Canada

Canadian Community Health Survey - Annual Component, 2017-2018

Study Documentation

September 20, 2020

Metadata Production

Metadata Producer(s)	Statistics Canada , Statistics Canada
Production Date	September 4, 2020
Version	

Table of Contents

Canadian Community Health Survey - Annual Component, 2017-2018 (CCHS 2017-2018, Annual Component)

Enquête sur la santé dans les collectivités canadiennes - Composante annuelle, 2017-2018

Overview

Type	Canadian Community Health Survey
Identification	CCHS-82M0021X-E-2017-2018-Annual
Version	Production Date: 2018-11-14
Series	The central objective of the Canadian Community Health Survey (CCHS) is to gather health-related data at the sub-provincial levels of geography (health region or combined health regions).

Abstract

In 1991, the National Task Force on Health Information cited a number of issues and problems with the health information system. To respond to these issues, the Canadian Institute for Health Information (CIHI), Statistics Canada and Health Canada joined forces to create a Health Information Roadmap. From this mandate, the Canadian Community Health Survey (CCHS) was conceived. The CCHS is a cross-sectional survey that collects information related to health status, health care utilization and health determinants for the Canadian population. The survey is offered in both official languages. It relies upon a large sample of respondents and is designed to provide reliable estimates at the health region level every 2 years. The CCHS has the following objectives: - Support health surveillance programs by providing health data at the national, provincial and intra-provincial levels; - Provide a single data source for health research on small populations and rare characteristics; - Timely release of information easily accessible to a diverse community of users; - Create a flexible survey instrument that includes a rapid response option to address emerging issues related to the health of the population. The CCHS produces an annual microdata file and a file combining two years of data. The CCHS collection years can also be combined by users to examine populations or rare characteristics. The primary use of the CCHS data is for health surveillance and population health research. Federal and provincial departments of health and human resources, social service agencies, and other types of government agencies use the information collected from respondents to monitor, plan, implement and evaluate programs to improve the health of Canadians. Researchers from various fields use the information to conduct research to improve health. Non-profit health organizations and the media use the CCHS results to raise awareness about health, an issue of concern to all Canadians. The survey began collecting data in 2001 and was repeated every two years until 2005. Starting in 2007, data for the Canadian Community Health Survey (CCHS) were collected annually instead of every two years. While a sample of approximately 130,000 respondents were interviewed during the reference periods of 2001, 2003 and 2005, the sample size was changed to 65,000 respondents each year starting in 2007. In 2012, CCHS began work on a major redesign project that was completed and implemented for the 2015 cycle. The objectives of the redesign were to review the sampling methodology, adopt a new sample frame, modernize the content and review the target population. Consultations were held with federal, provincial and territorial share partners, health region authorities and academics. As a result of the redesign, the 2015 CCHS has a new collection strategy, is drawing the sample from two different frames and has undergone major content revisions. With all these factors taken together, caution should be taken when comparing data from previous cycles to data released for the 2015 cycle onwards.

Kind of Data	Survey data

Scope & Coverage

Keywords	PUMFFILE, Diseases and health conditions, Health, Health care services, Lifestyle and social conditions, Mental health and well-being, Chronic health problems, Health information, Health status indicators
Topics	Diseases and health conditions, Health, Health care services, Lifestyle and social conditions, Mental health and well-being
Time Period(s)	2017-2018
Countries	Canada

Universe

The CCHS covers the population 12 years of age and over living in the ten provinces and the three territories. Excluded from the survey's coverage are: persons living on reserves and other Aboriginal settlements in the provinces; full-time members of the Canadian Forces; the institutionalized population, children aged 12-17 that are living in foster care, and persons living in the Quebec health regions of Région du Nunavik and Région des Terres-Cries-de-la-Baie-James. Altogether, these exclusions represent less than 3% of the Canadian population aged 12 and over. In the north, the frame for the CCHS covers 94% of the targeted population in the Yukon, 96% in the Northwest Territories and 93% in Nunavut.

Producers & Sponsors

Other Producer(s)	Statistics Canada , Government of Canada

Sampling

Sampling Procedure

To provide reliable estimates at the health region (HR) level, a sample of 130,000 respondents is required on a two years basis: 120,000 respondents to cover the population aged 18 and over and 10,000 respondents to cover the population aged 12 to 17 years. Since 2015, a multi-stage sample allocation strategy is used to give relatively fair sample distribution to the HRs and the provinces. For each age group (18 and over, 12 to 17), the sample is first allocated among the provinces using a power allocation of 0.75 according to the size of their respective population. Each province's sample is then allocated among its HRs using a power allocation of 0.35 according to the size of the population in each HR. From 2015 onwards, the CCHS sample is selected using two different frames: an Area frame and the Canadian Child Tax Benefit (CCTB) frame. Using the Area frame, a sample of dwellings is selected to target the population aged 18 and over. During collection, all members of the dwelling are listed and a person aged 18 years or over is automatically selected using various selection probabilities based on age and household composition. The CCTB frame is used to sample persons aged 12 to 17 years. One child is then pre-selected to complete the survey. The area frame is mainly designed to serve the Labour Force Survey (LFS). Thus, the sampling plan of the LFS must be considered in selecting the CCHS dwelling sample. The LFS plan is a complex two stage stratified design in which each stratum is formed of clusters. The LFS first selects clusters using a sampling method with a probability proportional to size (PPS), and then the final sample is chosen using a systematic sampling of dwellings in the cluster. For CCHS, LFS clusters are grouped in each HR. Then, a sample of clusters and systematic dwellings are selected in each HR. The process maximizes the overlap between the clusters selected by both surveys and ensures that the same dwelling is selected only once. For the CCTB frame, an HR is assigned to each child in the target population based on the address. The CCTB frame is then stratified by HR. A simple random sample (SRS) of children aged 12 to 17 is selected within each HR. The size of the sample is enlarged during the selection process to account for non responses and units outside the coverage (for example, vacant dwellings, institutions, children not eligible due to age or death, etc.).

Weighting

In order for estimates produced from survey data to be representative of the covered population, and not just the sample itself, users must incorporate the survey weights in their calculations. A survey weight is given to each respondent included in the final sample. This weight corresponds to the number of persons in the entire population that are represented by the respondent. As described above, the CCHS uses two sampling frames for its sample selection: an area frame for the Canadian population aged 18 and over, and a frame of telephone numbers from Canada Child Benefit (CCB) records for the 12-17 population. The weighting strategy treats both the area and CCB frames independently to come up with separate person-level weights for each of the frames used. The adjustments applied to the initial weights are based on modeling probabilities of response (at the household level and person level). Variables derived from the collection paradata as well as characteristics of the units are used to create the models. Then these probabilities are used to create groups of respondents and nonrespondents in which to transfer the weights of the nonrespondents to the respondents. The person-level weights from the two frames are then combined into a single set of weights, jointly undergo two more adjustments (Winsorization and Calibration to known population totals such as by geography and age and sex), and become the final person-level weights. Bootstrap weights are created through resampling the original sample and applying similar adjustments to the bootstrap weights as to the sample weights. The steps for weighting are described in chapter 8 of the CCHS User Guide. The sample design used for this survey was not self-weighting. That is to say, the sampling weights are not identical for all individuals in the sample. When producing simple estimates, including the production of ordinary statistical tables, users must apply the proper sampling weight. Estimates of the number of people with a certain characteristic are obtained from the data file by summing the final weights of all records possessing the characteristic of interest. Proportions and ratios are obtained by summing the final

weights of records having the characteristic of the numerator and the denominator, and then dividing the first estimate by the second.

Data Collection

Data Collection Dates	start 2017-01-02 end 2018-12-24
Time Period(s)	start 2017-01-01 end 2018-12-31
Data Collection Mode	Computer Assisted Telephone Interview [cati]

Questionnaires

Each component of the CCHS questionnaire is developed in collaboration with specialists from Statistics Canada, other federal and provincial departments and/or academic fields. The CCHS questions are designed for computer-assisted interviewing (CAI), meaning that, as the questions were developed, the associated logical flow into and out of the questions was programmed. This includes specifying the type of answer required, the minimum and maximum values, on-line edits associated with the question and what to do in case of item non-response. CCHS content is comprised of three components: common content (core content and theme content), optional content and Rapid Response content. Common content is asked of all respondents with some modules being collected each year. Other modules from the common content are collected for one or two years and alternate every two to four years. The optional content fulfils the unique data needs of each province or territory and may vary from year to year. The Rapid Response component is offered to organizations interested in national estimates on an emerging or specific issue related to the population's health. Provincial estimates may also be yielded from a Rapid Response, however they may be of limited quality. A Rapid Response component may be added to the survey in each three-month collection period. The data will be released about six months after the collection period via an announcement in The Daily. New modules and revisions to existing CCHS content are tested using different methods. Qualitative tests using individual cognitive interviews or, more rarely, focus groups are used to ensure that questions and concepts are appropriately worded. The computer application for data collection is extensively tested in-house each time changes are made. The objective of these tests is to identify any errors in the program flow and text before the start of the main survey.

Data Collector(s)	Statistics Canada (StatCan) , Government of Canada

Data Processing & Appraisal

Data Editing

Error Detection: Most editing of the data was performed at the time of the interview by the computer-assisted interviewing (CAI) application. It was not possible for interviewers to enter out-of-range values and flow errors were controlled through programmed skip patterns. For example, CAI ensured that questions that did not apply to the respondent were not asked. In response to some types of inconsistent or unusual reporting, warning messages were invoked but no corrective action was taken at the time of the interview. Where appropriate, edits were instead developed to be performed after data collection at Head Office. Inconsistencies were usually corrected by setting one or both of the variables in question to "not stated".
Imputation: Household income data in the 2018 CCHS is imputed. Missing values due to either respondent refusal or respondent's lack of knowledge of household income are replaced using a nearest neighbour imputation method based on a modeled household income.

Estimates of Sampling Error

The quality of estimates produced with CCHS data is measured with the coefficient of variation (CV), produced using bootstrap weights. The CV magnitude will depend on the domain of interest and the prevalence of the characteristic.

Accessibility

Access Authority	Statistics Canada (Government of Canada) , https://www.statcan.gc.ca/eng/microdata/data-centres , statcan.infostats-statcan.mad-damdam-mad.statcan@canada.ca
Distributor(s)	Statistics Canada

Access Conditions

RDCs provide researchers with access, in a secure university setting, to microdata from population and household surveys. The centres are staffed by Statistics Canada employees. They are operated under the provisions of the Statistics Act in accordance with all the confidentiality rules and are accessible only to researchers with approved projects who have been sworn in under the Statistics Act as "deemed employees". To access the microdata housed in the Research Data Centres (RDCs), apply online.

Citation Requirements

Rights & Disclaimer

Copyright	Copyright © Statistics Canada, 2018

Files Description

Dataset contains 1 file(s)

CCHS_Annual_2017_2018	
# Cases	113290
# Variable(s)	1051

Variables Group(s)

Dataset contains 66 group(s) total - showing a subset of 1

Group GEN: General health

#	Name	Label	Type	Format	Valid	Invalid	Question
5	GEN_020	Perceived life stress	discrete	numeric-1.0	112774	516	Thinking about the amount of stress in your life, would you say that most of your days are...?

Variables Description

Dataset contains 1051 variable(s) total - showing a subset of 10

File : CCHS_Annual_2017_2018

DHH_SEX: Sex

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=113290 / 31274372.01] [Invalid=0 / 0]
Universe	All respondents
Literal question	Is [respondent name] male or female?
Concepts	Sex

Value	Label	Cases	Weighted	Percentage (Weighted)
1	Male	52402	15432670.6	49.3%
2	Female	60888	15841701.4	50.7%
6	Valid skip	0	0.0	
7	Don't know	0	0.0	
8	Refusal	0	0.0	
9	Not stated	0	0.0	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

DHHGMS: Marital status

Information	[Type= discrete] [Format=numeric] [Range= 1-4] [Missing=*]
Statistics [NW/ W]	[Valid=113030 / 31221883.35] [Invalid=260 / 52488.66]
Universe	All respondents
Literal question	What is your marital status? Are you... ?
Concepts	Marital status
Notes	This variable has been grouped as a form of disclosure control.

Value	Label	Cases	Weighted	Percentage (Weighted)
1	Married	45945	14451998.8	46.3%
2	Common-law	11470	3703751.1	11.9%
3	Widowed/Divorced/Separated	23075	3682577.5	11.8%
4	Single	32540	9383556.0	30.1%
6	Valid skip	0	0.0	
7	Don't know	0	0.0	
8	Refusal	0	0.0	
9	Not stated	260	52488.7	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

DHHDGHSZ: Household size - (D)

Information	[Type= discrete] [Format=numeric] [Range= 1-5] [Missing=*]
Statistics [NW/ W]	[Valid=113202 / 31262891.26] [Invalid=88 / 11480.75]
Universe	See documentation on derived variables
Source	Derived from SAMPLEID, PERSONID
Concepts	Household size - (D)
Notes	This variable has been grouped as a form of disclosure control.

Value	Label	Cases	Weighted	Percentage (Weighted)
1	Exact number of persons living in household	32357	4636211.2	14.8%
2	Exact number of persons living in household	40433	10522030.2	33.7%
3	Exact number of persons living in household	15150	5801368.3	18.6%

File : CCHS_Annual_2017_2018

DHHDGHSZ: Household size - (D)

Value	Label	Cases	Weighted	Percentage (Weighted)
4	Exact number of persons living in household	15616	6123841.2	19.6%
5	Grouped - 5 or more persons live in the household	9646	4179440.4	13.4%
6	Valid skip	0	0.0	
7	Don't know	0	0.0	
8	Refusal	0	0.0	
9	Not stated	88	11480.8	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

DHHGAGE: Age

Information	[Type= discrete] [Format=numeric] [Range= 1-16] [Missing=*]
Statistics [NW/ W]	[Valid=113290 / 31274372.01] [Invalid=0 / 0]
Universe	All respondents
Literal question	What is your age?
Concepts	Age
Notes	This variable has been grouped as a form of disclosure control.

Value	Label	Cases	Weighted	Percentage (Weighted)
1	Age between 12 and 14	4478	1163284.4	3.7%
2	Age between 15 and 17	4176	1087783.2	3.5%
3	Age between 18 and 19	1905	871453.0	2.8%
4	Age between 20 and 24	5126	2176384.6	7.0%
5	Age between 25 and 29	6942	2491230.1	8.0%
6	Age between 30 and 34	8155	2714195.8	8.7%
7	Age between 35 and 39	7583	2458687.9	7.9%
8	Age between 40 and 44	7284	2320065.8	7.4%
9	Age between 45 and 49	7042	2387507.2	7.6%
10	Age between 50 and 54	8126	2487554.2	8.0%
11	Age between 55 and 59	9898	2614997.6	8.4%
12	Age between 60 and 64	10177	2476495.6	7.9%
13	Age between 65 and 69	10733	2075076.5	6.6%
14	Age between 70 and 74	8322	1607969.1	5.1%
15	Age between 75 and 79	5800	1076343.3	3.4%
16	Age 80 and older	7543	1265343.6	4.0%
96	Valid skip	0	0.0	
97	Don't know	0	0.0	
98	Refusal	0	0.0	
99	Not stated	0	0.0	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

GEN_020: Perceived life stress

Information	[Type= discrete] [Format=numeric] [Range= 1-5] [Missing=*]
Statistics [NW/ W]	[Valid=112774 / 31132808.26] [Invalid=516 / 141563.75]
Universe	Respondents with DOGEN = 1
Literal question	Thinking about the amount of stress in your life, would you say that most of your days are...?

File : CCHS_Annual_2017_2018

GEN_020: Perceived life stress

Concepts	Perceived life stress

Value	Label	Cases	Weighted	Percentage (Weighted)
1	Not at all stressful	17120	4168831.1	13.4%
2	Not very stressful	29357	7578780.8	24.3%
3	A bit stressful	44311	12734104.0	40.9%
4	Quite a bit stressful	18640	5664055.2	18.2%
5	Extremely stressful	3346	987037.2	3.2%
6	Valid skip	0	0.0	
7	Don't know	451	126067.0	
8	Refusal	65	15496.8	
9	Not stated	0	0.0	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

HWTDGBMI: Body mass index (self-reported) - (D)

Information	[Type= continuous] [Format=numeric] [Range= 12.07-54.42] [Missing=*]
Statistics [NW/ W]	[Valid=105109 / 28915403.93] [Invalid=8181 / 2358968.08] [Mean=26.465 / 26.032] [StdDev=5.325 / 5.144]
Universe	See documentation on derived variables
Source	Derived from ADM_PRX, HWT_005, HWT_020, HWT_025, HWT_030, HWT_035, MAC_025, DHHTAGM, AGET1
Concepts	Body mass index (self-reported) - (D)
Notes	This variable has been grouped as a form of disclosure control. BMI values have been regrouped to a minimum of 12 and a maximum of 58.

Value	Label	Cases	Weighted	Percentage (Weighted)
999.96	Valid skip	25	8753.1	
999.97	Don't know	0	0.0	
999.98	Refusal	0	0.0	
999.99	Not stated	8156	2350215.0	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

SLPG005: Number of hours per night usually spent sleeping

Information	[Type= discrete] [Format=numeric] [Range= 1-11] [Missing=*]
Statistics [NW/ W]	[Valid=55051 / 14620954.95] [Invalid=58239 / 16653417.06]
Universe	Respondents with DOSLP = 1
Literal question	How long do you usually spend sleeping each night?
Concepts	Number of hours per night usually spent sleeping
Notes	Was not asked in proxy interviews. This variable has been grouped as a form of disclo- sure control.

Value	Label	Cases	Weighted	Percentage (Weighted)
1	Under 3 hours	317	63207.9	0.4%
2	3 hours to less than 4 hours	751	160755.3	1.1%
3	4 hours to less than 5 hours	2289	558039.4	3.8%
4	5 hours to less than 6 hours	5632	1481713.2	10.1%
5	6 hours to less than 7 hours	13193	3632169.6	24.8%
6	7 hours to less than 8 hours	17096	4728433.8	32.3%
7	8 hours to less than 9 hours	11806	3055149.7	20.9%
8	9 hours to less than 10 hours	2619	626652.6	4.3%

File : CCHS_Annual_2017_2018

SLPG005: Number of hours per night usually spent sleeping

Value	Label	Cases	Weighted	Percentage (Weighted)
9	10 hours to less than 11 hours	975	236166.3	1.6%
10	11 hours to less than 12 hours	213	42448.4	0.3%
11	12 hours or more	160	36218.6	0.2%
96	Valid skip	56615	16178931.8	
97	Don't know	0	0.0	
98	Refusal	0	0.0	
99	Not stated	1624	474485.3	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ALC_015: Drank alcohol - frequency - 12 mo

Information	[Type= discrete] [Format=numeric] [Range= 1-7] [Missing=*]
Statistics [NW/ W]	[Valid=85179 / 23591176.12] [Invalid=28111 / 7683195.89]
Universe	Respondents who answered ALC_010 = 1
Literal question	During the past 12 months, how often did you drink alcoholic beverages?
Concepts	Drank alcohol - frequency - 12 mo

Value	Label	Cases	Weighted	Percentage (Weighted)
1	Less than once a month	19332	5041980.4	21.4%
2	Once a month	9204	2625351.5	11.1%
3	2 to 3 times a month	11361	3212387.5	13.6%
4	Once a week	13193	3854065.4	16.3%
5	2 to 3 times a week	18136	5240059.2	22.2%
6	4 to 6 times a week	6086	1746680.9	7.4%
7	Every day	7867	1870651.2	7.9%
96	Valid skip	27633	7543161.1	
97	Don't know	321	90915.1	
98	Refusal	37	7241.6	
99	Not stated	120	41878.1	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

LBFDGHPW: Total usual hours worked per week - (D)

Information	[Type= continuous] [Format=numeric] [Range= 0-99] [Missing=*]
Statistics [NW/ W]	[Valid=57675 / 18410259.63] [Invalid=55615 / 12864112.38] [Mean=39.046 / 39.171] [StdDev=14.484 / 13.944]
Universe	See documentation on derived variables
Source	Derived from DHH_AGE, LBF_16, LBF_035
Concepts	Total usual hours worked per week - (D)
Notes	This variable has been capped at 99 as a form of disclosure control.

Value	Label	Cases	Weighted	Percentage (Weighted)
996	Valid skip	54338	12352720.8	
997	Don't know	0	0.0	
998	Refusal	0	0.0	
999	Not stated	1277	511391.6	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

INCDGHH: Total household income - all sources - (D)

Information	[Type= discrete] [Format=numeric] [Range= 1-5] [Missing=*]
Statistics [NW/ W]	[Valid=113131 / 31254706.82] [Invalid=159 / 19665.19]
Universe	See documentation on derived variables
Source	Derived from INC_021
Concepts	Total household income - all sources - (D)
Notes	This variable has been grouped as a form of disclosure control.

Value	Label	Cases	Weighted	Percentage (Weighted)
1	No income or less than $20,000	9890	2041157.5	6.5%
2	$20,000 to $39,999	19260	4028103.1	12.9%
3	$40,000 to $59,999	17985	4446405.8	14.2%
4	$60,000 to $79,999	15085	4015075.8	12.8%
5	$80,000 or more	50911	16723964.5	53.5%
6	Valid skip	0	0.0	
7	Don't know	0	0.0	
8	Refusal	0	0.0	
9	Not stated	159	19665.2	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.