

# TP3

## TP 3.1 : Nettoyage initial des données avec Pandas

### Objectifs :

- Utiliser les principales fonctionnalités de Pandas pour manipuler et analyser des données
- Appliquer les notions de nettoyage de données à des données réelles

### Énoncé

Vous êtes chargé de nettoyer les données d'un fichier CSV contenant les informations suivantes sur les ressources de blé dans le monde :

- La région
- L'année
- La production de blé (en millions de tonnes)

Les données sont disponibles dans le fichier `ressources_ble.csv`.

- Charger le fichier CSV dans un DataFrame Pandas
- Vérifier la cohérence des données
- Remplacer les valeurs manquantes
- Corriger les erreurs de saisie
- Normaliser les données

### Résultats attendus :

Vous devez produire un document expliquant les résultats de votre nettoyage. Ce document doit inclure les éléments suivants :

- Une introduction qui présente les objectifs du nettoyage
- Une description des données utilisées
- Les résultats du nettoyage, avec des explications et des commentaires
- Une conclusion qui résume les principaux résultats

## TP 3.2 : Mise en œuvre de Z-score, DBSCAN et isolation forest avec Pandas, NumPy et SciPy

### Objectifs :

- Utiliser les principales fonctionnalités de Pandas, NumPy et SciPy pour manipuler et analyser des données
- Appliquer les notions de Z-score, DBSCAN et isolation forest à des données réelles

### Énoncé :

Vous êtes chargé de mettre en œuvre les algorithmes de détection d'anomalies Z-score, DBSCAN et isolation forest sur les données d'un fichier CSV contenant les informations suivantes sur les ressources de blé dans le monde :

- La région
- L'année
- La production de blé (en millions de tonnes)

Les données sont disponibles dans le fichier `ressources_ble.csv`.

- Charger le fichier CSV dans un DataFrame Pandas
- Calculer les Z-scores des données
- Appliquer l'algorithme DBSCAN aux données
- Appliquer l'algorithme isolation forest aux données

### Résultats attendus :

Vous devez produire un document expliquant les résultats de votre mise en œuvre. Ce document doit inclure les éléments suivants :

- Une introduction qui présente les objectifs de la mise en œuvre
- Une description des données utilisées
- Les résultats de la mise en œuvre, avec des explications et des commentaires
- Une conclusion qui résume les principaux résultats