

27 AVRIL 2025

SAE OUTILS DÉCISIONNELS

CITY FIGHTING

MEHDI BOUSSEDRA
CLEMENT TANG
BUT3 SD VCOD
Groupe 33

City Fighting

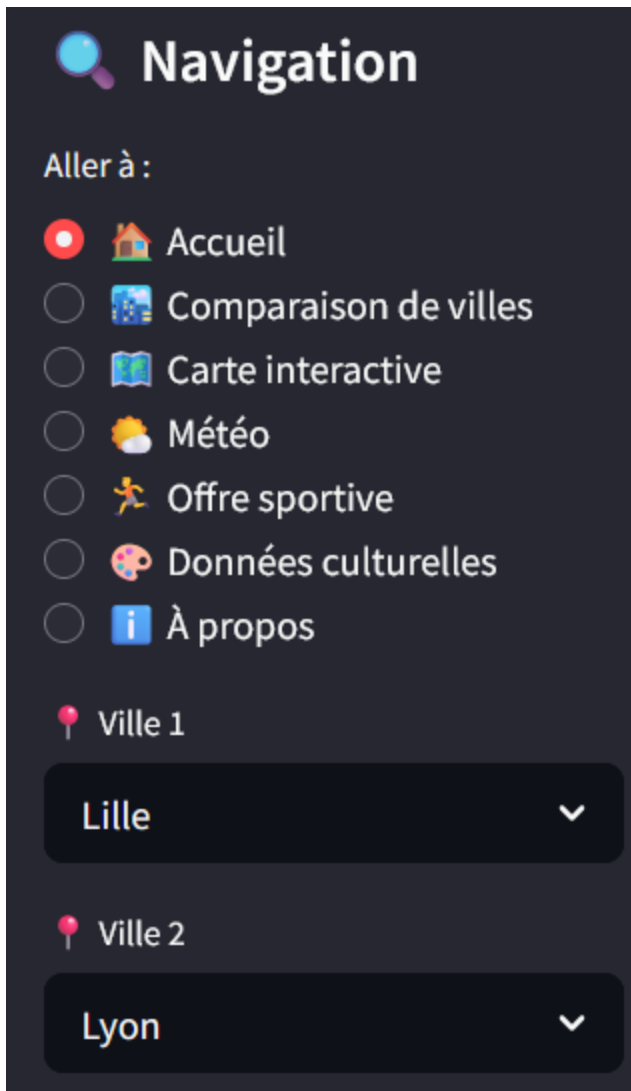
1. Présentation du projet

Introduction

Ce projet a été réalisé dans le cadre de la SAE Outils Décisionnels. L'objectif est de développer une application web permettant de comparer deux villes françaises de plus de 20 000 habitants selon des critères démographiques, économiques, sportifs et météorologiques. Le projet repose sur l'utilisation de Python, Streamlit et diverses sources de données ouvertes.

2. Fonctionnalités de l'application

- Comparaison de deux villes sur la population, l'emploi, le logement, les équipements sportifs.
- Accès à la météo actuelle et prévisions météo sur le nombre de jours souhaités.
- Carte interactive des villes avec indicateurs.
- Présentation des équipements sportifs par catégorie.
- Présentation du nombre d'infrastructures culturelles par catégorie.



3. Choix des sources de données

Sources de données :

INSEE - Population

Fichier Excel contenant la population municipale des communes françaises entre 2013 et 2021. Récupéré depuis le site de l'INSEE. Utilisé pour filtrer les villes de plus de 20 000 habitants.

Lien : <https://www.data.gouv.fr/fr/datasets/population-municipale-des-communes-france-entiere/>

INSEE - Emploi et Logement

Deux fichiers CSV (emploi et logement par commune) extraits du site de l'INSEE. Ces données sont agrégées par ville regroupée.

Lien : <https://www.insee.fr/fr/statistiques/7632867?sommaire=7632977>

Lien : <https://www.insee.fr/fr/statistiques/7631186?sommaire=7631713>

Data ES

Fichier CSV issu de l'open data des équipements sportifs, permettant de catégoriser les infrastructures sportives.

Lien : <https://equipements.sports.gouv.fr/explore/dataset/data-es>

Data Gouv

Fichier des coordonnées géographiques : récupéré sur data.gouv.fr pour obtenir les latitudes/longitudes nécessaires à la carte.

Fichier CSV sur les données culturelles, permettant de catégoriser les infrastructures culturelles.

Lien : <https://www.data.gouv.fr/fr/datasets/communes-de-france-base-des-codes-postaux/>

Lien : <https://www.data.gouv.fr/fr/datasets/base-des-lieux-et-equipements-culturels-basilic/>

WeatherAPI

API gratuite utilisée pour obtenir les conditions météo actuelles et les prévisions pour chaque ville. L'API a été choisie car elle permet la recherche par nom de ville, sans nécessiter les coordonnées GPS.

Wikipédia

Utilisé pour intégrer un lien vers la page de chaque ville.

Choix des années 2013 à 2021

Les années 2013 à 2021 ont été retenues car ce sont les dernières années pour lesquelles l'INSEE propose des données comparables et cohérentes à l'échelle communale. Les années antérieures à 2013 étaient soit incomplètes, soit présentaient des formats

différents. Après 2021, les données ne sont pas encore complètement disponibles pour toutes les communes.

4. Difficultés rencontrées

Plusieurs difficultés ont été rencontrées au cours de ce projet. Tout d'abord, la gestion et la structuration des données issues de différentes sources a nécessité un travail important de nettoyage, de normalisation et de fusion. Les formats csv étaient parfois corrompu ou à encodage spécifique (UTF-8, Latin-1).

Il a également fallu adapter certains traitements pour regrouper les arrondissements de grandes villes comme Paris, Lyon et Marseille.

```
# === 2. Fonction pour regrouper les arrondissements ===
def regrouper_arrondissements(nom):
    if isinstance(nom, str):
        if "Paris" in nom:
            return "Paris"
        elif "Lyon" in nom:
            return "Lyon"
        elif "Marseille" in nom:
            return "Marseille"
        else:
            return nom.strip()
    return nom

df_filtered["ville_regroupee"] = df_filtered["libgeo"].apply(regrouper_arrondissements)
```

Une difficulté majeure a été liée à la taille des bases de données. Quatre fichiers essentiels notamment ceux concernant l'emploi, le logement, les équipements sportifs et culturels étaient trop volumineux pour être hébergés directement sur Github. Par conséquent, ces fichiers ont été stockés sur Google Drive, et leur téléchargement automatique a dû être implémenté dans l'application. Cela a engendré plusieurs problématiques lors du déploiement, notamment :

- Des erreurs de lecture liées aux droits d'accès ou à la connectivité au moment de l'exécution ;
- Des difficultés pour garantir que les fichiers soient présents au bon moment et au bon format ;
- Un temps de chargement plus long au démarrage de l'application ;
- Des tests de robustesse supplémentaires pour gérer les cas où un fichier ne pouvait être téléchargé.

```

import gdown
import os

# === 1. Télécharger les fichiers Google Drive dans le dossier actuel ===
fichiers_drive = {
    "base-cc-emploi-pop-active-2020_v2.CSV": "1ZSzHZwxcsoDn4VxyjsPbj7K0zmoJYPdL",
    "base-cc-logement-2020.CSV": "1LU57jvjNQSOgwnGM6_cNbENVYeZ1FKzP",
    "data-es.csv": "1nn1TT2_2hNXLYDFOLA7StDYLOR0UE3NN",
    "base-des-lieux-et-des-equipements-culturels.csv": "1SAGJ_bxmCx4G4FdwD3hcXlQNwYaHC_OM"
}

for nom_fichier, file_id in fichiers_drive.items():
    if not os.path.exists(nom_fichier):
        url = f"https://drive.google.com/uc?id={file_id}"
        print(f"📄 Téléchargement de {nom_fichier}...")
        gdown.download(url, nom_fichier, quiet=False)
    else:
        print(f"✅ {nom_fichier} déjà présent.")

```

Enfin, la récupération des liens Wikipédia et leur intégration dans l'interface a nécessité de nombreuses adaptations pour s'assurer qu'elles s'affichent correctement selon la ville sélectionnée.

Pour la récupération des données météo historiques non disponible sans version payante. La météo avec coordonnées GPS dans certaines API a été complexe à gérer : nous avons donc opté pour WeatherAPI, qui permet une recherche directe par nom de ville.

Sources de données

Nous avons utilisé plusieurs sources de données fiables pour alimenter notre application :

- ****INSEE**** (<https://www.insee.fr/fr/statistiques>) : Fournit des fichiers CSV et Excel sur la population, l'emploi et le logement.
 - ***Population :** * Données de 2013 à 2021 par commune.
 - ***Emploi :** * Base-cc-emploi-pop-active-2020_v2.CSV.
 - ***Logement :** * Base-cc-logement-2020.CSV.
 - ***Culturelle :** * Base-des-lieux-et-des-equipements-culturels.csv.
 - ***Téléchargement :** * Directement depuis les pages de statistiques sur le site de l'INSEE.
- ****Data ES (Équipements sportifs)**** (<https://equipements.sports.gouv.fr/explore/dataset/data-es>) : Fichier CSV regroupant les équipements par commune.
- ****WeatherAPI**** (<https://www.weatherapi.com/>) : API météo gratuite permettant de récupérer les conditions actuelles et les prévisions.

- *Avantage :* Permet de rechercher une ville par son nom (sans coordonnées), ce qui facilite l'intégration.

5. Conclusion

L'application City Fighting offre une interface fluide et intuitive pour explorer les différences entre les grandes villes françaises. Elle allie visualisation, données ouvertes et interactivité. Ce projet a permis de mettre en œuvre des compétences en traitement de données, développement web et intégration d'APIs externes.

Lien Github :

<https://github.com/MehdiBoussedra/cityfighter.git>

Lien de l'application :

<https://cityfighter.streamlit.app/>