

Vanilla

BiProfiler

Outil d'analyse de sources de données



Charles Martin et Patrick Beaucamp

BPM Conseil

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

Octobre 2011

Document : BPM_Vanilla_BiProfiler_v4.0_FR.odt

Version 4.0

Table des matières

Introduction.....	3
Vue générale du package BiProfiler.....	3
Ecran principal.....	4
A propos de la documentation.....	4
Démarrer avec le BiProfiler.....	5
Définition de la connexion.....	5
Informations de connexion JDBC.....	6
Informations de connexion FMDT (FreeMetadata).....	7
Création de l'analyse.....	8
Définition de l'analyse.....	9
Création de Rules sets (Jeux de règles).....	10
Exécution de l'analyse.....	14
Historique.....	17
Mots clés.....	17

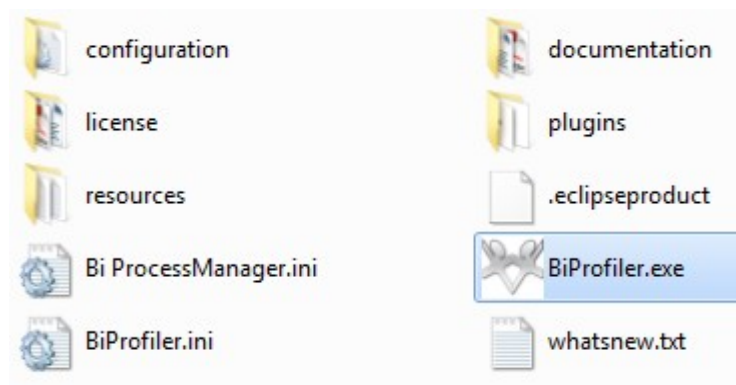
Introduction

Le package BiProfilier propose des interfaces pour des recherches et manipulations de profiling sur vos sources de données utilisées dans les référentiels Vanilla.

Vue générale du package BiProfilier

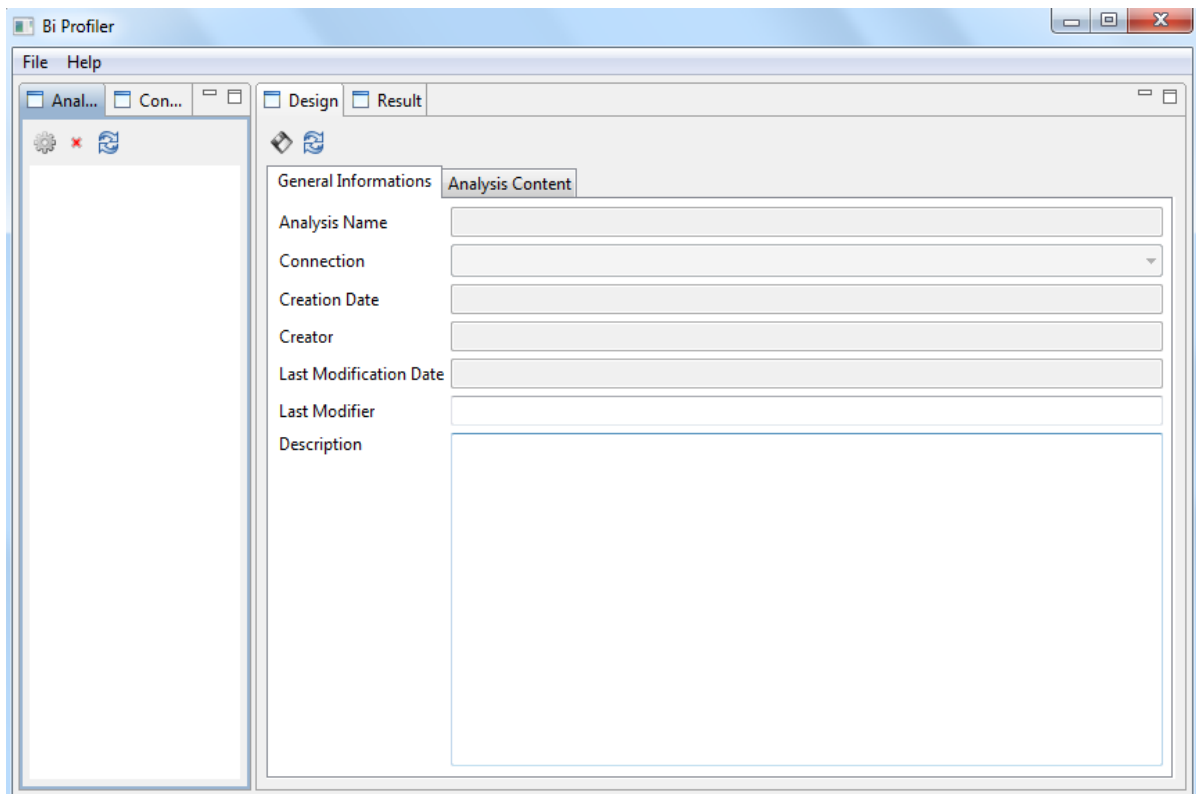
Une fois, le package téléchargé :

- Dézippez le package BiProfilier dans un répertoire dont le nom ne contient pas d'espace
- Exécutez BiProfilier.exe



Ecran principal

La fenêtre principale du module BiProfiler est la suivante :



Les fonctions et manipulations sont expliquées et détaillées dans la suite de ce document.

A propos de la documentation

Cette documentation porte sur le module BiProfiler (version 4.0, sortie Janvier 2012).

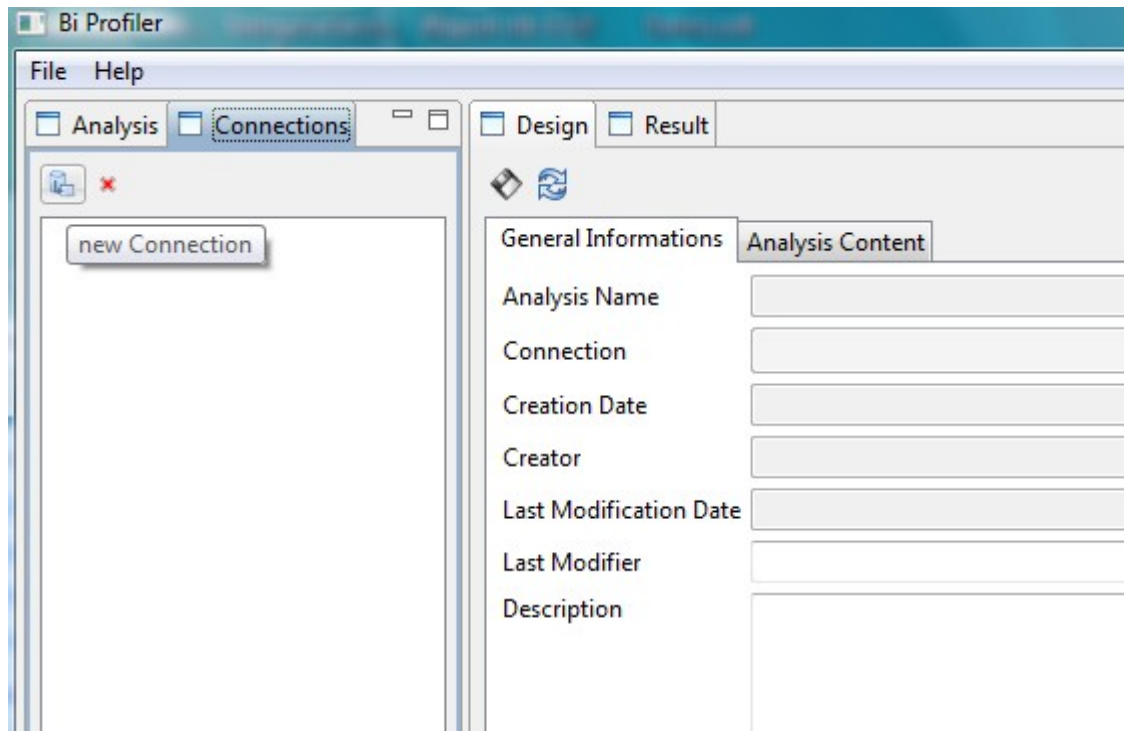
Elle décrit les fonctions principales et l'utilisation générale du package BiProfiler sous réserve de bugs mineurs.

Démarrer avec le BiProfiler

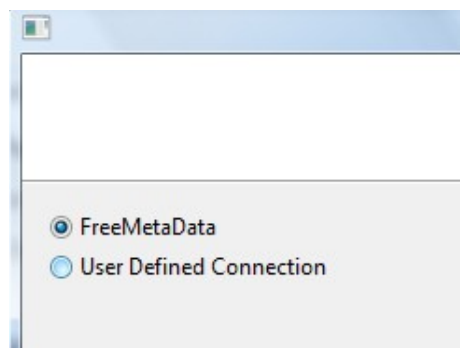
Définition de la connexion

La première étape de création d'une analyse est la création d'une source de données. La source de données créée peut être une connexion JDBC ou bien un document FMDT (FreeMetadata, voir la documentation relative) sauvé dans un référentiel Vanilla.

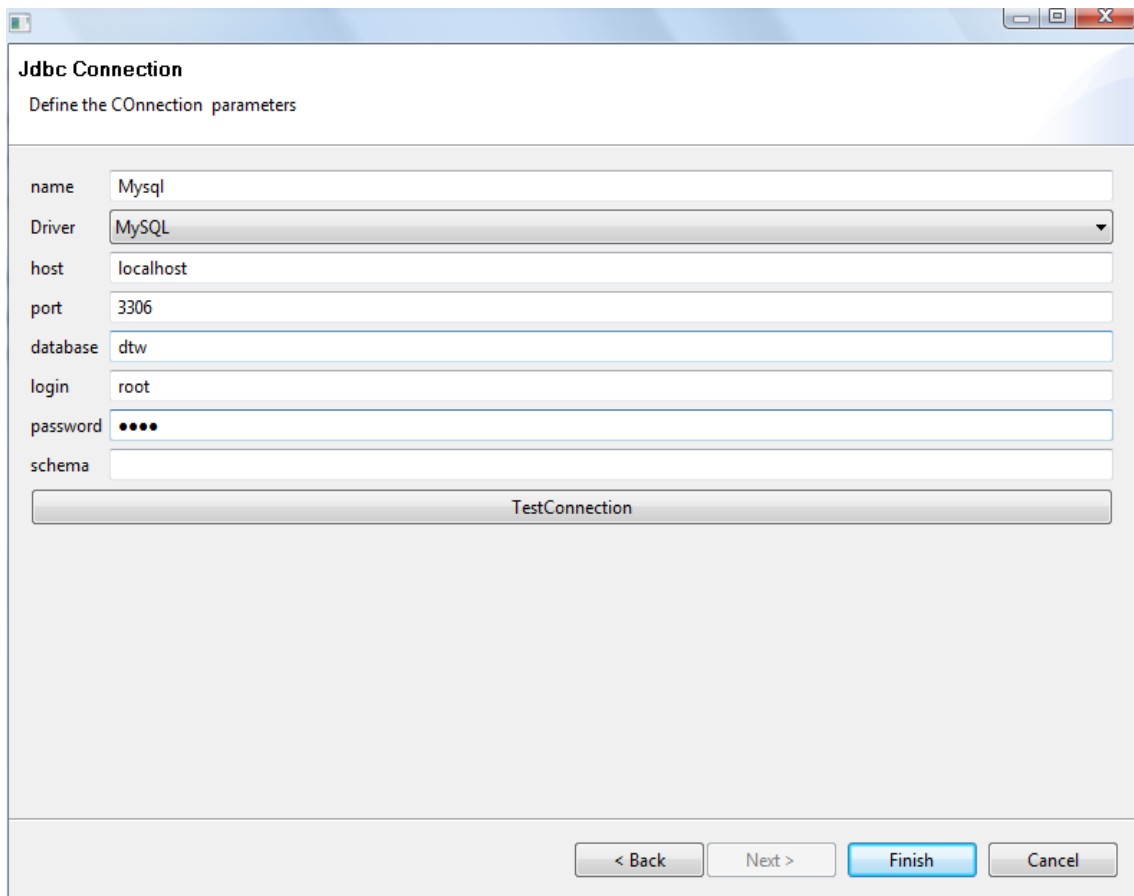
Cliquez sur le bouton 'new Connection' afin d'ouvrir les interfaces de création.



Choisissez un type de source de données :



Informations de connexion JDBC



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 right are four navigation buttons: '< Back', 'Next >', 'Finish' (highlighted in blue), and 'Cancel'.

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

TestConnection

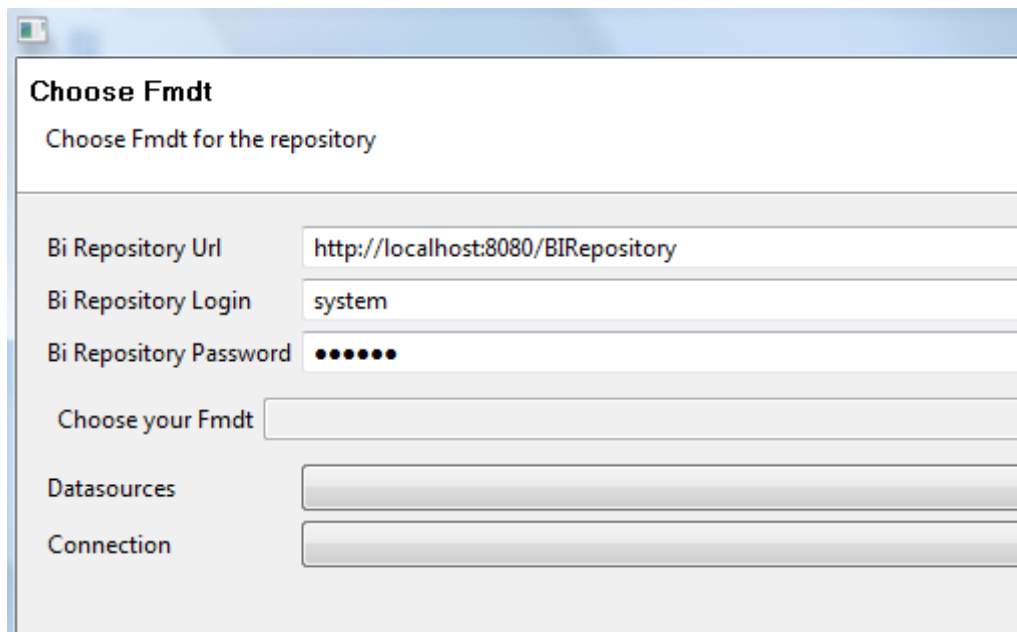
< Back Next > Finish Cancel

Informations de connexion FMDT (FreeMetadata)

Renseignez l'URL de votre référentiel et les informations de connexion.

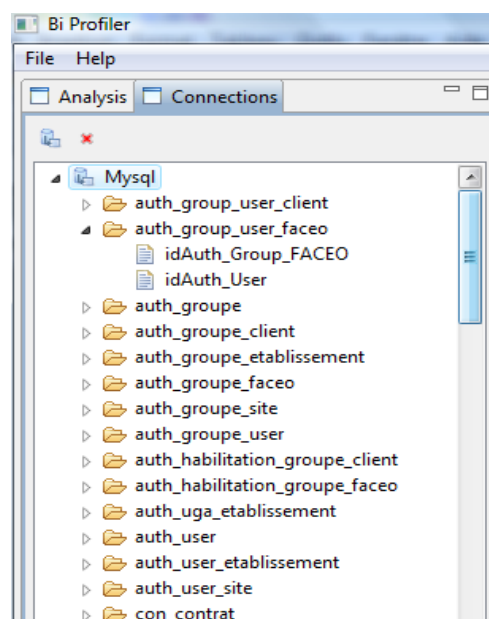
Sélectionnez ensuite votre modèle FMDT qui contient la source de données à analyser.

Une fois le modèle choisi, vous devez choisir quelle source de données à l'intérieur même du modèle vous souhaitez utiliser.



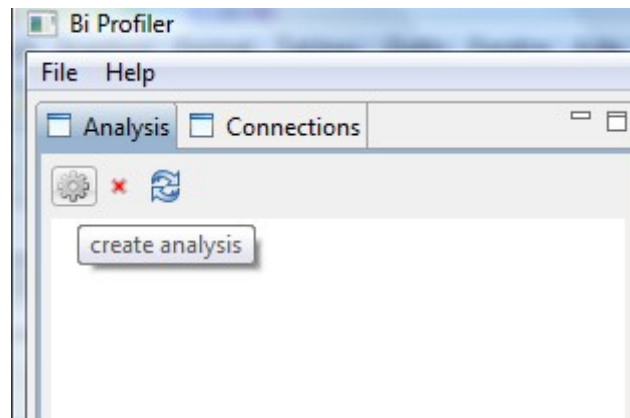
The screenshot shows a window titled "Choose Fmdt" with the subtitle "Choose Fmdt for the repository". It contains several input fields: "Bi Repository Url" with the value "http://localhost:8080/BIRepository", "Bi Repository Login" with the value "system", and "Bi Repository Password" with masked characters.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

Une fois la connexion créée (JDBC ou FMDT), les tables et les colonnes sont affichées dans le panneau 'Connection'.

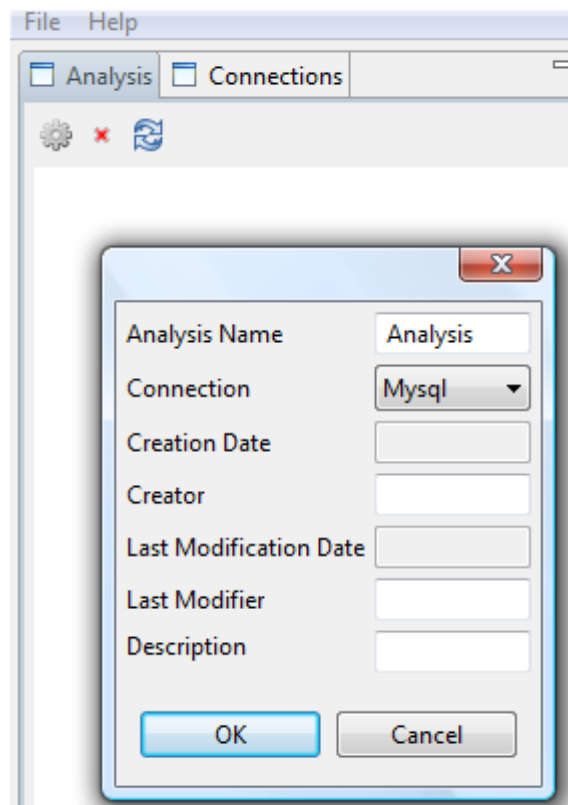


Création de l'analyse

La seconde étape dans la création d'une analyse se déroule depuis l'onglet 'Analysis'.



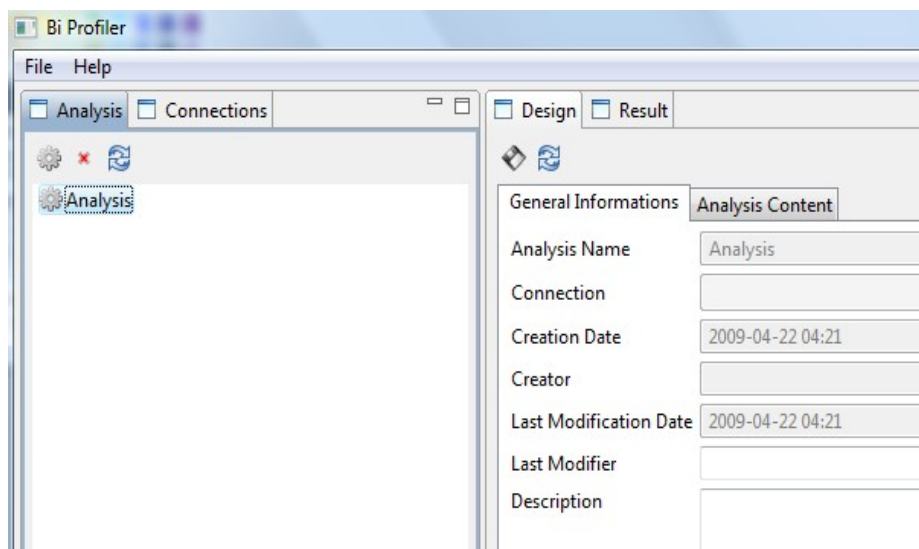
Cliquez sur le bouton 'create Analysis'.



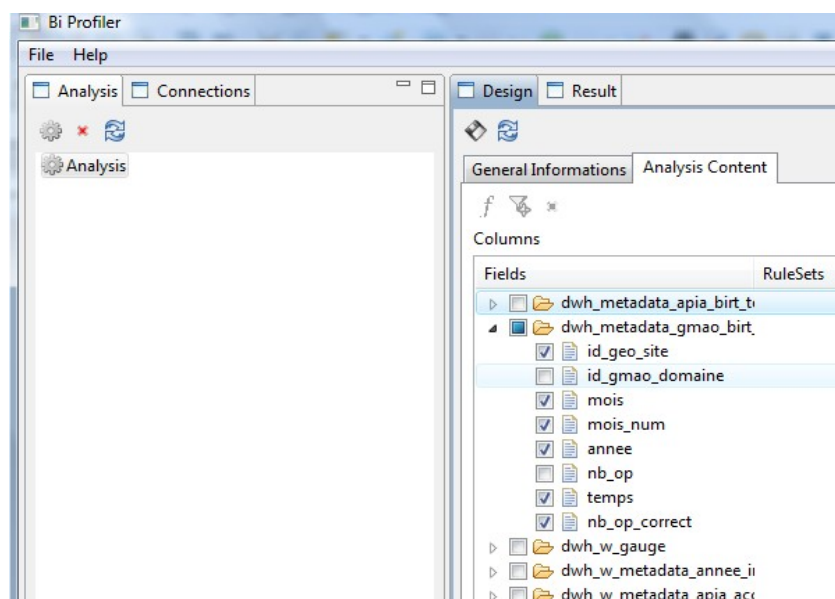
Renseignez les différents champs et sélectionnez la connexion utilisée pour l'analyse.

Définition de l'analyse

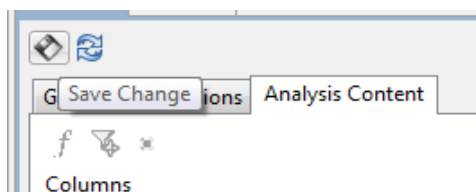
La prochaine étape se déroule depuis l'onglet 'Design'.



Cochez les colonnes que vous désirez ajouter à l'analyse.



Une fois les colonnes cochées, sauvez le contenu de votre analyse afin de pouvoir ensuite l'exécuter.



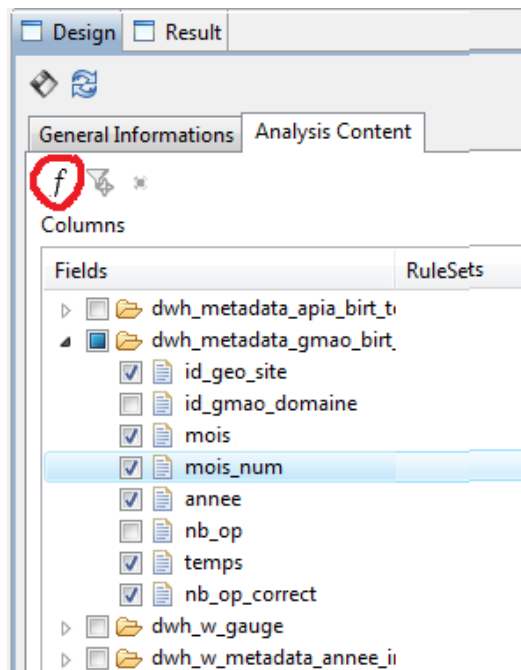
Création de Rules sets (Jeux de règles)

Vous pouvez créer un jeu de règles sur une colonne.

Chaque jeu de règles est une combinaison de conditions qui seront analysées pour la colonne associée.

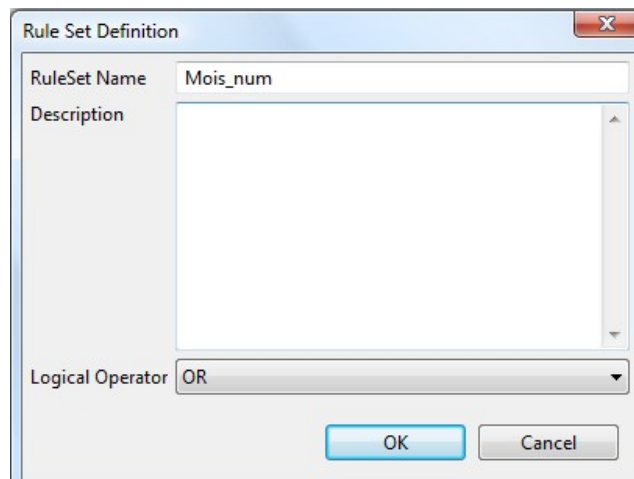
Note : vous ne pouvez créer des jeux de règles seulement sur les colonnes sélectionnées (cochées) et sauveées précédemment.

Pour ajouter une règle, sélectionnez une colonne cochée et cliquez sur le bouton entouré :

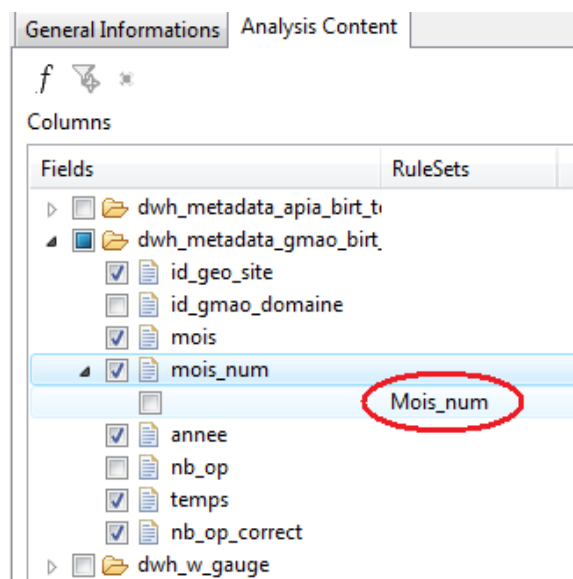


Définissez vos informations concernant votre jeu de règles.

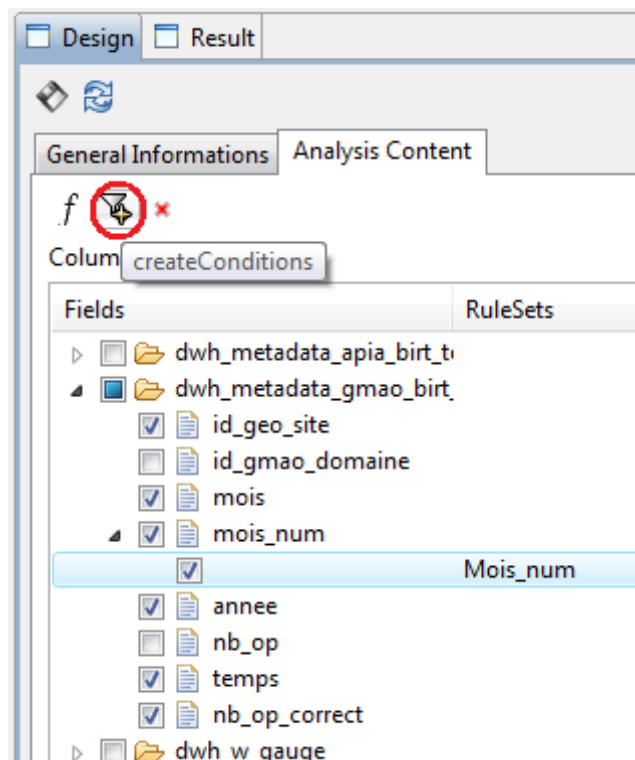
L'opérateur logique est utilisé afin d'autoriser de multiples conditions sur un jeu de règles.



Une fois validé (clic sur 'OK'), votre jeu de règles apparaît comme élément fils de votre colonne.

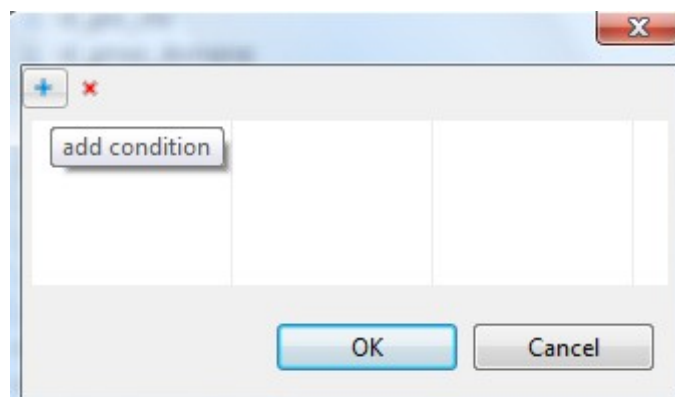


Vous devez maintenant définir les conditions de votre jeu de règles.



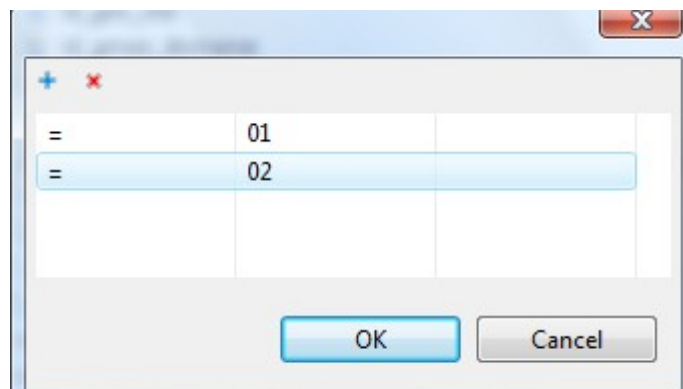
Sélectionnez votre jeu de règles et cliquez sur le bouton 'create Conditions'.

Puis, cliquez sur le bouton 'add condition' pour ajouter une condition.



Il y a 3 champs dans la table :

- Operateur
- Valeur 1 utilisée par l'opérateur
- Valeur 2 utilisée par l'opérateur

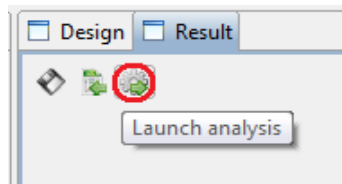


Si le type du champ est 'String' ou 'Date', il n'y a pas besoin de mettre l'expression entre quotes, elles seront ajoutées par défaut à l'exécution de l'analyse.

Exécution de l'analyse

La dernière étape est l'exécution de l'analyse.

Vous pouvez l'exécuter depuis le panneau 'Result' grâce au bouton 'Launch analysis'.



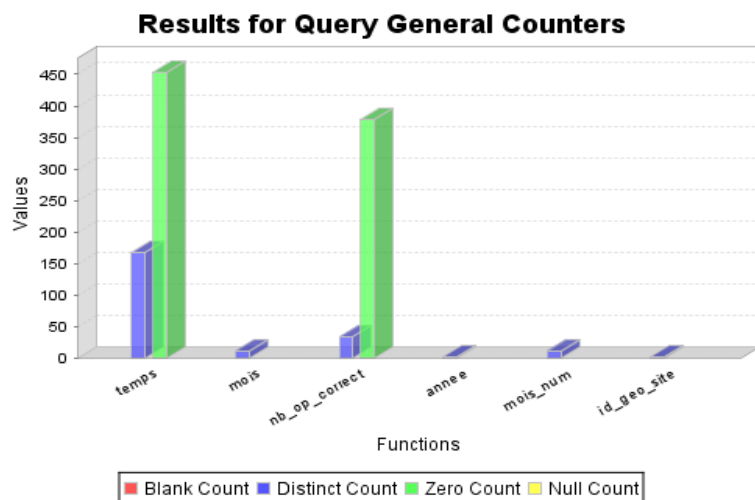
Vous pouvez à présent explorer les résultats au travers diverses interfaces.

Un onglet contient les informations générales des résultats de l'analyse.

Un second onglet contient les résultats par jeu de règles.

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



La table affiche les informations pour chaque champ mentionné dans l'analyse.

General Results

RuleSet Mois_num

Rule Set Informations

RuleSet Name

Mois_num

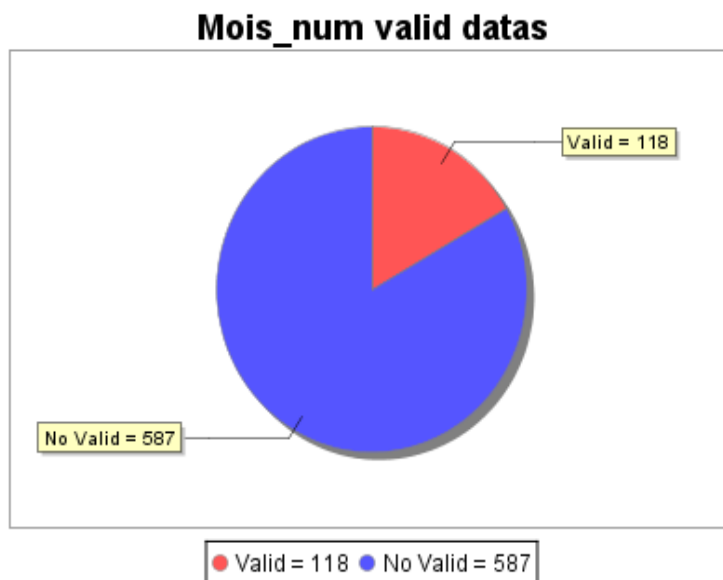
Description

Logical Operator

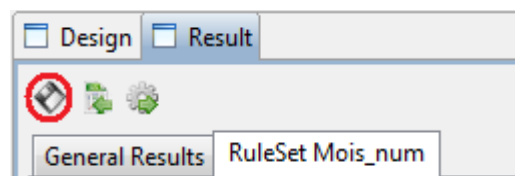
OR

Rule Set Results

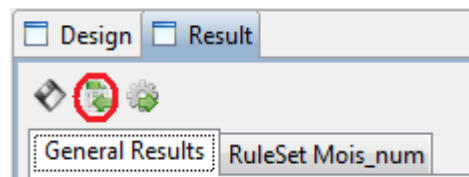
Operator	Value1	ValidCount	ValidPercent	DistinctValidCount	DistinctValidPercent
=	01	59	8 %	1	8 %
=	02	59	8 %	1	8 %
ALL		118	17 %	2	17 %



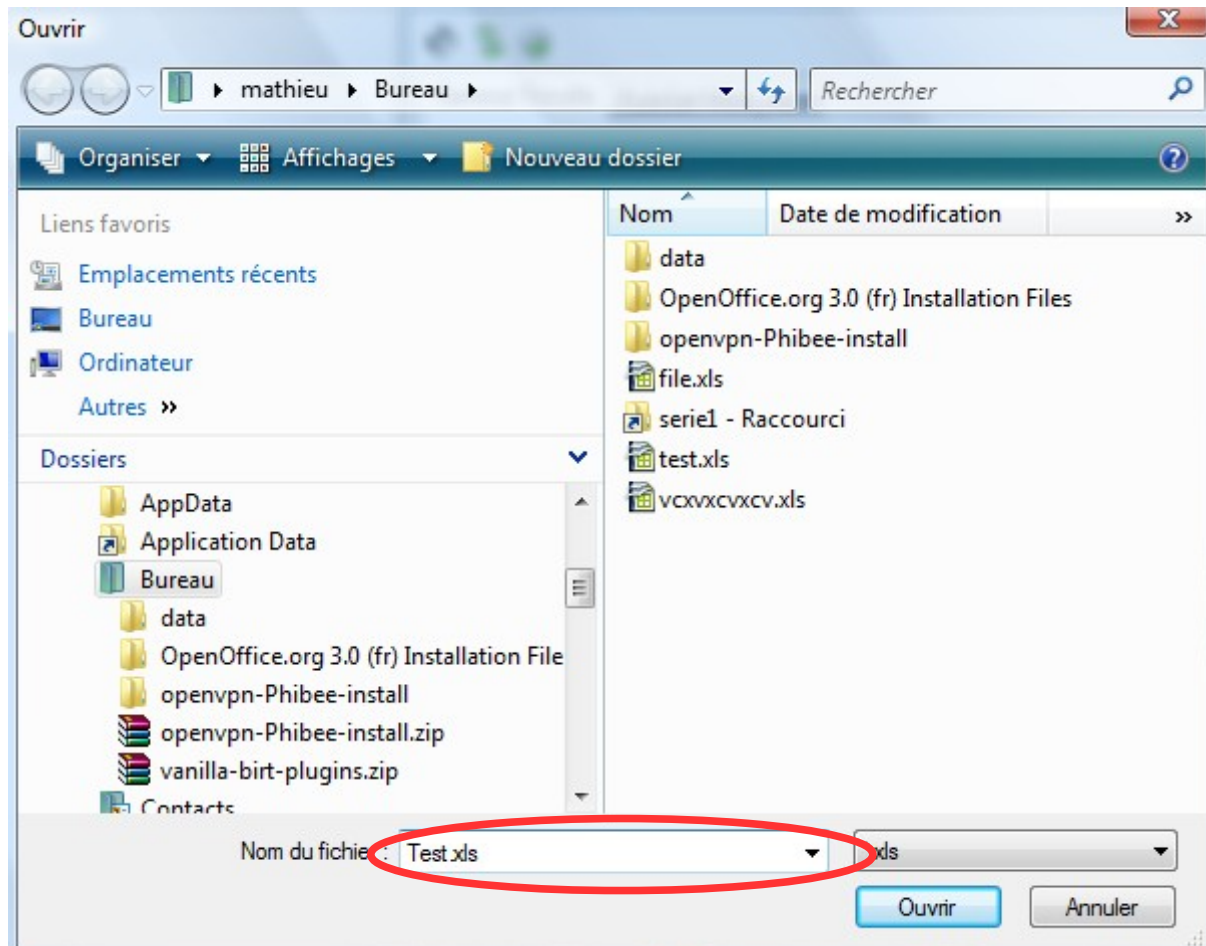
Le bouton 'save' vous permet de sauver les résultats de votre analyse en base de données.



Vous pouvez également exporter vos résultats dans des fichiers Excel.

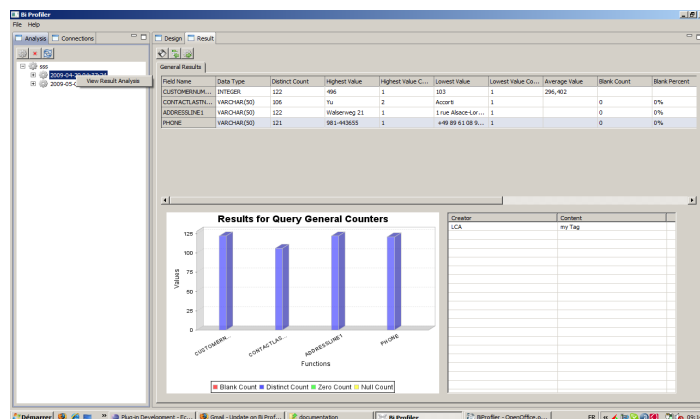


Définissez l'emplacement où sera sauvé votre fichier Excel et son nom.



Historique

Si vous avez déjà sauvé les résultats d'une analyse, vous pouvez consulter l'historique d'une analyse, puis l'éditer.



Mots clés

Vous pouvez ajouter des mots clés sur l'exécution d'une analyse avant de la sauver.

Pour cela, utilisez le clic droit sur la table affichant les résultats.

Notez que vous pouvez également ajouter des mots clés aussi bien sur les résultats généraux que sur les résultats par jeu de règles.

