



Précautions techniques

LA FICHE EN BREF

Les données DVF permettent de faire une multitude de traitements. Il y a néanmoins plusieurs précautions à prendre pour utiliser les fichiers restitués par la DGFIP. En effet, les données sont complexes et le risque d'une interprétation erronée est bien présent. Il convient donc de préciser les contraintes techniques et les méthodes à prendre en compte.

Il n'y a pas une méthode unique pour traiter DVF. Il est donc nécessaire de mettre en place une coordination nationale pour éviter les discordances d'analyse et faire que DVF puisse véritablement jouer son rôle de référentiel commun, permettant notamment de comparer les données sur différents territoires.

LE SENS

Une base de données est une série structurée d'informations. Il faut toutefois lui donner du sens. Cela implique de bien maîtriser la base.

LE TRAITEMENT

Le traitement doit tenir compte de la méthode de construction de la base. Il passe par une intégration en tableur ou en base de données.

LA COHERENCE

La mise en commun d'outils et de méthodes permet de maintenir le caractère comparable des traitements au niveau national.

7.1 Donner du sens à la donnée

7.1.1 Un référentiel partagé



L'un des principaux intérêts de DVF est de pouvoir disposer d'un référentiel commun qui permette de comparer les territoires entre eux et de produire des études selon des méthodes homogènes. Cela suppose que tous les utilisateurs parlent de la même chose. Il y a donc un travail important à mener pour donner du sens à la donnée et pour que tous interprètent de la même manière les informations.

Cela passe d'abord par une clarification du vocabulaire (voir le glossaire associé à ce guide).

Par exemple les terrains à bâtir sont définis dans DVF au sens fiscal : la définition des terrains à bâtir retenue à la DGFIP est précisée au 1° du 2 du I de l'article 257 du Code Général des Impôts comme étant les terrains sur lesquels des constructions peuvent être autorisées en application d'un plan local d'urbanisme, d'un autre document d'urbanisme en tenant lieu, d'une carte communale ou de l'article L. 111-1-2 du code de l'urbanisme. Cette définition juridique a vocation à être confrontée aux situations opérationnelles locales.

7.1.2 Un effort d'interprétation

L'enjeu principal du traitement des données est de bien comprendre ce qu'elles signifient. Il y a ainsi un travail conséquent à fournir pour bien cerner DVF et éviter de sortir du champ de fiabilité de chaque variable renseignée.

Il convient de faire preuve de prudence car il y a un risque important d'erreur d'interprétation, notamment pour les mutations multilignes, qui comportent plusieurs dispositions.

7.2 Des mutations décrites sur plusieurs lignes

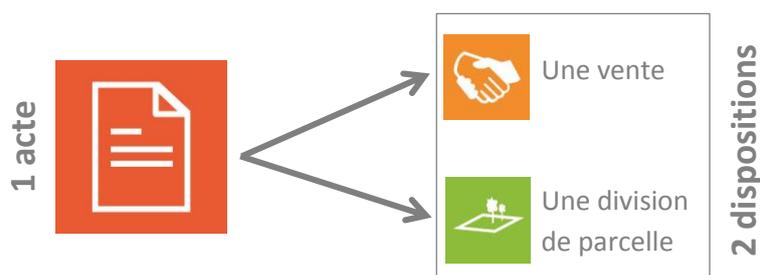
7.2.1 Qu'est-ce qu'une disposition ?

Certaines ventes, faisant l'objet d'un acte, présentent différentes parties juridiques. Elles sont appelées « dispositions ». Dans DVF, seules les dispositions faisant l'objet d'un prix sont restituées.

Une vente simple (environ 95 % des cas) est représentée par une seule disposition rattachée à un prix. Il y a alors concordance mutation/disposition.

Pour un même acte, plusieurs dispositions onéreuses peuvent coexister (c'est le cas pour environ 5 % des mutations), elles figurent donc dans DVF de manière distincte.

Ainsi, un acte comportant une vente et une division de parcelle donne lieu à la création de deux dispositions, l'une concernant la division et l'autre la vente car il s'agit de deux unités d'analyse juridique distinctes.



Dans cet exemple, la disposition relative à la division de parcelle ne sera pas restituée dans DVF, car elle ne fait pas l'objet d'un prix.

7.2.2 Les mutations simples décrites sur une seule ligne

Une mutation simple comprend une seule disposition représentée sur une seule ligne, comme dans l'exemple fictif suivant :

Reference document	No disposition	Valeur foncière	Section	No parcelle	Type local	Identifiant local	Nature culture	Nature culture spéciale	Surface terrain
2012P00174	1	179000	C	236	Maison	58639	Sol		780

↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓

Une mutation Une disposition = un prix Une parcelle pour la disposition 1 Un local pour la parcelle C236 Une nature de culture pour la parcelle C236 Une surface par nature de culture

En synthèse **la mutation 2012P00174** présentée dans DVF est d'un montant de **179 000 €**. Elle comprend **une maison** sur un terrain de 780 m².

Disposition 1



Maison
Local 58639
Parcelle C236

Mutation 2012P00174

7.2.3 Les mutations multilignes

Beaucoup de mutations comportent plusieurs lignes. Il faut alors gérer la duplication des informations et le dédoublement de certaines mutations.

La DGFIP décompose les mutations en plusieurs lignes pour les causes suivantes :

- Plusieurs dispositions
- Plusieurs parcelles dans la disposition
- Plusieurs locaux dans la parcelle
- Plusieurs natures de cultures (types de terrain du point de vue fiscal, aussi appelées subdivisions fiscales) dans la parcelle
- Plusieurs natures de cultures spéciales dans la nature de culture

A NOTER

Il résulte de ces mutations multilignes une impossibilité de réaliser des statistiques simples (nombre de mutations, prix moyen,...) sur la base brute. Une intégration en tableur ou en base de données est donc nécessaire afin de restituer une ligne par mutation. Voir fiche 9 *Structurer les données*

L'exemple fictif suivant permet d'illustrer ces mutations multilignes :

La disposition 1 n'est pas incluse dans DVF car il ne s'agit pas d'une disposition correspondant à un prix

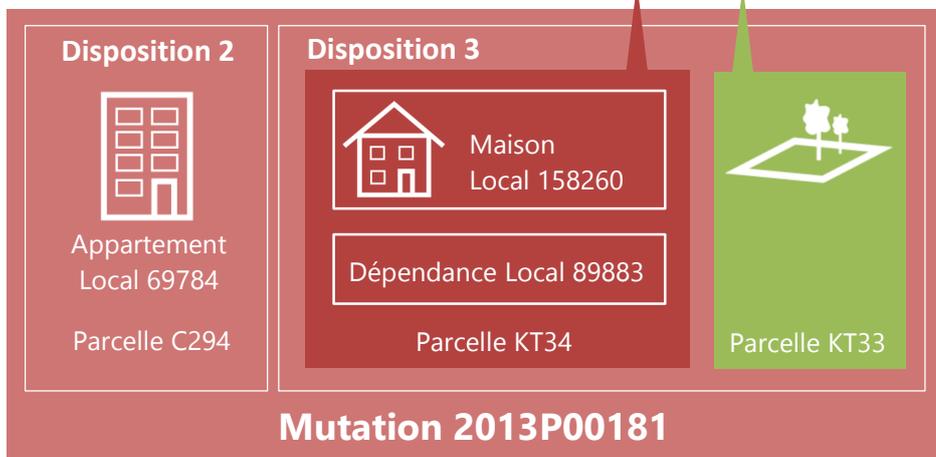
Pour la disposition 2, il n'y a pas de renseignement de surface : la parcelle n'est pas vendue avec l'appartement

Reference document	No disposition	Valeur foncière	Section	No parcelle	Type local	Identifiant local	Nature culture	Nature culture spéciale	Surface terrain
2013P00181	2	180000	C	294	Appartement	69784			
2013P00181	3	317000	KT	33			Jardin		1368
2013P00181	3	317000	KT	34	Dépendance	89883	Sol		1000
2013P00181	3	317000	KT	34	Dépendance	89883	Agrément	Sport	800
2013P00181	3	317000	KT	34	Dépendance	89883	Agrément	Chasse	3633
2013P00181	3	317000	KT	34	Maison	158260	Sol		1000
2013P00181	3	317000	KT	34	Maison	158260	Agrément	Sport	800
2013P00181	3	317000	KT	34	Maison	158260	Agrément	Chasse	3633



Parcelle composée de Sol pour 1000 m², de terrain d'agrément Sport pour 800 m², et de terrain d'agrément Chasse pour 3633 m²

Parcelle composée d'un jardin de 1368 m²



En synthèse **la mutation 2013P00181** présentée dans DVF est d'un montant de **497 000 €** (somme des valeurs foncières de la disposition 2 et de la disposition 3). Elle comprend d'une part **un appartement**, et d'autre part **une maison et sa dépendance**, sur un terrain de 6801 m² composé de deux parcelles.

A NOTER

Les natures de cultures sont affectées aux subdivisions fiscales de la parcelle, et ne sont donc pas liées aux locaux. Pour une même parcelle, la DGFIP ne propose aucune information permettant de savoir si le local est sur telle ou telle subdivision fiscale. Il n'y a donc pas de lien géographique dans MAJIC entre les locaux et les natures de culture. DVF va donc ventiler le local en autant de natures de culture et de natures de culture spéciales qu'il trouvera dans MAJIC sur la parcelle concernée.

7.3 Comment intégrer les données ?

7.3.1 Quels outils utiliser ?

Pour procéder au traitement des fichiers restitués par la DGFiP, plusieurs options sont possibles. Les solutions logicielles proposées ci-dessous sont issues des retours des utilisateurs. La recommandation principale est d'utiliser les scripts du Modèle Conceptuel de Données élaboré par le

Groupe national DVF (voir fiche 9 *Structurer les données*). Ces scripts peuvent être implémentés sur différents logiciels de gestion de bases de données selon les ressources et compétences de chaque requérant et de chaque structure.

	Solution logicielle	Exemples
La base	Traitement sur tableur	Microsoft Excel, Open/Libre Office Calc, etc.
Pour croiser DVF avec des bases de données lourdes	Utilisation d'un ETL (Extract-Transform-Load) permettant d'effectuer des synchronisations massives d'informations d'un format de données vers un autre.	FME, Talend, etc.
Pour aller plus loin	Traitement avec un système de gestion de base de données (SGBD) relationnel	PostgreSQL, scripts du GnDVF, Microsoft Access, open/Libre Office Base, etc.

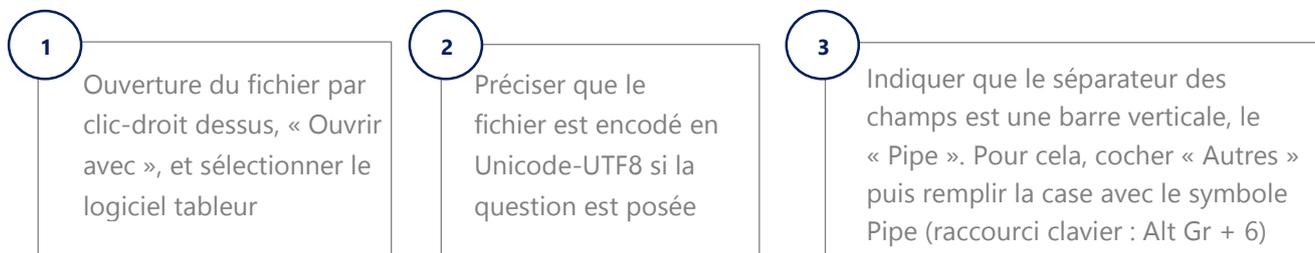
A NOTER

Le CEREMA a développé un outil permettant de structurer les données DVF sous forme de base de données « DVF+ » à l'aide d'un logiciel téléchargeable : [ImportDVF+](#).

7.3.2 Pour un usage de la donnée avec un tableur

Pour un usage ponctuel de la donnée ou en l'absence de moyens informatiques ou humains permettant de gérer des bases de données, une étude sur les prix ou la surface des terrains peut-être réalisée.

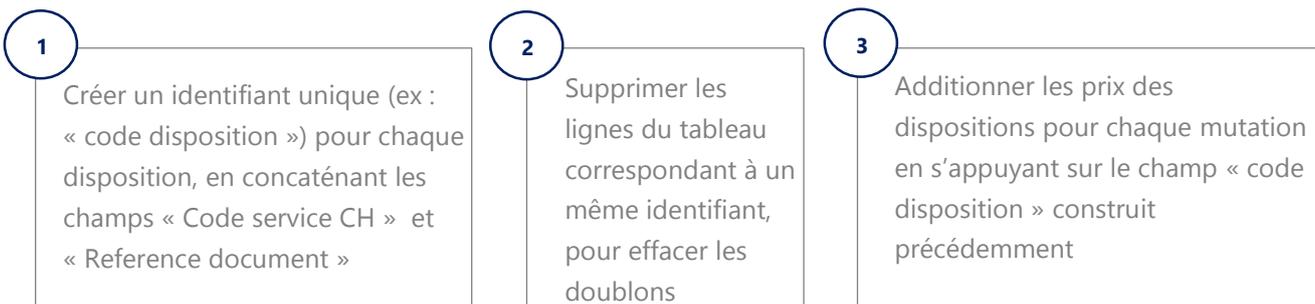
a. Ouvrir le fichier texte avec tout tableur (par exemple, Microsoft Excel ou Open Office Calc)



Dans les anciennes versions de Microsoft Excel, il est possible que la boîte de dialogue ne s'ouvre pas. Il faut alors utiliser le menu Données/Convertir.

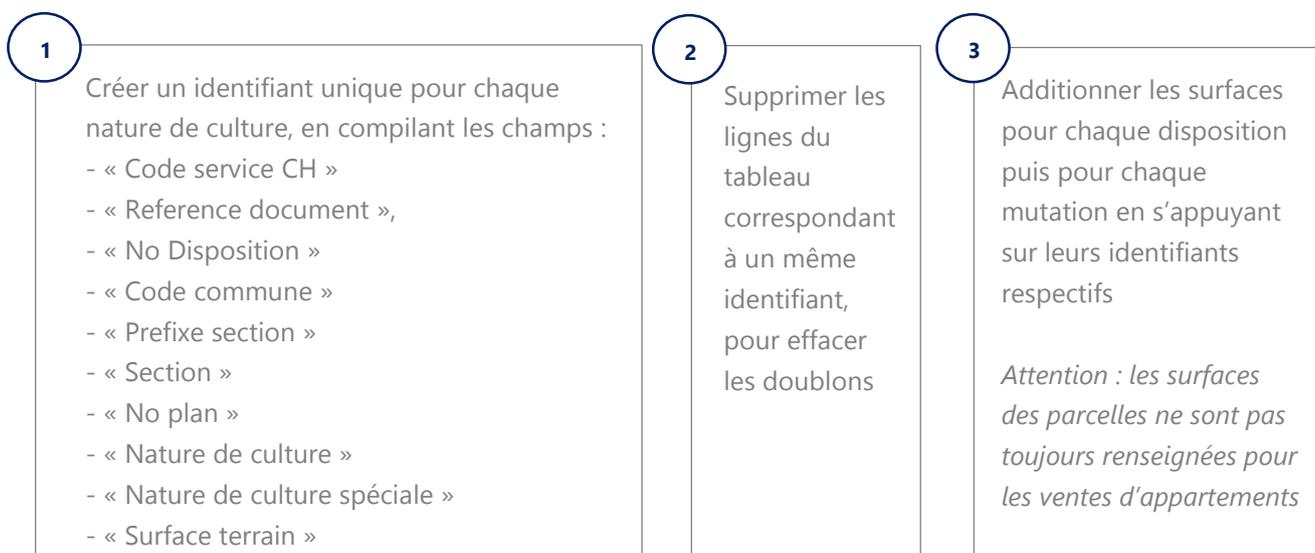
b. Etudier les prix (en perdant les informations sur les surfaces et sur les locaux)

Des analyses de prix peuvent être faites avec prudence via un tableur.



c. Etudier les surfaces de terrain

Des analyses de surface de terrains peuvent être faites avec prudence via un tableur.



7.4 Prise en compte des écarts entre les millésimes

Le décalage entre la date de la mutation et la date de sa publication par le service de publicité foncière engendre des écarts dans les données d'une même année récupérées à des moments différents. Le délai de publication de l'acte en service de publicité foncière est variable (de

plusieurs jours à plusieurs mois). Par ailleurs, certaines données des années antérieures étant enrichies, les utilisateurs peuvent récupérer tous les millésimes disponibles à chaque demande (antériorité de 5 ans).



Le fichier récupéré à l'étape 3 sera différent du fichier récupéré à l'étape 1 car :

- il contiendra en plus toutes les mutations publiées à l'étape 2,
- certaines mutations en VEFA se verront complétées du bien construit postérieurement à la transaction.

Certaines données sont enrichies d'une année sur l'autre

Année de remplissage de la base DVF	Nombre de mutations 2009
2009	51 133
2010	9 212
2011	30
2012	7
2013	1
TOTAL	60 383

La Région Aquitaine a étudié l'écart entre l'année de signature de l'acte (4 derniers caractères du champ date_mutation) et l'année de publication de l'acte au service de publicité foncière (4 premiers caractères du champ reference-document).

Lors du téléchargement des mutations de l'année 2013, les 5 dernières années ont été demandées (2009 à 2013).

Le tableau illustre les résultats issus du fichier des actes signés en 2009. Sur les 60 383 mutations réalisées en 2009, 51 133 ont été publiées par le service de publicité foncière en 2009, 9 212 en 2010, ...

Cette analyse illustre bien l'intérêt de récupérer tous les millésimes disponibles à chaque demande puisque le fichier des mutations restituées dans DVF pour une même année (ici 2009) est enrichi au fur et à mesure des enregistrements réalisés par le service de publicité foncière.

Une analyse valide le délai entre la date de la mutation et son apparition dans DVF

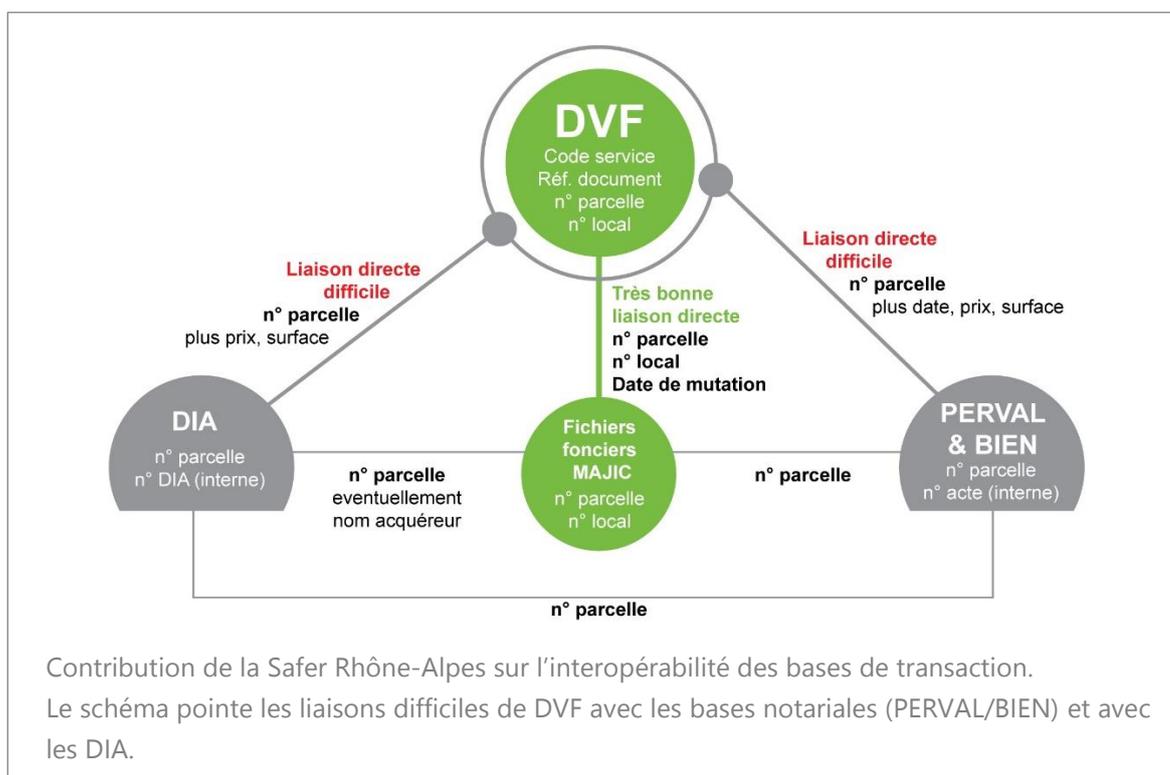
7.5 Interopérabilité des bases de données

Certaines bases de données sont plus faciles à croiser avec DVF que d'autres

Le traitement des données issues du service DVF amène très souvent un croisement avec d'autres bases de données. Se pose alors la question de l'interopérabilité entre les bases.

Si l'appariement avec la base MAJIC est opérationnelle (car les données DVF sont elles-mêmes en partie issues de la base MAJIC), le croisement avec d'autres bases de données n'est pas aussi aisé.

En effet, l'identifiant local, référence cadastrale, correspond à l'invariant de la matrice cadastrale, il est présent à la fois dans DVF et dans MAJIC, ce qui permet de croiser très facilement les données. Or, cet identifiant n'est pas présent dans les autres bases comme les DIA ou les bases notariales qui sont construites autour des bases d'adresses.



Croiser avec le cadastre ou les POS/PLU suppose de choisir les bons millésimes

Un autre aspect des appariements avec d'autres bases de données concerne la problématique de l'historicisation, notamment lors de croisements avec le cadastre ou avec les PLU/POS et autres documents d'urbanisme. Il faut alors veiller à croiser les données DVF avec le bon millésime de l'autre source d'informations.

Par exemple, les documents d'urbanisme peuvent utilement compléter DVF pour connaître la vocation ou la constructibilité d'un terrain ayant fait l'objet d'une vente. Il faut donc s'assurer de disposer de l'information correspondant à la date de transaction. Ainsi, une vente de 2010 peut ne pas correspondre à

un règlement d'urbanisme postérieur (ou même antérieur) à cette date.

En effet, les POS et PLU évoluent. Ils sont différents d'année en année. Par exemple les données DVF de 2012 peuvent correspondre à une date de POS/PLU antérieure à 2010.

La question de la date du plan cadastral est importante également car il y a un risque de perdre la géométrie des parcelles. Le plan cadastral n'est plus le même d'une année sur l'autre, donc il faut travailler avec un historique pour pérenniser la localisation de la parcelle (voir fiche 10 *La spatialisation des données DVF*).

7.6 Coordination des méthodes et des filtres

A NOTER

Le Groupe national assure une harmonisation au niveau national des méthodes et outils de traitement des données DVF par la diffusion de contributions d'utilisateurs et leur valorisation en tant que références.

Le Groupe national souhaite favoriser la mise en place de référents et de groupes de travail régionaux qui auront notamment pour mission de coordonner les traitements au niveau local. Dans certaines régions, cette dynamique en synergie existe déjà et permet la construction d'une homogénéité régionale.

Selon les usages des données DVF, les méthodes et les filtres utilisés ne sont pas les mêmes. Il y a donc un enjeu important d'harmonisation de ces traitements.

L'enjeu est de garantir la cohérence de la diffusion d'indicateurs qui sont issus d'une même source de données, à savoir DVF.

En ayant des méthodes de traitement, de redressement et des filtres communs, les structures peuvent mener des comparaisons entre différents territoires.

La coordination passe par le partage en réseau des méthodes de traitement et des filtres employés par les différentes structures.

Il est essentiel par exemple que chaque étude diffusée mentionne le processus de structuration et d'analyse utilisé.

La cohérence se crée tant par les synergies locales que par les travaux nationaux

La Communauté d'agglomération Lannion-Trégor a réalisé une étude sur le foncier bâti du centre-ville de Lannion, introduite par une partie sur la méthodologie employée.

Méthodologie employée

Le recensement des transactions a été permis par la fourniture par la DGfip (direction des Impôts) du fichier des valeurs foncières à l'échelle de l'Agglomération.

Ce fichier recense l'ensemble des transactions immobilières à l'échelle de territoires sur des périodes données avec la possibilité de revenir 5 ans en arrière.

Le traitement exposé au sein de ce document propose une réflexion à l'échelle de 4 sections cadastrales proches de la notion de centre-ville élargi.

Il a d'abord été utilisé cette échelle large, regroupant l'ensemble des sections cadastrales définissant le centre-ville et sa périphérie proche pour finir par des traitements pour chaque échelle cadastrale et sur certains secteurs ciblés avec un nombre de transactions suffisant pour pouvoir établir des statistiques pertinentes.

Pour fiabiliser ce fichier, il a également été comparé au fichier des DIA (Déclarations d'Intentions d'Aliéner) fournie par la ville de Lannion.

Près de **337 biens** ont été vendus entre 2008 et 2011 à l'échelle du centre-ville élargi. N'ont pas été comptabilisés les acquisitions de logements neufs liés notamment aux lois de défiscalisation, car marché fonctionnant différemment.

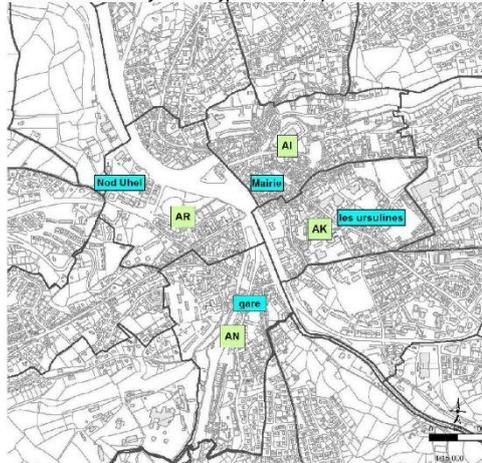
A noter que les chiffres 2012 seront intégrés dès réception du fichier des impôts (septembre) ainsi qu'une étude sur les logements neufs commercialisés.

3 types de biens différents sont recensés :

- ⇒ L'habitat individuel
- ⇒ L'habitat collectif
- ⇒ Les locaux commerciaux et industriels.

Cette étude doit permettre de soulever un certain nombre d'interrogations et d'éléments de contexte pour préparer des entretiens avec les acteurs de l'immobilier et peut-être faire des études plus précises sur des éléments posant question.

Egalement, l'étude porte sur le marché des biens vendus et non des biens en vente. Devra donc être menée en parallèle, une étude sur le marché de la vente (recensement de l'ensemble des offres immobilières et analyse avec type de bâtis,...).



L'étude commence par une page sur la méthode adoptée