

## Introduction to SPSS

We begin by first starting SPSS [Statistical Package for the Social Sciences]. You may choose to either

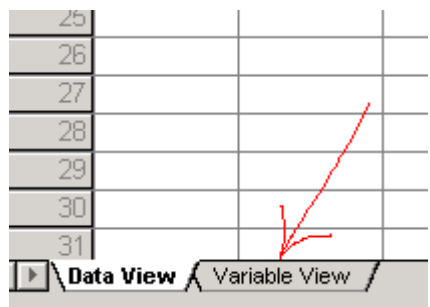
1. click on the SPSS icon in the corner of your screen
2. go to the lower left and hit **Start**, then from **Programs**, then **SPSS for windows**, then **SPSS 17.0**
3. A window will open asking you what you want to do. Either cancel this window or click the “type in data” button.
4. A new spreadsheet should appear on the screen.

Below is a sample data set that we will be using for today’s exercise. It lists the heights & weights for 10 men and 12 women.

Male		Female	
Height	Weight	Height	Weight
69	192	65	110
70	148	61	105
65	140	67	136
72	190	65	135
76	248	70	187
70	197	62	125
70	170	63	147
66	137	60	118
68	160	66	128
73	185	66	175
		65	147
		64	120

### Entering the data

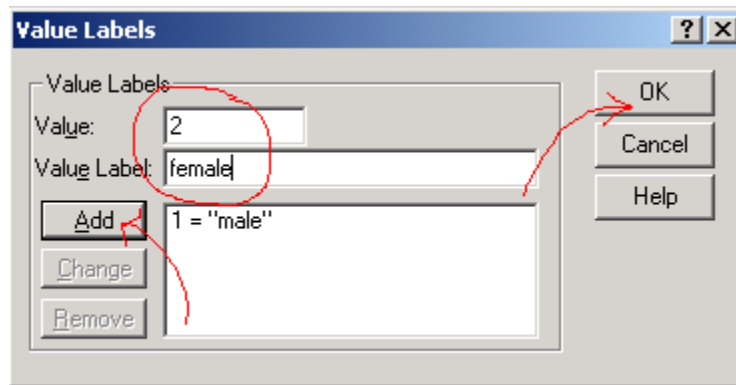
We begin by entering the data. The first task, after opening up SPSS, is to name the columns that we will be using. We do this by clicking on the variable view tab in the lower left hand corner of the spreadsheet.



We can now enter the column names “height”, “weight” & “gender” into rows 1, 2 & 3 in the upper left hand area of the variable spreadsheet. Note that SPSS will fill out the rest of the table assuming numerical values.

	Name	Type	Width	Decimals	Label	Values	Mis
1	height	Numeric	8	0		None	None
2	weight	Numeric	8	0		None	None
3	gender	Numeric	8	0		None	None
4							

After entering the column names, I would suggest making a few helpful changes to the automatic settings. If you click on the cells in the decimal column, you can change them to 0 decimals which will result in integers in the data spreadsheet. Also I strongly suggest entering in codes for the values of the factors so that you will have meaningful levels for your factors. You do this by clicking on the cell in the “values” column and then clicking on the small gray box that appears.



This will open the **Value Labels** window. For each of the levels of the factor you enter the value, the value label, and then hit **Add**. I suggest that you use “1” for males and “2” for females. When you have defined all of the levels hit the **OK** button.

After coding the values, we return to the data view, by clicking on the data tab in the lower right hand corner of the spreadsheet. We are now ready to enter our data. For each of the 22 data points we have a height, weight and a gender. These data will be entered into the spreadsheet so that we will have 3 columns of data, each with 22 rows. A row will correspond to one individual (or experimental unit). The height column will consist of the raw heights for the subjects, the weight column will consist of the raw weights and the gender will consist of a 1 or a 2, depending on the gender of the subject. [a section of the table is shown below]

	height	weight	gender
1	69	192	1
2	70	148	1
3	65	140	1
4	72	190	1
5	78	248	1
6	70	197	1
7	70	170	1
8	66	137	1
9	68	160	1
10	75	185	1
11	66	110	2
12	61	105	2

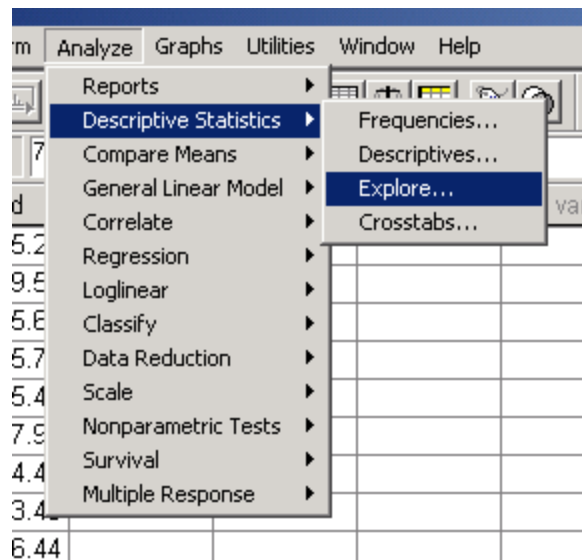
### Saving the data

To Save your data: (with the spreadsheet showing on the screen)

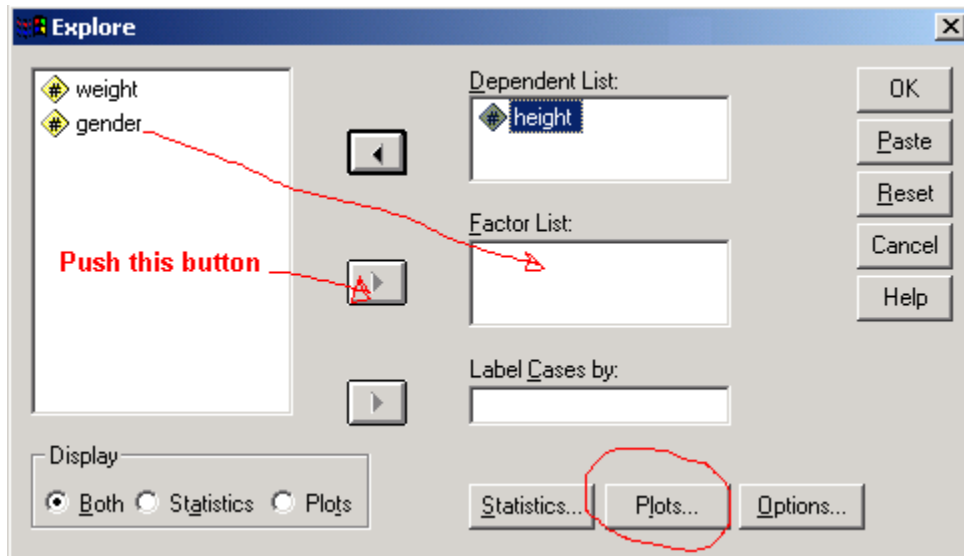
Click on **File** and **Save as** and locate your home directory. Type in the “filename” (pick a name that means something to you about the data). Click on **OK** ( this will save the data with a .sav extension)

### Descriptive Statistics and Graphical Analysis

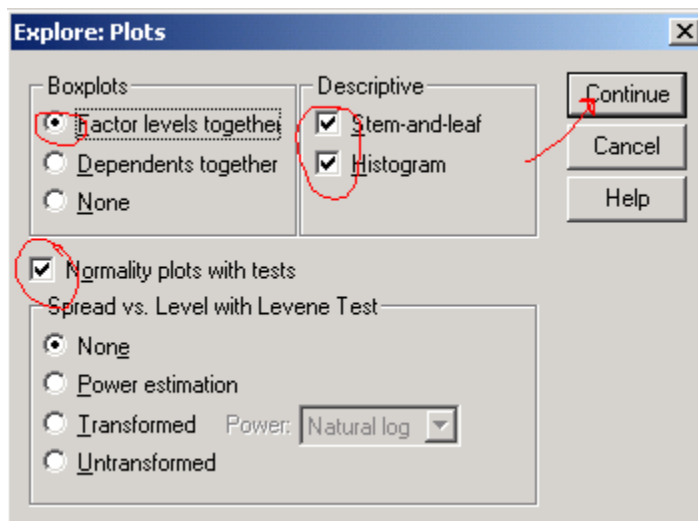
We begin our analysis of the data by calculating summary statistics, such as means and variances, and by creating several basic plots. These options can be found by going to the **Analyze-Descriptive Statistics-Explore** pull down menus.



This will take you to the **Explore** window. We need to place “height” in the dependent field as it is our response. “Gender” should be placed in the factor field as it is our independent variable/factor in this data. They can be put into the fields by first clicking on them and then clicking on the button for the field. Below we are indicating how to move gender into the factor list.



Having defined our dependent/response variable and our independent/factor variable we can now set up our plots. By pushing on the **Plots** button, found on the image above, we move into the **Explore Plots** window.

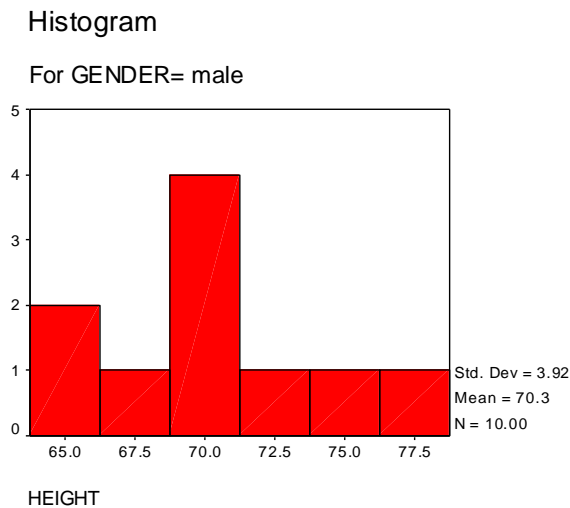


We need only to make sure that the four boxes are checked and then we can hit the **continue** button. This returns us to the Explore window where we click on the **OK** button and review the output from the analysis.

## Output

The output screen will be divided into two parts (left and right). On the right hand side is the output window, which contains all of the plots requested in the analysis. On the left will be a listing of what output is produced, by name, with an icon beside it that if you look really close is an open book. If you double click on any output in the left screen the book will close and the output will disappear from the right output window. (It is not gone, only hidden). Double click on the same icon and the output reappears.

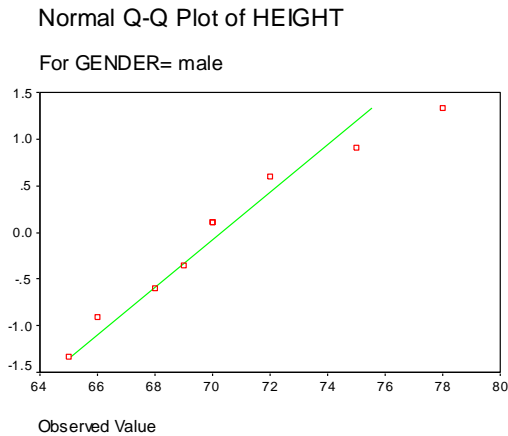
You should find a lot of output in the SPSS output window. The first part of the output will be various descriptive statistics such as averages, variances, confidence intervals, etc. Also among the information are the histograms, stem & leaf plots, normal plots and boxplots [partial examples shown below].



HEIGHT Stem-and-Leaf Plot for  
GENDER= female

Frequency	Stem & Leaf
5.00	6 . 01234
5.00	6 . 66678
2.00	7 . 04

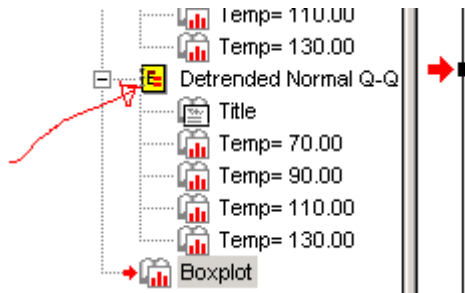
Stem width: 10  
Each leaf: 1 case(s)



All of these outputs can be edited to make them more readable. You can double click on any of the outputs to bring up an editor window that will allow you to edit the image. For most graphs you are able to change colors, symbols, line types, titles, or even rotate axes. For tables it is possible to remove unwanted lines in the output and even add additional text that wasn't originally included in the table. [note - next last section contains information on editing histograms].

### Printing

If you are planning on printing out the entire output you may wish to turn off certain parts [such as detrended normal plots] of the output screen. As was mentioned before, to do this click on the appropriate symbol in the left output screen twice and the selected output will disappear from the right output screen. If you double click on it again the output will reappear in the right window.



To print the entire output window make sure you are in the output window and then click on **File**, then **Print**. (hit **Print Preview** if you would like to see what it will look like before it prints). To print out only a certain segment of the output window click on the table or graph that you want and a red arrow will appear to the left. Now click on **File**, then **Print**.

I find it easier to cut and paste individual pieces into a word document. To do this you click on the image and use the copy command from the main tool bar. I suggest that when you copy use "copy" rather than "copy object" and use "paste special" when pasting.

### Editing a Histogram

To modify this histogram, double-click on the graph, wait for the editor window to appear, and follow the directions below.

*To change the class sizes:*

- Double-click on the numbers below the horizontal axis.
- Under Intervals, click Custom, then Define...
- Enter the minimum, maximum and interval width.
- Click Continue and OK enough times to get out of the dialog boxes.

*To change the vertical Scale:*

- Double-click the numbers to the left of the vertical axis.
- In the window that comes up, you can change the range and the number of tick marks on the vertical axis, among other things.

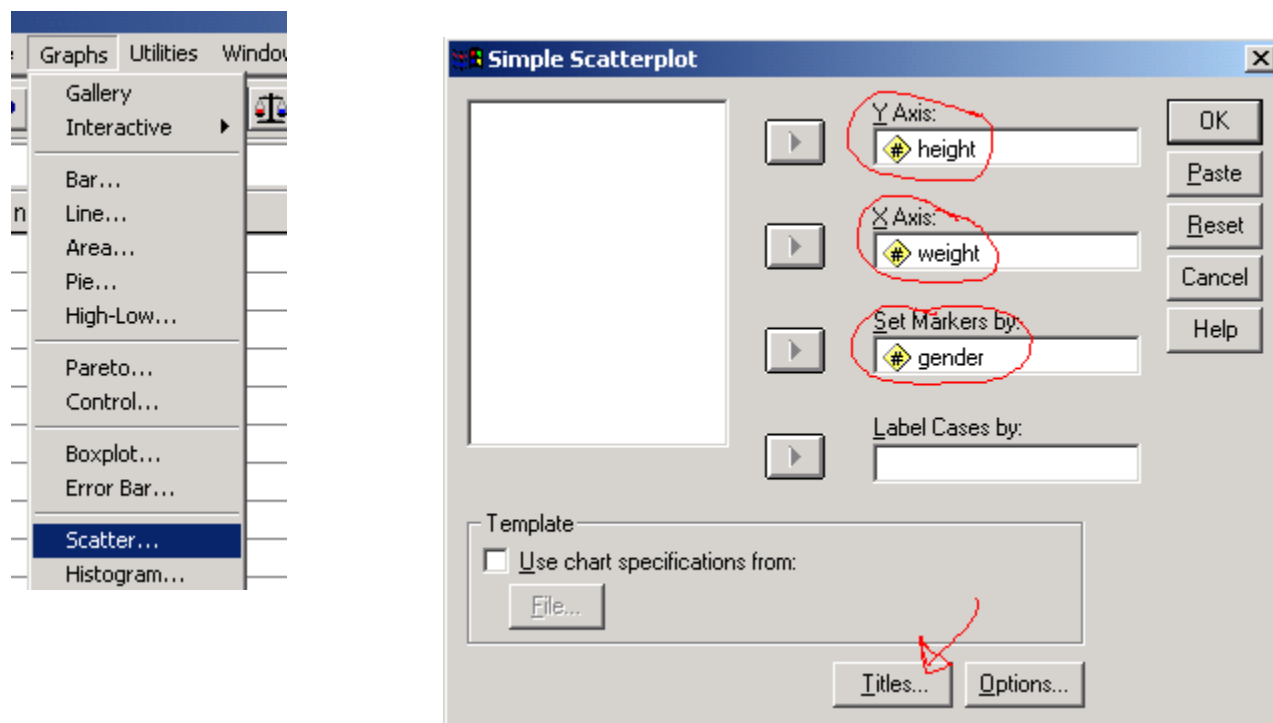
*To create a relative frequency histogram:*

- Double-click the numbers to the left of the vertical axis.
- Click Labels...
- Enter the number of data points (which is 22 in the case but varies depending on your data set) in the box marked Scaling Factor.
- Change the number of Decimal Places to at least 2.
- Click out of the dialog boxes.

As was mention earlier, there are lots of other parts of the graph you can change, like the titles, axis labels, etc. Double-clicking on each part of the graph will usually bring up the appropriate menu. When you have the histogram edited the way you like it, close the editor, and you'll be back at the graph output window.

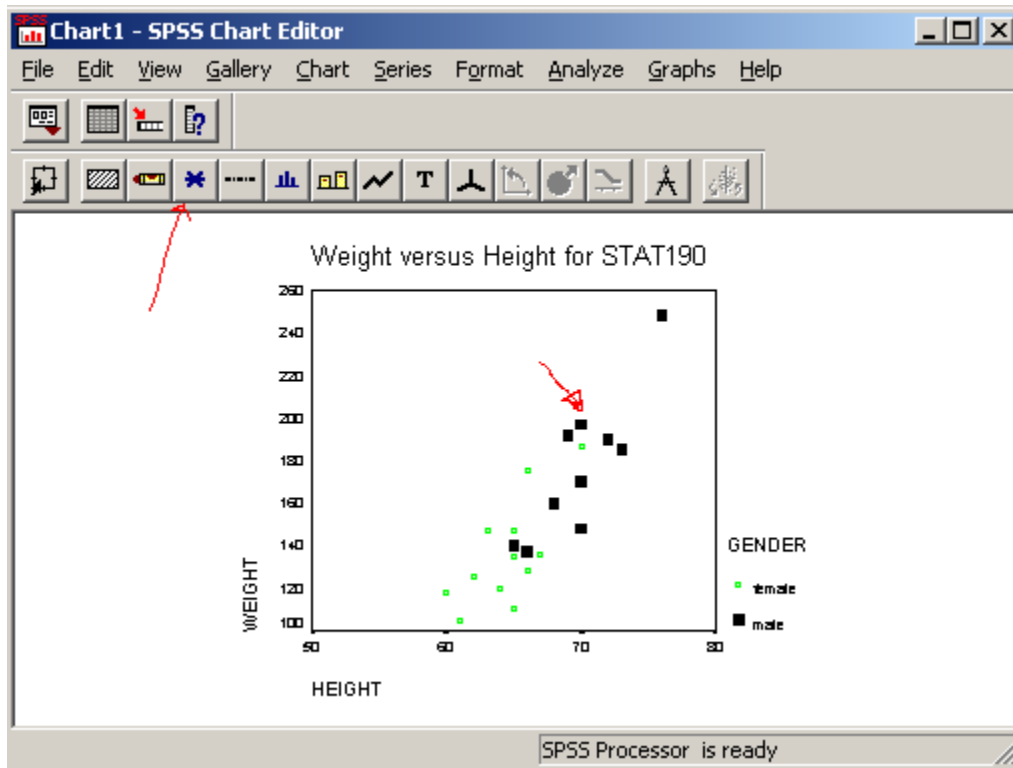
### Creating a scatterplot

If we want to compare the heights and weights for the sample in a quick and easy way to summarize the data is to present it in graphical format. By clicking on the graph and scatterplot button, the scatterplot window will be opened. Choose the simple and click on define.



We then enter the columns into the appropriate areas and the graph will appear in the output window.

As the default setting for SPSS is to use color to differentiate the groups [and most prints are in black and white] you should immediately edit the graph and change the symbols for the groups. To do this, double click on the graph to bring it into edit mode. Click on a male symbol on the graph (some of them will become highlighted) then click on the \* on the tool bar. Choose a symbol and hit apply. The male symbol should change. Now close the window and the graph is done.



You will now be able to differentiate the groups if you do a plot in black and white.

