Catalogue no. 13-20-0001
ISBN 978-0-660-79827-1

Canadian Food Environment Dataset: User Guide

Release date: December 19, 2025

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

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, the Agency has developed standards of service which its employees observe in serving its clients. 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."

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

Canadian Food Environment Dataset: User Guide

1 Background and Overview of Dataset

1.1 Background

The retail food environment is a modifiable element of the built environment that holds potential to influence the diet of Canadians at the population level. Unfavourable neighbourhood retail food environments are a target for intervention because they can contribute to poor diets (Roberto et al. 2015). In Canada, neighbourhoods with easy access to healthier food options or limited access to less healthy food options have been associated with better indicators of health, such as better diet (Minaker et al. 2013), lower body mass index (BMI) (Stevenson et al. 2019), and lower levels of Type 2 diabetes (Polsky et al. 2016). However, measurement error in the assessment of food environments, as well as of diet and health outcomes can make it difficult to draw meaningful conclusions from past work.

Until recently, there was no high-quality Canada-wide dataset of businesses that could be used to create food environment measures that are accessible, free to use, and relevant for public health. In 2021, Statistics Canada created the Canadian Food Environment Dataset (Stevenson et al. 2022), providing accessible and high-quality food environment access measures at the national level, as well as advancing the quality and accuracy of food environment analysis and data linkage—including national health surveys, health administrative data and investigator-led cohort studies—by researchers and public health communities (Stevenson et al. 2023; Li 2025).

The purpose of this document is to describe the development and user considerations for the Can-FED v2, a pan-Canadian dataset of retail food environment measures at the dissemination area (DA) level based on food outlet data from the 2024 Statistics Canada Business Register. Data are accessible through Statistics Canada's Research Data Centre Network and available for use by researchers and other potential users, including academics and students, and practitioners from both the public and private sector.

This document is mainly focused on the Researcher Can-FED, which includes densities of 20 retail food outlet types measured as counts per square kilometer within street network buffer areas and two relative food outlet proportion measures, and comprises two main sections. First, Can-FED data available in the Statistics Canada's Research Data Centre Network are described, including the classifications of outlet types, field names, and measures (Section 1). Second, the methodology for deriving Can-FED is detailed in Section 2.

A General Use Can-FED which includes categorical neighbourhood food access is available to download online.

1.2 About Can-FED

The Canadian Food Environment Dataset (Can-FED) is a geographic-based set of measures that represents the food environment of Canadian communities. The primary envisioned use of the Can-FED is research and analysis of the relationship between the local food environment, dietary intake and health outcomes. By using geography conversion tools such as the Standard Geographical Classification (Statistics Canada 2024b), it is possible to link Can-FED measures to individual-level health data from national-level survey platforms (e.g., the Canadian Health Measures Survey and the Canadian Community Health Survey) or to investigator-led cohorts.

A diversity of food categories was chosen to represent an inclusive set of different and overarching food outlet types, such as healthy food retailers (e.g., supermarkets, grocery stores), convenient food options (e.g., corner stores, fast-food outlets), and independent and specialty retailers (e.g., bakeries, fish markets, coffee shops), among other options.

1.2.1 Geographic unit of analysis

Retail food environment access measures were calculated in Valhalla version 3.4.0 (Various authors 2025) using network buffers around the population-weighted centroids (also known as representative points) of 57,936 dissemination areas (DAs) in Canada. Population-weighted centroids are weighted using the population mean center and provide a better indication of where the majority of populations live and the environments they interact with, compared with the administrative boundary or actual mean center of a DA.

DAs are the smallest standard geographic units defined by Statistics Canada for which census data are disseminated across Canada and have populations of 400 to 700 people (Statistics Canada 2022a). DA-level food environments capture a good indication of neighbourhood-level food access, as opposed to larger areas such as census tract boundaries, which can help with the identification of target areas for intervention strategies.

1.2.2 Reference years

This second cycle release of the Can-FED represents the Canadian food environment in August 2024. Road network and DA data were used from the most recent available years and correspond with 2024 and 2021 geographies, respectively.

1.2.3 File format and naming convention

The Can-FED is available as both a Comma Separated Values file (.csv) and parquet file (.parquet)—either as CanFED_1000m (using the 1-km buffer) and CanFED_3000m (using the 3-km buffer).

1.2.4 Dataset completeness

Of the total 57,936 DAs in Canada, 28 DAs were not included due to errors that arose when calculating their network buffers. The errors were due to the inability to locate a road network from the centroid of the DA (i.e., DAs with boundaries that follow features other than roads, such as water features or power transmission lines or DAs that have lower population counts, including 0). For the 1-km and 3-km network buffer, 99.95% of DAs in Canada were included. Appendix 1 details both the number and percentage of DAs that were excluded for each province and territory.

1.3 Key Changes Between Can-FED Cycles

The aim, scope and methodology of the second cycle of Can-FED (v2) was designed to be as similar to the first cycle (v1) as possible to maintain comparability. While the original methodology is no longer reproducible in its entirety due to software deprecation, this second cycle provides an open source-based framework—in both the road network buffer computation (OpenStreetMap, Valhalla) and the process of classifying and geolocating businesses (Python)—to allow for greater ease of reproducibility in the future.

The NAICS codes used to sort businesses have been updated to the 2022 NAICS version 1.0 (Statistics Canada 2022b), the most recent version as of the implementation of Can-FED cycle 2.

Additionally, some small changes were made to the filtering of business names into specific NAICS categories.

1.4 Can-FED Food Outlet Data

1.4.1 Source data

Food outlet data were extracted from the Statistics Canada Business Register (Statistics Canada 2024a), a central repository of baseline information on businesses operating in Canada. Information on businesses is compiled from mandatory tax data collected by the Canada Revenue Agency. All outlets are consistently classified with the 2022 North American Industry Classification System (NAICS) (Statistics Canada 2022b) that identifies the primary function of a business. Ongoing quality evaluation and assurance is performed by staff at Statistics Canada.

1.4.2 Data extraction & Construct definition and classification

Food outlets were extracted from the Business Register according to their NAICS code. Food outlets were extracted if they had an operating outlet as of August 2024 and if the level 5 (six-digit) NAICS code associated with the outlet indicated that it was a type of retail food outlet. These were then classified based primarily on NAICS code and, for some food outlet types, further classified using a name-based assignment method.

Businesses that needed to be disaggregated from their assigned NAICS code were further classified based on the outlet name. For example, outlets with NAICS code 722512 (“Limited-service eating places”) were classified into one of either “fast-food outlets”, “cafés”, or “other limited-service eating places.” Regular expression

searches were used on both the English and French business name, as well as the NAICS description, to classify businesses by certain keywords; for example, businesses with “burger” in their name were sorted into the “fast-food outlets” category, and those with “café” into the “café” category. All remaining businesses were classified into the “other limited-service eating place” category.

Table 1 provides an overview of the categorization methods used for each food outlet type.

Table 1
NAICS codes and categorization methods

Business Category	NAICS Code (2022)	Assignment Method
Warehouse clubs & supercenters	455211	Club brand name
Grocery stores	445110	Not a chain grocery store brand name
Superstores	445110	Chain grocery store brand name
Convenience stores	445131	N/A
Gas station stores	457110	N/A
Bakeries	311811	N/A
Fruit & vegetable markets	445230	N/A
Meat markets	445240	N/A
Fish markets	445250	N/A
Specialty retailers	445298	N/A
Confectionary & nut retailers	445292	N/A
Restaurants	722511	N/A
Limited service eating places (fast food)	722512	Fast food brand name
Limited service eating places (cafes)	722512	Café brand name
Limited service eating places (other)	722512	Not a fast food or café brand name
Bars	722410	N/A
Liquor stores	455320	N/A
Miscellaneous	455219	N/A
Pharmacies	456110	Pharmacy brand name
Cannabis stores	459993	N/A

1.5 Can-FED Access Metrics and Field Directory

Two buffer sizes were generated from the population-weighted centroid of each DA: 1-km and 3-km network buffers. The 1-km buffers represent a 10- to 15-minute walking distance from the DA to a food outlet (Papas et al. 2007). The 3-km buffers capture greater distances that may be travelled in more rural DAs, and also represent larger actual or perceived food environment boundaries.

Using a spatial query, food outlets were joined to the buffers and subsequently tabulated by each food outlet type. Business density[1] and two relative measures (mRFEI and Rmix—see the next Section) were calculated using these counts. The full list of names, labels and brief definitions of the variables are provided in Table 2.

1. Defined as the number of businesses of a given type per square kilometer.

Table 2
Can-FED v2 metadata and derivation sources

Field name	Full name	Definition	Derivation sources
Province or Territory	Province or territory	Major political and geographical unit of Canada	Business Register
DAUID	Dissemination area identifier	Dissemination area identifier	Statistics Canada standard geographies
den_01_warehouse_clubs_supercenters	Density of warehouse clubs and supercenters	Number of stores per square kilometer of buffer	Business Register
den_02_grocery_stores	Density of grocery stores	Number of stores per square kilometer of buffer	Business Register
den_03_superstores	Density of superstores	Number of stores per square kilometer of buffer	Business Register
den_04_convenience_stores	Density of convenience stores	Number of stores per square kilometer of buffer	Business Register
den_05_gas_station_stores	Density of gas station stores	Number of stores per square kilometer of buffer	Business Register
den_06_bakeries	Density of bakeries	Number of stores per square kilometer of buffer	Business Register
den_07_fruit_vegetable_markets	Density of fruit and vegetable markets	Number of stores per square kilometer of buffer	Business Register
den_08_meat_markets	Density of meat markets	Number of stores per square kilometer of buffer	Business Register
den_09_fish_markets	Density of fish markets	Number of stores per square kilometer of buffer	Business Register
den_10_specialty_retailers	Density of specialty retailers	Number of stores per square kilometer of buffer	Business Register
den_11_confectionery_nut_retailers	Density of confectionary and nut retailers	Number of stores per square kilometer of buffer	Business Register
den_12_restaurants	Density of restaurants	Number of stores per square kilometer of buffer	Business Register
den_13_limited_service_fast_food	Density of fast food eating places	Number of stores per square kilometer of buffer	Business Register
den_14_limited_service_other	Density of other limited service eating places	Number of stores per square kilometer of buffer	Business Register
den_15_limited_service_cafe	Density of cafes	Number of stores per square kilometer of buffer	Business Register
den_16_bars	Density of bars	Number of stores per square kilometer of buffer	Business Register
den_17_liquor_stores	Density of liquor stores	Number of stores per square kilometer of buffer	Business Register
den_18_misc_stores	Density of miscellaneous stores	Number of stores per square kilometer of buffer	Business Register
den_19_pharmacies	Density of pharmacies	Number of stores per square kilometer of buffer	Business Register
den_20_cannabis_retailers	Density of cannabis retailers	Number of stores per square kilometer of buffer	Business Register
Rmix	Fast food restaurant mix	The proportion of fast food eating places that operate within each buffer	Business Register
mRFEI	Modified retail food environment	The proportion of healthy outlets that operate within each buffer	Business Register

1.5.1 Food environment variables defined: Modified Retail Food Environment Index (mRFEI) and fast food restaurant mix (Rmix)

Relative food environment variables were derived to represent the proportion of healthy and less healthy food retailers for each DA. The modified Retail Food Environment Index (mRFEI) calculates the proportion of food outlets that sell a wide selection of fresh and nutritious food that operate within each buffer, defined as

$$mRFEI = \frac{\text{Supermarket chains} + \text{Grocery stores} + \text{Fruit \& vegetable markets}}{\text{Supermarket chains} + \text{Grocery stores} + \text{Fruit \& vegetable markets} + \text{Fast foot outlets} + \text{Convenience stores} + \text{Gas station stores}}$$

The fast-food restaurant mix (Rmix) calculates the proportion of fast-food restaurants relative to both fast-food and full-service restaurants, defined as

$$Rmix = \frac{\text{Fast food outlets}}{\text{Fast foot outlets} + \text{Restaurants}}$$

Note that in both mRFEI and Rmix equations, only those businesses sorted under the “fast food outlet” (NAICS code 722512) were included in these calculations.

1.5.2 General-use Can-FED access measures

The general use Can-FED file includes the densities of grocery stores, superstores, convenience stores, fruit and vegetable markets, restaurants, fast food limited service eating places, and the two relative density measures (mRFEI & Rmix).

Five categories were created for each absolute density and relative density measure on the Can-FED general use file: zero, and groups 1 to 4. All non-zero densities were sorted according to k-medians clustering thresholds (SAP HANA, n.d.), each density falling into one of four categories based on the threshold cutoffs for that business type, with 1 being the lowest and 4 the highest. Business densities of zero were assigned the category of 0, whereas invalid mRFEI & Rmix values (i.e., having a denominator of zero) were assigned the “..” value.

A k-medians approach finds k-number of cluster centers to minimize within-group variation of observations (DAs) and to maximize between-group variation, and are more resistant to outliers than k-means clustering. Zeros were isolated as a separate category because zero is a unique and meaningful value for retail food establishments that represents no access to a given outlet type in an area. Any value above zero indicates that an outlet type is present in the buffer (i.e., there is some access). This represents two different types of food environments that would not have been captured if zero counts were included in the k-median groupings.

1.6 Guidelines for Use

1.6.1 Caution for analysis of certain DAs in rural areas

1-km and 3-km buffers may be too small to capture some residents’ food shopping behaviour, particularly in rural DAs. Since the buffer is drawn around the population-weighted centroid of each DA, the buffers in rural areas tend to be closer to the population center, which is often where food stores operate. This may mean that the access to a food store is overestimated for those residents who live further from the population center.

2 Protocol for Deriving the Can-FED Measures

2.1 Software

The Can-FED v2 dataset was produced using open source software where possible:

- Road data: OpenStreetMap
- Buffer network: Valhalla
- Food data: Business Register (not open source)
- Processing & Analysis: R & Python

2.2 Acquiring Data & Preparation

DA representative points were taken from the Statistics Canada Geographic Attribute File, which contains geographic data at the dissemination block level. The geographical projection systems used were EPSG:3347 (Lambert projection) and EPSG:4326 (Mercator projection).

Food outlets were extracted from the Statistics Canada Business Register if they had an operating outlet as of August 2024 and matched the relevant NAICS codes for food outlets (see Table 2).

Data source descriptions and reference URLs

Data Sources and Tools: Representative Points (population-weighted centroids)

Description

The representative points for dissemination areas (DAs) are weighted using the population mean centre. DAs are the smallest standard geographic area for which census data are disseminated across Canada and have populations of 400 to 700 people.

Reference URL

https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/az/definition-eng.cfm?ID=geo040

Download URL

Geographic Attribute File https://www150.statcan.gc.ca/n1/en/catalogue/92-151-X

File Format

Shapefile (.shp)

Data Sources and Tools: Business Register (BR)

Description

The BR is a central repository of baseline information on businesses and institutions operating in Canada. All outlets are consistently classified with a NAICS code that identifies the primary function of a business.

Reference URL

https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=1105

Data Sources and Tools: OpenStreetMap

Description

OpenStreetMap is an open source database of map and road network information; contributors use aerial imagery, GPS devices, and low-tech field maps to verify that OSM is accurate and up to date.

Reference URL

https://www.openstreetmap.org/about

Data Sources and Tools: Valhalla

Description

Valhalla is an open source routing engine and accompanying libraries for use with OpenStreetMap data.

Reference URL

https://github.com/valhalla/valhalla

2.3 Code Overview

This section aims to give a brief overview of the process used to produce the Can-FED v2 dataset, as the update to open source software represents the biggest change from v1 to v2. Where possible, the v1 methodology was replicated using these methods.

2.3.1 Road network buffer

One- and three-kilometer buffers from each DA centroid were calculated using Valhalla, drawing from OpenStreetMap data for road networks.

2.3.2 Querying the Business Register

All businesses matching the relevant NAICS codes and which are considered in operation were extracted from the Business Register, as well as necessary information including operating name and address longitude and latitude.

2.3.3 Data processing

Certain business types were then further sorted into NAICS codes sub-categories based on keyword strings found in the business name or NAICS description.

Using Statistics Canada geography concordance data to join relevant geography data to each business, all businesses were spatially joined with the 1- and 3-km road network buffers to determine which businesses lay in each buffer in each DA. The resulting geographical data was joined according to the Lambert projection for all of Canada, since, given the road network buffers are small areas, province-by-province projections would have minimal to no effect on the business locations relative to the buffers.

The business densities, as well as the mRFEI and Rmix metrics, were then calculated for each buffer for each DA.

Appendix 1

Table A1
DAs with invalid buffer networks by province

Province or Territory	Total DAs	Invalid DAs	Excluded DAs
	number		percent
Newfoundland & Labrador	1,082	3	0.3
Prince Edward Island	319	0	0.0
Nova Scotia	1,670	0	0.0
New Brunswick	1,470	0	0.0
Quebec	13,806	8	0.1
Ontario	20,468	0	0.0
Manitoba	2,223	1	0.0
Saskatchewan	2,625	1	0.0
Alberta	6,203	2	0.0
British Columbia	7,848	1	0.0
Yukon	74	1	1.4
Northwest Territories	98	7	7.1
Nunavut	50	4	8.0
Total	57,936	28	0.0

References

Li, Yannan. 2025. “Characterizing the Food Environment and Its Association with Ultra-Processed Food Intake in British Columbia, Canada.” PhD thesis, University of British Columbia.

Minaker, Leia M, Kim D Raine, T Cameron Wild, Candace IJ Nykiforuk, Mary E Thompson, and Lawrence D Frank. 2013. “Objective Food Environments and Health Outcomes.” *American Journal of Preventive Medicine* 45 (3): 289–96.

Papas, Mia A, Anthony J Alberg, Reid Ewing, Kathy J Helzlsouer, Tiffany L Gary, and Ann C Klassen. 2007. “The Built Environment and Obesity.” *Epidemiologic Reviews* 29 (1): 129–43.

Polsky, Jane Y, Rahim Moineddin, Richard H Glazier, James R Dunn, and Gillian L Booth. 2016. “Relative and Absolute Availability of Fast-Food Restaurants in Relation to the Development of Diabetes: A Population-Based Cohort Study.” *Canadian Journal of Public Health* 107 (Suppl 1): eS27–33.

Roberto, Christina A, Boyd Swinburn, Corinna Hawkes, Terry TK Huang, Sergio A Costa, Marice Ashe, Lindsey Zwicker, John H Cawley, and Kelly D Brownell. 2015. “Patchy Progress on Obesity Prevention: Emerging Examples, Entrenched Barriers, and New Thinking.” *The Lancet* 385 (9985): 2400–2409.

SAP HANA. n.d. “Clustering Algorithms - K-Medians.”

Statistics Canada. 2022a. “Illustrated Glossary - Dissemination area (DA).”

Statistics Canada. 2022b. *North American Industry Classification System (NAICS) Canada 2022 Version 1.0*.

Statistics Canada. 2024a. “Business Register (BR).”

Statistics Canada 2024b. “Standard Geographical Classification.”

Stevenson, Andrew C, Anne-Sophie Brazeau, Kaberi Dasgupta, and Nancy A Ross. 2019. “Evidence Synthesis-Neighbourhood Retail Food Outlet Access, Diet and Body Mass Index in Canada: A Systematic Review.” *Health Promotion and Chronic Disease Prevention in Canada: Research, Policy and Practice* 39 (10): 261.

Stevenson, Andrew C, Rachel C Colley, Kaberi Dasgupta, Leia M Minaker, Mylene Riva, Michael J Widener, and Nancy A Ross. 2023. “Neighborhood Fast-Food Environments and Hypertension in Canadian Adults.” *American Journal of Preventive Medicine* 65 (4): 696–703.

Stevenson, Andrew C, Clara Kaufmann, Rachel C Colley, Leia M Minaker, Michael J Widener, Thomas Burgoine, Claudia Sanmartin, and Nancy A Ross. 2022. “A Pan-Canadian Dataset of Neighbourhood Retail Food Environment Measures Using Statistics Canada’s Business Register.” *Health Reports*.

Various authors. 2025. *Valhalla*.