TECHNOLOGY SUPPORT

# SAS Enterprise Guide - PART I

## SHORTCOURSE HANDOUT

TEXAS TECH UNIVERSITY
Information Technology Division
Technology Support

Texas Tech University | Heide Mansouri

# Table of Contents

# SAS Enterprise Guide 7.1 – Part I

# ShortCourse Handout



**Introduction**

SAS Enterprise Guide (EG) is the most powerful module in SAS. It is a Point-and-click version of SAS.

SAS Enterprise Guide is an easy-to-use Windows client application that provides:

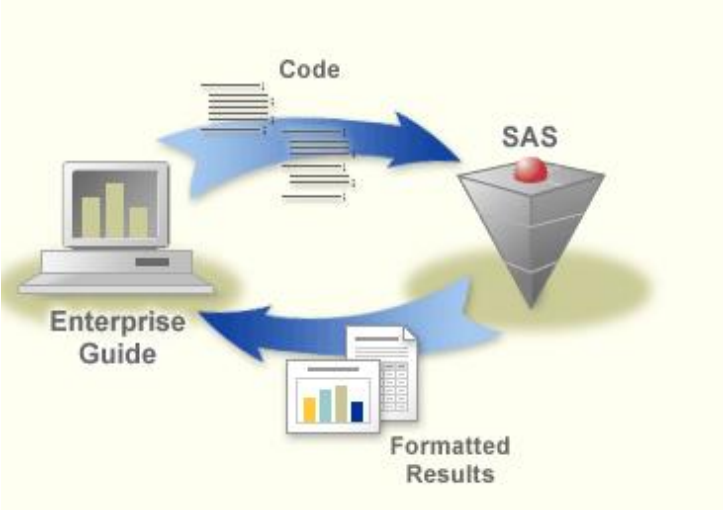
- Access to much of the functionality of SAS;
- Ready-to-use tasks for analysis and reporting;
- Easy ways to export data and results to other applications;
- Scripting and automation; and
- Code editing facility.

SAS Enterprise Guide is designed to be used by novice SAS users as well as researchers and business decision makers who are already familiar with SAS programming for exporting data and results to other applications and creating **reports**, **queries**, and **charts**.

When you use SAS Enterprise Guide, you are also using SAS software behind the scenes. That is SAS Enterprise Guide is not a standalone program; SAS itself needs to be installed, for Enterprise Guide to work.

SAS Enterprise Guide can connect to SAS on your local computer, or it can connect to SAS on another computer called a SAS server. As you access data and build tasks, SAS Enterprise Guide generates SAS code for you. When you run a task, the generated code is sent to SAS for processing and the results are returned to SAS Enterprise Guide.

For this ShortCourse, it is assumed that you have completed the Base SAS ShortCourse or already know how to create SAS data sets. It is also assumed that you are familiar with Elementary Statistics.




**SAS Press and SAS Documentation Example Code and data**

We will be using some of the SAS sample data sets included in ***Statistics Using SAS Enterprise Guide*** book, authored by James B. Davis http://www.sascommunity.org/wiki/Statistics_Using_SAS_Enterprise_Guide . We will also be using **SAS Enterprise Guide 7.1** and **SAS 9.4** under **Windows**.

**Course Objectives**

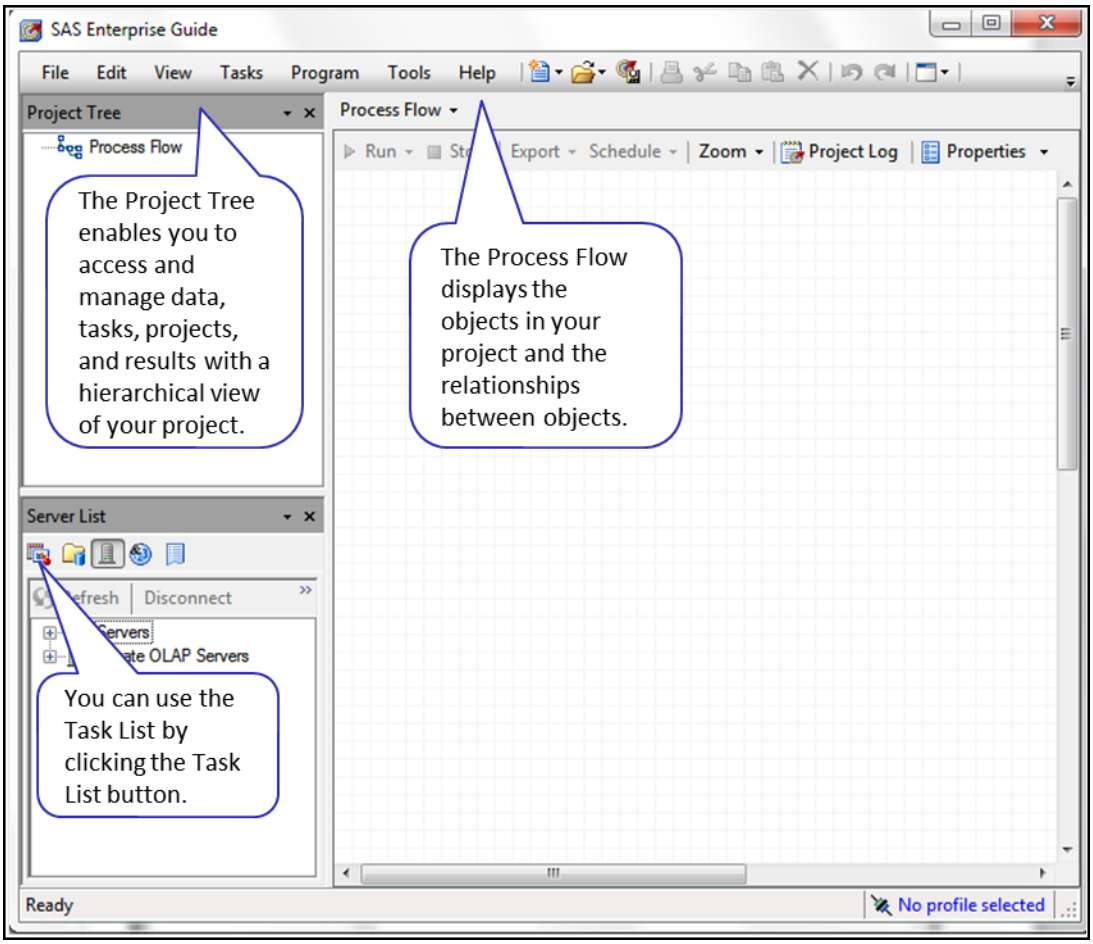
After completing this ShortCourse, you should be able to:

- Create and save projects;
- Create a SAS data sets;
- Create Scatter plots;
- Calculate Summary statistics;
- Import Excel files;
- Create Frequency distribution for discrete data;
- Perform One-Way ANOVA;
- Create Summary tables;
- Perform Two-Way ANOVA with balanced data using Linear models;
- Make inference on a population proportion;
- Calculate Pearson Correlation coefficient; and
- Perform simple linear regression analysis.

**Starting a New Project**

- From the **Start** menu choose **All Programs**
- Select **SAS ->** Click on **Enterprise Guide 7.1**
- In the **Welcome to SAS Enterprise Guide** window, click **New Project**.

  ***Note:*** You can have only one project open at a time in SAS Enterprise Guide.



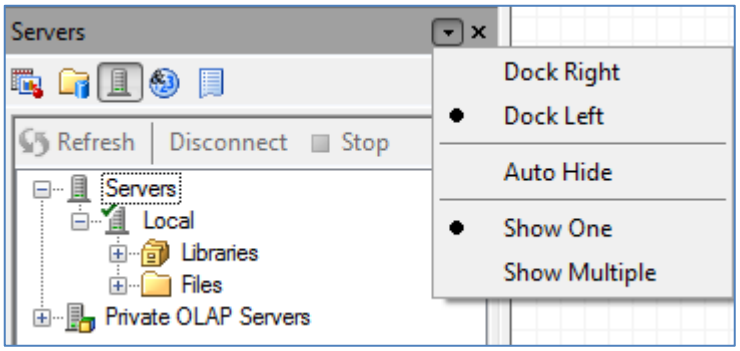

**SAS Enterprise Guide Interface**

- The **Project Tree** - lists the data, tasks, code, and notes in your project.
- The **Workspace** area- displays your **data**, **code**, **logs**, **task results**, and **process flows**.
- The **Servers List** window – lists different tasks that can be opened by clicking on the Task Lists button.
- **Resources pane** - enables you to access the
  - **Task List**,
  - **SAS Folders**,

  - **Server List**, and
  - The **Prompt Manager**.
  - We use Tasks to define what we want.
- **Task Status** - displays the status, queue position, and server of any task that is currently running.
- **Project Log** - displays an aggregated log of the entire project.

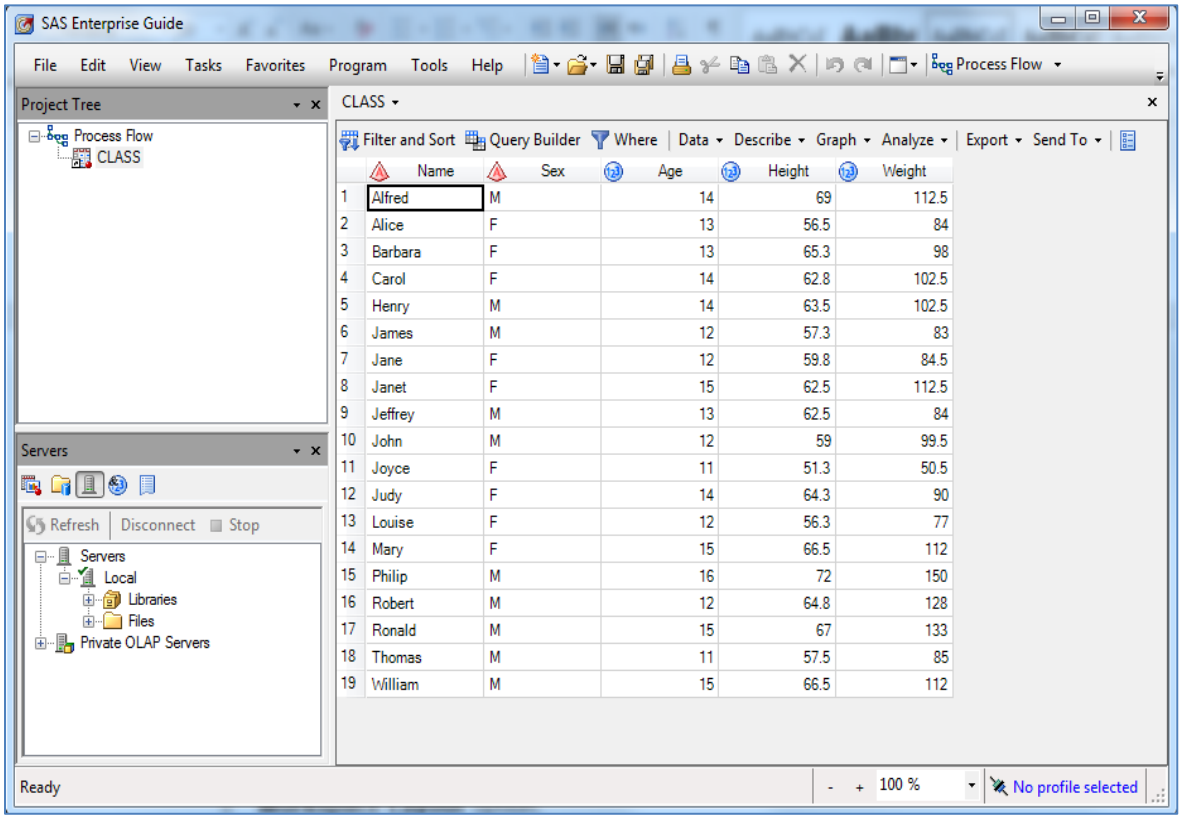
The Resources pane is displayed by default in the lower-left corner of the SAS Enterprise Guide window, and it provides access to the Task List, SAS Folders, the Server List, and the Prompt Manager. By default, the Resources pane displays the Server List.

| | |
|---|---|
| Task List | displays a list of all of the available tasks, either sorted by category or alphabetically. You can also view task templates. Double-click a task name to begin using it. |
| SAS Folders | displays a list of all of your stored processes, information maps, and projects. You can select an item from this list and open it. |
| Server List | displays a list of all the available SAS servers. |
| Prompt Manager | displays a list of all the available prompts. |

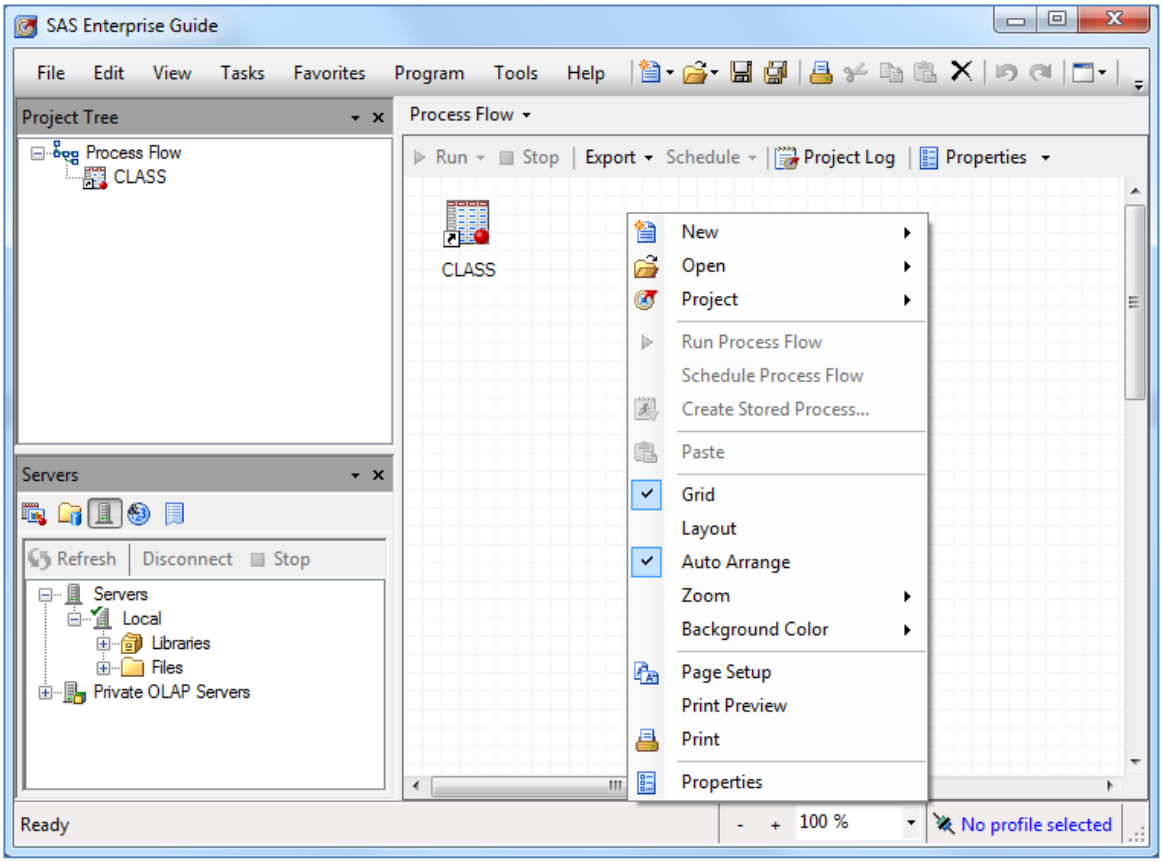


**The workspace**

- The workspace is the main area of the SAS Enterprise Guide application and is used to display your **data**, **code**, **logs**, **task results**, and **process flows**.
- The workspace is always displayed and cannot be minimized.
- There are ***four*** important features in the workspace:
  - **Tabbed access to task** and query items (**when data set is open**);
  - **Context-sensitive toolbar**;
  - **Recently viewed items** button; and
  - **Workspace Layout** option (when Right-clicking on Workspace).




| | Name | Sex | Age | Height | Weight |
|---|---|---|---|---|---|
| 1 | Alfred | M | 14 | 69 | 112.5 |
| 2 | Alice | F | 13 | 56.5 | 84 |
| 3 | Barbara | F | 13 | 65.3 | 98 |
| 4 | Carol | F | 14 | 62.8 | 102.5 |
| 5 | Henry | M | 14 | 63.5 | 102.5 |
| 6 | James | M | 12 | 57.3 | 83 |
| 7 | Jane | F | 12 | 59.8 | 84.5 |
| 8 | Janet | F | 15 | 62.5 | 112.5 |
| 9 | Jeffrey | M | 13 | 62.5 | 84 |
| 10 | John | M | 12 | 59 | 99.5 |
| 11 | Joyce | F | 11 | 51.3 | 50.5 |
| 12 | Judy | F | 14 | 64.3 | 90 |
| 13 | Louise | F | 12 | 56.3 | 77 |
| 14 | Mary | F | 15 | 66.5 | 112 |
| 15 | Philip | M | 16 | 72 | 150 |
| 16 | Robert | M | 12 | 64.8 | 128 |
| 17 | Ronald | M | 15 | 67 | 133 |
| 18 | Thomas | M | 11 | 57.5 | 85 |
| 19 | William | M | 15 | 66.5 | 112 |




**Creating a new project**

You can start a new project either when you first start up SAS Enterprise Guide or when you close the open project and start another one. *You can have only **one project** open at a time in SAS Enterprise Guide.*

**To start a new project when you invoke SAS Enterprise Guide:**

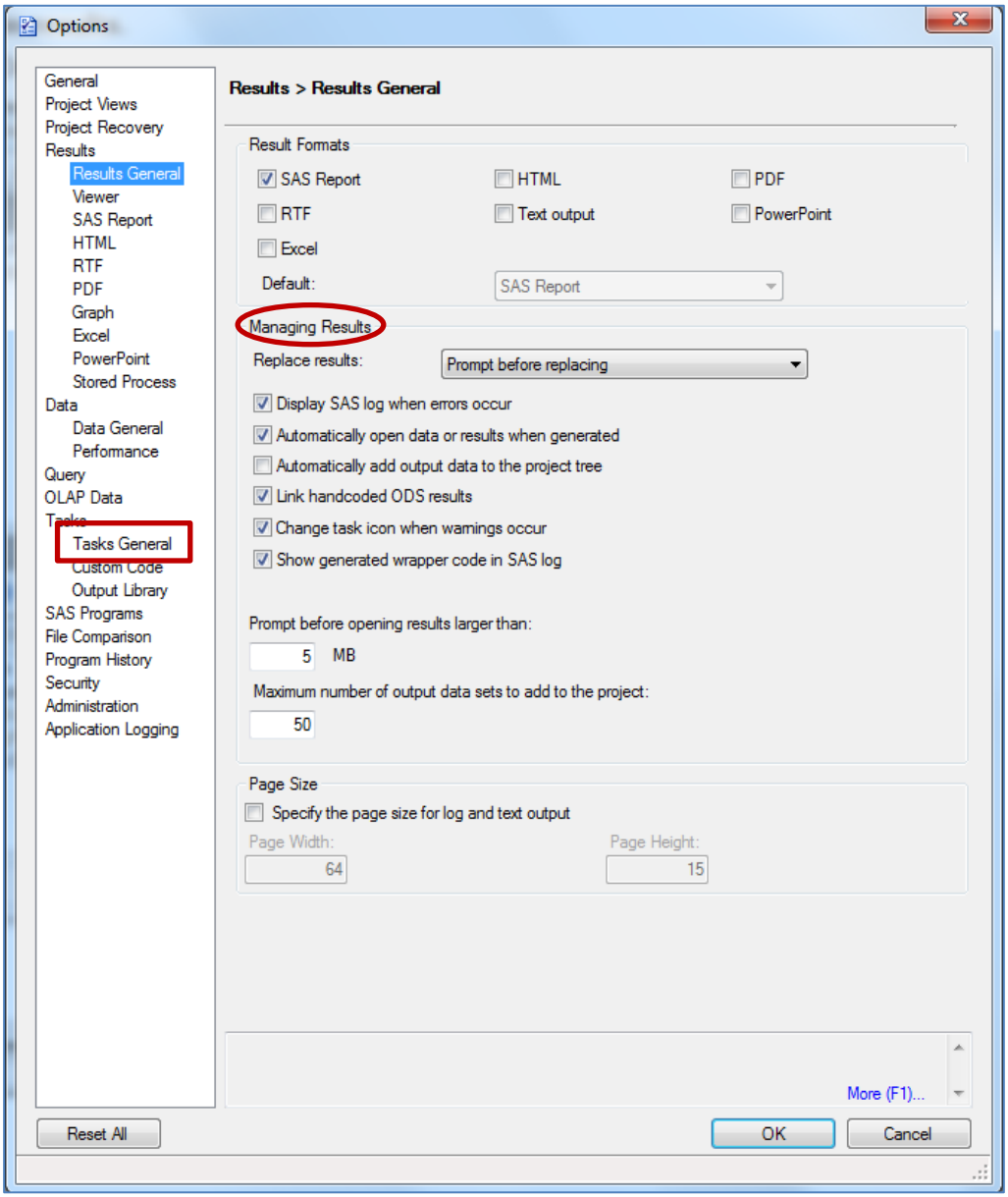
- Click **New Project** in the **Welcome to SAS Enterprise Guide** window.

**To start a new project when SAS Enterprise Guide is already running:**

- Select **File -> New**-> **Project**.
- You can also create a new project by clicking in the toolbar and then selecting **Project** from the drop-down list.

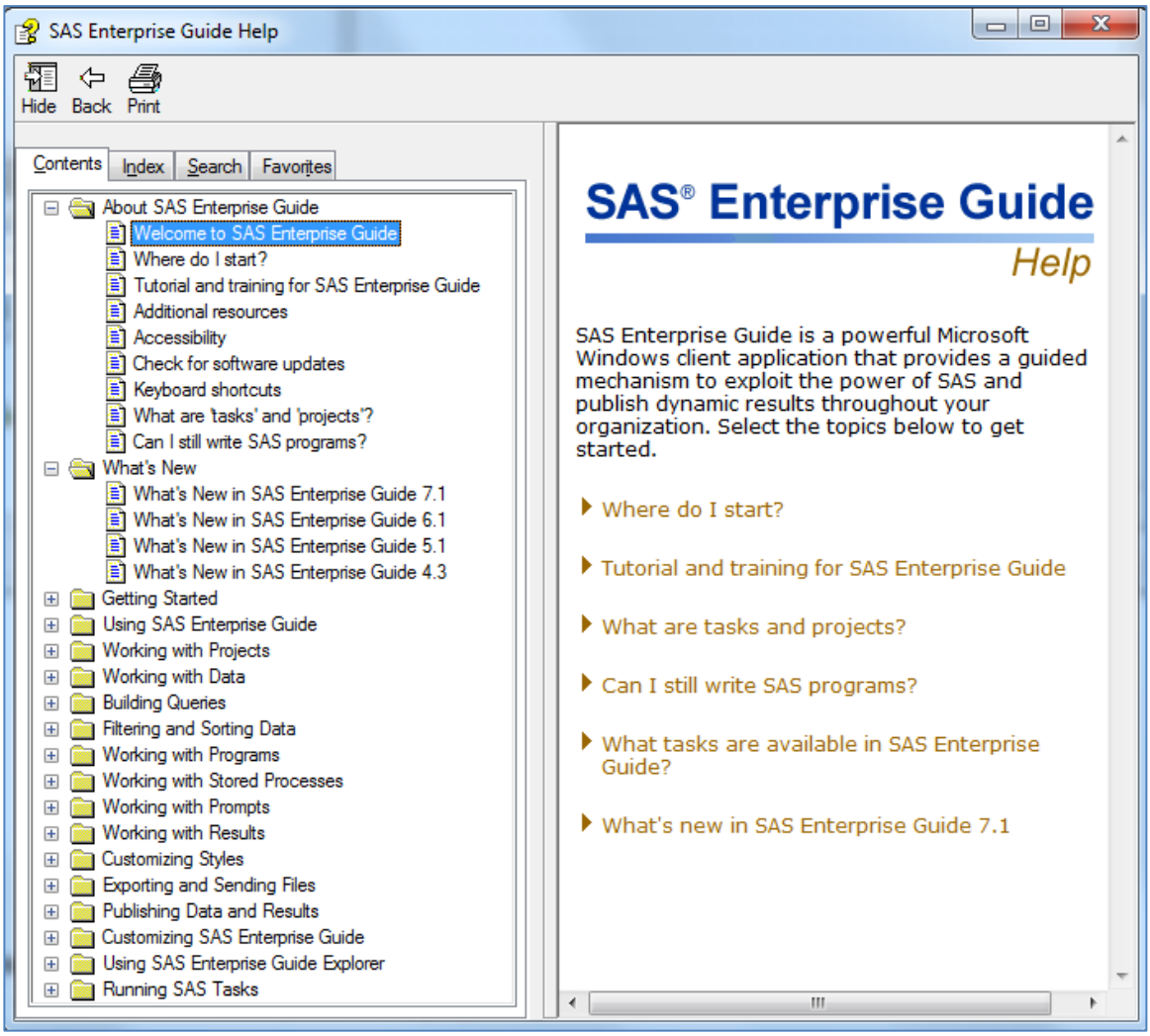
**Setting Options**

- Select **Tools** -> **Options**.
- On the **Results** tab, click the **Results General**, check the **HTML**.
- On the **Tasks** tab, click the **Tasks General**, deselect and delete the **Default footnote text for task output**, and deselect **Include SAS procedure title in results**.
- On the **Managing Results**, check mark the **Automatically add output data to the project tree.**
- Click **OK**.



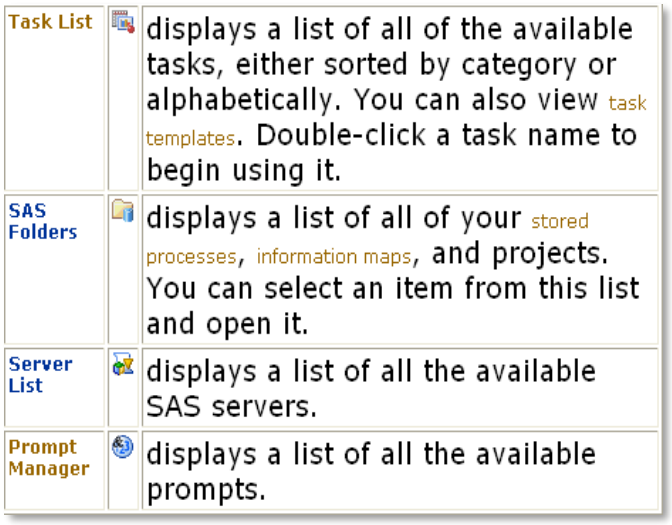
**Enterprise Guide Help**

- From the **Help** menu, select **SAS Enterprise Guide Help**.
- On **Contents** tab, double-click the folder icons to display the help topics.
- Or- from the **Help** menu, select **Getting Started Tutorial**.
- Click the **Begin Tutorial** link (Online).
- **Help** -> **Check for Updates.**
- **Close** this window.



**Working with Projects**

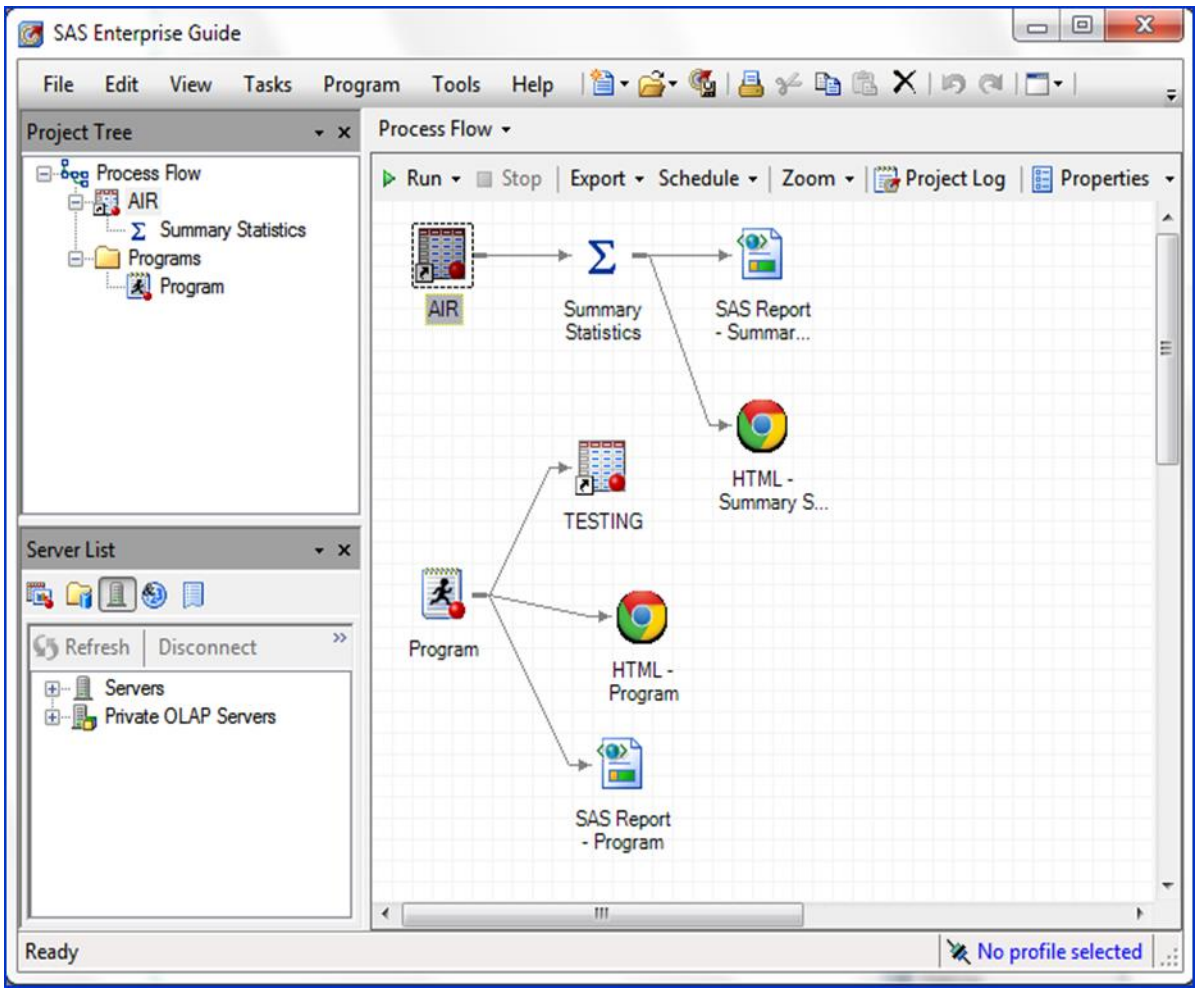
- The **Resources** pane is displayed by default in the lower-left corner of the SAS Enterprise Guide window, and it provides access to the **Task List**, **SAS Folders**, the **Server List**, and the **Prompt Manager**.
- By default, the Resources pane displays the **Server List**. You can view the other windows in the Resources pane by clicking the appropriate icon in the Resources pane toolbar or by using the **View** menu.



| | |
|---|---|
| Task List | displays a list of all of the available tasks, either sorted by category or alphabetically. You can also view task templates. Double-click a task name to begin using it. |
| SAS Folders | displays a list of all of your stored processes, information maps, and projects. You can select an item from this list and open it. |
| Server List | displays a list of all the available SAS servers. |
| Prompt Manager | displays a list of all the available prompts. |

**Working with Process Flows**

- A process flow consists of one or more objects and could contain a **project tree**.
- An object is represented by an icon in the process flow, such as a **data set**, a **task**, and your **results**.
- A process tree shows the relationship between two or more objects.
- SAS Enterprise Guide creates a default process flow when you start a project.
- For a single project, you can have **multiple process flows**.
- To view the process flow, select **View Process Flow**, or click **Process Flow** in the project tree.




**Differences between the process flow and the project tree**

- You can customize the process flow by turning ***grid lines*** on and off, by zooming, or by changing the background color.
    - Right-click on the **Process Flow**, and then select to change the background color.
- You can create multiple **Process Flows** and copy and paste or move objects among them.

**Customizing the View of Process Flows**

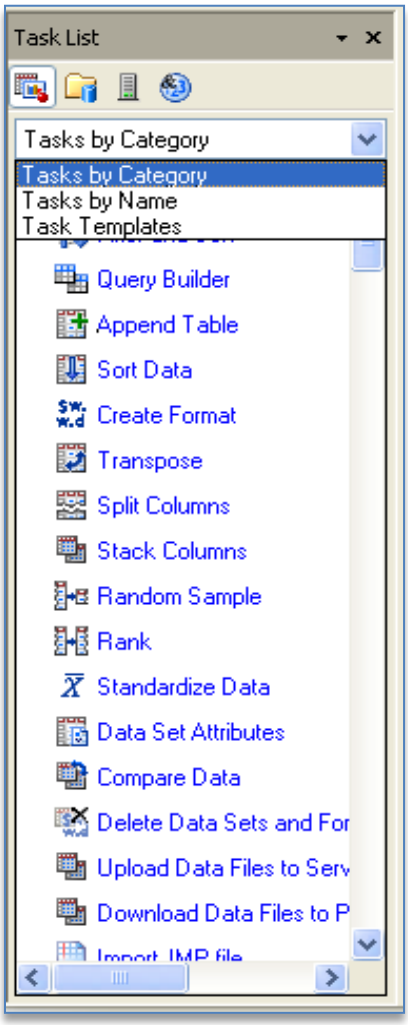
- **To turn off the grid line**
    - Right-click in a process flow and select Grid.
- **To view all of the objects in your process flow at one time**
    - Right-click in a process flow and select Zoom.

  - Select the size (in percent) of the window.
- **To change the background color of the process flow**
  - Right-click in a process flow and select Background Color.
  - Select the color that you want to use. By default, the background color is white.

**Task List**

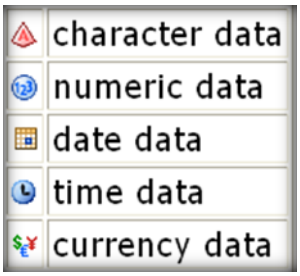
We use the **Task List** to tell SAS Enterprise Guide what task we want to be performed. The SAS Enterprise Guide ***tasks*** generate **SAS code** and **formatted results** for you. The **Task List** window can be viewed in the **Resources pane** and contains a list of the analysis and reporting tasks that you can run on your data.

- To begin a task, double-click the task name in the Task List window. If there is actively selected data that is opened in **Update mode** when you run a task, you are prompted to change it to **Read-only mode** before you can run the task.
- To open the Task List, select **View** -> **Task List**. You can also choose tasks from the **Tasks** menu in the SAS Enterprise Guide window.
- You can use the drop-down list to view the tasks in two different ways and to view task templates:
  - **Tasks by Category** - displays a list of tasks that are grouped by category.
  - **Tasks by Name** - displays an alphabetical list of tasks. The alphabetical list also includes the name of the SAS Procedure that is related to the task. To sort by the Procedure name, click the **SAS Procedures** column heading.
  - **Task Templates** - displays a list of your task templates.

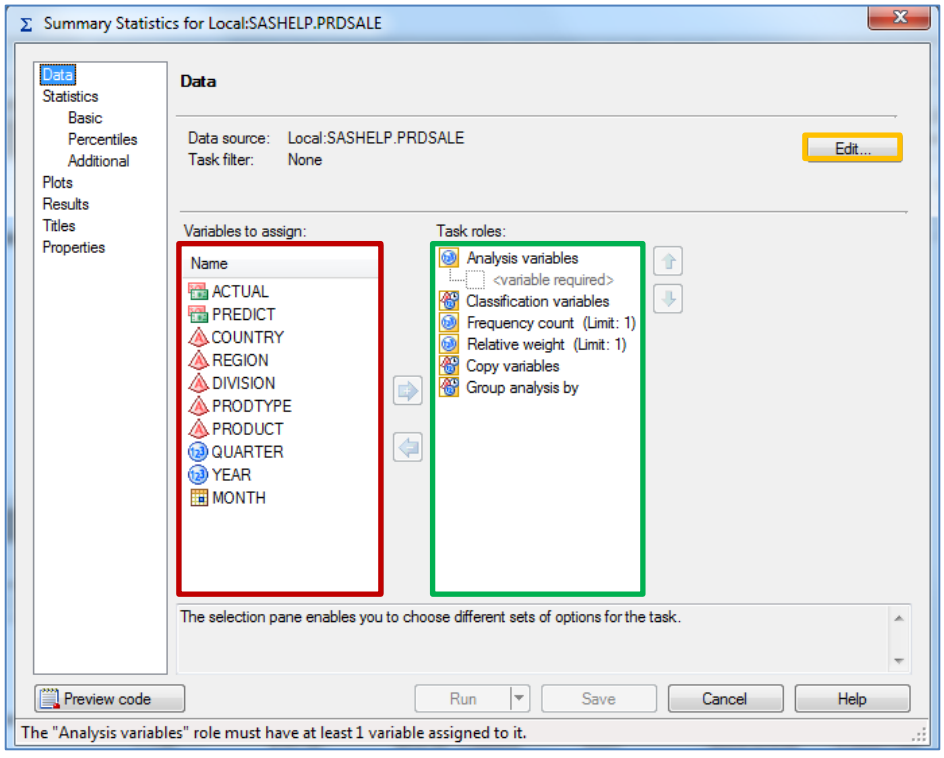


**Using Tasks in SAS Enterprise Guide**

- Use the selection pane on the left side of task window to select groups of options for the task (Data, Options, Titles, etc.).
- The **Variables to assign** box lists the variables (or columns) in the data set.
- The icon next to each variable represents the variable's type:




- All tasks have a **Data** area where you assign variables to roles in the task.
- The icon next to each task role indicates the type of variable that you can assign to the role.
- You assign variables to roles by dragging them from the **Variables to assign** list and dropping them on the appropriate role in the Data list. You can also use the arrows to assign and remove variables from roles. Variables can be assigned to more than one role.



**Task Status Window**

- The Task Status window displays messages about the status of tasks as they process.
- The Process Designer window also displays a different background color around the task item for each phase of processing.

**Data in SAS Enterprise Guide**

- SAS Enterprise Guide requires all Data that it accesses to be in **table format**.
- **Rows** (also called **observations**) in a table are collections of Data values related to an object.
- **Columns** (also called **variables**) in a table are collections of values that describe a characteristic.

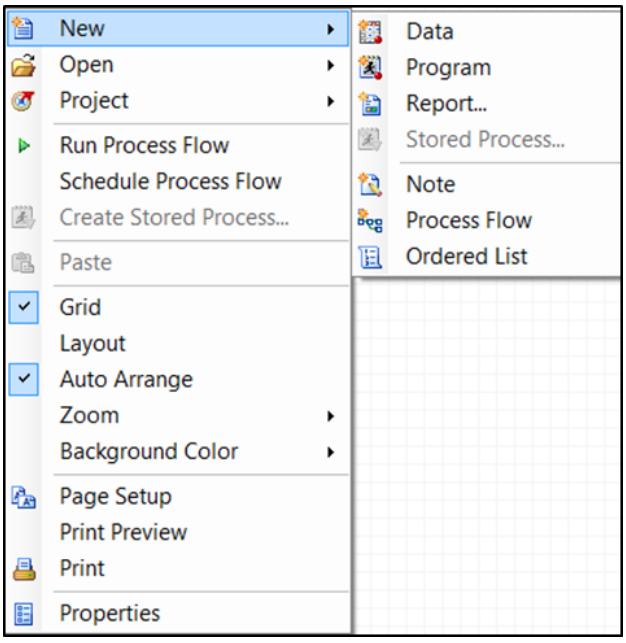
**Column (Variable) Properties**

- **Name -** can be **1 to 32 characters** long. They must begin with a **letter (A-Z, either uppercase or lowercase**) or an underscore **( _ ).** They can continue with any combination of numbers, letters, or underscores, **including blanks.** However, when SAS Enterprise Guide processes a variable name, SAS Enterprise Guide internally converts it to uppercase. You cannot, therefore, use the same variable name with a different combination of uppercase and lowercase letters to represent different variables. For example, cat, Cat, and CAT all represent the same variable.
- **Type** - is either **character** or **numeric**.
- **Character variables** can contain **any values**. Missing character values are represented by a blank.
- **Numeric variables** can contain **only numeric values** (the digits 0 through 9, +, -, . , and E for scientific notation).
- **Currency, date, and time data are stored as numeric variables.**
- **Missing numeric values are represented by a period.**
- **Length** - A column's **length** (the number of bytes that are used to store it) is related to its type.
- Character variables have a default length of **12 bytes.**
- All numeric variables have a default length of **8 bytes**.
- **Format** - affects how data values are displayed. SAS data formats include **character**, **numeric**, and **date** and **time**.

- **Informat** - determines how data values are read into a SAS data set. You **must** use informats to read numeric values that contain letters or other special characters.
- **Label** - A variable can have a **label**, which consists of descriptive text up to 256 characters in length.

**To create a new program**

- Select **File** -> **New Program**. The new program opens in the workspace.
- You can also **right-click** in the process flow and select New -> Program
- The autocompleting, or code completion, feature in the program editor can predict the next word that you want to type before you actually type it completely.
- To use the autocompletion feature, in the program editor, type the first one or more letters of the word that you want to use. A window opens with a list of suggested keywords that begin with those letters.
- Navigate to the keyword that you want to use.
    - Continue to type until the correct keyword is selected.
    - Or- Scroll through the list
    - Double-click the selected keyword.
    - Or- Press the spacebar, ENTER, or TAB keys.



***Example:***

- File -> New -> Program
- Type the following program:

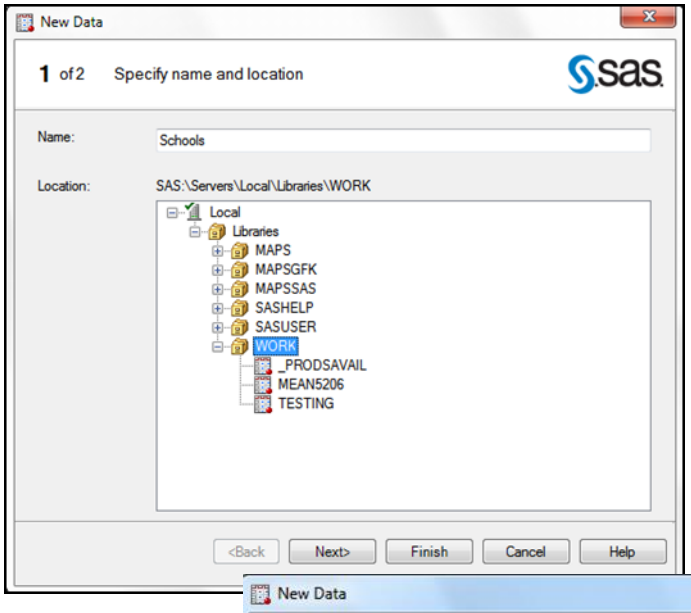
```
data test;
input x;
datalines;
1
2
3
4
5
;
run;
proc print data=test;
run;
```

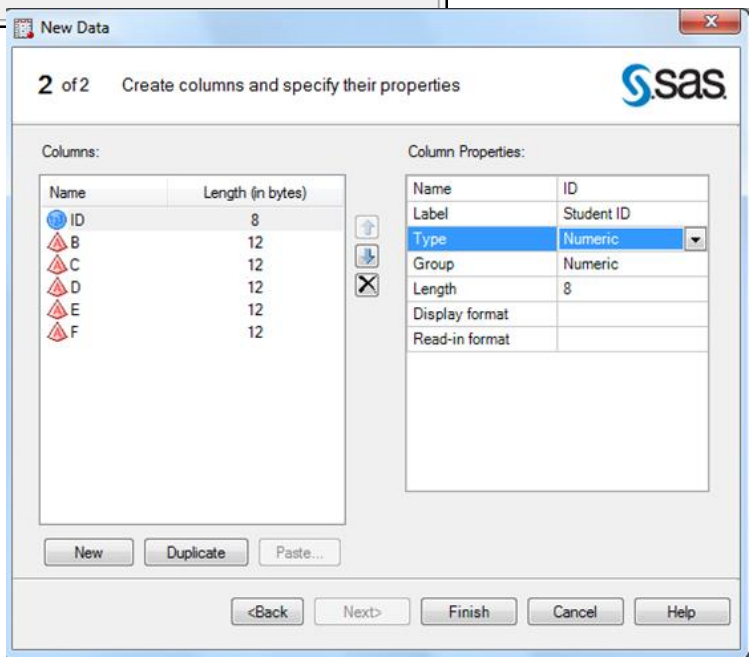
- Click on the **Run** button.
- View different outputs (PDF, HTML, simple SAS Output), if the settings are selected (Tools -> Options -> Results General).

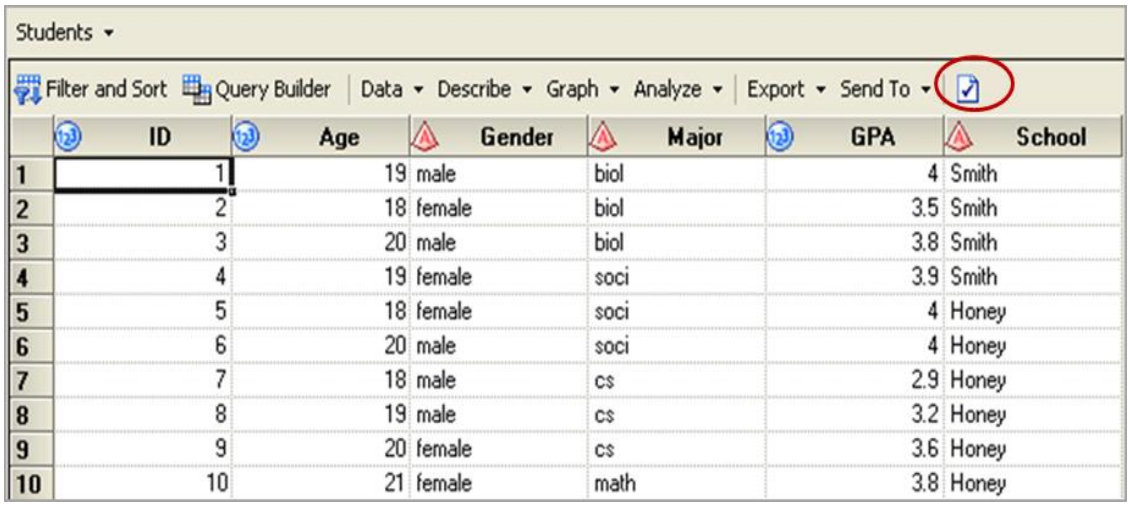
**Exercise #1: Creating a New data Set**

- Select **File > New > data**. The **New data wizard** opens to step you through the process of specifying the name and location of the new data and creating the initial columns.
- Name your data **Schools**, and click **Next.**
- Under **Column Properties:** select and change column Name **A** to **ID.**
- For this column's **Label**, type **Student ID.**
- Change this column's data **Type** to **Numeric**.
- Add the following column headings (**variable names**) to the table, selecting columns from the **Columns List** and changing their names and **properties** as specified:
    - **Age** Numeric
    - **Gender** Character
    - **Major** Character
    - **GPA** Numeric
    - **School** Character
- To create a new column name, click **New**.
- To rearrange the order of the existing columns, select the column in the **Columns** list that you want to move and click  or  .

- To delete an existing column, select the column in the **Columns** list that you want to delete and click  .
- Close this new data window.
- When completed, Click **Finish**.
- Type the data from the following table into each cell in the grid by using the Tab or arrow keys to navigate around the grid. (you can Copy & Paste data, one entry at a time)
- **Delete** unused rows and columns. **View** Table **Properties**.
- **Don't Close** this new data window.




Students

Filter and Sort | Query Builder | Data | Describe | Graph | Analyze | Export | Send To

| | ID | Age | Gender | Major | GPA | School |
|---|---|---|---|---|---|---|
| 1 | 1 | 19 | male | biol | 4 | Smith |
| 2 | 2 | 18 | female | biol | 3.5 | Smith |
| 3 | 3 | 20 | male | biol | 3.8 | Smith |
| 4 | 4 | 19 | female | soci | 3.9 | Smith |
| 5 | 5 | 18 | female | soci | 4 | Honey |
| 6 | 6 | 20 | male | soci | 4 | Honey |
| 7 | 7 | 18 | male | cs | 2.9 | Honey |
| 8 | 8 | 19 | male | cs | 3.2 | Honey |
| 9 | 9 | 20 | female | cs | 3.6 | Honey |
| 10 | 10 | 21 | female | math | 3.8 | Honey |

**Scatter Plots**

- A scatter plot (also called **Scatter Diagram**, or **Scatter Chart**) is type of graph that is useful when you are plotting one multi-value (paired or 3-D) variable against a second multi-value variable.
- Each data point in a scatter plot, represents a single observation (typically a human subject).
- You should always create a Scatter plot for a given pair of variables prior to computing the correlation coefficient between those variables.
- This is because The Pearson Correlation Coefficient is appropriate **ONLY** if the relationship between the variables is **LINEAR**.

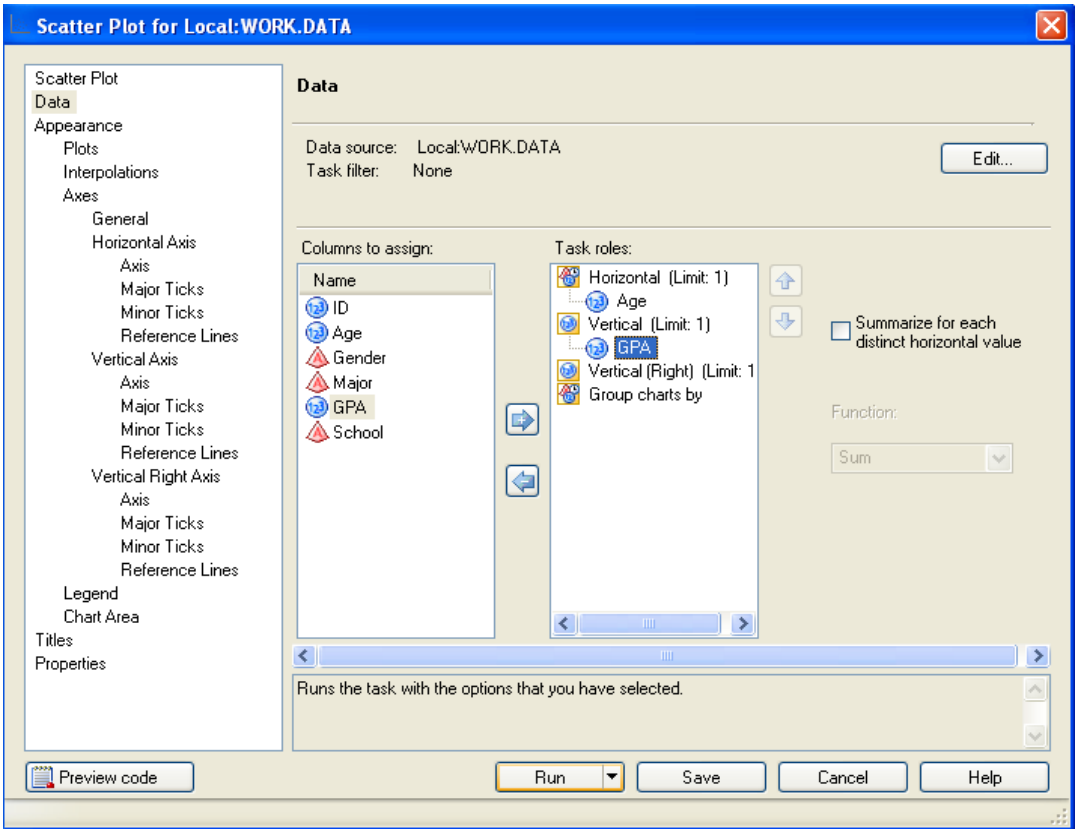
**When to Use a Scatter diagram?**

- ✓ When you have paired numerical data; and
- ✓ When trying to determine whether the two variables are related.

**Exercise #2: Creating a Scatter Plot (to see the pattern of data)**

- While the table is open, from the **Graph** menu, select **Scatter Plot . . .**
- Click yes, to protect your data (if you are asked).
- Select **2D Scatter Plot**.
- Click the **Data** tab.
- Select **Age** and drag it to **Horizontal** (axis) and select **GPA** and drag it to **Vertical** (axis) Data.

- On the **Plots** tab (under **Appearance** on the **Plot** tab), on the Data Point Marker, choose the **Symbol: Star**, and a **Color**.
- On the **Titles** tab, uncheck and delete the default text.
- On the **Titles** tab, clear the **Footnote**.
- When finished assigning variables to roles and selecting options, click **Run**. Each pair of observations becomes a data point on the plot.
- View the **HTML** output and the close this window.

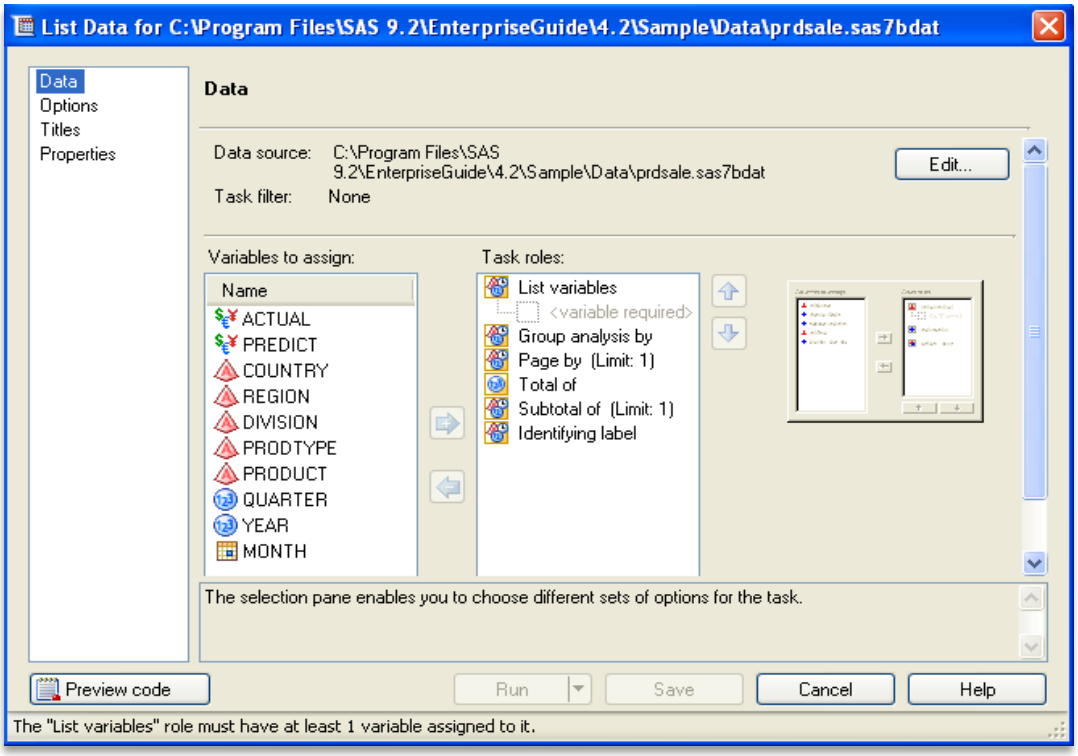



**To change the protection mode, to edit data files**

- Open your data set
- Select **Edit** > **Protect Data** to toggle the **Update mode** on and off.
- You are prompted for confirmation when you change a file from **Read-only** mode to **Update mode**.

**Assigning Data**

- To create ***reports*** and run ***analyses*** on the data, you select tasks from the **Task List**, either sorted by category or alphabetically.
- A ***task*** is an **analysis** that you perform or a **report** that you create from your data.
- Many Enterprise Guide tasks will require an analysis and /or classification variable.
- The available roles vary depending on the task , but some of the most common are:
    - ***Analysis Variables*** *are variables for which statistics will be produced, e.g. mean, standard deviation.* ***An analysis variable must be numeric.***
    - ***Dependent variable*** *is the response variable.*
    - ***Classification variables*** *are variables by which subjects (participants) will be classified within an. Classification variables can be either* ***character*** *(categorical) or* ***numeric****.*
    - ***Group analysis by*** *variables; are also discrete, numeric, or character variables. For example Gender variable, with values male, and female.*



**Exercise #3: Creating a List data for Schools data set**

- Double–click to open the Schools data set (if it is not open).
- From the **Describe** menu, select **List Data . . .**
- **Shift + Click** on first and last variables names to select them all.
- On the **Data** tab, drag all the variables to the right and drop them under the List variables heading, to assign procedures to variables.
- On the **Options** tab, uncheck mark **Print the Row number** box.
- On the Titles tab, for the **Report Titles**, uncheck mark the **Use default text** box, and delete the text for **Footnote**.
- Rename the Title to "**Schools Report**"
- Click the **Preview Code** button.
- Close the **Code Preview for Task** window.
- On the **Titles** tab, clear the Footnote.
- When finished assigning variables to roles and selecting options, click **Run**.
- View the **HTML** output and the close this window.

**School Report**

| Row number | ID | Age | Gender | Major | GPA | School |
|---|---|---|---|---|---|---|
| 1 | 1 | 19 | m | b | 4.0 | Smith |
| 2 | 2 | 18 | f | b | 3.5 | Smith |
| 3 | 3 | 20 | m | b | 3.8 | Smith |
| 4 | 4 | 19 | f | s | 3.9 | Smith |
| 5 | 5 | 18 | f | s | 4.0 | Honey |
| 6 | 6 | 20 | m | s | 4.0 | Honey |
| 7 | 7 | 18 | m | c | 2.9 | Honey |
| 8 | 8 | 19 | m | c | 3.2 | Honey |
| 9 | 9 | 20 | f | c | 3.6 | Honey |
| 10 | 10 | 21 | f | m | 3.8 | Honey |

**Editing SAS program generated by a Task**

- To make changes to the programs generated by SAS Enterprise Guide, 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.

- To preview the code generated by a Task, click the **Preview code** button in the lower-left corner of any **Task Window**.
- And then click the **Insert Code …** button, to insert codes at specific points in the SAS program, by **double-clicking to insert code**.

**Enterprise Guide and SAS programs**

- **To create a new SAS program**, from the File menu, click **New** -> **Program**. The program editor opens and it is **syntax-sensitive**.
- **To open an existing SAS program**, from the File menu, click **Open** -> **Program**. Navigate to the existing SAS program and click Open.
- **To save a program in a file:**
    - Any new programs you write are automatically embedded in your project (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 As from the menu.
- **To run your program**, click Run on the workspace toolbar for the Program window.

**Changing the Report Style**

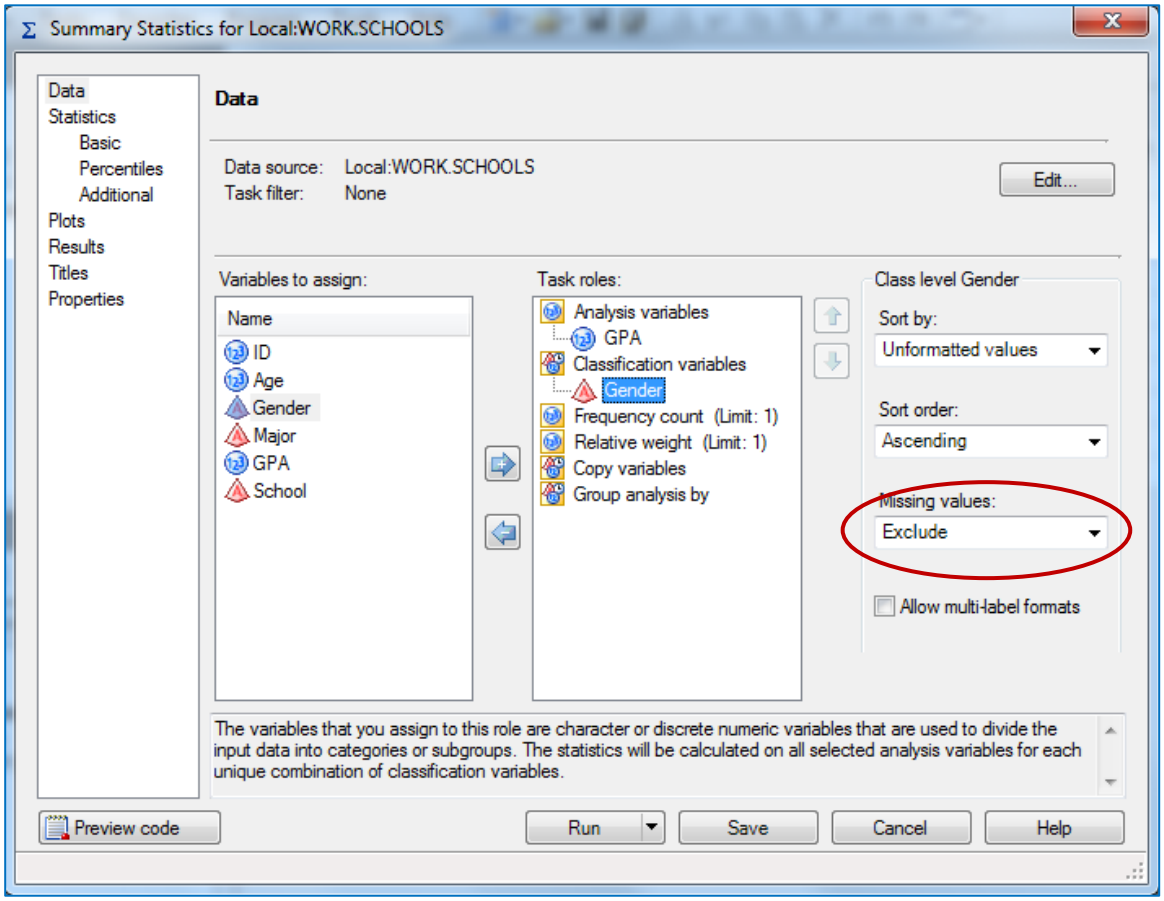
- From the **Tools** menu, select **Options**
- On the **Results** tab, select **HTML, if it is not selected.**
- On the HTML tab, change the **Appearance Style**, if you wish.

**Exercise #4: Calculating Summary Statistics (the Means Procedure in SAS)**

- Double-click to open the Schools data set.
- From the **Describe** menu, select **Summary Statistics . . .**
- Select **GPA** variable, and drag it to **Analysis Variables**.
- On **Statistics** tab, for **Basic statistics**, select **3** from the **Maximum decimal places** drop down list.
- On the **Data** tab, drag the **Gender** to **Classification variable**. This gives you additional options such as **Missing values**:
    - The Missing Values drop-down menu lists **Include** and **Exclude**.

- Missing values can be set to include or exclude missing values from analysis. The default for Missing values is "**Exclude**".

- On the **Titles** tab, clear the Footnote.
- Click **Run**.
- View the HTML output and the close this window.




**Frequency Distribution for Discrete Data**

- Discrete data can be summarized using the following distributions:
  - The ***frequency distribution***, which lists the number of occurrences of each discrete value (character values, counts, and numeric classifications).
  - The ***percent distribution***, which lists the percentage of occurrences of each distinct value in the sample.

  - The ***cumulative frequency distribution***, which lists a running total of frequencies.
  - The cumulative percent distribution lists a running total of the percentages.

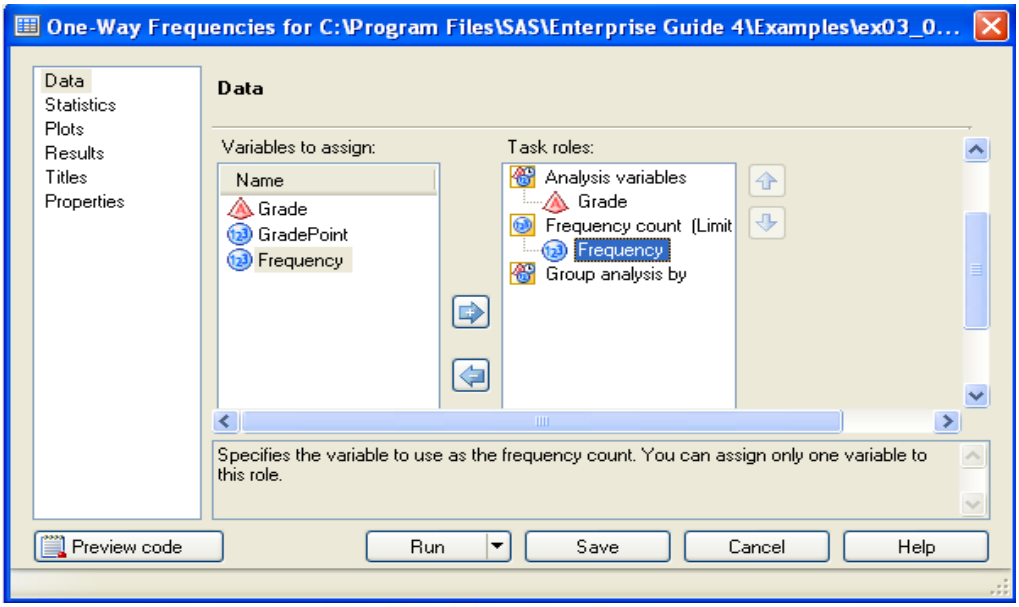
- Frequency distribution lists the number of occurrences of each distinct value (discrete data), in the sample. The underlying SAS procedure is Proc FREQ.

**Exercise #5: Creating Frequency Distribution and Frequency Histogram for Discrete Data**

Using the ex03_07.sas7bdat file, create frequency distribution for discrete data, as follows:

- File -> Open -> Data … -> Local Computer -> My Computer\ShortCourses Material\SAS and SPSS\SAS\SAS Enterprise Guide -> Examples -> ex03_07.sas7bdat -> Open
- From the **Describe** menu, choose **One-Way Frequencies**
- On the **Data** tab, assign the **Grade** variable to the **Analysis variables** Data.
- Assign **Frequency** to the **Frequency Count** role.
- On the **Statistics** tab, check mark the "**Frequencies and percentages with cumulative**" radio button.
- On the **Plots** tab, choose **Vertical**.
- When finished assigning variables to roles and selecting options, click **Run**.
- To make changes to the chart, Using the HTML output for this report, right-click inside the graph, and then choose options from the pop-up menus.
- View the HTML output and the close this window.

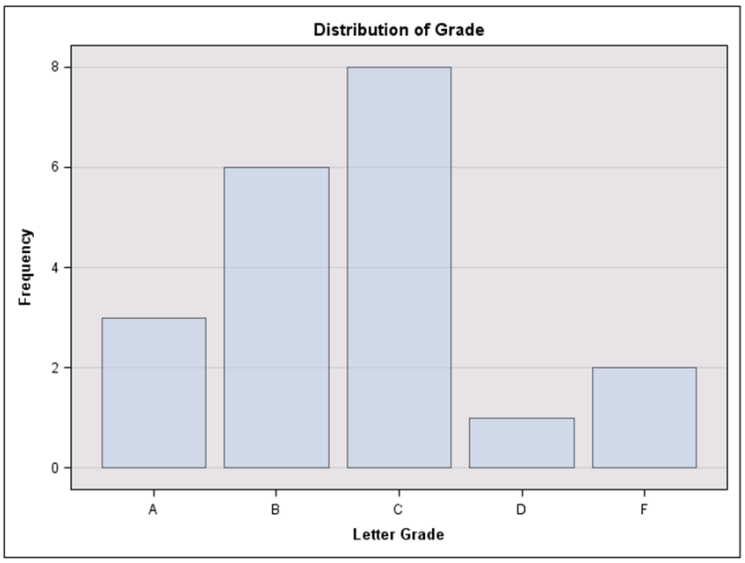
| | Grade | GradePoint | Frequency |
|---|---|---|---|
| 1 | A | 4 | 3 |
| 2 | B | 3 | 6 |
| 3 | C | 2 | 8 |
| 4 | D | 1 | 1 |
| 5 | F | 0 | 2 |




## One-Way Frequencies
## Results

***The FREQ Procedure***

| Letter Grade | | | | |
|---|---|---|---|---|
| Grade | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| A | 3 | 15.00 | 3 | 15.00 |
| B | 6 | 30.00 | 9 | 45.00 |
| C | 8 | 40.00 | 17 | 85.00 |
| D | 1 | 5.00 | 18 | 90.00 |
| F | 2 | 10.00 | 20 | 100.00 |



**Exercise #6: Performing One-Way ANOVA Test**

48 women (35-45 years old) participated in a study that examines the effectiveness of certain interventions in increasing physical activity. Their daily energy expenditure (DEE) is at most 35 kcal/kg/day (considered inactive). Each woman was randomly assigned to one of 4 groups (Advice, Assistance, Counseling, and Control). After 6 months, the DEE for each woman is measured.

- File -> Open -> Data … -> Local Computer -> My Computer\ShortCourses Material\SAS and SPSS\SAS\SAS Enterprise Guide -> Examples -> ex06_01.sas7bdat -> Open
- Analyze > ANOVA > One-Way ANOVA …
- Assign the DEE variable to Dependent variables Task role.
- Assign the Group variable to Independent variable Task role.
- On the **Means** tab, click the **Comparison**, and then select the **Fisher's least significant -difference test (LSD)**. Or, select **Tukey's Studentized range test (HSD)** which has a higher Type II error rate.
- On the **Means** tab, click the Breakdown, and then select Mean, Standard deviation, Variance, and Number of non-missing observations.
- On the Plots tab, check both Box and Whisker, and Means
- On the Titles tab, clear the Footnote.
- When finished assigning variables to roles and selecting options, click Run.
- View the HTML output and the close this window.

**One-Way Analysis of Variance Results**

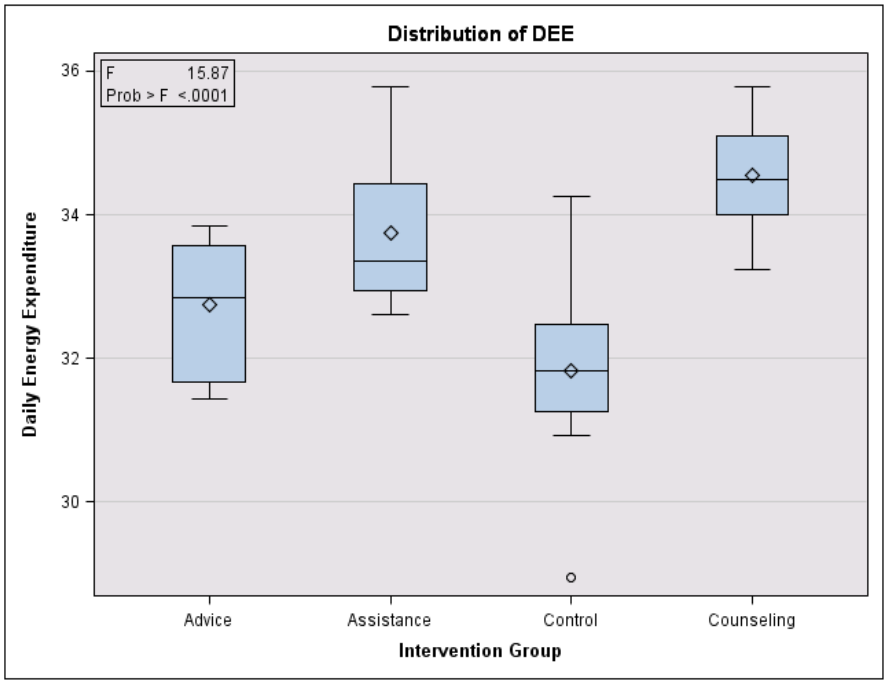
***The ANOVA Procedure***

***Dependent Variable: DEE Daily Energy Expenditure***

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 3 | 50.31750000 | 16.77250000 | 15.87 | <.0001 |
| Error | 44 | 46.50620000 | 1.05695909 | | |
| Corrected Total | 47 | 96.82370000 | | | |

| R-Square | Coeff Var | Root MSE | DEE Mean |
|---|---|---|---|
| 0.519682 | 3.095011 | 1.028085 | 33.21750 |

| Source | DF | Anova SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Group | 3 | 50.31750000 | 16.77250000 | 15.87 | <.0001 |




## Results

### *The ANOVA Procedure*

### *t Tests (LSD) for DEE*

**Note:** This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

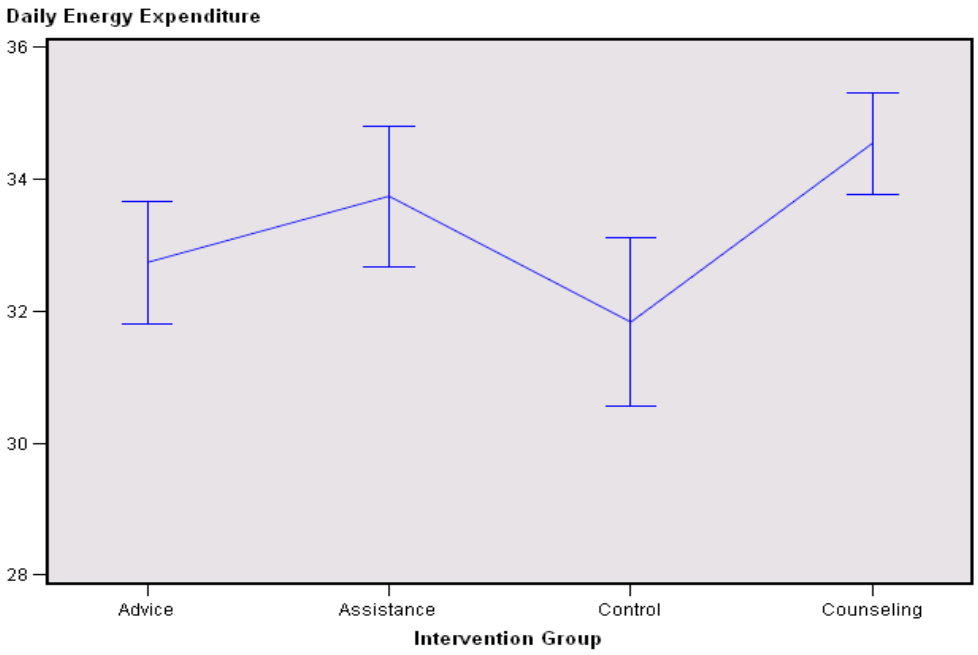
| | |
|---|---|
| **Alpha** | 0.05 |
| **Error Degrees of Freedom** | 44 |
| **Error Mean Square** | 1.056959 |
| **Critical Value of t** | 2.01537 |
| **Least Significant Difference** | 0.8459 |

| Means with the same letter are not significantly different. | | | |
|---|---|---|---|
| **t Grouping** | **Mean** | **N** | **Group** |
| A | 34.5500 | 12 | Counseling |
| A | | | |
| A | 33.7450 | 12 | Assistance |
| | | | |
| B | 32.7400 | 12 | Advice |
| | | | |
| C | 31.8350 | 12 | Control |

## One-Way Analysis of Variance
## Results
## Means and Descriptive Statistics

| Intervention Group | Mean of DEE | Std. Dev. of DEE | Variance of DEE | Number of non-missing values for DEE |
|---|---|---|---|---|
| | 33.2175 | 1.4352974338 | 2.06008 | 48 |
| Advice | 32.74 | 0.927881262 | 0.86096 | 12 |
| Assistance | 33.745 | 1.064608336 | 1.13339 | 12 |
| Control | 31.835 | 1.2811465036 | 1.64134 | 12 |
| Counseling | 34.55 | 0.7695098794 | 0.59215 | 12 |

## Means Plot of DEE by Group




**Results**

- Independent variable is Group (k=4)
- Dependent variable is DEE
- Coefficient of determination $R^2$ = 0.519682
- $\alpha$ = 0.05
- Degrees of freedom = n-k =48-4 = 44

- MS (Error) = 1.056959
- $T_{.025}$ = 2.01537
- The Control group has the lowest sample mean. All other groups are significantly higher.

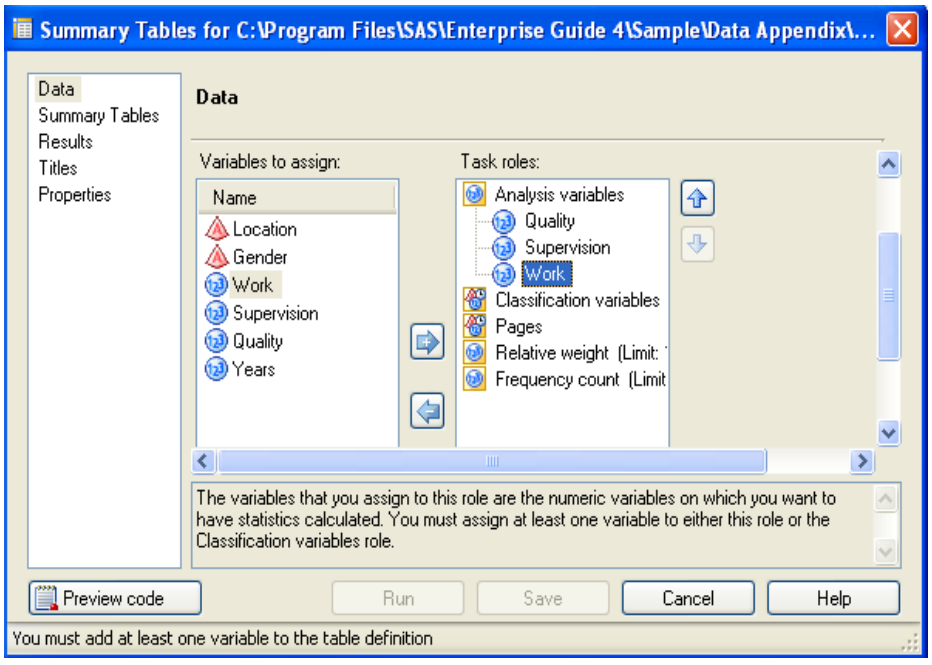
**Exercise #7: Creating Summary Tables**

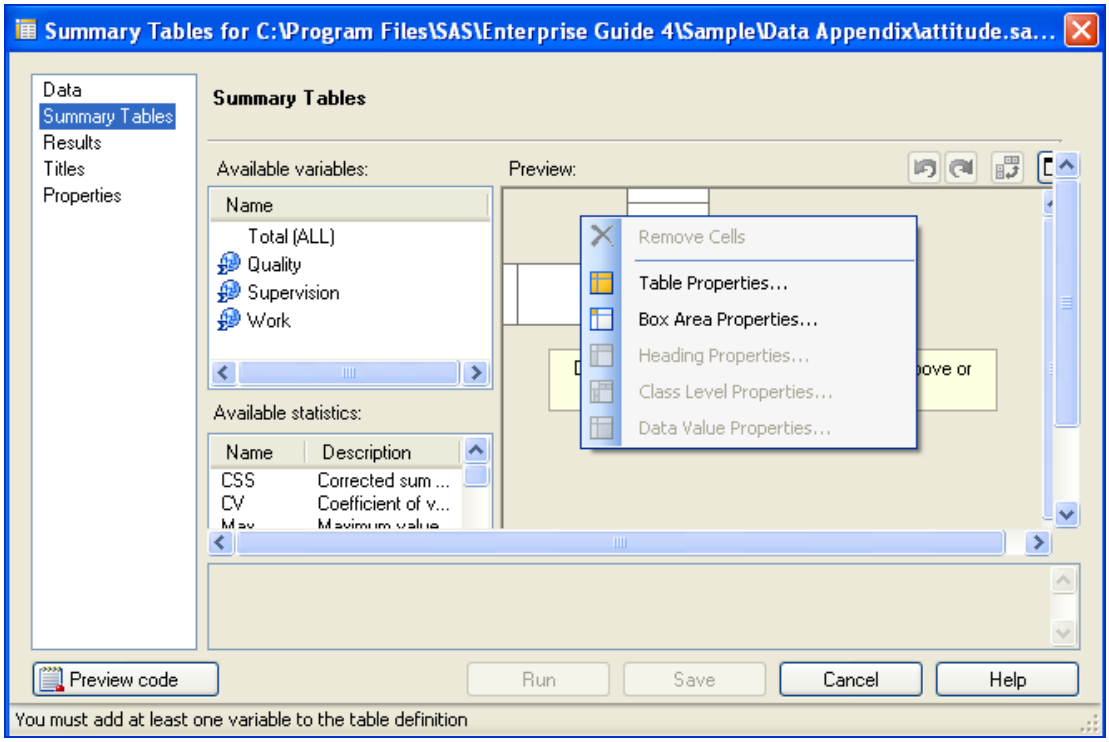
Sometimes **Summary Tables** are preferred over the **Summary Statistics** and One-Way Frequencies. Summary Tables need to be set up, and they do not produce graphs. Using the SAS Attitude data set, create a Summary Table as follows:

- File -> Open -> Data-> Local Computer -> My Computer-> ShortCourses -> ShortCourses Material\SAS and SPSS\SAS\SAS Enterprise Guide -> Sample -> Data Appendix -> attitude.sas7bdat -> Open
- From the **Describe** menu, choose **Summary Tables**…
- On the **Data** tab, assign **Quality**, **Supervision**, and **Work** variables to **Analysis** variables Data.
- On the **Summary Tables** tab, right-click on the **Preview Table**, and then select **Table Properties**.
- Click the **Format** tab, and then select **Numeric** from the **Categories**, and **BESTw.d** from the **Formats** box.
- Click **OK**.
- To add text to the table's box area:
    - Right-click on the Preview table, and select Box Area Properties from the drop down menu.
    - On the General tab, select Use the following text, and type your text (Attitude, for example).
    - Click **OK**.
- On the **Summary Tables** tab,
    - **From Available Variables:** one-by-one drag the 3 variables (Quantity, Supervision, and Work) into the cells, under the Box Area. *Note:* A summary table is constructed by dragging variables and statistics from the available boxes to the left and top portions of the Preview table.
    - **From Available Statistics:** drag Max, Mean, Min, N, and StdDev, to the right border of the last cell in the upper portion of the Preview table.
    - On the **Titles** tab, clear the Footnote.

- When finished assigning variables to roles and selecting options, click **Run**.
- View the HTML output and the close this window.

| | Location | Gender | Work | Supervision | Quality | Years |
|---|---|---|---|---|---|---|
| 1 | Central | F | 54 | 50 | 46 | 1 |
| 2 | Central | F | 53 | 52 | 51 | 4 |
| 3 | Central | F | 49 | 49 | 52 | 6 |
| 4 | Central | F | 51 | 52 | 53 | 5 |
| 5 | Central | F | 50 | 46 | 54 | 7 |
| 6 | Central | M | 47 | 35 | 54 | 5 |
| 7 | Central | M | 48 | 43 | 60 | 12 |
| 8 | North | F | 51 | 51 | 45 | 1 |
| 9 | North | F | 52 | 52 | 48 | 4 |
| 10 | North | F | 51 | 53 | 51 | 5 |
| 11 | North | F | 59 | 44 | 52 | 6 |
| 12 | North | F | 58 | 53 | 53 | 8 |
| 13 | North | M | 52 | 52 | 40 | 0 |
| 14 | North | M | 51 | 57 | 50 | 5 |
| 15 | North | M | 51 | 48 | 52 | 7 |
| 16 | South | F | 59 | 61 | 45 | 1 |
| 17 | South | F | 57 | 49 | 48 | 2 |
| 18 | South | F | 56 | 55 | 50 | 6 |
| 19 | South | F | 56 | 53 | 52 | 5 |
| 20 | South | F | 58 | 55 | 53 | 8 |
| 21 | South | M | 57 | 59 | 45 | 2 |
| 22 | South | M | 55 | 50 | 47 | 3 |
| 23 | South | M | 51 | 54 | 48 | 4 |
| 24 | South | M | 57 | 56 | 49 | 4 |




Summary Tables for C:\Program Files\SAS\Enterprise Guide 4\Sample\Data Appendix\attitude.sa...
Data
Summary Tables
Results
Titles
Properties
Summary Tables
Available variables:
Preview:
Name
Total (ALL)
Quality
Supervision
Work
Remove Cells
Table Properties...
Box Area Properties...
Heading Properties...
Class Level Properties...
Data Value Properties...
Available statistics:
Name
Description
CSS
Corrected sum ...
CV
Coefficient of v...
Preview code
Run
Save
Cancel
Help
You must add at least one variable to the table definition

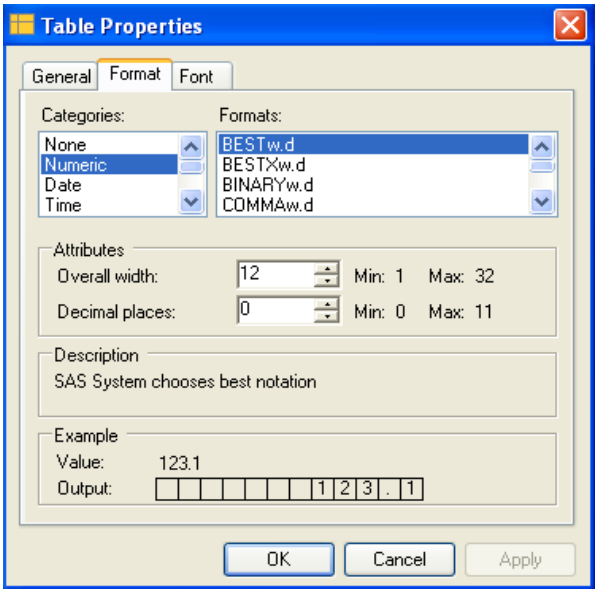
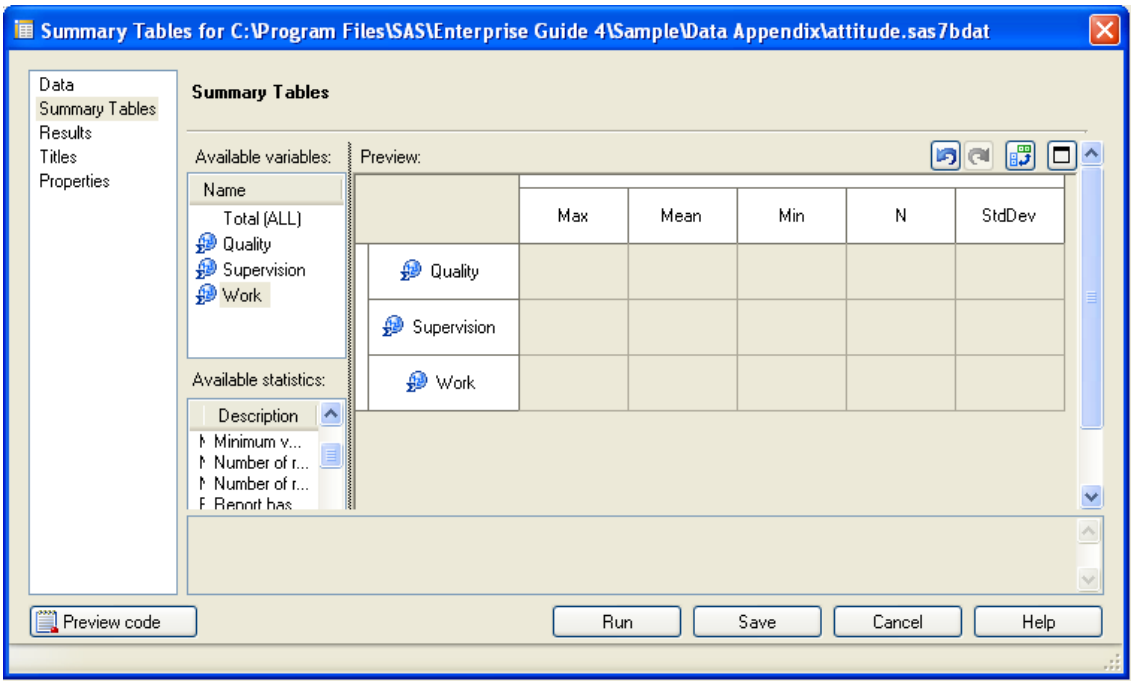


Table Properties
General
Format
Font
Categories:
Formats:
None
Numeric
Date
Time
BESTw.d
BESTXw.d
BINARYw.d
COMMAw.d
Attributes
Overall width:
12
Min: 1
Max: 32
Decimal places:
0
Min: 0
Max: 11
Description
SAS System chooses best notation
Example
Value:
123.1
Output:
1 2 3 . 1
OK
Cancel
Apply




**Summary Tables**

| Attitude | Max | Mean | Min | N | StdDev |
|---|---|---|---|---|---|
| Quality | 60 | 49.916666667 | 40 | 24 | 4.1169497644 |
| Supervision | 61 | 51.208333333 | 35 | 24 | 5.4929137486 |
| Work | 59 | 53.458333333 | 47 | 24 | 3.61132478 |

**Investigating a new Headache Drug**

The effectiveness of a new drug for the treatment of headaches is investigated in a clinical trial. The new drug is administered at two dosages: 5 milligrams and 10 milligrams (denoted by **N05** and **N10**). Female and male patients are randomly assigned to receive one of the two new dosages. After one month, the patients completed a quality of life (QOL) questionnaire. The QOL scores are on a 100-point scale. The standardized mean score is 50. The data is given in SAS Headache data set. We want to test the claim that the mean responses are the same for all populations associated with the treatments. Let $\alpha = 0.05$. We will use the Linear Model task to apply two-way factorial analysis of variance to this data.

This example involves four population means: $\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$

- $\mu_{11} = \mu + \alpha_1 + \beta_1 + (\alpha\beta)_{11}$ is the mean response for all female patients who take the new drug at 5-milligram dosage.
- $\mu_{12} = \mu + \alpha_1 + \beta_2 + (\alpha\beta)_{12}$ is the mean response for all male patients who take the new drug at 5-milligram dosage.
- $\mu_{21} = \mu + \alpha_2 + \beta_1 + (\alpha\beta)_{21}$ is the mean response for all male patients who take the new drug at 5-milligram dosage.
- $\mu_{22} = \mu + \alpha_2 + \beta_2 + (\alpha\beta)_{22}$ is the mean response for all male patients who take the new drug at 5-milligram dosage.

Treatment main effects are $\alpha_i$: the difference between the mean score of all patients who use the N05 or N10 dosage ($\mu_{i.}$) and the mean score of all patients who use either new dosage ($\mu$).

Sex main effects are $\beta_j$: the difference between the mean score of female or male who use the N05 or N10 dosage ($\mu_{.j}$) and the mean score of all patients who use either new dosage ($\mu$).

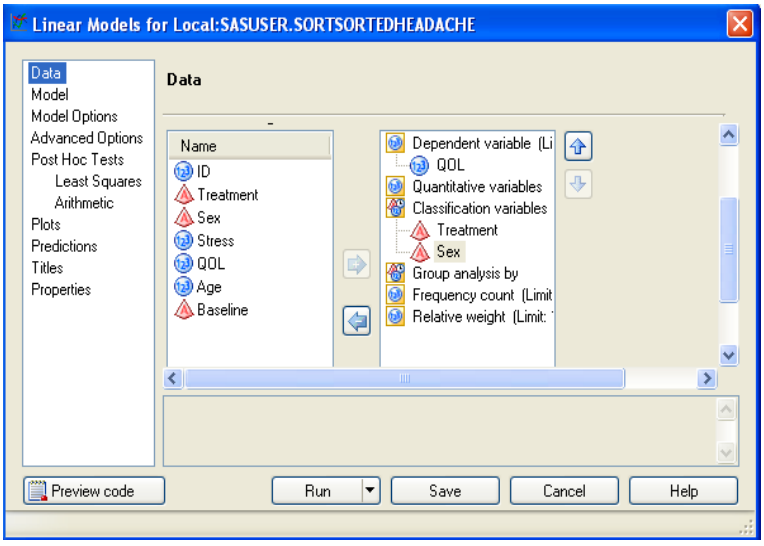
Treatment * Sex interaction effects are $(\alpha\beta)_{ij}$

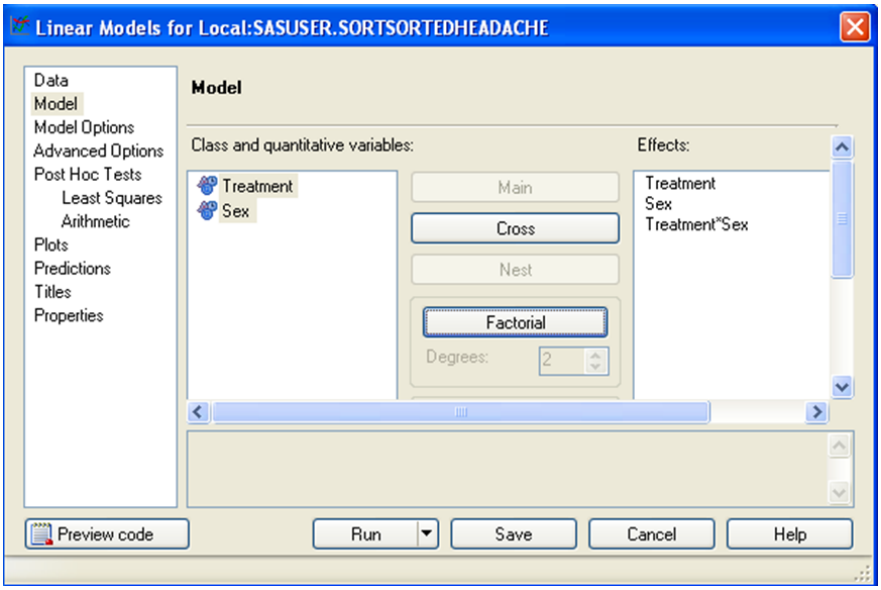
$H_0$: $\mu_{11} = \mu_{12} = \mu_{21} = \mu_{22}$
$H_1$: At least two means are not equal

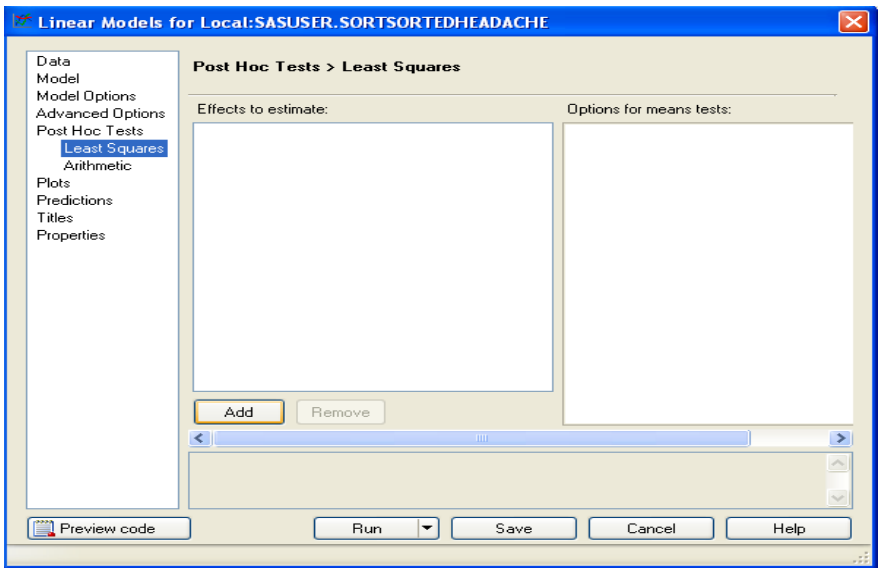
**Exercise #8: Two-Way Factorial ANOVA with Balanced data using Linear Models**

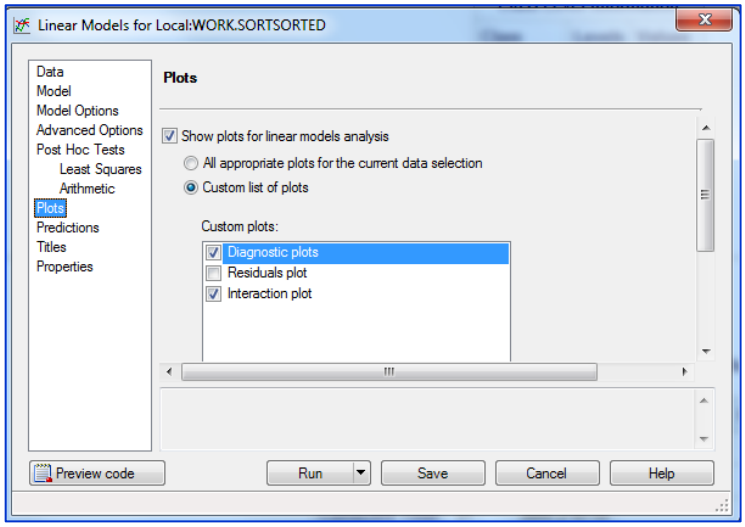
- File -> Open -> Data … -> Local Computer -> My Computer\ShortCourses Material\SAS and SPSS\SAS\SAS Enterprise Guide -> Sample-> Data Appendix -> headache.sas7bdat -> **Open**
- From **Data** menu, **Sort Data** by treatment-> click **Run**.
- **Delete** rows for other treatments (other than **N05**, and **N10**). ***Note:*** To delete unwanted rows: **Edit** > **Protect Data** > Delete rows, and then Protect Data again (only the variables **Treatment**, **sex**, and **QOL** are selected for this analysis).
- From the **Analyze** menu, select **ANOVA** -> **Linear Models …**
- Assign **QOL** to **Dependent Variable** Task role.

- Assign **Treatment** and **Sex** variables to **Classification** variables (**Sex** and **Treatment** main effects and a **Sex * Treatment** interaction effects).
- On the **Model** tab, **Ctrl** + Click both **Treatment** and **Sex** variables, and then click **Factorial** (the resulting model has a main effect for each independent variables and all possible interactions).
- On the **Post Hoc Tests**, click **Least Squares** (Least-Square means for comparisons on both main effects and interaction effects).
- Click the **Add** button (you may need to adjust the dialog box to see the Add button).
- On the **Options for means tests:**
    - At **Class effects to use:**, use the drop-down menu, to change **False** to **True** for each desired comparison (**Treatment** and **Sex** for example).
    - At **Comparisons:**
        - For **Show p-values for differences**, select **All pairwise differences**.
        - For **Adjustment method for comparison**, select a method computing p-values (**Tukey** for example).
- On **Plots** tab, click the **Custom List of Plots** radio button, and then check the **Interaction plot.**
- Check **Dependent means for main effects**, and **Dependent means for two-way effects.**
- On the **Titles** tab, clear the Footnote.
- When finished assigning variables to roles and selecting options, click **Run**.
- View the HTML output and the close this window.




Linear Models for Local:SASUSER.SORTSORTEDHEADACHE
Data
Model
Model Options
Advanced Options
Post Hoc Tests
Least Squares
Arithmetic
Plots
Predictions
Titles
Properties
Model
Class and quantitative variables:
Treatment
Sex
Main
Cross
Nest
Factorial
Degrees:
2
Effects:
Treatment
Sex
Treatment*Sex
Preview code
Run
Save
Cancel
Help


Linear Models for Local:SASUSER.SORTSORTEDHEADACHE
Data
Model
Model Options
Advanced Options
Post Hoc Tests
Least Squares
Arithmetic
Plots
Predictions
Titles
Properties
Post Hoc Tests > Least Squares
Effects to estimate:
Options for means tests:
Add
Remove
Preview code
Run
Save
Cancel
Help


Linear Models for Local:WORK.SORTSORTED
Data
Model
Model Options
Advanced Options
Post Hoc Tests
Least Squares
Arithmetic
Plots
Predictions
Titles
Properties
Plots
Show plots for linear models analysis
All appropriate plots for the current data selection
Custom list of plots
Custom plots:
Diagnostic plots
Residuals plot
Interaction plot
Preview code
Run
Save
Cancel
Help

## Linear Models

*The GLM Procedure*

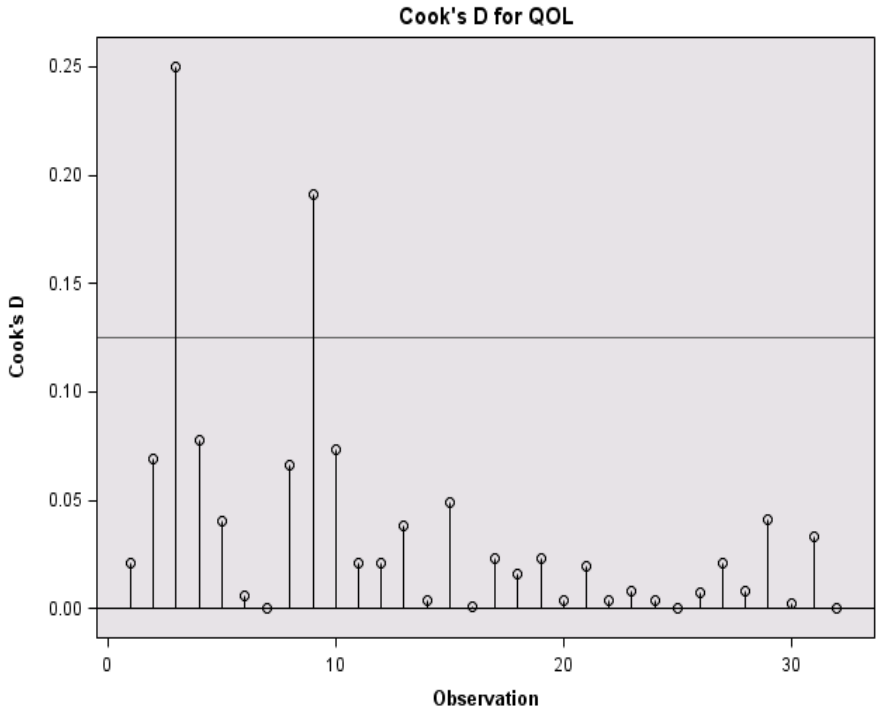
*Dependent Variable: QOL Quality of Life*

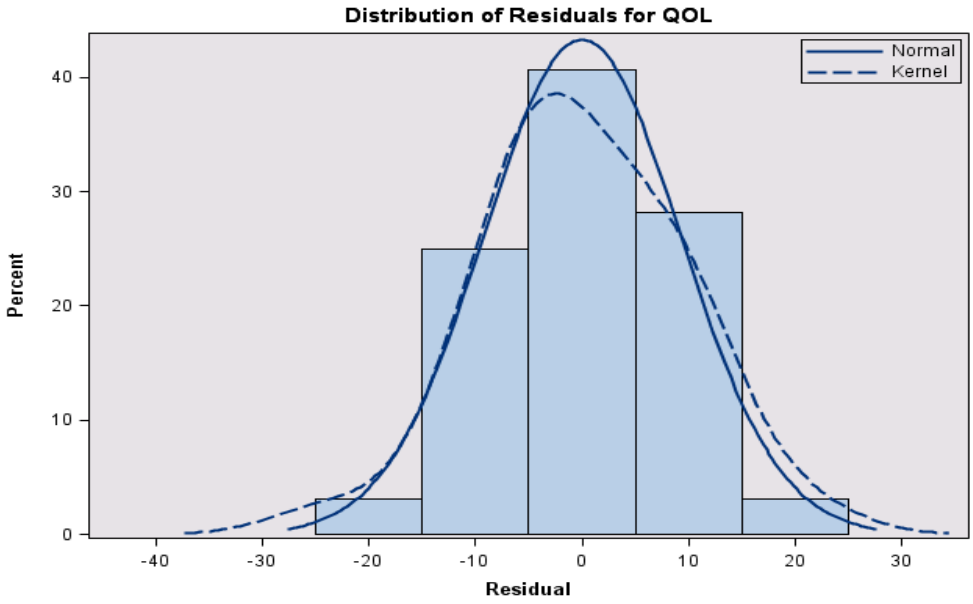
| Source | DF | Sum of Squares | Mean Square | F Value | Pr > |
|---|---|---|---|---|---|
| Model | 3 | 1206.843750 | 402.281250 | 4.28 | 0.013 |
| Error | 28 | 2633.375000 | 94.049107 | | |
| Corrected Total | 31 | 3840.218750 | | | |

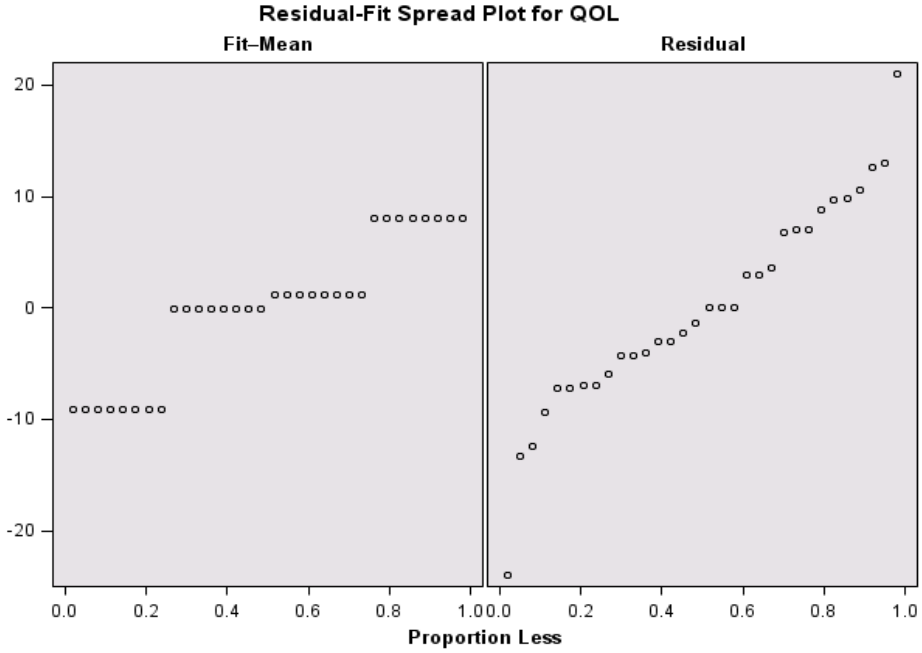
| R-Square | Coeff Var | Root MSE | QOL Mean |
|---|---|---|---|
| 0.314264 | 16.67558 | 9.697892 | 58.15625 |

| Source | DF | Type I SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Treatment | 1 | 504.0312500 | 504.0312500 | 5.36 | 0.0282 |
| Sex | 1 | 9.0312500 | 9.0312500 | 0.10 | 0.7589 |
| Treatment*Sex | 1 | 693.7812500 | 693.7812500 | 7.38 | 0.0112 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Treatment | 1 | 504.0312500 | 504.0312500 | 5.36 | 0.0282 |
| Sex | 1 | 9.0312500 | 9.0312500 | 0.10 | 0.7589 |
| Treatment*Sex | 1 | 693.7812500 | 693.7812500 | 7.38 | 0.0112 |






## Results

- Independent variables are: Treatment with a=2 and levels N05, N10, Sex with b=2 and levels F and M
- Dependent variable is QOL
- N..=32

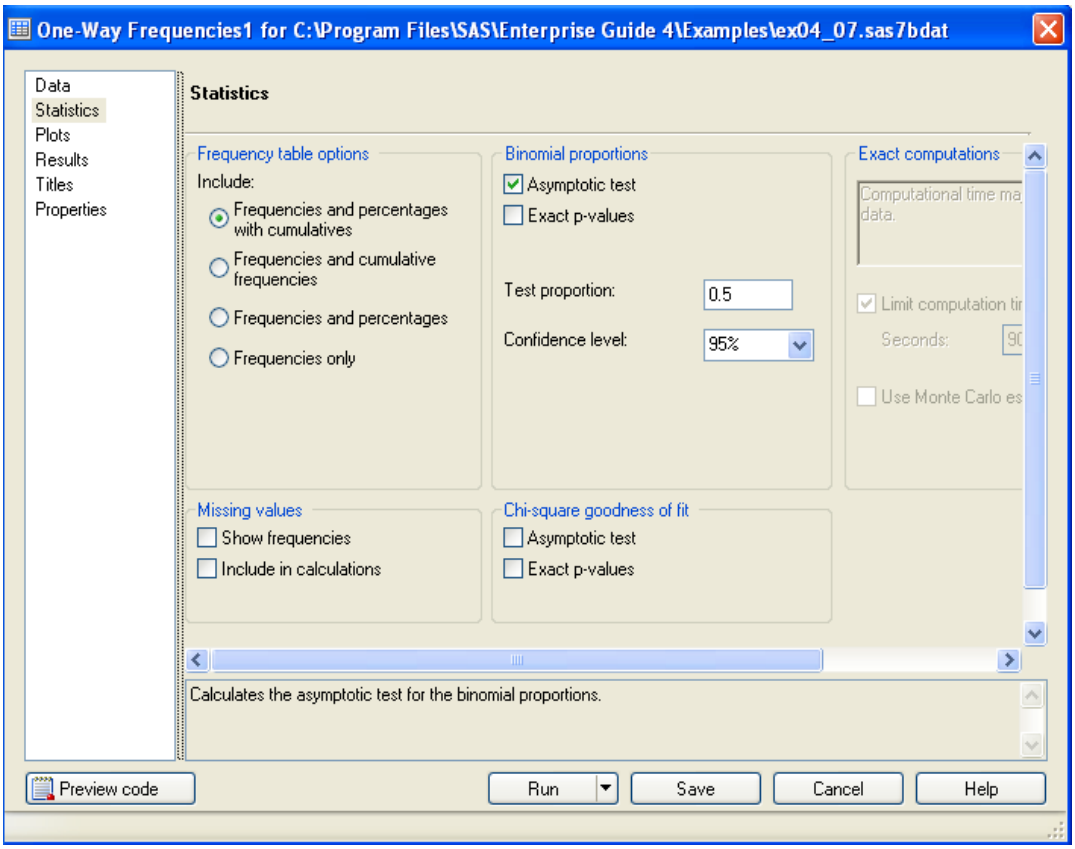
- Coefficient of determination $R^2$ = 0.314264
- Degrees of freedom = n-k =24 -3 = 21
- P-value = 0.0132 < 0.05, reject $H_0$. There is sufficient evidence to reject the claim that the mean responses are the same for all populations associated with the treatments. At least two treatment-sample means are significantly different.

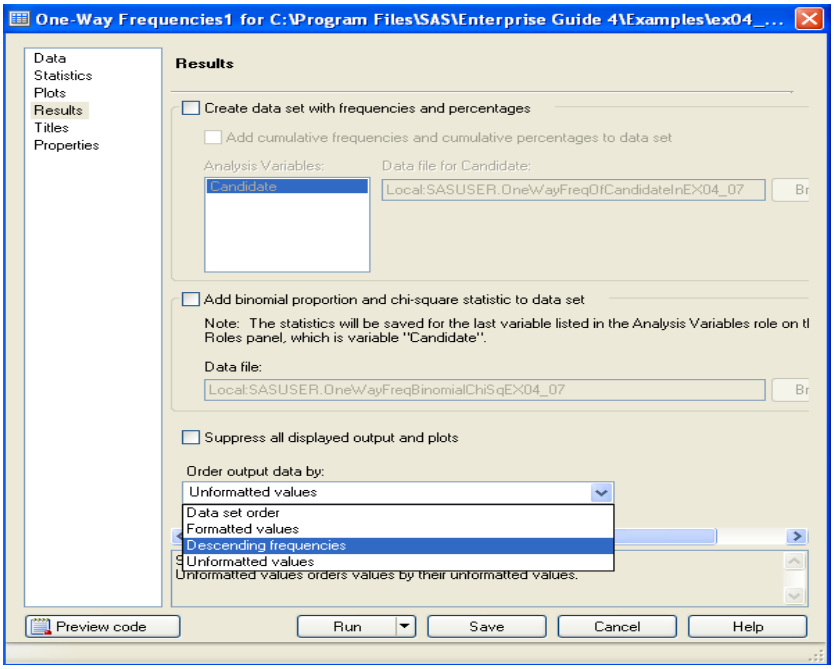
**Exercise #9: Making Inference on a Population Proportion, using One-Way Frequencies**

A poll of 423 likely voters examines Candidates Support a week before an election. There are three candidates in the election. The responses are summarized in table EX04_07. Test the claim that Smith has support among a majority of likely voters. Let $\alpha$ = 0.05

- File -> Open -> Data … -> Local Computer -> My Computer\ShortCourses Material\SAS and SPSS\SAS\SAS Enterprise Guide -> Examples -> EX04_07.sas7bdat -> Open
- From the **Describe** menu, select **One-Way Frequencies…**
- On the **Data**, assign Candidate to Analysis variables role, and Support to Frequency count role
- On the **Statistics** tab, on the Binomial Proportions, check the Asymptotic test (used when $n\hat{p}$ and $n\hat{p}(1-\hat{p}) \geq 5$). That is, the condition is true when n is large and/or sample proportion $\hat{p} = \frac{x}{n}$ is not too close to 0.0 or to 1.0.
- At Confidence Level, enter 95%
- On the **Results** tab, From Order Output data by drop down list, select Descending Frequencies (so that Smith is first in the list).
- On the **Titles** tab, clear the Footnote.
- When finished assigning variables to roles and selecting options, click **Run**.
- View the HTML output and the close this window.

| | Candidate | Support |
|---|---|---|
| 1 | Dodd | 58 |
| 2 | Lee | 121 |
| 3 | Smith | 227 |
| 4 | Undecided | 17 |


One-Way Frequencies1 for C:\Program Files\SAS\Enterprise Guide 4\Examples\ex04_07.sas7bdat
Data
Statistics
Plots
Results
Titles
Properties
Statistics
Frequency table options
Include:
Frequencies and percentages with cumulatives
Frequencies and cumulative frequencies
Frequencies and percentages
Frequencies only
Binomial proportions
Asymptotic test
Exact p-values
Test proportion:
0.5
Confidence level:
95%
Exact computations
Computational time ma
data.
Limit computation ti
Seconds:
Use Monte Carlo es
Missing values
Show frequencies
Include in calculations
Chi-square goodness of fit
Asymptotic test
Exact p-values
Calculates the asymptotic test for the binomial proportions.
Preview code
Run
Save
Cancel
Help


One-Way Frequencies1 for C:\Program Files\SAS\Enterprise Guide 4\Examples\ex04_...
Data
Statistics
Plots
Results
Titles
Properties
Results
Create data set with frequencies and percentages
Add cumulative frequencies and cumulative percentages to data set
Analysis Variables:
Candidate
Data file for Candidate:
Local:SASUSER.OneWayFreqOfCandidateInEX04_07
Add binomial proportion and chi-square statistic to data set
Note: The statistics will be saved for the last variable listed in the Analysis Variables role on tl
Roles panel, which is variable "Candidate".
Data file:
Local:SASUSER.OneWayFreqBinomialChiSqEX04_07
Suppress all displayed output and plots
Order output data by:
Unformatted values
Data set order
Formatted values
Descending frequencies
Unformatted values
Unformatted values orders values by their unformatted values.
Preview code
Run
Save
Cancel
Help

## One-Way Frequencies
## Results

*The FREQ Procedure*

| Candidate | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| Smith | 227 | 53.66 | 227 | 53.66 |
| Lee | 121 | 28.61 | 348 | 82.27 |
| Dodd | 58 | 13.71 | 406 | 95.98 |
| Undecided | 17 | 4.02 | 423 | 100.00 |

| Binomial Proportion for Candidate = Smith | |
|---|---|
| Proportion | 0.5366 |
| ASE | 0.0242 |
| 95% Lower Conf Limit | 0.4891 |
| 95% Upper Conf Limit | 0.5842 |
| | |
| Exact Conf Limits | |
| 95% Lower Conf Limit | 0.4878 |
| 95% Upper Conf Limit | 0.5849 |

| Test of H0: Proportion = 0.5 | |
|---|---|
| ASE under H0 | 0.0243 |
| Z | 1.5073 |
| One-sided Pr > Z | 0.0659 |
| Two-sided Pr > \|Z\| | 0.1317 |

*Sample Size = 423*

**Analysis:**

$H_0 : p \leq 0.5$

$H_1 : p > 0.5$

P-value = 0.0659 (see the output)

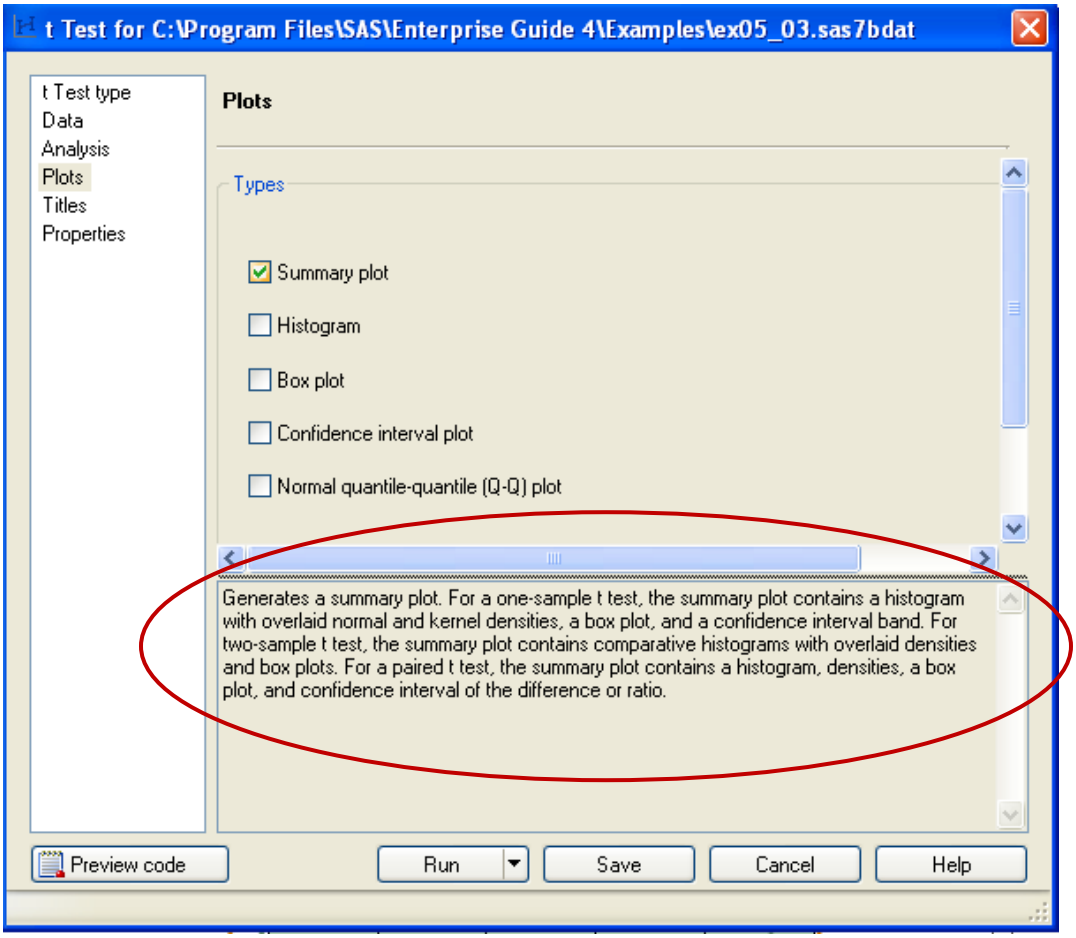
Since the p-value ≥ 0.05, do not reject $H_0$. There is not sufficient evidence to support the claim that Smith has a majority of support among all likely voters. That is, the sample proportion is not significantly greater than 0.5. ***Note:*** p-value is a measure of the likelihood that the sample comes from the process where $H_0$ is true. That is, the more the data agrees with $H_1$ – the smaller the p-value.

**Exercise #10: Make inference on mean of a population of differences (paired t-test)**

A university assesses the quality of its education program by testing each student's intellectual skills before their education course, and the after completion of the program. The results from 22 randomly selected students are included in **EX05_03** data set.

- Using **EX05_03** data set, from the **Analyze** menu, select ANOVA -> **t Test…**
- On the **t Test type**, select the **Paired**
- On the **Data**, assign **Pre** and **Post** variables to the **Paired variables** role
- On the **Plots** tab, check the **Summary plot** and **Box plot** types.
- On the **Titles** tab, clear the Footnote.
- When finished assigning variables to roles and selecting options, click **Run**.
- View the HTML output and the close this window.

| | Pre | Post |
|---|---|---|
| 1 | 14 | 30 |
| 2 | 24 | 29 |
| 3 | 19 | 40 |
| 4 | 25 | 42 |
| 5 | 20 | 24 |
| 6 | 21 | 38 |
| 7 | 15 | 46 |
| 8 | 16 | 42 |
| 9 | 38 | 75 |
| 10 | 21 | 34 |
| 11 | 16 | 31 |
| 12 | 23 | 38 |
| 13 | 15 | 28 |
| 14 | 25 | 63 |
| 15 | 15 | 16 |
| 16 | 16 | 44 |
| 17 | 13 | 35 |
| 18 | 37 | 70 |
| 19 | 8 | 9 |
| 20 | 42 | 51 |
| 21 | 17 | 40 |
| 22 | 28 | 43 |




## t Test

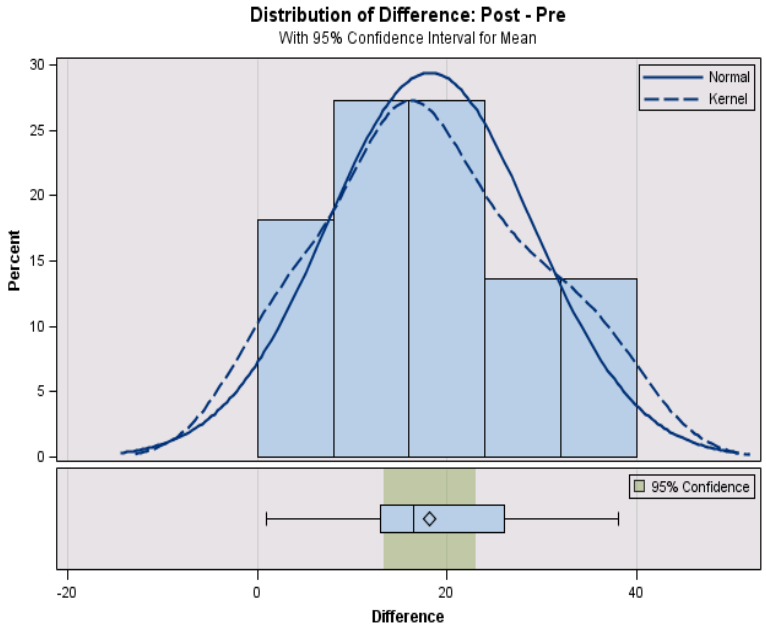
*The TTEST Procedure*

*Difference: Post - Pre*

| N | Mean | Std Dev | Std Err | Minimum | Maximum |
|---|---|---|---|---|---|
| 22 | 18.1818 | 10.8568 | 2.3147 | 1.0000 | 38.0000 |

| Mean | 95% CL Mean | | Std Dev | 95% CL Std Dev | |
|---|---|---|---|---|---|
| 18.1818 | 13.3682 | 22.9955 | 10.8568 | 8.3527 | 15.5151 |

| DF | t Value | Pr > \|t\| |
|---|---|---|
| 21 | 7.86 | <.0001 |

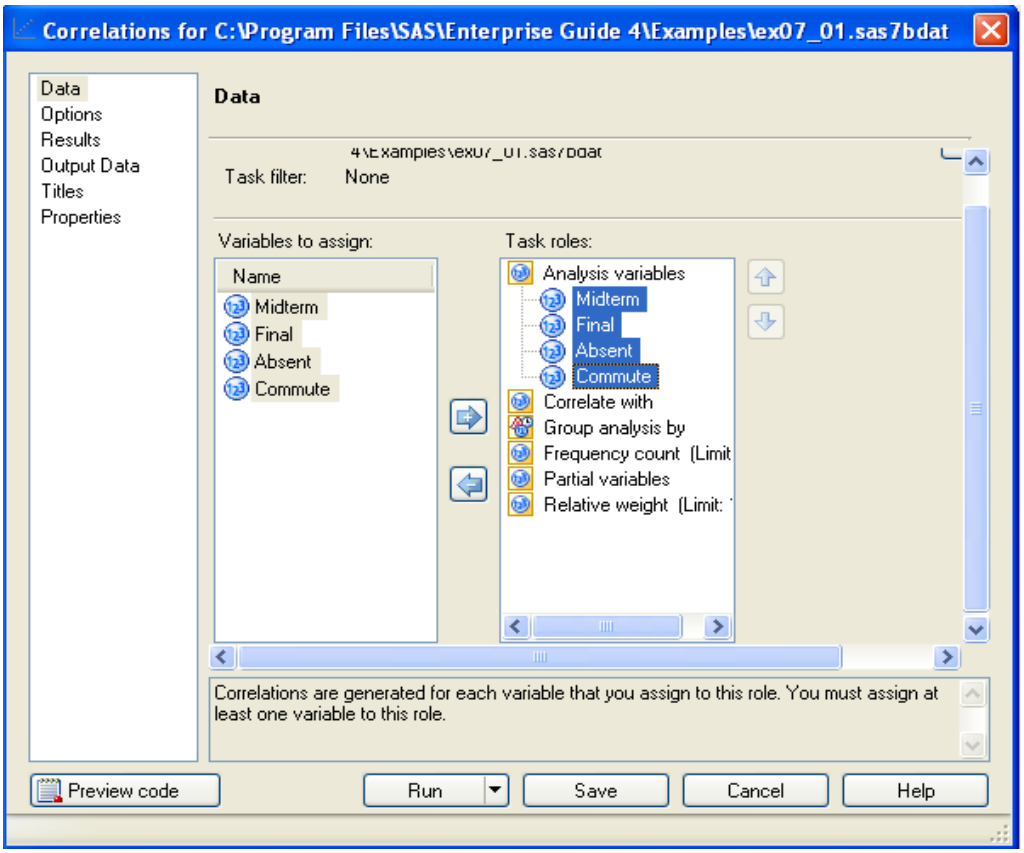



**Exercise #11: Calculating Pearson Correlations Coefficient**

The Pearson correlation coefficient, represented by r, is a measure of strength of the linear relationship between two numeric variables. We will use Ex07_01 SAS data set to determine the Pearson correlation coefficient for pairs of measurements in this data set.

- File -> Open -> Data … -> Local Computer -> My Computer\ShortCourses Material\SAS and SPSS\SAS\SAS Enterprise Guide 4 -> Examples -> EX07_01.sas7bdat -> Open
- Analyze > Multivariate > Correlations
- On the **Data** tab, Assign all the variables to Analysis Variables roles
- On the **Results** tab, check Create scatter plot for each correlation pair
- On the **Titles** tab, clear the Footnote.
- When finished assigning variables to roles and selecting options, click **Run**.
- View the HTML output and the close this window.

| | Midterm | Final | Absent | Commute |
|---|---|---|---|---|
| 1 | 64 | 63 | 5 | 10 |
| 2 | 64 | 68 | 5 | 20 |
| 3 | 66 | 68 | 4 | 5 |
| 4 | 68 | 72 | 3 | 50 |
| 5 | 72 | 69 | 5 | 5 |
| 6 | 73 | 80 | 3 | 10 |
| 7 | 78 | 74 | 2 | 10 |
| 8 | 81 | 85 | 3 | 20 |
| 9 | 91 | 96 | 2 | 25 |
| 10 | 92 | 90 | 3 | 60 |




## Correlation Analysis

*The CORR Procedure*

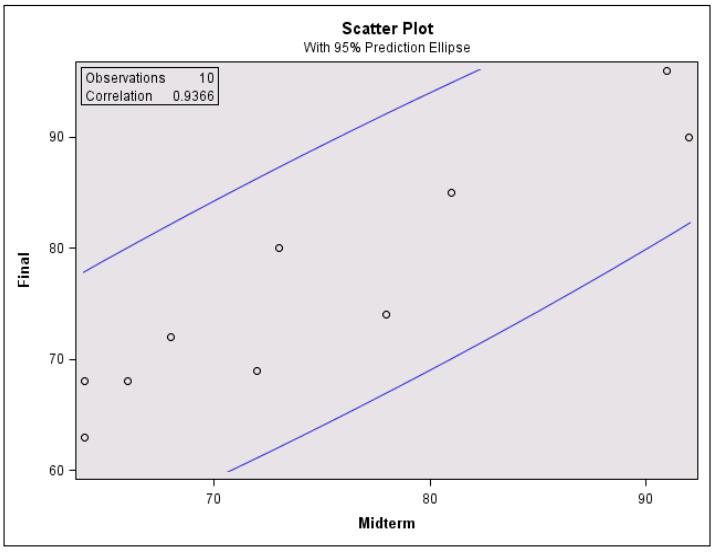
**4 Variables:** Midterm Final Absent Commute

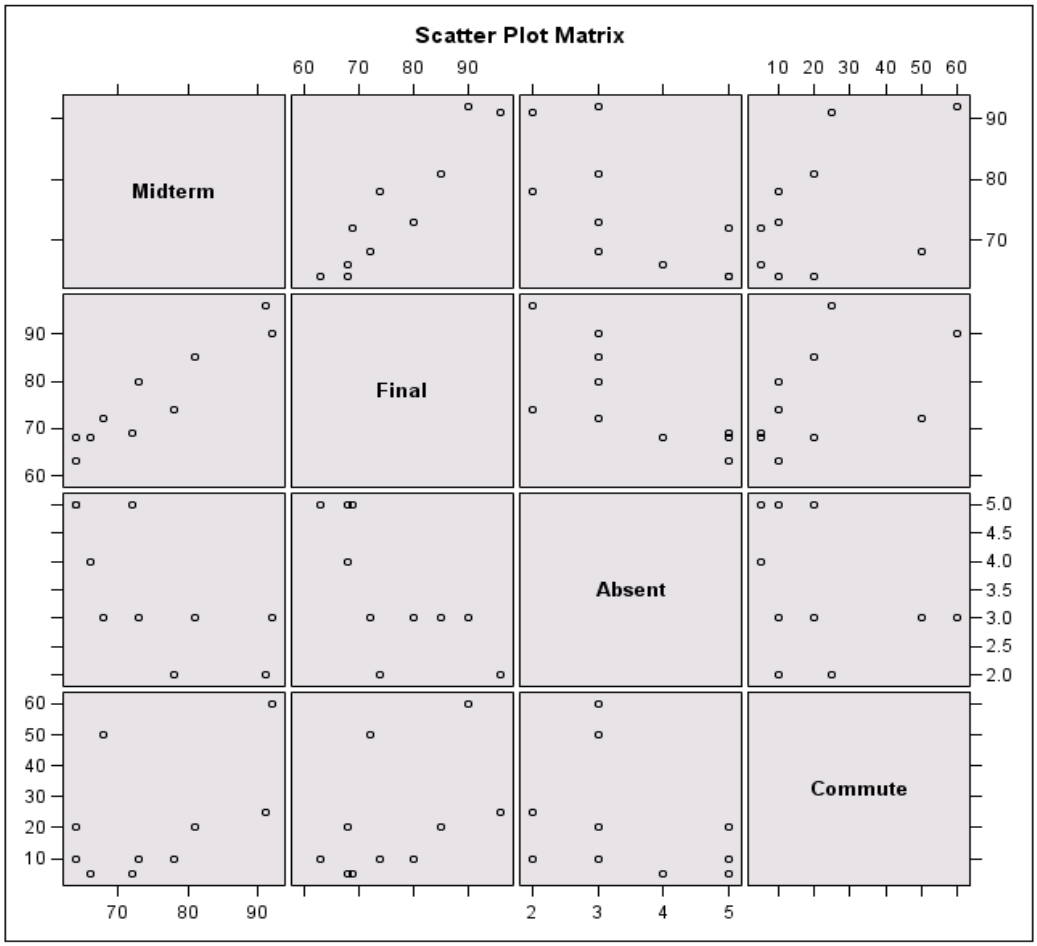
| Simple Statistics | | | | | | |
|---|---|---|---|---|---|---|
| Variable | N | Mean | Std Dev | Sum | Minimum | Maximum |
| Midterm | 10 | 74.90000 | 10.40780 | 749.00000 | 64.00000 | 92.00000 |
| Final | 10 | 76.50000 | 10.83462 | 765.00000 | 63.00000 | 96.00000 |
| Absent | 10 | 3.50000 | 1.17851 | 35.00000 | 2.00000 | 5.00000 |
| Commute | 10 | 21.50000 | 19.01023 | 215.00000 | 5.00000 | 60.00000 |

| Pearson Correlation Coefficients, N = 10<br>Prob > \|r\| under H0: Rho=0 | | | | |
|---|---|---|---|---|
| | Midterm | Final | Absent | Commute |
| Midterm | 1.00000 | 0.93656<br><.0001 | -0.70205<br>0.0236 | 0.46695<br>0.1736 |
| Final | 0.93656<br><.0001 | 1.00000 | -0.73530<br>0.0154 | 0.46798<br>0.1726 |
| Absent | -0.70205<br>0.0236 | -0.73530<br>0.0154 | 1.00000 | -0.33477<br>0.3444 |
| Commute | 0.46695<br>0.1736 | 0.46798<br>0.1726 | -0.33477<br>0.3444 | 1.00000 |

## Correlation Analysis

*The CORR Procedure*



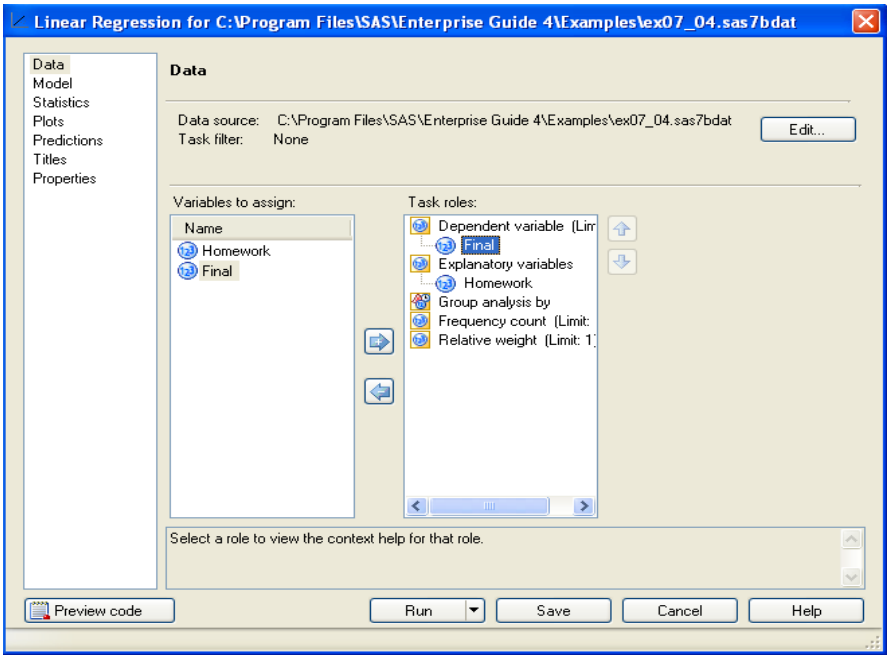
**Results**

- The top number is the Pearson correlation coefficient r.
- The second number is the p-value for the test of
  - $H_0: \rho = 0$
- n=10 for each pair of variables
- For **Midterm** and final r=0.93656 and the p-value < 0.0001
- For **Final** and **Absent**, r=-0.73530
- For Final and Commute, r=0.46798 and the p-value =0.1726

**Exercise #12: Simple Linear Regression**

- File -> Open -> Data … -> Local Computer -> My Computer\ShortCourses Material\SAS and SPSS\SAS\SAS Enterprise Guide 4 -> Examples ->ex07_04.sas7bdat -> Open
- **Analyze** > **Regression** > **Linear…**
- Assign the **Final** variable to **Dependent** variables Task role.
- Assign the **Homework** variable to **Explanatory** variables Task role.
- On **Statistics** tab, check the **Confidence limits for parameter estimates**.
- On the **Plots** tab, select **Observed vs Independent** and select the radio button for **Confidence limits**.
- On **Predictions** tab, select **Original sample**, **Prediction limits**, and **Show predictions** check boxes.
- On the **Titles** tab, clear the Footnote.
- When finished assigning variables to roles and selecting options, click **Run**.
- View the HTML output and the close this window.

| | Homework | Final |
|---|---|---|
| 1 | 47 | 63 |
| 2 | 59 | 68 |
| 3 | 47 | 68 |
| 4 | 62 | 72 |
| 5 | 51 | 69 |
| 6 | 70 | 80 |
| 7 | 54 | 74 |
| 8 | 63 | 85 |
| 9 | 78 | 96 |
| 10 | 76 | 90 |




## Linear Regression Results

*The REG Procedure*
*Model: Linear_Regression_Model*
*Dependent Variable: Final*

| | |
|---|---|
| Number of Observations Read | 10 |
| Number of Observations Used | 10 |

**Analysis of Variance**

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 1 | 866.20073 | 866.20073 | 36.41 | 0.0003 |
| Error | 8 | 190.29927 | 23.78741 | | |
| Corrected Total | 9 | 1056.50000 | | | |

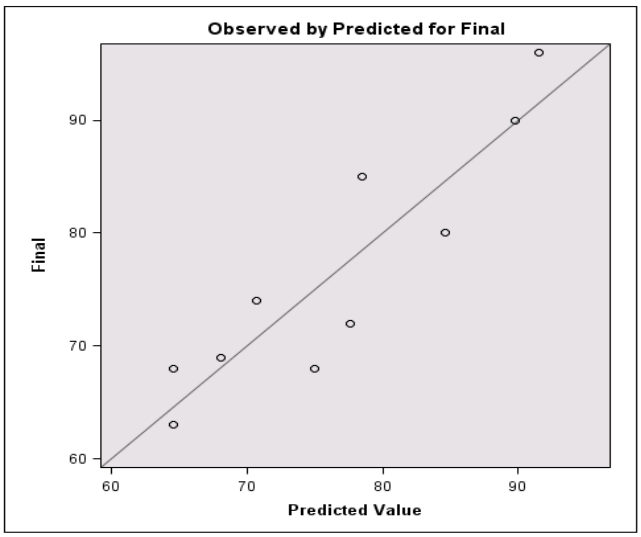
| | | | |
|---|---|---|---|
| Root MSE | 4.87723 | R-Square | 0.8199 |
| Dependent Mean | 76.50000 | Adj R-Sq | 0.7974 |
| Coeff Var | 6.37547 | | |

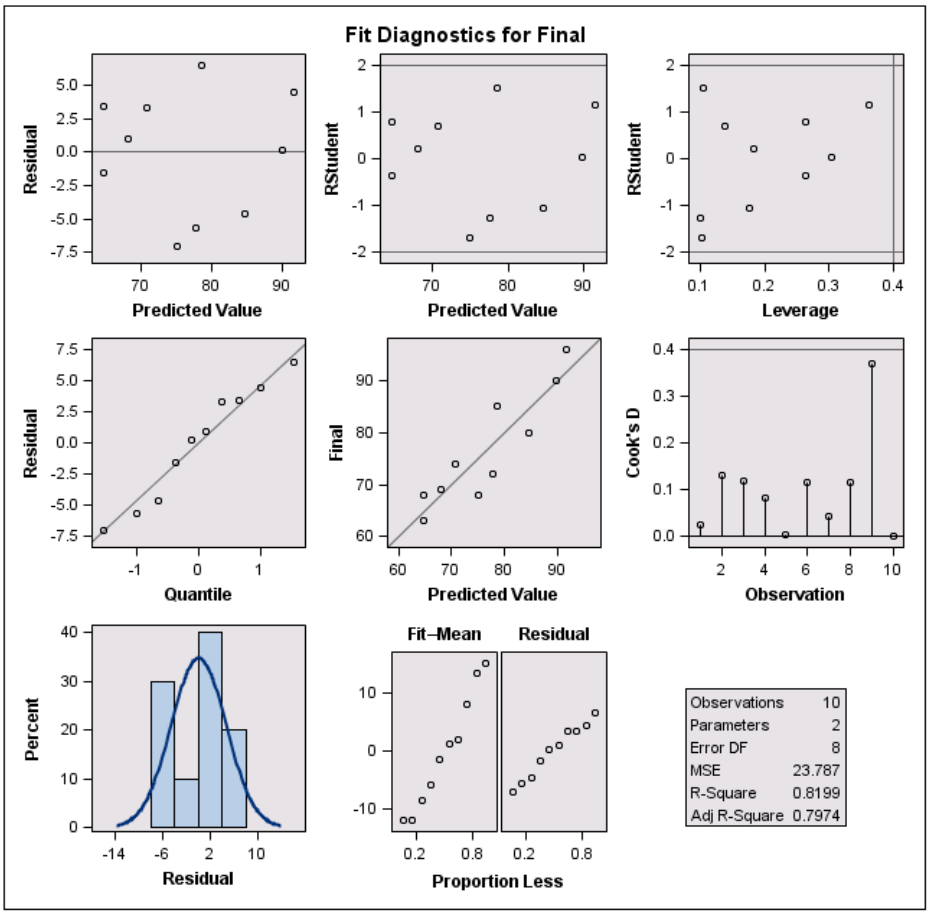
**Parameter Estimates**

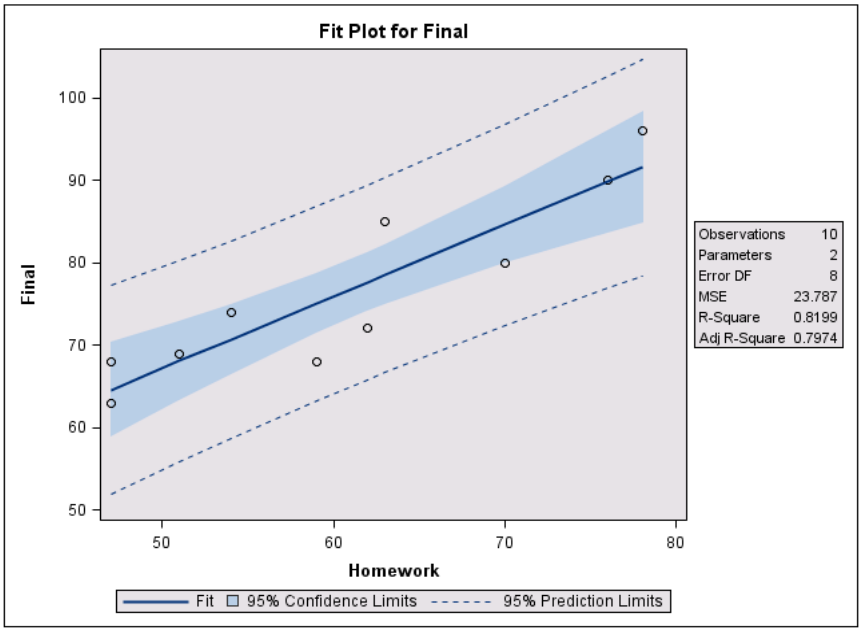
| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > \|t\| | 95% Confidence Limits | |
|---|---|---|---|---|---|---|---|
| Intercept | 1 | 23.68394 | 8.88731 | 2.66 | 0.0286 | 3.18976 | 44.17813 |
| Homework | 1 | 0.87012 | 0.14419 | 6.03 | 0.0003 | 0.53761 | 1.20262 |

## Linear Regression Results

*The REG Procedure*
*Model: Linear_Regression_Model*
*Dependent Variable: Final*

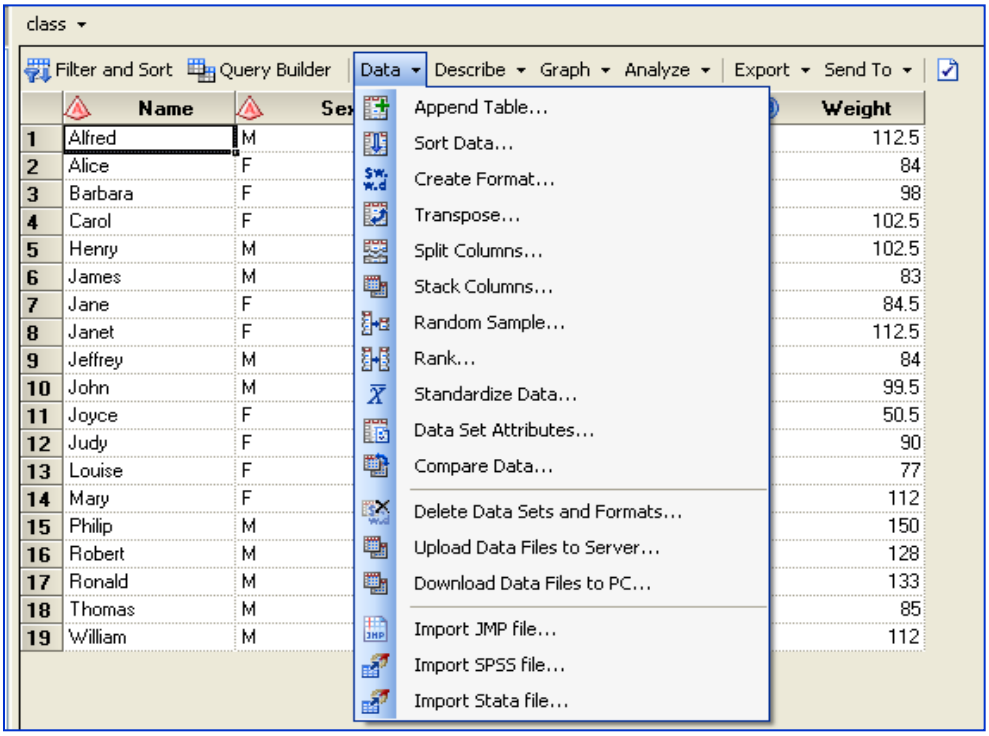


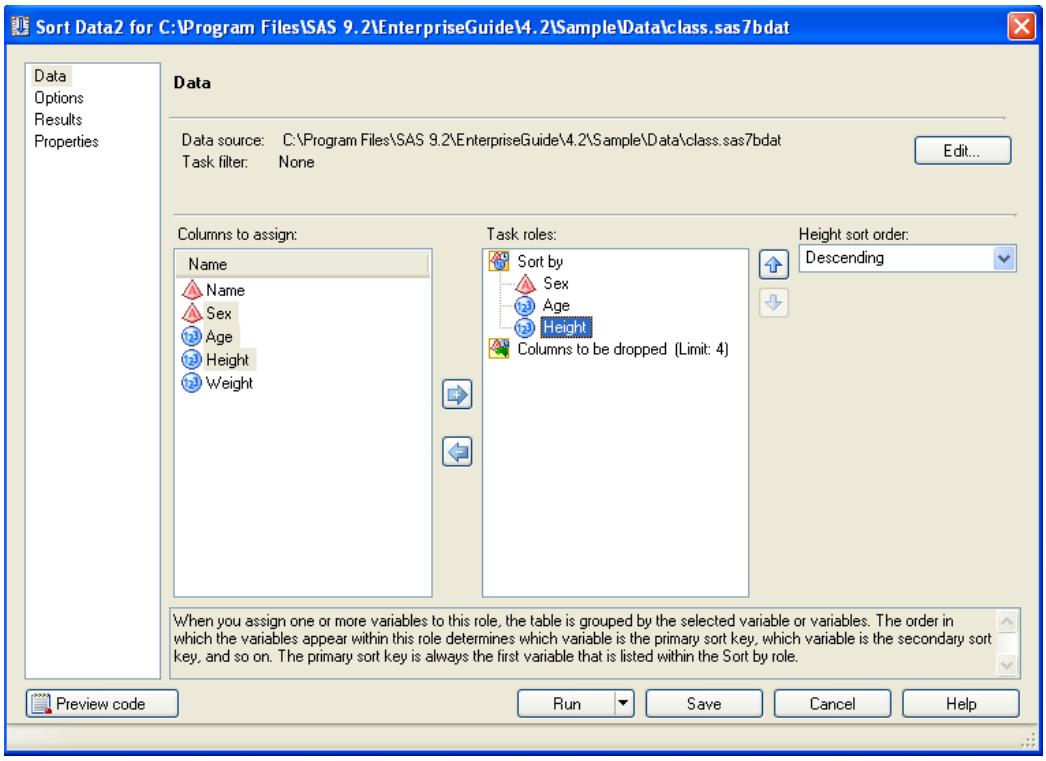






**Exercise #13: Finding the largest value in a BY-Group Using the SAS Class data set**

- **File -> Open -> data … -> Servers -> Local -> SASUSER (or SASHELP)> Class > Open**
- From the **data** menu, select **Sort data …**
- Move the SEX, AGE, and HEIGHT variables over from Columns To Assign to Data and put them under the Sort by role
- Select HEIGHT, and change the Height sort order from Ascending to Descending. This will ensure that the largest height is in the first observation for each age group
- Run the Sort task.
- Close the data set.
- Select the sorted data set and bring up the Sort data task again (data > Sort data …).
- This time, sort by **SEX** and **AGE**. Under **Options**, go to **Duplicate Records** and select **keep only the first record for each 'Sort by'** group.
- On the **Titles** tab, clear the Footnote.
- When finished assigning variables to roles and selecting options, click **Run**.
- View the HTML output and the close this window.


class
Filter and Sort
Query Builder
Data
Describe
Graph
Analyze
Export
Send To
Name
Sex
Weight
1 Alfred M 112.5
2 Alice F 84
3 Barbara F 98
4 Carol F 102.5
5 Henry M 102.5
6 James M 83
7 Jane F 84.5
8 Janet F 112.5
9 Jeffrey M 84
10 John M 99.5
11 Joyce F 50.5
12 Judy F 90
13 Louise F 77
14 Mary F 112
15 Philip M 150
16 Robert M 128
17 Ronald M 133
18 Thomas M 85
19 William M 112
Append Table...
Sort Data...
Create Format...
Transpose...
Split Columns...
Stack Columns...
Random Sample...
Rank...
Standardize Data...
Data Set Attributes...
Compare Data...
Delete Data Sets and Formats...
Upload Data Files to Server...
Download Data Files to PC...
Import JMP file...
Import SPSS file...
Import Stata file...


Sort Data2 for C:\Program Files\SAS 9.2\EnterpriseGuide\4.2\Sample\Data\class.sas7bdat
Data
Options
Results
Properties
Data
Data source: C:\Program Files\SAS 9.2\EnterpriseGuide\4.2\Sample\Data\class.sas7bdat
Task filter: None
Edit...
Columns to assign:
Name
Name
Sex
Age
Height
Weight
Task roles:
Sort by
Sex
Age
Height
Columns to be dropped (Limit: 4)
Height sort order:
Descending
When you assign one or more variables to this role, the table is grouped by the selected variable or variables. The order in which the variables appear within this role determines which variable is the primary sort key, which variable is the secondary sort key, and so on. The primary sort key is always the first variable that is listed within the Sort by role.
Preview code
Run
Save
Cancel
Help

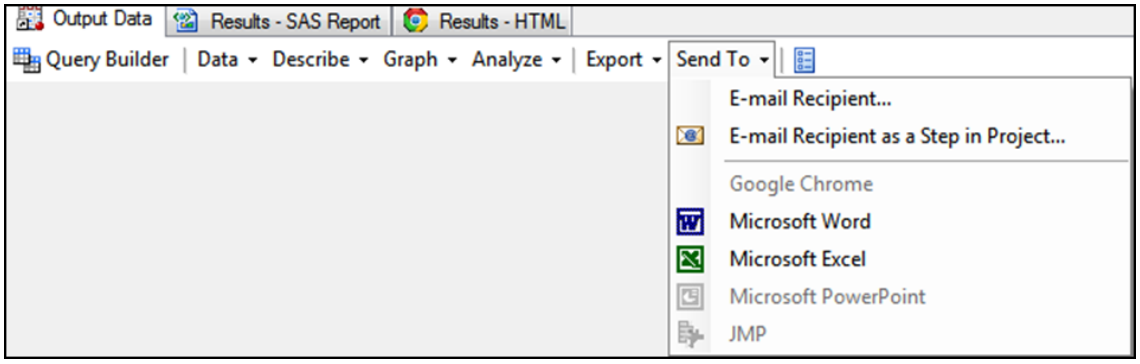
| | Name | Sex | Age | Height | Weight |
|---|---|---|---|---|---|
| 1 | Joyce | F | 11 | 51.3 | 50.5 |
| 2 | Jane | F | 12 | 59.8 | 84.5 |
| 3 | Barbara | F | 13 | 65.3 | 98 |
| 4 | Judy | F | 14 | 64.3 | 90 |
| 5 | Mary | F | 15 | 66.5 | 112 |
| 6 | Thomas | M | 11 | 57.5 | 85 |
| 7 | Robert | M | 12 | 64.8 | 128 |
| 8 | Jeffrey | M | 13 | 62.5 | 84 |
| 9 | Alfred | M | 14 | 69 | 112.5 |
| 10 | Ronald | M | 15 | 67 | 133 |
| 11 | Philip | M | 16 | 72 | 150 |

**Saving a project**

- **File** -> **Save 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.

**To send data to Microsoft Word, or Excel**

- You can send a copy of a file in a SAS Enterprise Guide project to another user through electronic mail. Files that you can send include data, programs, logs, results, and notes.
- You can also send a copy of the active file to Microsoft **Word**, **Excel**, or **PowerPoint**. ***Note:*** In order to send results in the **SAS Report** format to any Microsoft Office application, you must have the **SAS Add-In for Microsoft Office** installed on your computer.
- In the project tree, process flow or workspace, select the file that you want to send to another user.
    - Select **File** -> **Send To E-mail Recipient**.
    - Select **File** -> **Send To** -> **Microsoft Word (or Excel)**. The active data source is inserted as a table into a new Microsoft Word document.
- Or- on the **Output Data** tab, click Send To, and then
    - **Word**
    - **Excel**




**Online Resources**

- **Statistics Using SAS Enterprise Guide** by James B. Davis, 2007. Preview this book at http://books.google.com
- **Basic Statistics Using SAS Enterprise Guide**, a Primer, by Geoff Der, and Brian S. Everitt, 2007. Preview this book at http://books.google.com
- **The Little SAS® Book: a primer 5th edition,** by: Susan Slaughter and Lora Delwiche, Feb. 11, 2013 available from TTU Safari e-Books http://library.ttu.edu
- **SAS Online Resources** for Statistics Education http://support.sas.com/statlibrary
- **SAS Programming for Enterprise Guide Users**, by Neil, 2007 available from TTU Safari e-Books.

**Where to Get Help**

- If you need help from me, please e-mail heide.mansouri@ttu.edu, or call 834-2935 to make an appointment.



Please e-mail your comments or suggestions to: heide.mansouri@ttu.edu