
Couplage agriculture–population : rapport sur la qualité des données, 2021



Date de diffusion : le 25 août 2023

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- | | |
|---|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-514-283-9350 |

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « [Normes de service à la clientèle](#) ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Publication autorisée par le ministre responsable de Statistique Canada

© Sa Majesté le Roi du chef du Canada, représenté par le ministre de l'Industrie, 2023

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Date de diffusion : le 25 août 2023

N° 32260006 au catalogue, numéro 2021001

Table des matières

Au sujet du Couplage agriculture–population 1

Nouveautés pour 2021 1

Sources d’erreur 2

Processus d’appariement automatisé 2

Taux de réponse et d’imputation 2

Échantillonnage et pondération 4

Qualité et suppression des données 4

Au sujet du Couplage agriculture–population

L'un des principaux avantages de mener le Recensement de l'agriculture en même temps que le Recensement de la population est la possibilité de coupler les renseignements tirés de ces sources au moyen d'un processus d'appariement automatisé, en vue de créer la base de données du Couplage agriculture–population. Cette base de données comprend certaines variables du Recensement de l'agriculture et certaines variables du questionnaire détaillé du Recensement de la population (p. ex. niveau de scolarité, profession). La base de données du Couplage agriculture–population permet d'effectuer des totalisations croisées de caractéristiques socioéconomiques des exploitants agricoles et des membres de leur famille (p. ex. âge, niveau de scolarité et revenu des exploitants) avec des caractéristiques agricoles des exploitations (p. ex. type d'exploitation, forme juridique, superficie agricole, revenus agricoles totaux bruts, dépenses d'exploitations agricoles totales et capital d'exploitation agricole total).

La première version de la base de données du Couplage agriculture–population a été élaborée pour la première fois en 1971, fondée sur les données du Recensement de 1971. Celle-ci a été suivie par une version fondée sur les données du Recensement de 1981, puis par de nouvelles versions pour les recensements subséquents, de même que pour le Recensement de 2021. La base de données de 2021 couvre les exploitants agricoles qui ont été dénombrés lors du Recensement de l'agriculture de 2021 et les membres de leurs familles, à l'exception des résidents des trois territoires du Canada et de [logements collectifs](#).

La base de données du Couplage agriculture–population est un regroupement de renseignements provenant de deux sources de données. Jusqu'en 2006, les données démographiques étaient tirées du questionnaire détaillé du Recensement de la population à participation obligatoire, qui était distribué à environ 20 % des ménages canadiens. En 2011, les données démographiques ont été tirées de l'Enquête nationale auprès des ménages, une enquête à participation volontaire qui a été distribuée à environ 33 % des ménages privés canadiens, à l'exclusion des résidents de logements collectifs. Depuis 2016, les données sont de nouveau tirées du questionnaire détaillé du Recensement de la population à participation obligatoire, lequel est distribué à environ 25 % des ménages privés canadiens. Par conséquent, les résidents de logements collectifs ne sont pas représentés dans la base de données du Couplage agriculture–population à partir de 2016.

Comme la base de données du Couplage agriculture–population regroupe des renseignements provenant de deux sources de données, les utilisateurs sont invités à consulter les documents de référence du [Recensement de la population](#) et du [Recensement de l'agriculture](#) pour obtenir plus de renseignements sur les méthodes utilisées aux fins de la collecte, du traitement et de la diffusion des données.

Nouveautés pour 2021

La création de la base de données du Couplage agriculture–population de 2021 suit la même méthodologie que celle utilisée en 2016 avec des changements touchant les variables suivantes.

1. Une variable sur le genre remplacera la variable du sexe à la naissance qui avait été utilisée pour les recensements de 2016 et des années antérieures. Compte tenu de la petite taille de la population non binaire, il est nécessaire d'agrégier les données dans une variable sur le genre à deux catégories pour protéger la confidentialité des réponses fournies. Dans ces cas, les personnes dans la catégorie « personnes non binaires » sont réparties dans les deux autres catégories de genre et celles-ci sont désignées par le symbole « + ». Dans la base de données du Couplage agriculture–population, la variable du genre est agrégée selon la forme binaire, « femmes+ » et « hommes+ ». La catégorie « femmes+ » comprend les femmes (et les filles) et peut aussi comprendre certaines personnes non binaires. La catégorie « hommes+ » comprend les hommes (et les garçons) et peut aussi comprendre certaines personnes non binaires.
2. La variable de l'état matrimonial a une nouvelle catégorie distincte pour « vivant en union libre ».
3. La variable de la forme juridique de l'exploitation agricole comprend seulement une catégorie pour les sociétés de personnes. Les recensements précédents comportaient deux catégories : société de personnes avec contrat écrit et société de personnes sans contrat écrit.

Sources d'erreur

Dans une enquête par sondage, comme le questionnaire détaillé du Recensement de la population, deux types d'erreurs peuvent se produire : des erreurs dues à l'échantillonnage et des erreurs non dues à l'échantillonnage. Dans un recensement comme le Recensement de l'agriculture, seules les erreurs non dues à l'échantillonnage peuvent se produire.

Des erreurs dues à l'échantillonnage se produisent lorsqu'on estime une caractéristique d'une population donnée à partir d'une partie seulement de la population plutôt que de l'ensemble de celle-ci. Ce type d'erreur peut être contrôlé par la taille de l'échantillon, le plan d'échantillonnage et la méthode d'estimation.

Les erreurs non dues à l'échantillonnage sont celles qui ne concernent pas l'échantillonnage. Il peut s'agir d'erreurs dans la base à partir de laquelle l'échantillon est extrait, de l'utilisation d'outils de collecte inadéquats, de la non-réponse à l'enquête ou encore d'erreurs liées à la saisie, à la vérification, au codage ou à d'autres étapes du traitement des données. Au cours des phases de planification, des mesures ont été prises pour réduire les erreurs non dues à l'échantillonnage, dont la mise à l'essai des questionnaires, la formation des intervieweurs, le contrôle de la qualité de la saisie et du codage des données et de nombreuses autres mesures.

Processus d'appariement automatisé

Les principes de base du processus d'appariement automatisé des données agriculture–population sont simples. Un exploitant agricole remplit un questionnaire du Recensement de l'agriculture et un questionnaire du Recensement de la population. L'exploitant peut aussi être sélectionné pour remplir le questionnaire détaillé du Recensement de la population, distribué à environ le quart de l'ensemble des ménages. Au début du traitement des données du Recensement de l'agriculture, les exploitants agricoles sont couplés aux bases de données du Recensement de la population à l'aide de renseignements communs aux deux questionnaires, comme le nom, le sexe, la date de naissance et l'adresse, pour établir une correspondance individuelle. Cette correspondance permet de coupler les données du Recensement de l'agriculture aux données du questionnaire détaillé du Recensement de la population, qui est requis par le processus de création de la Base de données du Couplage agriculture–population. Les recensements de l'agriculture de 1991 à 2021 permettaient aux répondants de déclarer un maximum de trois exploitants par exploitation agricole, et tous les exploitants agricoles ont été compris dans le processus d'appariement. Grâce à ces renseignements supplémentaires, il est possible d'analyser la relation entre les membres de la famille qui habitent dans le même ménage et qui travaillent dans la même exploitation agricole. De plus, il est possible d'inclure dans cette analyse les exploitants de différents ménages qui gèrent la même exploitation.

Dans certains cas, le couplage entre les données sur l'exploitant tirées du Recensement de l'agriculture et les données sur le ménage tirées du questionnaire détaillé du Recensement de la population a échoué en raison de la non-réponse au questionnaire détaillé du Recensement de la population. Le couplage pourrait également échouer en raison de renseignements incomplets sur l'exploitant dans le Recensement de l'agriculture, des exploitations agricoles spéciales qui sont comprises avec certitude dans la base de données du Couplage agriculture–population, mais qui sont exclues du questionnaire détaillé du Recensement de la population, ou d'autres facteurs. Dans de tels cas, les données du questionnaire détaillé ont été imputées dans la base de données du Couplage agriculture–population au moyen de renseignements provenant d'un ménage répondant semblable du Recensement de la population ou, dans de rares cas, les poids ont été ajustés pour tenir compte des enregistrements non couplés.

Taux de réponse et d'imputation

Les tableaux suivants présentent le taux de réponse pondéré pour l'ensemble du questionnaire détaillé du Recensement de la population et le taux d'imputation pondéré pour la population de la base de données du Couplage agriculture–population de 2021.

Couplage agriculture–population : rapport sur la qualité des données, 2021

Tableau 1

Taux de réponse pondéré au questionnaire détaillé du Recensement de la population, Canada et provinces^{1,2}

Région géographique	Taux de réponse pondéré au questionnaire détaillé du Recensement de la population pourcentage
Canada³	95,7
Terre-Neuve-et-Labrador	95,6
Île-du-Prince-Édouard	96,8
Nouvelle-Écosse	96,1
Nouveau-Brunswick	95,7
Québec	96,3
Ontario	96,2
Manitoba	94,4
Saskatchewan	93,5
Alberta	94,4
Colombie-Britannique	95,1

1. En 2021, le taux de non-réponse totale a remplacé le taux global de non-réponse dans la mesure de la non-réponse. Pour obtenir plus de détails, consultez la section [Taux de non-réponse totale](#) dans le chapitre 9 du *Guide du Recensement de la population, 2021*.

2. Pour obtenir plus de renseignements sur le calcul des taux de réponse pondérés du Recensement de la population, consultez la section [Taux de réponse au Recensement de la population de 2021](#) dans le chapitre 9 du *Guide du Recensement de la population, 2021*.

3. Le taux de réponse au questionnaire détaillé du Recensement de la population du Canada comprend les ménages des trois territoires du Canada.

On observe un biais de non-réponse lorsque les caractéristiques d'intérêt des non-répondants d'une enquête sont différentes de celles des répondants. Ainsi, plus la non-réponse à une enquête est élevée, plus le risque d'un biais de non-réponse est grand. La qualité des estimations peut être réduite si un tel biais est observé.

Tableau 2

Taux d'imputation pondéré de la base de données du Couplage agriculture–population, Canada et provinces

Région géographique	Taux d'imputation pondéré de la base de données du Couplage agriculture–population pourcentage
Canada¹	0,5
Terre-Neuve-et-Labrador	5,0
Île-du-Prince-Édouard	2,1
Nouvelle-Écosse	0,9
Nouveau-Brunswick	1,5
Québec	0,5
Ontario	0,4
Manitoba	0,4
Saskatchewan	0,7
Alberta	0,4
Colombie-Britannique	0,7

1. Le taux d'imputation pondéré pour le Canada dans la base de données du Couplage agriculture–population exclut les résidents de [ménages collectifs](#) et des trois territoires du Canada.

Le taux d'imputation pondéré de la base de données du Couplage agriculture–population représente le pourcentage pondéré des ménages de la base de données pour lesquels les données du questionnaire détaillé du Recensement de la population ont été imputées. Les taux d'imputation pondérés de la base de données du Couplage agriculture–population sont très faibles, ce qui réduit le risque de biais causé par l'imputation des données.

Échantillonnage et pondération

La base de données du Couplage agriculture–population contient des données agricoles (sur les exploitations agricoles et les exploitants agricoles) et des données démographiques (sur les personnes, les ménages, les familles de recensement et les familles économiques). Comme seul un échantillon de ménages canadiens a été sélectionné pour recevoir le questionnaire détaillé du Recensement de la population, des poids ont été attribués aux enregistrements de la base de données du Couplage agriculture–population afin de représenter l'ensemble de la population agricole. Des poids ont été calculés au niveau des exploitations agricoles, des personnes, des ménages, des familles de recensement et des familles économiques.

Les poids ont été calculés de façon indépendante dans chaque province. Pour la plupart des enregistrements¹, un poids initial a été obtenu pour chaque ménage à partir du Recensement de la population en fonction du nombre de ménages dans une région infraprovinciale et du nombre de personnes ayant rempli le questionnaire détaillé du Recensement de la population. Des caractéristiques appelées « contraintes » ont ensuite été déterminées. Il s'agit de caractéristiques démographiques et agricoles qui ont fait l'objet d'un dénombrement complet soit dans le cadre du Recensement de la population ou du Recensement de l'agriculture, et qui sont de première importance pour les utilisateurs de données.

Cinq poids de base ont été créés (un pour chacun des niveaux suivants : niveau de l'exploitation agricole, niveau de la personne, niveau du ménage, niveau de la famille de recensement et niveau de la famille économique). Les contraintes associées au niveau du poids de base ont été utilisées dans le calcul dudit poids de base. Par exemple, seules les contraintes au niveau du ménage ont servi à calculer le poids de base au niveau du ménage. Pour chaque province, une méthode appelée « régression généralisée » a fait en sorte que les estimations de ces contraintes tirées de la base de données du Couplage agriculture–population correspondent aux chiffres de population connus à l'échelle de la province. À partir de ces cinq poids de base, un total de six poids ont été créés. Encore une fois, le poids à utiliser variait en fonction du type d'analyse.

Quelle que soit la région géographique, les totaux ou sous-totaux pondérés de la population, des ménages, des familles ou des exploitations agricoles peuvent différer des estimations semblables présentées dans les précédentes diffusions de données du Recensement de l'agriculture. Cela est attribuable au fait que, dans le cadre du Recensement de l'agriculture, les données sont recueillies auprès de toutes les exploitations agricoles, tandis que les estimations de la base de données du Couplage agriculture–population proviennent d'un échantillon. Les écarts relatifs aux variables hautement corrélées avec au moins l'une des variables utilisées pour définir une contrainte devraient être plutôt faibles en général. Quant aux autres variables, les écarts dépendent de la relation avec la variable utilisée pour définir une contrainte et pourraient être considérables s'il n'existe aucune relation. Les estimations de la base de données du Couplage agriculture–population peuvent aussi différer de celles du Recensement de l'agriculture, car les logements collectifs (et les exploitations agricoles associées à ces logements) sont exclus de la base de données.

Qualité et suppression des données

Les résultats de la base de données du Couplage agriculture–population peuvent être supprimés pour deux raisons : 1) protéger la confidentialité des données des répondants individuels; 2) limiter la diffusion des données de mauvaise qualité (notion représentée par l'expression « qualité des données » dans la suite du présent document).

La confidentialité des données des répondants est gérée au moyen de l'arrondissement aléatoire, qui transforme toutes les estimations des dénombrements en chiffres arrondis de façon aléatoire en base 5. Selon cette technique, tous les chiffres de chaque tableau, y compris les totaux, sont arrondis de façon aléatoire à la hausse ou à la baisse

1. Un petit nombre d'enregistrements de la base de données du Couplage agriculture–population n'a pas été pondéré de la façon décrite dans le présent document : ces enregistrements ont automatiquement reçu le poids de 1. Il s'agit de ménages associés aux exploitations agricoles ayant des caractéristiques particulières.

pour obtenir un multiple de 5. Cette procédure assure une protection contre la divulgation sans ajouter d'erreur significative aux données. L'algorithme d'arrondissement aléatoire s'appuie sur une valeur de départ aléatoire pour générer la formule d'arrondissement pour les tableaux. Lorsque ces routines sont appliquées, la méthode utilisée pour générer la formule peut faire en sorte que les mêmes chiffres du même tableau soient arrondis à la hausse dans une exécution et à la baisse dans l'exécution suivante.

La qualité des estimations tirées de la base de données du Couplage agriculture–population sera décrite au moyen du coefficient de variation (c.v.). Le c.v. d'une estimation est le ratio de l'estimation de l'erreur-type et de l'estimation elle-même, exprimé en pourcentage. Plus le c.v. est faible, plus l'estimation est précise. Le c.v. est une mesure de variabilité intéressante étant donné qu'elle ne dépend pas de l'unité de mesure de l'estimation. Cela permet de comparer la précision des estimations ayant différentes unités de mesure. Il est important de noter que le c.v. tient compte seulement des erreurs d'échantillonnage et non des erreurs non dues à l'échantillonnage (comme une erreur causée par l'imputation).

Pour faciliter l'interprétation, dans les tableaux publiés, la valeur du c.v. est remplacée par une lettre. Plus précisément, les estimations dont le c.v. se situe entre 25,0 % à 49,9 % sont accompagnées de la lettre « E » en exposant pour indiquer à l'utilisateur de faire preuve de prudence au moment d'interpréter ces résultats, car ceux-ci présentent un niveau modéré de variabilité. Les estimations dont le c.v. est de 50,0 % et plus ne sont pas publiées en raison des préoccupations qu'elles soulèvent en ce qui concerne la qualité des données. Elles sont remplacées par la lettre « F » dans les tableaux.

Tableau 3
Intervalles du coefficient de variation et lettres associées utilisées pour les estimations de la base de données du Couplage agriculture–population

Lettre	Intervalle du coefficient de variation	Interprétation de la qualité des données
[Vide]	0 % à 24,9 %	Acceptable ou supérieur
^E (exposant)	25,0 % à 49,9 %	À utiliser avec prudence
F	50,0 % et plus	Trop peu fiable pour être publié