

Méthodologie

ENQUÊTE SUR LA FRÉQUENTATION DES ÉTABLISSEMENTS D'HÉBERGEMENT DU QUÉBEC

1. Historique

Jun 1983

L'enquête sur la fréquentation des établissements d'hébergement du Québec¹, réalisée par l'Institut de la statistique du Québec (ISQ) à la demande du ministère du Tourisme², a été implantée en 1983. De façon générale, l'enquête visait à mesurer l'intensité de l'activité des établissements d'hébergement dans les différentes régions touristiques.

Le mandat consistait à réaliser une enquête auprès d'un échantillon d'établissements d'hébergement du Québec selon trois objectifs, à savoir fournir, sur une base mensuelle et pour chacune des 18 régions touristiques existant alors :

1. le nombre d'unités de location disponibles;
2. le nombre d'unités de location occupées;
3. le taux d'occupation moyen.

Ces indicateurs devaient être connus :

1. pour chaque mois de l'année pour les six régions touristiques suivantes : Montréal, Québec, Estrie, Outaouais, Charlevoix et Laurentides;
2. uniquement pour les quatre mois de l'été (juin, juillet, août et septembre) dans le cas des douze autres régions touristiques.

En outre, le taux d'occupation devait être désagrégé selon trois catégories de taille d'établissements :

1. les établissements de moins de 40 unités;
2. les établissements de 40 à 199 unités inclusivement;
3. les établissements de 200 unités et plus.

Au cours des années subséquentes, les méthodes mises en place pour atteindre ces objectifs ont fait l'objet de modifications et d'ajouts.

Oct 1986

À la demande des organismes du milieu, le ministère du Tourisme a fait ajouter la région touristique de Lanaudière aux six régions touristiques pour lesquelles l'enquête était menée pour chaque mois de l'année.

¹ L'enquête sur la fréquentation des établissements d'hébergement du Québec est aussi connu sous le nom d'enquête sur la fréquentation des établissements hôteliers du Québec.

² Le nom du ministère du Tourisme englobe les différents organismes gouvernementaux qui ont eu la responsabilité du tourisme au cours des dernières années.

Juin 1987

À la demande des organismes du milieu, le ministère du Tourisme a fait étendre aux dix-huit régions touristiques la réalisation de l'enquête pour chaque mois de l'année.

Septembre 1988

À la demande d'un certain nombre d'utilisateurs et après différentes consultations, notamment auprès de l'Association des hôteliers du Québec, l'enquête a été remodelée en septembre 1988. Aux trois objectifs spécifiques initialement retenus, le ministère du Tourisme a fait ajouter sur une base mensuelle :

1. le nombre moyen de personnes par unité occupée;
2. le prix de location moyen des unités.

La dix-neuvième région touristique, soit celle de Laval, a été ajoutée au plan de sondage de l'enquête de même que les résultats agglomérés selon la communauté urbaine de Montréal (CUM) et selon la communauté urbaine de Québec (CUQ).

Un autre ajout a été fait pour connaître le taux d'occupation moyen et le prix moyen de location des unités non seulement selon les trois catégories de taille citées précédemment, mais aussi selon les catégories de classification de « pignon » et de « lys ». Ces éléments ont été éliminés en 1991, à la suite du changement du système de classification des établissements d'hébergement.

Le dernier ajout a consisté à réaliser une enquête complémentaire auprès des établissements de 150 unités et plus, afin de connaître le but de voyage et l'origine géographique de leur clientèle. Cette enquête a été abandonnée après quelques mois d'expérimentation, en raison de l'impossibilité pour de nombreux hôteliers à fournir ces informations.

Juin 1994

La méthode de mise à jour de l'univers visé par l'enquête est passée d'une approche annuelle à une approche mensuelle. Cette approche est plus précise, puisqu'elle permet de tenir compte des ouvertures et des fermetures d'établissements s'effectuant chaque mois (l'ancienne méthode permettait uniquement de tenir compte des fermetures sur une base annuelle). Pour certaines régions, une partie de l'écart observé en ce qui a trait aux unités disponibles et occupées, peut provenir de cette modification. Ses effets sont plus susceptibles de se faire sentir durant les derniers mois de l'année 1994 (octobre, novembre et décembre).

1998

Le mois d'avril 1998 marque un point tournant dans la réalisation de l'enquête, car une importante amélioration a été apportée à la méthode d'estimation. L'ancien estimateur utilisé, *Horwitz-Thompson*, a été remplacé par un *estimateur par le quotient* qui utilise une donnée connue auprès de tous les établissements d'hébergement visés par l'enquête, soit le nombre d'unités disponibles. Cette modification, qui ne concerne que les estimations du nombre d'unités disponibles et occupées, a plusieurs avantages. En plus d'améliorer la précision des estimations, elle permet également de réduire l'impact du renouvellement de l'échantillon qui survient tous les mois de mai. Ainsi, on peut s'attendre à obtenir des séries chronologiques plus stables.

À partir de mai 1998, les données pour la Région métropolitaine de recensement (RMR) de Montréal ont été rendues disponibles.

Mai 1999

La population visée par l'enquête a subi un changement mineur. Les établissements d'enseignement ont été retirés de cette population. Cette modification a affecté les mois de juin et de juillet puisque ces établissements offrent de l'hébergement de type hôtelier uniquement durant ces deux mois de l'année.

Mai 2000

À la suite du découpage territorial de la région touristique de la Mauricie–Bois-Francs, deux nouvelles régions touristiques distinctes ont été créées, soit la Mauricie et le Centre-du-Québec. Le plan de sondage de l'enquête a été modifié afin d'y ajouter ces deux régions.

La taille de l'échantillon de l'enquête a été augmentée, en vue d'accroître la précision des estimations.

2002

En mars 2002, les données se rapportant à la Communauté métropolitaine de Québec (CMQ) ont été substituées à celles de la Communauté urbaine de Québec (CUQ).

La variable portant sur le nombre de personnes par unité a été retirée du questionnaire en mai 2002.

À la demande de l'Office du tourisme et des congrès de Québec (OTCQ), l'ISQ a augmenté la taille d'échantillon de l'enquête pour la région touristique de Québec. Depuis, le plan de sondage de l'enquête prévoit le recensement systématique des établissements d'hébergement de cette région, ce qui a pour effet d'accroître la précision des estimations.

2004

À la demande du ministère du Tourisme, l'ISQ produit des données estimées sur une base désagrégée par niveau de classification (0 étoile, 1 étoile, 2 étoiles, 3 étoiles, 4 étoiles, 5 étoiles et en cours d'évaluation) pour l'ensemble du Québec, à partir des données recueillies pour l'enquête sur la fréquentation des établissements d'hébergement du Québec.

À la demande de l'Association touristique régionale (ATR) de Charlevoix, l'ISQ a augmenté la taille d'échantillon de l'enquête pour cette région touristique. Depuis, le plan de sondage de l'enquête prévoit le recensement systématique des établissements d'hébergement de cette région, ce qui a pour effet d'accroître la précision des estimations.

Mai 2005

À la suite du découpage territorial de la région touristique du Nord-du-Québec, deux nouvelles régions touristiques distinctes ont été créées, soit la Baie-James et le Nunavik. Le plan de sondage de l'enquête a été de nouveau modifié afin d'y ajouter ces deux régions et prévoit le

recensement systématique des établissements d'hébergement de ces deux régions, ce qui a pour effet d'accroître la précision des estimations.

Janvier 2008

À la suite du découpage territorial de la région touristique de la Baie-James en 2007, deux nouvelles régions touristiques distinctes ont été créées, soit la Baie-James et l'Eeyou Istchee. Le plan de sondage de l'enquête a été de nouveau modifié afin d'y ajouter ces deux régions et prévoit le recensement systématique des établissements d'hébergement de ces deux régions, ce qui a pour effet d'accroître la précision des estimations.

2. Rappel de l'objectif de l'enquête

L'objectif de l'enquête est de produire, sur une base mensuelle, une série d'indicateurs sur l'industrie de l'hébergement du Québec. Ces indicateurs sont :

1. le nombre d'établissements d'hébergement (hôtels et résidences de tourisme de quatre unités et plus);
2. le nombre d'unités disponibles (chambres, condos et chalets);
3. le nombre d'unités occupées;
4. le taux d'occupation moyen des établissements;
5. le prix quotidien moyen de location des unités;
6. les revenus bruts de location des unités.

3. Définition des termes et concepts

Les principaux termes et concepts utilisés sont les suivants :

- Unités disponibles quotidiennement (chambres, condos, chalets)

Il s'agit des unités effectivement mises quotidiennement à la disposition de la clientèle par les établissements d'hébergement en exploitation durant chaque mois de l'enquête. Cela élimine les unités ou les parties des établissements fermés pour cause de rénovation ou pour toutes autres raisons.

- Unités occupées quotidiennement (chambres, condos, chalets)

Il s'agit des unités occupées quotidiennement par la clientèle ainsi que celles offertes à titre gracieux.

- Taux d'occupation moyen

Il s'agit du rapport entre le nombre estimé d'unités occupées par la clientèle d'un établissement d'hébergement et le nombre estimé d'unités disponibles.

- Prix quotidien moyen de location des unités

Il est calculé à partir des revenus bruts de location des unités occupées. Cela exclut les revenus découlant du service aux unités de location, la TPS, la TVQ et la taxe sur l'hébergement.

4. Couverture et plan de sondage

Jusqu'en 1991, un permis était délivré aux exploitants d'établissements d'hébergement ayant un minimum de 5 unités. Des modifications apportées à la réglementation en 1992 ont fait passer ce seuil à 4 unités. Depuis l'adoption de la Loi sur les établissements d'hébergement touristique en 2000, les exploitants doivent obtenir une attestation de classification du ministère du Tourisme, laquelle est le résultat du niveau de classification déterminé par la Corporation de l'industrie touristique du Québec (CITQ).

La population visée par cette enquête est l'ensemble des établissements hôteliers et des résidences de tourisme du Québec comptant quatre unités de location ou plus (chambres, condos, chalets). Cette population est de l'ordre de 2 300 établissements d'hébergement.

La base de sondage pour cette enquête est constituée des établissements ayant une attestation de classification établie par la CITQ ainsi que des établissements en cours d'évaluation pour l'obtention de cette classification. Juste avant le début de l'année d'enquête (fin mars ou début avril), l'ISQ procède au tirage de l'échantillon à partir de la liste de tous les établissements visés par l'enquête, qu'ils soient actifs ou non à ce moment.

Les établissements de l'échantillon à enquêter sont déterminés à chaque mois. Tous les établissements saisonniers ayant une date d'ouverture inférieure au 15 du mois et une date de fermeture supérieure ou égale au 15 du mois visé ainsi que tous les établissements non-saisonniers qui sont ouverts ou en opération durant le mois devront faire partie de l'enquête. C'est à partir de la liste des établissements ouverts ou en opération pour chaque mois que sont déterminés les établissements enquêtés. Les établissements échantillonnés faisant partie de cette « population active » sont alors interrogés pour le mois en question. Des productions informatisées permettent la validation de l'appariement de la population et de l'intégrité de l'échantillon pour le mois enquêté. Cette procédure est ainsi répétée pendant les douze mois de l'enquête, jusqu'au renouvellement partiel de l'échantillon, au printemps suivant. Cette taille d'échantillon est de l'ordre de 1 150 établissements d'hébergement, qui peut varier un peu d'un mois à l'autre selon les saisons, et permet de produire des estimations avec un niveau de confiance à 95 %.

Le tirage de l'échantillon est effectué selon un plan de sondage stratifié. Ainsi, afin de choisir les unités échantillonnées, l'univers statistique a été divisé en différentes strates d'échantillonnage de façon à assurer une bonne représentativité de chacun de ces groupes. Ces strates d'échantillonnage sont :

- les 22 régions touristiques;
- les 3 catégories de tailles d'établissement;

Certaines strates d'échantillonnage sont systématiquement recensées. C'est le cas :

- des établissements de 150 unités et plus;
- des strates comprenant 4 établissements ou moins;
- des strates comprenant les régions touristiques suivantes : Îles-de-la-Madeleine, Charlevoix, Québec, Laval, Baie-James, Nunavik et Eeyou Istchee.

Les autres strates sont échantillonnées et les établissements sélectionnés font l'objet d'une certaine rotation afin d'éviter aux mêmes exploitants le fardeau de répondre au questionnaire de l'enquête plusieurs années consécutives.

Depuis mai 2000, la taille de l'échantillon de l'enquête sur la fréquentation des établissements d'hébergement a été augmentée dans le but d'atteindre certains objectifs de précision au niveau des estimations du « nombre d'unités occupées », à savoir :

- Obtenir un coefficient de variation (CV) de 1,75 % ou moins pour les estimations québécoises.
- Obtenir un CV de 3,5 % ou moins pour les estimations de la région touristique de Québec.
- Obtenir un CV de 4,0 % ou moins pour les estimations de la région touristique de Montréal.
- Obtenir un CV de 10,5 % ou moins pour les estimations des autres régions touristiques du Québec.

5. Collecte des données

Une fois par année, un questionnaire est envoyé à chaque répondant. Ce questionnaire mesure chaque mois trois variables, soit le nombre d'unités disponibles, le nombre d'unités occupées et les revenus de location des unités. Il permet ainsi aux répondants de recueillir quotidiennement l'information nécessaire.

La collecte d'information se fait par téléphone et s'amorce le premier jour ouvrable du mois suivant pour se terminer dans les environs du 20. Plusieurs tentatives (à des heures et jours différents) sont effectuées pour tenter de rejoindre un établissement sélectionné. Les répondants ont également la possibilité de laisser leurs informations sur la boîte téléphonique consacrée exclusivement à cette enquête.

Règle générale, les taux de réponses de l'enquête aux niveaux provincial et régional sont dans l'ensemble satisfaisants, dépassant souvent 80 %.

6. Ajustements aux données

En cours de collecte et de traitement, les opérations usuelles de validation des données ont lieu. Les procédures d'imputation utilisées sont l'imputation pour la non-réponse globale, l'imputation pour la non-réponse partielle et la pondération.

6.1 Imputation pour la non-réponse globale

Il y a quelques années, une procédure d'imputation pour contrer la non-réponse globale a été mise en place de façon à assurer qu'il y ait, dans chaque strate, un nombre minimal d'établissements pour lesquels il y a des réponses. Initialement, pour des raisons de calculs, il fallait avoir au moins deux de ces établissements. Par la suite, ce nombre a été fixé à six, évidemment lorsque la taille de la strate le permet.

L'imputation est effectuée par la *méthode du plus proche voisin*. Cette méthode consiste à trouver un donneur pour chaque établissement qui doit subir une imputation en se basant sur des critères (connus dans la base de sondage) de proximité comme la région, le nombre d'unités de location inscrites sur l'attestation, etc.

Dans le cadre de la présente enquête, le taux d'imputation global est maintenu très bas pour tous les mois.

6.2 Imputation pour la non-réponse partielle

L'imputation pour la non-réponse partielle touche la variable « revenus de location des unités » et est réalisée à l'aide de la méthode du plus proche voisin. Elle est effectuée sur une base provinciale et régionale.

6.3 Pondération

Il s'agit d'une pondération inférentielle qui permet de considérer la probabilité de sélection de l'établissement et la non-réponse dans sa strate.

La pondération a pour but de permettre l'inférence des résultats observés à la population étudiée en fonction du plan de sondage qui a permis de recueillir les données. En fait, le poids accordé à un établissement d'hébergement reflète le nombre d'établissements que celui-ci « représente » dans la population. Ce poids est donc fonction de deux réalités. Initialement, il représente l'inverse de la probabilité de sélection de l'établissement dans sa strate. Ensuite, il tient compte de la structure de la réponse observée dans la strate à laquelle celui-ci appartient, en supposant que le profil des non-répondants de cette strate est semblable au profil des répondants.

Soit N_h = nombre total d'établissements de la strate h .

Soit n_h = nombre d'établissements sélectionnés pour faire partie de l'échantillon de la strate h .

Soit r_h = nombre d'établissements répondants de la strate h .

Alors le poids Wr_h accordé aux répondants de la strate h se calcule ainsi :

$$Wr_h = \frac{N_h}{n_h} \times \frac{n_h}{r_h} = \frac{N_h}{r_h}$$

Étant donné que la non-réponse partielle est éliminée par imputation, il n'y a pas lieu de pondérer à nouveau chacune des estimations.

7. Estimation et précision des résultats

7.1 Méthode d'estimation

Dans le cadre de l'enquête sur la fréquentation des établissements d'hébergement du Québec, des estimations de totaux et de taux sont produits, le tout à partir de données recueillies provenant d'un plan de sondage stratifié. Ces estimations sont produites pour la province de Québec, par région touristique, par catégorie de taille et par niveau de classification. De plus, elles sont produites également pour la Région métropolitaine de recensement (RMR) de Montréal et la Communauté métropolitaine de Québec (CMQ).

Tout d'abord, l'estimation de totaux s'effectue à l'aide de l'*estimateur par le quotient*. Ce type d'estimateur permet d'ajuster l'estimation obtenue à l'aide des répondants en utilisant une variable corrélée qui est connue pour l'ensemble de la population d'établissements. De plus, l'utilisation de cet estimateur permet d'améliorer la précision des estimations produites. Notons que les estimations de totaux portant sur l'ensemble des strates d'analyse sont calculées à l'aide de l'*estimateur par le quotient combiné*.

Ensuite, afin d'estimer les taux, l'équation suivante est utilisée :

$$\hat{T} = \frac{\hat{y}}{\hat{x}}$$

où \hat{T} pourrait être, par exemple, l'estimation du taux d'occupation; \hat{y} l'estimation du nombre d'unités occupées; et \hat{x} l'estimation du nombre d'unités disponibles. Les estimations \hat{y} et \hat{x} correspondent aux sommes pondérées du nombre d'unités occupées et du nombre d'unités disponibles des établissements répondants. Dans ce cas, l'estimateur par le quotient n'est pas utilisé pour estimer ces deux totaux.

7.2 Précision des estimations

Puisque les résultats de l'enquête sont établis à partir des réponses d'un échantillon de répondants, ils sont entachés d'une certaine imprécision. La précision d'une estimation est mesurable à l'aide de la marge d'erreur et du coefficient de variation (CV).

Tout d'abord, la marge d'erreur se définit, en fonction d'un niveau de confiance (1- α), comme suit :

$$\text{Marge d'erreur} = z_{\alpha/2} * \sqrt{\text{variance de l'estimation}}$$

Où $z_{\alpha/2}$ est le percentile d'une loi normale (0,1). Pour un niveau de confiance de 95 %, $z_{\alpha/2} = 1,96$.

À partir de cette marge d'erreur, il est possible de définir l'intervalle de confiance (IC) de niveau (1- α), qui est associé à l'estimation :

$$\text{IC} = \text{estimation} \pm \text{marge d'erreur}$$

Cela signifie que, si l'on refaisait le sondage à plusieurs reprises, 19 intervalles sur 20 contiendraient la valeur réelle du paramètre estimé. Notons que la marge d'erreur calculée ici repose sur l'hypothèse que les moyennes ou les totaux estimés suivent une loi normale. Ce résultat, qui est asymptotiquement démontré par le théorème de la limite centrale, peut ne pas s'appliquer dans le cas où la taille de l'échantillon est petite et que la distribution des valeurs mesurées est très étendue ou asymétrique ou bien dans le cas de l'estimation d'une petite proportion.

Par ailleurs, le coefficient de variation (CV) est une mesure qui permet de quantifier la précision relative de l'estimation et il se définit comme suit :

$$\text{CV} = \frac{\sqrt{\text{variance (estimation)}}}{\text{estimation}}$$

Cette mesure contribue à faciliter l'interprétation quant à la précision d'une estimation. Plus le CV est élevé, moins précise est l'estimation et vice versa. Le CV a l'avantage de permettre de comparer la précision de différentes estimations entre elles. Le tableau ci-dessous présente la relation entre le CV et la précision de l'estimation utilisée par l'ISQ. Notamment, il faut être très prudent lors de l'utilisation des estimations dont le CV est très élevé (supérieur à 25 %), ce qui indique une très grande variabilité relative de l'estimation. Ces estimations ne devraient être utilisées qu'à titre indicatif.

Relation entre la valeur d'un coefficient de variation et la précision de l'estimation associée	
CV	Précision de l'estimation
≤ 5 %	excellente
5 % < CV ≤ 10 %	très bonne
10 % < CV ≤ 15 %	bonne
15 % < CV ≤ 25 %	passable
> 25 %	faible, à utiliser avec circonspection

8. Confidentialité des données

La Loi de l'ISQ édicte les conditions de fonctionnement de l'Institut. Elle impose notamment :

- pour le répondant, de fournir l'information demandée;
- pour l'ISQ, de garantir la confidentialité des informations fournies par le répondant.

C'est ainsi que seuls les employés de l'ISQ ont accès aux données individuelles. Elles ne peuvent être transmises à quiconque.

9. Comparabilité des données dans le temps

Compte tenu des modifications qui ont jalonné le développement de l'enquête depuis 1983, la comparaison des données dans le temps n'est pas toujours possible.

L'adoption, en 1990, d'un nouveau découpage territorial pour plusieurs régions touristiques, a également pour conséquence de rendre les données de 1990 et celles des années suivantes totalement impossibles à comparer avec celles de 1989 et des années antérieures. Dans le cas de certaines autres régions touristiques, les changements dans le découpage territorial ont été mineurs.

Par contre, les modifications apportées à la méthodologie de l'enquête en 1999 créent une coupure complète dans les données. Ainsi, les résultats concernant les nombres d'unités disponibles et occupées de mai 1998 à décembre 1999 ne sont pas comparables avec celles des années précédentes.

Mise à jour : 29 avril 2008