TECHNOLOGY SUPPORT

SAS STUDIO

SHORTCOURSE HANDOUT

Texas Tech University | Heide Mansouri

TEXAS TECH UNIVERSITY
Information Technology Division
Technology Support

Table of Contents

SAS Studio

ShortCourse Handout

Adapted from:

- SAS Studio User's Guide (SAS Studio 3.1 with SAS 9.4), as of September 15, 2017 http://support.sas.com/documentation/onlinedoc/sasstudio/ accessed on 11/30/2017
- Performing Statistical Analysis http://support.sas.com/documentation/cdl/en/webeditorug/66932/HTML/default/viewer.htm#p1s681zspnfwagn1hpfix3h9d51y.htm , accessed on 12/6/2017

Introduction

SAS Studio is a module in SAS. SAS Studio is different because it is a tool that you can use to write and run **SAS** code through your web browser. Two free SAS programs: **SAS University Edition** and **SAS OnDemand** also run SAS Studio. With SAS Studio, you can access your data files, libraries, and existing SAS programs and write new SAS programs.

SAS Studio was first released in March 2014, and is included in installs of SAS version 9.4. It can run locally, with a URL that points to your local SAS install, or users can access a server by nearly any web browser from any operating system by using their server credentials.

When you use SAS Studio, you are also using SAS software behind the scenes. SAS Studio connects to a SAS server in order to process SAS commands. In the SAS University Edition, the SAS server is included in the **vApp** (Virtual Application) program that is installed on your local machine. After the code is processed by the SAS server, the results are returned to SAS Studio in your browser. SAS Studio supports multiple web browsers, such as **Microsoft Internet Explorer**, **Apple Safari**, **Mozilla Firefox**, and **Google Chrome**.

When you run a program or task, SAS Studio connects to a SAS server to process the SAS code. The SAS server can be a hosted server in a cloud environment, a server in your local environment, or a copy of SAS on your local machine. After the code is processed, the results are returned to SAS Studio in your browser.

Predefined tasks are included with SAS Studio to analyze your data. The tasks are based on SAS System procedures and provide access to some of the most commonly used graph and analytical procedures.

This ShortCourse assumes that users are familiar with **Elementary Statistics**.

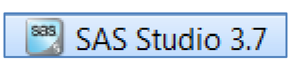
Course Objectives

After completing this ShortCourse, you should be able to:

- Create a SAS data set;
- Perform some basic data and file management tasks;
- Write, Submit and store a SAS program;
- Import an Excel file into a SAS data set.
- Generating Descriptive Statistics for Continuous Variables
- Performing One-Sample Tests
- Performing Paired t Test
- Performing a One-Way Analysis of Variance

Starting SAS Studio

- From the **Start** menu, click **All Programs -> SAS -> SAS Studio 3.7**

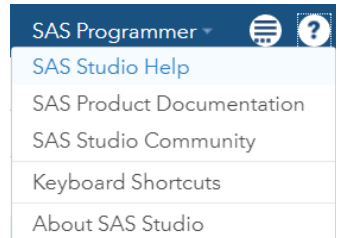


About Using SAS Studio

When you sign on to SAS Studio, the main SAS Studio window appears with a blank program window so that you can start programming immediately. You also have access to all five sections of the navigation pane.

Note: To sign out of SAS Studio, click **Sign Out** on the toolbar. **Do not use the Back button on your web browser.**

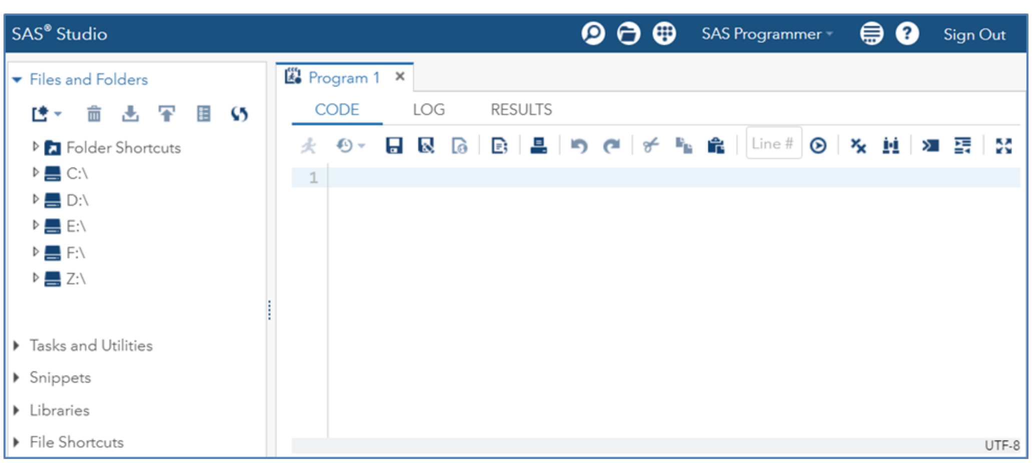
Accessing SAS Studio Help



- Pop-up help is available for some options in the application, which you can access by clicking next to the option.
- Comprehensive help for SAS Studio is available by clicking above the work area and selecting SAS Studio Help.

SAS Studio Interface

- The main window of SAS Studio consists of a **Navigation pane** on the left and a **Work area** on the right.
- The navigation pane provides access to files on your server and your folder shortcuts, your **tasks** and **snippets** (small blocks of reusable codes), the **libraries** that you have access to, and your file shortcuts.
- A task is an **XML code file** that generates SAS code and formats results for you. Tasks include SAS procedures from simple data listings to complex analytical procedures.
- The **Server Files** and **Folders** section is displayed by default. When you click on any of the choices in the Navigation pane, it expands and moves higher in the list.
- The work area is used to display your **data**, **code**, **logs**, and **results**. When you first open SAS Studio, the work area displays a new program window. When you open data and run tasks, other windows open in the work area with a tabbed interface.
- Comprehensive help for SAS Studio is available by clicking above the work area and selecting **SAS Studio Help**.



Exploring the libraries Tab

- When you click the **Libraries** to expand it, under **My Libraries**, you see a list of libraries. SAS Studio comes with some libraries already installed.
- The **Work Library** is a temporary library.
- All datasets and programs placed in Work library will be deleted, when you exit SAS Studio.

Rearranging the Tabs in the Work Area

- In the work area, you can rearrange the tabs by using a **drag-and-drop** operation to move them to the left or right.
- You can also dock a tab on the right side or bottom of the work area to view more than one tab at a time.
- To rearrange a tab:
  - Select the tab that you want to move.

Exploring the Tasks and Utilities Tab

- When you expand the Tasks and Utilities tab, you see two tabs: **Tasks**, and **Utilities**.
- Expanding the Utilities will display three sub-tabs: **Import Data**, **Query**, and **SAS Program**.

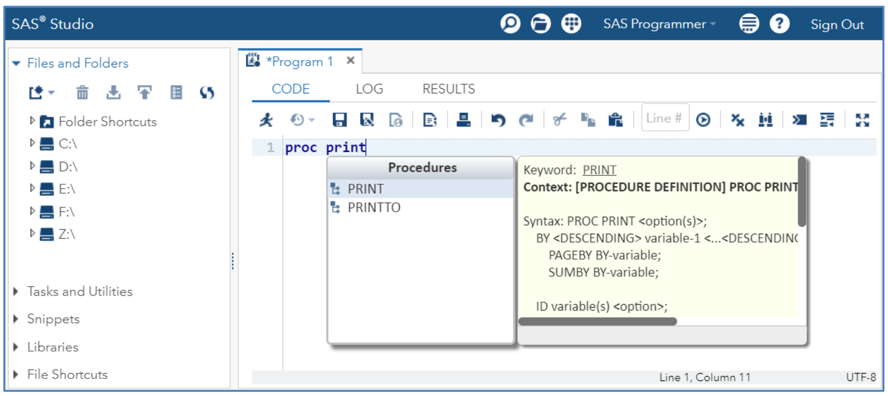
How to Run a Task

- In the navigation pane, click the **Tasks and Utilities** section.
- Expand the folder that contains the task.
- **Right-click** the task name and select Open. Alternatively, you can **double-click** the task to open it.
- The task opens to the right of the work area.
- If the **Data** tab is available, specify an input data source and select columns for the roles in the data source. A role is a description of a variable's purpose in the task. To add a column to a role, click . A list of available columns for that role appears. If only one column can be assigned to the role, you select a column and the list disappears. If multiple columns can be assigned, you can press Ctrl or Shift to select multiple columns from the list and click OK.

- On the remaining tabs, specify any other required options, which are denoted with a red asterisk. As you assign values to the task, the relevant SAS code is generated.
- To run the task, click .
- If the task generates output data, the table opens in the **Output Data** tab.

Writing a SAS Program

SAS Studio has several features to help you reduce your programming time, including **autocompletion** for hundreds of SAS statements and procedures as well as built-in syntax help that includes links to the more extensive SAS Product Documentation.



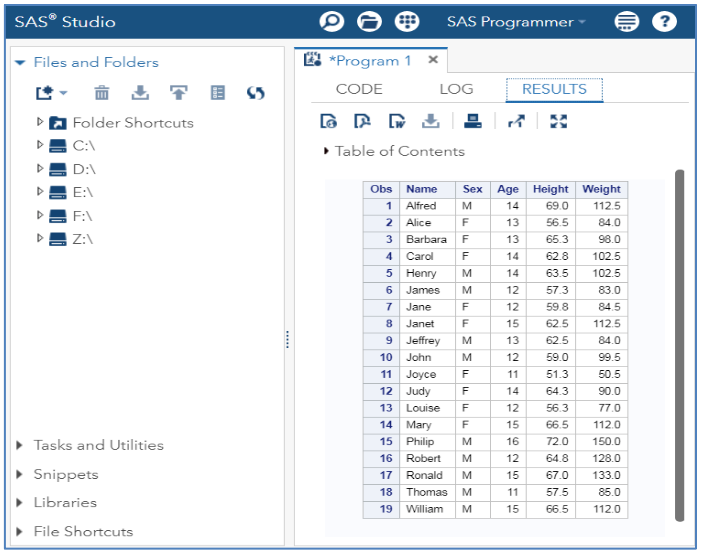
- Open SAS Studio and enter the following code in the Program 1 window that is created automatically for you:

```
proc print data=sashelp.class;
run;
```

  Notice that each SAS statement must end with a semicolon.

  Note: If you need to create a new program window, open the Server Files and Folders section of the navigation pane. Then click and select **SAS Program**.

- To run the code, click on the toolbar.



Some Basic Rules of SAS Programs

- Every SAS statement ends with a **semicolon** (**;**).
- You can place several statements on a single line if you end each statement with a semicolon. (***Note***: This practice is not recommended—it makes programs hard to read.)
- SAS names are a **maximum of 32 characters** in length. They must begin with a letter or underscore (_). The remaining characters in a SAS name are letters, numbers, or underscores.
- Spaces ***are not*** allowed in any SAS name.

Saving Your Results

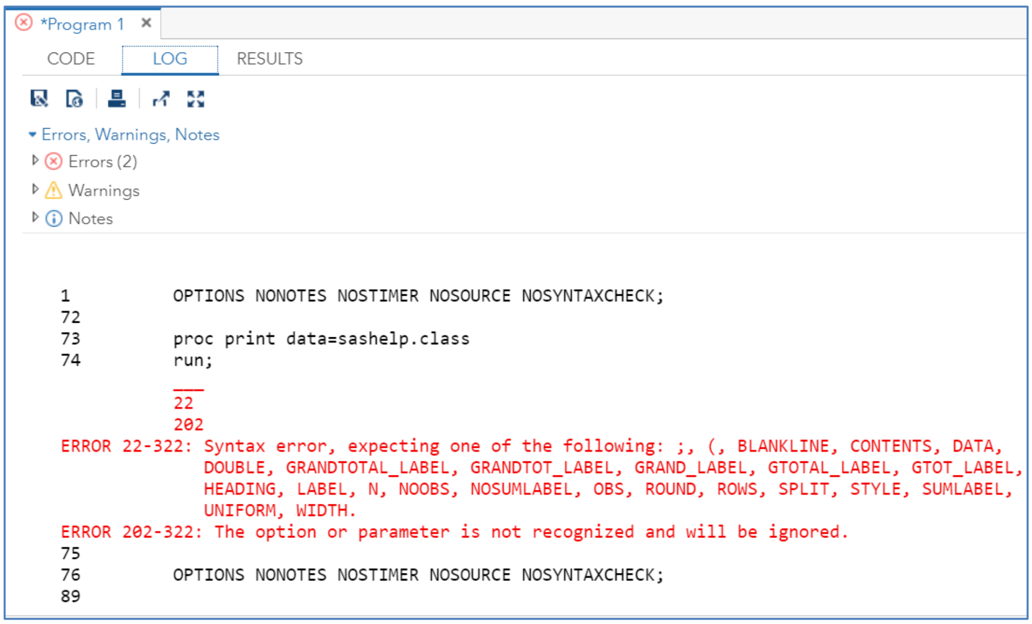
- By clicking a button on the **Results** toolbar, you can download your results to any of three different formats and save them or open them in the default application for that format:
  - HTML file
  - PDF file
  - RTF file
- You can also print your results from SAS Studio by clicking . The results open in a separate browser window, and you can then use the default printer controls for that browser.

SAS Program Errors

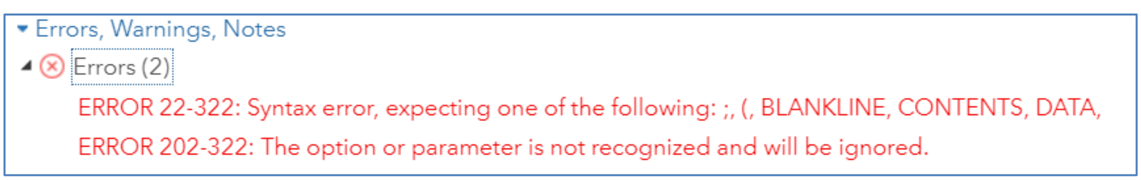
- Let’s add an error to the original program and see how it works.
- Click the **Code** tab to open your program. Delete the semicolon from the first line of the program.

```
proc print data=sashelp.class
run;
```

- Click to run the program. This time, the **Log** tab opens automatically to show you that you have an error.



- In the Errors, Warnings, Notes section, expand **Errors** to view a description of the error. Click the error message, and SAS Studio highlights it for you in the log so that you can see exactly where the error occurred.



- You can return to your program and correct the error. However, if you have a very long and complicated program that has a lot of errors, you might want to return to an earlier version of your program in which you knew all the code was correct.
- SAS Studio maintains a log, or submission history, with entries for each time you run a program, so you can easily return to an earlier version of a program.
- To find an earlier version of your program, click the **Code** tab to view the current version of your program. On the toolbar, click  and then click the first version of the program.

- The original version of your program opens in a new window from which you can copy and paste the error-free code into your original program or into a new program.

Add Column Names to Your Program

- SAS Studio adds code for the dragged items to your program for you.
- To see how this works, let's go back to the original program that you started with:

```
proc print data=sashelp.class;
run;
```

- Next, add the VAR statement to the program to specify which variables, or columns, to include in the results. After the first line of code, add the following new line of code:

```
var
```

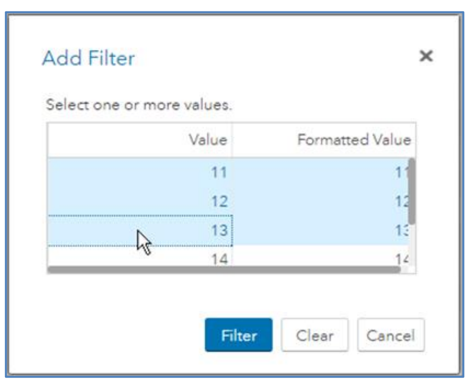
- Now you can use the Libraries section to help complete the VAR statement. Click the **Libraries** section in the navigation pane and expand the Sashelp library. Locate the Class table and expand it to view the columns.
- Hold down the Ctrl key and select the **Name**, **Age**, and **Height** columns and then drag them to the end of the VAR statement in your program. A green check mark icon indicates where you can drop the selected columns.
- When you drop the selected columns, SAS Studio adds the column names to your program.

- The SAS programming language requires that each statement end with a semicolon. To avoid another error when you run your program, you must add a semicolon to the end of the VAR statement.
- Click RUN.

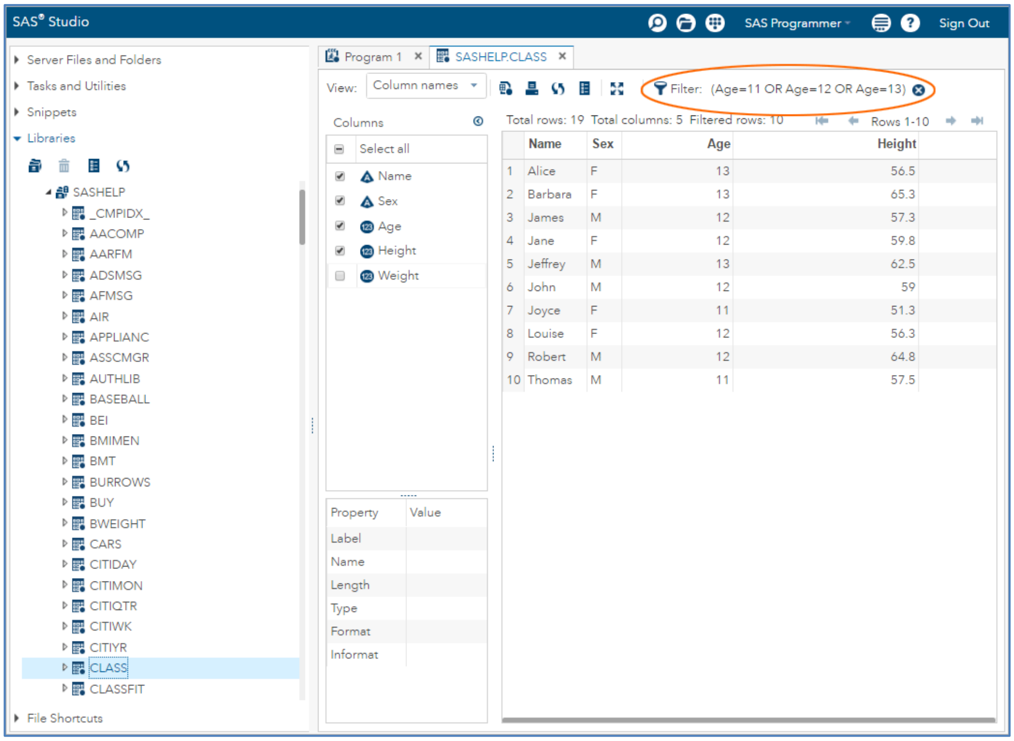
```
proc print data=sashelp.class;
var Age Height Name;
run;
```

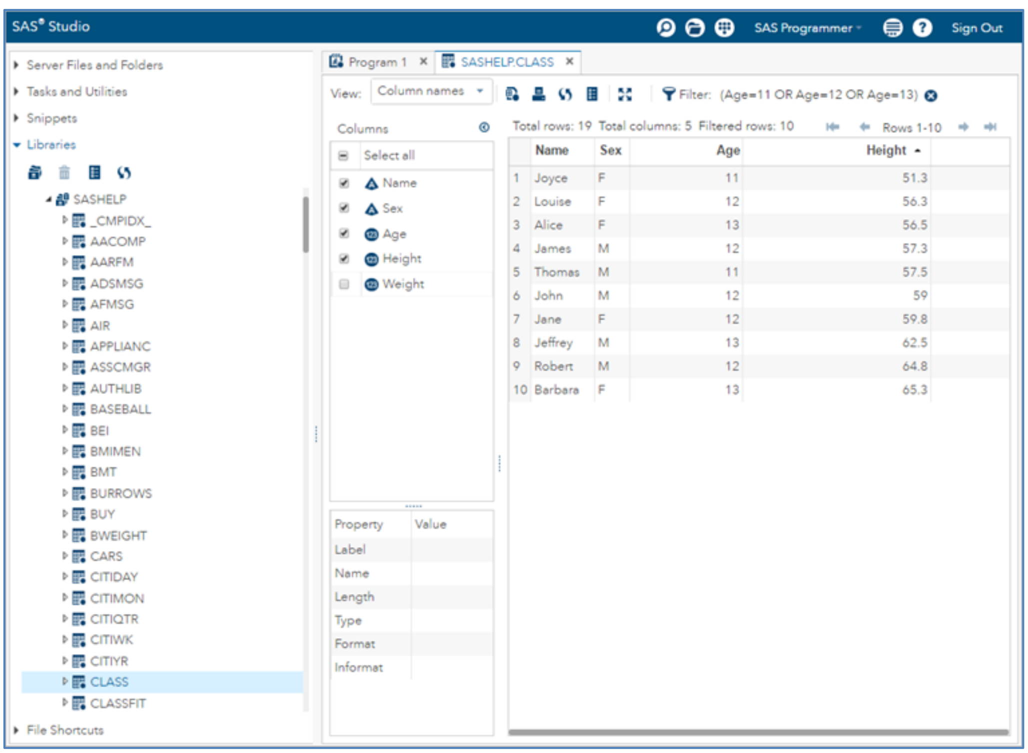
Let SAS Studio Do the Programming for You!

- From the Libraries section, double-click the **Class** table to open it in the table viewer (**Libraries** -> **My Libraries** -> **SASHELP** -> **CLASS**).
- In the **Columns** area of the table viewer, all of the columns are selected by default. Clear the **Weight** column and notice that it is immediately removed from the table viewer.
- To filter and sort the data.
    - Right-click the **Age** column heading, and select **Add Filter**. From the list of column values, hold down the Ctrl key and select three values: **11, 12, 13**.
    - Click **Filter**. The table viewer is updated and now displays only the rows for ages 11, 12, and 13. ***Note:*** The filter criteria are displayed at the top of the table viewer. You can click to edit the filter and to delete the filter (Clear the Filter – Apply).

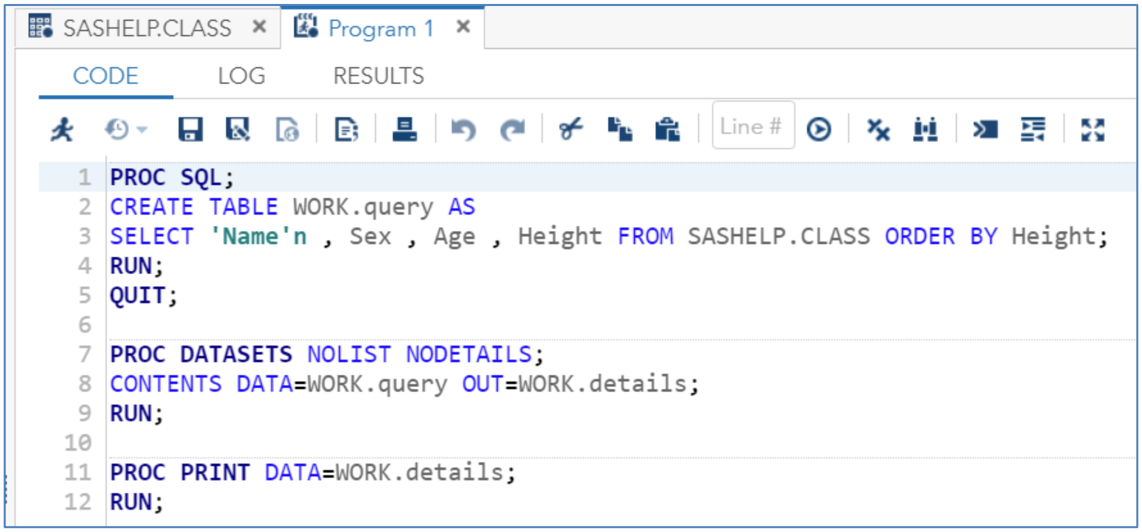


- Right-click the **Height** column heading and select **Sort Ascending**. The table is sorted by the values of the Height column from smallest to largest.





- To view the code, click on the toolbar. A new program window appears with the code that was used to create the view of the table in the table viewer.



- This program is a copy of the code that SAS Studio created and is no longer associated with the table viewer.
- Editing this program does not affect the data that is displayed in the table viewer, and modifying the table viewer does not affect the contents of this code. You can edit this code, or use it as the basis for another program.

Working with Snippets

- The **Snippets** section of the navigation pane enables you to access your code snippets.
- ***Code snippets*** are samples of commonly used SAS code that you can insert into your SAS program.
- SAS Studio is shipped with several predefined code snippets that you can use.

Working with Libraries

Icon	Type of Column
	Character
	Numeric
	Date
	Datetime

- The **Libraries** section of the navigation pane enables you to access your SAS libraries.
- SAS tables are stored in SAS libraries.
- From the **Libraries** section, you can open SAS tables and add them to your programs.
- You can use the **Libraries** section to expand a table and view the columns in that table. The icon in front of the column name indicates the type.
- You can drag tables and columns from the **Libraries** section to a program, and SAS Studio adds code for the dragged items to your program.

Note: The Sasuser library is read only, as in any SAS server environment. You cannot save content to this library.

To create a new library:

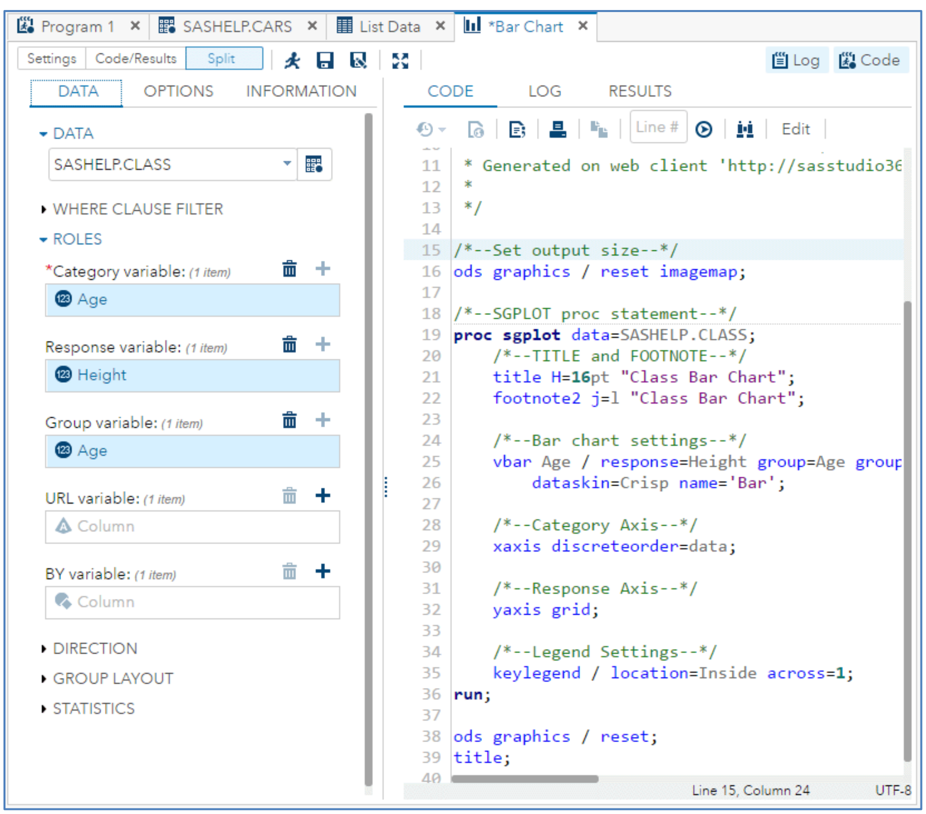
- Click **Libraries** in the navigation pane and then click .
- In the **Name** box, enter the libref for the library. The libref must be **eight characters or fewer.**
- In the **Path** box, enter the physical path where the library resides or click **Browse** to select a location.
- To create a new folder when you are selecting a location, click .
- In the **Options** box, specify any configuration options that you need. For the appropriate options, see the documentation for your operating environment.
- If you want to access this library each time you use SAS Studio, select **Re-create this library at start-up**.
- Click **OK** to create the library. The new library is added to the list of libraries in the navigation pane.
- In the SAS Studio Mid-Tier (the enterprise edition) deployment, you can assign unassigned metadata libraries by clicking . The libraries that you can assign must already be defined in your metadata. If you want to access the selected libraries each time you use SAS Studio, select **Assign selected libraries at start-up**. If a library is unassigned, then you cannot access the tables in that library.

Customizing the Navigation Pane

- By default, all five sections of the navigation pane are displayed when you open SAS Studio in the SAS Programmer perspective.
- To customize which sections are displayed, click  and select **View**.
- Select or clear any sections that you want to add or remove. The navigation pane is updated immediately. ***Note:*** The **File Shortcuts** section is not displayed by default in the Visual Programmer perspective.

Using the Work Area

- The work area is the main portion of the SAS Studio application for accessing programs and tasks and for viewing data.
- The work area is always displayed and cannot be minimized.
- When you open a **program**, **task**, or **table**, the windows open as new tabs in the work area.
- The **code**, **log**, and **results** that are associated with programs and tasks are grouped together under the main tab for the program or task.

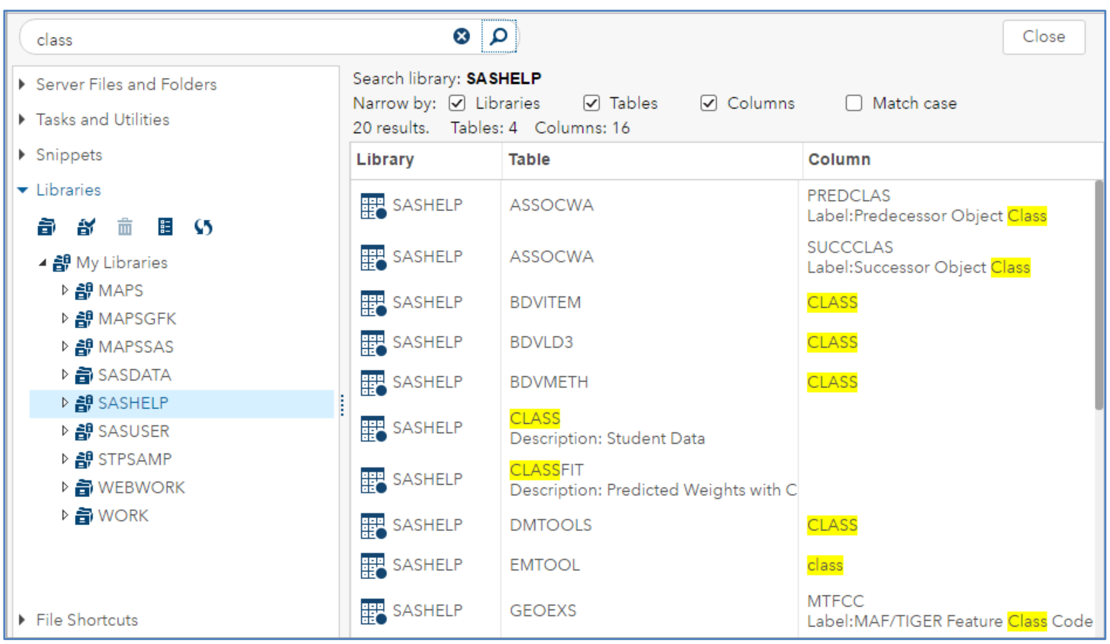


Customizing the Work Area

- By default, the work area is displayed beside the navigation pane, but you can maximize the work area and hide the navigation pane.
- You can also close all of the tabs in the work area at once.
- To maximize the work area, click on the toolbar. ***Note:*** To reopen the navigation pane, click again.
- To close all tabs that are open in the work area, click and select **Close All Tabs**. You are prompted to save any unsaved programs or tasks.

Searching in SAS Studio

- To access the Search feature, click . The search box appears so that you can enter the text that you want to search for.
- You can limit the scope of your search by selecting or clearing any of the **Narrow by** options. These options vary depending on what is selected in the navigation pane. By default, the search is case sensitive. To search for both uppercase and lowercase text, clear the **Match case** check box.
- The following example shows a search for “class” in the Sashelp library. The search includes all tables and columns in the Sashelp library, and it is not case sensitive.



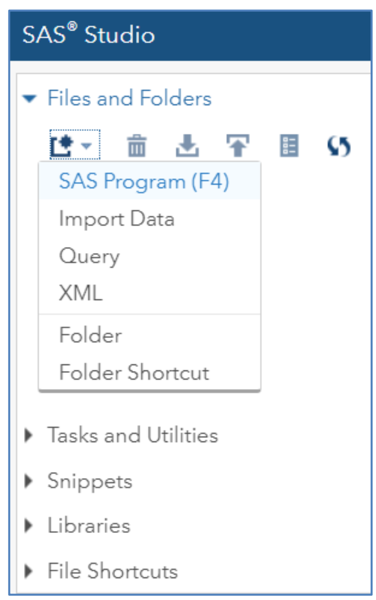
- You can open an item from your results by double-clicking it. To return to your search results list, click again.
- To clear the search results, click .

Opening a Program

- You can open SAS programs from the **Server Files and Folders** section of the navigation pane.
- To open a program, expand the appropriate folder and double-click the program that you want to open, or drag it into the work area. The program opens on a new tab in the work area. ***Note:*** Opening very large program files can affect your performance.
- If you open a program file that is greater than ***10 MB***, you are prompted to confirm whether you want to continue opening the file.
- If you are using Internet Explorer, you are prompted for confirmation if you open a program file that is greater than 3 MB.

Creating a Program

- You can create a SAS program from the **Server Files and Folders** section of the navigation pane. To create a program, click and select **SAS Program**. ***Note:*** You can also click on the main application toolbar and select **New SAS Program**.
- You can save the program, by clicking on the Save As button , and follow the steps.



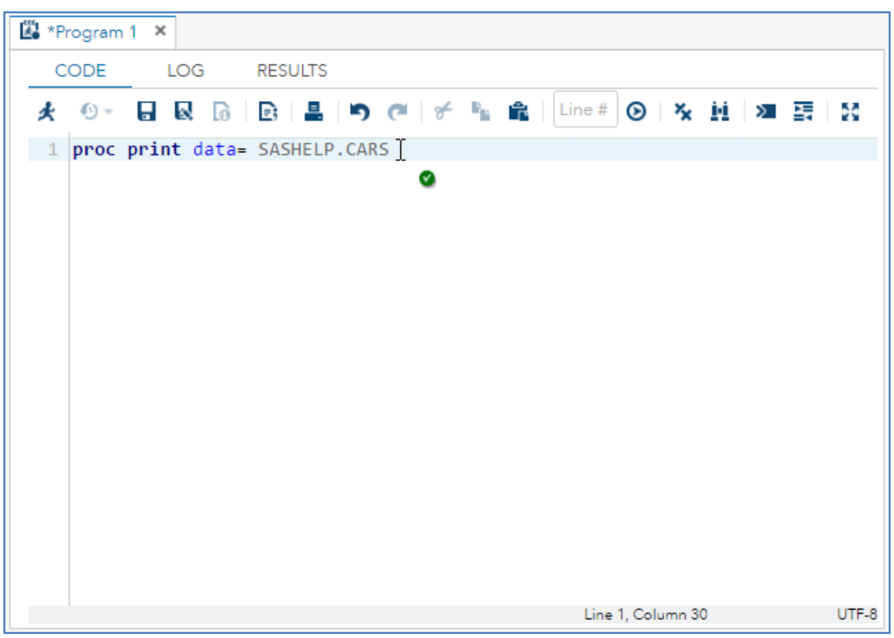
Running a Program

- To run the entire program, click .
- To run a portion of the program, select the lines of code that you want to run and then click .
- If there are no errors, the results open automatically.
- If there are errors, the **Log** tab opens by default. You can expand the **Errors**, **Warnings,** and **Notes** sections to view the messages.
- When you click a message, SAS Studio highlights it for you in the log so that you can see exactly where the message occurs in the log.

Note: Because you are working in a server environment, do not include the **ENDSAS** statement in your SAS programs. If you run a program that contains ENDSAS, reset your SAS session by clicking and selecting **Reset SAS Session**.

Adding Table Names and Column Names

- From the **Libraries** section of the navigation pane, you can use a drag-and-drop operation to move table names and column names into the SAS code.
- For example, you can move the **Sashelp.Cars** table into the DATA option for the PRINT procedure.
- When you release the mouse, the fully qualified name for the table appears in your code.



Describing Missing Data for SAS Dataset, BASEBALL

- In the **Tasks** section, expand the **Data** folder (Tasks and Utilities -> Tasks -> Data), and then double-click **Describe Missing Data**.
- On the **Data** tab, Click .
  - In the Select a Table window, expand the library that contains the data set that you want to use (**SASHELP**).
  - Select the **BASEBALL** data set and click **OK**. The selected data set should now appear in the drop-down list.
  - To the **Analysis variables** role, assign **Salary** and **Div,** using the button.
  - To run the task, click .

Missing Data Frequencies
Legend: ., A, B, etc = Missing

1987 Salary in $ Thousands		
Salary	Frequency	Percent
.	59	18.32
Non-missing	263	81.68

League and Division		
Div	Frequency	Percent
Non-missing	322	100.00

Missing Data Patterns across Variables
Legend: ., A, B, etc = Missing

1987 Salary in $ Thousands	League and Division	Frequency	Percent
.	Non-missing	59	18.3230
Non-missing	Non-missing	263	81.6770

Results:

- Under the Missing Data Frequencies heading, the first table shows 59 observations in the input data set have a missing value for the **Salary** variable.
- The second table shows that there are no missing values for the **League and Division** variable.
- Under the **Missing Data Patterns across Variables** heading, the table shows the pattern of missing values across the variables. In this case, 59 observations have a missing value for the Salary variable.
- The **League and Division** variables have no missing values. Therefore, the remaining 263 observations in the input data set do not have any missing values for the two variables.

Creating Reports, using the SAS CAR Sample Dataset

- In the **Tasks** section, expand the **Data** folder, and then double-click **List Data**.
- On the **Data** tab, select the **SASHELP.CARS** data set.
  - If the data set is not available from the drop-down list, click .
  - In the Select a Table window, expand the library that contains the data set that you want to use. Select the data set for the example and click **OK**. The selected data set should now appear in the drop-down list.
  - Assign columns to these roles:
    - **DriveTrain**, **MSRP**, and **EngineSize** Column Names to the **List Variables** Roles.
    - **Type** variable to the **Group analysis by**, role.
  - Click .

Results:

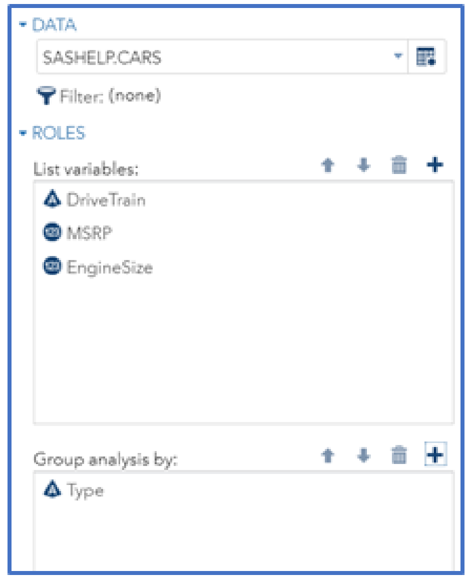
List Data for SASHELP.CARS

Type=Hybrid

Obs	DriveTrain	MSRP	Engine Size (L)
1	Front	$20,140	1.4
2	Front	$19,110	2.0
3	Front	$20,510	1.5

Type=SUV

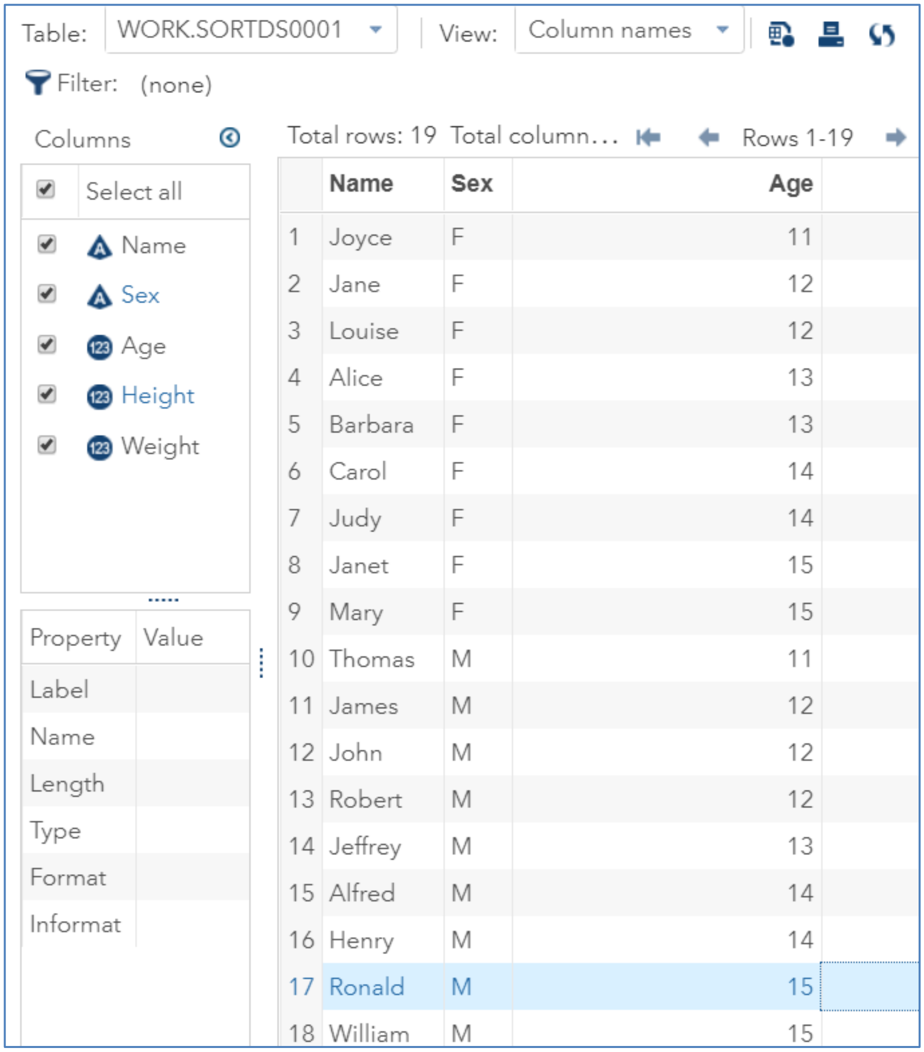
Obs	DriveTrain	MSRP	Engine Size (L)
4	All	$36,945	3.5
5	All	$37,000	3.0
6	All	$52,195	4.4
7	All	$37,895	4.2
8	Front	$26,545	3.4
9	Front	$52,795	5.3
10	Front	$46,995	4.6
11	Front	$42,735	5.3
12	All	$41,465	5.3
13	Front	$30,295	4.2
14	Front	$20,255	2.5
15	All	$32,235	4.7
16	All	$41,475	6.8
17	Front	$34,560	4.6



Sorting the SAS Dataset CLASS by Sex and Age

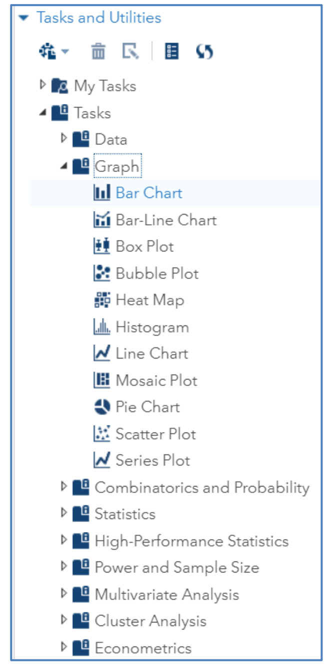
- In the **Tasks** section, expand the **Data** folder, and then double-click **Sort Data**. The user interface for the Sort Data task opens.
- On the **Data** tab, select the **SASHELP.CLASS** data set.
  - If the data set is not available from the drop-down list, click .
  - In the Select a Table window, expand the library that contains the data set that you want to use. Select the data set for the example and click **OK**. The selected data set should now appear in the drop-down list.
- To the **Sort by:** role, assign the **Name**, **Sex** and **Age** columns.

- Click .
- The newly created **WORK.SortDS** data set is available from the **Output Data** tab.
- In the output data, the observations are first sorted by Sex (whether Female or Male). And then, within each group, the observations are sorted in ascending order by age.



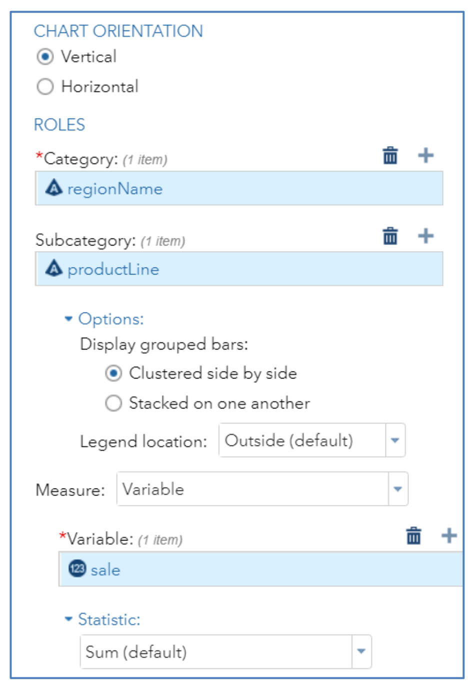
Creating a Bar Chart, using the SAS Dataset Pricedata

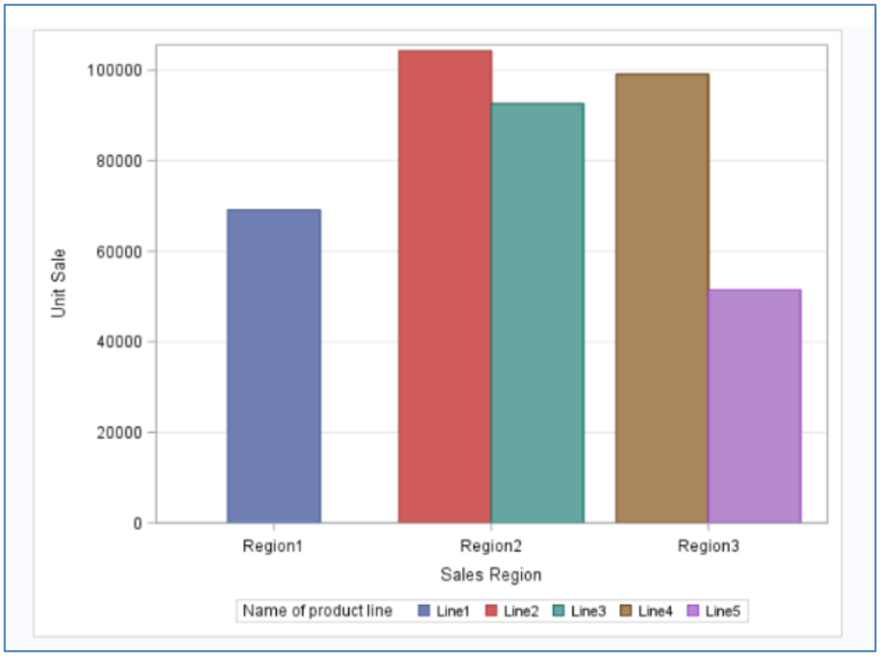
Create a bar chart that compares the total amount of sales for each product line in the **Sashelp.Pricedata** data set.



- In the **Tasks and Utilities** section, expand the **Graph** folder, and then double-click **Bar Chart**. The user interface for the Bar Chart task opens.
- On the **Data** tab, select the **SASHELP.PRICEDATA** data set.
    - If the data set is not available from the drop-down list, click .
    - In the **Select a Table** window, expand the library that contains the data set that you want to use.
    - Select the data set for the example and click **OK**. The selected data set should now appear in the drop-down list.

- Assign **regionName** to the **Category** role.
- Assign **productLine** to the **Subcategory** role.
- From the **Measure** drop-down list, select **Variable**.
    - Assign **sale** to the **Variable** role.
    - From the **Statistic** drop-down list, select **Sum (default)**.
- Click .

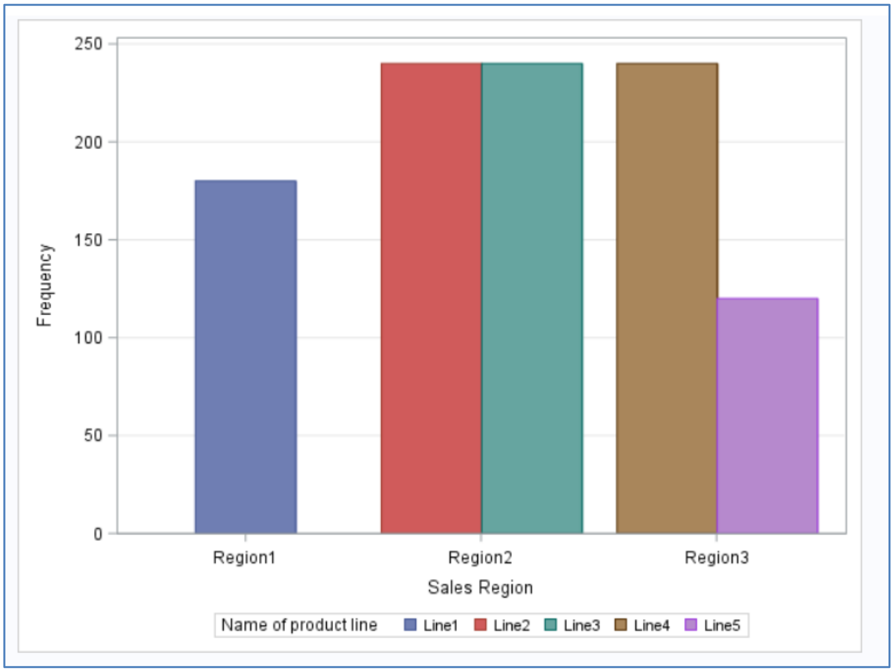




To create Total Sales by Frequency, Using the SAS Pricedata Dataset

Create a bar chart that shows the total sales for each product line by frequency.

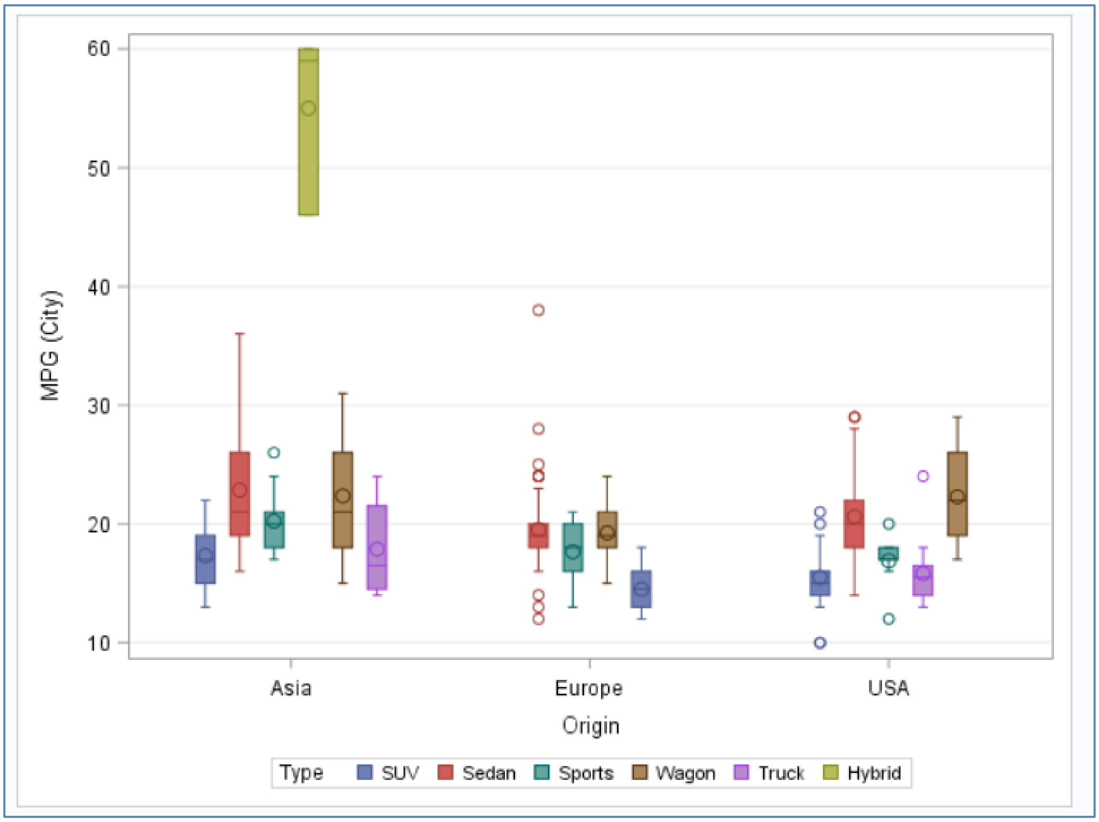
- In the **Tasks and Utilities** section, expand the **Graph** folder, and then double-click **Bar Chart**. The user interface for the Bar Chart task opens.
- On the **Data** tab, select the **SASHELP.PRICEDATA** dataset.
    - If the data set is not available from the drop-down list, click  .
    - In the Select a Table window, expand the library that contains the data set that you want to use.
    - Select the data set for the example and click **OK**. The selected data set should now appear in the drop-down list.
- Assign **regionName** to the **Category** role.
- Assign **productLine** to the **Subcategory** role.
- From the **Measure** drop-down list, select **Frequency count (default)**.
- Click .



Creating Box Plots , using the SAS Cars dataset, to Compare MPG (City) for Cars

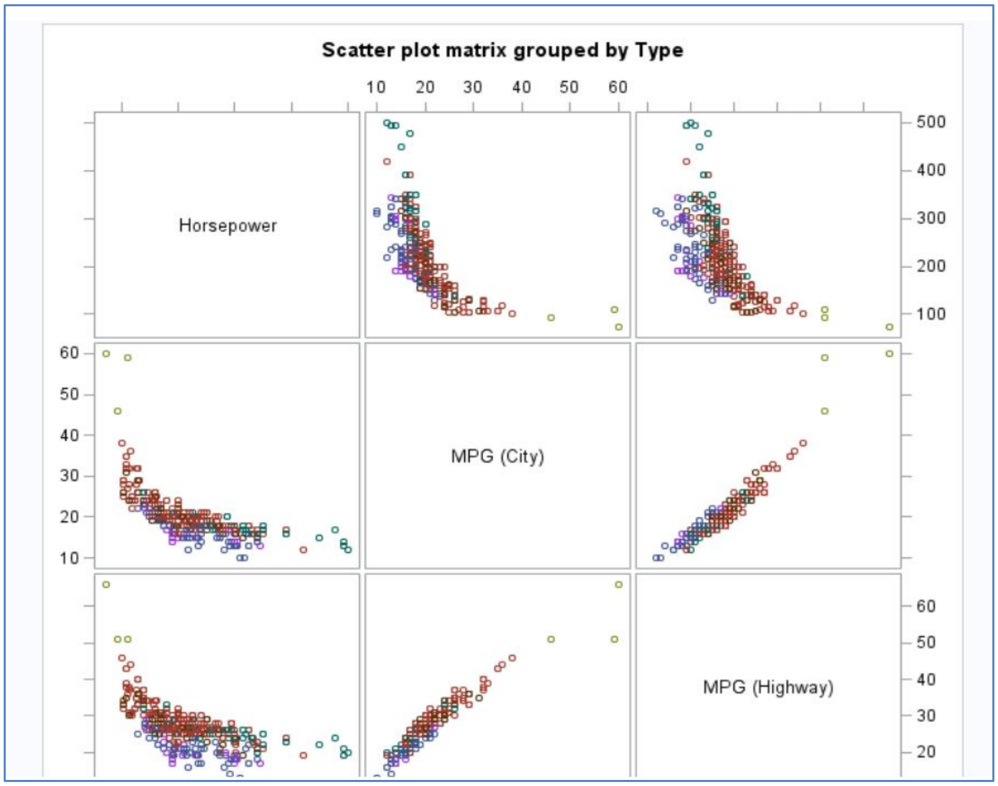
Creates three box plots that compare how many miles per gallon (city) cars consume depending on their area of origin (Asia, Europe, and the United States).

- In the **Tasks and Utilities** section, expand the **Graph** folder (Tasks -> Graph), and then double-click **Box Plot**.
- On the **Data** tab, select the **SASHELP.CARS** data set.
    - If the data set is not available from the drop-down list, click .
    - In the Select a Table window, expand the library that contains the data set that you want to use.
    - Select the data set for the example and click **OK.** The selected data set should now appear in the drop-down list.
- Assign columns to these roles:
    - **MPG_City** to **Analysis variables**
    - **Origin** to **Category**
    - **Type** to **Subcategory**
- Click .

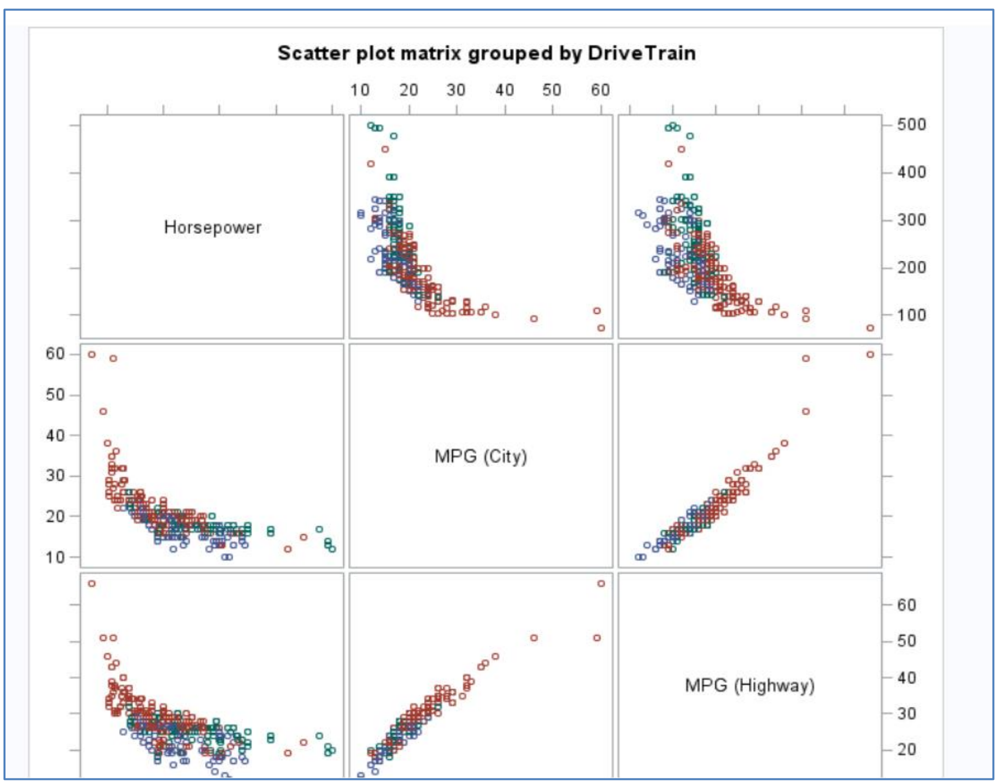


Exploring the SASHELP.CARS Dataset

- In the **Tasks and Utilities** section, expand the **Statistics** folder (Tasks -> Statistics), and then double-click **Data Exploration**.
- On the **Data** tab, select the **SASHELP.CARS** data set.
    - If the data set is not available from the drop-down list, click .
    - In the Select a Table window, expand the library that contains the data set that you want to use.
    - Select the data set for the example and click **OK.** The selected data set should now appear in the drop-down list.
- Assign columns to these roles:
    - **Horsepower**, **MPG_City**, and **MPG_Highway** variables to the **Continuous variables** role.
    - **Type**, and **DriveTrain** variables to the **Classification variables** role.
- Click .



Scatter plot matrix grouped by Type
Horsepower
MPG (City)
MPG (Highway)
10 20 30 40 50 60
500 400 300 200 100
60 50 40 30 20 10
60 50 40 30 20



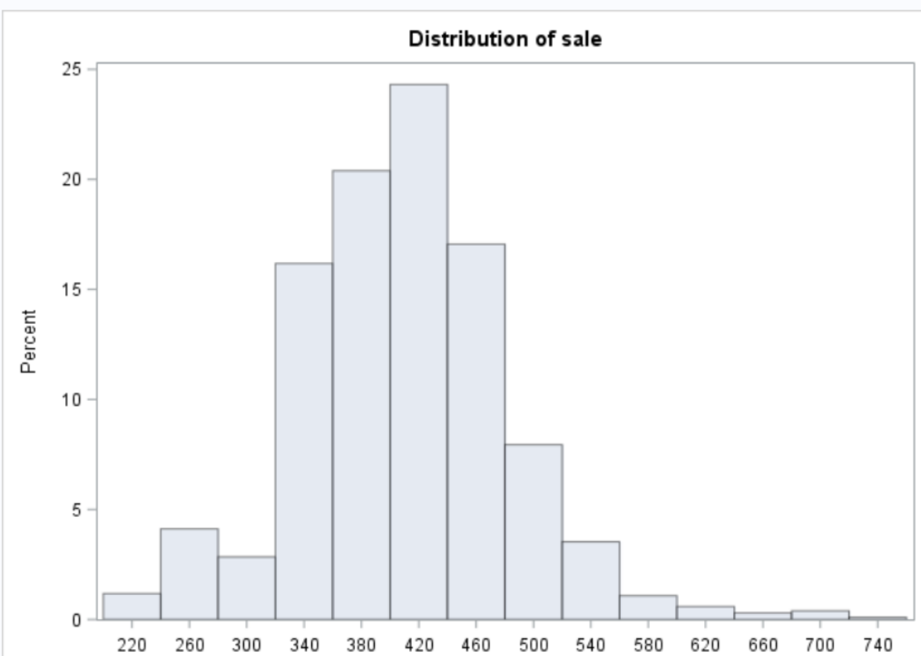
Scatter plot matrix grouped by DriveTrain
Horsepower
MPG (City)
MPG (Highway)
10 20 30 40 50 60
500 400 300 200 100
60 50 40 30 20 10
60 50 40 30 20

To Create Summary Statistics of Unit Sales, using the SAS Pricedata Dataset

To analyze unit sales, and to display a histogram of the distribution:

- In the **Tasks and Utilities** section, expand the **Statistics** folder, and then double-click **Summary Statistics**. The user interface for the Summary Statistics task opens.
- On the **Data** tab, select the **SASHELP.PRICEDATA** data set.
    - If the data set is not available from the drop-down list, click .
    - In the Select a Table window, expand the library that contains the data set that you want to use.
    - Select the data set for the example and click **OK**. The selected data set should now appear in the drop-down list.
- To the **Analysis variables** role, assign the **sale** column.
- On the **Options** tab, expand the **Plots** section and select the **Histogram** check box.
- Click .

Analysis Variable : sale Unit Sale				
Mean	Std Dev	Minimum	Maximum	N
408.5558824	73.0840041	203.0000000	747.0000000	1020



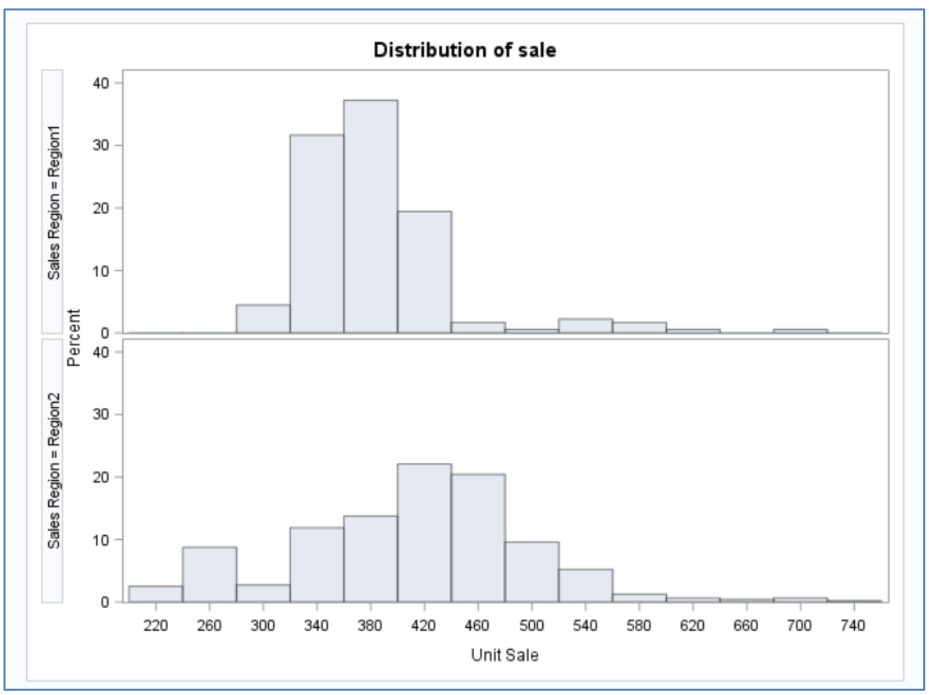
Distribution Analysis of Sales for Each Region, using the SAS Pricedata Dataset

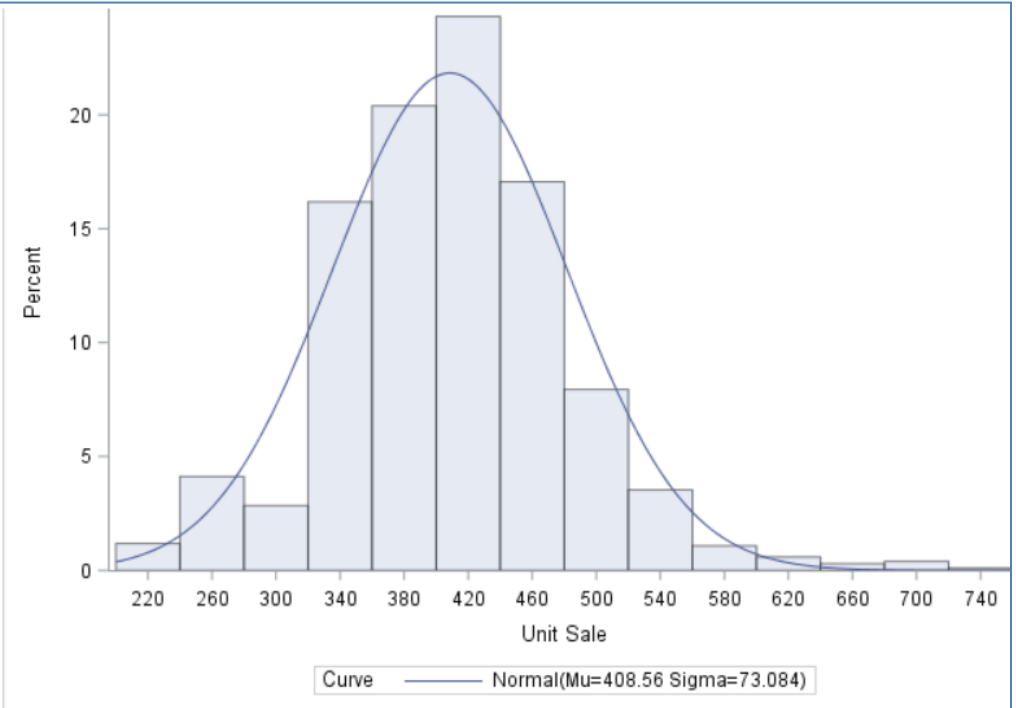
To analyze the sales for each region (because the data contains three regions, you get three sets of results):

- In the **Tasks and Utilities** section, expand the **Statistics** folder, and then double-click **Distribution Analysis**. The user interface for the Distribution Analysis task opens.
- On the **Data** tab, select the **SASHELP.PRICEDATA** data set.
  - If the data set is not available from the drop-down list, click .
  - In the Select a Table window, expand the library that contains the data set that you want to use.
  - Select the data set for the example and click **OK**. The selected data set should now appear in the drop-down list.
- To the **Analysis variables** role, assign the **sale** variable.
- On the **Options** tab, set these options:

  - In the **Exploring Data** group, assign the **regionName** variable to the **Classification variables** role.
  - In the **Checking for Normality** group, select the **Histogram and goodness-of-fit tests** and **Normal quantile-quantile plot** options.

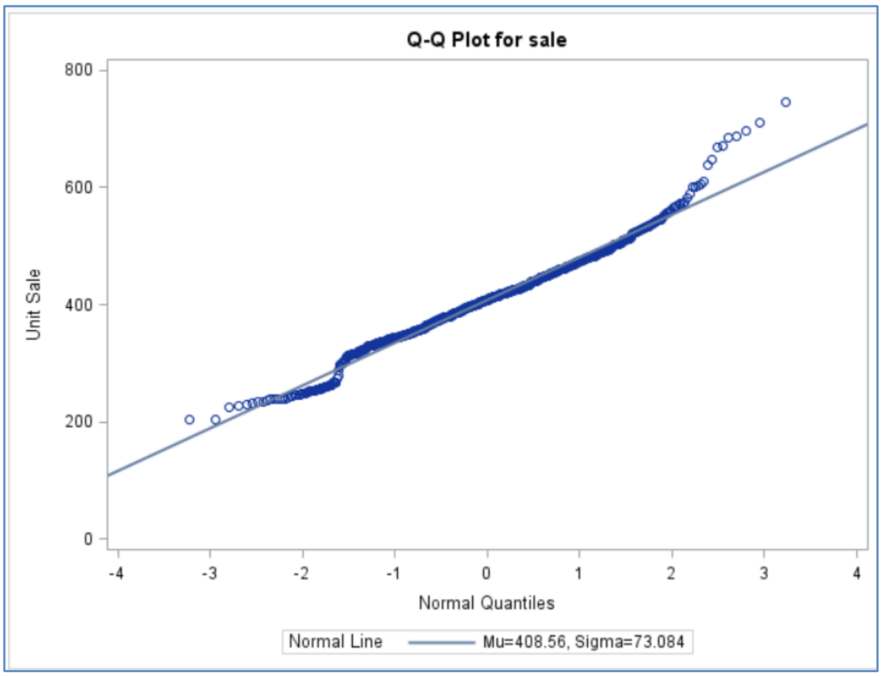
  Click .





Fitted Normal Distribution for sale (Unit Sale)

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.04121623	Pr > D	<0.010
Cramer-von Mises	W-Sq	0.44027199	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	3.46017068	Pr > A-Sq	<0.005



Performing One-Way Frequencies of Unit Sales, using the SAS Pricedata Dataset

To analyze unit sales for each sales region:

- In the **Tasks and Utilities** section, expand the **Statistics** folder, and then double-click **One-Way Frequencies**. The user interface for the One-Way Frequencies task opens.
- On the **Data** tab, select the **SASHELP.PRICEDATA** data set.
    - If the data set is not available from the drop-down list, click .
    - In the Select a Table window, expand the library that contains the data set that you want to use.
    - Select the data set for the example and click **OK.** The selected data set should now appear in the drop-down list.
- Assign columns to these roles:
    - **Sale** variable to **Analysis variables** role
    - **regionName** to **Group analysis by**, role (Under Additional Roles).
- Click .

Sales Region=Region1

Unit Sale

sale	Frequency	Percent	Cumulative Frequency	Cumulative Percent
298	1	0.56	1	0.56
300	1	0.56	2	1.11
301	1	0.56	3	1.67
307	1	0.56	4	2.22
308	1	0.56	5	2.78
314	1	0.56	6	3.33
316	1	0.56	7	3.89
318	1	0.56	8	4.44
320	1	0.56	9	5.00
321	1	0.56	10	5.56
322	2	1.11	12	6.67
323	1	0.56	13	7.22
324	2	1.11	15	8.33
328	1	0.56	16	8.89
331	3	1.67	19	10.56
332	2	1.11	21	11.67
333	1	0.56	22	12.22
334	2	1.11	24	13.33
335	1	0.56	25	13.89
337	1	0.56	26	14.44
338	4	2.22	30	16.67
339	1	0.56	31	17.22

To Perform Correlations Analysis using the SAS Cars Dataset

- In the **Tasks** section, expand the **Statistics** folder, and then double-click **Correlation Analysis**. The user interface for the Correlation Analysis task opens.
- On the **Data** tab, select the **SASHELP.CARS** data set.
  - If the data set is not available from the drop-down list, click .
  - In the Select a Table window, expand the library that contains the data set that you want to use.
  - Select the data set for the example and click **OK**. The selected data set should now appear in the drop-down list.
- Assign columns to these roles:
  - **EngineSize** and **Horsepower** variables to **Analysis Variables** roles
  - **Cylinders** and **MPG_Highway** variables to **Correlate with**, role.
- Click .

2 With Variables:	Cylinders MPG_Highway
2 Variables:	EngineSize Horsepower

Pearson Correlation Coefficients Number of Observations		
	EngineSize	Horsepower
Cylinders	0.90800 426	0.81034 426
MPG_Highway MPG (Highway)	-0.71730 428	-0.64720 428

Performing One-Sample t Test for Horsepower , using the SAS Cars Dataset

- In the **Tasks and Utilities** section, expand the **Statistics** folder, and then double-click **t Tests**.
- On the **Data** tab, select the **SASHELP.cars** data set.
  - If the data set is not available from the drop-down list, click .
  - In the Select a Table window, expand the library that contains the data set that you want to use.
  - Select the data set for the example and click **OK**. The selected data set should now appear in the drop-down list.
- From the **t test** drop-down list, select **One-sample test**.

- To the **Analysis variable** role, assign the **Horsepower** column.
- On the **Options** tab, enter 300 in the **Alternative hypothesis** field.
- Click .

Variable: Horsepower

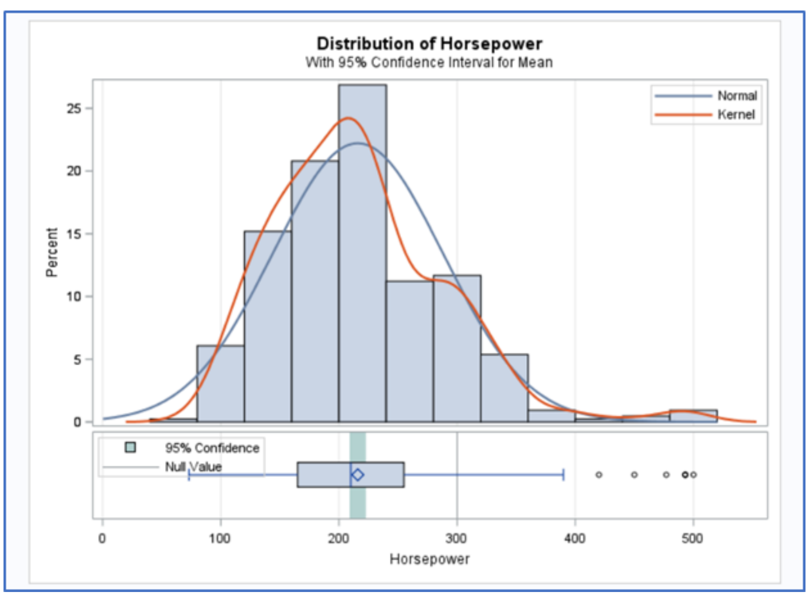
Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.949922	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.090516	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.589806	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	3.685805	Pr > A-Sq	<0.0050

Variable: Horsepower

N	Mean	Std Dev	Std Err	Minimum	Maximum
428	215.9	71.8360	3.4723	73.0000	500.0

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
215.9	209.1	222.7	71.8360	67.3244	77.0007

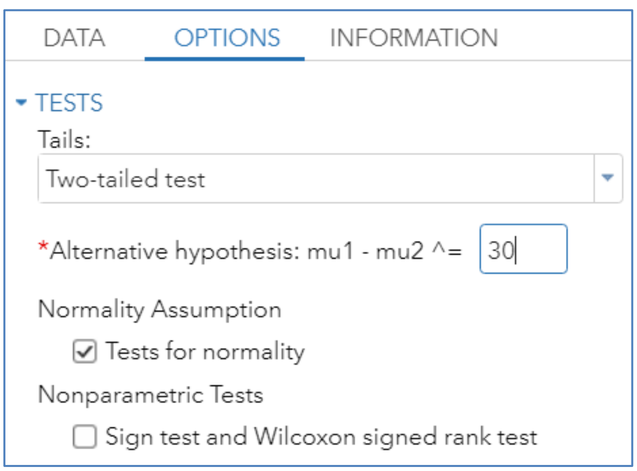
DF	t Value	Pr > |t|
427	-24.22	<.0001



Performing Paired t-Test, to compare the means of differences in price and cost, using the SAS Pricedata Dataset

To compare the means of differences in price and cost in the Sashelp.Pricedata data set. Assuming that the null hypothesis for this test is 30 (Determining the Distribution of (Price – Cost)):

- In the **Tasks** section, expand the **Statistics** folder, and then double-click **t Tests**. The user interface for the t Tests task opens.
- On the **Data** tab, select the **SASHELP.PRICEDATA** data set.
- From the **t test** drop-down list, select **Paired test**.
- Assign columns to these roles (using the button):
  - **Price** variable to Group 1 variable
  - **Cost** variable to Group 2 variable
- On the **Options** tab, enter 30 in the **Alternative hypothesis** field.
- Click . ***Note:*** "**^=**" is NOT Equal Operator, in SAS.



Variable: _Difference_ (Difference: price - cost)

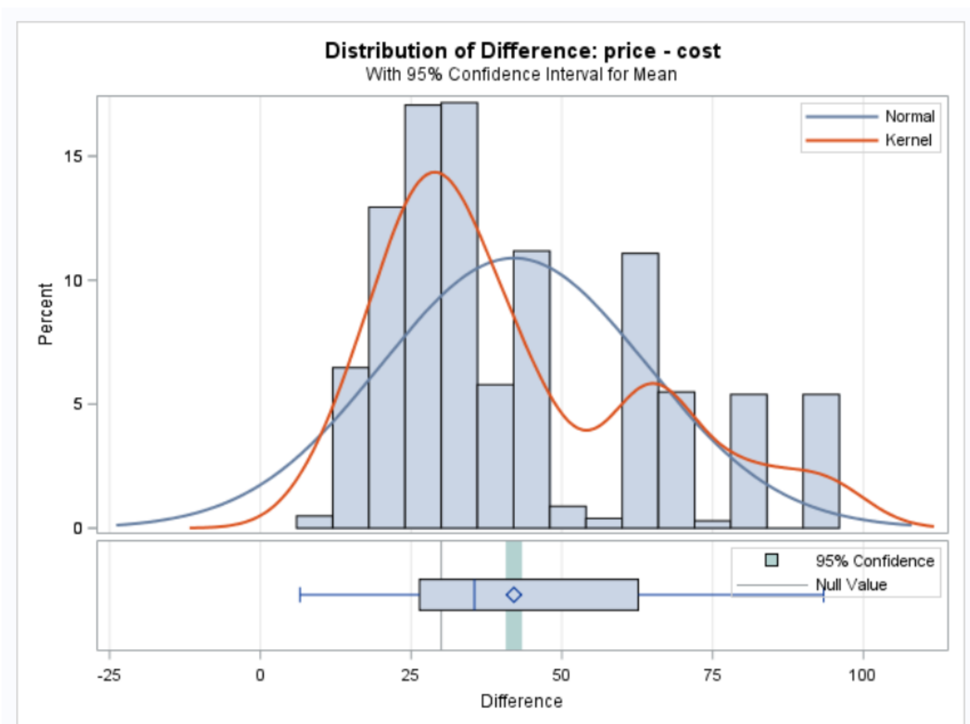
Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.896986	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.1888	Pr > D	<0.0100
Cramer-von Mises	W-Sq	7.159388	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	39.28743	Pr > A-Sq	<0.0050

Difference: price - cost

N	Mean	Std Dev	Std Err	Minimum	Maximum
1020	42.0448	21.9813	0.6883	6.5700	93.4000

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
42.0448	40.6942	43.3954	21.9813	21.0671	22.9791

DF	t Value	Pr > |t|
1019	17.50	<.0001



Performing One-Way ANOVA, to Test for Differences in the Means for MPG_Highway by Car Type, using the SAS Cars Dataset

To study the differences in the means for the number of highway miles per gallon for six car types:

- In the **Tasks and Utilities** section, expand **Tasks**, the **Statistics** folder and then double-click **One-Way ANOVA**.
- On the **Data** tab, select the **SASHELP.CARS** data set.
- Assign columns to these roles:
    - **MPG_Highway** variable to **Dependent variable** Role
    - **Type** variable to **Categorical variable** Role.
- Click .

Class Level Information		
Class	Levels	Values
Type	6	Hybrid SUV Sedan Sports Truck Wagon

Number of Observations Read	428
Number of Observations Used	428

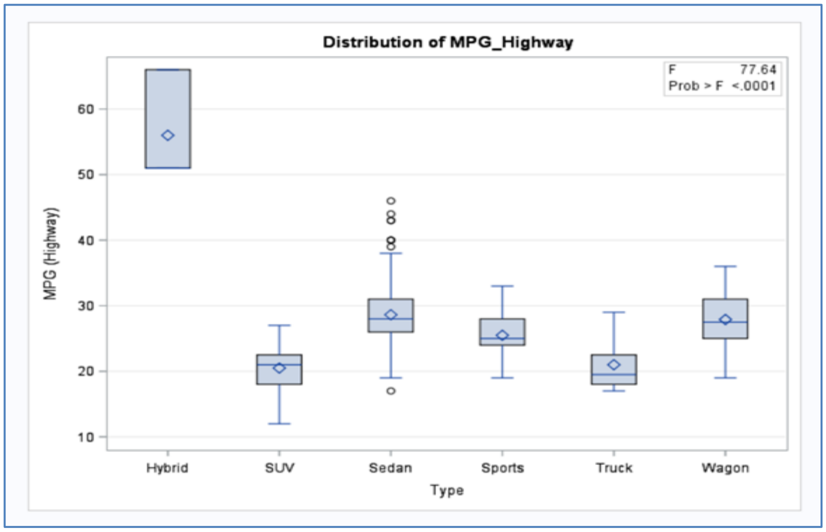
Dependent Variable: MPG_Highway MPG (Highway)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	6743.47900	1348.69580	77.64	<.0001
Error	422	7331.03268	17.37212		
Corrected Total	427	14074.51168			

R-Square	Coeff Var	Root MSE	MPG_Highway Mean
0.479127	15.52701	4.167987	26.84346

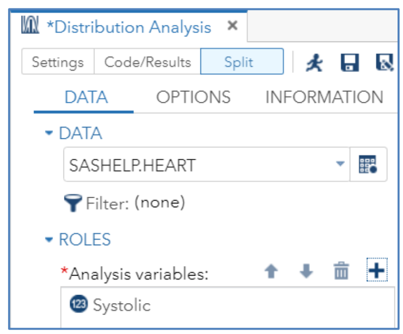
Source	DF	Type I SS	Mean Square	F Value	Pr > F
Type	5	6743.478998	1348.695800	77.64	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Type	5	6743.478998	1348.695800	77.64	<.0001

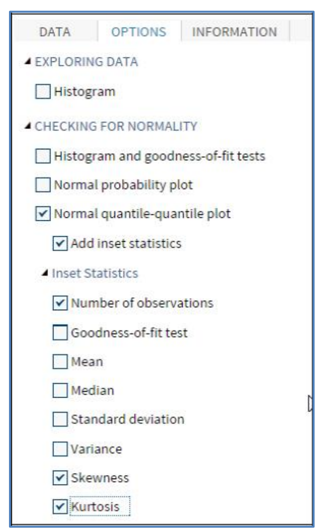


Investigating the Distribution for Systolic Blood Pressure, using the SAS Heart dataset 

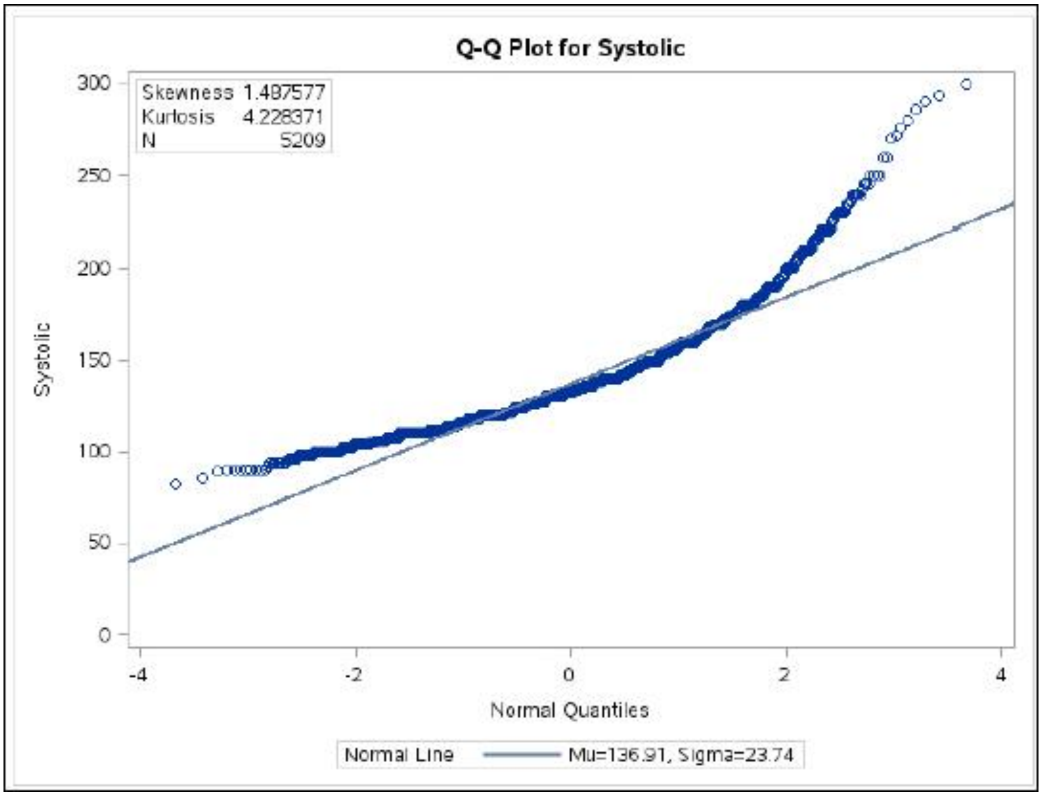
- **Library** -> **My Libraries** -> **SASHELP** -> **Heart**. The selected data set should now appear in the drop-down list.
- To investigate the distribution of systolic blood pressure, select **Distribution Analysis** from the list of **Statistics** tasks. Make sure that the Heart data set is selected on the **DATA** tab and **Systolic** is selected as the **Analysis variable**.



- Click the **OPTIONS** tab to bring up its menu:
  - Deselect the box next to Histogram.
  - On the **Checking For Normality** Tab, check the **Normal quantile-quantile plot** box, and **Add inset statistics** boxes.
    - Then, check the **Skewness**, and **Kurtosis** boxes (expanding the **Inset Statistics** tab).



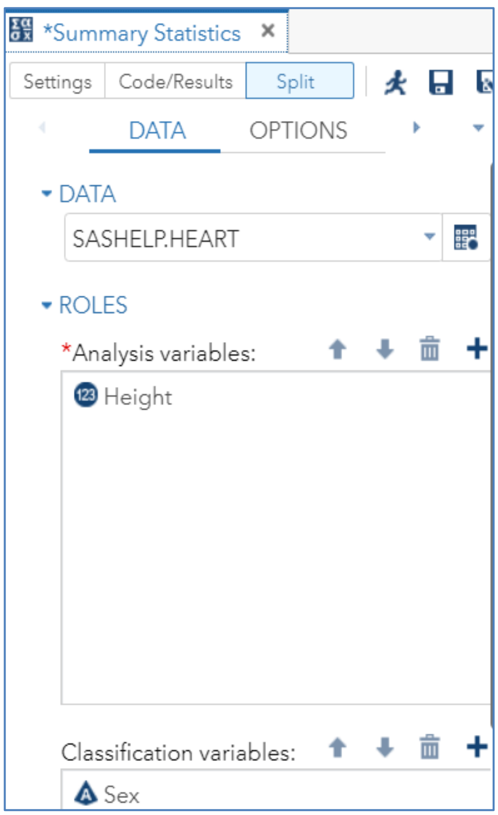
- Click .

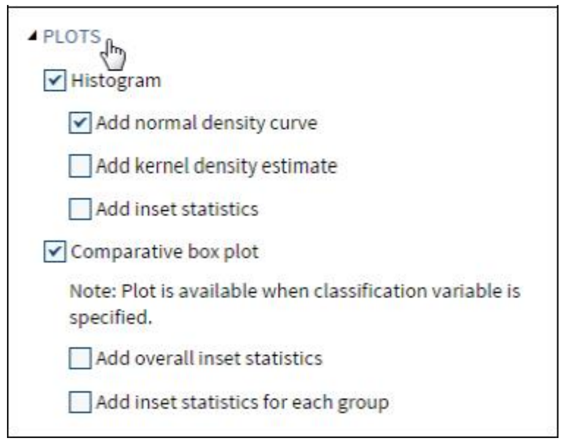


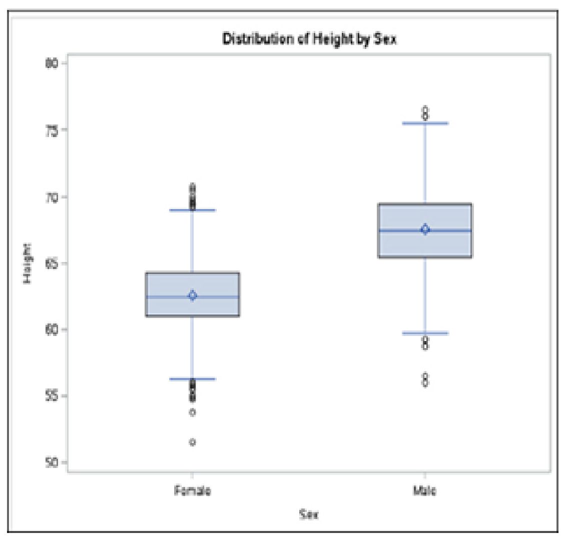
- The straight line on the plot represents a normal distribution with the same mean and standard deviation as the variable Systolic.
- The circles on the plot represent values of systolic blood pressure from your sample data. At the bottom of the Q-Q plot, you see that the theoretical normal distribution has a mean (Mu) equal to 136.91 and a standard deviation (Sigma) equal to 23.74.
- To help you understand this Q-Q plot, look at the right side of the plot. The circles above the straight line on this part of the plot indicate that your sample data includes values of systolic blood pressure that are higher (more extreme) than you would expect if the systolic blood pressures were normally distributed. This confirms the strong positive skewness that you saw in the histogram.
- Values for Skewness and Kurtosis close to zero result from distributions that are close to normal. Positive values for skewness, as in this plot, indicate a positively skewed distribution (extreme values in the right tail). Positive values for kurtosis (as in this example) indicate both that the distribution is too peaked and that the tails are too heavy. Negative values for kurtosis indicate that the distribution is too flat and that the tails are too light.

Adding a Classification Variable in the Summary Statistics Tab (using the SAS Heart dataset)

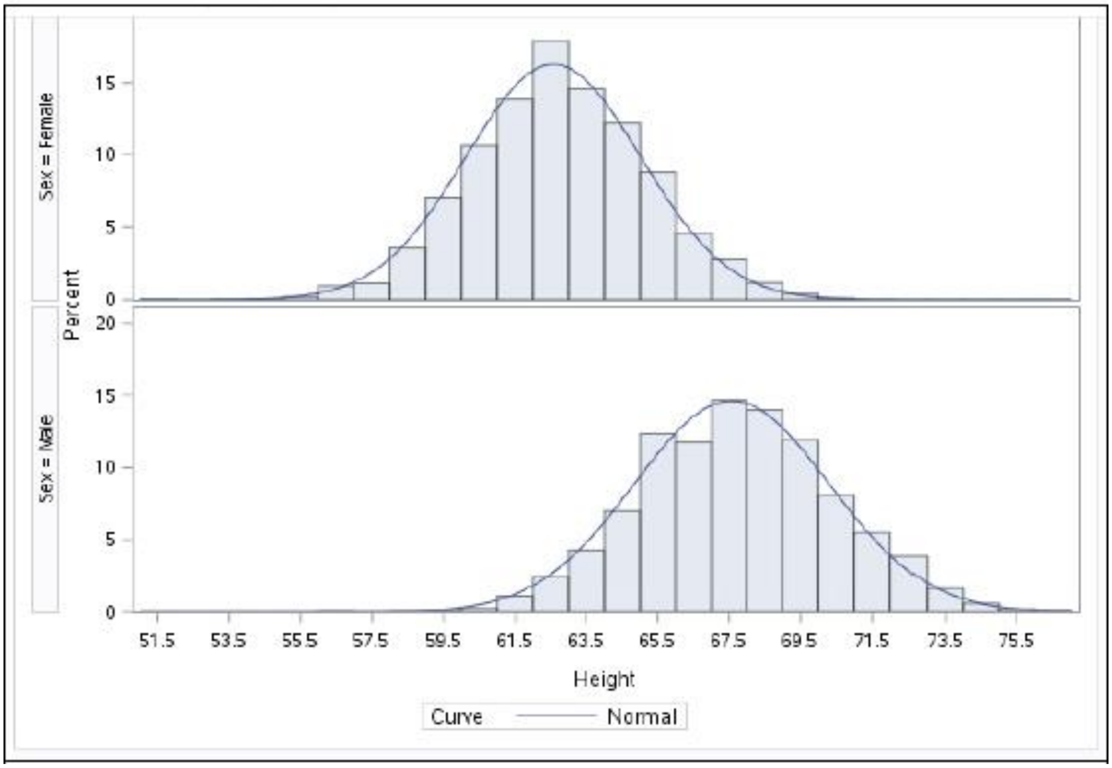
- To compare the variable **Height** for **males** and **females**:
    - **Tasks** -> **Statistics** -> **Summary Statistics**
    - On the **Analysis variables** section of the **DATA** tab add the variable **Height** and, in the box labeled **Classification** variables, add the variable **Sex**.
- On the Options Tab, under the **Plots** option, you can select a **histogram** and **Add normal density curve** boxes.
- Also select **Comparative box plot** to display distributions of Height for males and females.
- Click .







Analysis Variable : Height						
Sex	N Obs	Mean	Std Dev	Minimum	Maximum	N
Female	2873	62.5725863	2.4524112	51.5000000	70.7500000	2869
Male	2336	67.5673736	2.7321366	56.0000000	76.5000000	2334



Note: As expected, the center of the distribution for Sex=Male (bottom histogram) is shifted to the right compared to the distribution for Sex=Female.

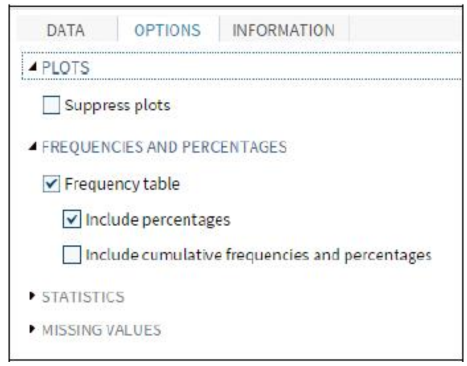
The box plots also show more outliers in the female distribution of Height compared to the male distribution. This may be partly due to the smaller interquartile range (the distance from the top to the bottom of the box) for the females compared to the males.

Describing Categorical Variables, using the SAS Heart dataset

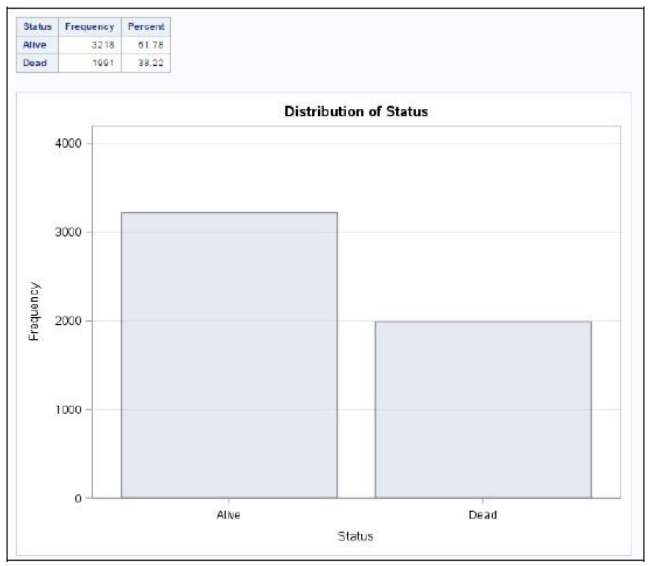
Descriptive statistics for most categorical variables consist of **frequency** tables and **bar charts.** You may also want to display **two-way tables** to investigate relationships between two categorical variables.

- On the **Data** tab, select the SAS **Heart dataset.**
- The first step in generating frequency tables and bar charts is to double-click the **One-Way Frequencies** selection on the **Statistics** tab (Tasks and Utilities -> Tasks -> Statistics), and then double-click **One-Way Frequencies.**

- Select the three variables, **Status**, **Sex**, and **Chol_Status,** as **Analysis variables**.
- On the **Options** tab -> check the **Include Percentages** box.
- Click .

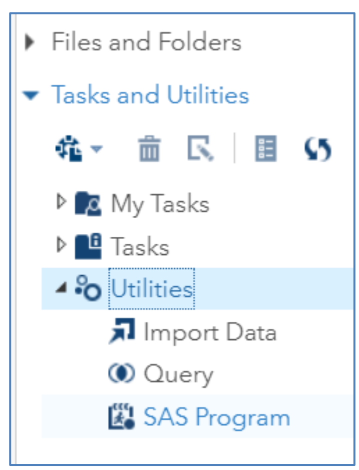


The frequency table shows the number of people in each category (Alive versus Dead) as well as a bar chart displaying the frequencies graphically.



Importing an Excel Worksheet:

- Click **Files and Folders** in the navigation pane and expand the **Tasks** and Utilities.
- Expand the **Utilities** -> **Import Data**.
- Right-click the file that you want to import and select Import Data.
- Click .
- **Save** the imported file.
- For more information, visit https://support.sas.com/software/products/university-edition/faq/SAS_accessdata_Excel.htm



SAS® University Edition &

SAS OnDemand for Academics

The two FREE SAS software are:

1. SAS University Edition

The SAS University Edition is a free version of SAS that will allow Faculty, Staff, and Students to use the SAS product to perform statistical analyses. ***This is not a full version of SAS program.*** SAS University Edition runs as a **virtual machine** on your native operating system and access is via a sweb browser to run statistics. This is explained in the installation and user guides. Three different versions are available using **Oracle Virtual Box**, **VMWare Player**, and **VMWare Fusion**. VirtualBox and VMWare player are free. This version of SAS may be run on **Windows**, **Mac OS X**, and **Linux**. SAS University Edition **is not** restricted to Faculty, Staff, and Students only.

Steps to download and install SAS University Edition, using Virtual Box

- System Requirements
    - Microsoft Windows version 7 or later
    - 32 / 64-bit hardware with a minimum of 1GB of RAM. For 32-bit hardware, you need to tweak some settings. Check out this link - 32-bit hardware SAS® University Edition via Virtual Box
    - One of the following virtualization software packages:
        - VMware Player
        - Oracle VirtualBox
    - One of the following web browsers:
        - Microsoft Internet Explorer 9, 10 or 11
        - Mozilla Firefox 21 or later
        - Google Chrome 27 or later

To download SAS University Edition, please visit
https://www.sas.com/en_us/software/university-edition.html

2. SAS OnDemand for Academics : SAS Studio

SAS OnDemand for Academics is available for everyone - Individual Learners, Teachers and Students. **It does not require any installation.** It **runs on cloud** so internet access is required to use this software. It comes with various modules of SAS such as SAS Operation Research module, SAS Enterprise Guide (SAS EG), SAS Enterprise Miner, SAS Credit Scoring etc. which are not available in SAS University Edition.

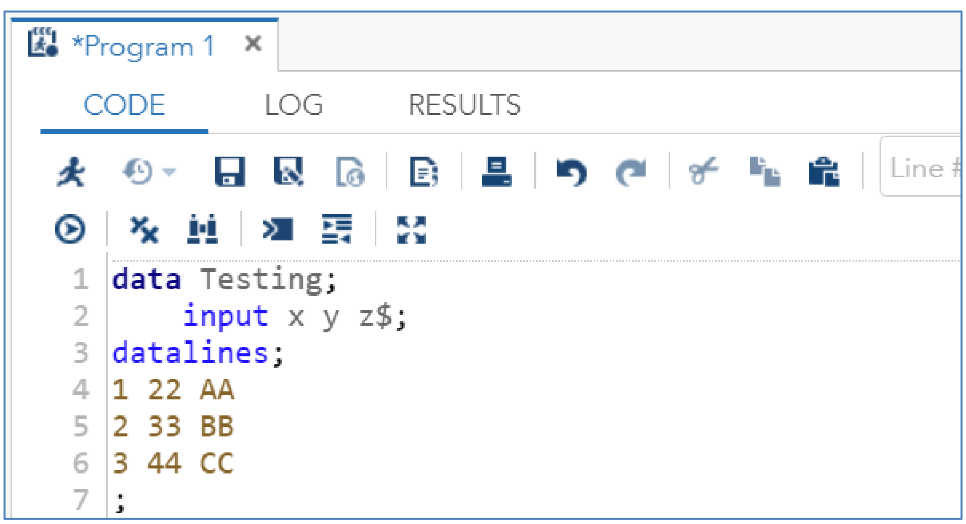
- To register, visit https://odamid.oda.sas.com/SASODARegistration/
- After completing your registration, you will receive a User ID and a URL to access SAS Studio

Comparisons of SAS University Edition & SAS OnDemand for Academics

	SAS University Edition	SAS OnDemand for Academics
Cost	Free	Free
Installation	Installation Required	No Installation Required
Cloud	Run on local machine	Run on cloud so access from everywhere
Internet	No Internet required to run	Internet required to run
Data Size	Huge data handling capability	Medium
Modules	Supports popular modules	Supports more modules than University Edition

Interface of both SAS University Edition, and SAS OnDemand for Academics

(both programs run SAS Studio)



Free SAS e-Learning

- **SAS Studio video library** http://video.sas.com/#category/videos/sas-studio
- **SAS Programming 1 e-course** https://support.sas.com/edu/schedules.html?ctry=us&crs=PROG1
- **Statistics 1 e-course** https://support.sas.com/edu/schedules.html?ctry=us&crs=STAT1
- **Library of SAS eLearning Courses** https://support.sas.com/edu/elearning.html?ctry=us&productType=library
-

Online Resources

- **LEARNING CENTER / TRAINING –** including SAS Self-Paced e-Learning and free tutorials http://www.sas.com/apps/elearning/elearning_category_welcome.jsp
- **Safari eBooks** from http://library.ttu.edu
  - **Biostatistics by Example Using SAS Studio,** by: Ron Cody, September 22, 2016
  - **An Introduction to SAS University Edition, by:** Ron Cody, September 16, 2015

Where to Get Help

- Texas Tech University has site licenses for SAS software, and as a member of the TTU community, you have the right to have full 24 hours support from SAS. You can do that via the e-mail: mailto:support@sas.com . For any problem you have you can get in touch with them through this e-mail. When you send an e-mail message, remember to insert at the beginning of the message **TTU's SAS site number.**

- If you need a Statistical Consultant, please visit the **Statistical Consulting Lab** (**SCL**) of the Department of Mathematics & Statistics www.math.ttu.edu/~SCL
- 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