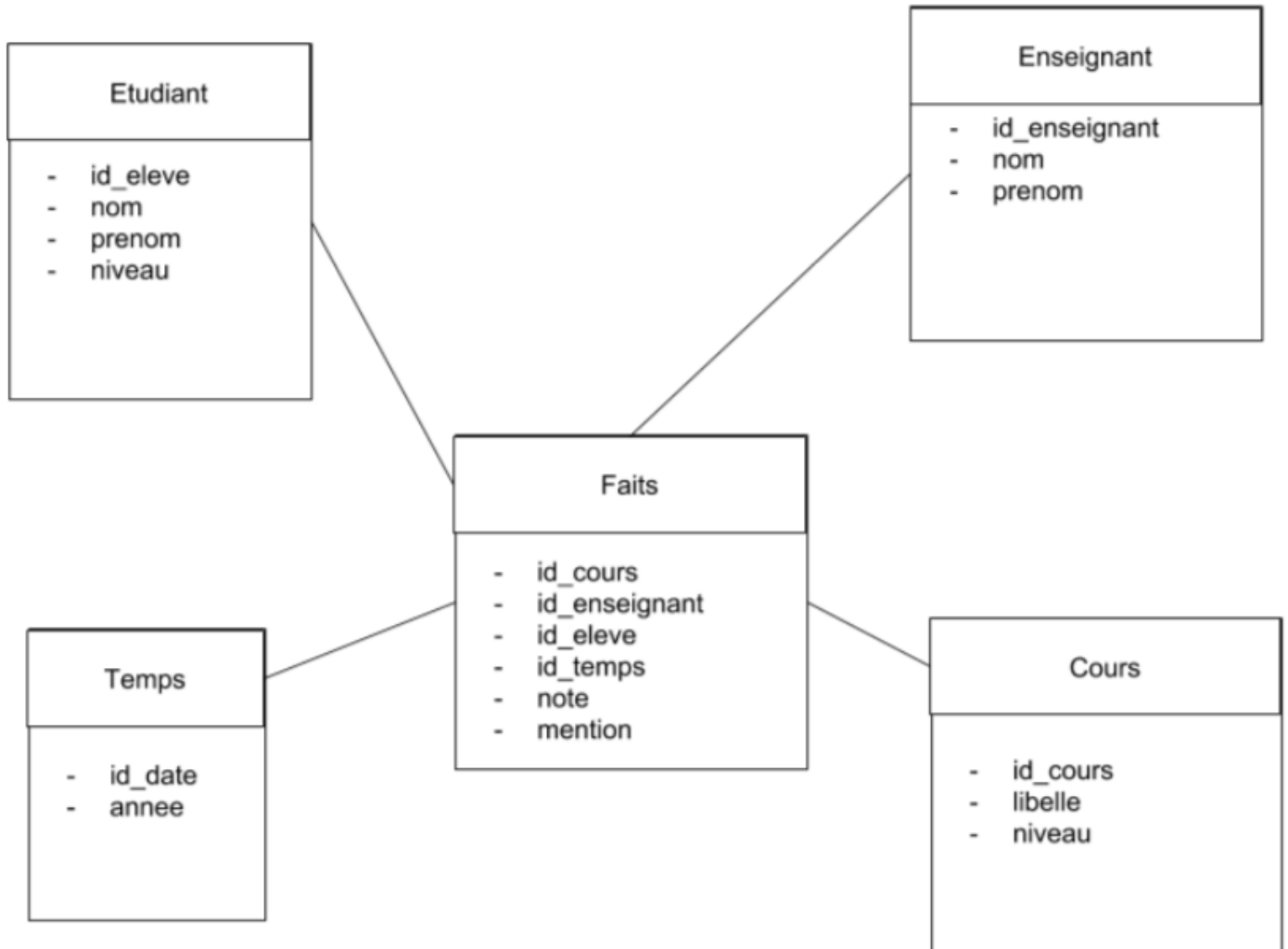


# TD Pentaho

## Partie 1 - Création d'entrepôt et ETL

### 1 - Schéma Conceptuel



Il y a donc 5 dimensions : Etudiants, Faits, Enseignant, Temps et Cours

Ces informations sont tirées du TD Médiateur réalisé plus tôt dans le cours

## Dimension Etudiant

Forme de la table:

Etudiant
<ul style="list-style-type: none"><li>- id_eleve</li><li>- nom</li><li>- prenom</li><li>- niveau</li></ul>

### Source XML

Dimension Etudiant (Etu)	Source 3: XML (S3) /BaseDeDonnee/Etudiants(XE)	Critère
Etu.id_etudiant: Type = Integer	XE.NumEt : Type = Byte	Conversion du type
Etu.nom: Type = String	XE.nom : Type = String	/
Etu.prenom: Type = String	'Source 3'	N'existe pas dans le schéma de la source 3: fichier
Etu.niveau: Type = String	XE.Niveau_insertion Type = String	/

Pour réaliser la table, il a juste fallu enlever les champs qui était en trop dans la source et renommer les champs.

### Source XLS

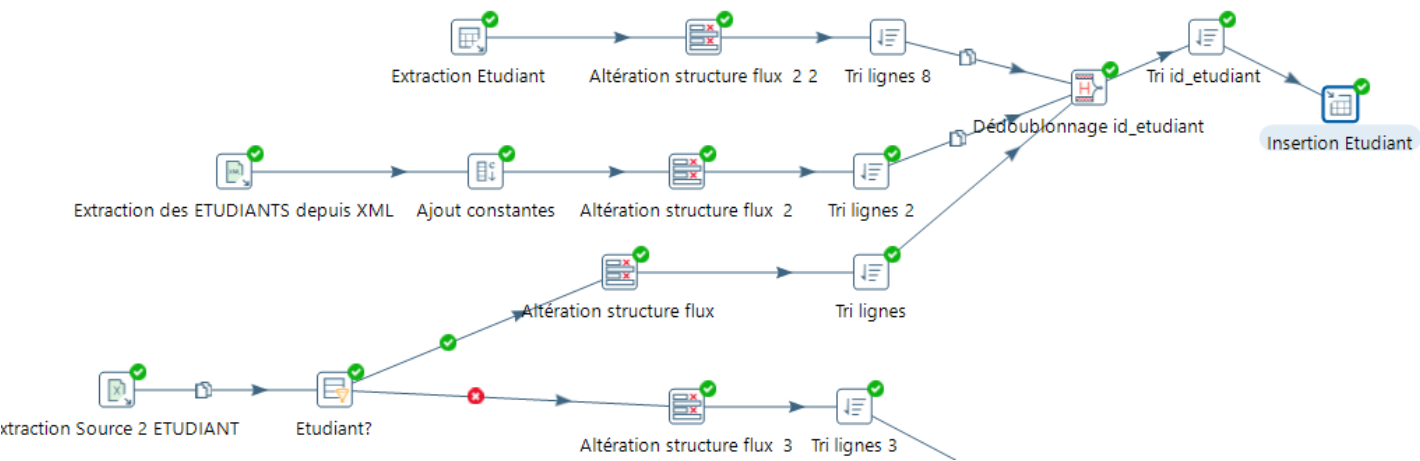
Dimension Etudiant (Etu)	Source 1: Excel (S1) Feuille Excel (FE)	Critère
Etu. id_etudiant: Type = Integer	FE.ID: Type=Integer	FE.Statut == etudiant
Etu.nom: Type = String	FE.Nom: Type = String	FE.Statut == etudiant
Etu.prenom: Type = String	Non présent dans la source	Non présent dans la source
Etu.niveau: Type = String	FE.NiveauInsertion: Type = String	FE.Statut == etudiant

Il a fallu combiner les 2 années qui était présentent sur 2 feuilles différentes, filtrer en fonction du statut puis il a fallu supprimer les champs qui ne concernait pas cette table (notamment à propos du cours) puis de trier et supprimer les doublons.

Source Oracle

Dimension Etudiant (Etu)	Source 2: Oracle (S2) Etudiant (BE)	Critère
Etu.id_etudiant: Type = Integer	BE.ID_Etudiant: Type = Integer	/
Etu.nom: Type = String	BE.nom: Type = varchar	Conversion du type
Etu.prenom: Type = String	BE.prenom: Type = varchar	Conversion du type

Fusion



## Dimension Enseignant

Forme de la table :

Enseignant
<ul style="list-style-type: none"><li>- id_enseignant</li><li>- nom</li><li>- prenom</li></ul>

### Source XML

Dimension Enseignant (Ens)	Source 3: XML (S3) /BaseDeDonnee/Enseignant(XEns)	Critère
Ens.id_enseignant: Type = Integer	XEns.NumEns: Type = Byte	Conversion du type
Ens.nom: Type = String	XEns.Nom: Type = String	/
Ens.prenom: Type = String	XEns.Prenom: Type = String	/

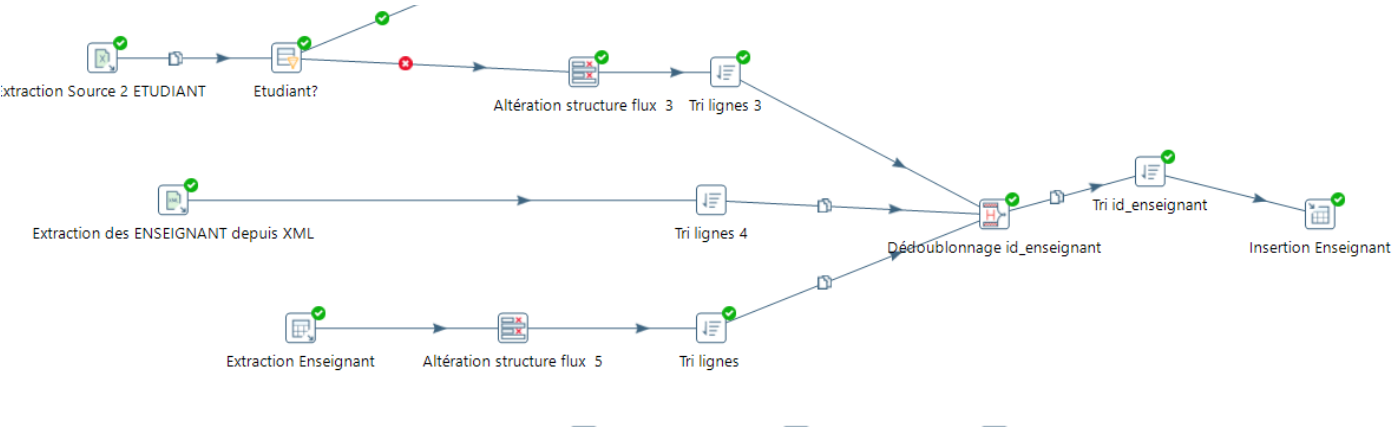
### Source XLS

Dimension Enseignant (Ens)	Source 1: Excel (S1) Feuilles Excel (FE)	Critère
Ens.id_enseignant: Type = Integer	FE.ID: Type = Integer	FE.Statut == enseignant
Ens.nom: Type = String	FE.Nom Type = String	FE.Statut == enseignant
Ens.prenom: Type = String	FE.Prenom Type = String	FE.Statut == enseignant

Source Oracle

Dimension Enseignant (Ens)	Source 2: Oracle (S2) Enseignant (BEns)	Critère
Ens.id_enseignant: Type = Integer	BEns.ID_ens: Type = Integer	/
Ens.nom: Type = String	BEns.nom: Type = varchar	Conversion du type
Ens.prenom: Type = String	BEns.prenom: Type = varchar	Conversion du type

Fusion



## Dimension Cours

Forme de la table :

Cours
<ul style="list-style-type: none"><li>- id_cours</li><li>- libelle</li><li>- niveau</li></ul>

### Source XML

Dimension Cours (Co)	Source 3: XML (S3) /BaseDeDonnee/Cours(XC)	Critère
Co.id_cours: Type = Integer	XC.ID_cours: Type = byte	Conversion du type
Co.libelle: Type = String	N'existe pas dans le schéma de la source 3: fichier XML	/
Co.niveau: Type = String	XC.Niveau: Type = String	/

Pour réaliser l'intégration de cette partie, il a fallu créer une constante String libellé pour que la forme de la base soit respectée.

### Source XLS

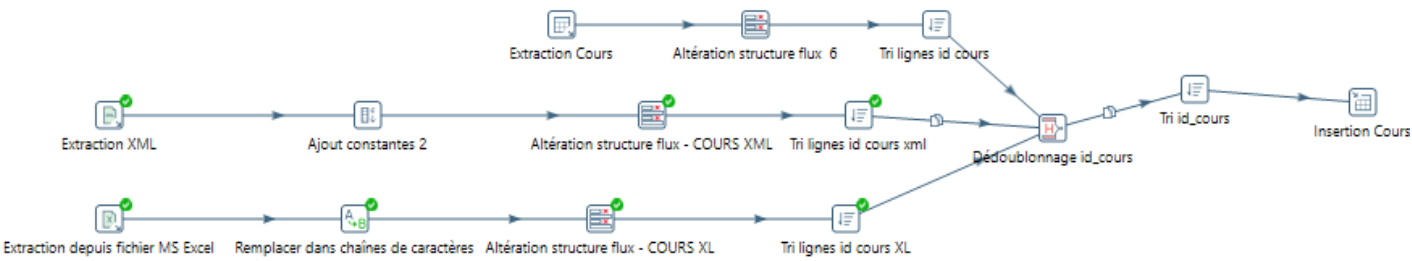
Dimension Cours (Co)	Source 1: Excel (S1) Feuilles Excel (FE)	Critère
Co.id_cours: Type = Integer	FE.ID_Cours: Type = Integer	/
Co.libelle: Type = String	FE.Libelle_Cours: Type = String	/
Co.niveau: Type = String	FE.Niveau_Cours: Type = String	/

Pour plus de simplicité, toutes les sources XLS proviennent du même alimentateur, il y a juste eu des suppressions de champs pour cette partie.

Source Oracle

Dimension Cours (Co)	Source 2: Oracle (S2) Cours (BC)	Critère
Co.id_cours: Type = Integer	BC.numCours: Type = Integer	/
Co.libelle: Type = String	BC.libele: Type = varchar	Conversion du type
Co.niveau: Type = String	BC.niveau: Type = varchar	Conversion du type

Fusion



## Dimension Temps

Forme de la table :

Temps
<ul style="list-style-type: none"><li>- id_date</li><li>- annee</li></ul>

### Source XML

Dimension Temps (Temps)	Source 3: XML (S3) /BaseDeDonnee/Inscription(XI)	Critère
Temps.id_temps: Type = Integer	N'existe pas	/
Temps.annee: Type = Integer	XI.Annee_universitaire Type = short	Conversion du type

### Source XLS

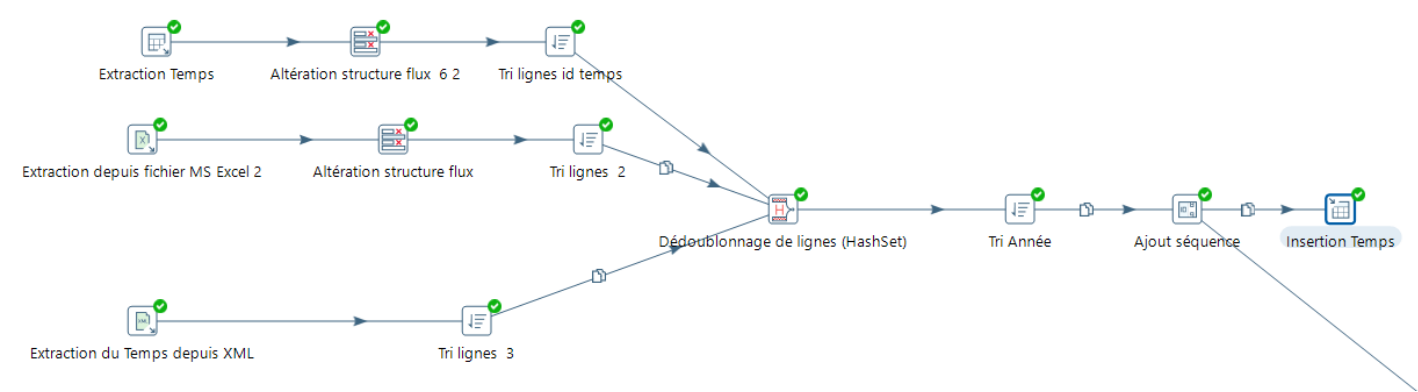
Dimension Temps (Temps)	Source 1: Excel (S1) Feuilles Excel (FE)	Critère
Temps.id_temps: Type = Integer	N'existe pas	/
Temps.annee: Type = Integer	Nom de la feuille S1	Récupérer le nom de la feuille

### Source Oracle

Dimension Temps (Temps)	Source 2: Oracle (S2) Inscription(BI)	Critère
Temps.id_temps: Type = Integer	N'existe pas	/
Temps.annee: Type = Integer	BI.Annee: Type = Integer	/



Fusion



## Dimension Fait

Forme de la table :

Faits
<ul style="list-style-type: none"><li>- id_cours</li><li>- id_enseignant</li><li>- id_eleve</li><li>- id_temps</li><li>- note</li><li>- mention</li></ul>

### Source XML

Dimension Fait (Fait)	Source 3: XML (S3) /BaseDeDonnee/Cours (XC) /BaseDeDonnee/Enseignant (XEns) /BaseDeDonnee/Etudiants(XE) /BaseDeDonnee/Inscription(XI)	Critère
Fait.id_cours: Type = Integer	XC.ID_cours: Type = byte	Conversion du type
Fait.id_enseignant: Type = Integer	XEns.NumEns: Type = Byte	Conversion du type
Fait.id_eleve: Type = Integer	XE.NumEt : Type = Byte	Conversion du type
Fait.id_temps: Type = Integer	N'existe pas	Faire une jointure avec la sortie de la dimension temps pour récupérer l'id de XI.Annee_universitaire
Fait.note Type = Number	XI.Note_cours Type = byte	Conversion du type
Fait.mention Type = String	Non présent dans la source	Note<12 : NULL 12<=Note<14 : AB 14<= Note <16 : B Note>=16 : TB

### Source XLS

Dimension Fait (Fait)	Source 1: Excel (S1) Feuille Excel (FE)	Critère
Fait.id_cours: Type = Integer	FE.ID_Cours: Type = Integer	/
Fait.id_enseignant: Type = Integer	FE.ID: Type = Integer	Statut= enseignant Faire une jointure avec l'id du

		cours pour avoir l'id
Fait.id_eleve: Type = Integer	FE.ID: Type=Integer	Statut= etudiant
Fait.id_temps: Type = Integer	N'existe pas	Faire une jointure avec la sortie de la dimension temps pour récupérer l'id correspondant au nom de la feuille
Fait.note Type = Number	FE.Note: Type = Integer	/
Fait.mention Type = String	Non présent dans la source	Note<12 : NULL 12<=Note<14 : AB 14<= Note <16 : B Note>=16 : TB

### Source Oracle

Dimension Fait (Fait)	Source 2: Oracle (S2) Cours (BC) Enseignant (BEns) Etudiant (BE) Inscription(BI)	Critère
Fait.id_cours: Type = Integer	BC.numCours: Type = Integer	/
Fait.id_enseignant: Type = Integer	BEns.ID_ens: Type = Integer	/
Fait.id_eleve: Type = Integer	BE.ID_Etudiant: Type = Integer	/
Fait.id_temps: Type = Integer	N'existe pas	Faire une jointure avec la sortie de la dimension temps pour récupérer l'id de BI.Annee
Fait.note Type = Number	BI.Note_cours: Type = Integer	/
Fait.mention Type = String	Non présent dans la source	Note<12 : NULL 12<=Note<14 : AB 14<= Note <16 : B Note>=16 : TB

### Fusion

Il a bien fallu à faire attention à ce que l'id année soit le même que la table Temps et que, dans le cas où il y a un enseignant qui enseigne un cours dans une BDD, de faire en sorte que cet enseignant enseigne le cours lorsqu'il y a une jointure.

Date

