

Notice relative à l'utilisation des données DV3F

Juin 2019



DREAL Occitanie



SOMMAIRE

1 PRÉSENTATION DE LA BASE DE DONNÉES DV3F	3
1.1 De DVF à DV3F : intérêt et limites de la base de données DV3F	
1.2 Entrons dans la base de données DV3F	
1.2.1 Tables principales	5
1.2.2 La typologie des biens	5
1.2.3 Les principales variables (ou champs)	9
1.2.4 La superficie des biens	11
1.2.5 La géolocalisation dans DV3F	12
1.2.6 La variable devenir pour évaluer la destination du bien	13
1.2.7 La variable filtre pour identifier des mutations particulières	15
1.2.8 La variable segmtab et l'identification des terrains à bâtir	17
1.2.9 La typologie des vendeurs et acquéreurs	18
2 PRISE EN MAIN DE LA BASE SOUS POSTGRESQL	21
2.1 Installer une base de données DV3F	21
2.2 Recommandations méthodologiques pour l'analyse des marchés immobiliers	21
2.2.1 Constituer de nouvelles tables	21
2.2.2 Constituer un segment de marché	21
2.2.3 Analyser le marché du logement	22
2.2.4 Estimer le devenir du bien en fonction du type d'acquéreur	25
2.2.5 Analyser le marché des terrains à bâtir	26

Le présent document constitue une notice d'utilisation de la base de données DV3F développée par le Cerema, réalisée dans le cadre d'une mission confiée par la DREAL Occitanie au Cerema.

La documentation complète relative à DV3F est consultable sur le site https://datafoncier.cerema.fr/

1 Présentation de la base de données DV3F

1.1 De DVF à DV3F : intérêt et limites de la base de données DV3F

La donnée DVF de la DGFIP est issue de 2 sources de données fiscales : FIDJI, Fichier Informatisé des Données Juridiques Immobilières, qui provient des informations remontées par le service de la publicité foncière concernant les transactions immobilières à titre onéreux et MAJIC, Mise à Jour des Informations Cadastrales. Elle est actualisée et accessible à un rythme semestriel.

La base de données DV3F développée par le Cerema structure la donnée source et l'enrichit par croisement avec les Fichiers Fonciers pour permettre une exploitation de la donnée dans une optique d'observation foncière. Elle comporte un grand nombre de champs qui permettent de caractériser les transactions immobilières.

La géolocalisation des données

Un des principaux intérêts de DV3F réside dans la géolocalisation des données sources. L'apport de la géométrie parcellaire accroît fortement les potentiels d'usage : représentations cartographiques, croisements avec d'autres périmètres (PLU par exemple), requêtes géométriques variées, etc.

Les géométries proviennent d'un historique de parcelles constitué majoritairement à partir de différents millésimes de la BD Parcellaire (IGN). Cela permet, dans DV3F, de géolocaliser des parcelles qui auraient disparu à la suite d'une vente et ainsi d'obtenir des taux de géolocalisation très satisfaisants.

La caractérisation des biens bâtis

Les caractéristiques des biens renseignés dans la donnée source DVF sont peu détaillées. Pour les locaux notamment, DVF fournit uniquement le type du local, la surface du bâti et le nombre de pièces principales.

Dans DV3F (avec les variables rapatriées des Fichiers fonciers), les informations sur les biens échangés sont plus précises et permettent ainsi d'aller plus loin dans la caractérisation de certains biens. Pour les locaux notamment, DV3F permet d'obtenir l'année de construction, de préciser le nombre et les types de dépendance ou encore d'identifier les locaux considérés comme du logement.

NB : Parfois, certaines caractéristiques du bien ne sont pas disponibles dans les données sources telles que la date de construction et apparaissent ainsi comme « indéterminées» dans la base DV3F.

La caractérisation des terrains nus

Les informations permettant de caractériser la nature des terrains nus vendus et leur usage envisagé par l'acquéreur n'entrent pas en ligne de compte dans le calcul des taxes de mutation. Par conséquent, ces informations sont peu voire mal renseignées dans les actes notariés. Il existe néanmoins une classification des terrains dans DV3F dans laquelle figure une catégorie « terrains à bâtir ».

Un terrain nu est considéré comme un terrain à bâtir dans DV3F lorsque différents champs renseignés dans la donnée source laissent penser qu'il s'agit bien d'un terrain à bâtir. Par conséquent, lorsqu'un terrain est classé dans cette catégorie, on estime que l'information est globalement fiable. En revanche, cette catégorie est loin d'être exhautive. Cela a l'inconvénient de sous-estimer le nombre de terrains à bâtir et de biaiser les prix des autres marchés fonciers.

Une information sur les acheteurs et les vendeurs

DV3F vise à pallier l'absence d'informations sur les acquéreurs et vendeurs dans DVF. En travaillant à partir de plusieurs millésimes des Fichiers Fonciers, il est possible de retrouver le propriétaire d'un bien avant et après une vente et ainsi d'en déduire le vendeur et l'acheteur.

Les informations sur les propriétaires sont donc rapatriées dans DV3F à partir des millésimes disponibles des Fichiers Fonciers et accompagnées d'une estimation de la fiabilité liée au rapatriement. Ces informations sont anonymisées pour les personnes physiques.

NB : cette information n'est accessible qu'avec un décalage temporel lié à la disponibilité et à l'intégration du dernier millésime des Fichiers Fonciers. En outre, il arrive que l'acquéreur reste indéterminé.

La temporalité

L'enregistrement des actes de vente par les services fiscaux peut prendre plusieurs mois ce qui implique une apparition tardive dans la donnée DVF. Ainsi, lorsque l'on effectue des analyses quantitatives (nombre de mutations ou volume de transactions, par exemple), il faut éviter de regarder l'évolution sur les derniers trimestres.

1.2 Entrons dans la base de données DV3F

1.2.1 Tables principales

Dans la base de données DV3F, les informations essentielles sont concentrées sur 2 tables principales qui comptent 225 variables :

- la table « mutation »,130 variables
- la table « **local** », 95 variables.

En pratique pour les requêtes, nous utiliserons principalement la **table mutation** et de façon plus marginale la **table local**.

1.2.2 La typologie des biens

Une typologie de biens classifie les mutations selon le ou les types de biens vendus. Cette typologie définie par le groupe national DVF (GnDVF) est homogène au niveau national.

Elle possède au plus 5 niveaux (voir schémas ci-dessous) :

- le niveau 1 différencie les biens bâtis des biens non bâtis,
- le niveau 2 présente les grandes catégories de biens immobiliers et fonciers,
- le niveau 3 présente les sous-catégories de biens immobiliers et fonciers (avec notamment les types d'activités ou l'occupation du sol),
- le niveau 4 intègre essentiellement les notions d'ancienneté des logements ou le détail de l'occupation des terrains agricoles,
- le niveau 5 intègre les notions de nombre de pièces pour les appartements.

A chaque catégorie de bien correspond un libellé et un code qui constituent deux variables essentielles pour interroger la base de données :

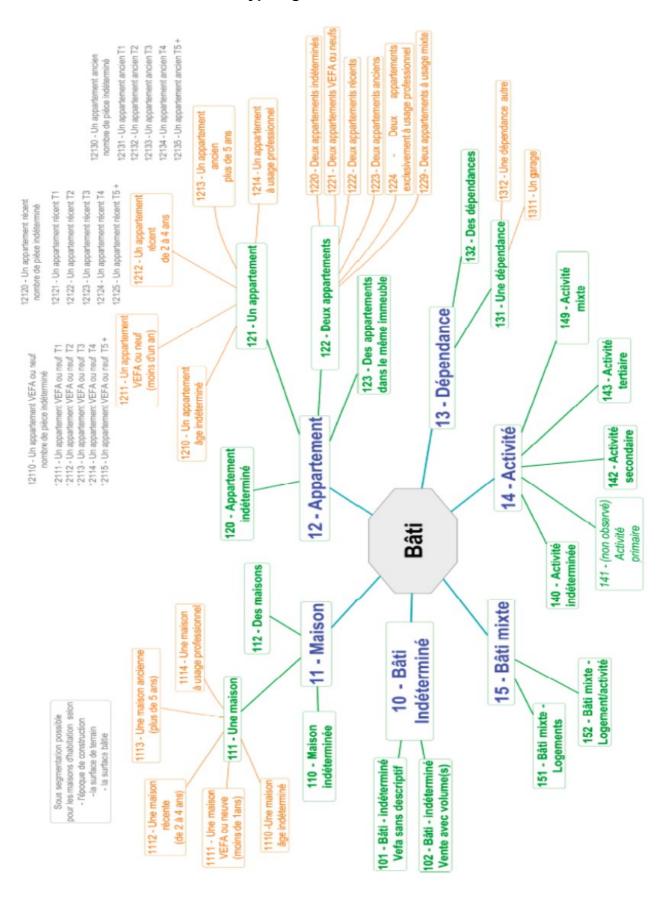
- **Libtypbien** (libellé du type de bien) : par exemple « un local d'activité secondaire »,
- Codtypbien (code du type de bien) : chaque chiffre du code correspond à un niveau de la typologie.

Exemple « 142 » correspond à un local d'activité secondaire avec

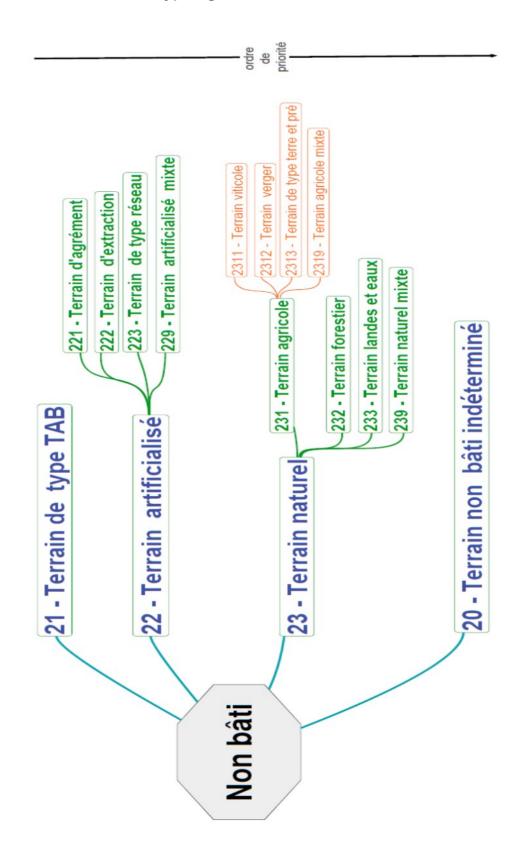
```
niveau 1 = « 1 » = bien bâti
niveau 2 = « 14 » = activité
niveau 3 = « 142 » = activité secondaire
```

Les deux schémas ci-dessous, explicient pour le « bâti » et le « non bâti », l'ensemble des typologies de biens identifiées dans la base de données, ainsi que le code correspondant.

Typologie des biens bâtis



Typologie des biens non bâtis



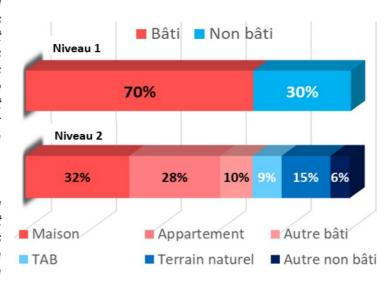
A titre d'exemple, le graphique ci-dessous illustre la répartition des mutations d'Occitanie entre 2010 et 2016, par typologies principales de biens identifiées dans DV3F, sur les niveaux 1 et 2.

Au niveau 1, on constate que 70 % des mutations comprennent au moins un élément bâti (regroupant toutes les typologies dont le codtypbien commence par le chiffre 1), contre 30 % de mutations liées uniquement au foncier (dont le codtypbien commence par le chiffre 2).

Plus finement, au niveau 2, 32 % des mutations concernent une ou plusieurs maisons (codtypbien commençant par 11), 28 % un ou plusieurs appartements (codtypbien commençant par 12) et 10 % d'autres éléments bâtis (parmi lesquels notamment les dépendances, les locaux d'activité ou des mutations concernant plusieurs catégories de bâti (ex: un appartement et une maison dans la même vente)).

Les 30 % de mutations relatives au non bâti concernent essentiellement des terrains naturels (15 %) et des terrains à bâtir (9 %), classés comme tels à partir des données fiscales.

Répartition de l'ensemble des mutations d'Occitanie entre 2010 et 2016



A titre indicatif, en interrogeant les niveaux 4 et 5, le type de mutations majoritaire concernant un bien immobilier est la vente d'une maison de plus de 5 ans (29 % des mutations totales, codtypbien = 1113).

1.2.3 Les principales variables (ou champs)

La syntaxe des variables

Quelques règles syntaxiques ont été mises en place sur les variables pour donner des repères à l'utilisateur :

Si la variable commence par	Alors il s'agit d'un(e)
id	identifiant
cod	code (postal, insee, etc)
S	surface
nb	dénombrement
I_	liste (type array)
lib	Libellé
ff	donnée directement issue des Fichiers fonciers
I_ff	liste (type array) de données issues des Fichiers fonciers
Si la variable se termine par	Alors il s'agit d'une
a	donnée liée à l'acheteur
V	donnée liée au vendeur

A noter que toutes les variables comportent au maximum 10 caractères de manière à faciliter l'exportation au format shp.

Les principales variables de la table mutation

Libellé variable	Description
anneemut	Année de signature de l'acte notarié relatif à la transaction
codtypbien	Code de la typologie des biens du GnDVF
codtyppoa	Code de la typologie des propriétaires acheteurs
datemut	Date de signature de l'acte de vente
devenir	Indicateur de changement de typologie ou d'usage des biens vendus après la vente
filtre	Code alphanumérique permettant d'exclure des mutations particulières pour l'analyse
idnatmut	Identifiant unique permettant d'identifier les différentes natures de mutation définies dans la donnée source DVF (1 : Vente, 2 : VEFA, 3 : Expropriation, 4 : Vente terrain à bâtir, 5 : Adjudication, 6 : Echange de biens). Les valeurs {1,2,4 } correspondant à {vente, VEFA, vente de TAB} permettent d'identifier toutes les transactions de type « Ventes ».
I_dcnt	Liste ordonnée des surfaces des 13 catégories de subdivisions fiscales (natures de culture).
l_idpar	Liste des identifiants de parcelles concernées par la mutation (idpar). Permet la géolocalisation du bien.
I_noma	Nom des acheteurs (anonymisés pour les personnes physiques)
I_nomv	Noms des vendeurs (anonymisés pour les personnes physiques)
libnatmut	Libellé de la nature de mutation (1 : Vente, 2 : VEFA, 3 : Expropriation, 4 : Vente terrain à bâtir, 5 : Adjudication, 6 : Echange de biens)
libtypbien	Libellé de la typologie des biens du GnDVF
nblocact	Nombre de locaux d'activités ayant muté
nblocapt	Nombre d'appartements ayant muté
nblocdep	Nombre de dépendances ayant muté
nblocmai	Nombre de maisons ayant muté
nblocmut	Nombre de locaux ayant muté
nbparmut	Nombre de parcelles ayant muté (attention : contient les parcelles nues mutés et les parcelles support d'un bien muté (hors immeuble d'appartement))
sbati	Surface de l'ensemble du bâti ayant muté. Permet le calcul du prix au m² du bâti
segmtab	Indicateur du segment des terrains à bâtir
sterr	Surface du terrain (= 0 dans le cas d'un appartement muté). Permet le calcul du prix au $\rm m^2$ de foncier
valeurfonc	Prix ou évaluation déclarée dans le cadre d'une mutation onéreuse ; prix de l'ensemble des biens faisant l'objet de la transaction
vefa	Si la mutation est une Vente en l'État Futur d'Achèvement (VEFA), la variable est égale à « vrai »

NB : Le champ vefa repère la majorité des VEFA, mais il reste des VEFA non repérées.

La variable anciennete de la table local

La table <u>local</u> comporte un champ **ancienneté** qui indique l'ancienneté du local au moment de la mutation. Ce champ est basé sur la différence entre l'année de mutation DVF et l'année de construction du local (champ **ffancst**) issue des fichiers fonciers.

Valeur du champ	Description
ancien	le local a au moins 5 ans au moment de la vente
neuf	le local a au plus 1 an au moment de la vente
non defini	le local n'a pas de date de construction défini. Il est impossible de connaître son ancienneté au moment de la vente.
recent	le local a plus d'1 an et moins de 5 ans au moment de la vente

La date de construction du bien n'est pas toujours mentionnée dans les fichiers fonciers. Par conséquent, lorsque l'on segmente les biens bâtis vendus en catégories neufs/récents/anciens, le volume total des biens vendus est égal à la somme des biens neufs, récents, anciens et « non défini ».

Il est à noter que les biens faisant l'objet de Vente en l'État Futur d'Achèvement (VEFA) sont inclus dans les biens « neufs ».

En pratique, lorsque l'on travaille sur un indicateur de prix, la valeur du champ **libtypbien** de la table <u>mutation</u> indique l'ancienneté du bien lorsqu'elle est connue.

A titre d'exemple, les ventes (idnatmut ={1,2,4 }) en Occitanie entre 2011 et 2015 sont réparties comme suit :

	Neuf (dont Vefa) (moins d'1 an)	Récent (de 1 à 5 ans)	Ancien (plus de 5 ans)	Non renseigné
Une maison seule	3% (1% vefa)	4%	91%	2%
Un appartement seul	19% (16% vefa)	3%	72%	6%

La très grande majorité des ventes concernent des logements de 5 ans et plus. On constate cependant que le poids des Vefa sur les appartements n'est pas négligeable et que la VEFA est le mode de vente privilégié pour un appartement neuf.

1.2.4 La superficie des biens

Surface d'une parcelle

La surface <u>ffsparc</u> est la surface officielle au cadastre et devrait être la référence à prendre pour la surface de la parcelle.

Il existe deux autres variables décrivant la surface du terrain ayant muté dans DV3F et qui normalement sont équivalentes à la surface de la parcelle :

sterr qui correspond à la somme de toutes les surfaces des suf (subdivisions

fiscales) issues de DVF ayant muté pour une même mutation,

• <u>ffsterr</u> qui correspond à la somme de toutes les surfaces des suf issues des Fichiers fonciers ayant muté pour une même mutation.

Il peut exister des différences de surface entre <u>sterr</u>, <u>ffsterr</u>, <u>ffsparc</u> mais cela reste marginal et <u>sterr</u> est systématiquement rempli, contrairement à <u>ffsparc</u>. De ce fait, il est conseillé d'utiliser en priorité <u>sterr</u>.

Par contre, il est peut-être intéressant de s'assurer lorsqu'on travaille sur des échantillons ou de très petits nombres de mutations que ces variables sont cohérentes entre elles, surtout si l'on travaille sur l'occupation du sol (suf).

Surface réelle bâtie

Dans DV3F, la surface bâtie représente la surface mesurée au sol du plancher entre murs ou séparations. Le résultat est arrondi au mètre carré inférieur.

Elle peut se décomposer en une surface dite "habitée", une surface professionnelle (ou liée à une activité) et une surface de dépendances.

Dans DV3F, la notion de surface réelle bâtie (<u>sbati</u>),est la somme de la surface habitée et de la surface professionnelle et la notion de surface réelle bâtie provenant des Fichiers fonciers (<u>ffsbati</u>),est la somme de la surface habitée, de la surface professionnelle et des surfaces des dépendances. Ainsi, lorsque le rapatriement est bien effectué, on peut écrire les égalités suivantes :

- sbati = ffshab + ffspro,
- ffsbati = ffshab + ffspro + ffsdep.

Le détail des différentes composantes de surface selon l'occupation est disponible.

Pour l'ensemble des calculs, il est préférable de se servir de sbati (qui ne comptabilise pas la surface des dépendances).

1.2.5 La géolocalisation dans DV3F

DV3F permet la localisation des biens selon plusieurs moyens :

- via les informations administratives : code du département, code insee de la commune, code de section cadastrale, numéro des parcelles concernées par la vente ;
- via la géolocalisation à la parcelle : un localisant ponctuel et un contour vectorisé ;
- via l'adresse du bien : numéro de la voie, libellé de la voie, code postal et libellé de commune.

Des géométries sont présentes dans trois tables de DV3F :

- dans la table disposition_parcelle : le contour et le localisant de chaque parcelle est restitué respectivement par geompar et geomloc;
- dans la table *local*: le localisant de la parcelle sur laquelle se situe le local au moment de la vente est restitué par <u>geomloc</u>;

 dans la table mutation: la variable geomparmut correspond au contour de l'ensemble des parcelles ayant muté, la variable geomlocmut correspond aux localisants des parcelles sur lesquelles se situent les locaux mutés et la variable geompar correspond au contour de l'ensemble des parcelles concernées.

Dans l'illustration ci-dessous, on retrouve, délimitées par des pointillés rouges, les géométries de toutes les parcelles vendues ou sur lesquelles un bien a été vendu (geompar). Lorsque la parcelle a été vendue, elle est ici colorée en vert (geomparmut). Lorsque qu'au moins un local a été vendu sur la parcelle, le localisant de la parcelle est



représenté par un rond vert (geomlocmut).

1.2.6 La variable devenir pour évaluer la destination du bien

Il ne sera jamais possible de connaître l'intention d'un acquéreur au moment de la transaction. Dans DV3F, avec un recul temporel suffisant, il reste toutefois possible de repérer certaines évolutions du bien.

Ces évolutions sont retranscrites dans la variable « **Devenir** » qui permet notamment d'identifier :

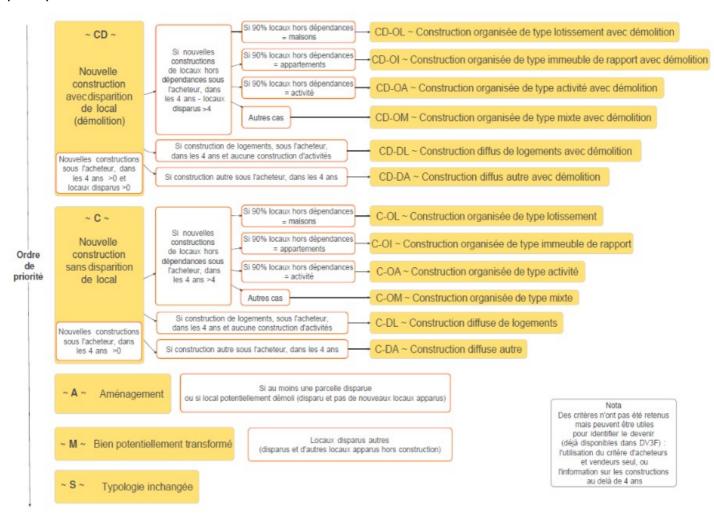
- la démolition et re-construction après la transaction,
- la construction après la transaction,
- l'aménagement après la transaction,
- la potentielle modification du bien,
- l'absence d'évolution constatée.

A titre d'exemple, sur deux périodes différentes, la répartition de l'évolution des mutations en Occitanie est la suivante :

	Absence d'évolution constatée	Aménagement après la vente	Construction (avec ou sans démolition) après la vente	Autre évolution
2010-2012	84%	5%	10%	1%
2013-2015	89%	3%	7%	1%

Le champ « devenir » peut évoluer jusque 4 ans après la signature de l'acte de la mutation. Ainsi, on peut imaginer que l'écart de 5 points sur la 1ère colonne du tableau de l'exemple est dû au fait que, pour certaines mutations de 2013 à 2015, des compléments peuvent encore être apportés, précisant le devenir du bien acquis.

Le schéma ci-après détaille les modalités proposées par cette variable avec leurs principes de construction :



Cette construction se traduit dans la table mutation par les valeurs suivantes :

Valeur du champ	Description
А	Aménagement
C-DA	Construction diffuse autre
CD-DA	Construction diffuse autre avec démolition
CD-DL	Construction diffuse de logements avec démolition
C-DL	Construction diffuse de logements
CD-OA	Construction organisée de type activité avec démolition
CD-OI	Construction organisée de type immeuble de rapport avec démolition
CD-OL	Construction organisée de type lotissement avec démolition
CD-OM	Construction organisée de type mixte avec démolition
C-OA	Construction organisée de type activité
C-OI	Construction organisée de type immeuble de rapport
C-OL	Construction organisée de type lotissement
C-OM	Construction organisée de type mixte
М	Bien potentiellement transformé
S	Typologie inchangée

Cette variable est particulièrement utile pour qualifier les terrains nus vendus et identifier en particulier les terrains à bâtir (voir plus loin).

1.2.7 La variable <u>filtre</u> pour identifier des mutations particulières

Dans la construction d'un indicateur de prix, il est important d'exclure certaines mutations ayant des conditions de ventes particulières. Si l'on vend un bien à un proche, le prix risque de ne pas être le même que si c'est une vente aux enchères, une expropriation, ou une vente simple.

De même, si la typologie classifie les biens, elle ne retranscrit pas certaines spécificités qui peuvent changer considérablement le prix d'une vente. Par exemple, la vente d'un terrain de 5 ha, avec une maison, sera considérée dans la typologie comme une maison. Ceci est peut-être réducteur.

DV3F propose, pour cela, un indicateur unique, <u>filtre</u>, qui alerte sur des biens rares ou des conditions de ventes spécifiques :

Chaque code de catégorie est défini par un unique caractère (une lettre ou un chiffre) et peut être associé, par concaténation, à d'autres si la vente répond à plusieurs catégories. L'ordre des éléments de la chaîne de caractère est alphabétique puis par ordre numérique croissant.

La modalité '0' correspond à une transaction de type vente sans particularités. C'est le

cas le plus courant.

Code	Description
0	Mutation ne comportant aucun des critères suivants
1	Mutation d'un terrain bâti de plus de 1 ha/local vendu
5	Mutation d'un terrain bâti de plus de 5 ha/local vendu
А	Mutation de type Adjudication
В	Mutation d'un appartement avec terrain
D	Mutation dont au moins un des biens a été vendu une deuxième fois le même jour
Е	Mutation de type Échange
Н	Mutation dont le prix est null, 0 ou 1 euro
L	Mutation comportant un bien exceptionnel parmi : logement de plus de 8 pièces strictement, maison de plus de 300m2 de surface bâtie, appartement de plus de 200m2 de surface bâtie
M	Mutation de plusieurs biens distancés de 1km ou appartenant à plusieurs communes
S	Mutation à caractère social ¹
Т	Transfert de propriétés entre deux opérateurs de logement social
Χ	Mutation de type Expropriation

Il est conseillé d'utiliser cette variable pour « évincer » l'ensemble des mutations particulières lors de l'élaboration d'indicateurs de prix :

SELECT * FROM dvf.mutation WHERE filtre = '0'

A titre d'exemple, 90 % des mutations d'Occitanie entre 2011 et 2015 correspondaient à une transaction de type vente sans particularités.

Notice d'utilisation des données DV3F

¹ Vente impliquant un bailleur, vente de logements sociaux ou vente d'un terrain sur lequel a été bâti du logement social.

1.2.8 La variable segmtab et l'identification des terrains à bâtir

Un terrain nu est considéré comme un « terrain à bâtir » dans la base DV3F (« codtypbien » = « 21 » et « libtypbien » = « terrain de type TAB », dans la typologie des biens) lorsque les informations issues de DVF permettent de le supposer :

 le champ 'nature de mutation' indique que c'est un terrain à bâtir (libnatmut = « vente terrain à bâtir », idnatmut=4)

ou

• il est soumis à une fiscalité relative aux terrains à bâtir (articles CGI : '1594D*2','257-7-1*3','278 sexies I.1','691bis', '1594OGA')

ou

la nature de culture est de type 'Terrains à bâtir' (« l_dcnt[10] > 0 »)

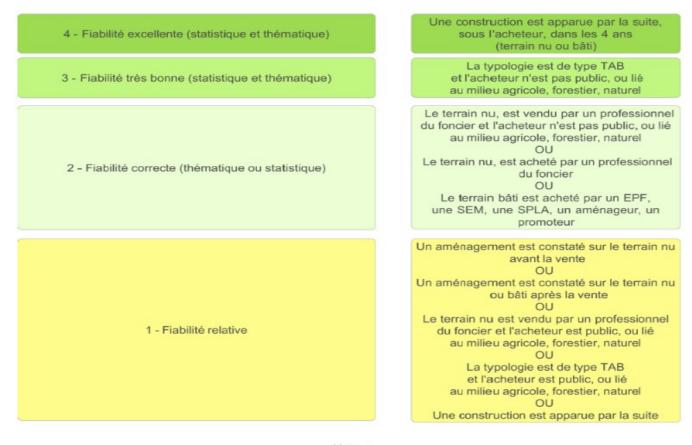
Il s'agit donc d'une catégorie « terrain à bâtir » <u>au sens fiscal</u>. Malgré tout, cette information n'est pas complètement fiable. En effet, on constate que certains biens classés dans la catégorie « terrains à bâtir » n'en sont pas réellement mais surtout que de nombreux terrains à bâtir n'ont pas été classés dans cette catégorie, car la nature de mutation, la fiscalité ou les natures de cultures ne le laissent pas supposer.

Aussi, pour identifier au mieux les terrains à bâtir, DV3F propose un indicateur de segmentation prêt-à-l'emploi selon une méthodologie élaborée par le Cerema. Une définition large du terrain-à-bâtir (TAB) est retenue : un TAB est un terrain acquis dans la perspective de pouvoir y construire immédiatement, avec/ou sans transformation, par l'acquéreur ou un tiers. Cela concerne entre autres les particuliers qui veulent construire une maison, et les acteurs de l'aménagement et de la construction: lotisseur, constructeur de maison individuelle (lorsqu'il est acquéreur), aménageur, promoteur, etc.

Un TAB est souvent un terrain non bâti, mais il peut être bâti au préalable. Par exemple, une maison acquise pour être démolie et permettre la construction d'un immeuble. Dès lors, un TAB peut être plus ou moins avancé dans la chaîne de transformation du foncier et sa valeur répondre à des raisonnements économiques différents (capacité d'emprunt foncier + construction de maison, compte à rebours promoteur, compte à rebours aménageur, etc).

L'indicateur <u>segmtab</u> détaillé ci-après, dont la valeur va de 1 à 4, permet d'évaluer avec une fiabilité plus ou moins élevée si le terrain vendu peut-être considéré comme un TAB. Cet indicateur est implémenté avec les informations disponibles et en particulier avec les Fichiers Fonciers. Parfois, aucune information ne permet de qualifier plus finement le terrain vendu. C'est notamment le cas des terrains vendus récemment non couverts par le dernier millésime de Fichiers Fonciers intégré à la base DV3F. De plus, l'indice de fiabilité 4 (fiabilité excellente) est basé sur le fait qu'une construction est apparue dans les 4 ans suivant la transaction. Pour les transactions les plus récentes, l'information n'est donc pas encore connue.

Valeurs de l'indicateur « segmtab »



Nota:

- Les terrains bâtis sans aménagement futurs ne rentrent pas dans l'observation du segment terrain-à-bâtir.
 - Un des critères de "fiabilité relative" implique à ce jour que tous les aménagements sont compris : routiers, public, etc.

1.2.9 La typologie des vendeurs et acquéreurs

Le nom des acquéreurs et des vendeurs n'est pas connu dans DVF mais est reconstitué dans DV3F par croisement de la donnée DVF avec les Fichiers Fonciers (champs « codtypproa » et « codtypprov »). Par suite, une typologie des propriétaires a été définie dans DV3F par le Cerema pour faciliter l'identification de certaines catégories d'acheteurs ou de vendeurs.

Dans la base DV3F, chaque code de catégorie est défini par 2 caractères : une lettre et un chiffre. La lettre correspond au premier niveau de la typologie (acteurs publics, professionnels du foncier et de l'immobilier, structures en milieu agro-naturel, propriétaires de réseau, personnes physiques). Le chiffre correspond à une souscatégorie de ce premier niveau.

Lorsque plusieurs catégories d'acheteurs (ou de vendeurs) ont pu être identifiées dans la même transaction, le code de catégorie comporte les différents codes concernés par concaténation.

Valeur des champs codtyppro(a et v)	Description
A1	PROPRIETAIRE ET EXPLOITANT DU FONCIER NATUREL AGRICOLE OU FORESTIER- STRUCTURE AGRICOLE
A2	PROPRIETAIRE ET EXPLOITANT DU FONCIER NATUREL AGRICOLE OU FORESTIER- STRUCTURE FORESTIERE
A3	PROPRIETAIRE ET EXPLOITANT DU FONCIER NATUREL AGRICOLE OU FORESTIER- STRUCTURE DU FONCIER ENVIRONNEMENTALE
A4	PROPRIETAIRE ET EXPLOITANT DU FONCIER NATUREL AGRICOLE OU FORESTIER-SAFER
F1	PROFESSIONNEL FONCIER IMMOBILIER- ORGANISME HLM
F2	PROFESSIONNEL FONCIER IMMOBILIER- ETABLISSEMENT PUBLIC FONCIER
F4	PROFESSIONNEL FONCIER IMMOBILIER- SEM / SPLA
F5	PROFESSIONNEL FONCIER IMMOBILIER- AMENAGEUR FONCIER
F6	PROFESSIONNEL FONCIER IMMOBILIER- PROMOTEUR IMMOBILIER PRIVE
F7	PROFESSIONNEL FONCIER IMMOBILIER- INVESTISSEUR PROFESSIONEL
G1	ORGANISATION DE GESTION FONCIERE ET IMMOBILIERE- SOCIETE CIVILE IMMOBILIERE
G2	ORGANISATION DE GESTION FONCIERE ET IMMOBILIERE- PROPRIETE DIVISEE EN LOT
MO	PERSONNE MORALE- AUTRE PERSONNE MORALE
PO	PUBLIC- AUTRE PERSONNE MORALE PUBLIQUE
P1	PUBLIC- ETAT
P2	PUBLIC- REGION
Р3	PUBLIC- DEPARTEMENT
P4	PUBLIC- INTERCOMMUNALITE
P5	PUBLIC- COMMUNE
P6	PUBLIC- COLLECTIVITE TERRITORIALE AUTRE
RO	PROPRIETAIRE ET EXPLOITANT DE RESEAU- PROPRIETAIRE DE RESEAU AUTRE
R1	PROPRIETAIRE ET EXPLOITANT DE RESEAU- CONCESSIONNAIRE AUTOROUTIER
R2	PROPRIETAIRE ET EXPLOITANT DE RESEAU- RESEAU FERRE
R3	PROPRIETAIRE ET EXPLOITANT DE RESEAU- STRUCTURE AERIENNE
R4	PROPRIETAIRE ET EXPLOITANT DE RESEAU- STRUCTURE FLUVIALE OU MARITIME
R5	PROPRIETAIRE ET EXPLOITANT DE RESEAU- RESEAU ELECTRIQUE OU GAZ
R6	PROPRIETAIRE ET EXPLOITANT DE RESEAU- RESEAU EAU OU ASSAINISSEMENT
R7	PROPRIETAIRE ET EXPLOITANT DE RESEAU- RESEAU DE TELECOMMUNICATION
XO	PERSONNE PHYSIQUE- PERSONNE PHYSIQUE

La fiabilité de ces classements est très liée à la qualité des variables concernant le propriétaire issues des différents millésimes des Fichiers fonciers. Ces variables s'améliorent de millésime en millésime mais certains classements ont dû être élaborés spécifiquement pour DV3F. La fiabilité de cette variable est fluctuante entre certaines catégories. Les premiers retours d'expérience montrent que la fiabilité est moindre pour les catégories aménageur et promoteur. Ceci reste à confirmer avec d'autres retours d'expérience. Par ailleurs, l'identification des acheteurs/vendeurs dépend du dernier millésime de Fichiers Fonciers intégré à la base de données. L'information n'est donc pas disponible pour les transactions les plus récentes.

Exemple : Top 5 des typologies de vendeurs et acquéreurs

Si l'on étudie l'ensemble des mutations en Occitanie entre 2012 et 2015, on constate que les typologies ayant principalement vendu et acquis sont les suivantes :

Vendeurs	Pourcentage de ventes sur l'ensemble des mutations
Personne physique (X0)	81%
Personne morale (M0)	5%
Promoteur immobilier privé (F6)	4%
Société civile immobilière (G1)	3%
Commune (P5)	2%

Acquéreurs	Pourcentage de ventes sur l'ensemble des mutations
Personne physique (X0)	85%
Société civile immobilière (G1)	5%
Personne morale (M0)	3%
Commune (P5)	3%
Département (P3)	1%

Les personnes physiques sont très majoritairement les principaux vendeurs et acquéreurs.

2 Prise en main de la base sous postgreSQL

2.1 Installer une base de données DV3F

- Installer un serveur PostgreSQL version 9.x (DV3F a été développé sous la version 9.3)
- Installer l'extension spatiale PostGIS version 2.x (DV3F a été développé sous la version 2.1)
- Créer une base de données vide
- Importer le jeu de données fourni dans la base de données vide

Le déroulé pas à pas se trouve à l'adresse : http://doc-datafoncier.cerema.fr/dv3f/tuto/installation_dv3f

2.2 Recommandations méthodologiques pour l'analyse des marchés immobiliers

2.2.1 Constituer de nouvelles tables

La constitution de nouvelles tables peut être utile pour réaliser des analyses. Par exemple, la constitution d'une table « périmètre » peut permettre de constituer des périmètres d'analyse qui pourront ensuite être appelés dans les requêtes.

De même, un travail d'identification des terrains à bâtir (TAB) peut entraîner la création d'une nouvelle variable qui regroupera tous les TAB identifiés.

Dans tous les cas, la table <u>dvf.mutation doit rester dans sa version originelle</u> (celle de la livraison) afin de pouvoir faire des comparaisons à posteriori (si nécessaire). De plus, sa modification peut entraîner un dysfonctionnement de l'outil Appdvf qui s'y réfère.

2.2.2 Constituer un segment de marché

L'analyse des marchés fonciers et immobiliers implique la création de segments de marchés, cad de catégories regroupant des transactions homogènes de biens, acquis avec la même finalité. L'intérêt est de définir un ensemble homogène où les biens sont comparables en termes de prix notamment. Par exemple, le marché des maisons anciennes, le marché des appartements T4 neufs, etc.

Nb : Si la segmentation est nécessaire pour l'observation des prix, ce n'est pas nécessairement le cas pour l'observation de volumes de ventes ou de biens.

La création d'un segment de marché passe alors par la combinaison de différentes variables dans DV3F :

• une typologie de bien, par exemple : "maison" ou "appartement" (ou plus détaillé selon la convenance), cad codtypbien = 11** ou 12**

- une valeur de la variable filtre égale à "0" pour exclure toutes les mutations spécifiques et ne s'intéresser qu'aux ventes classiques*,
- une valeur de la variable devenir égale à "S" pour ne conserver que les biens stables ou inchangés (logique de non changement d'usage),
- des acheteurs et vendeurs de particulier (X0) à particulier (X0), ou bien la vente de promoteur (F6) à particulier (X0), le cas échéant,
- une échelle d'observation ou un territoire délimité (agrégation à la commune conseillée pour avoir une épaisseur statistique suffisante).

Par ailleurs, si les indicateurs DV3F apportent déjà beaucoup d'éléments, l'intégration de données exogènes peut s'avérer utile. En particulier, dans le cas des transactions de foncier, pour l'occupation du sol ou l'usage du bien, mal restitués dans DV3F.

* Il est important de faire la distinction entre une **transaction** ou **mutation** et une **vente** qui est un type de transaction (le plus courant). Lorsque l'on s'intéresse aux volumes de biens échangés, on peut considérer tous les types de transactions. En revanche, lorsque l'on construit un indicateur de prix, on ne s'intéresse plus qu'aux **ventes** et aux biens **vendus**, car les biens mutés à l'occasion d'une expropriation ou d'une vente aux enchères ne reflètent pas forcément les prix du marché.

En SQL, cela se traduit :

- par la variable idnatmut égale à 1, 2 ou 4 [idnatmut in (1,2,4)] qui ne retient que les transactions de type ventes **ou**
- par la variable filtre égale à 0 [filtre = '0'] qui supprime tous les types de transactions différents des ventes classiques et toutes les ventes atypiques **ou**
- par une combinaison de valeurs pour la variable filtre [filtre not like '%A %' and filtre not like '%E %' and filtre not like '%X %'] Cf. paragraphe 1.2.7 pour la variable filtre.

Exemple: En Occitanie, en 2016:

Nombre total de transactions d'appartements : 39 736

- Nombre total de ventes : 39 384
 - Nombre total de ventes hors ventes atypiques (filtre = 0) : 34 853
 - Nombre total de ventes hors ventes atypiques et sans évolution du bien : 34 788
 - Nombre total de ventes de 1 appartement hors ventes atypiques et sans évolution du bien : 33 542

2.2.3 Analyser le marché du logement

Considérer les ventes à l'unité

Pour analyser les prix du marché du logement, il est conseillé de ne considérer que les transactions suivantes :

- vente de 1 appartement déclaré comme logement avec une superficie >9m²,
- vente de 1 maison déclarée comme logement et superficie >9m².

En SQL, cela correspond à :

```
codtypbien like '121 %' and (ffnbloch<>0 OR ffnbloch IS NULL) AND sbati > 9 pour un appartement codtypbien like '111 %' and (ffnbloch<>0 OR ffnbloch IS NULL) AND sbati > 9 pour une maison
```

avec ffnbloch = Nombre de locaux dédiés à l'habitation, indépendamment de leurs formes physiques (maison, activité, appartement).

En effet, lorsque plusieurs biens sont concernés par une même vente, DVF n'indique qu'un montant global pour la vente.

Exemple:

Nombre d'appartements vendus en Occitanie en 2016 : 49 066

Nombre d'appartements vendus seuls en Occitanie en 2016 (nblocapt) : 35 046

Nombre d'appartements dédiés aux logements vendus seuls en Occitanie en 2016 : 34 844

Indicateur de prix en fonction de l'ancienneté du bien :

L'ancienneté du bâti influe grandement sur la valeur du bien. Lorsque le nombre de transactions est élevé, la classification suivante peut être retenue pour les appartements ou les maisons :

- VEFA: pour créer ce segment de marché, il faut extraire les VEFA des biens neufs.
- Neuf : extraire les VEFA
- Récent
- Ancien

En SQL, cela correspond à :

case

```
when codtypbien = '1111' and vefa then 'VEFA'::text -- on extrait les vefa du neuf when codtypbien ='1111' and not vefa then 'neuf'::text when codtypbien ='1112' then 'recent'::text when codtypbien ='1113' then 'ancien'::text else null::text
```

end as ancien,

Les biens neufs (hors VEFA) et récents peuvent aussi être regroupés dans une même catégorie.

La catégorie des biens anciens peut également être sous segmentée par grande période de construction.

NB1 : Tous les biens dont la date de construction n'est pas connue sont exclus de fait de cette classification.

NB2 : Le marché de la maison neuve ne représente pas la production de maisons neuves car la majeure partie des maisons neuves sont construites par une personne physique suite à l'achat d'un terrain à bâtir.

Appliquer le secret statistique¹

Le secret statistique garantit le respect :

- -de la confidentialité due à la vie privée, personnelle et familiale, pour les personnes physiques ;
- -du secret commercial et des affaires, pour les entreprises.
- « **S'agissant des particuliers**, il est interdit de publier des données qui permettraient une identification indirecte d'une personne ou même, sans pouvoir l'identifier, d'obtenir une information à son sujet. Ces règles limitent la finesse des informations disponibles en diffusion. Des règles très strictes sont fixées spécifiquement pour les recensements de la population.

S'agissant des entreprises, aucun résultat n'est publié s'il concerne moins de trois entreprises, ni aucune donnée pour laquelle une seule entreprise représente 85% ou plus de la valeur obtenue. Cependant, il est admis que la diffusion de listes extraites du répertoire des entreprises ou des établissements peut mentionner l'activité économique, une classe d'effectifs et une tranche de chiffres d'affaires. »

"Pour la diffusion de tableaux tirés d'informations fiscales, la règle est qu'aucune case ne doit comporter moins de onze individus"

En pratique, aucun résultat d'indicateur ne doit être diffusé s'il comporte moins de 11 ventes, 11 biens vendus, etc.

Cette application du secret statistique aux données DV3F sera certainement assouplie du fait de la mise en open-data des données DVF.

Néanmoins, pour créer des références de prix, un nombre minimal de ventes doit être considéré (par exemple 30 ventes).

Vérifier les valeurs extrêmes

Il est possible que des erreurs de saisie soient présentes dans la base DV3F.

L'erreur de saisie sur le montant de la transaction est une des erreurs possibles. Elle n'est détectable que si l'erreur aboutit à un prix très élevé sans rapport avec le bien vendu.

Aussi, il est recommandé de vérifier les valeurs de transaction les plus élevées afin de voir si elles semblent cohérentes.

select * from dvf.mutation where valeurfonc is not null order by valeurfonc desc limit 50

Avec cette requête, on identifie dans le département du Lot une vente de terrain nu de 11 000 m² pour un montant de 244 000 000 €, soit 22 234 €/m². Ce prix ne semble pas réaliste. Par conséquent, il serait bon de supprimer cette transaction (en repérant son idmutation) lorsqu'on calcule le montant total de transactions.

¹ source: https://www.insee.fr/fr/information/1300624

select count(*) as nb_mut_foncier, sum(valeurfonc) as montant_foncier from dvf.mutation where anneemut between 2010 and 2016 and codtypbien like '2%' and idmutation != 7561....

Ainsi, le montant global des mutations de foncier nu dans le département du Lot entre 2010 et 2016 passe de 281 650 382 € à 37 650 382 €.

Privilégier la médiane des prix

Dans les analyses de marchés, il est conseillé d'utiliser le prix médian des ventes (la moitié des biens se sont vendus en deçà du prix médian et l'autre moitié des biens se sont vendus plus chers que le prix médian) car la valeur médiane est peu sensible aux valeurs extrêmes, contrairement à la valeur moyenne.

2.2.4 Estimer le devenir du bien en fonction du type d'acquéreur

Le type d'acquéreur du bien peut donner des indications intéressantes sur l'usage ou le devenir du bien. Par exemple, un terrain acheté par un agriculteur aura certainement une vocation agricole, une maison acquise par un promoteur immobilier en centre urbain fera certainement l'objet d'une démolition/recontruction pour de l'habitat collectif. De même, on peut considérer que l'achat d'un terrain à bâtir par une personne physique sera voué à la construction de la maison individuelle de l'acquéreur. Pour ce type d'analyse, on utilisera les variables correspondant au type d'acquéreur (codtypproa).

Toutefois, le nom de l'acquéreur n'est pas toujours identifié.

Lorsque l'on obtient une version de DV3F, il est important de connaître le dernier millésime de fichiers fonciers rapatrié dans la version en question avant d'exploiter les résultats sur les acheteurs/vendeurs. En effet, le vendeur ne sera connu au mieux qu'un an après la vente. A titre d'exemple, une version de DV3F d'octobre 2018 intégrant le millésime 2017 des Fichiers Fonciers comporte les informations sur les propriétaires des biens au 1er janvier 2017. Par conséquent, les acheteurs/vendeurs d'une transaction ayant eu lieu à compter de la fin 2016 ne sont pas connus. Ainsi, lorsque l'on souhaite travailler sur les logiques d'acteurs, il est préférable de travailler sur la période couverte par les Fichiers fonciers.



Sur les années 2016 et 2017, les données sont partielles. Ce paramètre est à prendre en compte lors d'une analyse des typologies de vendeurs ou acquéreurs dans une étude.

2.2.5 Analyser le marché des terrains à bâtir

L'analyse des marchés fonciers est rendue difficile par le manque d'information concernant le devenir des biens dans DVF. En particulier la catégorie « terrain de type terrain à bâtir » est insuffisamment fiable et très partielle. Elle est donc à manier avec précaution.

Les variables devenir et segmtab de DV3F permettent toutefois de mieux identifier les TAB pour créer un segment de marché TAB et extraire les terrains en question des autres segments de marchés.

Dans l'état actuel des connaissances, pour une analyse rapide concernant les terrains à bâtir, il nous paraît opportun de considérer comme TAB tous les terrains nus dont la valeur de segmtab est égale à 4 ou 3 (fiabilité excellente ou très bonne) quelle que soit la typologie des terrains dans DV3F.

<u>Le nombre de TAB restera toutefois sous-estimé</u> notamment pour <u>les terrains vendus récemment</u> dont on n'a pas encore pu observer de changement d'usage dans les fichiers fonciers.

En effet, une version DV3F livrée en octobre 2018 dispose des fichiers fonciers jusqu'au millésime 2017 (cad au 1^{er} janvier 2017). Or, le champ devenir considère les changements d'usage dans les 4 ans après la vente. Pour les terrains vendus après le 1^{er} janvier 2013, une incertitude demeure donc sur la construction dans les 4 ans après la vente. Cette information ne sera connue qu'avec les prochains millésimes de fichiers fonciers.

Pour une analyse approfondie des marchés fonciers, un croisement de la donnée DV3F avec des données exogènes peut s'avérer utile (zonage des PLU, mode d'occupation des sols, données Sitadel recensant les permis de construire déposés, zonages agricoles déclarés à la PAC, etc.), de même que des analyses de terrain (interprétation de photos aériennes, visites de sites, entretiens d'acteurs, etc.).

Concernant les analyses de prix, la valeur du terrain à bâtir dépend de la nature du terrain et de sa destination : terrain aménagé vendu pour une construction de maison individuelle, terrain nu non aménagé qui fera l'objet d'un aménagement, voire d'un allotissement avant revente et constructions, terrain nécessitant une remise en état avant construction (dépollution par exemple), terrain acheté par un promoteur pour une construction de logements collectifs (droit à bâtir), etc.

Lorsque l'on s'intéresse au marché du logement, on considère qu'une personne physique souhaitant se loger opte pour un appartement, une maison ou l'achat d'un TAB viabilisé pour y construire sa maison. C'est pourquoi, il peut être intéressant d'analyser le segment de marché « terrain à bâtir acquis par une personne physique » en supposant que le terrain est prêt à bâtir.

A titre d'exemple, si l'on étudie les ventes (idnatmut ={1,2,4 }) de terrains à bâtir en Occitanie entre 2012 et 2015, il est possible de définir différentes sélections de mutations. Avant toute analyse, il conviendra de comprendre les requêtes exécutées afin d'éviter de possibles erreurs d'interprétation.

Ainsi, les résultats de trois requêtes possibles sont affichés dans ce tableau :

	Nombre de ventes	Médiane de la valeur des ventes	Médiane de la valeur des ventes
1 - Ensemble des TAB fiscaux (codtypbien=21)	36 000	76 000 €	121 €/m²
2 - Ensemble des ventes foncières (codtypbien like '2%') ET Segmtab = (3 et 4)	45 000	74 000 €	104 €/m²
3 - Ensemble des ventes foncières (codtypbien like '2%') ET Segmtab = (1, 2, 3 et 4)	76 000	54 000 €	54 €/m²

La 1ère requête est déconseillée compte tenu du manque de fiabilité de la donnée TAB fiscaux seule (cf. 1.2.8). Il est par contre conseillé de retenir les valeurs de la 2° requête, comme expliqué dans le paragraphe ci-dessus. Ces valeurs peuvent être complétées par une analyse des acquéreurs des terrains :

Ensemble des ventes foncières (codtypbien like '2%') ET Segmtab = (3 et 4)	Nombre de ventes	Surface médiane des terrains	Médiane de la valeur des ventes	Médiane de la valeur des ventes
Acquis par des particuliers (codtypproa LIKE 'X0')	39 000	650 m²	71 000 €	114 €/m²
Acquis par des « non particuliers » (codtypproa not LIKE 'XO')	6 000	1 600 m²	90 000 €	60 €/m²
Acquis par des professionnels de l'immobilier (codtypproa LIKE '%F%')	1 000	2 600 m ²	180 000 €	68 €/m²

Sur les 45 000 ventes, 39 000 sont des acquisitions réalisées par des particuliers. Si l'achat de terrains à bâtir par les « non particuliers » est globalement plus élevé, le montant de la vente est quasiment deux fois moins important lorsqu'il est rapporté au m².

Distinction nombre de transactions/nombre de parcelles pour le foncier nu

Une transaction de foncier nu peut concerner plusieurs parcelles. Or, DVF indique un montant unique et la superficie totale de terrain pour la transaction quel que soit le nombre de parcelles concernées par celle-ci, ce qui permet de calculer le prix au m². Pour l'indicateur de prix, il est donc conseillé de prendre en compte la vente et non pas le nombre de parcelles vendues (ce qui revient à considérer qu'une vente correspond à un terrain vendu quel que soit le nombre de parcelles).

Exemple :

En Occitanie, en 2016, le nombre de ventes de foncier nu est de 36 232. L'ensemble de ces ventes ont concerné 103 314 parcelles nues.

Comme l'indique le diagramme ci-contre, la majorité des ventes ne contenait qu'une seule parcelle.

Cependant, 44 % des ventes en impliquaient plusieurs, et 12 % plus de 5. Il convient donc nombre de ventes et nombre de parcelles vendues.

