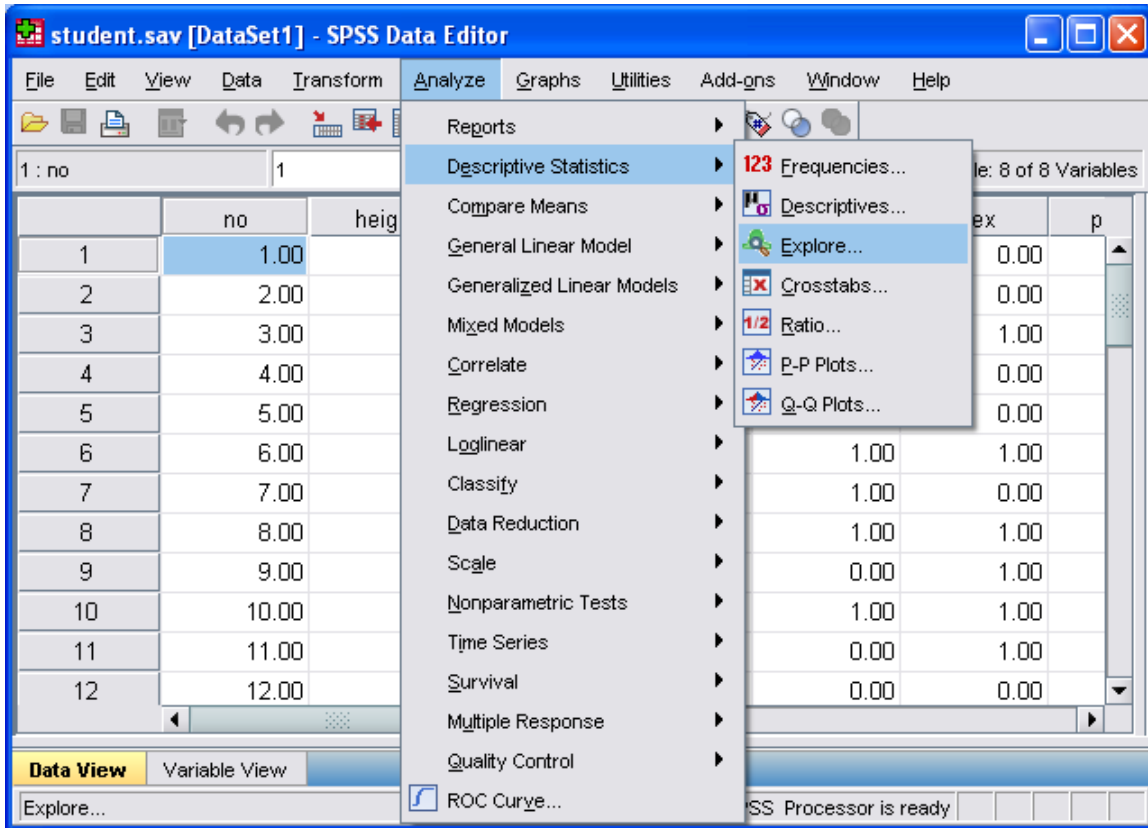


Data used in this guide: [studentp.sav](http://www.cc.ysu.edu/~ghchang/stat/studentp.sav) (<http://www.cc.ysu.edu/~ghchang/stat/studentp.sav>)

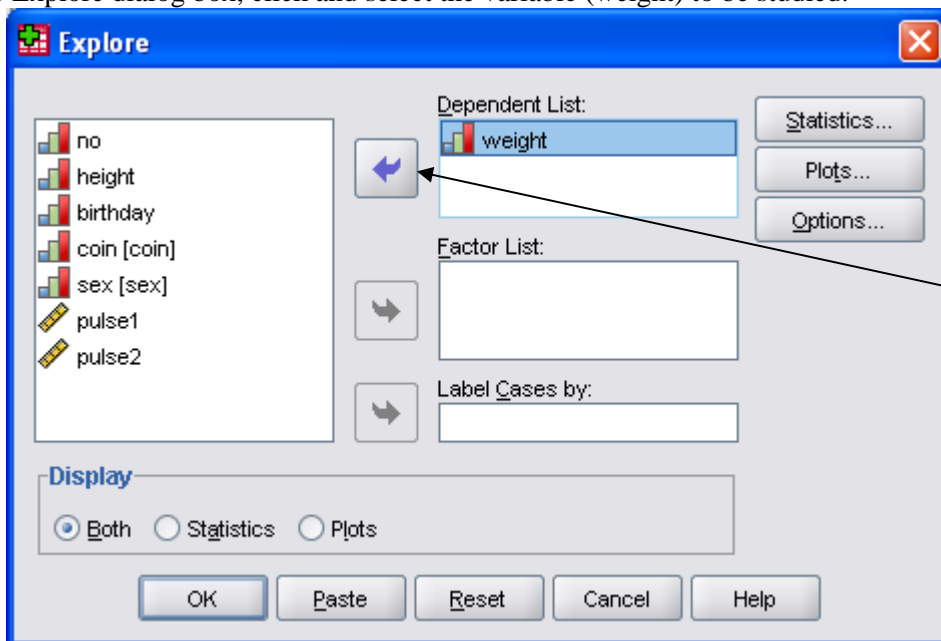
### Organize and Display One Quantitative Variable (Descriptive Statistics, Boxplot & Histogram)

1. Move the mouse pointer to **Analyze**, click the left button of the mouse and move through the following menu selections:

**Analyze** ⇒ **Descriptive Statistics** ⇒ **Explore ...** (To perform Exploratory Analysis)

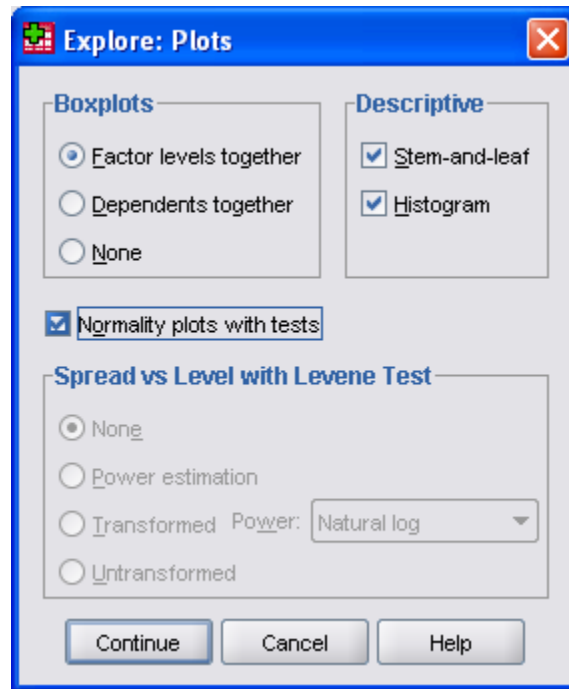


2. In the Explore dialog box, click and select the variable (weight) to be studied.



Click the variable to be selected from the list of variables on the left for analysis and click the select button, the button with an arrow shape in it, to select the variable into Dependent List box.

- In the Explore dialog box, click **Plots...** button. In the Explore: Plots dialog box, check the **Histogram** and **Normality plots with tests**, if they are needed, and click on **Continue** button. If Sig. value (or  $p$ -value) in the normality test table is less than .05, it implies that data may **not** be from normally distributed population. The values .200 and .236 are  $p$ -values calculated based on two different tests.



**Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
WEIGHT	.127	22	.200*	.938	22	.236

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

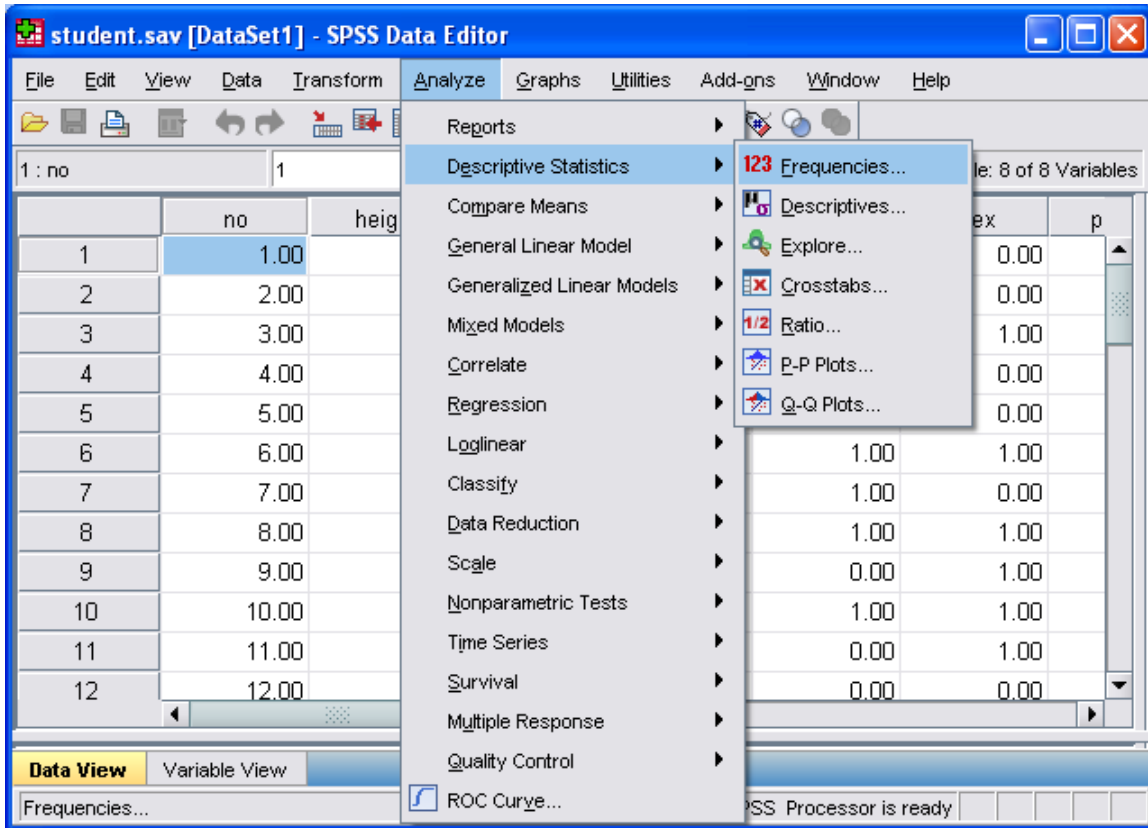
The  $p$ -values are both greater than 0.05. The distribution which the data was sampled from is not significantly different from normal, at 5% level of significance.

- In the Explore dialog box, click on **OK** button. The SPSS will put the results, histogram, stemplot and descriptive statistics such as mean, standard deviation, confidence interval for mean, in the OUTPUT window.
- If, in the Explore dialog box, one click on **Statistics** button and check on **Percentiles** box and click on **Continue**, SPSS will produce quartiles (25<sup>th</sup> and 75<sup>th</sup> percentiles are good for getting 5-number summary for the data), and some special percentiles for the active data set.
- If one wishes to explore the quantitative variable for separate categories of a qualitative variable, select that qualitative variable and put it in the **Factor List** and click **OK**. (See the example in last three pages of this document.)

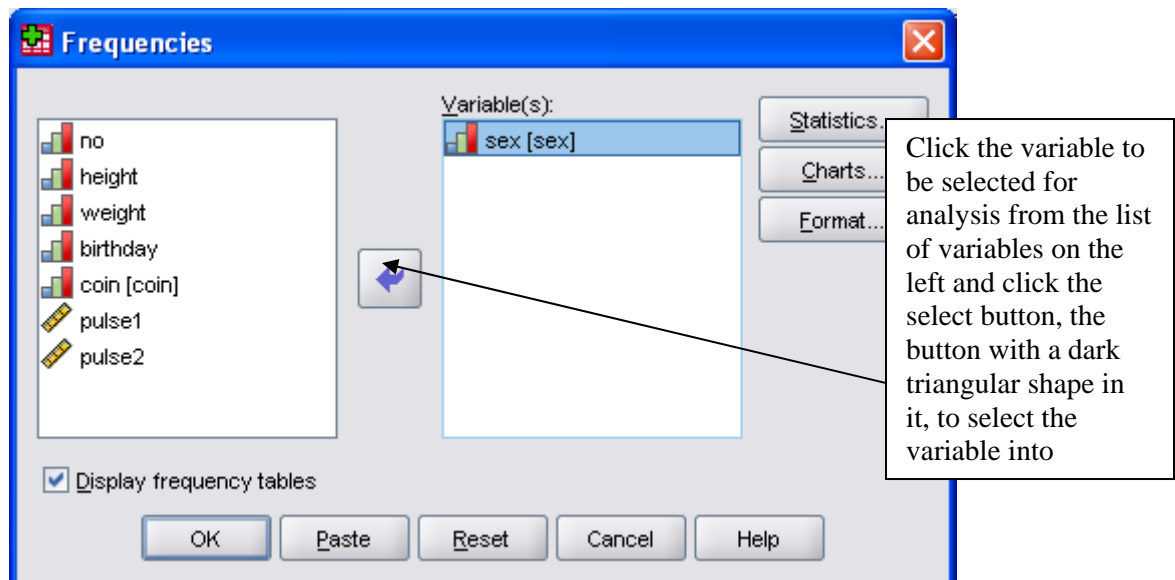
### Organize and Display One Qualitative(Categorical) Variable (Pie or bar charts)

1. Move the mouse pointer on **Analyze**, click the left button of the mouse and move through the following menu selections:

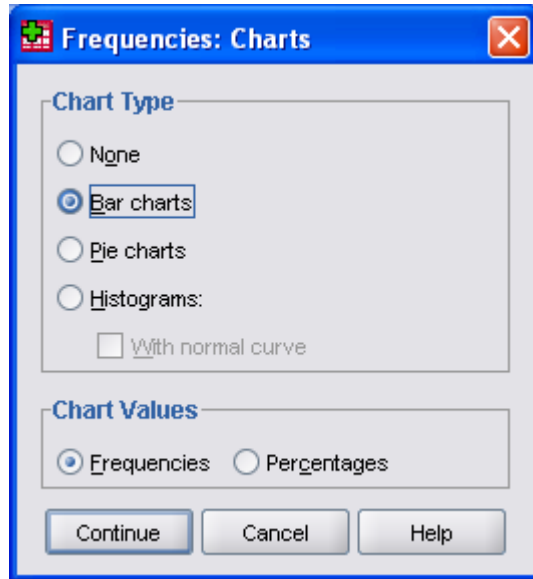
**Analyze ⇒ Descriptive Statistics ⇒ Frequencies ...**



2. In the Frequencies dialog box, click and select the variable (sex) to be studied.



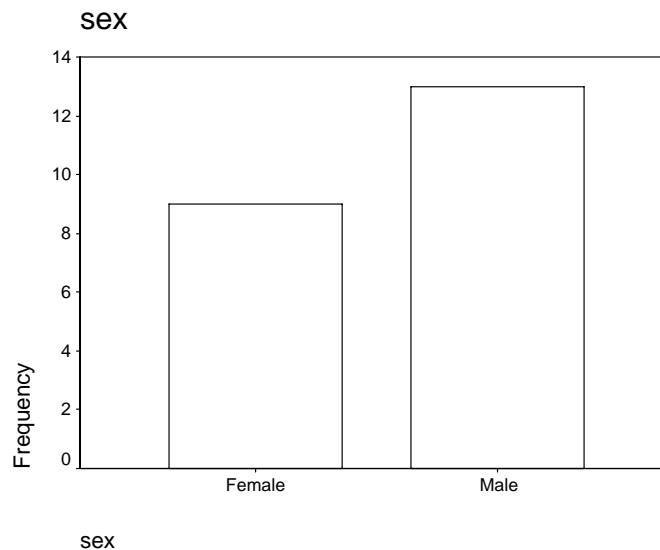
- In the Frequencies dialog box, click **Charts...** button, if one wishes to display chart. In the Frequencies: Charts dialog box check on the desired chart and select either Frequencies or Percentages to be displayed and click **Continue** button.



- In the Frequencies dialog box, click on OK button. The SPSS will put the results, frequency distribution table and bar chart (if checked), in the OUTPUT window.

**sex**

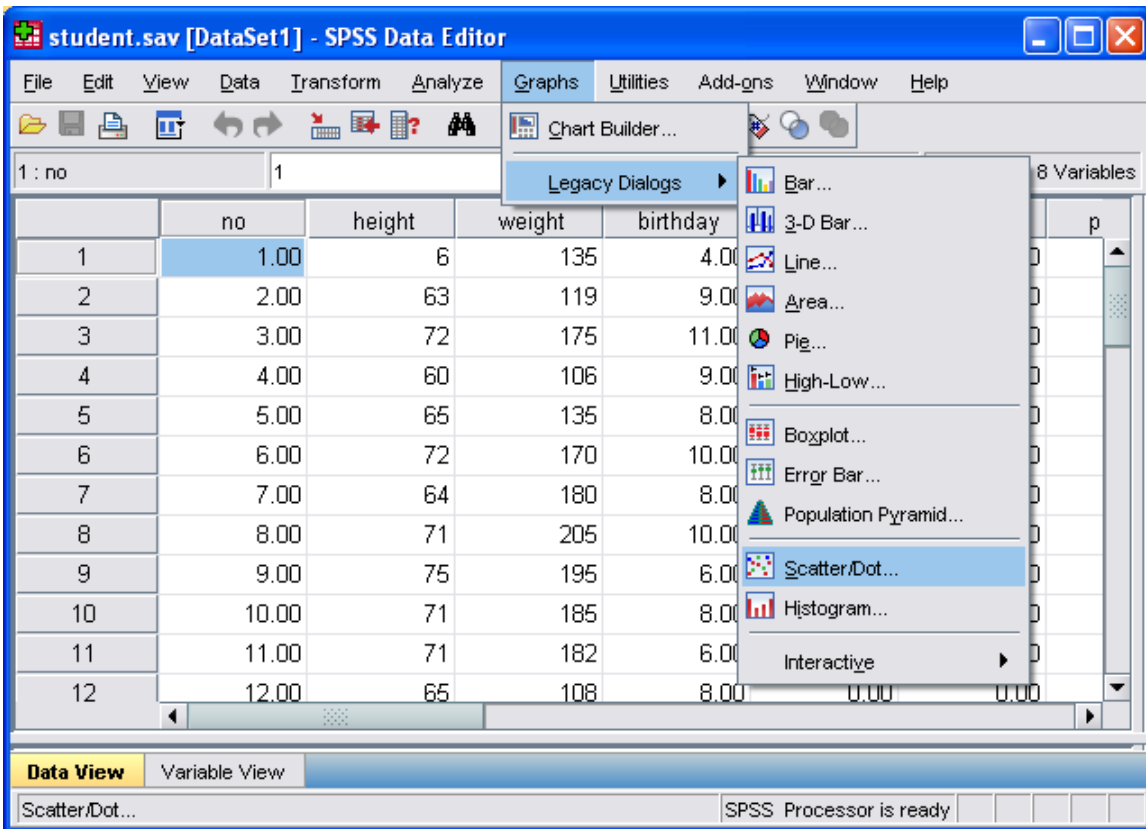
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	9	40.9	40.9	40.9
	Male	13	59.1	59.1	100.0
	Total	22	100.0	100.0	



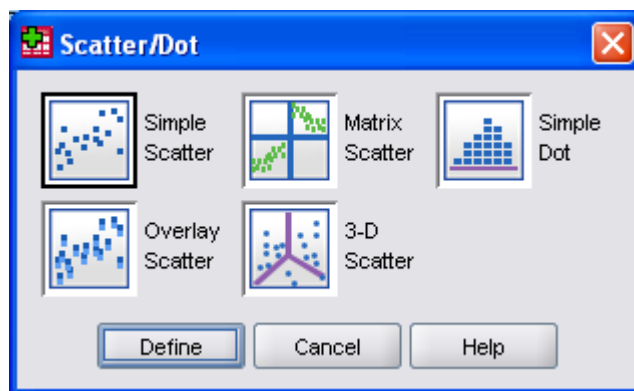
### Examine Relation Between Two Quantitative Variables by Chart (Scatter Plot)

1. Click and move through the following menu selections:

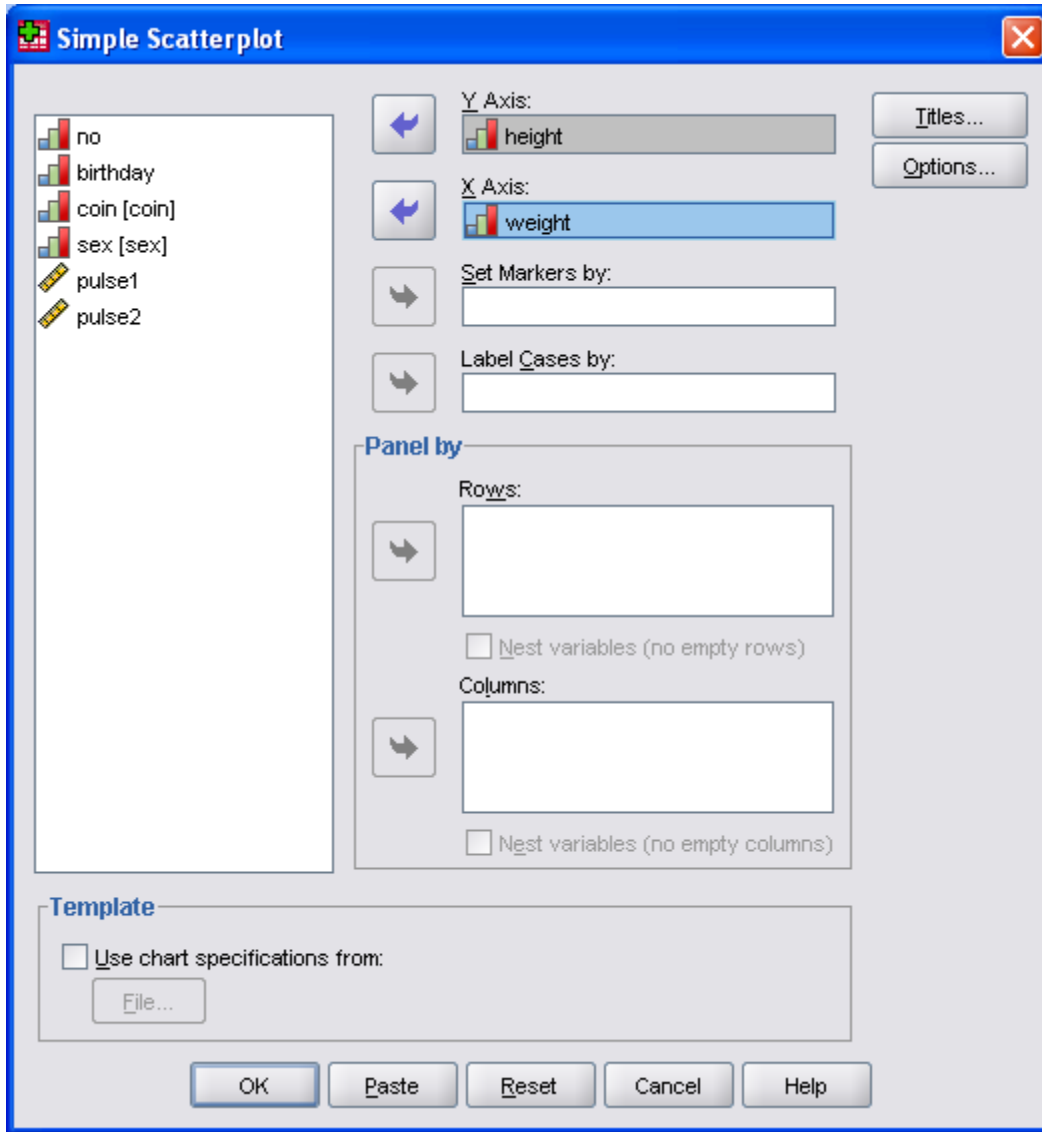
**Graphs ⇒ Legacy Dialogs ⇒ Scatter ...**



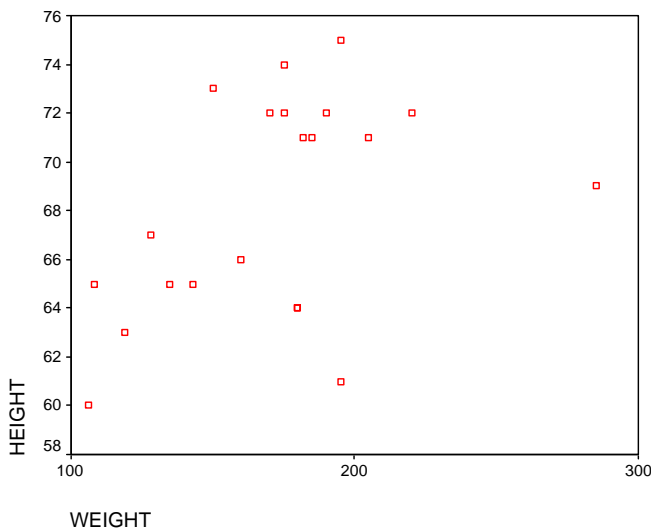
2. In Scatterplot dialog box, click the **Simple** option and click **Define** button.



3. In **Simple Scatterplot** dialog box, select the two variables (height and weight) to be studied. If one wishes to build a regression model for predicting height using weight variable, usually choose height variable for Y Axis (as response variable) and choose weight variable for X Axis (as explanatory variable). One can select sex variable for the **Set Markers by:** field to make scatter plot display scatter dots with different color for different sex.



- In the Simple Scatterplot dialog box, click on **OK** button. The SPSS will put the scatter plot in the OUTPUT window. The following scatter plot is based on the data in **studentp.sav** file with the first case dropped, since the height information for the first case is incorrect.

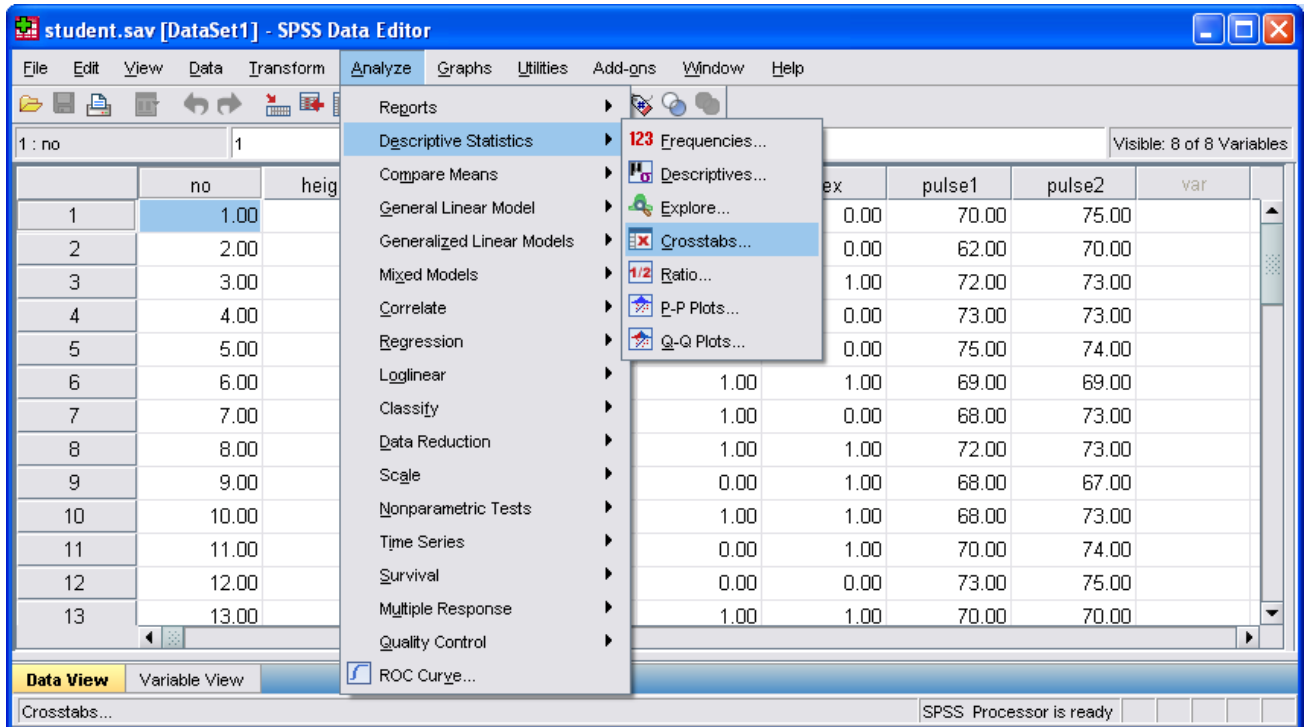


- One can double click on the any part of the chart in the SPSS output window to bring up the chart editor for editing the scatter plot. A fitted line can be added to the chart using the **Chart** option in the chart editor menu bar.

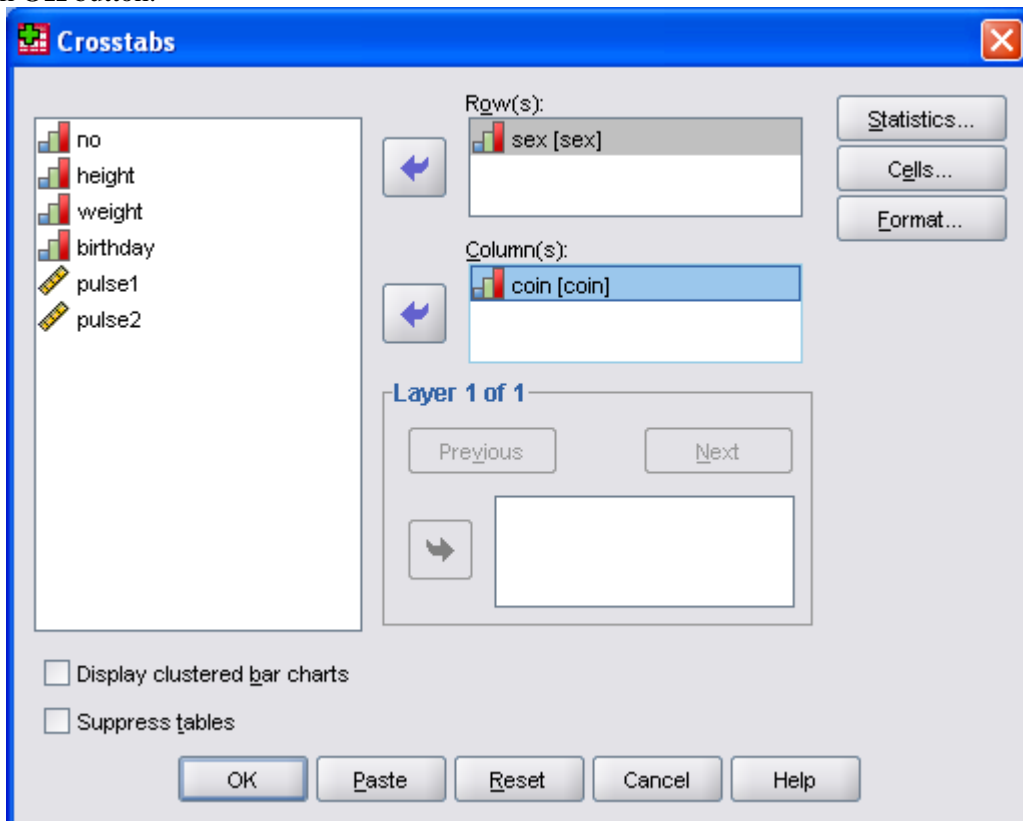
## Examine Relation Between Two Qualitative Variables (Contingency Table & Cluster Bar Chart)

1. Click and move through the following menu selections:

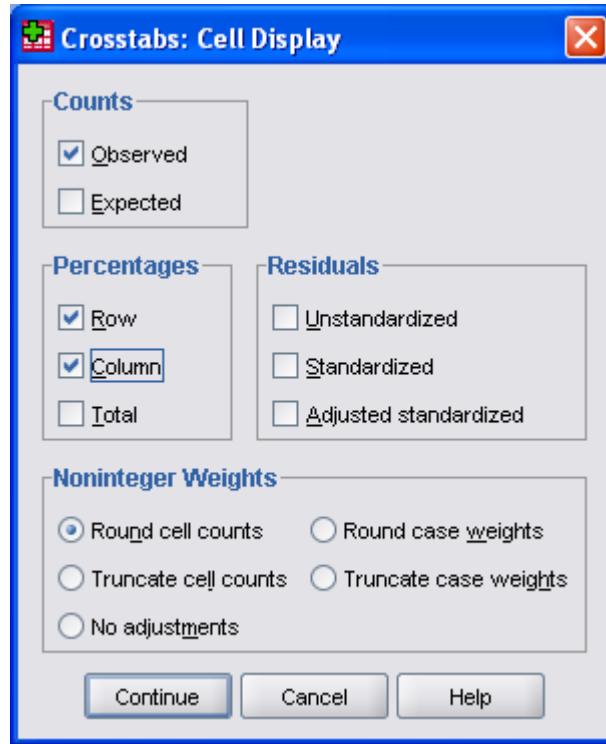
**Analyze ⇒ Descriptive Statistics ⇒ Crosstabs ...**



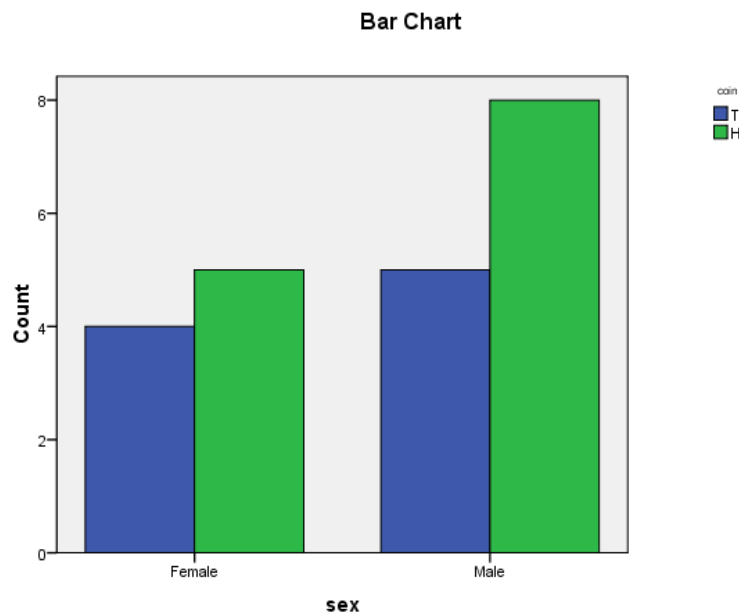
2. In Crosstabs dialog box, select the categorical variables for Row variable and Column variable, and click **OK** button.



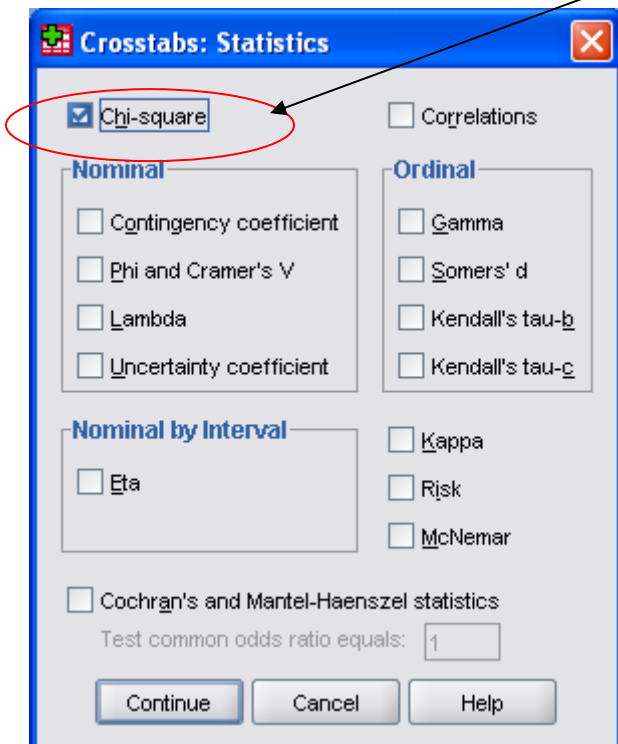
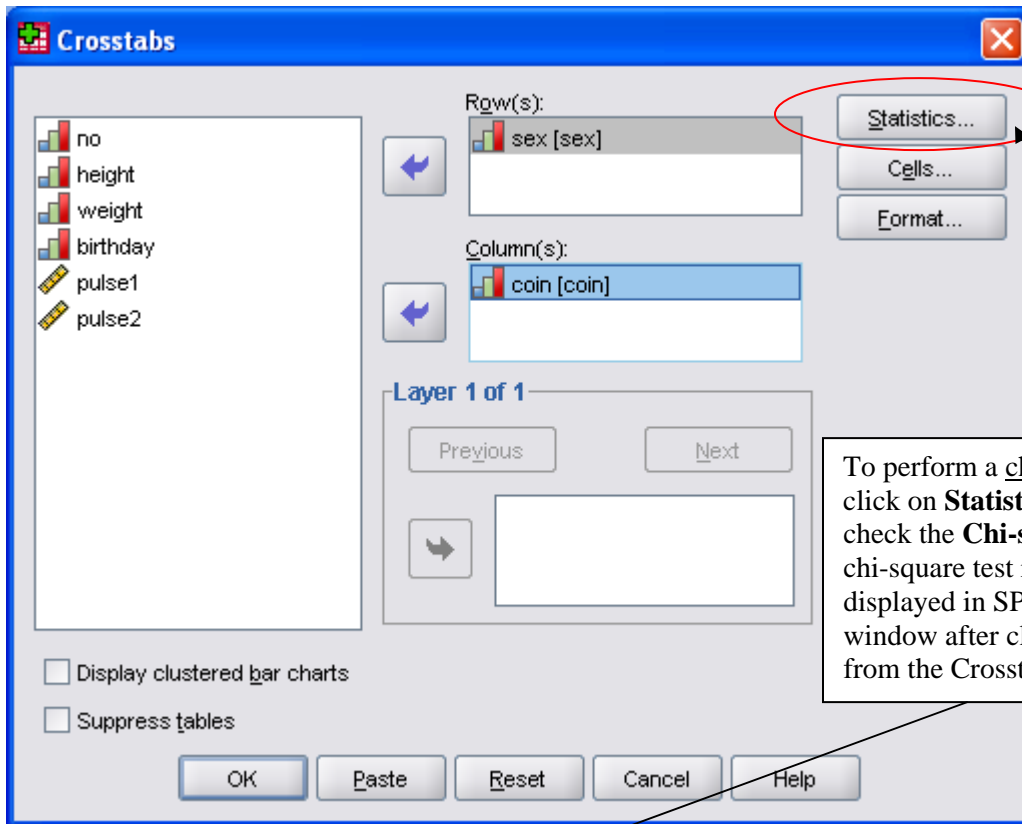
- In Crosstabs dialog box, click on **Cells ...** button to specify whether or not to display the percentage information and then click **Continue** button to go back to Crosstabs dialog box.



- One can also check on **Display clustered bar charts** option in the Crosstabs dialog box to display clustered bar chart with only frequency (count) information. If percentages are needed be displayed, go through Graphs options.
- In the Crosstabs dialog box, click on **OK** button. The SPSS will put a contingency table and also a clustered bar chart in the OUTPUT window if the clustered bar chart box is checked.



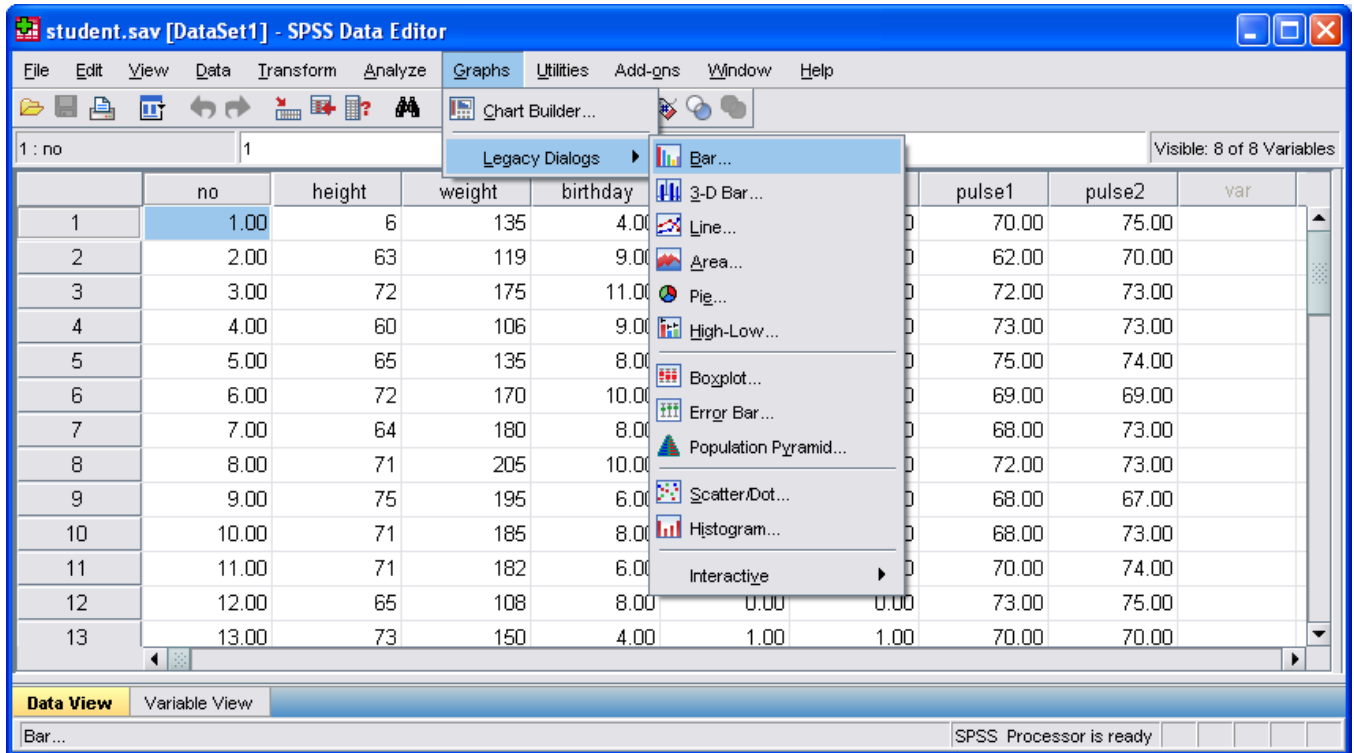
- The Chi-square test and other statistics can be done by clicking on Statistics button and choose the desired option.



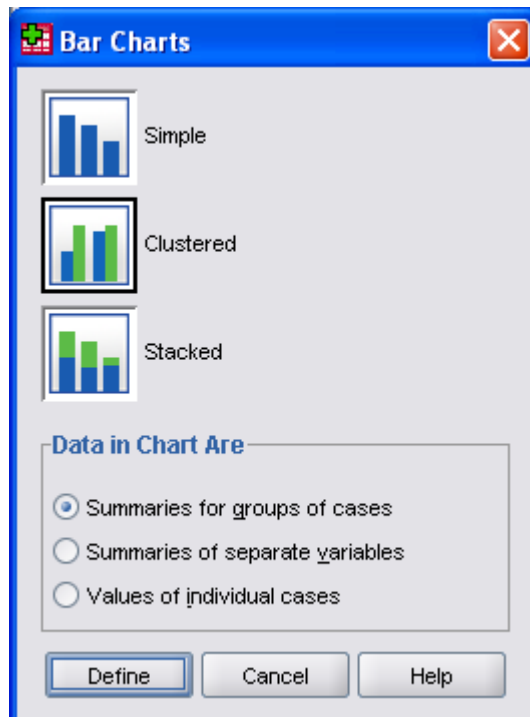
### Clustered Bar Chart (Make a separate cluster bar chart)

1. Move the mouse pointer on **Graphs**, click the left button of the mouse and move through the following menu selections:

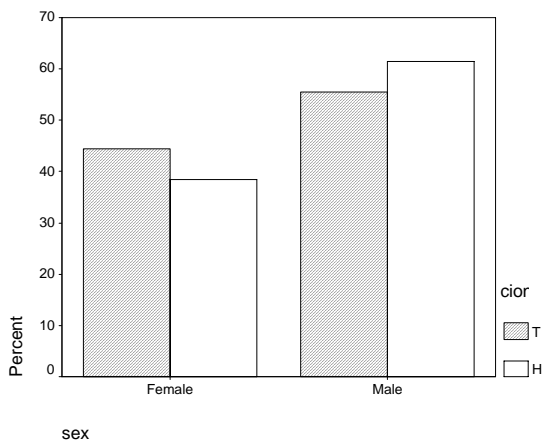
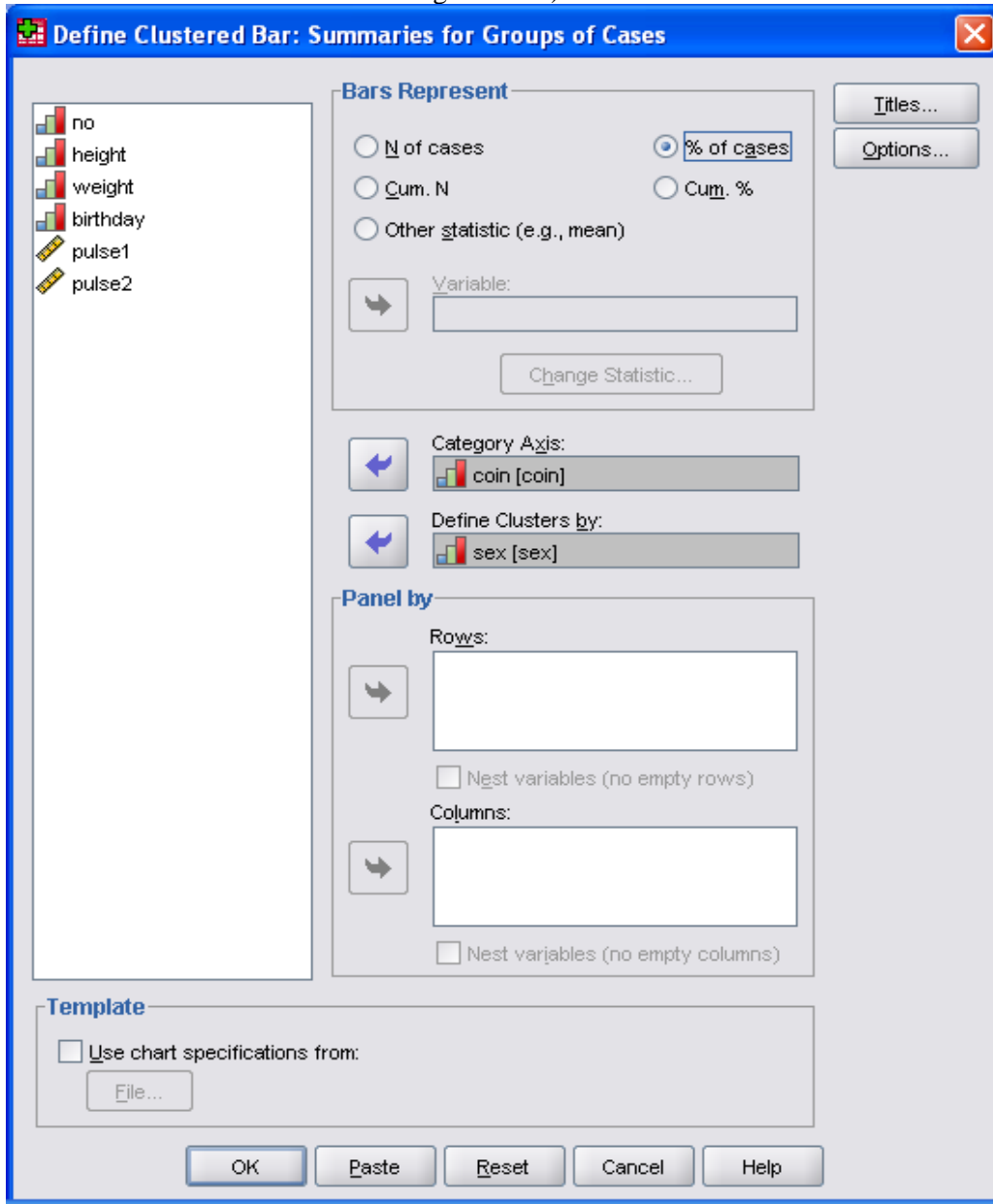
**Graphs** ⇨ **Legacy Dialogs** ⇨ **Bar ...**



2. In Bar Charts dialog box, click the **Clustered** option. Check the **Data in Chart Are** option in Summaries for groups of cases and click **Define** button.



Select the two variables (sex and coin) to be studied. One can select sex variable for the Category Axis and coin variable for Clusters and check of **% of cases** and click **OK**. (Sometimes, percentage information is better for understanding the data.)

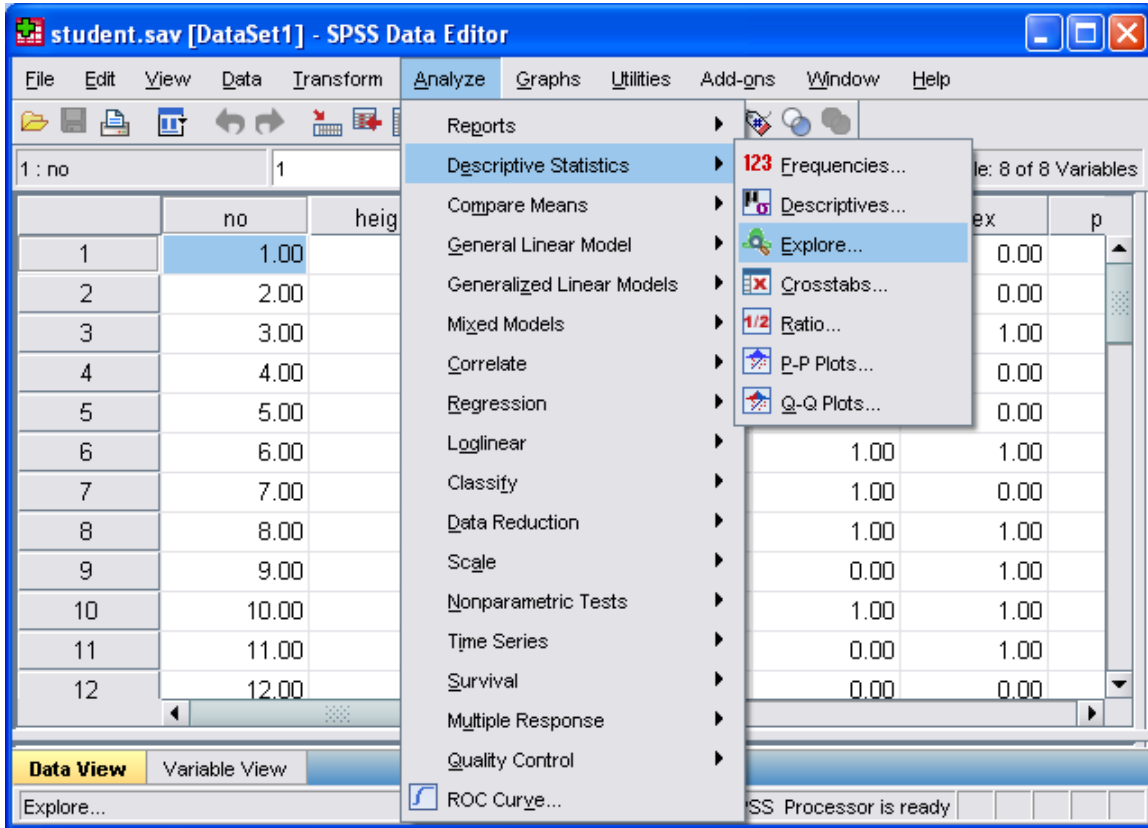


\* Chart editor can be used to modify the chart and change the color or pattern in the chart. To activate the chart editor, one can simply double click any part of the chart in the SPSS Output window.

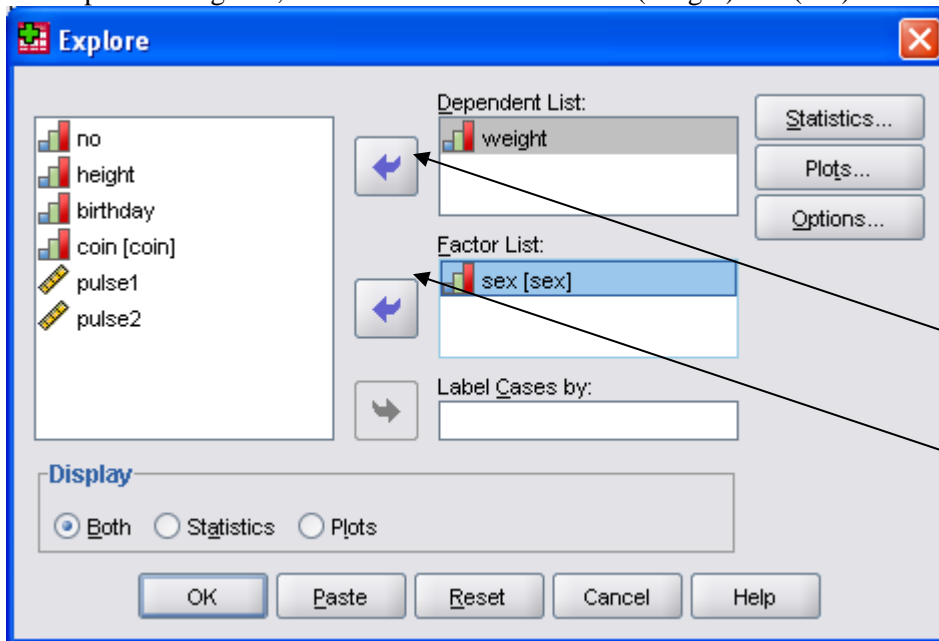
**Examine Relation Between One *Quantitative* Variable with One *Qualitative* Factor Variable (Side-by-side boxplot, descriptive measures for sub-categories.)**

1. Move the mouse pointer on **Analyze**, click the left button of the mouse and move through the following menu selections:

**Analyze ⇒ Descriptive Statistics ⇒ Explore ...** (To perform Exploratory Analysis)

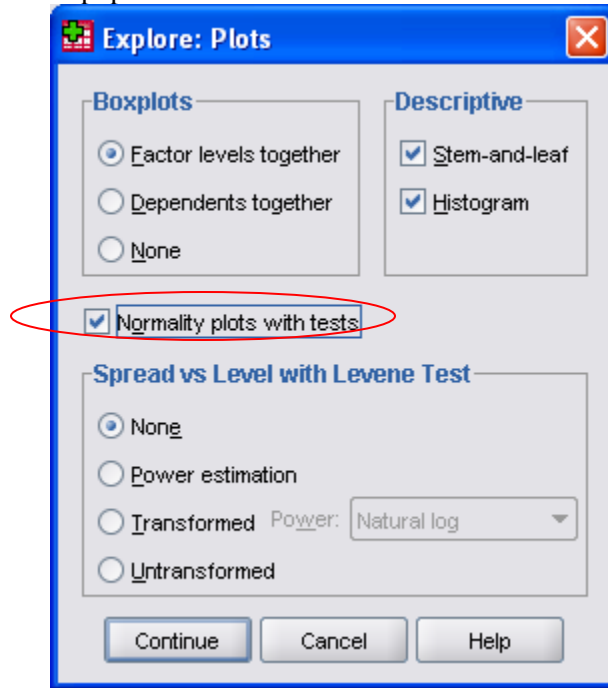


2. In the Explore dialog box, click and select the variables (weight) and (sex) to be studied.



Click the variable to be selected (weight) from the list of variables on the left for analysis and click the select button, the button with a dark triangular shape in it, to select the variable into Dependent List box. Select the sex variable in Factor List box to observe the difference between the weights from both genders.

- In the Explore dialog box, click **Plots...** button. In the **Explore: Plots** dialog box, check the **Histogram** and **Normality plots with tests**, if they are needed, and click on **Continue** button. If Sig. value (or *p*-value) in the normality test table is less than .05, it implies that data may **not** be from normally distributed population.



- In the Explore dialog box, click on **OK** button. The SPSS will put the results, histograms, stemplots, descriptive statistics, and side-by-side boxplot in the OUTPUT window.

**Descriptives**

sex				Statistic	Std. Error
weight	Female	Mean		134.89	7.983
		95% Confidence Interval for Mean	Lower Bound	116.48	
			Upper Bound	153.30	
		5% Trimmed Mean		133.99	
		Median		135.00	
		Variance		573.611	
		Std. Deviation		23.950	
		Minimum		106	
	Maximum		180		
	Range		74		
	Interquartile Range		38		
	Skewness		.705	.717	
	Kurtosis		.173	1.400	
	Male	Mean		192.85	9.035
		95% Confidence Interval for Mean	Lower Bound	173.16	
			Upper Bound	212.53	
5% Trimmed Mean			190.11		
Median			185.00		
Variance			1061.141		
Std. Deviation			32.575		
Minimum			150		
Maximum		285			
Range		135			
Interquartile Range		25			
Skewness		1.990	.616		
Kurtosis		5.449	1.191		

\* Tests of normality for weight variable, one for male and one for female.

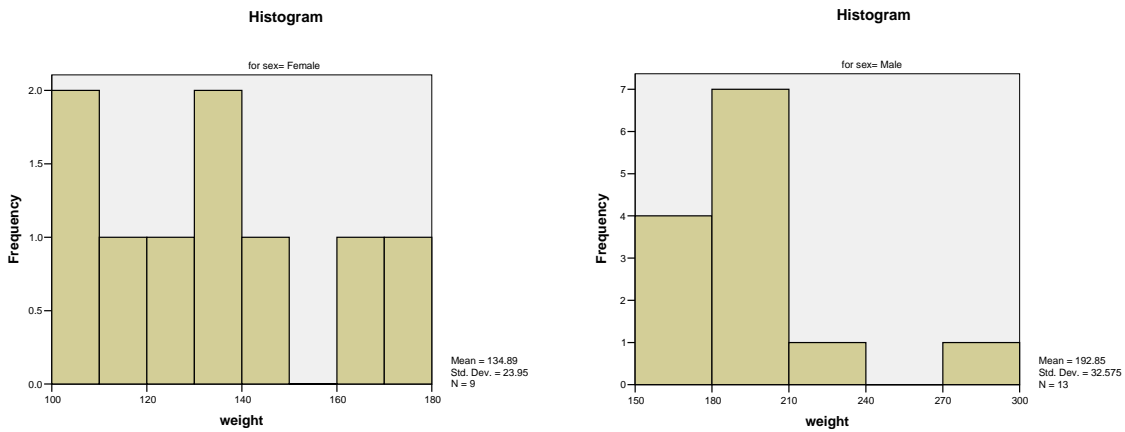
**Tests of Normality**

sex	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
weight Female	.165	9	.200*	.945	9	.633
weight Male	.243	13	.035	.812	13	.010

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

\* Two histograms for weights, one for male and one for female.



\* Side-by-side box plot for comparing weight between male and female students.

