Centre for Demography

Methodology / Notice to users:

Annual postcensal population, Census subdivisions, Canada

Mai 2024

1. Background

Statistics Canada's Centre for Demography produces annual population estimates at the census subdivision (CSD) level. The reference date for these population estimates is July 1 of each year. Population estimates are produced using the 2021 Standard Geographical Classification (SGC).

Since the components of population growth are not produced at the CSD level, the population of CSDs is not estimated using the cohort component method, as is done with higher levels of geography.

The method consists in using T1FF administrative data to calculate the change in the population of CSDs between two annual periods.[1,2,3,4,5] These changes, expressed as a percentage, are then applied to the adjusted census populations of the CSDs. The process is repeated for subsequent years. Finally, calibration is performed to ensure consistency between the CSD and census division (CD) population estimates.

> The T1 family file (T1FF) is derived from the Canada Revenue Agency (CRA) T1 file by the Centre for Income and Socioeconomic Well-being Statistics of Statistics Canada.

2. Base population of census subdivisions

A base population is the population at the beginning of a period, used as a starting point for the estimation process. At the CSD level, the base population is the population count by age and gender for five-year censuses, adjusted for coverage errors. The census data are adjusted as follows:

- Adjustment of the population for census net undercoverage (CNU).

---

[1] The T1FF is used for calculating all postcensal estimates, with the exception of the preliminary estimates, which are based on the T1. The main difference between those two files is that the T1FF includes taxfilers and their dependents, while the T1 only includes taxfilers.

[2] Population estimates for CSDs in Quebec are provided by l'*Institut de la statistique du Québec* (ISQ).

[3] Territorial-level population estimates for Yukon are produced by Statistics Canada, while census subdivision population estimates are supplied by the Yukon Bureau of Statistics for July 1, 2019 and after. Variations in total population estimates at the territorial level between both sources are due to methodological differences.

[4] Population estimates for census subdivisions in Northwest Territories are provided by the Northwest Territories Bureau of Statistics for July 1, 2001, and after. Data for unorganized regions is suppressed, resulting in total population estimate variations when comparing to estimates produced by Statistics Canada for other levels of geography.

[5] Population estimates for CSDs in Alberta are provided by Alberta Office of Statistics and Information (OSI) in current geography for July 1, 2022 and after. They are converted to the appropriate standard geographical classification (SGC) using geographical relationships provided by the OSI. Due to this geographical difference, small discrepancies may exist between CSD populations released by Statistics Canada for Alberta and those found directly on the OSI's website. CSD population estimates for Alberta prior to July 1, 2022 are produced using the same method adopted for provinces and territories that do not provide their own estimates.

Given that coverage studies are not designed to produce subprovincial-level estimates of CNU, provincial and territorial rates by age and gender are used.

- Addition of independent population estimates for incompletely enumerated Indian reserves by age and gender.

- Integration of population count amendments.

The population universe of the 2021 Census includes the following groups:

- Canadian citizens (by birth or by naturalization), landed immigrants (permanent residents), and (since 1991) non-permanent residents. Non-permanent residents are persons who have claimed refugee status [asylum claimants], or persons who hold a work or study permit and their family members living with them. All such persons are included in the population provided they have a usual place of residence in Canada.
- The total population also includes certain Canadian citizens and landed immigrants (permanent residents) living outside the country: government employees working outside Canada; embassy staff posted to other countries; members of the Canadian Armed Forces stationed outside Canada; and Canadian crew members of merchant vessels and their families. Together, they are referred to as 'persons living outside Canada.'
- Foreign residents are excluded from census data: for example, residents of another country visiting Canada temporarily, government representatives of another country posted in Canada and members of the armed forces of another country stationed in Canada.

3. Preparing the data used to produce CSD population estimates

Before CSD population estimates are produced, T1FF microdata are prepared to remove unnecessary records (emigrants, deceased persons, etc.). Then, the addresses of taxfilers and dependents are geocoded using PCCF+[6] so that they are assigned a CSD geographic code based on the basis of the postal code.

4. Producing CSD population estimates using T1FF administrative data

There are four steps to estimate the population of CSDs: (1) determine the number of people in each CSD based on the T1FF for each annual file; (2) calculate the annual population changes for each CSD; (3) apply the annual population changes to the base population of each CSD; and (4) calibrate the postcensal estimates to ensure consistency between the CSD and CD population estimates.

[6] https://www150.statcan.gc.ca/n1/en/catalogue/82F0086X

4.1. Step 1 – Calculating the number of people per CSD using the T1FF

The first step consists in determining the number of people in each CSD based on the T1FF, which was geocoded using postal code information. The coverage rate of the population in the T1FF is not necessarily 100%. As a result, CSD population estimates are not based directly on these counts. Instead, they are used to calculate annual changes in the population (see next step). The assumption is made that the coverage rate of the population in the administrative data remains constant over the postcensal period.

4.2. Step 2 – Calculating annual population changes for each CSD

On the basis of two consecutive annual frequency tables (produced in step 1), annual change in the population is calculated as follows:

$$ChangeCSD^{x}_{(t,t+1)} = \frac{(PopCSD\_T1FF^{x}_{(t+1)} - PopCSD\_T1FF^{x}_{(t)})}{PopCSD\_T1FF^{x}_{(t)}}$$

where

$ChangeCSD^{x}_{(t,t+1)}$ = Annual population change of CSD *x* between time *t* and *t+1*;

$PopCSD\_T1FF^{x}_{(t)}$ = Population of CSD *x* based on T1FF at time *t*.

If, for a given CSD, the denominator used to calculate the population change equals 0, the change in the CD in which that CSD is located is imputed.

The annual population changes calculated in step 2 are not necessarily used in full to calculate population estimates. Although the administrative data used are of high quality—as they were not collected specifically for the purpose of measuring population growth—they can generate abnormal variations, particularly for small CSDs. To avoid applying inaccurate population growth numbers, an algorithm was developed to replace annual changes deemed incorrect or too unreliable.

Annual changes in the population of CSDs are used in full when the following conditions are met:

- The adjusted census base population of the CSD must be equal to or higher than 250.
- The coverage rate of the CSD, expressed as the ratio of the T1FF population to the adjusted census base population, must be within an acceptable range.
- The population change must be within an acceptable range for the annual reference period and for the previous and subsequent annual periods.

The acceptable thresholds and levels were determined using several tests aimed at minimizing the error of closure, one of the main quality indicators of population estimates.

When at least one of the conditions listed above is not met, the annual change in the CSD population is not used and is replaced by the annual population change of the corresponding CD.

4.3 Step 3 – Applying the annual population changes to the base population of each CSD

First, the annual change is applied to the adjusted census population at the beginning of the cycle. For example, to estimate the postcensal population as of July 1, 2021, the annual change in the population for the 2020/2021 period is applied to the adjusted census population for 2021. Since this base population is as of May 11, 2021, a ratio corresponding to the fraction of the year between the census date and July 1 is also applied.

To finalize the process of producing postcensal population estimates as of July 1, 2021, it is necessary to calibrate to the CD population estimates (see step 4).

Steps 3 and 4 are then repeated for each subsequent reference date for which estimates must be calculated.

Estimating the CSD postcensal population as of July 1, 2021:

$$PopCSD^{x}_{(t+1)} = PopCSD^{x}_{(census)} + \left[(PopCSD^{x}_{(census)} * ChangeCSD^{x}_{(t,t+1)}) * \frac{Days_{(census,t+1)}}{Days_{(t,t+1)}}\right]$$

Estimating the CSD postcensal population for subsequent years:

$$PopCSD^{x}_{(t+1)} = PopCSD^{x}_{(t)} + (PopCSD^{x}_{(t)} * ChangeCSD^{x}_{(t,t+1)})$$

where

$PopCSD^{x}_{(t)}$	=	Estimate of the population of CSD *x* at time *t*;
$PopCSD^{x}_{(census)}$	=	Adjusted population of CSD *x* at census date;
$ChangeCSD^{x}_{(t,t+1)}$	=	Annual population change of CSD *x* between time *t* and *t+1*;
$Days_{(census,t+1)}$	=	Number of days between census date and time *t+1*;
$Days_{(t,t+1)}$	=	Number of days between time *t* and *t+1*.

4.4 Step 4 – Calibrating the CSD population estimates to the CD population estimates

Finally, to ensure consistency between the CSD and CD population estimates, the CSD population estimates are adjusted using calibration.

5. Estimating the CD population

The CD postcensal population estimates are produced using the cohort component method. It uses the following formula:

$$P_{(t+1)} = P_{(t)} + B_{(t,t+1)} - D_{(t,t+1)} + I_{(t,t+1)} - E_{(t,t+1)} + RE_{(t,t+1)} + \Delta NPR_{(t,t+1)} + \Delta Ninter_{(t,t+1)} + \Delta Nintra_{(t,t+1)}$$

where, for each CD:

(t,t+1)	=	interval between time *t* and *t+1*;
$P_{(t+1)}$	=	estimates of population at time *t+1*;
$P_{(t)}$	=	base population at time *t* (census counts adjusted for net census undercoverage or the most recent estimate);
B	=	number of births;
D	=	number of deaths;
I	=	number of immigrants;
E	=	number of emigrants;
RE	=	number of returning emigrants;
ΔNPR	=	net non-permanent residents;
ΔNinter	=	net interprovincial migration;
ΔNintra	=	net intraprovincial migration;

To ensure concordance between the subprovincial estimates and the provincial and territorial estimates by age and gender, two-way raking is used.

For more information, contact the Centre for Demography client services:

statcan.demography-demographie.statcan@canada.ca