

2021 Canadian Social Environment Typology User Guide

by Vijata Sharma, Rajendra Subedi and Nicole Aitken

Release date: March 19, 2026

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.

Published by authority of the Minister responsible for Statistics Canada

© His Majesty the King in Right of Canada, as represented by the Minister of Industry, 2026

Use of this publication is governed by the Statistics Canada [Open Licence Agreement](#).

An [HTML version](#) is also available.

Cette publication est aussi disponible en français.

Table of contents

1 Introduction	4
2 Background	4
3 Uses of the Canadian Social Environment Typology (CanSET)	5
4 Data used to create the CanSET	5
5 Methodology	6
5.1 Variable Selection	6
5.2 Number of Clusters	6
5.3 Cluster Characteristics and Uses.....	7
5.3.1 2016 and 2021 Cluster Comparability	8
5.3.3 Ten clusters solution.....	10
5.3.4 Twenty-clusters solution	12
6 How to integrate the CanSET with your data	16
7 Summary and Conclusion	16
Appendix A	17
References	19

2021 Canadian Social Environment Typology User Guide

by Vijata Sharma, Rajendra Subedi and Nicole Aitken

1 Introduction

The Canadian Social Environment Typology (CanSET) is a geographic classification tool which compares neighbourhoods across Canadian Census Metropolitan Areas and Census Agglomerations. It provides three levels of neighbourhood classifications based on combinations of 30 socioeconomic, demographic, ethnocultural and housing variables from the Census of population. Each social environment cluster is a group of similar dissemination areas and represents a unique neighbourhood type. The CanSET data comes with definitions of each neighbourhood type so that users can compare health and social outcomes by neighbourhood characteristics.

The CanSET was first developed using the 2016 data from the Census of the Population. A new version has now been created using the 2021 data.

The purpose of this user guide is to introduce the 2021 Canadian Social Environment Typology (CanSET) as a standardized way to classify Dissemination Areas (DAs) in Canada to study neighbourhood level social and health inequalities.

The objectives are to:

1. Explain how the 2021 CanSET was developed
2. Describe the concept and structure of DA clusters within CanSET
3. Provide guidance on how to use CanSET to analyze area-level inequalities

2 Background

In recent years, there has been a growing demand for relevant health information at the neighbourhood level. However, the absence of a standardized classification approach has made it difficult to compare and contrast small areas within a city or among the cities in Canada. To overcome this shortcoming, a Dissemination Area (DA) classification approach called the Canadian Social Environment Typology (CanSET) was developed. The first version of CanSET was developed using 2016 data from the Census of the Population. The new version of CanSET, which utilizes 2021 data from the Census of Population, is now available.

The CanSET is a hierarchical clustering of DAs within Canadian Census Metropolitan Areas (CMA)¹ and Census Agglomerations (CA).² The typology generated through this cluster analysis can be taken as a social unit of analysis that indicates the geographic distribution of different kinds of combinations of population characteristics throughout CMAs and CAs. In other word, each social environment cluster includes similar DAs scattered across Canada. Therefore, the CanSET is not defining neighbourhoods but defining areas of common social composition which may be entirely different than the neighbourhoods defined by different local authorities in Canada. The CanSET data can be used to better understand health inequalities in relation to the composition of sub-city geographies in Canada. The 2021 CanSET data should not be used for direct comparisons with the 2016 CanSET data.

This user guide is designed to facilitate the readers on how to use the new social environment clusters to understand social and health inequalities in more populous areas in Canada. The social environment clusters are developed by using 30 different socioeconomic, demographic, ethnocultural and housing variables from the 2021 Census aggregated at the DA level. All the variables for this product were taken directly from the 2021 Census of Population microdata except the population density variable, which was derived from a Census of Population 2021 profile table.

1. A CMA is a geographic unit with a total population of at least 100,000 of which 50,000 or more must live in the core based on adjusted data from the previous Census of Population Program.
2. A CA is a geographic unit with a core population of at least 10,000 based on adjusted data from the previous Census of Population Program.

3 Uses of the Canadian Social Environment Typology (CanSET)

The individual-level socioeconomic, demographic, ethnocultural and housing characteristics are useful to understand health and social outcomes but are often not available in the most commonly used datasets like surveys or administrative data. Therefore, area-based socioeconomic and demographic measures have been widely used to measure inequalities. Area-level analysis helps to understand whether living in a socioeconomically disadvantaged area is associated with additional health risk beyond the individual-level socioeconomic status.

The CanSET portrays the complex social composition of Canadian CMAs and CAs. While many previous studies focused only on marginalization or health inequalities, the CanSET enables a more comprehensive exploration as it includes a variety of dimensions of urban social environment and covers all populous areas of Canada. CanSET is the first measure to include all DAs from the CMAs and CAs simultaneously in a single typology rather than analyzing them separately for each CMA or CA. In addition, it includes the microdata from the Census of Population 2021 which covers all population of the study area. Therefore, the CanSET is very important for urban research as it allows comparisons within cities or between cities to set benchmarks and track progress in combating health and social inequalities.

For instance, the inequalities in COVID-19 mortality rates were studied using the 2016 CanSET classification tool at the 6-cluster level (Subedi & Aitken, 2022). This study showed the combined effect of neighbourhood level socioeconomic, demographic, ethnocultural and housing characteristics on COVID-19 mortality risks. It was found that COVID-19 mortality rates were highest for neighbourhoods characterized by a high proportion of low-income families, seniors, institutionalized populations and immigrants. Additionally, it was found that age-standardized COVID-19 mortality rate for men was 1.7 times higher than women in those senior or institutionalized population neighbourhoods. The findings help support and inform the public health efforts in protecting more vulnerable population groups on Canada.

4 Data used to create the CanSET

Data selected and analyzed for the CanSET were sourced from the Census of Population 2021 microdata for all the DAs within the 41 CMAs and 111 CAs in Canada. The DA was used as a geographic unit of analysis. Statistics Canada defines a DA as a small, relatively stable geographic unit composed of one or more adjacent dissemination blocks with an average population of 400 to 700 persons (Statistics Canada, 2021). The DAs are the smallest geographic units for which Statistics Canada disseminates its Census data for public use. The DAs are easy to aggregate small geographic units as they can be combined to form other larger geographic units. Although there are 57,936 DAs in Canada (2021 Census Geography), only the 44,518 (77%) DAs that belong to the CMAs or CAs were included in the CanSET because of their high population densities, and due to the availability of the variables of interest.

The DAs for which Census data (i.e., either short form or long form) were not released due to confidentiality or data quality issues were not included in the creation of the CanSET. Some DAs associated with Indian reserves were excluded from the analysis because some Census questions were either not asked or the concepts were not applicable, or they were incompletely enumerated.

Table 1
Dissemination Area (DA), average population and households by provinces and territories

Province/Territory	Total number of DAs	Number of DAs in CanSET	Average population per DA	Average number of households per DA
	number			
Newfoundland and Labrador	1,082	441	611	291
Prince Edward Island	319	167	579	276
Nova Scotia	1,670	1,032	642	305
New Brunswick	1,470	823	592	279
Quebec	13,806	10,775	645	311
Ontario	20,468	17,689	723	300
Manitoba	2,223	1,430	675	288
Saskatchewan	2,625	1,195	624	278
Alberta	6,203	4,613	760	320
British Columbia	7,848	6,284	704	310
Yukon	74	34	938	424
Northwest Territories	98	35	581	233
Nunavut	50	0

... not applicable

Source: Canadian Census of Population 2021.

5 Methodology

The CanSET was created using hierarchical cluster analysis of DAs into three levels of nested social environment clusters. Cluster analysis attempts to assign observations to groups (clusters) based on their similarity or differences using a measure of statistical (i.e., Euclidian) distance from each other. Observations within each group are found to be similar to one another with respect to variables or attributes of interest. In other words, the goal is to group the observations into homogeneous and distinct clusters. A hierarchical cluster analysis method was applied with a ‘Fast Ward’ option in JMP® 17 (JMP Statistical Discovery LLC) analytical software. The methodology used in 2016 was applied again for the creation of the 2021 clusters. Details of the 2016 methodology can be found in the “*Canadian Social Environment Typology: Methodology Guide*”, available upon request.

5.1 Variable Selection

The selection of variables used to construct the clusters for the 2021 CanSET was informed by those used in the 2016 CanSET, which had been identified through stakeholder consultations and an extensive literature review. There were 30 final variables selected as shown in Appendix A.

All the variables were aggregated at the DA level geography and median values were used. The variables chosen cover a wide range of subjects including demographic structure (age structure, family size, etc.), socioeconomic status (income, education, labour market status, etc.), ethnocultural background status (Indigenous status, immigration status, visible minority status, etc.) and housing characteristics (housing conditions, dwelling ownership, value etc.). Health-related variables were deliberately not used in the creation of the typology to make the typology equally applicable in other fields. General socioeconomic and demographic variables were derived directly from the short form Census, whereas specific variables that were not available in short form Census (e.g., level of education, immigration status, occupation etc.) were derived from the long form Census and weights were applied. Some variables were available at the person level whereas others were available at census family, economic family or household level.

The variables selected for the cluster analysis were measured on different scales, or on a common scale with differing variances. Therefore, they were standardized in order to mitigate the effect of these differences among the variables. All the 30 variables were standardized with mean 0 and variance 1 prior to performing the cluster analysis. Some variables with highly skewed distributions were still skewed after normalization. So, all variables were capped at their 99th percentile. The clustering algorithm was applied to the variables at DA level to group similar DAs into clusters.

5.2 Number of Clusters

Respecting the 2016 CanSET methodology, the optimal number of clusters were kept at six, ten and twenty for the first, second and third level of hierarchy respectively. The clusters are defined as the set of DAs that are similar in terms of the selected characteristics (variables). A nested hierarchy of six, ten and twenty clusters was created as outlined in Table 2.

Table 2
Number of optimal clusters in the 2021 CanSET by hierarchy

Cluster Levels	Total number of Dissemination Areas
A	15,707
A1	7,463
A11	3,132
A12	4,331
A2	6,151
A21	4,079
A22	2,072
A3	2,093
A31	2,093
B	11,697
B1	8,171
B11	2,912
B12	5,259
B2	3,526
B21	3,526
C	5,131
C1	5,131
C11	2,334
C12	393
C13	613
C14	1,791
D	1,614
D1	1,614
D11	1,614
E	4,143
E1	4,143
E11	1,481
E12	1,142
E13	1,520
F	6,226
F1	4,174
F11	2,786
F12	1,388
F2	2,052
F21	1,035
F22	1,017

Note: The alphabets A to F represent six clusters of the first level of the hierarchy; the alphanumeric values of A1 to F2 represent 10 clusters of the second level of the hierarchy; and the alphanumeric values of A11 to F22 represent 20 clusters of the third level of the hierarchy.

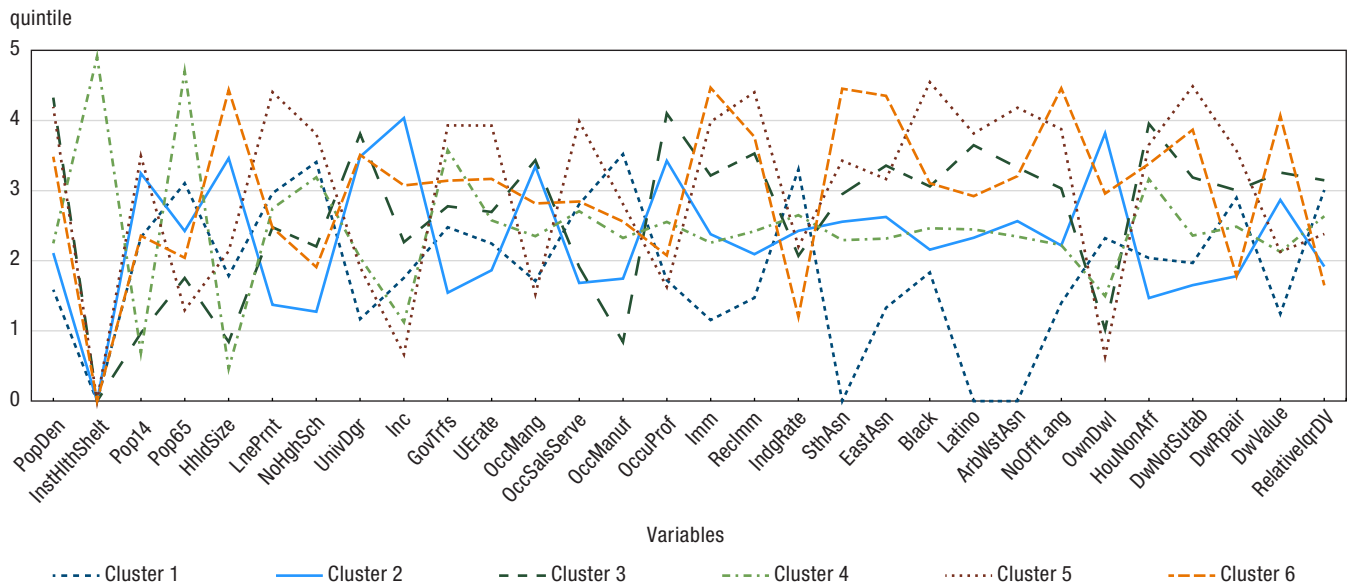
Source: 2021 Canadian Social Environment Typology Data.

5.3 Cluster Characteristics and Uses

The three levels of nested clusters were created for the CanSET using a hierarchical clustering method. The median values were calculated for each variable for each cluster solution and specific characteristics were developed based on the position of the median value of each cluster in a quintile distribution of all the DAs in the analysis. The distribution of the median values for all the variable at the six-cluster level can be seen in figure 1.

The distribution of the variables within the cluster was described using the quintile value, where DAs with a very high proportion of population with a certain characteristic were in quintile 5 and those DAs with a very low proportion of population with certain characteristics were in quintile 1. This method was used to characterize each cluster by highlighting those key variables in each cluster which were either in the high to very high range (quintile 4 to 5) or in the low to very low range (quintile 2 to 1). Additionally, all 30 variables are classified under 4 themes: demographic, socioeconomic, ethnocultural and housing characteristics (as seen in appendix A). All the key variables highlighted for a cluster were placed in the corresponding theme to structure each cluster description. The descriptions of the clusters at each cluster level can be seen in tables 3 to 5.

Chart 1
Quintile Distribution of six cluster median



Source: Statistics Canada, Census of Population microdata 2021.

The CanSET data table comes with a Dissemination Area Unique ID (DAUID), its population (DAPOP2021) and associated cluster membership for each of the three cluster solutions. Users can select the cluster solution that distinguishes their area of interest and can use the respective value to classify the neighbourhoods (e.g. A, A3, A32, etc.). For example, one DA may fall into cluster A when we use the six cluster classification but the same DA may fall into cluster A3 and A31 respectively for the ten and twenty cluster classification.

Users can select the cluster solution to use depending on the socioeconomic, demographic, ethnocultural and housing characteristics of a CMA or CA of interest. For instance, if users want to use the CanSET data to understand health inequalities in large CMAs like Toronto, Montreal or Vancouver, they may use all six, ten or twenty cluster solutions as all type of DAs are present in large CMAs having large population. However, only a few cluster types may be present in small CMAs or CAs. In that case, they may end up using only six or ten clusters to compare health outcomes.

5.3.1 2016 and 2021 Cluster Comparability

Although the same variables and methodology were used in the creation of 2016 and 2021 CanSET, the clusters are not directly comparable. The primary reason is that the data source differs: the 2016 CanSET was based on the 2016 Census of Population, while the 2021 CanSET used data from the 2021 Census of Population. Even though the same variables were applied, population characteristics such as population density, income, unemployment rates and immigration all changed over time (from 2016 to 2021).

Secondly, the clusters were created at the DA level through hierarchical clustering and were categorized into three different levels (Table 2), which also varies from 2016 to 2021. The number of DAs within each cluster and the specific cluster assigned to each level vary depending on the version. As a result, the 2021 CanSET clusters

are not directly comparable to 2016 clusters. The DAs included in the 2021 clusters were not the same DAs in the 2016 clusters, they are distinct. The clustering is based on the distance between variable means so that the observations within each cluster are similar to one another.

The 2021 CanSET data should not be used for direct comparisons with the 2016 CanSET data.

5.3.2 Six clusters solution

Level one of the hierarchy creates six clusters as the optimum cluster solution. The number of DAs in each cluster ranges from 1,614 to 15,707 whereas the median population density ranges from 1,975 to 7,511 people per square kilometre. In this solution, each DA is given a value of 1 to 6 and a cluster name from A to F based on its membership. Following are some of the major characteristics of each cluster for this solution. See Appendix A for variable definitions.

Table 3
Name and description of first level of 2021 clusters

Cluster No.	Cluster Name	Cluster Description
Cluster 1	A	<p>The majority of DAs in this category have low socioeconomic status and are mostly distributed in suburban regions. Demographic characteristics indicate a high proportion of population 65 years and above and one parent households but a very low population density.</p> <p>Socioeconomic characteristics indicate very high proportion of labour force concentrated in manufacturing occupations but low levels of educational attainment, and a very low proportion of households with a university degree at bachelors level or above.</p> <p>Ethnocultural characteristics indicate the cluster has limited ethnic diversity, with a very low proportion of immigrants and racialized population groups, but a higher proportion of Indigenous population.</p> <p>Housing characteristics indicate that housing conditions tend to be poorer and with very low dwelling values.</p>
Cluster 2	B	<p>The majority of DAs in this category have high socioeconomic status and are mostly distributed in suburban regions. Demographic characteristics indicate a high proportion of young population aged 14 years and below, large household size, and very low proportion of one parent families.</p> <p>Socioeconomic variables indicate a very high median household income and a high proportion of the population employed in professional and managerial occupations. The proportion of the population employed in sales occupations is very low, along with the unemployment rate and percent of population receiving government transfers.</p> <p>Ethnocultural characteristics indicates the cluster has low proportion of recent immigrants, Black and Latino population groups. Where there is a low proportion of population with no knowledge of either official language.</p> <p>Housing variables indicate a very high rate of dwelling ownership and very low proportion of households in need of major repairs or dwellings not suitable for the occupancy. A very low proportion of the population spends more than 30% of income on housing costs.</p>
Cluster 3	C	<p>In this cluster most of the DAs are located in core urban areas of major CMA/CAs in Canada.</p> <p>Demographic characteristics indicate that this cluster represent neighbourhoods with a very high population density and a small household size.</p> <p>Socioeconomic variables indicate a very high proportion of households with a university degree, a very high proportion of population working in management, administrative and professional occupations, and very low proportions working in manufacturing occupations.</p> <p>Ethnocultural characteristics indicates the cluster has high proportions of East Asian, Latino, Arab and West Asian population groups but low proportion of Indigenous population.</p> <p>Housing variables indicate that DAs in this cluster have a very high proportion of population spending more than 30% of income on housing costs and very high variability in housing values. The dwelling ownership rates are low.</p>
Cluster 4	D	<p>This cluster represents low socioeconomic status senior and institutionalized neighbourhoods in Canada.</p> <p>Demographic characteristics indicate DAs in this cluster have very high proportion of institutionalized populations and proportion of elderly population aged 65 and above. There is a very low proportion of population aged 14 years and below and clusters have a very small household size.</p> <p>Socioeconomic characteristics indicate DAs in this cluster have low median household incomes.</p> <p>Ethnocultural characteristics indicate that the proportion of immigrants and most racialized population groups such as South Asian, East Asian, Arab and West Asian are low.</p> <p>Housing characteristics indicate that dwelling values are low.</p>

Table 3
Name and description of first level of 2021 clusters

Cluster No.	Cluster Name	Cluster Description
Cluster 5	E	<p>Most DAs in this cluster represents multicultural urban neighbourhoods with low socioeconomic status. DAs in this cluster can be found in major cities from Quebec, Ontario, British Columbia and Alberta.</p> <p>Demographic characteristics indicate DAs in this cluster have a very high proportion of young population aged 14 years and below, with very high proportions of one parent households. There is a very low proportion of elderly population aged 65 and above.</p> <p>Socioeconomic characteristics indicate that there is a very high proportion of households where the highest level of education attained is less than high school. These DAs have a very high proportion of low-income households and those that receive government transfers of payments. The unemployment rate is very high and most employment is found in the sales and service occupations, whereas the employment in professional and managerial occupations is very low.</p> <p>Ethnocultural characteristics indicate that the proportion of recent immigrants and most racialized population groups are high.</p> <p>Housing characteristics indicate that a very high proportion of people live in households in need of major repairs and dwelling which are not suitable for the occupancy. The dwelling ownership rates in these DAs are very low.</p>
Cluster 6	F	<p>This cluster represents sub-urban neighbourhoods where DAs are mostly found in major CMAs such as Montreal, Toronto, Vancouver and Calgary.</p> <p>Demographic characteristics indicate that these neighbourhoods have very large household sizes.</p> <p>Ethnocultural characteristics indicate the immigrant population is high, with a very high proportion belonging to South Asian and East Asian racialized groups. DAs in this cluster have a very high proportion of population with no knowledge of either official language. In contrast, the proportion of people who identify as Indigenous peoples is very low.</p> <p>Housing characteristics indicate very high median dwelling values, with very low variability in dwelling values.</p>

Source: 2021 Canadian Social Environment Typology Data.

5.3.3 Ten clusters solution

Level two of the hierarchy creates ten clusters as an optimum cluster solution (Table 2). The number of DAs varies from a minimum of 1,614 to the maximum of 8,171. The median population density ranges from 625 to the maximum of 7,511 per square kilometre. The ten clusters developed are nested in the six clusters described above (see Table 2). Therefore, the clusters in each level share some characteristics with the higher level of cluster to which they belong. The ten clusters have the characteristics as outlined below. See Appendix A for variable definitions.

Table 4
Name and description of second level of 2021 clusters

Cluster No.	Cluster Name	Cluster Description
Cluster 1	A1	<p>Demographic characteristics indicate DAs in this cluster are characterized by a high proportion of seniors aged 65 and above but a very low population density.</p> <p>Socioeconomic characteristics indicate a very high proportion of the labour force in manufacturing occupations but a very low unemployment rate and a very low proportion of population receiving government transfers of payment.</p> <p>Ethnocultural characteristics indicate a very high proportion of Indigenous population but a very low proportion of immigrants, recent immigrants and racialized population groups.</p> <p>Housing characteristics indicate a very high variability of dwelling values. There is a very low proportion of households which are not suitable for the occupancy and population spending more than 30% of income on housing costs.</p>
Cluster 2	A2	<p>Demographic characteristics indicate DAs in this cluster have a very high proportion of seniors aged 65 and above and a high proportion of one parent households.</p> <p>Socioeconomic characteristics indicate very high proportion of population with below high school education and a very high proportion of the population working in sales and service occupations. Whereas a very low proportion of the population are in managerial occupations and a very low adjusted family income.</p> <p>Ethnocultural characteristics indicate DAs in this cluster have a very high proportion of Indigenous population and very low proportions of some racialized groups, such as South Asian, East Asian and Arab and west Asian groups. The proportion of recent immigrant population is very low.</p> <p>Housing characteristics indicate very low dwelling values.</p>

Table 4
Name and description of second level of 2021 clusters

Cluster No.	Cluster Name	Cluster Description
Cluster 3	A3	<p>Demographic characteristics indicate DAs in this cluster have a very high proportion of young population aged 14 years and below. The proportion of one parent households is very high but the population density is very low.</p> <p>Socioeconomic characteristics indicate there is a very high proportion of the population working in manufacturing occupations and very low proportion of those working in managerial and professional occupations. There is a very low proportion of households with a university degree at bachelor's level or above.</p> <p>Ethnocultural characteristics indicate there are very high proportion of Indigenous population but a very low proportion of immigrant population and racialized population groups are also very low with Latino, Arab and West Asian groups being the lowest.</p> <p>Housing characteristics indicate the proportion of dwellings in need of major repair are very high and median dwelling value is very low.</p>
Cluster 4	B1	<p>Demographic characteristics indicate DAs in this cluster have a low population density and very low proportion of one parent households.</p> <p>Socioeconomic characteristics indicate very high median household income and very low proportion of government transfer of payment recipients. These DAs are further characterized by a very low proportion of population with below high school education and very low unemployment rate.</p> <p>Ethnocultural characteristics indicate there is a high proportion of Indigenous and Latino population groups.</p> <p>Housing characteristics indicate very high dwelling ownership rates, with very low proportion of population spending more than 30% of income on housing costs and very low variability in dwelling values.</p>
Cluster 5	B2	<p>Demographic characteristics indicate DAs in this cluster have a very high median household income. There is a very low proportion of one parent families and a low population density.</p> <p>Socioeconomic characteristics indicate a very high proportion of the population is working in managerial, professional, and administration occupations, with very high proportions of households with a university degree. Whereas a very low proportion are working in sales and manufacturing occupations.</p> <p>Ethnocultural characteristics indicate high proportion of the South Asian population however a very low proportion of Latino population.</p> <p>Housing characteristics indicate very high dwelling values and very low proportion of households in need of major repairs or that are not suitable for the occupancy.</p>
Cluster 6	C1	<p>Demographic characteristics indicate DAs in this cluster have very high population density, but very small household size, with very low proportion of population aged 14 and below.</p> <p>Socioeconomic characteristics indicate a very high proportion of the population working in professional occupations and a very low proportion working in sales and manufacturing occupations.</p> <p>Ethnocultural characteristics indicate a high proportions of racialized population groups, with Latinos representing the largest proportion.</p> <p>Housing characteristics indicate a very high proportion of population spending more than 30% of income on housing costs and very high variation in the median dwelling value, while the dwelling ownership rates are very low.</p>
Cluster 7	D1	<p>Demographic characteristics indicate DAs in this cluster have very high proportion of institutionalized population and elderly population aged 65 and above. The one parent households are also high. The household size is very small and the proportion of population aged 14 and under is very low.</p> <p>Socioeconomic characteristics indicate a very high proportion of government transfer of payment recipients. The proportion of households with less than high school education is high.</p> <p>Ethnocultural characteristics indicate high proportions of the Indigenous population and the Latino population groups.</p> <p>Housing characteristics indicate a high proportion of the population spending more than 30% of income on housing costs and high variation in the median dwelling value.</p>
Cluster 8	E1	<p>Demographic characteristics indicate DAs in this cluster have a very high population density, with a very high proportion of population aged 14 years and below and a very high proportion of one parent households.</p> <p>Socioeconomic characteristics indicate the proportion of government transfer of payment recipients is very high. These DAs are characterized by a very high unemployment rate and proportion of the labour force working in sales and service-related occupations. A very low proportion of labour force working in professional occupations and very low adjusted family income.</p> <p>Ethnocultural characteristics indicate a very high proportion of the recent immigrant population and the proportions of racialized population groups are also high, where Black, Latino and Arab and West Asian groups are very high.</p> <p>Housing characteristics indicate proportions of dwellings not suitable for the occupancy and in need of major repairs are very high, while the dwelling ownership rates are very low.</p>

Table 4
Name and description of second level of 2021 clusters

Cluster No.	Cluster Name	Cluster Description
Cluster 9	F1	<p>Demographic characteristics indicate DAs in this cluster have a very large household size and very low proportion of elderly population aged 65 and above. These DAs have high population density.</p> <p>Socioeconomic characteristics indicate a very high unemployment rate.</p> <p>Ethnocultural characteristics indicate the proportion of population not speaking either of the official languages of Canada is very high, along with the proportions of immigrant, recent immigrants and racialized population groups. There is a very low proportion of Indigenous population.</p> <p>Housing characteristics indicate high median dwelling values and very high proportion of dwellings not suitable for the occupancy. There is a very low proportion of dwellings in need of repair and very low variation in dwelling values.</p>
Cluster 10	F2	<p>Demographic characteristics indicate DAs in this cluster have very large household size and a high population density.</p> <p>Socioeconomic characteristics indicate the proportion of the population working in management and administration occupations is very high, along with a high proportion of households with a university degree.</p> <p>Ethnocultural characteristics indicate there is a very high proportion of the population not speaking either of the official languages of Canada. The proportions of immigrants and racialized population groups, such as South and East Asian, are very high. The proportions of the Indigenous population and Black population are very low.</p> <p>Housing characteristics indicate very high dwelling value and a high proportion of households spending more than 30% of income on housing costs.</p>

Source: 2021 Canadian Social Environment Typology Data.

5.3.4 Twenty-clusters solution

Level three of the hierarchy creates twenty clusters as an optimum cluster solution. The twenty clusters developed are nested within the ten clusters, which in turn are nested within the six clusters described above (see Table 2). Therefore, the clusters in each level share some characteristics with the previous level of cluster to which they are nested. In the twenty-cluster solution, the number of DAs in each cluster ranges from 393 to 5,259 whereas the median population density ranges from 179 to 29,447 per square kilometre. The twenty clusters have distinct characteristics that differentiate them from one another. See Appendix A for variable definitions.

Table 5
Name and description of third level of 2021 clusters

Cluster No.	Cluster Name	Cluster Description
Cluster 1	A11	<p>Demographic characteristics indicate DAs in this cluster have very low population density.</p> <p>Socioeconomic characteristics indicate a very low unemployment rate and very low proportion of government transfer of payment recipients.</p> <p>Ethnocultural characteristics indicate a high proportion of Indigenous population and those who do not speak either of the official languages of Canada. There are very low proportions of immigrant, recent immigrants and racialized population groups.</p> <p>Housing characteristics indicate there is a high dwelling ownership rate but very low dwelling values. There are very low proportions of households spending more than 30% on housing cost and dwellings not suitable for the occupancy.</p>
Cluster 2	A12	<p>Demographic characteristics indicate DAs in this cluster have a very high proportion of seniors aged 65 and above but a very low population density.</p> <p>Socioeconomic characteristics indicate a very high proportion of the labour force working in manufacturing occupations whereas a very low proportion working in sales and professional occupations. There is a very low proportion of households having a member with a university degree and a very low unemployment rate.</p> <p>Ethnocultural characteristics indicate very low proportions of immigrants, recent immigrants and racialized population groups. There is a very high proportion of the Indigenous population.</p> <p>Housing characteristics indicate a very high dwelling ownership rate and very high variability of the dwelling values.</p>
Cluster 3	A21	<p>Demographic characteristics indicate DAs in this cluster have high proportions of seniors aged 65 and above and one parent households.</p> <p>Socioeconomic characteristics indicate DAs in this cluster have a very high proportion of population with below high school education and high unemployment rate.</p> <p>Ethnocultural characteristics indicate low proportions of racialized population groups, where the proportion of Arab and West Asian populations are very low. The proportion of the Indigenous population is high.</p> <p>Housing characteristics indicate low dwelling values and high proportion of dwellings that need major repairs.</p>

Table 5
Name and description of third level of 2021 clusters

Cluster No.	Cluster Name	Cluster Description
Cluster 4	A22	<p>Demographic characteristics indicate DAs in this cluster have a very high proportion of seniors aged 65 and above. A very high proportion of one parent families are found in this cluster, whereas the average household size is very small.</p> <p>Socioeconomic characteristics indicate a very high proportion of government transfer of payment recipients and a very high proportions of the labour force in manufacturing, and sales and service occupations. The proportion of households having a member with a university degree is very low and a very low proportion employed in managerial and professional occupations.</p> <p>Ethnocultural characteristics indicate very low proportions of immigrants, recent immigrants and racialized population groups.</p> <p>Housing characteristics indicate very high variability of the median dwelling values but the dwelling values are very low.</p>
Cluster 5	A31	<p>Demographic characteristics indicate DAs in this cluster have a very high proportion of the population aged 14 and below whereas the population density is very low.</p> <p>Socioeconomic characteristics indicate a very high proportion of the labour force in manufacturing occupations. The proportion of households having a member with a university degree is very low, along with those employed in managerial occupations.</p> <p>Ethnocultural characteristics indicate a very high proportion of the Indigenous population. The proportion of immigrants is very low, along with racialized population groups, particularly Latino and Arab and West Asians groups are very low.</p> <p>Housing characteristics indicate a very high proportion of dwellings in need of major repair and very low median dwelling values.</p>
Cluster 6	B11	<p>Demographic characteristics indicate DAs in this cluster have a very high proportion of population aged 14 and under and a very low proportion of one parent households.</p> <p>Socioeconomic characteristics indicate very high adjusted median household income and high proportion of households having a member with a university degree. There is a very low proportion of recipients of government transfer of payment and very low unemployment rate.</p> <p>Ethnocultural characteristics indicate a high proportion of Arab and West Asian population but very low proportions of South Asian and East Asian populations. There is a very low proportion of recent immigrants.</p> <p>Housing characteristics indicate a very high dwelling ownership rate. There is a very low proportions of dwellings which are not suitable for the occupancy and a very low proportion spending more than 30% of income on housing costs.</p>
Cluster 7	B12	<p>Demographic characteristics indicate DAs in this cluster have a very high proportion of the population aged 14 and below. The household size is large and there is a very low population density.</p> <p>Socioeconomic characteristics indicate a very high adjusted family income.</p> <p>Ethnocultural characteristics indicate a very high proportion of the Indigenous population and a low proportion of racialized population groups, particularly Arab and West Asian population is very low.</p> <p>Housing characteristics indicate a very high dwelling ownership rate and very low variability in dwelling values.</p>
Cluster 8	B21	<p>Demographic characteristics indicate DAs in this cluster have a very low population density and a very low proportion of one parent families.</p> <p>Socioeconomic characteristics indicate a very high adjusted household income and a very high proportion of households having a member with a university degree. A very high proportion of the labour force is working in managerial and professional occupations whereas the proportions of those working in sales and manufacturing occupations are very low.</p> <p>Ethnocultural characteristics indicate that there are very low proportions of Black and Latino population groups. There are low proportions of recent immigrants and Indigenous population groups.</p> <p>Housing characteristics indicate a very high household ownership rate and very high dwelling value, whereas the proportion of the population spending more than 30% of income on housing costs and the proportion of dwellings in need of major repairs are very low.</p>
Cluster 9	C11	<p>Demographic characteristics indicate DAs in this cluster have a very high population density. There are very low proportions of population aged 14 and under, as well as population aged 65 and above.</p> <p>Socioeconomic characteristics indicate a high proportion of households with at least a member having a university degree and a very high proportion of the population working in professional occupations. The proportion working in manufacturing occupations is very low.</p> <p>Ethnocultural characteristics indicate high proportions of racialized population groups, where the Latino group is very high.</p> <p>Housing characteristics indicate a very high variation in dwelling values and very high proportion of dwellings in need of major repair. The proportion of the population owning their own dwellings is very low.</p>

Table 5
Name and description of third level of 2021 clusters

Cluster No.	Cluster Name	Cluster Description
Cluster 10	C12	<p>Demographic characteristics indicate DAs in this cluster have a very high population density and the average household size is very small. There are very low proportions of the population aged 14 years and below, as well as seniors aged 65 and above, and one parent households.</p> <p>Socioeconomic characteristics indicate a very high proportion of the population working in professional and management occupations, along with a high proportion of households where at least one member holds a university degree. The proportions working in sales and manufacturing occupations are very low.</p> <p>Ethnocultural characteristics indicate high proportions of immigrant and racialized populations groups, particularly very high recent immigrants, East Asian, Latino, Arab and West Asian groups.</p> <p>Housing characteristics indicate a very high proportion spending more than 30% of income on household costs and a very low proportion of the population own their dwellings.</p>
Cluster 11	C13	<p>Demographic characteristics indicate DAs in this cluster have a very high proportion of senior population aged 65 and above. The average household size is very small. DAs in this cluster have very low proportion of one parent households.</p> <p>Socioeconomic characteristics indicate very high proportions of the population working in management and administration occupations, whereas the population working in sales and service occupations is also very low. The unemployment rate is also very low.</p> <p>Ethnocultural characteristics indicate a low proportion of the Black population group.</p> <p>Housing characteristics indicate a very high variability in dwelling values and very low proportion of dwellings needing major repairs.</p>
Cluster 12	C14	<p>Demographic characteristics indicate DAs in this cluster have a high population density.</p> <p>Socioeconomic characteristics indicate a very high proportion of households having a member with a university degree and a very high proportion of the labour force working in professional occupations, whereas a very low proportion of the population is working in manufacturing occupations.</p> <p>Ethnocultural characteristics indicate the proportions of immigrants and racialized population groups are high. There is a low proportion of the Indigenous population.</p> <p>Housing characteristics indicate a very high proportion of dwellings are not suitable for the occupancy and very high median dwelling values.</p>
Cluster 13	D11	<p>Demographic characteristics indicate DAs in this cluster have a very high proportion of seniors and institutionalized populations. The average household size is very small and a very low proportion of population is aged 14 years and below.</p> <p>Socioeconomic characteristics indicate the proportion of government transfer of payment recipients is high and a high proportion of the population has below high school education.</p> <p>Ethnocultural characteristics indicate a high proportion of the Indigenous population.</p> <p>Housing characteristics indicate a high variability in dwelling value. A low proportion of dwellings are not suitable for the occupancy and there is a low dwelling ownership.</p>
Cluster 14	E11	<p>Demographic characteristics indicate DAs in this cluster have a very low proportion of population aged 65 and above and very high proportion of one parent households.</p> <p>Socioeconomic characteristics indicate there is a very high unemployment rate and a very high proportion of the population working in sales and service occupations. The proportion of households with below high school education and recipients of government transfer of payment are very high. The adjusted family income is very low and a very low proportion of the population is working in professional and managerial positions.</p> <p>Ethnocultural characteristics indicate very high proportions of recent immigrants, Indigenous population and racialized population groups, such as the Black population.</p> <p>Housing characteristics indicate very high proportions of dwellings not suitable for the occupancy and those in need of major repairs. The dwelling values are very low.</p>
Cluster 15	E12	<p>Demographic characteristics indicate DAs in this cluster have a very high proportion of population aged 14 years and below and a very high proportion of one-parent households. The population aged 65 and above are a very low proportion.</p> <p>Socioeconomic characteristics indicate this cluster has a very high unemployment rate and very high proportion of recipients of government transfer of payments. There are very high proportions of the population with below high school education and those working in sales and service occupations. There is a very low proportion of the population working in managerial positions and the adjusted family income is very low.</p> <p>Ethnocultural characteristics indicate a high proportion of the immigrant population and very high proportions of racialized population groups, such as Black, Latino, Arab and West Asian groups.</p> <p>Housing characteristics indicate there is a very high proportion of dwellings in need of major repair and those deemed not suitable for the occupancy. The dwelling ownership rate is very low.</p>

Table 5
Name and description of third level of 2021 clusters

Cluster No.	Cluster Name	Cluster Description
Cluster 16	E13	<p>Demographic characteristics indicate DAs in this cluster have a very high population density and very high proportion of one parent households.</p> <p>Socioeconomic characteristics indicate these DAs have a very high unemployment rate and the proportion of recipients of government transfer of payment is very high. There is a very high proportion of those working in sales and service occupations. The median household income is very low.</p> <p>Ethnocultural characteristics indicate very high proportions of immigrants, recent immigrants and racialized population groups, such as the South Asian population. A very high proportion of the population does not speak either of the official languages of Canada and there is very a low proportion of the Indigenous population.</p> <p>Housing characteristics indicate a high proportion of dwellings that are not suitable for the occupancy. There is a very low household ownership rate.</p>
Cluster 17	F11	<p>Demographic characteristics indicate DAs in this cluster have a high population density and very large household size.</p> <p>Socioeconomic characteristics indicate very high unemployment rate.</p> <p>Ethnocultural characteristics indicate very high proportions of racialized population groups, particularly South Asian, East Asian, Black, Arab and West Asian groups.</p> <p>Housing characteristics indicate a very low variability in dwelling value and a very low proportion of dwellings in need of major repairs.</p>
Cluster 18	F12	<p>Demographic characteristics indicate DAs in this cluster have very large household sizes and a high population density.</p> <p>Socioeconomic characteristics indicate a very high proportion working in manufacturing occupations but a very low proportion working in professional occupations.</p> <p>Ethnocultural characteristics indicate very high proportions of immigrants and population not speaking either of the official languages of Canada. The proportions of racialized population groups are high, with the South Asian group very high. The proportion of the Indigenous population is very low.</p> <p>Housing characteristics indicate a very high proportion of dwellings not suitable for the occupancy and a very low proportion of households in need for major repairs. The variability in dwelling value is very low.</p>
Cluster 19	F21	<p>Demographic characteristics indicate DAs in this cluster have a very large household size.</p> <p>Socioeconomic characteristics indicate very high median household income. These DAs have a very high proportion of households having a member with a university degree, a very high proportion of population working in management and administration occupations and a very low proportion of government transfer of payment recipients.</p> <p>Ethnocultural characteristics indicate very high proportions of the immigrant population and the East Asian population group. The proportions of the Black and Indigenous population groups are very low.</p> <p>Housing characteristics indicate very high dwelling values and very high proportion of population spending more than 30% of income on housing costs.</p>
Cluster 20	F22	<p>Demographic characteristics indicate DAs in this cluster have high population density and very large household size.</p> <p>Socioeconomic characteristics indicate a high proportion of households having a member with a university degree and a high unemployment rate. There are high proportions of population working in sales, management and administration occupations, whereas low proportions working in professional and manufacturing occupations.</p> <p>Ethnocultural characteristics indicate very high proportions of the immigrant population, as well as South Asian and East Asian population groups. There is a very high proportion of population not speaking either of the official languages of Canada. There is a very low proportion of the Indigenous population.</p> <p>Housing characteristics indicate very high dwelling values and a very high proportion of population spending more than 30% of income on housing costs. The variability in median dwelling value is very low.</p>

Source: 2021 Canadian Social Environment Typology Data.

6 How to integrate the CanSET with your data

The CanSET is designed to be use with any dataset for which the DA or postal code information is available. The CanSET data can be easily integrated with other data sources using the dissemination area unique identification (DAUID) number from the Census of population 2021. DAUID from any other Census of population may need a correspondence file to match them with the 2021 Census DAUID. Any datasets that do not have DA information but have the postal code information can also be integrated with the CanSET data using the Postal Code^{DM} Conversion File plus (PCCF+).

The CanSET data comes with a table that includes DAUID, DA population and the hierarchal cluster membership for each of the three cluster solutions. The table contains nine columns of DAUID, PRUID, PRNAME, CMAUID, CMANAME, DAPOP2021, CanSET2021_6, CanSET2021_10 and CanSET2021_20 with cluster name and alpha-numeric value of A to F, A1 to F10 and A11 to F20 respectively as shown in a sample table (Table 6) below.

Match the 2021 DAUID from the CanSET data to the DAUID (or postal code) on your data file to obtain the CanSET cluster number for the cluster solution of interest. Review the cluster descriptions provided in this user guide to interpret the key socioeconomic, demographic, ethnocultural and housing characteristics in the DA.

By comparing the same CanSET number across geographies, it is possible to compare similar types of DAs with similar types of characteristics across Canada. Additionally, by comparing the distribution of CanSET numbers within a selected geography, it is possible to describe the different DAs and their characteristics.

Table 6
Example rows of CanSET data

DAUID	PRUID	PRNAME	CMAUID	CMANAME	DAPOP2021	CanSET2021_6	CanSET2021_10	CanSET2021_20
11020140	11	Prince Edward Island	105	Charlottetown	307	A	A1	A11
11020141	11	Prince Edward Island	105	Charlottetown	407	B	B1	B11
11020142	11	Prince Edward Island	105	Charlottetown	449	A	A2	A21
11020143	11	Prince Edward Island	105	Charlottetown	490	A	A1	A11
11020144	11	Prince Edward Island	105	Charlottetown	753	E	E1	E11
11020145	11	Prince Edward Island	105	Charlottetown	568	A	A1	A11
11020146	11	Prince Edward Island	105	Charlottetown	511	D	D1	D11
11020147	11	Prince Edward Island	105	Charlottetown	628	A	A2	A21

Source: Canadian Census of Population 2021 and 2021 Canadian Social Environment Typology Data.

7 Summary and Conclusion

This user guide provides an overview of the methodology used to develop the CanSET and information on how to use these social environment clusters with the users' own health or social data. To create the CanSET, Canadian DAs were grouped together based on their population's demographic structure, socioeconomic status, ethnocultural background and housing characteristics. The CanSET used 30 different variables from the Census of Population 2021 to classify Canadian DAs into three nested levels of clusters using the Fast Ward hierarchical clustering method. The optimal number of clusters (six, ten and twenty) determined using the Davies-Bouldin index for the previous version of the CanSET (2016) was applied to the 2021 CanSET. The median values were calculated for each variable for each cluster solution and specific characteristics and names were given to each cluster based on the distribution of median value of the variables in each clusters.

The CanSET portrays the complex social composition of Canadian urban geography. It is a comprehensive tool including a variety of dimensions of urban social environment and covers all populous areas of Canada. This area-level measure is unique as it includes all DAs from the CMAs and CAs simultaneously in a single typology rather than analyzing them separately for each CMA or CA. The analysis is based on the microdata from the Census of Population 2021, which covers all population of the study area. Therefore, the CanSET data allows cities to make comparison among neighbourhoods within themselves or with other cities and facilitates them to set benchmarks and track progress in combating health and social inequalities.

Appendix A

Table A
Description of the variables included in the CanSET cluster analysis

Theme	Variable name	Description	Universe and Remarks
Demographic	PopDen	Population density; total number of people per square kilometre of area.	Includes all population
	InstHlthShelt	Percent of population that are institutionalized and residing in medical or long-term care facilities or shelters	Includes all population
	Pop14	Percent of population aged 0 to 14 years	Includes all population
	Pop65	Percent of population aged 65 and above	Includes all population
	HhldSize	Average number of people in a household	Includes occupied private dwellings and units in senior residence but excludes other collective dwellings
	LnePrnt	Percent of one-parent census families	Includes occupied private dwellings and units in senior residence but excludes other collective dwellings
Socioeconomic	NoHghSch	Percent of private households with the highest level of education of all its members "no certificate, degree or diploma"	Includes number of people aged 15 and over in occupied private households
	UnivDgr	Percent of private households with at least one member of the household having "university certificate, diploma or degree at bachelor level or above"	Includes number of people aged 15 and over in private households
	Inc	Household size adjusted median household income after tax	The after-tax income of a household is the sum of the after-tax incomes of all members of that household.
	GovTrfs	Percent of population receiving specific government transfers.	Includes population aged 15 and over in an occupied private dwellings and units in senior residence but excludes other collective dwellings
	UErate	Unemployment rate for population aged 15 years and above	Includes population aged 15 and above who were available for work but unemployed in the census reference week
	OccMang	Percent of employed labour force in management and administration occupation	Includes population aged 15 and above who were available for work in the census reference week
	OccSalsServe	Percent of employed labour force in sales and service occupations.	Includes population aged 15 and above who were available to work in the census reference week
	OccManuf	Percent of employed labour force in manufacturing, construction and trade related occupation	Includes population aged 15 and above who were available for work in the census reference week
OccuProf	Percent of employed labour force in professional occupation	Includes population aged 15 and above who were available for work in the census reference week	

Table A
Description of the variables included in the CanSET cluster analysis

Theme	Variable name	Description	Universe and Remarks
Ethnocultural	Imm	Percent of immigrant population	Includes persons in occupied private households. This question is not asked in Indian Reserves.
	Reclmm	Percent of recent immigrant population (landed in Canada between 2016 and 2021)	Includes persons in occupied private households. This question is not asked in Indian Reserves.
	IndgRate	Percent of people who identified themselves with the Indigenous peoples of Canada	Includes respondents to this question in occupied private households.
	SthAsn	Percent of people belonging to South Asian visible minority group	Includes respondents to the visible minority question in occupied private households.
	EastAsn	Percent of people belonging to Chinese, Filipino, Southeast Asian, Korean or Japanese visible minority groups	Includes respondents to the visible minority question in occupied private households.
	Black	Percent of people belonging to Black visible minority group	Includes respondents to the visible minority question in occupied private households.
	Latino	Percent of people belonging to Latin American visible minority group	Includes respondents to the visible minority question in occupied private households.
	ArbWstAsn	Percent of people belonging to Arab or West Asian visible minority groups	Includes respondents to the visible minority question in occupied private households.
	NoOffLang	Percent with no knowledge of either of official languages	Includes all population
Housing	OwnDwl	Percent of private dwellings occupied by owner	Includes households that own or rent their private dwelling. Excluded is shelter occupancy on Indian reserves or settlements.
	HouNONAff	Percent of households spending more than 30% of its average total income on shelter costs	This variable is calculated for private households living in owned or rented dwellings who reported a total household income greater than zero. Excluded are band housing, farm dwellings.
	DwNotSutab	Percent of private households living in not suitable accommodations (according to the National Occupancy Standard)	A dwelling is considered suitable if it has enough bedrooms for the size and composition of the household. Includes only private households.
	DwRpair	Percent of occupied private dwellings in need of major repairs	Includes only occupied private dwellings
	DwValue	Median value of privately owned dwellings	Includes only occupied private dwellings
	Relative_iqrDV	Measure of variability in dwelling value within a DA	Includes only occupied private dwellings

Source: Canadian Census of Population 2021.

References

[JMP Statistical Discovery](https://www.jmp.com/content/dam/jmp/documents/en/support/jmp17/automation-reference.pdf) LLC SAS Campus Drive Cary, North Carolina 27513-2414; Retrieved from: <https://www.jmp.com/content/dam/jmp/documents/en/support/jmp17/automation-reference.pdf>

Johnson, R., & Wicheren, D. (2007). *Applied Multivariate Statistical Analysis* (6th ed.). Prentice Hall.

Statistics Canada. (2018). Postal Code Conversion File Plus (PCCF+) Version 7B, Reference Guide. Ottawa: Statistics Canada.

Statistics Canada. (2021). [Dictionary, Census of Population](https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/az/Definition-eng.cfm?ID=geo021). Ottawa: Statistics Canada. Retrieved from <https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/az/Definition-eng.cfm?ID=geo021>

Subedi, R., & Aitken, N. (2022). [Inequalities in COVID-19 mortality rates by neighbourhood types in Canada](https://www150.statcan.gc.ca/n1/pub/45-28-0001/2022001/article/00006-eng.htm). Statistics Canada. Retrieved from: <https://www150.statcan.gc.ca/n1/pub/45-28-0001/2022001/article/00006-eng.htm>