

# Analyse spatiale et territoriale de données de recensement

Formation Carthageo-Geoprisme 2021 / 1ere journée

C.GRASLAND

15/10/2020

## Section 1

# Données RPLS

Le répertoire des logements locatifs des bailleurs sociaux (RPLS) a pour objectif de dresser l'état global du parc de logements locatifs de ces bailleurs sociaux au 1er janvier d'une année. Il est alimenté par les informations transmises par les bailleurs sociaux. La transmission des informations pour la mise à jour annuelle du répertoire des logements locatifs est obligatoire. Les données sont ensuite géolocalisées à l'adresse et mis à disposition des utilisateurs sur le **site du ministère de la transition écologique**

Les fichiers sont disponibles en général par régions mais livrés par départements dans le cas de l'Île de France. Nous allons utiliser ici le fichier du 1er janvier 2020 accessible à l'adresse suivante

<https://www.statistiques.developpement-durable.gouv.fr/le-parc-locatif-social-au-1er-janvier-2020-0>

# Métadonnées

Le fichier de données brutes au format .csv est accompagné d'un document excel précisant le code des variables et la façon dont elles ont été obtenues.

B	C	D	E
Complément d'identification du bâtiment autre	<u>COMPLGEO</u>		
Lieu dit	<u>LIEUDIT</u>		
Code logement situé en Quartier prioritaire de la politique de la ville	<u>QPV</u>	1 = Logement en QPV 2 = Logement hors QPV	Cette donnée est issue de la déclaration des bailleurs et non de la géolocalisation.
Code type de construction	<u>TYPECONST</u>	I = Individuel C = Collectif E = Étudiant NC = Non conforme	Non conforme : valeur renseignée autre que I, E ou C
Code nombre de pièces	<u>NBPIECE</u>	1 = 1 pièce 2 = 2 pièces 3 = 3 pièces 4 = 4 pièces 5 = 5 pièces 6 = 6 pièces 7 = 7 pièces 8 = 8 pièces 9 = 9 pièces ou plus 0 = Non renseigné NC = Non conforme	Non conforme : valeur renseignée autre que 1 à 9
Surface habitable (m²) - pour calcul	<u>SURFHAB</u>	Exemple : 50 0 : Non renseigné Non conforme	<ul style="list-style-type: none"> <li>Pour les logements dont la surface habitable est renseignée : <b>valeur surface</b></li> <li>Non conforme : lorsque la surface habitable est un nombre de plus de 9</li> <li>Pour les logements dont la surface habitable est non renseignée : <b>0 : Non renseigné</b></li> </ul>
Année d'achèvement de la construction	<u>CONSTRUCT</u>	Exemple : 1950 Non renseigné Non conforme	Non conforme : Pour les logements construits avant 1000
Année de première mise en location du logement	<u>LOCAT</u>	Exemple : 2010 Non renseigné Non conforme	Non conforme : Pour les logements construits avant 1850
Année d'entrée du logement dans patrimoine locatif du bailleur	<u>PATRIMOINE</u>	Exemple : 2010 Non renseigné Non conforme	Non conforme : Pour les logements construits avant 1850

Le fichier indique pour chaque logement sa localisation précise en terme d'adresse mais aussi d'étage dans un immeuble. A partir de ces données qualitatives, l'INSEE a procédé à un géocodage qui aboutit à la création de deux champs :

- coordonnées de latitude et longitude non projetées
- coordonnées de position en projection Lambert officielle

Selon les analyses on peut utiliser l'une ou l'autre de ces coordonnées. Mais la meilleur solution consiste à **créer un fichier de type sf (spatial features)** en coordonnées WGS94 qu'on pourra ensuite reprojeter dans le système de son choix.

Avant toute exploitation du fichier il est fortement recommandé d'analyser en détail les métadonnées et de définir une stratégie d'analyse.

- ❶ **choisir une première zone d'étude** de petite taille et localisée de préférence dans un espace que l'on connaît bien.
- ❷ **choisir des variables intéressantes** dont l'on connaît bien la signification et dont on a analysé en détail les métadonnées
- ❸ **vérifier la qualité des données** en regardant notamment le nombre de valeurs manquantes, le degré de précision, etc.
- ❹ **sélectionner des données auxiliaires** issues d'autres sources que l'on souhaite croiser avec celles du RPLS en s'assurant de leur compatibilité (espace, temps, définition, ...)
- ❺ **Ajouter les coordonnées spatiales** et stocker le résultat dans un fichier de type sf comportant les indications de projection.

# Importation du fichier

Le fichier initial au format .csv a été importé à l'aide de la fonction `fread()` du package `data.table` car elle est rapide et relativement robuste face aux erreurs de codage. Le résultat a été stocké dans un fichier .RDS pour réduire sa taille. On peut désormais le charger rapidement avec `readRDS()`