



# Collecte automatisé de base de données SIRENE



## Contexte

Dans ce projet, j'ai enrichi la base de données SIRENE avec des coordonnées géographiques et des détails supplémentaires sur les établissements en utilisant les données de Google Maps.

Objectif était d'enrichir la base de données SIRENE avec des informations géographiques et des détails complémentaires pour améliorer sa précision et sa valeur informative.

Ce projet a été réalisé en binôme sur une période de 2 mois.

## La démarche suivie

### Nettoyage de la base des données

J'ai utilisé l'outil Python et la bibliothèque pandas pour importer, nettoyer, et transformer les données du fichier. L'objectif était d'assurer la qualité et la lisibilité des données avant leur enrichissement.

### Collecte des données

J'ai utilisé l'API pour obtenir des coordonnées géographiques précises. Puis j'ai intégré des coordonnées obtenues dans le DataFrame existant.



### Extraction d'Informations

J'ai employé Selenium et Chromedriver pour automatiser la navigation sur Google Maps.

J'ai ensuite extrait les détails supplémentaires, tels que les avis et les heures d'ouverture des établissements.

### Enrichissement de la Base de Données

Le projet a permis d'ajouter des coordonnées géographiques et des informations web aux entrées existantes, améliorant ainsi la précision et l'utilité de la base de données SIRENE.

## Conclusion



Ce projet a démontré l'importance de l'enrichissement des bases de données pour des applications pratiques. J'ai acquis des compétences en manipulation de données avec Python, en utilisation des API, et en automatisation de la navigation web avec Selenium. Le travail en binôme a également renforcé mes compétences en collaboration et en gestion de projet.