

# **Quelle est la répartition des valeurs immobilières en France ?**







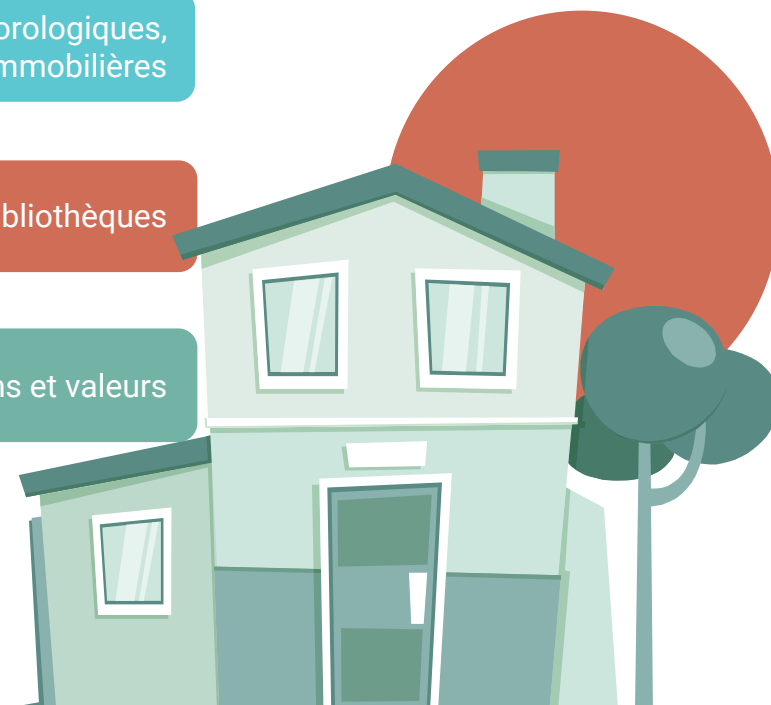
**Module : OpenData**

**Master 1 Miage . Année 2020-2021**

**Elèves : Mélissa Mérat, Gaëtan Pellerin, Samuel Rigaud**

# Sommaire

	01	<b>Organisation de l'équipe</b>	Membres, postes et missions
	02	<b>Choix du sujet</b>	Données météorologiques, de localisation et immobilières
	03	<b>Traitement des données</b>	Processus, logiciel et bibliothèques
	04	<b>Résultats</b>	Cartes, répartitions et valeurs



## 01 Organisation de l'équipe

# Equipe



Gaëtan PELLERIN

Mélissa Mérat

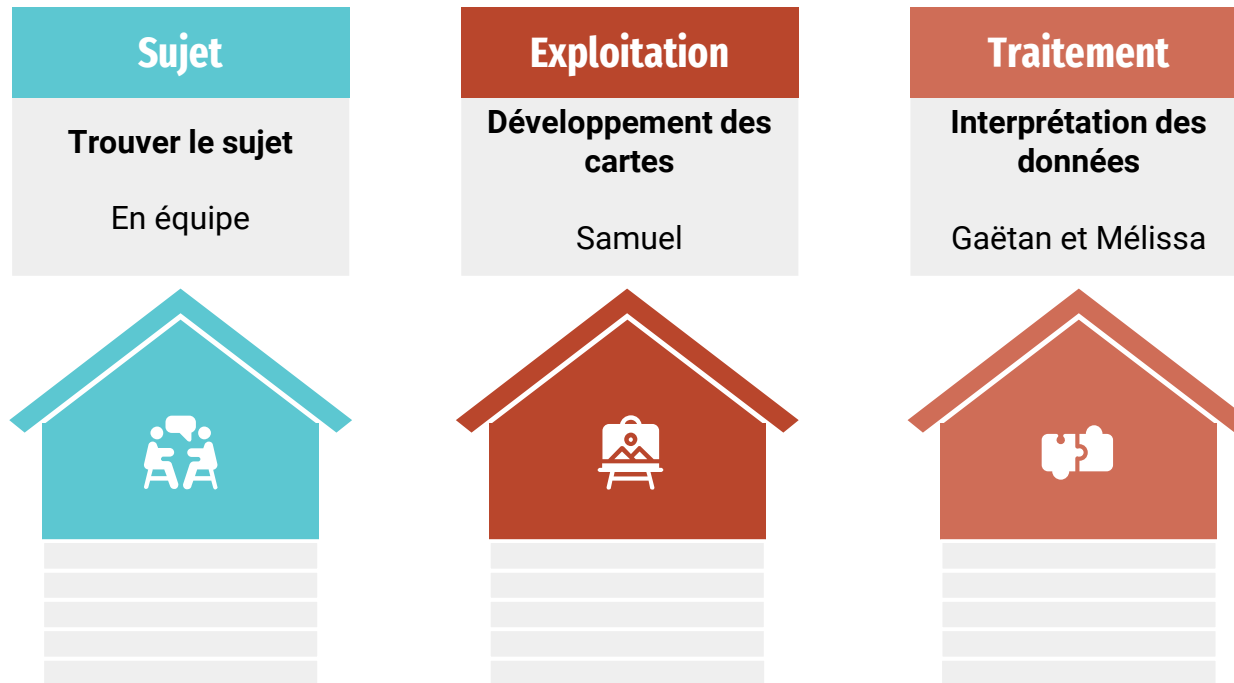


Samuel Rigaud



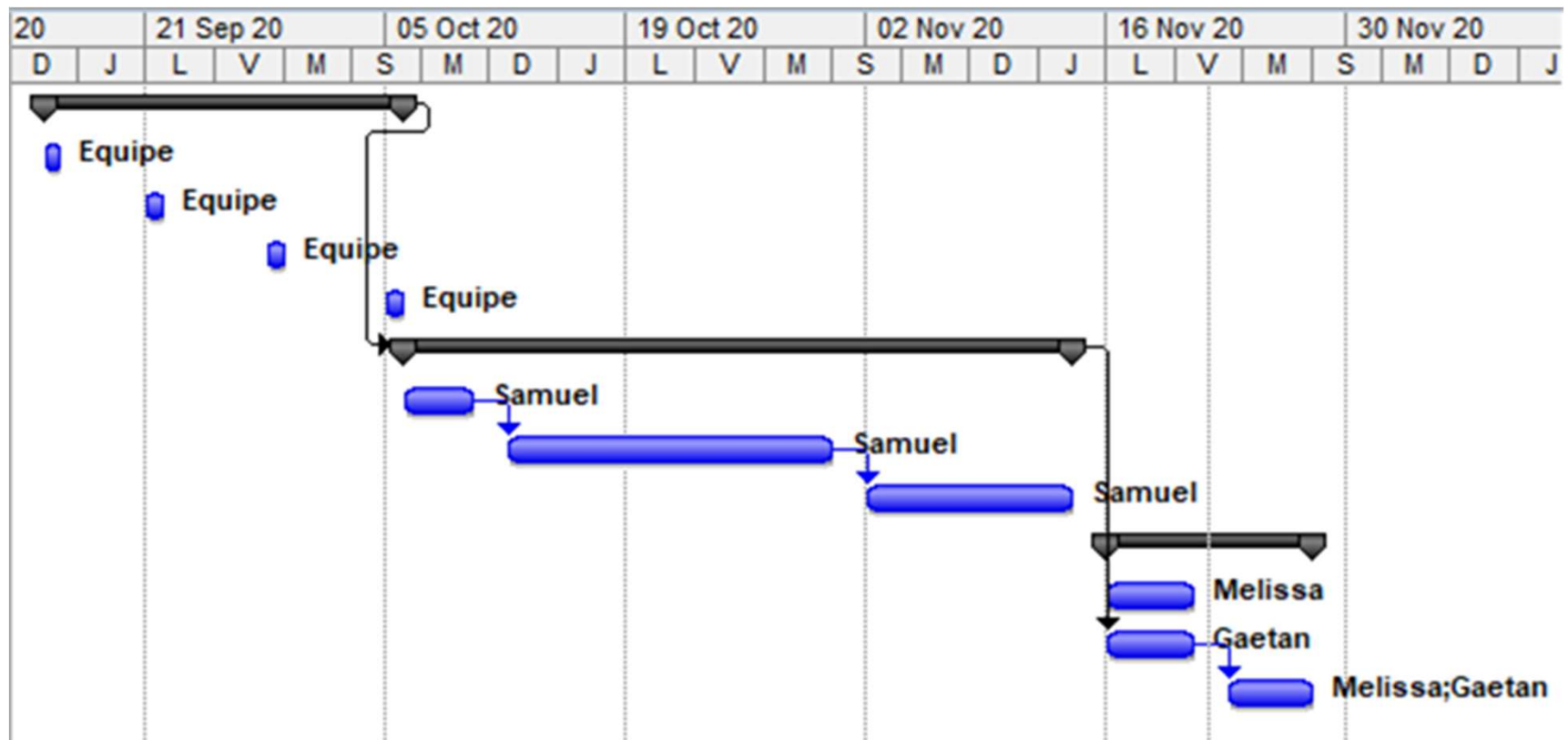
# 01 Organisation de l'équipe

## Répartition des missions



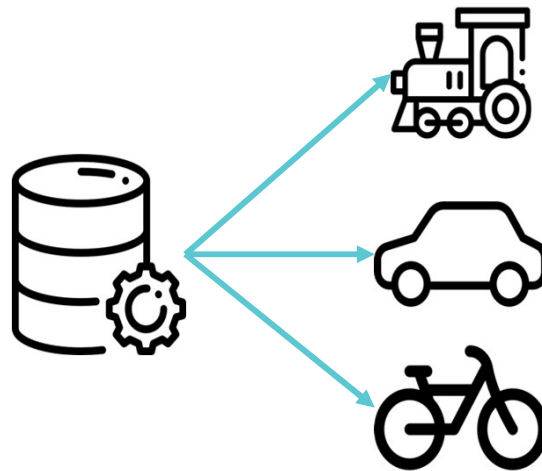
# 01 Organisation du projet

## Et concrètement...



## 02 Choix du sujet

# Le choix du sujet



02 Choix du sujet  
**L'accord trouvé**



**⇒ Etudes des valeurs immobilières des propriétés françaises**

## 02 Choix du sujet

# Meta-datas du jeu de données initial



**Licence  
ouverte**



**Données disponibles  
pour cinq ans**



**Données mise à jour de  
manière semestrielle**



**Demande valeurs foncières (DVF)**

**Normes CSV**

**2 500 000 évaluations immobilières pour 2019**

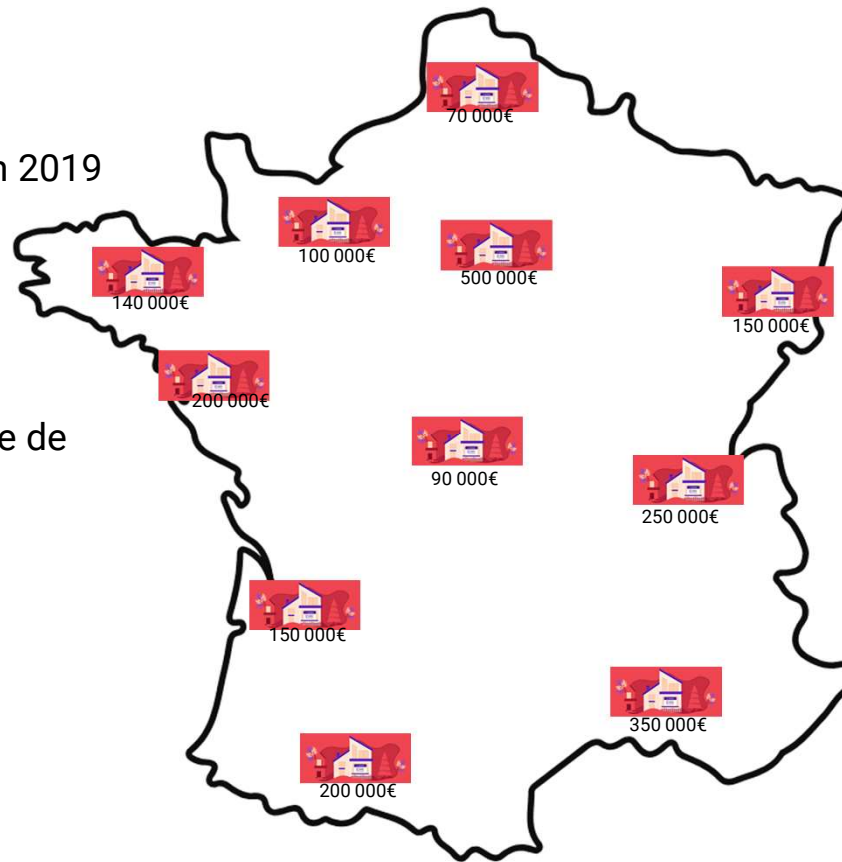
**Lors de mutations immobilières**



## 02 Choix du sujet

# Le sujet en détails

- 2 500 000 biens vendus en 2019
- Adresse complète
- Valeurs immobilières
- D'autres informations (type de bien, date, surface, etc..)



## 02 Choix du sujet

# Les données

### Nature

- |Nature culture
- |Code nature culture
- |Nature culture speciale
- |Code nature culture speciale

### Lot

- |1er lot (L1)
- |Surface Carrez L1
- |2eme lot (L2)
- |Surface Carrez L2
- |3eme lot (L3)
- |Surface Carrez L3
- |4eme lot (L4)
- |Surface Carrez L4
- |5eme lot (L5)
- |Surface Carrez L5
- |Nombre de lots

### Immobilisation

- |ID mutation
- |Date mutation
- |Numero disposition
- |Nature mutation
- |Valeur foncière
- |Surface reelle bati
- |Nombre pieces principales
- |Surface terrain
- |Numéro Volume
- |Code type local
- |Type local

### Localisation

- |Adresse numero
- |Adresse suffixe
- |Adresse code voie
- |Adresse nom voie
- |Code postal
- |Nom Commune
- |Ancien nom commune
- |Code departement
- |Code commune
- |Ancien code commune
- |ID parcelle
- |Ancien ID parcelle

## 02 Choix du sujet

# Les données en détails

### Nature

|Nature culture  
|Code nature culture  
|Nature culture speciale  
|Code nature culture speciale

### Lot

|1er lot (L1)  
|Surface Carrez L1  
|2eme lot (L2)  
|Surface Carrez L2  
|3eme lot (L3)  
|Surface Carrez L3  
|4eme lot (L4)  
|Surface Carrez L4  
|5eme lot (L5)  
|Surface Carrez L5  
|Nombre de lots

### Immobilisation

|ID mutation  
|Date mutation  
|Numero disposition  
|Nature mutation  
|Valeur foncière  
  
|Surface reelle bati  
|Nombre pieces principales  
|Surface terrain  
|Numéro Volume  
  
|Code type local  
|Type local

### Localisation

|Adresse numero  
|Adresse suffixe  
|Adresse code voie  
|Adresse nom voie  
|Code postal  
|Nom Commune  
|Ancien nom commune  
|Code departement  
|Code commune  
|Ancien code commune  
|ID parcelle  
|Ancien ID parcelle

### Légende:

- **Données** : Données toujours présentes
- **Données** : Données souvent présentes
- **Données** : Données souvent absentes
- **Données** : Données toujours absentes

## 02 Choix du sujet

# Les données utilisées

### Nature

Nature culture  
Code nature culture  
Nature culture speciale  
Code nature culture speciale

### Lot

1er lot (L1)  
Surface Carrez L1  
2eme lot (L2)  
Surface Carrez L2  
3eme lot (L3)  
Surface Carrez L3  
4eme lot (L4)  
Surface Carrez L4  
5eme lot (L5)  
Surface Carrez L5  
Nombre de lots

### Immobilisation

ID mutation  
Date mutation  
Numero disposition  
Nature mutation  
Valeur foncière  
  
Surface reelle bati  
Nombre pieces principales  
Surface terrain  
Numéro Volume  
  
Code type local  
Type local

### Localisation

Adresse numero  
Adresse suffixe  
Adresse code voie  
Adresse nom voie  
Code postal  
Nom Commune  
Ancien nom commune  
Code departement  
Code commune  
Ancien code commune  
ID parcelle  
Ancien ID parcelle

### Légende:

- Données : Données que nous avons utilisées
- Données : Données que nous pourrions utiliser plus tard
- Données : Données non exploitées

## 02 Choix du sujet

# Validation des préceptes de l'Open Data

- Complètes



## 02 Choix du sujet

# Validation des préceptes de l'Open Data

- **Complètes**
- **Brutes**



## 02 Choix du sujet

# Validation des préceptes de l'Open Data

- **Complètes** ✓
- **Brutes** ✓
- **À jour** ✓

## 02 Choix du sujet

# Validation des préceptes de l'Open Data

- **Complètes** ✓
- **Brutes** ✓
- **À jour** ✓
- **Accessibles à tous** ✓



## 02 Choix du sujet

# Validation des préceptes de l'Open Data

- **Complètes** ✓
- **Brutes** ✓
- **À jour** ✓
- **Accessibles à tous** ✓
- **Traitement peut être automatisé** ✓

## 02 Choix du sujet

# Validation des préceptes de l'Open Data

- **Complètes** ✓
- **Brutes** ✓
- **À jour** ✓
- **Accessibles à tous** ✓
- **Traitement peut être automatisé** ✓
- **Accès non discriminant** ✓

## 02 Choix du sujet

# Validation des préceptes de l'Open Data

- **Complètes** ✓
- **Brutes** ✓
- **À jour** ✓
- **Accessibles à tous** ✓
- **Traitement peut être automatisé** ✓
- **Accès non discriminant** ✓
- **Format non propriétaire** ✓

## 02 Choix du sujet

# Validation des préceptes de l'Open Data

- **Complètes** ✓
- **Brutes** ✓
- **À jour** ✓
- **Accessibles à tous** ✓
- **Traitement peut être automatisé** ✓
- **Accès non discriminant** ✓
- **Format non propriétaire** ✓
- **Licence libre** ✓

## 02 Choix du sujet

# Validation des préceptes de l'Open Data

- **Complètes** ✓
- **Brutes** ✓
- **À jour** ✓
- **Accessibles à tous** ✓
- **Traitement peut être automatisé** ✓
- **Accès non discriminant** ✓
- **Format non propriétaire** ✓
- **Licence libre** ✓
- **Permanente** ✓

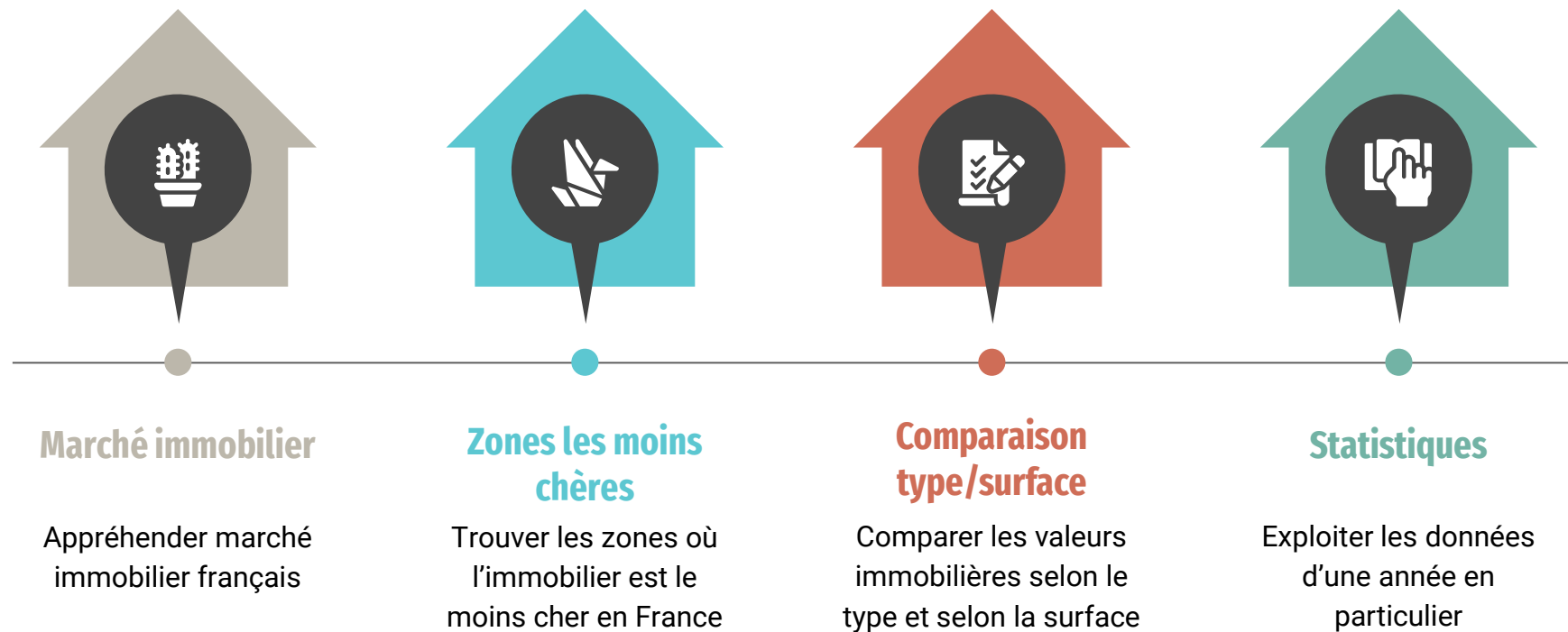
## 02 Choix du sujet

# Validation des préceptes de l'Open Data

- **Complètes** ✓
- **Brutes** ✓
- **À jour** ✓
- **Accessibles à tous** ✓
- **Traitement peut être automatisé** ✓
- **Accès non discriminant** ✓
- **Format non propriétaire** ✓
- **Licence libre** ✓
- **Permanente** ✓
- **Gratuite** ✓

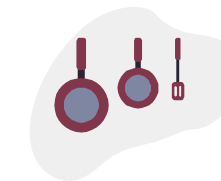
## 03 Traitement des données

# Les enjeux du projet



## 03 Traitement des données

# Les enjeux du projet



### Une carte avec l'emplacement de chaque bien

Coordonnées de tous les points  
+ Filtrage

### Une carte avec la valeur médiane de chaque département

Regroupement des lignes  
+ Mise en forme dynamique

### Une carte avec la valeur médiane de chaque région

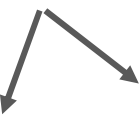
Regroupement des lignes  
+ Mise en forme dynamique



## 03 Traitement des données

# Les enjeux du projet

**Riche :**  
**Plusieurs évaluations**

  
**Nombre**      **Valeur**

### Immobilisation

- |ID mutation
- |Date mutation
- |Numero disposition
- |Nature mutation
- |Valeur foncière
- |Surface reelle bati
- |Nombre pieces principales
- |Surface terrain
- |Numéro Volume
- |Code type local
- |Type local

### Localisation

- |Adresse numero
- |Adresse suffixe
- |Adresse code voie
- |Adresse nom voie
- |Code postal
- |Nom Commune
- |Ancien nom commune
- |Code departement
- |Code commune
- |Ancien code commune
- |ID parcelle
- |Ancien ID parcelle

## 03 Traitement des données

# Applications similaires



**SeLoger**



**Le Bon Coin**



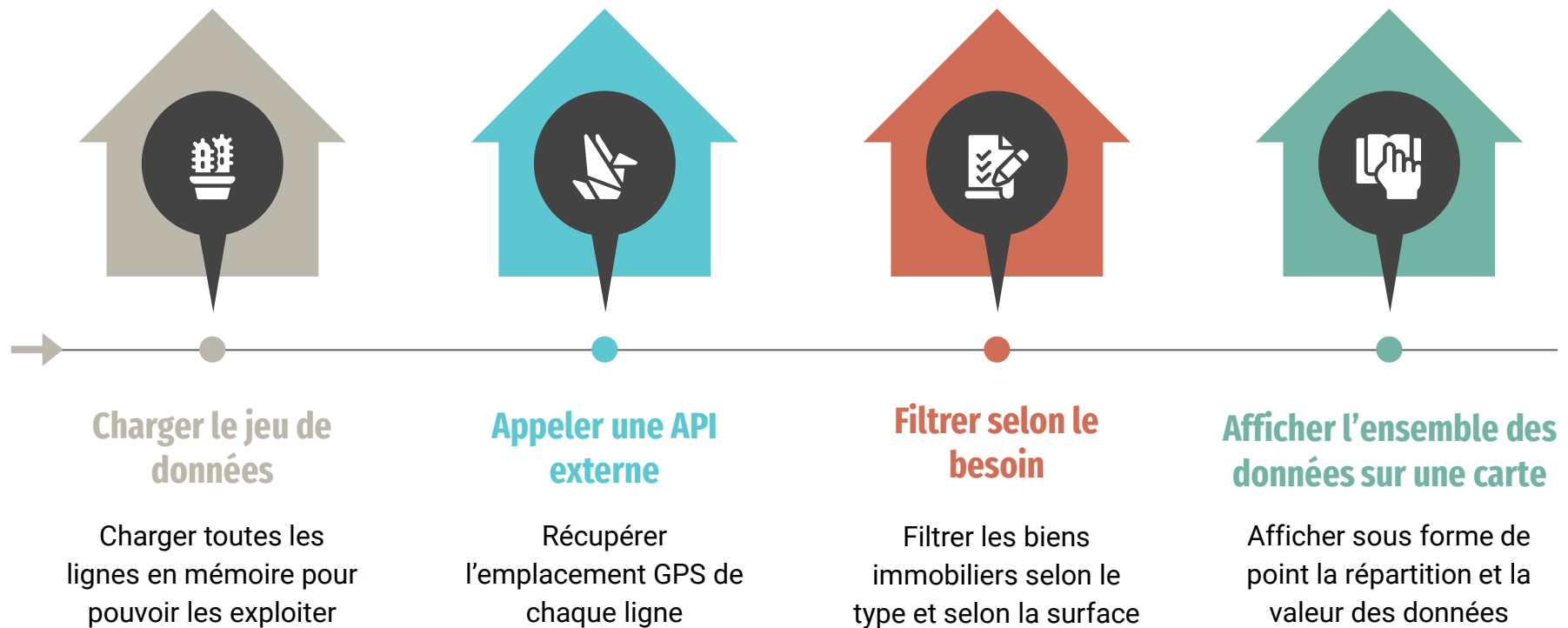
**Prix Immo**



**[app.dvf.etalab.gouv.fr/](https://app.dvf.etalab.gouv.fr/)**

## 03 Traitement des données

# Conception initiale

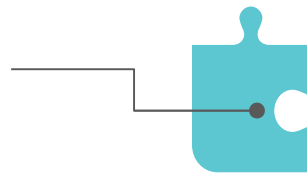


## 03 Traitement des données

# Déroulement général

- Plusieurs fonctions

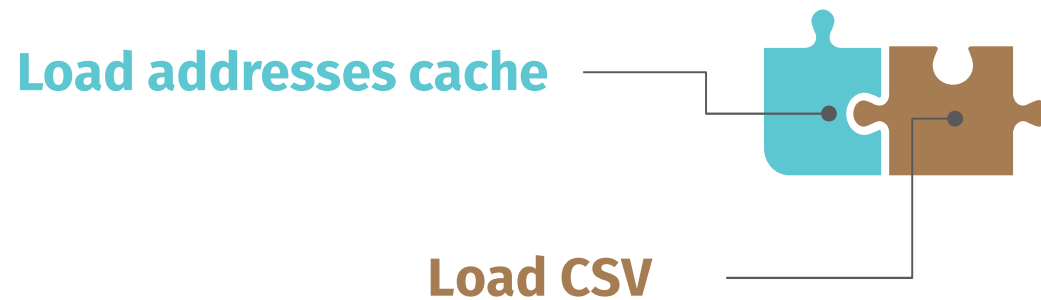
Load addresses cache



## 03 Traitement des données

# Déroulement général

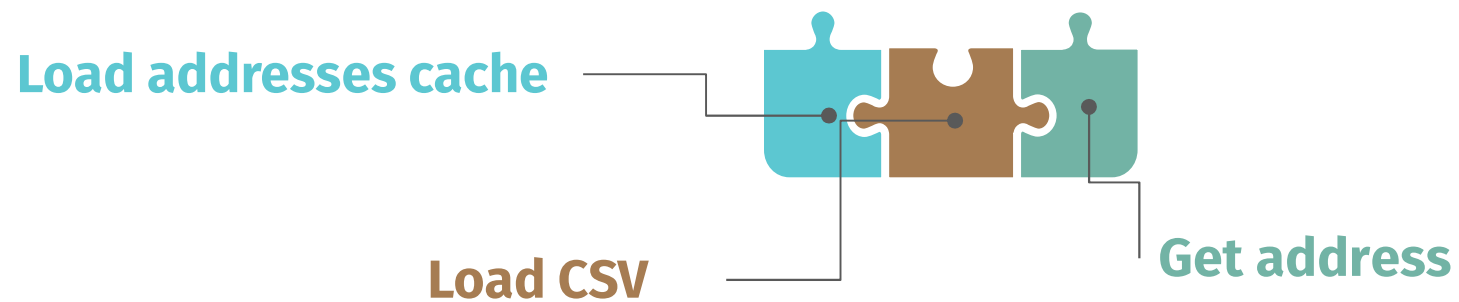
- Plusieurs fonctions



## 03 Traitement des données

# Déroulement général

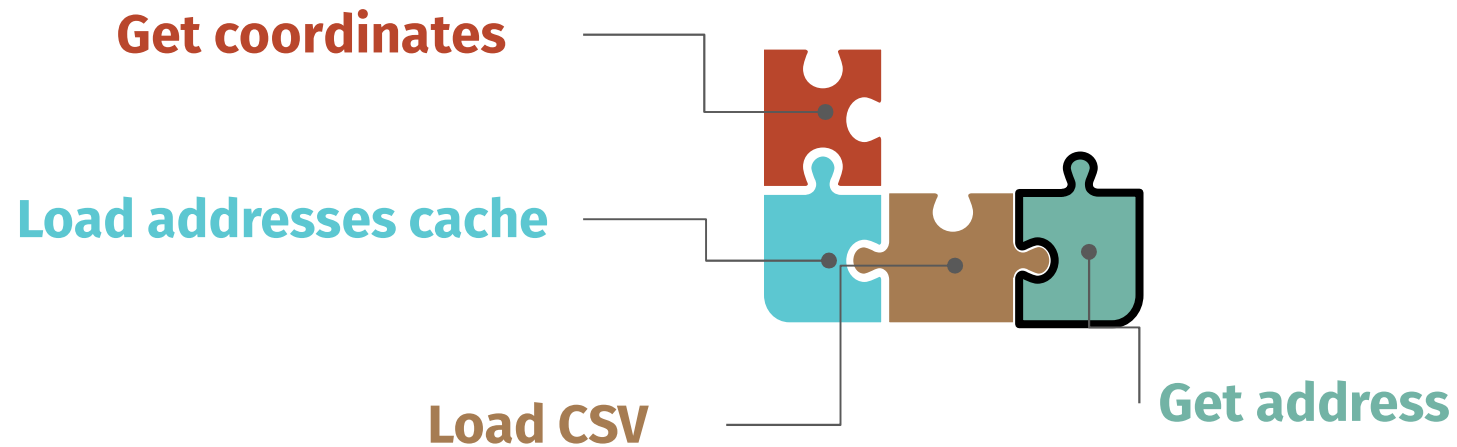
- Plusieurs fonctions



## 03 Traitement des données

# Déroulement général

- Plusieurs fonctions

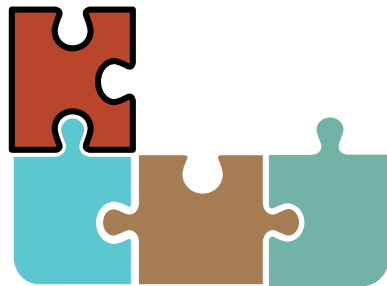


## 03 Traitement des données

# Déroulement général

### API de coordonnées GPS

#### Get coordinates



API Adresse

#### Adresse au format standard

Ex: 2, rue de la Houssinière  
44322 Nantes Cedex 3



#### Coordonnées GPS

EX: **score:** 0.5118764335664335,  
**x:** 355581.01,  
**y:** 6692055.75,



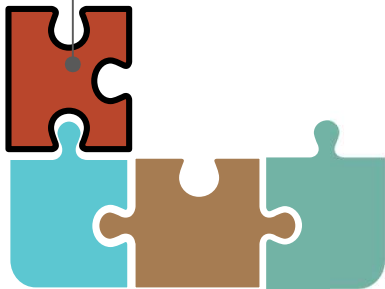
## 03 Traitement des données

# Déroulement général

### Prototype



Get coordinates



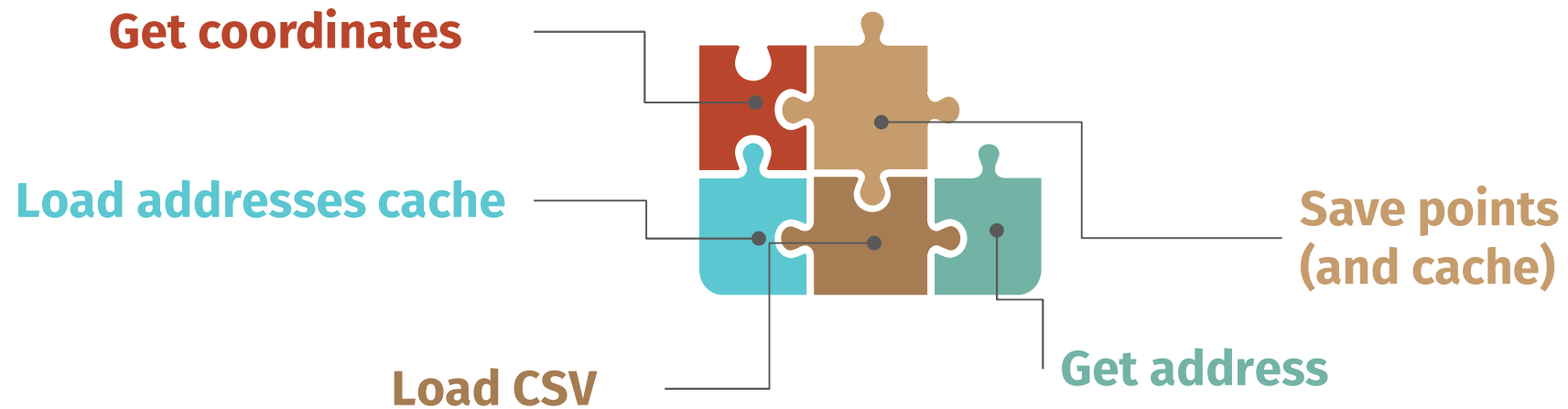
```
import pandas as pd
import requests

df = pd.read_csv(
    os.path.join(data_directory, "valeursfoncieres-2019.txt"),
    delimiter="|",
    encoding="utf-8",
)
first_address = get_address_from_row(df.head(0))
>> 8 rue sylvestre chauvelot 21200
response = requests.get(f"https://api-adresse.data.gouv.fr/search?q={first_address}")
>> <Response : 200 >
coordinates = extract_coordinates(response)
>> {"longitude": 5.22565, "latitude": 4.25232}
```

## 03 Traitement des données

# Déroulement général

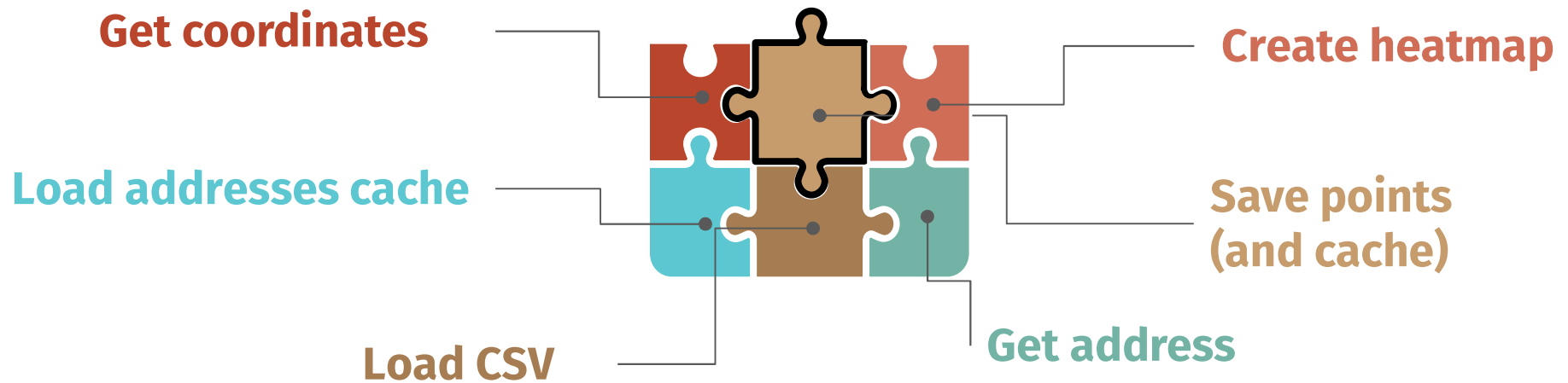
- Plusieurs fonctions



## 03 Traitement des données

# Déroulement général

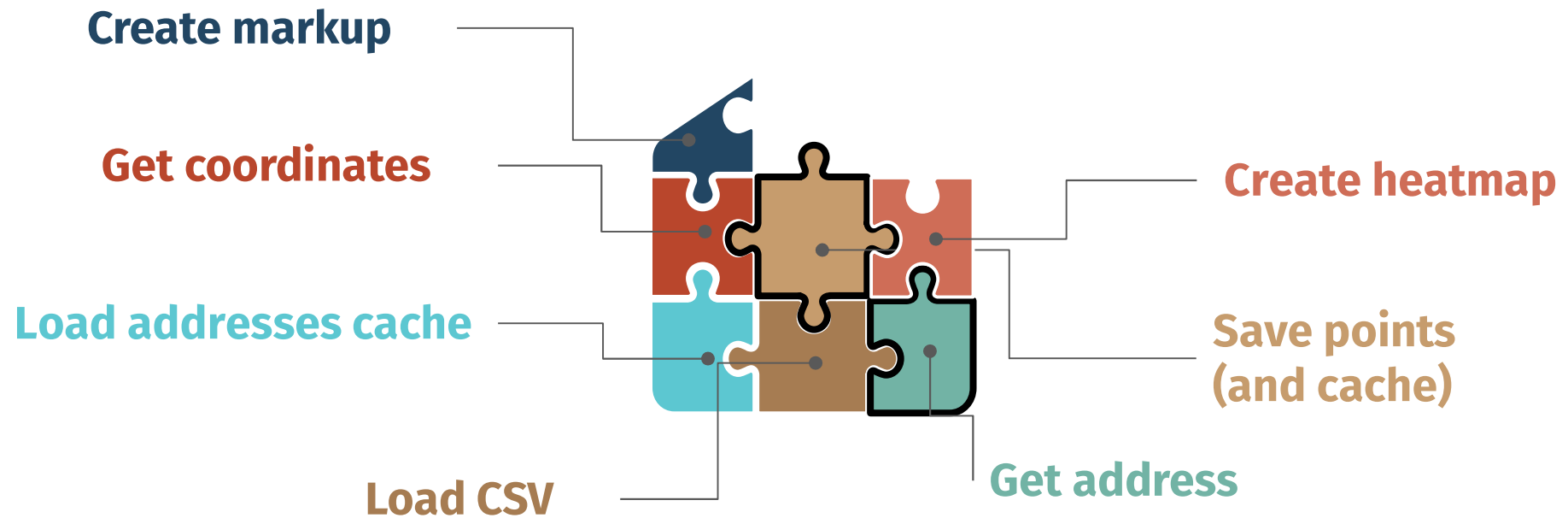
- Plusieurs fonctions



## 03 Traitement des données

# Déroulement général

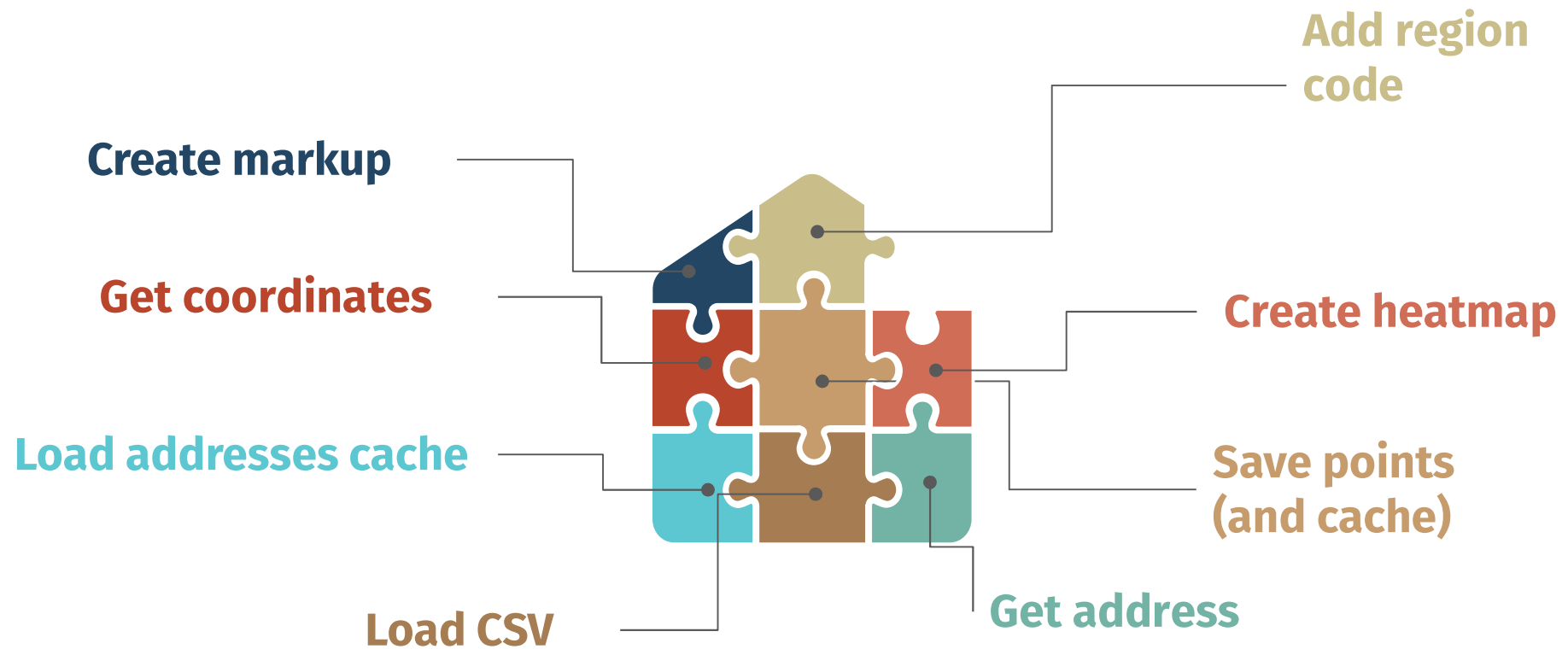
— Plusieurs fonctions



### 03 Traitement des données

## Déroulement général du processus

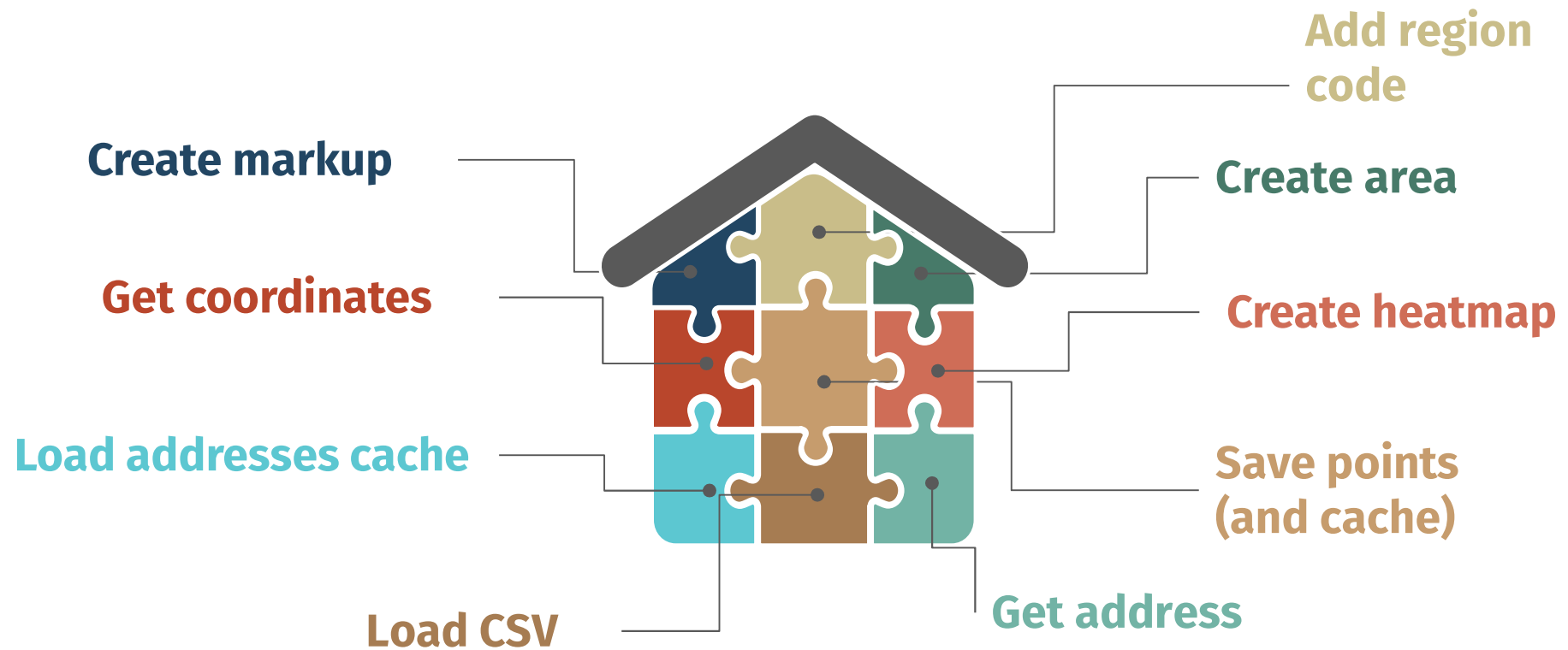
— Plusieurs fonctions



## 03 Traitement des données

# Déroulement général

— Plusieurs fonctions



## 03 Traitement des données

# Déroulement général



- Complexité en  $O(n)$

## 03 Traitement des données

# Problèmes rencontrés

### Problèmes

- Redondance des données
- Données partiellement exploitables
- Lenteur du chargement des données

### Solutions apportées

- Utilisation d'un cache
- Bibliothèque pandas



## 03 Traitement des données

# Première exploitation

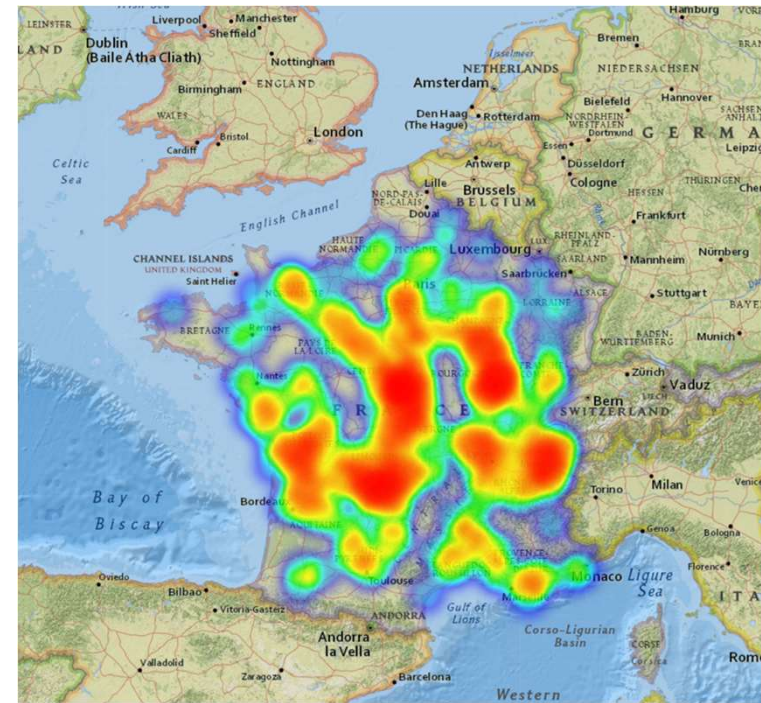
### Répartition des biens immobiliers recensés

```
import folium
from folium import plugins

def create_distribution_heatmap(name: str, lats: list, lons: list):
    url_base = "http://server.arcgisonline.com/ArcGIS/rest/services/"
    service = "NatGeo_World_Map/MapServer/tile/{z}/{y}/{x}"
    tileset = url_base + service

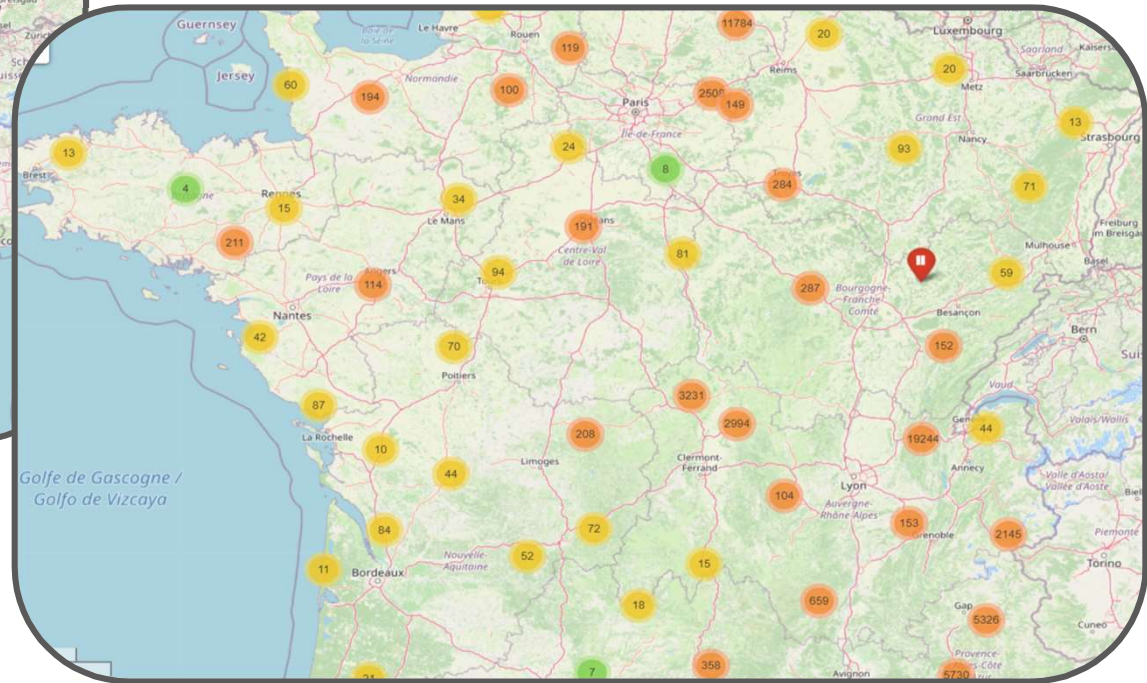
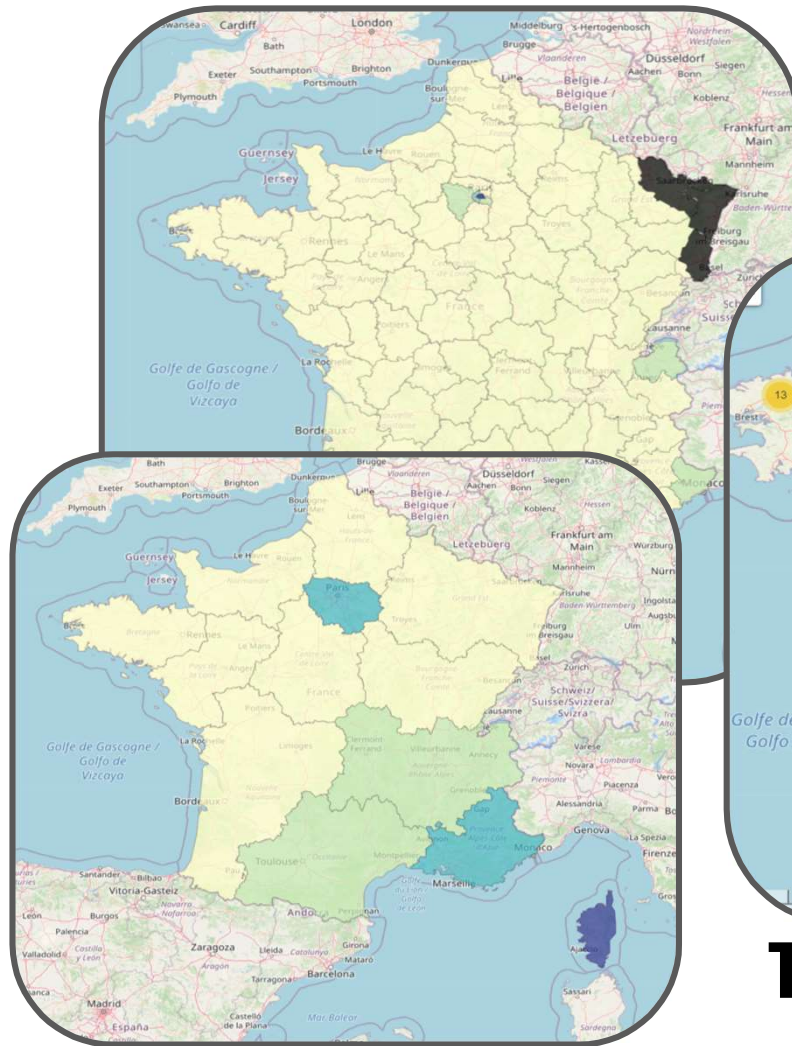
    heatmap_map = folium.Map(
        location=[48.52, 2.19],
        zoom_start=2,
        control_scale=True,
        tiles=tileset,
        attr="USGS style",
    )
    data = [lon_lat for lon_lat in zip(lats, lons)]

    heatmap_map.add_child(plugins.HeatMap(data))
    heatmap_map.save(name)
```



## 04 Résultats

# Nouvelles cartes



**Tout de suite ⇒ Démonstration !**

## 03 Traitement des données

# Evolutions possibles



01

### Evolution dans le temps

Une carte montrant l'évolution des données sur plusieurs années



02

### Filtre sur les types

Une carte permettant de voir pour chaque type les valeurs des mutations



03

### Filtre sur les types et surfaces

Une carte filtrant par type, et selon la surface



**Merci de votre attention !**

