

Systeme d'Analyse de l'Accessibilité des Soins en France

Identification des Déserts Médicaux

Projet d'analyse des soins neurologiques
avec calcul de l'Indicateur Composite d'Accessibilité (ICA)



JARI Jenna
GUILLOU Floriane
KARABAJAKIAN Fred
KAMGA Armand

Sommaire

PARTIE 1 : INTRODUCTION

PARTIE 2 : ARCHITECTURE ET OBJECTIFS

PARTIE 3 : LES 5 ÉTAPES DU PROJET + DEMONSTRATION

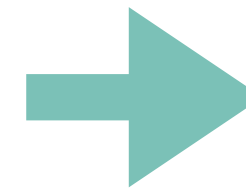
PARTIE 4 : RÉSULTATS ET CONCLUSION

Contexte & problématique



Le problème des déserts médicaux en France

- Inégalités d'accès aux soins selon les régions
- Manque de praticiens dans certaines zones
- Population vieillissante avec besoins croissants
- Difficulté à identifier les zones prioritaires



Objectif du projet :

Développer un système complet pour analyser l'accessibilité des soins neurologiques et identifier les déserts médicaux de manière objective et quantifiable.

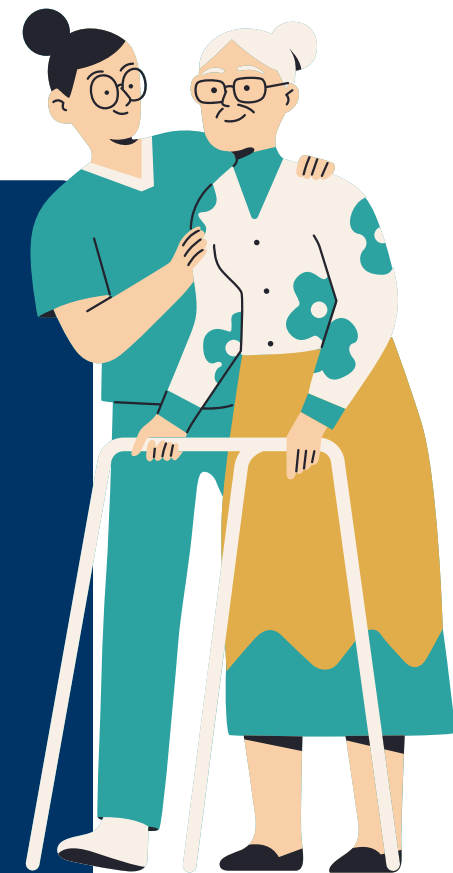
Objectifs du projet

Objectif principal :

Créer un système d'analyse complet de l'accessibilité des soins neurologiques en France

Objectifs spécifiques :

- Collecter des données sur les praticiens et leur localisation
 - Calculer des indicateurs d'accessibilité géographique
 - Analyser la demande démographique (population 65+)
- Créer un indicateur composite (ICA) combinant plusieurs dimensions
 - Visualiser les résultats de manière interactive
 - Exposer les données via une API



Architecture Globale du projet



Technologies principales :

- Python, MySQL, Streamlit, Flask
- Pandas, BeautifulSoup, Plotly

Sources :

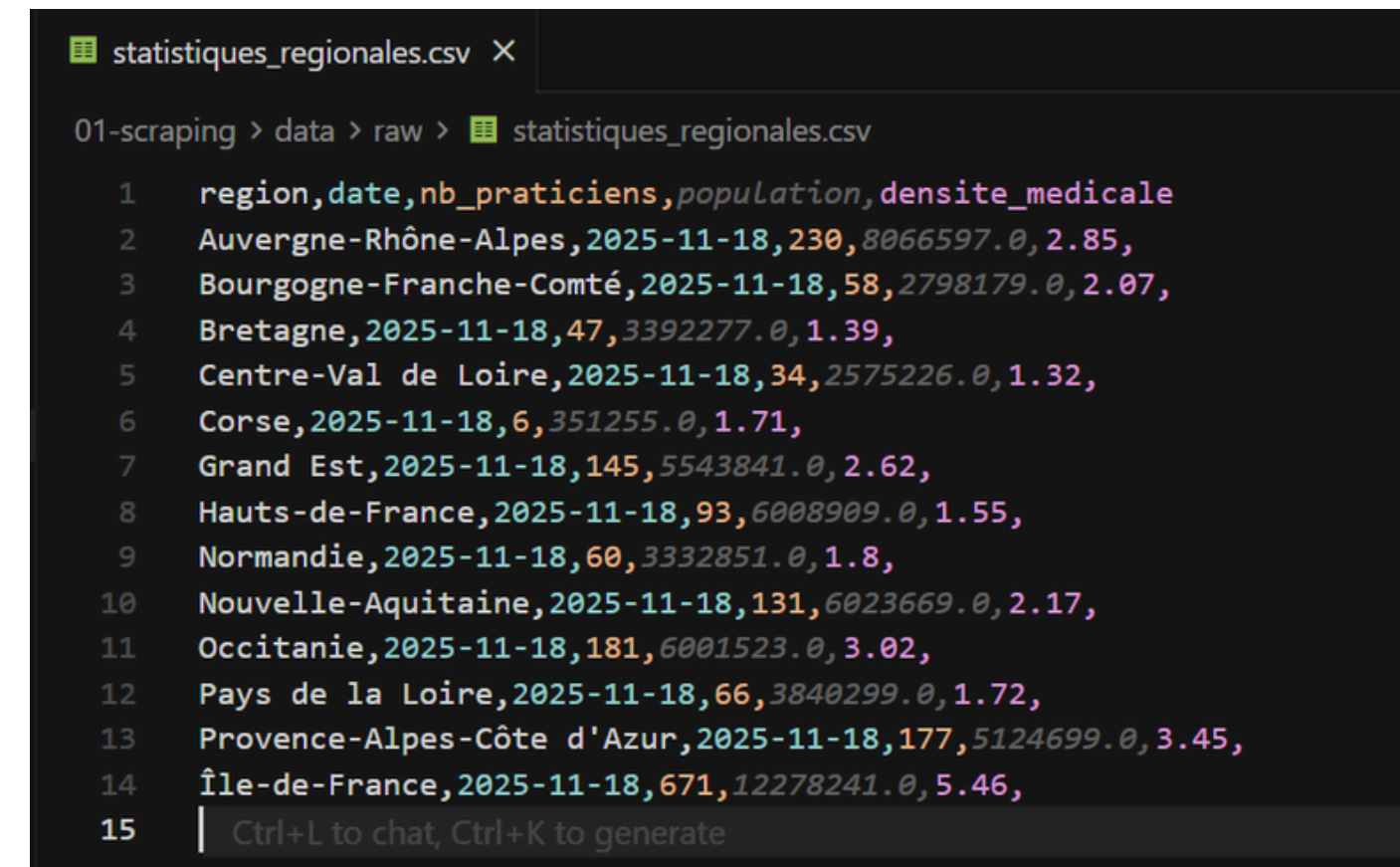
- Doctolib
- data.gouv.fr
- INSEE

Etape 1 - Scraping

Objectif : Collecter les données sur les praticiens

Sources de données :

- Doctolib : Praticiens neurologues avec localisation
- data.gouv.fr : Statistiques régionales et démographiques
- INSEE : Données sur la population 65+ par commune



```
statistiques_regionales.csv X
01-scraping > data > raw > statistiques_regionales.csv
1  region,date,nb_praticiens,population,densite_medicale
2  Auvergne-Rhône-Alpes,2025-11-18,230,8066597.0,2.85,
3  Bourgogne-Franche-Comté,2025-11-18,58,2798179.0,2.07,
4  Bretagne,2025-11-18,47,3392277.0,1.39,
5  Centre-Val de Loire,2025-11-18,34,2575226.0,1.32,
6  Corse,2025-11-18,6,351255.0,1.71,
7  Grand Est,2025-11-18,145,5543841.0,2.62,
8  Hauts-de-France,2025-11-18,93,6008909.0,1.55,
9  Normandie,2025-11-18,60,3332851.0,1.8,
10 Nouvelle-Aquitaine,2025-11-18,131,6023669.0,2.17,
11 Occitanie,2025-11-18,181,6001523.0,3.02,
12 Pays de la Loire,2025-11-18,66,3840299.0,1.72,
13 Provence-Alpes-Côte d'Azur,2025-11-18,177,5124699.0,3.45,
14 Île-de-France,2025-11-18,671,12278241.0,5.46,
15 | Ctrl+L to chat, Ctrl+K to generate
```

Exemple de csv obtenu

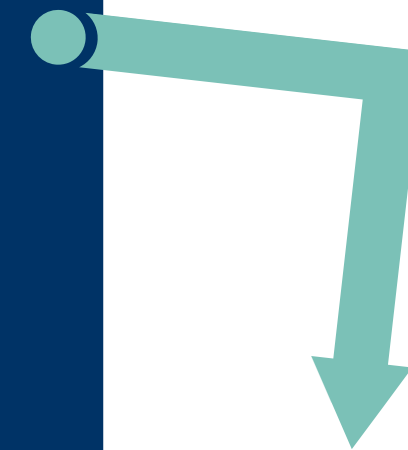
Etape 1 - Détails techniques

Scraping Doctolib :

- Navigation automatisée avec Selenium+Pandas
- Extraction : nom, spécialité, ville, code postal
- Geocoding pour obtenir latitude/longitude
- Délais entre requêtes pour respecter les limites

Scraping Data.gouv.fr :

- Téléchargement de datasets publics
- Statistiques régionales (population, densité médicale)

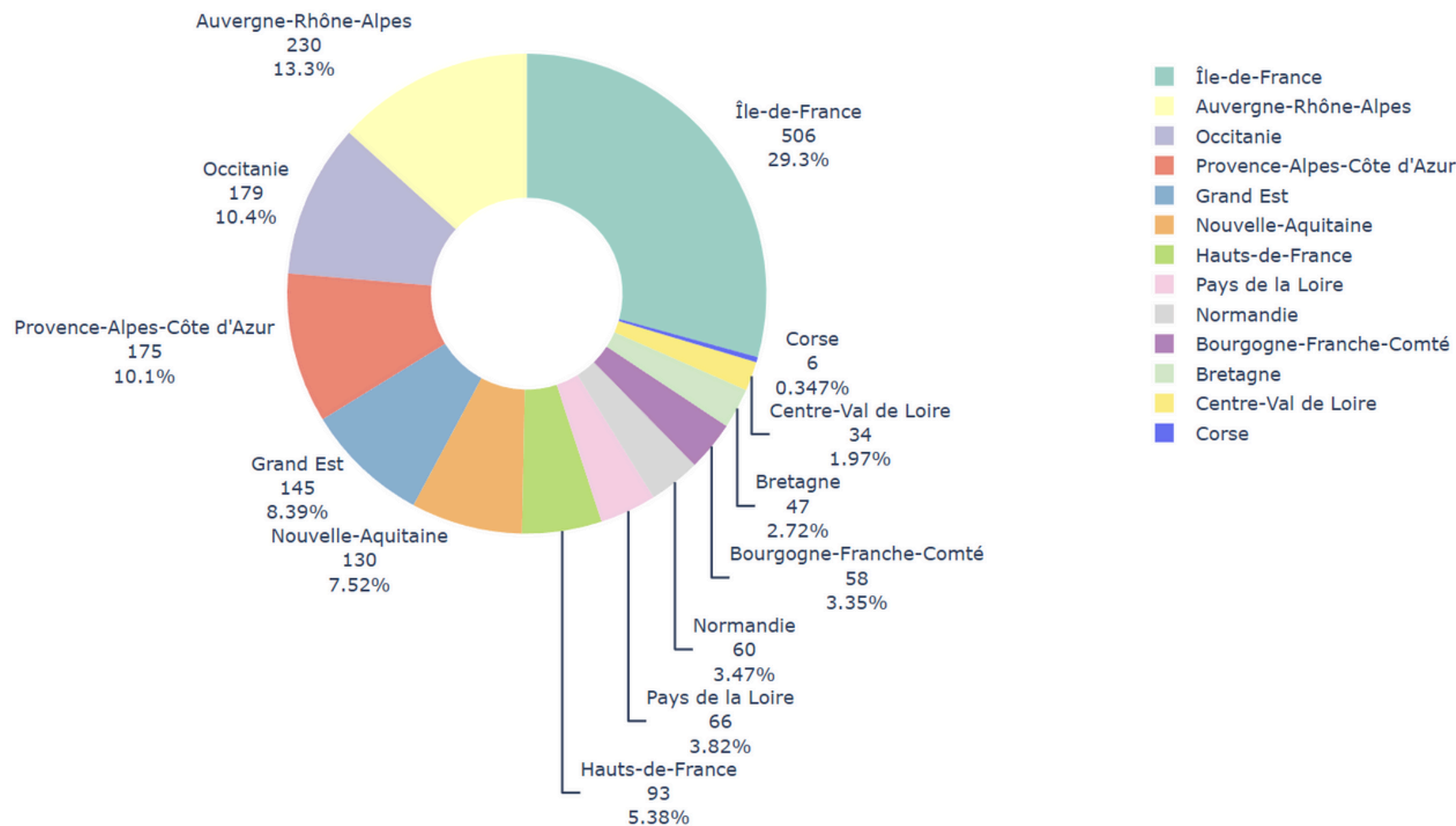


Données collectées :

- Praticiens neurologues
- Coordonnées géographiques
- Statistiques démographiques⁷

ETAPE 1- PREMIERE VISUALISATION

Répartition des praticiens par région

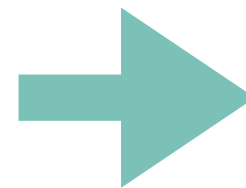


Etape 2 - Base de Données

Objectif : Stocker et structurer les données collectées

Base de données :

- MySQL
- Schéma relationnel optimisé
- Index pour performances



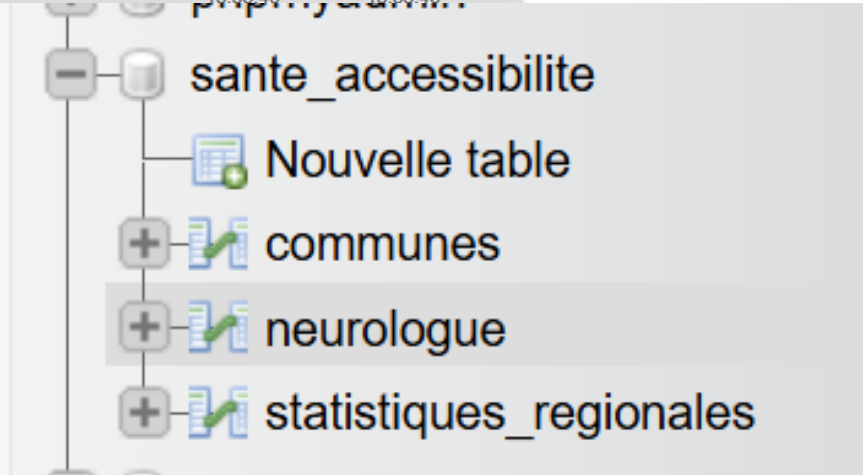
Tables principales :

- praticiens : Informations sur les praticiens
- statistiques_regionales : Données agrégées par région
- communes : Informations géographiques

Etape 2 - Schéma de la base de données

```
CREATE TABLE IF NOT EXISTS praticiens (  
  id INT PRIMARY KEY AUTO_INCREMENT,  
  specialite VARCHAR(100) NOT NULL,  
  ville VARCHAR(100) NOT NULL,  
  code_postal VARCHAR(10) NOT NULL,  
  region VARCHAR(100) NOT NULL,  
  latitude DECIMAL(10, 8),  
  longitude DECIMAL(11, 8),  
  delai_rdv_jours INT,  
  date_scraping DATE NOT NULL,  
  created_at TIMESTAMP DEFAULT CURRENT_TIMESTAMP,  
  updated_at TIMESTAMP DEFAULT CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP,  
  INDEX idx_region (region),  
  INDEX idx_specialite (specialite),  
  INDEX idx_ville (ville),  
  INDEX idx_code_postal (code_postal),  
  INDEX idx_delai (delai_rdv_jours)| TAB to jump here  
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_unicode_ci;
```

	id	specialite	ville	code_postal	region	latitude	longitude	delai_rdv_jours	date_scraping	created_at	updated_at
<input type="checkbox"/> Éditer <input type="checkbox"/> Copier <input type="checkbox"/> Supprimer	1	neurologue	Paris	75017	Île-de-France	48.85888970	2.32004100	NULL	2025-11-17	2025-11-21 15:58:46	2025-11-21 15:58:46
<input type="checkbox"/> Éditer <input type="checkbox"/> Copier <input type="checkbox"/> Supprimer	2	neurologue	Paris	75019	Île-de-France	48.88905790	2.39081000	NULL	2025-11-17	2025-11-21 15:58:46	2025-11-21 15:58:46
<input type="checkbox"/> Éditer <input type="checkbox"/> Copier <input type="checkbox"/> Supprimer	3	neurologue	Paris	75006	Île-de-France	48.85888970	2.32004100	N			
<input type="checkbox"/> Éditer <input type="checkbox"/> Copier <input type="checkbox"/> Supprimer	4	neurologue	Paris	75008	Île-de-France	NULL	NULL	N			
<input type="checkbox"/> Éditer <input type="checkbox"/> Copier <input type="checkbox"/> Supprimer	5	neurologue	Paris	75017	Île-de-France	NULL	NULL	N			



- Points clés :
- Index sur région, spécialité, ville pour performances
 - Clés étrangères pour intégrité référentielle
 - Timestamps pour traçabilité

Etape 3 - Agrégation et ICA

Un indicateur qui combine plusieurs dimensions

1. ACCESSIBILITÉ GÉOGRAPHIQUE (60%)

- Distance au praticien le plus proche
- Couverture (nombre de praticiens dans 30 km)
- Densité locale

2. OFFRE/DEMANDE DÉMOGRAPHIQUE (40%)

- Offre : Nombre de praticiens / 1000 habitants 65+
- Demande : Proportion 65+ × prévalence neurologique

$$\text{ICA}_{\text{total}} = w_{\text{geo}} \times \text{ICA}_{\text{geo}} + w_{\text{demo}} \times \text{ICA}_{\text{demo}}$$

Étape 3 - Détails du Calcul ICA

Étape 1 - Normalisation :

- Toutes les métriques sont normalisées entre 0 et 1
- Utilisation de percentiles pour robustesse

Étape 3 - Indice final :

$$ICA_{total} = w_{geo} \times ICA_{geo} + w_{demo} \times ICA_{demo}$$

(avec $w_{geo} + w_{demo} = 1$)

Étape 2 - Sous-indices :

$$ICA_{geo} = 0.5 \times distance_{norm} + 0.3 \times coverage_{norm} + 0.2 \times density_{norm}$$
$$ICA_{demo} = \alpha \times offre_{norm} + (1-\alpha) \times demande_{norm}$$

Étape 4 - Classification :

- Désert sévère : $ICA < 0.3$
- Désert modéré : $0.3 \leq ICA < 0.5$
- Zone tendue : $0.5 \leq ICA < 0.7$
- Zone normale : $ICA \geq 0.7$

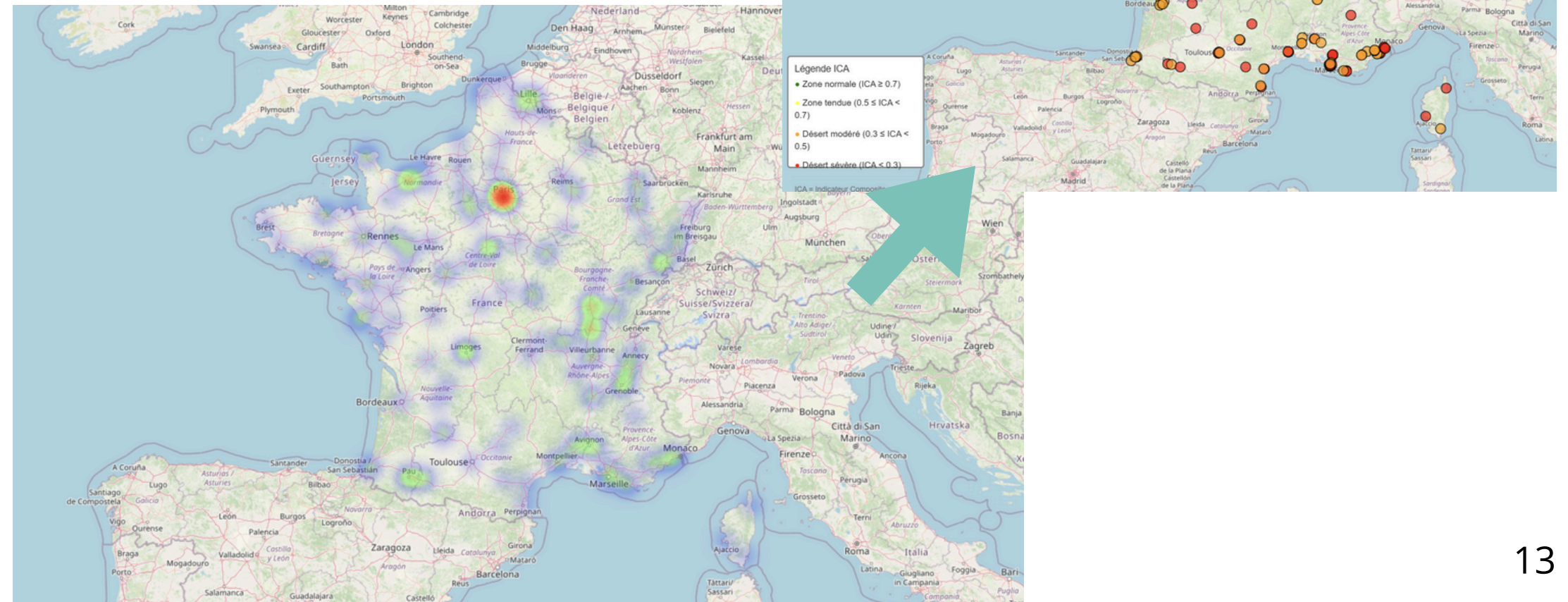
PARTIE 3 : LES 5 ÉTAPES DU PROJET

Etape 3 - Résultats de l'Agrégation

```
Agen,47000,Nouvelle-Aquitaine,44.20897008,0.61398943,0.55,12,80000,17600,0.22,0.05,0.682,0.231,0.8351516604723654,0.007751937984496124,0.0,0.0,0.421,0.0,0.253,Désert sév
Aix-en-Provence,13080,Provence-Alpes-Côte d'Azur,43.53337723,5.44514841,0.94,93,140000,30800,0.22,0.05,3.019,0.231,0.7182592015345883,0.07751937984496124,0.006436810735
Aix-en-Provence,13090,Provence-Alpes-Côte d'Azur,43.536531,5.43122565,0.89,93,40000,8800,0.22,0.05,10.568,0.231,0.7332454142189186,0.07751937984496124,0.0272290590192476
Aix-en-Provence,13100,Provence-Alpes-Côte d'Azur,43.54271292,5.45894707,1.51,90,220000,48400,0.22,0.05,1.86,0.231,0.5474163769332216,0.07493540051679587,0.00324457126488
Ajaccio,20000,Corse,41.9557665,8.72860483,1.66,9,60000,13200,0.22,0.05,0.682,0.231,0.5024577388802304,0.00516795865633075,0.0,0.0,0.254,0.0,0.152,Désert sévère
Albi,81000,Occitanie,43.91316907,2.16623638,1.37,42,280000,58800,0.21,0.05,0.714,0.221,0.5893777724493467,0.03359173126614987,8.813775931781351e-05,0.15873015873015883,6
Albi,81030,Occitanie,43.9277552,2.147899,0.0,42,20000,4200,0.21,0.05,10.0,0.221,1.0,0.03359173126614987,0.02566461379135589,0.15873015873015883,0.517,0.092,0.347,Désert
Amiens,80000,Hauts-de-France,49.88536235,2.30118421,1.63,30,200000,36000,0.18,0.05,0.833,0.189,0.5114494664908287,0.023255813953488372,0.00041590005178093335,0.666666666
Angers,49000,Pays de la Loire,47.48685391,-0.54695008,1.38,39,160000,32000,0.2,0.05,1.219,0.21,0.5863805299124807,0.031007751937984496,0.001479061773552062,0.3333333333
Angers,49100,Pays de la Loire,47.47030472,-0.546145,0.49,39,100000,20000,0.2,0.05,1.95,0.21,0.853135115693562,0.031007751937984496,0.003492458712968369,0.33333333333333
Angoulême,16000,Nouvelle-Aquitaine,45.66776384,0.14739924,0.89,15,100000,22000,0.22,0.05,0.682,0.231,0.7332454142189186,0.0103359173126615,0.0,0.0,0.372,0.0,0.223,Désert
```

Fichiers générés :

- ICA_par_commune.csv : ICA pour chaque commune
- densite_medicale.csv : Densité par région
- deserts_medicaux.csv : Liste des déserts identifiés
- praticiens_par_region_specialite.csv : Répartition



Etape 4 - API Rest

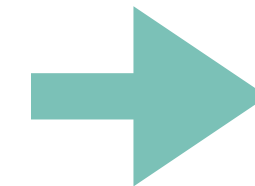
Objectif : Exposer les données via une API

Framework :

- Flask (Python)
- Architecture RESTful

Endpoints principaux :

- /api/praticiens : Liste des praticiens
- /api/indicateurs : Indicateurs par région
- /api/accessibilite : Données d'accessibilité
- /api/health : Vérification de santé



Utilité :

- Permet l'intégration avec d'autres systèmes
- Accès programmatique aux données
- Séparation des couches (données / présentation)

Etape 5 - Visualisation

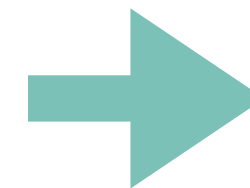
Objectif : Créer des dashboards interactifs

Framework :

- Streamlit (application web Python)
- Plotly (graphiques interactifs)

Dashboards créés :

- Dashboard général (vue d'ensemble)
- Dashboard ICA (analyse détaillée de l'accessibilité)



Fonctionnalités :

- Filtres interactifs
- Graphiques dynamiques
- Export de données
- Visualisations géographiques

DEMONSTRATION

Résultats et Insights principaux



Résultats clés :

- Identification objective des déserts médicaux
- Analyse multi-dimensionnelle (géographique + démographique)
- Comparaison inter-régionale
- Visualisation interactive des données

Insights :

- Certaines régions sont plus touchées que d'autres
- La distance n'est pas le seul facteur
- La démographie joue un rôle important
- L'ICA permet une vision globale et nuancée

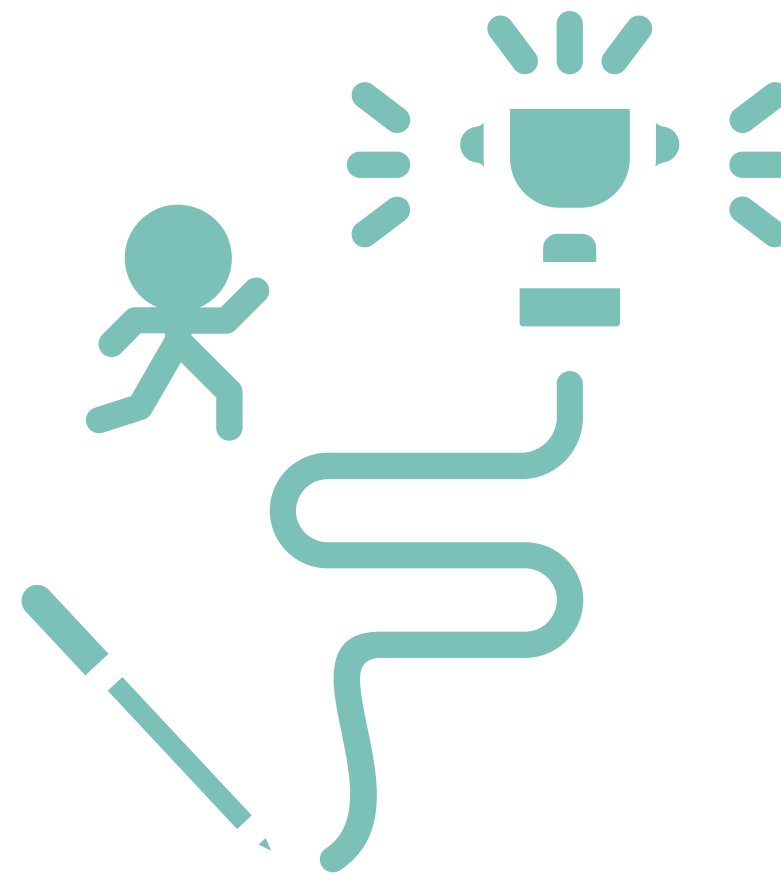
Défis rencontrés

DÉFI 1 - SCRAPING

Rate limiting de Doctolib
→ Délais entre requêtes

DÉFI 2 - GEOCODING

Besoin de coordonnées précises
→ API Nominatim avec gestion d'erreurs



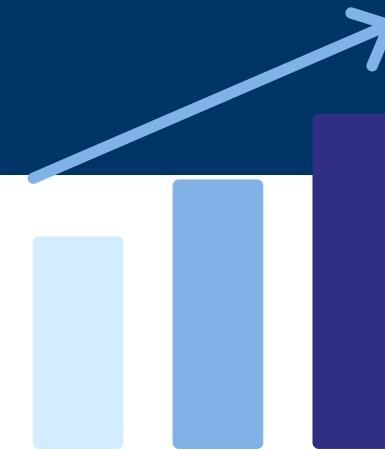
DÉFI 3 - CALCUL ICA

Normalisation robuste
→ Utilisation de percentiles

DÉFI 4 - PERFORMANCE

Calculs géospatiaux coûteux
→ Index sur BDD, cache Streamlit

Améliorations futures



- Mise à jour régulière des données
- Extension à d'autres spécialités médicales
- Dimension temporelle
- Système d'alertes
- Intégration avec d'autres systèmes

Comment réduire les déserts médicaux neurologiques ?



Cibler les zones prioritaires

Grâce à l'ICA, identifier précisément les communes où l'impact d'une intervention serait le plus fort.



Optimiser la mobilité des patients

Améliorer les transports médicaux, organiser des navettes ou partenariats locaux.



Favoriser l'installation de praticiens

Incitations financières, aides au logement, allègement de charges, infrastructures adaptées.



Planifier les ressources médicales

Ajuster les ouvertures de postes, les stages, et la formation selon les besoins des zones sous-dotées.

Conclusion



Ce qui a été accompli :

Système complet de collecte de données
Base de données structurée et optimisée
Calcul d'indicateurs d'accessibilité (ICA)
API REST pour exposer les données
Dashboard interactif pour visualisation

Valeur ajoutée :

- Outil d'aide à la décision pour les politiques de santé
- Identification objective des déserts médicaux
- Analyse multi-dimensionnelle de l'accessibilité
- Visualisation intuitive des résultats

Questions

