



Atelier 1 : ETL Talend et MySQL

Les objectifs de cet atelier sont les suivants :

1. **Compréhension de l'ETL (Extract, Transform, Load)** : Les participants devraient comprendre le concept général de l'ETL et son importance dans le processus d'intégration de données. Cela inclut la compréhension des différentes étapes du processus ETL, à savoir l'extraction des données à partir de sources variées, leur transformation selon les besoins métier spécifiques, et enfin leur chargement dans une base de données cible.
2. **Maîtrise de Talend** : Les participants devraient acquérir une connaissance pratique de Talend en tant qu'outil d'intégration de données. Cela implique de comprendre l'interface de Talend Studio, la manière de créer et de gérer des projets, ainsi que la manipulation des composants et des connexions pour concevoir des flux ETL efficaces.
3. **Manipulation des composants** : Les participants devraient apprendre à utiliser des composants spécifiques tels que TMap, LogRow et MySQLInput dans le cadre de la création de flux ETL. Cela comprend la compréhension des fonctionnalités de chaque composant, leur configuration appropriée et leur intégration dans des flux de données complexes.
4. **Gestion des erreurs et débogage** : Les participants devraient être capables de gérer les erreurs et les exceptions qui peuvent survenir lors de l'exécution des flux ETL. Cela implique l'utilisation de techniques de débogage telles que l'enregistrement des données à différentes étapes du processus, la gestion des exceptions et la mise en place de stratégies de récupération en cas d'échec.

1. Télécharger Talend

- https://perso.isima.fr/~jofontan/Install/TOS_DI-Win32-20200219_1130-V7.3.1.exe

2. Préparer une instance MySQL avec les données sources

Assurez-vous d'avoir le fichier SQL ([gestion_livres_init.sql](#)) dans le même répertoire que votre fichier `docker-compose.yml`.

Pour exécuter ce fichier `docker-compose.yml`, suivez ces étapes :

- a) Placez votre `docker-compose.yml` et vos fichiers SQL [gestion_livres_init.sql](#) dans un même répertoire.



Talend for Data Integration



Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com

```
CREATE DATABASE IF NOT EXISTS `gestion_livres` DEFAULT CHARACTER SET utf8;
USE `gestion_livres`;

-- Table `auteur`
DROP TABLE IF EXISTS `auteur`;
CREATE TABLE `auteur` (
  `NUMERO_A` int(10) unsigned NOT NULL AUTO_INCREMENT,
  `NOM` varchar(450) DEFAULT NULL,
  `PRENOM` varchar(450) DEFAULT NULL,
  `DOMICILE` varchar(450) DEFAULT NULL,
  PRIMARY KEY (`NUMERO_A`)
) ENGINE=InnoDB AUTO_INCREMENT=8547586 DEFAULT CHARSET=latin1;

-- Données pour table `auteur`
INSERT INTO `auteur` VALUES
(8547,'Chambord','Emilie','Nice'),
(52136,'Fabiere','Sylvie','Bordeaux'),
(78545,'TINTIN','Thierry','Clermont'),
(85478,'Castafiore','Emilie','Paris'),
(542536,'Dupont','Pierre','Avignon'),
(8547585,'Momo','Roland','Toulouse');

-- Table `livre`
DROP TABLE IF EXISTS `livre`;
CREATE TABLE `livre` (
  `NUMERO_L` int(10) unsigned NOT NULL AUTO_INCREMENT,
  `NUMERO_A` int(10) unsigned NOT NULL,
  `TITRE` varchar(128) DEFAULT NULL,
  `PRIX` float DEFAULT NULL,
  PRIMARY KEY (`NUMERO_L`)
) ENGINE=InnoDB AUTO_INCREMENT=11324 DEFAULT CHARSET=latin1;

-- Données pour table `livre`
INSERT INTO `livre` VALUES
(10101,85478,'aaaaaa',10),
(11111,8547,'eeeeee',54),
```

```
(11122,52136,'fffff',12),  
(11231,542536,'bbbbbb',34),  
(11321,8547585,'eeeeeee',29),  
(11323,8547585,'ccccccccc',45);
```

- b) Ouvrez un terminal dans ce répertoire.
- c) Exécutez la commande suivante pour démarrer les services définis dans le fichier `docker-compose.yml`:

```
version: '3.8'

services:
  mysql_db:
    image: mysql:latest
    container_name: mysql_db
    ports:
      - "3306:3306"
    environment:
      MYSQL_ROOT_PASSWORD: mysecret
      MYSQL_DATABASE: gestion_livres
    volumes:
      - ./gestion_livres_init.sql:/docker-entrypoint-initdb.d/gestion_livres_init.sql
    restart: always # Ajout de cette ligne pour redémarrer le conteneur en cas d'échec

  networks:
    - mysql_network

networks:
  mysql_network:
    driver: bridge
```



Talend for Data Integration

Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com



```
mysql> UNLOCK TABLES;
Query OK, 0 rows affected (0.00 sec)

mysql> show tables;
+-----+
| Tables_in_gestion_livres |
+-----+
| auteur                   |
| livre                    |
+-----+
2 rows in set (0.00 sec)

mysql> select * from livre
->
+-----+
| NUMERO_A | NUMERO_B | TITRE   | PRIX |
+-----+
| 10101   | 85478   | aaaaaa  | 16  |
| 11111   | 8547    | eeeee   | 54  |
| 11122   | 52136   | ffffff  | 12  |
| 11231   | 542536  | bbbbb   | 34  |
| 11321   | 8547585 | eeeeeff | 29  |
| 11323   | 8547585 | cccdddf | 45  |
+-----+
6 rows in set (0.00 sec)

mysql> select * from auteur
->
+-----+
| NUMERO_A | NOM      | PRENOM | DOMICILE |
+-----+
| 8547    | Chambord | Emilie | Nice     |
| 52136   | Fabriere | Sylvie | Bordeaux |
| 78545   | TINTIN   | Thierry | Clermont |
| 85478   | Castaflore | Emilie | Paris     |
| 542536  | Dupont   | Pierre  | Avignon  |
| 8547585 | Momo     | Roland  | Toulouse |
+-----+
6 rows in set (0.00 sec)

mysql>
```

3. Créer un projet

Talend Open Studio for Data Integration

Select what you want to do next:

Create a new project: Local_Project

Import a demo project

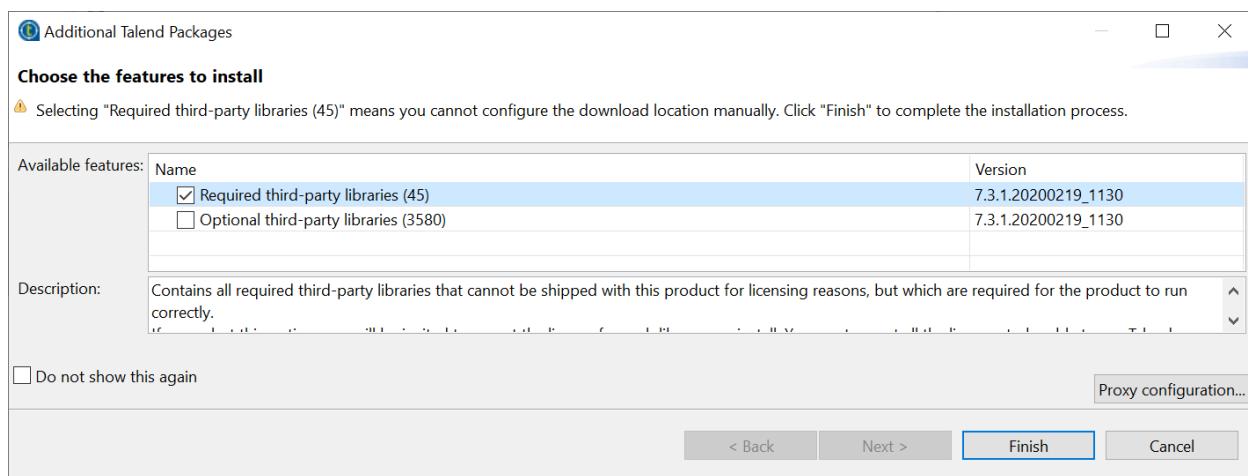
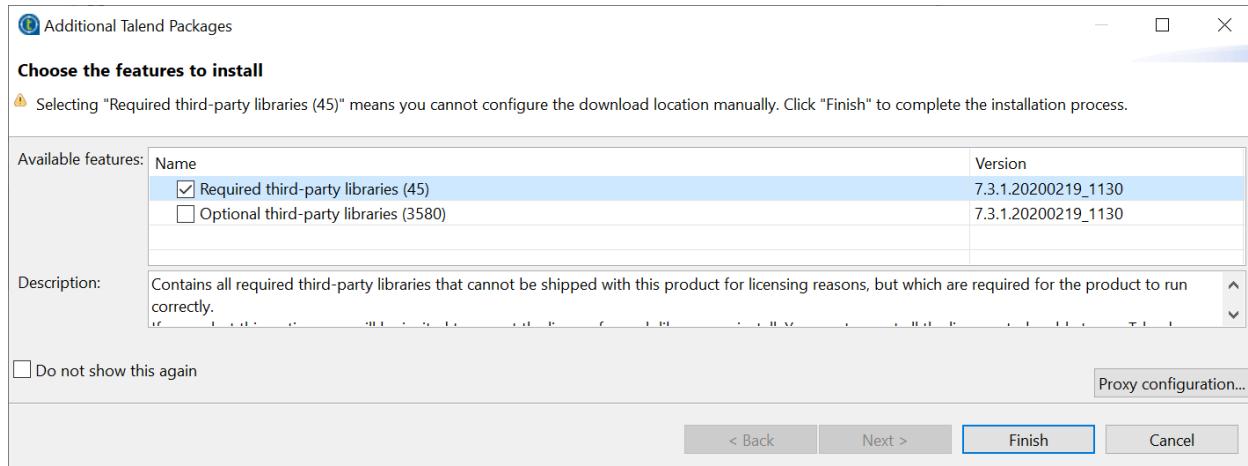
Import an existing project

Manage Connections

Always ask me at startup

Finish

4. Installer les fonctionnalités avancées

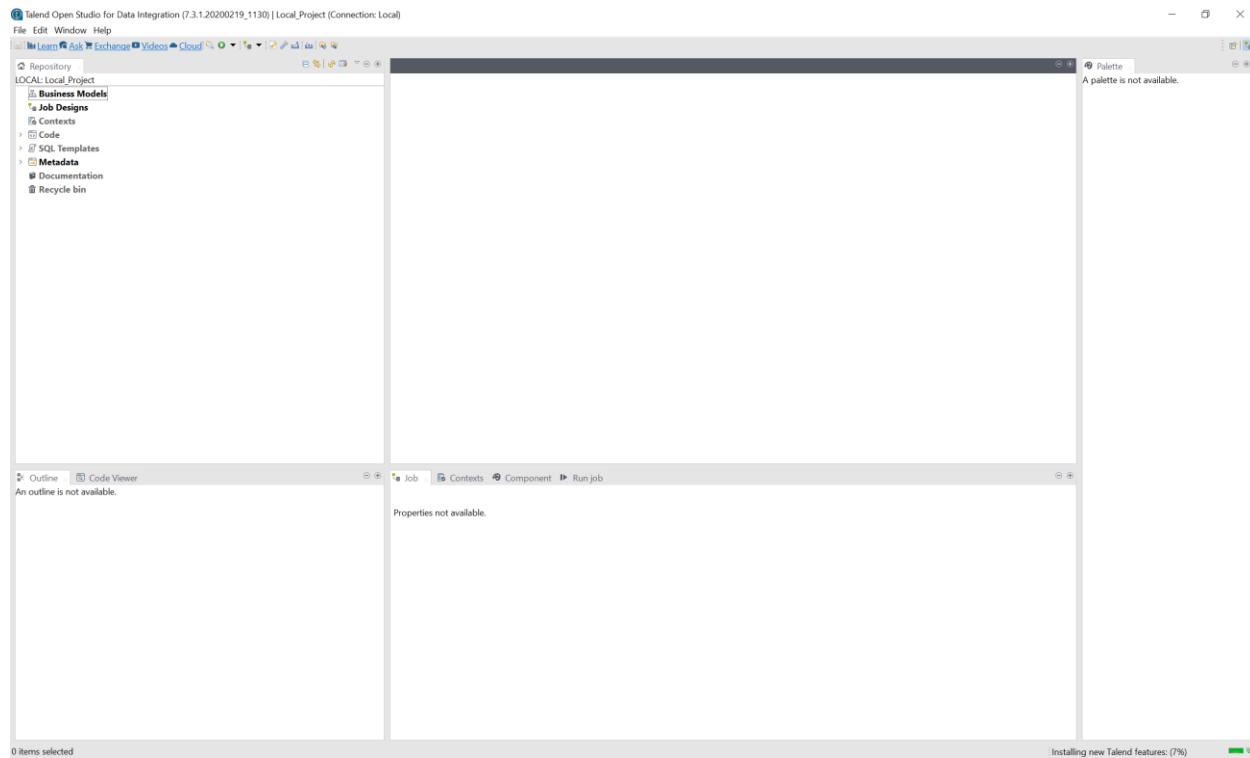


5. Interface d'accueil de Talend

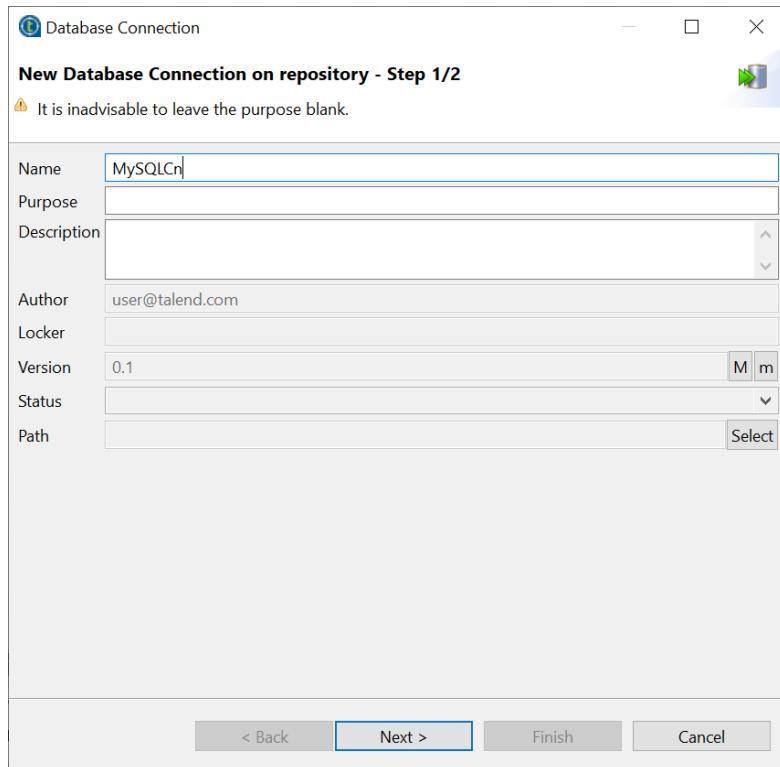


Talend for Data Integration

Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com



6. Création d'une connexion à une base MySQL :

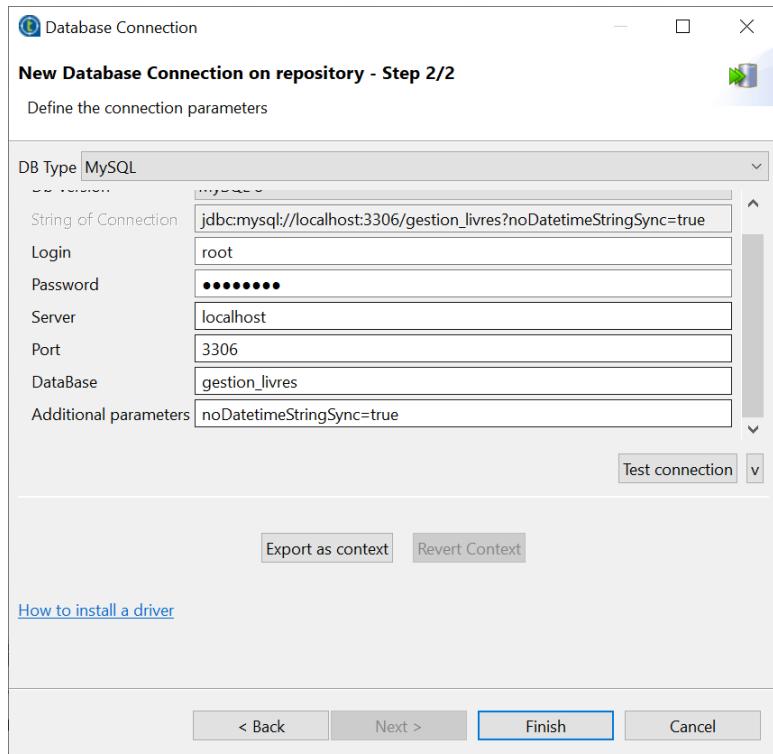


7. Configurer les paramètres de connexion à la base de données

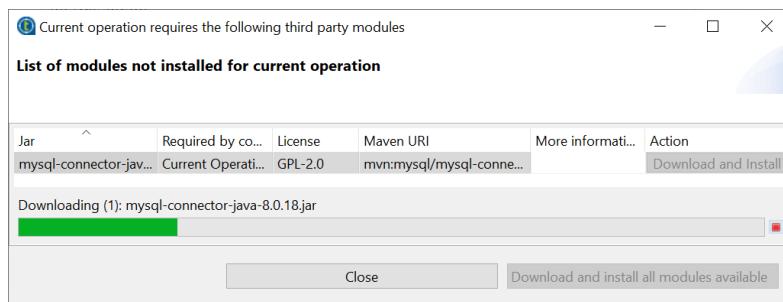


Talend for Data Integration

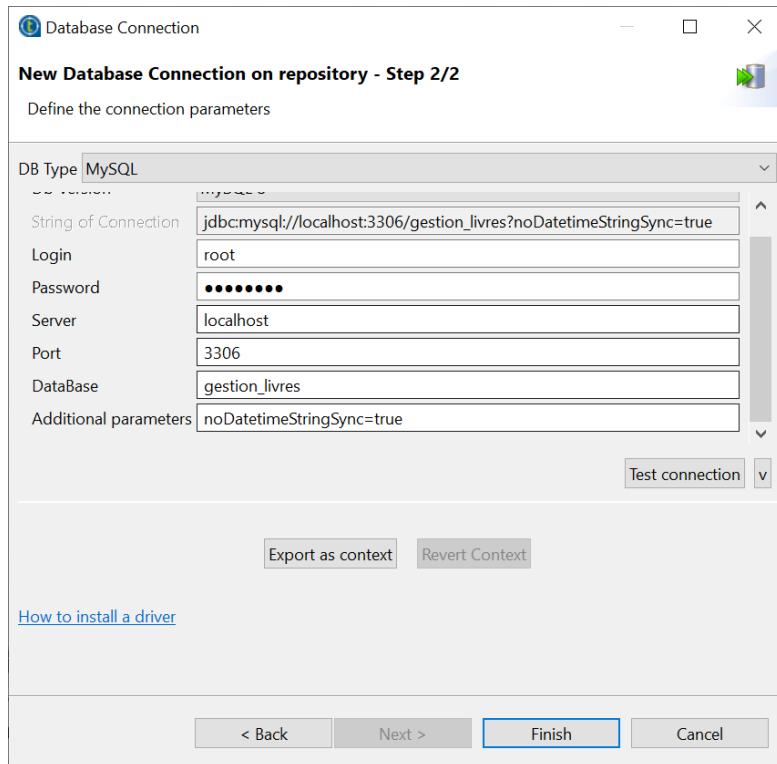
Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com



8. Installer les jars nécessaires

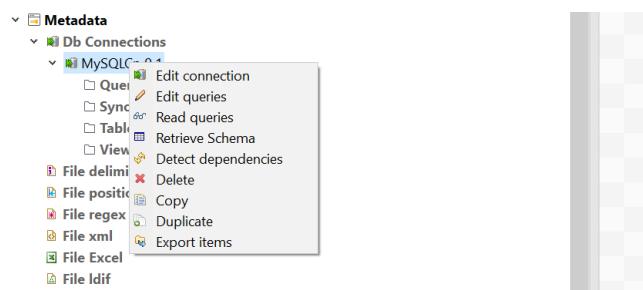


9. Tester la connexion



10. Création d'une connexion à une base MySQL :

– Récupération du schéma des tables



11. Cocher Table et cliquer « Next »



Talend for Data Integration

Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com



Schema

Filter for the Table.

Select Filter Conditions
 Use the Name Filter Use the Sql Filter

Select Types
 TABLE VIEW SYNONYM

Set the Name Filter:
% New...
Edit...
Remove...

Set the Sql Filter:
SELECT TNAME FROM TAB WHERE TNAME LIKE 'BAL%'

< Back Next > Finish Cancel

12. Sélectionner la base de données « gestion livres »

Schema

New Schema in connection "MySQLCn"

Add a Schema on repository

Select Schema to create

Name Filter:

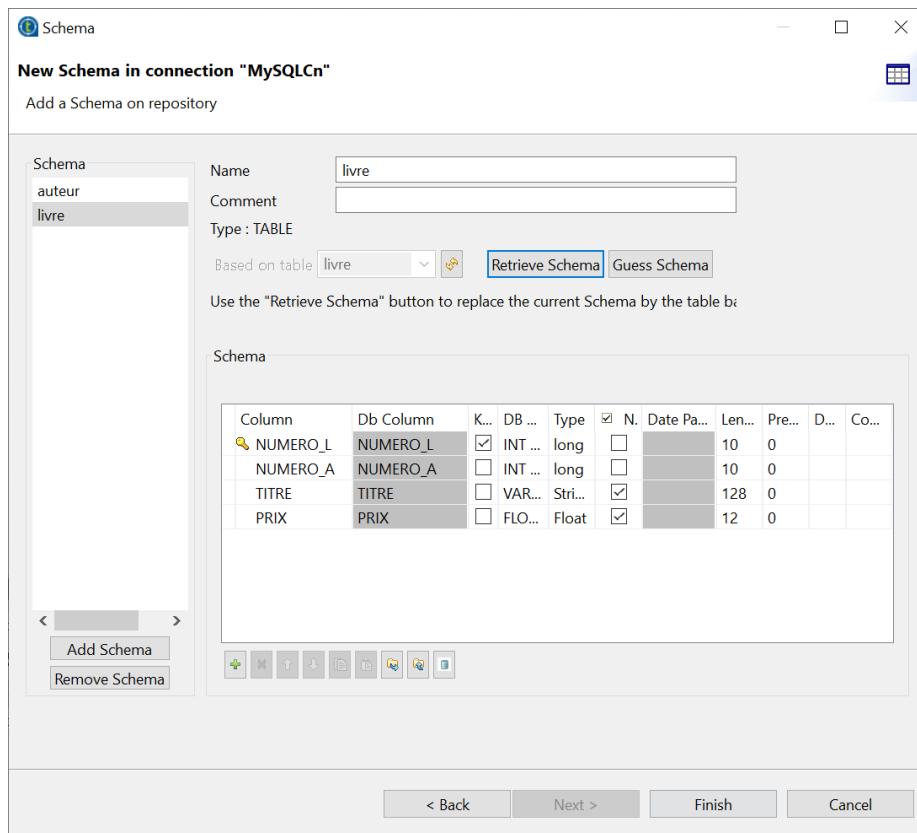
Name	Type	Column nu...	Creation status
▼ <input checked="" type="checkbox"/> gestion_livres	CATALOG		
<input checked="" type="checkbox"/> auteur	TABLE		
<input checked="" type="checkbox"/> livre	TABLE		

Select All Select None Check Connection

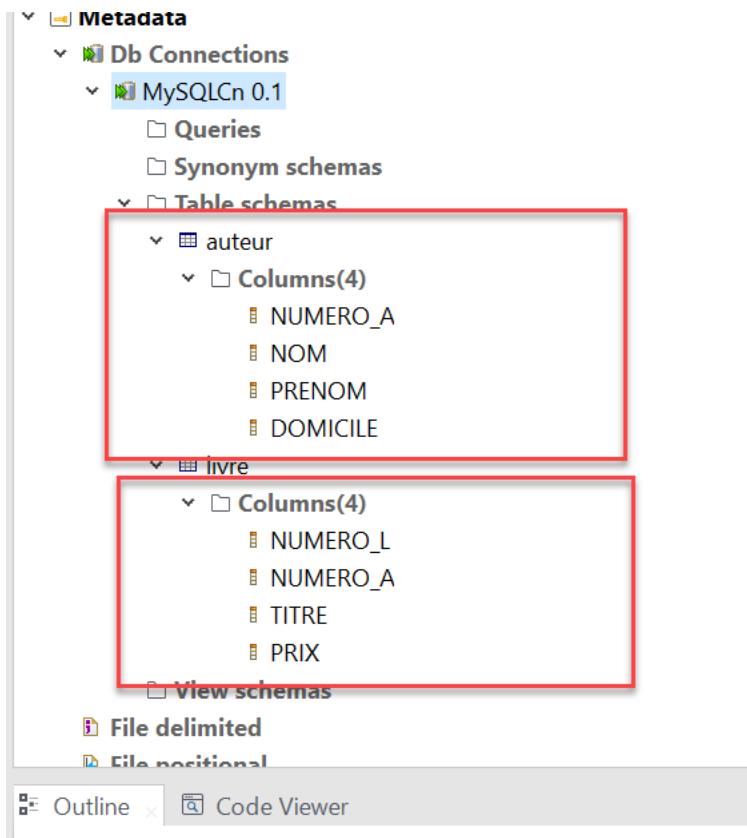
< Back Next > Finish Cancel

Name	Type	Column nu...	Creation status
▼ <input checked="" type="checkbox"/> gestion_livres	CATALOG		
<input checked="" type="checkbox"/> auteur	TABLE		
<input checked="" type="checkbox"/> livre	TABLE		

13. Choisir les colonnes désirées en cas de besoins et cliquer « Finish »

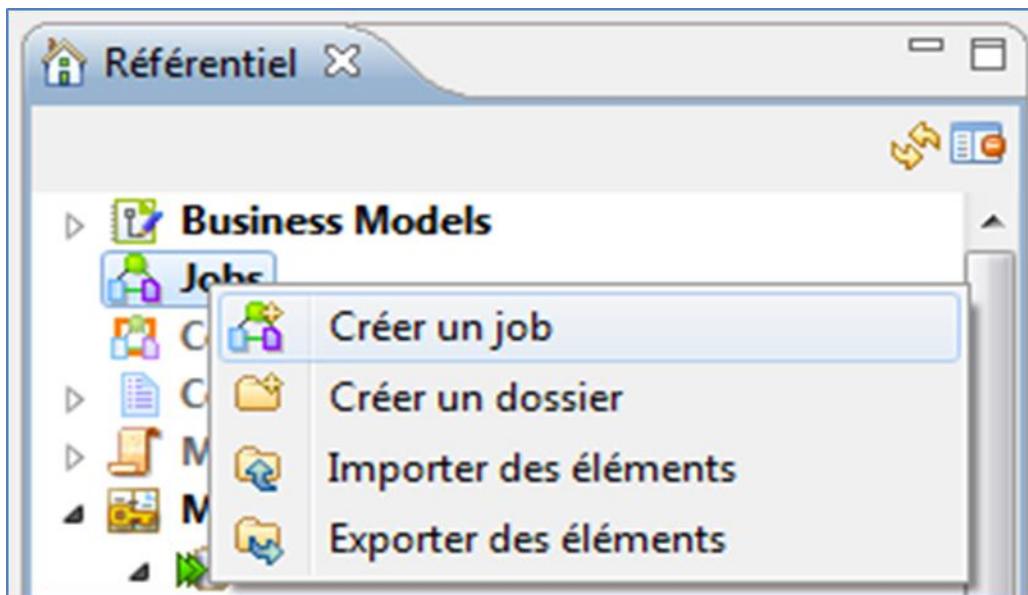


14. Vérifier les schémas de chaque table



The screenshot shows the Talend Studio interface for managing database metadata. The left pane displays a tree structure of database connections, schemas, and tables. Two specific tables are highlighted with red boxes: 'auteur' and 'livre'. The 'auteur' table schema is shown with four columns: NUMERO_A, NOM, PRENOM, and DOMICILE. The 'livre' table schema is shown with five columns: NUMERO_L, NUMERO_A, TITRE, and PRIX. The bottom navigation bar includes tabs for 'Outline' and 'Code Viewer'.

15. Création d'un premier job :



16. Création d'un premier job : ajout de la connexion

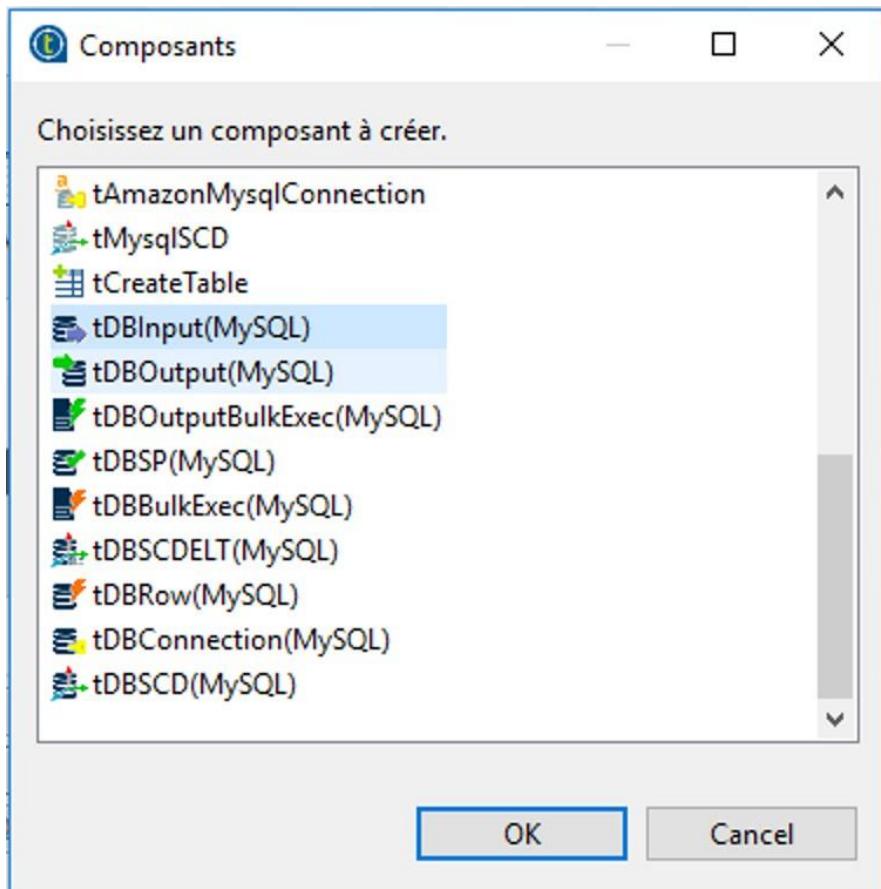
– Faire glisser le tDBInput (MySQL) dans la partie Design avec la table Auteur

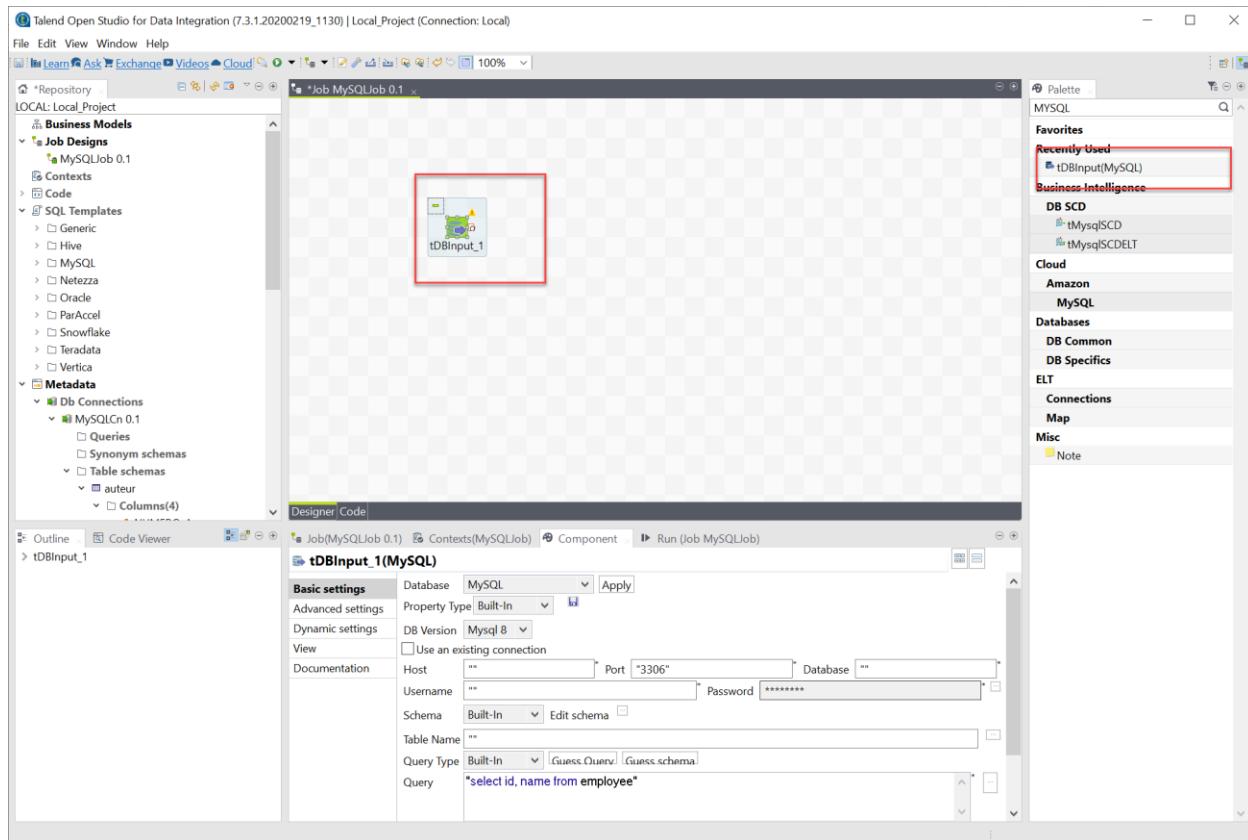


Talend for Data Integration



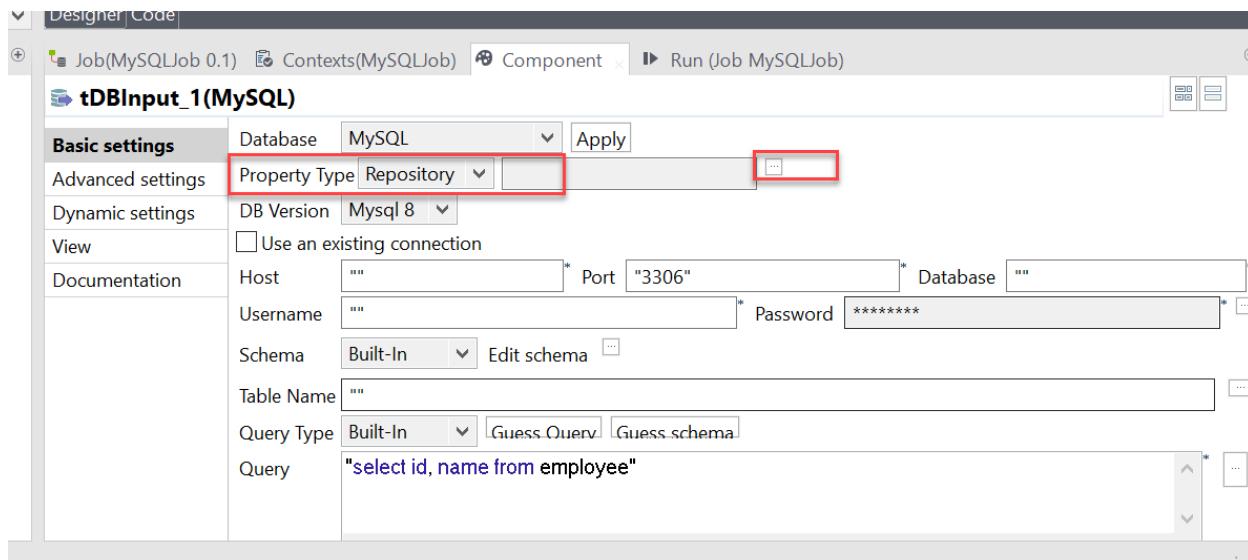
Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com



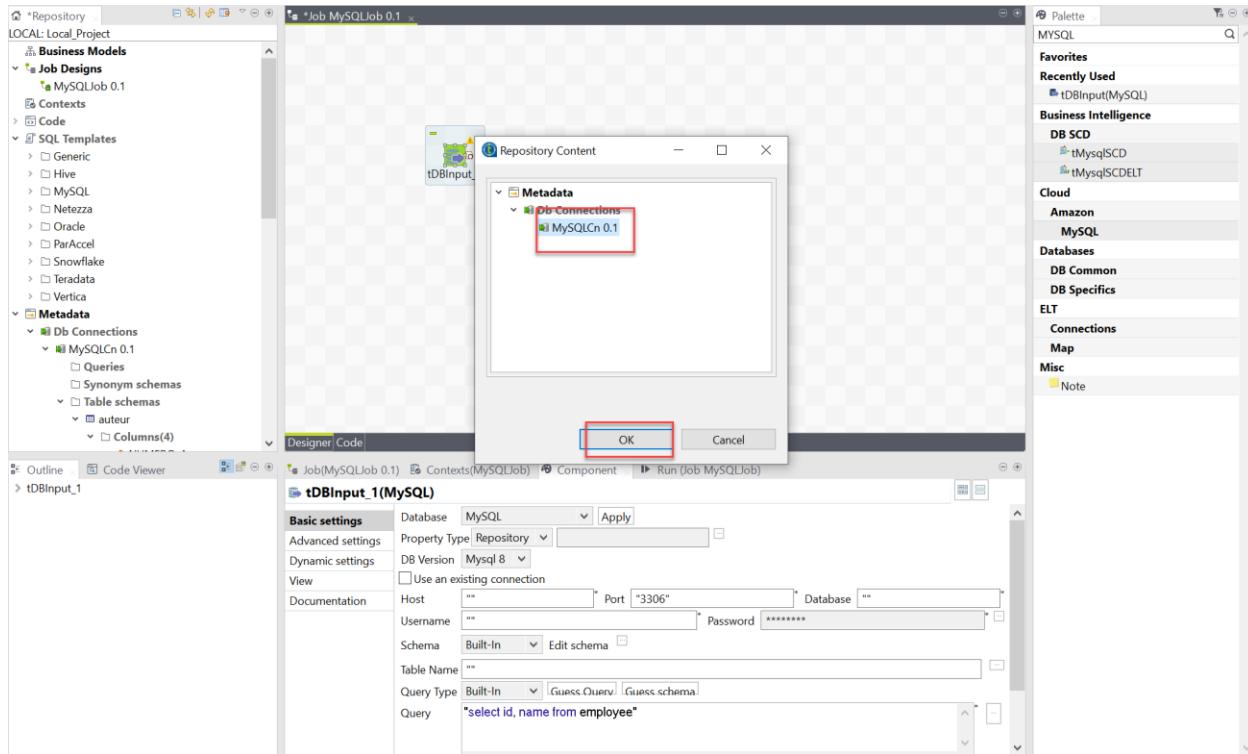


17. Configurer le composant tDBInput(MYSQL)

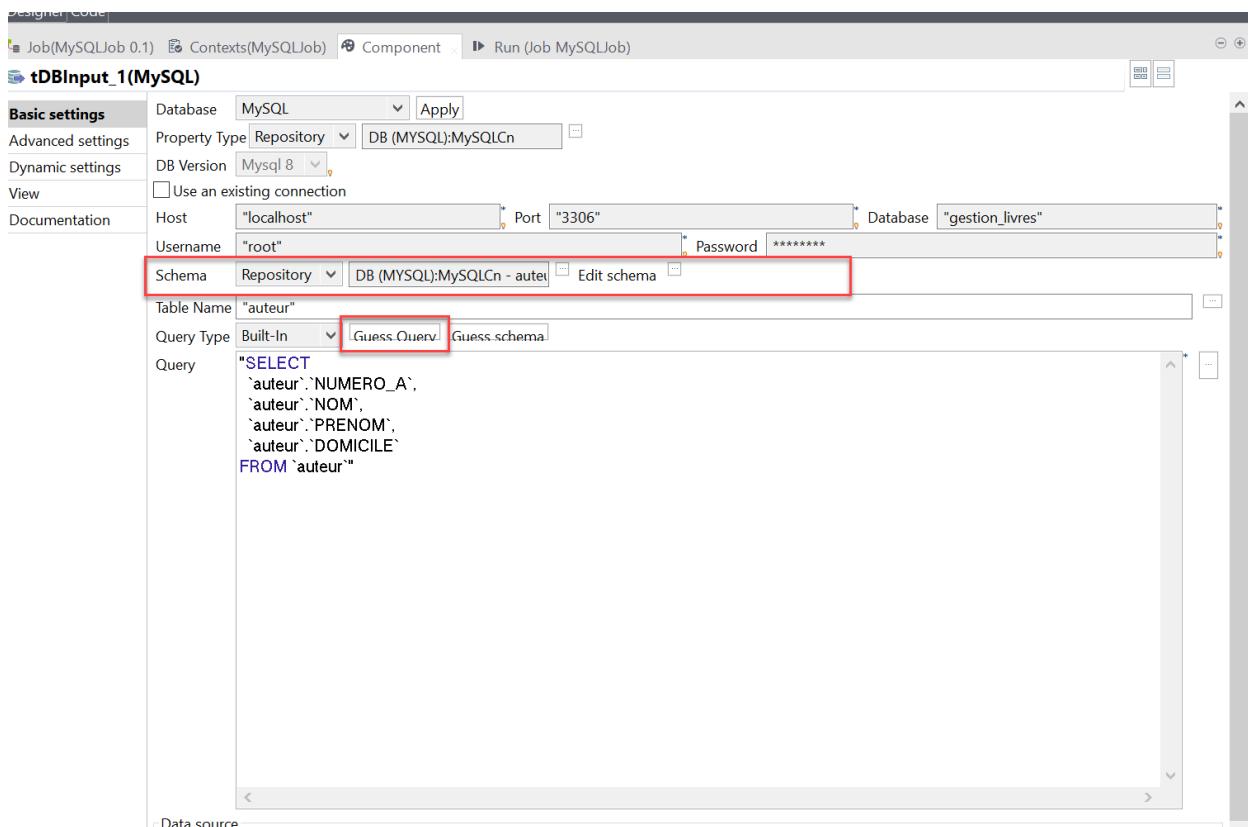
Selectionner le type « Property Type » la valeur « Repository » et cliquer sur



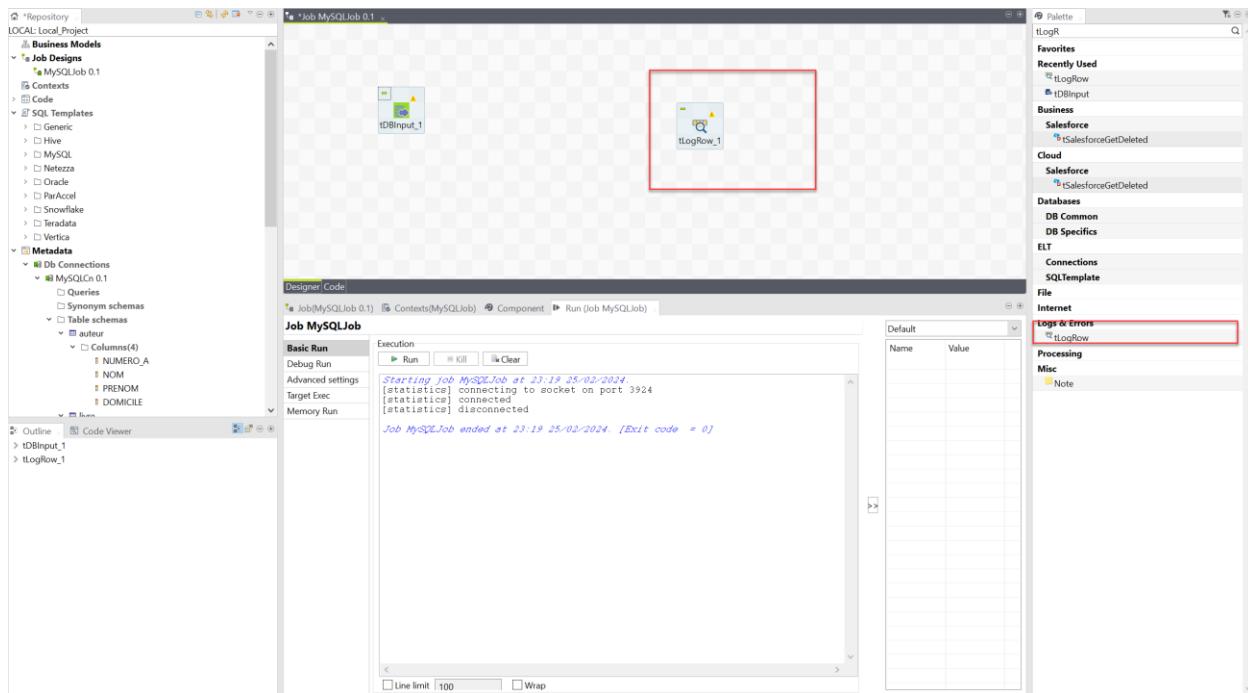
Sélectionner le type « MYSQLCN 0.1et cliquer sur **OK**



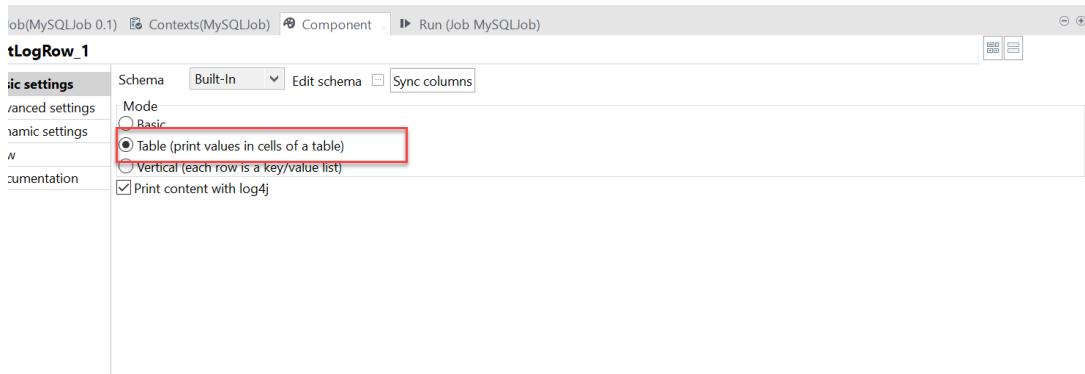
configurer le schéma et la requête à exécuter en cliquant sur Gues Query pour générer une requête de sélection avec les champs nécessaires



18. Création d'un premier job : ajout d'un composant de log

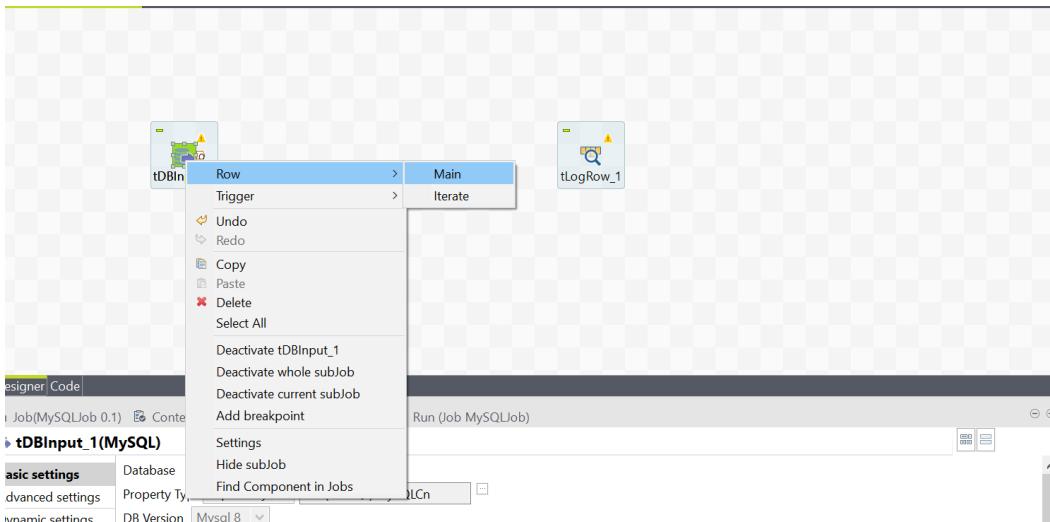


19. Configurer le composant de log



20. Liaison des composants

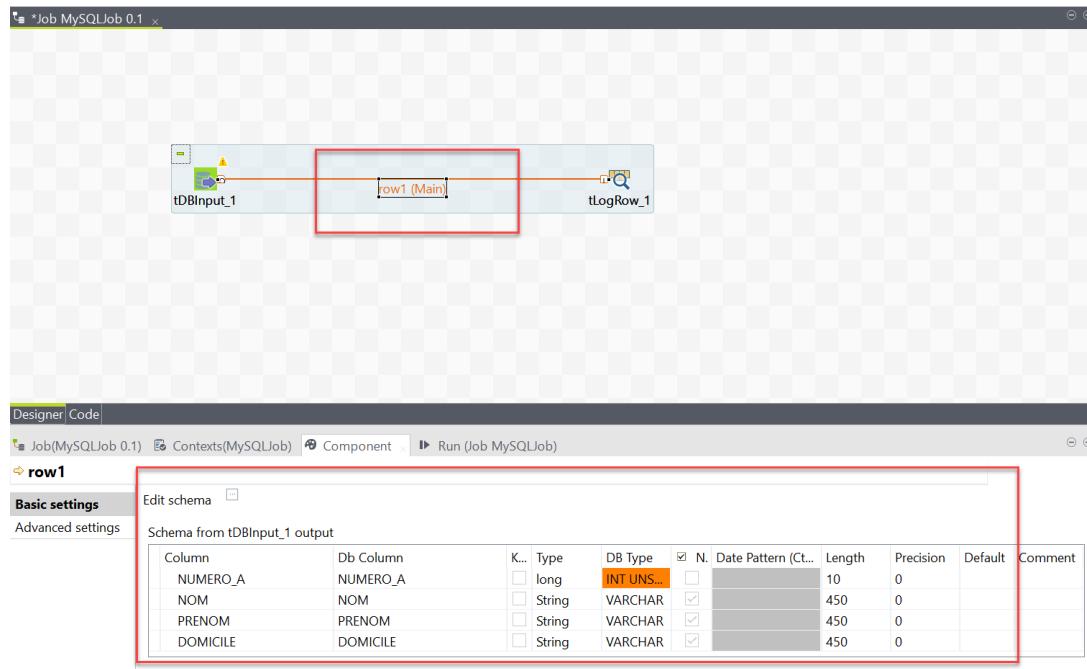
- Cliquer sur le composant tbDinput(MySQL) avec le bouton droit et afficher le menu contextuel



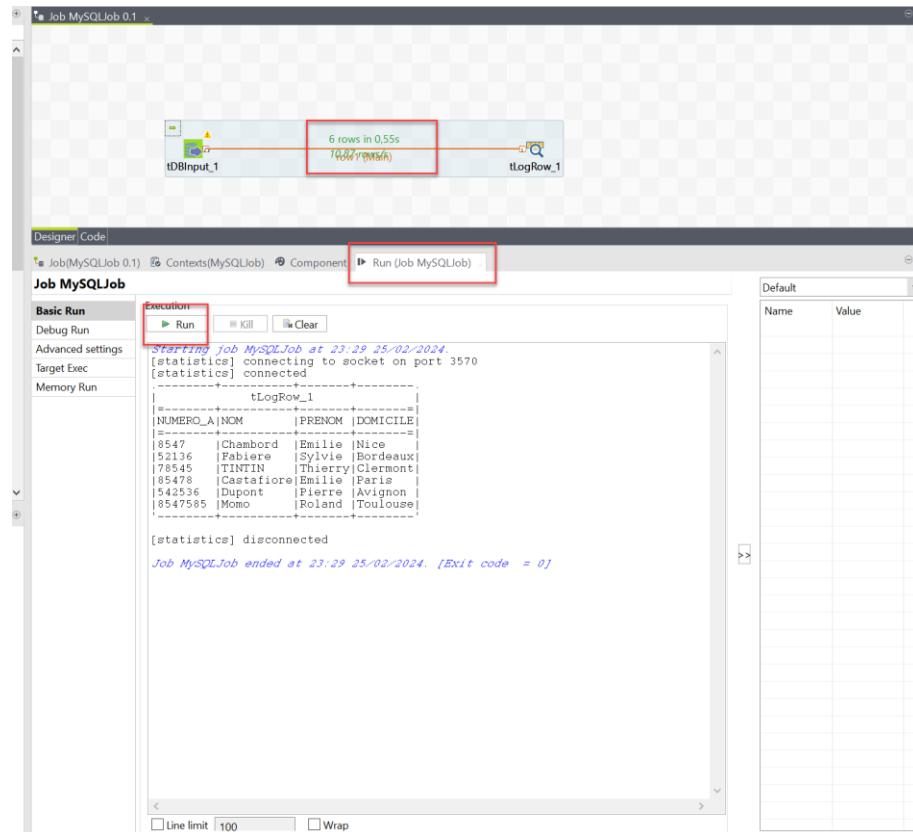
- Choisir Row -> Main et cliquer sur le composant tLogRow_1 pour établir la liaison



- La liaison est établie en fournissant l'output attendu

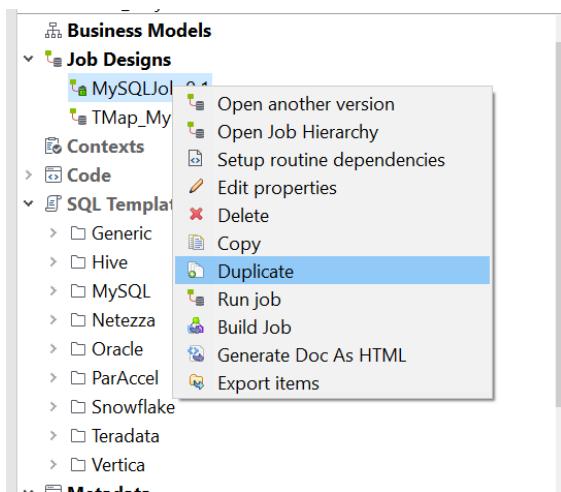


21. Exécuter le Job

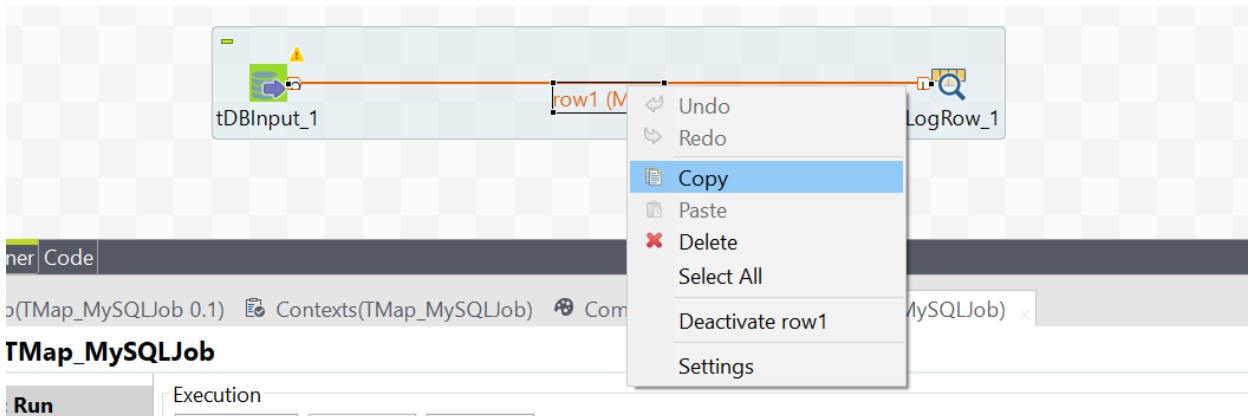


22. Modification de ce job : ajout d'un composant d'aiguillage tMap

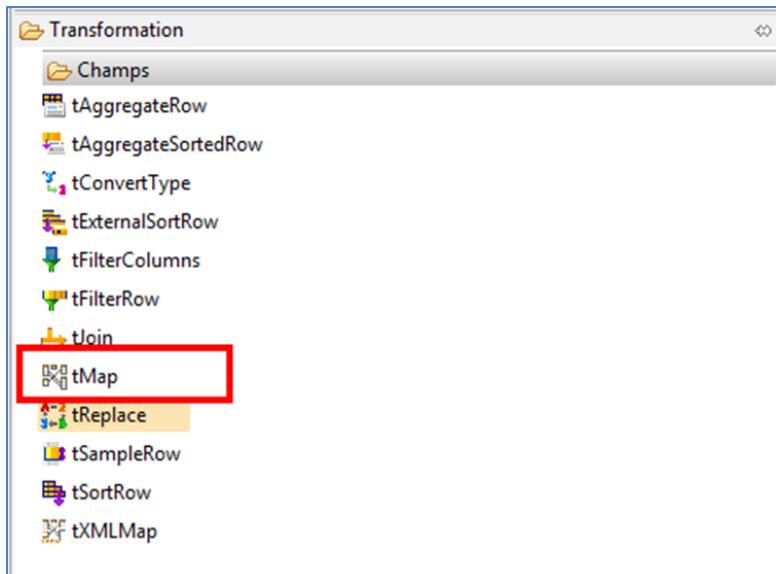
- Duplicer le projet



- Supprimer la liaison dans le nouveau job



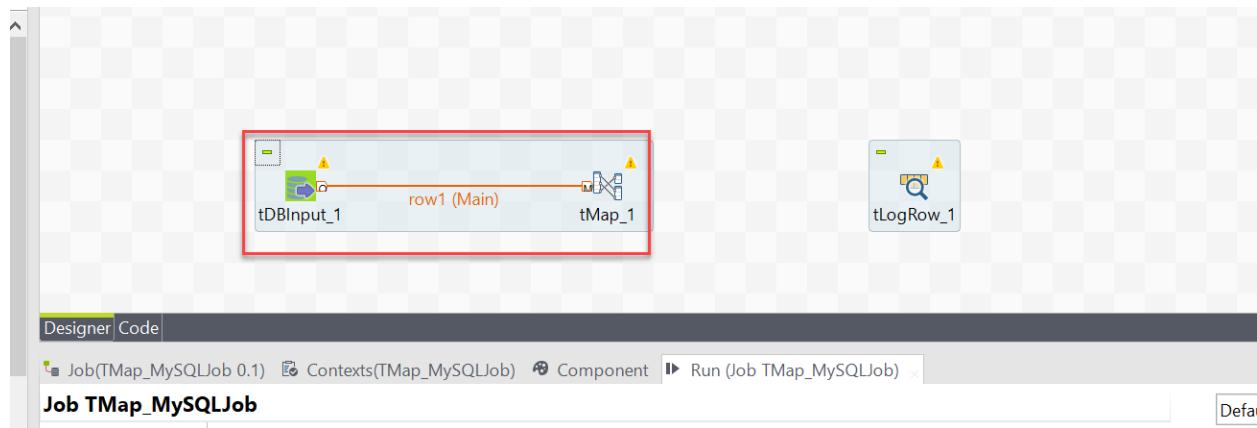
- Ajouter un composant tMap



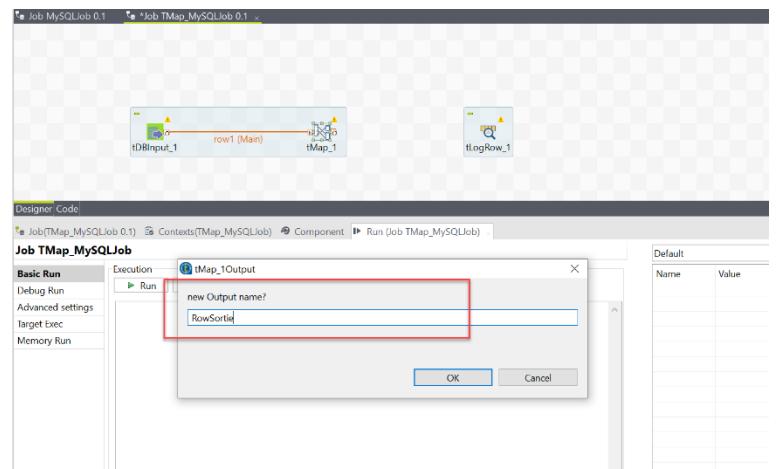
- Modification de ce job : liaison avec les autres composants

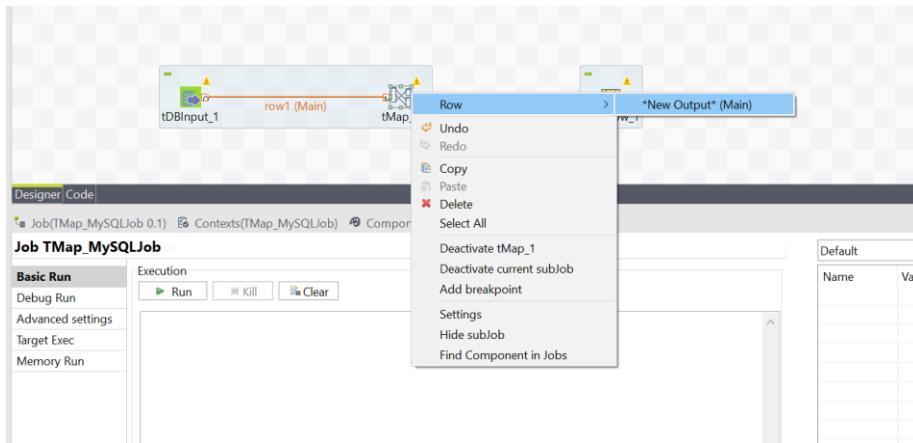


- Modification de ce job : Ajout liaison entre liaison Input(MySQL) entre tMap

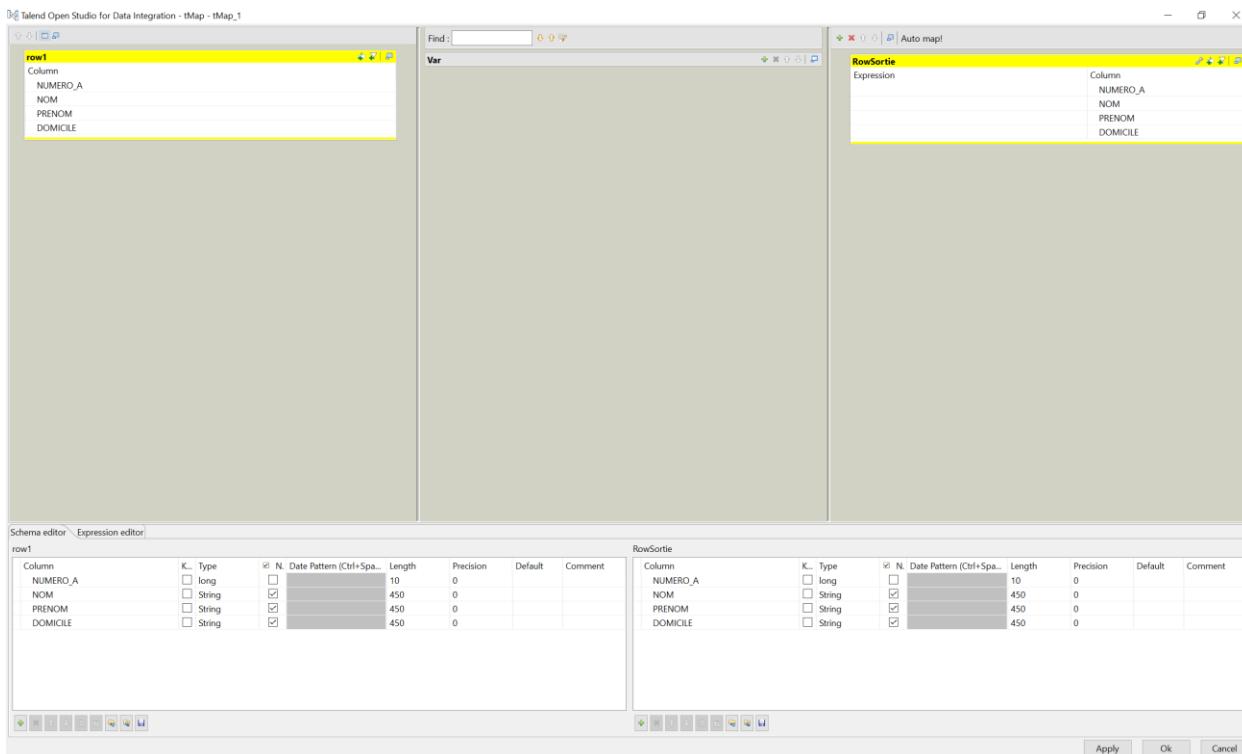


- Etablir liaison entre tMap et LogRow en spécifiant le nom de l'output



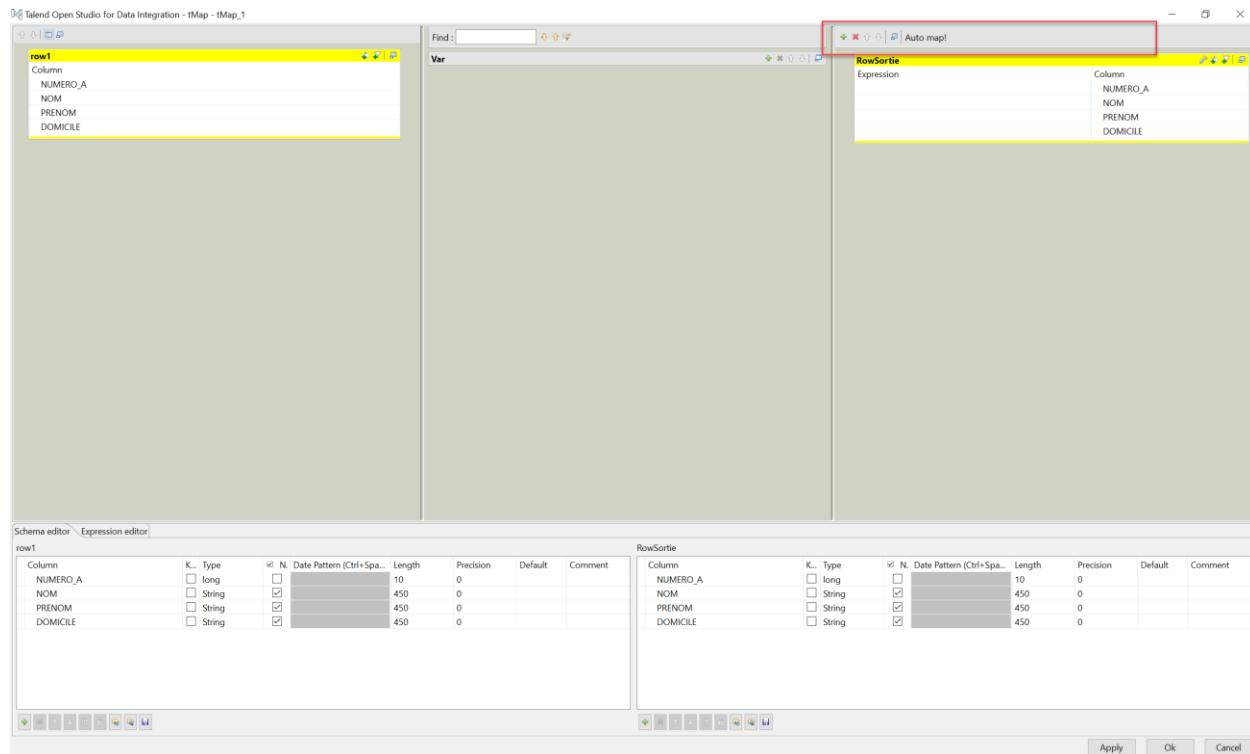


- Ouvrir le tMap

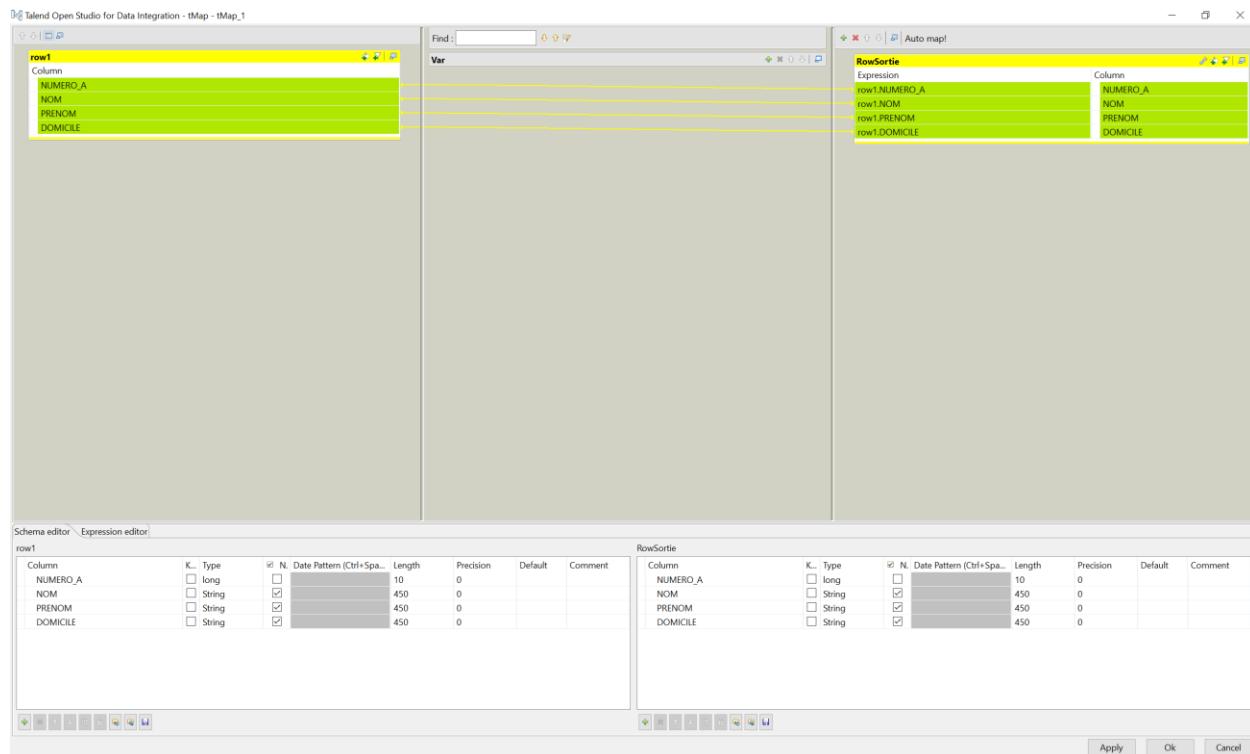


- Puis appuyer sur mapping auto en haut à droite

Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com

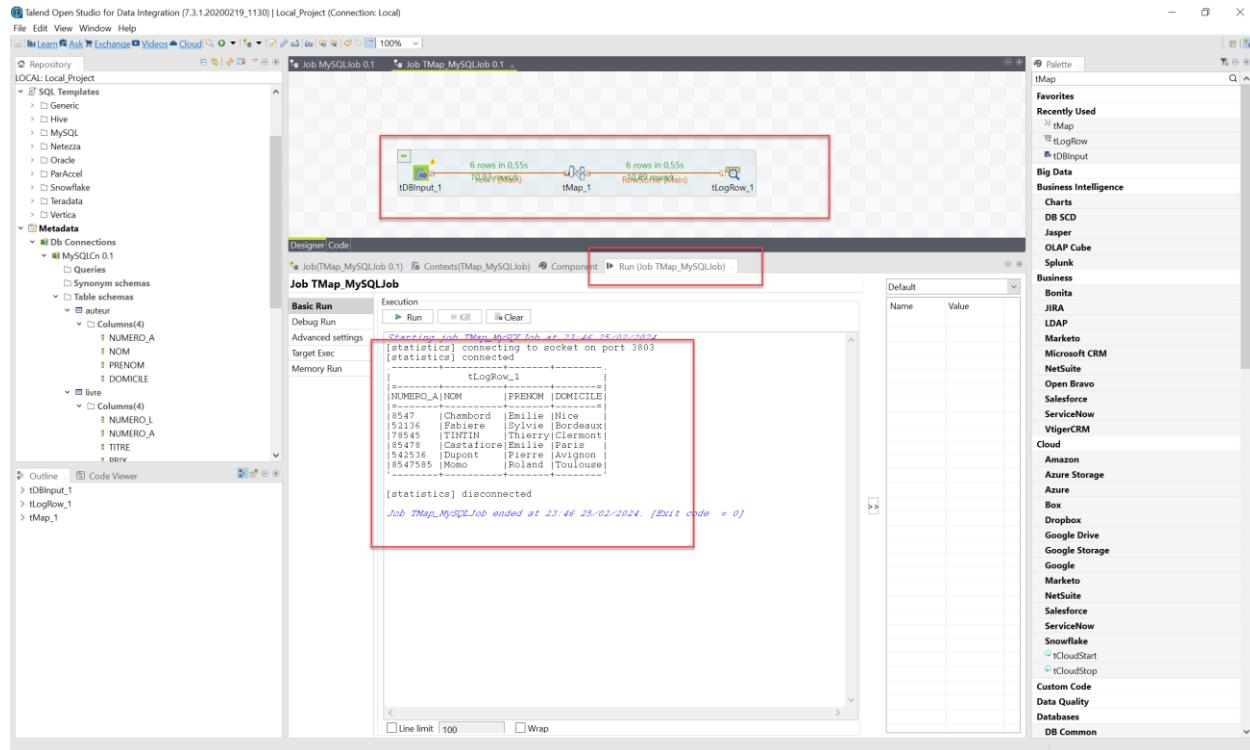


- Une correspondance sera établie automatiquement



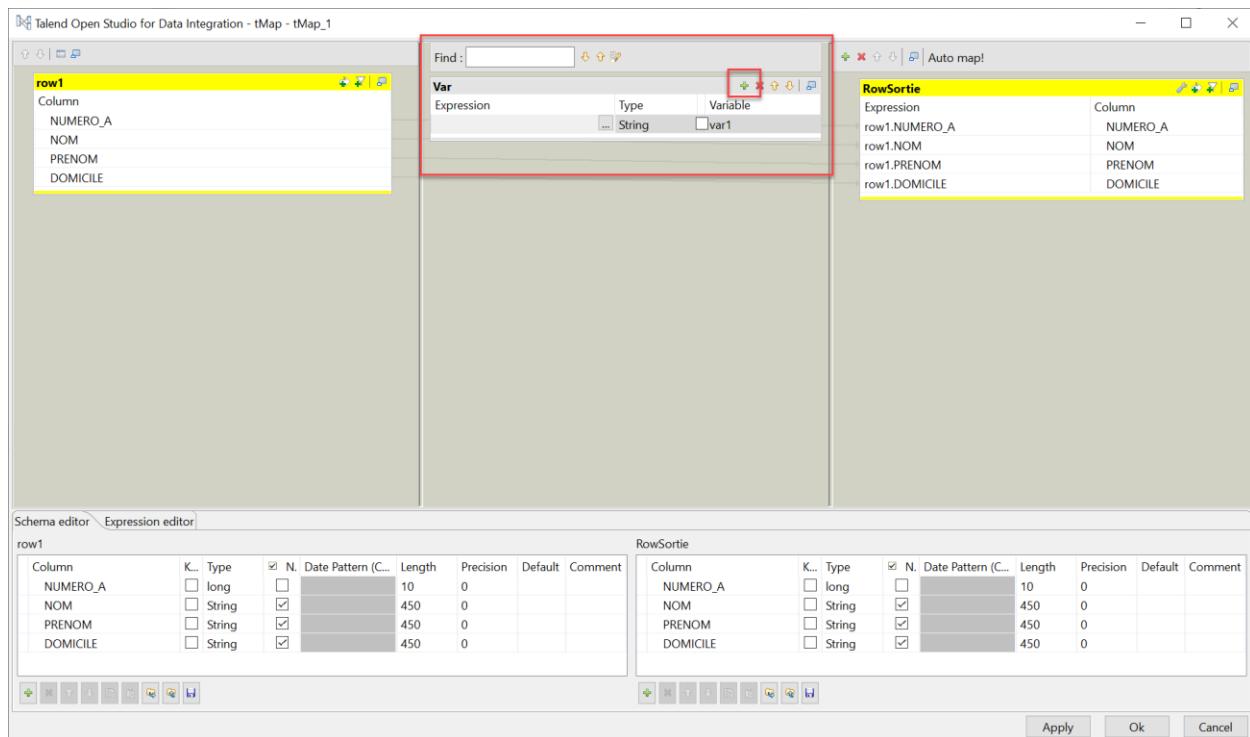
Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com

- Exécuter le job à nouveau (normalement il n'y a pas de changement)



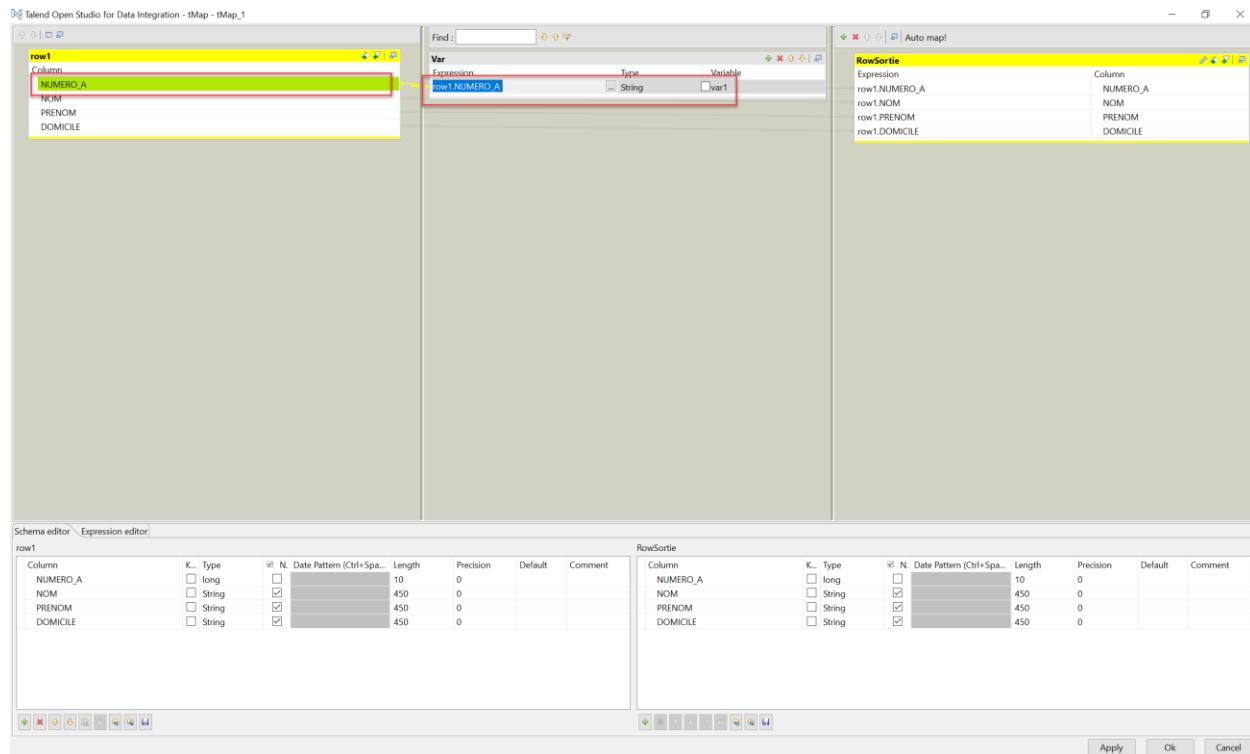
23. Modification de ce job :

- Ouvrir à nouveau le tMap et créer le champ var1 dans la partie droite.



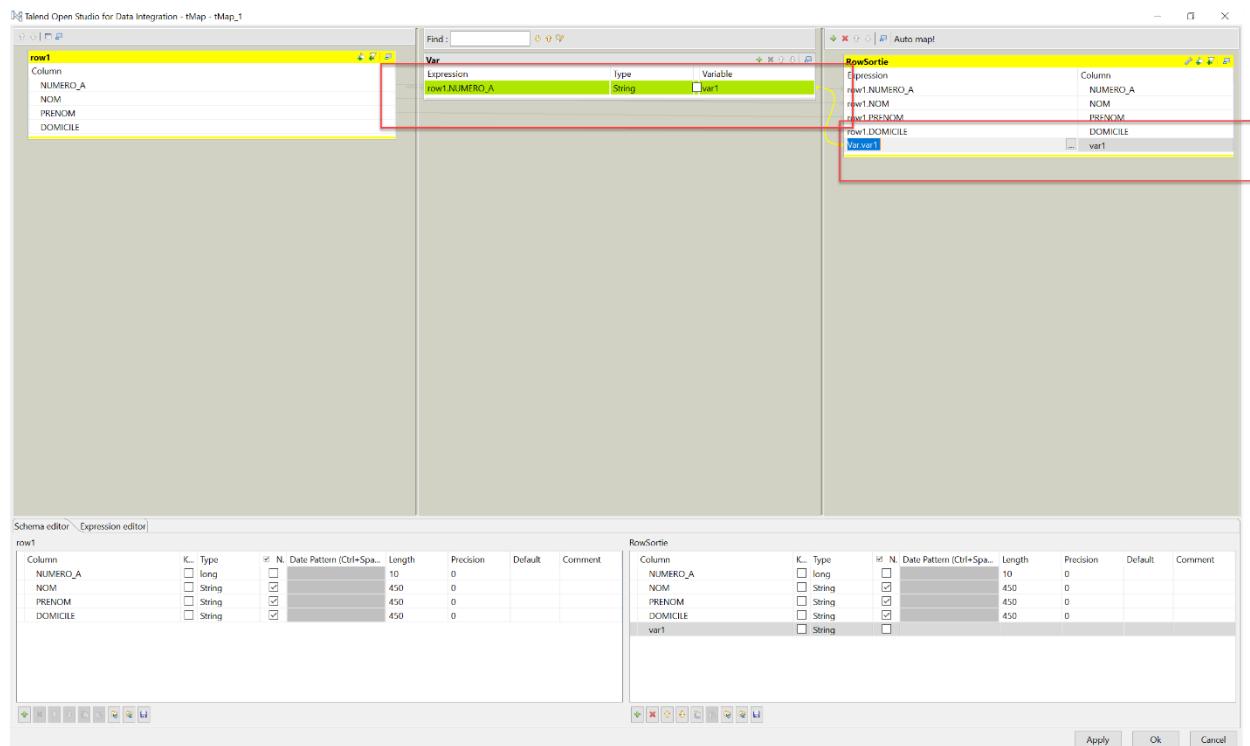
- Etablir une liaison entre la variable et Numéro en glissant le Numero_A vers la variable

Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com



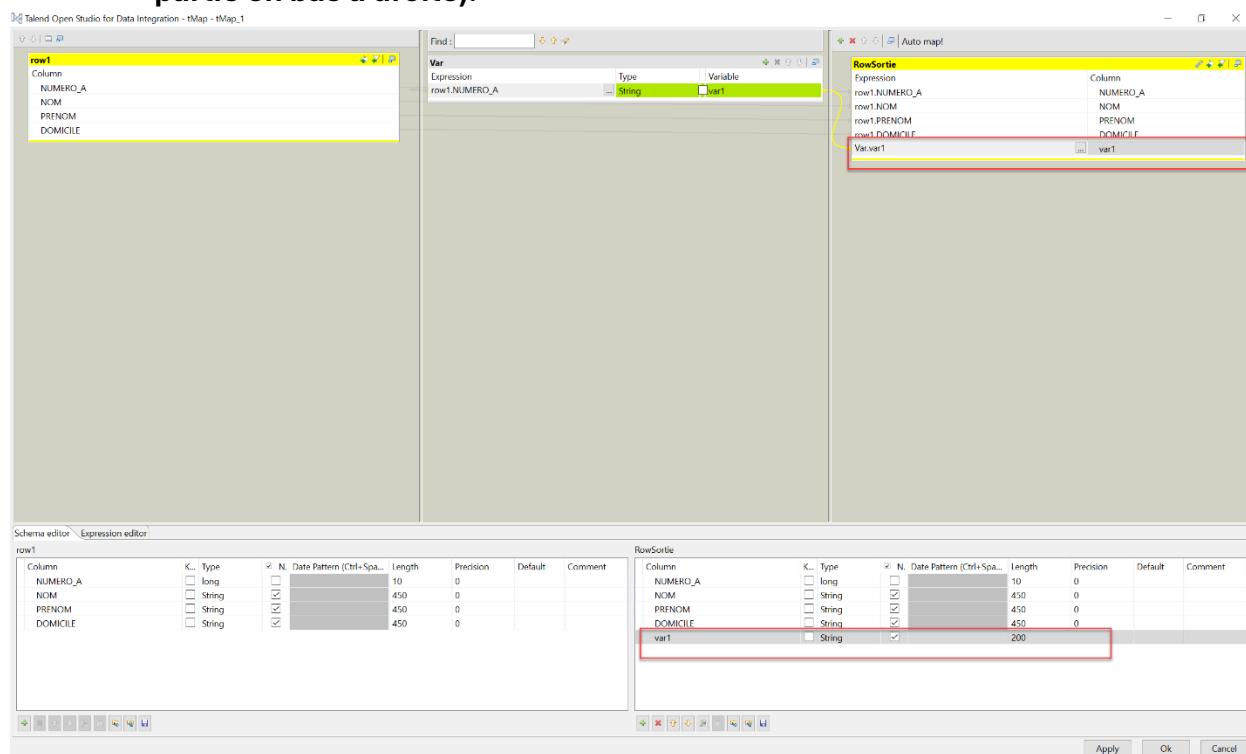
The screenshot shows the Talend Open Studio interface for the iMap component. The central panel displays a mapping rule where the expression `row1.NUMERO_A` is mapped to the variable `var1`. The 'RowSortie' panel on the right lists the output columns: NUMERO_A, NOM, PRENOM, and DOMICILE, all associated with the variable `var1`.

- Glisser cette variable vers la sortie**



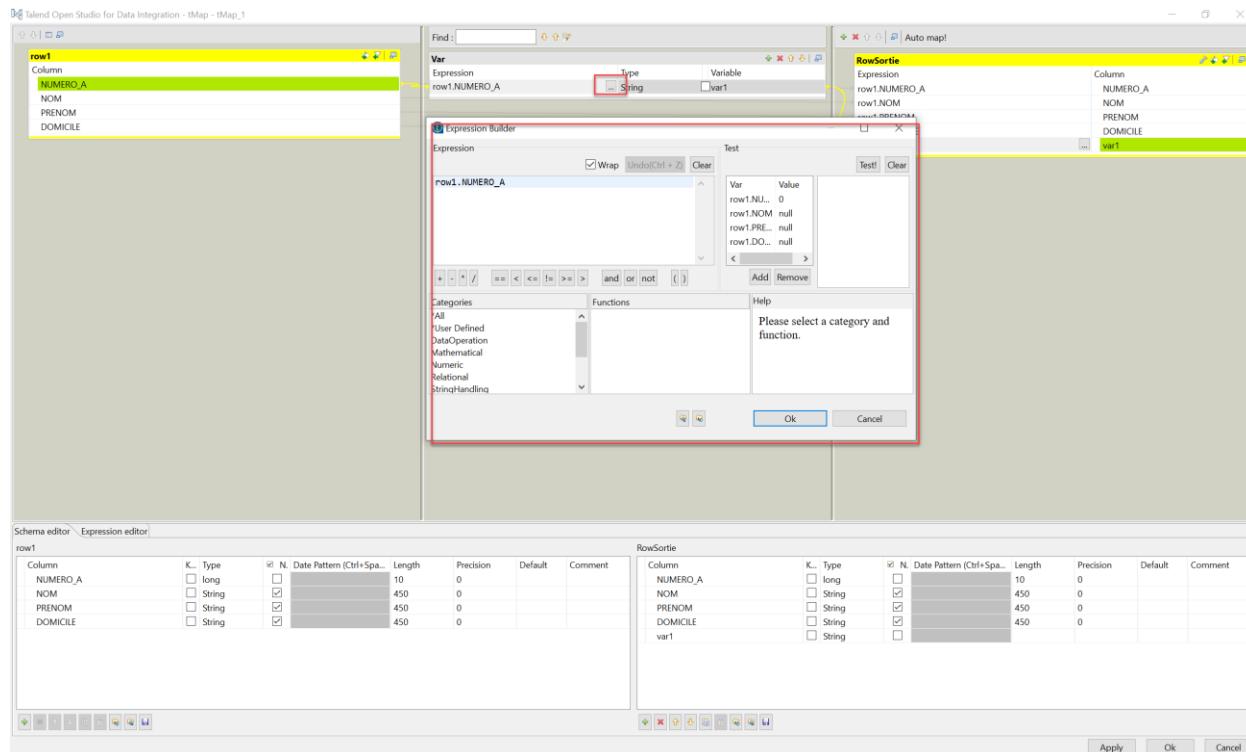
The screenshot shows the Talend Open Studio interface for the iMap component after dragging the variable `var1` from the central panel to the 'RowSortie' panel. The 'RowSortie' panel now lists the output columns: NUMERO_A, NOM, PRENOM, and DOMICILE, all associated with the variable `var1`.

- Cocher que celui-ci peut être null et qu'il peut avoir une taille de 200 (cf. partie en bas à droite).**

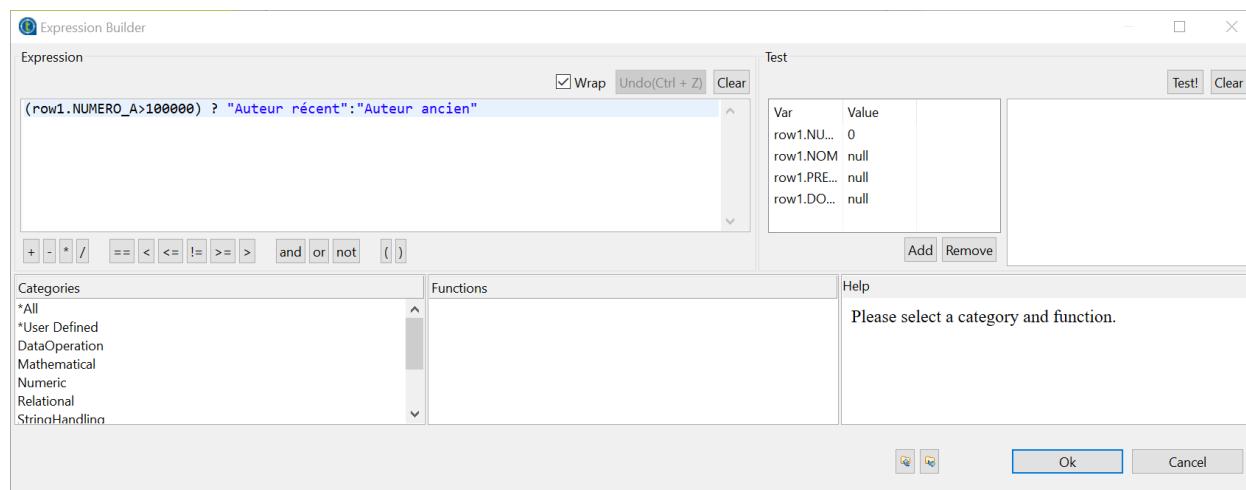


- Créer un nouveau champ dans la colonne du milieu puis cliquer sur les « ... » pour éditer l'expression :**

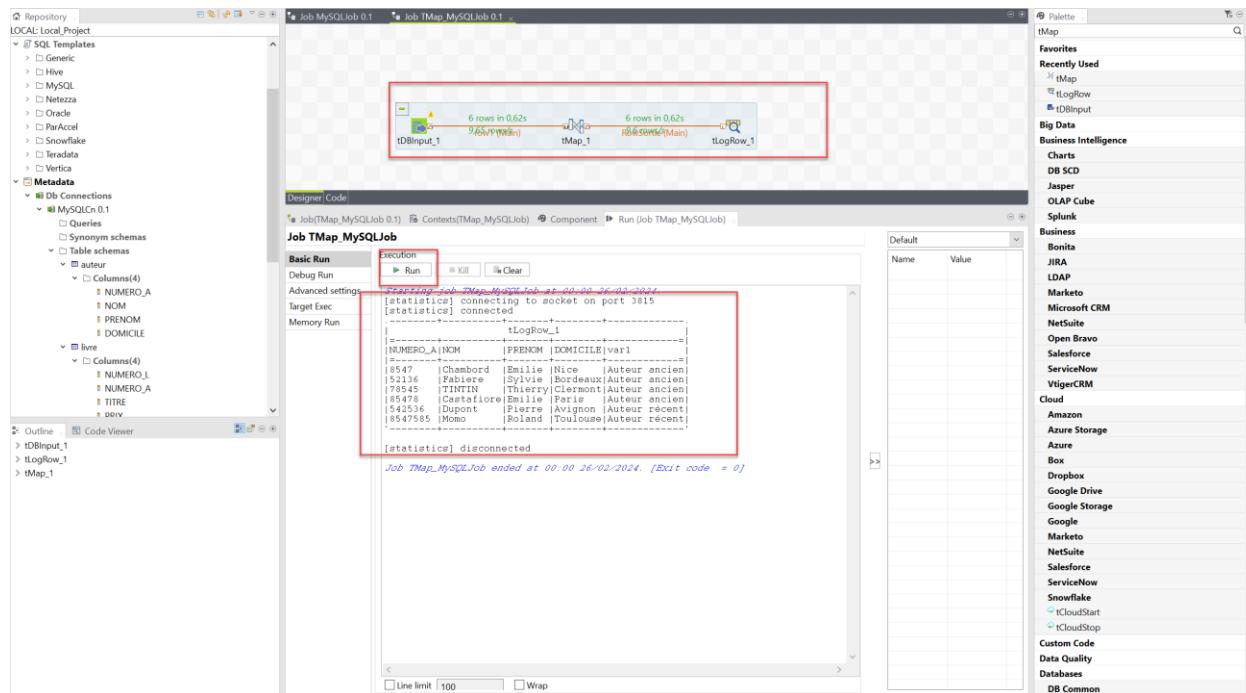
Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com



- Ajouter le test ternaire dans le nouveau champ :
- `(row1.NUMERO_A>100000)?"Auteur récent":"Auteur ancien"`

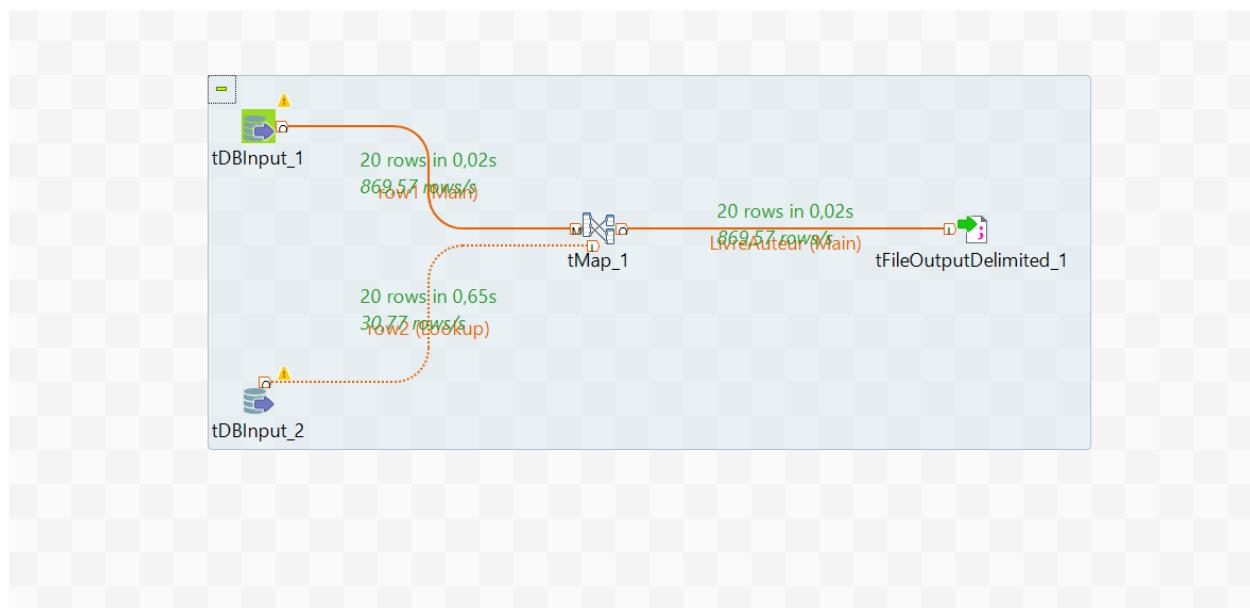


- Exécuter à nouveau le job, vous devriez obtenir ce type de résultat



Exercice à faire

L'objectif de cet exercice est de créer une ETL permettant de fusionner les tables "auteurs" et "livres" afin de stocker les données dans un fichier CSV nommé "BookAuteur.csv".



Votre job Talend devrait ressembler à quelque chose comme cela :

tMySQLInput ---(main)--> tMap ---(main)--> tFileOutputDelimited | ---(lookup)--> (Lookup sur la table auteur)

1. Creation de la base de données :

Pour créer une nouvelle base de données en utilisant le fichier SQL fourni, vous pouvez suivre les étapes suivantes :

- Modifier le fichier docker compose
- Exécution du script SQL : Copiez le contenu du fichier SQL dans un fichier nommé, par exemple, schema.sql.
- Utilisation de Docker pour MySQL : Utilisez Docker pour créer un conteneur MySQL dans lequel vous exécuterez le script SQL pour créer la nouvelle base de données et ses tables.

La nouvelle base de données à créer est décrite dans le fichier SQL fourni.

```
-- Create database
CREATE DATABASE IF NOT EXISTS `bookdb` DEFAULT CHARACTER SET utf8;
USE `bookdb`;

-- Table `auteur`
DROP TABLE IF EXISTS `auteur`;
CREATE TABLE `auteur` (
  `NUMERO_A` int(10) unsigned NOT NULL AUTO_INCREMENT,
  `NOM` varchar(450) DEFAULT NULL,
  `PRENOM` varchar(450) DEFAULT NULL,
  `DOMICILE` varchar(450) DEFAULT NULL,
  PRIMARY KEY (`NUMERO_A`)
) ENGINE=InnoDB AUTO_INCREMENT=1 DEFAULT CHARSET=latin1;

-- Insert sample authors
INSERT INTO `auteur` (`NOM`, `PRENOM`, `DOMICILE`) VALUES
('Smith', 'John', 'New York'),
('Johnson', 'Emily', 'San Francisco'),
('Garcia', 'Miguel', 'Madrid'),
('Brown', 'Jessica', 'London'),
('Chen', 'Wei', 'Beijing'),
('Dubois', 'Marie', 'Paris'),
('Kim', 'Seong', 'Seoul'),
('Müller', 'Hans', 'Berlin'),
```

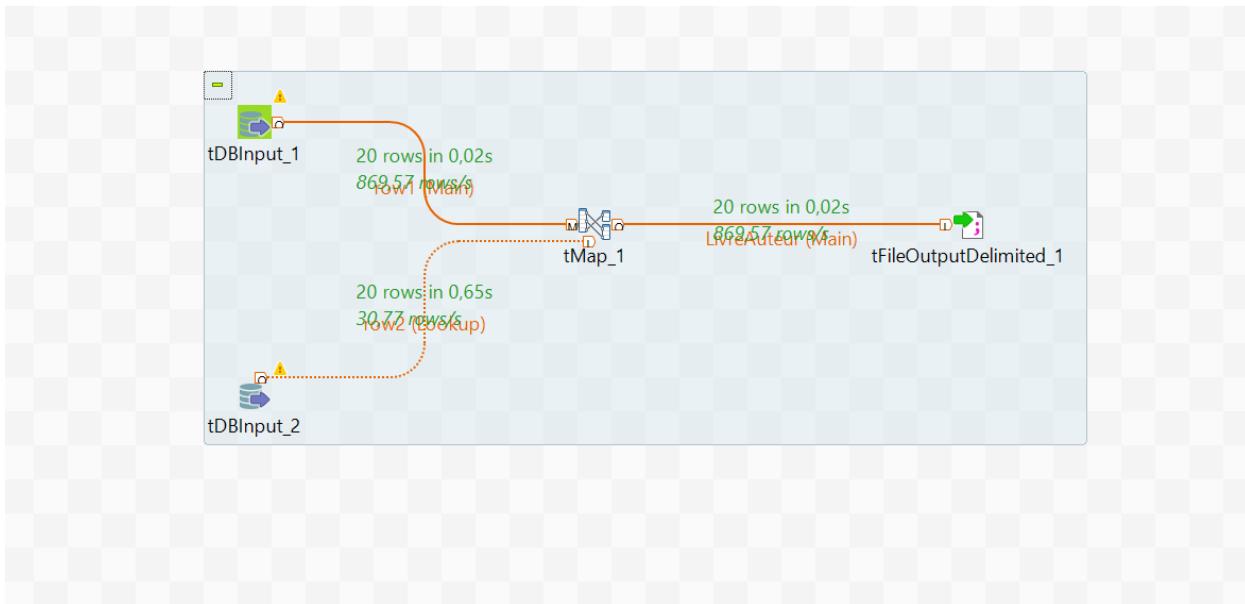
```
('Rossi', 'Marco', 'Rome'),
('Sato', 'Takashi', 'Tokyo');

-- Table `livre`
DROP TABLE IF EXISTS `livre`;
CREATE TABLE `livre` (
  `NUMERO_L` int(10) unsigned NOT NULL AUTO_INCREMENT,
  `NUMERO_A` int(10) unsigned NOT NULL,
  `TITRE` varchar(128) DEFAULT NULL,
  `PRIX` float DEFAULT NULL,
  PRIMARY KEY (`NUMERO_L`),
  CONSTRAINT `fk_auteur_numero` FOREIGN KEY (`NUMERO_A`) REFERENCES `auteur`(`NUMERO_A`)
) ENGINE=InnoDB AUTO_INCREMENT=1 DEFAULT CHARSET=latin1;

-- Insert sample books
INSERT INTO `livre`(`NUMERO_A`, `TITRE`, `PRIX`) VALUES
(1, 'Data Engineering for Dummies', 19.99),
(2, 'Practical Data Engineering', 29.99),
(3, 'Mastering Data Engineering', 34.99),
(4, 'Advanced Techniques in Data Engineering', 39.99),
(5, 'Big Data Infrastructure', 49.99),
(6, 'Data Pipelines and ETL Processing', 24.99),
(7, 'Data Science Essentials', 29.99),
(8, 'Machine Learning in Practice', 39.99),
(9, 'Cloud Data Solutions', 44.99),
(10, 'Data Governance and Compliance', 34.99),
(1, 'Data Warehousing Fundamentals', 29.99),
(2, 'Real-time Analytics', 49.99),
(3, 'Data Quality Management', 39.99),
(4, 'Data Mining Techniques', 34.99),
(5, 'Introduction to Hadoop', 29.99),
(6, 'Spark in Action', 49.99),
(7, 'Python for Data Analysis', 39.99),
(8, 'R Programming for Data Science', 34.99),
(9, 'SQL Mastery', 29.99),
(10, 'NoSQL Databases', 49.99);
```

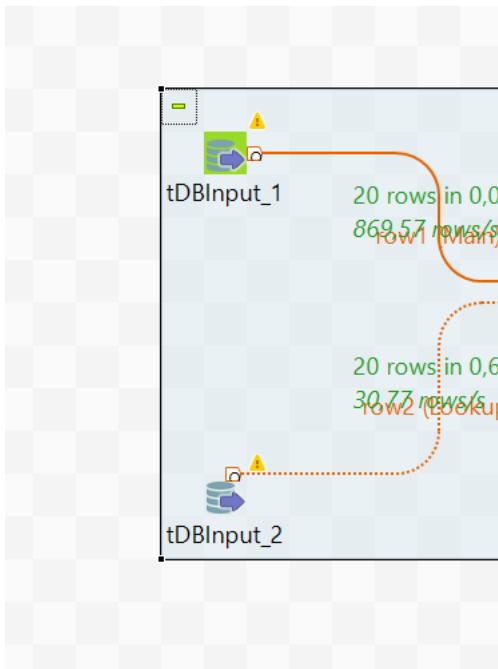
2. Creation de l'ETL :

Pour générer un fichier CSV avec les données fournies, vous pouvez utiliser Talend Data Integration. Voici comment vous pouvez le faire :



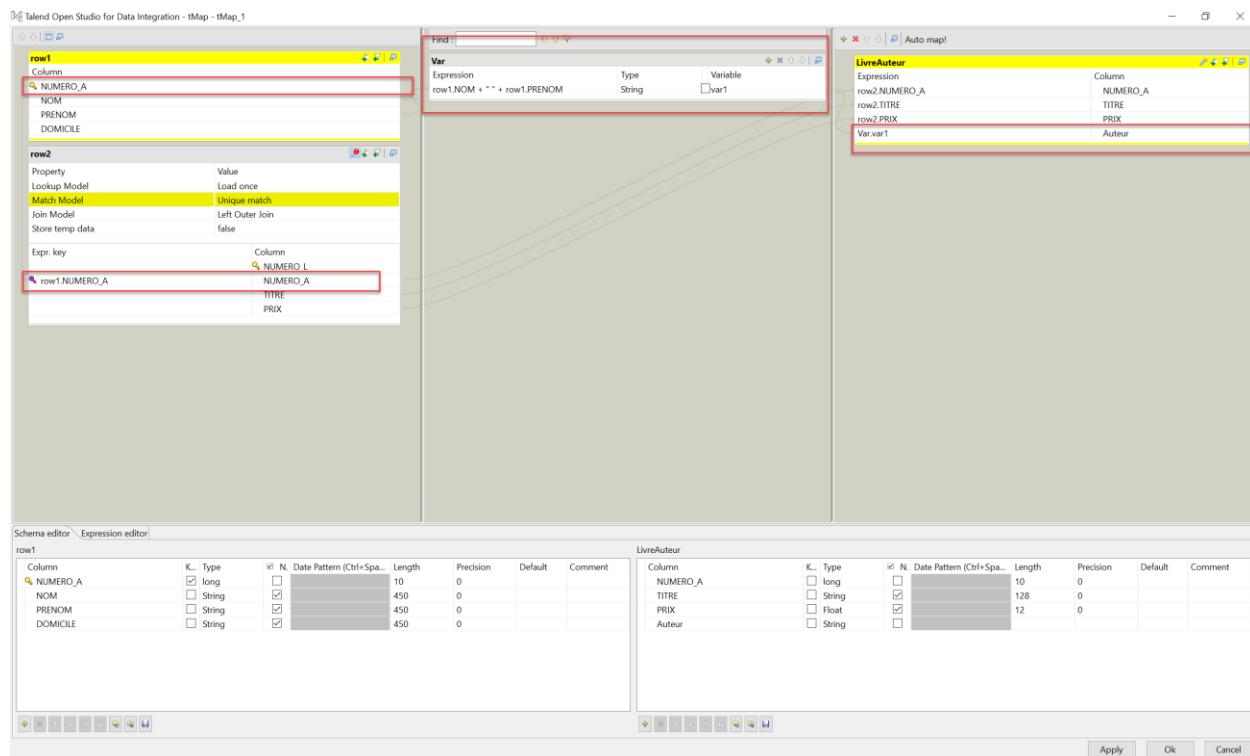
3. Extraction des données :

- Utilisez le composant **tMySQLInput** pour extraire les données des tables **livre** et **auteur** de la base de données **bookdb**.



4. Transformation des données :

- Dans le composant **tMap**, créez une colonne supplémentaire contenant le nom complet de l'auteur en concaténant les colonnes **NOM** et **PRENOM**.
- Réorganisez les colonnes selon le format requis pour le fichier CSV : **NUMERO_L, TITRE, PRIX, Auteur**.



Dans le composant **tMap**, vous pouvez utiliser une expression Java pour concaténer les colonnes **NOM** et **PRENOM**. Par exemple :

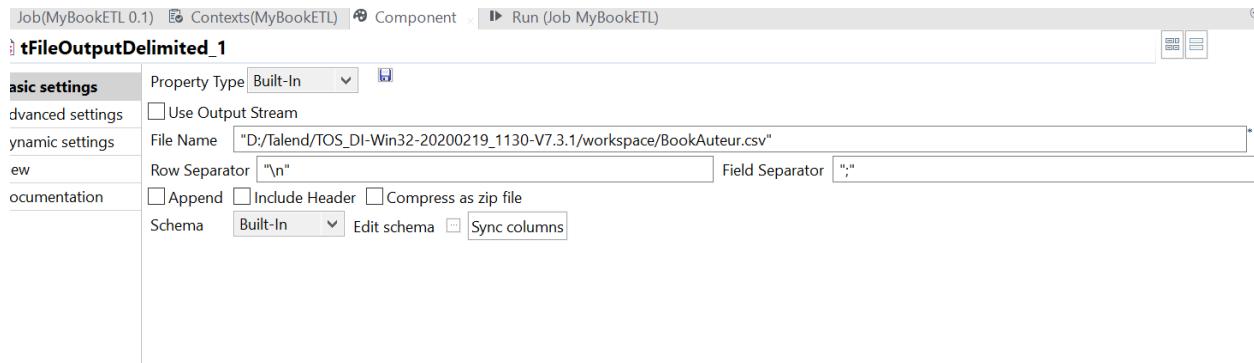
```
row1.NOM + " " + row1.PRENOM
```

Assurez-vous de mapper les colonnes appropriées de la table **livre** et de la table **auteur** dans le composant **tMap**.

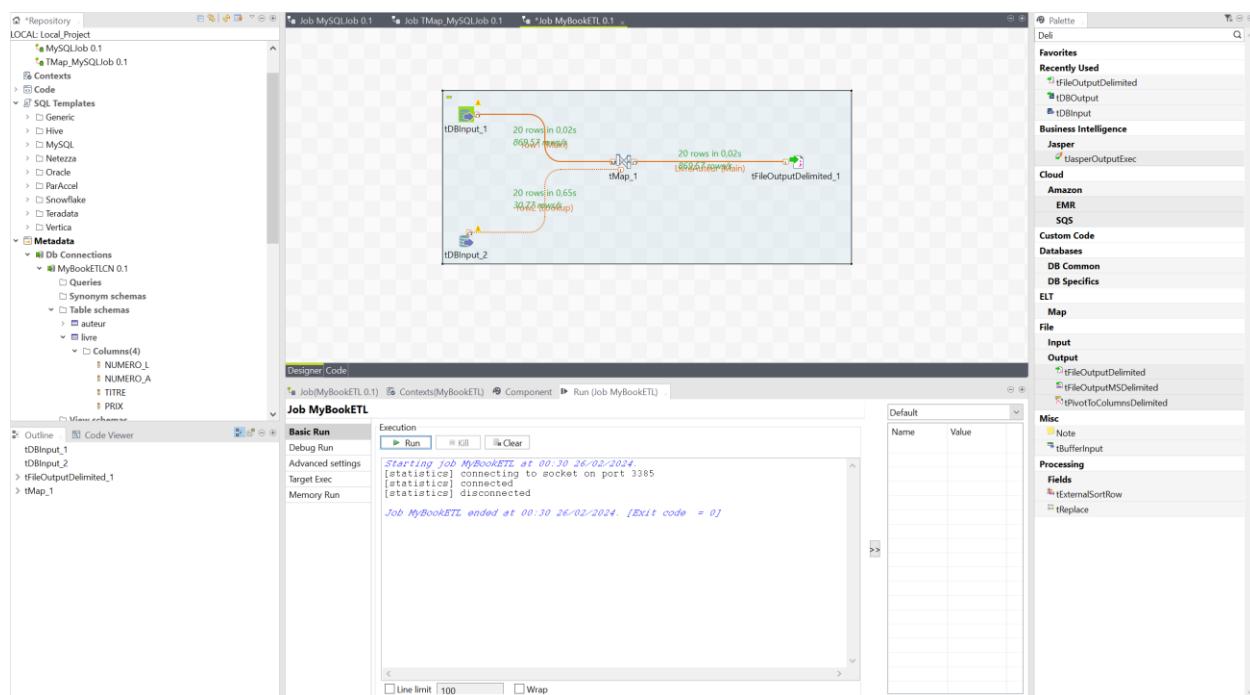
5. Chargement des données :

- Utilisez un composant **tFileOutputDelimited** pour écrire les données transformées dans un fichier CSV.
- Configurez le composant pour spécifier le chemin d'accès et le nom de fichier pour le fichier CSV de sortie.

Enfin, dans le composant **tFileOutputDelimited**, configurez le chemin d'accès et le nom de fichier pour votre fichier CSV de sortie, ainsi que les délimiteurs appropriés pour séparer les champs.



Exécuter le job et vérifier par la suite la bonne exécution :



En suivant ces étapes, vous pourrez générer un fichier CSV avec les données fournies,



Talend for Data Integration  **ESSIN**
ÉCOLE SUPÉRIEURE DES SCIENCES
ET DE L'INGÉIERIE NUMÉRIQUE.

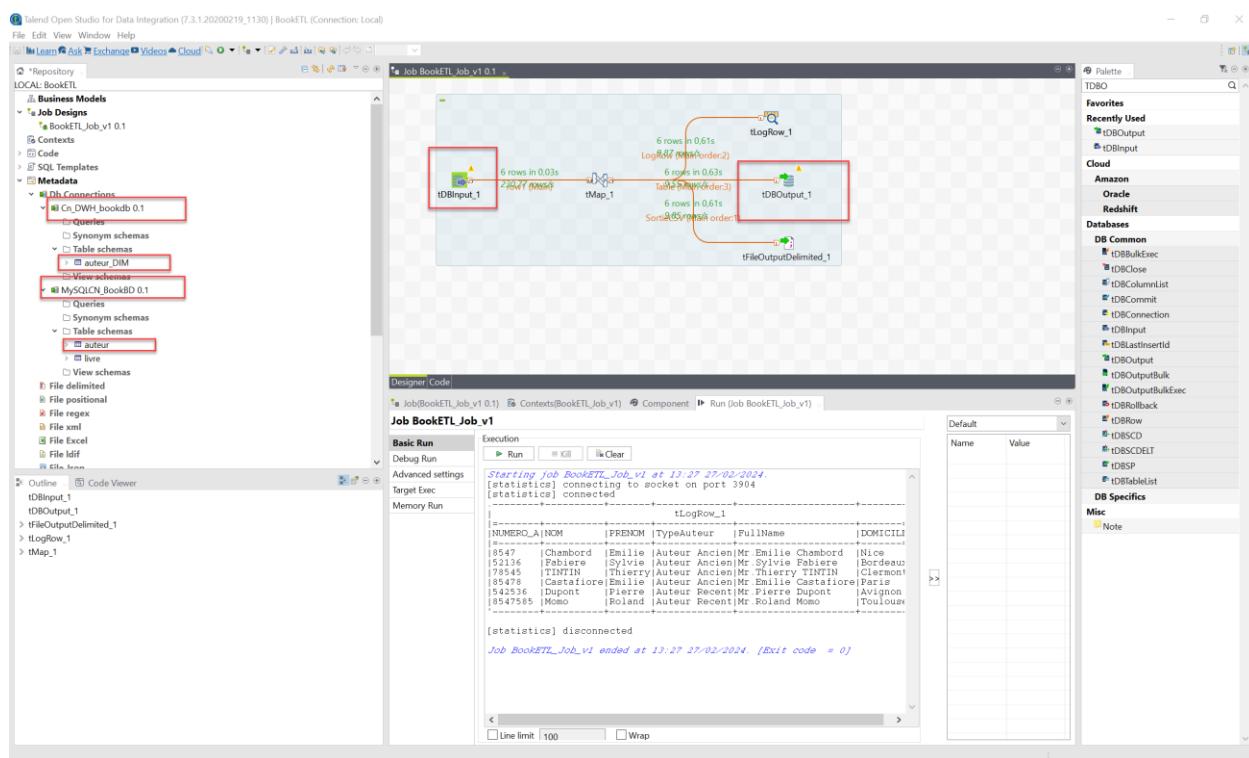
Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com



Exercice 2 :

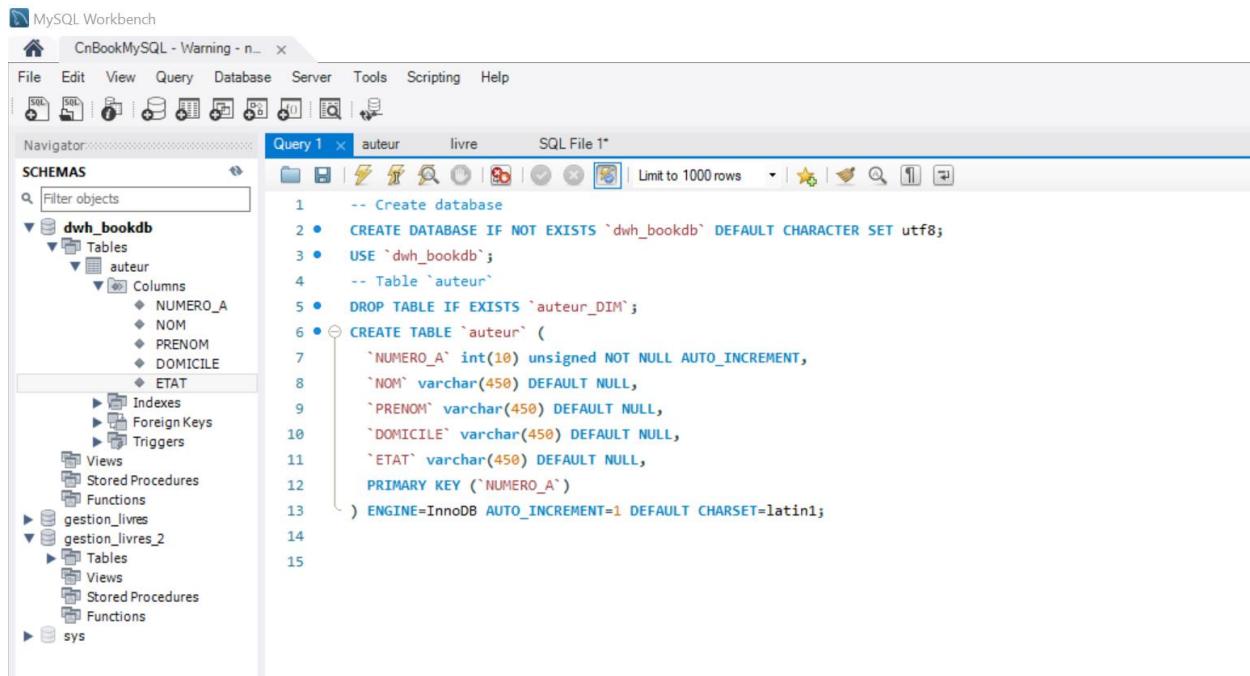
Chargement des données dans une bases de données MSYQL

Dans cet exemple on duplique le job et on ajoute un autre composant pour insérer des données dans la base de données cible



6. Préparation de la bases cibles :

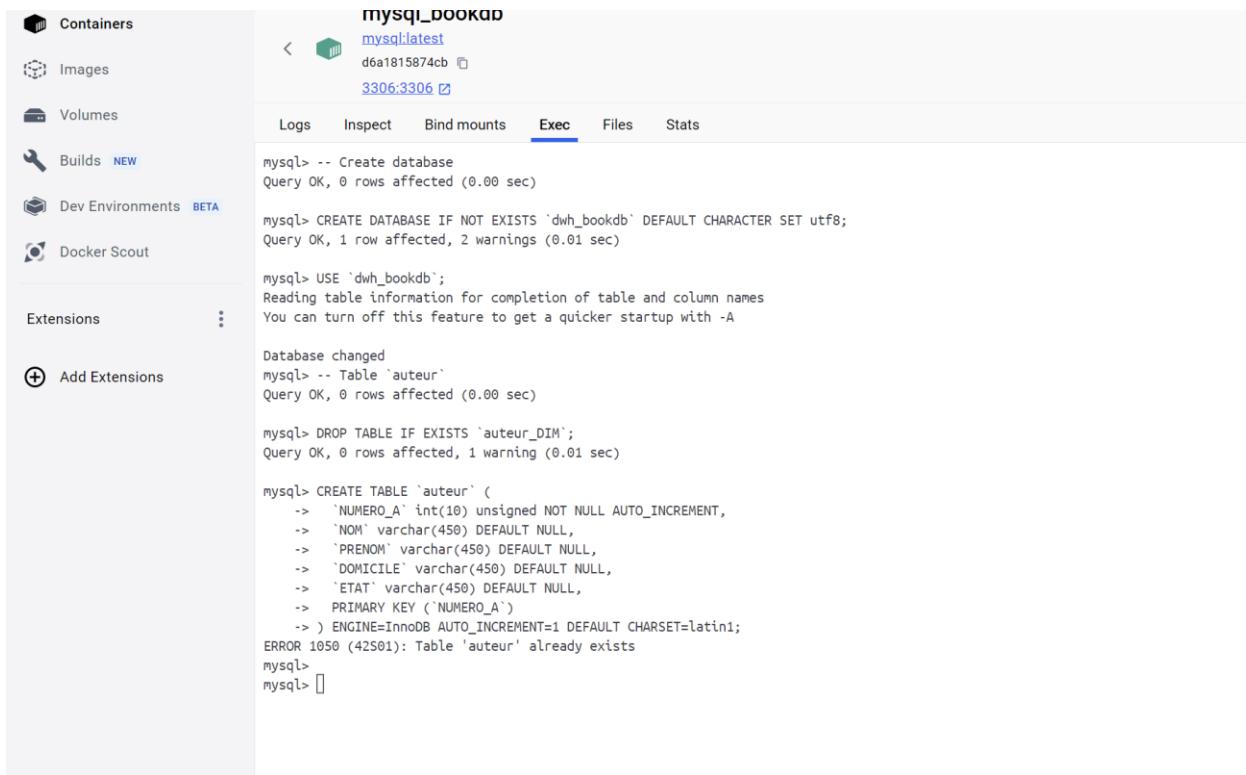
Créer la base de données Target « dwh_books » avec le script suivant en utilisant workbench ou Docker Desktop



The screenshot shows the MySQL Workbench interface. On the left, the Navigator pane displays the schema structure of the 'dwh_bookdb' database, including tables like 'auteur' and 'livre'. The main area, titled 'Query 1', contains the following SQL code:

```
1 -- Create database
2 • CREATE DATABASE IF NOT EXISTS `dwh_bookdb` DEFAULT CHARACTER SET utf8;
3 • USE `dwh_bookdb`;
4 -- Table `auteur`
5 • DROP TABLE IF EXISTS `auteur_DIM`;
6 • CREATE TABLE `auteur` (
7     `NUMERO_A` int(10) unsigned NOT NULL AUTO_INCREMENT,
8     `NOM` varchar(450) DEFAULT NULL,
9     `PRENOM` varchar(450) DEFAULT NULL,
10    `DOMICILE` varchar(450) DEFAULT NULL,
11    `ETAT` varchar(450) DEFAULT NULL,
12    PRIMARY KEY (`NUMERO_A`)
13 ) ENGINE=InnoDB AUTO_INCREMENT=1 DEFAULT CHARSET=latin1;
14
15
```

Avec Docker



```

mysql> -- Create database
Query OK, 0 rows affected (0.00 sec)

mysql> CREATE DATABASE IF NOT EXISTS `dwh_bookdb` DEFAULT CHARACTER SET utf8;
Query OK, 1 row affected, 2 warnings (0.01 sec)

mysql> USE `dwh_bookdb`;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql> -- Table `auteur`
Query OK, 0 rows affected (0.00 sec)

mysql> DROP TABLE IF EXISTS `auteur_DIM`;
Query OK, 0 rows affected, 1 warning (0.01 sec)

mysql> CREATE TABLE `auteur` (
    >     `NUMERO_A` int(10) unsigned NOT NULL AUTO_INCREMENT,
    >     `NOM` varchar(450) DEFAULT NULL,
    >     `PRENOM` varchar(450) DEFAULT NULL,
    >     `DOMICILE` varchar(450) DEFAULT NULL,
    >     `ETAT` varchar(450) DEFAULT NULL,
    >     PRIMARY KEY (`NUMERO_A`)
    > ) ENGINE=InnoDB AUTO_INCREMENT=1 DEFAULT CHARSET=latin1;
ERROR 1050 (42S01): Table 'auteur' already exists
mysql>
mysql> []

```

```

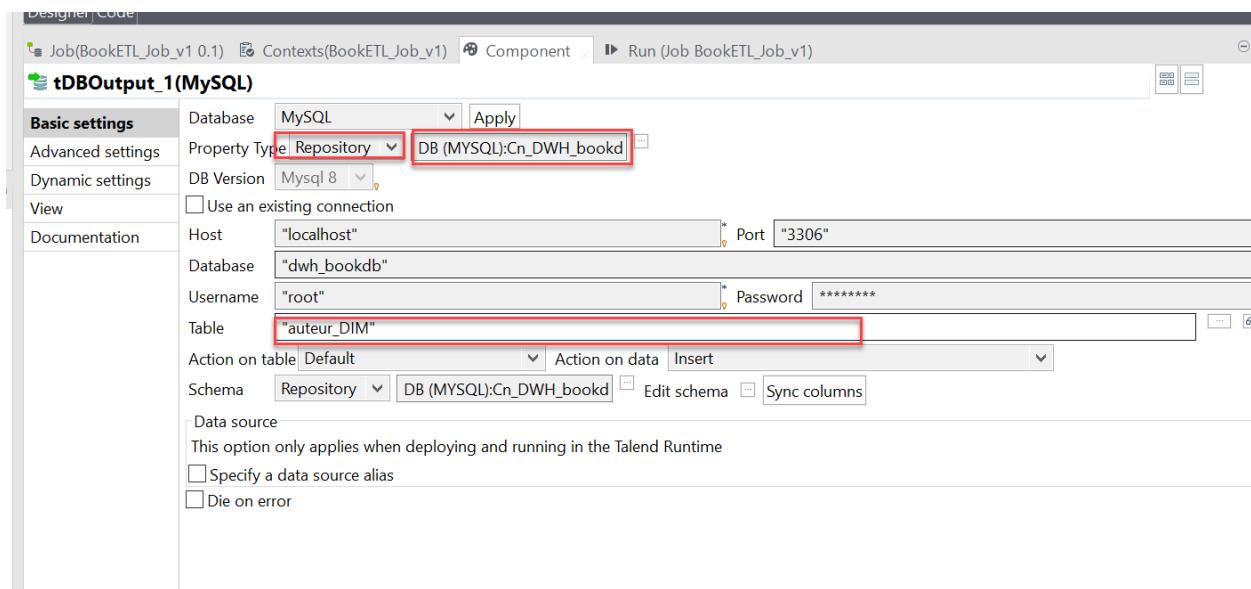
-- Create database
CREATE DATABASE IF NOT EXISTS `dwh_bookdb` DEFAULT CHARACTER SET utf8;
USE `dwh_bookdb`;
-- Table `auteur`
DROP TABLE IF EXISTS `auteur_DIM`;
CREATE TABLE `auteur_DIM` (
    `NUMERO_A` int(10) unsigned NOT NULL AUTO_INCREMENT,
    `NOM` varchar(450) DEFAULT NULL,
    `PRENOM` varchar(450) DEFAULT NULL,
    `DOMICILE` varchar(450) DEFAULT NULL,
    `ETAT` varchar(450) DEFAULT NULL,
    PRIMARY KEY (`NUMERO_A`)
) ENGINE=InnoDB AUTO_INCREMENT=1 DEFAULT CHARSET=latin1;

```

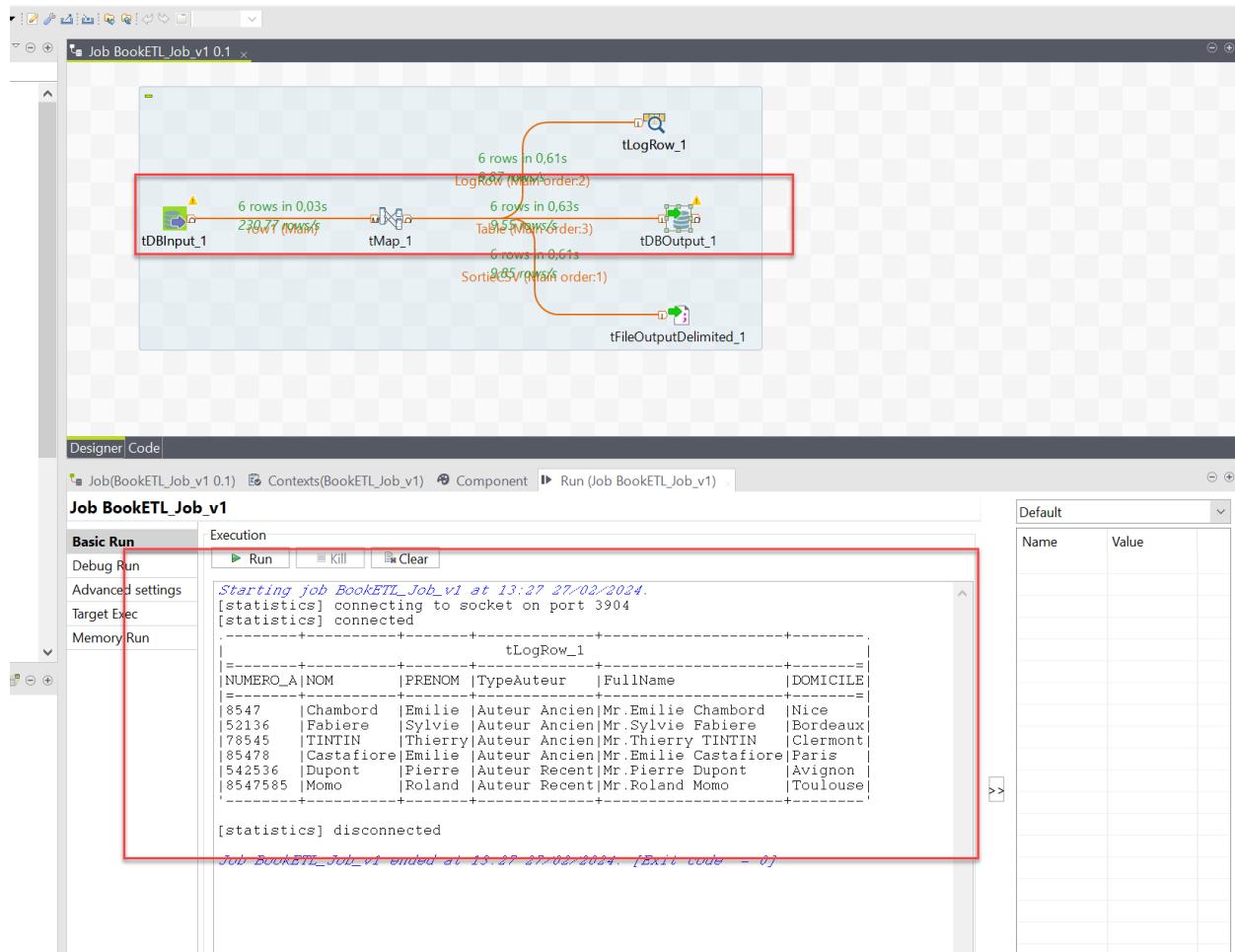
1. Chargement des données dans une table “Auteur_DIM”:

- Utilisez un composant **tDBOutput** pour écrire les données transformées dans une table mysql « auteur_dim ».

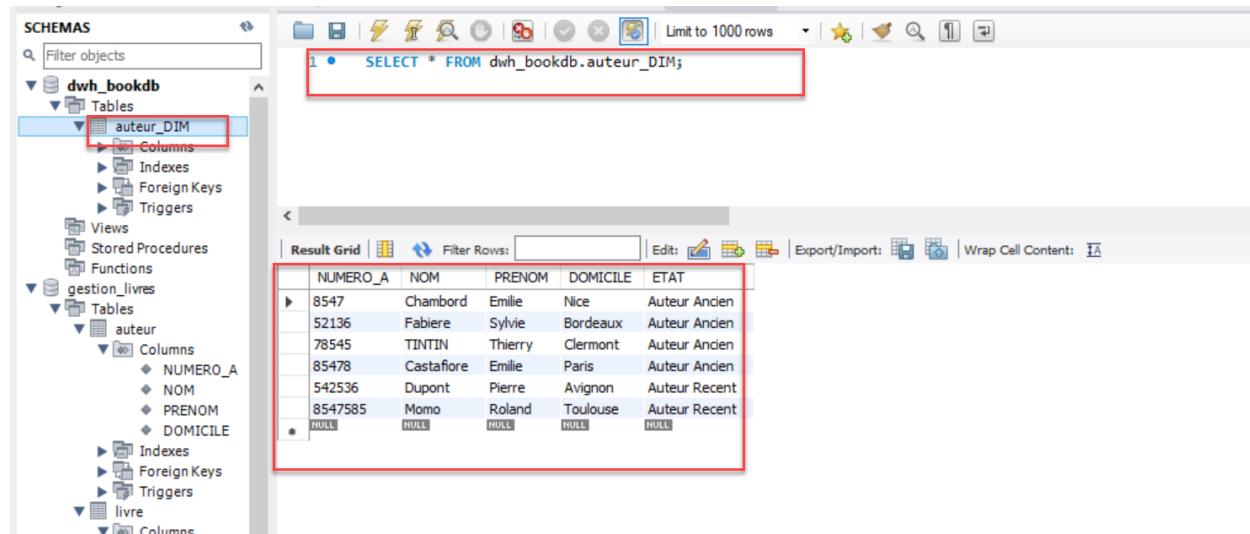
Enfin, dans le composant **tDBOutput**, configurez les paramètres de l'accès à la bases de données et de sortie, ainsi que le nom de la table « Auteur_DIM » appropriée.



Exécuter le job et vérifier par la suite la bonne exécution ;

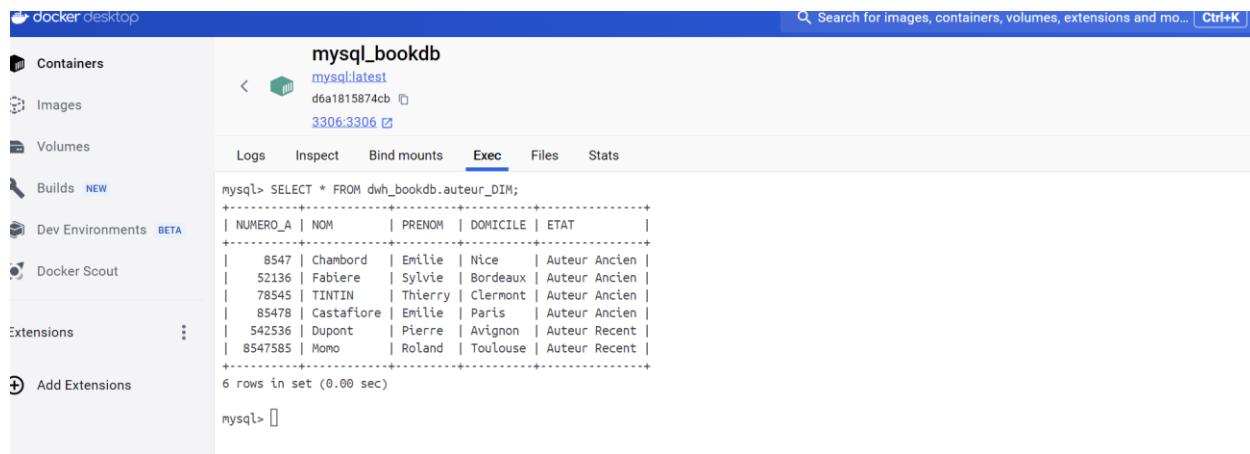


En suivant ces étapes, vous pourrez charger dans la table « Auteur_DIM » avec les données fournies, avec Workbench ou Docker Desktop « Exec »



The screenshot shows the Talend Studio interface. On the left, the 'SCHEMAS' tree view displays the 'dwh_bookdb' schema, with the 'auteur_DIM' table selected and highlighted with a red box. On the right, the 'Result Grid' window shows the output of the query 'SELECT * FROM dwh_bookdb.auteur_DIM;'. The result grid also has a red box around it, highlighting the data rows. The data in the grid is as follows:

	NUMERO_A	NOM	PRENOM	DOMICILE	ETAT
▶	8547	Chambord	Emilie	Nice	Auteur Ancien
	52136	Fabiere	Sylvie	Bordeaux	Auteur Ancien
	78545	TINTIN	Thierry	Clermont	Auteur Ancien
	85478	Castafiore	Emilie	Paris	Auteur Ancien
	542536	Dupont	Pierre	Avignon	Auteur Recent
	8547585	Momo	Roland	Toulouse	Auteur Recent
●	NULL	NULL	NULL	NULL	NULL



The screenshot shows the Docker Desktop interface. A MySQL container named 'mysql_bookdb' is running, as indicated by the status bar. The 'Logs' tab is selected, showing the command 'SELECT * FROM dwh_bookdb.auteur_DIM;' and its execution results, which match the data shown in the Talend screenshot above.

```

mysql> SELECT * FROM dwh_bookdb.auteur_DIM;
+-----+-----+-----+-----+-----+
| NUMERO_A | NOM | PRENOM | DOMICILE | ETAT |
+-----+-----+-----+-----+-----+
| 8547 | Chambord | Emilie | Nice | Auteur Ancien |
| 52136 | Fabiere | Sylvie | Bordeaux | Auteur Ancien |
| 78545 | TINTIN | Thierry | Clermont | Auteur Ancien |
| 85478 | Castafiore | Emilie | Paris | Auteur Ancien |
| 542536 | Dupont | Pierre | Avignon | Auteur Recent |
| 8547585 | Momo | Roland | Toulouse | Auteur Recent |
+-----+-----+-----+-----+-----+
6 rows in set (0.00 sec)

mysql>

```

1. Préparation des tables cibles dans dwh_bookdb

Assurez-vous que les tables suivantes existent dans votre data warehouse cible :

- **Abonne_DIM** : Pour stocker les informations des abonnés.
- **Emprunt_FACT** : Pour les emprunts effectués par les abonnés.
- **Dim_Calendar** : Pour stocker des informations de date (si elle n'est pas déjà peuplée, voir l'exemple précédent pour la création et le peuplement).

Voici les scripts pour créer les tables si ce n'est pas encore fait :

Table abonne_DIM

```
USE `dwh_bookdb`;  
DROP TABLE IF EXISTS `Abonne_DIM`;  
CREATE TABLE `Abonne_DIM` (  
  `AbonneKey` INT AUTO_INCREMENT PRIMARY KEY,  
  `NUMERO_AB` INT NOT NULL,  
  `NOM` VARCHAR(100),  
  `PRENOM` VARCHAR(100),  
  `ADRESSE` VARCHAR(255)  
);
```

Table de faits emprunt_FACT

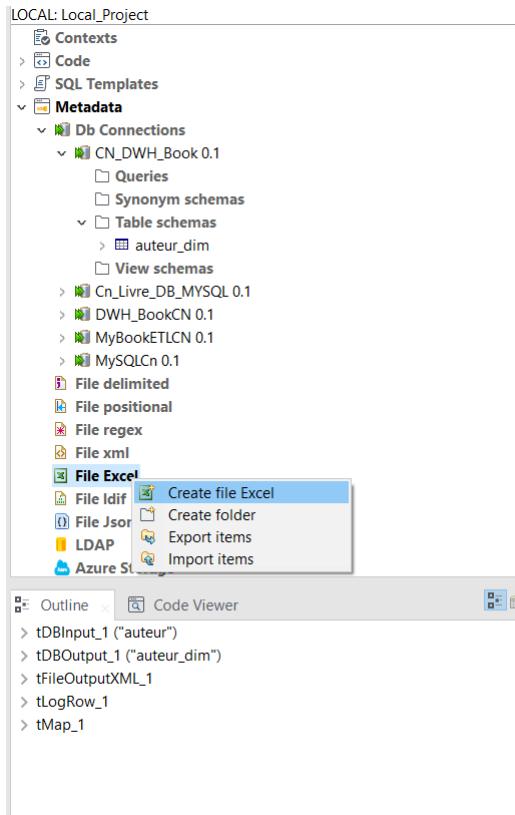
```
-- Table Emprunt_FACT  
DROP TABLE IF EXISTS `Emprunt_FACT`;  
CREATE TABLE `Emprunt_FACT` (  
  `EmpruntKey` INT AUTO_INCREMENT PRIMARY KEY,  
  `AbonneKey` INT,  
  `LivreKey` INT,  
  `DateEmpruntKey` INT,  
  `DateRetourKey` INT,  
  FOREIGN KEY (`AbonneKey`) REFERENCES `Abonne_DIM`(`AbonneKey`),  
  FOREIGN KEY (`DateEmpruntKey`) REFERENCES `Dim_Calendar`(`DateKey`),  
  FOREIGN KEY (`DateRetourKey`) REFERENCES `Dim_Calendar`(`DateKey`)  
);
```

Étape 2 : Créer une connexion pour le fichier Excel dans Talend**1. Créer des Métadonnées pour les Fichiers Excel**

Pour le fichier abonnes.xlsx

1. Accédez au Répertoire des Métadonnées :

- Dans Talend Studio, dans le panneau Repository, développez la section Metadata et faites un clic droit sur File Excel.
- Sélectionnez Create file Excel.



2. Créer une Connexion pour le Fichier Excel :

- Donnez un nom à la connexion, comme Abonnes_Source_Excel.
- Sélectionnez le fichier Books.xlsx comme chemin de fichier.
- Cliquez sur Next.

3. Configurer les Propriétés de la Feuille :

- Dans l'onglet suivant, sélectionnez la feuille contenant les données d'abonnés (par exemple, Abonnes).

 New Excel File

File - Step 2 of 4

Add a Metadata File on repository
Define the path of the file and the format settings

File Settings

Server: Localhost 127.0.0.1

File: D:/Formation/Wevops/ABData-Fitec/Talend Big Data/Cours/MySQL/Books TP/Book1.xlsx

Read excel2007 file format(xlsx)

Generation mode: Memory-consuming(User mode)

File Viewer and Sheets setting

Set sheets parameters

Please select sheet (Sheet structure as schema guide) **Abonnees**

A	B	C	D
NUM...	NOM	PREN...	ADR...
1.0	Martin	Alice	15 Ru...
2.0	Dupo...	Jean	32 Av...
3.0	Lefev...	Claire	54 B...
4.0	Bern...	Pierre	78 Ru...
5.0	Petit	Emma	12 Ru...
6.0	More...	Paul	45 Ru...
7.0	Garni...	Marie	23 Av...

< Back Finish Cancel

4.

- o Spécifiez la ligne d'en-tête si votre fichier en contient une (par exemple, la première ligne).

New Excel File

File - Step 3 of 4

Add a Metadata File on repository
Define the setting of the parse job



File Settings

Encoding	UTF-8
<input type="checkbox"/> Advanced separator(for number)	
Thousands separator:	,
Decimal separator:	.

Rows To Skip

If any rows must be ignored, specify the following parameters

<input checked="" type="checkbox"/> Header	1
<input type="checkbox"/> Footer	

Metadata column setting

First column:	1
Last column:	

Limit Of Rows

If the number of lines must be limited, specify this number

<input type="checkbox"/> Limit	
--------------------------------	--

Preview **Output**

Set heading row as column names 

NUMERO_AB	NOM	PRENOM	ADRESSE
1	Martin	Alice	15 Rue de Paris
2	Dupont	Jean	32 Avenue de la Liberté
3	Lefevre	Claire	54 Boulevard Saint-Michel
4	Bernard	Pierre	78 Rue des Lilas
5	Petit	Emma	12 Rue du Commerce

Export as context **Revert Context**

< Back **Next >** **Finish** **Cancel**

New Excel File

File - Step 4 of 4

Add a Schema on repository
 Define the Schema

Name: Abonnees

Comment:

Schema

Click to update schema preview | Guess

Description of the Schema

Column	Key	Type	N.	Date Pattern (Ctrl+Space ...)	Length	Precision	Default	Comment
NUMERO_AB	<input checked="" type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		2	0		
NOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		9	0		
PRENOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		6	0		
ADRESSE	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		25	0		

< >

Finish Cancel

- Cliquez sur Finish pour enregistrer la connexion.
- > DWH_BookCN 0.1
 - > MyBookETLCN 0.1
 - > MySQLCn 0.1
 - File delimited
 - File positional
 - File regex
 - File xml
 - File Excel
 - Abonnees_Excel 0.1
 - Abonnees
 - Columns(4)
 - ADRESSE
 - NOM
 - NUMERO_AB
 - PRENOM
 - File Idif



Talend for Data Integration



Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com

3. Créer les Métadonnées pour les « Emprunts »

1. Accédez au Répertoire des Métadonnées :

- Dans le panneau Repository (à gauche de l'interface), trouvez la section Metadata.

2. Créez une Connexion pour le Fichier Excel :

- Faites un clic droit sur File Excel dans le répertoire Metadata.
- Sélectionnez Create file Excel.

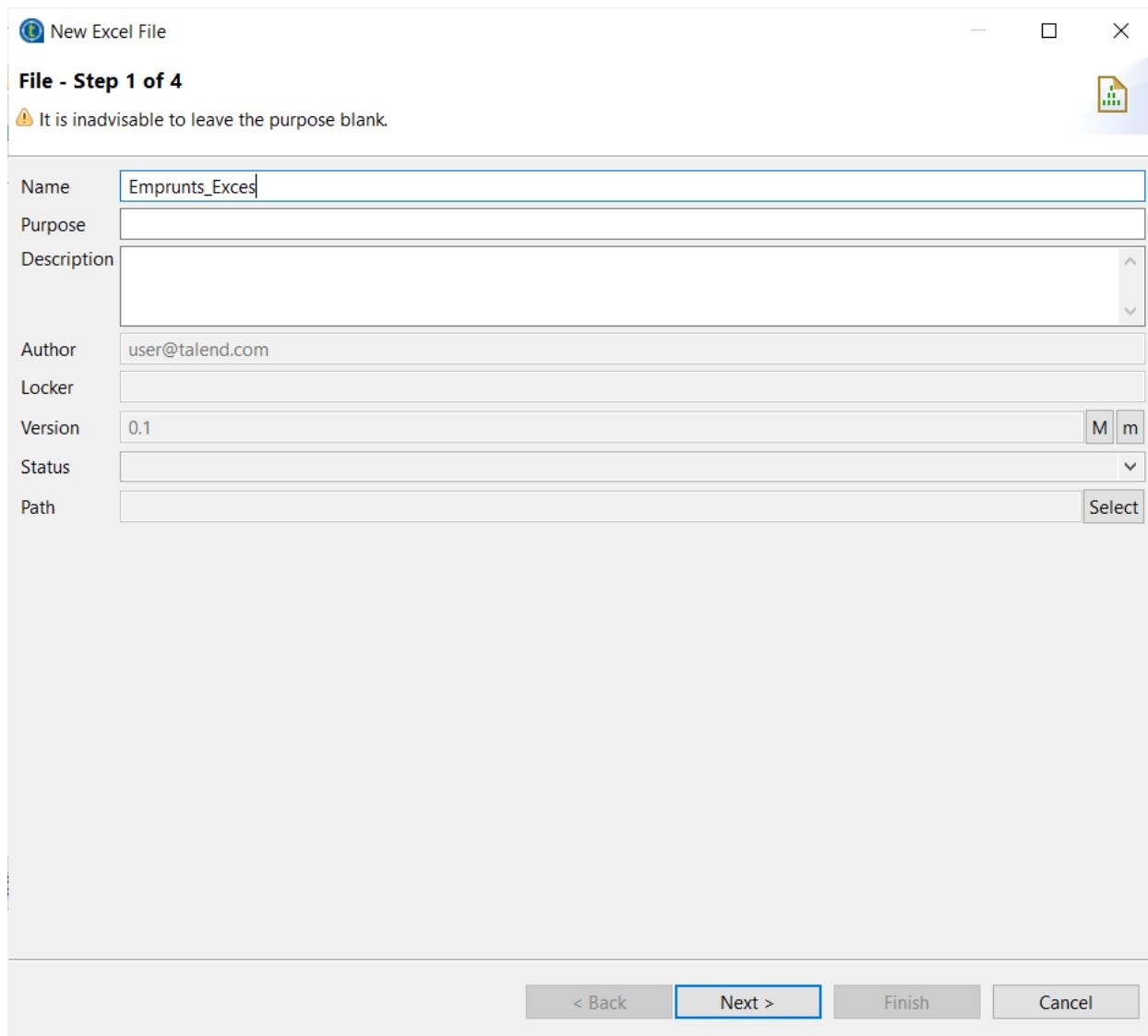
New Excel File

File - Step 1 of 4

⚠ It is inadvisable to leave the purpose blank.

Name	Emprunts_Exces
Purpose	
Description	
Author	user@talend.com
Locker	
Version	0.1
Status	
Path	Select

< Back Next > Finish Cancel

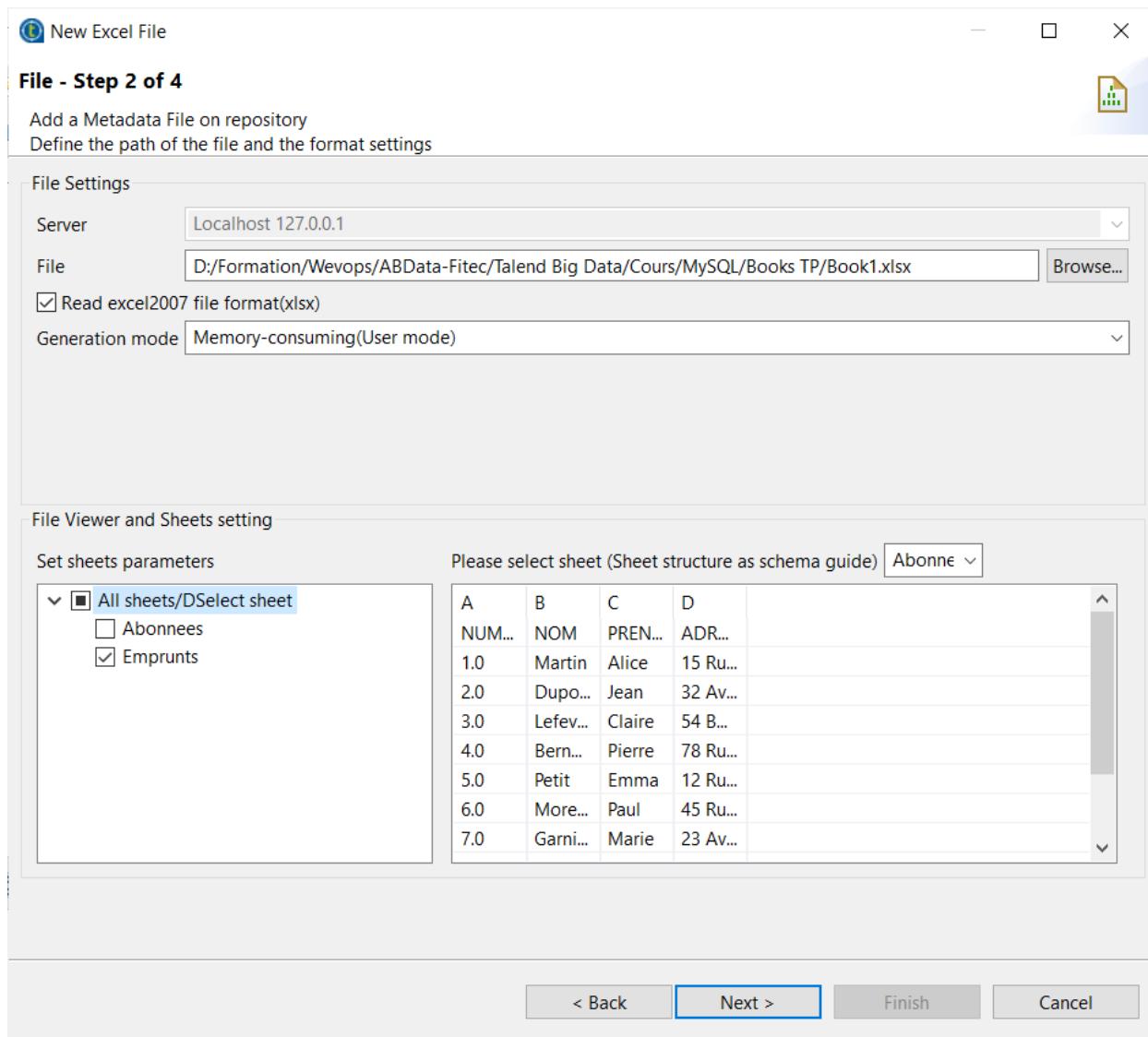


3. Configurer la Connexion au Fichier :

- Dans la fenêtre qui s'ouvre, donnez un nom à votre connexion, par exemple Emprunts_Source_Excel.
- Cliquez sur le bouton Browse pour sélectionner le fichier emprunts.xlsx sur votre disque.
- Cliquez sur Next.

4. Configurer les Propriétés de la Feuille :

- Sélectionnez la feuille qui contient les données d'emprunts (par exemple, Emprunts).



File - Step 2 of 4

Add a Metadata File on repository
Define the path of the file and the format settings

File Settings

Server: Localhost 127.0.0.1

File: D:/Formation/Wevops/ABData-Fitec/Talend Big Data/Cours/MySQL/Books TP/Book1.xlsx

Read excel2007 file format(xlsx)

Generation mode: Memory-consuming(User mode)

File Viewer and Sheets setting

Set sheets parameters

Please select sheet (Sheet structure as schema guide)

A	B	C	D
NUM...	NOM	PREN...	ADR...
1.0	Martin	Alice	15 Ru...
2.0	Dupo...	Jean	32 Av...
3.0	Lefev...	Claire	54 B...
4.0	Bern...	Pierre	78 Ru...
5.0	Petit	Emma	12 Ru...
6.0	More...	Paul	45 Ru...
7.0	Garni...	Marie	23 Av...

< Back Finish Cancel

- Vérifiez que la première ligne contient les en-têtes de colonne, puis cliquez sur Next.

New Excel File

File - Step 3 of 4

Add a Metadata File on repository
Define the setting of the parse job

File Settings

- Encoding: UTF-8
- Advanced separator(for number)
- Thousands separator: ;
- Decimal separator: ,

Rows To Skip

If any rows must be ignored, specify the following parameters

Header 1

Footer

Metadata column setting

First column: 1

Last column:

Limit Of Rows

If the number of lines must be limited, specify this number

Limit

Preview **Output**

Set heading row as column names **Refresh Preview**

NUMERO_E	NUMERO_AB	NUMERO_L	DATE_EMPRUNT
1	1	3	Wed Jan 10 00:00:00 CET 2024
2	2	7	Fri Jan 12 00:00:00 CET 2024
3	3	2	Mon Feb 05 00:00:00 CET 2024
4	4	5	Thu Feb 08 00:00:00 CET 2024
5	5	10	Fri Mar 01 00:00:00 CET 2024

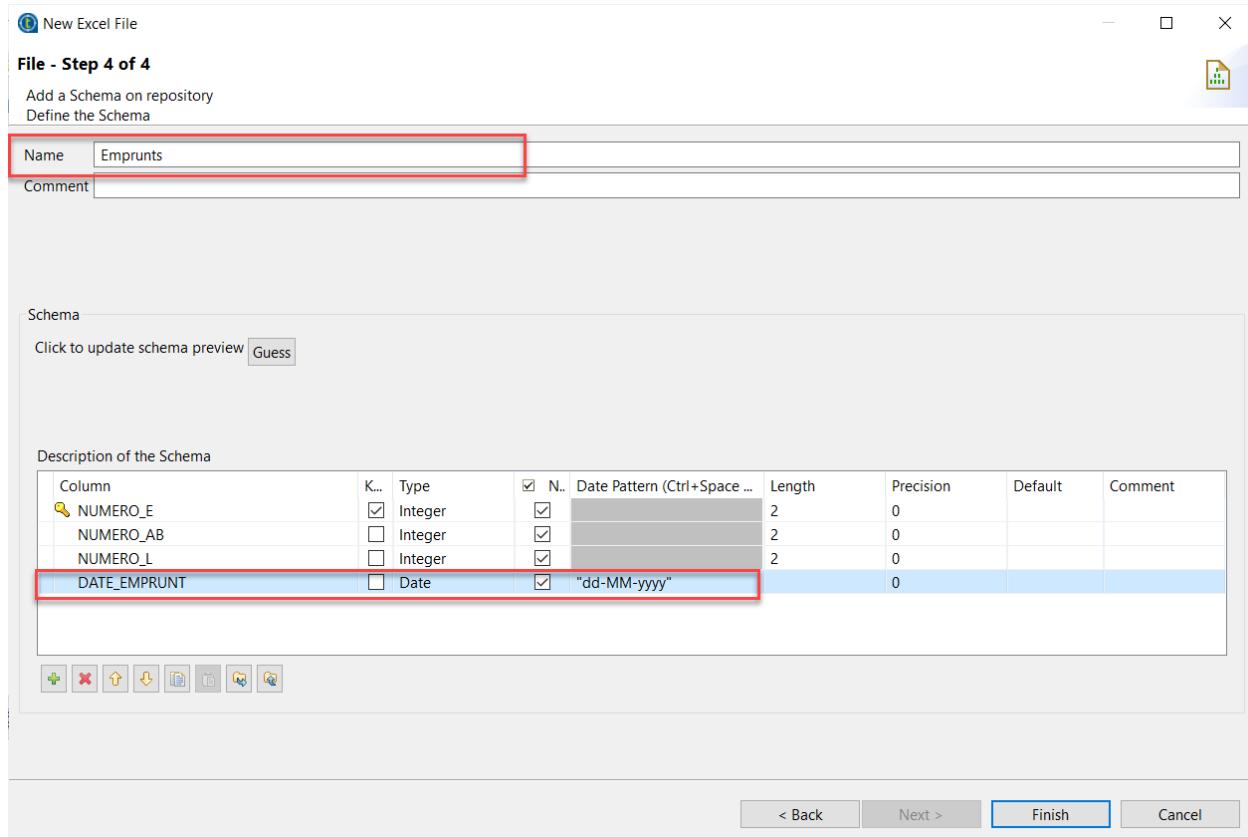
Export as context **Revert Context**

< Back **Next >** Finish Cancel

5. Définir les Colonnes :

- Talend va automatiquement détecter les colonnes basées sur les en-têtes.
- Vérifiez que les types de données détectés sont corrects. Par exemple :
 - **NUMERO_E : Integer**
 - **NUMERO_AB : Integer**
 - **NUMERO_L : Integer**
 - **DATE_EMPRUNT : Date**

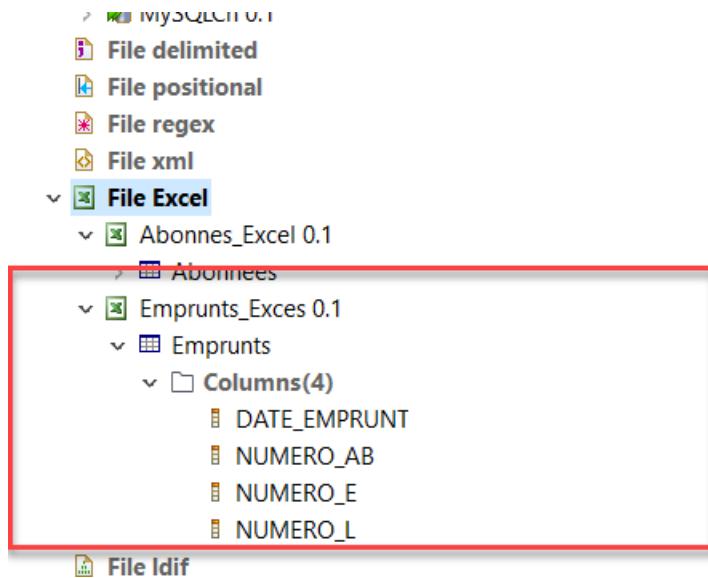
- **Si nécessaire, modifiez les types de données.**
- **Cliquez sur Finish pour enregistrer la connexion.**



The screenshot shows the 'File - Step 4 of 4' dialog for defining a schema. The 'Name' field is set to 'Emprunts'. In the 'Schema' section, there is a table titled 'Description of the Schema' with columns for Column, Type, N., Date Pattern, Length, Precision, Default, and Comment. The 'DATE_EMPRUNT' column is highlighted with a red box, showing it is of type Date with a length of 0 and a date pattern of "dd-MM-yyyy".

Column	Type	N.	Date Pattern (Ctrl+Space ...)	Length	Precision	Default	Comment
NUMERO_E	Integer	<input checked="" type="checkbox"/>		2	0		
NUMERO_AB	Integer	<input type="checkbox"/>	<input checked="" type="checkbox"/>	2	0		
NUMERO_L	Integer	<input type="checkbox"/>	<input checked="" type="checkbox"/>	2	0		
DATE_EMPRUNT	Date	<input checked="" type="checkbox"/>	"dd-MM-yyyy"	0			

Vérifier le schema de Emprunts



Travail à faire : Terminer les Jobs pour Alimenter l'entrepôt de données « DWH_BOOK »

Étape 3 : Créer un Nouveau Job dans Talend

1. Créez un Nouveau Job :

- Dans le panneau Repository, faites un clic droit sur Job Designs et sélectionnez Create Job.
- Nommez le job par exemple ETL_Dim_Abonnes et ajoutez une description si nécessaire.
- Cliquez sur Finish.

2. Créez un Nouveau Job ETL_Dim_Livres

- Dans le panneau Repository, faites un clic droit sur Job Designs et sélectionnez Create Job.
- Nommez le job par exemple ETL_Dim_Livre et ajoutez une description si nécessaire.
- Cliquez sur Finish.

3. Créer un Nouveau Job ETL_FACT_Emprunts

- Dans le panneau Repository, faites un clic droit sur Job Designs et sélectionnez Create Job.
- Nommez le job par exemple ETL_FACT_Emprunts et ajoutez une description si nécessaire.
- Cliquez sur Finish.

Prérequis

- Fichier Excel : Vous avez besoin d'un fichier Excel contenant des dates (par exemple, Emprunts.xlsx) avec une colonne Date Emprunt.
- Base de données SQL Server : Avoir une base de données SQL Server avec une table cible dim_calendar.

Étapes détaillées

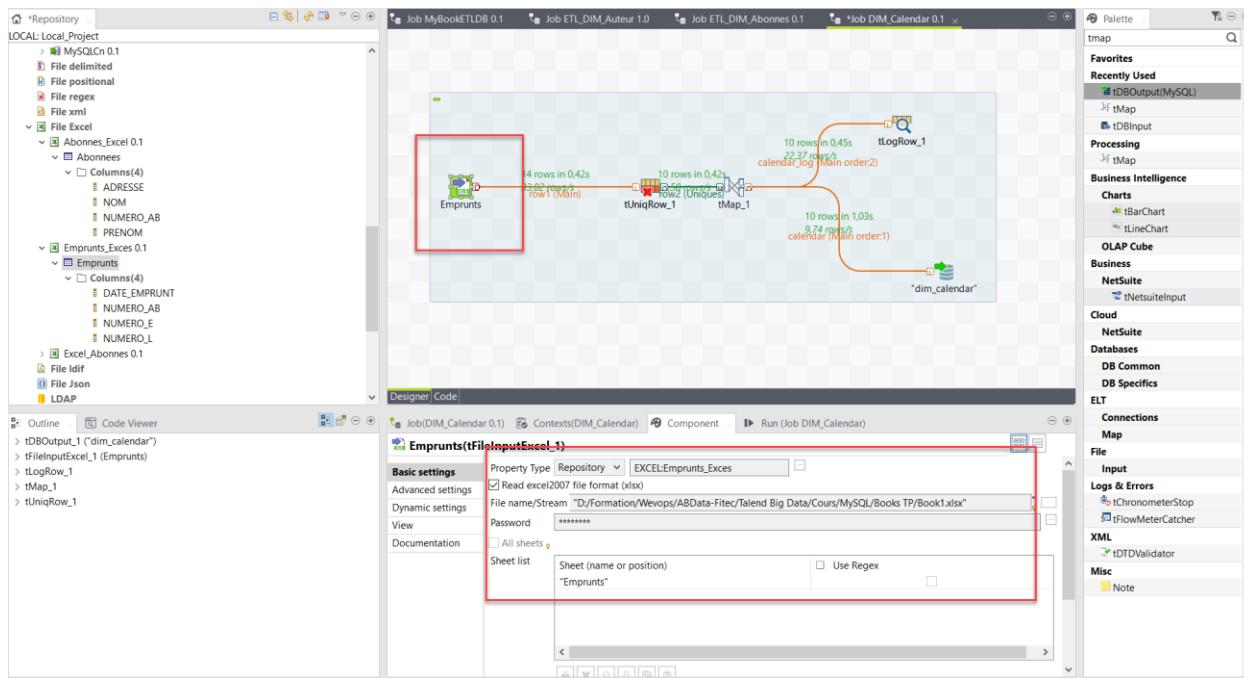
1. Créer le job ETL dans Talend

1. Créer un nouveau Job : Ouvrez Talend Studio, puis allez dans le panneau Repository. Cliquez avec le bouton droit sur Job Designs et sélectionnez Create Job. Donnez un nom à votre job, comme Job_DimCalendar.
-

2. Ajouter et configurer les composants

a) tFileInputExcel - Lecture du fichier Excel

1. Ajouter tFileInputExcel : Depuis le panneau Palette, faites glisser le composant tFileInputExcel dans votre espace de travail.



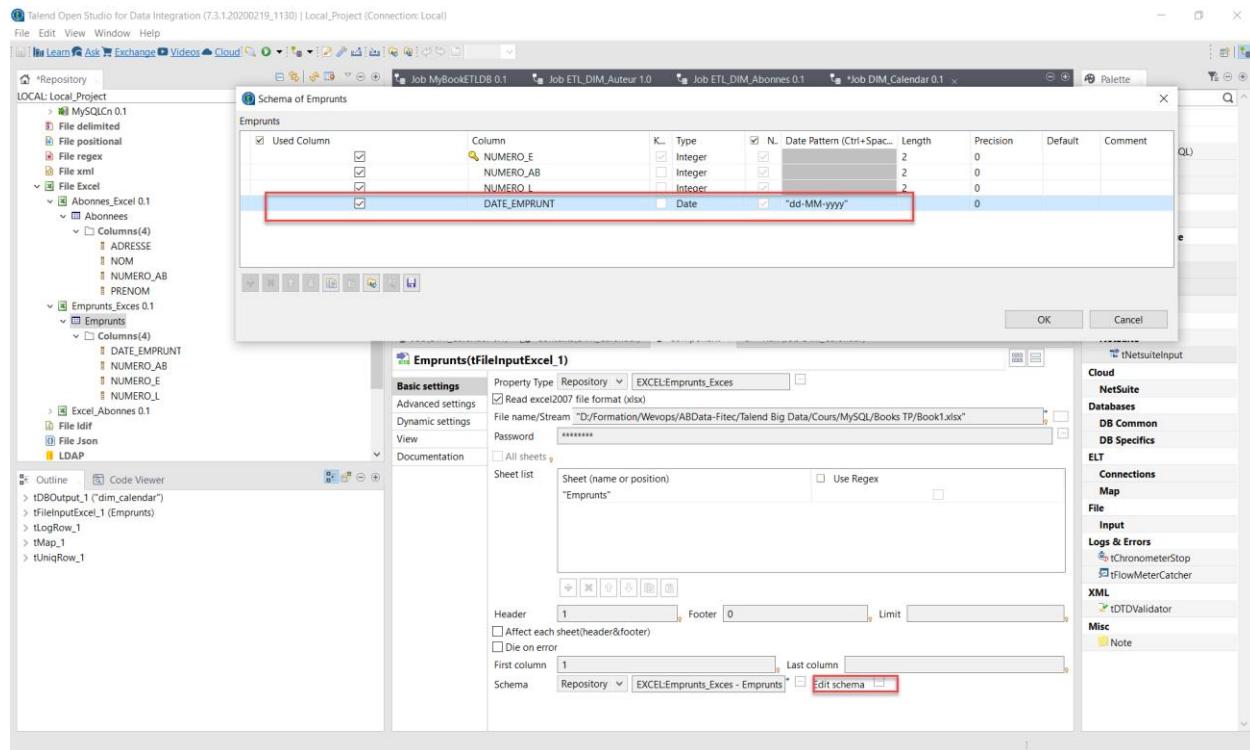
2. Configurer tFileInputExcel :

- Dans l'onglet Component de tFileInputExcel, renseignez les propriétés suivantes :
 - **File Name/Stream** : Indiquez le chemin de votre fichier Excel (**Books.xlsx**).
 - **Sheet name** : Entrez le nom de la feuille contenant les données (par exemple, **Emprunts**).
 - **Header** : Mettez **1** si la première ligne contient les en-têtes de colonnes.
 - **Set encoding** : Choisissez **UTF-8** pour l'encodage.

3. Définir le schéma :

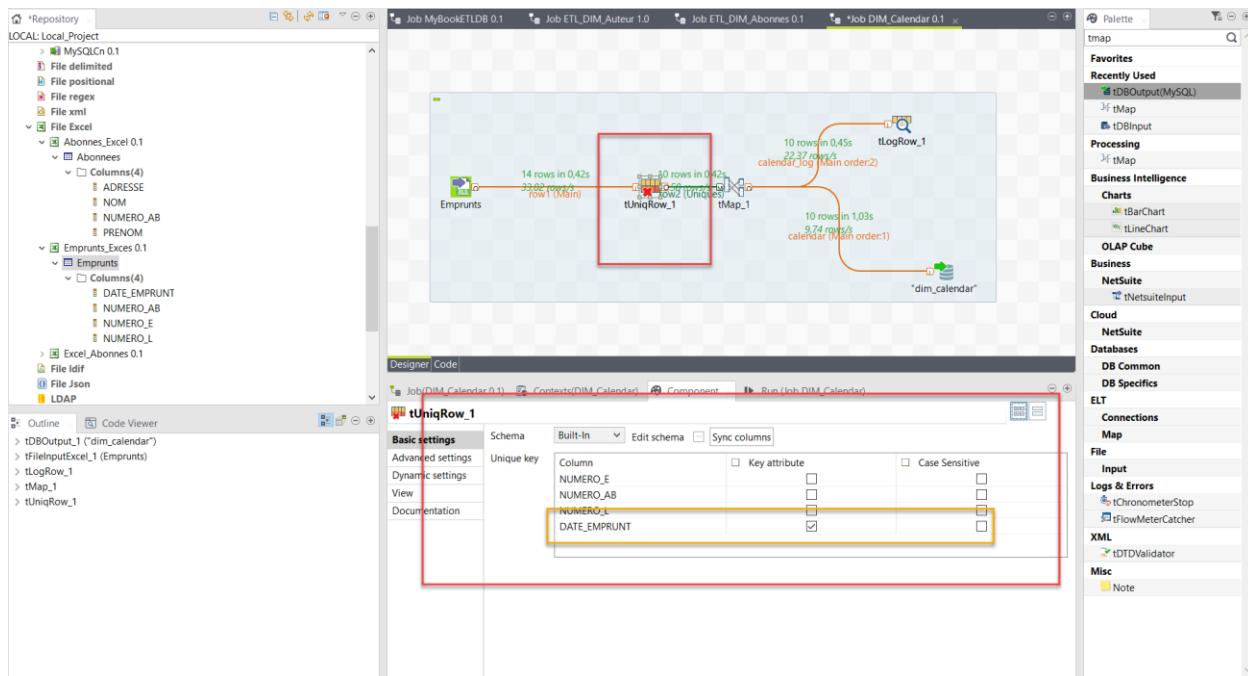
- Cliquez sur **Edit Schema** et définissez la colonne **Date Emprunt** en tant que **Date**. Configurez le format de date selon votre fichier, par exemple **dd-MM-yyyy**.

Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com



b) tUniqRow - Filtrage des valeurs uniques

1. Ajouter tUniqRow : Faites glisser tUniqRow depuis la palette et connectez tFileInputExcel à tUniqRow avec un lien de type Main.

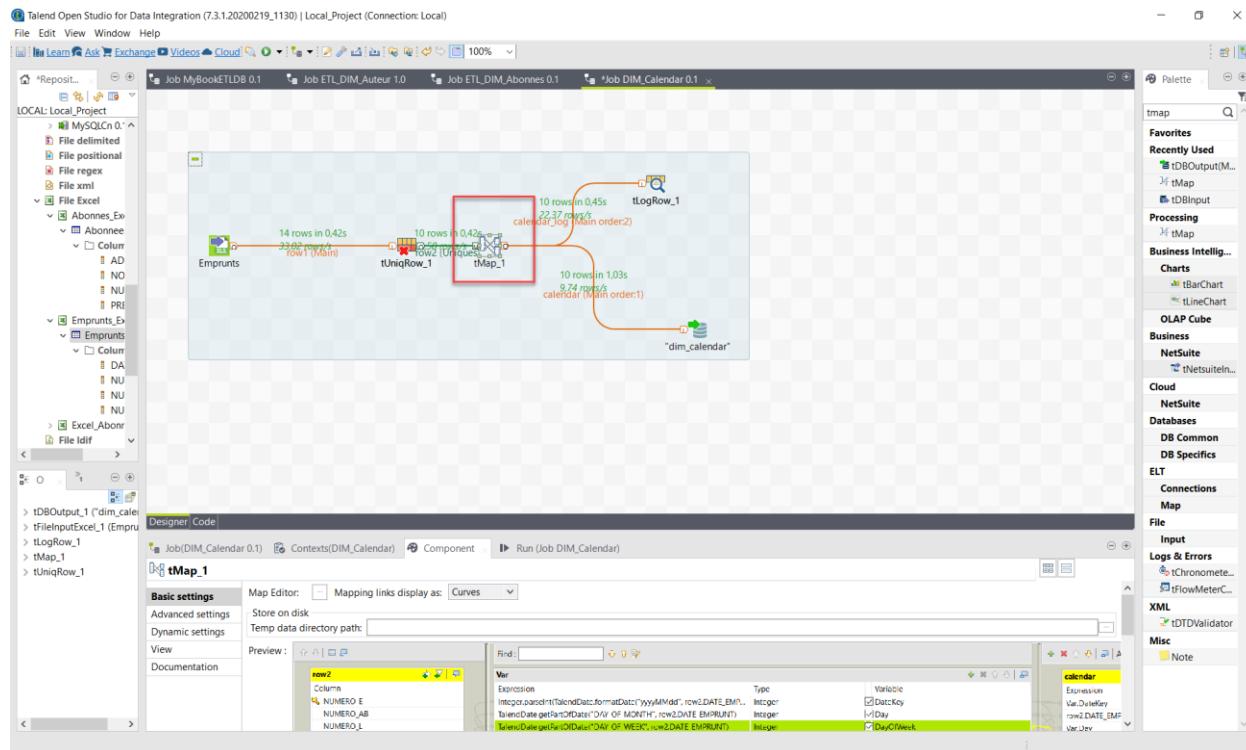


2. Configurer tUniqRow :

- Dans l'onglet Component de tUniqRow, cochez la colonne Date Emprunt dans la section Key columns pour ne conserver que des dates uniques.

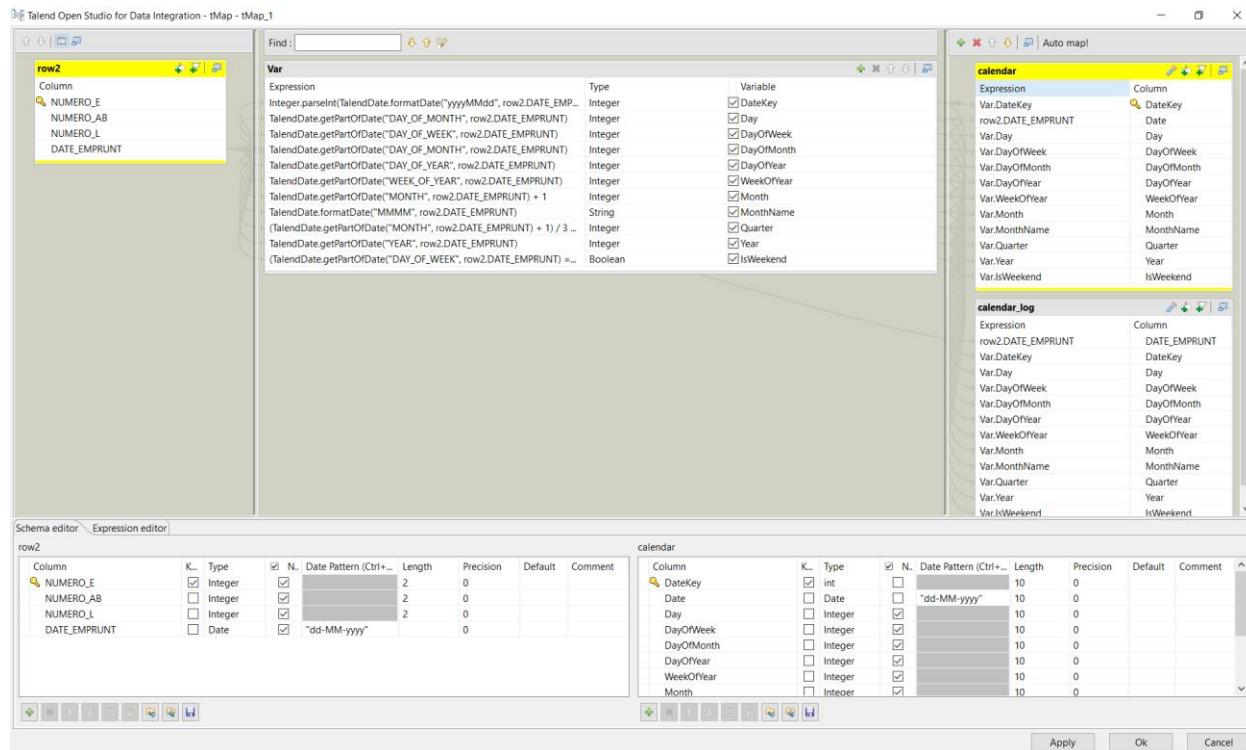
c) tMap - Transformation des données

- Ajouter tMap : Faites glisser le composant tMap et connectez tUniqRow à tMap via un lien Main.



2. Configurer tMap :

- Double-cliquez sur tMap pour ouvrir l'éditeur de mappage.
- Dans l'onglet Output de tMap, ajoutez une sortie appelée calendar pour préparer les données pour dim_calendar.



- Dans cette sortie, ajoutez les champs suivants et leurs transformations :

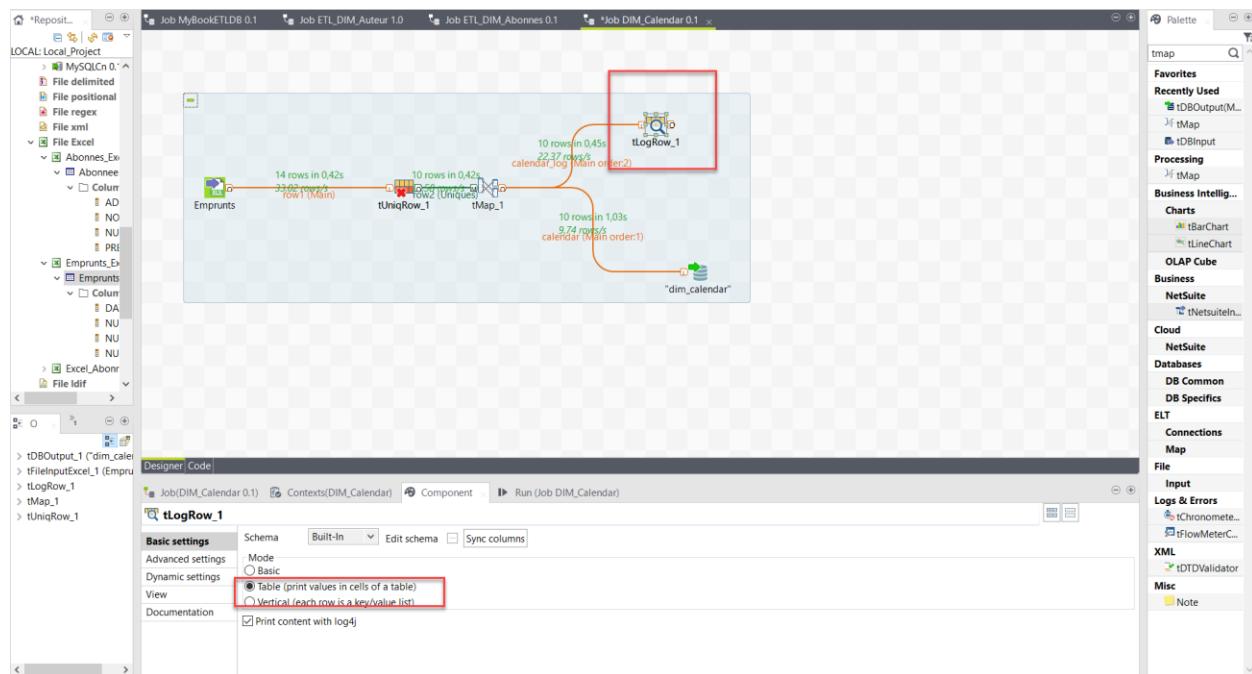
Champ	Expression
DateKey	Integer.parseInt(TalendDate.formatDate("yyyyMMdd", row2.DATE_EMPRUNT))
Day	TalendDate.getDate("DAY_OF_MONTH", row2.DATE_EMPRUNT)
DayOfWeek	TalendDate.getDay("DAY_OF_WEEK", row2.DATE_EMPRUNT)
DayOfMonth	TalendDate.getDate("DAY_OF_MONTH", row2.DATE_EMPRUNT)
DayOfYear	TalendDate.getDate("DAY_OF_YEAR", row2.DATE_EMPRUNT)
WeekOfYear	TalendDate.getDate("WEEK_OF_YEAR", row2.Date_Emprunt)
Month	TalendDate.getMonth("MONTH", row2.DATE_EMPRUNT)+1
MonthName	TalendDate.formatDate("MMMM", row2.DATE_EMPRUNT)
Quarter	(TalendDate.getMonth("MONTH", row2.DATE_EMPRUNT)+1) / 3 + 1
Year	TalendDate.getYear("YEAR", row2.DATE_EMPRUNT)
IsWeekend	(TalendDate.getDay("DAY_OF_WEEK", row1.Date_Emprunt) == 7 TalendDate.getDay("DAY_OF_WEEK", row1.Date_Emprunt) == 6)

WeekOfMonth **(TalendDate.getDate("DAY_OF_WEEK", row2.DATE_EMPRUNT) == 1 || TalendDate.getDate("DAY_OF_WEEK", row2.DATE_EMPRUNT) == 7)**

- Cliquez sur OK pour enregistrer les modifications et fermer tMap.

d) tLogRow - Vérification des données

1. Ajouter tLogRow : Faites glisser tLogRow et connectez-le à la sortie calendar de tMap.

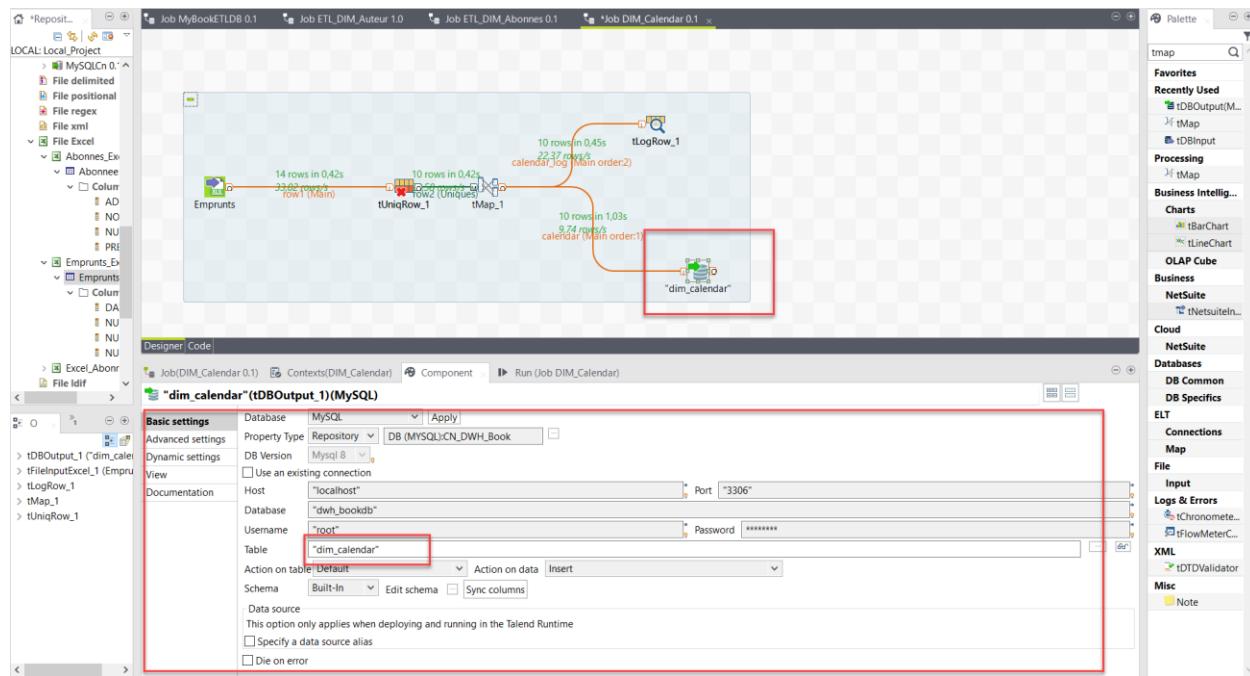


2. Configurer tLogRow :

- Dans l'onglet Component de tLogRow, sélectionnez Table comme Mode pour afficher les données sous forme de tableau dans la console lors de l'exécution du job.

e) tOutputMSQL - Chargement dans la base de données

1. Ajouter tOutputMSQL : Faites glisser tOutputMSQL et connectez-le à la sortie calendar de tMap.



2. Configurer tOutputMSQL :

- Dans l'onglet Component, configurez la connexion à votre base de données SQL Server.
 - Table : dim_calendar
 - Action on table : None (ou Drop and Create si vous voulez recréer la table à chaque exécution).
 - Action on data : Insert pour insérer les données transformées.
- Cliquez sur Edit Schema pour définir le schéma de la table dim_calendar en fonction des champs configurés dans tMap.

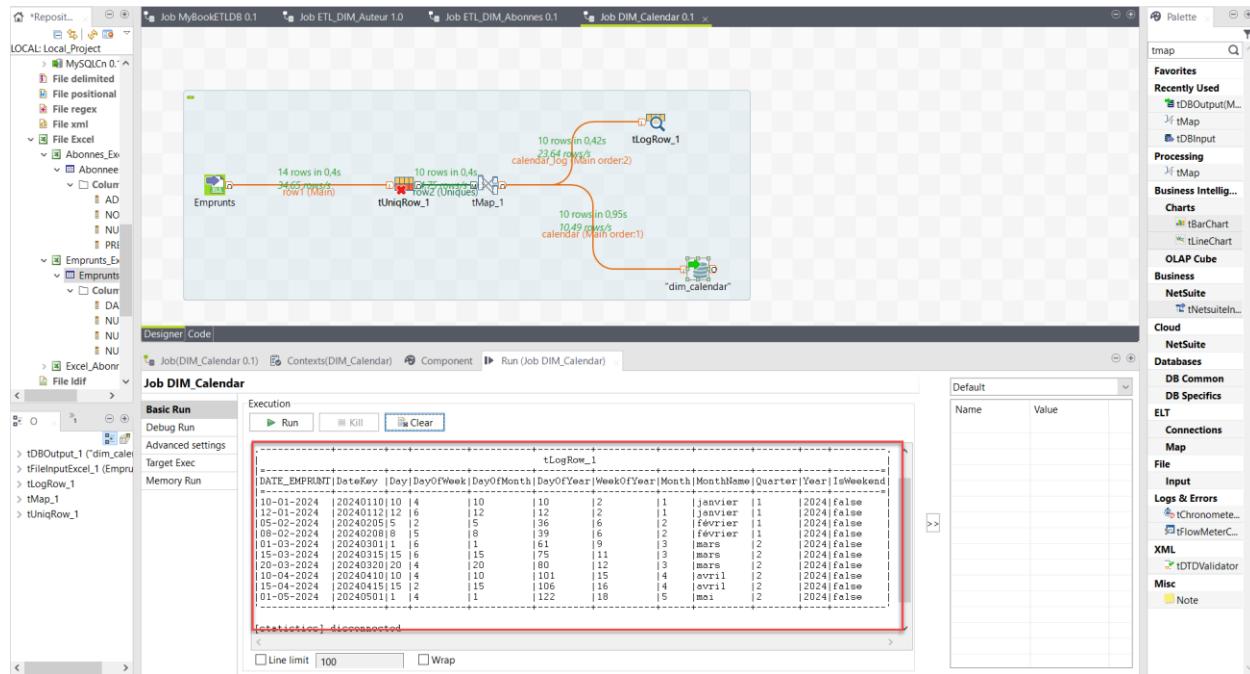
3. Exécuter et tester le job

1. Lancer le job : Cliquez sur le bouton Run pour exécuter le job.

2. Vérifier les logs :

- Consultez tLogRow pour vérifier que les transformations sont correctes.

Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com



3. Vérifier dans SQL Server :

- Après l'exécution, ouvrez SQL Server et interrogez la table dim_calendar pour confirmer que les données ont été correctement chargées.

The screenshot shows the MySQL Workbench interface. The left sidebar displays the "Schemas" tree, with the "dwh_bookdb" schema selected. Under "Tables", the "dim_calendar" table is highlighted with a red box. The main pane shows the results of a query: "SELECT * FROM dwh_bookdb.dim_calendar;". The result grid displays the data from the "dim_calendar" table, matching the structure shown in the Talend preview. A red box highlights the result grid area.

DateKey	Date	Day	DayOfWeek	DayOfMonth	DayOfYear	WeekOfYear	Month	MonthName	Quarter	Year	IsWeekend
20240110	2024-01-10	10	4	10	10	2	1	janvier	1	2024	0
20240111	2024-01-11	11	5	11	11	3	2	janvier	1	2024	0
20240112	2024-01-12	12	6	12	12	4	1	janvier	1	2024	0
20240113	2024-01-13	13	7	13	13	5	2	janvier	1	2024	0
20240114	2024-01-14	14	1	14	14	6	3	janvier	1	2024	0
20240115	2024-01-15	15	2	15	15	7	4	janvier	1	2024	0
20240116	2024-01-16	16	3	16	16	8	5	janvier	1	2024	0
20240117	2024-01-17	17	4	17	17	9	6	janvier	1	2024	0
20240118	2024-01-18	18	5	18	18	10	7	janvier	1	2024	0
20240119	2024-01-19	19	6	19	19	11	8	janvier	1	2024	0
20240120	2024-01-20	20	7	20	20	12	9	janvier	1	2024	0
20240121	2024-01-21	21	1	21	21	13	10	janvier	1	2024	0
20240122	2024-01-22	22	2	22	22	14	11	janvier	1	2024	0
20240123	2024-01-23	23	3	23	23	15	12	janvier	1	2024	0
20240124	2024-01-24	24	4	24	24	16	13	janvier	1	2024	0
20240125	2024-01-25	25	5	25	25	17	14	janvier	1	2024	0
20240126	2024-01-26	26	6	26	26	18	15	janvier	1	2024	0
20240127	2024-01-27	27	7	27	27	19	16	janvier	1	2024	0
20240128	2024-01-28	28	1	28	28	20	17	janvier	1	2024	0
20240129	2024-01-29	29	2	29	29	21	18	janvier	1	2024	0
20240130	2024-01-30	30	3	30	30	22	19	janvier	1	2024	0
20240131	2024-01-31	31	4	31	31	23	20	janvier	1	2024	0
20240201	2024-02-01	1	5	1	1	24	21	fevrier	1	2024	0
20240202	2024-02-02	2	6	2	2	25	22	fevrier	1	2024	0
20240203	2024-02-03	3	7	3	3	26	23	fevrier	1	2024	0
20240204	2024-02-04	4	1	4	4	27	24	fevrier	1	2024	0
20240205	2024-02-05	5	2	5	5	28	25	fevrier	1	2024	0
20240206	2024-02-06	6	3	6	6	29	26	fevrier	1	2024	0
20240207	2024-02-07	7	4	7	7	30	27	fevrier	1	2024	0
20240208	2024-02-08	8	5	8	8	31	28	fevrier	1	2024	0
20240209	2024-02-09	9	6	9	9	1	1	mars	2	2024	0
20240210	2024-02-10	10	7	10	10	2	2	mars	2	2024	0
20240211	2024-02-11	11	1	11	11	3	3	mars	2	2024	0
20240212	2024-02-12	12	2	12	12	4	4	mars	2	2024	0
20240213	2024-02-13	13	3	13	13	5	5	mars	2	2024	0
20240214	2024-02-14	14	4	14	14	6	6	mars	2	2024	0
20240215	2024-02-15	15	5	15	15	7	7	mars	2	2024	0
20240216	2024-02-16	16	6	16	16	8	8	mars	2	2024	0
20240217	2024-02-17	17	7	17	17	9	9	mars	2	2024	0
20240218	2024-02-18	18	1	18	18	10	10	mars	2	2024	0
20240219	2024-02-19	19	2	19	19	11	11	mars	2	2024	0
20240220	2024-02-20	20	3	20	20	12	12	mars	2	2024	0
20240221	2024-02-21	21	4	21	21	13	13	mars	2	2024	0
20240222	2024-02-22	22	5	22	22	14	14	mars	2	2024	0
20240223	2024-02-23	23	6	23	23	15	15	mars	2	2024	0
20240224	2024-02-24	24	7	24	24	16	16	mars	2	2024	0
20240225	2024-02-25	25	1	25	25	17	17	mars	2	2024	0
20240226	2024-02-26	26	2	26	26	18	18	mars	2	2024	0
20240227	2024-02-27	27	3	27	27	19	19	mars	2	2024	0
20240228	2024-02-28	28	4	28	28	20	20	mars	2	2024	0
20240229	2024-02-29	29	5	29	29	21	21	mars	2	2024	0
20240301	2024-03-01	1	6	1	1	22	22	avril	1	2024	0
20240302	2024-03-02	2	7	2	2	23	23	avril	1	2024	0
20240303	2024-03-03	3	1	3	3	24	24	avril	1	2024	0
20240304	2024-03-04	4	2	4	4	25	25	avril	1	2024	0
20240305	2024-03-05	5	3	5	5	26	26	avril	1	2024	0
20240306	2024-03-06	6	4	6	6	27	27	avril	1	2024	0
20240307	2024-03-07	7	5	7	7	28	28	avril	1	2024	0
20240308	2024-03-08	8	1	8	8	29	29	avril	1	2024	0
20240309	2024-03-09	9	2	9	9	30	30	avril	1	2024	0
20240310	2024-03-10	10	3	10	10	1	1	mai	2	2024	0
20240311	2024-03-11	11	4	11	11	2	2	mai	2	2024	0
20240312	2024-03-12	12	5	12	12	3	3	mai	2	2024	0
20240313	2024-03-13	13	6	13	13	4	4	mai	2	2024	0
20240314	2024-03-14	14	7	14	14	5	5	mai	2	2024	0
20240315	2024-03-15	15	1	15	15	6	6	mai	2	2024	0
20240316	2024-03-16	16	2	16	16	7	7	mai	2	2024	0
20240317	2024-03-17	17	3	17	17	8	8	mai	2	2024	0
20240318	2024-03-18	18	4	18	18	9	9	mai	2	2024	0
20240319	2024-03-19	19	5	19	19	10	10	mai	2	2024	0
20240320	2024-03-20	20	6	20	20	11	11	mai	2	2024	0
20240321	2024-03-21	21	7	21	21	12	12	mai	2	2024	0
20240322	2024-03-22	22	1	22	22	13	13	mai	2	2024	0
20240323	2024-03-23	23	2	23	23	14	14	mai	2	2024	0
20240324	2024-03-24	24	3	24	24	15	15	mai	2	2024	0
20240325	2024-03-25	25	4	25	25	16	16	mai	2	2024	0
20240326	2024-03-26	26	5	26	26	17	17	mai	2	2024	0
20240327	2024-03-27	27	6	27	27	18	18	mai	2	2024	0
20240328	2024-03-28	28	7	28	28	19	19	mai	2	2024	0
20240329	2024-03-29	29	1	29	29	20	20	mai	2	2024	0
20240330	2024-03-30	30	2	30	30	21	21	mai	2	2024	0
20240331	2024-03-31	31	3	31	31	22	22	mai	2	2024	0

Repository LOCAL: BookETL

- Business Models
- Job Designs
 - BookETL_Job_v1 0.1
- Contexts
- Code
- SQL Templates
- Metadata
 - Db Connections
 - Cn_DWH_bookd[0.1]
 - MySQLCN_BookE
 - File delimited
 - File positional
 - File regex
 - File xml
 - File Excel
 - File Idif
 - File Json
 - LDAP
 - Azure Storage
 - Google Drive
 - Marketo
 - Salesforce
 - Snowflake
 - Generic schemas
 - Talend MDM
 - Rules Management
 - Web Service
 - FTP



Talend for Data Integration



Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com

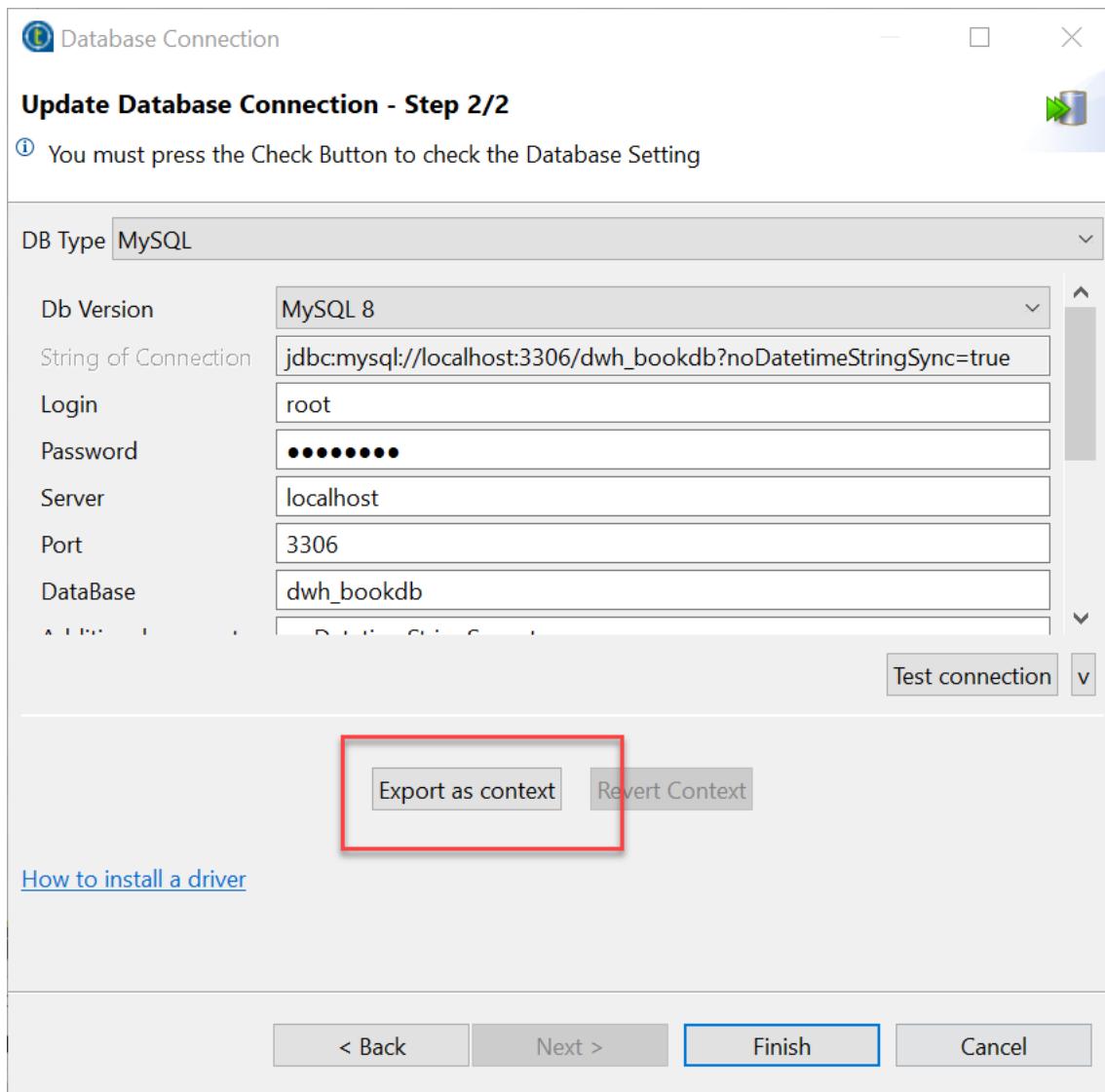
Database Connection

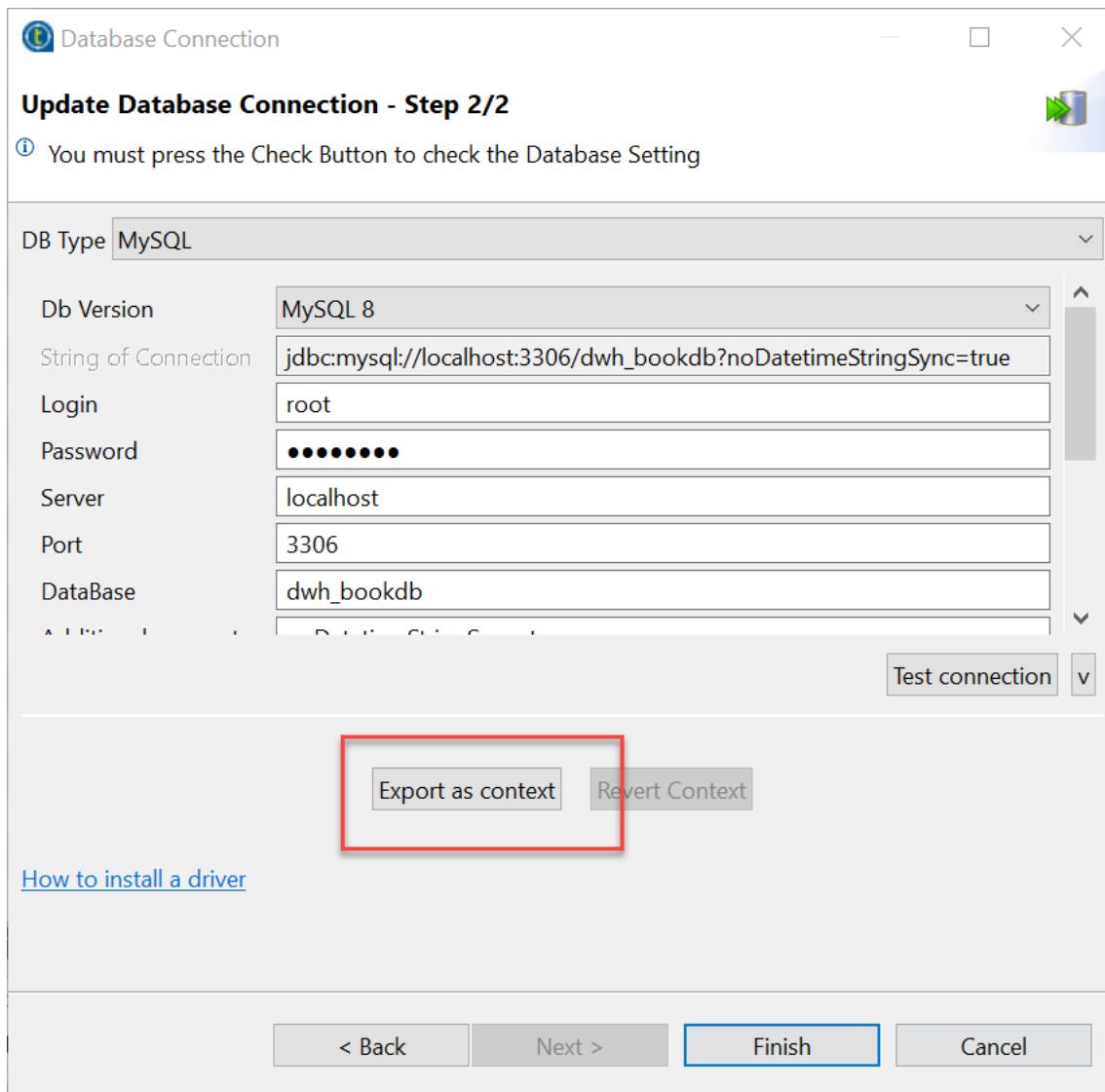
Update Database Connection - Step 1/2

Update properties

Name	Cn_DWH_bookdb		
Purpose	Cn_DWH_bookdb		
Description	Cn_DWH_bookdb		
Author	user@talend.com		
Locker	user@talend.com		
Version	0.1	M	m
Status			
Path			

< Back Next > **Finish** Cancel





Create / Reuse a context group

Step 1 of 2

Add any required information

Name	Cn_DWH_bookdb	
Purpose		
Description		
Author	user@talend.com	
Locker		
Version	0.1	M m
Status		▼
Path		Select

< Back Next > **Finish** Cancel

Create / Reuse a context group

Step 2 of 2

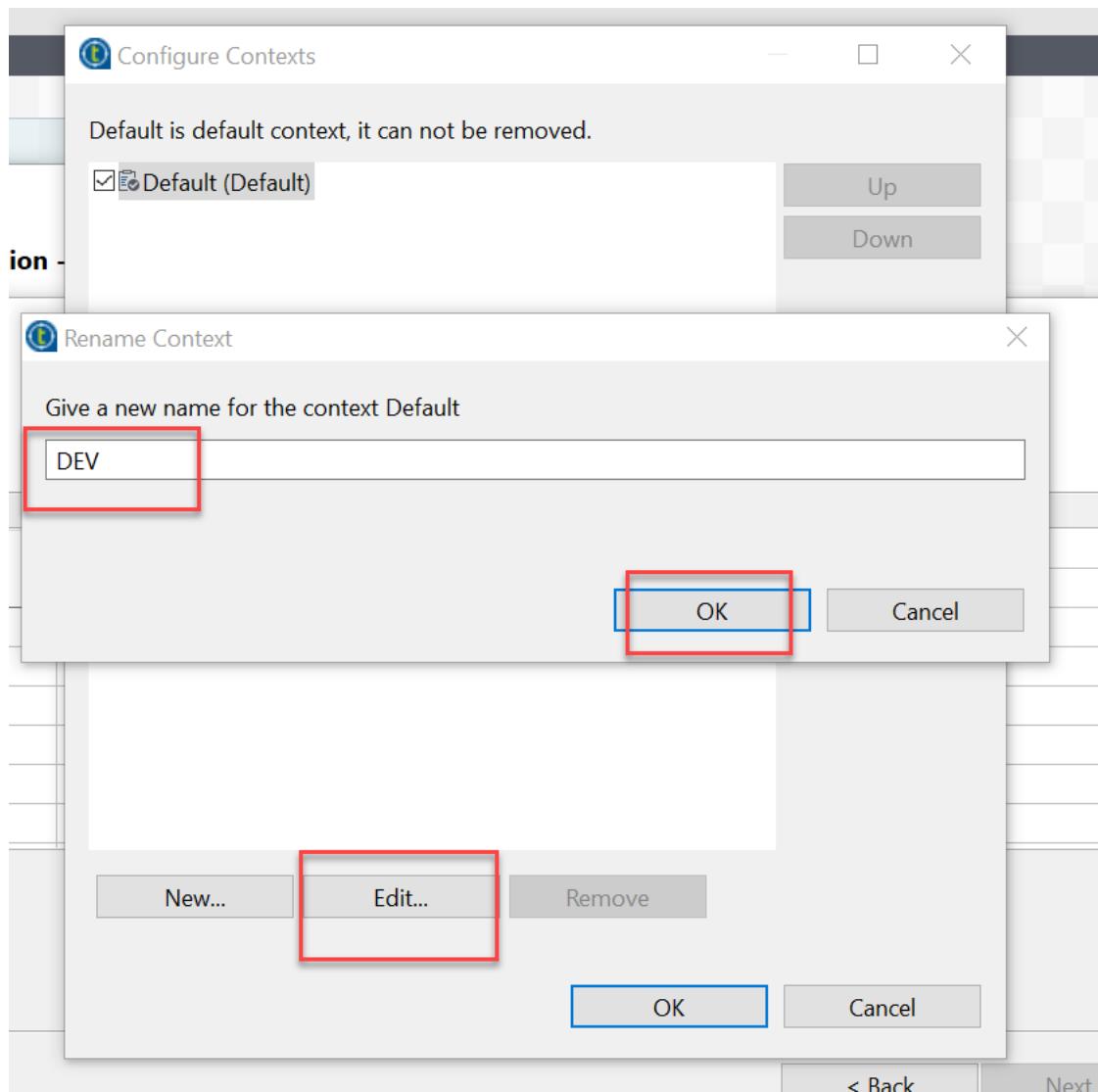
Define the contexts, variables and values

	Name	Type	Comment	Default Value
1				
2	Cn_DWH_bookdb_Password	Password		*****
3	Cn_DWH_bookdb_Login	String		root
4	Cn_DWH_bookdb_AdditionalParams	String		noDatetimeStringSync=true
5	Cn_DWH_bookdb_Server	String		localhost
6	Cn_DWH_bookdb_Port	String		3306

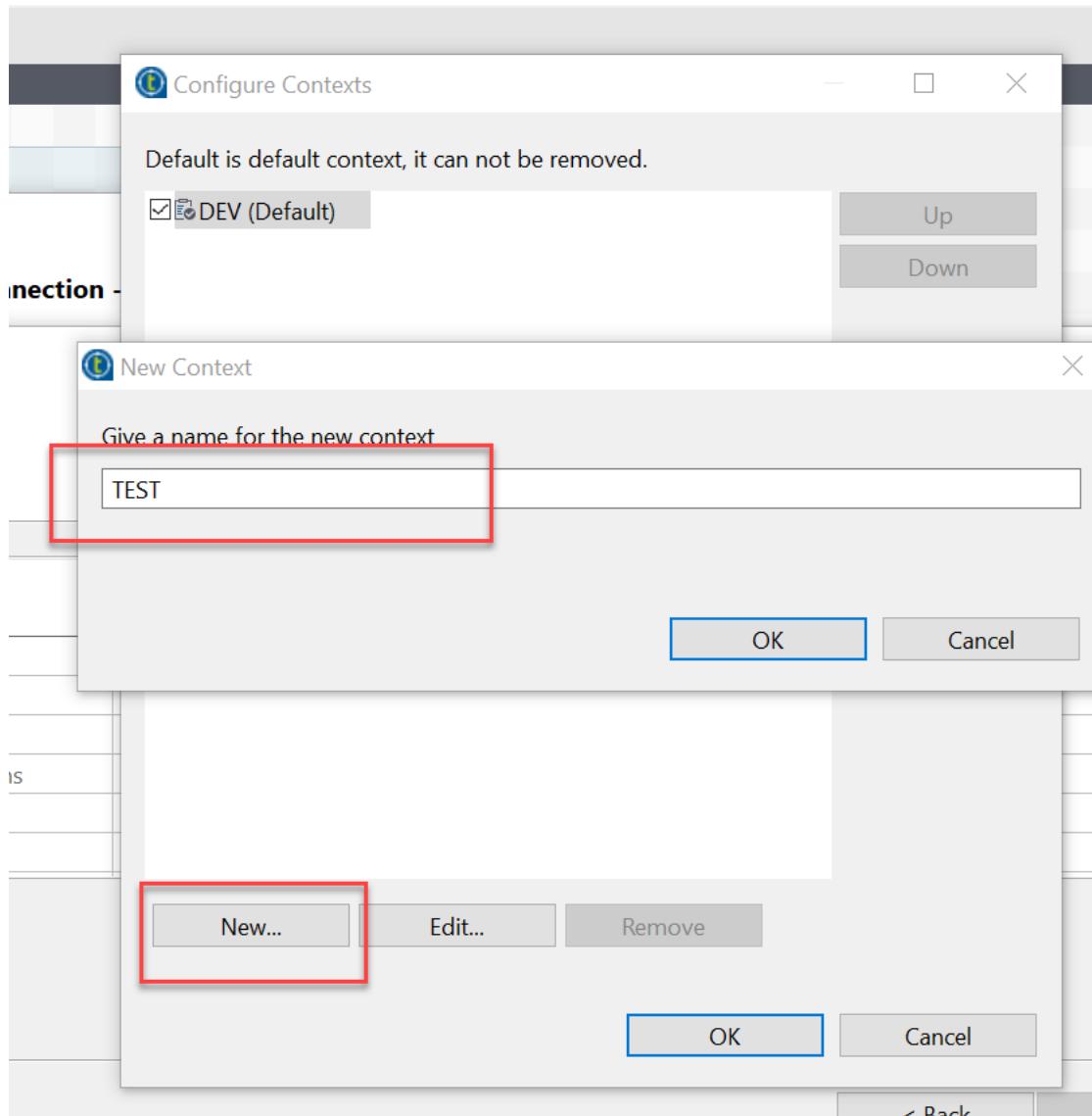
+ X Default context environment **Default**

< Back Next > **Finish** Cancel

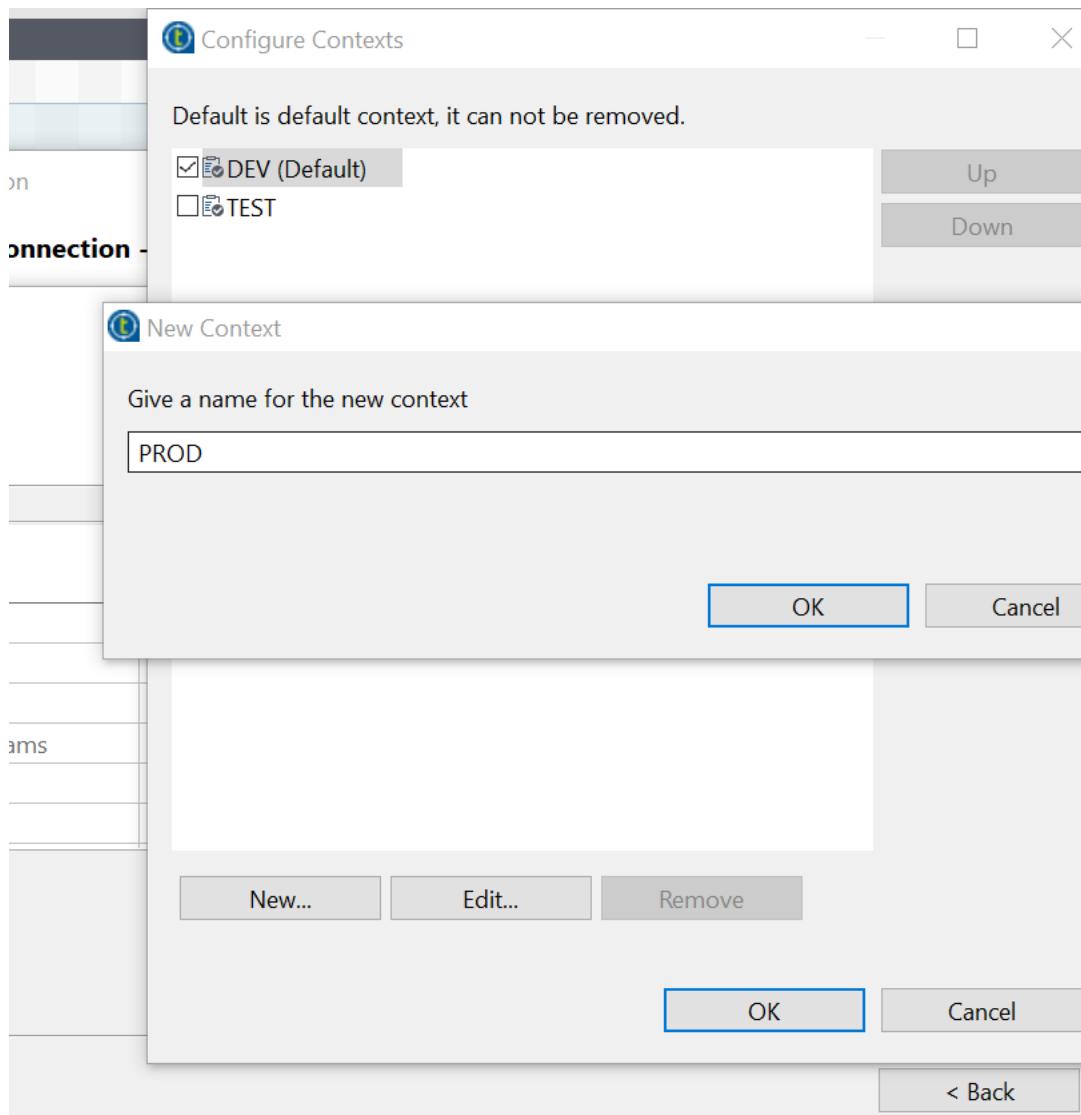
Editer le context Default et Renommer « DEV »



Ajouter un nouveau Context : « TEST »



Ajouter un nouveau Context : « PROD »



3 Environnement



Talend for Data Integration



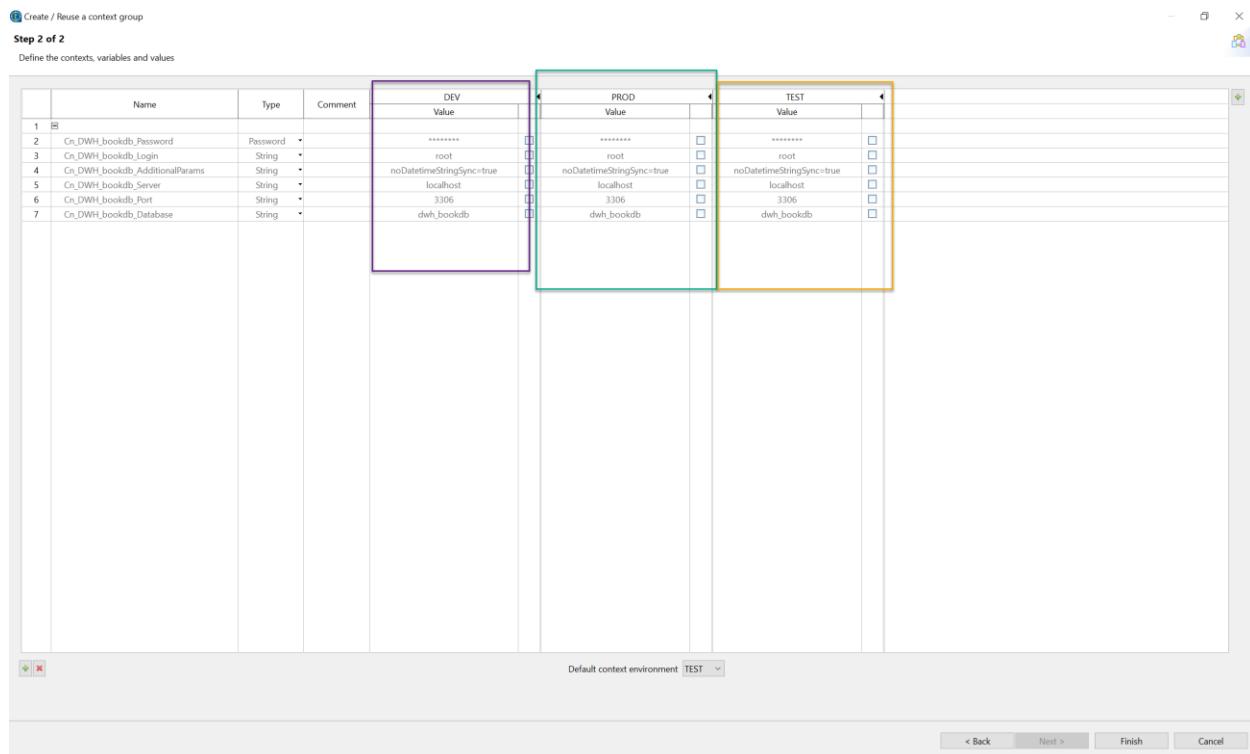
Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com

Create / Reuse a context group
Step 2 of 2
Define the contexts, variables and values

	Name	Type	Comment	DEV	PROD	TEST
1	Cn_DWH_bookdb.Password	Password		*****	*****	*****
2	Cn_DWH_bookdb.Login	String		root	root	root
3	Cn_DWH_bookdb.AdditionalParams	String		noDatetimeStringSync=true	noDatetimeStringSync=true	noDatetimeStringSync=true
4	Cn_DWH_bookdb.Server	String		localhost	localhost	localhost
5	Cn_DWH_bookdb.Port	String		3306	3306	3306
6	Cn_DWH_bookdb.Database	String		dwh_bookdb	dwh_bookdb	dwh_bookdb

Default context environment TEST

< Back Next > Finish Cancel



Cliquer et Accepter la modification des composants reliées au Contexte "Default" >

Modifier les ports respectant vos serveurs de bases de données en cas vous utiliser des serveurs différents pour chaque environnement



Talend for Data Integration



Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com

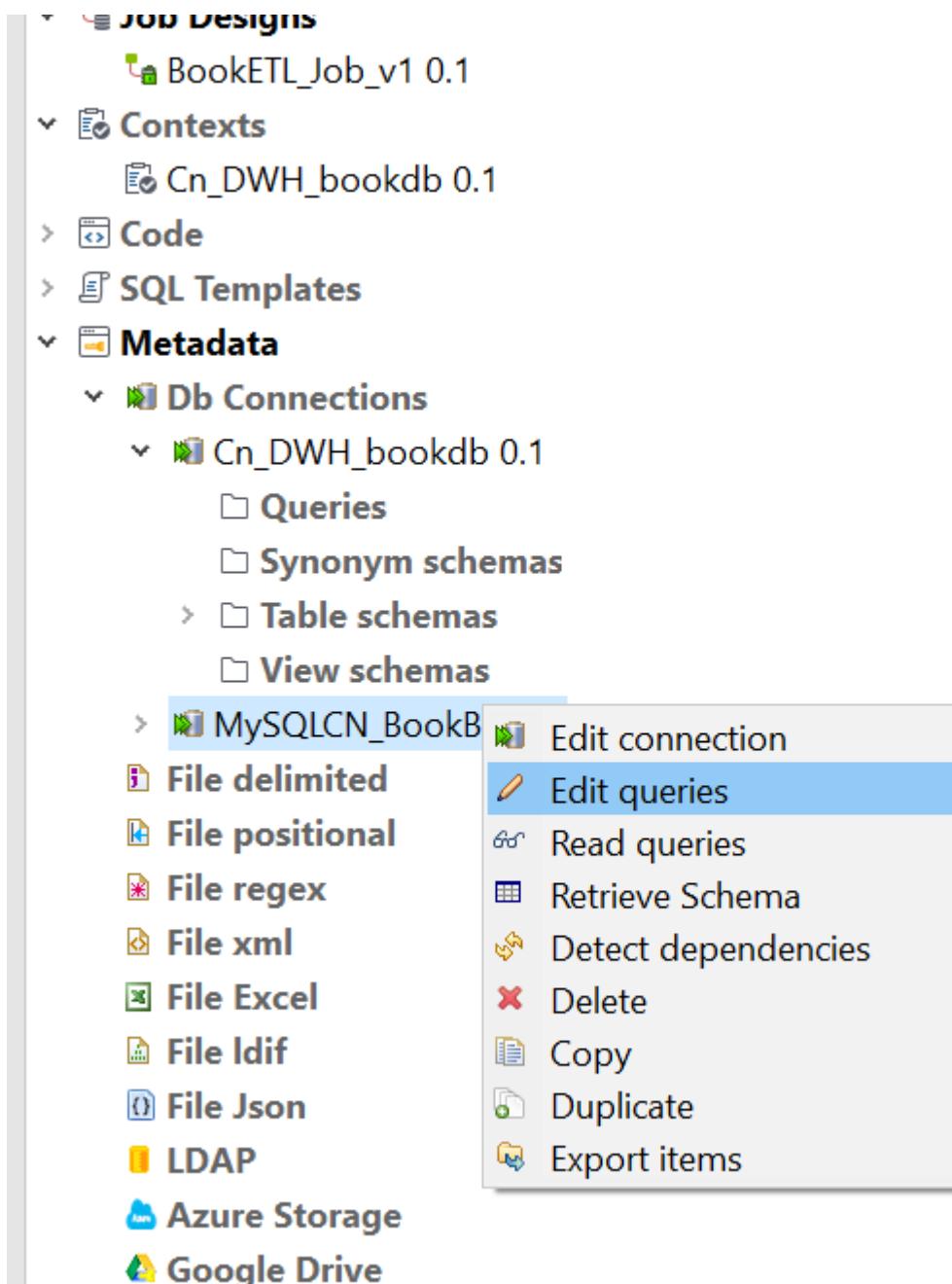
Create / Edit a context group
Step 2 of 2
Define the contexts, variables and values

Name	Type	Comment	DEV	PROD	TEST
			Value	Value	Value
1 Cn_DWH_bookdb_Password	Password		*****	*****	*****
2 Cn_DWH_bookdb_Login	String		root	root	root
3 Cn_DWH_bookdb_AdditionalParams	String		noDatetimeStringSync=true	noDatetimeStringSync=true	noDatetimeStringSync=true
4 Cn_DWH_bookdb_Server	String		localhost	localhost	localhost
5 Cn_DWH_bookdb_Port	String		3306	3308	3307
6 Cn_DWH_bookdb_Database	String		dwh_bookdb	dwh_bookdb	dwh_bookdb

Default context environment TEST

< Back Next > Finish Cancel

Editer maintenant la deuxième connexion pour l'ajouter aux contextes DEV ;TEST, PROD

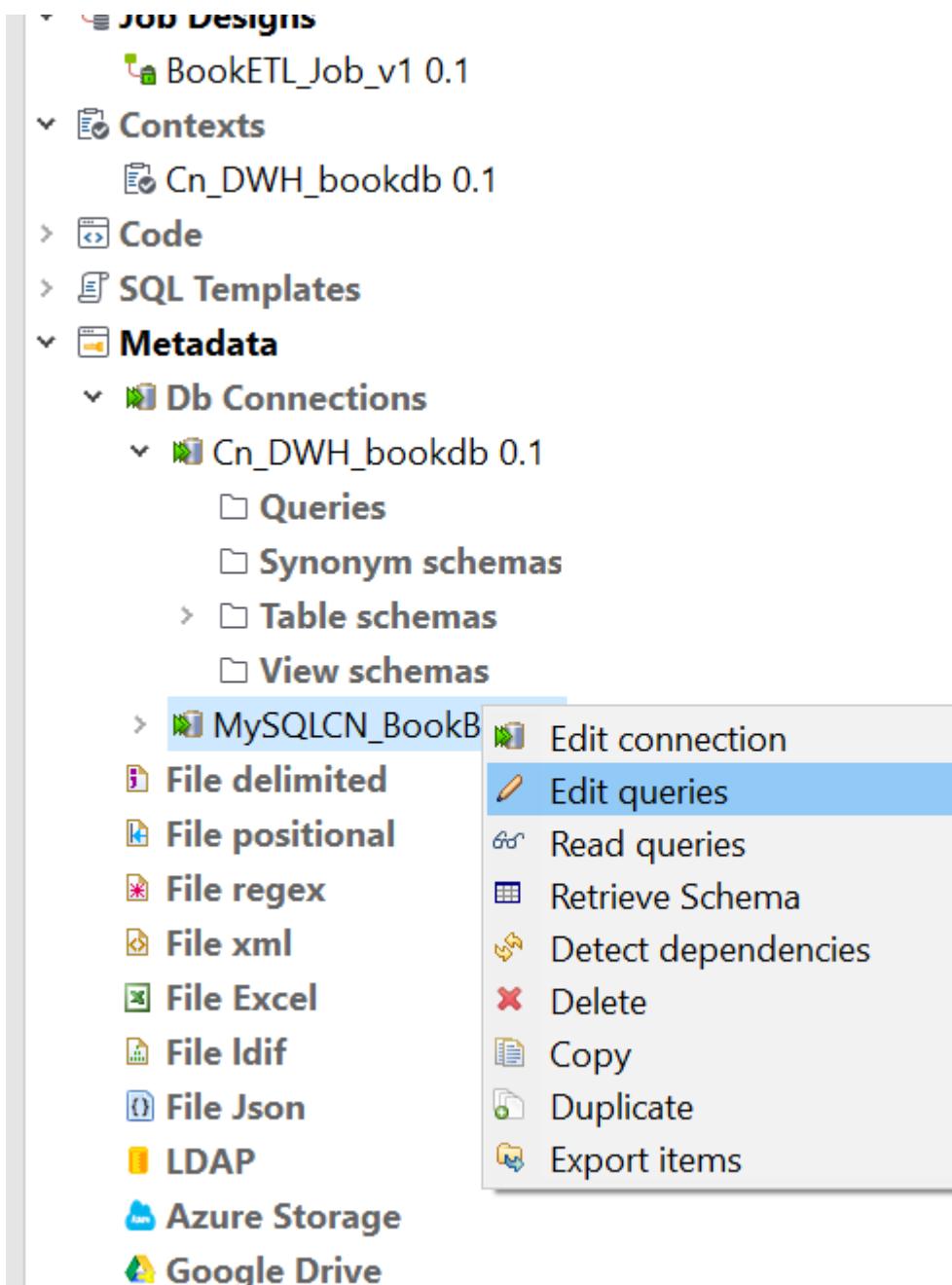


The screenshot shows the Talend Job Explorer interface. The left sidebar displays a tree structure of project components:

- JOB Designs**: Contains a job named BookETL_Job_v1 0.1.
- Contexts**: Contains a context named Cn_DWH_bookdb 0.1.
- Code**: A folder icon.
- SQL Templates**: A folder icon.
- Metadata**:
 - Db Connections**:
 - Cn_DWH_bookdb 0.1
 - Queries
 - Synonym schemas
 - Table schemas
 - View schemas
 - MySQLCN_BookB (selected)
 - File delimited
 - File positional
 - File regex
 - File xml
 - File Excel
 - File Idif
 - File Json
 - LDAP
 - Azure Storage
 - Google Drive

A context menu is open over the MySQLCN_BookB connection entry, listing the following options:

- Edit connection
- Edit queries** (highlighted in blue)
- Read queries
- Retrieve Schema
- Detect dependencies
- Delete
- Copy
- Duplicate
- Export items



The screenshot shows the Talend for Data Integration interface with the following structure in the Job Explorer:

- JOB Designs**: Contains a job named BookETL_Job_v1 0.1.
- Contexts**: Contains a context named Cn_DWH_bookdb 0.1.
- Code**: A folder icon.
- SQL Templates**: A folder icon.
- Metadata**:
 - Db Connections**:
 - Cn_DWH_bookdb 0.1
 - Queries
 - Synonym schemas
 - Table schemas
 - View schemas
 - MySQLCN_BookB (highlighted)
 - File delimited
 - File positional
 - File regex
 - File xml
 - File Excel
 - File Idif
 - File Json
 - LDAP
 - Azure Storage
 - Google Drive

A context menu is open over the MySQLCN_BookB connection entry, displaying the following options:

- Edit connection
- Edit queries** (selected)
- Read queries
- Retrieve Schema
- Detect dependencies
- Delete
- Copy
- Duplicate
- Export items

Database Connection

Update Database Connection - Step 1/2

Update properties

Name	MySQLCN_BookBD
Purpose	MySQLCN_BookBD
Description	MySQLCN_BookBD
Author	user@talend.com
Locker	user@talend.com
Version	0.1
Status	M m
Path	

< Back Next > Next > Finish Cancel

Cliquer sur export as contexte

 Database Connection

Update Database Connection - Step 2/2

ⓘ You must press the Check Button to check the Database Setting 

DB Type	MySQL
Db Version	MySQL 8
String of Connection	<code>jdbc:mysql://192.168.183.90:3306/gestion_livres?noDatetimeStringSync=true</code>
Login	root
Password	*****
Server	192.168.183.90
Port	3306
DataBase	gestion livres





[How to install a driver](#)



Talend for Data Integration



Formateur : Sellami Mokhtar
mokhtar.sellami@gmail.com

Database Connection

Update Database Connection - Step 2/2

Create / Reuse a context group

Select context group

Select a context group and click Next.

Available Context Groups
Cn_DWH_bookdb

< Back **Next >** Finish Cancel

Export as context Revert Context

[How to install a driver](#)

< Back Next > **Finish** Cancel

