

Semaine mutualisée Master 2 Carthagéo-Géoprisme

Hadrien Commenges, Claude Grasland, Pierre Pistre

Du 09/10/2023 au 13/10/2023

Objectifs de la semaine

- Développer les connaissances en analyse quantitative des données dans l'espace
- Exploiter des données individuelles (recensement de la population, demandes de valeurs foncières)
- Approfondir l'utilisation du logiciel R
- Travailler en groupe et préparer une restitution commune
- Répondre à une "commande" dans un temps limité

Organisation de la semaine

Planning des intervenants et contenu des journées

- 09/10 (matin, Pierre Pistre) : Organisation de la semaine & présentation des données
- 09/10 (après-midi, Claude Grasland) : Prise en main des données de recensement (fichier LOGEMT)
- 10/10 (matin, Pierre Pistre) : Prise en main des données de valeurs foncières (DVF)
- 10/10 (après-midi, Claude Grasland) : extractions des données pour la commande, et approfondissements statistiques & cartographiques
- 11/10 : Accompagnement des projets de groupe (Hadrien Commenges)
- 12/10 : Accompagnement des projets de groupe (Hadrien Commenges)
- 13/10 : Finalisation des projets de groupe (matin) et restitution orale (après-midi)

Horaires et planning des salles (Campus des Grands Moulins, Université Paris Cité)

- Du 09/10 au 11/10 : salle de cours 209 (Bâtiment Olympe de Gouges, 2ième étage)
- 12/10 : salle informatique 377 (Bâtiment Olympe de Gouges, 3ième étage)
- 13/10 : salle de cours 209 (Bâtiment Olympe de Gouges, 2ième étage)

Données utilisées

(sources) Base de données principales :

- Fichier "Logements ordinaires" (LOGEMT) 2019 (localisations : IRIS, communes...), produit par l'INSEE à partir du recensement de la population. Données en accès libre : <https://www.insee.fr/fr/statistiques/6544344?sommaire=6456104>
- "Demandes de valeurs foncières" (version Etalab et extraction avec Opendatasoft), produit par le Ministère de l'Économie, des Finances et de la Souveraineté industrielle et numérique : <https://www.data.gouv.fr/fr/datasets/demandes-de-valeurs-foncieres-geolocalisees/>

(sources) Fichiers complémentaires :

- Table d'appartenance géographique des communes (INSEE) : <https://www.insee.fr/fr/information/2028028>
- Shapefile ADMIN-EXPRESS-COG® (IGN) des découpage administratifs de la France métropolitaine : <https://geoservices.ign.fr/adminexpress#telechargement>

- Shapefile Contours IRIS® (IGN) : <https://geoservices.ign.fr/contoursiris>

Exercice d'application en groupe

Périmètres de l'exercice

- Objet d'étude : les dynamiques immobilières dans les principales agglomérations administratives de la "mégarégion parisienne" (hors Paris)
- Population d'étude : les logements vendus récemment et leurs habitants
- Espace(s) d'étude : les agglomérations ayant le statut de "métropole" ou de "communauté urbaine" (hors Paris)
- Echelles et mailles géographiques : département, commune, IRIS

Consignes de la "commande"

- Contexte général : il est connu que Paris influence la dynamique des territoires bien au-delà de ses territoires les plus proches, à commencer par les principales agglomérations urbaines de la "mégarégion parisienne" (au sens élargi de : <https://atlas-paris-mega-region.univ-rouen.fr/>). Le secteur immobilier est particulièrement concerné du fait d'une progression sensible des prix des logements dans la plupart de ces territoires au cours à minima de la dernière décennie et d'un potentiel renforcement avec la crise COVID par une progression des installations résidentielles. Pour autant, la géographie des dynamiques immobilières internes à ces agglomérations urbaines reste mal connue de même que les facteurs explicatifs (influence parisienne, spécificités locales et régionales...) et les tendances récentes.
- Commande : après avoir fait le constat de dynamiques immobilières souvent proches, les intercommunalités de statut "métropole" et "communauté urbaine" de la "mégarégion parisienne" ont décidé de s'associer pour demander aux différentes antennes régionales de l'Insee la réalisation d'études rigoureuses et fines spatialement sur la situation immobilière de leurs différentes agglomérations et périphéries proches. Le bilan de ces études sera présenté dans des publications synthétiques, intelligibles par le plus grand nombre et par les acteurs concernés, sous la forme de plusieurs 4 pages telle qu'utilisée régulièrement par l'Insee pour publier les résultats de ses études (par exemple, <file:///home/pierre/T%C3%A9l%C3%A9chargements/ip1715.pdf> ; file:///home/pierre/T%C3%A9l%C3%A9chargements/IR134_Notaires-IPLA_1T-23.pdf).

Modalités de l'exercice, restitution du travail et organisation durant la semaine

1. Constitution de 6 (voire 7) groupes de 4-5 étudiant.e.s, mélangeant les profils de Master (= Lundi matin)
2. Réflexion sur un angle problématique pour chaque cas d'étude, ainsi que les possibilités d'analyse et d'exploration des données : variables pertinentes, traitements envisagés... (= Lundi à Mercredi)
3. Réalisation des analyses statistiques et cartographiques (= Mardi à Jeudi)
4. Organisation, mise en forme des analyses et rédaction de la note écrite (= Mardi à Vendredi matin)
5. Présentation orale (environ 10 minutes, sans autre support visuel que la note écrite) devant un "jury" composé des intervenants de la semaine (= Vendredi après-midi)

Cas d'étude

- 1) En priorité :
 - Caen (communauté urbaine)
 - Le Mans (communauté urbaine)
 - Orléans (métropole)
 - Reims (communauté urbaine)
 - Rouen (métropole)
 - Tours (métropole)
- 2) Secondairement :

- Alençon (communauté urbaine)
- Le Havre (communauté urbaine)
- Mantes-la-Jolie (communauté urbaine)

Format du rendu (type 4 pages Insee)

- Introduction et définition (thématique et statistique) de l’objet : thème et cas d’étude
- Etat des lieux général des prix immobiliers dans les cas d’étude (données : DVF ; échelle : “agglomération” (rayon 50km) ; maille : commune)
- Approfondissements sur les prix immobiliers par type de biens ou de localisation (données : DVF ; échelles : “agglomération” (rayon 50km) et zooms ; mailles : commune, IRIS)
- Profils des acheteurs au sens des nouveaux arrivants (moins de 3 ans) qui sont propriétaires (données : recensement ; échelles : “agglomération” (rayon 50km) et zooms ; mailles : commune, IRIS)

Présentation des données DVF

La base Demandes de Valeurs Foncières (DVF) est produite par la direction générale des finances publiques. Les données sont disponibles en *open data* depuis 2018, elles contiennent l’ensemble des mutations foncières et immobilières qui ont lieu en France métropolitaine et outre-mer, à l’exception de l’Alsace, de la Moselle et de Mayotte. Pour une édition annuelle, elles contiennent toutes les mutations sur les 5 années précédentes.

Il existe actuellement trois bases principales :

- les données brutes publiées par la direction des finances publiques et disponibles sur data.gouv : <https://www.data.gouv.fr/fr/datasets/demandes-de-valeurs-foncieres/>
- les données retravaillées par Etalab, aussi disponibles sur data.gouv. Ces données sont nettoyées, enrichies et normalisées : renommage des colonnes, normalisation des codes Insee ou des noms de voies, géolocalisation long-lat.
- les données retravaillées par le Cerema, disponibles sur son site : Ces données sont mises à disposition dans un format de fichier spatial, sous forme de points et de polygones (parcelles cadastrales). Elles subissent un traitement plus important que les données Etalab, en particulier les opérations qui comportent plusieurs mutations y sont agrégées pour que chaque ligne du tableau contienne une seule opération.

Liste des champs des données Etalab :

- id_mutation : Identifiant de mutation (non stable, sert à grouper les lignes)
- date_mutation : Date de la mutation au format ISO-8601 (YYYY-MM-DD)
- numero_disposition : Numéro de disposition
- nature_mutation : Nature de la mutation
- valeur_fonciere : Valeur foncière (séparateur décimal = point)
- adresse_numero : Numéro de l’adresse
- adresse_suffixe : Suffixe du numéro de l’adresse (B, T, Q)
- adresse_code_voie : Code FANTOIR de la voie (4 caractères)
- adresse_nom_voie : Nom de la voie de l’adresse
- code_postal : Code postal (5 caractères)
- code_commune : Code commune INSEE (5 caractères)
- nom_commune : Nom de la commune (accentué)
- ancien_code_commune : Ancien code commune INSEE (si différent lors de la mutation)
- ancien_nom_commune : Ancien nom de la commune (si différent lors de la mutation)
- code_departement : Code département INSEE (2 ou 3 caractères)
- id_parcelle : Identifiant de parcelle (14 caractères)
- ancien_id_parcelle : Ancien identifiant de parcelle (si différent lors de la mutation)
- numero_volume : Numéro de volume
- lot_1_numero : Numéro du lot 1
- lot_1_surface_carrez : Surface Carrez du lot 1

- lot_2_numero : Numéro du lot 2
- lot_2_surface_carrez : Surface Carrez du lot 2
- lot_3_numero : Numéro du lot 3
- lot_3_surface_carrez : Surface Carrez du lot 3
- lot_4_numero : Numéro du lot 4
- lot_4_surface_carrez : Surface Carrez du lot 4
- lot_5_numero : Numéro du lot 5
- lot_5_surface_carrez : Surface Carrez du lot 5
- nombre_lots : Nombre de lots
- code_type_local : Code de type de local
- type_local : Libellé du type de local
- surface_reelle_bati : Surface réelle du bâti
- nombre_pieces_principales : Nombre de pièces principales
- code_nature_culture : Code de nature de culture
- nature_culture : Libellé de nature de culture
- code_nature_culture_speciale : Code de nature de culture spéciale
- nature_culture_speciale : Libellé de nature de culture spéciale
- surface_terrain : Surface du terrain
- longitude : Longitude du centre de la parcelle concernée (WGS-84)
- latitude : Latitude du centre de la parcelle concernée (WGS-84)