

Vanilla

BiProfiler

Profiling package



Charles Martin and Patrick Beaucamp

BPM Conseil

Contact : charles.martin@bpm-conseil.com, patrick.beaucamp@bpm-conseil.com

October 2011

Document : BPM_Vanilla_BiProfiler_v4.0_EN.odt

4.0 Version

Summary

Introduction.....	3
BiProfiler Package.....	3
Main Frame.....	4
About this documentation.....	4
Start with BiProfiler.....	5
Connection definition.....	5
Settings for JDBC Connection.....	6
Setting for FMDT Connection.....	7
Analysis creation.....	8
Analysis definition.....	9
Rules sets creation.....	10
Analysis execution.....	13
Historic.....	17
Tags.....	17

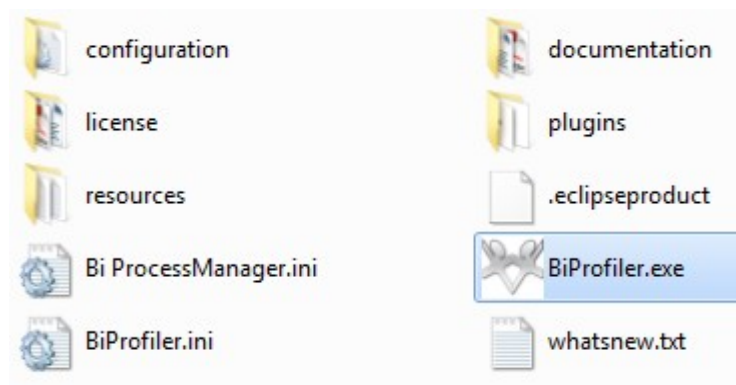
Introduction

The BiProfiler package offers interface for profiling manipulations about your datasources defined in Vanilla repositories.

BiProfiler Package

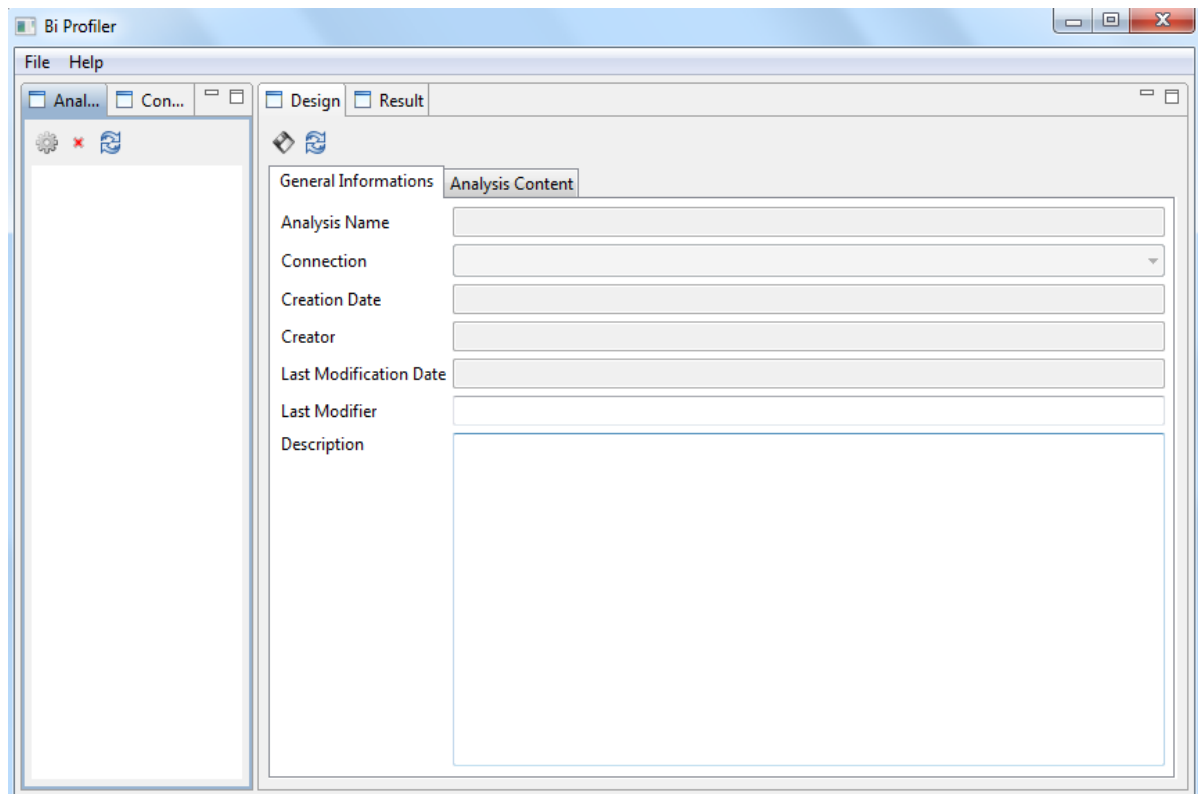
Once download :

- Unzip the BiProfiler package in a directory whose name do not contain a space character
- Run BiProfiler.exe



Main Frame

The main frame of the BiProfiler package is the following one :



The functions and manipulations are explained and detailed in this document.

About this documentation

This documentation is about the BiProfiler package (version 4.0, released January 2012).

It describes the main functions and general use of the BiProfiler package and puts aside minor bugs.

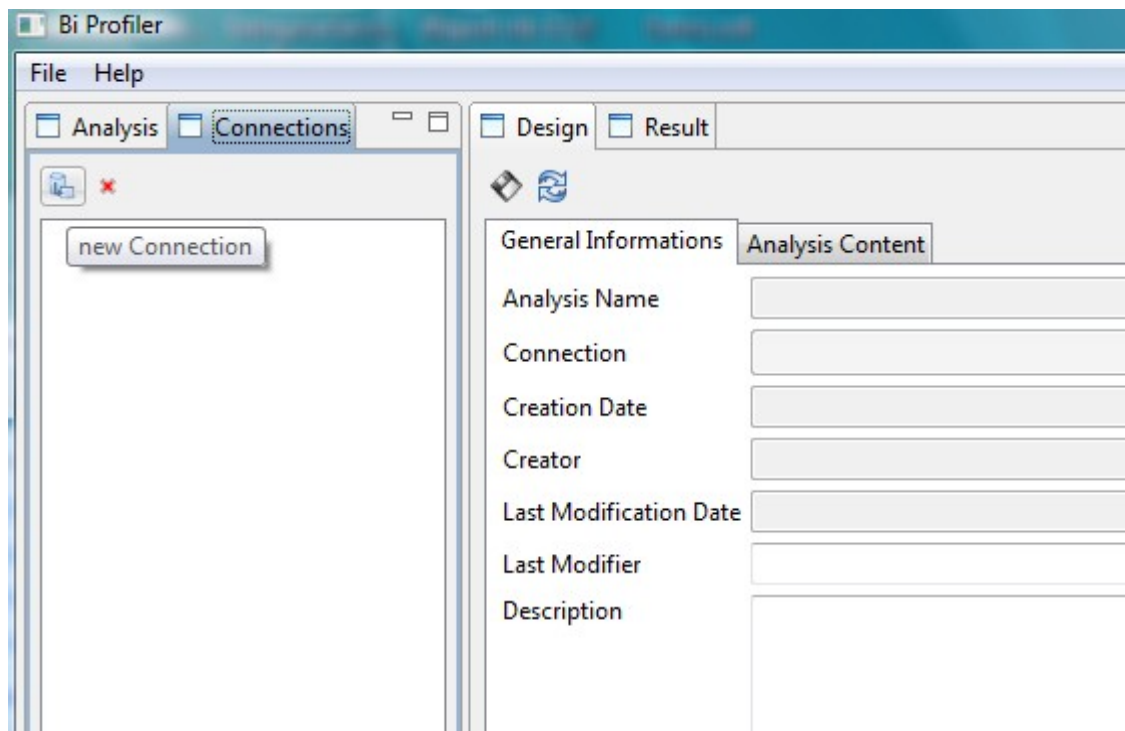
Start with BiProfiler

Connection definition

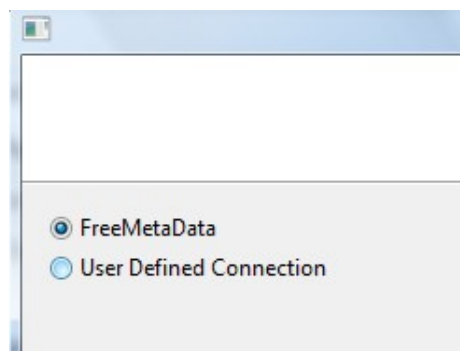
The first step to design an Analysis is to create a datasource.

The DataSource can be a native JDBC connection or a FMDT document stored in a Vanilla repository platform.

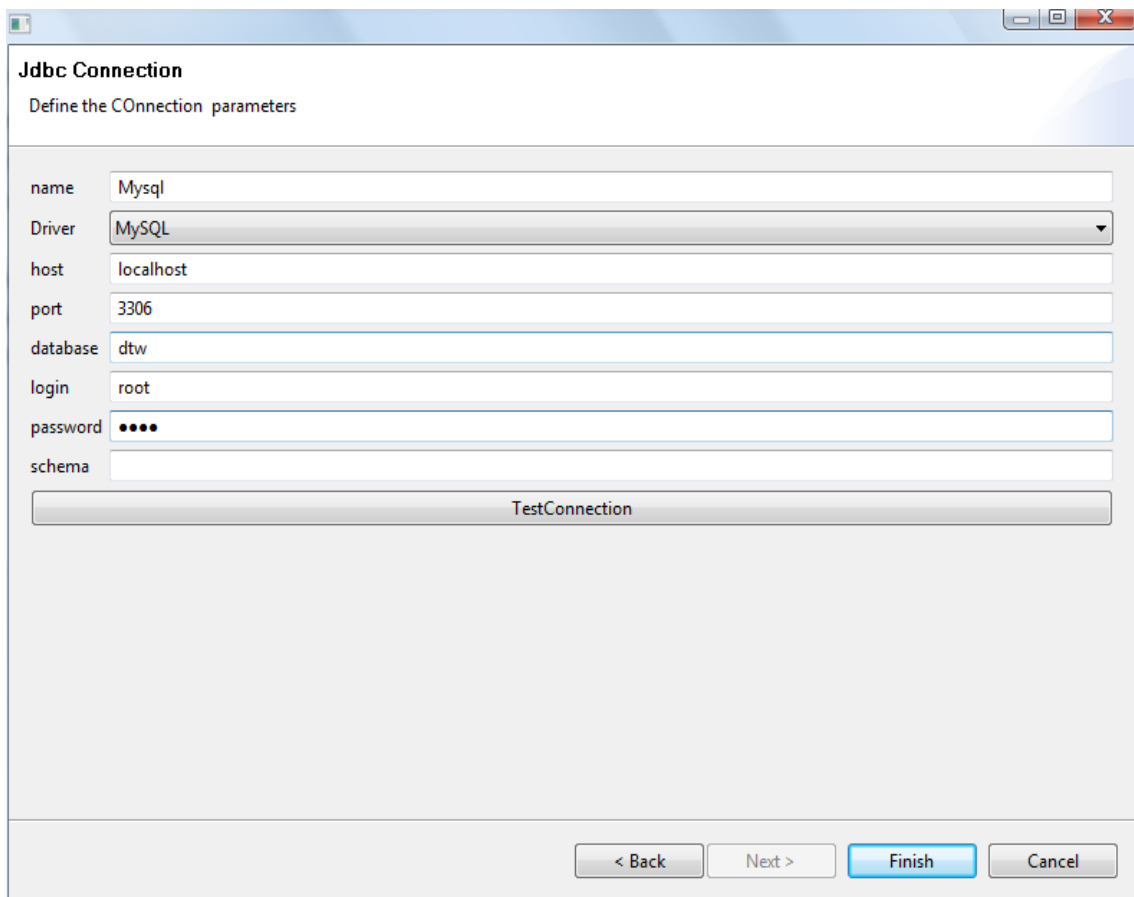
You just have to click on the “new Connection” button to launch the Connection Wizard.



Choose the datasource type :



Settings for JDBC Connection



The image shows a 'Jdbc Connection' dialog box with the subtitle 'Define the Connection parameters'. It contains several input fields for configuring a database connection. The fields are: 'name' (text box with 'Mysql'), 'Driver' (dropdown menu with 'MySQL'), 'host' (text box with 'localhost'), 'port' (text box with '3306'), 'database' (text box with 'dtw'), 'login' (text box with 'root'), 'password' (password field with four dots), and 'schema' (empty text box). Below these fields is a 'TestConnection' button. At the bottom of the dialog are four navigation buttons: '< Back', 'Next >', 'Finish' (highlighted in blue), and 'Cancel'.

Field	Value
name	Mysql
Driver	MySQL
host	localhost
port	3306
database	dtw
login	root
password	••••
schema	

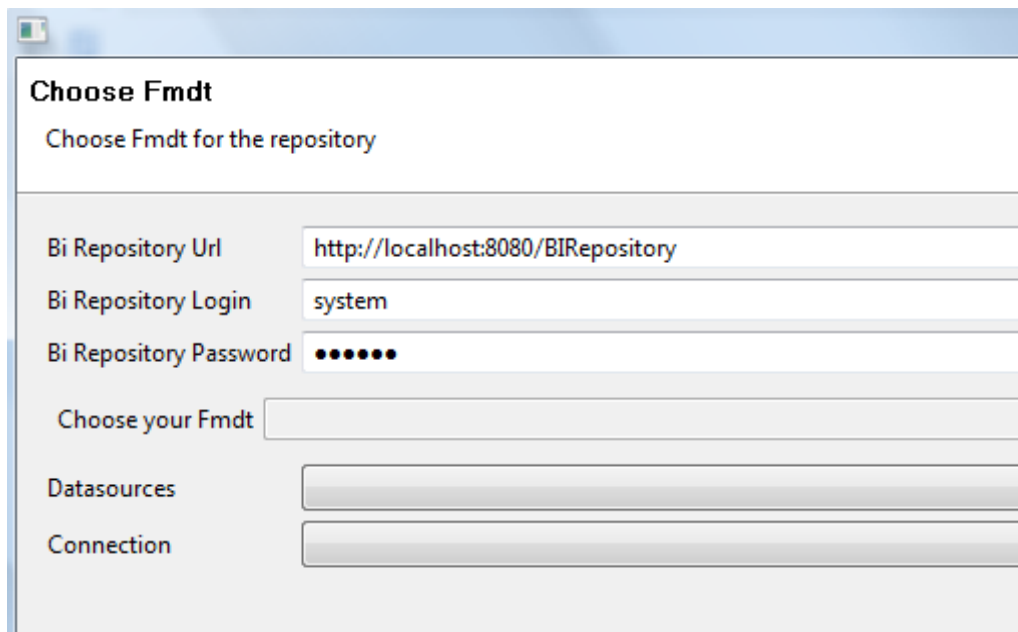
TestConnection

< Back Next > **Finish** Cancel

Setting for FMDT Connection

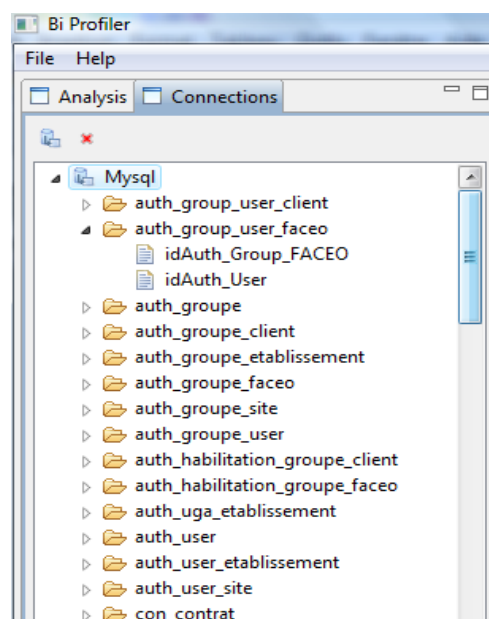
You just have to define the URL connection to your vanilla repository. Then select the FMDT that contains the datasource you want to analyse.

Once the model is loaded, you have to choose which datasource inside the FMDT you want to use, and its connection.



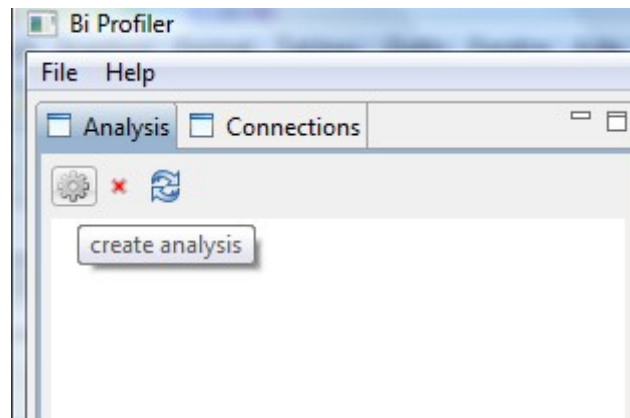
The screenshot shows a window titled "Choose Fmdt" with the subtitle "Choose Fmdt for the repository". It contains several input fields: "Bi Repository Url" with the value "http://localhost:8080/BIRepository", "Bi Repository Login" with the value "system", and "Bi Repository Password" with masked characters "••••••". Below these are three empty input fields labeled "Choose your Fmdt", "Datasources", and "Connection".

Once the connection is created, we can see in the Connection panel the connection and all the tables and columns that are inside.

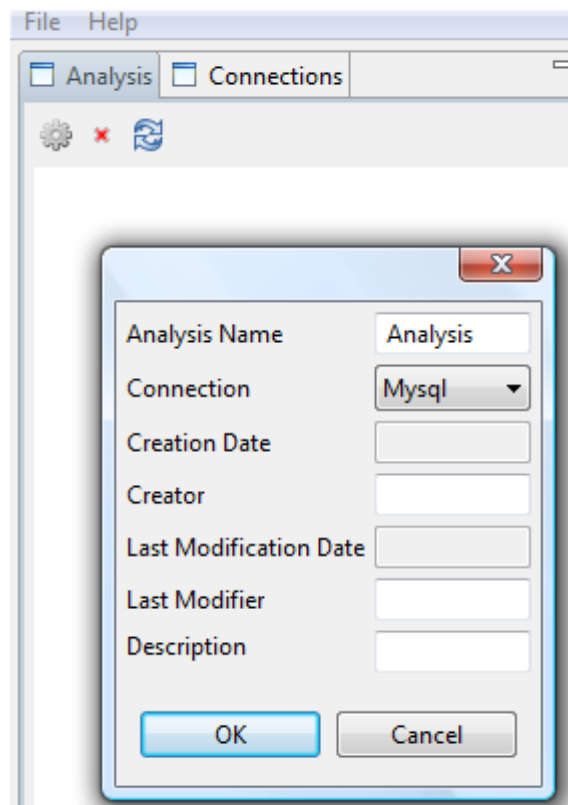


Analysis creation

The second step is now to create an Analysis from the Analysis panel.



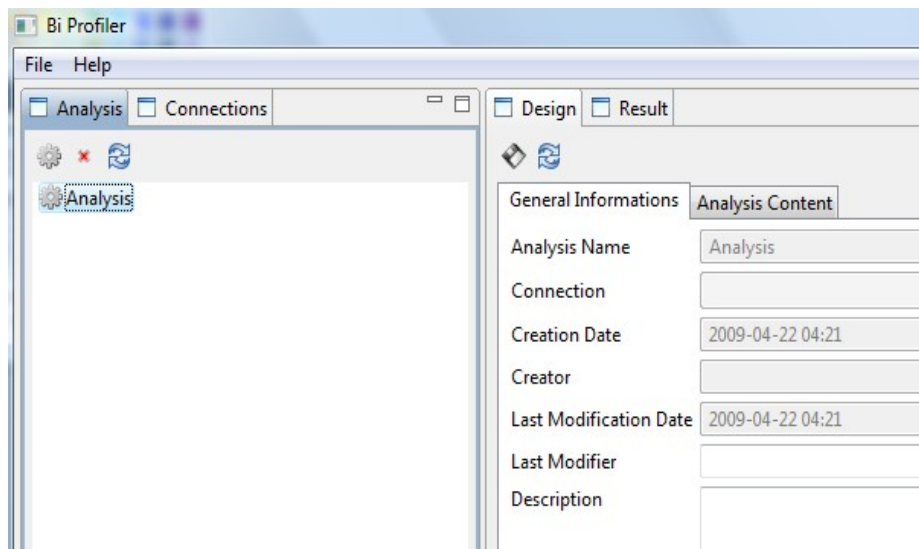
Just click on the “create Analysis” button.



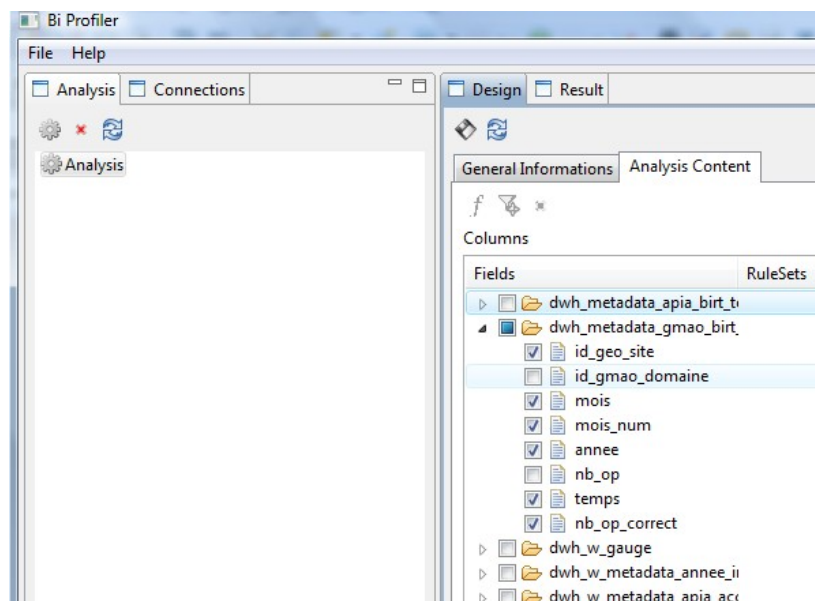
Fill the different fields and select the Connection that will support this Analysis

Analysis definition

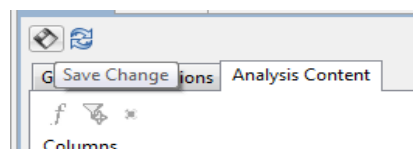
The next step is to define the analysis content from the Design panel.



You just have to check the Columns you want to analyse in the Analysis Content tab.



Once the columns are checked, you have to save the Analysis Content to be able to run the Analysis.



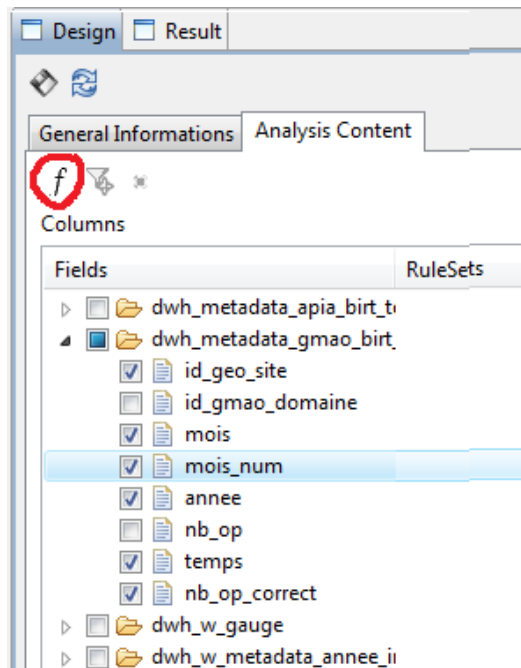
Rules sets creation

You can create some RuleSets on a column.

Each RuleSet is a combination of conditions that will be check on the data for the associated column.

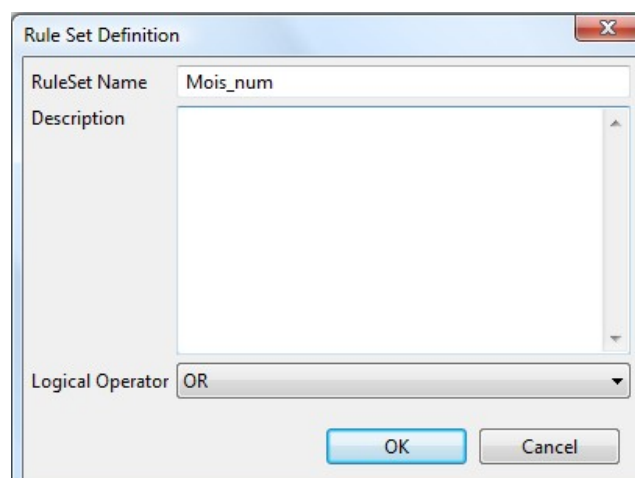
Note: you cannot create a RuleSet on a column you have checked but not saved.

To add a rule, select a checked column, and use the following button.

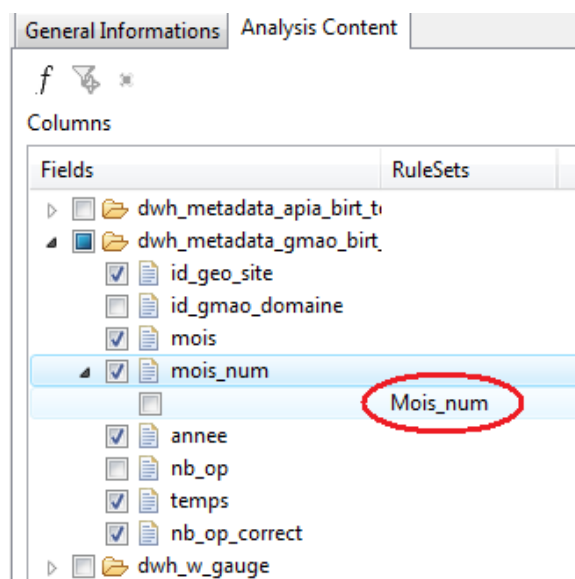


Define your RuleSet information.

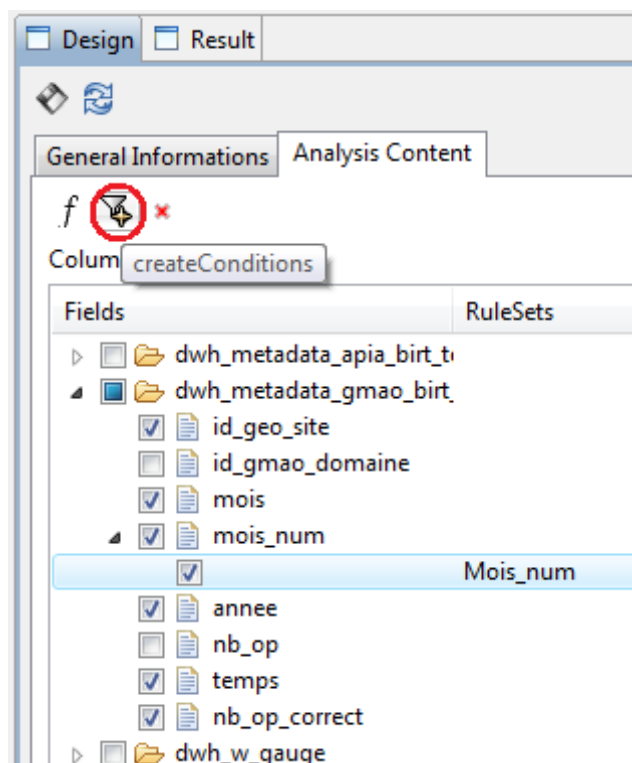
The logical operator is useful to perform mutliple conditions on the same RuleSet.



Once you click Ok, you should see the created RuleSet as a child of your column.

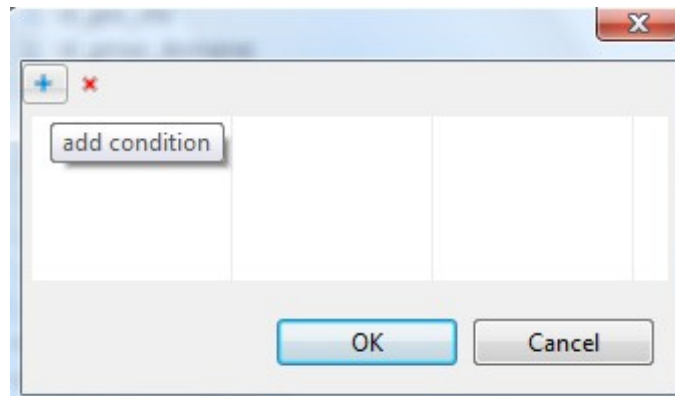


All we need to do is to add conditions in our RuleSet.



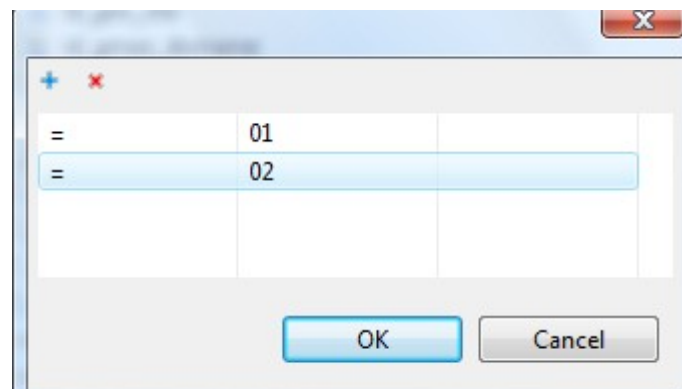
Select the RuleSet, then click on the “create Conditions” button

Then, click on the add button to add a condition.



There are 3 fields in the table :

- Operator
- Value 1 used by the operator
- Value 2 used by the operator

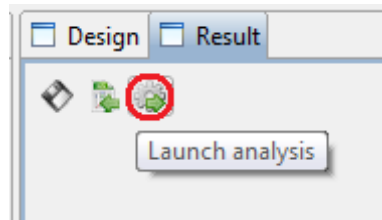


If the type is String or Date, there is no need to add some quote to enter values, they will be generated automatically at runtime.

Analysis execution

The last step is to run our Analysis.

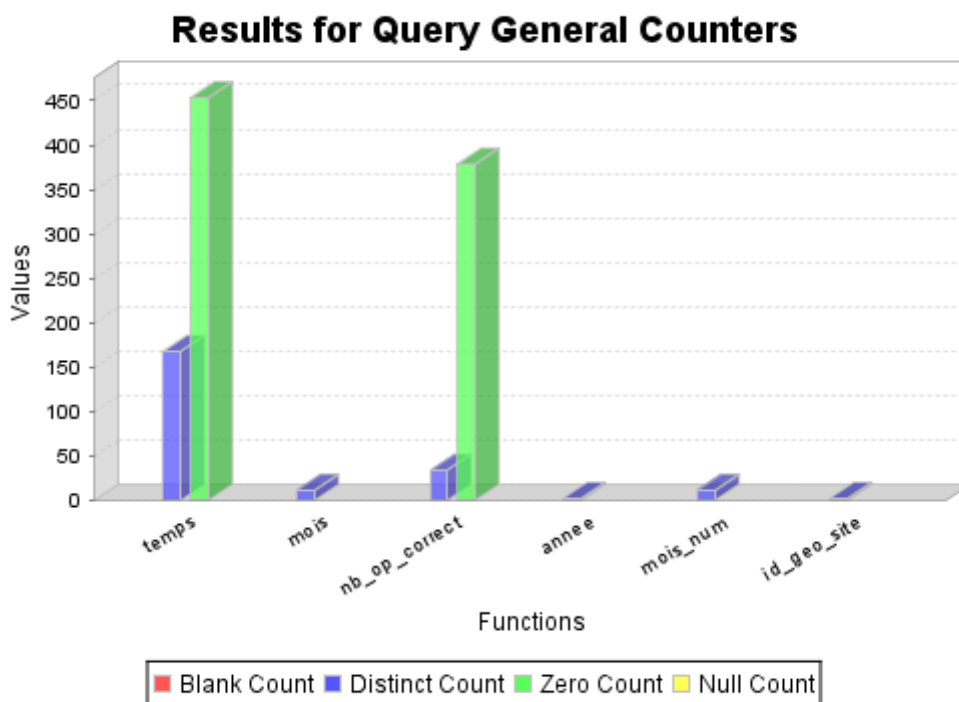
You can do it from the Result panel with the “Launch analysis” button.



You can browse the result in several folders, there will be one folder for the general information and one by RuleSet used by the Analysis.

General Results						
RuleSet Mois_num						
Field Name	Data Type	Distinct Count	Highest Value	Highest Value ...	Lowest Value	Lowest Value C...
temps	DECIMAL	167	1009.08	1	0.00	453
mois	VARCHAR(45)	12	Septembre	59	Aout	57
nb_op_correct	BIGINT	34	51	1	0	378
annee	VARCHAR(4)	2	2008	348	2007	357
mois_num	VARCHAR(2)	12	12	60	01	59
id_geo_site	INTEGER UNSIG...	2	1275	327	1198	378

Average Value	Blank Count	Blank Percent	Null count	Null Percent	Zero count
39,133			0	0 %	453
	0	0 %	0	0 %	
3,077			0	0 %	378
	0	0 %	0	0 %	
	0	0 %	0	0 %	
1 233,715			0	0 %	0



The table displays the different information on the data for each field mentioned in the analysis.

General Results

RuleSet Mois_num

Rule Set Informations

RuleSet Name

Mois_num

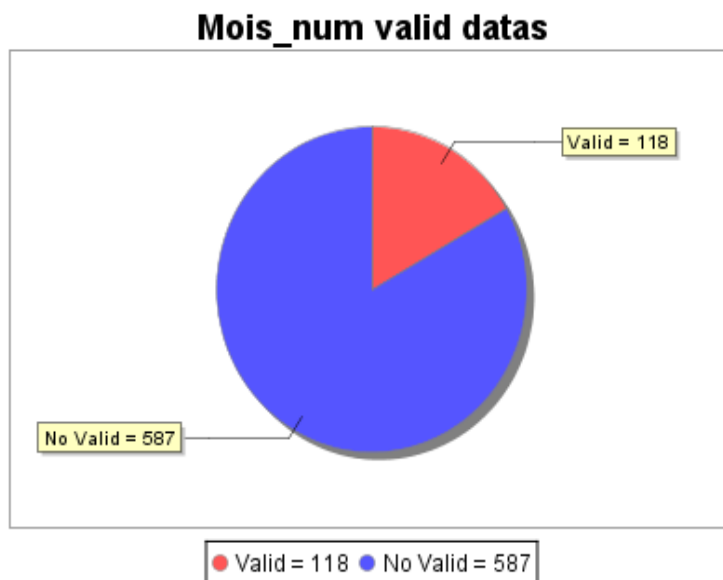
Description

Logical Operator

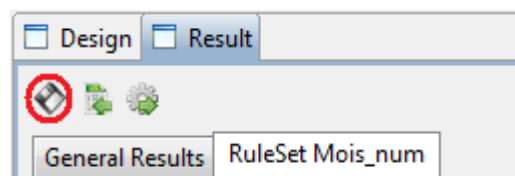
OR

Rule Set Results

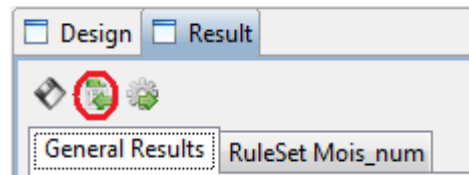
Operator	Value1	ValidCount	ValidPercent	DistinctValidCount	DistinctValidPercent
=	01	59	8 %	1	8 %
=	02	59	8 %	1	8 %
ALL		118	17 %	2	17 %



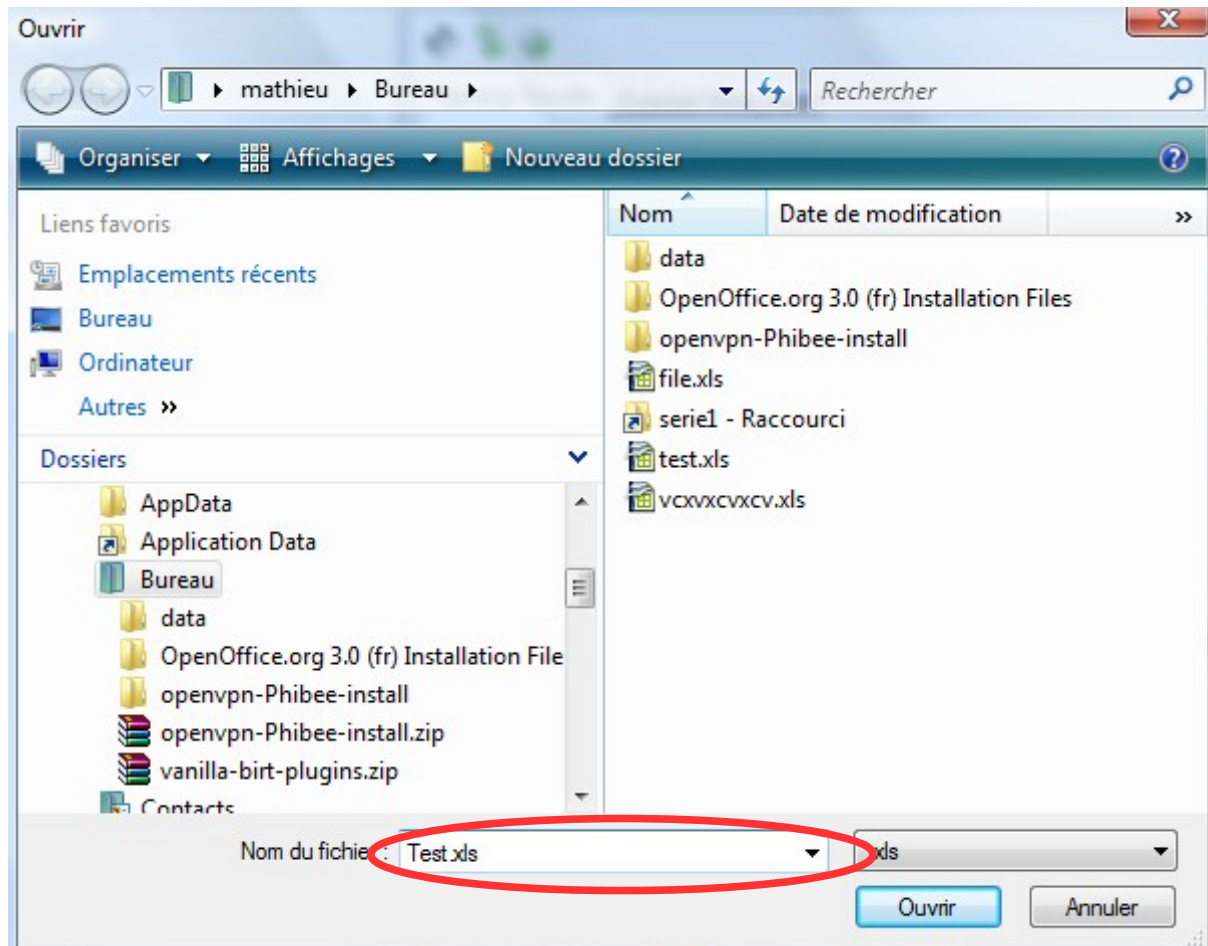
The 'save' button will save the results in the database.



You can also export the information into an Excel file.

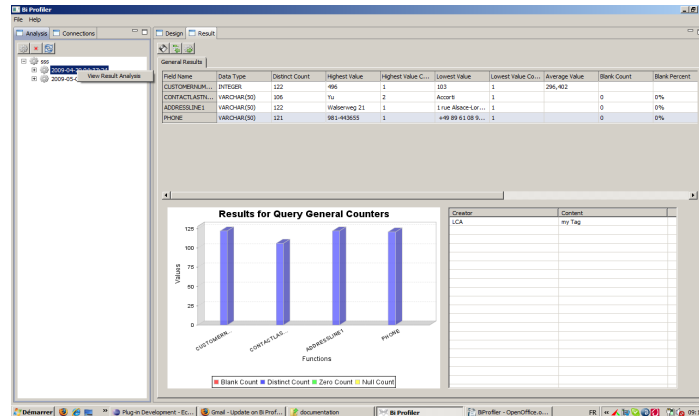


Define the place to save your excel file and its name.



Historic

If you already saved an analysis, you can now see the historic of the old results and edit them.



Tags

You can add some tags on each run of analysis before saving them by right clicking on the grid showing results. The tags can be added either on the general panel results or on the RuleSets grids.

