

# Project DALAS

## Partie 1 : Définition du problème

Mathilde GAUTEUR (21500128) – Xinyi LIANG (xxxxxxx)

Septembre 2025

### Problématique :

Les recettes végétariennes ont-elles un impact écologique moindre que les recettes à base de viande en France ?

### Contexte :

La consommation de viande est régulièrement citée comme une source importante d'émissions de CO<sub>2</sub>. Dans un contexte où la communauté végétarienne est grandissante et où les enjeux climatiques se retrouvent au cœur de tous les débats, comparer l'empreinte carbone des recettes végétariennes et carnées apparaît pertinent et socialement fort.

### Données :

- **Recettes** filtrées par catégorie (végé vs non-végé) et par popularité (likes, commentaires) pour obtenir les plus réalisées qui ont donc un impact : site Marmiton.
- **Empreinte carbone** : associée à chaque ingrédient : base Agribalyse.

### Analyses envisagées :

- **Exploratoires** : visualisations, statistiques descriptives.
- **Comparaison** : moyenne des émissions CO<sub>2</sub> végé vs non-végé, tests de significativité.
- **Clustering** : regroupement des recettes selon leur empreinte carbone et composition en ingrédients (PCA, t-SNE).
- **Modèles prédictifs** :
  - *Régression linéaire / Ridge / Lasso* : adaptées pour estimer l'empreinte carbone totale en fonction des ingrédients, tout en gardant un modèle simple et interprétable.
  - *Forêts aléatoires (Random Forest)* : utiles pour capturer les interactions complexes entre ingrédients et émissions, là où les modèles linéaires pourraient être limités.
  - Encodage des ingrédients en variables binaires (one-hot) pour transformer chaque recette en vecteur d'ingrédients.

Ces modèles permettent à la fois d'évaluer les facteurs les plus polluants et de prédire l'impact d'une nouvelle recette.

- (Optionnel) **Recommandations de substitution** : suggérer des alternatives moins émettrices pour certains ingrédients.

### Étapes techniques :

1. **Scraping** des données Marmiton et Agribalyse (Python, BeautifulSoup/Selenium).
2. **Nettoyage** : normalisation des ingrédients, classification végé/non-végé.
3. **Analyse** : calcul des émissions totales, visualisations, modèles prédictifs.

### Résultats attendus :

Les recettes végétariennes devraient présenter une empreinte carbone moyenne significativement plus faible que celles à base de viande.