

# DATA ACCESS DIVISION NEWSLETTER

## SPRING 2021 EDITION

### A MESSAGE TO OUR STAFF AND CLIENTS

With the arrival of spring and warmer weather comes a sense of hope with the development of new vaccinations across the globe as we move beyond the one-year mark since the first COVID-19 lockdown. The Data Access Division (DAD) would like to take a moment to thank its beloved staff. The success of our program comes from the hard work and dedication that each member has continued to show collectively throughout these changing times. We could not have achieved our advancements without your great efforts and continued collaboration. We would also like to thank our clients and friends for their continued patience and support as we are constantly reminded of how fortunate we are to be part of such a strong community. We remain devoted to continuing our work to ensure that you are provided with the real-time data and services that you need.

### CELEBRATING ACCOMPLISHMENTS AND FOCUS FOR THE UPCOMING YEAR

DAD would like to highlight and celebrate some of its greatest accomplishments within the last few months. The Self-Serve Access (SSA) section provided virtual access to clients and successfully onboarded new users to the Public Use Microdata File (PUMF) Collection, provided free access and created accounts for research data centre (RDC) researchers for Real Time Remote Access (RTRA), and had 82 Data Liberation Initiative (DLI) member institutions. The Virtual Data Lab (VDL) team successfully onboarded its first set of users in the new cloud environment for its first pilot project back in February, bringing the team one step closer to a production environment. DAD, in partnership with the

Canadian Research Data Centre Network (CRDCN), recently opened two new RDCs! One centre is located at Carleton University in Ottawa, and a new satellite centre opened its doors to researchers at the University of Calgary. In addition, the first Business Research Microdata (BRM) file that uses real, not synthetic, data was released to the RDCs. To help support researchers working in RDCs and in the VDL, we have produced a series of short training videos to help support researchers in producing their statistical output for release.

For the upcoming year, DAD will continue to focus on collaboration efforts with various teams and partners. We will focus on leveraging new technologies to help drive Statistics Canada's modernization efforts by developing new and innovative ways to access microdata, such as developing the Virtual Research Data Centre (vRDC) and the VDL, increasing granularity while meeting researcher needs, and continuing to provide the research community increased and faster access to data to support better decision making for policies and programs across the country. In the RDC Program, we will be opening another centre to researchers, as well as increasing our business data holdings.



# DATA ACCESS DIVISION NEWSLETTER

## SELF-SERVE ACCESS

### DATA LIBERATION INITIATIVE TEAM UPDATES

Welcome to another DLI program membership year.

The SSA section has added the following new services to the program:

- one free RTRA account per institution starting April 1, 2021
- limited number of free custom tabulations
- access to training offered by Regional Services.

### EXTERNAL ADVISORY COMMITTEE

The External Advisory Committee (EAC) sent a call-out in February to the Listserv for two volunteers to represent the Atlantic and Ontario regions.

The SSA section and the EAC would like to welcome Jane Fry as an Ontario Region representative and to thank two members who have stepped down from the committee: Peter Webster, Co-Chair of the EAC and Atlantic Region representative, and Claire Wollen, Ontario Region representative.

### PROFESSIONAL DEVELOPMENT COMMITTEE

The chair of the committee, Alex Cooper, sent an email to the Listserv confirming to the DLI community that face-to-face training has been cancelled again this year, but we will be doing national training again with regional sessions.

The Professional Development Committee (PDC) sent a call-out to the Listserv in March for a volunteer to represent the Quebec Region.

The SSA section and the PDC would like to take this opportunity to welcome Vivek Jadon from McMaster University as the new Ontario Regional Training Coordinator.

The PDC is working on several initiatives:

- Contacts and Alternates Survey – a working group is in place to revise the survey
- DLI Training Repository – the committee is looking at options
- colleges – a sub-committee met with college representatives from each region to discuss their needs
- training – a working group is in place to discuss training needs and coordinate with other data-centric organizations, such as the Canadian Research Data Centre Network and Portage.

### STATCAN WEB REDESIGN PROJECT

The newly designed [DLI website](#) is now live! The website includes the updated DLI Contact and Alternate's Survival Guide, information on the program, and resources. You can also easily navigate through the different data access programs.

### PUBLIC USE MICRODATA FILES ONLINE PROJECT

We are working on putting PUMFs online in a downloadable format. Newly released PUMFs are being added to the website as they become available, and older PUMFs are being added in phases. As part of this project, digital object identifiers are being assigned to PUMFs.

# DATA ACCESS DIVISION NEWSLETTER

## SELF-SERVE ACCESS

### DATA RELEASES TO DLI SINCE JANUARY 2021

- National Travel Survey (NTS) 2019 PUMF
- General Social Survey (GSS) Cycle 33, 2018 PUMF
- Canadian Perspectives Survey Series (CPSS) 5 PUMF
- Labour Force Survey (LFS) January 2021 PUMF
- Provincial Symmetric Input-Output Tables (2016 and 2017)
- December 2020 Business Counts
- Input-Output Multipliers Link 1961
- Hate Crimes (Province) Table E and Table F
- Human Trafficking Data Table
- Postal Code Conversion File Plus (PCCF+) Version 7D, November 2020
- Labour Force Survey (LFS) February 2021 PUMF
- Canadian Housing Survey (CHS) 2018 and 2019 PUMF

A list of all DLI products is available on the website:

<https://www.statcan.ca/eng/microdata/dli>.

### REAL TIME REMOTE ACCESS UPDATES

RDC researchers have had their access extended to March 31, 2022.

### SAS ASSISTANT

The graphical user interface has been launched! The number of surveys available is currently limited. However, more surveys will be added throughout the year.

The SAS Assistant will help users with little SAS experience to generate successful tables. You will be able to use buttons and dropdown menus to build your SAS code, and your code is created as you select the variables.

### DATA RELEASES TO RTRA SINCE JANUARY 2021

- Labour Force Survey (LFS) – monthly
- Registered Apprenticeship Information System (RAIS) 2019 (January 2021)
- Crowdsourcing 1: Impacts of COVID-19 on Canadians – All weeks
- Crowdsourcing 2: Impacts of COVID-19 on Canadians – Your Mental Health
- Crowdsourcing 3: Impacts of COVID-19 on Canadians – Perceptions of Safety
- Crowdsourcing 4: Impacts of COVID-19 on Canadians – Trust in Others
- Crowdsourcing 5: Impacts of COVID-19 on Canadians – Parenting During the Pandemic
- Crowdsourcing 6: Impacts of COVID-19 on Canadians – Living with Long-term Conditions and Disabilities
- Crowdsourcing 7: Impacts of COVID-19 on Canadians – Experiences of Discrimination
- Survey of Household Spending (SHS) 2005
- Survey of Household Spending (SHS) 2007
- Survey of Household Spending (SHS) 2009

A list of all RTRA products is available on the website:

<https://www.statcan.ca/eng/microdata/rtra>.

# DATA ACCESS DIVISION NEWSLETTER

## RESEARCH DATA CENTRES

### RESEARCH DATA CENTRE UPDATES

While RDCs are still operating under reduced capacity because of COVID-19 restrictions, we are excited to announce the opening of a new centre at Carleton University and new satellite centre at the University of Calgary. Researchers now have access to two sites in both Ottawa and Calgary to facilitate demand for access across the cities.

The high-level Joint Task Force, co-chaired by Martin Taylor (Executive Director, CRDCN) and Jacques Fauteux (Assistant Chief Statistician, StatCan), focused on developing and aligning data access strategies for academic researchers and provided a report to the CRDCN executive board and the Chief Statistician in February. The report gave an overview of the technical infrastructure of the vRDC, identified possible intersections with the VDL platform, and highlighted critical business questions that will be explored in partnership between the CRDCN and StatCan in the coming months. Presentations on the report and next steps will be provided to StatCan staff and CRDCN academic directors in April.

We are pleased to announce that our pilot to test the VDL access platform with academic researchers was launched in March! Starting with selected projects in four universities, a StatCan cloud infrastructure platform will make StatCan microdata securely accessible outside RDC facilities. This pilot project involves the University of Toronto, Université de Montréal, McMaster University, and University of Calgary to start, but will be implemented on a broader scale after pilot testing is complete and the vRDC infrastructure is available. For more information on the pilot testing, please see the Modernization of Access section below.

### NEW RESEARCH DATA CENTRE HOLDINGS

On February 23, the CRDCN hosted a very well-attended webinar on the BRM. This will be the first business dataset released to the RDCs using real, not synthetic, data. It is expected that it will take extra time for researchers to become familiar with the data. As well, vetting will take longer than for a typical social data project because we will be using the BRM to create and test business data vetting rules. For these reasons, it is not recommended that students undertake their thesis or dissertation work with the BRM. Researchers can start applying for access in April.

A total of 25 products were added to our data holdings in the fourth quarter of the 2020/2021 fiscal year. These include two new surveys (Survey of Digital Technology and Internet Use and Impacts of COVID-19 on Health Care Workers: Infection Prevention and Control), one new linked data file (Survey of Approaches to Educational Planning linked to Postsecondary Student Information System—Registered Apprenticeship Information System—T1 Family File), as well as updated survey cycles and administrative files.

### NEW DATA ACCESS TRAINING VIDEO SERIES

#### PARTIAL LIST OF DATA FILES UPDATED FROM JANUARY TO MARCH 2021

- Canadian Health Survey on Children and Youth (CHSCY) 2019
- Vital Statistics - Death Database (VSDD) 2019
- Survey of Household Spending (SHS) 2017
- General Social Survey (GSS) – Caregiving and Care Receiving, Cycle 33
- General Social Survey (GSS) – Victimization, Cycle 34
- Longitudinal Immigration Database (IMDB) 2019
- Canada Education Savings Programs (CESP) linked to 2016 Census
- Registered Apprenticeship Information System (RAIS)

For a complete list of data available in RDCs and government access centres, visit [www.statcan.gc.ca/eng/rdc/data](http://www.statcan.gc.ca/eng/rdc/data).

# DATA ACCESS DIVISION NEWSLETTER

## RESEARCH DATA CENTRES

We will soon be releasing our Data Access Training Video Series! These short videos will provide practical instruction to StatCan microdata users on a variety of topics related to access and data analysis. Our initial series of videos will cover how to prepare your output for confidentiality vetting, such as applying rounding techniques and testing for homogeneity and dominance using a variety of statistical software packages.

### GOVERNMENT DATA ACCESS

#### FEDERAL RESEARCH DATA CENTRES

The Government Data Access team has started plans to merge the Federal Research Data Centre (FRDC) and the Social Data Access Centre and the Business Data Access Centre (formerly known as CDER), both located at Tunney's Pasture, into one location by the end of spring 2021. This new centre will provide access to both social and business data for federal government users. The integration of business and social data access into one physical location is a significant step in completing the full integration of the Business Data Access program under the FRDC umbrella.

DAD is also working with Employment and Social Development Canada (ESDC) to facilitate access and address research needs for the emergency response data linked to StatCan datasets. ESDC will undergo an accreditation process and training, and data access will be available from home for researchers.

### PROVINCIAL SECURE ACCESS POINTS

Two provincial secure access points continue to operate in British Columbia and Alberta. Two more sites will open in spring 2021 in Ontario at the Ministry of Finance and the Ministry of Children, Community and Social Services. For more information about these initiatives, please contact [statcan.dad-ssap-dad-pass.statcan@canada.ca](mailto:statcan.dad-ssap-dad-pass.statcan@canada.ca).

### WELCOME SHELLEY JEGLIC!

We would like to welcome Shelley Jeglic, who has joined DAD as the new chief of the Government Data Access Program. Shelley previously worked in the Centre for Population Health Data and is excited to make the switch from client to service provider.



# DATA ACCESS DIVISION NEWSLETTER

## MODERNIZATION OF ACCESS

### PILOT PROJECTS AND TESTING

The VDL team is excited to announce that eight researchers from the University of Toronto, University of Calgary, McMaster University and Université de Montréal, and four researchers from the Public Health Agency of Canada are now accessing the VDL environment as part of the academic pilot! They are accessing microdata files with low to medium levels of sensitivity. This is the first set of users to be granted access to anonymized microdata in the cloud environment using the VDL. The VDL team has been working hard to establish and expand governance and system capabilities to support this virtual access. This is a big milestone for the team, and a step forward in enabling secure virtual access to data for more accessible research, and, in turn, better decision making. We could not have reached this level of success on this project and our pilots without the collaborative effort of many teams and individuals.

Going forward, the VDL project will continue to onboard users identified for our pilots. The established pilots will help evaluate the nuances of onboarding different types of researchers, as well as help inform how the user experience of the environment can be improved leading up to production. Once the VDL team has successfully conducted the pilots using data with low to medium levels of sensitivity, the team will conduct additional pilots using data with medium to high levels of sensitivity, as this is more representative of the data that are typically used by researchers. StatCan will select pilots based on several criteria to learn about and improve the VDL process to meet the project's objectives.

Overall, with the VDL, StatCan will be better positioned to advance its user-centricity by introducing this new mode of access and contribute to the agency's modernization efforts.

### VIRTUAL DATA LAB PROJECT UPDATES

The VDL will vastly improve access to statistical information for researchers by providing users 24/7 remote access to data housed at StatCan using a secure IT connection and a protected cloud environment. Progress is ongoing on a number of key initiatives to increase virtual data access and promote collaboration. These include the development of analytics platforms and monitoring capabilities, and continued assessment and development of the Client Relationship Management System (CRMS) and the Microdata Search Tool.

A number of monitoring mechanisms have been approved and are available to StatCan for deemed employees to access protected microdata in the cloud environment. Staff from DAD will use a variety of mechanisms to monitor for potential security incidents and follow the established incident protocol when required, which supplements the Information and Privacy Breach Protocol at StatCan.

Development on the CRMS corporate project continues under the Dissemination Division. The Dissemination team is currently working with IT to establish development priorities and DAD for pilot assessments. Stay tuned for more information!

**Questions or comments? Visit our website:** [www.statcan.gc.ca/eng/help/microdata](http://www.statcan.gc.ca/eng/help/microdata).

**Check out the StatCan Blog:** [www.statcan.gc.ca/eng/blog](http://www.statcan.gc.ca/eng/blog).

**Don't forget to follow us on social media!**



[Facebook](#) | [Instagram](#) | [LinkedIn](#) | [Reddit](#) | [Twitter](#) | [YouTube](#)



Statistics  
Canada

Statistique  
Canada

Canada