

AMIGOS

Rapport

Talend open studio

2022

Contents

Contents

Informations sur le travail

01

Introduction

02

Collection

03

Démentions

04

Les faits

05

Conclusion

06

Ce rapport parle de
comment utiliser Talend
Open Studio



Informations sur le travail



Date d'initiation:

Date de livraison:



Type travail Individual.



Nom du créateur: Boufares
Mohamed Karim

Introduction

Talend Studio est un outil de développement intégré (IDE) pour la conception et la mise en œuvre de solutions d'intégration de données. Il aide les développeurs à créer des flux de données entre différents systèmes, applications et bases de données.

Talend Studio offre une interface graphique intuitive et facile à utiliser pour la création de jobs, qui sont des processus de traitement de données. Les développeurs peuvent utiliser un ensemble de composants prédéfinis pour effectuer des tâches telles que l'extraction, la transformation et le chargement de données (ETL). Talend Studio prend en charge un large éventail de sources et de destinations de données, notamment les bases de données relationnelles, les fichiers, les applications cloud, les APIs et plus encore.

Talend Studio fournit également des fonctionnalités de test et de débogage pour les jobs, ainsi que des fonctionnalités de suivi des erreurs et de journalisation pour aider les développeurs à diagnostiquer les problèmes. Enfin, Talend Studio peut être intégré à d'autres outils et technologies, comme Apache Spark et Apache Kafka, pour une intégration plus poussée.

En résumé, Talend Studio est un outil puissant pour les développeurs souhaitant automatiser les processus d'intégration de données et améliorer la qualité des données utilisées par leurs applications et systèmes.

1-Defintion :

une collection est un type de variable qui permet de stocker plusieurs éléments de données dans un seul objet. Les collections peuvent être utilisées pour regrouper des données provenant de différentes sources, telles que des bases de données, des fichiers ou des API, et pour les manipuler de manière efficace dans des jobs Talend.

Il existe plusieurs types de collections dans Talend Open Studio, notamment les tableaux (arrays), les listes (lists), les ensembles (sets) et les map (maps). Chacun de ces types de collections a ses propres caractéristiques et utilisations, et peut être utilisé en fonction des besoins de chaque job.

Les collections peuvent être créées et initialisées dans Talend Open Studio à l'aide de l'interface graphique ou du code. Les éléments de données peuvent être ajoutés, modifiés et supprimés à partir d'une collection à l'aide de différentes opérations, telles que des boucles, des filtres et des transformations. Les collections peuvent également être utilisées pour stocker les résultats de requêtes SQL, d'interrogations d'API et d'autres opérations de traitement de données.

En résumé, les collections sont un élément clé de Talend Open Studio pour la manipulation efficace de grandes quantités de données dans des jobs. Elles permettent aux développeurs de regrouper des données provenant de différentes sources, de les traiter de manière efficace et de les stocker sous forme de variables pour une utilisation ultérieure.

2-Les étapes pour créer et utiliser une collection dans Talend Open Studio sont les suivantes :

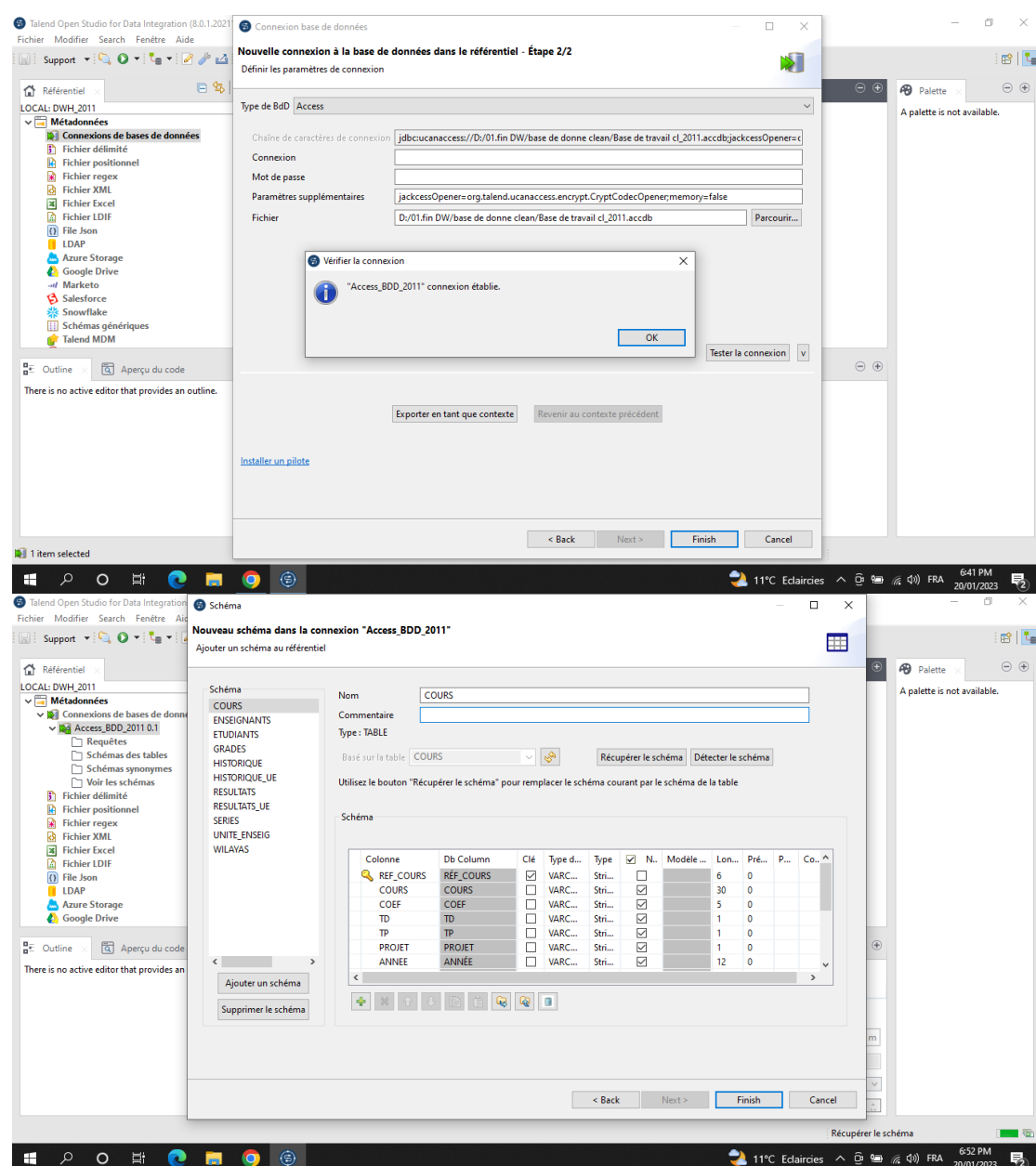
a-Choisissez le type de collection à utiliser

b-Initialisez la collection

c-Ajoutez des éléments de données à la collection

d-Manipulez les données de la collection

e-Stockez et utilisez les données de la collection



Talend Open Studio for Data Integration (8.0.1.2021)

Fichier Modifier Search Fenêtre Aide

Support

LOCAL: DWH_2011

Jobs

Contextes

Code

Modèles SQL

Métadonnées

Connexions de bases de données

Access_BDD_2011 0.1

Fichier délimité

Fichier positionnel

Fichier regex

Fichier XML

Fichier Excel

Fichier LDIF

File Ism

LDAP

Azure Storage

Google Drive

Outline

Aperçu du code

There is no active editor that provides an outline.

Connexion base de données

Nouvelle connexion à la base de données dans le référentiel - Étape 2/2

Définir les paramètres de connexion

Type de Bdd MySQL

Version de la base de données MySQL 8

Chaine de caractères de connexion jdbc:mysql://localhost:3306/dwh_2011?noDatetimeStringSync=true&enabledTLSProtocols=TLSv1

Connexion root

Mot de passe

Serveur localhost

Port 3306

Base de données dwh_2011

Paramètres supplémentaires noDatetimeStringSync=true&enabledTLSProtocols=TLSv1.2,TLSv1.1,TLSv1

Vérifier la connexion

"MySQL_BDD_2011" connexion établie.

OK

Tester la connexion

Exporter en tant que contexte

Revenir au contexte précédent

Installer un pilote

< Back

Next >

Finish

Cancel

1 item selected

11°C Eclaircies

6:55 PM

20/01/2023

Talend Open Studio for Data Integration - tMap - tMap_1

Find:

Var

Auto map!

row1

Column

REF_COURS

COURS

COEF

TD

TP

PROJET

ANNEE

NBREEMD

SEMESTRE

CREDIT

COEFEXAM

COEFCC

CODEUE

OBLIGATOIRE

cours

Expression

Column

row1.REF_COURS

REF_COURS

row1.COURS

COURS

row1.COEF

COEF

row1.TD

TD

row1.TP

TP

row1.PROJET

PROJET

row1.ANNEE

ANNEE

row1.NBREEMD

NBREEMD

row1.SEMESTRE

SEMESTRE

row1.CREDIT

CREDIT

row1.COEFEXAM

COEFEXAM

row1.COEFCC

COEFCC

row1.CODEUE

CODEUE

row1.OBLIGATOIRE

OBLIGATOIRE

Éditeur de schéma

Éditeur d'expression

row1

Colonne

Clé

Type

N.

Modèle de da...

Longu...

Précision

Par d...

Comm...

REF_COURS

String

String

6

0

COURS

String

String

30

0

COEF

String

String

5

0

TD

String

String

1

0

cours

Colonne

Clé

Type

N.

Modèle de da...

Longu...

Précision

Par d...

Comm...

REF_COURS

String

String

6

0

COURS

String

String

30

0

COEF

String

String

5

0

TD

String

String

1

0

Appliquer

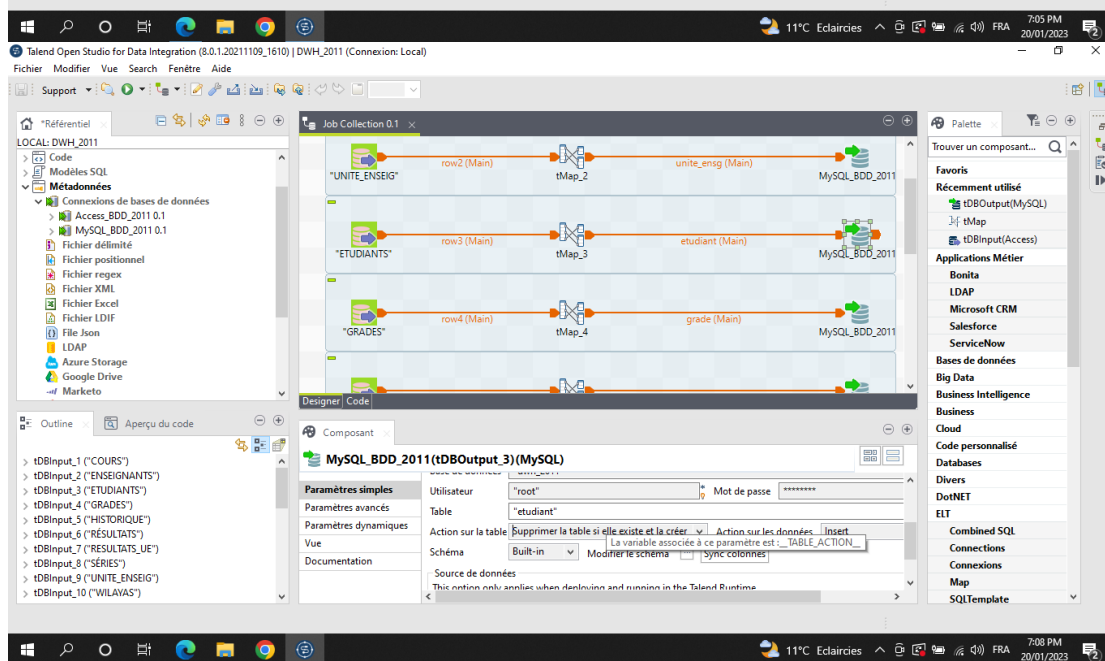
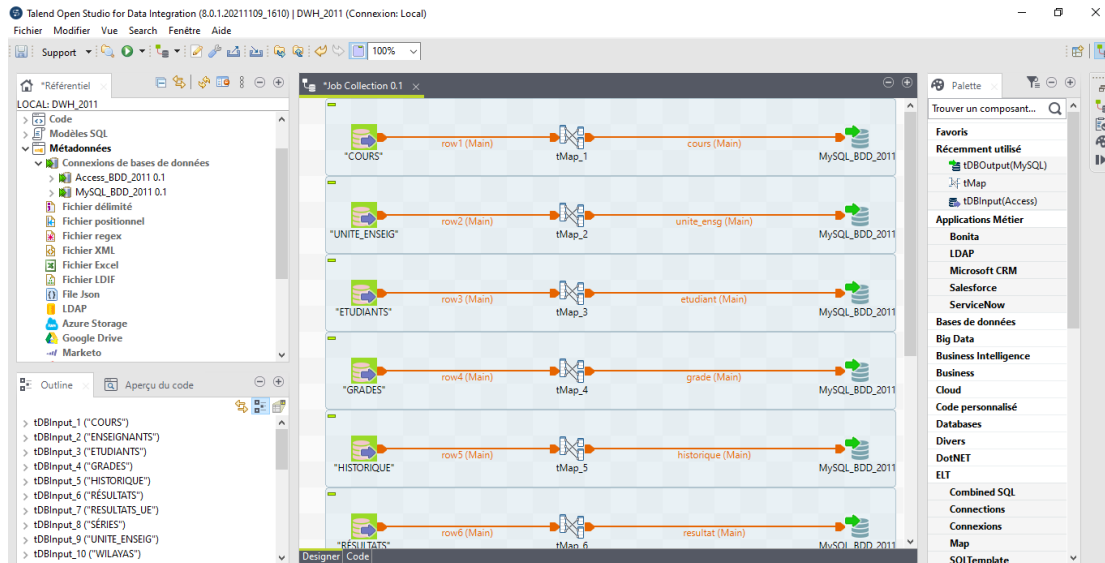
Ok

Annuler

11°C Eclaircies

7:00 PM

20/01/2023



1-Definition

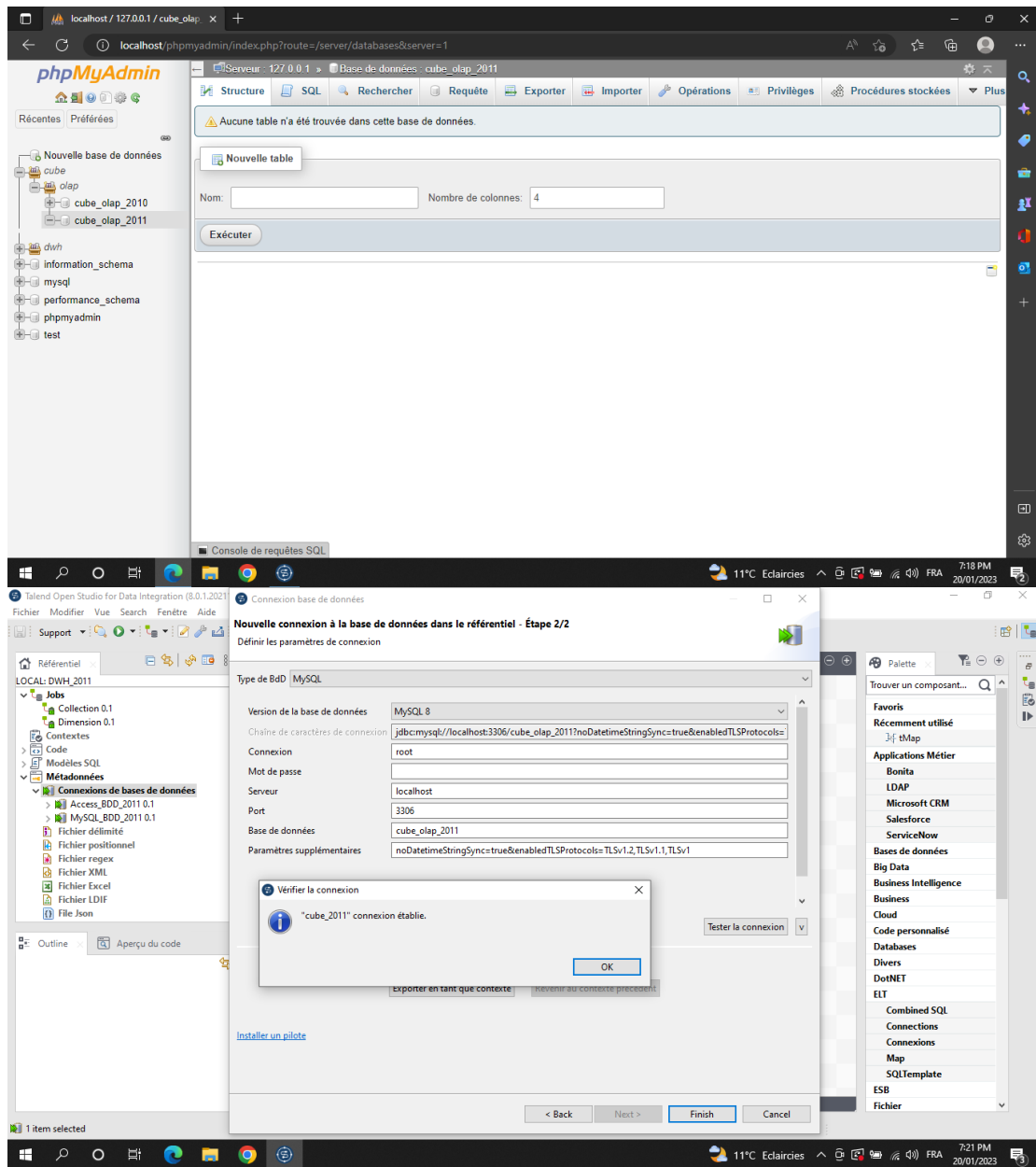
Dans le contexte de Talend Open Studio, une dimension est un ensemble de données structurées qui décrit une caractéristique particulière d'un ensemble de données plus large. Les dimensions sont souvent utilisées en conjonction avec des mesures pour permettre une analyse détaillée des données.

Les dimensions peuvent être utilisées pour créer des vues dimensionnelles de données, qui permettent d'analyser les données à différents niveaux de détail. Elles peuvent également être utilisées pour créer des sous-totaux, des totaux et des graphiques qui présentent les données d'une manière plus facile à comprendre.

Dans Talend Open Studio, vous pouvez utiliser des composants tels que "tMap" et "tAggregateRow" pour travailler avec des dimensions et des mesures. Vous pouvez également utiliser des fonctions de regroupement pour agréger les données à différents niveaux de détail et créer des vues dimensionnelles de données.

2- Les étapes pour travailler avec des dimensions dans Talend Open Studio :

- a-Préparation des données
- b-Définition des dimensions
- c-Agrégation des données
- d-Stockage des données agrégées
- e-Analyse des données



Talend Open Studio for Data Integration - th

Constructeur d'expression

Expression

☒ Wrap Undo(Ctrl + Z) Effacer

TalendDate.formatDate("MM",row1.DATE_NAISSANCE)

Test

Test Effacer

Var Vale

row1.NOINS	null
row1.DATE_NAI...	null
row1.SEXE	null
row1.ADRRESSE	null
row1.NO_BAC	null
row1.REDOUBL...	null
row1.NBRE	null
row1.ANNEE_B...	null
row1.N_SERIE...	null
row1.MENTIO...	null

Ajouter Supprimer

Catégories

- *Défini par l'utilisateur
- *Tous
- DataOperation
- Mathematical
- Numeric
- Relational
- StringHandling
- TalendDataGenerator
- TalendDate
- TalendString

Fonctions

Aide

Sélectionnez une catégorie et une fonction.

Ok Annuler

Appliquer

Ok Annuler

Éditeur de schéma Éditeur d'expression

row1

Colonne	Clé	Type
NOINS	<input checked="" type="checkbox"/>	String
DATE_NAISSANCE	<input type="checkbox"/>	Date
SEXE	<input type="checkbox"/>	String
ADRESSE	<input type="checkbox"/>	String

Talend Open Studio for Data Integration - thMap - thMap_1

Find :

Var

Auto map !

dem_etudiant

Expression

Column

row1.NOINS

TalendDate.formatDate("yyyy",row1.DATE...

TalendDate.formatDate("MM",row1.DATE...

TalendDate.formatDate("dd",row1.DATE...

NOINS

Annee

Mois

Jour

Éditeur de schéma Éditeur d'expression

row1

Colonne	Clé	Type	<input checked="" type="checkbox"/> N.	Modèle de da...	Longu...	Précision	Par d...	Comm...
NOINS	<input checked="" type="checkbox"/>	String	<input type="checkbox"/>		12	0		
DATE_NAISSANCE	<input type="checkbox"/>	Date	<input type="checkbox"/>	"dd-MM-yyyy"	19	0	'curr...	
SEXE	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		1	0		
ADRESSE	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		50	0		

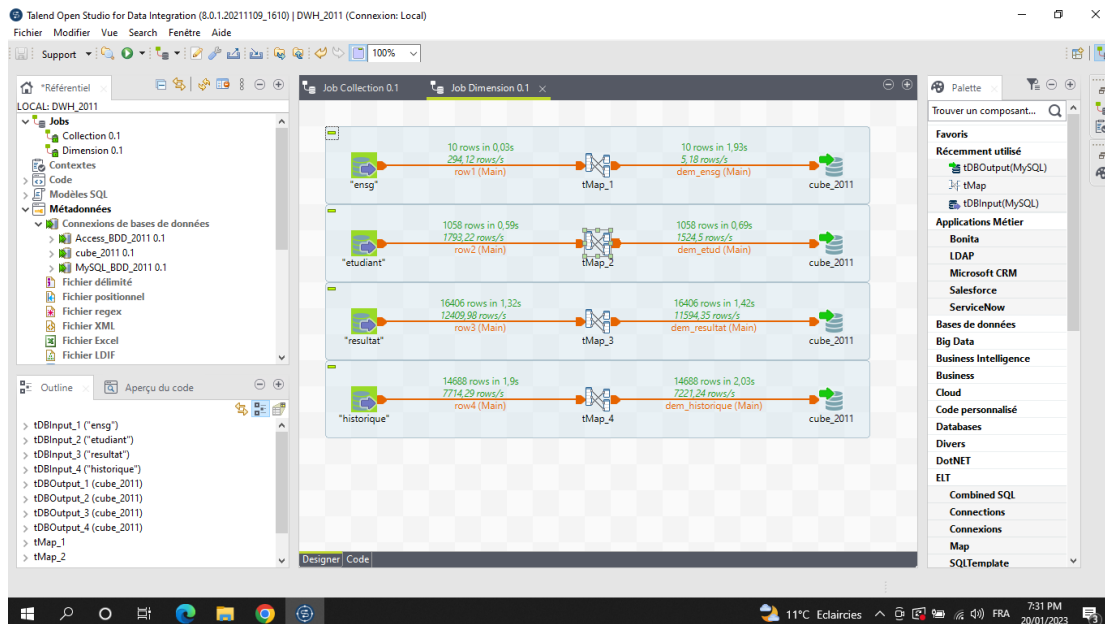
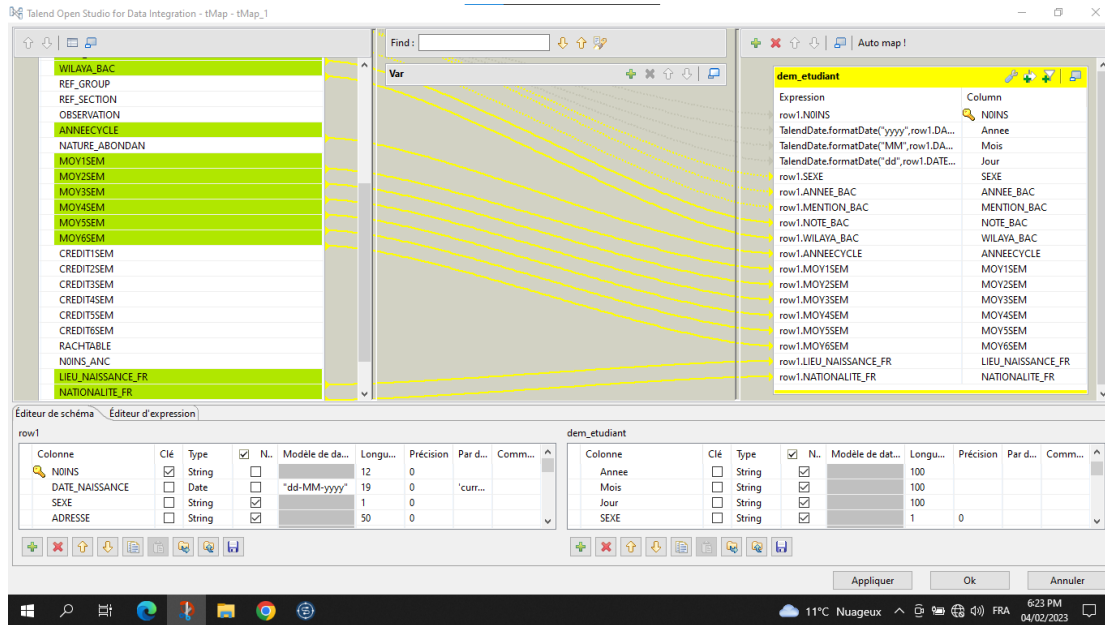
dem_etudiant

Colonne	Clé	Type	<input checked="" type="checkbox"/> N.	Modèle de dat...	Longue...	Précision	Par d...	Comm...
NOINS	<input checked="" type="checkbox"/>	String	<input type="checkbox"/>		12	0		
Annee	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>					
Mois	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>					
Jour	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>					

Appliquer

Ok Annuler

11°C Nuageux 6:08 PM 04/02/2023



Les faits

1-Définition

Les faits dans Talend Open Studio sont des mesures quantitatives qui peuvent être utilisées pour mesurer les résultats d'une entreprise ou d'un processus d'affaires. Ils peuvent être utilisés pour fournir des informations sur les ventes, les coûts, les profits, les stocks, etc. Les faits sont souvent utilisés en conjonction avec des dimensions pour fournir une vue plus complète et plus détaillée de l'entreprise ou du processus d'affaires.

Dans Talend Open Studio, les faits peuvent être extraits à partir de sources de données telles que des bases de données relationnelles, des fichiers CSV ou Excel, et peuvent être agrégés et stockés dans un format de sortie tel que CSV, Excel ou une base de données relationnelle. Les données de faits peuvent ensuite être analysées pour en tirer des conclusions et des insights en utilisant des outils tels que des graphiques et des tableaux croisés dynamiques.

1-Les étapes pour travailler avec des faits dans Talend Open Studio :

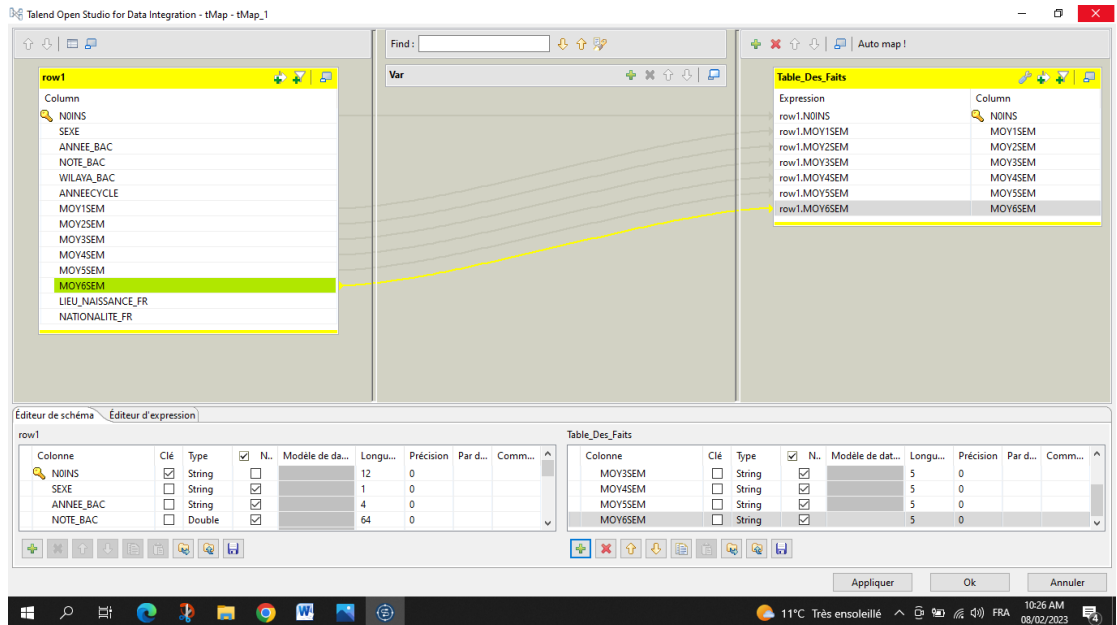
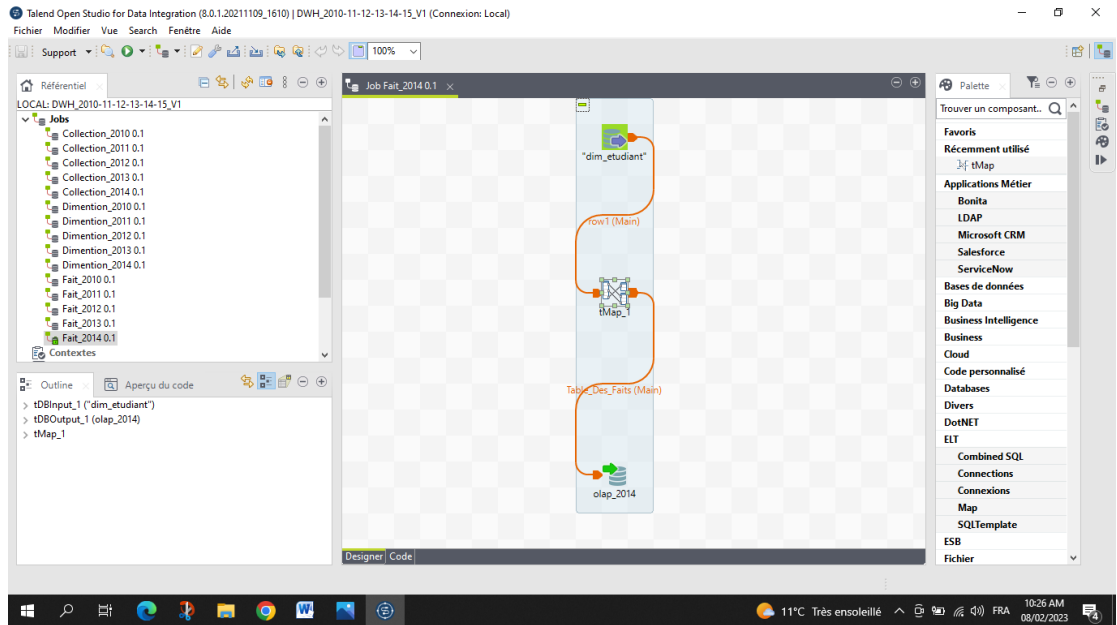
a-Préparation des données

b-Définition des faits

c-Agrégation des données

d-Stockage des données agrégées

e-Analyse des données



Conclusion

Talend Studio est un outil puissant et polyvalent pour les développeurs d'intégration de données. Il offre une grande variété de composants pour connecter, extraire, transformer et charger des données depuis diverses sources vers divers systèmes de destination. Talend Studio est facile à utiliser et peut être utilisé pour automatiser des tâches complexes d'intégration de données, telles que la fusion de données provenant de sources multiples, la transformation de données, le nettoyage de données et la synchronisation de données.

En outre, Talend Studio fournit des outils de qualité de données pour assurer la qualité et l'intégrité des données lors du processus d'intégration. De plus, il prend en charge une large gamme de systèmes d'exploitation, de bases de données et de formats de fichiers, ce qui le rend facile à utiliser pour les développeurs dans un large éventail d'industries.

En fin de compte, Talend Studio est un choix idéal pour les entreprises qui souhaitent optimiser leur processus d'intégration de données pour améliorer la qualité et la fiabilité de leurs informations commerciales.