

RENNES DATA CHALLENGE

1ere édition

20 & 21 Janvier 2023

SUJET N*1

ÉQUIPE : *IZY_CODE*

Taoufiq OUEDRAOGO

Wiyao TAKOUGNADI

Farid GNANKAMBARY

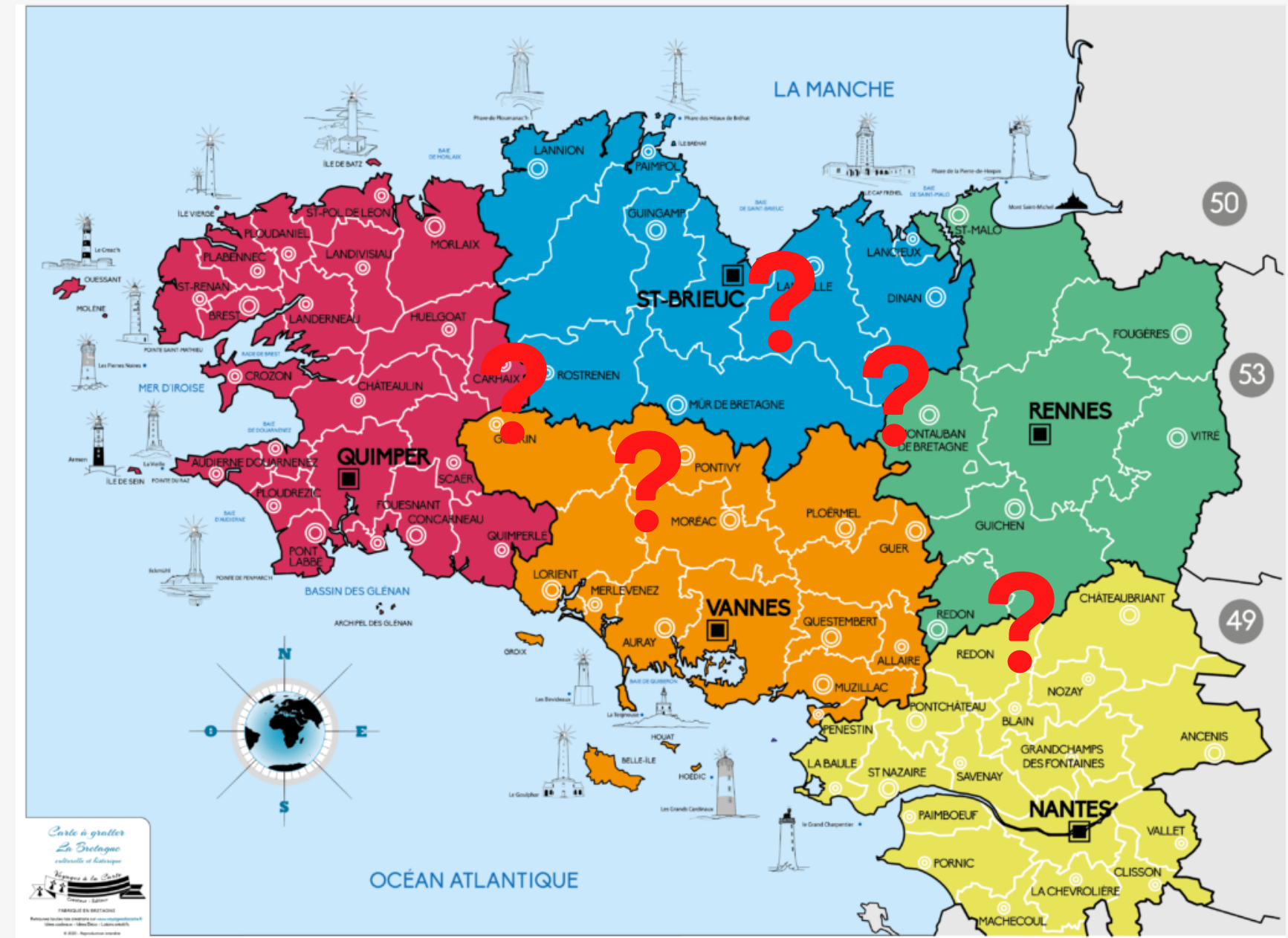
izy_code



PROBLÉMATIQUE

Pour mener à bien ses activités en Bretagne, ENEDIS a besoin de bases à partir desquels intervenir chez ses clients en cas de besoin.

Ces bases doivent être réparties sur le territoire de façon optimale en tenant compte en tenant compte de certains facteurs comme le temps de trajet d'une commune à l'autre, le nombre de demande d'interventions dans chaque commune...



DÉMARCHE

1

Pré-traiter le jeu de données pour en extraire les données les plus pertinentes pour notre étude

Création d'un dataframe contenant les niveaux d'intervention pondérés, les codes INSEE, les noms et les coordonnées de chaque commune.

2

Trouver les emplacements optimum des bases en appliquant un modèle de **clustering** aux données

Méthode des k-means avec attribution de poids aux différents niveaux d'intervention et aux temps de trajet (k-means pondérés)

3

Appliquer notre modèle au jeu de données, estimer la performance du modèle

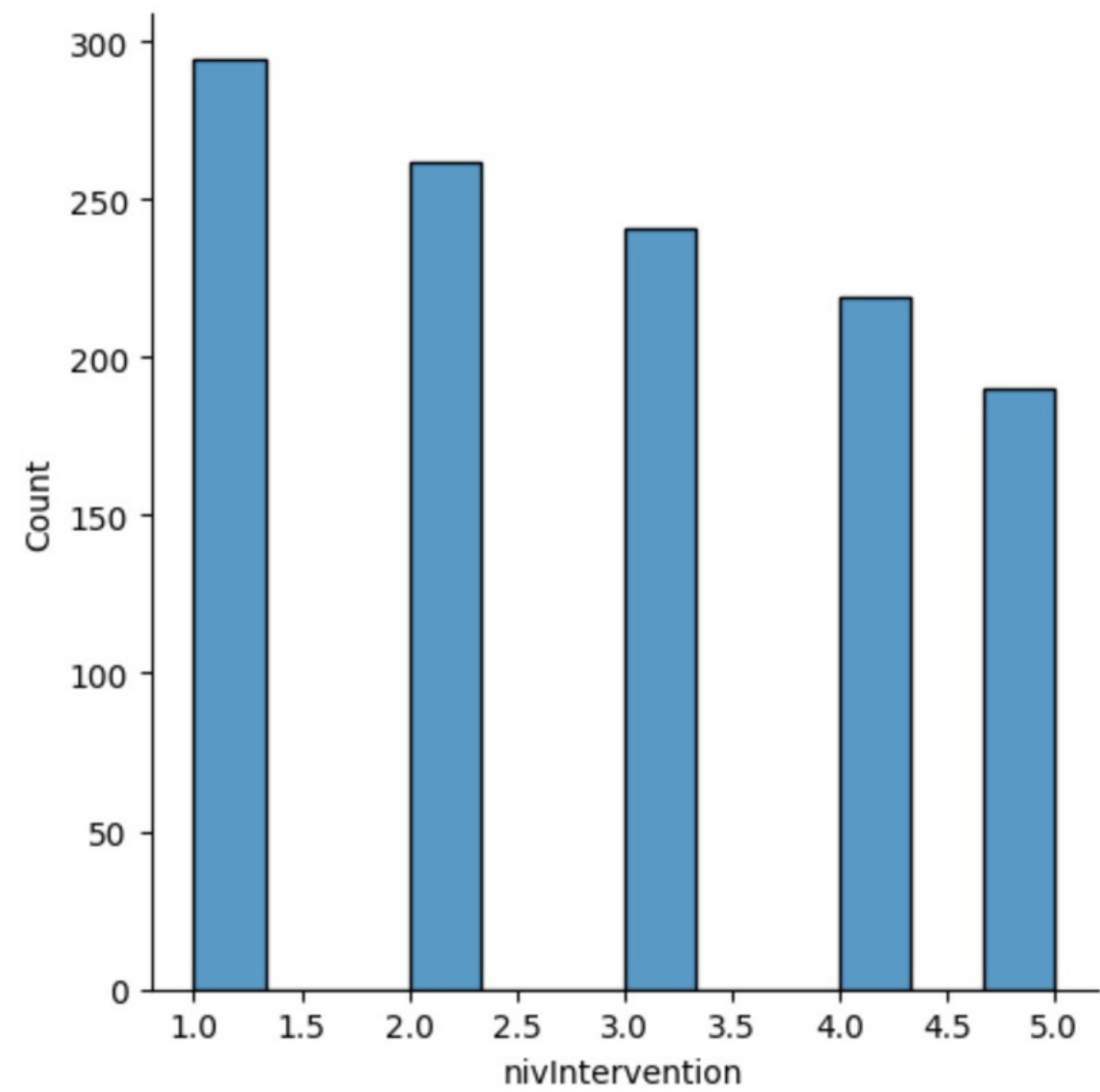
Obtention d'une carte indiquant les différentes bases opérationnelles et les communes qui leurs sont rattachées

HYPOTHÈSE RETENUE

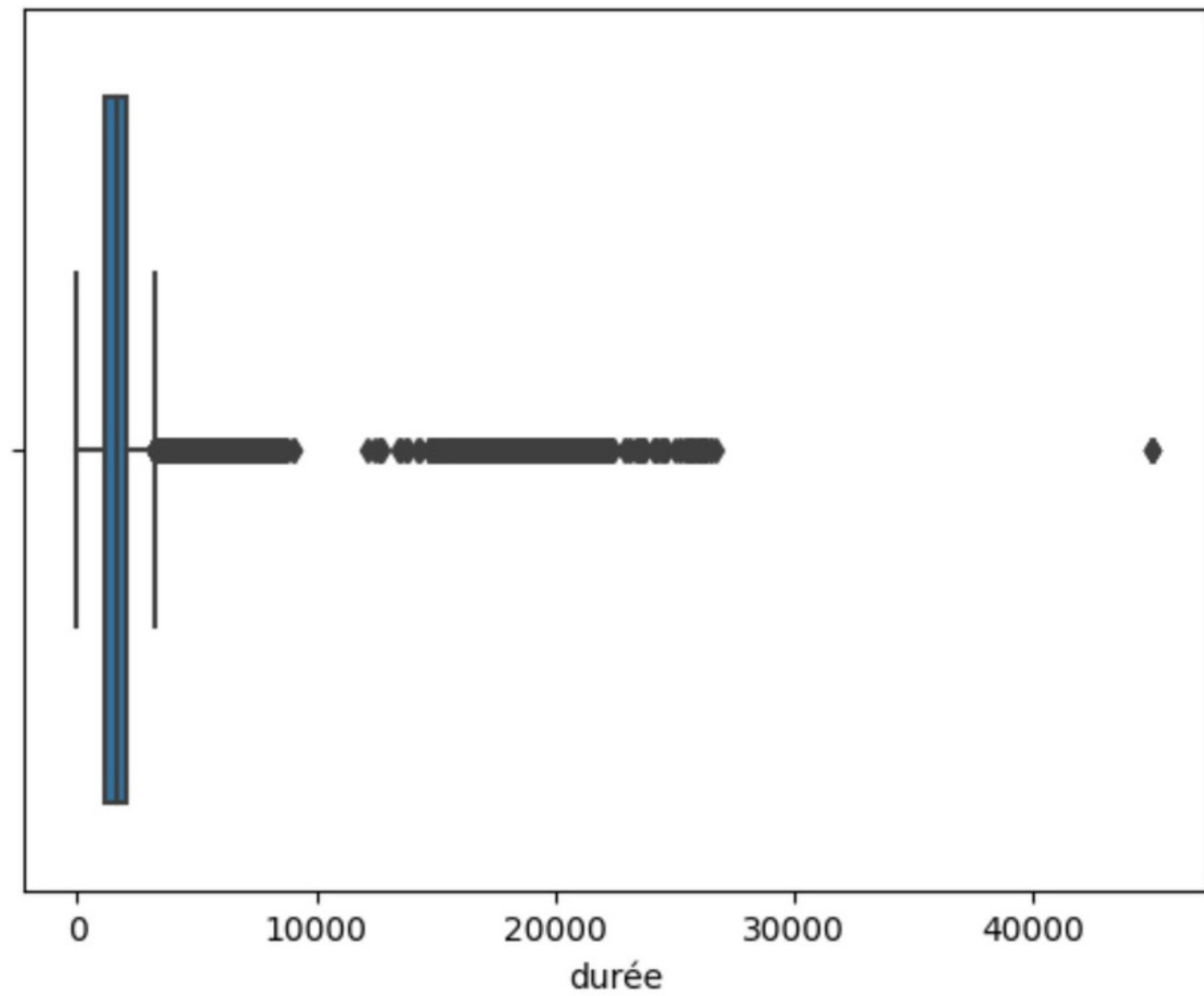
- Niveaux d'intervention
Les niveaux d'intervention clients et réseau pour chaque commune sont également importants
On préférera installer un centre dans une commune à niveau d'intervention élevé (≥ 4)



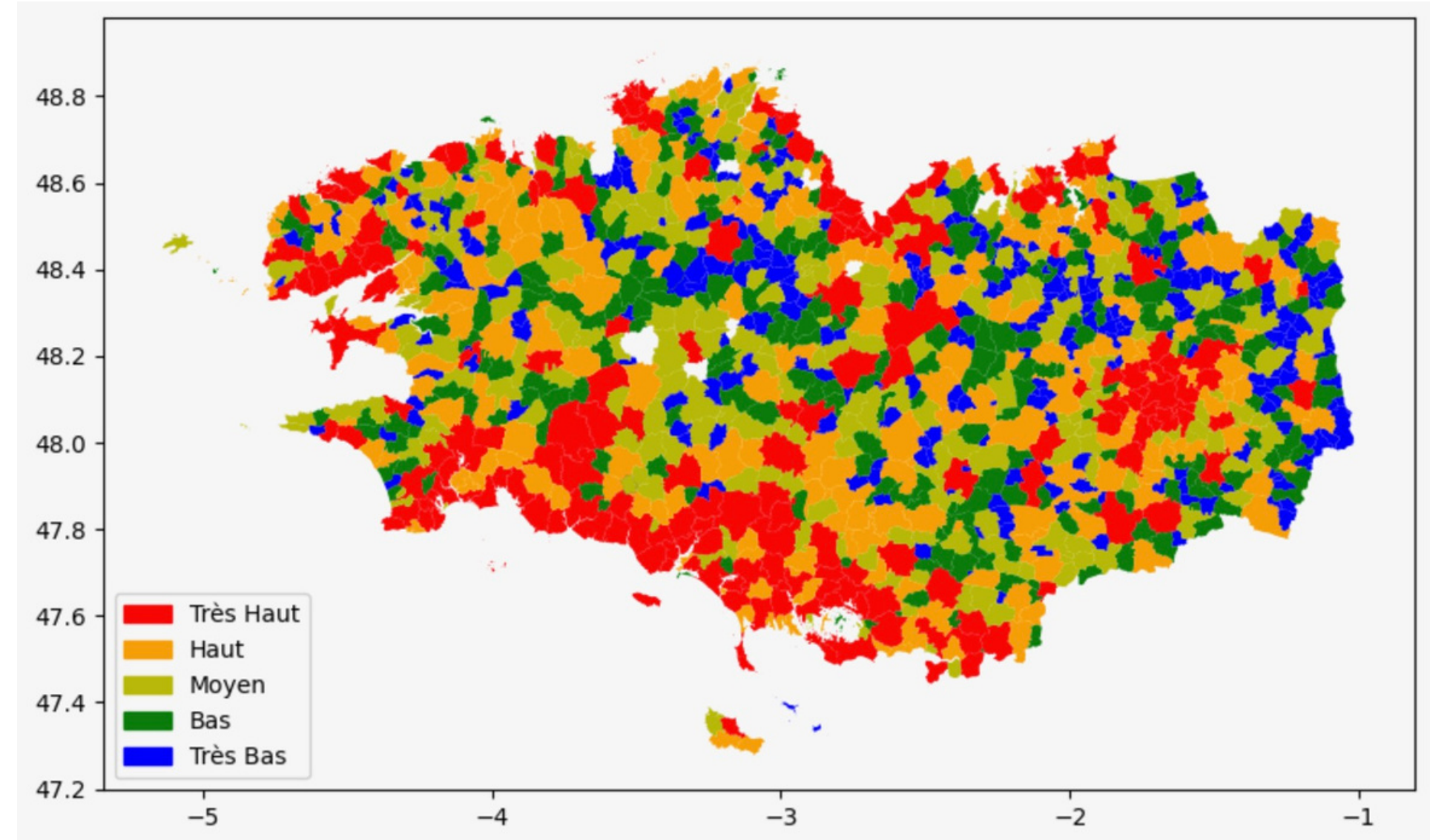
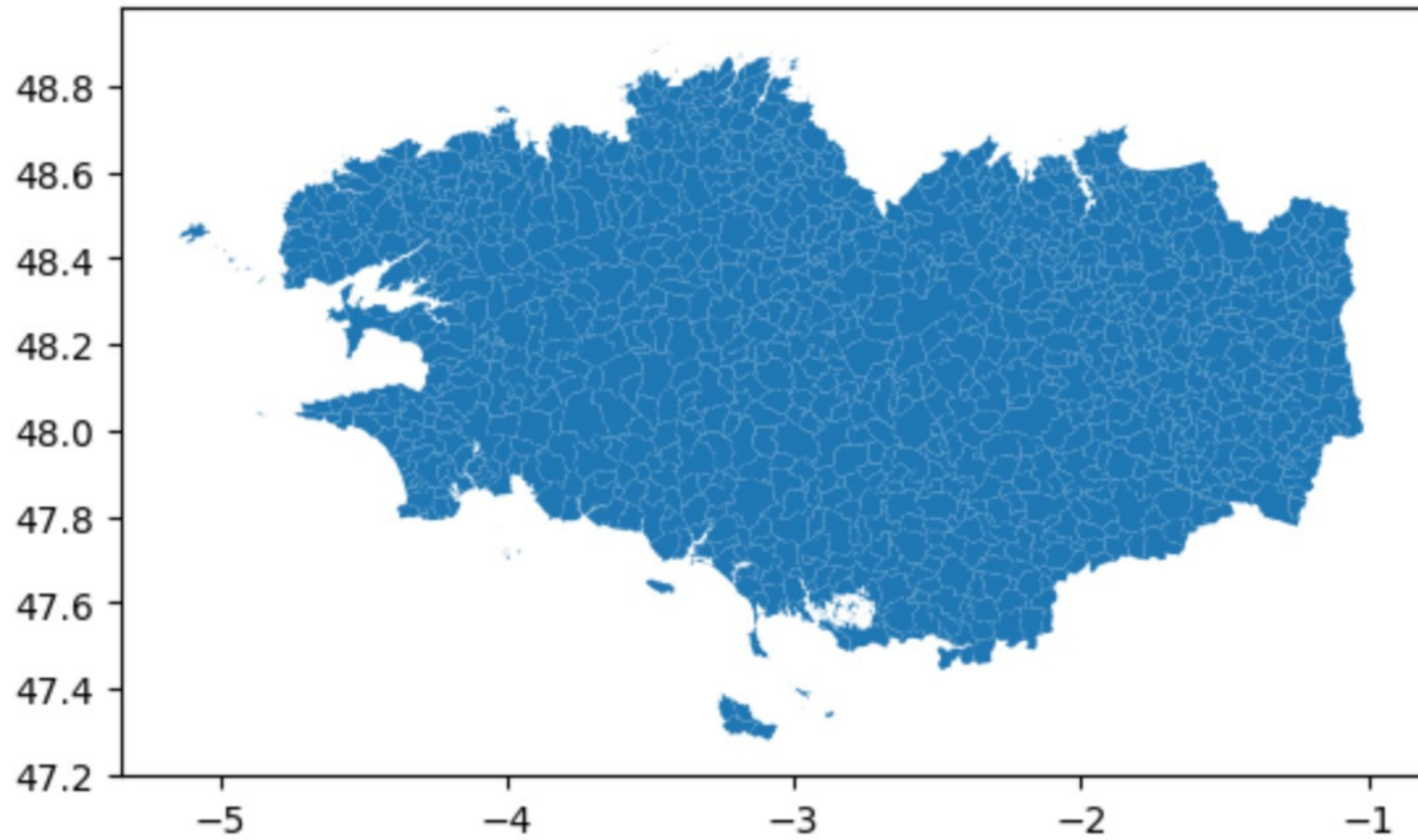
<seaborn.axisgrid.FacetGrid at 0x7f9ebb21bcd0>



<AxesSubplot: xlabel='durée'>



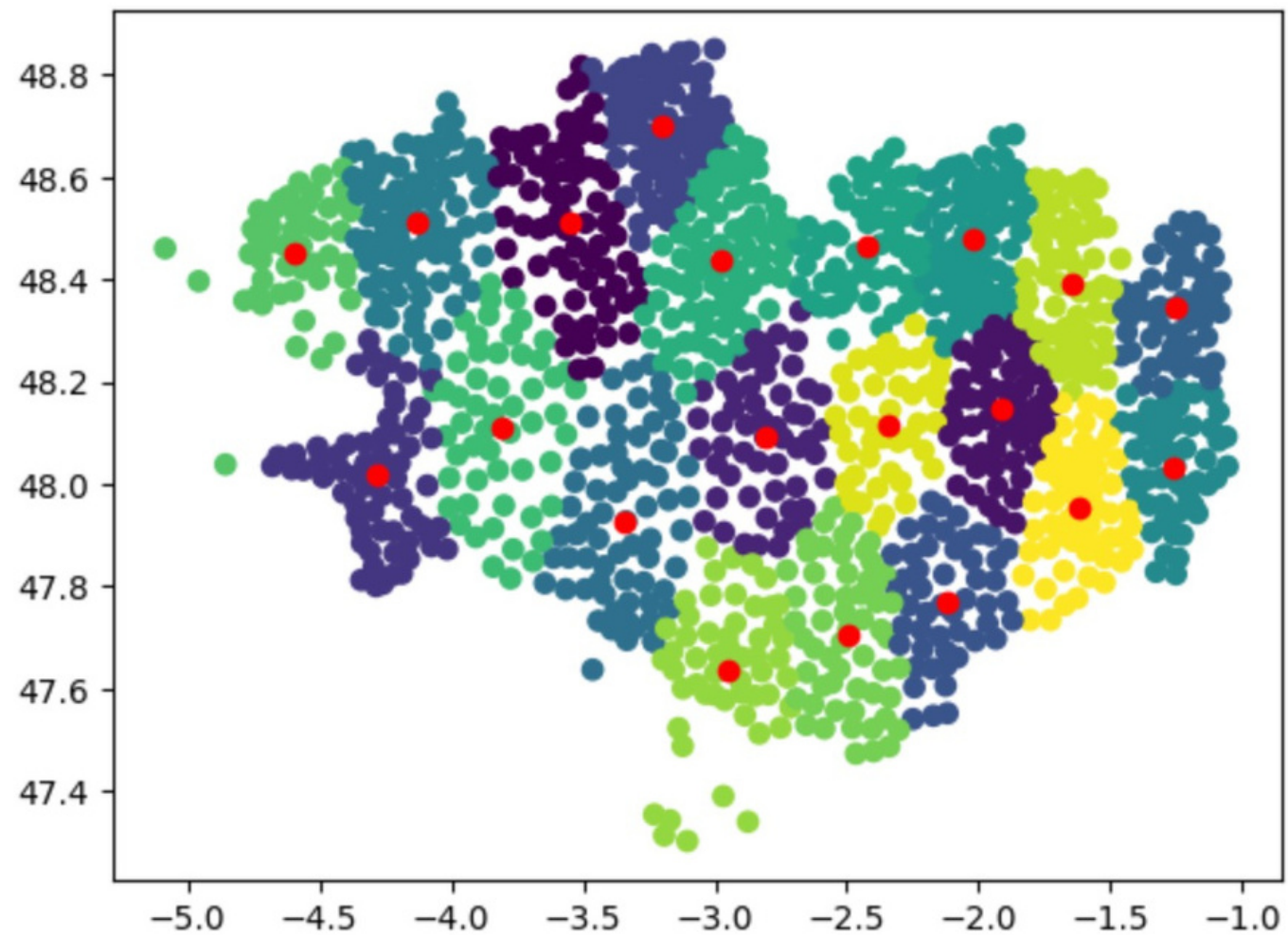
RÉSULTATS



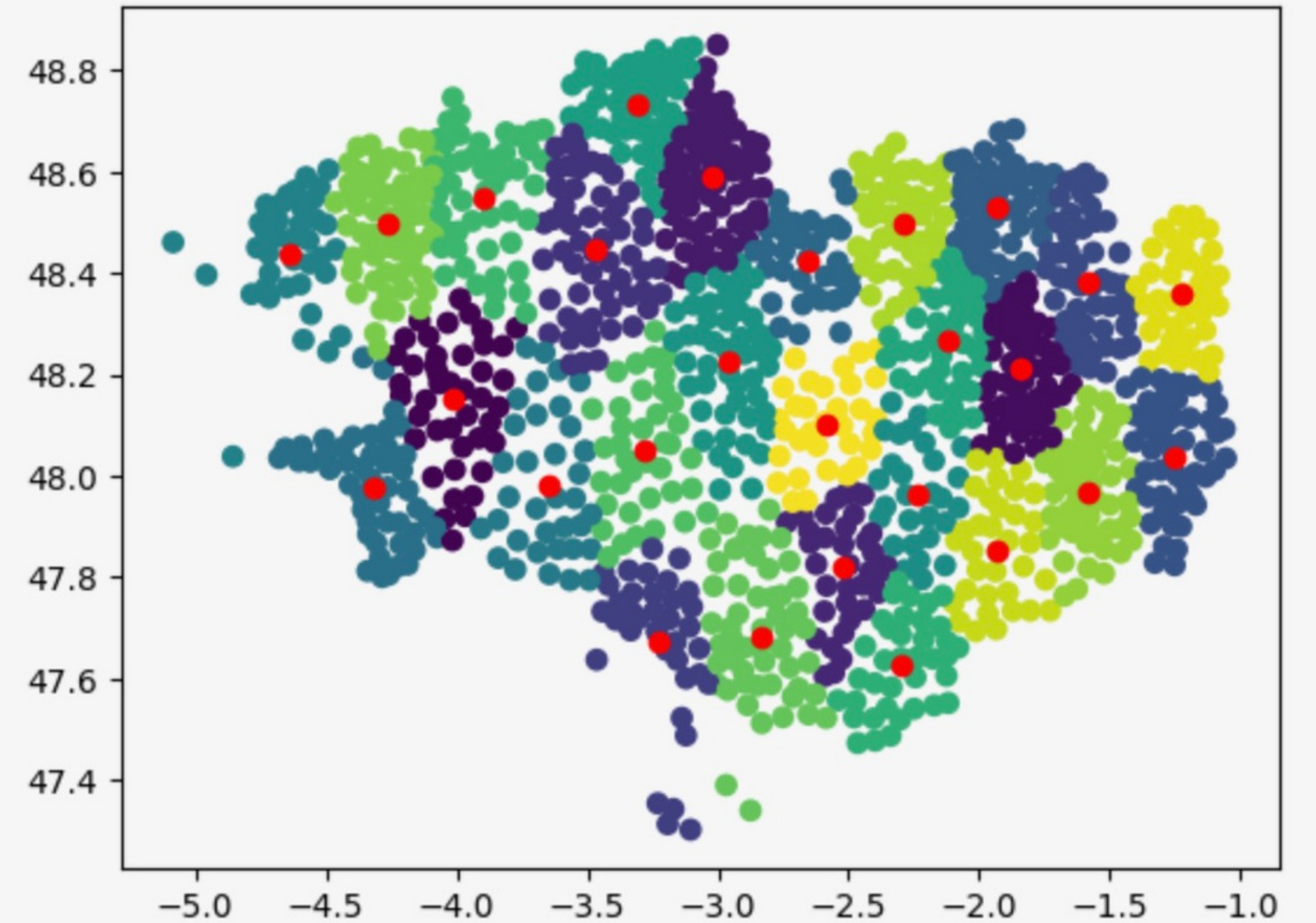
RÉSULTATS

```
1 plt.scatter(X[:,0], X[:,1], c=kmeans.labels_)
2 plt.scatter(kmeans.cluster_centers[:,0], kmeans.cluster_centers[:,1], c='r')
```

<matplotlib.collections.PathCollection at 0x7f9ece0a3e50>

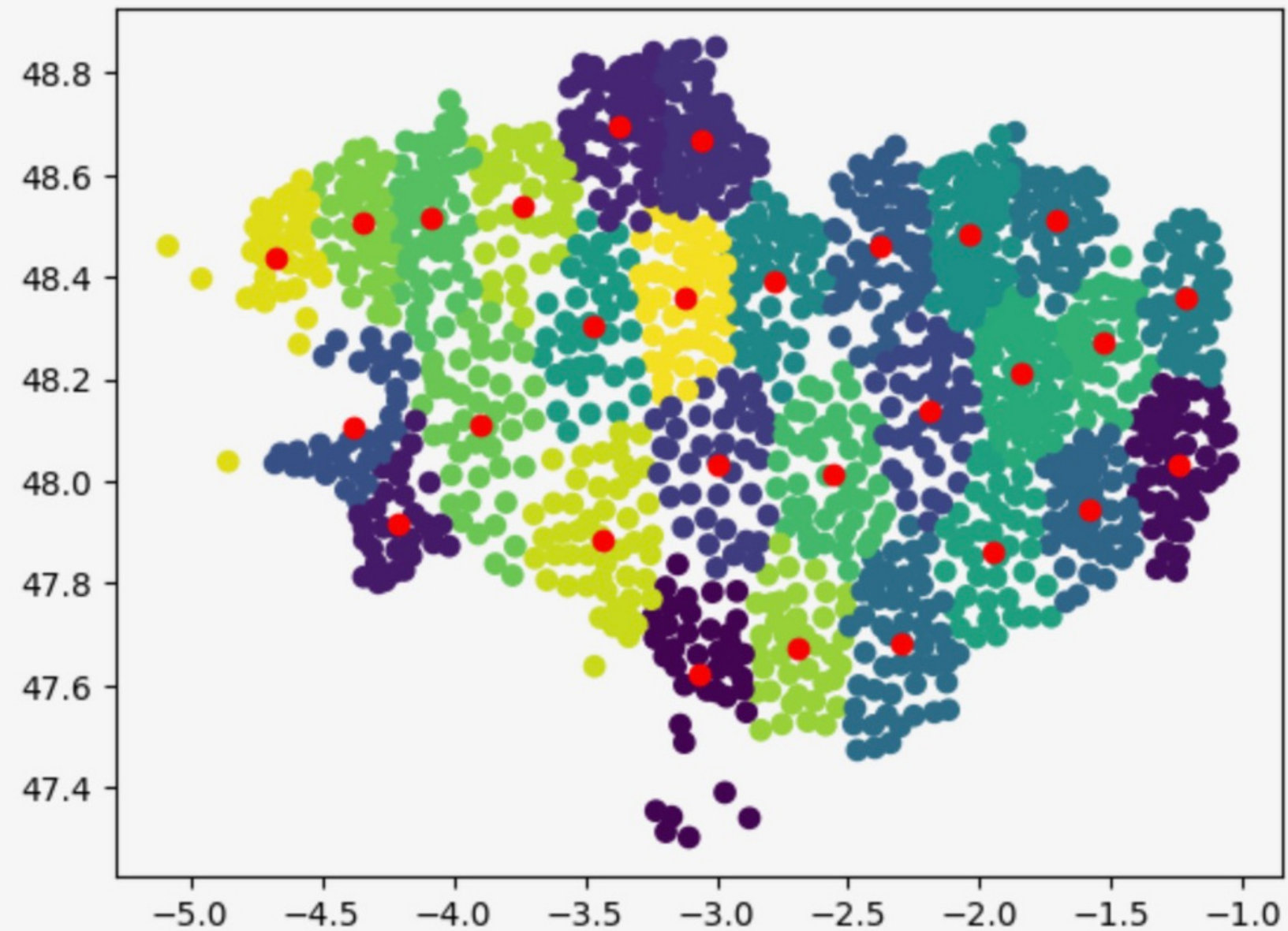


score calinski = 2067.987434227649
score silouhette = 0.3411430377309233

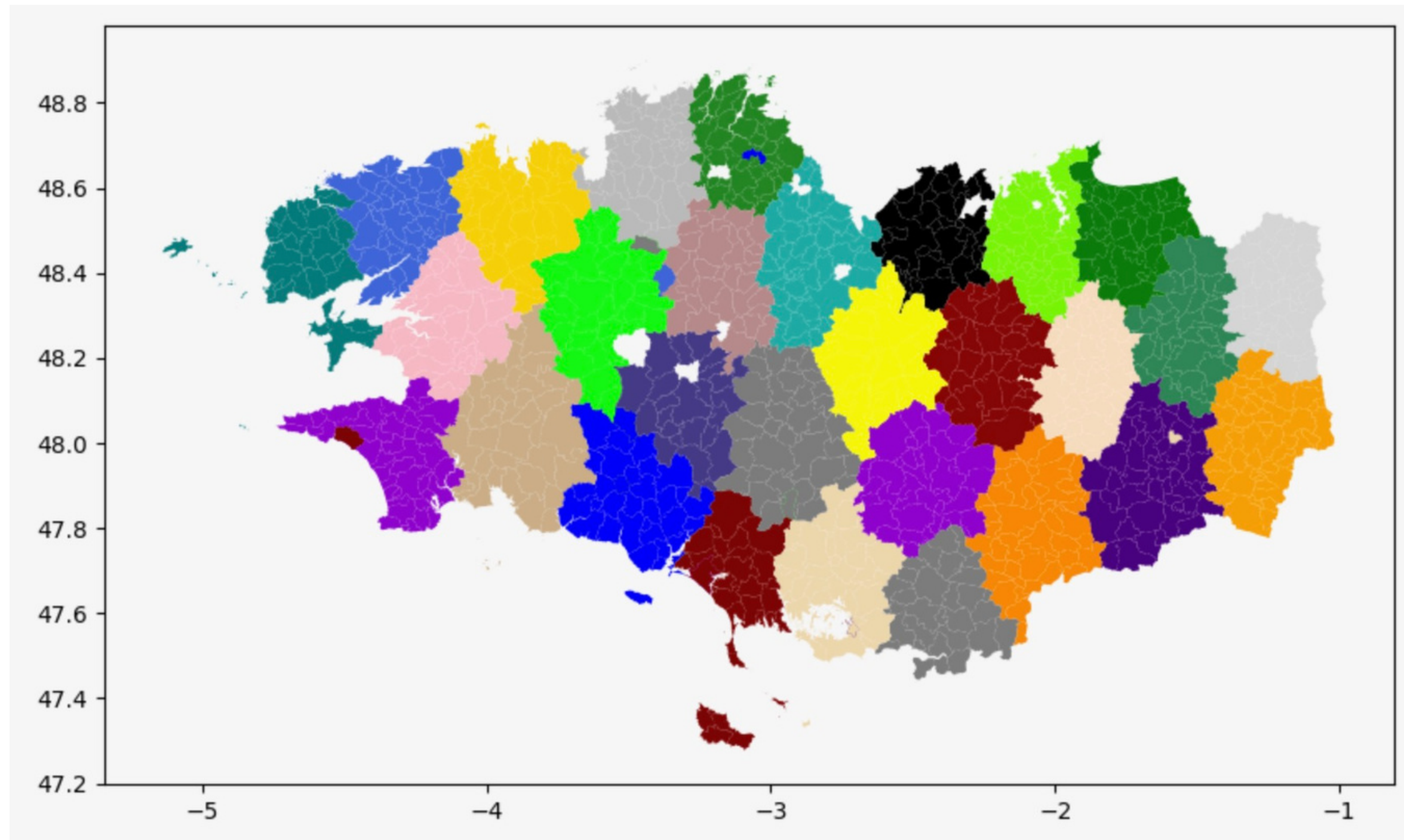


RÉSULTATS

```
GridSearchcv best params: {'algorithm': 'lloyd', 'init': 'random', 'n_clusters': 28}  
score calinski = 2149.0056785708084  
score silouhette = 0.34778633776259293
```



RÉSULTATS



**MERCI DE
VOTRE
ATTENTION !**

