

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 recensement, qui dépendent en outre de la répartition de la population correspondante sur le territoire considéré.

1. Le sondage dans le recensement

Le recensement est basé sur un plan de sondage qui est décrit en détail dans le document « Comprendre le recensement de la population » ([Insee Méthodes - Hors série - mai 2005](#)). Dans les communes de 10 000 habitants ou plus, le sondage sur un échantillon d'environ 40 % des logements sur cinq ans entraîne une marge d'incertitude sur les résultats. Cette fiche donne des informations permettant d'apprécier l'ampleur de cette marge d'incertitude sur le calcul de la population vivant dans des ménages. Cette incertitude est mesurée par le coefficient de variation (noté CV) qui renseigne sur l'écart relatif possible entre la valeur publiée (celle donnée par le recensement, ici celui de 2006) 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 :

$$[\text{valeur du RP2006} \times (1 - 2\text{CV}) ; \text{valeur du RP2006} \times (1 + 2\text{CV})].$$

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 « véritable » population se situera entre :

$$50\,000 \times (1 - (2 \times 0,6)/100) = 49\,400 \text{ et } 50\,000 \times (1 + (2 \times 0,6)/100) = 50\,600.$$

2 - Précisions de la population au niveau national, régional et départemental

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. Comme précisé en introduction, ce chiffre ne retrace pas l'ensemble des paramètres d'imprécision possibles autour du résultat national du recensement qui, au total, génèrent une incertitude nettement plus élevée. Il illustre le fait que, au niveau national, l'impact du sondage sur le résultat final est négligeable.

Le tableau suivant donne la précision indicative de la variable population des ménages obtenue aux niveaux régional et départemental.

La précision de la variable population aux niveaux régional et départemental

	Distribution du CV de la variable population au niveau région	Distribution du CV de la variable population au niveau département
75% Q3	0,20	0,53
50% Médiane	0,16	0,35
25% Q1	0,12	0,25

Comment lire ce tableau ?

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.

3. Précision des données communales

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 exhaustivement, leur part dans l'ensemble des adresses est susceptible d'influer sur la précision.

Le tableau suivant donne la précision indicative de la variable population des ménages obtenue au niveau communal.

Coefficients de variation de la population des ménages pour les communes de 10 000 habitants ou plus

	Communes de 10000 à 19999 habitants (50% des grandes communes)	Communes de 20000 à 49999 habitants (36% des grandes communes)	Communes de 50000 à 100000 habitants (9% des grandes communes)	Communes de 100000 habitants et plus (5% des grandes 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

Comment lire ce tableau ?

Pour la moitié des communes de 10 000 à 19 999 habitants, le **coefficient de variation** (CV) 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 de sa tranche de taille.