INTRODUCTORY LAB: DOING STATISTICS WITH SPSS 21

This section covers the basic structure and commands of SPSS for Windows Release 21. It is not designed to be a comprehensive review of the most important SPSS features. It only covers those features of SPSS that are essential for using SPSS for the data analyses in the labs. Some specific commands and more advanced software features will be explained in the labs where they are needed for the statistical analysis.

1. Introduction

SPSS for Windows is one of the most widely used and well-liked statistical computing packages available for the PC. It can take data from almost any type of file and use them to generate summary reports, charts, descriptive statistics, and complex statistical analyses. Most tasks can be accomplished simply by pointing and clicking the mouse. The version 21 you will be working with is designed to operate on computer systems running Windows 7, XP Professional or Vista.

2. Getting Started with Windows 7

Before you start working with SPSS and its various features, you need to know a few things about Windows. Though SPSS can run on any of the three platforms (PC, MAC, and UNIX), almost all campus workstations use Windows 7 and also the vast majority of students use the system on their home computers. The instructions in this section is intended for those students who have had a very limited computer experience and are not familiar with the Windows operating system. All other students can skip the section and move to the next section.

Turn on the computer in your computer lab. The graphical user interface of the University of Alberta *Instructional Laboratories* is displayed.

Although you can communicate with your workstation with some combination of keystrokes, many pointing or choosing tasks are more easily done by using a pointer device such as a mouse, trackball, or touchpad. Moving a pointer device moves a mouse pointer, an onscreen graphic that in its most common form takes the shape of an arrow. Moving a mouse pointer over another object and pressing one of the buttons on the pointer device defines a mouse operation.

To select an on screen object, move the mouse pointer (by moving the mouse) directly over an object and then quickly press and release the left mouse buton. **To drag** or move an object, first move the mouse pointer over an object and then, while holding down the left mouse button, move the mouse. After the object has been dragged, release the mouse button. This technique is useful for moving windows and icons about on the screen. **To double-click** an object, move the mouse pointer directly over an object and press the left mouse button twice in rapid succession.

If you are not familiar with the above mouse operations and Windows, click *Start* button in the bottom left corner on the desktop with the left mouse button, and select *Help and Support* option in the left panel. The *Help and Support Center* dialog window is displayed. Select *Windows Basics*. The program covers the most important mouse and windows operations. Just follow the instructions.

There are several icons displayed on your desktop: *Computer*, *Recycle Bin*, *AICT Help Desk*, *uAlberta Google Apps, Trouble Reporter, Password Maintenance* and *Authenticate*.

To see what any icon does, you double-click it. If that icon is a program, the program is started, and you see the program window. If that icon is a container icon like the *Recycle Bin* or *My Computer*, you see the contents of the icon in a window. In other words, anything you do in Windows is displayed in a window.

The *Computer* program provides an easy access to computer resources and files. From this window, you can get to your computer drives, printers, network, and other resources. You can view what files, folders, and programs are stored on the hard and floppy disks. Files deleted from an application or from the *Computer* list end up in the *Recycle Bin.* You should click on this bin and empty it before signing off from the computer.

In order to access a web browser (Mozilla Firefox, Internet Explorer or Google Chrome) or Google Apps you must authenticate by double-clicking on the icon *Authenticate* and entering your CCID (Campus Computing ID) and the password. *My AFS Disk Space* allows you to store and access your files in your personal directory on the AICT server.

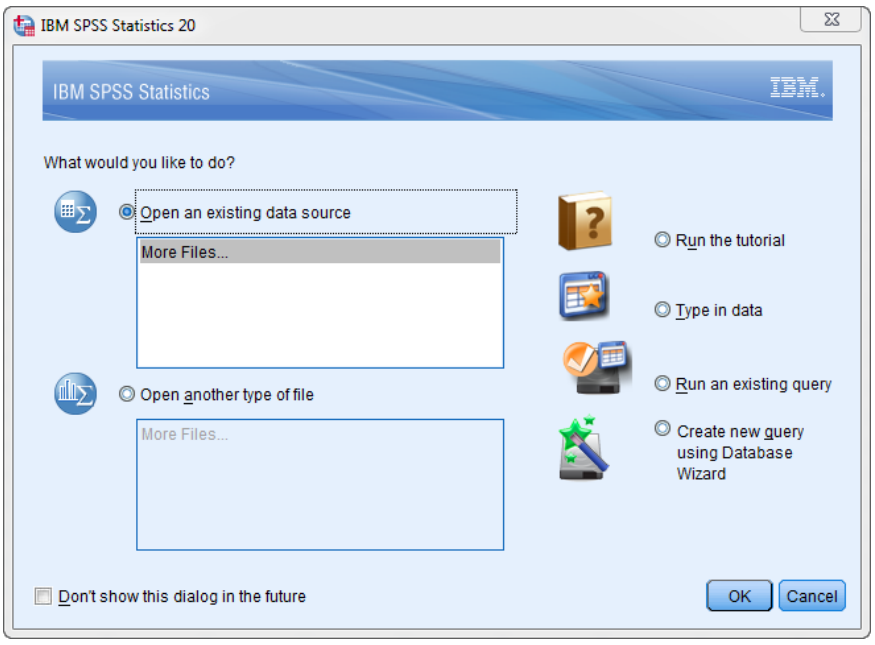
The great thing about windows is that they all work the same way; they all have the same set of controls. You open, close, move, resize, and scroll all windows the same way.

To resize a window, put the pointer on the window edge and drag. **To minimize** a window, click the *Minimize* button. The window is displayed as a taskbar button. **To maximize** a window, click the *Maximize* button. The window is enlarged to fill the entire screen. When a window is maximized, the maximize button changes to the *Restore* button. You can click this button to restore the window to its original size. To *move* a window, place your mouse pointer on the *Title bar* of the window and drag it to a new location.

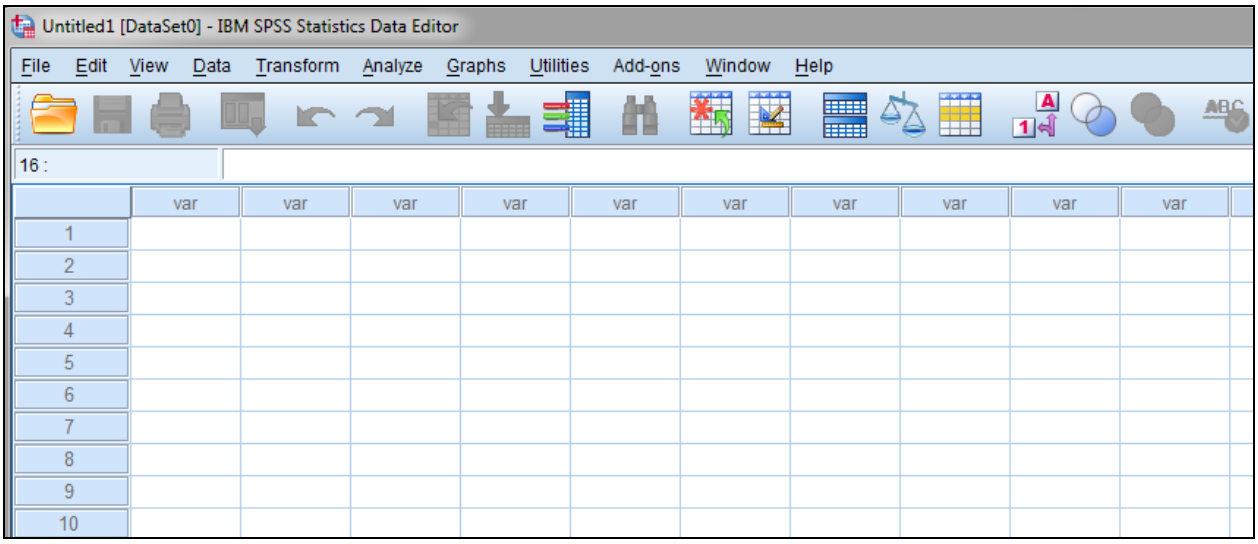
The window with which you are working at any given time is called the *active* window. When more than one window is open, the windows are displayed so as to ovelap. The active window appears in the foreground. You can change a window to active simply by clicking on the edge of the desired window.

3. Launching SPSS

There are many ways to launch SPSS. The easiest way is to start it from the *Start* button located at the bottom of the Windows desktop. Click the *Start* button, then click *All Programs*, select *SPSS Inc*, and finally IBM SPSS Statistics 21 . The following start dialog window opens to instruct SPSS what you intend to do:



You may select *Run the tutorial* to have a tour of SPSS most basic features. If you select *Type in data* the *SPSS Data Editor* opens with the window looking approximately as the picture displayed below:



4. Windows in SPSS

In running SPSS, you will encounter several windows. The four most common windows in SPSS are:

Data Editor. This window displays the contents of the current (working) data file. You can create new data files or modify existing ones with the *Data Editor*. The *Data Editor* window opens automatically when you start an SPSS session. You can have only one data file open at a time.

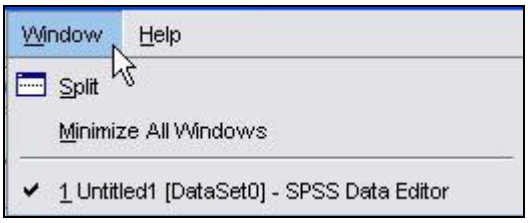
Viewer. This window displays the results of any statistical procedures you run and other text. In particular, tables, statistics, and charts are displayed in the *Viewer* window. A *Viewer* window opens automatically the first time you run a procedure that generates output. The window is not accessible until after output has been generated.

Chart Editor. This window is used to edit charts and plots. It is only displayed after SPSS has been requested to produce a plot. You can use the window to change the colors, select different type fonts or sizes, rotate axes, change the chart type, and the like. The window can be accessed by double-clicking on any graph displayed in the Viewer.

Syntax Editor. Most SPSS commands are accessible from the SPSS menus and dialog boxes. However, some commands and options are available only by using the SPSS command language. In this case the Syntax Window is used. You will also use this window if you wish to run SPSS commands instead of clicking on the pull-down menus.

Each window in SPSS has its own menu bar with menu selections appropriate for that window type. The *Analyze* and *Graphs* menus are available both in Data Editor and Viewer windows, making it easy to generate new output without having to switch windows. Moreover, each SPSS window has its own toolbar that provides quick, easy access to common tasks.

You can change a window to active simply by clicking on the edge of the desired window. You can also activate windows by selecting *Window* from the menu bar on any of the above windows. The bottom of the menu lists all currently open windows. To practice, make the *Data Editor* window active and click *Window* in the menu bar. Notice that the *Viewer* window is not listed in the menu because no statistical procedure has been run yet.



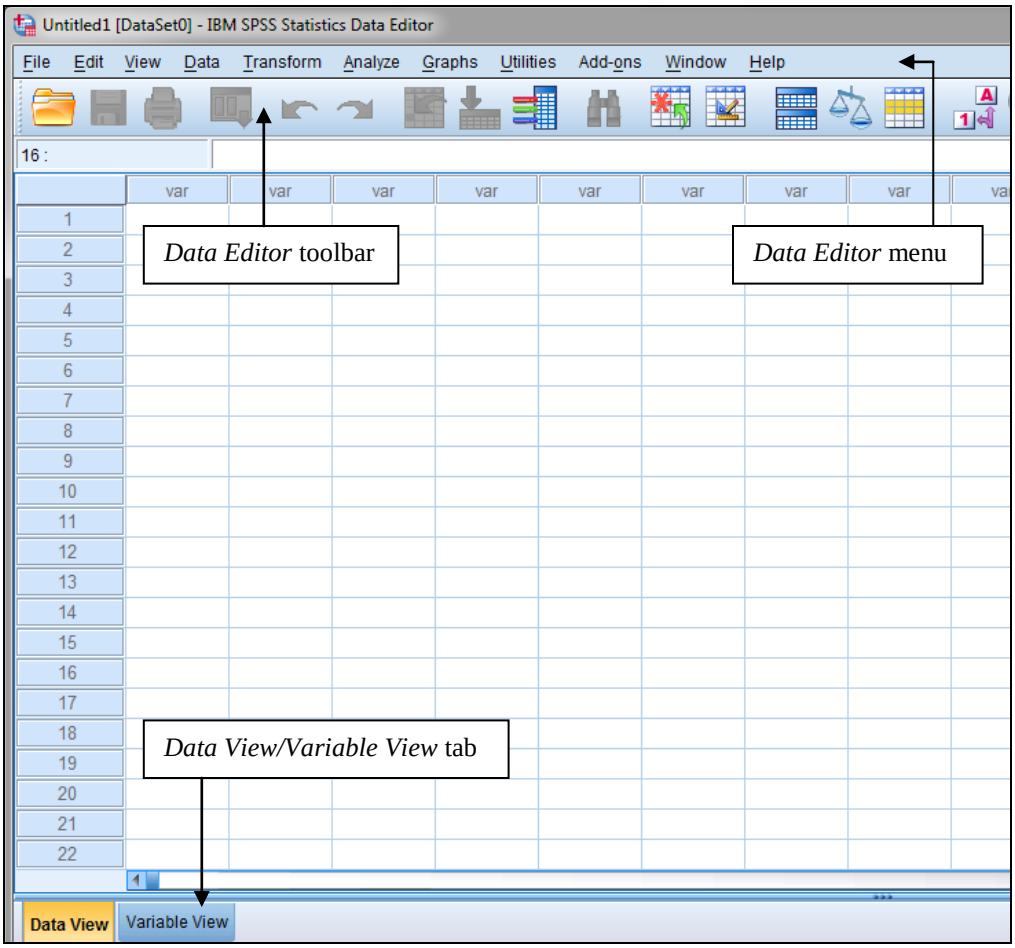
The active window is marked with a check mark. After running any statistical procedure or producing a graphical output, the Viewer window will be added to the list of windows in the Window menu.

If you want to keep the active cell where it is but view another part of the window, use the scroll arrows along the right and bottom sides of the workbook window. To practice, click the arrow in the direction you want to move in the *Data Editor* window.

Then click the down scroll arrow in the verical scroll bar. The worksheet scrolls down one row. Then click the up scroll arrow in the vertical scroll bar. The worksheet scrolls up one row. Similarly, the worksheet scrolls left by one column by clicking the left scroll arrow in the horizontal scroll bar.

5. Elements of Data Editor Window

The *Data Editor* window opens automatically when you start an SPSS session. It provides a simple, spreadsheet-like method for creating and editing data files. The most important components of the *Data Editor* window are menus, toolbar, and status bar. The components are displayed in the picture below:

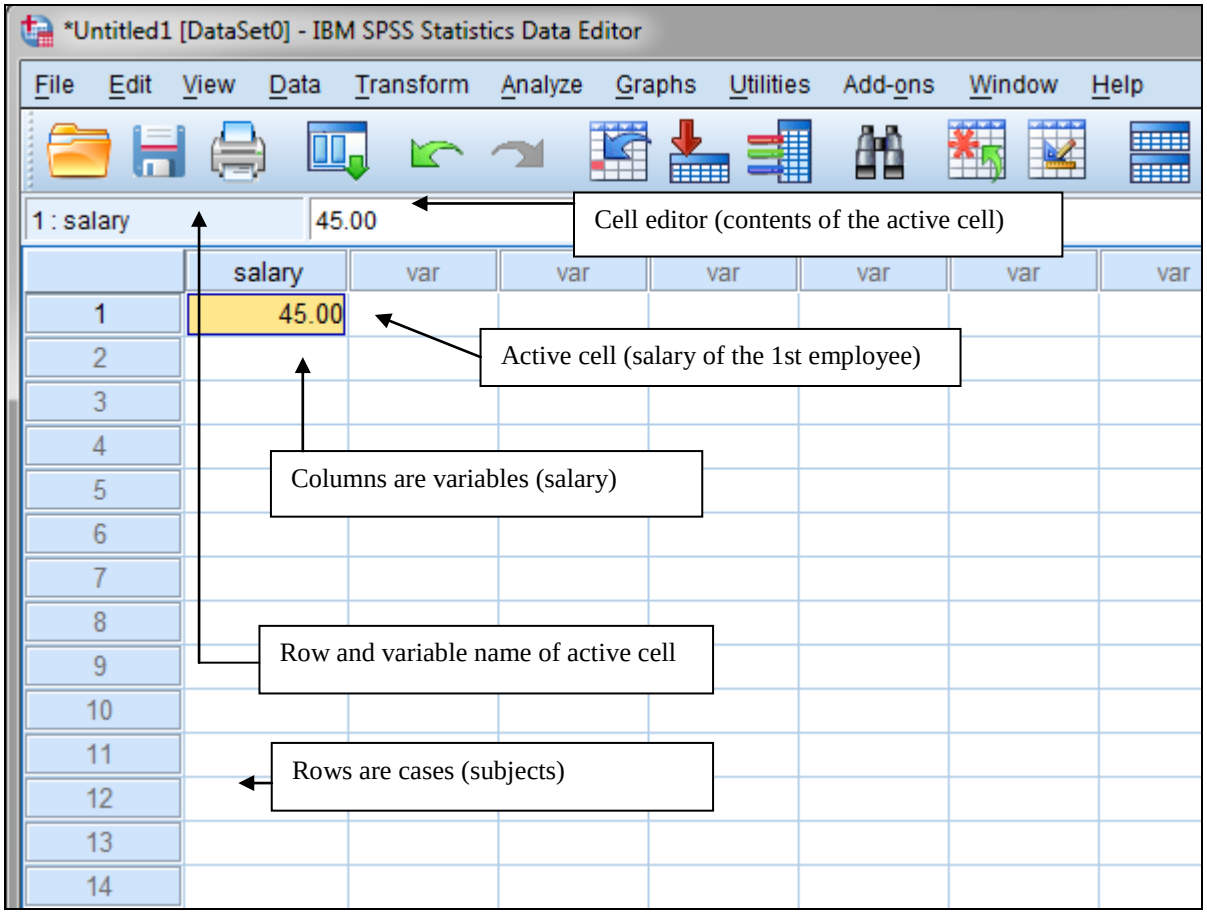


The *Data Editor* window can be displayed in one of the two views: *Data View* or *Variable View*. The *Data View* displays the contents of the data file in the form of a spreadsheet. The *Variable View* defines all variables in the data file. Switching from one view to the other can be done by clicking the appropriate tab (*Data View* or *Variable View*) at the bottom of the *Data Editor* window (see the picture above).

The *Data View* window is a grid, whose rows represent subjects (or cases) and whose columns contain values of the variables (gender, salary, age etc.) for each subject. Each cell of the grid, therefore, will usually contain the score of one particular subject on one particular variable. For example, the salaries of employees in a company can be presented in a column, and then each employee is a case. In general, each column represents a variable or characteristic that is being measured.

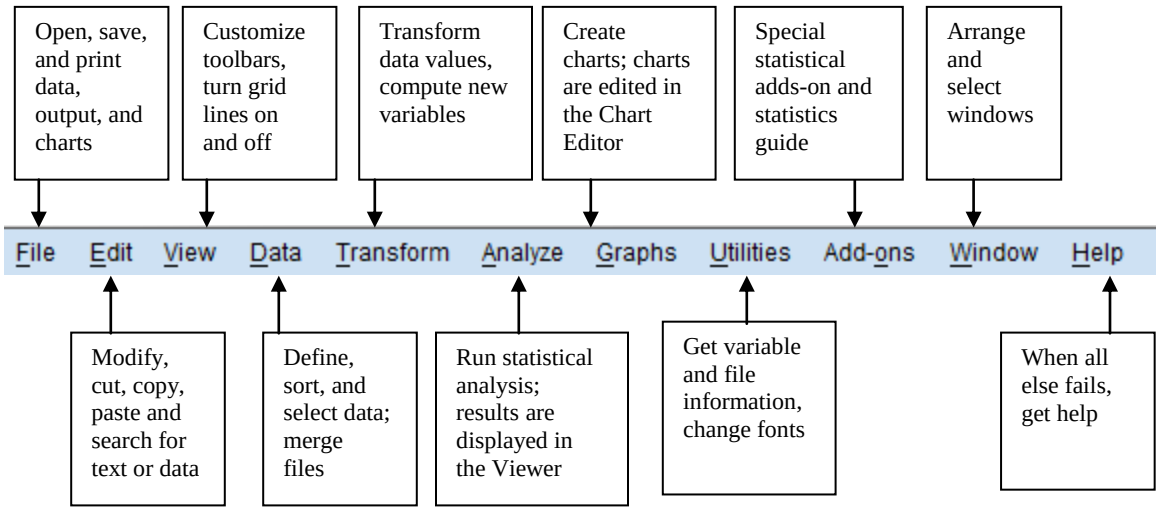
The cell is the intersection of the case and the variable. Cells contain only data values. Unlike spreadsheet programs (Excel, Lotus), cells in the *Data Editor* cannot contain formulas. You can enter data in any cell.

The data file is rectangular. The dimensions of the data file are determined by the number of cases and variables. Initially, every column in the *Data Editor* has the heading *var*, and all the cells are empty.



(a) Data Editor Menus

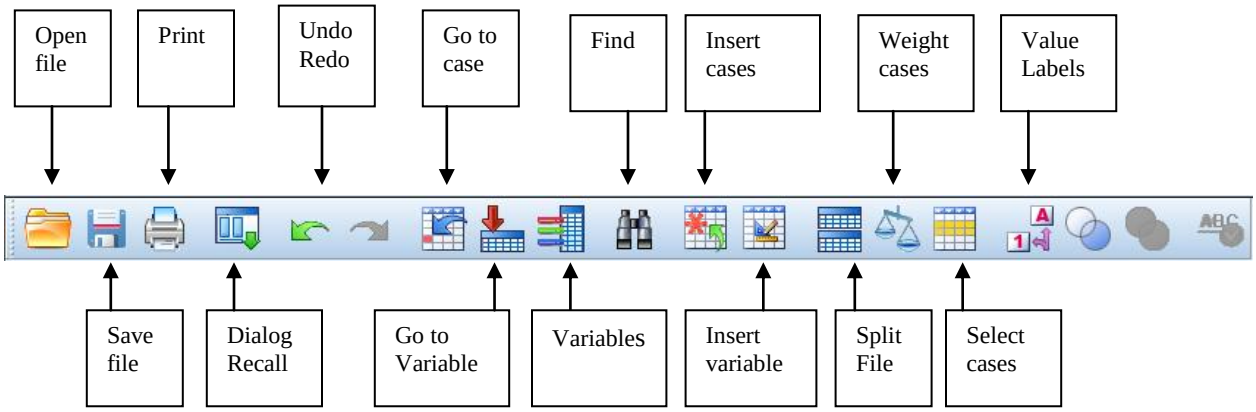
The menu bar provides easy access to most SPSS features. It consists of ten drop-down menus:



(b) Data Editor Toolbar

The toolbar provides quick and easy access to many useful features that you may use frequently. SPSS displays a toolbar below the menu bar on the *Data Editor* window. Clicking once on any of these buttons allows you to perform an action, such as opening a data file, or selecting a chart for editing.

In order to determine the function of a tool, place the mouse pointer over the corresponding button, but don't click the mouse button. SPSS displays a brief description of the tool in the *Status Bar*.



(c) Status Bar

A status bar at the bottom of the SPSS application window indicates the current status of the SPSS processor. If the processor is running a command, it displays the command name and a case counter indicating the current case number being processed. When the statement *SPSS Processor is ready* appears in the Status Bar, SPSS is ready to receive your instructions. The status bar also provides information such as command status, filter status, weight status, and split file status.

6. Working with the Data Editor

We will demonstrate the basic SPSS features using the following example.

Example The Framinhgham Heart Study followed a cohort of 5209 men and women for over 25 years. The study has been important in identifying risk factors associated with cardiovascular disease. The following is a description of the variables we have selected from the study for our purpose:

Column	**Description of Variable**
1	Sex (Gender: M-Male, F-Female),
2	Age (30-64 years),
3	Systolic blood pressure (82-300 mm).

The table below contains data for a random sample of 28 subjects from the study. The data file can also be downloaded from *Statistics 252 or 337 Labs* web site by clicking on *Data* in the *Introductory Lab* panel.

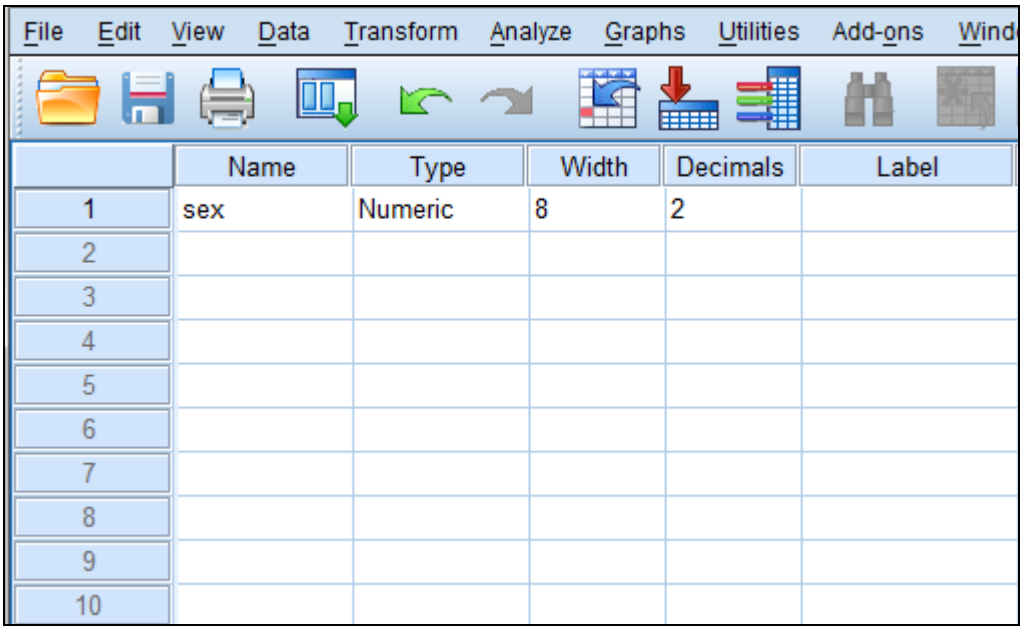
Number	**Sex**	**Age**	**Systolic**	**Number**	**Sex**	**Age**	**Systolic**
1	F	59	170	15	F	61	156
2	M	35	130	16	F	49	170
3	M	46	136	17	M	32	120
4	F	43	96	18	F	54	162
5	M	53	120	19	F	33	110
6	M	50	110	20	F	41	145
7	M	33	100	21	M	56	134
8	M	57	145	22	F	36	104
9	F	41	132	23	M	56	126
10	F	40	112	24	F	40	100
11	M	54	140	25	F	37	116
12	M	53	148	26	M	38	132
13	F	53	165	27	F	46	132
14	M	49	100	28	M	38	124

Now you will enter the data into the *Data Editor* window. Do not enter the names of the variables at the top of each column yet. Follow the instructions below.

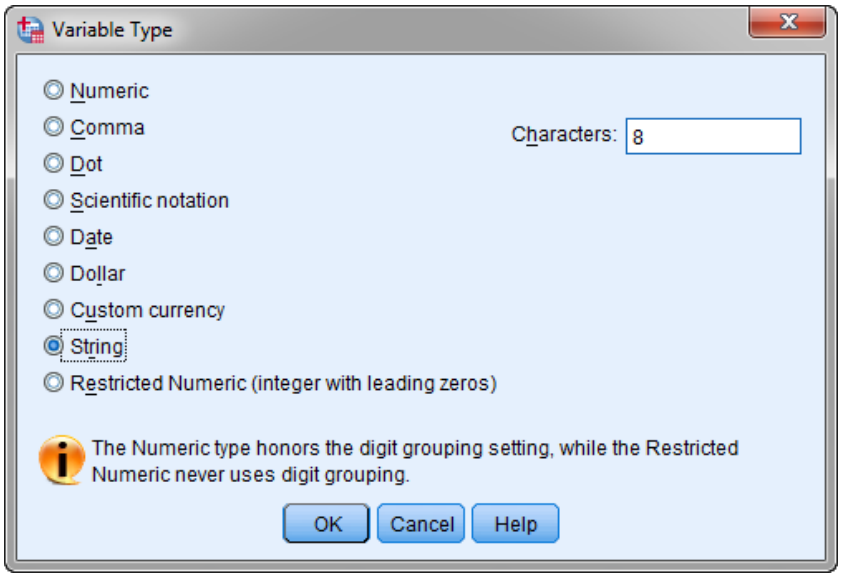
(a) Variables:

The variable name must begin with a letter and cannot end with a period. The length of the name cannot exceed 8 characters. To define a variable make the *Variables View* the active window (click the *Variable View* tab at the bottom of the *Data Editor* window). This will obtain the *Variable View* window.

Enter the new variable name in the column *Name* in any blank row. For example, enter the name *gender* in the first row. After entering the name, the default attributes (*Type*, *Width*,...) are automatically assigned.



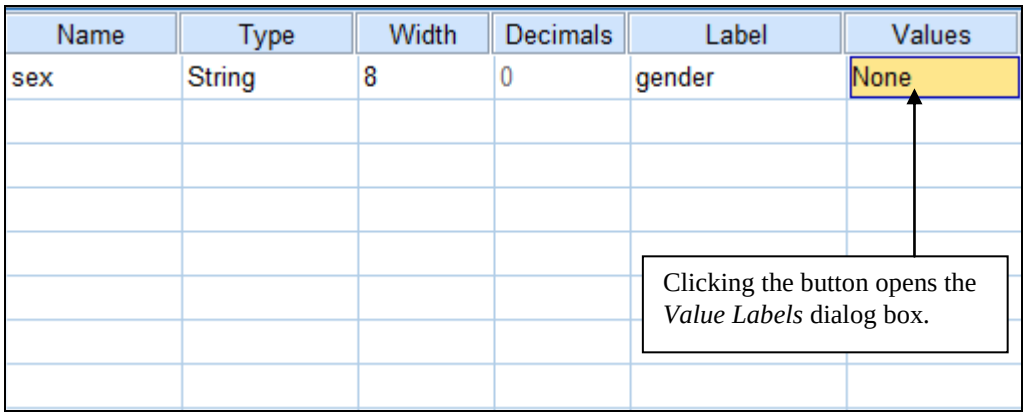
Clicking the cell *Numeric* and then the button in the cell opens the *Variable Type* dialog box.



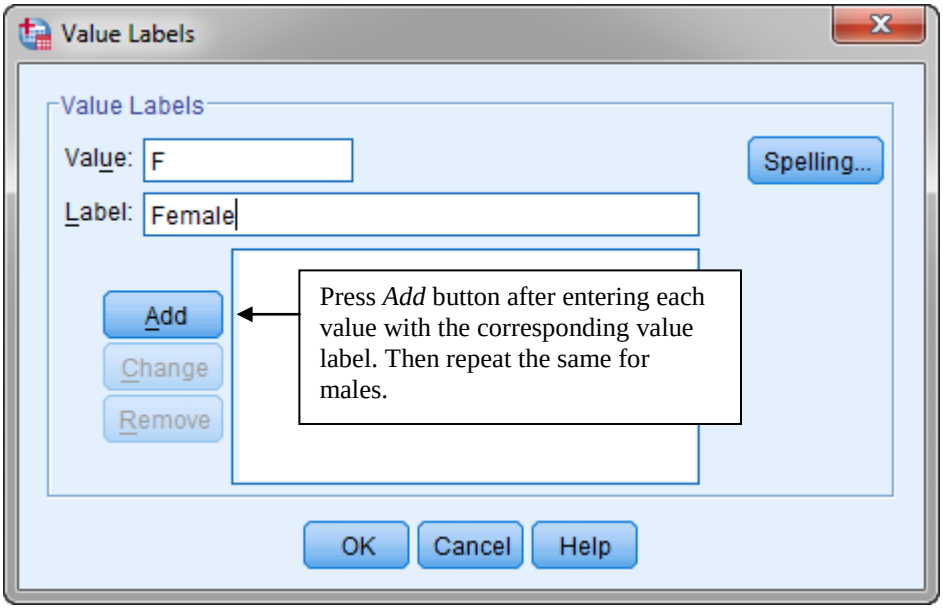
As *sex* is a categorical variable, click the radio button for *String*.

In our example there are three variables: sex (categorical), age (numeric), and systolic blood pressure (numeric). You may provide a description of the variable listed in each row of the Viewer window in the Label column. For example, we may assign the label *gender* to the *sex* variable. Enter gender in the *Label* column corrresponding to the variable *sex*.

To define possible values of the variable *sex* (possible values M for male and F for female) click the *Values* cell in the row for the variable, and then click the button in the cell.



Define the value labels for the variable *sex* as follows:



In the same way enter the remaining variables *Age* and *Systolic*. *Age* should be defined as a numeric variable with two digits (to minimize the chances of transcription error) and *Systolic* as a numeric variable with 3 digits.

To delete a variable (row), select the row number that you wish to delete, click *Edit*, and then on *Clear*. The selected variable will be deleted and all variables to the right of the deleted variable will shift to the left. Alternatively, you can select the row and press *Delete* key on your keyboard.

To insert a new variable (row) between existing variables: click on the row that is below the row where you wish to enter a new variable, click *Data* on the menu bar, and then click *Insert Variable* from the pull-down menu.

(b) Entering Data

Switch from the *Variables View* window to the *Data View* window. The three variables *gender*, *age*, and *systolic* are represented as columns. Now we are ready to enter the values of the three variables from the Framingham Heart Study on page 10.

Each row represents a case or an observation. For example, the gender, age, and systolic blood pressure of a particular subject in the Framingham Heart Study data file is recorded as one row.

Clicking any cell will highlight it (active cell) and its contents will appear in the *cell editor*. You can enter the data in any order. Data values are not recorded until you press *Enter* or select another cell. Unlike spreadsheet programs, cells in the *Data Editor* cannot contain formulas.

Enter the values for all cases on one variable (column) and then repeat the procedure for all values in the remaining columns. You will learn how to save the data in Section 7.

(c) Editing Data

To delete the old value and enter a new value: click the cell, enter the new value, press *Enter*. To modify a data value: click the cell, click the cell editor, edit the data value, and press *Enter*. To delete the values in a range, select (highlight) the area concerned and press *Delete*.

Use the *Undo* command in *Edit* to undo any action you just performed. For example, use the *Undo* command to delete the value you have just entered in the *Data Editor* window.

(d) Adding Cases

To insert a new case (row) in between cases that already exist in your data file: click the row below the row where you wish to enter the new case,click *Data* on the menu bar, click *Insert Case* from the pull-down menu.

(e) Deleting Cases

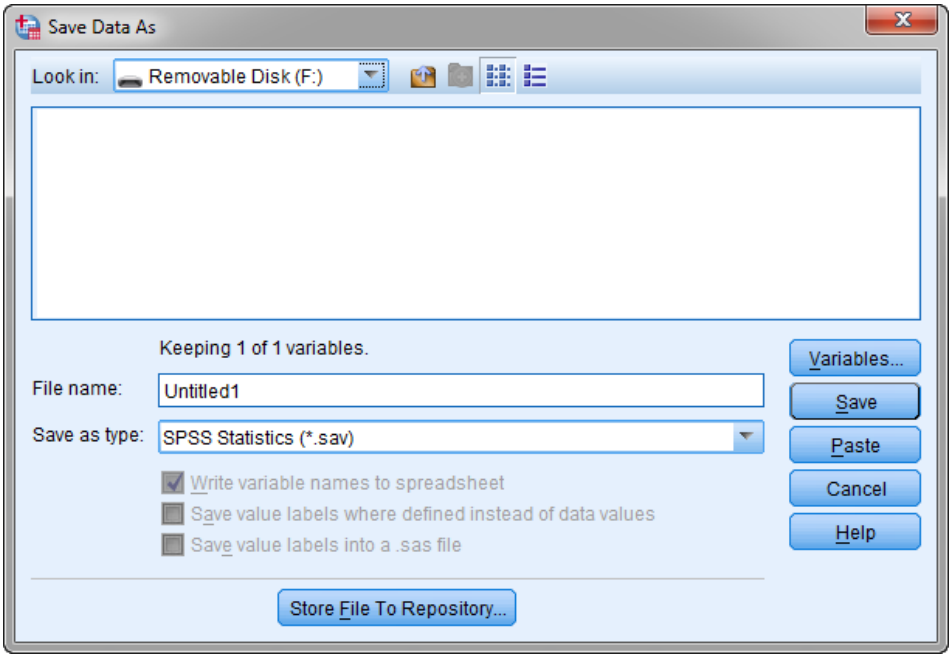
To delete a case, click the case number that you wish to delete, click *Edit* from the menu, and then on *Clear*. The selected case will be deleted and the rows below will shift upward.

7. Saving and Reading Data Files in SPSS

We will first demonstrate how to save the data from our example, and then how to open an existing data file.

7.1 Saving Data Files

To save a new SPSS data file or save data in a different format make the *Data Editor* the active window and from the main menu choose *File* and then *Save As*…. The following *Save Data As* dialog box is displayed:

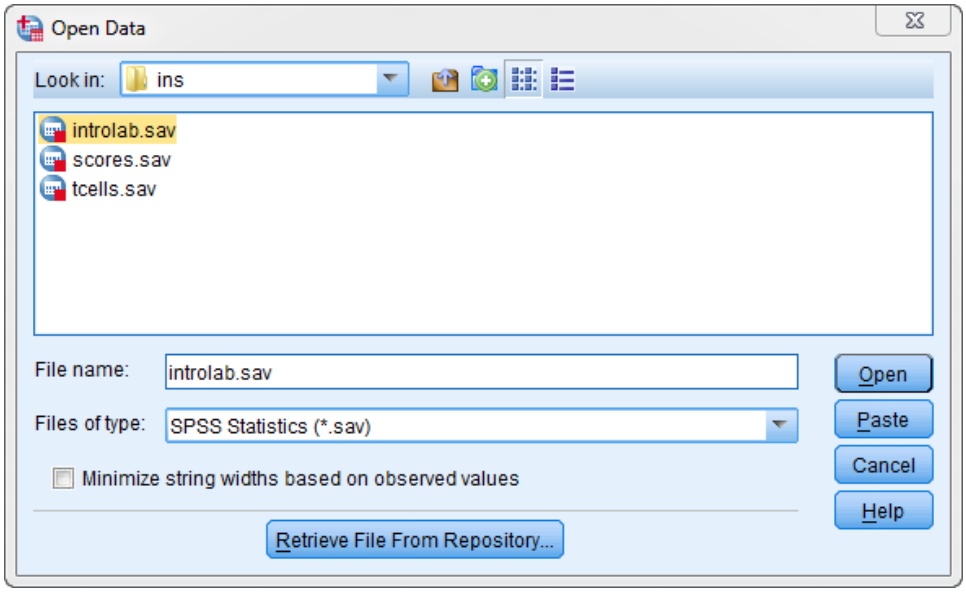


Choose the appropriate directory in the *Look in:* box to save your file. Then type the name of the data file in the *File name:* box. No extension (i.e. a dot followed by three letters) is required. SPSS automatically adds the proper extension, which depends on the type of file.

To save changes to an SPSS data file make the *Data Editor* the active window and from the menus choose *File* and then *Save*. The modified data file is saved, overwriting the previous version of the file. By default, this will save the data file as an SPSS data file.

7.2 Reading SPSS Data Files

SPSS for Windows can read different types of data files. To read data files, click *File* in the menu bar, then on *Open*, and finally on *Data*. The *Open Data* dialog box is displayed:

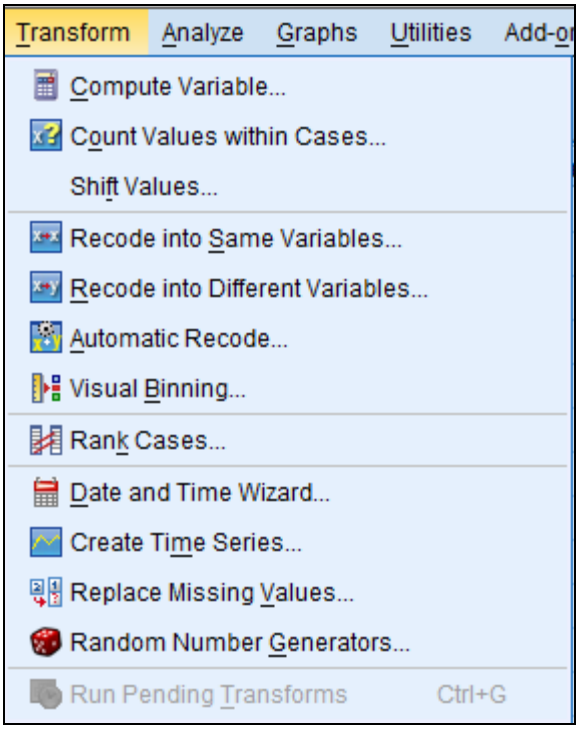


SPSS data files are easily identified since by default each file name is followed by ".sav" extension. SPSS data files contain not only the actual data but also some information about the data such as variable names and formats. These files are written in a special code that is read and interpreted by the SPSS program.

To read an SPSS data file, click *File* in the menu bar, and then click *Open*. This opens the *Open File* dialog box. Point the arrow to the data file you wish to open and click on it. If necessary use the up and down arrows to scroll through files until locate your file. Click OK.

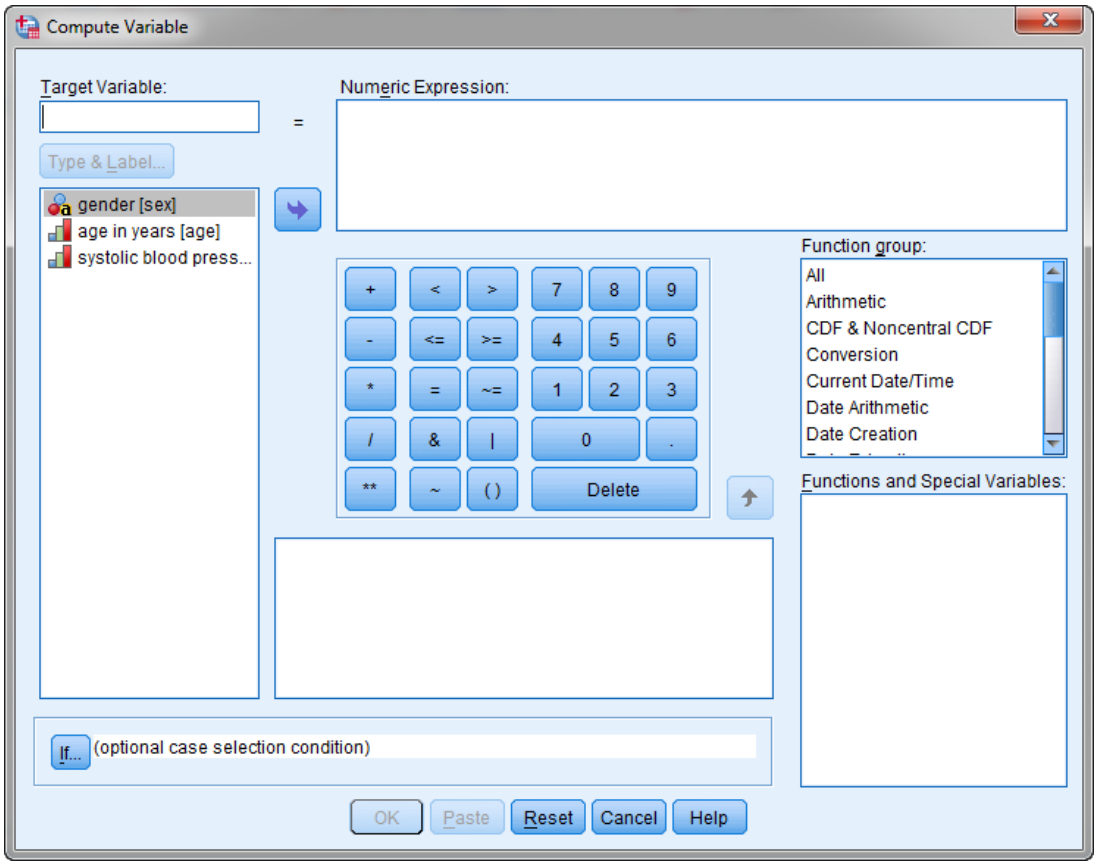
8. Data Transformations

After a data set has been entered into SPSS, it may be necessary to modify it in certain ways. With SPSS, you can perform data transformations ranging from simple tasks, such as combining categories for analysis, to more advanced tasks, such as creating new variables based on complex equations.



8.1 Computing New Variables

To create a new variable click *Transform* in Data Editor menu, and then on *Compute* from the pull-down menu. This opens the *Compute Variable* dialog box.



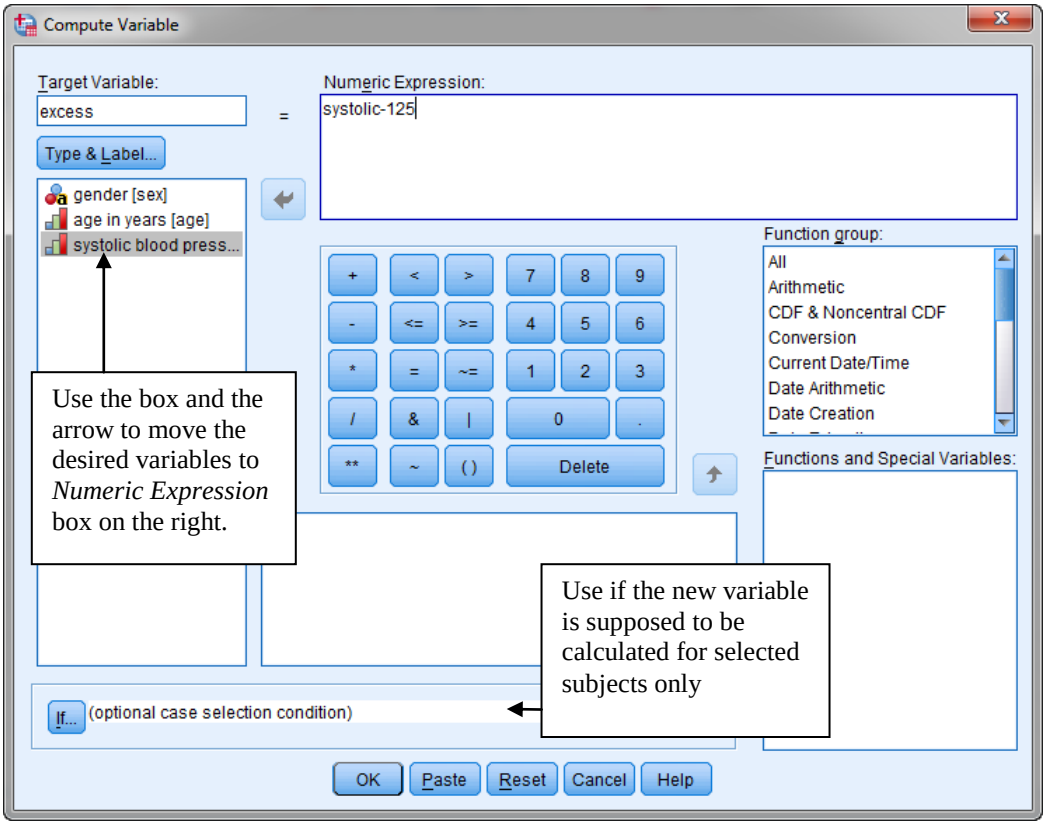
Numeric expresion can be typed directly or assembled by clicking arrows in the *Variable* and *Function group* boxes. Observe that there is a numerical variable icon (of a histogram shape) at the variable *age* and the variable *systolic*. In fact, all numeric variables (*age*, and *systolic* are numeric) are identified with the icon. On the other hand, all string variables (i.e. sex) are identified by a categorical variable icon (two circles) with the letter a.

For information about a variable, click the left mouse button on the variable name to select it, and then the right mouse button and choose *Variable Information* from the pop-up menu.

Enter the name of the new variable in the *Target Variable* box. To build an expression, either paste components into the *Expression* field or type directly in the *Expression* field. The *If...* dialog box allows you to apply data transformations to selected subsets of cases.

For example, to calculate the new variable EXCESS which is excess systolic blood pressure for males defined as excess = systolic - 125, your *Compute Variable* dialog box should look like the box shown below.

When you have completed the expression, click OK to close the *Compute Variable* dialog box. You will see the message "Running Execute" at the bottom of the application window indicating that SPSS is computing the new variable. When the computations are complete, this message will be replaced with "SPSS Processor is Ready" and your new variable will appear in the first empty column in the data editor window.



If you wish to obtain the excess systolic blood pressure for males over 45 only, you should click *If ...* button and enter "age>45" in the appropriate text box.

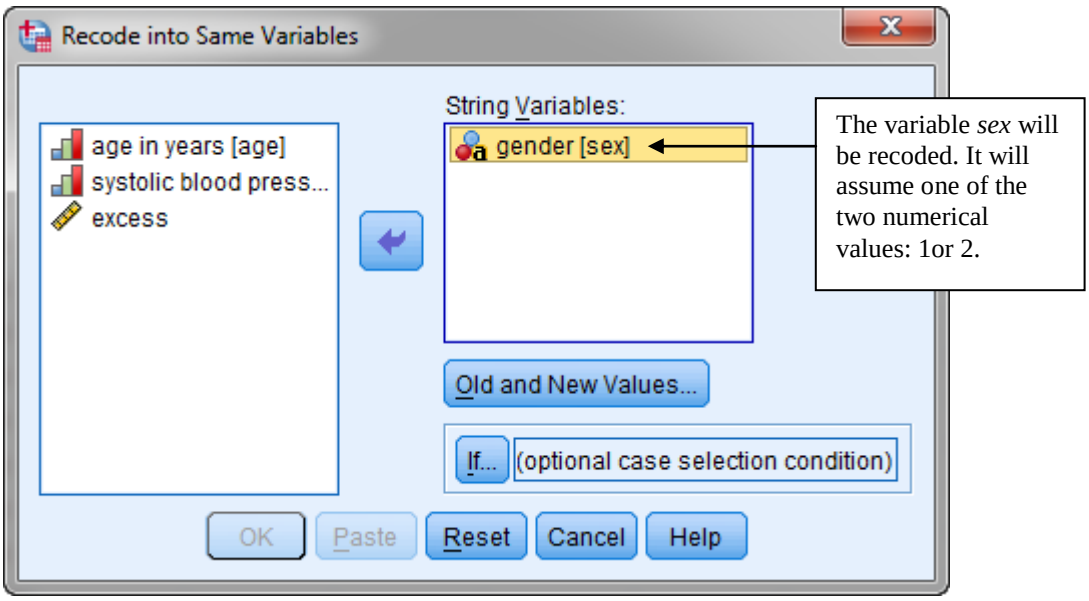
8.2 Recoding Variables

Some of the analyses to be performed in SPSS will require that a categorical variable be entered into SPSS as a numeric variable. If the variable has already been defined in SPSS as a string, you can easily create a new variable that contains the same information as the string variable but is numeric simply by recoding the variable. For example for the Framingham Heart study data, you might want to recode the variable *sex* into a numeric variable, called *sexnum*, by assigning F=1, and M=2.

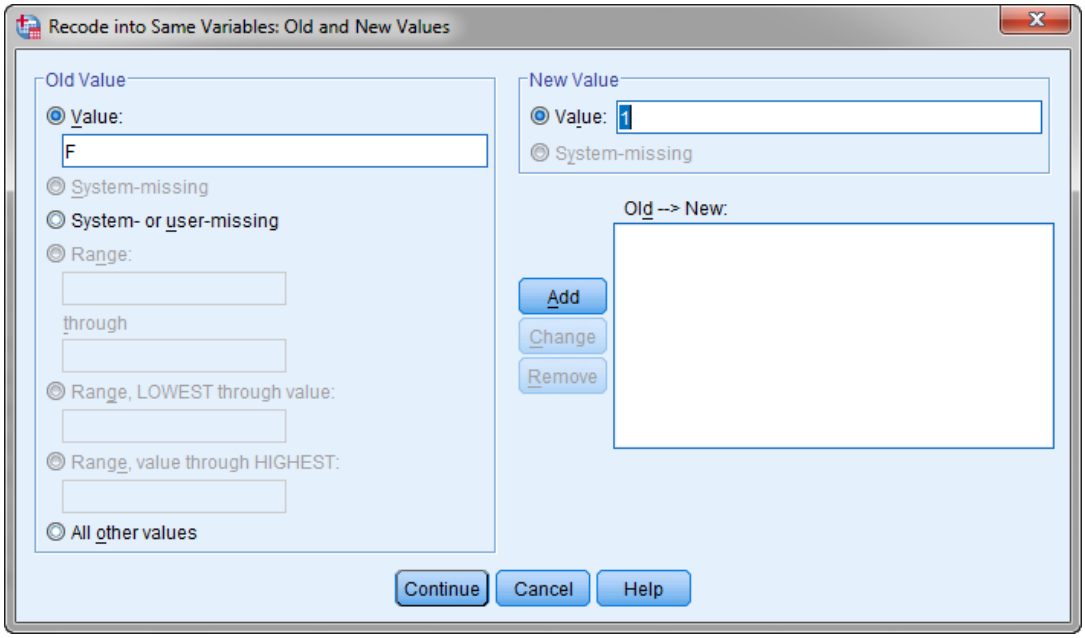
You have two options available for recoding variables. You may recode values into the same variable, which eliminates all record of the original values. You also have the option to create a new variable containing the recoded values. This preserves the original values of the variable.

(a) Recoding into the Same Variable

To recode into the same variable, click *Transform* from the main menu, then on *Recode* from the pull-down menu, and finally on *Into Same Variable*. This opens the *Recode into same variable* dialog box. For information about a variable, click the right mouse button on the variable name. Then select the name of the variable to be recoded (*sex*), and move it to the *String Variables* box with the right arrow button.



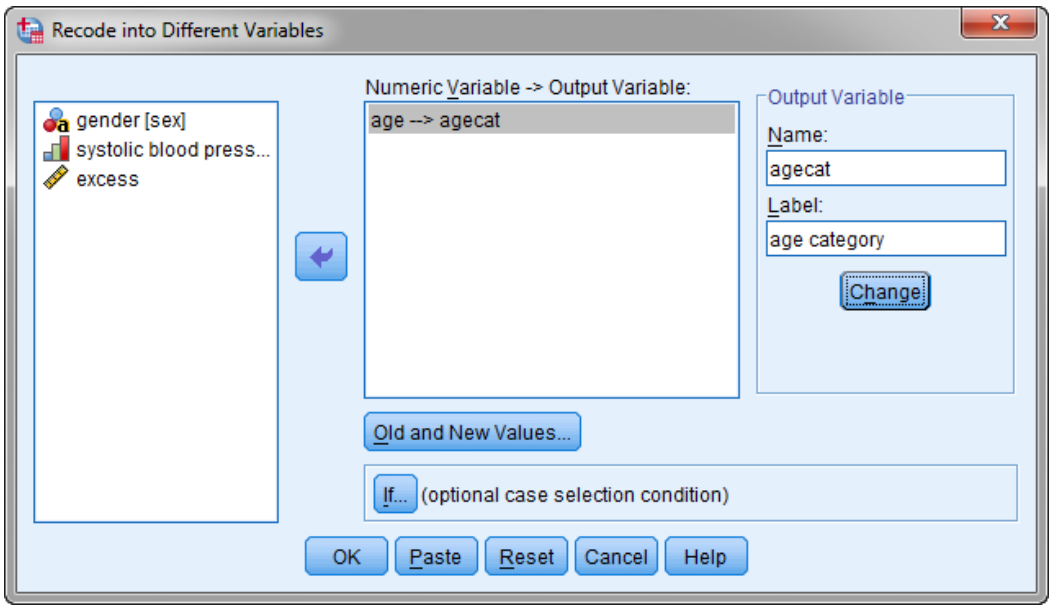
Then click *Old and New Values*. You will obtain the following box:



The old value of the variable *sex* is "F", and the new value is "1". Then click *Add* tab to recode the old value and its new value. Similarly enter "M" as the old value, "2" as the new value, and click *Add* tab. When you have indicated all the recode instructions, click *Continue* to close the above dialog box. Then click OK to close the *Recode Into Same Variables* dialog box. Now *sex* no longer is expressesed as either "F" or "M", but it is one of two integers 1, or 2. Close the file without saving the changes you have made and retrieve the original data file *introlab.sav*.

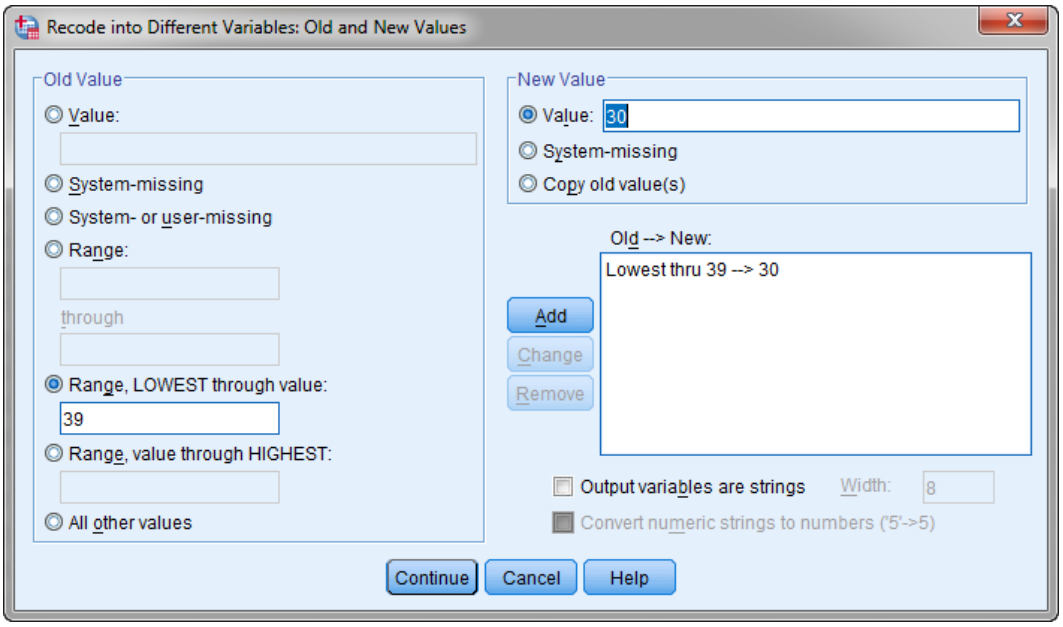
(b) Recoding into Different Variables

To recode into a different variable, click *Transform* from the main menu, then on *Recode* from the pull-down menu, and finally on *Into Different Variable*. This opens the *Recode into Different Variables* dialog box.



We might use the numerical variable age to define a new categorical variable *agecat* with the following four age categories (strings): "30" (30s), "40" (40s), "50" (50s), and "60" (60s). In order to achieve that select the variable *age* in the left panel and move it to the *Input Variable* → *Output Variable* entry box. In the *Output Variable* right panel, enter the name of the output variable *agecat* and the corresponding label: *age category*. Click the *Change* tab.

Click *Old and New Values*. Fill out the corresponding dialog box as shown below:

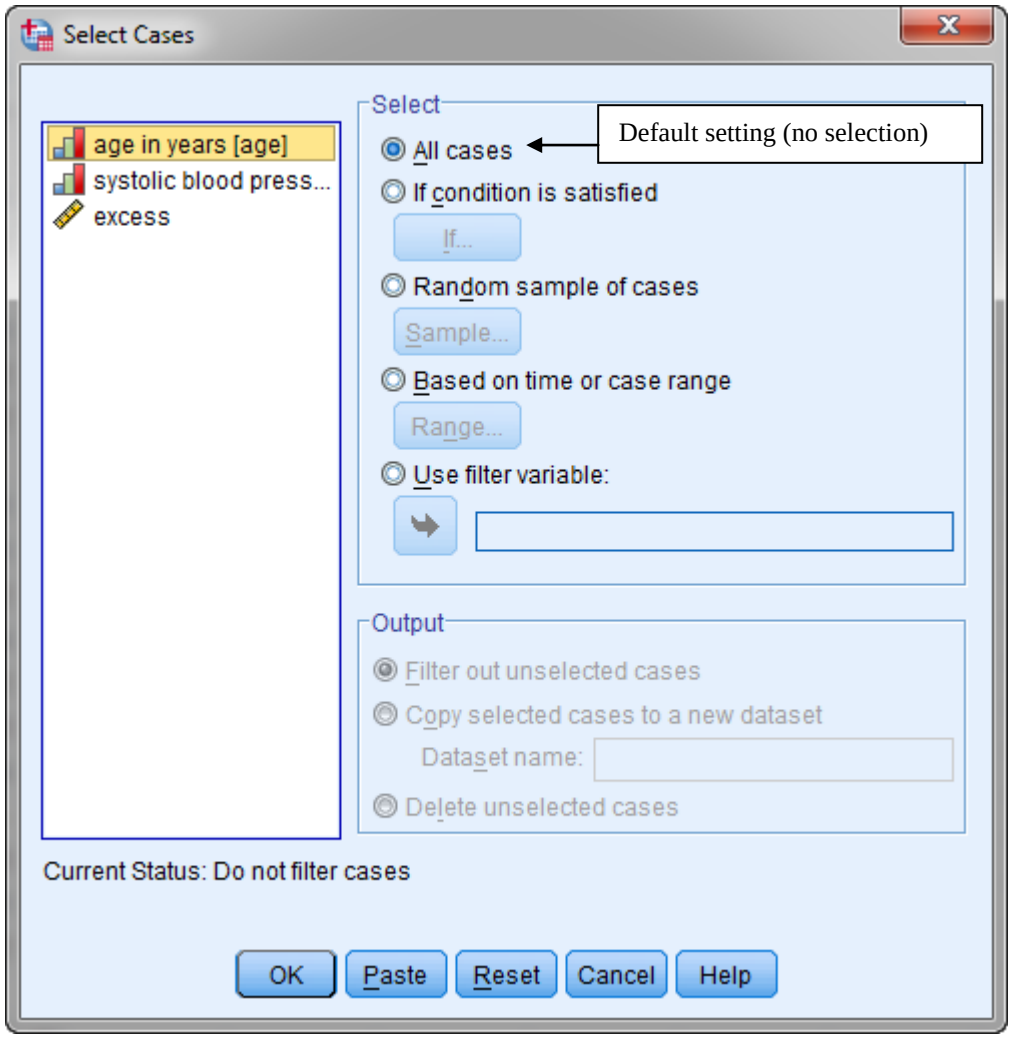


Click the *Add* tab to recode the range to its new value "30s". Repeat the procedure by entering in the *Range, LOWEST through value* entry box the values "49" (to recode to 40s), "59" (to recode to 50s), and "69" (to recode to 60s). Click OK. The new column *agecat* will be added to the data file in the *Data Editor*.

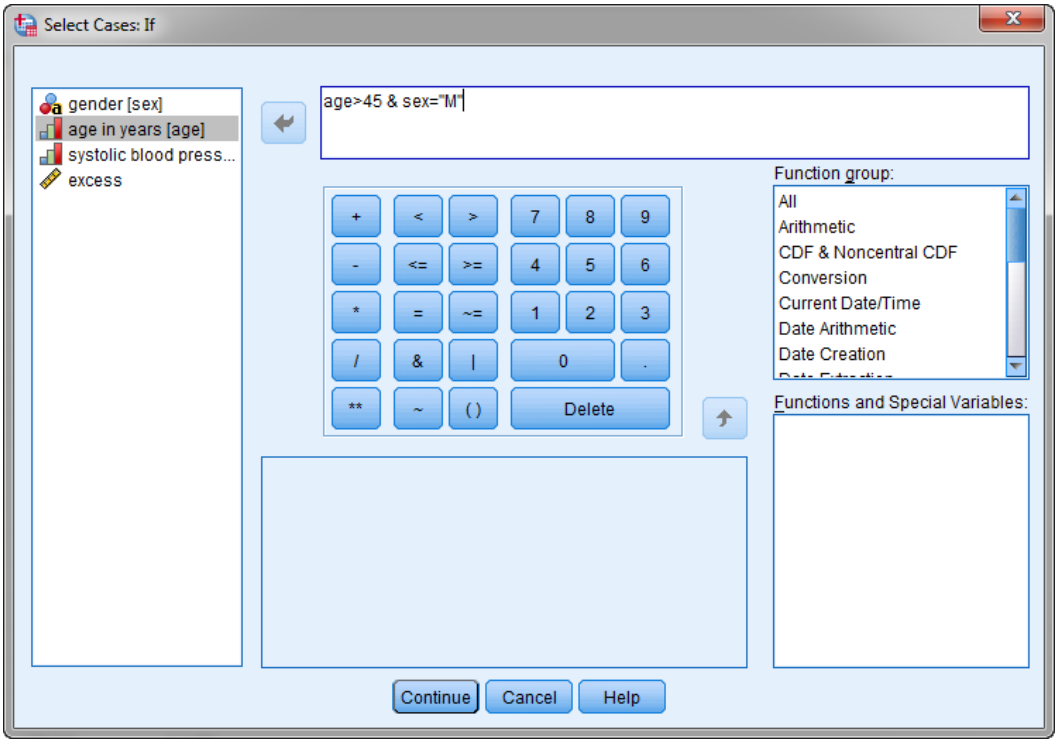
8.3 Selecting Cases

There are occasions on which you will want to select a subset of cases from your data file for a particular analysis. You may need to select the subset based on a formally defined criteria or randomly in case of a very large data file.

To select subset of cases, click *Data* in the main menu and then on *Select Cases* from the pull-down menu. This opens the *Select Cases* dialog box.



Click *If condition is satisfied* radio button, and then on *If…*tab. The *Select Cases If* dialog box will be displayed. Supose we want to select all mean over 45 years old in the data. In order to enter the appropriate logical expression ("age>45 & sex="M""), highlight the variable age and click the arrow to transfer the variable into the upper box on the right (expression box). Then complete the expression by using the buttons in the keyboard located below the expression box. The dialog box should be filled out as shown on the next page:

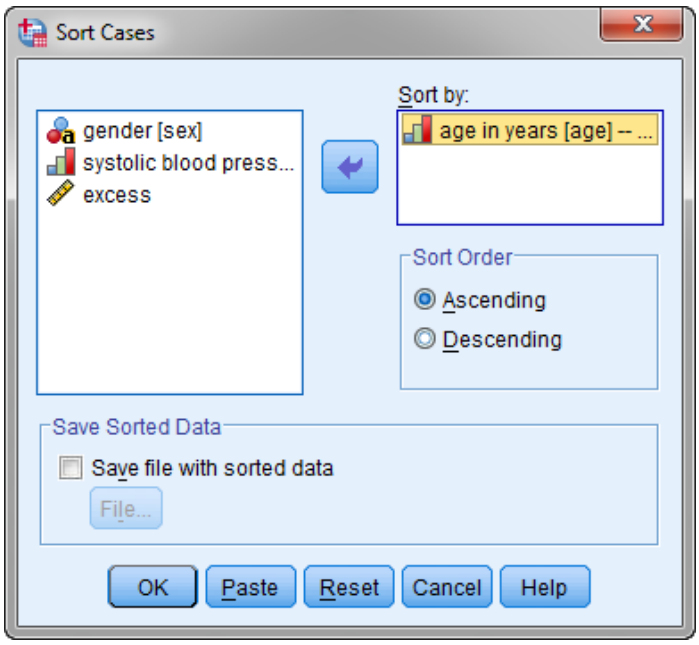


When the condition has been completed, click *Continue* and then on *OK* to return to the *Data Editor* window, where it will be noticed that a new column labeled filter_$, and containing 1s and 0s, has appeared. The 1s and 0s represent the selected and unselected cases, respectively. The row numbers of the unselected cases have also been marked with an oblique bar. This is a useful indicator of case selection status. Any further analyses of the data set will include the selected cases only. The case selection can be cancelled by returning to the *Select Cases* dialog box, clicking *All cases*, and then on OK.

	sex	age	systolic	filter_$
1	F	59.0	170.00	0
2	M	35.0	130.00	0
3	M	46.0	136.00	1
4	F	43.0	96.00	0
5	M	53.0	120.00	1
6	M	50.0	110.00	1
7	M	33.0	100.00	0

8.4 Sorting Data

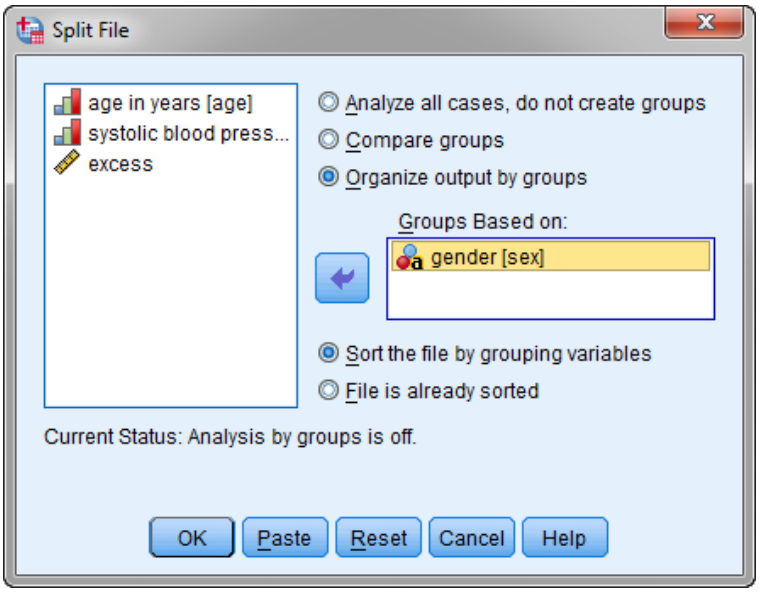
Suppose that we would like to sort the data in the data file *introlab.sav* according to the age of the subjects enrolled in the study. In order to sort the data, from the menus choose *Data*, and then *Sort Cases*. The following dialog box will be displayed:



In order to sort the subjects according to the age, select *age* and move it to the *Sort by* box. You can sort cases in ascending or descending order. If you select multiple sort variables, cases are sorted by each variable within category of the prior variable on the *Sort by:* list. For example, if you select *gender* as the first sorting variable and *age* as the second sorting variable, cases will be sorted by age classification within each gender category.

8.5 Splitting File

Split File feature can be accessed from the *Data* menu. The feature split the data file into separate groups for analysis based on the values of one or more grouping variables. If you select multiple grouping variables, cases are grouped by each variable within categories of the prior variable. For example, if you select sex as the first grouping variable and agecat as the second grouping variable, cases will be grouped by age category classification within each gender group.



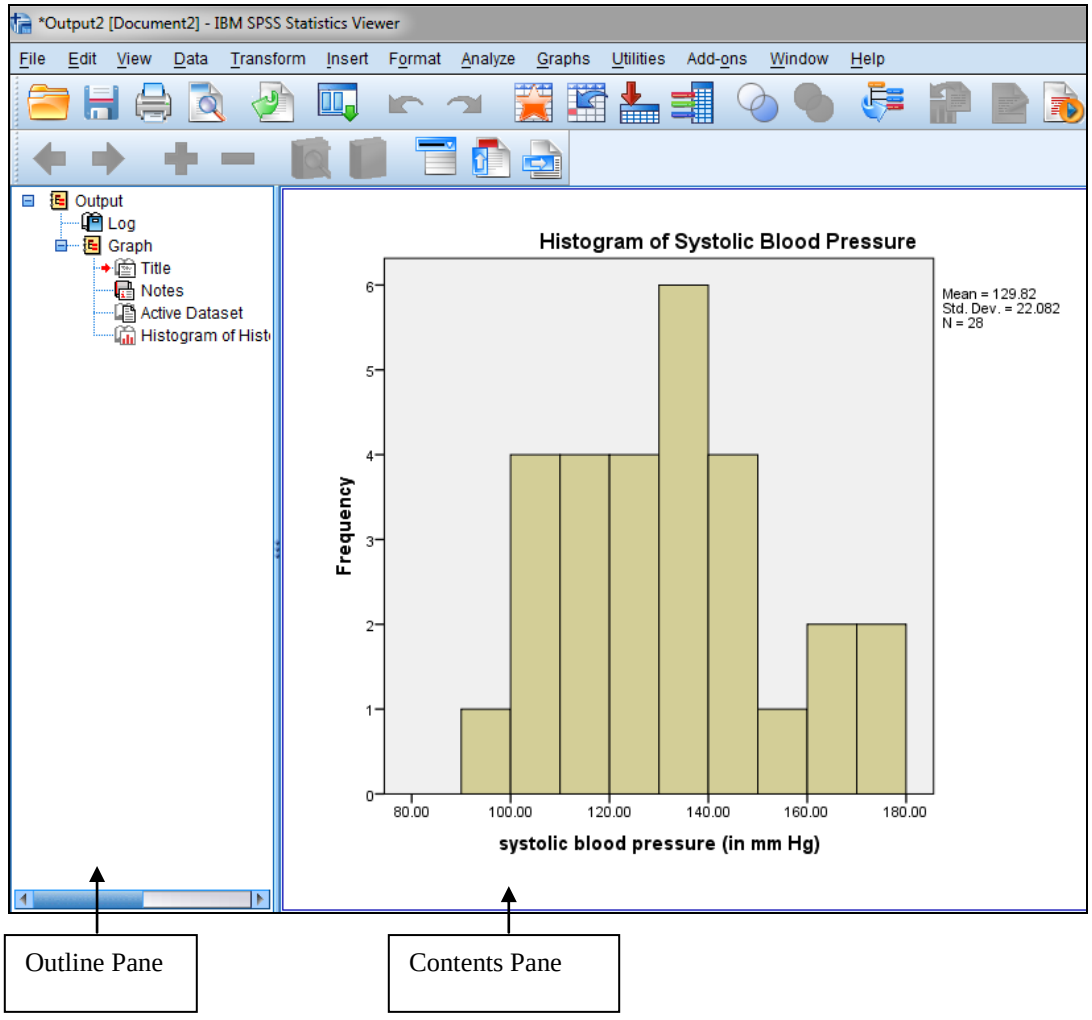
If you check *Organize output by groups* radio button, all results from each statistical procedure will be

displayed separately for each split-file group (in this case gender).

9. Working with the Viewer

All statistical results, tables, and charts are displayed in the *Viewer*. A *Viewer* window opens automatically the first time you run a procedure that generates output. In this window, you can easily navigate to whichever part of the output you want to see. You can also use the *Viewer* to show or hide selected tables or results from an entire procedure. This is useful when you want to shorten the amount of visible output in the contents pane.

The *Viewer* is divided into two panes: the left pane of the *Viewer* contains an outline of all of the information stored in the Viewer (*Outline pane*), the right pane contains statistical tables, charts, and text output (*Contents pane*).



The output in this case has been obtained by running only one procedure *Frequencies* to the variables *Systolic* (*Descriptive Statistics* in the *Analyze* menu). It consists of the five items: *Title*, *Notes*, *Active Dataset*, *Statistics*, *Systolic*, and *Histogram*. You can click and drag the right border of the outline pane to change the width of the outline pane.

You can use the scroll bars in the display pane to browse the results or you can click an item in the outline pane to go directly to the corresponding table or chart. To practice, click each item in the outline pane to display it in the contents pane.

An arrow in front of an open book icon in the outline pane indicates that the corresponding item is currently visible in the contents pane.

An open book icon in the outline pane indicates that the object is currently visible in the *Viewer*. To hide a table or chart, double-click its book icon in the outline pane. The open book icon changes to a closed book icon, showing that the information associated with the item is now hidden. To redisplay the hidden output, double-click the closed book icon.

To change the position of tables or charts in the display, click on the items in the outline pane, and drag them where you want to put them. Try it. You can switch between the *Data Editor* and *Viewer* windows by clicking *Window* in either menu.

10. Missing Values

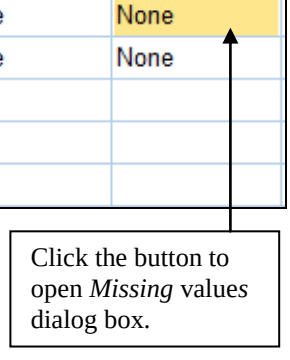
In many situations, data files do not have complete data on all variables, that is, there are missing values. You need to inform SPSS when you have missing values so that all computations are performed correctly. With SPSS, there are two forms of missing values: system-missing and user-defined missing.

System-missing values are those that SPSS automatically treats as missing. The most common form of this type of value is when there is a "blank" in the data file. For example, a value for a variable may not be entered in the data file if the information was not provided. When SPSS reads this variable, it will read a blank, and thus treat the value as though it is missing. Any further computations involving this variable will proceed without the missing information (computing the average without the missing value).

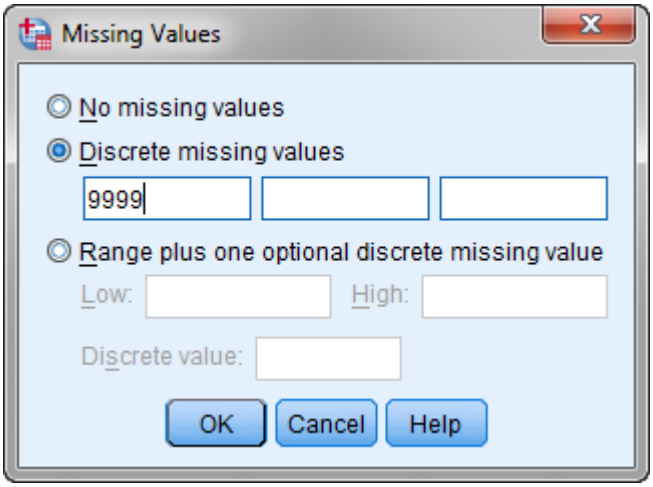
User-defined missing values are those that the user specifically informs SPSS to treat as missing. Rather than leaving a blank in the data file, numbers are often entered that are meant to represent data. For instance, if the systolic blood pressure for some subjects in our data set is unknown, we could use the number 9999 to represent those cases that were missing information on the variable.

You need to inform SPSS that 9999 is to be treated as a missing value, otherwise it will treat it as valid. More precisely, make the *Data Editor* the active window, select the *Variables View*, and click the *Missing* cel for the variable (systolic). A button in the cell appears.

Name	Type	Width	Decimals	Label	Values	Missing
sex	String	1	0	gender	{F, Female}...	None
age	Numeric	2	1	age in years	None	None
systolic	Numeric	3	2	systolic blood p...	None	None
excess	Numeric	8	2		None	None



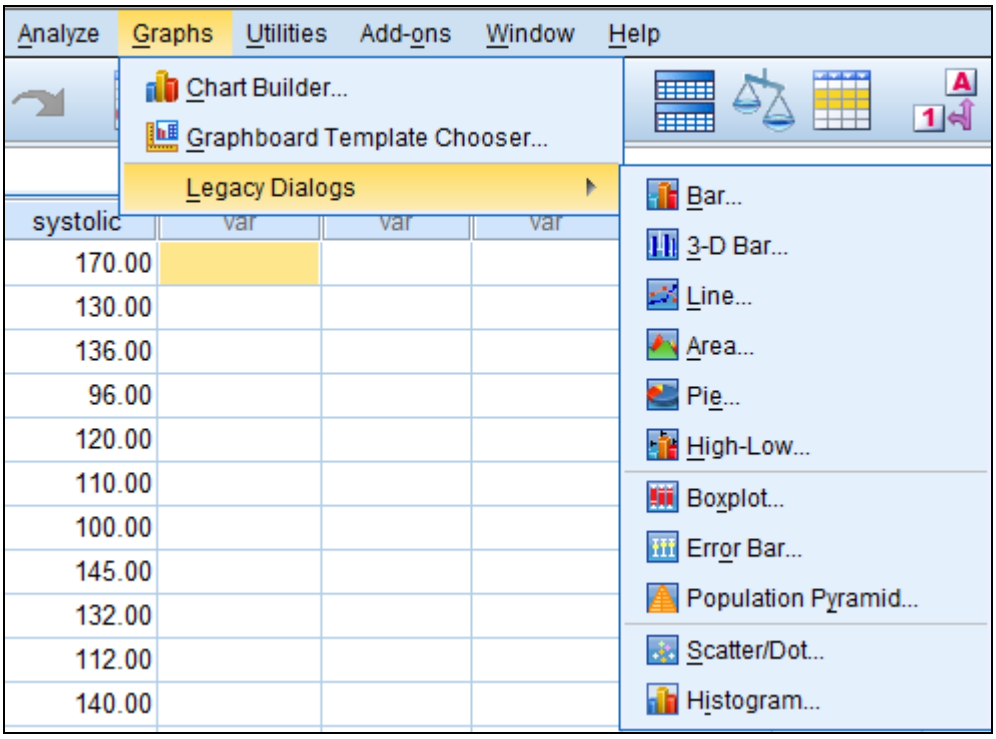
Suppose you define the missing values as displayed in the *Missing* values dialog box:



With this definition of the missing values for the variable *systolic*, SPSS will treat 9999 as a missing value of the variable and not include it in any computations involving the systolic blood pressure variable.

11. Creating and Editing Charts in SPSS

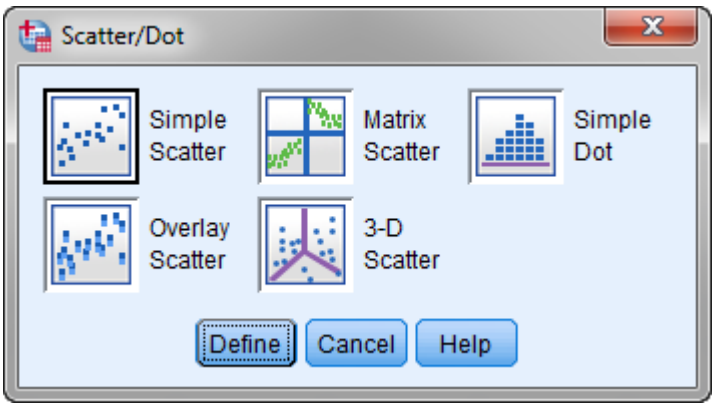
The charts in SPSS can be obtained either by using one of the descriptive procedures discussed above or directly by selecting *Graphs* in the main menu. The resulting pull-down menu is displayed below.



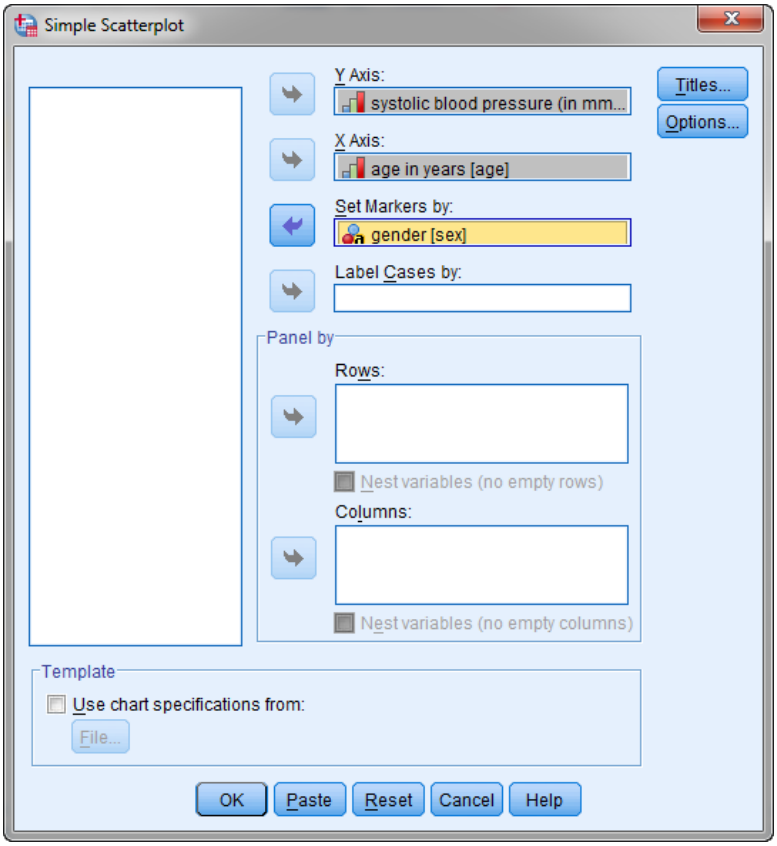
The *Graphs* menu contains two options to produce graphs in SPSS: Chart Builder and Legacy Dialogs. The Chart Builder dialog box is an interactive window that allows you to preview how a

chart will look while you build it. Most tasks can be accomplished simply by pointing, dragging selected graph components and clicking the mouse. The other option in the *Graphs* menu, the Legacy Dialogs allow you to access quickly the graphical procedure you need.

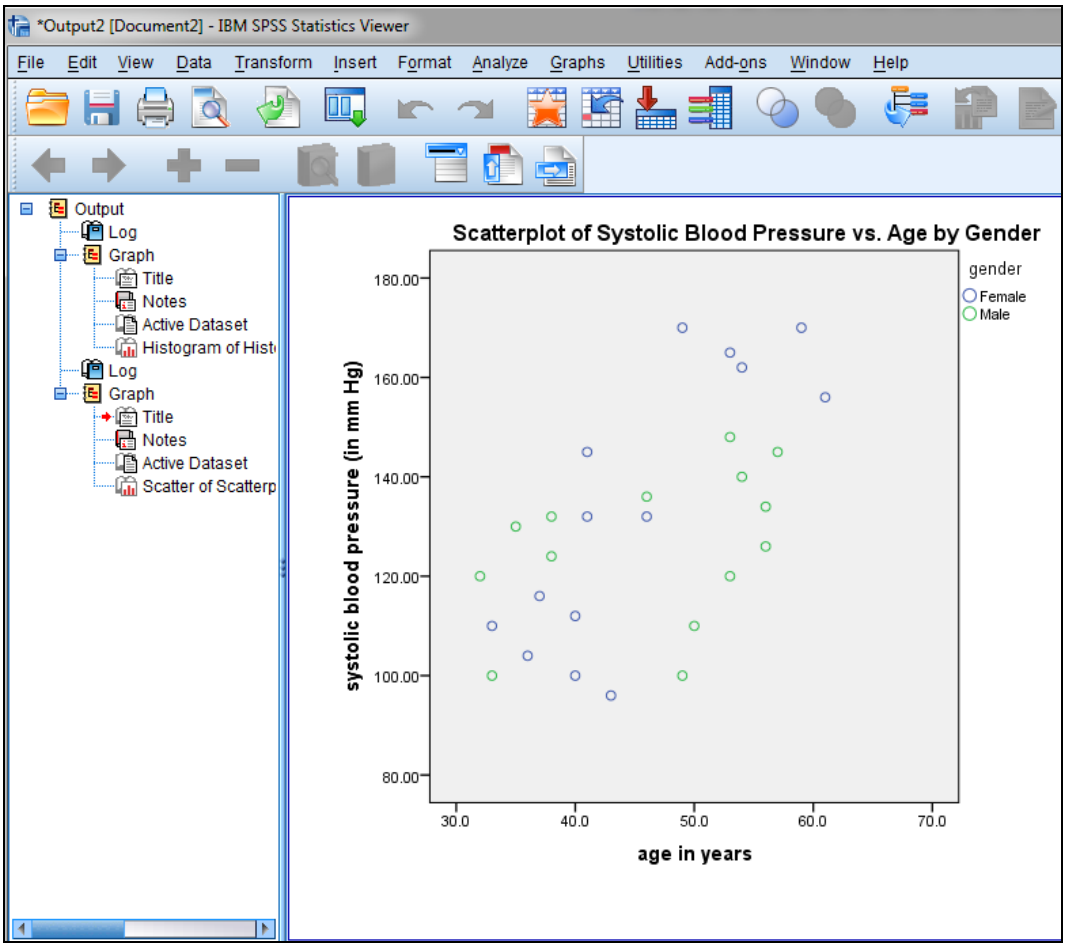
We will demonstrate obtaining and editing charts in SPSS using the scatterplot of the systolic blood pressure vs. age for males and females. Select *Graphs* in the menu, then *Legacy Dialogs* and finally *Scatter/Dot*.



Choose *Simple Scatter* as the form of the desired scatterplot in the obtained dialog box. Then click *Define*. You will obtain the *Simple Scatterplot* dialog box. In order to obtain a scatterplot of systolic blood pressure vs. age by sex, fill out the dialog box as follows:

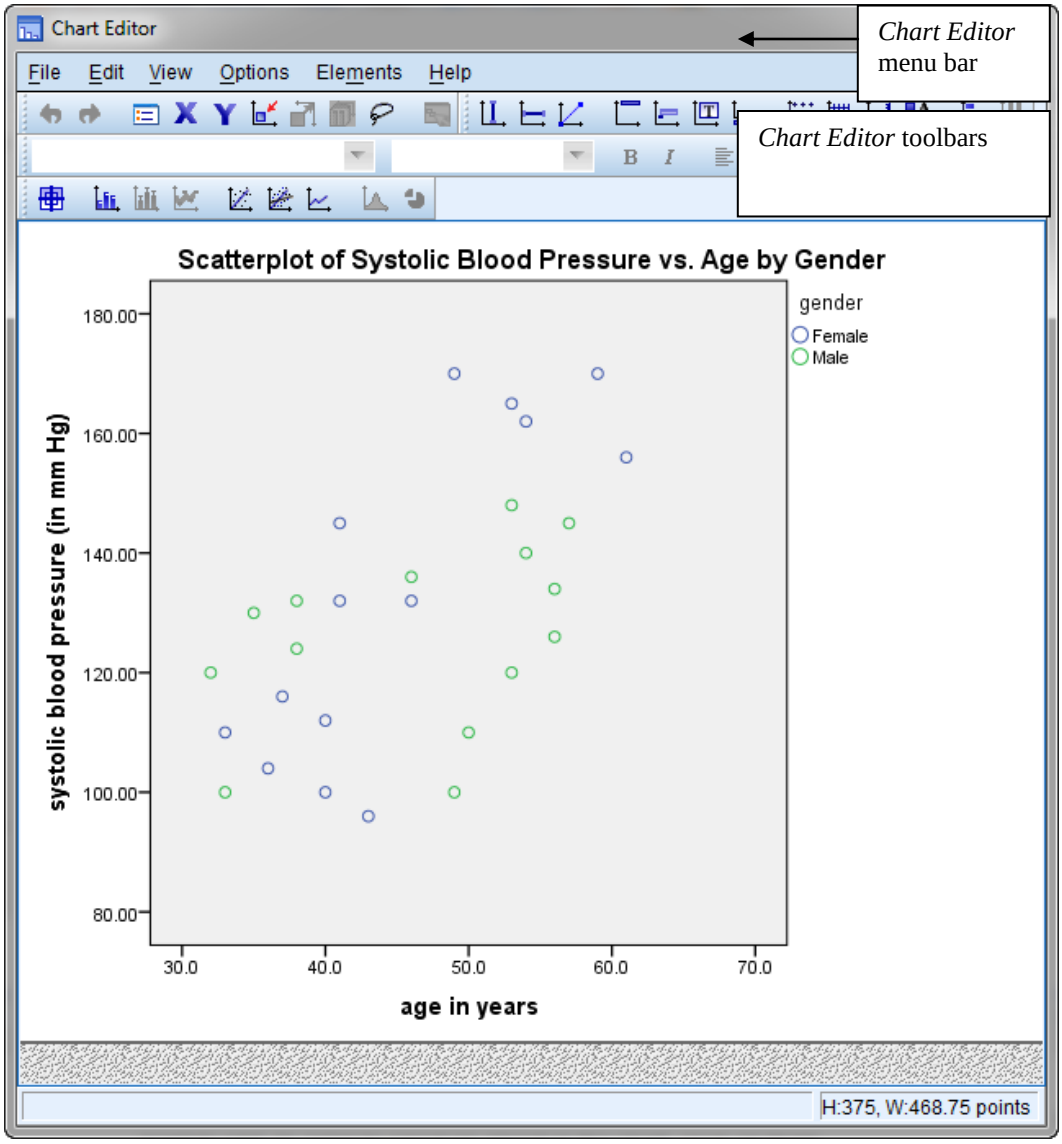


Click the *Titles* tab and enter the title of the plot. Click OK. The following chart will be displayed in the contents pane of the *Viewer* window.



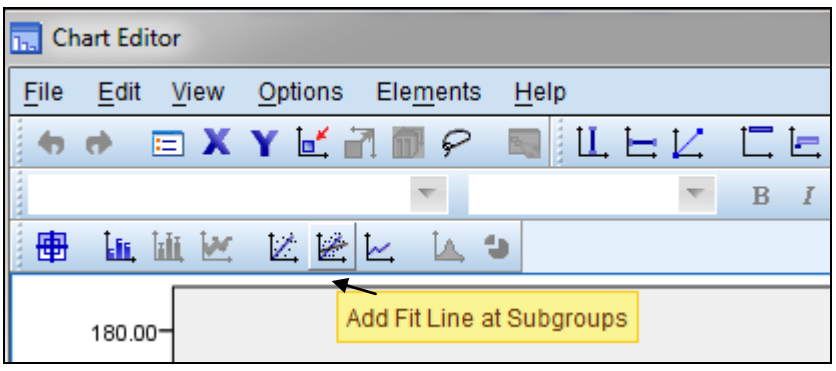
Once you have a chart, there are many attributes you can edit to change its appearance. You can change the title, labelling, fonts, colours, marking symbols, or the scale axis range.

To edit a chart, double-click the chart in the *Viewer*. This displays the chart in a chart window. The chart window for the scatterplot of systolic blood pressure vs. age by sex is shown below:



You can edit the chart from the menus, from the toolbar, or by double-clicking on the object you want to edit. In order to select an object in the chart, click on them.

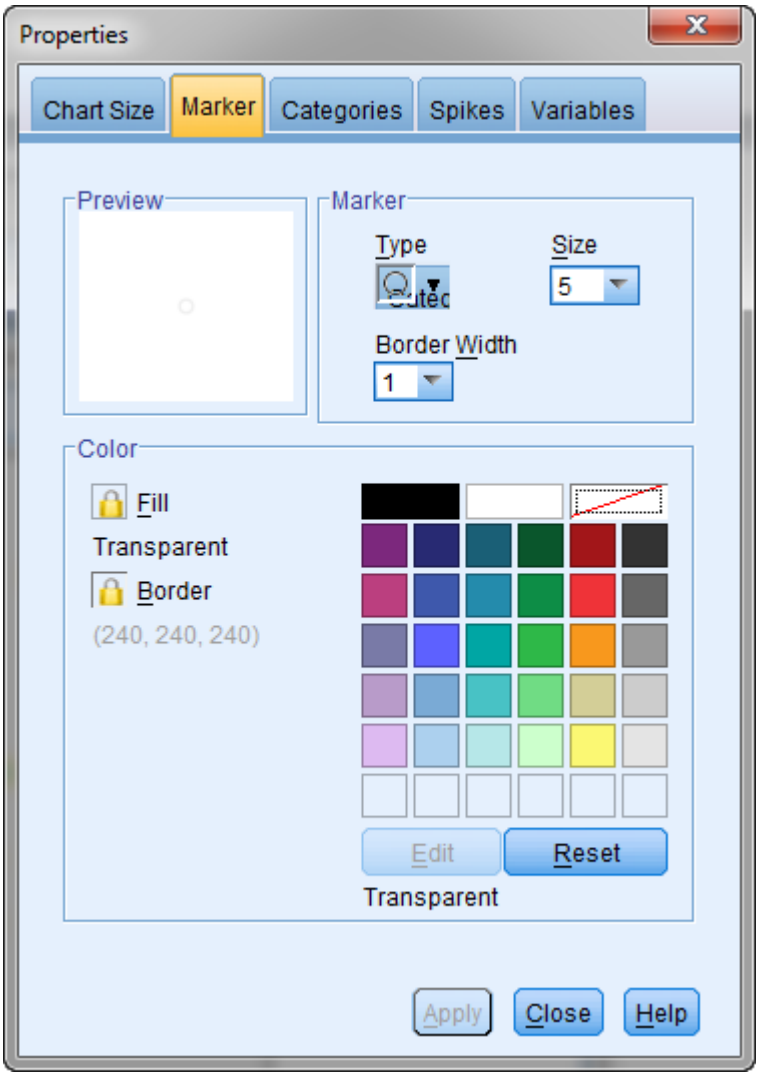
The *Chart Editor* window is displayed below. In order to determine the function of a tool, place the mouse pointer over the corresponding button, but don't click the mouse button. SPSS displays a brief description or the name of the tool.



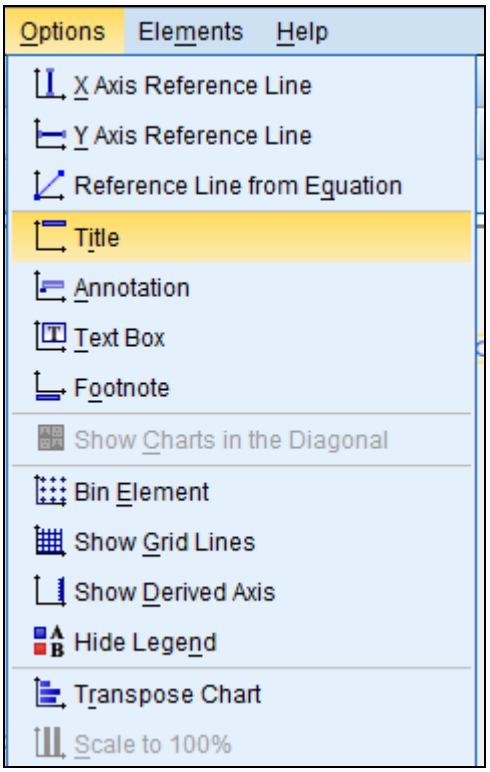
You must select an object in the chart to change its attributes with a tool in the toolbar. To select an object on a chart, click on it. For example, clicking on the legend title puts a selection box around it. To change the attributes of the selected object, double-click on it. This opens the *Properties* dialog box. Select the font and size you want, and click *Apply* to display the edited chart.

To practice, we will change the marker style for the males in our scatterplot using the appropriate tool in the toolbar. Indeed, the male observations are marked with green small circles and female observations are marked with blue small circles. We will replace the small circles for males with red small squares to better distinguish between male and female observations.

Click on one of the small circles corresponding to the male observations in the chart window or equivalently click on the small circle for the male observations in the legend on the right. and then select the *marker style* tool. In the resulting dialog box choose a tiny triangle to mark the male observations, and click *Apply*. The following *Properties Window* will be shown:



Now you can choose the colour, shape, size and border for the markers corresponding to the two gender groups. The *Properties Window* can also be accessed from the *Edit* menu. In a similar way you may edit the labels at the horizontal and vertical axes or the title of the plot. The title can also be changed from the *Options* menu.



12. Saving Results. Using Results in Other Applications

In this section you will learn how statistical results and charts can be saved, and copied and pasted into another Windows application, such as a word processing program or a spreadsheet.

12.1 Saving Results

The results from running a statistical procedure are displayed in the Viewer. The output produced can be statistical tables, charts, graphs, or text. The contents of the *Viewer* can be saved to a Viewer document. The saved document includes both panes of the Viewer window (the outline and the contents).

To save a Viewer document, from the Viewer menus choose *File*, and the *Save* from the pop-up menu. Enter the name of the document and click *Save*. All Viewer files have the extension ".spv". You can open the Viewer file either from *Data Editor* window or the *Viewer* window by clicking on *File*, and then *Open*. To save results in external formats (for example, HTML, text), use *Export* on the *File* menu.

12.2 Pasting Results as Metafiles

To copy output into another application as a metafile (picture) or unformatted text, click the item in the outline or contents pane of the *Viewer* to select it. Select *Copy* from the *Edit* menu. In the other application, select *Paste Special* from the *Edit* menu. In the *Paste Special* dialog box, select *Picture*. The metafile retains all the font characteristics and border styles of the items at the time you copied them.

To practice, paste the *scatterplot of systolic blood pressure vs. age by gender* in the *Viewer* window into a Word document.

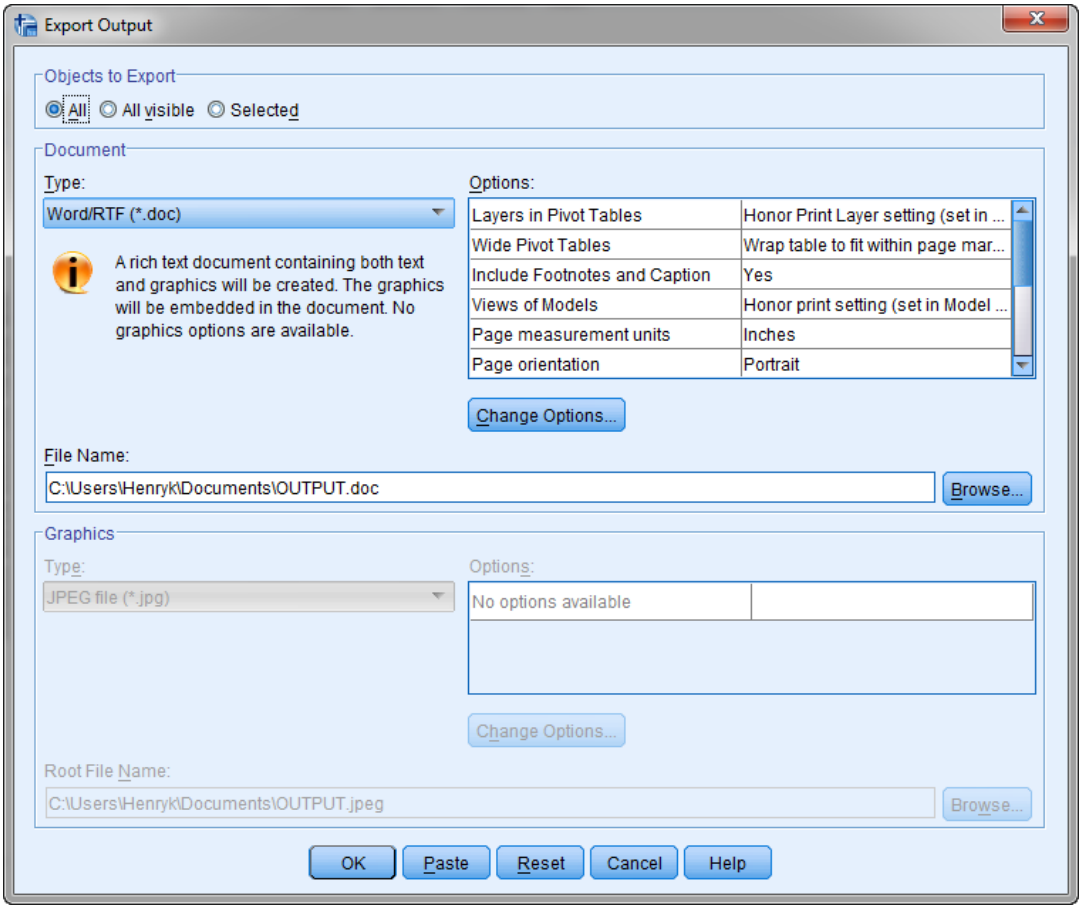
12.3 Pasting Results as Unformatted Text

To copy output into another application as unformatted text, click the item in the outline or contents pane of the *Viewer* to select it. Select *Copy* from the *Edit* menu. In the other application, select *Paste Special* from the *Edit* menu. In the *Paste Special* dialog box, select *Unformatted Text*. Unformatted text contains tabs between columns.

12.4 Exporting Output

Export option in the File menu (Viewer) saves the Viewer output in HTML, text, Word/RTF, Excel, PowerPoint, and PDF formats. Charts can also be exported in a a variety of common formats used by other applications.

To export output, make the *Viewer* the active window (click anywhere in the window), anf from the menus choose File, and then Export. The following dialog box is displayed:



13. Getting Help

SPSS has an extensive online help system that is available at any time. For general help and information about SPSS select *Help* from the SPSS menu bar. Moreover, *Help* is available from any SPSS window with the Help menu. In particular, if you need help on how to use a dialog box to run a procedure, click on the *Help* button in the dialog box. This takes you directly to the Help topic for that dialog box.