

Recensement de la population

La précision du chiffre de population dans les grandes communes de métropole

La qualité des résultats d'un recensement dépend de multiples facteurs, et en premier lieu de la qualité de la collecte. Elle dépend aussi de l'exhaustivité des fichiers et répertoires utilisés pour préparer et exploiter les enquêtes, ainsi que de la qualité des différents traitements mis en œuvre : saisie, contrôles, redressement des anomalies et codification des variables.

Dans les standards internationaux, on considère qu'un recensement est de très bonne qualité dès lors qu'il atteint une précision de l'ordre de 1 % pour le chiffre de l'ensemble de la population d'un pays. La méthode française de recensement est de nature à assurer ce niveau de qualité, grâce en particulier à la professionnalisation des acteurs, à la répartition de la charge de collecte dans le temps et à la mise à jour annuelle d'un répertoire d'immeubles dans les grandes communes.

La méthode du recensement utilise les techniques de sondage dans les communes de 10 000 habitants et plus, dites « grandes communes », qui regroupent la moitié de la population du pays. Cette fiche décrit l'imprécision de la variable « population » liée à cette particularité. Elle illustre le fait que l'ordre de grandeur de cette imprécision est négligeable par rapport aux autres facteurs influençant la qualité des résultats du recensement.

Cette fiche complète celle portant sur la précision des autres variables du recensement¹, qui dépendent en outre de la répartition de la population correspondante sur le territoire considéré.

1. Le sondage dans les grandes communes de France métropolitaine

Le recensement est basé sur un plan de sondage qui est décrit en détail dans le document « Comprendre le recensement de la population »². Dans les communes de 10 000 habitants ou plus, pour estimer la population des ménages, le sondage sur un échantillon d'environ 40 % des logements sur cinq ans entraîne une marge d'incertitude sur les résultats. Celle-ci est mesurée par le **coefficient de variation**, noté **CV**, qui renseigne sur l'écart relatif possible entre la valeur estimée par le recensement et la « vraie » valeur de la population. En termes statistiques il se calcule comme le rapport de l'écart-type à la moyenne. La vraie valeur sera comprise, dans 95 % des cas, dans la plage de valeurs possibles suivante :

[valeur au recensement x (1 - 2CV) ; valeur au recensement x (1 + 2CV)]

Par exemple, si la population d'une commune calculée par le recensement est de 50 000 habitants et que son coefficient de variation est de 0,6 %, sa « vraie » population se situera entre 49 400 ($50\,000 * (1 - 2 * 0,006)$) et 50 600 ($50\,000 * (1 + 2 * 0,006)$).

2. Précision de la population au niveau régional, départemental et communal

La précision de la population des ménages en grandes communes de métropole est calculée en tenant compte du plan de sondage et du calage réalisé à l'Iris³. En revanche, elle ne retrace pas l'ensemble des

¹ Fiche thématique sur la précision des résultats du recensement

² Insee Méthodes - Hors série - mai 2005

³ Fiche thématique sur les pondérations

paramètres d'imprécision possibles autour du résultat national du recensement qui, au total, génèrent une incertitude nettement plus élevée (Brilhault et Caron, 2016)⁴.

Au niveau national, l'erreur aléatoire introduite par le sondage dans les grandes communes conduit à une imprécision de 0,02 %. Par conséquent, au niveau de l'ensemble de la population française, l'imprécision liée au sondage est de + ou – 15 800 personnes. Ce chiffre illustre le fait que, au niveau national, l'impact du sondage sur le résultat final est négligeable.

Le tableau suivant, tiré de l'article de Brilhault et Caron (2016), indique la distribution du coefficient de variation associé à la variable population au niveau régional, départemental, et communal.

**Distribution des coefficients de variation (en %) de la variable population
aux niveaux régional, départemental, et communal**

| | Région | Département | Commune |
|--------------|--------|-------------|---------|
| Maximum | 0,45 | 1,15 | 6,62 |
| 99 % | 0,45 | 1,15 | 2,07 |
| 95 % | 0,26 | 0,94 | 1,42 |
| 90 % | 0,25 | 0,70 | 1,22 |
| 75 % Q3 | 0,20 | 0,53 | 1,05 |
| 50 % Médiane | 0,16 | 0,35 | 0,88 |
| 25 % Q1 | 0,12 | 0,25 | 0,71 |
| 10 % | 0,09 | 0,15 | 0,55 |
| 5 % | 0,05 | 0,13 | 0,46 |
| 1 % | 0,03 | 0,03 | 0,32 |
| Minimum | 0,03 | 0,03 | 0,23 |
| Moyenne | 0,17 | 0,40 | 0,91 |

Lecture : au niveau régional, le CV de la variable population varie entre 0,03 % et 0,45 % selon les régions. Pour 95 % des régions, le CV de la variable population est inférieur à 0,26 % ; pour la moitié d'entre elles, ce CV est inférieur à 0,16 %, etc. En moyenne, ce CV est de 0,17 %.

Champ : communes de 10 000 habitants ou plus de France métropolitaine.

Source : recensement de la population de 2006.

Région et département

Pour la moitié des régions le CV est inférieur à 0,16 %. Pour un quart d'entre elles, il est même inférieur à 0,12 %. Seul un quart des régions (les moins peuplées) ont un CV supérieur à 0,20 %.

Pour la moitié des départements le CV est inférieur à 0,35 %. Pour un quart d'entre eux, il est même inférieur à 0,25%. Seul un quart des départements ont un CV supérieur à 0,53 %.

En l'absence de connaissance sur les spécificités de la région ou du département, il convient d'estimer la précision de sa population en prenant la valeur médiane.

L'erreur aléatoire introduite par le sondage dans les grandes communes sur la variable population diffère selon les régions et selon les départements, celle-ci dépendant en particulier du nombre de grandes communes composant la région ou le département considéré.

⁴ Brilhault et Caron (2016), « Le passage à une enquête par sondage : quel impact sur la précision du recensement ? » in *Économie et Statistique* n°483-484-485, avril 2016.

Commune

Pour la moitié des grandes communes, le CV est inférieur à 0,88 %. Pour un quart d'entre elles, il est même inférieur à 0,71 %. Seul un quart des grandes communes ont un CV supérieur à 1,05 %.

L'imprécision introduite par le sondage sur la variable population diffère selon la commune considérée. D'une part, elle varie avec la taille de la commune : exprimée en pourcentage de la population, elle est plus faible dans les plus grandes communes. D'autre part, les grandes adresses et les adresses nouvelles étant enquêtées exhaustivement⁵, leur part dans l'ensemble des adresses est susceptible d'influer sur la précision.

3. Précision de la population par commune selon la taille de la commune

Le tableau suivant fournit des éléments indicatifs sur la précision de la variable population au niveau communal selon la tranche de taille de la commune.

Distribution des coefficients de variation (en %) de la variable population par tranche de taille

| | 10 000 – 19 999 habitants (452 communes) | 20 000 – 49 999 habitants (318 communes) | 50 000 – 99 999 habitants (79 communes) | Au moins 100 000 habitants (43 communes) |
|--------------|---|---|--|---|
| 75 % Q3 | 1,16 | 0,87 | 0,71 | 0,43 |
| 50 % Médiane | 1,02 | 0,78 | 0,56 | 0,39 |
| 25 % Q1 | 0,90 | 0,68 | 0,50 | 0,34 |

Champ : communes de 10 000 habitants ou plus de France métropolitaine.

Source : recensement de la population de 2006.

Pour la moitié des communes de 10 000 à 19 999 habitants, le coefficient de variation est inférieur à 1,02 %. Pour l'autre moitié, il est supérieur à cette valeur. Pour un quart d'entre elles, il est même supérieur à 1,16 %, alors que pour le quart des communes où ce coefficient est le plus faible, il est inférieur à 0,90 %.

En l'absence de connaissance sur les spécificités de la commune, il convient d'estimer la précision de sa population en prenant la valeur médiane correspondant à sa tranche de taille.

⁵ Pour plus de détails sur le plan de sondage, voir : Insee Méthodes - Hors série - mai 2005