Virtual Data Lab –
A more flexible approach to access Statistics Canada microdata

Kelly Cranswick, Sara Tumpane and Tina Hotton
Microdata Access Division
October 29, 2019

Delivering insight through data, for a better Canada
Legal context for the collection of national statistics

- Providing national statistics is a federal responsibility
- Under the Statistics Act, Statistics Canada is required to "collect, compile, analyze, abstract and publish statistical information relating to the commercial, industrial, financial, social, economic and general activities and conditions of the people of Canada."
- Legal authority for deemed employee access
  - Deemed employee – is any person, not currently employed by Statistics Canada, but who provides a service to Statistics Canada and requires access to information protected by the Statistics Act to perform the stated service.
- Providing access under Section 5 of the Statistics Act
  - Academic clients
  - Federal government partnerships
  - Other groups
- Providing access under Section 10 of the Statistics Act
  - Provincial/territorial government partnerships
Continuum of Access Vision

Microdata Search Tool

- Non-Secure
  - Less Sensitive Data
  - Downloadable PUMFs
  - Real Time Tabulations

- Secure
  - More Sensitive Data
  - Accreditation-Based Access
    - Authorized Workspace (Unsupervised)
    - Designated Certified Room (On-site Access)

Self-Service Based Access

- Authorized Workspace (Unsupervised)
- Designated Certified Room (On-site Access)

Downloadable PUMFs
Real Time Tabulations
Authorized Workspace (Unsupervised)
Designated Certified Room (On-site Access)
Virtual Data Lab Framework

Access framework
- Start with designated certified room (DCR) with 24/7 access for more sensitive microdata files
- Move to authorized workspaces outside DCR
- Goal is access outside of the workplace for less sensitive files

Access based on sensitivity of data
- Confidentiality Classification Tool (CCT) used in conjunction with accreditation level

Access based on accreditation level of researcher
- Shared risk framework
- Partner liability and accountability
Monitoring Access in the vDL

Statistics Canada

Protected Cloud Infrastructure

LOGIN

LOGOUT

Designated Secure Room
(highly sensitive data)

Safe Office Setting
(less-sensitive data)
Safeguarding Data: The 5 Safes Framework

The 5 Safes are considered in combination and on a sliding scale depending on type of researcher, access and data.

- Safe People
- Safe Projects
- Safe Data
- Safe Setting
- Safe Outputs
Safe Data: Confidentiality Classification Tool quantify disclosure risk

<table>
<thead>
<tr>
<th>Sensitivity (Impact scale)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>12</td>
<td>16</td>
<td>20</td>
<td>24</td>
<td>28</td>
<td>32</td>
<td>36</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>12</td>
<td>15</td>
<td>18</td>
<td>21</td>
<td>24</td>
<td>27</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

Potential Disclosure Risk

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Highly confidential

Less confidential
Safe Data: Granularity versus Utility on types of microdata files

All Microdata are anonymized (stripped of personal identifiers)
Sharing Responsibility with Host Organizations

Safe Organization
- Organizational Agreements (MOUs)
- Trusted Individual Responsible for Access Controls (TIRAC)
- Identify trustworthy applicants
- Shared maintenance of physical and IT security
- Shared monitoring and investigations

Safe People
- Security Clearance & Researcher Contracts
- Test of awareness and knowledge of security measures
- Attestation of compliance with protocols
- Accreditation

NEW

DELIVERING INSIGHT THROUGH DATA FOR A BETTER CANADA
DataLab End State and Timeline

• Remote data access (24-7);
• Secure IT connection, using a protected Cloud application;
• Data access for all levels of government and organizations, beyond the academic community;
• Work with Organizations to share the responsibility for protecting confidential data.

DataLab promotes evidence-based research while maintaining confidentiality of data.
THANK YOU!

For more information about access programs:

http://www.statcan.gc.ca/eng/help/microdata?HPA=1

Tina Hotton
Acting Chief, Microdata Access Division
Tina.Hotton@Canada.ca

Kelly Cranswick
Chief, Microdata Access Division
Kelly.Cranswick@Canada.ca