Joint UNECE/Eurostat work session on statistical data confidentiality
(Ottawa, Canada, 28-30 October 2013)

Topic (v): Confidentiality issues and case studies

Disclosure control for tables using administrative justice data

Prepared by S. Franklin and A. Reicker, Statistics Canada
Disclosure control for tables using administrative justice data

S. Franklin * and A. Reicker **

* Household Survey Methods Division, Statistics Canada, sarah.franklin@statcan.gc.ca
** Household Survey Methods Division, Statistics Canada, mailto:alexander.reicker@statcan.gc.ca

Abstract: The Canadian Centre for Justice Statistics (CCJS) has made administrative datasets available through Research Data Centres (RDCs) across Canada. Disclosure control rules were developed for tabular data for the Uniform Crime Reporting Survey and the Homicide Survey. The disclosure method for tables produced in the RDCs uses a scoring approach whereby all survey variables are assigned a sensitivity score. A table’s score is the sum of the scores of the contributing variables. If the score falls above a threshold value, then the table cannot be released. This approach is also being applied to tables that CCJS produces for its own publications as well as for custom requests that it receives from external clients. CCJS is also making data available through the Real Time Remote Access (RTRA) tool which allows individuals to submit SAS code to produce tables. The RTRA uses controlled rounding to mask sensitive data. This paper will discuss the two methods used by CCJS to protect tabular data: the scoring approach and additive controlled rounding.

1 Introduction

In order to make justice data more accessible, the Canadian Centre for Justice Statistics (CCJS) has placed microdata from justice surveys in Research Data Centres (RDCs) across Canada. In 2013, CCJS placed two police-reported administrative surveys in the RDCs: the Uniform Crime Reporting Survey (UCR) and the Homicide Survey. These surveys are also now available on a platform known as Real Time Remote Access (RTRA) which allows users to remotely send SAS queries to Statistics Canada in order to generate frequency tables.

Prior to making these datasets available through the RDCs and RTRA, CCJS’ relied solely on subject matter expertise as a means of assessing whether a table was releasable or not. Continuing with this approach would not be an effective means of vetting the large volume of data tables produced through RDC and RTRA platforms. Consequently, it was deemed essential to develop simple disclosure control rules in order to automate as well as simplify the vetting process.

The approach that was developed for the RDCs was a scoring approach first introduced by the Institut de la Statistique du Québec (ISQ) for their vital statistics (Boudreau et al
This method assigns a sensitivity score to each variable in a dataset. When a table of frequencies is generated, if the sum of the scores of the variables used to produce the table fall below a given threshold, the table may be released. If the table’s score is too high, the RDC researcher may choose to produce a table with a lower score using more aggregated variables. For the UCR and Homicide Surveys, researchers may alternatively use additive controlled rounding which is used by the RTRA.

For the UCR and Homicide Surveys, the scoring approach was found to mimic the thought processes used by CCJS subject matter experts to vet tables of frequency counts and crime rates. It has since become CCJS’ standard method of disclosure control for these two surveys. This paper describes how the ISQ’s scoring approach was adapted for police-reported crimedata.

2 Research Data Centres and Real Time Remote Access

Data collected by Statistics Canada under the Statistics Act must be protected to prevent the disclosure of confidential information. Specifically, section 17 of the Statistics Act states that:

‘no person […] shall disclose or knowingly cause to be disclosed, by any means, any information obtained under this Act in such a manner that it is possible from the disclosure to relate the particulars obtained from any individual return to any identifiable individual person, business or organization.’

Statistics Canada uses a variety of measures to protect the confidentiality of its data: employees swear an oath of confidentiality, physical measures are used to protect the data, and statistics are vetted prior to their release.

Statistics Canada is making more datasets more widely available to Canadians. While increasing access, the confidentiality of these datasets must still be preserved. Research Data Centres (RDCs) are secure settings, housed in universities across Canada where researchers can access microdata from Statistics Canada surveys. In order to gain access to an RDC, a researcher must first submit a proposal. Once the research proposal is approved, the researcher is sworn in as a deemed Statistics Canada employee and must adhere to all confidentiality measures. To further protect the data, each RDC has a Statistics Canada analyst to ensure that disclosure control procedures are followed.

The RTRA is a new platform developed by Statistics Canada which allows users to generate tables from a remote location. Users must obtain a licence and sign a contract which outlines the confidentiality measures that they must abide by. End user frequency
requests are submitted remotely from their computer to Statistics Canada which then runs
the request and applies a rounding program to the tabular output.

Microdata that are available in the RDCs or through RTRA are stripped of personally
identifying information such as name, birth date, fingerprint identifier, social insurance
number, driver’s licence, residential address, email address, etc. The disclosure control
rules that are applied in the RDCs and on RTRA depend on the nature of the data and
how they were collected (e.g., sample survey data or administrative data).

CCJS attempted to increase access to the Homicide Survey in 2009 when the data were
placed in the RDCs as a pilot project. However, this first attempt was not successful.
Feedback from end users and CCJS staff noted that the disclosure control procedures
employed at the time lacked clarity. RDC analysts were unable to assess the disclosure
risk of tables and consequently sent the majority of output directly to CCJS for review.
This process caused delays for researchers and the added workload was burdensome for
CCJS. Moving forward, CCJS focused on developing disclosure rules that were
straightforward, simple and easily administered without relying on subject matter
expertise to apply and these are detailed here.

This paper focuses on the disclosure control of frequency tables produced using
administrative data sets. While the goal of the RDCs is to permit researchers access to
data files in order to perform more sophisticated analysis than the production of
frequency tables, researchers often want such tables in their analysis. Frequency counts
are commonly used in justice statistics. Crime rates are simply frequency counts per
100,000 people in the population. Tables of frequency counts are of particular concern
from the perspective of disclosure control since they are susceptible to divulging sensitive
information.

3 The sensitive nature of crime data

The Uniform Crime Reporting Survey (UCR) and the Homicide Survey are two of CCJS’
flagship surveys. Both collect data from all police services (approximately 150) in
Canada making them censuses of police-reported crime. These surveys capture crimes
that fall under the Criminal Code of Canada that are reported to, and substantiated by, the
police. The Homicide Survey collects data on homicides (1st and 2nd degree murder,
manslaughter and infanticide) while the UCR collects data on all Criminal Code crimes.

Both surveys store their data in three separate files: an incident file with one record per
incident, a victim file with one record per victim for a given incident, and an accused file
with one record per accused person for a given incident. These files can be used
independently or linked together in order to respond to a specific research question.
Both surveys contain highly sensitive personal characteristics. The Homicide Survey is particularly sensitive since it contains more confidential information and unlike most crimes virtually all homicides come to the attention of the police. Homicides are also rare occurrences in Canada. Homicide data are consequently more at risk of disclosure than UCR data.

The risk of identifying an individual (identity disclosure) is reduced in the RDCs and on RTRA by removing personally identifying characteristics from the datasets (e.g., names, fingerprints, etc). The risk of divulging sensitive characteristics (attribute disclosure) in a table carries a higher risk (see section 4.1).

3.1 The Uniform Crime Reporting Survey

Many UCR variables describe the criminal incident rather than the victim or the accused. In general, incident-level characteristics are less sensitive than victim or accused characteristics. Each incident record contains up to four violations which are assigned a 4-digit code. The most serious violation in an incident is used by CCJS to produce official crime statistics. Often, only the accused and the police officer know what was considered to be the most serious violation in an incident. For example, an outsider may know that an incident of assault was reported to the police, but is unlikely to know if the police record indicates it was a level 1, level 2 or level 3 assault.

On the UCR, the relationship variable is one of the most sensitive variables and is also of great interest to researchers. It is on the victim file and records the relationship of each victim to the accused person. Other UCR variables which are potentially sensitive include the level of injury suffered by the victim and how the accused was cleared (charged or cleared otherwise).

3.2 The Homicide Survey

The types of variables collected by the Homicide Survey are similar to those collected by the UCR, however the former includes variables that are more detailed and potentially sensitive. Some of the highly sensitive personal variables collected by the Homicide Survey include:

- relationship between the victim and the accused person (while this variable appears on the UCR, the degree of detail on the Homicide is far greater)
- drug and alcohol use at the time of the incident
- mental health of accused
- whether the incident was gang related
- history of family violence between the victim and accused
- first to use physical force in the incident
- occupation of victim, accused

The occupation of the homicide victim or the person accused of homicide may be sensitive since it includes illegal occupations such as drug dealer and prostitute.

All (or almost all) homicides in Canada come to the attention of the police. This makes the Homicide Survey a census of all homicides in Canada, unlike most other crimes. A full cell in the Homicide data is a true full cell, while in the UCR data a full cell may appear that way only because other incidents were not reported to the police. Therefore, it is easier for an outside attacker to identify a particular homicide incident in a table.

The Homicide Survey also collects information on unsolved homicides. CCJS must take special care to ensure that sensitive information on unsolved homicides is not disclosed.

Thankfully, homicide is very rare in Canada, with 543 homicides in 2012. Each year, several Census Metropolitan Areas and even entire provinces have no homicides. A table showing a single homicide in a CMA but revealing nothing else is not considered to be sensitive. It might be possible to identify a particular homicide, but the table does not reveal any confidential characteristics. This is the principle behind the scoring approach: so long as the table is safe, the data may be released unperturbed.

4 Developing disclosure rules for CCJS crime data

The release of microdata files to the RDCs and RTRA is relatively new. Most CCJS data are released in CANSIM tables or Juristat articles, both of which are available through Statistics Canada website, or through custom requests on demand. All tables are vetted by subject matter experts to ensure that no confidential data are released. Cell suppression has sometimes been used to protect justice data in a table, but otherwise the true frequency counts are released without any perturbation.

CCJS does not produce public-use microdata files or synthetic microdata files (i.e., anonymised microdata files with noise added) for any of its administrative surveys, nor does it perform data swapping. One of its administrative surveys – a survey of child support payments – uses additive controlled rounding.

In 2009 (after the Homicide Survey was placed in the RDCs), Statistics Canada developed disclosure control guidelines for administrative data (Tambay, 2009) based primarily on the scoring approach used by the ISQ. All administrative surveys available in the RDCs follow these guidelines although how the scoring approach is adapted
depends on the needs of the survey. Many Statistics Canada administrative surveys use a combination of scoring, suppression and additive controlled rounding.

4.1 Disclosure control issues

There are two types of disclosure: identity disclosure and attribute disclosure. Identity disclosure involves the ability to identify an individual in a table, typically because the person is the only one in a table with a certain set of characteristics. This is the small cell problem, where the count in a cell is one or two. However, if no other information is revealed, this is not necessarily a breach of confidentiality. CCJS regularly publishes tables with cells of one or two, such as a single homicide in a city.

Attribute disclosure is the disclosure of characteristics that individuals possess. This is the full cell problem, where only one cell in a row or column is not zero. If an outsider knows which row contains a particular incident, and only one column in that row is not zero, then the outsider has learned which column contains the incident. While something may be disclosed, it is not always a breach of confidentiality. UCR tables often have harmless full cells. For example, a table of violations versus weapon used will indicate that for every traffic violation the weapon was a vehicle.

4.2 Residual disclosure and existing published tables

While a single table may not disclose any confidential information, there may be a risk that combining it with another published table could disclose something sensitive. This is known as residual disclosure. For example, residual disclosure could occur if different tables use different aggregations of a variable and combining the tables divulges a particular, sensitive category.

Statistics Canada is required to control for possible residual disclosure, so any disclosure control approach should account for data that have been previously published by CCJS.

The risk of residual disclosure is low for the UCR since the published UCR tables are generally high-level, aggregate counts of frequencies that would pass the disclosure control rules used with the scoring approach. In addition, as mentioned earlier, the fact that not all criminal incidents are included in the UCR provides some protection. In the case of homicide data, most of CCJS’ published homicide tables are at low risk of causing residual disclosure since the homicide statistics often are aggregated across a 10-year period.
With both the UCR and Homicide Surveys, the main risk of residual disclosure comes not from the scoring approach, but from rounded tables. With a rounded table, it is possible to determine that the rounded frequency was based on only one incident, one victim or one accused person. Several measures have been taken to prevent such disclosure from rounded tables produced in the RDCs and RTRA. These measures include aggregating violations that have low numbers, collapsing other variables to hide characteristics that are not common into larger groupings, and not allowing the RTRA to produce tables of homicides in a single year (homicides are aggregated across 10 years for the RTRA).

4.3 Data protection strategies

There are several methods to ensure that published frequency tables do not reveal sensitive characteristics. In the past, CCJS has sometimes used cell suppression for small cells and full cells. Cell suppression is not recommended where many tables are being produced, as is the case in the RDCs and RTRA, since there is a risk of residual disclosure by combining multiple tables.

Instead of suppressing a cell, it may be possible to collapse two or more rows or columns together. This increases cell counts, disguises sensitive categories and protects against identity and attribute disclosure.

The scoring approach (explained below) is a method that quantifies CCJS’ subject matter expertise so that it can be applied consistently to all frequency tables produced for a particular survey. Each table is assigned a total score indicating the table’s sensitivity. If the table score is below a given threshold, then the table may be released.

The presence of full cells or small cell counts does not figure into the decision to pass or fail a table: if the table passes, any small cells or full cells are innocuous. The drawback of the scoring approach is that it is relatively conservative and errs on the side of caution by suppressing tables that do not divulge any confidential information.

Rounding the values in a table offers protection against an outside attacker since no cells have small counts (the minimum is the rounding base) and any full cell may not be a true full cell since the zero cells in the row or column may be nonzero values which have been rounded down. However, rounding necessarily distorts the data which is problematic for rare crimes. There are different rounding algorithms, but additive controlled rounding is used by the RTRA and offered to RDC researchers. Additive controlled rounding is a type of random rounding which is constrained by the marginal totals (Boudreau et al, 2004).
4.4 The scoring approach for crime data

Since 2003, the ISQ has used a scoring approach to protect the confidentiality of tables of vital statistics. This approach was adapted in order to mirror how a CCJS’s subject matter analyst would vet frequency tables: if a table is deemed to be innocuous then the table may be released without any distortion or suppression of the data.

The procedure is straightforward and can be implemented in the RDCs without CCJS subject matter assistance. Each survey variable is given a score between zero and eight, and the table’s score is the sum of the variable scores. If the table score is below the threshold value of eight, then the table is deemed not to involve any risk of disclosure and can be published. In addition to the table score, there is a limit to the number of variables that can be used to produce a table (eight for UCR tables and three for homicide tables).

If the table score is too high, then it cannot be released, however it may be feasible to generate the table using more aggregated versions of variables. For instance, a table by age where the individual ages are listed might be rejected, but the same table using 5-year age categories may pass. Variables cannot be collapsed arbitrarily, since this creates a risk of residual disclosure. Different aggregations of variables have been defined and scored.

As an alternative to scoring, RDC researchers may choose to use additive controlled rounding which may be less appealing due to the distortion that it creates particularly for small cells.

4.5 Variable aggregations

When developing the scoring approach for a dataset, the bulk of the work centres around determining the score for each variable, the aggregations of variables and the threshold value for a table. Determining these scores and thresholds works best when there are many historical published tables, as with the UCR and Homicide Surveys: these tables establish the standard for disclosure control. The scores and thresholds should also be developed with the assistance of subject matter experts who are knowledgeable about the sensitivity and quality of each variable.

For consistency and ease of implementation, the threshold for UCR and Homicide tables was set to be the same. Tables with scores greater than seven may not be released in the RDCs. Also, whenever possible, for similar variables on the two surveys, the scores were
set to be the same. Significantly, the number of dimensions is very different between the two surveys: UCR tables may use up to eight variables but Homicide tables may only use up to three variables. This reflects the fact that, when working with homicide data, the risk of attribute disclosure is greater than with UCR data.

Both datasets contain variables which are not released into the RDCs. As well as being stripped of personal identifiers (e.g., names, fingerprint identifiers), other variables were removed if they were deemed to be too sensitive or of poor quality.

Other variables were aggregated before being released to the RDCs or RTRA because the original variable had sensitive, or too finely detailed, categories. For example, the variable ‘Apparent Cause of Death’ contains values such as ‘Poisoning or Lethal Injection’.

The RDC files contain several aggregated versions of a variable which the researchers can choose from (e.g., there are three versions of the variable ‘relationship between the victim and accused’). The more aggregated the variable, the lower the score. If the researcher chooses to derive another aggregation, then this is scored as the most detailed version of the variable.

Once the variables to be released in the RDCs, along with their aggregations, were determined, the variables were scored and the table thresholds set. Many UCR variables are not considered sensitive and consequently received a score of zero. These variables count towards the number of variables used to construct a table but they do not increase a table’s score.

The process of determining the scores was iterative and time consuming. The scores reflect CCJS subject matter expertise that is used to vet tables for disclosure risks and the decisions that have historically been used to release tables. Once an initial set of scores was developed, numerous tables were produced and their disclosure risk assessed. Special attention was paid to full cells and the possibility of attribute disclosure. Historical CCJS tables were used extensively to assess the robustness of the UCR and homicide variable scores.

For details on the scoring approach for the UCR and Homicide Surveys, refer to the UCR RDC Manual and the Homicide RDC Manual (Statistics Canada, 2013, a,b).

5 Status of disclosure control for police-reported data and future work

In 2013, CCJS made the UCR Survey (years 2007 to 2011) and the Homicide Survey (years 1961 to 2011) available in the RDCs and on RTRA on a pilot basis. A vetting committee was created for the pilot, consisting of RDC analysts, CCJS subject matter
experts and methodologists, in order to fine-tune the disclosure control rules as need be. So far, the scoring approach is working well and CCJS is using it to publish UCR and Homicide tables.

At the end of the two year pilot, CCJS will evaluate the effectiveness of the scoring approach and the documentation that is provided to researchers and make any necessary revisions before placing the files in the RDCs for all researchers.

Future disclosure control work for CCJS includes developing a scoring approach for the other justice administrative datasets, in particular the court and correctional services surveys.

6 Conclusion

The scoring approach proposed by the ISQ works well for crime-reported data and effectively mimics the subjective vetting of tables by subject matter experts. It works particularly well on census administrative data with an extensive history of published tables that set the standard for releasing tables. The scoring approach also works well when there are a manageable number of variables and categories within variables. While determining the scores and thresholds is time-consuming, once developed, the scoring approach is easy to apply and allows tables to be released without any distortion.

References


