

Hackaviz 2021 : quartiers prioritaires (QP) d'Occitanie : caractéristiques et transactions immobilières.



Figure 1: 14 avril 2011, un immeuble en train d'être repeint à Empalot. Olybrius, Wikimedia Commons

Cette année, nous proposons dans un jeu de données inédit :

- les caractéristiques démographiques et économiques des quartiers prioritaires
- les transactions foncières du 1er janvier 2016 au 31 décembre 2020.
- les contours géographiques des communes et des quartiers prioritaires

Ce jeu de données est réparti sur trois fichiers. Il est possible de faire de belles visualisations à partir d'un seul de ces fichiers. Vous pouvez aussi les combiner, mais il n'est pas certain que la plus belle histoire ait besoin de toutes ces données.

Ces données sont extraites de sources pour lesquelles certaines valeurs ne sont pas renseignées. Avant de vous lancer dans des calculs d'indicateurs statistiques, pensez à examiner la qualité des données et le nombre d'échantillons considérés !

N'hésitez pas à nous contacter sur le discord du Toulouse DataViz pour discuter entre participants, si vous avez besoin d'aide à propos des données ou pour rapporter des bugs.

Bonne chance !

Description générale des données

Ces données sont disponibles sous forme de 4 fichiers, déclinés sous forme de données tabulaires au format `.csv` et géographique au format `.geojson`. Pour télécharger un de ces fichiers, placez la souris sur le nom du fichier, faites un clic droit et choisir “Enregistrer la cible du lien sous ...”.

Fichier de données	Contenu
<code>qp.csv</code>	caractéristiques des 105 quartiers prioritaires d’Occitanie
<code>foncier_qp.csv</code>	transactions immobilières dans ces quartiers entre 2016 et 2020
<code>communes.geojson</code>	contour géographique des communes d’Occitanie
<code>qp.geojson</code>	contour géographique des quartiers de ces communes

Une archive (`.zip`) est aussi disponible pour télécharger toutes les données.

Format de fichiers CSV

- Les fichiers sont encodés en UTF-8.
- Les fichiers `csv` utilisent le séparateur de colonne “,” et le caractère décimal “.”.
- Ces `csv` ont été exportés depuis à un paramètre régional anglais. Pensez à modifier les paramètres d’import de votre logiciel préféré !
- Si vous avez des difficultés à importer ces `csv`, nous vous proposons une alternative sous forme de fichier Microsoft Excel.

Fichier de données	Contenu
<code>qp.xlsx</code>	caractéristiques des 105 quartiers prioritaires d’Occitanie
<code>foncier_qp.xlsx</code>	transactions immobilières dans ces quartiers entre 2016 et 2020

Format de fichiers JSON

- Les latitudes et longitudes sont exprimées dans la projection `World Geodetic System 1984` (WGS 84).
- Elles sont tronquées sur 4 décimales.

Jointures

- Nous avons pris le soin d’unifier les noms de colonnes des différents fichiers afin que vous puissiez faire des jointures.
- Vous n’avez aucune obligation à fusionner ces fichiers : pensez d’abord à l’histoire que vous voulez raconter !
- Pour joindre deux `csv`, vous pouvez utiliser un modèle `ObservableHQ`, des librairies comme `Pandas` ou le coder directement.
- Pour fusionner des données géographiques `geojson` et un `csv`, si vous ne connaissez pas `QGIS` ou `R`, nous vous suggérons `Geo Data Merger` qui vous permet de faire cela via une interface graphique simple. Sélectionnez un fichier `geojson` puis un `csv` qui va venir enrichir le `geojson`. Désignez ensuite une colonne “clé” sur les 2 fichiers pour faire la jointure.

Format des valeurs

- Les codes de départements et de communes ont été formatés respectivement sur 2 et 5 caractères conformément au code officiel géographique.
- Les codes de quartiers sont formés par le préfixe “QP” et 6 chiffres.
- Les dates sont au format ISO-8601 (YYYY-MM-DD) ex: 2020-12-31.
- Les valeurs foncières sont exprimées en euros.
- Les surfaces sont en mètres carrés.

Description détaillée des données

Description du fichier qp.csv

Caractéristiques démographiques et économiques des quartiers prioritaires d’Occitanie

- Définitions
 - Quartier prioritaire (QP) de la politique de la ville : territoires en difficulté, cibles prioritaires de la politique de la ville, définis par la part de la population ayant un revenu inférieur à 11 250 € par an.
 - Communes englobantes : pour un quartier prioritaire, l’unité urbaine (une ou plusieurs communes) à laquelle le quartier appartient.
 - Population municipale : comprend les personnes ayant leur résidence habituelle sur le territoire de la commune, dans un logement ou une communauté; détenues dans les établissements pénitentiaires de la commune ; les sans-abri recensées sur le territoire de la commune ; résidant habituellement dans une habitation mobile recensée sur le territoire de la commune.
 - Population étrangère : toute personne résidant en France qui n’a pas la nationalité française.
 - Part parmi la population : calculé par rapport à la population municipale du QP.
 - Personnes couvertes par au moins une prestation CAF : allocataire + conjoint + enfants et autres personnes à charge.
 - Taux de pauvreté au seuil de 60% : part de la population sous le seuil de 60% du niveau de vie médian métropolitain. Les revenus sont calculés depuis les revenus fiscaux.
 - Taux de bas revenus : part de la population sous le seuil de 60 % du revenu déclaré par unité de consommation médian métropolitain. Les revenus sont issus des fichiers de la CAF. En savoir plus sur la différence entre le taux de pauvreté et de bas revenu : DOSSIER D’ETUDE N° 107 Août 2008 CAF, pp24-25.
 - Emploi précaire : contrat d’apprentissage, placé par une agence d’intérim, emplois-jeunes, CES, contrats de qualification, stagiaires rémunérés en entreprise, autres emplois à durée limitée
- Note sur la source des données
 - Le comptage de la population municipale (colonne ‘pop_com’) est une estimation de l’Insee et de la Caisse nationale des allocations familiales (Cnaf) en 2015. Ces valeurs diffèrent de celles du recensement INSEE effectué en 2016.
 - Les données sont calculées pour la période 2015-2020. Elles ont été publiées par l’INSEE pour l’année 2019, sauf pour les données concernant l’insertion professionnelle. Pour ces dernières, l’année 2020 est considérée, le fichier de l’année 2019 étant corrompu.

Cette description est disponible sous forme de fichier `.csv` `meta_qp.csv`.

colonne	description	type_valeur	exemple
code_qp	Code du quartier prioritaire (QP)	chaîne de caractères	QP082003
nom_qp	Libellé du QP	chaîne de caractères	Centre Ville
nom_commune	Libellé géographique de la (des) commune(s) englobante(s)	chaîne de caractères	Moissac
pop_mun	Population municipale du QP	entier	2470
pop_com_qp	Cumul de la population municipale de tous les QP de la (des) commune(s) englobante(s)	entier	3672
pop_com	Population municipale de la (des) commune(s) englobante(s)	entier	12564
ind_jeune	Indice de jeunesse : population de 0 à 19 ans / population de 60 ans et plus	nombre rationnel	0.9
tx_tot_et	Part des étrangers parmi la population	nombre rationnel	25.4
tx_f	Part des femmes parmi la population	nombre rationnel	50.6
tx_f_et	Part des étrangères parmi les femmes	nombre rationnel	22.8
percou	Nombre de personnes couvertes par au moins une prestation CAF	entier	1307
pmimp	Part des ménages imposés	nombre rationnel	24.7
part_chomprinc	Part des ménages dont l'origine principale du revenu déclaré repose sur des indemnités de chômage	nombre rationnel	8.1
decuc_q2	Médiane (en euros) du revenu déclaré par unité de consommation	entier	10136
decuc_d1	Premier décile du revenu déclaré par unité de consommation	entier	1404
decuc_d9	Neuvième décile du revenu déclaré par unité de consommation	entier	26388
disp_q2	Médiane (en euros) du revenu disponible par unité de consommation	entier	13305
disp_d1	Premier décile (en euros) du revenu disponible par unité de consommation	entier	6421
disp_d9	Neuvième décile (en euros) du revenu disponible par unité de consommation	entier	25058
tp60	Taux de pauvreté	nombre rationnel	43.9
brev	Taux de bas revenus	nombre rationnel	59.1
a	Nombre total de foyers allocataires percevant au moins une prestation CAF	entier	642
tx_tot_empl	Part des personnes de 15 à 64 ans actives ayant un emploi parmi les personnes de 15 à 64 ans	nombre rationnel	47.5
tx_f_empl	Part des femmes de 15 à 64 ans actives ayant un emploi parmi les femmes de 15 à 64 ans	nombre rationnel	41
tx_et_empl	Part des étrangers de 15 à 64 actifs ayant un emploi parmi les étrangers de 15 à 64 ans	nombre rationnel	36.7
tx_tot_eprec	Part des personnes en emploi précaire parmi les personnes ayant un emploi.	nombre rationnel	34.4
tx_f_eprec	Part des femmes en emploi précaire parmi les femmes ayant un emploi.	nombre rationnel	31.9
tx_et_eprec	Part des étrangers en emploi précaire parmi les étrangers ayant un emploi.	nombre rationnel	59.3
abcde	Nombre total de demandeurs d'emploi en fin de mois	entier	774
abcde_f	Nombre total de demandeurs d'emploi femmes	entier	327
abcde_h	Nombre total de demandeurs d'emploi hommes	entier	447
abc_capbep	Nombre de demandeurs d'emploi de catégorie A, B ou C de niveau de formation CAP-BEP	entier	187
abc_bac	Nombre de demandeurs d'emploi de catégorie A, B ou C de niveau de formation BAC	entier	111
abc_supbac	Nombre de demandeurs d'emploi de catégorie A, B ou C de niveau de formation supérieur au BAC	entier	64
ec_mat	Nombre d'écoles maternelles à moins de 100m du QP	entier	1
ec_elem	Nombre d'écoles élémentaires à moins de 100m du QP	entier	2
ec_elem_pri	Nombre d'écoles élémentaires privées à moins de 100m du QP	entier	1
coll	Nombre de collèges à moins de 300m du QP	entier	2
coll_pri	Nombre de collèges privés à moins de 300m du QP	entier	1
nbetab	Nombre d'établissements	entier	282

colonne	description	type_valeur	exemple
ens_sante	Nombre d'établissements de service aux particuliers : enseignement, santé et action sociale	entier	37
serv_par	Nombre d'établissements de service aux particuliers	entier	72
serv_tot	Nombre d'établissements de service : services aux entreprise et services aux particuliers	entier	137
auto_ent	Nombre d'auto-entrepreneurs parmi les créations d'établissements	entier	3

Description du fichier `foncier_qp.csv`

Transactions foncières du 1er janvier 2016 au 31 décembre 2020

- Afin de simplifier les données, nous n'avons conservé que les transactions sur les maisons, les appartements et les locaux industriels. Les dépendances ainsi que les transactions de terrains non bâtis ont été supprimées.
- Ce fichier contient des transactions entre particuliers mais aussi celles effectuées par des investisseurs privés (groupe immobilier) ou institutionnels (office du logement). Selon l'histoire que vous voulez raconter, c'est à vous de trouver un moyen de différencier ou d'éliminer des transactions.

Cette description est disponible sous forme de fichier `.csv` `meta_foncier_qp.csv`.

colonne	description	type_valeur	exemple
code_departement	Code département INSEE	chaîne de caractères	09
nom_departement	Nom du département	chaîne de caractères	Ariège
code_commune	Code commune INSEE	chaîne de caractères	09261
nom_commune	Nom de la commune	chaîne de caractères	Saint-Girons
code_qp	Code du quartier prioritaire	chaîne de caractères	QP009001
nom_qp	Libellé du quartier prioritaire	chaîne de caractères	Coeur De Ville
id_mutation	Identifiant de vente	chaîne de caractères	2019-100563
id_parcelle	Identifiant de parcelle (14 caractères)	chaîne de caractères	092610000D0167
longitude	Longitude du centre de la parcelle concernée	nombre rationnel	1.1438
latitude	Latitude du centre de la parcelle concernée	nombre rationnel	42.9831
nombre_lot	Nombre de lots	entier	0
date_mutation	Date de la vente	date	2019-11-25
jour_mutation	Jour de la vente	chaîne de caractères	25
mois_mutation	Mois de la vente	chaîne de caractères	11
annee_mutation	Année de la vente	chaîne de caractères	2019
nature_mutation	Type de mutation	chaîne de caractères	Vente
valeur_fonciere	Valeur foncière	entier	85000
code_type	Code de type de local	chaîne de caractères	1
type_local	Libellé du type de local	chaîne de caractères	Maison
surface_reelle_bati	Surface réelle du bâti	entier	75
nombre_pieces_principales	Nombre de pièces principales	entier	4
surface_terrain	Surface du terrain	entier	60

Description du fichier `communes.geojson`

Contour des communes

- Toutes nos excuses auprès des Séméacaises et Séméacais. Le contour de la commune a été omis dans le jeu de données. Ce contour de commune concerne le quartier prioritaire “Tarbes Est”, à cheval sur Tarbes et Séméac.

Cette description est disponible sous forme de fichier `.csv` `meta_communes.csv`.

nom colonne	description	type_valeur	exemple
code_commune	Code commune INSEE	chaîne de caractères	81021
code_departement	Code département INSEE	chaîne de caractères	81
nom_commune	Nom de la commune	chaîne de caractères	Aussillon

Description du fichier `qp.geojson`

Contour des quartiers prioritaires

Cette description est disponible sous forme de fichier `.csv` `meta_contours_qp.csv`.

nom colonne	description	type_valeur	exemple
code_qp	Code du quartier prioritaire	chaîne de caractères	QP066007
nom_qp	Nom du quartier prioritaire	chaîne de caractères	Bas-Vernet Nouveau QPV
commune_qp	Nom de la commune principale du QP	chaîne de caractères	Perpignan
code_commune	Code commune INSEE	chaîne de caractères	66136
nom_commune	Nom de ou des communes englobantes	chaîne de caractères	Perpignan
code_departement	Code département INSEE	chaîne de caractères	66

Sources des données

- Caractéristiques des quartiers prioritaires: données produites par l’INSEE, démographie 2015 et revenus 2017.
- Transactions foncières: données publiées et produites par la direction générale des finances publiques. Ces données énumèrent les transactions intervenues au cours des cinq dernières années sur le territoire métropolitain et les DOM-TOM, à l’exception de l’Alsace, de la Moselle et de Mayotte. Les données contenues sont issues des actes notariés et des informations cadastrales. Les données brutes sont disponibles [ici](#). Elles peuvent être explorées par l’application Etalab.
- Liste des communes: liste publiée par l’INSEE [ici](#).
- Contours des quartiers prioritaires: fichier `.shapefile` est disponible [ici](#).

Historique des versions

- Juin 2021: construction du jeu de données avec Python, Pandas et jupyter-lab et création des fichiers

- Juillet 2021: première version de la consolidation des fichiers et étude autour d'aggrégats sous R et Rstudio
- Août 2021: consolidation des fichiers et génération du readme et export sous R et Visual Studio Code