USER GUIDE
Knowledge Blocks
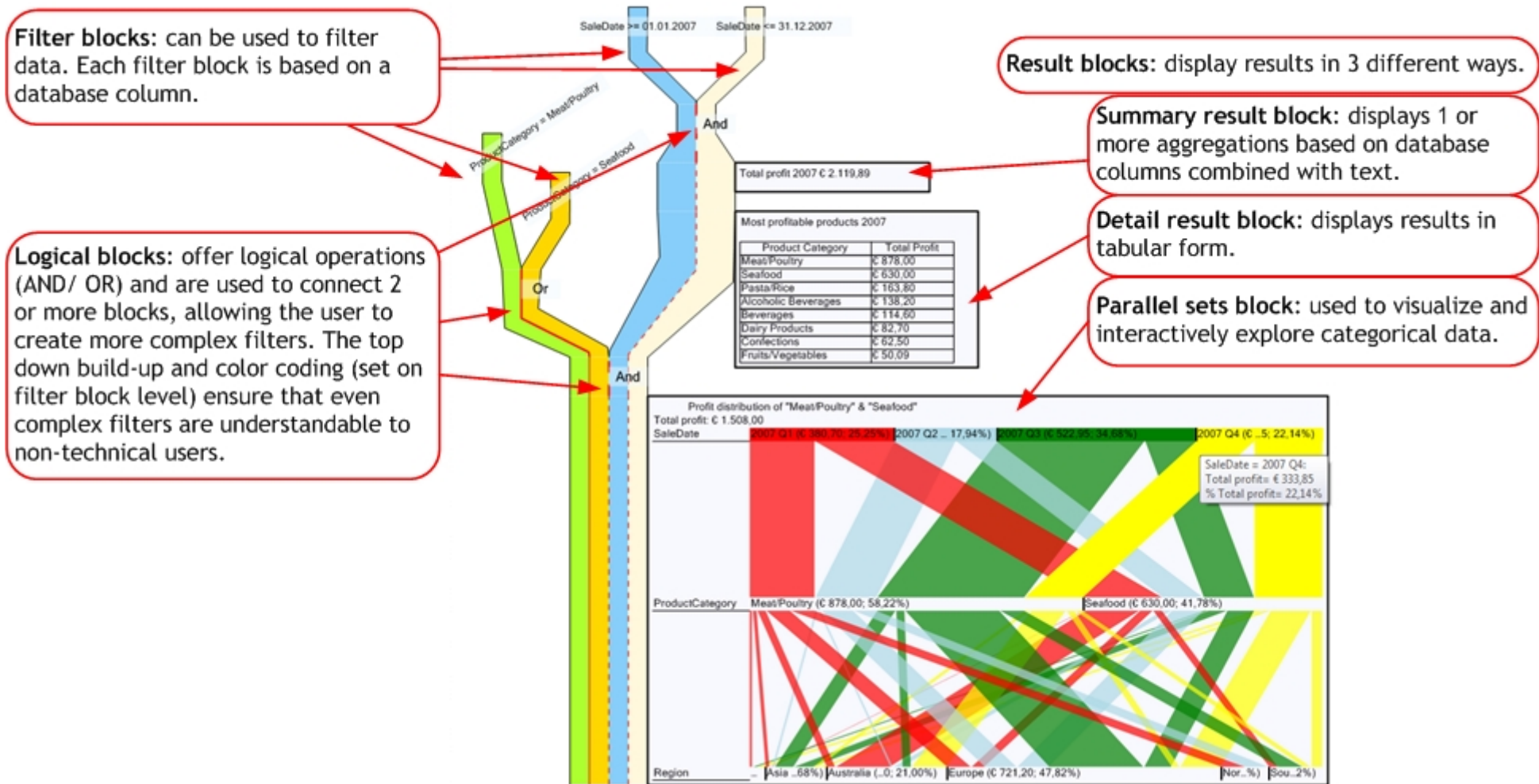
tawil eu

Table of contents

1 Introduction

What is Knowledge Blocks?
Knowledge blocks is a new type of data analysis software that is designed for non-technical business professionals. The intuitive system helps you create complex queries without having SQL knowledge or requiring IT support. The system presents you with 3 basic building blocks and a drag and drop interface to connect and configure blocks and build reports.

Filter blocks: can be used to filter data. Each filter block is based on a database column.

Logical blocks: offer logical operations (AND/ OR) and are used to connect 2 or more blocks, allowing the user to create more complex filters. The top down build-up and color coding (set on filter block level) ensure that even complex filters are understandable to non-technical users.

SaleDate >= 01.01.2007   SaleDate <= 31.12.2007

ProductCategory = Meat/Poultry

ProductCategory = Seafood

And

Or

And

Total profit 2007 € 2.119,89

Most profitable products 2007

Product Category	Total Profit
Meat/Poultry	€ 878,00
Seafood	€ 630,00
Pasta/Rice	€ 163,80
Alcoholic Beverages	€ 138,20
Beverages	€ 114,60
Dairy Products	€ 82,70
Confections	€ 62,50
Fruits/Vegetables	€ 50,09

Profit distribution of "Meat/Poultry" & "Seafood"
Total profit: € 1.508,00
SaleDate: 2007 Q1 (€ 380,70; 25,25%) 2007 Q2 .. 17,94%) 2007 Q3 (€ 522,95; 34,68%) 2007 Q4 (€ ..5; 22,14%)

SaleDate = 2007 Q4:
Total profit= € 333,85
% Total profit= 22,14%

ProductCategory: Meat/Poultry (€ 878,00; 58,22%)   Seafood (€ 630,00; 41,78%)

Region: .. Asia ..68%) Australia (..0; 21,00%) Europe (€ 721,20; 47,82%)   Nor..%) Sou..2%)

Result blocks: display results in 3 different ways.

Summary result block: displays 1 or more aggregations based on database columns combined with text.

Detail result block: displays results in tabular form.

Parallel sets block: used to visualize and interactively explore categorical data.

Why Knowledge blocks?
Knowledge blocks has been designed specifically for non-technical business professionals.

- **Visual presentation of queries**: Filter criteria are presented in a top down manner that can be easily interpreted. The color coding assigned to filter blocks is inherited to all other connected blocks, allowing the user to easily identify which filters are active on result blocks. The following is the SQL equivalent of the previous diagram. The SQL code is generated by knowledge blocks, allowing the user to concentrate on data analysis and not writing SQL statements.

```
SELECT Format([ProductSales].[SaleDate] , 'yyyy "Q"q') AS
[GroupYearQuarter_ProductSales_SaleDate_0], [Products].[ProductCategory] ,
[Stores].[Region] , Sum([ProductSales].[Profit] ) AS [Sum_ProductSales_Profit_3]
FROM [Stores] INNER JOIN ([Products] INNER JOIN [ProductSales] ON [Products].[ProductID]
= [ProductSales].[ProductID] ) ON [Stores].[StoreID]  = [ProductSales].[StoreID]
WHERE (((([Products].[ProductCategory]  = @P1  )  OR ([Products].[ProductCategory]  = @P2
)  ) AND (([ProductSales].[SaleDate]  >= @P3   )  AND ([ProductSales].[SaleDate]  <= @P4
)  ) )
Group By Format([ProductSales].[SaleDate] , 'yyyy "Q"q') , [Products].[ProductCategory]
, [Stores].[Region]

@P1 (WChar) = Meat/Poultry
@P2 (WChar) = Seafood
@P3 (Date) = 01.01.2007 00:00:00
@P4 (Date) = 31.12.2007 00:00:00
```

- **Database Relationship**: Database relationships can be challenging to non-technical users. Knowledge blocks identifies and imports all relevant database relationships. Using the library designer, users can edit existing and create new relationships. Knowledge blocks also supports relationships between excel tables or between text files.
  The table relationship information is used by the system to determine what filter combinations are valid and which columns can be added to result blocks, thus ensuring that only valid queries are generated.

- **Multiple Result blocks with different filter criteria in one document:** Knowledge blocks allows you attach multiple result blocks to defined filters, refine your filter and attach further result blocks. This can be seen in the previous diagram:
  1. We first set a filter on all sales records to see only the 2007 sales and attached a summary result block to calculate the profit for 2007.
  2. We then attached a detail result block to identify the most profitable product categories.
  3. To analyse the distribution of profit for the 2 most profitable product categories, we defined 2 further filters, attached them to our current analysis and added a parallel set analysis result block.

Application elements overview

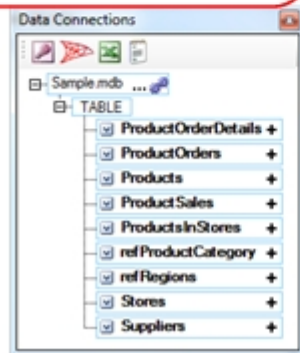**Library Designer:** create new or edit existing relationships. Toggle column visibilty.
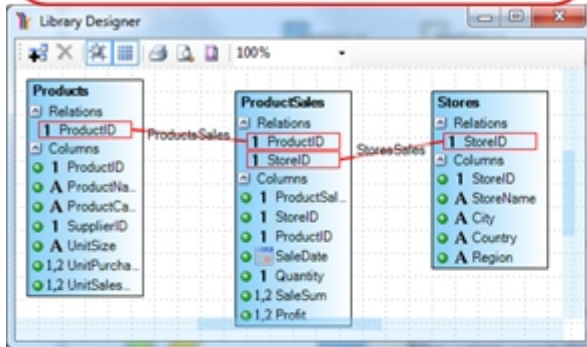
Library Panel gives you access to all knowledge blocks. Use drag and drop to add knowledge blocks to reports.

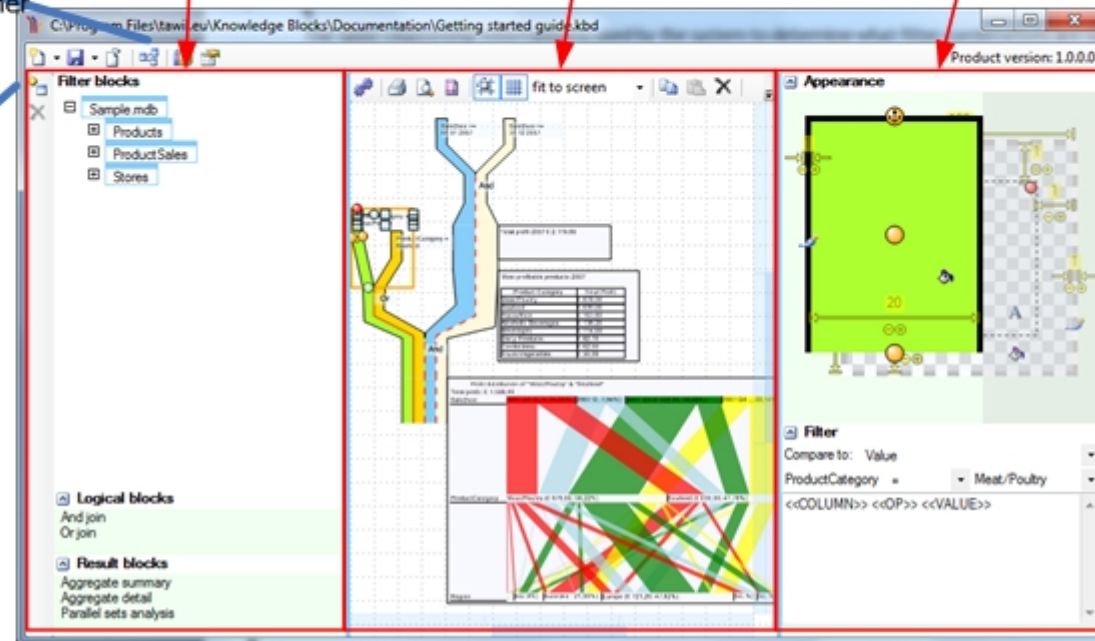**Document Panel:** connect knowledge blocks to create reports.

Properties Panel: Select 1 or more knowledge blocks in document panel to set visual properties (color, width, font), filter criteria and result columns.

Library designer

Connections manager: Add or modify database connections. Select database objects for analysis.

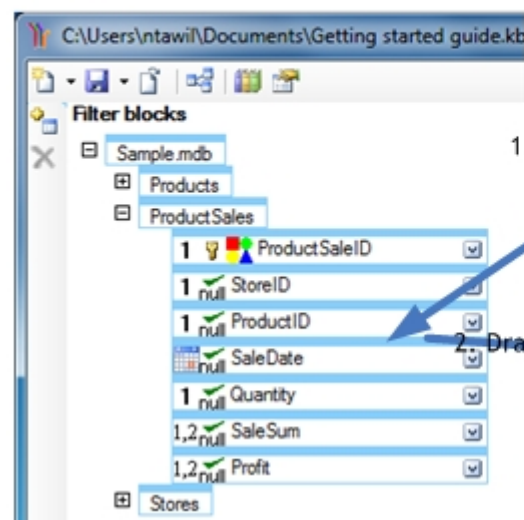Connections manager

2 Getting started

2.1 Add a data source

Before starting an analysis, we need to add a database connection and select the tables that contain the relevant data

Knowledge blocks

Filter blocks

1: Click to define a database connection

Start here

2: Add a database

3: Select „Sample.mdb"

Microsoft Access Database

Select or enter a database file:

C:\KnowledgeBlocks\Sample.mdb   Browse ...

Log on to database
- Standard security
- Database password:
  - Password:
- Workgroup security:
  - User name: Admin
  - Password:
- Workgroup file: Browse ...
- Save password

Test connection   OK   Cancel

Data Connections

- Sample.mdb ...
  - TABLE
    - ProductOrderDetails +
    - ProductOrders +
    - Products +
    - ProductSales +
    - ProductsInStores +
    - refProductCategory +
    - refRegions +
    - Stores +
    - Suppliers +

4.1: Add „Products"

4.2: Add „ProductSales"

4.3: Add „Stores"

Lets take a brief look at the tables we selected

List of all products

Products
- ▲ Relations
  - 1 ProductID
- ▲ Columns
  - ○ 1 ProductID
  - ○ A ProductName
  - ○ A ProductCategory
  - ○ 1 SupplierID
  - ○ A UnitSize
  - ○ 1,2 UnitPurchasePrice
  - ○ 1,2 UnitSalesPrice

ProductsSales

Sales information

ProductSales
- ▲ Relations
  - 1 ProductID
  - 1 StoreID
- ▲ Columns
  - ○ 1 ProductSaleID
  - ○ 1 StoreID
  - ○ 1 ProductID
  - ○ SaleDate
  - ○ 1 Quantity
  - ○ 1,2 SaleSum
  - ○ 1,2 Profit

StoresSales

List of stores

Stores
- ▲ Relations
  - 1 StoreID
- ▲ Columns
  - ○ 1 StoreID
  - ○ A StoreName
  - ○ A City
  - ○ A Country
  - ○ A Region

2.2 Profit 2007

Lets start with something simple.
What was the profit in the year 2007?

First, you need to add 2 „SaleDate" filter blocks

Repeat for 2nd filter

Click this to set blocks back color

C:\Users\ntawil\Documents\Getting started guide.kb

Filter blocks

Sample.mdb
  Products
  ProductSales
    1 ProductSaleID
    1 StoreID
    1 ProductID
    SaleDate
    1 Quantity
    1,2 SaleSum
    1,2 Profit
  Stores

1. Click & hold

2. Drag

3. Release

4. Set block properties

SaleDate >= 01.01.2007

SaleDate <= 31.12.2007

Filter
Compare to: Value
SaleDate >= 01.01.2007
<<COLUMN>> <<OP>> <<VALUE>>

Enter the value to compare to in this field

Use this pull down to select the compare operation

Now join the 2 filter blocks using an „and join"

2. Click & hold

3. Drag

4. Release

Logical blocks
And join
Or join

1. Add an „And join" block

SaleDate >= 01.01.2007

SaleDate <= 31.12.2007

And
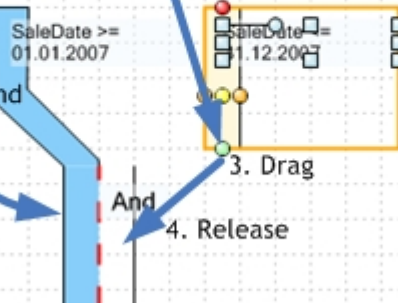
Finally add a result block

Result blocks
Aggregate summary
Aggregate detail
Parallel sets analysis

1. Add an „Aggregate summary" block

SaleDate >= 01.01.2007

SaleDate <= 31.12.2007

And

2. Connect the join and result blocks

3. Set block properties

Total profit 2007 <<Sum_ProductSales_Profit>>

4. Click here to run query and view results

Click here and select „ProductSales → Profit → Sum" to add the function „sum of profit"
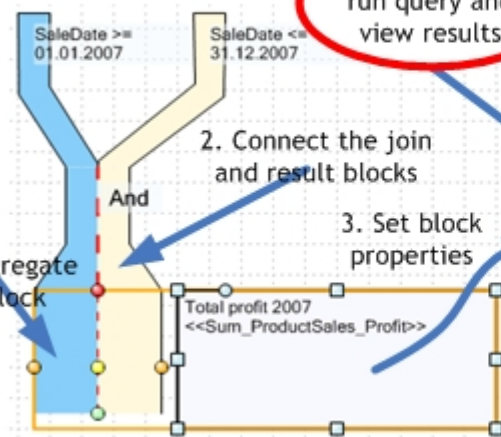
Aggregate summary
+ 1
Total profit 2007 <<Sum_ProductSales_Profit>>

Enter text to display in result block in this field

2.3 Most profitable product categories

Now that we know how much profit was made in 2007, lets find out which product categories were most profitable.

1. Add an „Aggregate detail" block

Result blocks
Aggregate summary
Aggregate detail
Parallel sets analysis

SaleDate >= 01.01.2007
SaleDate <= 31.12.2007

And

Total profit 2007 2119,89

2. Connect the aggregate summary & detail blocks

3. Set block properties

3.1. Add result columns: Products→ ProductCategory & ProductSales→ Profit

3.2 Click to activate functions

4. Update results

Aggregate detail
Σ Distinct Show max. 10 rows.
Title: Most profitable products 2007

Column	Sort	Function	Title
Products.Product...	none	Group	Product Category
ProductSales.Profit	Descending	Sum	Total profit

3.5 Enter a title for the table

3.4 Select „Descending" to see most profitable product categories at top of list

3.3 Select „Sum" function

Finally, a bit of formatting

SaleDate >= 01.01.2007
SaleDate <= 31.12.2007

And

Total profit 2007 2119,89

Most profitable products 2007
Product...	Sum of...
Meat/P...	878
Seafood	630
Pasta/...	163,8
Alcoholi...	138,2
Bevera...	114,6
Dairy P...	82,7
Confect...	62,5
Fruits/V...	50.09

5. Drag to change column width

6. Edit Column titles

7. Expand to edit column properties

8. Format „Profit" column as Currency

1,2 null SaleSum
1,2 null Profit
☐ Use list of values
Currency (€ 123.456,79)

General (123456,789)
Formatted Number (123.456,79)
Percent (12.345.678,90%)
Scientific (1,234568E+005)
Currency (€ 123.456,79)

SaleDate >= 01.01.2007
SaleDate <= 31.12.2007

And

Total profit 2007 € 2.119,89

Most profitable products 2007
Product Category	Total Profit
Meat/Poultry	€ 878,00
Seafood	€ 630,00
Pasta/Rice	€ 163,80
Alcoholic Beverages	€ 138,20
Beverages	€ 114,60
Dairy Products	€ 82,70
Confections	€ 62,50
Fruits/Vegetables	€ 50.09

2.4 Parallel sets analysis

Knowing that the 2 most profitable product categories are „Meat/Poultry" & „Seafood", we can analyse these categories more closely using parallel sets analysis.
We first need to narrow down the data to include only products of the categories „Meat/Poultry" and „Seafood". This is acomplished by following the first 3 steps described below.

What does this tooltip tell me?
The profit made from „Meat/Poultry" sales in the 2nd quarter 2007 was **162,50€**. That equals:
- 10,78% of the total profit of 2007 sales for the categories „Meat/Poultry" & „Seafood" together.
- Thats 60,07% of the Q2 profits
- 18,51% of the total „Meat/Poultry" profit for 2007.

1. Add 2 „Products→ ProductCategory" filter blocks. Set the filter conditions to „Meat/Poultry" & „Seafood"

2. Join them using an „Or join" block

3. Use an „And join" to include the date filters set above

4. Add & join a „Parallel sets analysis block

5. Set properties

Add result columns

Select „ProductSales → Profit → Sum" to set the main aggregate

Total Profit= 162,50 €
% Total Profit= 10,78%
% SaleDate = 2007 Q2: 60,07%
% Product Category = Meat/Poultry: 18,51%

Move mouse over connector to display details in a tooltip

Click and drag the labels to rearrange the columns

The bottom half of the graph shows the profit distribution in the sales regions.

Group Saledate by Year and Quarter

3 Data connections

Before starting an analysis, you need to connect to a database. Knowledge blocks allows you to connect to different types of databases. The connection manager can be used to add or edit connections. Each data connection is assigned a node where the different database objects are grouped according to object type (table, view, …).
Objects can be added to the library by clicking on the "+" sign next to the objects name. When adding an object, knowledge blocks also identifies and adds all relevant database relationships that exist between the selected objects. This information is used to ensure that only valid SQL statements are generated.
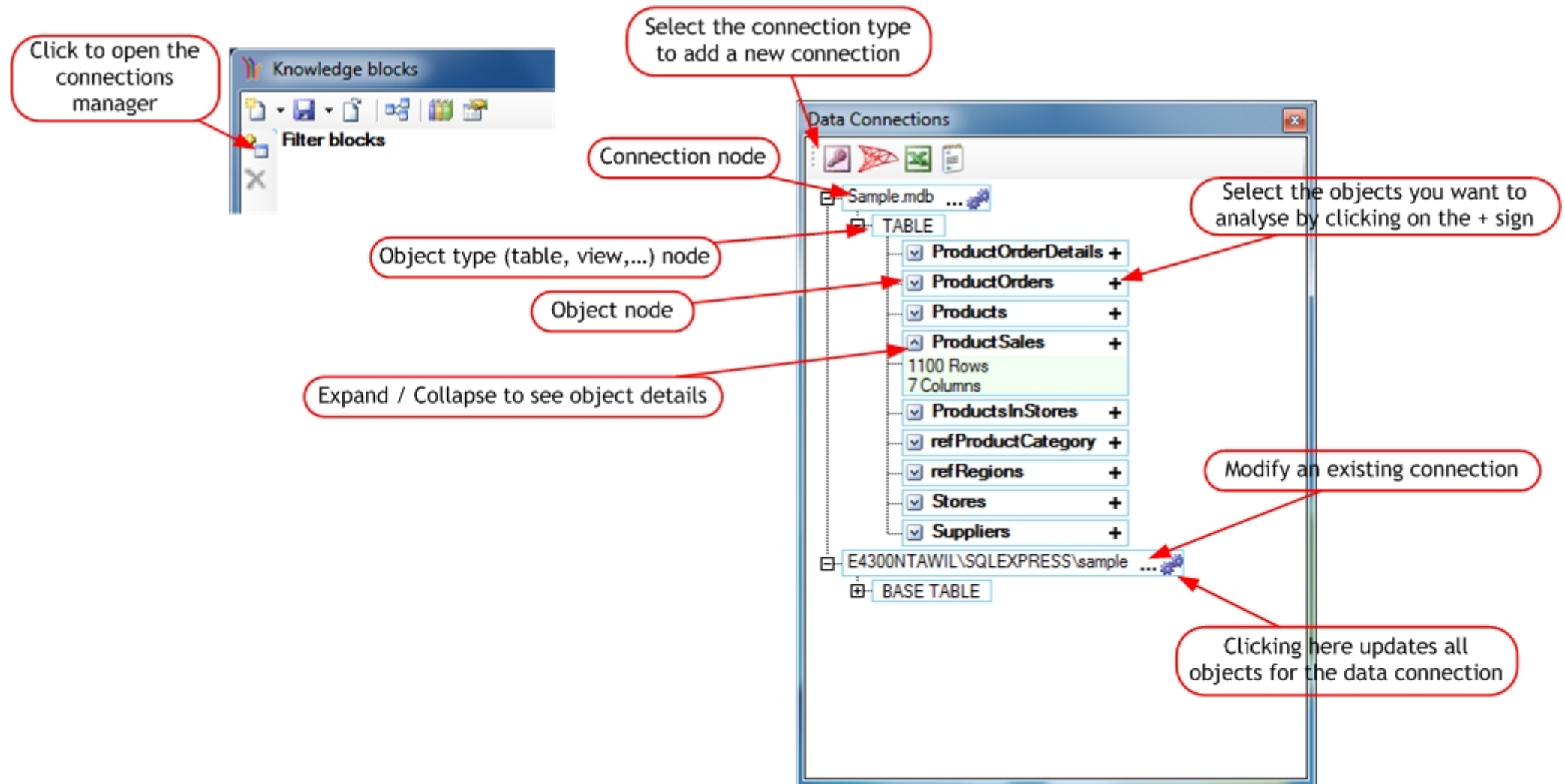
3.1 Microsoft Access

Knowledge blocks supports all versions of Microsoft Access. However, MS Access 2007 requires the "2007 Office System Driver" which contains the data connectivity components. This is a free download offered on the Microsoft website.

To add a Microsoft Access connection click on the first "Browse" button and select the MS Access database file.

MS Access supports 3 security modes:
- Standard security: Use this mode to connect without a username and password.
- Database password: If your database is protected by a password, then select this mode and enter the password.
- Workgroup security: This mode requires a username, password and a workgroup file.

Checking the "Save password" checkbox will also save the password in the knowledge blocks document.

**Caution: passwords are not encrypted before saving!**
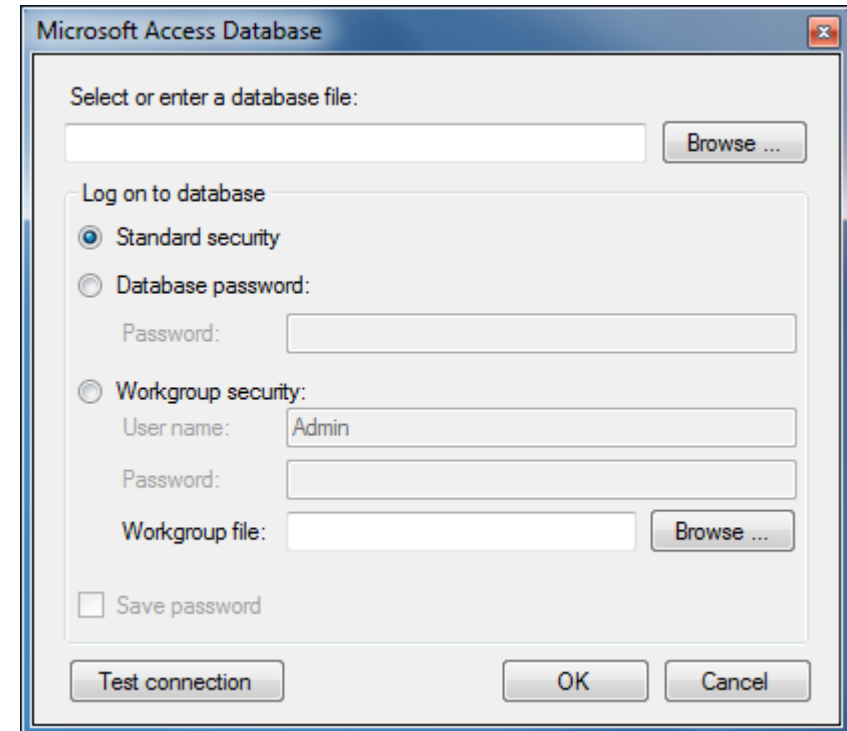
3.2 Microsoft SQL Server

The SQL Server connection dialog supports connections to SQL Servers or SQL database files. The connection type can be selected using the first dropdown list in the dialog.

Connecting to SQL Server:
Use the refresh buttons to populate the server and database dropdown lists
or enter Server and database names manually.
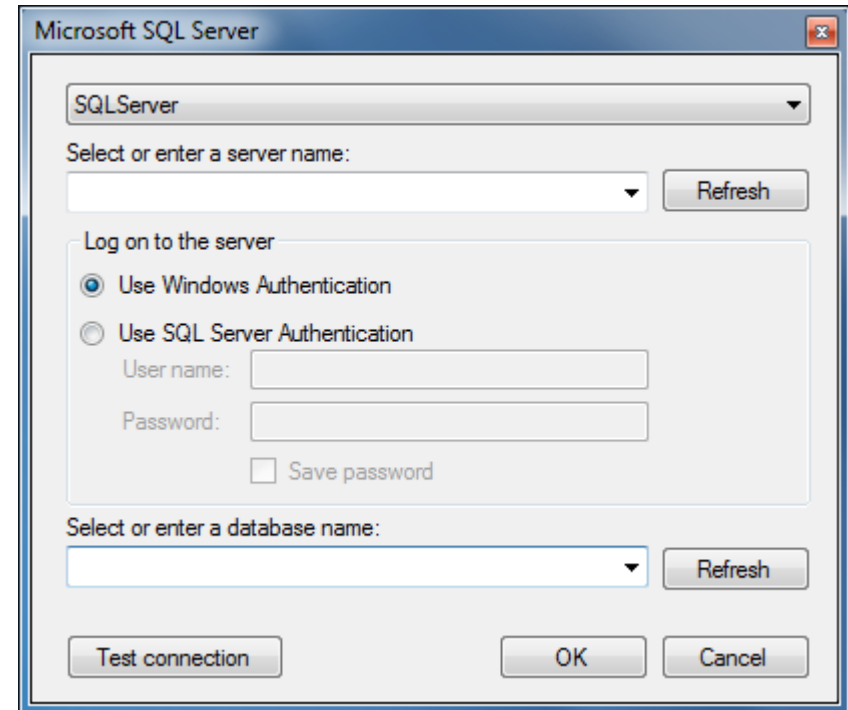
Security mode:
- Use Windows Authentication: This method uses the username and password that is used to log into the Windows Domain network.
- Use SQL Server Authentication: SQL Server authentication ensures that the username and password is checked against a login within the Microsoft SQL Server database.

Checking the "Save password" checkbox will also save the password in the knowledge blocks document.

Caution: passwords are not encrypted before saving!

Connecting to SQL Database file:
Use this selection to attach a database file to a local Microsoft SQL Server instance.

Select the database file by clicking on the browse button or enter the database file name manually in the textbox.

Security mode:
- Use Windows Authentication: This method uses the username and password that is used to log into the Windows Domain network.
- Use SQL Server Authentication: SQL Server authentication ensures that the username and password is checked against a login within the Microsoft SQL Server database.

Checking the "Save password" checkbox will also save the password in the knowledge blocks document.

Caution: passwords are not encrypted before saving!
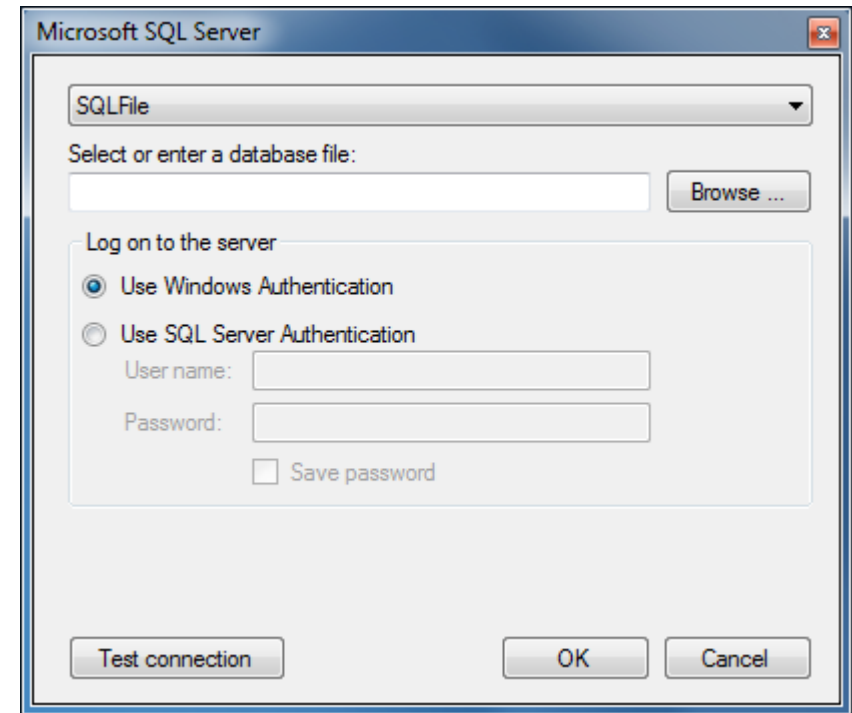
3.3 Microsoft Excel

Knowledge blocks also supports connections to Microsoft Excel files. Each table within the excel file is treated as a database table. Using the library designer, you can define relationships between tables.

Select the MS Excel file by clicking on the browse button or enter the file name manually in the textbox.

Check the "First row contains headers" checkbox if the column names are defined in the first row of the table.

Caution: Connections to password protected excel files is not possible!
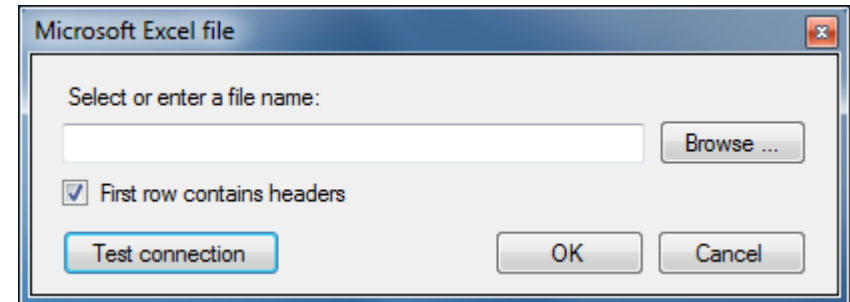
3.4 Text files

Use the Text files connection dialog to connect to text files. The selected folder is treated as a database and the individual text files in the folder are treated as database tables. Using the Library designer, you can define relationships between text files.

Select a folder by clicking on the browse button or enter the folder name manually in the textbox.

Check the "First row contains headers" checkbox if the column names are defined in the first row of the text file.
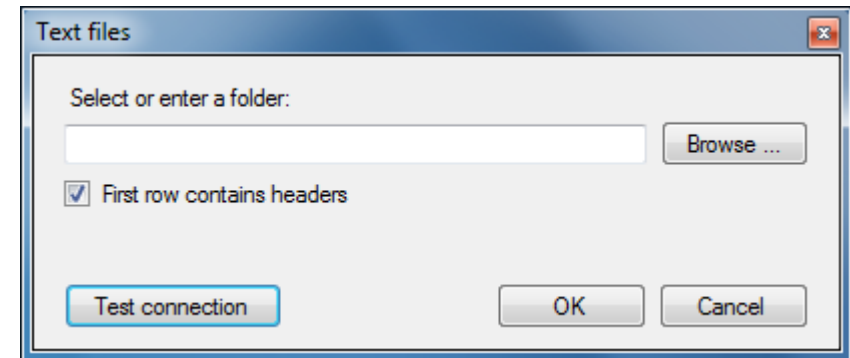
Text file format:
The text file format (file extension, delimiter character, …) are specified in the registry at the following location:
HKEY_LOCAL_MACHINE \ SOFTWARE \ Microsoft \ Jet \ 4.0 \ Engines \ Text

The 2 most important keys are:
- Extensions: defines which extensions are treated as text files. Default value: txt, csv, tab and asc
- Format: defines the delimiter character. Default value: CSVDelimited

You can change the default settings by editing the registry keys or defining a schema information file (schema.ini).

Schema ini file:
To access text files with different formatting or override registry keys, you need to create a schema information file (named Schema.ini) and place it in the same directory as the text data files.
The schema information file provides information about the general format of the file, the column name and data type information, and several other data characteristics.

Main Schema.ini characteristics:
- **The text file name:** The first entry in Schema.ini is always the name of the text file enclosed in square brackets. E.g. [Sample.txt]

- **The file format:** The Format option in Schema.ini specifies the format of the text file. The following table lists the valid values for the **Format** option.

Format specifier	Table format	Schema.ini entry
Tab Delimited	Fields in the file are delimited by tabs.	Format=TabDelimited
CSV Delimited	Fields in the file are delimited by commas (comma-separated	Format=CSVDelimited

	values).	
Custom Delimited	Fields in the file are delimited by any character you choose. All except the double quotation marks (") are allowed, including blank.	Format=Delimited(*custom character*)
Fixed Length	Fields in the file are of a fixed length.	Format=FixedLength

- **The field names, widths, and types:** You can specify field names in a character-delimited text file in two ways; Include the field names in the first row of the table and set **ColNameHeader** to **True or** specify each column by number and designate the column name and data type (Col*n*= *ColumnName type* Width #). The following table describes each part of the **Col***n* entry.

Parameter	Description
ColumnName	The text name of the column. If the column name contains spaces, you must enclose it in double quotation marks.
type	Bit Byte Short Long Currency Single Double DateTime Text Memo
#	Width of the column.

For detailed information about schema.ini files visit the Microsoft website http://msdn.microsoft.com/en-us/library/ms709353(VS.85).aspx

Schema.ini sample: Suppose we have 4 text files in our folder.

File 1: Sample1.txt
First row contains headers: Yes
Delimiter character: |

File 2: Sample2.log
First row contains headers: No
Delimiter character: " " (blank)

File 3: Sample3.txt
First row contains headers: Yes
Delimiter character: tab

File 4: Sample4.tab
First row contains headers: No

Format: Fixed length

Column name	Column width	Data type
ID	10	Number
CustomerName	50	Text
BirthDate	10	Date

Schema.ini file:

```
[Sample1.txt]
ColNameHeader=True
Format=Delimited(|)

[Sample2.log]
ColNameHeader=False
Format=Delimited( )

[Sample3.txt]
ColNameHeader=True
Format=TABDelimited

[Sample4.tab]
ColNameHeader=False
Format=FixedLength
Col1= ID Short Width 10
Col2= CustomerName Text Width 50
Col3= BirthDate DateTime Width 10
```

4 Library

Objects (tables, views,…) that are selected in the connection manager can be accessed from the library panel. Each column of each object is displayed as a filter block, which are grouped by object and by connection. The library also gives you access to the other two knowledge block types logical blocks and result blocks. Knowledge blocks can be added to the document panel by dragging and dropping.

Filter block characteristics:
The library also provides information about the characteristics of a filter block. A filter block is always based on a database column.

Data type:

Icon	Data type
A	Text
(calendar)	Date and time
1	Integer number
1,2	Decimal number
T/F	Boolean: True, False
(no entry)	Some columns (e.g. Binary object,…) cannot be filtered and are therefore not displayed in the library. These columns can be seen in the "Library Designer".

Column characteristics:

Icon	Data type
(key)	Primary key: The underlying column is the primary key in the table.
null	Nullable: The underlying column can contain empty (null) values.
(colored)	Unique: The underlying column contains unique values.

Editable characteristics:

- **Column name**: When adding filter blocks to a document, the column name is displayed in the filter expression. Double click on the column name in the library to edit the displayed name.
- **Use list of values**: Checking "Use list of values" will retrieve a list of all values in the table. This list can be used to select the compare value when setting the filter expression. The checkbox is automatically checked for indexed, non primary key, non unique columns.
- **Format**: The format drop down list is only available for blocks of type number or date and determines the formatting of the results in the document.
  Number formats: General, Formatted Number, Percent, Scientific or Currency.
  Date formats: Date and Time, Long date, short date
  The settings in the Windows "Regional Settings" determine the exact build-up of the formatting categories.
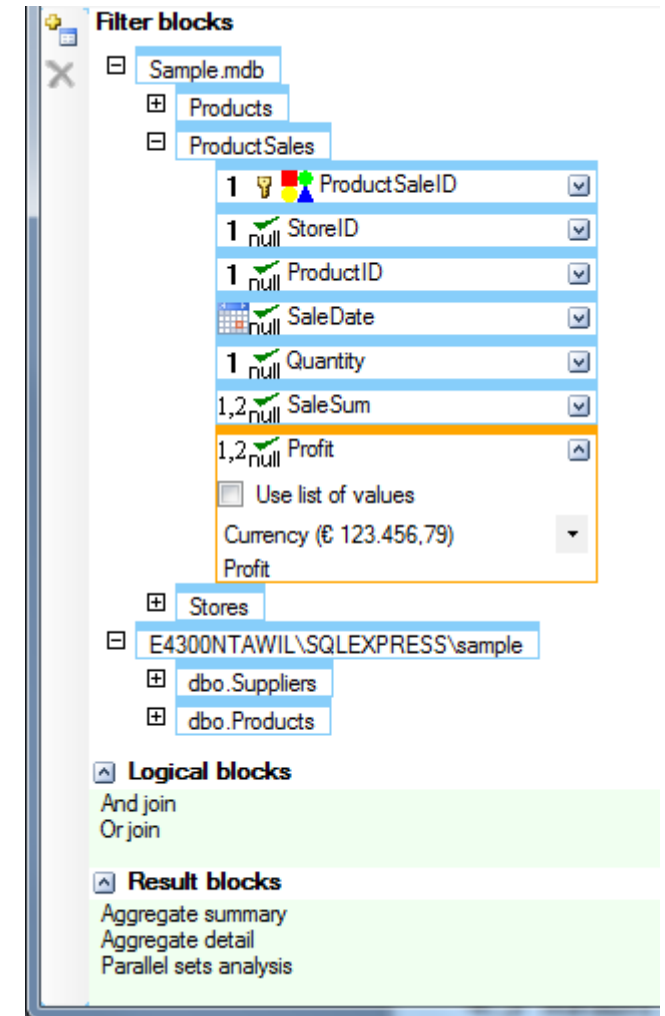
5 Library designer

The library designer shows all selected objects and all relationships between these objects. Here you can create new, edit existing or deactivate relationships. The library designer is also where you determine which columns are available for analysis.

Relationship information is used when connecting blocks and when adding aggregations to result blocks to ensure that only valid queries are generated. Filter blocks define filter criteria on database columns. When joining two or more filter blocks, knowledge blocks checks the library and only allows connections between blocks the are based on database columns belonging to one table or, if more than one table is involved, that all concerned tables are connected via database relationships. The same applies when connecting to result blocks or when adding aggregations to a result block.

Column visibility:
The library designer is used control which columns available in the library. By clicking on the circle left of the column name, the user can change the visibility of the columns. Green indicates that users can see the column in the library and can use it as a filter or can add the column to a result block.

Columns that have a red circle are excluded from the library panel and cannot be accessed by users.

Certain column types (e.g. binary columns) are always red and cannot be set to visible.
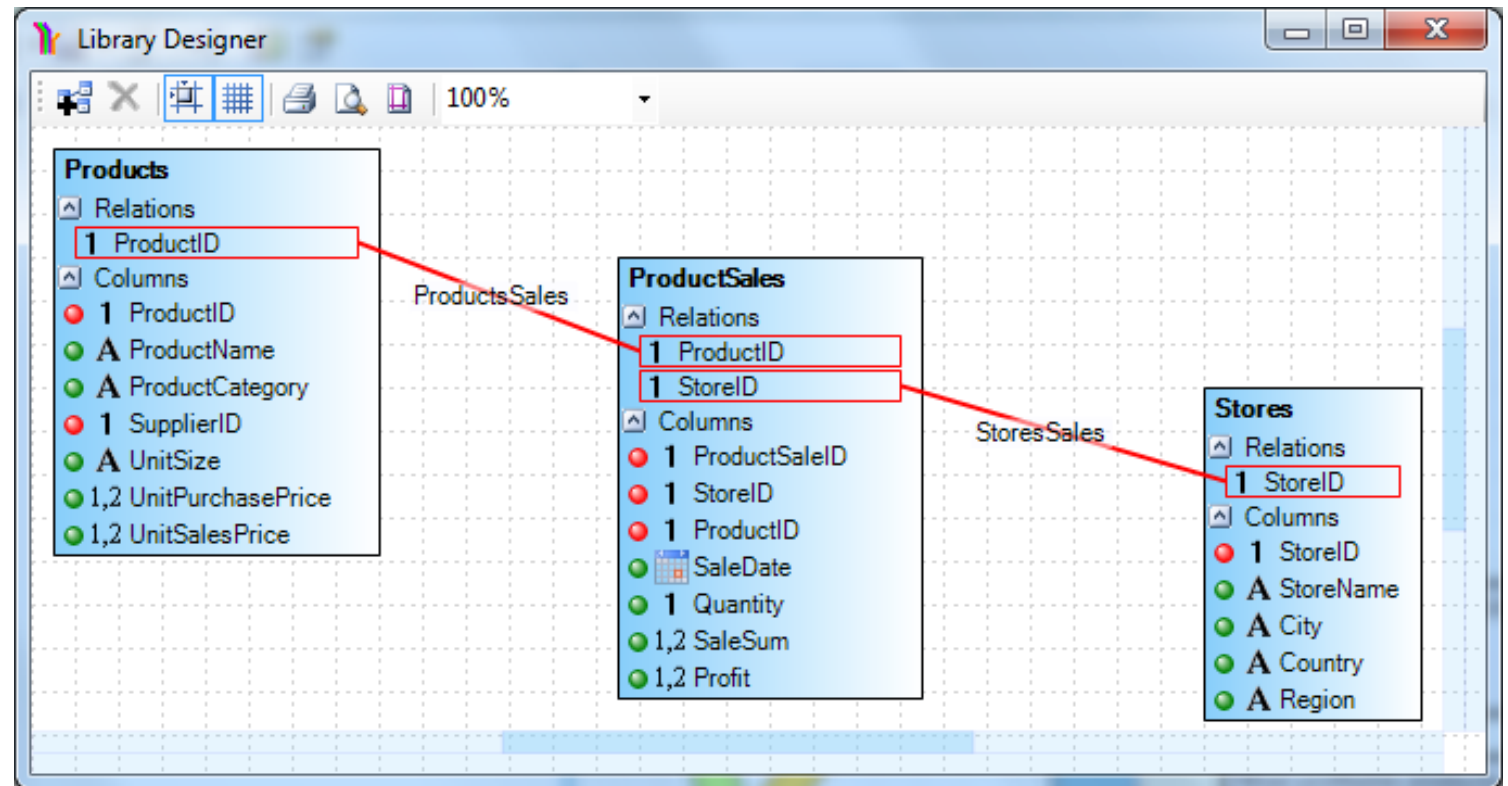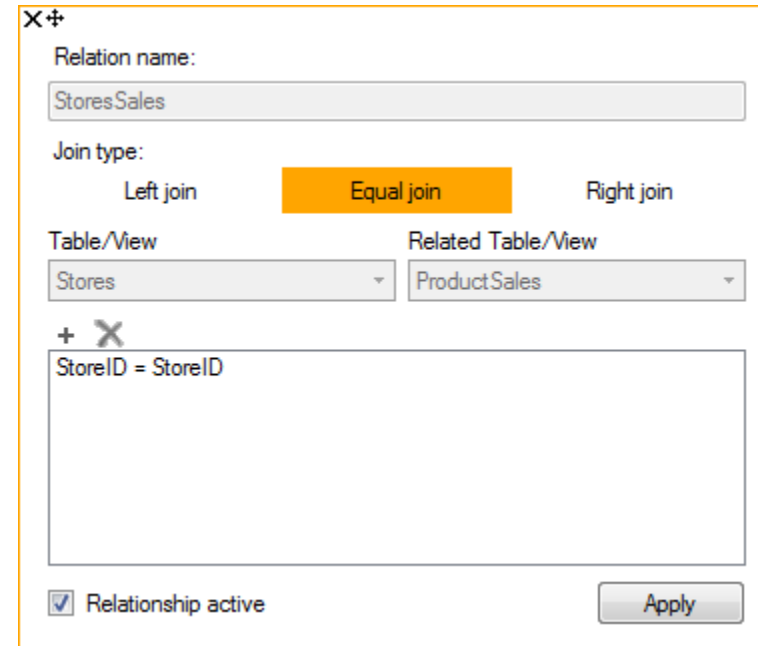
Relationships:
When adding objects to the library, the system also identifies and imports all relevant relationships. These relationships can be viewed and edited in the library designer. In addition, users can also define new relationships.

New relationships can be created by clicking on the "Add relationship" button in the designer toolbar. To edit an existing relationship, double click on the relationship name.

Relationship properties:
Knowledge blocks differentiates between two types of Relationships. The first are imported relationships; these are defined in the database and only certain properties can be edited. The second type is relationships the user creates using the library designer.

- **Relation name:** A name to identify the relationship. This property cannot be changed for imported relationships.
- **Join type:** Determines how data is retrieved from the tables;
    - **Left join:** Return all rows from the left table, even if there are no matches in the right table
    - **Equal join:** Return rows when there is at least one match in both tables
    - **Right join:** Return all rows from the right table, even if there are no matches in the left table
- **Table/View:** Sets the left table in the relationship. This property cannot be changed for imported relationships.
- **Related Table/View:** Sets the right table in the relationship. This property cannot be changed for imported relationships.
- **Columns:** List of all database columns used to connect the two tables. This property cannot be changed for imported relationships. The two buttons above the list can be used to add or delete columns.
- **Relationship active:** Used to activate / deactivate a relationship. Deactivated relationships are excluded when validating queries.

Relation name:

StoresSales

Join type:

| Left join | Equal join | Right join |

Table/View

Stores

Related Table/View

ProductSales

+ ✕

StoreID = StoreID

☑ Relationship active   Apply

Join type: The following is an example to show how the join type affects the rows retrieved. Suppose we have two tables that are connected using the "storeID" column.

Left table: "Stores" defines all stores:

storeID	country	region
1	Austria	Europe
2	France	Europe
3	Japan	Asia
4	USA	North America
5	Brazil	South America

Right table: "Orders" contains the orders submitted by the stores:

orderID	storeID	orderDate	orderSum
1	2	02.05.2009	182,22
2	1	12.05.2009	200,00
3	2	04.12.2008	502,50
4	5	06.09.2008	387,23
5	8	01.01.2007	798,00
6	4	07.07.2007	392,50

Using knowledge blocks, the user can drag a detail result block to the document panel and view all columns from both tables together. The join type set in the library designer determines the result set.

- **Equal Join:** is the default setting for a relationship and returns all stores that have placed at least one order:

Stores (Left table)

storeID	country	region
1	Austria	Europe
2	France	Europe
3	Japan	Asia
4	USA	North America
5	Brazil	South America

Orders (Right table)

orderID	storeID	orderDate	orderSum
1	2	02.05.2009	182,22
2	1	12.05.2009	200,00
3	2	04.12.2008	502,50
4	5	06.09.2008	387,23
5	8	01.01.2007	798,00
6	4	07.07.2007	392,50

Results

storeID	country	region	orderID	storeID	orderDate	orderSum
1	Austria	Europe	2	1	12.05.2009	200,00
2	France	Europe	1	2	02.05.2009	182,22
2	France	Europe	2	2	04.12.2008	502,50
4	USA	North America	6	4	07.07.2007	392,50
5	Brazil	South America	4	5	06.09.2008	387,23

- **Left join:** Return all stores and their orders. Stores (e.g. Japan) that have not placed any orders will have empty values in the order columns of the result.

Stores (Left table)

storeID	country	region
1	Austria	Europe
2	France	Europe
3	Japan	Asia
4	USA	North America
5	Brazil	South America

Orders (Right table)

orderID	storeID	orderDate	orderSum
1	2	02.05.2009	182,22
2	1	12.05.2009	200,00
3	2	04.12.2008	502,50
4	5	06.09.2008	387,23
5	8	01.01.2007	798,00
6	4	07.07.2007	392,50

Results

storeID	country	region	orderID	storeID	orderDate	orderSum
1	Austria	Europe	2	1	12.05.2009	200,00
2	France	Europe	1	2	02.05.2009	182,22
2	France	Europe	2	2	04.12.2008	502,50
3	Japan	Asia				
4	USA	North America	6	4	07.07.2007	392,50
5	Brazil	South America	4	5	06.09.2008	387,23

- **Right join:** Return all orders and the stores that placed the orders. Orders (e.g. Order number 5) that cannot be assigned to a store will have empty values in the stores columns.

Stores (Left table)

storied	country	region
1	Austria	Europe
2	France	Europe
3	Japan	Asia
4	USA	North America
5	Brazil	South America

Orders (Right table)

orderID	storeID	orderDate	orderSum
1	2	02.05.2009	182,22
2	1	12.05.2009	200,00
3	2	04.12.2008	502,50
4	5	06.09.2008	387,23
5	8	01.01.2007	798,00
6	4	07.07.2007	392,50

Results

storeID	country	region	orderID	storeID	orderDate	orderSum
1	Austria	Europe	2	1	12.05.2009	200,00
2	France	Europe	1	2	02.05.2009	182,22
2	France	Europe	2	2	04.12.2008	502,50
4	USA	North America	6	4	07.07.2007	392,50
5	Brazil	South America	4	5	06.09.2008	387,23
			5	8	01.01.2007	798,00

6   Knowledge blocks

Knowledge blocks are the basic elements used to analyse data. The library panel gives the user access to all knowledge blocks. Using drag and drop, users can add, connect and configure blocks in the document panel.

The system distinguishes between 3 categories of knowledge blocks:
- **Filter blocks** are used to set filters on individual columns or to directly compare data between 2 columns.
- **Logical blocks** allow the user to join 2 or more filter/logical blocks to create more complex filters.
- **Result blocks** are used to present the results.

6.1 Knowledge block structure

A knowledge block consists of a data lane and a text area. Knowledge blocks can be connected together by dragging the bottom handle (green circle) of the data lane and dropping it on another knowledge blocks data lane.

Filter blocks define the starting points of an analysis. A typical analysis has several starting points.

Each filter block data lane is assigned a color. This color is inherited to all following knowledge blocks, allowing users to easily identify which filter are active at the different stages of the analysis.

Logical and result blocks also have a data lane and text area. Result blocks can also be a starting point of an analysis. This can be done if you need to access all records without filtering.

The data lane section of logical and result blocks inherit their colors from all filter blocks further up the analysis. The color of each filter block is presented as a sub-lane in the block.

Logical blocks data lanes are also separated by a line to help identify the logical operation. AND operations are separated by a solid line. OR operations by a dashed line. This information is also inherited to all following blocks.
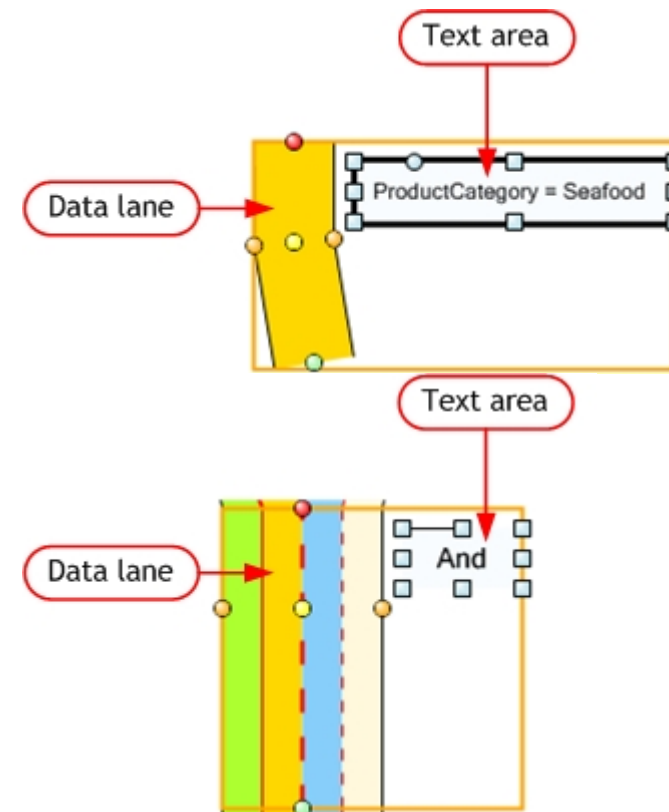
Common knowledge block Properties
The right panel (properties panel) of the applications main window can be used to set a knowledge block's properties. Properties are group together in categories. The first grouping displayed in the panel shows the "Appearance properties". These properties are shared by all knowledge block types. By selecting multiple blocks in the document panel, the user can set the appearance properties for all selected blocks.

Appearance properties: The following appearance properties can be set for all types of knowledge blocks.
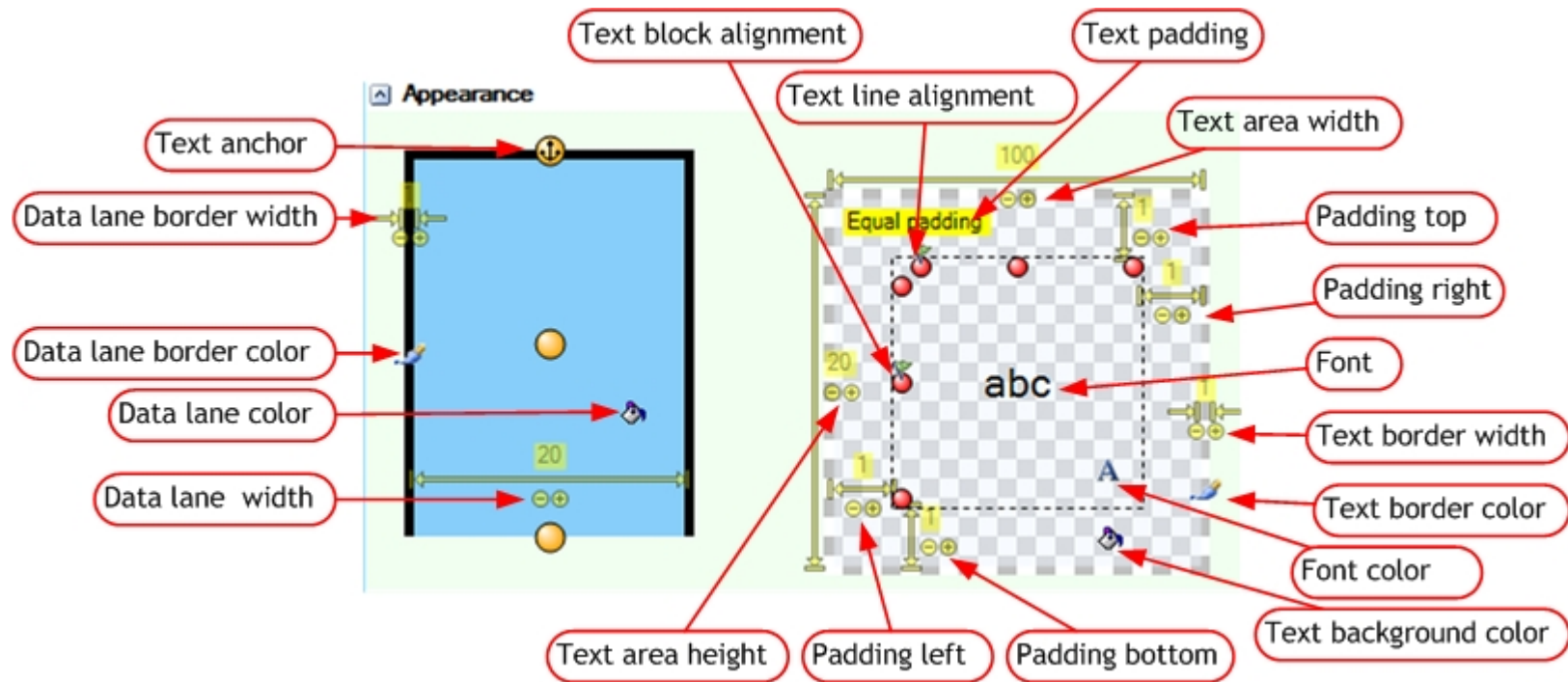
- **Text anchor**: Determines which part (top, centre or bottom) of the data lane the text component is anchored to. Moving the anchor point of a data lane will also move the text part of the block accordingly. Clicking on the orange circles changes the anchor position.
- **Data lane border width**: Click on the displayed width and enter a new value or click on the + / - signs to change the width.
- **Data lane border color**: Click on the icon to open the color selection menu and to set transparency.
- **Data lane color**: Click on the icon to open the color selection menu and to set transparency.
- **Data lane width**: Click on the displayed width and enter a new value or click on the + / - signs to change the width.
- **Text block alignment**: Click on one of the 3 red circles on the left to set the block alignment top, centre or bottom.
- **Text line alignment**: Click on one of the top 3 red circles to set the line alignment left, centre or right.
- **Text padding**: Click to switch between equal or individual padding values. Equal means that the same value is used for top, bottom, left and right padding.
- **Text area width**: Click on the displayed width and enter a new value or click on the + / - signs to change the text area minimum width.

- **Padding top**: Click on the displayed padding value and enter a new value or click on the + / - signs to change the top padding. If Equal padding is selected, then the new value is used for all 4 padding values.
- **Padding right**: Click on the displayed padding value and enter a new value or click on the + / - signs to change the right padding. If Equal padding is selected, then the new value is used for all 4 padding values.
- **Font**: Click on abc to change the font, style and size.
- **Text border width**: Click on the displayed width and enter a new value or click on the + / - signs to change the text border width.
- **Text border color**: Click on the icon to open the color selection menu and to set transparency.
- **Font color**: Click on the icon to open the color selection menu and to set transparency.
- **Text background color**: Click on the icon to open the color selection menu and to set transparency.
- **Padding bottom**: Click on the displayed padding value and enter a new value or click on the + / - signs to change the bottom padding. If Equal padding is selected, then the new value is used for all 4 padding values.
- **Padding left**: Click on the displayed padding value and enter a new value or click on the + / - signs to change the left padding. If Equal padding is selected, then the new value is used for all 4 padding values.
- **Text area height**: Click on the displayed width and enter a new value or click on the + / - signs to change the text area minimum height.

Additional appearance properties (Logical and Result blocks):
- **Inherit back color:** The inherit back color property can only be set for logical or result blocks. Click to activate / deactivate the property. When active, the blocks back color is determined by the color(s) set in the previous blocks.
  If you want to set the color of the block manually and inherit the new color to all following blocks, then deactivate the property.
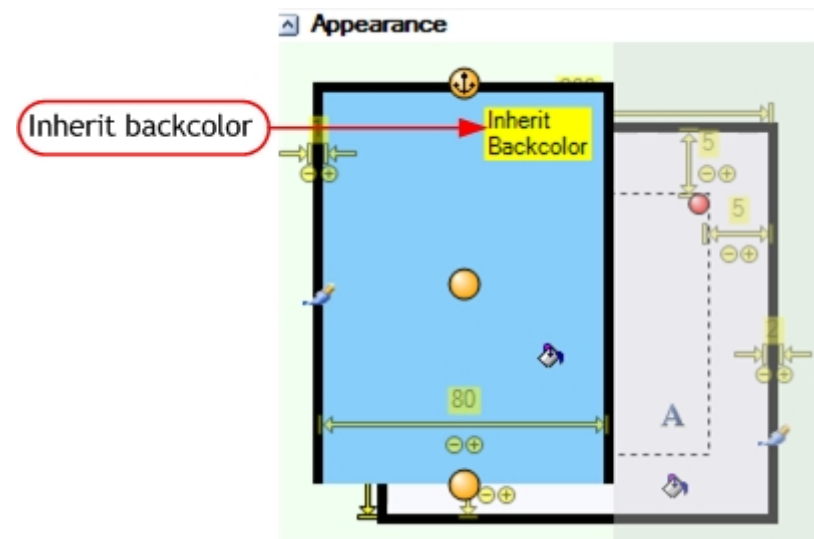
Additional appearance properties (Logical blocks):

The type of logical operation applied by a logical block can be identified by looking at the line running down the middle of the data lane.

- Dashed line indicates an AND operation.
- Solid line indicates an OR operation.

The two properties "Line width" and "Line color" set the width and color of the line and can only be applied to logical blocks. These settings are inherited to all following blocks.
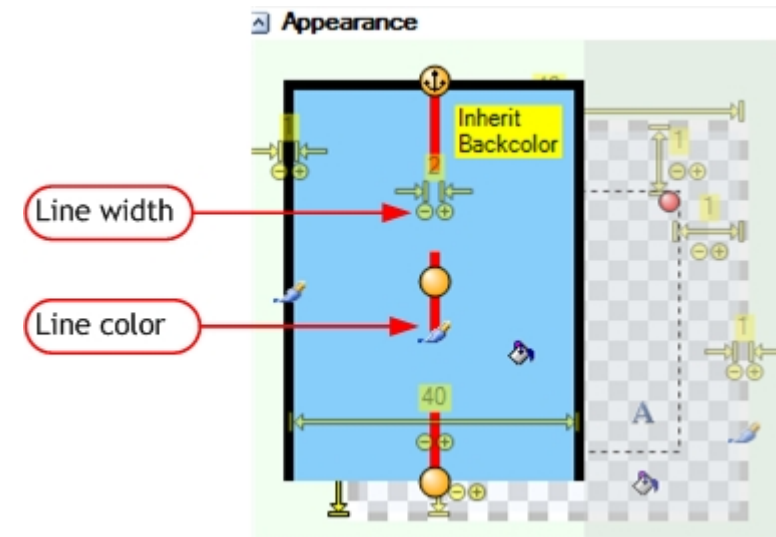
6.2 Filter block

A filter block defines a filter criteria on one database column. The comparison can be against a value or against another database column. When comparing to another database column, the values of both columns are compared at row level.

Filter block properties

- Compare type: Use the drop down list to select the compare type.
    - Value: compares the database column to a value.
    - Column: compares the content of the database column with the content of another column.
- Compare operation: can be set by selecting a value from the drop down list. The available compare operations depend on the column type:

Column type	Allowed operations
Text	=, >, >=, <, <=, <>, Starts with, Contains, Ends with, Is NULL, Is not NULL
Date	=, >, >=, <, <=, <>, Is NULL, Is not NULL
Number	=, >, >=, <, <=, <>, Is NULL, Is not NULL
Boolean	=, <>

- Compare value: text field to enter the value. Checking the "Use list of values" checkbox in the Library panel, changes the text box into a drop down list that contains all values in the database column.
- Date filter: this button is only available for date columns and opens a menu that allows you to select the compare date using a calendar or setting a variable date.
    - Variable date settings: use the drop down list to set the reference date (Current date, First day of week (Monday), First day of month, First day of Quarter, First day of Year)
    - +/-: the numbers entered in the five number fields are added / subtracted from the reference date.

    The variable date example to the right will always set the compare value to Friday 1 week ago. If today's date is Wednesday 3rd March 2010, then the compare value will be Friday 19th February 2010.
- Text area: The text displays the filter criteria set. The default setting is "<<COLUMN>> <<OP>> <<VALUE>>":
    - <<COLUMN>>: Displays the column name
    - <<OP>>: Displays the filter operation.
    - <<VALUE>>: Displays the compare value entered.

   The user can edit the text to help clarify the filter.
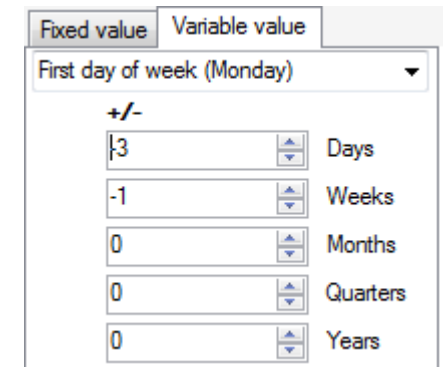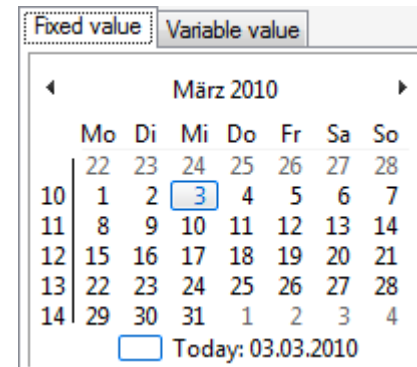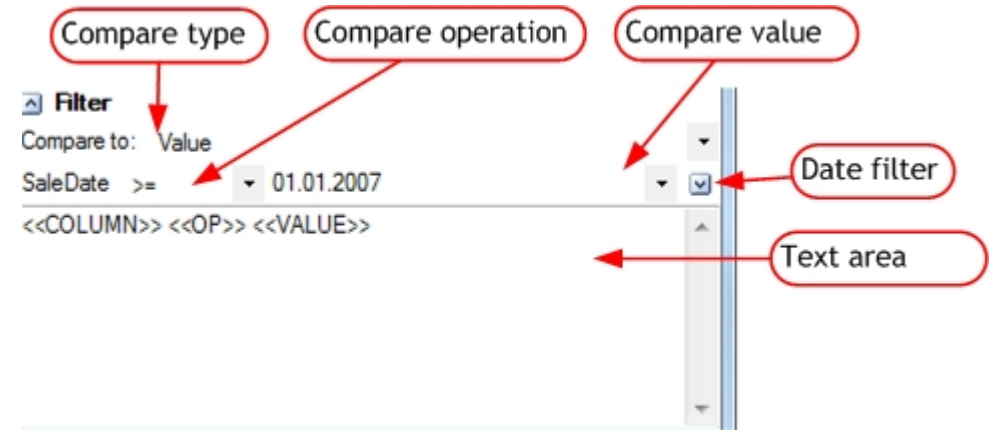
6.3 Logical block

Logical blocks provide logical operations that are used to connect two or more filter blocks. Knowledge blocks provides two kinds of logical blocks, AND and OR logical blocks.

Logical block properties
Besides the appearance properties described previously, logical blocks have an editable text area. The default text displayed is the logical operation. The properties panel can be used to edit the default text.

The logical operation is determined by the logical block type selected in the library panel and cannot be changed.
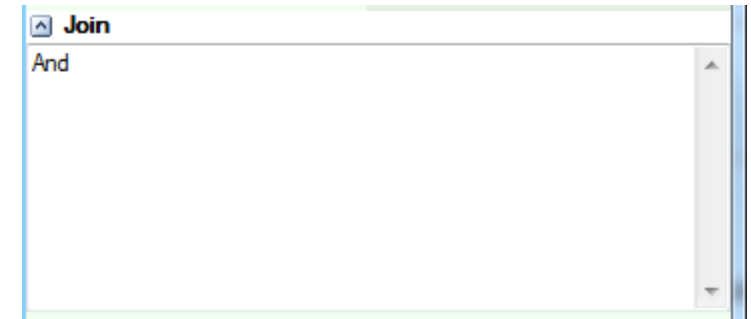
6.4 Result block

Result blocks are used to present the query results. The system provides three types of result blocks:

Summary result block
Summary blocks display one or more aggregations based on database columns.

- Text area: is used to combine aggregations with text. An aggregation handle is surrounded by "<<" and ">>" and formatted as FUNCTION_TABLE_COLUMN. The user can add text before and after an aggregation. Changing any of the characters in an aggregation will convert it to normal text.
- Add aggregation: click and navigate to the required column and select the aggregate function.
- Delete aggregation: Aggregations can be deleted by deleting the aggregation handle in the text are.
- Number of aggregations defined: this field shows how many aggregations have been defined. Aggregations that are used more than once within a block are only counted once.
- Update results: Click to query the database and replaces the aggregation handles with actual values.
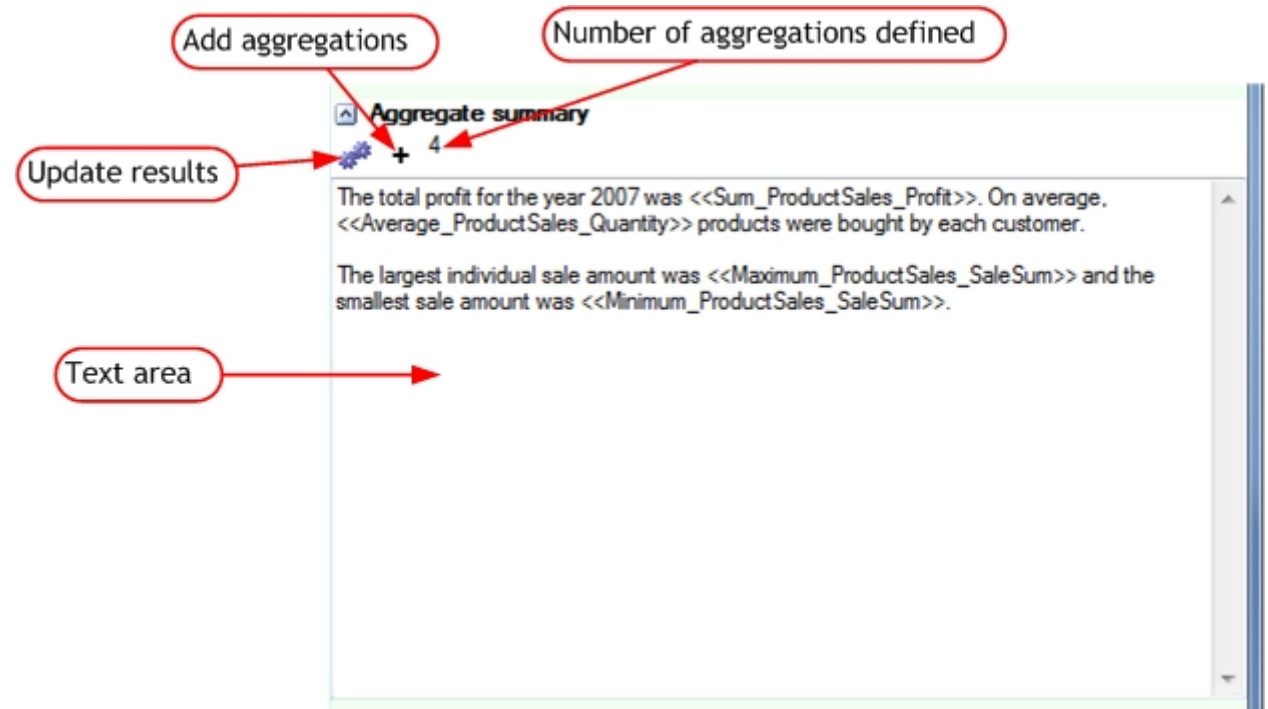
The example to the right will generate the following text block in the document panel:
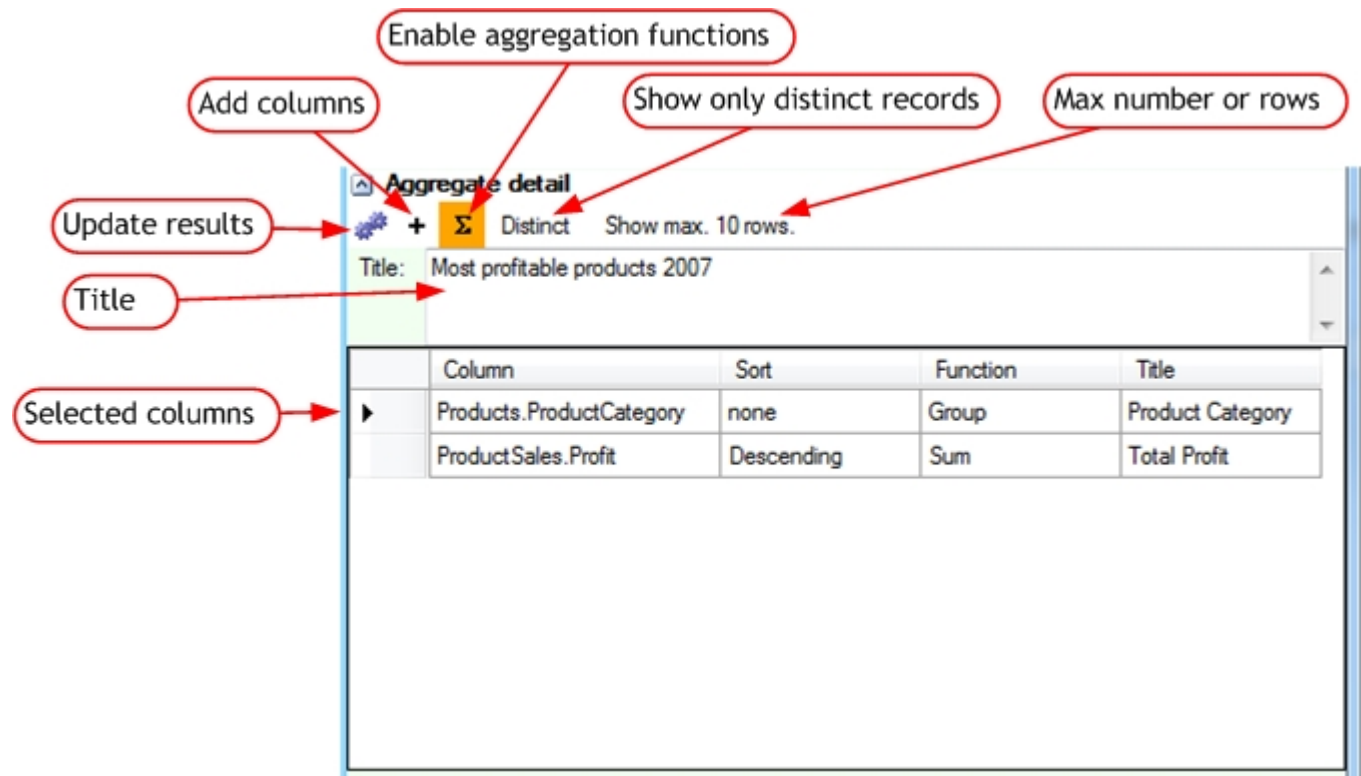
The total profit for the year 2007 was € 2.119,89. On average, 5,46 products were bought by each customer.

The largest individual sale amount was € 306,00 and the smallest sale amount was € 0,40.

Detailed result block

Detailed result blocks present data in tabular form. The data can be presented as individual records or as groupings with aggregations.

- Enable aggregation functions: Toggles the grouping and aggregation functions on or off.
- Show only distinct records: A queries result set can contain the same record more than once. If you want to display duplicate records only once, then activate the function by clicking on the button.
- Max number of rows: Specifies the maximum number of records displayed in the document panel. Select a predefined value from the list or enter the maximum number of rows manually.
- Title: Enter a title for the table in the text box.
- Add columns: click and navigate to the required column to add to the result set. The selected columns can be configured in the lower part of the properties dialog:
  - Column: Displays the column name and cannot be changed.
  - Sort: select the sorting type using the drop down list. When more than one sorting is defined, the order is then determined by the column order.
  - Function: Grouping and aggregation functions are only visible if the "Enable aggregation functions" is active. The available functions depends on the column type.
  - Title: The text box can be used to change the column header displayed in the result block.
- Delete columns: Select the entire row by clicking left of the column name and press the del key on the keyboard.
- Update results: click to query the database and update the results in the document panel.

The above example will generate the following result block:

The ordering and width of the columns can be changed by dragging the column headers and separators in the document panel.

Parallel sets analysis block

Parallel sets is a method for the visualization and interactive exploration of categorical data that shows data frequencies instead of the individual data points. The method is based on the axis layout of parallel coordinates, with boxes representing the categories and parallelograms between the axes showing the relations between categories.

- Add columns: click and navigate to the required columns to add as dimensions to the graph. Selected columns are displayed in the lower section of the properties panel, where they can be configured.
    - o Column: Displays the selected column and cannot be modified.
    - o Sort: If required, a sorting can be defined which sorts the data along the parallel axis.
    - o Group by: Date columns can be grouped by distinct values, year, year and quarter, year and month, year and week or year and day.
- Delete columns: Select the entire row by clicking left of the column name and press the del key on the keyboard.
- Main aggregate: the default aggregate is number of records. To change, click and navigate to the required column and select the aggregate function.
- Show results: When activated, the values of the main aggregate are also displayed in the graph.
- Show results as %: When activated, the values of the main aggregate are also displayed as percentage of totals in the graph.
- Main aggregate label: Edit the default value to display a different label.
- Title: Click in the text box and enter a title for the graph.
- Update results: click to query the database and update the result block in the document panel.

The above example will generate the following result block:
To reorder the dimensions, click and drag the dimension labels.

To interactively explore the data in the result block, move the
mouse over the boxes along the parallel axis or over the
parallelograms connecting the boxes to view detailed results.

Drag to reorder columns