#### Cahiers québécois de démographie



#### Recommandations pour le recensement de 1981 : Mémoire préparé pour Statistique Canada par l'Association des démographes du Québec

Victor Piché

Volume 4, Number 2, 1975

URI: https://id.erudit.org/iderudit/305800ar DOI: https://doi.org/10.7202/305800ar

See table of contents

Publisher(s)

Association des démographes du Québec

**ISSN** 

0380-1721 (print) 1705-1495 (digital)

Explore this journal

Cite this note

Piché, V. (1975). Recommandations pour le recensement de 1981 : Mémoire préparé pour Statistique Canada par l'Association des démographes du Québec. *Cahiers québécois de démographie*, 4(2), 38–77. https://doi.org/10.7202/305800ar

Tous droits réservés  ${\mathbb C}$  Association des démographes du Québec, 1975

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

https://apropos.erudit.org/en/users/policy-on-use/



Cahiers québécois de démographie Vol. 4, no 2

### Recommandations pour le recensement de 1981

Mémoire préparé pour

Statistique Canada par l'Association

des démographes du Québec

(Rédigé par Victor Piché)

#### INTRODUCTION

Chaque recensement apporte avec lui des changements tant au niveau de la collecte que de la diffusion des données, changements qui sont perçus par les responsables comme des améliorations par rapport à ce qui se faisait antérieurement. Malheureusement, ces changements ne sont pas toujours perçus de la même façon par les usagers, surtout ceux qui oeuvrent aux niveaux provincial, municipal, etc. Par exemple, la technique d'arrondissement aléatoire peut très bien s'appliquer au niveau de toute la population du Canada tout en compliquant la vie des analystes de petites unités de population telles que les comtés ou les divisions administratives du Québec.

L'ADQ a donc décidé de mettre sur pied un comité ayant comme mandat de préparer un mémoire sur les problèmes d'utilisation du recensement de 1971 et sur les recommandations possibles pour celui de 1981. A cette fin, deux réunions ont eu lieu: une à Québec le 11 novembre et une autre à Montréal le 26 novembre 1974. C'est à partir de ces discussions et aussi des textes soumis par certains membres de l'ADQ que le présent rapport a été rédigé.

Dans une première partie, nous exposerons les problèmes principaux que rencontrent les utilisateurs du recensement de 1971. La deuxième partie sera consacrée aux recommandations en vue d'une meilleure préparation et utilisation du recensement de 1981. Enfin, nous présentons en annexe les textes soumis au comité du recensement de l'ADQ: il faut insister sur le fait que ces textes par leur importance font partie intégrante de ce rapport.

#### I. PRINCIPAUX PROBLEMES

#### A. Problèmes généraux

Il s'agit en fait d'une série de problèmes soulevés par l'ensemble des utilisateurs.

#### 1. Découpages géographiques

Le découpage spécial par le système "géocodage" a reçu beaucoup de critique: compte tenu de la publicité faite sur les avantages du géocodage, les résultats semblent décevants. Par
exemple, il a fallu 7 mois à un utilisateur pour obtenir la
structure par âge et la population totale d'une unité géographique quelconque - encore que ces données étaient entachées
de plusieurs erreurs. Il est de plus très difficile de passer
des divisions du recensement aux régions administratives du
Québec, comme l'indique le commentaire suivant:

"Enfin, nous contestons le contenu des publications reliées aux données du recensement, car en bien des cas, nous ne possédons que des résultats par division de recensement, où l'addition ne peut correspondre, dans la majorité des cas, au découpage administratif du Québec"(1)

#### 2. Délais

Les problèmes de délais sont très importants, surtout pour les utilisateurs qui oeuvrent dans des services administratifs et qui doivent répondre rapidement à des questions d'information statistique. En ce qui concerne le recensement de 1971, il y a eu des délais de publication inexplicables, sans parler des délais considérables à répondre à des demandes particulières. Certains ont dû en fait retarder leurs travaux à cause de ces délais. (2)

#### 3. Questionnaires

Plusieurs membres se sentent très gênés par les changements dans les définitions des concepts et des questions qui rendent très difficiles, parfois impossibles, les comparaisons des données d'un recensement à l'autre. De plus, la définition de certains concepts (v.g. main-d'oeuvre) n'est pas toujours pertinente lorsque l'on passe du cadre canadien au cadre québécois. Certains concepts tels l'état matrimonial et la langue maternelle ont requ une attention particulière. Par exemple, il semblerait que le mot "maternelle" a prêté à confusion de sorte que certaines personnes ont indiqué la langue maternelle de leur mère plutôt que la leur. C'est cette confusion qui pourrait expliquer la situation absurde de certains jeunes enfants (0-4 ans) dont la langue maternelle serait différente de la langue d'usage. (3)

<sup>(1)</sup> Voir annexe, le texte de José A. Carmona et Pierre Levasseur, p. 57 (2) Ibid. p. 57

<sup>(3)</sup> Voir en annexe, texte de Robert Maheu, pp. 52-53.

Jacques Henripin a également souligné certaines anomalies particulièrement en ce qui concerne le rapport entre l'origine ethnique et la langue maternelle. (4)

#### 4. Autres problèmes

Mentionnons, en terminant cette section, que des personnes ont soulevé certaines questions sur l'échantillonnage. Par exemple, jusqu'à quel point les données sont-elles représentatives lorsque l'on travaille au niveau régional, surtout lorsqu'il s'agit de petites régions? De plus, compte tenu que le questionnaire était envoyé par la poste, on se demande quel est le degré de sous-enregistrement des québécois. Enfin, question de polissage des données: on ne connaît pas toujours les procédures utilisées, etc.

#### B. Problème particulier: l'arrondissement aléatoire

Il semble que ce soit de loin le problème qui préoccupe le plus les utilisateurs. D'une part, le sentiment général penchait plutôt vers le scepticisme quant à la qualité des données du recensement de 1971. En fait, il semble que ce sont les données sommaires dont la qualité est remise en question d'une façon générale:

"...la qualité douteuse des données premières reçues via le Bureau de la Statistique du Québec, nous a obligé à nous adresser directement à Statistique Canada, parce que le B.S.Q. ne possède que les bandes sommaires et que les compilations, à partir de ces bandes, entraînent des difficultés majeures... (5)

<sup>(4)</sup> Jacques Henripin, l'immigration et le déséquilibre linguistique, étude réalisée pour le Ministère de la Main d'Oeuvre et de l'Immigration, 1974, pp. 41-44.

<sup>(5)</sup> Voir en annexe, le texte de Carmona et Levasseur, p.57

Donnons quelques exemples des difficultés rencontrées à ce sujet. Si on compare les données du B.S.Q. et de Statistique Canada (demande spéciale) pour certaines régions administratives, on arrive aux résultats suivants: (1)

- Au niveau de la population totale, les différences sont minimes.
- 2. La valeur des pourcentages varie rapidement d'un âge à l'autre, peu importe la région. Il n'y a donc aucune correction systématique possible.
- 3. Plus la région est petite, en population, plus les différences sont importantes au niveau de la distribution par âge. La cause peut être liée au fait que le nombre restreint de secteurs de dénombrement, dans une petite région, ne permet pas l'annulation des écarts dus aux arrondis telle que souhaitée par Statistique Canada.
- 4. C'est surtout aux âges élevés que les écarts deviennent pénibles à supportes.

Un autre problème porte sur la cohérence des données d'un tableau à l'autre. Donnons comme exemple le tableau tiré de l'étude de Carmona et Levasseur (p.58):

## POPULATION DU QUEBEC DISTRIBUEE PAR SEXE RECENSEMENT CANADIEN JUIN 1971

	Sexe masculin	Sexe féminin	Total
Catalogue 72-717	2,994,550	3,033,215	6,027,760
Catalogue 92-733	2,997,225	3,030,540	6,027,765

<sup>(1)</sup> Ibid, p. 59

Et leur commentaire judicieux: "Sûrement que Statistique Canada a de bonnes raisons de faire un tel transfert de sexes, mais c'est pour le moins gênant pour nous, qui avons à nous y référer". (p. 59). Pour un autre exemple d'incohérence, nous vous référons à l'étude de Louis Duchesne (en annexe).

Il semble bien que la qualité douteuse des données des bandes sommaires soit liée à la technique de l'arrondissement aléatoire. Selon cette technique, "tous les tableaux mis à la disposition du public sont arrondis à la base 5 et le sens vers lequel on arrondit est déterminé par le hasard plutôt que par une série explicite de règles". (1). L'arrondissement aléatoire est justifié pour assurer la confidentialité des données. Une excellente étude de Yolande Lavoie (voir annexe) montre bien les problèmes que rencontre l'utilisateur lorsqu'il veut travailler au niveau régional.

#### II. RECOMMANDATIONS

#### A. Diffusion et publication

- Que le système de diffusion soit suffisamment bien pourvu et équipé pour répondre adéquatement et dans des délais raisonnables aux demandes spéciales:
  - recoupements originaux de variables
  - découpages géographiques particuliers
  - etc...
- 2. Que les données du recensement soient validées à des niveaux très raffinés et que les utilisateurs soient informés du degré de fiabilité de l'information à mesure que celle-ci est disponible.

<sup>(1)</sup> voir en annexe texte de Yolande Lavoie, p. 67

- Que l'onformation transmise, quel qu'en soit le mode de diffusion, soit accompagnée de la documentation nécessaire à sa bonne interprétation.
- 4. Que dorénavant on s'en tienne aux groupes d'ages quinquennaux traditionnels qui facilitent l'analyse démographique.
- 5. Que dans tous les tableaux portant sur les subdivisions de recensement, on publie les données pour chacune des réserves indiennes et pour chacun des territoires non municipalisés.
- 6. Que soit publié, sur le même modèle que la publication de 1971 "langue selon le groupe ethnique", un bulletin sur la langue d'usage selon la langue maternelle, ceci pour faciliter l'étude des transferts linguistiques.

Nous aimerions voir aussi un tableau où la population (au moins au niveau des provinces) serait répartie selon l'état matrimonial, le sexe et <u>l'année d'âge</u>. Ceci, à notre avis, serait fort important pour les études sur la nuptialité ou sur la fécondité selon l'état matrimonial.

#### B. Enquêtes post et inter-censitaires

7. Qu'une série d'enquêtes sur des sujets particuliers (fécondité, nuptialité, etc.) viennent s'ajouter au recensement.

Pour répondre à de nouveaux besoins, on ne peut songer à
alourdir indéfiniment un questionnaire déjà long et à partir
duquel est publiée au bout de quelques années une information
déjà partiellement désuête.

#### 7.1 Fécondité

Le nombre très limité de questions posées au recensement sur la fécondité ne permet qu'une analyse très partielle de ce phénomène. Puisqu'il semble impossible d'y ajouter les questions nécessaires à une bonne analyse de la fécondité, il est proposé que la collecte des données concernant l'histoire rétrospective de la fécondité des femmes canadiennes se fasse dans un cadre autre que celui du recensement.

Une enquête par sondage serait en effet plus appropriée, parce qu'elle permettrait d'établir non seulement une bonne mesure de la fécondité mais également les liens entre cette variable et plusieurs autres de type socio-culturel et économique. De plus, elle pourrait être réalisée plus souvent qu'un recensement.

8. Que des enquêtes post-censitaires soient entreprises en vue de tester la validité des données. Nous croyons que la critique des données du recensement n'est pas assez poussée et, lorsqu'elle existe les résultats ne sont pas suffisamment connus.

#### C. Echantillonnage

9. Que les données provenant des enquêtes au tiers ou au cinquième soient recueillies de façon à assurer leur représentativité, non seulement au niveau national ou provincial, mais aussi au niveau régional, c'est-à-dire, des régions administratives et des unités urbaines de 15,000 habitants

et plus. Il faudrait pour faciliter les analyses, que les caractéristiques de l'échantillon correspondent à celles de la population totale de chaque unité territoriale.

#### D. Questionnaire

10. Contrairement à ce qui s'est fait en 1971, nous souhaitons qu'on ne considère pas comme "mariées" les personnes vivant en union consensuelle. Il faudrait s'en tenir à des catégories légales, par opposition à des catégories sociologiques.

Ce désir est fondé sur le lien qui doit exister entre les statistiques sur l'état de la population (recensement) et celles sur le mouvement de la population (état civil). L'état civil enregistre, ou pourrait enregistrer, des mariages, des nullités, des divorces ou des séparations légales. Mais par définition, il est exclu que la formation ou la dissolution d'une union consensuelle donne lieu à un enregistrement à l'état civil. Compte tenu de ce qui précède, nous recommandons qu'on retienne les catégories suivantes, accompagnées de définitions légales:

- célibataire;
- marié (e);
- veuf (ve);
- divorcé (e);
- séparé (e) légalement.

Si certains sont intéressés à connaître d'autres types de relations existant entre les personnes ceci pourrait être fait soit par une question distincte, soit par le cadre de la question sur le lien avec le chef de ménage.

- 11. Compte tenu de la confusion possible avec la question de la langue maternelle, que l'on étudie la possibilité de parler plutôt de "langue de l'enfance".
- 12. Il arrive pour certaines questions que l'on fournisse une liste de réponses: religions, langues ou origines. Nous recommandons qu'en 1981 ces réponses soient données selon l'ordre alphabétique français pour bien montrer que la langue
  française peut aussi, à l'occasion, avoir sa place au recensement canadien.
- 13. Si toutefois pour des raisons d'ordre pratique une enquête spéciale sur la fécondité n'était pas possible, nous considérons qu'au minimum une question distincte devrait être adressée à toutes les femmes, la question existante devant porter uniquement sur le nombre d'enfants issus du premier mariage. La première question permettrait donc l'étude de l'intensité de la fécondité générale par génération et la seconde l'étude de l'intensité de la fécondité légitime par promotion de premier mariage en tenant compte de la variable âge au mariage.

Les questions devant servir à l'analyse de la fécondité seraient libellées comme suit:

1. A toutes les femmes: combien d'enfants nés-vivants avez-vous eus?

- 2. A toutes les femmes mariées ou qui l'ont déjà été: quelle est la date de votre premier mariage?
- 3. A toutes les femmes mariées ou qui l'ont déjà été: combien d'enfants nés-vivants avez-vous eus lors de votre premier mariage?

Les tableaux publiés seraient:

Tableau A: Répartition des femmes selon le nombre d'enfants mis au monde par groupe quinquennal de généra-tions.

Tableau B: Répartition des femmes déjà mariées selon le nombre d'enfants issus du premier mariage par groupe quinquennal de promotions de premier mariage et par groupe quinquennal d'âge au mariage.

Les mêmes tableaux devraient être disponibles par génération et par promotion de premier mariage, et de plus cestableaux ne devraient pas contenir une catégorie semi-ouverte (10 enfants et plus) mais le nombre exact d'enfants.

#### E. Collecte

14. Les agents recenseurs devraient être recrutés auprès des centres de main-d'oeuvre. La formation des agents recenseurs devrait être plus poussée.

#### F. Décentralisation

15. Que le traitement des données soit décentralisé. Ainsi,

le B.S.Q. pourrait avoir accès aux données réelles et pourrait lui-même assurer la confidentialité et procéder aux arrondissement voulus.

#### G. Consultation

- 16. Que s'établisse une consultation permanente entre le B.S.Q. et Statistique Canada sur toutes les étapes du recensement: préparation des questionnaires, collecte, publication, etc.
- 17. Que l'on respecte les régions administratives du Québec dans la compilation des données, ceci, quelle que soit la source de l'information: bulletins, bandes d'ordinateur, microfilms, imprimés d'ordinateur, microfiches, etc.

#### H. Arrondissement aléatoire

- 18. Que l'argument du "secret statistique" soit repensé surtout pour certaines variables comme l'âge, le sexe, etc.
- 19. Que l'arrondissement aléatoire soit supprimé, de façon à ce qu'on puisse agréger en toute sécurité des données raffinées ou alors que les bandes sommaires soient transmises sans arrondissement, mais accompagnées du programme d'arrondissement aléatoire qu'on n'appliquerait qu'après avoir agrégé les données au niveau voulu.

- 20. Cette expérience avec le recensement de 1971 nous amène à proposer un certain nombre de corrections au procédé actuel ou tout simplement le rejet de la technique de l'arrondissement aléatoire. Parmi les solutions aux problèmes créés par l'arrondissement aléatoire, mentionnons (dans le cas où on le conserve):
  - l'accès direct des services de diffusion provinciaux aux données du recensement - les services provinciaux contrôlent alors l'arrondissement et les délais de production;
  - 2. La préparation, par Statistique Canada, de bandes d'ordinateur incluant au moins les différents niveaux de découpage géographique les plus utilisés - cette solution suppose une enquête auprès des utilisateurs -, on éviterait ainsi d'agréger des séries de données déjà arrondies;
  - 3. La mise à la disposition des services de diffusion provinciaux des données brutes du recensement et des programmes de traitement des données et d'arrondissement aléatoire.

Il nous semble important que des moyens d'éliminer les ennuis créés par l'arrondissement aléatoire soient étudiés et mis au point avant que les plans de traitement des données du recensement de 1981 soient définitivement arrêtés.

#### I. Priorités

Nous recommandons que le problème majeur, celui de l'arrondissement aléatoire et de la qualité douteuse des données sommaires, soit sérieusement pris en considération en consultation avec les institutions québécoises de recherches démographiques, surtout le B.S.Q. De toute façon, nous pensons que la <u>décentralisation</u> demeure la seule solution aux nombreuses difficultés que rencontrent les utilisateurs québécois. Nous souhaitons que le recensement de 1981 sera <u>vraiment</u> le résultat d'une oeuvre collective, ce qui implique, pour le Québec, que le B.S.Q. <u>soit</u> mêlé à toutes les décisions majeures concernant le prochain recensement.

#### III. ANNEXES: ETUDES DE CAS

- A. Suggestions pour le recensement de 1981 par Robert Maheu
  - 1. Formulaires et directives
    - 1.1 Etat matrimonial

Contrairement à ce qui s'est fait en 1971, nous souhaitons qu'on ne considère pas comme "mariées" les personnes vivant en union consensuelle. Il faudrait s'en tenir à des catégories légales, par opposition à des catégories sociologiques.

Ce désir est fondé sur le lien qui doit exister entre les statistiques sur l'état de la population (recensement) et celles sur le mouvement de la population (état civil). L'état civil enregistre, ou pourrait enregistrer, des mariages, des nullités, des divorces ou des séparations légales. Mais par définition, il est exclu que la formation ou la dissolution d'une union consensuelle donne lieu à un enregistrement à l'état civil. Compte tenu de ce qui précède. nous

recommandons que retienne les catégories suivantes, accompagnées de définition légales:

- célibataire;
- marié (e);
- veuf (ve);
- divorcé (e);
- séparé (e) légalement.

Si certains son intéressés à connaître d'autres types de relations existant entre les personnes ceci pourrait être fait soit par une question distincte, soit dans le cadre de la question sur le lien avec le chef de ménage.

#### 1.2 Langue maternelle

Nous avons certaines raisons de penser que le mot "maternelle" a prêté à confusion dans le passé et que certaines personnes ont indiqué la langue maternelle de leur mère plutôt que la leur. Ainsi on trouve dans certaines provinces des différences importantes dans le groupe d'âge 0-4 ans entre le nombre de personnes qui ont le français comme langue maternelle et celles qui ont le français comme langue d'usage.

Chez de si jeunes enfants, une telle situation semble absurde et relève plus probablement d'une mauvaise qualité des données que de la réalité. Une petite enquête faite auprès d'un nombre très restreint de personnes laisse croire que certaines personnes pensent que l'expression "langue maternelle" signifie "langue de votre mère". Dès lors, l'explication fournie sur le formulaire du recensement de 1971 immédiatement

sous "langue maternelle": "La première langue parlée et encore comprise", a pu faire croire à certaines personnes qu'on demandait la première langue parlée par leur mère. Si notre explication est la bonne, on pourrait reformuler la question sur la langue maternelle pour parler plutôt de "langue de votre enfance".

#### 1.3 Enumération de réponses

Il arrive pour certaines questions que l'on fournisse une liste de réponses: religions, langues ou origines. Nous recommandons qu'en 1981 ces réponses soient données selon l'ordre alphabétique français pour bien montrer que la langue française peut aussi, à l'occasion, avoir sa place au recensement canadien.

#### 1.4 Questions sur la fécondité

Le nombre très limité de questions posées au recensement sur la fécondité ne permet qu'une analyse très partielle de ce phénomène (1). Puisqu'il semble impossible d'y ajouter les questions nécessaires à une bonne analyse de la fécondité (2), il est proposé que la collecte des données concernant l'histoire rétrospective de la fécondité des femmes canadiennes se fasse dans un cadre autre que celui du recensement.

<sup>(1)</sup> La principale faiblesse étant qu'aucune donnée n'est recueillie sur le calendrier de la fécondité.

<sup>(2)</sup> Ces questions devraient porter sur le nombre d'enfants et la date de naissance de ces enfants, la date du ou des mariages, et la date de la ou des ruptures d'union par veuvage, divorce ou séparation légale.

Une enquête par sondage serait en effet plus appropriée, parce qu'elle permettrait d'établir non seulement une bonne mesure de la fécondité mais également les liens entre cette variable et plusieurs autres de type socioculturel et économique. De plus, elle pourrait être réalisée plus souvent qu'un recensement.

Si toutefois, pour des raisons d'ordre pratique, cette solution n'était pas acceptée, nous considérons qu'au
minimum une question distincte devrait être adressée
à toutes les femmes, la question existante devant porter
uniquement sur le nombre d'enfants issus du premier
mariage. La première question permettrait donc l'étude de l'intensité de la fécondité générale par génération et la seconde l'étude de l'intensité de la fécondité légitime par promotion du premier mariage en tenant
compte de la variable âge au mariage.

Les questions devant servir à l'analyse de la fécondité seraient libellées comme suit:

- A toutes les femmes: combien d'enfants nés-vivants avezvous eus?
- 2. A toutes les femmes mariées ou qui l'ont déjà été: quelle est la date de votre premier mariage?
- 3. A toutes les femmes mariées ou qui l'ont déjà été: combien d'enfants nés-vivants avez-vous eus lors de votre premier mariage?

Les tableaux publiés seraient:

Tableau A: Répartition des femmes selon le nombre d'enfants mis au monde par groupe quinquennal de générations.

Tableau B: Répartition des femmes déjà mariées selon le nombre d'enfants issus du premier mariage par groupe quinquennal de promotions de premier mariage et par groupe quinquennal d'âge au mariage.

Les mêmes tableaux devraient être disponibles par génération et par promotion de premier mariage, et de plus ces tableaux ne devraient pas contenir une catégorie semi-ouverte (10 enfants et plus) mais le nombre exact d'enfants.

#### II. Collecte

Nous considérons que des relations privilégiées avec un député ou un de ses organisateurs ne sont pas un substitut à la compétence. Les agents recenseurs devraient plutôt être recrutés auprès des centres de main-d'oeuvre. La formation des agents recenseurs devrait être plus poussée.

#### III. Tableaux

#### III.1 Groupe d'âge

Plusieurs tableaux ont été publiés selon les groupes d'âge suivants dans le passé:

0- 4		35-44
5- 9		45-54
10-14		55-64
15-19	24	65-69
20-24		70 et +
25-34		III

Nous souhaitons qu'on s'en tienne dorénavant aux groupes d'âge quinquennaux traditionnels qui facilitent l'analyse démographique.

III.2 Réserves indiennes et territoires non municipalisés.

Nous aimerions que dans tous les tableaux portant sur les sub-

divisions de recensement, on publie les données pour chacune des réservec indiennes et pour chacun des territoires non municipalisés.

#### III.3 Nouveaux tableaux

On devrait publier, sur le même modèle que la publication de 1971 "langue selon le groupe ethnique", un bulletin sur la langue d'usage selon la langue maternelle, ceci pour faciliter l'étude des transferts linguistiques.

Nous aimerions aussi un tableau où la population (au moins au niveau des provinces) serait répartie selon l'état matrimonial, le sexe et <u>l'année d'âge</u>. Ceci, à notre avis, serait fort important pour les études sur la nuptialité ou sur la fécondité selon l'état matrimonial.

B. Recensement de Juin 1971 par Josée A. Carmona et Pierre Levasseur C'est en mon nom personnel et en celui de l'un des membres de l'A.D.Q., soit monsieur José A. Carmona, que je soulèverai quelques difficultés rencontrées par l'utilisation des données du dernier recensement, étant donné que le ministère des Affaires sociales (notre employeur) a déjà soumis un rapport à cet effet aux responsables du recensement canadien. Nous nous contentons d'exposer quelques-uns des multiples problèmes rencontrés, vu le peù de temps mis à notre disposition.

Que ce soit pour obtenir une population de départ pour nos perspectives démographiques oupour compiler des statistiques devant faire l'objet de publication, nous sommes dépendants de la qualité des données fournies par le recensement.

Avec le recensement de 1971, plusieurs problèmes se sont présen-

tés à nous. Premièrement, nous avons dû retarder certains de nos travaux, à cause des délais de publications des statistiques fournies par le recensement. Deuxièmement, la qualité douteuse des données premières reçues via le Bureau de la Statistique du Québec, nous a obligé à nous adresser directement à Statistique Canada, parce que le B.S.Q. ne possède que les bandes sommaires et que les compilations, à partir de ces bandes, entraînent des difficultés majeures, comme nous le verrons plus loin. Troisièmement, l'hétérogénéité des résultats relatifs à des données qui devraient nous offrir une certaine similitude, nous cause également des problèmes importants. Nous donnerons, plus loin, des exemples des difficultés que ce dernier fait nous a amenées. Enfin, nous contestons le contenu des publications reliées aux données du recensement, car en bien des cas, nous ne possédons que des résultats par division de recensement, où l'addition ne peut correspondre, dans la majorité des cas, au découpage administratif du Québec.

Nous donnerons maintenant deux exemples de difficultés que nous avons rencontrées lors de l'élaboration de nos travaux.

C'est ainsique la plupart des démographes oeuvrant au Québec ont eu ou auront à travailler avec les données relatives à la population des régions administratives ou socio-sanitaires distribuée par âge et par sexe. Par contre, nous arrivons à deux résultats différents, selon que l'on possède des statistiques provenant du B.S.Q., et donc tirées des bandes sommaires, ou que l'on ait en main des données fournies par Statistique Canada,

grâce à une demande spéciale. Cette dernière se réfère à la sommation des populations, dont les éléments ne sont pas arrondis, répartie par âge, par sexe et par secteur de dénombrement. L'arrondissement aléatoire nécessaire à la confidentialité est effectué seulement au niveau de la région administrative ou sociosanitaire.

Nous nous sommes "amusés" à comparer les résultats (sexe masculin) de nos deux sources pour quatre régions socio-sanitaires:

- A) Bas St-Laurent/Gaspésie (niveau de 100,000 h.)
- B) Québec (niveau de 500,000 h.)
- C) Montréal Métropolitain (niveau de 1,000,000 h.)
- D) Laurentide Lanaudière (niveau de 200,000 h.)

Les résultats de nos recherches ont été portés au tableau 1, où nous pouvons facilement vérifier les quelques remarques suivantes:

- Au niveau de la population totale, les différences sont minimes.
- 2. La valeur des pourcentages varie rapidement d'un âge à l'autre, peu importe la région. Il n'y a donc aucune correction systématique possible.
- 3. Plus la région est petite, en population, plus les différences sont importantes au niveau de la distribution par âge. La cause peut être liée au fait que le nombre restreint de secteurs de dénombrement, dans une petite région, ne permet pas l'annulation des écarts dus aux arrondis telle que souhaitée par Statistique Canada.

4. C'est surtout aux âges élevés que les écarts deviennent pénibles à supporter.

Passon maintenant à un second point que nous aimerions soulever. Pour ce faire, nous nous référons aux catalogues 92-717 et 92-733 publiés par Statistique Canada.

# TABLEAU 2 POPULATION DU QUEBEC DISTRIBUEE PAR SEXE RECENSEMENT CANADIEN JUIN 1971

	Sexe masculin	Sexe féminin	Total
Catalogue 72-7171	2,994,550	3,033,215	6,027,760
Catalogue 92-733 <sup>2</sup>	2,997,225	3,030,540	6,027,765

Sûrement que Statistique Canada a de bonnes raisons de faire un tel transfert de sexes, mais c'est pour le moins gênant pour nous, qui avons à nous y référer.

l Population par état matrimonial

<sup>2</sup> Population selon la langue, par groupe d'âge

# TABLEAU I TABLEAU COMPARATIF DE LA POPULATION DE CERTAINES REGIONS DU QUEBEC (JUIN 1971)

## SEXE MASCULIN (POURCENTAGES)

AGE	BAS ST-LAURENT	QUEBEC	MONTREAL METRO	LAURENTIDES
0	0,560	-0,342	-0,381	-0,341
1	0,501	-0,602	-0,063	-2,031
2	0,500	-0,997	0,032	4,314
3	1,909	0,786	1,350	-1,189
4	-0,216	2,006	1,040	0,266
5	1,370	0,605	-0,437	0,119
6	-1,920	0,492	-0,870	0,215
<b>7</b> 8	-1,727	1,796	0,372	0,213
8	-2,754	0,558	-0,311	0,607
9	-0,161	0,328	0,189	0,522
10	0,316	0,277	0,142	0,392
11	0,472	0,462	0,095	-0,100
12	0,467	-0,093	-0,311	0,497
13	1,495	-0,910	-0,360	0,615
14	-0,597	-0,593	-0,639	-0,208
15	1,236	-0,784	0,099	0,666
16	0,000	-0,412	-0,244	0,222
17	0,940	-0,096	0,890	0,838
18	1,596	0,857	-0,313	1,013
19	-0,958	-0,791	0,575	-0,552
20	-0,217	-0,269	-0,432	-2,564
21	-0,690	-0,111	<b>-1,</b> 238	-1,264

(1) Résultats portés sur ce tableau=  $\frac{A - B}{A}$  X 100

où

- A = Population, distribuée par age et par sexe, obtenue de Statistique Canada en effectuant les arrondis au niveau de la région.
- B = Population, distribuée par âge et par sexe, obtenue du B.S.Q. où les arrondis ont été faits au niveau du secteur de dénombrement. Les résultats par région proviennent de l'addition des arrondis.

# TABLEAU I (suite)

\$	AGE	
	BAS ST-LAURENT	
nonoof+74744000000000000000000000000000000000	QUEBEC	
1000001146 1000001146 1000001146 1000001146 1000001146 1000001146 1000001146 1000001146 1000001146 1000001146 1000001146 10000001146 1000000001146 10000000000	MONTREAL METRO	
1515200240000000000000000000000000000000	LAURENTIDES	

TABLEAU I (suite)

AGE	BAS ST-LAURENT	QUEBEC	MONTREAL METRO	LAURENTIDES
66 67 68 69 70 71 72 73 74 75 77 78	-8,730 -6,140 -2,778 -2,128 -2,222 4,167 15,854 10,256 16,883 2,500 16,393 1,786 14,583 -2,326	1,224 2,412 -3,687 0,000 0,548 -2,571 -2,736 -2,694 -2,273 1,961 3,084 2,577 7,303 0,000	-1,240 -2,752 0,286 -1,259 3,175 -1,264 0,840 -3,030 3,871 0,181 -2,222 1,129 5,897 2,849	-5,603 -4,327 3,448 0,000 6,579 -4,375 -0,709 -1,399 -9,402 -4,237 -4,000 -4,124 6,250 -4,478
80+ TOTAL	-3,020 0,348	5,316 0,016	0,002	1,050 -0,221

Des exemples de cet acabit pourraient se multiplier, mais nous nous arrêterons pour l'instant à celui-ci. Par contre, nous sommes limités dans nos comparaisons par le fait que nous ne possédons que la donnée arrondie et non la donnée exacte. Seul, Statistique Canada pourrait effectuer cette comparaison et ils seraient à même de constater les lacunes de leur système.

Pour terminer, nous souhaitons que ces difficultés liées à la conversion géographique et à la technique d'arrondissement aléatoire, qui nous empêchent de faire un travail valable, soient corrigées pour au moins le recensement de 1981 et d'ici là, nous continuerons à propager des statistiques de mauvaises qualité.

C. Certaines caractéristiques de la population féminine du Québec selon 3 bulletins de recensement, 1971 par Louis Duchesne.

Commentaires sur le tableau ci-dessous.

Les écarts entre les séries provenant de l'échantillon et les séries "officielles" sont parfois nuls et parfois assez importants.

Exemple: Population féminine, tout état matrimonial

65-69 ans 86 895 et 86 890

20-24 ans 277 425 et 276 475

Ces variations ne gênent pas, en général, l'analyse des résultats.

Toutefois, dans certains cas, il faut repondérer au complet toutes les données publiées. Par exemple, si l'on veut calculer, à partir du bulletin 1-5-10 (92-750) portant sur l'age au premier mariage, le pourcentage de femmes non célibataires à 20 ans, pour les femmes âgées de 40-44 ans en 1971, il faut ramener les 160 300 femmes du bulletin 92-750 au nombre "officiel" de 160 970 et donc repondérer les effectifs des âges au mariage puisque l'on divisera le nombre de femmes déjà mariées avant 20 ans, non pas par le nombre de femmes déjà mariées provenant de l'échantillon mais par le nombre de femmes de tout état matrimonial provenant des données "officielles". Il faut donc, pour faire certaines analyses, repondérer toutes les données publiées dans le bulletin 92-750.

## Certaines caractéristiques de la population féminine d'après 3 bulletins de recensement.

BULLETIN:	92-730			92 <b>-7</b> 50	92-731
Groupe d'âges	tout état matrimonial	Célibataires	Déjà mariées	Déjà mariées	Tout état matrimonial
9- 4 5- 9 10-14	(234,325) (309,565) (329,105)	0==			234,215 309,590 328,945
0-14 15-19 20-24 25-29	872,995 306,260 277,425 238,605	872,990 293,685 147,510 48,610	12,575 129,915 189,995	12,545 129,735 190,020	305,815 2 <b>7</b> 6,475
25 <b>-</b> 34 30 <b>-</b> 34 35 <b>-</b> 39	427,130 188,525 180,620	24,625 20,390	163,900 160,230	163,960 160,990	426,600
35-44 40-44 45-49	360,820 180,200 172,440	19,230 18,785	160,970 153,655	160,300 153,490	360,680
45 <b>-</b> 54 50 <b>-</b> 54 55 <b>-</b> 59 55 <b>-</b> 64	317,715 145,275 130,000 238,420	17,550 18,200	127,725 111,800	127,930 111,765	317,700 238,620
6 <b>9</b> -64 65-69 70-74	108,420 86,895 63,445	16,765 13,930 10,070	91,655 72,965 53,375	91,775 73,100	86,890
70 75-79 80-84 85-89 90-94 95	42,405 24,715 10,990 3,190 815	6,850 3,890 1,640 490 130	35,555 20,825 9,350 2,700 685	121,855	145,010
TOTAL	3,033,215	1,535,335	1,497,880	1,497,470	3,030,540

92-750 et 92-731: Résultats provenant de l'échantillon. Les écarts importants sont encadrés. D. Les implications de l'arrondissement aléatoire dans l'exploitation des données du recensement: un exemple par Yolande Lavoie

Statistique Canada met à la disposition des "utilisateurs" des tableaux sommaires sur bande d'ordinateur, afin qu'ils puissent en tirer les analyses de leur choix. C'est ainsi que les services de diffusion provinciaux<sup>(1)</sup> ont pu se procurer les bandes sommaires Al DEM OO1 et Al DEM OO2 présentant, sous une forme abrégée (OO1) ou sous une forme détaillée (OO2), les caractéristiques démographiques de la population canadienne suivant les secteurs de dénombrement. Ces secteurs sont ainsi définis: "Unité géographique recensée par un représentant du recensement et pouvant comprendre jusqu'à 300 ménages, mais toujours moins de 100 fermes, c'est-à-dire une population d'environ 500 personnes".

Il ne fait donc pas de doute que peu d'utilisateurs, qu'ils soient chercheurs, planificateurs ou technocrates, font directement usage de données présentées suivant ce découpage géographique très fin. En général, les utilisateurs agrégeront les secteurs de dénombrement pour reconstituer les unités géographiques qui conviennent à leurs besoins.

L'utilisation des bandes sommaires et l'arrondissement aléatoire
Selon la technique de l'arrondissement aléatoire, tous les

<sup>(1)</sup> Au Québec: Bureau de la Statistique du Québec, division de la diffusion des données du recensement.

tableaux mis à la disposition du public sont arrondis à la base 5 et le sens vers lequel on arrondit est déterminé par le hasard plutôt que par une série explicite de règles. Le chiffre 5 d'un tableau indique que la valeur réelle se situe entre 1 et 9. (1)

Imaginons maintenant la façon dont se présentera un tableau répartissant 500 personnes environ suivant le sexe, puis, à l'intérieur de cette première division, suivant l'âge (année d'âge). Il est évident que peu de cellules contiendront des nombres supérieurs à 10 puisqu'une telle distribution répartit les 500 personnes dans quelque 200 cellules. Qu'on se représente un tel tableau après qu'il aura été soumis à l'arrondissement aléatoire dont "l'influence sur chaque cellule de tableau est inversement proportionnelle à la taille de cette cellule "(2). Il est évident que n'y apparaîtront que les 3 ou 4 premiers multiples de "5" et des "0"

Un tel tableau ne permet de calculer ni une distribution relative suivant le sexe et l'âge (qu'on songe à l'allure de la pyramide des âges qui en découlerait), ni des rapports de masculinité pour ne citer que les calculs les plus simples et les plus courants.

<sup>(1) &</sup>quot;L'arrondissement aléatoire: Technique de protection contre la divulgation illégale de renseignements dans les tableaux de recensement", Statistique Canada, Division du recensement, note de recherche préparée par E. M. Murphy, 2 août 1972, 11 p.

<sup>(2)</sup> ibid, p. 11

Or, si Statistique Canada met les bandes sommaires à la disposition du public par le truchement des services de diffusion provinciaux, c'est qu'il juge que ces bandes constituent une banque d'information plus détaillée et plus raffinée que le contenu des brochures du recensement. Il arrive en effet que la reconstitution de certaines unités géographiques ne soit possible qu'à partir des secteurs de dénombrement qui constituent le niveau le plus fin de découpage territorial.

Voyons ce qui se passera si, à partir de la bande Al DEM 002 (secteurs de dénombrement et année d'âge), on veut reconstituer la population d'un comté, en l'occurence les Iles-de-la Madeleine, et la présenter sous la forme d'un tableau distribuant les effectifs suivant le sexe et le groupe quinquennal d'âges.

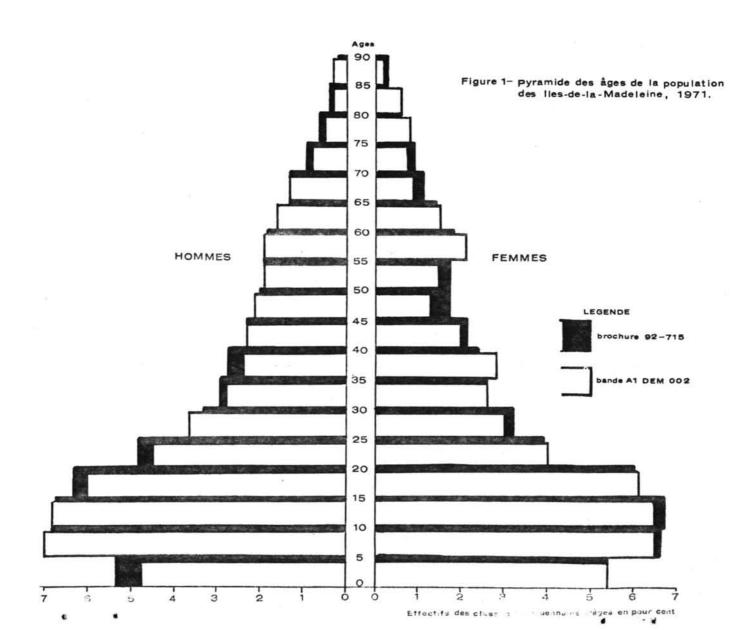
Rappelons-nous que l'arrondissement a été fait au niveau des âges individuels, par sexe, pour chacun des secteurs de dénombrement.

Le tableau 1 présente d'une part, les données provenant de la bande Al DEM 002 et d'autre part, celles qui sont contenues dans la publication de recensement "Population. Groupes d'â-ge. cat. 92-715, vol: 1- part: 2" où 1'arrondissement a été fait au niveau du comté, par sexe et groupe quinquennal d'âges. Quant au tableau 2, il contient les rapports de masculinité correspondant à chacune des distributions du tableau 1.

TABLEAU I

Iles-de-la-Madeleine: Population suivant le sexe et le groupe d'âges quinquennal. Recensement de 1971.

Groupes d' <b>â</b> ges	В	ande Al	DEM 002		Pu	blicati	on 92-715	
	Nombres absolus			Pour 100 de po- pulation totale				Pour 100 de po- pulation totale
	Н	F	Н	F	Н	F	Н	F
0-4 5-9 10-14 15-19 20-24 25-29 30-34 35-39 40-49 50-54 50-64 65-69 70-79 80-89 90-99 100+	680 935 910 805 605 480 380 315 275 250 255 215 180 105 60 35 35 15	715 870 865 815 535 395 345 270 170 205 285 195 110 80 20	5.1 7.0 8.0 5.6 8.4 7.6 8.4 7.6 8.4 7.6 8.4 7.6 7.6 7.6 7.6 7.6 7.6 7.6 7.6	5.6.5.1.0.0.6.8.0 5.6.5.1.5.9.8.6.2.1 5.6.6.2.2.1.5.9.8.6.2.1 5.6.6.2.1.5.9.8.6.2.1	710 905 895 845 845 355 300 255 240 170 115 60 30 155 -	665 875 890 805 515 425 320 225 225 235 150 120 835 155 -	5.87383973 0983396421	5.6.70 92.6 41 77.8 41.96.331
	6,850	6,485	51.4	48.6	6,885	6,435	51.7	48.3
Grand total "officiel"	6,865	6,440			6,870	6,435	41	



<u>Tableau 2</u>

Rapports de masculinité

Iles-de-la-Madeleine, 1971

Ensemble	100 100 100 100 100 100 100 100 100 100	Groupes d'âges
106.6	107.5 107.5 110.1	Bande Al DEM 002
106.8	103.8 103.8 110.6 110.6 110.6 110.6 100.6 100.6 100.6 100.6	Publication 92-715

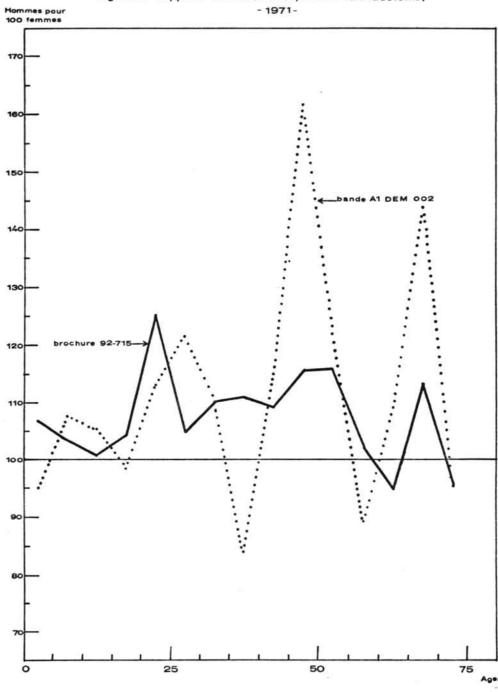


Figure 2- Rapports de masculinité, lles-de-la-Madeleine,

D'énormes distorsions apparaissent dans la distribution qui résulte de l'agrégation des secteurs de dénombrement de la bande Al DEM OO2. La pyramide des âges dressée à partir des données du tableau l'manifeste bien ces distorsions de la structure par âge et par sexe (figure 1). De même la figure 2 indique que la répartition suivant le sexe est faussée et que l'interprétation qu'on en aurait pu faire concernant la migration différentielle selon le sexe ou concernant les marchés matrimoniaux serait loin de traduire la réalité.

L'importance des écarts entre la série Al DEM 002 et la série 92-715 figure au tableau 3.

<u>Tableau 3</u>: Ecarts en nombres absolus et en nombres relatifs entre les résultats de la bande Al DEM 002 et ceux de la brochure 92-715 Iles-de-la-Madeleine, 1971.

Groupe d'ages	Nombres	absolus	Nombres relatifs (%)		
	н	F	н	F	
0- 4 5- 9 10-14 15-19 20-24 25-29 30-34 35-39 40-44 45-49 50-54 55-59 60-64 65-69 70-74 75-79 80-84 85-89 90-94	-30 +30 +35 -35 -40 +35 -40 +15 -15 +15 +10 -15 -15 +10 -25 +0	+50 -25 +10 +20 -30 -55 -55 -55 -150 +20 -55 -150 +155 -155 -155 -155 -155 -155 -155 -155	- 4324681136269690270 - 4324681136269690270	+ 8 - 1 - 3 + 1 + 7 - 1 + 17 - 24 - 27 + 21 + 17 - 8 + 18 + 129 - 43 - 33	

Le sens des écarts, tout comme celui de l'arrondissement. est aléatoire. Dans plus de la moitié des cas, les écarts entre les deux séries sont supérieurs à 5%. Si on excepte les 70 ans et plus, dont les effectifs sont très faibles, on constate que les différences entre la bande et la brochure dépassent dans quelques cas 20%. Il est évident que l'analyse des données provenant de l'agrégation des secteurs de dénombrement, préalablement soumis à l'arrondissement aléatoire, constitue une entreprise pleine d'embûches. En effet, si dans notre exemple et c'est volontaire, nous disposons d'une série de données justes, il n'en serait pas de même si nous avions reconstitué non plus un comté, mais un bassin hydrographique, une région administrative, un district judiciaire ou toute autre unité territoriale pour laquelle le recensement ne publie pas de statistiques. Nous serions alors privés d'un terme de comparaison permettant de juger de la validité des données.

Il ressort de ce test sur un comté, les Iles-de-la-Madeleine, que les bandes sommaires sont d'une utilité douteuse quand
on doit reconstituer des territoires en agrégeant des secteurs
de dénombrement. Il est dommage que cet effort du Service de diffusion des données soit si peu rentable uniquement à cause de
l'arrondissement aléatoire.

L'exemple que nous avons choisi, le comté des Iles-de-la-Madeleine, n'a rien d'un cas exceptionnel. Cependant, du fait de la population restreinte de ce comté, les écarts entre la série Al DEM 002 et la brochure 92-715 prennent une importance relative considérable. Il ne faut toutefois pas oublier que nous avons souvent à reconstituer des ensembles de cette taille.

D'autre part, même au niveau de l'ensemble du Québec, les distorsions entre les deux séries sont gênantes. Les écarts, à certains âges, ne sont pas négligeables (voir tableau 4). La somme des groupes d'âges s'écartent sensiblement du total provenant de la sommation des totaux des secteurs de dénombrement (+1,905 pour le sexe féminin); le calcul des distributions relatives s'en trouve faussé. De plus, on possède pour le Québec autant de chiffres de population totale et de répartitions de la population qu'il y a de niveaux d'arrondissement. Il s'ensuit qu'on doit multiplier les mises en garde, les notes explicatives, les estimations des marges d'erreur dans toutes les études faites à partir des données préalablement arrondies.

Les citoyens canadiens peuvent dormir tranquilles, même les astuces des plus rusés d'entre les utilisateurs des données du recensement ne sauraient arracher aux tableaux provenant du recensement de 1971 le moindre indice permettant d'identifier un individu. Par contre, Statistique Canada s'expose à une avalanche de demandes spéciales dont le résultat, nous l'expérimentons actuellement, est d'étendre les délais entre la formulation de la demande et la réception des résultats. Dans un deuxième temps, on assistera peutêtre à un désintéressement croissant des utilisateurs qui se seront résignés à user d'expédients pour arriver à satisfaire leurs besoins.

TABLEAU 4:

Ecarts en nombres absolus et en nombres relatifs entre les résultats de la bande Al DEM 002 et ceux de la brochure 92-715. Ensemble du Québec, 1971

	NOMBRES AB	SOLUS	POUR CENT		
Groupes d'Ages	н	F	н	F	
0- 4 5- 9 10-14 15-19 20-24 25-29 30-34 35-39 40-49 50-54 55-59 60-69 70-74 75-79 80-84 85-89 90-94	-310 + 80 +835 -1,030 - 20 - 45 + 270 - 635 - 540 + 795 + 305 + 360 - 125 + 115 + 485 + 355 - 495 + 15 	- 15 +250 +910 -195 -305 -105 +500 +540 -110 -320 + 65 +745 + 25 + 50 -630 +585 + 65 + 50 - 50 - 50	-0.13 +0.02 +0.24 -0.33 -0.01 -0.02 +0.14 -0.35 -0.30 +0.48 +0.22 +0.29 -0.13 +0.16 -0.98 +1.14 -2.92 +0.21	+0.01 +0.08 +0.28 -0.06 -0.11 -0.04 +0.27 +0.30 -0.06 -0.19 +0.04 +0.57 +0.02 +0.06 -0.99 +1.38 +0.26 +0.45 -1.57	
Total "offi- ciel"	- 300	+ 15	-0.01	; <del></del> :	
≤des groupes d'âges	+ 415	+1,905	+0.01	+0.06	

Cette expérience avec le recensement de 1971 nous amène à proposer un certain nombre de corrections au procédé actuel ou tout simplement le rejet de la technique de l'arrondissement aléatoire. Parmi les solutions aux problèmes créés par l'arrondissement aléatoire, mentionnons, dans le cas où on le conserve:

- l'accès direct des services de diffusion provinciaux aux données du recensement — les services provinciaux contrôlent alors l'arrondissement et les délais de production —;
- 2. la préparation, par Statistique Canada, de bandes d'ordinateur incluant au moins les différents niveaux de découpage géographique les plus utilisés — cette solution suppose une enquête auprès des utilisateurs —, on éviterait ainsi d'agréger des séries de données déjà arrondies;
- 3. la mise à la disposition des services de diffusion provinciaux des données brutes du recensement et des programmes de traitement des données et d'arrondissement aléatoire.

Il nous semble important que des moyens d'éliminer les ennuis créés par l'arrondissement aléatoire soient étudiés et mis au point avant que les plans de traitement des données du recensement de 1976 soient définitivement arrêtés.