

Open Building Population Layer for Canada (v3)

Technical Document (2026-04-21)

Table of Contents

| | |
|--|----|
| Introduction | 1 |
| What's new in v3 | 4 |
| Methodology | 5 |
| Data dictionary | 7 |
| Quality Flags | 8 |
| Download | 10 |
| QA/QC | 13 |
| Limitations and potential improvements | 14 |
| License and attribution | 15 |

Note

The Open Building Population Layer (Canada) is a personal project by Maxim Fortin. It is not an official Government of Canada data source for building footprints or population distribution. This layer is intended for high-level analysis and approximation, not for determining the population of a specific building or address. No warranty is given as to the accuracy or completeness of the information provided.

Introduction

How might 2021 census population be distributed across buildings in Canada? This dataset provides a building-level population approximation derived from census counts and building footprints. Population density products have long been used to quantify demographic information and to assess relationships with hazards, ecosystems, human health and infrastructure — but proprietary datasets can be expensive and public datasets lack the resolution required for most analyses. This open layer fills that gap, and both the data and code can be downloaded and used freely under open licenses.

The Open Building Population Layer for Canada was first released in 2022 (v1), then updated in 2024 (v2). Version 3 (2026) brings three major upgrades: approximately 13.7 million building footprints from Public Safety Canada's Canada Structures dataset, population allocation at the Dissemination Block level (498,547 units) using 2021 Census data, and an orphan rescue mechanism that recovers population from Dissemination Blocks with no buildings. Both Points and Footprints layers are provided in the EPSG:3979 coordinate reference system.



Open Building Population Layer (Canada) in Victoria, BC

Explore the data live: this interactive map shows a sample area in Ottawa, Canada. Toggle between building points and footprints, or switch the basemap, using the layer control in the top-right corner.



Interactive map sample: Sample area in Ottawa, ON

Intersecting building population estimates with geographic zones or features enables rapid analysis for hazard assessment, preliminary planning and other applications. For exploratory screening and preliminary analysis, the layer can help approximate questions like:

- How many people live in a floodplain?
- How many residents would a new transit line serve?
- What population is exposed to air pollution from an industrial facility?
- How many people would be affected by a planned power outage?
- What is the catchment population for a hospital or grocery store?



Left: 2008 Saint John River flood extents (Floods in Canada Archive, 2026); Right: 100, 500 and 1000 meters buffers on a new fictional transit line in Ottawa, ON

These examples are intended for exploratory and high-level analysis, not for estimating the true population of a specific building or for decision-making that requires site-level accuracy.

What's new in v3

Version 3 moves to Dissemination Blocks for a nearly ninefold increase in resolution over v2's Dissemination Areas, introduces an orphan rescue mechanism to redistribute population from Dissemination Blocks with census population but zero buildings, and closes the conservation gap from -1.626% to +0.069%.

| Aspect | v2 (2024) | v3 (2026) |
|---------------------|-----------------------------|--|
| Primary census unit | Dissemination Area (57,932) | Dissemination Block (498,547) |
| Fallback chain | DA → CSD → CD → PT | DB → DA → CSD → CD |
| Building footprints | Microsoft Bing (12.3M) | Public Safety Canada Structures (13.7M) |
| Orphan rescue | None | Population from 0-building DBs redistributed |

| Aspect | v2 (2024) | v3 (2026) |
|-----------------------------|------------|---|
| Quality flags | None | QF_O (orphan pop exceeds DB pop for that building), QF_V (implausible pop/height) |
| Population conservation gap | -1.626% | +0.069% |
| Output precision | Full float | 3 decimal places |
| File format | GeoPackage | GeoPackage & GeoParquet |

Methodology

The layer combines two open data sources:

1. **[Public Safety Canada — Canada Structures building footprints](#)**: Approximately 13.7 million building footprints from the Canada Structures dataset, freely available under the Open Government License (OGL). Each footprint includes a building height estimate where available. The product is provided in vector format created by Public Safety Canada (PS). This dataset combines and harmonizes different building footprint sources into one (Canada Open Database of Buildings, OpenStreetMap and Microsoft Bing Canada Building Footprints), with a notable increase in coverage compared to the Microsoft Bing Canada dataset previously used in v1 and v2 (+1.4 million buildings).
2. **[2021 Canada Census](#)**: Population distribution at the smallest available census geographical units, in a hierarchical fallback chain:
 3. Dissemination Blocks - DB (498,547 units)
 4. Dissemination Areas - DA (57,932 units)
 5. Census Subdivisions - CSD (5,161 units)
 6. Census Divisions - CD (293 units)

The building population is estimated through the following pipeline:

1. Load census geospatial data at the DB/DA/CSD/CD hierarchy and reproject to EPSG:3979
2. Validate and repair census geometries
3. Load Canada Structures building footprints and reproject to EPSG:3979
4. Perform spatial joins of building representative points to each census level
5. Calculate population-per-building ratios, with fallbacks when no results are available at the detailed level:
6. $DB_RATIO = DB_POP / COUNT_DB$

7. $DA_RATIO = DA_POP / COUNT_DA$ (fallback)
8. $CSD_RATIO = CSD_POP / COUNT_CSD$ (fallback)
9. $CD_RATIO = CD_POP / COUNT_CD$ (fallback)
10. Orphan population rescue: redistribute population from DBs with census population > 0 but zero buildings to buildings in the same DA
11. Assign: $BLDG_POP = best_available_ratio + ORPHAN_RATIO$, rounded to 3 decimal places
12. Assign flags documenting potential data quality concerns
13. Output: Points and Footprints layers per province and nationally

HEIGHT is included as an attribute from Canada Structures and used for QA/QC only; it is not used to estimate BLDG_POP.

The orphan population rescue mechanism introduced in v3 is important. Without it, census population in Dissemination Blocks with zero building footprints would be entirely excluded from the dataset (approx. 290,000 people nationally). Moving to finer census granularity (DBs) would actually result in *more* missed population compared to the coarser DA approach if this complementary process was not introduced (since it is much more likely to find 0 buildings in a smaller DB, than it is in the larger DA). The orphan rescue enables more detailed census data while accepting that some population estimates become partially dependent on redistribution from neighboring areas.

The entire pipeline is automated in Python using GeoPandas and Pyogrio. The code is available on [Github](#).

The image below shows an example of the v3 process in action at the Dissemination Block level.



Close-up view with population labels and Dissemination Block census boundaries in Chilliwack, BC

Data dictionary

Data fields provided with this data layer are described below. The main data field is "BLDG_POP" which provides the best available population estimate for each building.

View data dictionary (16 fields)

| Field | Type | Description |
|----------|-------|---|
| BLDG_ID | int | Building index unique identifier, consistent with Canada Structures input dataset |
| HEIGHT | float | Building height in metres (3 decimal places) |
| DBUID | str | Dissemination Block unique identifier |
| DB_RATIO | float | Population per building at DB level (3 decimal places) |
| DAUID | str | Dissemination Area unique identifier |
| DA_RATIO | float | Population per building at DA level (3 decimal places) |
| CSDUID | str | Census Subdivision unique identifier |

| Field | Type | Description |
|------------------|-------|--|
| CSD_RATIO | float | Population per building at CSD level (3 decimal places) |
| CDUID | str | Census Division unique identifier |
| CD_RATIO | float | Population per building at CD level (3 decimal places) |
| ORPHAN_RATIO | float | Bonus population from orphaned DBs located in the same DA (3 decimal places) |
| HAS_ORPHAN_BONUS | int | 1 if building receives orphan bonus, 0 otherwise |
| BLDG_POP | float | Estimated building population (3 decimal places). Calculated from the best available census ratio (DB, DA, CSD, or CD), plus ORPHAN_RATIO where applicable. |
| POP_SOURCE | str | Ratio used: DB_RATIO, DA_RATIO, CSD_RATIO, CD_RATIO, or NONE |
| QF_O | int | 1 if (ORPHAN_RATIO > DB_RATIO) AND (DB_RATIO > 0), 0 otherwise. Flags buildings where the orphan bonus exceeds the DB allocation. |
| QF_V | int | 1 if BLDG_POP > 100 AND 0 < HEIGHT < 5m, 0 otherwise. Flags physically implausible combinations of high population and low building height. |

Quality Flags

Two boolean columns identify buildings with potential data quality concerns.

QF_O — Orphan Bonus Exceeds DB Allocation

`QF_O = 1` when `ORPHAN_RATIO > DB_RATIO AND DB_RATIO > 0`.

The orphan rescue mechanism redistributes population from Dissemination Blocks with census population but zero building footprints. This flag identifies buildings where the orphan bonus exceeds the building's own DB allocation — meaning most of the population estimate comes from redistribution rather than the building's local DB. **24,139 buildings** (0.18%) are flagged with QF_O.

Why does this happen? The mismatch between census population and available building footprints has several causes: - **Incomplete footprint coverage** — the most common cause. Some areas have census population but missing building footprints. - **Temporal misalignment** — census data is from 2021, while footprint data was compiled from multiple sources collected at different times. Areas with very recent land and housing developments tend to have more flags, as census population exists for these areas but buildings may not

yet be captured in the footprint data. - **Source compilation gaps** — Canada Structures combines footprints from multiple datasets with varying completeness.

QF_V — Implausible Population for Building Height

`QF_V = 1` when `BLDG_POP > 100` AND `0 < HEIGHT < 5`.

This flag identifies buildings with physically implausible combinations: over 100 residents but confirmed heights under 5 metres. **830 buildings** (0.006%) are flagged with QF_V.

Why does this happen? - Height data errors — some buildings have incorrect height values, visible when compared against satellite imagery. - **Small dissemination blocks with high population density** — a DB with only 4-5 high-rise buildings may distribute high population to small auxiliary structures (garages, sheds) nearby.

Only buildings with confirmed heights are flagged (`HEIGHT = 0` or `NaN` excluded). Note: 66.5% of buildings in the footprint dataset have height data, and there is high data coverage for urban areas since height was originally estimated in Canada Structures from national lidar data holdings.

| Flag | Buildings | % of Dataset |
|--------------------|------------|--------------|
| QF_O only | 24,032 | 0.175% |
| QF_V only | 723 | 0.005% |
| Both QF_O and QF_V | 107 | 0.001% |
| Neither | 13,735,892 | 99.819% |

Flags per Province/Territory

| Province/Territory | QF_O | QF_V | % Flag Instances |
|--------------------|-------|------|------------------|
| AB | 3,701 | 35 | 0.196% |
| BC | 1,105 | 97 | 0.078% |
| MB | 863 | 15 | 0.122% |
| NB | 10 | 3 | 0.003% |
| NL | 668 | 0 | 0.233% |
| NS | 29 | 2 | 0.006% |

| Province/Territory | QF_O | QF_V | % Flag Instances |
|--------------------|--------|------|------------------|
| NT | 0 | 0 | 0.000% |
| NU | 0 | 0 | 0.000% |
| ON | 13,833 | 585 | 0.305% |
| PE | 0 | 0 | 0.000% |
| QC | 3,332 | 91 | 0.123% |
| SK | 448 | 2 | 0.061% |
| YT | 150 | 0 | 0.936% |

Download

The following files are available for download in zipped [GeoPackage](#) format and [GeoParquet](#) format (EPSG:3979 coordinate reference system).

File packages are provided for each province and territory (GeoPackage), along with a complete country-wide package (GeoPackage and GeoParquet).

Building footprints

| Province/Territory | Size |
|---|--------|
| Alberta | 245 MB |
| British Columbia | 206 MB |
| Manitoba | 89 MB |
| New Brunswick | 50 MB |
| Newfoundland and Labrador | 33 MB |
| Northwest Territories | 3.4 MB |
| Nova Scotia | 60 MB |
| Nunavut | 1.2 MB |
| Ontario | 641 MB |

| Province/Territory | Size |
|---|---------------|
| Prince Edward Island | 9.5 MB |
| Quebec | 354 MB |
| Saskatchewan | 93 MB |
| Yukon | 1.8 MB |
| Canada (all) - Zipped GeoPackage | 1.8 GB |
| Canada (all) - GeoParquet | 1.4 GB |

Building points

| Province/Territory | Size |
|---|---------------|
| Alberta | 102 MB |
| British Columbia | 79 MB |
| Manitoba | 40 MB |
| New Brunswick | 23 MB |
| Newfoundland and Labrador | 16 MB |
| Northwest Territories | 892 KB |
| Nova Scotia | 26 MB |
| Nunavut | 476 KB |
| Ontario | 238 MB |
| Prince Edward Island | 4.5 MB |
| Quebec | 160 MB |
| Saskatchewan | 42 MB |
| Yukon | 740 KB |
| Canada (all) - Zipped GeoPackage | 723 MB |
| Canada (all) - GeoParquet | 332 MB |

QA/QC

QA/QC focused on internal consistency by comparing aggregated BLDG_POP totals against 2021 Census Dissemination Block population totals. This checks population conservation and fallback behaviour, but does not independently validate population at the individual-building level. Additional quality control was also done to validate features, geometries and results by comparison between v2 and v3 layers, as well as deep-dives on the behaviour of the orphan rescue implementation.

The overall population conservation gap is +0.069% at the national level (computed as the sum of all provincial differences divided by the total census population). Some provinces show slight over-allocation, others under-allocation. This is primarily due to privacy suppressed DBs (843 nationally) receiving population via fallback ratios rather than DB_RATIO, plus minor rounding effects. The orphan population redistribution itself conserves population exactly and is not the source of these variations.

Key metrics: - Approximately 99.8% of buildings receive population via DB_RATIO (Dissemination Block level) - 1,808,449 buildings receive an orphan bonus (13.1%) - 843 Dissemination Blocks have suppressed (null) population data due to privacy reasons - these receive population estimates from fallback ratios - 24,139 buildings (0.18%) flagged with QF_O — orphan bonus exceeds DB allocation - 830 buildings (0.006%) flagged with QF_V — implausible population/height combination - Combined: 24,862 buildings (0.18%) have at least one quality flag

View per-province comparison table

| Province & territory | Census DB population 2021 | Sum of BLDG_POP | Difference | Difference (%) |
|----------------------|---------------------------|-----------------|------------|----------------|
| AB | 4,262,635 | 4,269,294.6 | +6,659.6 | +0.156 |
| BC | 5,000,874 | 5,002,649.5 | +1,775.5 | +0.036 |
| MB | 1,342,153 | 1,346,313.6 | +4,160.6 | +0.310 |
| NB | 775,610 | 776,061.0 | +451.0 | +0.058 |
| NL | 510,545 | 511,272.7 | +727.7 | +0.143 |
| NS | 969,383 | 970,161.1 | +778.1 | +0.080 |
| NT | 41,070 | 41,075.9 | +5.9 | +0.014 |
| NU | 36,858 | 36,876.4 | +18.4 | +0.050 |
| ON | 14,223,942 | 14,222,563.0 | -1,379.0 | -0.010 |
| PE | 154,331 | 154,466.2 | +135.2 | +0.088 |

| Province & territory | Census DB population 2021 | Sum of BLDG_POP | Difference | Difference (%) |
|----------------------|---------------------------|---------------------|------------------|----------------|
| QC | 8,501,833 | 8,514,167.0 | +12,334.0 | +0.145 |
| SK | 1,132,505 | 1,132,249.1 | -255.9 | -0.023 |
| YT | 40,232 | 40,191.0 | -41.0 | -0.102 |
| Total | 36,991,971 | 37,017,340.9 | +25,369.9 | +0.069 |

Limitations and potential improvements

Found an error, or have an idea for improvement? Please send me a message on [LinkedIn](#) or using the [contact form](#) in the site menu.

Population is averaged per building. No differentiation is made between residential, commercial, industrial, or institutional buildings — population is distributed using population-per-building ratios, primarily at the Dissemination Block level, with fallback to coarser census units where needed. **This layer is suitable for high-level assessments and approximations, but not for determining the population of a specific building or address.**

Northern and remote areas may have limited footprint coverage. The Canada Structures dataset has gaps in Yukon, the Northwest Territories, and particularly Nunavut — use with caution in these regions. Building footprint age varies across sources, and height data quality varies by source; neither should be relied upon for site-specific analysis. Additionally, recent constructions or changes in land use may affect building footprint gaps.

Census data is from 2021. A new census should be conducted in 2026 by Statistics Canada, and detailed census data profiles will likely be available around 2028. There are 843 Dissemination Blocks with privacy-suppressed population that receive values through fallback ratios at coarser census levels.

License and attribution

The Open Building Population Layer (Canada) v3 has been produced by Maxim Fortin as open and free data. A summary of this webpage is available as a [PDF document](#).

Dataset license

The dataset is released under the [Open Database License \(ODbL\) v1.0](#). This license allows you to freely use, distribute, and modify the database, and produce works from it, as long as you attribute the source and share any adapted database under the same license.

When using the dataset, please attribute as follows: *Fortin, Maxim (2026). Open Building Population Layer — Canada v3.* <https://maximfortin.com/project/obpl-ca-v3/>

This database incorporates data from the Open Database of Buildings (ODB), OpenStreetMap (OSM) buildings, and Microsoft Building Footprints (MSB) via Canada Structures (Public Safety Canada), all released under their respective open licenses. Census data is from Statistics Canada under the Open Government License – Canada. All upstream attributions must be preserved. A comprehensive list of all data providers and their attributions for the Open Database of Buildings (ODB) is available in the [attribution table \(CSV\)](#).

No warranty is given as to the accuracy or completeness of the information provided. While the dataset provides estimates of population for each building in Canada, please note that these are derived approximations and do not represent exact population counts. Use caution when interpreting and utilizing the data for decision-making purposes.

Code license

The code is released under the [Apache 2.0 License](#). This license allows you to freely use, distribute, and modify the code for both commercial and non-commercial purposes, with limited liability and warranty.

When using the code, please attribute as follows: *Fortin, Maxim (2026). Python code for the Open Building Population Layer — Canada v3.* <https://maximfortin.com/project/obpl-ca-v3/>

No warranty is given as to the accuracy or completeness of the information provided. We provide no warranties and assume no responsibility for any liabilities associated with the use of this code.