

SigmaPlot Statistics

Chapter

1

Using SigmaStat Statistics in SigmaPlot

SigmaStat provides a wide range of powerful yet easy to use statistical analyses specifically designed to meet the needs of researchers, without requiring in-depth knowledge of the math behind the procedures performed. You can find SigmaStat on SigmaPlot's **Analysis** tab.

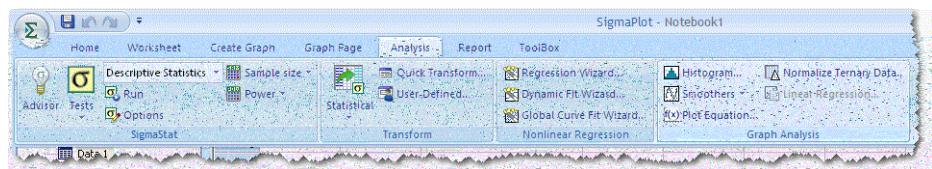


Figure 1: SigmaStat appears as its own group on the Analysis tab.

The tests and features described in this section include:

- [The Advisor Wizard](#) on page 5.
- [Using Statistical Procedures](#) on page 15.
- [Using SigmaStat Statistics in SigmaPlot](#) on page 3.
- [Comparing Two or More Groups](#) on page 53.
- [Comparing Repeated Measurements of the Same Individuals](#).
- [Comparing Frequencies, Rates, and Proportions](#) on page 187.
- [Principal Components Analysis](#) on page 215.
- [Prediction and Correlation](#) on page 225
- [Survival Analysis](#) on page 307.
- [Computing Power and Sample Size](#) on page 349.
- [Report Graphs](#) on page 373.

Chapter

2

The Advisor Wizard

Topics:

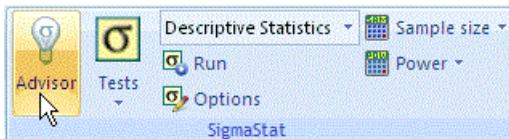
- Using the Advisor Wizard
- Select what you need to do
- How are the data measured?
- Does your data include potential risk factors that may affect survival time?
- Did you apply more than one treatment per subject?
- How many groups or treatments are there?
- What kind of data do you have?
- What kind of prediction do you want to make?
- What kind of curve do you want to use?
- How do you want to specify the independent variables?
- How do you want SigmaPlot to select the independent variable?
- Is your study retrospective or prospective?

Use the Advisor Wizard to help you to determine the appropriate test to use to analyze your data.

Using the Advisor Wizard

To use the Advisor Wizard:

1. On the **Analysis** tab, in the **SigmaStat** group, click **Advisor**.



2. When the **Advisor Wizard** appears, answer the questions about what you want to do and the format of your data. Click **Next** to go to the next panel, **Back** to go to the preceding panel, **Finish** to view the suggested test, or **Cancel** to close the **Advisor Wizard**.
3. After the **Advisor Wizard** suggests a test, click **Run** to perform the test. The **Select Data** panel for the suggested test appears prompting you to select the worksheet columns with the data you want to test.

The remainder of this section describes the answers for each dialog box.

Select what you need to do

The first step in assigning a test appropriate to your data is defining what you want to accomplish.

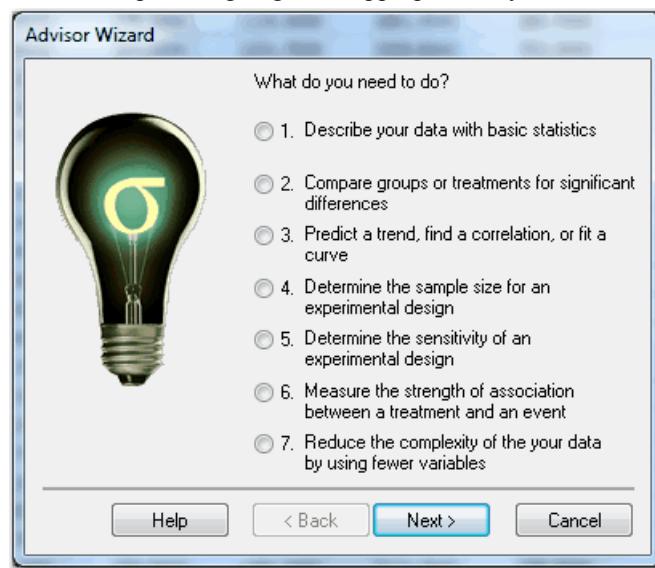


Figure 2: The Advisor Wizard

The **Advisor Wizard** begins by asking you if you need to:

- **Describe your data with basic statistics.** Select this option if you want to view a list of descriptive statistics for one or more columns of data.

For more information, see [Describing Your Data with Basic Statistics](#) on page 18.

- **Compare groups or treatments for significant differences.** Select this option if you want to compare data for significant differences, for example, if you want to compare the mean blood pressure of people who are receiving different drug treatments. The data to be compared can be the data collected from different groups, the data for different treatments on the same subjects, or the distributions or proportions of different groups. Click **Next**. You are asked to describe how your data is measured. For more information, see [Compare groups or treatments for significant differences](#).

- **Predict a trend, find a correlation, or fit a curve.** Select this option if you want to use regression to predict a dependent variable from one or more independent variables, or describe the strength of association between two variables with a correlation coefficient. For example, select this option if you want to see if you can predict the average caloric intake of an animal from its weight. Click **Next**. You are asked to describe how your data is measured. For more information, see [How are the data measured?](#) on page 7.
- **Determine the sample size for an experimental design.** Select this option if you want to determine the desired sample size for an experiment you intend to perform. Click **Next**. You are asked to describe how your data is measured. For more information, see [Determine the sample size for an experimental design](#).
- **Determine the sensitivity of an experimental design.** Select this option to determine the power or ability of a test to detect an effect for an experiment you want to perform. Click **Next**. You are asked to describe how your data is measured. For more information, see [Determine the sensitivity of an experimental design](#).
- **Measure the strength of association between a treatment and an event.** Select this option if you want to measure the strength of association between a treatment or risk factor and a specified event that occurs in members of a population. Click **Next**. You are asked if your study is retrospective or prospective. For more information, see [Measure the strength of association between a treatment and an event](#).
- **Reduce the complexity of your data by using fewer variables.** Select this option if your data is high dimensional and you want to approximate it using fewer dimensions. Click **Finish**. For more information, see [Reduce the complexity of your data by using fewer variables](#).

How are the data measured?

You need to define how your data are measured to determine which test to perform for most procedures.

There are four ways data can be measured:

1. **By numeric values.** Select **By numeric values** if your data are measured on a continuous scale using numbers. Examples of numeric values include height, weight, concentrations, ages, or any measurement where there is an arithmetic relationship between values.
 - If you are comparing groups or treatments for differences, you are asked if [you have repeated observations on the same individuals](#). See [Did you apply more than one treatment per subject?](#) on page 8 for more information.
 - If you are predicting a trend, you are prompted to [select the type of prediction you want to perform](#). See [What kind of prediction do you want to make?](#) on page 10 for more information.
 - If you are determining the sample size of or the sensitivity of an experimental design, you are asked [how many groups or treatments you have](#). See [How many groups or treatments are there?](#) on page 8 for more information.
2. **By order or rank.** Select **By order or rank** if your data are measured on a rank scale that has an ordering relationship, but no arithmetic relationship, between values.

For example, clinical status is often measured on an ordinal scale, such as: Healthy = 1; Feeling ill = 2; Sick = 3; Hospitalized = 4; and Dead = 5. These ratings show that being dead is worse than being healthy, but they do not indicate that being dead is five times worse than being healthy.

- If you are comparing groups or treatments for differences, you are asked [if you have repeated observations on the same individuals](#). See [Did you apply more than one treatment per subject?](#) on page 8 for more information.
 - If you are predicting a trend, click **Finish**. The Advisor suggests computing the [Spearman Rank Correlation](#). See [Spearman Rank Order Correlation](#) on page 299 for more information.
3. **By proportion or number of observations (for example, male vs. female).** Select **By proportion or number of observations** in categories if your data is measured on a nominal scale, which counts the number or proportions that fall into categories, and where there is no relationship between the categories (such as Democrat versus Republican).
 4. **By survival time.** Select **By survival time** if you have measurements that correspond to the time to an event. This event is typically a death but other events like the time to motor failure or the time to vascular graft closure are equally valid.

Does your data include potential risk factors that may affect survival time?

- **No.** Select **No** if all data is considered equally accurate and the **Advisor Wizard** will suggest use of the **LogRank test**. For more information, see [Survival LogRank Analysis](#) on page 316.
- **Yes.** Select **Yes** if you think the later survival times are less accurate than the early times. This might occur, for example, when there are many more late censored values. In this case the **Advisor Wizard** will suggest use of the **Gehan-Breslow test**. For more information, see [Gehan-Breslow Survival Analysis](#) on page 324.

Did you apply more than one treatment per subject?

If you are comparing groups or treatments, or determining sample size or power and your data is measured on a continuous numeric scale, you must specify whether the observations were, or are to be made, on the same or different subjects. Select **Yes** or **No**, then click **Next**.

- **Yes.** Answer **Yes** if the observations are different treatments made on the same subjects. Select **Yes** when you are comparing the same individuals before and after one or more different treatments or changes in condition.

For example, you would select **Yes** if you were testing the effect of changing diet on the cholesterol level of experimental subjects, or if you were taking an opinion poll of the same voters before and after a political debate.

- If you are comparing groups on an arithmetic or rank scale, you are asked to [specify the number of groups or treatments](#). See [How many groups or treatments are there?](#) on page 8 for more information.
- If you are comparing group proportions or distribution in categories, click **Finish**. SigmaPlot suggests performing **McNemar's Test** which you can learn about in [McNemar's Test](#) on page 203. There are also descriptions available of the [results for this procedure](#) which you can read more about in [Interpreting Results of McNemar's Test](#) on page 207.
- **No.** Answer **No** if each observation was obtained from a different subject. If you are seeing if there is a difference between different groups, such as comparing the weights of three different populations of elephants, you are not repeating observations. You should only select **Yes** if you are comparing the same individuals before and after one or more treatments.
 - If you are comparing groups on an arithmetic or rank scale, you are asked to [specify the number of groups or treatments](#). See [How many groups or treatments are there?](#) on page 8.
 - If you are comparing group proportions or distribution in categories, you are asked [what kind of data you have](#). See [What kind of data do you have?](#) on page 10.

How many groups or treatments are there?

When comparing groups or treatments or determining sample size or power and your data is measured on a continuous numeric or rank scale, SigmaPlot asks you how many treatments or conditions are involved. After specifying the number of groups, you are asked more questions, or a test is suggested.

- i** **Tip:** Click **Finish** to view the suggested test, then **Run** to perform it. You can also click **Back** to return to the previous dialog box, **Cancel** to return to the worksheet, or click **Help** for information on using the **Advisor Wizard**.

Select one of the following:

One. Select this option if you have only one different experimental group.

Two. Select this option if you have two different experimental groups or if your subjects underwent two different treatments.

For example, if you are comparing differences in hormone levels between men and women, or if you are measuring the change in individuals before and after a drug treatment, there are two groups.

- If you are comparing two different groups on an arithmetic scale, SigmaPlot suggests the [independent t-test](#) which you can read more about in [Unpaired t-Test](#) on page 57. You can read descriptions of the results for this procedure in [Interpreting t-Test Results](#) on page 62.
- If you are [determining sample size or power](#) for a comparison of two groups on an arithmetic scale, SigmaPlot suggests that you perform t-test sample size or power computations. You can also determine the power. See [Computing Power and Sample Size](#) on page 349.
- If you are comparing the same subjects undergoing two different treatments on an arithmetic scale, SigmaPlot suggests performing the [Paired t-test](#). You can learn about this in [Paired t-Test](#) on page 141. You can also read descriptions of the results for this procedure. See [Interpreting Paired t-Test Results](#) on page 146.
- If you are determining sample size or power for a comparison of the same subjects undergoing two treatments on an arithmetic scale, SigmaPlot suggests performing Paired t-test sample size or power computations. You can also read directions on determining power. See [Computing Power and Sample Size](#) on page 349.
- If you are comparing two different groups on a rank scale, SigmaPlot suggests performing the [Mann-Whitney Rank Sum Test](#). See [Mann-Whitney Rank Sum Test](#) on page 66. You can also read descriptions of the results for this procedure in [Interpreting Rank Sum Test Results](#) on page 71.
- If you are comparing the same subjects undergoing two different treatments on a rank scale, SigmaPlot suggests performing the [Wilcoxon Signed Rank Test](#). You can also read descriptions of the results for this procedure in [Interpreting Signed Rank Test Results](#) on page 154.

Three or more. Select this option if your group has three or more different groups to compare, or are comparing the response of the same subjects to three or more different treatments.

For example, if you collected ethnic diversity data from five different cities, or subjected individuals to a series of four dietary changes and measured change in serum cholesterol, you are analyzing three or more groups.

- If you are comparing three or more different groups on an arithmetic scale, SigmaPlot suggests performing [One Way ANOVA](#). See [One Way Analysis of Variance \(ANOVA\)](#) on page 73.
- If you are determining sample size or power for a comparison of three or more different groups on an arithmetic scale, SigmaPlot suggests performing One Way ANOVA sample size computations. You can also perform power computations. See [Computing Power and Sample Size](#) on page 349.
- If you are comparing the same subjects undergoing three or more different treatments on an arithmetic scale, SigmaPlot suggests performing [One Way Repeated Measures ANOVA](#). See [One Way Repeated Measures Analysis of Variance \(ANOVA\)](#) on page 157. You can also read descriptions of the results for this procedure in [Interpreting One Way Repeated Measures ANOVA Results](#) on page 163.
- If you are comparing three or more different groups on a rank scale, SigmaPlot suggests the [Kruskal-Wallis ANOVA on Ranks](#). See [Kruskal-Wallis Analysis of Variance on Ranks](#) on page 114. You can also read descriptions of the results for this procedure in [Interpreting ANOVA on Ranks Results](#) on page 120.
- If you are comparing the same subjects undergoing three or more different treatments on a rank scale, SigmaPlot suggests the [Friedman Repeated Measures ANOVA on Ranks](#). See [Friedman Repeated Measures Analysis of Variance on Ranks](#) on page 180. You can also read descriptions of the results for this procedure in [Interpreting Repeated Measures ANOVA on Ranks Results](#) on page 183.

There are two combinations of groups or treatments to consider (for example, males and females from different cities). Select this option if each experimental subject is affected by two different experimental factors or underwent two different treatments simultaneously. Note that different levels of a factor, such as male and female for gender, are not considered to be different factors.

For example, if you were comparing only males and females, you would have only one factor; however, if you compared males and females from different countries, there would be two factors, gender and nationality.

- If you are comparing three or more different groups on an arithmetic scale, SigmaPlot suggests performing [Two Way ANOVA](#). See [Two Way Analysis of Variance \(ANOVA\)](#) on page 89. To read descriptions of the results for this procedure, see [Interpreting Two Way ANOVA Results](#) on page 102.
- If you are comparing the same subjects undergoing three or more repeated treatments on an arithmetic scale, SigmaPlot suggests [Two Way Repeated Measures ANOVA](#). See [Two Way Analysis of Variance \(ANOVA\)](#) on page 89. Note that either one or both factors can be repeated treatments. To read descriptions of the results for this procedure, see [Interpreting Two Way ANOVA Results](#) on page 102.

There are three combinations of groups to consider. Select this option if each experimental subject is affected by three different experimental factors or underwent three different treatments simultaneously. Note that different levels of a factor, such as male and female for gender, and Italian and German for nationalities are not considered to be different factors.

For example, if you are comparing only males and females, from Italy and Germany, you have only two factors. However, if you are comparing males and females from different countries, with different diets, there are three factors, gender, nationality, and diet.

If you select this option, SigmaPlot suggests you run a Three Way ANOVA. See [Three Way Analysis of Variance \(ANOVA\)](#) on page 104.

This is a measure of the association between two variables. If you are determining power or sample size, this option also appears. SigmaPlot suggests performing power or sample size computations for a correlation coefficient.

There are two or more with covariates to reduce the random sampling variability. Select this option to:

- Reduce the unexplained variance in your dependent variable data, improving the precision of results.
-

Select one of the following:

- **Fit a straight line through the data.** Select this answer to find the slope and the intercept of the line $y=p_0+p_1x$ that most closely describes the relationship of your data, where y is the dependent variable and x is the independent variable.

If you select this option, click **Finish** to view the suggested test. SigmaPlot suggests performing a [Linear Regression](#). You can also read descriptions of the results for this procedure in [Interpreting Spearman Rank Correlation Results](#) on page 300.

- **Fit a curved line through the data.** Select this answer to find an equation that predicts the dependent variable from an independent variable without assuming a straight line relationship. If you select to fit a curved line through your data, SigmaPlot asks you [what kind of curve you want to use](#). See [What kind of curve do you want to use?](#) on page 11.
- **Predict a dependent variable from several independent variables.** Select this option if you want to predict a dependent variable from more than one independent variable using the linear relationship $y=b_0+b_1x^1+b_2x^2+b_3x^3+\dots b_kx^k$ where y is the dependent variable, $x_1, x_2, x_3, \dots, x_k$ are the k independent variables, and $b_1, b_2, b_3, \dots, b_k$ are the regression coefficients. As the values for x_1 vary, the corresponding value for y either increases or decreases proportionately.

If you select this option, SigmaPlot asks how you want to [specify the independent variables](#). See [How do you want to specify the independent variables?](#) on page 11.

- **Measure the strength of association between pairs of variables.** Select this option to find how closely the value of one variable predicts the value of another (for example, the likelihood that a variable increases or decreases when the other variable increases or decreases), without specifying which is the dependent and independent variable.

If you select this option, click **Finish**. SigmaPlot suggests computing the Pearson Product Moment Correlation.

What kind of curve do you want to use?

If you are trying to predict one variable from one or more other variables using a curved line, you are asked what kind of curve you want to use.

Select one of the following:

- **A polynomial curve with one independent variable.** Select this option if you want to use a k_{th} order polynomial curve of the form $k y=b_0+b_1x^1+b_2x^2+b_3x^3+\dots b_kx^k$ to predict the dependent variable y from the independent variable x , where $y=b_0+b_1x_1+b_2x_2+b_3x_3+\dots b_kx^k$, ..., b_k are the regression coefficients.

If you select this option, click **Finish**. SigmaPlot suggests [performing Polynomial Regression](#). See [Polynomial Regression](#) on page 260. You can also read descriptions of the results for this procedure in [Interpreting Incremental Polynomial Regression Results](#) on page 265. You can also read about interpreting Order Only Polynomial Results in [Interpreting Order Only Polynomial Regression Results](#) on page 267.

- **A general nonlinear equation.** Select this option if you want to describe your data with a nonlinear function. Common nonlinear functions include rising and falling exponential and log curves, logistic sigmoid curves, and hyperbolic curves that approach a maximum or minimum.

If you select this option, click **Finish**. SigmaPlot suggests using Nonlinear Regression.

Nonlinear Regression uses a dialog box to specify any general nonlinear equation with up to ten independent variables, then uses an iterative least squares algorithm to estimate the parameters in the regression model.

How do you want to specify the independent variables?

If you chose to predict a dependent variable from several independent variables, you can select the independent variables using two methods. The dependent variable and independent variables are selected as columns from the worksheet when the regression procedure is performed.

Select one of the following:

- **Include all selected independent variables in the equation.** Select this option if you want to compute a single equation using all independent variables you select for the equation, regardless of whether they contribute significantly to predicting the dependent variable.
If you select this option, click **Finish**. SigmaPlot suggests performing a [Multiple Linear Regression](#). See [Multiple Linear Regression](#) on page 237. You can also read descriptions of the results for this procedure in [Interpreting Multiple Linear Regression Results](#) on page 244.
- **Let SigmaPlot select the "best" variables to include in the equation.** Select this option if you want SigmaPlot to screen the potential independent variables you select and only include ones that significantly contribute to predicting the dependent variable. You are then asked how you want to [select the independent variables](#). See [How do you want SigmaPlot to select the independent variable?](#) on page 12.

How do you want SigmaPlot to select the independent variable?

If you are predicting the value of one variable from other variables, and you want SigmaPlot to screen potential variables for their contribution to the predictive value of the regression equation, you can select three different methods.

- **Sequentially add new independent variables to the equation.** Select this option to select the independent variables for the equation by starting with no independent variables, then adding variables until the ability to predict the dependent variable is no longer improved. The variables are added in order of the amount of predictive ability they add to the model.

The predictive ability of models produced with forward stepwise regression is measured by their ability to reduce the residual sum of squares in the regression equation.

If you select this option, click **Finish**. SigmaPlot suggests [Forward Stepwise Regression](#). See [Stepwise Linear Regression](#) on page 270. You can also read descriptions of the results for this procedure in [Interpreting Stepwise Regression Results](#) on page 283.

- **Sequentially remove independent variables from the equation.** Select this option to select the independent variables for the equation by starting with all independent variables in the equation, then deleting variables one at a time. The variable that contributes the least to the prediction of the dependent variable is deleted from the equation first. This elimination process continues until the ability of the model to predict the dependent variable is reduced below a specified level.

The predictive ability of models produced with backwards stepwise regression is measured by their ability to reduce the residual sum of squares in the regression equation.

If you select this option, click **Finish**. SigmaPlot suggests the [Backward Stepwise Regression](#). See [Stepwise Linear Regression](#) on page 270. You can also read descriptions of the results for this procedure in [Interpreting Stepwise Regression Results](#) on page 283.

- **Consider all possible combinations of the independent variable and select the best subset.** Select this option if you want SigmaPlot to evaluate all possible regression models, and isolate the models that "best" predict the dependent variable.

If you select this option, click **Finish**. SigmaPlot suggests the [Best Subset Regression](#). You can also read descriptions of the [results for this procedure](#).

SigmaPlot selects the sets of independent variables that "best" predict the dependent variable using criteria specified in the **Best Subsets Regression Options** dialog box.

Is your study retrospective or prospective?

When measuring the strength of association between a treatment and an event, the treatment effect can be determined after the event has been observed or a treatment and control group can be sampled before the event is observed.

Select one of the following:

- **A treatment effect is to be determined after an event has been observed.** Select this option for a retrospective study, when the treatment effect is determined after the event has been observed.
- **A treatment and a control group have been sampled before an event is observed.** Select this option for a prospective study, when a treatment and a control group are sampled before the event is observed.

Chapter

3

Using Statistical Procedures

Topics:

- [Running Procedures](#)
- [Choosing the Procedure to Use](#)
- [Describing Your Data with Basic Statistics](#)
- [Describing Your Data with Frequency Tables](#)
- [Choosing the Group Comparison Test to Use](#)
- [Choosing the Repeated Measures Test to Use](#)
- [Choosing the Rate and Proportion Comparison to Use](#)
- [Choosing the Prediction or Correlation Method](#)
- [Choosing the Survival Analysis to Use](#)
- [Testing Normality](#)
- [Determining Experimental Power and Sample Size](#)

The statistical procedure you use to analyze a given data set depends on the goals of your analysis and the nature of your data. The [Advisor Wizard](#) asks you questions about your goals and your data, then selects the appropriate test. For more information, see [The Advisor Wizard](#) on page 5.

Running Procedures

In general, the steps to run a test or procedure are:

1. Entering or importing and arranging your data appropriately in the worksheet.
2. Determining and choosing the test you want to perform.
3. If desired, setting the test options using the selected test's **Options** dialog box.
4. Running the test by picking the worksheet columns with the data you want to test using the **Select Data** panel of the Test Wizard.
5. Viewing, generating, and interpreting, the test reports and graphs.

Arranging Worksheet Data

The method you use to enter or arrange data in the worksheet depends on the type of test you are running. Some data formats include:

- Data format for group comparison tests.
- Data format for repeated measures tests.
- Data format for rate and proportion tests.

Selecting a Test

To select a statistical test in SigmaPlot:

Setting Test Options

You can configure almost all statistics procedures with a set of options. Use these settings to perform additional tests and procedures. You may wish to enable or disable some of these options or change assumption checking parameters; all changes are saved between sessions.

To change option settings before you run a test:

1. Select the test.
2. Click **Options**. The **Options** dialog box for the test appears.
3. Click the tab of the options you want to view. Select a check box to include an option in the test. Clear a check box if you do not want to use that test option.

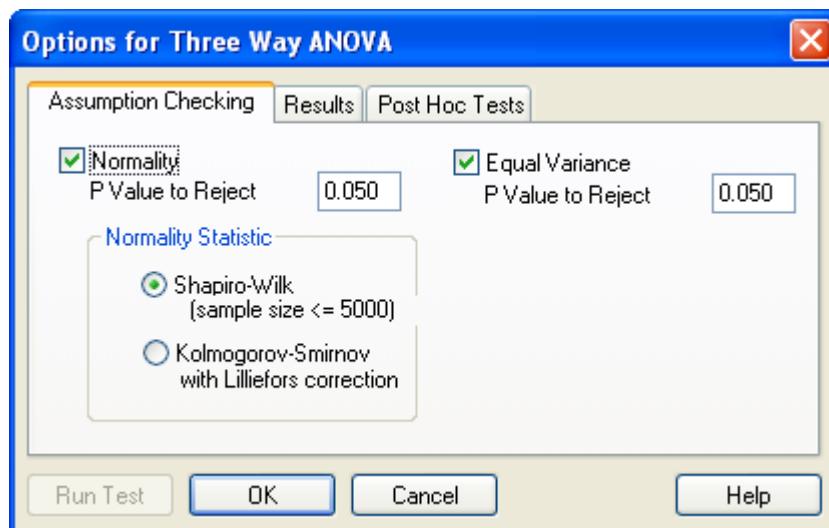


Figure 3: An Example of a Test Options Dialog Box. Each test has its own settings.

4. Click **Run Test** to continue the test.

Selecting the Data to Test

When you run a test and if you can arrange your data in more than one format, use the **Select Data** panel of the Test Wizard to select the worksheet columns with the data you want to test and to specify how your data is arranged in the worksheet.

1. Select the appropriate format from the **Data Format** drop-down list, then click **Next**.

If the test you are running uses only one type of data format, the **Select Data** panel appears prompting you to select the columns with the data you want to test (see the following step).

2. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data** drop-down list.

The dialog box indicates the type of data you are selecting.

The first selected column is assigned to the first entry in the **Selected Columns** list, and all successively selected columns are assigned to successive entries in the list. The number or title of selected columns appear in each entry. The number of columns you can select depends on the test you are running and the format of your data.

3. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.

4. If you are running a **Forward** or **Backward Stepwise Regression**, click **Next**.

The **Select Data dialog box** appears.

5. Click **Finish** to perform the test on the data in the selected columns.

After the computations are completed, the report appears.

Reports and Result Graphs

Test reports automatically appear after a test has been performed.

To generate a result graph:

1. Make sure the report is the active window. If it isn't, click the **Report** tab.
2. Click **Create Result Graph**.

The **Create Result Graph** dialog box appears, from which you can select an available **Graph Type** and create a graph.

SigmaPlot does not create graphs for rates and proportion tests, best subset and incremental polynomial regression reports and normality reports.

 **Attention:** If you close a report without generating or saving a graph, the graph is not recoverable.

Editing, Saving, and Opening Reports and Graphs

You can edit reports and graphs using the **Format** and **Tools** groups on the **Graph Page** tab, and **Graph Properties**. You can also export reports as non-notebook files and edit them in other applications.

Repeating Tests

Repeating a test involves running the last test you performed, using the same worksheet columns. To repeat a test using new data columns, click **Run** on the **Analysis** tab.

To repeat a test using the same worksheet columns:

1. Make sure the last test you performed is displayed in the drop-down list in the **SigmaStat** group. The **Rerun** button appears enabled below the **Run** button.
2. If desired, edit the data in the columns used by the test. You can add data and change values and column titles.
3. **To change the option settings before you rerun the test**, click **Options**, change the desired options, then click **OK** to accept the changes and close the dialog box.
4. Click **Rerun**.

The **Select Data** panel box appears with the columns used in the last procedure selected.

- Click **Finish** to repeat the procedure using these columns. After the computations are complete, a new report appears.

Choosing the Procedure to Use

You can use SigmaPlot to perform a wide range of statistical procedures. [The Advisor can suggest which test to use](#). You can also determine the appropriate test yourself. The type of procedure to choose depends on the kind of analysis you want to perform.

- Use descriptive statistics to compute a number of commonly used statistical values for the selected data.
- Use group comparison tests to analyze two or more different sample groups for statistically significant differences.
- Use repeated measures comparisons to test the differences in the same individuals before and after one or more treatments or changes in condition.
- Use rate and proportion analysis to compare the distribution of groups that are divided or fall into different categories or classes (for example, male versus female, or reaction versus no reaction).
- Use [survival](#) to determine statistics about the time to an event and to compare two or more time-to-event data sets.
- Use [power and sample size determination](#) to calculate the sensitivity, or power, of an experimental test, or to compute the experimental sample size required to achieve a desired sensitivity.
- Use [Odds Ratio](#) or [Relative Risk](#) to measure the strength of association between some event and a treatment or risk factor.

Table 1: Procedures to Use for Statistical Tests

Scale of Measurement	Type of Experiment				
	Two groups of different individuals	Three or more groups of different individuals	Same individuals before and after a single treatment	Same individuals after multiple treatments	Predict a variable or find an association between variables
Numeric, normally distributed with equal variances	Unpaired <i>t</i> -test	One Way or Two Way ANOVA	Paired <i>t</i> -test	One Way or Two Way Repeated Measures ANOVA	Regression or Pearson Product Moment Correlation
By rank or order, or numeric, but non-normally distributed or with unequal variances	Mann-Whitney Rank Sum Test	Kruskall-Wallis ANOVA on Ranks	Wilcoxon Signed Rank Test	Friedman Repeated Measures ANOVA on Ranks	Spearman Rank Order
By distribution in different categories	Chi-Square Analysis of Contingency Tables	Chi-Square Analysis of Contingency Tables	McNemar's Test	Not Available	Not Available

All statistical procedure commands are found on the **SigmaStat** group on the **Analysis** tab.

Describing Your Data with Basic Statistics

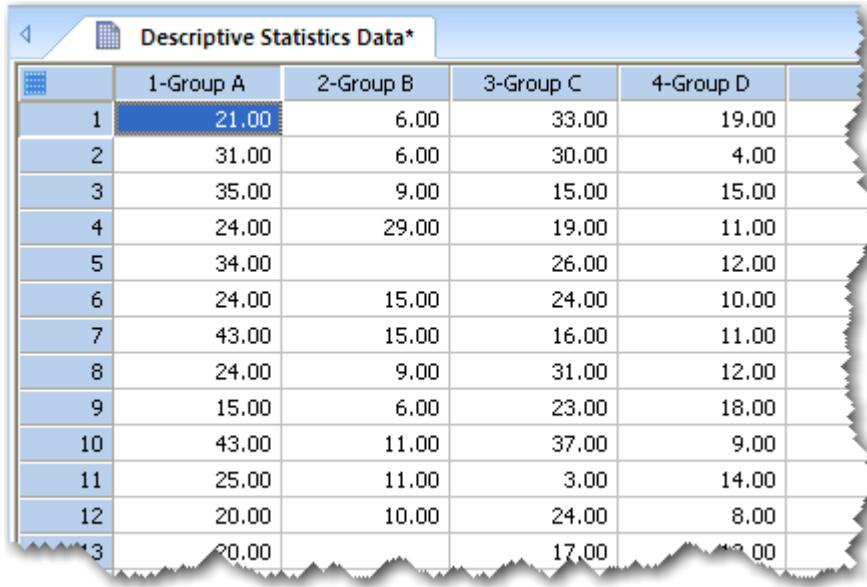
You can use SigmaPlot to describe your data by computing basic statistics, such as the mean, median, standard deviation, percentiles, etcetera, that summarize the observed data.

Describing your data involves:

- Arranging your data in the appropriate format.
- Setting descriptive statistic options.
- Selecting the columns you want to compute the statistics for.
- Viewing the descriptive statistics results.

Arranging Descriptive Statistics Data

Descriptive Statistics are performed on columns of data, so you should arrange the data for each group or variable you want to analyze in separate columns.



	1-Group A	2-Group B	3-Group C	4-Group D
1	21.00	6.00	33.00	19.00
2	31.00	6.00	30.00	4.00
3	35.00	9.00	15.00	15.00
4	24.00	29.00	19.00	11.00
5	34.00		26.00	12.00
6	24.00	15.00	24.00	10.00
7	43.00	15.00	16.00	11.00
8	24.00	9.00	31.00	12.00
9	15.00	6.00	23.00	18.00
10	43.00	11.00	37.00	9.00
11	25.00	11.00	3.00	14.00
12	20.00	10.00	24.00	8.00
13	0.00		17.00	2.00

Figure 4: Data Arrangement with Treatments or Groups in Columns

Selecting Data Columns

You can calculate statistics for entire columns or only a portion of columns. When running the descriptive statistics procedure, you can:

- Select the columns or block of data before you run the test, or
- Select the columns while running the test.

i **Tip:** **To calculate statistics for only a range of data,** select the data before you run the test. You can select a minimum of one column and a maximum of 32 columns when describing data.

Setting Descriptive Statistics Options

You select the statistics that you would like to calculate in the **Descriptive Statistics Options** dialog box.

To change descriptive statistics test options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. **To open the Options for Descriptive Statistics dialog box,** on the **Analysis** tab, click **Descriptive Statistics** from the drop-down list in the **SigmaStat** group.

3. Click Options.

The **Options for Descriptive Statistics** dialog box appears.

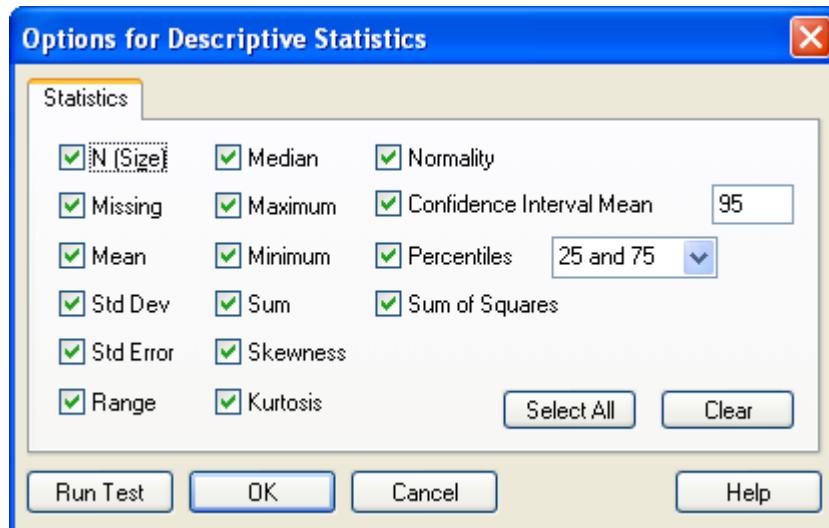


Figure 5: The Options for Descriptive Statistics dialog box

4. Clear any of the **selected statistics settings** you do not want to include in the report.

The specific summary statistics that are appropriate for a given data set depend on the nature of the data. If the observations are normally distributed, then the mean and standard deviation provide a good description of the data. If not, then the median and percentiles often provide a better description of the data.

5. **To change the confidence interval**, enter any number from 1 to 99 (95 and 99 are the most commonly used intervals) into the **Confidence Interval Mean** box.

6. **To change the percentile or confidence intervals computed**, edit the values in the **Percentile** box.

7. **To select all statistics options**, click **Select All**. To clear all selections, click **Clear**.

8. Click **Run Test** to perform the test with the selected options settings.

Tip: To set the number of decimal places displayed, click the **Sigma Button**, and then click **Options**. In the **Options** dialog box, click the **Report** tab, and select **Number of significant digits**.

Running the Descriptive Statistics Test

If you want to select your data before you run the procedure, drag the pointer over your data.

To describe your data:

1. On the **Analysis** tab, in the **SigmaStat** group, click the **Tests** drop-down list, and then select **Describe Data**. The **Descriptive Statistics - Select Data** dialog box appears prompting you to specify a data format.

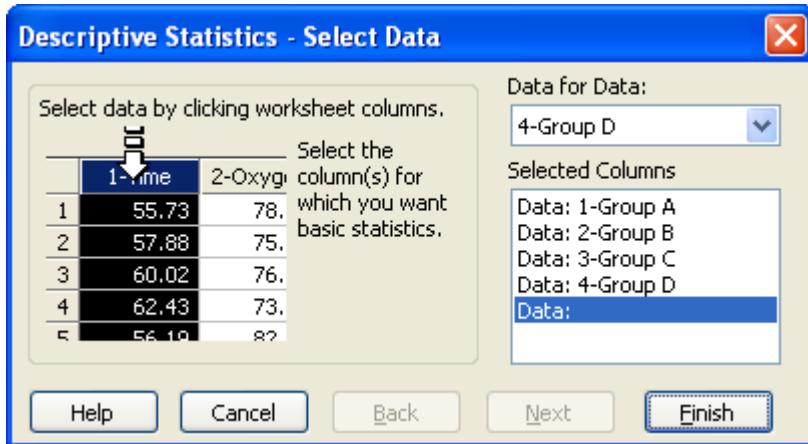


Figure 6: The Descriptive Statistics - Select Data Dialog Box

- i** **Tip:** If you selected columns before you chose the test, the selected columns automatically appear in the **Selected Columns** list.
2. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list. The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The number or title of selected columns appear in each row. You can select up to 64 columns of data for the **Descriptive Statistics Test**.
 3. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
 4. Click **Finish** to describe the data in the selected columns. After the computations are completed, the report appears.

Descriptive Statistics Results

The following statistics can be calculated and displayed in the results report. These values are calculated for each column selected. Select the specific statistics to compute in the **Options for Descriptive Statistics** dialog box.

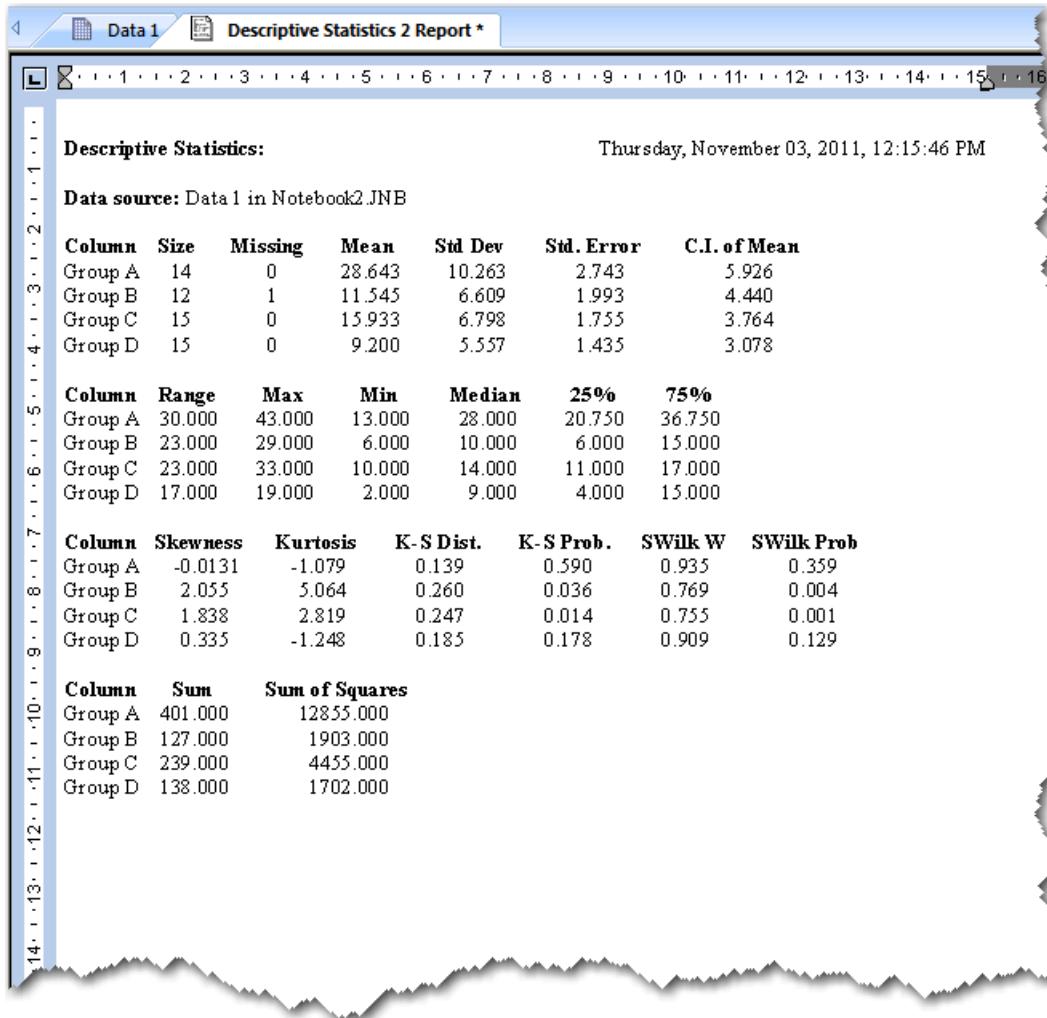


Figure 7: Descriptive Statistics Results Report

- **Size.** This is the number of non-missing observations in a worksheet column.
- **Missing.** This is the number of missing observations in a worksheet column.
- **Mean.** The mean is the average value for a column. If the observations are normally distributed, the mean is the center of the distribution.
- **Standard Deviation.** Standard deviation is a measure of data variability about the mean.
- **Standard Error of the Mean.** The standard error of the mean is a measure of how closely the sample mean approximates the true population mean.
- **Range.** The range is the minimum values subtracted from the maximum values.
- **Maximum.** Maximum is the largest observation.
- **Minimum.** Minimum is the smallest observation.
- **Median.** The medi

- **Skewness.** Skewness is a measure of how symmetrically the observed values are distributed about the mean. A normal distribution has skewness equal to zero.
- **Kurtosis.** Kurtosis is a measure of how peaked or flat the distribution of observed values is, compared to a normal distribution. A normal distribution has Kurtosis equal to zero.
- **K-S Distance.** The Kolmogorov-Smirnov distance is the maximum distance between the sample cumulative distribution function of your data and the theoretical Gaussian distribution function based on the mean and variance estimates from your data.
- **K-S Probability.** The Kolmogorov-Smirnov probability is the probability that the distance between a histogram of randomly selected data from the population and the theoretical Gaussian distribution curve is greater than the K-S Distance based on your sample.
- **Shapiro-Wilk W.** The Shapiro-Wilk W-statistic tests the null hypothesis that your data was sampled from a normal distribution. Small values of W indicate a departure from normality.
- **Shapiro-Wilk Probability.** The Shapiro-Wilk probability is the significance probability associated with the W-statistic.
- **Normality.** Normality tests the observations for normality using either the Shapiro-Wilk or the Kolmogorov-Smirnov test.

Descriptive Statistics Result Graphs

You can generate up to five graphs using the results from a descriptive statistics graph. They include a:

- **Bar chart of the column means.** The Descriptive Statistics bar chart plots the group means as vertical bars with error bars indicating the standard deviation.
- **Scatter plot with error bars of the column means**
- . The Descriptive Statistics scatter plot graphs the column means as single points with error bars indicating the standard deviation.
- **Point plot of the column data.** The Descriptive Statistics point plot graphs all values in each column as a point on the graph.
- **Point plot of the column data with error bars plotting the column means.** The Descriptive Statistics point and column means plot graphs all values in each column as a point on the graph with error bars indicating the column means and standard deviations of each column.
- **Box plot of the percentiles and median of column data.** The Descriptive Statistics test box plot graphs the percentiles and the median of column data.

Creating a Descriptive Statistics Result Graph

To generate a graph of Descriptive Statistics report data:

1. Make sure that the **Descriptive Statistics** report is in view.

2. On the **Report** tab, in the **Result Graphs** group, click **Create Result** Graph.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the **Descriptive Statistics** report.

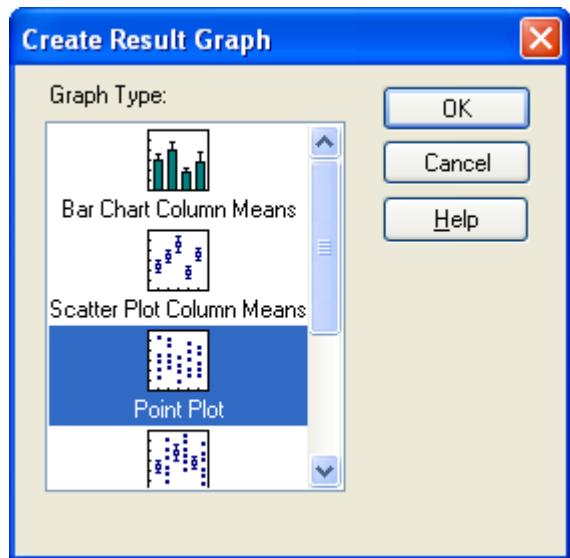


Figure 8: The Create Result Graph Dialog Box

3. Select the type of graph you want to create from the **Graph Type** list and click **OK**. The specified graph appears in a graph window or in the report.

Tip: You can also double-click the desired graph in the list.

Describing Your Data with Frequency Tables

Frequency tables describe the distribution of values in a data set in ascending order of magnitude with their corresponding frequencies.

Table 2: Example of a Frequency Table

Category	Frequency	Percent	Cum. Frequency	Cum. Percent
46 to 60	64	25	64	25
30 to 45	80	31.250	144	56.250
18 to 30	79	30.859	223	87.109
Over 60	33	12.891	256	100

Describing your data with frequency tables involves:

- Arranging your data in the appropriate format.
- Setting frequency table options.
- Selecting the columns you want to compute the statistics for.
- Viewing the frequency table results.

Arranging Frequency Table Data

Frequency tables are created from [category columns of data](#).

	1-Age	2-Gender	3
1	46 to 60	Male	
2	30 to 45	Female	
3	46 to 60		

Figure 9: Category Data Arrangement for a Frequency Table

Selecting Data Columns

You can calculate statistics for entire columns or only a portion of columns. When running the descriptive statistics procedure, you can:

- Select the columns or block of data before you run the test, or
- Select the columns while running the test.

Tip: To calculate statistics for only a range of data, select the data before you run the test. You can select a minimum of one column and a maximum of 32 columns when describing data.

Setting Frequency Table Options

You select the statistics that you would like to calculate in the **One Way Frequency Tables Options** dialog box.

To change frequency table test options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. To open the Options for One Way Frequency Tables dialog box, on the **Analysis** tab, click **Descriptive Statistics** from the drop-down list in the **SigmaStat** group.

3. Click Options.

The **Options for One Way Frequency Tables** dialog box appears.

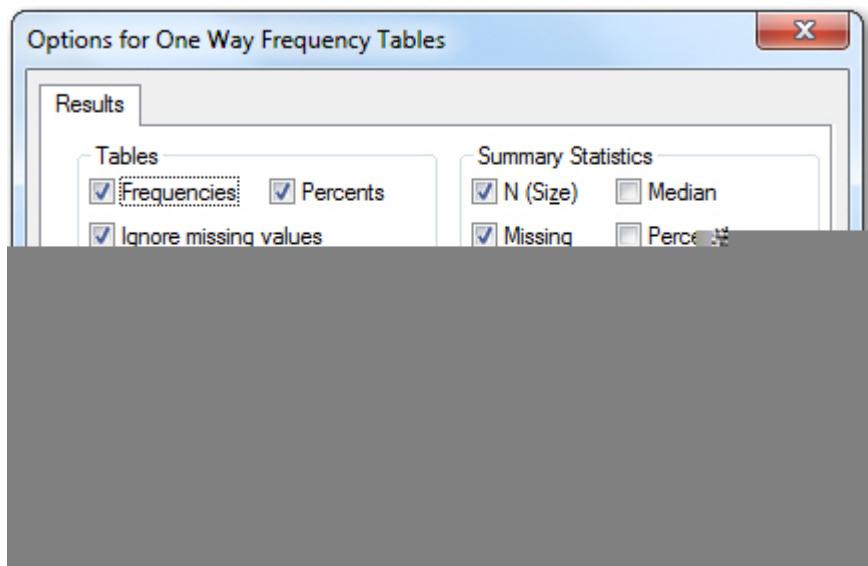


Figure 10: The Options for One Way Frequency Tables dialog box

4. Clear any of the **selected statistics settings** you do not want to include in the report.

The specific summary statistics that are appropriate for a given data set depend on the nature of the data. If the observations are normally distributed, then the mean and standard deviation provide a good description of the data. If not, then the median and percentiles often provide a better description of the data.

5. **To change the confidence interval**, enter any number from 1 to 99 (95 and 99 are the most commonly used intervals) into the **Confidence Interval Mean** box.
6. **To change the percentile or confidence intervals computed**, edit the values in the **Percentile** box.
7. **To select all statistics options**, click **Select All**. To clear all selections, click **Clear**.
8. Click **Run Test** to perform the test with the selected options settings.

Tip: To set the number of decimal places displayed, click the **Sigma Button**, and then click **Options**. In the **Options** dialog box, click the **Report** tab, and select **Number of significant digits**.

Running the Frequency Table Test

If you want to select your data before you run the procedure, drag the pointer over your data.

To describe your data:

1. On the **Analysis** tab, in the **SigmaStat** group, click the **Tests** drop-down list, and then select **One Way Frequency Table**.

The **One Way Frequency Table - Select Data** dialog box appears.

2. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

Tip: If you selected columns before you chose the test, the selected columns automatically appear in the **Select Columns** list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The number or title of selected columns appear in each row. You can select up to 64 columns of data for the **One Way Frequency Tables Test**.

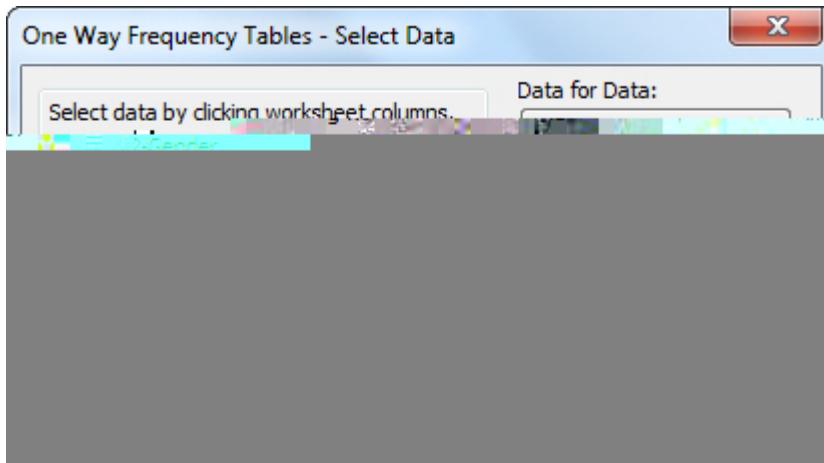


Figure 11: One Way Frequency Tables - Select Data Dialog Box

3. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
4. Click **Finish** to describe the data in the selected columns. After the computations are completed, the report appears.

Frequency Table Results

The following statistics can be calculated and displayed in the results report. These values are calculated for each column selected. Select the specific statistics to compute in the **Options for One Way Frequency Tables** dialog box.

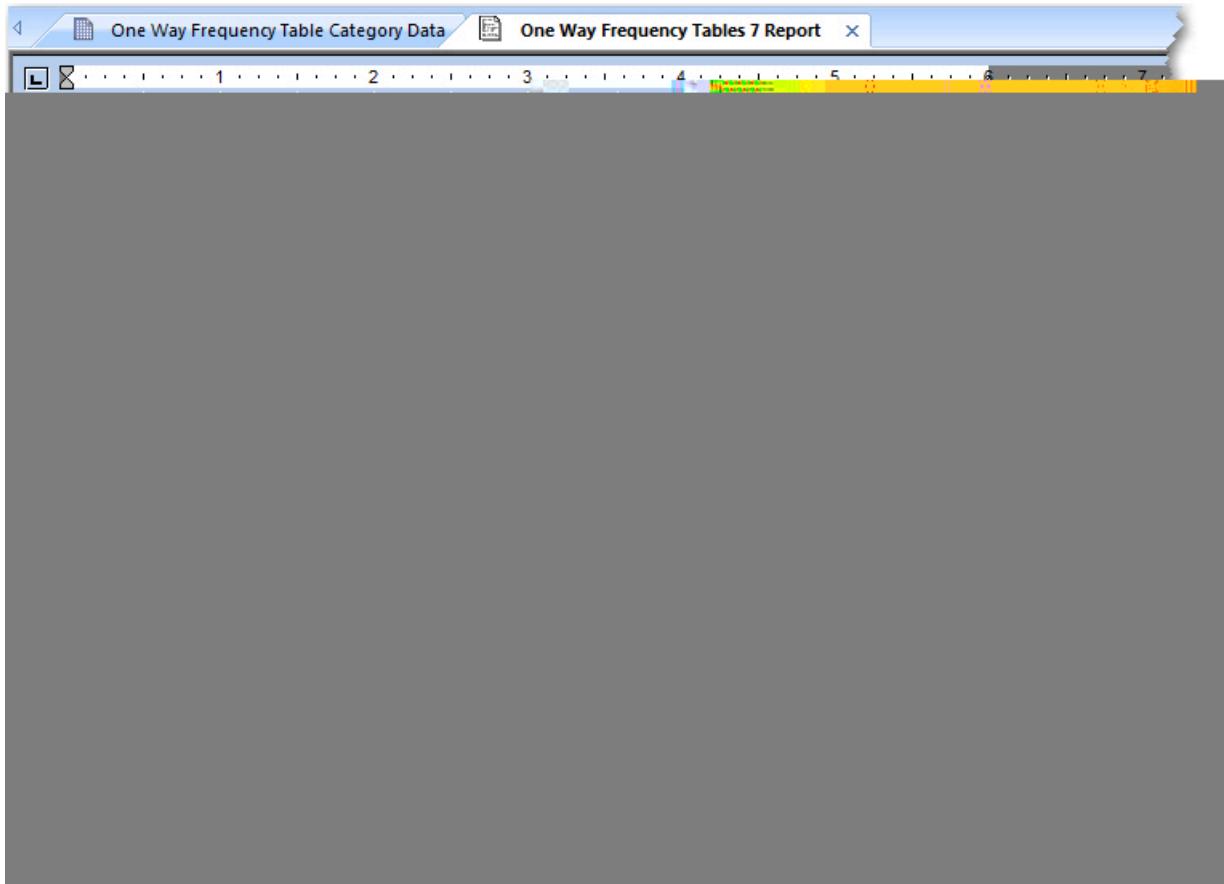


Figure 12: One Way Frequency Tables Results Report

Header. This includes the name of the analysis, date stamp, and data source, as for all other statistical procedures.

Column Statistics. This section includes one or two tables of descriptive statistics for the numeric data in each column selected for analysis. The types of statistics displayed are determined by the settings in the Test Options dialog. The *mode* is the only statistic which can have a text category as a value (see example report below). If a selected column has no numeric data, then missing values will be entered for the other statistics.

If no statistics are selected in the Test Options dialog box, then this section does not appear.

Dataset sections. There is one section in the report for each data set (worksheet column) that is analyzed. The heading of each section has the form *Data Set <number>*. This heading is followed by the column title of the data set. The columns are presented in the same order as they were chosen in the Test Wizard. Each data set section is divided into one to three subsections, depending on the settings in the Test Options dialog box:

- **Frequency Table.** This subsection displays the frequency table of the data. This subsection is always shown, although the number of columns is affected by the settings in the Test Options dialog box.
- **Confidence Intervals for Proportions.** This subsection displays the confidence intervals, in terms of percentages, for each category to measure the accuracy of the percent contribution of that category in the data set to the percent contribution of that category in the population. The name of the method that is selected for computing confidence intervals in the Test Options dialog box is shown after the heading for this subsection. If **Confidence** is left clear in the Test Options dialog box, which is the default, then this subsection does not appear.
- **(Pearson) Chi-Square Test for the Equality of Proportions.** This subsection displays the value of the statistic, degrees of freedom, and P value to test the hypothesis that the proportions of the categories in the underlying population for the data set are equal. The statistic asymptotically follows a chi-square distribution. There is a statement that follows the numeric results of this test to indicate the presence of a significant difference in the proportions (or percentages) of the categories in the population. If **Chi-square** is left clear in the Test Options dialog box, which is the default, then this subsection does not appear.

Frequency Tables Results Graphs

You can generate a frequency bar chart from a one way frequency table report.

A frequency bar chart is a bar chart with category names on the x-axis and the contribution of each category to the data set shown on the y-axis.

You can scale the y-axis as either Frequency or Percent. The dialog box will also contain a combo box for selecting the data column that will be represented in the graph. An example of this graph is shown below. There may be a limit on the number of categories that can be represented.

Creating a One Way Frequency Table Result Graph

To generate a graph of One Way Frequency Table report data:

1. Make sure that the **One Way Frequency Table** report is in view.
2. On the **Report** tab, in the **Result Graphs** group, click **Create Result Graph**. The **Create Result Graph** dialog box appears displaying the type of graph available for the **One Way Frequency Table** report.
3. Select the type of graph you want to create from the **Graph Type** list and click **OK**. The specified graph appears in a graph window or in the report.

 **Tip:** You can also double-click the desired graph in the list.

Choosing the Group Comparison Test to Use

Use the various group comparison procedures to test sample means or medians for differences.

The Advisor Wizard prompts you to answer questions about your data and goals, then selects the appropriate test; however, if you are already familiar with the comparison requirements, you can go directly to the appropriate test. The criteria used to select the appropriate procedure include:

- **The number of groups to compare.** Are you comparing two different groups or many different groups?
- **The distribution of the sample data.** Is the source population for your sample distributed along a normal "bell" (Gaussian) curve, or not? Comparisons of samples from normal populations use parametric tests, which are based on the mean and standard variation parameters of a normally distributed population. If the populations are not normal, a nonparametric, or distribution-free test must be used, which ranks the values along a new ordinal scale before performing the test.

 **Tip:** SigmaPlot can automatically test for assumptions of normality and equal variance.

When to Compare Two Groups

If you collected data from two different groups of subjects (for example, two different species of fish or voters from two different parts of the country), use a two group comparison to test for a significant difference beyond what can be attributed to random sampling variation.

When to Use a t-test versus a Mann-Whitney Rank Sum Test

You can perform two kinds of two group comparison tests: an unpaired **t-test** and the Mann-Whitney **Rank Sum Test**.

- Choose the unpaired **t-test** if your samples were taken from normally distributed populations and the variances of the two populations are equal. The **unpaired t-test** is a parametric test which directly compares the sample data.
- If your samples were taken from populations with non-normal distribution and/or unequal variances, choose the Mann-Whitney **Rank Sum Test**. The Mann-Whitney **Rank Sum Test** arranges the data into sets of rankings, then performs an unpaired t-test on the sum of these ranks, rather than directly on the data.

- If your samples are already ordered according to qualitative ranks, such as poor, fair, good, and very good, use the Mann-Whitney **Rank Sum Test**.

The advantage of the t-test is that, assuming normality and equal variance, it is slightly more sensitive (for example, it, has greater power) than the Mann-Whitney Rank Sum Test. When these assumptions are not met, the Mann-Whitney **Rank Sum Test** is more reliable.

- i** **Tip:** You can tell SigmaPlot to analyze your data and test for normal distribution and equal variance. If assumptions of normality and equal variance are violated, the alternative parametric or nonparametric test is suggested. Activate and configure assumption tests in the t-test and Mann-Whitney Rank Sum Test Options dialog boxes.

SigmaPlot tests for normality using either the Shapiro-Wilk or the Kolmogorov-Smirnov test, and for equal variance using the Levene Median test.

When to Compare Many Groups

If you collected data from three or more different groups of subjects, use one of the ANOVA (analysis of variance) procedures to test if there is difference among the groups beyond what can be attributed to random sampling variation.

There are four procedures available:

- [The single factor or One Way ANOVA](#).
- [The Two Way ANOVA](#).
- [The Three Way ANOVA](#).
- [The Kruskal-Wallis Analysis of Variance on Ranks](#).
- Choose One, Two, or Three Way ANOVA if the samples were taken from normally distributed populations and the variances of the populations are equal. The One, Two, and Three Way ANOVAs are parametric tests which directly compare the samples arithmetically.
- If your samples were taken from populations with non-normal distribution and/or unequal variance, choose the Kruskall-Wallis ANOVA on ranks, which is the nonparametric analog of the one way ANOVA. The Kruskall-Wallis ANOVA on ranks arranges the data into sets of rankings, then performs an analysis of variance based on these ranks, rather than directly on the data, so it does not require assuming normality and equal variance.

The advantage of parametric ANOVAs are that, when the normality and equal variance assumptions are met, they are slightly more sensitive (for example, have greater power) than the analysis based on ranks. When the assumptions are not met, the Kruskall-Wallis ANOVA on ranks is more reliable.

- 🚫 Restriction:** SigmaPlot does not have a two factor analysis of variance based on ranks.

Note that also you can tell SigmaPlot to analyze your data and tests for normal distribution and equal variance. If assumptions of normality and equal variance are violated, the alternative parametric or nonparametric test is suggested. These tests are specified in the Options dialog boxes.

To open the dialog box for the current test, click the Current Test Options button:

SigmaPlot tests for normality using either the Shapiro-Wilk or the Kolmogorov-Smirnov test, and for equal variance using the Levene Median test.

When to Use One, Two, and Three Way ANOVAs

The difference between a One, Two, and Three Way ANOVA lies in the design of the experiment that produced the data.

- Use a [One Way ANOVA](#) if there are several different experimental groups that received a set of related but different treatments (for example, *one factor*). This design is essentially the same as an [unpaired t-test](#) (a one way ANOVA of two groups obtains exactly the same P value as an unpaired t-test).
- Use a [Two Way ANOVA](#) if there were two experimental *factors* that are varied for each experimental group.
- Use a [Three Way ANOVA](#) if there are three experimental *factors* which are varied for each experimental group.

An example of when to use a **One Way ANOVA** would be when comparing biology teachers from three different states for their knowledge of evolution. The factor varied is state.

An example of when to use **Two Way ANOVA** would be when comparing teachers from the three states and with different education levels for their knowledge of evolution -- the two different factors are state and years of education. The two factor design can test three hypotheses about the state and education levels:

- There is no difference in opinion of the teachers among states.
- There is no difference in knowledge among education levels.
- There is no interaction between state and education in terms of knowledge; any differences between differing levels of education are the same in all states.

An example of when to use a **Three Way ANOVA** would be when comparing teachers male and female teachers from three different states, with different levels of education for their knowledge of evolution—the three different factors are gender, state, and years of education. The three factor design can test that:

- There is no difference in opinion of the teachers among gender.
- There is no difference in opinion of the teachers among states.
- There is no difference in knowledge among education levels.
- There is no interaction between gender, state, and education in terms of knowledge; any differences between differing levels of education are the same for all genders in all states.

How to Determine Which Groups are Different

Analysis of variance techniques (both parametric and nonparametric) test the hypothesis of no differences between the groups, but do not indicate what the differences are. You can use the *multiple comparison procedures* (post-hoc tests) provided by SigmaPlot to isolate these differences.

To always test for differences among the groups select **Always Perform** on the **Post Hoc Tests** tab in the ANOVA options dialog boxes. You can also specify to use multiple comparisons to test for a difference only when the ANOVA P value is significant by selecting the **Only When ANOVA P Value is Significant** option, then select the desired P value.

The specific multiple comparisons procedures to use for each ANOVA are selected in the **Multiple Comparison Options** dialog box. To open:

On the **Analysis** tab, in the **Statistic** group, click **Options**.

Choosing the Repeated Measures Test to Use

Use repeated measures tests to determine the effect a treatment or condition has on the same individuals by observing the individuals before and after the treatments or conditions.

By concentrating on the changes produced by the treatment instead of the values observed before and after the treatment, repeated measures tests eliminate the differences due to individual reactions, which gives a more sensitive (or more powerful) test for finding an effect.

The **Advisor Wizard** prompts you to answer questions about your data and goals, then selects the appropriate test. However, if you are already familiar with the comparison requirements, you can go directly to the appropriate test. The criteria used to select the appropriate procedure include:

- The number of treatments to compare. Are you comparing the effect before and after a single treatment, or after two or more different treatments?
- The distribution of the treatment effects. Are the individual effects distributed along a normal "bell" (Gaussian) curve, or not? Comparisons of treatments effects with normal distributions use parametric tests, which are based on the mean and standard deviation parameters of a normally distributed population. If the effect distributions are not normal, a nonparametric, or distribution-free test must be used, which ranks the values along a new ordinal scale before performing the test.

 **Tip:** SigmaPlot can automatically test for assumptions of normality and variance.

When to Compare Effects on Individuals Before and After a Single Treatment

If data was collected from the same group of individuals (for example, patients before and after a surgical treatment, or rats before and after training), use Before and After comparison to test for a significant difference beyond what can be attributed to random individual variation.

When to use a Paired t-test versus a Wilcoxon Signed Rank Test

You can use two different tests to compare observations before and after an intervention in the same individuals: the Paired t-test and the Wilcoxon Signed Rank Test.

- Choose the Paired t-test if your samples were taken from a population in which the changes to each subject are normally distributed. Choose the **Paired t-test** if your samples were taken from a population in which the changes to each subject are normally distributed. The Paired t-test is a parametric test which directly compares the sample data.
- If your sample effects are not normally distributed, choose the Wilcoxon Signed Rank Test. If your sample effects are not normally distributed, choose the **Wilcoxon Signed Rank Test**. The Wilcoxon Signed Rank Test arranges the data into sets of rankings, then performs a Paired t-test on the sum of these ranks, rather than directly on the data.
- If your samples are already ordered according to qualitative ranks, such as poor, fair, good, and very good, use the Wilcoxon Signed Rank Test.

The advantage of the paired t-test is that, assuming normality and equal variance, it is slightly more sensitive (for example, has greater power) than the Wilcoxon Signed Rank Test. When these assumptions are not met, the Wilcoxon Signed Rank Test is more reliable.

- i** **Tip:** You can tell SigmaPlot to analyze your data and test for normality. If the assumption of normality is violated, the alternative parametric or nonparametric test is suggested. Assumption tests are activated and configured in the Paired t-Test and Wilcoxon Options dialog boxes.

SigmaPlot tests for normality using either the Shapiro-Wilk or Kolmogorov-Smirnov test.

When to Compare Effects on Individuals After Multiple Treatments

If you collected data on the same individuals undergoing three or more different treatments or conditions, use one of the Repeated Measures ANOVA (analysis of variance) procedures to test if there is difference among the effects of the treatments beyond what can be attributed to random individual variation.

There are three procedures available: the single factor or One Way Repeated Measures ANOVA (analysis of variance), the Two Way Repeated Measures ANOVA, and the Friedman Repeated Measures ANOVA on Ranks.

- Choose a **One Way** or **Two Way** Repeated ANOVA if the treatment effects are normally distributed with equal variances. The one and two way ANOVAs are parametric tests which **directly compare the two samples arithmetically**.
- If the treatment effects are not normally distributed and/or have unequal variances, choose the Friedman **Repeated Measures ANOVA on Ranks**, which is the nonparametric analog of the One Way ANOVA. The Friedman Repeated Measures ANOVA on Ranks arranges the data into sets of rankings, then performs an analysis of variance based on these ranks, rather than directly on the data, so it does not require assuming normality and equal variances.

The advantage of parametric Repeated Measures ANOVAs are that, when the normality and equal variance assumptions are met, they are slightly more sensitive (for example, have greater power) than the analysis based on ranks. When the assumptions are not met, the Repeated Measures Friedman ANOVA on ranks is more reliable.

- 🚫 Restriction:** SigmaPlot does not have a two factor analysis of variance based on ranks.

Note that you can tell SigmaPlot to analyze your data and test for normal distribution and equal variance. If assumptions of normality and equal variance are violated, the alternative parametric or nonparametric test is suggested. These tests are specified in the repeated measures one and two way and Friedman options dialog boxes.

SigmaPlot tests for normality using either the Shapiro-Wilk or Kolmogorov-Smirnov test, and for equal variance using the Levene Median test.

When to Use One and Two Way Repeated Measures ANOVA

The difference between a one factor and two factor repeated measures ANOVA lies in the design of the experiment that produced the data.

- Use a **One Way RM ANOVA** if the individuals received a set of related but different treatments (for example, one factor). This design is essentially the same as a paired t-test (a one way repeated measures ANOVA of two groups obtains exactly the same P value as a paired t-test).
- Use a **Two Way RM ANOVA** if there were two experimental factors that are varied for the individuals. One or both of the factors can be repeated on the individuals.

An example of when to use One Way Repeated Measures ANOVA would be when comparing the reading skills of the same students after grade school, high school, and college. The repeated factor is education.

An example of when to use Two Way Repeated Measures ANOVA would be when comparing reading skills at different education levels, but the students attended different schools. This example has repeated measures on education level only, with school as the unrepeatable second factor. If you changed the schools so that all students attended all schools as well, then the school factor is also repeated.

The two factor design can test three hypotheses about the education levels and schools: (1) there is no difference in reading skill at different education levels; (2) there is no difference in reading skill at different schools or after changing schools; and (3) there is no interaction between education level and school in terms of reading skill; any effect of levels of education are the same in all schools.

-  **Tip:** SigmaPlot automatically determines if one or both factors have repeated observations in a two way repeated measures ANOVA.

How to Determine Which Treatments Have an Effect

Repeated measures analysis of variance techniques (both parametric and nonparametric) test the hypothesis of no effect among treatments, but do not indicate which treatments have an effect. You can use the multiple comparison procedures provided by SigmaPlot to isolate the differences in effect.

To always test for differences among the groups, select **Always Perform** on the **Post Hoc Tests** tab in the ANOVA options dialog boxes. You can also specify to use multiple comparisons to test for a difference only when the ANOVA P value is significant by selecting **Only When ANOVA P Value is Significant**, then select the **Multiple Comparisons** tab on the **Post Hoc Tests** tab on the **Options** for ANOVA Options dialog box.

Select the specific multiple comparisons procedures to use for each ANOVA under **Multiple Comparisons** on the **Post Hoc Tests** tab on the **Options** for ANOVA Options dialog box.

To open:

1. Select the appropriate test from the test drop-down list in the **SigmaStat** group on the **Analysis** tab.
2. Click **Options**.

Choosing the Rate and Proportion Comparison to Use

Frequency, rate, and proportion tests compare percent \square re

- Use (χ^2) analysis of contingency tables **Chi-Square** to compare the numbers of individuals of two or more groups that fall into two or more different categories.
- Use the **Fisher Exact Test** if you have two groups falling into two categories (a 2 x 2 contingency table) with a small number of expected observations in any category.
- Use **McNemar's Test** to compare the number of individuals that fall into different categories before and after a single treatment or change in condition.
- Use the **Relative Risk Test** to measure the strength of association between a treatment or risk factor and a specified event that occurs in members of a population in a prospective study.
- Use the **Odds Ratio Test** to measure the strength of association between a treatment or risk factor and a specified event that occurs in members of a population in a retrospective study.

i **Tip:** SigmaPlot automatically analyzes your data for its suitability for Chi-Square or Fisher Exact Test, and suggests the appropriate test.

Choosing the Prediction or Correlation Method

When you want to predict the value of one variable from one or more other variables, you can use regression methods to estimate the predictive equation, and compute a correlation coefficient to describe how strongly the value of one variable is associated with another.

When to Use Regression to Predict a Variable

Regression methods are used to predict the value of one variable (the *dependent* variable) from one or more *independent* variables by estimating the coefficients in a mathematical model. Regression assumes that the value of the dependent variable is always determined by the value of independent variables. Regression is also known as fitting a line or curve to the data.

Regression is a parametric statistical method that assumes that the *residuals* (differences between the predicted and observed values of the dependent variables) are normally distributed with constant variance.

The type of regression procedure to use depends on the number of independent variables and the shape of the relationship between the dependent and independent variables. You can perform regression using Simple Linear Regression, Multiple Linear Regression, Multiple Logistic Regression, Polynomial Regression, and Nonlinear Regression.

- Use a **Simple Linear Regression** procedure if there is a single independent variable, and the dependent variable changes in proportion to changes in the independent variable (for example, linearly).
- Use **Multiple Linear Regression** when there are several independent variables, and the dependent variable changes in proportion to changes in each independent variable (for example, linearly).
- Use **Multiple Logistic Regression** when you want to predict a qualitative dependent variable, such as the presence or absence of a disease, from observations of one or more independent variables, by fitting a logistic function to the data.
- Use **Polynomial Regression** for curved relationships that include powers of the independent variable in the regression equation.
- Use **Nonlinear Regression** to fit any general equation to the observations.

You can determine whether or not a possible independent variable contributes to a multiple linear regression model using **Forward and Backward Stepwise Regression** or **Best Subset Regression**. Use these procedures if you are unsure of the contribution of a variable to the value of the independent variable in a **Multiple Linear Regression**.

- Use **Backwards Stepwise Regression** to begin with all selected independent variables, and delete the variables that least contribute to predicting the dependent variable, until only variables with real predictive value remain in the model.
- Use **Forward Stepwise Regression** to start with zero independent variables, then add variables that contribute to the prediction of the dependent variable, until (ideally) all variables that contribute have been added to the model.
- Use **Best Subset Regression** to evaluate all possible models of the regression equation, and identify those with the best predictive ability (according to a specified criterion).

- Tip:** You can use these procedures to find **Multiple Linear Regression** models. Choose **Polynomial** or **Nonlinear Regression** for curved data sets.

When to Use Correlation

Compute the correlation coefficient if you want to quantify the relationship between two variables without specifying which variable is the dependent variable and which is the independent variable. Correlation does not predict the value of one variable from another; it only quantifies the strength of association between the value of one variable with another.

You can compute two kinds of correlation coefficients: the **Pearson Product Moment Correlation** coefficient, and the **Spearman Rank Order Correlation** coefficient.

- Choose Pearson **Product Moment Correlation** if the residuals are normally distributed and the variances are constant. The Pearson Product Moment Correlation is a parametric test which assumes that data were drawn from a normal population.
- If the residuals are not normally distributed and/or have non-constant variances, choose Spearman **Rank Order Correlation**. The Spearman **Rank Order Correlation** is a nonparametric test that constructs a measure of association based on ranks rather than on arithmetic values.
- If your samples are already ordered according to qualitative ranks, such as poor, fair, good, and very good, choose Spearman rank order correlation.

The advantage of the Pearson **Product Moment Correlation** is that, assuming normality and constant variance, it is slightly more sensitive (for example, has greater power) than the Spearman **Rank Order Correlation**.

Choosing the Survival Analysis to Use

Use survival analysis to generate the probability of the time to an event and the associated statistics such as the median survival time.

- Use **Single Group** to determine the survival time statistics and graph for a single data set (group). This may also be used to generate a single survival curve graph and statistics for all data sets combined in a multi-group data set provided that the data is in Indexed format. To do this, select the survival time a status columns and ignoring the group column.
- Use **LogRank** to determine the survival time statistics and graph for multi-group data sets. The **LogRank** statistic and one of two multiple comparison procedures will be used to determine which groups are significantly different. The **LogRank** statistic assumes that all survival times are equally accurate.
- Use **Gehan-Breslow** for exactly the same situation as the **LogRank** case except that the later survival times are assumed to be less accurate and are given less weight. Many censored values with large survival times provide an example of this situation.
- Use **Cox Regression** to study the impact of potential risk factors on survival time. For a single group, use the Proportional Hazards model. For multiple groups, use the Stratified model.

Testing Normality

A normal population follows a standard, "bell" shaped Gaussian distribution. Parametric tests assume normality of the underlying population or residuals of the dependent variable, and can become unreliable if this assumption is violated. SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test (with Lilliefors' correction) to test data for normality of the estimated underlying population.

When to Test for Normality

Normality is assumed for all parametric tests and regression procedures. SigmaPlot can automatically perform a normality test when running a statistical procedure that makes assumptions about the population parameters. This assumption testing is enabled in the Options dialog for each test. If the data fails the assumptions required for a particular test, SigmaPlot will suggest the appropriate test that can be used instead.

If you want to perform a parametric test and your data fails the normality test, you can transform your data using **Transforms** commands so that it meets the normality requirements. To make sure transformed data now follows a normal distribution pattern, you can run a normality test on the data before performing the parametric procedure again.

Performing a Normality Test

To run a normality test:

1. Enter, transform, or import the data to be tested for normality into data worksheet columns.
2. If desired, set the P value used to pass or fail the data on the **Report** tab of the **Options** dialog box.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, select **Normality** from the **Tests** drop-down list.
The **Normality - Select Data** dialog box appears.
5. Select the worksheet columns with the data you want to test.
6. View and interpret the Normality test report, and generate the [report graphs](#).

Setting the P Value for the Normality Test

The Shapiro-Wilk and Kolmogorov-Smirnov tests use a P value to determine whether the data passes or fails. Set this P value on the **Report** tab of the **Options** dialog box.

To set the P value for the Normality test:

1. Click the **Sigma Button**, and then click **Options**.

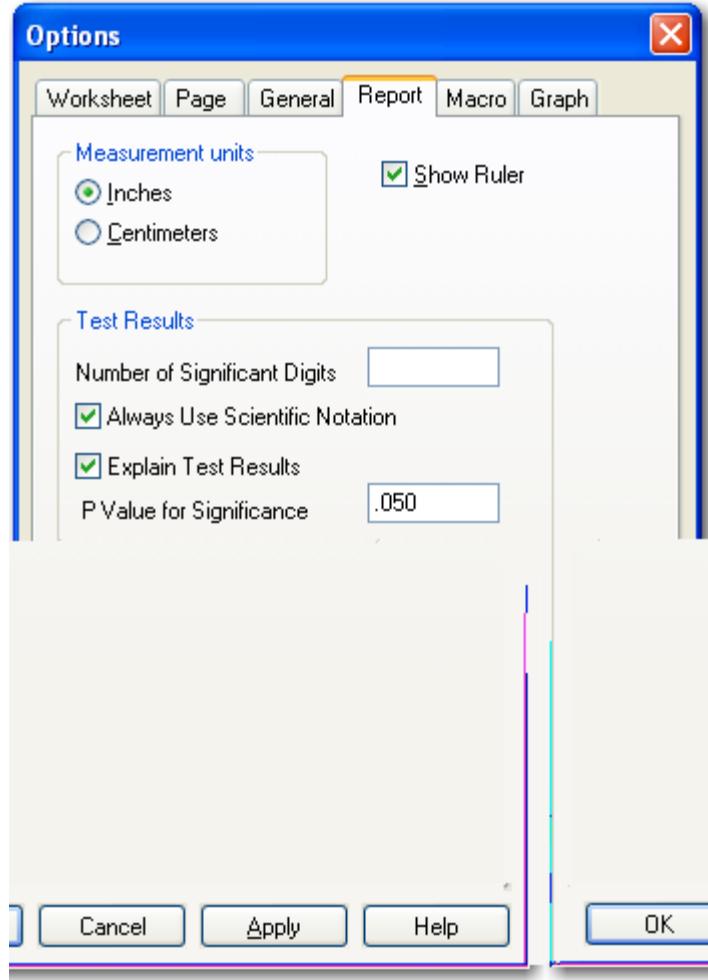
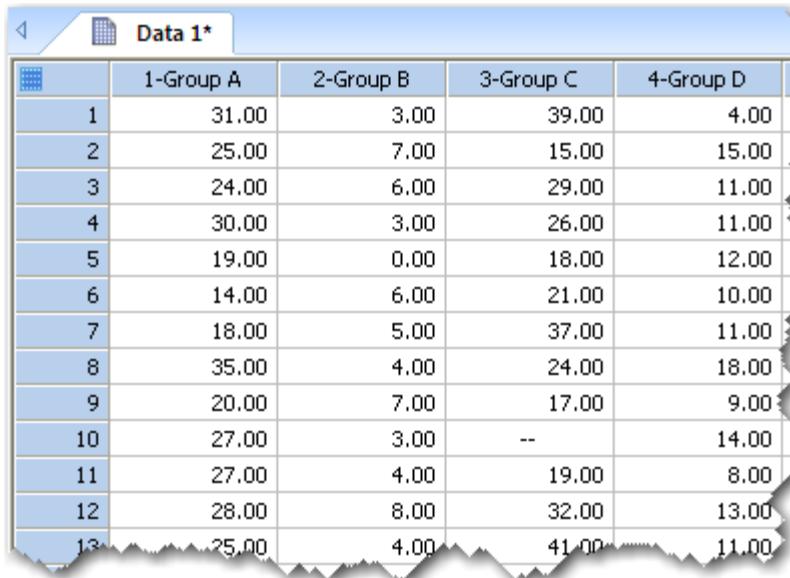


Figure 13: The Reports tab of the Options dialog box

2. On the **Options** dialog box, click the **Report** tab.
3. Click **OK** when finished. For more information, see [Report Graphs](#) on page 373.

Arranging Normality Test Data

Normality test data must be in raw data format, with the individual observations for each group, treatment or level in separate columns. You can test up to 640 columns of data for normality.



	1-Group A	2-Group B	3-Group C	4-Group D	
1	31.00	3.00	39.00	4.00	
2	25.00	7.00	15.00	15.00	
3	24.00	6.00	29.00	11.00	
4	30.00	3.00	26.00	11.00	
5	19.00	0.00	18.00	12.00	
6	14.00	6.00	21.00	10.00	
7	18.00	5.00	37.00	11.00	
8	35.00	4.00	24.00	18.00	
9	20.00	7.00	17.00	9.00	
10	27.00	3.00	--	14.00	
11	27.00	4.00	19.00	8.00	
12	28.00	8.00	32.00	13.00	
13	25.00	4.00	41.00	11.00	

Figure 14: Valid Data Format for Normality Testing

Running a Normality Test

To run a Normality test, you need to select the data to test. Use the **Select Data** panel of the Test Wizard is used to select the worksheet columns with the data you want to test.

To run a Normality test:

1. If you want to select your data before you run the test, drag the pointer over your data.
2. Click the **Analysis** tab, and then in the **SigmaStat** group, select **Normality** from the **Tests** drop-down list. The **Normality - Select Data** dialog box appears. If you selected columns before you chose the test, the selected columns automatically appear in the **Selected Columns** list.
3. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list. The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The number or title of selected columns appear in each row. You can select up to 64 columns of data for the **Normality** test.
4. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
5. Click **Finish** to describe the data in the selected columns. After the computations are completed, the **report** appears.

Interpreting Normality Test Results

Depending on the normality testing method you have selected in Options, the results of a Normality test display either the Shapiro-Wilk W statistics or the Kolmogorov-Smirnov K-S distances and P values computed for each column, and whether or not each column selected passed or failed the test.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this explanatory text in **Reports** tab of the **Options** dialog box.

- i** **Tip:** The number of decimal places displayed is also controlled in **Reports** tab of the **Options** dialog box.

K-S Distance

The Kolmogorov-Smirnov distance is the maximum distance between the sample cumulative distribution function of your data and the theoretical Gaussian distribution function based on the mean and variance estimates from your data.

P Values

The P values represent the observations for normality using either the Shapiro-Wilk or Kolmogorov-Smirnov test. If the P computed by the test is greater than the P set in the appropriate Report Options dialog, your data can be considered normal.

Normality Report Graphs

You can generate two graphs using the results from a Normality report. They include a:

- **Histogram of the residuals.** The Normality histogram plots the raw residuals in a specified range, using a defined interval set.
- **Normal probability plot of the residuals.** The Normality probability plot graphs the frequency of the raw residuals.

Creating a Normality Report Graph

To generate a graph of Normality report data:

1. Click the **Report** tab, and then in the **Result Graphs** group, click **Create Result Graph**. The **Create Graph** dialog box appears displaying the types of graphs available for the Normality report.
2. Select the type of graph you want to create from the **Graph Type** list, then click **OK**. The specified graph appears in a graph window or in the [report](#).

Determining Experimental Power and Sample Size

The power, or sensitivity of a statistical hypothesis test depends on the alpha (a) level, or risk of a false positive conclusion, the size of the effect or difference you wish to detect, the underlying population variability, and the sample size.

The sample size for an intended experiment is determined by the power, alpha (a), the size of the difference, and the population variability. For more information, see [Computing Power and Sample Size](#) on page 349.

When to Compute Power and Sample Size

Use *power and sample size* computations to determine the parameters for an intended experiment, before the experiment is carried out. Use these procedures to help improve the ability of your experiments to test the desired hypotheses.

You can determine power or sample size for:

- Paired and Unpaired t-Tests.
- One Way ANOVA.
- z-Test comparison of proportions.
- Chi-Square analysis of contingency tables.
- Correlation Coefficients.

How to Determine the Power of an Intended Test

1. Click the **Analysis** tab, and then in the **SigmaStat** group, click **Power** and then select a test.
2. When the dialog box **Power** appears, specify the remaining parameters of the data.

How To Estimate the Sample Size Necessary to Achieve a Desired Power

1. Click the **Analysis** tab, and then in the **SigmaStat** group, select **Sample Size** from the **Tests** drop-down list.
2. When the dialog box **Sample Size** appears, specify the power and the remaining parameters of the data.

Chapter

4

Single Group Analysis

Topics:

- One-Sample t-Test
- One-Sample Signed Rank Test

One-Sample t-Test

Use the One-Sample t-Test when you want to test the hypothesis that the mean of a sampled normally-distributed population equals a specified value.

About the One-Sample t-Test

Use the One-Sample t-Test when you want to determine if the group that you have measured came from a different population than the one specified. Your null hypothesis is that the group has been sampled from a population with the specified mean.

If you calculate a sample mean from that group and it is different from a specified value then this can occur for two reasons:

- Your sample mean came from a different population than the one specified.
- Your sample mean came from the same population but by chance the measured and specified values are different.

In the first case you reject the null hypothesis. In the second you fail to reject it.

You use the One-Sample t-test to perform this analysis. It will determine the mean of your group and generate the probability that this value occurred by chance (the P value).

Performing a One-Sample t-Test

To perform an a One-Sample t-test:

1. [Enter or arrange your data appropriately in the worksheet.](#)
2. If desired, set the t-test options.
3. Click the **Analysis** tab, and then in the **SigmaStat** group, from the **Tests** drop-down list, select:
Single Group > One-Sample t-test
4. Run the test.
5. Generate report graphs. [One-Sample t-Test Report Graphs](#) on page 47

Arranging One-Sample t-Test Data

The format of the data to be tested can be:

- **Raw.** The raw data format uses separate worksheet columns for the data in each group.
- **Mean, size, standard deviation.** This data format places the mean, sample size, and standard deviation in separate worksheet columns.
- **Mean, size, standard error.** This data format places the mean, sample size, and standard error in separate worksheet columns.

Setting One-Sample t-Test Data Options

Use the One-Sample t-Test Options to:

- Adjust the P value to relax or restrict the testing of your data for normality.
- Select the method for testing normality.
- Display the statistics summary and the confidence interval for the data in the report and save residuals to a worksheet column.
- Specify the alpha value for computing the power or sensitivity of the test.

To set One-Sample t-test options:

1. Click the **Analysis** tab.

2. In the **SigmaStat** group, click **Options**. The One-Sample t-test Options dialog box appears with four tabs:
 - **Assumption Checking.** Adjust the P value to relax or restrict the testing of your data for normality. Select which normality test to use – Shapiro-Wilk (recommended) or Kolmogorov-Smirnov.
 - **Results.** Display the statistics summary and the confidence interval for the data in the report and save residuals to a worksheet column.
 - **Post Hoc Tests.** Select the alpha value to compute the power or sensitivity of the test.

-  **Tip:** If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over the column title to select the data column.

Options settings are saved between SigmaPlot sessions.

3. To continue the test, click **Run Test**. The **Select Data** panel of the Test Wizard appears.

Options for One-Sample t-Test: Assumption Checking

The normality assumption test checks for a normally distributed population.

- **Normality.** SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.
- **P Values to Reject.** The P value determines the probability of being incorrect in concluding that the data is not normally distributed (P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.010 requires greater deviations from normality to flag the data as non-normal than a value of 0.050.

-  **Note:** There are extreme conditions of data distribution that these tests cannot take into account; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options for One-Sample t-Test: Results

Summary Table. Displays the number of observations for a column or group, the number of missing values for a column or group, the average value for the column or group, the standard deviation of the column or group, and the standard error of the mean for the column or group.

Confidence Intervals. Displays the confidence interval for the difference of the means. To change the interval, enter any number from 1 to 99 (95 and 99 are the most commonly used intervals).

Residuals in Column. Displays residuals in the report and to save the residuals of the test to the specified worksheet column. Edit the number or select a number from the drop-down list.

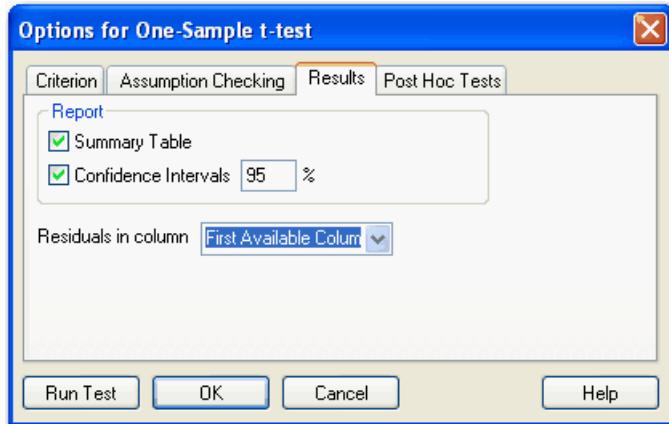


Figure 15: The Options for One-Sample t-Test Dialog Box Displaying the Summary Table, Confidence Intervals, and Residuals Options

Options for One-Sample t-Test: Post Hoc Tests

Power. The power or sensitivity of a test is the probability that the test will detect a difference between the groups if there is really a difference.

Use Alpha Value. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

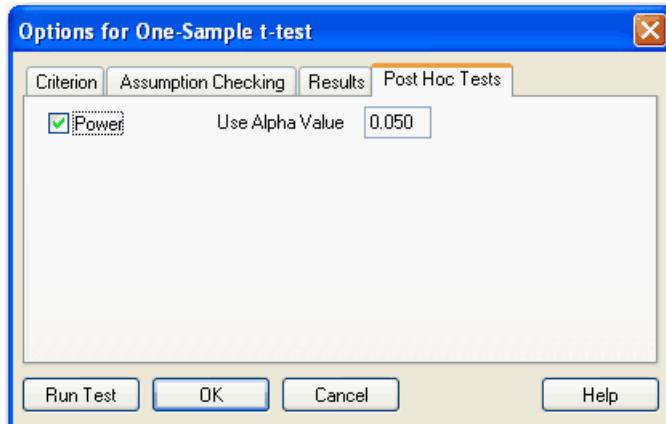


Figure 16: The Options for One-Sample t-Test Dialog Box Displaying the Power Option

Running a One-Sample t-Test

If you want to select your data before you run the test, drag the pointer over your data.

1. Click the **Analysis** tab, and then in the **SigmaStat** group, from the **Tests** drop-down list, select: **Single Group > One-Sample t-test**

The **Data Format** panel of the Test Wizard appears prompting you to specify a data format.

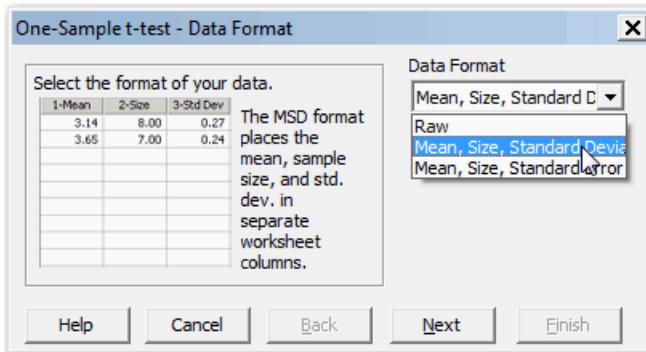


Figure 17: The Data Format Panel of the Test Wizard Prompting You to Specify a Data Format

2. Select the appropriate data format from the **Data Format** drop-down list.
3. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.
4. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw and indexed data, you are prompted to select two worksheet columns. For statistical summary data you are prompted to select three columns.

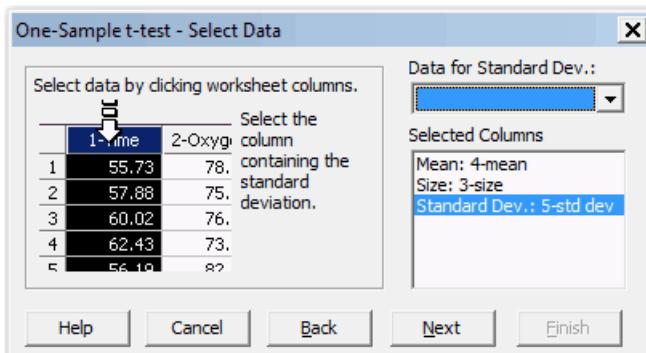


Figure 18: The Select Data for One-Sample t-test Panel Prompting You to Select Data Columns

5. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.

6. Click **Next** to enter the value of the population mean for the test.

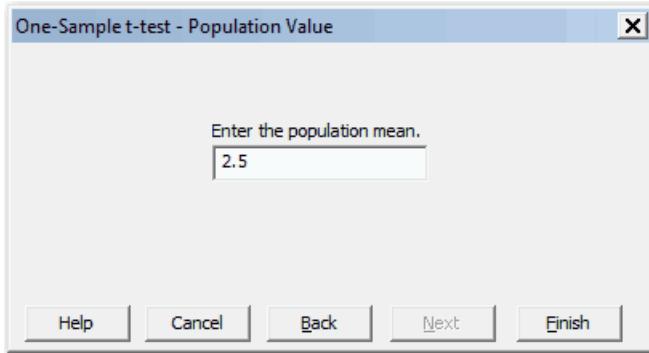


Figure 19: The Population Value for One-Sample t-Test Prompting You to Enter the Population Mean

7. Click **Finish** to run the t-test on the selected columns. After the computations are completed, the report appears.

Interpreting One-Sample t-Test Results

The One-Sample t-test calculates the t statistic, degrees of freedom, and P value of the specified data. These results are displayed in the One-Sample t-Test report which automatically appears after the One-Sample t-Test is performed. The other results displayed in the report are enabled and disabled in the **Options for t-Test** dialog box.

The null hypothesis is that the mean of the sampled population equals the user supplied value. The sample mean of the selected data is compared with the hypothesized population mean supplied by the user by computing:

$$t = \frac{\bar{x} - \mu}{\sigma} \sqrt{n}$$

where

\bar{x} = sample mean

μ = hypothesized population mean

σ = sample standard deviation

n = sample size

By random sampling of the population, assuming the null hypothesis is true, this quantity defines a random variable T , whose distribution is Student's central T-distribution with $n - 1$ degrees of freedom. The (two-sided) P-value for this test is computed as $P(|T| > |t|)$, where P denotes the probability distribution for T . This P-value is then compared to the significance level α that is set by you. If the value is less than α , there is a significant difference between the mean of the sampled population and μ .

The $(1 - \alpha)100\%$ confidence interval for the true population mean is:

$$\bar{x} \pm \sigma t_{\alpha/2, n-1}$$

where x and σ are defined above $t_{\alpha/2, n-1}$, and $t_{\alpha/2, n-1}$ is the value that satisfies $P(|T| > t_{\alpha/2, n-1}) = \alpha$.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can enable or disable this explanatory text in the **Options** dialog box. For more information, see [Report Graphs](#) on page 373.

Normality Test. Normality test results show whether the data passed or failed the test of the assumption that the samples were drawn from a normal population and the P value calculated by the test. All parametric tests require normally distributed source populations.

Summary Table. SigmaPlot can generate a summary table listing the size N for the sample, number of missing values, means, standard deviations, and the standard error of the mean (SEM). This result is displayed unless you disable **Summary Table** in the **Options for t-test** dialog box.

- **N (Size).** The number of non-missing observations for that column or group.
- **Missing.** The number of missing values for that column or group.
- **Mean.** The average value for the column. If the observations are normally distributed the mean is the center of the distribution.
- **Standard Deviation.** A measure of variability. If the observations are normally distributed, about two-thirds will fall within one standard deviation above or below the mean, and about 95% of the observations will fall within two standard deviations above or below the mean.
- **Standard Error of the Mean.** A measure of the approximation with which the mean computed from the sample approximates the true population mean.

One-Sample t-Test Report Graphs

You can generate up to three graphs using the results from a One-Sample t-Test. They include a:

- **Scatter plot with error bars of the column means.** The One-Sample t-test scatter plot graphs the group means as single points with error bars indicating the standard deviation.
- **Histogram of the residuals.** The One-Sample t-test histogram plots the raw residuals in a specified range, using a defined interval set.
- **Normal probability plot of the residuals.** The One-Sample t-test probability plot graphs the frequency of the raw residuals.

Creating a Graph of the One-Sample t-Test Data

1. Select the One-Sample t-Test report.
2. On the **Report** tab, click **Create Result Graph** in the **Result Graphs** group.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the One-Sample t-Test results.

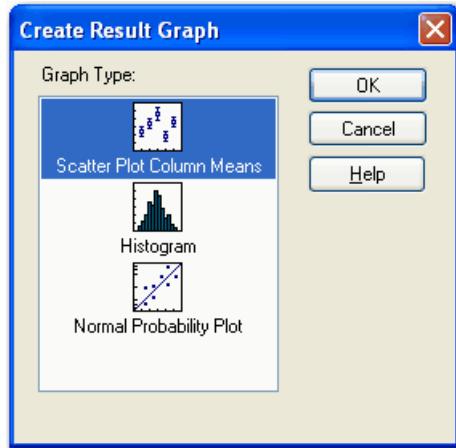
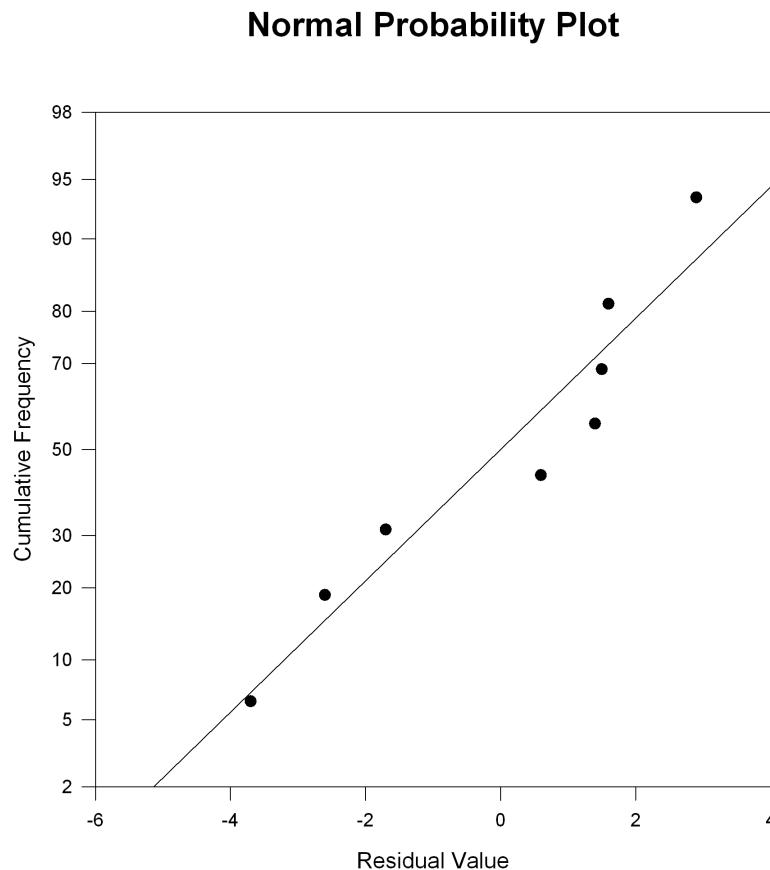


Figure 20: The Create Graph Dialog Box for the One-Sample t-test Report

3. Select the type of graph you want to create from the **Graph Type** list, then click **OK**, or double-click the desired graph in the list.

The selected graph appears in a graph window.



[Report Graphs](#) on page 373

One-Sample Signed Rank Test

The One-Sample Signed Rank Test tests the null hypothesis that the *median* of a population, rather than the mean (as seen in the One-Sample t-Test), is equal to a specified value,

About the One-Sample Signed Rank Test

Use the One-Sample Signed Rank Test when you want to determine if the group that you have observed came from a different population than the one specified. This test is generally more robust and powerful than the One-Sample t-Test for measuring the central tendency of the population when the group is sampled from a population that is not normally distributed. The basic assumption in using this test is that the underlying distribution of the population be symmetric about the median. Your null hypothesis is that the group has been sampled from a population with the specified median.

If you calculate a sample median from that group and it is different from a specified value then this can occur for two reasons:

- Your sample median came from a different population than the one specified.
- Your sample median came from the same population but the difference is due to random sample variability.

In the first case you reject the null hypothesis. In the second you fail to reject it.

Use the One-Sample Signed Rank Test to perform this analysis. It determines the median of your group and generates the probability that this value is significantly different from the specified value (the P value).

Performing a One-Sample Signed Rank Test

To perform a One-Sample Signed Rank Test:

1. Enter or arrange your data appropriately in the worksheet.
2. If desired, set the t-test options.
3. Click the **Analysis** tab, and then in the **SigmaStat** group, from the **Tests** drop-down list, select:
Single Group > One-Sample Signed Rank
4. Run the test.

Arranging One-Sample Signed Rank Test Data

The data for the One-Sample Signed Rank Test consists of a single worksheet column.

Setting One-Sample Signed Rank Test Options

Use the One-Sample Signed Rank Test Options to:

- Adjust the P value to relax or restrict the testing of your data for normality.
- Select the method for testing normality.
- Display the statistics summary for the data in the report.
- Display the confidence interval for the Test Median in the report.

To set One-Sample Signed Rank Test options:

1. Click the **Analysis** tab.
2. In the **SigmaStat** group, click **Options**. The One-Sample Signed Rank Test Options dialog box appears with three tabs:
 - **Criterion.** Specify the population median that you are testing.
 - **Assumption Checking.** Adjust the P value to relax or restrict the testing of your data for normality. Select which normality test to use, Shapiro-Wilk (recommended) or Kolmogorov-Smirnov.
 - **Results.** Display the statistics summary for the data and the confidence interval for the median in the report. Also, determine if the Yates-correction factor should be applied.

i **Tip:** If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over the column title to select the data column.

Options settings are saved between SigmaPlot sessions.

3. To continue the test, click **Run Test**. Select Data panel of the Test Wizard appears.

Options for One-Sample Signed Rank Test: Assumption Checking

The normality assumption test checks for a normally distributed population.

- **Normality.** SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.
- **P Values to Reject.** The P value determines the probability of being incorrect in concluding that the data is not normally distributed (P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.010 requires greater deviations from normality to flag the data as non-normal than a value of 0.050.

- i** **Tip:** There are extreme conditions of data distribution that these tests cannot take into account; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options for One-Sample Signed Rank: Results

Summary Table. Select to place a summary table in the report. This table displays the number of observations for the group, the number of missing values, the computed median value, and percentiles. Text boxes are available to enter two percentile values for the data. By default, the summary table check box is selected and the percentile values are given as 25% and 75%.

Confidence Intervals. Select to display the confidence interval for the population median. The confidence level can be any number from 1 to 99 (95 and 99 are the most commonly used). By default, the check box is selected and the confidence level is set to 95%.

Yates Correction Factor. When the sample size exceeds 20, the normal distribution is used to approximate the P-value for the test. The P-value is smaller than it should be since the actual distribution for the test statistic is discrete whereas the normal distribution is continuous. The Yates continuity correction adjusts the statistic to compensate for this discrepancy.

Click the selected check box to turn the Yates Correction Factor on or off.

For descriptions of the derivation of the Yates correction, you can reference any appropriate statistics reference.

Running a One-Sample Signed Rank Test

If you want to select your data before you run the test, drag the pointer over your data.

1. Click the **Analysis** tab, and then in the **SigmaStat** group, from the **Tests** drop-down list, select:

Single Group > One-Sample Signed Rank

The **One-Sample Signed Rank Test - Select Data** dialog box appears prompting you to select one column of data to test.

2. **To assign the desired worksheet columns to the Selected Columns list**, select the column in the worksheet, or select the column from the **Data for Data** drop-down list.
3. **To change your selection**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
4. Click **Next** to enter the value of the population median for the test.
5. Click **Finish** to run the test. After the computations are completed, [the report appears](#).

Running an Alternate One Sample Signed Rank Test

After clicking **Finish** in the Test Wizard, the normality test specified in **Test Options** tests the normality of the data. If the normality test passes (the data is consistent with a normal distribution), a message box appears with the option to select the one-sample t-test as an alternate test. If you select this option, a report for the one-sample t-test appears instead of the report for the one-sample signed rank test. Likewise, if the normality test fails when running the one-sample t-test, a message box appears giving you the option to switch to the one-sample signed rank test.

Interpreting One-Sample Signed Rank Test Results

The computational results are displayed in the One-Sample Signed Rank Test report which automatically appears after the One-Sample Signed Rank Test is performed.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can enable or disable this explanatory text in the **Options** dialog box.

Normality Test. Normality test results show whether the data passed or failed the test of the assumption that the samples were drawn from a normal population and the *P* value calculated by the test.

Summary Table. A single-line summary table of the basic statistics for the input data. The name of the group column, the number of cases and missing values, the sample median, and the lower and upper percentiles.

Hypothesized Population Median. The value of the **Test Median** that was entered on the **Criterion** tab of the **Options for One Sample Signed Rank Test** dialog box.

Test Statistics and P-Values. Values of the rank sums, their difference W, the Z-statistic (normal approximation), and the estimated and exact significance probabilities (the P-value). The exact P-value is based on the Wilcoxon distribution. The estimated P-value is based on the normal approximation to the Wilcoxon distribution. The Yates correction is used to adjust the estimated P-value. The sample size necessary for the normal approximation to hold varies among sources, with most agreeing that the sample size should be at least 20. Both P-values are reported if the number of values sample size that are different from the hypothesized median is 20 or less. Otherwise, only estimated P-value is reported.

Yates Correction. A statement indicating whether the Yates correction was used.

Confidence Interval. The lower and upper limits of the confidence interval of the population median.

Interpretation of P-Value. An interpretation of the significance probability that differs depending on whether the result is positive (significant) or not. The significance level is set on the **Assumption Checking** tab of the **Options for One Sample Signed Rank Test** dialog box. The default significance level is .05.

Chapter

5

Comparing Two or More Groups

Topics:

- About Group Comparison Tests
- Data Format for Group Comparison Tests
- Unpaired t-Test
- Mann-Whitney Rank Sum Test
- One Way Analysis of Variance (ANOVA)
- Two Way Analysis of Variance (ANOVA)
- Three Way Analysis of Variance (ANOVA)
- Kruskal-Wallis Analysis of Variance on Ranks
- One Way Analysis of Covariance
- Performing a Multiple Comparison

Use group comparison tests to compare random samples from two or more different groups for differences in the mean or median values that cannot be attributed to random sampling variation.

If you are comparing the effects of different treatments on the same individuals, use [repeated measures procedures](#).

About Group Comparison Tests

Group comparisons test two or more different groups for a significant difference in the mean or median values beyond what can be attributed to random sampling variation.

Parametric and Nonparametric Tests

Parametric tests assume samples were drawn from normally distributed populations with the same variances (or standard deviations). Parametric tests are based on estimates of the population means and standard deviations, the parameters of a normal distribution.

Nonparametric tests do not assume that the samples were drawn from a normal population. Instead, they perform a comparison on ranks of the observations. Rank Sum Tests automatically rank numeric data, then compare the ranks rather than the original values.

Comparing Two Groups

You can compare two groups using:

- [An Unpaired t-test \(a parametric test\).](#)
- [A Mann-Whitney Rank Sum Test \(a nonparametric test\).](#)

Comparing Many Groups

You can compare three or more groups using the:

- [One Way ANOVA \(analysis of variance\).](#) A parametric test that compares the effect of a single factor on the mean of two or more groups.
- [Two Way ANOVA.](#) A parametric test that compares the effect of two different factors on the means of two or more groups.
- [Three Way ANOVA.](#) A parametric test that compares the effect of three different factors on the means of two or more groups.
- [Kruskal-Wallis Analysis of Variance on Ranks.](#) This is the nonparametric analog of One Way ANOVA.

If you are using one of these procedures to compare multiple groups, and you find a statistically significant difference, you can use several multiple comparison procedures (also known as post-hoc tests) to determine exactly which groups are different and the size of the difference. These procedures are described for each test.

Data Format for Group Comparison Tests

You can arrange data in the worksheet as:

- Columns for each group (raw data).
- Data indexed to other column(s).

For *t*-Tests and One Way ANOVAs, you can also use:

- The sample size, mean, and standard deviation for each group.
- The sample size, mean, and standard error of the mean (SEM) for each group.

Columns 1 and 2 are arranged as raw data. Columns 3, 4, and 5 are arranged as descriptive statistics using the sample size, mean, and standard deviation. Columns 6 and 7 are arranged as group indexed data, with column 6 as the factor column and column 7 as the data column.

	1-male	2-female	3-mean	4-size	5-std dev	6-gender	7-response
1	3.0000	3.4000	3.1375	8.0000	0.2722	female	3.9000
2	3.5000	3.9000	3.5571	7.0000	0.2371	male	3.5000
3	3.5000	3.8000				female	3.5000
4	2.8000	3.5000				female	3.6000
5	3.1000	3.2000				male	2.8000
6	3.0000	2.5000				female	3.8000
7	3.2000	3.6000				male	2.9000
8	2.8000					male	3.2000
9						female	3.0000
10						male	3.4000
11						female	3.2000
12						female	3.5000
13							

Figure 21: Valid Data Formats for an Unpaired t-test

Descriptive Statistics

If your data is in the form of statistical values (sample size, mean, standard deviation, or standard error of the mean), the sample sizes (N) must be in one worksheet column, the means in another column, and the standard deviations (or standard errors of the mean) in a third column, with the data for each group in the same row. When comparing two groups, there should be exactly two rows of data.

Arranging Data for t-Tests and ANOVAs

There are several formats of data that can be analyzed by t-tests, analysis of variances (ANOVA), repeated measures ANOVAs, and their nonparametric analogs, including:

- **Raw data**, which places the data for each group in separate columns; this is the format used by SigmaPlot.
- **Indexed data**, which places the group names in one column, and the corresponding data for each group in another column.
- **Statistical summary data**, which can be used by unpaired t-tests and One Way ANOVAs.

Set data format in the Select Data panel of the Test Wizard that appears when you run the current test. For more information, see [Statistical summary data](#).

Messy and Unbalanced Data

SigmaPlot automatically handles missing data points (indicated with an "--") for all situations. If a two factor ANOVA is missing entire cells, the appropriate steps are suggested, and the desired procedure is performed.

Raw Data

The raw data format is the most common format, where your data have not yet been analyzed or transformed. It places the data for each group to be compared or analyzed in separate columns. Use column titles to identify the groups, as the titles will also be used in the analysis report.

You can use raw data for all tests except Two and Three Way ANOVAs.

 **Note:** SigmaPlot tests accept messy and unbalanced data and do not require equal sample sizes in the groups being compared. There are no problems associated with missing data or uneven columns. Missing values are represented by empty cells or double dashes ("--"). Text items may be considered missing also provided the test expects only numeric values.

t-Tests and Rank Tests

The groups to be compared are always placed in two columns.

Paired t-tests and signed rank tests (both repeated measures tests) assume that the data for each subject is in the same row.

One Way ANOVA and One Way ANOVA on Ranks

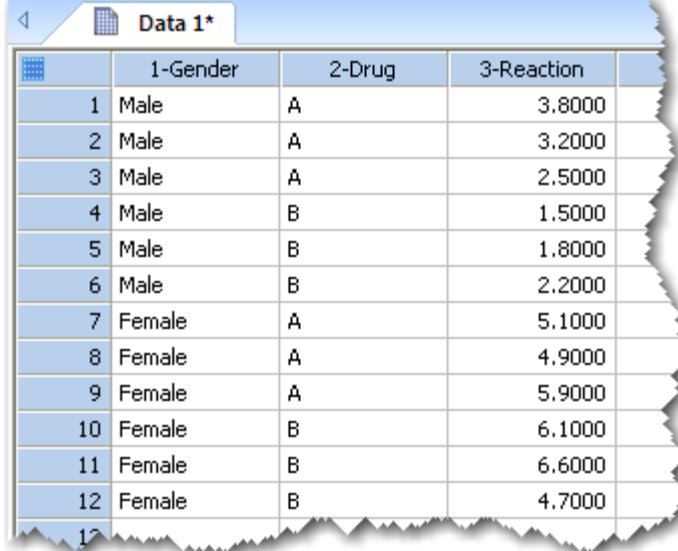
Data for each group is placed in separate columns, with as many columns as there are groups. One way repeated measures ANOVA and one way repeated measures ANOVA on ranks assume that the data for each subject is in the same row.

Indexed Data

Indexed data consists of a factor column, which contains the names of the groups or levels, and a data column containing the data points in corresponding rows. The data does not have to be organized in any particular order.

 **Note:** Data for a Two Way ANOVA is always assumed to be indexed.

Column 1 is the first factor column, column 2 is the second factor column, and column 3 contains the data.



	1-Gender	2-Drug	3-Reaction
1	Male	A	3.8000
2	Male	A	3.2000
3	Male	A	2.5000
4	Male	B	1.5000
5	Male	B	1.8000
6	Male	B	2.2000
7	Female	A	5.1000
8	Female	A	4.9000
9	Female	A	5.9000
10	Female	B	6.1000
11	Female	B	6.6000
12	Female	B	4.7000

Figure 22: Data Format for a Two Way ANOVA with Two Factor Indexed Data

Two way ANOVAs require two factor columns and one data column. Three Way ANOVAs require three factor columns and one data column, and Repeated measures ANOVAs require an additional subject column to identify the subject of the measurement.

The order of the rows containing the index and data does not matter, for example, they do not have to be grouped or sorted by factor level or subject.

 **Note:** If you are analyzing entire columns of data, the location in the worksheet of the factor, subject, and data columns does not matter. For more information, see [Indexed Data](#) on page 56.

Independent t-test and Mann-Whitney rank sum test. The group index is in a factor column, and the corresponding data points to be compared are in a second column.

Paired t-test and Wilcoxon signed rank test. Repeated measures comparisons require an additional subject index column, which indicates the subject for each level and data point.

One way ANOVA and Kruskall-Wallis ANOVA on ranks. The factor column contains the group index, and the data column contains the corresponding data points. Indexed data for one way ANOVA contains only two columns.

Two way ANOVA. Two factor columns are required for Two Way ANOVAs, one for each level of the observation. Each data point should be represented by different combinations of the factors. For example, the factors in a drug treatment test are Gender and Drug, and the levels are Male/Female and Drug A/Drug B.

Three way ANOVA. Three factors are required for Three Way ANOVAs, one for each level of observation. Each data point should be represented by different combinations of the factors. For more information, see [Arranging Three Way ANOVA Data](#) on page 105.

Repeated measures ANOVA. These tests require an additional subject column, which identifies the data points for each subject.

A Two Way Repeated Measures ANOVA requires both a subject column and two factor columns, as well as a data column.

Statistical Summary Data

Unpaired t-tests and one way ANOVAs can be performed on summary statistics of the data. These statistics can be in the form of:

- The sample size, mean, and standard deviation for each group, or
- The sample size, mean, and standard error of the mean (SEM) for each group

The sample sizes (N) must be in one worksheet column, the means in another column, and the standard deviations (or standard errors of the mean) in a third column, with the data for each group in the same row.

If you plan to compare only a portion of the data by selecting a block, put the sample sizes in the left column, the means in the middle column, and the standard deviations or SEMs in the right column.

Unpaired t-Test

Use an Unpaired t-test when:

- You want to see if the means of two different samples are significantly different.
- Your samples are drawn from normally distributed populations with the same variances.

If you know that your data was drawn from a non-normal population, use the [Mann-Whitney Rank Sum Test](#). When there are more than two groups to compare, do a [One Way Analysis of Variance](#).

- i** **Tip:** Depending on your t-test options settings, if you attempt to perform a t-test on non-normal populations or populations with unequal variances, SigmaPlot will inform you that the data is unsuitable for a t-test, and suggest the Mann-Whitney Rank Sum Test instead.

About the Unpaired t-Test

The Unpaired t-test is a parametric test based on estimates of the mean and standard deviation parameters of the normally distributed populations from which the samples were drawn. It tests for a difference between two groups that is greater than what can be attributed to random sampling variation. The null hypothesis of an unpaired t-test is that the means of the populations that you drew the samples from are the same. If you can confidently reject this hypothesis, you can conclude that the means are different.

Performing an Unpaired t-Test

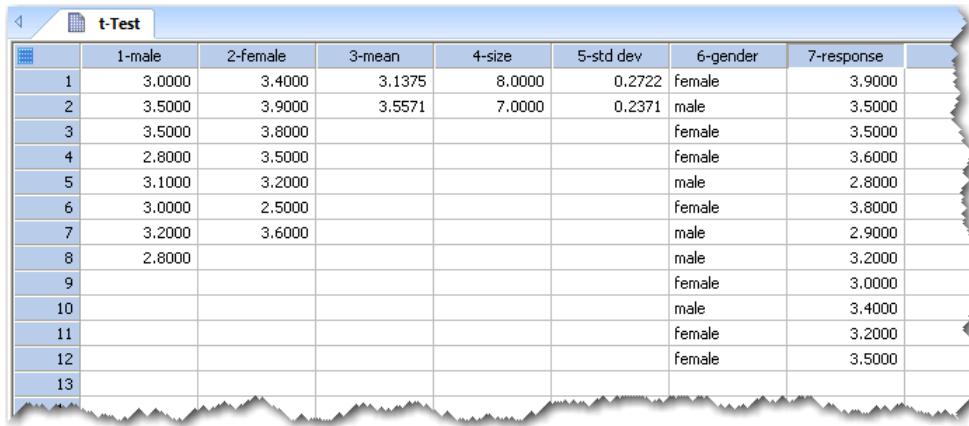
To perform an Unpaired t-test:

1. [Enter or arrange your data appropriately in the worksheet.](#)
2. If desired, set the t-test options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Compare Two Groups > t-test
5. Run the test.
6. [Generate report graphs.](#)

Arranging t-Test Data

The format of the data to be tested can be raw, indexed, or summary statistics. For raw and indexed data, the data is placed in two worksheet columns. Statistical summary data is placed in three worksheet columns.

Columns 1 and 2 are arranged as raw data. Columns 3, 4, and 5 are arranged as descriptive statistics using the sample size, mean, and standard deviation. Columns 6 and 7 are arranged as group indexed data, with column 6 as the factor column and column 7 as the data column.



	1-male	2-female	3-mean	4-size	5-std dev	6-gender	7-response
1	3.0000	3.4000	3.1375	8.0000	0.2722	female	3.9000
2	3.5000	3.9000	3.5571	7.0000	0.2371	male	3.5000
3	3.5000	3.8000				female	3.5000
4	2.8000	3.5000				female	3.6000
5	3.1000	3.2000				male	2.8000
6	3.0000	2.5000				female	3.8000
7	3.2000	3.6000				male	2.9000
8	2.8000					male	3.2000
9						female	3.0000
10						male	3.4000
11						female	3.2000
12						female	3.5000
13							

Figure 23: Valid Data Formats for an Unpaired t-test

Setting t-Test Options

Use the t-test options to:

- Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
- Display the statistics summary and the confidence interval for the data in the report and save residuals to a worksheet column.
- Compute the power or sensitivity of the test

To set t-test options:

1. Select **t-test** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.
2. Click **Current Test Options**.

The **Options for t-test** dialog box appears with three tabs:

- **Assumption Checking.** Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
- **Results.** Display the statistics summary and the confidence interval for the data in the report and save residuals to a worksheet column.
- **Post Hoc Tests.** Compute the power or sensitivity of the test.

Options settings are saved between SigmaPlot sessions.

i **Tip:** If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.

3. To continue the test, click **Run Test**. The **Select Data** panel of the Test Wizard appears.
4. To accept the current settings and close the options dialog box, click **OK**.

Options for t-Test: Assumption Checking

The normality assumption test checks for a normally distributed population. The equal variance assumption test checks the variability about the group means.

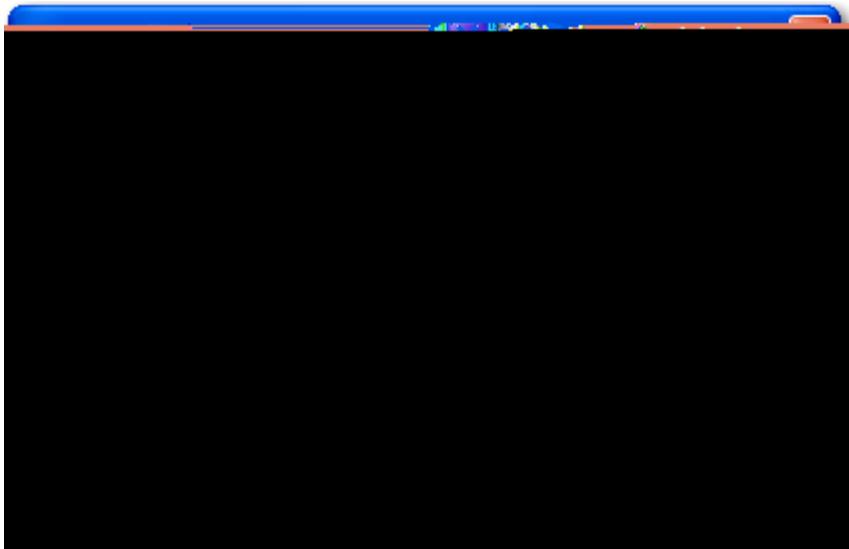


Figure 24: The Options for t-test Dialog Box Displaying the Assumption Checking Options

- **Normality Testing.** SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.
- **Equal Variance Testing.** SigmaPlot tests for equal variance by checking the variability about the group means.
- **P Values for Normality and Equal Variance.** The P value determines the probability of being incorrect in concluding that the data is not normally distributed (P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and equal variance, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality and equal variance, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.010 requires greater deviations from normality to flag the data as non-normal than a value of 0.050.

🚫 **Restriction:** There are extreme conditions of data distribution that these tests cannot take into account. For example, the Levene Median test fails to detect differences in variance of several orders of magnitude; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options for t-Test: Results

Summary Table. Displays the number of observations for a column or group, the number of missing values for a column or group, the average value for the column or group, the standard deviation of the column or group, and the standard error of the mean for the column or group.

Confidence Intervals. Displays the confidence interval for the difference of the means. To change the interval, enter any number from 1 to 99 (95 and 99 are the most commonly used intervals).

Test for Equal Means.

- **Student's t-test (Equal variances assumed).** The Student's t-test tests the null hypothesis that the population *means* of the two groups are equal. SigmaPlot computes the P-values for the test for both one-tailed and two-tailed alternative hypotheses. The standard assumptions of Student's t-test are that the data values within and between the two groups are independent, the populations from which the two groups are sampled are normally distributed, and the population variances of the two groups are equal. This is also known as *homoscedasticity*.
- **Welch's t-test (Equal variances not assumed).** The Welch's t-test is an adaption of the Student's t-test. It is regarded as the more reliable of the two when the two samples have unequal variances and unequal sample sizes.

P-values.

- **Two-tailed.** Select to test for a significant difference between the means that is either positive or negative.
- **One-tailed.** Select to test for a significant difference in only one direction.
- **Both.** Select Both to test for both two-tailed and one-tailed significant differences.

Residuals in Column. Displays residuals in the report and to save the residuals of the test to the specified worksheet column. Edit the number or select a number from the drop-down list.

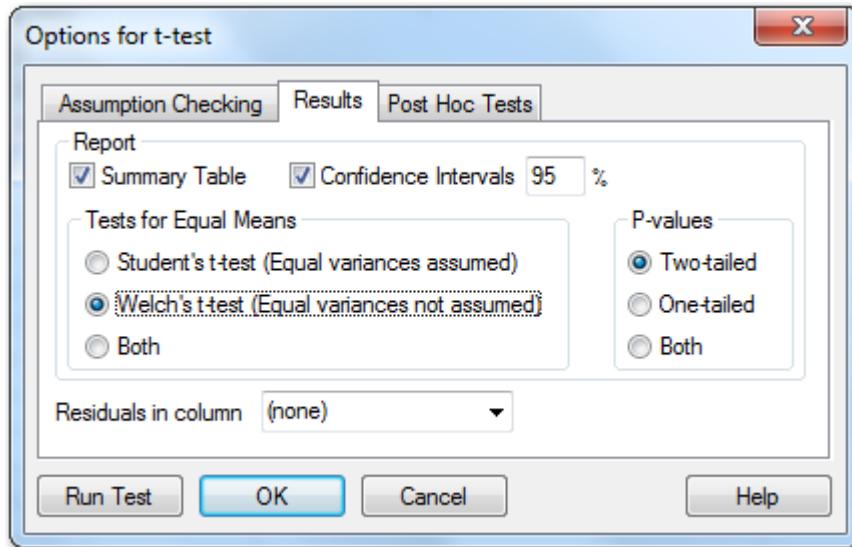


Figure 25: The Options for t-test Dialog Box Displaying the Summary Table, Confidence Intervals, and Residuals Options

Options for t-Test: Post Hoc Tests

Power. The power or sensitivity of a test is the probability that the test will detect a difference between the groups if there is really a difference.

Use Alpha Value. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

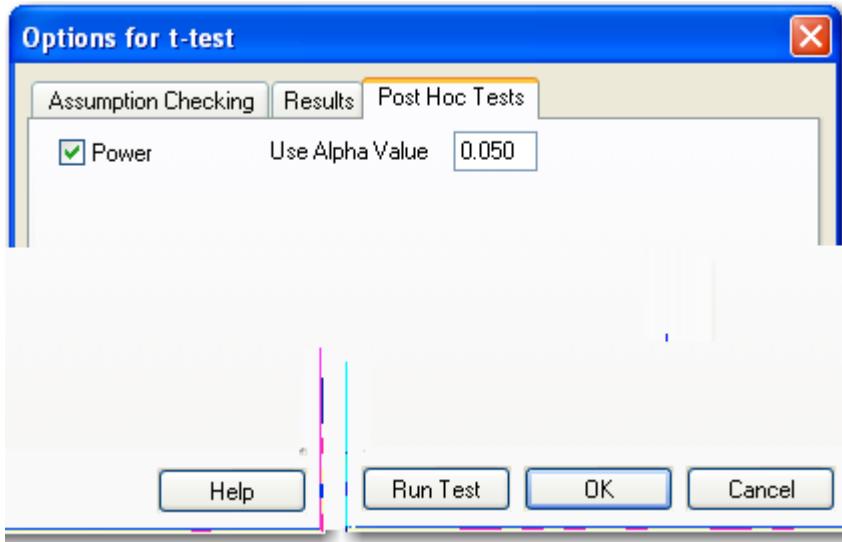


Figure 26: The Options for t-test Dialog Box Displaying the Power Option

Running a t-Test

If you want to select your data before you run the test, drag the pointer over your data.

1. Click the **Analysis** tab.
2. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Compare Two Groups > t-test The **Data Format** panel of the Test Wizard appears prompting you to specify a data format.

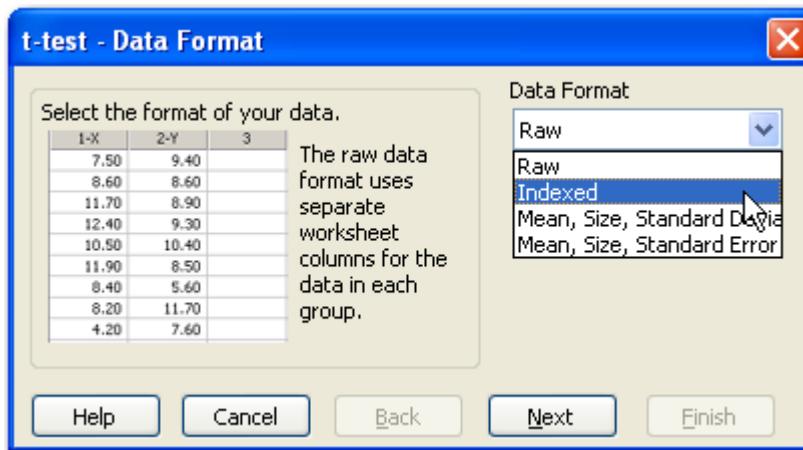


Figure 27: The Data Format for t-test Panel Prompting You to Specify a Data Format

3. Select the appropriate data format (**Raw** or **Indexed**) from the **Data Format** drop-down list.
4. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.
5. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw and

indexed data, you are prompted to select two worksheet columns. For statistical summary data you are prompted to select three columns.

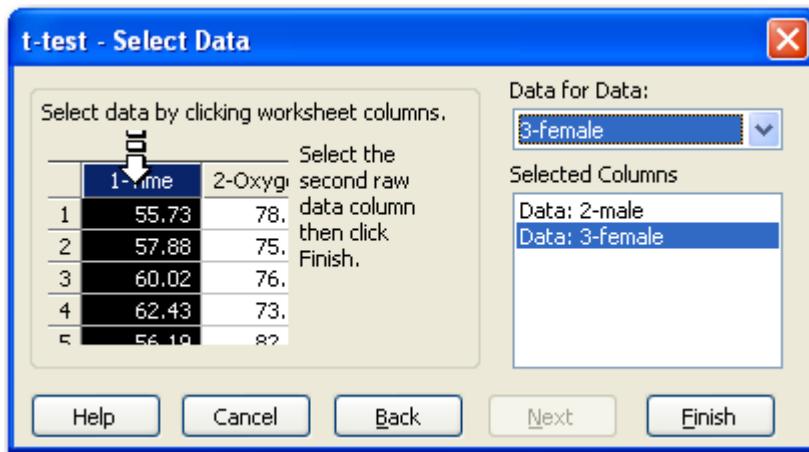


Figure 28: The Select Data Panel of the Test Wizard Prompting You to Select Data Columns

6. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
7. Click **Finish** to run the t-test on the selected columns. After the computations are completed, the report appears.

Interpreting t-Test Results

The t-test calculates the t statistic, degrees of freedom, and P value of the specified data. These results are displayed in the t-test report which automatically appears after the t-test is performed. The other results displayed in the report are enabled and disabled in the Options for t-test dialog box.

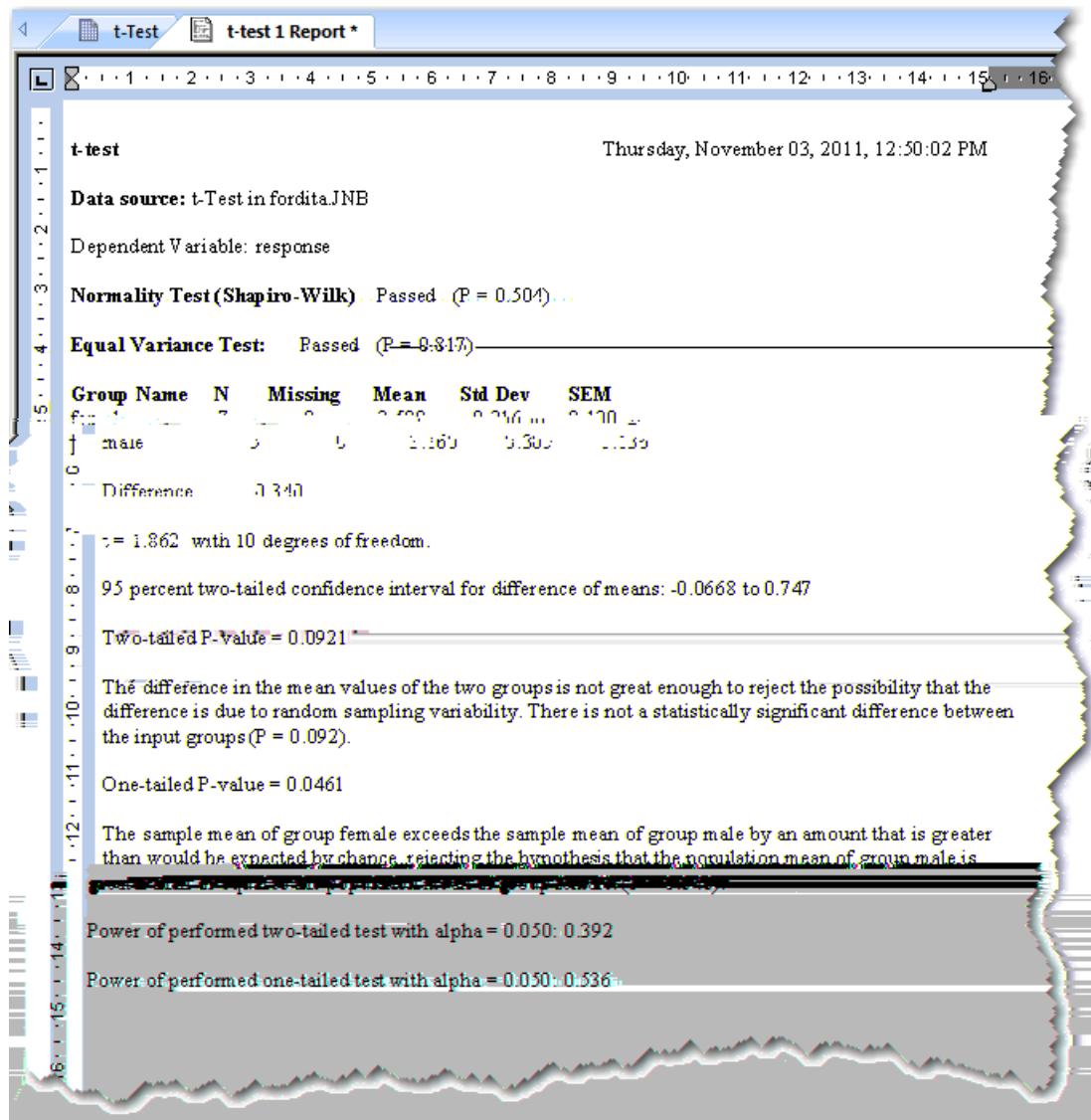


Figure 29: The t-test Report

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can enable or disable this explanatory text in the **Options** dialog box.

Normality Test. Normality test results show whether the data passed or failed the test of the assumption that the samples were drawn from normal populations and the P value calculated by the test. All parametric tests require normally distributed source populations.

This result is set in the **Options for t-test** dialog box.

Equal Variance Test. Equal Variance test results display whether or not the data passed or failed the test of the assumption that the samples were drawn from populations with the same variance and the P value calculated by the test. Equal variance of the source population is assumed for all parametric tests.

Summary Table. SigmaPlot can generate a summary table listing the sizes N for the two samples, number of missing values, means, standard deviations, and the standard error of the means (SEM). This result is displayed unless you disable **Summary Table** in the **Options for t-test** dialog box.

- **N (Size).** The number of non-missing observations for that column or group.
- **Missing.** The number of missing values for that column or group.

- **Mean.** The average value for the column. If the observations are normally distributed the mean is the center of the distribution.
- **Standard Deviation.** A measure of variability. If the observations are normally distributed, about two-thirds will fall within one standard deviation above or below the mean, and about 95% of the observations will fall within two standard deviations above or below the mean.
- **Standard Error of the Mean.** A measure of the approximation with which the mean computed from the sample approximates the true population mean.

t Statistic. The t-test statistic is the ratio:

$$t = \frac{\text{difference between the means of the two groups}}{\text{standard error of the difference between the means}}$$

The standard error of the difference is a measure of the precision with which this difference can be estimated.

You can conclude from "large" absolute values of t that the samples were drawn from different populations. A large t indicates that the difference between the treatment group means is larger than what would be expected from sampling variability alone (for example, that the differences between the two groups are statistically significant). A small t (near 0) indicates that there is no significant difference between the samples.

- **Degrees of Freedom.** Degrees of freedom represents the sample sizes, which affect the ability of the t-test to detect differences in the means. As degrees of freedom (sample sizes) increase, the ability to detect a difference with a smaller t increases.
- **P Value.** The P value is the probability of being wrong in concluding that there is a true difference in the two groups (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on t). The smaller the P value, the greater the probability that the samples are drawn from different populations. Traditionally, you can conclude there is a significant difference when $P < 0.05$.

Confidence Interval for the Difference of the Means. If the confidence interval does not include zero, you can conclude that there is a significant difference between the proportions with the level of confidence specified. This can also be described as $P < \alpha$ (alpha), where α is the acceptable probability of incorrectly concluding that there is a difference.

The level of confidence is adjusted in the **Options for t-test** dialog box; this is typically $100(1-\alpha)$, or 95%. Larger values of confidence result in wider intervals and smaller values in smaller intervals. For a further explanation of α , see Power below. This result is set **Options for t-test** dialog box.

Power. The power, or sensitivity, of a t-test is the probability that the test will detect a difference between the groups if there really is a difference. The closer the power is to 1, the more sensitive the test.

t-test power is affected by the sample size of both groups, the chance of erroneously reporting a difference, α (alpha), the difference of the means, and the standard deviation.

This result is set in the **Options for t-test** dialog box.

Alpha. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. An α error is also called a Type I error (a Type I error is when you reject the hypothesis of no effect when this hypothesis is true).

The α value is set in the **Options for t-test** dialog box; a value of $\alpha = 0.05$ indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference but also increase the risk of reporting a false positive (a Type I error).

t-Test Report Graphs

You can generate up to five graphs using the results from a t-test. They include a:

- **Bar chart of the column means.** The t-test bar chart plots the group means as vertical bars with error bars indicating the standard deviation.
- **Scatter plot with error bars of the column means.** The t-test scatter plot graphs the group means as single points with error bars indicating the standard deviation.

- **Point plot of the column means.** The t-test point plot graphs all values in each column as a point on the graph.
- **Histogram of the residuals.** The t-test histogram plots the raw residuals in a specified range, using a defined interval set.
- **Normal probability plot of the residuals.** The t-test probability plot graphs the frequency of the raw residuals.

How to Create a Graph of the t-test Data

1. Select the **t-test** report.
2. Click the **Report** tab.
3. In the **Results Graphs** group, click **Create Result Graph**.

The **Create Graph** dialog box appears displaying the types of graphs available for the t-test results.

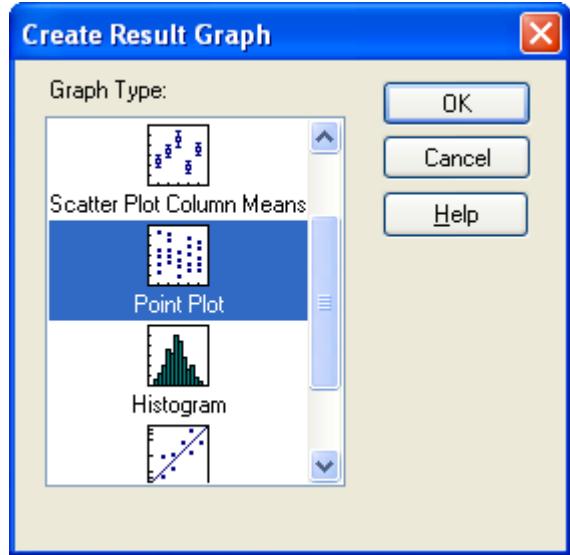


Figure 30: The Create Graph Dialog Box for the t-test Report

4. Select the type of graph you want to create from the **Graph Type** list, then click **OK**, or double-click the desired graph in the list.

The selected [graph](#) appears in a graph window.

Point Plot

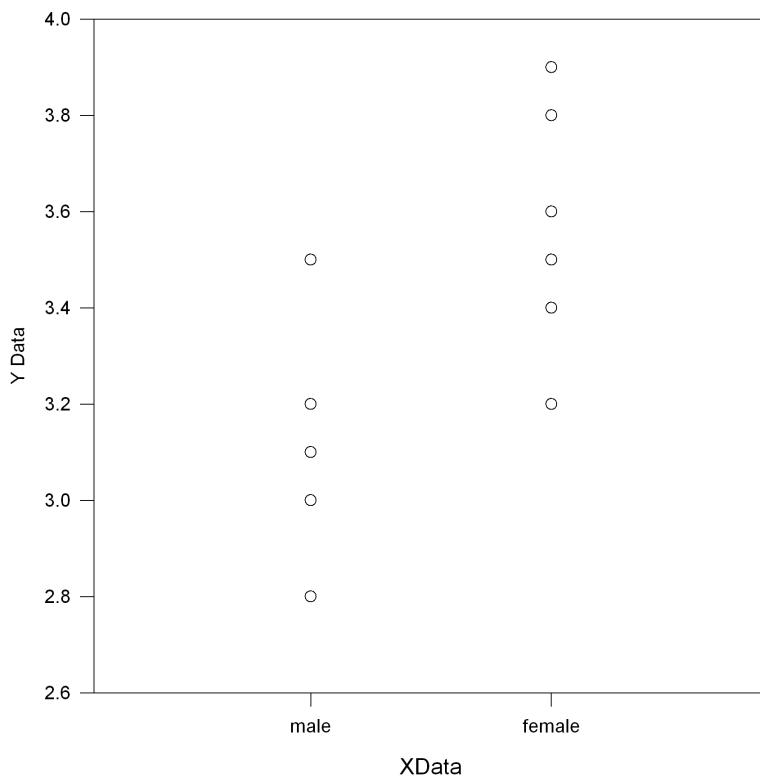


Figure 31: A Point Plot of the Result Data for a t-test

Mann-Whitney Rank Sum Test

Use the **Rank Sum Test** when:

- You want to see if the medians of two different samples are significantly different.
- The samples are not drawn from normally distributed populations with the same variances, or you do not want to assume that they were drawn from normal populations.

If you know your data was drawn from a normally distributed population, use the [Unpaired t-test](#). When there are more than two groups to compare, run a [Kruskal-Wallis ANOVA on Ranks test](#).

 **Note:** Depending on your Rank Sum Test options settings, if you attempt to perform a rank sum test on normal populations with equal variances, SigmaPlot informs you that the data can be analyzed with the more powerful Unpaired t-test instead.

About the Mann-Whitney Rank Sum Test

Use the **Mann-Whitney Rank Sum Test** to test for a difference between two groups that is greater than what can be attributed to random sampling variation. The null hypothesis is that the two samples were not drawn from populations with different medians.

The **Rank Sum Test** is a nonparametric procedure, which does not require assuming normality or equal variance. It ranks all the observations from smallest to largest without regard to which group each observation comes from. The ranks for each group are summed and the rank sums compared.

If there is no difference between the two groups, the mean ranks should be approximately the same. If they differ by a large amount, you can assume that the low ranks tend to be in one group and the high ranks are in the other, and conclude that the samples were drawn from different populations (for example, that there is a statistically significant difference).

Performing a Mann-Whitney Rank Sum Test

To perform a **Mann-Whitney Rank Sum Test**:

1. [Enter or arrange your data appropriately in the worksheet.](#)
2. If desired, set the Rank Sum options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Compare Two Groups > Rank Sum Test
5. Run the test.
6. [Generate report graphs.](#)

Arranging Rank Sum Data

The format of the data to be tested can be raw data or indexed data; in either case, the data is found in two worksheet columns.

Columns 1 and 2 are arranged as raw data. Columns 3 and 4 are arranged as group indexed data, with column 3 as the factor column.

	1-Species A	2-Species B	3-Species C	4-# melanistic	
1	3.0000	3.0000	A	198.0000	
2	1.0000	4.0000	B	301.0000	
3	12.0000	6.0000	B	4.0000	
4	228.0000	291.0000	A	3.0000	
5	238.0000	256.0000	A	22.0000	
6	198.0000	279.0000	B	291.0000	
7	217.0000	282.0000	A	229.0000	
8	22.0000		A	1.0000	
9			B	279.0000	
10			A	217.0000	
11			B	6.0000	
12			A	12.0000	
13			B	292.0000	
14					

Figure 32: Valid Data Formats for a Mann-Whitney Rank Sum Test

Setting Mann-Whitney Rank Sum Test Options

1. Select **Rank Sum Test** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.

- Click Current Test Options. The **Options for Rank Sum Test** dialog box appears with two tabs:
 - Assumption Checking.** Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
 - Results.** Display the statistics summary for the data in the report. Also use this tab to enable the Yates Correction Factor.

Tip: If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.

Options for Rank Sum Test: Assumption Checking

The normality assumption test checks for a normally distributed population. The equal variance assumption test checks the variability about the group means.

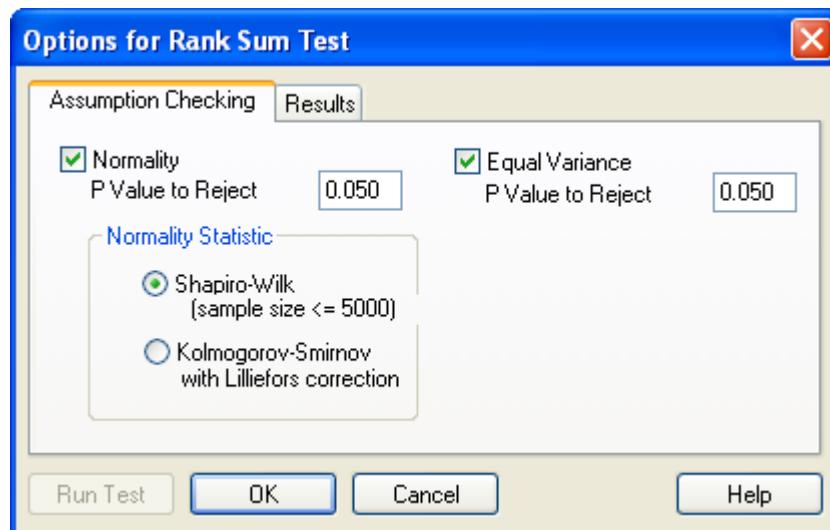


Figure 33: The Options for Rank Sum Test Dialog Box Displaying the Assumption Checking Options

- Normality Testing.** SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.
- Equal Variance Testing.** SigmaPlot tests for equal variance by checking the variability about the group means.
- P Values for Normality and Equal Variance.** The P value determines the probability of being incorrect in concluding that the data is not normally distributed (P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and/or equal variance, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality and/or equal variance, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.010 requires greater deviations from normality to flag the data as non-normal than a value of 0.050.

- Restriction:** There are extreme conditions of data distribution that these tests cannot take into account. For example, the Levene Median test fails to detect differences in variance of several orders of magnitude; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options for Rank Sum Test: Results

Summary Table. Displays the number of observations for a column or group, the number of missing values for a column or group, the average value for the column or group, the standard deviation of the column or group, and the standard error of the mean for the column or group.

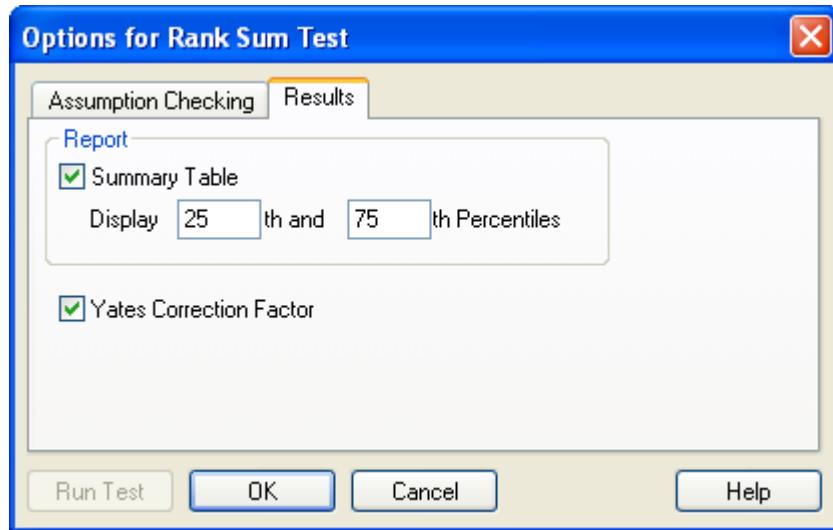


Figure 34: The Options for Rank Sum Test Dialog Box Displaying the Summary Table Options

Yates Correction Factor. When a statistical test uses a χ^2 distribution with one degree of freedom, such as analysis of a 2×2 contingency table or McNemar's test, the χ^2 calculated tends to produce P values which are too small, when compared with the actual distribution of the χ^2 test statistic. The theoretical χ^2 distribution is continuous, whereas the distribution of the χ^2 test statistic is discrete. Use the Yates Correction Factor to adjust the computed χ^2 value down to compensate for this discrepancy. Using the Yates correction makes a test more conservative; for example, it increases the P value and reduces the chance of a false positive conclusion. The Yates correction is applied to 2×2 tables and other statistics where the P value is computed from a χ^2 distribution with one degree of freedom. For descriptions of the derivation of the Yates correction, you can reference any appropriate statistics reference.

Running a Rank Sum Test

If you want to select your data before you run the test, drag the pointer over your data.

1. Click the **Analysis** tab.

2. In the SigmaStat group, from the **Tests** drop-down list, select:

Compare Two Groups > Rank Sum Test

The **Data Format** panel of the Test Wizard appears prompting you to specify a data format.

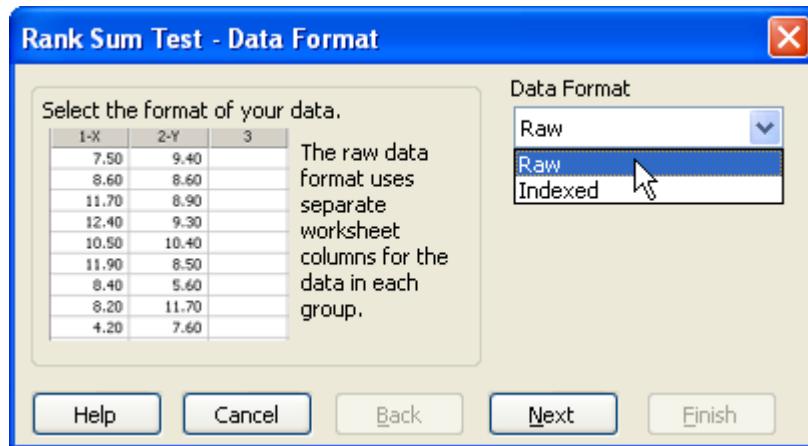


Figure 35: The Rank Sum Test - Data Format Dialog Box Prompting You to Specify a Data Format

3. Select the appropriate data format from the **Data Format** drop-down list.
4. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

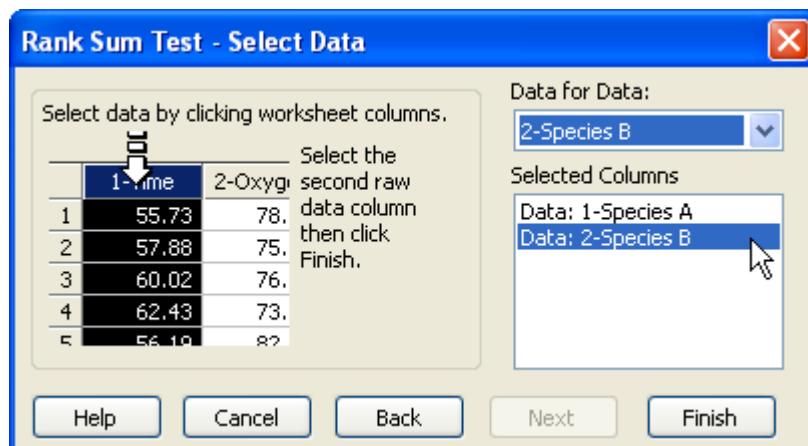


Figure 36: The Rank Sum Test - Select Data Dialog Box Prompting You to Select Data Columns

5. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the Selected Columns list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw and indexed data, you are prompted to select two worksheet columns.

6. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
7. Click **Finish** to run the **Rank Sum Test** on the selected columns.

If you elected to test for normality and equal variance, SigmaPlot performs the test for normality (Shapiro-Wilk or Kolmogorov-Smirnov) and the test for equal variance (Levene Median). If your data pass both tests, SigmaPlot informs you and suggests continuing your analysis using an **unpaired t-test**.

After the computations are completed, the report appears.

Interpreting Rank Sum Test Results

The Rank Sum Test computes the Mann-Whitney T statistic and the P value for T. These results are displayed in the rank sum report which appears after the rank sum test is performed. The other results displayed in the report are enabled and disabled in the Options for Rank Sum Test dialog box.

Mann-Whitney Rank Sum Test Thursday, November 03, 2011, 1:02:46 PM

Data source: Rank Sum Test in fordita.JNB

Dependent Variable: # melanistic

Normality Test (Shapiro-Wilk) Failed ($P < 0.050$)

Group	N	Missing	Median	25%	75%
A	7	0	22.000	3.000	217.000
B	6	0	285.000	5.500	294.250

Mann-Whitney U Statistic= 10.000

T = 53.000 n(small)= 6 n(big)= 7 P(est.)= 0.134 P(exact)= 0.138

The difference in the median values between the two groups is not great enough to exclude the possibility that the difference is due to random sampling variability; there is not a statistically significant difference ($P = 0.138$)

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can enable or disable this explanatory text in the **Options** dialog box.

Normality Test. Normality test results display whether the data passed or failed the test of the assumption that they were drawn from a normal population and the P value calculated by the test. For nonparametric procedures, this test can have failed, as nonparametric tests do not assume normally distributed source populations. This result is set in the Options for Rank Sum Test dialog box.

Equal Variance Test. Equal Variance test results display whether or not the data passed or failed the test of the assumption that the samples were drawn from populations with the same variance and the P value calculated by the test. Nonparametric tests do not assume equal variance of the source populations. This result is set in the Options for Rank Sum Test dialog box.

Summary Table. SigmaPlot generates a summary table listing the sample sizes N, number of missing values, medians, and percentiles unless you disable the Display Summary Table option in the Options for Rank Sum Test dialog box.

- **N (Size).** The number of non-missing observations for that column or group.
- **Missing.** The number of missing values for that column or group.
- **Medians.** The "middle" observation as computed by listing all the observations from smallest to largest and selecting the largest value of the smallest half of the observations. The median observation has an equal number of observations greater than and less than that observation.
- **Percentiles.** The two percentile points that define the upper and lower tails of the observed values.
- **T Statistic.** The T statistic is the sum of the ranks in the smaller sample group or from the first selected group, if both groups are the same size. This value is compared to the population of all possible rankings to determine the possibility of this T occurring.

- **P Value.** The *P* value is the probability of being wrong in concluding that there is a true difference in the two groups (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on *T*). The smaller the *P* value, the greater the probability that the samples are drawn from different populations.

Traditionally, you can conclude there is a significant difference when $P < 0.05$.

Rank Sum Test Report Graphs

You can generate up to two graphs using the results from a Rank Sum Test. They include a:

- **Box plot of the percentiles and median of column data.** The Rank Sum Test box plot graphs the percentiles and the median of column data. The ends of the boxes define the 25th and 75th percentiles, with a line at the median and error bars defining the 10th and 90th percentiles.
- **Point plot of the column data.** The Rank Sum Test point plot graphs all values in each column as a point on the graph.

How to Create a Rank Sum Test Report Graph

1. Select the **Rank Sum Test** report.
2. Click the **Report** tab.
3. In the **Results Graphs** group, click **Create Result Graph**.

The **Create Graph** dialog box appears displaying the types of graphs available for the Rank Sum Test results.

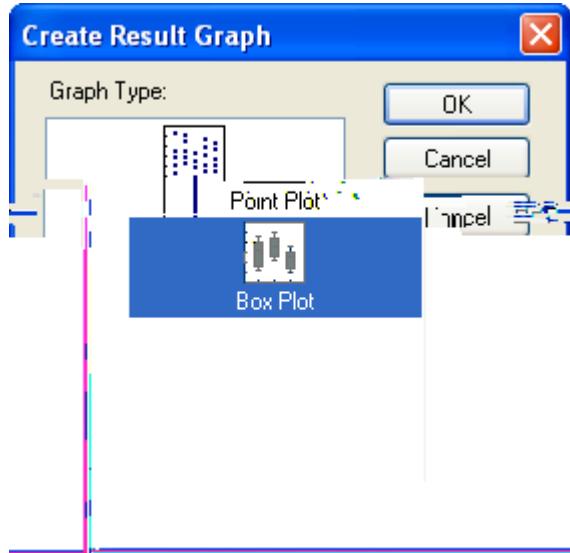


Figure 37: The Create Graph Dialog Box for the Rank Sum Test Report

4. Select the type of graph you want to create from the **Graph Type** list, then click **OK**, or double-click the desired graph in the list.

The selected [graph](#) appears in a graph window.

Box Plot

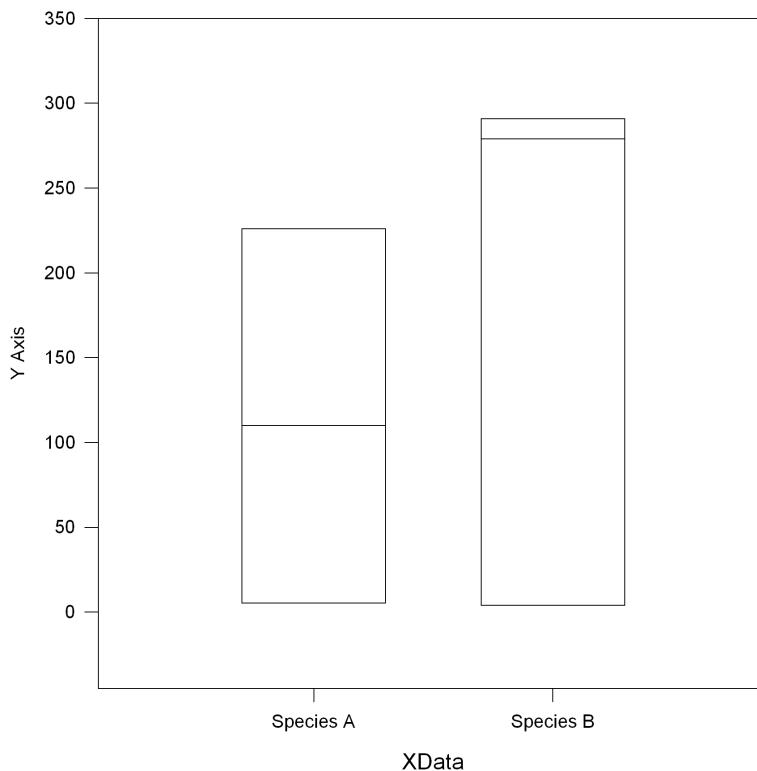


Figure 38: A Box Plot of the Result Data for a Rank Sum Test

One Way Analysis of Variance (ANOVA)

One Way Analysis of Variance is a parametric test that assumes that all the samples are drawn from normally distributed populations with the same standard deviations (variances).

Use a One Way or One Factor ANOVA when:

- You want to see if the means of two or more different experimental groups are affected by a single factor.
- Your samples are drawn from normally distributed populations with equal variance.

If you know that your data was drawn from non-normal populations, use the [Kruskal-Wallis ANOVA on Ranks](#) test.

If you want to consider the effects of two factors on your experimental groups, use Two Way ANOVA. [Two Way Analysis of Variance \(ANOVA\)](#) on page 89

When there are only two groups to compare, you can do a t-test (depending on the type of results you want). Performing an ANOVA for two groups yields exactly the same P value as an [unpaired t-test](#).



Note: Depending on your ANOVA options settings, if you attempt to perform an ANOVA on non-normal populations or populations with unequal variances, SigmaStat informs you that the data is unsuitable for a parametric test, and suggests the Kruskal-Wallis ANOVA on Ranks.

About One Way ANOVA

The design for a One Way ANOVA is the same as an unpaired t-test except that there can be more than two experimental groups. The null hypothesis is that there is no difference among the populations from which the samples were drawn.

Performing a One Way ANOVA

To perform a One Way ANOVA:

1. Enter or arrange your data appropriately in the worksheet.
2. If desired, set One Way ANOVA options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Compare Many Groups > One Way ANOVA
5. Run the test.
6. Generate report graphs.

Arranging One Way ANOVA Data

Arrange data as raw data, indexed data, or summary statistics. Place raw data in as many columns as there are groups, up to 640; each column contains the data for one group. Place indexed data in two worksheet columns. Place statistical summary data in three columns.

Columns 1 through 3 are arranged as groups in columns. Columns 4, 5, and 6 are arranged as descriptive statistics using the mean, standard deviation, and size. Columns 7 and 8 are arranged as group indexed data, with column 7 as the factor column.

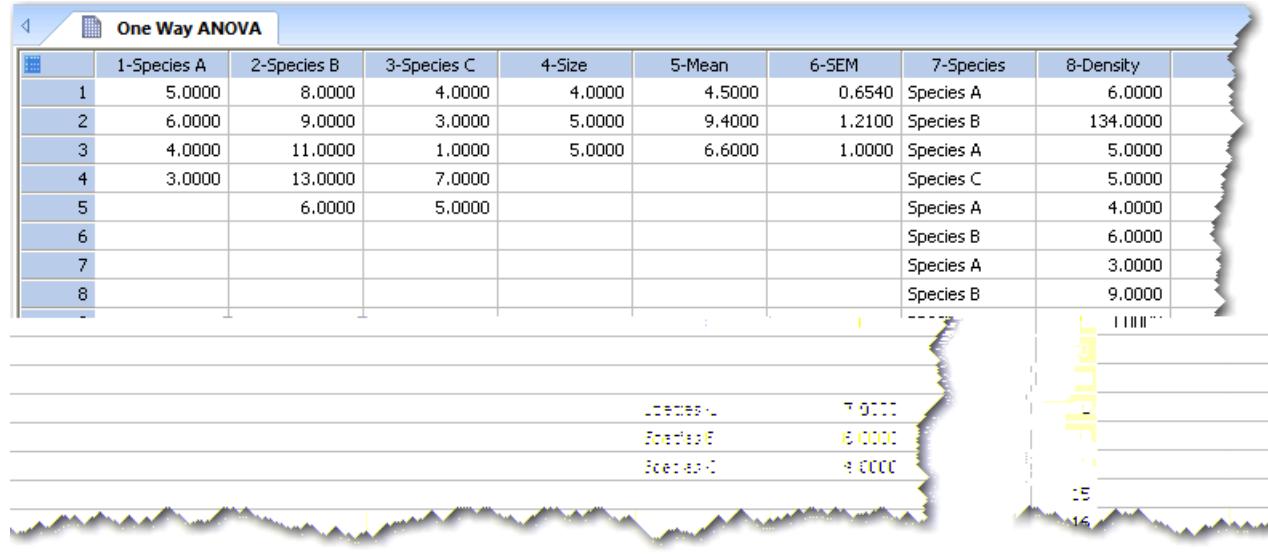


Figure 39: Valid Data Formats for a One Way ANOVA

Setting One Way ANOVA Options

1. Select **One Way ANOVA** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.

- Click Current Test Options. The **Options for One Way ANOVA** dialog box appears with three tabs:
 - Assumption Checking.** Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
 - Results.** Display the statistics summary for the data in the report and save residuals to a worksheet column.
 - Post Hoc Test.** Compute the power or sensitivity of the test and enable multiple comparisons.

Tip: If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
- To continue the test, click **Run Test**.
- To accept the current settings and close the options dialog box, click **OK**.

Options for One Way ANOVA: Assumption Checking

The normality assumption test checks for a normally distributed population. The equal variance assumption test checks the variability about the group means.

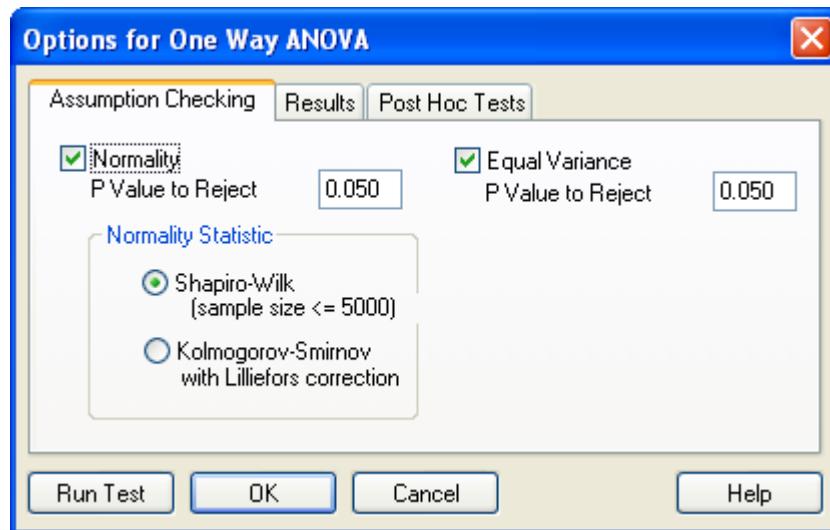


Figure 40: The Options for One Way ANOVA Dialog Box Displaying the Assumption Checking Options

- Normality Testing.** SigmaPlot uses the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.
- Equal Variance Testing.** SigmaPlot tests for equal variance by checking the variability about the group means.
- P Values for Normality and Equal Variance.** The P value determines the probability of being incorrect in concluding that the data is not normally distributed (P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and/or equal variance, decrease the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality and/or equal variance, increase P . Requiring larger values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.100 requires greater deviations from normality to flag the data as non-normal than a value of 0.050.

- Restriction:** There are extreme conditions of data distribution that these tests cannot take into account. For example, the Levene Median test fails to detect differences in variance of several orders of magnitude; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options for One Way ANOVA: Results

Summary Table. Select to display the number of observations for a column or group, the number of missing values for a column or group, the average value for the column or group, the standard deviation of the column or group, and the standard error of the mean for the column or group.

Residuals in Column. Select to display residuals in the report and to save the residuals of the test to the specified worksheet column. Edit the number or select a number from the drop-down list.

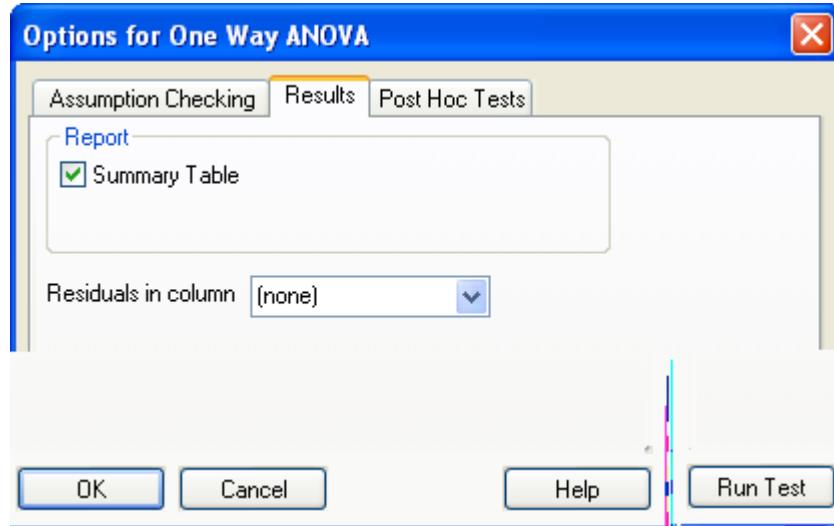


Figure 41: The Options for One Way ANOVA Dialog Box Displaying the Summary Table and Residuals Options

Options for One Way ANOVA: Post Hoc Tests

Power. The power or sensitivity of a test is the probability that the test will detect a difference between the groups if there is really a difference.

Use Alpha Value. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

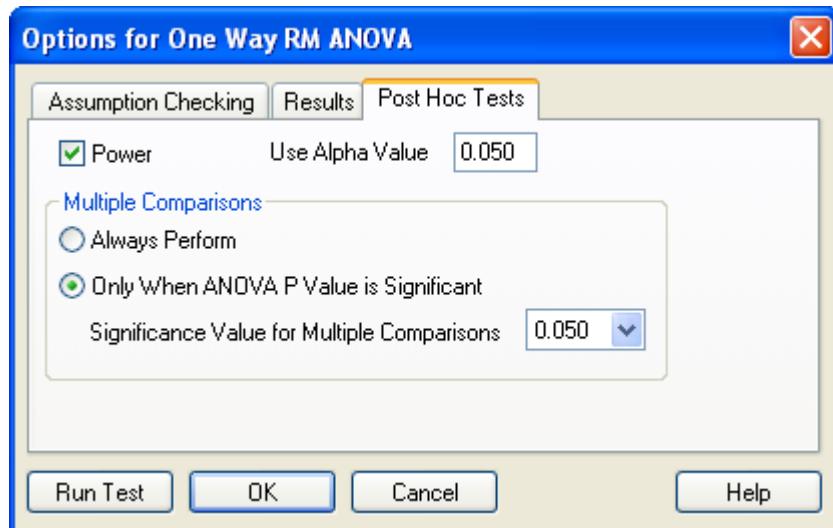


Figure 42: The Options for One Way ANOVA Dialog Box Displaying the Power and Multiple Comparison Options

Multiple Comparisons

One-Way ANOVAs test the hypothesis of no differences between the several treatment groups, but do not determine which groups are different, or the sizes of these differences. Multiple comparison procedures isolate these differences. You can choose to always perform multiple comparisons or to only perform multiple comparisons if a One Way ANOVA detects a difference.

The *P* value used to determine if the ANOVA detects a difference is set on the **Report** tab of the **Options** dialog box. If the *P* value produced by the One Way ANOVA is less than the *P* value specified in the box, a difference in the groups is detected and the multiple comparisons are performed.

- **Always Perform.** Select to perform multiple comparisons whether or not the ANOVA detects a difference.
- **Only When ANOVA P Value is Significant.** Select to perform multiple comparisons only if the ANOVA detects a difference.
- **Significance Value for Multiple Comparisons.** Select either .05 or .01 from the Significance Value for Multiple Comparisons drop-down list. This value determines the that the likelihood of the multiple comparison being incorrect in concluding that there is a significant difference in the treatments.

A value of .05 indicates that the multiple comparisons will detect a difference if there is less than 5% chance that the multiple comparison is incorrect in detecting a difference.

 **Note:** If multiple comparisons are triggered, the **Multiple Comparison Options** dialog box appears after you pick your data from the worksheet and run the test, prompting you to choose a multiple comparison method.

Running a One Way ANOVA

If you want to select your data before you run the test, drag the pointer over your data.

1. Click the **Analysis** tab.

2. In the SigmaStat group, from the **Tests** drop-down list, select:

Compare Many Groups > One Way ANOVA

The **Data Format** panel of the Test Wizard appears prompting you to specify a data format.

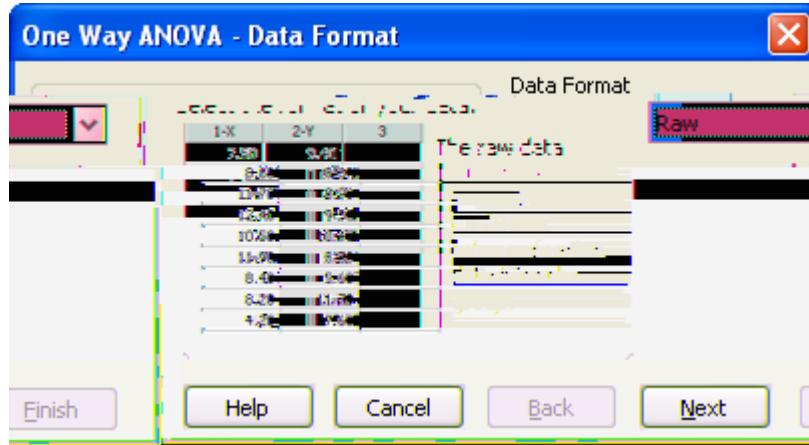


Figure 43: The One Way ANOVA - Data Format Dialog Box Prompting You to Specify a Data Format

3. Select the appropriate data format from the **Data Format** drop-down list.
4. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

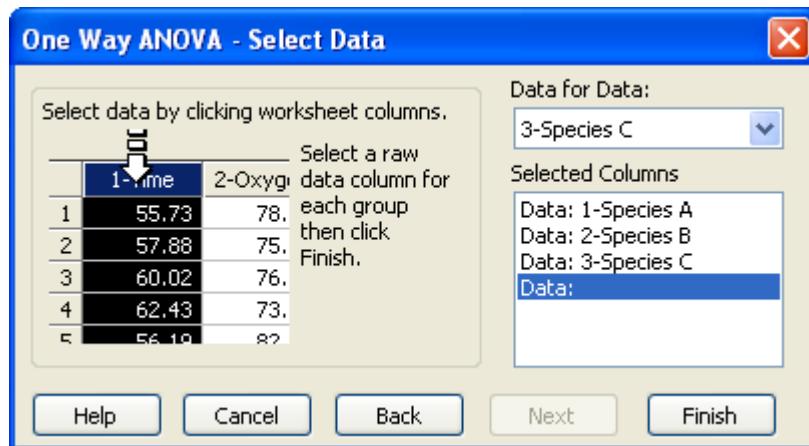


Figure 44: The Select Data Panel of the Test Wizard Prompting You to Select Data Columns

5. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw and indexed data, you are prompted to select two worksheet columns.

6. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
7. Click **Finish** to run the **One Way ANOVA** on the selected columns.

If you elected to test for normality and equal variance, SigmaPlot performs the test for normality (Shapiro-Wilk or Kolmogorov-Smirnov) and the test for equal variance (Levene Median).

If your data pass both tests, SigmaPlot informs you and suggests continuing your analysis using a **Paired t-Test** on page 141.

After the computations are completed, the **report** appears.

8. Click **Finish** to perform the One Way ANOVA.

If you elected to test for normality and equal variance, and your data fails either test, SigmaPlot warns you and suggests continuing your analysis using the nonparametric [Kruskal-Wallis ANOVA on Ranks](#).

If you selected to run multiple comparisons only when the P value is significant, and the P value is not significant, the [One Way ANOVA report](#) appears after the test is complete.

If the P value for multiple comparisons is significant, or you selected to always perform multiple comparisons, the [Multiple Comparisons Options](#) dialog box appears prompting you to select a multiple comparison method.

Multiple Comparison Options for a One Way ANOVA

The **One Way ANOVA** tests the hypothesis of no differences between the several treatment groups, but do not determine which groups are different, or the sizes of these differences. Multiple comparison tests isolate these differences by running comparisons between the experimental groups.

If you selected to run multiple comparisons only when the P value is significant, and the ANOVA produces a P value equal to or less than the trigger P value, or you selected to always run multiple comparisons in the **Options for One Way ANOVA** dialog box, the **Multiple Comparison Options** dialog box appears prompting you to specify a multiple comparison test. The P value produced by the ANOVA is displayed in the upper left corner of the dialog box.

There are seven multiple comparison tests to choose from for the One Way ANOVA:

- [Holm Sidak test](#).
- [Tukey Test](#).
- [Student-Newman-Keuls Test](#).
- [Bonferroni t-test](#).
- [Fisher's LSD](#).
- [Dunnet's Test](#).
- [Duncan's Multiple Range Test](#).

There are two types of multiple comparisons available for the One Way ANOVA. The types of comparison you can make depends on the selected multiple comparison test.

- All pairwise comparisons compare all possible pairs of treatments.
- Multiple comparisons versus a control compare all experimental treatments to a single control group.

The **Tukey** and **Student-Newman-Keuls** tests are recommended for determining the difference among all treatments. If you have only a few treatments, you may want to select the simpler **Bonferroni t-test**.

The **Dunnett's test** is recommended for determining the differences between the experimental treatments and a control group. If you have only a few treatments or observations, you can select the simpler **Bonferroni t-test**.

 **Note:** In both cases the Bonferroni t-test is most sensitive with a small number of groups. Dunnett's test is not available if you have fewer than six observations.

Interpreting One Way ANOVA Results

The One Way ANOVA report displays an ANOVA table describing the source of the variation in the groups. This table displays the sum of squares, degrees of freedom, and mean squares of the groups, as well as the F statistic and the corresponding P value. The statistical summary table of the data and other results displayed in the report are enabled and disabled in the Options for One Way ANOVA dialog box.

You can also generate tables of multiple comparisons. Multiple Comparison results are also specified in the Options for One Way ANOVA dialog box. The test used to perform the multiple comparison is selected in the Multiple Comparison Options dialog box.

One Way ANOVA One Way ANOVA 3 Report *

One Way Analysis of Variance Thursday, November 03, 2011, 1:12:39 PM

Data source: One Way ANOVA in fordita.JNB

Dependent Variable: Density

Normality Test (Shapiro-Wilk) Failed (P < 0.050)

Equal Variance Test: Passed (P = 0.395)

Group Name N Missing Mean Std Dev SEM

Species A	4	0	4.500	1.291	0.645
Species B	5	0	33.600	56.154	25.113
Species C	5	0	4.000	2.236	1.000

Source of Variation DF SS MS F P

Between Groups	2	2774.657	1387.329	1.207	0.336
Within Groups	11	12628.200	1148.020		

The differences in the mean values among the treatment groups are not great enough to exclude the possibility that the difference is due to random sampling variability, there is not a statistically significant difference (P = 0.336).

Power of performed test with $\alpha = 0.05$ is 0.74.

The power of the performed test (0.74) is below the desired power of 0.800.

Less than desired power indicates you are less likely to detect a difference when one actually exists.

Negative results should be interpreted cautiously.

Result Explanations

In addition to the numerical results, [expanded explanations of the results](#) may also appear. You can turn off this text on the [Options](#) dialog box.

Normality Test. Normality test results display whether the data passed or failed the test of the assumption that they were drawn from a normal population and the P value calculated by the test. Normally distributed source populations are required for all parametric tests. Set this result in the Options for One Way ANOVA dialog box.

Equal Variance Test. Equal Variance test results display whether or not the data passed or failed the test of the assumption that the samples were drawn from populations with the same variance, and the P value calculated by the test. Equal variance of the source populations is assumed for all parametric tests. This result appears is set in the Options for One Way ANOVA dialog box.

Summary Table. If you enabled this option in the Options for One Way ANOVA dialog box, SigmaPlot generates a summary table listing the sample sizes N, number of missing values, mean, standard deviation, differences of the means and standard deviations, and standard error of the means.

- **N (Size).** The number of non-missing observations for that column or group.
- **Missing.** The number of missing values for that column or group.
- **Mean.** The average value for the column. If the observations are normally distributed the mean is the center of the distribution.
- **Standard Deviation.** A measure of variability. If the observations are normally distributed, about two-thirds will fall within one standard deviation above or below the mean, and about 95% of the observations will fall within two standard deviations above or below the mean.

- **Standard Error of the Mean.** A measure of the approximation with which the mean computed from the sample approximates the true population mean.

Power. The power of the performed test is displayed unless you disable this option in the Options for One Way ANOVA dialog box.

The power, or sensitivity, of a One Way ANOVA is the probability that the test will detect a difference among the groups if there really is a difference. The closer the power is to 1, the more sensitive the test.

ANOVA power is affected by the sample sizes, the number of groups being compared, the chance of erroneously reporting a difference a (alpha), the observed differences of the group means, and the observed standard deviations of the samples.

Alpha. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. An α error is also called a Type I error. A Type I error is when you reject the hypothesis of no effect when this hypothesis is true.

Set this value in the Options for One Way ANOVA dialog box; the suggested value is $\alpha = 0.05$ which indicates that a one in twenty chance of error is acceptable. Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference but also increase the risk of seeing a false difference (a Type I error).

ANOVA Table. The ANOVA table lists the results of the one way ANOVA.

DF (Degrees of Freedom). Degrees of freedom represent the number of groups and sample size which affects the sensitivity of the ANOVA.

- The degrees of freedom between groups is a measure of the number of groups
- The degrees of freedom within groups (sometimes called the error or residual degrees of freedom) is a measure of the total sample size, adjusted for the number of groups
- The total degrees of freedom is a measure of the total sample size

SS (Sum of Squares). The sum of squares is a measure of variability associated with each element in the ANOVA data table.

- The sum of squares between the groups measures the variability of the average differences of the sample groups
- The sum of squares within the groups (also called error or residual sum of squares) measures the underlying variability of all individual samples
- The total sum of squares measures the total variability of the observations about the grand mean (mean of all observations)

MS (Mean Squares). The mean squares provide two estimates of the population variances. Comparing these variance estimates is the basis of analysis of variance.

The mean square between groups is:

$$\frac{\text{sum of squares between groups}}{\text{degrees of freedom between groups}} = \frac{SS \text{ between}}{DF \text{ between}} = MS \text{ between}$$

The mean square within groups (also called the residual or error mean square) is:

$$\frac{\text{sum of squares within groups}}{\text{degrees of freedom within groups}} = \frac{SS \text{ within}}{DF \text{ within}} = MS \text{ within}$$

F Statistic. The F test statistic is the ratio:

$$\frac{\text{estimated population variance between groups}}{\text{estimated population variance within groups}} = \frac{MS \text{ between}}{MS \text{ within}} = F$$

If the F ratio is around 1, you can conclude that there are no significant differences between groups (for example, the data groups are consistent with the null hypothesis that all the samples were drawn from the same population).

If F is a large number, you can conclude that at least one of the samples was drawn from a different population (for example, the variability is larger than what is expected from random variability in the population). To determine exactly which groups are different, examine the multiple comparison results.

P Value. The P value is the probability of being wrong in concluding that there is a true difference between the groups (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on F). The smaller the P value, the greater the probability that the samples are drawn from different populations. Traditionally, you can conclude that there are significant differences when $P < 0.05$.

Multiple Comparisons. If you selected to perform multiple comparisons, a table of the comparisons between group pairs is displayed. The multiple comparison procedure is activated in the Options for One Way ANOVA dialog box. The tests used in the multiple comparison procedure is selected in the Multiple Comparison Options dialog box.

Multiple comparison results are used to determine exactly which treatments are different, since the ANOVA results only inform you that two or more of the groups are different. The specific type of multiple comparison results depends on the comparison test used and whether the comparison was made pairwise or versus a control.

- All pairwise comparison results list comparisons of all possible combinations of group pairs; the all pairwise tests are the Tukey, Student-Newman-Keuls, Fisher LSD, Duncan's test and the Bonferroni t-test.
- Comparisons versus a single control group list only comparisons with the selected control group. The control group is selected during the actual multiple comparison procedure. The comparison versus a control tests are the Bonferroni t-test and the Dunnett's, Fishers LSD, and Duncan's tests.

For descriptions of the derivations of parametric multiple comparison procedure results, you can reference any appropriate statistics reference.

Bonferroni t-test Results. The Bonferroni t-test lists the differences of the means for each pair of groups, computes the t values for each pair, and displays whether or not $P < 0.05$ for that comparison. The Bonferroni t-test can be used to compare all groups or to compare versus a control.

You can conclude from "large" values of t that the difference of the two groups being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of erroneously concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The difference of the means is a gauge of the size of the difference between the two groups.

Tukey, Student-Newman-Keuls, Fisher LSD, Duncan's, and Dunnett's Test Results. The Tukey, Student-Newman-Keuls (SNK), Fisher LSD, and Duncan's tests are all pairwise comparisons of every combination of group pairs. While the Tukey Fisher LSD, and Duncan's can be used to compare a control group to other groups, they are not recommended for this type of comparison.

Dunnett's test only compares a control group to all other groups. All tests compute the q test statistic, and display whether or not $P < 0.05$ or < 0.01 for that pair comparison.

You can conclude from "large" values of q that the difference of the two groups being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of being incorrect in concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The Difference of the Means is a gauge of the size of the difference between the two groups.

p is a parameter used when computing q. The larger the p, the larger q needs to be to indicate a significant difference. p is an indication of the differences in the ranks of the group means being compared. Groups means are ranked in order from largest to smallest, and p is the number of means spanned in the comparison. For example, if you are comparing four means, when comparing the largest to the smallest p = 4, and when comparing the second smallest to the smallest p = 2. For the Tukey test, the p is always equal to the total number of groups.

If a group is found to be not significantly different than another group, all groups with p ranks in between the p ranks of the two groups that are not different are also assumed not to be significantly different, and a result of DNT (Do Not Test) appears for those comparisons.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the **Options** dialog box.

You can also set the number of decimal places to display the **Options** dialog box.

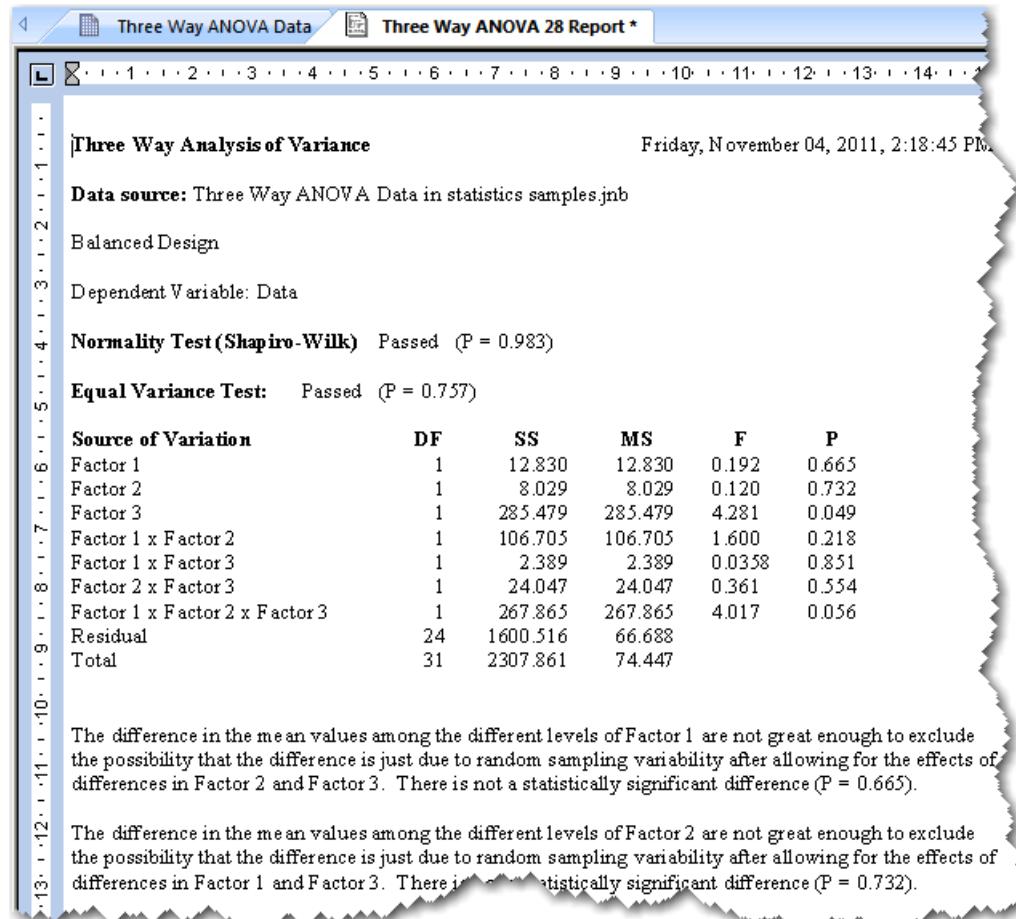


Figure 45: Three Way ANOVA Report

If There Were Missing Data Cells

If your data contained missing values but no empty cells, the report indicates the results were computed using a general linear model.

If your data contained empty cells, you either analyzed the problem assuming either no interaction or treated the problem as a Two or One Way ANOVA.

- If you choose no interactions, no statistics for factor interaction are calculated.
- If you performed a Two or One Way ANOVA, the results shown are identical to Two and One Way ANOVA results. For more information, see [Interpreting One Way ANOVA Results](#) on page 79.

Dependent Variable

This is the data column title of the indexed worksheet data you are analyzing with the Three Way ANOVA. Determining if the values in this column are affected by the different factor levels is the objective of the Three Way ANOVA.

Normality Test

Normality test results display whether the data passed or failed the test of the assumption that they were drawn from a normal population and the P value calculated by the test. Normally distributed source populations are required for all parametric tests.

This result appears if you enabled normality testing in the Options for Three Way ANOVA dialog box.

Equal Variance Test

Equal Variance test results display whether or not the data passed or failed the test of the assumption that the samples were drawn from populations with the same variance and the P value calculated by the test. Equal variance of the source population is assumed for all parametric tests.

This result appears if you enabled equal variance testing in the Options for Three Way ANOVA dialog box.

ANOVA Table

The ANOVA table lists the results of the Three Way ANOVA.

 **Note:** When there are missing data, the best estimate of these values is automatically calculated using a *general linear model*.

DF (Degrees of Freedom)

Degrees of freedom represent the number of groups in each factor and the sample size, which affects the sensitivity of the ANOVA.

- The degrees of freedom for each factor is a measure of the number of levels in each factor.
- The interaction degrees of freedom is a measure of the total number of cells.
- The error degrees of freedom (sometimes called the residual or within groups degrees of freedom) is a measure of the sample size after accounting for the factors and interaction.
- The total degrees of freedom is a measure of the total sample size .

SS (Sum of Squares)

The sum of squares is a measure of variability associated with each element in the ANOVA data table.

- The factor sums of squares measure the variability within between the rows or columns of the table considered separately.
- The interaction sum of squares measures the variability of the average differences between the cell in addition to the variation between the rows and columns, considered separately—this is a gauge of the interaction between the factors.
- The error sum of squares (also called residual or within group sum of squares) is a measure of the underlying random variation in the data, for example, the variability not associated with the factors or their interaction.
- The total sum of squares is a measure of the total variability in the data; if there are no missing data, the total sum of squares equals the sum of the other table sums of squares.

MS (Mean Squares)

The mean squares provide different estimates of the population variances. Comparing these variance estimates is the basis of analysis of variance.

The mean square for each factor:

$$\frac{\text{sum of squares for the factor}}{\text{degrees of freedom for the factor}} = \frac{SS_{\text{factor}}}{DF_{\text{factor}}} = MS_{\text{factor}}$$

is an estimate of the variance of the underlying population computed from the variability between levels of the factor.

The interaction mean square:

$$\frac{\text{sum of squares for the interaction}}{\text{degrees of freedom for the interaction}} = \frac{SS_{\text{inter}}}{DF_{\text{inter}}} = MS_{\text{inter}}$$

is an estimate of the variance of the underlying population computed from the variability associated with the interactions of the factors.

The error mean square (residual, or within groups):

$$\frac{\text{error sum of squares}}{\text{error degrees of freedom}} = \frac{SS_{\text{error}}}{DF_{\text{error}}} = MS_{\text{error}}$$

is an estimate of the variability in the underlying population, computed from the random component of the observations.

F Statistic

The F test statistic is provided for comparisons within each factor and between the factors.

The F ratio to test each factor is:

$$\frac{\text{mean square for the factor}}{\text{mean square of the error}} = \frac{MS_{\text{factor}}}{MS_{\text{error}}} = F_{\text{factor}}$$

The F ratio to test the interaction is:

$$\frac{\text{mean square for the interaction}}{\text{mean square of the error}} = \frac{MS_{\text{inter}}}{MS_{\text{error}}} = F_{\text{inter}}$$

If the F ratio is around 1, you can conclude that there are no significant differences between factor levels or that there is no interaction between factors (for example, the data groups are consistent with the null hypothesis that all the samples were drawn from the same population).

If F is a large number, you can conclude that at least one of the samples for that factor or combination of factors was drawn from a different population (for example, the variability is larger than what is expected from random variability in the population). To determine exactly which groups are different, examine the multiple comparison results.

P Value

The P value is the probability of being wrong in concluding that there is a true difference between the groups (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on F). The smaller the P value, the greater the probability that the samples are drawn from different populations.

Traditionally, you can conclude there are significant differences if $P < 0.05$.

Power

The power, or sensitivity, of a Three Way ANOVA is the probability that the test will detect the observed difference among the groups if there really is a difference. The closer the power is to 1, the more sensitive the test. The power for the comparison of the groups within the two factors and the power for the comparison of the interactions are all displayed. These results are set in the Options for Three Way ANOVA dialog box.

ANOVA power is affected by the sample sizes, the number of groups being compared, the chance of erroneously reporting a difference a (alpha), the observed differences of the group means, and the observed standard deviations of the samples.

Alpha

Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. An α error also is called a *Type I error* (a Type I error is when you reject the hypothesis of no effect when this hypothesis is true).

Set the value in the Options for Three Way ANOVA dialog box; the suggested value is $\alpha = 0.05$ which indicates that a one in twenty chance of error is acceptable. Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a *Type II error*). Larger values of α make it easier to conclude that there is a difference, but also increase the risk of seeing a false difference (a Type I error).

Summary Table

The least square means and standard error of the means are displayed for each factor separately (summary table row and column), and for each combination of factors (summary table cells). If there are missing values, the least square means are estimated using a general linear model.

Mean. The average value for the column. If the observations are normally distributed the mean is the center of the distribution.

Standard Error of the Mean. A measure of the approximation with which the mean computed from the sample approximates the true population mean.

When there are no missing data, the least square means equal the cell and marginal (row and column) means. When there are missing data, the least squared means provide the best estimate of these values, using a general linear model. These means and standard errors are used when performing multiple comparisons. For more information, see [Multiple Comparisons](#) on page 86.

Multiple Comparisons

If a difference is found among the groups, multiple comparison tables can be computed. Multiple comparison procedures are activated in the Options for Three Way ANOVA dialog box. The tests used in the multiple comparisons are set in the Multiple Comparisons Options dialog box.

Use multiple comparison results to determine exactly which groups are different, since the ANOVA results only inform you that three or more of the groups are different. Three factor multiple comparison for a full Three Way ANOVA also compares:

- Groups within each factor without regard to the other factor (this is a *marginal* comparison; for example, only the columns or rows in the table are compared).
- All combinations of factors (all cells in the table are compared with each other).

The specific type of multiple comparison results depends on the comparison test used and whether the comparison was made pairwise or versus a control.

- All pairwise comparison results list comparisons of all possible combinations of group pairs; the all pairwise tests are the Tukey, Student-Newman-Keuls, Fisher LSD, Duncan's, and Dunnett's, and Bonferroni t-test.
- Comparisons versus a single control group list only comparisons with the selected control group. The control group is selected during the actual multiple comparison procedure. The comparison versus a control tests are a Bonferroni t-test and Dunnett's test.

Bonferroni t-test Results The Bonferroni t-test lists the differences of the means for each pair of groups, computes the t values for each pair, and displays whether or not $P < 0.05$ for that comparison. The Bonferroni t-test can be used to compare all groups or to compare versus a control.

You can conclude from "large" values of t that the difference of the two groups being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of erroneously concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The Difference of Means is a gauge of the size of the difference between the levels or cells being compared.

The degrees of freedom DF for the marginal comparisons are a measure of the number of groups (levels) within the factor being compared. The degrees of freedom when comparing all cells is a measure of the sample size after accounting for the factors and interaction. This is the same as the error or residual degrees of freedom.

Tukey, Student-Newman-Keuls, Fisher LSD, Duncan's, and Dunnett's Test Results The Tukey, Student-Newman-Keuls (SNK), Fisher LSD, and Duncan's tests are all pairwise comparisons of every combination of group pairs. While the Tukey Fisher LSD, and Duncan's can be used to compare a control group to other groups, they are not recommended for this type of comparison.

Dunnett's test only compares a control group to all other groups. All tests compute the q test statistic and display whether or not $P < 0.05$ for that pair comparison.

You can conclude from "large" values of q that the difference of the two groups being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of being incorrect in concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

p is a parameter used when computing q . The larger the p , the larger q needs to be to indicate a significant difference. p is an indication of the differences in the ranks of the group means being compared. Groups means are ranked in order from largest to smallest, and p is the number of means spanned in the comparison. For example, when comparing four means, comparing the largest to the smallest $p = 4$, and when comparing the second smallest to the smallest $p = 2$.

If a group is found to be not significantly different than another group, all groups with p ranks in between the p ranks of the two groups that are not different are also assumed not to be significantly different, and a result of DNT (Do Not Test) appears for those comparisons.

The Difference of Means is a gauge of the size of the difference between the groups or cells being compared.

The degrees of freedom DF for the marginal comparisons are a measure of the number of groups (levels) within the factor being compared. The degrees of freedom when comparing all cells is a measure of the sample size after accounting for the factors and interaction (this is the same as the error or residual degrees of freedom).

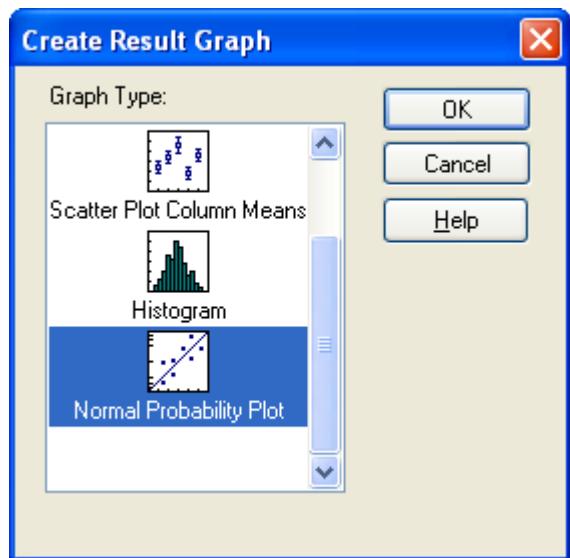
One Way ANOVA Report Graphs

You can generate up to five graphs using the results from a One Way ANOVA. They include a:

- **Bar chart of the column means.** The One Way ANOVA bar chart plots the group means as vertical bars with error bars indicating the standard deviation.
- **Scatter plot with error bars of the column means.** The One Way ANOVA scatter plot graphs the group means as single points with error bars indicating the standard deviation.
- **Histogram of the residuals.** The One Way ANOVA histogram plots the raw residuals in a specified range, using a defined interval set.
- **Normal probability plot of the residuals.** The One Way ANOVA probability plot graphs the frequency of the residuals against the expected frequency under a normal distribution.

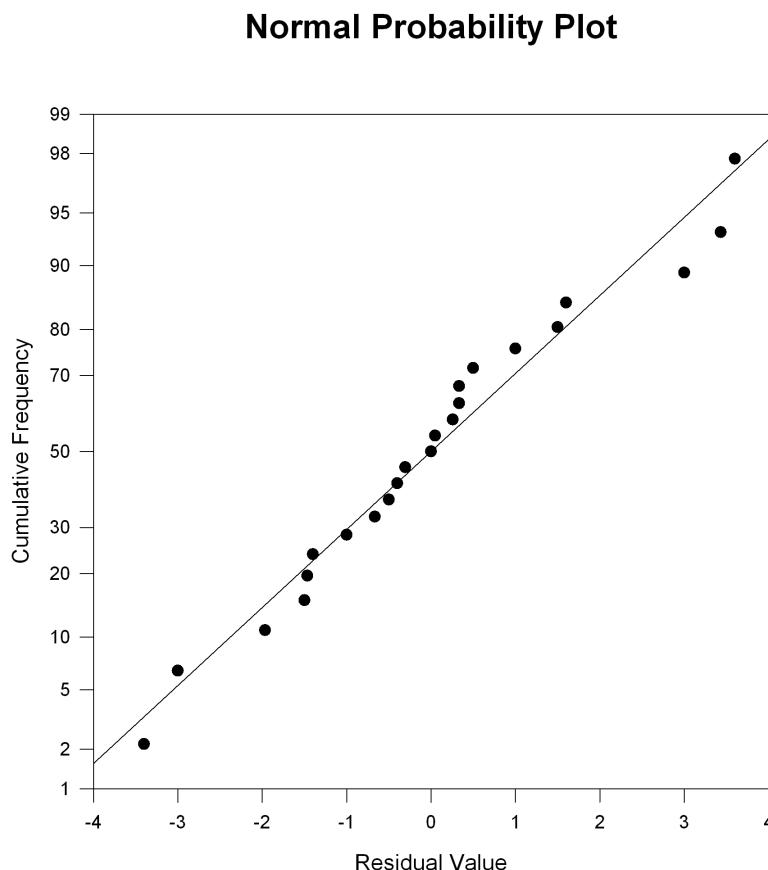
3. In the **Results Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the One Way ANOVA results.



4. Select the type of graph you want to create from the **Graph Type** list, then click **OK**, or double-click the desired graph in the list.

The selected graph appears in a [graph window](#).



Two Way Analysis of Variance (ANOVA)

Use a Two Way or Two Factor ANOVA (analysis of variance) when:

- You want to see if two or more different experimental groups are affected by two different factors which may or may not interact.
- Samples are drawn from normally distributed populations with equal variances.

If you want to consider the effects of only one factor on your experimental groups, use the [One Way ANOVA](#). If you are considering the effects of three factors on your experimental graphs, use the [Three Way ANOVA](#). SigmaPlot has no equivalent nonparametric two or three factor comparison for samples drawn from a non-normal population. If your data is non-normal, you can transform the data to make them comply better with the assumptions of analysis of variance using Transforms. If the sample size is large, and you want to do a nonparametric test, use the Rank Transform to convert the observations to ranks, then run a Two or Three Way ANOVA on the ranks.

About the Two Way ANOVA

In a two way or two factor analysis of variance, there are two experimental factors which are varied for each experimental group. A two factor design is used to test for differences between samples grouped according to the levels of each factor and for interactions between the factors.

A two factor analysis of variance tests three hypotheses:

- There is no difference among the levels of the first factor.
- There is no difference among the levels of the second factor.
- There is no interaction between the factors; for example, if there is any difference among groups within one factor, the differences are the same regardless of the second factor level.

Two Way ANOVA is a parametric test that assumes that all the samples were drawn from normally distributed populations with the same variances.

Performing a Two Way ANOVA

To perform a Two Way ANOVA:

1. [Enter or arrange your data appropriately in the worksheet.](#)
2. If desired, set **Two Way ANOVA** options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Compare Many Groups > Two Way ANOVA
5. Run the test.
6. [Generate report graphs.](#)

Arranging Two Way ANOVA Data

The Two Way ANOVA tests for differences between samples grouped according to the levels of each factor and the interactions between the factors.

For example, in an analysis of the effect of gender on the action of two different drugs, gender and drug are the factors, male and female are the levels of the gender factor, drug types are the levels for the drug factor, and the different combinations of the levels (gender and drug) are the groups, or cells.

Table 3: How to Arrange Two Way ANOVA Data

The factors are gender and drug, and the levels are Male/Female and Drug A/Drug B.

Gender	Drug
--------	------

	Drug A	Drug B
Male	3.83.23.5	1.51.82.2
Female	5.14.95.5	5.96.16.6

If your data is missing data points or even whole cells, SigmaPlot detects this and provides the correct solutions. For more information, see [Missing Data and Empty Cells](#) on page 171.

Indexing Raw Data for a Two-Way ANOVA

The Two-Way ANOVA test requires that the data be entered as indexed data. If your data is in a raw format, you can use a transform to convert it into an indexed format and then run the ANOVA.

In any Two-Way ANOVA, there are two factors, each divided into a number of levels. For example, Gender could be one factor with two levels: male and female. Drug Treatment could be another factor with three levels: Drug A, Drug B, Drug C.

Each combination of two levels, one from each factor, is called a *cell*. For example, all of the data measured for males receiving Drug A would be a cell. When the data for each cell is written into a column of the worksheet, this is known as a "raw data format" for Two-Way ANOVA. The number of columns equals the number of cells. Since each column gives the data for combining two factor levels, then the title of each column uses the names of the two levels.

The example above is a worksheet containing raw data for a Two-Way ANOVA. Note that the title of each column is composed of two names separated by a hyphen. The names refer to levels from different factors. There are six columns, and so there are six cells in the ANOVA.

To convert this data to Indexed format:

1. Click the **Analysis** tab.
 2. In the **Transforms** group, from the **Statistical** drop-down list, select:
Indexed > Two-Way
The **Select Data** panel of the Test Wizard appears.
 3. Select **column 7 (or First Empty** from the **Data for Output** drop-down list) as the **Output:** column.
 4. Select the first six columns for the input groups (this appears as **Group:** in the **Selected Columns** list).
i **Tip:** You can either select the columns from the worksheet, or you can select each column individually from the **Data for Group** drop-down list.
 5. Click **Finish**.
- The data appears as indexed data in columns 7 through 9.

Missing Data and Empty Cells Data

Ideally, the data for a Two Way ANOVA should be completely balanced. For example, each group or cell in the experiment has the same number of observations and there are no missing data; however, SigmaPlot properly handles all occurrences of missing and unbalanced data automatically.

Missing Data Points

If there are missing values, SigmaPlot automatically handles the missing data by using a general linear model approach. This approach constructs hypothesis tests using the marginal sums of squares (also commonly called the Type III or adjusted sums of squares).

Table 4: Data for a Two Way ANOVA with a Missing Value in the Male/Drug A Cell

A general linear model approach is used in these situations.

Gender	Drug	
	Drug A	Drug B
Male	3.8–3.5	1.51.82.2

Female

5.14.95.5

5.96.16.6

Empty Cells

When there is an empty cell, for example, there are no observations for a combination of two factor levels, SigmaPlot stops and suggests either analysis of the data using a two way design with the added assumption of no interaction between the factors, or a [One Way ANOVA](#).

Table 5: Data for a Two Way ANOVA with a Missing Data Cell (Male/Drug A)

You can use either one factor analysis or assume no interaction between factors.

Gender	Drug	
	Drug A	Drug B
Male	—	1.51.82.2
Female	5.14.95.5	5.96.16.6

Assumption of no interaction analyzes the main effects of each treatment separately.

 **DANGER:** It can be dangerous to assume there is no interaction between the two factors in a Two Way ANOVA. Under some circumstances, this assumption can lead to a meaningless analysis, particularly if you are interested in studying the interaction effect.

If you treat the problem as a One Way ANOVA, each cell in the table is treated as a different level of a single experimental factor. This approach is the most conservative analysis because it requires no additional assumptions about the nature of the data or experimental design.

Connected versus Disconnected Data

The no interaction assumption does not always permit a two factor analysis when there is more than one empty cell. The non-empty cells must be *geometrically connected* in order to do the computation. You cannot perform Two Way ANOVAs on disconnected data.

Arrange data in a two-dimensional grid, where you can draw a series of straight vertical and horizontal lines connecting all occupied cells, without changing direction in an empty cell, is guaranteed to be connected.

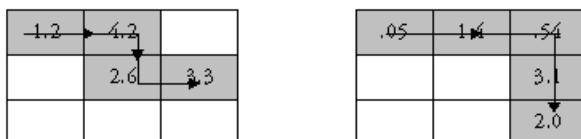


Figure 46: Example of Drawing Straight Horizontal and Vertical Lines Through Connected Data

It is important to note that failure to meet the above requirement does not imply that the data is disconnected. The data in the table below, for example, is connected.

1.2				8.3
	2.4			6.2
		5.8		1.0
			4.8	.98

Figure 47: Example of Connected Data that You Can't Draw a Series of Straight Vertical and Horizontal Lines Through

SigmaPlot automatically checks for this condition. If disconnected data is encountered during a Two Way ANOVA, SigmaPlot suggests treatment of the problem as a [One Way ANOVA](#).

Table 6: Disconnected Data

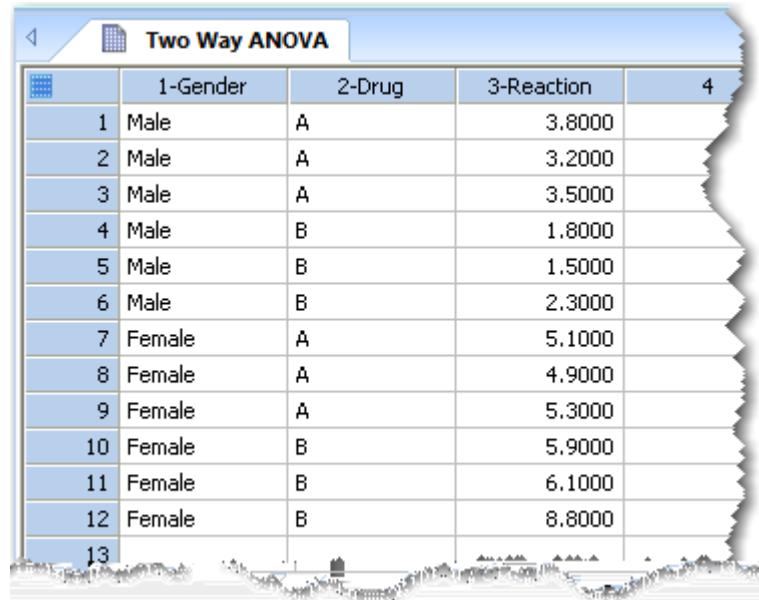
Gender	Drug	
	Drug A	Drug B
Male	—	1.51.82.2
Female	5.14.95.5	—

Because this data is not geometrically connected (the data shares no factor levels in common) a two way ANOVA cannot be performed, even assuming no interaction.

Entering Worksheet Data

A Two Way ANOVA can only be performed on two factor indexed data. Two factor indexed data is placed in three columns; a data point indexed two ways consists of the first factor in one column, the second factor in a second column, and the data point in a third column.

Column 1 is the first factor index, column 2 is the second factor index, and column 3 is the data.



	1-Gender	2-Drug	3-Reaction	4
1	Male	A	3.8000	
2	Male	A	3.2000	
3	Male	A	3.5000	
4	Male	B	1.8000	
5	Male	B	1.5000	
6	Male	B	2.3000	
7	Female	A	5.1000	
8	Female	A	4.9000	
9	Female	A	5.3000	
10	Female	B	5.9000	
11	Female	B	6.1000	
12	Female	B	8.8000	
13				

Figure 48: Valid Data Formats for a Two Way ANOVA

Setting Two Way ANOVA Options

Use the Two Way ANOVA options to:

- Adjust the parameters of the test to relax or restrict the testing of your data for normality and equal variance.
- Display the statistics summary table for the data.
- Compute the power, or sensitivity, of the test.
- Enable multiple comparison testing.

To change Two Way ANOVA options:

1. If you are going to run the test after changing test options and want to select your data before you run the test, drag the pointer over the data.
2. Select **Two Way ANOVA** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.

- Click Current Test Options. The **Options for Two Way ANOVA** dialog box appears with three tabs:
 - Assumption Checking.** Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
 - Results.** Display the statistics summary for the data in the report and save residuals to a worksheet column.
 - Post Hoc Tests.** Compute the power or sensitivity of the test.
- Tip:** If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.

Options settings are saved between SigmaPlot sessions.

- To continue the test, click **Run Test**. The **Select Data** panel of the Test Wizard appears.
- To accept the current settings and close the options dialog box, click **OK**.

Options Two Way ANOVA: Assumption Checking

Select the **Assumption Checking** tab from the options dialog box to view the options for **Normality** and **Equal Variance**. The normality assumption test checks for a normally distributed population. The equal variance assumption test checks the variability about the group means.

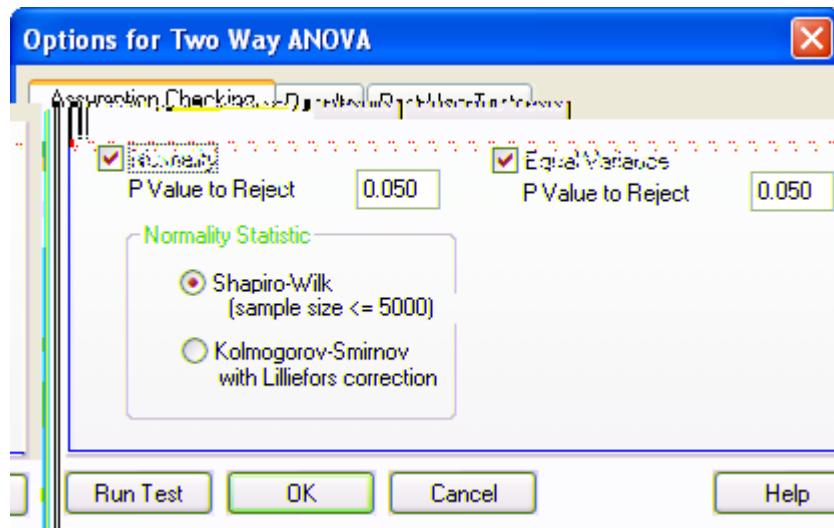


Figure 49: The Options for Two Way ANOVA Dialog Box Displaying the Assumption Checking Options

Normality Testing. SigmaPlot uses the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

Equal Variance Testing. SigmaPlot tests for equal variance by checking the variability about the group means.

P Values for Normality and Equal Variance. Enter the corresponding P value in the P Value to Reject box. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (the P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P value computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and equal variance, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality and/or equal variance, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.050 requires greater deviations from normality to flag the data as non-normal than a value of 0.100.



Restriction: There are extreme conditions of data distribution that these tests cannot take into account. For example, the Levene Median test fails to detect differences in variance of several orders of magnitude; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options Two Way ANOVA: Results

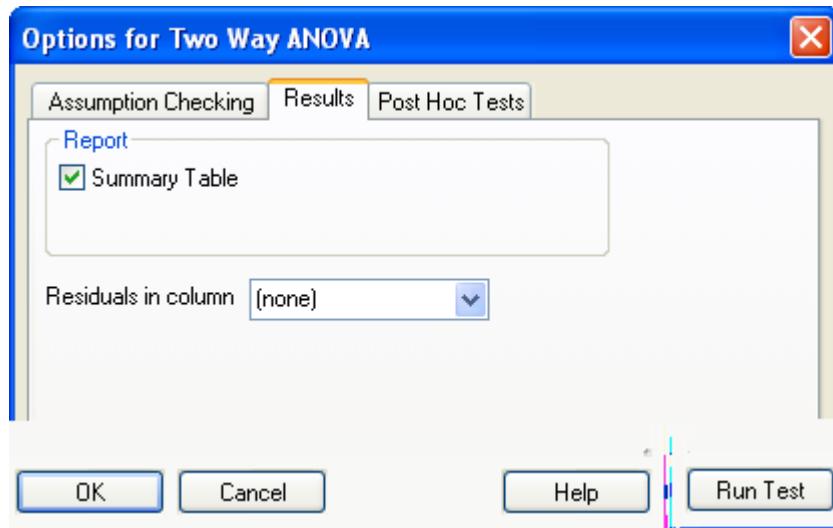
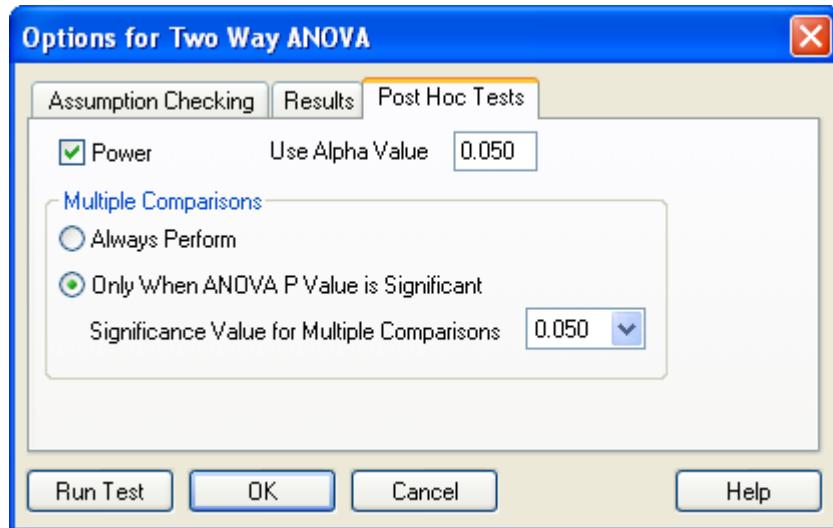


Figure 50: The Options for Two Way ANOVA Dialog Box Displaying the Summary Table, Confidence Intervals, and Residuals Options

Summary Table. Select **Summary Table** under **Report** to display the number of observations for a column or group, the number of missing values for a column or group, the average value for the column or group, the standard deviation of the column or group, and the standard error of the mean for the column or group.

Residuals in Column. The **Residuals in Column** drop-down list displays residuals in the report and to save the residuals of the test to the specified worksheet column. Edit the number or select a number from the drop-down list.

Options Two Way ANOVA: Post Hoc Tests



Power. The power or sensitivity of a test is the probability that the test will detect a difference between the groups if there is really a difference.

Use Alpha Value. Change the alpha value by editing the number in the Alpha Value box. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a

one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

Multiple Comparisons

Two Way ANOVAs test the hypothesis of no differences between the several treatment groups, but do not determine which groups are different, or the sizes of these differences. Use multiple comparisons to isolate these differences whenever a Two Way ANOVA detects a difference.

The P value used to determine if the ANOVA detects a difference is set in the **Report** tab of the **Options** dialog box. If the P value produced by the Two Way ANOVA is less than the P value specified in the box, a difference in the groups is detected and the multiple comparisons are performed.

- **Always Perform.** Select to perform multiple comparisons whether or not the Two Way ANOVA detects a difference.

Only When ANOVA P Value is Significant. Perform multiple comparisons only if the ANOVA detects a difference.

Significance Value for Multiple Comparisons. Select either .05 or .01. This value determines the likelihood of the multiple comparison being incorrect in concluding that there is a significant difference in the treatments. A value of .05 indicates that the multiple comparisons will detect a difference if there is less than 5% chance that the multiple comparison is incorrect in detecting a difference. A value of .01 indicates that the multiple comparisons will detect a difference if there is less than 1% chance that the multiple comparison is incorrect in detecting a difference.

 **Note:** If multiple comparisons are triggered, the Multiple Comparison Options dialog box appears after you pick your data from the worksheet and run the test, prompting you to choose a multiple comparison method.

Running a Two Way ANOVA

If you want to select your data before you run the test, drag the pointer over your data.

1. Click the **Analysis** tab.
2. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Compare Many Groups > Two Way ANOVA

The **Select Data** panel of the Test Wizard appears.

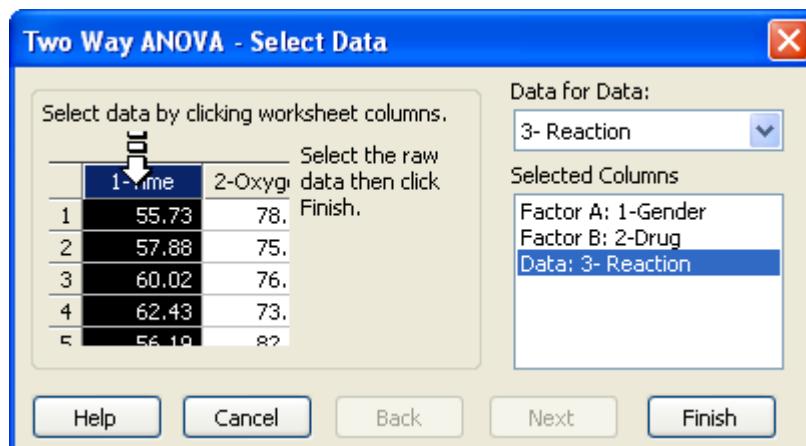


Figure 51: The Select Data Panel of the Test Wizard Prompting You to Select Data Columns

3. If the P value for multiple comparisons is significant, or you selected to always perform multiple comparisons, the **Multiple Comparisons Options** dialog box appears prompting you to select a multiple comparison method.

4. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the Data drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The number or title of selected columns appear in each row. You are prompted to pick a minimum three worksheet columns.

5. If the P value for multiple comparisons is significant, or you selected to always perform multiple comparisons, the **Multiple Comparisons Options** dialog box appears prompting you to select a multiple comparison method.
6. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the Selected Columns list.
7. Click **Finish** to perform the Two Way ANOVA.

The Two Way ANOVA report appears if you:

- Elected to test for normality and equal variance, and your data passes both tests.
 - Your data has no missing data points, cells, or is not otherwise unbalanced.
 - Selected not perform multiple comparisons, or if you selected to run multiple comparisons only when the P value is significant, and the P value is not significant.
8. If the P value for multiple comparisons is significant, or you selected to always perform multiple comparisons, the **Multiple Comparisons Options** dialog box appears prompting you to select a multiple comparison method.
 9. If you elected to test for normality and equal variance, and your data fails either test, either continue or transform your data, then perform the Two Way ANOVA on the transformed data.
 - If your data is missing data points, missing cells, or is otherwise unbalanced, you are prompted to perform the appropriate procedure.
 - If you are missing data points, but still have at least one observation in each cell, SigmaPlot automatically proceeds with the Two Way ANOVA using a general linear model.
 - If you are missing a cell, but the data is connected, you can proceed by either performing a two way analysis assuming no interaction between the factor, or converting the problem into a one way design with each non-empty cell a different level of a single factor.
 - If your data is not geometrically connected, you cannot perform a Two Way ANOVA. Either treat the problem as a **One Way ANOVA**, or cancel the test.
 10. If the P value for multiple comparisons is significant, or you selected to always perform multiple comparisons, the **Multiple Comparisons Options** dialog box appears prompting you to select a multiple comparison method.

Multiple Comparison Options for a Two Way ANOVA

If you selected to run multiple comparisons only when the P value is significant, and the ANOVA produces a P value, for either of the two factors or the interaction between the two factors, equal to or less than the trigger P value, or you selected to always run multiple comparisons in the Options for Two Way ANOVA dialog box, the **Multiple Comparison Options** dialog box appears prompting you to specify a multiple comparison test.

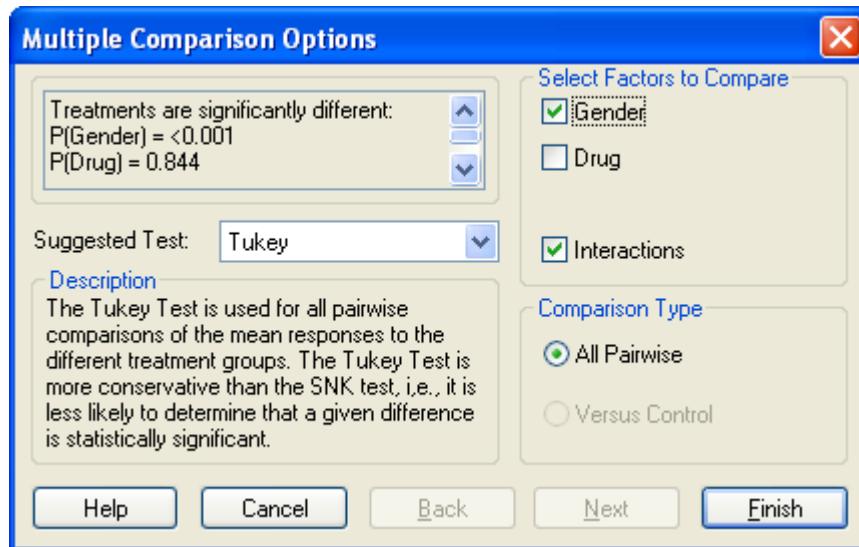


Figure 52: The Multiple Comparison Options Dialog Box for a Two Way ANOVA

This dialog box displays the P values for each of the two experimental factors and of the interaction between the two factors. Only the options with P values less than or equal to the value set in the **Options** dialog box are selected. You can disable multiple comparison testing for a factor by clicking the selected option. If no factor is selected, multiple comparison results are not reported.

There are seven multiple comparison tests to choose from for the Two Way ANOVA. You can choose to perform the

- **Holm-Sidak test.**
- **Tukey Test.**
- **Student-Newman-Keuls Test.**
- **Bonferroni t-test.**
- **Fisher's LSD.**
- **Dunnet's Test.**
- **Duncan's Multiple Range Test.**

The **Tukey** and **Student-Newman-Keuls** tests are recommended for determining the difference among all treatments. If you have only a few treatments, you may want to select the simpler **Bonferroni t-test**.

The **Dunnett's test** is recommended for determining the differences between the experimental treatments and a control group. If you have only a few treatments or observations, you can select the simpler **Bonferroni t-test**.

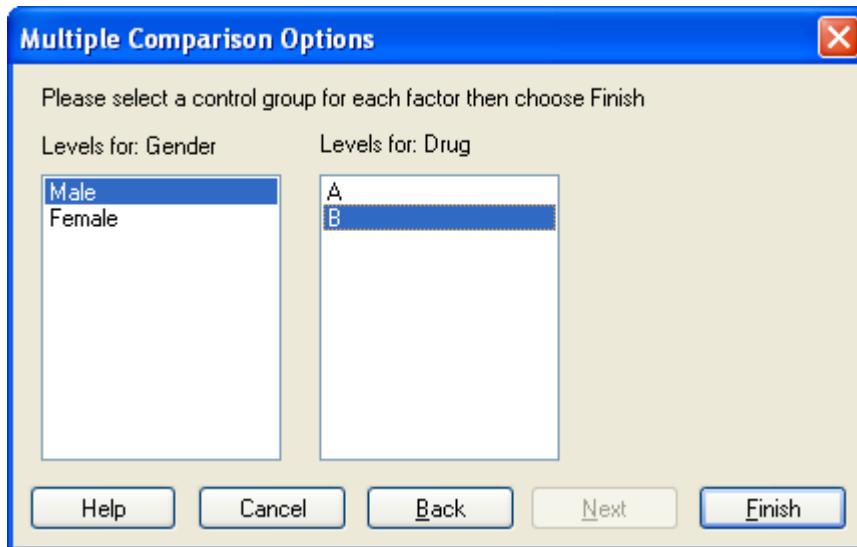


Figure 53: The Multiple Comparison Options Dialog Box Prompting You to Select Control Groups

 **Note:** In both cases the Bonferroni t-test is most sensitive with a small number of groups. Dunnett's test is not available if you have fewer than six observations.

There are two types of multiple comparison available for the Two Way ANOVA. The types of comparison you can make depends on the selected multiple comparison test. For more information, see [Two Way Analysis of Variance \(ANOVA\)](#) on page 89.

- All pairwise comparisons test the difference between each treatment or level within the two factors separately (for example, among the different rows and columns of the data table)
- Multiple comparisons versus a control test the difference between all the different combinations of each factors (for example, all the cells in the data table)

When comparing the two factors separately, the levels within one factor are compared among themselves without regard to the second factor, and vice versa. These results should be used when the interaction is not statistically significant.

When the interaction is statistically significant, interpreting multiple comparisons among different levels of each experimental factor may not be meaningful. SigmaPlot also suggests performing a multiple comparison between all the cells.

The result of all comparisons is a listing of the similar and different group pairs, for example, those groups that are and are not detectably different from each other. Because no statistical test eliminates uncertainty, multiple comparison procedures sometimes produce ambiguous groupings.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the **Options** dialog box.

You can also set the number of decimal places to display the **Options** dialog box.

If There Were Missing Data Cells

If your data contained missing values but no empty cells, the report indicates the results were computed using a general linear model.

If your data contained empty cells, you either analyzed the problem assuming either no interaction or treated the problem as a [One Way ANOVA](#).

- If you choose no interactions, no statistics for factor interaction are calculated.
- If you performed a One Way ANOVA, the results shown are identical to [One Way ANOVA results](#).

Dependent Variable

This is the data column title of the indexed worksheet data you are analyzing with the Two Way ANOVA. Determining if the values in this column are affected by the different factor levels is the objective of the Two Way ANOVA.

Normality Test

Normality test results display whether the data passed or failed the test of the assumption that they were drawn from a normal population and the P value calculated by the test. Normally distributed source populations are required for all parametric tests.

This result appears if you enabled normality testing in the Two Way ANOVA Options dialog box.

Equal Variance Test

Equal Variance test results display whether or not the data passed or failed the test of the assumption that the samples were drawn from populations with the same variance and the P value calculated by the test. Equal variance of the source population is assumed for all parametric tests.

This result appears if you enabled equal variance testing in the Two Way ANOVA Options dialog box.

ANOVA Table

The ANOVA table lists the results of the Two Way ANOVA.

-  **Tip:** When there are missing data, the best estimate of these values is automatically calculated using a general linear model.

DF (Degrees of Freedom)

Degrees of freedom represent the number of groups in each factor and the sample size, which affects the sensitivity of the ANOVA.

- The degrees of freedom for each factor is a measure of the number of levels in each factor.
- The interaction degrees of freedom is a measure of the total number of cells.
- The error degrees of freedom (sometimes called the residual or within groups degrees of freedom) is a measure of the sample size after accounting for the factors and interaction.
- The total degrees of freedom is a measure of the total sample size.

SS (Sum of Squares)

The sum of squares is a measure of variability associated with each element in the ANOVA data table.

Power

The power, or sensitivity, of a Two Way ANOVA is the probability that the test will detect the observed difference among the groups if there really is a difference. The closer the power is to 1, the more sensitive the test. The power for the comparison of the groups within the two factors and the power for the comparison of the interactions are all displayed. These results are set in the Options for Two Way ANOVA dialog box.

ANOVA power is affected by the sample sizes, the number of groups being compared, the chance of erroneously reporting a difference a (alpha), the observed differences of the group means, and the observed standard deviations of the samples.

Alpha

Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. An α error also is called a Type I error (a Type I error is when you reject the hypothesis of no effect when this hypothesis is true).

The α value is set in the Options for Two Way ANOVA dialog box; the sug

You can conclude from "large" values of t that the difference of the two groups being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of erroneously concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The Difference of Means is a gauge of the size of the difference between the levels or cells being compared.

The degrees of freedom DF for the marginal comparisons are a measure of the number of groups (levels) within the factor being compared. The degrees of freedom when comparing all cells is a measure of the sample size after accounting for the factors and interaction. This is the same as the error or residual degrees of freedom.

Tukey, Student-Newman-Keuls, Fisher LSD, Duncan's, and Dunnett's Test Results

The Tukey, Student-Newman-Keuls (SNK), Fisher LSD, and Duncan's tests are all pairwise comparisons of every combination of group pairs. While the Tukey Fisher LSD, and Duncan's can be used to compare a control group to other groups, they are not recommended for this type of comparison.

Dunnett's test only compares a control group to all other groups. All tests compute the q test statistic, the number of means spanned in the comparison p , and display whether or not $P < 0.05$ for that pair comparison.

You can conclude from "large" values of q that the difference of the two groups being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of being incorrect in concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

p is the parameter used when computing q . The larger the p , the larger q needs to be to indicate a significant difference. p is an indication of the differences in the ranks of the group means being compared. Groups means are ranked in order from largest to smallest, and p is the number of means spanned in the comparison. For example, when comparing four means, comparing the largest to the smallest $p = 4$, and when comparing the second smallest to the smallest $p = 2$.

If a group is found to be not significantly different than another group, all groups with p ranks in between the p ranks of the two groups that are not different are also assumed not to be significantly different, and a result of DNT (Do Not Test) appears for those comparisons.

The Difference of Means is a gauge of the size of the difference between the groups or cells being compared.

The degrees of freedom DF for the marginal comparisons are a measure of the number of groups (levels) within the factor being compared. The degrees of freedom when comparing all cells is a measure of the sample size after accounting for the factors and interaction (this is the same as the error or residual degrees of freedom).

Performing a One Way ANOVA on Two Way ANOVA Data

When your data is missing too many observations to perform a valid Two Way ANOVA, you can still analyze your data using a One Way ANOVA.

To perform a One Way ANOVA:

1. Click the **Analysis** tab.
2. In the **Transforms** group, from the **Statistical** drop-down list, select:
Unindex > Two Way
3. Select the output columns, then run a One Way ANOVA.
4. Select your two way indexed data columns as the input columns.
5. Select the output columns, then run a One Way ANOVA.

Interpreting Two Way ANOVA Results

A full Two Way ANOVA report displays an ANOVA table describing the variation associated with each factor and their interactions. This table displays the degrees of freedom, sum of squares, and mean squares for each of the elements in the data table, as well as the F statistics and the corresponding P values.

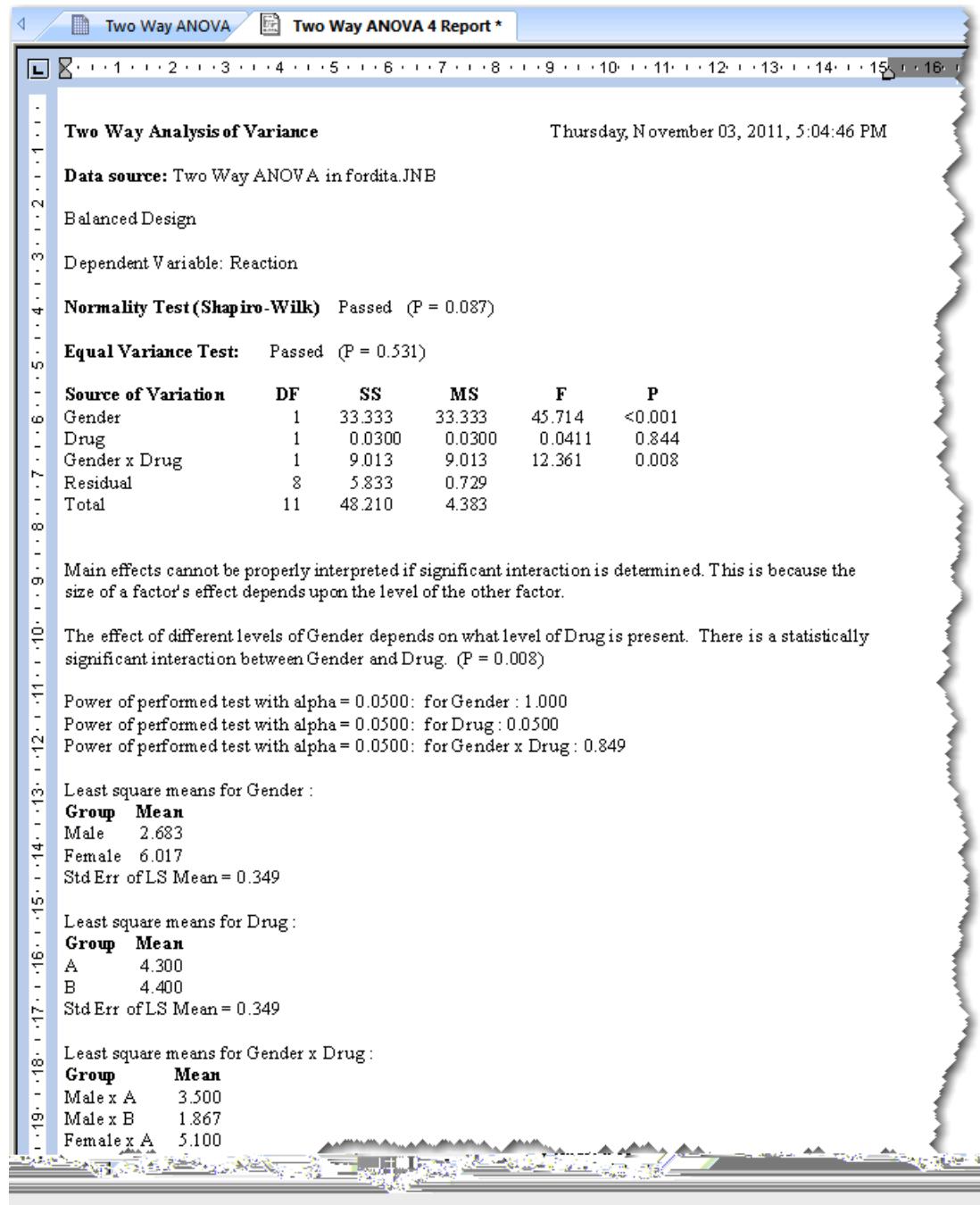


Figure 54: Two Way ANOVA Report

Summary tables of least square means for each factor and for both factors together can also be generated. This result and additional results are enabled in the **Options for Two Way ANOVA** dialog box. Select an option to enable or disable it. All options are saved between SigmaPlot sessions.

You can also generate tables of multiple comparisons. Multiple Comparison results are also specified in the **Options for Two Way ANOVA** dialog box. The tests used in the multiple comparisons are selected in the **Multiple Comparisons Options** dialog box.

Two Way ANOVA Report Graphs

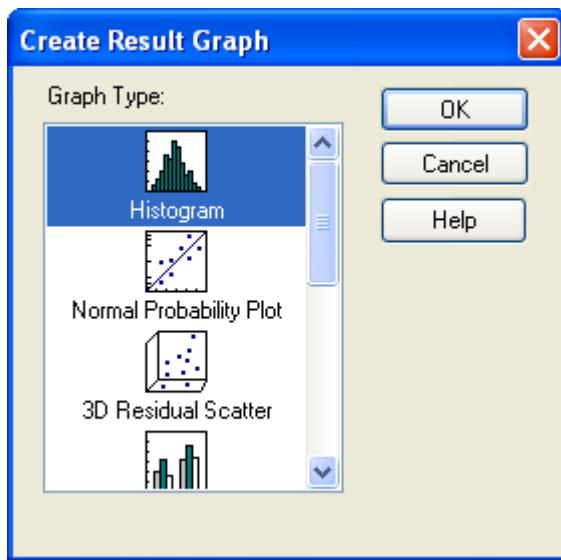
You can generate up to seven graphs using the results from a Two Way ANOVA. They include a:

- **Histogram of the residuals.**
- **Normal probability plot of the residuals.**
- **3D plot of the residuals.**
- **Grouped bar chart of the column means.**
- **3D category scatter plot.**
- **Multiple comparison graphs.**
- **Profile plots.**

How to Create a Two Way ANOVA Report Graph

1. Select the **Two Way ANOVA** test report.
2. Click the **Report** tab.
3. In the **Results Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the Two Way ANOVA results.



4. Select the type of graph you want to create from the **Graph Type** list.

- Click **OK**, or double-click the desired graph in the list.

The selected graph appears in a [graph window](#).

Histogram

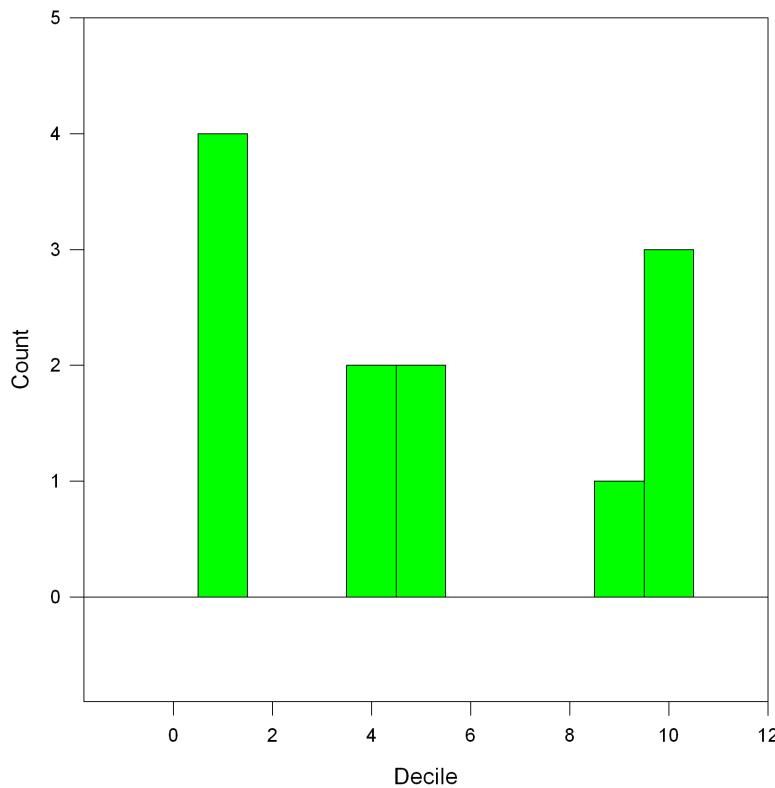


Figure 55: A Histogram of the Two Way ANOVA

Three Way Analysis of Variance (ANOVA)

Use a Three Way or three factor ANOVA (analysis of variance) when:

- You want to see if two or more different experimental groups are affected by three different factors which may or may not interact.
- Samples are drawn from normally distributed populations with equal variances.

To consider the effects of only one or two factors on your experimental groups, use a [One Way ANOVA](#). You can also use a [Two Way ANOVA](#). SigmaPlot has no equivalent nonparametric three factor comparison for samples drawn from a non-normal population. If your data is non-normal, you can transform the data to make them comply better with the assumptions of analysis of variance using Transforms. If the sample size is large, and you want to do a nonparametric test, use the Rank Transform to convert the observations to ranks, then run a Three Way ANOVA on the ranks.

About the Three Way ANOVA

In a three way or three factor analysis of variance, there are three experimental *factors* which are varied for each experimental group. A three factor design is used to test for differences between samples grouped according to the *levels* of each factor and for *interactions* between the factors.

A three factor analysis of variance tests four hypotheses:

- There is no difference among the levels of the first factor.
- There is no difference among the levels of the second factor.

- There is no difference among the levels of the third factor.
- There is no interaction between the factors; for example, if there is any difference among groups within one factor, the differences are the same regardless of the second and third factor levels.

Three Way ANOVA is a parametric test that assumes that all the samples were drawn from normally distributed populations with the same variances.

Performing a Three Way ANOVA

To perform a Three Way ANOVA:

1. [Enter or arrange your data appropriately in the worksheet.](#)
2. If desired, [set the Three Way ANOVA options.](#)
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Compare Many Groups > Three Way ANOVA
5. Run the test.
6. [Generate report graphs.](#)

Arranging Three Way ANOVA Data

The Three Way ANOVA tests for differences between samples grouped according to the levels of each factor and the interactions between the factors.

For example, in an analysis of the effect of gender on the action of two different drugs over a different periods of time, gender, drugs, and time period are the factors, male and female are the levels of the gender factor, drug types are the levels for the drug factor, days are the levels of the time period factor, and the different combinations of the levels (gender, drug, and time period) are the groups, or cells.

Table 7: Data for a Three Way ANOVA

Gender		Male						Female					
Drug		Drug A			Drug B			Drug A			Drug B		
Time Period		Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Reaction		11325	21426	31527	41628	51729	61830	71931	82032	92133	102234	112335	122436

The factors are gender, drug, and time period. The levels are Male/Female, Drug A/Drug B, and Day 1, 2, and 3.

If your data is missing data points or even whole cells, SigmaPlot detects this and provides the correct solutions. For more information, see [Missing Data and Empty Cells Data](#) on page 106.

	Three Way ANOVA Data			
	1-Factor 1	2-Factor 2	3-Factor 3	4-Data
1	Level A	Level C	Level E	22.2570
2	Level A	Level C	Level E	14.2345
3	Level A	Level C	Level E	18.3640
4	Level A	Level C	Level E	41.7464
5	Level A	Level C	Level F	84.7543
6	Level A	Level C	Level F	23.5897
7	Level A	Level C	Level F	88.6524
8	Level A	Level C	Level F	43.6743
9	Level A	Level D	Level E	32.9764
10	Level A	Level D	Level E	12.7357
11	Level A	Level D	Level E	87.8367
12	Level A	Level D	Level E	38.9126
13	Level A	Level D	Level F	91.7286
14	Level A	Level D	Level F	48.6385
15	Level A	Level D	Level F	54.8642
16	Level A	Level D	Level F	32.9189
17	Level B	Level C	Level E	56.4532
			Level E	87.8578

Figure 56: Valid Data Formats for a Three Way ANOVA

Column 1 is the first factor index, column 2 is the second factor index, column 3 is the third factor index, and column 4 is the data.

Missing Data and Empty Cells Data

Ideally, the data for a Three Way ANOVA should be completely balanced. For example, each group or cell in the experiment has the same number of observations and there are no missing data; however, SigmaPlot properly handles all occurrences of missing and unbalanced data automatically.

Missing Data Points. If there are missing values, SigmaPlot automatically handles the missing data by using a *general linear model* approach. This approach constructs hypothesis tests using the marginal sums of squares (also commonly called the Type III or adjusted sums of squares).

Table 8: Data for a Three Way ANOVA with a Missing Value in the Male, Drug A, Day 1 Cell

Gender	Male						Female					
	Drug	Drug A			Drug B			Drug A	Drug B			
Day		Day 1	Day 2	Day 3	Day 1	Day 2	Day 3		Day 1	Day 2	Day 3	Day 1
Reaction	1--25	21426	31527	41628	51729	61830	71931	82032	92133	102234	112335	122436

Use a general linear model approach in these situations.

Empty Cells. When there is an empty cell, for example, there are no observations for a combination of three factor levels, a dialog box appears asking you if you want to analyze the data using a two way or a one way design. If you select a two way design, SigmaPlot attempts to analyze your data using two interactions. If there are no observations with two interactions, SigmaPlot runs a [One Way ANOVA](#).

If you treat the problem as a [Two Way ANOVA](#), a dialog box appears prompting you to remove one of the factors. Select the factor you want to remove, then click OK. The Two Way ANOVA is performed.

If you treat the problem as a One Way ANOVA, each cell in the table is treated as a single experimental factor. This approach is the most conservative analysis because it requires no additional assumptions about the nature of the data or experimental design.

Table 9: Data for a Three Way ANOVA with a Missing Cell (Male/Drug A, Day 1)

You can use either a two factor analysis or assume no interaction between factors.

Assumption of no interaction analyzes the main effects of each treatment separately.

DANGER: It can be dangerous to assume there is no interaction between the three factors in a Three Way ANOVA. Under some circumstances, this assumption can lead to a meaningless analysis, particularly if you are interested in studying the interaction effect.

Connected versus Disconnected Data

The no interaction assumption does not always permit a two factor analysis when there is more than one empty cell. The non-empty cells must be geometrically connected in order to do the computation. You cannot perform Three Way ANOVAs on disconnected data.

Data arranged in a two-dimensional grid, where you can draw a series of straight vertical and horizontal lines connecting all occupied cells, without changing direction in an empty cell, is guaranteed to be connected.

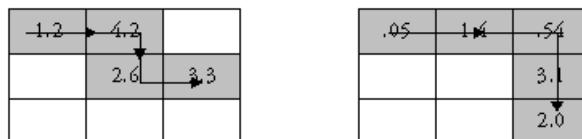


Figure 57: Example of Drawing Straight Horizontal and Vertical Lines through Connected Data

It is important to note that failure to meet the above requirement does not imply that the data is disconnected. The data in the table below, for example, is connected.

1.2				8.3
	2.4			6.2
		5.8		1.0
			4.8	.98

Figure 58: Example of Connected Data that You Can't Draw a Series of Straight Vertical and Horizontal Lines Through

SigmaPlot automatically checks for this condition. If disconnected data is encountered during a Three Way ANOVA, SigmaPlot suggests treatment of the problem as a Two Way ANOVA. If the disconnected data is still encountered during a Two Way ANOVA, a One Way ANOVA is performed.

Table 10: Disconnected Data

Gender	Male	Female
Drug	Drug A	Drug B

Day	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Reaction	—	—	—	41628	51729	61830	71931	82032	92133	—	—	—
	—	—	—	—	—	—	—	—	—	—	—	—

Because this data is not geometrically connected (they share no factor levels in common), a Three Way ANOVA cannot be performed, even assuming no interaction.

Entering Worksheet Data

A Three Way ANOVA can only be performed on three factor indexed data. Three factor indexed data is placed in four columns; a data point indexed three ways consists of the first factor in one column, the second factor in a second column, the third factor in a third column, and the data in a forth column.

Setting Three Way ANOVA Options

Use the Three Way ANOVA options to:

- Adjust the parameters of the test to relax or restrict the testing of your data for normality and equal variance.
- Include the statistics summary table for the data in the report, and save residuals to the worksheet.
- Compute the power, or sensitivity, of the test.
- Enable multiple comparison testing.

To set Three Way ANOVA options:

1. If you are going to run the test after changing test options and want to select your data before you run the test, drag the pointer over the data.
2. On the **Analysis** tab, in the **SigmaStat** group, click **Options**. The Options for Three Way ANOVA dialog box appears with three tabs:
 - **Assumption Checking.** Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
 - **Results.** Display the statistics summary for the data in the report and save residuals to a worksheet column.
 - **Post Hoc Tests.** Compute the power or sensitivity of the test.

Tip: If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.

Options settings are saved between SigmaPlot sessions.

3. To continue the test, click **Run Test**. Select Data panel of the Test Wizard appears.
4. To accept the current settings and close the options dialog box, click **OK**.

Options for Three Way ANOVA: Assumption Checking

Select the Assumption Checking tab from the options dialog box to view the Normality and Equal Variance options. The normality assumption test checks for a normally distributed population. The equal variance assumption test checks the variability about the group means.

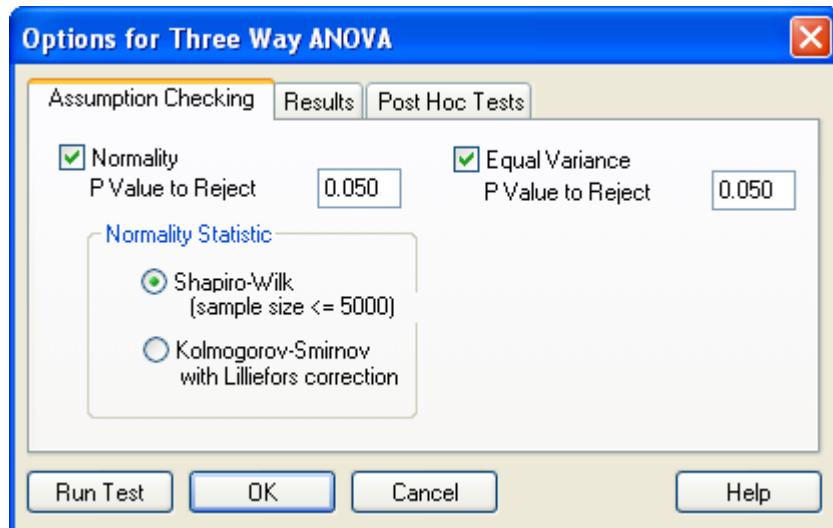


Figure 59: The Options for Three Way ANOVA Dialog Box Displaying the Assumption Checking Options

Normality Testing. SigmaPlot uses the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

Equal Variance Testing. SigmaPlot tests for equal variance by checking the variability about the group means.

P Values for Normality and Equal Variance. Type the corresponding P value in the P Value to Reject box. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (the P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P value computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and/or equal variance, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality and/or equal variance, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.050 requires greater deviations from normality to flag the data as non-normal than a value of 0.100.

 **Note:** There are extreme conditions of data distribution that these tests cannot take into account. For example, the Levene Median test fails to detect differences in variance of several orders of magnitude; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options for Three Way ANOVA: Results

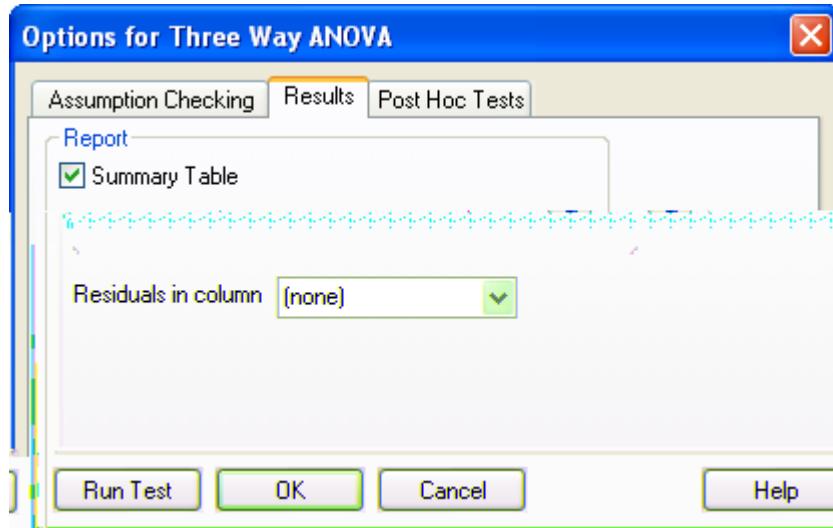


Figure 60: The Options for Three Way ANOVA Dialog Box Displaying the Summary Table and Residual Options

Summary Table. Select **Summary Table** under **Report** to display the number of observations for a column or group, the number of missing values for a column or group, the average value for the column or group, the standard deviation of the column or group, and the standard error of the mean for the column or group.

Residuals in Column. The **Residuals in Column** drop-down list displays residuals in the report and to save the residuals of the test to the specified worksheet column. Edit the number or select a number from the drop-down list.

Options for Three Way ANOVA: Post Hoc Tests

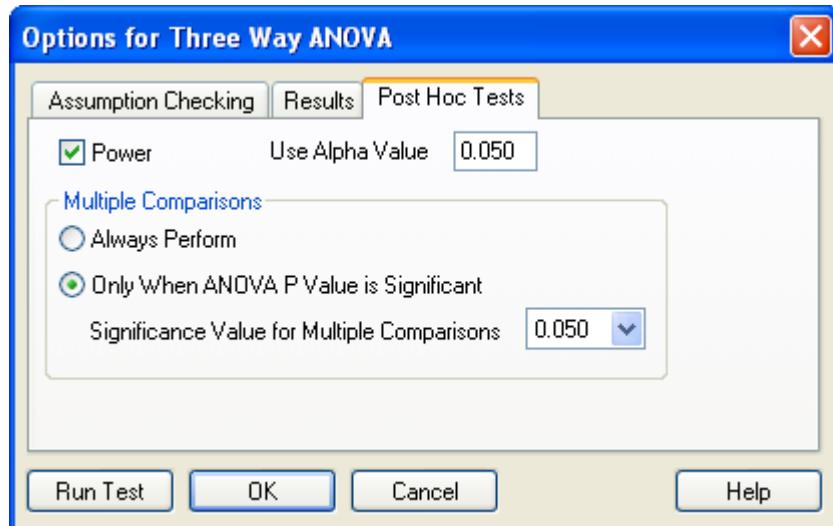


Figure 61: The Options for Three Way ANOVA Dialog Box Displaying the Power and Multiple Comparisons Options

Power. The power or sensitivity of a test is the probability that the test will detect a difference between the groups if there is really a difference.

Use Alpha Value. Change the alpha value by editing the number in the Alpha Value box. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

Multiple Comparisons

Three Way ANOVAs test the hypothesis of no differences between the several treatment groups, but do not determine which groups are different, or the sizes of these differences. Multiple comparisons isolate these differences whenever a Three Way ANOVA detects a difference.

The P value used to determine if the ANOVA detects a difference is set in the **Report** tab of the **Options** dialog box. If the P value produced by the Three Way ANOVA is less than the P value specified in the box, a difference in the groups is detected and the multiple comparisons are performed.

- **Always Perform.** Select to perform multiple comparisons whether or not the Two Way ANOVA detects a difference.

Only When ANOVA P Value is Significant. Perform multiple comparisons only if the ANOVA detects a difference.

Significant Multiple Comparison Value. Select either .05 or .10 from the Significance Value for Multiple Comparisons drop-down list. This value determines that the likelihood of the multiple comparison being incorrect in concluding that there is a significant difference in the treatments.

A value of .05 indicates that the multiple comparisons will detect a difference if there is less than 5% chance that the multiple comparison is incorrect in detecting a difference. A value of .10 indicates that the multiple comparisons will detect a difference if there is less than 10% chance that the multiple comparison is incorrect in detecting a difference.

 **Note:** If multiple comparisons are triggered, the Multiple Comparison Options dialog box appears after you pick your data from the worksheet and run the test, prompting you to choose a multiple comparison test.

Running a Three Way ANOVA

If you want to select your data before you run the test, drag the pointer over your data.

1. Click the **Analysis** tab.
2. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Compare Many Groups > Three Way ANOVA

The **Three Way ANOVA - Select Data** panel of the Test Wizard appears.

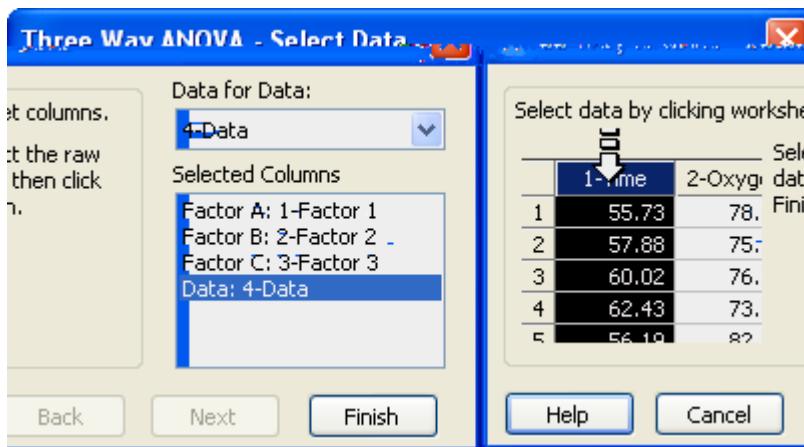


Figure 62: The Three Way ANOVA - Select Data Dialog Box

3. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The number or title of selected columns appear in each row. You are prompted to pick a minimum three worksheet columns.

4. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
5. Click **Finish** to perform the Three Way ANOVA. The Three Way ANOVA report appears if you:
 - Elected to test for normality and equal variance, and your data passes both tests.
 - Your data has no missing data points, cells, or is not otherwise unbalanced.
 - Selected not to perform multiple comparisons, or if you selected to run multiple comparisons only when the P value is significant, and the P value is not significant

If you elected to test for normality and equal variance, and your data fails either test, either continue or transform your data, then perform the Three Way ANOVA on the transformed data. If your data is missing data points, missing cells, or is otherwise unbalanced, you are prompted to perform the appropriate procedure.

Multiple Comparison Options for a Three Way ANOVA

If you enabled multiple comparisons in the Three Way ANOVA Options dialog box, and the ANOVA produces a P value, for either of the three factors or the interaction between the three factors, equal to or less than the trigger P value, the Multiple Comparison Options dialog box appears.

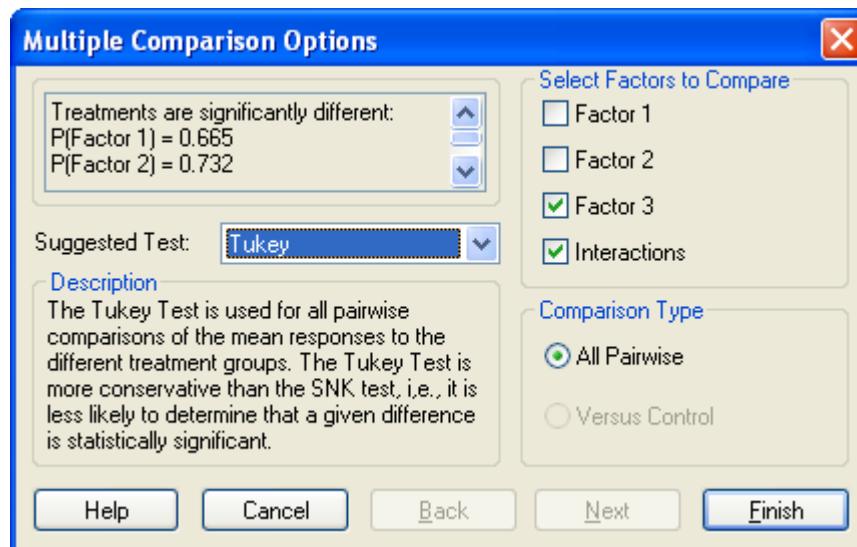


Figure 63: The Multiple Comparison Options Dialog Box for a Three Way ANOVA

This dialog box displays the P values for each of the experimental factors and of the interaction between the factors. Only the options with P values less than or equal to the value set in the Options dialog box are selected. You can disable multiple comparison testing for a factor by clicking the selected option. If no factor is selected, multiple comparison results are not reported.

There are seven multiple comparison tests to choose from for the Three Way ANOVA. You can choose to perform the:

- **Holm-Sidak test.**
- **Tukey Test.**
- **Student-Newman-Keuls Test.**
- **Bonferroni t-test.**
- **Fisher's LSD.**
- **Dunnet's Test.**
- **Duncan's Multiple Range Test.**

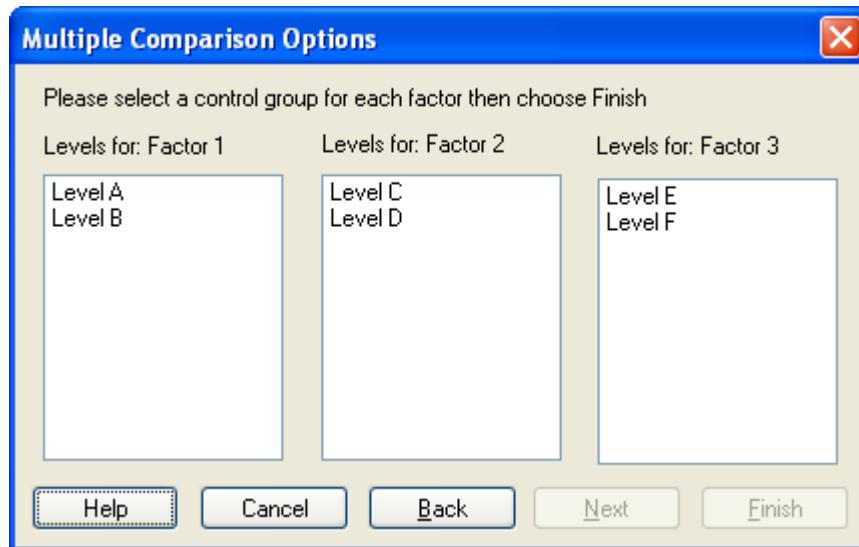


Figure 64: The Multiple Comparison Options Dialog Box Prompting You to Select a Control Group

There are two types of multiple comparison available for the Three Way ANOVA. The types of comparison you can make depends on the selected multiple comparison test.

- All pairwise comparisons test the difference between each treatment or level within the two factors separately (for example, among the different rows and columns of the data table).
- Multiple comparisons versus a control test the difference between all the different combinations of each factors (for example, all the cells in the data table).

All pairwise comparisons test the difference between each treatment or level within the two factors separately (for example, among the different rows and columns of the data table). Multiple comparisons versus a control test the difference between all the different combinations of each factors (for example, all the cells in the data table).

When comparing the two factors separately, the levels within one factor are compared among themselves without regard to the second factor, and vice versa. These results should be used when the interaction is not statistically significant.

When the interaction is statistically significant, interpreting multiple comparisons among different levels of each experimental factor may not be meaningful. SigmaPlot also suggests performing a multiple comparison between all the cells.

The result of both comparisons is a listing of the similar and different group pairs, for example, those groups that are and are not detectably different from each other. Because no statistical test eliminates uncertainty, multiple comparison procedures sometimes produce ambiguous groupings.

Interpreting Three Way ANOVA Results

A full Three Way ANOVA report displays an ANOVA table describing the variation associated with each factor and their interactions. This table displays the degrees of freedom, sum of squares, and mean squares for each of the elements in the data table, as well as the F statistics and the corresponding P values.

Summary tables of least square means for each factor and for all three factors together can also be generated. This result and additional results are enabled in the Options for Three Way ANOVA dialog box. Select an option to enable or disable it. All options are saved between SigmaPlot sessions.

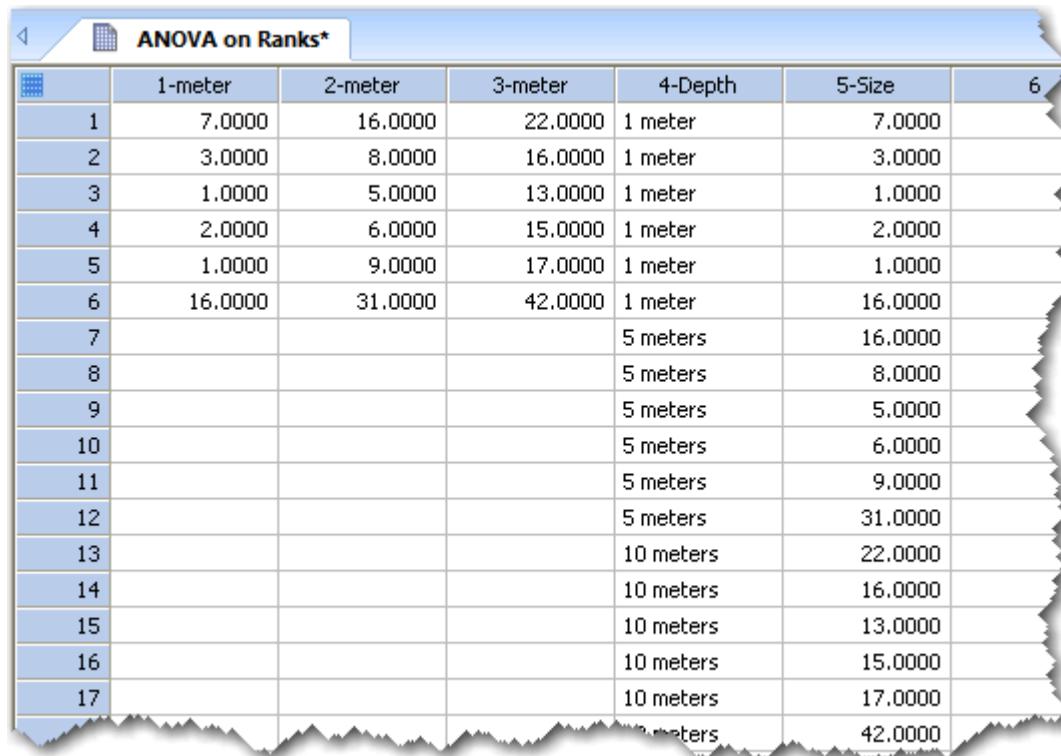
You can also generate tables of multiple comparisons. Multiple Comparison results are also specified in the Options for Three Way ANOVA dialog box. The tests used in the multiple comparisons are selected in the Multiple Comparisons Options dialog box.

Three Way ANOVA Report Graphs

You can generate up to four graphs using the `ThreeWayAnova` command.

Arranging ANOVA on Ranks Data

The format of the data to be tested can be raw data or indexed data. Raw data is placed in as many columns as there are groups, up to 640; each column contains the data for one group. Indexed data is placed in two worksheet columns with at least three treatments. If you have less than three treatments you should use the [Rank Sum Test](#).



	1-meter	2-meter	3-meter	4-Depth	5-Size	6
1	7.0000	16.0000	22.0000	1 meter	7.0000	
2	3.0000	8.0000	16.0000	1 meter	3.0000	
3	1.0000	5.0000	13.0000	1 meter	1.0000	
4	2.0000	6.0000	15.0000	1 meter	2.0000	
5	1.0000	9.0000	17.0000	1 meter	1.0000	
6	16.0000	31.0000	42.0000	1 meter	16.0000	
7				5 meters	16.0000	
8				5 meters	8.0000	
9				5 meters	5.0000	
10				5 meters	6.0000	
11				5 meters	9.0000	
12				5 meters	31.0000	
13				10 meters	22.0000	
14				10 meters	16.0000	
15				10 meters	13.0000	
16				10 meters	15.0000	
17				10 meters	17.0000	
				10 meters	42.0000	

Figure 65: Valid Data Formats for an ANOVA on Ranks

Columns 1 through 3 are arranged as raw data. Columns 4 and 5 are arranged as indexed data, with column 4 as the factor column and column 5 as the data column.

Setting the ANOVA on Ranks Options

Use the ANOVA on Ranks options to:

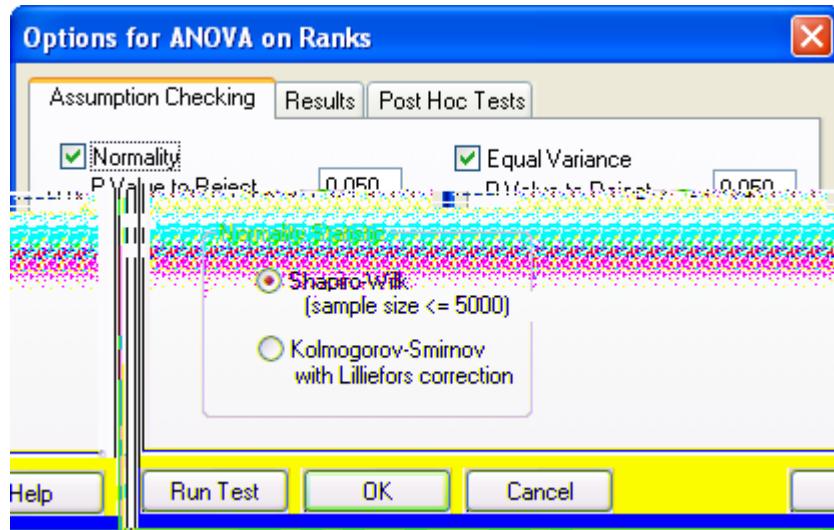
- Adjust the parameters of the test to relax or restrict the testing of your data for normality and equal variance.
- Enable multiple comparison testing.
- Display the summary table.

To change the ANOVA on Ranks options:

1. Select **ANOVA on Ranks** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.
2. Click **Current Test Options**. The **Options for ANOVA on Ranks** dialog box appears with three tabs:
 - **Assumption Checking**. Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
 - **Results**. Display the statistics summary for the data in the report.
 - **Post Hoc Test**. Compute the power or sensitivity of the test and enable multiple comparisons.
3. To continue the test, click **Run Test**.
4. To accept the current settings, click **OK**.

Options for ANOVA on Ranks: Assumption Checking

Click the Assumption Checking tab from the options dialog box to view the Normality and Equal Variance options. The normality assumption test checks for a normally distributed population. The equal variance assumption test checks the variability about the group means.



- **Normality.** SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.
- **Equal Variance.** SigmaPlot tests for equal variance by checking the variability about the group means.
- **P Values for Normality and Equal Variance.** Enter the corresponding P value in the **P Value to Reject** box. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (the P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P value computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and/or equal variance, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality and equal variance, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.050 requires greater deviations from normality to flag the data as non-normal than a value of 0.100.

 **Note:** There are extreme conditions of data distribution that these tests cannot take into account. For example, the Levene Median test fails to detect differences in variance of several orders of magnitude; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options for ANOVA on Ranks: Results

The **Summary Table** for a Rank Sum Test lists the medians, percentiles, and sample sizes N in the Rank Sum test report. If desired, change the percentile values by editing the boxes. The 25th and the 75th percentiles are the suggested percentiles.

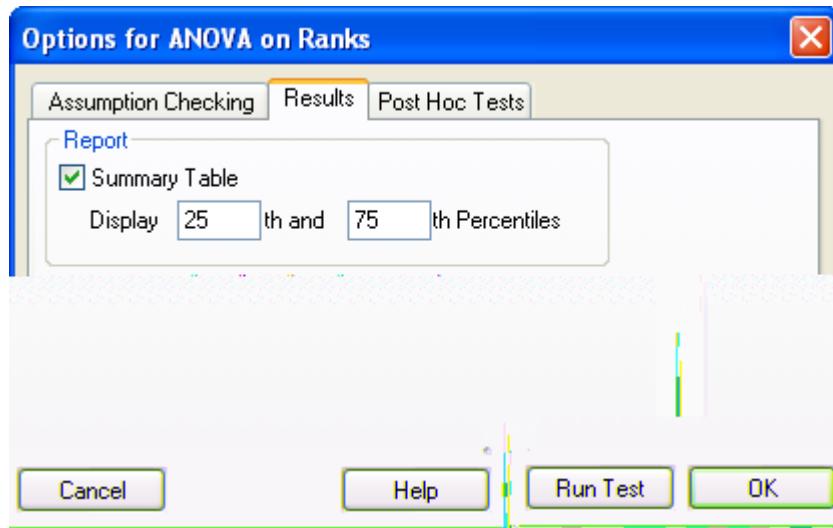
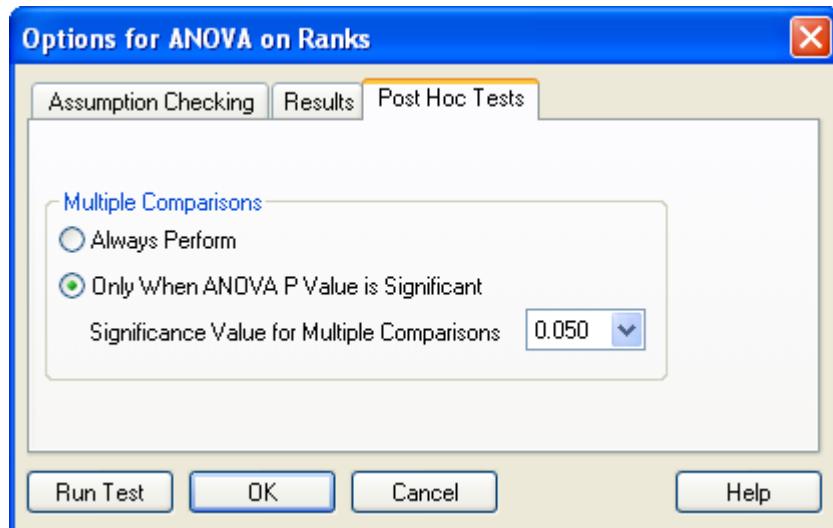


Figure 66: The Options for ANOVA on Ranks Dialog Box Displaying the Summary Table Option

Options for ANOVA on Ranks: Post Hoc Tests

Select the **Post Hoc Test** tab in the **Options** dialog box to view the multiple comparisons options. An ANOVA on Ranks tests the hypothesis of no differences between the several treatment groups, but does not determine which groups are different, or the size of these differences. Multiple comparisons isolate these differences.

The P value used to determine if the ANOVA detects a difference is set in the Report Options dialog box. If the P value produced by the ANOVA on Ranks is less than the P value specified in the box, a difference in the groups is detected and the multiple comparisons are performed.



Multiple Comparisons. You can choose to always perform multiple comparisons or to only perform multiple comparisons if the ANOVA on Ranks detects a difference.

- **Always Perform.** Select to perform multiple comparisons whether or not the ANOVA detects a difference.
- **Only When ANOVA P Value is Significant.** Select to perform multiple comparisons only if the ANOVA detects a difference.
- **Significance Value for Multiple Comparisons.** Select a value from the **Significance Value for Multiple Comparisons** drop-down list. This value determines the that the likelihood of the multiple comparison being incorrect in concluding that there is a significant difference in the treatments.

A value of .05 indicates that the multiple comparisons will detect a difference if there is less than 5% chance that the multiple comparison is incorrect in detecting a difference.

-  **Note:** If multiple comparisons are triggered, the **Multiple Comparison Options** dialog box appears after you pick your data from the worksheet and run the test, prompting you to choose a multiple comparison method.
-  **Attention:** Because no statistical test eliminates uncertainty, multiple comparison tests sometimes produce ambiguous groupings.

Running an ANOVA on Ranks

If you want to select your data before you run the test, drag the pointer over your data.

To run an ANOVA on Ranks:

1. Click the **Analysis** tab.
2. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Compare Many Groups > ANOVA on Ranks

The **ANOVA on Ranks - Data Format** panel of the Test Wizard appears prompting you to specify a data format.

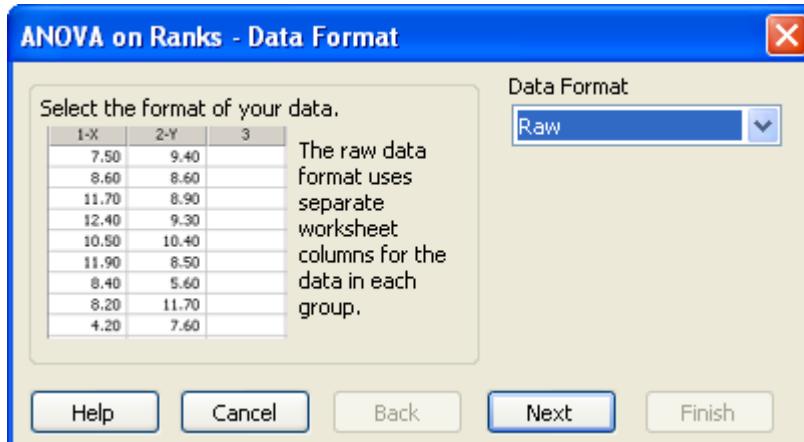


Figure 67: The ANOVA on Ranks - Data Format Panel of the Test Wizard Prompting You to Specify A Data Format

3. If the P value for multiple comparisons is significant, or you selected to always perform multiple comparisons, the **Multiple Comparisons Options** dialog box appears prompting you to select a multiple comparison method.
4. Select the appropriate data format from the **Data Format** drop-down list.
5. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

6. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list.

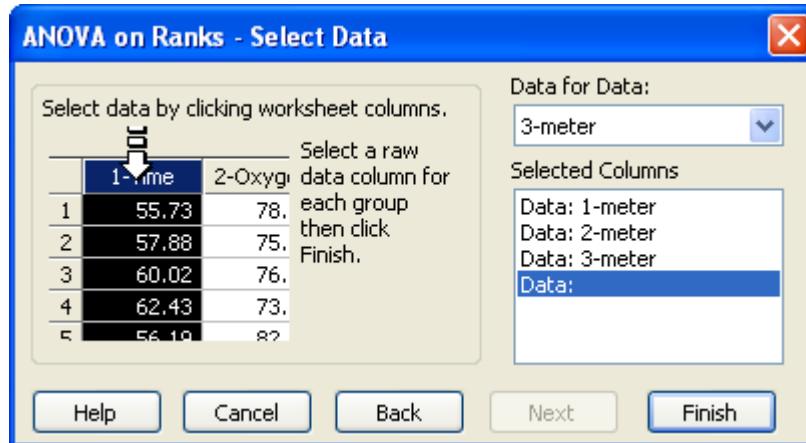


Figure 68: The Select Data Panel of the Test Wizard Prompting You to Select Data Columns

The number or title of selected columns appear in each row. You are prompted to pick a minimum of two and a maximum of 640 columns for raw data and two columns with at least three treatments are selected for indexed data. If you have less than three treatments, a message appears telling you to use the [Mann-Whitney Rank Sum Test](#) on page 66.

7. If the P value for multiple comparisons is significant, or you selected to always perform multiple comparisons, the [Multiple Comparisons Options](#) dialog box appears prompting you to select a multiple comparison method.
8. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the Selected Columns list.
9. If you elected to test for normality and equal variance, and your data fails either test, either continue or transform your data, then perform the Two Way ANOVA on the transformed data.
10. Click **Finish** to perform the ANOVA on Ranks. The **ANOVA on Ranks** report appears if you:
 - Elected to test for normality and equal variance, and your data passes both tests
 - Selected not perform multiple comparisons, or if you selected to run multiple comparisons only when the P value is significant, and the P value is not significant. For more information, see [Interpreting ANOVA on Ranks Results](#) on page 120.
11. If the P value for multiple comparisons is significant, or you selected to always perform multiple comparisons, the [Multiple Comparisons Options](#) dialog box appears prompting you to select a multiple comparison method.

Multiple Comparison Options for ANOVA on Ranks

If you selected to run multiple comparisons only when the P value is significant, and the ANOVA produces a P value, for either of the two factors or the interaction between the two factors, equal to or less than the trigger P value, or you selected to always run multiple comparisons in the Options for ANOVA on Ranks dialog box, the Multiple Comparison Options dialog box appears prompting you to specify a multiple comparison test.

This dialog box displays the P values for each of the two experimental factors and of the interaction between the two factors. Only the options with P values less than or equal to the value set in the Options dialog box are selected. You can disable multiple comparison testing for a factor by clicking the selected option. If no factor is selected, multiple comparison results are not reported.

There are four multiple comparison tests to choose from for the ANOVA on Ranks. You can choose to perform the:

- [Dunn's Test](#).
- [Dunnett's Test](#).
- [Tukey Test](#).

- **Student-Newman-Keuls Test.**

There are two types of multiple comparison available for the ANOVA on Ranks. The types of comparison you can make depends on the selected multiple comparison test.

- Multiple comparisons versus a control test the difference between all the different combinations of each factors (for example, all the cells in the data table).
- All pairwise comparisons test the difference between each treatment or level within the two factors separately (for example, among the different rows and columns of the data table)

Interpreting ANOVA on Ranks Results

The ANOVA on Ranks report displays the H statistic (corrected for ties) and the corresponding P value for H. The other results displayed in the report are enabled and disabled in the Options for ANOVA on Ranks dialog box.

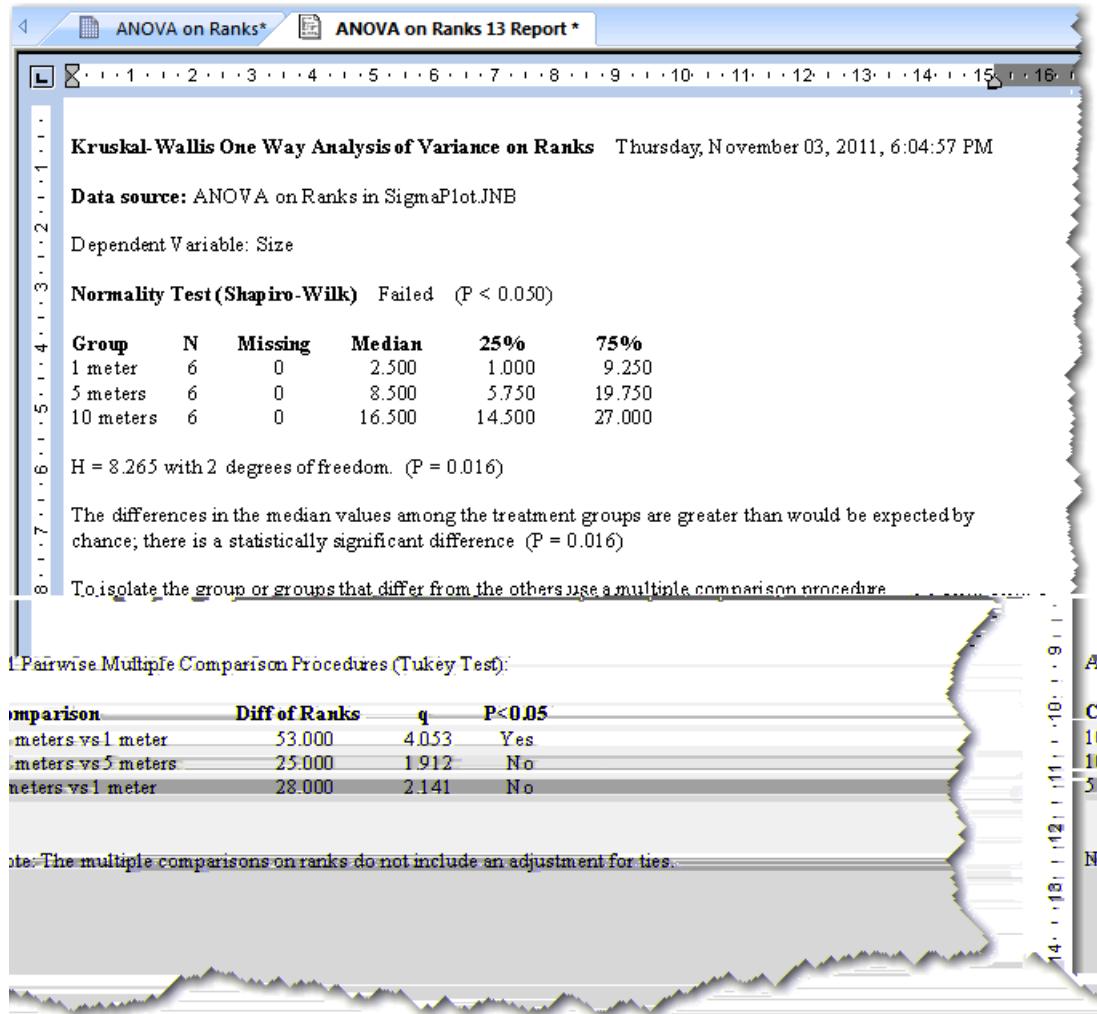


Figure 69: The ANOVA on Ranks Results Report

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the **Options** dialog box.

You can also set the number of decimal places to display the **Options** dialog box.

Normality Test

Normality test results display whether the data passed or failed the test of the assumption that it was drawn from a normal population and the *P* value calculated by the test. For nonparametric procedures, this test can fail, since nonparametric tests do not assume normally distributed source populations.

These results appear unless you disabled normality testing in the Options for ANOVA on Ranks dialog box.

Equal Variance Test

Equal Variance test results display whether or not the data passed or failed the test of the assumption that the samples were drawn from populations with the same variance and the *P* value calculated by the test. Nonparametric tests do not assume equal variances of the source populations.

These results appear unless you disabled equal variance testing in the Options for ANOVA on Ranks dialog box.

Summary Table

If you selected this option in the Options for ANOVA on Ranks dialog box, SigmaPlot generates a summary table listing the medians, the percentiles defined in the Options dialog box, and sample sizes *N*.

N (Size). The number of non-missing observations for that column or group.

Missing. The number of missing values for that column or group.

Median. The "middle" observation as computed by listing all the observations from smallest to largest and selecting the largest value of the smallest half of the observations. The median observation has an equal number of observations greater than and less than that observation.

Percentiles. The two percentile points that define the upper and lower tails of the observed values.

H Statistic

The ANOVA on Ranks test statistic *H* is computed by ranking all observations from smallest to largest without regard for treatment group. The average value of the ranks for each treatment group are computed and compared.

For large sample sizes, this value is compared to the chi-square distribution (the estimate of all possible distributions of *H*) to determine the possibility of this *H* occurring. For small sample sizes, the actual distribution of *H* is used.

If *H* is small, the average ranks observed in each treatment group are approximately the same. You can conclude that the data is consistent with the null hypothesis that all the samples were drawn from the same population (for example, no treatment effect).

If *H* is a large number, the variability among the average ranks is larger than expected from random variability in the population, and you can conclude that the samples were drawn from different populations (for example, the differences between the groups are statistically significant).

P Value. The *P* value is the probability of being wrong in concluding that there is a true difference in the groups (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on *H*). The smaller the *P* value, the greater the probability that the samples are significantly different. Traditionally, you can conclude there are significant differences when *P* < 0.05.

Multiple Comparisons

If a difference is found among the groups, and you requested and elected to perform multiple comparisons, a table of the comparisons between group pairs is displayed. The multiple comparison procedure is activated in the Options for ANOVA on Ranks dialog box. The test used in the multiple comparison procedure is selected in the Multiple Comparison Options dialog box.

Multiple comparison results are used to determine exactly which groups are different, since the ANOVA results only inform you that two or more of the groups are different. The specific type of multiple comparison results depends on the comparison test used and whether the comparison was made pairwise or versus a control.

- All pairwise comparison results list comparisons of all possible combinations of group pairs: the all pairwise tests are the Tukey, Student-Newman-Keuls test and Dunn's test.

- Comparisons versus a single control list only comparisons with the selected control group. The control group is selected during the actual multiple comparison procedure. The comparison versus a control tests are Dunnett's test and Dunn's test.

Tukey, Student-Newman-Keuls, and Dunnett's Test Results The Tukey and Student-Newman-Keuls (SNK) tests are all pairwise comparisons of every combination of group pairs. Dunnett's test only compares a control group to all other groups. All tests compute the q test statistic. They also display the number of rank sums spanned in the comparison p, and display whether or not $P < 0.05$ or < 0.01 for that pair comparison.

You can conclude from "large" values of q that the difference of the two groups being compared is statistically significant.

If the P value for the comparison is less than 0.05, the probability of being incorrect in concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The Difference of Ranks is a gauge of the size of the real difference between the two groups.

p is a parameter used when computing q or. The larger the p, the larger q needs to be to indicate a significant difference. p is an indication of the differences in the ranks of the group means being compared. Group rank sums are ranked in order from largest to smallest in an SNK or Dunnett's test, so p is the number of rank sums spanned in the comparison. For example, when comparing four rank sums, comparing the largest to the smallest p = 4, and when comparing the second smallest to the smallest p = 2.

If a group is found to be not significantly different than another group, all groups with ranks in between the rank sums of the two groups that are not different are also assumed not to be significantly different, and a result of DNT (Do Not Test) appears for those comparisons.

Dunn's Test Results Dunn's test is used to compare all groups or to compare versus a control. Dunn's test lists the difference of rank means, computes the Q test statistic, and displays whether or not $P < 0.05$, for each group pair.

You can conclude from "large" values of Q that the difference of the two groups being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of being incorrect in concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The Difference of Rank Means is a gauge of the size of the difference between the two groups.

ANOVA on Ranks Report Graphs

You can generate up to three graphs using the results from an ANOVA on Ranks. They include a:

- [Point plot of the column data](#).
- [Box plot](#).
- [Multiple comparison graphs](#).

How to Create an ANOVA on Ranks Graph

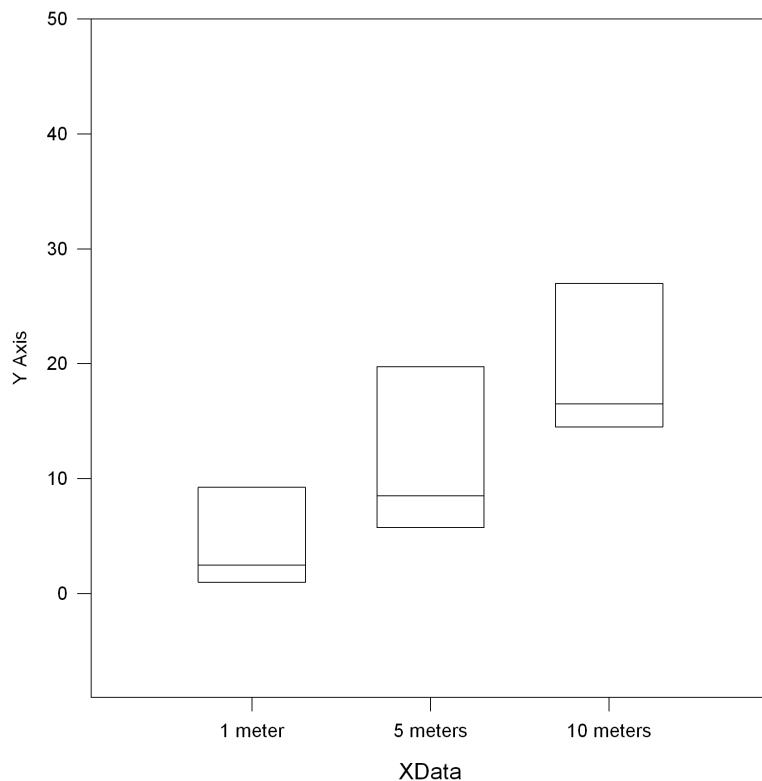
How to Create an ANOVA on Ranks Graph

1. Select the **ANOVA on Ranks** test report.
2. Click the **Report** tab.
3. In the **Results Graphs** group, click **Create Result Graph**.

4. Select the type of graph you want to create from the **Graph Type** list, then click **OK**.

The selected [graph](#) appears in a graph window.

Box Plot



One Way Analysis of Covariance

ANCOVA (Analysis of Covariance) is an extension of ANOVA where the model contains one or more additional variables called *covariates*. The covariates are variables outside the investigator's control that affect the observations within one or more factor group. With ANCOVA you can:

- Reduce the unexplained variance in your dependent variable data, improving the precision of results.
- Increase the sensitivity of the test, achieving higher statistical power than the standard ANOVA model.

About One Way Analysis of Covariance

A single-factor ANOVA model is based on a *completely randomized design* in which the subjects of a study are randomly sampled from a population and then each subject is randomly assigned to one of several factor levels or treatments so that each subject has an equal probability of receiving a treatment. A common assumption of this design is that the subjects are *homogeneous*. This means that any other variable, where differences between the subjects exist, does not significantly alter the treatment effect and need not be included in the model. However, there are often variables, outside the investigator's control, that affect the observations within one or more factor groups, leading to necessary adjustments in the group means, their errors, the sources of variability ,and the P-values of the group effect, including multiple comparisons. These variables are called *covariates*. They are typically continuous variables, but can also be categorical. Since they are usually of secondary importance to the study and, as mentioned above, not controllable by the investigator, they do not represent additional main-effects factors, but can still be included into the model to improve the precision of the results. Covariates are also known as *nuisance* variables or *concomitant* variables.

The ANCOVA data in a SigmaPlot worksheet is arranged in the *indexed* data format, where one column represents the factor groups and one column represents the dependent variable (the observations) as in an ANOVA design. In addition, you have one column for each covariate. When using a model that includes the effects of covariates, there is more explained variability in the value of the dependent variable. This generally reduces the unexplained variance that is attributed to random sampling variability, which increases the sensitivity of the ANCOVA as compared to the same model without covariates (the ANOVA model). Higher test sensitivity means that smaller mean differences between treatments will become significant as compared to a standard ANOVA model, thereby increasing statistical power.

As a simple example of using ANCOVA, consider an experiment where students are randomly assigned to one of three types of teaching methods and their achievement scores are measured. The goal is to measure the effect of the different methods and determine if one method achieves a significantly higher average score than the others. The methods are Lecture, Self-paced, and Cooperative Learning. Performing a One Way ANOVA on this hypothetical data gives the results in the table below, under the ANOVA column heading. We conclude there is no significant difference among the teaching methods. Also note that the variance unexplained by the ANOVA model which is due to the random sampling variability in the observations is estimated as 35.17.

It is possible that students in our study may benefit more from one method than the others, based on their previous academic performance. Suppose you refine the study to include a covariate that measures some prior ability, such as a state-sanctioned Standards Based Assessment (SBA). Performing a One Way ANCOVA on this data gives the results in the table below, under the ANCOVA column heading.

ANOVA		ANCOVA		
Method	Mean	Std. Error	Adjusted Mean	Std. Error
Coop	79.33	2.421	82.09	0.782
Self	83.33	2.421	82.44	0.751
Lecture	86.83	2.421	84.97	0.764
	P = .124		P = 0.039	
	MSres = 35.17		MSres = 3.355	

The adjusted mean that is given in the table for each method is a correction to the group mean to control for the effects of the covariate. The results show the adjusted means are significantly different with the Lecture method as the more successful. Notice how the standard errors of the means have decreased by almost a factor of three while the variance due to random sample variability has decreased by a factor of ten. A reduction in error is the usual consequence of introducing covariates and performing an ANCOVA analysis.

Covariates can be used to extend various treatment and design structures other than One Way ANOVA for the purpose of increasing the precision of results. For example, there are models for multi-factor ANCOVA and repeated measures ANCOVA. There are also alternatives to ANCOVA, such as the use of *randomized block designs*.

Performing a One Way Analysis of Covariance

To perform an One Way Analysis of Covariance:

1. Enter or arrange your data appropriately in the worksheet.
2. If desired, set the One Way Analysis of Covariance options.
3. Click the Analysis tab.
4. In the SigmaStat group, from the Tests drop-down list, select:
Compare Many Groups > One Way ANCOVA
5. Run the test.
6. Interpret results.
7. Generate report graphs.

Arranging One Way Analysis of Covariance Data

One Way Analysis of Covariance data uses the indexed data format, where one column is for the Factor variable, one column is for the Data (dependent variable), and one column for each Covariate variable.

Figure 70: The One Way ANCOVA test requires that you enter data using the indexed data

	1-Gender	2-Height	3-Weight	4-Age	5
1	f	56.3000	85.0000	143.0000	
2	f	62.5000	112.5000	191.0000	
3	f	62.0000	94.5000	160.0000	
4	f	64.5000	123.5000	157.0000	
5	f	65.3000	107.0000	191.0000	
6	f	61.8000	85.0000	141.0000	
7	f	63.3000	101.0000	185.0000	
8	f	65.5000	140.0000	210.0000	
9	f	64.3000	110.5000	149.0000	
10	f	62.3000	99.5000	169.0000	
11	f	62.8000	102.5000	173.0000	
12	f	61.3000	94.0000	150.0000	
13	f	59.5000	93.5000	144.0000	
14	f	60.0000	109.0000	146.0000	
15	f	61.3000	107.0000	155.0000	
16	f	64.5000	102.5000	183.0000	
17	f	60.0000	114.0000	154.0000	
18	f	60.5000	105.0000	152.0000	
19	f	60.5000	84.5000	148.0000	
20	f	65.3000	98.0000	164.0000	
21	f	61.3000	81.0000	177.0000	
22	f	66.5000	112.0000	183.0000	
23	f	65.5000	133.0000	182.0000	
24	f	55.5000	67.0000	165.0000	
25	f	56.5000	84.0000	163.0000	
26	f	63.0000	84.0000	171.0000	

format.

Setting One Way Analysis of Covariance Options

Use the ANCOVA options to:

- Set assumption checking options.
- Specify the residuals to display and save them to the worksheet.
- Display confidence intervals and save them to the worksheet.
- Specify tests to identify outlying or influential data points.
- Display power.

To set ANCOVA options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. Click the Analysis tab.
3. In the SigmaStat group, click Options. The Options for One Way ANCOVA dialog box appears with four tabs:

Table 11: Options for One Way ANCOVA

Option	Description
Assumption Checking	Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance

Option	Description
Residuals	Click the Residuals tab to view the residual options.
More Results	Click the More Results tab to view confidence intervals, plus Leverage, Cooks Distance, and other options.
Post Hoc Tests	Compute the power or sensitivity of the test and enable multiple comparisons.

4. Select and option to enable or disable it. SigmaPlot saves options between SigmaPlot sessions. For more information, see [Interpreting Simple Linear Regression Results](#).
5. To continue the test, click **Run Test**.
6. To accept the current settings and close the options dialog box, click **OK**.

Options for One Way ANCOVA: Assumption Checking

The normality assumption test checks for a normally distributed population. The equal variance assumption test checks the variability about the group means.

- **Normality Testing.** SigmaPlot uses the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.
- **Equal Variance Testing.** SigmaPlot tests for equal variance by checking the variability about the group means.
- **P Values for Normality and Equal Variance.** The *P* value determines the probability of being incorrect in concluding that the data is not normally distributed (*P* value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the *P* computed by the test is greater than the *P* set here, the test passes.

To require a stricter adherence to normality and/or equal variance, decrease the *P* value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of *P* (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality and/or equal variance, increase *P*. Requiring larger values of *P* to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a *P* value of 0.100 requires greater deviations from normality to flag the data as non-normal than a value of 0.050.

- **Restriction:** There are extreme conditions of data distribution that these tests cannot take into account. For example, the Levene Median test fails to detect differences in variance of several orders of magnitude; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.
- **Equality of Slopes.** **Equality of Slopes** tests the assumption that there is no interaction between the factor variable (the treatment levels) and the covariate variables. In other words, the coefficient of each covariate in the model is assumed to be the same for all treatments.
- **Report Parametric Statistics.** If you leave **Equality of Slopes** clear, then SigmaPlot assumes equality of the slopes and the analysis for the report focuses on the results of fitting the ANCOVA model to your data. If you do select **Equality of Slopes**, then the interaction model fits to the data to determine if any of the interactions is significant. If an interaction between the factor and any covariate is significant (so the equality of slopes test fails), then the analysis stops, but SigmaPlot will still provide regression equations for each group. If there is no significant interaction between the factor and any covariate (so the equality of slopes test passes), then the report continues to provide the results of the *equal slopes model*.

 **Note:** The ANCOVA model can also be called the *equal slopes model*.

Options for One Way ANCOVA: Residuals

Predicted Values. Use this option to calculate the predicted value of the dependent variable for each observed value of the independent variable(s), then save the results to the worksheet. Clear the selected option if you do not want to include raw residuals in the worksheet.

To assign predicted values to a worksheet column, select the worksheet column you want to save the predicted values to from the corresponding drop-down list. If you select **None** and you've also selected **Predicted Values**, the values appear in the report but are not assigned to the worksheet.

Raw Residuals. The raw residuals are the differences between the predicted and observed values of the dependent variables. To include raw residuals in the report, make sure this check box is selected. Clear the selected option if you do not want to include raw residuals in the worksheet.

To assign the raw residuals to a worksheet column, select the number of the desired column from the corresponding drop-down list. If you select **None** from the drop-down list and you've also selected **Raw**, the values appear in the report but are not assigned to the worksheet.

Standardized Residuals. The standardized residual is the residual divided by the standard error of the estimate. The standard error of the residuals is essentially the standard deviation of the residuals, and is a measure of variability around the regression line. To include standardized residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include raw residuals in the worksheet.

Flag Values >. SigmaPlot automatically flags data points lying outside of the confidence interval specified in the corresponding box. These data points are considered to have "large" standardized residuals, for example, outlying data points. You can change which data points are flagged by editing the value in the Flag Values > edit box. The suggested residual value is 2.5.

Studentized Residuals. Studentized residuals scale the standardized residuals by taking into account the greater precision of the regression line near the middle of the data versus the extremes. The Studentized residuals tend to be distributed according to the Student t distribution, so the t distribution can be used to define "large" values of the Studentized residuals. SigmaPlot automatically flags data points with "large" values of the Studentized residuals, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

To include studentized residuals in the report, make sure that you select this option. Clear this options if you do not want to include studentized residuals in the worksheet.

Studentized Deleted Residuals. Studentized deleted residuals are similar to the Studentized residual, except that the residual values are obtained by computing the regression equation without using the data point in question.

To include Studentized deleted in the report, make sure you select this option. Clear this options if you do not want to include Studentized deleted residuals in the worksheet.

SigmaPlot can automatically flag data points with "large" values of the Studentized deleted residual, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

 **Note:** Both Studentized and Studentized deleted residuals use the same confidence interval setting to determine outlying points.

Report Flagged Values Only. To include only the flagged standardized and Studentized deleted residuals in the report, make sure you've selected **Report Flagged Values Only**. Clear this option to include all standardized and Studentized residuals in the report.

Options for One Way ANCOVA: More Results

Confidence Intervals for Model. Displays confidence and prediction intervals for the true values of the model in the report and in the worksheet. The value in the **Confidence Level** applies to the confidence intervals computed for the regression parameters as well as the intervals computed for the mean response of the observations in your data. The confidence intervals for the parameters are included in the report if you select **Parameter Statistics**. If the Equality of Slopes test passes, the confidence intervals will be computed for the **Equal Slopes Model**; otherwise, the intervals are computed for the Interaction Model.

Summary of Covariates. Select **Summary of Covariates** to place basic statistics for each covariate, including sample mean, sample standard deviation, minimum value, and maximum value in the report. These statistics are computed within each group and also across all groups.

Covariance Matrix. Select **Covariate Matrix** to place the covariance matrix for the parameters in the **Equal Slopes Model** section the report.

Parameter Statistics. Select **Parameter Statistics** to place two tables of statistics for all regression parameters in the **Equal Slopes Model** in the report. One table displays the parameter values, their standard errors, and results for testing the hypothesis that the true value of the parameter is zero. The other table gives the confidence intervals for the parameters.

- **Variance Inflation Factor.** **Variance Inflation Factor** disabled if don't select **Parameter Statistics**. Select **Variance Inflation Factor** to append a column titled *VIF* to the first table of **Parametric Statistics**. This column contains the *variance inflation factors* of the regression parameters in the **Equal Slopes Model**, excluding the constant term. The variance inflation factor of a parameter measures the multicollinearity of the data variable for which it is the coefficient to the other data variables in the model.
- **Flag Values.** Here you can also set the threshold level for flagging values. If you exceed this threshold, the report displays statements to indicate which variables are causing the multicollinearities.

Leverage. **Leverage** measures influential points, including both vertical and horizontal outliers, of the regression.

Cook's Distance. **Cook's Distance** also measures influential points, including both vertical and horizontal outliers, of the regression.

- **Flag values.** Select to set the threshold level for flagging values.

Report Flagged Values Only. If you select **Report Flagged Values Only**, certain statistics and diagnostics that appear in the report will only have those values displayed if they exceed the various thresholds specified in **Flag values**. These include Standardized, Studentized, and Studentized Deleted residuals, as well as the diagnostics Leverage and Cook's Distance.

Options for One Way ANCOVA: Post Hoc Tests

Power. The power or sensitivity of a test is the probability that the test will detect a difference between the groups if there is really a difference.

Use Alpha Value. Change the alpha value by editing the number in the Alpha Value box. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

Multiple Comparisons

A One Way ANCOVA tests the hypothesis of no differences between the several treatment groups, but do not determine which groups are different, or the sizes of these differences. Multiple comparisons isolate these differences whenever a One Way ANCOVA detects a difference.

The P value used to determine if the One Way ANCOVA detects a difference is set in the **Report** tab of the **Options** dialog box. If the P value produced by the One Way ANCOVA is less than the P value specified in the box, a difference in the groups is detected and the multiple comparisons are performed.

- **Always Perform.** Select to perform multiple comparisons whether or not the One Way ANCOVA detects a difference.

Only When ANCOVA P Value is Significant. Perform multiple comparisons only if the ANCOVA detects a difference.

Significant Multiple Comparison Value. Select either .05 or .10 from the Significance Value for Multiple Comparisons drop-down list. This value determines that the likelihood of the multiple comparison being incorrect in concluding that there is a significant difference in the treatments.

A value of .05 indicates that the multiple comparisons will detect a difference if there is less than 5% chance that the multiple comparison is incorrect in detecting a difference. A value of .10 indicates that the multiple comparisons will detect a difference if there is less than 10% chance that the multiple comparison is incorrect in detecting a difference.

-  **Note:** If multiple comparisons are triggered, the Multiple Comparison Options dialog box appears after you pick your data from the worksheet and run the test, prompting you to choose a multiple comparison test.

Running a One Way Analysis of Covariance

If you want to select your data before you run the test, drag the pointer over your data.

1. Click the **Analysis** tab.
2. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Compare Many Groups > One Way ANCOVA

The **One Way ANCOVA - Select Data** panel of the Test Wizard appears.

3. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Factor** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The number or title of selected columns appear in each row. You are prompted to pick a minimum three worksheet columns.

4. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
5. Click **Finish** to perform the One Way ANCOVA. The One Way ANCOVA report appears if you:
 - Elected to test for normality and equal variance, and your data passes both tests.
 - Your data has no missing data points, cells, or is not otherwise unbalanced.
 - Selected not to perform multiple comparisons, or if you selected to run multiple comparisons only when the P value is significant, and the P value is not significant

-  **Note:** If you elected to test for normality and equal variance, and your data fails either test, either continue or transform your data, then perform the One Way ANCOVA on the transformed data. If your data is missing data points, missing cells, or is otherwise unbalanced, you are prompted to perform the appropriate procedure.

Interpreting One Way Analysis of Covariance Results

Basic numeric results always appear in an ANCOVA report, unless equality of slopes is being tested (turned on in Assumption Checking panel of the Test Options dialog box) and the test fails. In this case, the results for the ANCOVA model do not appear.

You also have the option to display [other results](#).

Basic Numeric Results

- The following sections of a report are always displayed, unless equality of slopes is being tested (turned on in Assumption Checking panel of the Test Options dialog box) and the test fails. In this case, the results for the ANCOVA model are not displayed.
 - **Header.** This includes the name of the test, date stamp, and data source, as for all other tests.
 - **Dependent Variable.** This gives the column name for the dependent variable (observations or responses).
 - **Descriptive Statistics for Groups.** Table with basis statistics and summary data for the factor groups. The layout of the table is given below (grid lines below will not be shown in the report).

Group

Total

- **Notification of Eliminated Groups.** A group is eliminated from the analysis if it has no valid subjects (all cases for this group are labeled as missing). A subject is valid only if the dependent variable value and the associated value of each covariate are finite numbers. A statement is placed in the report indicating which groups, if any, have been removed from the analysis. Note that if fewer than two valid groups exist, then the Test Wizard will display a message when Finish is pressed to indicate an analysis cannot be performed. In this case, you are returned to the Wizard to repick data.
- **Notification of Eliminated Covariates.** The set of covariates may be severely ill-conditioned, either precluding a unique solution to the regression problem that determines the regression planes for each group, or giving a solution in which the parameters have very large errors. To avoid this difficulty, covariates are selectively removed until a well-conditioned set of covariates remains. Problematic covariates include those which are constant on each group or those which are linear combinations of other covariates. Removing these covariates has no effect on sums of squares estimates of effects or on the predicted values of the regression. A statement will be placed in the report indicating which covariates, if any, have been removed from the analysis. If all covariates have been removed, then the report will suggest using a One-Way ANOVA test to analyze the design.

The remaining basic numeric results in the report are for the main ANCOVA model, or the *Equal Slopes* model. This model assumes that the slope coefficients are the same across all groups for each covariate. The main hypothesis testing using the equal slopes model is to test the equality of the adjusted group means (the means obtained by controlling for the effects of the covariates). This is equivalent to testing the equality of the intercepts of the regression equations for the groups since the slopes are equal. If the P-value for the effect of the factor is significant, then multiple comparisons can be performed to determine the most significant differences in the adjusted means between pairs of groups.

Suppose your ANCOVA treatment structure has total valid observations, valid groups, and valid covariates. The analysis of the ANCOVA model can only proceed *if*; however, this condition is selected in the **Test Wizard** before processing begins. If the condition is violated, then the Test Wizard displays a message when you click **Finish** to indicate the analysis cannot be performed. In this case, you are returned to the **Test Wizard** to repick the data.

- **The Analysis of Variance Table.** The Analysis of Variance table gives the sums of squares for the main effects. Using this table, we can determine the significance of the differences between the adjusted means. We can also determine the significance of each covariate in the model. If no covariate is significant, then a statement is added to the report to inform you. SigmaPlot will suggest that you use a One-Way ANOVA model to analyze the data. The layout of this table is shown below. Note that the table is preceded with a statement giving the computations of the correlation coefficient *R*, the coefficient of determination *Rsq*, and adjusted *Rsq* (*Adj Rsq*).

R=	Rsq=	Adj Rsq=			
Analysis of Variance:					
Source of Variation	DF	SS	MS	F	P
Factor					
Cov 1					
⋮	⋮				
Cov <i>m</i>					
Residual				None	None
Total (Corrected)				None	None

- **Adjusted Means Table.** This table reports the adjusted mean for each factor group together with the standard error and confidence intervals for each mean. There will be a statement following the table indicating how the adjusted means are computed. The layout of this table is given below:

Adjusted Means of the Groups

Group	Adjusted Mean	Std. Error	Conf-Lower	Conf-Upper
Group 1				
⋮				
Group p				

The adjusted means for the groups are obtained by evaluating the model at the grand mean of each covariate.

- **Regression Equations for Each Group.** For each group, the regression equation is given, where the model's predicted values are expressed as a function of the covariates.
- • **Optional Numeric Results.** The appearance of the following sections of the report depends upon the settings in the **Test Options** dialog box. SigmaPlot defaults are [described above](#).
- **Descriptive Statistics for Covariates.** Tables with basic statistics and summary data for the covariates your data. The layout of each table is given below (grid lines are not shown in the report).

Group 1:

Covariate	Mean	Std Dev	Min	Max
Cov 1				
⋮	⋮	⋮	⋮	⋮
Cov m				

Group p:

Covariate	Mean	Std Dev	Min	Max
Cov 1				
⋮	⋮	⋮	⋮	⋮
Cov m				

Overall:

Covariate	Mean	Std Dev	Min	Max
Cov 1				
⋮	⋮	⋮	⋮	⋮
Cov m				

The above tables of the covariate statistics are computed for valid rows of data only.

These tables appear immediately after the *Descriptive Statistics for Groups* table.

- **The Interaction Model (Testing Equality of Slopes): Analysis of Variance Table.** Suppose your ANCOVA treatment structure has total valid observations, valid groups, and valid covariates. If the option for Equality of Slopes on the Assumption Checking panel in the Test Options dialog box is turned on and if, then results will be given for testing the equality of slopes across factor groups for each covariate. If results are not given (or cannot be given), then analysis proceeds with the *Equal Slopes* model. The Analysis of Variance table gives the sums of squares for the main effects and for interactions of the factor with the covariates. Using this table, we can determine if any interaction is significant. If at least one interaction is significant, an equation will be given for the regression plane for each group and the report will end. If no interaction is significant, then the test for equality of slopes passes and the analysis continues with the ANCOVA model (equal slopes model), producing the remaining results of the report. The layout of this table is shown below. I have inserted mathematical expressions for the

degrees of freedom associated with each effect in terms,, and. Note that the table is preceded by a line giving the computations of the correlation coefficient R , the coefficient of determination $Rsqr$, and adjusted $Rsqr$ (Adj $Rsqr$).

R=	Rsqr=	Adj Rsqr=			
Analysis of Variance:					
Source of Variation	DF	SS	MS	F	P
Factor					
Cov 1					
⋮	⋮				
Cov m					
Factor x Cov 1					
⋮	⋮				
Factor x Cov m					
Residual				None	None
Total (Corrected)				None	None

- **The Interaction Model (Testing Equality of Slopes): Parameter Statistics.** In addition to the Analysis of Variance table, the report will display the values and statistics of the parameters in the interaction model if the additional options were turned on the Test Options dialog box. These tables will be placed in the report immediately before the Analysis of Variance table discussed in the previous item. The layouts of the tables are given below:

Effects coded group variables – Group p is the reference group				
Parameter	Value	Std. Error	t	P
Constant				
Group 1				
⋮				
Group p-1				
Cov 1				
⋮				
Cov m				
Group 2 - Cov 1				
⋮				
Group (p-1) - Cov 1				
⋮				
Group 2 - Cov m				
⋮				
Group (p-1) - Cov m				

Parameter	Value	Conf-Lower	Conf-Upper
Constant			

Group 1
⋮
Group p-1
Cov 1
⋮
Cov m
Group 2-Cov 1
⋮
Group p-1-Cov 1
⋮
Group 2-Cov m
⋮
Group p-1-Cov m

- **Normality.** As in several other statistical tests, we will test the normality of the residuals obtained from the ANCOVA regression analysis for the Equal Slopes model. The settings in the Test Options dialog box on the Assumption Checking panel will determine which normality test to perform and at what significance level. The

Constant
Group 1
⋮
Group p-1
Cov 1
⋮
Cov m

- **Parameter Covariance Matrix.** This matrix gives the estimates for the covariance among pairs of the parameter estimates in the Equal Slopes model.
- **Power.** This is the observed (retrospective or achieved) power of the test for equality of the adjusted means. This value is only computed so you can observe the relationship between the significance of the test and the Power (the probability of correctly rejecting the null hypothesis that the means are equal); however, the power value, being computed from the data, provides no new information about the analysis.
- **Multiple Comparisons.** The multiple comparison procedures that are available to One Way ANOVA should also be available to ANCOVA for pairwise comparisons of adjusted means. The major difference involves the computation of the standard error of the difference in the adjusted means. For ANOVA, the standard error computation for each pair of groups is based on the common variance between groups, estimated by the mean residual sum of squares. For ANCOVA, the standard error is dependent on the pair of groups being compared.

The following output provides influence diagnostics for fitting the ANCOVA model to your data as well as confidence intervals for the means values corresponding to each observation.

- **Predicted Values and Residuals.** The predicted values from the regression as well as various normalizations of the residuals (raw, standardized, Studentized, Studentized deleted) to determine influential data points.
- **Cook's Distance and Leverage.** Other outlier detection methods in addition to residual analysis. Cook's distance is measuring a change in the parameters relative to a particular norm when an observation is deleted. Leverage detects potential horizontal outliers (outliers relative to the independent variables).
- **Confidence Intervals for Observation Means.** Using the predicted values for the observations, confidence intervals and prediction intervals will be given for locating the true mean values and estimating intervals containing additional observations.

Other Results

There are several other results consisting of explanations and interpretations of the numeric results.

To omit these from the report, select **Main button > Options > Report**. Then clear **Explain Test Results**.

- If the Equality of Slopes assumption is being tested and the test fails, thus concluding that the slopes differ significantly between groups for at least one covariate, then a statement will follow the Analysis of Variance table for the Interaction Model to indicate which covariates have a significant interaction and a statement that the analysis of the ANCOVA model will not be performed.
- Statements are made following the Analysis of Variance table for the ANCOVA model to indicate whether there are or are not significant differences between the adjusted means of the treatment groups and to indicate the relevance of the covariates. If the P-value corresponding to the F-test for the factor effect is significant (look at the Factor row in the Analysis of Variance table), then is a significant difference between the adjusted means of the groups. A statement is also made indicating the significance or non-significance of each covariate (for each covariate, you are testing the hypothesis that the slope is zero). If no covariate makes a significant contribution to the model, then SigmaPlot suggests that you may want to run a One Way ANOVA on the data with the covariates removed.
- Following the Adjusted Means table, a statement is made to indicated that the adjust means are the predicted values of the model where each covariate variable is evaluated at the grand mean of its sampled values.

One Way Analysis of Covariance Report Graphs

You can generate up to five graphs using the results from a One Way ANCOVA. They include a:

- Normal probability plot of the residuals.
- Regression lines for groups.
- Adjusted means with confidence intervals.
- Multiple comparisons.
- Scatter plot of the residuals.

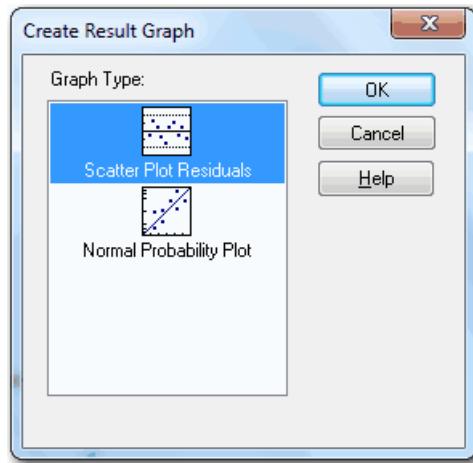
How to Create a One Way ANCOVA Report Graph

1. Select the **One Way ANCOVA** test report.

2. Click the **Report** tab.

3. In the **Results Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the One Way ANCOVA results.



4. Select the type of graph you want to create from the **Graph Type** list.

 **Note:** If you select **Scatter Plot Residuals**, the **Select a Covariate** dialog box appears. Select a covariate from the list to use for the horizontal axis.

- Click **OK**, or double-click the desired graph in the list.

The selected graph appears in a [graph window](#).

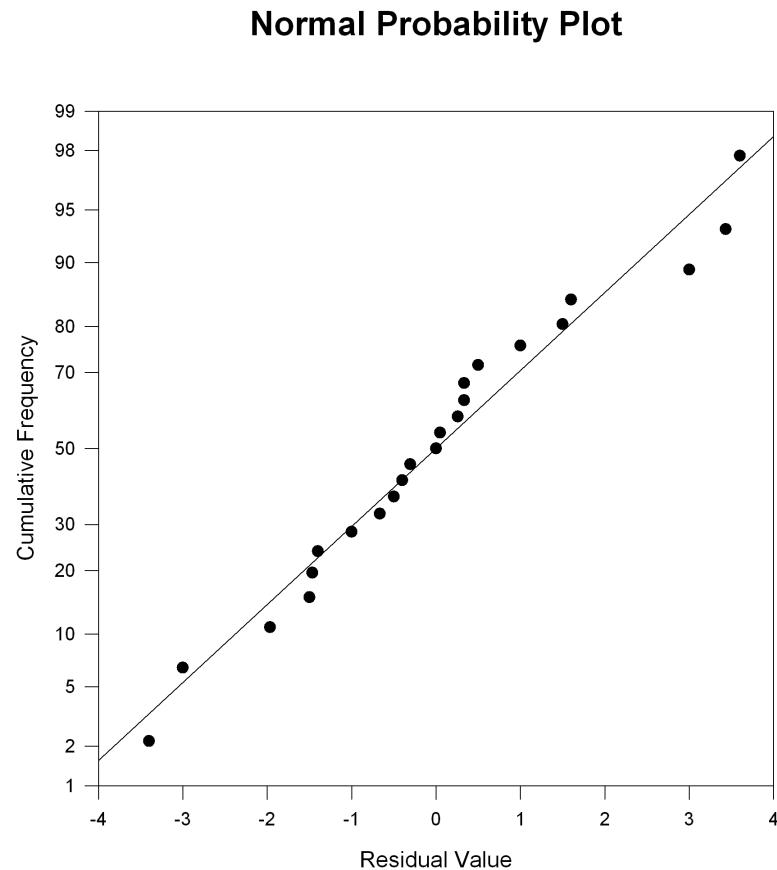


Figure 71: A Normal Probability Plot of the One Way ANCOVA

Performing a Multiple Comparison

The multiple comparison test you choose depends on the treatments you are testing. Click **Cancel** if you do not want to perform a multiple comparison test.

To perform a multiple comparison test:

- Select which factors you wish to compare under **Select Factors to Compare**.

This option is automatically selected if the P value produced by the ANOVA (displayed in the upper left corner of the dialog box) is less than or equal to the P value set in the **Options** dialog box, and multiple comparisons are performed. If the P value displayed in the dialog box is greater than the P value set in the **Options** dialog box, multiple comparisons are not performed.

- Select the desired multiple comparison test from the **Suggested Test** drop-down list.
- Select a **Comparison Type**. The types of comparisons available depend on the selected test. **All Pairwise** compares all possible pairs of treatments and is available for the Tukey, Student-Newman-Keuls, Bonferroni, Fisher LSD, and Duncan's tests. **Versus Control** compares all experimental treatments to a single control group and is available for the Tukey, Bonferroni, Fisher LSD, Dunnett's, and Duncan's tests. It is not recommended for the Tukey, Fisher LSD, or Duncan's test.
- If you select **Versus Control**, you must also select the control group from the list of groups.
- If you selected an all pairwise comparison test, click **Finish** to continue with the test and view the report. For more information, see [Interpreting One Way ANOVA Results](#) on page 79.

6. If you selected a multiple comparisons versus a control test, click Next. The Multiple Comparisons Options dialog box prompts you to select a control group. Select the desired control group from the list, then click Finish to continue the test and view the report.

Holm-Sidak Test

Use the Holm-Sidak Test for both pairwise comparisons and comparisons versus a control group. It is more powerful than the Tukey and Bonferroni tests and, consequently, is able to detect differences that these other tests do not. It is recommended as the first-line procedure for pairwise comparison testing.

When performing the test, the P values of all comparisons are computed and ordered from smallest to largest. Each P value is then compared to a critical level that depends upon the significance level of the test (set in the test options), the rank of the P value, and the total number of comparisons made. A P value less than the critical level indicates there is a significant difference between the corresponding two groups.

Tukey Test

The Tukey Test and the Student-Newman-Keuls test are conducted similarly to the Bonferroni t-test, except that they use a table of critical values that is computed based on a better mathematical model of the probability structure of the multiple comparisons. The Tukey Test is more conservative than the Student-Newman-Keuls test, because it controls the errors of all comparisons simultaneously, while the Student-Newman-Keuls test controls errors among tests of k means. Because it is more conservative, it is less likely to determine that a given difference is statistically significant and it is the recommended test for all pairwise comparisons.

Student-Newman-Keuls (SNK) Test

The Student-Newman-Keuls Test and the Tukey Test are conducted similarly to the Bonferroni t-test, except that they use a table of critical values that is computed based on a better mathematical model of the probability structure of the multiple comparisons. The Student-Newman-Keuls Test is less conservative than the Tukey Test because it controls errors among tests of k means, while the Tukey Test controls the errors of all comparisons simultaneously. Because it is less conservative, it is more likely to determine that a given difference is statistically significant. The Student-Newman-Keuls Test is usually more sensitive than the Bonferroni t-test, and is only available for all pairwise comparisons.

Bonferroni t-Test

The Bonferroni t-test performs pairwise comparisons with paired t -tests. The P values are then multiplied by the number of comparisons that were made. It can perform both all pairwise comparisons and multiple comparisons versus a control, and is the most conservative test for both each comparison type. For less conservative all pairwise comparison tests, see the Tukey and the Student-Newman-Keuls tests, and for the less conservative multiple comparison versus a control tests, see the Dunnett's Test.

Fisher's Least Significance Difference Test

Fisher's Least Significant Difference (LSD) Test is the least conservative of the all-pairwise comparison tests. Unlike the Tukey and Student-Newman-Keuls tests, it controls the error rate of individual comparisons and does not control the family error rate, where the "family" is the whole set of comparisons. Because of this it is not recommended.

Dunnett's Test

Dunnett's test is the analog of the Student-Newman-Keuls Test for the case of multiple comparisons against a single control group. It is conducted similarly to the Bonferroni t-test, but with a more sophisticated mathematical model of the way the error accumulates in order to derive the associated table of critical values for hypothesis testing. This test is less conservative than the Bonferroni Test, and is only available for multiple comparisons versus a control.

Dunn's test

Dunn's test must be used for ANOVA on Ranks when the sample sizes in the different treatment groups are different. You can perform both all pairwise comparisons and multiple comparisons versus a control with the Dunn's test. The all pairwise Dunn's test is the default for data with missing values.

Duncan's Multiple Range

The Duncan's Test is the same way as the Tukey and the Student-Newman-Keuls tests, except that it is less conservative in determining whether the difference between groups is significant by allowing a wider range for error rates. Although it has a greater power to detect differences than the Tukey and the Student-Newman-Keuls tests, it has less control over the Type 1 error rate, and is, therefore, not recommended.

Chapter

6

Comparing Repeated Measurements of the Same Individuals

Topics:

- [About Repeated Measures Tests](#)
- [Data Format for Repeated Measures Tests](#)
- [Paired t-Test](#)
- [Wilcoxon Signed Rank Test](#)
- [One Way Repeated Measures Analysis of Variance \(ANOVA\)](#)
- [Two Way Repeated Measures Analysis of Variance \(ANOVA\)](#)
- [Friedman Repeated Measures Analysis of Variance on Ranks](#)

Use repeated measures procedures to test for differences in same individuals before and after one or more different treatments or changes in condition.

When comparing random samples from two or more groups consisting of different individuals, use group comparison tests. For more information, see [Choosing the Procedure to Use](#) on page 18.

About Repeated Measures Tests

Repeated measures tests are used to detect significant differences in the mean or median effect of treatment(s) within individuals beyond what can be attributed to random variation of the repeated treatments. Variation among individuals is taken into account, allowing concentration of the effect of the treatments rather than the differences between individuals. For more information, see [Choosing the Repeated Measures Test to Use](#) on page 31.

Parametric and Nonparametric Tests

Parametric tests assume treatment effects are normally distributed with the same variances (or standard deviations). Parametric tests are based on estimates of the population means and standard deviations, the parameters of a normal distribution.

Nonparametric tests do not assume that the treatment effects are normally distributed. Instead, they perform a comparison on ranks of the observed effects.

Comparing Individuals Before and After a Single Treatment

Use before and after comparisons to test the effect of a single experimental treatment on the same individuals. There are two tests available:

- [The Paired t-test](#). This is a parametric test.
- [Wilcoxon Signed Rank Test](#). This is a nonparametric test.

Comparing Individuals Before and After Multiple Treatments

Use repeated measures procedures to test the effect of more than one experimental treatment on the same individuals. There are three tests available:

- [One Way Repeated Measures ANOVA](#). A parametric test comparing the effect of a single series of treatments or conditions.
- [Two Way Repeated Measures ANOVA](#). A parametric test comparing the effect of two factors, where one or both factors are a series of treatments or conditions.
- [Friedman One Way Repeated Measures ANOVA on Ranks](#). The nonparametric analog of One Way Repeated Measures ANOVA.

When using one of these procedures to compare multiple treatments, and you find a statistically significant difference, you can use several multiple comparison procedures to determine exactly which treatments had an effect, and the size of the effect. These procedures are described for each test.

Data Format for Repeated Measures Tests

You can arrange repeated measures test data in the worksheet as:

- [Columns for each treatment \(raw data\)](#).
- [Data indexed to other column\(s\)](#).

You cannot use the summary statistics for repeated measures tests.

-  **Tip:** You can perform repeated measures tests on a portion of the data by selecting a block on the worksheet before choosing the test. If you plan to do this, make sure that all data columns are adjacent to each other.

Paired t-test						
	1-Before Class	2-After Class	3-Student	4-Education	5-Score	
1	55.0000	69.0000	John	Before	55.0000	
2	51.0000	76.0000	John	After	69.0000	
3	64.0000	88.0000	Scott	Before	51.0000	
4	72.0000	82.0000	Scott	After	76.0000	
5	66.0000	70.0000	Emma	Before	64.0000	
6	48.0000	91.0000	Emma	After	88.0000	
7	72.0000	100.0000	Yvonne	Before	72.0000	
8			Yvonne	After	82.0000	
9			Dick	Before	66.0000	
10			Dick	After	70.0000	
11			Pam	Before	48.0000	
12			Pam	After	91.0000	
13			Oliver	Before	72.0000	
14			Oliver	After	100.0000	

Figure 72: Valid Data Formats for a Paired t-test

Columns 1 and 2 in the worksheet above are arranged as raw data. Columns 3, 4, and 5 are arranged as indexed data, with column 3 as the subject column, column 4 as the factor column, and column 5 as the data column.

Raw Data

To enter data in raw data format, enter the data for each treatment in separate worksheet columns. You can use raw data for all tests except Two Way ANOVAs.

Important: The worksheet columns for raw data must be the same length. If a missing value is encountered, that individual is either ignored or, for parametric ANOVAs, a general linear model is used to take advantage of all available data.

Indexed Data

Indexed data contains the treatments in one column and the corresponding data points in another column. A One Way Repeated Measures ANOVA requires a subject index in a third column. Two Way Repeated Measures ANOVA requires an additional factor column, for a total of four columns.

If you plan to compare only a portion of the data, put the treatment index in the left column, followed by the second factor index (for Two Way ANOVA only), then the subject index (for Repeated Measures ANOVA), and finally the data in the right-most column.

Tip: You can index raw data or convert indexed data to raw data.

Paired t-Test

The Paired t-test is a parametric statistical method that assumes the observed treatment effects are normally distributed. It examines the changes which occur before and after a single experimental intervention on the same individuals to determine whether or not the treatment had a significant effect. Examining the changes rather than the values observed before and after the intervention removes the differences due to individual responses, producing a more sensitive, or powerful, test.

Use Paired t-test when:

- You want to see if the effect of a single treatment on the same individual is significant.
- The treatment effects (for example, the changes in the individuals before and after the treatment) are normally distributed.

If you know that the distribution of the observed effects are non-normal, use the [Wilcoxon Signed Rank Test](#). If you are comparing the effect of multiple treatments on the same individuals, do a [Repeated Measures Analysis of Variance](#).

Performing a Paired t-test

To perform a Paired t-test:

1. [Enter or arrange your data in the worksheet](#).
2. If desired, set the Paired t-test options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Before and After > Paired t-test
5. Run the test.
6. Generate report graphs.

Arranging Paired t-Test Data

The format of the data to be tested can be raw data or indexed data. The data is placed in two worksheet columns for raw data and three columns (a subject, factor, and data column) for indexed data. The columns for raw data must be the same length. If a missing value is encountered, that individual is ignored. You cannot use statistical summary data for repeated measures tests.

	1-Before Class	2-After Class	3-Student	4-Education	5-Score	
1	55.0000	69.0000	John	Before	55.0000	
2	51.0000	76.0000	John	After	69.0000	
3	64.0000	88.0000	Scott	Before	51.0000	
4	72.0000	82.0000	Scott	After	76.0000	
5	66.0000	70.0000	Emma	Before	64.0000	
6	48.0000	91.0000	Emma	After	88.0000	
7	72.0000	100.0000	Yvonne	Before	72.0000	
8			Yvonne	After	82.0000	
9			Dick	Before	66.0000	
10			Dick	After	70.0000	
11			Pam	Before	48.0000	
12			Pam	After	91.0000	
13			Oliver	Before	72.0000	
14			Oliver	After	100.0000	

Figure 73: Valid Data Formats for a Paired t-test

Columns 1 and 2 in the worksheet above are arranged as raw data. Columns 3, 4, and 5 are arranged as indexed data, with column 3 as the subject column, column 4 as the factor column, and column 5 as the data column.

Setting Paired t-Test Options

Use the Paired t-test options to:

- Adjust the parameters of a test to relax or restrict the testing of your data for normality.
- Display the statistics summary for the data.
- Compute the power, or sensitivity, of the test.

Options settings are saved between SigmaPlot sessions.

To change the Paired t-test options:

1. Select **Paired t-test** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.
2. Click **Options**. The **Options for Paired t-test** dialog box appears with three tabs:
 - **Assumption Checking**. Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
 - **Results**. Display the statistics summary for the data in the report and save residuals to a worksheet column.
 - **Post Hoc Tests**. Compute the power or sensitivity of the test.

i **Tip:** If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.

Options settings are saved between SigmaPlot sessions.

3. To continue the test, click **Run Test**. The **Select Data** panel of the Test Wizard appears.

To accept the current settings and close the options dialog box, click **OK**.

Options for Paired t-test: Assumption Checking

The normality assumption test checks for a normally distributed population.

- i** **Tip:** **Equal Variance** is not available for the Paired t-test because Paired t-tests are based on changes in each individual rather than on different individuals in the selected population, making equal variance testing unnecessary.

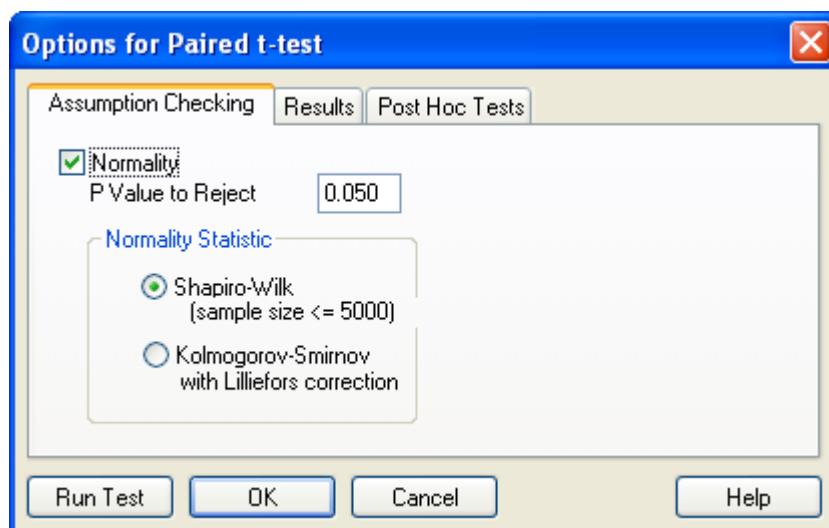


Figure 74: The Options for Paired t-test Dialog Box Displaying the Assumption Checking Options

Normality. SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

- **P Value to Reject.** Enter the corresponding P value in the **P Value to Reject** box. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (the P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P value computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.050 requires greater deviations from normality to flag the data as non-normal than a value of 0.100.

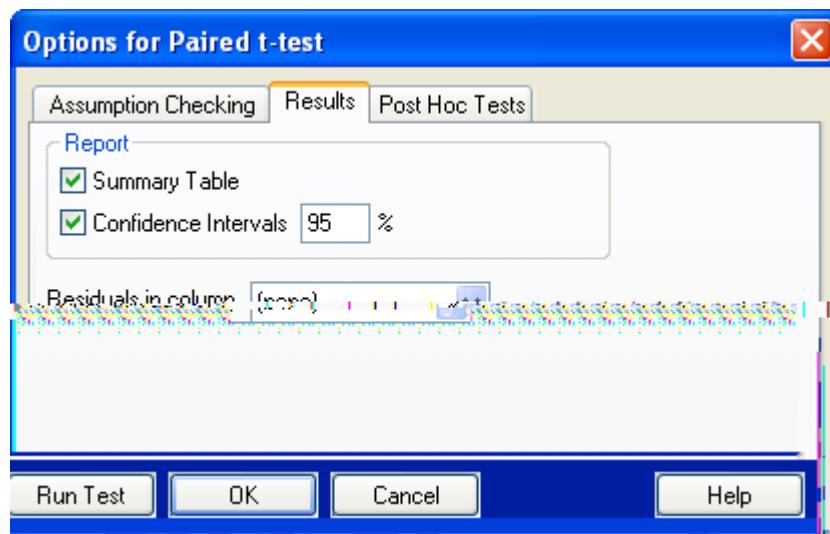
- 🚫 **Restriction:** Although the normality test is robust in detecting data from populations that are non-normal, there are extreme conditions of data distribution that this test cannot take into account; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption test.

Options for Paired t-Test: Results

Summary Table. Displays the number of observations for a column or group, the number of missing values for a column or group, the average value for the column or group, the standard deviation of the column or group, and the standard error of the mean for the column or group.

Confidence Intervals. Displays the confidence interval for the difference of the means. To change the interval, enter any number from 1 to 99 (95 and 99 are the most commonly used intervals).

Residuals in Column. Displays residuals in the report and to save the residuals of the test to the specified worksheet column. Edit the number or select a number from the drop-down list.



Options for Paired t-Test: Post Hoc Tests

Power. The power or sensitivity of a test is the probability that the test will detect a difference between the groups if there is really a difference.

Use Alpha Value. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

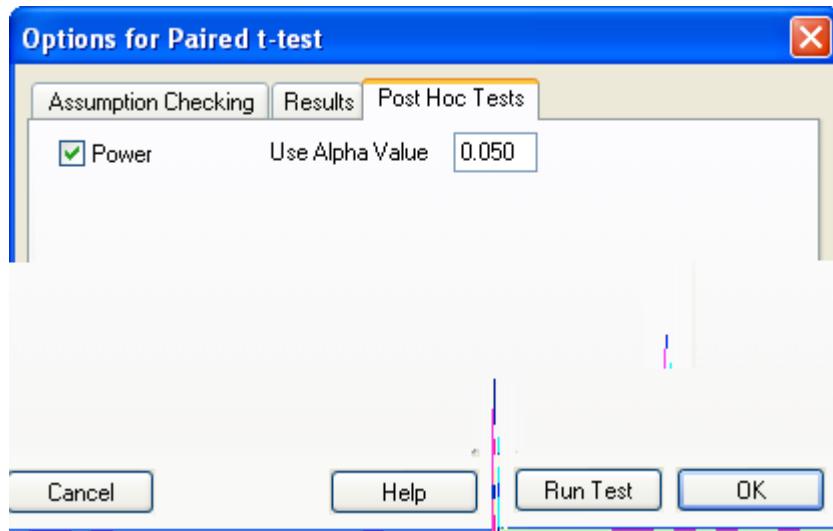


Figure 75: The Options for Paired t-test Dialog Box Displaying the Power Option

Running a Paired t-Test

If you want to select your data before you run the test, drag the pointer over your data.

1. Click the **Analysis** tab.
2. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Before and After > Paired t-test

The **Paired t-test - Data Format** panel of the Test Wizard appears prompting you to specify a data format.

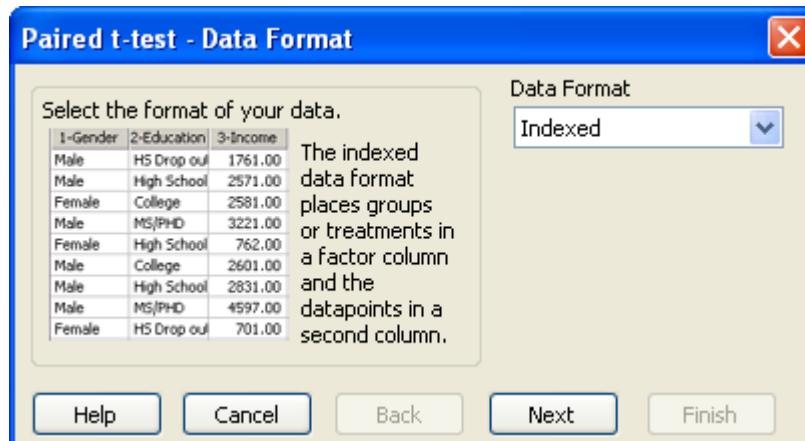


Figure 76: The Paired t-test - Data Format Panel of the Test Wizard Prompting You to Specify a Data Format

3. Select the appropriate data format (**Raw** or **Indexed**) from the **Data Format** drop-down list. For more information, see [Data Format for Repeated Measures Tests](#) on page 140.

4. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

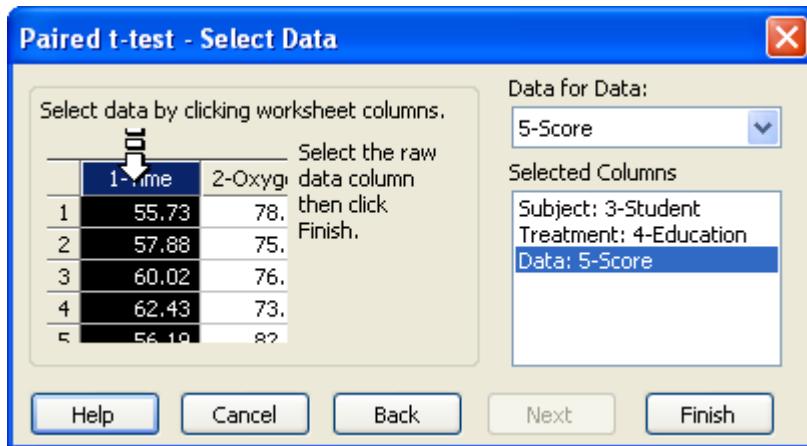


Figure 77: The Select Data Panel of the Test Wizard Prompting You to Select Data Columns

5. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw and indexed data, you are prompted to select two worksheet columns. For statistical summary data you are prompted to select three columns.

6. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
7. Click **Finish** to run the t-test on the selected columns. After the computations are completed, the report appears. For more information, see [Interpreting Paired t-Test Results](#) on page 146.

Interpreting Paired t-Test Results

The Paired t-test report displays the t statistic, degrees of freedom, and *P* value for the test. The other results displayed in the report are selected in the **Options for Paired t-test** dialog box.

Paired t-test 14 Report*

Paired t-test: Thursday, November 03, 2011, 6:10:01 PM

Data source: Paired t-test in SigmaPlot.JNB

Normality Test (Shapiro-Wilk) Passed (P = 0.852)

Treatment Name	N	Missing	Mean	Std Dev	SEM
Before Class	7	0	61.143	9.839	3.719
After Class	7	0	82.286	11.470	4.335
Difference	7	0	-21.143	13.018	4.920

t = -4.297 with 6 degrees of freedom.

95 percent two-tailed confidence interval for difference of means: -33.183 to -9.103

Two-tailed P-value = 0.00511

The change that occurred with the treatment is greater than would be expected by chance; there is a statistically significant change (P = 0.005)

One-tailed P-value = 0.00255

The sample mean of treatment After Class exceeds the sample mean of treatment Before Class by an amount that is greater than would be expected by chance, rejecting the hypothesis that the nomination

Reversed 6 performed two-tailed test with alpha = 0.050; 0.044

Reversed 6 performed one-tailed test with alpha = 0.050; 0.023

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the **Options** dialog box. You can also set the number of decimal places to display in the **Options** dialog box.

Normality Test

Normality test results display whether the data passed or failed the test of the assumption that the changes observed in each subject is consistent with a normally distributed population, and the P value calculated by the test. A normally distributed source is required for all parametric tests.

This result appears unless you disabled normality testing in the **Paired t-test Options** dialog box.

Summary Table

SigmaPlot can generate a summary table listing the sample size N, number of missing values (if any), mean, standard deviation, and standard error of the means (SEM). This result is displayed unless you disabled it in the **Paired t-test Options** dialog box.

N (Size). The number of non-missing observations for that column or group.

Missing. The number of missing values for that column or group.

Mean. The average value for the column. If the observations are normally distributed the mean is the center of the distribution.

Standard Deviation. A measure of variability. If the observations are normally distributed, about two-thirds will fall within one standard deviation above or below the mean, and about 95% of the observations will fall within two standard deviations above or below the mean.

Standard Error of the Mean. A measure of the approximation with which the mean computed from the sample approximates the true population mean.

Difference

The difference of the group before and after the treatment is described in terms of the mean of the differences (changes) in the subjects before and after the treatment, and the standard deviation and standard error of the mean difference.

The standard error of the mean difference is a measure of the precision with which the mean difference estimates the true difference in the underlying population.

t Statistic

The t-test statistic is computed by subtracting the values before the intervention from the value observed after the intervention in each experimental subject. The remaining analysis is conducted on these differences.

The t-test statistic is the ratio:

$$t = \frac{\text{mean difference of the subjects before after}}{\text{standard error of the mean difference}}$$

You can conclude from large (bigger than ~2) absolute values of t that the treatment affected the variable of interest (you reject the null hypothesis of no difference). A large t indicates that the difference in observed value after and before the treatment is larger than one would be expected from effect variability alone (for example, that the effect is statistically significant). A small t (near 0) indicates that there is no significant difference between the samples (little difference in the means before and after the treatment).

Degrees of Freedom. The degrees of freedom is a measure of the sample size, which affects the ability of t to detect differences in the mean effects. As degrees of freedom increase, the ability to detect a difference with a smaller t increases.

P Value. The P value is the probability of being wrong in concluding that there is a true effect (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on t). The smaller the P value, the greater the probability that the treatment effect is significant. Traditionally, you can conclude there is a significant difference when $P < 0.05$.

Confidence Interval for the Difference of the Means

If the confidence interval does not include a value of zero, you can conclude that there is a significant difference with that level of confidence. Confidence can also be described as $P < \alpha$, where α is the acceptable probability of incorrectly concluding that there is an effect.

The level of confidence is adjusted in the **Options for Paired t-test** dialog box; this is typically 100(1- α), or 95%. Larger values of confidence result in wider intervals.

This result is displayed unless you disabled it in the **Options for Paired t-test** dialog box.

Power

The power, or sensitivity, of a Paired t-test is the probability that the test will detect a difference between treatments if there really is a difference. The closer the power is to 1, the more sensitive the test.

Paired t-test power is affected by the sample sizes, the chance of erroneously reporting a difference a (alpha), the observed differences of the subject means, and the observed standard deviations of the samples.

This result is displayed unless you disabled it in the **Options for Paired t-test** dialog box.

Alpha. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. An α error is also called a Type I error. A Type I error is when you reject the hypothesis of no effect when this hypothesis is true.

Set the value in the **Options for Paired t-test** dialog box; the suggested value is $\alpha = 0.05$ which indicates that a one in twenty chance of error is acceptable. Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference but also increase the risk of seeing a false difference (a Type I error).

Paired t-Test Report Graphs

- **Before and after line graph.** The Paired t-test graph uses lines to plot a subject's change after each treatment.
- **Normal probability plot of the residuals.** The Paired t-test probability plot graphs the frequency of the raw residuals.
- **Histogram of the residuals.** The Paired t-test histogram plots the raw residuals in a specified range, using a defined interval set.

Creating a Graph of the Paired t-test Data

1. Select the Paired *t*-Test report.
2. On the **Report** tab, in the **Result Graphs** group, click **Create Result Graph**.

The **Create Graph** dialog box appears displaying the types of graphs available for the Paired t-test results.

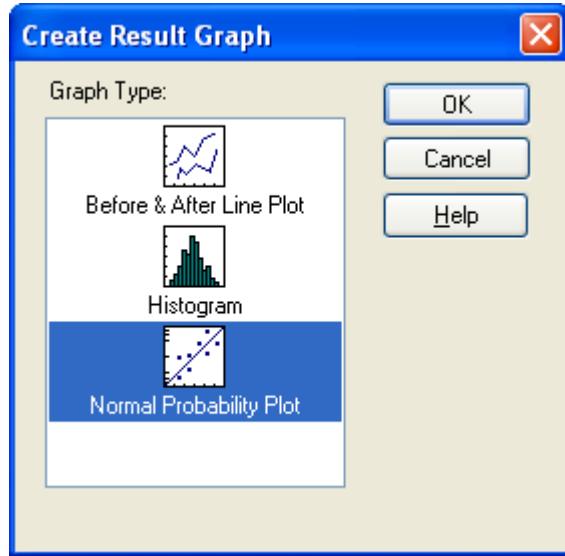


Figure 78: The Create Graph Dialog Box for Paired t-test Report Graphs

3. Select the type of graph you want to create from the **Graph Type** list, then click **OK**, or double-click the desired graph in the list.

The selected graph appears in a graph window.

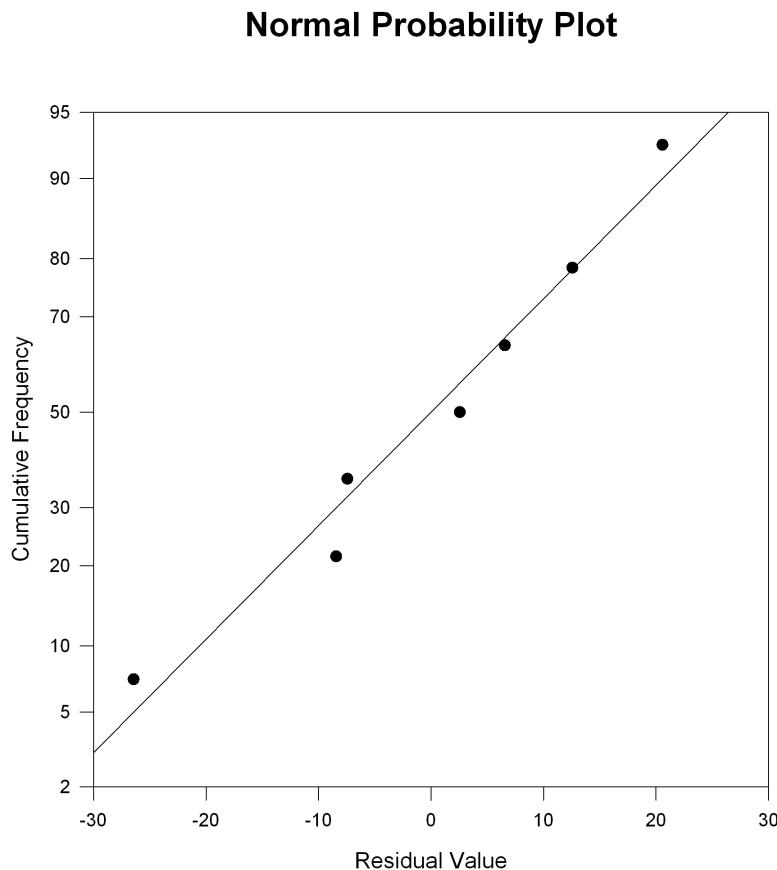


Figure 79: A Normal Probability Plot of the Report Data

Wilcoxon Signed Rank Test

The Signed Rank Test is a nonparametric procedure which does not require assuming normality or equal variance. Use a Signed Rank Test when:

- You want to see if the effect of a single treatment on the same individual is significant.
- The treatment effects are *not* normally distributed with the same variances.

If you know that the effects are normally distributed, use the [Paired t-test](#). When there are multiple treatments to compare, do a [Friedman Repeated Measures ANOVA on Ranks](#).

- Tip:** Depending on your Signed Rank Test option settings, if you attempt to perform a Signed Rank Test on a normal population, SigmaPlot suggests that the data can be analyzed with the more powerful Paired t-test instead.

About the Signed Rank Test

A Signed Rank Test Ranks all the observed treatment differences from smallest to largest without regard to sign (based on their absolute value), then attaches the sign of each difference to the ranks. The signed ranks are summed and compared. This procedure uses the size of the treatment effects and the sign.

If there is no treatment effect, the positive ranks should be similar to the negative ranks. If the ranks tend to have the same sign, you can conclude that there was a treatment effect (for example, that there is a statistically significant difference before and after the treatment).

The Wilcoxon Signed Rank Tests the null hypothesis a treatment has no effect on the subject.

Performing a Signed Rank Test

To perform a Signed Rank Test:

1. Enter or arrange your data in the data worksheet.
2. If desired, set the Signed Rank Test options.
3. On the **Analysis** tab, in the **SigmaStat** group, from the **Tests** drop-down list select:
Before and After > Signed Rank Test
4. Generate report graphs.
5. Run the test.

Arranging Signed Rank Data

The format of the data to be tested can be raw data or indexed data; in either case, the data is found in two worksheet columns.

Columns 1 and 2 are arranged as raw data. Columns 3 and 4 are arranged as indexed data, with column 3 as the factor column.

	1-Before Dose	2-After Dose	3-Subject	4-Treatment	5-Size	
1	22.0000	2.0000	A	Before	22.0000	
2	24.0000	23.0000	B	Before	24.0000	
3	4.0000	16.0000	C	Before	4.0000	
4	8.0000	5.0000	D	Before	8.0000	
5	21.0000	22.0000	E	Before	21.0000	
6	22.0000	1.0000	F	Before	22.0000	
7	17.0000	16.0000	G	Before	17.0000	
8			A	After	2.0000	
9			B	After	23.0000	
10			C	After	16.0000	
11			D	After	5.0000	
12			E	After	22.0000	
13			F	After	1.0000	

Figure 80: Valid Data Formats for a Wilcoxon Signed Rank Test

Setting Signed Rank Test Options

Use the Signed Rank Test options to:

- Adjust the parameters of the test to relax or restrict the testing of your data for normality.
- Display the summary table.
- Enable the Yates Correction Factor.

Options settings are saved between SigmaPlot sessions.

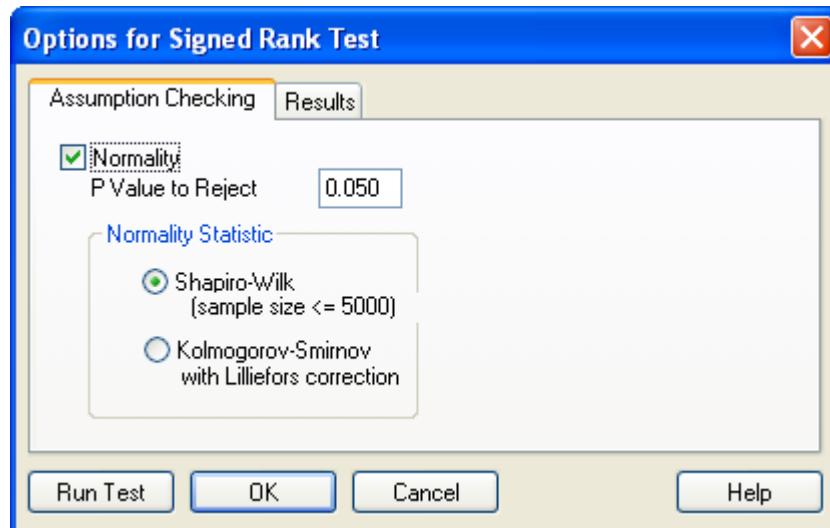
To change the Signed Rank Test options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.

2. On the **Analysis** tab, in the **SigmaStat** group, click **Options**. The **Options for Signed Rank Test** dialog box appears with two tabs:
 - **Assumption Checking.** Adjust the parameters of a test to relax or restrict the testing of your data for normality.
 - **Results.** Display the statistics summary and the confidence interval for the data in the report.
3. To continue the test, click **Run Test**. The **Select Data** panel of the **Test Wizard** appears.
4. To accept the current settings and close the options dialog box, click **OK**.

Options for Signed Rank Test: Assumption Checking

Click the **Assumption Checking** tab on the **Options for Signed Rank Test** dialog box to set Normality. The normality assumption test checks for a normally distributed population.



Note: Equal Variance is not available for the Signed Rank Test because Signed Rank Tests are based on changes in each individual rather than on different individuals in the selected population, making equal variance testing unnecessary.

Normality. SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

P Value to Reject. Enter the corresponding P value in the P Value to Reject box. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (the P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P value computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.050 requires greater deviations from normality to flag the data as non-normal than a value of 0.100.

Restriction: Although this assumption test is robust in detecting data from populations that are non-normal, there are extreme conditions of data distribution that this test cannot take into account; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption test.

Options for Signed Rank Test: Results

Summary Table. The summary table for a Signed Rank Test lists the medians, percentiles, and sample sizes N in the Rank Sum test report. If desired, change the percentile values by editing the boxes. The 25th and the 75th percentiles are the suggested percentiles.

Yates Correction Factor. When a statistical test uses a χ^2 distribution with one degree of freedom, such as analysis of a 2 x 2 contingency table or McNemar's test, the χ^2 calculated tends to produce P values which are too small, when compared with the actual distribution of the χ^2 test statistic. The theoretical χ^2 distribution is continuous, whereas the distribution of the χ^2 test statistic is discrete.

Use the Yates Correction Factor to adjust the computed χ^2 value down to compensate for this discrepancy. Using the Yates correction makes a test more conservative; for example, it increases the P value and reduces the chance of a false positive conclusion.

The Yates correction is applied to 2 x 2 tables and other statistics where the P value is computed from a χ^2 distribution with one degree of freedom.

Running a Signed Rank Test

To run a test, you need to select the data to test by dragging the pointer over your data. Then use the **Select Data** panel of the Test Wizard to select the worksheet columns with the data you want to test and to specify how your data is arranged in the worksheet.

To run a Signed Rank Test:

1. On the **Analysis** tab, in the **SigmaStat** group, from the tests drop-down list select;

Before and After > Signed Rank Test

The **Signed Rank Test — Data Format** panel of the Test Wizard appears prompting you to specify a data format.

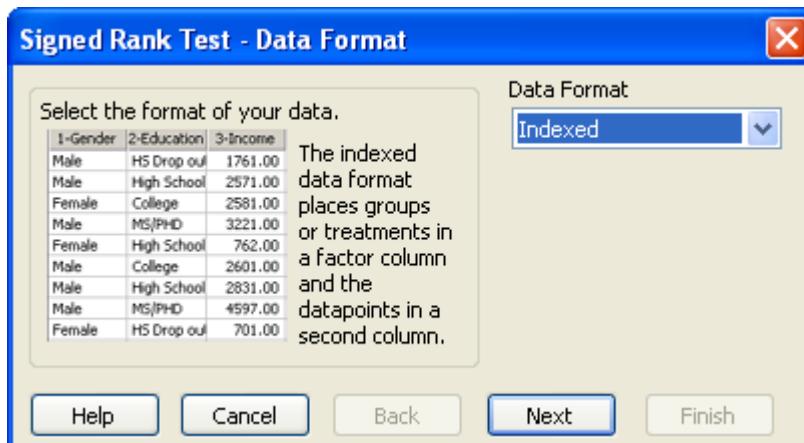


Figure 81: The Signed Rank Test — Data Format Panel of the Test Wizard Prompting You to Specify a Data Format

2. Select the appropriate data format from the **Data Format** drop-down list.

If your data is grouped in columns, select **Raw**. If your data is in the form of a group index column(s) paired with a data column(s), select **Indexed**.

3. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

4. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list.

The number or title of selected columns appear in each row. You are prompted to pick two columns for raw data and three columns for indexed data.

- To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
- Click **Finish** to perform the test. If you elected to test for normality, SigmaPlot performs the test for normality (Shapiro-Wilk or Kolmogorov-Smirnov). If your data pass the test, SigmaPlot informs you and suggests continuing your analysis using a Paired t-test.

When the test is complete, the report appears displaying the results of the Signed Rank Test.

Interpreting Signed Rank Test Results

The Signed Rank Test computes the Wilcoxon W statistic and the P value for W. Additional results to be displayed are selected in the **Options for Signed Rank Test** dialog box.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the **Options** dialog box. You can also set the number of decimal places to display in the **Options** dialog box.

Wilcoxon Signed Rank Test Thursday, November 03, 2011, 6:12:31 PM

Data source: Wilcoxon Signed Rank Test in SigmaPlot.JNB

Normality Test (Shapiro-Wilk) Passed (P = 0.226)

Test execution ended by user request, Paired t-test begun

Paired t-test: Thursday, November 03, 2011, 6:12:31 PM

Data source: Wilcoxon Signed Rank Test in SigmaPlot.JNB

Treatment Name	N	Missing	Mean	Std Dev	SEM
Before Dose	7	0	16.857	7.798	2.947
After Dose	7	0	12.143	9.335	3.528
Difference	7	0	4.714	11.842	4.476

t = 1.053 with 6 degrees of freedom.

95 percent two-tailed confidence interval for difference of means: -6.238 to 15.667

Two-tailed P-value = 0.333

The change that occurred with the treatment is not great enough to exclude the possibility that the difference is due to chance (P = 0.333)

One-tailed P-value = 0.166

The sample mean of treatment Before Dose does not exceed the sample mean of the treatment After Dose. The null hypothesis that the population mean of treatment After Dose is less than or equal to the population mean of treatment Before Dose cannot be rejected. (P = 0.333)

Power of performed two-tailed test with alpha = 0.050: 0.145

The power of the performed test (0.145) is below the desired power of 0.800. Less than desired power indicates you are less likely to detect a difference when one actually exists. Negative results should be interpreted cautiously.

Power of performed one-tailed test with alpha = 0.050: 0.239

The power of the performed test (0.239) is below the desired power of 0.800. Less than desired power indicates you are less likely to detect a difference when one actually exists. Negative results should be interpreted cautiously.

Figure 82: The Wilcoxon Signed Rank Test Results Report

Normality Test

Normality test results display whether the data passed or failed the test of the assumption that the difference of the treatment originates from a normal distribution, and the P value calculated by the test. For nonparametric procedures this test can fail, since nonparametric tests do not require normally distributed source populations. This result appears unless you disabled normality testing in the Options for Signed Rank Test dialog box.

Summary Tables

SigmaPlot generates a summary table listing the sample sizes N, number of missing values (if any), medians, and percentiles. All of these results are displayed in the report unless you disable them in the Signed Rank Test Options dialog box.

N (Size). The number of non-missing observations for that column or group.

Missing. The number of missing values for that column or group.

Medians. The "middle" observation as computed by listing all the observations from smallest to largest and selecting the largest value of the smallest half of the observations. The median observation has an equal number of observations greater than and less than that observation.

Percentiles. The two percentile points that define the upper and lower tails of the observed values.

W Statistic

The Wilcoxon test statistic W is computed by ranking all the differences before and after the treatment based on their absolute value, then attaching the signs of the difference to the corresponding ranks. The signed ranks are summed and compared.

If the absolute value of W is "large", you can conclude that there was a treatment effect (for example, the ranks tend to have the same sign, so there is a statistically significant difference before and after the treatment).

If W is small, the positive ranks are similar to the negative ranks, and you can conclude that there is no treatment effect.

P Value. The *P* value is the probability of being wrong in concluding that there is a true effect (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on W). The smaller the *P* value, the greater the probability that there is a treatment effect.

Traditionally, you can conclude there is a significant difference when $P < 0.05$.

Signed Rank Test Report Graphs

You can generate a line scatter graph of the changes after treatment for a Signed Rank Test report.

- **Before and After Line Graph.** The Signed Rank Test graph uses lines to plot a subject's change after each treatment.

Creating a Graph of the Signed Rank Test Data

1. Select the **Signed Rank Test** report.

2. On the **Report** tab in the **Result Graphs** group, click **Create Result Graph**.

The **Create Graph** dialog box appears displaying the types of graphs available for the Signed Rank Test results.

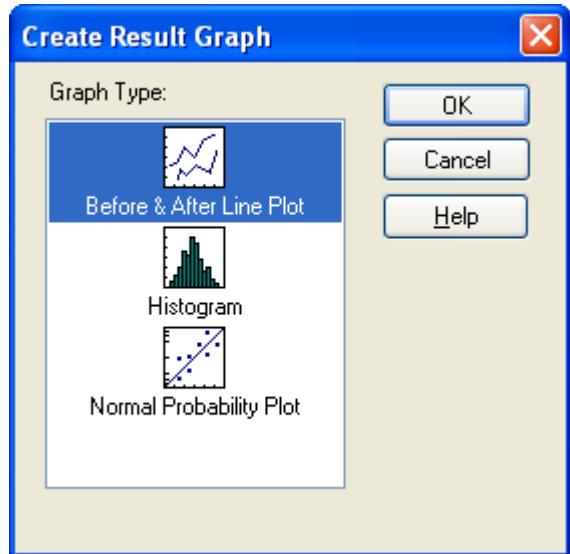


Figure 83: The Create Graph Dialog Box for the Signed Rank Test Report

3. Select the type of graph you want to create from the **Graph Type** list.

4. Click **OK**, or double-click the desired graph in the list. The specified graph appears in a graph window or in the report. For more information, see [Report Graphs](#) on page 373.

Before & After Line Plot

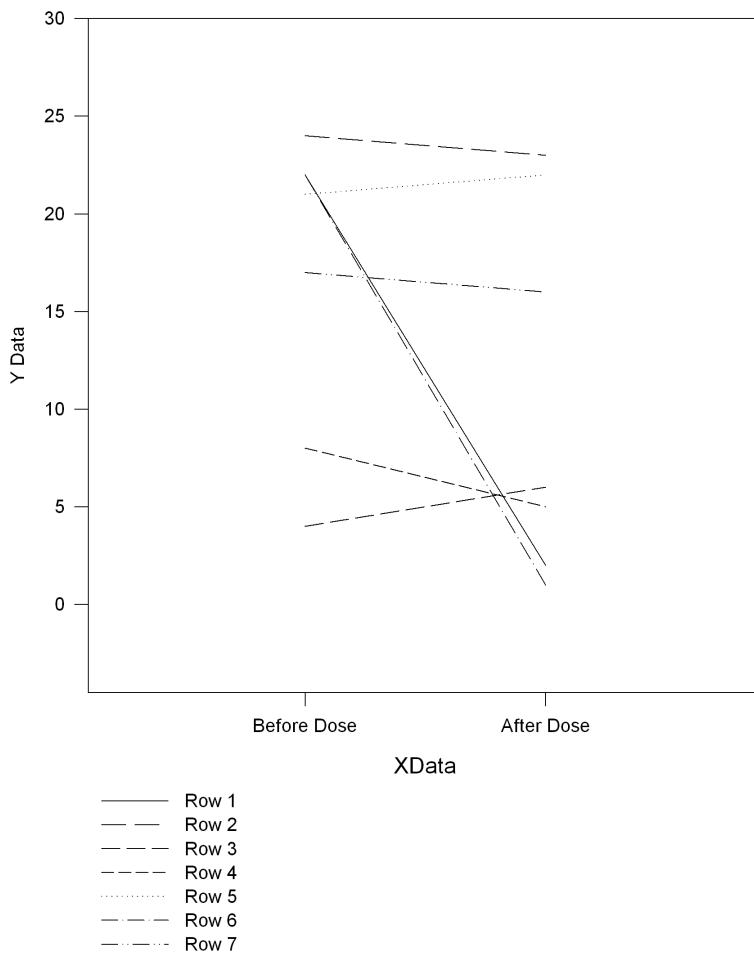


Figure 84: A Before & After Scatter Graph

One Way Repeated Measures Analysis of Variance (ANOVA)

Use a one way or one factor repeated measures ANOVA (analysis of variance) when:

- You want to see if a single group of individuals was affected by a series of experimental treatments or conditions.
- Only one factor or one type of intervention is considered in each treatment or condition.
- The treatment effects are normally distributed with the same variances.

If you know that the treatment effects are not normally distributed, use the Friedman Repeated Measures ANOVA on Ranks. If you want to consider the effects of an additional factor on your experimental treatments, use Two Way Repeated Measures ANOVA. When there is only a single treatment, you can do a Paired t-test (depending on the type of results you want).

- Tip:** Depending on your One Way Repeated Measures ANOVA options settings if you attempt to perform an ANOVA on a non-normal population, SigmaPlot informs you that the data is unsuitable for a parametric test, and suggest the Friedman ANOVA on Ranks instead.

About the One Way Repeated Measures ANOVA

A One Way or One Factor Repeated Measures ANOVA tests for differences in the effect of a series of experimental interventions on the same group of subjects by examining the changes in each individual. Examining the changes rather than the values observed before and after interventions removes the differences due to individual responses, producing a more sensitive (or more powerful) test.

The design for a One Way Repeated Measures ANOVA is essentially the same as a Paired t-test, except that there can be multiple treatments on the same group. The null hypothesis is that there are no differences among all the treatments.

One Way Analysis of Variance is a parametric test that assumes that all treatment effects are normally distributed with the same standard deviations (variances).

Performing a One Way Repeated Measures ANOVA

To perform a One Way Repeated Measures ANOVA:

1. [Enter or arrange your data in the worksheet.](#)
2. If desired, set One Way Repeated Measures ANOVA options.
3. On the **Analysis** tab, in the **SigmaStat** group, from the **Tests** drop-down list click select:
Repeated Measures > One Way Repeated Measures ANOVA
4. [Generate report graphs.](#)
5. Run the test.

Arranging One Way Repeated Measures ANOVA Data

The format of the data to be tested can be raw data or indexed data. Place raw data in as many columns as there are treatments, up to 64; each column contains the data for one treatment. The columns for raw data must be the same length.

Place Indexed data in two worksheet columns. You cannot use statistical summary data for repeated measures tests.

	1-Placebo	2-Hormone A	3-Hormone B	4-Subject	5-Treatment	6-6% Body Fat
1	32.0000	33.0000	21.0000	1.0000	Placebo	32.0000
2	33.0000	46.0000	31.0000	2.0000	Placebo	33.0000
3	45.0000	51.0000	32.0000	3.0000	Placebo	45.0000
4	61.0000	60.0000	44.0000	4.0000	Placebo	61.0000
5	19.0000	23.0000	17.0000	5.0000	Placebo	19.0000
6				1.0000	Hormone A	33.0000
7				2.0000	Hormone A	46.0000
8				3.0000	Hormone A	51.0000
9				4.0000	Hormone A	60.0000
10				5.0000	Hormone A	23.0000
11				1.0000	Hormone B	21.0000
12				2.0000	Hormone B	31.0000

Figure 85: Valid Data Formats for a One Way Repeated Measures ANOVA

Columns 1 through 3 in the worksheet above are arranged as raw data. Columns 4, 5, and 6 are arranged as indexed data, with column 4 as the treatment index column and column 5 as the subject index column.

Missing Data Points

If there are missing values, SigmaPlot automatically handles the missing data by using a general linear model. This approach constructs hypothesis tests using the marginal sums of squares (also commonly called the Type III or adjusted sums of squares); however, the columns must still be equal in length.

Setting One Way Repeated Measures ANOVA Options

Use the One Way Repeated Measures ANOVA options to:

- Adjust the parameters of the test to relax or restrict the testing of your data for normality and equal variance.
- Display the statistics summary table for the data, and assign residuals to a worksheet column.
- Enable multiple comparisons.
- Compute the power, or sensitivity, of the test.

To change the One Way Repeated Measures ANOVA options:

i **Tip:** If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.

1. On the **Analysis** tab, in the **SigmaStat** group, from the **Tests** drop-down list select: **Repeated Measures > One Way Repeated Measures ANOVA**
2. Click **Options** in the **SigmaStat** group. The **Options for One Way RM ANOVA** dialog box appears with three tabs:
 - **Assumption Checking.** Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
 - **Results.** Display the statistics summary for the data in the report and save residuals to a worksheet column.
 - **Post Hoc Test.** Compute the power or sensitivity of the test and enable multiple comparisons.
3. To continue the test, click **Run Test**.
4. To accept the current settings and close the options dialog box, click **OK**.

Options for One Way Repeated Measures ANOVA: Assumption Checking

The normality assumption test checks for a normally distributed population. The equal variance assumption test checks the variability about the group means.

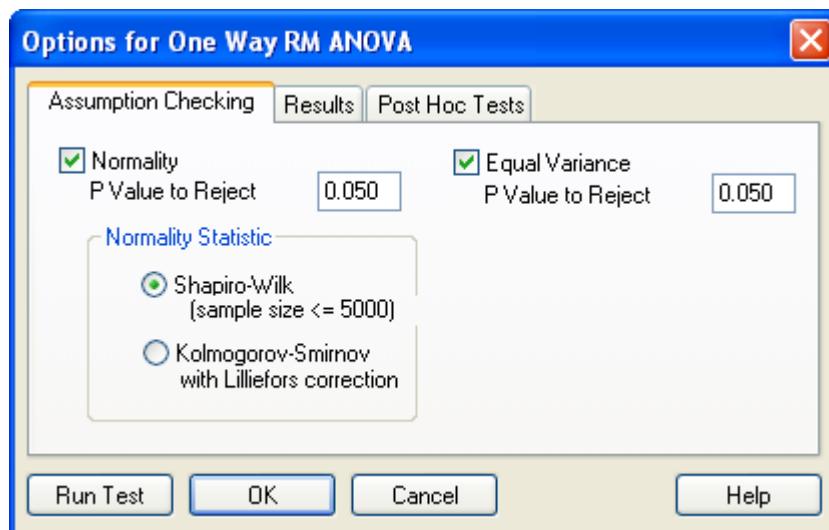


Figure 86: The Options for One Way RM ANOVA Dialog Box Displaying the Assumption Checking Options

- **Normality Testing.** SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

- **Equal Variance Testing.** SigmaPlot tests for equal variance by checking the variability about the group means.
- **P Values for Normality and Equal Variance.** The *P* value determines the probability of being incorrect in concluding that the data is not normally distributed (*P* value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the *P* computed by the test is greater than the *P* set here, the test passes.

To require a stricter adherence to normality and/or equal variance, increase the *P* value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of *P* (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality and/or equal variance, decrease *P*. Requiring larger values of *P* to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a *P* value of 0.010 requires greater deviations from normality to flag the data as non-normal than a value of 0.050.

 **Note:** There are extreme conditions of data distribution that these tests cannot take into account. For example, the Levene Median test fails to detect differences in variance of several orders of magnitude; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options for One Way RM ANOVA: Results

Summary Table. Select to display the number of observations for a column or group, the number of missing values for a column or group, the average value for the column or group, the standard deviation of the column or group, and the standard error of the mean for the column or group.

Residuals in Column. Select to display residuals in the report and to save the residuals of the test to the specified worksheet column. Edit the number or select a number from the drop-down list.

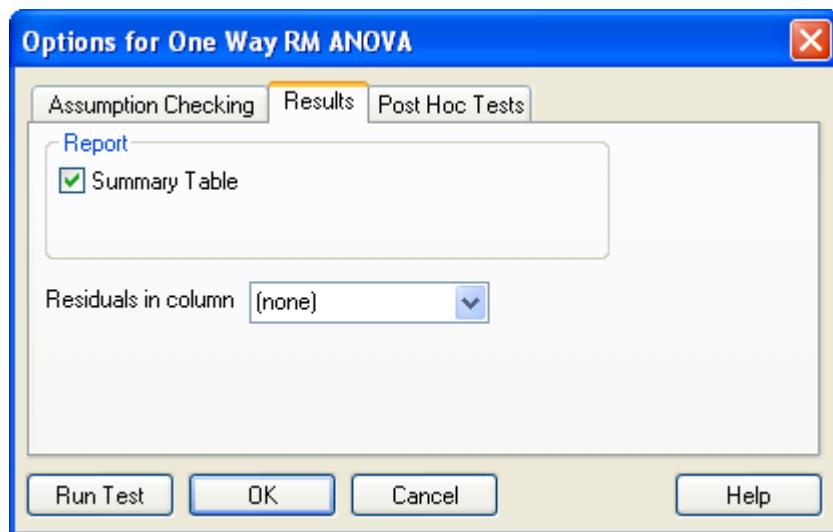


Figure 87: The Options for One Way ANOVA Dialog Box Displaying the Summary Table Options

Options for One Way RM ANOVA: Post Hoc Tests

Power. The power or sensitivity of a test is the probability that the test will detect a difference between the groups if there is really a difference.

Use Alpha Value. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

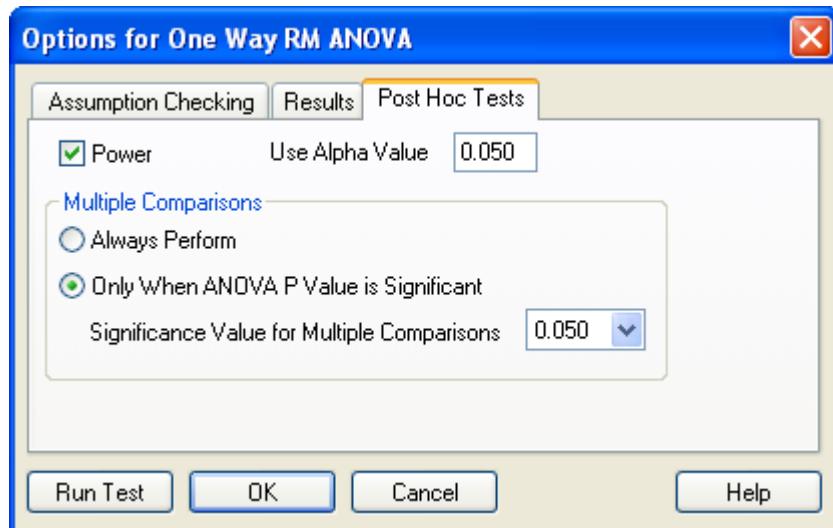


Figure 88: The Options for One Way ANOVA Dialog Box Displaying the Power and Multiple Comparison Options

Multiple Comparisons

A One Way Repeated Measures ANOVA tests the hypothesis of no differences between the several treatment groups, but does not determine which groups are different, or the sizes of these differences. Multiple comparison procedures isolate these differences.

The P value used to determine if the ANOVA detects a difference is set on the **Report** tab of the **Options** dialog box. If the P value produced by the One Way ANOVA is less than the P value specified in the box, a difference in the groups is detected and the multiple comparisons are performed. For more information, see [report.dita](#).

- **Always Perform.** Select to perform multiple comparisons whether or not the ANOVA detects a difference.
- **Only When ANOVA P Value is Significant.** Select to perform multiple comparisons only if the ANOVA detects a difference.
- **Significance Value for Multiple Comparisons.** Select either .05 or .01 from the Significance Value for Multiple Comparisons drop-down list. This value determines the that the likelihood of the multiple comparison being incorrect in concluding that there is a significant difference in the treatments.

A value of .05 indicates that the multiple comparisons will detect a difference if there is less than 5% chance that the multiple comparison is incorrect in detecting a difference. A value of .10 indicates that the multiple comparisons will detect a difference if there is less than 10% chance that the multiple comparison is incorrect in detecting a difference.

 **Note:** If multiple comparisons are triggered, the **Multiple Comparison Options** dialog box appears after you pick your data from the worksheet and run the test, prompting you to choose a multiple comparison method.

Running a One Way Repeated Measures ANOVA

If you want to select your data before you run the test, drag the pointer over your data.

1. On the **Analysis** tab, in the **SigmaStat** group, from the **Tests** drop-down list select:

Repeated Measures > One Way Repeated Measures ANOVA

The **One Way RM ANOVA — Data Format** panel of the Test Wizard appears prompting you to specify a data format.

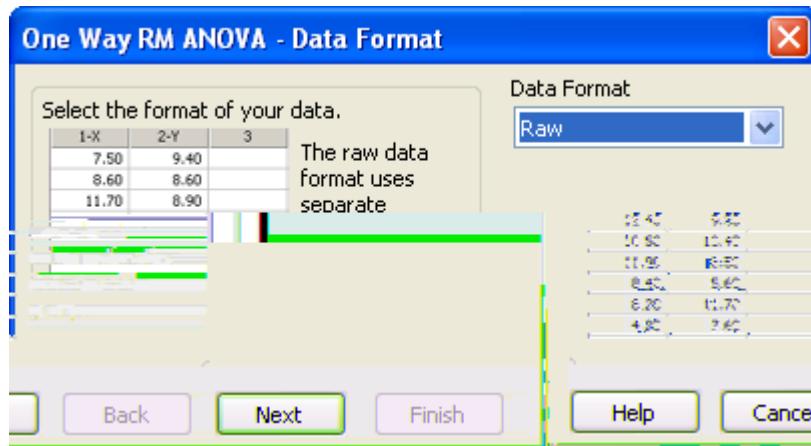


Figure 89: The One Way RM ANOVA — Data Format Panel of the Test Wizard Prompting You to Specify a Data Format

2. Select the appropriate **data format** from the **Data Format** drop-down list.
3. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

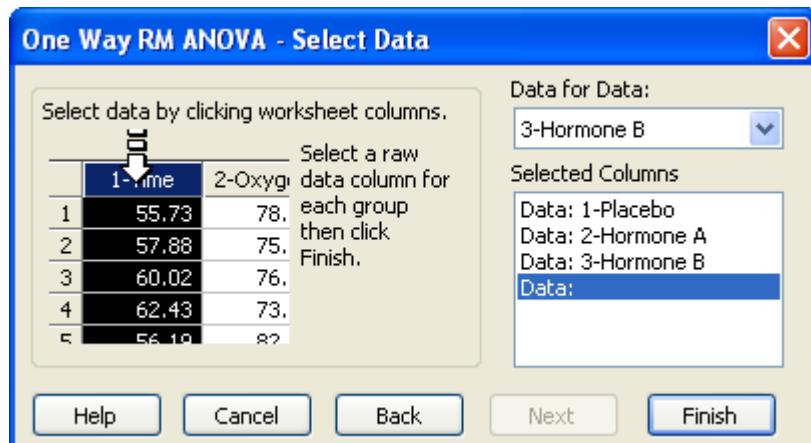


Figure 90: The One Way RM ANOVA — Select Data Panel of the Test Wizard Prompting You to Select Data Columns

4. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw and indexed data, you are prompted to select two worksheet columns.

5. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.

6. Click **Finish** to run the **One Way RM ANOVA** on the selected columns.

If you elected to test for normality and equal variance, and your data fails either test, SigmaPlot warns you and suggests continuing your analysis using the nonparametric [Friedman Repeated Measures ANOVA on Ranks](#).

If you selected to run multiple comparisons only when the P value is significant, and the P value is not significant, the [One Way ANOVA report](#) appears after the test is complete.

If the P value for multiple comparisons is significant, or you selected to always perform multiple comparisons, the [Multiple Comparisons Options](#) dialog box appears prompting you to select a multiple comparison method.

Multiple Comparison Options (One Way RM ANOVA)

The One Way Repeated Measures ANOVA tests the hypothesis of no differences between the several treatment groups, but do not determine which groups are different, or the sizes of these differences. Multiple comparison tests isolate these differences by running comparisons between the experimental groups.

If you selected to run multiple comparisons only when the P value is significant, and the ANOVA produces a P value equal to or less than the trigger P value, or you selected to always run multiple comparisons in the Options for One Way RM ANOVA dialog box, the Multiple Comparison Options dialog box appears prompting you to specify a multiple comparison test. The P value produced by the ANOVA is displayed in the upper left corner of the dialog box. For more information, see [Interpreting One Way Repeated Measures ANOVA Results](#) on page 163.

There are seven kinds of multiple comparison tests available for the One Way Repeated Measures ANOVA, including:

- [Holm-Sidak Test](#).
- [Tukey Test](#).
- [Student-Newman-Keuls Test](#).
- [Bonferroni t-test](#).
- [Fisher's LSD](#).
- [Dunnett's Test](#).
- [Duncan's Multiple Range Test](#).

There are two types of multiple comparisons available for the One Way Repeated Measures ANOVA. The types of comparison you can make depends on the selected multiple comparison test. The tests are:

- All pairwise comparisons compare all possible pairs of treatments.
- Multiple comparisons versus a control compare all experimental treatments to a single control group.

Interpreting One Way Repeated Measures ANOVA Results

The One Way Repeated Measures ANOVA report generates an ANOVA table describing the source of the variation in the treatments. This table displays the degrees of freedom, sum of squares, and mean squares of the treatments, as well as the F statistic and the corresponding P value. The other results displayed are in the Options for One Way RM ANOVA dialog box.

You can also generate tables of multiple comparisons. Multiple Comparison results are also specified in the Options for One Way RM ANOVA dialog box. The test used to perform the multiple comparison is selected in the Multiple Comparison Options dialog box.

For descriptions of the derivations for One Way RM ANOVA results, you can reference any appropriate statistics reference.

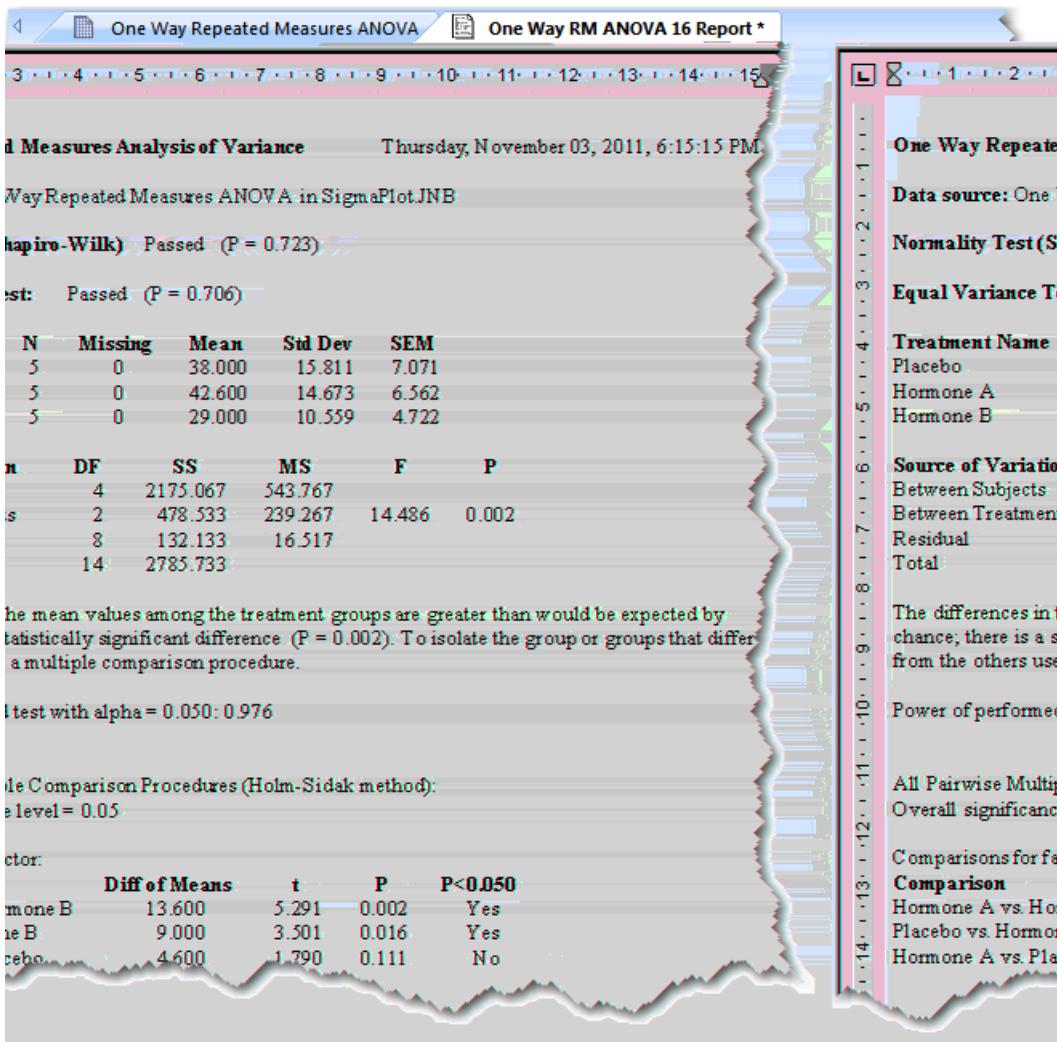


Figure 91: Example of the One Way Repeated Measures ANOVA Report

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

If There Were Missing Data Cells

If your data contained missing values, the report indicates the results were computed using a general linear model. The ANOVA table includes the degrees of freedom used to compute F, the estimated mean square equations are listed, and the summary table displays the estimated least square means.

For descriptions of the derivations for One Way Repeated Measures ANOVA results, you can reference an appropriate statistics reference.

Normality Test

Normality test results display whether the data passed or failed the test of the assumption that the differences of the changes originate from a normal distribution, and the P value calculated by the test. Normally distributed source populations are required for all parametric tests.

This result appears unless you disabled equal variance testing in the Options for One Way RM ANOVA dialog box.

Equal Variance Test

Equal Variance test results display whether or not the data passed or failed the test of the assumption that the differences of the changes originate from a population with the same variance, and the P value calculated by the test. Equal variances of the source populations are assumed for all parametric tests.

This result appears unless you disabled equal variance testing in the Options for One Way RM ANOVA dialog box.

Summary Table

If you enabled this option in the Options for One Way RM ANOVA dialog box, SigmaPlot generates a summary table listing the sample sizes N, number of missing values, mean, standard deviation, differences of the means and standard deviations, and standard error of the means.

- **N (Size).** The number of non-missing observations for that column or group.
- **Missing.** The number of missing values for that column or group.
- **Mean.** The average value for the column. If the observations are normally distributed the mean is the center of the distribution.
- **Standard Deviation.** A measure of variability. If the observations are normally distributed, about two-thirds will fall within one standard deviation above or below the mean, and about 95% of the observations will fall within two standard deviations above or below the mean.
- **Standard Error of the Mean.** A measure of the approximation with which the mean computed from the sample approximates the true population mean.

Power

The power of the performed test is displayed unless you disable this option in the Options for One Way RM ANOVA dialog box.

The power, or sensitivity, of a One Way Repeated Measures ANOVA is the probability that the test will detect a difference among the treatments if there really is a difference. The closer the power is to 1, the more sensitive the test.

Repeated measures ANOVA power is affected by the sample sizes, the number of treatments being compared, the chance of erroneously reporting a difference α (alpha), the observed differences of the group means, and the observed standard deviations of the samples.

Alpha (α). Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. An α error is also called a Type I error. A Type I error is when you reject the hypothesis of no effect when this hypothesis is true.

Set this value in the Options for One Way RM ANOVA dialog box; the suggested value is $\alpha = 0.05$ which indicates that a one in twenty chance of error is acceptable. Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference but also increase the risk of seeing a false difference (a Type I error).

ANOVA Table

The ANOVA table lists the results of the One Way Repeated Measures ANOVA.

DF (Degrees of Freedom). Degrees of freedom represent the number of groups and sample size which affects the sensitivity of the ANOVA.

- The degrees of freedom between subjects is a measure of the number of subjects
- The degrees of freedom within subjects is a measure of the total number of observations, adjusted for the number of treatments
- The degrees of freedom for the treatments is a measure of the number of treatments
- The residual degrees of freedom is a measure of the difference between the number of observations, adjusted for the number of subjects and treatments
- The total degrees of freedom is a measure of both number of subjects and treatments

SS (Sum of Squares). The sum of squares is a measure of variability associated with each element in the ANOVA data table.

- The sum of squares between the subjects measures the variability of the average responses of each subject.
- The sum of squares within the subjects measures the underlying total variability within each subject.
- The sum of squares of the treatments measures the variability of the mean treatment responses within the subjects.
- The residual sum of squares measures the underlying variability among all observations after accounting for differences between subjects.
- The total sum of squares measures the total variability.

MS (Mean Squares). The mean squares provide two estimates of the population variances. Comparing these variance estimates is the basis of analysis of variance.

The mean square of the treatments is:

$$\frac{\text{sum of squares between groups}}{\text{degrees of freedom between groups}} = \frac{SS_{\text{between}}}{DF_{\text{between}}} = MS_{\text{between}}$$

The residual mean square is

$$\frac{\text{sum of squares within groups}}{\text{degrees of freedom within groups}} = \frac{SS_{\text{within}}}{DF_{\text{within}}} = MS_{\text{within}}$$

F Statistic

The F test statistic is a ratio used to gauge the differences of the effects. If there are no missing data, F is calculated as:

$$\frac{\text{estimated population variance between groups}}{\text{estimated population variance within groups}} = \frac{MS_{\text{between}}}{MS_{\text{within}}} = F$$

If the F ratio is around 1, you can conclude that there are no differences among treatments (the data is consistent with the null hypothesis that there are no treatment effects).

If F is a large number, the variability among the effect means is larger than expected from random variability in the treatments, you can conclude that the treatments have different effects (the differences among the treatments are statistically significant).

P Value. The P value is the probability of being wrong in concluding that there is a true difference between the groups (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on F). The smaller the P value, the greater the probability that the samples are drawn from different populations. Traditionally, you can conclude that there are significant differences when $P < 0.05$.

Expected Mean Squares

If there was missing data and a general linear model was used, the linear equations for the expected mean squares computed by the model are displayed. These equations are displayed only if a general linear model was used.

Multiple Comparisons

If you selected to perform multiple comparisons, a table of the comparisons between group pairs is displayed. For more information, see [Multiple Comparison Options \(One Way RM ANOVA\)](#) on page 163. The multiple comparison procedure is activated in the Options for One Way RM ANOVA dialog box. The tests used in the multiple comparison procedure is selected in the Multiple Comparison Options dialog box.

Multiple comparison results are used to determine exactly which treatments are different, since the ANOVA results only inform you that two or more of the groups are different. The specific type of multiple comparison results depends on the comparison test used and whether the comparison was made pairwise or versus a control.

- All pairwise comparison results list comparisons of all possible combinations of group pairs; the all pairwise tests are the Holm Sidak, Tukey, Student-Newman-Keuls, Fisher LSD, Duncan's test and the Bonferroni t-test.
- Comparisons versus a single control group list only comparisons with the selected control group. The control group is selected during the actual multiple comparison procedure. The comparison versus a control tests are the Bonferroni t-test and the Dunnett's, Fishers LSD, and Duncan's tests.

For descriptions of the derivation of parametric multiple comparison procedure results, you can reference an appropriate statistics reference.

Holm-Sidak Test Results. The Holm-Sidak Test can be used for both pairwise comparisons and comparisons versus a control group. It is more powerful than the Tukey and Bonferroni tests and, consequently, it is able to detect differences that these other tests do not. It is recommended as the first-line procedure for pairwise comparison testing.

When performing the test, the P values of all comparisons are computed and ordered from smallest to largest. Each P value is then compared to a critical level that depends upon the significance level of the test (set in the test options), the rank of the P value, and the total number of comparisons made. A P value less than the critical level indicates there is a significant difference between the corresponding two groups.

Bonferroni t-test Results. The Bonferroni t-test lists the differences of the means for each pair of groups, computes the t values for each pair, and displays whether or not $P < 0.05$ for that comparison. The Bonferroni t-test can be used to compare all groups or to compare versus a control.

You can conclude from "large" values of t that the difference of the two treatments being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of erroneously concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The difference of the means is a gauge of the size of the difference between the two treatments.

Tukey, Student-Newman-Keuls, Fisher LSD, Duncan's, and Dunnett's Test Results. The Tukey, Student-Newman-Keuls (SNK), Fisher LSD, and Duncan's tests are all pairwise comparisons of every combination of group pairs. While the Tukey, Fisher LSD, and Duncan's can be used to compare a control group to other groups, they are not recommended for this type of comparison.

Dunnett's test only compares a control group to all other groups. All tests compute the θ test statistic, and display whether or not $P < 0.05$ or < 0.01 for that pair comparison.

You can conclude from "large" values of θ that the difference of the two groups being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of being incorrect in concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The Difference of the Means is a gauge of the size of the difference between the two groups.

ϕ is a parameter used when computing θ . The larger the ϕ , the larger θ needs to be to indicate a significant difference. ϕ is an indication of the differences in the ranks of the group means being compared. Groups means are ranked in order from largest to smallest in an SNK test, so ϕ is the number of means spanned in the comparison. For example, when comparing four means, comparing the largest to the smallest $\phi = 4$, and when comparing the second smallest to the smallest $\phi = 2$.

If a treatment is found to be not significantly different than another treatment, all treatments with ϕ ranks in between the ϕ ranks of the two treatments that are not different are also assumed not to be significantly different, and a result of DNT (Do Not Test) appears for those comparisons.

One Way Repeated Measures ANOVA Report Graphs

You can generate up to three graphs using the results from a One Way RM ANOVA. They include a:

- **Before and after line graph.** The One Way Repeated Measures ANOVA uses lines to plot a subject's change after each treatment.
- **Histogram of the residuals.** The One Way Repeated Measures ANOVA histogram plots the raw residuals in a specified range, using a defined interval set.
- **Normal probability plot of the residuals.** The One Way Repeated Measures ANOVA probability plot graphs the frequency of the raw residuals.
- **Multiple comparison graphs.** The One Way Repeated Measures ANOVA multiple comparison graphs a plot significant differences between levels of a significant factor.

Creating a One Way Repeated Measures ANOVA Report Graph

1. Select the One Way Repeated Measures ANOVA test report.
2. On the **Report** tab, in the **Result Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the One Way Repeated Measure ANOVA results.

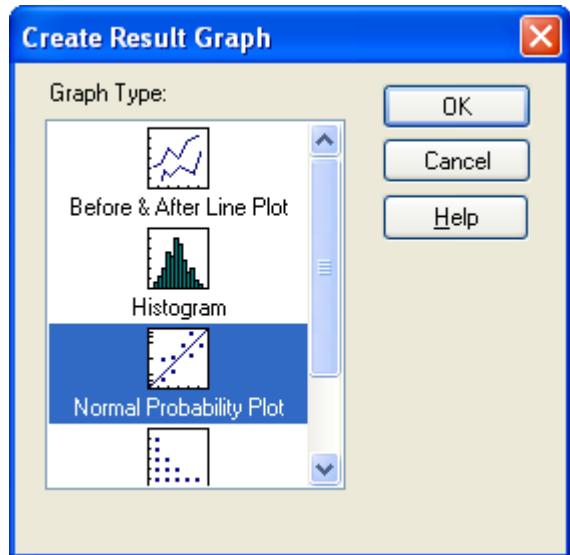


Figure 92: The Create Graph Dialog Box for a One Way RM ANOVA Report

3. Select the type of graph you want to create from the **Graph Type** list, then click **OK**, or double-click the desired graph in the list.

The selected graph appears in a graph window.

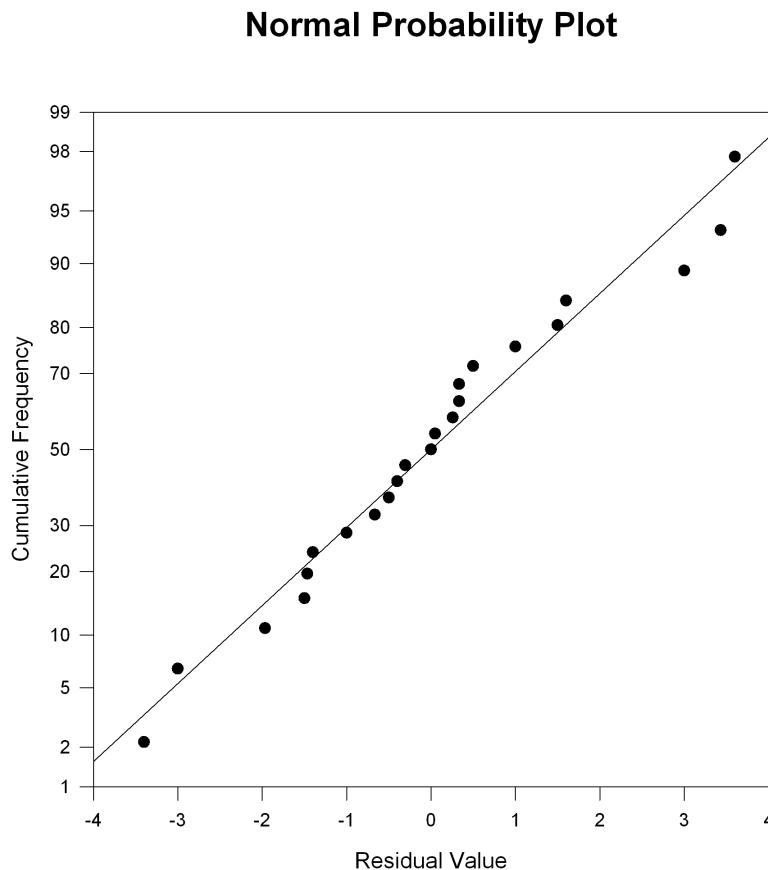


Figure 93: A Normal Probability Plot for a One Way RM ANOVA

Two Way Repeated Measures Analysis of Variance (ANOVA)

Use Two Way or two factor Repeated Measures ANOVA (analysis of variance) when:

- You want to see if the same group of individuals are affected by a series of experimental treatments or conditions.
- You want to consider the effect of an additional factor which may or may not interact, and may or may not be another series of treatments or conditions.
- The treatment effects are normally distributed with equal variances.



Note: SigmaPlot performs Two Way Repeated Measures ANOVAs for one factor repeated or both factors repeated. SigmaPlot automatically determines if one or both factors are repeated from the data, and uses the appropriate procedures.

If you want to consider the effects of only one factor on your experimental groups, use [One Way Repeated Measures ANOVA](#).

There is no equivalent in SigmaPlot for a two factor repeated measure comparison for samples drawn from a non-normal populations. If your data is non-normal, you can transform the data to make it comply better with the assumptions of analysis of variance using transforms. If the sample size is large, and you want to do a nonparametric test, use **Rank** transform (available in the **Transform** group on the **Analysis** tab) to convert the observations to ranks, then do a Two Way ANOVA on the ranks.

About the Two Way Repeated Measures ANOVA

In a two way or two factor repeated measures analysis of variance, there are two experimental factors which may affect each experimental treatment. Either or both of these factors are repeated treatments on the same group of individuals. A two factor design tests for differences between the different levels of each treatment and for interactions between the treatments. For more information, see [Arranging Two Way Repeated Measures ANOVA Data](#) on page 170.

A two factor analysis of variance tests three hypotheses: (1) There is no difference among the levels or treatments of the first factor; (2) There is no difference among the levels or treatments of the second factor; and (3) There is no interaction between the factors, for example, if there is any difference among treatments within one factor, the differences are the same regardless of the second factor.

Two Way Repeated Measures ANOVA is a parametric test that assumes that all the treatment effects are normally distributed with the same variance. SigmaPlot does not have an automatic nonparametric test if these assumptions are violated.

Performing a Two Way Repeated Measures ANOVA

To perform a Two Way Repeated Measures ANOVA:

1. [Enter or arrange your data in the data worksheet.](#)
2. Set the Two Way Repeated Measures ANOVA options.
3. On the **Analysis** tab, in the **SigmaStat** group, from the **Tests** drop-down list select:
Repeated Measures > Two Way Repeated Measures ANOVA
4. [Generate report graphs.](#)
5. Run the test.

Arranging Two Way Repeated Measures ANOVA Data

Either or both of the two factors used in the Two Way Repeated Measures ANOVA can be repeated on the same group of individuals. For example, if you analyze the effect of changing salinity on the activity of two different species of shrimp, you have a two factor experiment with a *single repeated treatment* (salinity). Different salinity treatment and shrimp type are the levels.

Table 12: Data for a Two Way Repeated Factor ANOVA with one repeated factor (salinity)

Species	Subject	Salinity
		10
<i>Artemia sp. 1</i>	ABC	108.59.0
<i>Artemia sp. 2</i>	DEF	5.57.57.0
		12.513.010.5
		5.58.06.5

If you wanted to test the effect of different salinities and temperatures on the activity on a single species of shrimp, you have a two factor experiment with two repeated treatments, salinity and temperature. In both cases, the different combinations of treatments/factors levels are the cells of the comparison. SigmaPlot automatically handles both one and two repeated treatment factors.

Table 13: Data for a Two Way Repeated Factor ANOVA with two repeated factors (temperature and salinity)

Temperature	Subject	Salinity
		10
25°	ABC	8.58.59.5
30°	DEF	9.09.010.0
		11.010.512.0
		12.511.513.0

Missing Data and Empty Cells

Ideally, the data for a Two Way ANOVA should be completely balanced, for example, each group or cell in the experiment has the same number of observations and there are no missing data. However, SigmaPlot properly handles all occurrences of missing and unbalanced data automatically.

Missing Data Point(s). If there are missing values, SigmaPlot automatically handles the missing data by using a general linear model. This approach constructs a hypothesis tests using the marginal sums of squares (also commonly called the Type III or adjusted sums of squares).

Table 14: Data for a Two Way Repeated Factor ANOVA with one repeated factor (salinity) and a missing data point

Species	Subject	Salinity
		10
<i>Artemia sp. 1</i>	ABC	108.59.0
<i>Artemia sp. 2</i>	DEF	5.57.57.0
		5.5- -6.5

SigmaPlot uses a general linear model to handle missing data points.

Empty Cell(s). When there is an empty cell, for example, there are no observations for a combination of two factor levels, but there is still at least one repeated factor for every subject, SigmaPlot stops and suggests either analysis of the data assuming no interaction between the factors, or using One Way ANOVA.

Assumption of no interaction analyzes the effects of each treatment separately.

 **DANGER:** Assuming there is no interaction between the two factors in Two Way ANOVA can be dangerous. Under some circumstances, this assumption can lead to a meaningless analysis, particularly if you are interested in studying the interaction effect.

Table 15: Data for a Two Way Repeated Factor ANOVA with two repeated factors (temperature and salinity) and a missing cell

Temperature	Subject	Salinity
		10
25°	ABC	8.58.59.5
30°	DEF	9.09.010.0

Data with missing cells that still have repeated factor data for every subject can be analyzed either by assuming no interaction or a One Way ANOVA.

If you treat the problem as One Way ANOVA, each cell in the table is treated as a different level of a single experimental factor. This approach is the most conservative analysis because it requires no additional assumptions about the nature of the data or experimental design.

Connected versus Disconnected Data

The no interaction assumption requires that the non-empty cells must be geometrically connected in order to do the computation of a two factor no interaction model. You cannot perform Two Way Repeated Measures ANOVA on data disconnected by empty cells.

Table 16: Data for a Two Way Repeated Factor ANOVA with geometrically disconnected data.

Temperature	Subject	Salinity
		10
		15

25°	ABC	- - - -	11.010.512.0
30°	DEF	9.09.010.0	- - - -

This data cannot be analyzed with a Two Way Repeated Measures ANOVA.

When the data is geometrically connected, you can draw a series of straight vertical and horizontal lines connecting all cells containing data without changing direction in any empty cells. SigmaPlot automatically checks for this condition. If disconnected data is encountered during Two Way Repeated Measures ANOVA, SigmaPlot suggests treatment of the problem as a One Way Repeated Measures ANOVA.

For descriptions of the concept of connectivity, you can reference an appropriate statistics reference.

Missing Factor Data for One Subject

Another case of an empty cell can occur when both factors are repeated, and there are no data for one level for one of the subjects. SigmaPlot automatically handles this situation by converting the problem to a One Way Repeated Measures ANOVA.

Table 17: Data for a Two Way Repeated Factor ANOVA with two factors repeated and no data for one level for a subject.

Temperature	Subject	Salinity
25°	ABC	10
30°	DEF	9.09.010.0

This data cannot be analyzed as a Two Way Repeated Measures ANOVA problem.

Entering Worksheet Data

You can only perform a Two Way Repeated Measures ANOVA on data indexed by both subject and two factors. The data is placed in four columns; the first factor is in one column, the second factor is in a second column, the subject index is in a third column, and the actual data is in a fourth column.

 **Note:** SigmaPlot performs two way repeated measures for one factor repeated or both factors repeated. SigmaPlot automatically determines if one or both factors are repeated from the data, and uses the appropriate procedures.

Setting Two Way Repeated Measures ANOVA Options

Use the Two Way Repeated Measures ANOVA to:

- Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
- Display the statistics summary table for the data and assign residuals to the worksheet.
- Compute the power, or sensitivity, of the test.
- Enable multiple comparison testing.

To change the Two Way Repeated Measures ANOVA options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. On the **Analysis** tab, in the **SigmaStat** group, from the **Tests** drop-down list select:

Repeated Measures > Two Way Repeated Measures ANOVA

3. Click **Options**. The **Options for Two Way RM ANOVA** dialog box appears with three tabs:
 - **Assumption Checking.** Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
 - **Results.** Display the statistics summary for the data in the report and save residuals to a worksheet column.
 - **Post Hoc Test.** Compute the power or sensitivity of the test and enable multiple comparisons.
4. To continue the test, click **Run Test**.
5. To accept the current settings and close the options dialog box, click **OK**.

Options for Two Way Repeated Measures ANOVA: Assumption Checking

Click the **Assumption Checking** tab to view options for normality and equal variance. The normality assumption test checks for a normally distributed population. The equal variance assumption test checks the variability about the group means.

Normality Testing. SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

Equal Variance Testing. SigmaPlot tests for equal variance by checking the variability about the group means.

P Values for Normality and Equal Variance. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (the P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P value computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and/or equal variance, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality and/or equal variance, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.01 for the normality test requires greater deviations from normality to flag the data as non-normal than a value of 0.05.

 **Note:** Although the assumption tests are robust in detecting data from populations that are non-normal or with unequal variances, there are extreme conditions of data distribution that these tests cannot take into account. For example, the Levene Median test fails to detect differences in variance of several orders of magnitude; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options for Two Way Repeated Measures ANOVA: Results

Click the **Results** tab to view options for the summary table and residuals.

Summary Table. Select **Summary Table** to display the number of observations for a column or group, the number of missing values for a column or group, the average value for the column or group, the standard deviation of the column or group, and the standard error of the mean for the column or group.

Residuals. Select **Residuals** to display residuals in the report and to save the residuals of the test to the specified worksheet column. To change the column the residuals are saved to, edit the number in or select a number from the drop-down list.

Options for Two Way Repeated Measures ANOVA: Post Hoc Tests

Click the **Post Hoc Tests** tab to view options for power and multiple comparisons.

Power. The power or sensitivity of a test is the probability that the test will detect a difference between the groups if there is really a difference.

Alpha (α). Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

Multiple Comparisons

The Two Way Repeated Measures ANOVA tests the hypothesis of no differences between the several treatment groups, but do not determine which groups are different, or the sizes of these differences. Multiple comparison procedures isolate these differences.

The P value used to determine if the ANOVA detects a difference is set in the Report Options dialog box. If the P value produced by the Two Way RM ANOVA is less than the P value specified in the box, a difference in the groups is detected and the multiple comparisons are performed.

Performing Multiple Comparisons. You can choose to always perform multiple comparisons or to only perform multiple comparisons if a Two Way Repeated Measures ANOVA detects a difference.

Select Always Perform to perform multiple comparisons whether or not the ANOVA detects a difference.

Select Only When ANOVA P Value is Significant to perform multiple comparisons only if the ANOVA detects a difference.

Significant Multiple Comparison Value. Select either .05 or .10 from the Significance Value for Multiple Comparisons drop-down list. This value determines the that the likelihood of the multiple comparison being incorrect in concluding that there is a significant difference in the treatments.

A value of .05 indicates that the multiple comparisons will detect a difference if there is less than 5% chance that the multiple comparison is incorrect in detecting a difference. A value of .10 indicates that the multiple comparisons will detect a difference if there is less than 10% chance that the multiple comparison is incorrect in detecting a difference.

 **Note:** If multiple comparisons are triggered, the Multiple Comparison Options dialog box appears after you pick your data from the worksheet and run the test, prompting you to choose a multiple comparison method.

Running a Two Way Repeated Measures ANOVA

To run a test, you need to select the data to test. If you want to select your data before you run the test, drag the pointer over your data.

1. On the **Analysis** tab, in the **SigmaStat** group, from the **Tests** drop-down list select:

Repeated Measures > Two Way Repeated Measures ANOVA

The **Data Format** panel of the Test Wizard appears prompting you to specify a data format.

2. Select the appropriate data format from the **Data Format** drop-down list. For more information, see [Data Format for Repeated Measures Tests](#) on page 140.
3. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.
4. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw and indexed data, you are prompted to select two worksheet columns.

5. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
6. Click **Finish** to run the **Two Way RM ANOVA** on the selected columns.
7. If you elected to test for normality and equal variance, SigmaPlot performs the test for normality (Shapiro-Wilk or Kolmogorov-Smirnov) and the test for equal variance (Levene Median). If your data fail either test, SigmaPlot informs you. You can either continue, or transform your data, then perform a Two Way Repeated Measures ANOVA on the transformed data.

8. If your data have empty cells, you are prompted to perform the appropriate procedure.
 - If you are missing a cell, but the data is still connected, you may have to proceed by either assuming no interaction between the factors, or by performing a one factor analysis on each cell.
 - If your data is not geometrically connected, or if a subject is missing data for one level, you cannot perform a Two Way Repeated Measures ANOVA. Continue using a One Way ANOVA, or cancel the test.
 - If you are missing a few data points, but there is still at least one observation in each cell, SigmaPlot automatically proceeds. For more information, see [Arranging Two Way Repeated Measures ANOVA Data](#) on page 170.
9. If you selected to run multiple comparisons only when the P value is significant, and the P value is not significant the One Way ANOVA report appears after the test is complete.
If the P value for multiple comparisons is significant, or you selected to always perform multiple comparisons, the Multiple Comparisons Options dialog box appears prompting you to select a multiple comparison method.

Multiple Comparison Options (Two Way Repeated Measures ANOVA)

The Two Way Repeated Measures ANOVA tests the hypothesis of no differences between the several treatment groups, but do not determine which groups are different, or the sizes of these differences. Multiple comparison tests isolate these differences by running comparisons between the experimental groups.

If you selected to run multiple comparisons only when the P value is significant, and the ANOVA produces a P value equal to or less than the trigger P value, or you selected to always run multiple comparisons in the Options for Two Way RM ANOVA dialog box the Multiple Comparison Options dialog appears prompting you to specify a multiple comparison test. The P value produced by the ANOVA is displayed in the upper left corner of the dialog box.

There are six multiple comparison tests to choose from for the Two Way Repeated Measures ANOVA. You can choose to perform the:

- [Holm-Sidak Test](#).
- [Tukey Test](#).
- [Student-Newman-Keuls Test](#).
- [Bonferroni t-test](#).
- [Fisher's LSD](#).
- [Dunnet's Test](#).
- [Duncan's Multiple Range Test](#).

There are two types of multiple comparisons available for the Two Way Repeated Measures ANOVA. The types of comparison you can make depends on the selected multiple comparison test.

- All pairwise comparisons compare all possible pairs of treatments.
- Multiple comparisons versus a control compare all experimental treatments to a single control group.

When comparing the two factors separately, the treatments within one factor are compared among themselves without regard to the second factor, and vice versa. These results should be used when the interaction is not statistically significant.

When the interaction is statistically significant, interpreting multiple comparisons among different levels of each experimental factor may not be meaningful. SigmaPlot also performs a multiple comparison between all the cells.

The result of both comparisons is a listing of the similar and different treatment pairs, for example, those treatments that are and are not different from each other. Because no statistical test eliminates uncertainty, multiple comparison procedures sometimes produce ambiguous groupings.

Interpreting Two Way Repeated Measures ANOVA Results

A Two Way Repeated Measures ANOVA of one repeated factor generates an ANOVA table describing the source of the variation among the treatments. This table displays the sum of squares, degrees of freedom, and mean squares for the subjects, for each factor, for both factors together, and for the subject and the repeated factor. The corresponding F statistics and the corresponding P values are also displayed.

A Two Way Repeated Measures ANOVA of two repeated factors includes the sum of squares, degrees of freedom, and mean squares for the subjects with both factors, since both factors are repeated. Corresponding F statistics and the corresponding P values are also displayed.

Tables of least square means for each for the levels of factor and for the levels of both factors together are also generated for both one and two factor two way repeated measures ANOVA.

Additional results for both forms of Two Way Repeated Measure ANOVA can be disabled and enabled in the Options for Two Way RM ANOVA dialog box. Multiple comparisons are enabled in the Options for Two Way RM ANOVA dialog box.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the **Options** dialog box. You can also set the number of decimal places to display in the Options dialog box.

If There Were Missing Data or Empty Cells

If your data contained missing values but no empty cells, the report indicates the results were computed using a general linear model. The ANOVA table includes the approximate degrees of freedom used to compute F, the estimated mean square equations are listed, and the summary table displays the estimated least square means.

If your data contained empty cells, you either analyzed the problem assuming no interaction, or treated the problem as a One Way ANOVA.

- If you choose no interactions, no statistics for factor interaction are calculated.
- If you performed a One Way ANOVA, the results shown are identical to one way ANOVA results.

For more information, see [Interpreting Two Way Repeated Measures ANOVA Results](#) on page 175.

Dependent Variable

This is the column title of the indexed worksheet data you are analyzing with the Two Way Repeated Measures ANOVA. Determining if the values in this column are affected by the different factor levels is the objective of the Two Way Repeated Measures ANOVA.

Normality Test

Normality test results display whether the data passed or failed the test of the assumption that the differences of the changes originate from a normal distribution, and the P value calculated by the test. A normally distributed source is required for all parametric tests.

This result appears if you enabled normality testing in the Options for Two Way RM ANOVA dialog box.

Equal Variance Test

Equal Variance test results display whether or not the data passed or failed the test of the assumption that the differences of the changes originate from a population with the same variance, and the P value calculated by the test. Equal variance of the source is assumed for all parametric tests.

This result appears if you enabled equal variance testing in the Options for Two Way RM ANOVA dialog box.

ANOVA Table

The ANOVA table lists the results of the two way repeated measures ANOVA. The results are calculated for each factor, and then between the factors.

DF (Degrees of Freedom). The degrees of freedom are a measure of the numbers of subjects and treatments, which affects the sensitivity of the ANOVA.

- Factor degrees of freedom are measures of the number of treatments in each factor (columns in the table).
- The factor x factor interaction degrees of freedom is a measure of the total number of cells.
- The subjects degrees of freedom is a measure of the number of subjects (rows in the table).
- The subject x factor degrees of freedom is a measure of the number of subjects and treatments for the factor.

- The residual degrees of freedom is a measure of difference between the number of subjects and the number of treatments after accounting for factor and interaction.

SS (Sum of Squares). The sum of squares is a measure of variability associated with each element in the ANOVA table.

- Factor sum of squares measures variability of treatments in each factor (between the rows and columns of the table, considered separately).
- The factor x factor interaction sum of squares measures the variability of the treatments for both factors; this is the variability of the average differences between the cell in addition to the variation between the rows and columns, considered separately.
- The subjects sum of squares measures the variability of all subjects.
- The subject x factor sum of squares is a measure of the variability of the subjects within each factor.
- The residual sum of squares is a measure of the underlying variability of all observations.

MS (Mean Squares). The mean squares provide estimates of the population variances. Comparing these variance estimates is the basis of analysis of variance.

The mean square for each factor

$$\frac{\text{sum of squares for the factor}}{\text{degrees of freedom for the factor}} = \frac{SS_{\text{factor}}}{DF_{\text{factor}}} = MS_{\text{factor}}$$

is an estimate of the variance of the underlying population computed from the variability between levels of the factor.

The interaction mean square

$$\frac{\text{sum of squares for the interaction}}{\text{degrees of freedom for the interaction}} = \frac{SS_{\text{inter}}}{DF_{\text{inter}}} = MS_{\text{inter}}$$

is an estimate of the variance of the underlying population computed from the variability associated with the interactions of the factors.

The error mean square (residual, or within groups)

$$\frac{\text{error of sum of squares}}{\text{error degrees of freedom}} = \frac{SS_{\text{error}}}{DF_{\text{error}}} = MS_{\text{error}}$$

is an estimate of the variability in the underlying population, computed from the random component of the observations.

F Test Statistic. The F test statistic is provided for comparisons within each factor and between the factors

If there are no missing data, the F statistic within the factors is

$$\frac{\text{mean square for the factor}}{\text{error mean square for the factor}} = \frac{MS_{\text{factor}}}{MS_{\text{error}}} = F_{\text{factor}}$$

and the F ratio between the factors is

$$\frac{\text{mean square for the interaction}}{\text{error mean square for the interaction}} = \frac{MS_{\text{inter}}}{MS_{\text{error}}} = F_{\text{inter}}$$

 **Note:** If there are missing data or empty cells, SigmaPlot automatically adjusts the F computations to account for the offsets of the expected mean squares.

If the F ratio is around 1, the data is consistent with the null hypothesis that there is no effect (for example, no differences among treatments).

If F is a large number, the variability among the means is larger than expected from random variability in the population, and you can conclude that the samples were drawn from different populations (for example, the differences between the treatments are statistically significant).

P value. The P value is the probability of being wrong in concluding that there is a true difference between the treatments (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based

on F). The smaller the P value, the greater the probability that the samples are drawn from different populations. Traditionally, you can conclude there are significant differences if $P < 0.05$.

Approximate DF (Degrees of Freedom). If a general linear model was used, the ANOVA table also includes the approximate degrees of freedom that allow for the missing value(s). See DF (Degrees of Freedom) above for an explanation of the degrees of freedom for each variable.

Power

The power of the performed test is displayed unless you disable this option in the Options for Two Way RM ANOVA dialog box.

The power, or sensitivity, of a Two Way Repeated Measures ANOVA is the probability that the test will detect a difference among the treatments if there really is a difference. The closer the power is to 1, the more sensitive the test.

Repeated Measures ANOVA power is affected by the sample sizes, the number of treatments being compared, the chance of erroneously reporting a difference a (alpha), the observed differences of the group means, and the observed standard deviations of the samples.

Alpha (α). Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. An α error is also called a Type I error. A Type I error is when you reject the hypothesis of no effect when this hypothesis is true.

Set the value in the Options for Two Way RM ANOVA dialog box; the suggested value is $\alpha = 0.05$ which indicates that a one in twenty chance of error is acceptable. Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference but also increase the risk of seeing a false difference (a Type I error).

Expected Mean Squares

If there were missing data and a general linear model was used, the linear equations for the expected mean squares computed by the model are displayed. These equations are displayed only if a general linear model was used.

Summary Table

The least square means and standard error of the means are displayed for each factor separately (summary table row and column), and for each combination of factors (summary table cells). If there are missing values, the least square means are estimated using a general linear model.

Mean. The average value for the condition or group.

Standard Error of the Mean. A measure of uncertainty in the mean.

The Least Squares Mean and associated Standard Error are computed based on all the data. These values can differ from the values computed from the data in the individual cells. In particular, if the design is balanced, all the least square errors will be equal for all cells. (If the sample sizes in different cells are different, the least squares standard errors will be different, depending on the sample sizes, with larger standard errors associated with smaller sample sizes.) These standard errors will be different than the standard errors computed from each cell separately.

This table is generated if you select to display summary table in the Options for Two Way RM ANOVA dialog box.

Multiple Comparisons

If SigmaPlot finds a difference among the treatments, then you can compute a multiple comparison table. Multiple comparisons are enabled in the Options for Two Way Repeated Measures ANOVA dialog box.

Use the multiple comparison results to determine exactly which treatments are different, since the ANOVA results only inform you that two or more of the treatments are different. Two factor multiple comparison for a full Two Way ANOVA also compares:

- Treatments within each factor without regard to the other factor (this is a marginal comparison, for example, only the columns or rows in the table are compared).
- All combinations of factors (all cells in the table are compared).

The specific type of multiple comparison results depends on the comparison test used and whether the comparison was made pairwise or versus a control.

- All pairwise comparison results list comparisons of all possible combinations of group pairs; the all pairwise tests are the Tukey, Student-Newman-Keuls, Fisher LSD, Duncan's, and Dunnett's, and Bonferroni t-test.
- Comparisons versus a single control group list only comparisons with the selected control group. The control group is selected during the actual multiple comparison procedure. The comparison versus a control tests are a Bonferroni t-test and Dunnett's test.

Bonferroni t-test Results. The Bonferroni t-test lists the differences of the means for each pair of treatments, computes the t values for each pair, and displays whether or not $P < 0.05$ for that comparison. The Bonferroni t-test can be used to compare all treatments or to compare versus a control.

You can conclude from "large" values of t that the difference of the two treatments being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of erroneously concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The Difference of Means is a gauge of the size of the difference between the treatments or cells being compared.

The degrees of freedom DF for the marginal comparisons are a measure of the number of treatments (levels) within the factor being compared. The degrees of freedom when comparing all cells is a measure of the sample size after accounting for the factors and interaction (this is the same as the error or residual degrees of freedom).

Tukey, Student-Newman-Keuls, Fisher LSD, Duncan's, and Dunnett's Test Results. The Tukey, Student-Newman-Keuls (SNK), Fisher LSD, and Duncan's tests are all pairwise comparisons of every combination of group pairs. While the Tukey, Fisher LSD, and Duncan's can be used to compare a control group to other groups, they are not recommended for this type of comparison.

Dunnett's test only compares a control group to all other groups. All tests compute the θ test statistic, the number of means spanned in the comparison ϕ , and display whether or not $P < 0.05$ for that pair comparison.

You can conclude from "large" values of θ that the difference of the two treatments being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of being incorrect in concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

α is parameter used when computing θ . The larger the α , the larger θ needs to be to indicate a significant difference. α is an indication of the differences in the ranks of the group means being compared. Groups means are ranked in order from largest to smallest in an SNK test, so α is the number of means spanned in the comparison. For example, when comparing four means, comparing the largest to the smallest $\alpha = 4$, and when comparing the second smallest to the smallest $\alpha = 2$.

If a treatment is found to be not significantly different than another treatment, all treatments with α ranks in between the α ranks of the two treatments that are not different are also assumed not to be significantly different, and a result of DNT (Do Not Test) appears for those comparisons.

 **Note:** SigmaPlot does not apply the DNT logic to all pairwise comparisons because of differences in the degrees of freedom between different cell pairs.

The Difference of Means is a gauge of the size of the difference between the treatments or cells being compared.

The degrees of freedom DF for the marginal comparisons are a measure of the number of treatments (levels) within the factor being compared. The degrees of freedom when comparing all cells is a measure of the sample size after accounting for the factors and interaction (this is the same as the error or residual degrees of freedom).

Two Way Repeated Measures ANOVA Report Graphs

You can generate up to five graphs using the results from a Two Way Repeated Measures ANOVA. They include a:

- [Histogram of the residuals](#).

- Normal probability plot of the residuals.
- 3D scatter plot of the residuals.
- 3D category scatter plot.
- Multiple comparison graphs.

Creating a Two Way Repeated Measures ANOVA Report Graph

1. Select the Two Way Repeated Measures ANOVA test report.
2. On the **Report** tab in the **Result Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the Two Way Repeated Measure ANOVA results.

3. Select the type of graph you want to create from the **Graph Type** list, then click **OK**, or double-click the desired graph in the list.

The selected graph appears in a graph window.

Friedman Repeated Measures Analysis of Variance on Ranks

Use a Repeated Measures ANOVA (analysis of variance) on Ranks when:

- You want to see if a single group of individuals was affected by a series of three or more different experimental treatments, where each individual received treatment.
- The treatment effects are not normally distributed.

If you know the treatment effects are normally distributed, use One Way Repeated Measures ANOVA. If there are only two treatments to compare, do a Wilcoxon Signed Rank Test. There is no two factor test for non-normally distributed treatment effects.

of one subject. Indexed data is placed in three worksheet columns: a factor column, a subject index column, and a data column.

The columns for raw data must be the same length. If a missing value is encountered, that individual is ignored.

Setting the Repeated Measures ANOVA on Ranks Options

Use the Repeated Measures ANOVA on Ranks options to:

- Adjust the parameters of the test to relax or restrict the testing of your data for normality and equal variance.
- Display the summary table.
- Enable and disable multiple comparison testing.

To change the Repeated Measures ANOVA on Ranks options:

1. On the **Analysis** tab, in the **SigmaStat** group, select **RM ANOVA on Ranks**.
2. Click **Options**. The Options for RM ANOVA on Ranks dialog box appears with three tabs:
 - **Assumption Checking**. Select the **Assumption Checking** tab to view the Normality and Equal Variance options.
 - **Results**. Select the **Results** tab to view the Summary Table option.
 - **Post Hoc Tests**. Select the **Post Hoc Test** tab to view the multiple comparisons options.
3. To continue the test, click **Run Test**.
4. To accept the current settings and close the options dialog box, click **OK**.

Options for Repeated Measures ANOVA on Ranks: Assumption Checking

The normality assumption test checks for a normally distributed population. The equal variance assumption test checks the variability about the group means.

- **Normality Testing**. SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.
- **Equal Variance Testing**. SigmaPlot Tests for equal variance by checking the variability about the group means.
- **P Values for Normality and Equal Variance**. Enter the corresponding P value in the P Value to Reject box. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (the P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P value computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and/or equal variance, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality and/or equal variance, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.01 for the normality test requires greater deviations from normality to flag the data as non-normal than a value of 0.05.

- 🚫 **Restriction:** Although the assumption tests are robust in detecting data from populations that are non-normal or with unequal variances, there are extreme conditions of data distribution that these tests cannot take into account. For example, the Levene Median test fails to detect differences in variance of several orders of magnitude; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption tests.

Options for Repeated Measures ANOVA on Ranks: Results

The Summary Table for a ANOVA on Ranks lists the medians, percentiles, and sample sizes N in the ANOVA on Ranks report. If desired, change the percentile values by editing the boxes. The 25th and the 75th percentiles are the suggested percentiles.

Options for Repeated Measures ANOVA on Ranks: Post Hoc Tests

Select the Post Hoc Test tab in the Options dialog box to view the multiple comparisons options. Repeated Measures ANOVA on Ranks test the hypothesis of no differences between the several treatment groups, but do not determine which groups are different, or the sizes of these differences. Multiple comparison procedures isolate these differences.

The P value used to determine if the ANOVA detects a difference is set in the Report Options dialog box. If the P value produced by the One Way ANOVA is less than the P value specified in the box, a difference in the groups detected and the multiple comparisons are performed.

Performing Multiple Comparisons. You can choose to always perform multiple comparisons or to only perform multiple comparisons if the Two Way ANOVA detects a difference.

Select the Always Perform option to perform multiple comparisons whether or not the ANOVA detects a difference.

Select the Only When ANOVA P Value is Significant option to perform multiple comparisons only if the ANOVA detects a difference.

Significant Multiple Comparison Value. Select either .05 or .10 from the Significance Value for Multiple Comparisons drop-down list. This value determines the that the likelihood of the multiple comparison being incorrect in concluding that there is a significant difference in the treatments.

A value of .05 indicates that the multiple comparisons will detect a difference if there is less than 5% chance that the multiple comparison is incorrect in detecting a difference. A value of .10 indicates that the multiple comparisons will detect a difference if there is less than 10% chance that the multiple comparison is incorrect in detecting a difference.

 **Note:** If multiple comparisons are triggered, the [Multiple Comparison Options](#) dialog box appears after you pick your data from the worksheet and run the test, prompting you to choose a multiple comparison method.

Running a Repeated Measures ANOVA on Ranks

To run an Repeated Measures ANOVA on Ranks, you need to select the data to test. If you want to select your data before you run the test, drag the pointer over your data.

1. On the **Analysis** tab, in the **SigmaStat** group, from the **Tests** drop-down list select:

Repeated Measures > Repeated Measures ANOVA on Ranks

The **Data Format** panel of the Test Wizard appears prompting you to specify a data format.

2. Select the appropriate data format from the **Data Format** drop-down list. For more information, see [Data Format for Repeated Measures Tests](#) on page 140.
3. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.
4. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the Selected Columns list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw and indexed data, you are prompted to select two worksheet columns.

5. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
6. Click **Finish** to run the RM ANOVA on Ranks test on the selected columns.

If you elected to test for normality and equal variance, SigmaPlot performs the test for normality (Shapiro-Wilk or Kolmogorov-Smirnov) and the test for equal variance (Levene Median). If your data passes both tests, SigmaPlot informs you and suggests continuing your analysis using One Way Repeated Measures ANOVA.

If you did not enable multiple comparison testing in the Options for RM ANOVA on Ranks dialog box, the Repeated Measures ANOVA on Ranks report appears after the test is complete.

If you did enable the [Multiple Comparisons](#) option in the options dialog box, the Multiple Comparison Options dialog box appears prompting you to select a multiple comparison method.

Multiple Comparison Options (Repeated Measures ANOVA on Ranks)

If you selected to run multiple comparisons only when the P value is significant, and the ANOVA produces a P value, for either of the two factors or the interaction between the two factors, equal to or less than the trigger P value, or you selected to always run multiple comparisons in the Options for RM ANOVA on Ranks dialog box, the Multiple Comparison Options dialog box appears prompting you to specify a multiple comparison test.

This dialog box displays the P values for each of the two experimental factors and of the interaction between the two factors. Only the options with P values less than or equal to the value set in the Options dialog box are selected. You can disable multiple comparison testing for a factor by clicking the selected option. If no factor is selected, multiple comparison results are not reported.

There are four multiple comparison tests to choose from for the ANOVA on Ranks. You can choose to perform the:

- Dunn's Test.
- Dunnett's Test.
- Tukey Test.
- Student-Newman-Keuls Test.

There are two kinds of multiple comparison procedures available for the Repeated Measures ANOVA on Ranks.

- All pairwise comparisons test the difference between each treatment or level within the two factors separately (for example, among the different rows and columns of the data table)
- Multiple comparisons versus a control test the difference between all the different combinations of each factors (for example, all the cells in the data table)

Interpreting Repeated Measures ANOVA on Ranks Results

The Friedman Repeated Measures ANOVA on Ranks report displays the results for χ^2 , degrees of freedom, and P. The other results displayed are selected in the Options for RM ANOVA on Ranks dialog box. Multiple comparisons are enabled in the Options for RM ANOVA on Ranks dialog box. The test used to perform the multiple comparison is selected in the Multiple Comparisons Options dialog box.

Result Explanations. In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

Normality Test

Normality test results display whether the data passed or failed the test of the assumption that the differences of the treatments originate from a normal distribution, and the P value calculated by the test. For nonparametric procedures this test can fail, as nonparametric tests do not require normally distributed source populations. This result appears unless you disabled normality testing in the Options for RM ANOVA on Ranks dialog box.

Equal Variance Test

Equal Variance test results display whether or not the data passed or failed the test of the assumption that the differences of the treatments originate from a population with the same variance, and the P value calculated by the test. Nonparametric tests do not assume equal variance of the source. This result appears unless you disabled equal variance testing in the Options for RM ANOVA on Ranks dialog box.

Summary Table

SigmaPlot can generate a summary table listing the sample sizes N, number of missing values, medians, and percentiles defined in the Options for RM ANOVA on Ranks dialog box.

N (Size). The number of non-missing observations for that column or group.

Missing. The number of missing values for that column or group.

Medians. The "middle" observation as computed by listing all the observations from smallest to largest and selecting the largest value of the smallest half of the observations. The median observation has an equal number of observations greater than and less than that observation.

Percentiles. The two percentile points that define the upper and lower tails of the observed values.

These results appear in the report unless you disable them in the Options for RM ANOVA on Ranks dialog box.

Chi-Square Statistic

The Friedman test statistic x_r^2 is used to evaluate the null hypothesis that all the rank sums are equal. If the value of x_r^2 is large, you can conclude that the treatment effects are different (for example, that the differences in the rank sums are greater than would be expected by chance).

Values of x_r^2 near zero indicate that there is no significant difference in treatments; the ranks within each subject are random.

x_r^2 is computed by ranking all observations for each subject from smallest to largest without regard for other subjects. The ranks are summed for each treatment and x_r^2 is computed from the sum of squares.

Degrees of Freedom. The degrees of freedom is an indication of the sensitivity of x_r^2 . It is a measure of the number of treatments.

P value. The P value is the probability of being wrong in concluding that there is a true difference in the treatments (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on x_r^2). The smaller the P value, the greater the probability that the samples are significantly different.

Traditionally, you can conclude there are significant differences when $P < 0.05$.

Multiple Comparisons

If a difference is found among the groups, and you requested and elected to perform multiple comparisons, a table of the comparisons between group pairs is displayed. The multiple comparison procedure is activated in the Options for ANOVA on Ranks dialog box. The test used in the multiple comparison procedure is selected in the Multiple Comparison Options dialog box.

Multiple comparison results are used to determine exactly which groups are different, since the ANOVA results only inform you that two or more of the groups are different. The specific type of multiple comparison results depends on the comparison test used and whether the comparison was made pairwise or versus a control.

- All pairwise comparison results list comparisons of all possible combinations of group pairs: the all pairwise tests are the Tukey, Student-Newman-Keuls test and Dunn's test.
- Comparisons versus a single control list only comparisons with the selected control group. The control group is selected during the actual multiple comparison procedure. The comparison versus a control tests are Dunnett's test and Dunn's test.

Tukey, Student-Newman-Keuls, and Dunnett's Test Results. The Tukey and Student-Newman-Keuls (SNK) tests are all pairwise comparisons of every combination of group pairs. Dunnett's test only compares a control group to all other groups. All tests compute the θ test statistic, the number of rank sums spanned in the comparison ϕ , and display whether or not $P < 0.05$ for that pair comparison.

You can conclude from "large" values of θ that the difference of the two treatments being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of being incorrect in concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The rank sums is a gauge of the size of the difference between the two treatments.

ϕ is parameter used when computing θ . The larger the ϕ , the larger θ needs to be to indicate a significant difference. ϕ is an indication of the differences in the ranks of the rank sums being compared. Group rank sums are ranked in order from largest to smallest in an SNK test, so ϕ is the number of ranks spanned in the comparison. For example, when comparing four rank means, comparing the largest to the smallest $\phi = 4$, and when comparing the second smallest to the smallest $\phi = 2$.

If a treatment is found to be not significantly different than another treatment, all treatments with ϕ ranks in between the ϕ ranks of the two treatments that are not different are also assumed not to be significantly different, and a result of Do Not Test appears for those comparisons.

 **Note:** SigmaPlot does not apply the DNT logic to all pairwise comparisons because of differences in the degrees of freedom between different cell pairs.

Dunn's Test Results. Dunn's test is used to compare all treatments or to compare versus a control when the group sizes are unequal. Dunn's test lists the difference of ranks, computes the Q test statistic, and displays whether or not $P < 0.05$, for each treatment pair.

You can conclude from "large" values of Q that the difference of the two treatments being compared is statistically significant.

If the P value for the comparison is less than 0.05, the likelihood of being incorrect in concluding that there is a significant difference is less than 5%. If it is greater than 0.05, you cannot confidently conclude that there is a difference.

The rank sums is a gauge of the size of the difference between the two treatments.

A result of DNT (do not test) appears for those comparison pairs whose difference of rank means is less than the differences of the first comparison pair which is found to be not significantly different. For more information, see [Repeated Measures ANOVA on Ranks Report Graphs](#) on page 185.

Repeated Measures ANOVA on Ranks Report Graphs

You can generate up to three graphs using the results from a Repeated Measures ANOVA on Ranks. They include a:

- [Box plot of the column means](#).
- [Line graph of the changes after treatment](#).
- [Multiple comparison graphs](#).

Creating a Repeated Measures ANOVA on Ranks Report Graph

1. Select the Repeated Measures ANOVA on Ranks test report.
2. On the **Report** tab, in the **Result Graphs** group, select **Create Report Graph**.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the **One Way Repeated Measure ANOVA** results.

3. Select the type of graph you want to create from the **Graph Type** list, then click **OK**, or double-click the desired graph in the list.

The selected graph appears in a graph window.

Chapter

7

Comparing Frequencies, Rates, and Proportions

Topics:

- [About Rate and Proportion Tests](#)
- [Comparing Proportions Using the z-Test](#)
- [Chi-square Analysis of Contingency Tables](#)
- [The Fisher Exact Test](#)
- [McNemar's Test](#)
- [Relative Risk Test](#)
- [Odds Ratio Test](#)

Use rate and proportion tests to compare two or more sets of data for differences in the number of individuals that fall into different classes or categories. You can find all of these tests by going to the menus and selecting: If you are comparing groups where the data is measured on a numeric scale, use the appropriate group comparison or repeated measures tests. For more information, see [Choosing the Procedure to Use](#) on page 18.

About Rate and Proportion Tests

Rate and proportion tests are used when the data is measured on a *nominal scale*. Rate and proportion comparisons test for significant differences in the categorical distribution of the data beyond what can be attributed to random variation. For more information, see [Choosing the Rate and Proportion Comparison to Use](#) on page 33.

Contingency Tables

Many rate and proportion tests utilize a *contingency* table which lists the groups and/or categories to be compared as the table column and row titles, and the number of observations for each combination of category or group as the table *cells*. A contingency table is used to determine whether or not the distribution of a group is contingent on the categories it falls in.

A 2 x 2 contingency table has two groups and two categories (for example, two rows and two columns). A 2 x 3 table has two groups and three categories or three groups and two categories, and so on.

Comparing the Proportions of Two Groups in One Category

Use a z-test to compare the proportions of two groups found within a single category for a significant difference. To perform a z-Test:

1. Click the **Analysis** tab.
2. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Rates and Proportions > z-Test

Comparing Proportions of Multiple Groups in Multiple Categories

You can use analysis of contingency tables to test if the distributions of two or more groups within two or more categories are significantly different.

- Use Chi-Square χ^2 analysis of contingency if there are more than two groups or categories, or if the expected number of observations per cell in a 2 x 2 contingency table are greater than five.
- Use the Fisher Exact Test when the expected number of observations is less than five in any cell of a 2 x 2 contingency table.

SigmaPlot automatically checks your data during a Chi-Square analysis and suggests the Fisher Exact Test when applicable. Note than you can perform the Fisher Exact Test on any 2 x 2 contingency table.

 **Note:** SigmaPlot computes a two-tailed Fisher Exact Test.

Comparing Proportions of the Same Group to Two Treatments

You can test for differences in the proportions of the responses in the same individuals to a series of two different treatments using McNemar's Test for changes.

Yates Correction

The Yates Correction for continuity can be automatically applied to the z-test and for all tests using 2 x 2 tables or comparisons with the χ^2 distribution with one degree of freedom. It is generally accepted that the Yates Correction yields a more accurately computed P value in these cases.

Application of the Yates Correction Factor is selected in the **Options** dialog box for each test.

Comparing Proportions Using the z-Test

Compare proportions with a z-test when:

- You have two groups to compare.

- You know the total sample size (number of observations) for each group.
- You have the proportions p for each group that falls within a single category.

If you have data for the numbers of observations for each group that fall in two categories perform LogRank Survival Analysis of contingency tables instead. This will produce the same P value as the z-test. You can also run the LogRank Survival Analysis of contingency tables if you have more than two groups or categories.

About the z-Test

The z-test comparison of proportions is used to determine if the proportions of two groups within one category or class are significantly different. The z-test assumes that:

- Each observation falls into one of two mutually exclusive categories.
- All observations are independent.

Performing a z-Test

To perform a z-test:

1. [Enter or arrange your data in the data worksheet.](#)
2. If desired, set the z-test options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Rates and Proportions > z-test
5. Run the test.
6. [View and interpret the z-test report.](#)

Arranging z-Test Data

To compare two proportions, enter the two sample sizes in one column and the corresponding observed proportions p in a second column. There must be exactly two rows and two columns. The sample sizes must be whole numbers and the observed proportions must be between 0 and 1.

Setting z-Test Options

Use the Compare Proportion options to:

- Display the confidence interval for the data in Compare Proportion test reports.
- Display the power of a performed test for Compare Proportion tests in the reports.
- Enable the Yates Correction Factor.

To change z-test options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. Select **z-test** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis**.

3. Click Current Test Options.

The **Options for z-test** dialog box appears. For more information, see [Options for z-Test](#) on page 190.

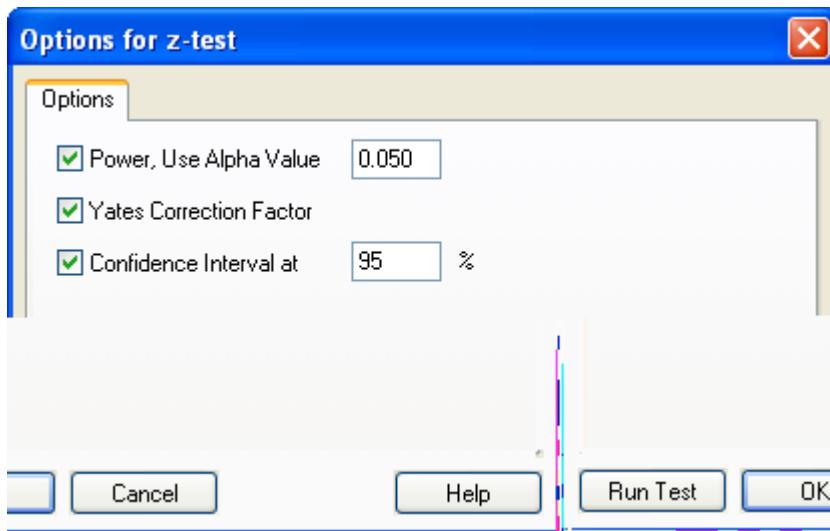


Figure 94: The Options for z-test Dialog Box

4. Click a check box to select a test option. All options are saved between SigmaPlot sessions.
5. **To continue the test, click Run Test.**
6. **To accept the current settings and close the options dialog box, click OK.**

Options for z-Test

Power, Use Alpha Value. Select to detect the sensitivity of the test. The power or sensitivity of a test is the probability that the test will detect a difference between the proportions of two groups if there is really a difference.

Change the alpha value by editing the number in the Alpha Value box.

Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

The Yates Correction Factor. When a statistical test uses a χ^2 distribution with one degree of freedom, such as analysis of a 2×2 contingency table or McNemar's test, the χ^2 calculated tends to produce P values which are too small, when compared with the actual distribution of the χ^2 test statistic. The theoretical χ^2 distribution is continuous, whereas the distribution of the χ^2 test statistic is discrete.

Use the Yates Correction Factor to adjust the computed LogRank Survival Analysis value down to compensate for this discrepancy. Using the Yates correction makes a test more conservative, for example, it increases the P value and reduces the chance of a false positive conclusion. The Yates correction is applied to 2×2 tables and other statistics where the P value is computed from a χ^2 distribution with one degree of freedom.

Click the selected check box to turn the Yates Correction Factor on or off.

Confidence Interval. This is the confidence interval for the difference of proportions. To change the specified interval, select the box and type any number from 1 to 99 (95 and 99 are the most commonly used intervals).

Running a z-Test

To run a test, you need to select the data to test. The **Select Data** panel of the Test Wizard is used to select the worksheet columns with the data you want to test and to specify how your data is arranged in the worksheet.

To run a z-test:

1. If you want to select your data before you run the test, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Rates and Proportions > z-test

The **Select Data** panel of the Test Wizard appears. If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog box prompts you to pick your data.

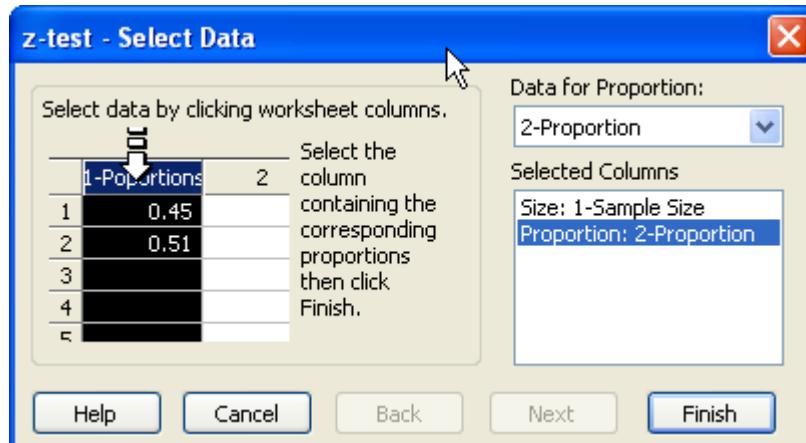
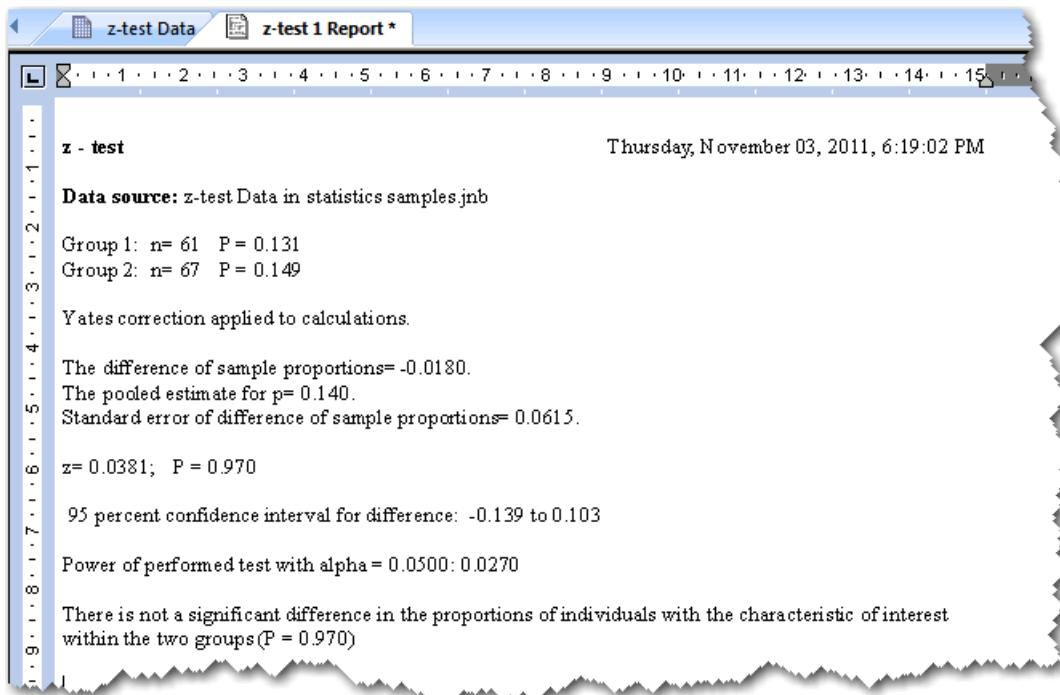


Figure 95: The z-test - Select Data Dialog Box Prompting You to Select Data Columns

4. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Size or Proportion** drop-down list. The first selected column is assigned to the **Size** row in the **Selected Columns** list, and the second column is assigned to **Proportion** row in the list. The title of selected columns appear in each row. You can only select one **Size** and one **Proportion** data column.
5. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
6. Click **Finish** to perform the test. The report appears displaying the results of the z-test. For more information, see [Interpreting Proportion Comparison Results](#) on page 191.

Interpreting Proportion Comparison Results

The z-test report displays a table of the statistical values used, the z statistic, and the P for the test. You can also display a confidence interval for the difference of the proportions using the Options for z-test dialog box.



```

z - test                                         Thursday, November 03, 2011, 6:19:02 PM

Data source: z-test Data in statistics samples.jnb

Group 1: n= 61  P = 0.131
Group 2: n= 67  P = 0.149

Yates correction applied to calculations.

The difference of sample proportions= -0.0180.
The pooled estimate for p= 0.140.
Standard error of difference of sample proportions= 0.0615.

z= 0.0381;  P = 0.970

95 percent confidence interval for difference: -0.139 to 0.103

Power of performed test with alpha = 0.0500: 0.0270

There is not a significant difference in the proportions of individuals with the characteristic of interest
within the two groups(P = 0.970)

```

Figure 96: The z-test Comparison of Proportions Results Report

Results Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box. For more information, see [Report Graphs](#) on page 373.

Statistical Summary

The summary table for α z-test lists the sizes of the groups n and the proportion of each group in the category p . These values are taken directly from the data.

Difference of Proportions. This is the difference between the p proportions for the two groups.

Pooled Estimate for P. This is the estimate of the population proportion p based on pooling the two samples to test the hypothesis that they were drawn from the same population. It depends on both the nature of the underlying population and the specific samples drawn.

Standard Error of the Difference. The standard error of the difference is a measure of the precision with which this difference can be estimated.

z Statistic

The z statistic is

$$\frac{\text{difference of the sample proportions}}{\text{standard error of the sample proportions}} = z$$

You can conclude from "large" absolute values of z that the proportions of the populations are different. A large z indicates that the difference between the proportions is larger than what would be expected from sampling variability alone (for example, that the difference between the proportions of the two groups is statistically significant). A small z (near 0) indicates that there is no significant difference between the proportions of the two groups.

If you enabled the Yates correction in the **Options for z-test** dialog box, the calculation of z is slightly smaller to account for the difference between the theoretical and calculated values of z .

P Value. The P value is the probability of being wrong in concluding that>>

	Tundra	Foothills	Treeline
Lowland	125	16	6
Alpine	7	19	117

The χ^2 test uses the percentages of the row and column totals for each cell to compute the expected number of observations per cell if the treatment had no effect. The χ^2 statistic summarizes the difference between the expected and the observed frequencies.

Performing a Chi-Square Test

To perform a Chi-Square Test:

1. [Enter or arrange your data appropriately in the data worksheet.](#)
2. If desired, set the Chi-Square options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Rates and Proportions > Chi-Square
5. Run the test.
6. [View and interpret the Chi-Square report.](#)

Arranging Chi-Square Data

The data can be arranged in the worksheet as either contingency table data or as raw data.

Tabulated Data Tabulated data is arranged in a contingency table showing the number of observations for each cell. The worksheet rows and columns correspond to the groups and categories. The number of observations must always be an integer.

Note that the order and location of the rows or columns corresponding to the groups and categories is unimportant. You can use the rows for category and the columns for group, or vice versa.

Table 19: A Contingency Table describing the number of Lowland and Alpine species found at different locations.

Species	Location		
	Tundra	Foothills	Treeline
Lowland	125	16	6
Alpine	7	19	117

Raw Data You can report the group and category of each individual observation by placing the group in one worksheet column and the corresponding category in another column. Each row corresponds to a single observation, so there should be as many rows of data as there are total numbers of observations.

SigmaPlot automatically cross tabulates these data and performs the χ^2 analysis on the resulting contingency table. For more information, see [Arranging Chi-Square Data](#) on page 194.

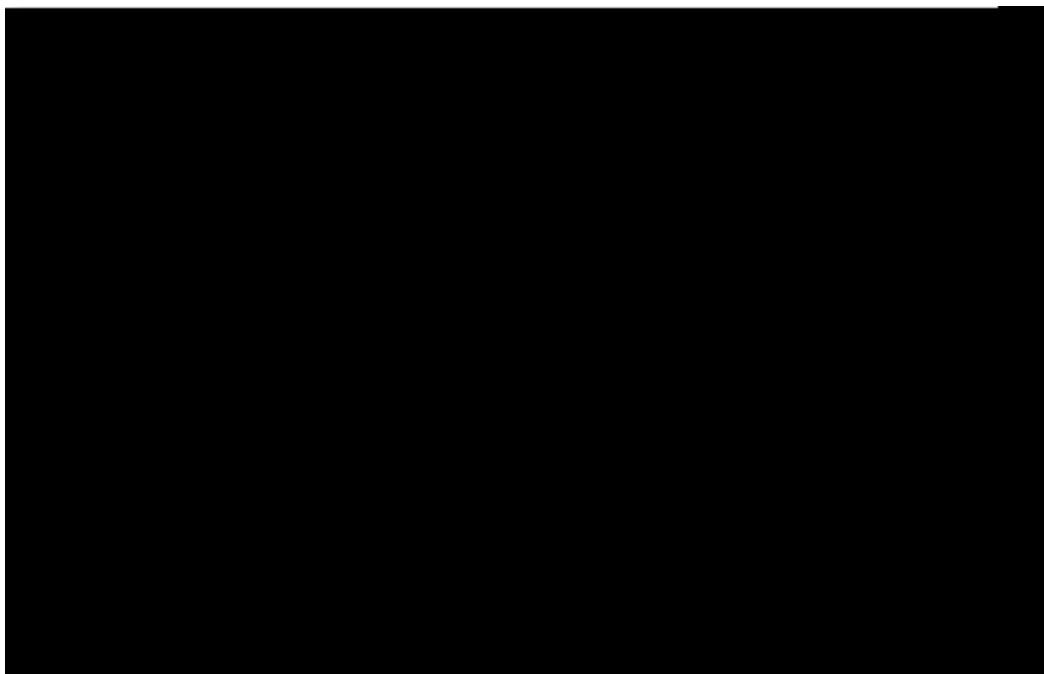


Figure 97: Worksheet Data Arrangement for Contingency Table Data from the Table above

Columns 1 through 3 in the worksheet above are in tabular format, and columns 4 and 5 are raw data.

Setting Chi-Square Options

Use the Chi-Square options to:

- Display the power of a performed test for Compare Proportion tests in the reports.
- Enable the Yates Correction Factor.

To change Chi-Square options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. Select **Chi-square** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis**.

The **Options for Chi-Square** dialog box appears.

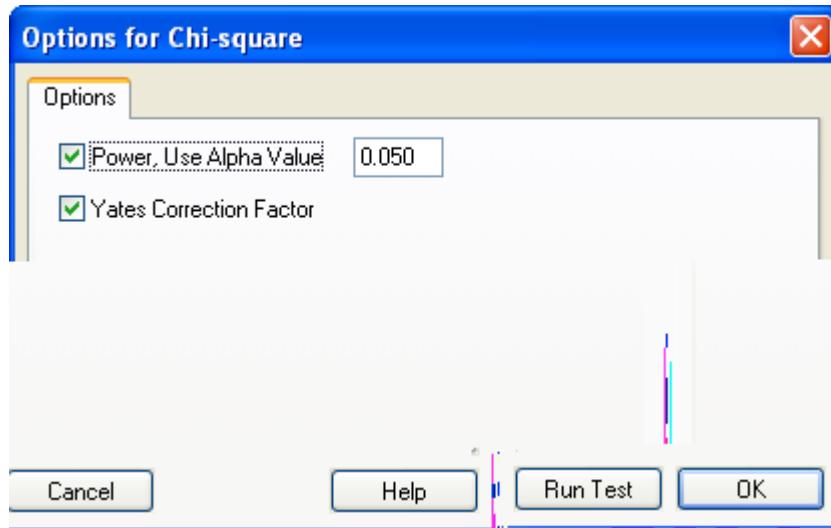


Figure 98: The Options for Chi-Square Dialog Box

3. Click a check box to enable or disable a test option. All options are saved between SigmaPlot sessions.
4. To continue the test, click **Run Test**.
5. To accept the current settings and close the options dialog box, click **OK**.

Options for Chi Square

Power, Use Alpha Value. Select to detect the sensitivity of the test. The power or sensitivity of a test is the probability that the test will detect a difference between the proportions of two groups if there is really a difference.

Change the alpha value by editing the value in the Alpha Value box.

Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

The Yates Correction Factor. When a statistical test uses a χ^2 distribution with one degree of freedom, such as analysis of a 2 x 2 contingency table or McNemar's test, the χ^2 calculated tends to produce P values which are too small, when compared with the actual distribution of the χ^2 test statistic. The theoretical χ^2 distribution is continuous, whereas the χ^2 produced with real data is discrete.

You can use the *Yates Continuity Correction* to adjust the computed χ^2 value down to compensate for this discrepancy. Using the Yates correction makes a test more conservative, for example, it increases the P value and reduces the chance of a false positive conclusion. The Yates correction is applied to 2 x 2 tables and other statistics where the P value is computed from a χ^2 distribution with one degree of freedom.

Click the check box to turn the Yates Correction Factor on or off.

Running a Chi-Square Test

To run a test, you need to select the data to test. Use the **Select Data** panel of the Test Wizard to select the worksheet columns with the data you want to test and to specify how your data is arranged in the worksheet.

To run a Chi-Square Test:

1. If you want to select your data before you run the test, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Rates and Proportions > Chi-Square

The **Chi-Square - Data Format** dialog box appears prompting you to specify a data format.

4. Select the appropriate data format from the **Data Format** drop-down list. If you are testing contingency table data, select **Tabulated**. If your data is arranged in raw format, select **Raw**. For more information, see [Arranging Chi-Square Data](#) on page 194.

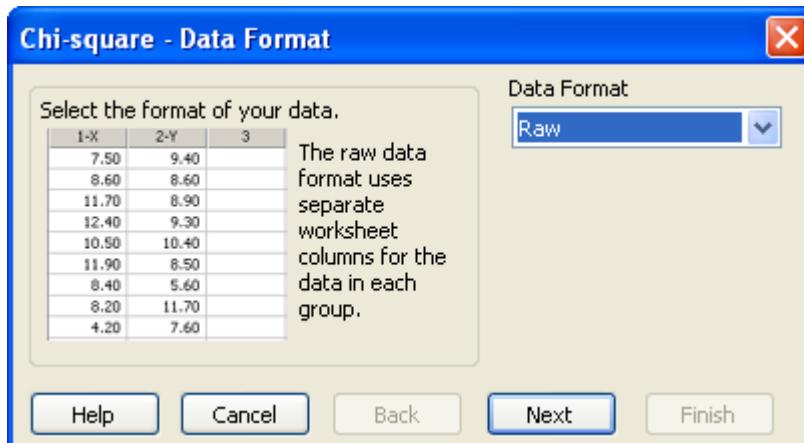


Figure 99: The Chi-Square - Data Format Dialog Box Prompting You to Select a Data Format

5. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.
If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog box prompts you to pick your data.
6. To assign the desired worksheet columns to the Selected Columns list, select the columns in the worksheet, or select the columns from the **Data for Observations** or **Category** drop-down list.
The first selected column is assigned to the first **Observation** or **Category** row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw data, you are prompted to select two worksheet columns. For tabulated data you are prompted to select up to 640 columns.
7. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the Selected Columns list.

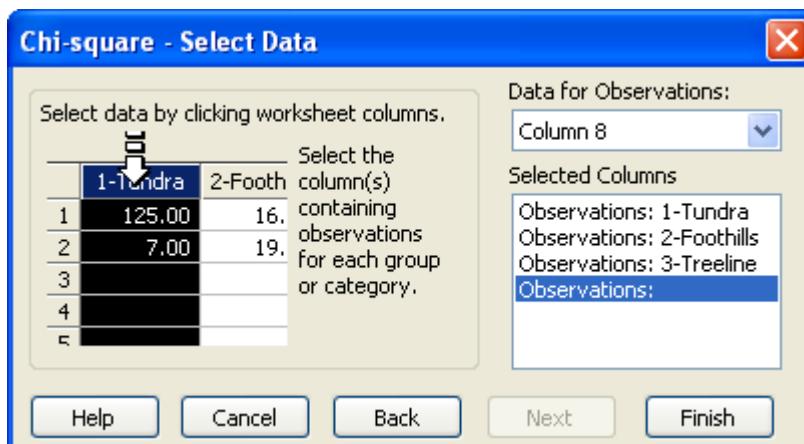


Figure 100: The Chi-Square - Select Data Dialog Box Prompting You to Select Data Columns

8. Click **Finish** to run the test. If there are too many cells in a contingency table with expected values below 5, SigmaPlot either:

- Suggests that you redefine the groups or categories in the contingency table to reduce the number of cells and increase the number of observations per cell.
- Suggests the Fisher Exact Test if the table is a 2 x 2 contingency table.

When there are many cells with expected observations of 5 or less, the theoretical χ^2 distribution does not accurately describe the actual distribution of the χ^2 test statistic, and the resulting P values may not be accurate.

Fisher Exact Test computes the exact two-tailed probability of observing a specific 2 x 2 contingency table, and does not require that the expected frequencies in all cells exceed 5. When the test is complete, the χ^2 test report appears. For more information, see [Interpreting Results of a Chi-Squared Analysis of Contingency tables](#) on page 198.

Interpreting Results of a Chi-Squared Analysis of Contingency tables

The report for a χ^2 test lists a summary of the contingency table data, the χ^2 statistic calculated from the distributions, and the P value for χ^2 .

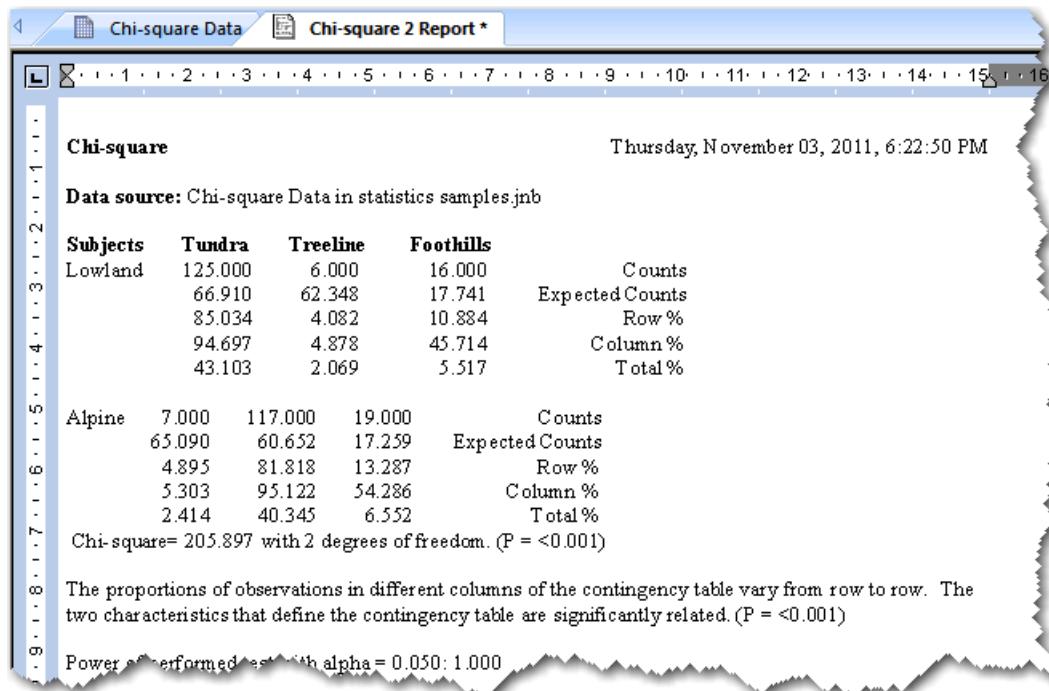


Figure 101: A Chi-Square Test Results Report

Results Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

Contingency Table Summary

Each cell in the table is described with a set of statistics.

Observed Counts. These are the number of observations per cell, obtained from the contingency table data.

Expected Frequencies. The expected frequencies for each cell in the contingency table, as predicted using the row and columns percentages.

Row Percentage. The percentage of observations in each row of the contingency table, obtained by dividing the observed frequency counts in the cells by the total number of observations in that row.

Column Percentage. The percentage of observations in each column of the contingency table, obtained by dividing the observed frequency counts in the cells by the total number of observations in that column.

Total Cell Percentage. The percentage of total number observations in the contingency table, obtained by dividing the observed frequency in the cells by the total number of observations in the table.

Chi-Square

χ^2 is the summed squared differences between the observed frequencies in each cell of the table and the expected

$$\chi^2 = \sum \frac{(observed - expected\ numbers\ per\ cell)^2}{expected\ numbers\ per\ cell}$$

frequencies, or

This computation assumes that the rows and columns are independent.

If the value of χ^2 is large, you can conclude that the distributions are different (for example, that there is a large differences between the expected and observed frequencies, indicating that the rows and columns are independent).

Values of χ^2 near zero indicate that the pattern in the contingency table is no different from what one would expect if the counts were distributed at random.

Yates Correction. The Yates correction is used to adjust the χ^2 and therefore the P value for 2 x 2 tables to more accurately reflect the true distribution of χ^2 . The Yates correction is enabled in the Options for Chi-Square dialog box, and is only applied to 2 x 2 tables.

P Value. The P value is the probability of being wrong in concluding that there is a true difference in the distribution of the numbers of observations (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on χ^2). The smaller the P value, the greater the probability that the samples are drawn from populations with different distributions among the categories. Traditionally, you conclude that there are significant differences when P < 0.05.

Power

The power, or sensitivity, of a Chi-Square test is the probability that the test will detect a difference among the groups if there really is a difference. The closer the power is to 1, the more sensitive the test. Chi-Square power is affected by the sample size and the observed proportions of the samples. This result is displayed if you selected this option in the Options for Chi-Square dialog box.

Alpha. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. An α error is also called a Type I error (a Type I error is when you reject the hypothesis of no effect when this hypothesis is true).

Set the α value is set in the **Power Option** dialog box. The suggested value is $\alpha = 0.05$, which indicates that a one in twenty chance of error is acceptable. Smaller values of α result in stricter requirements before concluding there is a difference in distribution, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference, but also increase the risk of seeing a false difference (a Type I error).

The Fisher Exact Test

Use the Fisher Exact Test to compare the distributions in a 2 x 2 contingency table that has 5 or less expected observations in one or more cells.

If no cells have less than five expected observations, you can use a χ^2 test.

About the Fisher Exact Test

The Fisher Exact Test determines the exact probability of observing a specific 2 x 2 contingency table (or a more extreme pattern). Use the Fisher Exact Test instead of χ^2 analysis of a 2 x 2 contingency table when the expected frequencies of one or more cells is less than 5.

 **Note:** SigmaPlot automatically suggests the Fisher Exact Test when a χ^2 analysis of a 2 x 2 contingency table is performed and less than 5 expected observations are encountered in any cells.

Performing a Fisher Exact Test

To perform a Fisher Exact Test:

1. Enter or arrange your data in the data worksheet.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Rates and Proportions > Fisher Exact Test
4. Run the test.
5. View and interpret the Fisher Exact Test report.

Arranging Fisher Exact Test Data

The data must form a 2×2 contingency table, with the number of observations in each cell. You can test tabulated data or raw data observations.

Table 20: A 2×2 Contingency Table describing the number of harbor seals and sea lions found on two different islands.

Pinniped Species		Island	
		Island 1	Island 2
Sea Lions		5	1
Harbor Seals		2	7

Tabulated Data. Tabulated data is arranged in a contingency table showing the number of observations for each cell. The worksheet rows and columns correspond to the groups and categories. The number of observations must always be an integer.

Raw Data. A group identifier is placed in one worksheet column and the corresponding category in another column. There must be exactly two kinds of groups and two types of categories. Each row corresponds to a single observation, so there should be as many rows of data as there are total numbers of observations.

SigmaPlot automatically cross-tabulates this data and performs the Fisher Exact Test on the resulting contingency table. For more information, see [Arranging Fisher Exact Test Data](#) on page 200.

Fisher Exact Test*				
	1-Island 1	2-Island 2	3-Pinniped	4-Island
1	5.0000	1.0000	sea lion	Island 1
2	2.0000	7.0000	sea lion	Island 1
3			sea lion	Island 1
4			sea lion	Island 1
5			sea lion	Island 1
6			harbor seal	Island 1
7			harbor seal	Island 1
8			sea lion	Island 2
9			harbor seal	Island 2
10			harbor seal	Island 2
11			harbor seal	Island 2
12			harbor seal	Island 2
13			harbor seal	Island 2

Figure 102: Data Formats for a Fisher Exact Test

Columns 1 and 2 in the worksheet above are in tabular format and columns 3 and 4 are raw data observations. A Fisher Exact Test requires data for a 2 x 2 table.

Running a Fisher Exact Test

To run a test, you need to select the data to test. Use the Test Wizard to select the worksheet columns with the data you want to test and to specify how your data is arranged in the worksheet.

To run a Fisher Exact Test:

1. If you want to select your data before you run the test, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Rates and Proportions > Fisher Exact Test

The **Fisher Exact - Data Format** dialog box appears prompting you to specify a data format.

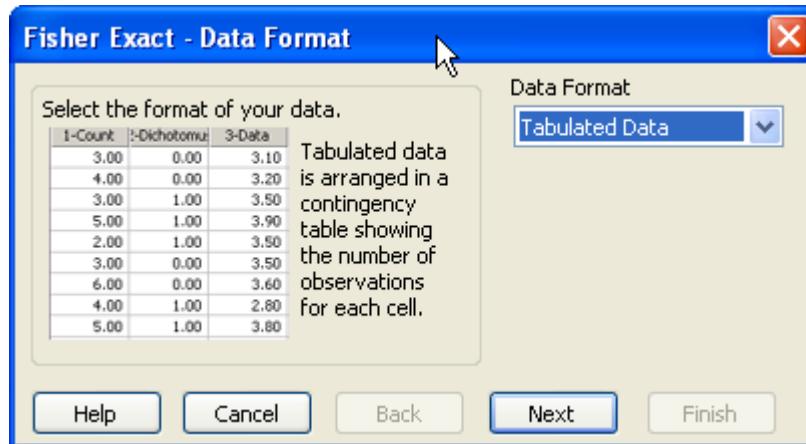


Figure 103: The Fisher Exact - Data Format Dialog Box Prompting You to Specify a Data Format

4. Select the appropriate data format from the **Data Format** drop-down list. If you are testing contingency table data, select **Tabulated**. If your data is arranged in raw format, select **Raw**. For more information, see [Arranging Fisher Exact Test Data](#) on page 200.
5. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog box prompts you to pick your data.

6. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Observations** or **Category** drop-down list.

The first selected column is assigned to the first **Observation** or **Category** row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw data, you are prompted to select up to two worksheet columns. For tabulated data you are prompted to select up to 640 columns.

7. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.

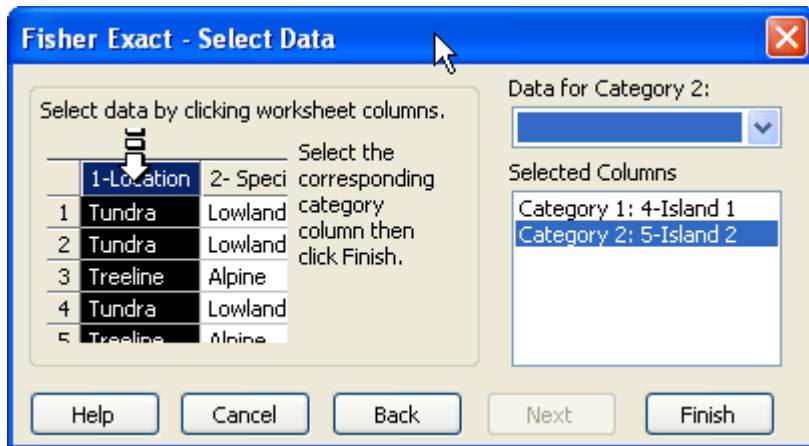


Figure 104: The Fisher Exact - Select Data Dialog Box Prompting You to Select Data Columns

8. Click **Finish** to run the test. If there are no cells in the table with expected values below 5, SigmaPlot suggests the χ^2 test instead. (You can use the Fisher Exact Test, but it takes longer to compute.)

 **Note:** The Fisher Exact Test computes the exact *two-tailed* probabilities of observing a specific 2×2 contingency table, and does not require that the expected frequencies in all cells exceed 5.

The Fisher Exact Test is performed. When the test is complete, the Fisher Exact Test report appears. For more information, see [Interpreting Results of a Fisher Exact Test](#) on page 202.

Interpreting Results of a Fisher Exact Test

Fisher Exact Test computes the two-tailed P value corresponding to the exact probability distribution of the table.

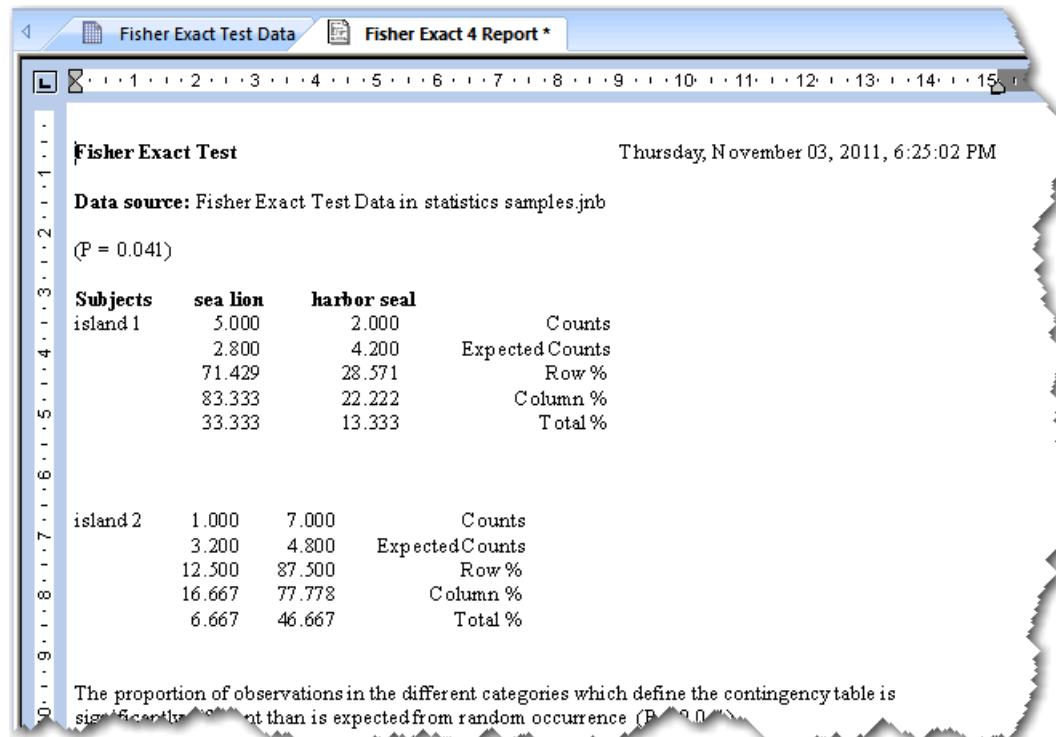


Figure 105: A Fisher Exact Test Results Report

Results Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

P Value

The *P* value is the two-tailed probability of being wrong in concluding that there is a true difference in the distribution of the numbers of observations (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error). The smaller the *P* value, the greater the probability that the samples are drawn from populations with different distributions among the two categories.

Traditionally, you conclude that there are significant differences when $P < 0.05$.



Note: The Fisher Exact Test computes *P* directly using a two tailed probability.

Contingency Table Summary

Each cell in the table is described with a set of statistics.

Observed Counts. These are the number of observations per cell, obtained from the contingency table data.

Total Cell Percentage. The percentage of total number of observations in the contingency table, obtained by dividing the observed frequency in the cells by the total number of observations in the table.

Row Percentage. The percentage of observations in each row of the contingency table, obtained by dividing the observed frequency counts in the cells by the total number of observations in that row.

Column Percentage. The percentage of observations in each column of the contingency table, obtained by dividing the observed frequency counts in the cells by the total number of observations in that column.

McNemar's Test

Use McNemar's Test when you are:

- Making observations on the same individuals.
- Counting the distributions in the same categories after two different treatments or changes in condition.

About McNemar's Test

McNemar's Test is an analysis of contingency tables that have repeated observations of the same individuals. These table designs are used when:

- Determining whether or not an individual responded to a treatment or change in condition, which uses observations before and after the treatment.
- Comparing the results of two different treatments or conditions that result in the same type of responses; for example, surveying the opinion (approve, disapprove, or don't know) of the same people before and after a report.

McNemar's Test is similar to a regular analysis of a contingency table. However, it ignores individuals who responded the same way to the same treatments, and calculates the expected frequencies using the remaining cells as the average number of individuals who responded differently to the treatments.

Performing McNemar's Test

To perform McNemar's Test:

1. Enter or arrange your data appropriately in the data worksheet.
2. View and interpret the McNemar Test report.

Arranging McNemar Test Data

The data must form a table with the same number of rows and columns, since the both treatments must have the same number of categories. You can test tabulated data or raw data observations.

Tabulated Data. Tabulated data is arranged in a contingency table showing the number of observations for each cell. The worksheet rows and columns correspond to the two groups of categories. The number of category types must be the same for both groups, so that the contingency table is square. The number of observations must always be an integer.

Table 21: A 3 x 3 Contingency Table describing the effect of a report on the opinion of surveyed people.

Before Report		After Report		
	Approve	Approve	Disapprove	Don't Know
Approve	12	24	6	
Disapprove	5	32	3	
Don't Know	4	6	4	

Raw Data A category identifier is placed in one worksheet column and the corresponding category in another column. There must be the same number of the types of categories. Each row corresponds to a single observation, so there should be as many rows of data as there are total numbers of observations.

SigmaPlot automatically cross tabulates this data and performs McNemar's Test on the resulting contingency table. For more information, see [Arranging McNemar Test Data](#) on page 204.

	1-Approve	2-Disapprove	3-Don't Know	4-Before	5-After
1	12.0000	24.0000	6.0000	Disapprove	Disapprove
2	5.0000	32.0000	3.0000	Approve	Approve
3	4.0000	6.0000	4.0000	Disapprove	Disapprove
4				Disapprove	Disapprove
5				Disapprove	Disapprove
6				Don't Know	Approve
7				Disapprove	Disapprove
8				Don't Know	Don't Know
9				Don't Know	Approve
10				Don't Know	Disapprove

Figure 106: Data Formats for McNemar's Test

Columns 1 through 3 in the worksheet above are in tabular format, and columns 4 through 6 are raw data observations. McNemar's Test requires data for tables with equal numbers of columns and rows—here a 3 x 3 table.

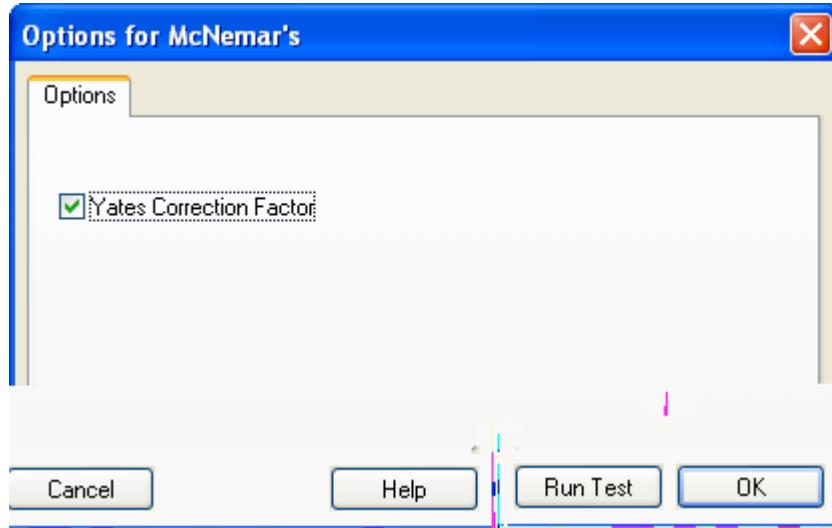
Setting McNemar's Options

Use the McNemar Test options to enable the Yates Correction Factor.

To change McNemar Test options:

1. If you are going to run the test after changing test options and want to select your data before you run the test, drag the pointer over your data.
2. Select **McNemar Test** from the **Tests** drop-down list in the **SigmaStat** group on the **Analysis** tab.
3. Click **Options**.

The **Options for McNemar's** dialog box appears.



4. Select **Yates Correction Factor** to include the Yates Correction Factor in the test report. For more information, see [Setting McNemar's Options](#) on page 204.
5. To continue the test, click **Run Test**.
6. To close the options dialog box and accept the current settings without continuing the test, click **OK**.

Options for McNemar's

Yates Correction Factor. When a statistical test uses a χ^2 distribution with one degree of freedom, such as analysis of a 2 x 2 contingency table or McNemar's test, the χ^2 calculated tends to produce P values which are too small when compared with the actual distribution of the χ^2 test statistic. The theoretical χ^2 distribution is continuous, whereas the χ^2 produced with real data is discrete.

You can use the *Yates Continuity Correction* to adjust the computed χ^2 value down to compensate for this discrepancy. Using the Yates correction makes a test more conservative; for example, it increases the P value and reduces the chance of a false positive conclusion. The Yates correction is applied to 2 x 2 tables and other statistics where the P value is computed from a χ^2 distribution with one degree of freedom.

Running McNemar's Test

To run the McNemar's Test, you need to select the data to test. Use the **Select Data** panel of the Test Wizard to select the worksheet columns with the data you want to test and to specify how your data is arranged in the worksheet.

To run McNemar's Test:

1. If you want to select your data before you run the test, drag the pointer over your data.
2. Click the **Analysis** tab.

3. In the SigmaStat group, from the **Tests** drop-down list, select:

Rates and Proportions > McNemar's Test

The **Data Format** panel of the Test Wizard appears prompting you to specify a data format.

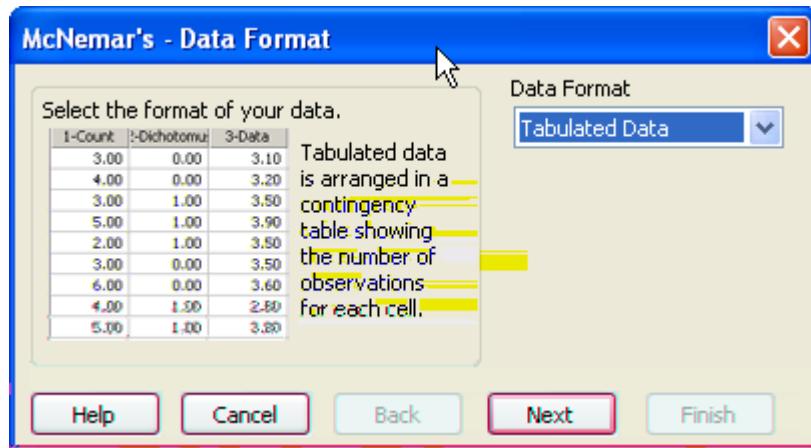


Figure 107: The McNemar's ' Data Format Dialog Box Prompting You to Specify a Data Format

- Select the appropriate data format from the **Data Format** drop-down list. If you are testing contingency table data, select Tabulated. If your data is arranged in raw format, select Raw. For more information, see [Arranging McNemar Test Data](#) on page 204.
- Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the Selected Columns list.
If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog box prompts you to pick your data.
- To assign the desired worksheet columns to the Selected Columns list, select the columns in the worksheet, or select the columns from the **Data for Observations** or **Category** drop-down list.
The first selected column is assigned to the first **Observation** or **Category** row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw data, you are prompted to select two worksheet columns. For tabulated data you are prompted to select up to 640 worksheet columns.
- To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.

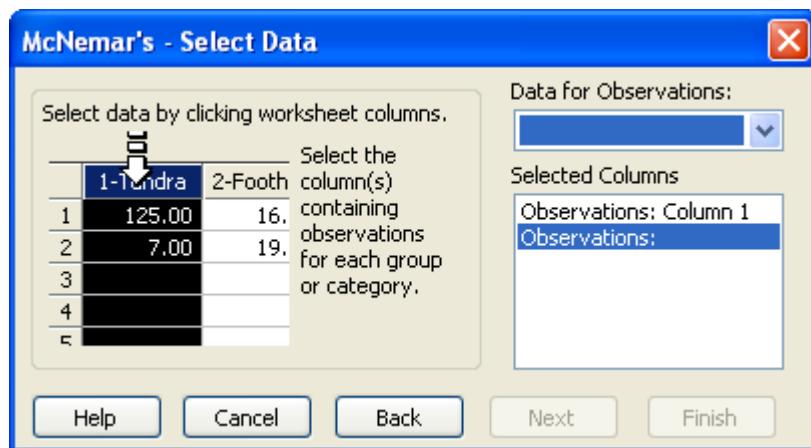


Figure 108: The McNemar's -- Select Data Dialog Box Prompting You to Select Data Columns

- Click **Finish** to run the test. The McNemar's test report appears. For more information, see [Interpreting Results of McNemar's Test](#) on page 207.

Interpreting Results of McNemar's Test

The report for McNemar's Test lists a summary of the contingency table data, the χ^2 statistic calculated from the distributions, and the P value.

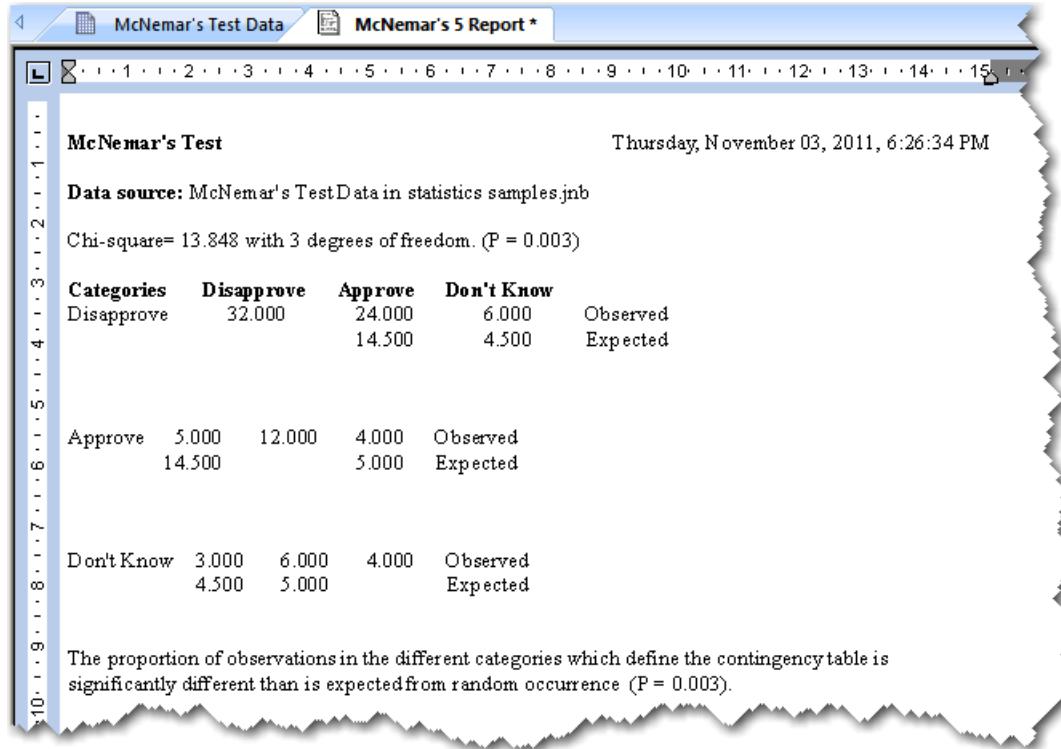


Figure 109: A McNemar Test Results Report

Results Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of tttcc

Contingency Table Summary

Each cell in the table is described with a set of statistics for that cell.

Observed Counts. These are the number of observations per cell, obtained from the contingency table data.

Expected Frequencies. The expected frequencies for each cell in the contingency table, as predicted using the row and columns percentages.

Relative Risk Test

Use the Relative Risk Test to determine if a treatment or risk factor has a significant effect on the occurrence of some event. It is usually computed for prospective studies in which the investigator initially selects subjects at random for a treatment group and a control group. At the end of the study period, the number of subjects from each group who experienced the event is counted.

About the Relative Risk Test

The Relative Risk *RR* is defined as the probability of the event occurring in the treatment group divided by the probability of the event occurring in the control group, where each probability is estimated as the relative frequency of the event in the group.

$$RR = \frac{\text{probability of event in treatment group}}{\text{probability of event in control group}}$$

The null hypothesis for the Relative Risk Test is that the value of *RR* for the entire population is 1. If the computed value of *RR* is significantly different from 1, then the treatment either significantly increases or decreases the risk of the event in the population.

The data for a Relative Risk Test can always be represented in a 2x2 contingency table. The probability of significance calculation for the test is based on the chi-square statistic for this table. If the expected number of observations for any cell of the table is less than 5, then the Fisher-Exact test is used to compute the probability. For more information, see [About the Odds Ratio Test](#) on page 212.

Performing the Relative Risk Test

To perform Relative Risk Test:

1. [Enter or arrange your data appropriately in the data worksheet.](#)
2. If desired, set the Relative Risk options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Rates and Proportions > Relative Risk
5. Run the test.
6. [View and interpret the Relative Risk Test report.](#)

Arranging Relative Risk Test Data

You can run a relative risk test using data from a contingency table entered in the worksheet or from two columns of raw data observations.

Specify the data format to use in the test in the Test Wizard dialog box.

Tabulated Data. Tabulated data is arranged in a 2x2 contingency table using the worksheet rows and columns as the groups and categories. The first column selected always represents the observations that experienced the event of interest. The two rows represent the treatment group and the control group. The group that corresponds to the first row is determined by a setting in the **Test Options** dialog box.

Raw Data The first column contains the two levels for the event (event versus no event). The second column represents the two levels of treatment (treatment versus control, or risk versus no risk). The number of rows is the total number of observations in the study. Any labels can be used in the two columns to denote the two levels for event and the two levels for treatment. To distinguish the two levels in the event column, the label in the first row will always represent the event. In the treatment column, the label that represents the treatment is determined by a setting in the **Test Options** dialog box.

Setting Relative Risk Test Options

Use the Relative Risk options to:

- Display the power of the Relative Risk Test in the report.
- Enable the Yates Correction Factor.
- Display the confidence interval for the relative risk of the population in the report.
- Determine whether the first row of the selected data represents the treatment group or the control group.

To change Relative Risk options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. Select **Relative Risk** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.
3. Click **Current Test Options**. The **Options for Relative Risk** dialog box appears.
4. Click a check box to enable or disable a test option. All options are saved between SigmaPlot sessions.
5. To continue the test, click **Run Test**.
6. To accept the current settings and close the options dialog box, click **OK**.

Options for Relative Risk

Power, Use Alpha Value. Select to detect the sensitivity of the test. The power or sensitivity of a test is the probability that the test will detect a difference between the proportions of two groups if there is really a difference.

Change the alpha value by editing the value in the Alpha Value box.

Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater probability of concluding there is no difference when one actually exists. Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive.

Yates Correction Factor. When a statistical test uses a chi-square distribution with one degree of freedom, such as analysis of a 2×2 contingency table or McNemar's test, the calculated value of the chi-square statistic tends to produce P values which are too small when compared with the actual distribution of the chi-square test distribution. The theoretical chi-square distribution is continuous, whereas the chi-square produced with real data is discrete.

The Yates continuity correction is used to adjust the chi-square statistic so that it more accurately computes P -values based on the chi-square probability distribution.

Confidence Interval. This is the confidence interval for the population value of the relative risk. To change the specified interval, select the box and type any number from 1 to 99 (95 and 99 are the most commonly used intervals).

Use the first row of the selected data as the treatment group. The Relative Risk Test assumes the population has been sampled into two groups, where the members of one group receive a treatment and the members of the other group do not. The groups are specified in the worksheet by the selection of this check box. The default is to have the check box selected so that the first row of data represents the treatment group. This option applies to both tabulated and raw data formats.

Running the Relative Risk Test

To run the Relative Risk Test, you need to select the data to test. Use the Test Wizard to select the worksheet columns with the data you want to test and to specify how your data is arranged in the worksheet.

To run Relative Risk's Test:

1. If you want to select your data before you run the test, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Rates and Proportions > Relative Risk Test

The **Relative Risk - Data Format** panel of the **Test Wizard** appears prompting you to specify a data format.

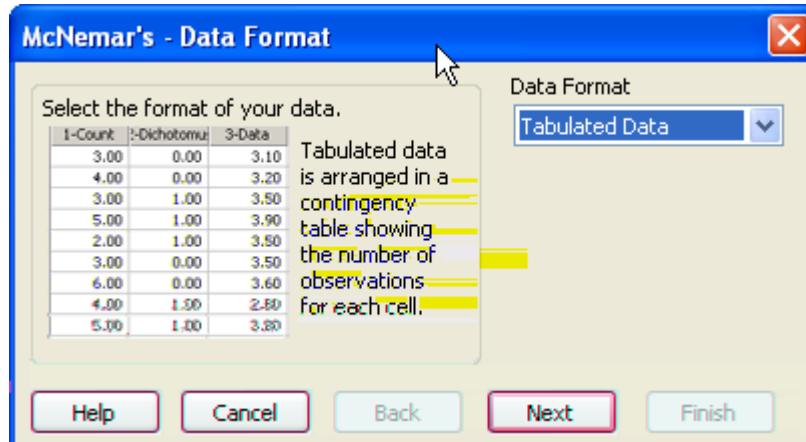


Figure 110: The Relative Risk - Data Format Dialog Box Prompting You to Specify a Data Format

4. Select the appropriate data format from the **Data Format** drop-down list. If you are testing contingency table data, select Tabulated. If your data is arranged in raw format, select Raw. For more information, see [Arranging Relative Risk Test Data](#) on page 208.
5. Click **Next** to pick the data columns for the test. If you selected columns before you chose the test, the selected columns appear in the Selected Columns list.

If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog box prompts you to pick your data.

6. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Observations** or **Category** drop-down list.

If you selected the tabulated data format, the first selected column is assigned to the **Event** row in the **Selected Columns** list and the second selected column is assigned to the **No Event** row in the list. If the raw data format was selected, the first selected column is assigned to the **Event** row in the **Selected Columns** list and the second selected column is assigned to the **Group** row in the list.

7. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.

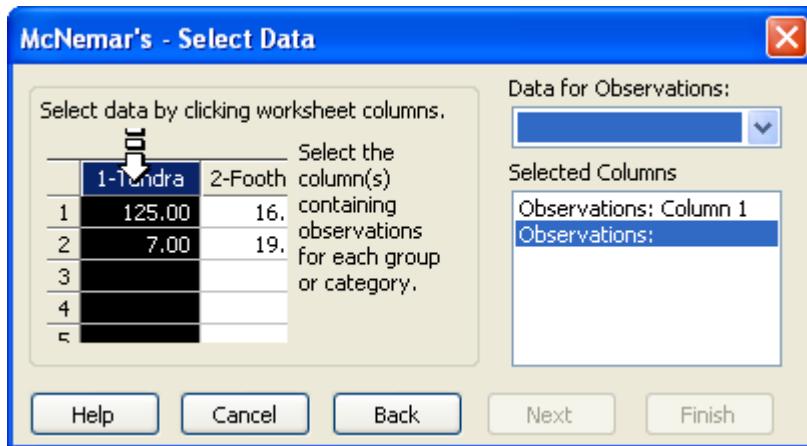


Figure 111: The Relative Risk - Select Data Dialog Box Prompting You to Select Data Columns

8. Click **Finish** to run the test. The Relative Risk test report appears. For more information, see [Interpreting Results of the Relative Risk Test](#) on page 211.

Interpreting Results of the Relative Risk Test

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the *P*-value for the Significance Level of the test and the number of decimal places to display in the Options dialog box.

Results Explanations

In addition to the usual header information, the report includes the following:

- **A 2x2 contingency table of the data.** The row and column titles are taken from the worksheet. The type of group, control or treatment, is also noted for each of the two rows.
- **The computed relative risk value.** For example, if the relative risk is 1.88, then the interpretation is that the risk of experiencing the event is estimated to be 1.88 times higher for those receiving the treatment.
- **The chi-square statistic and corresponding P-value.** If an expected number of observations in any cell of the contingency table is less than 5, then the Fisher-Exact test is used instead. For more information, see [Interpreting Results of a Chi-Squared Analysis of Contingency tables](#) on page 198.
- **The confidence interval for the population value of the relative risk.** This interval provides an alternate way of testing the null hypothesis that the relative risk equals one.
- **An interpretation of the significance probability that is different depending upon whether there is a positive result or not.** For example, if the significance level for the test is .05 and the computed *P*-value is .007, then the risk of experiencing the event is significantly different between the treatment group and the control group.
- **The value of observed power for the test based on the results for the selected data.** For more information, see [Computing Power and Sample Size](#) on page 349.

Odds Ratio Test

Use the Odds Ratio Test to determine if a treatment or risk factor has a significant effect on the occurrence of some event. The test is frequently used in case-control studies. This type of study is done retrospectively, in which the investigator samples two groups of subjects according to who did or did not experience the event. The number of subjects from each group who were exposed to the risk factor is then counted.

About the Odds Ratio Test

Odds ratio is frequently used in case-control studies. This type of study is done retrospectively, in which the investigator samples two groups of subjects from the population according to whether a subject did or did not experience the event. The number of subjects from each group who were exposed to the treatment or risk factor is then noted. The odds ratio OR is defined by:

$$OR = \frac{\text{odds of event in treatment group}}{\text{odds of event in control group}}$$

The odds ratio is an estimate of how much more likely the event occurs for an individual in the population exposed to the risk factor as compared to an individual not exposed to the risk factor.

Performing the Odds Ratio Test

To perform an Odds Ratio Test:

1. [Enter or arrange your data appropriately in the data worksheet.](#)
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select: **Rates and Proportions > Odds Ratio**
4. Run the test.
5. [View and interpret the Odds Ratio Test report.](#)

Arranging Odds Ratio Test Data

You can run an Odds Ratio test using data from a contingency table entered in the worksheet or from two columns of raw data observations.

Specify the data format to use in the test in the **Data Format** panel of the Odd Ratio Test Wizard.

Tabulated Data. Tabulated data is arranged in a 2 x 2 contingency table using the worksheet rows and columns as the groups and categories. The first column selected always represents the observations that experienced the event of interest. The two rows represent the treatment group and the control group. The group that corresponds to the first row is determined by a setting in the Test Options dialog box.

Raw Data The first column contains the two levels for the event (for example, event versus no event, or cases versus controls). The second column represents the two levels of treatment (treatment versus control, or risk versus no risk). The number of rows is the total number of observations in the study. Any labels can be used in the two columns to denote the two levels for event and the two levels for treatment. To distinguish the two levels in the event column, the label in the first row will always represent the event. In the treatment column, the label that represents the treatment is determined by a setting in the Test Options dialog box.

Setting Odds Ratio Test Options

Use the Odds Ratio options to:

- Display the power of the Relative Risk Test in the report.
- Enable the Yates Correction Factor.
- Display the confidence interval for the relative risk of the population in the report.
- Change whether the first row represents the treatment group or the control group.

To change Odds Ratio options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Current Test Options

The [Options for Odds Ratio](#) dialog box appears.

4. Click a check box to enable or disable a test option. All options are saved between SigmaPlot sessions.
5. To continue the test, click **Run Test**.
6. To accept the current settings and close the options dialog box, click **OK**.

Options for Odds Ratio

Power, Use Alpha Value. Select to detect the sensitivity of the test. The power or sensitivity of a test is the probability that the test will detect a difference between the proportions of two groups if there is really a difference.

Change the alpha value by editing the value alu

was selected, the first selected column is assigned to the **Event** row in the **Selected Columns** list and the second selected column is assigned to the **Group** row in the list.

7. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
8. Click **Finish** to run the test. The [Odds Ratio test report](#) appears.

Interpreting Results of the Odds Ratio Test

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box. For more information, see [Setting Report Options](#).

Results Explanations

In addition to the usual header information, the report includes the following:

- **A 2 x 2 contingency table of the data.** The row and column titles are taken from the worksheet. The type of group, control or treatment, is also noted for each of the two rows.
- **The computed odds ratio value.** For example, if the odds ratio is 2.91, then the interpretation is that exposure to the treatment increases the odds of experiencing the event by an estimated 2.91 time among the population.
- **The chi-square statistic and corresponding P-value.** If an expected number of observations in any cell of the contingency table is less than five, then the Fisher-Exact test is used instead. For more information, see [Interpreting Results of a Chi-Squared Analysis of Contingency tables](#) on page 198.
- **The confidence interval for the population value of the odds ratio.** This interval provides an alternate way of testing the null hypothesis that the odds ratio equals one.
- **An interpretation of the significance probability that is different depending upon whether there is a positive result or not.** For example, if the significance level for the test is .05 and the computed P-value is .007, then the risk of experiencing the event is significantly greater for those receiving the treatment.
- **The value of observed power for the test based on the results for the selected data.**

Chapter

8

Principal Components Analysis

Topics:

- About Principal Components Analysis
- Performing a Principal Components Analysis
- Arranging Data For Principal Components Analysis
- Setting Principal Components Analysis Options
- Running Principal Components Analysis
- Interpreting Principal Components Analysis Results
- Principal Components Analysis Report Graphs

Principle Components Analysis (PCA) is a multivariate technique for analyzing the complexity of high-dimensional data sets. Use Principle Components Analysis when you need to:

- Explain the sources of variability in your data.
- Approximate the data with fewer dimensions to reduce the complexity of your data.

You can use Principle Components Analysis in a number of statistical procedures, including predictive modeling and cluster analysis. *Factor Analysis*, in comparison, is a statistical method that has many similarities to PCA mathematically, but the objectives are somewhat different. The aim of Factor Analysis is to discover latent variables (or common factors) to help explain the correlations that exist between the original variables. The primary goal of Principal Components Analysis is to explain the sources of variability in the data and to represent the data with fewer variables while preserving most of the total variance.

About Principal Components Analysis

In PCA, the sampled variables forming your data are transformed into a new set of variables called *principal components*. The first component accounts for as much variation in the data as possible. Each subsequent component accounts for as much of the remaining variation as possible and is orthogonal to all of the previous components. There are as many principal components as there are variables in your data set; however, if most of the variability in your data exists in a low-dimensional subset, then you may be able to reformulate (and simplify) the model for your study in terms of only a few principal components.

The input data consists of a set of n observations where each observation is a multivariate sample of m correlated variables. Each multivariate observation is assumed to be sampled from a larger population having a multivariate normal distribution with a common population mean and covariance. The variables appear as worksheet columns and the observations as worksheet rows.

In preparation for the analysis, SigmaPlot preprocesses the raw data according to whether you've chosen to obtain the principal components by analyzing the sample covariance matrix or the sample correlation matrix. If you choose the covariance matrix is, SigmaPlot centers the column data for each variable about the mean. If choose the correlation matrix, SigmaPlot standardizes the column data for each variable to have unit sample variance. The principal component variables that the analysis creates are generated from these adjusted variables. For simplicity's sake, you can call these adjustments in the raw data as the *original variables* of your study.

The main construction in Principal Components Analysis is to derive a set of uncorrelated variables, called principal components (PC's), from the set of original variables in such a way that the following conditions are satisfied:

- Each PC is a linear combination of the original variables and, conversely, each original variable is a linear combination of the PC's. Moreover, the sum of squares of the coefficients in each linear combination is 1.
 - The total number of PC variables is usually the same as the number of original variables, but may be fewer if the set of original variables is not linearly independent.
 - The sum of the variances of the PC variables equals the total variance of the original variables.
-  **Note:** The total variance of a multivariate data set is defined as the sum of the sample variances of the variables involved.
- The PC's are ordered by their sample variances (in descending order).

From these conditions, you can explain the sources of variance for the original variables in terms of the principal components. These conditions also facilitate the interpretation of the PC's in reference to the original variables which is important in many applications. For example, a PC may represent the overall effect of all or some subset of the original variables. Several PC's may lead to interpretations that divide up the original variables into a number of distinct categories.

SigmaPlot constructs the PC's that satisfy the above conditions by obtaining the spectral decomposition of the covariance (or correlation) matrix. The *eigenvalues* in the spectral decomposition equal the variances of the PC's. The *eigenvectors* in the spectral decomposition provide the coefficients in the linear combinations described in the first condition (see above). These provide the interpretation of the PC's.

 **Note:** An eigenvalue λ of a square matrix \mathbf{S} is any number that satisfies the equation $\mathbf{S}\mathbf{x} = \lambda\mathbf{x}$ for some non-zero vector \mathbf{x} . The vector \mathbf{x} is called an eigenvector of \mathbf{S} .

Probably the most important application of PCA is data reduction. The idea is to apply some criterion for selecting a small subset of those PC's having the highest variances. Several selection criteria are available in SigmaPlot. The selected components are called the *in-model* PC's. One requirement of this subset is that the residual per observation in approximating each original variable by the best-fit linear combination of the in-model PC's be acceptably small.

 **Note:** Because of how principal components are constructed, the coefficients of the best-fit linear combination are the same as the coefficients of the in-model PC's that appear in the linear combination referenced in condition 1.

Another requirement is that most of the total variance is explained by the in-model PC's. If there are too few PC's, the residual may be too large to adequately represent the data. If there are too many PC's, the goal of simplicity may be compromised and the PC's may be difficult to interpret.

Performing a Principal Components Analysis

To perform a Principal Components Analysis:

1. Enter or arrange your data in the worksheet.
2. If desired, set the Principal Component Analysis options.
3. Click the **Analysis** tab.
4. In the **Statistics** group, from the **Tests** drop-down list, select:
 - Principal Component Analysis**
5. Run the test.
6. Generate report graphs.

Arranging Data For Principal Components Analysis

PCA data consists of a set of n observations where each observation is a multivariate sample of m correlated variables. The variables are columns in the worksheet and the observations are rows.

Figure 112: Sample Worksheet with Principal Components Data

Before the major analysis begins, SigmaPlot preprocesses the raw data according to whether or not you've chosen to analyze the *sample covariance matrix* or the *sample correlation matrix*. You [set this](#) on the **Criterion** tab on the **Options for Principal Components Analysis** dialog box.

If you select a covariance matrix, the column data for each variable is centered about the mean. If you select correlation matrix, the column data for each variable is standardized to have unit sample variance. The principal component variables created by the analysis are generated from these preprocessed variables. Any reference to the original variables of your study will refer to your data after it has been preprocessed.

Setting Principal Components Analysis Options

Use the Principal Component Analysis options to:

- Set the analysis matrix options.
- Set eigenvalues.
- Set assumption checking options.

- Specify the residuals to display and save them to the worksheet.
- Display the statistics summary for the data.
- Compute the power, or sensitivity, of the test.

Options settings are saved between SigmaPlot sessions.

To change the Paired t-test options:

1. Select **Principal Component Analysis** from the **Select Test** drop-down list in the **Statistics** group on the **Analysis** tab.
2. Click **Options**. The **Options for Principal Component Analysis** dialog box appears with four tabs:

Option	Description
Criterion	Set the matrix for analysis and the selection method for components.
Assumption Checking	Adjust the parameters of a test to relax or restrict the testing of your data for normality and equal variance.
Residuals	Select column worksheets in which you can display residuals and component scores.
Results	Display the statistics summary for the data in the report and save residuals to a worksheet column.

- i** **Tip:** If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.

Options settings are saved between SigmaPlot sessions.

3. To continue the test, click **Run Test**. The **Select Data** panel of the Test Wizard appears.
4. To accept the current settings and close the options dialog box, click **OK**

Options for Principal Components Analysis: Criterion

Click the **Criterion** tab from the options dialog box to view the **Matrix for Analysis** and **Selection Method for Components** options.

Matrix for Analysis. The matrix you choose depends on the scaling of the original variables.

- **Correlation.** Select **Correlation** if the observations are scaled differently between variables. Because **Correlation** is more common it is the program default.
- **Covariance.** Select **Covariance** if the observations are scaled approximately the same between variables. This option provides more statistics.

Selection Method for Components. The principal components that appear in the model correspond to the eigenvalues of the covariance or correlation matrix that lie above a certain threshold, where the contribution to total variability in the data is the greatest.

- **Average eigenvalue.** This is the default.
- **Minimum eigenvalue.** Select **Minimum eigenvalue** if you want all eigenvalues of the matrix greater than or equal to this value to be included in the data reduction model. The default value is zero, which includes all eigenvalues.
- **Minimum percent of total.** Select **Minimum percent of total** if you want the eigenvalues of the matrix, beginning with the largest, to be added until the sum exceeds or equals this value. Each added eigenvalue is included in the data reduction model. The default value is zero, which includes all eigenvalues.
- **Number of components.** When this option is chosen and the value entered is n, then the components corresponding to the largest n eigenvalues will be included in the data reduction model.

Significance level for hypothesis. Set the P-value. The default value is .05. Use for testing equality of eigenvalues, Hotelling's T-squared statistic for scores, and the Q-statistic for residuals.

Confidence level. Set a percentage for the confidence level and to set confidence limits for eigenvalues. Valid values are integers between 1 and 99, inclusive. The default value is 95%.

Options for Principal Components Analysis: Assumption Checking

The normality assumption test checks for a normally distributed population.

- **Tip: Equal Variance** is not available for Principal Components Analysis because Principal Components Analysis tests are based on changes in each individual rather than on different individuals in the selected population, making equal variance testing unnecessary.

Normality. For Principal Components Analysis, SigmaPlot uses either the Henze-Zinkler test or Mardia's test for skewness and kurtosis.

- **P Value to Reject.** Enter the corresponding P value in the **P Value to Reject** box. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (the P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P value computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality, decrease P. Requiring smaller values of P to reject the normality assumption means that your are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.050 requires greater deviations from normality to flag the data as non-normal than a value of 0.100.

- **Restriction:** Although the normality test is robust in detecting data from populations that are non-normal, there are extreme conditions of data distribution that this test cannot take into account; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption test.

Options for Principal Components Analysis: Residuals

Component Scores. If you'd like the results to appear in the worksheet, select this option, and then from the **in Column** drop-down list, select either **(none)** or **First Available Column**.

Residuals. For residuals to appear in the worksheet, after selecting this option, from the **in Column** drop-down list, select either **(none)** or **First Available Column**.

Options for Principal Components Analysis: Results

Summary Table. Displays the number of observations for a column or group, the number of missing values for a column or group, the average value for the column or group, the standard deviation of the column or group, and the standard error of the mean for the column or group.

Confidence Intervals. Displays the confidence interval for the difference of the means. To change the interval, enter any number from 1 to 99 (95 and 99 are the most commonly used intervals).

Residuals in Column. Displays residuals in the report and to save the residuals of the test to the specified worksheet column. Edit the number or select a number from the drop-down list.

Running Principal Components Analysis

To run a Principal Components Analysis you need to select the data to test. Use the **Select Data** panel of the Test Wizard to select the worksheet columns with the data you want to test.

To run a Principal Components Analysis:

1. If you want to select your data before you run the regression, drag the pointer over your data.

2. Click the **Analysis** tab.

3. In the **SigmaStat** group, from the **Tests** drop-down list, select **Principal Components**.

The **Select Data** panel of the **Test Wizard** appears. If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog box prompts you to pick your data.

4. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Variable** drop-down list.

The first selected column is assigned to the variable row in the **Selected Columns** list, and the second column is assigned to variable row in the list. The title of selected columns appear in each row.

 **Note:** To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.

5. Click **Next** to select **Labels**.

6. Click **Finish**.

Interpreting Principal Components Analysis Results

Basic numeric results typically appear in a Principal Components Analysis report, although there will be some differences in the output depending upon whether you are analyzing the covariance matrix or the correlation matrix for the main results.

Optional numeric results only appear depending on what you've set in the **Test Options** dialog box.

Basic Numeric Results

- **Header.** This includes the name of the test, date stamp, and data source, as for all other tests.
- **Descriptive Statistics.** A table with the mean and standard deviation for each input variable (worksheet column).
- **Total number, missing, and valid number of observations used for analysis.** A small table with three values in a column. An observation is defined as missing if at least one cell in that row for that observation is non-numeric.
- **Eigenvalues of the Covariance Matrix or Eigenvalues of the Correlation Matrix.** A table with five columns. The first column is untitled and contains an order number for the eigenvalue. The remaining four columns are titled *Eigenvalue*, *Difference*, *Proportion*, and *Cumulative*. All eigenvalues are listed in the *Eigenvalue* column in descending order of magnitude. The *Difference* column displays the differences between successive values in the *Eigenvalue* column. The *Proportion* column lists the proportion of the total variance explained by each eigenvalue. The *Cumulative* column lists the proportion of the total variance explained by all the eigenvalues before and including the current row.
- **Chi-Square Test that All Eigenvalues are Equal.** A line in the report are displayed three values: the chi-square statistic, the degrees of freedom, and the p-value. If the p-value is not significant, then the principal components will not be well-defined and the statistics will be invalid. In this case, a principal components analysis cannot be conducted. Note that the hypothesis that all eigenvalues are equal is equivalent to having all of the original variables uncorrelated with a common variance.
- **Chi-Square Test that the Last n Eigenvalues are Equal.** A line in the report will display three values: the chi-square statistic, the degrees of freedom, and the p-value. The value of *n* in the section title is a placeholder for the difference between the total number of variables in the study and the number of in-model principal components. If the p-value is not significant, then you should not add more PC's to the model. In fact, you may want to consider fewer PC's since one of the in-model x's may not be significantly different from the ones that were omitted.
- **Asymptotic Confidence Intervals for the Eigenvalues.** A table with three columns titled *Eigenvalue*, *p% Lower Limits*, and *p% Upper Limits*. All eigenvalues will be listed. The value of *p* depends upon the settings in the Test Option dialog, with the default equal to 95. This output is not available if using a correlation matrix for the analysis.
- **Unbiased Estimates of the Population Eigenvalues.** A three column table where the first column is the order number of the eigenvalue, the second column is expected value of the eigenvalue, and the third column is the standard error of the eigenvalue. This output is not available if using a correlation matrix for the analysis.
- **Eigenvectors.** A table of the eigenvectors corresponding to the in-model principal components selected for analysis. The first column is the list of the variable names (same as column titles) of the columns selected in

the Test Wizard. The number of remaining columns equals the number of in-model components. The first row contains the names of the in-model components, using the notation *PC 1*, *PC 2*, etc. The number of remaining rows equals the number of original variables.

- **Standard Errors for the Eigenvector Entries.** A table of the standard errors for each entry in each in-model eigenvector. This table will have the same size and layout as the *Eigenvectors* table above.

Optional Numeric Results

The appearance of the following sections of the report depends upon the settings in the **Test Options** dialog box. The program defaults are described in [Basic Numeric Results](#) and are restated here.

- **Normality Test.** Output for the results of the normality test. The value of *test name* in the section title is a placeholder for the name of the test, either *Mardia* or *Henze-Zirkler*. For Mardia's test, there are two lines of output, one for Skewness and one for Kurtosis. Each line will have the value of the test statistic and the word *Passed* or *Failed*, followed by the P-value. For the *Henze-Zirkler* test, there is one line of output listing the value of the test statistic, *Passed* or *Failed*, and the P-value. Displayed by default.
- **Covariance Matrix or Correlation Matrix.** The output of the matrix whose eigenvalues and eigenvectors are used to produce the principal components and statistical results. Since the matrix is symmetric, it is only necessary to display the lower triangular part of it, including the main diagonal. Displayed by default.
- **Component Loadings.** A table of the loadings for each pair of an original variable and an in-model principal component. Each original variable is a linear combination of all principal components (including out-of-model). If the principal components are standardized to have unit variance, then the coefficients in the linear combination are called *loadings*. If a correlation matrix is analyzed, then the loadings equal the correlations between the original variables and the principal components. The sum of squares of the loadings for each original variable is the variance of that variable. Displayed by default.
- **Percentage of the Variance Explained by the In-Model Components.** A table of the percentage of variance for each variable that is explained by each in-model principal component. The first column of the table is a list of the original variables. There follows a column for each in-model principal component. There is a final column titled *Unexplained Variance*. Not displayed by default.
- **Fitted Covariance Matrix or Fitted Correlation Matrix.** This is an approximation to the original covariance matrix using the in-model principal components. Each original variable is approximated by a linear combination of the in-model principal components (this is the same as projecting or regressing the original variable onto the in-model principal components). These approximations are used to create the fitted matrix. Displayed by default.
- **Original minus Fitted Covariance (or Correlation) Matrix.** This is the difference between the original covariances and the fitted covariances described above. This is so you can more easily see the errors in using the in-model principal components. Not displayed by default.
- **Component Scores.** A table of the in-model component scores for each observation in the data, together with Hotelling's T-square statistic and probability for each observation. The first column is the row number or name of the observation. There follows a column for each in-model principal component. The last two columns are the T-square statistics and their P-values. A significant P-value indicates that an observation is a possible outlier. The significant P-values will be flagged. Not displayed by default.
- **Residuals.** A table of the residuals between the original variables and the approximation of those variables using the in-model principal components. There is one column of residuals for each original variable at every observation. The first column is the row number or name of the observation. The last column is the list of Q-statistics for each observation. The Q-statistic is the sum of squares for the residuals for a given multivariate observation. A value of Q is significant if it exceeds the corresponding critical value and this is indicated by flagging the value in the Q-column. A significant value indicates either that the current set of in-model components do not fit the observation or that the observation may be an outlier. Not displayed by default.

Principal Components Analysis Report Graphs

Depending on the number of in-model principal components, the Principal Component Analysis can generate up to three graph types:

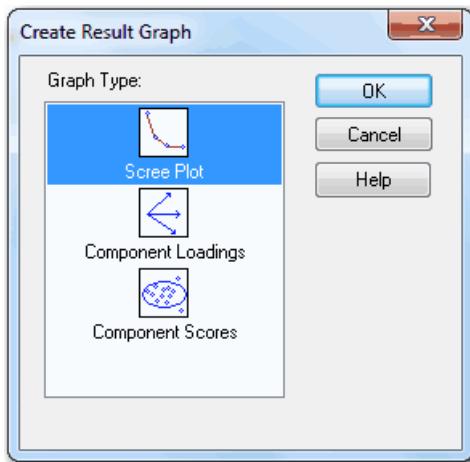
- [A scree plot](#)

- A component loadings plot
- A component scores plot

Creating a Graph for Principal Components Analysis

1. Select the **Principal Components Analysis** test report.
2. Click the **Report** tab.
3. In the **Results Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the **Principal Components Analysis** results.

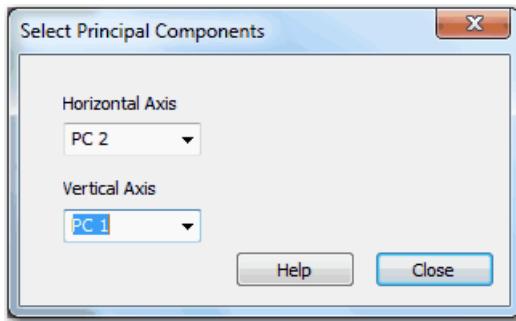


4. Select the type of graph you want to create from the **Graph Type** list.

Important:

If you select **Component Loadings** graph type, you need to select the components to represent the horizontal and vertical axes in the **Select Principal Components** dialog box.

If you are creating a **Component Scores** graph, then you have the additional option of adding a **Prediction Ellipse** which indicates the possible outliers in the distribution of the data.



5. Click **OK**, or double-click the desired graph in the list.

The selected graph appears in a [graph window](#).

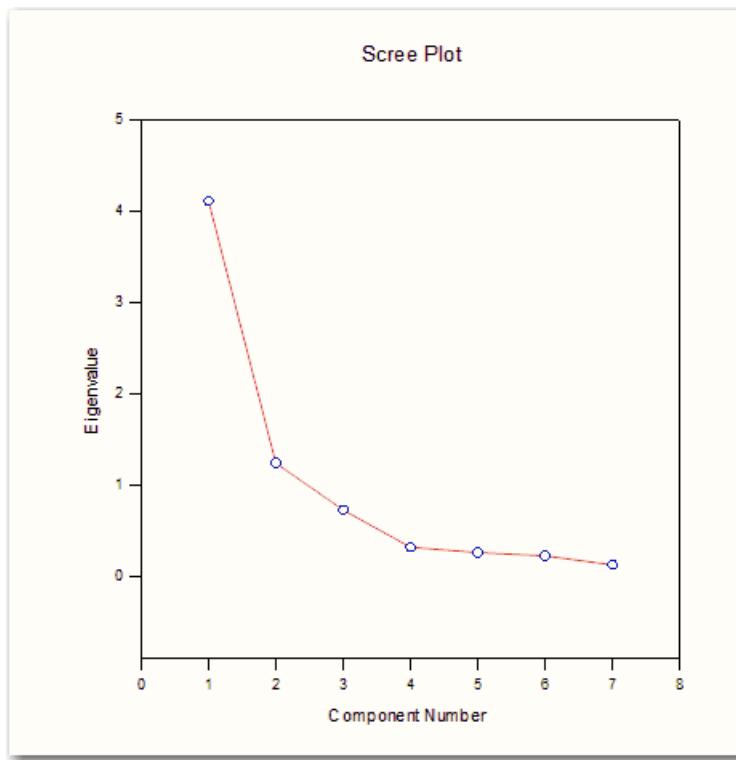


Figure 113: A Scree Plot of the Principal Component Analysis

Chapter

9

Prediction and Correlation

Topics:

- About Regression
- Simple Linear Regression
- Multiple Linear Regression
- Multiple Logistic Regression
- Polynomial Regression
- Stepwise Linear Regression
- Best Subsets Regression
- Pearson Product Moment Correlation
- Spearman Rank Order Correlation
- Deming Regression

Prediction uses regression and correlation techniques to describe the relationship between two or more variables. For more information, see [Choosing the Prediction or Correlation Method](#) on page 34.

About Regression

Regression procedures use the values of one or more independent variables to predict the value of a dependent variable. The independent variables are the known, or predictor, variables. When the independent variables are varied, they result in a corresponding value for the dependent, or response, variable.

You can perform regressions using seven different methods.

- [Simple Linear Regression](#).
- [Multiple Linear Regression](#).
- [Multiple Linear Logistic Regression](#).
- [Polynomial Regression](#).
- [Stepwise Regression, both forwards and backwards](#).
- [Best Subset Regression](#).
- [Deming Regression](#).

Regression assumes an association between the independent and dependent variables that, when graphed on a Cartesian coordinate system, produces a straight line, plane, or curve. Regression finds the equation that most closely describes the actual data.

For example, Simple Linear Regression uses the equation for a straight line $y=b_0+b_1x$ where y is the dependent variable, x is the independent variable, b_0 is the intercept, or constant term (the value of the dependent variable when $x=0$, the point where the regression line intersects the Y axis), and b_1 is the slope, or regression coefficient (increase in the value of Y per unit increase in X). As the values for X increase by 1, the corresponding values for Y either increase or decrease by b_1 , depending on the sign of b_1 .

Multiple Linear Regression is similar to simple linear regression, but uses multiple independent variables to fit the general equation for a multidimensional plane $y=b_0+b_1x_1+b_2x_2+b_3x_3+\dots+b_kx_k$ where y is the dependent variable, $x_1, x_2, x_3, \dots, x_k$ are the k independent variables, and $b_1, b_2, b_3, \dots, b_k$ are the k regression coefficients. As the values for x_1 increase by 1, the corresponding value for y either increases or decreases by b_k depending on the sign of b_k .

Regression is a parametric statistical method that assumes that the residuals (differences between the predicted and observed values of the dependent variables) are normally distributed with constant variance.

Because the regression coefficients are computed by minimizing the sum of squared residuals, this technique is often called least squares regression.

Correlation

Correlation procedures measure the strength of association between two variables, which can be used as a gauge of the certainty of prediction. Unlike regression, it is not necessary to define one variable as the independent variable and one as the dependent variable.

The correlation coefficient r is a number that varies between -1 and $+1$. A correlation of -1 indicates there is a perfect negative relationship between the two variables, with one always decreasing as the other increases. A correlation of $+1$ indicates there is a perfect positive relationship between the two variables, with both always increasing together. A correlation of 0 indicates no relationship between the two variables.

There are two types of correlation coefficients.

- The [Pearson Product Moment Correlation](#), a parametric statistic which assumes a normal distribution and constant variance of the residuals.
- The [Spearman Rank Order Correlation](#), a nonparametric association test that does not require assuming normality or constant variance of the residuals.

Data Format for Regression and Correlation

Data for all regression and correlation procedures consists of the dependent variables (usually the " y " data) in one column, and the independent variables (usually the " x " data) in one or more additional columns, one column for each independent variable.

Regression ignores rows containing missing data points within columns of data. Blank cells, double dashes ("--"), and text items are considered missing values. All the columns must be of equal length, including missing values, or you will receive an error message.

If you plan to test blocks of data instead of picking columns, the columns must be adjacent, and the left-most column is assumed to be the dependent variable.

See the Selecting Data Columns sections under each test for information on selecting blocks of data instead of entire columns.

Simple Linear Regression

Use Linear Regression when:

- You want to predict a trend in data, or predict the value of a variable from the value of another variable, by fitting a straight line through the data.
- You know there is exactly one independent variable.

The independent variable is the known, or predicted, variable, such as time or temperature. When the independent variable is varied, it produces a corresponding value for the dependent, or response, variable. If you know there is more than one independent variable, use multiple linear regression.

About Simple Linear Regression

Linear Regression assumes an association between the independent and dependent variable that, when graphed on a Cartesian coordinate system, produces a straight line. Linear Regression finds the straight line that most closely describes, or predicts, the value of the dependent variable, given the observed value of the independent variable.

The equation used for a Simple Linear Regression is the equation for a straight line, or $y=b_0+b_1x$ where y is the dependent variable, x is the independent variable, b_0 is the intercept, or constant term (value of the dependent variable when $x=0$, the point where the regression line intersects the y axis), and b_1 is the slope, or regression coefficient (increase in the value of y per unit increase in x). As the values for x increase, the corresponding value for y either increases or decreases by b_1 is the slope, or regression coefficient (increase in the value of y per unit increase in x). As the values for x increase, the corresponding value for y either increases or decreases by b_1 , depending on the sign of b_1 .

Linear Regression is a parametric test, that is, for a given independent variable value, the possible values for the dependent variable are assumed to be normally distributed with constant variance around the regression line.

Performing a Simple Linear Regression

To perform a Simple Linear Regression:

1. [Enter or arrange your data in the worksheet.](#)
2. If desired, set the Linear Regression options.
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Regression > Linear
5. [Generate report graphs.](#)
6. Run the test.

Arranging Simple Linear Regression Data

Place the data for the observed dependent variable in one column and the data for the corresponding independent variable in a second column. Observations containing missing values are ignored, and both columns must be equal in length.

Setting Simple Linear Regression Options

Use the Linear Regression options to:

- Set assumption checking options.
- Specify the residuals to display and save them to the worksheet.
- Display confidence intervals and save them to the worksheet.
- Display the PRESS Prediction Error and standardized regression coefficients.
- Specify tests to identify outlying or influential data points.
- Display power.

To change Linear Regression options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. Select the **Analysis** tab.
3. Select a check box to enable or disable a test option. Options settings are saved between SigmaPlot sessions. For more information, see [Interpreting Simple Linear Regression Results](#) on page 232.
4. In the **SigmaPlot** group, click **Options**. The Options for Linear Regression dialog box appears with four tabs:
 - **Assumption Checking.** Click the **Assumption Checking** tab to return to the Normality, Constant Variance, and Durbin-Watson options.
 - **Residuals.** Click the **Residuals** tab to view the residual options.
 - **More Statistics.** Click the **More Statistics** tab to view the confidence intervals, PRESS Prediction Error, and Standardized Coefficients options.
 - **Other Diagnostics.** Click the **Other Diagnostics** tab to view the Influence and Power options.
5. To continue the test, click **Run Test**.
6. To accept the current settings and close the options dialog box, click **OK**.

Options for Linear Regression: Assumption Checking

Select the **Assumption Checking** tab from the options dialog box to view the **Normality**, **Constant Variance**, and **Durbin-Watson** options. These options test your data for its suitability for regression analysis by checking three assumptions that a linear regression makes about the data. A linear regression assumes:

- That the source population is normally distributed about the regression.
- The variance of the dependent variable in the source population is constant regardless of the value of the independent variable(s).
- That the residuals are independent of each other.

All assumption checking options are selected by default. Only disable these options if you are certain that the data was sampled from normal populations with constant variance and that the residuals are independent of each other.

Normality Testing. SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

Constant Variance Testing. SigmaPlot tests for constant variance by computing the Spearman rank correlation between the absolute values of the residuals and the observed value of the dependent variable. When this correlation is significant, the constant variance assumption may be violated, and you should consider trying a different model (for example, one that more closely follows the pattern of the data), or transforming one or more of the independent variables to stabilize the variance.

P Values for Normality and Constant Variance. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and/or constant variance, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.05. Larger values of P (for example, 0.10) require less evidence to conclude that the residuals are not normally distributed or the constant variance assumption is violated.

To relax the requirement of normality and/or constant variance, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.01 for the normality test requires greater deviations from normality to flag the data as non-normal than a value of 0.05.

- i** **Tip:** Although the assumption tests are robust in detecting data from populations that are non-normal or with non-constant variances, there are extreme conditions of data distribution that these tests cannot detect; however, these conditions should be easily detected by visually examining the data without resorting to the automatic assumption tests.

Durbin-Watson Statistic. SigmaPlot uses the Durbin-Watson statistic to test residuals for their independence of each other. The Durbin-Watson statistic is a measure of serial correlation between the residuals. The residuals are often correlated when the independent variable is time, and the deviation between the observation and the regression line at one time are related to the deviation at the previous time. If the residuals are not correlated, the Durbin-Watson statistic will be 2.

Difference from 2 Value. Enter the acceptable deviation from 2.0 that you consider as evidence of a serial correlation in the Difference for 2.0 box. If the computed Durbin-Watson statistic deviates from 2.0 more than the entered value, SigmaPlot warns you that the residuals may not be independent. The suggested deviation value is 0.50, for example, Durbin-Watson Statistic values greater than 2.5 or less than 1.5 flag the residuals as correlated.

To require a stricter adherence to independence, decrease the acceptable difference from 2.0.

To relax the requirement of independence, increase the acceptable difference from 2.0.

Options for Linear Regression: Residuals

Select the Residuals tab in the options dialog box to view the Predicted Values, Raw, Standardized, Studentized, Studentized Deleted, and Report Flagged Values Only options.

Predicted Values. Use this option to calculate the predicted value of the dependent variable for each observed value of the independent variable(s), then save the results to the worksheet. Click the selected check box if you do not want to include raw residuals in the worksheet.

To assign predicted values to a worksheet column, select the worksheet column you want to save the predicted values to from the corresponding drop-down list. If you select none and the Predicted Values check box is selected, the values appear in the report but are not assigned to the worksheet.

Raw Residuals. The raw residuals are the differences between the predicted and observed values of the dependent variables. To include raw residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include raw residuals in the worksheet.

To assign the raw residuals to a worksheet column, select the number of the desired column from the corresponding drop-down list. If you select none from the drop-down list and the Raw check box is selected, the values appear in the report but are not assigned to the worksheet.

Standardized Residuals. The standardized residual is the residual divided by the standard error of the estimate. The standard error of the residuals is essentially the standard deviation of the residuals, and is a measure of variability around the regression line. To include standardized residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include raw residuals in the worksheet.

SigmaPlot automatically flags data points lying outside of the confidence interval specified in the corresponding box. These data points are considered to have "large" standardized residuals, for example, outlying data points. You can change which data points are flagged by editing the value in the Flag Values > edit box. The suggested residual value is 2.5.

Studentized Residuals. Studentized residuals scale the standardized residuals by taking into account the greater precision of the regression line near the middle of the data versus the extremes. The Studentized residuals tend to be distributed according to the Student t distribution, so the t distribution can be used to define "large" values of the Studentized residuals. SigmaPlot automatically flags data points with "large" values of the Studentized residuals, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

To include Studentized residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include Studentized residuals in the worksheet.

Studentized Deleted Residuals. Studentized deleted residuals are similar to the Studentized residual, except that the residual values are obtained by computing the regression equation without using the data point in question.

To include Studentized deleted residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include Studentized deleted residuals in the worksheet.

SigmaPlot can automatically flag data points with "large" values of the Studentized deleted residual, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

 **Note:** Both Studentized and Studentized deleted residuals use the same confidence interval setting to determine outlying points.

Report Flagged Values Only. To include only the flagged standardized and Studentized deleted residuals in the report, make sure the Report Flagged Values Only check box is selected. Clear this option to include all standardized and Studentized residuals in the report.

Options for Linear Regression: More Statistics

Click the **More Statistics** tab in the **Options for Linear Regression** dialog box to view the confidence interval options. You can set the confidence interval for the population, regression, or both and then save them to the worksheet.

Confidence Interval for the Population. The confidence interval for the population gives the range of values that define the region that contains the population from which the observations were drawn.

To include confidence intervals for the population in the report, select **Population**. Clear the check box if you do not want to include the confidence intervals for the population in the report.

Confidence Interval for the Regression. The confidence interval for the regression line gives the range of values that defines the region containing the true mean relationship between the dependent and independent variables, with the specified level of confidence.

To include confidence intervals for the regression in the report, select **Regression**, then specify a confidence level by entering a value in the percentage box. The confidence level can be any value from 1 to 99. The suggested confidence level for all intervals is 95%. Clear the **Regression** check box if you do not want to include the confidence intervals for the population in the report.

Saving Confidence Intervals to the Worksheet. To save the confidence intervals to the worksheet, select the column number of the first column you want to save the intervals to from the **Starting in Column** drop-down list. The selected intervals are saved to the worksheet starting with the specified column and continuing with successive columns in the worksheet.

PRESS Prediction Error. The PRESS Prediction Error is a measure of how well the regression equation predicts the observations. Leave this check box selected to evaluate the fit of the equation using the PRESS statistic. Clear the selected check box if you do not want to include the PRESS statistic in the report.

Options for Linear Regression: Other Diagnostics

Click the **Other Diagnostics** tab in the **Options for Linear Regression** dialog box to view the **Influence** options.

Influence. Influence options automatically detect instances of influential data points. Most influential points are data points which are outliers, that is, they do not "line up" with the rest of the data points. These points can have a potentially disproportionately strong influence on the calculation of the regression line. You can use several influence tests to identify and quantify influential points.

- **DFFITS.** DFFITS i is the number of estimated standard errors that the predicted value changes for the i th data point when it is removed from the data set. It is another measure of the influence of a data point on the prediction used to compute the regression coefficients.

Predicted values that change by more than two standard errors when the data point is removed are considered to be influential.

Select **DFFITS** to compute this value for all points and flag influential points, for example, those with DFFITS greater than the value specified in the **Flag Values** > edit box. The suggested value is 2.0 standard errors, which indicates that the point has a strong influence on the data. To avoid flagging more influential points, increase this value; to flag less influential points, decrease this value.

- **Leverage.** Leverage is used to identify the potential influence of a point on the results of the regression equation. Leverage depends only on the value of the independent variable(s). Observations with high leverage tend to be at the extremes of the independent variables, where small changes in the independent variables can have large effects on the predicted values of the dependent variable.

The expected leverage of a data point is $\frac{1}{n}$, where there are k independent variables and n data points. Observations with leverages much higher than the expected leverages are potentially influential points.

Select **Leverage** to compute the leverage for each point and automatically flag potentially influential points, for example, those points that could have leverages greater than the specified value times the expected leverage. The suggested value is 2.0 times the expected leverage for the regression. To avoid flagging more potentially influential points, increase this value; to flag points with less potential influence, lower this value.

- **Cook's Distance.** Cook's distance is a measure of how great an effect each point has on the estimates of the parameters in the regression equation. Cook's distance assesses how much the values of the regression coefficients change if a point is deleted from the analysis. Cook's distance depends on both the values of the independent and dependent variables.

Select **Cook's Distance** to compute this value for all points and flag influential points, for example, those with a Cook's distance greater than the specified value. The suggested value is 4.0. Cook's distances above 1 indicate that a point is possibly influential. Cook's distances exceeding 4 indicate that the point has a major effect on the values of the parameter estimates. To avoid flagging more influential points, increase this value: to flag less influential points, lower this value.

Report Flagged Values Only. To only include only the influential points flagged by the influential point tests in the report, make sure you've selected **Report Flagged Values Only**. Clear this check box to include all influential points in the report.

What to Do About Influential Points

Influential points have two possible causes:

- There is something wrong with the data point, caused by an error in observation or data entry.
- The model is incorrect.

If you made a mistake in data collection or entry, correct the value. If you do not know the correct value, you may be able to justify deleting the data point. If the model appears to be incorrect, try regression with different independent variables, or a Nonlinear Regression.

Power. The power of a regression is the power to detect the observed relationship in the data. The alpha (α) is the acceptable probability of incorrectly concluding there is a relationship.

Select **Power** to compute the power for the linear regression data. Change the alpha value by editing the number in the **Alpha Value** edit box. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant relationship when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant relationship, but a greater possibility of concluding there is no relationship when one exists. Larger values of α make it easier to conclude that there is a relationship, but also increase the risk of reporting a false positive.

Running a Simple Linear Regression

To run a Simple Linear Regression, you need to select the data to test. You use the **Select Data** panel of the Test Wizard to select the worksheet columns with the data you want to test.

To run a Linear Regression:

1. If you want to select your data before you run the test, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Regression > Linear

The **Select Data** panel of the Test Wizard appears. If you selected columns before you chose the test, the columns appear in the **Selected Columns** list. If you have not selected columns, the dialog box prompts you to pick your data.

4. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Dependent** or **Data for Independent** drop-down list.

The first selected column is assigned to the dependent row in the **Selected Columns** list, and the second column is assigned to independent row in the list. The title of selected columns appear in each row. You can only select one dependent and one independent data column.

5. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
6. Click **Finish** to run the regression. If you elected to test for normality, constant variance, and independent residuals, SigmaPlot performs the tests for normality (Shapiro-Wilk or Kolmogorov-Smirnov), constant variance, and independent residuals. If your data fail either of these tests, SigmaPlot warns you. When the test is complete, the **Simple Linear Regression** report appears.

If you selected to place predicted values and residuals the worksheet, they are placed in the specified column and are labeled by content and source column.

Interpreting Simple Linear Regression Results

The report for a Linear Regression displays the equation with the computed coefficients for the line, R , R^2 , and adjusted R^2 , a table of statistical values for the estimate of the dependent variable, and the P values for the regression equation and for the individual coefficients.

The other results displayed in the report are enabled and disabled **Options for Linear Regression** dialog box.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box. For more information, see [Report Graphs](#) on page 373.

Regression Equation

This is the equation for a line with the values of the coefficients—the intercept (constant) and the slope—in place.

This equation takes the form: $y = b_0 + b_1x$ where y is the dependent variable, x is the independent variable, b_0 is the constant, or intercept (value of the dependent variable when $x = 0$, the point where the regression line intersects the y axis), and b_1 is the slope (increase in the value of y per unit increase in x).

The number of observations N , and the number of observations containing missing values (if any) that were omitted from the regression, are also displayed.

R, R Squared, and Adj R Squared

R , the correlation coefficient, and R^2 , the coefficient of determination, are both measures of how well the regression model describes the data. R values near 1 indicate that the straight line is a good description of the relation between the independent and dependent variable.

R equals 0 when the values of the independent variable do not allow any prediction of the dependent variables, and equals 1 when you can perfectly predict the dependent variable from the independent variable.

Adjusted R Squared. The adjusted R^2_{adj} is also a measure of how well the regression model describes the data, but takes into account the number of independent variables, which reflects the degrees of freedom. Larger values (nearer to 1) indicate that the equation is a good description of the relation between the independent and dependent variables.

Standard Error of the Estimate

The standard error of the estimate s_{yx} is a measure of the actual variability about the regression line of the underlying population. The underlying population generally falls within about two standard errors of the observed sample.

Statistical Summary Table

Coefficients. The value for the constant (intercept) and coefficient of the independent variable (slope) for the regression model are listed.

Standard Error. The standard errors of the intercept and slope are measures of the precision of the estimates of the regression coefficients (analogous to the standard error of the mean). The true regression coefficients of the underlying population generally fall within about two standard errors of the observed sample coefficients. These values are used to compute t and confidence intervals for the regression.

t Statistic. The t statistic tests the null hypothesis that the coefficient of the independent variable is zero, that is, the independent variable does not contribute to predicting the dependent variable. t is the ratio of the regression coefficient to its standard error, or

$$t = \frac{\text{regression coefficient}}{\text{standard error of regression coefficient}}$$

You can conclude from "large" t values that the independent variable can be used to predict the dependent variable (for example, that the coefficient is not zero).

P Value. P is the P value calculated for t . The P value is the probability of being wrong in concluding that there is a true association between the variables (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on t). The smaller the P value, the greater the probability that the independent variable can be used to predict the dependent variable.

Traditionally, you can conclude that the independent variable can be used to predict the dependent variable when $P < 0.05$.

Beta (Standardized Coefficient b)

$$\pi_1 = b_1 \frac{s_x}{s_y}$$

This is the coefficient of the independent variable standardized to dimensionless values, where b_1 = regression coefficient, s_x = standard deviation of the independent variable x , and s_y = standard deviation of dependent variable y .

This result is displayed unless the Standardized Coefficients option is disabled in the Options for Linear Regression dialog box.

Analysis of Variance (ANOVA) Table

The ANOVA (analysis of variance) table lists the ANOVA statistics for the regression and the corresponding F value.

DF (Degrees of Freedom). Degrees of freedom represent the number of observations and variables in the regression equation.

- The regression degrees of freedom is a measure of the number of independent variables in the regression equation (always 1 for simple linear regression)
- The residual degrees of freedom is a measure of the number of observations less the number of terms in the equation
- The total degrees of freedom is a measure of total observations

SS (Sum of Squares). The sum of squares are measures of variability of the dependent variable.

- The sum of squares due to regression (SSreg) measures the difference of the regression line from the mean of the dependent variable
- The residual sum of squares (SSres) is a measure of the size of the residuals, which are the differences between the observed values of the dependent variable and the values predicted by regression model

- The total sum of squares (SStot) is a measure of the overall variability of the dependent variable about its mean value

MS (Mean Square). The mean square provides two estimates of the population variances. Comparing these variance estimates is the basis of analysis of variance.

The mean square regression is a measure of the variation of the regression from the mean of the dependent variable, or

$$\frac{\text{sum of squares due to regression}}{\text{regression degrees of freedom}} = \frac{SS_{\text{reg}}}{DF_{\text{reg}}} = MS_{\text{reg}}$$

The residual mean square is a measure of the variation of the residuals about the regression line, or

$$\frac{\text{residual sum of squares}}{\text{residual degrees of freedom}} = \frac{SS_{\text{res}}}{DF_{\text{res}}} = MS_{\text{res}}$$

The residual mean square is also equal to s_{yx}^2 .

F Statistic. The F test statistic gauges the contribution of the independent variable in predicting the dependent variable. It is the ratio

$$\frac{\text{regression variation from the dependent variable mean}}{\text{residual variation about the regression line}} = \frac{MS_{\text{reg}}}{MS_{\text{res}}} = F$$

If F is a large number, you can conclude that the independent variable contributes to the prediction of the dependent variable (for example, the slope of the line is different from zero, and the "unexplained variability" is smaller than what is expected from random sampling variability). If the F ratio is around 1, you can conclude that there is no association between the variables (for example, the data is consistent with the null hypothesis that all the samples are just randomly distributed about the population mean, regardless of the value of the independent variable).

P Value. The P value is the probability of being wrong in concluding that there is an association between the dependent and independent variables (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on F). The smaller the P value, the greater the probability that there is an association.

Traditionally, you can conclude that the independent variable can be used to predict the dependent variable when $P < 0.05$.

- **Tip:** In simple linear regression, the P value for the ANOVA is identical to the P value associated with the t of the slope coefficient, and $F=t^2$, where t is the t value associated with the slope.

PRESS Statistic

PRESS, the Predicted Residual Error Sum of Squares, is a measure of how well a regression model predicts the observations.

The PRESS statistic is computed by summing the squares of the prediction errors (the differences between predicted and observed values) for each observation, with that point deleted from the computation of the estimated regression model.

One important use of the PRESS statistics is for model comparison. If several different regression models are applied to the same data, the one with the smallest PRESS statistic has the best predictive capability.

Durbin-Watson Statistic

The Durbin-Watson statistic is a measure of correlation between the residuals. If the residuals are not correlated, the Durbin-Watson statistic will be 2; the more this value differs from 2, the greater the likelihood that the residuals are correlated. This result appears if it was selected in the Regression Options dialog box.

Regression assumes that the residuals are independent of each other; the Durbin-Watson test is used to check this assumption. If the Durbin-Watson value deviates from 2 by more than the value set in the Options for Linear Regression dialog box, a warning appears in the report. The suggested trigger value is a difference of more than 0.50 (for example, if the Durbin-Watson statistic is below 1.5 or over 2.5).

Normality Test

Normality test result displays whether the data passed or failed the test of the assumption that the source population is normally distributed around the regression line, and the P value calculated by the test. All regressions assume a source population to be normally distributed about the regression line. When this assumption may be violated, a warning appears in the report. This result appears unless you disabled normality testing in the Options for Linear Regression dialog box.

Failure of the normality test can indicate the presence of outlying influential points or an incorrect regression model.

Constant Variance Test

The constant variance test result displays whether or not the data passed or failed the test of the assumption that the variance of the dependent variable in the source population is constant regardless of the value of the independent variable, and the P value calculated by the test. When the constant variance assumption may be violated, a warning appears in the report.

If you receive this warning, you should consider trying a different model (for example, one that more closely follows the pattern of the data), or transforming the independent variable to stabilize the variance and obtain more accurate estimates of the parameters in the regression equation.

Power

This result is displayed if you selected this option in the options dialog box. The power, or sensitivity, of a performed regression is the probability that the model correctly describes the relationship of the variables, if there is a relationship.

Regression power is affected by the number of observations, the chance of erroneously reporting a difference α (alpha), and the correlation coefficient r associated with the regression.

Alpha (α). Alpha (α) is the acceptable probability of incorrectly concluding that the model is correct. An α error is also called a Type I error (a Type I error is when you reject the hypothesis of no association when this hypothesis is true).

Set the value in the Power Options dialog box; the suggested value is $\alpha = 0.05$ which indicates that a one in twenty chance of error is acceptable. Smaller values of α result in stricter requirements before concluding the model is correct, but a greater possibility of concluding the model is bad when it is really correct (a Type II error). Larger values of α make it easier to conclude that the model is correct, but also increase the risk of accepting a bad model (a Type I error).

Regression Diagnostics

The regression diagnostic results display only the values for the predicted values, residual results, and other diagnostics selected in the Options for Regression dialog box. All results that qualify as outlying values are flagged with a $<$ symbol. The trigger values to flag residuals as outliers are set in the Options for Linear Regression dialog box.

If you selected Report Cases with Outliers Only, only those observations that have one or more residuals flagged as outliers are reported; however, all other results for that observation are also displayed.

Row. This is the row number of the observation.

Predicted Values. This is the value for the dependent variable predicted by the regression model for each observation.

Residuals. These are the raw residuals, the difference between the predicted and observed values for the dependent variables.

Standardized Residuals. The standardized residual is the raw residual divided by the standard error of the estimate S_{yx} .

If the residuals are normally distributed about the regression line, about 66% of the standardized residuals have values between -1 and +1, and about 95% of the standardized residuals have values between -2 and +2. A larger standardized residual indicates that the point is far from the regression line; the suggested value flagged as an outlier is 2.5.

Studentized Residuals. The Studentized residual is a standardized residual that also takes into account the greater confidence of the predicted values of the dependent variable in the "middle" of the data set. By weighting the values of the residuals of the extreme data points (those with the lowest and highest independent variable values), the Studentized residual is more sensitive than the standardized residual in detecting outliers.

Both Studentized and Studentized deleted residuals that lie outside a specified confidence interval for the regression are flagged as outlying points; the suggested confidence value is 95%.

This residual is also known as the internally Studentized residual because the standard error of the estimate is computed using all data.

Studentized Deleted Residuals. The Studentized deleted residual, or externally Studentized residual, is a Studentized residual which uses the standard error of the estimate $s_{yx(-i)}$, computed after deleting the data point associated with the residual. This reflects the greater effect of outlying points by deleting the data point from the variance computation.

Both Studentized and Studentized deleted residuals that lie outside a specified confidence interval for the regression are flagged as outlying points; the suggested confidence value is 95%.

The Studentized deleted residual is more sensitive than the Studentized residual in detecting outliers, since the Studentized deleted residual results in much larger values for outliers than the Studentized residual.

Influence Diagnostics

The influence diagnostic results display only the values for the results selected in the Options dialog box under the Other Diagnostics tab. All results that qualify as outlying values are flagged with a < symbol. The trigger values to flag data points as outliers are also set in the Options dialog box under the Other Diagnostics tab.

If you selected Report Cases with Outliers Only, only observations that have one or more observations flagged as outliers are reported; however, all other results for that observation are also displayed.

Row. This is the row number of the observation.

Cook's Distance. Cook's distance is a measure of how great an effect each point has on the estimates of the parameters in the regression equation. It is a measure of how much the values of the regression equation would change if that point is deleted from the analysis.

Values above 1 indicate that a point is possibly influential. Cook's distances exceeding 4 indicate that the point has a major effect on the values of the parameter estimates. Points with Cook's distances greater than the specified value are flagged as influential; the suggested value is 4.

Leverage. Leverage values identify potentially influential points. Observations with leverages a specified factor greater than the expected leverages are flagged as potentially influential points; the suggested value is 2.0 times the expected leverage.

$$\frac{k+1}{n}$$

The expected leverage of a data point is $\frac{k+1}{n}$, where there are k independent variables and n data points.

Because leverage is calculated using only the dependent variable, high leverage points tend to be at the extremes of the independent variables (large and small values), where small changes in the independent variables can have large effects on the predicted values of the dependent variable.

DFFITS. The DFFITS statistic is a measure of the influence of a data point on regression prediction. It is the number of estimated standard errors the predicted value for a data point changes when the observed value is removed from the data set before computing the regression coefficients.

Predicted values that change by more than the specified number of standard errors when the data point is removed are flagged as influential; the suggested value is 2.0 standard errors.

Confidence Intervals

These results are displayed if you selected them in the Regression Options dialog box. If the confidence interval does not include zero, you can conclude that the coefficient is different than zero with the level of confidence specified.

This can also be described as $P < \alpha$ (alpha), where α is the acceptable probability of incorrectly concluding that the coefficient is different than zero, and the confidence interval is $100(1 - \alpha)$.

The specified confidence level can be any value from 1 to 99; the suggested confidence level for both intervals is 95%.

Row. This is the row number of the observation.

Predicted. This is the value for the dependent variable predicted by the regression model for each observation.

Regression. The confidence interval for the regression line gives the range of variable values computed for the region containing the true relationship between the dependent and independent variables, for the specified level of confidence.

Population. The confidence interval for the population gives the range of variable values computed for the region containing the population from which the observations were drawn, for the specified level of confidence.

Simple Linear Regression Report Graphs

You can generate up to five graphs using the results from a Simple Linear Regression. They include a:

- [Histogram of the residuals](#).
- [Scatter plot of the residuals](#).
- [Bar chart of the standardized residuals](#).
- [Normal probability plot of residuals](#).
- [Line/scatter plot of the regression with confidence and prediction intervals](#).

Creating a Linear Regression Report Graph

To generate a graph of Linear Regression report data:

1. With the report in view, click the **Report** tab.
2. In the **Result Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the **Linear Regression** results.

3. Select the type of graph you want to create from the **Graph Type** list, then click **OK**.

The specified graph appears in a graph window or in the report. For more information, see [Report Graphs](#) on page 373.

Multiple Linear Regression

Use a Multiple Linear Regression to when you want to:

- Predict the value of one variable from the values of two or more other variables, by fitting a plane (or hyperplane) through the data, and
- You know there are two or more independent variables and want to find a model with these independent variables.

The independent variables are the known, or predictor, variables. When the independent variables are varied, they produce a corresponding value for the dependent, or response, variable.

If you know there is only one independent variable, use Simple Linear Regression. If you are not sure if all independent variables should be used in the model, use Stepwise or Best Subsets Regression to identify the important independent variables from the selected possible independent variables.

If the relationship is not a straight line or plane, use Polynomial or Nonlinear Regression, or use a variable transformation.

About the Multiple Linear Regression

Multiple Linear Regression assumes an association between the dependent and k independent variables that fits the general equation for a multidimensional plane:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_k x_k$$

where y is the dependent variable, $x_1, x_2, x_3, \dots, x_k$ are the k independent variables, and $b_1, b_2, b_3, \dots, b_k$ are the k regression coefficients.

As the values x_1 vary, the corresponding value for y either increases or decreases, depending on the sign of the associated regression coefficient b_1 .

Multiple Linear Regression finds the $k+1$ dimensional plane that most closely describes the actual data, using all the independent variables selected.

Multiple Linear Regression is a *parametric* test, that is, for a given set of independent variable values, the possible values for the dependent variable are assumed to be normally distributed and have constant variance about the regression plane.

Performing a Multiple Linear Regression

To perform a Multiple Linear Regression:

1. Enter or arrange your data appropriately in the worksheet.
2. Generate report graphs.

Arranging Multiple Linear Regression Data

Place the data for the observed dependent variable in one column and the data for the corresponding independent variables in two or more columns.

Setting Multiple Linear Regression Options

Use the Multiple Linear Regression options to:

- Set assumption checking options.
- Specify the residuals to display and save them to the worksheet.
- Display confidence intervals and save them to the worksheet.
- Display the PRESS Prediction Error and standardized regression coefficients.
- Specify tests to identify outlying or influential data points.
- Set the variance inflation factor.
- Display power.

To change Multiple Linear Regression options:

1. If you are going to run the test after changing test options and want to select your data before you run the test, drag the pointer over the data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, select **Multiple Linear Regression** from the **Tests** drop-down list.
4. Click **Options**.

The **Options for Multiple Linear Regression** dialog box appears with four tabs:

- **Assumption Checking.** Click the **Assumption Checking** tab to view the Normality, Constant Variance, and Durbin-Watson options.
- **Residuals.** Click the **Residuals** tab to view the residual options.
- **More Statistics.** Click the **More Statistics** tab to view the confidence intervals, PRESS Prediction Error, Standardized Coefficients options.
- **Other Diagnostics.** Click **Other Diagnostics** to view the Influence, Variance Inflation Factor, and Power options.

5. Select or clear a check box to enable or disable a test option. Options settings are saved between SigmaPlot sessions. For more information, see [Interpreting Multiple Linear Regression Results](#) on page 244.
6. To continue the test, click **Run Test**.
7. To accept the current settings and close the options dialog box, click **OK**.

Options for Multiple Linear Regression: Assumption Checking

Click the **Assumption Checking** tab from the options dialog box to view the **Normality**, **Constant Variance**, and **Durbin-Watson** options. These options test your data for its suitability for regression analysis by checking three assumptions that a multiple linear regression makes about the data. A Multiple Linear Regression assumes

- That the source population is normally distributed about the regression.
- The variance of the dependent variable in the source population is constant regardless of the value of the independent variable(s).
- That the residuals are independent of each other.

All assumption checking options are selected by default. Only disable these options if you are certain that the data was sampled from normal populations with constant variance and that the residuals are independent of each other.

Normality Testing. SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

Constant Variance Testing. SigmaPlot tests for constant variance by computing the Spearman rank correlation between the absolute values of the residuals and the observed value of the dependent variable. When this correlation is significant, the constant variance assumption may be violated, and you should consider trying a different model (for example, one that more closely follows the pattern of the data), or transforming one or more of the independent variables to stabilize the variance.

P Values for Normality and Constant Variance. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and/or constant variance, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.05. Larger values of P (for example, 0.10) require less evidence to conclude that the residuals are not normally distributed or the constant variance assumption is violated.

To relax the requirement of normality and/or constant variance, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.01 for the normality test requires greater deviations from normality to flag the data as non-normal than a value of 0.05.

- Tip:** Although the assumption tests are robust in detecting data from populations that are non-normal or with non-constant variances, there are extreme conditions of data distribution that these tests cannot detect; however, these conditions should be easily detected by visually examining the data without resorting to the automatic assumption tests.

Durbin-Watson Statistic. SigmaPlot uses the Durbin-Watson statistic to test residuals for their independence of each other. The Durbin-Watson statistic is a measure of serial correlation between the residuals. The residuals are often correlated when the independent variable is time, and the deviation between the observation and the regression line at one time are related to the deviation at the previous time. If the residuals are not correlated, the Durbin-Watson statistic will be 2.

Difference from 2 Value. Enter the acceptable deviation from 2.0 that you consider as evidence of a serial correlation in the Difference for 2.0 box. If the computed Durbin-Watson statistic deviates from 2.0 more than the entered value, SigmaPlot warns you that the residuals may not be independent. The suggested deviation value is 0.50, for example, Durbin-Watson Statistic values greater than 2.5 or less than 1.5 flag the residuals as correlated.

To require a stricter adherence to independence, decrease the acceptable difference from 2.0.

To relax the requirement of independence, increase the acceptable difference from 2.0.

Options for Multiple Linear Regression: Residuals

Click the **Residuals** tab in the options dialog box to view the **Predicted Values**, **Raw**, **Standardized**, **Studentized**, **Studentized Deleted**, and **Report Flagged Values Only** options.

Predicted Values. Use this option to calculate the predicted value of the dependent variable for each observed value of the independent variable(s), then save the results to the data worksheet.

To assign predicted values to a worksheet column, select the worksheet column you want to save the predicted values to from the corresponding drop-down list. If you select none and the Predicted Values check box is selected, the values appear in the report but are not assigned to the worksheet.

Raw Residuals. The raw residuals are the differences between the predicted and observed values of the dependent variables. To include raw residuals in the report, make sure this check box is selected.

To assign the raw residuals to a worksheet column, select the number of the desired column from the corresponding drop-down list. If you select none from the drop-down list and the Raw check box is selected, the values appear in the report but are not assigned to the worksheet.

Standardized Residuals. The standardized residual is the residual divided by the standard error of the estimate. The standard error of the residuals is essentially the standard deviation of the residuals, and is a measure of variability around the regression line. To include standardized residuals in the report, make sure this check box is selected.

SigmaPlot automatically flags data points lying outside of the confidence interval specified in the corresponding box. These data points are considered to have "large" standardized residuals, for example, outlying data points. You can change which data points are flagged by editing the value in the Flag Values > edit box.

Studentized Residuals. Studentized residuals scale the standardized residuals by taking into account the greater precision of the regression line near the middle of the data versus the extremes. The Studentized residuals tend to be distributed according to the Student t distribution, so the t distribution can be used to define "large" values of the Studentized residuals. SigmaPlot automatically flags data points with "large" values of the Studentized residuals, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

To include Studentized residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include Studentized residuals in the worksheet.

Studentized Deleted Residuals. Studentized deleted residuals are similar to the Studentized residual, except that the residual values are obtained by computing the regression equation without using the data point in question.

To include Studentized deleted residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include Studentized deleted residuals in the worksheet.

SigmaPlot can automatically flag data points with "large" values of the Studentized deleted residual, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

 **Tip:** Both Studentized and Studentized deleted residuals use the same confidence interval setting to determine outlying points.

Report Flagged Values Only. To include only the flagged standardized and Studentized deleted residuals in the report, select Report Flagged Values Only. Clear this option to include all standardized and Studentized residuals in the report.

Options for Multiple Linear Regression: More Statistics

Click the **More Statistics** tab in the options dialog box to view the confidence interval options. You can set the confidence interval for the population, regression, or both and then save them to the data worksheet.

Confidence Interval for the Population. The confidence interval for the population gives the range of values that define the region that contains the population from which the observations were drawn.

To include confidence intervals for the population in the report, make sure the Population check box is selected. Click the selected check box if you do not want to include the confidence intervals for the population in the report.

Confidence Interval for the Regression. The confidence interval for the regression line gives the range of values that defines the region containing the true mean relationship between the dependent and independent variables, with the specified level of confidence.

To include confidence intervals for the regression in the report, make sure the Regression check box is selected, then specify a confidence level by entering a value in the percentage box. The confidence level can be any value from 1 to 99. The suggested confidence level for all intervals is 95%. Click the selected check box if you do not want to include the confidence intervals for the population in the report.

Click the selected check box if you do not want to include the confidence intervals for the population in the report.

Saving Confidence Intervals to the Worksheet. To save the confidence intervals to the worksheet, select the column number of the first column you want to save the intervals to from the Starting in Column drop-down list. The selected intervals are saved to the worksheet starting with the specified column and continuing with successive columns in the worksheet.

PRESS Prediction Error. The PRESS Prediction Error is a measure of how well the regression equation predicts the observations. Leave this check box selected to evaluate the fit of the equation using the PRESS statistic. Clear the selected check box if you do not want to include the PRESS statistic in the report.

Standardized Coefficients. These are the coefficients of the regression equation standardized to dimensionless

$$\pi_i = b_i \frac{s_x}{s_y}$$
 values, where b_i = regression coefficient, s_x = standard deviation of the independent variable x_i and s_y standard deviation of dependent variable y .

To include the standardized coefficients in the report, select **Standardized Coefficients**. Clear this option if you do not want to include the standardized coefficients in the worksheet.

Options for Multiple Linear Regression: Other Diagnostics

Click the **Other Diagnostics** tab in the **Options for Multiple Linear Regression** dialog box to view the Influence options. Influence options automatically detect instances of influential data points. Most influential points are data points which are outliers, that is, they do not do not "line up" with the rest of the data points. These points can have a potentially disproportionately strong influence on the calculation of the regression line. You can use several influence tests to identify and quantify influential points.

DFFITS. $DFFITS_i$ is the number of estimated standard errors that the predicted value changes for the i th data point when it is removed from the data set. It is another measure of the influence of a data point on the prediction used to compute the regression coefficients.

Predicted values that change by more than two standard errors when the data point is removed are considered to be influential.

Select **DFFITS** to compute this value for all points and flag influential points, for example, those with DFFITS greater than the value specified in the **Flag Values** > box. The suggested value is 2.0 standard errors, which indicates that the point has a strong influence on the data. To avoid flagging more influential points, increase this value; to flag less influential points, decrease this value. For more information, see [What to Do About Influential Points](#) on page 243.

Leverage. Select **Leverage** to identify the potential influence of a point on the results of the regression equation. Leverage depends only on the value of the independent variable(s). Observations with high leverage tend to be at the extremes of the independent variables, where small changes in the independent variables can have large effects on the predicted values of the dependent variable.

$$\frac{k+1}{n}$$

The expected leverage of a data point is $\frac{k+1}{n}$, where there are k independent variables and n data points. Observations with leverages much higher than the expected leverages are potentially influential points.

Select **Leverage** to compute the leverage for each point and automatically flag potentially influential points, for example, those points that could have leverages greater than the specified value times the expected leverage. The

$$\frac{2(k+1)}{n}$$

suggested value is 2.0 times the expected leverage for the regression (for example, $\frac{2(k+1)}{n}$). To avoid flagging more potentially influential points, increase this value; to flag points with less potential influence, lower this value. For more information, see [What to Do About Influential Points](#) on page 243.

Cook's Distance. Cook's distance is a measure of how great an effect each point has on the estimates of the parameters in the regression equation. Cook's distance assesses how much the values of the regression coefficients change if a point is deleted from the analysis. Cook's distance depends on both the values of the independent and dependent variables.

Select **Cook's Distance** to compute this value for all points and flag influential points, for example, those with a Cook's distance greater than the specified value. The suggested value is 4.0. Cook's distances above 1 indicate that a point is possibly influential. Cook's distances exceeding 4 indicate that the point has a major effect on the values of the parameter estimates. To avoid flagging more influential points, increase this value: to flag less influential points, lower this value. For more information, see [What to Do About Influential Points](#) on page 243.

Report Flagged Values Only. To only include only the influential points flagged by the influential point tests in the report, select **Report Flagged Values Only**. Clear this option to include all influential points in the report.

Power. The power of a regression is the power to detect the observed relationship in the data. The alpha (α) is the acceptable probability of incorrectly concluding there is a relationship.

Select **Power** to compute the power for the multiple linear regression data. Change the alpha value by editing the number in the **Alpha Value** edit box. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant relationship when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant relationship, but a greater possibility of concluding there is no relationship when one exists. Larger values of α make it easier to conclude that there is a relationship, but also increase the risk of reporting a false positive.

Variance Inflation Factor. Select **Variance Inflation Factor** to measure the multicollinearity of the independent variables, or the linear combination of the independent variables in the fit.

Regression procedures assume that the independent variables are statistically independent of each other, for example, that the value of one independent variable does not affect the value of another. However, this ideal situation rarely occurs in the real world. When the independent variables are correlated, or contain redundant information, the estimates of the parameters in the regression model can become unreliable.

The parameters in regression models quantify the theoretically unique contribution of each independent variable to predicting the dependent variable. When the independent variables are correlated, they contain some common information and "contaminate" the estimates of the parameters. If the multicollinearity is severe, the parameter estimates can become unreliable. For more information, see [What to Do About Multicollinearity](#) on page 243.

There are two types of multicollinearity:

- **Structural Multicollinearity.** Structural multicollinearity occurs when the regression equation contains several independent variables which are functions of each other. The most common form of structural multicollinearity occurs when a polynomial regression equation contains several powers of the independent variable. Because these powers (for example, x^2 , and so on) are correlated with each other, structural multicollinearity occurs. Including interaction terms in a regression equation can also result in structural multicollinearity.
- **Sample-Based Multicollinearity.** Sample-based multicollinearity occurs when the sample observations are collected in such a way that the independent variables are correlated (for example, if age, height, and weight are collected on children of varying ages, each variable has a correlation with the others).

SigmaPlot can automatically detect multicollinear independent variables using the variance inflation factor.

Flagging Multicollinear Data. Use the value in the **Flag Values** > edit box as a threshold for multicollinear variables. The default threshold value is 4.0, meaning that any value greater than 4.0 will be flagged as multicollinear. To make this test more sensitive to possible multicollinearity, decrease this value. To allow greater correlation of the independent variables before flagging the data as multicollinear, increase this value.

When the variance inflation factor is large, there are redundant variables in the regression model, and the parameter estimates may not be reliable. Variance inflation factor values above 4 suggest possible multicollinearity; values above 10 indicate serious multicollinearity.

Report Flagged Values Only. To only include only the points flagged by the influential point tests and values exceeding the variance inflation threshold in the report, select **Report Flagged Values**. Clear this option to include all influential points in the report.

What to Do About Influential Points

Influential points have two possible causes:

- There is something wrong with the data point, caused by an error in observation or data entry.
- The model is incorrect.

If a mistake was made in data collection or entry, correct the value. If you do not know the correct value, you may be able to justify deleting the data point. If the model appears to be incorrect, try regression with different independent variables, or a Nonlinear Regression.

What to Do About Multicollinearity

Sample-based multicollinearity can sometimes be resolved by collecting more data under other conditions to break up the correlation among the independent variables. If this is not possible, the regression equation is over parameterized and one or more of the independent variables must be dropped to eliminate the multicollinearity.

You can resolve structural multicollinearities by centering the independent variable before forming the power or interaction terms.

Running a Multiple Linear Regression

To run a Multiple Linear Regression, you need to select the data to test. Use the **Select Data** panel of the Test Wizard to select the worksheet columns with the data you want to test.

To run a Multiple Linear Regression:

1. If you want to select your data before you run the regression, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Regression > Multiple Linear

The **Select Data** panel of the Test Wizard appears. If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog box prompts you to pick your data.

4. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet or from the **Data for Dependent** or **Independent** drop-down list.

The first selected column is assigned to the **Dependent** row in the **Selected Columns** list, and all successively selected columns are assigned to the **Independent** rows in the list. The title of selected columns appear in each row. You can select up to 64 independent columns.

5. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
6. Click **Finish** to run to perform the regression. If you elected to test for normality, constant variance, and/or independent residuals, SigmaPlot performs the tests for normality (Shapiro-Wilk or Kolmogorov-Smirnov), constant variance, and independent residuals. If your data fails either of these tests, SigmaPlot warns you. When the test is complete, the report appears displaying the results of the Multiple Linear Regression.

If you selected to place residuals and other test results in the worksheet, they are placed in the specified column and are labeled by content and source column.

Interpreting Multiple Linear Regression Results

The report for a Multiple Linear Regression displays the equation with the computed coefficients, R, R^2 , and the adjusted R^2 , a table of statistical values for the estimate of the dependent variable, and the P value for the regression equation and for the individual coefficients.

The other results displayed in the report are enabled or disabled in the Options for Multiple Linear Regression dialog box.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

Regression Equation

This is the equation with the values of the coefficients in place. This equation takes the form:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_k x_k$$

where y is the dependent variable, $x_1, x_2, x_3, \dots, x_k$ are the independent variables, and b_1, b_2, b_3 are the regression coefficients.

The number of observations N , and the number of observations containing missing values (if any) that were omitted from the regression, are also displayed.

R, R Squared, and Adjusted R Squared

R and R^2 . R, the correlation coefficient, and R^2 , the coefficient of determination for multiple regression, are both measures of how well the regression model describes the data. R values near 1 indicate that the equation is a good description of the relation between the independent and dependent variables.

R equals 0 when the values of the independent variable do not allow any prediction of the dependent variables, and equals 1 when you can perfectly predict the dependent variables from the independent variables.

Adjusted R Squared. The adjusted R^2 , R^2_{adj} , is also a measure of how well the regression model describes the data, but takes into account the number of independent variables, which reflects the degrees of freedom. Larger R^2_{adj} values (nearer to 1) indicate that the equation is a good description of the relation between the independent and dependent variables.

Standard Error of the Estimate

The standard error of the estimate s_{yx} is a measure of the actual variability about the regression plane of the underlying population. The underlying population generally falls within about two standard errors of the estimate of the observed sample.

Statistical Summary Table

Coefficients. The value for the constant and coefficients of the independent variables for the regression model are listed.

Standard Error. The standard errors of the regression coefficients (analogous to the standard error of the mean). The true regression coefficients of the underlying population generally fall within about two standard errors of the observed sample coefficients. Large standard errors may indicate multicollinearity.

These values are used to compute t and confidence intervals for the regression.

Beta

$$\beta_i = b_i \frac{s_x}{s_y}$$

These are the coefficients of the regression equation standardized to dimensionless values, where b_i = regression coefficient, s_x = standard deviation of the independent variable x_i , and s_y = standard deviation of dependent variable y .

These results are displayed if the Standardized Coefficients option was selected in the Regression Options dialog box.

t Statistic. The t statistic tests the null hypothesis that the coefficient of the independent variable is zero, that is, the independent variable does not contribute to predicting the dependent variable. t is the ratio of the regression coefficient to its standard error, or:

$$t = \frac{\text{regression coefficient}}{\text{standard error of regression coefficient}}$$

You can conclude from "large" t values that the independent variable can be used to predict the dependent variable (for example, that the coefficient is not zero).

P value. P is the P value calculated for t . The P value is the probability of being wrong in concluding that there is a true association between the variables (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on t). The smaller the P value, the greater the probability that the variables are correlated.

Traditionally, you can conclude that the independent variable contributes to predicting the dependent variable when $P < 0.05$.

VIF (Variance Inflation Factor). The variance inflation factor is a measure of multicollinearity. It measures the "inflation" of the standard error of each regression parameter (coefficient) for an independent variable due to redundant information in other independent variables.

If the variance inflation factor is 1.0, there is no redundant information in the other independent variables. If the variance inflation factor is much larger, there are redundant variables in the regression model, and the parameter estimates may not be reliable.

Variance inflation factor values for independent variables above the specified value are flagged with a $>$ symbol, indicating multicollinearity with other independent variables. The suggested value is 4.0.

Analysis of Variance (ANOVA) Table

The ANOVA (analysis of variance) table lists the ANOVA statistics for the regression and the corresponding F value.

SS (Sum of Squares). The sum of squares are measures of variability of the dependent variable.

- The sum of squares due to regression measures the difference of the regression plane from the mean of the dependent variable.
- The residual sum of squares is a measure of the size of the residuals, which are the differences between the observed values of the dependent variable and the values predicted by regression model.
- The total sum of squares is a measure of the overall variability of the dependent variable about its mean value.

DF (Degrees of Freedom). Degrees of freedom represent the number observations and variables in the regression equation.

- The regression degrees of freedom is a measure of the number of independent variables.
- The residual degrees of freedom is a measure of the number of observations less the number of terms in the equation.
- The total degrees of freedom is a measure of total observations.

MS (Mean Square). The mean square provides two estimates of the population variances. Comparing these variance estimates is the basis of analysis of variance.

The mean square regression is a measure of the variation of the regression from the mean of the dependent variable, or:

$$\frac{\text{sum of squares due to regression}}{\text{regression degrees of freedom.}} = \frac{SS_{\text{reg}}}{DF_{\text{reg}}} = MS_{\text{reg}}$$

The residual mean square is a measure of the variation of the residuals about the regression plane, or:

$$\frac{\text{residual sum of squares}}{\text{residual degrees of freedom.}} = \frac{SS_{\text{res}}}{DF_{\text{res}}} = MS_{\text{res}}$$

The residual mean square is also equal to

$$\frac{\text{regression variation from the dependent variable mean}}{\text{residual variation about the regression curve}} = \frac{MS_{\text{reg}}}{MS_{\text{res}}} = F \text{ overall}$$

F Statistic. The F test statistic gauges the ability of the regression equation, containing all independent variables, to predict the dependent variable. It is the ratio

$$\frac{\text{regression variation from the dependent variable mean}}{\text{residual variation about the regression line}} = \frac{MS_{\text{reg}}}{MS_{\text{res}}} = F$$

If F is a large number, you can conclude that the independent variables contribute to the prediction of the dependent variable (for example, at least one of the coefficients is different from zero, and the "unexplained variability" is smaller than what is expected from random sampling variability about the mean value of the dependent variable). If the F ratio is around 1, you can conclude that there is no association between the variables (for example, the data is consistent with the null hypothesis that all the samples are just randomly distributed).

P Value. The P value is the probability of being wrong in concluding that there is an association between the dependent and independent variables (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on F). The smaller the P value, the greater the probability that there is an association.

Traditionally, you can conclude that the independent variable can be used to predict the dependent variable when $P < 0.05$.

Incremental Sum of Squares

SSincr. SSincr, the incremental or Type I sum of squares, is a measure of the new predictive information contained in an independent variable, as it is added to the equation.

The incremental sum of squares measures the increase in the regression sum of squares (and reduction in the sum of squared residuals) obtained when that independent variable is added to the regression equation, after all independent variables above it have been entered.

You can gauge the additional contribution of each independent variable by comparing these values.

SSmarg. SSmarg, the marginal or Type III sum of squares, is a measure of the unique predictive information contained in an independent variable, after taking into account all other independent variables. You can gauge the independent contribution of each independent variable by comparing these values.

The marginal sum of squares measures the reduction in the sum of squared residuals obtained by entering the independent variable last, after all other variables in the equation have been entered.

PRESS Statistic

PRESS, the Predicted Residual Error Sum of Squares, is a measure of how well a regression model predicts the observations.

The PRESS statistic is computed by summing the squares of the prediction errors (the differences between predicted and observed values) for each observation, with that point deleted from the computation of the estimated regression model.

One important use of the PRESS statistics is for model comparison. If several different regression models are applied to the same data, the one with the smallest PRESS statistic has the best predictive capability.

Durbin-Watson Statistic

The Durbin-Watson statistic is a measure of correlation between the residuals. If the residuals are not correlated, the Durbin-Watson statistic will be 2; the more this value differs from 2, the greater the likelihood that the residuals are correlated. This results appears if it was selected in the Regression Options dialog box.

Regression assumes that the residuals are independent of each other; the Durbin-Watson test is used to check this assumption. If the Durbin-Watson value deviates from 2 by more than the value set in the Regression Options dialog box, a warning appears in the report. The suggested trigger value is a difference of more than 0.50, for example, the Durbin-Watson statistic is below 1.50 or above 2.50.

Normality Test

Normality test result displays whether the data passed or failed the test of the assumption that the source population is normally distributed around the regression, and the P value calculated by the test. All regressions require a source population to be normally distributed about the regression line. When this assumption may be violated, a warning appears in the report. This result appears unless you disabled normality testing in the Regression Options dialog box.

Failure of the normality test can indicate the presence of outlying influential points or an incorrect regression model.

Constant Variance Test

The constant variance test result displays whether or not the data passed or failed the test of the assumption that the variance of the dependent variable in the source population is constant regardless of the value of the independent variable, and the P value calculated by the test. When the constant variance assumption may be violated, a warning appears in the report.

If you receive this warning, you should consider trying a different model (for example, one that more closely follows the pattern of the data), or transforming the independent variable to stabilize the variance and obtain more accurate estimates of the parameters in the regression equation.

Power

This result is displayed if you selected this option in the Options for Multiple Linear Regression dialog box.

The power, or sensitivity, of a regression is the probability that the regression model can detect the observed relationship among the variables, if there is a relationship in the underlying population.

Regression power is affected by the number of observations, the chance of erroneously reporting a difference a (alpha), and the slope of the regression.

Alpha (α). Alpha (α) is the acceptable probability of incorrectly concluding that the model is correct. An a error is also called a Type I error (a Type I error is when you reject the hypothesis of no association when this hypothesis is true).

Set the value in the Power Options dialog box; the suggested value is $\alpha = 0.05$ which indicates that a one in twenty chance of error is acceptable. Smaller values of α result in stricter requirements before concluding the model is correct, but a greater possibility of concluding the model is bad when it is really correct (a Type II error). Larger values of α make it easier to conclude that the model is correct, but also increase the risk of accepting a bad model (a Type I error).

Regression Diagnostics

The regression diagnostic results display only the values for the predicted values, residuals, and other diagnostic results selected in the Options for Multiple Linear Regression dialog box. All results that qualify as outlying values are flagged with a < symbol. The trigger values to flag residuals as outliers are set in the Options for Multiple Linear Regression dialog box.

If you selected Report Cases with Outliers Only, only those observations that have one or more residuals flagged as outliers are reported; however, all other results for that observation are also displayed.

Row. This is the row number of the observation.

Predicted Values. This is the value for the dependent variable predicted by the regression model for each observation.

Residuals. These are the raw residuals, the difference between the predicted and observed values for the dependent variables.

Standardized Residuals. The standardized residual is the raw residual divided by the standard error of the estimate s_{yx} .

If the residuals are normally distributed about the regression, about 66% of the standardized residuals have values between -1 and +1, and about 95% of the standardized residuals have values between -2 and +2. A larger standardized residual indicates that the point is far from the regression; the suggested value flagged as an outlier is 2.5.

Studentized Residuals. The Studentized residual is a standardized residual that also takes into account the greater confidence of the predicted values of the dependent variable in the "middle" of the data set. By weighting the values of the residuals of the extreme data points (those with the lowest and highest independent variable values), the Studentized residual is more sensitive than the standardized residual in detecting outliers.

Both Studentized and Studentized deleted residuals that lie outside a specified confidence interval for the regression are flagged as outlying points; the suggested confidence value is 95%.

This residual is also known as the internally Studentized residual, because the standard error of the estimate is computed using all data.

Studentized Deleted Residual. The Studentized deleted residual, or externally Studentized residual, is a Studentized residual which uses the standard error of the estimate, computed after deleting the data point associated with the residual. This reflects the greater effect of outlying points by deleting the data point from the variance computation.

Both Studentized and Studentized deleted residuals that lie outside a specified confidence interval for the regression are flagged as outlying points; the suggested confidence value is 95%.

The Studentized deleted residual is more sensitive than the Studentized residual in detecting outliers, since the Studentized deleted residual results in much larger values for outliers than the Studentized residual.

Influence Diagnostics

The influence diagnostic results display only the values for the results selected in the Options dialog box under the Other Diagnostics tab. All results that qualify as outlying values are flagged with a < symbol. The trigger values to flag data points as outliers are also set in Options dialog box under the Other Diagnostics tab.

If you selected Report Cases with Outliers Only, only observations that have one or more observations flagged as outliers are reported; however, all other results for that observation are also displayed.

Row. This is the row number of the observation.

Cook's Distance. Cook's distance is a measure of how great an effect each point has on the estimates of the parameters in the regression equation. It is a measure of how much the values of the regression coefficients would change if that point is deleted from the analysis.

Values above 1 indicate that a point is possibly influential. Cook's distances exceeding 4 indicate that the point has a major effect on the values of the parameter estimates. Points with Cook's distances greater than the specified value are flagged as influential; the suggested value is 4.

Leverage. Leverage values identify potentially influential points. Observations with leverages a specified factor greater than the expected leverages are flagged as potentially influential points; the suggested value is 2.0 times the expected leverage.

$$\frac{k+1}{n}$$

The expected leverage of a data point is $\frac{k+1}{n}$ where there are k independent variables and n data points.

Because leverage is calculated using only the dependent variable, high leverage points tend to be at the extremes of the independent variables (large and small values), where small changes in the independent variables can have large effects on the predicted values of the dependent variable.

DFFITS. The DFFITS statistic is a measure of the influence of a data point on regression prediction. It is the number of estimated standard errors the predicted value for a data point changes when the observed value is removed from the data set before computing the regression coefficients.

Predicted values that change by more than the specified number of standard errors when the data point is removed are flagged as influential; the suggested value is 2.0 standard errors.

Confidence Intervals

These results are displayed if you selected them in the Options for Multiple Linear Regression dialog box. If the confidence interval does not include zero, you can conclude that the coefficient is different than zero, concluding that the

The specified confidence level can be any value from 1 to 99; the suggested confidence level for both intervals is 95%.

Row. This is the row number of the observation.

Predicted. This is the value for the dependent variable predicted by the regression model for each observation.

Regression. The confidence interval for the regression gives the range of variable values computed for the region containing the true relationship between the dependent and independent variables, for the specified level of confidence.

Population. The confidence interval for the population gives the range of variable values computed for the region containing the population from which the observations were drawn, for the specified level of confidence.

Multiple Linear Regression Report Graphs

You can generate up to six graphs using the results from a Multiple Linear Regression. They include a:

- [Histogram of the residuals.](#)
- [Scatter plot of the residuals.](#)
- [Bar chart of the standardized residuals.](#)
- [Normal probability plot of the residuals.](#)
- [Line/scatter plot of the regression variable and confidence and prediction intervals with one independent.](#)
- [3D scatter plot of the residuals.](#)

Creating Multiple Linear Regression Report Graphs

To generate a report graph of Multiple Linear Regression data:

1. With the Multiple Linear Regression report in view, click the **Report** tab.
2. In the **Graph Results** group, click **Create Result Graph**.
3. The **Create Result Graph** dialog box appears displaying the types of graphs available for the Multiple Linear Regression results.
4. Select the type of graph you want to create from the **Graph Type** list, then click **OK**, or double-click the desired graph in the list.

If you select **Scatter Plot Residuals**, **Bar Chart Std Residuals**, **Regression, Conf. & Pred**, a dialog box appears prompting you to select the column with independent variables you want to use in the graph.

If you select **3D Scatter & Mesh**, or **3D Residual Scatter**, and you have more than two columns of independent variables, a dialog box appears prompting you to select the two columns with the independent variables you want to plot.

5. Select the columns with the independent variables you want to use in the graph, then click **OK**. The graph appears using the specified independent variables. For more information, see [Report Graphs](#) on page 373.

Multiple Logistic Regression

Use a Multiple Logistic Regression when you want to predict a qualitative dependent variable, such as the presence or absence of a disease, from observations of one or more independent variables, by fitting a logistic function to the data.

The independent variables are the known, or predictor, variables. When the independent variables are varied, they produce a corresponding value for the dependent, or response, variable. SigmaPlot's Logistic Regression requires that the dependent variable be dichotomous or take two possible responses (dead or alive, black or white) represented by values of 0 and 1.

If your dependent variable data does not use dichotomous values, use a Simple Linear Regression if you have one independent variable and a Multiple Linear Regression if you have more than one independent variable.

About the Multiple Logistic Regression

Multiple Logistic Regression assumes an association between the dependent and k independent variables that fits the general equation for a multidimensional plane:

$$P(y=1) = \frac{1}{1+e^{-(\beta_0+\beta_1x_1+\beta_2x_2+\dots+\beta_kx_k)}}$$

where y is the dependent variable, $P(y=1)$ is the predicted probability that the dependent variable is positive response or has a value of 1, β_0 through β_k are the $k+1$ regression coefficients, and x_1 through x_k are the independent variables.

As the values x_i vary, the corresponding estimated probability that $y=1$ increases or decreases, depending on the sign of the associated regression coefficient β_i .

Multiple Logistic Regression finds the set of values of the regression coefficients most likely to predict the observed values of the dependent variable, given the observed values of the independent variables.

Performing a Multiple Logistic Regression

To perform a Multiple Logistic Regression:

1. [Enter or arrange your data appropriately in the worksheet.](#)
2. Set the Logistic Regression options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Regression > Multiple Logistic
5. [View and interpret the Multiple Logistic Regression report.](#)
6. Run the test.

Arranging Multiple Logistic Regression Data

Place the data for the observed dependent variable in one column and the data for the corresponding independent variables in one or more columns.

Logistic Regression data is entered into the worksheet in a raw data format in which the data for the observed dependent variable is in one column and the data for the corresponding independent variables are in one or more other columns. You must also enter dependent variable data as dichotomous data and independent variable data must be entered in numerical format.

If you have continuous numerical data or as text as your dependent variable data, or if you are using categorical independent variables, you must convert them into an equivalent set of dummy variables using reference coding.

Observations containing missing values are ignored, and all columns must be equal in length.

Setting Multiple Logistic Regression Options

Use the Multiple Logistic Regression options to:

- Set options used to determine how well the logistics regression equation fits the data.
- Estimate the variance inflation factors for the regression coefficients.
- Specify the residuals to display and save them to the worksheet.
- Calculate the standard error coefficient, Wald statistic, odds ratio, odds ratio confidence, and coefficients P value.
- Specify tests to identify outlying or influential data points.

To change Multiple Logistic Regression options:

1. If you are going to run the test after changing test options and want to select your data before you run the test, drag the pointer over the data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, select **Multiple Logistic Regression** from the **Select Test** drop-down list.

4. Click **Options**. The **Options for Multiple Logistic Regression** dialog box appears with three tabs:

- **Criterion.** Click the **Criterion** tab to view the criterion options.
- **More Statistics.** Click the **More Statistics** tab to view the Standard Error Coefficients, Wald Statistic, Odds Ratio, Odds Ratio Confidence, and Coefficients P Values, Predicted Values, and Variance Inflation Factor options.
- **Residuals.** Click the **Residuals** tab to view the residual and influence options.

Option settings are saved between SigmaPlot sessions.

5. To continue the test, click **Run Test**.

6. To accept the current settings and close the options dialog box, click **OK**.

Options for Multiple Logistic Regression: Criterion

Click the **Criterion** tab in the **Options for Multiple Logistic Regression** dialog box to set the criterion options. Use these options to specify the criterion you want to use to test how well your data fits the logistic regression equation.

Hosmer-Lemshow Test Statistic. The Hosmer-Lemshow statistic tests the null hypothesis that the logistic equation fits the data by comparing the number of individuals with each outcome with the number expected based on the logistic equation.

Threshold probability for goodness of fit. Small P values indicate that you can reject the null hypothesis that the logistic equation fits the data and try should try an equation with different independent variables. Large P values indicate a good fit between the logistic equation and the data. The default value is 0.2. Setting the P value to larger values requires smaller deviations between the values predicted by the logistic equation and the observed values of the dependent variable to accept the equation as a good fit to the data. To change the P value, type a new value in the edit box.

Pearson Chi-Square Statistic. The Pearson Chi-Square statistic tests how well the logistic regression equation fits your data by summing the squares of the Pearson residuals. Small values of the Pearson Chi-Square statistic indicate a good agreement between the logistic regression equation and the data. Large values of the Pearson Chi-Square indicate a poor agreement.

Likelihood Ratio Test Statistic. The Likelihood Ratio Test statistic tests how well the logistic regression equation fits your data by summing the squares of the deviance residuals. It compares the your full model against a model that uses nothing but the mean of the dependent variable. Small P values indicate a good fit between the logistic regression equation and your data.

Classification Table. The classification table tests the null hypothesis that the data follow the logistic equation by comparing the number of individuals with each outcome with the number expected based on the logistic equation. It summarizes the results of whether the data fits the logistic equation by cross-classifying the actual dependent response variables with predicted responses and identifying the number of different combinations of the independent variables.

Threshold probability for positive classification. The predicted responses are assigned dichotomous variables derived by comparing estimated logistic probabilities to the probability value specified in the **Threshold probability for positive classification** edit box.

If the estimated probability exceeds the specified probability value, the predicted variable is assigned a positive response (value of 1); probabilities less than or equal to the specified value are assigned a value of 0 or a reference value. The default threshold is 0.5. The resulting contingency table can be analyzed with a Chi-Square test. As with the Hosmer-Lemshow statistic, a large P value indicates a good fit between the logistic regression equation and the data. For more information, see [Interpreting Multiple Logistic Regression Results](#) on page 255.

Number of Independent Variable Combinations. If the number of unique combinations of the independent variables is not large compared to the number of independent variables, your logistic regression results may be unreliable. To calculate the number of independent variable combinations and warn if there are not enough combinations as compared to the independent variables, select the Number of Independent Variable Combinations check box. If the calculated independent combination is less than the value in the corresponding edit box, a dialog box appears warning you that the number of independent variable combinations are too small, and asks if you want to continue. If you select Yes, the warning message appears in the report.

Options for Multiple Logistic Regression: Statistics

Click the **More Statistics** tab in the **Options** dialog box to view the statistics options. These options help determine how well your data fits the logistic regression equation using maximum likelihood as the estimation criterion.

Standard Error Coefficients. The Standard Error Coefficients are measures of the precision of the estimates of the regression coefficients. The true regression coefficients of the underlying population generally fall within two standard errors of the observed sample coefficients.

Wald Statistic. The Wald Statistic compares the observed value of the estimated coefficient with its associated standard error. It is computed as the ratio:

$$z = \frac{b_i^2}{s_{b_i}^2}$$

where b_i is the observed value of the estimated coefficient, and s_{b_i} is the standard error of the coefficient.

Select **Wald Statistic** to include the ratio of the observed coefficient with the associated standard error in the report. The Wald statistic can also be used to determine how significant the independent variables are in predicting the dependent variable.

Odds Ratio. The odds of any event occurring can be defined by

$$Odds = \Omega = \frac{P}{1-p}$$

where P is the probability of the event happening. The odds ratio for an independent variable is computed as

$\Omega_{\beta_1} = e^{\beta_1}$ where β_1 is the regression coefficient. The odds ratio is an estimate of the increase (or decrease) in the odds for an outcome if the independent variable value is increased by 1.

Odds Ratio Confidence. The odds ratio confidence intervals are defined as

$$\Omega = e^{(b_i \pm z_{1-\frac{\alpha}{2}} s_{b_i})}$$

where b_i is the coefficient, s_{b_i} is the standard error of the coefficient, and $z_{1-\frac{\alpha}{2}}$ is the point on the axis of the standard normal distribution that corresponds to the desired confidence interval.

The default confidence used is 95%. To change the confidence used, change the percentage in the corresponding edit box.

Coefficients P Value. The **Coefficients P Value** determines the probability of being incorrect in concluding that the each independent variable has a significant effect on determining the dependent variable. The smaller the P value, the more likely the independent variables actually predicts the dependent variables.

Use the **Wald Statistic** to test whether the coefficients associated with the independent variables are significantly different from zero. The significance of independent variables is tested by comparing the observed value of the coefficients with the associated standard error of the coefficient. If the observed value of the coefficient is large compared to the standard error, you can conclude that the coefficients are significantly different from zero and that the independent variables contribute significantly to predicting the dependent variables. For more information, see [Interpreting Multiple Logistic Regression Results](#) on page 255.

Predicted Values. Use this option to calculate the predicted value of the dependent variable for each observed value of the independent variable(s), then save the results to the data worksheet.

For logistic regression the predicted values indicate the probability of a positive response. For more information, see [Interpreting Multiple Logistic Regression Results](#) on page 255.

To assign predicted values to a worksheet column, select the worksheet column you want to save the predicted values to from the corresponding drop-down list. If you select none and the Predicted Values check box is selected, the values appear in the report but are not assigned to the worksheet.

Variance Inflation Factor. Use this option to measure the multicollinearity of the independent variables, or the linear combination of the independent variables in the fit.

Regression procedures assume that the independent variables are statistically independent of each other, for example, that the value of one independent variable does not affect the value of another. However, this ideal situation rarely occurs in the real world. When the independent variables are correlated, or contain redundant information, the estimates of the parameters in the regression model can become unreliable.

The parameters in regression models quantify the theoretically unique contribution of each independent variable to predicting the dependent variable. When the independent variables are correlated, they contain some common information and "contaminate" the estimates of the parameters. If the multicollinearity is severe, the parameter estimates can become unreliable.

There are two types of multicollinearity.

- **Sample-Based Multicollinearity.** Sample-based multicollinearity occurs when the sample observations are collected in such a way that the independent variables are correlated (for example, if age, height, and weight are collected on children of varying ages, each variable has a correlation with the others). This is the most common form of multicollinearity.
- **Structural Multicollinearity.** Structural multicollinearity occurs when the regression equation contains several independent variables which are functions of each other. An example of this is when a regression equation contains several powers of the independent variable. Because these powers (for example, x , x^2) are correlated with each other, structural multicollinearity occurs. Including interaction terms in a regression equation can also result in structural multicollinearity.

Flag values >. Use the value in the **Flag Values >** edit box as a threshold for multicollinear variables. The default threshold value is 4.0, meaning that any value greater than 4.0 will be flagged as multicollinear. To make this test more sensitive to possible multicollinearity, decrease this value. To allow greater correlation of the independent variables before flagging the data as multicollinear, increase this value. For more information, see [What to Do About Multicollinearity](#) on page 253.

When the variance inflation factor is large, there are redundant variables in the regression model, and the parameter estimates may not be reliable. Variance inflation factor values above 4 suggest possible multicollinearity; values above 10 indicate serious multicollinearity.

Report Flagged Values Only. To only include only the points flagged by the influential point tests and values exceeding the variance inflation threshold in the report, select **Report Flagged Values Only**. Clear this option to include all influential points in the report.

What to Do About Multicollinearity

You can sometimes resolve sample-based multicollinearity by collecting more data under other conditions to break up the correlation among the independent variables. If this is not possible, the regression equation is over parameterized and one or more of the independent variables must be dropped to eliminate the multicollinearity.

You can resolve structural multicollinearities by centering the independent variable before forming the power or interaction terms.

Options for Multiple Logistic Regression: Residuals

Click the **Residuals** tab in the options dialog box to view the Residual Type, Raw, Standardized, Studentized, Studentized Deleted, and Report Flagged Values Only options.

Residual Type. Residuals are not reported by default. To include residuals in the report select either Pearson or Deviance from the Residual Type drop-down list. Select None from the drop-down list if you don't want to include residuals in the report.

Deviance residuals are used to calculate the likelihood ratio test statistic to assess the overall goodness of fit of the logistic regression equation to the data. The likelihood ratio test statistic is the sum of squared deviance residuals. The deviance residual for each point is a measure of how much that point contributes to the likelihood ratio test statistic. Larger values of the deviance residual indicate a larger difference between the observed and predicted values of the dependent variable.

Pearson residuals are calculated by dividing the raw residual by the standard error. The standard error is defined as the observed value of the dependent variable (0 or 1) divided by the probability of a positive response (for example $y = 1$) outcome that is estimated from the Logistic Regression equation. Pearson residuals are the default residual type used

to calculate the goodness of fit for the logistic regression equation because the Chi-Square goodness of fit statistic is the sum of squared Pearson residuals.

Raw Residuals. The raw residuals are the differences between the predicted and observed values of the dependent variables. To include raw residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include raw residuals in the worksheet.

To assign the raw residuals to a worksheet column, select the number of the desired column from the corresponding drop-down list. If you select none from the drop-down list and the Raw check box is selected, the values appear in the report but are not assigned to the worksheet.

Studentized Residuals. Studentized residuals take into account the greater precision of the regression estimates near the middle of the data versus the extremes. The Studentized residuals tend to be distributed according to the Student t distribution, so the t distribution can be used to define "large" values of the Studentized residuals. SigmaPlot automatically flags data points with "large" values of the Studentized residuals, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

To include Studentized residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include Studentized residuals in the worksheet.

Studentized Deleted Residuals. Studentized deleted residuals are similar to the Studentized residual, except that the residual values are obtained by computing the regression equation without using the data point in question.

To include Studentized deleted residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include Studentized deleted residuals in the worksheet.

SigmaPlot can automatically flag data points with "large" values of the Studentized deleted residual, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

 **Note:** Both Studentized and Studentized deleted residuals use the same confidence interval setting to determine outlying points.

Report Flagged Values Only. To only include the flagged standardized and Studentized deleted residuals in the report, select **Report Flagged Values Only**. Clear this option to include all standardized and Studentized residuals in the report.

Influence

Influence options automatically detect instances of influential data points. Most influential points are data points which are outliers, that is, they do not "line up" with the rest of the data points. These points can have a potentially disproportionately strong influence on the calculation of the regression line. You can use several influence tests to identify and quantify influential points.

Leverage. Leverage is used to identify the potential influence of a point on the results of the regression equation. Leverage depends only on the value of the independent variable(s). Observations with high leverage tend to be at the extremes of the independent variables, where small changes in the independent variables can have large effects on the predicted values of the dependent variable.

$$\frac{k+1}{n}$$

The expected leverage of a data point is $\frac{k+1}{n}$, where there are k independent variables and n data points. Observations with leverages much higher than the expected leverages are potentially influential points.

Select Leverage to compute the leverage for each point and automatically flag potentially influential points, for example, those points that could have leverages greater than the specified value times the expected leverage. The

$$\frac{2(k+1)}{n}$$

suggested value is 2.0 times the expected leverage for the regression $\frac{2(k+1)}{n}$. To avoid flagging more potentially influential points, increase this value; to flag points with less potential influence, lower this value.

Cook's Distance. Cook's distance is a measure of how great an effect each point has on the estimates of the parameters in the regression equation. Cook's distance assesses how much the values of the regression coefficients

change if a point is deleted from the analysis. Cook's distance depends on both the values of the independent and dependent variables.

Select Cook's Distance to compute this value for all points and flag influential points, for example, those with a Cook's distance greater than the specified value. The suggested value is 4.0. Cook's distances above 1 indicate that a point is possibly influential. Cook's distances exceeding 4 indicate that the point has a major effect on the values of the parameter estimates. To avoid flagging more influential points, increase this value: to flag less influential points, lower this value. For more information, see [What to Do About Influential Points](#) on page 243.

Influential Points

Influential points have two possible causes:

- There is something wrong with the data point, caused by an error in observation or data entry.
- The model is incorrect.

If a mistake was made in data collection or entry, correct the value. If you do not know the correct value, you may be able to justify deleting the data point. If the model appears to be incorrect, try regression with different independent variables, or a Nonlinear Regression.

Running a Multiple Logistic Regression

To run a Multiple Logistic Regression, you need to select the data to test. Use the **Select Data** panel of the Test Wizard select the worksheet columns with the data you want to test.

To run a Multiple Logistic Regression:

1. If you want to select your data before you run the regression, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Regression > Multiple Logistic

The **Select Data** panel of the Test Wizard appears. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

4. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Dependent, Independent, or Count** drop-down list.

Select the column with the values indicating the number of time a dependent and independent combination is repeats as the Count column. The title of selected columns appears in each row.

5. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
6. Click **Finish** to run the regression. If you elected to test for normality, constant variance, and/or independent residuals, SigmaPlot performs the tests for normality (Shapiro-Wilk or Kolmogorov-Smirnov), constant variance, and independent residuals. If your data fails either of these tests, SigmaPlot warns you. When the test is complete, the report appears displaying the results of the Multiple Logistic Regression.

If you selected to place residuals and other test results in the worksheet, they are placed in the specified column and are labeled by content and source column.

Interpreting Multiple Logistic Regression Results

The report for a Multiple Logistic Regression displays the logistic equation with the computed coefficients, their standard errors, the number of observations in the test, estimation criterion used to fit the logistic equation to your data, the worksheet column with the dependent variable data, the values representing the positive and reference responses, and the Hosmer-Lemeshow and Chi Square goodness of fit statistics.

The other results displayed in the report are enabled or disabled in the Options for Multiple Logistic Regression dialog box.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the **Options** dialog box. You can also set the number of decimal places to display in the **Options** dialog box. For more information, see [Report Graphs](#) on page 373.

Regression Equation

The logistic regression equation is:

$$P = \frac{1}{1+e^{-(b_0+b_1x_1+b_2x_2+\dots+b_kx_k)}}$$

where P is the probability of a "positive" response (for example, value of the dependent variable equal to 1) and $x_1, x_2, x_3, \dots, x_k$ are the independent variables and $b_1, b_2, b_3, \dots, b_k$ are the regression coefficients. The equation can be rewritten by applying the logit transformation to both sides of this equation.

$$\text{Logit } P = \ln\left(\frac{P}{1-P}\right)$$

Number of Observations

The number of observations N , and the number of observations containing missing values (if any) that were omitted from the regression, are also displayed.

Estimation Criterion

Logistic regression uses the maximum likelihood approach to find the values of the coefficients (b_i) in the Logistic Regression Equation that were most likely to fit the observed data.

 **Note:** The regression coefficients computed by minimizing the sum of squared residuals in Multiple Logistic Regression are also the maximum likelihood estimates.

Dependent Variable

This section of the report indicates which values in the dependent variable column represent the positive response (1) and which value represents the reference response (0).

Number of Unique Independent Variable Combinations

This value represents the number of unique combinations of the independent variables and appears if you have the Number of Independent Variable Combinations option in the Options for Logistic Regression dialog box selected. The number of unique independent variable combinations is compared to the actual number of independent variables. If this value is less than the value specified for the Number of Independent Variable Combinations option, a warning message appears in the report that your results may be unreliable.

Hosmer-Lemshow P Value

The Hosmer-Lemshow P value indicates how well the logistic regression equation fits your data by comparing the number of individuals with each outcome with the number expected based on the logistic equation. It tests the null hypothesis that the logistic equation describes the data. Thus, small P values indicate a poor fit of the equation to your data (for example, you reject the null hypothesis of agreement). Large P values indicate a good fit between the logistic equation and the data. The critical Hosmer-Lemshow P value option is set in the Options for Multiple Logistic Regression dialog box.

When the dataset is small, goodness of fit measures for the logistic regression should be interpreted with great caution. All of the P values are based on a chi-square probability distribution, which is not recommended for use with small numbers of observations.

Pearson Chi-Square Statistic

The Pearson Chi-Square statistic is the sum of the squared Pearson residuals. It is a measure of the agreement between the observed and predicted values of the dependent variable using a Chi-Square test statistic. The Chi-Square test statistic is analogous to the residual sum of squares in ordinary linear regression. Small values of the Chi-Square

(and corresponding large values of the associated P value) indicate a good agreement between the logistic regression equation and the data and large values of Chi-Square (and small values of P) indicate a poor agreement. The Pearson Chi-Square option is set in the Options for Multiple Logistic Regression dialog box.

Likelihood Ratio Test Statistic

The Likelihood Ratio Test statistic is derived from the sum of the squared deviance residuals. It indicates how well the logistic regression equation fits your data by comparing the likelihood of obtaining observations if the independent variables had no effect on the dependent variable with the likelihood of obtaining the observations if the independent variables had an effect on the dependent variables.

This comparison is computed by running the logistic regression with and without the independent variables and comparing the results. If the pattern of observed outcomes is more likely to have occurred when independent variables affect the outcome than when they do not, a small coefficients of P value is reported, indicating a good fit between the logistic regression equation and your data.

Log Likelihood Statistic

The $-2 \log$ likelihood statistic is a measure of the goodness of fit between the actual observations and the predicted probabilities. It is the summation:

$$-2 \sum_i y_i \ln(\mu_i) + (1 - y_i) \ln(1 - \mu_i)$$

where the y_i and μ_i are respectively the observed and predicted values of the dependent variable, and n is the number of observations. Note that $\ln(1)$ is zero and the observed values must be 0 or 1. Thus the closer the predicted values are to the observed, the closer this sum will be to zero.

The $-2 \log$ likelihood is also equal to the sum of the squared deviance residuals.

The $-2 \log$ likelihood (LL) statistic is related to the likelihood ratio (LR): $LR = LL - LL_0$ where LL_0 is the $-2 \log$ likelihood of a regression model having none of the independent variables, just a constant term. In viewing this relationship note that both LL_0 and LL are positive, and LL must be closer to zero reflecting a better fit. (At the extremes, LL will be zero when there is a perfect fit, and LL will equal LL_0 when there is no fit whatsoever). Thus the larger the LR the larger the implied explanatory power of the independent variables for the given dependent variable.

Threshold Probability for Positive Classification

The threshold probability value determines whether the response predicted by the logistic model in the classification and probability tables (see following sections) is a positive or a reference response. If the estimated probability in the probability table exceeds the specified threshold probability value, the predicted variable is assigned a positive response (value of 1); probabilities less than or equal to the specified value are assigned a value of 0 or a reference value. The threshold probability value is set in the options dialog box.

Classification Table

The classification table summarizes the results by cross-classifying the observed dependent response variables with predicted and identifying the number of correctly and incorrectly classified cases.

The responses classified by the logistic model are derived by comparing estimated logistic probabilities in the Probability Table to the specified threshold probability value (see preceding section).

This table appears in the report if the Classification Table option is selected in the Options dialog box.

Probability Table

The Probability Table lists the actual responses of the dependent variable, the estimated logistic probability of a positive response (a value of 1), and the predicted response of the dependent variables. The predicted responses are assigned values of 1 (positive response) or 0 (reference response) derived by comparing estimated logistic probabilities to the specified threshold probability value (see preceding section).

This table appears in the report if the Predicted Values option is selected in the Options dialog.

Statistical Summary Table

The summary table lists the coefficient, standard error, Wald Statistic, Odds Ratio, Odds Ratio Confidence, P value, and VIF for the independent variables.

Coefficients. The value for the constant and coefficients of the independent variables for the regression model are listed.

Standard Error. The standard errors of the regression coefficients (analogous to the standard error of the mean). The true regression coefficients of the underlying population generally fall within about two standard errors of the observed sample coefficients. Large standard errors may indicate multicollinearity.

Use these values to compute the Wald statistic and confidence intervals for the regression coefficients.

Wald Statistic. The Wald statistic is the regression coefficient divided by the standard error. It is computed as the ratio:

$$z = \frac{b_i}{s_{b_i}}$$

where z is the Wald Statistics, b_i is the observed value of the estimated coefficient, and s_{b_i} is the standard error of the coefficient.

P value. P is the P value calculated for the Wald statistic. The P value is the probability of being wrong in concluding that there is a true association between the variables. The P value is based on the chi-square distribution with one degree of freedom. The smaller the P value, the greater the probability that the independent variables affect the dependent variable.

Traditionally, you can conclude that the independent variable contributes to predicting the dependent variable when P < 0.05.

Odds Ratio. The odds ratio for an independent variable is computed as $\Omega = e^{\beta_1}$ where β_1 is the regression coefficient. The odds ratio is an estimate of the increase (or decrease) in the odds for an outcome if the independent variable value is increased by 1.

Odds Ratio Confidence. These two values represent the lower and upper ends of the confidence interval in which the true odds ratio lies. The level of confidence (95%) is specified in the options dialog.

VIF (Variance Inflation Factor). The variance inflation factor is a measure of multicollinearity. It measures the "inflation" of the standard error of each regression parameter (coefficient) for an independent variable due to redundant information in other independent variables.

If the variance inflation factor is 1.0, there is no redundant information in the other independent variables. If the variance inflation factor is much larger, there are redundant variables in the regression model, and the parameter estimates may not be reliable.

Variance inflation factor values for independent variables above the specified value are flagged with a $>$ symbol, indicating multicollinearity with other independent variables.

The presence of serious multicollinearity indicates that you have too many redundant independent variables in your regression equation. To improve the quality of the regression equation, you should delete the redundant variables. The cutoff value for flagging multicollinearity is set in the Options dialog box. The suggested value is 4.0.

Residual Calculation Method

The residual calculation method indicates how the residuals for the logistic regression are calculated. You can choose Pearson or Deviance residuals from the Options for Logistic Regression dialog. This choice does not affect the logistic regression itself, which minimizes the deviance residuals squared, but does affect how the Studentized residuals are calculated.

The Pearson residual is defined as:

$$\frac{y_i - \mu_i}{\sqrt{\mu_i(1-\mu_i)}}$$

where y_i and μ_i are respectively the observed and predicted values of the dependent variable for the i^{th} case.

The deviance residual is defined as:

$$\frac{y_i - \mu_i}{\sqrt{\mu_i(1-\mu_i)}}$$

Residuals Table

The residuals table displays the raw, Pearson or Deviance, Studentized, and Studentized deleted residuals if the associated options are selected in the options dialog. All residuals that qualify as outlying values are flagged with a < symbol. The trigger values to flag residuals as outliers are also set in the Options for Multiple Logistic Regression dialog.

If you selected Report Flagged Values Only, only those observations that have one or more residuals flagged as outliers are reported; however, all other results for that observation are also displayed. The way the residuals are calculated depend on whether Pearson or Deviance is selected as the residual type in the Options dialog box.

Row. This is the row number of the observation. Note that if your data has a case with a value missing, the corresponding row is entirely omitted from the table of residuals.

Pearson/Deviance Residuals. The Residual table displays either Pearson or Deviance residuals, depending on the Residual Type option setting in the Options for Logistic Regression dialog box.

Both Pearson and Deviance residuals indicate goodness of fit between the logistic equation and the data, with smaller values indicating a better fit. These two residual types are calculated differently and affect the way the Studentized residuals in the table are calculated.

Pearson residuals, also known as standardized residuals, are the raw residuals divided by the standard error. Deviance residuals are a measure of how much each point contributes to the likelihood function being minimized as part of the maximum likelihood procedure.

Raw Residuals. Raw residuals are the difference between the predicted and observed values for each of the subjects or cases.

Studentized Residuals. The Studentized residual is a standardized residual that also takes into account the greater confidence of the predicted values of the dependent variable in the "middle" of the data set.

This residual is also known as the internally Studentized residual, because the standard error of the estimate is computed using all data.

Studentized Deleted Residual. The Studentized deleted residual, or externally Studentized residual, is a Studentized residual which uses the standard error, computed after deleting the data point associated with the residual.

Both Studentized and Studentized deleted residuals that lie outside a specified confidence interval for the regression are flagged as outlying points; the suggested confidence value is 95%.

The Studentized deleted residual is more sensitive than the Studentized residual in detecting outliers, since the Studentized deleted residual results in much larger values for outliers than the Studentized residual.

Influence Diagnostics

The influence diagnostic results display only the values for the results selected in the Options dialog under the More Statistics tab. All results that qualify as outlying values are flagged with a < symbol. The trigger values to flag data points as outliers are also set in the Options dialog under the More Statistics tab.

If you selected Report Cases with Outliers Only, only observations that have one or more observations flagged as outliers are reported; however, all other results for that observation are also displayed.

Row. This is the row number of the observation.

Cook's Distance. Cook's distance is a measure of how great an effect each point has on the estimates of the parameters in the regression equation. It is a measure of how much the values of the regression coefficients would change if that point is deleted from the analysis.

Values above 1 indicate that a point is possibly influential. Cook's distances exceeding 4 indicate that the point has a major effect on the values of the parameter estimates. Points with Cook's distances greater than the specified value

are flagged as influential; the suggested value is 4. The Cook's Distance value used to flag "large" values is set in the Options dialog box.

Leverage. Leverage values identify potentially influential points. Observations with leverages a specified factor greater than the expected leverages are flagged as potentially influential points; the suggested value is 2.0 times the expected leverage.

$$\frac{k+1}{n}$$

The expected leverage of a data point is $\frac{k+1}{n}$ where there are k independent variables and n data points.

Because leverage is calculated using only the dependent variable, high leverage points tend to be at the extremes of the independent variables (large and small values), where small changes in the independent variables can have large effects on the predicted values of the dependent variable.

Polynomial Regression

Use Polynomial Regression to when you:

- Want to predict a trend in the data, or predict the value of one variable from the value of another variable, by fitting a curve through the data that does not follow a straight line, and
- Know there is only one independent variable

The *independent* variable is the known, or predictor, variable. When the independent variable is varied, a corresponding value for the *dependent*, or response, variable is produced.

If the relationships between the independent variables and the dependent variables is first order (a straight line), use *Multiple Linear Regression*. If the relationship is not a linear polynomial (for example, a log or exponential function), use *Nonlinear Regression*.

About the Polynomial Regression

Polynomial Regression assumes an association between the independent and dependent variables that fits the general equation for a polynomial of order k $y=b_0+b_1x^1+b_2x^2+b_3x^3+\dots+b_kx^k$ where y is the dependent variable, x is the independent variable, and b_1, b_2, b_3 are the regression coefficients. As the value for x varies, the corresponding value varies according to a polynomial function.

The *order* of the polynomial k is the highest exponent of the independent variable; a first order polynomial is a straight line, a second order (quadratic) polynomial is a parabola, and so on.

Polynomial Regression is a parametric test, that is, for a given independent variable value, the possible values for the dependent variable are assumed to be normally distributed and have equal variance.

- i** **Tip:** If you are fitting a polynomial to data, the polynomial regression procedure yields more reliable results than simply performing a Multiple Linear Regression using x, x^2 , and so on, as the independent variables.

Performing a Polynomial Regression

To perform a Polynomial Regression:

1. [Enter or arrange your data in the worksheet.](#)
2. Set the polynomial regression options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Regression > Polynomial
5. Run the test.
6. [View and interpret the incremental polynomial regression reports.](#)
7. [View and interpret the order only polynomial regression reports.](#)
8. [Generate report graphs.](#)

Arranging Polynomial Regression Data

Place the data for the dependent variable in one column and the corresponding data for the observed independent variable in another column.

Observations containing missing values are ignored, and all columns must be equal in length.

Setting Polynomial Regression Options

Use the Polynomial Regression options to:

- Set the polynomial order.
- Specify the type of polynomial regression you want to perform (incremental evaluation or order only).
- Set the assumption checking options.
- Specify the residuals to display and save them to the worksheet.
- Display confidence intervals and save them to the worksheet.
- Display the PRESS prediction error and the standardized coefficients.
- Display the power.

To change Polynomial Regression options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, select **Polynomial Regression** from the **Select Test** drop-down list.
4. Click **Options**. The **Options for Polynomial Regression** dialog box appears. If you select **Incremental Order** as the regression type, only the **Criterion** options are available. If you select **Order Only**, then the following tabs appear:
 - **Criterion**. Click the **Criterion** tab to return to the Normality, Constant Variance, and Durbin-Watson options.
 - **Assumption Checking**. Click the **Assumption Checking** tab to view the Normality, Constant Variance, and Durbin-Watson options.
 - **Residuals**. Click the **Residuals** tab to view the residual options.
 - **More Statistics**. Click the **More Statistics** tab to view the confidence intervals, PRESS Prediction Error, Standardized Coefficients options.
 - **Post Hoc**. Click the **Post Hoc Tests** tab to view the Power options.

Options settings are saved between SigmaPlot sessions.

5. To continue the test, click **Run Test**.
6. To accept the current settings and close the dialog box, click **OK**.

Options for Polynomial Regression: Criterion

Select the Criterion tab from the options dialog to view the Polynomial Order and Regression options. Use these options to specify the polynomial order to use and the type of polynomial to use to evaluate your data.

Polynomial Order. Select the desired polynomial order from the Polynomial Order drop-down list. You can also type the desired value on the drop-down box. This value is used either as the maximum order to evaluate or the specific order to compute.

Order Only. Select Order Only from the Regression drop-down list to fit only the order specified in the Polynomial Order edit box to the data.

Incremental Evaluation. Select Incremental Evaluation if you need to find the order of polynomial to use. This option evaluates each polynomial order equation starting at zero and increasing to the value specified in the Polynomial Order box.

Note this option does not display all regression results; instead, it is used to evaluate the order for the best model to use. Once the order is determined, run an order only polynomial regression to obtain complete regression results.

Model Selection Criteria. Select any combination of four criteria to measure the quality of fit when using Incremental Evaluation to determine the best model for your data. The four criteria are R-Square, Adjusted R-Square, Predicted R-Square, and the corrected AIC (Akaike Information Criterion). The program's default selection is R-Square.

Options for Polynomial Regression: Assumption Checking

Click the **Assumption Checking** tab from the options dialog to view the Normality, Constant Variance, and Durbin-Watson options. These options test your data for its suitability for regression analysis by checking three assumptions that a polynomial regression makes about the data. A polynomial regression assumes:

- That the source population is normally distributed about the regression.
- The variance of the dependent variable in the source population is constant regardless of the value of the independent variable(s).
- That the residuals are independent of each other.

All assumption checking options are selected by default. Only disable these options if you are certain that the data was sampled from normal populations with constant variance and that the residuals are independent of each other.

Normality Testing. SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

Constant Variance Testing. SigmaPlot tests for constant variance by computing the Spearman rank correlation between the absolute values of the residuals and the observed value of the dependent variable. When this correlation is significant, the constant variance assumption may be violated, and you should consider trying a different model (for example, one that more closely follows the pattern of the data), or transforming one or more of the independent variables to stabilize the variance.

P Values for Normality and Constant Variance. The P value determines the probability of being incorrect in concluding that the data is not normally distributed (P value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the P computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality and/or constant variance, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.05. Larger values of P (for example, 0.10) require less evidence to conclude that the residuals are not normally distributed or the constant variance assumption is violated.

To relax the requirement of normality and/or constant variance, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.01 for the normality test requires greater deviations from normality to flag the data as non-normal than a value of 0.05.



Note: Although the assumption tests are robust in detecting data from populations that are non-normal or with non-constant variances, there are extreme conditions of data distribution that these tests cannot detect; however, these conditions should be easily detected by visually examining the data without resorting to the automatic assumption tests.

Durbin-Watson Statistic. SigmaPlot uses the Durbin-Watson statistic to test residuals for their independence of each other. The Durbin-Watson statistic is a measure of serial correlation between the residuals. The residuals are often correlated when the independent variable is time, and the deviation between the observation and the regression line at one time are related to the deviation at the previous time. If the residuals are not correlated, the Durbin-Watson statistic will be 2.

Difference from 2 Value. Enter the acceptable deviation from 2.0 that you consider as evidence of a serial correlation in the Difference for 2.0 box. If the computed Durbin-Watson statistic deviates from 2.0 more than the entered value, SigmaPlot warns you that the residuals may not be independent. The suggested deviation value is 0.50, for example, Durbin-Watson Statistic values greater than 2.5 or less than 1.5 flag the residuals as correlated.

To require a stricter adherence to independence, decrease the acceptable difference from 2.0.

To relax the requirement of independence, increase the acceptable difference from 2.0.

Options for Polynomial Regression: Residuals

Click the **Residuals** tab in the **Options for Polynomial Regression** dialog box to view the Predicted Values, Raw, Standardized, Studentized Deleted, and Report Flagged Values Only options.

Predicted Values. Use this option to calculate the predicted value of the dependent variable for each observed value of the independent variable(s), then save the results to the worksheet. Click the selected check box if you do not want to include raw residuals in the worksheet.

To assign predicted values to a worksheet column, select the worksheet column you want to save the predicted values to from the corresponding drop-down list. If you select none and the **Predicted Values** check box is selected, the values appear in the report but are not assigned to the worksheet.

Raw Residuals. The raw residuals are the differences between the predicted and observed values of the dependent variables. To include raw residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include raw residuals in the worksheet.

To assign the raw residuals to a worksheet column, select the number of the desired column from the corresponding drop-down list. If you select none from the drop-down list and the Raw check box is selected, the values appear in the report but are not assigned to the worksheet.

Standardized Residuals. Select Standardized Residuals to include them in the report. The standardized residual is the residual divided by the standard error of the estimate. The standard error of the residuals is essentially the standard deviation of the residuals, and is a measure of variability around the regression line.

SigmaPlot automatically flags data points lying outside of the confidence interval specified in the corresponding box. These data points are considered to have "large" standardized residuals, for example, outlying data points. You can change which data points are flagged by editing the value in the **Flag Values > edit** box. The suggested residual t distribution, so the

Studentized Residuals. Select Studentized Residuals to include them in the report. Studentized residuals scale the t distribution, so the

distribution can be used to define "large" values of the Studentized residuals.

SigmaPlot automatically flags data points with "large" values of the Studentized residuals.

Studentized Deleted Residuals. Studentized deleted residuals are similar to the Studentized residual, except that the t distribution can be used to define "large" values of the Studentized deleted residual, for example,

Report Flagged Values Only. To only include only the flagged standardized and Studentized deleted residuals in the report, select Report Flagged Values Only.

Options for Polynomial Regression: More Statistics

Click the **More Statistics** tab in the options dialog to view the confidence interval options. You can set the confidence interval for the population, regression, or both and then save them to the worksheet.

Confidence Interval for the Population. The confidence interval for the population gives the range of μ s that

Population

The confidence interval for the regression line gives the range of μ s

To include confidence intervals for the regression in the report, select **Regression** and then specify a confidence level by entering a value in the percentage box. The confidence level can be any value from 1 to 99. The suggested confidence level for all intervals is 95%.

Clear the selected check box if you do not want to include the confidence intervals for the population in the report.

Saving Confidence Intervals to the Worksheet. To save the confidence intervals to the worksheet, select the column number of the first column you want to save the intervals to from the **Starting in Column** drop-down list. The selected intervals are saved to the worksheet starting with the specified column and continuing with successive columns in the worksheet.

PRESS Prediction Error. The PRESS Prediction Error is a measure of how well the regression equation predicts the observations. Leave this check box selected to evaluate the fit of the equation using the PRESS statistic. Clear the selected check box if you do not want to include the PRESS statistic in the report.

Standardized Coefficients. These are the coefficients of the regression equation standardized to dimensionless

values, $\hat{b}_i = b_i \frac{s_x}{s_y}$ where b_i = regression coefficient, s_x = standard deviation of the independent variable x_i , and s_y = standard deviation of dependent variable y .

To include the standardized coefficients in the report, make sure that you select **Standardized Coefficients**. Clear that option if you do not want to include the standardized coefficients in the worksheet.

Options for Polynomial Regression: Post Hoc Tests

Click the **Post Hoc Tests** tab on the **Options for Polynomial Regression** dialog box to view the Power options.

The power of a regression is the power to detect the observed relationship in the data. The alpha (α) is the acceptable probability of incorrectly concluding there is a relationship.

Select **Power** to compute the power for the polynomial regression data. Change the alpha value by editing the number in the **Use Alpha Value** edit box. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant relationship when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant relationship, but a greater possibility of concluding there is no relationship when one exists. Larger values of α make it easier to conclude that there is a relationship, but also increase the risk of reporting a false positive.

Running a Polynomial Regression

To run a Polynomial Regression you need to select the data to test. Use the **Select Data** panel of the Test Wizard to select the worksheet columns with the data you want to test.

To run a Polynomial Regression:

1. If you want to select your data before you run the regression, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Regression > Polynomial

The **Select Data** panel of the Test Wizard appears. If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog prompts you to pick your data.

4. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Dependent** and **Independent** drop-down list.

The first selected column is assigned to the **Dependent Variable** row in the **Selected Columns** list, and the second column is assigned to the **Independent Variable** row. The title of selected columns appears in each row. You are only prompted for one dependent and one independent variable column.

5. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.

6. Click **Finish** to run the regression. If you elected to test for normality, constant variance, and/or independent residuals, SigmaPlot performs the tests for normality (Shapiro-Wilk or Kolmogorov-Smirnov), constant variance, and independent residuals. If your data fail either of these tests, SigmaPlot warns you. When the test is complete, the report appears displaying the results of the Polynomial Regression.

If you are performing a regression using one order only, and selected to place predicted values, residuals, and/or other test results in the worksheet, they are placed in the specified data columns and are labeled by content and source column.

 **Remember:** Worksheet results can only be obtained using order only polynomial regression.

Interpreting Incremental Polynomial Regression Results

Incremental Order Polynomial Regression results display the regression equations for each order polynomial, starting with zero order and increasing to the specified order. The residual and incremental mean square, and incremental and overall R^2 , F value, and P value for each order equation are listed.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

Regression Equation

These are the regression equations for each order, with the values of the coefficients in place. The equations take the form: $k y = b_0 + b_1 x^1 + b_2 x^2 + b_3 x^3 + \dots + b_k x^k$ where y is the dependent variable, x is the independent variable, and b_1, b_2, b_3 are the regression coefficients.

The order k of the polynomial is the largest exponent of the independent variable.

For incremental polynomial regression, all equations from zero order up to the maximum order specified in the Options for Polynomial Regressions dialog box are listed.

Incremental Results

MSres (Residual Mean Square). The residual mean square is a measure of the variation of the residuals about the regression line.

$$\frac{\text{residual sum of squares}}{\text{residual degrees of freedom}} = \frac{SS_{\text{res}}}{DF_{\text{res}}} = MS_{\text{res}}$$

MSincr (Incremental Mean Square). The incremental mean square is a measure of the reduction in variation of the residuals about the regression equation gained with this order polynomial

$$\frac{\text{incremental sum of squares}}{\text{incremental degrees of freedom}} = \frac{SS_{\text{incr}}}{DF_{\text{incr}}} = MS_{\text{incr}}$$

The sum of squares are measures of variability of the dependent variable.

The residual sum of squares is a measure of the size of the residuals, which are the differences between the observed values of the dependent variable and the values predicted by regression model.

The incremental or Type I sum of squares, is a measure of the new predictive information contained in the added power of the independent variable, as it is added to the equation.

It is a measure of the increase in the regression sum of squares (and reduction in the sum of squared residuals) obtained when the highest order term of the independent variable is added to the regression equation, after all lower order terms have been entered. Since one order is added in each step, $DF_{\text{incr}} = 1$.

Rsq. R^2 , the coefficient of determination, is a measure of how well the regression model describes the data.

- The incremental R^2 is the gain in R^2 obtained with this order polynomial over the previous order polynomial
- The overall R^2 is the actual R^2 of this order polynomial

Overall R^2 values nearer to 1 indicate that the curve is a good description of the relation between the independent and dependent variables. R^2 is near 0 when the values of the independent variable poorly predict the dependent variables.

F Value. The F test statistic gauges the ability of the independent variable in predicting the dependent variable.

- The incremental F value gauges the increase in contribution of each added order of the independent variable in predicting the dependent variable. It is the ratio

$$\frac{\text{incremental variation from the dependent variable mean}}{\text{residual variation about the regression curve}} = \frac{MS_{\text{incr}}}{MS_{\text{res}}}$$

If the incremental F is large and the overall F jumps to a large number, you can conclude that adding that order of the independent variables predicts the dependent variable significantly better than the previous model. The "best" order polynomial to use is generally the highest order polynomial that produces a marked improvement in predictive ability.

- Overall F value gauges the contribution of all orders of the independent variable in predicting the dependent variable. It is the ratio

$$\frac{\text{regression variation from the dependent variable mean}}{\text{residual variation about the regression curve}} = \frac{MS_{\text{reg}}}{MS_{\text{res}}}$$

When the overall F ratio is around 1, you can conclude that there is no association between the independent variables (for example, the data is consistent with the null hypothesis that all the samples are just randomly distributed).

P Value. P is the P value calculated for F. The P value is the probability of being wrong in concluding that there is a true association between the dependent and independent variables (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on F). The smaller the P value, the greater the probability that there is an association.

- The incremental P value is the change in probability of being wrong that the added independent variable order improves the prediction of the dependent variable.
- The overall P value is the probability of being wrong that order of polynomial correctly predicts the dependent variable.

Traditionally, you can conclude that the independent variable can be used to predict the dependent variable when $P < 0.05$.

Although large incremental values for the coefficient of determination and the F-statistics can reveal the best model for your data, they are not always simple to interpret and may not reliably account for overfitting that results when using a model with too many parameters. Overfitting can be a serious problem because it increases the model's complexity, increases the sensitivity of the model's parameters to changes in the input data, and can ruin the model's predictive capability. There are other measures of quality of fit that are frequently better indicators of the best model for your data and are simpler to interpret.

Adj Rsqr. Adjusted is a modified version of that is adjusted for the number of coefficients in the model.

$$Adj R^2 = 1 - \frac{MS_{\text{res}}}{MS_{\text{tot}}}$$

The value of $Adj R^2$ is always less than R^2 , but it can be negative. The value of $Adj R^2$ generally increases only if additional terms added to the model, thus increasing its order, significantly improves the fit.

- The incremental $Adj R^2$ is the gain in $Adj R^2$ obtained with this order polynomial over the previous order polynomial.
- The overall $Adj R^2$ is the actual value for this order polynomial. A higher overall value of $Adj R^2$ indicates a higher quality of fit to your data.

Pred Rsqr. Predicted R^2 measures the quality of predictability of a model by showing how stable the model is when observations are removed. This measure also guards against overfitting.

$$Pred R^2 = 1 - \frac{PRESS}{SST_p}$$

$$\text{where } \text{PRESS} = \sum_{i=1}^n (y_i - \hat{y}_{(i)})^2 \quad \text{and} \quad \text{SST}_p = \sum_{i=1}^n (y_i - \bar{y}_{(i)})^2 = \left(\frac{n}{n-1}\right)^2 \text{SS}_{\text{res}}$$

The value $\hat{y}_{(i)}$ is the predicted value of the observation y_i after fitting the model to the data with this observation removed. The value $\bar{y}_{(i)}$ is the mean of observations after removing y_i

- The incremental is the gain in obtained with this order polynomial over the previous order polynomial.
- The overall is the actual value for this order polynomial. A higher overall value of indicates a higher quality of fit to your data.

AICc. The Akaike Information Criterion provides a relative measure of the information loss when modelling your data that provides a reliable trade-off between maximizing the likelihood of the estimated model and minimizing overfitting.

$$AIC = 2k + n \ln \left(\frac{\text{SS}_{\text{res}}}{n} \right)$$

Here, k is the number of coefficients in the model plus one. The corrected version $AICc$, which works well for small sample sizes, is the value computed in reports.

$$AICc = AIC + \frac{2k(k+1)}{n-k-1}$$

- The incremental $AICc$ is the gain in $AICc$ obtained with this order polynomial over the previous order polynomial.
- The overall $AICc$ is the actual value for this order polynomial. A lower overall value of $AICc$ indicates a higher quality of fit to your data.

Assumption Testing

Normality. Normality test result displays whether or not the polynomial model passed or failed the test of the assumption that the source population is normally distributed around the regression curve, and the P value calculated by the test. All regression requires a source population to be normally distributed about the regression curve.

When this assumption may be violated, a warning appears in the report. Failure of the normality test can indicate the presence of outlying influential points or an incorrect regression model.

Constant Variance. The constant variance test results list whether or not that polynomial model passed the test for constant variance of the residuals about the regression, and the P value computed for that order polynomial. All regression techniques require a normal distribution of the residuals about the regression curve.

Choosing the Best Model

The smaller the residual sum of squares and mean square, the closer the fit curve matches the data at those values of the independent variable. The first model that has a significant increase in the incremental F value is a good indicator of the best model to use. Because the R^2 value increases as the order increases, you also want to use the simplest model corresponding to the largest incremental increase in R^2 . The other measures we provide for quality of fit are generally simpler to interpret and are designed specifically to avoid overfitting. Evaluation of the best model using these measures is obtained by examining the overall results table of the report. For $Adj R^2$ and $Pred R^2$, locate the largest value that occurs and choose the smallest order with this value as the best model for your data. For $AICc$, locate the smallest value that occurs and choose the smallest order with this value as the best model for your data.

Interpreting Order Only Polynomial Regression Results

The report for an order only Polynomial Regression displays the equation with the computed coefficients for the curve, R and R^2 , mean squares, F , and the P value for the regression equation.

The other results displayed in the report are selected in the Options for Polynomial Regression dialog.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

Regression Equation

This is the equation with the values of the coefficients in place. This equation takes the form: $y = b_0 + b_1x^1 + b_2x^2 + b_3x^3 + \dots + b_kx^k$ where y is the dependent variable, x is the independent variable, and b_1, b_2, b_3 are the regression coefficients.

The order of the polynomial is the exponent of the independent variable. The number of observations N is also displayed, with the missing values, if any.

Analysis of Variance (ANOVA)

MSres (Residual Mean Square). The mean square provides an estimate of the population variance. The residual mean square is a measure of the variation of the residuals about the regression curve, or

$$\frac{\text{residual sum of squares}}{\text{residual degrees of freedom}} = \frac{SS_{\text{res}}}{DF_{\text{res}}} = MS_{\text{res}}$$

Rsq. The coefficient of determination R^2 is a measure of how well the regression model describes the data.

R^2 values near 1 indicate that the curve is a good description of the relation between the independent and dependent variables. R^2 values near 0 indicate that the values of the independent variable do not predict the dependent variables.

F Statistic. The F test statistic gauges the contribution of the regression equation to predict the dependent variable. It is the ratio

$$\frac{\text{regression variation from the dependent variable mean}}{\text{residual variation about the regression curve}} = \frac{MS_{\text{reg}}}{MS_{\text{res}}}$$

If F is a large number, you can conclude that the independent variable contributes to the prediction of the dependent variable (for example, the "unexplained variability" is smaller than what is expected from random sampling variability of the dependent variable about its mean). If the F ratio is around 1, you can conclude that there is no association between the variables (for example, the data is consistent with the null hypothesis that all the samples are just randomly distributed).

P Value. P is the P value calculated for F . The P value is the probability of being wrong in concluding that there is a true association between the variables (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on F). The smaller the P value, the greater the probability that the variables are correlated.

Standard Error of the Estimate

The standard error of the estimate s_{yx} is a measure of the actual variability about the regression line of the underlying population. The underlying population generally falls within about two standard errors of the observed sample.

PRESS Statistic. PRESS, the Predicted Residual Error Sum of Squares, is a measure of how well a regression model predicts new data. The smaller the PRESS statistic, the better the predictive ability of the model.

The PRESS statistic is computed by summing the squares of the prediction errors (the differences between predicted and observed values) for each observation, with that point deleted from the computation of the regression equation.

Durbin-Watson Statistic

The Durbin-Watson statistic is a measure of correlation between the residuals. If the residuals are not correlated, the Durbin-Watson statistic will be 2; the more this value differs from 2, the greater the likelihood that the residuals are correlated. This result appears if it was selected in the Options for Polynomial Regression dialog.

Normality Test

The normality test results display whether or not the polynomial model passed or failed the test of the assumption that the source population is normally distributed around the regression curve, and the P value calculated by the test. All regression requires a source population to be normally distributed about the regression curve.

When this assumption may be violated, a warning appears in the report. Failure of the normality test can indicate the presence of outlying influential points or an incorrect regression model.

This result appears unless you disabled normality testing in the **Options for Polynomial Regression** dialog box.

Constant Variance Test

The constant variance test result displays whether or not the polynomial model passed or failed the test of the assumption that the variance of the dependent variable in the source population is constant regardless of the value of the independent variable, and the P value calculated by the test. When the constant variance assumption may be violated, a warning appears in the report.

If you receive this warning, you should consider trying a different model (for example, one that more closely follows the pattern of the data), or transforming the independent variable to stabilize the variance and obtain more accurate estimates of the parameters in the regression equation.

This result appears unless you disabled constant variance testing in the Options for Polynomial Regression dialog box.

Regression Diagnostics

The regression diagnostic results display only the values for the predicted values, residual results, and other diagnostics selected in the Options for Polynomial Regression dialog. All results that qualify as outlying values are flagged with a < symbol. The trigger values to flag residuals as outliers are set in the Options for Polynomial Regression dialog.

If you selected Report Cases with Outliers Only, only those observations that have one or more residuals flagged as outliers are reported; however, all other results for that observation are also displayed.

Row. This is the row number of the observation.

Residuals. These are the raw residuals, the difference between the predicted and observed values for the dependent variables.

Standardized Residuals. The standardized residual is the raw residual divided by the standard error of the estimate S_{yx} .

If the residuals are normally distributed about the regression line, about 66% of the standardized residuals have values between -1 and +1, and about 95% of the standardized residuals have values between -2 and +2. A larger standardized residual indicates that the point is far from the regression line; the suggested value flagged as an outlier is 2.5.

Confidence Intervals

These results are displayed if you selected them in the **Options for Polynomial Regression** dialog box. If the confidence interval does not include zero, you can conclude that the coefficient is different than zero with the level of confidence specified. This can also be described as $P < \alpha$ (alpha), where α is the acceptable probability of incorrectly concluding that the coefficient is different than zero, and the confidence interval is $100(1 - \alpha)$.

The specified confidence level can be any value from 1 to 99; the suggested confidence level for both intervals is 95%.

Row. This is the row number of the observation.

Predicted. This is the value for the dependent variable predicted by the regression model for each observation.

Regression. These are the values that define the region containing the true relationship between the dependent and independent variables, for the specified level of confidence, centered at the predicted value.

This result is displayed if you selected it in the **Options for Polynomial Regression** dialog box. The specified confidence level can be any value from 1 to 99; the suggested confidence level is 95%.

Population Confidence Interval. These are the values that define the region containing the population from which the observations were drawn, for the specified level of confidence, centered at the predicted value.

This result is displayed if you selected it in the **Options for Polynomial Regression** dialog box. The specified confidence level can be any value from 1 to 99; the suggested confidence level is 95%.

Polynomial Regression Report Graphs

You can generate up to five graphs using the results from a Polynomial Regression. They include a:

- [Histogram of the residuals](#).
- [Scatter plot of the residuals](#).
- [Bar chart of the standardized residuals](#).
- [Normal probability plot of the residuals](#).
- [Line/scatter plot of the regression with one independent variable and confidence and prediction intervals](#).

Creating Polynomial Regression Report Graphs

To generate a report graph of Polynomial Regression report data:

1. With the Polynomial Regression report in view, click the **Report** tab.
2. In the **Result Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the Polynomial Regression report.

3. Select the type of graph you want to create from the **Graph Type** list, then click **OK**, or double-click the desired graph in the list.

The selected graph appears in a graph window. For more information, see [Report Graphs](#) on page 373.

Stepwise Linear Regression

Use Stepwise Linear Regression when you:

- Want to predict a trend in the data, or predict the value of one variable from the values of one or more other variables, by fitting a line or plane (or hyperplane) through the data.
- Do not know which independent variables contribute to predicting the dependent variable, and you want to find the model with suitable independent variables by adding or removing independent variables from the equation.

If you already know the independent variables you want to include, use Multiple Linear Regression. If you want to find the few best equations from all possible models, use Best Subsets Regression. If the relationship is not a straight line or plane, use Polynomial or Nonlinear Regression.

About Stepwise Linear Regression

Stepwise Regression is a technique for selecting independent variables for a Multiple Linear Regression equation from a list of candidate variables. Using Stepwise Regression instead of regular Multiple Linear Regression avoids using extraneous variables, or under specifying or over specifying the model.

Stepwise Regression assumes an association between the one or more independent variables and a dependent variable that fits the general equation for a multidimensional plane: $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_kx_k$ where y is the dependent variable, $x_1, x_2, x_3, \dots, x_k$ are the independent variables, and $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_kx_k$ are the regression coefficients. The independent variable is the known, or predicted, variable. As the values for x_1 vary, the corresponding value for y either increases or decreases, depending on the sign of b_1 . Stepwise Regression determines which independent variables to use by adding or removing selected independent variables from the equation.

There are two approaches to Stepwise Regression:

- **Forward Stepwise Regression.** In Forward Stepwise Regression, the independent variable that produces the best prediction of the dependent variable (and has an F value higher than a specified F-to-Enter) is entered into the equation first, the independent variable that adds the next largest amount of information is entered second, and so

on. After each variable is entered, the F value of each variable already entered into the equation is checked, and any variables with small F values (below a specified F-to-Remove value) are removed. This process is repeated until adding or removing variables does not significantly improve the prediction of the dependent variable.

- **Backward Stepwise Regression.** In Backward Stepwise Regression, all variables are entered into the equation. The independent variable that contributes the least to the prediction (and has an F value lower than a specified F-to-Remove) is removed from the equation, the next least important independent variable is removed, and so on. After each variable is removed, the F value of each variable removed from the equation is checked, and any variables with large F values (above a specified F-to-Enter value) are reentered into the equation. This process is repeated until removing or adding variables does not significantly improve the prediction of the dependent variable.

 **Note:** Forward and Backward Stepwise Regression using the same potential variables do not necessarily yield the same final regression model when there is multicollinearity among the possible independent variables.

Performing a Stepwise Linear Regression

To perform a Stepwise Linear Regression:

1. [Enter or arrange your data in the worksheet.](#)
2. If desired, set the Stepwise Regression options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Regression > Stepwise > Forward
or
Regression > Stepwise > Backward
5. Run the test.
6. [Generate report graphs.](#)

Arranging Stepwise Linear Regression Data

The data format for a Stepwise Linear Regression consists of the data for the independent variables in one or more columns and the corresponding data for the observed dependent variable in a single column. Any observations containing missing values are ignored, and the columns must be equal in length.

Setting Forward Stepwise Linear Regression Options

Use the Stepwise Regression options to:

- Specify which independent variables entered, replaced, deleted, and/or removed into or from a regression equation during forward or backwards stepwise regression.
- Set the number of steps permitted before the stepwise algorithm stops.
- Set assumption checking options.
- Specify the residuals to display and save them to the worksheet.
- Set confidence interval options.
- Display the PRESS statistic error.
- Display standardized regression coefficients.
- Display the power of the regression.

To change the Forward Stepwise Regression options:

1. If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
2. Select **Forward Stepwise Regression** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.

3. Click **Options**. The **Options for Forward Stepwise Regression** dialog box appears with five tabs:

- **Criterion.** Click the **Criterion** tab to return to the F-to-Enter, F-to-Remove, and Number of Steps options.
- **Assumption Checking.** Click the **Assumption Checking** tab to view the Normality, Constant Variance, and Durbin-Watson options.
- **Residuals.** Click the **Residuals** tab to view the residual options.
- **More Statistics.** Click the **More Statistics** tab to view the confidence intervals, PRESS Prediction Error, Standardized Coefficients options.
- **Other Diagnostics.** [Options for Forward Stepwise Regression: Other Diagnostics on page 275](#) Click the **Post Hoc Tests** tab to view the Power options.

Options settings are saved between SigmaPlot sessions.

4. To continue the test, click **Run Test**.

5. To accept the current settings and close the dialog box, click **OK**.

Options for Forward Stepwise Regression: Criterion

Click the **Criterion** tab from the options dialog box to view the F-to-Enter, F-to-Remove, and Number of Steps options. Use these options to specify the independent variables that are entered into, replaced, or removed from the regression equation during the stepwise regression, and to specify when the stepwise algorithm stops.

F-to-Enter Value. The F-to-Enter value controls which independent variables are entered into the regression equation during forward stepwise regression or replaced after each step during backwards stepwise regression.

The F-to-Enter value is the minimum incremental F value associated with an independent variable before it can be entered into the regression equation. All independent variables producing incremental F values above the F-to-Enter value are added to the model.

The suggested F-to-Enter value is 4.0. Increasing F-to-Enter requires a potential independent variable to have a greater effect on the ability of the regression equation to predict the dependent variable before it is accepted, but may stop too soon and exclude important variables.

 **Tip:** The F-to-Enter value should always be greater than or equal to the F-to-Remove value, to avoid cycling variables in and out of the regression model.

Reducing the F-to-Enter value makes it easier to add a variable, because it relaxes the importance of a variable required before it is accepted, but may produce redundant variables and result in multicollinearity.

 **Note:** If you are performing backwards stepwise regression and you want any variable that has been removed to remain deleted, increase the F-to-Enter value to a large number, for example, 100000.

F-to-Remove Value. The F-to-Remove value controls which independent variables are deleted from the regression equation during backwards stepwise regression, or removed after each step in forward stepwise regression.

The F-to-Remove is the maximum incremental F value associated with an independent variable before it can be removed from the regression equation. All independent variables producing incremental F values below the F-to-Remove value are deleted from the model.

The suggested F-to-Remove value is 3.9. Reducing the F-to-Remove value makes it easier to retain a variable in the regression equation because variables that have smaller effects on the ability of the regression equation to predict the dependent variable are still accepted. However, the regression may still contain redundant variables, resulting in multicollinearity.

 **Remember:** The F-to-Remove value should always be less than or equal to the F-to-Enter value, to avoid cycling variables in and out of the regression model.

Increasing the F-to-Remove value makes it easier to delete variables from the equation, as variables that contain more predictive value can be removed. Important variables may also be deleted, however.

 **Tip:** If you are performing forwards stepwise regression and you want any variable that has been entered to remain in the equation, set the F-to-Remove value to zero.

Number of Steps. Use this option to set the maximum number of steps permitted before the stepwise algorithm stops. Note that if the algorithm stops because it ran out of steps, the results are probably not reliable. The suggested number of steps is 20 added or deleted independent variables.

Options for Forward Stepwise Regression: Assumption Checking

Click the **Assumption Checking** tab from the options dialog box to view the Normality, Constant Variance, and Durbin-Watson options. These options test your data for its suitability for regression analysis by checking three assumptions that a Stepwise Linear Regression makes about the data. A Stepwise Linear Regression assumes:

- That the source population is normally distributed about the regression.
- The variance of the dependent variable in the source population is constant regardless of the value of the independent variable(s).
- That the residuals are independent of each other.

All assumption checking options are selected by default. Only disable these options if you are certain that the data was sampled from normal populations with constant variance and that the residuals are independent of each other.

Normality Testing. SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

Constant Variance Testing. SigmaPlot tests for constant variance by computing the Spearman rank correlation between the absolute values of the residuals and the observed value of the dependent variable. When this correlation is significant, the constant variance assumption may be violated, and you should consider trying a different model (for example, one that more closely follows the pattern of the data), or transforming one or more of the independent variables to stabilize the variance.

P Values for Normality and Constant Variance

The *P* value determines the probability of being incorrect in concluding that the data is not normally distributed (*P* value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the *P* computed by the test is greater than the *P* set here, the test passes.

To require a stricter adherence to normality and/or constant variance, increase the *P* value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.05. Larger values of *P* (for example, 0.10) require less evidence to conclude that the residuals are not normally distributed or the constant variance assumption is violated.

To relax the requirement of normality and/or constant variance, decrease *P*. Requiring smaller values of *P* to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a *P* value of 0.01 for the normality test requires greater deviations from normality to flag the data as non-normal than a value of 0.05.



Note: Although the assumption tests are robust in detecting data from populations that are non-normal or with non-constant variances, there are extreme conditions of data distribution that these tests cannot detect; however, these conditions should be easily detected by visually examining the data without resorting to the automatic assumption tests.

Durbin-Watson Statistic. SigmaPlot uses the Durbin-Watson statistic to test residuals for their independence of each other. The Durbin-Watson statistic is a measure of serial correlation between the residuals. The residuals are often correlated when the independent variable is time, and the deviation between the observation and the regression line at one time are related to the deviation at the previous time. If the residuals are not correlated, the Durbin-Watson statistic will be 2.

Difference from 2 Value Enter the acceptable deviation from 2.0 that you consider as evidence of a serial correlation in the Difference for 2.0 box. If the computed Durbin-Watson statistic deviates from 2.0 more than the entered value, SigmaPlot warns you that the residuals may not be independent. The suggested deviation value is 0.50, for example, Durbin-Watson Statistic values greater than 2.5 or less than 1.5 flag the residuals as correlated.

To require a stricter adherence to independence, decrease the acceptable difference from 2.0.

To relax the requirement of independence, increase the acceptable difference from 2.0.

Options for Forward Stepwise Regression: Residuals

Click the **Residuals** tab in the options dialog box to view the Predicted Values, Raw, Standardized, Studentized, Studentized Deleted, and Report Flagged Values Only options.

Predicted Values. Select this option to calculate the predicted value of the dependent variable for each observed value of the independent variable(s), then save the results to the data worksheet. Click the selected check box if you do not want to include raw residuals in the worksheet.

To assign predicted values to a worksheet column, select the worksheet column you want to save the predicted values to from the corresponding drop-down list. If you select none and the Predicted Values check box is selected, the values appear in the report but are not assigned to the worksheet.

Raw Residuals. The raw residuals are the differences between the predicted and observed values of the dependent variables. To include raw residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include raw residuals in the worksheet.

To assign the raw residuals to a worksheet column, select the number of the desired column from the corresponding drop-down list. If you select none from the drop-down list and the Raw check box is selected, the values appear in the report but are not assigned to the worksheet.

Standardized Residuals. The standardized residual is the residual divided by the standard error of the estimate. The standard error of the residuals is essentially the standard deviation of the residuals, and is a measure of variability around the regression line. To include standardized residuals in the report, make sure this check box is selected.

SigmaPlot automatically flags data points lying outside of the confidence interval specified in the corresponding box. These data points are considered to have "large" standardized residuals, for example, outlying data points. You can change which data points are flagged by editing the value in the Flag Values > edit box.

Studentized Residuals. Studentized residuals scale the standardized residuals by taking into account the greater precision of the regression line near the middle of the data versus the extremes. The Studentized residuals tend to be distributed according to the Student *t* distribution, so the *t* distribution can be used to define "large" values of the Studentized residuals. SigmaPlot automatically flags data points with "large" values of the Studentized residuals, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

To include Studentized residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include Studentized residuals in the worksheet.

Studentized Deleted Residuals Studentized deleted residuals are similar to the Studentized residual, except that the residual values are obtained by computing the regression equation without using the data point in question.

To include Studentized deleted residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include Studentized deleted residuals in the worksheet.

SigmaPlot can automatically flag data points with "large" values of the Studentized deleted residual, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

 **Note:** Both Studentized and Studentized deleted residuals use the same confidence interval setting to determine outlying points.

Report Flagged Values Only. To only include only the flagged standardized and Studentized deleted residuals in the report, make sure the Report Flagged Values Only check box is selected. Clear this option to include all standardized and Studentized residuals in the report.

Options for Forward Stepwise Regression: More Statistics

Click the **More Statistics** tab in the options dialog to view the confidence interval options. You can set the confidence interval for the population, regression, or both, and then save them to the worksheet.

Confidence Interval for the Population. The confidence interval for the population gives the range of values that define the region that contains the population from which the observations were drawn.

To include confidence intervals for the population in the report, make sure the Population check box is selected. Click the selected check box if you do not want to include the confidence intervals for the population in the report.

Confidence Interval for the Regression. The confidence interval for the regression line gives the range of values that defines the region containing the true mean relationship between the dependent and independent variables, with the specified level of confidence.

To include confidence intervals for the regression in the report, make sure the Regression check box is selected, then specify a confidence level by entering a value in the percentage box. The confidence level can be any value from 1 to 99. The suggested confidence level is 95%. Click the selected check box if you do not want to include the confidence intervals for the population in the report.

Clear the selected check box if you do not want to include the confidence intervals for the population in the report.

Saving Confidence Intervals to the Worksheet. To save the confidence intervals to the worksheet, select the column number of the first column you want to save the intervals to from the Starting in Column drop-down list. The selected intervals are saved to the worksheet starting with the specified column and continuing with successive columns in the worksheet.

PRESS Prediction Error. The PRESS Prediction Error is a measure of how well the regression equation predicts the observations. Leave this check box selected to evaluate the fit of the equation using the PRESS statistic. Clear the selected check box if you do not want to include the PRESS statistic in the report.

Standardized Coefficients. These are the coefficients of the regression equation standardized to dimensionless

$$\beta_i = b_i \frac{s_x}{s_y}$$
 values, where b_i = regression coefficient, s_x = standard deviation of the independent variable x_i , and s_y = standard deviation of dependent variable y .

To include the standardized coefficients in the report, select **Standardized Coefficients**. Clear the check box if you do not want to include the standardized coefficients in the worksheet.

Options for Forward Stepwise Regression: Other Diagnostics

Click the **Other Diagnostics** tab in the options dialog box to view the Influence, Variance Inflation Factor and Power options. If Other Diagnostic is hidden, click the right pointing arrow to the right of the tabs to move it into view. Use the left pointing arrow to move the other tabs back into view.

Influence options automatically detect instances of influential data points. Most influential points are data points which are outliers, that is, they do not "line up" with the rest of the data points. These points can have a potentially disproportionately strong influence on the calculation of the regression line. You can use several influence tests to identify and quantify influential points.

DFFITS. DFFITS i is the number of estimated standard errors that the predicted value changes for the i th data point when it is removed from the data set. It is another measure of the influence of a data point on the prediction used to compute the regression coefficients.

Predicted values that change by more than two standard errors when the data point is removed are considered to be influential.

Select the DFFITS check box to compute this value for all points and flag influential points, for example, those with DFFITS greater than the value specified in the Flag Values > edit box. The suggested value is 2.0 standard errors, which indicates that the point has a strong influence on the data. To avoid flagging more influential points, increase this value; to flag less influential points, decrease this value.

Leverage. Leverage is used to identify the potential influence of a point on the results of the regression equation. Leverage depends only on the value of the independent variable(s). Observations with high leverage tend to be at the extremes of the independent variables, where small changes in the independent variables can have large effects on the predicted values of the dependent variable.

The expected leverage of a data point is $\frac{(k+1)}{n}$, where there are k independent variables and n data points. Observations with leverages much higher than the expected leverages are potentially influential points.

Select the Leverage check box to compute the leverage for each point and automatically flag potentially influential points, for example, those points that could have leverages greater than the specified value times the expected leverage. The suggested value is 2.0 times the expected leverage for the regression (for example, $\frac{2(k+1)}{n}$). To avoid flagging more potentially influential points, increase this value; to flag points with less potential influence, lower this value.

Cook's Distance. Cook's distance is a measure of how great an effect each point has on the estimates of the parameters in the regression equation. Cook's distance assesses how much the values of the regression coefficients change if a point is deleted from the analysis. Cook's distance depends on both the values of the independent and dependent variables.

Select the Cook's Distance check box to compute this value for all points and flag influential points, for example, those with a Cook's distance greater than the specified value. The suggested value is 4.0. Cook's distances above 1 indicate that a point is possibly influential. Cook's distances exceeding 4 indicate that the point has a major effect on the values of the parameter estimates. To avoid flagging more influential points, increase this value: to flag less influential points, lower this value.

Report Flagged Values Only. To only include only the influential points flagged by the influential point tests in the report, make sure the Report Flagged Values Only check box is selected. Clear this option to include all influential points in the report.

Variance Inflation Factor

The Variance Inflation Factor option measures the multicollinearity of the independent variables, or the linear combination of the independent variables in the fit.

Regression procedures assume that the independent variables are statistically independent of each other, for example, that the value of one independent variable does not affect the value of another. However, this ideal situation rarely occurs in the real world. When the independent variables are correlated, or contain redundant information, the estimates of the parameters in the regression model can become unreliable.

The parameters in regression models quantify the theoretically unique contribution of each independent variable to predicting the dependent variable. When the independent variables are correlated, they contain some common information and "contaminate" the estimates of the parameters. If the multicollinearity is severe, the parameter estimates can become unreliable.

There are two types of multicollinearity.

- **Structural Multicollinearity.** Structural multicollinearity occurs when the regression equation contains several independent variables which are functions of each other. The most common form of structural multicollinearity occurs when a polynomial regression equation contains several powers of the independent variable. Because these powers (for example, x , x^2 , and so on) are correlated with each other, structural multicollinearity occurs. Including interaction terms in a regression equation can also result in structural multicollinearity.
- **Sample-Based Multicollinearity.** Sample-based multicollinearity occurs when the sample observations are collected in such a way that the independent variables are correlated (for example, if age, height, and weight are collected on children of varying ages, each variable has a correlation with the others).

SigmaPlot can automatically detect multicollinear independent variables using the variance inflation factor. Click the Other Diagnostics tab in the Options dialog to view the Variance Inflation Factor option.

Flagging Multicollinear Data. Use the value in the Flag Values > edit box as a threshold for multicollinear variables. The default threshold value is 4.0, meaning that any value greater than 4.0 will be flagged as multicollinear. To make this test more sensitive to possible multicollinearity, decrease this value. To allow greater correlation of the independent variables before flagging the data as multicollinear, increase this value.

When the variance inflation factor is large, there are redundant variables in the regression model, and the parameter estimates may not be reliable. Variance inflation factor values above 4 suggest possible multicollinearity; values above 10 indicate serious multicollinearity.

What to Do About Multicollinearity. Sample-based multicollinearity can sometimes be resolved by collecting more data under other conditions to break up the correlation among the independent variables. If this is not possible, the

regression equation is over parameterized and one or more of the independent variables must be dropped to eliminate the multicollinearity.

Structural multicollinearities can be resolved by centering the independent variable before forming the power or interaction terms.

Report Flagged Values Only. To only include only the points flagged by the influential point tests and values exceeding the variance inflation threshold in the report, make sure the Report Flagged Values Only check box is selected. Clear this option to include all influential points in the report.

Power

Select the Other Diagnostics tab in the options dialog to view the Power options. If Other Diagnostic is hidden, click the right pointing arrow to the right of the tabs to move it into view. Use the left pointing arrow to move the other tabs back into view.

The power of a regression is the power to detect the observed relationship in the data. The alpha (α) is the acceptable probability of incorrectly concluding there is a relationship.

Check the Power check box to compute the power for the stepwise linear regression data. Change the alpha value by editing the number in the Alpha Value edit box. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant relationship when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant relationship, but a greater possibility of concluding there is no relationship when one exists. Larger values of α make it easier to conclude that there is a relationship, but also increase the risk of reporting a false positive.

Setting Backward Stepwise Linear Regression Options

Use the Backward Stepwise Regression options to:

- Specify which independent variables entered, replaced, deleted, and/or removed into or from a regression equation during forward or backward stepwise regression.
- Set the number of steps permitted before the stepwise algorithm stops.
- Set assumption checking options.
- Specify the residuals to display and save them to the worksheet.
- Set confidence interval options.
- Display the PRESS statistic error.
- Display standardized regression coefficients.
- Display the power of the regression.

To change the Backward Stepwise Regression options:

- If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
- Select **Backward Stepwise Regression** from the **Tests** drop-down list in the **SigmaStat** group on the **Analysis** tab.
- Click **Options**. The **Options for Backward Stepwise Regression** dialog box appears with five tabs:
 - Criterion.** Click the Criterion tab to return to the F-to-Enter, F-to-Remove, and Number of Steps options.
 - Assumption Checking.** Click the Assumption Checking tab to view the Normality, Constant Variance, and Durbin-Watson options.
 - Residuals.** Click the Residuals tab to view the residual options.
 - More Statistics.** Click the More Statistics tab to view the confidence intervals, PRESS Prediction Error, Standardized Coefficients options.
 - Other Diagnostics.** Click the Post Hoc Tests tab to view the Power options.

Options settings are saved between SigmaPlot sessions.

- To continue the test, click **Run Test**.
- To accept the current settings and close the dialog box, click **OK**.

Options for Backward Stepwise Regression: Criterion

Click the **Criterion** tab from the options dialog box to view the F-to-Enter, F-to-Remove, and Number of Steps options. Use these options to specify the independent variables that are entered into, replaced, or removed from the regression equation during the stepwise regression, and to specify when the stepwise algorithm stops.

F-to-Enter Value. The F-to-Enter value controls which independent variables are entered into the regression equation during forward stepwise regression or replaced after each step during backwards stepwise regression.

The F-to-Enter value is the minimum incremental F value associated with an independent variable before it can be entered into the regression equation. All independent variables producing incremental F values above the F-to-Enter value are added to the model.

The suggested F-to-Enter value is 4.0. Increasing F-to-Enter requires a potential independent variable to have a greater effect on the ability of the regression equation to predict the dependent variable before it is accepted, but may stop too soon and exclude important variables.

 **Remember:** The F-to-Enter value should always be greater than or equal to the F-to-Remove value, to avoid cycling variables in and out of the regression model.

Reducing the F-to-Enter value makes it easier to add a variable, because it relaxes the importance of a variable required before it is accepted, but may produce redundant variables and result in multicollinearity.

 **Tip:** If you are performing backwards stepwise regression and you want any variable that has been removed to remain deleted, increase the F-to-Enter value to a large number, for example, 100000.

F-to-Remove Value. The F-to-Remove value controls which independent variables are deleted from the regression equation during backwards stepwise regression, or removed after each step in backward stepwise regression.

The F-to-Remove is the maximum incremental F value associated with an independent variable before it can be removed from the regression equation. All independent variables producing incremental F values below the F-to-Remove value are deleted from the model.

The suggested F-to-Remove value is 3.9. Reducing the F-to-Remove value makes it easier to retain a variable in the regression equation because variables that have smaller effects on the ability of the regression equation to predict the dependent variable are still accepted. However, the regression may still contain redundant variables, resulting in multicollinearity.

 **Remember:** The F-to-Remove value should always be less than or equal to the F-to-Enter value, to avoid cycling variables in and out of the regression model.

Increasing the F-to-Remove value makes it easier to delete variables from the equation, as variables that contain more predictive value can be removed. Important variables may also be deleted, however.

 **Tip:** If you are performing backward stepwise regression and you want any variable that has been entered to remain in the equation, set the F-to-Remove value to zero.

Number of Steps. Use this option to set the maximum number of steps permitted before the stepwise algorithm stops. Note that if the algorithm stops because it ran out of steps, the results are probably not reliable. The suggested number of steps is 20 added or deleted independent variables.

Options for Backward Stepwise Regression: Assumption Checking

Click the **Assumption Checking** tab from the options dialog box to view the Normality, Constant Variance, and Durbin-Watson options. These options test your data for its suitability for regression analysis by checking three assumptions that a Stepwise Linear Regression makes about the data. A Stepwise Linear Regression assumes:

- That the source population is normally distributed about the regression.
- The variance of the dependent variable in the source population is constant regardless of the value of the independent variable(s).
- That the residuals are independent of each other.

All assumption checking options are selected by default. Only disable these options if you are certain that the data was sampled from normal populations with constant variance and that the residuals are independent of each other.

Normality Testing. SigmaPlot uses either the Shapiro-Wilk or Kolmogorov-Smirnov test to test for a normally distributed population.

Constant Variance Testing. SigmaPlot tests for constant variance by computing the Spearman rank correlation between the absolute values of the residuals and the observed value of the dependent variable. When this correlation is significant, the constant variance assumption may be violated, and you should consider trying a different model (for example, one that more closely follows the pattern of the data), or transforming one or more of the independent variables to stabilize the variance.

P Values for Normality and Constant Variance

The *P* value determines the probability of being incorrect in concluding that the data is not normally distributed (*P* value is the risk of falsely rejecting the null hypothesis that the data is normally distributed). If the *P* computed by the test is greater than the *P* set here, the test passes.

To require a stricter adherence to normality and/or constant variance, increase the *P* value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.05. Larger values of *P* (for example, 0.10) require less evidence to conclude that the residuals are not normally distributed or the constant variance assumption is violated.

To relax the requirement of normality and/or constant variance, decrease *P*. Requiring smaller values of *P* to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a *P* value of 0.01 for the normality test requires greater deviations from normality to flag the data as non-normal than a value of 0.05.



Note: Although the assumption tests are robust in detecting data from populations that are non-normal or with non-constant variances, there are extreme conditions of data distribution that these tests cannot detect; however, these conditions should be easily detected by visually examining the data without resorting to the automatic assumption tests.

Durbin-Watson Statistic. SigmaPlot uses the Durbin-Watson statistic to test residuals for their independence of each other. The Durbin-Watson statistic is a measure of serial correlation between the residuals. The residuals are often correlated when the independent variable is time, and the deviation between the observation and the regression line at one time are related to the deviation at the previous time. If the residuals are not correlated, the Durbin-Watson statistic will be 2.

Difference from 2 Value Enter the acceptable deviation from 2.0 that you consider as evidence of a serial correlation in the Difference for 2.0 box. If the computed Durbin-Watson statistic deviates from 2.0 more than the entered value, SigmaPlot warns you that the residuals may not be independent. The suggested deviation value is 0.50, for example, Durbin-Watson Statistic values greater than 2.5 or less than 1.5 flag the residuals as correlated.

To require a stricter adherence to independence, decrease the acceptable difference from 2.0.

To relax the requirement of independence, increase the acceptable difference from 2.0.

Options for Backward Stepwise Regression: Residuals

Click the **Residuals** tab in the options dialog box to view the Predicted Values, Raw, Standardized, Studentized, Studentized Deleted, and Report Flagged Values Only options.

Predicted Values. Select this option to calculate the predicted value of the dependent variable for each observed value of the independent variable(s), then save the results to the data worksheet. Click the selected check box if you do not want to include raw residuals in the worksheet.

To assign predicted values to a worksheet column, select the worksheet column you want to save the predicted values to from the corresponding drop-down list. If you select none and the Predicted Values check box is selected, the values appear in the report but are not assigned to the worksheet.

Raw Residuals. The raw residuals are the differences between the predicted and observed values of the dependent variables. To include raw residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include raw residuals in the worksheet.

To assign the raw residuals to a worksheet column, select the number of the desired column from the corresponding drop-down list. If you select none from the drop-down list and the Raw check box is selected, the values appear in the report but are not assigned to the worksheet.

Standardized Residuals. The standardized residual is the residual divided by the standard error of the estimate. The standard error of the residuals is essentially the standard deviation of the residuals, and is a measure of variability around the regression line. To include standardized residuals in the report, make sure this check box is selected.

SigmaPlot automatically flags data points lying outside of the confidence interval specified in the corresponding box. These data points are considered to have "large" standardized residuals, for example, outlying data points. You can change which data points are flagged by editing the value in the Flag Values > edit box.

Studentized Residuals. Studentized residuals scale the standardized residuals by taking into account the greater precision of the regression line near the middle of the data versus the extremes. The Studentized residuals tend to be distributed according to the Student t distribution, so the t distribution can be used to define "large" values of the Studentized residuals. SigmaPlot automatically flags data points with "large" values of the Studentized residuals, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

To include Studentized residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include Studentized residuals in the worksheet.

Studentized Deleted Residuals Studentized deleted residuals are similar to the Studentized residual, except that the residual values are obtained by computing the regression equation without using the data point in question.

To include Studentized deleted residuals in the report, make sure this check box is selected. Click the selected check box if you do not want to include Studentized deleted residuals in the worksheet.

SigmaPlot can automatically flag data points with "large" values of the Studentized deleted residual, for example, outlying data points; the suggested data points flagged lie outside the 95% confidence interval for the regression population.

 **Note:** Both Studentized and Studentized deleted residuals use the same confidence interval setting to determine outlying points.

Report Flagged Values Only. To only include only the flagged standardized and Studentized deleted residuals in the report, make sure the Report Flagged Values Only check box is selected. Clear this option to include all standardized and Studentized residuals in the report.

Options for Backward Stepwise Regression: More Statistics

Click the **More Statistics** tab in the options dialog to view the confidence interval options. You can set the confidence interval for the population, regression, or both, and then save them to the worksheet.

Confidence Interval for the Population. The confidence interval for the population gives the range of values that define the region that contains the population from which the observations were drawn.

To include confidence intervals for the population in the report, make sure the Population check box is selected. Click the selected check box if you do not want to include the confidence intervals for the population in the report.

Confidence Interval for the Regression. The confidence interval for the regression line gives the range of values that defines the region containing the true mean relationship between the dependent and independent variables, with the specified level of confidence.

To include confidence intervals for the regression in the report, make sure the Regression check box is selected, then specify a confidence level by entering a value in the percentage box. The confidence level can be any value from 1 to 99. The suggested confidence level is 95%. Click the selected check box if you do not want to include the confidence intervals for the population in the report.

Clear the selected check box if you do not want to include the confidence intervals for the population in the report.

Saving Confidence Intervals to the Worksheet. To save the confidence intervals to the worksheet, select the column number of the first column you want to save the intervals to from the Starting in Column drop-down list.

The selected intervals are saved to the worksheet starting with the specified column and continuing with successive columns in the worksheet.

PRESS Prediction Error. The PRESS Prediction Error is a measure of how well the regression equation predicts the observations. Leave this check box selected to evaluate the fit of the equation using the PRESS statistic. Clear the selected check box if you do not want to include the PRESS statistic in the report.

Standardized Coefficients. These are the coefficients of the regression equation standardized to dimensionless

values, $\beta_i = b_i \frac{s_x}{s_y}$ where b_i = regression coefficient, s_x = standard deviation of the independent variable x_i , and s_y = standard deviation of dependent variable y .

To include the standardized coefficients in the report, select Standardized Coefficients. Clear the check box if you do not want to include the standardized coefficients in the worksheet.

Options for Backward Stepwise Regression: Other Diagnostics

Click the Other Diagnosticstab in the options dialog box to view the Influence, Variance Inflation Factor and Power options. If Other Diagnostic is hidden, click the right pointing arrow to the right of the tabs to move it into view. Use the left pointing arrow to move the other tabs back into view.

Influence options automatically detect instances of influential data points. Most influential points are data points which are outliers, that is, they do not "line up" with the rest of the data points. These points can have a potentially disproportionately strong influence on the calculation of the regression line. You can use several influence tests to identify and quantify influential points.

DFFITS. DFFITS i is the number of estimated standard errors that the predicted value changes for the i th data point when it is removed from the data set. It is another measure of the influence of a data point on the prediction used to compute the regression coefficients.

Predicted values that change by more than two standard errors when the data point is removed are considered to be influential.

Select the DFFITS check box to compute this value for all points and flag influential points, for example, those with DFFITS greater than the value specified in the Flag Values > edit box. The suggested value is 2.0 standard errors, which indicates that the point has a strong influence on the data. To avoid flagging more influential points, increase this value; to flag less influential points, decrease this value.

Leverage. Leverage is used to identify the potential influence of a point on the results of the regression equation. Leverage depends only on the value of the independent variable(s). Observations with high leverage tend to be at the extremes of the independent variables, where small changes in the independent variables can have large effects on the predicted values of the dependent variable.

$$\frac{k+1}{n}$$

The expected leverage of a data point is $\frac{k+1}{n}$, where there are k independent variables and n data points. Observations with leverages much higher than the expected leverages are potentially influential points.

Select the Leverage check box to compute the leverage for each point and automatically flag potentially influential points, for example, those points that could have leverages greater than the specified value times the expected

$$\frac{2(k+1)}{n}$$

leverage. The suggested value is 2.0 times the expected leverage for the regression (for example, $\frac{2(k+1)}{n}$). To avoid flagging more potentially influential points, increase this value; to flag points with less potential influence, lower this value.

Cook's Distance. Cook's distance is a measure of how great an effect each point has on the estimates of the parameters in the regression equation. Cook's distance assesses how much the values of the regression coefficients change if a point is deleted from the analysis. Cook's distance depends on both the values of the independent and dependent variables.

Select the Cook's Distance check box to compute this value for all points and flag influential points, for example, those with a Cook's distance greater than the specified value. The suggested value is 4.0. Cook's distances above 1

indicate that a point is possibly influential. Cook's distances exceeding 4 indicate that the point has a major effect on the values of the parameter estimates. To avoid flagging more influential points, increase this value: to flag less influential points, lower this value.

Report Flagged Values Only. To only include only the influential points flagged by the influential point tests in the report, make sure the Report Flagged Values Only check box is selected. Clear this option to include all influential points in the report.

Variance Inflation Factor

The Variance Inflation Factor option measures the multicollinearity of the independent variables, or the linear combination of the independent variables in the fit.

Regression procedures assume that the independent variables are statistically independent of each other, for example, that the value of one independent variable does not affect the value of another. However, this ideal situation rarely occurs in the real world. When the independent variables are correlated, or contain redundant information, the estimates of the parameters in the regression model can become unreliable.

The parameters in regression models quantify the theoretically unique contribution of each independent variable to predicting the dependent variable. When the independent variables are correlated, they contain some common information and "contaminate" the estimates of the parameters. If the multicollinearity is severe, the parameter estimates can become unreliable.

There are two types of multicollinearity.

- **Structural Multicollinearity.** Structural multicollinearity occurs when the regression equation contains several independent variables which are functions of each other. The most common form of structural multicollinearity occurs when a polynomial regression equation contains several powers of the independent variable. Because these powers (for example, x , x^2 , and so on) are correlated with each other, structural multicollinearity occurs. Including interaction terms in a regression equation can also result in structural multicollinearity.
- **Sample-Based Multicollinearity.** Sample-based multicollinearity occurs when the sample observations are collected in such a way that the independent variables are correlated (for example, if age, height, and weight are collected on children of varying ages, each variable has a correlation with the others).

SigmaPlot can automatically detect multicollinear independent variables using the variance inflation factor. Click the **Other Diagnostics** tab in the **Options** dialog to view the **Variance Inflation Factor** option.

Flagging Multicollinear Data. Use the value in the Flag Values > edit box as a threshold for multicollinear variables. The default threshold value is 4.0, meaning that any value greater than 4.0 will be flagged as multicollinear. To make this test more sensitive to possible multicollinearity, decrease this value. To allow greater correlation of the independent variables before flagging the data as multicollinear, increase this value.

When the variance inflation factor is large, there are redundant variables in the regression model, and the parameter estimates may not be reliable. Variance inflation factor values above 4 suggest possible multicollinearity; values above 10 indicate serious multicollinearity.

What to Do About Multicollinearity. Sample-based multicollinearity can sometimes be resolved by collecting more data under other conditions to break up the correlation among the independent variables. If this is not possible, the regression equation is over parameterized and one or more of the independent variables must be dropped to eliminate the multicollinearity.

Structural multicollinearities can be resolved by centering the independent variable before forming the power or interaction terms.

Report Flagged Values Only. To only include only the points flagged by the influential point tests and values exceeding the variance inflation threshold in the report, make sure the Report Flagged Values Only check box is selected. Clear this option to include all influential points in the report.

What to Do About Influential Points. Influential points have two possible causes:

- There is something wrong with the data point, caused by an error in observation or data entry.
- The model is incorrect.

If a mistake was made in data collection or entry, correct the value. If you do not know the correct value, you may be able to justify deleting the data point. If the model appears to be incorrect, try regression with different independent variables, or a Nonlinear Regression.

Power

Click the **Other Diagnostics** tab in the options dialog to view the Power options. If Other Diagnostic is hidden, click the right pointing arrow to the right of the tabs to move it into view. Use the left pointing arrow to move the other tabs back into view.

The power of a regression is the power to detect the observed relationship in the data. The alpha (α) is the acceptable probability of incorrectly concluding there is a relationship.

Select **Power** to compute the power for the stepwise linear regression data. Change the alpha value by editing the number in the **Alpha Value** box. The suggested value is $\alpha = 0.05$. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant relationship when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant relationship, but a greater possibility of concluding there is no relationship when one exists. Larger values of α make it easier to conclude that there is a relationship, but also increase the risk of reporting a false positive.

Running a Stepwise Linear Regression

To run a Stepwise Regression you need to select the data to test. Use the **Select Data** panel of the Test Wizard to select the worksheet columns with the data you want to test.

To run a Stepwise Regression:

1. If you want to select your data before you run the regression, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Regression > Stepwise > Forward

or

Regression > Stepwise > Backward

The **Select Data** panel of the Test Wizard appears. If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog prompts you to pick your data.

4. To assign the desired worksheet columns to the **Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Dependent** and **Independent** drop-down list.

The first selected column is assigned to the **Dependent Variable** row in the **Selected Columns** list, and the second column is assigned to the **Independent Variable** row. The title of selected columns appears in each row. You are only prompted for one dependent and one independent variable column.

5. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
6. Click **Finish** to run the regression. If you elected to test for normality, constant variance, and/or independent residuals, SigmaPlot performs the tests for normality (Shapiro-Wilk or Kolmogorov-Smirnov), constant variance, and independent residuals. If your data fail either of these tests, SigmaPlot warns you. When the test is complete, the report appears displaying the results of the Stepwise Regression.

If you are performing a regression using one order only, and selected to place predicted values, residuals, and/or other test results in the worksheet, they are placed in the specified data columns and are labeled by content and source column.



Note: Worksheet results can only be obtained using order only stepwise regression.

Interpreting Stepwise Regression Results

The report for both Forward and Backward Stepwise Regression displays the variables that were entered or removed for that step, the regression coefficients, an ANOVA table, and information about the variables in and not in the

model. Regression diagnostics, confidence intervals, and predicted values are listed for the final regression model if these options were selected in the **Options for Forward or Backward Regression** dialog box.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

F-to-Enter, F-to-Remove

This is the worksheet column used as the dependent variable in the regression computation.

These are the F values specified in the Options for Stepwise Regression dialog boxes.

F-to-Enter. The F-to-Enter value controls which independent variables are entered into the regression equation during forward stepwise regression, or replaced after each step during backwards stepwise regression. It is the minimum incremental F value associated with an independent variable before it can be entered into the regression equation.

All independent variables with incremental F values above the F-to-Enter value are added to the model. The suggested F-to-Enter value is 4.0.

F-to-Remove. The F-to-Remove value controls which independent variables are deleted from the regression equation during Backwards Stepwise Regression, or removed after each step in Forward Stepwise Regression. It is the maximum incremental F value associated with an independent variable before it can be removed from the regression equation.

All independent variables with incremental F values below the F-to-Remove value are deleted from the model. The suggested F-to-Remove value is 3.9.

Step

The step number, variable added or removed, R , R^2 and the adjusted R^2 for the equation, and standard error of the estimate are all listed under this heading.

R and R Squared. R , the multiple correlation coefficient, and R^2 , the coefficient of determination for Stepwise Regression, are both measures of how well the regression model describes the data. R values near 1 indicate that the equation is a good description of the relation between the independent and dependent variables.

R equals 0 when the values of the independent variable does not allow any prediction of the dependent variables, and equals 1 when you can perfectly predict the dependent variables from the independent variables.

Adjusted R Squared. The adjusted R^2 , R^2_{adj} , is also a measure of how well the regression model describes the data, but takes into account the number of independent variables, which reflects the degrees of freedom. Larger R^2_{adj} values (nearer to 1) indicate that the equation is a good description of the relation between the independent and dependent variables.

Standard Error of the Estimate. The standard error of the estimate s_{yx} is a measure of the actual variability about the regression plane of the underlying population. The underlying population generally falls within about two standard errors of the observed sample. This statistic is displayed for the results of each step.

Analysis of Variance (ANOVA) Table

The ANOVA (analysis of variance) table lists the ANOVA statistics for the regression and the corresponding F value for each step.

SS (Sum of Squares). The sum of squares are measures of variability of the dependent variable.

- The sum of squares due to regression measures the difference of the regression plane from the mean of the dependent variable
- The residual sum of squares is a measure of the size of the residuals, which are the differences between the observed values of the dependent variable and the values predicted by regression model

DF (Degrees of Freedom). Degrees of freedom represent the number observations and variables in the regression equation.

- The regression degrees of freedom is a measure of the number of independent variables
- The residual degrees of freedom is a measure of the number of observations less the number of terms in the equation.

MS (Mean Square). The mean square provides two estimates of the population variances. Comparing these variance estimates is the basis of analysis of variance.

The mean square regression is a measure of the variation of the regression from the mean of the dependent variable, or

$$\frac{\text{sum of squares due to regression}}{\text{regression of degrees of freedom}} = \frac{SS_{\text{reg}}}{DF_{\text{reg}}} = MS_{\text{reg}}$$

The residual mean square is a measure of the variation of the residuals about the regression plane, or

$$\frac{\text{residual sum of squares}}{\text{residual degrees of freedom}} = \frac{SS_{\text{res}}}{DF_{\text{res}}} = MS_{\text{res}}$$

The residual mean square is also equal to s_{yx}^2 .

F Statistic. The F test statistic gauges the contribution of the independent variables in predicting the dependent variable. It is the ratio

$$\frac{\text{regression variation from the dependent variable mean}}{\text{residual variation about the regression curve}} = \frac{MS_{\text{reg}}}{MS_{\text{res}}} = F \text{ overall}$$

If F is a large number, you can conclude that the independent variables contribute to the prediction of the dependent variable (for example, at least one of the coefficients is different from zero, and the "unexplained variability" is smaller than what is expected from random sampling variability of the dependent variable about its mean). If the F ratio is around 1, you can conclude that there is no association between the variables (for example, the data is consistent with the null hypothesis that all the samples are just randomly distributed).

P Value. The P value is the probability of being wrong in concluding that there is an association between the dependent and independent variables (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on F). The smaller the P value, the greater the probability that there is an association.

Traditionally, you can conclude that the independent variable can be used to predict the dependent variable when $P < 0.05$.

Variables in Model

Information about the independent variables used in the regression equation for the current step are listed under this heading. The value of the variable coefficients, standard errors, the F-to-Remove, and the corresponding P value for the F-to-Remove are listed. These statistics are displayed for each step. An asterisk (*) indicates variables that were forced into the model.

Coefficients. The value for the constant and coefficients of the independent variables for the regression model are listed.

Standard Error . The standard errors are estimates of the regression coefficients (analogous to the standard error of the mean). The true regression coefficients of the underlying population generally fall within about two standard errors of the observed sample coefficients. Large standard errors may indicate multicollinearity.

F-to-Enter. The F-to-Enter gauges the increase in predicting the dependent variable gained by adding the independent variable to the regression equation. It is the ratio

regression variation
from the dependent
variable mean associated
with adding x when
 x_1, \dots, x_{j-1} are already
in the equation
residual variation about the $\frac{MS_{x_1, x_2, \dots, x_{j-1}}}{MS_{res|x_1, \dots, x_j}} = F_{enter}$
regression containing x_1, \dots, x_j

If the F-to-Enter for a variable is larger than the F-to-Enter cutoff specified with the Stepwise Regression options, the variable remains in or is added back to the equation.

 **Note:** The F-to-Remove value is the cutoff that determines if a variable is removed from or stays out of the equation.

P Value. P is the P value calculated for the F-to-Enter value. The P value is the probability of being wrong in concluding that adding the independent variable contributes to predicting the dependent variable (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on F-to-Enter). The smaller the P value, the greater the probability that adding the variable contributes to the model.

Traditionally, you can conclude that the independent variable can be used to predict the dependent variable when $P < 0.05$.

Variables not in Model

The variables not entered or removed from the model are listed under this heading, along with their corresponding F-to-Remove and P values.

F-to-Remove. The F-to-Remove gauges the increase in predicting the dependent variable gained by removing the independent variable from the regression equation.

If the F-to-Remove for a variable is larger than the F-to-Remove cutoff specified with the stepwise regression options, the variable is removed from or stays out of the equation.

 **Remember:** It is the F-to-Enter value that determines which variable is reentered into or remains in the equation.

P Value. P is the P value calculated for the F-to-Remove value. The P value is the probability of being wrong in concluding that removing the independent variable contributes to predicting the dependent variable (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on F-to-Enter). The smaller the P value, the greater the probability that removing the variable contributes to the model.

Traditionally, you can conclude that the independent variable can be used to predict the dependent variable when $P < 0.05$.

PRESS Statistic

PRESS, the Predicted Residual Error Sum of Squares, is a measure of how well a regression model predicts the observations.

The PRESS statistic is computed by summing the squares of the prediction errors (the differences between predicted and observed values) for each observation, with that point deleted from the computation of the estimated regression model.

One important use of the PRESS statistics is for model comparison. If several different regression models are applied to the same data, the one with the smallest PRESS statistic has the best predictive capability.

Durbin-Watson Statistic

The Durbin-Watson statistic is a measure of correlation between the residuals. If the residuals are not correlated, the Durbin-Watson statistic will be 2; the more this value differs from 2, the greater the likelihood that the residuals are correlated. This results appears if it was selected in the Options for Stepwise Regression dialog box.

Regression assumes that the residuals are independent of each other; the Durbin-Watson test is used to check this assumption. If the Durbin-Watson value deviates from 2 by more than the value set in the Options for Stepwise Regression dialog, a warning appears in the report. The suggested trigger value is a difference of more than 0.50, for example, when the Durbin-Watson statistic is less than 1.5 or greater than 2.5.

Normality Test

The Normality test result displays whether the data passed or failed the test of the assumption that the source population is normally distributed around the regression, and the P value calculated by the test. All regression requires a source population to be normally distributed about the regression. When this assumption may be violated, a warning appears in the report. This result appears unless you disabled normality testing in the Options for Best Subset Regression dialog box.

Failure of the normality test can indicate the presence of outlying influential points or an incorrect regression model.

Constant Variance Test

The constant variance test result displays whether or not the data passed or failed the test of the assumption that the variance of the dependent variable in the source population is constant regardless of the value of the independent variable, and the P value calculated by the test. When the constant variance assumption may be violated, a warning appears in the report.

If you receive this warning, you should consider trying a different model (for example, one that more closely follows the pattern of the data), or transforming the independent variable to stabilize the variance and obtain more accurate estimates of the parameters in the regression equation.

Power

This result is displayed if you selected this option in the Options for Stepwise Regression dialog box.

The power, or sensitivity, of a regression is the probability that the model correctly describes the relationship of the variables, if there is a relationship.

Regression power is affected by the number of observations, the chance of erroneously reporting a difference α (alpha), and the slope of the regression.

Alpha. Alpha (α) is the acceptable probability of incorrectly concluding that the model is correct. An α error is also called a Type I error (a Type I error is when you reject the hypothesis of no association when this hypothesis is true).

The α value is set in the Power Options dialog box; the suggested value is $\alpha = 0.05$ which indicates that a one in twenty chance of error is acceptable. Smaller values of α result in stricter requirements before concluding the model is correct, but a greater possibility of concluding the model is bad when it is really correct (a Type II error). Larger values of α make it easier to conclude that the model is correct, but also increase the risk of accepting a bad model (a Type I error).

Regression Diagnostics

The regression diagnostic results display only the values for the predicted and residual results selected in the Options for Stepwise Regression dialog. All results that qualify as outlying values are flagged with a $<$ symbol. The trigger values to flag residuals as outliers are set in the Options for Stepwise Regression dialog box.

If you selected Report Cases with Outliers Only, only those observations that have one or more residuals flagged as outliers are reported; however, all other results for that observation are also displayed.

Predicted Values. This is the value for the dependent variable predicted by the regression model for each observation. If these values were saved to the worksheet, they may be used to plot the regression using SigmaPlot.

Residuals. These are the raw residuals, the difference between the predicted and observed values for the dependent variables.

Standardized Residuals. The standardized residual is the raw residual divided by the standard error of the estimate.

If the residuals are normally distributed about the regression, about 66% of the standardized residuals have values between -1 and +1, and about 95% of the standardized residuals have values between -2 and +2. A larger standardized residual indicates that the point is far from the regression; the suggested value flagged as an outlier is 2.5.

Studentized Residuals. The Studentized residual is a standardized residual that also takes into account the greater confidence of the data points in the "middle" of the data set. By weighting the values of the residuals of the extreme data points (those with the lowest and highest independent variable values), the Studentized residual is more sensitive than the standardized residual in detecting outliers.

Both Studentized and Studentized deleted residuals that lie outside a specified confidence interval for the regression are flagged as outlying points: the suggested confidence value is 95%.

This residual is also known as the internally Studentized residual, because the standard error of the estimate is computed using all data.

Studentized Deleted Residual. The Studentized deleted residual, or externally Studentized residual, is a Studentized residual which uses the standard error of the estimate $s_{yx(-i)}$, computed by deleting the data point associated with the residual. This reflects the greater effect of outlying points by deleting the data point from the variance computation.

Both Studentized and Studentized deleted residuals that lie outside a specified confidence interval for the regression are flagged as outlying points; the suggested confidence value is 95%.

The Studentized deleted residual is more sensitive than the Studentized residual in detecting outliers, since the Studentized deleted residual results in much larger values for outliers than the Studentized residual.

Influence Diagnostics

The influence diagnostic results display only the values for the results selected in the Options dialog under the Other Diagnostics tab. All results that qualify as outlying values are flagged with a < symbol. The trigger values to flag data points as outliers are also set in the Options dialog under the Other Diagnostics tab.

If you selected Report Cases with Outliers Only, only observations that have one or more observations flagged as outliers are reported; however, all other results for that observation are also displayed.

Cook's Distance. Cook's distance is a measure of how great an effect each point has on the estimates of the parameters in the regression equation. It is a measure how much the values of the regression equation would change if that point is deleted from the analysis.

Values above 1 indicate that a point is possibly influential. Cook's distances exceeding 4 indicate that the point has a major effect on the values of the parameter estimates. Points with Cook's distances greater than the specified value are flagged as influential; the suggested value is 4.

Leverage. Leverage values identify potentially influential points. Observations with leverages a specified factor greater than the expected leverages are flagged as potentially influential points; the suggested value is 2.0 times the expected leverage.

The expected leverage of a data point is $\frac{(k+n)}{n}$ where there are k independent variables and n data points.

Because leverage is calculated using only the dependent variable, high leverage points tend to be at the extremes of the independent variables (large and small values), where small changes in the independent variables can have large effects on the predicted values of the dependent variable.

DFFITS. The DFFITS statistic is a measure of the influence of a data point on regression prediction. It is the number of estimated standard errors the predicted value for a data point changes when the observed value is removed from the data set before computing the regression coefficients.

Predicted values that change by more than the specified number of standard errors when the data point is removed are flagged as influential: the suggested value is 2.0 standard errors.

Confidence Intervals

These results are displayed if you selected them in the Options for Stepwise Regression dialog. If the confidence interval does not include zero, you can conclude that the coefficient is different than zero with the level of confidence

specified. This can also be described as $P < \alpha$ (alpha), where α is the acceptable probability of incorrectly concluding that the coefficient is different than zero, and the confidence interval is $100(1 - \alpha)$.

The specified confidence level can be any value from 1 to 99; the suggested confidence level for both intervals is 95%.

Pred (Predicted Values). This is the value for the dependent variable predicted by the regression model for each observation.

Mean. The confidence interval for the regression gives the range of variable values computed for the region containing the true relationship between the dependent and independent variables, for the specified level of confidence.

Obs (Observations). The confidence interval for the population gives the range of variable values computed for the region containing the population from which the observations were drawn, for the specified level of confidence.

Stepwise Regression Report Graphs

You can generate up to five graphs using the results from a Simple Linear Regression. They include a:

- [Histogram of the residuals.](#)
- [Scatter plot of the residuals.](#)
- [Bar chart of the standardized residuals.](#)
- [Normal probability plot of residuals.](#)
- [Line/scatter plot of the regression with confidence and prediction intervals.](#)
- [3D scatter plot of the residuals.](#)

Creating Stepwise Regression Report Graphs

To generate a graph of Stepwise Regression report data:

1. With the report in view, click the **Report** tab.
2. In the **Result Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying the types of graphs available for the **Stepwise Regression** results.

3. Select the type of graph you want to create from the **Graph Type** list, then click **OK**. For more information, see [Report Graphs](#) on page 373.

The specified graph appears in a graph window or in the report. For more information, see [Report Graphs](#) on page 373.

Best Subsets Regression

Use Linear Best Subsets Regression when you:

- Need to predict a trend in the data, or predict the value of one variable from the values of one or more other variables, by fitting a line or plane (or hyperplane) through the data.
- Do not know which independent variables contribute to the prediction of the dependent variable, and you want to find the subsets of independent variables that best contribute to predicting the dependent variable.

The independent variable is the known, or predicted, variable. When the independent variable is varied, a corresponding value for the dependent, or response, variable is produced.

If you already know which independent variables to use, use Multiple Linear Regression. If you want to select the equation model by incrementally adding or deleting variables from the model, use Stepwise Regression. If the relationship is not a straight line or plane, use Polynomial or Nonlinear Regression.

About Best Subset Regression

Best Subsets Regression is a technique for selecting variables in a multiple linear regression by systematically searching through the different combinations of the independent variables and selecting the subsets of variables that best contribute to predicting the dependent variable.

Best Subset Regression assumes an association between the independent and dependent variables that fits the general equation for a multidimensional plane: $y=b_0+b_1x_1+b_2x_2+b_3x_3+\dots+b_kx_k$

where y is the dependent variable, $x_1, x_2, x_3, \dots, x_k$ are the independent variables, and $b_1, b_2, b_3, \dots, b_k$ are the regression coefficients. As the values for x_1 vary, the corresponding value for y either increases or decreases. Best subsets regression searches for those combinations of the independent variables that give the "best" prediction of the dependent variable. There are several criteria for "best," and the results depend on which criterion you select. These criteria are specified in the **Options for Best Subset Regression** dialog box.

No predicted values, residuals, graphs, or other results are produced with a best subsets regression. To view results, note which independent variables were used for the desired model, then perform a multiple linear regression using only those independent variables.

Best Subsets Criteria

There are three statistics that can be used to evaluate which subsets of variables best contribute to predicting the dependent variable.

R Squared. R^2 , the coefficient of determination for multiple regression, is a measure of how well the regression model describes the data. The larger the value of R^2 , the better the model predicts the dependent variable.

However, the number of variables used in the equation is not taken into account. Consequently, equation with more variables will always have higher R^2 values, whether or not the additional variables really contribute to the prediction.

Adjusted R Squared. The adjusted R^2 , R^2_{adj} , is a measure of how well the regression model describes the data based on R^2 , but takes into account the number of independent variables.

Mallows. C_p is a gauge of the size of the bias introduced into the estimate of the dependent variable when independent variables are omitted from the regression equation, as computed from the number of parameters plus a measure of the difference between the predicted and true population means of the dependent variable.

The optimal value of C_p is equal to the number of parameters (the independent variables used in the subset plus the constant), or: $C_p=p=k+1$

where p is the number of parameters and k is the number of independent variables.

The closer the value of C_p is to the number of parameters, the less likely a relevant variable was omitted. Note that the fully specified model will always have a $C_p=p$.

Performing a Best Subset Regression

To perform a Best Subset Regression:

1. [Enter or arrange your data in the worksheet.](#)
2. If desired, set the Best Subset Regression options.
3. Click the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Regression > Best Subsets
5. [View and interpret the Best Subset Regression report.](#)

Arranging Best Subset Regression Data

Place the data for the observed dependent variable in a single column and the corresponding data for the independent variables in one or more columns. Rows containing missing values are ignored, and the columns must be of equal length.

Setting Best Subset Regression Options

Use the Best Subset Regression options to

- Specify the criterion to use to predict the dependent variable and the number of subset used in the equation.
- Enable the variance inflation factor to identify potential difficulties with the regression parameter estimates (multicollinearity).

To change Best Subset Regression options:

- If you are going to run the test after changing test options, and want to select your data before you run the test, drag the pointer over your data.
 - Select **Best Subset Regression** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.
 - Click **Options**.
- The **Options for Best Subset Regression** dialog box appears with the Criterion tab in view.
- Options settings are saved between SigmaPlot sessions.
- To continue the test, click **Run Test**.
 - To accept the current settings and close the dialog box, click **OK**.

Options for Best Subset Regression: Criterion

Use the Best Criterion option to select the criterion used to determine the best subsets and the Number of Subsets option to specify the number of subsets to list.

Best Criterion. Select the criterion to determine the best subsets from this drop-down list.

- Mallows.** Select **Mallows Cp** from the Best Criterion drop-down list to use a gauge of the bias introduced when variables are omitted to quickly screen large numbers of potential variables and produce a few subsets that include only the relevant variables. The number of subsets listed is equal to the number set with the Number of Subsets option.
- R Squared.** Select **R Squared** (R^2) from the Best Criterion drop-down list to use the largest coefficient of determination to find the best fitting subset. R^2 contains no information on the number of variables used, so subsets are listed for each number of possible variables (for example, one independent variable, two variables, and so on, up to all variables selected). The maximum number of subsets listed for each number of possible variables is equal to the Number of Subsets option.
- Adjusted R Squared.** Select **Adjusted R Squared** (Adjusted R^2) from the Best Criterion drop-down list to use the largest values to select the best regressions. R^2_{adj} takes into account the loss of degrees of freedom when additional independent variables are added to the regression equation. The number of subsets listed is equal to the number set with the Number of Subsets option.

Number of Subsets. Use this option to specify the number of most contributing variable groups to list by entering the desired value in the Number of Subsets edit box. For C_p this is the total number of subsets. For R^2 , this is the number of variable subsets listed for each number of independent variables in the equation.

Variance Inflation Factor. Use Variance Inflation Factor option to measure the multicollinearity of the independent variables, or the linear combination of the independent variables in the fit.

Regression procedures assume that the independent variables are statistically independent of each other, for example, that the value of one independent variable does not affect the value of another. However, this ideal situation rarely occurs in the real world. When the independent variables are correlated, or contain redundant information, the estimates of the parameters in the regression model can become unreliable.

The parameters in regression models quantify the theoretically unique contribution of each independent variable to predicting the dependent variable. When the independent variables are correlated, they contain some common information and "contaminate" the estimates of the parameters. If the multicollinearity is severe, the parameter estimates can become unreliable.

There are two types of multicollinearity.

- Structural Multicollinearity.** Structural multicollinearity occurs when the regression equation contains several independent variables which are functions of each other. The most common form of structural multicollinearity

occurs when a polynomial regression equation contains several powers of the independent variable. Because these powers (for example, x , x^2 , and so on) are correlated with each other, structural multicollinearity occurs. Including interaction terms in a regression equation can also result in structural multicollinearity.

- **Sample-Based Multicollinearity.** Sample-based multicollinearity occurs when the sample observations are collected in such a way that the independent variables are correlated (for example, if age, height, and weight are collected on children of varying ages, each variable has a correlation with the others).

Report Flagged Values Only. To only include only the points flagged by the influential point tests and values exceeding the variance inflation threshold in the report, make sure the Report Flagged Values Only check box is selected. Clear this option to include all influential points in the report. For more information, see [Flagging Multicollinear Data](#) on page 292.

Flagging Multicollinear Data

Use the value in the Flag Values > edit box as a threshold for multicollinear variables. The default threshold value is 4.0, meaning that any value greater than 4.0 will be flagged as multicollinear. To make this test more sensitive to possible multicollinearity, decrease this value. To allow greater correlation of the independent variables before flagging the data as multicollinear, increase this value.

When the variance inflation factor is large, there are redundant variables in the regression model, and the parameter estimates may not be reliable. Variance inflation factor values above 4 suggest possible multicollinearity; values above 10 indicate serious multicollinearity.

What to Do About Multicollinearity. Sample-based multicollinearity can sometimes be resolved by collecting more data under other conditions to break up the correlation among the independent variables. If this is not possible, the regression equation is over parameterized and one or more of the independent variables must be dropped to eliminate the multicollinearity.

Structural multicollinearities can be resolved by centering the independent variable before forming the power or interaction terms.

Running a Best Subset Regression

To run a Best Subset Regression, you need to select the data to test. You use the **Select Data** panel of the Test Wizard to select the worksheet columns with the data you want to test.

To run a Best Subset Regression:

1. If you want to select your data before you run the regression, drag the pointer over your data.
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Regression > Best Subsets

The **Select Data** panel of the Test Wizard appears. If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog prompts you to pick your data.

4. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Dependent** and **Independent** drop-down list.

The first selected column is assigned to the **Dependent Variable** row in the **Selected Columns** list, and the second column is assigned to the **Independent Variable** row. The title of selected columns appears in each row. You are only prompted for one dependent and one independent variable column.

5. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
6. Click **Finish** to run the regression. The Best Subset Regression is performed. When the test is complete, the Best Subset regression report appears.

- i** **Tip:** No predicted values, residuals and other test results are computed or placed in the worksheet. To view results for models, note which independent variables were used for that model, then perform a Multiple Linear Regression using only those independent variables.

Interpreting Best Subset Regression Results

A Best Subsets Regression report lists a summary table of the "best" criteria statistics for all variable subsets, along with the error mean square and the specific member variables of the subset. Detailed results for each subset regression equation are then listed individually.

Note that the number of subsets listed is determined by the number of subsets selected in the Options for Best Subsets Regression dialog, and the criterion used to select the best subsets.

- If you used R^2 , the maximum number of subsets reported for each number of variables included is the number set in the Best Subsets Regression Options dialog box.
- If you used R^2 or C_p , the number of subset results reported is the number set in the Options for Best Subsets Regression dialog box.

Tip: You cannot generate report graphs for Best Subsets Regression. To view a graph, perform a Multiple Linear Regression using the variables in the subset(s) of interest, and graph those results. For more information, see [Multiple Linear Regression](#) on page 237.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

Summary Table

Variables. The variables included in the subset are noted by asterisks (*) which appear below the variable symbols on the right side of the table.

Mallows. C_p is a gauge of the bias introduced into the estimate of the dependent variable when independent variables are omitted from the regression equation. The optimal value of C_p is equal to the number of parameters (the independent variables used in the subset plus the constant), or $C_p=p=k+1$

where p is the number of parameters and k is the number of independent variables. The closer the value of C_p is to the number of parameters, the less likely a relevant variable was omitted. Subsets with low orders that also have C_p values close to $k + 1$ are good candidates for the best subset of variables.

R Squared. R^2 , the coefficient of determination for multiple regression, is a measure of how well the regression model describes the data. The closer the value of R^2 to 1, the better the model predicts the dependent variable. However, because the number of variables used is not taken into account, higher order subsets will always have higher R^2 values, whether or not the additional variables really contribute to the prediction.

Adjusted R Squared. The adjusted R^2 , R^2_{adj} , is a measure of how well the regression model describes the data based on R^2 , but takes into account the number of independent variables.

Larger values (nearer to 1) indicate that the equation is a good description of the relation between the independent and dependent variables. Note that the subset that includes all variables always has a $C_p=p$.

MSerr (Error Mean Square). The error mean square (residual, or within groups):

$$\frac{\text{error sum of squares}}{\text{error degrees of freedom}} = \frac{SS_{\text{error}}}{DF_{\text{error}}} = MS_{\text{error}}$$

is an estimate of the variability in the underlying population, computed from the random component of the observations.

Residual Sum of Squares. The residual sum of squares is a measure of the size of the residuals, which are the differences between the observed values of the dependent variable and the values predicted by regression model.

Subsets Results

Tables of statistical results are listed for each regression equation identified in the summary table.

Coefficient. The value for the constant and coefficients of the independent variables for the regression model are listed.

Std Err (Standard Error). The standard errors are estimates of these regression coefficients (analogous to the standard error of the mean). The true regression coefficients of the underlying population generally fall within about two standard errors of the observed sample coefficients. Large standard errors may indicate multicollinearity. These values are used to compute t for the regression coefficients.

t Statistic. The t statistic tests the null hypothesis that the coefficient of each independent variable is zero, that is, the independent variable does not contribute to predicting the dependent variable. t is the ratio of the regression coefficient to its standard error, or:

$$t = \frac{\text{regression coefficient}}{\text{standard error of regression coefficient}}$$

You can conclude from "large" t values that the independent variable(s) can be used to predict the dependent variable (for example, that the coefficient is not zero).

P Value. P is the P value calculated for t . The P value is the probability of being wrong in concluding that there is a true association between the variables (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on t). The smaller the P value, the greater the probability that the independent variable helps predict the dependent variable.

Traditionally, you can conclude that the independent variable can be used to predict the dependent variable when $P < 0.05$.

VIF (Variance Inflation Factor). The variance inflation factor is a measure of multicollinearity. It measures the "inflation" of a regression parameter (coefficient) for an independent variable due to redundant information in other independent variables.

If the variance inflation factor is at or near 1.0, there is no redundant information in the other independent variables. If the variance inflation factor is much larger, there are redundant variables in the regression model, and the parameter estimates may not be reliable.

This result appears unless it was disabled in the **Options for Best Subset Regression** dialog box.

Pearson Product Moment Correlation

Use Pearson Product Moment Correlation when:

- You want to measure the strength of the association between pairs of variables without regard to which variable is dependent or independent.
- You want to determine if the relationship, if any, between the variables is a straight line.
- The residuals (distances of the data points from the regression line) are normally distributed with constant variance.

The Pearson Product Moment Correlation coefficient is the most commonly used correlation coefficient.

If you want to predict the value of one variable from another, use Simple or multiple Linear Regression. If you need to find the correlation of data measured by rank or order, use the nonparametric Spearman Rank Order Correlation.

Pearson Product Moment Correlation Report Graph

The Pearson Moment Correlation matrix is a series of scatter graphs that plot the associations between all possible combinations of variables.

The first row of the matrix represents the first set of variables or the first column of data, the second row of the matrix represents the second set of variables or the second data column, and the third row of the matrix represents the third set of variables or third data column. The X and Y data for the graphs correspond to the column and row of the graph in the matrix.

For example, the X data for the graphs in the first row of the matrix is taken from the second column of tested data, and the Y data is taken from the first column of tested data. The X data for the graphs in the second row of the matrix is taken from the first column of tested data, and the Y data is taken from the second column of tested data. The X

data for the graphs in the third row of the matrix is taken from the second column of tested data, and the Y data is taken from the third column of tested data, and so on. The number of graph rows in the matrix is equal to the number of data columns being tested.

Creating the Pearson Product Moment Report Graph

To generate a report graph of Pearson Product Moment report data:

1. With the Pearson Product Moment report in view, click the **Report** tab.

2. In the **Result Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying a Scatter Matrix graph.

3. Click **OK**.

[The selected graph appears in a graph window.](#)

About the Pearson Product Moment Correlation Coefficient

When an assumption is made about the dependency of one variable on another, it affects the computation of the regression line. Reversing the assumption of the variable dependencies results in a different regression line.

The Pearson Product Moment Correlation coefficient does not require the variables to be assigned as independent and dependent. Instead, only the strength of association is measured.

Pearson Product Moment Correlation is a parametric test that assumes the residuals (distances of the data points from the regression line) are normally distributed with constant variance.

Computing the Pearson Product Moment Correlation Coefficient

To compute the Pearson Product Moment Correlation coefficient:

1. [Enter or arrange your data appropriately in the data worksheet.](#)

2. Click the **Analysis** tab.

3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Correlation > Pearson Product Moment

4. [View and interpret the Pearson Product Moment Report.](#)

5. [Pearson Product Correlation Report Graph](#) on page 298

6. Run the test by selecting the worksheet columns with the data you want to test using the **Select Data** panel of the Test Wizard.

Arranging Pearson Product Moment Correlation Data

Place the data for each variable in a column. You must have at least two columns of variables, with a maximum of 64 columns. Observations containing missing values are ignored, including missing values created by columns of unequal length.

Setting Pearson Correlation Options

Use the Pearson Correlation options to:

- Set the assumption checking options for normality
- Set the display format for results in reports.

To change Pearson Correlation options:

1. If you are going to run the test after changing test options and want to select your data before you run the test, drag the pointer over your data.
2. Select **Pearson Correlation** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.

3. Click **Options**. The **Options for Pearson Correlation** dialog box appears with two tabs:

- **Assumption Checking.** Click the **Assumption Checking** tab to view the Normality options.
- **Results.** Click the **Results** tab to view formatting options for reports

Options settings are saved between SigmaPlot sessions.

4. To continue the test, click **Run Test**.

5. To accept the current settings and close the dialog box, click **OK**.

Options for Pearson Correlation: Assumption Checking

Hypothesis testing for zero correlation in Pearson Correlation assumes the data for each selected pair of variables have been sampled from populations that have a joint bivariate normal distribution. This assumption can be checked for each pairwise combination of the variables selected in the Test Wizard. If this assumption fails for one or more pairs of variables, then the corresponding P-values may be unreliable.

Normality. For Pearson Correlation, SigmaPlot uses either the **Henze-Zirkler** test or **Mardia's test for skewness and kurtosis**.

- The Henze-Zirkler test is more difficult conceptually as it measures a distance between the empirical distribution and the bivariate normal distribution using Fourier transforms. Mardia's test is simpler and more intuitive but is generally not as powerful as the Henze-Zirkler test. This means that the latter test can detect smaller deviations from the hypothesis that the joint distribution for the data is bivariate normal.
- **P Value to Reject.** Enter the corresponding P value in the P Value to Reject box. The P value determines the probability of being incorrect in concluding that a pair of variables does not have a joint bivariate normal distribution. If the P value computed by the test is greater than the P set here, the test passes.

To require a stricter adherence to normality, increase the P value. Because the parametric statistical methods are relatively robust in terms of detecting violations of the assumptions, the suggested value in SigmaPlot is 0.050. Larger values of P (for example, 0.100) require less evidence to conclude that data is not normal.

To relax the requirement of normality, decrease P. Requiring smaller values of P to reject the normality assumption means that you are willing to accept greater deviations from the theoretical normal distribution before you flag the data as non-normal. For example, a P value of 0.050 requires greater deviations from normality to flag the data as non-normal than a value of 0.100.



Note: Although the normality test is robust in detecting data from populations that are non-normal, there are extreme conditions of data distribution that this test cannot take into account; however, these conditions should be easily detected by simply examining the data without resorting to the automatic assumption test.

Options for Pearson Correlation: Results

Report Display Format. There are two options for formatting results in the report for normality and correlation for each pair of sampled data columns in the worksheet.

- **Matrix.** When this option is selected, all results for normality (if chosen) and correlation are arranged in a matrix format with the rows and columns labeled with the names of the variables that were selected in the Test Wizard. The results for each pair of variables are found by locating one variable's row and the other variable's column. This is the default format.
- **Table.** When this option is selected, the results are given in separate tables for normality and correlation. Each row of a table corresponds to a pair of the variables that were selected in the Test Wizard. This format is particularly useful if a large number of variables were selected.

Running a Pearson Product Moment Correlation

To run a Pearson Product Moment test, you need to select the data to test. The **Select Data** panel of the Test Wizard is used to select the worksheet columns with the data you want to test.

To run a Pearson Product Moment Correlation:

1. If you want to select your data before you run the regression, drag the pointer over your data.
2. Click the **Analysis** tab.

3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Correlation > Pearson Product Moment

The **Select Data** panel of the Test Wizard appears. If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog box prompts you to pick your data.

4. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Variable** drop-down list.

The selected columns are assigned to the **Variables** row in the **Selected Columns** list in the order they are selected from the worksheet. The title of selected columns appears in each row. You can select up to 64 variable columns. SigmaPlot computes the correlation coefficient for every possible pair.

5. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
6. Click **Finish**. The correlation coefficient is computed. When the test is complete, the Pearson Product Moment Correlation Coefficient report appears.

Interpreting Pearson Product Moment Correlation Results

The report for a Pearson Product Moment Correlation displays the correlation coefficient r , the P value for the correlation coefficient, and the number of data points used in the computation, for each pair of variables.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

Correlation Results

Correlation Coefficient

The correlation coefficient r quantifies the strength of the association between the variables. r varies between -1 and +1. A correlation coefficient near +1 indicates there is a strong positive relationship between the two variables, with both always increasing together. A correlation coefficient near -1 indicates there is a strong negative relationship between the two variables, with one always decreasing as the other increases. A correlation coefficient of 0 indicates no relationship between the two variables.

P Value

The P value is the probability of being wrong in concluding that there is a true association between the variables (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error). The smaller the P value, the greater the probability that the variables are correlated.

Traditionally, you can conclude that the independent variable can be used to predict the dependent variable when $P < 0.05$.

Number of Samples

This is the number of data points used to compute the correlation coefficient. This number reflects samples omitted because of missing values in one of the two variables used to compute each correlation coefficient.

Normality Results

Mardia's Test

There are two sets of results:

- **Skewness.** The skewness results contain of the values of bivariate skewness, a test statistic, and the P-value for the probability of being incorrect in rejecting the hypothesis that the joint distribution of a given variable pair is bivariate normal.

- **Kurtosis.** The kurtosis results contain bivariate kurtosis, a test statistic, and the corresponding P-value for the probability of being incorrect in rejecting the hypothesis that the joint distribution of a given variable pair is bivariate normal.

If you use the table format to display results, the last column shows if the bivariate normality test passed or failed for each variable pair.

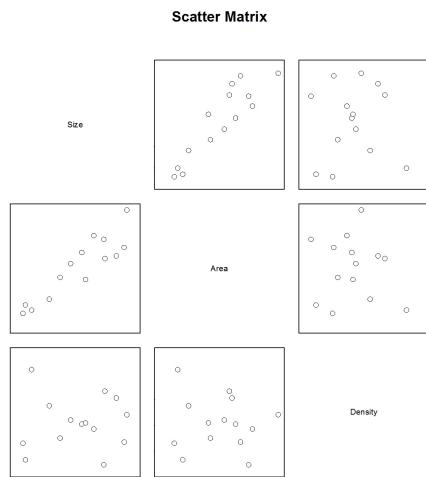
Henze-Zirkler

The results contain of the values of the test statistic and the P-value for the probability of being incorrect in rejecting the hypothesis that the joint distribution of a given variable pair is bivariate normal.

If you use the table format to display results, the last column shows if the bivariate normality test passed or failed for each variable pair.

Pearson Product Correlation Report Graph

The Pearson Moment Correlation matrix is a series of scatter graphs that plot the associations between all possible combinations of variables.



The first row of the matrix represents the first set of variables or the first column of data, the second row of the matrix represents the second set of variables or the second data column, and the third row of the matrix represents the third set of variables or third data column. The X and Y data for the graphs correspond to the column and row of the graph in the matrix.

For example, the X data for the graphs in the first row of the matrix is taken from the second column of tested data, and the Y data is taken from the first column of tested data. The X data for the graphs in the second row of the matrix is taken from the first column of tested data, and the Y data is taken from the second column of tested data. The X data for the graphs in the third row of the matrix is taken from the second column of tested data, and the Y data is taken from the third column of tested data, and so on. The number of graph rows in the matrix is equal to the number of data columns being tested.

Creating the Pearson Product Moment Report Graph

To generate a report graph of Pearson Product Moment report data:

1. With the Pearson Product Moment report in view, click the **Report** tab.
2. In the **Result Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying a Scatter Matrix graph.

3. Click **OK**.

The selected graph appears in a graph window.

Spearman Rank Order Correlation

Use Spearman Rank Order Correlation when:

- You want to measure the strength of association between pairs of variables without specifying which variable is dependent or independent.
- The residuals (distances of the data points from the regression line) are not normally distributed with constant variance.

If you want to assume that the value of one variable affects the other, use some form of regression. If you need to find the correlation of normally distributed data, use the parametric Pearson Product Moment Correlation.

About the Spearman Rank Order Correlation Coefficient

When an assumption is made about the dependency of one variable on another, it affects the computation of the regression line. Reversing the assumption of the variable dependencies results in a different regression line.

The Spearman Rank Order Correlation coefficient does not require the variables to be assigned as independent and dependent. Instead, only the strength of association is measured.

The Spearman Rank Order Correlation coefficient is computed by ranking all values of each variable, then computing the Pearson Product Moment Correlation coefficient of the ranks.

Spearman Rank Order Correlation is a nonparametric test that does not require the data points to be linearly related with a normal distribution about the regression line with constant variance.

Computing the Spearman Rank Order Correlation Coefficient

To compute the Spearman Rank Order Correlation coefficient:

1. [Enter or arrange your data appropriately in the worksheet.](#)
2. Click the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Correlation > Spearman Rank Order
4. [View and interpret the Spearman rank order correlation report.](#)
5. [Generate the report graph.](#)
6. Run the test.

Arranging Spearman Rank Order Correlation Coefficient Data

Place the data for each variable in a column. You must have at least two columns of variables, with a maximum of 64 columns. Observations containing missing values are ignored. However, rank order correlations require columns of equal length.

Running a Spearman Rank Order Correlation

To run a Spearman Rank Order Correlation test, you need to select the data to test. The **Select Data** panel of the Test Wizard is used to select the worksheet columns with the data you want to test and to specify how your data is arranged in the worksheet.

To run a Spearman Rank Order Correlation:

1. If you want to select your data before you run the regression, drag the pointer over your data.
2. Click the **Analysis** tab.

3. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Correlation > Spearman Correlation

The **Select Data** panel of the Test Wizard appears. If you selected columns before you chose the test, the selected columns appear in the column list. If you have not selected columns, the dialog box prompts you to pick your data.

4. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for Variable** drop-down list.

The selected columns are assigned to the **Variables** row in the **Selected Columns** list in the order they are selected from the worksheet. The title of selected columns appears in each row. You can select up to 64 variable columns. SigmaPlot computes the correlation coefficient for every possible pair.

5. To change your selections, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
6. Click **Finish**. The correlation coefficient is computed. When the test is complete, the Spearman Rank Order Correlation Coefficient report appears.

Interpreting Spearman Rank Correlation Results

The report for a Spearman Rank Order Correlation displays the correlation coefficient rs , the P value for the correlation coefficient, and the number of data points used in the computation, for each pair of variables. If Normality is selected in the Options for Spearman Correlation, then values are also displayed for the test statistic(s) and the P-value(s) for normality.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

Correlation Results

Spearman Correlation Coefficient rs. The Spearman correlation coefficient rs quantifies the strength of the association between the variables. rs varies between -1 and +1. A correlation coefficient near +1 indicates there is a strong positive relationship between the two variables, with both always increasing together. A correlation coefficient near -1 indicates there is a strong negative relationship between the two variables, with one always decreasing as the other increases. A correlation coefficient of 0 indicates no relationship between the two variables.

P Value. The P value is the probability of being wrong in concluding that there is a true association between the variables (for example, the probability of falsely rejecting the null hypothesis, or committing a Type I error). The smaller the P value, the greater the probability that the variables are correlated.

Traditionally, you can conclude that the independent variable can be used to predict the dependent variable when $P < 0.05$.

Number of Samples. This is the number of data points used to compute the correlation coefficient. This number reflects samples omitted because of missing values in one of the two variables used to compute each correlation coefficient.

Normality Results

Mardia's Test

There are two sets of results:

- **Skewness.** The skewness results contain of the values of bivariate skewness, a test statistic, and the P-value for the probability of being incorrect in rejecting the hypothesis that the joint distribution of a given variable pair is bivariate normal.
- **Kurtosis.** The kurtosis results contain bivariate kurtosis, a test statistic, and the corresponding P-value for the probability of being incorrect in rejecting the hypothesis that the joint distribution of a given variable pair is bivariate normal.

If you use the table format to display results, the last column shows if the bivariate normality test passed or failed for each variable pair.

Henze-Zirkler

The results contain of the values of the test statistic and the P-value for the probability of being incorrect in rejecting the hypothesis that the joint distribution of a given variable pair is bivariate normal.

If you use the table format to display results, the last column shows if the bivariate normality test passed or failed for each variable pair.

Spearman Rank Order Correlation Report Graph

The Spearman Rank Order Correlation matrix of scatter graphs is a series of scatter graphs that plot the associations between all possible combinations of variables.

The first row of the matrix represents the first set of variables or the first column of data, the second row of the matrix represents the second set of variables or the second data column, and the third row of the matrix represents the third set of variables or third data column. The X and Y data for the graphs correspond to the column and row of the graph in the matrix.

For example, the X data for the graphs in the first row of the matrix is taken from the second column of tested data, and the Y data is taken from the first column of tested data. The X data for the graphs in the second row of the matrix is taken from the first column of tested data, and the Y data is taken from the second column of tested data. The X data for the graphs in the third row of the matrix is taken from the second column of tested data, and the Y data is taken from the third column of tested data, and so on. The number of graph rows in the matrix is equal to the number of data columns being tested. For more information, see [Report Graphs](#) on page 373.

Creating the Spearman Correlation Report Graph

To generate the graph of Spearman Correlation report data:

1. With the Spearman Correlation report in view, click the **Report** tab.
2. In the **Result Graphs** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying a Scatter Matrix graph.

3. Click **OK**.

The selected graph appears in a graph window. For more information, see [Report Graphs](#) on page 373.

Deming Regression

Use Deming Regression when:

- You want to determine the slope and intercept of the best-fit line.

About Deming Regression

The main objective of Deming Regression is to determine the slope and intercept of the best-fit line when there are uncertainties in both the independent and dependent variables. If the slope of the line is denoted by b and the intercept denoted by a , then maximizing the likelihood of obtaining the given observations is equivalent to the following minimization problem:

$$\min_{a,b} \sum_k \frac{(a + bx_k - y_k)^2}{\sigma_{y_k}^2 + b^2 \sigma_{x_k}^2}$$

where, (x_k, y_k) is the observation, σ_{x_k} is the standard deviation of x_k , and σ_{y_k} is the standard deviation of y_k .

There are two types of Deming Regression: Simple Deming Regression and General Deming Regression. Use Simple Deming Regression when the data errors are constant among all measurements for each of the two variables. If the

two constant error values for the independent and dependent variables are equal to each other, then Simple Deming Regression is often called *Orthogonal Regression*.

Use General Deming Regression to allow arbitrary values for the error at each observation.

Performing a Deming Regression

To perform a Deming Regression:

1. [Enter or arrange your data in the worksheet](#).
2. If desired, [set the Deming Regression options](#).
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Regression > Deming
5. [Generate report graphs](#).
6. Run the test.

Arranging Deming Regression Data

Enter your Deming Regression data using two columns to represent pairs of observation, where the first coordinate refers to the independent variable and second coordinate refers to the dependent variable.

Setting Deming Regression Options

Report

Standard errors of parameters.

- **Apply correction factor estimated by the reduced chi-square** Select if all of the data errors, for both X and Y, are only known up to some common, but unknown, scaling factor σ . This scaling factor is estimated as the square root of the reduced chi-square statistic that results from the fit analysis. If the best-fit parameters for the intercept and the slope are given by a and b , respectively, and the number of observations is n , then the estimate is:

$$\sigma = \sqrt{\frac{1}{n-2} \sum_k \frac{(a + b x_k - y_k)^2}{\sigma_{y_k}^2 + b^2 \sigma_{x_k}^2}}$$

where (x_k, y_k) is the k^{th} observation; σ_{x_k} is the standard of the deviation of x_k ; and σ_{y_k} is the standard deviation of y_k .

This option has no effect on the computation of the best-fit parameter values, but does affect the computation of the standard errors of both parameters.

Add table of predicted means. Select to put a table of the predicted means (or true values) for every observation in both the X and Y variables in the report. This table also includes the values of the residuals between the observations and the predicted means.

Add parameter covariance matrix. Select to put the two-by-two variance-covariance matrix for the parameters in the report.

Confidence Intervals.

- **Confidence level.** Set the percent confidence level for the confidence intervals for the parameters and for the confidence bands for the regression line. The default value is 95%.

Graph

Create new graph. Select to create a result graph of Deming Regression after the report is created. By default, the result graph consists of the regression line and the raw data. Although this graph does not appear with error bars, you can add them later by modifying the graph.

Graph features.

- **Scatter plot of predicted means.** Select to place a scatter plot of the predicted means for each data point on top of the regression line. Data for the predicted means also appears in the worksheet.
- **Confidence Bands.** Select to add pair of curves on the graph that represent the lower and upper limit of confidence intervals for the predicted means at specific values of the independent variable over the range of the input data. Data for the confidence bands appears in the worksheet.

Running a Deming Regression

If you want to select your data before you run the test, drag the pointer over your data.

1. Click the **Analysis** tab, and then in the **SigmaStat** group, from the **Tests** drop-down list, select:

Regression > Deming

The **Deming Regression - Data Format** panel of the Regression Wizard appears prompting you to specify a data format.

2. Click **Finish** to run the t-test on the selected columns. After the computations are completed, the report appears.
3. Select either **XY Pair** or **XY Pair-Errors** from the **Data Format** drop-down list. For more information, see [Arranging Deming Regression Data](#) on page 302.
4. Click **Next** to pick the data columns for the test. If you selected columns before you the test, the selected columns appear in the **Selected Columns** list.
5. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for Data** drop-down list.

The first selected column is assigned to the first row in the **Selected Columns** list, and all successively selected columns are assigned to successive rows in the list. The title of selected columns appears in each row. For raw and indexed data, you are prompted to select two worksheet columns. For statistical summary data you are prompted to select three columns.

6. **To change your selections**, select the assignment in the list, then select new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
7. Click **Finish** to run the t-test on the selected columns. After the computations are completed, the report appears.

Interpreting Deming Regression Results

The Deming Regression report contains the numerical results of the regression analysis as well as information on the data source used for input. It begins with the usual header information (date, title, and data source) contained in reports for other statistical tests. The information in the report is divided into sections as follows:

Report Title. The report is titled Simple Deming Regression if the data format XY Pair was chosen in the wizard and titled General Deming Regression if the data format XY Pair-Errors was chosen.

Input Data Columns. The column selections for the X- and Y- measurements are first given. For Simple Deming Regression, the next report entries are the two constant standard deviations for the X- and Y- measurements. If General Deming Regression is used, then instead we give the column selections for the standard deviations of both the X and Y variables.

Data Summary. This section simply shows the total number of selected observations and the number of missing data rows. What constitutes a missing data row was discussed above in the description of the Deming Regression Wizard.

Regression Measures. There are four values that measure the strength of association between the two data variables and the error variance scaling used in the model.

- The *correlation coefficient* measures the linear correlation between the data for the independent and dependent variables. In Simple Deming Regression, a high enough correlation coefficient ($> .975$) is sometimes used as a criterion for using simple linear regression as an alternative to Deming Regression.
- The *chi-square statistic* is the optimal value of the sum that was minimized to obtain the best-fit parameters (equivalent to maximizing the log-likelihood function). If you provided exact (or nearly so) standard deviations of the data, then this statistic has an approximate chi-square distribution with $N-2$ degrees of freedom, where N is the number of pairs of observations.

- The *reduced chi-square statistic* is simply the value of the chi-square statistic divided by the degrees of freedom. It estimates the variance scaling for the data, assuming you selected the option to apply a scaling factor to the data errors.
- The *degrees of freedom* is the integer $N-2$, where N is the number of pairs of observations: $N-2 = (2N \text{ observations for } X \text{ and } Y) - (N \text{ linear constraints on the means for both } X \text{ and } Y) - (2 \text{ parameters})$.

Parameter Estimates. This is a table of the estimated values and statistics for the intercept and slope parameters for the best-fit line, and asymptotic values of certain statistics. The statistical values reported are the standard error of the parameters and their individual confidence intervals.

The statistics are affected not only by the data set, but by options that you selected. In the Options dialog box, you can choose to interpret the data errors either absolutely or relative to some scaling factor. Also, there is a setting in the SigmaPlot's configuration file (spw.ini) that allows you to choose between two estimation methods for the standard errors - York and Williamson's method which is similar to the so-called delta method and is based directly on the values of the observations, and another method based on maximum likelihood theory (MLE method) which is based on using the predicted means.

Covariance Matrix. This is a two-by-two matrix whose diagonal entries are the variances of the parameters and whose off-diagonal entry is the covariance between both parameters.

Hypothesis Testing. Two F-tests are used to test the hypotheses that the slope is 0 and that the slope is 1.

Predicted Means. This is a table of the predicted (estimated) means for the distributions from which the data is sampled. There are two predicted means given for the X and Y measurements in each observed data point. The residual difference between the measured value of each variable and its predicted mean is also given.

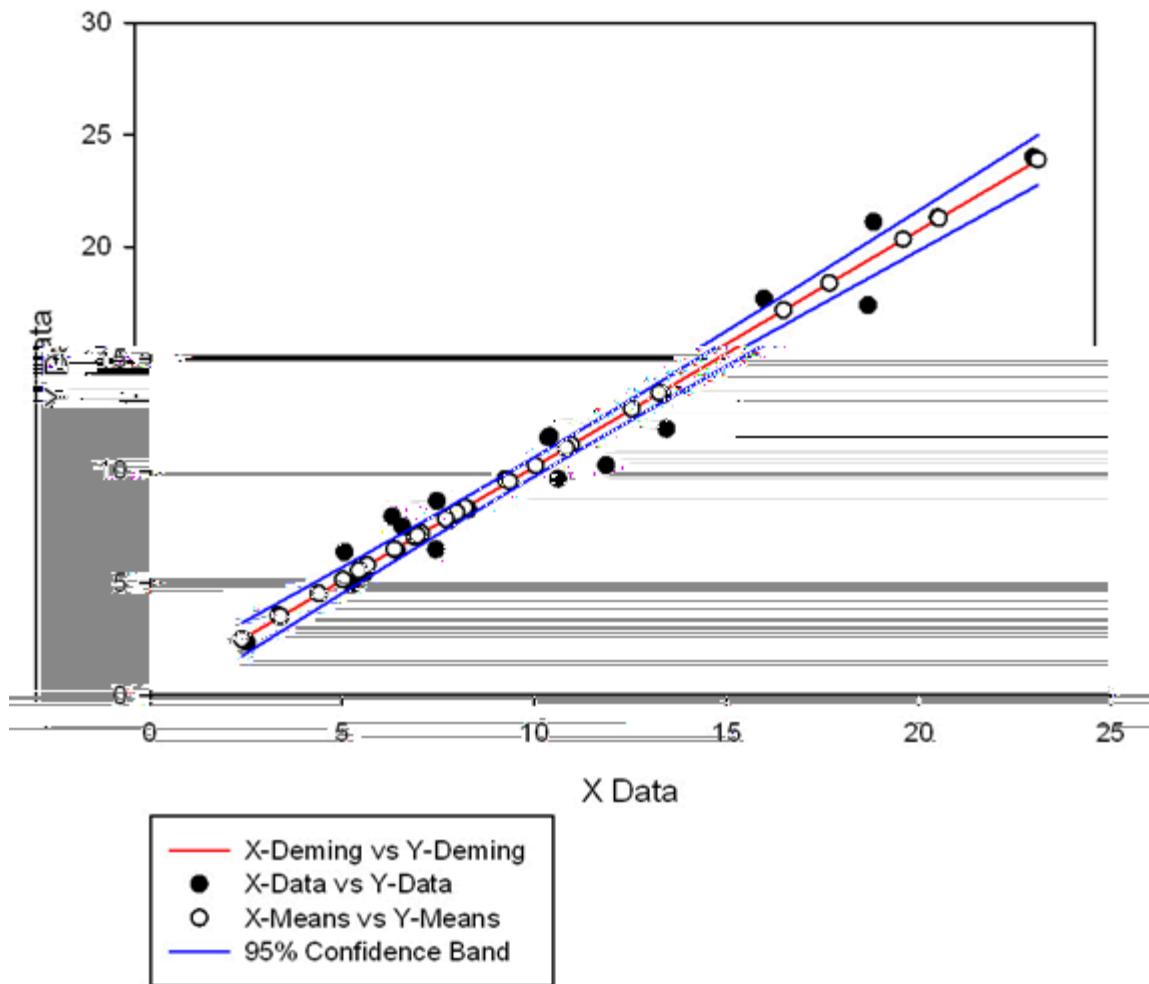
Deming Regression Result Graph

The default graph for Deming Regression contains a plot of the regression line together with a scatter plot of the raw data. There are options available in the Options dialog for creating additional plots on the graph.

One option is to add a scatter plot of the pairs of estimated means for each observation. Since the regression line gives the linear relationship between these means, this scatter plot will lie on top of the regression line.

Confidence bands for the predicted values of the dependent variable can also be added to the graph. These bands measure the accuracy of the predicted values for Y assuming specified values for X .

Deming Graph



Chapter

10

Survival Analysis

Topics:

- [Five Survival Tests](#)
- [Data Format for Survival Analysis](#)
- [Single Group Survival Analysis](#)
- [Survival LogRank Analysis](#)
- [Gehan-Breslow Survival Analysis](#)
- [Cox Regression](#)
- [Survival Curve Graph Examples](#)
- [Editing Survival Graphs Using Graph Properties](#)
- [Using Test Options to Modify Graphs](#)

Survival analysis studies the variable that is the time to some event. The term *survival* originates from the event *death*. But the event need not be death; it can be the time to any event. This could be the time to closure of a vascular graft or the time when a mouse footpad swells from infection. Of course it need not be medical or biological. It could be the time a motor runs until it fails. For consistency we will use survival and death (or failure) here.

Sometimes death doesn't occur during the length of the study or the patient dies from some other cause or the patient relocates to another part of the country. Though a death did not occur, this information is useful since the patient survived up until the time he or she left the study. When this occurs the patient is referred to as *censored*. This comes from the expression *censored from observation* – the data has been lost from view of the study. Examples of censored values are patients who moved to another geographic location before the study ended and patients who are alive when the study ended. Kaplan-Meier survival analysis includes both failures (death) and censored values.

Five Survival Tests

Use the Survival statistic to obtain one of the following five tests:

- **Single Group.** Use this to analyze and graph one survival curve.
- **LogRank.** Use this to compare two or more survival curves. The LogRank test assumes that all survival time data is equally accurate and all data will be equally weighted in the analysis.
- **Gehan-Breslow.** Use this to compare two or more survival curves when you expect the early data to be more accurate than later. Use this, for example, if there are many more censored values at the end of the study than at the beginning.
- **Cox Regression – Proportional Hazards Model.** Use this to study the impact of potential risk factors on the survival time of a population when the survival data is sampled from a single group.
- **Cox Regression – Stratified Model.** Use this to study the impact of potential risk factors on the survival time of a population when the survival data is sampled from a multiple groups.

Data Format for Survival Analysis

Survival data consists of three variables:

- Survival time
- Status
- Group

The survival times are the times when the event occurred. They must be positive and all non-positive values will be considered missing values. Survival time or group need not sort the data.

The status variable defines whether the data is a failure or censored value. You are allowed to use multiple names for both failure and censored. These can be text or numeric.

The group variable defines each individual survival data set (and curve).

Arrange data in the worksheet in either of two formats:

- **Raw data format.** Column pairs of survival time and status value for each group.
- **Indexed data format.** Data indexed to a group column.

Raw Data

To enter the data in Raw data format, enter the survival time in one column and the corresponding status in a second column. Do this for each group. If you wish, you can identify each group with a column title in the survival time column. If you do this then these group titles will be used in the graph and report.

	1-Affected Node	2-Status	3-Total Node	4-Status	
1	346.0000	failure	1699.0000	censored	
2	151.0000	failure	2177.0000	censored	
3	296.0000	failure	1969.0000	censored	
4	1953.0000	censored	1889.0000	censored	
5	1375.0000	failure	173.0000	failure	
6	822.0000	failure	2070.0000	censored	
7	2052.0000	censored	1972.0000	censored	
8	836.0000	failure	1897.0000	censored	
9	1910.0000	censored	2022.0000	censored	
10	419.0000	failure	1879.0000	censored	
11	107.0000	failure	1726.0000	censored	
12	570.0000	failure	1807.0000	censored	
13	312.0000	failure	615.0000	failure	
		1818.0000			

Figure 114: Raw Data Format for a Survival Analysis with Two Groups

In the graph above, columns 1 and 2 are the survival time and status values for the first group - Affected Node. Columns 3 and 4 are the same for the second group - Total Node. The report and the survival curve graph will use the text strings ("Affected Node", "Total Node") found in the survival time column titles.

Important: The worksheet columns for each group must be the same length. If not then the cells in the longer length column will be considered missing. All non-positive survival times will also be considered missing. All status variable values not defined as either a failure or a censored value will be considered missing.

Indexed Data

Indexed data is a three-column format. The survival time and status variable in two columns are indexed on the group names in a third column. Informative column titles are not necessary but are useful when selecting columns in the wizard.

	1-Group	2-Time	3-Status
1	Affected Node	346.0000	failure
2	Affected Node	151.0000	failure
3	Affected Node	296.0000	failure
4	Affected Node	1953.0000	censored
5	Affected Node	1375.0000	failure
6	Affected Node	822.0000	failure
7	Total Node	1699.0000	censored
8	Total Node	2177.0000	censored
9	Total Node	1969.0000	censored
10	Total Node	1889.0000	censored
11	Total Node	173.0000	failure
12	Total Node	2070.0000	censored
13	Total Node	1972.0000	censored
14	Total Node	1897.0000	censored
	Node	2022.0000	

Figure 115: Indexed Data Format - a Three-Column Format Consisting of Group, Survival Time, and Status

In the example above, group is in column 1, survival time is in column 2 and the status variable is in column 3.

Note: The Index and Unindex transforms are not designed for converting between survival analysis data formats. To use these features you must index and unindex the survival time and status variables separately and then reorganize the resulting columns.

Single Group Survival Analysis

Single Group Survival Analysis analyzes the survival data from one group, and then creates a report and a graph with a single survival curve. There is no statistical test performed but statistics associated with the data, such as the median survival time, are calculated and presented in the report.

Performing a Single Group Survival Analysis

1. Enter or arrange your data in the worksheet.
2. If desired set the Single Group options.
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Survival > Kaplan-Meier > Single Group
5. Select the two worksheet columns with the survival times and status values in the **Select Data** panel of the Test Wizard.
6. Interpret the Single Group survival analysis report and curve.

Arranging Single Group Survival Analysis Data

Two data columns are required, a column with survival times and a column with status labels. These can be just two columns in a worksheet or two columns from a multi-group data set. You can select a single pair of columns from the multiple groups in the Raw data format.

Tip: Use this option to analyze all groups as a single group from an indexed format data set. For example, select the last two columns in the worksheet to analyze both groups as one group. You cannot do this directly

with Raw data format since the groups are not concatenated in two columns. You would need to use the Stack transform in Transforms to concatenate the columns.

Setting Single Group Test Options

Use the Survival Curve Test Options to:

- Specify attributes of the generated survival curve graph.
- Customize the post-test contents of the report and worksheet.

To change the Survival Curve options:

- If you are going to analyze your survival curve after changing test options, and want to select your data before you create the curve, then drag the pointer over your data.
- Select **Survival Single Group** from the **Tests** drop-down list in the **SigmaStat** group.
- Click **Options** in the **SigmaStat** group. The **Options for Survival Single Group** dialog box appears with two tabs:
 - Graph Options.** Click the **Graph Options** tab to view the graph symbol, line and scaling options. You can select additional statistical graph elements here.
 - Results.** Click the **Results** tab to specify the survival time units and to modify the content of the report and worksheet.

SigmaPlot saves the options settings between sessions.

- To continue the test, click **Run Test**.

The **Select Data** panel of the Test Wizard appears.

- To accept the current settings and close the dialog box, click **OK**.

Tip: All options in these dialog boxes are "sticky" and remain in the state that you have selected until you change them.

Options for Survival Single Group: Graph Options

Status Symbols. All graph options apply to graphs that are created when the analysis is run. You can use Graph Properties to modify the attributes of the survival curves after they have been created.

- Censored.** Click the **Graph Options** tab from the **Options for Survival Single Group** dialog box to view the status symbols options. Censored symbols are graphed by default. Clear this option to not display the censored symbols.
- Failures.** Select **Failures to display symbols** at the failure times. These symbols always occupy the inside corners of the steps in the survival curve. As such they provide redundant information and need not be displayed.

Group Color. The color of the objects in a survival curve group may be changed with this option. All objects (for example, survival line, symbols, confidence interval lines) are changed to the selected color. Use Graph Properties to modify individual object colors after the graph has been created.

Survival Scale. You can display the survival graph either using fractional values (probabilities) or percents. Select one of the following:

- Fraction.** If you select this then the Y-axis scaling will be from 0 to 1.
- Percent.** Selecting this will result in a Y-axis scaling from 0 to 100.

Note: The results in the report are always expressed in fractional terms no matter which option is selected for the graph.

Additional Plot Statistics. You can add two different types of graph elements to your survival curve from the Type drop-down list:

- 95% Confidence Intervals.** Selecting adds the upper and lower confidence lines in a stepped line format.
- Standard Error Bars.** Selecting this will add error bars for the standard errors of the survival probability. These are placed at the failure times. All of these elements will be graphed with the same color as the survival curve. You may change these colors, and other graph attributes, in **Graph Properties** after creating the graph.

Options for Single Group Survival: Results

Report.

- **Cumulative Probability Table.** Clear this option to exclude the cumulative probability table from the report. This reduces the length of the report for large data sets.

Worksheet.

- **95% Confidence Intervals.** Select this to place the survival curve upper and lower 95% confidence interval values into the worksheet. These are placed into the first empty worksheet columns.

Time Units. Select a time unit from the drop-down list or enter a unit. These units are used in the graph axis titles and the survival report.

Running a Single Group Survival Analysis

To run a single group survival analysis you need to select survival time and status data columns to analyze. Use the Select Data panel of the Test Wizard to select these two columns in the worksheet.

To run a Single Group analysis:

1. Specify any options for your graph and report.
2. If you want to select your data before you run the test then drag the pointer over your data. The Survival Time column must precede and be adjacent to the Status column.
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Survival > Kaplan-Meier > Single Group

The **Select Data** panel of the Test Wizard appears prompting you to select your data columns. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

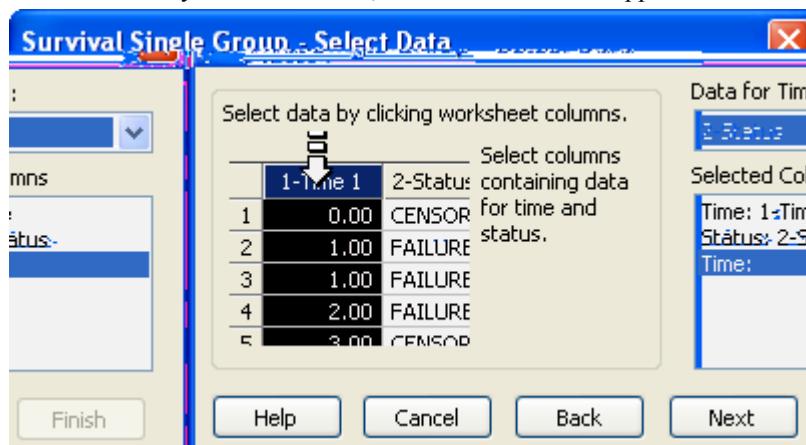


Figure 116: The Select Data for Survival Single Group Panel Prompting You to Select Time and Status Columns

5. **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for** drop-down list. The first selected column is assigned to the first row (Time) in the **Selected Columns** list, and the next selected column is assigned to the next row (Status) in the list. The number or title of selected columns appears in each row.
6. **To change your selections**, select the assignment in the list and then select a new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.

7. Click **Next** to choose the status variables. The status variables found in the columns you selected are shown in the Status labels in selected columns window. Select these and click the right arrow buttons to place the event variables in the **Event** window and the censored variable in the **Censored** window.

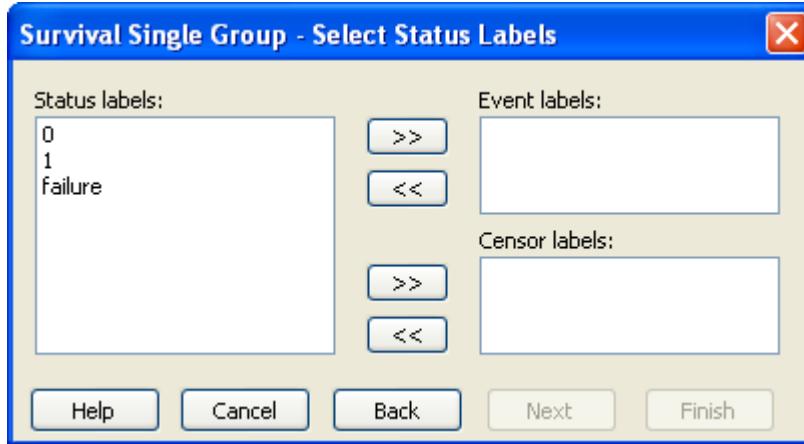
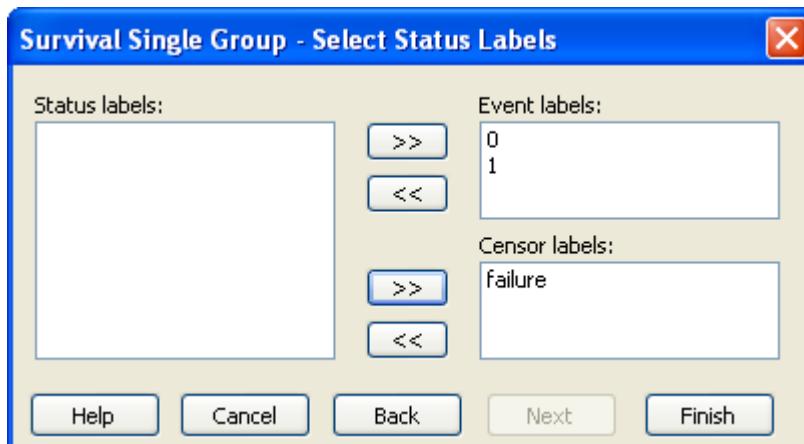


Figure 117: The Select Status Labels for Survival Single Group Panel Prompting You to Select the Status Variables.

You can have more than one **Event** label and more than one **Censored** label. You must select one Event label in order to proceed. You need not select a censored variable, though, and some data sets will not have any censored values. You need not select all the variables; any data associated with cleared status variables will be considered missing.

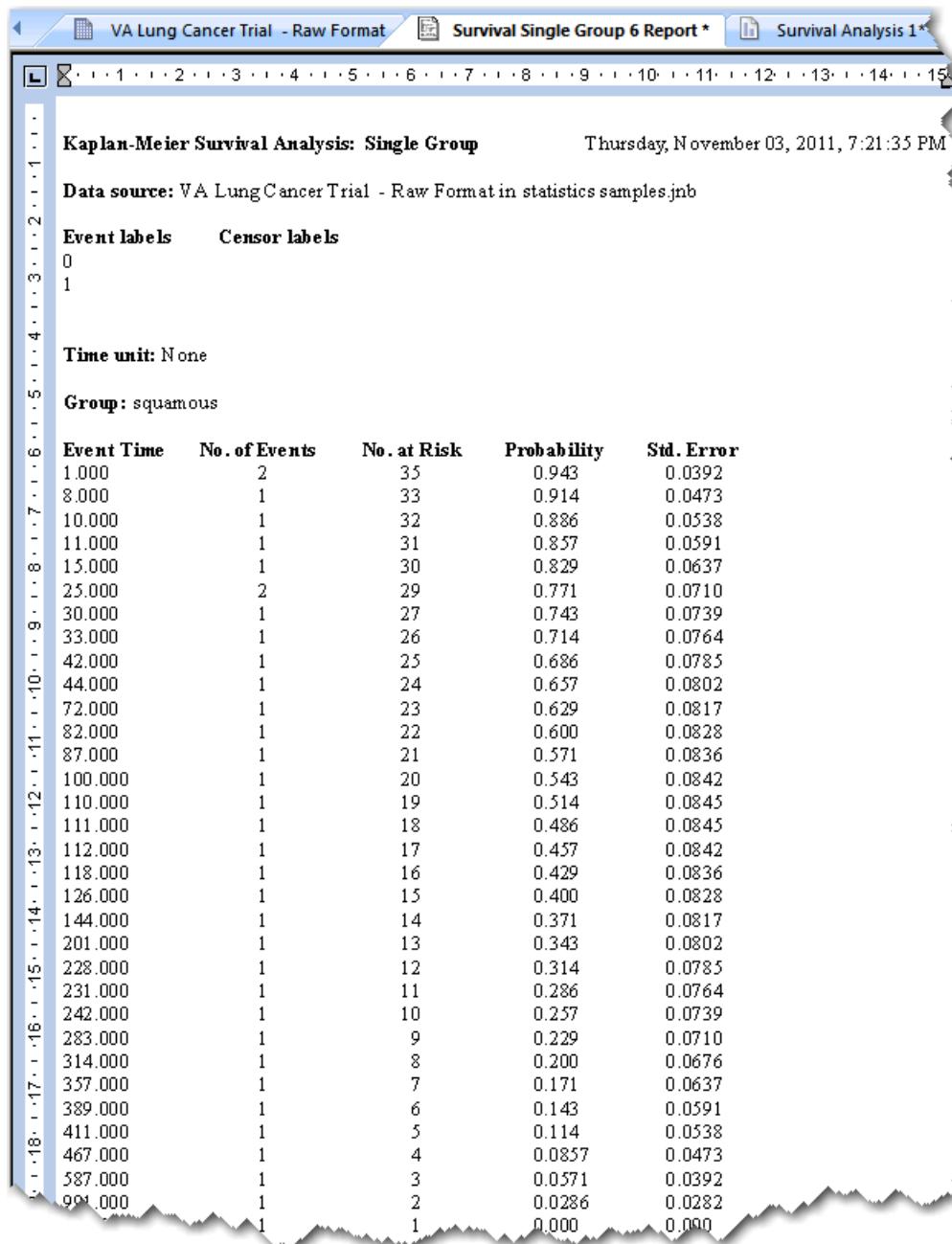


8. Click the back arrows to remove labels from the **Event** and **Censored** windows. This places them back in the Status labels in selected columns window.
- SigmaPlot saves the **Event** and **Censored** labels that you selected for your next analysis. If the next data set contains exactly the same status labels, or if you are reanalyzing your present data set, then the saved selections appear in the **Event** and **Censored** windows.
9. Click **Finish** to create the survival graph and report. The results you obtain depend on the Test Options that you selected.

Interpreting Single Group Survival Results

The Single Group survival analysis report displays information about the origin of your data, a table containing the cumulative survival probabilities and summary statistics of the survival curve.

For descriptions of the derivations for survival curve statistics see Hosmer & Lemeshow or Kleinbaum.



Results Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display.

Report Header Information

The report header includes the date and time that the analysis was performed. The data source is identified by the worksheet title containing the data being analyzed and the notebook name. The event and censor labels used in this analysis are listed. Also, the time units used are displayed.

Survival Cumulative Probability Table

The survival probability table lists all event times and, for each event time, the number of events that occurred, the number of subjects remaining at risk, the cumulative survival probability and its standard error. The upper and lower

95% confidence limits are not displayed but these may be placed into the worksheet. Failure times are not shown but you can infer their existence from jumps in the Number at Risk data and the summary table immediately below this table

You can turn the display of this table off by clearing this option in the Results tab of Test Options. This is useful for large data sets.

Data Summary Table

The data summary table shows the total number of cases. The sum of the number of events, censored and missing values, shown below this, will equal the total number of cases.

Statistical Summary Table

The mean and percentile survival times and their statistics are listed in this table. The median survival time is commonly used in publications.

Single Group Survival Graph

Visual interpretation of the survival curve is an important component of survival analysis. For this reason SigmaPlot always generates a survival curve graph. This is different from the other statistical tests where you select a report graph a posteriori.

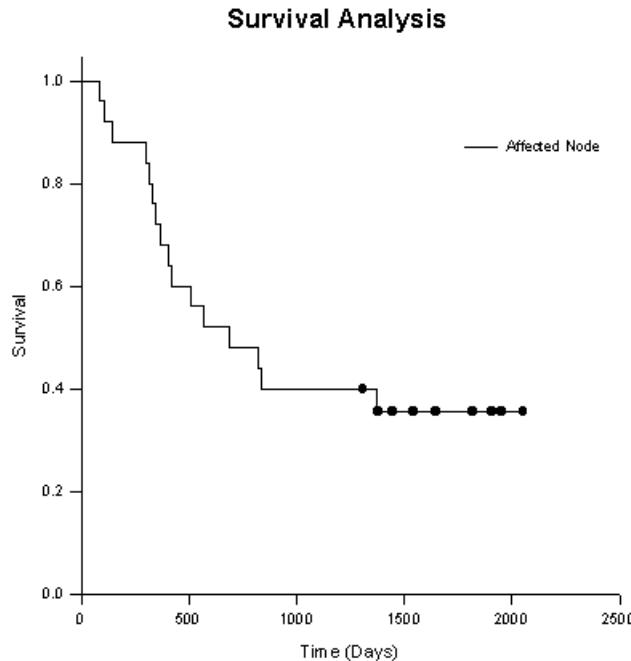


Figure 118: A Single Group Survival Curve

You can control the graph in two ways:

- You can set the graph options to become the default values until they are changed.
- After the graph is created you can modify it using SigmaPlot's Graph Properties. Each object in the graph is a separate plot (for example, survival curve, failure symbols, censored symbols, upper confidence limit, etc.) so you have considerable control over the appearance of your graph.

Survival LogRank Analysis

LogRank Survival Analysis analyzes survival data from multiple groups and creates a report and a graph showing multiple survival curves. Statistics associated with each group, such as the median survival time, are calculated and presented in the report.

You can also perform the LogRank test to determine whether survival curves are significantly different. It is a nonparametric test that uses a chi-square statistic to reject the null hypothesis that the survival curves came from the same population. The LogRank statistic is approximately the same as the familiar chi-square statistic, and is formed from the sum across groups of the square of the difference of the actual and estimated number of events for each

group (censored values removed) divided by the estimated number of events $(\sum(O_i - E_i)^2/E_i)$. It generates a P value that is the probability of the chance occurrence of survival curves as different (or more so) as those observed.

The LogRank test assumes that there is no difference in the accuracy of the data at any given time. This is different from the Gehan-Breslow test that weights the early data more since it assumes that this data is more accurate.

Performing a Survival LogRank Analysis

1. [Enter or arrange your data in either Indexed or Raw data format in the worksheet.](#)
2. If desired set the LogRank options.
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Survival > Kaplan-Meier > LogRank
5. [View the LogRank survival analysis graph.](#)

Arranging Survival LogRank Analysis Data

Multiple Time, Status column pairs (two or more) are required for Raw data format. Indexed data format requires three columns for Group, Time and Status. You can preselect the data to have the column selection panel automatically select the Time, Status column pairs if you organize your worksheet with the Time column preceding the Status column and have all columns be adjacent. For Indexed data format, placing the Group, Time and Status variables in adjacent columns and in that order also allows automatic column selection.

Setting Survival LogRank Options

Use the Survival LogRank Test Options to:

- Specify attributes of the generated survival curve graph
- Customize the post-test contents of the report and worksheet
- Select the multiple comparison test and its options

To change the Survival Curve options:

1. If you are going to analyze your survival curve after changing test options, and want to select your data before you create the curve, then drag the pointer over your data.
2. Select **Survival LogRank** from the **Tests** drop-down list in the **SigmaStat** group.
3. Click **Options** in the **SigmaStat** group. The **Survival LogRank** Options for dialog box appears with three tabs:
 - **Graph Options.** Click the **Graph Options** tab to view the graph symbol, line and scaling options. Additional statistical graph elements may also be selected here.
 - **Results.** Click the Results tab to specify the survival time units and to modify the content of the report and worksheet.
 - **Post Hoc Tests.** Click the Post Hoc Tests tab to modify the multiple comparison options.

SigmaPlot saves options settings between sessions.

4. **To continue the test, click Run Test.**

- To accept the current settings and close the options dialog box, click **OK**.

Options for Survival LogRank: Graph Options

Status Symbols. All graph options apply to graphs that are created when the analysis is run. Use Graph Properties to modify the attributes of the survival curves after they have been created.

- Censored Symbols.** Select the Graph Options tab on the Options dialog box to view the status symbols options. Censored symbols are graphed by default. Clear this option to not display the censored symbols.
- Failures Symbols.** Selecting this box displays symbols at the failure times. These symbols always occupy the inside corners of the steps in the survival curve. As such they provide redundant information and need not be displayed.

Group Color. The color of the objects in a survival curve group may be changed with this option. All objects, for example, survival line, symbols, confidence interval lines, will be changed to the selected color or color scheme. A four density gray scale color scheme is used as the default. You may change this to black, where all survival curves and their attributes will be black, or incrementing that is a multi-color scheme. Use Graph Properties to modify individual object colors after the graph has been created.

Survival Scale. You can display the survival graph either using fractional values (probabilities) or percents.

- Fraction.** If you select this then the Y axis scaling will be from 0 to 1.
- Percent.** Selecting this will result in a Y axis scaling from 0 to 100.

 **Note:** The results in the report are always expressed in fractional terms no matter which option is selected for the graph.

Additional Plot Statistics. Two different types of graph elements may be added to your survival curves. You can select one of two **Types**:

- 95% Confidence Intervals.** Selecting this will add the upper and lower confidence lines in a stepped line format.
- Standard Error Bars.** Selecting this will add error bars for the standard errors of the survival probability. These are placed at the failure times. All of these elements will be graphed with the same color as the survival curve. You may change these colors, and other graph attributes, in Graph Properties after the graph has been created.

Options for Survival Log Rank: Results

Report.

- Cumulative Probability Table.** Clear this option to exclude the cumulative probability table from the report. This reduces the length of the report for large data sets.
- P values for multiple comparisons.** Select this to show both the P values from the pairwise multiple comparison tests and the critical values against which the pairwise P values are tested. The critical values for the Holm-Sidak test will vary for each pairwise test. If this is selected for the Bonferroni test, the critical values will be identical for all pairwise tests.

 **Tip:** You can change the critical P value for the LogRank test on the Options dialog box. For more information, see [. This is a global setting for the critical P value and affects all tests in SigmaPlot. Report Graphs](#) on page 373

Time Units. Select a time unit from the drop-down list or enter a unit. These units will be used in the graph axis titles and the survival report.

Worksheet.

- 95% Confidence Intervals.** Select this to place the survival curve upper and lower 95% confidence intervals into the first empty worksheet columns.

Options for Survival LogRank: Post Hoc Tests

Multiple Comparisons. You can select when multiple comparisons are to be computed and displayed in the report. LogRank tests the hypothesis of no differences between survival groups but does not determine which groups are different, or the sizes of these differences. Multiple comparison procedures isolate these differences.

- **Always Perform.** Select this option to always display multiple comparison results in the report. If the original comparison test is not significant then the multiple comparison results will also be not significant and will just clutter the report. The multiple comparison test is a separate computation from the original comparison test so it is possible to obtain significant results from the multiple comparison test when the original test was insignificant.
 - **Only when Survival P Value is Significant.** Select this to place multiple comparison results in the report only when the original comparison test is significant. The significance level can be set to either 0.05 or 0.01 using the Significance Value for Multiple Comparisons drop-down list.
- i** **Tip:** If multiple comparisons are triggered, the report will show the results of the comparison. You may elect to always show them by clearing **Only when Survival P Value is Significant**.

Running a Survival LogRank Analysis

To run a LogRank survival analysis you need to select data in the worksheet and specify the status variables.

To run a LogRank Survival analysis:

1. If you want to select your data before you run the test then drag the pointer over your data. The columns must be adjacent and in the correct order (Time, Status for Raw data and Group, Time Status for Indexed data). For more information, see [Arranging Survival LogRank Analysis Data](#) on page 316.
2. Select the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Survival > Kaplan-Meier > LogRank
The Survival LogRank – Data Format panel of the Test Wizard appears.
4. From the Data Format drop-down list select either:
 - Raw data format when you have groups of data in multiple Time, Status column pairs.
 - Indexed data format when you have the groups specified by a column.

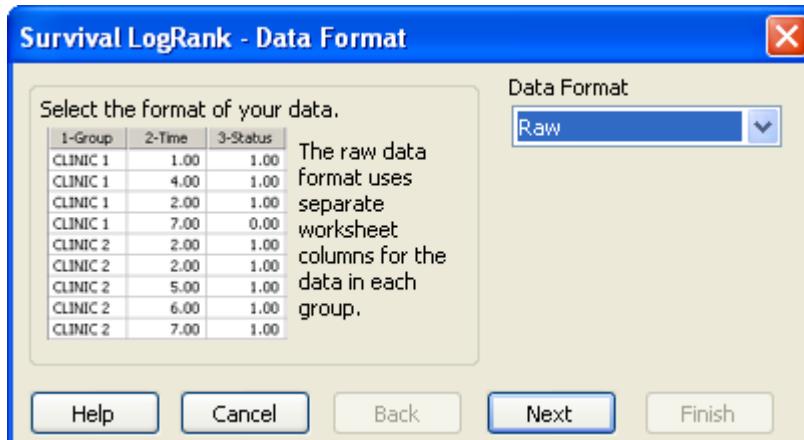


Figure 119: The Data Format Panel With Raw Data Format Selected

5. Click **Next** to display the **Select Data** panel that prompts you to select your data columns. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

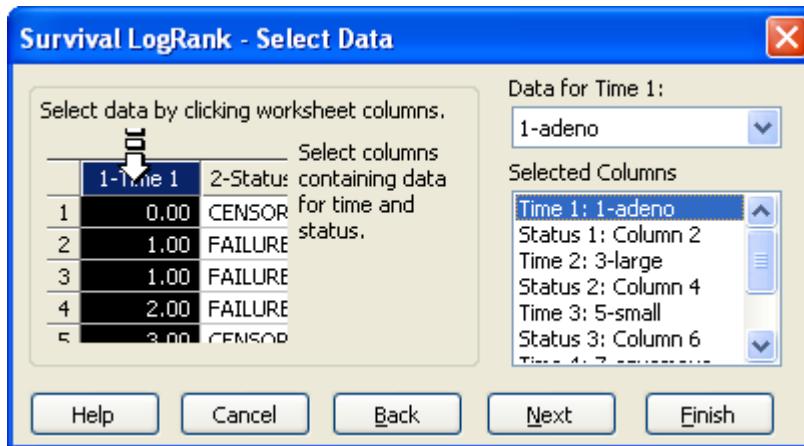


Figure 120: The Select Data Panel for Survival LogRank Raw Data Format Prompting You to Select Multiple Time and Status Columns

6. To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for** drop-down list.

The first selected column is assigned to the first row (Time 1) in the Selected Columns list, and the next selected column is assigned to the next row (Status 1) in the list. The number or title of selected columns appears in each row. Continue selecting Time, Status columns for all groups that you wish to analyze.

7. To change your selections, select the assignment in the list and then select a new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
8. Click **Next** to choose the status variables. The status variables found in the columns you selected are shown in the **Status labels in selected columns:** box. Select these and click the right arrow buttons to place the event variables in the **Event:** window and the censored variable in the **Censored:** window.

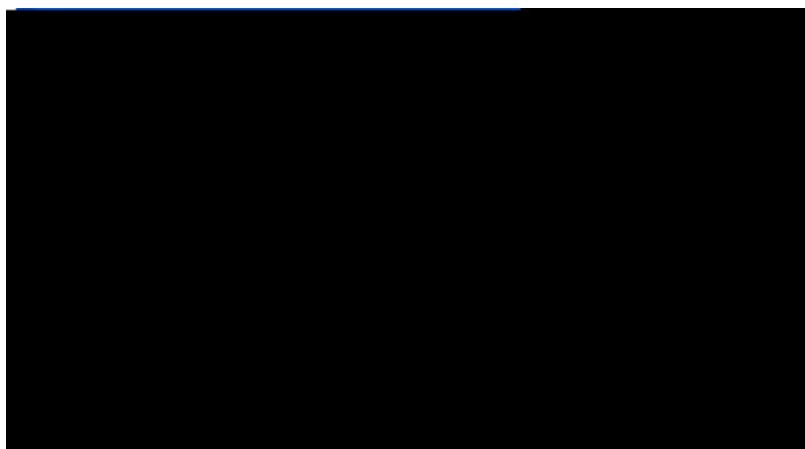


Figure 121: The Survival LogRank — Select Status Labels Panel of the Test Wizard Prompting You to Select the Status Variables

You can have more than one Event label and more than one Censored label. You must select one Event label in order to proceed. You need not select a censored variable, though, and some data sets will not have any censored values. You need not select all the variables; any data associated with unselected status variables will be considered missing.

9. Click the back arrow keys to remove labels from the **Event:** and **Censored:** windows. This places them back in the **Status labels in selected columns:** window.

SigmaPlot saves the Event and Censored labels that you selected for your next analysis. If the next data set contains exactly the same status labels, or if you are re-analyzing your present data set, then the saved selections appear in the **Event:** and **Censored:** windows.

10. Click **Finish** to create the survival graph and report. The results you obtain depend on the **Test Options** that you selected.

If you selected Indexed data format then the **Select Data** panel of the Test Wizard asks you to select the three columns in the worksheet for your Group, Time and Status.

11. Click **Next** to select the groups you want to include in the analysis. If you want to analyze all groups found in the Group column then select **Select all groups**. Otherwise select groups from the **Data for Group** drop-down list. You can select subsets of all groups and select them in the order that you wish to see them in the report.

12. Click **Next** to select the status variables as described above and then continue to complete the analysis to create the report and graph.

Multiple Comparison Options

LogRank tests the hypothesis of no differences between the several survival groups, but does not determine which groups are different, or the sizes of the differences. Multiple comparison tests isolate these differences by running comparisons between the experimental groups.

If you selected to run multiple comparisons only when the P value is significant, and LogRank produces a P value equal to or less than the trigger P value, or you selected to always run multiple comparison in the Options for LogRank dialog, the multiple comparison results are displayed in the Report.

There are two multiple comparison tests to choose from for the LogRank survival analysis:

- [Holm-Sidak](#).
- [Bonferroni](#).

Holm-Sidak Test

The Holm-Sidak Test can be used for both pairwise comparisons and comparisons versus a control group. It is more powerful than the Bonferroni test and, consequently, it is able to detect differences that the Bonferroni test does not. It is recommended as the first-line procedure for pairwise comparison testing.

When performing the test, the P values of all comparisons are computed and ordered from smallest to largest. Each P value is then compared to a critical level that depends upon the significance level of the test (set in the test options), the rank of the P value, and the total number of comparisons made. A P value less than the critical level indicates there is a significant difference between the corresponding two groups.

Multiple Comparisons:

All Pairwise Multiple Comparison Procedures (Holm-Sidak method)

Significance level = 0.05

Comparisons	Statistic	Unadjusted P Value	Critical Level	Significant?
adeno vs. large	17.669	0.0000263	0.00851	Yes
squamous vs. adeno	12.045	0.000519	0.0102	Yes
squamous vs. small	11.574	0.000669	0.0127	Yes
small vs. large	9.371	0.00220	0.0170	Yes
squamous vs. large	0.823	0.364	0.0253	No
small vs. adeno	0.0968	0.756	0.0500	No

Figure 122: Holm-Sidak Multiple Comparison Results for VA Lung Cancer Study

Bonferroni Test

The Bonferroni test performs pairwise comparisons with paired chi-square tests. It is computationally similar to the Holm-Sidak test except that it is not sequential (the critical level used is fixed for all comparisons). The critical level is the ratio of the family P value to the number of comparisons. It is a more conservative test than the Holm-Sidak test in that the chi-square value required to conclude that a difference exists becomes much larger than it really needs to be.

The critical level is constant at $0.05/6 = 0.00833$. Since the critical level does not increase, as it does for the Holm-Sidak test, there will tend to be fewer comparisons with significant differences.

Multiple Comparisons:

All Pairwise Multiple Comparison Procedures (Bonferroni method)

Significance level = 0.05

Comparisons	Statistic	Unadjusted P Value	Critical Level	Significant?
squamous vs. small	11.574	0.000669	0.00833	Yes
squamous vs. adeno	12.045	0.000519	0.00833	Yes
squamous vs. large	0.823	0.364	0.00833	No
small vs. adeno	0.0968	0.756	0.00833	No
small vs. large	9.371	0.00220	0.00833	Yes
adeno vs. large	17.669	0.0000263	0.00833	Yes

Figure 123: Bonferroni Multiple Comparison Results for VA Lung Cancer Study

Interpreting Survival LogRank Results

The LogRank survival analysis report displays information about the origin of your data, tables containing the cumulative survival probabilities for each group, summary statistics for each survival curve and the LogRank test of significance. Multiple comparison test results will also be displayed provided significant differences were found or the Post Hoc Tests Options were selected to display them.

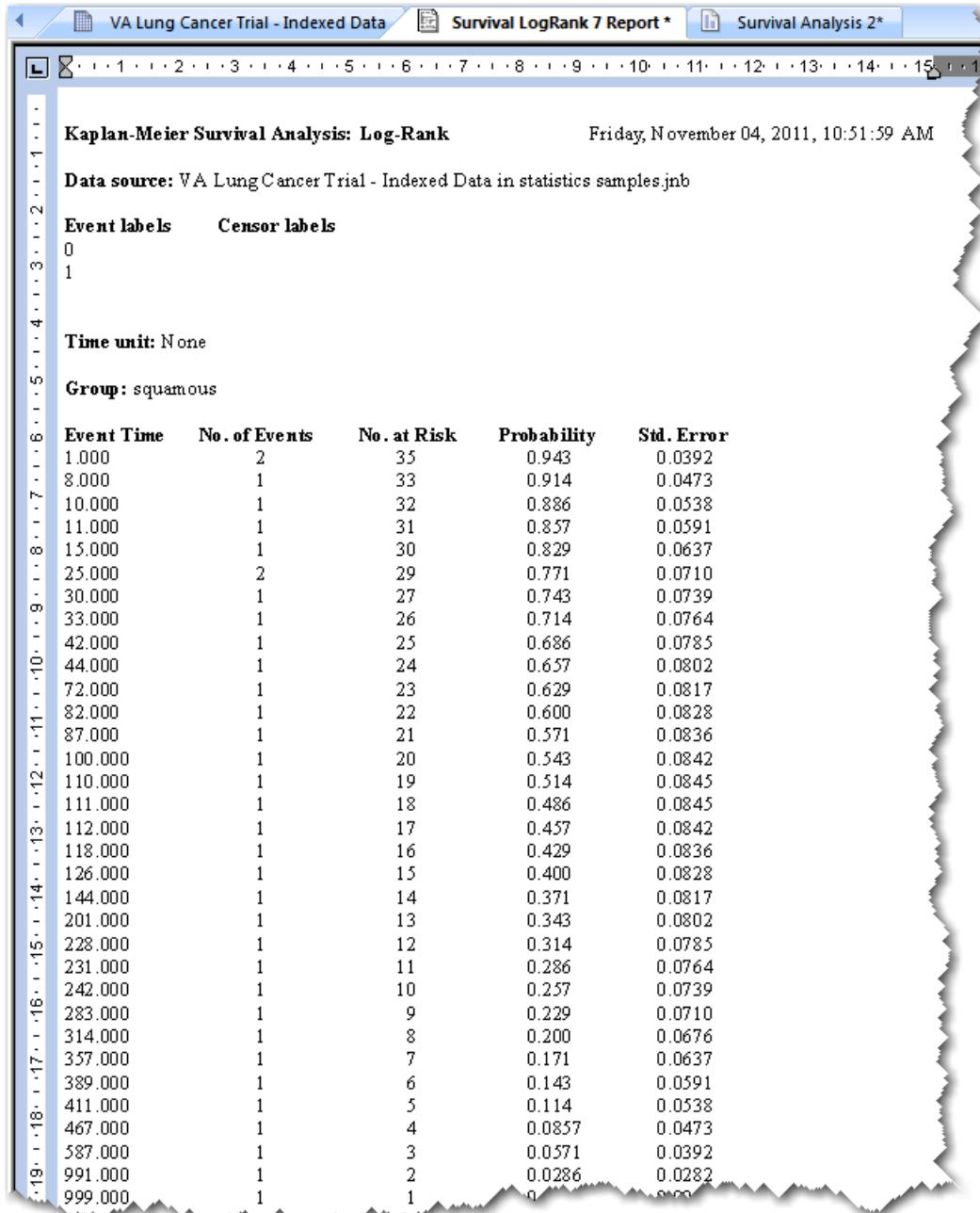


Figure 124: The LogRank Survival Analysis Results Report

Results Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display in the Options dialog box.

Report Header Information

The report header includes the date and time that the analysis was performed. The data source is identified by the worksheet title containing the data being analyzed and the notebook name. The event and censor labels used in this analysis are listed. Also, the time units used are displayed.

Survival Cumulative Probability Table

The survival probability table lists all event times and, for each event time, the number of events that occurred, the number of subjects remaining at risk, the cumulative survival probability and its standard error. The upper and lower 95% confidence limits are not displayed but these may be placed into the worksheet. Failure times are not shown but you can infer their existence from jumps in the Number at Risk data and the summary table immediately below this table.

You can turn the display of this table off by clearing this option in the Results tab of Test Options. This is useful to keep the report a reasonable length when you have large data sets.

Data Summary Table

The data summary table shows the total number of cases. The sum of the number of events, censored and missing values, shown below this, will equal the total number of cases.

Statistical Summary Table

The mean and percentile survival times and their statistics are listed in this table. The median survival time is commonly used in publications.

Survival LogRank Graph

Visual interpretation of the survival curve is an important component of survival analysis. For this reason SigmaPlot always generates a survival curve graph. This is different from the other statistical tests where you select a report graph a posteriori.

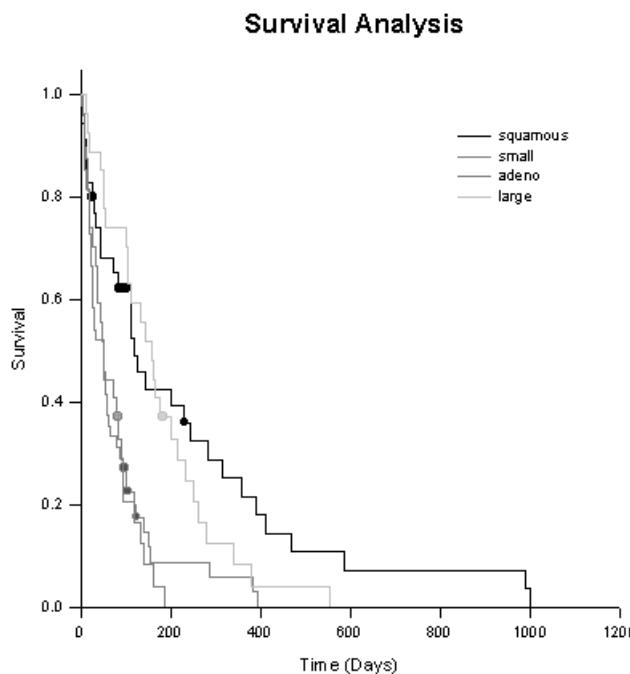


Figure 125: LogRank Survival Curves

In the graph above, the default Test Options, gray scale colors, solid circle symbols, was used. Squamous and large cell carcinomas do not appear to be significantly different (as well as small cell and adenocarcinoma). This is confirmed by the LogRank test.

You can control the graph in two ways:

- You can set the graph options to become the default values until they are changed.

- After the graph is created you can modify it using SigmaPlot's Property Browser. Each object in the graph is a separate plot (for example, survival curve, failure symbols, censored symbols, upper confidence limit, etc.) so you have considerable control over the appearance of your graph.

Gehan-Breslow Survival Analysis

The Gehan-Breslow option analyzes survival data from multiple groups, creates a report and a graph showing multiple survival curves. Statistics associated with each group, such as the median survival time, are calculated and presented in the report.

The Gehan-Breslow test is also performed to determine whether survival curves are significantly different. It is a nonparametric test that uses a chi-square statistic to reject the null hypothesis that the survival curves came from the same population. It generates a P value that is the probability of the chance occurrence of survival curves as different (or more so) as those observed.

The Gehan-Breslow test assumes that the early survival times are known more accurately than later times and weights the data accordingly. As an example, you would want to use Gehan-Breslow if there were many late-survival-time censored values. This is different from the LogRank test that assumes there is no difference in the accuracy of the survival times.

Performing a Gehan-Breslow Analysis

1. Enter or arrange your data in either **Indexed** or **Raw** data format in the worksheet. For more information, see [Arranging Gehan-Breslow Survival Analysis Data](#) on page 324.
2. If desired set the Gehan-Breslow options.
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Survival > Kaplan-Meier > Gehan-Breslow
5. Run the test.
6. Generate a report graph. See [Gehan-Breslow Survival Graph](#) on page 332 for more information.

Arranging Gehan-Breslow Survival Analysis Data

Multiple Time, Status column pairs (two or more) are required for Raw data format. Indexed data format requires three columns for Group, Time and Status. You can preselect the data to have the column selection panel automatically select the Time, Status column pairs if you organize your worksheet with the Time column preceding the Status column and have all columns be adjacent.

Setting Gehan-Breslow Survival Options

Use the Survival Gehan-Breslow Test Options to:

- Specify attributes of the generated survival curve graph.
- Customize the post-test contents of the report and worksheet.
- Select the multiple comparison test and its options.

To change the Survival Curve options:

1. If you are going to analyze your survival curve after changing test options, and want to select your data before you create the curve, then drag the pointer over your data.
2. Select the **Analysis** tab.
3. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Survival > Kaplan-Meier > Gehan-Breslow

4. Click **Options**. The **Options for Survival Gehan-Breslow** dialog box appears with three tabs:

- **Graph Options.** Click the **Graph Options** tab to view the graph symbol, line and scaling options. You can select additional statistical graph elements here.
- **Results.** Click the **Results** tab to specify the survival time units and to modify the content of the report and worksheet.
- **Post Hoc Tests.** Click the **Post Hoc Tests** tab to modify the multiple comparison options.

SigmaPlot saves the options settings between sessions.

5. To continue the test, click **Run Test**. The **Select Data** panel of the Test Wizard appears.

6. To accept the current settings, click **OK**.

Options for Survival Gehan-Breslow: Graph Options

Status Symbols. All graph options apply to graphs that are created when the analysis is run. Use Graph Properties to modify the attributes of the survival curves after they have been created.

- **Censored Symbols.** Select the **Graph Options** tab on the **Options** dialog box to view the status symbols options. Censored symbols are graphed by default. Clear this option to not display the censored symbols.
- **Failures Symbols.** Selecting this box displays symbols at the failure times. These symbols always occupy the inside corners of the steps in the survival curve. As such they provide redundant information and need not be displayed.

Group Color. The color of the objects in a survival curve group may be changed with this option. All objects, for example, survival line, symbols, confidence interval lines, will be changed to the selected color or color scheme. A four density gray scale color scheme is used as the default. You may change this to black, where all survival curves and their attributes will be black, or incrementing (a multi-color scheme). Use Graph Properties to modify individual object colors after the graph has been created.

Survival Scale. You can display the survival graph either using fractional values (probabilities) or percents.

- **Fraction.** If you select this then the Y axis scaling will be from 0 to 1.
- **Percent.** Selecting this will result in a Y axis scaling from 0 to 100.

 **Note:** The results in the report are always expressed in fractional terms no matter which option is selected for the graph.

Additional Plot Statistics. Two different types of graph elements may be added to your survival curves. You can select one of two **Types**:

- **95% Confidence Intervals.** Selecting this will add the upper and lower confidence lines in a stepped line format.
- **Standard Error Bars.** Selecting this will add error bars for the standard errors of the survival probability. These are placed at the failure times. All of these elements will be graphed with the same color as the survival curve. You may change these colors, and other graph attributes, from **Graph Properties** after the graph has been created.

Options for Survival Gehan-Breslow: Results

Report.

- **Cumulative Probability Table.** Clear this option to exclude the cumulative probability table from the report. This reduces the length of the report for large data sets.
- **P values for multiple comparisons.** Select this to show both the P values from the pairwise multiple comparison tests and the critical values against which the pairwise P values are tested. The critical values for the Holm-Sidak test will vary for each pairwise test. If this is selected for the Bonferroni test, the critical values will be identical for all pairwise tests.

 **Note:** You can also change the critical P value for the Gehan-Breslow test on the Options dialog box. This is a global setting for the critical P value and affects all tests in SigmaPlot. For more information, see [Report Graphs](#) on page 373.

Time Units. Select a time unit from the drop-down list or enter a unit. These units will be used in the graph axis titles and the survival report.

Worksheet.

- **95% Confidence Intervals.** Select this to place the survival curve upper and lower 95% confidence intervals into the first empty worksheet columns.

Options for Survival Gehan-Breslow: Post Hoc Tests

Multiple Comparisons. You can select when multiple comparisons are to be computed and displayed in the report. Gehan-Breslow tests the hypothesis of no differences between survival groups but does not determine which groups are different, or the sizes of these differences. Multiple comparison procedures isolate these differences.

- **Always Perform.** Select this option to always display multiple comparison results in the report. If the original comparison test is not significant then the multiple comparison results will also be not significant and will just clutter the report. The multiple comparison test is a separate computation from the original comparison test so it is possible to obtain significant results from the multiple comparison test when the original test was insignificant.
 - **Only when Survival P Value is Significant.** Select this to place multiple comparison results in the report only when the original comparison test is significant. The significance level can be set to either 0.05 or 0.01 using the Significance Value for Multiple Comparisons drop-down list.
- Tip:** If multiple comparisons are triggered, the report shows the results of the comparison. You may elect to always show them by clearing **Only when Survival P Value is Significant**.

Running a Gehan-Breslow Survival Analysis

To run a Gehan-Breslow survival analysis you need to select data in the worksheet and specify the status variables.

To run a Gehan-Breslow Survival analysis:

1. Specify any options for your graph, report and post-hoc tests.
2. If you want to select your data before you run the test then drag the pointer over your data. The columns must be adjacent and in the correct order, for example: Time, Status for Raw data and Group, Time Status for Indexed data.
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Survival > Kaplan-Meier > Gehan-Breslow
5. Click **Run**.

The **Survival Gehan-Breslow — Data Format** panel of the Test Wizard appears.

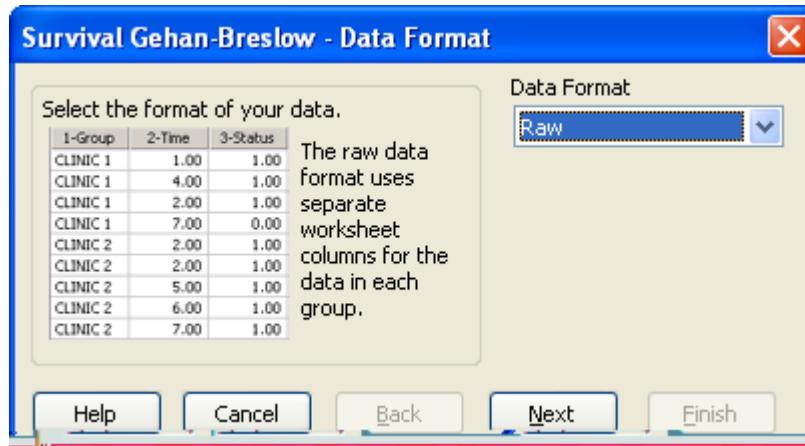


Figure 126: The Data Format Panel With Raw Data Format Selected

6. From the **Data Format** drop-down list select either:
 - **Raw.** Select the **Raw** data format if you have groups of data in multiple Time, Status column pairs.
 - **Indexed.** Select the **Indexed** data format when you have the groups specified by a column.
7. If you select **Raw Data**, click **Next**.
If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

8. To assign the desired worksheet columns to the Selected Columns list:

- a) Select the columns in the worksheet, or select the columns from the **Data for** drop-down list.

The first selected column is assigned to the first row (Time 1) in the **Selected Columns** list, and the next selected column is assigned to the next row (Status 1) in the list. The number or title of selected columns appears in each row. Continue selecting Time, Status columns for all groups that you wish to analyze.

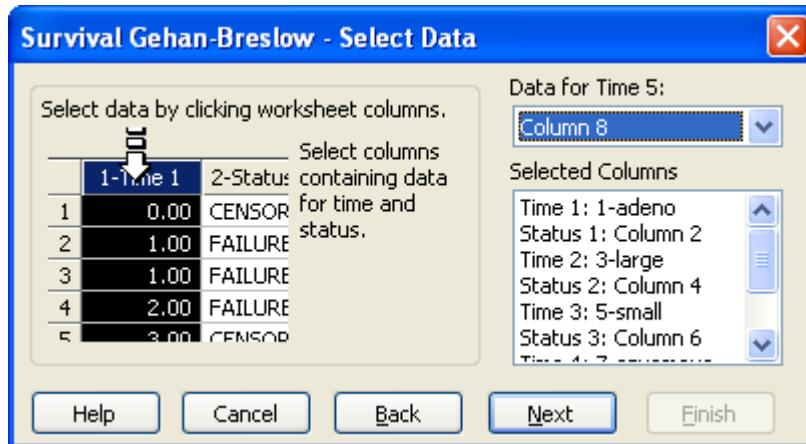


Figure 127: The Select Data Panel Prompting You to Select Multiple Time and Status Columns

- b) To assign the desired worksheet columns to the **Selected Columns** list, select the columns in the worksheet, or select the columns from the **Data for** drop-down list.

To change your selections, select the assignment in the list and then select a new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.



Figure 128: The Survival Gehan-Breslow — Select Status Labels Panel Prompting You to Select the Status Variables

- c) Click **Next** to choose the status variables. The status variables found in the columns you selected are shown in the **Status labels:** window. Select these and click the right arrow buttons to place the event variables in the **Event:** window and the censored variable in the **Censored:** window.

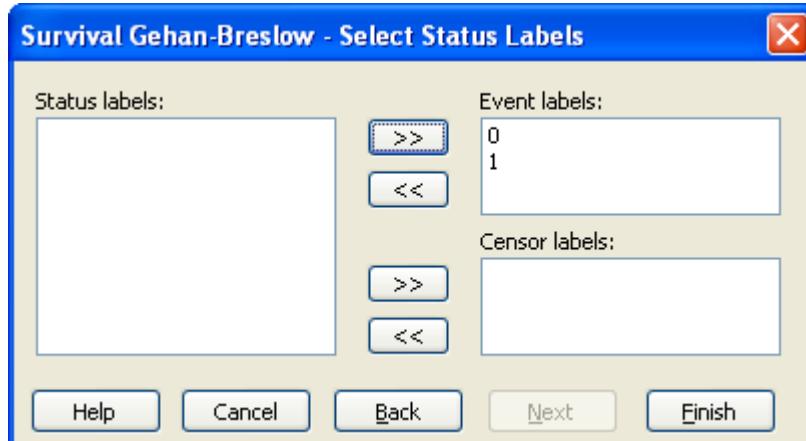


Figure 129: The Survival Gehan-Breslow — Select Status Labels Panel Showing the Results of Selecting the Status Variables

You can have more than one **Event:** label and more than one **Censored:** label. Select one **Event:** label to proceed. You don't need to select a censored variable, though, and some data sets will not have any censored values. You also don't need to select all the variables; any data associated with cleared status variables are considered missing.

- d) Click the back arrow keys to remove labels from the **Event:** and **Censored:** windows. This places them back in the **Status labels in selected columns:** window.
- SigmaPlot saves the **Event** and **Censored** labels that you selected for your next analysis. If the next data set contains exactly the same status labels, or if you are analyzing your present data set again, then the saved selections appear in the **Event** and **Censored** windows.
- e) Click **Finish** to create the survival graph and report. The results you obtain will depend on the **Test Options** that you selected.

9. If you select **Indexed Data**, then the **Select Data** panel of the Test Wizard asks you to select the three columns in the worksheet for your **Group**, **Time** and **Status**.

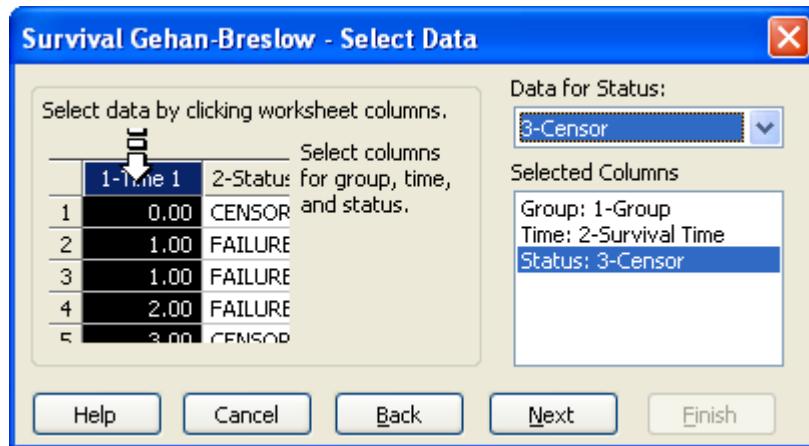


Figure 130: The Select Data Panel for Survival Gehan-Breslow Indexed Data Format Prompting You to Select Group, Time and Status Columns

- a) Click **Next** to select the groups you want to include in the analysis. If you want to analyze all groups found in the Group column then select **Select all groups**. Otherwise select groups from the **Data for Group** drop-down list. You can select subsets of all groups and select them in the order that you wish to see them in the report.

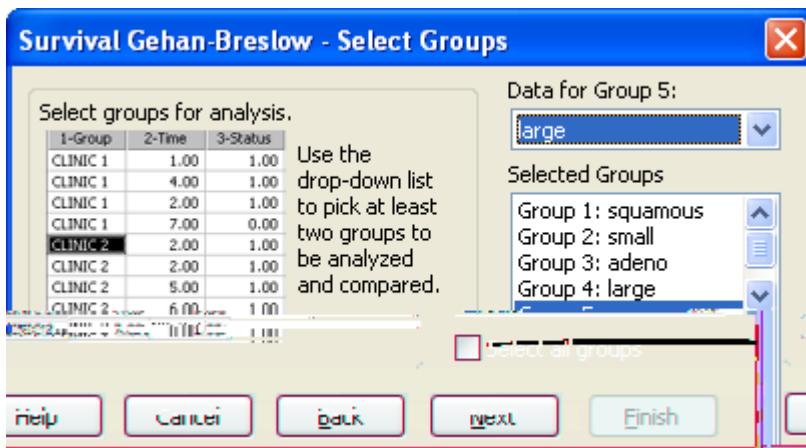


Figure 131: The Group Selection Panel for Survival Gehan-Breslow Indexed Data Format Prompting You to Select Groups to Analyze

- b) Click **Next** to select the status variables as described above and then to complete the analysis to create the report and graph.

Multiple Comparison Options

Gehan-Breslow tests the hypothesis of no differences between the several survival groups, but does not determine which groups are different, or the sizes of the differences. Multiple comparison tests isolate these differences by running comparisons between the experimental groups.

If you selected to run multiple comparisons only when the P value is significant, and Gehan-Breslow produces a P value equal to or less than the trigger P value, or you selected to always run multiple comparison in the Options for Gehan-Breslow dialog box, the multiple comparison results are displayed in the Report.

There are two multiple comparison tests to choose from for the Gehan-Breslow survival analysis.

- **Holm-Sidak.**
- **Bonferroni.**

Holm-Sidak Test

The Holm-Sidak Test can be used for both pairwise comparisons and comparisons versus a control group. It is more powerful than the Bonferroni test and, consequently, it is able to detect differences that Bonferroni test does not. It is recommended as the first-line procedure for pairwise comparison testing.

When performing the test, the P values of all comparisons are computed and ordered from smallest to largest. Each P value is then compared to a critical level that depends upon the significance level of the test (set in the test options), the rank of the P value, and the total number of comparisons made. A P value less than the critical level indicates there is a significant difference between the corresponding two groups.

Multiple Comparisons:

All Pairwise Multiple Comparison Procedures (Holm-Sidak method)

Significance level = 0.05

Comparisons	Statistic	Unadjusted P Value	Critical Level	Significant?
adeno vs. large	13.605	0.000226	0.00851	Yes
small vs. large	12.402	0.000429	0.0102	Yes
squamous vs. small	7.688	0.00556	0.0127	Yes
squamous vs. adeno	6.087	0.0136	0.0170	Yes
squamous vs. large	0.0520	0.820	0.0253	No
small vs. adeno	0.0510	0.821	0.0500	No

Figure 132: Holm-Sidak Multiple Comparison Results for VA Lung Cancer Study

Bonferroni Test

The Bonferroni test performs pairwise comparisons with paired chi-square tests. It is computationally similar to the Holm-Sidak test except that it is not sequential (the critical level used is fixed for all comparisons). The critical level for the Bonferroni test is the ratio of the family P value to the number of comparisons. It is a more conservative test than the Holm-Sidak test in that the chi-square value required to conclude that a difference exists becomes much larger than it really needs to be.

The critical level is constant at $0.05/6 = 0.00833$. Since the critical level does not increase, as it does for the Holm-Sidak test, there will tend to be fewer comparisons with significant differences. This occurs here with three significant comparisons as compared to four for the Holm-Sidak case.

Multiple Comparisons:

All Pairwise Multiple Comparison Procedures (Bonferroni method)

Significance level = 0.05

Comparisons	Statistic	Unadjusted P Value	Critical Level	Significant?
squamous vs. small	7.688	0.00556	0.00833	Yes
squamous vs. adeno	6.087	0.0136	0.00833	No
squamous vs. large	0.0520	0.820	0.00833	No
small vs. adeno	0.0510	0.821	0.00833	No
small vs. large	12.402	0.000429	0.00833	Yes
adeno vs. large	13.605	0.000226	0.00833	Yes

Figure 133: Bonferroni Multiple Comparison Results for VA Lung Cancer Study

Interpreting Gehan-Breslow Survival Results

The Gehan-Breslow survival analysis report displays information about the origin of your data, tables containing the cumulative survival probabilities for each group, summary statistics for each survival curve and the Gehan-Breslow test of significance. Multiple comparison test results will also be displayed provided significant differences were found or the Post Hoc Tests Options were selected to display them.

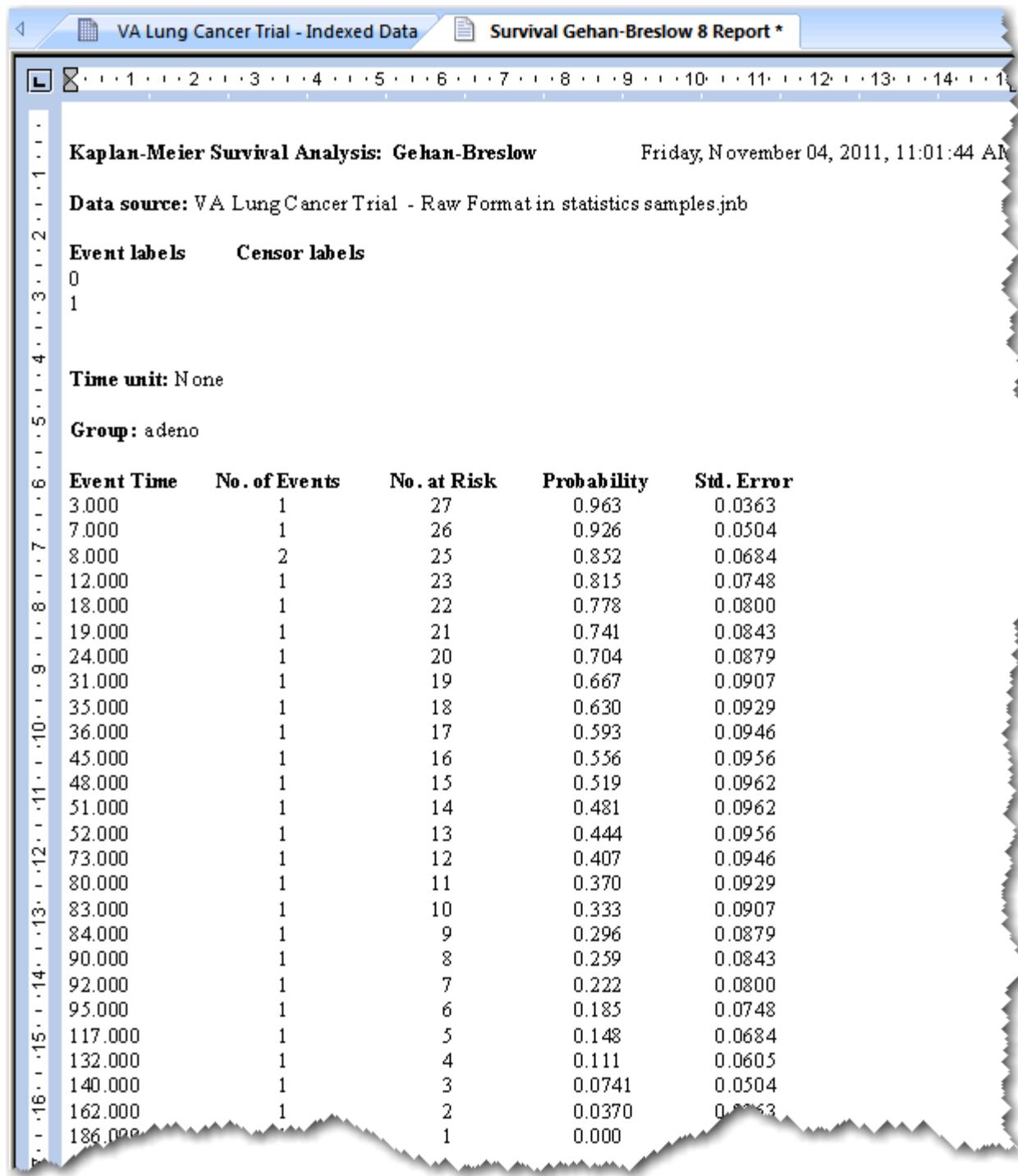


Figure 134: The Gehan-Breslow Survival Analysis Results Report

Results Explanations

The number of significant digits displayed in the report may be set in the Report Options dialog box. For more information, see [Report Graphs](#) on page 373.

Report Header Information

The report header includes the date and time that the analysis was performed. The data source is identified by the worksheet title containing the data being analyzed and the notebook name. The event and censor labels used in this analysis are listed. Also, the time units used are displayed.

Survival Cumulative Probability Table

The survival probability table lists all event times and, for each event time, the number of events that occurred, the number of subjects remaining at risk, the cumulative survival probability and its standard error. The upper and lower 95% confidence limits are not displayed but these may be placed into the worksheet. Failure times are not shown but you can infer their existence from jumps in the Number at Risk data and the summary table immediately below this table

You can turn the display of this table off by clearing this option in the Results tab of Test Options. This is useful to keep the report a reasonable length when you have large data sets.

Data Summary Table

The data summary table shows the total number of cases. The sum of the number of events, censored and missing values, shown below this, will equal the total number of cases.

Statistical Summary Table

The mean and percentile survival times and their statistics are listed in this table. The median survival time is commonly used in publications.

Gehan-Breslow Survival Graph

Visual interpretation of the survival curve is an important component of survival analysis. For this reason SigmaPlot always generates a survival curve graph. This is different from the other statistical tests where you select a report graph a posteriori.

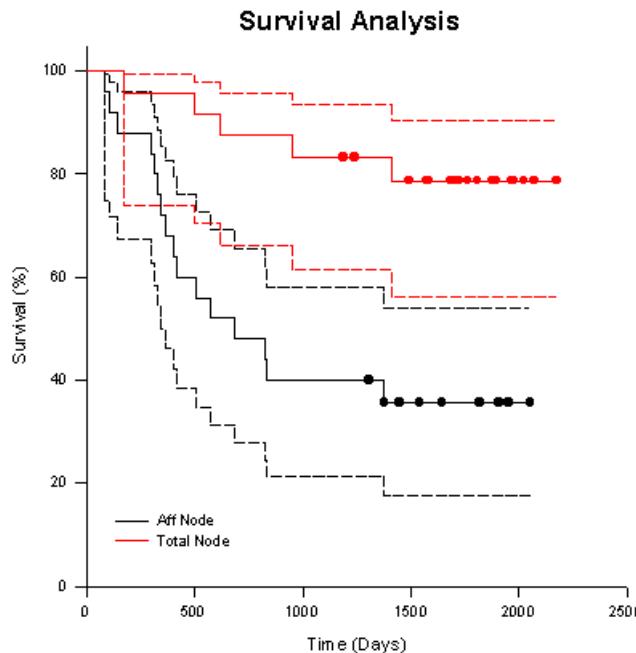


Figure 135: Gehan-Breslow Survival Curves

In the graph above, incrementing colors, percent survival and 95% confidence interval options were selected from Test Options. The Holm-Sidak test showed these two curves to be significantly different at the 0.001 level.

You can control the graph in two ways:

- You can set the graph options to become the default values until they are changed.
- After the graph is created you can modify it using SigmaPlot's Property Browser. Each object in the graph is a separate plot (for example, survival curve, failure symbols, censored symbols, upper confidence limit, etc.) so you have considerable control over the appearance of your graph.

Cox Regression

Cox Regression is a part of Survival Analysis that studies the impact of potential risk factors on the survival time of a population.

SigmaPlot has two Cox Regression tests:

- Proportional Hazards.
- Stratified Model.

About Cox Regression

Cox Regression is a part of Survival Analysis that studies the impact of potential risk factors, or *covariates*, on the survival time of a population. (The risk factors are also often called *predictors* or *explanatory variables*.)

Consider the possible effects of gender, age, and two types of drug therapy on the survival of a population suffering from some form of cancer. The survival time may decrease as age increases. Death rates among males may be higher than for females. Finally, drug A may increase survival time more than drug B. In this study, Gender, Age, and Drug Therapy are the covariates that affect the survival experience. In Cox Regression defines the model that describes the relationship between the covariates and survival time. This model helps to predict the likelihood of survival at each point in time for any values of the covariates. It also allows us to determine the significant effect of each covariate.

There are two types of covariates. The above covariates, Gender and Drug Therapy, each have two categories of non-numeric values and are called *categorical covariates*. Since the covariate Age can assume a continuous range of numeric values, it is called a *continuous covariate*. Frequently, a categorical covariate has numeric values assigned to its categories, but these values are only used for naming purposes and are not used to indicate a measurement.

The simplest way to visualize the effect of covariates on survival time is to construct a survival curve. A survival curve plots the relationship between each value of time and the probability of surviving beyond that value. This relationship is called the survival function (or survivorship function). In Kaplan-Meier survival analysis, one survival function is defined that is independent of any covariates. In Cox survival analysis, specific values for each of the covariates lead to one estimated survival function for the population. The graph of such a function is called a covariate-adjusted survival curve.

In Cox Regression, the primary object of study is the hazard function of the population, as estimated from the sampled survival data. This function is closely related to the survival function. The hazard function (sometimes known as the conditional failure rate, hazard rate, or just the hazard) is defined as the instantaneous rate of change in the likelihood of failure at each point in time, given survival up to that point. As an example, suppose h is the hazard function and suppose $h(t) = .1$ at some time t , then an interpretation of this value is that there is approximately a 10% chance that a subject will fail within the next unit time period, given the subject has survived up to time t .

Another function, the cumulative hazard function, is defined at each value of time as the integral of the hazard over all previous values of time. It provides a smoothed alternative to the hazard function as estimates of the hazard function itself can be too "noisy" for practical use. If H denotes the cumulative hazard function, then the above definitions can be used to show that the survival function S is defined at each time t by: $S(t) = \exp(-H(t))$

All of the functions discussed above are not only functions of time, but also depend upon the covariates in the survival study. In the Cox model, the hazard function assumes a specific form given by:

$$h(t, X_1, X_2, X_n) = h_0(t) \exp(b_1 X_1 + b_2 X_2 + \dots + b_n X_n)$$
 where X_1, X_2, \dots, X_n are the covariates in the study. The function h_0 is called the baseline hazard function and only depends upon time. The exponential factor on the right-hand side of the equation involves the covariates, but does not depend on time. In our implementation of Cox Regression, we are assuming that every covariate is time-independent and so its value for each subject remains constant over time (it is possible, however, to extend Cox Regression to include time-dependent covariates).

The coefficients b_1, b_2, b_n in our model are constants, independent of both time and the covariates, and their values are determined from the regression analysis by maximizing a quantity known as the partial likelihood function. The resulting values of the coefficients are called the best-fit coefficients or, sometimes, the maximum likelihood estimates. Once the coefficients are determined, there is a procedure that estimates the values of the baseline survival function at the sampled event times. The baseline survival function is defined by setting all covariates to zero. Denoting this function by S_0 , the covariate-adjusted survival functions and cumulative hazard functions are determined for each event time t by:

$$H_0(t) = -\log(S_0(t))$$

$$H(t, X_0, X_n) = H_0(t)_{\exp}(b_0 X_0 + b_n X_n)$$

$$S(t, X_0, X_n) = S_0(t)_{\exp}(b_0 X_0 + b_n X_n)$$

Our model of the hazard function shows that if there are two specifications for the values of the covariates, then the corresponding values of the hazards are proportional over time. This is the reason the Cox model is called a *proportional hazards model*. It is possible that a potential covariate for the model does not satisfy this assumption. For example, suppose we have the covariate *Gender* in a survival study. If males are dying at twice the rate of females during the first month of a study, and both genders die at the same rate during the next month of the study, then the ratio of the hazards, or the hazard ratio, for males to females is not constant over time and the proportionality assumption fails. Such a covariate cannot be included in the hazard model.

A covariate may also be omitted from the model because its value is based on the design of the study and has secondary importance as a risk factor for survival. For example, when a study is performed at two different clinics to determine the impact of age and drug therapy on patient recovery, then the variable *Clinic* is such a covariate.

Any variable whose values have been included in the survival data but is not included as a covariate in the hazard model for the reasons described above is called a *stratification variable*. Each value or level of such a variable is called a *stratum*; collectively, the levels are the *strata*. When a stratification variable is present, then the survival study is partitioned into groups, one for each stratum, where each group has its own survival function that is determined from the regression analysis. The best-fit coefficients are the same for each stratum, but the baseline time-dependent factors in the model are different.

Performing a Cox Regression Proportional Hazards Model

1. [Enter or arrange you data in the worksheet](#).
2. If desired set the Cox Regression Proportional Hazards options.
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Survival > Cox Regression > Proportional Hazards
5. Select the worksheet columns with the survival times, status values, and covariates in the **Select Data** panel.
6. [Interpret the Cox Regression results](#).

Performing a Cox Regression Stratified Model

1. Enter or arrange your data in the worksheet. [Arranging Cox Regression Data](#) on page 335
2. If desired set the Cox Regression Stratified options.
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:
Survival > Cox Regression > Stratified Model
5. Select the worksheet columns with the strata, survival times, status, and covariates in the **Select Data** panel.
6. Interpret the Cox Regression results. [Interpreting Cox Regression Results](#) on page 341

Arranging Cox Regression Data

Cox Regression in SigmaPlot consists of two separate tests, Proportional Hazards and Stratified Model. Each test requires at least three data columns: a time column, status column, and any number of covariate columns. In the Stratified Model test, you also select the worksheet column containing the strata.

Setting Cox Regression Proportional Hazards Options

Use the Cox Regression Proportional Hazards Options to:

- Specify the type of regression analysis to perform.
- Specify which results are included in the report.
- Specify attributes for the Cox result graphs.

To change the Cox Regression Proportional Hazards options:

1. Select **Cox PH Model** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.
2. Click **Options**. The **Options for Cox PH Model** dialog box appears with three tabs:
 - **Criterion**. Click the **Criterion** tab to specify variable selection and convergence options.
 - **Results**. Click the **Results** tab to specify the survival time units and to modify the content of the report and worksheet.
 - **Graph Options**. Click the **Graph Options** tab to view the graph symbol, line and scaling options. You can select additional statistical graph elements here.

SigmaPlot saves the options settings between sessions.

3. To continue the test, click **Run Test**.

The **Select Data** panel of the Test Wizard appears.

4. To accept the current settings and close the dialog box, click **OK**.

 **Note:** All options in these dialog boxes are "sticky" and remain in the state that you have selected until you change them.

Options for Cox Regression Proportional Hazard: Criterion

Variable Selection Method. There are two methods for determining which covariates will be included in the optimized regression model for the survival data.

- **Complete.** If you select Complete, the regression enters all covariates into the model and optimizes the partial likelihood function using an iterative damped-Newton method. The values of the likelihood function at all iterations appear in the report.
- **Stepwise.** If you select Stepwise, a forward stepwise procedure optimizes the partial likelihood function using the most significant covariates. Initially, no covariates are entered into the model. At each subsequent step, the covariate that contributes most to increasing the value of the likelihood function is added and any previously added covariates that are no longer significant are removed. This procedure continues until all covariates have been entered into the model or until each covariate not in model makes no significant contribution.

P-to-Enter. This value establishes the criterion for entering a covariate into the hazard model. A covariate enters the model only if there is a significant change in the likelihood function by adding the covariate. A change is significant if the probability associated with this change (the P-value) is less than the P-to-Enter value. The default value is .05.

P-to-Remove This value establishes the criterion for removing a covariate from the hazard model. A covariate is removed from the model only if there is no significant change in the likelihood function by adding the covariate. A change is not significant if the probability associated with this change (the P-value) is greater than the P-to-Remove value. The default value is .10.

To prevent the regression algorithm from cycling, the P-to-Remove value must be greater than the P-to-Enter value.

Maximum Steps. This integer value is the largest number of steps allowed for entering covariates. If this value is attained in the regression process, then the algorithm exits regardless of the stopping criteria indicated above.

Convergence. These options control the behavior of the regression algorithm

- **Tolerance**: The **Tolerance** value determines the upper limit for the quantities that measure convergence. One quantity is the coordinate of the gradient of the likelihood function with largest absolute value. The other quantity is ~~the~~ **absolute** measure of the model's coefficients between two consecutive iterations. ~~The~~ The default value is ~~1e-6~~.
- **Step Length**: The **Step Length** refers to the initial value of the parameter that controls the direction and size of the change in coefficients between two consecutive iterations. This value should not be changed.

- Specify which results are included in the report.
- Specify attributes for the Cox result graphs.

To change the Cox Regression Stratified Model options:

1. If you are going to analyze your survival curve after changing test options, and want to select your data before you create the curve, then drag the pointer over your data.
2. Select **Cox Stratified Model** from the **Select Test** drop-down list in the **SigmaStat** group on the **Analysis** tab.
3. Click **Options**. The **Options for Cox Stratified Model** dialog box appears with three tabs:
 - **Criterion.** Click the **Criterion** tab to specify variable selection and convergence options.
 - **Graph Options.** Click the **Graph Options** tab to view the graph symbol, line and scaling options. You can select additional statistical graph elements here.
 - **Results.** Click the **Results** tab to specify the survival time units and to modify the content of the report and worksheet.

SigmaPlot saves the options settings between sessions.

4. To continue the test, click **Run Test**.

Options for Cox Regression Stratified Model: Criterion

Variable Selection Method. There are two methods for determining which covariates will be included in the optimized regression model for the survival data.

- **Complete.** When this option is selected, all covariates that you select in the Test Wizard are entered into the hazard model when the regression algorithm is applied. The optimization procedure used in this algorithm is based upon an iterative damped-Newton type method to determine the best-fit coefficients in the hazard model. This is the default method.
- **Report likelihood values at each iteration.** When this option is selected, the values of the partial likelihood function for the Cox model are reported for all iterations of the optimization procedure. This option is not selected by default.
- **Stepwise.** When this option is selected, a forward stepwise procedure is applied to optimize the partial likelihood function using the most significant covariates. Initially, no covariates are entered into the model. At each subsequent step, the covariate that contributes most to increasing the value of the likelihood function is added and any previously added covariates that are no longer significant are removed. This procedure continues until all covariates have been entered into the model or until each covariate not in model makes no significant contribution.
- **P-to-Enter.** This value establishes the criterion for removing a covariate from the hazard model. A covariate is removed from the model only if there is no significant change in the likelihood function by adding the covariate. A change is not significant if the probability associated with this change (the P-value) is greater than the P-to-Remove value. The default value is .10.

To prevent the regression algorithm from cycling, the P-to-Remove value must be greater than the P-to-Enter value.

- **P-to-Remove.** This value establishes the criterion for removing a covariate from the hazard model. A covariate is removed from the model only if there is no significant change in the likelihood function by adding the covariate. A change is not significant if the probability associated with this change (the P-value) is greater than the P-to-Remove value. The default value is .10.

To prevent the regression algorithm from cycling, the P-to-Remove value must be greater than the P-to-Enter value.

- **Maximum Steps.** This integer value is the largest number of steps allowed for entering covariates. If this value is attained in the regression process, then the algorithm exits regardless of the stopping criteria indicated above.

Convergence. These options control the behavior of the regression algorithm.

- **Tolerance.** This value determines the upper limit for the two quantities that measure convergence. One quantity is the coordinate of the gradient of the likelihood function with largest absolute value. The other quantity is a distance measure of the model's coefficients between two consecutive iterations. The default value is 1e-008.
- **Step Length.** This value refers to the initial value of the parameter that controls the direction and size of the change in coefficients between two consecutive iterations. This value should not be changed unless there is a problem with obtaining convergence. The default value is 1.0.

Maximum Iterations. This integer value is the largest number of improved changes to the coefficients that are allowed in order to obtain convergence. If this value is exceeded in the regression process, then the algorithm exits regardless of whether the convergence criterion (determined by the **Tolerance**) has been satisfied. The default value is 20.

Options for Cox Regression Stratified Model: Graph Options

Status Symbols.

- **Censored.** Symbols for censored observations on the covariate-adjusted survival curves are graphed by default. Clear this option to not display the censored symbols.
- **Failures.** Symbols for failures (events) always occupy the inside corners of the steps in each covariate-adjusted survival curve. This option is cleared by default.

Group Color.

2. If you want to select your data before you run the test then drag the pointer over your data. Your data must be selected in contiguous columns with the (survival) Time column first, followed by the Status column, and then one or more Covariate columns. From the menus select:
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Survival > Cox Regression > Proportional Hazards

The **Cox PH Model - Select Data** panel of the **Test Wizard** appears, prompting you to select your data columns. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

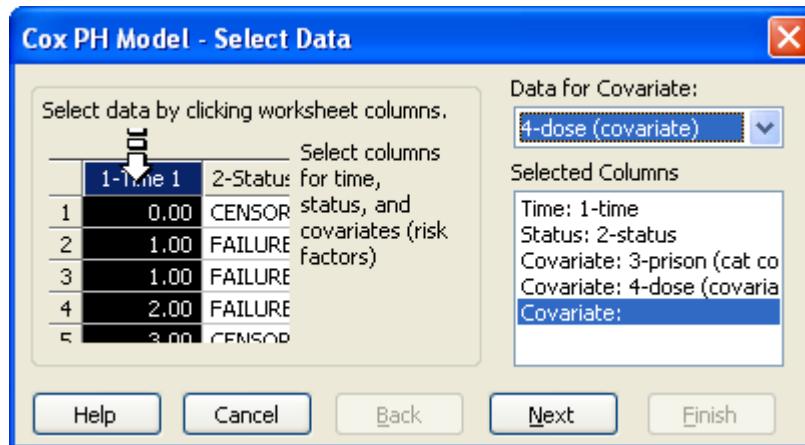


Figure 136: The Cox PH Model - Select Data Panel Prompting You to Select Time, Status, and Covariate Columns

- a) **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for** drop-down list. The first selected column is assigned to the first row (Time) in the **Selected Columns** list, the next selected column is assigned to the next row (Status) in the list, and then the next column is assigned to the next row (Covariate). The number or title of selected columns appears in each row.
- b) **To change your selections**, select the assignment in the list and then select a new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
5. Click **Next** to choose the **Categorical Covariates**. The drop-down list displays all of the covariates that you had selected on the Select Data panel. To select a categorical covariate, the clicks an item in this list and the selection will be entered in the Selected Covariates list. If a covariate column is not listed as categorical and it contains a non-numeric data entry, then this entry is treated as missing. Making a selection on this panel is optional as there may be no categorical covariates in the study.

6. Click **Next** to choose the status variables. The status variables found in the columns you selected are shown in the Status labels in selected columns window. Select these and click the right arrow buttons to place the event variables in the **Event** window and the censored variable in the **Censored** window.



Figure 137: The Cox PH Model - Select Status Labels Panel Prompting You to Select the Status Variables.

You can have more than one **Event** label and more than one **Censored** label. You must select one **Event** label in order to proceed. You need not select a censored variable, though, and some data sets will not have any censored values. You need not select all the variables; any data associated with cleared status variables will be considered missing.



7. Click the back arrows to remove labels from the **Event** and **Censored** windows. This places them back in the Status labels in selected columns window.
- SigmaPlot saves the **Event** and **Censored** labels that you selected for your next analysis. If the next data set contains exactly the same status labels, or if you are reanalyzing your present data set, then the saved selections appear in the **Event** and **Censored** windows.
8. Click **Finish** to create the survival graph and report. The results you obtain depend on the Test Options that you selected.

Running a Cox Regression Stratified Model

To run a Cox Regression Stratified Model analysis you need to select strata, survival time, status, and covariate data columns to analyze. Use the Select Data panel of the Test Wizard to select these columns from the worksheet.

The Strata column contains the various levels of the stratification variable that are used to separate the survival study into groups, each of which has its own baseline survival curve. The term *baseline* refers to computations that result when all covariates are set to zero.

To run a Cox Regression Stratified Model analysis:

1. Specify any options for your graph and report.
2. If you want to select your data before you run the test then drag the pointer over your data. Your data must be selected in contiguous columns with the Strata column first, followed by (survival) Time column, the Status column, and then one or more Covariate columns.
3. Select the **Analysis** tab.
4. In the **SigmaStat** group, from the **Tests** drop-down list, select:

Survival > Cox Regression > Stratified Model

The **Cox Stratified Model - Select Data** panel of the **Test Wizard** appears prompting you to select your data columns. If you selected columns before you chose the test, the selected columns appear in the **Selected Columns** list.

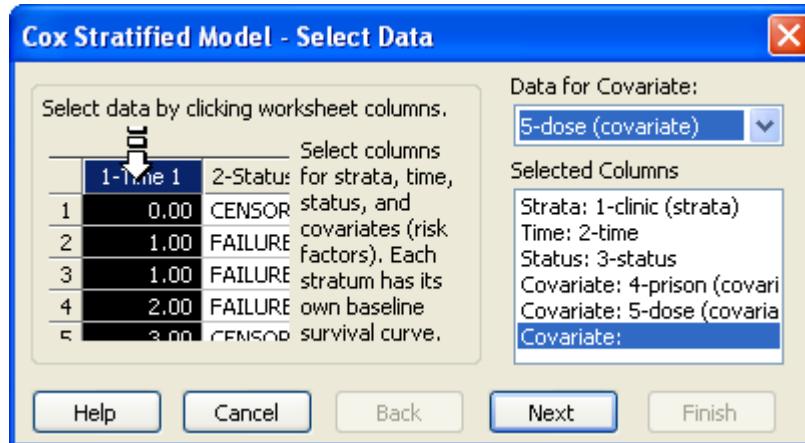


Figure 138: The Select Data for Cox Stratified Model Panel Prompting You to Select Time, Status, and Covariate Columns

- a) **To assign the desired worksheet columns to the Selected Columns list**, select the columns in the worksheet, or select the columns from the **Data for** drop-down list. The first selected column is assigned to the first row (Strata) in the **Selected Columns** list, the next selected column is assigned to the next row (Time) in the list, the next selected column is assigned to the next row (Status), and then the next column is assigned to the next row (Covariate). The number or title of selected columns appears in each row.
 - b) **To change your selections**, select the assignment in the list and then select a new column from the worksheet. You can also clear a column assignment by double-clicking it in the **Selected Columns** list.
5. Click **Next** to choose the **Categorical Covariates**. The drop-down list displays all of the covariates that you selected on the Select Data panel. To select a categorical covariate, the click an item in this list and the selection will be entered in the Selected Covariates list. If a covariate column is not listed as categorical and it contains a non-numeric data entry, then this entry is treated as missing. Making a selection on this panel is optional as there may be no categorical covariates in the study.
 6. Click **Next** to choose the status variables. The status variables found in the columns you selected are shown in the Status labels in selected columns window. Select these and click the right arrow buttons to place the event variables in the **Event** window and the censored variable in the **Censored** window.
- You can have more than one **Event** label and more than one **Censored** label. You must select one Event label in order to proceed. You need not select a censored variable, though, and some data sets will not have any censored values. You need not select all the variables; any data associated with cleared status variables will be considered missing.
7. Click **Finish** to create the survival graph and report. The results you obtain depend on the Test Options that you selected.

Interpreting Cox Regression Results

The Cox Regression report displays statistical information about the input data, the best-fit results for the hazard model, the results of hypothesis testing, confidence intervals, and covariate-adjusted survival probabilities. For

more in-depth discussions of the statistics and performance measures reported for Cox Regression see Hosmer & Lemeshow or Kleinbaum.

The screenshot shows a software window titled 'Cox Stratified Model 9 Report'. The report details a Cox Regression analysis on the 'addicts - stratified' dataset. It includes sections for Event labels (censored, failure), Time unit (None), Stratification Variable (clinic), and Categorical Covariates (prison, dose). A table summarizes the data by strata (1.000, 2.000, Total) across columns for Cases, Missing, Events, Censored, and % Censored.

Strata	Cases	Missing	Events	Censored	% Censored
1.000	161	0	161	0	0
2.000	77	0	77	0	0
Total	238	0	238	0	0

Results Explanations

In addition to the numerical results, expanded explanations of the results may also appear. You can turn off this text on the Options dialog box. You can also set the number of decimal places to display. For more information, see [Report Graphs](#) on page 373.

Header. This includes the name of the test, date stamp, and data source, as for all other tests.

Event and Censor Labels. is a listing of the labels that you've selected in the Test Wizard. There can be more than one label of each type.

Time Unit. This information comes from a setting in the Test Options dialog box and is used to indicate the unit of survival time on result graphs.

Stratification Variable. This is the worksheet column (by title) to stratify the data if using the Stratified Model test. This section does not appear if you're using the Proportional Hazards test.

Basic summary of time-event data over strata. This is a table whose first column is a list of the strata for the stratification variable. The remaining columns have integer entries and are titled: Cases, Missing, Events, Censored, and % Censored. The last row of the table gives the total over all strata. If there is no stratification variable, then the table has one row of data.

Regression analysis. This section contains the results of maximizing the partial log-likelihood function for the Cox Proportional Hazards Model to obtain the maximum likelihood estimates of the coefficients. The partial likelihood function is based on the Breslow method for resolving ties.

The coefficient values found by the regression are used to represent the hazard as a function of time and the covariates. Each categorical covariate in the model is replaced by one or more reference coded dummy variables, each with its own coefficient, before the regression analysis is performed. The optimization process uses an iterative Damped-Newton method with zero as the starting value for each coefficient.

The output of the analysis depends upon the variable selection method that is specified in the Test Options dialog, either Complete or Stepwise. The default method is Complete, where all covariates that you selected are used to model the hazard function. When the default method is used, the results show the maximum value of the log-likelihood function, the number of iterations to convergence, and the tolerance used in the criterion for convergence.

If the Stepwise method is chosen, then only the covariates that contribute most to increasing the value of the likelihood function are included in the hazard model. The included covariates are determined using a step-by-step procedure. More details on the stepwise-regression results are given later.

Testing the Global Null Hypothesis. This is the hypothesis that all coefficients in the hazard model are zero. SigmaPlot provides two tests: the (partial) Likelihood Ratio test and the Global Chi-Square test (also called Score test). The statistic for each test has a chi-square distribution with p degrees of freedom, where p is the number of covariates. The default significance level of the test is .05, which can be changed on the Report tab of the Tools/Options dialog box. A significant result means that at least one of the covariates has a significant effect on survival time. If the result is not significant, then no covariate significantly influences the survival time and a Kaplan-Meier analysis should be considered for computing survival probabilities.

Both tests are used by many survival software applications and they usually agree in their determination of significance. In the event they disagree, then the result of the Likelihood Ratio test should be used as it is more accurate.

Model Estimates. This is a table of the best-fit coefficient values and their basic statistics. It has five columns. The

Creating a Cox Regression Graph

1. Select the **Report** tab.
2. In the **Results** group, click **Create Result Graph**.

The **Create Result Graph** dialog box appears displaying the available graphs for the Cox Regression report.

3. Select the report graph you want to create, then click **OK**, or double-click the graph in the list.

The **Covariate Values for Plot** dialog box appears, in which you can select the covariates values to use to specify the graph data.

4. From the **Select value type** drop-down list, select either:
 - Mean
 - Median
 - Baseline
 - User-Defined

The names of the covariates appear in the first column of two in the list below. These are the covariates that you specified earlier the Test Wizard. The values in the second column correspond to the value type option that you selected in the **Select value type** drop-down list.

If you select **User-Defined**, the covariate values default to all zeroes (same meaning **Baseline**), and you can enter values into the **Enter value:** box.

5. To change a covariate's value, select a covariate from the **Covariate** column, enter a new value in the **Enter value:** box, and click **Change**. After selecting the value type and completing any changes to the covariate values, click **Close**.

A graph appears.

Survival Curve Graph Examples

You can modify survival curve attributes using Test Options or Graph Properties.

Failures, Censored Values, and Ties

The relationship between failures, censored values and ties effects the shape of a survival curve. Some rules that characterize survival curves are:

- A step decrease occurs at every failure
- Larger step decreases result from multiple failures occurring at the same time (ties).
- The curve does not decrease at a censored value.
- Tied failure (and failure and censored) values superimpose at the appropriate inside corner of the step survival curve.
- It is useful to display symbols for censored values.
- It is not necessary to display symbols for failures.
- The survival curve decreases to zero if the largest survival time is a failure.
- Censored values cause the survival curve to decrease more slowly.

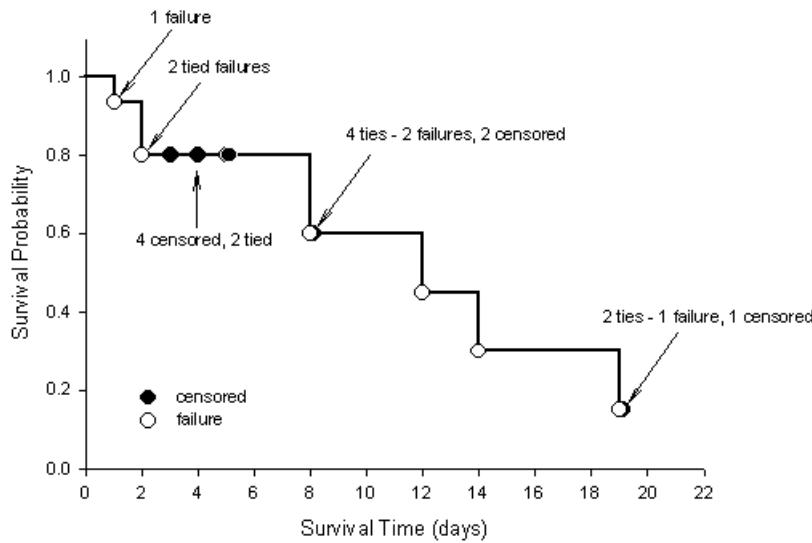


Figure 139: A contrived survival curve with various combinations of failures, censored values and tied data that graphically shows the effects of these rules.

Failures and censored values are shown above as open and filled circles, respectively. A single failure is shown at time = 1.0. It is located at the inner corner of the step curve. All failures occur at the inner corners so it is not necessary to display failure symbols. You can display failure symbols in SigmaPlot, but by default they are not visible. Two tied failures are shown at time = 2.0. They superimpose at the inner corner of the step that has decreased roughly twice as much as the step for a single failure. Four censored values, two of which are tied, are shown in the time interval between 2.0 and 8.0. Censored values do not cause a decrease in the survival curve and nothing unusual occurs at tied censor values. Four tied values, two failures and two censored, are shown at time = 8.0 (the censored values are slightly displaced for clarity). They occur at the inside corner of the step since that is where failures are located. The censored value at time = 19.0 prevents the survival curve from touching the X-axis.

Editing Survival Graphs Using Graph Properties

This example shows modifications made from Graph Properties to a survival curve with both symbol types and 95% confidence intervals.

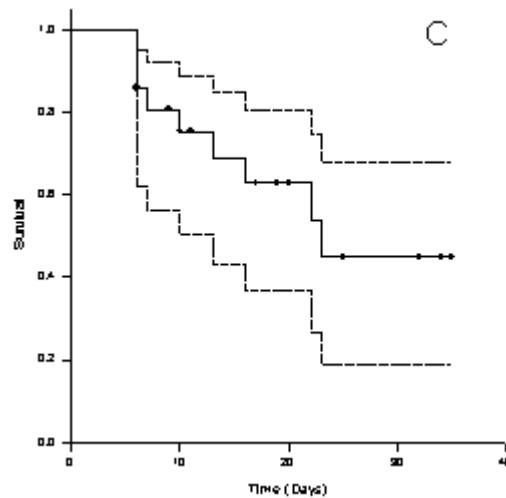


Figure 140: Survival Curve with both Symbol Types and 95% Confidence Intervals

The confidence interval lines were changed from small gray dashed to solid blue. The censored symbol type was also changed from a solid circle to a square.

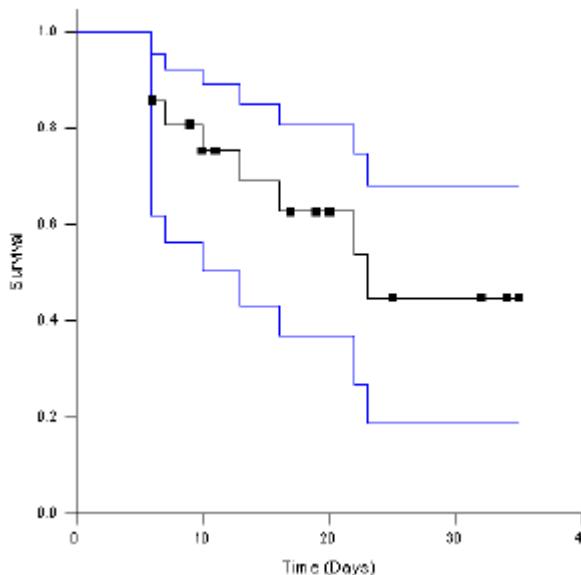


Figure 141: Modifications made using Graph Properties to a Survival Curve with both Symbol Types and 95% Confidence Interval

Using Test Options to Modify Graphs

The examples below show four variations that can be achieved by modifying the test options for survival curves. Once you've selected a test from the **SigmaStat** toolbar, you can open this dialog box by selecting from the menus:

SigmaStat > Current Test Options

The options used to create the examples below appear on the **Graph Options** tab of any of the **Options for Survival** dialog boxes.

Survival curve with censored symbols. Under **Status Symbols**, select **Censored**.

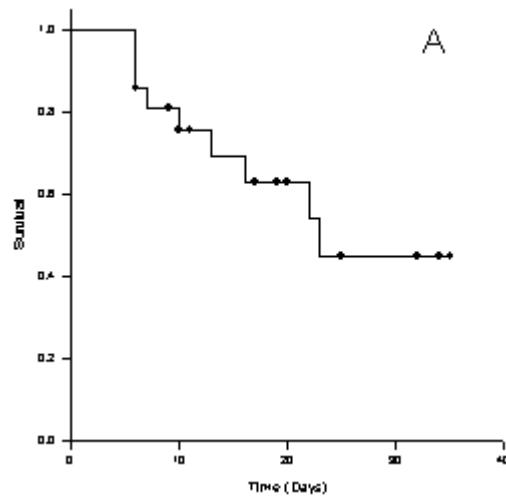


Figure 142: Survival Curve with Censored Symbols

Survival curve with censored and failure symbols. Under **Status Symbols**, select both **Censored** and **Failures**.

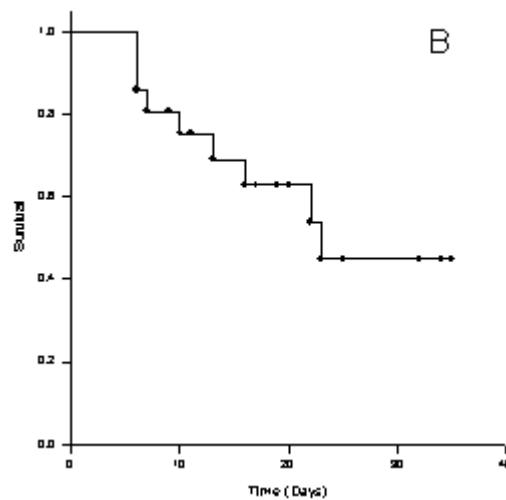


Figure 143: Survival Curve with Censored and Failure Symbols

Survival curve with both symbol types and 95% confidence intervals. To add 95% confidence intervals:

1. Select **Additional Plot Statistics**.

2. From the Type drop-down list, select **95% Confidence Intervals**.

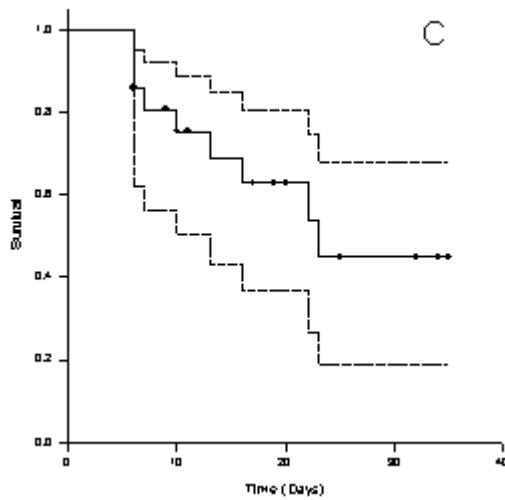


Figure 144: Survival Curve with both Symbol Types and 95% Confidence Intervals

Survival curve with standard error bars. To add standard error bars:

3. Select **Additional Plot Statistics**.
4. From the **Type** drop-down list, select **Standard Error Bars**.

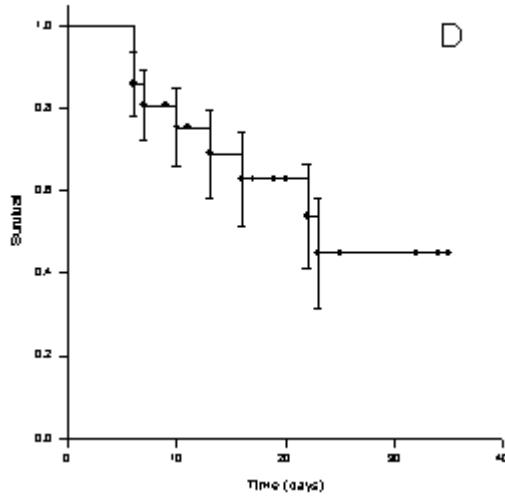


Figure 145: Survival Curve with Standard Error Bars

Chapter

11

Computing Power and Sample Size

Topics:

- [About Sample Size](#)
- [About Power](#)

SigmaPlot provides two experimental design aids: experimental power and sample size computations. Use these procedures to determine the power of an intended test or to determine the minimum sample size required to achieve a desired level of power.

Power and sample size computations are available for:

- Unpaired and Paired t-tests
- A z-test comparison of proportions
- One way ANOVAs
- Chi-Square Analysis of Contingency Tables
- Correlation Coefficient

About Sample Size

You can estimate how big the sample size has to be in order to detect the treatment effect or difference with a specified level of statistical significance and power. All else being equal, the larger the sample size, the greater the power of the test.

Determining the Minimum Sample Size for a t-Test

You can determine the minimum sample size for an intended t-test. Unpaired t-tests are used to compare two different samples from populations that are normally distributed with equal variances among the individuals. For more information, see [Unpaired t-Test](#) on page 57.

To determine the sample size for a t-test, you need to specify the:

- Expected difference of the means of the groups you want to detect.
- Expected standard deviation of the underlying populations.
- Desired power of the t-test.
- Alpha level (α) used for determining the sample size.

To determine the sample size of a t-test:

1. With the worksheet in view, click the **Analysis** tab.
2. In the **SigmaStat** group, select:

Sample Size > t-test

The **t-test Sample Size** dialog box appears.

3. Enter the size of the difference between the means of the two groups to be detected in the **Expected Difference in Means** box. This can be size you expect to see, as determined from previous samples or experiments, or just an estimate.
4. Enter the estimated standard deviation of the underlying population in the **Expected Standard Deviation** box. This can be size you expect to see, as determined from previous samples or experiments, or just an estimate.
5. Enter the desired power, or test sensitivity in the **Desired Power** box. Power is the probability that the t-test will detect a difference if there really is a difference. The closer the power is to 1, the more sensitive the test. Traditionally, you want to achieve a power of 0.80, which means that there is an 80% chance of detecting a difference with 1- confidence (for example, a 95% confidence when $\alpha = 0.05$).
6. Enter the desired alpha level in the **Alpha** box. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference.

The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive (a Type I error).

7. Click = to see the required sample size for a t-test at the specified conditions. The sample size calculation appears at the top of the dialog. The sample size is the size of each of the groups. If desired, you can change any of the settings and click = again to view the new sample size as many times as desired.

8. Click **Save to Report** to save the sample size computation settings and resulting sample size to the current report.

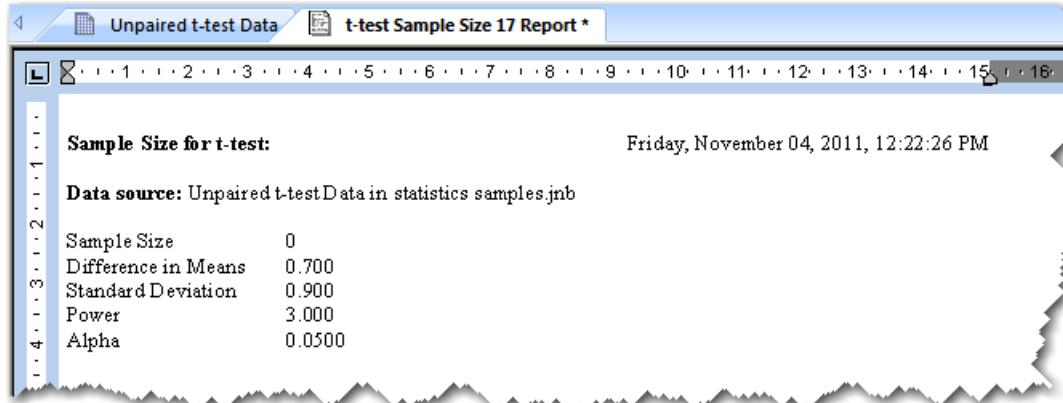


Figure 146: The t-test Sample Size Results Viewed in the Report

9. Click **Close** to exit from t-test sample size computation.

Determining the Minimum Sample Size for a Paired t-Test

You can determine the sample size for a Paired t-test. Use Paired t-tests to see if there is a change in the same individuals before and after a single treatment or change in condition. The sizes of the treatment effects are assumed to be normally distributed. For more information, see [Paired t-Test](#) on page 141.

To determine the sample size for a Paired t-test, you need to estimate the:

- Difference of the means you wish to detect.
- Estimated standard deviation of the changes in the underlying population.
- Desired power or sensitivity of the test.
- Alpha (α) used to determine the sample size.

To find the sample size for a Paired t-test:

1. With the worksheet in view, click the **Analysis** tab.
2. In the **SigmaStat** group, select:

Sample Size > Paired t-test

The **Paired t-test Sample Size** dialog box appears.

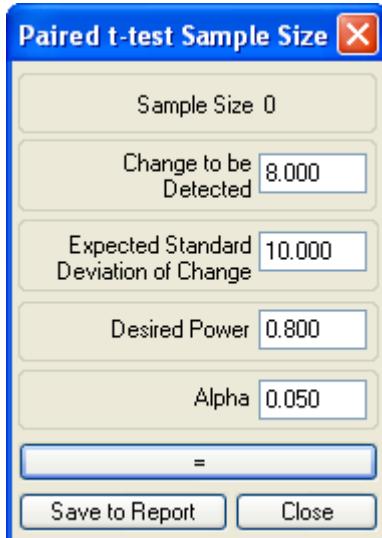


Figure 147: The t-test Sample Size Results Viewed in the Report

3. Enter the size of the change before and after the treatment in the **Change to be Detected** box. This can be size of the treatment effect you expect to see, as determined from previous experiments, or just an estimate.
 4. Enter the size of standard deviation of the change in **Expected Standard Deviation of Change**. This can be size you expect to see, as determined from previous experiments, or just an estimate.
 5. Enter the desired power, or test sensitivity. Power is the probability that the paired t-test will detect an effect if there really is an effect. The closer the power is to 1, the more sensitive the test. Traditionally, you want to achieve a power of 0.80, which means that there is an 80% chance of detecting an effect with $1 - \alpha$ confidence (for example, a 95% confidence when $\alpha = 0.05$).
 6. Enter the desired alpha level. Alpha (α) is the acceptable probability of incorrectly concluding that there is an effect. The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant treatment difference when $P < 0.05$.
- Smaller values of α result in stricter requirements before concluding there is a significant effect, but a greater possibility of concluding there is no effect when one exists (a Type II error). Larger values of α make it easier to conclude that there is an effect, but also increase the risk of reporting a false positive (a Type I error).
7. Click = to see the required sample size for a Paired t-test at the specified conditions. The sample size calculation appears at the top of the dialog box. If desired, you can change any of the settings and click = again to view the new sample size as many times as desired.
 8. Click **Save to Report** to save the sample size computation settings and resulting sample size to the current report.

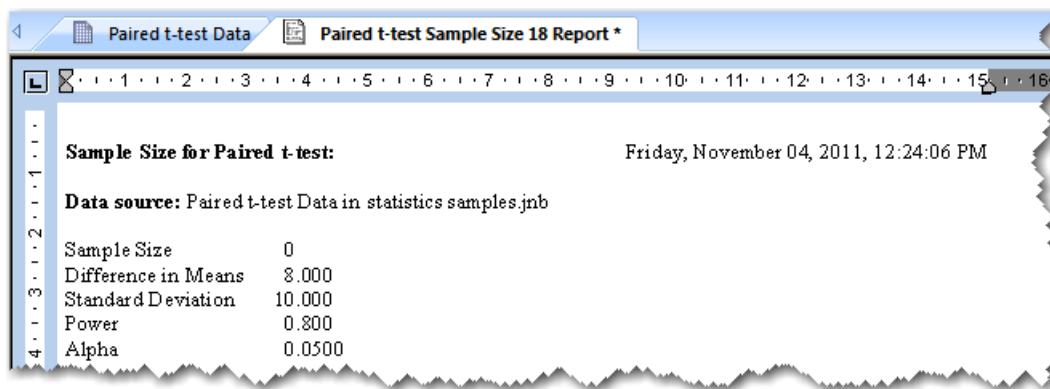


Figure 148: The Paired t-test Sample Size Dialog Box

9. Click **Close** to exit from paired t-test sample size computation.

Determining the Minimum Sample Size for a Proportions Comparison

You can determine the sample size for a z-test comparison of proportions. A comparison of proportions compares the difference in the proportion of two different groups that falls within a single category. For more information, see [Comparing Proportions Using the z-Test](#) on page 188.

To determine the sample size for a proportion comparison, you need to specify the:

- Proportion of each group that falls within the category.
- Desired power or sensitivity of the test.
- Alpha (α) used to determine the sample size.

To find the sample size for a z-test proportion comparison:

1. With the worksheet in view, click the **Analysis** tab.

2. In the **SigmaStat** group, select:

Sample Size > Proportions

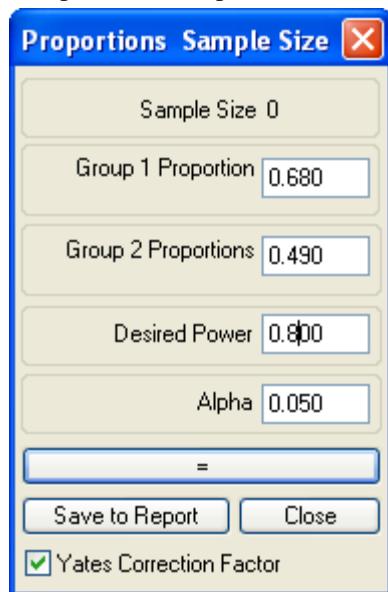


Figure 149: The Proportions Sample Size Dialog Box

The **Proportions Sample Size** dialog box appears.

3. Enter the expected proportions that fall into the category for each group in the **Group 1 and 2 Proportion** boxes. This can be the distribution you expect to see, as determined from previous experiments, or just an estimate.
4. Enter the desired power, or test sensitivity. Power is the probability that the proportion comparison will detect a difference if there really is a difference in proportion. The closer the power is to 1, the more sensitive the test. Traditionally, you want to achieve a power of 0.80, which means that there is an 80% chance of detecting a difference with $1 - \alpha$ confidence (for example, a 95% confidence when $\alpha = 0.05$).
5. Enter the desired alpha level. Alpha (α) is the acceptable probability of incorrectly concluding that there is an effect. The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant distribution difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference in distribution when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive (a Type I error).

6. Click = to see the required sample size for a proportion comparison at the specified conditions. The calculated sample size appears at the top of the dialog. If desired, you can change any of the settings and click = again to view the new sample size as many times as desired.

 **Note:** The Yates correction factor is used if this option was selected in the **Options for z-Test** dialog box.

7. Click **Save to Report** to save the sample size computation settings and resulting sample size to the current report. The estimated sample size is the sample size for each group.

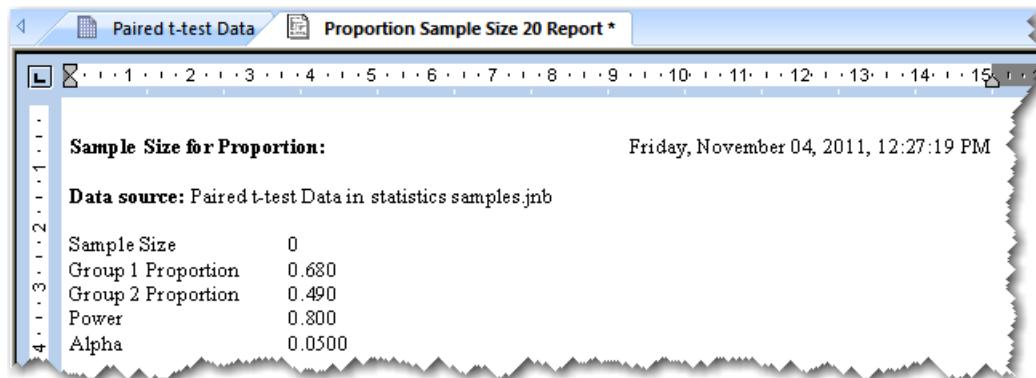


Figure 150: The Proportions Sample Size Results Viewed in the Report

8. Click **Close** to exit from proportion comparison sample size computation.

Determining the Minimum Sample Size for a One Way ANOVA

You can determine the group sample size for a One Way ANOVA (analysis of variance). One Way ANOVAs are used to see if there is a difference among two or more samples taken from populations that are normally distributed with equal variances among the individuals. For more information, see [One Way Analysis of Variance \(ANOVA\)](#) on page 73.

To determine the sample size for a One Way ANOVA, you need to specify the:

- Minimum difference in between group means to be detected.
- Estimated standard deviation of the underlying populations.
- Number of groups.
- Desired power or sensitivity of the ANOVA.
- Alpha (α) used to determine the sample size.

To find the sample size for a One Way ANOVA:

1. With the worksheet in view, click the **Analysis** tab.

2. In the **SigmaStat** group, select:

SampleSize > ANOVA

The ANOVA Sample Size dialog box appears.

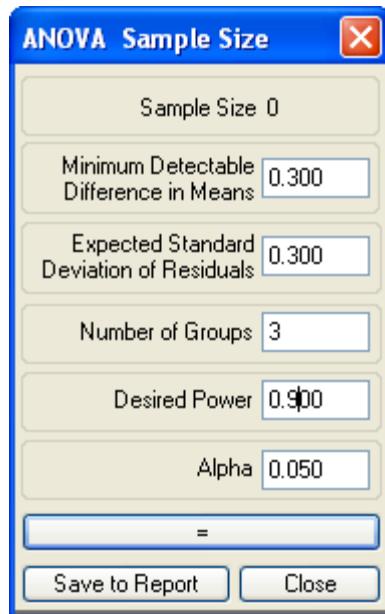


Figure 151: The ANOVA Sample Size Dialog Box

- Enter the size of the minimum expected difference of group means in the **Minimum Detectable Difference** box. This can be size of a difference you expect to see, as determined from previous experiments, or just an estimate. The minimum detectable difference is the minimum difference between the largest and smallest means.
- Enter the size of standard deviation of the residuals. This can be size you expect to see, as determined from previous experiments, or just an estimate. Note that one way ANOVA assumes that the standard deviations of the underlying normally distributed populations are equal. Then enter the expected number of groups.
- Enter the desired power, or test sensitivity. Power is the probability that the ANOVA will detect a difference if there really is a difference among the groups. The closer the power is to 1, the more sensitive the test. Traditionally, you want to achieve a power of 0.80, which means that there is an 80% chance of detecting a difference with $1 - \alpha$ confidence (for example, a 95% confidence when $\alpha = 0.05$).
- Enter the desired alpha level. Alpha (α) is the acceptable probability of incorrectly concluding that there is an effect. The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$. Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive (a Type I error).
- Click = to see the required sample size for a One Way ANOVA at the specified conditions. The sample size calculation appears at the top of the dialog. The sample size is the size of each group. If desired, you can change any of the settings and click = again to view the new sample size as many times as desired.

8. Select **Save to Report** to save the sample size computation settings and resulting sample size to the current report, and then click **Close**.

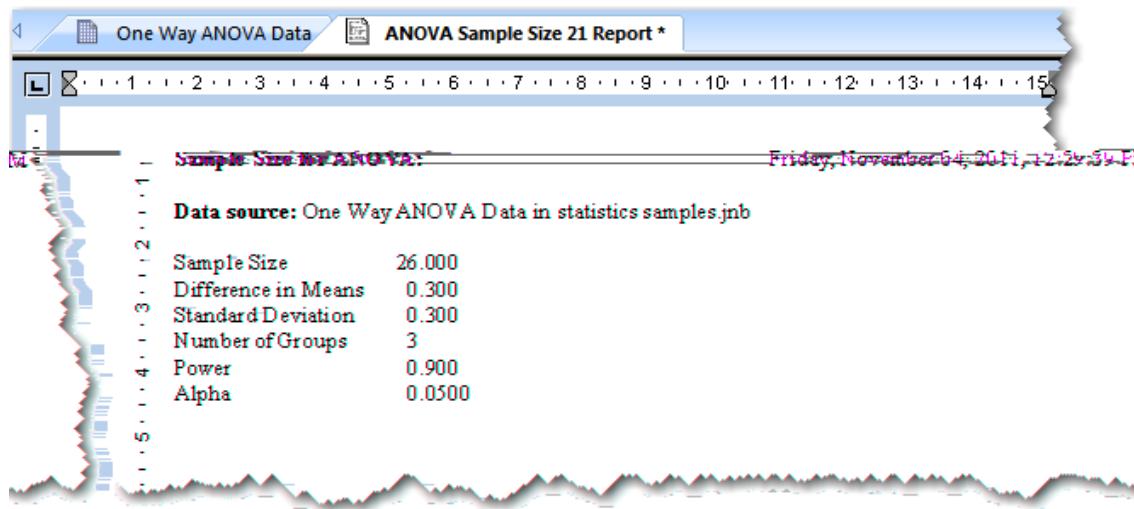


Figure 152: The ANOVA Sample Size Results Viewed in the Report

Determining the Minimum Sample Size for a Chi-Square Test

You can determine the sample size for a chi-square χ^2 analysis of a contingency table. A Chi-square test compares the difference between the expected and observed number of individuals of two or more different groups that fall within two or more categories.

Table 22: The Contingency Table with Expected Numbers of Observations of Two Groups in Three Categories

Group	Categories		
	Category 1	Category 2	Category 3
Group 1	15	15	35
Group 2	15	30	10

The sample size for a chi-square analysis contingency table is determined by the estimated relative proportions in each category for each group. Because SigmaPlot uses numbers of observations to compute these estimated proportions, you need to enter a contingency table in the worksheet containing the estimated number of observations before you can compute the estimated proportions.

To find the sample size for a Chi-square test:

1. Enter a contingency table into the worksheet by placing the estimated number of observations for each table cell in a corresponding worksheet cell.

	1-Tundra	2-Foothills	3-Treeline	4
1	125.0000	16.0000	6.0000	T
2	7.0000	19.0000	117.0000	T
3				T
4				T

Figure 153: Contingency Table Data Entered into the Worksheet

The worksheet rows and columns correspond to the groups and categories. The number of observations must always be an integer.

Note that the order and location of the rows or columns corresponding to the groups and categories is unimportant. You can use the rows for category and the columns for group, or vice versa.

2. Select the **Analysis** tab.
3. In the **SigmaStat** group, select:

Sample Size > Chi-Square

The **Select Data** panel of the Test Wizard appears.

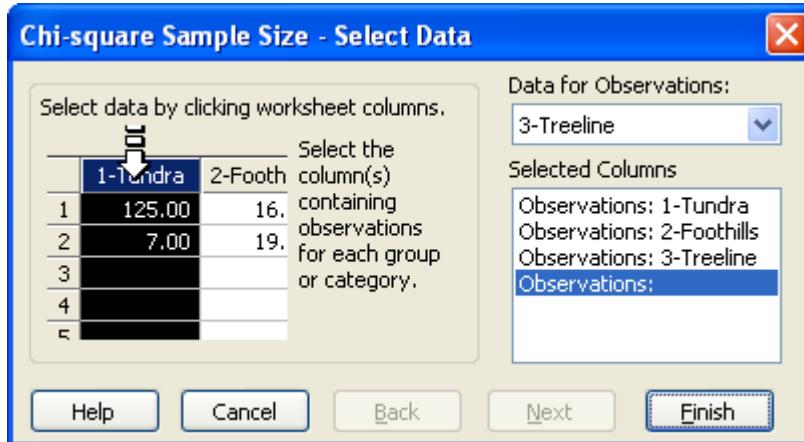


Figure 154: The Select Data Panel of the Test Wizard

4. Select the columns of the contingency table from the worksheet as prompted.

5. Click **Finish** when you have selected all three columns.

The **Chi-Square Sample Size** dialog box appears.

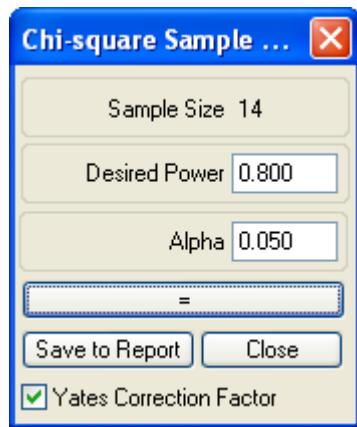


Figure 155: The Chi-square Sample Size Dialog Box

6. Enter the desired power, or test sensitivity. Power is the probability that the chi-square test will detect a difference in observed distribution if there really is a difference. The closer the power is to 1, the more sensitive the test. Traditionally, you want to achieve a power of 0.80, which means that there is an 80% chance of detecting an difference with $1 - \alpha$ confidence (for example, a 95% confidence when $\alpha = 0.05$).
 7. Enter the desired alpha level. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.
- Smaller values of α result in stricter requirements before concluding there is a significant difference, but increase the possibility of concluding there is no effect when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference, but also increase the possibility of concluding there is an effect when none exists.
8. Click **=** to see the required sample size for a Chi-Square test at the specified conditions. The sample size calculation appears at the top of the dialog. If desired, you can change any of the settings and click **=** again to view the new sample size as many times as desired. However, if you want to change the number of observations per category, you need to select **Close**, edit the table, then repeat the sample size computation.
 9. Click **Save to Report** to save the sample size computation settings and resulting sample size to the current report.

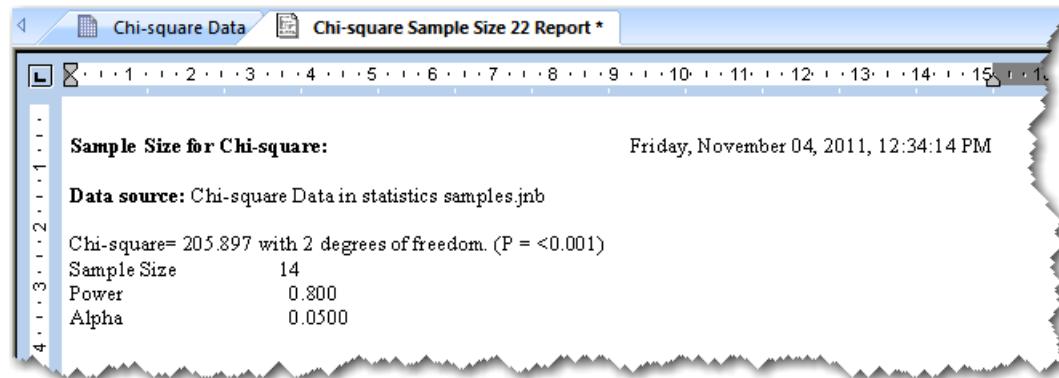


Figure 156: The Chi-square Sample Size Computation Results Viewed in the Report

10. Click **Close** to exit from Chi-Square test sample size computation.

Determining the Minimum Sample Size to Detect a Specified Correlation

You can determine the sample size necessary to detect a specified Pearson Product Moment Correlation Coefficient R . A correlation coefficient quantifies the strength of association between the values of two variables. A correlation

coefficient of 1 means that as one variable increases, the other increases exactly linearly. A correlation coefficient of -1 means that as one variable increases, the other decreases exactly linearly. For more information, see [Pearson Product Moment Correlation](#) on page 294.

To determine the sample size necessary to detect a specified correlation coefficient, you need to specify the:

- Expected value of the correlation coefficient.
- Desired power or sensitivity of the test.
- Alpha (α) used to determine the sample size.

To find the sample size required for a specific correlation coefficient:

1. With the worksheet in view, select the **Analysis** tab.

2. In the **SigmaStat** group, select:

Sample Size > Correlation

The **Correlation Sample Size** dialog box appears.

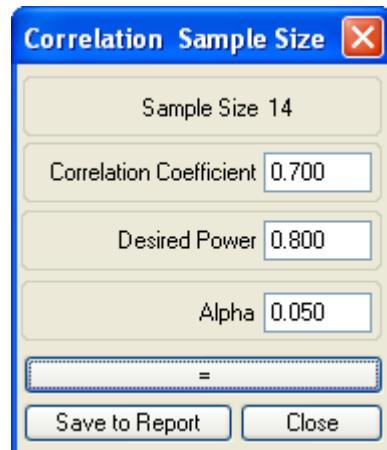


Figure 157: The Correlation Sample Size Dialog Box

3. Enter the expected correlation coefficient in the **Correlation Coefficient** box. This can be the correlation coefficient you expect to see, as determined from previous experiments, or just an estimate.
4. Enter the desired power, or test sensitivity. Power is the probability that the correlation coefficient quantifies an actual association. The closer the power is to 1, the more sensitive the test. Traditionally, you want to achieve a power of 0.80, which means that there is an 80% chance of detecting an association with $1 - \alpha$ confidence (for example, a 95% confidence when $\alpha = 0.05$).
5. Enter the desired alpha level. Alpha (α) is the acceptable probability of incorrectly concluding that there is an association. The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is an association when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a true association, but a greater possibility of concluding there is no relationship when one exists (a Type II error). Larger values of α make it easier to conclude that there is an association, but also increase the risk of reporting a false positive (a Type I error).

6. Click **=** to see the required sample size of a correlation coefficient at the specified conditions. The sample size calculation appears at the top of the dialog. If desired, you can change any of the settings and click **=** again to view the new sample size as many times as desired.

7. Click **Save to Report** to save the sample size computation settings and resulting sample size to the current report.

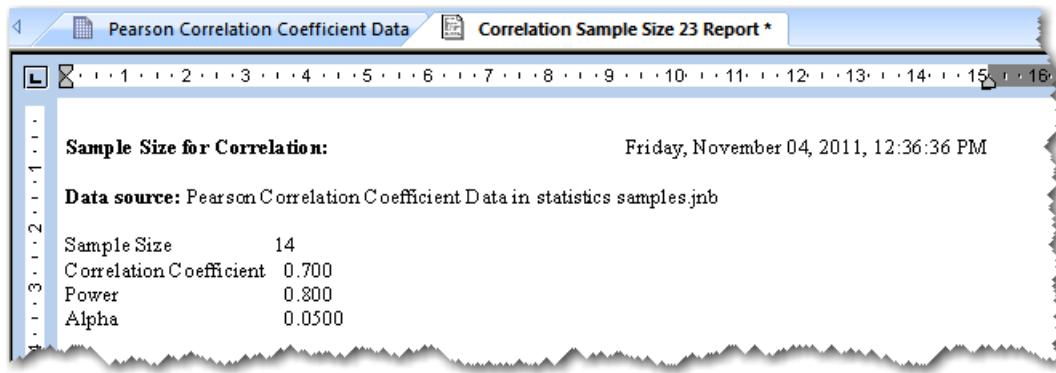


Figure 158: The Correlation Coefficient Sample Size Results Viewed in the Report

8. Click **Close** to exit from correlation coefficient sample size computation.

About Power

The power, or sensitivity, of a test is the probability that the test will detect a difference or effect if there really is a difference or effect. The closer the power is to 1, the more sensitive the test. Traditionally, you want to achieve a power of 0.80, which means that there is an 80% chance of detecting a specified effect with $1 - \alpha$ confidence (for example, a 95% confidence when $\alpha = 0.05$). Power less than 0.001 is noted as "P = < 0.001."

The power of a statistical test depends on:

- The specific test
- The alpha (α), or acceptable risk of a false positive
- The sample size
- The minimum difference or treatment effect to detect
- The underlying variability of the data

Determining the Power of a t-Test

You can determine the power of an intended t-test. Use unpaired t-tests to compare two different samples from populations that are normally distributed with equal variances among the individuals. For more information, see [Unpaired t-Test](#) on page 57.

To determine the power for a t-test, you need to set the:

- Expected difference of the means of the groups you want to detect.
- Expected standard deviation of the groups.
- Expected sizes of the two groups.
- Alpha (α) used for power computations.

To find the power of a t-test:

1. With the worksheet in view, click the Analysis tab.

2. In the **SigmaStat** group, select:

Power > t-test

The **t-test Power** dialog box appears.

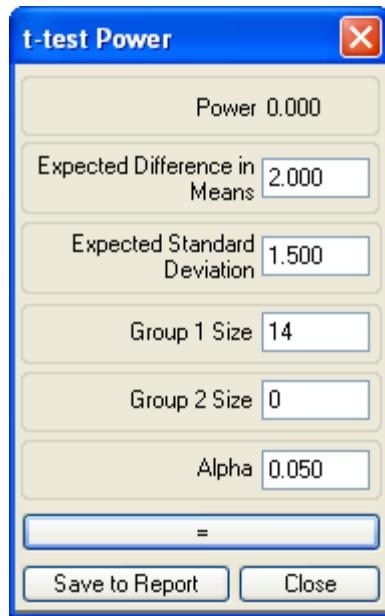


Figure 159: The t-test Power Dialog Box

- Enter the size of the difference between the means of the two groups you want to be able to detect in the **Expected Difference of Means** box. This can be the size you expect to see, as determined from previous samples or experiments, or just an estimate.
- Enter the estimated size of the standard deviation for the population your data will be drawn from in the **Expected Standard Deviation** box. This can be the size you expect to see, as determined from previous samples or experiments, or just an estimate.



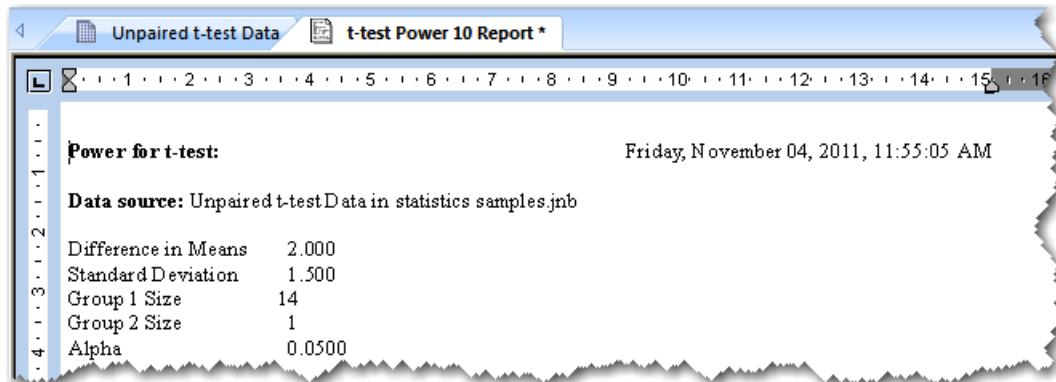
Note: *t*-Tests assume that the standard deviations of the underlying normally distributed populations are equal.

- Enter the expected sizes of each group in the **Group 1 Size** and **Group 2 Size** boxes.
- If desired, change the alpha level in the **Alpha** box. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. An α error is also called a Type I error (a Type I error is when you reject the hypothesis of no effect when this hypothesis is true). The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive (a Type I error).

- Click **=** to see the power of a *t*-test at the specified conditions. The Power calculation appears at the tip of the dialog box. If desired, you can change any of the settings and click **=** again to view the new power as many times as desired.

8. Click **Save to Report** to save the power computation settings and resulting power to the current report and click **Close** to exit from t-test power computation.



Determining the Power of a Paired t-Test

You can determine the power of a Paired t-test. Use Paired t-tests to see if there is a change in the same individuals before and after a single treatment or change in condition. The sizes of the treatment effects are assumed to be normally distributed. For more information, see [Paired t-Test](#) on page 141.

To determine the power for a Paired t-test, you need to set the:

- Expected change before and after treatment you want to detect.
- Expected standard deviation of the changes.
- Number of subjects.
- Alpha used for power computations

To find the power of a Paired t-test:

1. With the worksheet in view, click the **Analysis** tab.
2. In the **SigmaStat** group, select:

Power > Paired t-test

The **Paired t-test Power** dialog box appears.

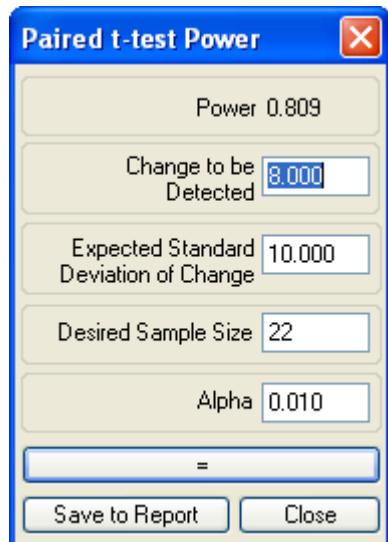


Figure 160: The Paired t-test Power Dialog Box

3. Enter the size of the change before and after the treatment in the **Change to be Detected** box. The size of the change is determined by the difference of the means. This can be size of the treatment effect you expect to see, as determined from previous experiments, or just an estimate.

4. Enter the size of standard deviation of the change in the **Expected Standard Deviation of Change** box. This can be the size you expect to see, as determined from previous experiments, or just an estimate.
5. Enter the expected (or estimated) number of subjects in the **Desired Sample Size** box.
6. Enter the desired alpha level. Alpha (α) is the acceptable probability of incorrectly concluding that there is an effect. The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant treatment difference when $P < 0.05$.
Smaller values of α result in stricter requirements before concluding there is a significant effect, but a greater possibility of concluding there is no effect when one exists (a Type II error). Larger values of α make it easier to conclude that there is an effect, but also increase the risk of reporting a false positive (a Type I error).
7. Click = to see the power of a Paired t-test at the specified conditions. If desired, you can change any of the settings and click = again to view the new power as many times as desired.
8. Select **Save to Report** to save the power computation settings and resulting power to the current report.

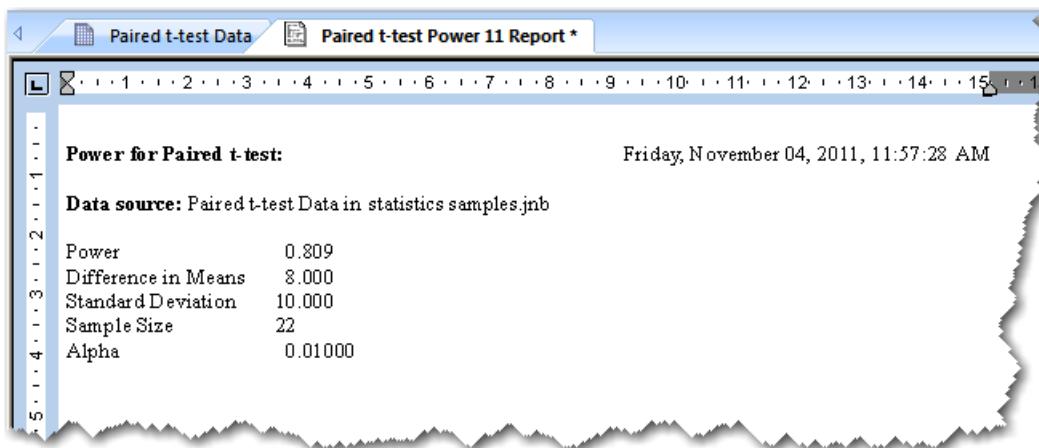


Figure 161: The Paired t-test Power Computation Results Viewed in the Report

9. Click Close.

Determining the Power of a z-Test Proportions Comparison

You can determine the power of a z-test comparison of proportions. A comparison of proportions compares the difference in the proportion of two different groups that fall within a single category. For more information, see [Comparing Proportions Using the z-Test](#) on page 188.

To determine the power for a proportion comparison, you need to set the:

- Expected proportion of each group that falls within the category.
- Size of each sample.
- Alpha (α) used for power computations.

To find the power of a z-test proportion comparison:

1. With the worksheet in view, click the **Analysis** tab.

2. In the **SigmaStat** group, select:

Power > Proportions

The **Proportions Power** dialog box appears.

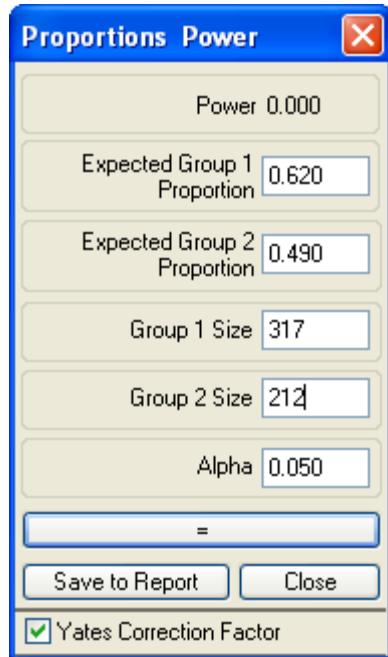


Figure 162: The Proportions Power Dialog Box

- Enter the expected proportions that fall into the category for each group. This can be the distribution you expect to see, as determined from previous experiments, or just an estimate.
- Enter the sizes of each group. This can be sample sizes you expect to obtain, or just an estimate.
- Enter the desired alpha level. Alpha (α) is the acceptable probability of incorrectly concluding that there is an effect. The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant distribution difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference in distribution when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive (a Type I error).

- Click = to see the power of a proportion comparison at the specified conditions. If desired, you can change any of the settings and click = again to view the new power as many times as desired.

 **Note:** SigmaPlot uses the Yates correction factor if this option is selected in the Options for z-Test dialog box.

7. Click **Save to Report** to save the power computation settings and resulting power to the current report.

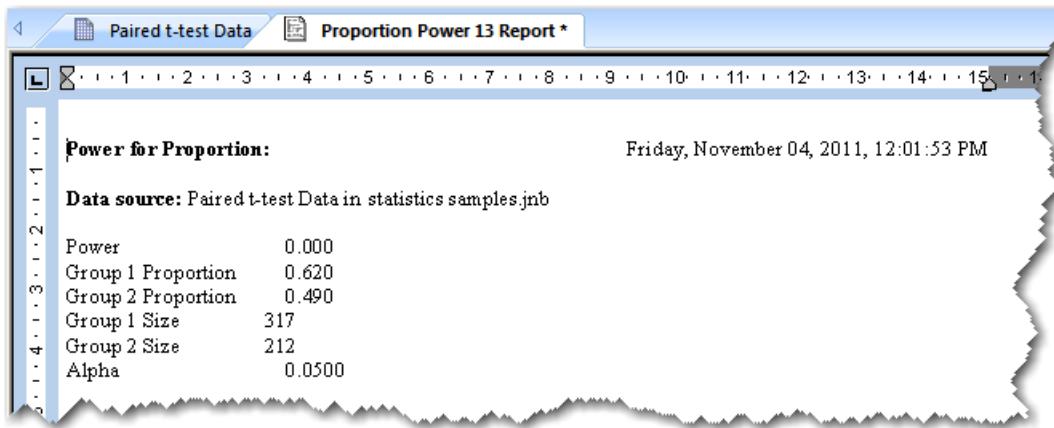


Figure 163: The Proportion Power Computation Results Viewed in the Report

8. Click **Close** to exit from proportion comparison power computation.

Determining the Power of a One Way ANOVA

You can determine the power of a One Way ANOVA (analysis of variance). Use One Way ANOVAs to see if there is a difference among two or more samples taken from populations that are normally distributed with equal variances among the individuals. For more information, see [One Way Analysis of Variance \(ANOVA\)](#) on page 73.

To determine the power for a One Way ANOVA, you need to specify the:

- Minimum difference between group means you want to detect.
- Standard deviation of the population from which the samples were drawn.
- Estimated number of groups.
- Estimated size of a group.
- Alpha (α) used for power computations.

To find the power of a One Way ANOVA:

1. With the worksheet in view, click the **Analysis** tab.

2. In the **SigmaStat** group, select:

Power > ANOVA

The **ANOVA Power** dialog box appears.

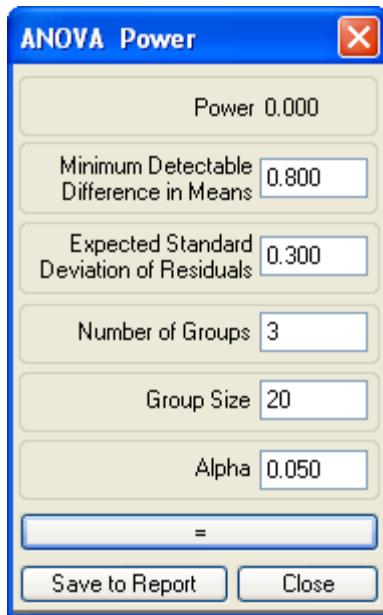


Figure 164: The ANOVA Power Dialog Box

3. Enter the minimum size of the expected difference of group means in the **Minimum Detectable Difference in Means** box. This can be size of a difference you expect to see, as determined from previous experiments, or just an estimate.

The minimum detectable difference is the minimum difference between the largest and smallest means.

4. Enter the estimated standard deviation of the population from which the samples will be drawn. This can be size you expect to see, as determined from previous experiments, or just an estimate.

5. Enter the expected number of groups and the expected size of each group.

6. Enter the desired alpha level. Alpha (α) is the acceptable probability of incorrectly concluding that there is an effect. The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, for example, you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive (a Type I error).

7. Click **=** to see the power of a One Way ANOVA at the specified conditions. The power calculation appears at the top of the dialog. If desired, you can change any of the settings and click **=** again to view the new power as many times as desired.

8. Select **Save to Report** to save the power computation settings and resulting power to the current report.

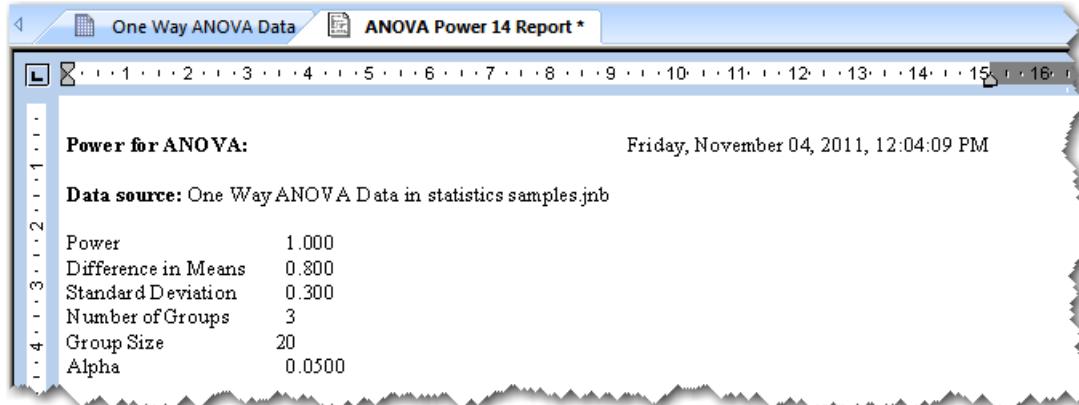


Figure 165: The ANOVA Power Computation Results Viewed in the Report

9. Click **Close** to exit from ANOVA power computation.

Determining the Power of a Chi-Square Test

You can determine the power of a chi-square χ^2 analysis of a contingency table. A χ^2 test compares the difference between the expected and observed number of individuals of two or more different groups that fall within two or more categories. For more information, see [Chi-square Analysis of Contingency Tables](#) on page 193.

The power of a χ^2 analysis contingency tables is determined by the estimated relative proportions in each category for each group. Because SigmaPlot uses numbers of observations to compute the estimated proportions, you need to enter a contingency table in the worksheet containing the estimated pattern in the observations before you can compute the estimated proportions.

Table 23: The Contingency Table with Expected Numbers of Observations of Two Groups in Three Categories

Group	Categories		
	Category 1	Category 2	Category 3
Group 1	15	15	35
Group 2	15	30	10

- Tip:** You only need to specify the *pattern* (distribution) of the number of observations. The absolute numbers in the cells do not matter, only their relative values.

To find the power of a chi-squared test:

1. Enter a contingency table into the worksheet by placing the estimated number of observations for each table cell in a corresponding worksheet cell. These observations are used to compute the estimated proportions.

	1-Tundra	2-Foothills	3-Treeline	4
1	125.0000	16.0000	6.0000	125.0000
2	7.0000	19.0000	117.0000	7.0000
3				
4				

Figure 166: Contingency Table Data Entered into the Worksheet

The worksheet rows and columns correspond to the groups and categories. The number of observations must always be an integer.

Tip: The order and location of the rows or columns corresponding to the groups and categories is not important.

2. With the worksheet in view, click the **Analysis** tab.

3. In the **SigmaStat** group, select:

Power > Chi-Square

The **Select Data** panel of the Test Wizard appears.

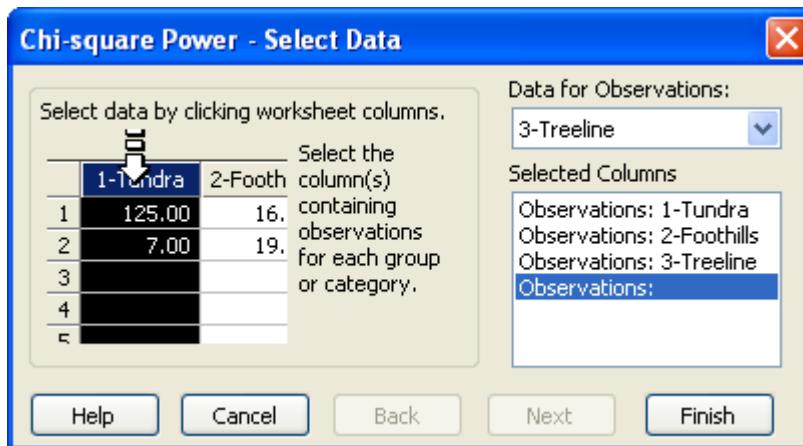


Figure 167: The Chi-square Power — Select Data Panel of the Test Wizard

4. Select the columns of the contingency table from the worksheet as prompted.

5. Click **Finish** when you've selected the desired columns.

The **Chi-Square Power** dialog box appears.



Figure 168: The Chi-square Power Dialog Box

6. Enter the total number of observations in the **Sample Size** box. This can be the number of observations you expect to see, as determined from previous experiments, or just an estimate.
7. Enter the desired alpha level. Alpha (α) is the acceptable probability of incorrectly concluding that there is a difference. The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.
Smaller values of α result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no effect when one exists (a Type II error). Larger values of α make it easier to conclude that there is a difference, but also increase the risk of reporting a false positive (a Type I error).
8. Click **=** to see the power of a chi-square test at the specified conditions. If desired, you can change any of the settings and click **=** again to view the new power as many times as desired. However, if you want to change the number of observations per category, you need to click **Cancel**, edit the table, then repeat the sample size computation.
9. Select **Save to Report** to save the power computation settings and resulting power to the current report file, and then click **Cancel** to exit from chi-square test power computation.

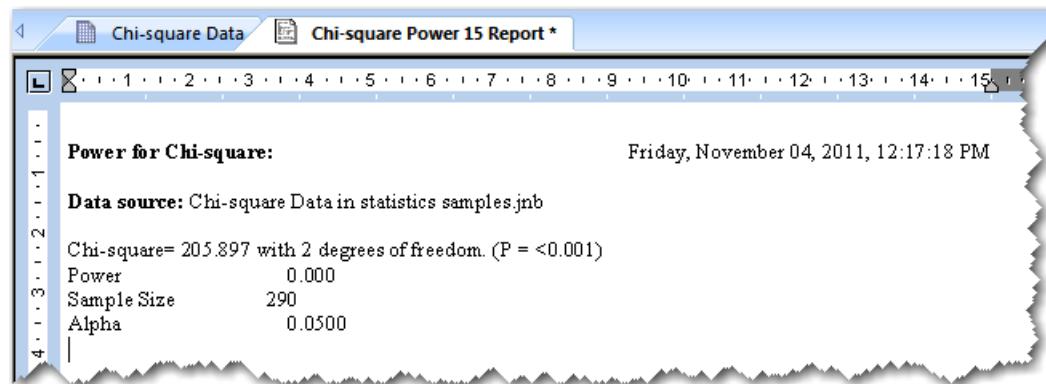


Figure 169: The Chi-square Power Computation Results Viewed in the Report

Determining the Power to Detect a Specified Correlation

You can determine the power to detect a given Pearson Product Moment Correlation Coefficient R . A correlation coefficient quantifies the strength of association between the values of two variables. A correlation coefficient of 1 means that as one variable increases, the other increases exactly linearly. A correlation coefficient of -1 means that as one variable increases, the other decreases exactly linearly. For more information, see [Pearson Product Moment Correlation](#) on page 294.

To determine the power of a correlation coefficient, you need to specify the:

- Correlation coefficient you want to detect.
- Desired sample size.
- Alpha (α) used for power computations.

To find the power to detect a correlation coefficient:

1. With the worksheet in view, click the **Analysis** tab.

2. In the **SigmaStat** group, select:

Power > Correlation

The **Correlation Power** dialog box appears.

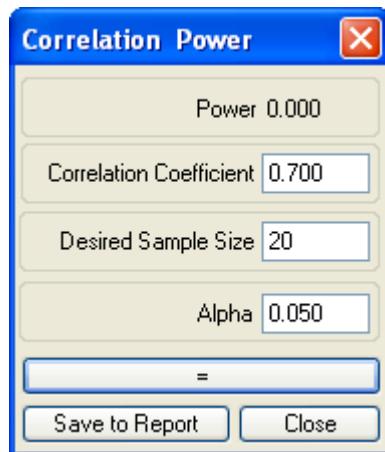


Figure 170: The Correlation Power Dialog Box

3. Enter the expected correlation coefficient. This can be the correlation coefficient you expect to see, as determined from previous experiments, or just an estimate.
4. Enter the desired number of data points. This can be the sample size you expect to obtain, or just an estimate.
5. Enter the desired alpha level. Alpha (α) is the acceptable probability of incorrectly concluding that there is an association. The traditional α value used is 0.05. This indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is an association when $P < 0.05$.

Smaller values of α result in stricter requirements before concluding there is a true association, but a greater possibility of concluding there is no relationship when one exists (a Type II error). Larger values of α make it easier to conclude that there is an association, but also increase the risk of reporting a false positive (a Type I error).

6. Click '=' to see the power of a correlation coefficient at the specified conditions. The power calculation appears at the top of the dialog box. If desired, you can change any of the settings and click '=' again to view the new power as many times as desired.

7. Click **Save to Report** to save the power computation settings and resulting power to the current report, and then click **Close** to exit from correlation coefficient power computation.

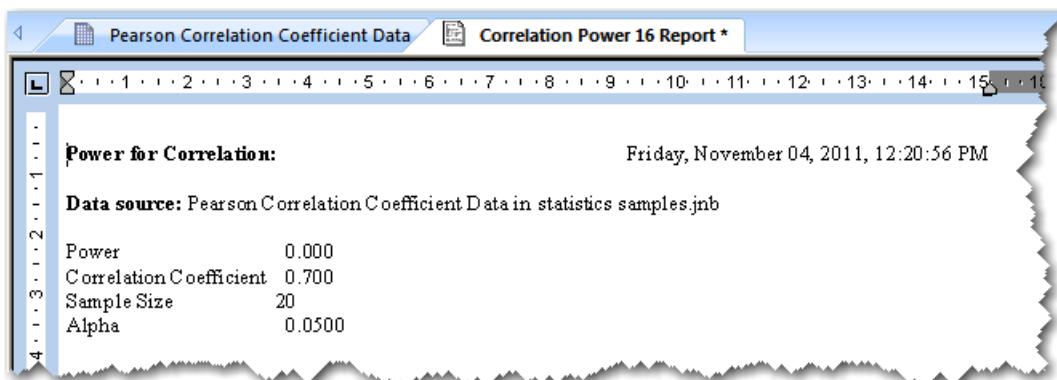


Figure 171: The Correlation Power Dialog Box

Chapter

12

Report Graphs

Topics:

- Generating Report Graphs
- Bar Charts of the Column Means
- Scatter Plot
- Point Plot
- Point Plot and Column Means
- Box Plot
- Scatter Plot of the Residuals
- Bar Chart of the Standardized Residuals
- Histogram of Residuals
- Normal Probability Plot
- 2D Line/Scatter Plots of the Regressions with Prediction and Confidence Intervals
- 3D Residual Scatter Plot
- Grouped Bar Chart with Error Bars
- 3D Category Scatter Graph
- Before and After Line Plots
- Multiple Comparison Graphs
- Scatter Matrix
- Profile Plots
- Scree Plot
- Component Loadings Plot
- Component Scores Plot
- Regression Lines for Groups
- Adjusted Means with 95% Confidence Intervals

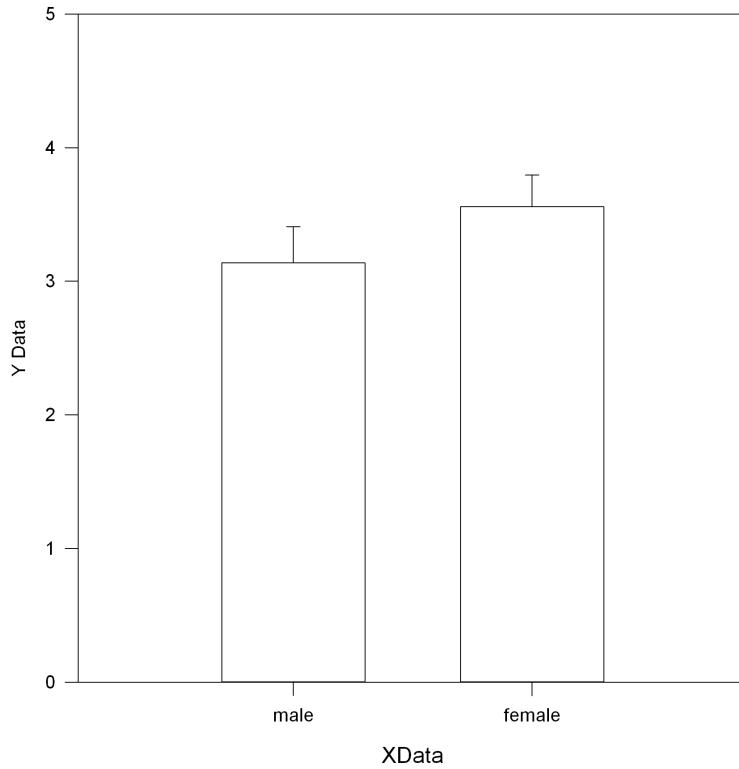
You can generate graphs for all test reports except rates and proportions tests, best subset and incremental polynomial regression, and multiple logistic reports

Generating Report Graphs

To generate a report graph:

1. Click the **Report** tab and then in the **Result Graphs** group, click **Create Result Graph**.
The **Create Result Graph** dialog box appears displaying the available graphs for the selected report.

Bar Chart Column Means



Scatter Plot

The scatter plot is available for the following tests:

- [Descriptive Statistics](#).
- [t-Test](#).
- [One Sample t-Test](#).
- [One Way ANOVA](#).

If the graph data is indexed, the levels in the factor column are used as the tick marks for the scatter plot points, and the column titles are used as the X and Y axis titles. If the graph data is in raw or statistical format, the column titles are used as the tick marks for the scatter plot points and default X Data and Y Data axis titles are assigned to the graph.

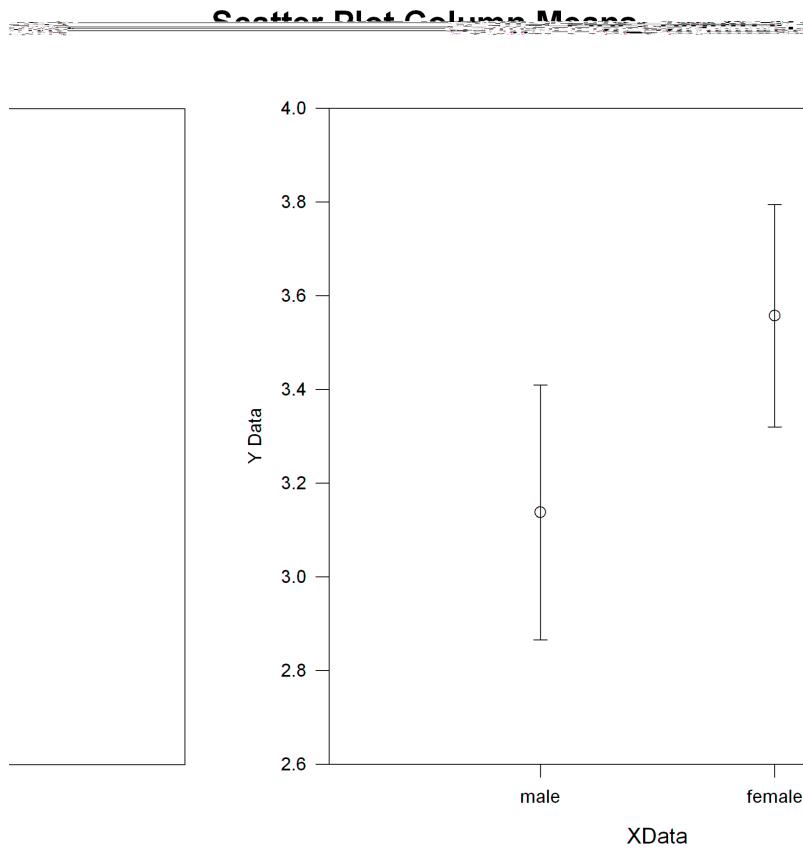


Figure 173: The scatter plot graphs the group means as single points with error bars indicating the standard deviation.

Point Plot

The point plot is available for the following tests:

- [Descriptive Statistics](#).
- [t-test](#).
- [Rank Sum Test](#).
- [ANOVA on Ranks](#).

If the graph data is indexed, the levels in the factor column are used as the tick marks for the plot points, and the column titles are used as the X and Y axis titles. If the graph data is in raw or statistical format, the column titles are used as the tick marks for the plot points and default X Data and Y Data axis titles are assigned to the graph.

Point Plot

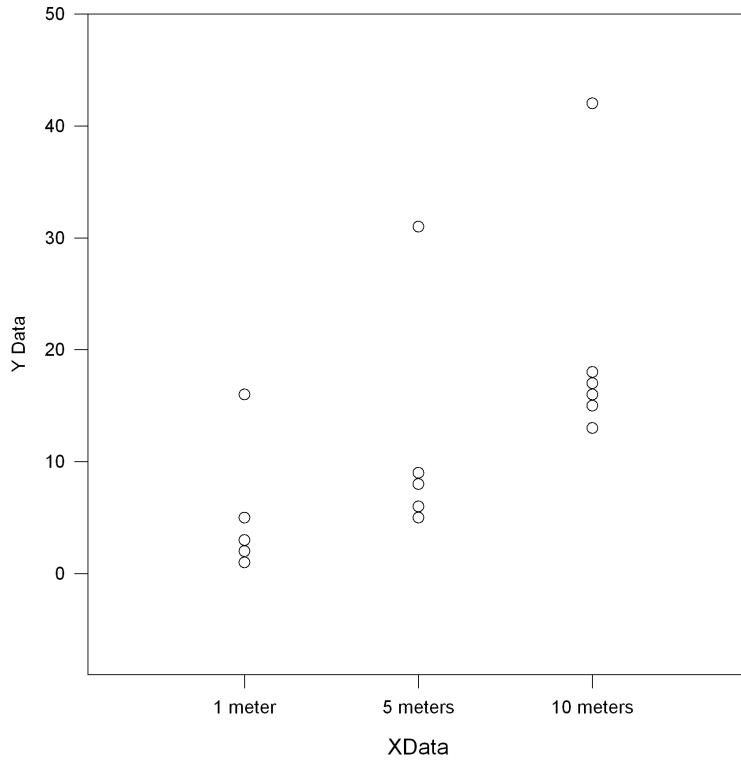
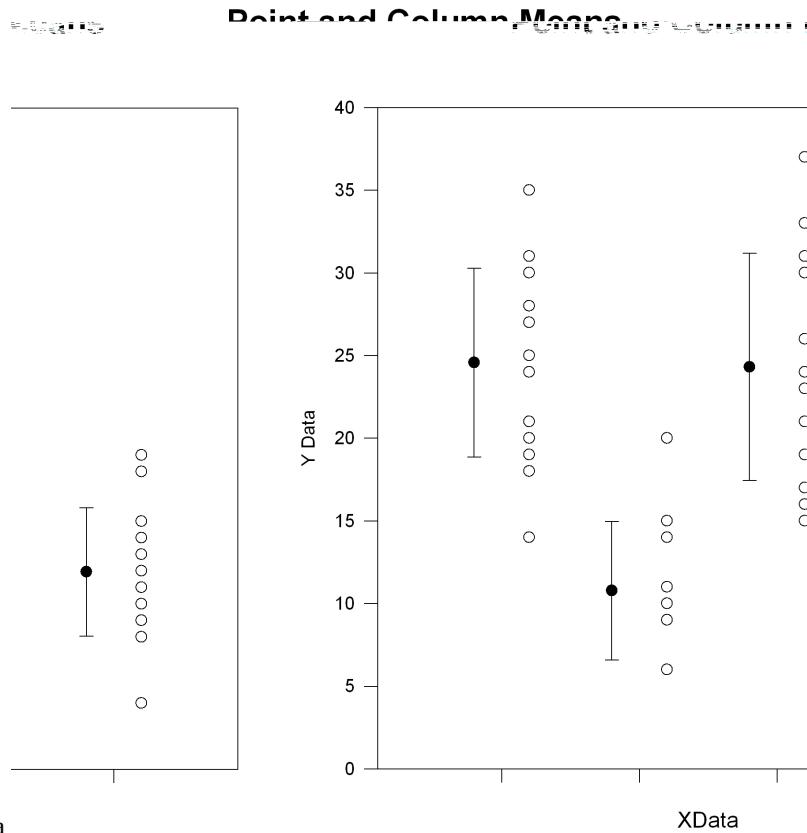


Figure 174: A Point Plot of the Result Data for an ANOVA on Ranks

Point Plot and Column Means

The point and column means plot is only available for Descriptive Statistics. The point and column means plot graphs all values in each column as a point on the graph with error bars indicating the column means and standard deviations of each column. For more information, see [Describing Your Data with Basic Statistics](#) on page 18.

The error bars plot the column means and the standard deviations of the column



data.

Figure 175: A Point and Column Means Plot of the Result Data for a Descriptive Statistics Test

Box Plot

The Rank Sum Test box plot graphs the percentiles and the median of column data. The ends of the boxes define the 25th and 75th percentiles, with a line at the median and error bars defining the 10th and 90th percentiles.

If the graph data is indexed, the levels in the factor column are used as the tick marks for the box plot boxes, and the column titles are used as the axis titles. If the graph data is in raw format, the column titles are used as the tick marks for the box plot boxes, and no axis titles are assigned to the graph.

The box plot is available for the following tests:

- **Descriptive Statistic.**
- **Rank Sum Test.**
- **ANOVA on Ranks.**
- **Repeated Measures ANOVA on Ranks.**

Box Plot

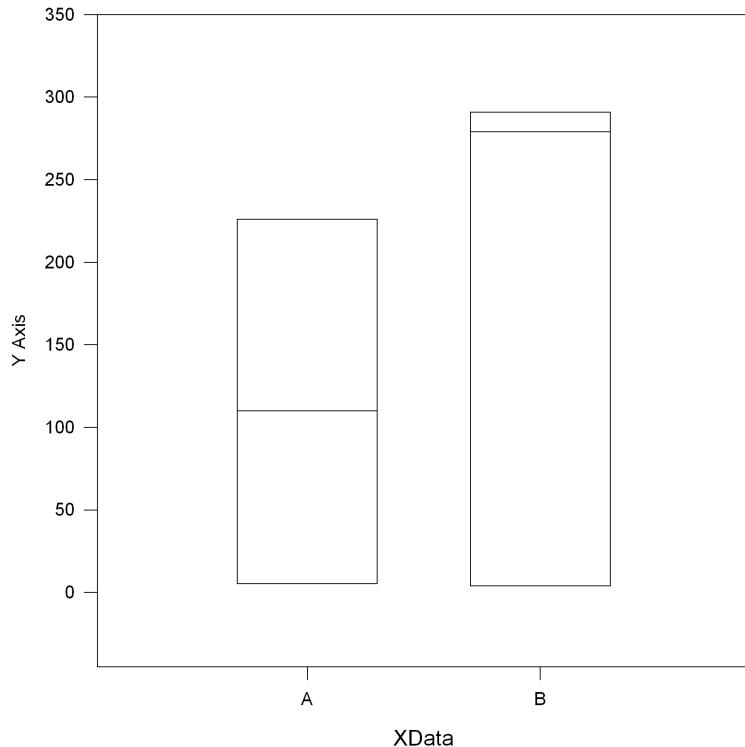


Figure 176: A Box Plot of the Result Data for the Rank Sum Test

Scatter Plot of the Residuals

The 2D scatter plot of the residuals is available for all of the regressions except the Multiple Logistic and the Incremental Polynomial Regressions.

The scatter plots of the residuals plot the raw residuals of the independent variables as points relative to the standard deviations. The X axis represents the independent variable values, the Y axis represents the residuals of the variables, and the horizontal lines running across the graph represent the standard deviations of the data. For more information, see [Prediction and Correlation](#) on page 225 .

Scatter Plot Residuals

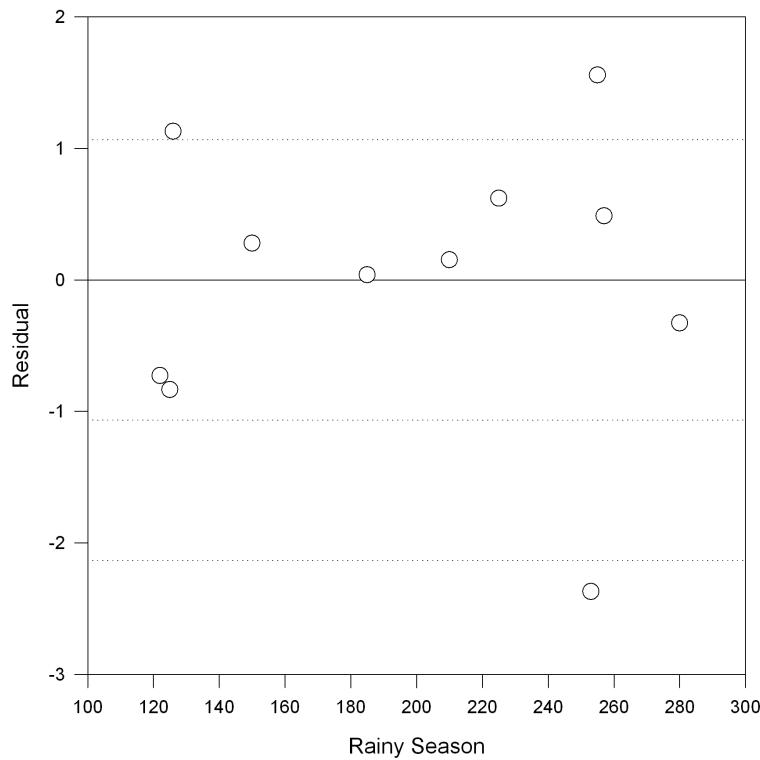


Figure 177: Scatter Plot of the Simple Linear Regression Residuals with Standard Deviation

Bar Chart of the Standardized Residuals

Bar charts of the standardized residuals are available for all regressions except the Multiple Logistic and the Incremental Polynomial Regressions. They plot the standardized residuals of the data in the selected independent variable column as points relative to the standard deviations. For more information, see [Prediction and Correlation](#).

Bar Chart Std Residuals

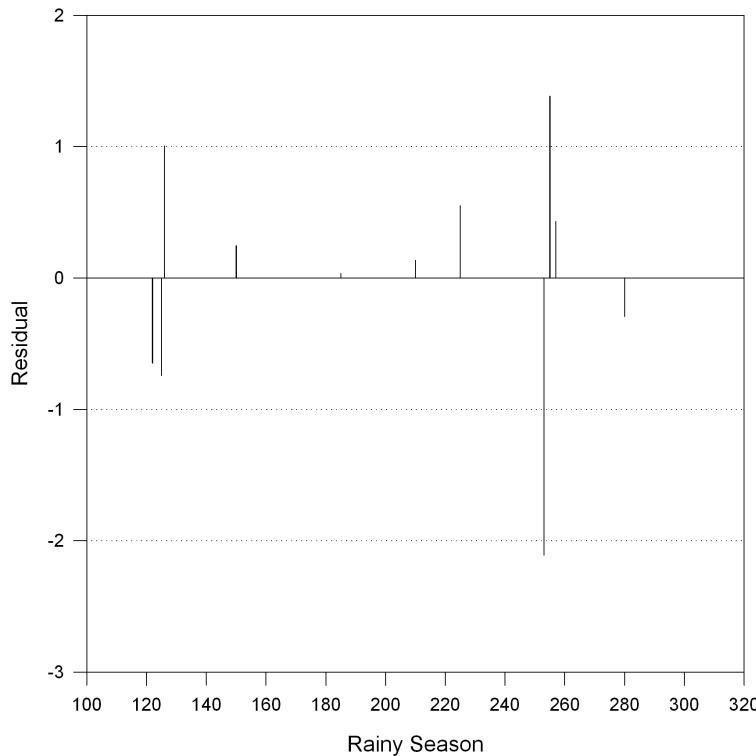


Figure 178: A Multiple Linear Regression Bar Chart of the Standardized Residuals with Standard Deviations Using One Independent Variable

Histogram of Residuals

The histogram plots the raw residuals as a bar chart in a specified range, using a defined interval set. The range of each interval is identical; the total range is the data minimum to the data maximum. The number of bars equals the number of groups or bins that are specified by the user to partition the data. The X axis represents the residual data to be binned, and the Y axis represents a scaling of the number of residuals in each group.

To open the **Histogram Options** dialog box, select **Histogram** from the **Create Result Graph** dialog box, and then click **OK**.

The **Histogram Options** dialog box appears.

The histogram of residuals graph is available for the following tests:

- [One Sample t-Test](#).
- [t-test](#).
- [One Way ANOVA](#).
- [Two Way ANOVA](#).
- [Three Way ANOVA](#).
- [Paired t-Test](#).
- [One Way Repeated Measures ANOVA](#).
- [Two Way Repeated Measures ANOVA](#).
- [Linear Regression](#).
- [Multiple Linear Regression](#).
- [Polynomial Regression](#).

- [Stepwise Regression](#).
- [Nonlinear Regression](#). [Normality Test](#).
- [Normality Test](#).

Histogram

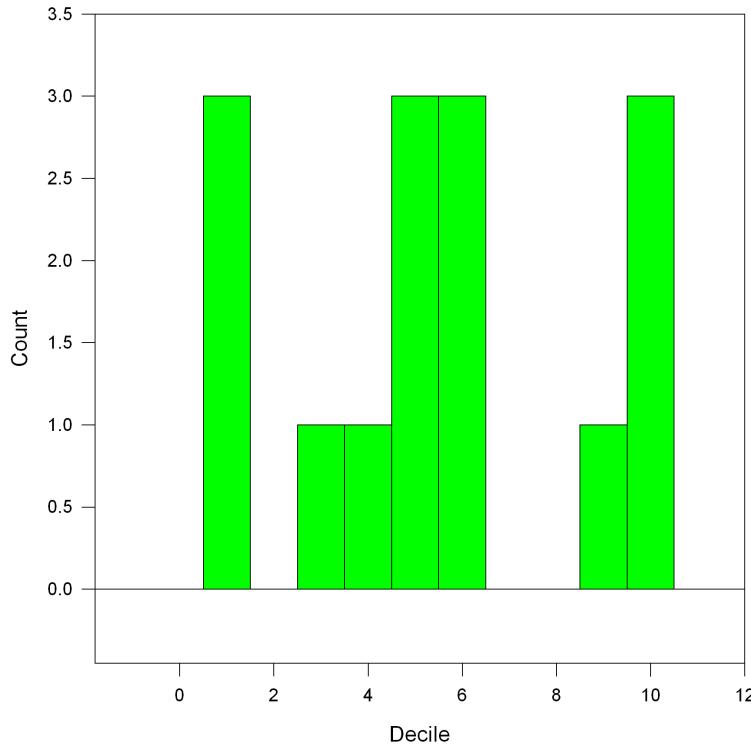


Figure 179: A Histogram of the Residuals for a t-Test

Normal Probability Plot

The normal probability plot graphs the frequency of the raw residuals. The residuals are sorted and then plotted as points around a curve representing the area of the Gaussian SigmaPlot plotted on a probability axis. Plots with residuals that fall along Gaussian curve indicate that your data was taken from a normally distributed population. The X axis is a linear scale representing the residual values. The Y axis is a probability scale representing the cumulative frequency of the residuals.

The normal probability plot is available for the following test reports:

- [One Sample t-Test](#).
- [t-test](#).
- [One Way ANOVA](#).
- [Two Way ANOVA](#).
- [Three Way ANOVA](#).
- [Paired t-Test](#).
- [One Way Repeated Measures ANOVA](#).
- [Two Way Repeated Measures ANOVA](#).
- [Linear Regression](#).
- [Multiple Linear Regression](#).
- [Polynomial Regression](#).
- [Stepwise Regression](#).

- [Nonlinear Regression.](#)
- [Normality Test.](#)

Normal Probability Plot

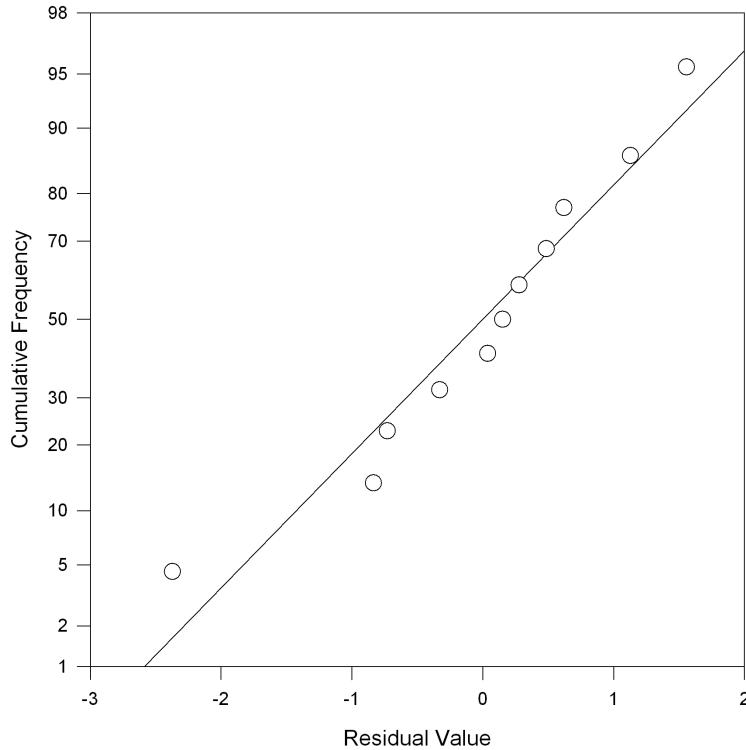


Figure 180: Normal Probability Plot of the Residuals

2D Line/Scatter Plots of the Regressions with Prediction and Confidence Intervals

The 2D line and scatter plots of the regressions are available for all of the regression reports, except Multiple Logistic and Incremental Polynomial Regressions. They plot the observations of the regressions as a line/scatter plot. The points represent the data dependent variables plotted against the independent variables, the solid line running through the points represents the regression line, and the dashed lines represent the prediction and confidence intervals. The X axis represents the independent variables and the Y axis represents the dependent variables. For more information, see [Prediction and Correlation](#).

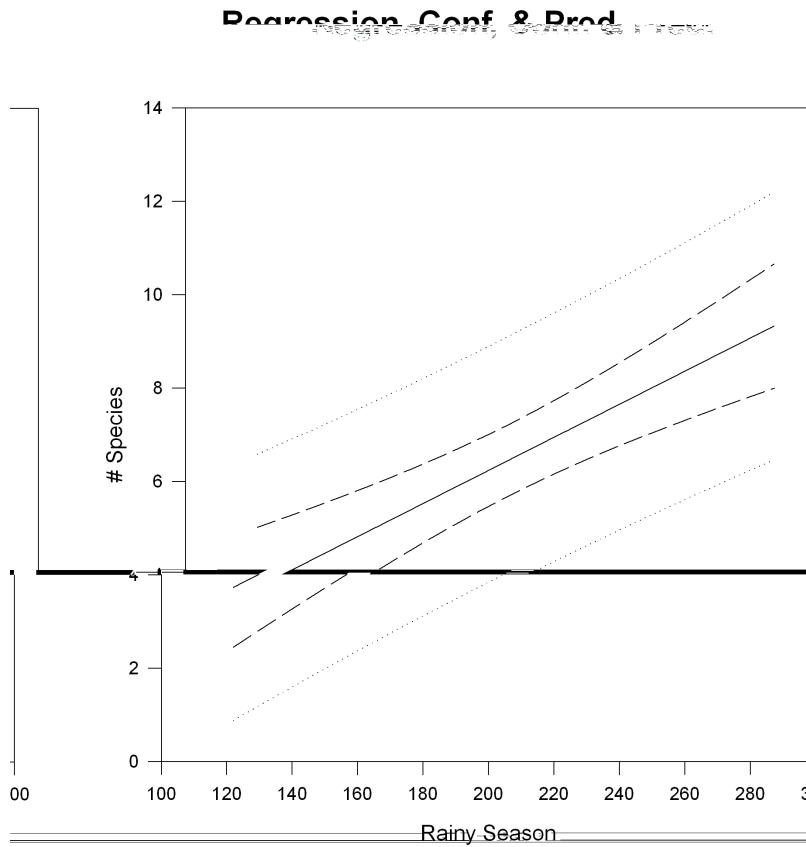


Figure 181: A Line/Scatter Plot of the Linear Regression Observations with a Regression and Confidence and Prediction Interval Lines

3D Residual Scatter Plot

The 3D residual scatter plots are available for the following test reports:

- [Two Way ANOVAs](#).
- [Two Way Repeated Measures ANOVA](#).
- [Multiple Linear Regression](#).
- [Stepwise Regression](#).

They plot the residuals of the two selected columns of independent variable data. The X and the Y axes represent the independent variables, and the Z axis represents the residuals.

3D Residual Scatter

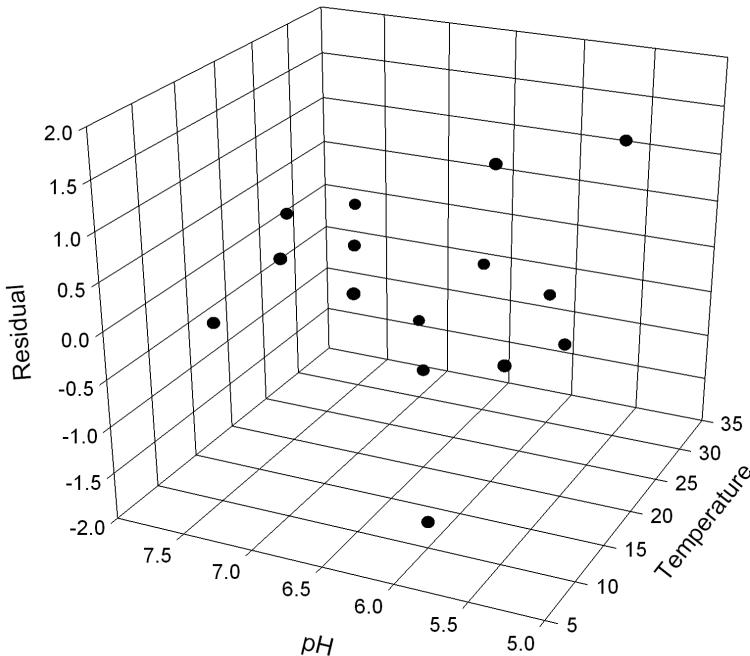


Figure 182: A Multiple Linear Regression 3D Residual Scatter Plot of the Two Selected Independent Variable Columns

Grouped Bar Chart with Error Bars

This graph is available for the Two Way ANOVA. It plots the least square means of the ANOVA cells with error bars indicating the standard errors of these means. The levels in the second factor column are used as the X axis tick marks, and the titles of the two factors and the title of the data column are used as the X and the Y axis titles. The first bar in the group represents the first level of the first factor column and the second bar in the group represents the second level in the first factor column.

For more information, see [Two Way Analysis of Variance \(ANOVA\)](#) on page 89.

Grouped Bar Chart

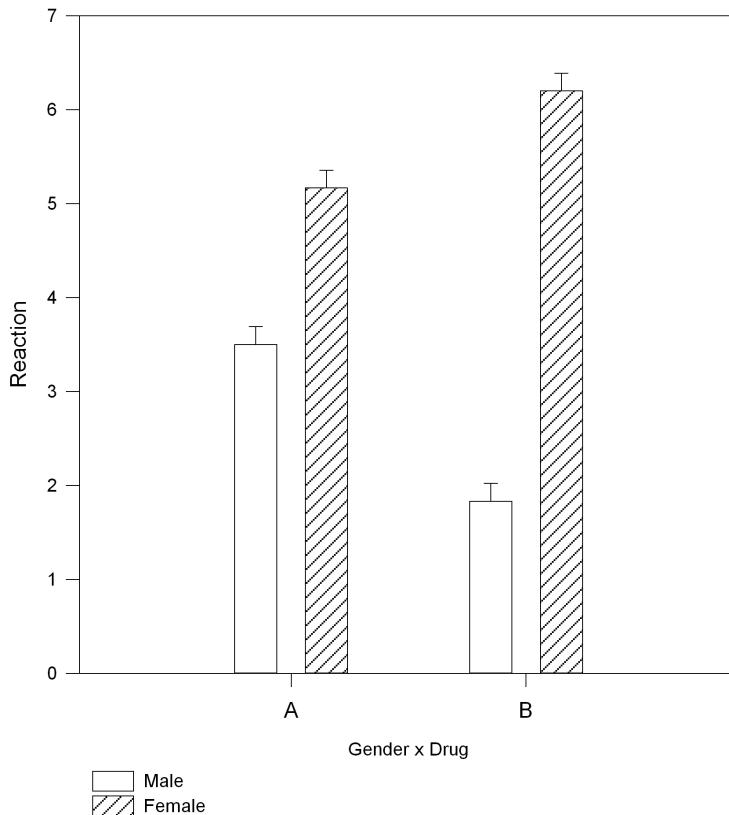


Figure 183: A Two Way ANOVA Grouped Bar Chart with Error Bars

3D Category Scatter Graph

This graph is available for the Two Way ANOVA and the Two Way Repeated Measures ANOVA. The 3D Category Scatter plot graphs the two factors from the independent data columns along the X and Y axes against the data of the dependent variable column along the Z axis. The tick marks for the X and Y axes represent the two factors from the independent variable columns, and the tick marks for the Z axis represent the data from the dependent variable column.

3D Category Scatter

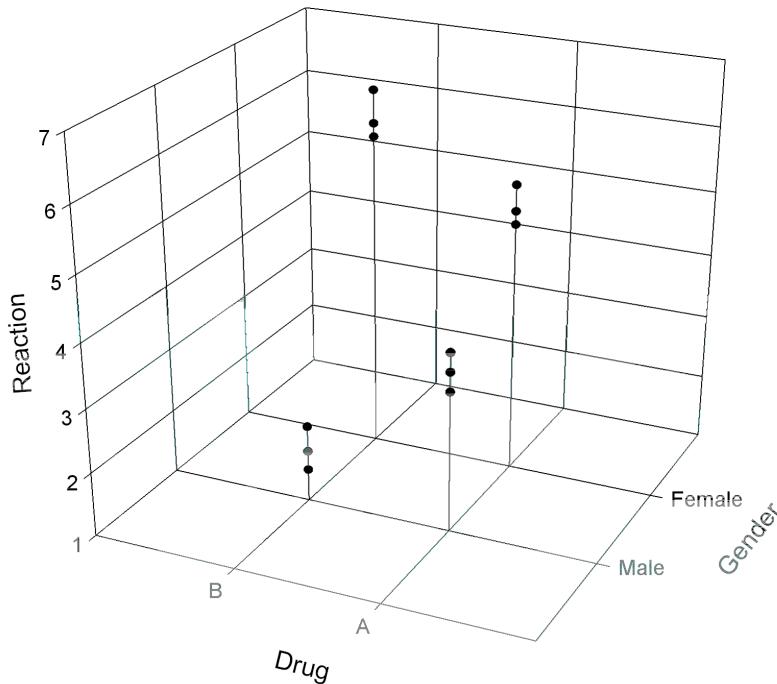


Figure 184: A Two Way ANOVA 3D Category Scatter Plot

Before and After Line Plots

The before and after line plot uses lines to plot a subject's change after each treatment. If the graph plots raw data, the lines represent the rows in the column, the column titles are used as the tick marks for the X axis and the data is used as the tick marks for the Y axis.

If the graph plots indexed data, the lines represent the levels in the subject column, the levels in the treatment column are used as the tick marks for the X axis, the data is used as the tick marks for the Y axis, and the treatment and data column titles are used as the axis titles.

The before and after line plot is available for the:

- [Paired t-test](#).
- [Repeated Measures ANOVA on Ranks. Signed Rank Test](#).
- [One Way Repeated Measures ANOVA](#).
-

Before & After Line Plot

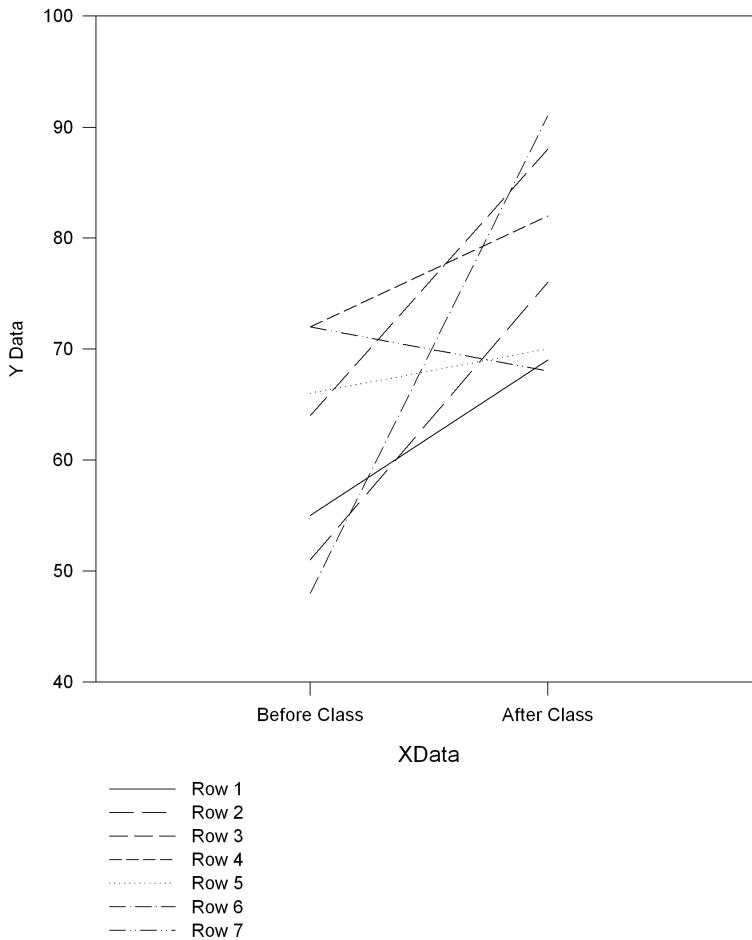


Figure 185: A Before and After Plot Displaying Data for a Paired t-Test

Multiple Comparison Graphs

The multiple comparison graphs are available for all ANOVA reports and the ANCOVA report. They plot significant differences between levels of a significant factor. There is one graph for every significant factor reported by the specified multiple comparison test. If there is one significant factor reported, one graph appears; if there are two significant factors, two graphs appear, and so on. If a factor is not reported as significant, a graph for the factor does not appear.

For an ANCOVA report, the multiple comparison graphs are available only if you are analyzing the equal slopes model. This means that you either cleared the option to do the [Equality of Slopes](#) test or the Equality of Slopes test passed.

Multiple Comparison Graph

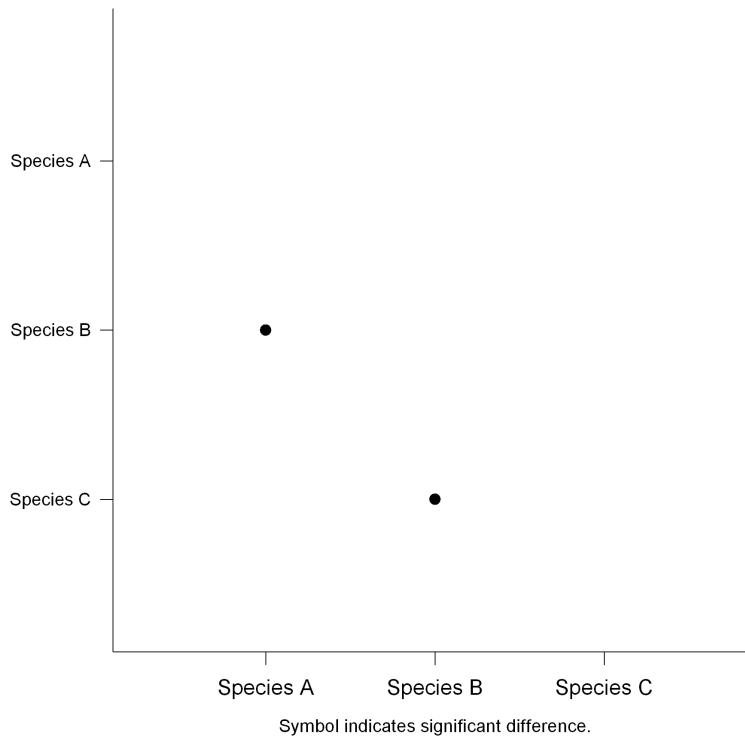


Figure 186: A Multiple Comparison Graph

Scatter Matrix

The matrix of scatter graphs is available for all the Pearson and the Spearman Correlation reports. The matrix is a series of scatter graphs that plot the associations between all possible combinations of variables.

The first row of the matrix represents the first set of variables or the first column of data, the second row of the matrix represents the second set of variables or the second data column, and the third row of the matrix represents the third set of variables or third data column, and so on. The X and Y data for the graphs correspond to the column and row of the graph in the matrix.

For example, the X data for the graphs in the first row of the matrix is taken from the second column of tested data, and the Y data is taken from the first column of tested data. The X data for the graphs in the second row of the matrix is taken from the first column of tested data, and the Y data is taken from the second column of tested data. The X data for the graphs in the third row of the matrix is taken from the second column of tested data, and the Y data is taken from the third column of tested data. The number of graph rows in the matrix is equal to the number of data columns being tested.

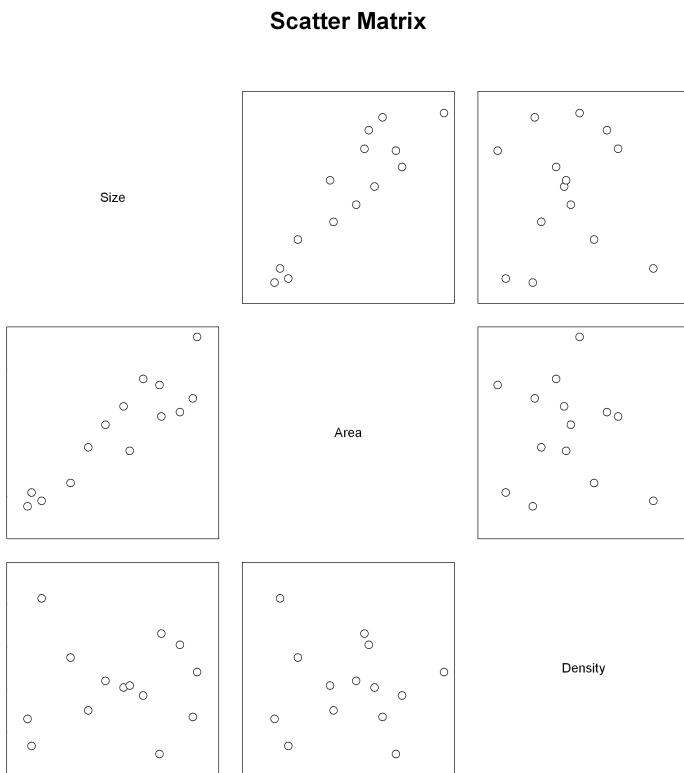


Figure 187: A Scatter Matrix for a Pearson Correlation

Profile Plots

Profile plots for ANOVA designs are line plots with the levels of one factor represented on the horizontal axis of the graph and the experiment's data represented on the vertical axis. The graph data used to create a profile plot comes from the least square means results found in the Summary section of the ANOVA report.

Profile plots are useful for when you want to compare the *least square means*, also called estimated marginal means, in a multifactor ANOVA model. Differences in the means, or effects, among the levels of a specified factor, when computed over a range of levels of the remaining factors, determine how the data is affected by that factor and its interaction with other factors. Profile plots provide a quick qualitative assessment of the various treatment effects so that the investigator can determine the impact of each factor on the data. The hypothesis testing in ANOVA reports quantifies these effects to determine if any of the differences are statistically significant.

In ANOVA analysis, the least square means are first computed for the individual cells. A cell is defined as the collection of observations made for a particular combination of levels, where one level is selected from each factor. Generally, the cell means are obtained as the predicted values in a regression model that is associated with the ANOVA model. The cell means determine the two-way interaction effects in a Two-Way ANOVA and the three-way interaction effects in a Three-Way ANOVA. If the cell means are averaged over all levels of one factor while fixing the levels of the remaining factors, you obtain lower-order effects. This is how the main effects are computed in Two-Way ANOVA and the two-way interaction effects are computed in Three-Way ANOVA. Finally, the main effects for a

given factor in a Three-Way ANOVA are determined by averaging the cell means over all levels of the remaining two factors while fixing each level of the given factor.

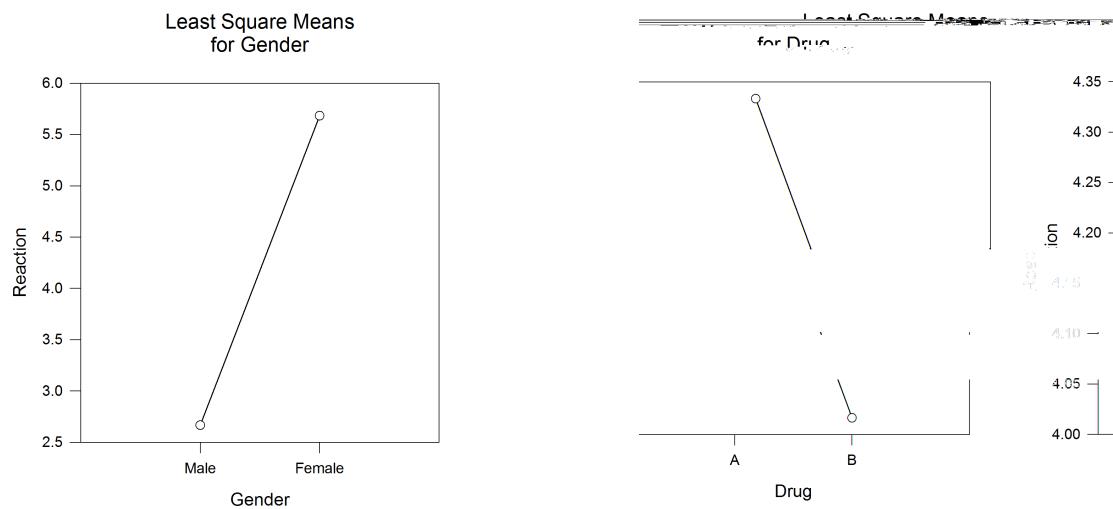
Main Effects

Main Effects graphs are available for the following tests:

- **Two Way Analysis of Variance (ANOVA).**
- **Three Way Analysis of Variance (ANOVA).**

For Main Effects, there is one plot per graph and the number of graphs equals the number of factors. For each graph, the levels of one factor are fixed, while cell means are averaged over all levels of the other factors (one other factor for Two-Way ANOVA, two other factors for Three-Way ANOVA).

Table 24: Main Effects Profile Plots for Two Way ANOVA



Two Way Effects

Two-Way Effects graphs are available for the following tests:

- **Two Way Analysis of Variance (ANOVA).**
- **Three Way Analysis of Variance (ANOVA).**

For Two Way Effects, there is one graph for each distinct pairwise-combination of factors (so there will be one graph for Two Way ANOVA and three graphs for Three Way ANOVA). Each of these graphs contains multiple profile plots, one for each level of one of the factors. For Three Way ANOVA, cell means are averaged over all levels of the remaining third factor (whichever factor not included in the pairwise-combination for the given Two Way Effects graph).

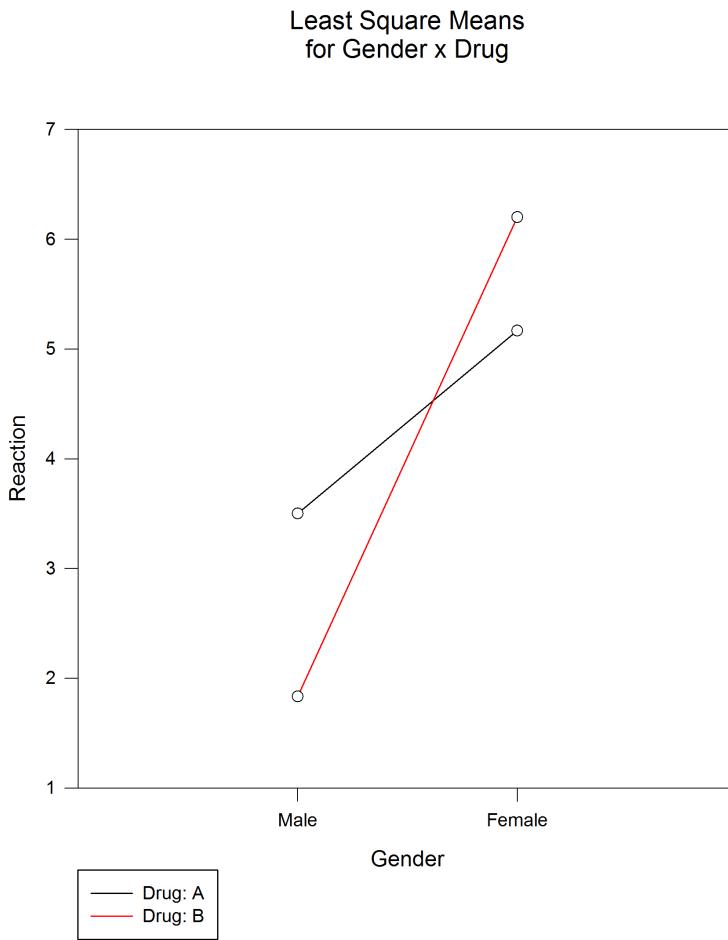


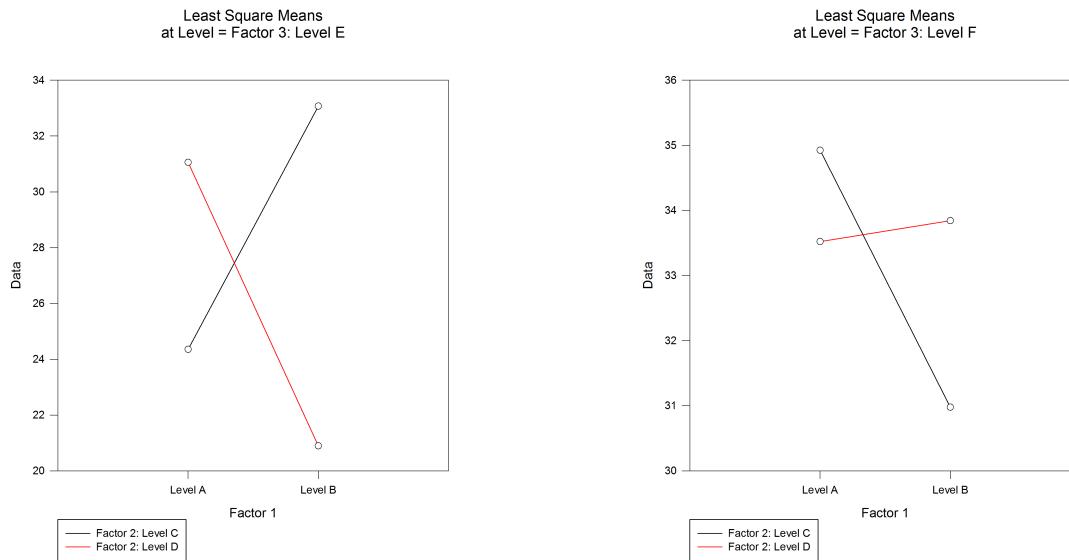
Figure 188: Two Way Effects Profile Plots for Two Way ANOVA

Three Way Effects

Three Way Effects graphs are available for the following test:

- [Three Way Analysis of Variance \(ANOVA\)](#).

For Three Way Effects in Three Way ANOVA, the number of graphs equals the number of levels of the third factor (which is the last factor that was selected for running the test). Each graph for Three Way Effects contains multiple profile plots, one for each level of one of the second factor (which is the factor that was selected second for running the test).

Table 25: Three Way Effects Profile Plots for Three Way ANOVA

Scree Plot

A Scree Plot is a plot of all eigenvalues for the covariance or correlation matrix (in descending order) versus the corresponding component number. This plot is sometimes used to select which principal components to include in your model. The usual rule is to select the components to the left of the “elbow” plus one, although the plot may have more than one break (or elbow) in it. The eigenvalues to the right of the elbow are the “scree”. In this study, the first three eigenvalues were chosen (for this and other reasons like small residuals). Recall that the variance of each principal component is equal to the corresponding eigenvalue.

The scree plot is available for the following tests:

- [Principal Component Analysis](#)

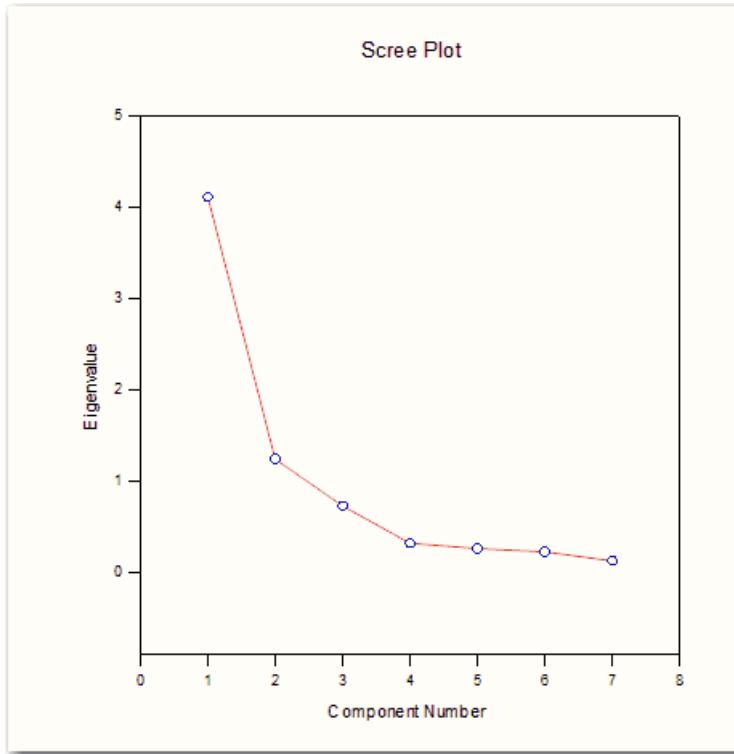


Figure 189: A Scree Plot

Component Loadings Plot

A Component Loadings plot is a [vector plot](#) showing a representation of the approximation of each original variable as the fit by any two in-model principal components that are selected by the user. There are as many vectors as there are original variables. The magnitude of each vector corresponds to the explained standard deviation of the variable. The cosine of the angle between any two vectors approximates the correlation between the corresponding variables. Thus, two variables are highly correlated if their vectors are close to pointing in the same or opposite directions. Two variables are highly uncorrelated if their vectors are close to perpendicular. The Axes labels show the percentage of total variance contributed by corresponding component.

The Components Loading Plot is available for the following tests:

- [Principal Component Analysis](#)

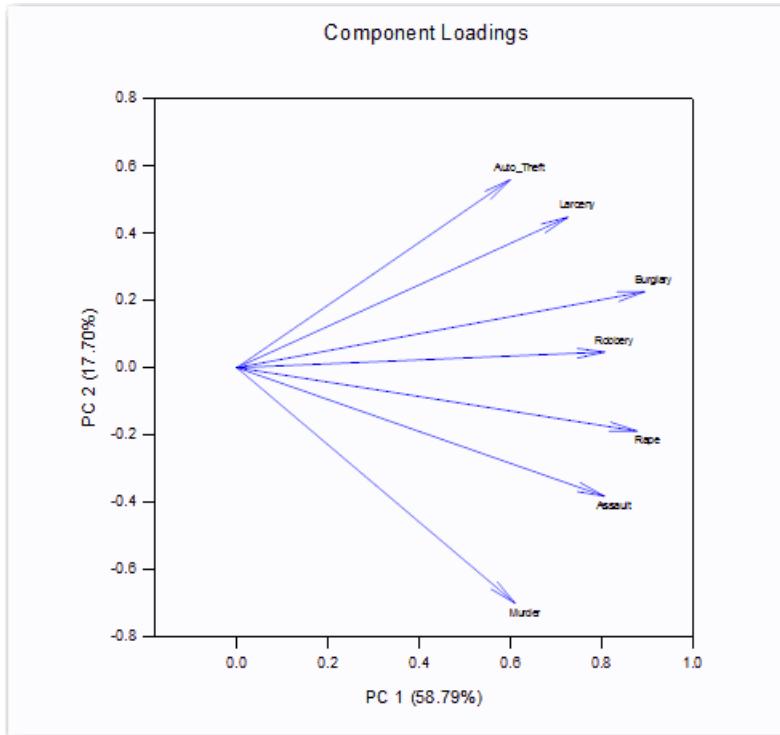


Figure 190: A Components Loading Plot

Component Scores Plot

A Component Scores plot is scatter plot of the coordinates of two principal components that you have selected. The Axes labels show the percentage of total variance contributed by corresponding component. The text label for each point in the graph refers to the row number of the observation corresponding to that point. The 95% prediction ellipse is used to indicate possible outliers in the distribution of the data.

The Component Scores Plot is available for the following tests:

- [Principal Component Analysis](#)

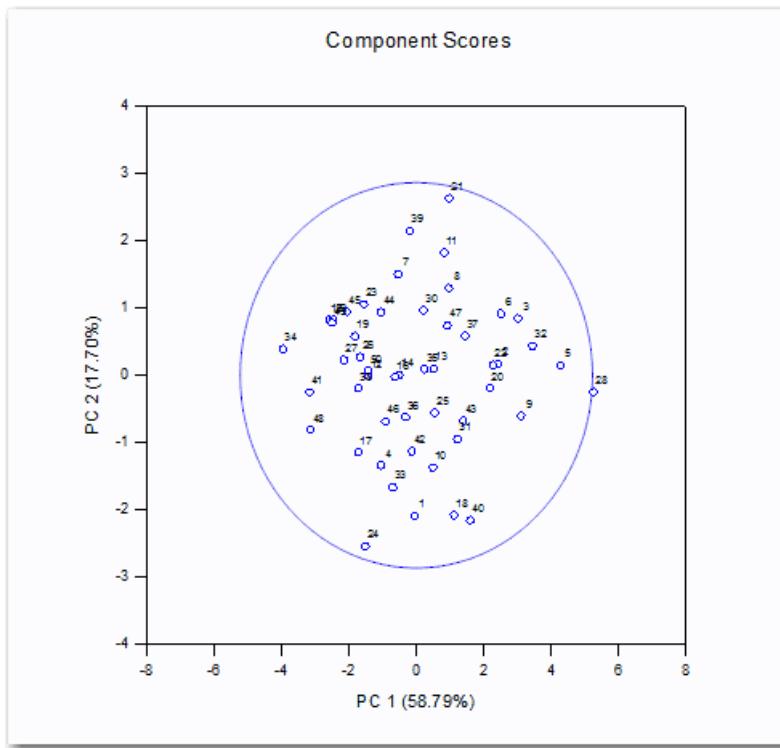


Figure 191: A Component Scores Plot

Regression Lines for Groups

A regression lines for groups graph is only available if you've used a [single covariate](#) while running a [One Way ANCOVA](#) test.

There are two types of regression lines for groups graphs:

- Equal slopes
- Interaction model

Each line is the graph of the ANCOVA regression equation for a particular group. The range of values on the horizontal axis corresponds to the data range of the covariate across all groups. In addition to the regression lines, there is a scatter plot of the observations for each group. The graph's legend shows the association of the line and symbol color with a group.

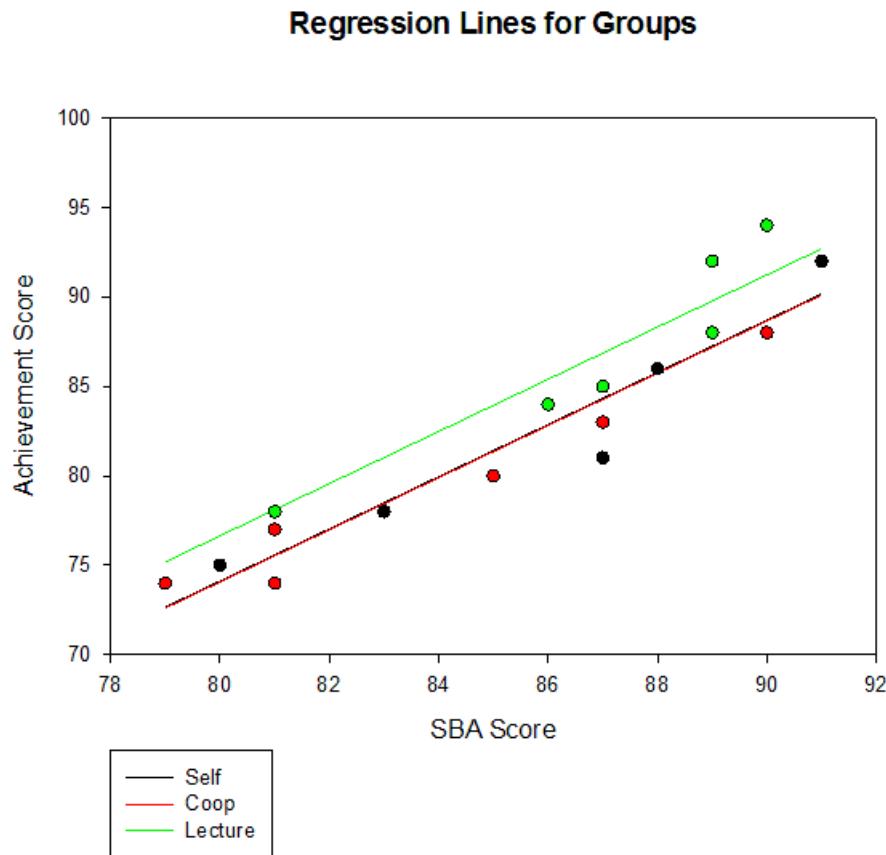


Figure 192: A Regression Line for Groups Graph -- Equal Slopes

Each line is the graph of the regression equation for the interaction model by specifying a particular group. This graph is only available if you are testing the assumption of equal slopes and the test fails. The range of values on the horizontal axis corresponds to the data range of the covariate across all groups. In addition to the regression lines, a scatter plot of the observations will be appears for each group. The graph will also contain a legend that shows the association of the line and symbol color with a group.

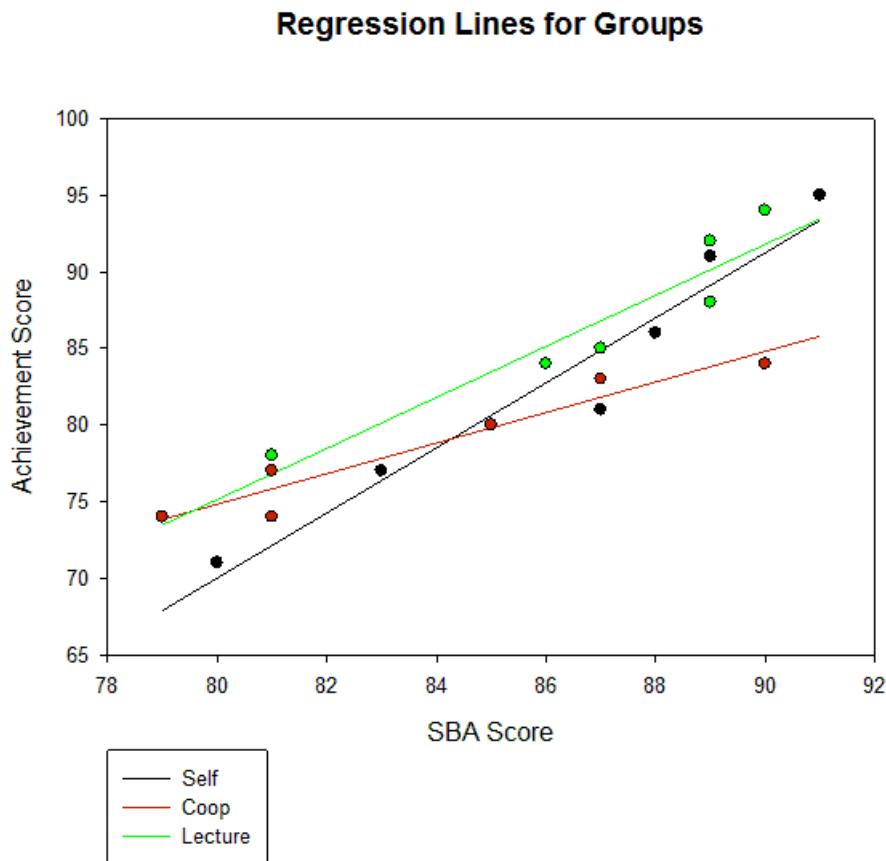


Figure 193: A Regression Line for Groups Graph -- Interaction Model

Adjusted Means with 95% Confidence Intervals

This report graph type is only available only if you are analyzing the equal slopes model while running a [One Way ANCOVA](#). This means that you've ever cleared the option to do the [Equality of Slopes Test](#) or the Equality of Slopes Test passed.

Each symbol is the adjusted mean corresponding to one of the groups. The error bars represent confidence intervals for the population values of the adjusted means. You select the confidence level, 95% as shown in this example, on the [More Results](#) tab of the [Options for One Way ANCOVA](#) dialog box.

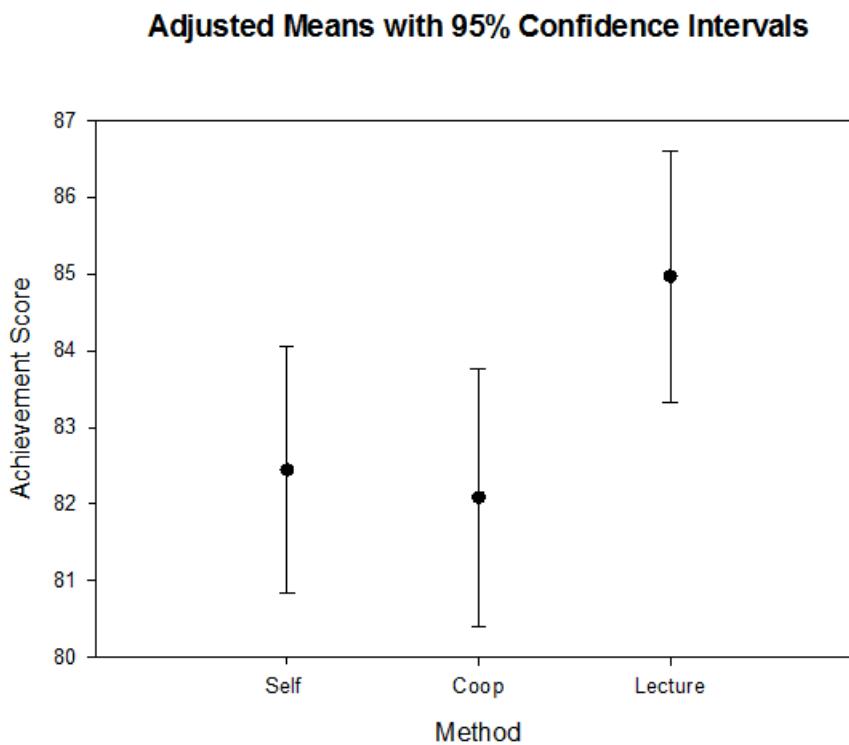


Figure 194: Adjusted Means with 95% Confidence Intervals Graph

Index

Numerics

- 3D category scatter graph
 - two way ANOVA 386
 - two way repeated measures ANOVA 386
- 3D residual scatter plot
 - multiple linear regression 384
 - stepwise regression 384
 - two way ANOVA 384
 - two way repeated measures ANOVA 384

A

- adj R2
 - multiple linear regression results 244
 - simple linear regression results 232
- adjusted means with 95% confidence intervals
 - one way ANCOVA 398
- Advisor Wizard
 - calculating power 8, 8, 8
 - calculating sample size 6, 8, 8
 - data format 10
 - defining your goals 6
 - determining sensitivity 6
 - independent variables 11
 - measuring data 7, 7
 - number of treatments 8
 - repeated observations 8
 - starting 5
 - using 5
 - viewing 5
- alpha value
 - in power 39
 - sample size 39
- ANCOVA
 - adjusted means with 95% confidence intervals 398
 - multiple comparison graphs 388
 - normal probability plot 382
 - regression lines for groups 396
 - scatter plot of the residuals 379
- ANOVA
 - arranging data 55
 - multiple comparison graphs 388
 - order only polynomial regression results 268
 - profile plots 390
- ANOVA on ranks
 - about 114
 - arranging data 115
 - box plot 378
 - multiple comparison options 119
 - report graphs 122
 - results 120
 - running 118
 - setting options 115
 - steps to run 118
 - when to use 8, 30
- ANOVA table
 - multiple linear regression results 245
 - simple linear regression results 232, 232, 232, 233, 233, 233, 233, 233, 234, 234, 235, 235, 235, 235, 236, 236
 - stepwise linear regression results 284
- arranging data
 - ANOVA on ranks 115
 - best subsets regression 290
 - chi-square test 194
 - Cox regression 335
 - Deming regression 302
 - descriptive statistics 19
 - Fisher exact test 200
 - Gehan-Breslow survival analysis 324
 - LogRank survival analysis 316
 - Mann-Whitney rank sum test 67
 - McNemar's test 204
 - multiple linear regression 238
 - multiple logistic regression 250
 - normality test 37
 - odds ratio test 212
 - one way ANCOVA 125
 - one way ANOVA 74
 - one way frequency tables 24
 - Pearson product moment correlation 295
 - polynomial regression 261
 - principal components analysis 217
 - rank sum test 67
 - relative risk test 208
 - simple linear regression 227
 - single group survival analysis 310
 - Spearman rank order correlation 299
 - stepwise linear regression 271
 - t-test 55
 - three way ANOVA 105
 - two way ANOVA 89
 - unpaired t-test 57
 - z-test 189
- assumption checking options
 - setting for backward stepwise regression 278
 - setting for forward stepwise regression 273
 - setting for linear regression 228
 - setting for multiple linear regression 239
 - setting for polynomial regression 262
 - setting for principal components analysis 218
- assumption testing
 - incremental polynomial regression results 267
- B
 - backward stepwise linear regression
 - setting options 277
 - backward stepwise regression
 - setting assumption checking options 278
 - setting criterion options 278
 - setting more statistics options 280
 - setting other diagnostics options 281
 - setting residual options 279
 - when to use 34

bar chart of the column means
 descriptive statistics 374
 one way ANOVA 374
 t-test 374

bar chart of the standardized residuals 380

bar charts
 descriptive statistics results 23
 one way frequency tables results 29

basic statistics
 describing your data with 18
 one way frequency tables 24

before & after procedures
 paired t-test 32
 signed rank test 32

before and after line plots
 one way repeated measures ANOVA 387
 paired t-test 387
 repeated measures ANOVA on ranks 387
 signed rank test 387

best subset regression
 setting criterion options 291
 when to use 12, 34

best subset regression results
 subsets results 293
 summary table 293

best subsets regression
 about 290
 arranging data 290
 criteria 290
 performing 290, 292
 running 292
 setting options 291

best subsets regression results
 interpreting 293

best subsets regression 289

beta
 multiple linear regression results

374

ANOVA on ranks
 387 P à ð € p @ 0` ð € à P € p ° @ à ð

dŒ \$BYñ †U

descriptive statistics results 1 dŒ \$BYñ †U à à P pp@ 8uXf
 @ € ð ö à @ @ ° 5 •&1fp

component loadings plot
 principal component analysis 394
 component scores plot
 principal component analysis 395
 computing
 calculating 360
 power 349
 sample size 349
 computing power
 chi-square analysis of contingency tables 349
 correlation efficient 349
 one way ANOVA 349
 paired t-test 349
 unpaired t-test 349
 z-test comparison of proportions 349
 computing sample size
 chi-square analysis of contingency tables 349
 correlation efficient 349
 one way ANOVA 349
 paired t-test 349
 unpaired t-test 349
 z-test comparison of proportions 349
 conditions
 number of 6, 8
 confidence interval
 descriptive statistics 19
 descriptive statistics results 21
 for the mean 21
 one way frequency tables 25
 confidence interval for the difference
 z-test results 193
 confidence intervals
 multiple linear regression results 244, 248
 order only polynomial regression results 269
 simple linear regression results 236
 stepwise linear regression results 288
 constant variance test
 multiple linear regression results 247
 order only polynomial regression results 269
 simple linear regression results 235
 stepwise linear regression results 287
 contingency table
 data format 10
 contingency table summary
 chi-square test results 198
 Fisher exact test results 203
 McNemar's test results 208
 contingency tables
 rate and proportion tests 188, 188
 continuous scale
 measuring data 7
 correlation
 data format 226
 Pearson product moment correlation 294
 Spearman rank order correlation 299
 when to use 35
 correlation coefficient
 best subset regression results 297
 calculating power 8
 correlation coefficients
 calculating power/sample size 39
 correlation efficient
 computing power 349
 computing sample size 349
 correlation method
 when to use 34
 correlation procedures
 Pearson Product Moment 35, 35
 Correlation procedures
 Spearman Rank Order 35
 correlation results 297
 Cox regression
 about 333
 arranging data 335
 performing 334, 334
 report graphs 343
 Cox regression proportional hazards
 performing 338
 running 338
 setting criterion options 335
 setting graph options 336
 setting options 335
 setting results options 336
 Cox regression results
 interpreting 341
 Cox regression stratified model
 performing 340
 running 340
 setting criterion options 337
 setting graph options 338
 setting options 336
 setting results options 338
 creating
 descriptive statistics report graph 23
 normality test report graph 39
 one way frequency tables report graph 29
 report graphs 374
 criterion options
 setting for backward stepwise regression 278
 setting for best subset regression 291
 setting for Cox regression proportional hazards 335
 setting for Cox regression stratified model 337
 setting for forward stepwise regression 272
 setting for multiple logistic regression 251
 setting for polynomial regression 261
 curve
 fitting through data 10
 polynomial 11

D

data
 arranging 19, 24
 data format 37
 describing 7, 11
 fitting curve through 10
 indexing for a Two-Way ANOVA 90
 measuring 7
 plotting residuals 39
 selecting to run a test 17
 data format
 contingency table 10
 correlation 226

- determining 10
 - group comparison tests 54
 - normality test 37
 - observed proportions 10
 - raw data 37
 - regression 226
 - survival analysis 308
- data summary table
 - Gehan-Breslow survival analysis results 332
 - LogRank survival analysis results 323
 - single group survival analysis results 315
- data:
 - observing 8
- Deming regression
 - about 301
 - arranging data 302
 - performing 302, 303
 - report graphs 304
 - running 303
- Deming regression results
 - interpreting 303
- dependent variable
 - multiple logistic regression results 256
- dependent variables
 - predicting 6, 10, 10
- descriptive statistics
 - arranging data for 19
 - bar chart of the column means 374
 - box plot 378
 - confidence interval 19
 - graphing data 23
 - group comparison tests 55
 - interpreting results 21
 - picking column data 20
 - point plot and column means 377
 - results 21
 - scatter plot 375
 - setting options 19, 21
 - viewing 7
- descriptive statistics results
 - bar chart 23
 - box plot 23
 - point and column means plot 23
 - point plot 23
- descriptive statistics results:
 - scatter plot 23
- determining to detect a specified correlation 358
- Durbin-Watson statistic
 - multiple linear regression results 246
 - order only polynomial regression results 268
 - simple linear regression results 234
 - stepwise linear regression results 286
- E**
 - editing
 - graphs 17
 - reports 17
 - equations
 - adding independent variables 12
 - nonlinear 11, 11
 - removing independent variables 12
- estimating
 - sample size for power 40
- estimation criterion
 - multiple logistic regression results 256
- example graphs
 - editing with the Graph Properties 345
 - survival analysis 344
 - using test options to modify graphs 346
- experimental power 349

F

- F-to-Enter
 - stepwise linear regression results 284
- F-to-Remove
 - stepwise linear regression results 284
- failures
 - survival analysis 344
- Fisher exact test
 - about 199
 - arranging data 200
 - performing 200, 201
 - running 201
 - when to use 33
- Fisher exact test results
 - contingency table summary 203
 - interpreting 202
 - P value 203
- fitting
 - curve through data 10
- forward stepwise linear regression
 - setting options 271
- forward stepwise regression
 - setting assumption checking options 273
 - setting criterion options 272
 - setting more statistics options 274
 - setting other diagnostics options 275
 - setting residual options 274
 - when to use 12, 34
- frequencies
 - comparing 187
- functions
 - nonlinear 11

G

- Gehan-Breslow survival analysis
 - arranging data 324
 - multiple comparison options 329
 - performing 324, 326
 - report graphs 332
 - running 326
 - setting graph options 325
 - setting options 324
 - setting post hoc test options 326
 - setting results options 325
- Gehan-Breslow survival analysis results
 - data summary table 332
 - interpreting 330
 - report header information 332
 - statistical summary table 332
 - survival cumulative probability table 332

goals

- defining 6
- predicting 10

graph options

- setting for Cox regression proportional hazards 336
- setting for Cox regression stratified model 338
- setting for Gehan-Breslow survival analysis 325
- setting for LogRank survival analysis 317
- setting for single group survival analysis 311

graphs

- 2D line and scatter plots of the regressions with prediction and confidence intervals 383
- 3D category scatter graph 386
- 3D residual scatter plot 384
- adjusted means with 95% confidence intervals 398
- bar chart of the column means 374
- bar chart of the standardized residuals 380
- before and after line plots 387
- box plot 378
- component loadings plot 394
- component scores plot 395
- creating 374
- descriptive statistics 23
- for ANCOVA reports 388
- for ANOVA reports 388
- grouped bar chart with error bars 385
- histogram of the residuals 381
- multiple comparison graphs 388
- normal probability plot 382
- one way frequency tables 29
- point plot 376
- point plot and column means 377
- profile plots 390
- profile plots — main effects 391
- profile plots — three way effects 392
- profile plots — two way effects 391
- regression lines for groups 396
- scatter matrix 389
- scatter plot 375
- scatter plot of the residuals 379
- scree plots 393

group comparison tests

- about 54
- choosing appropriate 29
- data format 54
- descriptive statistics 55
- nonparametric tests 54
- parametric tests 54
- when to use 29

grouped bar chart with error bars

- two way ANOVA 385

groups

- comparing many 30
- comparing two 29
- number of 8

H

Henze-Zirkler 297, 300

histogram of residuals

- normality test results 39

histogram of the residuals

- linear regression 381
- multiple linear regression 381
- nonlinear regression 381
- normality test 381
- one sample t-test 381
- one way ANOVA 381
- one way repeated measures ANOVA 381
- paired t-test 381
- polynomial regression 381
- stepwise regression 381
- t-test 381
- three way ANOVA 381
- two way ANOVA 381
- two way repeated measures ANOVA 381

Hosmer-Lemshow P value

- multiple logistic regression results 256

I

incremental polynomial regression results

- assumption testing 267
- choosing the best model 267
- incremental results 265
- interpreting 265
- regression equation 265

incremental results

- incremental polynomial regression results 265

incremental sum of squares

- multiple linear regression results 246

independent variables

- adding to equations 12
- predicting dependent variables 6, 10, 10, 34
- removing from equations 12
- selecting 12
- specifying 11

indexed data

- survival analysis 309

indexing data

- for a Two-Way ANOVA 90

influence diagnostics

- multiple linear regression results 248
- multiple logistic regression results 259
- stepwise linear regression results 288

influence diagnostics

- simple linear regression results 236

intercept

- finding for line 10

interpreting

- best subsets regression results 293
- chi-square test results 198
- Cox regression results 341
- Deming regression results 303
- Fisher exact test results 202
- Gehan-Breslow survival analysis results 330
- incremental polynomial regression results 265
- LogRank survival analysis results 321
- McNemar's test results 207
- multiple linear regression results 244
- multiple logistic regression results 255
- odds ratio test results 214
- order only polynomial regression results 267

Pearson product moment correlation results 297
 principal components analysis results 220
 proportion comparison results 191
 relative risk test results 211
 simple linear regression results 232
 single group survival analysis results 313
 Spearman rank order correlation results 300
 stepwise linear regression results 283
 z-test results 191
 interpreting results
 descriptive statistics 21
 one way frequency tables 27

K

K-S distance
 descriptive statistics results 21
 normality test results 39
 K-S Probability
 descriptive statistics results 21
 Kaplan-Meier survival curve 307
 kurtosis
 descriptive statistics results 21

L

least square means
 comparing using a profile plot 390
 likelihood ratio test statistic
 multiple logistic regression results 257
 line
 slope & intercept 10
 linear regression
 histogram of the residuals 381
 normal probability plot 382
 predicting variables 34
 setting assumption checking options 228
 setting more statistics options 230
 setting other diagnostics options 230
 setting residual options 229
 when to use 10, 34
 log likelihood statistic
 multiple logistic regression results 257
 LogRank survival analysis
 arranging data 316
 multiple comparison options 320
 performing 316, 318
 report graphs 323
 running 318
 setting graph options 317
 setting options 316
 setting post hoc test options 317
 setting results options 317
 LogRank survival analysis results
 data summary table 323
 interpreting 321
 report header information 322
 statistical summary table 323
 survival cumulative probability table 323

M

Mann-Whitney rank sum test
 about 66
 arranging data 67
 performing 67
 report graphs 237
 setting options 67
 Mann-Whitney Rank Sum Test
 when to use 29
 Mardia's test 297, 300
 maximum value
 descriptive statistics results 21
 McNemar's test
 about 203
 arranging data 204
 performing 203, 205
 running 205
 setting options 204
 McNemar's test results
 chi-square 207
 contingency table summary 208
 interpreting 207
 McNemar's test
 when to use 33
 mean
 describing your data with 18
 descriptive statistic results 21
 measuring data
 continuous scale 7
 nominal/ordinal scale 7, 7
 measuring sensitivity 360
 median
 describing your data with 18
 descriptive statistics results 21
 minimum sample size
 determining for chi-square test 356
 determining for one way ANOVA 354
 determining for paired t-test 351
 determining for proportions comparison 352
 determining for t-test 350
 minimum value
 descriptive statistics results 21
 missing values
 descriptive statistic results 21
 modifying graphs
 Graph Properties 345
 using test options 346
 more statistics options
 setting for backward stepwise regression 280
 setting for forward stepwise regression 274
 setting for linear regression 230
 setting for multiple linear regression 240
 setting for polynomial regression 263
 multiple comparison graphs
 ANCOVA 388
 ANOVA 388
 multiple comparison options
 ANOVA on ranks 119
 Gehan-Breslow survival analysis 329
 LogRank survival analysis 320
 one way ANOVA 79

setting 31, 33
 three way ANOVA 112
 two way ANOVA 96

multiple linear regression
 3D residual scatter plot 384
 about 238
 arranging data 238
 histogram of the residuals 381
 normal probability plot 382
 performing 238, 243
 report graphs 249
 running 243
 setting assumption checking options 239
 setting more statistics options 240
 setting options 238
 setting other diagnostics options 241
 setting residual options 240
 when to use 34

multiple linear regression results
 adj R2 244
 ANOVA table 245
 beta 244
 confidence intervals 244, 248
 constant variance test 247
 Durbin-Watson statistic 246
 incremental sum of squares 246
 influence diagnostics 248
 interpreting 244
 normality test 247
 power 247
 PRESS statistic 246
 R 244
 R2 244
 regression diagnostics 247
 standard error of the estimate 244
 statistical summary table 244

multiple logistic regression
 about 250
 arranging data 250
 performing 250, 255
 running 255
 setting criterion options 251
 setting options 250
 setting residual options 253
 setting statistics options 252
 when to use 34

multiple logistic regression results
 classification table 257
 dependent variable 256
 estimation criterion 256
 Hosmer-Lemeshow P value 256
 influence diagnostics 259
 interpreting 255
 likelihood ratio test statistic 257
 log likelihood statistic 257
 number of observations 256
 number of unique independent variable combinations 256
 Pearson chi-square statistic 256
 probability table 257
 regression equation 256
 residual calculation method 258
 residuals table 259

statistical summary table 258
 threshold probability for positive classification 257

N

N statistic
 descriptive statistic results 21

nominal (category) scale
 measuring data 7

non-normal populations
 testing 29, 32, 35

nonlinear equation
 describing data 11
 fitting curve through data 11
 when to use 11

nonlinear regression
 histogram of the residuals 381
 normal probability plot 382
 when to use 34

nonparametric tests
 signed rank test 32

normal probability plot
 linear regression 382
 multiple linear regression 382
 nonlinear regression 382
 normality test 382
 one sample t-test 382
 one way ANCOVA 382
 one way ANOVA 382
 one way repeated measures ANOVA 382
 paired t-test 382
 polynomial regression 382
 stepwise regression 382
 t-test 382
 three way ANOVA 382
 two way ANOVA 382
 two way repeated measures ANOVA 382

normality procedure
 normality test 36
 when to use 35

normality results 297, 300

normality test
 data format 37
 descriptive statistics results 21
 histogram of the residuals 381
 interpreting results 38
 multiple linear regression results 247
 normal probability plot 382
 normality procedure 36
 order only polynomial regression results 269
 performing 36
 picking data columns 38
 running 38
 setting P-value 36
 simple linear regression results 235
 stepwise linear regression results 287
 when to use 35

normality test results
 creating graphs 39
 histogram of residuals 39
 K-S distance 39
 normal probability plot of residuals 39

P value 39
 report graphs 39
 normally distributed populations
 testing 29, 30, 32, 32, 35
 number of observations
 multiple logistic regression results 256
 number of samples
 Spearman rank order correlation results 300
 number of unique independent variable combinations
 multiple logistic regression results 256
 numeric values
 measuring data 7

O

observations
 data 8
 repeated 8
 ochi-square test
 setting options 195
 odds ratio test
 about 212
 arranging data 212
 performing 212, 213
 running 213
 setting options 212
 Odds Ratio Test
 when to use 33
 odds ratio test results
 interpreting 214
 one sample t-test
 histogram of the residuals 381
 normal probability plot 382
 scatter plot 375
 one way analysis of covariance
 about 123
 one way analysis of variance (ANOVA)
 about 74
 performing 74
 one way ANCOVA
 about 123
 arranging data 125
 interpreting results 129
 performing 124
 report graphs 135
 running 129
 setting options 125
 steps to run 129
 one way ANOVA
 about 74
 arranging data 74
 bar chart of the column means 374
 calculating power/sample size 39
 determining minimum sample size 354
 determining power 365
 histogram of the residuals 381
 multiple comparison options 79
 normal probability plot 382
 performing 74
 report graphs 87
 results 79
 running 77

scatter plot 375
 setting options 74
 steps to run 77
 when to use 8, 30, 30, 33
 with two way ANOVA data 101
 one way frequency tables
 graphing data 29
 interpreting results 27
 picking column data 26
 results 27
 setting options 25, 27
 one way frequency tables results
 bar chart 29
 one way repeated measures ANOVA
 before and after line plots 387
 histogram of the residuals 381
 normal probability plot 382
 when to use 8, 32
 one-sample signed rank test
 about 48
 arranging data 49
 one-sample t-test 42
 One-Sample t-test
 about 42
 arranging data 42
 opening
 reports 17
 result graphs 17
 options
 ANOVA on ranks 115
 descriptive statistics 19
 multiple comparison 31, 33
 one way ANCOVA 125
 one way ANOVA 74
 one way frequency tables 25
 setting 16
 setting for backward stepwise linear regression 277
 setting for best subsets regression 291
 setting for chi-square test 195
 setting for Cox regression proportional hazards 335
 setting for Cox regression stratified model 336
 setting for forward stepwise linear regression 271
 setting for Gehan-Breslow survival analysis 324
 setting for LogRank survival analysis 316
 setting for Mann-Whitney rank sum test 67
 setting for McNemar's test 204
 setting for multiple linear regression 238
 setting for multiple logistic regression 250
 setting for odds ratio test 212
 setting for polynomial regression 261
 setting for principal components analysis 217
 setting for rank sum test 67
 setting for relative risk test 209
 setting for simple linear regression 228, 302
 setting for single group survival analysis 311
 setting for t-test 58
 setting for z-test 189
 three way ANOVA 108
 two way ANOVA 92
 order only polynomial regression results
 ANOVA 268
 confidence intervals 269

- constant variance test 269
 - Durbin-Watson statistic 268
 - interpreting 267
 - normality test 269
 - regression diagnostics 269
 - regression equation 268
 - standard error of the estimate 268
 - ordinal (rank) scale
 - measuring data 7
 - other diagnostics options
 - setting for backward stepwise regression 281
 - setting for forward stepwise regression 275
 - setting for linear regression 230
 - setting for multiple linear regression 241
- P**
- P value
 - best subset regression results 297
 - Fisher exact test results 203
 - normality test results 39
 - Spearman rank order correlation results 300
 - P-value
 - normality test 36
 - paired t-test
 - before and after line plots 387
 - computing power 349
 - computing sample size 349
 - determining minimum sample size 351
 - determining power 362
 - histogram of the residuals 381
 - normal probability plot 382
 - setting post hoc options 219
 - when to use 32
 - Paired t-test
 - when to use 32
 - parametric tests
 - paired t-test 32
 - Pearson chi-square statistic
 - multiple logistic regression results 256
 - Pearson product moment correlation
 - about 295
 - arranging data 295
 - computing its coefficient 295
 - detecting 358
 - performing 296
 - report graphs 294
 - running 296
 - scatter matrix 389
 - when to use 35
 - Pearson Product Moment Correlation
 - when to use 10
 - Pearson product moment correlation normality results
 - Henze-Zirkler 297
 - Mardia's test 297
 - Pearson product moment correlation results
 - correlation coefficient 297
 - interpreting 297
 - P value 297
 - percentiles
 - describing your data with 18
 - descriptive statistics results 21
 - performing
 - best subsets regression 290, 292
 - chi-square test 194, 196
 - Cox regression 334, 334
 - Cox regression proportional hazards 338
 - Cox regression stratified model 340
 - Deming regression 302, 303
 - Fisher exact test 200, 201
 - Gehan-Breslow survival analysis 324, 326
 - LogRank survival analysis 316, 318
 - McNemar's test 203, 205
 - multiple linear regression 238, 243
 - multiple logistic regression 250, 255
 - normality test 36
 - odds ratio test 212, 213
 - Pearson product moment correlation 296
 - polynomial regression 260, 264
 - power/sample size procedures 360
 - principal components analysis 217, 219
 - procedures 15
 - relative risk test 208, 210
 - simple linear regression 227, 231
 - single group survival analysis 310, 312
 - Spearman rank order correlation 299
 - stepwise linear regression 271, 283
 - z-test 189, 190
 - point and column means plots
 - descriptive statistics results 23
 - point plot
 - ANOVA on ranks 376, 376
 - descriptive statistics 376, 376
 - rank sum test 376, 376
 - t-test 376, 376
 - point plot and column means 377
 - point plots
 - descriptive statistics results 23
 - polynomial curve
 - fitting through data 11
 - polynomial regression
 - about 260
 - arranging data 261
 - histogram of the residuals 381
 - normal probability plot 382
 - performing 260, 264
 - report graphs 270
 - running 264
 - setting assumption checking options 262
 - setting criterion options 261
 - setting more statistics options 263
 - setting options 261
 - setting post hoc test options 264
 - setting residual options 263
 - when to use 11, 34
 - polynomial regression results
 - incremental 265
 - order only 267
 - post hoc options
 - setting for paired t-test 219
 - post hoc test options
 - setting for Gehan-Breslow survival analysis 326
 - setting for LogRank survival analysis 317
 - setting for polynomial regression 264

power
 alpha value 39
 calculating 6, 8, 8, 39
 chi-square test results 199
 computing 349
 determining for an intended test 39
 determining for chi-square test 367
 determining for one way ANOVA 365
 determining for paired t-test 362
 determining for z-test proportions comparison 363
 experimental 349
 multiple linear regression results 247
 performing procedure 360
 sample size 350
 simple linear regression results 235
 stepwise linear regression results 287
 t-test 360
 when to use 39
 z-test results 193

prediction method
 when to use 34

predicting
 goals 10
 variables and trends 6
 variables/trends 10, 34

prediction
 best subsets regression 289
 Deming regression 301
 multiple linear regression 237
 multiple logistic regression 249
 polynomial regression 260
 simple linear regression 227
 stepwise linear regression 270

prediction and confidence intervals
 2D line and scatter plots of the regressions 383

PRESS statistic
 multiple linear regression results 246
 simple linear regression results 234
 stepwise linear regression results 286

principal component analysis
 component loadings plot 394
 component scores plot 395
 scree plots 393

principal components analysis
 about 216
 arranging data 217
 performing 217, 219
 report graphs 221
 running 219
 scree plot 221
 setting assumption checking options 218
 setting options 217
 setting residuals options 219
 setting results options 219

principal components analysis results
 interpreting 220

probability plots
 normality test results 39

probability table
 multiple logistic regression results 257

procedures
 choosing appropriate 18

compare many groups 30
 compare two groups 29
 multiple comparison 31
 normality 35
 performing 15
 power 39, 360
 repeating 17
 running 16
 sample size 39, 360

profile plots
 ANOVA 390

profile plots — main effects
 three way ANOVA 391
 two way ANOVA 391

profile plots — three way effects
 three way ANOVA 392
 two way ANOVA 392

profile plots — two way effects
 three way ANOVA 391
 two way ANOVA 391

proportion comparison results
 interpreting 191

proportional hazards 335

proportions
 comparing 187
 measuring data by 7

proportions comparison
 determining minimum sample size 352

- relative risk test 208
- Yates correction 188
- z-test 188
- rates
 - comparing 187
- raw data
 - in normality tests 37
 - survival analysis 308
- regression
 - about 226
 - best subset 12
 - correlation 226
 - data format 226
 - defined 34
 - forward stepwise 12
 - linear 10
 - nonlinear 11
 - polynomial 11
 - stepwise 12
- regression diagnostics
 - multiple linear regression results 247
 - order only polynomial regression results 269
 - simple linear regression results 235
 - stepwise linear regression results 287
- regression equation
 - incremental polynomial regression results 265
 - multiple logistic regression results 256
 - order only polynomial regression results 268
 - simple linear regression results 232
- regression lines for groups
 - one way ANCOVA 396
- relative risk test
 - about 208
 - arranging data 208
 - performing 208, 210
 - running 210
 - setting options 209
- Relative Risk Test
 - when to use 33
- relative risk test results
 - interpreting 211
- repeated measures
 - choosing which test to use 31
- repeated measures ANOVA on ranks
 - before and after line plots 387
 - box plot 378
 - when to use 8, 32
- repeated observations 8
- repeating
 - procedures 17
- report graphs
 - ANOVA on ranks 122
 - Cox regression 343
 - Deming regression 304
 - Gehan-Breslow survival analysis 332
 - generating 374
 - LogRank survival analysis 323
 - Mann-Whitney rank sum test 237
 - multiple linear regression 249
 - normality test results 39
 - one way ANCOVA 135
 - one way ANOVA 87
- Pearson product moment correlation 294
- plotting residuals 39
- polynomial regression 270
- principal components analysis 221
- probability plots 39
- rank sum test 72
- single group survival analysis 315
- Spearman rank order correlation 301
- stepwise linear regression 289
- t-test 64
- two way ANOVA 103
- report header information
 - Gehan-Breslow survival analysis results 332
 - LogRank survival analysis results 322
 - single group survival analysis results 314
- reports
 - editing 17
 - generating 17
 - opening 17
 - saving 17
- residual calculation method
 - multiple logistic regression results 258
- residual options
 - setting for backward stepwise regression 279
 - setting for forward stepwise regression 274
 - setting for linear regression 229
 - setting for multiple linear regression 240
 - setting for multiple logistic regression 253
 - setting for polynomial regression 263
- residuals
 - defined 34
 - plotting 39
 - probability plots 39
- residuals options
 - setting for principal components analysis 219
- residuals table
 - multiple logistic regression results 259
- result graphs
 - editing 17
 - generating 17, 17
 - opening 17
 - saving 17
- results
 - ANOVA on ranks 120
 - descriptive statistics 21
 - interpreting for best subsets regression 293
 - interpreting for chi-square test 198
 - interpreting for Cox regression 341
 - interpreting for Deming regression 303
 - interpreting for Fisher exact test 202
 - interpreting for Gehan-Breslow survival analysis 330
 - interpreting for incremental polynomial regression 265
 - interpreting for LogRank survival analysis 321
 - interpreting for McNemar's test 207
 - interpreting for multiple linear regression 244
 - interpreting for multiple logistic regression 255
 - interpreting for odds ratio test 214
 - interpreting for order only polynomial regression 267
 - interpreting for Pearson product moment correlation 297
 - interpreting for principal components analysis 220
 - interpreting for relative risk test 211
 - interpreting for simple linear regression 232

- interpreting for single group survival analysis 313
- interpreting for Spearman rank order correlation 300
- interpreting for stepwise linear regression 283
- interpreting for t-test 62
- interpreting for z-test results 191
- interpreting proportion comparison results 191
- normality test 38
- one way ANCOVA 129
- one way ANOVA 79
- one way frequency tables 27
- rank sum test 71
- three way ANOVA 113
- two way ANOVA 102
- results options
 - setting for Cox regression proportional hazards 336
 - setting for Cox regression stratified model 338
 - setting for Gehan-Breslow survival analysis 325
 - setting for LogRank survival analysis 317
 - setting for principal components analysis 219
 - setting for single group survival analysis 312
- running
 - best subsets regression 292
 - chi-square test 196
 - Cox regression proportional hazards 338
 - Cox regression stratified model 340
 - Deming regression 303
 - descriptive test 20
 - Fisher exact test 201
 - Gehan-Breslow survival analysis 326
 - LogRank survival analysis 318
 - McNemar's test 205
 - multiple linear regression 243
 - multiple logistic regression 255
 - normality test 38
 - odds ratio test 213
 - one way frequency tables test 26
 - Pearson product moment correlation 296
 - polynomial regression 264
 - principal components analysis 219
 - procedures 15
 - relative risk test 210
 - simple linear regression 231
 - single group survival analysis 312
 - Spearman rank order correlation 299
 - stepwise linear regression 283
 - z-test 190
- S**
 - s
 - arranging data for 24
 - sample size
 - alpha value 39
 - calculating 6, 8, 8, 39
 - calculating for Chi-Square test 39
 - calculating for correlation coefficients 39
 - calculating for one way ANOVA 39
 - calculating for unpaired t-tests 39
 - calculating for z-tests 39
 - computing 349
 - defined 350
 - estimating to achieve a desired power 40
 - setting
 - assumption checking options
 - backward stepwise regression 278
 - forward stepwise regression 273
 - linear regression 228
 - multiple linear regression 239
 - polynomial regression 262
 - principal components analysis 218
 - criterion options
 - backward stepwise regression 278
 - best subset regression 291
 - Cox regression proportional hazards 335
 - Cox regression stratified model 337

forward stepwise regression 272
 multiple logistic regression 251
 polynomial regression 261
 setting graph options
 Cox regression proportional hazards 336
 Cox regression stratified model 338
 Gehan-Breslow survival analysis 325
 LogRank survival analysis 317
 single group survival analysis 311
 setting more statistics options
 backward stepwise regression 280
 forward stepwise regression 274
 linear regression 230
 multiple linear regression 240
 polynomial regression 263
 setting other diagnostics options
 backward stepwise regression 281
 forward stepwise regression 275
 linear regression 230
 multiple linear regression 241
 setting post hoc options
 paired t-test 219
 setting post hoc test options
 Gehan-Breslow survival analysis 326
 LogRank survival analysis 317
 polynomial regression 264
 setting residual options
 backward stepwise regression 279
 forward stepwise regression 274
 linear regression 229
 multiple linear regression 240
 multiple logistic regression 253
 polynomial regression 263
 setting residuals options
 principal components analysis 219
 setting results options
 Cox regression proportional hazards 336
 Cox regression stratified model 338
 Gehan-Breslow survival analysis 325
 LogRank survival analysis 317
 principal components analysis 219
 single group survival analysis 312
 setting statistics options
 multiple logistic regression 252
 setting test options 16
 settings
 descriptive statistics options 19
 multiple comparison options 31, 33
 one way frequency tables options 25
 Shapiro-Wilk Probability.
 descriptive statistics results 21
 Shapiro-Wilk W
 descriptive statistics results 21
 SigmaStat 3
 signed rank test
 before and after line plots 387
 when to use 32
 signed rank test:
 when to use 8
 simple linear regression
 about 227
 arranging data 227
 performing 227, 231
 running 231
 setting options 228, 302
 simple linear regression results
 adj R2 232
 ANOVA table 233
 beta 233
 confidence intervals 236
 constant variance test 235
 Durbin-Watson statistic 234
 influence diagnostics 236
 interpreting 232
 normality test 235
 power 235
 PRESS statistic 234
 R 232
 R2 232
 regression diagnostics 235
 regression equation 232
 standard error of the estimate 233
 statistical summary table 233
 single group analysis
 one-sample signed rank test 48
 one-sample t-test 42
 single group survival analysis
 arranging data 310
 performing 310, 312
 report graphs 315
 running 312
 setting graph options 311
 setting options 311
 setting results options 312
 single group survival analysis results
 data summary table 315
 interpreting 313
 report header information 314
 statistical summary table 315
 survival cumulative probability table 314
 skewness
 descriptive statistics results 21
 slope
 finding for line 10
 Spearman correlation coefficient rs
 Spearman rank order correlation results 300
 Spearman rank order correlation
 about 299
 arranging data 299
 computing its coefficient 299
 performing 299
 report graphs 301
 running 299
 scatter matrix 389
 when to use 35
 Spearman rank order correlation normality results
 Henze-Zirkler 300
 Mardia's test 300
 Spearman rank order correlation results
 interpreting 300
 number of samples 300
 P value 300
 Spearman correlation coefficient rs 300

standard deviation
 describing your data with 18
 descriptive statistics results 21

standard error
 descriptive statistic results 21

standard error of the estimate
 multiple linear regression results 244
 order only polynomial regression results 268
 simple linear regression results 233

statistical summary table
 simple linear regression results 233

statistical summary
 z-test results 192

statistical summary table
 Gehan-Breslow survival analysis results 332
 LogRank survival analysis results 323
 multiple linear regression results 244
 multiple logistic regression results 258
 single group survival analysis results 315

statistics
 ANOVA 73
 ANOVA on ranks 114
 comparing many groups 54
 comparing two or more groups 53, 54, 54
 descriptive 7
 group comparison tests 54, 54, 54, 54
 Mann-Whitney rank sum test 66
 nonparametric tests 54
 one way analysis of covariance 123
 one way analysis of variance (ANOVA) 73
 one way ANCOVA 123
 overview 3
 parametric tests 54
 random samples 53
 rank sum test 66
 running procedures 16
 selecting a test to run 16
 three way ANOVA 104
 two way ANOVA 89
 unpaired t-test 57
 using procedures 15

statistics options
 setting for multiple logistic regression 252

step
 stepwise linear regression results 284

steps to run a statistical test 16

stepwise linear regression
 about 270
 arranging data 271
 performing 271, 283
 report graphs 289
 running 283

stepwise linear regression results
 ANOVA table 284
 confidence intervals 288
 constant variance test 287
 Durbin-Watson statistic 286
 F-to-Enter 284
 F-to-Remove 284
 influence diagnostics 288
 interpreting 283
 normality test 287

power 287
 PRESS statistic 286
 regression diagnostics 287
 step 284
 variables in model 285
 variables not in model 286

stepwise regression
 3D residual scatter plot 384
 backward 34
 forward 34
 histogram of the residuals 381
 normal probability plot 382
 when to use 12, 34, 34

Student's t-test 59

subsets results
 best subset regression results 293

sum
 descriptive statistics results 21

sum of squares
 descriptive statistics results 21

summary table
 best subset regression results 293

survival analysis
 censored values 344
 Cox regression 333
 Cox regression — proportional hazards model 308
 Cox regression — stratified model 308
 data format 308
 example graphs 344
 failures 344
 Gehan-Breslow 308, 324
 Gehan-Breslow survival analysis 324
 indexed data 309
 LogRank 308, 316
 LogRank survival analysis 316
 raw data 308
 single group 308, 310
 single group survival analysis 310
 ties 344
 when to use 35

survival cumulative probability table
 Gehan-Breslow survival analysis results 332
 LogRank survival analysis results 323
 single group survival analysis results 314

T

t-test
 arranging data 55
 bar chart of the column means 374
 determining minimum sample size 350
 histogram of the residuals 381
 interpreting results 62
 normal probability plot 382
 report graphs 64
 running 61
 scatter plot 375
 setting options 58
 when to use 29

t-tests
 paired 32
 power 360

test goals
 defining 6
 predicting 10

testing
 non-normal populations 29, 35
 normally distributed populations 29, 30, 32, 32, 35

Testing
 non-normal populations 32

tests
 choosing appropriate 18
 defining goals 6
 group comparison 29
 measuring effect 6
 measuring strength of association between a treatment and an event 6
 normality 35
 rank sum 8
 repeating 17
 selecting data 17
 setting options 16
 signed rank 8

three way ANOVA
 about 104
 arranging data 105
 histogram of the residuals 381
 multiple comparison options 112
 normal probability plot 382
 performing 105
 profile plots — main effects 391
 profile plots — three way effects 392
 profile plots — two way effects 391
 report graphs 114, 114
 results 113
 running 111
 setting options 108
 steps to run 111
 when to use 30

threshold probability for positive classification
 multiple logistic regression results 257

ties
 survival analysis 344

treatments
 number of 6, 8

trends
 predicting 6

two way ANOVA
 3D category scatter graph 386
 3D residual scatter plot 384
 about 89
 arranging data 89
 grouped bar chart with error bars 385
 histogram of the residuals 381
 multiple comparison options 96
 normal probability plot 382
 performing 89
 profile plots — main effects 391
 profile plots — three way effects 392
 profile plots — two way effects 391
 report graphs 103
 results 102
 running 95
 setting options 92

steps to run 95
 when to use 8, 30, 30

two way repeated measures ANOVA
 3D category scatter graph 386
 3D residual scatter plot 384
 histogram of the residuals 381
 normal probability plot 382
 when to use 8, 32

two way RM ANOVA:
 when to use 33

Two-Way ANOVA
 indexing data 90

U

unpaired t-test
 about 57
 arranging data 57
 calculating power/sample size 39
 computing power 349
 computing sample size 349
 performing 57
 power 360

V

values
 alpha 39

variables
 measuring strength 10
 predicting 6, 10, 34
 quantifying strength of association 35
 selecting independent 12
 specifying independent 11

variables in model
 stepwise linear regression results 285

variables not in model
 stepwise linear regression results 286

viewing
 descriptive statistics 7

W

Welch's t-test 59

Wilcoxon signed rank test
 signed rank test 32
 when to use 32

Z

z statistic
 z-test results 192

z-test
 about 189
 arranging data 189
 calculating power/sample size 39
 determining power 363
 performing 189, 190
 running 190
 setting options 189
 when to use 33

z-test comparison of proportions
computing power 349
computing sample size 349

z-test results
confidence interval for the difference 193
interpreting 191
statistical power 193
statistical summary 192
z statistic 192