

## TP4 - Initiation à TALEND –

### Objectifs du TP

- Découvrir l'interface [Talend Open Studio](#) ( faut installer avec le [jdk22](#) )
- Comprendre les concepts ETL (Extract, Transform, Load) dans Talend
- Maîtriser les composants de base : tFileInputDelimited, tMap, tFileOutputDelimited et autres
- Gérer les schémas de données et les transformations

### Exercice Pratique 01: Système de calcul de revenus

#### Contexte

Vous êtes chargé de créer un job Talend qui :

1. Lit un fichier CSV de ventes
2. Calcule le revenu pour chaque transaction
3. Exporte les résultats dans un nouveau fichier

#### Données d'Entrée fournies (ventes.csv)

### Étapes Détaillées

#### Étape 1 : Création du Job

- Nouveau job "TP1\_Ventes"
- Description : "Calcul des revenus par produit"

#### Étape 2 : Configuration tFileInputDelimited

- **Composant** : tFileInputDelimited\_1
- **Fichier** : [ventes.csv](#)
- **Encodage** : ISO-8859-15
- **Séparateur** : ;
- **En-tête** : 1 ligne

- **Schéma :**

- Date (String)
- Produit (String)
- Quantité (Integer)
- Prix\_Unitaire (Float)

**Étape 3 : Transformation tMap**

- **Expression :** row1.Quantite \* row1.Prix\_Unitaire
- **Schéma de sortie :**
  - Date (String)
  - Produit (String)
  - Revenu (Float)

**Étape 4 : Export tFileOutputDelimited**

- **Fichier :** resultats\_ventes.csv
- **Séparateur :** ;
- **Inclure l'en-tête :** ✓

**Questions Pratiques**

1. Modifiez le job pour ajouter une colonne "TVA" (20%)
2. Créez un filtre pour n'exporter que les ventes > 500€
3. Analyses variées : produit haut de gamme **vs** moyenne gamme **vs** accessoires

≥ 300€ : "Haut de Gamme"

100€ à 299€ : "Milieu de Gamme"

< 100€ : "Accessoire"