

Talend Initiation

Talend OS for Data Intégration

Features

1. Jobs
2. Components
3. Metadata
4. Running jobs
5. Manipulating data using Tmap Component

TP - Pays et disponibilité énergétique

- 1- Lire le fichier de la dispo de gaz et afficher les résultats dans la console
- 2- Ajouter une nouvelle colonne qui affiche le total (reserves + exportations - importations) de la dispo énergétique
- 3- Ajouter une colonne qui affiche le statut des pays : “Vert” => total >200000 sinon “Rouge”
- 4- Trier les pays par statut “descending” (vert first)
- 5- Filtrer les pays “Rouge” dans un tableau différent
- 6- Exporter les tableaux générés dans 2 fichiers différents
- 7- Exporter les pays au statut “Rouge” dans une base de données MySQL

Stack Technique : tFileInputDelimited, tMap, tlogRow, tFileOutputDelimited, tMySQLOutputDelemited

TP - Tennis et grand chelems

- 1- Récupérer les fichiers player et player_grandslams du repo <https://github.com/dovene/talend-resources>
- 2- Fusionner et afficher toutes les données des 2 fichiers en un seul tableau
- 3- Stocker le résultat de la fusion dans un autre fichier
- 4- Importer les informations des 2 fichiers d'entrée dans une base MySQL
- 5- Reprendre le point 2 et 3 avec les input de base de données
- 6- Dupliquer le fichier des trophées(player_grandslams)
- 7- Concaténer le contenu des deux fichiers dans un autre fichier

Stack Technique et thèmes : tFileInputDelimited, tMap, tlogRow, tFileOutputDelimited, tMySQLOutputDelemited, tMySQLInputDelemited, tFileList

Thèmes : Base de données, Jointure, Concaténation, Fichiers

TP - Variables de contexte et Environnement

- 1- Récupérer le fichier country_population <https://github.com/dovene/talend-resources>
- 2- Dupliquer le fichier en 2 exemplaires, l'un avec le suffixe dev et l'autre prod
- 3 - Pour distinguer le contenu des 2 fichiers, garder uniquement 3 lignes dans le fichier dev en suffisant les nom de pays par dev
- 4- Créer 2 bases de données dans MySQL l'un nommé dev et l'autre prod
- 5- Créer un job qui permet selon le contexte d'importer le contenu du fichier dev/prod dans la BD correspondante
 - 5.1- Créer les différents contextes (dev et prod)
 - 5.2- Créer les variables représentant les noms des fichiers et les nom des bases de données
- 6- Mettre en oeuvre les différents composants en utilisant les noms de variables créées en remplacement des valeurs fixes.

Stack Technique et thèmes : tFileInputDelimited, tlogRow, tMySQLOutputDelemited, contexte, tMySQLConnection, tPreJob

Thèmes : Contexte, Environnement

TP - Pays et disponibilité énergétique

3- Ajouter une colonne qui affiche le statut des pays : “Vert” => total >200000
sinon “Rouge”

- new column with expression : (Var.var1 < 200000) ? "Rouge" : "Vert"

4- Trier les pays par statut ascendant

5- Filtrer les pays “Rouge” dans un tableau différent

6- Exporter les tableaux générés dans 2 fichiers différents

TP - Pays et disponibilité énergétique

1- Lire le fichier de la dispo de gaz et afficher les résultats dans la console

- Create new input file delimited from metadata section with file https://github.com/dovene/talend-resources/blob/main/pays_dispo_gaz.csv
- Create new job named dispo_gaz
- Drop previously created metadata as tFileInputDelimited to job layout
- Link (Right click main) the tFileInputDelimited to a tLogRow component
- Run the job

2- Ajouter une nouvelle colonne qui affiche le total (reserves + exportations - importations) de la dispo énergétique

- Create new tMap component on the previous link between input and logRow
- Setup (double click) tMap => add new column total in output (bottom right) section
- In column expression of the new colum add : `row1.Importations - row1.Exportations + row1.Reserves`
- Apply and Run the job

TP - Pays et disponibilité énergétique

1- Lire le fichier de la dispo de gaz et afficher les résultats dans la console

- Create new input file delimited from metadata section with file https://github.com/dovene/talend-resources/blob/main/pays_dispo_gaz.csv
- Create new job named dispo_gaz
- Drop previously created metadata as tFileInputDelimited to job layout
- Link (Right click main) the tFileInputDelimited to a tLogRow component
- Run the job

2- Ajouter une nouvelle colonne qui affiche le total (reserves + exportations - importations) de la dispo énergétique

- Create new tMap component on the previous link between input and logRow
- Setup (double click) tMap => add new column total in output (bottom right) section
- In column expression of the new colum add : `row1.Importations - row1.Exportations + row1.Reserves`
- Apply and Run the job

TP - Webservices

1- GET request - Retrieve countries with this web service : <https://countriesnow.space/api/v0.1/countries/iso>

- Add a new tRest component
- Fill the component details using URL Field and others
- Add a tExtractJSONFields component to parse Json result
- Configure tExtractJSONFields : Select Body for JSON Field, JSONPath = "\$.data[*]", Set Mapping => add output column in schema and "name" as JSON key
- Output Result with a tLog then Run the job

Do the same using a tRestClient :

- Unselect "Convert to Dom" in Advanced params for tRestClient
- Configure tExtractJSONFields : Select string for JSON Field

2- GET Request with query parameters Retrieve

- Use query parameters attribute of tRESTCient to set parameters

Stack Technique et thèmes : tRestClient, tRest, tExtractJSONFields

TP -Webservices