

MODERNISATION DU PROGRAMME DE LA STATISTIQUE DU TOURISME DE STATISTIQUE CANADA

Keven Bosa¹ & François Gagnon²

¹ *Statistique Canada, 100 promenade Tunney's Pasture, Ottawa, Ontario, K1A 0T6 et
keven.bosa@canada.ca*

² *Statistique Canada, 100 promenade Tunney's Pasture, Ottawa, Ontario, K1A 0T6 et
francois.gagnon2@canada.ca*

Résumé. La modernisation du Programme de la statistique du tourisme de Statistique Canada repose sur les cinq piliers suivants : (i) la prestation de services axés sur l'utilisateur; (ii) l'utilisation de méthodes d'intégration des données à la fine pointe de la technologie; (iii) le renforcement des capacités statistiques et leadership; (iv) le partage et la collaboration et (v) l'effectif moderne et le milieu de travail flexible. En plus des enquêtes par sondage, de nombreuses avenues innovatrices sont actuellement explorées afin de moderniser le Programme en fonction de ces cinq piliers. Parmi elles, notons l'utilisation des données de paiement et de cartes de crédit afin d'estimer les dépenses reliées aux voyages; l'utilisation de l'intelligence artificielle et des données volumineuses provenant des réseaux sociaux ou des compagnies de téléphonie mobile pour recueillir de l'information sur les voyages; l'utilisation accrue des données administratives provenant de l'Agence des services frontaliers du Canada; et la collaboration avec les organismes fédéraux, provinciaux et municipaux reliés au tourisme afin d'unir nos efforts pour la cueillette de données sur le tourisme. La présentation décrira le progrès réalisé par Statistique Canada sur chacune de ces initiatives et discutera du potentiel de chacune de ces avenues pour produire des statistiques de qualité, disponibles rapidement et qui répondront aux besoins des utilisateurs.

Mots-clés. Tourisme, données volumineuses et données administratives

Abstract. Modernization of Statistics Canada's Tourism Statistics Program is based on the following five pillars: (i) providing user-oriented service delivery; (ii) using state-of-the-art data integration methods; (iii) bolstering statistical capacity and leadership; (iv) sharing and collaboration, and (v) ensuring a modern workforce and flexible workplace. In addition to sample surveys, a number of innovative avenues are being explored to modernize the Program based on these pillars. Some of avenues include using payment and credit card data to estimate travel-related expenses; using artificial intelligence and big data from social media and mobile telephone companies to collect information on trips; increasing use of data from the Canada Border Services Agency; and, joining forces with federal, provincial and municipal tourism organizations to collect data on tourism. This presentation will describe the progress Statistics Canada has made in each of these initiatives and explore the potential of each avenue to produce quality, timely statistics that meet user needs.

Keywords. Tourism, big data and administrative data

1. Introduction

Statistique Canada évalue actuellement différentes options afin de moderniser et améliorer le Programme de la statistique du tourisme. La modernisation repose sur les cinq piliers suivants : (i) la prestation de services axés sur l'utilisateur; (ii) l'utilisation de méthodes d'intégration des données à la fine pointe de la technologie; (iii) le renforcement des capacités statistiques et leadership; (iv) le partage et la collaboration et (v) l'effectif moderne et le milieu de travail flexible. Dans ce court article, des améliorations au programme du tourisme ainsi que l'utilisation d'approches non traditionnelles et innovantes seront abordées.

Le programme actuel sera brièvement décrit à la Section 2. Par la suite, deux nouvelles approches considérées pour améliorer les statistiques sur le tourisme seront présentées à la Section 3. Une brève conclusion et les travaux futurs forment la dernière partie de ce document.

2. Description du programme actuel de tourisme

Le Programme de la statistique du tourisme à Statistiques Canada comprend deux composantes principales : le dénombrement aux frontières et deux enquêtes trimestrielles. Le Programme produit plusieurs estimations dont les plus importantes sont le nombre de voyageurs, le nombre de voyages, le nombre de nuitées ainsi que les dépenses associées aux voyages. L'information reliée au tourisme est principalement utilisée à Statistique Canada pour estimer des recettes et des paiements pour le compte des voyages de la balance des paiements et pour tenir à jour le Compte satellite du tourisme. Le Compte satellite du tourisme permet de mesurer l'importance des activités touristiques dans l'économie canadienne et de les comparer à d'autres activités économiques au Canada.

2.1 Dénombrement aux frontières

Cette partie du Programme de la statistique du tourisme utilise différentes sources de données administratives provenant de l'Agence des services frontaliers du Canada (ASFC). Ces données représentant le dénombrement aux frontières doivent être compilées afin de diffuser mensuellement le nombre total de voyages internationaux. Il y a deux types de voyages internationaux : les voyages faits par les résidents canadiens à l'extérieur du pays ; et les voyages des visiteurs étrangers au Canada. L'ASFC modernise également ses processus et Statistique Canada doit travailler en collaboration avec ce partenaire fédéral pour comprendre le changement dans les données et continuer de produire les mêmes estimations. Par exemple, la carte de déclaration que doit remplir voyageur international (Canadien ou étranger) entrant au Canada par voie aérienne à bord d'un transporteur commercial a été remplacée par une déclaration électronique sur un kiosque pour certains aéroports en 2017. Statistique Canada a dû modifier les systèmes en place pour traiter cette nouvelle source de données. Le fait que l'information provenant de ces données électroniques n'est pas exactement la même que celle obtenue des cartes de déclaration constitue un des principaux défis pour ce programme. Certaines améliorations ont été apportées récemment dont la production d'estimations mensuelles préliminaires diffusées à peine quelques jours suivant le mois de référence pour les voyageurs arrivant des États-Unis par automobile. Lorsque deux sources de données de l'ASFC sont disponibles, Statistique Canada favorise l'utilisation de données électroniques lorsque c'est possible ce qui fait en sorte d'accélérer la saisie et le traitement des données.

2.2 Enquête sur les voyages des visiteurs

L'Enquête sur les voyages des visiteurs (EVV) a débuté officiellement en janvier 2018 et c'est l'une des deux enquêtes trimestrielles sur le tourisme menées par Statistique Canada. Cette enquête remplace l'Enquête sur les voyages internationaux (EVI) qui s'est terminée en 2017. La différence principale entre les deux enquêtes est que les voyages faits par les Canadiens à l'international sont maintenant couverts par une autre enquête, soit l'Enquête nationale sur les voyages qui sera présentée à la section 2.3. L'EVV recueille de l'information sur les visiteurs internationaux ayant

fait un voyage au Canada. Cette enquête a deux méthodes de collecte : la distribution d'une carte d'invitation à participer à l'enquête en ligne ainsi que des interviews en personne dans le cadre de la composante sur les départs aériens (CDA). Pour le premier mode de collecte, les agents de l'ASFC remettent les cartes d'invitation à remplir le questionnaire en ligne aux visiteurs qui entrent au pays. La carte invite les voyageurs à remplir le questionnaire en ligne à la fin de leur voyage en utilisant l'adresse web fournie sur la carte. Statistique Canada détermine le nombre de cartes que reçoit chaque bureau de l'ASFC et la date à laquelle ces derniers doivent distribuer les cartes basées sur les données de l'année précédente. Des questionnaires supplémentaires sont obtenus par le biais de la CDA pour laquelle les intervieweurs de Statistique Canada interrogent en personne les voyageurs en attente de leur vol de retour dans certains pays visés. Ces interviews sont effectuées chaque mois et la période de collecte dure de 5 à 7 jours.

2.3 Enquête nationale sur les voyages

L'Enquête nationale sur les voyages (ENV) est une autre enquête trimestrielle qui a débuté officiellement en 2018. Elle remplace l'Enquête sur les voyages des résidents du Canada ainsi que la partie canadienne de l'EVI qui se sont terminées en 2017. L'ENV recueille de l'information sur les voyages domestiques et internationaux faits par les Canadiens âgés de 18 ans ou plus. Chaque mois, un échantillon de ménage est tiré et ces derniers reçoivent une lettre d'invitation à participer à l'enquête en ligne. Une des raisons d'être de l'ENV est l'amélioration souhaitée des estimations reliées aux voyages internationaux faits par les Canadiens. Contrairement à l'EVI, une base de sondage est disponible et Statistique Canada peut mener la collecte de façon autonome ce qui devrait augmenter la couverture de ce type de voyage. Avant l'implantation de cette enquête, deux enquêtes tests ont eu lieu afin de mettre à l'essai différentes versions du questionnaire ainsi que différentes méthodologies en utilisant des plans d'expérience.

Une première enquête pilote de l'ENV a eu lieu en 2016 où le contenu de l'enquête ainsi que différentes méthodes de sélection d'une personne à l'intérieur d'un ménage ont été mises à l'essai. Les différentes méthodes de sélection ont été mises à l'essai en utilisant un plan d'expérience en blocs randomisés. Plus de détails sur le plan d'expérience sont disponibles dans Poirier et Bosa (2017) ainsi que Vézina et Caron (2017). La méthode d'ordre d'âge des membres du ménage développée à Statistique Canada, voir Vézina et Caron (2017), est maintenant utilisée pour choisir le répondant de l'ENV. Les directives sont incluses dans la lettre d'invitation de sorte que le ménage peut ainsi identifier la personne du ménage qui a été sélectionnée. Auparavant, à Statistique Canada, la méthode traditionnelle de sélection aléatoire d'une personne pour les enquêtes en ligne auprès des ménages consistait à faire ladite sélection dans le questionnaire après que l'information contenant l'énumération des membres du ménage ait été obtenue. Un des avantages principaux de la méthode basée sur l'âge est qu'un seul répondant doit remplir le questionnaire en ligne contrairement à la méthode traditionnelle ou un premier répondant devait compléter l'énumération des membres de son ménage sur le questionnaire avant de demander à une autre personne de terminer le questionnaire – si une personne différente était aléatoirement choisie.

Une deuxième enquête pilote a eu lieu durant l'été 2017 où l'objectif principal était de mettre à l'essai les changements apportés au questionnaire et différentes stratégies de suivi en présence de non-réponse. Un plan d'expérience en blocs randomisés a été intégré à cette enquête afin de comparer différentes stratégies de collecte en présence de non-réponse. Trois différentes actions en présence de non-réponse ont été mises à l'essai : envoyer un rappel par la poste ; téléphoner au ménage et lui demander de remplir le questionnaire en ligne ; et téléphoner au ménage pour remplir le questionnaire au moyen d'une interview téléphonique assistée par ordinateur (ITAO). Six différentes stratégies de suivis auprès des non-répondants pouvant combiner deux actions de façon séquentielles ont été mises à l'essai. Un des faits saillants de cette expérience est le fait qu'appeler le répondant pour lui demander de remplir le questionnaire en ligne est moins efficace que

d'envoyer une lettre de rappel par la poste. L'option de conduire une ITAO est la meilleure option pour augmenter le taux de réponse, mais à des coûts plus élevés. Doering et coll. (2018) décrivent le plan d'expérience et présentent les faits saillants de cette étude. Pour l'ENV, il a été décidé d'envoyer un rappel par la poste pour convertir les non-répondants. Plus tard pendant la période de collecte, un échantillon des non-répondants restants est tiré dans le but de compléter ces cas à l'aide d'une ITAO. Cette décision a été prise afin de contenir le plus possible le biais potentiel dû à la non-réponse tout en respectant le budget de l'enquête.

3. Amélioration du Programme de tourisme

3.1 Données de traitement de paiements

Statistique Canada, en collaboration avec Destination Canada, explore présentement le potentiel des données de traitement de paiements reliées à des cartes de crédit et de débit émises par une institution à l'extérieur du Canada. L'objectif est de déterminer si ces données, en combinaison avec d'autres sources de données alternatives ou d'enquêtes permettraient d'améliorer les estimés de dépenses de voyage effectuées au Canada par les étrangers. Statistique Canada a actuellement accès aux données de deux fournisseurs ce qui ne couvre pas tous les paiements faits au Canada. Les données contiennent les transactions ayant été enregistrées par les terminaux des fournisseurs. Les données reçues sont agrégées et le Tableau 1 contient un exemple montrant le niveau de détail de ces dernières. Les colonnes «Code associé au marchand», «Code postal», «pays» et «Date» représentent respectivement la catégorie auquel le marchand est associé, le niveau géographique représentant les trois premiers caractères du code postal de l'emplacement du marchand, le pays duquel la carte a été émise et le mois pour lequel les transactions ont été effectuées. La variable «Présence de la carte» indique si la carte était présente (1) ou non (0) lors des transactions. Finalement, les variables «Nombre de transactions» et «Total des dépenses» sont respectivement le nombre de transactions et les dépenses nettes.

Code associé au marchand	Code postal	Pays	Date	Présence de la carte	Nombre de transactions	Total des dépenses
Location de voiture	L2A	US	2015-07	1	xx	xxxx
Divertissement	V3S	CH	2016-10	0	xx	xxxxxxx
Hotel	G6T	DE	2017-06	1	xx	xxxxxx
Épiceries	M6P	GB	2017-03	1	xx	xxxxxx
Essence	S7R	US	2016-09	1	xxx	xxxxx

Tableau 1 : Exemple de données de paiements.

Produire des estimations en utilisant de telles données comporte plusieurs défis : la base de données ne contient pas l'ensemble des dépenses ; certaines dépenses peuvent être effectuées avant le voyage ; les données peuvent contenir des dépenses qui ne sont pas associées au tourisme ; et le fichier peut contenir des dépenses qui ont été faites en ligne sans qu'il n'y ait eu un voyage.

Des résultats préliminaires montrent que la tendance associée aux dépenses totales observées sur les données de paiements d'un fournisseur pour la période allant du deuxième trimestre de 2015 au

deuxième trimestre de 2017 est similaire à la tendance observée dans l'EVI couvrant la même période. Statistique Canada continuera d'évaluer le potentiel de cette source de données au cours des prochains mois.

3.2 Statistiques du tourisme dans les territoires

Statistique Canada ne recueille pas de données sur le tourisme dans les trois territoires du Canada soient le Yukon, les Territoires du Nord-Ouest et le Nunavut. Ces régions nordiques du Canada sont éloignées et ont une faible densité de population ce qui explique pourquoi le programme actuel ne couvre pas ces derniers. Statistique Canada travaille actuellement en partenariat avec les trois territoires afin de développer un programme de la statistique durable satisfaisant aux besoins des différents utilisateurs. Plusieurs options sont envisagées en ce moment.

Dans un premier temps, les trois territoires seront sur le terrain en 2018 et Statistique Canada les aide à revoir les questionnaires de façon à harmoniser le contenu avec celui de l'ENV et de l'EVV. Les méthodologies utilisées par les partenaires territoriaux seront également revues de façon à bien couvrir les populations d'intérêt et être en mesure de faire l'inférence sur ces dernières. Les territoires seront responsables de la collecte des données alors que Statistique Canada s'occupera du traitement, de la pondération et de l'estimation. La pertinence de ces enquêtes sera évaluée ainsi que la comparabilité des résultats entre les territoires ainsi qu'avec les estimations des diverses provinces.

Statistique Canada évalue la possibilité de conduire l'ENV et l'EVV dans les trois territoires. Cependant, cette option comporte certains inconvénients, dont le coût élevé et la faible couverture du registre des adresses dans les territoires. Des solutions alternatives aux enquêtes traditionnelles pour produire les estimations liées au tourisme sont également envisagées. L'utilisation des données de traitement de paiements décrites à la section 3.1 est une des options envisagées, mais il sera nécessaire d'avoir les données des résidents canadiens, car la majorité des touristes dans les territoires sont canadiens.

4. Travaux futurs

Statistique Canada a franchi une première étape en modernisant le programme de dénombrement aux frontières et en remaniant les deux enquêtes du Programme de la statistique du tourisme. Des données alternatives pourraient être intégrées au modèle actuel dans le but de réduire la dépendance du programme face aux enquêtes traditionnelles et d'améliorer la qualité des différents produits. Les données de traitement de paiements sont une de ces sources de données alternatives qui a le potentiel de s'intégrer dans le programme. En parallèle, Statistique Canada est également en négociation afin d'acquérir les données de différents fournisseurs de cartes de crédit. Ces données offrent le potentiel d'estimer les dépenses liées aux voyages domestiques et internationaux des Canadiens ainsi qu'aux voyages des étrangers au Canada.

L'organisme est également intéressé à étudier les données détenues par les réseaux de téléphones cellulaires. Cette source de données permet de connaître les endroits où le téléphone a été utilisé et le lieu de résidence peut être déduit. Une des utilisations potentielles de cette source de données est l'estimation du nombre de voyages domestiques qui est inconnu contrairement au nombre de voyages internationaux qui est fourni par l'ASFC. Le développement d'une application mobile pouvant être téléchargée sur un téléphone figure également parmi les options pour améliorer le Programme de la statistique du tourisme. Cette dernière pourrait extraire des données de géolocalisation de l'appareil mobile et demander aussi des questions sur les dépenses ou le but de voyage à son utilisateur. Le forage des données provenant du web et des réseaux sociaux à l'aide de

l'intelligence artificielle constitue une autre option potentielle.

En somme, Statistique Canada tente d'acquies et analyser le plus de sources de données possibles reliées au tourisme afin d'améliorer et de moderniser son Programme de la statistique du tourisme. L'acquisition et l'analyse du potentiel de différentes sources de données prendront une place importante dans les prochains mois.

5. Bibliographie

Poirier, G. et Bosa, K. (2017), *Methods of selecting a person within a household: Mailed invitations to complete an electronic questionnaire*, Proceedings of the Statistical Society of Canada annual meeting, Survey Methods Section.

Vézina, G. et Caron, P. (2017), *Comparing Alternative Methods for the Random Selection of a Respondent within a Household for Online Surveys*, Proceedings of the Joint Statistical Meetings, Survey Research Methods Section.

Doering, J., Bosa, K. et Gagnon, F. (2018), *The Impact of Various Nonresponse Follow-up Strategies on the Response Rate of Online Surveys*, to appear in the 2018 Proceedings of the American Association for Public Opinion Research.