



Mode d'utilisation des données

1. Description des données

Outre les résultats statistiques, les bases contiennent les informations suivantes :

- l'identifiant de l'IRIS constitué du code département, du code commune et du code IRIS à l'intérieur de la commune. Pour Paris, Lyon et Marseille, le code commune identifie l'arrondissement municipal. Le code IRIS est '0000' pour les communes non découpées en IRIS ;
- les codes région, département, commune (ou arrondissement municipal) et unité urbaine 2010 qui permettent notamment d'extraire des sous-ensembles de la base ;
- les codes TRIRIS et "Grand quartier" qui permettent d'effectuer les regroupements d'IRIS correspondants ; le code TRIRIS vaut 'ZZZZZZ' et le code "Grand quartier" est composé du code département, du code commune et de '00' pour les communes non découpées en IRIS ou lorsque le regroupement n'est pas possible ;
- le type d'IRIS qui permet de distinguer les IRIS d'habitat 'H' des IRIS d'activité 'A' et des IRIS divers 'D' ; le type d'IRIS est 'Z' pour les communes non découpées en IRIS. Le type d'IRIS a été déterminé lors de la création des IRIS, en 1999. Depuis, les caractéristiques de l'IRIS ont pu évoluer (population, nombre d'emplois, ...) sans que le type d'IRIS n'ait toujours été modifié ;
- le label de l'IRIS qui renseigne sur la nature des informations fournies concernant la précision des données (cf. ci-après "Conseils d'utilisation") ;
- le code modification de l'IRIS qui permet de savoir si les contours de l'IRIS ou de la commune ont été modifiés depuis 1999 (cf. ci-après § 3) ;
- le libellé de la commune (ou de l'arrondissement municipal) et le libellé de l'IRIS.

On repère les variables :

- de l'exploitation principale grâce à la lettre "P" au début du code de la variable et à la mention "(princ)" dans l'intitulé de la colonne ;
- de l'exploitation complémentaire grâce à la lettre "C" au début du code de la variable et à la mention "(compl)" dans l'intitulé de la colonne.

Pour quelques IRIS, très peu peuplés, les données ne sont pas renseignées. Ce sont essentiellement des IRIS d'activité ou divers qui n'ont aucun logement tiré lors de l'exploitation complémentaire.

Les données ne sont pas des nombres entiers car elles sont issues de calculs liés aux techniques de sondage ; elles sont proposées avec six décimales. Afin d'éviter les erreurs d'arrondi, les six décimales doivent être conservées dans tous les calculs intermédiaires.

2. Privilégier la réalisation de classements ou de typologies d'IRIS

L'IRIS, maille de base de la géographie infra-communale, permet de conduire des analyses sur les disparités au sein d'un territoire communal ou supra-communal. Que ce soit à partir des résultats du recensement de la population ou à partir d'autres sources, le territoire de l'IRIS n'est pas destiné à être analysé en tant que tel, comme peut l'être un

grand quartier, en vue de produire une monographie. La réalisation de classements ou typologies d'IRIS sera donc privilégiée. Pour faciliter l'étude d'une zone composée de communes découpées et non découpées en IRIS, par exemple une unité urbaine, la base fournit, pour ces communes non découpées, les données au niveau communal.

3. Comparer avec les résultats du recensement de la population de 1999

Les communes de 10 000 habitants ou plus et la plupart des communes de 5 000 à 9 999 habitants ont été découpées en IRIS pour la diffusion des résultats du recensement de 1999. Depuis, les IRIS n'ont été modifiés qu'à la marge pour tenir compte, chaque année, des modifications des limites communales.

Par ailleurs, l'Insee a effectué, au cours du deuxième trimestre 2008, la retouche de quelques IRIS afin de prendre en compte les évolutions importantes de la voirie et de la démographie. Cette retouche a été réalisée en respectant les règles établies en 1999 avec la Commission Nationale de l'Informatique et des Libertés (CNIL) et en préservant au maximum la continuité des séries de diffusion.

Ces différentes modifications sont identifiées grâce au code modification de l'IRIS présent dans chaque base. Il se décline selon les modalités suivantes :

- 00 pas de modification,
- 1Z scission d'IRIS,
- 2A à 2AB déplacement de limites d'IRIS (du groupe 2A au groupe 2AB),
- 3A fusion de communes irisées,
- 3B rétablissement de communes issues de communes irisées,
- 4A échange de parcelles entre communes irisées,
- 4B échange de parcelles habitées entre communes irisées et non irisées,
- 4C échange de parcelles inhabitées entre communes irisées et non irisées.

Pour reconstituer, chaque fois que c'est possible, l'IRIS en géographie 1999 à partir du ou des IRIS en géographie 2011, il convient d'utiliser la table de passage mise à disposition en format Excel. Cette table concerne les IRIS modifiés depuis 1999 et contient le même code modification de l'IRIS que ci-dessus à l'exception de la modalité '00'. Ainsi par exemple, si un IRIS de 1999 a été scindé en deux IRIS en 2011 (code modification à '1Z'), la table assure, pour ces IRIS, la correspondance entre les deux numéros IRIS de 2011 et le numéro IRIS de 1999.

La table de passage est accessible, depuis le site "www.insee.fr", à partir du menu "Définitions et méthodes", rubrique "Code officiel géographique, zonages d'études ", en cliquant sur le lien "Découpage infra-communal en IRIS" de la partie "Zonages d'études".

4. Les contours des IRIS

Pour la France métropolitaine et les DOM, les contours des IRIS sont accessibles depuis le site "www.insee.fr", à partir du menu "Définitions et méthodes", rubrique "Code officiel géographique, zonages d'études ", en cliquant sur le lien "Découpage infra-communal en IRIS" de la partie "Zonages d'études".

Pour les COM, ils sont accessibles depuis le site "www.insee.fr", à partir du menu "Bases de données", "Les résultats des recensements de la population".

Conseils d'utilisation

Les conseils qui suivent concernent essentiellement la robustesse des données.

RÉSUMÉ

Pour les communes de France métropolitaine, la base de données à l'IRIS est accompagnée d'informations sur la précision pour toutes les variables et d'un label de qualité synthétique de l'IRIS.

Pour les DOM et les COM, seul le label de qualité de l'IRIS est disponible.

Les informations sur la précision se présentent sous la forme d'une table donnant, pour chaque variable, le coefficient de variation (en %) en fonction de la tranche d'effectif.

Cette table est utilisable pour les IRIS des communes de 10 000 habitants ou plus de France métropolitaine repérés par le label de l'IRIS valant '1' (LAB_IRIS = 1). Elle ne porte que sur la population des ménages.

Il est conseillé de regrouper les IRIS dont le label vaut '2' (LAB_IRIS = 2) avec au moins deux autres IRIS.

La présence, dans un IRIS, de population hors ménage (communautés, habitations mobiles, personnes sans abri, marinières) doit inciter à la vigilance car les données relatives à ces populations sont, en général, moins robustes que celles concernant la population des ménages.

Les données contenues dans les bases sont fournies avec un niveau de détail adapté à la finesse du découpage en IRIS afin de garantir une assez bonne robustesse. Pour autant, il peut se faire qu'au niveau de certains IRIS la précision de telle ou telle donnée soit faible. Ces données sont néanmoins diffusées pour trois raisons :

- diffuser les effectifs de toutes les modalités d'une variable ; par exemple, les effectifs par catégorie socioprofessionnelle (CS) sont diffusés pour les agriculteurs alors que, compte tenu de la faiblesse des effectifs concernés, cette donnée est très fragile ;
- permettre le regroupement de modalités ainsi que le calcul d'indicateurs pour des regroupements d'IRIS correspondant à des zones infra-communales personnalisées ou standard comme les grands quartiers ;
- ne pas préjuger du niveau de qualité requis par l'utilisateur sachant que ce niveau dépend de l'utilisation qui sera faite de la donnée.

Cette diffusion est assortie de la fourniture à l'utilisateur d'informations sur la qualité des données. La maille IRIS étant plus fine, les résultats qui s'y rapportent sont, par construction, plus fragiles que ceux produits aux niveaux communal ou supra-communal et l'exigence en termes de qualité est donc plus forte. En particulier, certains IRIS présentent des configurations qui rendent les résultats les concernant plus fragiles. Les indications sur la qualité des données fournies aident à bien positionner les limites de l'utilisation statistique et incitent à procéder, le cas échéant, à des regroupements d'IRIS.

1. Le coefficient de variation : indicateur de qualité privilégié

L'indicateur de qualité privilégié est le coefficient de variation (CV) : il renseigne sur la précision liée au sondage. Si pour un effectif "a", le coefficient de variation est "c", l'effectif a 95 % de chances de se trouver dans l'intervalle : $[a \pm 2 \cdot a \cdot c]$.

Exemple illustrant une utilisation du coefficient de variation :

la population âgée de 18 à 24 ans est de 250 avec un CV de 10 % dans l'IRIS 1 et de 320 avec un CV de 8 % dans l'IRIS 2.

Cela signifie que les effectifs ont 95 % de chances de se trouver respectivement dans les intervalles : 250 ± 50 et 320 ± 51 . Les deux intervalles se recouvrent en partie : on ne peut pas conclure avec certitude que l'effectif de l'IRIS 2 est supérieur à celui de l'IRIS 1.

En revanche, si l'effectif de l'IRIS 3 est de 400 avec un CV de 8%, l'intervalle à 95 % est : 400 ± 64 ; il est totalement disjoint de l'intervalle 250 ± 50 ; on peut donc conclure avec une très forte probabilité que l'effectif de l'IRIS 3 est supérieur à celui de l'IRIS 1.

2. Plusieurs cas repérés par le label de l'IRIS

L'information fournie sur la qualité des données est différente selon les communes et selon les IRIS. C'est le "Label" de l'IRIS qui indique dans quel cas on se trouve.

a) Les IRIS appartiennent à une commune de France métropolitaine de 10 000 habitants ou plus

Dans ces communes, trois cas sont à distinguer :

- pour les IRIS dont le label vaut '1', l'utilisation de la table fournissant les coefficients de variation pour estimer la précision d'une donnée à l'IRIS est possible. Ces IRIS sont les plus nombreux, ils représentent près de 90 % des IRIS d'habitat de ces communes. Au vu de la précision et de la stabilité de l'estimation de leur population, leur échantillon a été jugé d'une qualité qui autorise le calcul et l'utilisation de la table de précision (cf. §3. Les tables de précision) ;

- pour les IRIS dont le label vaut '2', la préconisation est de les regrouper au TRIRIS ou en ensembles de plusieurs IRIS équivalent à un TRIRIS. La table est alors utilisable pour estimer la précision des résultats ainsi regroupés (cf. §3. Les tables de précision). Pour ces IRIS, l'échantillon n'a pas été jugé suffisamment représentatif. C'est en général lié à une structure de l'habitat peu homogène. En regroupant ces IRIS avec d'autres, ils perdent de leur spécificité et l'échantillon de l'ensemble recouvre une représentativité correcte. Environ 10 % des IRIS d'habitat sont dans cette classe ;

- les IRIS dont le label vaut '3' sont atypiques. Ils sont mal représentés par l'échantillon parfois en raison du mode de tirage mais généralement en raison de leurs caractéristiques. Les informations les concernant peuvent être entachées d'une très forte imprécision et la table de précision n'est pas utilisable. Il s'agit essentiellement d'IRIS d'activité et d'IRIS divers. Quelques dizaines d'IRIS d'habitat, dont la structure est très particulière, sont également dans ce cas ;

- les IRIS dont le label vaut "Z" (quelques dizaines d'IRIS) sont des IRIS qui ont été définis pour une petite commune devenue grande. C'est le cas lors de fusion de communes ou lorsqu'une commune franchit à la hausse le seuil de 10 000 habitants et change alors de mode de recensement. Le label de ces IRIS n'a pas été calculé et il convient donc d'utiliser leurs données avec précaution.

b) Les IRIS appartiennent à une commune des départements d'outre-mer (DOM) de 10 000 habitants ou plus

Le calcul des coefficients de variation est disponible, pour la seule variable de population. L'équivalent des tables de précision pour les autres variables calculées pour les IRIS des communes de 10 000 habitants ou plus de France métropolitaine n'existe donc pas encore pour les IRIS des communes de 10 000 habitants ou plus des DOM. Compte tenu des spécificités du mode de tirage de l'échantillon dans les DOM, il n'est pas possible de se référer à celles calculées pour la France métropolitaine.

Néanmoins, un label de qualité a pu être calculé pour les IRIS des communes de 10 000 habitants ou plus des DOM selon le même principe que pour ceux de France métropolitaine, pour permettre de distinguer trois cas parmi ces IRIS :

- pour les IRIS dont le label vaut '1' (70 % des IRIS de ces communes), la précision et la stabilité de l'estimation de la population sont d'une qualité suffisamment bonne. Ceci autorise l'utilisation des informations fournies au niveau de ces IRIS ;

- pour les IRIS dont le label vaut '2' (environ 25 % des IRIS de ces communes), la précision et la stabilité de l'estimation de la population ne sont pas d'une qualité suffisamment bonne pour que les données fournies au niveau de ces IRIS soient utilisables. C'est en général lié à une structure de l'habitat peu homogène. La préconisation est de regrouper ces IRIS au TRIRIS ou en ensembles de plusieurs IRIS équivalent à un TRIRIS. En effet, la précision et la stabilité de l'estimation de la population, au niveau de ces TRIRIS ou équivalents, sont de bonne qualité. Pour cette catégorie d'IRIS, ceci autorise donc l'utilisation des informations au niveau TRIRIS ou ensembles équivalents ;

- les IRIS dont le label vaut '3' (environ 5 % des IRIS de ces communes) sont atypiques. Ils sont trop peu peuplés pour pouvoir être correctement représentés par l'échantillon. La précision et la stabilité de l'estimation de la population sont de mauvaise qualité, même en regroupant ces IRIS au TRIRIS ou en ensembles de plusieurs IRIS équivalent à un TRIRIS. Il s'agit essentiellement d'IRIS d'activité ou d'IRIS divers. Quelques IRIS d'habitation dont la structure est très particulière sont également dans ce cas ;

- les IRIS dont le label vaut "Z" (quelques IRIS) sont des IRIS qui ont été définis pour une petite commune devenue grande. C'est le cas lors de fusion de communes ou lorsqu'une commune franchit à la hausse le seuil de 10 000 habitants et change alors de mode de recensement. Le label de ces IRIS n'a pas été calculé et il convient donc d'utiliser leurs données avec précaution.

Il faut souligner que les trois premières modalités proposées, pour les communes de 10 000 habitants ou plus des DOM, n'ont pas tout à fait le même sens que pour celles de France métropolitaine, le niveau d'exigence sur la précision de l'estimation de la population étant moindre pour les IRIS des communes de 10 000 habitants ou plus des DOM.

c) Les IRIS appartiennent à une commune des collectivités d'outre-mer (COM) de 10 000 habitants ou plus

Le calcul des coefficients de variation n'a pas encore été réalisé, les tables de précision ne sont pas disponibles. Néanmoins, d'autres indicateurs de qualité permettent de distinguer trois cas parmi ces IRIS :

- pour les IRIS dont le label vaut '1', l'échantillon est de bonne qualité. Ceci autorise l'utilisation des informations fournies au niveau de ces IRIS ;

- pour les IRIS dont le label vaut '2', l'échantillon n'est pas jugé suffisamment représentatif. C'est en général lié à une structure de l'habitat peu homogène. La préconisation est de regrouper ces IRIS au TRIRIS ou en ensembles de plusieurs IRIS équivalent à un TRIRIS. Ce faisant, l'échantillon de l'ensemble recouvre une représentativité correcte ;

- les IRIS dont le label vaut '3' sont atypiques. Ils sont trop peu peuplés (moins de 1 000 habitants) pour pouvoir être correctement représentés par l'échantillon. Les informations les concernant peuvent être entachées d'une très forte imprécision. Il s'agit essentiellement d'IRIS d'activité ou d'IRIS divers. Quelques IRIS d'habitat dont la structure est très particulière sont également dans ce cas.

Il faut souligner que les trois modalités proposées pour les communes de 10 000 habitants ou plus des COM n'ont pas tout à fait le même sens que pour celles de France métropolitaine ou des DOM.

d) Les IRIS appartiennent à une commune de moins de 10 000 habitants

Pour ces IRIS le label vaut '4'.

Si la donnée provient de l'exploitation principale, la collecte et l'exploitation sont exhaustives, il n'y a pas d'imprécision liée au sondage.

Si la donnée provient de l'exploitation complémentaire, pour un effectif "a" le coefficient de variation est approché par la formule : $CV = 2/\sqrt{a}$.

Dans ces communes, l'actualisation de la population de la commune entre deux collectes peut induire une fragilité des données au niveau infra-communal. En effet, elle conduit à répartir l'évolution de la population entre les IRIS au prorata de leur population. Or, cette évolution peut être liée pour une part importante à des opérations immobilières localisées dans des IRIS déterminés. En conséquence, pour ces communes, surtout si elles ont fait l'objet d'opérations immobilières importantes, il est conseillé de privilégier les données correspondant à une année de collecte et les évolutions entre deux années de collecte.

e) Les communes de moins de 10 000 habitants ne sont pas découpées en IRIS

Pour ces communes dont le label vaut '5', la précision des données communales les concernant peut être obtenue de la même façon que pour les IRIS des communes de moins de 10 000 habitants (cf. alinéa précédent).

3. Les tables de précision (pour la France métropolitaine)

Les tables de précision fournissent, sauf exception, pour chaque variable le coefficient de variation (CV) en fonction de la tranche d'effectif de cette variable. Les tables comportent dix tranches d'effectif. Pour des raisons techniques, les coefficients de variation n'ont pas été calculés pour la tranche "Moins de 50". Par ailleurs, pour une variable donnée, dans certaines tranches d'effectif le nombre d'IRIS a été jugé insuffisant pour pouvoir calculer un coefficient de variation significatif. La précision est alors indiquée, en s'appuyant sur le coefficient de variation de la tranche voisine, par un majorant ou un minorant du coefficient de variation. Et lorsqu'il n'y a pas d'IRIS pour lesquels la variable se situe dans la tranche concernée, les tables contiennent la mention "sans objet" (so).

Considérons, par exemple, la variable "Population de 18 à 24 ans" :

- pour les 7 tranches de 50 à 1 999 : la table donne pour chaque tranche une valeur du CV ;
- pour la tranche "Moins de 50" : la table donne un minorant déduit du CV de la tranche "50 à 99" ;
- pour la tranche "2 000 à 2 999" : il y a trop peu d'IRIS tels que l'effectif des habitants âgés de 18 à 24 ans soit compris entre 2 000 et 2 999 pour calculer un CV. Un majorant déduit du CV de la tranche "1 000 à 1 999" est fourni ;
- pour la tranche "3 000 ou plus" : la table indique "Sans objet" (so) car aucun IRIS ne compte 3 000 habitants ou plus âgés de 18 à 24 ans.

Pour certaines variables de la base "Population" les coefficients de variation n'ont pu être calculés. La mention "nd" est alors indiquée.

Les DOM étant pour l'instant exclus du calcul des coefficients de variation, les tables ne fournissent pas de précision pour les variables "Logement" spécifiques aux DOM. Le calcul des coefficients de variation n'a pas été possible, non plus, pour quelques variables de la base "Logement" qui totalisent un nombre de pièces ou un nombre d'années. La mention "nd" est alors affichée.

Pour l'élaboration des tables de précision, seule la population des ménages des IRIS d'habitat est prise en compte. La présence de population hors ménage dans un IRIS (communautés, personnes résidant dans une habitation mobile, personnes sans abri, marinières) doit inciter à la vigilance (1). Ces IRIS sont repérables grâce à la variable "Population hors ménage" présente dans la base Population. En effet, les résultats concernant cette population sont moins robustes que ceux concernant la population des ménages.

S'agissant des communautés, elles sont présentes exhaustivement dans l'exploitation principale, les données qui en sont issues ne sont donc pas affectées d'imprécision liée au sondage. En revanche, l'exploitation complémentaire porte sur un quart des individus des communautés et les données correspondantes sont affectées d'une imprécision : pour un effectif "a" le coefficient de variation est approché par la formule : $CV = 2/\sqrt{a}$.

Outre le sondage, la collecte elle-même peut être source de fragilité. En effet, de par la nature même des communautés (maisons de retraite, établissements pénitentiaires, établissements sanitaires de long séjour, etc.) les bulletins sont souvent incomplètement renseignés même si le comptage des personnes est, lui, parfaitement réalisé. Les traitements informatiques affectent alors aux personnes concernées les caractéristiques d'une personne en principe proche. Cependant, à un niveau fin comme celui des IRIS, ces traitements automatiques ne peuvent exactement suppléer les informations manquantes.

Par ailleurs, les comparaisons avec les résultats du recensement de 1999 sont parfois délicates. En effet, certaines communautés de 1999 sont devenues des logements : c'est le cas notamment de résidences pour étudiants dès lors que les logements comportent une installation pour faire la cuisine.

La population des habitations mobiles et des personnes sans abri est, comme celle des communautés, présente exhaustivement à l'exploitation principale et au quart dans l'exploitation complémentaire. La fragilité tient plutôt à sa localisation au sein de la commune. Ainsi, les personnes sans abri sont localisées dans la zone de collecte où elles se trouvaient lorsqu'elles ont été recensées.

Quant aux marinières, recensés et traités exhaustivement, ils sont comptés dans leur commune de rattachement au sein de laquelle ils sont localisés de façon arbitraire dans l'IRIS le plus peuplé.

Pour plus d'information méthodologique, consulter, depuis le site "www.insee.fr", les fiches "Conseils pour l'utilisation des résultats du recensement" du menu "Bases de données > Les résultats des recensements de la population > Documentation sur les résultats du recensement", rubrique "Comprendre et utiliser le recensement". Elles présentent les nouveautés introduites à l'occasion de la rénovation du recensement de la population, leurs conséquences sur les résultats et des conseils d'utilisation. Elles précisent les changements affectant les principales variables et leur incidence sur la comparabilité avec les résultats du recensement de 1999. Une fiche sur la précision apporte des conseils sur la façon de calculer et d'utiliser le coefficient de variation.