

PRÉSENTATION DES DONNÉES (1/3)

Dans le cadre du développement de cette application web, nous avons collecté diverses données issues de sources variées afin d'alimenter les analyses et les comparaisons proposées. Ces données ont été réparties sur plusieurs onglets.

Onglet sur le choix des villes (recherche des villes) :

- **communes-france-2025.csv** : référentiel des communes de France obtenu sur le site [data.gouv](https://data.gouv.fr/). Cette base de données contient divers données relatives aux communes : code postal, code insee, région, département, population, densité, superficie, etc...

Onglet sur les informations générales :

- **data_tranche_age.csv** : données sur la répartition de la population par tranches d'âge (obtenu par extraction Web scraping avec le code "webscrapping_tranches_d-age.py" sur le site de l'[insee](https://www.insee.fr/))
- **data_sexe_pourcentage.csv** : données sur la répartition de la population par genre (obtenu par extraction par extraction Web scraping avec le code "webscrapping_sexe.py" sur le site de l'[insee](https://www.insee.fr/))
- **data_cat_socio.csv** : données sur la répartition des catégories socioprofessionnelles (obtenu par extraction par extraction Web scraping avec le code "webscrapping_cat_socio.py" sur le site de l'[insee](https://www.insee.fr/))
- **data_menage_composition.csv** : données sur la répartition des types de ménage (obtenu par extraction par extraction Web scraping avec le code "webscrapping_menage_composition.py" sur le site de l'[insee](https://www.insee.fr/))
- **data_statut_conj.csv** : données sur la répartition des types de statut conjugal (obtenu par extraction par extraction Web scraping avec le code "webscrapping_statut_conj.py" sur le site de l'[insee](https://www.insee.fr/))
- **data_famille.csv** : données sur la répartition des du nombre d'enfants par famille (obtenu par extraction par extraction Web scraping avec le code "webscrapping_famille.py" sur le site de l'[insee](https://www.insee.fr/))

PRÉSENTATION DES DONNÉES (2/3)

- **data_age_graph.csv** : données sur l'évolution du nombre d'habitants par année (obtenu par extraction Web scraping avec le code "webscrapping_age_graph.py" sur le site de l'[insee](#))
- **data_naissance_deces** : données sur l'évolution du nombre d'habitants par année (obtenu par extraction Web scraping avec le code "webscrapping_naissance_deces.py" sur le site de l'[insee](#))

Onglet sur les données d'emploi :

- Les données concernant la population active (%), l'évolution de la population active, la part des femmes actives et des hommes actifs, le taux de chômeurs, la part d'actifs ayant un emploi, la répartition par secteur d'activité ainsi que par tranche d'âge ont été obtenues via l'api [francetravail](#).
- **data_salaire_genre_categ_pro.csv** : données sur le salaire horaire par genre et catégorie socioprofessionnelle (obtenu par extraction Web scraping avec le code "webscrapping_statut_pro_genre.ipynb" sur le site de l'[insee](#))

Onglet sur les données de logement :

- **data_log_T2.csv** : données sur le nombre de logements principaux ou secondaires ainsi que le nombre d'appartements et de maisons (obtenu par extraction Web scraping avec le code "webscrapping_type_logement.ipynb" sur le site de l'[insee](#))
- **data_superficie_log.csv** : données sur la superficie moyenne des logements (obtenu par extraction Web scraping avec le code "webscrapping_superficie_log.py" sur le site de [linternaute](#))
- **data_dfv_2019-2023_reduit.csv** : données sur le type prix du m² moyen sur la période 2019 à 2023 (obtenu sur le site de [data.gouv](#))
- **data_nbr_pcs_logement.csv** : données sur le nombre de pièces moyen pour un logement (obtenu par extraction Web scraping avec le code "wescrapping_log_T3.ipynb" sur le site de l'[insee](#))

PRÉSENTATION DES DONNÉES (3/3)

- **data_nb_moyen_hab_log.csv** : données sur le nombre d'habitants moyen par logement (obtenu par extraction Web scraping avec le code "webscrapping_nb_moyen_hab_log.py" sur le site de l'[insee](#))
- **data_statut_occupation.csv** : données sur la part de propriétaire et locataire (obtenu par extraction Web scraping avec le code "webscrapping_logement.py" sur le site de l'[insee](#))
- **data_date_log.csv** : données sur la date de création des logements (obtenu par extraction Web scraping avec le code "webscrapping_data_log.py" sur le site de l'[insee](#))
- **data_anciennete_emmenagement.csv** : données sur l'ancienneté des déménagements (obtenu par extraction Web scraping avec le code "webscrapping_anciennete_emmenagement.py" sur le site de l'[insee](#))

Onglet sur les données météo :

- Les données relatives à la météo actuelle + aux prévisions sur 2 jours ont été récupérées via l'api [Prévisionmétéo](#)
- **climat.xlsx** : données sur type de climat des communes de France. (source : étude climatique réalisée par MaxBruciamacchie, lien : [github](#))
- **data_climat_saison.csv** : données sur la hauteur des pluies et les heures d'ensoleillement recensé (obtenu par extraction Web scraping avec le code "webscrapping_météo.ipynb" sur le site de [linternaute](#))

Onglet sur les données loisirs :

- **loisirs.csv** : données en lien avec les activités sportives, culturelle et informations sur l'hôtellerie (obtenu par extraction Web scraping avec le code "webscrapping_loisirs.ipynb" sur le site de [linternaute](#))
- **education1.csv et ecole.csv** : données sur la part d'étudiants scolarisés, du niveau de diplôme des habitants et du nombre d'établissements scolaires (obtenu par extraction Web scraping avec le code "webscrapping_education.ipynb" sur le site de [linternaute](#))

CHOIX DES SOURCES DE DONNÉES

Dans le cadre du développement de City-Fighting, notre priorité a été de privilégier des sources de données fiables et reconnues, plutôt que de rechercher systématiquement les données les plus récentes. Cet objectif visait à assurer la cohérence, la qualité et la comparabilité des informations exploitées.

Les données présentées dans l'application datent principalement des années 2021 à 2024. Les principales sources utilisées sont l'INSEE, France Travail et la plateforme Data.gouv, toutes reconnues pour la fiabilité et l'officialité des données qu'elles mettent à disposition.

Afin d'enrichir l'application et de compléter certaines informations non disponibles dans les bases officielles, nous avons également utilisé d'autres sources secondaires. Le site Linternaute : pour obtenir des données locales supplémentaires concernant les communes (équipements, cadre de vie, etc.). Par ailleurs, nous avons intégré certains travaux de l'utilisateur MaxBruciamacchie disponibles sur GitHub.