

Recensement de la population

Les exploitations principale et complémentaire

Les questionnaires renseignés par les personnes recensées et collectés par les agents recenseurs sont saisis par un dispositif de lecture optique et font l'objet d'un vidéo-codage. L'Insee procède ensuite à une série de traitements statistiques : contrôle de la cohérence des informations collectées, imputation des valeurs manquantes, codification des variables de diffusion, calcul des pondérations en vue d'obtenir un fichier de données individuelles, anonymes et apurées, prêt à être exploité.

Comme lors des recensements antérieurs, l'exploitation statistique est réalisée en deux temps : l'exploitation « principale » puis l'exploitation « complémentaire ». Mais, contrairement au passé, le délai qui sépare les deux exploitations est très court et les résultats qui en sont issus sont diffusés simultanément.

1. L'exploitation principale

L'exploitation « principale » porte, après leur réception et leur contrôle, sur l'ensemble des questionnaires collectés. Elle est exhaustive pour les communes de moins de 10 000 habitants et porte sur environ 40 % des logements dans les communes de 10 000 habitants ou plus. Elle permet de produire un « fichier détail » contenant l'ensemble des logements et des individus recensés. Ces observations sont assorties d'un poids spécifique de l'exploitation principale¹.

Les résultats statistiques issus de cette exploitation portent sur de nombreuses variables et peuvent être classés en 11 thèmes :

- Population (âge, sexe, état matrimonial, actifs ou inactifs...) ;
- Activité des résidents (taux d'activité par sexe et âge, chômage, statut professionnel...) ;
- Emploi au lieu de travail (salariés ou non, temps partiel ou complet, conditions d'emploi...) ;
- Déplacements domicile - lieu de travail ;
- Formation (scolarisation et lieu d'études, diplômes...) ;
- Migrations (lieu de résidence cinq ans plus tôt, lieu de naissance) ;
- Nationalité (Français de naissance ou par acquisition, nationalité selon le sexe et l'âge...) ;
- Immigration (pays de naissance, année d'arrivée en France...) ;
- Ménages (taille du ménage, sexe et âge de la personne de référence...) ;
- Logements (type, époque d'achèvement...) ;
- Résidences principales (statut d'occupation, date d'emménagement, nombre de pièces, surface, sanitaires, mode de chauffage, nombre de voitures...).

¹ Pour en savoir plus, se reporter à la fiche thématique sur « **Les pondérations** ».

2. L'exploitation complémentaire

La seconde phase de l'exploitation statistique, dite « complémentaire », est destinée à produire les variables dont l'élaboration est complexe. Il s'agit de celles qui décrivent :

- la structure familiale des ménages : détermination précise de la personne de référence du ménage, identification, le cas échéant, de familles au sein du ménage et composition de ces familles ;
- les secteurs d'activité dans lesquels les emplois sont exercés ;
- la profession et la catégorie socioprofessionnelle des personnes notamment de celles qui exercent un emploi.

L'élaboration de ces informations fait appel à des traitements lourds, par exemple la consultation du répertoire des entreprises et des établissements SIRENE, qui ont en commun de nécessiter le recours à des interventions humaines dans certains cas, complexes ou particuliers, pour lesquels les procédures automatiques sont insuffisantes. Ces traitements sont donc longs et coûteux, et c'est la raison pour laquelle ils ne portent que sur un échantillon des questionnaires collectés.

2.1. L'échantillon de l'exploitation complémentaire

L'échantillon destiné à l'exploitation complémentaire est constitué ainsi :

- pour les communes de 10 000 habitants ou plus : l'ensemble des résidences principales recensées et de leurs habitants, soit environ 40 % ;
- pour les communes de moins de 10 000 habitants : 25 % des résidences principales recensées et leurs habitants ;
- pour toutes les communes : 25 % des personnes recensées dans les communautés et les habitations mobiles et 25 % des personnes sans abri ;
- l'ensemble des mariniers.

En pratique, le tirage de l'échantillon de l'exploitation complémentaire est réalisé de la façon suivante.

- Sélection des unités statistiques participant au tirage :
 - Les logements ordinaires, y compris les logements de fonction situés dans l'enceinte d'une communauté, ayant au moins un bulletin individuel, c'est-à-dire les **résidences principales**.
 - Les communautés ayant au moins un bulletin individuel des communautés (BIC) ou un bulletin individuel d'établissement pénitentiaire (BIEP).
 - L'ensemble des habitations mobiles, des personnes sans abri (à chacune on associe un logement fictif) et des mariniers.
 - Les fiches de logement non enquêtées (FLNE) correspondant à des résidences principales.
- Tirage des unités statistiques :
 - Pour la population hors communauté, l'unité statistique de tirage est le logement ; tous les individus du logement tiré font partie de l'échantillon de l'exploitation complémentaire.
 - Pour la population qui réside en communauté, l'unité statistique de tirage est l'individu. Les communautés dont un BIC au moins est tiré donnent lieu à la création d'un logement fictif.

Cas particuliers :

Les logements et individus enquêtés dans les communes de moins de 10 000 habitants de Corse sont retenus exhaustivement dans l'exploitation complémentaire. Il en est de même pour Saint-Pierre-et-Miquelon qui n'est composé que de deux communes de moins de 10 000 habitants.

2.2. Le contenu du fichier de l'exploitation complémentaire

L'exploitation complémentaire consiste à traiter les réponses relatives à l'ensemble des questions posées notamment celles qui figurent au verso du bulletin individuel et dans les pages centrales de la feuille de logement. Elle produit un fichier qui contient trois types d'observations ou « unités statistiques » : **des logements** et **des individus**, comme dans le fichier issu de l'exploitation principale, ainsi que **des familles**.

Les logements retenus étant des **résidences principales**, le niveau logement accueille aussi les informations relatives au **ménage** : nombre de personnes du ménage, structure familiale du ménage et caractéristiques de la personne de référence du ménage.

Chaque observation est assortie d'un poids spécifique de l'exploitation complémentaire. Les individus et les familles résidant dans un logement ont le même poids, égal au poids du logement.

Le fichier issu de l'exploitation complémentaire contient trois types de variables :

- Les variables présentes dans le fichier de l'exploitation principale et qui, pour les logements et les individus tirés dans l'échantillon de l'exploitation complémentaire, sont recopiées sans modification.
- Les variables « nouvelles », c'est-à-dire spécifiques à l'exploitation complémentaire. Elles se répartissent en deux sous-ensembles :
 - celles relatives à la structure familiale du ménage et aux liens qui unissent les personnes appartenant à un même ménage ;
 - celles relatives à l'activité, à l'emploi et aux professions.
- Les variables de l'exploitation principale qui sont recalculées lors de l'exploitation complémentaire en utilisant les informations supplémentaires disponibles notamment sur la composition du ménage, la profession ainsi que le secteur d'activité et la localisation de l'établissement où l'activité est exercée. Ces informations supplémentaires permettent d'améliorer la qualité des variables ainsi recalculées.
 - Six variables présentes dans le fichier de l'exploitation principale font l'objet d'une nouvelle codification dans l'exploitation complémentaire : statut professionnel, condition d'emploi, indicateur du lieu de travail, commune du lieu de travail, commune frontalière du lieu de travail et pays du lieu de travail. Dans le fichier issu de l'exploitation complémentaire seules figurent les variables recodifiées ; les variables correspondantes de l'exploitation principale ne sont pas reportées.
 - La détermination de la personne de référence du ménage a lieu dans chacune des deux exploitations.

Lors de l'exploitation principale, la personne de référence du ménage est déterminée de façon fruste à partir des seules informations alors disponibles ; elle est repérée par la variable « Indicateur de la personne de référence du ménage ». Lors de l'exploitation complémentaire, la personne de référence du ménage est déterminée de façon plus sophistiquée à partir de la structure familiale du ménage ; elle est repérée grâce à la variable « Lien à la personne de référence du ménage ». Ces deux modes de détermination étant différents, la personne de référence d'un ménage n'est pas toujours la même selon l'exploitation considérée. Ce peut-être une source d'écart dans la répartition des ménages selon les caractéristiques de la personne de référence. L'indicateur calculé lors de l'exploitation principale n'est pas recopié dans le fichier de l'exploitation complémentaire.

Le choix de l'exploitation à privilégier pour ces variables résulte d'un arbitrage faisant intervenir les critères suivants :

- origine des autres variables : si un tableau fait intervenir des variables spécifiques de l'exploitation complémentaire (activité économique, catégorie socioprofessionnelle, structure familiale du ménage...) alors, pour des raisons de cohérence, il convient d'utiliser la variante « complémentaire » des variables recodifiées ;
- qualité de la variable : en particulier le lieu de travail est mieux localisé au complémentaire ;
- précision des données : pour les communes de moins de 10 000 habitants, la précision est supérieure à l'exploitation principale qui est exhaustive alors que l'exploitation complémentaire porte sur un échantillon au quart ; pour les communes de 10 000 habitants ou plus ce critère n'intervient pratiquement pas dans la mesure où le sondage est le même pour les ménages, ménages qui représentent, en moyenne, 98 % de la population.

De façon plus générale, les avantages respectifs de chacune des deux exploitations sont les suivants :

- **les résultats issus de l'exploitation principale ont une précision supérieure, surtout pour les communes de moins de 10 000 habitants et les zones qui en englobent²;**
- **l'exploitation complémentaire fournit des résultats plus complets et améliore la qualité de quelques variables.**

2.3. Les résultats tirés de l'exploitation complémentaire

Les résultats socio-économiques qui s'en dégagent peuvent être classés en 12 thèmes :

- huit thèmes de l'exploitation principale, enrichis de nouvelles variables d'analyse, notamment les catégories socioprofessionnelles, les secteurs d'activité, la structure familiale du ménage ;
- les thèmes emploi au lieu de travail, déplacements domicile-lieu de travail et ménages profondément modifiés (variables recodifiées) et enrichis (nouvelles variables) par rapport à l'exploitation principale ;
- un thème nouveau : **les familles** (couples, familles monoparentales, caractéristiques de la personne de référence de la famille, nombre et âge des enfants...).

Comme lors des recensements antérieurs, en raison du sondage, les résultats tirés de l'exploitation complémentaire ne sont significatifs que pour des **zones géographiques et/ou des sous-populations d'une taille suffisante**. La taille requise est d'au moins 2 000 habitants. C'est la raison pour laquelle les résultats de l'exploitation complémentaire ne font pas l'objet d'une diffusion standard pour les zones de moins de 2 000 habitants. Cependant, même pour une zone de plus de 2 000 habitants, il convient de s'assurer que la précision de la donnée considérée autorise bien l'analyse envisagée.

Cette contrainte est plus forte que pour les résultats issus de l'exploitation principale dans la mesure où l'ensemble des communes font l'objet d'un sondage et pas seulement les communes de 10 000 habitants ou plus.

² Pour en savoir plus, se reporter à la fiche thématique sur « **La précision des résultats du recensement** ».