



Fusion de deux sources de données sous Talend avec rejet

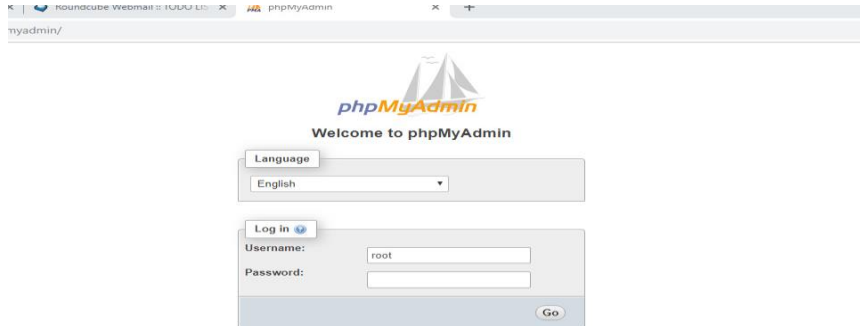
Introduction.

Ce tutoriel permet à tout lecteur ou internaute de récupérer les données dans un fichier Excel (par exemple liste des clients) et d'une table dans une base de données (par exemple les informations liées à des commandes), et de faire la jointure des deux sources de données et ressortir un résultat dans une table (listes des clients et des commandes passées) en rejetant les données qui ne sont pas liées (liste des clients qui n'ont pas passé de commande).

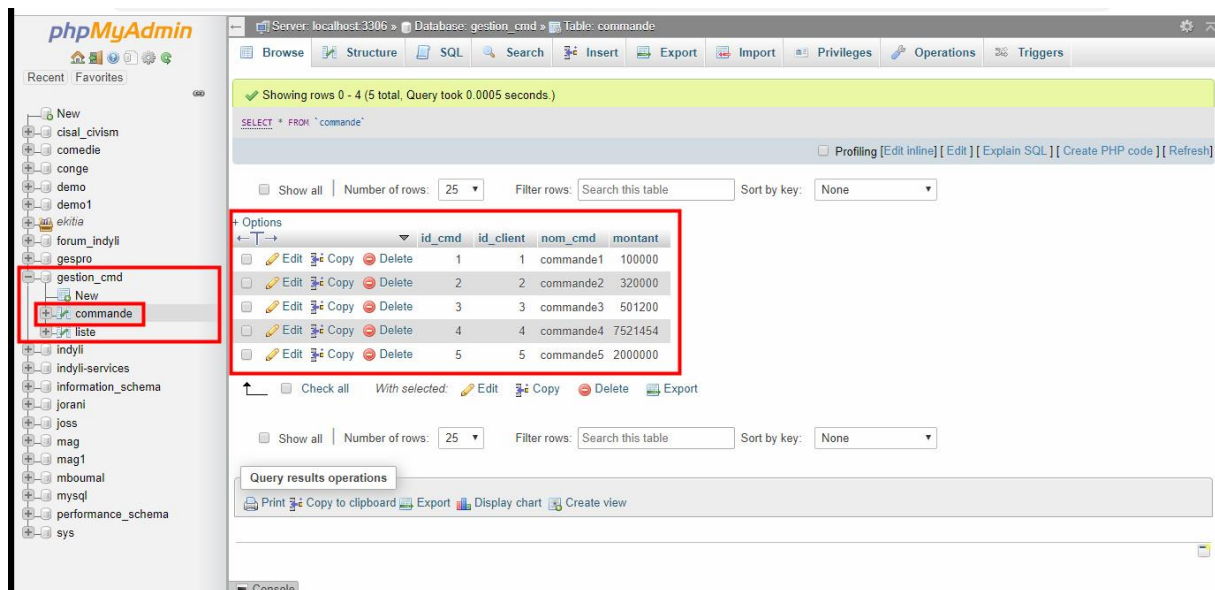
-Etape1 : préalable : création du fichier des données des clients (fichier Excel) avec les données.

id_client	nom_client	prenom	telephone	ville	email
1	Abesolo	yves	5698745	Douala	yves@yahoo.fr
2	Toupa	tia	4587965	Douala	tia@yahoo.com
3	Zonzon	zazine	78954632	Douala	zazine@gmail.com
4	Biboum	jack	95487445	Bafoussam	jack@gmail.com
5	Oumarou	ahmed	457896644	Garoua	ahmed@hotmail.com
6	Taphile	paul	458965555	yaounde	paul.tap@yahoo.fr
7	Toupa	tito	45789654	edea	tito@gmail.com

-Etape2 : Préalable : Création de la base de données avec les tables



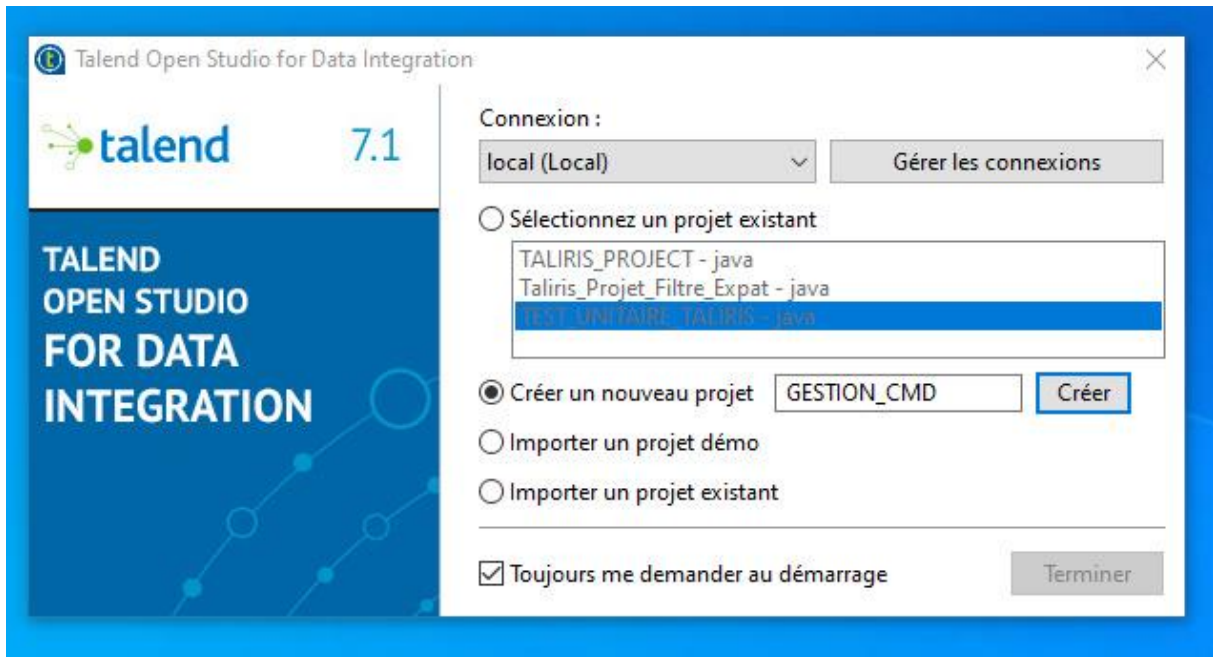
* 1 table (commande) comportant les données sur les commandes des clients avec un attribut faisant la liaison entre le client et la commande



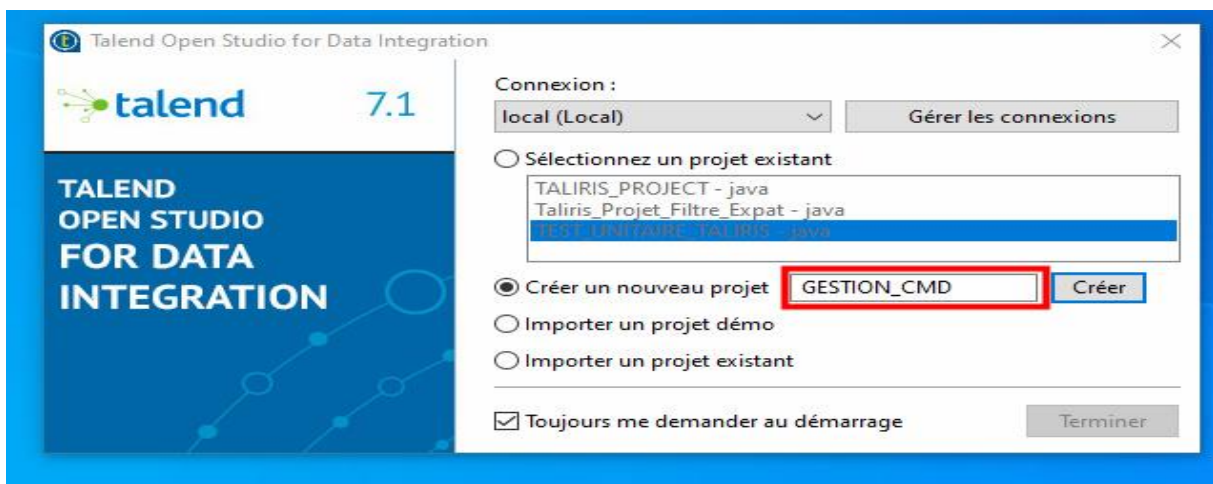
* 1 table (liste ou all) qui comportera les clients qui ont passé une commande

Etape3 : Création du projet

Au démarrage de talend, une boîte de dialogue se présente et nous pouvons choisir « **créer un projet** »

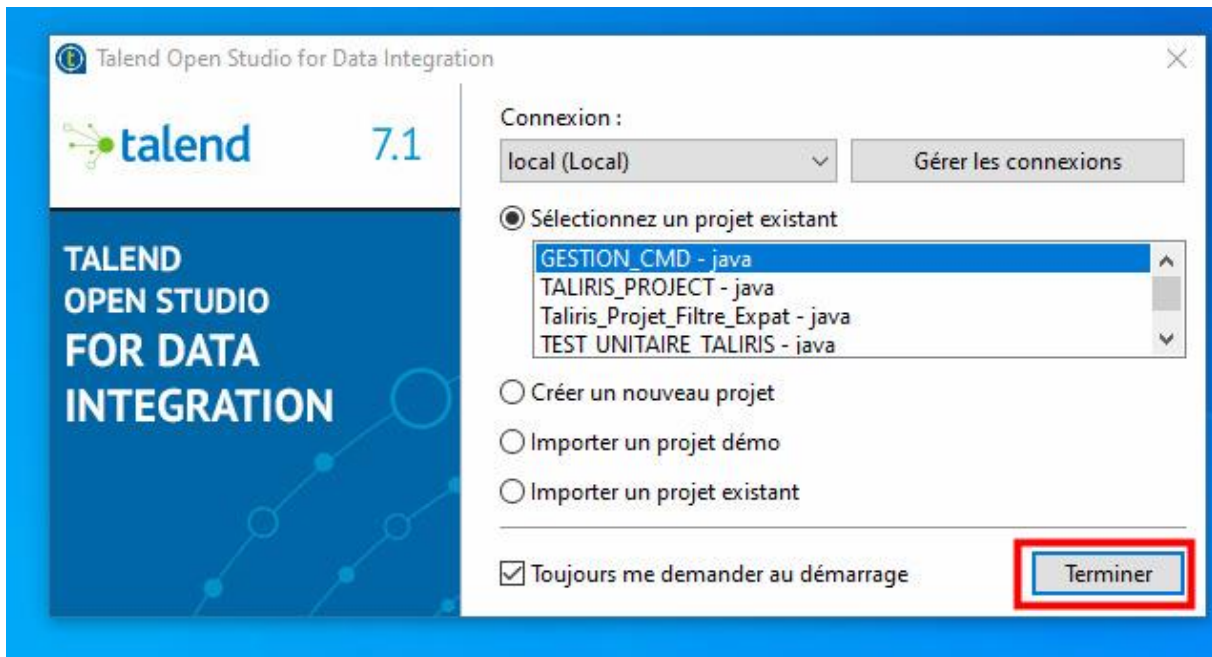


, en suite renseigner le **nom du projet** et de cliquer sur le bouton « **Créer** »





, et enfin sélectionner **le nouveau projet créé** et de cliquer sur « **Terminer** ».

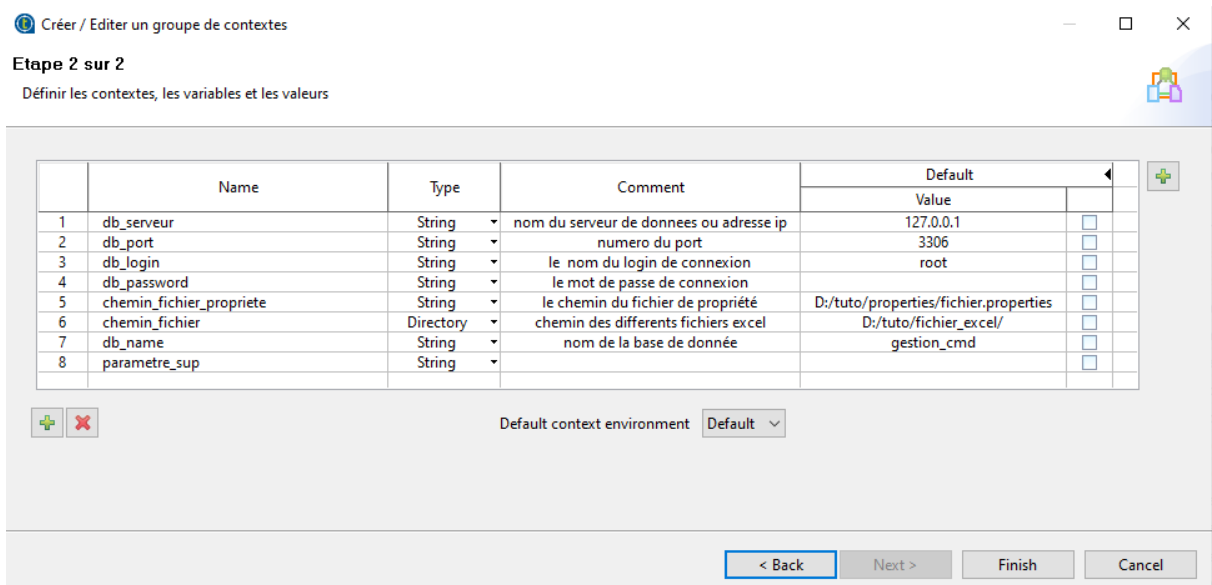
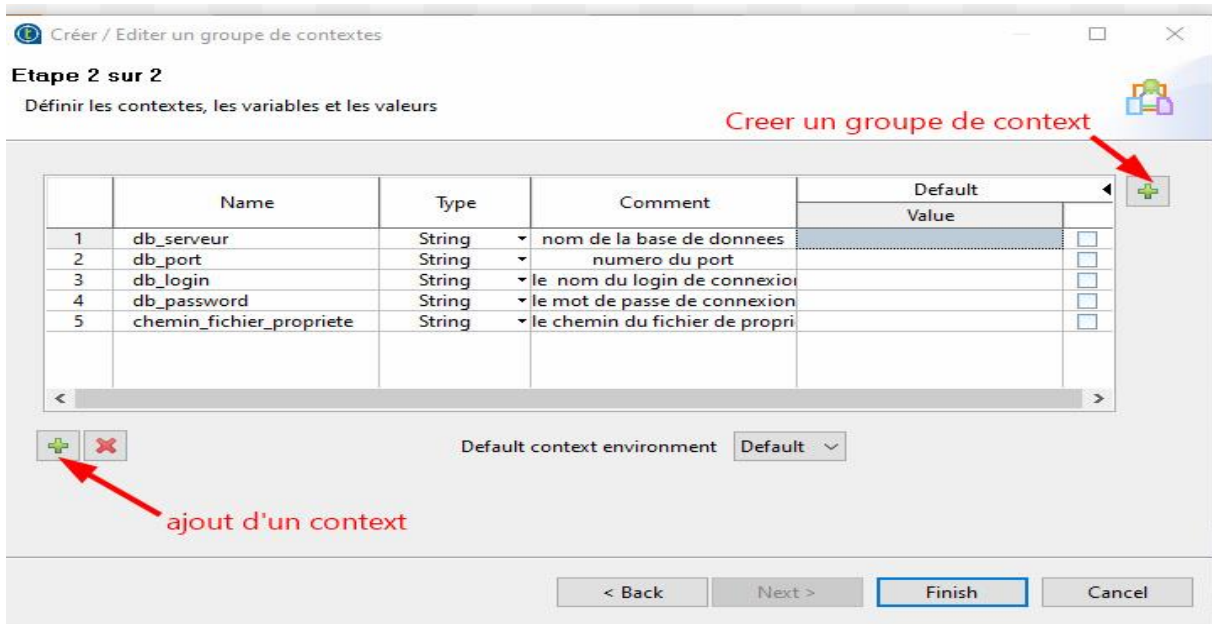


-Etape4 CREATION DES CONTEXTES DE CONNEXION ET DU FICHER DE PROPRIETE

La création des contextes de connexion permet de dynamiser les accès de connexion à la base de donnée à travers un fichier de configuration appelé fichier de propriété.

a) Création des contextes

- Faites un clic droit sur le menu « CONTEXTES » située sur la barre latérale gauche des menus, puis cliquer sur « Groupes de contexte » ; ensuite renseigner les variables de contexte, les types, et les valeurs des variables et enfin cliquer sur le bouton « finish ».



b) Création du fichier de propriété

Créer un fichier avec un éditeur de texte (Notepad++) et renseigner les valeurs, puis d'enregistrer se fichier avec l'extension « **.properties** »

```
1
2 #Groupe de contexte de connexion a la base de donnée
3
4 db_serveur=127.0.0.1
5 db_port=3306
6 db_login=root
7 db_password=
8 db_name=gestion_cmd
9 chemin_fichier=D:/tuto/fichier_excel/
10 parametre_sup=
11 chemin_fichier_propriete=D:/tuto/properties/fichier.properties
12
```

Properties file length : 264 lines : 12 Ln : 1 Col : 1 Sel : 0 | 0 Windows (CR LF) UTF-8 INS

-Etape5 CREATION D'UNE CONNEXION A LA BASE DE DONNEES

Faites un clic gauche sur le menu « métadonnées » située sur la barre latérale gauche des menus, puis faire un clic droit sur le menu « Connexion aux bases de données » et de cliquer sur « Créer une connexion »



Connexion à la base de données

Nouvelle connexion à une base de données dans le référentiel - Etape1/2

Le nom est vide.

Nom: (1)

Objectif: (2)

Description: (3)

Créé par: user@talend.com

Verrouillé par:

Version: 0.1

Statut:

Chemin d'accès:

< Back Next > Finish Cancel

Connexion à la base de données

Mettre à jour la connexion à une base de données - Etape1/2

Mettre à jour les propriétés

Nom:

Objectif:

Description:

Créé par: user@talend.com

Verrouillé par: user@talend.com

Version: 0.1

Statut:

Chemin d'accès:

< Back **Next >** Finish Cancel



Ensuite renseigner les informations ensuite cliquer sur suivant ou next, puis renseigner les variables des contextes précédé du mot clé « **context** » suivit d'un « . » puis suivit du **nom de la variable**.

Connexion à la base de données

Mettre à jour la connexion à une base de données - Etape 2/2

Type de BdD MySQL

Version de la base de données MySQL 5

Chaîne de caractères de connexion jdbc:mysql://127.0.0.1:3306/gestion_cmd?

Identifiant context.db_login

Mot de passe context.db_password

Serveur context.db_serveur

Port context.db_port

Base de données context.db_name

Paramètres supplémentaires context.parametre_sup

Tester la connexion

Exporter en tant que contexte Revenir au contexte précédent

[Comment installer un pilote](#)

< Back Next > Finish Cancel

Enfin vous pouvez cliquer sur le bouton « Tester la connexion ». Si une boîte de dialogue s'affiche avec un message « **connexion établie** » alors votre configuration est bonne.

ETAPE6 : CREATION DE LA METADONNEES EXCEL DANS TALEND

Faites un clic gauche sur le menu « métadonnées » située sur la barre latérale gauche des menus, puis faire un clic droit sur le menu « fichier excel » et renseigner les informations puis cliquer sur « suivant ».

Nouveau fichier Excel

Fichier - Etape 1 de 4

Le nom est vide.

Nom 1 Nom de votre composant du fichier

Objectif 2 Objectif (facultatif)

Description 3 description(facultative)

Créé par : user@talend.com

Verrouillé par :

Version 0.1 M m

Statut

Chemin d'accès Sélectionner

< Back Next > Finish Cancel

Nouveau fichier Excel

Fichier - Etape 1 de 4

Ajouter un fichier de métadonnées au référentiel
Définir les propriétés

Nom

Objectif

Description

Créé par : user@talend.com

Verrouillé par :

Version 0.1 M m

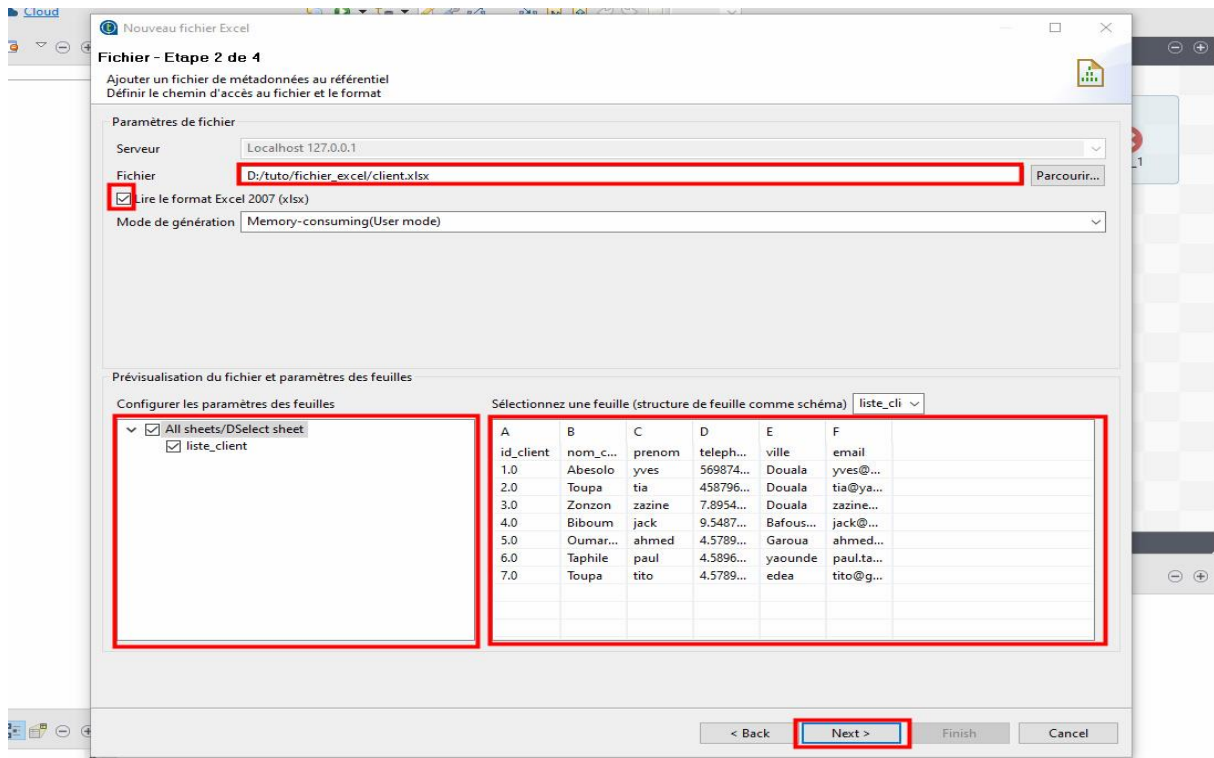
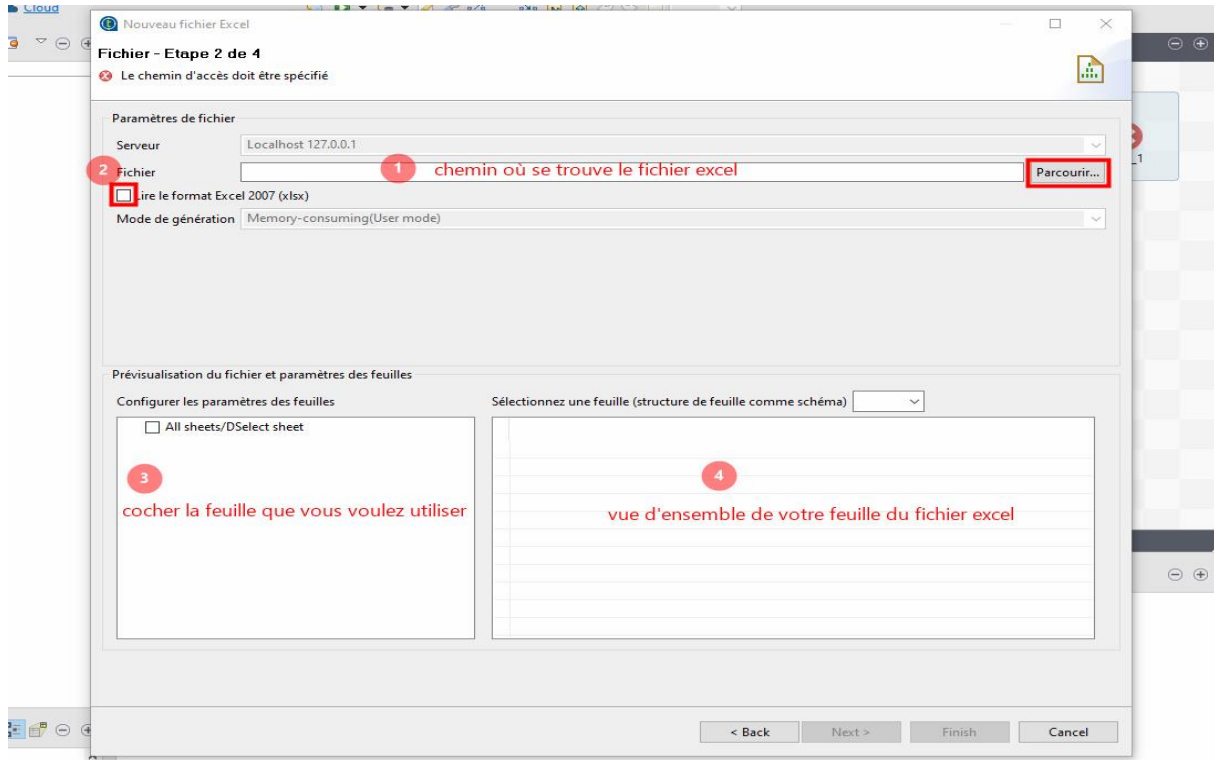
Statut

Chemin d'accès Sélectionner

< Back **Next >** Finish Cancel



Ensuite cliquer sur le bouton « parcourir » et choisissez le fichier Excel qui contient les données sur les clients, et cocher la case « lire le format excel 2007 » ; puis choisir la feuille de votre fichier excel qui contient les données, puis cliquer sur « next ».



Ensuite, laissez l'encodage utf8 par défaut, et cocher la case « En tête », puis cocher aussi la case « Définir les lignes d'en-tête comme nom de colonne » se trouvant sur l'onglet « aperçu »

Fichier - Etape 3 de 4
Add a Metadata File on repository
Define the setting of the parse job

Paramètres de fichier
Encodage: UTF-8
 Séparateur avancé (pour les nombres)
Séparateurs de milliers: ,
Séparateur décimal: .

Paramètre de la colonne de métadonnées
Première colonne: 1
Dernière colonne:

Lignes à ignorer
Si des lignes doivent être ignorées, spécifiez les paramètres suivants
En-tête:
Pied de page:

Limite de lignes
Si le nombre de lignes doit être limité, spécifiez ce nombre.
Limite:

Aperçu | Sortie
 Définir les lignes d'en-tête comme nom de colonnes Rafraîchir l'aperçu

A	B	C	D	E	F
id_client	nom_client	prenom	telephone	ville	email
1	Abesolo	yves	5698745	Douala	yves@yahoo.fr
2	Toupa	tia	4587965	Douala	tia@yahoo.com
3	Zonzon	zazine	78954632	Douala	zazine@gmail.com
4	Biboum	jack	95487445	Bafoussam	jack@gmail.com
5	Oumarou	ahmed	457896644	Garoua	ahmed@hotmail.com
6	Taphile	paul	458965555	yaounde	paul.tap@yahoo.fr

Exporter en tant que contexte Revenir au contexte précédent

< Back Next > Finish Cancel

Fichier - Etape 3 de 4
Add a Metadata File on repository
Define the setting of the parse job

Paramètres de fichier
Encodage: UTF-8
 Séparateur avancé (pour les nombres)
Séparateurs de milliers: ,
Séparateur décimal: .

Paramètre de la colonne de métadonnées
Première colonne: 1
Dernière colonne:

Lignes à ignorer
Si des lignes doivent être ignorées, spécifiez les paramètres suivants
En-tête: 1
Pied de page:

Limite de lignes
Si le nombre de lignes doit être limité, spécifiez ce nombre.
Limite:

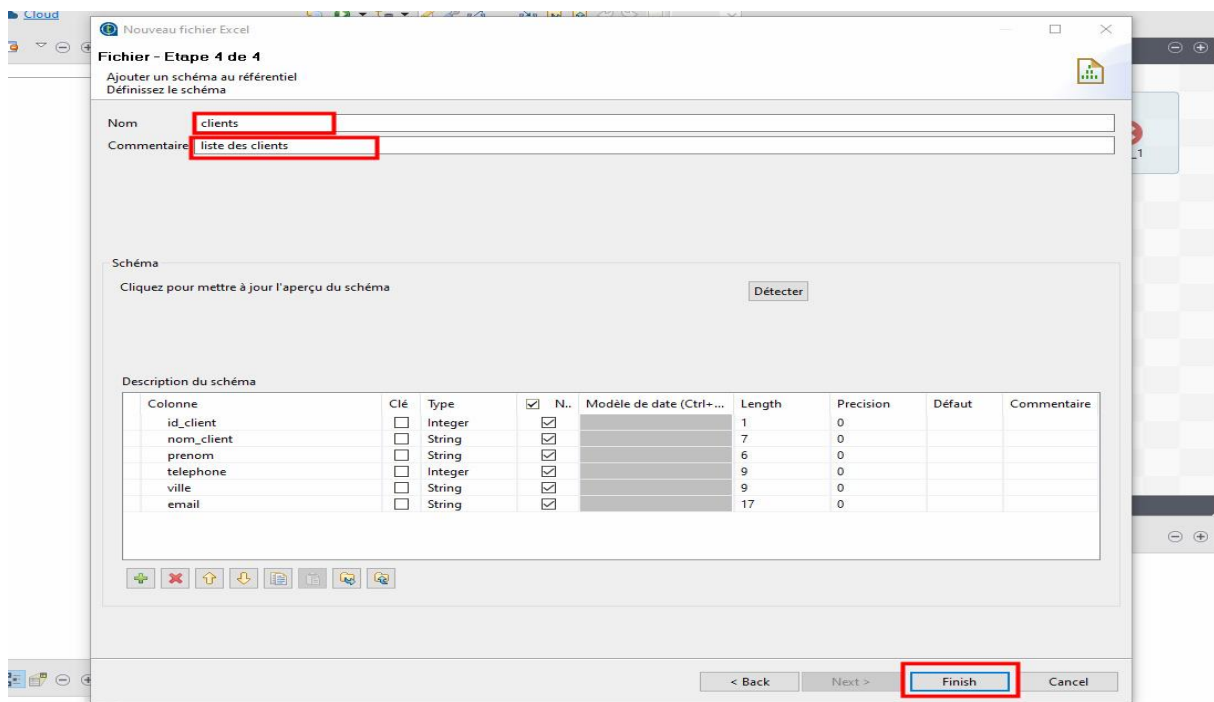
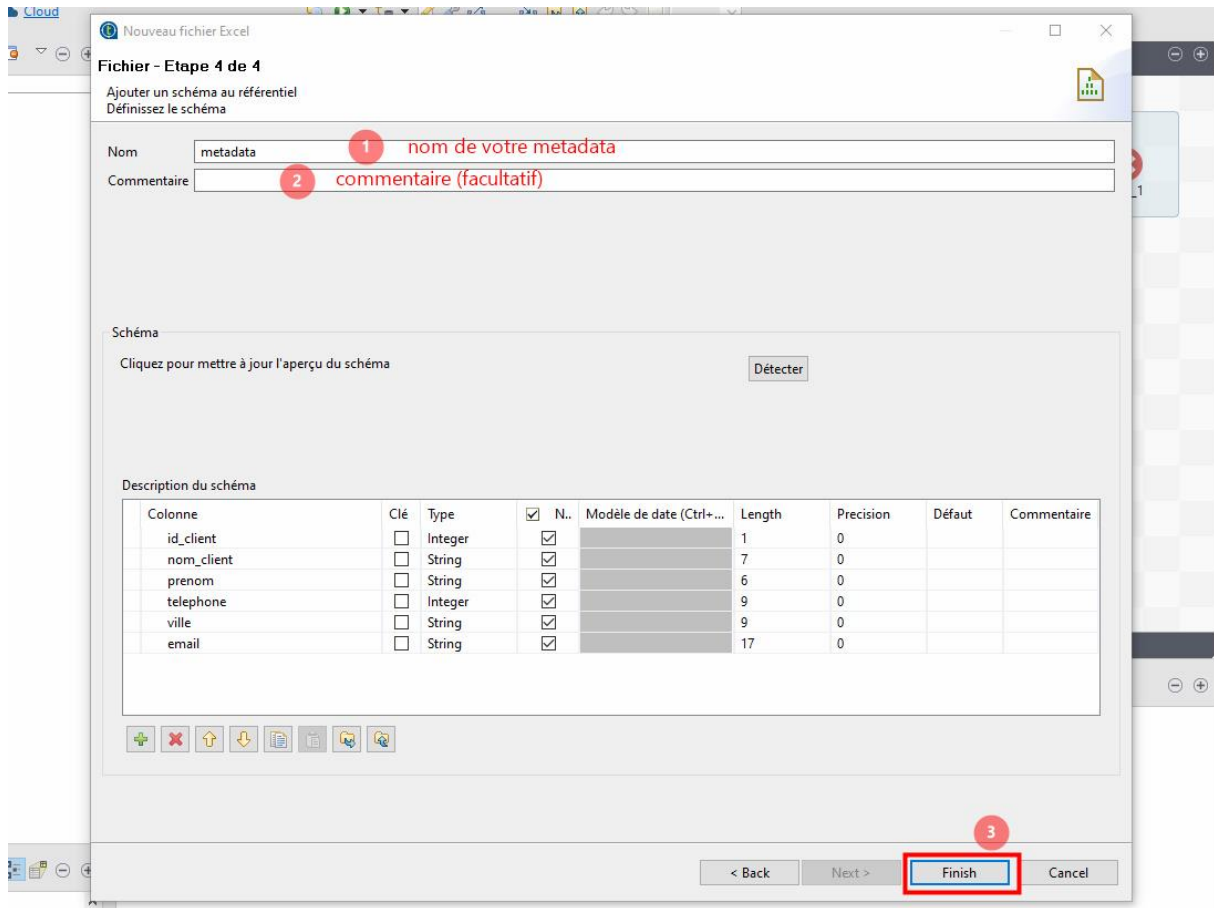
Aperçu | Sortie
 Définir les lignes d'en-tête comme nom de colonnes Rafraîchir l'aperçu

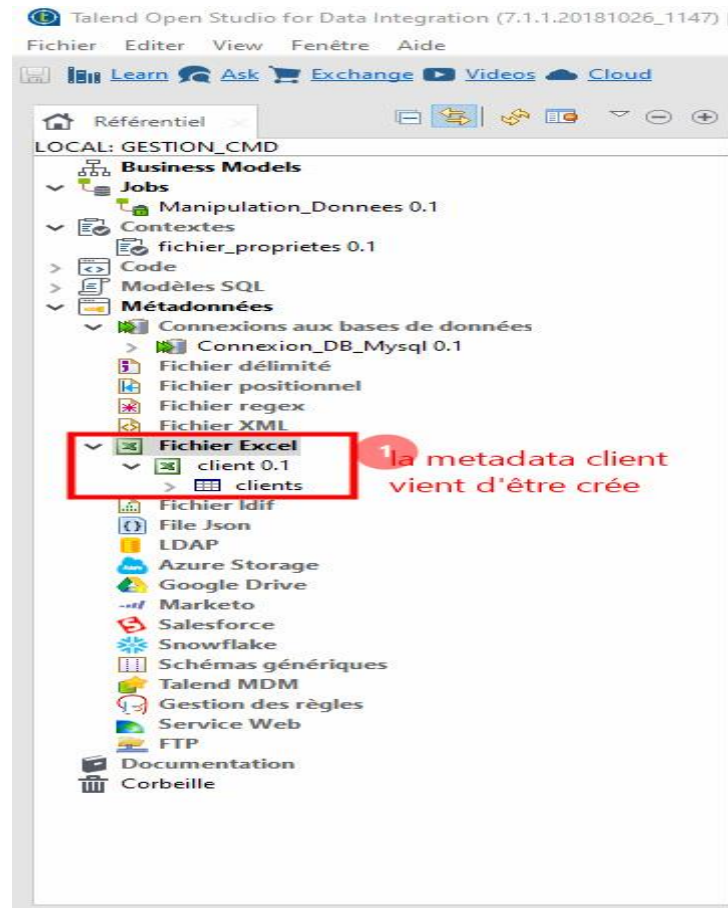
A	B	C	D	E	F
id_client	nom_client	prenom	telephone	ville	email
1	Abesolo	yves	5698745	Douala	yves@yahoo.fr
2	Toupa	tia	4587965	Douala	tia@yahoo.com
3	Zonzon	zazine	78954632	Douala	zazine@gmail.com
4	Biboum	jack	95487445	Bafoussam	jack@gmail.com
5	Oumarou	ahmed	457896644	Garoua	ahmed@hotmail.com
6	Taphile	paul	458965555	yaounde	paul.tap@yahoo.fr
7	Toupa	tito	45789654	ede	tito@gmail.com

Exporter en tant que contexte Revenir au contexte précédent

< Back Next > Finish Cancel

Enfin entrer le nom de votre metadonnées du fichier excel et de cliquer sur terminer.

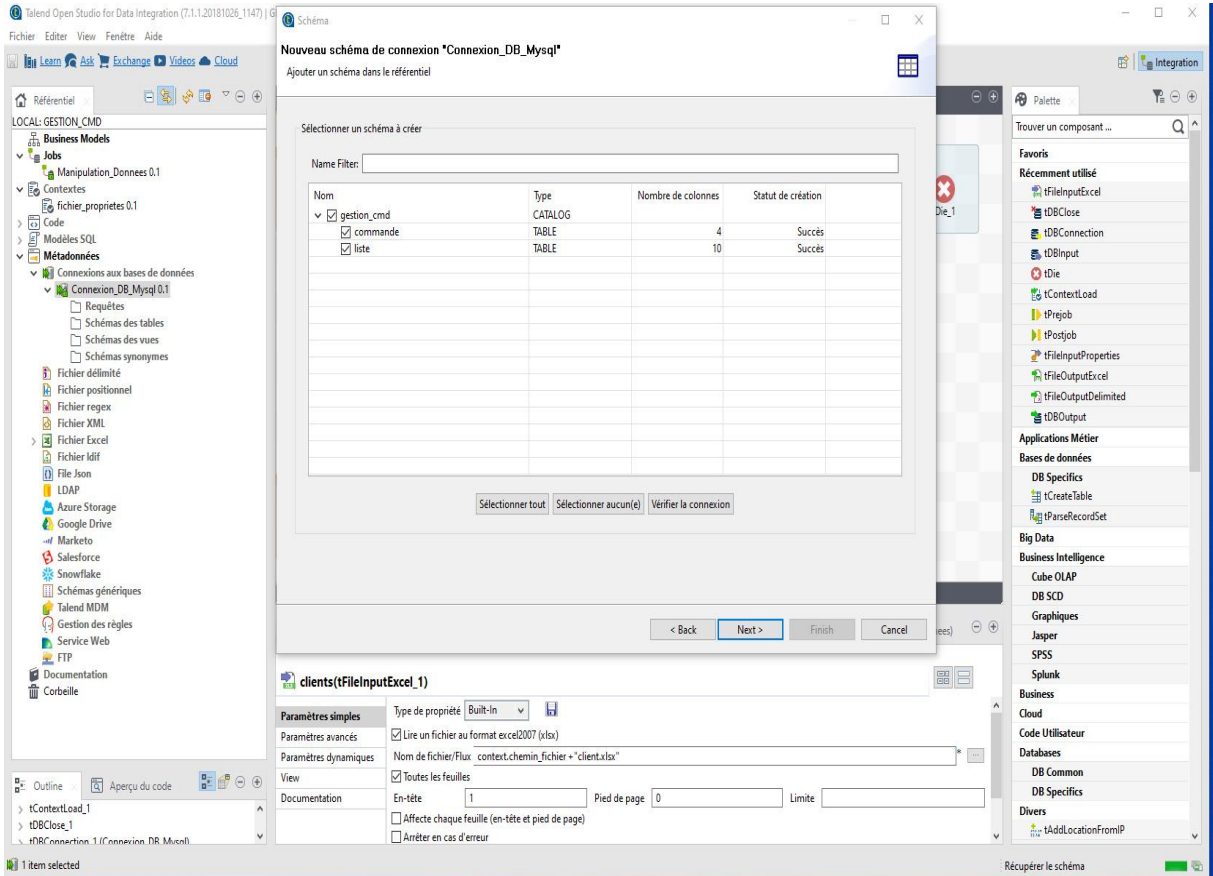
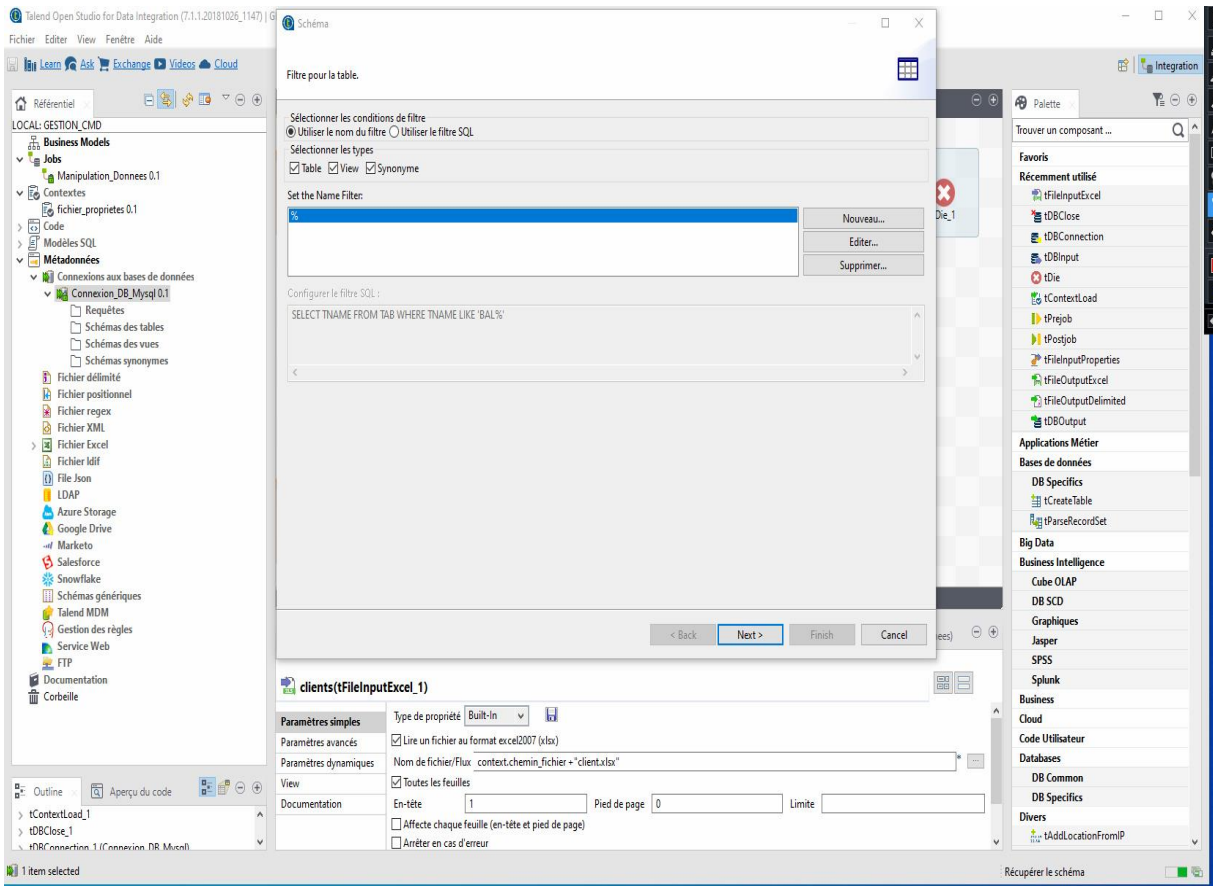




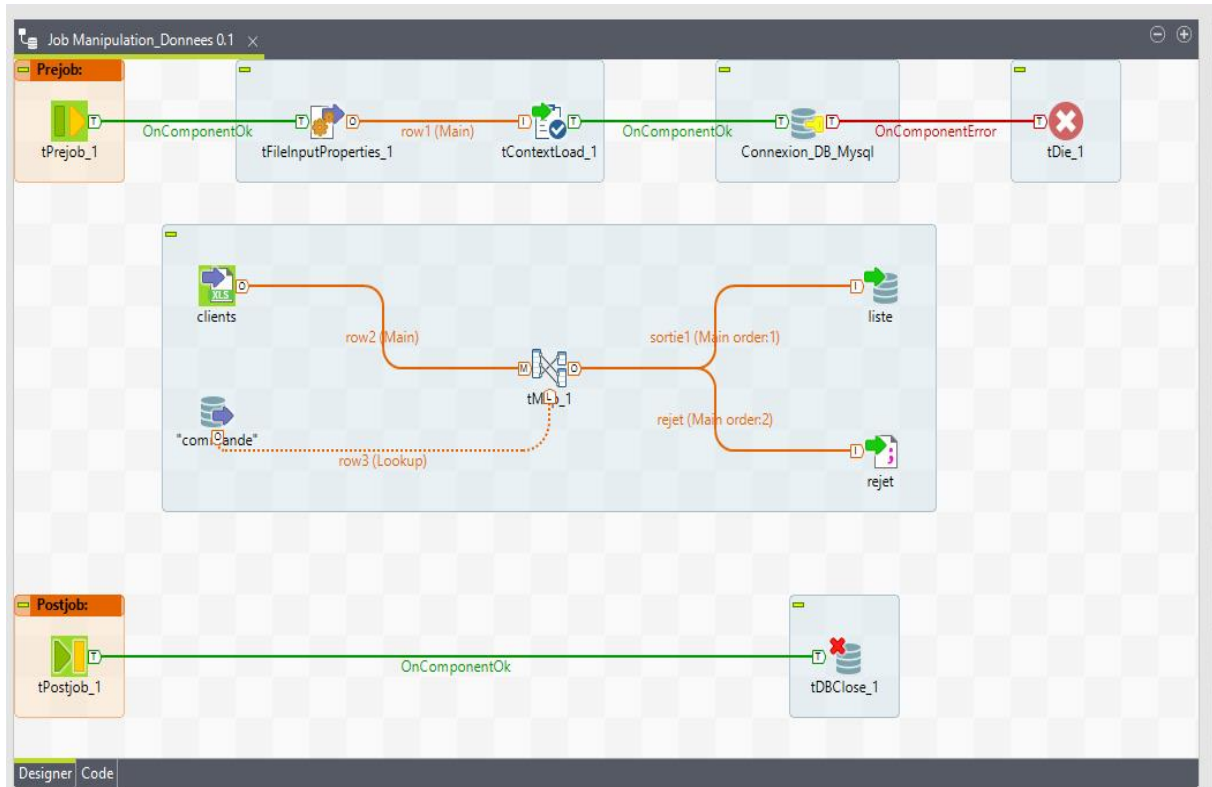
ETAPE7 : IMPORTER LA METADONNEES DES TABLES DANS TALEND

Faites un clic gauche sur le menu « métadonnées » située sur la barre latérale gauche des menus, ensuite faire un clic gauche sur le sous menu « Connexion aux bases de données » et de faire un clic droit sur votre connexion nouvellement créée, et de choisir « **Recupérer le schemas** ».

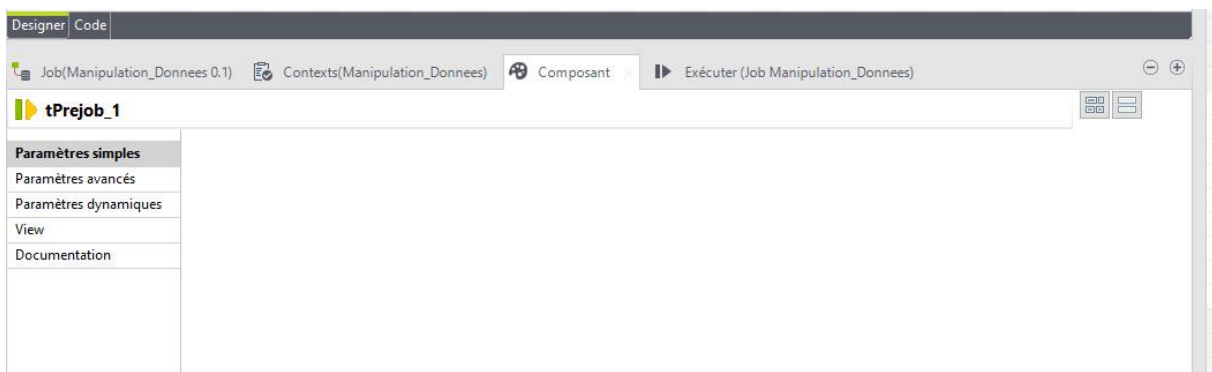
Laisser les éléments par défaut et cliquer sur « suivant » et enfin cocher les tables à importer dans la liste des métadonnées des tables de talend.



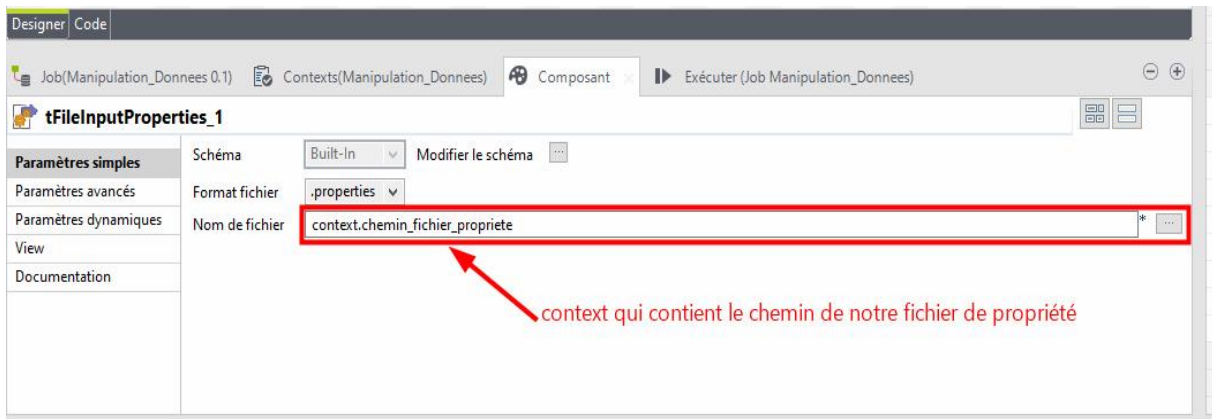
Etape8 : PREPARER LES COMPOSANT DU JOB ET LES CONFIGURER



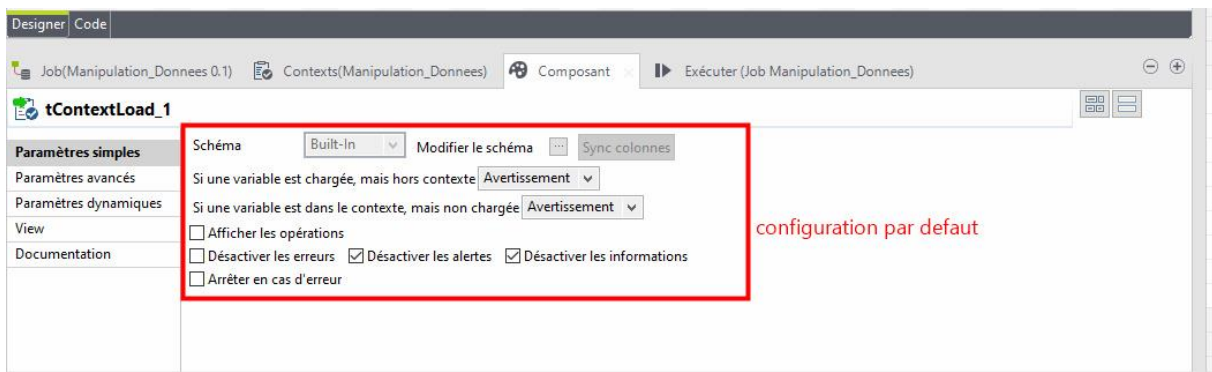
a) Déposer le composant « Tprejob » il est configuré par défaut comme l'image ci-dessous



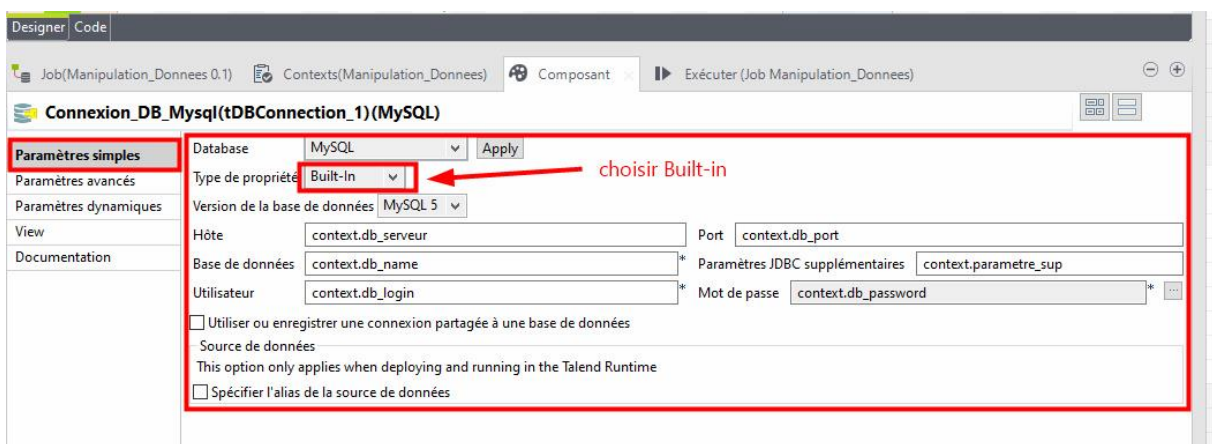
b) Déposer le composant « tFileinputProperies » et configurer ce composant comme l'image ci-dessous



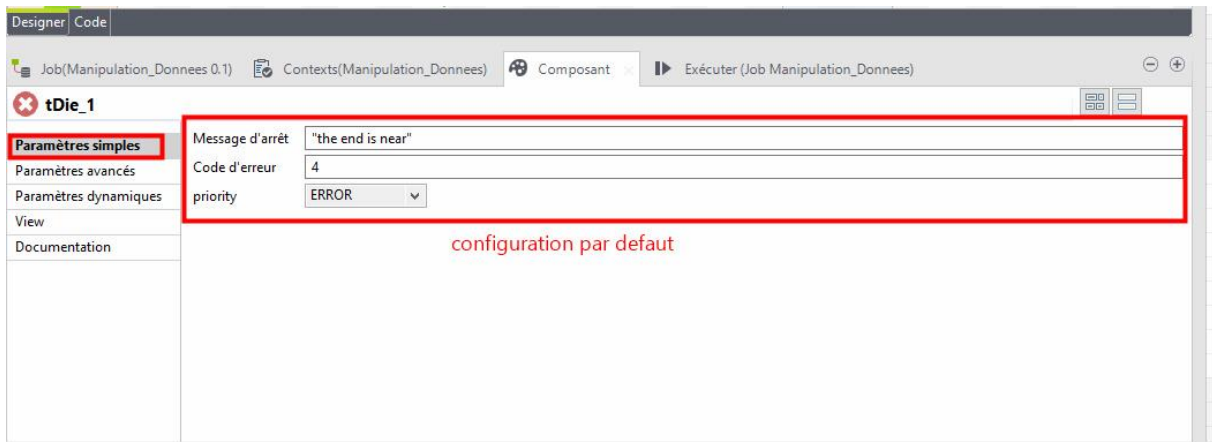
c) Déposer le composant « tContextLoader » qui permet d'enregistrer les valeurs des contextes en mémoire et configurer ce composant comme l'image ci-dessous



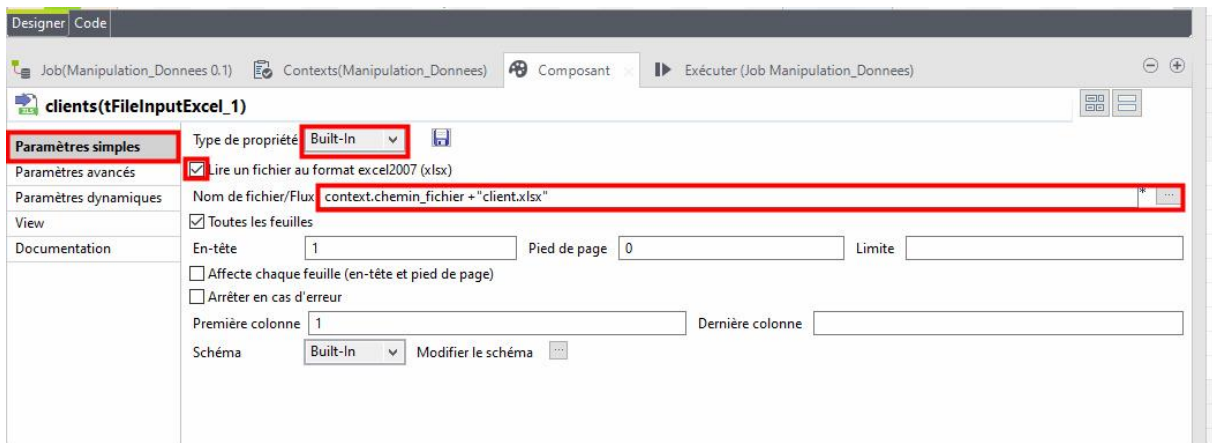
d) Déposer le composant « tDBConnexion » et configurer ce composant comme l'image ci-dessous



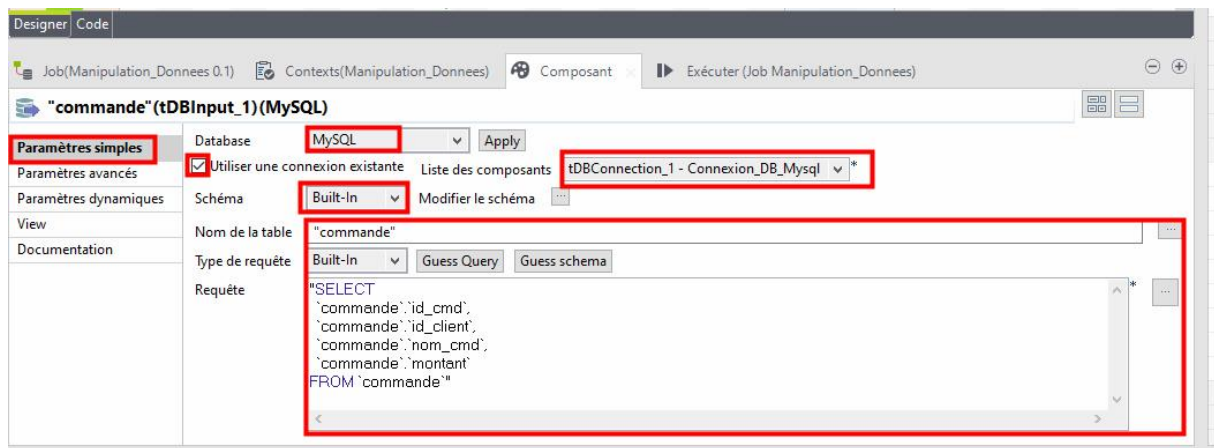
e) Déposer le composant « tDie » il est configuré par défaut comme l'image ci-dessous



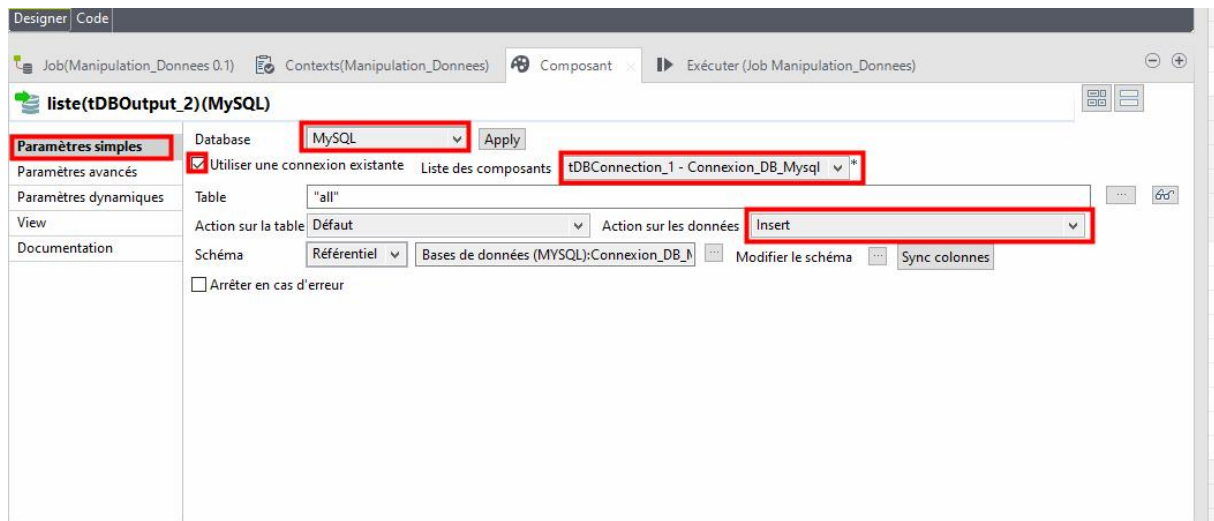
f) Déposer le composant de la métadonnées Excel (client) précédemment importé à l'étape 6 et configurer ce composant comme l'image ci-dessous



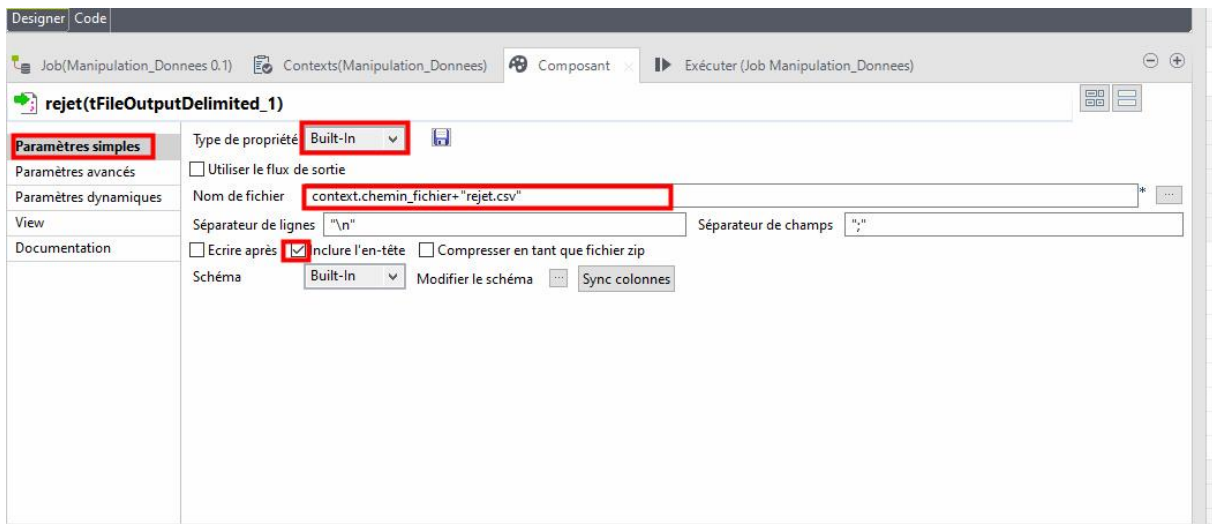
g) Déposer le composant de la métadonnée de la table (commande) précédemment importé à l'étape 7 en choisissant « tMysqlInput » et configurer ce composant comme l'image ci-dessous



h) Déposer le composant de la métadonnée de la table (all) précédemment importé à l'étape 7 en choisissant « tMysqlOutput » et configurer ce composant comme l'image ci-dessous



i) Déposer le composant **tFileOutputDelimited du fichier csv pour les rejets et configurer ce composant comme l'image ci-dessous**



- j) **Déposer le composant « tMap » en créant les différentes sorties** (une pour charger les données des clients et des commandes dans la table « all », l'autre sortie nommée rejet pour charger les données des clients qui n'auront pas de commande dans le fichier csv rejet).

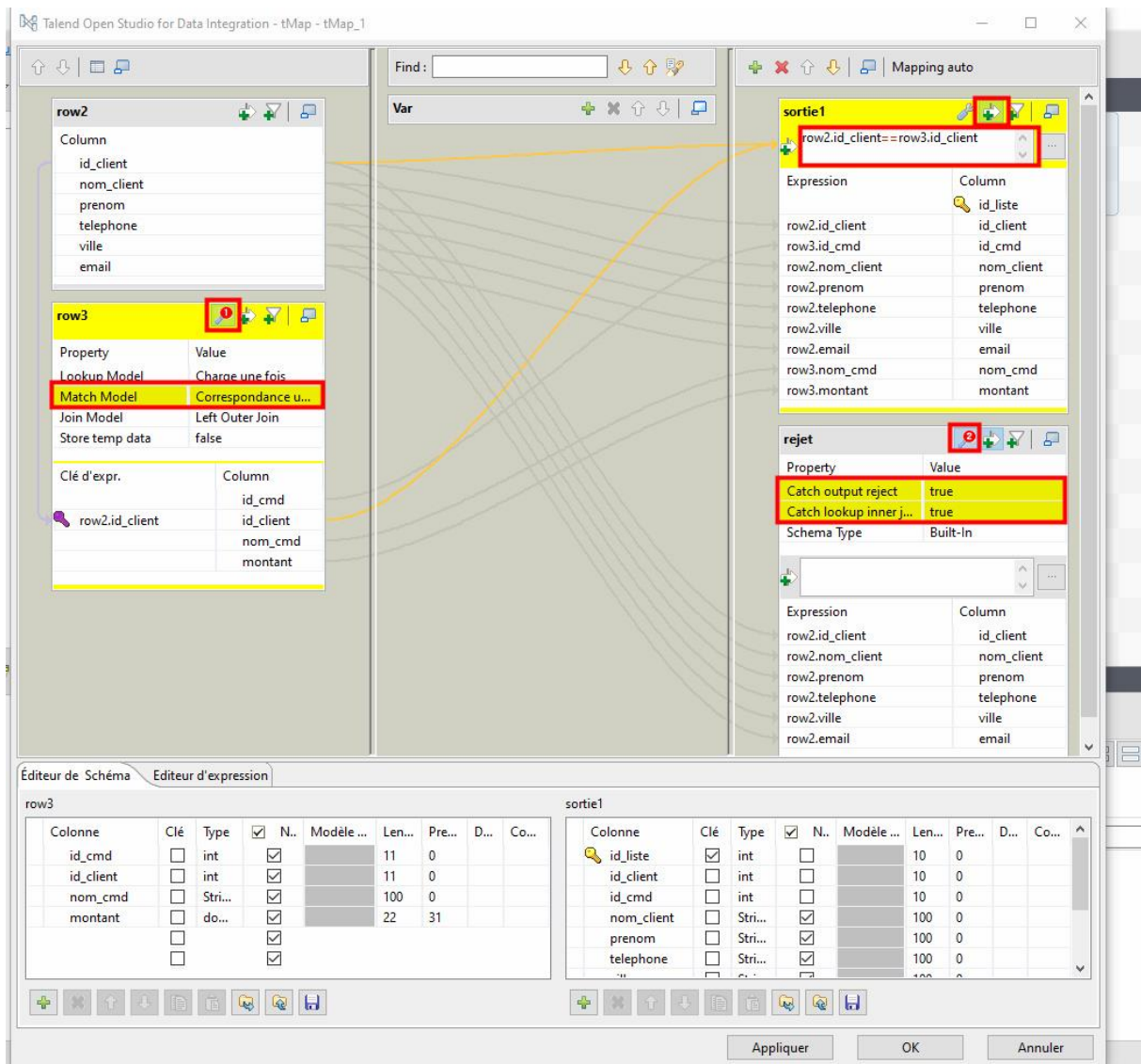
Une jointure est faite dans la **TMAP** entre l'entrée du fichier Excel et l'entrée de la table **commande** sur l'attribut « id_client ».

Les conditions sont également faites sur les différentes sorties

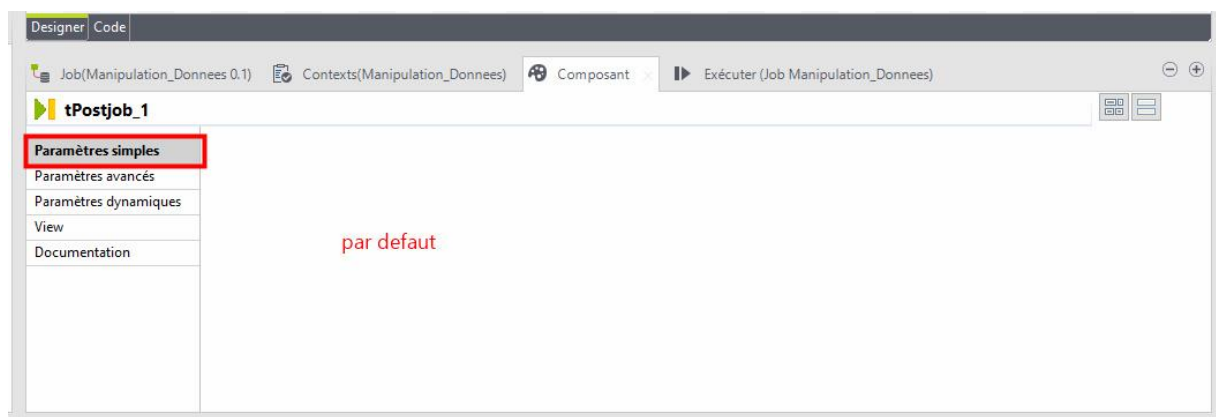
- id_client du fichier Excel doit être égale à id_client se trouvant dans la table commande

-les données qui ne respectent pas la condition de jointures seront directement placées dans le fichier de rejet.

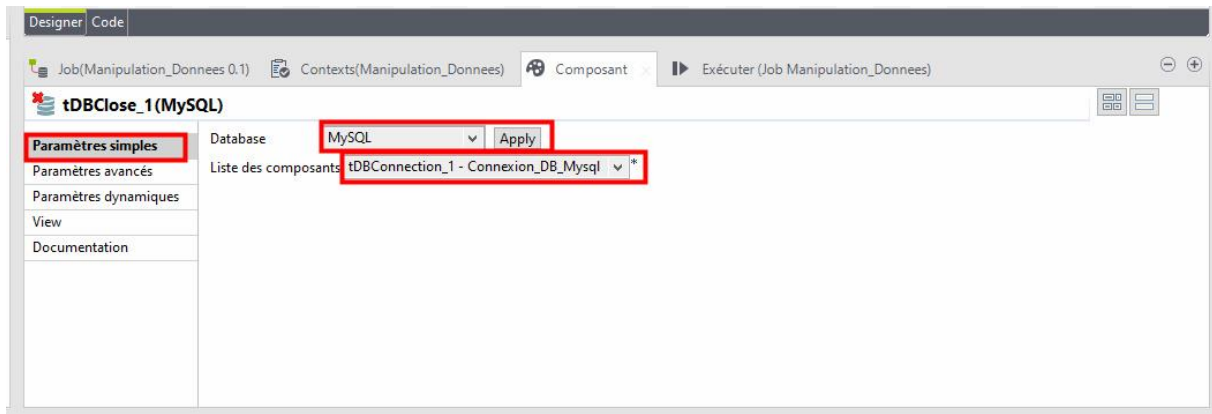
La configuration de ce composant comme l'image ci-dessous.



k) Déposer le composant « Tpostjob » il est configuré par défaut comme l'image ci-dessous

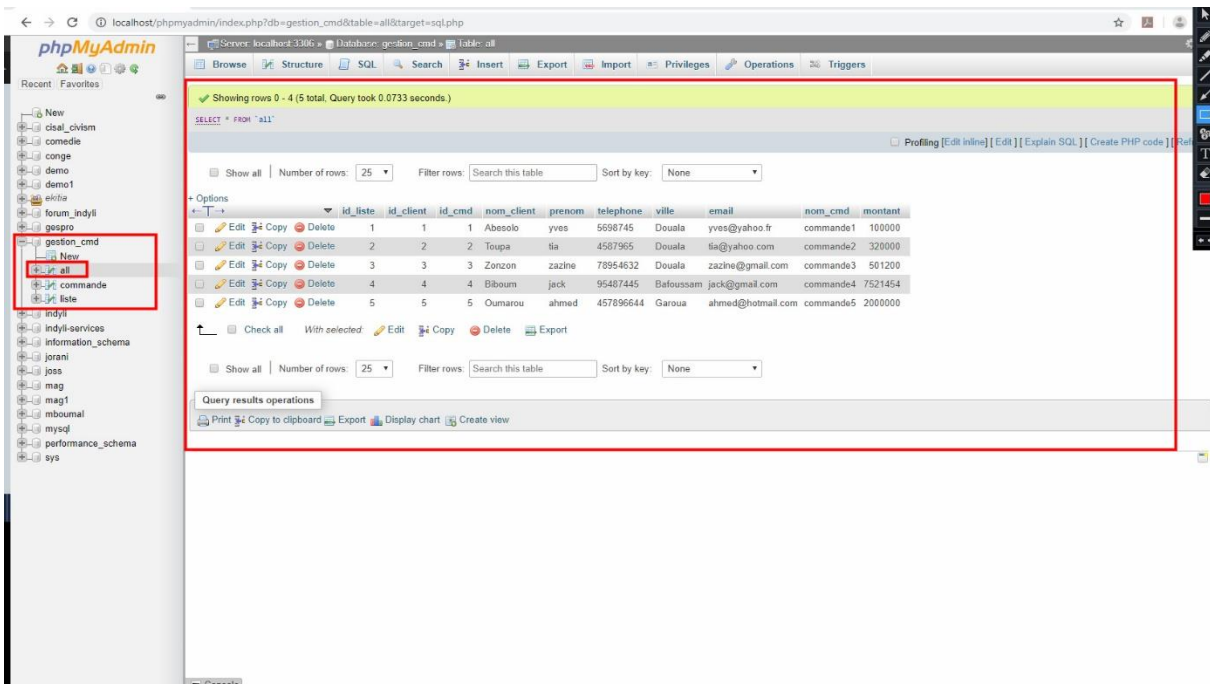


I) Déposer le composant « TDBclose » qui permet de fermer une connexion, il est configuré comme l'image ci-dessous



Après l'exécution du job, nous avons les résultats suivants dans notre base de données et dans notre fichier de rejet.

Résultat dans la Table all :



id_liste	id_client	id_cmd	nom_client	prenom	telephone	ville	email	nom_cmd	montant
1	1	1	Abeolo	yves	5698745	Douala	yves@yahoo.fr	commande1	100000
2	2	2	Toupa	tia	4587965	Douala	tia@yahoo.com	commande2	320000
3	3	3	Zonzon	zazine	78954632	Douala	zazine@gmail.com	commande3	501200
4	4	4	Biboum	jack	95487445	Bafoussam	jack@gmail.com	commande4	7521454
5	5	5	Oumarou	ahmed	457896644	Garoua	ahmed@hotmail.com	commande5	2000000



Resultat dans le fichier de rejet :

id_client	nom_client	prenom	telephone	ville	email
6	Taphile	paul	4,59E+08	yaounde	paul.tap@yahoo.fr
7	Toupa	tito	45789634	ede	tito@gmail.com

Et le tout est joué