



Calc Guide

Chapter 9 *Data Analysis*

Using Scenarios, Goal Seek, Solver, Statistics, others

Copyright

This document is Copyright © 2019 by the LibreOffice Documentation Team. Contributors are listed below. You may distribute it and/or modify it under the terms of either the GNU General Public License (<http://www.gnu.org/licenses/gpl.html>), version 3 or later, or the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), version 4.0 or later.

All trademarks within this guide belong to their legitimate owners.

Contributors

This book is adapted and updated from the *LibreOffice 4.1 Calc Guide*.

To this edition

Steve Fanning
Olivier Hallot

Kees Kriek
Jean Hollis Weber

Zachary Parliman

To previous editions

Barbara Duprey

Jean Hollis Weber

John A Smith

Feedback

Please direct any comments or suggestions about this document to the Documentation Team's mailing list: documentation@global.libreoffice.org



Note

Everything you send to a mailing list, including your email address and any other personal information that is written in the message, is publicly archived and cannot be deleted.

Publication date and software version

Published December 2019. Based on LibreOffice 6.2.

Using LibreOffice on macOS

Some keystrokes and menu items are different on macOS from those used in Windows and Linux. The table below gives some common substitutions for the instructions in this chapter. For a more detailed list, see the application Help.

<i>Windows or Linux</i>	<i>macOS equivalent</i>	<i>Effect</i>
Tools > Options menu	LibreOffice > Preferences	Access setup options
Right-click	Control + click or right-click depending on computer setup	Open a context menu
Ctrl (Control)	⌘ (Command)	Used with other keys
F5	Shift + ⌘ + F5	Open the Navigator
F11	⌘ + T	Open the sidebar Styles panel

Contents

Copyright	2
Contributors.....	2
To this edition.....	2
To previous editions.....	2
Feedback.....	2
Publication date and software version.....	2
Using LibreOffice on macOS.....	2
Introduction	5
Consolidating data	5
Consolidating data example.....	6
Consolidation settings.....	6
Consolidation example.....	7
Creating subtotals	8
Using the SUBTOTAL function.....	8
A SUBTOTAL example.....	8
Using Data > Subtotals.....	10
Using the Subtotals tool.....	11
Subtotal outline.....	11
Subtotals tool options.....	12
Reset and Remove.....	13
Using scenarios	13
Creating scenarios.....	14
Scenario settings.....	15
Changing scenarios.....	16
Changing scenario properties.....	16
Changing scenario cell values.....	16
Working with scenarios using the Navigator.....	17
Using the Multiple Operations tool	17
Multiple Operations with one formula and one variable.....	18
An example with one formula and one variable.....	18
Calculating with several formulas simultaneously.....	20
An example with two formulas and one variable.....	20
Multiple operations with two variables.....	22
Calculating with two variables.....	22
Using Goal Seek	23
Goal Seek example.....	24
Using the Solver	25
Solver dialog.....	26
Solver options.....	28
Solver example.....	28
Using Statistics tools	29
Sampling tool.....	30
Descriptive Statistics tool.....	31

Analysis of Variance (ANOVA) tool.....	33
Correlation tool.....	34
Covariance tool.....	35
Exponential Smoothing tool.....	36
Moving Average tool.....	37
Regression tool.....	39
Paired T-test tool.....	41
F-test tool.....	43
Z-test tool.....	44
Test of Independence (Chi-Square) tool.....	46

Introduction

Once you are familiar with functions and formulas, the next step is to learn how to use Calc's automated processes to perform useful analysis of your data quickly.

In addition to formulas and functions, Calc includes several tools for processing your data. These tools include features for copying and reusing data, creating subtotals, running what-if analysis, and performing statistical analysis. They can be found under the **Tools** and **Data** menus in the Menu bar. While they are not necessary to using Calc, they can help you save time and effort with handling large data sets or preserving your work for future review.



Note

A related tool, the Pivot Table, is not mentioned here, as it is complex enough to require its own chapter. See Chapter 8, Using Pivot Tables, for more information.

Consolidating data

The Consolidate tool allows you to combine and aggregate data spread across one or more worksheets into one place. This tool is useful if you need to quickly summarize a large, scattered set of data for review. For example, you could use it to consolidate multiple department budgets from different worksheets into a single company-wide budget contained in a master worksheet.

To access the Consolidate tool dialog, select **Data > Consolidate** from the Menu bar (Figure 1).

Consolidate

Function:
Sum

Consolidation ranges:
\$Year 1'.\$A\$1:\$E\$5
\$Year 2'.\$A\$1:\$E\$5

Source data ranges:
- undefined -

Copy results to:
- undefined - \$Consolidated Sales'.\$A\$1

Add Delete

Options

Consolidate by
 Row labels
 Column labels

Options
 Link to source data

Help OK Cancel

Figure 1: Consolidate tool dialog

Consolidating data example

To consolidate data:

- 1) Open the document containing the cell ranges to be consolidated.
- 2) Select **Data > Consolidate** from the Menu bar to open the Consolidate dialog (Figure 2).
- 3) Click on the *Source data ranges* field, then type a reference to a source data range, a named range or select it with the mouse. Use the associated **Shrink / Expand** button if you need to minimize the dialog while you select the range. Alternatively, select a named range from the drop-down list to the left of the field.
- 4) Click **Add**. The selected range is added to the *Consolidation ranges* list.
- 5) Repeat steps 3 and 4 to add additional source ranges.
- 6) To delete an entry in the *Consolidation ranges* list, select it and click **Delete**. The deletion is carried out without further confirmation.
- 7) Click on the *Copy results to* field, then type a reference to the first cell of the target range or select it with your mouse. You can also select a named range from the drop-down list to the left of the field.
- 8) Select a function to aggregate your data from the *Function* drop-down list. The default is Sum. Other available functions include Count, Average, Max, Min, Product, Count (numbers only), StdDev (sample), StDevP (population), Var (sample), and VarP (population).
- 9) Click **OK** to consolidate the ranges. Calc runs the function from step 8 on your source data ranges and populates the target range with the results.



Tip

If you are consolidating the same cell ranges repeatedly, consider converting them into reusable named ranges to make the process easier. For more information about named ranges, see Chapter 13, Calc as a Simple Database.

Consolidation settings

In the Consolidate dialog, click on the '+' sign next to *Options* to access the following settings (Figure 2):

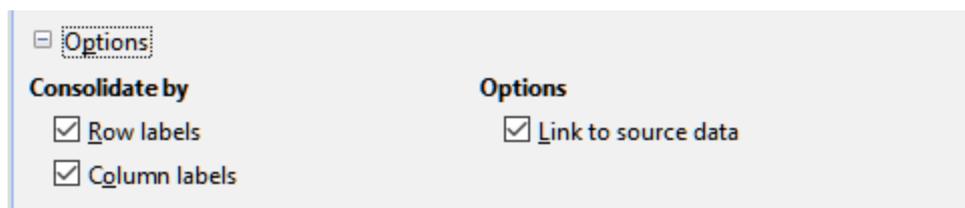


Figure 2: Consolidate tool options

Consolidate by

In this section, choose whether to consolidate your source data ranges by their range position or by matching labels. Consolidation labels must be contained within each range, and the text of corresponding row or column labels must match for the rows or columns themselves to be combined.

- *Row labels* – Consolidates rows by matching label. If this option is unchecked, the tool will consolidate rows by position instead.
- *Column labels* – Works the same as *Row labels*, but with columns instead.

Options

In this section, select *Link to source data* to add formulas in the target range that link back to values in the source ranges. When you select this option, any changes you make in the source ranges will automatically update values in the target range.



Note

If you use the *Link to source data* option, each source link is inserted into the target range, then ordered and hidden from view. Only the final results of consolidation are displayed by default.

Consolidation example

Figures 3, 4, and 5 show a simple example of consolidation using a spreadsheet with the worksheets *Year 1*, *Year 2*, and *Consolidated Sales*. Figure 3 shows the contents of the *Year 1* worksheet, with sales figures by region for each of the four product colors.

	A	B	C	D	E
1		Red	Orange	Green	Blue
2	North	\$84,213	\$87,445	\$5,930	\$31,653
3	East	\$92,609	\$94,959	\$44,466	\$28,606
4	South	\$67,987	\$19,169	\$92,682	\$28,253
5	West	\$29,628	\$95,213	\$74,880	\$10,334
6					
7					
8					
9					

< | < > > + | **Year 1** | Year 2 | Consolidated Sales

Figure 3: Year 1 sales by region

Figure 4 shows the *Year 2* worksheet, sales figures by region for each of four product colors. Note the different ordering of row and column labels between the two figures.

	A	B	C	D	E
1		Red	Green	Blue	Orange
2	North	\$11,716	\$94,891	\$68,566	\$71,120
3	South	\$86,039	\$20,239	\$1,194	\$83,642
4	East	\$18,326	\$69,337	\$59,089	\$48,803
5	West	\$60,579	\$25,280	\$75,349	\$47,605
6					
7					
8					
9					

< | < > > + | **Year 1** | **Year 2** | Consolidated Sales

Figure 4: Year 2 sales by region

Figure 5 shows the consolidated sales data, created by using the Consolidate dialog settings shown in Figure 2. Note that because the **Link to source data** option was selected, clicking on the plus sign (+) indicators to the left of the data will reveal formula links back to the source ranges.

	A	B	C	D	E
1		Red	Orange	Green	Blue
4	North	\$95,929	\$158,565	\$100,821	\$100,219
7	East	\$110,935	\$143,762	\$113,803	\$87,695
10	South	\$154,026	\$102,811	\$112,921	\$29,447
13	West	\$90,207	\$142,818	\$100,160	\$85,683
14					
15					
16					
17					

Figure 5: Consolidated sales by region

The source ranges and target range are saved as part of the document. If you later open a document with consolidated ranges, they will still be available in the *Consolidation ranges* list of the Consolidate dialog.

Creating subtotals

Calc offers two methods of creating subtotals: the SUBTOTAL function and the Subtotals tool.

Using the SUBTOTAL function

The SUBTOTAL function is listed under the Mathematical category of the Function Wizard, which is described in Chapter 7, Using Formulas and Functions. SUBTOTAL is a relatively limited method for generating a subtotal, and works best if used with only a few categories.

A SUBTOTAL example

To illustrate how to use the SUBTOTAL function, we use the sales data worksheet shown in Figure 6. The AutoFilter function is already applied to the sales data, and are indicated by the down arrow buttons at the head of each column. AutoFilters are described in Chapter 2, Entering, Editing, and Formatting Data.

	A	B	C	D	E
1	Date ▼	Sales Value ▼	Category ▼	Region ▼	Employee ▼
2	04/13/08	\$498	Sailing	North	Kurt
3	02/07/08	\$1,383	Sailing	South	Kurt
4	05/28/08	\$4,655	Tennis	East	Hans
5	04/20/08	\$3,993	Sailing	East	Fritz
6	02/28/08	\$3,377	Golf	South	Ute
7	01/27/08	\$2,095	Sailing	North	Brigitte
8	02/16/08	\$4,731	Sailing	East	Ute
9	04/15/08	\$4,330	Tennis	South	Kurt
10	04/20/08	\$3,664	Golf	South	Kurt

Figure 6: Sales data with AutoFilter applied (only the first few rows are shown)

To create a summation subtotal for the *Sales Value* field:

- 1) Select the cell to contain a subtotal. Typically, this cell is at the bottom of the column being subtotaled, which, for our example, is the *Sales Value* column.
- 2) Use one of the following methods to open the Function Wizard dialog:
 - The **Function Wizard** icon from the Formula bar
 - **Insert > Function** from the Menu bar
 - *Ctrl+F2* key combination

- 3) Select SUBTOTAL from the *Function* list in the Function Wizard dialog and click **Next**.
- 4) Enter the numeric code of a function into the *Function* field of the dialog. This code must be a value in the range 1 to 11 or 101 to 111, with the meaning of each value shown in Table 1.



Note

Values 1 to 11 include hidden values in the calculated subtotal, while values 101 to 111 do not. Hiding and showing data is described in Chapter 2, Entering, Editing, and Formatting Data. Filtered-out cells are always excluded by the SUBTOTAL function.

Table 1: SUBTOTAL function numbers

<i>Function index (includes hidden values)</i>	<i>Function index (ignores hidden values)</i>	<i>Function</i>
1	101	AVERAGE
2	102	COUNT
3	103	COUNTA
4	104	MAX
5	105	MIN
6	106	PRODUCT
7	107	STDEV
8	108	STDEVP
9	109	SUM
10	110	VAR
11	111	VARP

- 5) Click on the *Range* field, then type a reference to the *Sales Value* range or select the cells with your mouse (Figure 7). Use the **Shrink / Expand** button if you need to temporarily minimize the dialog while selecting the cells.

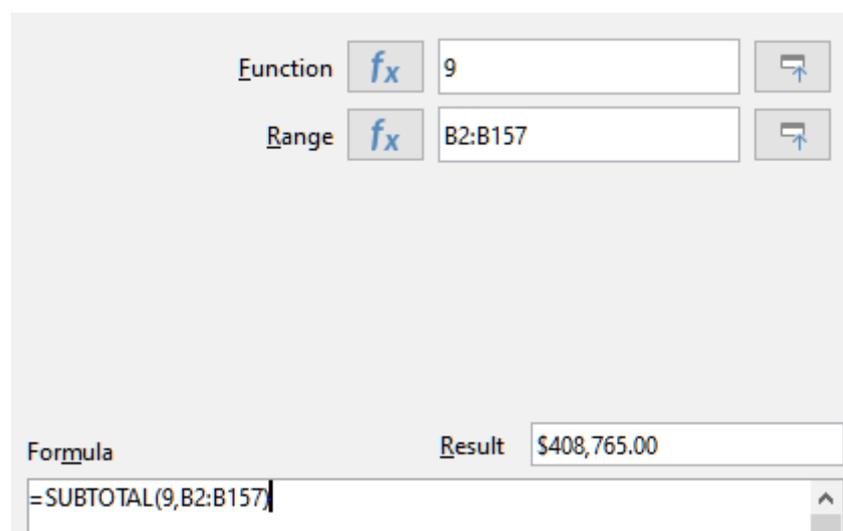


Figure 7: Function Wizard dialog

- 6) Click **OK** to close the Function Wizard dialog. The cell you selected in step 1 now contains the total sales value.

- 7) Click on the AutoFilter down arrow button at the top of the *Employee* column and remove all marks from the *Standard Filter* area except those next to *Brigitte* and (*empty*). The cell that you selected in step 1 should now reflect the sum of all of Brigitte's sales (Figure 8).

114	05/15/08	\$520	Sailing	North	Brigitte
115	05/06/08	\$575	Golf	West	Brigitte
118	06/02/08	\$2,213	Golf	East	Brigitte
119	06/03/08	\$1,819	Golf	East	Brigitte
122	04/18/08	\$3,060	Golf	West	Brigitte
124	02/20/08	\$1,660	Golf	South	Brigitte
127	04/02/08	\$1,355	Sailing	North	Brigitte
134	02/27/08	\$4,433	Sailing	West	Brigitte
138	01/25/08	\$3,116	Sailing	West	Brigitte
148	06/20/08	\$3,617	Sailing	North	Brigitte
151	02/27/08	\$1,790	Golf	East	Brigitte
158		\$69,104			

Figure 8: SUBTOTAL result for Brigitte's sales



Note

If the cell range used to calculate a subtotal contains other subtotals, these subtotals will not be counted in the final one. Similarly, if you use this function with AutoFilters, only the data satisfying the current filter selections will be displayed. Any filtered-out data is ignored.

Using Data > Subtotals

Calc offers the Subtotals tool as a more comprehensive alternative to the SUBTOTAL function. In contrast to SUBTOTAL, which only works on a single array, the Subtotals tool can create subtotals for up to three arrays arranged in labeled columns. It also groups subtotals by category and sorts them automatically, thereby eliminating the need to apply AutoFilters and filter categories by hand.

To use this tool, select **Data > Subtotals** from the Menu bar, which opens the Subtotals dialog.

Figure 9: Subtotals dialog

Using the Subtotals tool

To insert subtotal values into a worksheet:

- 1) Select the cell range for the subtotals that you want to calculate, and remember to include the column heading labels. Alternatively, click on a single cell within your data to allow Calc to automatically identify the range.
- 2) Select **Data > Subtotals** from the Menu bar to open the Subtotals dialog (Figure 9).
- 3) In the *Group by* drop-down list on the *1st Group* page, select a column by its label. Entries in the cell range from step 1 will be grouped and sorted by matching values in this column.
- 4) In the *Calculate subtotals for* box on the *1st Group* page, select a column containing values to be subtotaled. If you later change values in this column, Calc will automatically recalculate the subtotals.
- 5) In the *Use function* box on the *1st Group* page, select a function to calculate the subtotals for the column selected in step 4.
- 6) Repeat steps 4 and 5 to create subtotals for other columns on the *1st Group* page.
- 7) You can create two more subtotal categories by using the *2nd Group* and *3rd Group* pages and repeating steps 3 to 6. If you do not want to add more groups, then leave the *Group by* list for each page set to “- none -”.
- 8) Click **OK**. Calc will add subtotal and grand total rows to your cell range.

For our sales data example, a partial view of the results is shown in Figure 10. The group settings are identified in Table 2.

Table 2: Group settings used on Subtotals dialog for example sales data

Tab	Group by	Calculate subtotals for	Use function
1st Group	Employee	Sales Value	Sum
2nd Group	Category	Sales Value	Sum
3rd Group	- none -	-	-

Subtotal outline

When you use the Subtotals tool, Calc inserts an outline to the left of the row number column. This outline represents the hierarchical structure of your subtotals, and can be used to hide or show data at different levels in the hierarchy using the numbered column indicators at the top of the outline or the group indicators, denoted by plus (+) and minus (-) signs.

This feature is useful if you have many subtotals, as you can simply hide low-level details, such as individual entries, to produce a high-level summary of your data. For more information on how to use outlines, see Chapter 2, Entering, Editing, and Formatting Data.

To turn off outlines, select **Data > Group and Outline > Remove Outline** from the Menu bar. To reinstate them, select **Data > Group and Outline > AutoOutline**.

Figure 10 shows the outline for our sales data example. Column 1 represents the highest group level, the grand total over all employees. Columns 2 to 5 show descending group levels as follows:

- Column 2 represents the grand total over all categories.
- Column 3 represents the total for each employee.
- Column 4 represents the total for each category for an individual employee.
- Column 5 shows individual entries.

1	2	3	4	5	A	B	C	D	E	
					1	Date	Sales Value	Category	Region	Employee
-	-	-	-		2	01/19/08	\$2,032	Golf	West	Brigitte
					3	01/29/08	\$3,491	Golf	South	Brigitte
					4	02/20/08	\$1,660	Golf	South	Brigitte
					5	02/27/08	\$1,146	Golf	West	Brigitte
					6	02/27/08	\$1,790	Golf	East	Brigitte
					7	03/18/08	\$3,814	Golf	North	Brigitte
					8	04/04/08	\$4,045	Golf	West	Brigitte
					9	04/18/08	\$3,060	Golf	West	Brigitte
					10	04/29/08	\$1,316	Golf	West	Brigitte
					11	05/06/08	\$575	Golf	West	Brigitte
					12	06/02/08	\$2,213	Golf	East	Brigitte
					13	06/03/08	\$1,819	Golf	East	Brigitte
					14		\$26,961	Golf Sum		
				-	15	01/25/08	\$3,116	Sailing	West	Brigitte
					16	01/27/08	\$2,095	Sailing	North	Brigitte
					17	01/28/08	\$155	Sailing	West	Brigitte
					18	02/27/08	\$4,433	Sailing	West	Brigitte
					19	04/02/08	\$1,355	Sailing	North	Brigitte
					20	04/25/08	\$2,564	Sailing	North	Brigitte
					21	05/12/08	\$3,887	Sailing	South	Brigitte
					22	05/15/08	\$520	Sailing	North	Brigitte
					23	06/04/08	\$2,135	Sailing	East	Brigitte
					24	06/20/08	\$3,617	Sailing	North	Brigitte
					25	06/28/08	\$545	Sailing	South	Brigitte
					26		\$24,422	Sailing Sum		
				-	27	01/03/08	\$4,872	Tennis	East	Brigitte
					28	02/02/08	\$1,218	Tennis	West	Brigitte
					29	02/11/08	\$2,565	Tennis	West	Brigitte
					30	02/28/08	\$3,145	Tennis	West	Brigitte
					31	03/03/08	\$3,985	Tennis	North	Brigitte
					32	06/07/08	\$769	Tennis	West	Brigitte
					33	06/08/08	\$1,167	Tennis	West	Brigitte
					34		\$17,721	Tennis Sum		
					35		\$69,104			Brigitte Sum

Figure 10: Partial outlined view of sales data example with subtotals

Subtotals tool options

Click on the *Options* tab of the Subtotals dialog to access the following settings:

Groups

Defines how subtotal data is organized.

- *Page break between groups* – inserts page breaks between each subtotal group so that each group displays on a separate page when you print the data.
- *Case sensitive* – prevents the tool from grouping entries by data labels that differ by case. In our sales data example, entries with “Brigitte” and “brigitte” under the Employee column will not match if this option is selected.
- *Pre-sort area according to groups* – sorts entries by group before calculating subtotals. Disabling this option prevents the tool from grouping matching entries together. As a result, distinct subtotals will be created for matching entries if they do not appear on consecutive rows. For example, two entries under the “Golf” category will not count towards the same group subtotal if there is an entry for “Tennis” in between them.

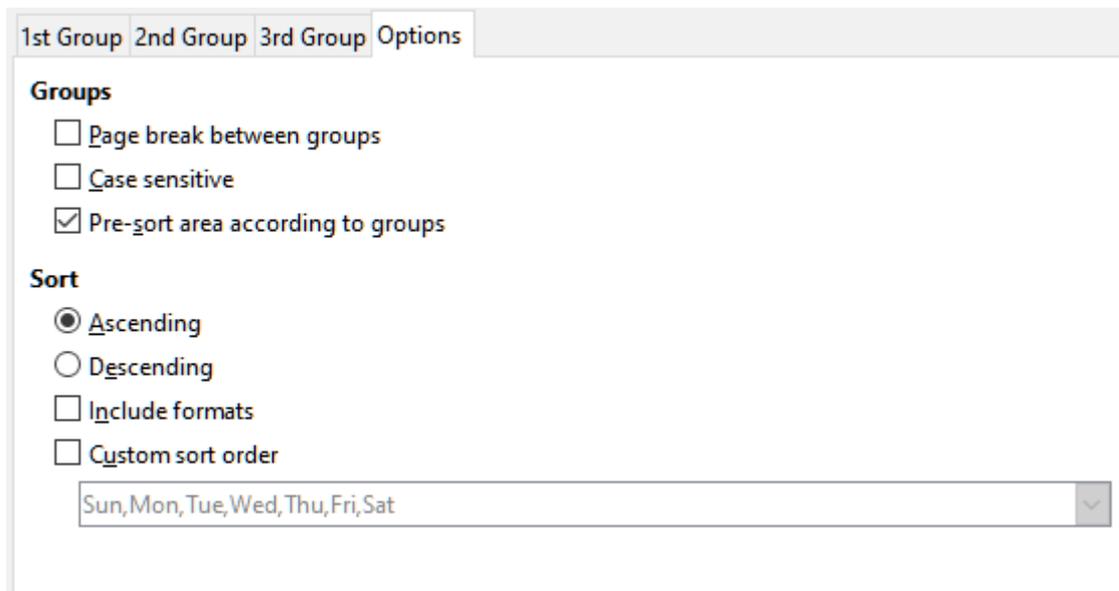


Figure 11: Options page of the Subtotals dialog

Sort

Defines how subtotal data is sorted. This section is disabled if the *Pre-sort area according to groups* is unchecked.

- *Ascending* or *Descending* – sorts entries by value from lowest to highest and highest to lowest, respectively. You can modify these sort rules by using **Data > Sort** in the Menu bar. For more detail, see Chapter 2, Entering, Editing, and Formatting Data.
- *Include formats* – carries over formatting, such as the currency format, from the data to the corresponding subtotals.
- *Custom sort order* – sorts your data according to one of the predefined custom sorts defined in **Tools > Options > LibreOffice Calc > Sort Lists** in the Menu bar. For more details about custom sort lists, see Chapter 2, Entering, Editing, and Formatting Data.

Reset and Remove

In the Subtotals dialog, use the **Reset** button to undo any changes made on the current tab page. Use the **Remove** button to remove any subtotals that have already been created using the Subtotals tool. Use these features with care, as no confirmation dialogs will be displayed.

Using scenarios

Scenarios are saved, named cell ranges that you can use to answer “what-if” questions about your data. You can create multiple scenarios for the same calculation set, then quickly swap between them to view the outcomes of each. This feature is useful if you need to test the effects of different conditions on your calculations, but do not want to deal with repetitive manual data entry. For example, if you wanted to test different interest rates for an investment, you could create scenarios for each rate, then switch between them to find out which rates work the best for you.

To use the Scenarios tool, select **Tools > Scenarios** from the Menu bar to open the Create Scenario dialog (Figure 12).

Figure 12: Create Scenario dialog

Creating scenarios

To create a new scenario:

- 1) Select the cells that contain the values that will change between scenarios. To select multiple ranges, hold down the *Ctrl* key as you click. You must select at least two cells.
- 2) Choose **Tools > Scenarios** from the Menu bar to open the Create Scenario dialog.
- 3) Enter a name for the new scenario in the *Name of Scenario* field.



Tip

For each scenario you create, use a unique name that clearly identifies and distinguishes it. This practice will save you time and headache if you have to work with a large set of scenarios.

- 4) Optionally, add information to the *Comment* field. The example in Figure 12 shows the default comment.
- 5) Click **OK** to close the dialog. The new scenario is automatically activated upon creation.
- 6) Repeat steps 1 to 5 to create additional scenarios. Select the same cell range that you used for the first scenario to have multiple scenarios for the same calculations.

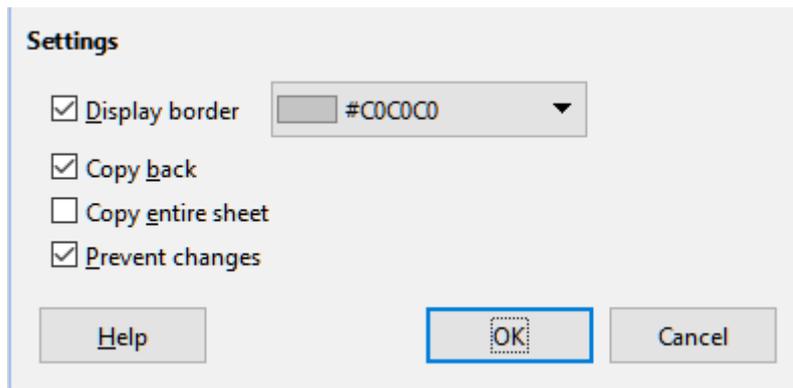


Tip

To keep track of what calculations are dependent on your scenarios, use **Tools > Detective > Trace Dependents** from the Menu bar on your scenario cells. Arrows will point from your scenario cells to the dependent formula cells. For more information about the Detective tool, see Chapter 7, Using Formulas and Functions.

Scenario settings

The *Settings* sections of the Create Scenario dialog contains the following options:



Settings

Display border #C0C0C0

Copy back

Copy entire sheet

Prevent changes

Help OK Cancel

Figure 13: Create Scenario settings

Display border

Places a colored border around the cell range that contains your scenario (Figure 14). To choose the color of the border, use the drop-down field to the right of this option. The border has a title bar displaying the name of the active scenario and a down arrow button that opens a drop-down list of all the scenarios defined for the current cell range. You can swap to a scenario by selecting it from this list.

	A	B
1	Fixed cost estimates	
2		
3		Estimate 2 ▾
4	Salaries	\$6,532
5	Sales & Marketing	\$6,117
6	Insurance	\$3,908
7	Rent	\$5,610
8	Other	\$1,017
9		
10	Total	\$23,184

Figure 14: Scenario cell range with border

Copy back

Copies any changes made to a scenario's cell values back into the active scenario. If you uncheck this option, then the scenario's initial values cannot be overwritten. The actual behavior of the **Copy back** setting depends on the current cell and sheet protections and the **Prevent changes** setting (see Table 4 on page 16).



Caution

When creating a new scenario from the cells of a scenario with **Copy back** enabled, be careful not to overwrite the old scenario. To avoid this situation, create the new scenario with **Copy back** enabled first, then change its values only once it is active.

Copy entire sheet

Creates a copy of the sheet with the new active scenario. The sheet is named after this scenario. Note that changing scenario values in the copy sheet will not affect the active scenario, even if it has **Copy back** enabled.

Prevent changes

Prevents changes to a scenario with **Copy back** enabled when the sheet is protected but the cells are not. Also prevents changes to the scenario's settings while the sheet is protected. This option and its effects are more fully explained in the next section.

Changing scenarios

Scenarios have two aspects that can be altered independently:

- Scenario properties (that is, its settings)
- Scenario cell values

The extent to which these aspects can be changed depends on the active scenario's properties and the current sheet and cell protections. For more detail about sheet and cell protections, see Chapter 2, Entering, Editing, and Formatting Data.

Changing scenario properties

Table 3 summarizes how sheet protection and the **Prevent changes** option affect your ability to change scenario properties.

Table 3: Changing scenario properties

Sheet protection	Prevent changes	Property changes
On	On	No scenario properties can be changed.
On	Off	Display border and Copy back can be changed. Prevent changes and Copy entire sheet cannot be changed.
Off	Any setting	All scenario parameters except for Copy entire sheet can be changed. In this case, the Prevent changes option has no effect.

Changing scenario cell values

Table 4 summarizes the interaction of various settings in making changes to scenario cell values.

Table 4: Changing scenario cell values

Sheet protection	Scenario cell protection	Prevent changes	Copy back	Change allowed
On	Off	On	On	Scenario cell values cannot be changed.
On	Off	Off	On	Scenario cell values can be changed, and the scenario is updated.
On	Off	Any setting	Off	Scenario cell values can be changed, but the scenario is not updated due to the Copy back setting.
On	On	Any setting	Any setting	Scenario cell values cannot be changed.
Off	Any setting	Any setting	Any setting	Scenario cell values can be changed and the scenario is updated or not, depending on the Copy back setting.

Working with scenarios using the Navigator

After you add scenarios to a spreadsheet, you can view a particular scenario by using the Navigator. Open it by selecting **View > Navigator** from the Menu bar, then click on the **Scenarios** icon in the Navigator and select a scenario from the list (Figure 15). Any defined scenarios are listed along with the comments that were entered when each scenario was created. For more information about the Navigator, see Chapter 1, Introducing Calc.

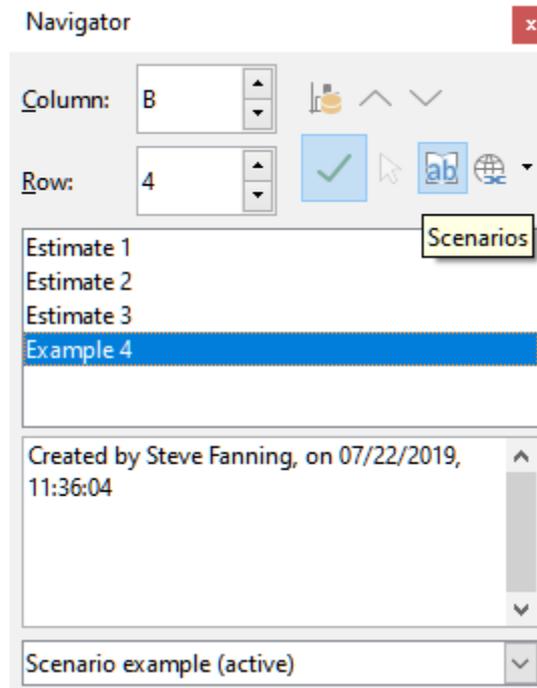


Figure 15: Scenarios in the Navigator

To apply a scenario to the current sheet, double-click the scenario name in the Navigator.

To delete a scenario, right-click the name in the Navigator and choose **Delete**. A confirmation dialog will be displayed.

To edit a scenario, right-click the name in the Navigator and choose **Properties**. Calc displays the Edit Scenario dialog, which is similar to the Create Scenario dialog (Figure 12) except for the **Copy entire sheet** option, which is disabled.

Using the Multiple Operations tool

Like scenarios, the Multiple Operations tool performs what-if analysis on your calculations. Unlike scenarios, which represent individual sets of values for multiple formula variables, this tool uses a whole range of values for just one or two variables. It then uses one or more formulas to create a matching range of solutions. Since each solution corresponds to one or two variable values, both the variable and solution ranges can be easily arrayed in tabular format. As a result, the Multiple Operations tool is well-suited to generating data that easy to read and share or visualize using graphs.



Tip

Exercising good organization can make using this tool relatively painless. For example, we recommend keeping your data together on one sheet and using labels to identify your formulas, variables, and table ranges.

To use this tool, select **Data > Multiple Operations** from the Menu bar to open the Multiple Operations dialog (Figure 16).

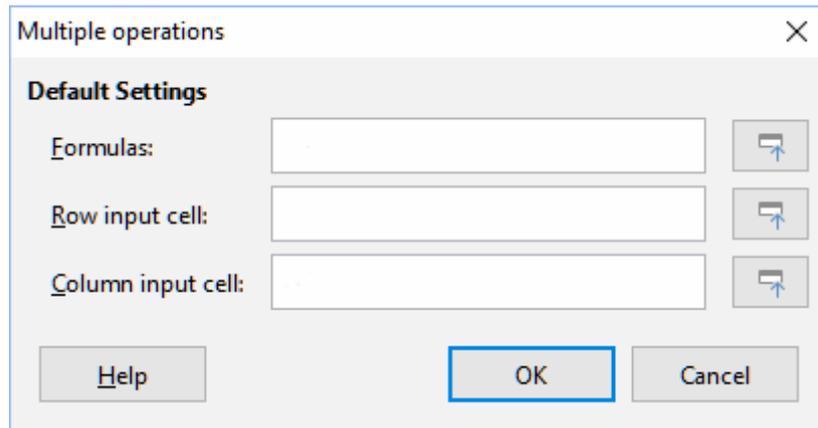


Figure 16: Multiple Operations dialog

Multiple Operations with one formula and one variable

The easiest way to learn how to use the Multiple Operations tool is with one formula and one variable. To learn how to use the tool with multiple formulas or with two variables, see “Calculating with several formulas simultaneously” (page 20) and “Multiple operations with two variables” (page 22), respectively.

To use the Multiple Operations tool with one formula and one variable:

- 1) In the cells of a worksheet, enter a formula and at least one variable that it uses.
- 2) In the same worksheet, enter values into a cell range that occupies a single column or row. These values will be used for one of variables of the formula that you defined in step 1.
- 3) With the mouse, select the range containing both the variable range that you defined in step 2 and the adjacent empty cells that follow it. Depending on how your variable range is arrayed, these empty cells will either be in the column to the right (if the range is in a column) or in the row immediately below (if it is in a row).
- 4) Select **Data > Multiple Operations** to open the Multiple Operations dialog.
- 5) Click on the *Formulas* field and type a cell reference to the formula you defined in step 1 or select the cell with the mouse. Use the associated **Shrink / Expand** button if you need to minimize the dialog while selecting the cell.
- 6) If the range from step 2 is arrayed in a column, then click on the *Column input cell* field and type a cell reference to the variable that you want to use or select the cell with the mouse. If the range is in a row, then use the *Row input cell* field instead.
- 7) Click **OK** to run the tool. The Multiple Operations tool will generate its results in the empty cells that you selected in step 3. Each result value corresponds to the variable value adjacent to it, and together they form the entries of a results table.

An example with one formula and one variable

Using the Multiple Operations tool is best explained by example. Suppose that you produce toys that you sell for \$10 each (cell B1 of a worksheet). Each toy costs \$2 to make (B2), and you have a fixed annual cost of \$10,000 (B3). What is the minimum number of toys that you must sell to break even? Suppose that our initial estimate of quantity sold is 2,000 (B4).

To answer this question:

- 1) Enter the following formula into B5: $=B4 * (B1 - B2) - B3$. This formula represents the equation $Profit = Quantity * (Selling price - Direct costs) - Fixed costs$. With this equation, our initial quantity produces a \$6,000 profit, which is higher than the break-even point.

- 2) In D2:D11, enter a range of alternate quantities from 500 to 5000 in steps of 500.
- 3) Select the range D2:E11 to define the results table. This range includes the alternate quantity values (column D) and the empty results cells (column E).
- 4) Select **Data > Multiple Operations** to open the Multiple Operations dialog.
- 5) Using the *Formulas* field, select the cell B5.
- 6) Using the *Column input cell* field, select the cell B4 to set the quantity as the variable for our calculations. Figure 17 shows the worksheet and Multiple Operations dialog at this point.

	A	B	C	D	E
1	Unit sale price	\$10		Quantity	Profit
2	Unit cost	\$2		500	
3	Fixed annual cost	\$10,000		1000	
4	Quantity sold	2,000		1500	
5	Profit	\$6,000		2000	
6				2500	
7				3000	
8				3500	
9				4000	
10				4500	
11				5000	

Multiple operations ×

Default Settings

Formulas: ↶

Row input cell: ↶

Column input cell: ↶

Figure 17: Inputs to Multiple Operations tool for one formula, one variable

- 7) Click **OK**. The profits for the different quantities are now shown in column E (Figure 18). We can see that the break-even point is between 1000 and 1500 toys sold – namely, 1250.

	A	B	C	D	E
1	Unit sale price	\$10		Quantity	Profit
2	Unit cost	\$2		500	-\$6,000
3	Fixed annual cost	\$10,000		1000	-\$2,000
4	Quantity sold	2,000		1500	\$2,000
5	Profit	\$6,000		2000	\$6,000
6				2500	\$10,000
7				3000	\$14,000
8				3500	\$18,000
9				4000	\$22,000
10				4500	\$26,000
11				5000	\$30,000

Figure 18: Results of Multiple Operations tool for one formula and one variable



Figure 19: XY scatterplot of profit over quantity of toys sold (example of visualization)

Calculating with several formulas simultaneously

Using the Multiple Operations tool with multiple formulas follows nearly the same process as with one formula, but with two important differences:

- 1) For each formula that you add, you must also add a corresponding column or row to the results table to contain the output of that formula.
- 2) How you initially arrange your formulas determines how their results will be displayed in the results table. For example, if you arrange the formulas *A*, *B*, and *C* in a single row in that order, then Calc will generate the results of *A* in the first results table column, the results of *B* in the second column, and the results of *C* in the third.



Note

The Multiple Operations tool only accepts formulas arranged in a single row or column, depending on how your results table is oriented. If the table is column-oriented – that is, the way it is in our sales data example – then your formulas must be arranged in a row. If the table is row-oriented, then your formulas must be in a column.



Caution

Be careful not to add empty cells between formulas, as they will create gaps in the results table and may cause some results not to appear if you don't select enough rows or columns for the table.

An example with two formulas and one variable

Using our sales data example, suppose that we want to calculate the annual profit per item sold in addition to the annual overall profit. To calculate the results:

- 1) In the sheet from the previous example, delete the results in column E.
- 2) Enter the following formula in C5: `=B5/B4`. You are now calculating the annual profit per item sold.
- 3) Select the range D2:F11 for the results table. Column F will contain the results of the annual profit per item formula in C5.

- 4) Select **Data > Multiple Operations** from the Menu bar to open the Multiple Operations dialog.
- 5) Using the *Formulas* field, select the range B5:C5.
- 6) Using the *Column input cell* field, select the cell B4. Figure 20 shows the worksheet and the dialog at this point.

	A	B	C	D	E	F
1	Unit sale price	\$10		Quantity	Profit	Per item
2	Unit cost	\$2		500		
3	Fixed annual cost	\$10,000		1000		
4	Quantity sold	2,000	Annual profit per item	1500		
5	Profit	\$6,000	\$3	2000		
6				2500		
7				3000		
8				3500		
9				4000		
10				4500		
11				5000		

Multiple operations ×

Default Settings

Formulas: ↕

Row input cell: ↕

Column input cell: ↕

Figure 20: Inputs to Multiple Operations tool for one variable and two formulas

- 7) Click **OK**. Now the profits are listed in column E and the annual profit per item in column F.

F11 fx Σ = =MULTIPLE.OPERATIONS(C\$5,\$B\$4,\$D11)

	A	B	C	D	E	F
1	Unit sale price	\$10		Quantity	Profit	Per item
2	Unit cost	\$2		500	-\$6,000	-\$12.00
3	Fixed annual cost	\$10,000		1000	-\$2,000	-\$2.00
4	Quantity sold	2,000	Annual profit per item	1500	\$2,000	\$1.33
5	Profit	\$6,000	\$3	2000	\$6,000	\$3.00
6				2500	\$10,000	\$4.00
7				3000	\$14,000	\$4.67
8				3500	\$18,000	\$5.14
9				4000	\$22,000	\$5.50
10				4500	\$26,000	\$5.78
11				5000	\$30,000	\$6.00

Figure 21: Results of Multiple Operations tool for one variable and two formulas

Multiple operations with two variables

When you use the Multiple Operations tool with two variables, it creates a two-dimensional results table. Each variable defines one of the table's dimensions such that the alternate values for both variables serve as the table's row and column headings, respectively. Each cell in the table corresponds to a distinct pair of row and column heading values. In turn, the results in each cell are created from these values for both variables.

Since you are using two variables, you must use both the *Column input cell* and *Row input cell* dialog fields to define them. The order is important; the *Column input cell* field corresponds to the row heading values, while the *Row input cell* field corresponds to the column heading values.



Tip

A good guideline to remember is that since column headings are in a row at the top of the table, they correspond to the *Row input cell* field. Likewise, row headings are in a column, so they correspond to the *Column input cell* field.



Note

If you use two variables, the Multiple Operations tool will not work with multiple formulas. It will allow you to enter the extra formulas, but will not generate the expected results for any formula beyond the first.

Calculating with two variables

Using our sales example, suppose that in addition to varying the quantity of toys sold, you also want to vary the unit sale price as well. To calculate the results:

- 1) Expand the sales data table by entering **\$8**, **\$10**, **\$15** and **\$20** in the range E1:H1.
- 2) Select the range D1:H11 for the results table.
- 3) Select **Data > Multiple Operations** from the Menu bar to open the Multiple Operations dialog.
- 4) Using the *Formulas* field, select cell B5.
- 5) Using the *Row input cell* field, select cell B1. The column headings – **\$8**, **\$10**, **\$15** and **\$20** – are now linked to the unit sale price variable defined in cell B1.
- 6) Using the *Column input cell* field, select cell B4. The row headings – **500**, **1000**, ... , **5000** – are now linked to the quantity sold variable defined in cell B4. Figure 22 shows the worksheet and dialog at this point.
- 7) Click **OK**. The profits for the different sale prices and quantities are now shown in the range E2:H11 (Figure 23).

	A	B	C	D	E	F	G	H
1	Unit sale price	\$10			\$8	\$10	\$15	\$20
2	Unit cost	\$2		500				
3	Fixed annual cost	\$10,000		1000				
4	Quantity sold	2,000	Annual profit per item	1500				
5	Profit	\$6,000	\$3	2000				
6				2500				
7				3000				
8				3500				
9				4000				
10				4500				
11				5000				

Multiple operations

Default Settings

Formulas:

Row input cell:

Column input cell:

Figure 22: Inputs to Multiple Operations tool for two variables

H11 fx Σ =MULTIPLE.OPERATIONS(\$B\$5,\$B\$4,\$D11,\$B\$1,H\$1)

	A	B	C	D	E	F	G	H
1	Unit sale price	\$10			\$8	\$10	\$15	\$20
2	Unit cost	\$2		500	-\$7,000	-\$6,000	-\$3,500	-\$1,000
3	Fixed annual cost	\$10,000		1000	-\$4,000	-\$2,000	\$3,000	\$8,000
4	Quantity sold	2,000	Annual profit per item	1500	-\$1,000	\$2,000	\$9,500	\$17,000
5	Profit	\$6,000	\$3	2000	\$2,000	\$6,000	\$16,000	\$26,000
6				2500	\$5,000	\$10,000	\$22,500	\$35,000
7				3000	\$8,000	\$14,000	\$29,000	\$44,000
8				3500	\$11,000	\$18,000	\$35,500	\$53,000
9				4000	\$14,000	\$22,000	\$42,000	\$62,000
10				4500	\$17,000	\$26,000	\$48,500	\$71,000
11				5000	\$20,000	\$30,000	\$55,000	\$80,000

Figure 23: Results of Multiple Operations tool for two variables

Using Goal Seek

In addition to scenarios and the Multiple Operations tool, Calc has a third “what-if” analysis tool: Goal Seek. Usually, you use a formula to calculate a result from existing values. In contrast, with Goal Seek, you work backwards from a result to discover what values produce it. This feature is useful if you already know the outcome you want, but need to answer questions such as how to reach it or how it could be changed if you altered conditions.

To use this tool, select **Tools > Goal Seek** from the Menu bar to open the Goal Seek dialog (Figure 24).



Note

Only one argument can be altered at a time in a single goal seek. If you need to test multiple arguments, then you must run a separate goal seek on each one.

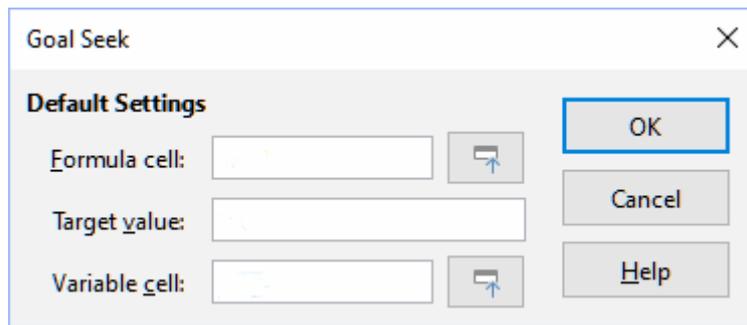


Figure 24: Goal Seek dialog

Goal Seek example

To illustrate how to use Goal Seek, suppose that we want to calculate the annual interest return for an account. To calculate annual interest (**I**), we must create a table with values for the capital (**C**), the interest period length in years (**n**), and the interest rate (**i**). The formula is $I = C * n * i$.

Suppose that the interest rate $i = 7.5\%$ (cell B3 of a worksheet) and the period length $n = 1$ (B2) remain constant. We want to know how much investment capital **C** is needed to achieve a return of $I = \$15,000$. Assume that our initial capital estimate is $C = \$100,000$ (B1).

To calculate the return:

- 1) Enter the return formula ($=B1 * B2 * B3$) into B4 and select the cell with the mouse.
- 2) Select **Tools > Goal Seek** from the Menu bar to open the Goal Seek dialog.
- 3) B4 should already be entered in the *Formula cell* field. However, if you want to select a different cell, use the associated **Shrink / Expand** button to minimize the dialog while you select the required cell.
- 4) Click on the *Variable cell* field, then type a reference to cell B1 or select it with the mouse to make the capital the variable in the current Goal Seek.
- 5) Enter the desired formula result in the *Target value* field. In this example, the value is 15000 . Figure 25 shows the cells and dialog fields at this point.

	A	B	C	D
1	Capital	\$100,000		
2	Years	1		
3	Interest rate	7.5%		
4	Annual return	\$7,500		
5				
6				
7				
8				
9				
10				
11				
12				
13				

Figure 25: Example setup for goal seek

- 6) Click **OK**. A dialog appears informing you that the goal seek was successful (Figure 26).

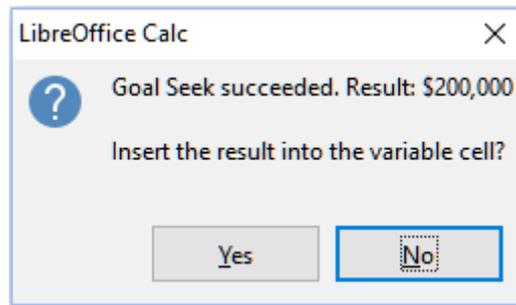


Figure 26: Goal seek result dialog

- 7) Click **Yes** to enter the goal value into the variable cell. The result is shown in Figure 27, indicating that a capital requirement of \$200,000 is needed to achieve a \$15,000 return.

	A	B
1	Capital	\$200,000
2	Years	1
3	Interest rate	7.5%
4	Annual return	\$15,000

Figure 27: Result of goal seek in worksheet



Note

Not every goal seek problem succeeds in returning a good result. It depends on the formula used, goal value, and initial value. The goal seek algorithm iterates internally several times converging to the goal.

Using the Solver

The Solver amounts to a more elaborate form of goal seek, which allows you to solve linear programming or optimization problems. A linear programming problem is concerned with minimizing or maximizing a linear function subject to a set of linear constraints. Such problems arise in many scientific, engineering, business, and other disciplines. A full discussion of linear programming is beyond the scope of this guide and the interested reader is referred to relevant Operations Research websites or text books for more information.

Currently, in version 6.2, LibreOffice offers the following linear solvers:

- The open source *lp_solve* solver (prior to version 4.3)
- The open source *CoinMP (CLP)* linear solver (since 4.3; default)

A benchmark study involving both linear solvers suggests that *CLP* is faster and more efficient than *lp_solve*. In general, however, either engine will produce an accurate solution relatively quickly.



Note

For more information on Calc's linear solvers, see the following links:

lp_solve – <http://lpsolve.sourceforge.net/5.5/>

CLP and the COIN-OR Foundation – <https://www.coin-or.org/>

Benchmark study including both *lp_solve* and *CLP*: <https://prod-ng.sandia.gov/techlib-noauth/access-control.cgi/2013/138847.pdf>



Caution

LibreOffice 6.2 offers a third solver engine: the *LibreOffice Swarm Non-Linear Solver (experimental)*. As it is an experimental tool, it may not be supported in future versions of Calc, and we recommend that you do not use this option unless you are familiar with nonlinear programming concepts.

In order to use the Solver to solve a linear programming problem, you must formulate the problem as follows:

- *Decision variables* – a set of n non-negative variables x_1, \dots, x_n . Decision variables may be real numbers, but generally tend to be integers in many real world problems.
- *Constraints* – a set of linear equalities or inequalities involving the n decision variables.
- *Objective function* – a linear expression involving the n decision variables.

The goal is usually to find values of the decision variables that satisfy the constraints and maximize or minimize the result of the objective function.

Solver dialog

After setting up the data for the problem in your Calc spreadsheet, select **Tools > Solver** from the Menu bar to open the Solver dialog (Figure 28).

Figure 28: Solver dialog



Note

Depending on the configuration of your computer, a message may be displayed the first time that you select **Tools > Solver** after starting Calc. The message will contain the text *LibreOffice requires a 64-bit Java runtime environment (JRE) to perform this task. Please install a JRE and restart LibreOffice.* Click the **OK** button to dismiss this dialog – the message does not affect operation of the two linear solvers.

Target cell

Type a cell reference to the objective function or select it with the mouse.

Optimize result to

Select **Maximum** to find the maximum result for the objective function, **Minimum** to find the minimum result, or **Value of** to set it to a specific value. If you select **Value of**, enter the required value or a reference to the cell containing that value.

By changing cells

Enter the locations of any cells that define your decision variables.

Limiting Conditions

Enter your constraints in the fields in this area:

- *Cell reference* – enter a cell reference to a decision variable.
- *Operator* – defines a parameter for a constraint. Available options include **<=** (less than or equal to), **=** (equal to), **=>** (greater than or equal to), **Integer** (values without decimals), and **Binary** (only 0 or 1).
- *Value* – enter a value or a cell reference to a constraint formula.
- *Remove* button – deletes the currently-defined constraint.



Tip

Remember that for any of these options, you can minimize the Solver dialog using the associated **Shrink / Expand** buttons if you need to select cells with the mouse.

Once you have finished setting up the Solver, click the **Solve** button to begin the process of adjusting values and calculating results. Depending on the complexity of the task, this may take some time. On successful completion, Calc presents a Solving Result dialog (Figure 29). This dialog includes buttons to save (**Keep Result**) or discard (**Restore Previous**) your results.

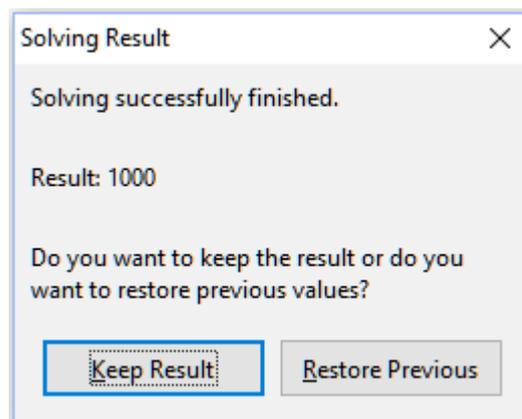


Figure 29: Solving Result dialog

The Solver dialog also has an Options button, which opens the Options dialog shown in Figure 30.

Solver options

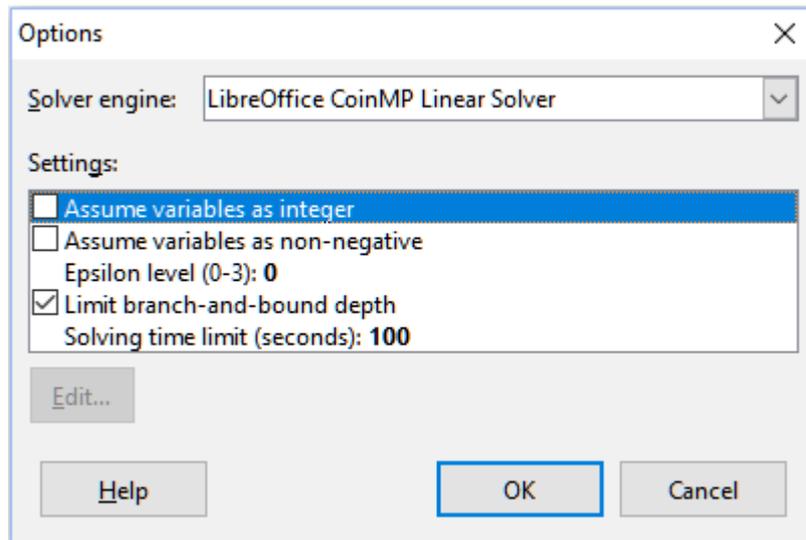


Figure 30: Solver Options dialog

Solver engine

With the standard LibreOffice 6.2 download, the *Solver* engine drop-down list provides three options:

- **LibreOffice CoinMP Linear Solver** (default)
- **LibreOffice Linear Solver** (*lp_solve*)
- **LibreOffice Swarm Non-Linear Solver (experimental)**

Settings

This area enables the user to adjust the detailed processing parameters of the selected solver engine. For the two linear solvers, these parameters include:

- *Assume variables are integers*
- *Assume variables are non-negative*
- *Epsilon level (0-3)* – This parameter determines the upper bound on relative error due to floating-point rounding. Value 0 corresponds to very tight epsilon values (default); value 3 corresponds to very loose epsilon values.
- *Limit branch-and-bound depth*
- *Solving time limit (secs)* – 100 seconds is the default.

Solver example

Suppose that you have \$10,000 that you want to invest in two mutual funds for one year. Fund X is a low risk fund with an 8% interest rate and Fund Y is a higher risk fund with a 12% interest rate. How much money should be invested in each fund to earn a total interest of \$1,000?

To find the answer using the Solver:

- 1) Enter the following labels and data into a worksheet:
 - Row labels: *Fund X*, *Fund Y*, and *Total* in cells A2, A3, and A4.
 - Column labels: *Interest earned*, *Amount invested*, *Interest rate*, and *Time period* in cells B1 thru E1.
 - Interest rates: 8% and 12% in cells D2 and D3.
 - Time period: 1 in cells E2 and E3.

- Total amount invested: **\$10000** in cell C4.
 - Enter an arbitrary value (**\$0** or leave blank) in cell C2 as the amount invested in Fund X.
- 2) Enter the following formulas:
- In cell C3, enter the formula **=C4-C2** (total amount – amount invested in Fund X) as the amount invested in Fund Y.
 - In cells B2 and B3, enter the formulas **=C2*D2*E2** (B2) and **=C3*D3*E3** (B3).
 - In cell B4, enter the formula **=B2+B3** as the total interest earned. Figure 31 shows the worksheet at this point.

	A	B	C	D	E
1		Interest earned	Amount invested	Interest rate	Time period
2	Fund X	\$0	\$0	8%	1
3	Fund Y	\$1,200	\$10,000	12%	1
4	Total	\$1,200	\$10,000		
5					

Figure 31: Solver example setup

- 3) Select **Tools > Solver** from the Menu bar to open the Solver dialog (Figure 28).
- 4) Using the *Target cell* field, select the cell that contains the target value. In this example, it is B4, which contains the total interest value.
- 5) Select **Value of** and enter **1000** in the field next to it. In this example, the target cell value is 1000 because your target is a total interest earned of \$1,000.
- 6) Using the *By changing cells* field, select cell C2 in the sheet. In this example, you need to find the amount invested in Fund X (cell C2).
- 7) Enter the following limiting conditions for the variables by using the *Cell reference*, *Operator*, and *Value* fields:
 - **C2 <= C4** – the amount invested in Fund X cannot exceed the total amount available.
 - **C2 => 0** – the amount invested in Fund X cannot be negative.
 - **C2 is an Integer** – specified for convenience.
- 8) Click **OK**. The result is shown in Figure 32.

	A	B	C	D	E
1		Interest earned	Amount invested	Interest rate	Time period
2	Fund X	\$400	\$5,000	8%	1
3	Fund Y	\$600	\$5,000	12%	1
4	Total	\$1,000	\$10,000		
5					

Figure 32: Solver example result

Using Statistics tools

Under **Data > Statistics** in the Menu bar, Calc features several tools for quick and easy statistical analysis of your data. These tools include:

- Sampling
- Descriptive Statistics
- Analysis of Variance (ANOVA)
- Correlation

- Covariance
- Exponential Smoothing
- Moving Average
- Regression
- Paired t-test
- F-test
- z-test
- Test of Independence (Chi-square)

Sampling tool

The Sampling tool creates a target table with data sampled from a source table. The sampling tool can pick samples randomly or on a periodic basis. Sampling is done row-wise, with whole rows of the source table copied into rows of the target table. To use this tool, select **Data > Statistics > Sampling** from the Menu bar to access the Sampling dialog (Figure 33).

Figure 33: Sampling dialog

Input range

Specifies the cell range containing the source table.

Results to

Specifies the top left cell of the target table. When you run the tool, it will fill out the rest of the target table starting from this cell.

Random

Enables random sampling mode when selected.

Sample size

Defines the number of lines to be sampled from the source table during random sampling. Only available if the **Random** option is selected.

Periodic

Enables periodic sampling mode when selected.

Period

Defines the number of lines to skip between samples during periodic sampling. For example, a value of 2 will cause alternate lines to be sampled from the source table, starting with the second line. Only available if the **Periodic** option is selected.



Tip

Use the **Shrink / Expand** buttons next to the *Input range* and *Results to* fields if you need to shrink the dialog while selecting cells with the mouse.

Figure 34 shows the source table (below the *Source Data* heading) and the corresponding target table (under the *Target Data* heading), sampled using the settings shown in Figure 33.

	A	B	C
1	Source Data		
2	11	21	31
3	12	22	32
4	13	23	33
5	14	24	34
6	15	25	35
7	16	26	36
8	17	27	37
9	18	28	38
10	19	29	39
11			
12	Target Data		
13	12	22	32
14	14	24	34
15	16	26	36
16	18	28	38

Figure 34: Example data for the Sampling tool

Descriptive Statistics tool

Given a set of data, the Descriptive Statistics tool creates a tabular report of a data set's primary statistical properties, such as information about its central tendency and variability. Select **Data > Statistics > Descriptive Statistics** in the Menu bar to access the Descriptive Statistics dialog (Figure 35).

Descriptive Statistics

Data

Input range:

Results to:

Grouped by

Columns Rows

Help OK Cancel

Figure 35: Descriptive Statistics dialog

Input range

Specifies the cell range containing the source data.

Results to

Specifies the top left cell of the results area. When you run the tool, it will fill out the rest of the report table starting at this cell.

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows.



Tip

Use the **Shrink / Expand** buttons next to the *Input range* and *Results to* fields if you need to shrink the dialog while selecting cells with the mouse.

Figure 36 shows a small data set comprising student examination results in three subjects.

	A	B	C
1	Maths	Physics	Biology
2	47	67	33
3	36	68	42
4	40	65	44
5	39	64	60
6		38	43
7	47	84	62
8	29	80	51
9	27	49	40
10	57	49	12
11	56	33	60
12	57		
13	26		

Figure 36: Input data for descriptive statistics analysis

Figure 37 shows the statistics report generated for this input data using the settings shown in Figure 35.

E	F	G	H
	Column 1	Column 2	Column 3
Mean	41.909091	59.7	44.7
Standard Error	3.561038	5.358379	4.768065
Mode	47	49	60
Median	40	64.5	43.5
First Quartile	32.5	49	40.5
Third Quartile	51.5	67.75	57.75
Variance	139.490909	287.122222	227.344444
Standard Deviation	11.810627	16.944681	15.077946
Kurtosis	-1.462168	-0.941599	1.418053
Skewness	0.015241	-0.222643	-0.976680
Range	31	51	50
Minimum	26	33	12
Maximum	57	84	62
Sum	461	597	447
Count	11	10	10

Figure 37: Results from Descriptive Statistics tool



Tip

For more information on descriptive statistics, refer to the corresponding Wikipedia article at https://en.wikipedia.org/wiki/Descriptive_statistics.

Analysis of Variance (ANOVA) tool

The Analysis of Variance (ANOVA) tool compares the means of two or more groups in a sample. Select **Data > Statistics > Analysis of Variance (ANOVA)** in the Menu bar to access the Analysis of Variance (ANOVA) dialog (Figure 38).

The screenshot shows the 'Analysis of Variance (ANOVA)' dialog box. It contains the following fields and options:

- Data**
 - Input range: SSheet1.\$A\$2:\$C\$13
 - Results to: SE\$1
- Type**
 - Single factor
 - Two factor
- Grouped by**
 - Columns
 - Rows
- Parameters**
 - Alpha: 0.05
 - Rows per sample: 1

Buttons: Help, OK, Cancel

Figure 38: Analysis of Variance (ANOVA) dialog

Input range

Specifies the cell range containing the source data.

Results to

Specifies the top left cell of the results area. When you run the tool, it will fill out the analysis of variance report table starting at this cell.

Single / Two factor

Determines whether the analysis is for single factor or two factor ANOVA.

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows. Only available if **Single factor** is selected.

Alpha

In this field, enter a significance level in the range 0.01 to 0.99. The default is 0.05.

Rows per sample

Defines how many rows a sample has. This option is always set to 1 in this version of Calc.

To illustrate how to use this tool, we use the input data set from Figure 36. Figure 39 shows the analysis of variance results generated for this data using the settings shown in Figure 38.

E	F	G	H	I	J	K
ANOVA - Single Factor						
Alpha	0.05					
Groups	Count	Sum	Mean	Variance		
Column 1	11	461	41.909091	139.490909		
Column 2	10	597	59.7	287.122222		
Column 3	10	447	44.7	227.344444		
Source of Variation	SS	df	MS	F	P-value	F critical
Between Groups	1876.568328	2	938.284164	4.360412	0.022461	3.340386
Within Groups	6025.109091	28	215.182468			
Total	7901.677419	30				

Figure 39: Results from Analysis of Variance (ANOVA) tool



Tip

For more information on analysis of variance, refer to the corresponding Wikipedia article at https://en.wikipedia.org/wiki/Analysis_of_variance.

Correlation tool

The Correlation tool calculates the correlation of two sets of numeric data and generates the resulting correlation coefficient. This coefficient is a value between -1 and +1 that indicates how strongly two variables are related to each other. A correlation coefficient of +1 indicates a perfect positive correlation (the data sets match) and a coefficient of -1 indicates a perfect negative correlation (the data sets are inverse to each other). Select **Data > Statistics > Correlation** in the Menu bar to access the Correlation dialog (Figure 40).

Figure 40: Correlation dialog

Input range

Specifies the cell range containing the source data.

Results to

Specifies the top left cell of the results area. When you run the tool, it will fill out the correlation coefficient table starting at this cell.

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows.

To illustrate how to use this tool, we again use the data set from Figure 36. Figure 41 shows the six correlation coefficients generated for this input data using the settings shown in Figure 40.

E	F	G	H
Correlations	Column 1	Column 2	Column 3
Column 1	1		
Column 2	-0.402925	1	
Column 3	-0.210764	0.230971	1

Figure 41: Correlation results



Tip

For more information on statistical correlation, refer to the corresponding Wikipedia article at https://en.wikipedia.org/wiki/Correlation_and_dependence.

Covariance tool

The Covariance tool measures how much two sets of numeric data vary together. Select **Data > Statistics > Covariance** in the Menu bar to access the Covariance dialog (Figure 42).

Figure 42: Covariance dialog

Input range

Specifies the cell range containing the source data.

Results to

Specifies the top left cell of the results area. When you run the tool, it will fill out the covariance table starting at this cell.

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows.

To illustrate how to use this tool, we again use the data set from Figure 36. Figure 43 shows the six covariance values generated for this input data using the settings shown in Figure 42.

E	F	G	H
Covariances	Column 1	Column 2	Column 3
Column 1	126.809917		
Column 2	-61.444444	258.41	
Column 3	-32	53.11	204.61

Figure 43: Covariance results



Tip

For more information on statistical covariance, refer to the corresponding Wikipedia article at <https://en.wikipedia.org/wiki/Covariance>.

Exponential Smoothing tool

The Exponential Smoothing tool filters a data set to produce smoothed results. It is used in domains such as stock market analysis and in sampled measurements. Select **Data > Statistics > Exponential Smoothing** in the Menu bar to access the Exponential Smoothing dialog (Figure 44).

The screenshot shows the 'Exponential Smoothing' dialog box. It has a title bar with a close button (X). The dialog is organized into several sections: 'Data' with 'Input range' (SSheet1.SAS1:SB\$13) and 'Results to' (SD\$1); 'Grouped by' with radio buttons for 'Columns' (selected) and 'Rows'; and 'Parameters' with a 'Smoothing factor' spinner set to 0.50. At the bottom are 'Help', 'OK', and 'Cancel' buttons.

Figure 44: Exponential Smoothing dialog

Input range

Specifies the cell range containing the source data.

Results to

Specifies the top left cell of the results area. When you run the tool, it will generate smoothed results starting at this cell.

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows.

Smoothing factor

Enter a value here between 0 and 1 (default is 0.2) that represents the damping factor alpha in the smoothing equation.

To illustrate how to use this tool, we use the data set shown in Figure 45. The table has two time series representing impulse functions at times $t=0$ and $t=2$.

	A	B
1	1	0
2	0	0
3	0	1
4	0	0
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	0
13	0	0

Figure 45: Input data set for exponential smoothing example

Figure 46 shows the smoothed results for this input data using the settings shown in Figure 44.

D	E
Alpha	
0.5	
Column 1	Column 2
1	0
1	0
0.5	0
0.25	0.5
0.125	0.25
0.0625	0.125
0.03125	0.0625
0.015625	0.03125
0.0078125	0.015625
0.00390625	0.0078125
0.001953125	0.00390625
0.0009765625	0.001953125
0.00048828125	0.0009765625
0.000244140625	0.00048828125

Figure 46: Results from Exponential Smoothing tool



Tip

For more information on exponential smoothing, refer to the corresponding Wikipedia article at https://en.wikipedia.org/wiki/Exponential_smoothing.

Moving Average tool

The Moving Average tool calculates the moving average of a time series data set. Select **Data > Statistics > Moving Average** in the Menu bar to access the Moving Average dialog (Figure 47).

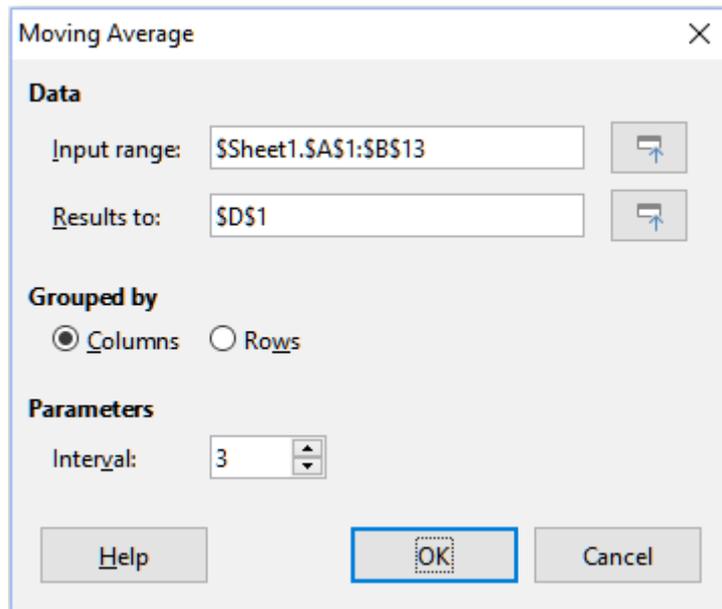


Figure 47: Moving Average dialog

Input range

Specifies the cell range containing the source data.

Results to

Specifies the top left cell of the results area. When you run the tool, it will generate the moving average results starting at this cell.

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows.

Interval

Specifies the number of samples used in the moving average calculation.

To illustrate how to use this tool, we again use the data set from Figure 45. Figure 48 shows the moving averages calculated for this input data using the settings shown in Figure 47.

D	E
Column 1	Column 2
#N/A	#N/A
0.333333	0.333333
0	0.333333
0	0.333333
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
#N/A	#N/A

Figure 48: Calculated moving averages



Tip

For more information on the moving average, refer to the corresponding Wikipedia article at https://en.wikipedia.org/wiki/Moving_average.

Regression tool

The Regression tool analyzes the relationship in a data set between one or more independent variables and a dependent variable. Select **Data > Statistics > Regression** in the Menu bar to access the Regression dialog (Figure 49).

Figure 49: Regression dialog

Independent variable(s) (X) range

Specifies the cell range containing the independent variables in the source data.

Dependent variable (Y) range

Specifies the cell range containing the dependent variable in the source data.

Both X and Y ranges have labels

Specifies whether the ranges above include data labels.

Results to

Specifies the top left cell of the results area. When you run the tool, it will generate the regression analysis table starting at this cell.

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows.

Linear Regression

Select this option to use linear regression. Linear regression finds a straight line in the form of $y = (a * x) + b$ that best fits the data, where a is the slope and b is the intercept.

Logarithmic Regression

Select this option to use logarithmic regression. Logarithmic regression finds a logarithmic curve in the form of $y = (a * \ln(x)) + b$ that best fits the data, where a is the slope, b is the intercept and $\ln(x)$ is the natural logarithm of x .

Power Regression

Select this option to use power regression. Power regression finds a power curve in the form of $y = (a * x) ^ b$ that best fits the data, where a is the coefficient and b is the exponent.

Confidence level

Specifies the confidence level, which is a value between 0 and 1. The default value is 0.95.

Calculate residuals

Specifies whether to calculate residuals.

Force intercept to be zero

Specifies whether to force the regression intercept to be 0.



Tip

Use the **Shrink / Expand** buttons next to the *Independent variable(s) (X) range*, *Dependent variable (Y) range*, and *Results to* fields if you need to shrink the dialog while selecting cells with the mouse.

To illustrate how to use this tool, we use the data set shown in Figure 50. This table contains measurements taken at 1 second intervals.

	A	B
1	Time	Measurement
2	1	2.7
3	2	4.0
4	3	4.4
5	4	7.1
6	5	4.9
7	6	3.6
8	7	4.0
9	8	0.6
10	9	1.0
11	10	4.3

Figure 50: Input data set for regression analysis

Figure 51 shows the regression outputs calculated for this input data using the settings shown in Figure 49.

D	E	F	G	H	I	J
Regression						
Regression Model	Linear					
LINEST raw output						
	-0.219394	4.866667				
	0.205798	1.276946				
	0.124390	1.869257				
	1.136489	8				
	3.971030	27.952970				
Regression Statistics						
R²	0.124390					
Standard Error	1.869257					
Count of x-variables	1					
Observations	10					
Adjusted R ²	0.014939					
Analysis of Variance (ANOVA)						
	df	SS	MS	F	Significance F	
Regression	1	3.971030	3.971030	1.136489	0.317509	
Residual	8	27.952970	3.494121			
Total	9	31.924				
Confidence level						
	0.95					
	Coefficients	Standard Error	t-Statistic	P-value	Lower 95%	Upper 95%
Intercept	4.866667	1.276946	3.811178	0.005154	1.922025	7.811308
Time	-0.219394	0.205798	-1.066062	0.317509	-0.693966	0.255178
Time	Predicted y	Measurement	Residual			
1	4.64727	2.7	-1.947273			
2	4.42788	4	-0.427879			
3	4.20848	4.4	0.191515			
4	3.98909	7.1	3.110909			
5	3.76970	4.9	1.130303			
6	3.55030	3.6	0.049697			
7	3.33091	4	0.669091			
8	3.11152	0.6	-2.511515			
9	2.89212	1	-1.892121			
10	2.67273	4.3	1.627273			

Figure 51: Linear regression outputs



Tip

For more information on regression analysis, refer to the corresponding Wikipedia article at https://en.wikipedia.org/wiki/Regression_analysis.

Paired T-test tool

The Paired T-test tool compares the population means of two related sample sets and determines the difference between the two. Select **Data > Statistics > Paired t-test** in the Menu bar to access the Paired t-test dialog (Figure 52).

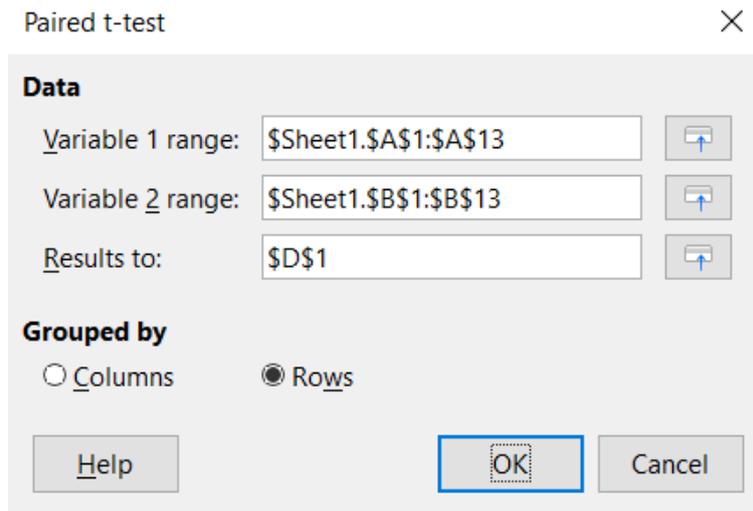


Figure 52: Paired t-test dialog

Variable 1 range

Specifies the cell range containing the first set of input data.

Variable 2 range

Specifies the cell range containing the second set of input data.

Results to

Specifies the top left cell of the results area. When you run the tool, it will generate the paired t-test table starting at this cell.

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows.

To provide an example of using this tool, we make use of the input data set shown in Figure 53. The data sets in columns A and B represent two sets of paired values referred to as *Variable 1* and *Variable 2*.

	A	B
1	28	19
2	26	13
3	31	12
4	23	5
5	20	34
6	27	31
7	28	31
8	14	12
9	4	24
10	0	23
11	2	19
12	8	10
13	9	33

Figure 53: Input data for paired t-test example

Figure 54 shows the paired t-test results calculated for this input data using the settings shown in Figure 52.

	D	E	F
Paired t-test			
Alpha		0.05	
Hypothesised Mean Difference		0	
		Variable 1	Variable 2
Mean		16.923077	20.461538
Variance		125.076923	94.435897
Observations		13	13
Pearson Correlation		-0.061754	
Observed Mean Difference		-3.538462	
Variance of the Differences		232.935897	
df		12	
t Stat		-0.835926	
P (T<=t) one-tail		0.209765	
t Critical one-tail		1.782288	
P (T<=t) two-tail		0.419530	
t Critical two-tail		2.178813	

Figure 54: Results from Paired T-test tool



Tip

For more information on paired t-tests, refer to the corresponding Wikipedia article at https://en.wikipedia.org/wiki/Student's_t-test.

F-test tool

The F-test tool calculates the F-Test of two data samples. The tool is used to test the hypothesis that the variance of two populations are equal. Click **Data > Statistics > F-test** in the Menu bar to access the F-test dialog shown in Figure 55 and define the required inputs to the tool.

Figure 55: F-test dialog

Variable 1 range

Specifies the cell range containing the first set of input data.

Variable 2 range

Specifies the cell range containing the second set of input data.

Results to

Specifies the top left cell of the results area. When you run the tool, it will generate the F-test table starting at this cell.

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows.



Tip

Use the **Shrink / Expand** buttons next to the *Variable 1 range*, *Variable 2 range*, and *Results to* fields if you need to shrink the dialog while selecting cells with the mouse.

To illustrate how to use this tool, we again use the data set from Figure 53. In this case, the data in columns A and B represent two independent sample sets, referred to as *Variable 1* and *Variable 2*. Figure 56 shows the F-test results calculated for this input data using the settings shown in Figure 55.

	D	E	F
F-test			
Alpha		0.05	
		Variable 1	Variable 2
Mean		16.923077	20.461538
Variance		125.076923	94.435897
Observations		13	13
df		12	12
F		1.324464	
P (F<=f) right-tail		0.317061	
F Critical right-tail		2.686637	
P (F<=f) left-tail		0.682939	
F Critical left-tail		0.372213	
P two-tail		0.634123	
F Critical two-tail		0.305131	3.277277

Figure 56: Results from F-test tool



Tip

For more information on F-tests, refer to the corresponding Wikipedia article at <https://en.wikipedia.org/wiki/F-test>.

Z-test tool

The Z-test tool calculates the Z-test of two data samples. The tool performs a two sample Z-test to test the null hypothesis that there is no difference between the means of the two data sets. Click **Data > Statistics > Z-test** in the Menu bar to access the z-test dialog shown in Figure 57 and define the required inputs to the tool.

Variable 1 range

Specifies the cell range containing the first set of input data.

Variable 2 range

Specifies the cell range containing the second set of input data.

Results to

Specifies the top left cell of the results area. When you run the tool, it will generate the Z-test table starting at this cell.

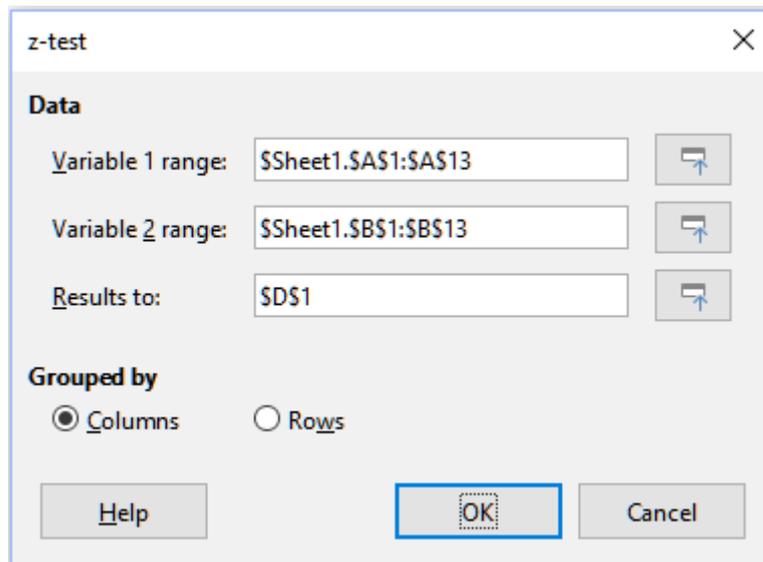


Figure 57: z-test dialog

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows.



Tip

Use the **Shrink / Expand** buttons next to the *Variable 1 range*, *Variable 2 range*, and *Results to* fields if you need to shrink the dialog while selecting cells with the mouse.

To provide an example of using this tool, we again make use of the input data set shown in Figure 53. In this case the data in columns A and B represent two data sets, referred to as *Variable 1* and *Variable 2*. Figure 58 shows the Z-test results calculated for this input data using the settings shown in Figure 57.

	D	E	F
z-test			
Alpha		0.05	
Hypothesised Mean Difference		0	
		Variable 1	Variable 2
Known Variance		0.000000	0.000000
Mean		16.923077	20.461538
Observations		13	13
Observed Mean Difference		-3.538462	
z		#DIV/0!	
P (Z<=z) one-tail		#DIV/0!	
z Critical one-tail		1.644854	
P (Z<=z) two-tail		#DIV/0!	
z Critical two-tail		1.959964	

Figure 58: Results from Z-test tool



Tip

For more information on Z-tests, refer to the corresponding Wikipedia article at <https://en.wikipedia.org/wiki/Z-test>.

Test of Independence (Chi-Square) tool

The Test of Independence (Chi-Square) tool calculates the chi-square test of a data sample, which determines how well a set of measured values fit a corresponding set of expected values. Select **Data > Statistics > Chi-square Test** in the Menu bar to access the Test of Independence (Chi-Square) dialog (Figure 59).

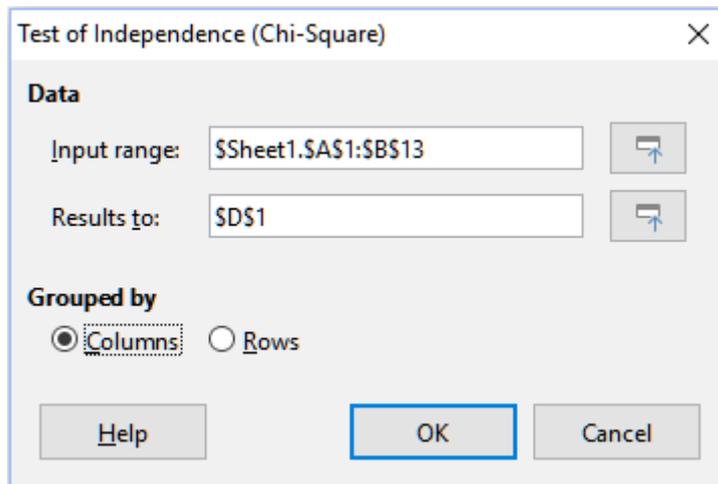


Figure 59: Test of Independence (Chi-Square) dialog

Input range

Specifies the cell range containing the source data.

Results to

Specifies the top left cell of the results area. When you run the tool, it will generate the Chi-square table starting at this cell.

Columns / Rows

Specifies whether the data to be analyzed is organized in columns or rows.

To provide an example of using this tool, we again make use of the input data set shown in Figure 53. In this case the data in column A is the observed data while the data in column B are the corresponding expected values. Figure 60 shows the chi-square results calculated for this input data using the settings shown in Figure 59.

D	E
Test of Independence (Chi-Square)	
Alpha	0.05
df	12
P-value	2.33E-14
Test Statistic	91.687006
Critical Value	21.026070

Figure 60: Results of chi-square test



Tip

For more information on chi-square tests, refer to the corresponding Wikipedia article at https://en.wikipedia.org/wiki/Chi-squared_test.