



N° 11-522-XIF au catalogue

**La série des symposiums internationaux
de Statistique Canada - Recueil**

Symposium 2003 : Défis reliés à la réalisation d'enquêtes pour la prochaine décennie

2003



Statistique
Canada

Statistics
Canada

Canada

Recueil du Symposium 2003 de Statistique Canada
Défis reliés à la réalisation d'enquêtes pour la prochaine décennie

RECHERCHE ET ANALYSE POUR MIEUX COMPRENDRE LES ACTIVITÉS DE COLLECTE DES DONNÉES

Judy Lee et Tony LaBillois¹

RÉSUMÉ

Il a toujours été difficile de quantifier l'efficacité des activités de collecte des données et difficile également de produire de façon normalisée et en temps opportun de l'information de gestion utile sur de nombreuses enquêtes de manière à améliorer les processus de collecte. Depuis l'avènement de l'interview assistée par ordinateur et la normalisation des logiciels utilisés pour recueillir les données d'enquête, cependant, on dispose d'une foule de renseignements sur le processus de collecte des données. Le système BLAISE enregistre les activités des intervieweurs qui entrent dans les cas et en sortent durant le cycle de collecte. Statistique Canada a créé le Groupe de travail chargé de l'analyse et de la recherche qui coordonne et dirige divers projets portant sur des questions clés liées à la collecte des données et élabore des outils de gestion de la collecte aux fins de planification. Le présent document expose la portée, les objectifs et les défis de la recherche et de l'analyse portant sur les opérations à Statistique Canada et donne certains exemples des travaux accomplis jusqu'ici.

MOTS CLÉS : Amélioration du processus, collecte de données, interview assistée par ordinateur.

1. INTRODUCTION

Il y a quelques années, Statistique Canada a consolidé ses ressources et ses activités liées à l'élaboration des systèmes d'interview assistée par ordinateur (IAO). En outre, le Bureau a décidé d'adopter un logiciel unique devant servir à l'élaboration de toutes les nouvelles applications IAO et à la conversion des applications existantes. BLAISE² a été choisi à cette fin et la normalisation s'est amorcée. On a également affecté au même moment les ressources nécessaires pour élargir les travaux de recherche et d'analyse liés aux activités de collecte des données.

Auparavant, les gestionnaires de la collecte pouvaient consulter de nombreux rapports de gestion portant sur une enquête donnée ou un type d'enquête donné pour évaluer les activités de collecte. Toutefois, seul un petit nombre d'études spéciales étaient fondées sur des renseignements quantitatifs sur les activités de collecte tandis que de nombreux renseignements non scientifiques et commentaires fondés sur l'expérience du personnel de gestion de la collecte étaient difficiles à vérifier, ou bien il était difficile d'en tirer des généralisations fondées sur des faits clairs donnant un aperçu réel du processus de collecte.

L'affectation de ressources aux travaux de recherche et d'analyse des activités de collecte des données, l'effort de normalisation et l'intérêt croissant des gestionnaires à obtenir plus de renseignements quantitatifs utiles aux fins de la prise de décisions touchant les activités de collecte sont autant de facteurs qui ont contribué à la création et au succès du Groupe de travail chargé de l'analyse et de la recherche. Il est vite devenu évident qu'on pouvait en apprendre beaucoup en comprenant et utilisant mieux ce type de renseignement.

¹ Judy Lee et Tony Labillois, Statistique Canada, Ottawa (Ontario), Canada, K1A 0T6. Judy.Lee@statcan.ca, Tony.Labillois@statcan.ca

² BLAISE est un progiciel de collecte de traitement de données d'enquête élaboré par le Bureau de la statistique des Pays-Bas et distribué par Westat en Amérique du Nord. Pour plus de renseignements, consulter les sites Web indiqués dans la bibliographie.

Le présent document comprend une brève description de la portée et des objectifs du programme de recherche sur les opérations, des sources de données utilisées aux fins de l'analyse et quelques exemples de projets examinés jusqu'ici. Enfin, nous y examinons les défis et les priorités pour l'avenir rapproché.

2. PORTÉE ET OBJECTIFS DU PROGRAMME DE RECHERCHE SUR LES OPÉRATIONS

Statistique Canada prend des mesures de contrôle de la qualité et d'assurance de la qualité depuis de nombreuses années. On procède à la surveillance des intervieweurs durant les activités de collecte ITAO dans les emplacements de collecte de données au moyen de méthodes d'échantillonnage et de contrôle de la qualité et les résultats sont compilés dans le Système de rétroaction du contrôle de la qualité (SRCQ) afin d'améliorer la qualité des données et d'aider à la formation de l'intervieweur. La portée de l'évaluation de nos opérations de collecte a été élargie considérablement il y a trois ans de manière à inclure la capacité d'évaluer de nombreux aspects de nos activités pour toutes nos enquêtes. Ces aspects sont liés à la disponibilité récente de données sur la collecte de données.

Le Groupe de travail vise principalement à recueillir et à compiler des renseignements quantitatifs sur différentes composantes du processus de collecte. Les données peuvent porter sur de nombreux aspects du processus : vérification et suivi à l'étape de collecte de données, historique des appels, nombre d'accès à un cas, raisons et résultats de l'accès à un cas, durée des interviews, nombre d'interruptions et de répétitions durant une question, etc.

En deuxième lieu, le Groupe de travail vise à fournir des outils aux gestionnaires de la collecte pour les aider à gérer les activités de collecte de données. Les gestionnaires auraient accès non seulement aux rapports d'étape réguliers sur la collecte, mais, à leur ordinateur de bureau, à une série d'outils dynamiques souples qui leur permettent d'examiner divers aspects du processus, y compris les caractéristiques du rendement des « intervieweurs ».

Le Groupe de travail chargé de l'analyse et de la recherche s'occupe également de projets de recherche particuliers pour montrer les fins auxquelles peuvent servir les renseignements relatifs à la collecte et pour répondre aux questions posées par les gestionnaires des opérations ou les représentants de domaines spécialisés. En outre, l'examen et l'analyse des données soulèvent de nouvelles questions et indiquent de nouveaux sujets de recherche. Ces activités suscitent de l'intérêt concernant le programme et leurs résultats engendrent souvent des discussions à propos des pratiques actuelles et des changements qui ont lieu dans l'environnement de production ou aboutissent à l'inclusion du nouveau processus d'examen des renseignements dans les autres activités de collecte de données d'enquête.

Nos clients comprennent les cadres supérieurs qui souhaitent réaliser des économies, les gestionnaires de la collecte qui veulent s'informer du rendement de leur personnel et les spécialistes qui ont besoin de savoir comment leur argent est dépensé ou si l'enchaînement des questions dans leur questionnaire est logique. Le programme aide à cerner les faits qui ont une incidence sur l'élaboration des systèmes de collecte à l'avenir, sur les procédures ou activités de collecte et sur les budgets alloués à ces tâches.

Bref, l'accent est mis sur l'optimisation de l'effort de collecte dans toutes nos enquêtes sur l'agriculture et auprès des entreprises ou des ménages et dans chacun de nos neuf emplacements de collecte de données. Le but visé est de produire des renseignements qui permettront d'améliorer les activités de collecte.

3. SOURCES DE DONNÉES

Pour procéder à l'analyse des activités de collecte de données, il faut disposer de fichiers de données fiables, actuelles, cohérentes, normalisées et d'accès facile. Étant donné les centaines d'enquêtes menées chaque année auprès des entreprises, des exploitations agricoles ou des ménages dans de nombreux emplacements de collecte, accumuler tous ces renseignements au fil du temps pour toutes les périodes de référence présente toujours un défi. Comme Statistique Canada est un organisme de statistique très centralisé et comme l'élaboration de nos applications assistées par ordinateur est centralisée dans la même division où se trouve le Groupe de travail chargé de l'analyse et

de la recherche, il est possible de coordonner ces activités en établissant des processus clairs, des normes et de la documentation. Il faut déployer des efforts constants pour veiller à ce que ces règles soient respectées et pour évaluer la qualité des sources de données de manière à garantir la disponibilité de fichiers qui sont « appropriés pour l'analyse ».

Jusqu'ici, nous utilisons principalement les fichiers de l'historique des transactions BLAISE (BLAISE Transaction History ou BTH) pour nos projets. Durant la période de collecte d'une enquête, chaque fois qu'un intervieweur accède à un cas au moyen de l'application IAO, l'opération est enregistrée dans le fichier BTH. Cet enregistrement contient une série de variables³ dont la date de la tentative d'appel, l'heure de la tentative d'appel, la durée, le code de résultat et le numéro de l'intervieweur. Ces données peuvent servir plus tard à la rédaction de rapports d'étape ou à l'examen de divers aspects du déroulement de la collecte. Par exemple, nous pouvons examiner le nombre de fois où l'on a eu accès à un cas et le temps consacré à un cas donné

En établissant les applications BLAISE, on peut activer une fonction piste de vérification. Lorsque cette fonction est activée, un bref enregistrement est produit pour chaque entrée faite durant la session. En combinant ces enregistrements, on peut déterminer, par exemple, le temps consacré en moyenne à remplir une partie donnée du questionnaire ou le nombre de fois où les intervieweurs ont dû revenir à une question donnée, ou encore surveiller le déroulement des interviews afin de détecter toute falsification éventuelle de réponses.

Pour certains projets, nous devons combiner les données liées à la collecte à celles provenant d'autres sources. Ces sources peuvent comprendre les fichiers des échantillons ou les fichiers de données d'enquête combinés aux fichiers BTH lorsqu'il s'agit de déterminer l'incidence éventuelle des caractéristiques des répondants sur la collecte des données. Les fichiers de contrôle des documents ou les fichiers de la saisie de données sur les cas qui sont réintégrés dans l'application IAO aux fins d'examen et de suivi éventuel sont utiles lorsqu'il s'agit d'évaluer l'exécution des vérifications dans le processus de collecte.

Nous cherchons actuellement la meilleure façon de procéder pour stocker les données découlant des activités de collecte et y accéder facilement. Il s'agit ultérieurement d'élaborer un modèle durable pour le répertoire des activités de collecte. Nous avons également élaboré des prototypes de systèmes de rapport et d'analyse des activités de collecte pour les enquêtes menées auprès des entreprises et des ménages. Ces prototypes élaborés en SAS/AF sont souvent utilisés pour répondre à des demandes ponctuelles spéciales et pour faire certaines analyses. Ultérieurement, il s'agit d'élaborer un outil très souple, facile à utiliser et opportun à l'intention du personnel du Groupe de travail ainsi que d'autres personnes qui participent à la gestion d'activités de collecte. Il reste encore beaucoup à faire sur le répertoire et les systèmes avant d'en arriver à ce point.

4. EXEMPLES DE RÉSULTATS

Malgré une infrastructure de recherche sur la collecte en évolution et un programme de recherche encore récent, nous avons entrepris de nombreux projets d'analyse dont certains ont révélé des changements importants dans nos processus de collecte de données ou ont entraîné de tels changements. Nous visons toujours à répondre aux questions d'actualité ou à fournir des éclaircissements sur de nouvelles questions. Ces projets aident à souligner le potentiel et la pertinence d'un programme de recherche et d'analyse consacré aux secteurs des opérations. Nous en donnons trois exemples ci-dessous.

Examen de l'effort de collecte de données pour l'Enquête annuelle des manufactures (EAM)

Dans le cadre d'une vaste initiative d'examen et de rationalisation lancée en 2002, Statistique Canada a créé un comité de rationalisation des enquêtes-entreprises chargé de recueillir des renseignements sur les coûts et les processus pour chaque étape de deux enquêtes-entreprises choisies comme études de cas. Le comité devait rédiger

³ Les fichiers BTH diffèrent légèrement selon le type d'enquête. On trouvera des renseignements sur la structure et le contenu des fichiers BTH dans les documents énumérés dans la bibliographie.

un rapport détaillé⁴ et visait à recueillir des renseignements quantitatifs sur les activités de collecte de l'Enquête annuelle des manufactures (EAM), puisqu'elles représentaient la plus grande partie du budget total de l'enquête.

Pour l'année de référence (AR) 2001, on a mis en oeuvre une fonction de score dans le cadre de certaines nouvelles procédures pour cerner les unités critiques. Cette fonction attribue un score aux unités de collecte selon leur incidence sur les estimations de l'année précédente. On accorde une priorité plus élevée en matière de suivi à certaines unités et une priorité moins élevée à d'autres, jugées moins importantes.

À l'aide des renseignements dans les fichiers BTH, nous avons pu comparer les activités enregistrées dans le système de collecte pour l'AR2000 et l'AR2001. Il importe de se rappeler que les activités telles que l'envoi des trousseaux par la poste, la préparation des questionnaires et le balayage optique ne sont pas enregistrées dans les systèmes. Il a donc été très important de discuter avec les gestionnaires de la collecte pour garantir que les résultats sont interprétés correctement et situés dans le contexte voulu.

Nous avons constaté une diminution de 40 % (plus de 8 000) des cas suivis en l'AR2001 comparativement à l'AR2000, malgré la taille semblable des échantillons des deux années. En l'AR2001, seulement 60 % des cas ont fait l'objet d'un suivi comparativement à 97 % en l'AR2000. En termes d'effort ou d'accès aux cas, l'effort a été réduit de 36 % en l'AR2001, ce qui s'est traduit par une réduction de temps de 40 % ou une réduction de ressources de quatre personnes par an.

On peut ventiler le nombre total d'accès et le temps total consacré selon leurs composantes : collecte ou suivi avant la réception d'un questionnaire, vérification ou suivi après la réception d'un questionnaire ou aucune donnée reçue, (où l'on a fait un suivi mais aucun questionnaire n'a été reçu). En comparant la distribution de ces composantes pour les deux années, nous avons constaté une augmentation de 3,5 et 4,9 points de pourcentage, respectivement, de l'effort de collecte et de vérification et une diminution de 9 points de pourcentage de l'effort associé à aucune donnée reçue. Cette diminution de l'effort qui aboutit à aucune donnée reçue s'est traduite par une diminution de temps de 14 points de pourcentage. Bref, nous réduisons notre « effort gaspillé » et nous améliorons nos efforts de collecte et de vérification.

Stratégie de vérification des données pour l'Enquête unifiée auprès des entreprises (EUE)

L'EUE inclut maintenant la plupart des enquêtes annuelles auprès des entreprises de Statistique Canada qui sont toutes intégrées dans un environnement unique de collecte et de traitement. Depuis l'AR2000, lorsque les questionnaires sont reçus, les données sont saisies dans une application appelée saisie rapide des données qui n'a pas de capacité de vérification. Les données sont ensuite chargées dans une application BLAISE pour examen, vérification et suivi éventuel. En comparant le fichier des données de la saisie rapide des données et le fichier des données sorties de la saisie rapide des données, et en utilisant les fichiers et les indicateurs BLAISE, nous pouvons surveiller les rejets à la vérification. Nous voulons analyser le rendement des règles de vérification et les taux de suivi ainsi que l'effet sur les estimations de l'enquête des changements dus à la vérification.

Les résultats de ce projet permanent ont entraîné le contrôle de la qualité du processus de saisie rapide des données. L'analyse des différentes tâches que comprend la vérification est fort intéressante : examen du taux de confirmation des données existantes et analyse au fil du temps à compter de l'AR2001, entre autres. En outre, les spécialistes ont entrepris un examen des règles de vérification pour les périodes de collecte futures, afin de tâcher de réduire le nombre de règles et d'ajuster certains paramètres.

Par exemple, nous avons constaté que, pour les enquêtes menées auprès des industries de services, le nombre de vérifications a été réduit, passant de 96 en l'AR2000 à 59 en l'AR2001 et à 51 en l'AR2002. En même temps, au cours de ces trois années, les taux de suivi (cas rejetés à au moins une vérification grave ou obligatoire et devant faire l'objet d'un suivi) ont diminué, passant de 89 % à 73 % et à 59 %, respectivement.

Les rejets à la vérification entraînent une vérification des données qui exige beaucoup d'effort et de temps. Nous sommes maintenant en mesure de surveiller les résultats de ces efforts de vérification. Par exemple, dans le cas des Services de réparation et d'entretien, en l'AR2001, 55 % des questionnaires ont été rejetés à au moins une

⁴ On trouvera plus de détails dans le rapport intitulé « Strategic Streamlining Initiative, Business Surveys Review ».

vérification obligatoire et 29 %, à au moins une vérification grave, ce qui a entraîné 1 240 et 2 658 vérifications, respectivement. Dans les cas des vérifications obligatoires, 61 % des cas ont été résolus, 7 % sont restés non résolus et, dans le cas d'un suivi auprès des répondants, 27 % de ces derniers ont répondu « ne sais pas » et 5 % ont « refusé » de valider les données douteuses. Dans le cas des vérifications graves, 50 % des cas ont été résolus et le suivi auprès des répondants a donné 34 % de réponses « ne sait pas », 3 % de « refus » et 13 % de confirmations des données douteuses. Les vérifications visent à réduire le nombre de confirmations et autres vérifications non résolues.

On peut procéder à ce type d'analyse pour tous les types de questionnaire et au niveau de la vérification dans le but de cerner les vérifications et les industries où le suivi des vérifications a entraîné peu de modification des données. En outre, la prochaine étape de l'analyse consisterait à déterminer l'ampleur de la modification des données dans le cas de modifications résultant du suivi d'une vérification.

Profils de l'achèvement des cas pour l'Enquête sur la dynamique du travail et le revenu (EDTR)

L'EDTR est une enquête longitudinale menée auprès des ménages qui a commencé dans les années 1990 avec deux périodes de collecte annuelles distinctes pour les variables de travail et les variables de revenu. Étant donné le grand échantillon et les coûts croissants de collecte, les spécialistes de cette enquête s'intéressent vivement à savoir comment on pourrait accroître l'efficacité des activités de collecte. Nous avons examiné jusqu'ici divers aspects de cette enquête liés à la collecte et voici certaines de nos constatations.

Nous nous sommes penchés, entre autres, sur le moment où les cas sont les plus susceptibles d'être résolus ou de se voir attribuer un code de résultat « final ». Si ce profil s'observe durant tout le cycle de production de l'enquête ou entre les cycles, alors on pourrait prévoir les opérations de collecte selon le scénario du meilleur moment. Par exemple, pour l'EDTR 2003, Revenu (AR2002), le nombre moyen de cas résolus par intervieweur était le plus élevé le samedi et la période la plus productive, cette journée, se situe entre 12 heures et 15 heures. Toutefois, le moment de la journée le plus productif pour la résolution des cas est le soir, entre 18 heures et 23 heures tous les jours de la semaine à l'exception du samedi. Le dimanche soir est le pire moment pour appeler dans l'espoir de finaliser et de résoudre un cas.

On a manifesté beaucoup d'intérêt pour l'effort que comprend la conversion des cas de refus pour les enquêtes-ménages. Pour l'EDTR 2003, Revenu (AR2002), le nombre moyen d'appels par cas pour convertir un refus en réponse complète allait de 2,9 à Halifax et à Sherbrooke à 4,2 à Toronto et à 5,5 à Edmonton. La conversion d'un refus en réponse complète a pris en moyenne 14,8 minutes à Sherbrooke, 16,4 minutes à Halifax, 18 minutes à Edmonton et 23,2 minutes à Toronto. Toutefois, lorsque nous examinons les taux de conversion, les préposés à Edmonton ont été les plus efficaces en ce qu'ils ont pu convertir 60 % de leurs refus en réponses complètes, suivis de Sherbrooke à 47 %, de Halifax à 41 % et de Toronto à 31 %.

Enfin, nous disposons maintenant de données recueillies au fil du temps nous permettant de comparer un cycle à l'autre. Par exemple, en comparant les années de collecte de l'EDTR, Travail et Revenu selon l'effort ou selon le nombre d'accès ou d'appels pour obtenir des données utilisables (codes de résultat, réponses complètes ou partielles) nous constatons que plus de 80 % des données utilisables ont été obtenues au quinzième appel. Au-delà de 15 tentatives d'appel, on obtient un autre 5 % à 10 % de données utilisables.

5. DÉFIS

La création d'une véritable équipe de recherche et d'analyse portant sur les activités de collecte commence à peine à porter fruit. Ce volet de nos activités semble voué à un avenir plein de défis, mais prometteur. Nous examinons dans la présente section certains des défis⁵ que le programme sera appelé à relever à l'avenir.

Le besoin de consolider notre infrastructure d'information est à la base des activités à venir. Nous devons veiller à ce qu'elle puisse intégrer avec succès des données tirées d'un nombre croissant d'enquêtes pour un nombre croissant

⁵ On trouvera une description plus détaillée des défis à relever dans le document intitulé « Quarterly Program Review, Operations Research and Development Division ».

de périodes de collecte ainsi qu'inclure des données d'enquête actuelles aux fins de couplage. Nous examinons différentes façons de construire cet énorme répertoire de données tout en améliorant nos systèmes d'extraction de données.

En outre, nous devons maintenir et accroître notre capacité et nos compétences spécialisées afin d'apporter des éclaircissements sur des questions particulières en réponse à des demandes spéciales. Même si notre équipe est très petite, nous devons assurer un apprentissage continu et veiller au partage des produits pertinents avec nos partenaires. Nous aiderons ainsi à maintenir l'intérêt manifesté actuellement pour les constatations et les questions sur lesquelles nous avons travaillé jusqu'ici. Il s'agit là d'un facteur clé susceptible de garantir la survie et la croissance de ce programme.

Étant donné notre vaste éventail d'utilisateurs et de clients, nous devons coordonner les exigences des utilisateurs pour déterminer les outils, les rapports périodiques et les indicateurs normalisés pertinents. Nous pouvons aider sur le plan de la coordination, mais nous avons besoin de la participation et de l'appui des utilisateurs pour leur fournir toutes les analyses dont ils ont besoin. À notre avis, nous devons susciter et maintenir leur intérêt afin de pouvoir examiner de près les processus et prendre des décisions sur les mesures concrètes qui permettront d'améliorer les activités de collecte ou de réaliser des économies à cet égard.

Les résultats des recherches sur les opérations doivent servir à améliorer les processus de collecte de façon générale. Ils ne sont pas destinés à permettre éventuellement aux spécialistes de faire de la microgestion. Les préposés aux opérations doivent apprendre à utiliser les outils pour accroître l'efficacité, communiquer leurs résultats et montrer qu'ils réagissent de la façon appropriée en optimisant leur utilisation des ressources disponibles.

RÉFÉRENCES

<http://www.cbs.nl/en/service/blaise/introduction.htm> (The BLAISE system homepage)

<http://www.westat.com/blaise/> (BLAISE overview)

<http://www.blaiseusers.org/> (International BLAISE User Group)

Blaise Transaction History Files for Business Surveys RY2001, RY2002, Operations Research and Development Division, Statistics Canada (internal document).

Blaise Transaction History File Requirements for Social Surveys, Operations Research and Development Division, Statistics Canada (internal document).

LaBillois, T. The integration of Electronic Data Reporting (EDR) and other new systems in the ongoing collection process, AMRADS Workshop on Automatic Data Collection, Rome, Italy, May 2002.

<http://amrads.jrc.cec.eu.int/>

Royce, D., Labillois, T., Leduc, J., Lepage J., Marcogliese, M., Zinnicker, D. – Strategic Streamlining Initiative, Business Surveys Review – Interim Report, January 9, 2003, Statistics Canada, internal document.

Quarterly program review of Operations Research and Development Division – April 2000 to November 2003 – November 2003, Statistics Canada (internal document).