

## Détection des erreurs:

Des procédures de contrôle de la qualité visant à garantir l'exactitude et la qualité des données déclarées sont suivies par Statistique Canada en collaboration avec le Comité de l'information et de la statistique policières (POLIS) de l'Association canadienne des chefs de police. Les problèmes relatifs à la qualité des données ont principalement trait à une mauvaise compréhension par le répondant des règles de déclaration ainsi qu'à des erreurs occasionnelles dans l'entrée des données. Un autre problème qui surgit relativement au Programme DUC2 tient à une erreur dans l'extraction des données.

Le processus de saisie et de correction des données provenant des deux programmes DUC est une activité permanente. Actuellement, les chiffres définitifs sont diffusés dans le public en juillet. Cette diffusion fait suite à un long cycle de production des données qui débute en décembre de l'année précédente avec la préparation de statistiques agrégées provisoires. Ces statistiques permettent de détecter très tôt tout problème dans les données.

Les problèmes sont repérés au moyen d'un modèle de « tolérance » statistique élaboré expressément pour comparer les données d'un répondant avec ses données antérieures et avec celles de répondants semblables. Ils sont également relevés au moyen d'un modèle de « détection des valeurs aberrantes » qui crée un certain nombre de ratios intérieurs à un service de police particulier et à une infraction particulière. Selon l'importance de la valeur aberrante, celle-ci est retenue pour un examen plus approfondi.

Les non-répondants et les anomalies dans les données sont facilement repérés au moyen de ces modèles, et le personnel peut ensuite vérifier auprès du service de police.

Pour créer le fichier provisoire ou le fichier maître, il faut convertir le fichier du programme fondé sur l'affaire afin de produire des chiffres comparables à ceux du programme à base de données agrégées. Les chiffres convertis, soit les « chiffres agrégés calculés relatifs aux infractions », sont combinés avec les chiffres du programme à base de données agrégées, afin de produire des tableaux de données nationales et provinciales. De cette façon, on maintient la continuité avec les séries chronologiques sur la criminalité.

La qualité des données constitue une préoccupation primordiale dans toute enquête. Le CCSJ applique plusieurs méthodes différentes pour garantir la qualité des données du Programme DUC. L'intégrité interne des enregistrements reçus des répondants et la vraisemblance des déclarations des répondants sont constamment évaluées au moyen d'une révision manuelle et électronique des données. Si possible, les données du Programme DUC sont vérifiées en regard des statistiques sur la criminalité que le répondant a produites de manière indépendante, comme les rapports annuels des corps policiers.

Toutefois, le CCSJ ne vérifie pas normalement les données des répondants. Comme l'enquête à base de données agrégées et l'enquête fondée sur l'affaire fonctionnent d'une façon très différente, les questions particulières entourant la qualité des données et les procédures requises pour garantir cette qualité sont différentes.

### **Programme DUC à base de données agrégées**

Comparativement au Programme DUC fondé sur l'affaire, le Programme DUC à base de données agrégées est assez simple à appliquer, valider, réviser et diffuser. Les fichiers maîtres de données agrégées sont continuellement mis à jour pendant toute l'année, et ils sont traités encore davantage, au besoin, pour produire des statistiques agrégées consolidées. Le cas échéant, les erreurs découvertes pendant la révision des données font l'objet d'un suivi auprès du répondant et sont corrigées. Étant donné que les données déclarées à l'enquête à base de données agrégées le sont sous forme de chiffres sommaires, les problèmes liés au dénombrement et à la déclaration ne peuvent être cernés qu'au moyen d'une évaluation de la vraisemblance des comptes du répondant. À cet égard, il est essentiel pour maintenir la qualité des données d'appuyer largement les répondants en leur offrant un guide d'utilisation, un soutien téléphonique et de la formation.

### **Programme DUC2 fondé sur l'affaire**

Dans le cadre du Programme DUC fondé sur l'affaire, on recueille une vaste gamme de données sur chaque affaire criminelle. De plus, les données du programme font l'objet d'une révision approfondie pendant leur saisie au CCSJ. Contrairement aux données agrégées où la révision peut seulement repérer des données manquantes et garantir que la somme des valeurs dans les cellules correspond aux totaux indiqués, les données sur les affaires sont assujetties à des vérifications très détaillées de la validité et de la logique, ainsi qu'à des imputations, qui permettent de déterminer si le répondant a fourni les données requises, et de garantir que seuls les fichiers qui répondent aux besoins en données de l'enquête sont disponibles pour être analysés.

Par conséquent, en plus de pouvoir déterminer le nombre total d'affaires, on peut évaluer plus en détail la qualité des données liées à chaque affaire. Les anomalies dans les pratiques de déclaration de la criminalité qui n'avaient pas été cernées auparavant sont maintenant corrigées. À mesure que d'autres services de police se joignent au Programme DUC2 fondé sur l'affaire, les données fournissent une image de plus en plus exacte de la nature des affaires déclarées à la police au Canada.

Les préoccupations particulières entourant la qualité des données comprennent les suivantes : la variété des interfaces de logiciel utilisées pour extraire les données du Programme DUC fondé sur l'affaire; la complexité de l'extraction et de la saisie des données; les différences dans les pratiques opérationnelles et administratives entre les diverses forces de police; et la nécessité de traduire fidèlement ces données en comptes agrégés.