



Commission économique pour l'Europe**Conférence des statisticiens européens****Soixante-huitième réunion plénière**

Genève, 22-24 juin 2020

Point 4 g) de l'ordre du jour provisoire

**Rapports, directives et recommandations élaborés
sous les auspices de la Conférence : Recensements
de la population et du logement au-delà de 2020****Nouvelles perspectives de recensement au-delà de 2020****Note du Groupe directeur sur les recensements de la population et
des habitations***Résumé*

Le présent document contient une analyse approfondie de l'avenir des recensements de la population et des habitations après le cycle de 2020, qui servira de base aux recommandations concernant le cycle de recensements de 2030.

Le document a été établi par le Groupe directeur sur les recensements de la population et des habitations, dont les membres viennent d'Arménie, du Canada (qui assurera la présidence jusqu'en mars 2020), des États-Unis d'Amérique, de France, de Géorgie, d'Italie, des Pays-Bas (qui assurera la présidence à partir de mars 2020), de Pologne, du Royaume-Uni, de Slovaquie et de Suisse. Des représentants du Comité inter-États de statistique de la Communauté d'États indépendants, de l'Association européenne de libre-échange, d'Eurostat, du Fonds des Nations Unies pour la population et de la Division de statistiques de l'ONU ont également participé aux travaux.

Le document a été envoyé à tous les membres de la Conférence des statisticiens européens (CSE) pour consultation. Si la consultation donne des résultats positifs, la CSE sera invitée à approuver le document à sa réunion plénière.



I. Introduction

1. Au cours des dernières décennies, les statisticiens n'ont cessé de faire évoluer les méthodes employées pour recenser la population et les habitations afin de les adapter à une société et à des besoins en information qui changent, de tirer parti des possibilités offertes par des sources de données nouvelles ou différentes telles que les registres et les dossiers administratifs et d'exploiter pleinement le potentiel des innovations dans les technologies habilitantes.
2. Cette évolution est particulièrement marquée dans de nombreux pays de la CEE. Il ressort clairement des informations disponibles sur les plans élaborés par les pays de la CEE pour le cycle de recensements de 2020 qu'une grande partie des pays de la région prévoient d'introduire d'importantes innovations méthodologiques et technologiques, ou étudient la possibilité de le faire, et adoptent d'autres approches fondées sur l'utilisation accrue des registres et des données administratives. De nombreux pays ont pour objectif final de passer, pour le cycle de 2020 ou le suivant, à un recensement entièrement fondé sur des registres. D'autres élargissent l'application des approches traditionnelles pour étendre la pratique de l'autodénombrement, en particulier par Internet, et pour tirer le meilleur parti des divers processus de recensement grâce aux nouvelles technologies.
3. Les changements évoqués plus haut ne s'opèrent pas sans difficultés. Les experts de la CEE et d'Eurostat ont examiné ces difficultés lors d'une réunion que le Groupe d'experts des recensements de la population et des habitations a tenue en 2017 (Genève, 4-6 octobre), notamment dans le cadre d'une table ronde sur l'avenir des recensements après 2020. Des travaux sur ce sujet ont également été menés au niveau de l'Union européenne (ECE/CES/GE.41/2017/21). Au cours du débat, les experts ont unanimement admis que la plupart des pays ne pourraient plus se contenter de fournir les données à intervalle de dix ans et qu'ils devraient les actualiser plus fréquemment. Ils ont également fait observer que le débat sur l'avenir des recensements ne devrait pas porter uniquement sur les améliorations progressives fondées sur le passé et que de nombreux pays devraient peut-être adopter une orientation entièrement nouvelle.
4. Le présent document a été établi à la demande du Bureau de la Conférence des statisticiens européens (CSE) par le Groupe directeur sur les recensements de la population et des habitations, représentant le Groupe d'experts des recensements de la population et des habitations. On y trouvera une analyse des nombreux enjeux et problèmes liés à l'avenir des recensements de la population et des habitations après le cycle de 2020. Cette analyse servira de base aux recommandations de la CSE pour le cycle de recensements de 2030¹.
5. Dans le cadre de l'analyse, les questions suivantes ont été abordées :
 - a) L'éventuelle nécessité d'utiliser différentes mesures de la population ;
 - b) Les méthodes innovantes en cours de développement et les approches multimodales fondées sur l'utilisation d'enquêtes, de données administratives et de nouvelles sources de données ;
 - c) L'importance croissante des informations géocodées ;
 - d) Le passage d'un recensement décennal à des collectes de données plus fréquentes, notamment annuelles ;
 - e) L'utilisation de mégadonnées et le risque de se retrouver en concurrence avec des producteurs de mégadonnées ;
 - f) La nécessité de prendre en compte les changements sociétaux tout en préservant la comparabilité dans le temps et entre les pays : les notions de « ménage » et de « lieu de résidence habituel », entre autres, resteront-elles pertinentes ?

¹ On trouvera les Recommandations de la CSE pour les recensements de la population et des habitations de 2020 à l'adresse suivante : <http://www.unecce.org/index.php?id=41398>.

6. L'analyse effectuée a permis de repérer un certain nombre de questions qui devraient être prises en compte dans le cadre de la révision des recommandations pour les recensements de la population et des habitations. On trouvera ci-après un aperçu de ces questions.

A. Aperçu des questions à prendre en compte dans le cadre de la révision des recommandations de la CSE pour les recensements de la population et des habitations de 2030

- a) Réviser les caractéristiques essentielles d'un recensement, telles que la simultanéité, afin de mieux répondre aux impératifs des nouvelles approches, qui peuvent s'appuyer sur des sources multiples ayant des dates de référence différentes ;
- b) Décrire de manière plus détaillée les difficultés liées à l'utilisation de sources multiples aux fins de la production des résultats du recensement, en particulier dans le cas des approches mixtes ;
- c) Élaborer des critères pour définir la résidence secondaire ;
- d) Élaborer des critères pour définir et isoler les logements qui sont utilisés régulièrement, mais ne sont pas occupés en permanence ;
- e) Améliorer la définition et la compréhension de la migration pendulaire, y compris les méthodes permettant de mesurer la couverture de ce phénomène dans le cadre de l'utilisation des registres établis à partir de sources administratives ;
- f) Mettre davantage l'accent sur la mesure de la qualité des résultats des recensements ;
- g) Mettre au point une approche cohérente et systématique permettant de combiner les données statistiques et géospatiales. Une telle approche pourrait s'appuyer sur le cadre statistico-géospatial mondial sur lequel la communauté internationale s'est accordée ;
- h) Formuler des recommandations sur l'utilisation de méthodes d'estimation aux fins de la création de variables de recensement et sur d'autres questions connexes ;
- i) Fournir des indications sur les différents aspects de l'utilisation de nouvelles sources de données non structurées à des fins de recensement ;
- j) Donner des orientations sur la manière de trouver un compromis entre les différents aspects de la qualité et la manière dont ces aspects changent à mesure que les besoins des utilisateurs évoluent.

II. Situation dans la région de la CEE : résumé des résultats de l'enquête

7. Dans le cadre des présents travaux, le Groupe directeur a mené, en septembre 2018, une enquête sur les futurs recensements au-delà du cycle de 2020 auprès des pays de la CEE. L'enquête avait pour objectif d'examiner ce que les pays envisageaient de faire pour le prochain recensement après le cycle de 2020, qui aura lieu en 2029, 2030 ou 2031 dans les pays procédant à un recensement décennal, ou en 2025 ou une autre année dans les pays où le recensement sera réalisé chaque année ou complété par des mises à jour annuelles. Par souci de simplicité, on désignera dans le présent document tous ces cas de figure par l'expression « cycle de 2030 ».

8. L'enquête a été réalisée par le secrétariat de la CEE en septembre 2018 au moyen d'un bref questionnaire en ligne. Quarante-sept pays de la CEE y ont répondu, ce qui représente un très bon taux de réponse de 90 %. Les informations recueillies dans le cadre de l'enquête et présentées dans la présente section doivent être considérées comme non officielles et provisoires puisque le cycle de 2030 est encore très éloigné.

9. Parmi les pays qui ont répondu, 13 (28 %) mènent un recensement traditionnel dans le cadre du cycle de 2020, 17 (37 %) un recensement mixte, mélange entre un dénombrement direct et l'utilisation de sources administratives, et 13 (28 %) un recensement entièrement fondé sur des registres. Le pourcentage de pays de la CEE qui procèdent à un recensement traditionnel a considérablement diminué, alors que cette méthode était de loin la plus utilisée dans le cadre du cycle de 2000 (83 % des pays) et de celui de 2010 (64 % des pays).

10. Pour le cycle de 2030, 19 pays (41 %) ont indiqué qu'ils envisageaient, même à titre provisoire, de modifier la méthode utilisée dans le cadre du cycle de 2020, et sept pays (15 %) ont répondu « peut-être » à la même question.

11. Sur les 26 pays qui envisagent d'apporter des modifications, 11 procèdent à un recensement traditionnel dans le cadre du cycle de 2020 (dont cinq prévoient de passer à un recensement mixte et deux à un recensement fondé sur des registres), treize pays mènent un recensement mixte (huit d'entre eux prévoient de passer à un recensement entièrement fondé sur des registres et cinq de conserver la méthode employée) et deux pays comptent effectuer un recensement fondé sur des registres dans le cadre des cycles de 2020 et 2030.

12. Sur les 13 pays qui n'envisagent actuellement pas de changer de méthode entre les cycles de 2020 et 2030 (28 % de tous les pays ayant répondu), 10 procèdent à un recensement fondé sur des registres dans le cadre du cycle de 2020, et trois à un recensement mixte.

13. Seuls sept pays (15 %) ont répondu qu'ils ne savaient pas s'ils allaient apporter des changements méthodologiques entre les cycles de 2020 et 2030, ce qui montre qu'une grande majorité des pays ont déjà une idée, qui peut encore évoluer, des méthodes qu'ils utiliseront dans le cadre du cycle de recensement de 2030.

14. Le tableau 1 montre la répartition des pays par méthode de recensement utilisée dans le cadre du cycle de 2020 et par méthode de recensement qu'il est prévu d'utiliser (en principe) dans le cadre du cycle de 2030. Les 13 pays qui procèdent à un recensement fondé sur des registres dans le cadre du cycle de 2020 prévoient, comme on pouvait s'y attendre, d'utiliser la même méthode dans le cadre du cycle de 2030, bien que deux d'entre eux envisagent d'apporter des changements méthodologiques, comme indiqué ci-dessus. Près de la moitié des pays qui procèdent à un recensement mixte dans le cadre du cycle de 2020 (8 sur 17) prévoient de passer à un recensement entièrement fondé sur des registres dans le cadre du cycle de 2030. Enfin, seuls deux des 13 pays qui mènent un recensement traditionnel dans le cadre du cycle de 2020 comptent poursuivre avec cette méthode dans le cadre du cycle de 2030. Parmi les autres pays, cinq prévoient de passer à un recensement mixte, deux à un recensement entièrement fondé sur des registres, et quatre pays ne se sont pas encore décidés.

15. Il a également été demandé aux pays s'ils envisageaient, en vue du cycle de 2030, d'apporter d'autres changements importants (autres que la principale méthode de recensement utilisée) ou des innovations. Seuls quatre pays, qui procèdent déjà à un recensement fondé sur des registres, ont indiqué qu'ils ne prévoyaient pas d'apporter des changements ou des innovations notables ; neuf pays ont répondu « peut-être », et 33 pays (72 %) ont indiqué qu'ils prévoyaient d'apporter un ou plusieurs changements ou innovations en vue du cycle de 2030.

Tableau 1
Méthodes de recensement qu'il est prévu d'utiliser dans le cadre du cycle de 2030, au regard des méthodes utilisées dans le cadre du cycle de 2020

Méthode utilisée dans le cadre du cycle de 2020	Méthode de recensement qu'il est prévu d'utiliser dans le cadre du cycle de 2030					Total
	Méthode traditionnelle	Méthode mixte	Méthode fondée sur des registres	Recensement en continu	Autre / ne sait pas	
Méthode traditionnelle	2	5	2		4	13
Méthode mixte		8	8		1	17
Méthode fondée sur des registres			13			13
Recensement en continu					1	1
Autre / ne sait pas		1			1	2
Total	2	14	23	-	7	46

16. Les changements auxquels les pays envisagent de procéder sont, par ordre décroissant, les suivants : passage d'un recensement décennal à un recensement plus fréquent (21 pays, soit 46 % du total des pays ayant répondu au questionnaire), ventilation géographique plus détaillée des résultats du recensement (18 pays), utilisation accrue d'informations géocodées aux fins de la collecte de données (15 pays), meilleure harmonisation entre le recensement et d'autres programmes statistiques (13 pays) et utilisation des mégadonnées ou d'autres sources (11 pays) (tableau 2).

Tableau 2
Changements ou innovations qu'il est prévu d'apporter en vue du cycle de 2030

Changement / innovation	Méthode de recensement qu'il est prévu d'utiliser dans le cadre du cycle de 2030					Total (46 pays)
	Méthode traditionnelle (2 pays)	Méthode mixte (14 pays)	Méthode fondée sur des registres (23 pays)	Autre / ne sait pas (7 pays)		
Passage d'un recensement décennal à un recensement plus fréquent		5	15	1		21
Ventilation géographique plus détaillée des résultats du recensement		6	11	1		18
Utilisation accrue d'informations géocodées aux fins de la collecte de données	1	7	5	2		15
Meilleure harmonisation entre le recensement et d'autres programmes statistiques		4	9			13
Utilisation de mégadonnées ou d'autres sources		6	5			11
Néant			4			4

17. Si l'on examine les changements et les innovations envisagés par méthode de recensement utilisée, les pays procédant à un recensement fondé sur des registres semblent

plus enclins à augmenter la fréquence des recensements que les pays menant un recensement mixte. Néanmoins, ces derniers semblent plus susceptibles d'utiliser davantage d'informations géocodées aux fins de la collecte de données. Ces deux résultats étaient prévisibles, car il est plus facile de réaliser fréquemment un recensement fondé sur des registres qu'un recensement mixte, tandis que l'efficacité de ce dernier peut facilement être améliorée par l'utilisation d'informations géocodées.

18. Il est intéressant de noter que plusieurs pays envisagent, pour le cycle de 2030, d'utiliser des mégadonnées dans le cadre d'un recensement mixte (six pays) ou d'un recensement entièrement fondé sur des registres (cinq pays).

19. Il a été demandé aux pays s'ils prévoient, en vue du cycle de 2030, la nécessité d'apporter des changements à des notions clefs. Seuls 14 pays (30 % du total) ont répondu par l'affirmative : huit pays qui prévoient de procéder à un recensement entièrement fondé sur des registres et six pays qui prévoient de mener un recensement mixte dans le cadre du cycle de 2030 (tableau 3).

20. La notion que la plupart des pays (11) souhaitent modifier est celle du « lieu de résidence habituel ». Bon nombre de ces pays (dont le Canada, Chypre, l'Espagne, Israël, la Russie et la Tchéquie) estiment qu'il peut être nécessaire d'apporter des modifications lorsque des données administratives sont utilisées dans le cadre d'une approche mixte ou d'une approche entièrement fondée sur des registres. Le Canada souhaite aussi élargir la notion de population recensée en raison des changements intervenus dans la société et dans les besoins des utilisateurs.

21. De l'avis de 6 pays, il faudrait éventuellement modifier la notion de « ménage ». Plusieurs pays (dont Chypre, la Hongrie, la Serbie et la Tchéquie) ont fait observer que cette notion était difficile à appliquer lorsque des données de registres étaient utilisées et qu'il fallait donc employer la notion de ménage-logement.

22. Selon quatre pays, il pourrait être nécessaire d'apporter des changements à certaines notions clefs liées aux caractéristiques essentielles. En particulier, l'Espagne et l'Islande ont noté que la notion de situation au regard de l'activité du moment, qui est fondée sur les définitions que donne l'Organisation internationale du Travail de l'emploi et du chômage, pose problème lorsque des données de registres sont utilisées.

Tableau 3

Notions clefs qu'il faudra éventuellement modifier en vue du cycle de 2030

Notions clefs qu'il faudra éventuellement modifier	Méthode de recensement prévue pour le cycle de 2030				Total (46 pays)
	Méthode traditionnelle (2 pays)	Méthode mixte (14 pays)	Méthode fondée sur des registres (23 pays)	Autre / ne sait pas (7 pays)	
Définition d'un ménage		1	5		6
Lieu de résidence habituel		5	6		11
Caractéristiques essentielles		1	3		4
Autre		2			2

23. Plusieurs raisons poussent les pays à envisager d'apporter des changements à leur méthode de recensement. Pour le cycle de 2030, il s'agit principalement des changements technologiques (11 pays), des changements sociétaux (8 pays), des besoins projetés des utilisateurs (5 pays) et des besoins exprimés par les utilisateurs (4 pays).

24. Enfin, il a été demandé aux pays si les Recommandations de la CSE pour les recensements de 2020 correspondaient à ce qu'ils envisageaient de faire pour le cycle de 2030. La grande majorité des pays (35) ont répondu par l'affirmative. Seul le Canada a répondu « non », indiquant notamment que les définitions de la population devaient être révisées (observation partagée par l'Allemagne) et estimant que les caractéristiques essentielles devaient être précisées dans le contexte des approches mixtes. La France a fait

observer que l'une des caractéristiques essentielles d'un recensement était que, dans la mesure du possible, chaque individu devait être enregistré en un seul lieu.

25. En résumé, l'enquête a permis de tirer les principales conclusions suivantes :

- a) La tendance à l'abandon du recensement traditionnel au profit du recensement mixte et du recensement entièrement fondé sur des registres devrait se poursuivre dans le cadre du cycle de 2030 ;
- b) Pour le cycle de 2030, plusieurs pays prévoient de passer d'un recensement traditionnel à un recensement mixte ou à un recensement entièrement fondé sur des registres (au moins sept), ou d'un recensement mixte à un recensement fondé sur des registres (au moins huit) ;
- c) Environ 70 % des pays envisagent, en vue du cycle de 2030, d'apporter des changements importants (autres que la principale méthode de recensement utilisée) ou des innovations, notamment le passage d'un recensement décennal à un recensement plus fréquent (21 pays), une ventilation géographique plus détaillée des résultats du recensement (18 pays) et une utilisation accrue d'informations géocodées aux fins de la collecte de données (15 pays) ;
- d) Au moins 11 pays envisageront d'utiliser, pour le cycle de 2030, des mégadonnées dans le cadre d'un recensement mixte ou fondé sur des registres ;
- e) Il a été jugé qu'il fallait modifier des notions clefs du recensement telles que le lieu de résidence habituel (11 pays), le ménage (6 pays) et la situation au regard de l'activité du moment. Pour la plupart des pays, la nécessité d'apporter des changements découle de l'utilisation de données provenant de registres ;
- f) Les principaux motifs de changement invoqués sont les technologies (11 pays), les mutations sociétales (8 pays), les besoins projetés des utilisateurs (5 pays) et les besoins exprimés par les utilisateurs (4 pays) ;
- g) La grande majorité des pays ont indiqué que les Recommandations de la CSE pour les recensements de 2020 étaient conformes à ce qu'ils envisageaient de faire pour le cycle de 2030, bien qu'il convenait de modifier certaines parties et notions clefs.

III. Évolution des méthodes de recensement de la population : faut-il changer d'orientation ?

26. Comme le montrent les résultats de l'enquête, de nombreux pays étudient ou adoptent de nouvelles approches de recensement et de dénombrement de la population dans le cadre des cycles de 2020 ou de 2030. À mesure que de nouvelles approches sont étudiées et utilisées, certaines difficultés à suivre les recommandations existantes pour les recensements de la population et des habitations ressortent, en particulier dans le cas des méthodes fondées sur des registres ou des sources administratives. Elles sont exacerbées par l'évolution de la société et par les changements qui en résultent dans les exigences des utilisateurs de données, ce qui fait peser une pression supplémentaire sur les statisticiens, qui doivent garantir la pertinence des programmes nationaux de recensement.

27. Certains pays jouissent d'un système de registres solide, dont ils extraient depuis un certain temps déjà les résultats de leurs recensements. C'est le cas notamment de la Suède, de la Finlande et des Pays-Bas. D'autres pays tels que l'Italie et Israël sont en train de construire de tels systèmes et de les utiliser pour produire leurs données de recensement.

28. Certains pays ne disposent d'aucun registre universel. Cela ne les empêche toutefois pas de produire des dénombrements de la population à partir de sources administratives, en particulier dans le cadre de l'approche mixte. À cette fin, certains pays tels que le Canada et le Royaume-Uni construisent des registres statistiques. Ces registres sont créés par la combinaison de plusieurs sources de données administratives, l'objectif étant de représenter un univers de personnes ou d'habitations. L'expression « registres statistiques » est utilisée pour indiquer que les registres sont construits à partir de diverses sources de données

administratives et ne serviront qu'à appuyer le mandat d'un organisme national de statistique consistant à fournir des produits statistiques dépersonnalisés et protégés.

29. Lorsqu'un pays décide de passer d'une méthode de recensement traditionnelle à une méthode de recensement fondée sur des registres (ou mixte), cette démarche s'inscrit généralement dans le cadre d'une décision stratégique de l'organisme national de statistique de construire ou d'adopter un système statistique fondé sur des registres qui restera au cœur de la production statistique nationale. Ce changement d'orientation implique souvent de repenser complètement le processus global de production statistique, notamment la fréquence des diverses activités statistiques et le bon moment pour les mener. De nombreux pays estiment qu'en passant d'une méthode de recensement traditionnelle à une méthode de recensement fondée sur des registres, il devient possible de produire des données à une fréquence plus élevée que dans le cas d'un recensement traditionnel (généralement réalisé à intervalle de dix ans) et de publier les résultats plus rapidement après la date de référence. Pour la plupart d'entre eux, l'objectif explicite est de produire chaque année ne serait-ce qu'un ensemble restreint de statistiques de base sur la population et les habitations. En outre, lorsque l'on choisit d'utiliser des registres ou des données administratives aux fins de recensement, il faut souvent repenser la population servant de base et les notions clés employées telles que le lieu de résidence et le ménage privé. Pour ce faire, il convient de tenir compte des caractéristiques des informations disponibles dans les registres, mais aussi de l'évolution de la société et des besoins en information des utilisateurs.

IV. Population servant de base : difficultés liées aux nouvelles méthodes fondées sur des sources administratives

A. Cohérence dans la production de données provenant de sources multiples

30. Les renseignements sur les individus et les habitations obtenus au moyen d'un recensement doivent généralement se rapporter à une période de référence unique et bien définie. Cette exigence concorde avec l'une des caractéristiques essentielles du recensement, à savoir la simultanéité. Lorsque diverses sources sont utilisées aux fins du dénombrement, il arrive souvent que les dates ou périodes de référence divergent. Par exemple, la période de référence varie souvent d'une source administrative à l'autre et peut ne pas correspondre à la date de référence de la partie du recensement provenant du dénombrement direct dans le cas d'une approche mixte. Il peut donc être difficile de choisir la meilleure source à utiliser pour le dénombrement. Une source pourrait avoir une date de référence similaire à la date de dénombrement, mais présenter des lacunes importantes. À mesure que de plus en plus de pays adoptent une approche partiellement ou entièrement fondée sur des registres ou une approche fondée sur des données administratives, il peut être souhaitable de réviser les caractéristiques essentielles du recensement dans la prochaine version des recommandations afin de mieux répondre aux impératifs des nouvelles approches.

B. Cohérence des notions relatives au lieu de résidence

31. Pour réaliser un dénombrement direct de la population servant de base, il est possible d'utiliser une approche *de jure* ou une approche *de facto*. La notion de lieu de résidence est alors définie sur la base de la notion de dénombrement. Les définitions du lieu de résidence peuvent varier selon si la population a été recensée par dénombrement direct ou à partir de registres ou de sources administratives. Si tel est le cas, ces différences peuvent être difficiles à concilier lors du dénombrement de la population servant de base dans le cadre d'une approche mixte. La tâche pourrait être encore plus complexe si plusieurs dossiers administratifs sont utilisés. Il pourrait être problématique de déterminer le lieu de résidence sur la base des notions existantes applicables aux groupes de population mobiles ou aux personnes ayant plus d'un lieu de résidence. Il peut également être difficile de savoir si la notion de « lieu de résidence habituel » employée dans les sources de

données administratives correspond à celle utilisée pour le dénombrement direct dans le cadre des méthodes de recensement traditionnelles.

32. Les difficultés liées à l'utilisation de sources de données administratives à des fins de recensement peuvent être exacerbées par un manque de stabilité dans la couverture de certaines de ces sources. Les notions employées dans les sources de données administratives pourraient évoluer rapidement, ce qui limiterait la possibilité d'évaluer leur incidence. Certains changements pourraient même passer inaperçus. L'utilisation de données administratives dans le cadre d'un recensement peut donc créer des failles, car des changements peuvent survenir dans la couverture, les dates de référence, les variables disponibles ou d'autres caractéristiques. De tels changements pourraient nuire à la comparabilité entre les différents cycles de recensement, qui est une exigence essentielle des utilisateurs de ces données.

33. Les futures recommandations devraient décrire de manière plus détaillée les différentes difficultés liées à l'utilisation de sources multiples dans la production des données de recensement, en particulier dans le cadre des approches mixtes.

C. Ménages ou personnes ayant plusieurs domiciles

34. De plus en plus de personnes ou de ménages ont plus d'un domicile ou peuvent vivre dans plusieurs domiciles au cours de l'année. Comme indiqué dans les Recommandations pour les recensements de la population et des habitations, l'unité de dénombrement utilisée dans le recensement doit être bien définie. En règle générale, on se fonde pour ce faire sur l'enregistrement de la résidence permanente, bien que celle-ci puisse être de plus en plus difficile à déterminer ou changer au fil du temps.

35. Pour rendre compte de l'évolution croissante des comportements de la population et pour garantir aux utilisateurs la pertinence des données, il serait souhaitable à l'avenir d'enregistrer, en plus de la résidence permanente, une résidence secondaire si la personne y séjourne régulièrement pendant une période suffisamment longue, par exemple plus de trois mois par an. Ces critères devraient être inclus dans les futures recommandations. Dans de nombreux recensements nationaux, un logement qui n'est pas une résidence permanente est considéré comme un logement inhabité ou vacant (c'est-à-dire inoccupé par les résidents habituels). Il pourrait être utile d'élaborer des critères permettant de définir et d'isoler les logements qui sont utilisés régulièrement, mais ne sont pas occupés en permanence.

D. Difficultés liées à la mesure de la migration

36. L'un des processus démographiques les plus complexes et les plus répandus est la migration, en particulier la migration internationale (migration entre pays ou franchissement de frontières nationales). Dans de nombreux pays, les informations sur les migrations proviennent de fichiers ou de registres de migration officiels, dans lesquels la migration, en particulier l'émigration, a tendance, dans la plupart des cas, à être sous-estimée. De nombreux émigrants ne signalent pas qu'ils quittent leur pays, et ce pour plusieurs raisons : ils ne savent pas qu'ils sont tenus de le faire ; lorsqu'une telle obligation existe, ils ont oublié de le faire ou considéré que c'était sans importance ; ou ils souhaitent conserver leur statut de résident permanent dans le pays de départ afin de continuer à bénéficier des avantages que leur accorde ce pays. La migration internationale se compose donc généralement de deux parties : les migrations déclarées et les migrations non déclarées. Si la migration internationale est essentiellement un mouvement unidirectionnel, le biais dû aux migrations non déclarées se cumule au fil des ans et finit par fausser l'évaluation du nombre total d'émigrants et de la population totale lorsque des données de registres sont utilisées.

37. Telle était la situation en Estonie avant le recensement réalisé en 2011 lorsque la population dénombrée à partir des données du registre de la population a été surestimée de près de 3 %, tous les départs du pays n'ayant pas été enregistrés. Certains types de migration qui ne coïncident pas avec la définition habituelle de la migration internationale compliquent davantage l'estimation de la migration et le dénombrement de la population.

1. Migration « pendulaire »

38. La migration « pendulaire », qui désigne les allers-retours répétés d'une personne entre deux endroits, n'est pas un phénomène nouveau. En Europe, elle est apparue dès le XIX^e siècle avec le développement du transport ferroviaire. La migration pendulaire internationale, quant à elle, est relativement nouvelle et peu répandue. Elle se produit principalement dans les zones frontalières, en particulier si la frontière sépare deux pays ayant des possibilités d'emploi et des niveaux de vie différents. Dans les statistiques démographiques, les migrations pendulaires sont généralement traitées comme un phénomène relativement localisé et rare, mais dans certains pays, notamment dans ceux de l'Europe de l'Est, elles ont atteint des niveaux assez élevés au cours des dernières décennies.

39. Bien qu'elle existe depuis un certain temps dans la société, la migration pendulaire est une notion statistique relativement nouvelle sur laquelle la communauté internationale ne s'est pas encore accordée, ni dans les recommandations pour les recensements ni dans les recommandations en matière de statistiques des migrations internationales². Toutefois, la définition généralement retenue est celle d'une migration dont la durée n'excède pas une semaine. On considère le plus souvent que le lieu de résidence des migrants pendulaires est celui où habite leur famille (selon la notion de lieu de résidence habituel), tandis que pour les personnes qui s'absentent de leur domicile pour raisons professionnelles pendant plus de 12 mois, mais passent la plupart de leurs jours de congé avec les membres de leur ménage, il s'agit du lieu de résidence de leur ménage.

40. La notion de migration pendulaire peut également être élargie par rapport à la définition donnée ci-dessus si l'on prend en compte les séjours d'une durée supérieure à une semaine effectués dans un autre pays. Il est généralement difficile, y compris pour les personnes concernées, d'appliquer les règles de la statistique démographique pour déterminer le lieu de résidence et le statut des personnes qui se déplacent d'un pays à l'autre en raison de leur travail ou de leurs études.

41. Ce phénomène continuera sans aucun doute à se développer, raison pour laquelle il sera encore plus difficile d'enregistrer les personnes au bon endroit. Pour parvenir à une définition commune et précise de la migration pendulaire, il faudrait que la communauté internationale statue sur les points suivants :

- Quelle doit être la durée du séjour pour qu'une migration soit considérée comme une migration pendulaire ?
- Selon les caractéristiques du travail, la période de migration peut également être de durée variable. Comment prendre en compte les périodes de résidence à durée variable ?
- Si le migrant pendulaire espace ses visites à son lieu de résidence habituel au-delà de la durée généralement admise, il peut être considéré, selon la réglementation nationale, comme un résident permanent du pays dans lequel il travaille.
- La définition de la migration pendulaire est principalement liée à l'emploi. De toute évidence, il faudrait étendre la définition de la migration pendulaire à la migration pour d'autres raisons telles que les études, en particulier pour ce qui est de la migration pendulaire interne, ou aux personnes retraitées qui passent une partie de l'année dans une autre région ou un autre pays.
- Les étudiants qui se rendent fréquemment au domicile de leurs parents, par exemple les week-ends, ne se font souvent pas immatriculer à l'endroit où ils étudient parce qu'ils considèrent la résidence étudiante ou l'appartement loué comme étant temporaire. Dans certains pays, les règles de recensement suivent d'ailleurs cette approche.

² Dans la publication de la CEE de 2016 intitulée *Defining and Measuring Circular Migration (Définir et mesurer les migrations circulaires)*, la migration pendulaire est considérée comme un type particulier de migration circulaire. Voir http://www.unece.org/fileadmin/DAM/stats/publications/2016/ECECESSTAT20165_E.pdf (p. 11,12 et 19).

2. Transnationalisme

42. Bien qu'en principe, le transnationalisme ne soit pas non plus un phénomène nouveau, ce n'est que récemment qu'on s'y est intéressé. Jusqu'à présent, la notion n'a pas été clairement définie. On entend généralement par transnationalisme le fait de vivre dans deux (ou plusieurs) pays, l'un d'entre eux étant considéré comme le pays d'origine ou de résidence, avec lequel la personne maintient un lien lorsqu'elle vit dans un autre (ou plusieurs autres) pays.

43. Il est clair que la migration pendulaire peut être considérée comme un cas particulier de transnationalisme. Dans le cas de la migration pendulaire, le lien des migrants avec leur pays d'origine est particulièrement fort parce que leur famille y habite. Le transnationalisme se définit au regard de la migration, bien que celle-ci ne soit pas toujours une condition *sine qua non* (en particulier dans le cas de grandes distances) ; un contact étroit avec des parents dans le pays d'origine peut suffire.

44. Actuellement, le plus grand problème lié au transnationalisme est l'ambiguïté de sa définition. Néanmoins, ce phénomène est important pour de nombreux pays, et la société y porte un intérêt non négligeable.

V. Nouvelles approches pour garantir la couverture et la qualité du recensement

45. L'exhaustivité du dénombrement de la population est l'un des éléments essentiels de la qualité d'un recensement. Il existe actuellement diverses méthodes pour évaluer l'exhaustivité de la couverture, la plus courante étant l'enquête postcensitaire. La transition vers des méthodes de recensement partiellement ou entièrement fondées sur des registres et des sources administratives peut nécessiter de revoir la méthodologie utilisée pour mesurer la couverture. Les Directives concernant l'utilisation de registres et de données administratives pour les recensements de la population et des logements³, que la CEE a publiées en 2018, abordent en partie ce sujet.

46. La qualité du recensement pourrait être évaluée à l'aune de l'objectif visé. Bien que les pays aient généralement tendance à suivre les principes et les recommandations applicables à l'échelon international lorsqu'ils procèdent à un recensement de la population et des habitations, chacun d'eux pourrait évaluer différemment si l'« objectif visé » a été atteint selon l'utilisation particulière qu'il fera des résultats du recensement. Un pays pourra les utiliser pour justifier des transferts de fonds entre différents échelons de son administration, un autre pourrait les mettre à profit dans le cadre d'un processus électoral, tandis qu'un troisième pays pourrait suivre une tout autre voie⁴. Un des volets du recensement qui est parfois négligé est celui qui consiste à déterminer comment la couverture de la population devrait être mesurée. Cet exercice peut devenir plus problématique et complexe si les résultats du recensement sont publiés plus fréquemment, par exemple tous les ans.

47. Un recensement vise généralement à dénombrer la population d'un territoire délimité aussi précisément que possible, ce qui nécessite des niveaux géographiques très détaillés. Les concepts mesurés doivent correspondre aux différentes utilisations prévues des données produites et être aussi cohérents que possible dans le temps, tout en tenant compte des phénomènes émergents. Si tous les aspects d'un recensement qui ont trait à l'actualité, à la pertinence, à la cohérence, à l'accessibilité et à l'interprétabilité sont importants, un accent particulier peut être mis sur l'exactitude.

³ Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses, texte disponible à l'adresse : <https://www.unece.org/index.php?id=50794&L=0>.

⁴ L'Équipe spéciale de l'évaluation de la qualité des sources administratives en vue de leur utilisation pour les recensements, qui relève de la CSE, se penche actuellement sur les critères permettant de définir la qualité des recensements fondés sur de nouvelles méthodes et sources de données ; elle rendra compte de ses conclusions au Bureau de la CSE en 2021.

48. Comment l'utilisation de données provenant de sources administratives aux fins d'un recensement peut-elle nuire à la couverture ? Théoriquement, de telles sources devraient garantir une couverture adéquate de l'ensemble de la population. Leur utilisation est particulièrement importante lorsqu'il n'existe pas de véritables registres (notamment pour la population), comme c'est le cas, par exemple, au Canada et au Royaume-Uni. Les sources administratives doivent être disponibles en temps utile et doivent également porter sur la période de référence concernée. Lorsque plusieurs sources sont utilisées, la période de référence différera généralement d'une source à l'autre. Il arrive même que divers attributs relevant d'une seule et même source se rapportent à des périodes de référence différentes. Par exemple, la période de référence d'un dossier d'impôt sur le revenu peut correspondre à l'année civile précédente, tandis que l'adresse d'une personne peut être son adresse actuelle. Dans pareils cas, le choix de la date ou de la période de référence du recensement devrait se faire en fonction de la date effective ou prévue à laquelle sera disponible la version la plus complète des registres ou des sources administratives qui seront utilisés aux fins du décompte de la population. Les pays qui envisagent d'utiliser une méthode de recensement mixte peuvent éprouver plus de difficultés à choisir une date de référence qui convienne, puisque la période optimale pour produire chaque segment peut varier.

49. L'utilisation de multiples registres ou sources administratives peut susciter des difficultés supplémentaires. Par exemple, il peut être compliqué d'intégrer ces sources par couplage des données. Un manque de précision ou de qualité des clés de couplage peut directement entraîner des erreurs de corrélation et, partant, avoir une incidence sur la couverture. Il convient de trouver un équilibre entre les erreurs de couplage (faux positifs entraînant un sous-dénombrement) et les omissions de couplage (faux négatifs entraînant un surdénombrement). Cette opération peut devenir plus complexe en fonction du nombre de sources utilisées, puisqu'il faudra établir davantage de liens, et en fonction des attributs de ces sources pour ce qui est des clés de couplage, de l'incidence sur la couverture et de la qualité de l'information.

50. Les difficultés peuvent être exacerbées par l'incertitude des signaux. L'absence d'informations sur une personne ne devrait pas être automatiquement interprétée comme signifiant que cette personne n'entre pas dans le dénombrement de la population, pas plus que la présence d'un signal ne devrait être considérée comme une indication certaine que la personne concernée fait partie de la population à dénombrer. L'utilisation de multiples sources administratives dans le but d'améliorer la qualité de la mesure concernant chaque individu entraînera très probablement un surdénombrement, puisqu'il est généralement plus difficile d'identifier les cas qui ne correspondent pas aux critères⁵.

51. Les études de la couverture sont généralement conçues de telle sorte qu'elles produisent des résultats agrégés, alors que les résultats d'un recensement portent habituellement sur de plus petites unités géographiques. Comme indiqué dans la section consacrée aux populations servant de base, l'évolution des schémas sociaux et migratoires rendra la situation encore plus complexe. Les prochaines recommandations pour les recensements de la population et des habitations devraient mettre davantage l'accent sur la mesure de la qualité des résultats des recensements.

⁵ Plusieurs publications de la CEE donnent des conseils sur les techniques et les considérations relatives à l'intégration de multiples sources de données administratives : *Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses* (2018, cité plus haut) ; *Guidance on Data Integration for Measuring Migration* (2019), disponible à l'adresse : <https://www.unece.org/index.php?id=51143> ; *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices* (2012), disponible à l'adresse : <http://www.unece.org/index.php?id=28898> ; et *Guide to Data Integration for Official Statistic* (2018), établi par le Groupe de haut niveau sur la modernisation de la statistique officielle, disponible à l'adresse : <https://statswiki.unece.org/x/OQGzBw>.

VI. Nécessité de moderniser les statistiques géospatiales

52. L'idée semble s'imposer que les enquêtes et les recensements traditionnels concernant des zones statistiques fixes ne répondent pas aux demandes des utilisateurs de données en matière de flexibilité territoriale. Afin de répondre à ces demandes croissantes, il est nécessaire d'introduire une nouvelle dimension territoriale dans la production de statistiques qui repose sur un système de production ponctuelle. En reliant des données de recensement à une localisation géographique définie par les coordonnées x et y , il est possible d'obtenir des informations non seulement au sein de zones administratives ou statistiques, mais aussi concernant un large éventail de zones fonctionnelles (villes ou villages, zones urbaines, littoral, etc.), y compris des petites zones comme les mailles statistiques (de 1 km^2 ou de forme hexagonale) ou des zones librement définies en fonction des intérêts des utilisateurs.

53. La demande en informations statistiques se rapportant à un lieu a fortement augmenté dans tous les secteurs de la plupart des pays : secteurs public, commercial, à but non lucratif, universitaire et civique. La demande croissante en statistiques géospatiales s'explique par le fait que les individus, les administrations, les organismes publics et d'autres organisations veulent mieux connaître certaines populations spécifiques ou un sujet qui suscite un intérêt croissant, et ce, à l'aide d'informations spatio-temporelles de plus en plus précises. Le lieu où vivent les populations étudiées est un aspect important des informations demandées, particulièrement au niveau local et cela dans de nombreux cas. En outre, l'accélération rapide de la mobilité des personnes, des biens et des services signifie qu'il est nécessaire de mieux comprendre cette dynamique complexe qui se déroule à l'intérieur des régions et entre celles-ci. Les hypothèses qui ont été formulées précédemment à propos du lieu où les gens vivent, travaillent, jouent ou accèdent aux services perdent de leur valeur dans le monde actuel.

54. Les informations sur les mouvements de populations et les flux de transactions commerciales entre des lieux, de même que les statistiques régionales plus traditionnelles sur la démographie et les entreprises, sont utilisées conjointement afin de suivre la tendance actuelle en faveur d'une « prise de décisions fondée sur le lieu ».

55. Créer un cadre de référence spatiale pour les statistiques pourrait être l'un des principaux moyens dont disposent l'ensemble des pays pour que leurs statistiques spatiales soient plus actuelles, plus précises et plus rentables. Des registres administratifs géocodés répertoriant les adresses, immeubles et logements sont au cœur d'un tel cadre de référence. Ces registres devraient fournir un cadre de référence pour le géocodage de tous les recensements à venir, ainsi que pour relier les sources de données administratives aux emplacements géographiques. Un aspect important du système de référence spatiale est qu'il doit être doté de marqueurs uniques, stables dans le temps et susceptibles d'être utilisés comme des clés uniques pour référencer toutes les informations pertinentes les concernant. Les données composant ce système de référence spatiale doivent également être facilement accessibles.

56. Il est indispensable de disposer d'informations de recensement à des niveaux géographiques précis et normalisés pour faciliter la référencement et pour obtenir une meilleure qualité de production de données statistiques géospatiales. À des fins de recensement, il est important d'avoir des unités de données géoréférencées, de préférence à l'aide des coordonnées x et y . Ces informations permettent aux organismes nationaux de statistique d'agrèger les données du recensement à n'importe quel niveau spatial.

57. Les opérations d'intégration de données géospatiales et statistiques sont très complexes. La difficulté consiste à comprendre quelle est la meilleure façon de réaliser cette intégration d'une manière efficace et cohérente. Dans le cadre des prochaines recommandations pour les recensements de la population et des habitations, ou même avant, il pourrait être souhaitable de mettre au point une approche cohérente et systématique permettant de combiner les données statistiques et géospatiales. Une telle approche pourrait s'appuyer sur le cadre statistico-géospatial mondial, sur lequel la communauté internationale s'est accordée. Les paragraphes suivants offrent un aperçu d'un

document établi par l'organisme national de statistique de la Pologne, Statistics Poland, qui fournit des informations sur la façon dont une telle intégration pourrait être réalisée.

A. Intégration accrue des données géospatiales dans le système statistique : l'exemple de la Pologne

58. En Pologne, le traitement des données statistiques liées à l'information géospatiale a été introduit pour la première fois lors du recensement de 2011. Au début, des données géodésiques de référence provenant des sources géodésiques et cartographiques nationales étaient utilisées pour l'ensemble du pays. Elles comprenaient les informations géométriques suivantes : limites administratives, noms géographiques, base de données topographiques, données cadastrales, registre des bâtiments, cartes orthophotographiques (télétection par satellite) et réseau de rues et de routes. Les sources statistiques utilisaient des documents tels que des croquis de situation, des cartes statistiques indiquant les limites des divisions statistiques (régions statistiques, zones de calcul) et des données descriptives provenant du Registre national officiel de la Division territoriale nationale (système TERYT) qui compile les identifications uniques assignées à tous les échelons de l'administration (ville, village, secteur statistique, rue, immeuble, appartement).

59. Toutes les informations susmentionnées ont été réunies dans la base de données statistico-spatiales qui avait servi à la collecte des données (géocodage) et ont été utilisées afin que des données de recensement soient disponibles aux fins de l'analyse de la dimension spatiale.

60. Les statisticiens ont eu accès aux bases de données de cartographie des adresses et se sont appuyés sur des cartes orthophotographiques et de cartes hypsométriques de bâtiments. Grâce aux possibilités de vérification, ils ont pu mettre à jour l'emplacement et les attributs des localisations géographiques (ville, rue, numéro d'habitation), ainsi que saisir manuellement de nouvelles adresses directement dans ces bases de données. Cela a permis de définir ou d'évaluer la localisation spatiale des adresses avec exactitude grâce aux coordonnées x et y, et de les ajouter au bon endroit.

61. Cet exemple venu de Pologne montre que la création d'une infrastructure de géocodage pour les statistiques et son intégration dans le processus de production statistique n'exige pas une refonte complète des architectures d'entreprise et des processus de production statistique. Des améliorations mineures et progressives sont possibles. Toutefois, l'intégration des informations statistiques et géospatiales reste la pierre angulaire de la modernisation des statistiques officielles.

VII. Dépendance accrue à l'égard des modèles d'estimation pour produire des résultats de recensement

62. À mesure que les pays délaissent leurs méthodes de recensement traditionnelles au profit d'approches fondées en grande partie sur des registres statistiques et des sources administratives, ils se heurtent à de nouvelles difficultés en matière de production ou de préservation de certains résultats de recensement. Il arrive que les informations trouvées dans les nouvelles sources ne soient pas aussi complètes ou précises pour certaines variables que les informations précédemment produites à l'aide du modèle traditionnel. Il peut également y avoir de légères différences entre les sources s'agissant des concepts mesurés.

63. C'est le cas, par exemple, de l'identification des ménages. Selon une approche traditionnelle, le ménage serait établi sur la base de questions et de concepts appliqués au moment du recensement. La notion de ménage n'existe souvent pas dans les sources de données administratives et elle doit être construite d'une manière ou d'une autre. La plupart des sources administratives contiennent des informations sur les personnes, et la relation entre les personnes résidant à une adresse donnée doit souvent être déduite puisqu'elle peut ne pas être facilement disponible dans ces sources. La notion de ménage peut également varier d'un pays à un autre. Les différences portent principalement sur les règles appliquées

pour déterminer à quel ménage appartiennent les enfants en garde partagée, les étudiants absents du logement pendant leurs études, les personnes vivant temporairement dans une institution (comme un hôpital), les personnes qui quittent temporairement leur logement pour des motifs professionnels, etc.

64. Dans les cas susmentionnés, il sera peut-être nécessaire d'appliquer diverses méthodes statistiques pour déduire ou estimer les variables de recensement requises. Le recours à des méthodes d'estimation peut également être nécessaire pour produire des résultats à variables multiples. Des méthodes différentes doivent être utilisées si les données manquent à des degrés divers dans les thèmes traités, comme cela est souvent le cas lorsque de multiples sources de données sont utilisées pour produire des résultats de recensement.

65. Les prochaines recommandations devraient tenir compte de la transformation que connaissent les méthodes de recensement et inclure des informations sur l'utilisation des méthodes d'estimation aux fins de la création de variables de recensement et sur d'autres questions connexes.

A. Système de recensement fondé sur des registres et micro-intégration : l'exemple des Pays-Bas

66. La clé de voûte du recensement aux Pays-Bas est le registre central de la population qui regroupe tous les registres municipaux de la population. Les données du registre central de la population au 1^{er} janvier 2011 (année du dernier recensement) ont servi de base pour élaborer un ensemble d'hypercubes (tableaux englobant de nombreuses dimensions). Les hypercubes mettent l'accent sur le dénombrement de fréquences et non sur des informations quantitatives. Les données non disponibles dans le registre central de la population ou qui ne pouvaient en être déduites ont été tirées d'autres registres. Toutes les variables issues de registres sont désormais disponibles dans le système regroupant l'ensemble des données statistiques sociales de Statistics Netherlands, et leur qualité a été améliorée grâce à des techniques consistant à intégrer des microdonnées. La micro-intégration consiste à vérifier les données et à corriger celles qui sont incorrectes. On estime généralement que les données micro-intégrées donnent des résultats plus fiables lorsque des registres sont utilisés, car elles reposent sur le plus grand nombre d'informations possible. Elles permettent également de mieux couvrir les sous-populations : si des données manquent dans une source, une autre source peut être mise à profit.

67. Dans le passé, la micro-intégration a joué un rôle très important dans le système de données statistiques sociales néerlandais, et en particulier dans les recensements. Toutefois, à l'avenir, le rôle de la micro-intégration sera plus limité étant donné qu'aux Pays-Bas, le nouveau système des registres de base est mieux intégré que l'ensemble des registres que Statistics Netherlands recevait auparavant. Le système des registres de base comprendra à terme 13 registres de la population (résidents et non-résidents), d'adresses et de bâtiments, d'entreprises, de biens immobiliers (limites, propriété, valeur, etc.), de topographie (cartes : terrain, eau, routes), d'automobiles (modèle, couleur, propriétaire, etc.), de revenus imposables, de travail (salaires, employeurs, avantages sociaux, etc.) et d'infrastructures souterraines (réseau d'assainissement, câbles, etc.). Le système des registres de base est défini par la législation. Un ministre est responsable de chaque registre de base et le Ministre de l'intérieur est responsable du système des registres de base.

68. Les utilisateurs des registres de base peuvent se fier à leur validité. Les données statistiques fondées sur les registres de base ne nécessitent que peu de vérifications. Dès lors que les registres de base sont reliés entre eux (par des numéros d'identification), les données statistiques sont donc également cohérentes. Contrairement aux données issues d'enquêtes par sondage, il n'est pas nécessaire de recourir à des méthodes d'estimation pour les statistiques concernant de petites zones et le problème des non-réponses ne se pose pas. Toutefois, certaines données sont disponibles avec retard, comme les informations sur les travailleurs indépendants. Il ne faut pas oublier que les registres ne sont pas conçus pour servir des fins statistiques et qu'ils ne contiennent pas toujours les concepts statistiques requis. Il arrive toutefois que des concepts statistiques différents soient acceptables, alors

que, d'autres fois, il faut essayer d'obtenir des informations sur le concept statistique requis en procédant par déduction. Si de telles déductions n'étaient pas satisfaisantes, on pourrait finalement décider de recourir – faute de mieux – à des enquêtes par sondage à l'aide du concept statistique requis, bien que les possibilités de publication des résultats soient moindres qu'avec les données de registres.

VIII. Utilisation des mégadonnées à des fins de recensement

69. Outre les registres et les sources administratives publiques, les organismes nationaux de statistique sont de plus en plus contraints de chercher d'autres sources de données pour la production de statistiques officielles, y compris de résultats de recensement. Beaucoup d'entre eux étudient ou exploitent déjà de telles sources dans le cadre de nombre de leurs programmes, comme les données scannées et d'autres données transactionnelles à des fins de production de statistiques économiques, par exemple.

70. Les méthodes permettant d'obtenir ces types de données sont également en train d'évoluer. Hormis le fait qu'elles permettent d'obtenir des fichiers de données directement auprès des producteurs et des fournisseurs de données, des méthodes comme l'extraction de contenus de sites Web et la télédétection gagnent du terrain.

71. L'utilisation des informations provenant de ces sources peut présenter des difficultés d'un nouvel ordre que le recours à des registres et à des dossiers administratifs de sources publiques ne pose pas. Il se peut que les informations ne soient pas accompagnées de métadonnées complètes, ce qui complique davantage l'évaluation de la qualité ou de l'exhaustivité de la source de données. Les concepts inhérents aux dossiers peuvent également ne pas correspondre exactement à ce qui est disponible dans certains registres et à ce qui avait été recueilli directement auprès des répondants au moyen d'un dénombrement direct traditionnel.

72. Les prochaines recommandations devraient aborder de manière plus complète les différents aspects de l'utilisation de nouvelles sources de données non structurées à des fins de recensement.

A. Utilisation des données mobiles de localisation pour déterminer le lieu de résidence : l'exemple de l'Estonie

73. En 2017, Statistics Estonia a mené un projet pilote visant à tester la possibilité d'utiliser les données mobiles de localisation dans le contexte d'un recensement basé sur les registres, afin de résoudre le problème lié à la différence entre lieux de résidence enregistrés et lieux de résidence effectifs, laquelle était source d'écarts dans les statistiques sur la population. Les données mobiles de localisation permettent de détecter l'emplacement d'un appareil mobile. De nombreuses technologies, qu'elles soient basées sur un réseau ou un téléphone portable, permettent d'obtenir des données mobiles de localisation.

74. Statistics Estonia a trouvé des volontaires qui ont été invités à participer au projet pilote, et un ensemble d'adresses potentielles a été constitué pour chaque participant à partir des registres. Ensuite, les points d'ancrage des lieux de résidence ont été estimés sur la base des données mobiles de localisation. Les points d'ancrage et autres informations auxiliaires ont ensuite été utilisés pour construire un modèle permettant de sélectionner le lieu de résidence le plus probable pour chaque personne à partir de l'ensemble des adresses. En outre, les volontaires ont indiqué leur lieu de résidence effectif pour qu'il soit possible de comparer leur adresse réelle avec les données de localisation obtenues.

75. L'analyse de la précision des points d'ancrage domicile et domicile-lieu de travail obtenus par rapport aux adresses communiquées par les personnes a été réalisée dans le cadre de l'analyse de l'indicateur de précision des résultats.

76. Un point d'ancrage domicile ou un point d'ancrage domicile-lieu de travail a pu être calculé pour 93 % des volontaires à l'aide des données mobiles de localisation. Pour certains individus, il n'a pas été possible de calculer le point d'ancrage à l'aide des données

mobiles de localisation et cette méthode n'a par conséquent pas pu être employée pour valider leur lieu de résidence. Cette situation s'explique par le fait que les individus en question faisaient un usage limité de leur téléphone ou par des problèmes techniques liés au traitement et à l'interprétation des données, notamment au calcul des points d'ancrage.

77. Pour 82 % des personnes concernées, l'adresse de leur lieu de résidence avait été située dans la zone de couverture du point d'ancrage domicile ou du point d'ancrage domicile-travail. L'observation du centroïde de la zone de couverture d'un point d'ancrage présente pour inconvénient que, normalement, la zone de couverture d'une antenne relais couvre plusieurs localités (particulièrement dans les zones faiblement peuplées), ce qui ne signifie pas nécessairement que le point d'ancrage a été mal calculé.

78. L'analyse comparative par échantillonnage a permis d'analyser les sources d'erreurs dans les calculs des points d'ancrage ainsi que les corrélations entre les résultats obtenus et divers facteurs. Ces connaissances peuvent aider à élaborer plus avant et à compléter la méthode des points d'ancrage, laquelle devrait permettre d'améliorer considérablement la validation des données.

79. L'équipe estonienne chargée du recensement a estimé que les nouvelles sources de données et les résultats étaient acceptables, mais elle n'a pas appliqué cette méthode aux processus statistiques du recensement à l'époque. L'absence de base juridique pour l'utilisation de ces informations en était la raison principale : l'utilisation de données mobiles dans le cadre d'un recensement nécessitait de modifier plusieurs textes juridiques, ce qui est une tâche complexe.

B. Conciliation des besoins des différents utilisateurs

80. Selon les Recommandations pour les recensements de la population et des habitations de 2020, il est largement accepté que la qualité des statistiques se mesure à l'aune de six critères : la pertinence, l'exactitude, l'actualité, l'accessibilité, l'interprétabilité et la cohérence⁶.

81. La prolifération des sources de données, dont les mégadonnées, les progrès enregistrés en matière de capacité et de méthodes d'exploitation des grands ensembles de données et la valeur accrue accordée aux données acquises dans un monde en constante évolution s'accompagnent de la demande de statistiques plus actuelles. Les discussions au niveau européen mettent l'accent sur cette nécessité de garantir l'actualité des cycles de recensements à venir, raison pour laquelle les méthodes statistiques basées en grande partie sur l'utilisation de registres administratifs et d'autres sources de données retiennent de plus en plus l'attention. L'exploitation de ces sources de données est souvent moins coûteuse et peut produire des informations régulières dans des délais plus courts que les méthodes basées sur le dénombrement direct.

82. Tenter de répondre aux demandes changeantes et croissantes des utilisateurs de données revient souvent à trouver un compromis entre l'actualité, l'exactitude et la comparabilité dans le temps, trois qualités qu'il est rarement possible de réunir simultanément. Lorsque l'on travaille avec des sources de données administratives, il est certes possible de produire et de diffuser des données très rapidement, mais, pour cela, il faut parfois sacrifier certains contrôles de la qualité ou renoncer à recourir à des techniques de modélisation pour pallier certaines lacunes des sources de données. Divers contrôles pourraient être effectués aux fins de l'assurance qualité, mais ils le seraient généralement au détriment de l'actualité, puisque les données seraient alors publiées plus longtemps après la période de référence.

83. Pour de nombreux organismes nationaux de statistique, les possibilités offertes par les nouvelles sources de données peuvent créer des conflits avec les attentes fluctuantes des utilisateurs. En mettant l'accent sur la production de statistiques officielles présentant un degré élevé d'exactitude, ces organismes ont souvent dû renoncer à l'actualité des données. La transition vers de nouvelles approches peut nécessiter un changement d'orientation dans

⁶ Chapitre IV, Gestion de la qualité, p. 79 et 80.

ce domaine. De nouveaux critères de qualité pourraient tenir compte en premier lieu de la satisfaction des besoins des utilisateurs. La difficulté consisterait à déterminer comment évaluer la satisfaction des utilisateurs afin de trouver le bon compromis entre l'exactitude et l'actualité.

84. L'une des difficultés que pose la demande de statistiques plus actuelles peut être liée au fait que les besoins des utilisateurs sont variés et dépendent de leur profil. Par exemple, les utilisateurs de données relatives aux politiques et de nombreux chercheurs peuvent avoir besoin de résultats très précis pour de petites zones ou de petits groupes, et garantir l'exactitude de ces résultats nécessitera généralement plus de temps. Les utilisateurs au niveau national peuvent préférer des données plus actuelles, même si celles-ci ne sont pas aussi précises qu'au niveau local.

85. Un autre aspect lié à l'évolution des méthodes de recensement est la capacité des utilisateurs de données de comprendre ces changements et leur éventuelle incidence sur les résultats du recensement. Certains utilisateurs peuvent exprimer des préoccupations quant aux méthodes utilisées pour produire des résultats, comme la modélisation, et s'inquiéter d'une réduction ou d'une perte éventuelle de comparabilité avec les recensements précédents et d'autres sources.

86. Les prochaines recommandations pour les recensements de la population et des habitations devront peut-être donner des orientations sur les compromis à envisager, et prendre position sur les questions suivantes :

- a) Quel est l'équilibre optimal entre l'actualité et l'exactitude ou l'exhaustivité lorsqu'on tient compte des besoins des utilisateurs ? Quels sont les critères à prendre en considération ?
- b) Est-il acceptable de modifier la définition d'une notion (celle de population servant de base, par exemple) pour permettre une utilisation plus facile des données administratives ou pour mieux accommoder les nouvelles sources de données afin d'obtenir des données plus actuelles ? Les définitions devraient-elles tenir compte de ce besoin d'actualité ou devraient-elles être arrêtées indépendamment de ce besoin ?
- c) S'il est possible de ne plus mesurer certaines caractéristiques en prenant l'intégralité de la population, mais en se limitant à un échantillon, cela peut donner aux utilisateurs l'impression que la qualité est moins bonne. Ce serait le cas, notamment, lorsque le dénombrement de la population est réalisé à partir de registres et que les caractéristiques traditionnellement recueillies directement grâce au recensement le sont désormais au moyen d'un échantillonnage, soit par dénombrement direct, soit par des sources administratives. Il deviendrait alors essentiel de gérer la façon dont les utilisateurs perçoivent les méthodes utilisées aux fins de l'assurance qualité.

IX. Aspects éthiques, protection de la vie privée et acceptabilité par le public

87. Pour la plupart des pays qui envisagent de transformer la façon dont ils conçoivent le recensement de leur population en s'appuyant sur des registres, la transition se fera sans heurt à condition d'avoir accès à un éventail plus large de sources de données. Dans de nombreux pays, les organismes nationaux de statistique disposent déjà de ce qu'ils considèrent comme des sources de données essentielles pour produire des dénombrements de la population. Il s'agit généralement de sources censées permettre une couverture élevée de la population, comme les registres fiscaux et les registres de sécurité sociale. Cependant, la nécessité de continuer à produire des estimations de recensement pour des caractéristiques plus larges de la population – qui relèvent souvent de variables multiples/de petites zones – obligera maints pays à développer de nouvelles microdonnées qui seront construites à partir de sources multiples et diverses.

88. Nombre de pays dont les travaux de recherche portent sur les recensements à venir ont tendance à mettre l'accent sur l'intégration des sources de données que les ministères

recueillent régulièrement. Toutefois, certains pays se tournent vers d'autres formes de données commerciales qui présentent un intérêt singulier pour des sujets spécifiques, comme les téléphones mobiles pour les flux pendulaires.

89. Il est également nécessaire de revoir le rôle que joueront les enquêtes sociales dans les prochains recensements, puisque de nombreux pays s'attendent encore à devoir recueillir des informations sur des sujets qui ne sont pas disponibles auprès d'autres sources, ou à devoir évaluer la couverture et la qualité des données administratives.

A. Protection de la vie privée

90. Un modèle de recensement qui fait fond sur l'intégration de multiples sources de données pose de nouvelles difficultés aux organismes nationaux de statistique en matière de protection de la vie privée. Lorsque l'organisme national de statistique n'est plus chargé de recueillir et de traiter des données de bout en bout, il doit prendre des mesures supplémentaires pour veiller à protéger les informations personnelles. Ces organismes doivent repenser leur modèle d'activité concernant toute une gamme de leurs fonctions, notamment l'importation et la gestion des données, l'accès aux données, le couplage des données et le contrôle de la divulgation des données.

91. Il arrive que les fournisseurs et l'organisme national de statistique ne soient pas par la suite en mesure de se communiquer des données tant que certaines conditions ne sont pas remplies. Au Royaume-Uni, par exemple, l'Office for National Statistics (ONS) progresse dans les travaux de recherche préliminaires qu'il mène en vue d'organiser un recensement fondé sur des données administratives, réalisé dans un environnement de recherche sécurisé et reposant sur un cryptage des données et sur un certain nombre d'autres outils destinés à préserver la confidentialité des informations permettant d'identifier une personne.

92. S'il est important de reconnaître que l'utilisation de sources multiples augmente les risques pour la vie privée des individus, tout changement radical apporté aux procédures peut avoir des effets négatifs sur la qualité des résultats statistiques. Par exemple, le cryptage des données peut entraîner des erreurs dans les procédures de couplage des données et fausser les résultats ainsi obtenus. D'autres tentatives visant à protéger la vie privée en séparant les attributs des ensembles de données ou en réduisant la granularité des unités géographiques peuvent restreindre les possibilités de produire des statistiques au niveau de détail et de précision requis. L'enjeu auquel les organismes nationaux de statistique doivent répondre consiste à trouver un équilibre leur permettant de protéger la vie privée tout en fournissant des produits qui ont encore une utilité en ce qu'ils répondent aux besoins des utilisateurs.

B. Acceptabilité par le public

93. La majeure partie des organismes nationaux de statistique qui souhaitent évoluer vers un recensement fondé sur des registres, ou qui mènent déjà des recensements sur cette base, peuvent s'appuyer sur une législation les autorisant à accéder à des ensembles de données pour soutenir la production de statistiques officielles. Or, les membres du public qui fournissent des informations les concernant lorsqu'ils interagissent avec des services publics ne savent généralement pas comment ces données sont utilisées. Comme pour les recensements traditionnels, la participation aux prochains recensements fondés sur des registres pourrait être considérée comme obligatoire puisque les possibilités pour une personne de demander que les informations la concernant ne soient pas retenues aux fins de l'analyse seraient limitées. Il est généralement admis que les organismes nationaux de statistique devront résister à toute pression visant à exiger d'eux qu'ils obtiennent le consentement éclairé des personnes concernées avant d'utiliser leurs données dans le cadre de prochains recensements fondés sur des données administratives. Un faible pourcentage d'individus ne voulant pas fournir d'informations personnelles à ces fins pourrait suffire à fausser dans une très large mesure les statistiques démographiques.

94. Les membres du public qui souhaitent obtenir des garanties quant aux motifs sous-tendant la collecte des données et aux méthodes de traitement des données peuvent

aisément avoir des informations à ce sujet lorsque des données sont recueillies aux fins d'un recensement traditionnel. Dans le cas d'un recensement fondé sur des registres, les utilisations qui sont faites des données n'ont plus grand-chose à voir avec les objectifs pour lesquels ces données avaient été recueillies dans un premier temps. Les organismes nationaux de statistique doivent se demander dans quelle mesure ils comptent informer le public de l'utilisation qui sera faite des sources de données recueillies pour les statistiques officielles. Statistics New Zealand, qui a relativement bien progressé dans le développement de son infrastructure de données intégrée, a pris des mesures importantes pour sensibiliser le public à la manière dont elle conçoit la réutilisation des données existantes à des fins de statistiques et de recherche⁷.

95. En fonction du niveau de sensibilité du débat public sur le partage de données, les organismes nationaux de statistique doivent se demander dans quelle mesure ils souhaitent promouvoir activement les utilisations de différentes sources de données dans les recensements à venir. Il est important d'expliquer aux parties prenantes l'intérêt que présentent ces utilisations, tout en évitant que celles proposées suscitent des échos défavorables. La stratégie qui est souvent privilégiée pour tenter de cerner les éventuelles préoccupations du public consiste à conduire des études à petite échelle sur l'acceptabilité par le public, à l'aide soit de groupes de discussion, soit d'enquêtes quantitatives. De telles études parviennent en général à la conclusion que le public est favorable au partage de données une fois qu'il a pris conscience de l'intérêt que cela présente, mais qu'il est de plus en plus compliqué pour les participants de comprendre les processus qui entrent en jeu lors de la compilation des statistiques provenant de sources multiples. Compte tenu de la portée apparemment illimitée des éventuelles études de recherche basées sur l'intégration des données, il est également difficile pour les participants d'appréhender tous les aspects éthiques qui sont pertinents dans des circonstances particulières. Pour cette raison, certains organismes nationaux de statistique cherchent à obtenir le soutien de la population pour produire de nouveaux résultats, en collaborant étroitement avec des groupes défendant les intérêts des personnes visées par l'étude.

C. Aspects éthiques

96. D'autres sources de données offrent de nouvelles possibilités aux organismes nationaux de statistique, même si certaines de ces sources soulèvent également des problèmes éthiques plus pointus. Des données sont généralement mises à la disposition des organismes nationaux de statistique à des fins statistiques uniquement, et à la condition que l'organisme concerné n'utilise aucune analyse qui en résulterait pour s'acquitter de fonctions directes qui sont considérées comme opérationnelles ou administratives dans la pratique. Dans les pays où des recensements fondés sur des registres ont déjà été mis en place, certains organismes nationaux de statistique utilisent des données administratives pour améliorer la qualité des collectes de données d'enquête. Par exemple, des agents recenseurs qui réalisent des enquêtes par téléphone ou en personne disposent des noms des individus inscrits dans les registres de la population et doivent vérifier auprès des répondants quelles sont les personnes qui résident encore à leur adresse.

97. L'utilisation de données administratives à l'appui de la collecte de données d'enquête est une technique parfois appelée « enquête longitudinale » (*dependent interviewing*). Cette technique présente un avantage particulier en ce qu'elle permet de mesurer le surdénombrement dans les données administratives, qui est un problème majeur pour les pays ne disposant pas d'un système de registres de la population. Les enquêtes longitudinales se heurtent toutefois à des obstacles importants puisqu'elles communiquent des informations sur les résidents précédents, enfreignent les accords de partage de données conclus avec des fournisseurs et portent atteinte à la réputation des pratiques de l'organisme national de statistique concerné.

98. Afin d'examiner les propositions de nouvelles utilisations des données, il y a lieu de mettre en place des mécanismes de contrôle éthique et de transparence. Par exemple,

⁷ Voir <https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/> et les informations auxquelles cette page renvoie.

l'Autorité de la statistique du Royaume-Uni (UK Statistics Authority) a créé un comité consultatif sur l'éthique des données chargé de fournir des avis au Statisticien en chef (National Statistician) tendant à ce que les modalités d'accès aux données, leur utilisation et leur partage aux fins de la recherche et de l'élaboration de statistiques pour le bien public soient conformes à l'éthique. Ce comité examine en détail chaque proposition novatrice et inédite d'utilisation des données que formulent l'ONS et, plus largement, le service public de la statistique (Government Statistical Service), avant de donner un avis indépendant au Statisticien en chef sur le point de savoir si ces propositions sont conformes à l'éthique. Dans le cadre de ses travaux, l'Autorité de la statistique du Royaume-Uni a fixé un certain nombre de principes éthiques et établi plusieurs processus transparents afin d'évaluer les recherches proposées à l'aune des principes éthiques suivants :

- a) L'identité du sujet des données (qu'il s'agisse d'une personne physique ou d'une organisation) est protégée, l'information reste confidentielle et sûre, et la question du consentement est dûment prise en considération ;
- b) Les risques et limites des nouvelles technologies sont pris en compte et la surveillance exercée par l'être humain est suffisante pour que les méthodes employées soient conformes aux normes d'intégrité et de qualité reconnues ;
- c) Les données utilisées et les méthodes employées sont conformes aux prescriptions légales ;
- d) Les vues du public sont prises en compte au regard des données utilisées et des avantages perçus des travaux de recherche ;
- e) L'accès aux données, leur utilisation et leur partage sont transparents, et communiqués clairement et de façon accessible au public.

X. Conclusion

99. Au cours des vingt dernières années, la manière dont les recensements de la population et des habitations sont menés à l'échelon international, et plus particulièrement dans la région de la CEE, a connu une évolution de plus en plus rapide. Étant donné qu'un nombre croissant de pays étudient et utilisent de nouvelles méthodes, on peut s'attendre à ce que cette cadence se maintienne jusqu'au prochain cycle de recensements.

100. Les recommandations en matière de recensement ont joué un rôle important en ce qu'elles guident l'élaboration et la mise en œuvre des concepts et méthodes propres au recensement. Dès lors que les recensements nationaux connaissent des transformations notables, il est plus important que jamais que celles-ci soient encadrées par un ensemble commun de principes et de lignes directrices, de sorte que cet examen statistique essentiel conserve sa pertinence et son efficacité au cours des nombreux autres cycles à venir. Sous l'égide des Principes fondamentaux des Nations Unies concernant la statistique officielle⁸, qui guident toutes les activités statistiques officielles, les recommandations en matière de recensement continueront de jouer ce rôle directeur. Étant donné qu'elles auront été révisées à la lumière des suggestions émises dans la présente analyse, leur utilité pour les recensements au-delà de 2020 en sera renforcée.

101. Il ressort de l'analyse des sujets abordés dans le présent document que la CSE devrait réexaminer, voire réviser, les points suivants en vue de la prochaine mouture de ses recommandations :

- a) Les moyens de veiller à la cohérence des notions et des données lorsque de multiples sources de données sont utilisées ;
- b) Les caractéristiques essentielles d'un recensement ;
- c) Les définitions de la population, du « lieu de résidence habituel » (voire de « résidence secondaire ») et du « ménage » ;

⁸ <http://www.unecce.org/stats/fps.html>

- d) Les paramètres de mesure de la couverture du recensement et de la qualité des résultats du recensement ;
 - e) La modernisation des statistiques géospatiales ;
 - f) La dépendance accrue à l'égard des modèles d'estimation pour produire des résultats de recensement ;
 - g) L'utilisation des mégadonnées à des fins de recensement ;
 - h) La manière de mieux concilier des différents besoins des utilisateurs et la multiplication des sources de données et des méthodes statistiques ;
 - i) Les aspects éthiques, la protection de la vie privée et l'acceptabilité par le public.
-