**Présentation du projet.**

Repo GitHub : https://github.com/Evgeni1494/bloc3/tree/main/Data\_expose\_api\_E1

#### **Contexte et Acteurs**

Ce projet vise à extraire, traiter et enrichir les données relatives aux émissions de gaz à effet de serre par état aux États-Unis. L'objectif principal est de fournir une base de données géolocalisée qui permet d'analyser les émissions par secteur et type de carburant. Les principaux acteurs impliqués dans ce projet incluent :

* Développeurs : Responsables de la mise en œuvre du script pour l'extraction et le traitement des données.
* Analystes de données : Utilisateurs finaux des données traitées pour des analyses et des visualisations.
* Administrateurs de bases de données : En charge de la gestion et de la maintenance de la base de données SQLite.

#### Objectifs Fonctionnels et Techniques

* Objectifs Fonctionnels :
  + Extraire les données géographiques des principales villes de chaque état américain.
  + Enrichir les données d'émissions de gaz à effet de serre avec des coordonnées géographiques.
  + Sauvegarder les données traitées dans une base de données SQLite pour une utilisation et une analyse ultérieures.
* Objectifs Techniques :
  + Utiliser l'API de géocodage d'OpenWeatherMap pour obtenir les coordonnées géographiques.
  + Assurer une intégration fluide des données géographiques et des données d'émissions.
  + Automatiser le processus de mise à jour des données et de leur sauvegarde dans une base de données.

#### Environnements et Contraintes Techniques

* Environnement de Développement :
  + Python 3.11 pour le développement du script.
  + Pandas pour la manipulation et le traitement des données.
  + SQLite pour le stockage des données traitées.
  + API d'OpenWeatherMap pour obtenir les coordonnées géographiques.
* Contraintes Techniques :
  + Gestion des erreurs et des exceptions pour assurer la robustesse du script.
  + Disponibilité des données de l'API d'OpenWeatherMap.
  + Performance du traitement des données pour gérer de grands volumes de données.

#### Budget

Le budget de ce projet inclut les coûts suivants :

* Ressources humaines : Temps des développeurs et des analystes de données.
* Infrastructure : Serveurs ou environnements de développement nécessaires pour exécuter et tester le script.

#### Organisation du Travail et Planification

* Étape 1 : Définition des besoins et spécifications techniques.
* Étape 2 : Développement du script pour l'extraction des coordonnées géographiques
* Étape 3 : Intégration et traitement des données d'émissions.
* Étape 4 : Mise en place de la base de données SQLite et insertion des données.
* Étape 5 : Tests et validation.
* Étape 6 : Documentation et déploiement.

### Spécifications Techniques

#### Technologies et Outils

* Langage de Programmation : Python 3.11
* Bibliothèques :
  + requests pour les appels à l'API d'OpenWeatherMap.
  + pandas pour la manipulation des données.
  + sqlite3 pour l'interaction avec la base de données SQLite.
* Outils :
  + CSV pour le stockage intermédiaire des données.
  + SQLite pour la base de données finale.

#### Services Externes

* OpenWeatherMap Geocoding API : Utilisée pour obtenir les coordonnées géographiques des villes principales de chaque état.

#### Exigences de Programmation

* Gestion des Erreurs : Implémenter des blocs try-except pour gérer les erreurs potentielles lors des appels API et du traitement des données.
* Automatisation : Le script doit être capable de mettre à jour la BDD en utilisant l’API OpenWeather et en agrégeant les données avec celles du fichier csv.

#### Accessibilité

* Disponibilité : Le script et la base de données doit étre fonctionnel et mis a disposition de la personne chargé de maintenance de la BDD.
* Accès : Les utilisateurs doivent avoir un accès aux données via la base de données SQLite.

### Périmètre des Spécifications Techniques

#### Extraction des Données

* Données d'Émissions : Extraire les données d'émissions de gaz à effet de serre à partir d'un fichier CSV.
* Données Géographiques : Utiliser l'API d'OpenWeatherMap pour obtenir les coordonnées géographiques des principales villes de chaque état.

#### Agrégation des Données

* Enrichissement : Ajouter les coordonnées géographiques aux données d'émissions.

#### Stockage des Données

* CSV : Stockage intermédiaire des données enrichies dans un fichier CSV.
* SQLite : Sauvegarde finale des données dans une base de données SQLite, avec création de tables pour les états et les émissions.