

I. Fonction

La fonction **tUniqRow** dans Talend Open Studio for Data Integration est utilisée pour éliminer les doublons dans un flux de données. Elle est souvent utilisée dans des scénarios où des données dupliquées peuvent causer des problèmes, tels que la fusion ou la comparaison de fichiers. Le fonctionnement de **tUniqRow** est assez simple. Elle prend en entrée un flux de données avec une ou plusieurs colonnes qui peuvent contenir des doublons. Elle compare ensuite chaque ligne de données avec toutes les lignes précédentes dans le flux pour détecter les doublons. Si un doublon est détecté, **tUniqRow** peut soit le supprimer, soit le renvoyer dans une sortie distincte, en fonction de la configuration choisie.

La fonction **tjoin** dans Talend Open Studio for Data Integration est utilisée pour joindre deux ou plusieurs flux de données en fonction d'une clé commune. Elle permet de combiner les enregistrements de deux tables ou fichiers différents en un seul ensemble de résultats. Le fonctionnement de **tjoin** est le suivant :

- Les flux de données à joindre sont définis en entrée.
- La clé de jointure est spécifiée, qui permettra de faire correspondre les enregistrements des deux flux.
- Les enregistrements correspondants sont combinés en un seul ensemble de résultats.
- Les options de jointure (inner join, left outer join, right outer join, full outer join) sont spécifiées pour déterminer comment les enregistrements non correspondants seront traités.
- Le résultat final est renvoyé en sortie

La fonction **tFilterRow** dans Talend Open Studio for Data Integration permet de filtrer les enregistrements d'un flux de données en fonction de critères spécifiques. Elle permet de sélectionner uniquement les enregistrements qui répondent à ces critères et de supprimer les autres. Le fonctionnement de **tFilterRow** est le suivant :

- Le flux de données à filtrer est défini en entrée.

- Les critères de filtrage sont spécifiés, qui permettront de déterminer quels enregistrements seront sélectionnés.
- Les enregistrements qui répondent aux critères de filtrage sont sélectionnés et renvoyés en sortie.
- Les enregistrements qui ne répondent pas aux critères de filtrage sont supprimés

La fonction **tFilterColumns** dans Talend Open Studio for Data Integration permet de sélectionner les colonnes d'un flux de données à conserver ou à supprimer en fonction de critères spécifiques. Elle permet ainsi de réduire la taille du flux de données et de ne conserver que les colonnes pertinentes pour une analyse ou un traitement ultérieur. Le fonctionnement de **tFilterColumns** est le suivant :

- Le flux de données à filtrer est défini en entrée.
- Les critères de filtrage sont spécifiés, qui permettront de déterminer quelles colonnes seront sélectionnées ou supprimées.
- Les colonnes qui répondent aux critères de filtrage sont sélectionnées et renvoyées en sortie.
- Les colonnes qui ne répondent pas aux critères de filtrage sont supprimées

La fonction **tConvertType** dans Talend Open Studio for Data Integration est une fonction utilisée pour convertir un type de données en un autre type de données. Grâce à cette fonction, vous pouvez convertir différentes données dans différents formats. Par exemple, vous pouvez convertir une chaîne de caractères en un nombre entier ou un nombre à virgule flottante (float), ou convertir une chaîne de caractères représentant une date en un type de données de date.numeric.

II. Fonction de la classe Numeric :

1. La fonction "**convertImplicit**" de la catégorie "numeric" dans Talend Open Studio est utilisée pour convertir les types de données numériques implicites en types de données explicites dans une chaîne ou un flux de

données. Les types de données implicites peuvent être des nombres entiers ou décimaux stockés sous forme de chaînes plutôt que de types de données numériques. Cette fonction accepte deux entrées, la première étant la chaîne ou le champ que vous souhaitez convertir et la seconde étant le type de données explicite que vous voulez utiliser. on peut choisir différents types de données numériques, tels que "Integer", "Double" ou "Float", en fonction de vos besoins.

2. La fonction **random** de la catégorie numeric dans Talend Open Studio for Data Integration est utilisée pour générer des nombres aléatoires dans une plage de valeurs spécifiée. Elle peut être utile pour diverses tâches telles que la génération de données de test, la création de clés primaires aléatoires pour les tables de base de données, etc.
3. La fonction **removeSequence** de la catégorie numeric dans Talend Open Studio for Data Integration est utilisée pour supprimer une séquence de nombres dans une chaîne de caractères. Par exemple, si vous avez une chaîne de caractères contenant "12345abc", la fonction **removeSequence** peut être utilisée pour supprimer la séquence de nombres "12345" pour obtenir la chaîne "abc"
4. La fonction "**resetSequence()**" dans la catégorie "numeric" de Talend Open Studio for Data Integration est utilisée pour réinitialiser une séquence numérique. Cette fonction prend en entrée deux paramètres : le premier est le nom de la séquence que vous souhaitez réinitialiser et le second est la valeur à laquelle vous voulez réinitialiser la séquence.
5. La fonction "**Sequence**" dans Talend Open Studio for Data Integration est utilisée pour générer une séquence de nombres dans un flux de données. Elle peut être utilisée pour attribuer des identifiants uniques à chaque ligne d'un ensemble de données ou pour générer des numéros de commande ou de facture. La fonction prend deux paramètres en entrée : le premier est le numéro de départ de la séquence et le second est l'incrément à ajouter à chaque nombre de la séquence