

2 – Données et champ

- 2.1 – Le cadre institutionnel
- 2.2 – Présentation des bases
- 2.3 – Champ des indices
- 2.4 – Variables utilisées et traitement des valeurs manquantes

2.1 – Le cadre institutionnel

Lors de la transaction d'un bien immobilier en France, un acte de vente doit être réalisé chez un notaire. Dans le cas d'un logement existant, cet acte doit obligatoirement mentionner les informations suivantes :

- concernant les parties :
 - les coordonnées du vendeur et de l'acquéreur ;
- concernant le bien :
 - l'adresse du bien ;
 - l'origine du bien (date du précédent acte de vente, nom du précédent propriétaire, acte notarié...) ;
 - le descriptif détaillé du bien, de ses équipements et annexes ;
 - l'existence d'une hypothèque et/ou d'une servitude ;
- concernant la vente :
 - le montant des honoraires du professionnel chargé de la vente et à qui en incombe le paiement ;
 - le prix de vente et ses modalités de paiement (avec ou sans l'aide d'un prêt immobilier) ;
 - la date de disponibilité du bien ;
 - les conditions suspensives lorsqu'il en existe.

Le notaire est un officier public, intervenant dans l'ensemble des domaines du droit incluant l'immobilier, mais également la famille, la fiscalité et le patrimoine, les entreprises, les zones rurales, les collectivités locales... Agissant pour le compte de l'État et nommé par arrêté du ministre de la justice, il possède de véritables prérogatives de puissance publique et confère aux actes qu'il rédige, par sa signature, l'authenticité.

Légalement, la mission de service public du notariat a été définie par l'article premier de l'ordonnance n° 45-2590 du 2 novembre 1945 : « Les notaires sont les officiers publics, établis pour recevoir tous les actes et contrats auxquels les parties doivent ou veulent faire donner le caractère d'authenticité attaché aux actes de l'autorité publique et pour en assurer la date, en conserver le dépôt, en délivrer des grosses et expéditions. »

Depuis la loi n° 2011-331 du 28 mars 2011, complétée par le décret d'application n° 2013-803 du 3 septembre 2013 et de ses arrêtés du 30 septembre 2016, la collecte d'informations statistiques sur les transactions de logement devient également partie de la mission de service public des notaires qui « contribuent à la diffusion des informations relatives aux mutations d'immeubles à

titre onéreux. Ils transmettent au Conseil supérieur du notariat³ les données nécessaires à l'exercice de cette mission de service public ».

Cette même loi reconnaît aussi comme mission de service public l'activité de centralisation et de diffusion des données collectées par le Conseil supérieur du notariat :

- « Le Conseil supérieur du notariat met gratuitement des résultats statistiques à la disposition du public sur le réseau internet. »
- « Le Conseil supérieur du notariat, ou son délégataire, transmet à toute personne qui le demande, pour au moins vingt mutations, les informations relatives à la transaction, au prix et aux caractéristiques essentielles de chaque bien. »
- « Le Conseil supérieur du notariat, ou son délégataire, transmet à toute personne qui le demande, moyennant le paiement d'une rémunération, un ou plusieurs tableaux de résultats statistiques (...). »

2.2 – Présentation des bases

Les indices Notaires-Insee sont calculés à partir des informations enregistrées dans les bases de données notariales, qui sont au nombre de deux :

- La base BIEN, gérée par PNS (Paris Notaires Services), couvre l'Île-de-France. Elle a été mise en place à partir de 1989 pour Paris, 1991 pour la petite couronne et 1996 pour la grande couronne ; elle est exploitable pour les indices depuis 1991.
- La base Perval, gérée par l'ADSN (Association pour le développement du service notarial, dépendant du notariat), couvre la province et les Drom. Elle a été mise en place et exploitée à partir de 1994.

Ces bases enregistrent les transactions de tous les types de biens : appartements et maisons, mais aussi immeubles, locaux d'activités, terrains, garages, vignobles et autres biens agricoles. Les informations sont récupérées dans les actes de ventes, permettant de recueillir des informations détaillées pour les biens échangés.

Les actes de vente de biens immobiliers sont transmis directement par les offices notariaux situés sur l'ensemble du territoire national. Les biens situés en province et dans les Drom hors Mayotte alimentent la base Perval, tandis que ceux situés en Île-de-France sont enregistrés dans la base BIEN.

Il existe deux filières d'alimentation de la base de données :

- l'envoi en version papier était le seul en vigueur jusqu'en 2009 : les offices notariaux transmettent aux gestionnaires des bases des copies des actes de vente complétées de fiches descriptives des biens. Les informations sont ensuite contrôlées et saisies par des codificateurs ;
- la télétransmission des actes est à la disposition des offices depuis 2010 : elle se fait à partir du logiciel de rédaction d'actes lors de la saisie de la vente dans l'office notarial. L'acte de vente est ensuite télétransmis *via* l'intranet de la profession pour intégrer la base de données immobilières, où sont effectués des contrôles qualité par les codificateurs du côté des gestionnaires des bases.

³ Le Conseil supérieur du notariat est un établissement d'utilité publique créé par l'ordonnance du 2 novembre 1945 ; il représente la profession auprès des pouvoirs publics, détermine sa politique générale, contribue à l'évolution du notariat et fournit des services collectifs aux notaires.

Le délai de réception, conditionnant la réactivité des indices, dépend du mode de transmission : en 2016, le délai est en moyenne de 18 jours pour les actes télétransmis et 59 jours pour les actes papiers en province (respectivement 19 jours et 47 jours en Île-de-France). La télétransmission permet ainsi un gain notable sur les délais. Elle se développe progressivement depuis 2010 et a vocation à devenir à terme le seul vecteur de transmission. Ce processus est déjà bien avancé : en 2016, 67 % des actes de province ont été télétransmis et 54 % en Île-de-France.

L'envoi des données par les notaires s'effectuait jusqu'à fin 2016 sur une base volontaire, le taux de couverture de la base de données n'était donc pas de 100 %. Depuis le 1^{er} janvier 2017, la loi⁴ oblige les notaires à alimenter les bases, aussi le taux de couverture a progressé, même si la phase de montée en charge n'est pas terminée. Il se situait autour de 60 % début 2017. Ce taux de couverture est estimé en comparant le montant des droits de mutation figurant dans les bases de données avec celui perçu par la direction générale des Finances publiques (DGFIP) (voir partie 4.4.1).

Fin 2018, les bases contenaient 18 millions d'enregistrements correspondant aux mutations des années 1990 à 2017 (25 % en Île-de-France et 75 % en province), dont 11 millions portaient sur des logements anciens au sens fiscal (voir partie 2.3). Pour l'année 2017, elles contenaient 910 000 mutations, dont 610 000 portaient sur des logements anciens au sens fiscal.

2.3 – Champ des indices

Les indices couvrent aujourd'hui l'ensemble de la France excepté Mayotte. Ils portent sur les maisons et appartements anciens au sens fiscal (c'est-à-dire non soumis en partie ou en totalité à la TVA) : les logements neufs ou ayant subi une rénovation lourde sont donc exclus. Sont également exclus les logements atypiques tels que chambres, greniers, lofts, ateliers, loges de gardien, châteaux, grandes propriétés, hôtels particuliers. De plus, les logements doivent être vendus de gré à gré, libres d'occupation au moment de la vente⁵, destinés à un usage strict d'habitation et acquis en pleine propriété par un particulier ou une société civile immobilière. Sur le parc des logements anciens, environ 12 % des appartements et 13 % des maisons en province (respectivement 11 % et 7 % en Île-de-France) sont exclus du champ des indices en appliquant ces critères (calcul effectué sur les données de 2012 à 2016).

En outre, afin d'éviter les enregistrements trop incomplets ou atypiques, voire faux, les transactions retenues doivent respecter les caractéristiques suivantes :

- Pour les appartements :
 - le nombre de pièces est renseigné et strictement inférieur à 9 ;
 - la surface habitable réelle ou estimée⁶ est comprise entre 10 m² et 200 m² (inclus) ;
 - le prix de vente est renseigné et compris entre 1 500 euros et 5 000 000 euros (exclus) ;

⁴ Les arrêtés du décret n°2013-803 du 3 septembre 2013 sont parus le 30 septembre 2016. Avec leur publication, la loi n°2011-331 du 28 mars 2011 relative à la modernisation des professions judiciaires et juridiques est entrée en vigueur, conférant aux notaires une nouvelle mission de service public au travers de la collecte et de la diffusion des statistiques immobilières issues de leurs contrats et avant-contrats de vente.

⁵ Les logements sont retirés lorsque la période d'occupation par un tiers ou par le vendeur excède six mois en considérant que, compte tenu de la réglementation en matière de baux locatifs, ces logements subissent en général une décote.

⁶ Lorsque la surface habitable n'est pas disponible, elle est estimée à partir d'autres informations. Voir partie 2.4.

- le prix au m² est strictement inférieur à 25 000 euros ;
- la date de transaction est correctement renseignée ;
- la commune de la transaction est bien présente dans le référentiel géographique.
- Pour les maisons :
 - le nombre de pièces est renseigné et strictement inférieur à 13 ;
 - la surface habitable réelle ou estimée est comprise entre 20 m² et 300 m² (inclus) ;
 - la surface du terrain est renseignée ;
 - le prix de vente est renseigné et compris entre 1 500 euros et 15 000 000 euros (exclus) ;
 - la date de transaction est correctement renseignée ;
 - la commune de la transaction est bien présente dans le référentiel géographique.

Le prix retenu est le prix net vendeur, hors droits de mutation, frais de notaire et commission d'agence.

Au total, après ces deux étapes, on conserve, dans le champ des indices, 87 % des appartements et 80 % des maisons en province (86 % pour les appartements et maisons d'Île-de-France).

2.4 – Variables utilisées et traitement des valeurs manquantes

Parmi les variables collectées dans les bases notariales, celles retenues dans le cadre des indices sont, au-delà du prix de vente et de la date de mutation, celles qui ont un impact sur le prix du bien. Elles doivent en outre être suffisamment bien renseignées pour permettre une estimation pertinente de l'équation des prix.

Les caractéristiques ainsi retenues sont :

- pour les appartements :
 - la localisation⁷ ;
 - la distance en kilomètres à la commune de 50 000 habitants la plus proche pour les strates de province, la distance à Paris pour les strates d'Île-de-France ;
 - la surface habitable ;
 - le nombre de pièces ;
 - l'époque de construction ;
 - le nombre de salles de bains ;
 - le nombre de garages ou parkings ;
 - l'état du bien⁸ ;
 - l'étiquette énergétique ;
 - l'étage ;
 - la présence d'un ascenseur (à partir du 4^e étage) ;
 - la présence d'une terrasse (ou balcon ou loggia) ;
 - la présence d'une cave.

⁷ La localisation, ayant un fort impact sur la détermination des prix, est prise en compte de plusieurs manières : *via* la stratification, puis au sein de chaque modèle. Ces points sont détaillés dans les chapitres suivants.

⁸ L'état, quand il est renseigné, est le résultat d'une appréciation faite par les acteurs de la transaction. Il peut prendre trois valeurs : bon état, travaux à prévoir, à rénover.

- pour les maisons :
 - la localisation ;
 - la distance en kilomètres à la commune de 50 000 habitants la plus proche pour les strates de province, la distance à Paris pour les strates d'Île-de-France ;
 - la surface habitable ;
 - la surface du terrain ;
 - le nombre de pièces ;
 - l'époque de construction ;
 - le nombre de salles de bains ;
 - le nombre de garages ou parkings ;
 - l'état du bien ;
 - l'étiquette énergétique ;
 - la présence d'une cave ou d'un sous-sol ;
 - le nombre de niveaux.

Les caractéristiques présentes dans les bases notariales peuvent parfois être incomplètes. Lorsqu'une transaction est enregistrée dans les bases mais qu'une ou plusieurs variables relatives au bien échangé sont manquantes⁹, un traitement est effectué sur ces dernières afin de pouvoir utiliser la transaction dans le calcul des indices. Selon les cas, les variables peuvent être imputées (par estimation ou recodage) ou bien la valeur manquante est considérée en tant que telle dans les modèles (ajout d'une modalité « inconnu »). Les règles retenues sont synthétisées dans les figures 2-1 et 2-2.

Figure 2-1 : Règles de traitement de la non-réponse pour les valeurs manquantes - Appartements

Appartements				
<i>Variable</i>	<i>Modalité attendue</i>	<i>Action en cas de non-réponse</i>	<i>Pourcentage de non-réponse en 2017</i>	
			<i>Île-de-France</i>	<i>Province</i>
Surface habitable	Valeur numérique	Estimation linéaire en fonction du nombre de salles de bains, du nombre de pièces, de l'époque de construction, de l'étage et du nombre de garages. Les coefficients sont remis à jour tous les 2 ans.	2 % (avant imputation)	2 % (avant imputation)
Nombre de salles de bains	Valeur numérique	Recodage à 1 salle de bains.	2 %	2 %

⁹ Les informations essentielles, à savoir la nature du bien (appartement ou maison), le prix, la date et le lieu de mutation, le nombre de pièces et (pour les maisons) la surface du terrain doivent au minimum être présentes. Les observations ne respectant pas ces critères ont été exclues lors de la sélection du champ (cf. partie 2.3).

Appartements				
Époque de construction	10 tranches	Si possible, imputation à partir des ventes réalisées à la même adresse sur les 10 dernières années. ¹⁰ Sinon, ajout d'une modalité « inconnu ».	25 % (après imputation)	17 % (après imputation)
Étage	Valeur numérique	Recodage à 0 (rez-de-chaussée).	1 %	2 %
Présence d'un ascenseur (à partir du 4 ^e étage)	Oui/non	Si possible, imputation à partir des ventes réalisées à la même adresse sur les 10 dernières années ¹¹ . Sinon, recodage à « oui ».	37 % (après imputation)	31 % (après imputation)
Présence d'une cave	Oui/non	Recodage à « non » ¹² .	Proche de 0 %	9 %
Nombre de garages ou parkings	Valeur numérique	Recodage à 0 (pas de garage ni parking) ¹³ .	3 %	8 %
Présence d'une terrasse (ou balcon ou loggia)	Oui/non	Ajout d'une modalité « inconnu ».	23 %	13 %
État du bien	3 modalités	Ajout d'une modalité « inconnu ».	82 %	70 %
Étiquette énergie	8 modalités	Ajout d'une modalité « inconnu ».	41 %	32 %

¹⁰ Le principe est de redresser ces variables en prenant les informations fournies par les autres ventes disponibles dans le même bâtiment au cours des 10 dernières années, lorsqu'il y en a, dans les bases immobilières notariales. Afin de s'assurer que l'information est suffisamment fiable, ce traitement est effectué uniquement s'il y a au moins cinq ventes à la même adresse, et dont au moins 60 % indiquent la même valeur. Ce traitement n'est fait qu'en appartements, en raison du peu de maisons dans les bases qui ont fait l'objet d'au moins deux ventes.

¹¹ Avec les mêmes contraintes que pour l'époque de construction.

¹² Un numéro de lot doit être indiqué sur l'acte s'il y a une cave.

¹³ Un numéro de lot doit être indiqué sur l'acte s'il y a un garage.

Figure 2-2 : Règles de traitement de la non-réponse pour les valeurs manquantes - Maisons

Maisons				
<i>Variable</i>	<i>Modalité attendue</i>	<i>Action en cas de non-réponse</i>	<i>Pourcentage de non-réponse en 2017</i>	
			<i>Île-de-France</i>	<i>Province</i>
Surface habitable ¹⁴	Valeur numérique	Estimation linéaire en fonction du nombre de salles de bains, du nombre de pièces, de l'époque de construction et du nombre de niveaux. Les coefficients sont remis à jour tous les 2 ans.	23 %	8 %
Nombre de salles de bains	Valeur numérique	Recodage à 1 salle de bains.	1 %	2 %
Époque de construction	10 tranches	Ajout d'une modalité « inconnu ».	32 %	13 %
Nombre de niveaux	2 modalités	Recodage à « 2 niveaux ou plus ».	0 %	2 %
Présence d'une cave ou d'un sous-sol	Oui/non	Recodage à « non ».	7 %	18 %
Nombre de garages ou parkings	Valeur numérique	Ajout d'une modalité « inconnu ».	15 %	12 %
État du bien	3 modalités	Ajout d'une modalité « inconnu ».	82 %	57 %
Étiquette énergie	8 modalités	Ajout d'une modalité « inconnu ».	32 %	31 %

¹⁴ La méthode utilisée pour estimer la surface habitable des maisons anciennes a fait l'objet d'une attention particulière, car le taux de non-réponse est assez élevé sur ce marché. En effet, la loi Carrez (entrée en vigueur le 19 juin 1997), obligeant le vendeur à déclarer la surface sur l'acte de vente, n'est valable que pour les appartements. Le modèle retenu est finalement proche de celui utilisé en version 3, à l'exception de l'ajout dans les modèles de quelques variables ainsi que de la mise à jour plus fréquente des coefficients.