Overview

About the program

The Data Liberation Initiative (DLI) is a partnership between postsecondary institutions and Statistics Canada with the goal of improving access to data resources. The DLI is a program within Statistics Canada's continuum of microdata access services. Over the years, the focus of the DLI has evolved from purchasing access to major Statistics Canada datasets to providing the training services and continuous support required for the proper understanding and use of an expanding data collection. For more information, including on the benefits of DLI membership and the history of the program, please visit the [DLI website](#).

Role of the DLI contact

As the DLI contact for your institution (a role shared by a network of contacts across Canada), you play an essential part in promoting access to Canadian data resources.

DLI contacts generally provide assistance in finding, accessing and analyzing Statistics Canada data and products. While contacts may not always be experts in statistical software and data analysis, they are able to refer users to others in their institution or to the DLI community (through the listserv) who can help answer technical and methodological questions.

DLI contacts have the following responsibilities:

User support: DLI contacts assist faculty members, staff and students with using Statistics Canada resources.
Licensing: DLI contacts ensure that the conditions of use of the DLI licence agreements are being abided by at their institutions.
Membership renewal: DLI contacts ensure that the annual membership fees are paid.
Liaison: The DLI contact communicates with member institutions through their network of contacts. These communications involve licensing changes, updates to member services, and outreach and professional development sessions.
Access: DLI contacts ensure that the program has current institutional intellectual property (IP) ranges to maintain IP-based access to resources.

Governance: DLI contacts may be asked to vote for their Regional Training Coordinator (RTC) if more than one person volunteers for an open position.

In addition to conducting a census every five years, Statistics Canada conducts a wide range of surveys on virtually all aspects of Canadian life. Statistics Canada makes data available to support research, industry and policy development. Understanding some basic data concepts will help inform what products are available through which access programs.

Data terminology

Aggregate data

Information derived directly from statistical microdata files or statistical aggregate files. Unlike statistical microdata files, aggregate statistics do not record information at the level of individual units of observation. In other words, they are the result of grouping data at an aggregate or macro level (e.g., people in a specific age group, businesses or organizations in a particular industry, or households in a particular region).

Microdata file

A structured file containing information on individuals, businesses or organizations. A microdata file may come from a census of all units or from only a sample of units. In addition, the file may be the product of direct collection for statistical purposes, an administrative file where the statistical uses are not the primary purpose of the collection or a combination of the two.

There are three types of microdata files:

- Master files: For each survey conducted, a master file is constructed, which contains all responses from each respondent recorded in the format specified on the questionnaire. Access to master files is available only through the research data centres (RDC) by application.
- Synthetic files: Continuing with the focus on offering new alternative access options, Statistics Canada is investing in researching methods for creating synthetic data. Synthetic data can take on a variety of forms and possess a range of quality characteristics, but the main goal is always to offer a microdata access option that poses little or no disclosure risk and—therefore—can be released to the general public.

- Public use microdata files (PUMFs): PUMFs consist of sets of records that contain information on individuals or households (microdata). They are non-aggregated data that are modified carefully then reviewed to ensure that no individual or business is identified directly or indirectly.

Documentation

Statistics Canada releases supporting documentation along with its microdata files. This documentation is needed for the use and interpretation of microdata files and can include survey questionnaires, instructions to interviewers, codebooks, user guides, record layouts, data dictionaries, frequency files and cv tables, among others.

Administrative data

Administrative data are information collected by governments or private-sector organizations as part of their ongoing operations, for example, records of births and deaths, taxation records, records of the flow of goods and people across borders, and data collected by satellites. Like most other statistical agencies, Statistics Canada uses administrative data instead of or in addition to survey data and to support statistical operations.

Data terminology resources

- Statistics Canada's Definitions, data sources and methods: This information is provided to ensure an understanding of the basic concepts that define the data, including variables and classifications, the underlying statistical methods and surveys, and key aspects of the data quality. Direct access to questionnaires is also provided.
- Statistics: Power from Data! glossary: These definitions provide information to those who have statistics-related questions but who do not require highly technical explanations.

Continuum of microdata access

Access to microdata is made available through a variety of dissemination channels.

The following table outlines the dissemination channels available for aggregated data and microdata.

Open statistics --> **Restricted data**

Statistics Canada website	Data Liberation Initiative	Product sales and customized tabulations	Real Time Remote Access (RTRA) program	Research data centres
Who can access: General public	**Who can access:** Students, faculty members and staff of member postsecondary institutions	**Who can access:** Individual members of organizations	**Who can access:** Individual members of organizations, postsecondary students, governments with a membership	**Who can access:** Approved researchers (individual members of organizations, postsecondary students, governments)
Conditions: Statistics Canada Open Licence	**Conditions:** The majority of products fall under the Statistics Canada Open Licence. Access to products outside of the open licence is restricted to teaching, research and statistical purposes. Please refer to Application process and guidelines for more information.	**Conditions:** Purchase confirmation between Statistics Canada and individual members of an organization	**Conditions:** RTRA agreement and Statistics Canada open data licence	**Conditions:** Deemed Statistics Canada employee status
Data available: Electronic standard data	**Data available:** Standard data products, public use microdata files,	**Data available:** Tables from confidential files that are specially processed by Statistics Canada for a fee.	**Data available:** "Dummy" microdata files for various social survey and administrative datasets,	**Data available:** Confidential microdata files and administrative datasets

products and publications	postal code data products, etc.			which provide statistical table outputs.	
Mode of access: Available on the Internet	**Mode of access:** Electronic file transfer service	**Mode of access:** Custom tabulation distributed to client		**Mode of access:** Electronic file transfer service	**Mode of access:** A secure research data centre

Not all surveys and statistical programs produce data products. Many divisions do not create PUMFs, as these are costly to produce and must be vetted by the Microdata Release Committee (Statistics Canada's confidentiality control for microdata files). Some divisions create only standard tables that are available through the Statistics Canada website and charge retrieval fees for more in-depth requests (e.g., custom tabulations). Although the data may be freely available, cost recovery charges apply for the analyst's time.

Governance

External Advisory Committee

The DLI is guided by its External Advisory Committee (EAC). The EAC meets biannually and is composed of representatives appointed from DLI member institutions, Statistics Canada and external organizations. For a current listing of DLI-EAC members, visit the Governance section of the DLI website.

Professional Development Committee

The DLI Professional Development Committee (PDC), which reports to the EAC, is responsible for the ongoing development of a data services curriculum for postsecondary staff who support the DLI at their institutions. The PDC consists of eight RTCs, one college representative, a chair and a DLI section representative. For a current listing of PDC members, visit the Governance section of the DLI website.

Regional Training Coordinators

Two RTCs for each of the four regions (Atlantic Canada, Quebec, Ontario and Western Canada) and one college representative sit on the DLI-PDC, and they are responsible for

- identifying training needs within their region

- communicating those needs to the PDC both for the purpose of budgeting for training and for coordinating national training activities
- organizing local training events
- developing their local region's training program.

DLI contact

Member institutions designate their DLI contact and alternate. The DLI contact is responsible for promoting and facilitating access to Statistics Canada resources and ensures that the DLI licence is followed. See the Manage your membership section below for more information on the DLI licence.

Member institutions must have one DLI contact, but the selection of an alternate is optional. The DLI contact and alternate need to be familiar with the DLI and Statistics Canada resources to be able to assist users with their data-related questions. Additionally, it is advised that the DLI contact be familiar with the resources available on campus to assist users with data-related questions, such as the use of statistical software, in the event the DLI contact does not have those skill sets.

See the User community for a list of contacts at each DLI member institution.

Changing a DLI contact

If your institution's DLI contact changes, please advise the Self-Serve Access section. The DLI contact's contact information should also be updated if the current DLI contact goes on extended leave (e.g., sabbatical, maternity leave). Please provide the date on which the change will be effective and the name of the new contact, as well as the person's position title, mailing address, email address, phone number and fax number.

Manage your membership

When a DLI contact is identified, they are provided with access to DLI resources, including the electronic file transfer (EFT) service and the mailing list (dlilist).

Electronic file transfer password

The DLI EFT site is a repository used to disseminate the DLI collection. Users of the EFT are limited to an institution's DLI contact or alternate. The EFT requires that each user have their own unique user ID and password. When a new contact has been

identified, the DLI unit sends the EFT account information via email. To request a password reset, contact the [Self-Serve Access section](#).

dlilist

The DLI listserv is used by DLI contacts to get information on the DLI collection and data licences, as well as to provide feedback about Statistics Canada products and services.

The dlilist is a subscription-based listserv, which means that only registered users can post and receive messages. Messages from the list are sent to all registered users by email.

The listserv home page is available at https://dli-idd-listserv.statcan.gc.ca/scripts/wa.exe?A0=DLILIST.

If you are attempting to connect off campus, you will need to connect using a VPN.

Subscribing and unsubscribing

- To subscribe to the dlilist, email the [Self-Serve Access section](#).
- To post a message to the dlilist, email dlilist@idd-dli.statcan.gc.ca.
- To unsubscribe from the list, send a blank email to DLILIST-signoff-request@idd-dli.statcan.gc.ca.

Disclaimer

The dlilist is an opt-in listserv. By using this service, you agree that your email address and any communications will be made available to the other dlilist users. All communications will be archived in Statistics Canada's mailing list archive. The opinions expressed are those of the dlilist users and are not representative of Statistics Canada.

Dlilist archives

Messages from the dlilist are archived and kept in a protected, searchable archive that can be accessed by DLI contacts.

2014 to present: [dlilist archives](#)

Membership renewal

A DLI annual membership runs from April 1 to March 31 of the following year. DLI memberships are renewed on an annual basis between April and June each year.

DLI contacts are responsible for making sure that the annual membership fees are paid. Some member institutions assign the paying of membership fees to a specific department in the library. Others have the invoicing fees sent to the DLI contact who coordinates payment internally. To change who the purchase confirmation and invoice should be directed to, contact the Invoicing section.

Learn

Training sessions

Every year, the DLI conducts one training session in each of its four regions: Atlantic Canada, Quebec, Ontario and Western Canada. These multi-day sessions are open to anyone who provides services for the DLI. However, priority goes to primary DLI contacts and alternates. The DLI hosts a national training session approximately every four years (usually in conjunction with IASSIST being held in Canada). This is the opportunity for the entire DLI community to meet.

RTCs are responsible for organizing the training in each of their regions with support from the DLI unit. Topics range from basic data service skills to advanced sessions that build on prior training. These training sessions allow DLI contacts to learn from one another and from Statistics Canada subject-matter experts.

Travel subsidies

Financial support for transportation to training is offered to each DLI contact or their representative to attend one training session per fiscal year. All travel requests must be approved by the DLI unit before being booked. If a contact or alternate gives a presentation at the training session, additional subsidies may be made available. For more information, visit the Governance section of the DLI website.

Statistics Canada data literacy training initiative

The data literacy training initiative provides a wealth of resources aimed at those who are new to data or those who have some experience with data but may need a refresher or want to expand their knowledge. The goal of this initiative is to provide learners with the basic concepts and skills related to a range of data literacy topics, including What is Data? An Introduction to Data Terminology and Concepts and Types of Data: Understanding and Exploring Data.

Data Access Division (DAD) Newsletter

The purpose of the DAD Newsletter is to inform subscribers and users of ongoing divisional initiatives. This includes providing updates on DLI projects and local data-related and modernization initiatives, as well as any updates on our other data access modes, such as RTRA and RDCs.

Feedback, ideas and DLI submissions for future issues of the newsletter are welcome. Please send them to the Self-Serve Access section.

Training Repository

The DLI Training Repository contains workshop presentations from DLI training and from conferences. The DLI unit is responsible for uploading presentations and materials to the repository after each training session. These materials are available for anyone to view and download.

For more information on the Training Repository, visit the repository web page. For more details about the history of the Training Repository, consult the presentation titled Creating a Repository of Training Materials: The Canadian Experience by Jane Fry from Carleton University (English only).

Data Interest Group for Reference Services

Hosted at the University of Alberta, Data Interest Group for Reference Services (DIGRS) content is based primarily on questions and answers from the DLI listserv from 2004 to the present. The content is presented in a user-friendly manner and information can be retrieved through keyword searches or by searching by date or category.

Citing data

The importance of citing data

Bibliographic references are important when using the data or ideas of others in your written work. References credit your sources and allow your readers to find those sources. Additional information is available here: How to Cite Statistics Canada Products.

Access

What is in the DLI collection?

The DLI collection is composed primarily of standard products produced by Statistics Canada, including PUMFs, aggregated data tables and boundary files. Licensed collections include sample files from the Discharge Abstract Database (DAD) from the Canadian Institute for Health Information (CIHI), postal code data products from Canada Post and the Social Policy Simulation Database and Model (SPSD/M).

Electronic file transfer site

About the site

The DLI EFT site is the data repository of the DLI collection. To ensure the absolute protection of data files, the EFT requires that each user have their own unique user ID and password.

The EFT service supports a file transfer protocol (FTP) standard for sending and receiving files. DLI contacts will need an FTP application, such as WS_FTP or filezilla, to access the EFT site.

Understanding the directory structure

The DLI EFT collection contains five subdirectories, which are outlined in the table below. Some DLI contacts may not be able to view all of the directories if your institution has not signed the appropriate DLI licences (e.g., Postal Code Conversion File [PCCF] or SPSD/M).

Safe name	Contents	Licence
MAD_PUMF_FMGD_DAM	Survey public use microdata files and metadata, organized according to their survey record number, acronym and year[1]	Statistics Canada Open Licence Agreement
MAD_DLI_IDD_DAM	DLI annual reports, DLI training materials, CD-ROM data products, geography files, Census of Population and Census of	Statistics Canada Open Licence Agreement

---

1. Readme-Key_Lisezmoi-cle.xls lists all PUMFs by survey name, acronym and record number for easier searching of data files.

	Agriculture files, aggregate data files,[2] and more	
MAD-PCCF_FCCP_DAM	Postal Code Conversion File, Postal Code Conversion File Plus and Postal Codes by Federal Ridings File	Section 1 – Postal Code OM Conversion File (PCCF) Access: PCCF Licence
MAD_CIHI_ICIS_DAM	Discharge Abstract Database from the Canadian Institute for Health Information	Section 4 – Discharge Abstract Database (DAD) Analytic File Access – DAD Licence
MAD_SPSDM_BDMSPS_DAM	Social Policy Simulation Database and Model	Section 3 – Social Policy Simulation Database and Model (SPSD/M) Access: SPSD/M Licence

MAD_PUMF_FMGD_DAM

Each year of a survey is usually contained in a separate subdirectory. The secondary level in the survey breaks down the information based on data (*data*) and documentation (*doc*). The readme file for the survey is also found at this level. The data folder contains a zipped file with the data. The data can take the form of microdata in ascii, SPSS, STATA or SAS format. The documentation folder includes the metadata, which is the information required to interpret and understand the microdata.

For example:

/MAD_PUMF_FMGD_DAM/Root/
 /3250_APS_EAPA
   /1991
   /2001

2. other-products_autres-produits.xls lists all aggregate data files by survey name, acronym and record number for easier searching of data files.

```
/2001-Children
/2006
  /age-06-14
  /age-15+
    /data
    /doc
    lisezeapa2006-age-15+.txt
    readaps2006-age-15+.txt
/3251_PALS_EPLA
```

With respect to the General Social Survey (GSS) folders, please consult the EFT key titled Readme-Key_Lisezmoi-clé.xls. A note of the GSS cycles (far right column) indicates the associated cycles.

	Title-ENG	Acronym-ENG	Record Number / Numéro d'enregistrement	Acronym-FRA	Titre-FRA	Notes
4502_GSS-Care_ESG-Soins	General Social Survey - Caregiving and Care Receiving	GSS-Care	4502	ESG-Soins	Enquête sociale générale - Les soins donnés et reçus	2012 Cycle 26; 2007 Cycle 21; 2002 Cycle 16; 1996 Cycle 11;
4503_GSS-Time_ESG-Temps	General Social Survey - Time Use	GSS-Time	4503	ESG-Temps	Enquête sociale générale - L'emploi du temps	2015-2016 Cycle 29; 2010 Cycle 24; 2005 Cycle 19; 1998 Cycle 12; 1992 Cycle 7; 1986 Cycle 2;

MAD_DLI_IDD_DAM

Census folders

Census folders are organized by census year. The way each census-year folder is organized varies from year to year. Generally, for the Census of Population, folders within a census year are organized either by data type (e.g., b2020, PUMF) or by topic (e.g., labour, income). For the Census of Agriculture, the way the folders are organized can vary by data type (e.g., Excel), geography (e.g., small area, agricultural region), or data and documentation. Sometimes the quickest way to find a census file is to email the DLI list asking where it is.

Geography folder

The geography folder is initially broken down by census year. The secondary level of breakdown identifies the type of information sought. For example, a user may be seeking reference maps, boundary files or a specific product. The readme once again becomes a critical tool for navigating the folder.

Reports folder

The reports folder contains materials of particular interest to DLI contacts, such as the EAC biannual report, DLI updates, and meeting minutes for both the EAC and PDC. In addition, users can find the images of the new DLI graphic identifier.

Other-Autres folder (e.g., data tables, CD-ROM products)

The other folder provides a listing of additional data products organized according to their survey record number or catalogue number and corresponding survey or product acronym. The DLI unit has begun using this naming convention to conserve server space and harmonize both official languages. In addition, users seeking clarity on the record numbers or acronyms are invited to use the Excel workbook housed in the folder (other-products_autres-produits.xls). The workbook functions are key for helping to explain the nomenclature. The CD-ROM products found in this folder are available in a zipped format for download. Occasionally, the user will be need to download the contents of the CD-ROM, unzip them and then burn them onto a CD-ROM (this will be noted in the file's readme). Many of the products have unusual proprietary structures and—as a result—must be run from a CD instead of being downloaded to a hard drive.

MAD_CIHI_ICIS_DAM

The CIHI safe contains sample files from the DAD. Data for 2009 onward are currently available in clearly labelled subfolders.

MAD_PCCF_FCCP_DAM

The PCCF safe is initially broken down by census year. The secondary level of breakdown identifies the postal code data product:

PCCF (folder: pccf-fccp)

Postal Codes by Federal Ridings File (folder: pcfrf-fcpcef)

Postal Code Conversion File Plus (folder: ppcf-fccp-plus)

Within each subfolder, a readme file provides a product description and a summary of changes to the product (e.g., starting in June 2013, the PCCF is available only as a standard package for Canada [no longer available at the province level] and is updated and released annually [previously released semiannually]). The readme file also lists the title of the product (e.g., PCCF for August 2015), the release date (e.g., February 12, 2016), the frequency of release and the directory.

MAD_SPSDM_BDMSPS_DAM

In 2016, the DLI unit created the SPSD/M safe. Subfolders are labelled by version, each containing its own unique install files and instructions. Please consult the readme files housed within the folders for more information.

File-naming convention

Files located on the EFT site follow a similar naming convention. When the files are received from the author division, they are renamed to fit the DLI naming convention. Therefore, a file produced by the subject-matter division may differ from the file located on the DLI EFT site.

Files are initially named according to the survey acronym followed by the year or cycle of the survey, then by type of document.

Documents and their extensions

If a file is updated or replaced by the subject-matter division, an additional extension will be added identifying the version number, for example,

- User Guide for the 2003 Household Internet Use Survey: hius2003gid.pdf
- Questionnaire for the 2009 Survey of Household Spending: shs2009que.pdf.

readme files

The readme file is a quick reference guide to the DLI EFT site. Once in a folder (e.g., survey, census, geography), the readme file provides a breakdown of the contents of the folder. This includes not only the file names, but also longer titles, which allows users to identify the file they are looking for. The readme file also includes the size and length of the data file for PUMFs to perform a quick verification that the file was transferred from the EFT to the user's computer successfully.

Retrieving files from the EFT site

Using specialized FTP software, access and log into the DLI EFT site. The hostname, userid and password are provided by the DLI unit.

When you have located the files you wish to download, mark them and transfer them to your computer. Make sure to select the receiving folder on your computer before initiating the transaction.

Another useful tip is to set the **transfer mode** based on the type of file you are transferring. A good idea is to set your default to **auto** so that the program selects the right transfer mode based on the file extension. As a general rule, all files should

be downloaded in **binary** except for files with the following extensions: .txt, .sps, .sas and .dat.

Once you have downloaded the data files, decompress (unzip) them as necessary. If the documentation is zipped, it will require unzipping.

If you encounter any problems, please contact the [Self-Serve Access section](#).

Requesting data not found in the DLI collection

If you identify a product that you think should be a part of the DLI collection, please submit your request on the dlilist.

Secondary data distributors

Computing in the Humanities and Social Sciences ([CHASS](#)) (institutional membership required)

CHASS (available in English only) is a computing facility in the Faculty of Arts and Science at the University of Toronto. It offers a collection of social sciences and general interest databases (e.g., International Monetary Fund and World Bank tables) and two databases with Statistics Canada data:
  ● [Canadian census – census profile data at various levels of geography back to 1961](#)

[CANSIM on CHASS](#): **Microdata analysis and subsetting with survey documentation and analysis (SDA) on CHASS** (institutional membership required)

[SDA @ CHASS](#) is a set of programs for the documentation and web-based analysis of survey data. SDA also has procedures for creating customized subsets of datasets.

Abacus Dataverse Network (institutional membership required)

[Abacus Dataverse Network](#) is the research data repository of the British Columbia Research Libraries' data services—a collaboration involving the data libraries at Simon Fraser University, the University of British Columbia, the University of Northern British Columbia and the University of Victoria.

Ontario Data Documentation, Extraction Service and Infrastructure Initiative (ODESI) (institutional membership required)

[ODESI](#) is a digital repository for social science data, including DLI data and a range of public opinion polls. It is a web-based data exploration, extraction and analysis tool created by Scholars Portal—a service of the [Ontario Council of University Libraries](#) (OCUL)—and is available to authorized users from Ontario universities and subscribing institutions. ODESI metadata are openly available to the world.

Scholars Portal Dataverse (institutional membership required)

[Scholars Portal Dataverse](#) is a repository primarily for research data collected by researchers and organizations provided by Scholars Portal on behalf of OCUL and other participating institutions, although anyone is welcome to use it to deposit, share and archive data.

Données statistiques et géographiques

[Données statistiques et géographiques](#) was developed by the Quebec university libraries and enables access to DLI geographic products and aggregate data. Access to data is restricted to students, researchers and professors at participating universities.

Open Data – Government of Canada

Search [open data](#) that are relevant to Canadians, learn how to work with datasets and see what people have done with open data across the country.

Canadian Century Research Infrastructure

The Canadian Century Research Infrastructure (CCRI) is a pan-Canadian, multidisciplinary and multi-institutional effort to develop a set of interrelated databases centred on data from Canadian censuses from 1911 to 1951. The [CCRI gateway website](#) is hosted by the University of Alberta and provides access to microdata, as well as a geographical framework developed to enable the location, selection, aggregation and analysis of census and contextual data.

Glossary

A

Abacus
    Abacus Dataverse Network is the research data repository of the British Columbia
    Research Libraries' data services—a collaboration involving the data libraries at Simon

Fraser University, the University of British Columbia, the University of Northern British Columbia and the University of Victoria.

Administrative data
> Information collected by government or private-sector organizations as part of their ongoing operations.

Aggregate data
> Statistics that are organized into a data structure and stored in a database or data file. The data structure is based on tabulations organized by time, geography or social content.

C

Canadian Century Research Infrastructure
> A pan-Canadian database centred on data from the Canadian census from 1911 to 1951.

Canadian Institute for Health Information
> An independent, not-for-profit organization that provides essential information on Canada's health systems and the health of Canadians. The institute has shared access to sample files from the Discharge Abstract Database with the Data Liberation Initiative community.

CCRI
> See Canadian Century Research Infrastructure.

CHASS
> See Computing in the Humanities and Social Sciences.

CIHI
> See Canadian Institute for Health Information.

Codebooks
> A generic term often used to describe the user guide, record layout and data dictionary or combinations of these documents.

Command files
> Defines the microdata file by providing the name of the dataset; the variable locations (column locations and widths), names and labels; and value labels and missing value locations.

Computing in the Humanities and Social Sciences
> This is a computing facility within the Faculty of Arts and Science at the University of Toronto.

Continuum of access
> A dissemination model used by Statistics Canada that helps users identify the various levels of data access at Statistics Canada, of which the Data Liberation Initiative is one avenue.

Cross tabulation
> Cross tabulation groups variables to understand the correlation between different variables. Also known as "cross tab".

Customized tabulations
> Tables from confidential files that are specially processed by Statistics Canada for a fee and are also known as "custom tabs."

CV tables

Coefficients of variation tables are used to assess the quality of the data.

D

DAD

See Discharge Abstract Database.

Data

A collection of data objects stored together in electronic form, according to one schema, and accessible by computer.

Data Access Division Newsletter

The purpose of this newsletter is to inform subscribers and users of ongoing divisional initiatives.

Data dictionaries

Data dictionaries are used to find general information on the variables in a survey, the codes for the variables, missing value assignments and frequency counts.

Data Interest Group for Reference Services

Hosted at the University of Alberta, Data Interest Group for Reference Services content is based primarily on questions and answers from the Data Liberation Initiative listserv from 2004 to present.

Data Liberation Initiative

The Data Liberation Initiative is a partnership between Statistics Canada and Canadian postsecondary institutions with the objective of promoting and facilitating the availability of Statistics Canada data and other Canadian data for educational and academic research purposes.

DIGRS

See Data Interest Group for Reference Services.

Discharge Abstract Database

The Discharge Abstract Database consists of research analytic files on patient discharges from hospital, including codes for treatments and diagnostics from the Canadian Institute for Health Information.

DLI

See Data Liberation Initiative.

DLI alternate

An alternate is identified by the Data Liberation Initiative (DLI) contact to assist them in their DLI role at the institution.

DLI contact

The primary point of contact between the end users and the Data Liberation Initiative program.

DLI licence

The formal agreement between Statistics Canada (Data Liberation Initiative program) and participating postsecondary institutions.

DLI unit

The Statistics Canada unit responsible for administering the Data Liberation Initiative partnership.

dlilist

The Data Liberation Initiative (DLI) listserv is the primary communication tool for the DLI community.

Dummy files

Dummy files are created by author divisions by reproducing the master file and distorting the data. Dummy files do not contain actual data and can never be used for analysis.

E

EAC

See External Advisory Committee.

EFT

See Electronic File Transfer.

Electronic File Transfer

A Statistics Canada service, where Data Liberation Initiative data can be accessed via a file transfer protocol.

External Advisory Committee

The committee's primary role is to advise the Data Liberation Initiative (DLI) on matters pertaining to the planning, promotion and implementation of DLI initiatives.

F

FTP

A file transfer protocol is a way to download files.

M

Master files

Master files are so-called "pure" datasets created by the author division. All variables and cases are available for analysis in the master file. The master file is not available to all users.

Member institutions

Member institutions are postsecondary institutions that subscribe to the Data Liberation Initiative.

Metadata

Documentation that accompanies and assists users in the interpretation of various data and geographic files.

Microdata

The data observed or collected directly from a specific unit of observation.

Microdata Release Committee

Statistics Canada's confidentiality control mechanism for the release of microdata files.

N

National training sessions

> National training sessions are held every four years and allow the Data Liberation Initiative (DLI) community to come together and learn about data services. They also assist postsecondary institution users with accessing and using DLI data.

O

ODESI

> See Ontario Data Documentation, Extraction Service and Infrastructure Initiative.

Ontario Data Documentation, Extraction Service and Infrastructure Initiative

> ODESI is a digital repository for social science data. It includes Data Liberation Initiative data, as well as a range of public opinion polls created by the Ontario Council of University Libraries. It is available to authorized users from Ontario universities, but the metadata are openly available anyone.

Open Licence Agreement

> Statistics Canada's Open Licence Agreement applies to standard and custom data products in the Data Liberation Initiative collection, but not to postal code files, the Discharge Abstract Database files or the Social Policy Simulation Database and Model.

P

PCCF

> See Postal Code Conversion File.

PDC

> See Professional Development Committee.

Postal Code Conversion File

> A digital file that provides a correspondence between the six-character postal code and geographical areas for which census data and other statistics are produced.

Professional Development Committee

> The Data Liberation Initiative Professional Development Committee reports to the External Advisory Committee and is tasked with planning and developing a training curriculum for regional workshops.

Public use microdata file

> A master file that has been modified to minimize the possibility of disclosing a respondent's identity.

PUMF

> See Public use microdata file.

Purchase confirmation

> This is the purchase confirmation provided by Statistics Canada and accepted by the client to confirm their order when the total amount of the transaction is less than CAN$20,000.

Q

Questionnaire

A questionnaire is a list of questions posed to the survey respondent.

R

RDC
> See Research data centre.

readme file
> A quick reference guide on the Data Liberation Initiative electronic file transfer site that provides a breakdown of the contents of a folder.

Real Time Remote Access
> An online remote access facility that allows subscribers to run SAS programs in real time against microdata sets located in a central and secure location.

Record layout
> A record layout provides the names of the variables and their column position in the microdata file.

Regional Training Coordinator
> A person selected to represent their region in identifying training priorities and carrying out Data Liberation Initiative training.

Regional training session
> This training is held in four regions across Canada, the purpose of which is to prepare data services staff to assist postsecondary institution users with accessing and using Data Liberation Initiative data.

Research data centre
> Research data centres provide researchers with access—in a secure university setting— to microdata from Statistics Canada's population and household surveys.

RTC
> See Regional Training Coordinator.

RTRA
> See Real Time Remote Access.

S

SAS
> The Statistical Analysis System is a statistical analysis software program.

Scholar's Portal Dataverse
> The Scholars Portal Dataverse is provided by Scholars Portal on behalf of the Ontario Council of University Libraries and other participating institutions.

SDA
> Survey Documentation and Analysis is a set of programs used for the documentation and web-based analysis of survey data.

Social Policy Simulation Database and Model
> A static microsimulation model available in the Data Liberation Initiative collection and subject to a licensing agreement.

SPSD/M

See Social Policy Simulation Database and Model.

SPSS

Statistical Package for the Social Sciences is a statistical analysis software program.

STATA

STATA is a statistical analysis software program.

Synthetic files

These files are created by author divisions by reproducing the master file and distorting the data. Synthetic files do not contain actual data and can never be used for analysis.

T

Training Repository

The Data Liberation Initiative (DLI) Training Repository contains training sessions and workshop presentations from the DLI and other national and international conferences over the years.

U

User guide

The user guide provides essential information to inform the use, analysis and interpretation of the microdata file.