

Name of collection: *Household Internet Use Survey* (survey ceased after 2003)

<p>Nature of collection If possible, use the classification of collection types shown above e.g. <i>ICT use collection – business</i>. For “other” collections, provide details e.g. <i>Other ICT collection – ISPs</i>.</p>	ICT use collection – Household
<p>Collection agency</p>	Statistics Canada http://www.statcan.ca
<p>General references to collection material Metadata, questionnaires etc</p>	Questionnaires and additional information can be found in the survey notes .
<p>Survey basis or vehicle E.g. Labour Force supplement, standalone survey, administrative byproduct data</p>	Labour Force Survey (LFS) supplement
<p>Frequency of collection</p>	Annual
<p>Collection history Reference dates and/or periods from the first to the latest collection</p>	1997-2003 Data for reference year 2004 were not collected. Collection was resumed for Reference Year 2005 with the new Canadian Internet Use Survey (CIUS).
<p>Whether collection is mandatory or voluntary</p>	Voluntary
<p>Scope and coverage of collection Target population in terms of size, industry, population groups etc</p>	Respondents report on behalf of their household. Eligible LFS respondents include all residents of Canada 15 years of age or older excluding: Residents of the Yukon, Northwest Territories and Nunavut, Inmates of Institutions, Persons living on Indian Reserves, Full time members of the Canadian Armed Forces. Exclusions constitute 2% of the population, 15 years of age and older.
<p>Main classifications used E.g. industry, size, commodity, occupation</p>	Use of the Internet at home, work, school, public library
<p>Collection methodology E.g. face-to-face, mail, Web, telephone interview</p>	Computer Assisted Telephone Interview (CATI) with some Computer Assisted Personal Interviews (CAPI)
<p>Reporting and Statistical units Enterprise, establishment, household, etc</p>	Households (Individual reports on behalf of their household).
<p>Sample frame used</p>	For the 2003 Reference Year, the HIUS used four of the six rotation groups in the January 2004 Labour Force Survey (LFS) sample. For the HIUS, the coverage of the LFS was set at the household level. However, unlike the LFS where information is collected for all eligible household members, the HIUS only collected information from one household member who reported about the household.
<p>Sampling method E.g. stratified random sampling, cluster sampling</p>	The LFS sample is based upon a stratified, multi-stage design employing probability sampling at all stages of the design. The design principles are the same for each province.
<p>Sample size</p>	34,674 households

For the most recent collection	
Response rate The responding proportion of the live in-scope population, most recent collection	66.7%
Methods for dealing with non-response (item and unit) Indicate whether imputations are made for non-response and a short description of methods used.	<p>Imputation was limited in the HIUS to item non-response for a few variables. Total non-respondents were dropped from the data file and accounted for in the weighting process. Imputation was performed for the “income” variable and for some of the e-commerce variables.</p> <p>A nearest neighbour imputation procedure was used to find donors from which data were transferred to the records requiring imputation (recipients). Donors were selected using a score function. Certain characteristics were compared between records requiring imputation and all plausible donors. Whenever the recipient and the donor shared the same characteristic, a value was added to the score function. The potential donors with the highest scores were then compared by the way of a distance function involving other collected variables. The record with the smallest distance from the recipient was chosen as the donor.</p> <p>Income Imputation</p> <p>The HIUS collected information on household income. Respondents were asked for a best numerical estimate of household income and, failing that, for the best categorical estimate from among 11 possible categories (from “Less than \$5,000” to \$100,000 or more”). If an estimate was not given, income was coded as missing. Households in the HIUS for which income was coded as missing were linked to the Canadian Travel Survey (CTS), a LFS supplement also conducted in January 2004. In the CTS, respondents were asked for the best estimate of household income among five broad categories (from “Less than \$20,000” to “\$80,000+”). If an estimate was not given, income was coded as missing. Overall, 58% of the households reported income as numerical, 21% as a HIUS category, and 2% as a CTS category. For 19% of the households, no income was available from the HIUS or the CTS.</p> <p>In order to produce income quartiles, categorical and missing income values were imputed to have numerical values. The imputation process was performed in three steps:</p> <ol style="list-style-type: none"> 1) income for a given household reporting a categorical HIUS value was substituted by the income of a household which reported a numerical HIUS value and, according to the score and distance functions, shared the most similar characteristics (eg., hourly earnings, geographic region), provided the numerical value was consistent with the HIUS category; 2) income for a given household reporting a categorical CTS value was substituted by the income of a household which reported a numerical HIUS value or whose income had been imputed via step 1) and shared the most similar characteristics, provided the numerical value was consistent

with the CTS category; and 3) missing income for a given household was substituted by the income of a household which reported a numerical HIUS value or whose income had been converted to a numerical value via step 1) or 2) and shared the most similar characteristics.

E-commerce Imputation

There are two types of e-commerce variables that were imputed:

1) the number of separate orders that the household placed over the Internet, and 2) the cost of these orders.

These variables were collected separately for two different categories; orders which were placed and paid for directly over the Internet with a credit card and those placed, but not paid for over the Internet. The HIUS first collected the total number of orders and the total cost of orders in each category. The HIUS then asked for the number and the cost of these reported orders which were placed with Canadian companies. In total there were eight e-commerce variables requiring imputation; two types of variables (number of orders, cost) for the two categories of orders (paid over the Internet versus paid through other means) for both Canadian companies and all companies. In order to make the imputation process consistent, two additional variables were also imputed.

They were the two introductory questions asking:

1) whether the respondent had placed any orders at all over the Internet which they paid for over the Internet with a credit card, and 2) whether the respondent had placed any orders at all which they did not pay for over the Internet.

Each record with at least one of the ten e-commerce variables of interest with a missing or invalid value was identified as requiring imputation. The imputation process was performed in three stages. In the first two stages, records were imputed which had one or more of the e-commerce variables missing but also had some of the e-commerce variables reported. The first two stages differed in the pattern of responses. The reported e-commerce variables along with variables from other sections of the questionnaire were used, by way of the score and distance functions, to determine the donors.

The pattern of responses and non-response affected the choice of variables included in the score function. The last stage of the imputation dealt with those records which had missing values for all of the e-commerce variables. Information from other sections of the questionnaire was used in the score and distance functions to find the donor. Records that were manually investigated as possible outliers (as described in the section above) were excluded from acting as donors during the imputation of electronic commerce variables. Only those respondents who were usual users of the Internet from any location were eligible for the e-commerce questions. In total, 61% of the HIUS

	<p>respondents were eligible for the e-commerce section. Of those eligible, 6.0% needed at least one of the ecommerce fields to be imputed.</p>
<p>Weighting of results Weighting method e.g. by employment, number of enterprises, revenue</p>	<p>Since the Household Internet Use Survey (HIUS) used a sub-sample of the Labour Force Survey (LFS) sample, the derivation of weights for the survey records is clearly tied to the weighting procedure used for the LFS. In the LFS, the final weight attached to each record is the product of the following factors: the basic weight, the cluster sub-weight, the stabilization weight, the balancing factor for non-response, and the province-age-sex and sub-provincial area ratio adjustment factor.</p> <p>The principles behind the calculation of the weights for the HIUS are nearly identical to those for the LFS. However, this survey is a household-weighted survey, not a person-weighted survey. Also, further adjustments are made to the LFS sub-weights in order to derive a final weight for each record on the HIUS microdata file.</p> <p>Details of the method used to calculate these weights are presented in Chapter 11.0 of the User Guide to the Household Internet Use Survey.</p>
<p>Relative standard errors (or coefficients of variation) on main aggregates For the most recent collection. These can be expressed as a range of values. For a given variable, the RSE or CV is equal to the ratio of the square root of the estimate of the sampling variance to the estimated value. It can be expressed as a fraction or a percentage.</p>	<p>Estimates at the national level for users and non-users of the Internet have a CV of 0.6% and 1.3% respectively. For more information on the calculation of coefficients of variation as well as CV tables, see Chapter 10 of the User Guide to the Household Internet Use Survey.</p>
<p>Known data quality issues with this collection E.g. non-response bias, comparability problems over time, definitional issues, coverage deficiencies, timeliness of frame, high item non-response (identify topics which are particularly problematic).</p>	<p>With the exception of income and electronic commerce (e-commerce) items, item non-response was very low for the HIUS. Most questions had non-response rates which were less than 1.0%.</p> <p>Direct comparisons with data from the most recent collection to 2001 can be made, but not for previous years. The 2001 electronic commerce component of the HIUS was redesigned to capture Internet shopping from households that regularly used the Internet from various locations, solely for household purposes. Previously, household e-commerce data were collected only if the Internet shopping was conducted from home. This constitutes a break in the data series.</p> <p>For more information on data quality, see Chapter 8 of the User Guide to the Household Internet Use Survey.</p>
<p>Output details Please list (or link to) relevant publications for this collection. You can also email relevant files to the OECD.</p>	<p>Annual information on household Internet use is released through The Daily. See also <i>Connectedness Series No. 1, No. 3, No. 4, No. 7, No. 10</i> and Cat. No. 56F0003XIE for research papers and articles, survey information and guides and data tables. See also Canadian Statistics.</p>
<p>Other comments</p>	<p>Data will no longer be collected at the household level. Individuals will report on behalf of themselves, and not their</p>

	households (see Canadian Internet Use Survey).
Contact/s Where available, provide names and email addresses.	Larry McKeown – Senior Analyst larry.mckeown@statcan.ca

Name of collection: *Canadian Internet Use Survey*

Nature of collection If possible, use the classification of collection types shown above e.g. <i>ICT use collection – business</i> . For “other” collections, provide details e.g. <i>Other ICT collection – ISPs</i> .	ICT use collection – Individual
Collection agency	Statistics Canada http://www.statcan.ca
General references to collection material Metadata, questionnaires etc	Questionnaires and additional information can be found in the survey notes .
Survey basis or vehicle E.g. Labour Force supplement, standalone survey, administrative byproduct data	Labour Force Survey (LFS) supplement
Frequency of collection	Biennial
Collection history Reference dates and/or periods from the first to the latest collection	2005 Next data collection scheduled for Reference Year 2007.
Whether collection is mandatory or voluntary	Voluntary
Scope and coverage of collection Target population in terms of size, industry, population groups etc	Respondents report on behalf of themselves. Eligible respondents include all residents of Canada 18 years of age or older excluding: Residents of the Yukon, Northwest Territories and Nunavut, Inmates of Institutions, Persons living on Indian Reserves, Full time members of the Canadian Armed Forces. Exclusions constitute 2% of the population, 18 years of age and older.
Main classifications used E.g. industry, size, commodity, occupation	Location of Internet use (e.g. at home, at work), frequency and intensity of use, specific uses of the Internet from the home, purchase of products and services (electronic commerce), and other issues related to Internet use (such as language of use and concerns over privacy). This content is supplemented by information on individual and household characteristics (e.g. age, income, education, family type) and some geographic detail (e.g. province, urban/rural, and CMA).
Collection methodology E.g. face-to-face, mail, Web, telephone interview	Computer Assisted Telephone Interview (CATI) with some Computer Assisted Personal Interviews (CAPI)
Reporting and Statistical units Enterprise, establishment, household, etc	Individuals report on behalf of themselves (non-proxy reporting).
Sample frame used	This is a sample survey with a cross-sectional design. The CIUS was administered to a sub-sample of the individuals already selected for the Labour Force Survey (LFS). Every

	<p>month, one-sixth of the LFS sample is replaced by a new "birth" panel of dwellings.</p> <p>The CIUS sample begins with the households from the five non-birth panels from the LFS. When contact is made with the households selected by the CIUS, the computer survey application randomly selects one eligible member, 18 years of age or older.</p>
<p>Sampling method E.g. stratified random sampling, cluster sampling</p>	<p>The LFS uses a stratified, multi-stage cluster design employing probability sampling at all stages to select a representative sample of households from Canada's ten provinces, excluding persons living on Indian Reserves, full-time members of the Canadian Armed Forces and inmates of institutions.</p>
<p>Sample size For the most recent collection</p>	<p>30,466 individuals</p>
<p>Response rate The responding proportion of the live in-scope population, most recent collection</p>	<p>70.7%</p>
<p>Methods for dealing with non-response (item and unit) Indicate whether imputations are made for non-response and a short description of methods used.</p>	<p>Once the data are received back at head office an extensive series of processing steps is undertaken to examine each record received. A top-down flow edit cleaned up any paths that may have been mistakenly followed during the interview. The editing and imputation phases of processing identify logically inconsistent or missing information items, and correct such errors related to incomes and on-line expenditures (for example, the amount spent on goods and services from companies in Canada is greater than the total amount spent on-line).</p> <p>In an additional step, suspiciously large reported values for personal or household incomes and the number of Internet orders or their value are identified as "outliers". These few records were treated by replacing the suspicious values by ones from respondents with similar demographic characteristics in the imputation step of processing.</p> <p>Imputation was limited to item non-response for household and personal incomes and for the number of orders and their value for e-commerce spending (from Canadian companies and in total). Imputation was done by the Impudon software using a nearest-neighbour method which searches for "donor" records from individuals with complete and consistent values. The recipient records are imputed by a donor chosen from a group of records with similar demographic characteristics. In the case of income, LFS data on the usual weekly earnings for a recipient and his household (when available) were used to select the donor with the most similar values.</p> <p>In total, over 19,000 respondents (63%) were eligible donors having reported both household and personal incomes. Respondents who did not provide a dollar estimate of their incomes were asked a series of questions in order to derive an income range. Almost 4,500 respondents (15%) did not provide any information on their incomes. The reported</p>

	<p>income ranges and the missing income information were imputed by the donor values in a series of steps, depending on the information available for other variables involved in forming the imputation groups. In a final step, all incomes, reported or imputed, were converted to quartile, quintile and decile categories for the purposes of analysis.</p>
<p>Weighting of results Weighting method e.g. by employment, number of enterprises, revenue</p>	<p>Estimates are produced using weights attached to each sampled unit, or individual. The weight of a sampled individual indicates the number of people in the population that the unit represents. The initial weight was provided by the LFS and incorporated the probability of selecting the individual in their sample, as well as other adjustments such as the treatment of non-response to the LFS.</p> <p>For the CIUS, the final individual weights reflect four adjustments to the LFS initial weight. A first adjustment reflects that only a sub-sample (i.e., 5 of 6 panels) of the LFS was used. A second adjustment accounts for the selection of a single household member, 18 years of age or older. A third adjustment accounts for non-respondents who did not participate in the CIUS but who did in the LFS. A fourth adjustment used Generalized regression (GREG) estimation to calibrate the CIUS weights - matching the age-sex distributions for each province and population counts for several CMAs.</p>
<p>Relative standard errors (or coefficients of variation) on main aggregates For the most recent collection. These can be expressed as a range of values. For a given variable, the RSE or CV is equal to the ratio of the square root of the estimate of the sampling variance to the estimated value. It can be expressed as a fraction or a percentage.</p>	<p>The CIUS estimates are based on a sample of Canadians. The results obtained from asking the same questions of all Canadians would differ to some known extent. The extent of this sampling error is quantified by the coefficient of variation (CV). Given the complexity of the CIUS design, Bootstrap replicate weights were used to calculate the CVs with the following guidelines:</p> <ul style="list-style-type: none"> - 16.5% and below Acceptable; - 16.6% to 33.3% Marginal, with cautionary note; and - Above 33.3% Unacceptable estimate. <p>Estimates that do not meet an acceptable level of quality are either flagged for caution or suppressed. CV tables are prepared by Statistics Canada and made available to help users understand the quality of individual estimates.</p>
<p>Known data quality issues with this collection E.g. non-response bias, comparability problems over time, definitional issues, coverage deficiencies, timeliness of frame, high item non-response (identify topics which are particularly problematic).</p>	<p>A comparison of social and demographic domains from CIUS was made with previous surveys to ensure consistency. Subject matter experts made selective data confrontations with other data sources (e.g. the General Social Survey and the Survey of Household Spending) and with information from regulatory agencies (e.g. the Canadian Radio-television and Telecommunications Commission) and international organizations (e.g. OECD).</p>
<p>Output details Please list (or link to) relevant publications for this collection. You can also email relevant files</p>	<p>Annual information on individual Internet use is released through The Daily (see Canadian Internet Use Survey and E-commerce: Shopping on the Internet). See also Canadian</p>

to the OECD.	Statistics, Connectedness Series, Innovation Analysis Bulletin and the Rural and Small Town Canada Analysis Bulletin . The 2007 Statistics Canada Socio-Economic Conference also includes a session on Indicators and Impacts of a Digital Society featuring CIUS-related research.
Other comments	Next data collection is scheduled for Reference Year 2007.
Contact/s Where available, provide names and email addresses.	Larry McKeown – Senior Analyst larry.mckeown@statcan.ca

Name of collection: *General Social Survey, cycle 14, Access and use of ICT*

Nature of collection If possible, use the classification of collection types shown above e.g. <i>ICT use collection – business</i> . For “other” collections, provide details e.g. <i>Other ICT collection – ISPs</i> .	ICT use collection – Household/individual
Collection agency	Statistics Canada http://www.statcan.ca
General references to collection material Metadata, questionnaires etc	See Questionnaire and survey documentation .
Survey basis or vehicle E.g. Labour Force supplement, standalone survey, administrative byproduct data	General Social Survey
Frequency of collection	One time (GSS cycle 14)
Collection history Reference dates and/or periods from the first to the latest collection	2000 (GSS cycle 14)
Whether collection is mandatory or voluntary	Voluntary
Scope and coverage of collection Target population in terms of size, industry, population groups etc	All persons aged 15 and older living in Canada, excluding residents of the Yukon, Northwest Territories and Nunavut and full-time residents of institutions.
Main classifications used E.g. industry, size, commodity, occupation	None
Collection methodology E.g. face-to-face, mail, Web, telephone interview	Computer-assisted telephone interview (CATI)
Reporting and Statistical units Enterprise, establishment, household, etc	Individual
Sample frame used	All households with a telephone.
Sampling method E.g. stratified random sampling, cluster sampling	Stratified random sampling.
Sample size For the most recent collection	25,000
Response rate The responding proportion of the live in-scope	80.8%

population, most recent collection	
Methods for dealing with non-response (item and unit) Indicate whether imputations are made for non-response and a short description of methods used.	Missing data items coded as “not stated”. Values imputed only for the rare cases of missing sex and number of household telephone lines.
Weighting of results Weighting method e.g. by employment, number of enterprises, revenue	Weighting is the inverse of the probability of selection adjusted for non-response. Used raking ratio procedure, then adjusted within each stratum for each month.
Relative standard errors (or coefficients of variation) on main aggregates These can be expressed as a range of values. For a given variable, the RSE or CV is equal to the ratio of the square root of the estimate of the sampling variance to the estimated value. It can be expressed as a fraction or a percentage.	For estimates of total Internet users and non-users, a CV of 0.6% is obtained. Detailed CV tables can be ordered as part of the Public Use Microdata File Documentation and User's Guide (refer to Contacts below).
Known data quality issues with this collection E.g. non-response bias, comparability problems over time, definitional issues, coverage deficiencies, timeliness of frame, high item non-response (identify topics which are particularly problematic).	None
Output details Please list (or link to) relevant publications for this collection. You can also email relevant files to the OECD.	Data and analysis from the General Social Survey, Cycle 14: Access to and Use of ICT can be found in The Daily, Cat. No. 56-505-XIE and Cat. No. 56F0006XIE . Other analysis has been published in Canadian Social Trends Cat. No. 11-008 .
Other comments	
Contact/s Where available, provide names and email addresses.	Heather Dryburgh, GSS Survey Manager heather.dryburgh@statcan.ca

Name of collection: *General Social Survey – Time use*

Nature of collection If possible, use the classification of collection types shown above e.g. <i>ICT use collection – business</i> . For “other” collections, provide details e.g. <i>Other ICT collection – ISPs</i> .	ICT use collection – Household/Individual
Collection agency	Statistics Canada http://www.statcan.ca
General references to collection material Metadata, questionnaires etc	See Questionnaire and survey documentation .
Survey basis or vehicle E.g. Labour Force supplement, standalone survey, administrative byproduct data	General Social Survey
Frequency of collection	Quinquennial (5 year)
Collection history Reference dates and/or periods from the first to the latest collection	1986, 1992, 1998, 2005 (GSS cycle 19)

Whether collection is mandatory or voluntary	Voluntary
Scope and coverage of collection Target population in terms of size, industry, population groups etc	The target population is non-institutionalized persons 15 years of age or older, living in the ten provinces.
Main classifications used E.g. industry, size, commodity, occupation	None
Collection methodology E.g. face-to-face, mail, Web, telephone interview	Computer-assisted telephone interview (CATI)
Reporting and Statistical units Enterprise, establishment, household, etc	Individual
Sample frame used	All households with a telephone (the samples for most GSS cycles are selected using random digit dialing telephone methods and the interviews are conducted by telephone).
Sampling method E.g. stratified random sampling, cluster sampling	Stratified random sampling.
Sample size For the most recent collection	25,000
Response rate The responding proportion of the live in-scope population, most recent collection	59%
Methods for dealing with non-response (item and unit) Indicate whether imputations are made for non-response and a short description of methods used.	<p>Missing data items are coded as "not stated".</p> <p>Values are imputed only for the rare cases of missing sex and number of household telephone lines.</p> <p>The flow editing carried out by head office followed a 'top down' strategy, in that whether or not a given question was considered "on path" was based on the response codes to the previous questions. If the response codes to the previous questions indicated that the current question was "on path", the responses, if any, to the current question were retained, though "don't know" was recoded as 9 (99 or 999, etc.) and refusals were recoded as "Not Stated", i.e. 8 (98 or 998, etc.). If the response codes to the previous questions indicated that the current question was "off path" because the respondent was clearly identified as belonging to a sub-population for which the current question was inappropriate or not of interest, the current question was coded as "Not Asked", i.e. 7 (97 or 997, etc.).</p>
Weighting of results Weighting method e.g. by employment, number of enterprises, revenue	Weighting is the inverse of the probability of selection adjusted for non-response. Used raking ratio procedure, then adjusted within each stratum for each month. In addition to the estimation weights, bootstrap weights have been created for the purpose of design-based variance estimation.
Relative standard errors (or coefficients of variation) on main aggregates These can be expressed as a range of values. For a given variable, the RSE or CV is equal to	<p>For estimates of total Internet users and non-users, a CV of 0.6% is obtained.</p> <p>Detailed CV tables can be ordered as part of the Public Use</p>

<p>the ratio of the square root of the estimate of the sampling variance to the estimated value. It can be expressed as a fraction or a percentage.</p>	<p>Microdata File Documentation and User's Guide (refer to Contacts below).</p> <p>The methodology of this survey has been designed to control errors and to reduce the potential effects of these. However, the results of the survey remain subject to error due to both sampling error and non-sampling error.</p> <p>Sampling error: As the data are based on a sample of persons they are subject to sampling error. That is, estimates based on a sample will vary from sample to sample, and typically they will be different from the results that would have been obtained from a complete census. The potential range of this difference has been estimated for key data and used to produce tables that can be used to estimate the sampling variability of many estimates. More precise estimates of the sampling variability of estimates can be produced with the bootstrap method using bootstrap weights that have been created for this survey. The bootstrap method was used to estimate the sampling variability for all of the estimates included in 'General Social Survey on Time Use, cycle 19: Overview of Time Use of Canadians'. Estimates with high sampling variability are indicated in this publication and all of the highlighted differences between subgroups of the population are significant at the 95% level.</p> <p>Non-sampling error: Even a census of the population of interest produces estimates subject to error. While these are called non-sampling errors, estimates from samples still contain errors of this type. Common sources of these errors are imperfect coverage, non-response, response errors, and processing errors.</p>
<p>Known data quality issues with this collection E.g. non-response bias, comparability problems over time, definitional issues, coverage deficiencies, timeliness of frame, high item non-response (identify topics which are particularly problematic).</p>	<p>None</p>
<p>Output details Please list (or link to) relevant publications for this collection. You can also email relevant files to the OECD.</p>	<p>Relevant data and analysis from the General Social Survey, Cycle 19: Time use can be found in The Daily, Connectedness Series No. 13, the Innovation Analysis Bulletin and Cat. No. 89-622-XWE and Cat. No. 12F0080XWE.</p>
<p>Other comments</p>	
<p>Contact/s Where available, provide names and email addresses.</p>	<p>Heather Dryburgh, GSS Survey Manager heather.dryburgh@statcan.ca</p>

Name of collection: *International Adult Literacy and Skills Survey*

Nature of collection If possible, use the classification of collection types shown above e.g. <i>ICT use collection – business</i> . For “other” collections, provide details e.g. <i>Other ICT collection – ISPs</i> .	ICT use collection – Household/Individual
Collection agency	Statistics Canada http://www.statcan.ca
General references to collection material Metadata, questionnaires etc	See Questionnaire and survey documentation .
Survey basis or vehicle E.g. Labour Force supplement, standalone survey, administrative byproduct data	International Adult Literacy and Skills Survey
Frequency of collection	Occasional
Collection history Reference dates and/or periods from the first to the latest collection	1994 (International Adult Literacy Survey), 2003
Whether collection is mandatory or voluntary	Voluntary
Scope and coverage of collection Target population in terms of size, industry, population groups etc	The target population is Canadian adults aged 16 and over not residing in institutions or on Aboriginal reserves. In addition to provincial and territorial estimates, the survey was designed to provide reliable estimates for a variety of special target populations such as recent and established immigrants, Francophones in New Brunswick, Manitoba and Ontario, Anglophones in Quebec, Urban Aboriginals in Manitoba and Saskatchewan, Youth in Quebec and British Columbia and Aboriginal residents in the three northern territories.
Main classifications used E.g. industry, size, commodity, occupation	None
Collection methodology E.g. face-to-face, mail, Web, telephone interview	Computer-assisted telephone interview (CATI)
Reporting and Statistical units Enterprise, establishment, household, etc	Individual
Sample frame used	The sampling unit was the household. The sampling frame was the 2001 Census of Population (reference date, May 15th).
Sampling method E.g. stratified random sampling, cluster sampling	A stratified multi-stage probability sample design was used to select the sample from the census frame.
Sample size For the most recent collection	A base sample of 16,000 dwellings was selected to cover the general population. An additional 24,000 dwellings were selected in supplementary samples targeting the following subpopulations: Francophones in New Brunswick, Ontario and Manitoba; Anglophones in Québec; immigrants in Québec, Ontario, Alberta and British Columbia; youth (16-24) in Québec, youth (16-29) in British Columbia; Aboriginal persons in urban Manitoba and Saskatchewan; Aboriginal persons and non-Aboriginal persons in Yukon, the Northwest

	Territories and Nunavut. The survey covered more than 23,000 individuals.
Response rate The responding proportion of the live in-scope population, most recent collection	The response rate was 65.6% nationally, with provincial rates ranging from 59.5% to 75.4%.
Methods for dealing with non-response (item and unit) Indicate whether imputations are made for non-response and a short description of methods used.	<p>Missing data were imputed for a small number of variables using the Impudon software. Nearest neighbour imputation was used, which involves copying information from a "donor" record with similar characteristics.</p> <p>The task language was imputed for respondents who did not complete a task booklet; the imputation was performed for approximately 9% of records. Exact personal income and household income were imputed for records which reported a range for income rather than a value. As well, age, level of education, year of immigration and Aboriginal status were imputed for a very small number of records, fewer than 0.1% of records overall.</p>
Weighting of results Weighting method e.g. by employment, number of enterprises, revenue	<p>The weights were calculated in several steps:</p> <ol style="list-style-type: none"> 1) An initial weight was calculated based on the probability of selecting the unit in the sample. 2) The weights were adjusted to account for non-response. 3) The weights from the base sample and the various supplementary samples were integrated using a multiple-frame method. 4) Generalized regression (GREG) estimation was used to calibrate the weights and to make them agree with external population counts.
Relative standard errors (or coefficients of variation) on main aggregates These can be expressed as a range of values. For a given variable, the RSE or CV is equal to the ratio of the square root of the estimate of the sampling variance to the estimated value. It can be expressed as a fraction or a percentage.	<p>The quality of the estimates is assessed using estimates of their coefficient of variation (CV). Given the complexity of the IALSS survey design, CVs cannot be calculated using a simple formula. Jackknife replicate weights are used to establish the CVs of the estimates.</p> <p>The following guidelines are recommended:</p> <ul style="list-style-type: none"> - If the CV is less than 16%, the estimate can be used without restriction; - If the CV is between 16% and 33%, the estimate should be used with caution; - If the CV is 33% or more, or if the estimate is based on fewer than 30 observations, then the estimate should not be released.
Known data quality issues with this collection E.g. non-response bias, comparability problems over time, definitional issues, coverage deficiencies, timeliness of frame, high item non-response (identify topics which are particularly problematic).	None

<p>Output details Please list (or link to) relevant publications for this collection. You can also email relevant files to the OECD.</p>	<p>Relevant data and analysis from the IALSS: International Report, National Data, National Report, The Daily, Connectedness Series No. 12, and the Innovation Analysis Bulletin.</p>
<p>Other comments</p>	<p>The 2003 International Adult Literacy and Skills Survey (IALSS) is the Canadian component of the Adult Literacy and Life Skills Survey (ALL). The main purpose of the survey was to find out how well adults used printed information to function in society. Survey data include background information (demographic, education, language, labour force, training, literacy uses, information and communication technology, income) and psychometric results of respondents' proficiency along four skill domains: prose and document literacy, numeracy and problem-solving.</p>
<p>Contact/s Where available, provide names and email addresses.</p>	<p>Client Services (1-800-307-3382; 613-951-7608; fax: 613-951-9040; educationstats@statcan.ca), Culture, Tourism and the Centre for Education Statistics.</p>

Name of collection: *Aboriginal Peoples Survey, module on access and use of ICT*

<p>Nature of collection If possible, use the classification of collection types shown above e.g. <i>ICT use collection – business</i>. For “other” collections, provide details e.g. <i>Other ICT collection – ISPs</i>.</p>	<p>ICT use collection – Household/individual</p>
<p>Collection agency</p>	<p>Statistics Canada http://www.statcan.ca</p>
<p>General references to collection material Metadata, questionnaires etc</p>	<p>See survey notes and questionnaires: APS Questionnaires See also Concepts and Methods Guide</p>
<p>Survey basis or vehicle E.g. Labour Force supplement, standalone survey, administrative byproduct data</p>	<p>Standalone survey; Sample drawn from 2001 Census</p>
<p>Frequency of collection</p>	<p>Occasional</p>
<p>Collection history Reference dates and/or periods from the first to the latest collection</p>	<p>2001</p>
<p>Whether collection is mandatory or voluntary</p>	<p>Voluntary</p>
<p>Scope and coverage of collection Target population in terms of size, industry, population groups etc</p>	<p>The target population comprises adults and children living in private dwellings in the 10 provinces and three territories who are North American Indian, Métis or Inuit, and/or are a Treaty Indian or a Registered Indian as defined by the Indian Act of Canada and/or are members of an Indian Band or First Nation and/or who have Aboriginal ancestry. All residents of collective dwellings are excluded from the survey.</p>
<p>Main classifications used E.g. industry, size, commodity, occupation</p>	
<p>Collection methodology E.g. face-to-face, mail, Web, telephone interview</p>	<p>Telephone contact for screening portion of questionnaire; all components of the survey were interviewer administered and a pencil and paper methodology was employed (A proportion</p>

	of interviews were completed on the telephone).
Reporting and Statistical units Enterprise, establishment, household, etc	Individuals
Sample frame used	<p>APS is a post-censal survey, meaning that respondents were selected based upon their responses to the 2001 Census.</p> <p>Four questions from the Census Long Questionnaire were used to determine the APS target population. The target population comprised all persons who gave a positive answer to question 18 (Aboriginal self-reporting question) OR question 20 (member of an Indian Band/First Nation) OR question 21 (Treaty or Registered Indian) OR who indicated they had Aboriginal ancestry in question 17 on the 2001 Census.</p> <p>Collection for the 2001 APS was conducted in two phases:</p> <ul style="list-style-type: none"> • Phase I, which took place from October to December 2001, focused on the “Aboriginal identity population”. This population is defined as all individuals with a positive answer to the Aboriginal identity in question 18 or Band/First Nation membership in question 20 or Registered Indian status in question 21 on the 2001 Census. • Phase II, which took place from April to June 2002, focused on people who reported in question 17 that they had Aboriginal ancestry, but who did not report Aboriginal identity in question 18 or Band/First Nation membership in question 20 or Treaty or Registered Indian status in question 21. This population is referred to as the “Aboriginal origin only population”. A small portion of the Aboriginal identity population was also covered in Phase II.
Sampling method E.g. stratified random sampling, cluster sampling	<p>Two sample designs were used in Phase I, one for Aboriginal people living “on-community” and another one for Aboriginal people living “offcommunity”. The term “on- community” includes the following participating communities:</p> <ul style="list-style-type: none"> • 123 First Nations communities (reserves) • 53 Inuit Communities in Arctic regions • 38 communities with a minimum Aboriginal population of 250 with a concentration of 40% or more Aboriginal people. 28 of these (including 8 Métis settlements in Alberta) have high concentrations of Métis people • 5 additional communities with a large number of Aboriginal people. <p>The term “off-community” excludes the areas mentioned above and consists of all other non-reserve rural and urban areas.</p> <p>The sampling design Phase I for all communities can be described as a two-phase stratified design. The 2001 Census long form sample was used in the first phase to identify Aboriginal people. Individuals were divided into groups (“strata”) by cross-classifying the communities by adults and children. In the second phase, a sample was selected in each stratum using a systematic random sampling design.</p>

	<p>The sample design for the Phase I – off-community portion can be considered a two-stage stratified design. In the first stage, EAs or groups of EAs (primary sampling units (PSUs)) were selected. In the second stage, all long form respondents (in the Aboriginal group for which the PSU was selected in the first stage) were selected.</p> <p>The sample was stratified by what is called the “domains of estimation”. The domains of estimation correspond to geographical regions for which estimates with an “acceptable” level of precision for a particular Aboriginal group (i.e. North American Indian ((NAI)), Métis, Inuit) are required. Once the strata/domains had been defined, a sample of PSUs was selected within each domain.</p> <p>Phase II. This portion of the survey used a two-phase stratified sample design. The first phase consisted of sampling about one in five households in each EA and the second phase consisted of taking a subsample of Aboriginal people selected in the first phase. The stratification used differed for the ancestry and identity samples and consisted of combinations of Aboriginal groups, subprovincial or provincial regions and adults/children. A systematic sample was taken independently from each stratum after the individuals had been sorted according to key variables such as subprovincial region (when the stratum was the province), age group, gender, and so on.</p>
<p>Sample size For the most recent collection</p>	<p>Phase I (on-community) 57,560 Phase I (off-community) 45,710 Phase II 13,971 Total 117,241</p>
<p>Response rate The responding proportion of the live in-scope population, most recent collection</p>	<p>Phase I (on-community) 87.9% Phase I (off-community) 84.1% Phase II 68.6% Total 84.1%</p>
<p>Methods for dealing with non-response (item and unit) Indicate whether imputations are made for non-response and a short description of methods used.</p>	<p>For APS, generally the only type of imputation done was deterministic imputation. Questions related to each other were edited simultaneously. Valid responses were imputed for missing responses if sufficient information was available in the related questions. Otherwise, they were coded to "not stated".</p> <p>The filter questions (i.e., questions 1 to 4 in the Identification section) and the date of birth and sex variables were sometimes imputed from the Census. The filter questions were NOT imputed for records corresponding to non-respondents. For respondents, question #1 was always cleaned using Census information. If a record had enough of the questionnaire completed to keep it BUT did NOT have a "yes" to one of the filters AND one or more of the filters were blank, Census responses were imputed (both "yes" and "no") in all blank questions.</p>

	Census data was also used to correct missing or invalid entries in the date of birth and sex fields.
<p>Weighting of results Weighting method e.g. by employment, number of enterprises, revenue</p>	<p>The weights were calculated in a three-stage process. The 1st stage was the assignment of an initial weight based on the sampling design. The initial weight was simply the inverse of the inclusion probability (probability of falling in the sample). The 2nd stage of the weighting process was the adjustment for nonresponse. The 3rd stage of the weighting adjustment was the post-stratification. This adjustment ensures that the sum of the final weights for the respondents is equal to the population counts from the Census. The adjustment was done for groups (called post-strata) defined by the combination of several variables.</p> <p>For more detail on weighting procedures, see Estimation: 2001 APS.</p>
<p>Relative standard errors (or coefficients of variation) on main aggregates These can be expressed as a range of values. For a given variable, the RSE or CV is equal to the ratio of the square root of the estimate of the sampling variance to the estimated value. It can be expressed as a fraction or a percentage.</p>	<p>Not available. For more information about CVs, please refer to section 8.1 of Concepts and Methods Guide.</p>
<p>Known data quality issues with this collection E.g. non-response bias, comparability problems over time, definitional issues, coverage deficiencies, timeliness of frame, high item non-response (identify topics which are particularly problematic).</p>	<p>Coverage errors occur when there are differences between the target population and the population sampled. Because the APS sample is selected from those who participated in the Census, the APS information is not available for those communities that were incompletely enumerated in the Census. In 2001, 30 Indian reserves and settlements were incompletely enumerated in the Census, and 22 communities, the majority of which were Indian reserves, were incompletely enumerated in APS.</p> <p>Differences in counts between the 2001 APS and the 2001 Census: While the post-stratification ensured that the total number of people with Aboriginal origins or Aboriginal identity is the same for the Census and APS, it did not ensure the counts for the Aboriginal groups would match. Indeed, the Census and APS produce different counts at the group level. This is due to changes in the way that respondents answered questions about their Aboriginal origins and Aboriginal identity from the time of the Census to the time of APS. Respondents may have changed their responses for a number of reasons, including: different modes of interview, different questionnaires, different context, and coverage and sampling methodology.</p>
<p>Output details Please list (or link to) relevant publications for this collection. You can also email relevant files to the OECD.</p>	<p>See Off-reserve Aboriginal Internet users</p> <p>Aboriginal Peoples Survey community profiles for children and youth</p> <p>Aboriginal Peoples Survey: Well-being of the non-reserve Aboriginal population;</p> <p>Aboriginal Peoples Survey: Children who live in non-reserve areas</p>

<p>Other comments</p>	<p>A profile of the complete reserve population cannot be done as the way the reserves were selected to participate, means that there is not a representative sample of the Canadian on-reserve population (analyses of on-reserve Aboriginals can only be done by individual communities).</p>
<p>Contact/s Where available, provide names and email addresses.</p>	<p>Francine Lavoie francine.lavoie@statcan.ca</p>