



Calc Guide

Chapter 8

Using the DataPilot

This PDF is designed to be read onscreen, two pages at a time. If you want to print a copy, your PDF viewer should have an option for printing two pages on one sheet of paper, but you may need to start with page 2 to get it to print facing pages correctly. (Print this cover page separately.)

Copyright

This document is Copyright © 2009 by its contributors as listed in the section titled **Authors**. You may distribute it and/or modify it under the terms of either the [GNU General Public License](#), version 3 or later, or the [Creative Commons Attribution License](#), version 3.0 or later.

All trademarks within this guide belong to their legitimate owners.

Authors

Stefan Weigel
Jean Hollis Weber
Sigrid Kronenberger
Andy Brown

Feedback

Please direct any comments or suggestions about this document to:
authors@user-faq.openoffice.org

Acknowledgments

The chapter is adapted from a German original written by Stefan Weigel and translated into English by Sigrid Kroneneberger. Andy Brown and Jean Hollis Weber edited the translation, and Andy Brown recaptured the illustrations from an English version of OpenOffice.org.

Publication date and software version

Published 18 May 2009. Based on OpenOffice.org 3.1.



*You can download
an editable version of this document from
<http://oooauthors.org/en/authors/userguide3/published/>*

Contents

Copyright.....	2
Introduction.....	5
Examples with step by step descriptions.....	5
Example 1: Sales volume overview.....	5
Practical problems and questions.....	6
Solution.....	6
Advantages.....	9
Example 2: Timekeeping.....	12
Practical problems and questions.....	14
Solution.....	14
Differences and advantages.....	19
Example 3: Frequency distribution.....	19
Solution with a matrix formula.....	21
Solution with the DataPilot.....	22
DataPilot functions in detail.....	27
The database (preconditions).....	27
Start.....	28
Data source.....	29
Calc sheet.....	29
Registered database.....	30
External data (OLAP).....	30
The DataPilot dialog.....	30
Basic layout.....	30
More options.....	32
More settings for the fields.....	34
Data fields.....	34
Row and column fields.....	36
Page fields.....	39
Working with the results of the DataPilot.....	39
Start the dialog.....	40

Change layout by using drag and drop.....	40
Grouping rows or columns.....	40
Grouping of categories with scalar values.....	40
Grouping of categories with date or time values.....	41
Grouping without the automatic creation of intervals.....	42
Sorting of the results.....	44
Sort manually by using drag and drop.....	44
Automatic sorting.....	44
Drill (show details).....	45
Filtering.....	47
Updating (refreshing) changed values.....	48
Cell formatting.....	48
Multiple data fields.....	49
Shortcuts.....	53
Function GETPIVOTDATA.....	53
Difficulty.....	53
Syntax.....	54
First syntax variation.....	55
Second syntax variation.....	55

Introduction

In requests for office software support there are regularly questions caused by some awkward solution strategies for day to day exercises in a spreadsheet. By not knowing simple but powerful concepts, users create very complicated formulas. A much more solid, simple, and flexible solution is based on using the DataPilot.

The DataPilot is not just a tool for advanced users. If beginners would know about this tool, it would save them many hours of work.

This chapter is divided into two sections:

- “Examples with step by step descriptions” uses three typical cases to demonstrate the advantages and applications of the DataPilot. You can follow these examples and learn by doing how easy it is to use the DataPilot.
- “DataPilot functions in detail” starting on page 27 describes the DataPilot in detail. You can use this part for looking up how a function is used.

Examples with step by step descriptions

This section demonstrates some of the possibilities of the DataPilot in some typical use cases. By following the step by step description, you should be able to recreate the examples yourself and learn about the power of the DataPilot.

Example 1: Sales volume overview

A typical introductory example in courses and books for beginners with spreadsheets is a simple sales volume overview. (See Figure 1.)

This example demonstrates the basic way to use the user interface and how to insert text and numbers into cells. Useful aids like *AutoFill* and *drag and drop* have been demonstrated in other chapters. The most important part is the connection between cells through formulas, for example addition with the plus operator or the *SUM* function.

This small exercise might be useful for a first contact with the program, but it shows only a very small fraction of the tasks in an office. To create such a sales overview, you also need the original data. That is, before you can use a spreadsheet for creating the sales overview, you need to add many single purchases from different lists and then enter the sums into the relevant fields **C5** to **F7**.

	A	B	C	D	E	F	G
1							
2				Experience-Travel			
3							
4			Jan	Feb	Mar	Apr	Total
5		Golf	1000	1100	1200	1300	4600
6		Tennis	2000	2200	2400	2600	9200
7		Sailing	3000	3300	3600	3900	13800
8		Total	=SUM(C5:C7)		7200	7800	27600
9							

Figure 1: Typical example for beginners

Practical problems and questions

- 1) For displaying additional values (for May, June, July, and so on) over time, you need to add extra columns. This means that you have to change the structure of the calculation sheet. This is not only somewhat inefficient from a workflow point of view, it also adds some practical questions: How do references react, if you add more columns or rows to the sum formulas?
- 2) The layout, where the timeline is displayed horizontally, might be less convenient if you add more months. For using space more efficiently, a vertical layout might be better. How can the table then be transposed? Do you have to enter everything again?
- 3) Assume that the management asks unexpected questions or for an additional subdivision for the different sales regions or a constraint of the sales for a given employee. In this case you again have to manually add all the sums and create different tables in many variations.
- 4) Is it really acceptable and realistic to create such an overview by adding the different values manually? This is really a lot of work and extremely error-prone!

Solution

The most important part of your task in the example is the addition of the sales per month field. This had to be done manually. To do this automatically with the program, just get the data into Calc. You can enter the single numbers by hand or you can import a file from your bookkeeping software. In any case we assume a continuous table, that keeps track of all sales in a somewhat primitive form.

	A	B	C	D	E
1	date	sales	category	region	employee
2	4/13/08	\$498.00	sailing	north	Kurt
3	2/7/08	\$1,383.00	sailing	south	Kurt
4	5/28/08	\$4,655.00	tennis	east	Hans
5	4/20/08	\$3,993.00	sailing	east	Fritz
6	2/28/08	\$3,377.00	golfing	south	Ute
7	1/27/08	\$2,095.00	sailing	north	Brigitte
8	2/16/08	\$4,731.00	sailing	east	Ute
9	4/15/08	\$4,330.00	tennis	south	Kurt
10	4/20/08	\$3,664.00	golfing	south	Kurt
11	6/7/08	\$769.00	tennis	west	Brigitte
12	2/29/08	\$293.00	sailing	east	Kurt
13	4/6/08	\$3,705.00	tennis	north	Kurt
14	6/19/08	\$1,747.00	golfing	west	Hans
15	5/14/08	\$449.00	tennis	west	Ute
16	5/30/08	\$1,795.00	golfing	south	Kurt
17	2/28/08	\$610.00	sailing	west	Ute
18	6/8/08	\$1,625.00	tennis	east	Fritz
19	5/12/08	\$3,887.00	sailing	south	Brigitte
20	1/30/08	\$2,458.00	tennis	west	Hans
21	5/9/08	\$4,369.00	golfing	south	Ute
22	3/15/08	\$1,785.00	golfing	east	Hans
23	3/27/08	\$669.00	sailing	south	Hans
24	1/28/08	\$155.00	sailing	west	Brigitte
25	2/5/08	\$2,657.00	golfing	south	Hans
26	2/28/08	\$3,145.00	tennis	west	Brigitte
27	2/4/08	\$2,433.00	tennis	south	Kurt
28	2/2/08	\$1,218.00	tennis	west	Brigitte
29	1/2/08	\$410.00	tennis	north	Kurt
30	2/6/08	\$3,443.00	golfing	north	Fritz
31	1/6/08	\$3,821.00	tennis	south	Fritz

Figure 2: Basic data in Calc

You can create the sales volume overview by following these instructions:

- 1) Select the cell **A1** (or any other single cell within the list).
- 2) Select **Data > DataPilot > Start** and click **OK**.
- 3) The DataPilot dialog (Figure 3) has four layout areas (large white fields) and five fields that look similar to buttons. These small fields are the titles of the different columns of your list.
 - Move the field **date** into the Layout area *Column fields* .
 - Move the field **sales** into the Layout area *Data fields*.
 - Move the field **category** into the Layout area *Row fields*.
- 4) Click **More**, to see more options in the lower part of the dialog.
- 5) Choose - **new sheet** - for *Results to*.
- 6) Click **OK**.

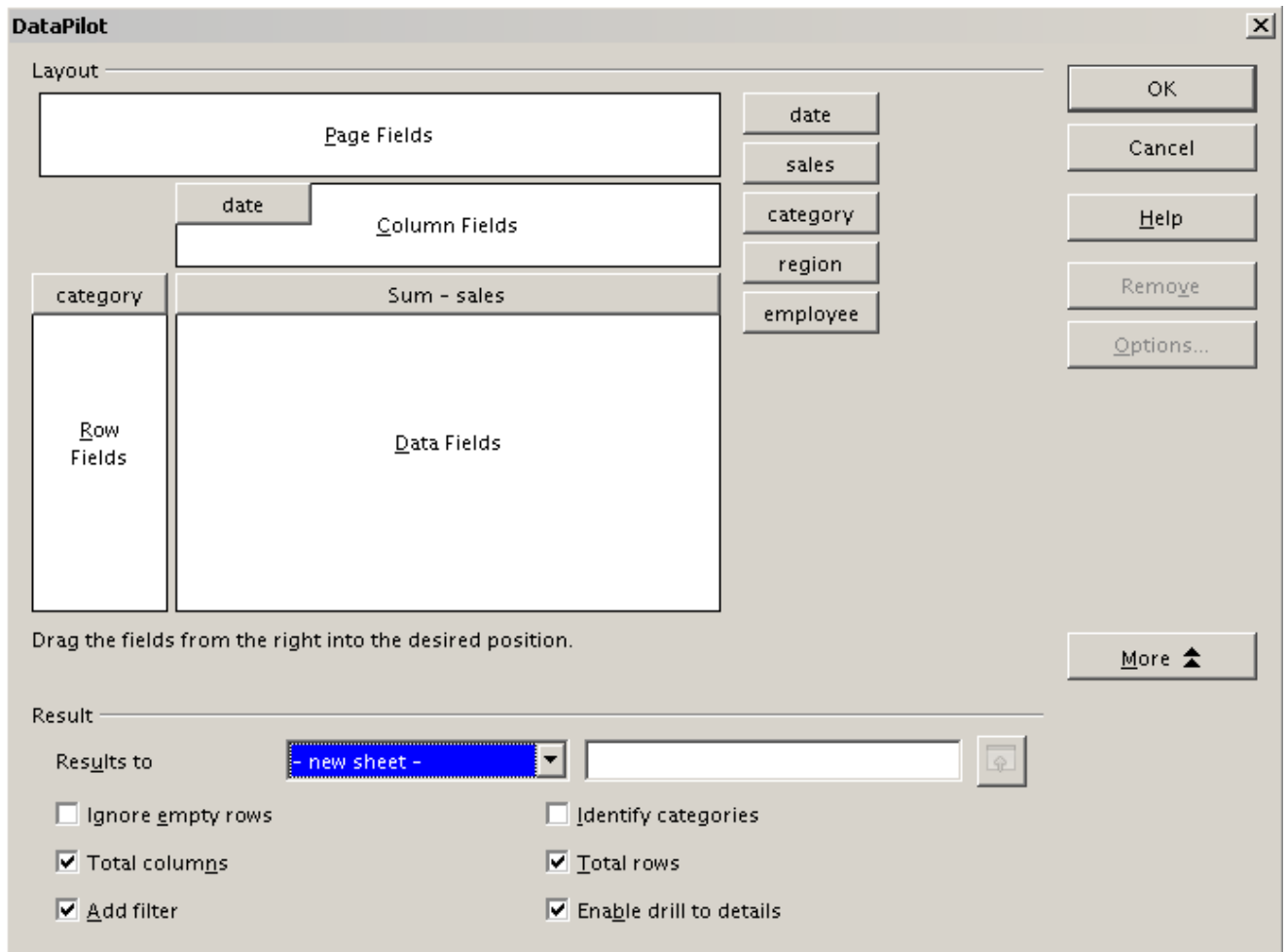


Figure 3: DataPilot dialog

- 7) The result appears on a new sheet. It has the wanted structure, but the columns are not yet grouped into months.

	A	B	C	D	E	F	G	H
1	Filter							
2								
3	Sum - sales	date						
4	category	01/02/08	01/03/08	01/06/08	01/07/08	01/08/08	01/10/08	01/14/08
5	golfing	\$1,508.00			\$3,739.00	\$4,195.00	\$2,023.00	
6	sailing	\$2,340.00						
7	tennis	\$410.00	\$4,872.00	\$6,444.00				\$2,804.00
8	Total Result	\$4,258.00	\$4,872.00	\$6,444.00	\$3,739.00	\$4,195.00	\$2,023.00	\$2,804.00
9								

Figure 4: DataPilot result without grouping

- 8) To group the columns. select cell **B4** or any other cell that contains a date. Then select **Data > Group and Outline > Group**. On the Grouping dialog (Figure 5), make sure *Intervals* and *Months* are selected, and click **OK**. The result is now grouped for months (Figure 6).

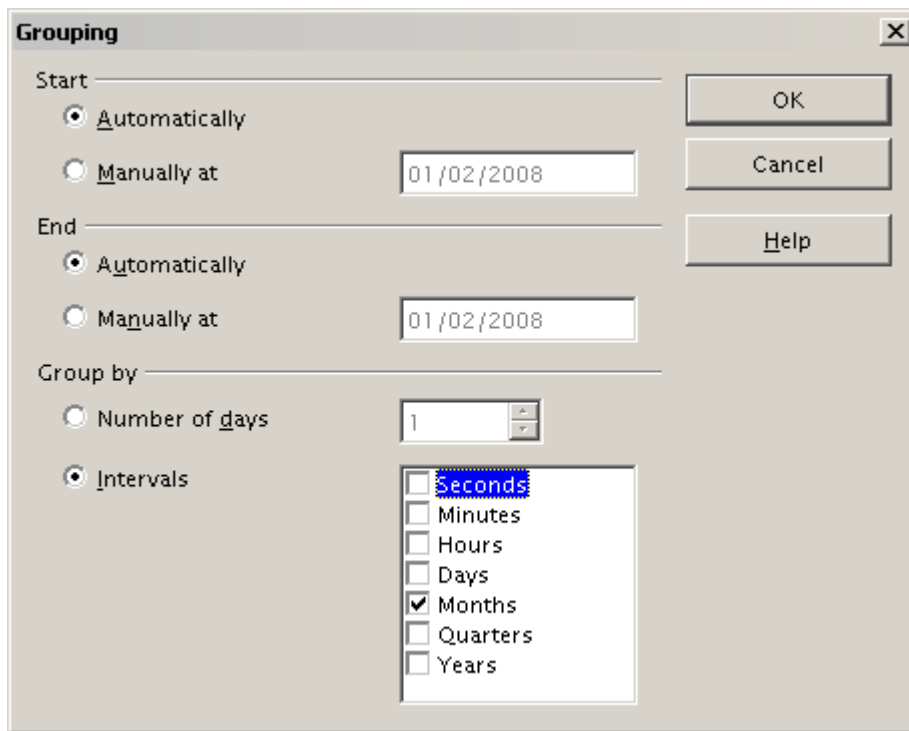


Figure 5: Grouping on months

	A	B	C	D	E	F	G	H
1	Filter							
2								
3	Sum - sales	date						
4	category	Jan	Feb	Mar	Apr	May	Jun	Total Result
5	golfing	\$26,180.00	\$30,444.00	\$31,714.00	\$24,747.00	\$24,686.00	\$18,443.00	\$156,214.00
6	sailing	\$13,979.00	\$15,625.00	\$17,409.00	\$19,769.00	\$21,799.00	\$21,271.00	\$109,852.00
7	tennis	\$39,206.00	\$23,710.00	\$12,097.00	\$31,918.00	\$11,439.00	\$24,329.00	\$142,699.00
8	Total Result	\$79,365.00	\$69,779.00	\$61,220.00	\$76,434.00	\$57,924.00	\$64,043.00	\$408,765.00
9								

Figure 6: DataPilot result grouped for months

In this result you will recognize the beginners' example. It is very easy to produce, without any further knowledge about the spreadsheet. (You do not have to enter any formulas.)

Advantages

- 1) No manual adding of single values necessary. No manual entering of the values. Less work and fewer errors.
- 2) The layout is very flexible: months horizontal and fields vertical or vice versa: two mouse clicks.
- 3) Additional differentiating factors are immediately available.
- 4) Many evaluation possibilities, for example: number or average instead of sum, accumulated values, comparisons and more.

We will now demonstrate some of these advantages.

Starting with the result of Figure 6, move the **Date** field to the left of the **Category** field. Now the summary is as shown in Figure 7.

	A	B	C
1	Filter		
2			
3	category	date	
4	golfing	Jan	\$26,180.00
5		Feb	\$30,444.00
6		Mar	\$31,714.00
7		Apr	\$24,747.00
8		May	\$24,686.00
9		Jun	\$18,443.00
10	sailing	Jan	\$13,979.00
11		Feb	\$15,625.00
12		Mar	\$17,409.00
13		Apr	\$19,769.00
14		May	\$21,799.00
15		Jun	\$21,271.00
16	tennis	Jan	\$39,206.00
17		Feb	\$23,710.00
18		Mar	\$12,097.00
19		Apr	\$31,918.00
20		May	\$11,439.00
21		Jun	\$24,329.00
22	Total Result		\$408,765.00
23			

Figure 7: Layout option with only one mouse click

To transpose the table completely, just move the field Category above the area of the displayed values, to **C3**. (See Figure 8.)

	A	B	C	D	E
1	Filter				
2					
3	Sum - sales	category			
4	date	golfing	sailing	tennis	Total Result
5	Jan	\$26,180.00	\$13,979.00	\$39,206.00	\$79,365.00
6	Feb	\$30,444.00	\$15,625.00	\$23,710.00	\$69,779.00
7	Mar	\$31,714.00	\$17,409.00	\$12,097.00	\$61,220.00
8	Apr	\$24,747.00	\$19,769.00	\$31,918.00	\$76,434.00
9	May	\$24,686.00	\$21,799.00	\$11,439.00	\$57,924.00
10	Jun	\$18,443.00	\$21,271.00	\$24,329.00	\$64,043.00
11	Total Result	\$156,214.00	\$109,852.00	\$142,699.00	\$408,765.00
12					

Figure 8: Transposed layout of Figure 7

In contrast to the beginners' example in Figure 1, it is now very simple to view or add different aspects of the underlying data. For example, to see the values for different regions just do the following:

- 1) Select the cell **A3** (or any other single cell that is part of the DataPilot result).

- 2) Select **Data > DataPilot > Start**, to start the DataPilot again. Drag the field **Region** into the layout area *Row fields*. Depending on the order you choose for the row fields, you get in the result either regions with date subdivisions or vice versa.
- 3) Click **OK**.

	A	B	C	D	E	F
1	Filter					
2						
3	Sum - sales		category			
4	date	region	golfing	sailing	tennis	Total Result
5	Jan	east	\$3,531.00	\$1,956.00	\$18,566.00	\$24,053.00
6		north	\$4,842.00	\$2,095.00	\$7,791.00	\$14,728.00
7		south	\$11,580.00	\$2,340.00	\$3,821.00	\$17,741.00
8		west	\$6,227.00	\$7,588.00	\$9,028.00	\$22,843.00
9	Feb	east	\$5,392.00	\$6,124.00		\$11,516.00
10		north	\$3,443.00		\$938.00	\$4,381.00
11		south	\$15,990.00	\$4,458.00	\$12,490.00	\$32,938.00
12		west	\$5,619.00	\$5,043.00	\$10,282.00	\$20,944.00
13	Mar	east	\$11,962.00	\$5,275.00	\$2,219.00	\$19,456.00
14		north	\$5,193.00		\$4,903.00	\$10,096.00
15		south	\$4,881.00	\$5,546.00		\$10,427.00
16		west	\$9,678.00	\$6,588.00	\$4,975.00	\$21,241.00
17	Apr	east	\$5,343.00	\$4,902.00	\$2,417.00	\$12,662.00
18		north		\$10,551.00	\$12,658.00	\$23,209.00
19		south	\$7,750.00	\$4,316.00	\$9,203.00	\$21,269.00
20		west	\$11,654.00		\$7,640.00	\$19,294.00
21	May	east	\$10,557.00	\$2,092.00	\$7,137.00	\$19,786.00
22		north	\$2,214.00	\$520.00		\$2,734.00
23		south	\$11,228.00	\$18,307.00	\$675.00	\$30,210.00
24		west	\$687.00	\$880.00	\$3,627.00	\$5,194.00
25	Jun	east	\$5,186.00	\$2,135.00	\$5,627.00	\$12,948.00
26		north	\$3,049.00	\$9,302.00	\$8,243.00	\$20,594.00
27		south	\$4,828.00	\$9,834.00	\$8,069.00	\$22,731.00
28		west	\$5,380.00		\$2,390.00	\$7,770.00
29	Total Result		\$156,214.00	\$109,852.00	\$142,699.00	\$408,765.00
30						

Figure 9: Additional subdivision into regions, added later

In another variation you want to add the mentioned employees.

- 1) Select the cell **A3** (or any other single cell that is part of the DataPilot result).
- 2) Select **Data > DataPilot > Start**, to start the DataPilot again.
 - You do not need the field **Region** in this case. Drag it out of the layout area.
 - Drag the field **Employee** into the layout area *Page fields*.
- 3) Click **OK**.

Fields that you use as page fields are placed in the result above the summary with the name *Filter*. You then have a drop-down list that you can use to show only the sums of a given employee:

	A	B	C	D	E
1	Filter				
2	employee	- all -			
3		- all -			
4	Sum - sales	Brigitte			
5	date	Fritz	sailing	tennis	Total Result
6	Jan	Hans	\$13,979.00	\$39,206.00	\$79,365.00
7	Feb	Kurt	\$15,625.00	\$23,710.00	\$69,779.00
8	Mar	Ute	\$17,409.00	\$12,097.00	\$61,220.00
9	Apr		\$19,769.00	\$31,918.00	\$76,434.00
10	May		\$21,799.00	\$11,439.00	\$57,924.00
11	Jun		\$21,271.00	\$24,329.00	\$64,043.00
12	Total Result		\$109,852.00	\$142,699.00	\$408,765.00
13					
14					

Figure 10: Selection of subtotals for several employees.

Up to here we have not seen the most powerful features of the DataPilot. The following examples will show you more.

Example 2: Timekeeping

This example is often used within the consultancy field and is also used in several variations in the user support area. The task is to provide a means for one or more users to keep track of working hours.

A typical way of doing this is to create a spreadsheet per month and a sum sheet with all the results of one year. For each employee, there is one file.

	A	B	C	D	E	F
1			Time sheet for Erika Mustermann			
2			January 2008			
3						
4		date	arrives	leaves	break	hours
5		01/01/08	08:00	17:30	00:30	9.00
6		01/02/08	07:45	14:45	00:30	6.50
7		01/03/08	09:00	18:00	00:30	8.50
8		01/04/08	07:15	17:45	01:00	9.50
9		01/05/08				
10		01/06/08				
11		01/07/08	08:15	19:30	01:00	10.25
12		01/08/08	08:15	20:00	00:30	11.25
13		01/09/08	07:45	16:45	00:30	8.50
14		01/10/08	08:15	13:45	00:00	5.50
15		01/11/08	08:00	15:15	00:30	6.75
16		01/12/08				
17		01/13/08				
18		01/14/08	08:00	13:45	00:00	5.75
19		01/15/08	07:30	13:00	00:00	5.50
20		01/16/08	08:45	18:45	00:30	9.50
21		01/17/08	08:45	20:30	01:00	10.75
22		01/18/08	07:30	13:45	00:30	5.75
23		01/19/08				
24		01/20/08				
25		01/21/08	08:15	18:45	01:00	9.50
26		01/22/08	08:15	16:45	00:30	8.00
27		01/23/08	08:45	15:00	00:30	5.75
28		01/24/08	08:30	19:30	01:00	10.00

Figure 11: One month of timekeeping for one employee

1			
2		Timekeeping for Erika Mustermann	
3			
4		2008	
5			
6		January	183.25
7		February	165.50
8		March	172.25
9		April	162.00
10		May	168.50
11		June	0.00
12		July	0.00
13		August	
14		September	
15		October	
16		November	
17		December	
18		Total	851.50
19			
20			
21			

Figure 12: Yearly sums for one employee

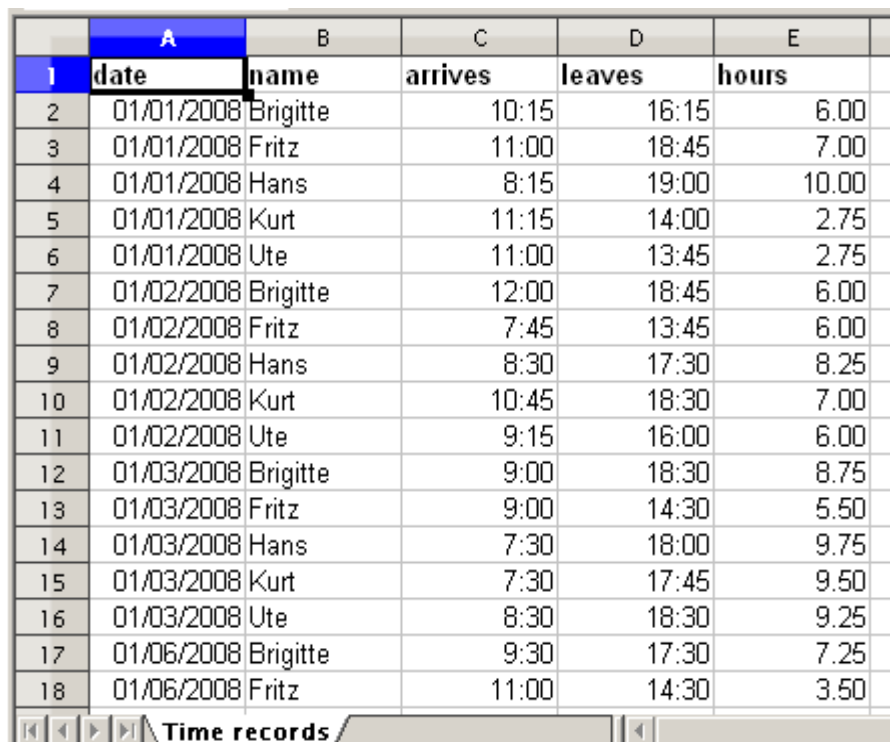
Practical problems and questions

- 1) It is very difficult and time-consuming to create the timekeeping table: 12 sheets that have to be copied from a raw template and adjusted for each month, and a sheet with all the yearly sums with references to all the other sheets. Users often search for a macro to make the creation easier.
- 2) The shown file contains only the data of one employee. How can you get all the data for all the employees, so that you can have a summary of all the work hours from all employees of a department or the whole company?
- 3) How can you compare employees and/or departments?
- 4) The shown file contains data for one year. How can you compare it with the data of the previous years?

Solution

To use the DataPilot for this task, collect all data into one single table. This can be done either manually or by importing a file from an (electronic) timekeeping machine.

In very simple cases each employee takes care of their own working hours. If you need calculations that cover several employees, departments, or the whole company, just copy everything into one huge table.



	A	B	C	D	E
1	date	name	arrives	leaves	hours
2	01/01/2008	Brigitte	10:15	16:15	6.00
3	01/01/2008	Fritz	11:00	18:45	7.00
4	01/01/2008	Hans	8:15	19:00	10.00
5	01/01/2008	Kurt	11:15	14:00	2.75
6	01/01/2008	Ute	11:00	13:45	2.75
7	01/02/2008	Brigitte	12:00	18:45	6.00
8	01/02/2008	Fritz	7:45	13:45	6.00
9	01/02/2008	Hans	8:30	17:30	8.25
10	01/02/2008	Kurt	10:45	18:30	7.00
11	01/02/2008	Ute	9:15	16:00	6.00
12	01/03/2008	Brigitte	9:00	18:30	8.75
13	01/03/2008	Fritz	9:00	14:30	5.50
14	01/03/2008	Hans	7:30	18:00	9.75
15	01/03/2008	Kurt	7:30	17:45	9.50
16	01/03/2008	Ute	8:30	18:30	9.25
17	01/06/2008	Brigitte	9:30	17:30	7.25
18	01/06/2008	Fritz	11:00	14:30	3.50

Figure 13: Data basis in Calc

Using the DataPilot just needs 12 mouse clicks and gives you within a few seconds a nice overview:

- 1) Select the cell **A1** (or any other single cell within the list).
- 2) Choose **Data > DataPilot > Start** and click **OK**.
- 3) On the DataPilot dialog:
 - Drag the field **date** into the layout area *Row fields*.
 - Drag the field **hours** into the layout area *Data fields*. Notice that it becomes **Sum - hours**.
 - Drag the field **name** into the layout area *Column fields*.
- 4) Click **More** to show more options in the lower part of the dialog.
- 5) Choose **- new sheet -** for *Results to*.
- 6) Click **OK**.

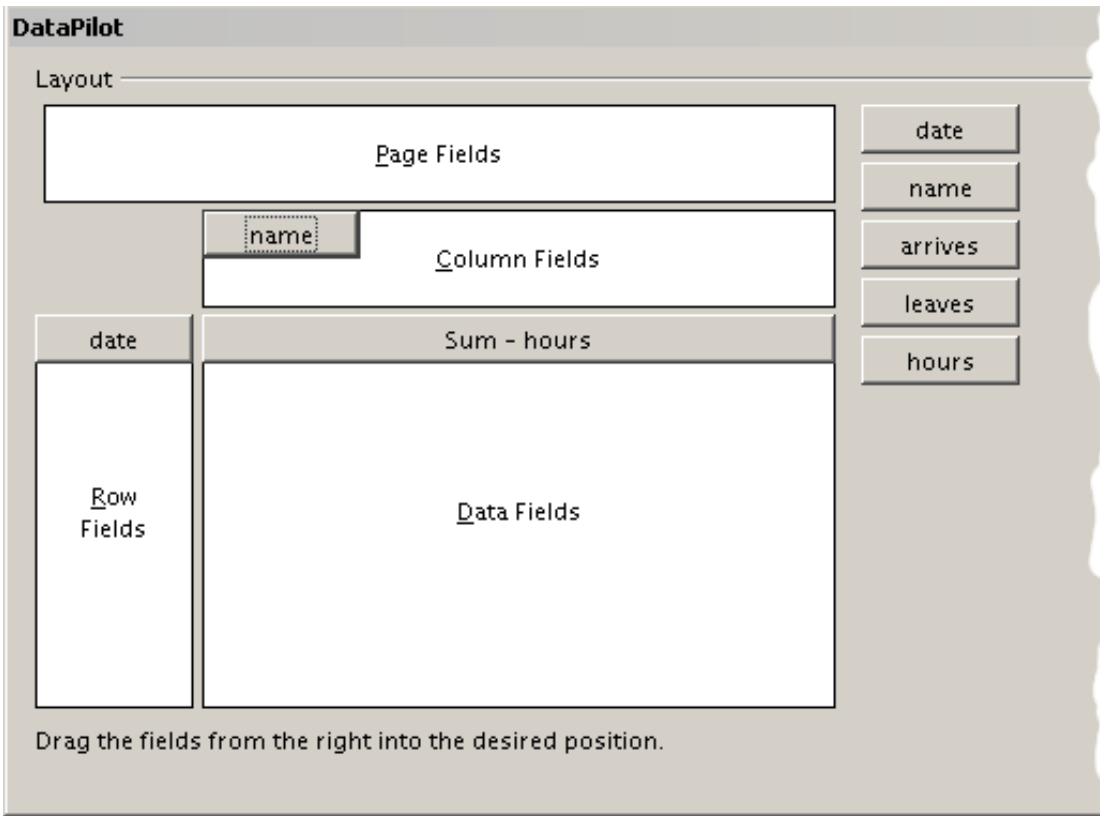


Figure 14: Part of DataPilot dialog

The result appears on a new sheet (Figure 15).

	A	B	C	D	E	F	G
1	Filter						
2							
3	Sum - hours	name					
4	date	Brigitte	Fritz	Hans	Kurt	Ute	Total Result
5	01/01/08	6.00	7.00	10.00	2.75	2.75	28.50
6	01/02/08	6.00	6.00	8.25	7.00	6.00	33.25
7	01/03/08	8.75	5.50	9.75	9.50	9.25	42.75
8	01/06/08	7.25	3.50	6.00	5.00	4.50	26.25
9	01/07/08	4.75	9.00	6.00	5.75	5.75	31.25
10	01/08/08	7.25	4.75	7.75	4.75	7.25	31.75
11	01/09/08	6.75	2.50	5.50	6.00	5.00	25.75
12	01/10/08	7.00	8.75	3.75	10.50	3.50	33.50
13	01/13/08	6.00	7.00	7.50	2.75	5.50	28.75
14	01/14/08	5.50	6.75	7.25	4.25	8.00	31.75
15	01/15/08	4.75	7.75	4.00	9.25	7.00	32.75
16	01/16/08	10.25	4.50	8.00	5.75	5.50	34.00
17	01/17/08	4.75	7.00	8.75	5.50	6.50	32.50
18	01/20/08	8.75	5.75	10.75	6.25	8.75	40.25
19	01/21/08	10.00	6.00	9.25	6.75	3.25	35.25
20	01/22/08	8.50	8.00	9.75	1.50	7.00	34.75
21	01/23/08	5.50	6.00	5.50	8.25	3.50	28.75
22	01/24/08	7.25	3.00	5.50	7.50	4.00	27.25

Figure 15: The evaluation, done within seconds with DataPilot

The result is much more powerful than the solution with the classic formula-based calculation. For example, you can summarize the daily results to a monthly result very easily.

- 1) To group together the rows, select the cell **A5** (or any other cell that contains a date).
- 2) Choose **Data > Group and Outline > Group** and click **OK**. The result is now grouped into months.

	A	B	C	D	E	F	G
1	Filter						
2							
3	Sum - hours	name					
4	date	Brigitte	Fritz	Hans	Kurt	Ute	Total Result
5	Jan	150.75	137.00	166.50	141.75	137.00	733.00
6	Feb	131.50	139.25	112.25	119.75	116.75	619.50
7	Mar	140.25	154.50	135.75	134.75	141.00	706.25
8	Apr	140.50	128.50	131.75	131.25	144.00	676.00
9	May	126.00	134.75	116.00	121.25	140.00	638.00
10	Jun	127.75	128.00	118.25	120.25	121.25	615.50
11	Total Result	816.75	822.00	780.50	769.00	800.00	3988.25
12							

Figure 16: Monthly sums

If you need a result with a percentage, start the DataPilot from this page.

- 1) Select the cell **A3** (or any other cell that contains a result of the DataPilot).
- 2) Choose **Data > DataPilot > Start**.
- 3) Double-click on the field **Sum - hours** to open the Data Field dialog (Figure 17).
- 4) Click on **More**, to see more options.
- 5) Switch the type of the displayed value to **% of column**.

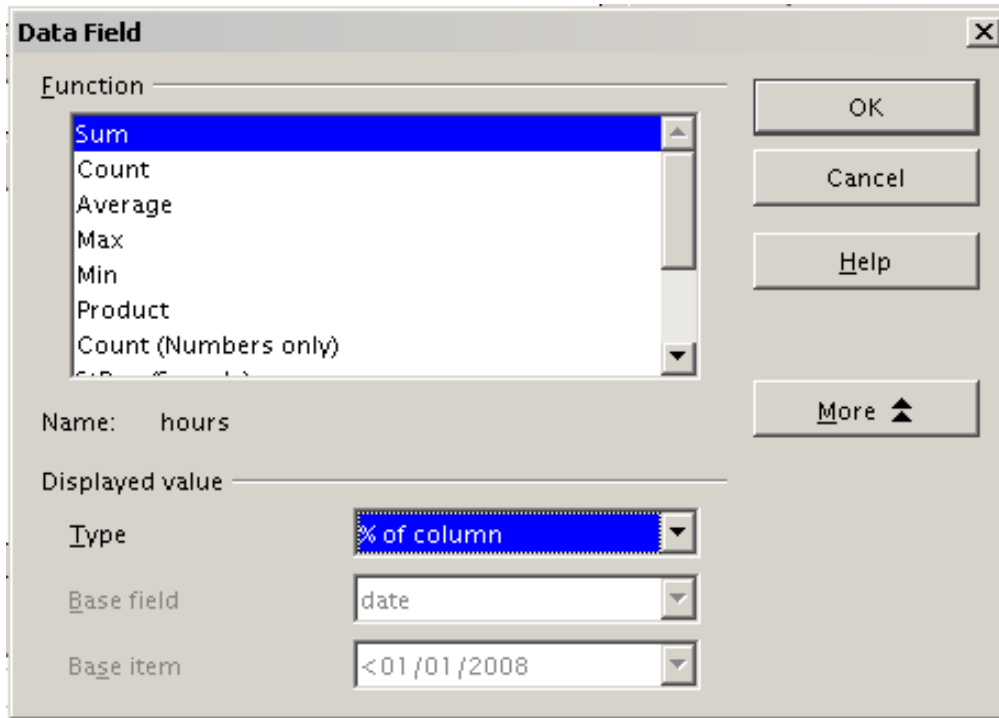


Figure 17: Properties of the data field

- 6) Click **OK** twice.

In the result the percentages are shown as a decimal number. If you prefer to have a percent format, select the cells and click on the % icon in the menu bar.

	A	B	C	D	E	F	G
1	Filter						
2							
3	Sum - hours	name					
4	date	Brigitte	Fritz	Hans	Kurt	Ute	Total Result
5	Jan	18.46%	16.67%	21.33%	18.43%	17.13%	18.38%
6	Feb	16.10%	16.94%	14.38%	15.57%	14.59%	15.53%
7	Mar	17.17%	18.80%	17.39%	17.52%	17.63%	17.71%
8	Apr	17.20%	15.63%	16.88%	17.07%	18.00%	16.95%
9	May	15.43%	16.39%	14.86%	15.77%	17.50%	16.00%
10	Jun	15.64%	15.57%	15.15%	15.64%	15.16%	15.43%
11	Total Result	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
12							

Figure 18: Result with percentages

To get a comparison between employees, start the DataPilot again:

- 1) Select the cell **A3** (or any other cell that contains a result of the DataPilot).
- 2) Choose **Data > DataPilot > Start**.
- 3) Double-click on the field **Sum - hours** to open the Data Field dialog.
- 4) Click **More**, to see more options.
 - Switch the *Type* of the displayed value to **Difference from**.
 - Switch the *Base field* to name.
 - Switch the *Base item* to Brigitte.
- 5) Click **OK** twice.

	A	B	C	D	E	F	G
1	Filter						
2							
3	Sum - hours	name					
4	date	Brigitte	Fritz	Hans	Kurt	Ute	Total Result
5	Jan		-13.75	15.75	-9.00	-13.75	
6	Feb		7.75	-19.25	-11.75	-14.75	
7	Mar		14.25	-4.50	-5.50	0.75	
8	Apr		-12.00	-8.75	-9.25	3.50	
9	May		8.75	-10.00	-4.75	14.00	
10	Jun		0.25	-9.50	-7.50	-6.50	
11	Total Result		5.25	-36.25	-47.75	-16.75	

Figure 19: Absolute comparison with Brigitte

As a last example we switch to an accumulated view; that is, continuing sums of all values:

- 1) Choose **Data > DataPilot > Start**.
- 2) Double-click on the field **Sums - hours** to open the Data Field dialog.
- 3) Click **More**, to see more options.
 - Switch the type of the displayed value to **Running total in**.
 - Switch the *Base field* to **Date**.
- 4) Click **OK** twice.

	A	B	C	D	E	F	G
1	Filter						
2							
3	Sum - hours	name					
4	date	Brigitte	Fritz	Hans	Kurt	Ute	Total Result
5	Jan	150.75	137.00	166.50	141.75	137.00	733.00
6	Feb	282.25	276.25	278.75	261.50	253.75	1352.50
7	Mar	422.50	430.75	414.50	396.25	394.75	2058.75
8	Apr	563.00	559.25	546.25	527.50	538.75	2734.75
9	May	689.00	694.00	662.25	648.75	678.75	3372.75
10	Jun	816.75	822.00	780.50	769.00	800.00	3988.25
11	Total Result						
12							

Figure 20: The DataPilot now shows accumulated values

Differences and advantages

These experiments show an important aspect of the DataPilot concept. With the classic method you have to collect your data according to the way you want the result to be represented. This means you have to use a specific structure and you're stuck with it.

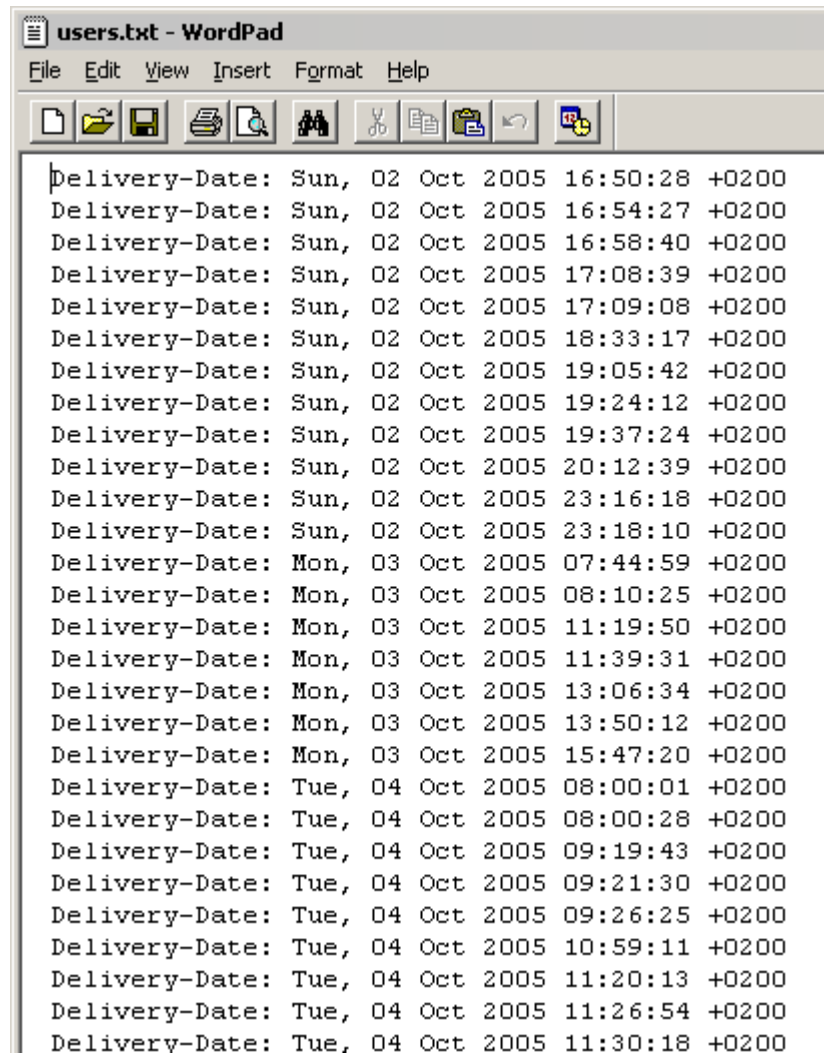
The DataPilot works more like a real database. The source data are collected in a regular form, a somewhat primitive but universal structure: a simple spreadsheet that contains all data. Only when you want to look at it, you select which part of the data you want to use.

Example 3: Frequency distribution

For showing the frequency of incidents, Calc uses the function FREQUENCY. This formula has to be used in a so-called matrix formula. This is an advanced feature. As an alternative you can use the DataPilot, which requires almost no knowledge about spreadsheets.

In our example we want to investigate the number of emails that go to the Germanophone support mailing list (users@de.openoffice.org). We want to know how the activity on the list is distributed during the day.

The starting point for this example is the mailbox file of the Thunderbird mail client. The command `grep 'Delivery-Date:' Inbox` creates a text file that contains for the timespan of more than 2 years one line for each message, with date and time (Figure 21).

A screenshot of a WordPad window titled 'users.txt - WordPad'. The window contains a list of 25 lines of text, each representing a message delivery record. The records are organized by day: Sunday, October 2, 2005 (12 records), Monday, October 3, 2005 (8 records), and Tuesday, October 4, 2005 (5 records). Each record follows the format 'Delivery-Date: Day, DD Oct YYYY HH:MM:SS +0200'.

```
Delivery-Date: Sun, 02 Oct 2005 16:50:28 +0200
Delivery-Date: Sun, 02 Oct 2005 16:54:27 +0200
Delivery-Date: Sun, 02 Oct 2005 16:58:40 +0200
Delivery-Date: Sun, 02 Oct 2005 17:08:39 +0200
Delivery-Date: Sun, 02 Oct 2005 17:09:08 +0200
Delivery-Date: Sun, 02 Oct 2005 18:33:17 +0200
Delivery-Date: Sun, 02 Oct 2005 19:05:42 +0200
Delivery-Date: Sun, 02 Oct 2005 19:24:12 +0200
Delivery-Date: Sun, 02 Oct 2005 19:37:24 +0200
Delivery-Date: Sun, 02 Oct 2005 20:12:39 +0200
Delivery-Date: Sun, 02 Oct 2005 23:16:18 +0200
Delivery-Date: Sun, 02 Oct 2005 23:18:10 +0200
Delivery-Date: Mon, 03 Oct 2005 07:44:59 +0200
Delivery-Date: Mon, 03 Oct 2005 08:10:25 +0200
Delivery-Date: Mon, 03 Oct 2005 11:19:50 +0200
Delivery-Date: Mon, 03 Oct 2005 11:39:31 +0200
Delivery-Date: Mon, 03 Oct 2005 13:06:34 +0200
Delivery-Date: Mon, 03 Oct 2005 13:50:12 +0200
Delivery-Date: Mon, 03 Oct 2005 15:47:20 +0200
Delivery-Date: Tue, 04 Oct 2005 08:00:01 +0200
Delivery-Date: Tue, 04 Oct 2005 08:00:28 +0200
Delivery-Date: Tue, 04 Oct 2005 09:19:43 +0200
Delivery-Date: Tue, 04 Oct 2005 09:21:30 +0200
Delivery-Date: Tue, 04 Oct 2005 09:26:25 +0200
Delivery-Date: Tue, 04 Oct 2005 10:59:11 +0200
Delivery-Date: Tue, 04 Oct 2005 11:20:13 +0200
Delivery-Date: Tue, 04 Oct 2005 11:26:54 +0200
Delivery-Date: Tue, 04 Oct 2005 11:30:18 +0200
```

Figure 21: Raw data for the frequency of the messages

To import the data into a table in Calc, open the text file *users.txt* with the raw data. The Text Import dialog appears; here you can choose your import options. Use the following properties:

Separator options: *Fixed width*

Divide into columns at position 20, 31 and 40

Column types:

Column 1: *Hide*

Column 2: *Date (DMY)*

Column 3: *Standard*

Column 4: *Hide*

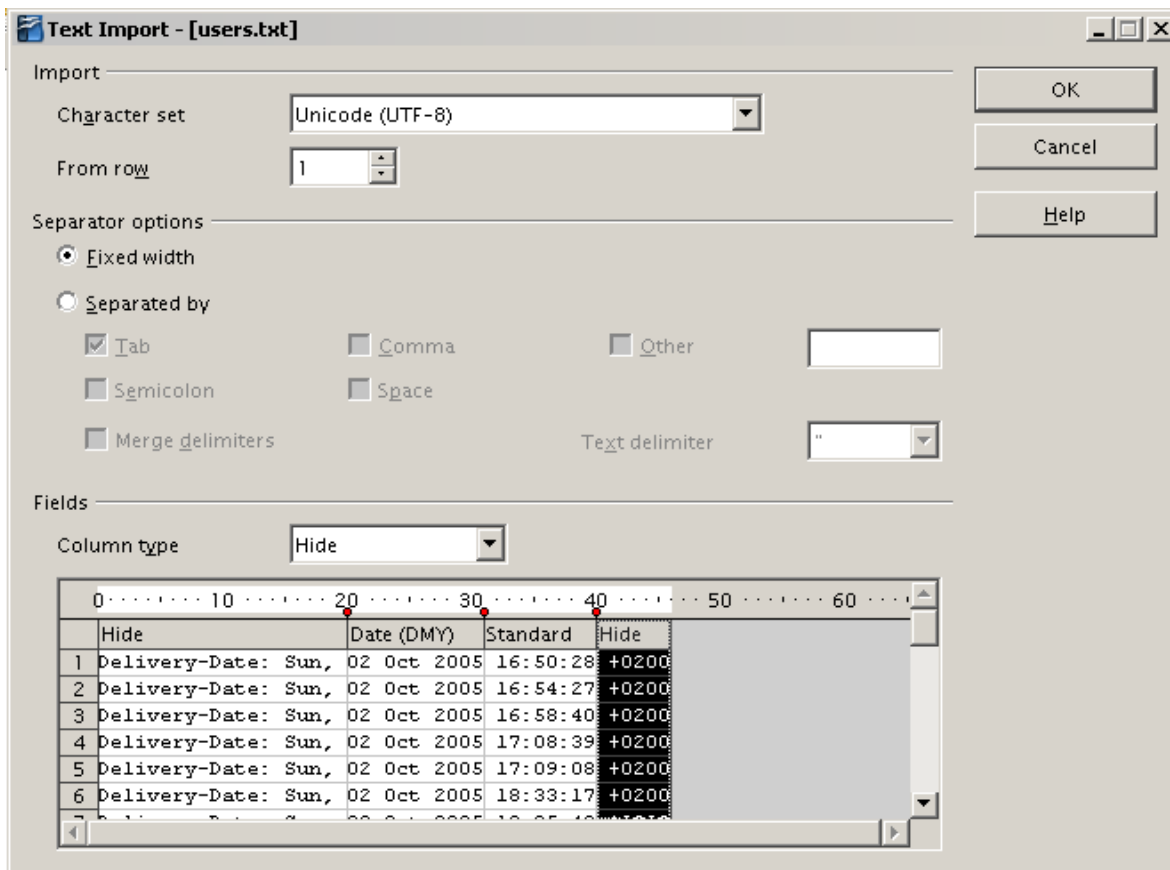


Figure 22: Import settings

Figure 23 shows the imported raw data with a line added at the top with titles for each column.

	A	B
1	Date	Time
2	10/02/05	16:50:28
3	10/02/05	16:54:27
4	10/02/05	16:58:40
5	10/02/05	17:08:39
6	10/02/05	17:09:08
7	10/02/05	18:33:17
8	10/02/05	19:05:42
9	10/02/05	19:24:12
10	10/02/05	19:37:24
11	10/02/05	20:12:39

Figure 23: Raw data in Calc

Solution with a matrix formula

To calculate the frequency you have to create 24 classes, one for each hour. In the next column you enter the number of emails with the function FREQUENCY.

={FREQUENCY(B2:B38096;E2:E25)}		
D	E	F
	01:00	922
	02:00	466
	03:00	204
	04:00	123
	05:00	97
	06:00	110
	07:00	272
	08:00	772
	09:00	1538
	10:00	1944
	11:00	2214
	12:00	2418
	13:00	2382
	14:00	2327
	15:00	2428
	16:00	2524
	17:00	2398
	18:00	2273
	19:00	2260
	20:00	2368
	21:00	2264
	22:00	2131
	23:00	2089
	00:00	1571

Figure 24: FREQUENCY function in a matrix formula

The first argument is the cell area with the times of all approximately 38,000 emails. The second argument is the cell area E2:E25 that describes the frequency classes. To enter the formula, first select the cell area F2:F25, then enter the formula. Then finish the formula by using the key combination *Shift+Ctrl+Enter*. This indicates to the program that you want to use a matrix formula. To indicate the matrix formula, the program uses curly brackets.

This technique is most often known to advanced users only.

Solution with the DataPilot

With the DataPilot you can achieve the same result much simpler and faster. The solution is also feasible for the less advanced user. Starting with the raw data (Figure 23), you need only a few mouse clicks:

- 1) Select the cell **A1** (or any other cell within the list).
- 2) Choose **Data > DataPilot > Start** and click **OK**.

- 3) In the DataPilot:
 - Drag the field **time** into the layout area *Row fields*.
 - Drag the field **date** into the layout area *Data fields*.
- 4) Click **More** to show more options in the lower part of the dialog.
- 5) Choose - **new sheet** - for *Results to*.

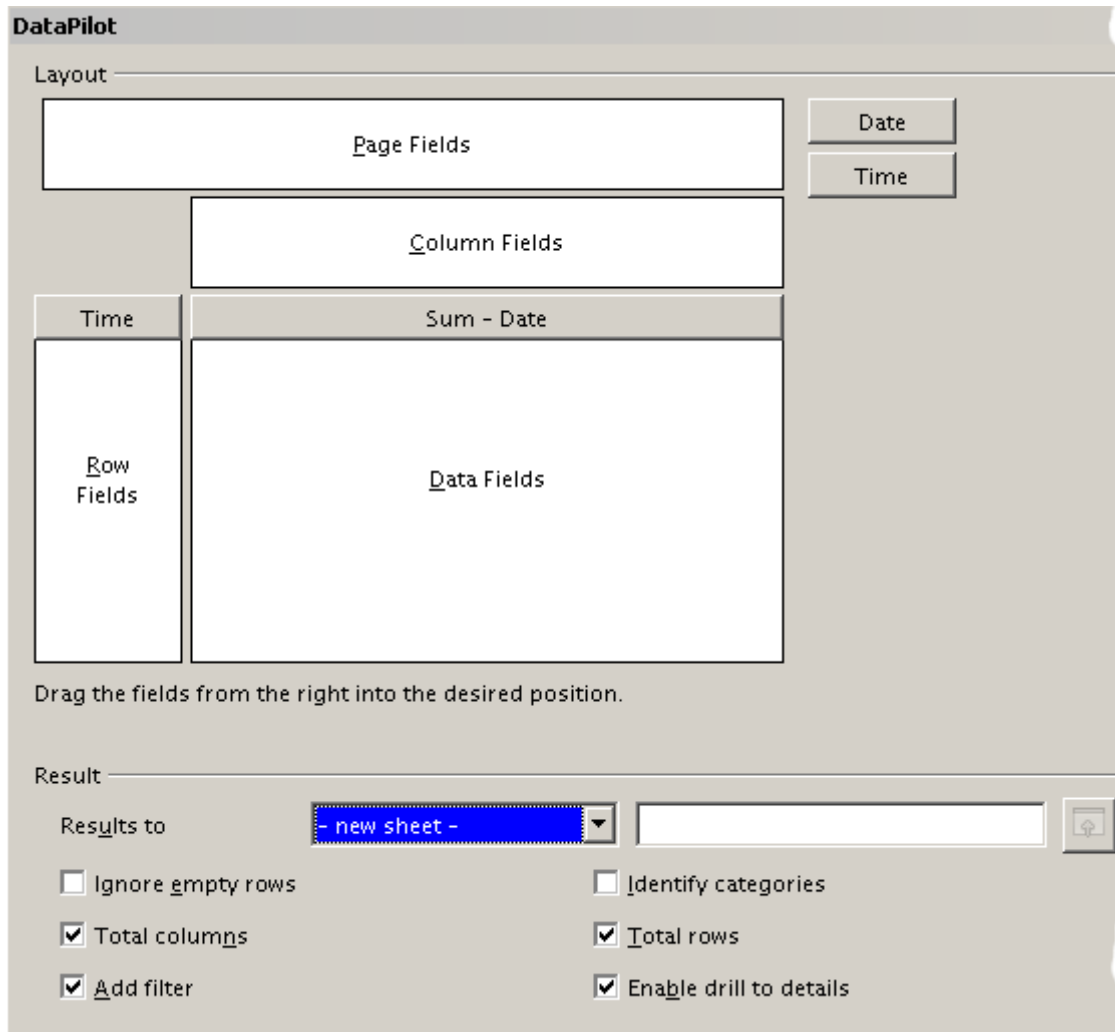


Figure 25: Part of DataPilot dialog

- 6) In this case we need to count the number of values, not their sum (summing dates is rather silly). Double-click on the field **Sum - Date** to open the Data Field dialog and select the function *Count*.

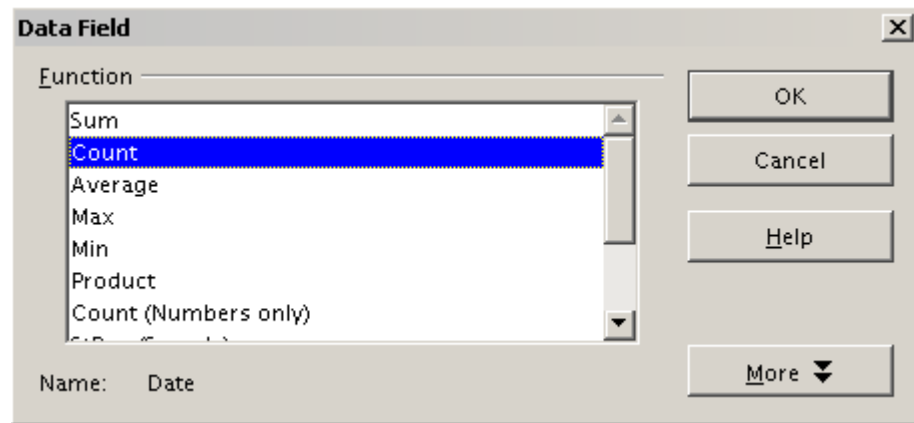


Figure 26: Properties of the data field

- 7) Click **OK**. As an intermediate result, you get a DataPilot table that has for every time within the raw data a separate line.

Note

The time needed for creating the table might be huge because of the large number of items. Please be patient. The time does not depend that much on the number of lines but on the number of rows needed for the table that contains the results.

- 8) For grouping the rows, select the cell **A4** or any other cell, that contains a time.
- 9) Choose **Data > Group and Outline > Group**, select for the interval *Hours* and click **OK**. The result is now grouped according to hours.

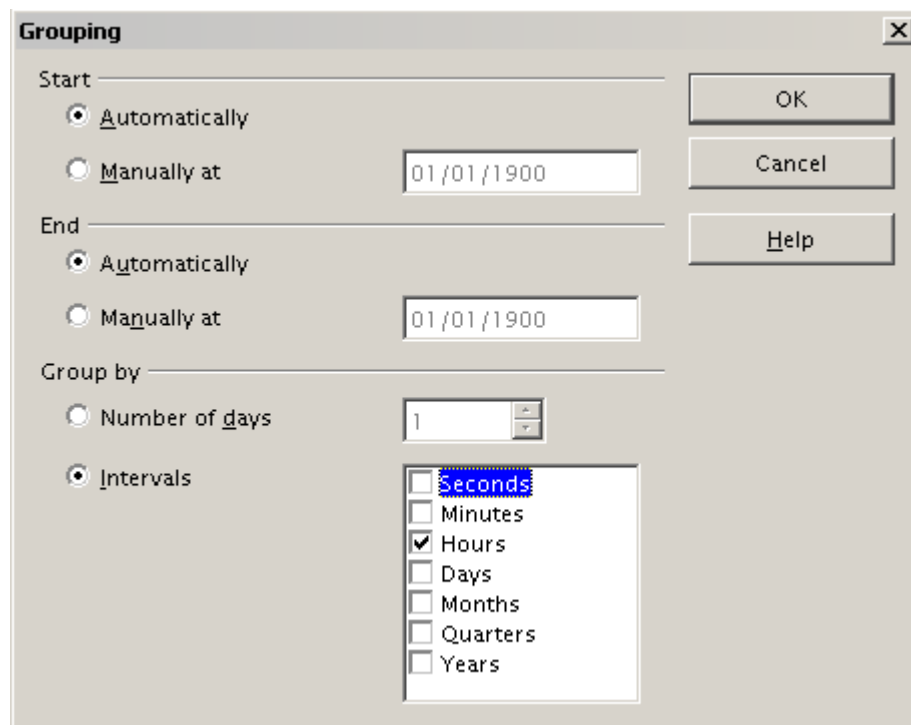


Figure 27: properties for grouping according to hours

- 10) Figure 29 shows the absolute occurrence. Figure 28 shows the Data Field dialog for the data field **Number - Date**. Click **More** and select as type *% of column*. The result is shown in Figure 30. Whether the relative values are shown as a decimal (0.1) or as a percentage (10%) depends only on the cell formatting itself and has nothing to do with the DataPilot.

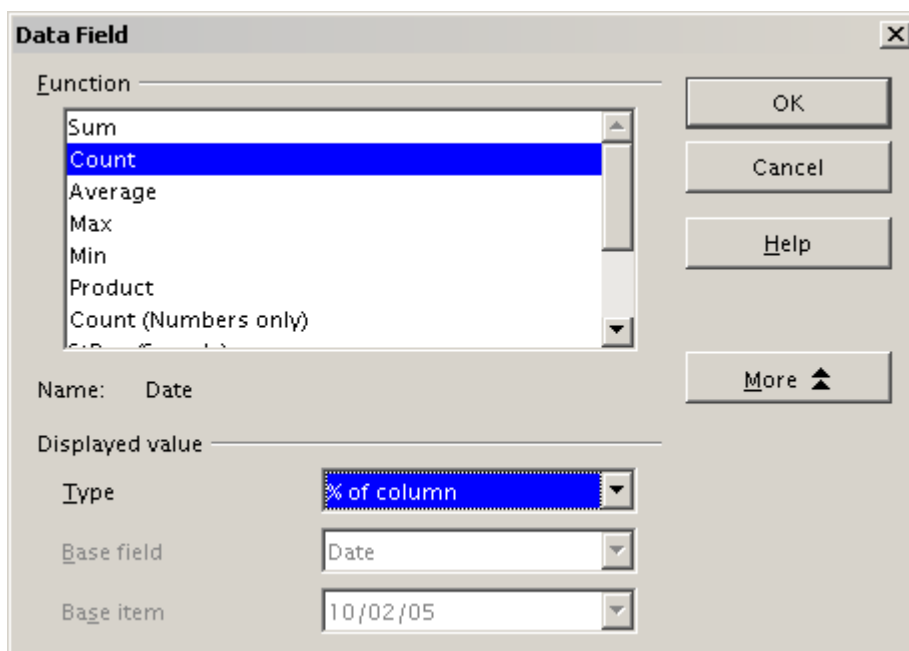


Figure 28: Data Field settings for relative values

	A	B
1	Filter	
2		
3	Time	
4	00	921
5	01	467
6	02	204
7	03	123
8	04	97
9	05	110
10	06	271
11	07	773
12	08	1537
13	09	1945
14	10	2212
15	11	2420
16	12	2382
17	13	2327
18	14	2427
19	15	2525
20	16	2397
21	17	2274
22	18	2259
23	19	2367
24	20	2266
25	21	2130
26	22	2089
27	23	1572
28	Total Result	38095

Figure 29: Frequency distribution with the DataPilot

	A	B
1	Filter	
2		
3	Time	
4	00	2.42%
5	01	1.23%
6	02	0.54%
7	03	0.32%
8	04	0.25%
9	05	0.29%
10	06	0.71%
11	07	2.03%
12	08	4.03%
13	09	5.11%
14	10	5.81%
15	11	6.35%
16	12	6.25%
17	13	6.11%
18	14	6.37%
19	15	6.63%
20	16	6.29%
21	17	5.97%
22	18	5.93%
23	19	6.21%
24	20	5.95%
25	21	5.59%
26	22	5.48%
27	23	4.13%
28	Total Result	100.00%

Figure 30: Relative occurrence

DataPilot functions in detail

This part describes the use and options of the DataPilot in detail.

The database (preconditions)

The basis for work with the DataPilot is always a list with your data in raw format. Such a list is comparable to a database table. The table consists of rows (data sets) and columns (data fields). The field names are in the first row above the list.

Later we will explain that the data source could be an external file or database. The simplest use case is that your data is contained in a Calc spreadsheet. For such a list, Calc offers sorting functions that do not depend on the DataPilot.

For processing data in lists, the program must know in which area of the spreadsheet the table is. A Calc table can be anywhere in the sheet, in any position. It is also possible, that a spreadsheet contains several unrelated tables.

Calc recognizes your lists automatically. It uses the following logic:

Starting from the cell you've selected (which must be within your list), Calc checks the surrounding cells in all 4 directions (left, right, above, below). The border is recognized if the program discovers an empty row or column, or if it hits the left or upper border of the spreadsheet.

This means that the described functions can only work correctly if there are no empty rows or columns in your list. Avoid empty lines (for example for formatting). You can format your list by using cell formats.

Rule No empty rows, no empty columns within lists!

If you select more than one single cell before you start sorting, filtering or calling the DataPilot, then the automatic list recognition is switched off. Calc assumes that the list matches exactly the cells you have selected. This might be useful in only very few cases.

Rule For sorting, filtering or using the DataPilot, always select only one cell.

A relatively big source for errors is that you might declare a list by mistake and you sort your list. If you select multiple cells (for example, a whole column) then the sorting mixes up the data that should be together in one row.

In addition to these formal aspects, the logical structure of your table is very important for using the DataPilot.

Rule	Calc lists must have the <i>normal form</i> ; that is, they must have a simple linear structure.
-------------	--

When entering the data, do not add outlines, groups, or summaries. This becomes clear when we think about what we could have done wrong in our Sales list example in the section “Examples with step by step descriptions” starting at page 5. This will give you a list of bad ideas, that you can find very often among spreadsheet users who are not informed about the possibilities of processing lists within a spreadsheet.

First bad idea: You could have made several sheets. For example, you could have made a sheet for each group of articles. Analyses are then only possible within each group. Analyses for several groups would then be a lot of hassle.

Second bad idea: In the Sales list, instead of only one column for the amount, you could have made a column for the amounts for each employee. The amounts then had to be entered into the appropriate column. An analysis with the DataPilot would not be possible any more. In contrast, one result of the DataPilot is that you can get results for each employee if you have entered everything in one column.

Third bad idea: You could have entered the amounts in chronological order. At the end of each month you could have made a sum. In this case, a sorting of the list for different criteria is not possible, and you could not use the DataPilot. The rows with the sums would be treated by the DataPilot like any other amount you have entered. Getting monthly results is one of the very fast and easy features of the DataPilot.

Start

You start the DataPilot with **Data > DataPilot > Start**. If the list to be analyzed is in a spreadsheet table, select one cell within this list. Calc recognizes and selects the list automatically for use with the DataPilot.

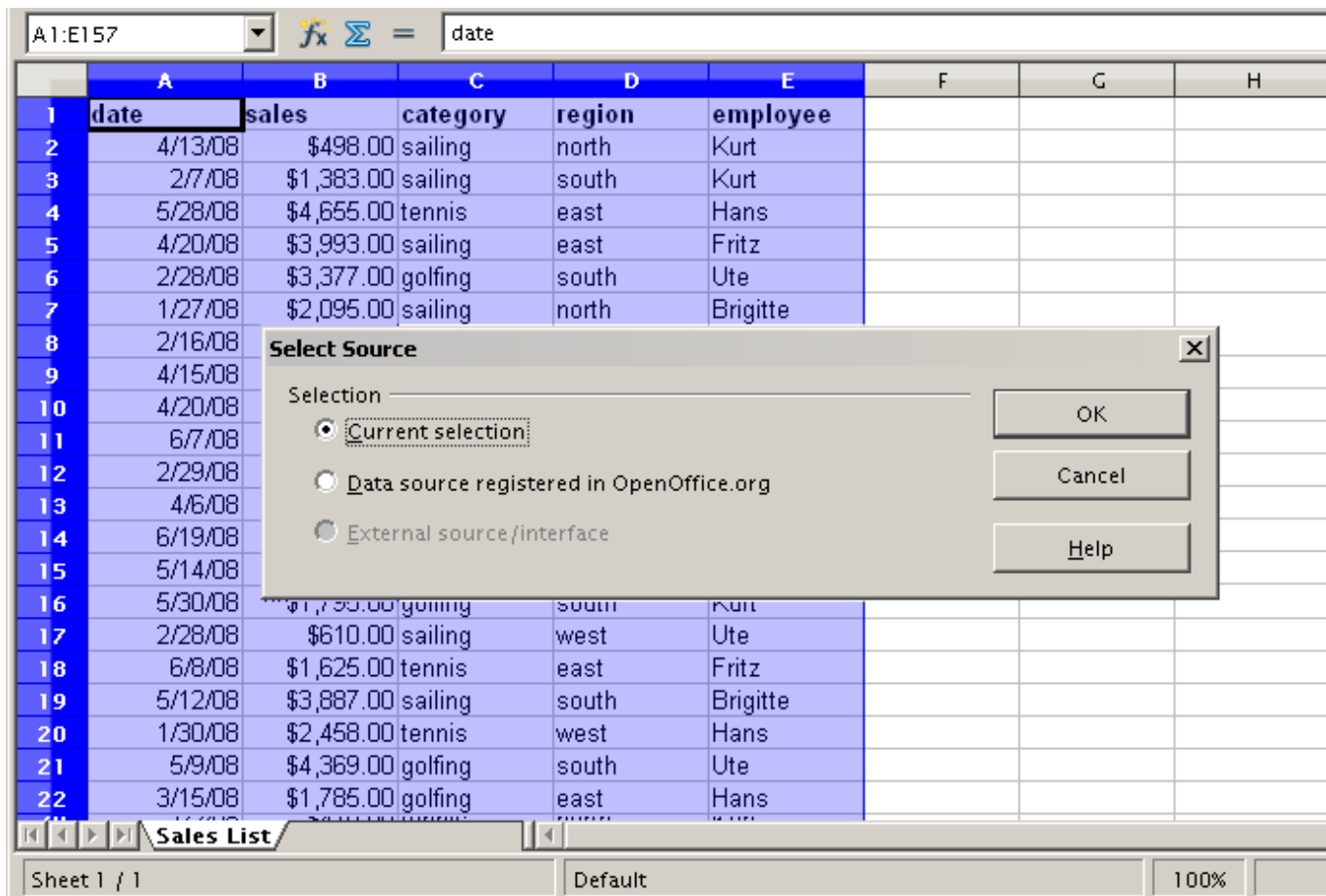


Figure 31: After the start of the DataPilot

Data source

There are usually three possibilities for the DataPilot to have a data source: a Calc sheet, an external data source that has to be registered in OpenOffice.org, and access to an OLAP system (not available at this time).

Calc sheet

The simplest and most often used case is analyzing a list in a Calc sheet. This list might be updated regularly or the data are imported from a different application.

A huge list can, for example, be selected in a different application and with copy and paste inserted into Calc. The behavior of Calc while inserting the data depends on which format the data is pasted to the clip board. If the format fits well, the data is copied into Calc directly. Often the data are in plain text. In this case a dialog appears with the text import assistant that also appears when you open a text file.

Calc can import data from a huge number of foreign data formats, for example from other spreadsheets (Excel, Lotus 1, 2, 3) from databases (like dBase) and from simple text files.

The drawback of copying or importing of foreign data is that it won't update automatically if there are changes in the source file. With a Calc file you are limited to 65535 rows.

Registered data source

A registered data source in OpenOffice.org is a connection to data held in a database outside Calc. This means that the data to be analyzed won't be saved in Calc. It always uses the data from the original source. Calc is able to use many different data sources and also databases that are created and maintained with OpenOffice.org Base. See Chapter 10 (Linking Calc Data) for more about this.

External data (OLAP)

The option *External Source/interface* is grayed out (deactivated) in the selection dialog. This does not mean that a connection to external data is not possible. It is a very specific option for an interface in OpenOffice.org, for which no implementation exists at the moment.

The DataPilot dialog

The function of the DataPilot is mainly managed in two places: first in the dialog of the DataPilot and second through manipulations of the result in the spreadsheet. This section describes the dialog in detail.

Basic layout

In the DataPilot are four big white areas, that indicate schematically the layout of the result. Beside these fields are button-like fields with the names of the fields in your data source. You can choose your layout by moving these fields with the mouse into the white areas.

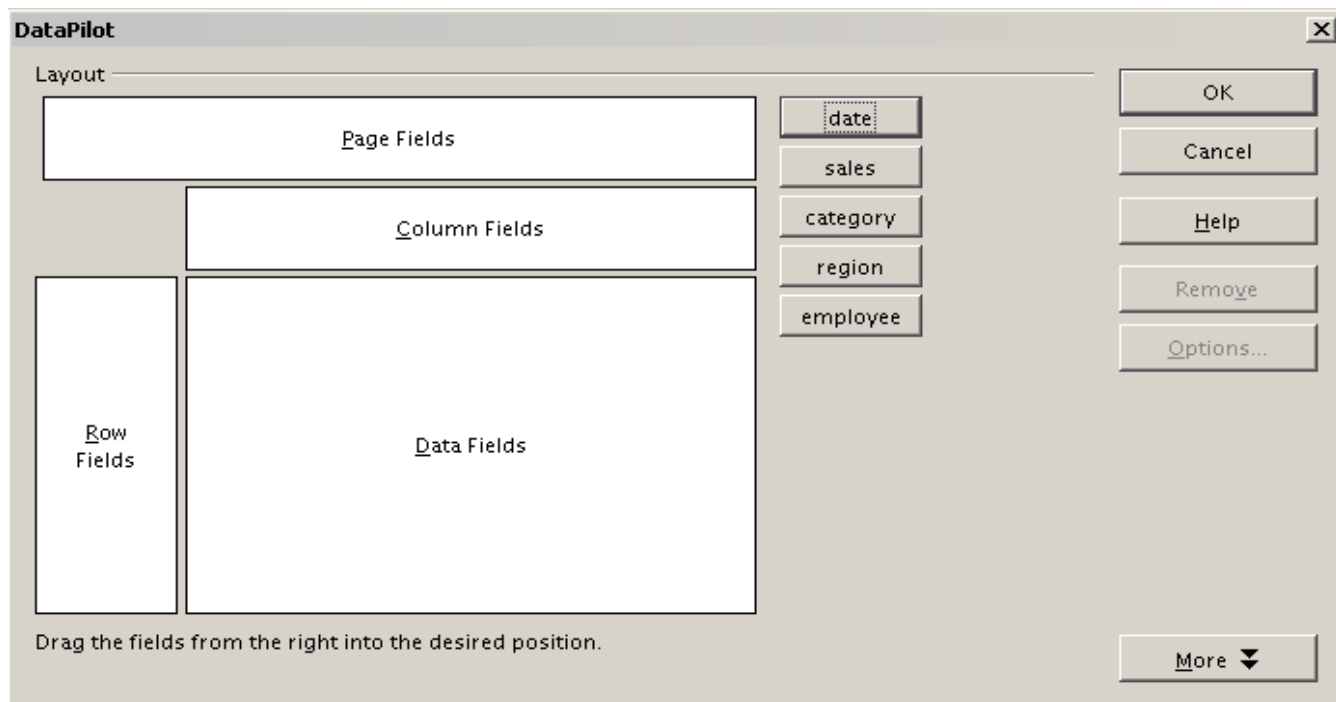


Figure 32: DataPilot dialog

The *Data Fields* area in the middle must contain at least one field. Normally it has exactly one field. Advanced users could use more than one field here. For the data field an aggregate function is used. For example, if you move the field **sales** into the area *Data fields*, it appears there as **Sum - sales**.

Row Fields and *Column Fields* indicate in which groups the result will be sorted in the columns and rows. If there are no entries in the row or column fields, then there won't be partial sums for the corresponding rows or columns. There are often more than one field at the same time used to get partial sums for rows or columns. The order of the fields gives you then the order of the sums from overall to specific.

If you drag for example **region** and **employee** into the area *Row Fields*, the result will be the following: in the vertical direction (row wise), the sum will be divided into the employees. Within the employees will be the listing for the different regions.

Fields that are placed into the area *Page Fields* appear in the result above as a drop down list. The summary in your result takes only that part of your base data into account that you've selected. For example, if you use the field **Employee** as a page field, you can filter the shown results for each employee.

To remove a field from the white layout area just move it with the mouse to the border and let it drop there. The cursor will change to a crossed symbol. As alternative you can use the button **Remove**.

More options

To expand the DataPilot and show more options, click **More**.

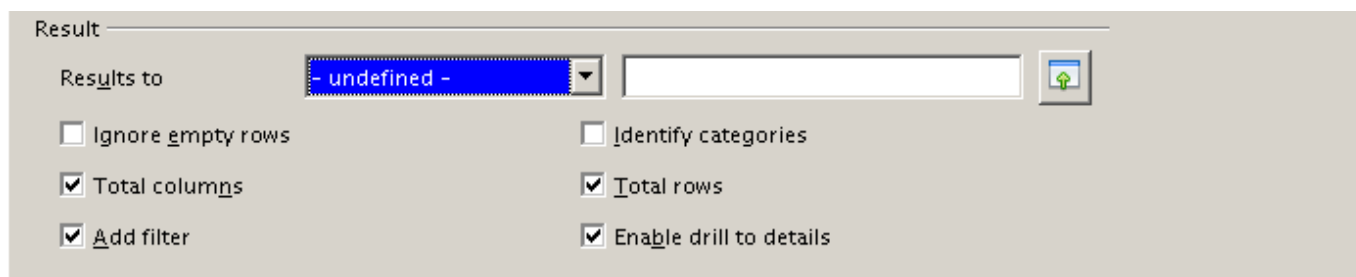


Figure 33: Expanded dialog of the DataPilot

Results to

Results to defines where your result will be shown. If you do not enter anything, the DataPilot will put your result below the list, that contains your data. If you leave - **undefined** - and you enter a cell reference, then the DataPilot will show the results at this place. It's useful instead of using the setting - **undefined** - to use the setting - **new sheet** -. Calc adds in this case a new sheet. This is useful, because it is much cleaner and makes sure that no areas with data are overwritten.¹

Ignore empty rows

By using this setting, completely empty rows in the source area will be ignored by the DataPilot. As mentioned in the requirements for the database, empty rows are against the rules for processing list-like data within a spreadsheet. Therefore, this setting is used on the rare occasions when the source data is not in the expected (recommended) form.

Identify categories

This option also has an effect if the basic data does not meet the recommended data structure. It takes effect when there are missing entries in a list, as shown in this example:

¹ The word - *undefined* - used in this place is wrong and misleading. The output position is also exactly defined. It is somewhat annoying, that the default setting is not -*new sheet* -, which is in almost all cases the needed setting.

	A	B	C
1	Product	Region	Quantity
2	Apples	Italy	6.2 t
3		Lake Constance	19.2 t
4		California	3.6 t
5	Pears	Italy	7.0 t
6		Lake Constance	22.0 t
7			

Figure 34: Example of Identify categories

The option *Identify categories* would make sure that in this case the lines 3 and 4 are added to the product Apples and the row 6 is added to Pears. Without category recognition, the DataPilot shows an empty category.

Sum - Quantity	Region			
Product	California	Italy	Lake Constance	Total Result
(empty)	3.6 t		41.2 t	44.8 t
Apples		6.2 t		6.2 t
Pears		7.0 t		7.0 t
Total Result	3.6 t	13.2 t	41.2 t	58.0 t

Figure 35: Example without Identify categories

From a purely logical point of view, the behavior without category recognition is more correct. A list with the shown missing entries is also less useful, because you cannot use other functions (like sorting or filtering).

Total columns / total rows

With this option you decide if the DataPilot will show an extra row with the sums of each column or if it adds on the very right a column with the sums of each row. In some cases, an added total sum is useless, for example if your entries are accumulated or in comparisons.

Add filter

This option allows the blending or hiding of the cell labelled Filter above the DataPilot results. This cell is a convenient button for additional filtering options within the DataPilot.

Enable drill to details

In the table area of the DataPilot result, a double-click has a special effect. Normally a double-click on any cell switches to the edit mode, but within the DataPilot a listing of the detailed data that give you the shown result will appear. By disabling this function the double-click will keep its usual function within a spreadsheet.

More settings for the fields

The discussed options are valid for the DataPilot in general. Additionally, you can change settings for every field, that you have added to the DataPilot layout. You can reach the dialog for the field settings either by clicking on the button **Options** in the DataPilot dialog box or by a double-click on the appropriate field.

There are differences between data fields, row or column fields, and page fields of the DataPilot.

Data fields

In the preferences dialog of a data field you can select the Sum function for accumulating the values from your data source. In many cases you will need the sum function, but other functions (like standard distribution or a counting function) are also available. For example, the counting function can be useful for non-numerical data fields.

On the Data Field dialog, click **More** to see more options.

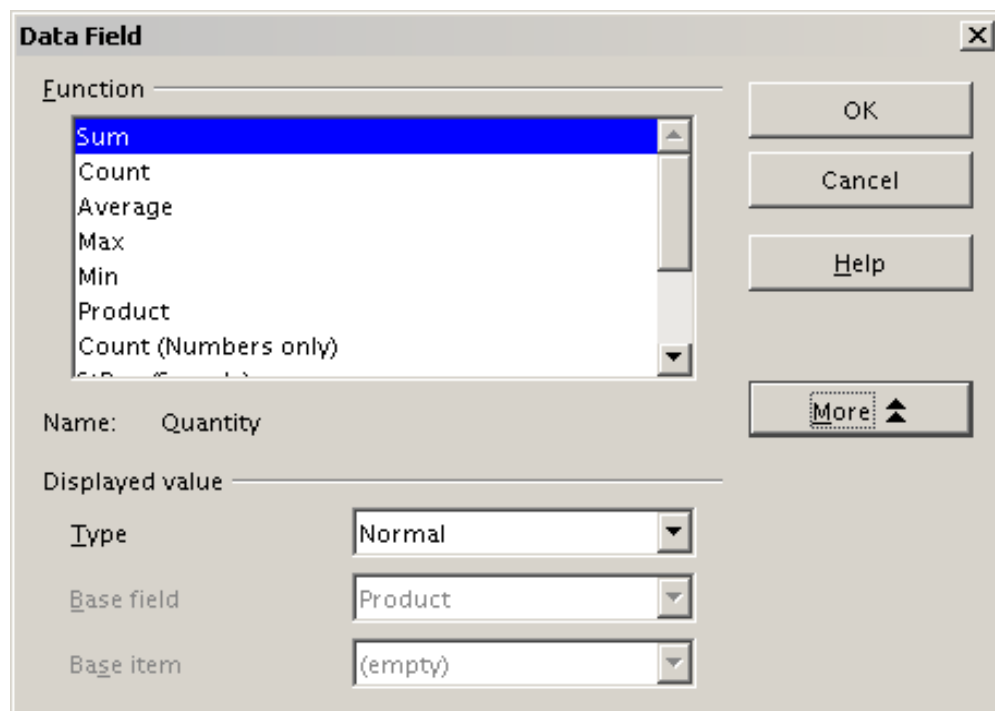


Figure 36: Expanded dialog for a data field

In the *Displayed value* section, you can choose more analyzing possibilities by using the aggregate function. Depending on the setting for **Type**, you have to choose a definition for **Base field** and **Base element**.

Type	Base field	Base element	Analysis
Normal	—	—	Simple use of the chosen aggregate function (for example, sum)
Difference from	Selection of a field from the data source of the DataPilot (for example, employee)	Selection of an element from the selected base field (for example, Brigitte)	Result as difference to the result of the base element (for example, Sales volume of the employees as difference of the sales volume of Brigitte)
% of	Selection of a field from the data source of the DataPilot (for example, employee)	Selection of an element from the selected base field (for example, Brigitte)	Result as a ratio based on the result of the base element (for example, Sales result of the employee relative to the sales result of Brigitte)
% difference from	Selection of a field from the data source of the DataPilot (for example, employee)	Selection of an element from the selected base field (for example, Brigitte)	Result as relative difference to the result of the base element (for example, Sales volume of the employees as relative difference of the sales volume of Brigitte)
Running total in	Selection of a field from the data source of the DataPilot (for example, date)	—	Result as a continuing sum (for example, Continuing sum of the sales volume for days or months)
% of row	—	—	Result as relative part of the result in the whole row (for example the row sum)

Type	Base field	Base element	Analysis
% of column	—	—	Result as relative part of the total column (for example, the column sum)
% of total	—	—	Result as relative part of the overall result (for example the total sum)
Index	—	—	Default result x total result / (row result x column result)

Row and column fields

In the preferences dialog for the row or column fields, you can choose if there are partial sums for each category. This is deactivated by default. It is only useful if the values in one row or column field can be divided into partial sums for another (sub)field.

Some examples are shown in the next three figures.

Sum - sales	category			
region	golfing	sailing	tennis	Total Result
east	\$41,971.00	\$22,484.00	\$35,966.00	\$100,421.00
north	\$18,741.00	\$22,468.00	\$34,533.00	\$75,742.00
south	\$56,257.00	\$44,801.00	\$34,258.00	\$135,316.00
west	\$39,245.00	\$20,099.00	\$37,942.00	\$97,286.00
Total Result	\$156,214.00	\$109,852.00	\$142,699.00	\$408,765.00

Figure 37: No subdivision with only one row or column field.

Sum - sales		category			
region	employee	golfing	sailing	tennis	Total Result
east	Brigitte	\$5,822.00	\$2,135.00	\$4,872.00	\$12,829.00
	Fritz	\$15,172.00	\$5,730.00	\$12,455.00	\$33,357.00
	Hans	\$5,316.00	\$909.00	\$12,220.00	\$18,445.00
	Kurt	\$9,707.00	\$6,475.00	\$2,417.00	\$18,599.00
	Ute	\$5,954.00	\$7,235.00	\$4,002.00	\$17,191.00
north	Brigitte	\$3,814.00	\$10,151.00	\$3,985.00	\$17,950.00
	Fritz	\$3,443.00	\$2,698.00	\$9,115.00	\$15,256.00
	Hans	\$3,049.00	\$3,008.00	\$5,361.00	\$11,418.00
	Kurt	\$2,214.00	\$3,485.00	\$10,499.00	\$16,198.00
	Ute	\$6,221.00	\$3,126.00	\$5,573.00	\$14,920.00
south	Brigitte	\$5,151.00	\$4,432.00		\$9,583.00
	Fritz	\$23,290.00	\$4,806.00	\$15,641.00	\$43,737.00
	Hans	\$4,196.00	\$9,263.00	\$3,858.00	\$17,317.00
	Kurt	\$11,681.00	\$7,270.00	\$14,759.00	\$33,710.00
	Ute	\$11,939.00	\$19,030.00		\$30,969.00
west	Brigitte	\$12,174.00	\$7,704.00	\$8,864.00	\$28,742.00
	Fritz	\$4,934.00	\$6,742.00	\$1,427.00	\$13,103.00
	Hans	\$5,380.00	\$880.00	\$9,028.00	\$15,288.00
	Kurt	\$4,744.00	\$3,584.00		\$8,328.00
	Ute	\$12,013.00	\$1,189.00	\$18,623.00	\$31,825.00
Total Result		\$156,214.00	\$109,852.00	\$142,699.00	\$408,765.00

Figure 38: Division of the regions for employees without partial sums

Sum - sales		category			
region	employee	golfing	sailing	tennis	Total Result
east	Brigitte	\$5,822.00	\$2,135.00	\$4,872.00	\$12,829.00
	Fritz	\$15,172.00	\$5,730.00	\$12,455.00	\$33,357.00
	Hans	\$5,316.00	\$909.00	\$12,220.00	\$18,445.00
	Kurt	\$9,707.00	\$6,475.00	\$2,417.00	\$18,599.00
	Ute	\$5,954.00	\$7,235.00	\$4,002.00	\$17,191.00
east Sum - sales		\$41,971.00	\$22,484.00	\$35,966.00	\$100,421.00
north	Brigitte	\$3,814.00	\$10,151.00	\$3,985.00	\$17,950.00
	Fritz	\$3,443.00	\$2,698.00	\$9,115.00	\$15,256.00
	Hans	\$3,049.00	\$3,008.00	\$5,361.00	\$11,418.00
	Kurt	\$2,214.00	\$3,485.00	\$10,499.00	\$16,198.00
	Ute	\$6,221.00	\$3,126.00	\$5,573.00	\$14,920.00
north Sum - sales		\$18,741.00	\$22,468.00	\$34,533.00	\$75,742.00
south	Brigitte	\$5,151.00	\$4,432.00		\$9,583.00
	Fritz	\$23,290.00	\$4,806.00	\$15,641.00	\$43,737.00
	Hans	\$4,196.00	\$9,263.00	\$3,858.00	\$17,317.00
	Kurt	\$11,681.00	\$7,270.00	\$14,759.00	\$33,710.00
	Ute	\$11,939.00	\$19,030.00		\$30,969.00
south Sum - sales		\$56,257.00	\$44,801.00	\$34,258.00	\$135,316.00
west	Brigitte	\$12,174.00	\$7,704.00	\$8,864.00	\$28,742.00
	Fritz	\$4,934.00	\$6,742.00	\$1,427.00	\$13,103.00
	Hans	\$5,380.00	\$880.00	\$9,028.00	\$15,288.00
	Kurt	\$4,744.00	\$3,584.00		\$8,328.00
	Ute	\$12,013.00	\$1,189.00	\$18,623.00	\$31,825.00
west Sum - sales		\$39,245.00	\$20,099.00	\$37,942.00	\$97,286.00
Total Result		\$156,214.00	\$109,852.00	\$142,699.00	\$408,765.00

Figure 39: Division of the regions for employees with partial sums

Choose the option **Automatically** to use the aggregate function for the partial results that can also be used for the data fields (see above). To set up the aggregate function for the partial results independently from the overall settings of the DataPilot, choose **User-defined**.

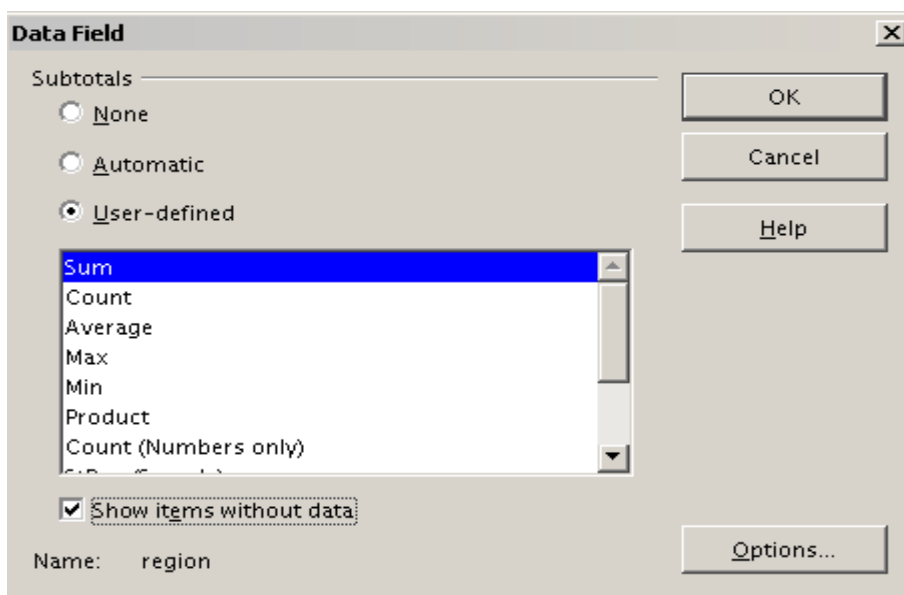


Figure 40: Preferences dialog of a row or column field

Normally, the DataPilot does not show a row or column for categories that have no entries in the underlying database. With the help of the **Show items with no data** option, you can force this.

For illustration purposes, the data was manipulated in such a way, that the employee Brigitte has no sales values for the category golfing.

employee	Brigitte			
Sum - sales	category			
region	sailing	tennis	Total Result	
east	\$2,135.00	\$4,872.00	\$7,007.00	
north	\$10,151.00	\$3,985.00	\$14,136.00	
south	\$4,432.00		\$4,432.00	
west	\$7,704.00	\$8,864.00	\$16,568.00	
Total Result	\$24,422.00	\$17,721.00	\$42,143.00	

Figure 41: Default setting

employee	Brigitte			
Sum - sales	category			
region	golfing	sailing	tennis	Total Result
east		\$2,135.00	\$4,872.00	\$7,007.00
north		\$10,151.00	\$3,985.00	\$14,136.00
south		\$4,432.00		\$4,432.00
west		\$7,704.00	\$8,864.00	\$16,568.00
Total Result		\$24,422.00	\$17,721.00	\$42,143.00

Figure 42: Setting "Show Items with no data"

Page fields

The preferences dialog for page fields is the same as for row and column fields, even though it appears to be useless to do the same settings as described for the row and column fields. With the flexible use of the DataPilot you can switch often the use of the different fields for pages, columns or rows. The fields keep the settings that you made for them. For this, the page field has the same properties as a row or column field. These settings only take effect when you use the field not as page field but as row or column field.

Working with the results of the DataPilot

One very important feature is the flexibility of the DataPilot. An analysis can be changed with only a few mouse clicks. Some functions of the DataPilot can be used only for the results of an analysis.

Start the dialog

Right-click in the area of the resulting table of the DataPilot. The command **Start** opens the DataPilot dialog with all current settings.

Change layout by using drag and drop

The easiest and fastest method to change the layout of the DataPilot is drag and drop. Within the result table of the DataPilot just move one of the page, column, or row fields in a different position. The section “Examples with step by step descriptions” starting on page 5 contains some examples.

You can remove a column, row, or page field from the DataPilot simply by dragging it out of the DataPilot with the mouse.

Grouping rows or columns

For many analyses or summaries, the categories have to be grouped. You can merge the results in classes or periods. In the DataPilot you do a grouping after you have first made an ungrouped DataPilot table.

You can access the grouping with the menu entry **Data > Group and Outline > Group** or by pressing *F12*. It is important that you select the correct cell area. The grouping function is mostly dominated by the kind of values that have to be grouped. You have to distinguish if you have scalar values, date or time values or other values like text that you want to group.

Note

The time needed for creating a DataPilot depends mostly on the number of columns and rows (and not of the size of the basic data). Through grouping you can produce the DataPilot with a small number of rows and columns. Before you can group you have to produce a DataPilot with ungrouped data as an intermediate step. This DataPilot can contain a lot of categories, depending on your data source. Please expect in such cases a computing time of several minutes!

Grouping of categories with scalar values

For grouping scalar values, select a single cell in the row or column of the category to be grouped.

km/h	
30	2
31	1
32	5
33	4
34	5
35	2
36	2
37	1
38	6
39	1
40	1
41	4
42	2
43	3

Figure 43: DataPilot without grouping (frequency of the km/h values of a radar control)

km/h	
30-39	29
40-49	22
50-59	23
60-69	23
70-79	18
80-89	23
90-99	18
100-109	27
110-119	35
Total Result	218

Figure 44: DataPilot with grouping (classes of 10 km/h each)

With the menu entry **Data > Group and Outline > Group** or by pressing *F12*, you get the following dialog.

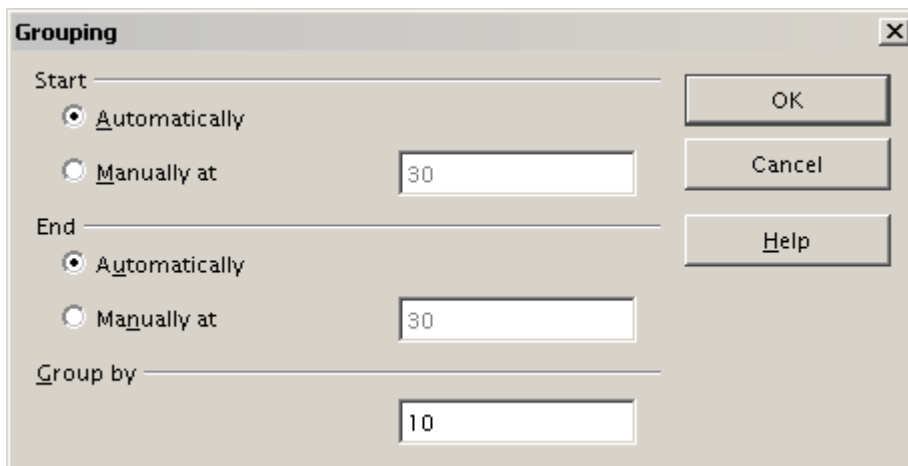


Figure 45: Grouping dialog with scalar categories

You can define in which value range (start/end) the grouping should take place. The default setting is the whole range from smallest to biggest value. In the field *Grouping after* you can enter the class size, this means the interval size (in this example groups of 10 km/h each).

Grouping of categories with date or time values

For grouping date or time values select a single cell in the column or row of the category that should be grouped. This was demonstrated in all three examples in the section “Examples with step by step descriptions” starting on page 5.

With the menu entry **Data > Group and Outline > Group** or by pressing *F12*, you get the following dialog.

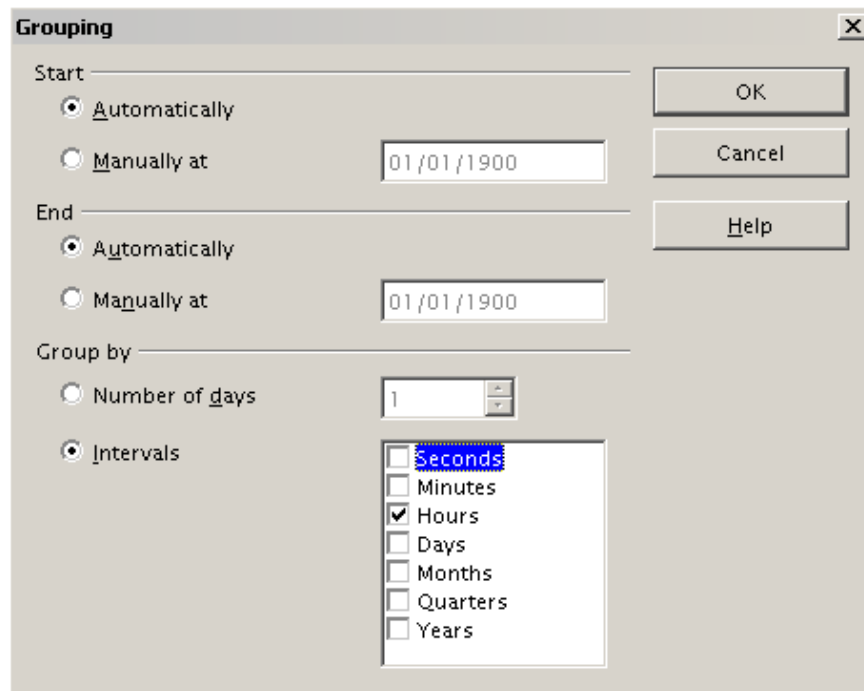


Figure 46: Grouping dialog for categories with dates or times

You can decide in which date or time area (start/end) the grouping should take place. The default setting is the whole area from the earliest to the latest value. In the field *Group by* you can enter the class size (the interval size), that should be used for grouping.

Possible intervals are: seconds, minutes, hours, days, months, quarters and years. These can be combined, for example grouping of years and within the years according to months.

As alternative you can enter any number of days as grouping interval.

Tip For grouping the DataPilot in calendar weeks, choose the beginning on a Sunday or Monday and enter manually the grouping interval of 7 days.

Grouping without the automatic creation of intervals

If the categories do not contain scalar values or date and time entries but instead text fields, then the automatic creation of intervals isn't possible. You can define for each category of any data type which categories you want to put together in one group.

Every time you use the menu entry **Data > Group and Outline > Group** or you press *F12* and you have more than one cell selected, then all the cells will be selected as one group.

Last name	First name	Department	Sick days
Meier	Hans	Sales	7
Muller	Karin	Accounting	7
Schuster	Josef	Purchasing	3
Huber	Erna	Purchasing	3
Aigner	Hermann	Production	7
Schulze	Josef	Production	7
Schroder	Gerhard	Production	4
Forster	Inge	Assembly	4
Meier	Gunter	Assembly	1
Gabriel	Juri	Warehouse	0
Schumacher	Helmut	Warehouse	5

Figure 47: Database with nonscalar categories (departments)

Department	
Accounting	7
Assembly	5
Production	18
Purchasing	6
Sales	7
Warehouse	5
Total Result	48

Figure 48: DataPilot with nonscalar categories

For grouping of non scalar categories, select in the result of the DataPilot all single categories that you want to put in one group.

Tip You can select several non-contiguous cells in one step by pressing and holding the *Control* key while left-clicking with the mouse.

Choose the menu entry **Data > Group and Outline > Group** or press *F12*. Repeat this for all groups that you want to create from the different categories.

Department2	Department	
Assembly	Assembly	5
Group1	Accounting	7
	Purchasing	6
	Sales	7
Production	Production	18
Warehouse	Warehouse	5
Total Result		48

Figure 49: Summary of single categories in one group

Department3	Department	
Group1	Accounting	7
	Purchasing	6
	Sales	7
Group2	Assembly	5
	Production	18
	Warehouse	5
Total Result		48

Figure 50: Grouping finished

You can change the automatically given names for the groups and the newly created group field to any name you like. The DataPilot will remember these settings, even if you change the layout later on. For the following pictures, the dialog was called again (with a right-click) and within the preferences menu the point Automatic was activated.

Department3	Department2	Department	
Group1	Group1	Accounting	7
		Purchasing	6
		Sales	7
Technical	Assembly	Assembly	5
	Production	Production	18
	Warehouse	Warehouse	5
Total Result			48

Figure 51: Renamed groups and partial results

Department3	
Group1	20
Group2	28
Total Result	48

Figure 52: Reduced to the new groups

Note

A good structured database makes manual sorting within the DataPilot obsolete. In the shown example you could add another column with the name Department, that has for each person the correct entry based on whether the employee's department belongs to the group Office or Technical. The mapping for this (1:n relationship) can be done easily with the VLOOKUP function from Calc.

Sorting of the results

The result of any DataPilot is sorted (categories) in columns and rows in an ascendent way. You have two possibilities, to change the sorting:

- Sort manually by using drag and drop.
- Sort automatically through choosing the options in the preferences dialog of the row or column field.

Sort manually by using drag and drop

Change the order within the categories just by moving the cells with the category values in the result table of the DataPilot.

Please be aware that in Calc a cell has to be selected. It is not enough that this cell contains the cell cursor. The background of a selected cell is marked with a different colour. To achieve this, click in one cell with no extra key pressed and redo this by pressing additionally the *Shift* or *Ctrl* key. Another possibility is to keep the mouse button pressed on the cell you want to select, move the mouse to a neighbor cell and move back to your original cell before you release the mouse button.

Automatic sorting

To manage automatic sorting, start the options of the preferences of the row or column field: right-click on the table area with the DataPilot result and choose **Start**. Open the preferences of the row or column field with a double-click onto the field. In the Data Field dialog, click **Options** to display the Data Field Options dialog (Figure 53).

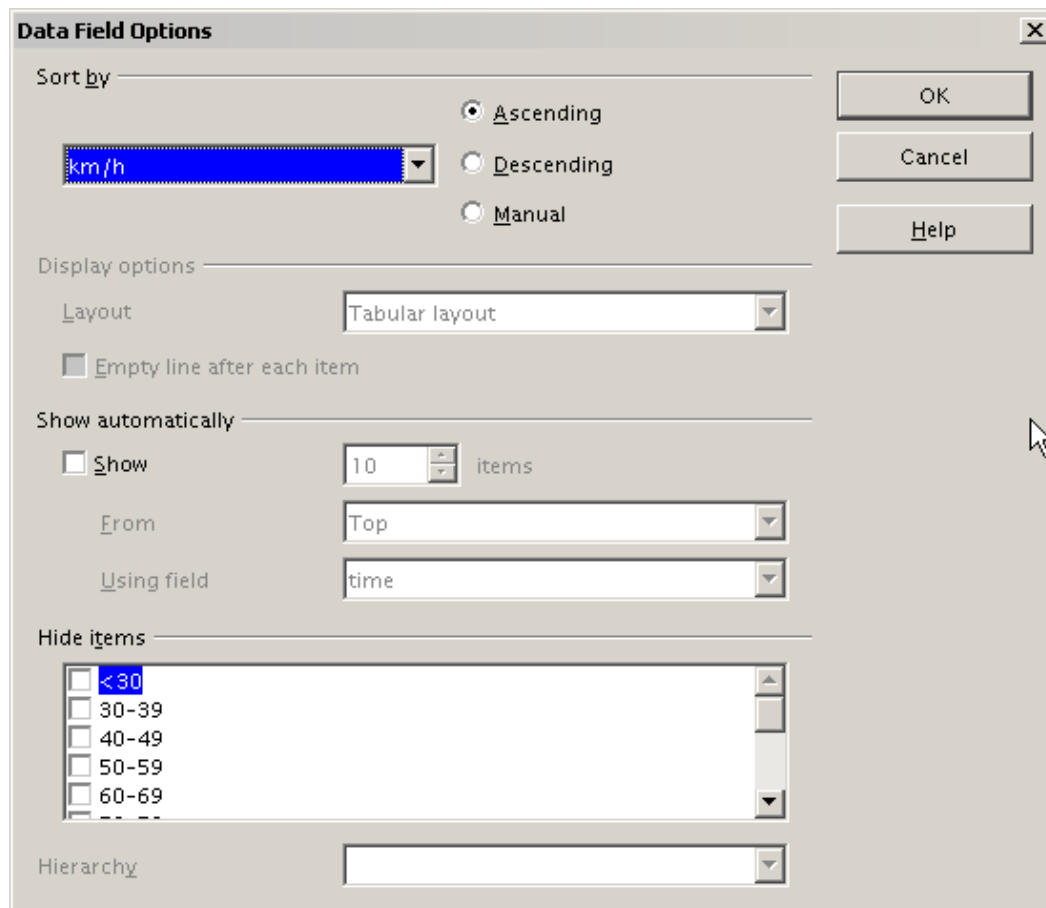


Figure 53: Options for a row or column field

For *Sort by* choose either *Ascending* or *Descending*. On the left side is a drop-down list where you can choose for which field this setting should apply. With this method you can specify that sorting does not happen according to the categories but according to the results of the data field.

Drill (show details)

With a drill you can show for the single, compressed values in the DataPilot result the related detailed data. You can activate a drill through a double-click on the cell or use **Data > Group and Outline > Show Details**. You have to distinguish two cases:

- 1) The active cell is the category of a row or column field.

In this case drill means an additional breakdown into the categories of another field.

For example, double-click on the cell with the value *golfing* in the row field **region**. In this case the values that are aggregated in the category *golfing* are subdivided according to another field.

Since there are more possibilities for a subdivision, a dialog appears where you can make your needed settings.

Sum - sales	region				
category	east	north	south	west	Total Result
golfing	\$41,971.00	\$18,741.00	\$56,257.00	\$39,245.00	\$156,214.00
sailing	\$22,484.00	\$22,468.00	\$44,801.00	\$20,099.00	\$109,852.00
tennis	\$35,966.00	\$34,533.00	\$34,258.00	\$37,942.00	\$142,699.00
Total Result	\$100,421.00	\$75,742.00	\$135,316.00	\$97,286.00	\$408,765.00

Figure 54: Before the drill down for the category golfing

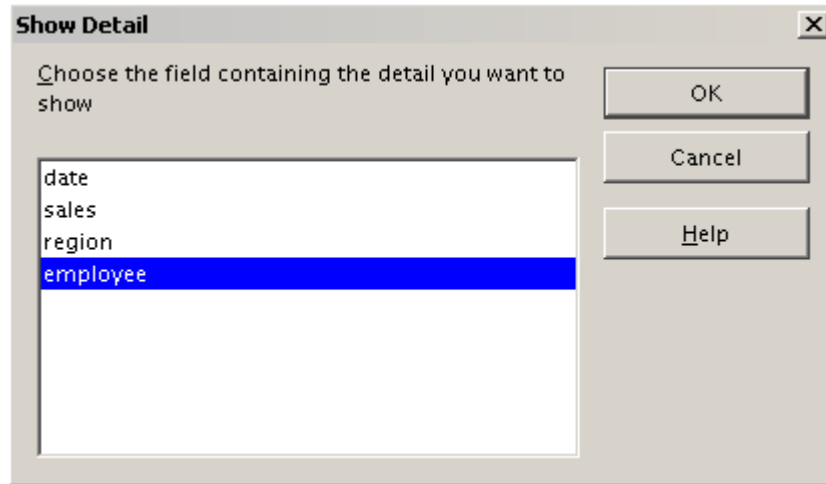


Figure 55: Selection of the field for the subdivision

Sum - sales		region				
category	employee	east	north	south	west	Total Result
golfing	Brigitte	\$5,822.00	\$3,814.00	\$5,151.00	\$12,174.00	\$26,961.00
	Fritz	\$15,172.00	\$3,443.00	\$23,290.00	\$4,934.00	\$46,839.00
	Hans	\$5,316.00	\$3,049.00	\$4,196.00	\$5,380.00	\$17,941.00
	Kurt	\$9,707.00	\$2,214.00	\$11,681.00	\$4,744.00	\$28,346.00
	Ute	\$5,954.00	\$6,221.00	\$11,939.00	\$12,013.00	\$36,127.00
sailing		\$22,484.00	\$22,468.00	\$44,801.00	\$20,099.00	\$109,852.00
tennis		\$35,966.00	\$34,533.00	\$34,258.00	\$37,942.00	\$142,699.00
Total Result		\$100,421.00	\$75,742.00	\$135,316.00	\$97,286.00	\$408,765.00

Figure 56: After the drill down

To hide the details again, double-click on the cell *golfing* or choose **Data > Group and Outline > Hide Details**.

The DataPilot remembers your selection (in our example the field **employee**), so that for the next drill down for a category in the field **region** the dialog does not appear. To remove the selection **employee**, open the DataPilot dialog by right-clicking and choosing **Start**, then delete the unwanted selection in the row or column field.

- 2) The active cell is a value of the data field.

In this case drilldown means a listing of all data entries of the data source that aggregates to this value.

Double-click on the cell with the value \$18,741 from Figure 54 and you'll get a list of all data sets that are included in this value. This list is shown in a new sheet.

	A	B	C	D	E
1	date	sales	category	region	employee
2	2/6/08	\$3,443.00	golfing	north	Fritz
3	3/18/08	\$3,814.00	golfing	north	Brigitte
4	1/17/08	\$4,842.00	golfing	north	Ute
5	6/28/08	\$3,049.00	golfing	north	Hans
6	3/6/08	\$1,379.00	golfing	north	Ute
7	5/30/08	\$2,214.00	golfing	north	Kurt
8					

Figure 57: New table sheet after the drilldown for a value in a data field.

Filtering

For limiting the analysis with the DataPilot to a subset of the information that is contained in the data basis, you can filter with the DataPilot.

Note

An Autofilter or default filter used on the sheet has no effect for the analysis of the DataPilot. The DataPilot always uses the complete list that was selected when it was started.

To use the filtering function that is implemented in the DataPilot, just click on the field **Filter** on the left side above the results.

	A	B	C	D	E	F	G
1	Filter						
2							
3	Sum - sales	employee					
4	region	Brigitte	Fritz	Hans	Kurt	Ute	Total Result
5	east	\$12,829.00	\$33,357.00	\$18,445.00	\$18,599.00	\$17,191.00	\$100,421.00
6	north	\$17,950.00	\$15,256.00	\$11,418.00	\$16,198.00	\$14,920.00	\$75,742.00
7	south	\$9,583.00	\$43,737.00	\$17,317.00	\$33,710.00	\$30,969.00	\$135,316.00
8	west	\$28,742.00	\$13,103.00	\$15,288.00	\$8,328.00	\$31,825.00	\$97,286.00
9	Total Result	\$69,104.00	\$105,453.00	\$62,468.00	\$76,835.00	\$94,905.00	\$408,765.00

Figure 58: Filter field in the upper left area of the DataPilot

In the Filter dialog, you can define of up to 3 filter options that are used in the same way as Calc's default filter.

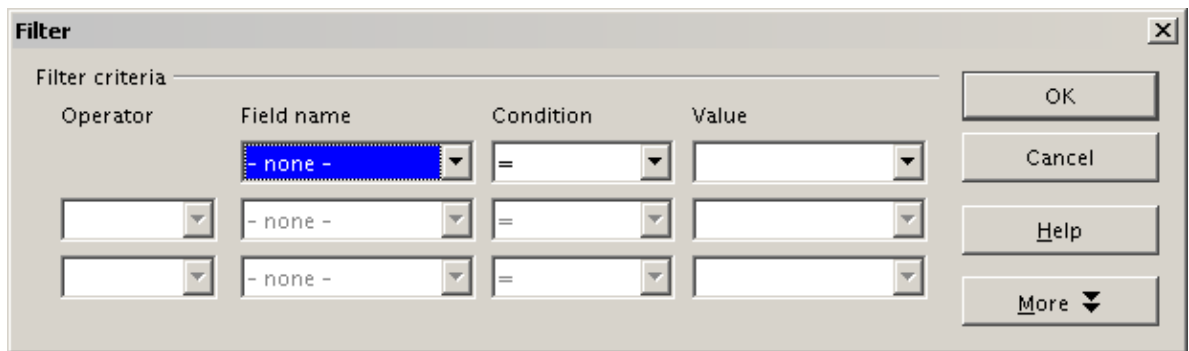


Figure 59: Dialog for defining the filter

Note Even if they are not called a filter, page fields are a practical way to filter the results. The advantage is that the filtering criteria used are clearly visible.

Updating (refreshing) changed values

After you have created the DataPilot, changes in the source data do not cause an automatic update in the resulting table. You have to always update (refresh) the DataPilot manually.

Changes in the source data could appear in two ways:

- 1) The content of existing data sets has been changed.
For example, you might have changed a sales value afterwards. To update the DataPilot, right-click in the result area and choose **Update** (or choose **Data > DataPilot > Refresh** from the menu bar).
- 2) You have added or deleted data sets in the original list.
In this case the change means that the DataPilot has to use a different area of the spreadsheet for its analysis. A change of the data reference for an existing DataPilot is not possible. In this case you have to redo the DataPilot.

Cell formatting

The cells in the results area of the DataPilot are automatically formatted with a simple format by Calc. You can use all the formatting options that are available for Calc if you wish them to have a different appearance. However, every change in the design of the DataPilot and each update changes your formatting back to the simple default formatting.

For the number format in the data field, Calc uses the number format that is used in the corresponding cell in the source list. In most cases, this is useful. For example, if the values are in the currency format, then the corresponding cell in the result area is also formatted as

currency. The DataPilot does not recognize that this might be useless, if the result is a fraction or a percentage because of the options used. Such results have to be without a unit or have to be displayed as a percentage. You can correct the number format manually, but the correction stays only until the next update.

Multiple data fields

Until now we have assumed that the layout of the DataPilot contains only one data field. However, it is possible to have several data fields in the middle of the layout. This makes summaries and analyses of multiple aspects possible.

You could, for example, list all the sales values per day and additionally give the number of entries per day.

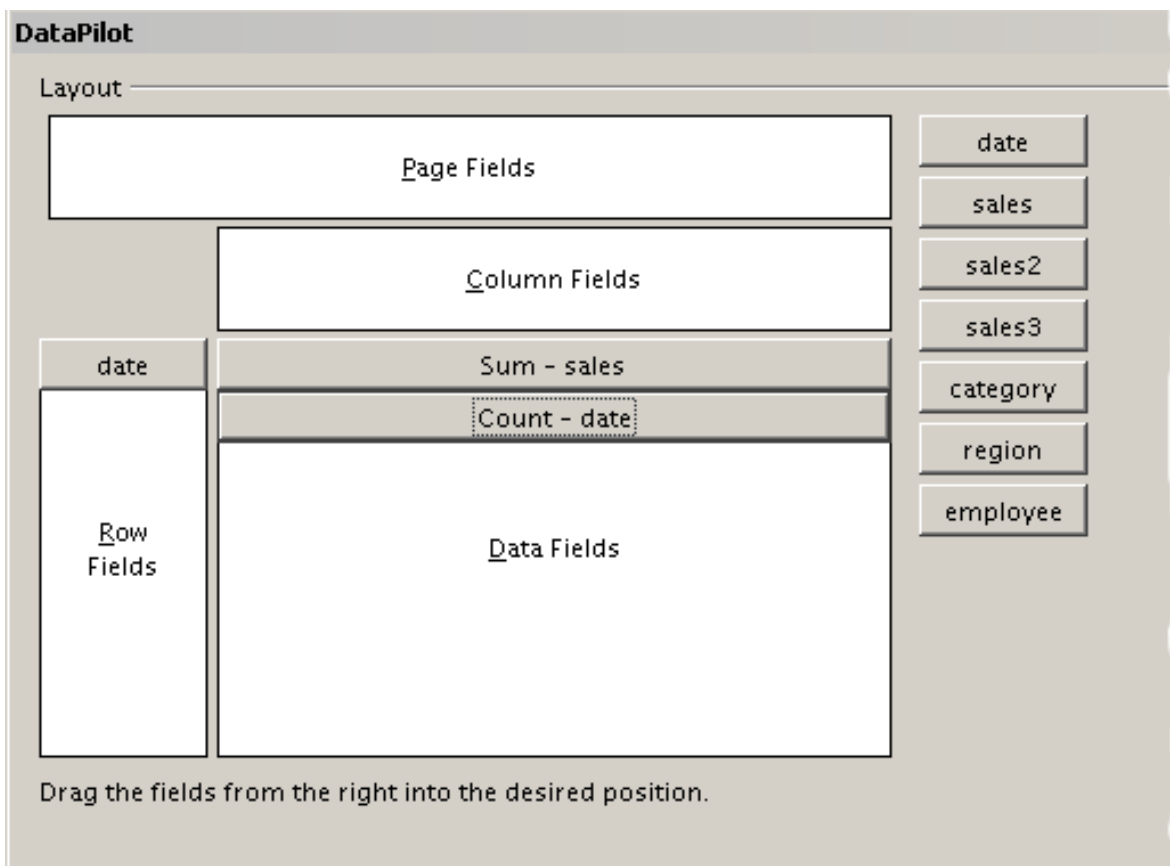


Figure 60: Multiple data fields in the DataPilot

For this, in addition to the **sales** field, add the **date** field into the Data Fields area. For the **date** field, choose the Count option for the aggregate function (Figure 60).

Since every entry has a specific date, this field will give you the number of entries for each date. If you group the values per month,

you get an overview with the sales value and the number of closed sales for each category and month (Figure 61).

		category			
date	Data	golfing	sailing	tennis	Total Result
Jan	Sum - sales	\$26,180.00	\$13,979.00	\$39,206.00	\$79,365.00
	Count - date	9	8	13	30
Feb	Sum - sales	\$30,444.00	\$15,625.00	\$23,710.00	\$69,779.00
	Count - date	11	8	10	29
Mar	Sum - sales	\$31,714.00	\$17,409.00	\$12,097.00	\$61,220.00
	Count - date	10	7	5	22
Apr	Sum - sales	\$24,747.00	\$19,769.00	\$31,918.00	\$76,434.00
	Count - date	8	8	9	25
May	Sum - sales	\$24,686.00	\$21,799.00	\$11,439.00	\$57,924.00
	Count - date	11	9	6	26
Jun	Sum - sales	\$18,443.00	\$21,271.00	\$24,329.00	\$64,043.00
	Count - date	8	7	9	24
Total Sum - sales		156214	109852	142699	408765
Total Count - date		57	47	52	156

Figure 61: DataPilot shows sales value and number of entries

For manipulating the existing data fields, DataPilot result contains a field called Data. You can move it just like row or column fields by using drag and drop. This is an easy way to archive a different structure for the results (Figures 62 and 63).

		category			
Data	date	golfing	sailing	tennis	Total Result
Sum - sales	Jan	\$26,180.00	\$13,979.00	\$39,206.00	\$79,365.00
	Feb	\$30,444.00	\$15,625.00	\$23,710.00	\$69,779.00
	Mar	\$31,714.00	\$17,409.00	\$12,097.00	\$61,220.00
	Apr	\$24,747.00	\$19,769.00	\$31,918.00	\$76,434.00
	May	\$24,686.00	\$21,799.00	\$11,439.00	\$57,924.00
	Jun	\$18,443.00	\$21,271.00	\$24,329.00	\$64,043.00
Count - date	Jan	9	8	13	30
	Feb	11	8	10	29
	Mar	10	7	5	22
	Apr	8	8	9	25
	May	11	9	6	26
	Jun	8	7	9	24
Total Sum - sales		\$156,214.00	\$109,852.00	\$142,699.00	\$408,765.00
Total Count - date		57	47	52	156

Figure 62: Layout option for presenting the sums and numbers of the sales values

	category	Data						
	golfing		sailing		tennis		Total Sum - sales	Total Count - date
date	Sum - sales	Count - date	Sum - sales	Count - date	Sum - sales	Count - date		
Jan	\$26,180.00	9	\$13,979.00	8	\$39,206.00	13	\$79,365.00	30
Feb	\$30,444.00	11	\$15,625.00	8	\$23,710.00	10	\$69,779.00	29
Mar	\$31,714.00	10	\$17,409.00	7	\$12,097.00	5	\$61,220.00	22
Apr	\$24,747.00	8	\$19,769.00	8	\$31,918.00	9	\$76,434.00	25
May	\$24,686.00	11	\$21,799.00	9	\$11,439.00	6	\$57,924.00	26
Jun	\$18,443.00	8	\$21,271.00	7	\$24,329.00	9	\$64,043.00	24
Total Result	\$156,214.00	57	\$109,852.00	47	\$142,699.00	52	\$408,765.00	156

Figure 63: Another layout option for presenting the sums and numbers of the sales values

If you want to put the different data fields in different columns and your DataPilot does not contain another column field or you sort the different data fields in different rows and don't have another row field, then it is useful to disable the view of the row or column sums

	Data			
date	Sum - sales	Count - date	Total Sum - sales	Total Count - date
Jan	\$79,365.00	30	\$79,365.00	30
Feb	\$69,779.00	29	\$69,779.00	29
Mar	\$61,220.00	22	\$61,220.00	22
Apr	\$76,434.00	25	\$76,434.00	25
May	\$57,924.00	26	\$57,924.00	26
Jun	\$64,043.00	24	\$64,043.00	24
Total Result	\$408,765.00	156	\$408,765.00	156

Figure 64: Unnessecary columns

date	Data	
Jan	Sum - sales	\$79,365.00
	Count - date	30
Feb	Sum - sales	\$69,779.00
	Count - date	29
Mar	Sum - sales	\$61,220.00
	Count - date	22
Apr	Sum - sales	\$76,434.00
	Count - date	25
May	Sum - sales	\$57,924.00
	Count - date	26
Jun	Sum - sales	\$64,043.00
	Count - date	24
Total Sum - sales		408765
Total Count - date		156

Figure 65: Disabled column sums

A frequent use case for multiple data fields is the aggregation of one value according to different aggregate functions at the same time. You can create a DataPilot that shows you the monthly sales values and shows you additionally the smallest and the biggest amount.

	Data		
date	Sum - sales	Min - sales2	Max - sales3
Jan	\$79,365.00	\$155.00	\$4,872.00
Feb	\$69,779.00	\$204.00	\$4,893.00
Mar	\$61,220.00	\$669.00	\$4,881.00
Apr	\$76,434.00	\$498.00	\$4,873.00
May	\$57,924.00	\$112.00	\$4,655.00
Jun	\$64,043.00	\$454.00	\$4,839.00
Total Result	\$408,765.00	\$112.00	\$4,893.00

Figure 66: Multiple analyses for the same data field

Unfortunately, Calc does not allow you to use the same data field multiple times. For such multiple analyses you have to duplicate in the source the respective column. This column is then available within the DataPilot; it just happens to have identical values.

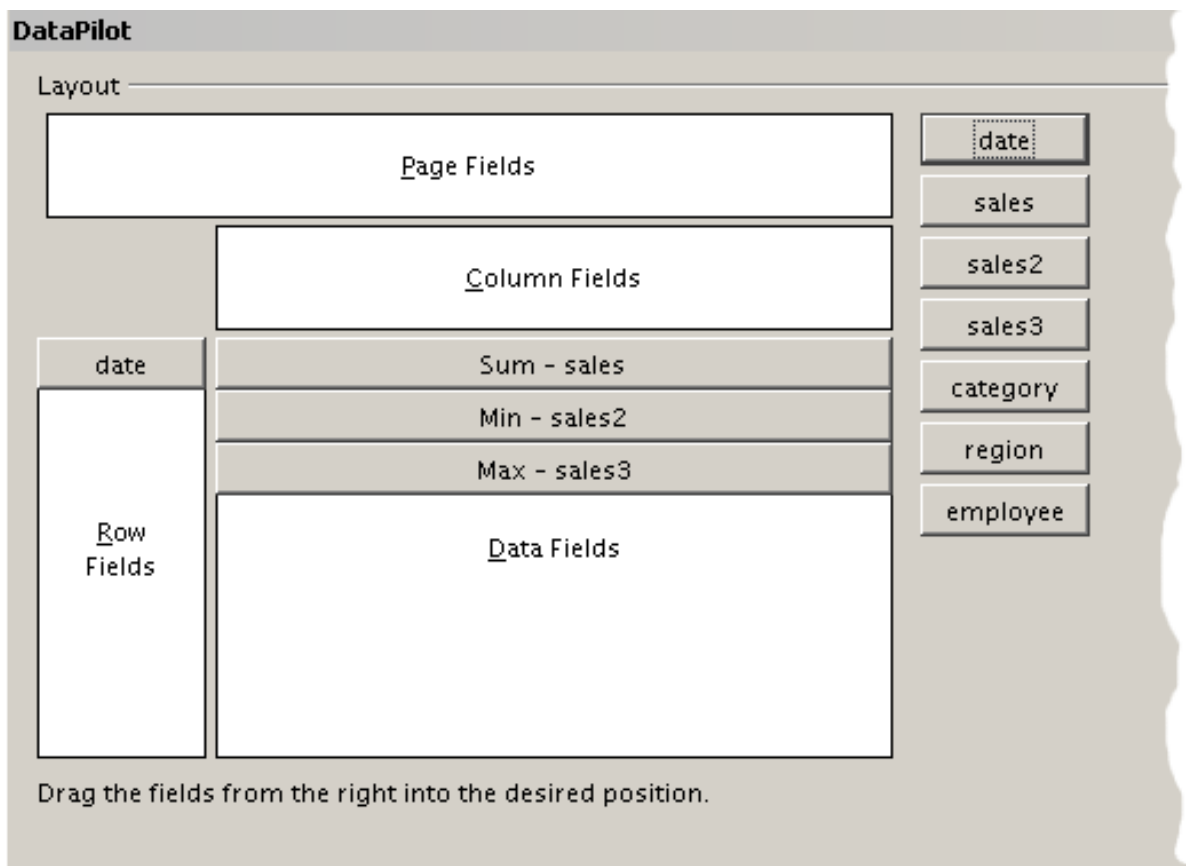


Figure 67: Multiple identical fields for sales value, that have been created as duplicates within your source.

Shortcuts

If you use the DataPilot very often, you might find the frequent use of the menu paths (**Data > DataPilot > Start** and **Data > Group and Outline > Group**) inconvenient.

For grouping, a shortcut is already defined: *F12*. For starting the DataPilot, you can define your own keyboard shortcut. If you prefer to have toolbar icons instead of keyboard shortcuts, you can create a user-defined symbol and add it to either your own custom made toolbar or the Standard toolbar.

For an explanation how to create keyboard shortcuts or add icons to toolbars, see Chapter 14 (Setting Up and Customizing Calc).

Function GETPIVOTDATA

The function GETPIVOTDATA can be used with formulas in Calc. Use this function if you want to reuse the results from the DataPilot somewhere else in your spreadsheet.

Difficulty

Normally you create a reference to a value by entering the address of the cell that contains the value. For example, the formula `=C6*2` creates a reference to cell **C6** and returns the doubled value.

If this cell is located in the results area of the DataPilot, it contains the result that was calculated by referencing specific categories of the row and column fields. In Figure 68, the cell **C6** contains the sum of the sales values of the employee Hans in the category Sailing. The formula in the cell **C12** uses this value.

	A	B	C	D	E
1	Filter				
2					
3	Sum - sales	category			
4	employee	golfing	sailing	tennis	Total Result
5	Brigitte	\$26,961.00	\$24,422.00	\$17,721.00	\$69,104.00
6	Hans	\$17,941.00	\$14,060.00	\$30,467.00	\$62,468.00
7	Kurt	\$28,346.00	\$20,814.00	\$27,675.00	\$76,835.00
8	Ute	\$36,127.00	\$30,580.00	\$28,198.00	\$94,905.00
9	Total Result	\$109,375.00	\$89,876.00	\$104,061.00	\$303,312.00
10					
11					
12			\$28,120.00		

Figure 68: Formula reference to a cell of the DataPilot

If the underlying data or the layout of the DataPilot changes, then you must take into account that the sales value for Hans might appear in a different cell. Your formula still references the cell **C6** and therefore uses a wrong value. The value that you really want to use can be found in a different location.

	A	B	C	D	E
1	Filter				
2					
3	Sum - sales	category			
4	employee	golfing	sailing	tennis	Total Result
5	Brigitte	\$26,961.00	\$24,422.00	\$17,721.00	\$69,104.00
6	Fritz	\$46,839.00	\$19,976.00	\$38,638.00	\$105,453.00
7	Hans	\$17,941.00	\$14,060.00	\$30,467.00	\$62,468.00
8	Kurt	\$28,346.00	\$20,814.00	\$27,675.00	\$76,835.00
9	Ute	\$36,127.00	\$30,580.00	\$28,198.00	\$94,905.00
10	Total Result	\$156,214.00	\$109,852.00	\$142,699.00	\$408,765.00
11					
12			\$39,952.00		

Figure 69: The value that you really want to use can be found now in a different location.

The function GETPIVOTDATA allows you to have a reference to a value inside the DataPilot by using the specific identifying categories for this value.

Syntax

The syntax has two variations:

GETPIVOTDATA(target field, DataPilot; [Field name / Element; ...])

GETPIVOTDATA(DataPilot; specification)

First syntax variation

Giving the **target field** is necessary to specify which data field of the DataPilot is used within the function. If your DataPilot has only one data field, this entry is ignored, but you have to enter it anyway.

If your DataPilot has more than one data field, then you have to enter the field name from the underlying data source (for example "sales value") or the field name of the data field itself (for example "sum – sales value").

The argument **DataPilot** specifies the DataPilot that you want to use. It is possible that your document contains more than one DataPilot. Enter here a cell reference that is inside the results area of your DataPilot. It might be a good idea to always use the upper left corner cell of your DataPilot, so you can be sure that the cell will always be within your DataPilot even if the layout changes.

Example: `GETPIVOTDATA("sales value";A1)`

If you enter only the first two arguments, then the function returns the total result of the DataPilot.

You can add more arguments as pairs with **field name** and **element** to retrieve specific partial sums. In the example in Figure 68, where we want to get the partial sum of Hans for sailing, the formula in cell **C12** would look like this:

`=GETPIVOTDATA("sales value";A1;"employee";"Hans";"category";"sailing")`

	A	B	C	D	E	F	G
1	Filter						
2							
3	Sum - sales	category					
4	employee	golfing	sailing	tennis	Total Result		
5	Brigitte	\$26,961.00	\$24,422.00	\$17,721.00	\$69,104.00		
6	Fritz	\$46,839.00	\$19,976.00	\$38,638.00	\$105,453.00		
7	Hans	\$17,941.00	\$14,060.00	\$30,467.00	\$62,468.00		
8	Kurt	\$28,346.00	\$20,814.00	\$27,675.00	\$76,835.00		
9	Ute	\$36,127.00	\$30,580.00	\$28,198.00	\$94,905.00		
10	Total Result	\$156,214.00	\$109,852.00	\$142,699.00	\$408,765.00		
11							
12			14060				

Figure 70: First syntax variation

Second syntax variation

The argument **DataPilot** has to be given in the same way as for the other syntax variation.

For the **specifications**, enter a list separated by spaces to specify the value you want from the DataPilot. This list must contain the name of the data field, if there is more than one data field, otherwise it is not required. To select a specific partial result, add more entries in the form of Field name[element].

In the example in Figure 68, where we want to get the partial sum of Hans for Sailing, the formula in cell **C12** would look like this:

```
=GETPIVOTDATA(A1;"sales value employee[Hans] category[sailing]")
```

	A	B	C	D	E	F	G
1	Filter						
2							
3	Sum - sales	category					
4	employee	golfing	sailing	tennis	Total Result		
5	Brigitte	\$26,961.00	\$24,422.00	\$17,721.00	\$69,104.00		
6	Fritz	\$46,839.00	\$19,976.00	\$38,638.00	\$105,453.00		
7	Hans	\$17,941.00	\$14,060.00	\$30,467.00	\$62,468.00		
8	Kurt	\$28,346.00	\$20,814.00	\$27,675.00	\$76,835.00		
9	Ute	\$36,127.00	\$30,580.00	\$28,198.00	\$94,905.00		
10	Total Result	\$156,214.00	\$109,852.00	\$142,699.00	\$408,765.00		
11							
12			14060				

Figure 71: Second syntax variation