

## ADAPTATION DES MÉTHODES DE CALCUL DES POPULATIONS DE RÉFÉRENCE ET DES RÉSULTATS STATISTIQUES DU RECENSEMENT DE LA POPULATION SUITE AU REPORT DE L'EAR 2021

Du fait de la situation sanitaire, la collecte de l'enquête annuelle de recensement (EAR) de 2021 a été reportée à 2022, à l'exception du territoire de Mayotte et du recensement des bateliers. Ce report est acté par **l'article 17 de la loi n°2021-689 relative à la gestion de la sortie de la crise sanitaire**. Consultés, l'Association des maires de France (AMF) et les représentants des communes au sein de la Commission nationale d'évaluation du recensement de la population (Cnerp) ont soutenu ce report.

Les conditions n'étaient en effet pas réunies pour réussir une collecte de qualité en janvier 2021. La collecte sur le terrain de l'enquête de recensement entraîne de nombreux déplacements et contacts avec les habitants ; même si ceux-ci sont courts et limités, ils étaient difficilement compatibles avec la situation sanitaire de janvier 2021. Une moindre adhésion de la population aurait entraîné de nombreux refus de répondre, engendrant une dégradation des résultats produits. Par ailleurs, de nombreuses communes ont rencontré des difficultés importantes dans la préparation de l'enquête du fait des mesures de confinement à l'automne, notamment pour trouver des candidats aux fonctions d'agents recenseurs, qui faisaient peser un risque sur le bon déroulement de la collecte sur le terrain.

Conformément à la loi et malgré ce report, l'Insee s'est engagé à produire et publier les populations de référence selon le calendrier habituel. En particulier, les populations relatives au 1<sup>er</sup> janvier 2019 seront publiées avant la fin de l'année 2021. Les résultats statistiques du recensement seront également publiés annuellement, selon le calendrier habituel<sup>1</sup>.

**Le report de l'EAR 2021 conduit donc à une adaptation des méthodes de calcul des populations de référence et des résultats statistiques.**

### Sommaire

1 Adaptation de la méthode de calcul des populations de référence.....	2
1.1 Dans les communes de moins de 10 000 habitants.....	2
1.2 Dans les communes de 10 000 habitants ou plus.....	5
1.3 Performances des méthodes adaptées de calcul des populations de référence.....	7
2 Estimation des résultats statistiques suite au report de l'EAR 2021.....	7

<sup>1</sup> Notamment fin juin : diffusion des résultats statistiques sur insee.fr

## 1 Adaptation de la méthode de calcul des populations de référence

L'adaptation des méthodes de calcul ne concerne que la population en logement ordinaire. En effet, les populations correspondant aux personnes vivant en communauté et celles des habitations mobiles et sans abri (HMSA) ont des modes de calcul différents, et plus simples. On conserve un an de plus la dernière population collectée.

### 1.1 Dans les communes de moins de 10 000 habitants

#### Rappel de la méthode habituelle

Chaque commune de moins de 10 000 habitants est recensée exhaustivement tous les 5 ans (toutes les communes de moins de 10 000 habitants sont ainsi réparties dans 5 groupes de rotation). Pour l'estimation de la population de l'année N, diffusée en fin d'année N+2, le mode de calcul dépend de l'année de la dernière collecte :

- la dernière collecte a eu lieu avant la date de référence : N-1 ou N-2. La population est alors estimée par extrapolation à partir de la dernière population issue d'une collecte sur le terrain.

**En métropole et à La Réunion**, l'extrapolation consiste à faire évoluer la population comme le nombre de logements dans les sources fiscales, en tenant compte d'un coefficient de décohabitation (variation du nombre moyen de personnes par résidence principale entre les deux dernières collectes).

**Aux Antilles-Guyane et dans les collectivités de Saint-Pierre-et-Miquelon et Saint-Barthélemy**, l'extrapolation consiste à faire évoluer la population au rythme observé entre les deux dernières enquêtes de recensement (prolongement de la tendance).

- la dernière collecte a eu lieu l'année de référence : N. Dans ce cas, les données de la collecte sont directement utilisées.
- la dernière collecte a eu lieu après l'année de référence : N+1 ou N+2. La population est alors estimée par interpolation linéaire entre la dernière population publiée (N-1) et la population collectée.

L'illustration 1 représente schématiquement la méthode habituelle de calcul des populations de référence des communes de moins de 10 000 habitants pour le recensement de la population (RP) de 2019 (en l'absence de report de l'EAR 2021).

*Illustration 1 : méthode habituelle de calcul des populations 2019 pour les communes de moins de 10 000 habitants (en l'absence de report de l'EAR 2021)*

	2016	2017	2018	2019	2020	2021
Communes du GR1				Collecte		
Communes du GR2					← Collecte	
Communes du GR3	Collecte				← Interpolation	Collecte
Communes du GR4		Collecte				
Communes du GR5			Collecte			

GR : groupe de rotation

### Adaptation de la méthode :

Le report de l'EAR 2021 induit un report en cascade d'un an de la collecte pour chaque commune de moins de 10 000 habitants (cf. décret n°2021-681 du 28 mai 2021 modifiant l'annexe au décret n°2003-561 du 23 juin 2003 portant répartition des communes pour les besoins du recensement de la population) :

- les communes devant être recensées en 2021 le seront en 2022 ;
- les communes devant être recensées en 2022 le seront en 2023 ;
- les communes devant être recensées en 2023 le seront en 2024 ;
- les communes devant être recensées en 2024 le seront en 2025 ;
- les communes devant être recensées en 2025 le seront en 2026 ;

Dans la pratique, les communes vont changer de groupe de rotation. L'écart inter-collecte va donc passer temporairement de 5 à 6 ans pour l'ensemble des communes de moins de 10 000 habitants.

Durant cette période temporaire, chaque année, la méthode de calcul des populations relative à l'année N sera adaptée uniquement pour un cinquième des communes : celles qui auraient dû réaliser une collecte l'année du calcul (N+2), et qui ne la réaliseront que l'année suivante (N+3). En effet, les autres communes se trouvent dans une situation classique, avec une dernière collecte réalisée l'une des années habituelles du cycle de 5 ans, à savoir en N-2, N-1, N, N+1 ou N+2. Pour ces communes, la méthode habituelle peut donc être appliquée.

Pour le cinquième de communes qui auraient dû faire une collecte en N+2 mais dont la collecte est reportée en N+3, la dernière collecte remonte à l'année N-3. La population de l'année N de ces communes sera calculée par extrapolation à partir de la dernière collecte disponible. **La méthode adaptée consiste donc à appliquer une année de plus la méthode par extrapolation utilisée habituellement pour les deux années de référence qui suivent la collecte.** L'année suivante, en N+3, ces communes réalisent une nouvelle enquête annuelle de recensement. Le calcul de leur population en N+1 mobilisera donc cette nouvelle information, par une interpolation classique.

**Le régime adapté ne concerne donc chaque année que 20 % des communes, et ne dure qu'un an pour chaque commune.** Au final, au cours du cycle temporaire de 6 ans, les populations sont estimées trois fois par extrapolation (une fois de plus qu'habituellement), deux fois par interpolation et une fois directement à partir de la population collectée. Toutes les communes de moins de 10 000 habitants connaîtront exactement une année de régime adapté pour les populations de la période 2019-2023.

L'illustration 2 représente schématiquement la méthode adaptée pour le calcul des populations de référence 2019. Les communes qui devaient faire une collecte en 2021 y sont en troisième extrapolation.

Illustration 2 : méthode adaptée de calcul des populations 2019 pour les communes de moins de 10 000 habitants

	2016	2017	2018	2019	2020	2021
Communes du GR1				Collecte		
Communes du GR2					Collecte	
<b>Communes du GR3</b>	Collecte					Pas de collecte
Communes du GR4		Collecte				
Communes du GR5			Collecte			

*Note: The table includes blue arrows indicating extrapolation from 2016 to 2019 for GR3 and from 2017 to 2019 for GR4, and from 2018 to 2019 for GR5. A green arrow points from 2020 to 2019 for GR2. A large orange oval highlights the GR3 row and the 2019 column.*

GR : groupe de rotation

L'illustration 3 représente schématiquement la méthode adaptée pour le calcul des populations de référence 2020. Les communes qui devaient faire une collecte en 2021 (ancien groupe de rotation 3) réaliseront une collecte en 2022 (report de celle de 2021) et seront donc en interpolation pour le calcul des populations 2020. En revanche, les communes qui auraient dû réaliser une collecte en 2022 (ancien groupe de rotation 4) auront une population 2020 calculée par troisième extrapolation.

Illustration 3 : méthode adaptée de calcul des populations 2020 pour les communes de moins de 10 000 habitants

	2016	2017	2018	2019	2020	2021	2022
Communes de l'ancien GR1				Collecte			
Communes de l'ancien GR2					Collecte		
<b>Communes de l'ancien GR3</b>	Collecte					Pas de collecte	Collecte
<b>Communes de l'ancien GR4</b>	Collecte						Pas de collecte
Communes de l'ancien GR5			Collecte				

*Note: The table includes blue arrows indicating extrapolation from 2016 to 2020 for GR3 and from 2017 to 2020 for GR4, and from 2018 to 2020 for GR5. A green arrow points from 2022 to 2020 for GR3, labeled 'Interpolation'. A large orange oval highlights the GR3 and GR4 rows and the 2020 column.*

GR : groupe de rotation

## 1.2 Dans les communes de 10 000 habitants ou plus

### Rappel de la méthode habituelle

En métropole, la population des communes de 10 000 habitants ou plus relative au 1<sup>er</sup> janvier de l'année N est obtenue par la multiplication du nombre de logements issus du répertoire d'immeubles localisés (RIL) par le nombre moyen de personnes par logement estimé à partir des EAR N-2 à N+2.

$$Pop_N = \text{nombre de logements du RIL} * \overline{NMPL}_{EAR}$$

où  $\overline{NMPL}_{EAR}$  est le nombre moyen de personnes par logement estimé sur l'ensemble des EAR N-2 à N+2.

Le RIL est un répertoire exhaustif des adresses d'habitation et de leurs caractéristiques en termes de logements qui est mis à jour chaque année via un partenariat entre l'Insee et les communes. Il permet notamment de déterminer le nombre de logements de chaque commune de 10 000 habitants ou plus. Malgré la crise sanitaire, ce répertoire a été mis à jour en 2020 et 2021, selon les mêmes exigences de qualité qu'habituellement : le nombre de logements est donc toujours bien disponible chaque année. En revanche, le report de l'EAR 2021 ne permet pas de calcul direct du nombre moyen de personnes par logement.

Dans les DOM, les communes de 10 000 habitants ou plus sont découpées en îlots pour les besoins du recensement. Ces îlots sont répartis en 5 groupes de rotation. Chaque année, 40 % des logements d'un groupe de rotation sont recensés. Le RIL est mis à jour chaque année sur un cinquième du territoire. En l'absence de RIL mis à jour simultanément sur l'ensemble des adresses d'habitation, la population des communes de 10 000 habitants ou plus relative au 1<sup>er</sup> janvier de l'année N est obtenue en sommant la population des adresses des cinq groupes de rotation, la population d'un groupe de rotation donné étant estimée à partir d'une EAR.

$$Pop_N = \sum_{i=N-2}^{N+2} \overline{PopEAR}_i$$

où  $\overline{PopEAR}_i$  est la population estimée par l'EAR du groupe de rotation recensé en i.

Dans le cas des DOM, l'information manquante suite au report de l'EAR 2021 est donc la population du groupe de rotation qui aurait dû être recensée en 2021. Le nombre de logements de ce groupe de rotation est toutefois disponible, par l'intermédiaire des enquêtes cartographiques et des travaux de la commune sur les RIL. Comme pour la métropole, c'est le nombre moyen de personnes par logement qui aurait été collecté en 2021 qui est manquant.

### Méthode adaptée :

La méthode adaptée consiste à créer une pseudo-EAR 2021 pour chaque commune de 10 000 habitants ou plus qui remplacera l'EAR 2021, absente, dans les estimations de population et les résultats statistiques pour les cinq années à venir. Pour les populations de référence, cette pseudo-EAR vise à refléter au mieux ce qu'aurait donné une collecte réalisée en 2021, dans des conditions normales, en termes de nombre de logements et de nombre moyen de personnes par logement.

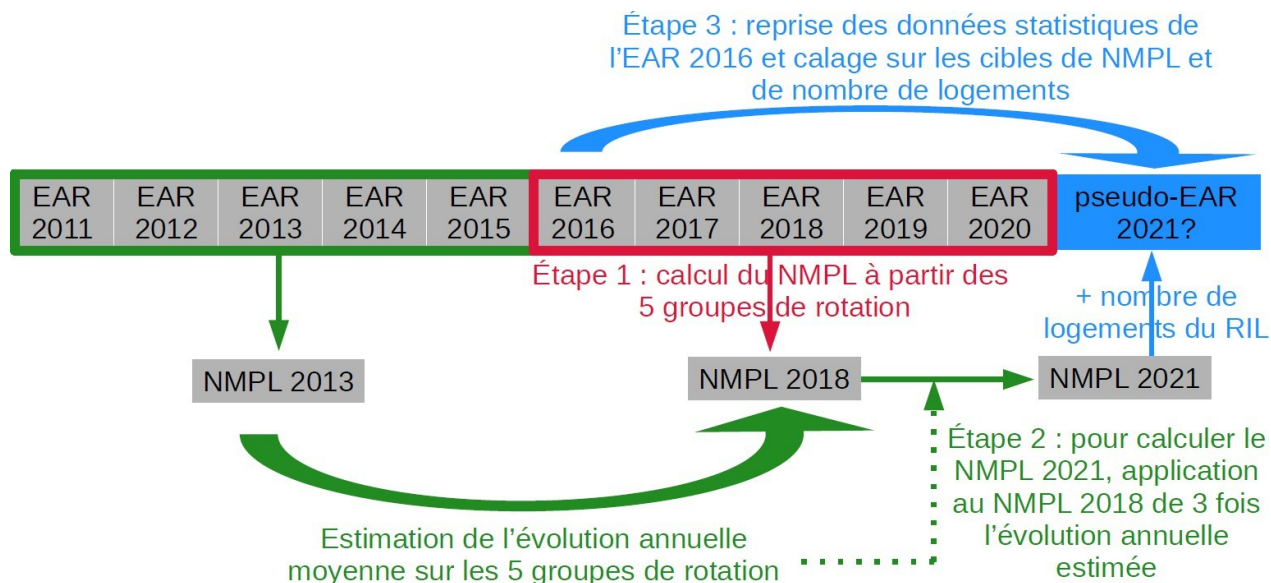
Avant de construire cette pseudo-EAR, une première étape consiste donc à déterminer un nombre cible de logements, et un nombre moyen de personnes par logement cible au 1<sup>er</sup> janvier 2021 :

- Pour le nombre de logements : on dispose du RIL qui fournit le nombre de logements habitables du groupe de rotation qui devait faire l'objet de l'enquête.
- Pour le nombre moyen de personnes par logement : on calcule dans un premier temps le nombre moyen de personnes par logement (NMPL) sur les cinq dernières EAR (2016 à 2020), ce qui fournit une estimation du NMPL au 1<sup>er</sup> janvier 2018. Pour le ramener au 1<sup>er</sup> janvier 2021, on lui applique sur 3 ans un coefficient annuel d'évolution, correspondant à l'évolution annuelle moyenne observée pour les 5 groupes de rotation.

Concrètement, la formule de calcul est la suivante :

$$\widehat{NMPL}_{\text{EAR } 2021} = \left( \prod_{a=2016}^{2020} NMPL_{\text{EAR } a} \right)^{\frac{1}{5}} \times \left( \frac{\left( \prod_{a=2016}^{2020} NMPL_{\text{EAR } a} \right)^{\frac{1}{5}}}{\left( \prod_{a=2011}^{2015} NMPL_{\text{EAR } a} \right)^{\frac{1}{5}}} \right)^{\frac{3}{5}}$$

Illustration 4 : détermination de la cible du nombre moyen de personnes par logement pour la pseudo-EAR 2021



Une fois ces cibles déterminées, les données individuelles de la pseudo-EAR 2021 sont produites à partir des données individuelles de l'EAR 2016<sup>2</sup> et d'un calage sur les cibles de nombre de logements et de nombre d'individus. Ces données individuelles de la pseudo-EAR 2021 seront intégrées aux bases de données du recensement de la population (RP) en lieu et place des données de l'EAR 2021 manquantes.

Cette méthode de création d'une pseudo-EAR sera appliquée de manière similaire pour la métropole et les DOM. La pseudo-EAR 2021 sera créée une seule fois, et servira pour les estimations de population et de résultats statistiques pour les cinq années à venir.

Pour la métropole, pour les cinq cycles RP à venir (de 2017-2021 à 2021-2025), le nombre moyen de personnes par logement sera estimé à partir des données des 4 EAR du cycle disponibles et de la pseudo-EAR 2021. Le nombre de logements de l'année médiane sera toujours calculé à partir du RIL.

Pour les DOM, pour les cinq cycles RP à venir (de 2017-2021 à 2021-2025), la population sera calculée en sommant la population de chaque groupe de rotation estimée à partir des 4 EAR du cycle disponibles et de la pseudo-EAR 2021.

<sup>2</sup> Si les groupes de rotation ne présentent pas de déséquilibre en termes de nombre moyen de personnes par logement, cela n'est pas garanti pour toutes les variables statistiques, si bien qu'il est préférable de répartir des données portant sur le même groupe de rotation.

### 1.3 Performances des méthodes adaptées de calcul des populations de référence

Pour évaluer la performance des méthodes adaptées de calcul des populations, l'Insee a simulé l'absence d'une EAR passée, et a comparé les populations estimées par les méthodes adaptées et celles obtenues par les méthodes habituelles (populations publiées). Les principaux résultats sont les suivants :

- les écarts sont très faibles aux niveaux national, régional et départemental entre les deux estimations : l'écart au niveau national est de 15 000 personnes pour l'ensemble des communes de 10 000 habitants ou plus et de 12 000 personnes pour l'ensemble des communes de moins de 10 000 habitants.
- au niveau communal, pour les communes de moins de 10 000 habitants :
  - les méthodes étant identiques pour 80 % des communes, aucun écart n'est observé pour ces communes
  - pour les 20 % de communes restantes, l'écart entre la population estimée par chacune des deux méthodes est inférieur à 2 % en valeur absolue pour 58 % des communes. Les écarts relatifs les plus grands concernent majoritairement les plus petites communes.
- au niveau communal, pour les communes de 10 000 habitants ou plus : l'écart entre les deux estimations est inférieur à 0,6 % en valeur absolue pour la moitié des communes. L'écart est supérieur à 1,2 % en valeur absolue pour 20 % des communes. L'écart relatif est d'autant plus grand que la commune est petite.

Au final, la population estimée par la méthode adaptée est incluse dans l'intervalle de confiance à 95 %<sup>3</sup> de la population publiée pour 92 % des communes de 10 000 habitants ou plus.

**En résumé, les méthodes adaptées sont de très bonne qualité pour les estimations agrégées et de bonne qualité pour les estimations au niveau des communes. Les populations ainsi estimées sont probablement de bien meilleure qualité que celles qui auraient été produites si une collecte avait eu lieu en 2021 compte-tenu des difficultés qui auraient été rencontrées en cette période de crise sanitaire avec un risque élevé d'augmentation nette de la non-réponse. Ces méthodes permettent donc de faire face à une situation exceptionnelle.**

## 2 Estimation des résultats statistiques suite au report de l'EAR 2021

En régime courant, dans les communes de moins de 10 000 habitants, la structure socio-démographique n'est modifiée qu'après une nouvelle enquête de recensement. Ainsi, cette structure est maintenue constante pendant 5 ans. L'écart inter-collecte passant temporairement à 6 ans pour toutes les communes de moins de 10 000 habitants, la structure va être maintenue constante pendant 6 ans, soit une année de plus qu'habituellement. Pour les comparaisons temporelles, il est ainsi recommandé de comparer des résultats du recensement sur des millésimes distants d'au moins 6 ans (contre au moins 5 ans précédemment).

Dans les communes de 10 000 habitants ou plus, l'absence de l'EAR 2021 est palliée par la création d'une pseudo-EAR 2021 qui mobilise les données individuelles de l'EAR 2016. Seules les variables de nombre de logements et de nombre moyen de personnes par logement sont actualisées avec pour date de référence 2021 (cf. partie 1-B). Ainsi, les structures socio-démographiques évolueront peu entre les RP 2018 et 2019 pour la quasi-totalité des variables. Seules les variables corrélées aux données actualisées (nombre de logements et nombre moyen de personnes par logement) pourront évoluer notablement : c'est le cas par exemple de la part de logements vacants ou de la taille moyenne des logements.

La non-évolution des données socio-démographiques pendant un an est acceptable dans la mesure où il est préconisé de comparer deux millésimes du recensement sur des périodes éloignées (au moins 5 ans actuellement et au moins 6 ans suite au report de l'EAR 2021), où les évolutions socio-démographiques sont par nature lentes, et où le recensement n'est pas un outil conçu pour appréhender des phénomènes conjoncturels. Ainsi, les messages que l'on tire des évolutions temporelles ne sont pas affectés par le report de l'EAR 2021.

<sup>3</sup> Intervalle de confiance consécutif à l'aléa de sondage dans les communes de 10 000 habitants ou plus.

Le tableau 1 en donne une illustration sur le champ des communes de 10 000 habitants ou plus : il compare les estimations produites par la méthode habituelle et par la méthode adaptée en simulant l'absence de l'EAR 2017 pour le calcul des résultats statistiques du recensement 2015.

Tableau 1 : Estimation de 4 indicateurs selon la méthode : méthode adaptée ou méthode habituelle – simulation d'absence de l'EAR 2017 pour le RP 2015

	RP 2015		Variation RP 2010-2015	
	Méthode adaptée	Méthode habituelle	Méthode adaptée	Méthode habituelle
Part de femmes	52,6%	52,6%	-0,1 pt	-0,1 pt
Part des 60 ans et plus	22,0%	22,2%	+1,2 pt	+1,5 pt
Part d'étrangers	9,6%	9,9%	+0,8 pt	+1,1 pt
Part de logements vacants	8,4%	8,3%	+0,8 pt	+0,8 pt

Lecture :

- la part de femmes est stable dans le temps. La non-évolution pendant un an de la structure n'a pas d'impact sur le niveau estimé, ni sur l'évolution estimée.
- la part des 60 ans et plus et la part d'étrangers croissent sur le passé récent. La non-évolution pendant un an de la structure socio-démographique engendre une légère sous-estimation du niveau 2015 avec la méthode adaptée. Cependant, le diagnostic de la hausse de ces deux parts reste vrai quelle que soit la méthode.
- la part de logements vacants est également orientée à la hausse sur le passé récent. Cette variable étant bien corrélée à la part de résidences principales et donc au nombre moyen de personnes par logement qui est actualisé lors de la création de la pseudo-EAR 2021, l'estimation par la méthode adaptée est proche de celle de la méthode habituelle.

Champ : communes de 10 000 habitants ou plus

Le report de l'EAR 2021 implique donc que les structures socio-démographiques des communes de moins de 10 000 habitants seront temporairement figées pendant 6 ans au lieu de 5 ans aujourd'hui. **Il est donc recommandé à partir de la diffusion des populations du recensement 2019 à la fin de l'année 2021 de faire des comparaisons de résultats entre deux millésimes distants d'au moins 6 ans, et ce jusqu'au recensement 2023 inclus.** Ce changement de recommandation permet de garder la rigueur statistique (si on reste sur une période de 5 ans, 20 % des communes de moins de 10 000 habitants n'auraient aucune variation de leurs structures socio-démographiques).