

Sommaire

1 L'existence de pondérations et leurs principes.....	2
1.1 Pourquoi des pondérations ?.....	2
1.2 Leur principe général.....	2
1.3 Leurs caractéristiques.....	3
1.3.1 Des poids à décimales.....	3
1.3.2 Identité du poids pour le logement et ses occupants.....	3
1.3.3 Population à l'exploitation principale = population légale municipale.....	3
2 Les poids à l'exploitation principale – Logements ordinaires et population des ménages.....	4
2.1 Les communes de 10 000 habitants ou plus.....	4
2.1.1 Pour la France métropolitaine.....	4
2.1.2 Dans les DOM.....	6
2.2 Les communes de moins de 10 000 habitants.....	7
2.2.1 Pour les communes enquêtées en 2015 (N+1) ou 2016 (N+2).....	7
2.2.2 Pour les communes enquêtées en 2014 (N).....	9
2.2.3 Pour les communes enquêtées en 2012 (N-2) ou 2013 (N-1).....	9
3 Les poids à l'exploitation principale – Population hors ménage.....	11
3.1 Les communautés.....	11
3.1.1 Pour les communautés enquêtées en 2012 (N-2) ou 2013 (N-1).....	12
3.1.2 Pour les communautés enquêtées en 2014 (N).....	12
3.1.3 Pour les communautés enquêtées en 2015 (N+1) ou 2016 (N+2).....	12
3.1.4 Pour les communautés existantes au 1er janvier 2014 (N) mais jamais encore enquêtées.....	13
3.1.5 Pour les communautés existantes au 1er janvier 2014 (N), non enquêtées au cours du cycle de 5 ans de 2012 (N-2) à 2016 (N+2) mais enquêtées au cours du cycle précédent.....	13
3.2 Les autres populations hors ménages.....	14
4 Les poids à l'exploitation complémentaire.....	14
4.1 Le tirage de l'échantillon de l'exploitation complémentaire.....	14
4.2 Le calcul des poids.....	15
4.3 Le calcul des poids pour les recensements millésimés 2021 à 2023.....	15
4.3.1 Pour les communes de 10 000 habitants ou plus.....	16
4.3.2 Pour les communes de moins de 10 000 habitants.....	16

1 L'existence de pondérations et leurs principes

Avec la formule de recensement par cycle d'enquêtes annuelles, **chaque observation (logement, ménage, individu, famille) est assortie d'un poids.**

Un système de pondération des observations est mis en place à l'exploitation principale, et complété à l'exploitation complémentaire pour les observations concernées¹.

L'existence de pondérations dans les fichiers des recensements n'est pas totalement nouvelle : aux recensements généraux de la population, elle était habituelle à l'exploitation complémentaire, mais n'était pas nécessaire à l'exploitation principale².

1.1 Pourquoi des pondérations ?

Le recensement de la population repose désormais sur une collecte annuelle d'information organisée par cycles de cinq ans. Les communes de moins de 10 000 habitants réalisent une enquête de recensement portant sur toute la population, à raison d'une commune sur cinq chaque année. Les communes de 10 000 habitants ou plus réalisent tous les ans une enquête par sondage auprès d'un échantillon d'adresses représentant environ 8 % de leurs logements. Les populations légales et les résultats statistiques détaillés sont calculés et diffusés à partir de cinq enquêtes annuelles successives et sont relatifs à l'année centrale de la période de cinq ans.

Ce sont ce calage à une même date et le recours pour partie à de l'échantillonnage qui expliquent l'existence d'un système de pondération appliqué aux unités statistiques enquêtées.

1.2 Leur principe général

Le fichier détail du recensement millésimé 2014, par exemple, est obtenu en empilant les fichiers des cinq enquêtes annuelles réalisées de 2012 à 2016. Il comprend une observation par individu et logement recensés au cours de la période de cinq ans considérée.

Le poids associé à chaque observation intègre :

- un poids initial de sondage (lié au taux de sondage et donc éventuellement égal à 1, pour les populations et logements enquêtés exhaustivement) ;
- un calage sur l'année médiane du cycle de cinq ans.

Ce « fichier cumul » permet, par sommation des poids, de calculer les populations légales municipales et d'obtenir l'ensemble des résultats statistiques de l'exploitation principale, du niveau infra-communal (Iris) jusqu'au niveau national.

Pour l'exploitation complémentaire, seule une partie des observations est conservée. Le poids des observations sélectionnées intègre alors une composante supplémentaire : le taux de sondage de l'exploitation complémentaire. Ce « fichier cumul » de l'exploitation complémentaire permet, par sommation des poids, d'obtenir l'ensemble des résultats statistiques de l'exploitation complémentaire, du niveau infra-communal (Iris) jusqu'au niveau national.

1 Pour plus de détails sur l'existence de deux exploitations, se référer à la fiche thématique « *Les exploitations principale et complémentaire* »

2 Au recensement général de la population de 1999, comme lors des recensements précédents, le fichier issu de l'exploitation principale n'était pas pondéré car la collecte de recensement était exhaustive et pratiquée à la même date sur l'ensemble du territoire. En revanche, le fichier issu de l'exploitation complémentaire était pondéré, car il portait sur un échantillon au quart. Les différentes unités statistiques (logement, ménage, famille, individu) étaient donc affectées d'un poids de 4, égal à l'inverse du taux de sondage. Ainsi, les résultats obtenus en totalisant les poids des unités statistiques reflétaient bien l'ensemble des logements, ménages, familles ou individus concernés.

1.3 Leurs caractéristiques

1.3.1 Des poids à décimales

Le poids attaché à chaque observation dépend de nombreux facteurs et peut ne pas être un nombre entier. Dans la majorité des cas, il comporte d'ailleurs de nombreuses décimales.

Les résultats obtenus en cumulant les observations pondérées sont calculés avec toutes ces décimales puis arrondis pour être affichés sur insee.fr.

Cela explique que dans un tableau un total puisse différer de quelques unités de la somme de ses composantes (l'arrondi d'un total n'étant pas nécessairement égal à la somme des arrondis de ses composantes).

Dans les bases téléchargeables associées aux tableaux, les données sont en revanche fournies avec toutes leurs décimales pour que tous les calculs à façon puissent être réalisés de manière autonome par leurs utilisateurs.

Quant à l'utilisateur des fichiers détail, il doit particulièrement veiller à utiliser la variable de pondération, avec toutes ses décimales, pour obtenir des résultats corrects.

1.3.2 Identité du poids pour le logement et ses occupants

Les poids associés aux observations répondent à une convention de principe : tous les individus et toutes les familles appartenant à un même ménage sont affectés du même poids, celui de leur logement.

Le respect de cette convention est une composante de la construction du système de pondération.

1.3.3 Population à l'exploitation principale = population légale municipale

Au niveau de la commune, la population issue des résultats statistiques de l'exploitation principale est strictement égale à la population municipale légale³, contrairement à ce qui prévalait dans les recensements jusqu'en 1999.

Ceci est possible car :

- le concept de population municipale a évolué et correspond désormais à la notion de population statistique (jusqu'en 1999, c'était la population sans doubles comptes et non la population municipale qui assurait cette correspondance de concept) ;
- c'est maintenant à partir du même fichier informatique que les populations légales sont établies et que les fichiers destinés aux exploitations statistiques sont produits. Dans le passé, jusqu'au recensement de 1999, les populations légales étaient établies à partir d'un comptage manuel des bulletins (appelé « dénombrement ») qui précédait la constitution du fichier statistique car les volumes d'informations à saisir auraient été trop importants pour assurer la diffusion des populations légales dans le délai prescrit, soit avant la fin de l'année du recensement. Dans certains cas, les populations issues du dénombrement et celles issues quelques mois plus tard de l'exploitation principale pouvaient différer de quelques unités.

Par ailleurs, la population statistique issue de l'exploitation principale est, depuis le recensement millésimé 2008, parfaitement égale à la population légale municipale, à tous les échelons communaux et supracommunaux (cantons, arrondissements, départements, régions, France).

Pour les recensements millésimés 2006 et 2007, cela n'était vrai que pour les communes. Aux échelons supracommunaux, il pouvait exister des écarts entre population statistique et population légale municipale, allant de quelques unités à quelques dizaines d'unités. Cela était lié aux arrondis (la population d'un département par exemple, une fois arrondie à l'entier pour publication au journal officiel,

³ Pour plus de détails sur la définition des populations légales, se reporter à la page « *Comprendre les populations légales* » sur insee.fr

n'était pas nécessairement égale à la somme des populations, arrondies à l'entier, des communes du département).

Depuis le recensement millésimé 2008, les poids bénéficient d'un calage supplémentaire de manière que population statistique à l'exploitation principale et population légale municipale soient identiques à tous les échelons communaux et supracommunaux (cantons, arrondissements, départements, régions, France).

La suite de cette fiche présente dans le détail :

- la méthode de calcul des poids à l'exploitation principale⁴, en distinguant d'une part les logements ordinaires et la population des ménages (qui représente environ 98 % de l'ensemble de la population) et d'autre part, la population hors ménage ;
- puis, comment en sont déduits les poids dans le fichier de l'exploitation complémentaire.

2 Les poids à l'exploitation principale – Logements ordinaires et population des ménages

La population des ménages est celle qui vit dans un logement ordinaire⁵.

2.1 Les communes de 10 000 habitants ou plus

2.1.1 Pour la France métropolitaine

Le calcul des poids se fait à partir du poids initial de sondage qui découle directement du plan de sondage et qui est égal à l'inverse de la probabilité de tirage.

La base de sondage des adresses (BSA) est obtenue à partir du répertoire des immeubles localisés (RIL), et constitue la référence sur le stock d'adresses et de logements de la commune.

Les adresses d'habitation de chaque commune sont réparties en quatre strates de tirage :

- les grandes adresses : ce sont des adresses auxquelles sont associées un grand nombre de logements, qu'elles correspondent à un immeuble collectif ou à un groupe de logements individuels. Le statut de « grande adresse » est attribué à partir d'un seuil minimal de nombre de logements, sous des conditions pouvant le faire varier d'une commune à l'autre et dans le temps :
 - lors de l'initialisation du dispositif (en 2004), ce seuil a été fixé à 60 logements sauf s'il conduisait à ce que la part des logements de la commune que représentaient les grandes adresses dépasse 10 % (dans ce cas, le seuil de nombre de logements a été relevé de telle sorte que cette part soit de 10 % au plus) ;
 - depuis, chaque année, ce seuil initial est revu à la hausse si la part de logements en grandes adresses représente plus de 25 % des logements du groupe de rotation.
 - par ailleurs, avant l'enquête annuelle de 2015, certains seuils trop élevés ont été revus à la baisse afin d'intégrer d'autres adresses regroupant un très grand nombre de logements, qui autrement auraient pu augmenter l'aléa de sondage si elles avaient été conservées dans la strate des autres adresses.
- les adresses nouvelles, qui correspondent à des constructions nouvelles ou des rénovations modifiant sensiblement le nombre de logements de l'adresse ;
- depuis l'enquête annuelle de recensement de 2016, les adresses correspondant à une structure touristique (hôtels, campings et résidences hôtelières), alors qu'auparavant elles n'étaient pas spécifiquement isolées des adresses d'habitation.
- les autres adresses, ni grandes, ni nouvelles, ni ne correspondant à une structure touristique.

⁴ Elle reprend la méthode de calcul des populations légales décrite dans la note sur « *La détermination de la population légale des communes* » accessible sur insee.fr

⁵ Se reporter à la fiche thématique « *Ménages et familles* » pour la définition de la population des ménages.

Ces adresses sont ensuite réparties en cinq groupes (un groupe par année du cycle de cinq ans) équilibrés au mieux, notamment quant au nombre de logements.

Puis un échantillon d'adresses est tiré dans le groupe d'adresses de l'année :

- les grandes adresses et adresses nouvelles sont toutes retenues dans l'échantillon. C'est le cas aussi des structures touristiques, du fait de leur mode d'occupation spécifique ;
- les autres adresses sont échantillonnées de sorte qu'au total environ 40 % des logements du groupe d'adresses de l'année (toutes strates d'adresses confondues) soient retenus.

Tous les logements et individus d'une adresse échantillonnée sont enquêtés exhaustivement.

Il en résulte les **poids initiaux de sondage** suivants, calculés à partir de la BSA de l'année⁶ :

- logements et individus des grandes adresses, adresses nouvelles et structures touristiques : poids égal à 1 ;
- logements et individus des autres adresses : poids égal à l'inverse du taux de sondage (en général compris entre 2,5 et 5 et plus souvent près de 3⁷).

En cumulant ces poids initiaux de sondage des observations obtenus au cours de cinq enquêtes annuelles de recensement (EAR) successives, une estimation moyenne du nombre de logements sur la période de cinq ans peut être produite. Comme dans la plupart des enquêtes, l'échantillon de logements est calé a posteriori sur la base de sondage. Dans le cas du recensement, il faut caler un échantillon final constitué du cumul de cinq échantillons annuels, tirés dans cinq bases de sondage successives. L'estimation produite étant datée du 1^{er} janvier de l'année centrale du cycle de cinq ans, c'est le stock de logements fourni par la BSA médiane⁸ qui va faire référence.

En conséquence, pour obtenir **les poids finaux** de chaque observation, les poids initiaux de sondage doivent être multipliés par le rapport suivant :

Nombre de logements de la BSA médiane

Nombre de logements estimé à partir de l'échantillon pondéré par les poids de sondage

Autrement dit, pour produire une estimation de la population, cela revient à multiplier le nombre moyen de personnes par logement obtenu avec les poids initiaux de sondage des cinq enquêtes annuelles de recensement successives, par le nombre de logements de référence issu de la BSA médiane.

Pour les adresses d'habitation, ce calage est réalisé non pas au niveau de la commune mais au niveau infracommunal de l'Iris⁹, sauf exceptions correspondant à des structures de l'habitat très atypiques, notamment les Iris d'activité ou les Iris divers. Ce calage est effectué toutes strates d'adresses confondues car les strates de tirage fluctuent avec les évolutions du parc de logements.

6 Chaque année, le RIL et donc la BSA font l'objet d'une mise à jour à l'aide de sources administratives et d'informations transmises par les communes : les immeubles détruits sont supprimés, les adresses qui apparaissent sont intégrées et classées en grandes adresses ou adresses nouvelles selon leur nombre de logements. Cette mise à jour intègre également les informations de la collecte du recensement, à partir des constats de terrain (adresse non trouvée, nombre de logements modifié, etc.). Enfin, les adresses nouvelles qui viennent d'être recensées rejoignent dans la BSA suivante la strate des « autres adresses ».

7 Puisque le poids moyen pour l'ensemble des logements et individus de la commune est de $1 / 0,40 = 2,5$ et que, par ailleurs, le poids des logements et individus des grandes adresses, nouvelles adresses et structures touristiques est de 1, les poids pour les autres adresses sont généralement supérieurs à 2,5.

8 Les BSA de chaque année étant fixées au 1^{er} juillet pour préparer la collecte se déroulant au mois de janvier suivant, le nombre de logements de référence au 1^{er} janvier de l'année médiane du cycle de 5 ans est obtenu en faisant la moyenne des deux BSA encadrant cette date, celle du mois de juillet précédent et celle du mois de juillet suivant.

9 Voir [définition](#) sur insee.fr

Un calage analogue au niveau communal est conduit pour isoler des adresses d'habitation les structures de tourisme comportant des logements (campings et résidences hôtelières). Les hôtels ne comportant pas de logement, ils ne sont pas associés à ce calage. Ils comptent à hauteur de leur collecte pendant les cinq années du cycle auxquelles leur enquête annuelle de recensement participe.

Ce traitement est appliqué indifféremment à toutes les catégories de logements, qu'ils soient des résidences principales, des résidences secondaires, des logements occasionnels ou des logements vacants¹⁰.

2.1.2 Dans les DOM

Sur la base du recensement de 1999¹¹, le territoire de chaque commune a été découpé en petites zones appelées îlots. Ces îlots ont été répartis en cinq groupes équilibrés, recensés successivement au cours de cinq ans. Pour les communes qui ont franchi à la hausse le seuil de 10 000 habitants depuis 1999, le découpage en îlots et leur répartition dans les groupes de rotation sont réalisés au moment du franchissement de seuil. Chaque année, les îlots d'un groupe font l'objet d'une enquête cartographique en vue de répertorier toutes les adresses, avec leur nombre de logements. Cette enquête permet de mettre à jour les adresses du RIL qui sont susceptibles d'être enquêtées pour la prochaine EAR. Une expertise du RIL est réalisée par chaque commune, avant le tirage d'échantillon. Le fichier qui en résulte est alors utilisé comme base de sondage.

L'échantillonnage de l'EAR pour les communes de 10 000 habitants ou plus des DOM suit un plan de sondage en deux phases. La première phase correspond à la constitution des groupes de rotations comme décrit ci-dessus. La séquence des groupes de rotation a été déterminée une fois pour toutes à l'initialisation et toutes les entités adressées d'un même îlot appartiennent au même groupe de rotation.

La seconde phase de tirage correspond à un tirage stratifié selon le type d'adresses. Jusqu'à l'EAR 2018 incluse, le tirage n'est pas stratifié par type d'entités adressées comme c'est le cas en métropole. À partir de l'EAR 2019, on introduit les strates exhaustives des grandes adresses et des adresses nouvelles. À partir de l'EAR 2020, on introduit la strate exhaustive dédiée aux établissements touristiques (hôtels, campings et résidences hôtelières), comme en métropole, et on ajoute une strate de logements en zone d'écart¹² qui sont à recenser exhaustivement. L'objectif est de recenser environ 40 % de logements du groupe de rotation annuel (hors établissements touristiques). On a donc cinq strates d'entités adressées (avec la strate des « autres adresses ») :

- Si l'entité adressée appartient à une zone d'écart, elle est affectée à la strate correspondante.
- Sinon, si l'entité adressée est un établissement touristique, elle est affectée à la strate correspondante.
- Sinon, si le nombre de logements de l'entité adressée est supérieur au seuil communal (en général 60 logements), alors elle est affectée à la strate des grandes adresses.
- Sinon :
 - Si l'entité adressée est considérée comme nouvelle dans la BSA, alors elle est affectée à la strate des adresses nouvelles.
 - Sinon l'entité adressée est affectée à la strate des « autres adresses ».

Cette seconde phase de tirage est légèrement différente pour les communes de 10 000 habitants ou plus de Mayotte. Les grandes communes de Mayotte ont intégré le protocole des EAR en 2021. Le plan de sondage est adapté pour répondre aux spécificités du terrain mahorais qui présente des zones où les adresses sont très difficilement identifiables à partir d'une adresse ou même d'un plan et où le bâti évolue rapidement.

10 Pour plus de détail sur les catégories de logement, se reporter à la fiche thématique « *Le logement* »

11 Sur la base du recensement exhaustif de 2017 pour Mayotte.

12 Les zones d'écart correspondent à des zones d'accès difficile où il est souvent impossible d'identifier précisément les entités adressées à l'aide de variables d'adressage (numéro de voie...). Ces zones se trouvent en Guyane dans certaines parties peu praticables du territoire (par exemple en bordure de fleuve) où des habitations ont pu être construites parfois de manière irrégulière.

Ces adresses correspondent à des habitations de fortune (habitations en tôle notamment), qui peuvent par ailleurs être construites et détruites rapidement. Ainsi, deux cas sont séparés :

- les habitations de fortune des îlots du groupe de rotation de l'année sont tirées exhaustivement et sont donc à recenser exhaustivement dès lors qu'elles sont situées dans un îlot appartenant au groupe de rotation recensé.
- pour les habitations en dur, la seconde phase de tirage est similaire à celle des autres DOM Sur ce champ, l'objectif est de recenser environ 40 % de logements du groupe de rotation annuel (hors établissements touristiques). Au total, le taux d'échantillonnage dans les communes de 10 000 habitants ou plus de Mayotte est plus élevé que dans les autres DOM.

Tous les logements et individus d'une adresse échantillonnée sont enquêtés exhaustivement.

Il en résulte les **poids initiaux de sondage** suivants, calculés à partir de la BSA de l'année :

- logements et individus des adresses appartenant à une zone d'écart, grandes adresses, adresses nouvelles et structures touristiques (et habitation de fortune pour Mayotte) : poids égal à 1 ;
- logements et individus des autres adresses : poids théorique égal à l'inverse du taux de sondage.

Dans le recensement de la population, les estimations sont réalisées à partir des 5 dernières enquêtes de recensement. Plus précisément, dans les communes de 10 000 habitants ou plus des DOM, chaque enquête fournit une estimation relative au groupe de rotation enquêté ; ces estimations sont ensuite sommées pour produire les résultats de l'année médiane du cycle sur chaque commune. Il paraît donc pertinent, comme dans la plupart des enquêtes, de caler l'échantillon de logements de chaque année sur la base de sondage.

Ainsi, pour améliorer la précision des estimations, on mobilise le nombre de logements du groupe de rotation recensé l'année donnée issu de la BSA, qui est exhaustif sur ce groupe de rotation via l'enquête cartographique et l'expertise de la commune.

En conséquence, **les poids finaux** de chaque logement et individu des autres adresses sont :

$$\frac{\text{Nombre de logements BSA du groupe de rotation annuel, hors strates exhaustives}}{\text{Nombre de logements recensés, hors strates exhaustives}}$$

Pour produire une estimation de la population dans chaque commune de 10 000 habitants ou plus des DOM, on somme le nombre d'individus pondérés (avec cette méthode du ratio) calculé avec les 5 dernières enquête de recensement.

2.2 Les communes de moins de 10 000 habitants

Dans les communes de moins de 10 000 habitants, la collecte est réalisée exhaustivement, donc les poids initiaux de sondage sont égaux à 1. Mais les communes ne sont pas toutes recensées la même année. Pour ramener à une même date des populations enquêtées sur un intervalle de cinq ans, il est procédé, selon les cas, à une interpolation ou à une extrapolation de la population recensée. Ce traitement se traduit par une modification des poids initiaux associés aux observations collectées.

Pour les rendre plus lisibles, les calculs de population sont explicités en prenant l'exemple du millésime 2014, année médiane du cycle d'enquêtes annuelles 2012-2016.

Dans les formules générales, l'année N est l'année médiane.

2.2.1 Pour les communes enquêtées en 2015 (N+1) ou 2016 (N+2)

Pour ces communes, il est procédé par interpolation : l'hypothèse est faite que la population varie chaque année du même nombre de personnes entre la date du dernier effectif calculé (celui millésimé 1^{er} janvier N-1) et le 1^{er} janvier de l'enquête de recensement la plus récente (réalisée en N+1 ou en N+2).

Nota Bene : cette interpolation ne fait donc aucune référence directe à la population de la précédente collecte réalisée 5 ans plus tôt, en N-4 ou en N-3.

La population des ménages au 1^{er} janvier 2014 est ainsi égale à :

- population calculée 2013 + (population enquêtée 2015 - population calculée 2013) x (1/2), pour les communes enquêtées en 2015 ;
- population calculée 2013 + (population enquêtée 2016 - population calculée 2013) x (1/3), pour les communes enquêtées en 2016.

Pour l'année N, cela équivaut à modifier les poids initiaux (égaux à 1) issus des enquêtes en N+1 ou N+2 de la manière suivante :

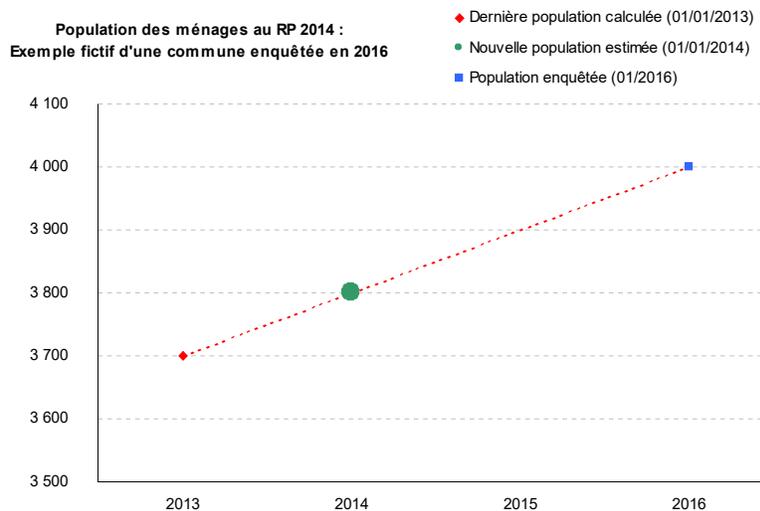
pour les communes enquêtées en N+1 : poids = $1/2 + (1/2) \times (\text{population N-1} / \text{population N+1})$
pour les communes enquêtées en N+2 : poids = $1/3 + (2/3) \times (\text{population N-1} / \text{population N+2})$

Pour les résidences principales, la convention retenue fixe qu'elles ont le même poids que la population.

Pour les logements autres que les résidences principales, l'interpolation est d'abord réalisée globalement puis son résultat est réparti selon la structure par catégorie de logement (résidences secondaires, logements occasionnels, logements vacants) observée lors de l'enquête la plus récente. Cela revient pour chaque logement en question recensé en N+1 ou N+2 à lui affecter un poids égal au nombre de logements non résidences principales calculé par interpolation pour l'année N, divisé par le nombre de logements non résidences principales effectivement constaté en N+1 ou N+2.

Si par exemple l'interpolation aboutit à estimer à 200 pour l'année N le nombre de logements non résidences principales et qu'à l'enquête en N+1 les résidences secondaires représentent 150 des 300 logements non résidences principales, alors le nombre de résidences secondaires estimé pour l'année N est de $200 \times (150/300) = 100$ et chacune des 150 résidences secondaires enquêtées en N+1 aura pour poids $200/300$ l'année N.

Illustration par un exemple fictif :



Pour les communes enquêtées en 2016, le point 2014 est estimé en faisant comme si la population avait suivi une progression régulière entre la dernière population calculée (celle de 2013) et le résultat de la collecte de 2016.

Dans cet exemple fictif, la population est de 4 000 personnes à la collecte de 2016 et avait été estimée à 3 700 personnes pour l'année 2013. La population de cette commune a donc augmenté de 300 personnes en 3 ans. L'hypothèse est faite que cette augmentation est régulière d'année en année sur cette période de 3 ans, donc de 100 personnes chaque année. Par conséquent, la dernière estimation de population au 1^{er} janvier 2013 est augmentée de 100 personnes pour fournir la nouvelle estimation de population au 1^{er} janvier 2014, soit 3 800 habitants (3 700 + 100).

En termes de pondération, les poids initiaux des observations collectées en 2016 (égaux à 1) doivent être modifiés de la manière suivante pour se caler à la date du 1^{er} janvier 2014 : $1/3 + (2/3) \times (3\ 700 / 4\ 000)$, soit 0,95 (équivalent à 3 800 / 4 000).

2.2.2 Pour les communes enquêtées en 2014 (N)

Pour ces communes, la population des ménages obtenue lors de la collecte est conservé telle quelle.

Pour l'année N, cela revient à laisser inchangés les poids initiaux issus de l'enquête en N :

pour les communes enquêtées en N : poids = 1

2.2.3 Pour les communes enquêtées en 2012 (N-2) ou 2013 (N-1)

Pour ces communes, il est procédé à une extrapolation de la dernière population calculée pour le 1^{er} janvier 2013, dont le principe est de se baser sur une évaluation de l'évolution du nombre de résidences principales entre 2013 et 2014, corrigée de l'évolution annuelle moyenne du nombre de personnes par ménage constatée entre les deux dernières enquêtes de recensement.

Le nombre de résidences principales à estimer pour le 1^{er} janvier 2014 est évalué à partir de l'évolution du nombre de logements fournie par les fichiers de la taxe d'habitation (TH) de 2013 et 2014. Il est obtenu en appliquant à ce nombre de logements actualisé la répartition par catégorie de logement (part des résidences principales parmi l'ensemble des logements) observée lors de la collecte de recensement la plus récente (2012 ou 2013 selon les cas).

Pour passer d'une évolution du parc de logements à une évolution du nombre d'habitants, le principe est de tenir compte du fait que, souvent, la population et le nombre de ménages ne croissent pas exactement au même rythme¹³. Pour cela, au taux d'évolution du nombre de résidences principales estimé en utilisant les fichiers de la taxe d'habitation est appliqué un coefficient correcteur correspondant à l'évolution annuelle moyenne de la taille des ménages observée entre les deux dernières enquêtes de recensement (2007-2012 ou 2008-2013 selon les cas).

Sans autre traitement, cette évolution différenciée aboutirait néanmoins à introduire une distorsion entre le poids des logements et le poids de leurs occupants, en contradiction avec la convention « *tous les individus et toutes les familles appartenant à un même ménage sont affectés du même poids, celui de leur logement* ».

Afin de réconcilier poids des logements et poids de leurs occupants, il est procédé à un calage sur marge qui conduit à multiplier les poids par un facteur de calage (dans le cas d'une tendance à la baisse de la taille moyenne des ménages, les poids des logements comptant peu d'occupants sont réhaussés alors que les poids des logements comptant beaucoup d'occupants sont abaissés). Ce calage n'est valide que si la commune comporte suffisamment de logements. En pratique, il n'est appliqué qu'aux communes de plus de 2 000 habitants.

¹³ En moyenne, et dans la plupart des communes, la taille des ménages tend à diminuer (le nombre de ménages augmente plus vite que celui des habitants).

Pour les autres communes, à défaut de pouvoir corriger les pondérations par calage sur marge, il est fait l'hypothèse que le nombre de résidences principales augmente au même rythme que la population, qui tient compte de l'évolution de la taille des ménages.

Pour le calcul de la population des ménages de l'année N, cela équivaut à modifier les poids initiaux (égaux à 1) issus des enquêtes en N-1 ou N-2 de la manière suivante :

pour les communes de plus de 2 000 habitants enquêtées en N-1 :

Poids =

(nombre logements selon la TH l'année N / nombre de logements selon la TH l'année N-1)
x (évolution annuelle moyenne du nombre de personnes par ménage entre N-6 et N-1)
x facteur de calage

pour les communes de plus de 2 000 habitants enquêtées en N-2 :

Poids =

(nombre logements selon la TH l'année N / nombre de logements selon la TH l'année N-2)
x (évolution annuelle moyenne du nombre de personnes par ménage entre N-7 et N-2)²
x facteur de calage

pour les communes de moins de 2 000 habitants enquêtées en N-1 :

Poids =

(nombre logements TH l'année N / nombre de logements TH l'année N-1)
x (évolution annuelle moyenne du nombre de personnes par ménage entre N-6 et N-1)

pour les communes de moins de 2 000 habitants enquêtées en N-2 :

Poids =

(nombre logements TH l'année N / nombre de logements TH l'année N-2)
x (évolution annuelle moyenne du nombre de personnes par ménage entre N-7 et N-2)²

L'extrapolation des logements non résidences principales (résidences secondaires, logements occasionnels, logements vacants) est également réalisée à l'aide de la TH (pour la France métropolitaine et la Réunion), en leur affectant les poids suivants :

- *pour une commune enquêtée en 2012 :
Nombre de logements TH 2014 / Nombre de logements TH 2012*
- *pour une commune enquêtée en 2013 :
Nombre de logements TH 2014 / Nombre de logements TH 2013*

Le cas particulier de la Guadeloupe, de la Martinique et de la Guyane

Le manque d'exhaustivité du fichier de la taxe d'habitation pour la Guadeloupe, la Martinique et la Guyane ne permet pas d'utiliser ses informations pour mener l'extrapolation comme en métropole et à La Réunion. En l'absence de cette possibilité, l'extrapolation est réalisée en prolongeant l'évolution tendancielle de la population constatée entre les deux derniers recensements, soit entre 2007 et 2012 ou entre 2008 et 2013 selon les cas.

La population des ménages au 1^{er} janvier 2014 est ainsi égale à :

- *pour les communes enquêtées en 2013
population enquêtée en 2013 x (population enquêtée en 2013 / population enquêtée en 2008)^(1/5)*
- *pour les communes enquêtées en 2012
population enquêtée en 2012 x (population enquêtée en 2012 / population enquêtée en 2007)^(2/5)*

Pour le calcul de la population des ménages de l'année N, cela équivaut à modifier les poids initiaux (égaux à 1) issus des enquêtes en N-1 ou N-2 de la manière suivante :

pour les communes enquêtées en N-1 : poids = (population N-1 / population N-6)^(1/5)

pour les communes enquêtées en N-2 : poids = (population N-2 / population N-7)^(2/5)

Le même coefficient est appliqué pour les logements, les résidences principales et la population.

3 Les poids à l'exploitation principale – Population hors ménage

La **population hors ménage** comprend :

les personnes résidant dans une communauté, hors logements de fonction (les occupants des logements de fonction font partie de la population des ménages) ;

les personnes vivant dans des habitations mobiles (y compris les bateliers) et les personnes sans abri.

La population hors ménage ne représente qu'une part très faible de l'ensemble de la population : environ 2 % pour les communautés et de l'ordre de 0,2 % pour les personnes vivant dans une habitation mobile ou sans abri.

Ces populations sont traitées de la même façon pour les communes de moins de 10 000 habitants et pour les communes de 10 000 habitants ou plus, ainsi que pour la métropole et les DOM.

Elles font l'objet d'un recensement exhaustif quelle que soit la taille de la commune ; elles ont donc un poids initial égal à 1.

3.1 Les communautés

Une communauté est un ensemble de locaux d'habitation relevant d'une **même autorité gestionnaire** et dont les habitants partagent à titre habituel un **mode de vie commun**. La population de la communauté comprend les personnes qui résident dans la communauté, à l'exception de celles qui résident dans des logements de fonction. Par ailleurs, les mineurs recensés dans un établissement d'enseignement étant par convention rattachés à leur résidence familiale, ils sont affectés d'un poids nul dans la communauté (leur bulletin individuel de communauté est conservé pour le calcul de la population légale dite « comptée à part » de la commune de l'établissement). Si leur résidence familiale a été enquêtée, leurs parents ont rempli un bulletin individuel pour eux et ils ont alors le poids affecté à leur résidence familiale.

Les catégories de communautés sont les suivantes :

- les services de moyen ou de long séjour des établissements publics ou privés de santé, les établissements sociaux de moyen et long séjour, les maisons de retraite, les foyers et résidences sociales ou assimilés ;
- les communautés religieuses ;
- les casernes, quartiers, bases ou camps militaires ou assimilés ;
- les établissements hébergeant des élèves ou des étudiants, y compris les établissements militaires d'enseignement ;
- les établissements pénitentiaires ;
- les établissements sociaux de court séjour ;
- les autres communautés.

Le recensement complet des communautés se réalise sur cinq ans :

- les communautés dans les communes de moins de 10 000 habitants sont enquêtées une fois tous les cinq ans, la même année que la population des ménages de la commune ;
- dans les communes de 10 000 habitants ou plus, elles sont, sauf exceptions, enquêtées tous les cinq ans une année donnée, qui n'est pas la même pour toutes les communes, de manière à ce que la charge globale de la collecte soit étalée sur une période de cinq ans.

Comme pour la population des ménages, il est nécessaire de ramener des populations enquêtées sur un cycle de cinq ans à une même date de référence, le 1^{er} janvier de l'année médiane du cycle. Cette opération est réalisée communauté par communauté, et aboutit à modifier les poids des individus résidant dans la communauté.

Pour les rendre plus lisibles, les calculs¹⁴ sont explicités en prenant l'exemple du millésime 2014, année médiane du cycle d'enquêtes annuelles 2012-2016.

Dans les formules générales, l'année N figure l'année médiane.

3.1.1 Pour les communautés enquêtées en 2012 (N-2) ou 2013 (N-1)

Pour ces communautés, l'actualisation consiste à prendre en compte les éventuelles suppressions de communautés intervenues entre la date de la dernière enquête de recensement¹⁵ et le 1^{er} janvier de l'année de référence N (01/01/2014 en l'occurrence dans cet exemple). Cette prise en compte s'appuie sur le répertoire des communautés, régulièrement mis à jour à l'aide de sources administratives et d'informations transmises par les communes.

- En cas de suppression avant le 1^{er} janvier de l'année N (2014) d'une communauté enquêtée en 2012 ou 2013, la collecte des individus recensés dans la communauté en question est annulée.
- Pour les communautés pérennes sur la période, le poids des individus concernés est maintenu égal à 1 car il n'existe pas d'information suffisamment fiable pour actualiser leur effectif entre 2012 (ou 2013) et le 1^{er} janvier 2014.

3.1.2 Pour les communautés enquêtées en 2014 (N)

Pour ces communautés, le résultat de la collecte est conservé tel quel et le poids des individus concernés est maintenu à 1.

3.1.3 Pour les communautés enquêtées en 2015 (N+1) ou 2016 (N+2)

Pour ces communautés, il est procédé à une interpolation analogue à celle pratiquée pour la population des ménages des communes de moins de 10 000 habitants, entre la population estimée l'année précédente pour 2013 (N-1) et la dernière collecte connue de 2015 (N+1) ou 2016 (N+2). Tous les occupants de la communauté enquêtée en 2015 (N+1) ou 2016 (N+2) sont conservés et ils sont pondérés par un coefficient ad hoc (voir encadré ci-dessous). Ce coefficient permet d'obtenir par agrégation la population estimée dans la communauté au 1^{er} janvier 2014 (N).

¹⁴ Il existe des traitements dérogatoires aux principes de calculs qui sont présentés ici. Ils sont en nombre très limité et concernent des décisions gouvernementales de création de prisons et de transfert de régiments dans certaines casernes, ayant un impact très fort sur les communes concernées.

¹⁵ Jusqu'en 2009, les communautés étaient recensées plus tardivement que les logements ; la date de référence de la collecte était le 1^{er} mars de l'année. Depuis 2010, les communautés sont recensées en même temps que les ménages, avec pour date de référence le troisième jeudi de janvier.

Pour le calcul de la population de la communauté au 1^{er} janvier de l'année N, cela équivaut à modifier les poids initiaux (égaux à 1) des individus de la communauté de la manière suivante :

Pour les communautés enquêtées en N+1 :

poids = $1/2 + (1/2) \times (\text{population estimée en N-1} / \text{population collectée en N+1})$

Pour les communautés enquêtées en N+2 :

poids = $1/3 + (2/3) \times (\text{population estimée en N-1} / \text{population collectée en N+2})$

Si la communauté n'existait pas au 1^{er} janvier N-1, la population estimée en N-1 est alors égale à zéro, et les poids respectivement de 1/2 et de 1/3.

3.1.4 Pour les communautés existantes au 1^{er} janvier 2014 (N) mais jamais encore enquêtées

Il s'agit de prendre en compte les créations de communautés existantes au 1^{er} janvier 2014 (N) mais n'ayant encore jamais été enquêtées. Cette prise en compte s'appuie sur le répertoire des communautés, régulièrement mis à jour à l'aide de sources administratives et d'informations transmises par les communes.

Dans ce cas, une population égale à la capacité d'accueil fournie par le répertoire des communautés est affectée à la communauté créée. Pour « alimenter » cette nouvelle communauté de l'effectif retenu, il est procédé à la recopie (« clonage ») des bulletins individuels d'autant d'individus que nécessaire recensés dans une communauté de même catégorie, aussi proche que possible géographiquement et recensée l'année du recensement des communautés de la commune concernée.

Le poids de chacun de ces individus « clonés » est fixé à 1.

3.1.5 Pour les communautés existantes au 1^{er} janvier 2014 (N), non enquêtées au cours du cycle de 5 ans de 2012 (N-2) à 2016 (N+2) mais enquêtées au cours du cycle précédent

Il s'agit du cas des communautés qui existaient lors de l'enquête de recensement en 2010 (N-4) ou 2011 (N-3) et pour lesquelles une population a été estimée pour l'année 2013 (N-1) mais qui ont fermé avant l'enquête de recensement suivante en 2015 (N+1) ou 2016 (N+2).

Malgré l'absence d'individus enquêtés en 2015 (N+1) ou 2016 (N+2), cette communauté doit participer aux populations légales de l'année 2014 (N).

L'effectif de la communauté en 2014 (N) est estimé en interpolant entre l'effectif de la collecte 2010 (N-4) ou 2011 (N-3) (qui correspond à celui estimé pour 2013 (N-1)) et celui de la collecte 2015 (N+1) ou 2016 (N+2), considéré comme nul.

Pour « alimenter » cette communauté de l'effectif obtenu, il est procédé à la recopie (« clonage ») de bulletins individuels recensés dans une autre communauté de même catégorie, aussi proche que possible géographiquement et qui, elle, a été recensée en 2015 (N+1) ou en 2016 (N+2). Sont ainsi « clonés » autant d'individus que la dernière collecte de 2010 (N-4) ou 2011 (N-3) avait constaté dans la communauté en question.

Afin d'interpoler le nombre de résidents de la communauté entre la population correspondant au cycle précédent et la collecte nulle du cycle actuel, les poids des individus clonés sont modifiés selon les formules de l'encadré ci-dessous.

Pour le calcul de la population de la communauté au 1^{er} janvier de l'année N, cela équivaut à créer des poids pour les individus de la communauté de la manière suivante :

Pour les communautés qui auraient dû être enquêtées en N+1 si elles n'avaient pas fermé avant :

$$\text{poids} = (1/3)$$

Pour les communautés qui auraient dû être enquêtées en N+2 si elles n'avaient pas fermé avant :

$$\text{poids} = (2/3)$$

3.2 Les autres populations hors ménages

La population des habitations mobiles terrestres, les bateliers et les personnes sans abri sont recensés exhaustivement : leur poids initial est donc de 1.

Ces enquêtes exhaustives ont lieu pour tout le territoire au même moment et sont répétées à intervalles quinquennaux (2006 / 2011 / 2016, etc.).

En l'absence de données permettant leur actualisation, les effectifs de ces populations sont maintenus constants pendant quatre ans, jusqu'à l'enquête de recensement suivante. Ils ne font l'objet ni d'extrapolation ni d'interpolation.

Le poids final des individus concernés est donc toujours égal à 1.

4 Les poids à l'exploitation complémentaire

4.1 Le tirage de l'échantillon de l'exploitation complémentaire¹⁶

Alors que l'exploitation principale porte sur l'ensemble des bulletins collectés, l'exploitation complémentaire porte dans certains cas sur un échantillon de cet ensemble.

Il est constitué de la façon suivante :

	Communes de moins de 10 000 habitants	Communes de 10 000 habitants ou plus
Ménages	20 % (*)	100 % des ménages enquêtés (soit environ 40 % du total des ménages)
Individus des communautés	20 % (*)	
Habitations mobiles terrestres et personnes sans abri	20 % (*)	100 %
Habitations mobiles fluviales (bateliers)	100 %	

(*) : jusqu'à l'enquête annuelle de recensement (EAR) de 2013, ce taux de sondage s'élevait à 25 %

¹⁶ Pour en savoir plus sur l'exploitation complémentaire, se reporter à la fiche thématique « Les exploitations principale et complémentaire »

4.2 Le calcul des poids

Le principe est de réutiliser les poids de l'exploitation principale et de les modifier par un coefficient correcteur directement lié au sondage supplémentaire réalisé pour l'exploitation complémentaire. Le coefficient correcteur est égal à l'inverse du taux de sondage de l'échantillon du complémentaire parmi les bulletins collectés. Les valeurs de ce **coefficient correcteur** sont donc :

- pour les logements et individus hors du champ de l'exploitation complémentaire : **0** ;
- pour les logements ou individus dont les caractéristiques impliquent qu'ils font automatiquement partie de l'échantillon de l'exploitation complémentaire : **1** ;
- pour les logements ou individus issus d'un tirage au cinquième (ou au quart avant l'EAR de 2014) : **5** (respectivement 4).

Pour une observation donnée, le poids à l'exploitation complémentaire sera donc :

$$\begin{array}{c} \text{poids à l'exploitation complémentaire} \\ = \\ \text{poids à l'exploitation principale} \\ \times \\ \text{coefficient correcteur} \end{array}$$

4.3 Le calcul des poids pour les recensements millésimés 2021 à 2023

En 2018, suite à une demande exprimée par le Conseil national de l'information statistique (Cnis), la feuille de logement du recensement de la population a été rénovée pour mieux décrire les situations de multi-résidence qui se développent¹⁷. Les liens familiaux qui unissent les personnes habitant un même logement et les lieux d'habitations des enfants de parents séparés sont ainsi mieux connus.

Concrètement, la feuille de logement rénovée permet de décrire finement les liens de parenté entre les habitants du logement deux à deux sans passer par la saisie d'un libellé. L'analyse ménages-familles a pu ainsi être généralisée à l'ensemble des logements (présents dans l'exploitation principale) car elle ne repose plus sur des traitements manuels.

Les résultats de l'analyse ménages-familles, issus de cette nouvelle feuille de logement, auraient dû être diffusés à l'exploitation principale du recensement de la population millésimé 2020, cumul des enquêtes annuelles de recensement de 2018 à 2022. En raison du report de l'enquête annuelle de recensement 2021, du fait de la crise sanitaire liée à la Covid-19, les résultats sont diffusés à l'exploitation complémentaire à partir du millésime 2021 du recensement de la population, et seront diffusés à l'exploitation principale à partir du millésime 2024.

En effet, **pour les communes de 10 000 habitants ou plus**, l'enquête de recensement 2021 n'a pas été simplement décalée d'un an : elle a été remplacée par une pseudo-EAR constituée à partir des données individuelles de l'enquête annuelle de recensement 2016 (avec des informations différentes de celles du questionnaire rénové). Tant que cette pseudo-EAR est utilisée dans la construction des résultats du recensement (jusqu'au millésime 2023), la diffusion de l'analyse ménages-familles à l'exploitation

¹⁷ Pour plus d'informations sur la nouvelle analyse ménages-familles, se reporter à la fiche thématique « *Ménages et familles* ».

principale n'est pas possible. Par conséquent, **la nouvelle analyse ménages-familles sera associée à l'exploitation complémentaire du recensement de la population des millésimes 2021 à 2023 avec des pondérations spécifiques** (cf paragraphe 4.3.1 et paragraphe 4.3.2).

4.3.1 Pour les communes de 10 000 habitants ou plus.

Dans les communes **de 10 000 habitants ou plus**, l'Insee a décidé d'utiliser, pour les millésimes 2021 à 2023, 4 enquêtes annuelles de recensement au lieu de 5, moyennant l'ajout d'une pondération spécifique (POIDS_COMPL_AMF) à l'exploitation complémentaire.

Ce poids AMF est le plus petit dénominateur commun sur lequel on peut exploiter conjointement les variables de l'AMF et les autres variables du recensement de la population.

$$\text{POIDS_COMPL_AMF} = \text{Poids de l'exploitation complémentaire (cf §4.2)} \times \text{Coefficient appliqué pour l'exploitation de l'analyse ménages-familles}$$

Toutes les observations d'une année de collecte donnée ont le même coefficient, compris entre 0,5 et 2, sauf celles de la pseudo-enquête de recensement 2021 (avec des informations correspondant à l'ancien questionnaire) dont le coefficient vaut 0.

Tableau : Coefficients appliqués pour l'exploitation de l'analyse ménages-familles par millésime du recensement (RP) et année de l'enquête annuelle de recensement (EAR) (pour les communes de 10 000 habitants ou plus)

	EAR 2019	EAR 2020	EAR 2021	EAR 2022	EAR 2023	EAR 2024	EAR 2025
RP 2021	1,25	1,25	0	1,25	1,25		
RP 2022		1,57	0	1,29	1,14	1,00	
RP 2023			0	2,00	1,50	1,00	0,50

4.3.2 Pour les communes de moins de 10 000 habitants

À partir du recensement de la population millésimé 2021, toutes les communes de moins de 10 000 habitants auront été recensées avec la nouvelle feuille de logement, ce qui signifie que les informations correspondent aux nouvelles variables de l'analyse ménages-familles.

La pondération spécifique (POIDS_COMPL_AMF) sera égale au poids de l'exploitation complémentaire, sans besoin d'introduire de coefficient supplémentaire .

Pour une observation donnée, le poids à l'**exploitation complémentaire** pour les **communes de moins de 10 000 habitants** sera :

$$\text{POIDS_COMPL_AMF} = \text{Poids de l'exploitation complémentaire (cf §4.2)}$$