



Calc Guide

Chapter 6

DataPilot:

Creating order from chaos

OpenOffice.org

Copyright

This document is Copyright © 2007 by its contributors as listed in the section titled **Authors**. You can distribute it and/or modify it under the terms of either the GNU General Public License, version 2 or later (<http://www.gnu.org/licenses/gpl.html>), or the Creative Commons Attribution License, version 2.5 (<http://creativecommons.org/licenses/by/2.5/>) or later.

All trademarks within this guide belong to their legitimate owners.

Authors

Peter Kupfer
Richard Detwiler

Feedback

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

Publication date and software version

Published 04 March 2007. Based on OpenOffice.org 2.1.



You can download an editable version of this document from
<http://oooauthors.org/en/authors/userguide2/published/>

Contents

Copyright.....	i
Authors.....	i
Feedback.....	i
Publication date and software version.....	i
What is DataPilot?.....	1
Example of use of DataPilot.....	1
Creating DataPilot tables.....	2
DataPilot layout fields.....	3
Page fields.....	3
Data fields.....	3
Column fields.....	4
Row fields.....	4
Completed layout.....	4
Selecting DataPilot output.....	5
New sheet.....	5
In a specific range of the current sheet.....	6
Filtering DataPilot tables.....	6
DataPilot filter criteria.....	7
Field name.....	7
Condition.....	7
Value.....	7
DataPilot filter examples.....	7
Text filter.....	7
Numerical filter.....	8
Making changes to the DataPilot table.....	9
Editing in the DataPilot wizard.....	9
Rearranging the DataPilot layout.....	9
Updating DataPilot tables.....	9
Deleting DataPilot tables.....	9
Advanced DataPilot use.....	10

Using Page fields.....	10
Grouping DataPilot tables.....	10

What is DataPilot?

The DataPilot (referred to as Pivot Table in Microsoft Office) is a tool for creating an interactive table where you can combine, compare, and analyze large amounts of data. You can view different summaries of the source data, display the details of areas of interest, and create reports. Data can be arranged, rearranged, or summarized according to different points of view.

This chapter shows how to make use of one of Calc's most powerful tools.

Throughout this chapter we will be using data borrowed from <http://www.wcape.school.za/subject/CS/PHS/excelpiv/index.htm>.

Example of use of DataPilot

Here is an example of a situation where DataPilot would be useful. The following data were compiled about the occupancy rates of three different hotels in a certain chain, over a five-month period.

	A	B	C
1	Hotel	Month	Occupancy rate
2	Hotel A	Jan	85%
3	Hotel A	Feb	83%
4	Hotel A	Mar	87%
5	Hotel A	Apr	88%
6	Hotel A	May	81%
7	Hotel B	Jan	63%
8	Hotel B	Feb	59%
9	Hotel B	Mar	69%
10	Hotel B	Apr	72%
11	Hotel B	May	70%
12	Hotel C	Jan	78%
13	Hotel C	Feb	79%
14	Hotel C	Mar	75%
15	Hotel C	Apr	81%
16	Hotel C	May	83%

Figure 1: Data on occupancy rate, by hotel and by month

The chain manager would like an easy-to-read summary that shows the average occupancy rate by hotel and by month, and also will show, in a neatly-organized table, the occupancy rate data shown in the above table.

DataPilot can quickly produce the summary shown in Figure 2.

This summary shows that the average occupancy rate is highest for Hotel A and lowest for Hotel B, and also shows that across all hotels, the occupancy rate gradually rose from January to April, and then fell slightly in May.

Average - Occupancy rate	Hotel			
Month	Hotel A	Hotel B	Hotel C	Total Result
Jan	85%	63%	78%	75%
Feb	83%	59%	79%	74%
Mar	87%	69%	75%	77%
Apr	88%	72%	81%	80%
May	81%	70%	83%	78%
Total Result	85%	67%	79%	77%

Figure 2: DataPilot summary of occupancy rate data

Creating DataPilot tables

To create a DataPilot table, first enter the data into Calc.

Our sample data will utilize the data shown, in part, in Figure 3. There are 40 lines of data in this table, which gives the diameter of widgets made at factories in four locations, on five days of the week. There are two different but similar product types, Widget A and Widget B. Our goal is to determine the differences between the two products, the location in which they were manufactured, and the day of the week they were manufactured.

Product name	Factory	Day of week	Diameter, cm
Widget A	Spain	Mon	2.41
Widget A	Germany	Mon	2.24
Widget A	UK	Mon	2.18
Widget A	Sweden	Mon	2.16
Widget A	Spain	Tues	3.02
Widget A	Germany	Tues	3.05
Widget A	UK	Tues	2.44
Widget A	Sweden	Tues	2.46
Widget A	Spain	Wed	2.53
Widget A	Germany	Wed	2.41

Figure 3: Some of the sample data

After you have entered the data:

- 1) Select all of the data, including column headings, by highlighting it. Optionally, if the data you want to analyze is contiguous (no blank columns or rows within the data range of interest), just select a single cell. The DataPilot will select the entire contiguous range in the next step.
- 2) Select **Data > DataPilot > Start**. The Select Source dialog (Figure 4) opens.

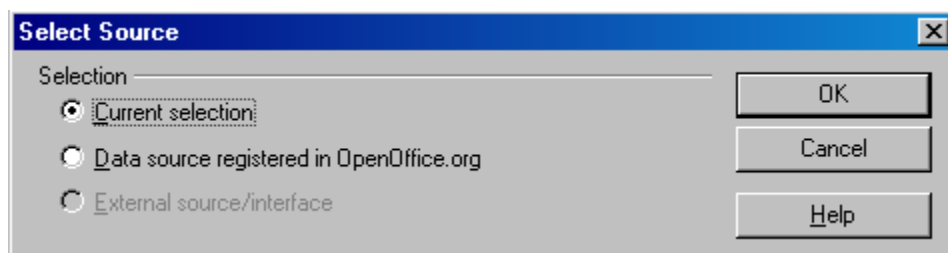


Figure 4: DataPilot—Select Source dialog

- 3) Choose **Current selection** unless you know you are choosing the data from a data source. Click **OK**. The DataPilot Layout dialog (Figure 5) appears.

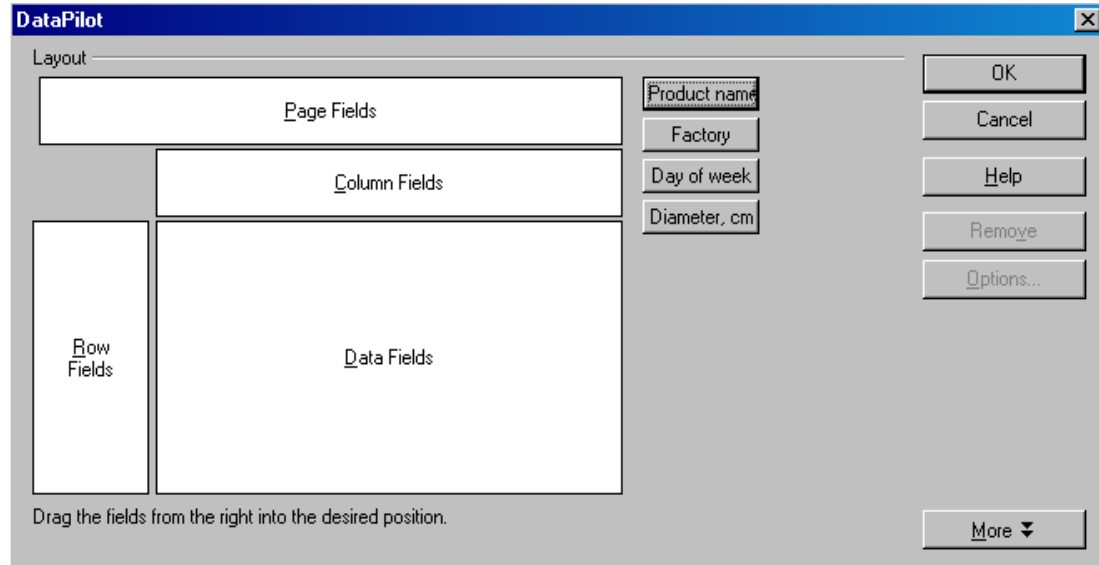


Figure 5: DataPilot—Layout dialog

- 4) The column headers from the data appear as buttons in the DataPilot Layout dialog. Those column headers can be placed in any of the four areas shown: Page Fields, Data Fields, Column Fields, and Row Fields.

DataPilot layout fields

Page fields

The optional Page Fields feature allows filtering the data by the selected item. In this case, we decide to use **Factory** as the Page Field. This allows us to display data by each factory in turn, by selecting that factory from a list. It also will be possible to select *all* factories, to show the data summary averaged across all factory locations.

Drag the **Factory** button to the Page Fields area to create a button and a list box on top of the generated DataPilot table. More than one field can be placed in the Page Fields portion of the layout. You can use drag-and-drop within the generated DataPilot table to use another page field as a filter.

Data fields

Data fields are calculated columns. In our example, the field **Diameter, cm** is the item of interest, so this is the field that will be used as the data field. Any column heading dropped in the Data Fields area will automatically be given a caption that also shows the formula used to calculate the data. The default is the *Sum* function; other options include *Count*, *Average*, *Max*, *Min*, *Standard Deviation*, and several other choices.

To change the formula, double-click on the field that has been dragged into the Data Fields area, which calls up the Data Field dialog. (You can also click the **Options** button.) The Data Field dialog is shown in Figure 6.

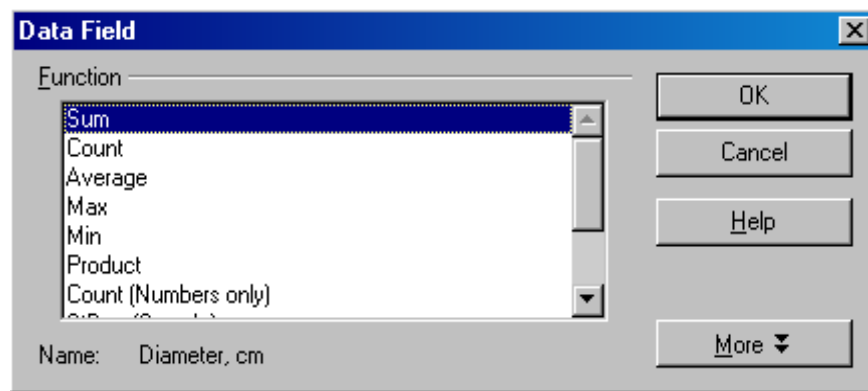


Figure 6: Data Field dialog

Use the Data Field dialog to select the calculations to be used for the data. To make a multiple selection, hold down the *Ctrl* key while clicking the desired calculation.

The order of the buttons can be changed at any time by moving them to a different position in the area with the mouse. Remove a button by dragging it back to the area of the other buttons at the right of the dialog.

To open the Data Field dialog, double-click one of the buttons in the Row or Column area. Use the dialog to select if and to what extent Calc calculates display subtotals.

Column fields

Items in the Column Fields area will appear on the top. In this example, we chose **Product Name** for the column field.

Row fields

Items in the Row Fields area will appear on the left. In this example, we chose **Day of Week** for the row field.

Completed layout

After dragging **Factory** into the *Page Fields* area, **Product name** into the *Column Fields* area, **Day of Week** into the *Row Fields* area, and **Diameter, cm** into the *Data fields* area, the result should look like Figure 7. (In this example, the function for the **Diameter, cm** field in the *Data fields* area was changed from *Sum* to *Average*.)

Note All items in the Data Fields area will be calculated values. As soon as you drag the item in, it will automatically change to a calculated value.

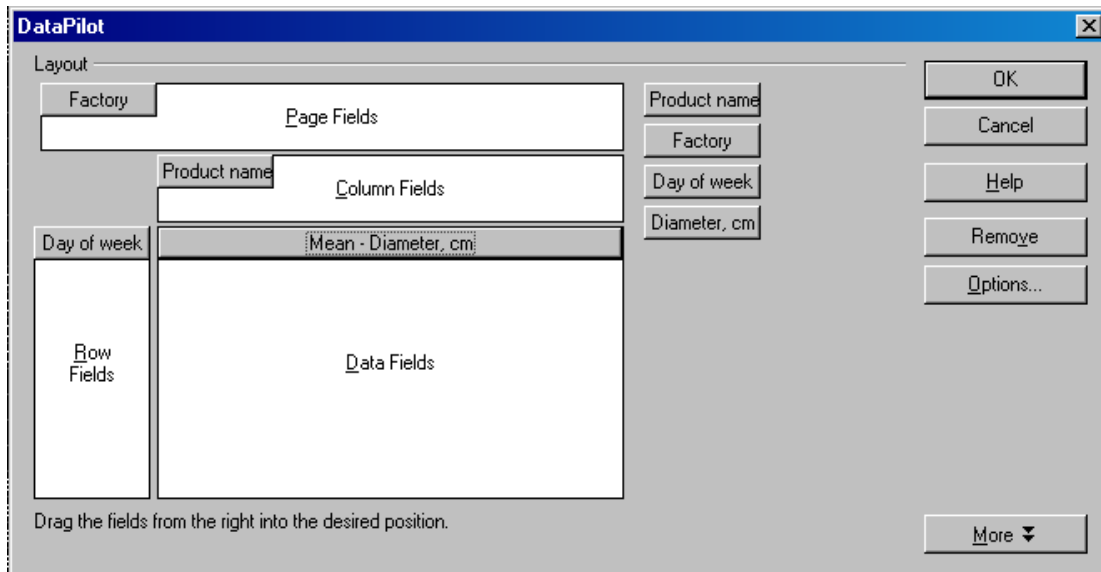


Figure 7: DataPilot—layout complete

Selecting DataPilot output

By default, the DataPilot output will be placed in the current sheet, immediately below the data range that is used. If this would overwrite existing data in the sheet, there will be a warning dialog. (Figure 8)

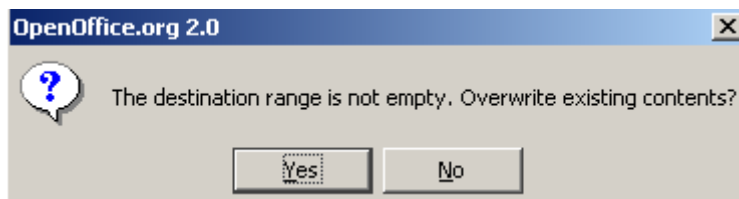


Figure 8: Warning dialog if the default destination range is not empty

Optionally, the output can be placed in a different existing sheet, in a new sheet, or in a specific range of the current sheet.

New sheet

To place the DataPilot table in a new sheet, click on the **More** button and choose *-new sheet-* in the *Results to* drop downlist, as shown in Figure 9.

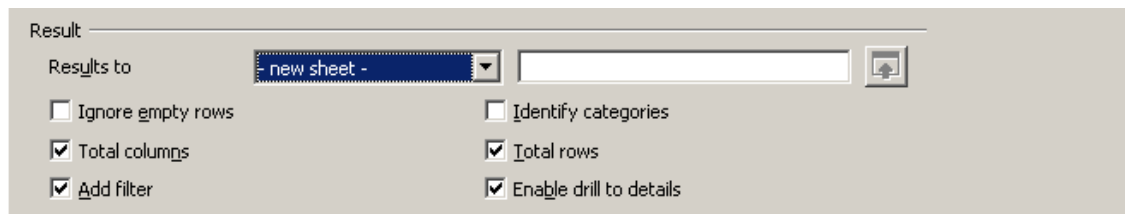


Figure 9: More Options - New Sheet

In a specific range of the current sheet

To place the DataPilot table in a specific location in the current sheet select *undefined* in the *Results to* drop down list and then enter the top left cell of the new DataPilot table's location in the next box, as shown in Figure 10.

Figure 10: Specifying a location in the current sheet

- 1) In our example, we want to put our DataPilot table on a new sheet. To do this, click the **More** button in the bottom left hand corner of Figure 7.
- 2) Select the new sheet option in the dropdown list shown in Figure 9.
- 3) Click **OK**. Your DataPilot table will open in a new sheet and will resemble the example in Figure 11.

	A	B	C	D
1	Filter			
2	Factory	- all -		
3				
4	Average - Di	Product name		
5	Day of week	Widget A	Widget B	Total Result
6	Fri	1.91	2.30	2.11
7	Mon	2.25	2.45	2.35
8	Th	2.18	2.54	2.36
9	Tues	2.74	2.91	2.82
10	Wed	2.31	2.41	2.36
11	Total Result	2.28	2.52	2.40
12				

Figure 11: Completed DataPilot Example

Filtering DataPilot tables

One important feature of the DataPilot table is that you can easily filter out the unwanted material without deleting the data. For example, in the “Widget” data set above, let’s say that the manager is interested in any differences between parts made on Monday and Friday, compared to parts made on Tuesday through Thursday. Using the filter feature allows the manager to obtain data summaries of the day-of-week ranges of interest.

There are two ways to call the filter dialog:

- Left-click on the **Filter** button in cell A1.
- Right-click anywhere in the output of the DataPilot, and select **Filter**.

Either method will open the Filter dialog shown in Figure 12.

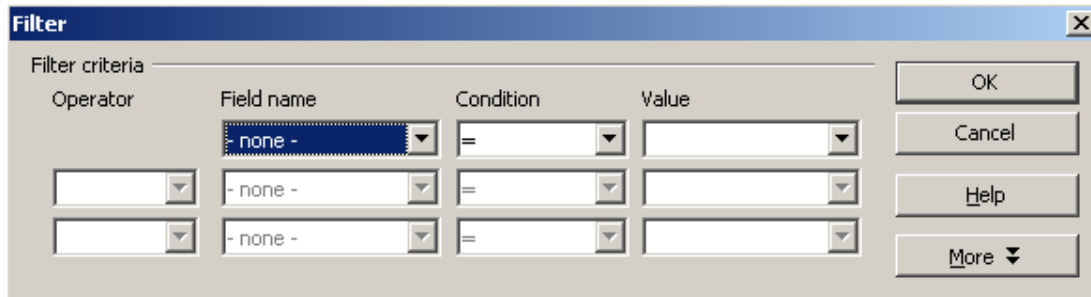


Figure 12: DataPilot Filter dialog

DataPilot filter criteria

Field name

Here is where all of the column headers in your data will appear. If you did not have column headers in your original data, columns labels (or letters) will appear.

Condition

The conditions are logical operators such as greater than, less than, equal to, or not equal to.

Value

Select the value that you want to compare to the selected field. If you chose a text field, a list of options will be in the drop-down list. If you chose a numerical field, a list of values will be in the numerical field. If you chose less than a specific numerical value, then only types with all of the items less than that value will be displayed.

DataPilot filter examples

Text filter

In Figure 13, we choose to display the diameters of widgets produced on Monday or Friday. This yields the DataPilot table in Figure 14.

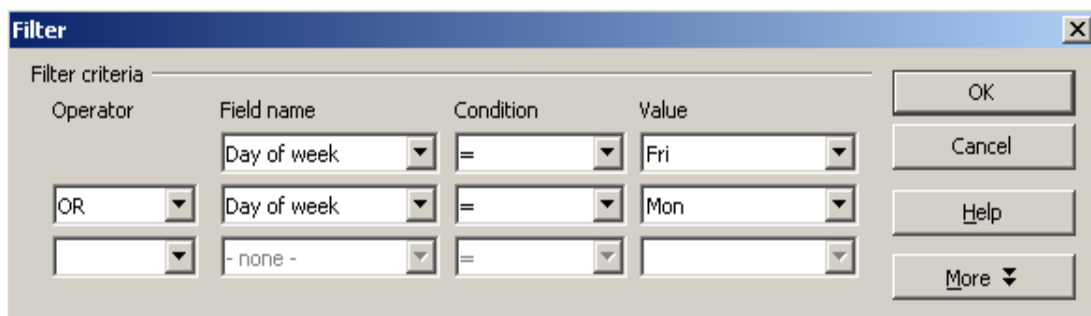


Figure 13: DataPilot filter for widgets produced on Monday or Friday

	A	B	C	D
1	Filter			
2	Factory	- all -		
3				
4	Average - Diameter, cm	Product name		
5	Day of week	Widget A	Widget B	Total Result
6	Fri	1.91	2.30	2.11
7	Mon	2.25	2.45	2.35
8	Total Result	2.08	2.38	2.23

Figure 14: DataPilot filter example for Monday and Friday

Numerical filter

For this example, a modification to the DataPilot used on the “widget” data set will be used. In this case, instead of showing the average diameter, the DataPilot is set up to show the count (the number of results from each widget type and day of the week). The filter in Figure 15 is then used to show the number of widgets that have a diameter less than 2.3 cm. The result is the DataPilot table shown in Figure 16.

Figure 15: DataPilot filter for diameters less than 2.30 cm

	A	B	C	D
1	Filter			
2	Factory	- all -		
3				
4	Count - Diameter, cm	Product name		
5	Day of week	Widget A	Widget B	Total Result
6	Fri	4	2	6
7	Mon	3	1	4
8	Th	2	1	3
9	Wed	1	1	2
10	Total Result	10	5	15

Figure 16: DataPilot filter for diameters less than 2.30 cm

You can mix and match many different filter combinations, but Calc limits you to only three controls in your filter.

Making changes to the DataPilot table

You can edit the DataPilot table in two ways:

- From the table itself
- By reopening the DataPilot dialog

Editing in the DataPilot wizard

To edit the DataPilot table, right-click a cell inside the DataPilot table and select **Start**, which displays the DataPilot dialog for the current DataPilot table. Alternatively, left-clicking on a cell inside the DataPilot table and selecting the menu path **Data > DataPilot > Start** also displays the DataPilot dialog for the current DataPilot table.

From the dialog you can add and remove buttons as you did when you created the DataPilot table.

Rearranging the DataPilot layout

Click one of the buttons in the table that the DataPilot has created and hold the mouse button down. A special symbol will appear next to the mouse pointer.

By dragging the button to a different position in the same row, you can alter the order of the columns. If you drag a button to the left edge of the table into the row headings area, you can change a column into a row.

Updating DataPilot tables

If the data of the source sheet has been changed, Calc does not automatically recalculate the DataPilot table. To recalculate the table, choose **Data > DataPilot > Refresh**, or right-click anywhere inside the table and select **Refresh**. Do the same after you have imported an Excel Pivot table into Calc.

Deleting DataPilot tables

To delete the DataPilot table, do one of the following:

- Right-click anywhere in the DataPilot table and select **Delete** from the pop-up menu.
- Place the cursor in the DataPilot table and select **Data > DataPilot > Delete** from the menu bar.

Advanced DataPilot use

This section explores some further abilities of the DataPilot. We will start with the same base table and try some of the other features.

Using Page fields

Page fields were described on page 3. A Page field entry allows you to filter the DataPilot table without opening the filter dialog every time. In the Widget example, **Factory** was selected as the Page field.

In Figure 11, the output of the DataPilot is shown with data from all factories; in other words, the Page field is not in use in Figure 11. To use the Page field, click the down arrow in cell B2 and select a factory location, as shown in Figure 17.

	A	B	C	D
1	Filter			
2	Factory	Germany ↓		
3				
4	Average - Diameter, cm	Product name		
5	Day of week	Widget A	Widget B	Total Result
6	Fri	2.05	2.61	2.33
7	Mon	2.24	2.71	2.48
8	Th	2.10	2.46	2.28
9	Tues	3.05	2.85	2.95
10	Wed	2.41	2.38	2.40
11	Total Result	2.37	2.60	2.49

Figure 17: Use of the Page field to select results from Germany only

Grouping DataPilot tables

Another way to potentially enhance the usability of the DataPilot table is to group some of the entries. In our widget example, let's again look at the situation of comparing parts made on Monday or Friday, with those made on other days of the week¹.

- 1) Select row headings or column headings that are to be grouped together. (*Ctrl* click to select non-contiguous headings.) For our example, we will first select Mon and Fri.
- 2) Select **Data > Outline > Group** or press *F12*.
- 3) Next, select the row headings Th, Tu, and Wed, and again select **Data > Outline > Group**.

Depending on the format of the selected cells, either a new group field is added to the DataPilot table, or you see one of the two Grouping dialogs, either for numeric values or for date values. Neither is the case here so the result should resemble Figure 18.

¹ Thanks to Erwin Tenhumberg's excellent tutorial on DataPilot for the help on the grouping feature. http://blogs.sun.com/roller/page/dancer?entry=just_switch_analyzing_data_with

	A	B	C	D	E	F
1	Filter					
2	Factory	- all -				
3						
4	Average - Diameter, cm			Product name		
5	Day of week3	Day of week2	Day of week	Widget A	Widget B	Total Result
6	Group1	Group1	Fri	1.91	2.30	2.11
7			Mon	2.25	2.45	2.35
8	Group2	Th	Th	2.18	2.54	2.36
9		Tues	Tues	2.74	2.91	2.82
10		Wed	Wed	2.31	2.41	2.36
11	Total Result			2.28	2.52	2.40

Figure 18: DataPilot grouping example

You can also rename the group to be more useful. To rename a group, click in the cell with the name, and type in a new name. For instance, in this case, we might rename the first group *M-F*, and the second group *Tu-W-Th*.

If you double-click on the group name, it will collapse down to one line as in Figure 19.

	A	B	C	D
1	Filter			
2	Factory	- all -		
3				
4	Average - Diameter, cm	Product name		
5	Day of week3	Widget A	Widget B	Total Result
6	M-F	2.08	2.38	2.23
7	Tu-W-Th	2.41	2.62	2.51
8	Total Result	2.28	2.52	2.40

Figure 19: DataPilot collapsed grouping example

To remove a grouping, click inside the group, then choose **Data > Outline > Ungroup**.