Data used in this guide: 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 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 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 a dark triangular shape in it, to select the variable into Dependent List box.
3. 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.

4. In the Explore dialog box, click on **OK** button. The SPSS will put the results, histogram, stemplot and descriptive statistics, in the **OUTPUT** window.

5. 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 two 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.

   Click the variable to be selected for analysis from the list of variables on the left and click the select button, the button with a dark triangular shape in it, to select the variable into Variables(s) box.
3. 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.

![Frequencies Charts dialog box](image)

4. 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.

<table>
<thead>
<tr>
<th>sex</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid Female</td>
<td>9</td>
<td>40.9</td>
<td>40.9</td>
<td>40.9</td>
</tr>
<tr>
<td>Male</td>
<td>13</td>
<td>59.1</td>
<td>59.1</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

![Bar chart](image)
Examine Relation Between Two Quantitative Variables by Chart (Scatter Plot)

1. Click and move through the following menu selections:

   **Graphs ➔ Scatter ...**

   ![Graphs window with Scatterplot selection](image)

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

   ![Scatterplot dialog box](image)

   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.
4. 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.

5. 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 and 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.
3. 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.

![Crosstabs: Cell Display dialog box](image)

4. 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.

![Crosstabs dialog box](image)

To perform a chi-square test, click on Statistics… button and check the Chi-square box. The chi-square test results will be displayed in SPSS output window after clicking on OK from the Crosstabs dialog box.

5. 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.

![Contingency table and clustered bar chart](image)
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 ⇒ Bar …

   ![Graphs Bar dialog box](image)

   1. 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.

   ![Bar Charts dialog box](image)
3. 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 are 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.
3. 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.

4. 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**.

---

<table>
<thead>
<tr>
<th>sex</th>
<th>Statistic</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight</td>
<td>Female</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% Confidence Interval for Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5% Trimmed Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std. Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maximum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interquartile Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skewness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kurtosis</td>
</tr>
<tr>
<td>Male</td>
<td>Mean</td>
<td>192.85</td>
</tr>
<tr>
<td></td>
<td>95% Confidence Interval for Mean</td>
<td>Lower Bound</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Upper Bound</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5% Trimmed Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std. Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maximum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interquartile Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skewness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kurtosis</td>
</tr>
</tbody>
</table>
Tests of Normality

<table>
<thead>
<tr>
<th>sex</th>
<th>Statistic</th>
<th>df</th>
<th>Sig.</th>
<th>Statistic</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight Female</td>
<td>.165</td>
<td>9</td>
<td>.200*</td>
<td>.945</td>
<td>9</td>
<td>.633</td>
</tr>
<tr>
<td>Male</td>
<td>.243</td>
<td>13</td>
<td>.035</td>
<td>.812</td>
<td>13</td>
<td>.010</td>
</tr>
</tbody>
</table>

* 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.