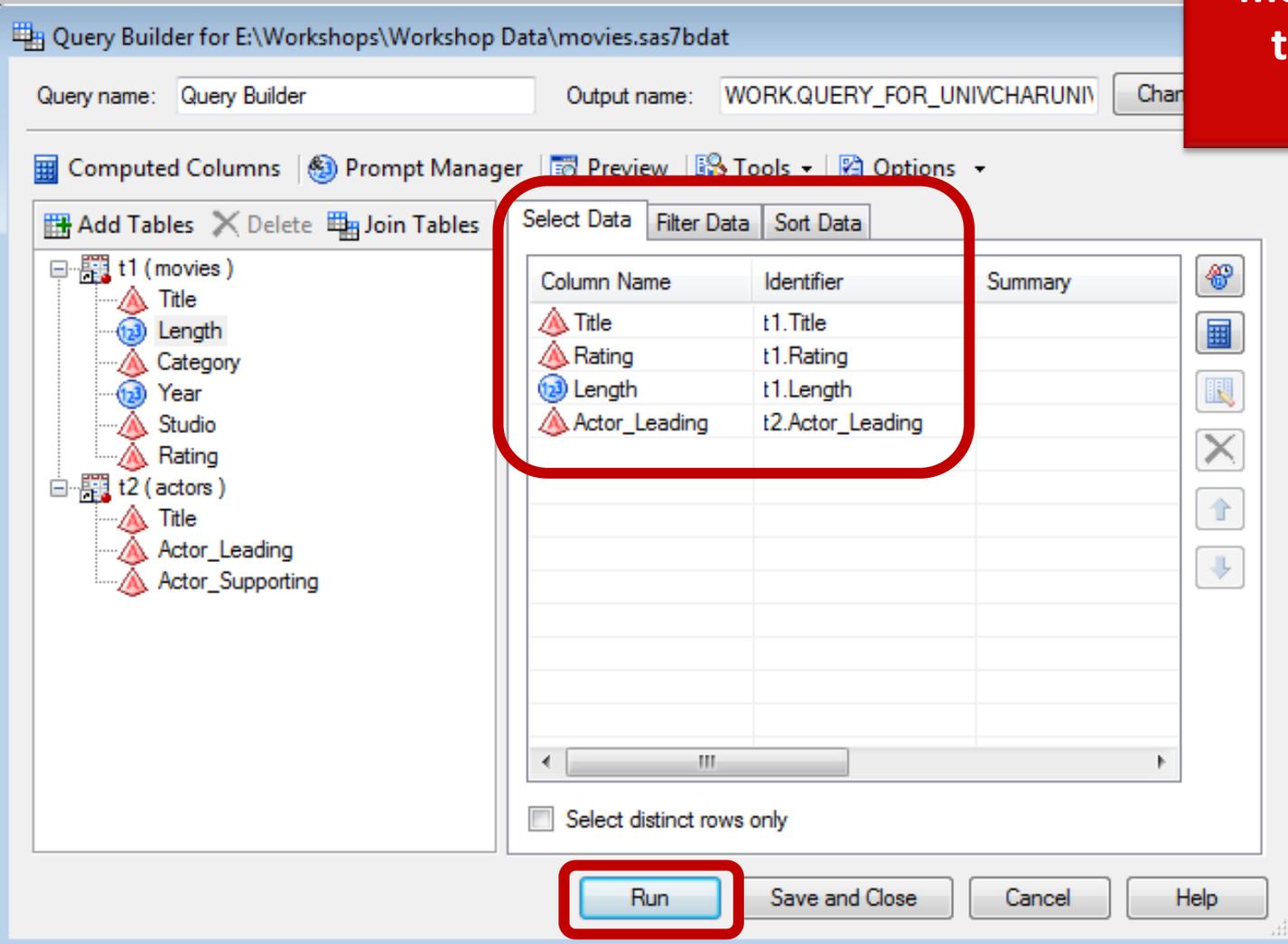
Add Tables    Delete    Join Tables

Select Data    Filter Data    Sort Data

Column Name	Identifier	Summary
Title	t1.Title	
Rating	t1.Rating	
Length	t1.Length	
Actor_Leading	t2.Actor_Leading	

Select distinct rows only

Run    Save and Close    Cancel    Help



# Exercise #8 (continued)

	 Title	 Rating	 Length	 Actor_Leading
1	Brave Heart	R	177	Mel Gibson
2	Christmas Vacati...	PG-13	97	Chevy Chase
3	Coming to Ameri...	R	116	Eddie Murphy
4	Forrest Gump	PG-13	142	Tom Hanks
5	Ghost	PG-13	127	Patrick Swayze
6	Lethal Weapon	R	110	Mel Gibson
7	Michael	PG-13	106	John Travolta
8	National Lampoo...	PG-13	98	Chevy Chase
9	Rocky	PG	120	Sylvester Stallone
10	Silence of the La...	R	118	Anthony Hopkins
11	The Hunt for Red...	PG	135	Sean Connery
12	The Terminator	R	108	Arnold Schwarze...
13	Titanic	PG-13	194	Leonardo DiCaprio

Output from  
Equi-Join

# Managing Projects and Flowcharts



# Managing Projects and Flowcharts

- Save and recall projects and all its contents
- Visually organize, view and manage projects
- EG-generated flow diagrams and source code
- Clickable hyperlinks
- Support documentation requirements.

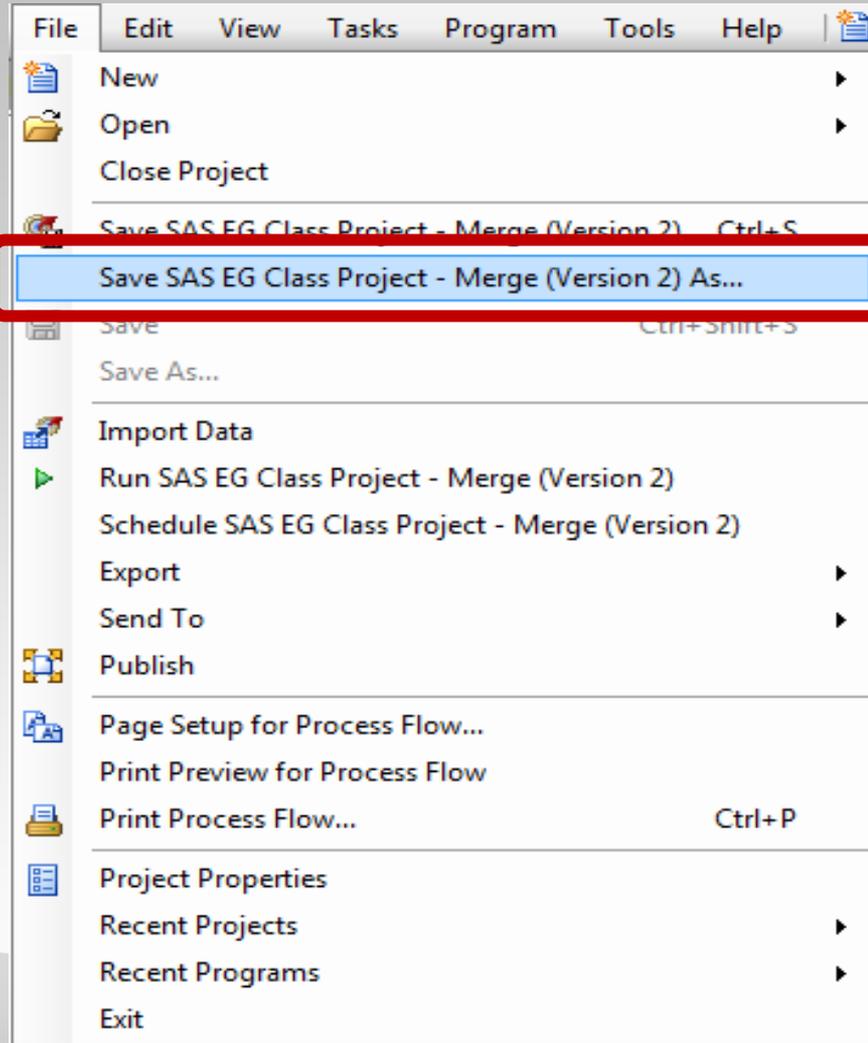
# Exercise #9 – Saving a Project

1. Click the “Save” tab on the main menu bar and select “Save As” from the drop-down list.
2. Enter the location and name for the project file.
3. Click the “Save” button.
4. Let’s discuss the exercise and the resulting saved project.



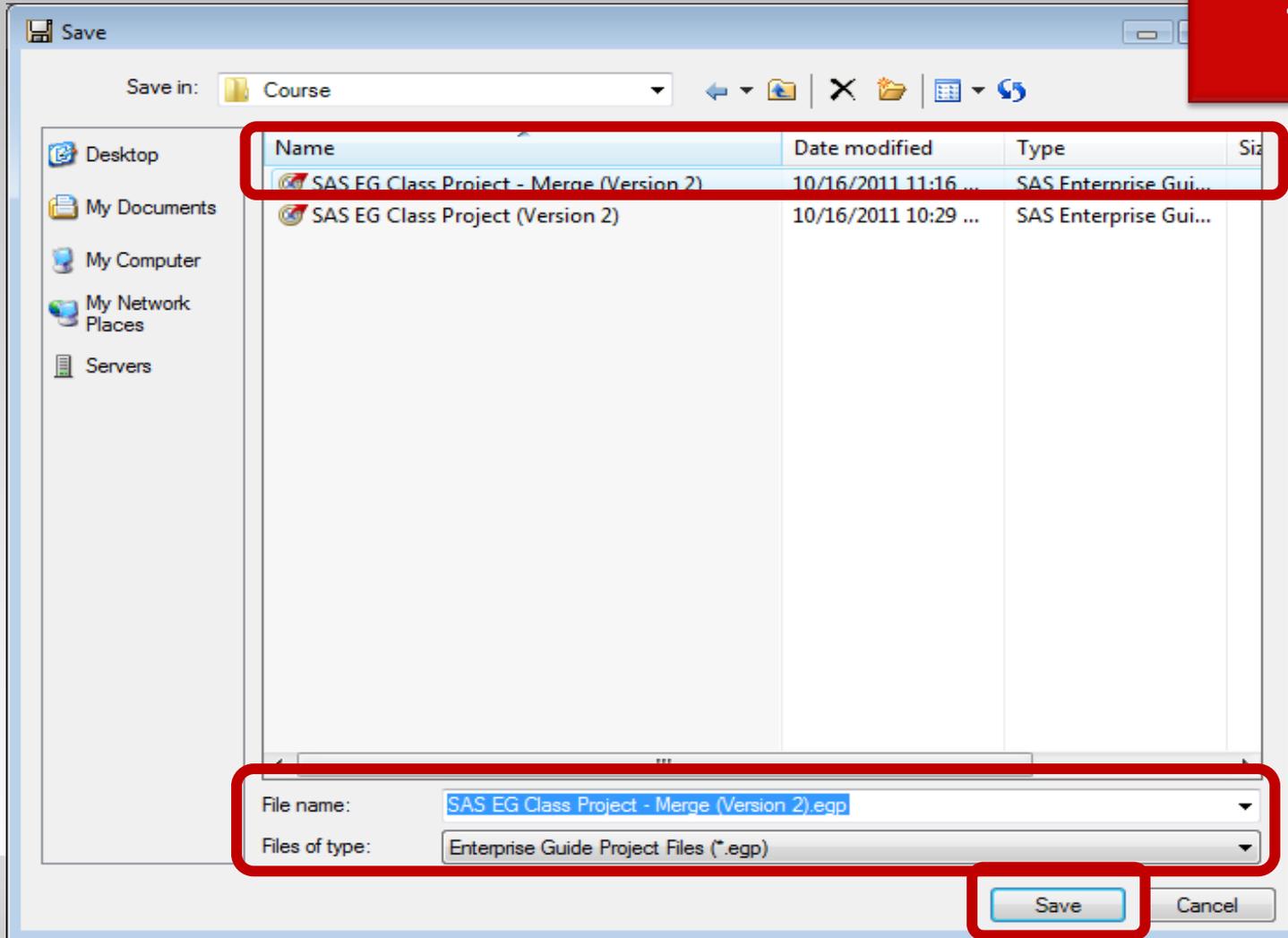
# Exercise #9 (continued)

Save As  
Project

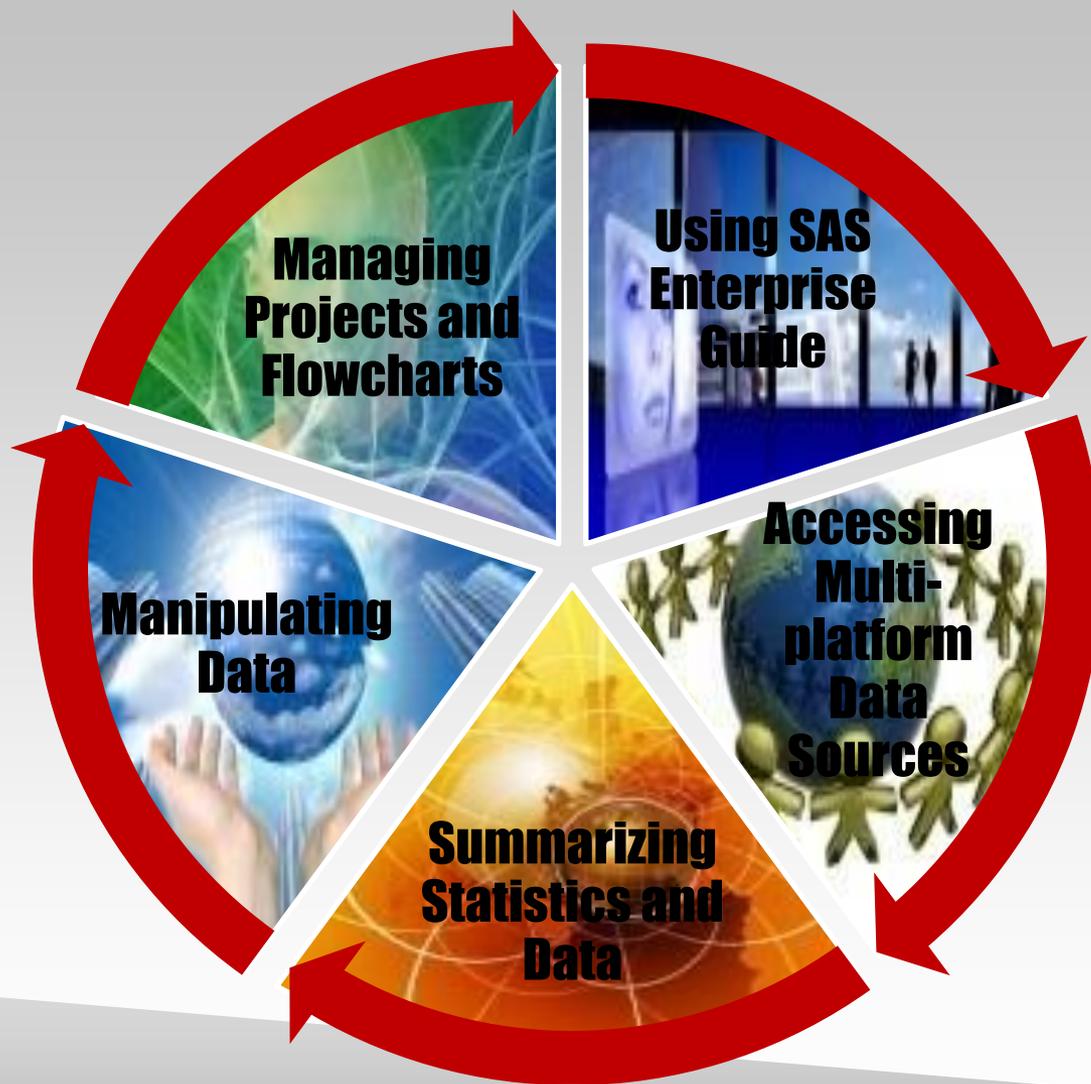


# Exercise #9 (continued)

Save As Project



# Conclusion



**SAS EG – A Powerful Tool for All SAS Users!**



1<sup>st</sup> Edition



**Tips, Tricks and Shortcuts for  
Better Searches and Better Results**

Kirk Paul Laffer  
Charles Edwin Shipp



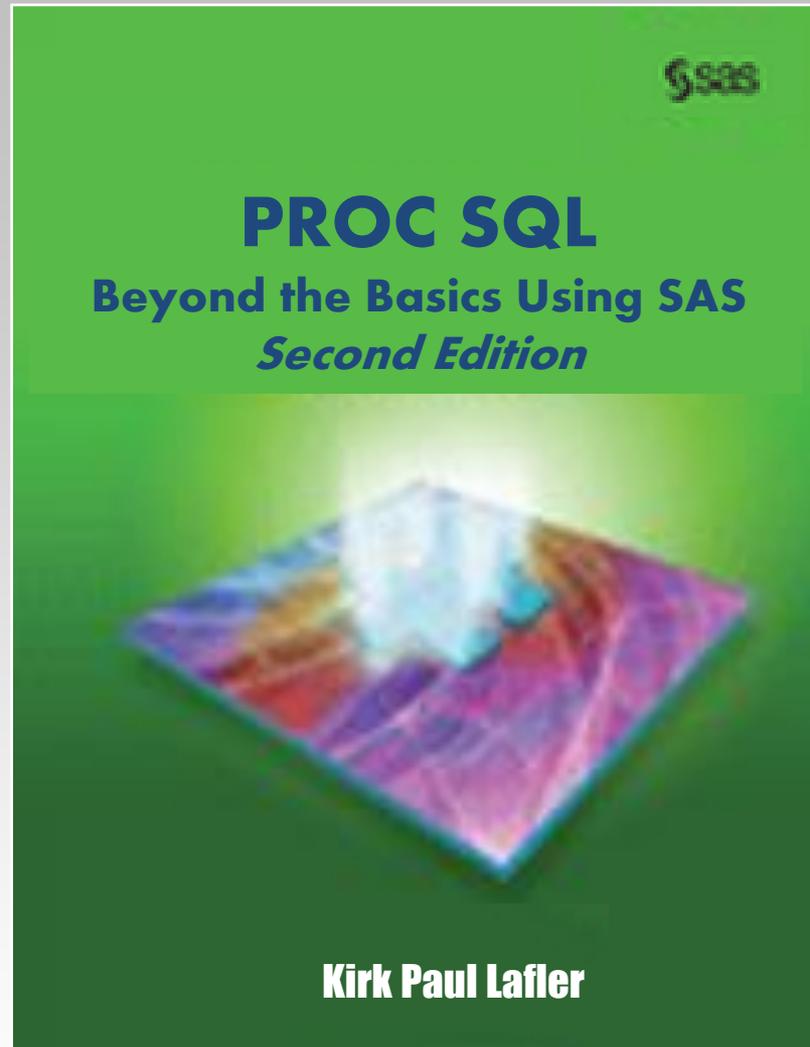
Copyright © 2014 by Kirk Paul Laffer and Charles Edwin Shipp  
All rights reserved

Odyssey  
Press

Published by Odyssey Press  
A Division of Software Intelligence Corporation

**Become a Search Pro!  
Filled with  
Tips and Techniques  
for Better Searches  
and Better Results**

**Available on Amazon.com!**



**An SQL Book  
with “behind the  
scenes” details,  
explanations and  
lots of examples**

**Available from SAS Press at [sas.com](http://sas.com)!**

**Thank you for attending!**

# Questions?

SAS® GLOBAL FORUM 2017

April 2 - 5 | Orlando, FL

*A Hands-on Workshop by*

**Kirk Paul Lafler**

**Mira Shapiro**

**Ryan Paul Lafler**