Nations Unies ECE/CES/2023/6



Distr. générale 11 avril 2023 Français

Original : anglais

Commission économique pour l'Europe

Conférence des statisticiens européens

Soixante et onzième réunion plénière
Genève, 22 et 23 juin 2023
Point 4 b) de l'ordre du jour provisoire
Coordination des activités statistiques internationales
dans la région de la Commission économique pour l'Europe :
Les groupes difficiles à atteindre dans le contexte
des sources administratives

Les groupes difficiles à atteindre dans le contexte des sources administratives

Document établi par le Danemark avec des contributions du Canada, des États-Unis d'Amérique, de l'Italie et de la Nouvelle-Zélande

Résumé

Le présent examen approfondi porte sur les principes et les moyens permettant de représenter les groupes difficiles à atteindre dans les sources administratives. Ce document résume l'expérience des organismes nationaux de statistique en la matière et décrit les problèmes et les enjeux liés à cette question. Il y est également proposé d'engager d'éventuels travaux complémentaires pour améliorer l'accès aux groupes difficiles à atteindre à l'aide des données administratives. La dernière section récapitule les délibérations de la réunion de février 2023 du Bureau de la Conférence des statisticiens européens.

À la suite de l'examen, le Bureau de la Conférence a souscrit à l'idée de poursuivre les travaux dans ce domaine et a décidé de créer une nouvelle équipe spéciale, comme le recommande le document. Le Danemark présidera cette équipe spéciale.

La Conférence est invitée à approuver les résultats de l'examen approfondi, à réfléchir aux principales conclusions et recommandations figurant dans le présent document et à apporter une contribution supplémentaire aux travaux de la nouvelle équipe spéciale.



I. Résumé analytique

- 1. Le présent examen approfondi, demandé par le Bureau de la Conférence des statisticiens européens, porte sur les principes et les moyens permettant de représenter les groupes difficiles à atteindre dans les sources administratives. Cette note résume l'expérience des organismes nationaux de statistique en la matière et décrit les problèmes et les enjeux liés à cette question. Il y est également proposé d'engager d'éventuels travaux complémentaires pour améliorer l'accès aux groupes difficiles à atteindre à l'aide des données administratives.
- La volonté de « ne laisser personne de côté », inscrite dans le Programme de développement durable à l'horizon 2030, constitue désormais la toile de fond de bon nombre des efforts déployés pour améliorer la qualité et la couverture des statistiques. Une désagrégation multiple des données est préconisée de toutes parts pour pouvoir analyser l'impact de facteurs de vulnérabilité qui se recoupent. Or de nombreuses dimensions en fonction desquelles les utilisateurs souhaiteraient que les données soient ventilées concernent des groupes de population considérés comme difficiles à atteindre, c'est-à-dire des groupes généralement sous-représentés, parce qu'ils sont numériquement très restreints, qu'ils sont difficiles à identifier faute de définitions normalisées ou de données sur les variables pertinentes, qu'ils choisissent de ne pas se faire identifier, par exemple en raison de la stigmatisation associée à l'appartenance à tel ou tel groupe, qu'ils sont systématiquement exclus des techniques de collecte et des bases d'échantillonnage standard (personnes vivant en institution, par exemple), qu'ils sont physiquement difficiles à atteindre (car vivant dans des régions reculées, n'ayant pas de domicile fixe, etc.), ou qu'ils sont difficiles à dénombrer même une fois identifiés et échantillonnés (personnes atteintes de démence, personnes ne parlant pas la langue du pays, jeunes enfants, etc.).
- 3. Le passage généralisé à l'utilisation de sources administratives pour les recensements et les enquêtes accentue certaines des difficultés rencontrées s'agissant de sauvegarder et d'améliorer les données et statistiques disponibles sur les groupes difficiles à atteindre. Le problème se pose par exemple dans les cas suivants : prise en compte des sans-abri et sans domicile fixe et des immigrants sans papiers dans les statistiques dérivées des registres administratifs; moyens de représenter les minorités (ethniques, religieuses, minorités de genre) marginalisées et les personnes handicapées dans les statistiques si les sources administratives ne sont pas en mesure de saisir ces caractéristiques ou ne le font pas systématiquement; et production de statistiques ventilées par sexe pour permettre une analyse des questions de genre sur des sujets cruciaux, lorsque les sources administratives utilisées pour fournir les informations sont collectées au niveau du ménage plutôt que de l'individu.
- 4. Cinq pays (le Canada, le Danemark, les États-Unis, l'Italie et la Nouvelle-Zélande) ont contribué à l'élaboration du présent document. Même si leurs systèmes statistiques diffèrent les uns des autres, ils se heurtent tous à des problèmes analogues pour identifier les groupes difficiles à atteindre dans leurs registres administratifs.
- 5. Le fait que des populations soient *difficiles à atteindre* dans le contexte des données administratives peut être interprété de différentes manières. L'une des interprétations tient au caractère incomplet des registres administratifs, ce qui rend certains groupes, tels que les enfants ou les personnes âgées, *difficiles à atteindre* et, partant, à décrire par des données. Il en est question dans les cas du Canada et des États-Unis. Selon une autre interprétation, il s'agit de sélectionner certains groupes, tels que les sans-abri, les immigrés clandestins ou les autochtones, puis de tenter de les faire apparaître dans des dossiers administratifs. Une telle approche est décrite par le Danemark, les États-Unis, l'Italie et la Nouvelle-Zélande.
- 6. La façon dont les registres administratifs sont actuellement utilisés pour accéder aux groupes difficiles à atteindre laisse entrevoir quelques pratiques permettant d'identifier certains de ces groupes, mais leur application aux statistiques varie d'un pays à l'autre. Des données sur les groupes difficiles à atteindre peuvent, dans certains cas, être extraites de dossiers administratifs, principalement étayés par des enquêtes dans tel ou tel domaine, mais il ne semble guère que ces données soient utilisées pour produire des statistiques régulières, liées à la législation. En outre, les exemples présentés montrent que l'identification des populations difficiles à atteindre dépend du contexte national, mais qu'elle s'appuie souvent

sur des enquêtes ou des recensements qui tentent de saisir des attributs tels que le numéro d'identification, l'adresse ou la date de naissance, pour aider à identifier ces populations dans les registres administratifs.

7. Il est recommandé de créer une équipe spéciale chargée d'examiner s'il est possible à ce stade d'engager des travaux visant à améliorer l'accès aux groupes difficiles à atteindre dans le contexte des registres administratifs.

II. Introduction

- 8. Dans tous les domaines des statistiques sociales ainsi que dans divers domaines des statistiques économiques et commerciales, il est désormais de plus en courant de recourir aux sources administratives soit pour compléter les sources traditionnelles (recensements et enquêtes), soit pour les remplacer. Certains pays procèdent déjà à un recensement complet à partir de sources administratives. Depuis l'adoption du Programme 2030, l'utilisation de sources administratives est privilégiée pour les statistiques, tendance qui s'est accélérée avec la pandémie de COVID-19. Il y a de nombreux avantages à cela, notamment une moindre charge de travail pour les personnes interrogées, la production plus rapide de statistiques et la réduction concomitante des coûts. Les inconvénients sont néanmoins bien connus, parmi lesquels la dépendance à l'égard du contenu des sources administratives et la difficulté de recueillir des données sur des caractéristiques subjectives ou d'autres variables qui ne sont pas consignées dans ces sources.
- 9. Pendant la pandémie, les organismes nationaux de statistique ont dû recourir aux sources administratives pour pallier l'absence ou les insuffisances des données collectées par les moyens traditionnels. En même temps, les nouvelles exigences en matière de données et le Programme 2030 nécessitent une meilleure désagrégation des données, à la fois pour mieux cerner différentes catégories de population et pour garantir le respect du principe consistant à « ne laisser personne de côté », inhérent au Programme 2030. La communication d'informations sur diverses composantes de la population est un défi d'un point de vue statistique, car il peut être difficile d'obtenir des données « d'enquête » sur elles, tant sur le plan administratif qu'au sens physique ou géographique. Un tel état de choses, conjugué à la baisse constante des taux de réponse aux enquêtes, fait qu'il est compliqué d'obtenir des éléments d'information suffisants pour les groupes difficiles à atteindre.
- 10. Le présent document décrit les initiatives en cours visant à fournir de meilleures informations sur les groupes difficiles à atteindre à partir de données administratives. Il indique quels sont les groupes actuellement considérés comme difficiles à atteindre, présente quelques pratiques suivies en la matière et envisage des dispositions éventuelles à prévoir. L'intention est non pas de passer en revue l'ensemble des modes de représentation des groupes difficiles à atteindre dans les statistiques, ni de souligner l'importance de la désagrégation pour les objectifs de développement durable (ODD), mais de se concentrer sur les sources administratives, l'accent étant mis sur la présentation des démarches possibles.
- 11. Il convient enfin de faire une distinction importante au sujet des groupes *difficiles à atteindre* dans le contexte des données administratives. La notion de groupe de population *difficile à atteindre* peut être interprétée de deux façons, puisqu'il s'agit :
- a) Soit de groupes *difficiles à atteindre* dans tout contexte statistique, c'est-à-dire dans les enquêtes ou les données administratives (sans-abri, immigrés clandestins, etc.);
- b) Soit de groupes *difficiles à atteindre* dans les registres administratifs en raison d'une sous-déclaration liée à un signalement tardif, comme dans le cas d'enfants, de catégories de population très mobiles (jeunes) ou de personnes âgées.

Le présent document décrit les deux cas de figure.

III. Domaine statistique visé

- 12. Le concept de groupe de population difficile à atteindre d'un point de vue statistique découle généralement du fait que bon nombre de techniques d'échantillonnage standard ne sont guère applicables, ou se révèlent souvent inopérantes, car les populations cibles ne sont pas accessibles par des cadres fondés sur des sources de données traditionnelles, telles qu'une liste des logements. Par exemple, si les membres d'une population cible sont peu nombreux ou font l'objet d'une stigmatisation dans l'ensemble de la population, il peut être coûteux et/ou difficile de les contacter en utilisant les approches probabilistes traditionnelles. L'autre solution consistant à recourir aux données administratives offre la possibilité d'améliorer la couverture du cadre pour certaines populations cibles, mais il se peut également que d'autres populations restent difficiles à atteindre ou « cachées » pour divers groupes d'intérêts.
- 13. En tout état de cause, l'identification transversale des groupes difficiles à atteindre n'est pas une tâche aisée, pas plus que leur prise en compte par les sources administratives, soit qu'ils échappent à ces sources, soit que l'accès en soit plus ou moins facile suivant les pays. La définition des groupes difficiles à atteindre diffère en outre d'un pays à l'autre, en dépit des tentatives faites au niveau international pour en fournir une description uniforme. Les raisons pour lesquelles les membres d'un groupe de population sont difficiles à atteindre peuvent varier en fonction de l'environnement national, géographique ou social.
- 14. Dans le contexte des sources administratives, les groupes *difficiles* à atteindre ne sont généralement pas couverts ni aisément identifiables, d'où la nécessité d'y suppléer au moyen d'enquêtes sous une forme ou une autre pour élargir le champ des statistiques ou des analyses. Le couplage de ces deux types de données (résultats des enquêtes et données administratives) peut être compliqué s'il faut disposer pour cela d'un « identifiant ». Les groupes *difficiles* à atteindre sont de fait peu accessibles.
- 15. La question de l'identification des groupes difficiles à atteindre et de leur couverture par les sources administratives est une véritable gageure. En ce qui concerne les données proprement dites, les sources administratives et les informations décrivant les groupes difficiles à atteindre ne se recoupent guère. La plupart des chevauchements découlent des données d'enquête qui saisissent l'identifiant unique de la personne en question, ce qui permet de les relier à des registres administratifs. Il y a néanmoins un large éventail de cas dans lesquels les données ne peuvent pas être utilisées de manière convaincante. Il peut s'agir par exemple d'une personne qui s'est vu attribuer un logement mais qui, en raison de troubles mentaux, vit comme un sans-abri. En pareil cas, elle ne sera pas prise en compte par les données administratives mais pourrait l'être par une enquête spécifique sur le sans-abrisme.

IV. Pratiques des pays

- 16. Les informations présentées ci-dessous récapitulent les réponses fournies par les organismes nationaux de statistique du Canada, du Danemark, des États-Unis, de l'Italie et de la Nouvelle-Zélande sur le sujet. Les pays retenus pour le présent document l'ont été sur la base d'une connaissance subjective de l'utilisation des registres administratifs dans la production de statistiques au niveau national.
- 17. Le principal intérêt des données provenant des registres administratifs tient au vaste sous-ensemble de la population qu'elles couvrent. Cependant, si des groupes particuliers n'y sont pas enregistrés, ils ne peuvent pas être décrits par les données provenant de cette source. En revanche, s'ils sont pris en compte, les données enregistrées peuvent ouvrir la voie à une analyse statistique plus large du groupe.
- 18. Les expériences de certains pays concernant les moyens d'identifier les groupes difficiles à atteindre dans les registres administratifs sont décrites ci-après.

A. Canada

- 19. Un certain nombre de personnes et de groupes de population sont considérés comme « difficiles à atteindre » dans le cadre d'un recensement traditionnel. L'analyse ci-dessous reprend des exemples tirés de recherches sur l'utilisation de données administratives dans le cadre du recensement de la population canadienne. Le Canada étant dépourvu de registre de population, les travaux portant sur une plus large utilisation des données administratives aux fins du recensement s'appuient sur des fichiers administratifs corrélés. En règle générale, le point de départ des données administratives est souvent l'information disponible au niveau individuel : les personnes concernées peuvent ainsi être échantillonnées à partir d'une base plus ciblée, plutôt que par des enquêtes fondées sur le logement.
- 20. Dans le contexte canadien, les autochtones vivant dans des réserves, les habitants des régions isolées ou septentrionales et les personnes vivant dans des logements collectifs (centres de long séjour) pourraient être considérés comme *difficiles à atteindre*. Dans les cas où les membres de ces groupes de population recourent aux services publics ou interagissent avec des organismes qui collectent des informations, les données administratives offrent la possibilité d'aider à dénombrer des personnes et des groupes qui sont difficiles à atteindre par un recensement traditionnel.
- 21. Il est également question de populations difficiles à atteindre dans le cas d'opérations de collecte perturbées à la suite d'une catastrophe naturelle ou d'une pandémie. Pour le recensement de 2021, Statistique Canada a élaboré un plan de contingence statistique fondé sur l'utilisation sûre, responsable et appropriée des données administratives à l'appui de la collecte dans le cas de telles perturbations. Bien que ce recensement ait été un succès, certaines unités de collecte (configuration géographique de la collecte des données de recensement) ont eu des taux de réponse plus faibles que prévu en raison de la pandémie de COVID-19, parmi d'autres problèmes. Le Canada compte environ 49 000 unités de collecte. Sur ce total, 1 045 unités affichaient de faibles taux de réponse tout en disposant de données administratives de bonne qualité, se prêtant à un processus d'imputation. Pour les logements non participants, l'âge, le sexe à la naissance et le nombre de résidents habituels ont ainsi été imputés. Cela correspond à une proportion minime de la population canadienne, mais Statistique Canada prévoit un recours accru aux données administratives pour le recensement de 2026¹.
- 22. L'utilisation accrue de données administratives pour les programmes de recensement et d'enquête pourrait nécessiter une redéfinition de la notion de personne et de population difficiles à atteindre. À la différence d'un recensement traditionnel, les groupes de population difficiles à atteindre dans les données administratives ne sont pas toujours définis en fonction de caractéristiques communes aux membres de ces groupes, telles que leurs caractéristiques ethniques, géographiques ou socioéconomiques. Du point de vue de la couverture des données administratives, les nouveaux types de populations difficiles à atteindre qui apparaissent peuvent ou non coïncider avec les groupes de population difficiles à atteindre dans un recensement traditionnel. Quelques exemples de nouveaux types de populations difficiles à atteindre dans le contexte canadien sont présentés ci-dessous.
- 23. Premièrement, une source de données administratives pourrait constituer un échantillon incomplet de l'ensemble de la population à prendre en considération, car les données ont été obtenues à des fins administratives et non statistiques. Cela soulève la question de savoir qui apparaît ou n'apparaît pas dans les sources de données administratives à un moment donné. Un travail de recherche utilisant des données administratives dans le cadre du recensement canadien de 2021 a par exemple fait apparaître des retards dans la réception des statistiques de l'état civil à l'approche du jour du recensement, le 11 mai 2021, d'où une surreprésentation des personnes âgées de 80 ans et plus, ainsi qu'une sous-représentation des enfants de moins de 1 an.

On trouvera de plus amples informations sur le plan de contingence statistique à partir du lien suivant : https://www12.statcan.gc.ca/census-recensement/2021/ref/98-304/2021001/app-ann1-7-eng.cfm.

- 24. Deuxièmement, les différences conceptuelles entre les types d'informations collectées dans les programmes de recensement/d'enquête et les types d'informations disponibles dans les données administratives pourraient faire apparaître de nouveaux groupes difficiles à atteindre. Au Canada, comme dans d'autres pays ne disposant pas d'un registre de population, des différences se manifestent souvent sur le plan conceptuel entre les informations administratives et les informations issues d'enquêtes. Par exemple, les informations sur les conjoints et les personnes à charge sont bien établies dans de nombreuses sources administratives, mais d'autres relations telles que l'union libre peuvent être sous-représentées dans les données administratives, compte tenu des différences entre les notions fiscales et celles d'un recensement traditionnel. Il est également à noter que les familles monoparentales sont moins bien représentées dans les données administratives que dans le recensement traditionnel. Tout attribut non collecté ni bien défini par un enregistrement administratif dans le cas d'une personne ou d'une composante de la population fait de celle-ci une population difficile à atteindre dans les données administratives.
- 25. Troisièmement, une personne peut, dans les données administratives, être considérée comme difficile à atteindre lorsque les informations la concernant sont brouillées par de multiples enregistrements contradictoires provenant de différentes sources. Au Canada, il est procédé au couplage probabiliste des données pour relier plusieurs fichiers administratifs entre eux, après quoi des modèles statistiques permettent d'identifier telle ou telle personne pour le recensement. Indépendamment des aspects statistiques des groupes de population difficiles à atteindre, il devient délicat, dans les pays dépourvus de registre de population, d'attribuer une adresse ou une localisation géographique à certaines personnes dans les sources de données administratives. En l'occurrence, l'intéressé n'est pas nécessairement difficile à atteindre, mais son individualisation et son identification dans des fichiers corrélés peuvent poser des problèmes. Lorsque des fichiers administratifs semblent contradictoires, il peut s'agir : a) de deux personnes différentes ; b) d'une même personne à différents moments; ou c) d'une seule personne ayant à un moment donné plusieurs adresses et numéros de téléphone. Dans le contexte canadien, la recherche sur les méthodes d'intégration statistique progresse, à l'aide de techniques d'estimation à petite échelle et de modèles linéaires hiérarchiques qui combinent des données d'enquête et des données administratives pour aplanir ces difficultés.
- <u>Groupes d'âge difficiles à atteindre</u>: 0-1 an, 18-24 ans et 80 ans et plus. Trois groupes d'âge distincts se révèlent difficiles à atteindre dans les données administratives pour différentes raisons. Les enfants en bas âge (de 0 à 1 an) sont sous-dénombrés car les données administratives relatives à la naissance sont reçues et intégrées avec retard. Les jeunes de 18 à 24 ans constituant une population très mobile, les sources administratives d'informations géographiques les concernant sont souvent contradictoires et il est difficile de leur attribuer un lieu de résidence habituel. Pour les sujets âgés de plus de 80 ans, deux problèmes se posent : premièrement, un surdénombrement dû aux retards dans la réception de certains actes de décès et, deuxièmement, un dénombrement incorrect dans les données administratives, car les plus de 80 ans peuvent résider dans des établissements de soins de longue durée tout en étant comptés comme résidents habituels par le compagnon ou la compagne resté à la maison. En outre, les conjoints, les comptables ou les enfants adultes de ces personnes les aident souvent à remplir leur déclaration de revenus et à s'acquitter d'autres tâches administratives, les personnes âgées en question étant ainsi rattachées par erreur à l'endroit où se trouve le déclarant (tel qu'un enfant adulte) et non à leur lieu de résidence habituel (un hébergement collectif, par exemple).
- 27. Émigrants difficiles à atteindre: le Canada n'exige pas des citoyens ou des résidents qu'ils s'enregistrent lorsqu'ils quittent le pays. Dans leur grande majorité, ceux-ci reviennent au bout de quelques jours ou de quelques semaines, mais une petite partie d'entre eux réside à l'étranger pendant une longue période tout en conservant des liens fiscaux avec le Canada. Les émigrants sont donc difficiles à détecter dans les données administratives, en raison du nombre élevé des cas de « faux positifs » concernant des Canadiens qui sont hors du pays pendant le recensement, mais qui résident au Canada et y retournent (ou ont l'intention d'y retourner) au cours de l'année. Ce problème se posant également pour les recensements traditionnels, le suivi des émigrants à l'aide de données administratives est envisagé pour les études de couverture.

B. Danemark

- 28. Le Danemark dispose d'un vaste système de données administratives, véritable « épine dorsale » du système statistique, qui repose sur trois registres principaux : le registre de la population, le registre des logements et le registre des entreprises. Ceux-ci sont fréquemment mis à jour par un dispositif de transfert automatisé de données, de façon à faire apparaître le dernier état disponible. Grâce à un identifiant unique, le système permet d'apparier les données, ce qui rend possible la production de séries chronologiques statistiques dans différents domaines. En outre, le Danemark est l'un des pays où une nette majorité des statistiques sociales sont produites à partir de données administratives. Chaque citoyen dispose d'un numéro personnel à des fins administratives. Les informations figurant dans le numéro personnel couvrent de nombreux aspects, tels que le sexe, l'âge et l'adresse. Le numéro personnel est utilisé comme identifiant unique pour relier les données à d'autres domaines statistiques, parmi lesquels l'éducation, la santé, le revenu ou l'emploi, ce qui permet d'obtenir des informations détaillées sur divers groupes de population de manière anonyme. Le Danemark procède à des recensements en s'appuyant sur des données administratives.
- 29. Cependant, en ce qui concerne les groupes difficiles à atteindre, ceux-ci ne sont pas facilement identifiables dans les registres administratifs, car il peut y avoir toutes sortes de circonstances dans lesquelles les données ne fournissent pas d'informations sur la situation récente de la personne en question. Tel peut être le cas par exemple d'une personne sans domicile fixe qui a une adresse ou d'une personne n'ayant pas d'adresse sans être considérée comme sans-abri.
- 30. Certains sous-groupes de sans-abri peuvent être pris en compte par les statistiques sur les centres d'hébergement et les foyers (enquête statistique menée par Statistics Denmark). Dans la plupart des cas, les personnes qui ont besoin de passer la nuit dans un centre d'hébergement ou un foyer sont invitées à fournir leur numéro personnel. Ces informations peuvent ensuite servir à l'établissement de statistiques. La question qui se pose ici est de savoir quelle est la part des sans-abri qui cherchent à être ainsi hébergés pour la nuit. Il est donc difficile de fournir une quelconque forme de codage pour les statistiques générales. Le Centre danois de recherche en sciences sociales réalise une estimation semestrielle des personnes sans abri au Danemark et Statistics Denmark y contribue au moyen des statistiques sur les centres d'hébergement et les foyers. Cependant, aucune autre statistique sur les sans-abri n'est produite à partir de cette estimation. L'autre facteur de complication est que les organisations bénévoles qui viennent en aide aux sans-abri le font souvent à condition qu'il ne soit pas obligatoire de collecter le numéro personnel des personnes en question.
- 31. Les immigrants clandestins, de par leur nature, ne peuvent pas être pris en compte dans les registres administratifs. Dans certains cas, il leur est éventuellement attribué un numéro d'identification temporaire, principalement à des fins administratives. Ce numéro ne contient pas d'autres informations que l'âge, le sexe et le pays d'origine et ne peut servir à des fins statistiques car il est généralement attribué une seule fois ; si l'intéressé entre à nouveau en contact avec les autorités, il se verra attribuer un autre numéro d'identification temporaire, sans lien avec le précédent. Ce n'est que dans des cas très sporadiques, lorsqu'un immigrant clandestin est reconnu avec certitude, qu'il se verra attribuer le même numéro d'identification qu'antérieurement, ce qui pourrait théoriquement servir à recueillir quelques informations succinctes à son sujet. Cependant, cela ne suffit pas pour produire des statistiques, même à titre expérimental.
- 32. Les registres administratifs offrent différentes possibilités d'identification des groupes difficiles à atteindre, quoique au prix d'efforts considérables. La production de statistiques sur les groupes de population difficiles à atteindre n'étant légalement requise ni au Danemark ni dans l'Union européenne, il reste encore à étudier tout l'éventail des moyens de produire de telles statistiques, notamment en corrélant les informations issues d'enquêtes et les registres administratifs.

C. Italie

- 33. Pour remplacer le recensement décennal, l'Institut national italien de statistique (Istat) a lancé en 2018 le recensement permanent de la population et des logements, conformément au programme de modernisation de l'Istat, qui place le système intégré de registres statistiques au cœur de la production statistique. L'élément central du recensement permanent est le registre de base de la population, dont les principales sources sont les registres locaux de la population des municipalités italiennes, tandis que deux enquêtes par sondage (enquête aréolaire et enquête par listes) sont menées chaque année afin de faciliter la tenue des registres, c'est-à-dire, au sens large, d'en évaluer la qualité, ainsi que d'ajouter les informations manquantes, incomplètes ou de qualité insuffisante. Cela permet de disposer chaque année de statistiques en matière de recensement.
- 34. Quant au dénombrement de la population, il a d'abord été obtenu en appliquant des coefficients de correction pour les erreurs de sous-couverture et de surcouverture aux individus inscrits comme résidents habituels dans le registre de base de la population (le modèle de capture-recapture a été adopté pour les estimations directes des erreurs de couverture du registre de base de la population, les données présentes dans le registre de la population constituant la « première capture » et les données recueillies sur le terrain la « deuxième capture »). En 2020, à la suite de l'annulation des enquêtes en raison de la pandémie de COVID-19, il a été fait appel à une méthode différente, entièrement fondée sur l'utilisation de « signes de vie » administratifs et l'application de critères de classification aux registres statistiques. Plus précisément, les résidents habituels (dans le registre de base de la population) sans signes de vie administratifs dans d'autres sources ont été assimilés à des individus en surnombre dans le registre de base de la population, tandis que les individus non inscrits dans le registre de base de la population comme résidents habituels mais ayant manifesté des signes de vie administratifs selon d'autres sources ont été considérés comme omis en raison d'une sous-couverture du registre. Cette correction a ensuite fait l'objet de micro-interventions, consistant à reclasser les fichiers individuels dans le registre de base de la population.
- 35. En raison de cette poussée contrainte vers une plus large utilisation des données administratives, il a fallu repenser le cadre statistique nécessaire à l'évaluation de la qualité des processus d'estimation du recensement permanent. Plus généralement, il a fallu aussi repenser la conception du recensement permanent en se servant des données d'enquête pour mesurer la qualité d'une estimation de la population désormais entièrement fondée sur des registres. Une enquête de contrôle devrait être menée tous les deux à quatre ans pour fournir des mesures de la qualité concernant l'estimation de la taille de la population fondée sur des registres, tandis qu'une enquête aréolaire à petite échelle pourrait être réalisée afin d'évaluer la sous-couverture des sources administratives qui, par définition, n'incluent pas des groupes de population tels que les immigrants sans papiers. En même temps, l'Istat s'efforce de mieux tirer parti des signes de vie administratifs dans le nouveau cycle du recensement permanent (après 2021) et de prendre également en compte l'erreur que constitue une mauvaise localisation dans le registre de population, qui n'a pas été évaluée jusqu'à présent. En outre, l'acquisition de nouvelles sources (par exemple, les archives des services collectifs de distribution, notamment les données sur la consommation d'énergie et celles des compteurs intelligents, qui fourniront très probablement des éléments d'évaluation objectifs concernant le lieu réel de résidence habituelle) devrait marquer un tournant.
- 36. Dans le cadre de cette conception générale, une distinction a été faite dans le cas de certains sous-groupes de population, dont les particularités nécessitent l'adoption d'une approche spécifique et qui ont donc été exclus du champ d'application des enquêtes de recensement susmentionnées. Ces groupes difficiles à atteindre sont : les sans-abri et les personnes sans domicile fixe qui, même dans les recensements traditionnels, représentent une population typique « difficile à dénombrer » en raison du risque élevé de sous-couverture ; les personnes vivant en institution ; et les personnes vivant dans des établissements de type formel, informel ou non autorisé.

- 37. Pour estimer la taille de ces sous-groupes de population, une enquête distincte sur les adresses des groupes difficiles à atteindre est donc menée chaque année à partir des registres locaux de la population. Plus précisément, les adresses correspondant à celles d'institutions, de sans-abri² et d'établissements formels/informels sont identifiées dans le registre de base de la population et soumises à la validation des bureaux de recensement des municipalités au moyen d'une plateforme en ligne. Pour chaque adresse confirmée (y compris les nouvelles adresses), il convient d'ajouter le nombre d'individus par sexe et par nationalité. Le registre de base de la population est ensuite mis à jour sur la base de ces informations : les personnes appartenant à l'un des trois groupes y sont identifiées grâce à leur adresse et signalées clairement afin que des données de base sur ces groupes puissent être tirées du registre de population.
- 38. Ce processus permet de dénombrer avec plus de précision les groupes difficiles à atteindre dans le cadre du dénombrement fondé sur les registres. Les personnes appartenant aux catégories enregistrées ont été identifiées dans le registre de base de la population et dénombrées en tant que telles, c'est-à-dire qu'elles n'ont pas été soumises à l'application de coefficients de correction (dénombrement de 2018-2019) ou à un reclassement sur la base des signes de vie administratifs (dénombrement de 2020 et recensement de la population).
- 39. Cela étant, surtout en ce qui concerne les sans-abri, il est évident qu'il n'y a qu'un chevauchement partiel entre la population enregistrée et la population réelle. Bon nombre de sans-abri (en particulier les sans-papiers) ne sont enregistrés dans aucune municipalité, tandis que d'autres ne le sont pas en tant que tels (car ils sont toujours officiellement enregistrés en tant que membres de leur ancien ménage); en même temps, certaines des personnes enregistrées à des adresses fictives sont des nomades ou des personnes qui n'ont pas de lieu de résidence habituel en raison de leur travail.
- 40. Étant donné la sous-représentation notoire des sans-abri et des sans domicile fixe dans les données administratives et, plus généralement, les problèmes liés à leur identification, et compte tenu de la nécessité de collecter des données supplémentaires en plus des données de base qui peuvent être dérivées des variables disponibles dans le registre (c'est-à-dire la localisation et la répartition par sexe, âge et nationalité), deux enquêtes de terrain sur les sansabri et les personnes sans domicile fixe sont actuellement en cours de planification. Une enquête sur les utilisateurs des cantines, dortoirs et autres services essentiels fournis aux sans-abri et aux personnes vivant dans des conditions d'extrême pauvreté sera menée dans un certain nombre de municipalités à travers le pays (celles où le nombre de sans-abri est le plus élevé), tandis qu'une enquête ponctuelle sera réalisée dans les 14 plus grandes municipalités. Cette dernière visera à produire une estimation quantitative des sans-abri, voire à faire concorder les données collectées sur le terrain avec les chiffres relatifs aux sans-abri enregistrés, tandis que la première recueillera des données d'ordre qualitatif sur le profil sociodémographique et les conditions de vie des personnes en situation d'extrême pauvreté.
- 41. En outre, en raison de problèmes de qualité, la prochaine vague de l'enquête administrative sur les groupes difficiles à atteindre sera menée au niveau du code du ménage pour les sans-abri et les personnes vivant dans des établissements formels/informels. En effet, si pour les personnes vivant en institution, l'information sur l'adresse est suffisante pour identifier avec précision la population cible dans le registre, il n'en va pas de même pour les deux autres agrégats. Les codes des ménages (au lieu des adresses) seront donc soumis à validation au moyen de la plateforme en ligne et il sera ainsi plus facile d'identifier la population concernée dans le registre de base de la population dans les cas où les données agrégées déclarées par les municipalités ne coïncident pas avec celles dérivées de ce registre. Enfin, il convient de mentionner le risque de sous-couverture pour les personnes vivant en institution (par exemple, les personnes âgées résidant dans des établissements de soins de longue durée), car dans de nombreux cas (pour des raisons fiscales ou autres), elles sont

² En effet, selon la législation italienne, les sans-abri et les personnes sans domicile fixe ont le droit d'être inscrits dans les registres locaux de la population, ce pour quoi les municipalités utilisent des adresses fictives. Dans certains cas, les adresses d'ONG venant en aide aux sans-abri, aux migrants et aux personnes démunies sont également utilisées.

toujours enregistrées à l'adresse de leur ancien domicile. Cette question sera traitée dans le cadre du problème plus large des erreurs de localisation mentionné ci-dessus.

D. Nouvelle-Zélande

- 42. L'infrastructure de données intégrée (IDI) de Stats NZ permet aux chercheurs d'accéder à des microdonnées dépersonnalisées. Elle comprend des données provenant de divers organismes publics, d'enquêtes de Stats NZ et du recensement de la population et des logements. Il s'agit notamment des données relatives aux enregistrements des naissances et des décès et aux mouvements internationaux aux frontières et, sur des sujets tels que l'éducation, le travail, les revenus, les versements de prestations, la justice et la santé, des données recueillies auprès des administrations publiques.
- 43. Il n'existe pas d'identifiant national de la population en Nouvelle-Zélande, ni de registre national de la population conçu à des fins administratives. Pour faciliter l'intégration des données dans l'infrastructure de données intégrée, il est établi une liste centrale des identités (« épine dorsale » de l'infrastructure de données intégrée) en combinant trois sources de données de qualité : les enregistrements des naissances, les immatriculations fiscales et les demandes de visa. Les sources de données sont rapprochées à l'aide d'un couplage probabiliste par paire. L'« épine dorsale » de l'infrastructure de données intégrée est conçue de façon à prendre en compte, dans la mesure du possible, toutes les personnes ayant vécu en Nouvelle-Zélande, y compris celles disposant d'un visa de travail ou d'étude : en 2021, 10 millions de personnes environ y étaient comptabilisées (la population actuelle de la Nouvelle-Zélande avoisine 5 millions de personnes). Un identifiant codé commun à tous les ensembles de données est attribué à chaque identité dans l'infrastructure de données intégrée. Toutes les autres sources de données sont ensuite reliées à l'infrastructure de données intégrée.
- 44. Les processus liés à l'infrastructure intégrée de données déterminent, dans la mesure du possible, des caractéristiques démographiques de base âge, « sexe ou genre », origine ethnique et lieu de résidence habituelle pour toutes les personnes figurant dans l'« épine dorsale » de cette infrastructure. Ces variables sont obtenues à partir d'une combinaison d'informations provenant d'une série de sources de données.
- 45. Une population résidente administrative à une date de référence donnée peut être construite à partir de l'« épine dorsale » de l'infrastructure de données intégrée. Il s'agit des personnes figurant dans celle-ci qui ont une activité dans certaines sources de données administratives sur une période de deux ans précédant la date de référence. Les personnes décédées avant la date de référence sont identifiées par un lien avec les données des registres de décès et sont exclues. Les données sur les mouvements internationaux aux frontières sont utilisées pour exclure toute personne qui ne résidait pas en Nouvelle-Zélande à la date de référence, par exemple un résident qui émigre pour vivre à l'étranger ou un visiteur de courte durée en Nouvelle-Zélande.
- 46. <u>Jeunes adultes</u>. Quasiment toutes les personnes figurant dans l'« épine dorsale » de l'infrastructure de données intégrée ont un âge connu, la variable est de qualité très satisfaisante et les jeunes adultes peuvent être bien identifiés. En ce qui concerne la population résidente, il a été constaté que l'écart entre les personnes dénombrées à partir de sources administratives et les chiffres officiels est relativement uniforme pour tous les âges, y compris les jeunes adultes. Cependant, il est plus difficile de fournir une adresse précise pour les jeunes adultes en utilisant uniquement des sources administratives. En comparant l'adresse de résidence habituelle dérivée des sources administratives avec les données de recensement, la concordance est plus faible pour les jeunes adultes, en particulier la tranche d'âge des 20 à 24 ans. Une tendance analogue est observée à toutes les échelles géographiques, même si elle est moins prononcée pour les zones relativement étendues.
- 47. <u>Autochtones</u>. Les Maoris sont le peuple autochtone de Nouvelle-Zélande. L'ascendance maorie et l'appartenance au groupe ethnique maori sont les deux principales façons de les identifier.

- 48. L'ascendance maorie, déterminée à partir d'enregistrements des naissances pour l'enfant et ses parents, est consignée depuis 1995. Le recensement de 2013 présente une bonne couverture pour les personnes nées en Nouvelle-Zélande avant 1995 et pour les migrants arrivés dans le pays avant 2013. Utilisé de cette manière, il peut être considéré comme une source de données administratives ponctuelle qui est précieuse pour fournir des informations historiques antérieures aux enregistrements des naissances. La qualité des données sur l'ascendance maorie provenant de ces deux sources est très satisfaisante dans le cas de personnes pour lesquelles une valeur est disponible, mais il manquait une valeur pour l'ascendance maorie dans 14 % des cas en 2018. La liste électorale consigne également l'ascendance maorie, et la combinaison des enregistrements des naissances et des données de la liste électorale (personnes âgées de 18 ans et plus qui s'inscrivent pour voter) aiderait à combler le manque de données et à fournir en continu des informations sur une grande partie de la population. L'adoption récente de la loi sur les données et les statistiques de 2022 permet à Stats NZ d'accéder aux données électorales et la priorité sera accordée à des travaux de recherche une fois que les données seront disponibles.
- 49. L'appartenance ethnique est une mesure de l'identité culturelle en Nouvelle-Zélande et un facteur social clef utilisé pour décrire la population. On estime à 17 % environ la part de la population néo-zélandaise d'origine maorie.
- 50. Des données sur l'appartenance ethnique sont régulièrement collectées par plusieurs organismes publics. Ceux-ci s'efforcent de les recueillir conformément à la norme statistique applicable, mais diverses contraintes liées à la collecte et au traitement des données font que la qualité des informations sur l'appartenance ethnique est variable. Il y a également des raisons légitimes pour lesquelles les réponses peuvent différer, car la vision qu'ont les gens du groupe ethnique auquel ils appartiennent peut changer au fil du temps, et ils ne donnent pas toujours les mêmes réponses sur leur appartenance ethnique dans des situations différentes. Pour harmoniser les informations provenant de différentes sources, l'infrastructure de données intégrée applique une méthode qui classe ces sources en fonction de la qualité de leurs données sur l'appartenance ethnique et sélectionne la meilleure source disponible.
- 51. En utilisant uniquement des sources administratives, il est possible de connaître l'appartenance ethnique de quasiment tout le monde. Bien que la collecte de données sur l'appartenance au groupe ethnique maori par certains organismes suscite quelques préoccupations, la méthode de classement selon la qualité donne généralement de bons résultats. La répartition par âge du recensement administratif de la population suit de près, mais à un niveau légèrement inférieur, les tendances des données officielles et les chiffres du recensement administratif de la population par âge sont similaires ou supérieurs à ceux du recensement de 2018.
- 52. <u>Sans-abri</u>. Même si les personnes sans abri sont identifiées par quelques ministères et organisations non gouvernementales qui leur fournissent des services, les données administratives sur ces personnes sont limitées et fragmentaires.
- 53. Stats NZ a établi une définition du sans-abrisme en distinguant quatre catégories : « vivant à la rue », « vivant dans un hébergement temporaire », « partageant un logement » et « occupant un logement inhabitable ». Il n'y a pas de méthode globale ou standard d'identification ou de dénombrement de la population sans-abri.
- 54. La plupart des estimations du sans-abrisme reposent sur le recensement quinquennal. Les données de recensement sont complétées par des données administratives pour les organisations s'occupant des asiles de nuit, des refuges pour femmes et d'autres lieux d'hébergement destinés aux personnes qui n'ont pas accès à un logement minimal adéquat. Les problèmes liés au recensement de 2018 ont pesé sur la qualité des informations relatives à la grave précarité en matière de logement. Cela montre bien combien il est essentiel d'élaborer d'autres sources de données pour observer ce phénomène en fournissant des données plus régulières et en faisant une comparaison avec les résultats du recensement.
- 55. <u>Migrants illégaux.</u> La Nouvelle-Zélande dispose de solides données administratives pour identifier les migrants qui séjournent illégalement sur son territoire.

- 56. Le pays est séparé de ses voisins les plus proches par des milliers de kilomètres de haute mer. Cet isolement géographique signifie qu'il n'y a guère de possibilités pour les demandeurs d'asile, les réfugiés ou tout autre migrant de franchir les frontières de la Nouvelle-Zélande sans passer par les points de contrôle officiels. Stats NZ estime que très peu de personnes entrent illégalement en Nouvelle-Zélande et que le problème le plus important est probablement celui des personnes qui sont admises dans le pays avec un visa valide mais qui y restent au-delà de la durée de validité de leur visa.
- 57. Les administrations publiques collectent et conservent des données sur les voyageurs internationaux et les migrants. Les douanes néo-zélandaises détiennent des données sur les passeports pour pratiquement toutes les arrivées et tous les départs à la frontière. Le Ministère du commerce, de l'innovation et de l'emploi gère les données relatives aux demandes de visa, aux approbations de visa et aux mouvements aux frontières (avec des données limitées sur les citoyens néo-zélandais et australiens qui bénéficient d'un droit de libre accès). Immigration New Zealand contrôle le respect des conditions d'obtention des visas et fournit des informations au Système d'observation permanente des migrations, notamment des données sur les expulsions et les départs volontaires. Les personnes admises en Nouvelle-Zélande en tant que réfugiés et demandeurs d'asile peuvent également être identifiées à partir des données de l'immigration contenues dans l'infrastructure de données intégrée et font l'objet d'une étude récente.
- 58. <u>Santé mentale et dépendance</u>. Les données sur la santé contenues dans l'infrastructure de données intégrée comprennent des informations sur la santé mentale et la dépendance dans le cas des personnes qui recourent aux services de santé. L'utilisation de ces services, telle qu'elle ressort des microdonnées administratives, sert de point de départ pour comprendre les problèmes de santé mentale et de dépendance. Pour aider les chercheurs, cinq sources essentielles de données sur la santé mentale et les addictions ont été regroupées dans l'infrastructure de données intégrée en un seul tableau des événements.
- 59. L'infrastructure de données intégrée fait apparaître deux niveaux de services de santé auxquels correspondent des événements clairement codés en matière de santé mentale et de dépendance. Il s'agit des services spécialisés enregistrés par le projet d'intégration des données sur la santé mentale et des hospitalisations consignées dans l'ensemble de données minimales nationales. Le troisième niveau, à savoir celui des services de santé des collectivités, n'est pas disponible dans l'infrastructure de données intégrée.
- 60. On compte en outre trois services de soutien qui recouvrent différents niveaux et types de services et fournissent des informations sur la santé mentale et les addictions. Il s'agit des analyses de laboratoire, de la délivrance de médicaments et des certificats médicaux utilisés pour l'aide sociale.
- 61. Dans le tableau d'ensemble, les événements liés à l'addiction se limitent à l'abus d'alcool et d'autres drogues (il est question dans ce cas de troubles liés à la consommation de substances psychoactives).
- 62. Si les données administratives de l'infrastructure de données intégrée portent surtout sur l'utilisation effective des services de santé, les besoins en matière de santé mentale et d'addiction ont été essentiellement déterminés au moyen d'enquêtes.
- 63. <u>Personnes handicapées.</u> Les principales informations sur les personnes handicapées proviennent du recensement et de l'enquête sur le handicap réalisée après celui-ci.
- 64. Rares sont les données administratives sur le handicap. Certaines organisations rendent compte des services qu'elles assurent en faveur de la communauté des personnes handicapées (par exemple dans les domaines de la santé, de l'éducation ou des aides de l'État). Le secteur du handicap a constamment encouragé les programmes et les services à collecter des données relatives au handicap, ce qui contribue à la quantité d'informations disponibles dans ce domaine. Le Bureau des services aux personnes handicapées (Office for Disability Services) a formulé des conseils à l'intention des organisations sur la manière de collecter des données sur le handicap de manière appropriée et efficace, et a réuni des exemples et des conseils concernant les moyens d'évaluer la qualité des données administratives et de les utiliser.

E. États-Unis

- 65. Aux États-Unis, le système statistique est décentralisé. À l'heure actuelle, le Census Bureau des États-Unis ne dispose pas d'informations globales sur la population du pays fondées sur les recensements antérieurs et les données administratives provenant de l'enregistrement des naissances et des décès, des impôts, du secteur de la santé et d'autres sources. Pour cette raison, il se peut que les groupes de population *difficiles à atteindre* sur le plan des registres administratifs ne soient pas les mêmes pour le Census Bureau des États-Unis que pour d'autres pays.
- 66. Aux États-Unis, certains secteurs de la population sont plus difficiles à atteindre que d'autres. L'un de ces groupes est celui des enfants. Les trois derniers recensements décennaux les ont sous-dénombrés, notamment dans le cas des jeunes enfants. Il a également été constaté que le sous-dénombrement de ces derniers dans les données administratives était plus important que dans le cas d'autres groupes d'âge. Cela peut avoir un impact sur le contenu et la qualité des statistiques, puisque les résultats des recensements ont servi à la fois de base au programme d'estimation de la population pour l'attribution de fonds fédéraux, d'éléments de vérification de la population pour les enquêtes démographiques et de dénominateurs pour le calcul des taux démographiques. Une des solutions a consisté à recourir aux estimations nationales du programme d'analyse démographique 2020, qui utilise les données relatives aux naissances, aux décès, à l'immigration et à l'émigration de façon à constituer une base mixte pour le programme d'estimation de la population.
- 67. Les personnes handicapées forment un autre groupe difficile à atteindre. Le Census Bureau se fonde sur des enquêtes telles que l'enquête sur les collectivités (American Community Survey) et l'enquête sur le revenu et la participation aux programmes (Survey of Income and Program Participation) qui fournissent des informations sur les personnes handicapées. Une analyse utilisant les estimations sur un an de l'American Community Survey a permis d'étudier les taux d'invalidité des enfants et les coûts monétaires et non monétaires associés à leur prise en charge. Le Census Bureau continue d'identifier et d'obtenir des données administratives pour ce groupe de population. La Social Security Administration gère le supplément de revenu garanti, qui est un programme d'aide pécuniaire soumis à des conditions de ressources. L'enquête sur le revenu et la participation aux programmes utilise ces informations aux fins de vérification et d'imputation. Les bénéficiaires du programme peuvent également bénéficier d'une couverture d'assurance maladie dans la plupart des États. En se procurant les données administratives du supplément de revenu garanti et de la couverture d'assurance médicale, le Census Bureau a entrevu une solution potentielle pour combiner ces données administratives avec ses enquêtes de façon à produire des analyses statistiques plus approfondies.
- Le Census Bureau a engagé un effort de modernisation qui associe la science des données aux méthodes d'enquête traditionnelles pour diversifier ses produits et placer les données au centre de l'approche suivie. L'un des aspects de cette démarche est la création du programme Frames, qui correspond à un éventail de plus en plus large d'ensembles de données reliés. Si le Census Bureau en compte déjà un certain nombre sous la forme d'entités autonomes, l'approche Frames permettra de regrouper ces ensembles de données, parmi d'autres, et offrira un moyen simple et efficace de les coupler à des fins à la fois prévisibles (pour constituer telle ou telle base d'échantillonnage, par exemple) et imprévues (pour pouvoir répondre à une question sur les emplois et les taux de vaccination contre la COVID-19, etc.). Ces ensembles de données couplés, augmentés et constamment actualisés fourniront un moyen plus complet de gérer et de mettre à jour l'inventaire des adresses, des emplois, des entreprises et des personnes et d'autres données corrélées. La centralisation et la possibilité d'établir des liens avec d'autres fichiers contribueront à améliorer l'efficacité, à réduire les chevauchements dans la tenue et la gestion des données et à accroître considérablement la capacité de répondre à des questions essentielles sur la population et l'économie du pays à de multiples niveaux géographiques. Une telle solution pourrait aider le Census Bureau à continuer de recenser et de mettre à profit des sources de données administratives pour combler les lacunes concernant les enfants, les personnes handicapées et d'autres groupes de population difficiles à atteindre.

V. Principales constatations

- 69. Le fait que des groupes de population soient difficiles à atteindre dans le contexte des données administratives peut être interprété de différentes manières et leur définition dépend de la situation du pays. L'une des interprétations tient au caractère incomplet des registres administratifs ou des bases de données administratives corrélées, certains groupes, tels que les enfants ou les personnes âgées, étant de ce fait difficiles à atteindre et, partant, à décrire par des données. Il en est question dans les cas du Canada et des États-Unis. Selon une autre interprétation, il s'agit de sélectionner certains groupes, tels que les sans-abri, les immigrés clandestins ou les autochtones, puis de tenter de les faire apparaître dans les registres administratifs pour obtenir des informations plus complètes. Les personnes handicapées sont un autre groupe souvent mentionné dans ce contexte, mais l'état des données les concernant ne permet pas une analyse plus large. Une telle approche est décrite par le Danemark, les États-Unis, l'Italie et la Nouvelle-Zélande.
- 70. Il n'a pas été constaté que des mécanismes bien établis permettaient d'identifier les groupes de population *difficiles à atteindre*, notamment les sans-abri ou les immigrés clandestins, dans les registres administratifs. Les exemples présentés montrent que l'identification des groupes *difficiles à atteindre* dépend du contexte national mais s'appuie souvent sur des enquêtes qui tentent de saisir des attributs tels que le numéro d'identification, l'adresse ou la date de naissance qui aident à reconnaître ces groupes dans les registres administratifs.
- 71. Ainsi qu'il ressort de ces exemples, les données administratives sont dans la pratique complétées par des enquêtes statistiques ou des données de recensement en vue d'obtenir des informations un peu plus globales sur les groupes *difficiles à atteindre*. Cela permet d'étoffer les informations disponibles et de relier les données entre elles.
- 72. Les pays considérés ont pris différentes initiatives pour améliorer l'accès aux groupes de population *difficiles à atteindre*. Pour autant que l'objectif consistant à faciliter l'accès à ces groupes dans les registres administratifs de manière transversale suscite un intérêt, une plus ample coopération s'avère nécessaire dans ce domaine.

VI. Recommandations

- 73. De nombreuses initiatives ont été engagées dans différents pays pour faire apparaître dans les données administratives les groupes de population *difficiles à atteindre*, concept dont la compréhension peut varier d'un pays à l'autre. Certains pays les considèrent comme des catégories de population spécifiques, telles que les sans-abri, d'autres comme des composantes particulières de la population qui ne sont pas entièrement couvertes par les données administratives.
- 74. Les divergences quant à la notion de population *difficile à atteindre* pourraient donner à penser que la question devraient faire l'objet de plus amples travaux de nature transversale, afin que les pays puissent bénéficier d'un cadre commun de concepts et de lignes directrices générales dans ce domaine.
- 75. L'une des étapes du développement des travaux relatifs aux groupes de population difficiles à atteindre dans les registres administratifs pourrait consister à recenser les questions transversales auxquelles les pays sont confrontés puis, en fonction des résultats, à circonscrire le sujet pour l'appliquer à des groupes précis, tels que les sans-abri, les immigrés clandestins ou d'autres groupes.
- 76. La délimitation des groupes de population *difficiles* à *atteindre* pourrait servir de point de départ à une étude plus approfondie visant à déterminer s'il existe un terrain d'entente pour une analyse des moyens de mieux identifier ces groupes dans les registres administratifs.
- 77. En l'absence d'un tel terrain d'entente, il pourrait être envisagé de dresser une liste des meilleures pratiques qui, dans les registres administratifs des différents pays, permettent d'accéder aux groupes *difficiles à atteindre*.

- 78. L'un des facteurs susceptibles de favoriser un travail plus large sur l'utilisation des données relatives aux groupes *difficiles à atteindre* pourrait être une demande clairement définie, par exemple sous la forme d'un cadre législatif ou de recommandations formulées au niveau international.
- 79. Compte tenu de ce qui précède, il est recommandé de créer une équipe spéciale chargée de définir les besoins de coopération future concernant les groupes *difficiles à atteindre* dans le contexte des données administratives.

VII. Délibérations du Bureau de la Conférence des statisticiens européens

- 80. Le Bureau a procédé à un examen approfondi de la question des groupes difficiles à atteindre dans le contexte des sources administratives à sa réunion de février 2023. Les observations ci-après ont été faites au cours du débat :
- a) Le document donne une bonne vue d'ensemble des questions liées à ce sujet fort important ;
- b) Un cadre conceptuel pourrait aider à déterminer ce que nous savons et ce que nous ne savons pas sur les groupes difficiles à atteindre. Pour certains groupes, il n'y a que des informations très succinctes (nombre total de personnes appartenant au groupe, par exemple), voire pas d'information du tout ;
- c) Il pourrait être utile de travailler sur des principes de base et de se concentrer sur certains groupes concernés par les politiques, tels que les personnes handicapées, les migrants, les minorités ethniques, les sans-abri, les enfants et les personnes âgées. Il importe de considérer la question sur le plan des politiques. Qui sont ces groupes ? Quelles sont leurs caractéristiques ? Comment sont-ils intégrés ?
- d) Les sources administratives ne font pas apparaître certaines personnes parce qu'elles ne répondent pas aux critères d'inclusion. Il convient de prêter attention à la conception des sources et au fait qu'elles peuvent être imparfaites. L'expérience des pays dotés de systèmes statistiques fondés sur les registres sera particulièrement importante. Toutes les sources devraient être prises en considération, et pas seulement les sources administratives. L'intégration des données est essentielle;
- e) Certaines personnes ne souhaitent pas être incluses et feront tout leur possible pour ne pas être identifiées. Il faudrait concevoir des méthodes particulières pour détecter ces personnes et recourir à de multiples sources.
- 81. En conclusion, le Bureau de la Conférence a souscrit à l'idée de poursuivre les travaux dans ce domaine et approuvé la création d'une nouvelle équipe spéciale, comme le recommande le document. Le Danemark présidera cette équipe spéciale. Outre les pays qui ont déjà contribué au document (Canada, États-Unis, Italie et Nouvelle-Zélande), l'Irlande, le Mexique, le Royaume-Uni, Eurostat, l'OCDE et la Division de statistique de l'ONU ont fait part de leur désir de participer aux travaux de l'équipe spéciale. Le secrétariat établira le mandat de la nouvelle équipe spéciale, pour examen par le Bureau à la réunion d'octobre 2023.

VIII. Remerciements

82. Statistics Denmark souhaite remercier tout particulièrement : Sylvain Delisle, Karelyn Davis et James Falconer, de Statistique Canada ; Donatella Zindato de l'Istituto Nazionale di Statistica (ISTAT) ; Christine Bycroft, Pubudu Senanayake, Nathaniel Matheson-Dunning, Hannes Diener de Statistics New Zealand ; et Vincent T Mule Jr et Andrew D Keller du Census Bureau des États-Unis.

IX. Bibliographie

Bowlby G., Morel J.: Utilisation des données administratives pour le recensement canadien, vingt-quatrième réunion du Groupe d'experts des recensements de la population et des habitations de la CEE, 21-23 septembre 2022 :

https://unece.org/sites/default/files/2022-07/ECE_CES_GE.41_2022_7-2210829F.pdf

Statistique Canada (2021). Guide du recensement de la population. Annexe 1.7 – Utilisation de données administratives pour imputer les ménages non répondants des régions où les taux de réponse sont faibles.

https://www12.statcan.gc.ca/census-recensement/2021/ref/98-304/2021001/app-ann1-7-fra.cfm

Bernardini A., Chieppa A., Cibella N., Solari F, Zindato D.: Évolution du recensement permanent de la population en Italie. Enseignements à retenir du premier cycle et de la conception du recensement permanent au-delà de 2021, vingt-quatrième réunion du Groupe d'experts des recensements de la population et des habitations de la CEE, 21-23 septembre 2022:

https://unece.org/sites/default/files/2022-07/ECE_CES_GE.41_2022_4-2209925F.pdf.

Falorsi, S.: The Italian experience on the Population and Housing Census: the Master Sample, dix-neuvième réunion du Groupe d'experts des recensements de la population et des habitations, 4-6 octobre 2017: https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2017/Meeting-Geneva-Oct/Day2_1130_Italy_falorsi_presentation.ppt__1_pdf.

Gallo, G., Zindato, D.: Annexe H. Italy case study, in UNECE (2018), Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses, https://unece.org/guidelines-use-registers-and-administrative-data-population-and-housing-censuses-0.

Gallo, G., Zindato, D.: Italy: The combined use of survey and register data for the Italian Permanent Population Census count in UNECE (2021), Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses, https://unece.org/statistics/publications/CensusAdminQuality

Stats NZ (2022a). Experimental administrative population census: Data sources, methods, and quality (second iteration). Tiré de https://www.stats.govt.nz/

Milne, B. J., D'Souza, S., Andersen, S. H., et Richmond-Rakerd, L. S. (2022). Use of population-level administrative data in developmental science. Annual Review of Developmental Psychology, 4. https://doi.org/10.1146/annurev-devpsych-120920-023709

Stats NZ, 2022b Information for Māori about the experimental administrative population census.

Amore K. (2016). Severe housing deprivation in Aotearoa/New Zealand: 2001-2013. He Kainga Oranga/Housing & Health Research Programme, Université d'Otago, Wellington.

Amore, K. Viggers H, Howden-Chapman P (2021). Severe housing deprivation in Aotearoa New Zealand, 2018: June 2021 Update. He Kāinga Oranga / Housing & Health Research Programme Department of Public Health Université d'Otago, Wellington. Tiré de https://www.hud.govt.nz/

Arezoo Malihi, Annie Chiang, Jay Marlowe, Barry Milne, Dan Exeter (2022) Navigating data in the IDI to unlock refugee settlement trajectories – le lien renvoie à un exposé, qui pourrait fait l'objet d'un article publié prochainement.

SWA's guide to using Integrated Data to understand mental health and addiction conditions.

Fernandez, L., Shattuck, R., et Noon, J. (2018), « The Use of Administrative Records and the American Community Survey to Study the Characteristics of Undercounted Young Children in the 2010 Census »

Center for Administrative Records Research and Application Working Paper Series #2018 - 05

https://www.census.gov/content/dam/Census/library/working-papers/2018/adrm/carra-wp-2018-05.pdf

Khubba, S., Heim, K. et Hong, J (2022), « National Census Coverage Estimates for People in the United States by Demographic Characteristics » 2020 Post-Enumeration Survey Estimation Report, PES20-G-01.

https://www2.census.gov/programs-surveys/decennial/coverage-measurement/pes/national-census-coverage-estimates-by-demographic-characteristics.pdf

Velkoff, V. et Hartley, C. (2022). Moving Forward With the U.S. Census Bureau's Annual Population Estimates Post-2020. Harvard Data Science Review, 4(4). https://doi.org/10.1162/99608f92.4ba61ca4

Young, N. (2021), « Childhood Disability in the United States: 2019, » ACSBR-006, American Community Survey Briefs, U.S. Census Bureau, Washington, DC, 2021.

https://www.census.gov/library/publications/2021/acs/acsbr-006.html

Thieme, M. (2022), « Technology Transformation at the Census Bureau : Building a Modern, Data-Centric Ecosystem »

https://www.census.gov/newsroom/blogs/research-matters/2022/10/technology-transformation.html