
Postal Code^{OM} Conversion File Plus (PCCF+) Version 8A, Reference Guide

December 2022 Postal codes^{OM}



Statistics
Canada

Statistique
Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

email at infostats@statcan.gc.ca

telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- | | |
|---|----------------|
| • Statistical Information Service | 1-800-263-1136 |
| • National telecommunications device for the hearing impaired | 1-800-363-7629 |
| • Fax line | 1-514-283-9350 |

Depository Services Program

- | | |
|------------------|----------------|
| • Inquiries line | 1-800-635-7943 |
| • Fax line | 1-800-565-7757 |

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on

www.statcan.gc.ca

under "Contact us" > [Standards of service to the public \(statcan.gc.ca\)](#)

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Table of Contents

	Page
What's new?	5
1. About this guide	6
2. Overview	7
3. About this tool	8
Purpose of PCCF+	8
Definitions and concepts	8
Contents	8
When to use PCCF+	9
Which routine to use: residential or institutional?	11
Methodology	11
Limitations	13
Comparison to other products/versions	14
Use with other products	14
Reference dates	15
4. Technical specifications	16
System requirements and installation	16
Variable descriptions (in order of output file)	20
File naming convention	36
5. Data quality	37
Lineage	37
Positional accuracy	37
Attribute accuracy	37
Logical consistency	37
Consistency with other products	38
Completeness	38
Appendix A: Glossary	39
Appendix B: Record Layouts for PCCF+ input files	43
Appendix C: Supplementary data file included with PCCF+	57
Appendix D: Hierarchy of standard geographic units for dissemination, 2021 Census	58
Appendix E: Derivation of neighbourhood income quantiles	59
Appendix F: Definition of "Usual place of residence"	59
Appendix G: Residential Routine Flowchart for PCCF+	61
Appendix H: Institutional Routine Flowchart for PCCF+	62
Appendix I: Additional notes on Delivery Mode Type (DMT)	63

Introduction

The Postal Code Project is responsible for linking the approximately 900,000 single postal codes in Canada to Statistics Canada's Census dissemination geography (presently 2021 Census geographic areas). This process is performed by using data provided by Canada Post Corporation and linking to Census Dissemination geography via the process of geocoding. The result is the creation of the Postal Code Conversion File (PCCF) which provides a correspondence between the six character postal code and Statistics Canada's standard geographical areas. The Postal Code Conversion File Plus (PCCF+) is a companion tool of the PCCF.

Roles and Responsibilities

This reference guide is intended for users of the PCCF+. The guide provides an overview of the file, the general methodology used in its creation and important technical information. On the Statistics Canada Web site, it is reported that the product was discontinued as an official release product. Therefore, officially, the reference guides are no longer distributed by Statistics Canada but continue to be distributed to our partners. The Canada Post Corporation (CPC) distributes to other individuals and organizations.

Licenses

The PCCF+ data product is available for license by Statistics Canada and CPC.

Both organizations agree that Statistics Canada manages licensing with Focal Points, Federal Government Agencies, Crown Corporations, and Cancer Registry end-users, while CPC manages licenses for value-added distributors/commercial, provincial or municipal government and non-government users.

Statistics Canada will process licensing for and distributes the Postal Code Data Products to the share partners permitted under the new revised licensing governance between Statistics Canada and CPC.

For internal users, to contact statistics Canada about the distribution of PCCF+ please use the following email address: statcan.haddas-dashad.statcan@statcan.gc.ca.

For external users, based on the agreement reached between Statistics Canada and CPC, you have to obtain the Postal Code End-use license agreement to continue to use and receive the Postal Code Data Products files directly from CPC.

For external users, the distribution of PCCF+ please use the following email address: data.solutionscentre@canadapost.ca

Postal Code^{OM} Conversion File Plus, Reference Guide

This reference guide is intended for users of the Postal Code^{OM} Conversion File Plus (PCCF+). The guide provides an overview of the file, the general methodology used in its creation and important technical information.

What's new?

The postal code^{OM} reference date for the Postal Code^{OM} Conversion File (PCCF) and the Postal Code^{OM} Conversion File Plus (PCCF+) is December 2022.

The postal codes^{OM} are linked to the geographic areas used in the 2021 Census of Population. The reference date for these geographic areas is January 1, 2021. The previous version (7E) linked to the 2016 Census of Population geographic areas.

The PCCF+ Version 8A includes a population-weighting file calculated from the 2021 Census population counts

Neighbourhood income quintiles and deciles have been calculated from 2021 Census population data.

The routine that allowed geocoding of historical postal codes ^{OM} in British Columbia (V1H, V9G, prior 1998) has been removed.

^{OM} Postal Code is an official mark of Canada Post Corporation (CPC).

1. About this guide

This reference guide is intended for users of the Postal Code^{OM} Conversion File Plus (PCCF+) Version 8A, a companion tool to the Postal Code^{OM} Conversion File (PCCF). The guide provides an overview of the files and software, the general methodology used in the creation, important technical information, and instruction on how and when to use the included SAS[®] programs.¹

This data and software tool is provided 'as-is,' and Statistics Canada makes no warranty, either express or implied, including but not limited to, warranties of merchantability and fitness for a particular purpose. In no event will Statistics Canada be liable for any direct, special, indirect, consequential or other damages, however caused.

¹ © Copyright 2002-2008 by SAS Institute Inc., Cary, NC, USA.

2. Overview

The PCCF+ is a SAS[®] control program and set of associated datasets derived from the PCCF, a 2021 postal code^{OM} population weight file, the Geographic Attribute File, Health Region boundary files, and other supplementary data. PCCF+ automatically assigns a range of Statistics Canada standard geographic areas and other geographic identifiers based on postal codes^{OM}.

PCCF+ differs from the PCCF in that it:

- Uses population-weighted random allocation for many postal codes^{OM} that link to more than one geographic area.
- Options are available for institutional postal codes^{OM}.
- Procedures are included to link partial postal codes^{OM} to geographic identifiers where possible.
- Problem records and diagnostics are provided in the program output, along with reference information for possible solutions.

The geographic coordinates, which represent the standard geostatistical areas linked to each postal code^{OM} on the PCCF, are commonly used to map the distribution of data for spatial analysis (e.g., clients, activities).

In April 1983, the Geography Division released the first version of the PCCF, which linked postal codes to 1981 census geographic areas and included geographic coordinates. PCCF+ was first created using the 1986 census and has been updated regularly with population weight files calculated for each census from 1991 through 2021.

3. About this tool

Purpose of PCCF+

The purpose of the PCCF+ is to provide a link between six-character postal codes^{OM} produced by the Canada Post Corporation (CPC), standard 2021 Census geographic areas (such as dissemination areas, census subdivisions, and census tracts) produced by Statistics Canada, and supplementary administrative areas and neighbourhood income quintiles.

Postal codes^{OM} do not respect census geographic boundaries and so may be linked to more than one standard geographic area, or assigned to more than one set of coordinates. Therefore, one postal code^{OM} may be represented by more than one record. The PCCF product, produced by Statistics Canada, provides links between postal codes and all recorded matches to census geography.

PCCF+ uses the PCCF but provides additional functionality in that it uses a population-weighted matching process for some residential postal codes^{OM} where more than one geographic code is possible. PCCF+ also provides routines for institutional postal codes^{OM} and for historic postal codes^{OM}.

The purpose of this product is *not* to validate postal codes^{OM} or the PCCF.

Definitions and concepts

Selected geographic terms and concepts are briefly defined in the glossary (*Appendix A*).

More detail can be found by referring to the definitions of the Geography universe from the *Dictionary, Census of Population, 2021*. <https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/index-eng.cfm>.

Additional reference guides include the following:

- *Postal Code^{OM} Conversion File (PCCF) Reference Guide*
- *Geographic Attribute File, Reference Guide, Census year 2021* ([Geographic Attribute File \(statcan.gc.ca\)](https://www12.statcan.gc.ca/geography/reference/geographic_attribute/index-eng.cfm))
- *Correspondence Files, 2021 Census* ([Correspondence Files, Reference Guide, Census year 2021 \(statcan.gc.ca\)](https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/index-eng.cfm))

Contents

This version of PCCF+ contains a SAS® control program, and program and data files for operation of the geocoding process. Table 3.1 provides a list of these files with descriptions. Several supplementary SAS® programs and datasets are also provided, and are listed in Appendix C, though these are not necessary for the main functionality of the program.

Table 3.1: List of main files included with PCCF+

Folder name	File name	Description
(main)	PCCFplus_8A.sas (FCCPplus_8A.sas)	Primary PCCF+ SAS® control program for geocoding for residential and institutional (French version)
SAS_InputFiles	input_georef.sas	SAS® input file for geographic data files
SAS_InputFiles	input_pccf.sas	SAS® input file for PCCF data files
DATA	cpcref.airstage.v2212.txt	CPCost air stage delivery list (6+ months per year)
DATA	cpcref.bldgnam.v2212.txt	Building names and addresses (mostly non-residential)

Folder name	File name	Description
DATA	cpcref.emgres.v2212.txt	Flag for possible non-residential postal codes ^{OM} (DMT = E, G, M)
DATA	cpcref.instflg.v2212.txt	Flag for potential institutional postal codes ^{OM}
DATA	cpcref.nadr.v2212.txt	Number of address records for each postal code ^{OM}
DATA	Georef21.dablkpnt21.txt	Pointer file for dissemination area and dissemination block
DATA	Georef21.gaf21.txt	Subset of geographic attribute file
DATA	Georef21.hrdef.txt	Health Region definitions, plus additional variables for PCCF+
DATA	Georef21.ses21.txt	Neighbourhood income quantiles
DATA	pccf2212.pccf.dups.txt	Duplicate postal codes ^{OM} from PCCF
DATA	pccf2212.pccf.pointdup.txt	Pointer file for duplicate postal codes ^{OM}
DATA	pccf2212.pccf.rpo.txt	Rural post office codes
DATA	pccf2212.pccf.uniq.txt	Unique postal codes ^{OM} from PCCF
DATA	pccf2212.wc2dups.txt	Weighting for first 2 characters of postal code ^{OM}
DATA	pccf2212.wc2point.txt	Pointer for 2-character weighting file
DATA	pccf2212.wc3dups.txt	Weighting for first 3 characters of Postal Code ^{OM} (FSA [®])
DATA	pccf2212.wc3point.txt	Pointer for 3-character weighting file (FSA [®])
DATA	pccf2212.wc4dups.txt	Weighting for first 4 characters of postal code ^{OM}
DATA	pccf2212.wc4point.txt	Pointer for 4-character weighting file
DATA	pccf2212.wc5dups.txt	Weighting for first 5 characters of postal code ^{OM}
DATA	pccf2212.wc5point.txt	Pointer for 5-character weighting file
DATA	pccf2212.wc6dups.txt	Weighting for postal code ^{OM}
DATA	pccf2212.wc6point.txt	Pointer for postal code ^{OM} weighting file
DATA_SAS		A pre-read SAS dataset of all the above text file in the DATA folder
DA_Nat_weight_file_2021		2021 Census Postal Code – Dissemination Area (population) Weight File (complete file)

When to use PCCF+

The PCCF+ is updated on an annual basis. The regular maintenance includes taking all postal code^{OM} changes continually introduced by CPC as updated in the PCCF, as well as any updates to health regions or other administrative files. Every five years, after each census, PCCF+ and the underlying population weight files are recalculated and aligned with the new vintage of census geographic areas.

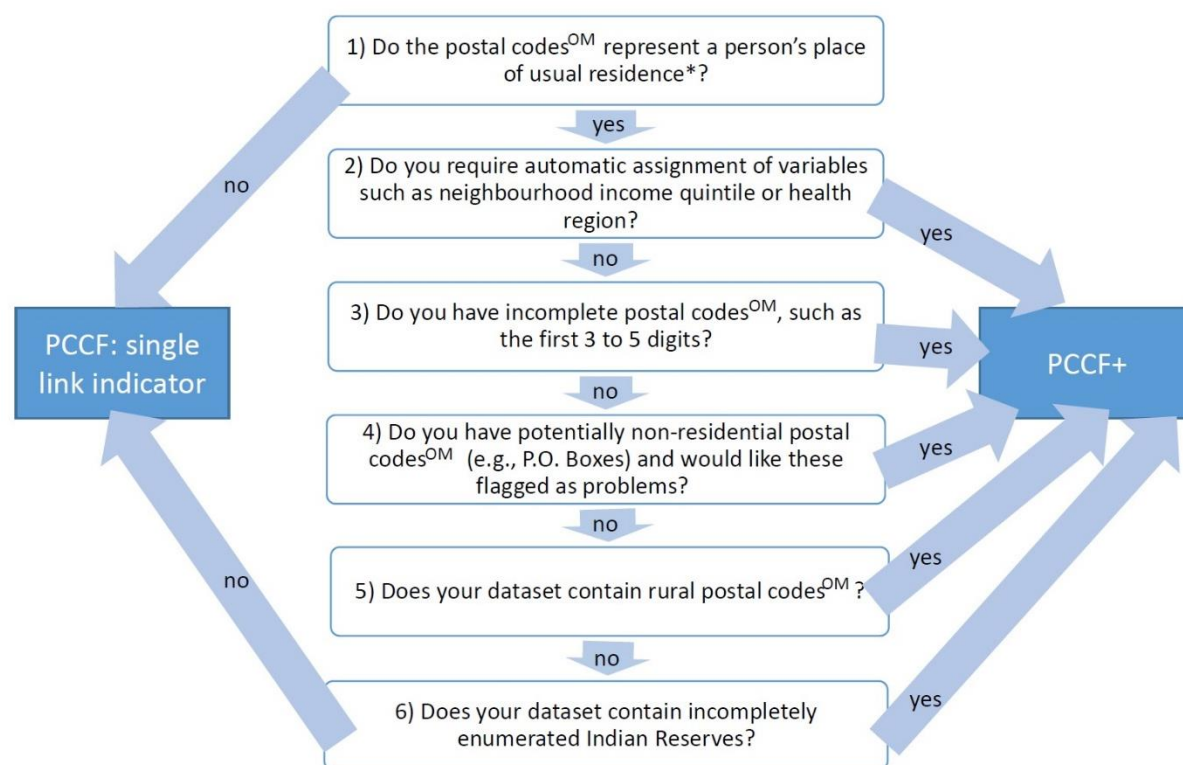
PCCF+ consists of a SAS[®] control program and a series of reference files derived from the PCCF, the geographic attribute file, a custom postal code^{OM} population weight file, health region boundaries, and other sources. PCCF+ automatically assigns a full range of geographic identifiers based on postal codes^{OM}. PCCF+ has been developed over a number of years for research studies at Statistics Canada with considerable input from the broader research community. With the help of user feedback, any

incorrect coding due to errors in the underlying reference files or programming can be identified and corrected.

The primary capabilities of PCCF+ are:

- Read and troubleshoot community mail boxes and other sources of duplicate records on the PCCF (Delivery Mode Type, or DMT A,B).
- Identify postal codes^{OM} which may be used by businesses or institutions rather than residential locations (DMT E,G,M).
- Provide geographically unbiased coding for small PO boxes at urban post offices (DMT K) and for General Delivery at urban post offices (DMT J).
- Provide better geographic coding for institutions using large PO boxes (DMT M).
- Incorporate retired postal codes^{OM}, taking into account problems related to previous DMT.
- Provide translation across different vintages of census geography, which is useful in geocoding datasets that span backwards in time (such as cohort studies).
- Use the first one, two, three, four, or five characters of the postal code^{OM} to impute or partially impute census geographic coding.
- Provide additional information to correct erroneous or problematic postal codes^{OM} and find geographic codes by other means.
- For postal codes^{OM} which may or may not refer to a place of business (DMT E,G,M), flag records known to serve non-residential addresses, and flag those known to serve residential addresses.
- For areas consisting primarily of collective dwellings, indicate the predominant dwelling type (hospital, nursing home, prison, etc.).
- Provide detailed diagnostics indicating how the coding was done, what problems were encountered, and how ambiguous or precise the coding was.

For research purposes, the PCCF and PCCF+ products provide different capabilities. The following flow chart provides some guidance to users to determine product suitability for their geocoding purposes.



*Place of usual residence: Definition in Appendix F

Flow chart. When to use PCCF or the PCCF+ product to geocode postal codes^{OM}.

Which routine to use: residential or institutional?

There are two alternative geocoding routines within the PCCF+ program, the residential routine and the institutional routine.

The PCCF+ residential routine (default) is to geocode records corresponding to persons at their *usual place of residence*. Canadians respond to census questionnaires at their usual place of residence and these postal codes are used to create the population weights used in the PCCF+ assignment. The census definition of the usual place of residence is provided in Appendix F. Note that some postal codes, such as university residences, may not be coded properly by the PCCF+ since census respondents are not enumerated at these locations. In situations where the postal code might not correspond to a usual place of residence, users should consider using the PCCF product instead.

Alternatively, the optional institutional coding routine can be used to geocode postal codes^{OM} corresponding to health facilities (long-term care facilities, nursing homes, hospitals, residential care homes, offices of health professionals) or other institutions.

Methodology

The general methodology for residential coding is outlined below, with procedures that apply only to institutional geocoding shown in *italics*. Please refer to Appendix G (residential) and H (institutional) for flowcharts depicting these processes.

1. Rural postal codes^{OM} and postal codes^{OM} served by rural route delivery or suburban services from urban post offices, or which indicate a group of post office boxes or a single post office box, are matched to a subset of the weighted conversion file. As most such codes serve more than one dissemination area (DA), the geographic codes are assigned randomly in proportion to the distribution of population with that postal code^{OM}, as seen in the weighting file. *For coding of office locations, etc., the institutional routine omits the rural postal codes^{OM} from this step, so that they can all be assigned to the same dissemination area as the rural post office.*
2. Remaining postal codes^{OM} which are unique on the PCCF (only linked to a single dissemination area, dissemination block, or block face) are matched to corresponding codes on the incoming file. *For institutional geocoding, rural postal codes^{OM} together with their corresponding post office geography are added at this point, as those records are also unique.*
3. Postal codes^{OM} which are not unique on the PCCF are matched to the remaining records from the input file. Most urban postal codes^{OM} and some rural postal codes^{OM} which are not unique on the PCCF (in the sense that they link to more than one dissemination area, dissemination block, or block-face) are nonetheless not ambiguous in terms of higher levels of geography such as census division, census subdivision, census metropolitan area, or census tract. To avoid "many-to-many" matching, the matching in this part of the program is done in two steps:
 - a. Each remaining input record (not already matched to the weighting file or to the PCCF unique file) is matched by postal code^{OM} to a pointer file which contains a single record for each postal code^{OM} which occurs more than once on the PCCF. The pointer file shows how many times the postal code^{OM} occurs and the observation number of the first occurrence of that postal code^{OM} on the pointer file.
 - b. The information on the pointer file is used to match successive records with the next occurrence of that postal code^{OM} on the pointer file. This has the effect of distributing events for such postal codes^{OM} across all possible dissemination areas, dissemination blocks, or block-faces which are served by that postal code^{OM} – with equal weight assigned to each PCCF record.
4. Missing dissemination block codes are assigned based on population-weighted imputation from the dissemination area code, if available.
5. Error records are then identified and processed as follows:
 - a. Any record with a postal code^{OM} which did not match on all 6 characters to the PCCF is identified as an error record (Link=0).
 - b. Records with postal codes^{OM} which partially matched to the PCCF or weighting file (i.e., DA='9999'), and whose DMT is H, J, K, M, T or X are also identified as error records (Link=1), since the PCCF is ambiguous with regards to residential location (e.g., post office location, rural route).
 - c. Using auxiliary files, an attempt is then made to assign highly probable census metropolitan area, census division, and census subdivision codes, plus census tract and dissemination area for urban postal codes^{OM}. Coding will be suggested based on the first 5, 4 or 3 characters of the postal code^{OM}, or failing that, based on the first 2 characters of the postal code^{OM}. The province may also be assigned based on the first character of the postal code^{OM}.
6. All records with their corresponding geography (to the extent found) are output to the specified output file. If some or all geographic codes could not be determined, those fields are set to missing values in the output file.
7. A problem output file is created, containing:
 - a. Records with postal codes^{OM} which could not be matched on all 6 characters (Link0: error);

- b. Records with postal codes^{OM} for a DMT which is only linked to post office location on the PCCF (Link 1: error), and for which census location data were not available on the weighted conversion file;
 - c. Records where the DMT frequently indicates a non-residential address (Link 3 and 4: warning);
 - d. Records for postal codes^{OM} known to indicate a non-residential address (Link 2: warning);
 - e. Records which could have been assigned more than one census subdivision based on the unweighted PCCF (Link 6: note); and,
 - f. Records which could have been assigned to more than one census subdivision based on the weighted conversion file (Link 7: note).
8. A summary of the geocoding process, including the number of records in each link type above is printed in the PDF output, together with suggestions as to what to do in each case. The summary also shows the distribution of records by the number of geographic codes which were assigned.
9. Frequency counts of the occurrence of each value of the main fields are output. This is done for both the geocoded output and for the problem subset.

Limitations

The PCCF contains multiple records for a postal code^{OM} when the postal code^{OM} straddles more than one blockface or dissemination block. It should be acknowledged that the CPC source data used to create the PCCF contains some postal codes^{OM} which have links to multiple address ranges.

Civic addresses are not available for some postal codes^{OM} such as those associated with rural routes. Many of these postal codes^{OM} tend to straddle several dissemination areas and often cross boundaries of standard geographic areas such as census tracts or census subdivisions. It is difficult, if not impossible, to identify the precise physical location of a rural postal code^{OM}.

Community mailboxes are also a source for multiple records per postal code^{OM} on the PCCF. In newer urban delivery areas, postal codes^{OM} are assigned to a community mailbox that may cover partial dissemination blocks, both sides of a street, and different streets within 200 metres of the community mailbox. These situations often result in multiple links being established between a postal code^{OM} and blockfaces, unlike the more traditional urban postal codes^{OM}, which correspond generally to a single blockface.

The single link indicator (SLI) was created to assist users in dealing with postal codes^{OM} having multiple records. The method used to establish the single link indicator identifies the geographic area with the majority of population assigned to a particular geographic area. Users should be aware that only a partial correspondence between the postal code^{OM} and other geographic areas is achieved when using the single link indicator. It should also be noted that the single link indicator is identified on both active and retired postal codes^{OM}. Users will find when working with both active and retired postal codes^{OM}, multiple SLIs will appear for a postal code^{OM} that has been retired and reintroduced.

The address associated with a postal code^{OM} does not always represent the location where those receiving mail using that postal code^{OM} actually reside. This is particularly the case in rural areas, where rural route service and post office pick-up are commonly used to deliver mail. The delivery mode type of 'W' (rural) and 'H' (rural route) on the PCCF identify postal codes^{OM} that are usually considered rural.

A typical rural route address, such as 'RR#6, Georgeville, QC', does not provide sufficient address information to identify a precise physical location. A rural post office address such as 'PO BOX 4001 STN A VICTORIA BC' is also imprecise and not explicitly attached to the dwellings served by that postal code^{OM}. Consequently, rural postal codes^{OM} cannot be used in the same manner as most urban postal codes^{OM} can to precisely geo-reference a physical location.

Similarly, postal codes^{OM} with a delivery mode type of 'K' (group of post office boxes) or 'M' (one post office box) may be linked to the location of the postal installation on the PCCF, as opposed to the physical location of customers who rent a post office box.

Among postal codes^{OM} with matches to multiple records, those with DMT = A, B, E, G or M (i.e., urban) with links to 4 or more dissemination areas were deemed problematic since they represent a higher degree of inaccurate geocoding. As a result, these postal codes^{OM} were removed from the PCCF duplicate records file. Instead, the PCCF+ processes them using the weighted conversion (WCF) files if the postal code^{OM} or partial postal code^{OM} was reported on the 2016 Census of Population, or using partial postal codes (e.g., first 5 digits) if a match to the weighted conversion file does not occur.

For geographic coding from postal codes^{OM}, any geographic coding which requires dissemination block level precision (PopCtrRAuid, PopCtrRAtype, PopCtrRAclass) is meaningless outside of urban cores (where block level precision is usually possible based on urban postal codes^{OM}). For most rural postal codes^{OM}, a portion of the area served will meet the census criteria for urban, however, most of the surrounding area would be rural; but rural postal codes^{OM} are inadequate to distinguish one from the other since the same postal code^{OM} serves both urban and rural areas, as defined by the census.

The health region correspondence to census geography, which is coded as part of PCCF+, is based on 2021 and 2016 census geographic units. The smallest geographic unit required has been used as the building block to define health regions. In general, the legislated limits respect these units, but not all respect dissemination areas or dissemination blocks once the legislated boundaries are digitized. In most provinces, the dissemination area was used to define health regions. However, in several instances, the actual physical legal limits split dissemination areas. In these cases, the dissemination block was used to improve the accuracy of these boundaries. The limits that did not respect Statistics Canada geometry (the splits) were digitized by utilizing maps, spatial layers and/or descriptions supplied by and with the cooperation of the authority for each province.

Comparison to other products/versions

Version 8 of PCCF+ is a major revision of the source data files and user input / output of the SAS[®] program. Version 8A uses 2021 census geography, with the 2021 Census population weight file created using rounded 100% population counts. PCCF+ source data files use postal codes^{OM} and associated data from CPC through to the end of December 2022.

- Version 1 – 1986 census; equal weight to duplicate records
- Version 2 – 1991 census; 2B (20% sample) household weights for duplicate records (DMT H-Z)
- Version 3 – 1996 census; 2A (100% count) population weights for duplicate records (DMT H-Z)
- Version 4 – 2001 census; 2A (100% count) population weights for duplicate records (DMT H-Z)
- Version 5 – 2006 census; 2A (100% count) population weights for duplicate records (DMT H-Z)
- Version 6 – 2011 census; 2A (100% count) population weights for duplicate records (DMT H-Z)
- Version 7 – 2016 census; 2A (100% count) population weights for duplicate records (DMT H-Z)
- Version 8 – 2021 census; 2A (100% count) population weights for duplicate records (DMT H-Z)

Use with other products

Output from PCCF+ can be linked to the 2021 census geographic attribute file and other 2021 Census products using the geographic unique identifiers. For completely geocoded records with a 2021 dissemination block identifier, unique geographic identifiers are provided for other census periods (1981

through 2016) and thus can be linked to other census products from those periods at the dissemination/enumeration area level.

Successfully geocoded health region unique identifiers can be linked to 2022 health region boundaries. Using the 2021 or 2016 dissemination block identifiers, geocoded records can be linked to health region boundaries from other periods as well.

Reference dates

The reference date for postal codes^{OM} contained in this product is December 2022.

The geographic reference date is a date determined by Statistics Canada to finalize the geographic framework for which the census data are collected, tabulated and reported. The geographic reference date for the 2021 Census is 1 January 2021. See the *Geographic Attribute File, Reference Guide, 2021 Census*. Statistics Canada Catalogue no. 92-151-G.

The reference date for health region boundaries is December 2022, with correspondence to 2021 and 2016 Census geography. See *Health Regions: Boundaries and Correspondence with Census Geography*. Statistics Canada Catalogue no. 82-402-X.

4. Technical specifications

System requirements and installation

Use of PCCF+ requires a Microsoft Windows® based system capable of running SAS® version 9 or higher, or SAS® Enterprise Guide version 4 or higher. Programs were developed using SAS® Version 9.4 for Windows 7.

To install PCCF+ and perform automated geocoding based on postal codes^{OM} using PCCF+, instructions are provided below.

1. Setting up PCCF+

The main component of PCCF+ consists of a single SAS® control file (PCCFplus_8A.sas) as well as reference files primarily derived from the PCCF, the geographic attribute file, and the weighted conversion file. SAS® input programs are also included as part of the PCCF+ installation folders. Data files are in ASCII text format and do not include any software nor instructions on how to use the product beyond the provision of record layouts in *Appendix B*.

To use PCCF+ all files and folders in must be copied to a directory accessible to SAS®. An example is provided in Figure 4.1.

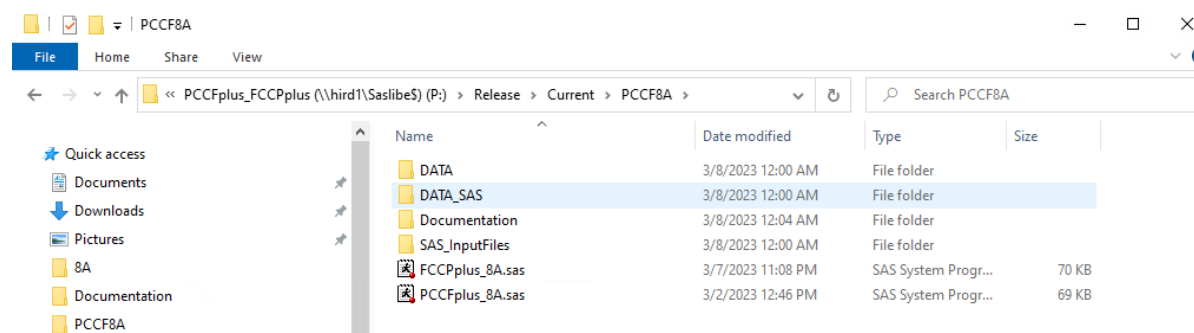


Figure 4.1: Example installation folder for PCCF+ Version 8A.

2. Preparation of the input file

Input data must be a SAS® (or text) data file, sorted or unsorted, with each logical record containing a unique identifier (ID) and a postal code^{OM} (PCODE) if available. The postal code^{OM} must be formatted to be 6 characters with no spaces or hyphens. Table 4.1 shows an example layout for the input file.

The ID should be in character format, up to 15 characters in length. Records with the same ID but different PCODEs will each be assigned geographic codes (that may differ). For records with the same ID and PCODE combination, only one example of each will be retained.

Table 4.1: Example record layout for a PCCF+ SAS input file.

Position	Size	Type	Field Name	Description
1	15	C	ID	User-supplied unique identifier
16	6	C	PCODE	User-supplied postal code ^{OM}

3. Specifying macro variables in the PCCF+ SAS program

The PCCF+ program has 7 macro variables that need to be specified before running. These variables are described in detail at the top of the SAS® program, but are also outlined here.

- Installation directory (`installDir`) – The directory where the PCCF+ program and source files are located. PCCF+ can be stored on any local or network directory; however, the directory name must **not** contain any spaces but can contain hyphens or underscores.
Example: `%let installDir = C:\PCCFplus_FCCPplus\;`
- Input data library (`inData`) – The folder path where the input dataset is located. This can be any directory accessible to SAS®.
Example: `libname inData = "C:\PCCFplus_FCCPplus\input\";`
- Input data file (`inFile`) – The name of the input SAS® data file to be coded. This will need to be produced according to the specifications above. Note that filename cannot begin with "GEO".
Example: `%let inFile = inputDataSet;`
- Output file name (`outName`) – The name of the output SAS® file (layout provided above). This is the geocoded file and includes *all* records, including those where there was a problem in geocoding. PCCF+ will automatically create an additional SAS® file for problem records, which has the same name as the geocoded output with the additional of a `_problem` suffix.
Example: `%let outName = outputDataSet;`
- PDF listing output (`pdfOutput`) – The path and filename for the coding output summary. This summarises the geocoding process and identifies any potential problems and errors in geocoding.
Example: `%let pdfOutput = "C:\PCCFplus_FCCPplus\output\output01.pdf";`
- Code version (`codeVersion`) – Whether to run residential (`codeVersion=0`) or institutional (`codeVersion=1`) geocoding. Residential coding is the default version, but if the data contains postal codes^{OM} of businesses or institutions (rather than individuals), then an institutional coding option is provided.

4. Running PCCF+

Once the macro variables listed above have been specified, PCCF+ can be run using the SAS® submit command. When submitting for the first time, it is recommended to closely watch the log window for any errors that appear immediately. In this case, the program can be halted and any errors corrected.

Once the geocoding process has completed, the first step is to examine the log for any additional errors or warnings. These are usually indicated in red or green respectively and should be addressed before examining the results.

5. Interpreting PCCF+ Results

PCCF+ produces two output datasets, one for all the geocoded data (including problem records), and a subset that contains only the problem records (errors, warnings, and notes). The first of these output files, the main output file, contains the ID, postal code^{OM}, the geographic codes that were successfully determined, and additional diagnostic codes that can aid in understanding how the coding proceeded for each case. Table 4.2 provides the record layout for the main SAS® output file.

Table 4.2: PCCF+ main output description and sources

Field Name	Source	Description
Coder	User selection	Residential or institutional run of PCCF+
Version	Fixed	Version of PCCF+,
ID	User-supplied	User-supplied unique identifier
PCODE	User-supplied postal code ^{OM}	Postal code ^{OM}
PR	Geographic attribute file	Province
DAuid	Geographic attribute file	Dissemination area unique identifier
DB	Geographic attribute file	Dissemination block code
DB_ir2021	Geographic attribute file	2021 census Indian reserve refusal flag
CSDuid	Geographic attribute file	Census subdivision unique identifier

Field Name	Source	Description
CSDname	Geographic attribute file	Census subdivision name
CSDtyp	Geographic attribute file	Census subdivision type
CMA	Geographic attribute file	Census metropolitan area / census agglomeration unique identifier
CMAtype	Geographic attribute file	Census metropolitan area type
CMAname	Geographic attribute file	Census metropolitan area / census agglomeration name
CTname	Geographic attribute file	Census tract name
Tracted	Geographic attribute file	Flag for tracted (in census tract) / non-tracted (not in census tract) area
SACcode	Geographic attribute file	Statistical area classification code
SACtype	Geographic attribute file	Statistical area classification type
CCSuid	Geographic attribute file	Census consolidated subdivision code
FEDuid	Geographic attribute file	Federal electoral district – 2013 Representation Order unique identifier
FEDname	Geographic attribute file	Federal electoral district name – 2013 Representation Order unique identifier
DPLuid	Geographic attribute file	Designated place identifier
DPLtype	Geographic attribute file	Designated place type
DPLname	Geographic attribute file	Designated place name
ERuid	Geographic attribute file	Economic region identifier
ERname	Geographic attribute file	Economic region name
CARuid	Census agricultural regions boundary file	Census agricultural region identifier
CARname	Census agricultural regions boundary file	Census agricultural region name
PopCtrRAPuid*	Geographic attribute file	Population centre/rural area code
PopCtrRAname*	Geographic attribute file	Population centre/rural area name
PopCtrRAtype*	Geographic attribute file	Population centre/rural area type
PopCtrRAclass*	Geographic attribute file	Population centre/rural area classification
CSize	PCCF+	Community size code (2021 CMA/CA population)
CSizeMIZ	PCCF+	Urban CMA/CA size and rural MIZ code
HRuid	Health Statistics Branch	Health region unique identifier
HRename	Health Statistics Branch	Health region name (English)
HRfname	Health Statistics Branch	Health region name (French)
AHRuid	Health Statistics Branch	Alternate health region unique identifier
AHRename	Health Statistics Branch	Alternate health region name (English)
AHRfname	Health Statistics Branch	Alternate health region name (French)
SLI	PCCF	Single link indicator
Rep_Pt_type	PCCF	Representative point type (PCCF)
RPF	PCCF/PCCF+	Representative point flag (PCCF+)
PCtype	PCCF	Postal code ^{OM} type
DMT	PCCF	Delivery mode type
H_DMT	PCCF	Historic delivery mode type
DMTDIFF	PCCF/PCCF+	Previous or alternate DMT (if applicable)
PO	PCCF	PCCF Delivery installation
QI	PCCF	PCCF Quality indicator
Source	PCCF	Source of geographic coding (PCCF)
Lat	PCCF	Latitude of lowest level geographic area
Long	PCCF	Longitude of lowest level geographic area
Link_Source	PCCF+	Source of geographic coding (PCCF+)
Link	PCCF+	Link type (nature of problem)

Field Name	Source	Description
Prec	PCCF+	Precision of representative point
nCD	Geographic attribute file	Number of census divisions
nCSD	Geographic attribute file	Number of census subdivisions
Comm_Name	CPC	CPC community name
AirLift	CPC	CPC air stage community
InstFlag	PCCF+	Institutional flag
ResFlag	PCCF+	Residence flag (where DMT=E,G,M)
InuitLands	PCCF+	Inuit Nunangat land claims settlement areas
BTIPPE	Health Analysis Division	Neighbourhood before tax income per single person equivalent (BTIPPE)
ATIPPE	Health Analysis Division	Neighbourhood after tax income per single person equivalent (ATIPPE)
QABTIPPE	Health Analysis Division	Neighbourhood income quintile before tax (in CMA/CA)
QNBTIPPE	Health Analysis Division	Neighbourhood income quintile before tax (national)
DABTIPPE	Health Analysis Division	Neighbourhood income decile before tax (in CMA/CA)
DNBTIPPE	Health Analysis Division	Neighbourhood income decile before tax (national)
QAATIPPE	Health Analysis Division	Neighbourhood income quintile after tax (in CMA/CA)
QNATIPPE	Health Analysis Division	Neighbourhood income quintile after tax (national)
DAATIPPE	Health Analysis Division	Neighbourhood income decile after tax (in CMA/CA)
DNATIPPE	Health Analysis Division	Neighbourhood income decile after tax (national)
IMPFLG	Health Analysis Division	Imputation flag for IPPE (mean household income imputed)
DA16uid	PCCF+	2016 dissemination area identifier
DB16uid	PCCF+	2016 dissemination block identifier
DA11uid	PCCF+	2011 dissemination area identifier
DB11uid	PCCF+	2011 dissemination block identifier
DA06uid	PCCF+	2006 dissemination area identifier
DA01uid	PCCF+	2001 dissemination area identifier
EA96uid	PCCF+	1996 enumeration area identifier
EA91uid	PCCF+	1991 enumeration area identifier
EA86uid	PCCF+	1986 enumeration area identifier
EA81uid	PCCF+	1981 enumeration area identifier

* See Limitations section above.

The second output file, indicated by the suffix _problem, contains a subset of the output records for any cases that had warnings, errors, or notes. To facilitate correction it is sorted by the type of problem (errors, followed by warnings, followed by notes), then by DMT, then by postal code^{OM}. In the event that none of the input records were identified as potential problems then the problem file would be empty. Table 4.3 provides the record layout for the problem SAS[®] output file.

Table 4.3: PCCF+ problem output description and sources

Field Name	Source	Description
Coder	User selection	Residential or institutional run of PCCF+
Version	Fixed	Version of PCCF+

Field Name	Source	Description
ID	User-supplied	User-supplied unique identifier
PCODE	User-supplied postal code ^{OM}	Postal code ^{OM}
Message	PCCF+	Text message indicating an error, warning, or note
DMT	PCCF	Delivery mode type
DMTDIFF	PCCF/PCCF+	Previous or alternate DMT (if applicable)
Link_Source	PCCF+	Source of geographic coding (PCCF+)
Link	PCCF+	Link type (nature of problem)
PR	Geographic attribute file	Province
DAuid	Geographic attribute file	Dissemination area unique identifier
DB	Geographic attribute file	Dissemination block code
CSDuid	Geographic attribute file	Census subdivision unique identifier
CSDtyp	Geographic attribute file	Census subdivision type
CMA	Geographic attribute file	Census metropolitan area / census agglomeration unique identifier
CTname	Geographic attribute file	Census tract name
InstFlag	PCCF+	Institutional flag
ResFlag	PCCF+	Residence flag (where DMT=E,G,M)
ADR	CPC	Building name and address, for business or institutional

Variable descriptions (in order of output file)

Coder

Indicates the user-selected routine of PCCF+. Residential routine is indicated by “R” and institutional routine indicated by “I”.

Version

Indicates version of the PCCF+ used.

ID

User-provided unique identification code.

Postal Code^{OM} (Pcode)

The postal code^{OM} is a six-character code defined and maintained by CPC for the purpose of sorting and delivering mail. The characters are arranged in the form 'ANA NAN', where 'A' represents an alphabetic character and 'N' represents a numeric character (such as K1A 0T6). The postal code^{OM} uses 18 alphabetic characters and 10 numeric characters. Six alphabetic characters (D, F, I, O, Q, and U) are not in use at the present time. In addition the first position does not make use of the letters W or Z.

The first three characters of the postal code^{OM} ('ANA') represent a set of defined and mostly stable areas known as Forward Sortation Areas (FSAs[®]). The FSA[®] represents a specific area within a major geographical region, a province or a territory. Rural postal codes^{OM} are identifiable by the presence of a zero (0) in the second position of the FSA[®] code. Urban postal codes^{OM} are composed of FSAs[®] with numerals 1 to 9 in the second position of the code. PCCF+ also identifies postal codes^{OM} with a DMT of H (rural route service) or T (suburban route service), which do not have a zero in the second position.

The last three characters of the postal code^{OM} ('NAN') identify routes known as local delivery units (LDUs[®]). In population centres, a single postal code^{OM} may correspond to the following types of LDUs[®]:

- a block-face (one side of a city street between consecutive intersections)
- a community mailbox (commonly called super mailboxes)
- an apartment building
- a business building
- a large firm or organization that does considerable business with CPC
- a federal government department, agency or branch
- a mail delivery route (rural, suburban or mobile)
- general delivery at a specific post office
- one or more post office boxes.

A community mailbox postal code^{OM} services both odd and even sides of the same street, or different streets, within a 200 metre radius of the community mailbox.

In rural FSAs[®], the LDU[®] generally refers to services which originate from a post office or postal station. These include rural routes, general deliveries, post office boxes, and suburban services. Often, in rural FSAs[®], the postal code^{OM} usually identifies a specific rural community.

Province (PR)

Province or territory of the postal code^{OM}.

Table 4.4: Province and Territory codes

PR	Description
10	Newfoundland and Labrador
11	Prince Edward Island
12	Nova Scotia
13	New Brunswick
24	Quebec
35	Ontario
46	Manitoba
47	Saskatchewan
48	Alberta
59	British Columbia
60	Yukon
61	Northwest Territories
62	Nunavut

Dissemination area unique identifier (DAuid)

The DAuid uniquely identifies a dissemination area. It is composed of the two-digit province or territory code, the two-digit census division code and the four-digit dissemination area code.

In PCCF+, this identifier can take additional forms depending on how much is known about the subfields. The census division missing value is set to '00' as '99' is a valid code.

99009999	Province (2), census division (2), and dissemination area (4) are unknown
nn009999	Census division (2) and dissemination area (4) are unknown
nnnn9999	Dissemination area (4) is unknown

Dissemination block (DB)

A dissemination block (DB) is an area bounded on all sides by roads and/or boundaries of standard geographic areas. Dissemination blocks cover all the territory of Canada. This code should be combined with the dissemination area unique identifier to uniquely identify the dissemination block within the country.

DB=nnn; 000=missing.

Dissemination block Indian Reserve refusal flag (DB_ir2016)

In 2021, there were a total of 63 Indian reserves and Indian settlements that were 'incompletely enumerated.' For these reserves or settlements, enumeration was either not permitted or could not be completed for various reasons. The 2021 Census population and dwelling counts are not available for the 63 incompletely enumerated Indian reserves and Indian settlements.

F=not an IR refusal area; T=IR refusal area; 9=missing.

Census subdivision unique identifier (CSDuid)

This uniquely identifies a census subdivision in the country. The province/territory, census division, and census subdivision (municipality) codes combine to represent the CSDuid.

In PCCF+, this identifier can take additional forms depending on how much is known about the subfields. The census division missing value is set to '00' as '99' is a valid code.

9900999	Province (2), census division (2), and census subdivision (3) are unknown
nn00999	Census division (2) and census subdivision (3) are unknown
nnnn999	Census subdivision (3) is unknown

Census subdivision name (CSDname)

This contains the name of the census subdivision (municipality) in effect as of 1 January 2021.

Census subdivision type (CSDtyp)

Census subdivisions are classified according to designations adopted by provincial, territorial or federal authorities (see Appendix D, or *Geographic Attributes File, Reference Guide*).

Census metropolitan area / census agglomeration unique identifier (CMA)

3-digit CMA code which uniquely identifies the census metropolitan area or census agglomeration. census metropolitan areas and census agglomerations are not limited by provincial boundaries. For assistance in identifying these codes, use accompanying variable CMAname.

In addition to the 3-digit CMA codes,

999	=CMA unknown
000	=CMA not applicable (not in any CMA or CA)

Census metropolitan area type (CMAtype)

A one-character field identifying whether the unit is a census metropolitan area, a tracted census agglomeration or a non-tracted census agglomeration.

Table 4.5: Census metropolitan area type

CMAtype	Description
B	Census metropolitan area
D	Census agglomeration with no census tracts
G	Strong metropolitan influence zone
H	Moderate metropolitan influence zone
I	Weak metropolitan influence zone
J	No metropolitan influence zone
K	Census agglomeration with census tracts
L	Territories, outside census agglomerations
9	Missing

Census metropolitan area / census agglomeration name (CMAname)

This field contains the census metropolitan area or census agglomeration name.

Census tract name (CTname)

This 7-character field identifies a census tract within a census metropolitan area or census agglomeration. To uniquely identify each census tract in its corresponding census metropolitan area or tracted census agglomeration, the three-digit census metropolitan area / census agglomeration code must precede the census tract 'name.' If a census tract is split into two or more parts due to a population increase, the number after the decimal point identifies the splits. For example, CT 0042.00 becomes CT 0042.01 and CT 0042.02. If CT 0042.01 is subsequently split, it becomes CT 0042.03 and CT 0042.04.

For areas within a census metropolitan area or census tracted census agglomeration where the census tract is missing, the CTname is coded to 9999.99. For areas not in any CMA or census tracted census agglomeration, CTname=0000.00 (not applicable).

Flag for tracted/non-tracted area (Tracted)

Indicates whether a postal code is within the tracted area (census metropolitan area or census-tracted agglomeration).

0=not tracted; 1=tracted; 9=unknown.

Statistical Area Classification code (SACcode)

The statistical area classification, groups census subdivisions according to whether they are a component of a census metropolitan area, a census agglomeration, a census metropolitan influence zone (strong metropolitan influence zone, moderate metropolitan influence zone, weak metropolitan influence zone or no metropolitan influence zone), or the territories (Yukon, Northwest Territories and Nunavut, outside of any census agglomeration).

Table 4.6: Statistical area classification code

SACcode	Description
000	Territories, outside of a census agglomeration
001 to 995	Census metropolitan area / census agglomeration unique identifier
996	Strong metropolitan influence zone
997	Moderate metropolitan influence zone
998	Weak metropolitan influence zone
999	No metropolitan influence zone

Blank	Missing
--------------	---------

Statistical Area Classification type (SACtype)

This identifies the type of statistical area classification in which the census subdivision is located.

Table 4.7: Statistical area classification type

SACtype	Description
1	Census subdivision within census metropolitan area
2	Census subdivision within census agglomeration with at least one census tract
3	Census subdivision within census agglomeration having no census tracts
4	Census subdivision outside of census metropolitan area and census agglomeration having strong metropolitan influence
5	Census subdivision outside of census metropolitan area and census agglomeration having moderate metropolitan influence
6	Census subdivision outside of census metropolitan area and census agglomeration having weak metropolitan influence
7	Census subdivision outside of census metropolitan area and census agglomeration having no metropolitan influence
8	Census subdivision within the territories, outside of census agglomerations
9	Missing

Census consolidated subdivision unique identifier (CCSuid)

The three-character CCS code identifies a census consolidated subdivision within a census division. The CDuid is combined with the CCS code to uniquely identify a census consolidated subdivision (CCSuid) in the country.

999=CCS unknown.

Federal electoral district – 2013 Representation Order unique identifier (FEDuid)

This uniquely identifies a federal electoral district – 2013 representation order. The first two digits of the FEDuid identify the province or territory.

nn999=FEDuid unknown.

Federal electoral district – 2013 Representation Order names (FEDname)

Names for federal electoral districts.

Blank=unknown (FED=999).

Designated place unique identifier (DPLuid)*

The 4-digit DPL code identifies a designated place within a province or territory. DPLuid=PR(2) + DPL(4). A missing or not applicable DPL code is left blank, resulting in DPLuid=nn+blank(4).

* Requires DB precision which is most often not available for these areas.

Designated place type (DPLtype)

The following is a list of designated place types:

Table 4.8: Designated place types

DPLtype	Description
CFA	Class IV area
DMU	Dissolved municipality
DPL	Designated place
IRI	Indian reserve / Réserve indienne
IST	Island trust
LNC	Localité non constituée
LSB	Local service board
LSD	Local service district
LUD	Local urban district
MDI	Municipalité dissoute
MDP	Municipal defined places
MET	Métis settlement
NCM	Northern community
NS	Northern settlement
NVL	Nisga'a village
OHM	Organized hamlet
RPC	Retired population centre
SE	Aboriginal settlement
UNP	Unincorporated place
UUC	Unincorporated urban centre
Blank	Not applicable or missing

Designated place name (DPLname)

Designated place name.

Blank=not applicable or unknown.

Economic region unique identifier (ERuid)

The 2-digit ER code identifies an economic region within a province or territory. Combined with the province or territory code it uniquely identifies an economic region.

missing ERuid=nn99.

Economic region name (ERname)

Economic region name.

Census agricultural region identifier (CARuid)

Census agricultural regions are used by the Census of Agriculture for disseminating agricultural statistics. Census agricultural regions are composed of groups of adjacent census divisions, except in Saskatchewan, where they are composed of groups of adjacent census consolidated subdivisions not respecting census division boundaries. Census agricultural regions are not defined for the territories (CAR='00'). The two-digit census agricultural region code is unique only when preceded by the province code. CARuid=PR+CAR.

nn99=missing; nn00=territories.

Census agricultural region name (CARname)

Census agricultural region name. This shows the name of each census agricultural region, including unofficial descriptive names for otherwise unnamed census agricultural regions.

Population centre/rural area code (POPCTRRAPuid)*

This 6-digit code uniquely identifies the provincial parts of each population centre and rural area (composed of the 2-digit province or territory unique identifier followed by the 4-digit population centre unique identifier). Rural areas = PR(2) + '99' + PR(2). For example, records in rural areas in Manitoba are assigned '469946.' Missing = PR(2) + '9999'.

* Note this code requires dissemination block precision which is unreliable for postal codes^{OM} used by about 25% of the population (primarily those living in urban fringe and rural areas).

Population centre/rural area name (POPCTRRAname)*

Population centre and rural area name.

* Note that this code requires dissemination block precision which is unreliable for postal codes^{OM} used by about 25% of the population (primarily those living in urban fringe and rural areas).

Population centre/rural area type (POPCTRRAtype)*

For population centres, the type code indicates the relationship of the population centre to the census metropolitan area and census agglomeration structure.

Table 4.9: Population centre / rural area type

POPCTRRAtype	Description
1	Core
2	Fringe
3	Rural area inside of a CMA or CA
4	Population centre outside of a CMA or CA
5	Rural area outside of a CMA or CA
6	Secondary core
9	Missing

* Note this code requires dissemination block precision which is unreliable for postal codes^{OM} used by about 25% of the population (primarily those living in urban fringe and rural areas).

Population centre and rural area classification (POPCTRRAClass)***Table 4.10: Population centre and rural area classification**

POPCTRRAClass	Description
1	Rural area
2	Small population centre (1,000 to 29,999)
3	Medium population centre (30,000 to 99,999)
4	Large urban population centre (100,000 or greater)
9	Missing

* Note that this code requires dissemination block precision which is unreliable for postal codes^{OM} used by about 25% of the population (primarily those living in urban fringe and rural areas).

Community Size (CSize)

Community Size is defined in terms of the 2021 census population in each census metropolitan area or census agglomeration. Community size 1 consists of Toronto, Montreal and Vancouver census

metropolitan areas. Community size 2 consists of Ottawa-Gatineau, Edmonton, Calgary, Québec, Winnipeg and Hamilton census metropolitan areas. Community size 3 includes all 18 other census metropolitan areas plus 7 of the larger census agglomerations. Community size 4 includes all 106 other census agglomerations. Community Size 5—"rural and small town Canada"--includes all places not included in any census metropolitan area or census agglomeration. (i.e., places with an urban area population less than about 10,000, plus rural areas.

Note that almost all records with a valid FSA[®] (whether or not the rest of the postal code is valid) can be assigned to a census metropolitan area or census agglomeration, and thus to a community size category. According to Statistics Canada's recommended definition, rural and small town Canada (*Plessis et al, 2001*) is defined as CSize='5'.

Table 4.11: Community size classification

CSize	Description (population)
1	1,500,000 +
2	500,000 – 1,499,999
3	100,000 – 499,999
4	10,000 – 99,999 (any CMACA < 100,000)
5	Less than 10,000 (any non-CMACA)
9	Missing

Community Size and Metropolitan Influence Zones (CSizeMIZ)

This variable is a combination of the CSize variable for urban areas, and of the SAcType variable for rural areas. See the definitions of each for more information.

Table 4.12: Community size classification, with MIZ

CSizeMIZ	Description
1	1,500,000 +
2	500,000 – 1,499,999
3	100,000 – 499,999
4	10,000 – 99,999 (any CMACA < 100,000)
5	Non-CMACA; Strong MIZ
6	Non-CMACA; Moderate MIZ
7	Non-CMACA; Weak / No MIZ, Territories outside of any CA
8	Non-CMACA; unknown MIZ
9	Missing (unknown if CMACA or not)

Health Region (HRuid)

Health regions are sub-provincial areas defined by provincial departments of health. In some cases, those definitions may split dissemination areas or dissemination blocks between two or more health regions, but to simplify the coding here, each dissemination area and dissemination block has been uniquely assigned to a single health region. Since each health region covers many dissemination areas, most of which are not split, this simplification should have little effect on the number of events coded to each health region. The two-character health region code is only unique within a given province (HRuid=Province (2) + health region (2)). Missing = PR(2) + '99'.

Where a province only uses a single digit to represent a health region, a zero has been added preceding that digit. In Ontario, the HRuid previously referred to as the Local Health Integration Networks (LHINs) is now renamed as Home and Community care Support Services. See *Health Regions: Boundaries and Correspondence with Census Geography*. Statistics Canada Catalogue no. 82-402-X for more information on other changes from other provinces. Note that the definitions used were generally those in effect as of December 2022, but the definitions may be changed by provinces at any time, particularly in provinces without a long history of producing data by health region.

Health Region Name – English (HRename)

English health region name.

Health Region Name – French (HRfname)

French health region name.

Alternate Health Region (AHRuid)

Alternate health regions are like health regions. They are for Public Health Units (PHU) defined by the province of Ontario. The alternate health region code is only unique within the province.

'3599' = Missing (but applicable); PR(2) + '99' = Missing (not applicable).

Alternate Health Region Name – English (AHRename)

English alternate health region name.

Alternate Health Region Name – French (AHRfname)

French alternate health region name.

Single link indicator (SLI)

The single link indicator (SLI) provides a geographic record for mapping a postal code^{OM} representative point. It can be used to force a one-to-one relationship between postal codes^{OM} and dissemination areas, dissemination blocks, or block-faces. The SLI has the value of '1' to flag one record of an active Postal Code^{OM}. Every set of retired records for a postal code^{OM}, for a given retirement date, has one SLI equal to '1.' The SLI value '0' indicates additional records. '9' = SLI unknown.

PCCF representative point type (Rep_Pt_Type)

This identifies whether the PCCF record uses a block-face, dissemination block, dissemination area or census subdivision representative point as the coordinate.

Table 4.13: Representative point type

Rep_Pt_Type	Rep_Pt_Type Description
1	Block-face
2	Dissemination block
3	Dissemination area
4	Census subdivision
9	unknown (postal code ^{OM} not found on PCCF)

PCCF+ representative point flag (RPF)

This variable expands on the PCCF Rep_Pt_Type variable, by including information on geocoding from PCCF+.

Table 4.14: Representative point flag

RPF	RPF Description
1	Block-face representative point (Link_Source=F,D)
2	Dissemination block representative point (Link_Source=F,D)
3	Dissemination block representative point imputed within a dissemination area (Link_Source=F,D)
4	Dissemination block representative point imputed within a set of possible DAs (Link_Source=C)
5	Dissemination area representative point imputed within a set of possible DAs (Link_Source=C)
6	Dissemination area representative point imputed from partial postal code ^{OM} (Link_Source=3,4,5)
8	Representative point imputed from first 1 or 2 characters (Link_Source=2,1)
9	Missing

Postal Code^{OM} type (PCTYPE)

This indicates the type of addresses used to identify the points of call served by the postal code^{OM}. This field was introduced by CPC after the creation of the original PCCF. Where possible, a value has been imputed by Statistics Canada for retired postal codes^{OM} using historical address information and delivery mode type.

Table 4.15: Postal Code^{OM} types

PCTYPE	Description
0	Any service from rural postal office (2 nd character of postal code ^{OM} = '0')
1	Street address with letter carrier service from urban post office
2	Street address with route service from urban post office
3	Post office box from urban post office
4	Route service from urban post office
5	General delivery from urban post office
9	Missing

Note: Postal code^{OM} types (PCTYPE) 3 and 5 represent mail service that can be accessed at the post office or other postal installation. Where possible, these PCCF records are coded to the corresponding post office or other postal installation, which is *not* where the addressee resides or does business.

Delivery mode type (DMT)

This is the delivery mode type as defined by CPC. Note that Statistics Canada assigns a DMT of 'W' to rural postal codes, which are left blank by CPC. Note that all DMT except 'W' are delivered from urban post offices. For additional considerations regarding DMT, see Appendix I.

Table 4.16: Delivery mode type (December 2022 PCCF)

DMT	Description	Number of Records
A	Letter carrier delivery to street address	1,164,334
B	Letter carrier delivery to an apartment building	28,384
E	Delivery to a business building	10,595
G	Delivery to a large volume receiver	8,093
H	Delivery via a rural route	7,087
J	General delivery	662
K	Delivery to a post office box (not a community mail box)	9,117
M	Delivery to a large volume receiver (post office box)	4,550
T	Delivery via a suburban service	10,314

DMT	Description	Number of Records
W	Rural postal code ^{OM} (the second character of the postal code is '0')	122,792
X	Delivery via a mobile route	0
Z	postal code ^{OM} is retired (no further delivery to this code)	33,419
9	Missing	6,086
Total		1,405,433

Historic delivery mode type (H_DMT)

The historic delivery mode retains the previous delivery mode type value, if known. Blank = Not applicable or unknown. See Table 4.16 for categorization.

Historic delivery mode type difference (DMTDIFF)

This field is for the previous Delivery mode type (DMT) if different from the current DMT. This usually occurs when the current DMT=Z (retired). Blank=Not applicable or unknown. See Table 4.16 for categorization.

PCCF delivery installation (PO)

This indicates whether the PCCF record represents coding to a post office where the mail can be accessed. The value '1' indicates this PCCF record was coded to a post office or other postal installation and the value '2' indicates 'unknown.' The value '0' indicates this PCCF record was coded to the area serviced by the postal code^{OM}.

PCCF quality indicator (QI)

The PCCF quality indicator provides an indicator of the quality of the geocoding that links the postal code^{OM} and its address information and that of the Geography Division's Spatial Data Infrastructure. The QI is established at the record level and is currently available only for the postal codes^{OM} that were originally geocoded using the automated geocoding system. A QI of 'AAA' indicates the highest quality and a QI of 'CCC' indicates the lowest quality.

Latitude (LAT)

This is the latitude, in decimal degrees, of the dissemination area, dissemination block, or block-face representative point. The decimal point is explicit.

Longitude (LONG)

This is the longitude, in decimal degrees, of the dissemination area, dissemination block, or block-face representative point. The decimal point is explicit.

Source of PCCF geographic coding (Source)

The PCCF source variable indicates the primary source of the geocoding used for the PCCF.

Table 4.17: Source of geocoding (PCCF)

Source	Description
1	Automated geocoding directly to Census geographic areas
2	Geocoded using Census response
3	Converted from geocoding done to previous Census geographic areas
4	Manually geocoded
5	Geocoded using latitude/longitude coordinates
6	Geocoded using imputation match
9	Missing

Source of PCCF+ geographic coding (Link_Source)

This field provides additional information on the source of geographic coding related to PCCF+ matching process. The possible values of this field are as follows:

Table 4.18: Source of geographic coding (PCCF+)

Link_Source	Description
F	A full set of geographic codes and latitude/longitude were derived from an exact match to a PCCF unique record.
D	A full set of geographic codes and latitude/longitude were derived from an exact match to a PCCF duplicate record.
C	A full set of geographic codes and latitude/longitude were derived from an exact match to a WCF record (for DMT of H, J, K, R, T, W, Z, and some M).
5	Full geography was imputed from the first 5 characters of a postal code ^{OM} (when DMT=9), using census population weights.
4	Full geography was imputed from the first 4 characters of a postal code ^{OM} (when DMT=9), using census population weights.
3	Partial geography was imputed from the first 3 characters of a postal code ^{OM} (when DMT=9 or most M), using census population weights.
2	A partial set of geographic codes were assigned based on only the first 2 characters of this postal code ^{OM} .
1	A province code was assigned based on only the first character of this postal code ^{OM} . No other geographic codes or latitude and longitude were assigned.
0	The first character of this postal code ^{OM} is not in the set used for Canadian postal codes ^{OM} . No geographic codes assigned.
9	Missing / Not Applicable

PCCF+ link type (Link)

The link type code identifies the type of problems encountered in coding. The link type codes (LINK) and corresponding messages (MESSAGE) are arranged in hierarchical order, starting with 0 for the most serious problems, and going to 9 for no problem at all (not even a Warning or Note). If more than one type of problem was present, only the most erroneous type is shown.

Table 4.19: Linkage type code (PCCF+)

Link	Message
0	Error: No match to PCCF (unique, duplicate, or weighted conversion file).
1	Error: Linked to PO geography
2	Warning: Non-residential. DMT=E, G or M and InstFlag=- (probable non-residential).
3	Warning: Business building (usually not a legitimate residence). DMT=E and InstFlag =blank.
4	Warning: Commercial or institutional (check if legitimate residence, and if pertinent to your study). DMT=G or M and InstFlag =blank.

Link	Message
5	Note: Retired postal code ^{OM} – expected and normal on administrative files. No further action required (slight chance of DMT problem prior to retirement, only if DMT=Z, and DMTDIFF=blank).
6	Note: Multiple match to dissemination area using unweighted allocation. No further action required.
7	Weighted allocation using census population weights (6-character WCF).
9	Not applicable (no error, warning or note). <i>Such records do not appear on the problem file.</i>

PCCF+ Precision (Prec)

Precision of representative point (latitude / longitude). 0 is the least precise coding, 9 is the most precise. For spatial studies it is recommended to review this variable before analysis of all records.

Table 4.20: Precision of representative point

Prec	Description
0	No geographic coding
1	Imputed from first 1 or 2 characters
2	Imputed from first 3 characters
3	Imputed from first 4 characters
4	Imputed from first 5 characters
5	1 or more dissemination areas (WC6) (DMT=H-X)
6	2 or more dissemination areas (DMT=A,B,E,G)
7	1 dissemination area (DMT=A,B,E,G)
8	1 dissemination block (DMT=A,B,E,G)
9	1 block-face (DMT=A,B,E,G)

Number of Census Divisions (nCD)

Number of Census Divisions (CDs) connected to the postal code.

Number of Census Subdivisions (nCSD)

Number of Census Divisions (CSDs) connected to the postal code.

CPC Community name (Comm_Name)

The community name, as defined by CPC, denotes any city, town or village in Canada that is recognised as a valid mailing address. The service area of a CPC community *never* corresponds exactly to the boundaries of a Statistics Canada census subdivision, even when the names are identical.

CPC air stage community (AirLift)

An air stage office is a post office to or from which all mail must be airlifted for more than six months of every year as a viable surface transportation alternative is not available. These offices are generally confined to remote or isolated communities. An office designated an air stage office is deemed to be air stage for the whole year.

'*' = air stage office; blank= not an air stage office or missing.

Institutional Flag (InstFlag)

This field is used to help identify records likely to be for institutional residents. It is usually blank. The categories should not be expected to correspond to the classification of facilities used by the Health Statistics Division, provincial or territorial authorities. See the problem file for the building name and address of these large volume receivers.

Table 4.21: Institutional flag

InstFlag	Description
E	School or university residence
G	Group home
H	Hospitals
M	Military bases
N	Nursing and group homes
P	Prisons, jails
R	Religious
S	Seniors residences
T	Hotels, motels
U	Other
Blank	Not applicable or unknown

Residential flag (ResFlag)

If the delivery mode type (DMT) is E, G or M, then ResFlag indicates postal codes^{OM} for possible or improbable residence addresses, or postal codes^{OM} for which the residential or non-residential nature is undetermined. If the DMT is not in E, G or M, then ResFlag will be blank. See the problem file output for CPC building name and address information, if available.

Table 4.22: Residential flag

ResFlag	Description
+	Possible residence
-	Improbable residence
?	DMT= E,G or M but residence status is undetermined
Blank	Not in DMT= E,G, or M

InuitLands

Inuit Nunangat land claims settlement regions.

Table 4.23: Land claims settlement regions

InuitLands	Description
0	Outside of Inuit Nunangat
1	Inuvialuit Settlement Region (Northwest Territories)
2	Nunavut Territory
3	Nunavik (Northern Quebec)
4	Nunatsiavut (Labrador)
9	Unknown

Neighbourhood before tax income per single person equivalent (BTIPPE)

Neighbourhood income per single person equivalent is a household size-adjusted measure of household income (before tax), based on 2021 Census Summary profile data at the DA level, and using

person-equivalents implied by the 2021 low income cut-offs (LICOs). See Appendix E for more detail on how these measures were created.

Neighbourhood after tax income per single person equivalent (ATIPPE)

Neighbourhood after tax income per single person equivalent.

Neighbourhood income quintile (CMA/CA) Before Tax (QABTIPPE)

Neighbourhood income quintiles (area-based) were created using BTIPPE. See Appendix E for more detail on how these measures were created.

4.24: Neighbourhood income per person equivalent (CMA/CA) Before Tax

QABTIPPE	Description
1	Lowest quintile
2	Medium-low quintile
3	Middle quintile
4	Medium-high quintile
5	Highest quintile
9	Missing

Neighbourhood income quintile (national) Before Tax (QNB TIPPE)

This variable is calculated similarly to the QABTIPPE, but where the *national* distribution for income quintiles is considered rather than area-based (census metropolitan area / census agglomeration) distributions.

Neighbourhood income decile (CMA/CA) Before Tax (DABTIPPE)

Calculated the same as the QABTIPPE, expressed as deciles instead of quintiles.

Neighbourhood income decile (national) Before Tax (DNBTIPPE)

Calculated the same as the QNB TIPPE, expressed as deciles instead of quintiles.

Neighbourhood income quintile (CMA/CA) After Tax (QAATIPPE)

This variable is calculated similarly to the QABTIPPE, but where income after-tax is considered rather than before tax.

Neighbourhood income quintile (national) After Tax (Q NATIPPE)

This variable is calculated similarly to the QAATIPPE, but where the *national* distribution for income quintiles is considered rather than area-based (census metropolitan area / census agglomeration) distributions.

Neighbourhood income decile (CMA/CA) After Tax (DAATIPPE)

Calculated the same as the QAATIPPE, expressed as deciles instead of quintiles.

Neighbourhood income decile (national) After Tax (DNATIPPE)

Calculated the same as the QNATIPPE, expressed as deciles instead of quintiles.

Imputation flag for IPPE (IMPFLG)

Mean household income imputed

2016 dissemination area (DA16uid)

This field shows the 2021 dissemination area (PR+CD+DA),, based on the Correspondence File Reference Guide Census year 2021

2011 dissemination area (DA11uid)

This field shows the 2011 dissemination area (PR+CD+DA), based on the Correspondence File Reference Guide Census year 2016

2011 dissemination block (DB11uid)

This field shows the 10-digit 2011 dissemination block unique identifier (PR+CD+DA+DB based on the Correspondence File Reference Guide Census year 2016

2006 dissemination area (DA06uid)

This field shows the 2006 dissemination area (PR+CD+DA), based on the based on the Correspondence File Reference Guide Census year 2016

2001 dissemination area (DA01uid)

This field shows the 2001 dissemination area (PR+CD+DA), based on the based on the Correspondence File Reference Guide Census year 2016

1996 enumeration area (EA96uid)

This field shows the 1996 enumeration area (PR+FED+EA), based on the based on the Correspondence File Reference Guide Census year 2016

1991 enumeration area (EA91uid)

This field shows the 1991 enumeration area (PR+FED+EA), based on the based on the Correspondence File Reference Guide Census year 2016

1986 enumeration area (EA86uid)

This field shows the 1986 enumeration area (PR+FED+EA), based on the based on the Correspondence File Reference Guide Census year 2016.

1981 enumeration area (EA81uid)

This field shows the 1981 enumeration area (PR+FED+EA), based on the based on the Correspondence File Reference Guide Census year 2016

File naming convention

Each PCCF+ primary version number (1 through 8) reflects a major update with geocoding to a different census period. Version 8 was developed using 2021 census geography and version 7 was created using 2016 census geography, and so on. The revision, indicated by a letter indicates the update version. PCCF+ is updated regularly to reflect new postal code^{OM} updates, updates to health region boundaries, and other programming updates suggested by users.

The SAS[®] control program file name for this release is PCCFplus_8A.sas.

5. Data quality

Linkage data quality elements provide information on the fitness-for-use of a linkage database by describing why, when, and how the data are created, and how accurate the data are. The quality elements include an overview reporting on lineage, positional accuracy, attribute accuracy, logical consistency and completeness. This information is provided to users for all linkage data products.

Lineage

Lineage describes the history of the linkage data, including descriptions of the source material from which the data were derived and the methods of derivation. It also contains the dates of the source material, and all transformations involved in producing the final digital files.

The PCCF+ is updated regularly via two processes. The first updates are done every five years, after each census, to align the source files to the latest census geographic areas and calculate new postal code population weights. The second is ongoing maintenance of the source files with updated postal codes^{OM} from CPC, updated health regions, and other updated coding as required. Updates to each revision are described in the *Version updates* section of this reference manual.

Positional accuracy

Positional accuracy refers to the absolute and relative accuracy of the positions of geographic features. Absolute accuracy is the closeness of the coordinate values in a dataset to values accepted as being true. Relative accuracy is the closeness of the relative positions of features to their respective relative positions accepted as or being true. Descriptions of positional accuracy include the quality of the final file or product after all transformations.

Geographic coordinates included in PCCF+ source files are either from the PCCF, from the Geographic Attribute File, or from FSA[®] centroids. Geographic coordinates (latitude and longitude) assigned to postal codes^{OM} are taken from the PCCF flat file. Where postal codes^{OM} are incompletely coded, coordinates are taken from dissemination block, dissemination area, or FSA[®] centroids, each calculated from the corresponding geographic product.

For more information on the method used to calculate representative points, see the documentation for the PCCF.

Attribute accuracy

Attribute accuracy refers to the accuracy of the quantitative and qualitative information attached to each feature (such as population for a population centre, street name, census subdivision name and code).

PCCF+ produces a flat file that includes data elements from the PCCF, geographic attribute file, health region boundaries, and other correspondence files. For many postal codes^{OM} there is not a one-to-one match to geographic attributes. In these cases, possible matches are drawn through population-weighting via a random selection process. For more details see the *Technical specifications* section.

Tests are run to ensure that certain basic data relationships were consistent within the set of records in PCCF+.

Logical consistency

Logical consistency describes the fidelity of relationships encoded in the data structure of the digital linkage data.

In some cases, including most rural areas, postal code^{OM} service areas do not respect dissemination area boundaries. In such cases, multiple records for a postal code^{OM} reflect the relationship between the postal code^{OM} and census geographic areas. Also, a postal code^{OM} can be linked to more than one block-face or dissemination block within the same dissemination area. PCCF+ uses several methods to geocode postal codes^{OM} with multiple matches, using either a population-weighting or a random allocation approach. Details of these methods are outlined in the *Technical specifications* section.

Consistency with other products

The PCCF+ maintains consistency with several other Statistics Canada products. These include:

- PCCF
- Geographic Attribute File (GAF)
- 2021 census forward sortation area boundary file
- Health region boundaries
- CPC air stage offices
- Correspondence files for standard geographical identifiers

PCCF+ is updated annually to reflect changes to these products, with details of updates provided in the *Version updates* section.

Completeness

Completeness refers to the degree to which geographic features, their attributes and their relationships are included or omitted in a dataset. It also includes information on selection criteria, definitions used, and other relevant mapping rules.

Completeness for PCCF+ is the degree to which all potential postal codes^{OM} are accounted for and all geographic codes and attributes are linked to a postal code^{OM}. PCCF+ includes all active postal codes^{OM} as well as all retired postal codes^{OM} within the geocoding process.

Indicators are provided on geocoded output for the quality of links, the precision of geographic coordinates, and other flags for errors and warnings. The details of these codes are provided in the *Technical specifications* section.

Appendix A: Glossary

Block-face

A block-face is one side of a street between two consecutive features intersecting that street. The features can be other streets or boundaries of standard geographic areas.

Block-faces are used for generating block-face representative points, which in turn are used for geocoding and census data extraction when the street and address information are available.

Census agricultural region

Census agricultural regions (CARs) are composed of groups of adjacent census divisions. In Saskatchewan, census agricultural regions are made up of groups of adjacent census consolidated subdivisions, but these groups do not necessarily respect census division boundaries.

Census consolidated subdivision

A census consolidated subdivision (CCS) is a group of adjacent census subdivisions. Generally, the smaller, more densely-populated census subdivisions (towns, villages, etc.) are combined with the surrounding, larger, more rural census subdivisions, in order to create a geographic level between the census subdivision and the census division.

Census division

Census division is the general term for provincially legislated areas (such as county, *municipalité régionale de comté* and regional district) or their equivalents. Census divisions are intermediate geographic areas between the province/territory and the municipality (census subdivision).

Census metropolitan area and census agglomeration

A census metropolitan area (CMA) or a census agglomeration (CA) is formed by one or more adjacent municipalities centred on a population centre (known as the core). A CMA must have a total population of at least 100,000 of which 50,000 or more must live in the core. A CA must have a core population of at least 10,000. To be included in the CMA or CA, other adjacent municipalities must have a high degree of integration with the core, as measured by commuting flows derived from previous census place of work data.

If the population of the core of a CA declines below 10,000, the CA is retired. However, once an area becomes a CMA, it is retained as a CMA even if its total population declines below 100,000 or the population of its core falls below 50,000. Small population centres with a population count of less than 10,000 are called fringe. All areas inside the CMA or CA that are not population centres are rural areas.

When a CA has a core of at least 50,000, it is subdivided into census tracts. Census tracts are maintained for the CA even if the population of the core subsequently falls below 50,000. All CMAs are subdivided into census tracts.

Metropolitan influence zone

The metropolitan influence zone (MIZ) is a concept that geographically differentiates the area of Canada outside census metropolitan areas (CMAs) and census agglomerations (CAs). Census subdivisions (CSDs) within provinces that are outside CMAs and CAs are assigned to one of four categories according to the degree of influence (strong, moderate, weak or no influence) that the CMAs or CAs have on them. CSDs within the territories that are outside CAs are assigned to a separate category.

Census subdivisions within provinces are assigned to a MIZ category based on the percentage of their resident employed labour force that commutes to work in the core(s) of any CMA or CA. CSDs with the same degree of influence tend to be clustered. They form zones around CMAs and CAs that progress through the categories from 'strong' to 'no' influence as distance from the CMAs and CAs increases. As many CSDs in the territories are very large and sparsely populated, the commuting flow of the resident employed labour force is unstable. For this reason, CSDs in the territories that are outside CAs are assigned to a separate category that is not based on their commuting flows.

Census subdivision

Census subdivision (CSD) is the general term for municipalities (as determined by provincial/territorial legislation) or areas treated as municipal equivalents for statistical purposes (such as Indian reserves, Indian settlements and unorganized territories).

Census tract

Census tracts (CTs) are small, relatively stable geographic areas that usually have a population between 2,500 and 8,000 persons. They are located in census metropolitan areas and in census agglomerations that had a core population of 50,000 or more in the previous census.

A committee of local specialists (for example, planners, health and social workers, and educators) initially delineates census tracts in conjunction with Statistics Canada. Once a census metropolitan area (CMA) or census agglomeration (CA) has been subdivided into census tracts, the census tracts are maintained even if the core population subsequently declines below 50,000.

Coordinate system

A coordinate system is a reference system based on mathematical rules for specifying positions (locations) on the surface of the earth. The coordinate values can be spherical (latitude and longitude) using angular units of measure such as degrees, minutes and seconds or planar (Lambert conformal conic) using linear units such as metres.

Cartographic boundary files, digital boundary files, representative points and road network files are disseminated in Lambert conformal conic projection.

Core, fringe and rural area*

The terms 'core,' 'fringe' and 'rural area' replace the terms 'urban core,' 'urban fringe' and 'rural fringe' for the 2016 census. These terms distinguish between population centres (POPCTRs) and rural areas (RAs) within a census metropolitan area (CMA) or census agglomeration (CA).

The term 'fringe' includes all population centres within a CMA or CA that have less than 10,000 persons and are not contiguous with the core.

All territory within a CMA or CA that is not classified as a core or fringe is classified as rural area.

* Note that this code requires dissemination block precision which is unreliable for postal codes^{OM} used by about 25% of the population (primarily those living in urban fringe and rural areas).

Designated place*

A designated place (DPL) is normally a small community or settlement that does not meet the criteria established by Statistics Canada to be a census subdivision (an area with municipal status) or a population centre. Designated places are created by provinces and territories, in cooperation with Statistics Canada, to provide data for sub-municipal areas.

* Note that this code requires dissemination block precision which is unreliable for postal codes^{OM} used by about 25% of the population (primarily those living in urban fringe and rural areas).

Dissemination area

A dissemination area (DA) is a small, relatively stable geographic unit composed of one or more adjacent dissemination blocks. It is the smallest standard geographic area for which all census data are disseminated. DAs cover all the territory of Canada.

Dissemination block

A dissemination block (DB) is an area bounded on all sides by roads and/or boundaries of standard geographic areas. The dissemination block is the smallest geographic area for which population and dwelling counts are disseminated. Dissemination blocks cover all the territory of Canada.

Economic region

An economic region (ER) is a grouping of complete census divisions (CDs) (with one exception in Ontario) created as a standard geographic unit for analysis of regional economic activity.

Federal electoral district

A federal electoral district (FED) is an area represented by a member of the House of Commons. The federal electoral district boundaries used for the 2016 census were based on the 2013 representation order.

Geocoding

Geocoding is the process of assigning geographic identifiers (codes or x,y coordinates) to map features and data records. The resulting geocodes permit data to be linked geographically to a place on the earth.

Geographic code

A geographic code is a numerical identifier assigned to a geographic area. The code is used to identify and access standard geographic areas for the purposes of data storage, retrieval and display.

Geographic reference date

The geographic reference date is a date determined by Statistics Canada for the purpose of finalizing the geographic framework for which census data will be collected, tabulated and reported. For the 2021 census, the geographic reference date was 1 January 2021.

Population centre*

A population centre (PopCtr) has a population of at least 1,000 and a population density of 400 persons or more per square kilometre, based on the current census population count. All areas outside population centres are classified as rural areas. Taken together, population centres and rural areas cover all of Canada.

* Note that this code requires dissemination block precision which is unreliable for postal codes^{OM} used by about 25% of the population (primarily those living in urban fringe and rural areas).

Population density

Population density is the number of persons per square kilometre.

Postal code^{OM}

The postal code^{OM} is a six-character code defined and maintained by CPC for the purpose of sorting and delivering mail.

Province or territory

'Province' and 'territory' refer to the major political units of Canada. Canada is divided into 10 provinces and 3 territories.

Representative point

A representative point is a coordinate point that represents a line or a polygon. The point is centrally located along the line, and centrally located or population weighted in the polygon.

On the PCCF, postal codes^{OM} are linked to block-face representative points (coordinates) when street address information is available; otherwise, they are linked to dissemination block (DB) representative points. In some cases, postal codes^{OM} are linked to dissemination area (DA) representative points when they cannot be linked to dissemination blocks.

Rural area*

Rural areas (RAs) include all territory lying outside population centres (PopCtr). Taken together, population centres and rural areas cover all of Canada.

Rural population includes all population living in rural areas of census metropolitan areas (CMAs) and census agglomerations (CAs), as well as population living in rural areas outside CMAs and CAs.

* Note that this code requires dissemination block precision which is unreliable for postal codes^{OM} used by about 25% of the population (primarily those living in urban fringe and rural areas).

Statistical Area Classification

The Statistical Area Classification (SAC) groups census subdivisions according to whether they are a component of a census metropolitan area, a census agglomeration or a metropolitan influence zone (MIZ). The MIZ classifies all CSDs in provinces and territories that are outside census metropolitan areas and census agglomerations.

Appendix B: Record Layouts for PCCF+ input files**Air stage office (cpcref.airstage.v2212.txt)****B.1: CPC airstage delivery (6+ months per year)**

Position	Size	Type	Read by PCCF+	Field Name	Description
1	6	C	Y	PCODE	Postal code ^{OM}
7	2	C	N	PR	Province
9	30	C	Y	Comm_Name	CPC community name

Building name and addresses (cpcref.bldgnam.v2212.txt)**B.2: Building names and addresses (possible non-residential) for manual resolution.**

Position	Size	Type	Read by PCCF+	Field Name	Description
2	3	N	Y	NumAdr	Number of address ranges at this postal code ^{OM}
6	1	N	N	Hosp	Hospital flag
8	6	C	Y	PCODE	Postal code ^{OM}
15	1	C	N	DMT	Delivery mode type
17	97	C	Y	NameAdr	Building name and address
116	26	C	Y	Comm_Name	CPC community name
143	2	C	N	PROV	Province

Residential Postal Code^{OM} file (cpcref.egmres.v2212.txt)**B.3: Flag for possible residential postal codes^{OM} despite institutional type**

Position	Size	Type	Read by PCCF+	Field Name	Description
1	6	C	Y	PCODE	Postal code ^{OM}
8	1	C	Y	ResFlag	Flag for residential (+) or possible non-residential (-)
10	1	C	N	DMT	Delivery mode type
11	1	C	N	DMTAny	Delivery mode type
12	1	C	N	H_DMT	Historical delivery mode type
14	97	C	Y	BldgName	Building Name

Institutional flag (cpcref.instflag.v2212.txt)**B.4: Institutional flag.**

Position	Size	Type	Read by PCCF+	Field Name	Description
1	6	C	Y	PCODE	Postal code ^{OM}
12	1	C	Y	InstFlag	Institutional flag

Address range file (cpcref.nadr.v2212.txt)**B.5: Number of address records for each postal code^{OM}.**

Position	Size	Type	Read by PCCF+	Field Name	Description
1	6	C	Y	PCODE	Postal code ^{OM}
8	1	C	N	DMT	Delivery mode type
9	3	N	Y	NumAdr	Number of address ranges for this postal code ^{OM}

PCCF+ duplicates file (pccf2212.pccf.dups.txt) – A tab delimited format (columns header included)**B.6: Duplicate postal codes^{OM} from PCCF.**

Type	Read by PCCF+	Field Name	Description
C	Y	PCODE	Postal code ^{OM}
C	Y	FSA	Forward Sortation Area [®]
C	Y	PR	Province code
C	Y	CDuid	Census Division code
C	Y	CSDuid	Census Subdivision code
C	Y	CSDname	Census Subdivision name
C	Y	CSDtype	Census Subdivision type
C	Y	CCScode	Census Consolidated Subdivision code
C	Y	SAC	Statistical Area Classification code
C	Y	SACtype	Statistical Area Classification type
C	Y	CTname	Census Tract name
C	Y	ER	Economic Region identifier
C	Y	DPL	Designated Place identifier
C	Y	FEDuid	Federal Electoral District (2003) unique identifier
C	Y	Pop_Cntr_RA	Population centre/rural area code
C	Y	Pop_Cntr_RA_type	Population centre/rural area type
C	Y	DAuid	Dissemination area code
C	Y	DB	Dissemination block code
C	Y	Rep_Pt_Type	Representative Point Type
N	Y	LAT	Latitude of lowest level geographic area
N	Y	LONG	Longitude of lowest level geographic area
C	Y	SLI	Single Link Indicator
C	Y	PCtype	Postal code ^{OM} type
C	Y	Comm_Name	CPC Community Name
C	Y	DMT	Delivery Mode Type
C	Y	H_DMT	Historic Delivery Mode Type
C	Y	DMTDIFF	Previous or alternate DMT (if applicable)
C	Y	Birth_Date	Birth date of postal code ^{OM} (yyyymmdd)
C	Y	Ret_Date	Retirement date of postal code ^{OM} (yyyymmdd)
C	Y	PO	Delivery installation
C	Y	QI	Quality indicator
C	Y	Source	Source of postal code geocoding (PCCF)

Type	Read by PCCF+	Field Name	Description
C	Y	Pop_Cntr_RA_Size_Class	Population centre/rural area classification
N	Y	nCD	Number of census divisions for this Postal Code ^{OM} (1-9+)
N	Y	nCSD	Number of census subdivisions for this postal code ^{OM} (1-9+)
N	Y	nDA	Number of dissemination areas for this postal code ^{OM} (1-9+)
N	Y	nDB	Number of dissemination blocks for this postal code ^{OM} (1-9+)
N	N	_keeprec	Postal code has both retired and rebirthed records

Duplicate pointer file (pccf2212.pccf.pointdup.txt)

B.7: Pointer file for duplicate Postal Codes^{OM}.

Position	Size	Type	Read by PCCF+	Field Name	Description
1	6	C	Y	PCODE	Postal code ^{OM}
7	4	N	Y	nPCODE	Number of records for this postal code ^{OM}
11	8	N	Y	ObsDup	Observation number for first occurrence on duplicates file

Rural Post Office (pccf2212.pccf.rpo.txt) – A tab delimited format (columns header included)

B.8: Coding for rural post offices.

Type	Read by PCCF+	Field Name	Description
C	Y	PCODE	Postal code ^{OM}
C	Y	FSA	Forward Sortation Area [®]
C	Y	PR	Province code
C	Y	CDuid	Census Division code
C	Y	CSDuid	Census Subdivision code
C	Y	CSDname	Census Subdivision name
C	Y	CSDtype	Census Subdivision type
C	Y	CCScode	Census Consolidated Subdivision code
C	Y	SAC	Statistical Area Classification code
C	Y	SACtype	Statistical Area Classification type
C	Y	CTname	Census Tract name
C	Y	ER	Economic Region identifier
C	Y	DPL	Designated Place identifier
C	Y	FED	Federal Electoral District (2013) unique identifier
C	Y	Pop_Cntr_RA	Population centre/rural area code
C	Y	Pop_Cntr_RA_type	Population centre/rural area type
C	Y	DAuid	Dissemination area code
C	Y	DB	Dissemination block code
C	Y	Rep_Pt_Type	Representative Point Type
N	Y	LAT	Latitude of lowest level geographic area
N	Y	LONG	Longitude of lowest level geographic area
C	Y	SLI	Single Link Indicator

Type	Read by PCCF+	Field Name	Description
C	Y	PCtype	Postal code ^{OM} type
C	Y	Comm_Name	CPC Community Name
C	Y	DMT	Delivery Mode Type
C	Y	H_DMT	Historic Delivery Mode Type
C	Y	DMTDIFF	Previous or alternate DMT (if applicable)
C	Y	Birth_Date	Birth date of postal code ^{OM} (yyyymmdd)
C	Y	Ret_Date	Retirement date of postal code ^{OM} (yyyymmdd)
C	Y	PO	Delivery installation
C	Y	QI	Quality indicator
C	Y	Source	Source of postal code ^{OM} geocoding (PCCF)
C	Y	POP_CNTR_RA_SIZE_CLASS	Population centre/rural area classification

Unique Postal Codes^{OM} (pccf2212.pccf.uniq.txt) – A tab delimited format (columns header included)

B.9: Unique postal codes^{OM}.

Type	Read by PCCF+	Field Name	Description
C	Y	PCODE	Postal code ^{OM}
C	Y	FSA	Forward Sortation Area [®]
C	Y	PR	Province code
C	Y	CDuid	Census Division code
C	Y	CSDuid	Census Subdivision code
C	Y	CSDname	Census Subdivision name
C	Y	CSDtype	Census Subdivision type
C	Y	CCScode	Census Consolidated Subdivision code
C	Y	SAC	Statistical Area Classification code
C	Y	SACtype	Statistical Area Classification type
C	Y	CTname	Census Tract name
C	Y	ER	Economic Region identifier
C	Y	DPL	Designated Place identifier
C	Y	FED	Federal Electoral District (2013) unique identifier
C	Y	Pop_Cntr_RA	Population centre/rural area code
C	Y	Pop_Cntr_RA_type	Population centre/rural area type
C	Y	DAuid	Dissemination area code
C	Y	DB	Dissemination block code
C	Y	Rep_Pt_Type	Representative Point Type
N	Y	LAT	Latitude of lowest level geographic area
N	Y	LONG	Longitude of lowest level geographic area
C	Y	SLI	Single Link Indicator
C	Y	PCtype	Postal code type
C	Y	Comm_Name	CPC Community Name
C	Y	DMT	Delivery Mode Type
C	Y	H_DMT	Historic Delivery Mode Type
C	Y	DMTDIFF	Previous or alternate DMT (if applicable)

Type	Read by PCCF+	Field Name	Description
C	Y	Birth_Date	Birth date of postal code ^{OM} (yyyymmdd)
C	Y	Ret_Date	Retirement date of postal code ^{OM} (yyyymmdd)
C	Y	PO	Delivery installation
C	Y	QI	Quality indicator
C	Y	Source	Source of postal code ^{OM} geocoding (PCCF)
C	Y	POP_CNTR_RA_SIZE_CLASS	Population centre/rural area classification
N	Y	nBLK	Number of dissemination blocks for this postal code ^{OM} (1-9+)

First 2 character unique and duplicates (pccf2212.pccf.wc2dups.txt) – A tab delimited format (columns header included)

B.10: Weighting for first 2 characters of postal code^{OM}.

Type	Read by PCCF+	Field Name	Description
C	Y	PCODE2	First 2 characters of postal code ^{OM}
C	Y	PR	Province code
C	Y	CD	Census Division code
C	Y	DA	Dissemination Area code
C	Y	CSDuid	Census subdivision code
C	Y	SAC	Statistical Area Classification code
C	Y	CT	Census Tract name
C	Y	Tracted	Flag for census tracted (DA)
N	Y	LAT	Latitude of dissemination area centroid
N	Y	LONG	Longitude of dissemination area centroid
C	Y	RPF	Representative point flag (PCCF+)
N	Y	nDA	Number of dissemination areas for PCODE2
N	Y	nCD	Number of census divisions for PCODE2
N	Y	nCSD	Number of census subdivisions for PCODE2
N	Y	PC2DAWT	Dissemination area level weight for PCODE2

Pointers for first 2 characters (pccf2212.pccf.wc2point.txt)

B.11: Pointer for 2-character weighting file.

Position	Size	Type	Read by PCCF+	Field Name	Description
1	2	C	Y	PCODE2	First 2 characters of postal code ^{OM}
7	7	N	Y	FirstObs	Pointer for first observation on wc2dups
14	5	N	Y	nOBS	Total number of observations for first 2 characters
19	4.2	N	Y	TWT	2 character population weight

Forward Sortation Area[®] unique and duplicates (pccf2212.pccf.wc3dups.txt) – A tab delimited format (columns header included)

B.12: Weighting for first 3 characters of postal code^{OM} (FSA[®]).

Type	Read by PCCF+	Field Name	Description
C	Y	PCODE3	First 3 characters of postal code ^{OM}
C	Y	PR	Province code
C	Y	CD	Census Division code
C	Y	DA	Dissemination Area code
C	Y	CSDuid	Census subdivision code
C	Y	SAC	Statistical Area Classification
C	Y	CT	Census Tract name
C	Y	Tracted	Flag for census tracted (DA)
N	Y	LAT	Latitude of dissemination area centroid
N	Y	LONG	Longitude of dissemination area centroid
C	Y	RPF	Representative point flag (PCCF+)
N	Y	nDA	Number of dissemination areas for PCODE3
N	Y	nCD	Number of census divisions for PCODE3
N	Y	nCSD	Number of census subdivisions for PCODE3
N	Y	PC3DAWT	Dissemination area level weight for PCODE3

Forward Sortation Area[®] pointers (pccf2212.pccf.wc3point.txt)

B.13: Pointer for 3-character weighting file (FSA[®]).

Position	Size	Type	Read by PCCF+	Field Name	Description
1	3	C	Y	PCODE3 (FSA [®])	Forward Sortation Area [®]
7	7	N	Y	FirstObs	Pointer for first observation on wc3dups
14	5	N	Y	nOBS	Total number of observations for FSA [®]
19	4.2	N	Y	TWT	3 character population weight

WCF unique and duplicates, 4-character (pccf2212.pccf.wc4dups.txt) – A tab delimited format (columns header included)

B.14: Weighting for first 4 characters of postal code^{OM}.

Type	Read by PCCF+	Field Name	Description
C	Y	PCODE4	First 4 characters of postal code ^{OM}
C	Y	PR	Province code
C	Y	CD	Census Division code
C	Y	DA	Dissemination Area code
C	Y	CSDuid	Census subdivision code
C	Y	SAC	Statistical Area Classification
C	Y	CT	Census Tract name
C	Y	Tracted	Flag for census tracted (DA)
N	Y	LAT	Latitude of dissemination area centroid

Type	Read by PCCF+	Field Name	Description
N	Y	LONG	Longitude of dissemination area centroid
C	Y	RPF	Representative point flag (PCCF+)
N	Y	nDA	Number of dissemination areas for PCODE4
N	Y	nCD	Number of census divisions for PCODE4
N	Y	nCSD	Number of census subdivisions for PCODE4
N	Y	PC4DAWT	Dissemination area level weight for PCODE4

WCF pointers, 4-character (pccf2212.pccf.wc4point.txt)

B.15: Pointer for 4-character weighting file.

Position	Size	Type	Read by PCCF+	Field Name	Description
1	4	C	Y	PCODE4	First 4 characters of postal code ^{OM}
7	7	N	Y	FirstObs	Pointer for first observation on wc4dups
14	5	N	Y	nOBS	Total number of observations for PCODE4
19	4.2	N	Y	TWT	Total weight

WCF unique and duplicates, 5-character (pccf2212.pccf.wc5dups.txt) – A tab delimited format (columns header included)

B.16: Weighting for first 5 characters of postal code^{OM}.

Type	Read by PCCF+	Field Name	Description
C	Y	PCODE5	First 5 characters of postal code ^{OM}
C	Y	PR	Province code
C	Y	CD	Census Division code
C	Y	DA	Dissemination Area code
C	Y	CSDuid	Census subdivision code
C	Y	SAC	Statistical Area Classification
C	Y	CT	Census Tract name
C	Y	Tracted	DA is census tracted
N	Y	LAT	Latitude of dissemination area centroid
N	Y	LONG	Longitude of dissemination area centroid
C	Y	RPF	Representative point flag
N	Y	nDA	Number of dissemination areas PCODE5
N	Y	nCD	Number of census divisions for PCODE5
N	Y	nCSD	Number of census subdivisions for PCODE5
N	Y	PC5DAWT	Dissemination area level weight for PCODE5

WCF pointers, 5-character (pccf2212.pccf.wc5point.txt)**B.17: Pointer for 5-character weighting file.**

Position	Size	Type	Read by PCCF+	Field Name	Description
1	5	C	Y	PCODE5	First 5 characters of postal code ^{OM}
7	7	N	Y	FirstObs	Pointer for first observation on wc5dups
14	5	N	Y	nOBS	Total number of observations for PCODE5
19	4.2	N	Y	TWT	Total weight

WCF unique and duplicates (pccf2212.pccf.wc6dups.txt) – A tab delimited format (columns header included)**B.18: Weighting for postal code^{OM}.**

Type	Read by PCCF+	Field Name	Description
C	Y	PCODE	6 character postal code ^{OM}
C	Y	PR	Province code
C	Y	CD	Census Division code
C	Y	DA	Dissemination Area code
C	Y	CSDuid	Census subdivision code
C	Y	SAC	Statistical Area Classification
C	Y	CT	Census Tract name
C	Y	Tracted	Census tracted DA
N	Y	LAT	Latitude of dissemination area centroid
N	Y	LONG	Longitude of dissemination area centroid
C	Y	DMT	Delivery Mode Type
C	Y	H_DMT	Historic Delivery Mode Type
C	Y	DMTDIFF	Previous DMT if different from current DMT
C	Y	Rep_Pt_Type	Representative Point Type
C	Y	PCtype	Postal code ^{OM} type
N	Y	nDA	Number of dissemination areas for PCODE
N	Y	nCD	Number of census divisions for PCODE
N	Y	nCSD	Number of census subdivisions for PCODE
N	Y	PC6DAWT	Dissemination area level weight for PCODE

WCF pointers, 6-character (pccf2212.pccf.wc6point.txt)**B.19: Pointer for postal code^{OM} weighting file.**

Position	Size	Type	Read by PCCF+	Field Name	Description
1	6	C	Y	PCODE	6 character postal code ^{OM}
7	7	N	Y	FirstObs	Pointer for first observation on wc6dups
14	5	N	Y	nOBS	Total number of observations for PCODE
19	4.2	N	Y	TWT	Total weight

DA-Block pointer file (georef21.dablkpnt21.txt)**B.20: Pointer file for dissemination area and dissemination block file.**

Position	Size	Type	Read by PCCF+	Field Name	Description
1	8	C	Y	DAuid	Dissemination area unique identifier
1	2	C	Y	PR	Province
3	2	C	Y	CD	Census division
5	4	C	Y	DA	Dissemination area
9	4	N	Y	nBLK	Number of dissemination blocks in DA
13	6	N	Y	FirstObs	Observation of first block occurrence
19	6	N	Y	DAPop2021	Sum of DB population within DA
26	3	C	Y	SACcode	Statistical Area Classification code
30	3	C	Y	CSD	Census subdivision

Partial geographic attribute file (georef21.gaf21.txt) – A tab delimited format (columns header included)**B.21: Subset of Geographic Attribute File.**

Type	Read by PCCF+	Field Name	Description
C	Y	PR	Province or territory code
C	Y	CD	Census division
C	Y	DA	Dissemination area
C	Y	DB	Dissemination block
C	Y	BLK	Block face
C	Y	PRuid	Uniquely identifies a province or territory
C	N	PRdguid	Dissemination Geography Unique Identifier for 2021 provinces or territories
C	Y	PRabbr	Province or territory name in English
C	Y	PRfabbr	Province or territory name in French
C	Y	CDuid	Uniquely identifies a census division (composed of the 2-digit province or territory unique identifier the 2-digit census division code)
C	N	CDdguid	Dissemination Geography Unique Identifier for 2021 census divisions
C	Y	CDname	Census division name
C	Y	FEDuid	Uniquely identifies a federal electoral district (composed of the 2-digit province or territory unique identifier followed by the 3-digit federal electoral district code)

C	N	FEDdguid	Dissemination Geography Unique Identifier for 2021 federal electoral districts
C	Y	FEDname	Federal electoral district name
C	Y	CSDuid	Uniquely identifies a census subdivision (composed of 2-digit province or territory unique identifier followed by the 2-digit census division code and 3-digit census subdivision code)
C	N	CSDdguid	Dissemination Geography Unique Identifier for 2021 census subdivisions
C	Y	CSDname	Census subdivision name
C	Y	CSDtype	Census subdivisions are classified according to designations adopted by provincial, territorial or federal authorities
C	Y	DPLuid	Uniquely identifies a designated place (composed of the 2-digit province or territory unique identifier followed by the 4-digit designated place code)
C	N	DPLdguid	Dissemination Geography Unique Identifier for 2021 designated places
C	Y	DPLname	Designated place name
C	Y	DPLtype	Designated place type
C	Y	ERuid	Uniquely identifies an economic region (composed of the 2-digit province or territory unique identifier followed by the 2-digit economic region code)
C	Y	ERdguid	Dissemination Geography Unique Identifier for 2021 economic regions
C	Y	ERname	Economic region name
C	Y	CCSuid	Uniquely identifies a census consolidated subdivision (composed of the 2-digit province or territory unique identifier followed by the 2-digit census division code and the 3 digit census consolidated subdivision code)
C	N	CCSdguid	Dissemination Geography Unique Identifier for 2021 census consolidated subdivisions
C	Y	CCSname	Census consolidated subdivision name
C	Y	SACtype	Statistical Area Classification
C	Y	SACcode	The 3-digit Statistical Area Classification code
C	Y	CMApuid	Uniquely identifies the provincial or territorial part of a census

			metropolitan area and census agglomeration
C	N	CMApdguid	Dissemination Geography Unique Identifier for 2021 Census metropolitan area or census agglomeration (provincial component)
C	Y	CMAuid	Uniquely identifies a census metropolitan area or census agglomeration
C	N	CMAadguid	Dissemination Geography Unique Identifier for 2021 census metropolitan areas or census agglomerations
C	Y	CMAname	Census metropolitan area or census agglomeration name
C	Y	CMAtype	A one-character field identifying whether the unit is a census metropolitan area, a tracted census agglomeration or a non-tracted census agglomeration
C	Y	CTuid	Uniquely identifies a census tract within a census metropolitan area or census agglomeration (composed of the 3-digit census metropolitan area or census agglomeration code followed by the 7-character census tract name)
C	N	CTdguid	Dissemination Geography Unique Identifier for 2021 census tracts
C	Y	CTcode	Uniquely identifies a census tract
C	Y	CTname	Every census tract is assigned a 7.2-character numeric 'name' (including leading zeros, a decimal point and trailing zeros)
C	Y	PopCtrRAPuid	Uniquely identifies the provincial parts of each population centre and rural area (composed of the 2-digit province or territory unique identifier followed by the 4-digit population centre unique identifier)
C	Y	PopCtrRAPDGuid	Dissemination Geography Unique Identifier for 2021 population centres and rural areas (provincial component)
C	Y	PopCtrRADGuid	Uniquely identifies a population centre and rural area
C	Y	PopCtrRAname	Population centre and rural area name
C	Y	PopCtrRAtype	Population centre and rural area type
C	Y	PopCtrRAclass	Distinguishes between small population centres, medium population centres, large urban population centres and rural areas

C	Y	DAuid	Uniquely identifies a dissemination area (composed of the 2-digit province or territory unique identifier followed by the 2-digit census division code, and the 4-digit dissemination area code)
C	Y	DAdguid	Dissemination Geography Unique Identifier for 2021 dissemination areas
C	Y	DB21dguid	Dissemination Geography Unique Identifier for 2021 dissemination blocks
C	Y	DB_ir2021	2021 Census Indian reserve refusal flag
C	Y	ADAuid	Uniquely identifies an aggregate dissemination area
C	Y	ADAguid	Dissemination Geography Unique Identifier for 2021 aggregate dissemination areas
C	Y	ADAcodes	Aggregate dissemination area code
C	Y	CSIZE	Community size code
C	Y	CSIZEMIZ	Community size with MIZ
C	Y	CARUID	Census Agricultural region unique identifier
C	Y	CARname	Census Agricultural region name
C	Y	InuitLands	Inuit Nunangat land claim regions
C	Y	DB16UID	Dissemination Geography Unique Identifier for 2016 dissemination blocks
C	Y	DB16DGUID	Dissemination Geography Unique Identifier for 2016 dissemination blocks
C	Y	DB11uid	Dissemination Geography Unique Identifier for 2011 dissemination blocks
C	Y	DB06uid	Dissemination Geography Unique Identifier for 2006 dissemination blocks
C	Y	DA06uid	Dissemination Geography Unique Identifier for 2006 dissemination areas
C	Y	DA01uid	Dissemination Geography Unique Identifier for 2001 dissemination areas
C	Y	EA96uid	Enumeration area (1996)
C	Y	EA91uid	Enumeration area (1991)
C	Y	EA86uid	Enumeration area (1986)
C	Y	EA81uid	Enumeration area (1981)
C	Y	da16uid	Dissemination Geography Unique Identifier for 2016 dissemination areas
C	Y	DBuid	Dissemination Geography Unique Identifier for 2021 dissemination blocks

N	Y	DALAT	Dissemination area latitude coordinate (decimal degrees)
N	Y	DALONG	Dissemination area longitude coordinate (decimal degrees)
N	Y	DBPOP2021	Dissemination Block population 2021 (rounded)
N	Y	DBtdwell2021	Dissemination Block total private dwelling 2021
C	Y	FED	Federal Electoral District (2013) unique identifier
C	Y	CSD	Census subdivision unique identifier
C	Y	DA11uid	Dissemination Geography Unique Identifier for 2011 dissemination areas

Health region definition (georef21.hrdef.txt)

B.22: Health region and alternate health region definition.

Position	Size	Type	Read by PCCF+	Field Name	Description
1	11	C	N	DBuid	Dissemination block code
1	2	C	Y	PR	Province code
3	2	C	Y	CD	Census Division code
5	4	C	Y	DA	Dissemination Area code
9	3	C	Y	DB	Dissemination Block code
12	8	C	N	DAuid	Dissemination Area code
20	3	C	Y	CSD	Census subdivision code
27	4	C	Y	HRuid	Health Region unique identifier
31	60	C	Y	HRename	Health Region name (English)
91	60	C	Y	HRfname	Health Region name (French)
151	4	C	Y	AHRuid	Alternate Health Region unique identifier
155	60	C	Y	AHRename	Alternate Health Region name (English)
215	60	C	Y	AHRfname	Alternate Health Region name (French)

SES reference file (georef21.ses21.txt) – A tab delimited format (columns header included)

B.23: Neighbourhood income quintiles (2021 DA identifiers).

Type	Read by PCCF+	Field Name	Description
C	Y	DAUID	Dissemination Area unique identifier
C	Y	AREA	CMA/CA identifier or: CMA/CA code or R+PR (Ex : R24 = Québec)
C	Y	IMPFLG	Imputation flag for IPPE (mean household income imputed)
N	Y	BTIPPE	Neighbourhood before tax income per single person equivalent (BTIPPE)
N	Y	ATIPPE	Neighbourhood after tax income per single person equivalent (ATIPPE)
C	Y	QABTIPPE	Neighbourhood income quintile before tax (in CMA/CA)
C	Y	QNBTPPE	Neighbourhood income quintile before tax (national)
C	Y	DABTIPPE	Neighbourhood income decile before tax (in CMA/CA)

Type	Read by PCCF+	Field Name	Description
C	Y	DNBTIPPE	Neighbourhood income decile before tax (national)
C	Y	QAATIPPE	Neighbourhood income quintile after tax (in CMA/CA)
C	Y	QNATIPPE	Neighbourhood income quintile after tax (national)
C	Y	DAATIPPE	Neighbourhood income decile after tax (in CMA/CA)
C	Y	DNATIPPE	Neighbourhood income decile after tax (national)

Appendix C: Supplementary data file included with PCCF+

The PCCF+ includes the complete 2021 Census-DA weight file.

File name: DA_Nat_wgt_2021.txt

Total record count: 750,604

Number of unique postal codes: 679,002

Number of unique FSAs: 1,644

Number of unique DAs: 56,541

Record layout:

@1	postal code	\$char6.
@7	dissemination area	8.
@15	weight	5.2

The content of this file has been subjected to all regular census data quality and area suppression rules.

Appendix D: Hierarchy of standard geographic units for dissemination, 2021 Census

See Figure 1.1 Hierarchy of standard geographic areas for dissemination, 2021 Census from the *Dictionary, Census of Population, 2021*. https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/fig/index-eng.cfm?ID=F1_1

Appendix E: Derivation of neighbourhood income quantiles

2021 Census Profile data available at the dissemination area (DA) geographic level were used to calculate neighborhood income quintiles and deciles. This analysis relied on single-person equivalent (SPE) to take into consideration household size. SPE is the person-weights implicit in the Statistics Canada low-income cut-offs needed to specify an appropriate multiplier for different household sizes (see Table below). This adjustment is necessary since it generally costs less per person for two or more persons living together compared to one person living alone. SPE values were based on low-income lines for 2020.

Table 1. Single-person equivalent weights for the calculation of neighbourhood income quintiles or deciles

	Before-tax income	After-tax income
Single-person household	1.00	1.00
2-person household	1.24	1.22
3-person household	1.53	1.52
4-person household	1.86	2.14
5+ person household	2.24	2.27

Low income cut-offs (LICOs) before and after tax by community size and family size, in current dollars. Retrieved 14 December 2022.

<https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1110024101&pickMembers%5B0%5D=2.1&cubeTimeFrame.startYear=2005&cubeTimeFrame.endYear=2020&referencePeriods=20050101%2C20200101>

For each DA, total income was calculated by multiplying the median income¹ (either before- or after-tax) by the number of households. That total income (for each DA) was then divided by the sum of single-person equivalent in the DA to obtain income per single person equivalent (IPSPEs). In the 2021 Statistics Canada census, there were a total of 57,936 DAs. Complete income and population data were available for 54,191 DAs. The IPSPEs for 3,262 DAs were imputed based on the average IPSPE values in adjacent DAs (as defined using a first-order queen spatial adjacency matrix). 483 DAs had missing income data as well as no adjacent DAs that could be used for imputation.

National-level neighbourhood income quintiles and deciles were constructed by ranking all DAs in Canada from the lowest IPSPE to the highest IPSPE, and then assigning each DA to one of five groups (for quintiles) and one of ten groups (for deciles). The assignment of DAs to national-level income quintiles and deciles was weighted by residential population, such that each group contained approximately one-fifth or one-tenth of the total in-scope population.

In addition to national-level quantiles, area-based income quintiles and deciles were created based on the set of DAs within each census metropolitan area (CMA), each census agglomeration (CA), and for the residual non-CMA and non-CA DAs within each province. To be clear, each DA was assigned into one geographical group; a CMA (if the DA was located in a CMA), a CA (if the DA was located in a tracted CA), or a province/territory (if the DA was located outside of a CMA or tracted CA). There were 63 geographical groups used to calculate area-based income quintiles and deciles (i.e., 63 different quintile and decile rankings). Area-based income quintiles and deciles were constructed by ranking all DAs within a geographical group from the lowest IPSPE to the highest IPSPE, and then assigning each DA to one of five groups (for quintiles) and to one of ten groups (for deciles). The assignment of DAs to area-based income quintiles and deciles was weighted by residential population, such that each group contained approximately one-fifth or one-tenth of the total in-scope population within a CMA, a CA, or the non-CMA and non-CA DAs in a province. The rationale for creating area-based income quantiles is that housing and other costs vary across Canada. For example, rents and house prices in some places, such as most of Quebec and the Atlantic provinces, have historically been lower than those in Toronto, Ontario or Vancouver, British Columbia.

Appendix F: Definition of “Usual place of residence”

The following is an excerpt from the Dictionary, Census of Population, 2021

“Usual place of residence in Canada” refers to the main dwelling in which the person lives most of the time. It is used to identify the person as a member of a particular household and, potentially, family (depending on the composition of the household).

For persons with one residence, that residence is their usual place of residence.

For persons with no residence, their usual place of residence is where they stayed on May 11, 2021.

For persons with more than one residence in Canada, their usual place of residence is where the person lives most of the time, with the following exceptions:

- The usual residence of spouses or common-law partners who stay elsewhere while working or studying is the residence where their family lives, if they return periodically.
- The usual residence of students is that of their parents, if they return to live with their parents during the year even if they live elsewhere while attending school or working at a summer job.
- The usual residence of persons who have lived in an institutional collective dwelling for six months or more is the institutional collective dwelling in which they currently reside. Institutional collective dwellings include correctional institutions, hospitals, nursing homes, chronic and long term care hospitals and related institutions. The usual residence of persons who have lived in an institutional collective dwelling for less than six months is the private dwelling to which they can return. If they do not have a private dwelling to which they can return, then the institutional collective dwelling is their usual place of residence.
- For persons who spend equal time at each residence (as in some cases of joint custody), their usual place of residence is determined by where they stayed on May 11, 2021.

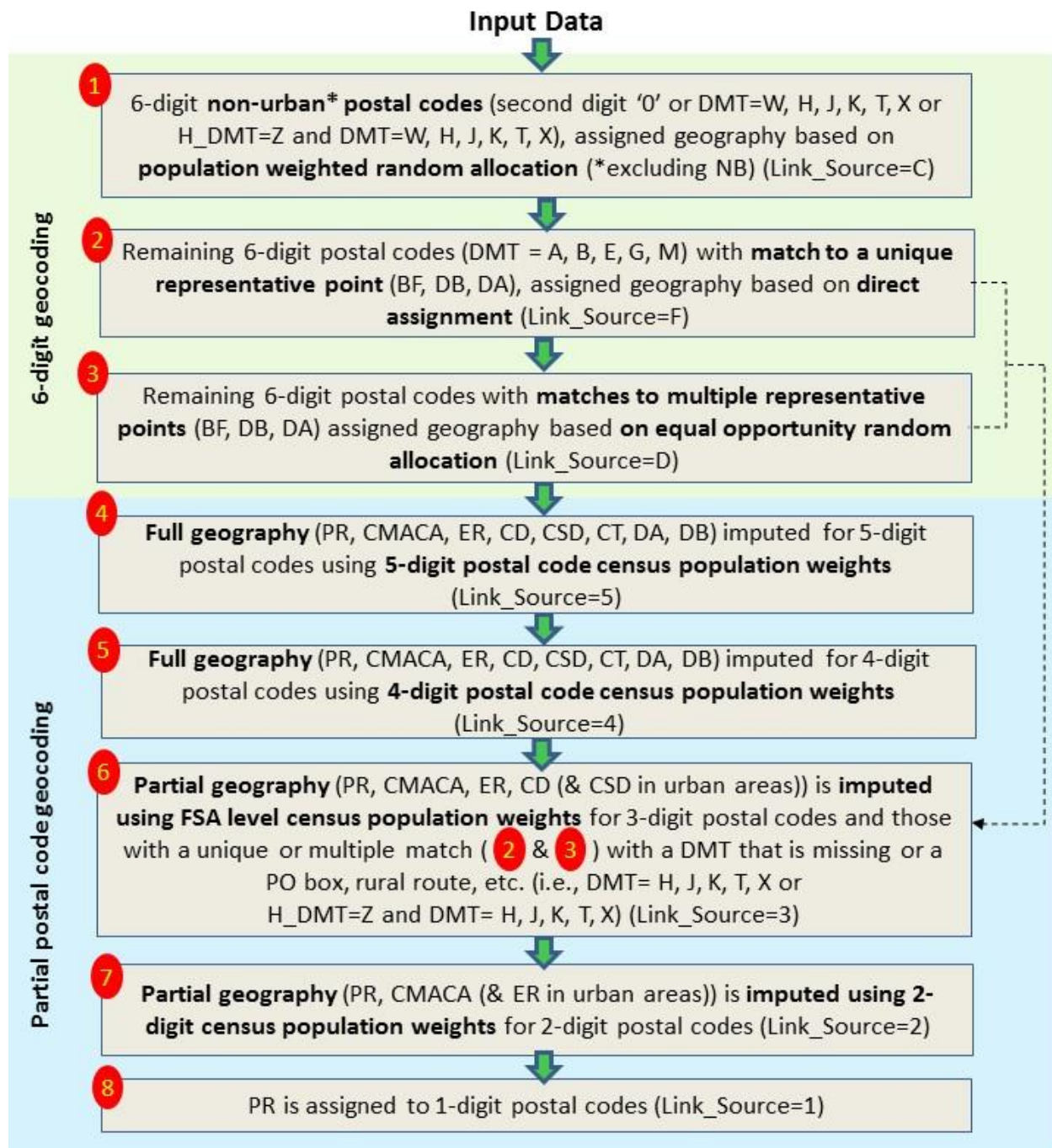
For persons with a residence in Canada and a residence outside Canada, their Canadian residence is their usual place of residence.

Canadian government employees, including Canadian Armed Forces personnel, residing outside Canada for all of the reference period are out of scope for most surveys. The census, which does include them, determines a geographic location for their usual place of residence using the address they used for election purposes or their last permanent address.

In the context of the Census of Population, ‘person’ refers to Canadian citizens (by birth or by naturalization), a landed immigrants (permanent resident), a person who has claimed refugee status (asylum claimant), and a person from another country with a work or study permit. Family members living with work or study permit holder are also included. Foreign residents are excluded.

Appendix G: Residential Routine Flowchart for PCCF+

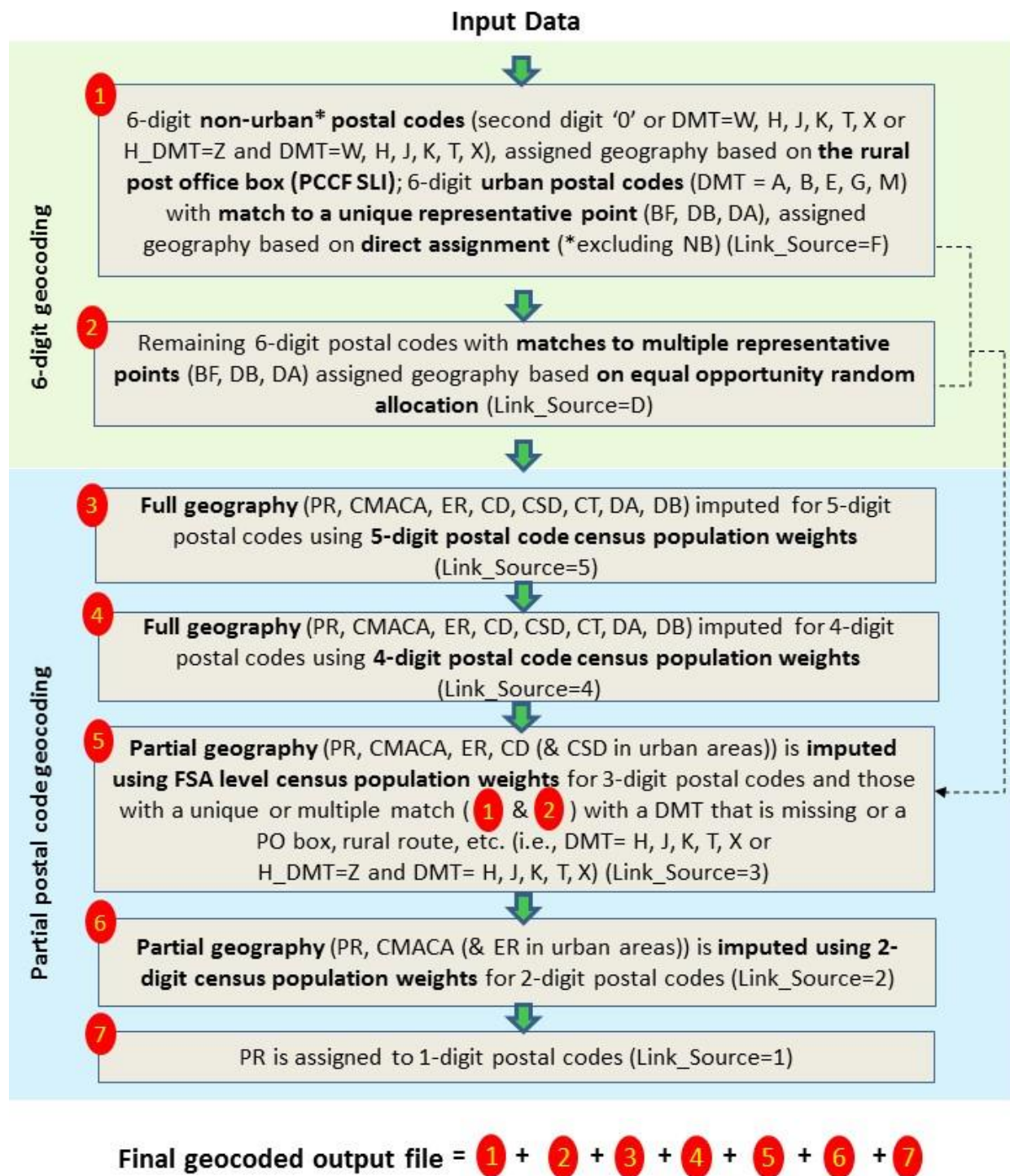
Geocoding of Residential Postal Codes in PCCF+



Final geocoded output file = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8

Appendix H: Institutional Routine Flowchart for PCCF+

Geocoding of Institutional Postal Codes in PCCF+



Appendix I: Additional notes on Delivery Mode Type (DMT)

- W Rural postal codes^{OM} (regardless of type of service) always have a DMT of 'W'. Where more than 1 census subdivision is served by the rural post office, this will result in a note to that effect on the problem file. No action is recommended in such cases, since manual coding would defeat the population-weighted allocation.
- A Ordinary household (including community mail boxes) served by letter carrier. The most common DMT; usually no problem.
- B Apartment building (large) served by letter carrier. No problem with this DMT.
- E Business buildings served by letter carrier. This DMT results in a warning message, with the suggestion to check postal code^{OM} and address, to see if they refer to a legitimate residence or office location. In most cases, the residential flag field will indicate whether the postal code^{OM} is probable or improbable as a place of residence. The building name and brief address are shown on the problem file. The legitimacy of a postal code^{OM} with this DMT may also depend on the nature of the records being coded: appropriate codes for offices are not necessarily appropriate for residences.
- G Large Volume Receiver served by letter carrier (includes many institutions). This DMT results in a warning message, with the suggestion to check postal code^{OM} and address, to see if they refer to a legitimate residence or office location. In most cases, the residential flag field will indicate whether the postal code^{OM} is probable or improbable as a place of residence. The building, company or institution name and brief address will be shown on the problem file. The legitimacy of postal codes^{OM} with this DMT may also depend on the nature of the records being coded: appropriate codes for offices are not necessarily appropriate for residences. For example, a postal code^{OM} for a nursing home may be reasonable for coding the place of usual residence on a death record, but it would be highly suspicious on a birth record. Even if it is a legitimate place of residence, consider whether an event at such a place is pertinent to your study. For example, if the study uses neighbourhood income as a proxy for individual or family-level socioeconomic position, it may not make sense to include college or university residences.
- H Rural route delivery from urban post office. For most rural routes, the weighted conversion file shows the 2016 census population weights associated with each postal code^{OM} and dissemination area combination. If the Link_Source is not equal to 'C', then only province and census metropolitan area will be imputed from FSA[®], since the service area of these postal codes^{OM} extends out into adjacent rural FSAs[®].
- J General delivery. Residence location may be available from census data (weighted conversion file, Link_Source=C). Otherwise, this DMT will result in an error.
- K Group of post office boxes. Residence location may be available from census data (WCF). Otherwise, this DMT will result in an error.
- M Single post office box. If present on the weighted conversion file (Link_Source =C), will be fully coded. In most cases, the RESFLAG field will indicate whether the postal code^{OM} is probable or improbable as a place of residence. The building, company or institution name and brief address will be shown on the problem file. If not present on the weighted conversion file, postal codes with this DMT will result in an error, since the PCCF only links postal codes^{OM} with this DMT to post office location.
- R Miscellaneous delivery services. Residence location may be available from census data (weighted conversion file). Otherwise, this DMT will result in an error, as the regular PCCF only links these postal codes^{OM} to post office location. *DMT R is no longer used by CPC, but it may appear in the field for previous DMT.*
- T Suburban service delivery (rare). Residence location may be available from the weighted conversion file. Otherwise, this DMT will result in an error, as the regular PCCF only links these to post office location.

- X Mobile route (urban industrial areas; rare). This DMT will result in an error, as the regular PCCF only links these to post office location. *However, since in such cases the first three characters of the postal code^{OM} are known to be valid, then a "most likely" province and census metropolitan area may often be imputed.*
- W Rural postal codes^{OM}. Usually geography for records with rural postal codes^{OM} will be derived from the weighted conversion file (Link_Source =C).
- Z Retired postal codes^{OM}. Usually the DMTDIFF field will show the previous DMT for retired postal codes^{OM}. If so, the Link_Source and other diagnostic codes make use of the DMTDIFF. However, if DMTDIFF is blank, then there is a slight chance that a currently retired postal code^{OM} may have formerly had a DMT of E, G, M or X, so this condition will result in output of the record to the problem file with a warning message to that effect.
- 9 Not applicable. No exact match to the PCCF or weighted conversion file, hence DMT is unknown. These will result in an error message as well as output to the problem file. A full or partial set of geographic codes may still be assigned based on the first 1 to 5 characters of the postal code (Link_Source =1, 2, 3, 4 or 5).

Special note concerning Delivery Mode Types H, J, K, M, R and T: Except on rare occasions, it is not necessary to manually recode records with a DMT of H (for rural route delivery from an urban post office), J (general delivery – pick up from an urban post office counter), K (pick-up from group of urban post office boxes), or T (suburban service delivery from an urban post office). Most postal codes^{OM} with those DMTs can now be assigned a full set of geographic codes by reference to the weighted conversion file (Link_Source =C). That also applies to many postal codes^{OM} with DMT of M (pick up from a single large urban post office box) and R (miscellaneous services; no longer used by CPC).