2019 Survey of Financial Security: Public Use Microdata File User Guide

Centre for Income and Socio-economic Well-being Statistics
Statistics Canada
Ottawa, ON K1A 0T6
Telephone: 613.951.7355
STATCAN.income-revenu.STATCAN@canada.ca

Statistics Canada Statistique Canada

Canada

Table of contents

Contents

1.0 Introduction 4

1.1 Overview 4

1.2 History 4

1.3 How to cite SFS in publications 4

2.0 Key findings 5

3.0 Concepts and definitions 8

3.1 Net worth 9

3.2 Assets 9

3.3 Debts 11

3.4 Family type 12

4.0 Survey methodology 13

4.1 The survey universe 13

4.2 Survey content and reference period 13

4.3 The sample 14

4.4 Data collection 14

4.5 Data processing and quality control 14

4.6 Imputation of missing data 14

4.7 Weighting 16

5.0 Data accuracy and quality 17

5.1 Sampling errors 17

5.2 Measures of Sampling Error 17

5.3 Non-sampling errors 20

5.3.1 Coverage errors 20

5.3.2 Response errors 20

5.3.3 Non-response errors 20

5.3.4 Processing errors 21

5.3.5 Estimation errors 21

5.3.6 Analysis error 21

5.4 Treatment of large values 21

5.5 Impact of sampling and non-sampling errors on SFS estimates 21
5.6 Comparability of data and related sources 22
5.7 Response rates 22
6.0 Record layout, data dictionary and univariate distributions 23
6.1 Columns of the record layout 23
6.2 Data dictionary with frequencies 23
6.3 Reserved codes 23
7.0 Guidelines for analysis and presentation 25
7.1 Applying weights 25
7.2 Rounding guidelines 25
7.3 Missing values and reserved codes 25
8.0 Guidelines for release 26
8.1 Introduction 26
8.2 Minimum sizes of estimates for release (Guideline for release) 26
8.3 Statistical Packages for Variance Estimation 27
8.4 Hypothesis tests provided by standard statistical software packages 27
9.0 Appendix A – SFS production totals VS PUMF totals 28
10.0 Appendix B – SFS production Quality Indicators VS PUMF Quality Indicators 29
11.0 Appendix C – SFS totals 30

1.0 Introduction

1.1 Overview

The 2019 Survey of Financial Security (SFS) provides a comprehensive picture of the net worth of Canadians. Information was collected on the value of all major financial and non-financial assets and on the money owing on mortgages, vehicles, credit cards, student loans and other debts. The value of these assets less the debts is referred to as net worth.

The public use microdata file (PUMF) is a collection of data on the incomes, assets, debts, wealth, and financial behaviours and attitudes of Canadian families. This file contains information collected from more than 10,000 family units residing in private households in Canada. All records have been thoroughly screened to ensure the anonymity of respondents.

This manual was produced as a reference guide to help users manipulate the microdata file of the survey results.

For more information, contact us toll-free at 1-800-263-1136 or 514-283-8300;
STATCAN.infostats-infostats.STATCAN@canada.ca.

1.2 History

Since the 1950s, Statistics Canada has conducted occasional surveys on the assets and debts of Canadians. Up to, and including, 1984, these surveys were supplements to the more regular income surveys, known as the Surveys of Consumer Finances (SCF). In 1999, the assets and debts component of the SCF was replaced by the Survey of Financial Security. There were other iterations of this survey in 2005, 2012, 2016 and most recently in 2019.

Over the three-year period between 2016 and 2019, a number of important factors influenced the evolution of the wealth distribution in Canada. The real estate market experienced strong growth over the period, with historically low interest rates and favorable economic conditions spurring the value of existing homes.

With the cost of borrowing at all-time lows, consumer debt continued to rise to unprecedented levels, while debt service ratios nonetheless have not increased in relation to income.

1.3 How to cite SFS in publications

"This analysis is based on Statistics Canada's Survey of Financial Security Public Use Microdata, 2019, which contains anonymous data collected in the Survey of Financial Security. All computations on these microdata were prepared by (Name of user). The responsibility for the use and interpretation of these data is entirely that of the author(s)".

2.0 Key findings

The 2019 survey results showed that:

Families in Canada are currently experiencing a profound disruption in their lives as a result of the COVID-19 pandemic, with many facing job loss and financial uncertainty. The pandemic has led to considerable work interruptions in Canada and around the world since mid-March 2020.

While the estimates from the 2019 Survey of Financial Security do not reflect the impacts of the COVID-19 pandemic, they do provide a baseline for assessing developments during the pandemic and shine light on which Canadian families may be more or less financially vulnerable. For example, statistics on net worth, indebtedness of families and their liquid financial asset holdings show that some families were more likely to be vulnerable than others. The data also allow an examination of the net worth of families by age and family structure.

The median net worth of Canadian families was $329,900 in 2019. The growth in net worth was slower over the 2016 to 2019 period compared with earlier in the decade.

On an annualized basis, the net worth of Canadian families was up 1.8% per year from 2016 to 2019, a growth rate that was not large enough to be statistically significant. By contrast, from 2012 to 2016, net worth grew 3.5% per year.

Canadian families include families of two or more persons, as well as unattached individuals. Net worth is the difference between a family's assets and debts.

Housing is both the largest asset and the largest debt for Canadians as has been the case since this survey was first launched in 1999. In 2019, about three-fifths (61.9%) of Canadian families reported a principal residence as an asset with a median value of $400,000, while about one-third (34.6%) reported holding a mortgage on their principal residence with a median outstanding value of $180,000.

The next largest asset was employer-sponsored pension plans (EPP). Just over half of Canadian families reported having an EPP in 2019 with a median value of $164,900.

Overall, just under one-third (30.2%) of Canadian families were debt-free in 2019.Families where no member had an employer pension plan, families who were renters, lone-parent families, younger families and unattached non-seniors had lower net worth than others.

Net worth highest in Ontario and British Columbia

Families in Ontario reported a median net worth in 2019 at $434,500, while those in British Columbia reported $423,700.

Families in Vancouver ($521,500) and Toronto ($467,900) had the highest median net worth among the eight largest census metropolitan areas. Québec ($352,800), Ottawa–Gatineau ($348,000), Calgary ($336,100), Winnipeg ($336,100) and Edmonton ($308,800) each had similar levels of net worth. Montréal had the lowest ($220,200).
Geographical differences in net worth can often be linked to strong housing markets in some parts of Canada. For example, the median value of principal residences in Vancouver rose from $366,000 in 1999 to $900,000 in 2019—over double the median value reported in Montréal ($350,000 in 2019 from $175,700 in 1999).

Seniors less likely to be debt-free during retirement compared with two decades ago

While net worth remained high in 2019, some groups reported being more financially vulnerable than in the past two decades. For example, fewer Canadians reported being debt-free during their retirement years. Almost one-third (30.2%) of Canadian families were debt-free in 2019. While senior-led families (56.7%) were most likely to be debt-free, this rate was down from 72.6% in 1999. Over 1 in 10 senior-led families (12.1%) still had a mortgage on their principal residence in 2019, compared with 6.6% in 1999. Furthermore, more senior-led families reported owing on a line of credit, carrying a credit card balance from month to month or having installment debt in 2019 (27.5%), than was the case in 1999.

Families in pre-retirement years were also more likely to be carrying debt in 2019 than in previous decades. Over one-quarter of Canadian families (28.1%) where the principal earner was aged 55 to 64 were debt-free in 2019, down 11.0 percentage points from 20 years earlier (39.1% in 1999).

Lone-parent families had low median net worth

The net worth of Canadian families also varied among family types, with lone-parent families and unattached non-seniors having lower net worth.

Lone-parent families reported median net worth of $83,100 in 2019, which was less than one-fifth of the median net worth of couples with children ($435,700). Lone-parent families were less likely to own their own home, have pension assets or to own a vehicle compared with couples with children.

Not surprisingly, given that net worth tends to increase with age, senior-led families reported the highest median net worth in 2019 ($840,900), while non-senior led couple-only families reported a median net worth of $459,400. Senior-led families had more pension and non-pension financial assets and less mortgage and vehicle debt than non-senior couple-only families.

Unattached seniors reported a median net worth of $322,300 in 2019, over six times higher compared with unattached non-seniors ($51,000).

Lone-parent families also held the least amount of liquid financial assets

In 2019, the median value of liquid financial assets held by families was $27,700. Senior-led families held $107,000 in these assets, while lone-parent families held $5,500.

Liquid financial assets are all assets held in chequing and saving accounts, term deposits, treasury bills, tax-free savings accounts, stocks and bonds (in mutual funds or not), and registered retirement savings plans. Liquid financial assets can provide some relief for families faced with a disruption to their primary source of income, so they can continue to meet their essential needs and financial obligations. Therefore, families with lower liquid financial assets may be more vulnerable to a disruption in income than families with higher liquid financial assets.

Three-quarters of families (75.0%) held sufficient liquid financial assets to replace one month of their own after-tax family income. Nearly two-thirds (60.5%) reported having at least three months-worth, while just under half (49.5%) reported having enough liquid financial assets to cover at least six months of their after-tax income.
Some families had little or no liquid financial assets to replace their income. In 2019, almost half of lone-parent families (45.7%) reported having less than enough to cover one month of their after-tax income, followed by unattached non-seniors (31.0%).

The median net worth for families with an employer-sponsored pension plan nearly seven times higher than for those without one

In 2019, just over half of Canadian families (50.4%) had at least one member with a current, deferred or in-pay EPP. Over half of the families where the major income earner was aged 35 or older had an EPP (ranging from 51.9% to 57.0% depending on the age cohort), compared with about one-third (35.7%) of families where the major income earner was under the age of 35.

The median net worth for those families with an EPP was $633,300 in 2019, nearly seven times higher than for those without an EPP ($91,200).

Families with an EPP had higher median asset values for every asset category except bonds and equity in business compared with those without an EPP. Those with an EPP also had higher median debt in every debt category except mortgage on principal residence, where both groups had a median of $180,000.

Renters nearing retirement had lower net worth than homeowners

Homeowners reported a higher median net worth ($685,400) compared with renters ($24,000).

While homeowners tend to be older than renters and therefore have had more time to accumulate net worth, a difference in median net worth between homeowners and renters remains when controlling for age.

For those nearing retirement age (55 to 64), the median net worth of homeowners was $952,100, compared with $40,000 for renters. This suggests that renters in this age group have not accumulated a large nest-egg in preparation for retirement.

Likewise, few young renters have accumulated as much net worth as their homeowner counterparts. Median net worth of homeowners under 35 years of age was $272,100, while median net worth of renters was $14,500 for the same age group. In 2019, 1 in 20 young renters (5.0%) had accumulated as much net worth as the median young homeowner.

3.0 Concepts and definitions

This chapter outlines the definitions of the main assets, debts and wealth concepts and their components.

Table 3-1 below illustrates the components of the net worth calculation accounted for by the Survey of Financial Security. The value of all assets less all debts is net worth. A family's net worth can be thought of as the amount of money they would have if they liquidated their assets and paid off all of their debts. The PUMF variable names appear in brackets.

Table 3-1 Components of net worth calculation

Assets **(PWATOTPT PWATOTPG)**	**less:**	**Total debts** **(PWDTOTAL)**	**equals**	**Net worth** **(PWNETWPT PWNETWPG)**
RRSPs/LIRAs **(PWARRSPL)** RRIF **(PWARRIF)** Employer Pension Plans (EPP) **(PWARPPT, PWARPPG)** Retirement funds, other **(PWAOTPEN)** Deposits in financial institutions **(PWASTDEP)** Mutual funds and other investment funds **(PWASTMUI)** Stocks **(PWASTSTK)** Bonds **(PWASTBND)** TFSAs **(PWATFS)** Other financial assets, non-pension **(PWASTOIN)** Principal residence **(PWAPRVAL)** Other real estate **(PWASTRST)** Vehicles **(PWASTVHE)** Other non-financial assets **(PWASTONF)** Equity in business **(PWBUSEQ)**		Mortgage on principal residence **(PWDPRMOR)** Mortgage on other real estate **(PWDSTOMR)** Line of credit **(PWDSTLOC)** Credit card and installment debt **(PWDSTCRD)** Student loans **(PWDSLOAN)** Vehicle loans **(PWDSTVHN)** Other debt **(PWDSTODB)**		

3.1 Net worth

The net worth (sometimes referred to as wealth) of a family unit is defined as the difference between the value of its total asset holdings and the amount of total indebtedness.

There are two types of net worth variables:

1. **PWNETWPT** – Net worth of the family unit. (Assets including current pensions valued on termination basis[1] (PWATOTPT) - debts (PWDTOTAL).)
2. **PWNETWPG** – Net worth of the family unit. (Assets including current pensions valued on going concern basis[1] (PWATOTPG) - debts (PWDTOTAL).)

Respondents were asked to provide the value of the asset or the amount of the debt at a time as close as possible to the date of the interview. Assets and debts were reported for the family unit as a whole and not for each person in the family. The assets and debts included in the survey are identified below.

3.2 Assets

Respondents were asked to report the market value of the asset that is the amount they would receive if they had sold the asset at the time of the survey. If available, respondents were encouraged to consult financial records. When the value could not be determined through an independent source, the respondent was asked to estimate the value. This is in itself prone to error. Values provided by respondents were not adjusted unless they were judged to be an error, resulting, for example, from data entry. If the respondent either over or underestimated the value of an asset by a relatively small proportion, this would not be readily apparent. However, extreme values were reviewed and adjusted if necessary.

The value of all invested assets was to include accrued earnings or interest. Respondents were asked to estimate the actual value, at the time of the survey. In one case, for the value of the contents of the principal residence, the respondent was able to select one of 16 ranges.

The definitions of the assets items identified in table 3-1 are:

Assets, total (PWATOTPT, P WATOTPG): Total value of all financial assets, non-financial assets and equity in business.

There are two types of total asset variables:

1. **PWATOTPT** - Total assets, including employer pension plans (current plans valued on termination[1] basis).
2. **PWATOTPG** - Total assets, including employer pension plans (current plans valued on a going concern[1] basis)

Bonds (PWASTBND): Bonds are the total value, including earnings, of federal and provincial savings bonds and other bonds issued by governments and corporations. Includes investment in foreign bonds but excludes the amount held within registered plans.

Deposits (PWASTDEP): Deposits are the total amount of all chequing and savings accounts and of other deposits such as term deposits, Guaranteed Investment Certificates and treasury bills. These amounts would generally be held in financial institutions such as chartered banks, trust companies, co-ops and caisses populaires. This item includes only the amount held outside of registered plans.

Employer pension plans (**PWARPPT, PWARPPG**): An employer pension plan (EPP) is an employer-sponsored

[1] Employer pension plan valuation is explained further in this chapter.

plan registered with Canada Revenue Agency and most commonly also with one of the pension regulatory authorities. The purpose of such plans is to provide employees with a regular income at retirement.

There are two commonly used approaches to valuing EPP assets: the **going concern** and the **termination** approach. The two EPP variables included on the PUMF are:

1. **PWARPPT** – Current pension plans valued on a termination basis.
2. **PWARPPG** – Current pension plans valued on a going concern basis

The main differences between the two valuation methods are:

(a) Although future service is not considered in either type of valuation, in a going concern valuation assumptions are made about future salary increases. As many EPPs base the amount of the pension on average earnings close to the time of retirement, assuming salary increases up to that time will obviously increase the value of the benefit. In a termination valuation, salary increases are not considered.

(b) Interest rates for a termination valuation are assumed based on current market rates. For a going concern valuation longer term interest rates are assumed.

(c) The going concern valuation method is applicable only for current members of certain types of EPPs. Those with deferred pensions (people who had previously belonged to an EPP) and those receiving benefits are no longer members of the plan so future salary increases need not be considered.

When analyzing SFS data the termination valuation approach is generally used. That approach is more consistent with the basis on which other assets are valued, in that future expectations are not taken into consideration and current market conditions are used to estimate the value. The termination approach, however, can underestimate the value of the benefit earned (accrued) as of the time of the survey because many employees will continue to participate in the plan, and therefore receive a pension based on their salary closer to the time of retirement. In order to allow users the option of selecting the value of the EPP that is most appropriate for their type of analysis both values have been produced and are available.

In valuing benefits for those respondents who belonged to a pension plan at the time of the survey, only plan membership up to the time of the survey has been considered. Therefore, in the case of a person who was 45 at the time of the survey and who had participated in an EPP for 10 years, the pension would be valued for the 10 years of known service.

For more information on employer pension valuation see M. Cohen, H. Frenken and K. Maser, *Survey of Financial Security: Methodology for estimating the value of employer pension plan benefits,* Statistics Canada, Catalogue 13F0026MIE-01003. It can be found on the Statistics Canada website at www.statcan.gc.ca/pub/13f0026m/13f0026m2001003-eng.pdf.

Equity in business (PWBUSEQ): The estimated amount the respondent would receive if the business were sold, after deducting any outstanding debts to be paid.

Locked-in Retirement Accounts (included in **PWARRSPL**): A Locked-In Retirement Account (LIRA) is an RRSP in which the money is locked-in until the person reaches a specified age. LIRAs are included in the RRSP category. This money would have been transferred from an employer pension plan after the individual terminated employment. For the most part, LIRAs came into use in the late 1980s, when revisions to pension regulatory legislation provided for enhanced portability of pension accruals on termination of employment.

Mutual funds and other investment funds (PWASTMUI): The total value, including investment earnings, of all holdings in mutual and investment funds. Excludes the amount held within registered plans.

Principal residence (**PWAPRVAL**): Market value, as estimated by the respondent, of the residence where the respondent lives. If the respondent has two residences, this would be the one where they most often live. If the respondent shares ownership of the home with someone outside the family, only the family's share is included. If

the property is a farm, the estimated value of the farmhouse is included; the value of the farmland would be included either with business equity or with other real estate, if no business were reported.

Real estate, other (PWASTRST): Estimated market value of real estate other than the respondent's principal residence. Included would be second homes, vacation homes, timeshares, rental property (residential or non-residential) or vacant lots. Real estate includes property in Canada or outside.

Registered Retirement Savings Plans (included in **PWARRSPL**): A Registered Retirement Savings Plan (RRSP) is a capital accumulation program designed to encourage saving for retirement. Contributions are tax-deductible within prescribed limits. Investment income earned in the RRSP is tax exempt, but benefits are taxable.

The RRSP could be held in deposits, mutual funds, stocks or bonds. As well, this includes the amount held in Locked-in retirement accounts (LIRAs); see definition above.

Registered retirement income funds (PWARRIF): A Registered Retirement Income Fund (RRIF) is intended to provide a regular income in retirement. Monies in RRSPs must be transferred to a RRIF or an annuity before the end of the year in which the owner of the RRSP turns 71. Payments from an RRIF may be varied, but a minimum amount must be withdrawn annually. Also includes monies in locked-in retirement income funds (LRIFs) and life income funds (LIFs); these plans are intended to receive amounts transferred from an employer pension plan.

Stocks (PWASTSTK): Total value, including earnings, of all publicly-traded common and preferred shares. Includes foreign stock but excludes the amount held within registered plans.

Tax free savings accounts (PWATFS): A TFSA is an account that lets deposits grow through tax-free compounding. No income tax is paid on investment returns earned in the account, and there are no taxes on the amounts withdrawn. Any Canadian resident 18 and over with a social insurance number can open a TFSA. The TFSA maximum annual contribution limit has changed several times since it was introduced in 2009. As of 2019, the cumulative maximum contribution limit was $63,500.

Vehicles (PWASTVHE): Estimated value of cars, trucks, vans, sport utility vehicles as well as motorcycles, mobile homes, boats and snowmobiles. Excludes vehicles owned by the respondent's business and vehicles that are leased.

3.3 Debts

The amount reported for debts is not intended to include interest owing, as this would most often not be known.

The debt items listed in table 3-1 comprise the following:

Debts, total (PWDTOTAL): Total of all debts for the family unit.

Credit card and installment debt (PWDSTCRD): For credit cards, the amount owing on the last bill, excluding any new purchases. Includes major credit cards (VISA, MasterCard, American Express) and retail store cards, gasoline station cards, etc. Installment debt is the total amount owing on deferred payment or installment plans where the purchased item is to be paid for over a period of time.

Line of credit (PWDSTLOC): Total amount owing on both a home equity line of credit and a regular line of credit. This does not refer to the credit limit on the LOC.

Mortgages, on principal residence (PWDPRMOR): Outstanding amount owing on the respondent's principal residence. If the respondent shares ownership of the home with someone outside the family, only the family's share of the mortgage is included. If the property is a farm, the mortgage owing on the farmhouse is included; the mortgage on the remainder of the farm would implicitly be included with business equity or would be included with mortgage owing on other real estate, if no business were reported.

Mortgages, on other real estate (PWDSTOMR): Respondent's share of the mortgage owing on second homes, vacation homes, timeshares, rental property (residential or non-residential) or vacant lots.

Student loans (PWDSLOAN): Amount owing on loans taken out to attend a post secondary education program. These loans are most often taken through the Canada Student Loan Program or one of the provincial student loan programs. This item also includes amounts owing on loans taken directly from a financial institution to attend school.

Vehicle loans (PWDSTVHN): Amount owing on loans for those vehicles listed under assets.

3.4 Family type

Within the family type classification, the following definitions apply:

Couples: Couples include legally married, common-law and same-sex relationships.

Couples with children: Couples living with a child or children (by birth, adopted, step, or foster) under age 18. Children aged 18 or over are considered to be "other relatives". Other relatives may also be in the family.

Economic family: An economic family is defined as a group of two or more persons who live in the same dwelling and are related to each other by blood, marriage, common law or adoption.

Senior/senior economic families: Person aged 65 and over. In the case of senior families, the major income recipient is aged 65 or over at the time of the interview.

Family units: Includes economic families and or a person not in an economic family (unattached individual).

Lone-parent families: One parent, who is the major income recipient, living with at least one child under age 18. Other relatives may also be in the family. Families where the parent is 65 years and older are excluded.

Other non-senior families: All other economic families where the major income recipient was aged 64 or under at the time of the interview and not included in the couples, couples with children, or lone-parent family types.

Person not in an economic family - Unattached individual: An unattached individual is a person living either alone or with others to whom he or she is unrelated, such as roommates or a lodger. The correct standard term for unattached individual is now *'person not in an economic family'*.

Major income recipient or earner: For each family unit, the major income recipient is the person with the highest income before tax. For persons with negative total income before tax, the absolute value of their income is used, to reflect the fact that negative incomes generally arise from losses "earned" in the market place and are not meant to be sustained. In the rare situations where two persons have exactly the same income, the older person is the major income recipient.

4.0 Survey methodology

4.1 The survey universe

The 2019 Survey of Financial Security was carried out in all ten provinces, the territories were not included. Those living on reserves and other Aboriginal settlements and Crown lands and official representatives of foreign countries living in Canada and their families were also excluded from the survey. Members of religious and other communal colonies, members of the Canadian Forces living in military camps and people living in residences for senior citizens were excluded, as were people living full time in institutions, for example, inmates of penal institutions and chronic care patients living in hospitals and nursing homes. The survey covers about 98% of the population in the ten provinces.

Information was not gathered from persons temporarily living away from their families (for example, students at university) because it would be gathered from their families if selected. In this way, double counting of such individuals was avoided.

4.2 Survey content and reference period

With a few exceptions, the reference period for the information was the time of data collection (September to December 2019). For the asset and debt information respondents were asked to provide an estimate of the value or amount as close to the survey date as possible, recognizing that their most recent statement may have been as of the end of the previous calendar year, or for the last financial quarter.

Some of the information was collected for each person in the family 15 years of age and over. The assets and debts, however, were collected for the family as a whole, because they often cannot easily be assigned to one person in the family. Specifically, the following information was collected:

From each family unit member 15 years of age and over:

- demographics (age, sex at birth, gender, marital status);
- ethno-cultural characteristics;
- education;
- current employment;
- income, for the calendar year 2018.

From each family member 25 years of age and over:

- previous employer pension plans;
- pension plan benefits.

From each family member 45 years of age and over:

- retirement information.

For the family unit as a whole:

- financial and non-financial assets;
- equity in business;
- debt in the form of mortgages, vehicle loans, credit card and line of credit debt, student loans and other debt;
- distribution of registered plans investments;
- distribution of mutual funds investments.

4.3 The sample

As in previous iterations, SFS 2019 used more than one frame in its design. In rural areas, the SFS design continues to use the Labour Force Survey (LFS**)** area frame. However, the Household Survey Frame Service (HSFS), specifically the Dwelling Universe File (DUF), was used in urban areas, to allow the frame to be stratified by predicted net worth, which increases the sampling efficiency. The rural and area frames are mutually exclusive, which greatly simplifies the overall design of SFS. The overall initial sample size was 20,010 dwellings, and the sample was selected as two independent samples from the two distinct frames.

The rural sample was a stratified, multi-stage sample of 6,575 dwellings selected from the LFS area frame. This frame provides good coverage of the entire target population. Dwellings selected for SFS had not previously participated in labour force or financial surveys conducted by Statistics Canada which also select their samples from this frame. The rural frame is stratified into provinces and the sample selection comprised two steps in each province: the selection of clusters (small geographic areas) from the LFS area frame followed by the selection of dwellings within the selected clusters. At the time that the SFS sample was selected the LFS area frame used 2016 Census geography. The drawback of this frame's clustered design is that it decreases the sampling efficiency.

To improve the efficiency of the overall SFS sample, a second sample of 13,435 dwellings was selected from the DUF. Using information from the Socioeconomic indicators File (SEF) and the 2017 and prior years of the T1FF T1 Family File (T1FF), the frame was stratified into groups of dwellings predicted to have similar net worth. A simple random sample is then independently selected within each province and net worth strata. Stratifying by the key survey variable improves the sampling efficiency and ensures that all levels of family net worth are well represented in the sample.

4.4 Data collection

The 2019 Survey of Financial Security was conducted from September 2019 to December 2019. Data were collected during a personal interview using a CAPI application (computer assisted personal interview). Separate interviews were conducted for each economic family within the sampled dwellings.

For families, the interview was usually held with the family member with most knowledge of the family's financial situation. If necessary, follow-up was done with other family members. Proxy responses were accepted. This allowed one family member to answer questions on behalf of any or all other members of the family, provided he or she was willing and able to do so.

4.5 Data processing and quality control

The 2019 Survey of Financial Security was processed using the Social Survey Processing Environment (SSPE) processing system. The SSPE consists of a set of generalized processes to be used in the processing activities of the Survey Life Cycle. The purpose of these processes is to allow subject matter and survey support staff to specify and run the processing of a survey in a timely fashion with high quality outputs.

The various SSPE processing steps and utilities were developed using SAS, and utilize other software including PFM (Process Flow Manager), Excel and SQL Server. The processing steps and utilities were programmed as SAS macros. All parameters and specifications are provided through Excel workbooks, and the metadata for all surveys are stored on the metadata repository. This involved the following standard processing steps, Receipt of Raw Data, Clean Up, Recode, Flow Edits, Coding Apply, Consistency Edits, Derived Variables, Final Processing File and Dissemination Files, as well as steps created for external files created by Subject Matter and Methodology with regards to Pension Processing variables as well as Income Processing variables.

4.6 Imputation of missing data

Missing responses were imputed for all key fields in the questionnaire. All dollar amounts relating to assets and debts (and therefore, net worth) have been imputed. The nearest-neighbour imputation method was used primarily. In some cases deductive imputation was also used. There are flags in the SFS database that allow the user to

distinguish imputed values from the reported values. Imputation rates for fields that were imputed, are available for all imputed variables in the SFS database.

Not all variables on the SFS database were imputed for non-response. Variables that were not imputed may, therefore, contain missing values in the form of "Don't Know", "Refusal" or "Not Stated". None of the variables in the Language, Immigration and Aboriginal sections have been imputed.

Table 4-1 provides the percentage of total assets or total debts that each asset or debt item comprises and the percentage of the total value of each asset or debt item that was determined through imputation. For example, it shows that principal residences constituted 32% of total assets and that 7% of the total amount for principal residences was imputed. Note that Employer Pension Plan values were not imputed – all EPP values were estimated as described in section 3.2. EPP values are treated however as 100% imputed for the purposes of this table. Also note that the table was produced using the internal SFS database variables prior to perturbation for the PUMF.

Table 4-1 Percentage of Asset and Debt Item Totals due to Imputation

	Percentage of Total Assets or Total Debts (after imputation)	Percentage of Value of Item Determined through Imputation
Assets (PWATOTPT)	**100**	**30**
Pension assets	**30**	**69**
RRSPs/LIRAs (PWARRSPL)	8	17
RRIFs & Other retirement funds (PWARRIF & PWAOTPEN)	3	22
EPPs (PWARPPT)	19	100 [(1)]
Financial assets, non-pension	**14**	**20**
Deposits in financial institutions (PWASTDEP)	3	22
TFSAs (PWATFS)	2	17
Mutual funds and investment funds (PWASTMUI)	4	24
Stocks & Bonds (PWASTSTK & PWASTBND)	4	20
Other investments or financial assets (PWASTOIN)	1	11
Non-financial assets	**46**	**8**
Real estate	**41**	**8**
Principal residence (PWAPRVAL)	32	7
Other real estate (PWASTRST)	9	10
Vehicles and other non-financial assets	**5**	**6**
Vehicles (PWASTVHE)	2	6
Other non-financial assets (WASTONF)	3	6
Equity in business (PWBUSEQ)	**10**	**31**
Debts (PWDTOTAL)	**100**	**11**
Mortgage debt	**80**	**12**
Mortgage on principal residence (PWDPRMOR)	62	13
Mortgage on other real estate (PWDSTOMR)	18	7
Non-mortgage debt	**20**	**9**
Lines of credit (PWDSTLOC)	9	7
Credit card and instalment debt (PWDSTCRD)	2	7
Student loans (PWDSLOAN)	2	9
Vehicle loans (PWDSTVHN)	6	14
Other debt (PWDSTODB)	2	5

[1] All Employer Pension Plan values were estimated.

4.7 Weighting

The estimation of population characteristics from a survey is based on the premise that each sampled unit represents, in addition to itself, a certain number of non-sampled units in the population. A survey design weight is determined for each sample unit based on its probability of selection to indicate the number of units in the population that the unit represents. Design weights for the SFS are determined separately for the samples selected from the two frames. Three types of adjustments are then applied to the survey design weights: non-response adjustment, calibration, and adjustments for influential values.

The survey design weights are first adjusted to compensate for non-response. Adjustments are made separately to the weights within groups of dwellings with similar characteristics so that the respondents within the group also represent the non-respondents. The adjustment groups are formed using design information and variables which explain the non-response pattern. After non-response adjustment, the weights for the samples from both frames are combined into a single set of weights.

Next, the weights undergo an initial calibration to ensure that certain weighted estimates respect population totals from reliable sources other than the survey. The population totals used for the SFS comprise the following: demographic and household size projections, the number of wage and salary earners by several wage categories, the number of persons in high total income classes and the number of people with registered pension plans. All calibration totals are at the provincial level, except for demographic projections which also include totals for CMAs. The demographic and household size projection are produced by Statistics Canada's Demography division based on the 2016 Census. The wage and income classes are based on the APIM 2018 Prelim administrative file and Pension Plans in Canada (PPIC) 2019 is the source of the pension totals.

Influential observations are then identified, and weights are reduced for a small number of extreme observations.

To complete the process, calibration is then repeated a second time with the condition that the influential observations do not have their weights adjusted any further. The resulting weighted units represent the entire SFS target population.

5.0 Data accuracy and quality

5.1 Sampling errors

Sampling errors arise from estimating a population characteristic by looking at only a portion of the population rather than the entire population. It refers to the difference between the estimate derived from a sample survey and the 'true' value that would result if a census of the whole population were taken under the same conditions. There are no sampling errors in a census because the calculations are based on the entire population. The sample design, the variability of the population characteristics measured by the survey, and the sample size affect the magnitude of the sampling error.

The quality of estimates produced with SFS data is measured using 95% confidence intervals. Confidence intervals describe sampling variability and give an indication of the precision of a given estimate. When comparing estimates, it is important to use confidence intervals to determine if differences between values are statistically significant. The calculation and interpretation of confidence intervals is covered in the following section.

5.2 Measures of Sampling Error

There are several related measures of sampling error, each of which can be used to calculate the others using simple mathematical operations. The purpose of these measures is to estimate the degree of variation introduced in the estimates by selecting one particular sample rather than another sample of the same size and design.

One common measure of sampling error is the standard error (SE). The standard error of an estimate Y is defined as the square root of the estimated variance of the estimate Y. Standard error is easier to interpret than estimated variance because it uses the same scale as the estimate of Y. However, the magnitude of an estimate's standard error must be considered relative to the size of the estimate itself: if the standard error is large relative to the estimate, then the estimate has poor precision and is unreliable. For example, a standard error of 100 would be considered large for measuring the average weight of people but would not be considered large for estimating their average annual income.

It is more useful in many situations to assess the size of the standard error relative to the estimate of the characteristic being measured. One such relative measure of sampling error is the coefficient of variation (CV), which is computed as the estimated standard error of an estimate Y as a percentage of the estimate Y (i.e. $100\% \times SE / Y$). The CV is usually expressed as a percentage (10% instead of 0.1). It is useful in comparing the precision of sample estimates, where their sizes or scales differ from one another.

The CV is, however, a less meaningful measure of sampling error in the case of variables with values that can be zero or negative and estimates of proportions, particularly proportions that are close to zero or one. In the context of SFS, CV is a suitable measure of precision for estimates of assets or of debts, but it is less so for estimates of net worth, particularly for estimates that are close to zero. Furthermore, CV is also not as suitable as a measure of quality for estimates of proportions (e.g. proportions of households holding a given asset). In the cases where CV is not appropriate, it is recommended to use a confidence interval to measure sampling error.

One commonly used method to calculate the confidence intervals associated with an estimate Y uses the standard error. Under certain conditions, if the sampling were repeated many times, the true population value would lie within the Y +/- $1.96 \times SE$ confidence interval 95 times out of 100. However, the conditions necessary for this method of calculating confidence intervals are not always satisfied. Alternative methods for calculating confidence intervals also exist, including some based on bootstrap replicates which are mentioned below.

To illustrate the relationship between the standard error, the confidence intervals and the coefficient of variation, let us take the following example. Suppose that the estimated median net worth for a given domain is $10,000, and that its corresponding standard error is $200. The coefficient of variation is therefore equal to 2.0%. The 95% confidence interval estimated from this sample ranges from $9,608 to $10,392, i.e. $10,000 +/- $392. This means that with a 95% degree of confidence, it can be asserted that the median net worth of the target population in the domain is between $9,600 and $10,400.

Since the SFS uses a multi-stage survey design and calibration, there is no simple formula that can be used to calculate variance estimates. Therefore, an approximate method, known as the generalized bootstrap, was used for the PUMF. The generalized bootstrap method applies a random bootstrap adjustment to the sampling weight to create *R* bootstrap replicates. The variation among the *R* replicates approximates the sampling variance of the estimator. For a detailed explanation of the generalized bootstrap refer to Beaumont and Patak[2]

Users should note that R=1,000 bootstrap weights are provided for the SFS PUMF and should be used for variance estimation for all estimates. Users should also note that the variance estimates from the PUMF bootstraps will be smaller than the variance estimates released by Statistics Canada for the SFS. The greatest difference occurs for variances calculated with less than 100 observations. The difference in the variance estimates is due to the different methodology used to produce the bootstrap replicates. The generalized bootstraps which are included in the PUMF are used to protect confidentiality of survey respondents. The CV can be calculated to determine the quality of estimates for all asset and debt categories and for estimates of net worth that are not near zero. For near-zero net worth estimates and for estimates of proportions, confidence intervals should be calculated to determine quality. For quality release guidelines and information on statistical packages for variance estimation, please see Section 8.

Table 5-1 shows the precision of the SFS estimates of asset and debt totals. The table presents coefficients of variation for totals at the national level (i.e. the ten provinces combined) and at the provincial level for a number of asset and debt items as well as for total assets, total debts, and total net worth (termination basis). At the national level, the estimates are generally reliable. However, users should exercise caution when producing detailed estimates at the provincial level.

Note that the coefficients of variation presented in Table 5-1 are computed using the internal SFS database variables prior to perturbation for the PUMF.

[2] Beaumont, J.-F., and Z. Patak. (2012). On the generalized bootstrap for sample surveys with special attention to Poisson sampling. *International Statistical Review,* 80(1), 127-148.

Table 5-1 Coefficients of variation for totals, at the national level and at the provincial level [1] [2] [3]

	All family units										
	Newfoundland and Labrador	Prince Edward Island	Nova Scotia	New Brunswick	Quebec	Ontario	Manitoba	Saskatchewan	Alberta	British Columbia	Canada
Total Assets (WATOTPT)	**5.8**	**9.2**	**6.0**	**5.0**	**3.0**	**2.7**	**4.6**	**6.0**	**3.6**	**3.2**	**1.5**
Registered Retirement Savings Plans (RRSPs) and Locked-in Retirement Accounts (LIRAs) (WARRSPL)	13.5	24.5 [4]	12.4	12.6	5.6	4.8	8.5	9.3	6.0	6.2	**2.7**
Registered Retirement Income Funds (RRIFs),(WARRIF)	----- [6]	24.7 [4]	20.4 [4]	36.7 [5]	12.4	9.4	25.9 [4]	16.3	14.7	50.9 [5]	**11.0**
Retirement funds, other (WAOTPEN)	----- [6]	----- [6]	----- [6]	----- [6]	28.8 [4]	22.6 [4]	35.2 [5]	24.5 [4]	27.3 [4]	26.5 [4]	**14.4**
Employer-sponsored Registered Pension Plans (EPPs) (termination basis) (WARPPT)	7.9	9.4	7.2	8.1	3.9	3.4	6.4	6.9	4.8	4.8	**1.9**
Deposits in financial institutions (WASTDEPT)	15.4	16.4	14.5	31.2 [4]	6.3	6.2	9.1	10.3	7.8	7.2	**3.2**
Mutual funds and other (WASTMUIC)	----- [6]	----- [6]	35.0 [5]	----- [6]	21.1 [4]	18.1 [4]	20.9 [4]	22.3 [4]	20.1 [4]	14.2	**9.9**
Stocks (WASTSTCK)	----- [6]	----- [6]	39.6 [5]	----- [6]	15.6	18.2 [4]	48.3 [5]	36.9 [5]	24.5 [4]	18.0 [4]	**10.2**
Bonds (WASTBOND)	----- [6]	----- [6]	----- [6]	----- [6]	26.8 [4]	50.3 [5]	----- [6]	----- [6]	40.0 [5]	----- [6]	**32.8** [4]
Tax Free Saving Accounts (TFSA) (WATFS)	15.5	18.4 [4]	14.4	14.5	4.7	4.0	8.9	8.2	5.8	6.1	**2.3**
Other financial assets, non-pension (WASTOINP)	17.4 [4]	55.9 [5]	22.4 [4]	27.6 [4]	29.6 [4]	18.1 [4]	33.2 [4]	26.2 [4]	22.2 [4]	21.6 [4]	**11.1**
Principal residence (WAPRVAL)	4.6	8.0	7.4	4.9	2.6	2.0	4.0	3.9	2.4	3.0	**1.2**
Other real estate (WASTREST)	16.1	21.3 [4]	21.4 [4]	21.6 [4]	10.2	11.6	18.0 [4]	14.0	9.3	10.4	**6.1**
Vehicles (WASTVHLE)	5.3	7.6	8.5	5.7	3.0	2.6	6.3	4.8	3.4	3.9	**1.4**
Other non-financial assets (WASTONOF)	10.1	9.9	7.9	9.9	4.3	5.4	8.7	10.3	5.5	9.3	**2.9**
Equity in business (WBUSEQ)	38.1 [5]	37.5 [5]	55.0 [5]	44.9 [5]	16.7 [4]	13.8	29.7 [4]	21.0 [4]	17.9 [4]	21.4 [4]	**7.6**
Total debts (WDTOTAL)	**7.3**	**8.0**	**8.1**	**6.2**	**4.1**	**3.5**	**6.9**	**5.5**	**3.2**	**4.9**	**1.9**
Mortgage on principal residence (WDPRMOR)	8.5	10.6	10.5	7.6	4.2	3.8	7.4	6.9	3.5	5.7	**2.2**
Mortgage on other real estate (WDSTOMOR)	22.1 [4]	----- [6]	----- [6]	----- [6]	15.5	11.2	26.3 [4]	21.9 [4]	12.1	14.0	**6.3**
Line of credit (WDSTLOC)	15.3	27.9 [4]	17.8 [4]	15.8	14.9	8.0	24.3 [4]	16.1	10.8	15.4	**5.4**
Credit card and instalment debt (WDSTCRED)	11.5	14.6	13.5	13.0	7.7	7.2	14.5	10.0	9.2	9.9	**3.8**
Student loans (WDSLOAN)	----- [6]	23.2 [4]	22.7 [4]	23.1 [4]	12.2	9.0	25.6 [4]	20.0 [4]	11.9	14.6	**5.7**
Vehicle loans (WDSTVHLN)	8.7	10.4	11.2	7.8	5.3	4.6	11.9	8.6	6.1	7.8	**2.5**
Other debt (WDSTODBT)	33.5 [5]	28.7 [4]	35.5 [5]	41.7 [5]	36.3 [5]	35.3 [5]	31.7 [4]	23.3 [4]	20.5 [4]	24.1 [4]	**14.4**
Net Worth (assets less debts) (WNETWPT)	**6.6**	**10.2**	**6.6**	**5.9**	**3.3**	**3.0**	**5.0**	**6.9**	**4.3**	**3.6**	**1.6**

[1] The SurveyMeans procedure in SAS has been used to calculate these CVs. Some variables have also been grouped differently from what we can find in the CANSIM tables. Therefore some discrepancies might exist between these counts and the ones in the CANSIM tables.

[2] Estimates with CVs exceeding 16.6% should be used with caution and estimates with a sample size of less than 30 units are considered unreliable.

[3] For all types of assets and debts, the total, average and median values, as well as the number and percentage of family units holding, are calculated using values not equal to $0. For net worth, the total, average and median values, as well as the number of family units holding, are calculated using all values including net worth equal to $0.

[4] Estimate with CV between 16.6% and 33.3%.

[5] Estimate with CV exceeding 33.3%.

[6] Estimate is based on fewer than 30 observations.

5.3 Non-sampling errors

Non-sampling errors can be defined as errors arising during the course of all survey activities other than sampling. Unlike sampling errors, they can be present in both sample surveys and censuses.

Non-sampling errors can be classified into two groups: random errors and systematic errors.

- **Random errors** are the unpredictable errors resulting from estimation. They are generally cancelled out if a large enough sample is used. However, when these errors do take effect, they often lead to an increased variability in the characteristic of interest (i.e., the greater the difference between the population units, the larger the sample size required to achieve a specific level of reliability).

- **Systematic errors** are those errors that tend to accumulate over the entire sample. For example, if there is an error in the questionnaire design, this could cause problems with the respondent's answers, which in turn, can create processing errors, etc. These types of errors often lead to a bias in the final results.

Non-sampling errors are extremely difficult, if not impossible, to measure. Since random errors have the tendency to cancel out, systematic errors are the principal cause for concern. Unlike sampling variance, bias caused by systematic errors cannot be reduced by increasing the sample size.

Non-sampling errors can occur because of problems in **coverage**, **response**, **non-response**, **data processing**, **estimation and analysis**.

5.3.1 Coverage errors

Coverage errors are omissions, erroneous additions, duplicates and errors of classification of units in the survey frame. They can result from incomplete listing and inadequate coverage of the population. They have an impact on each survey estimate and are therefore one of the most important types of error. They can introduce bias, and the impact can vary for different sub-groups of the population. These errors are often systematic and result more often in undercoverage than in overcoverage of the target population. Calibration of survey weights is often used to correct for coverage errors.

5.3.2 Response errors

Response errors may be due to many factors, such as faulty questionnaire design, interviewers' or respondents' misinterpretation of questions, or respondents' faulty reporting. Great effort is invested in the SFS to reduce the occurrence of response error. Measures undertaken to minimize response errors include the use of highly-skilled and well-trained interviewers, and supervision of interviewers to detect misinterpretation of instructions or problems with the questionnaire design. Response error can also be brought about by respondents who, willingly or not, provide inaccurate responses.

Questions about the value of assets and the amount of debt can be particularly prone to misreporting, as respondents may misinterpret the questions or may not be able to provide an accurate answer. Furthermore, because proxy response was accepted, one family member may have provided inaccurate information for another family member, believing that information to be accurate. When providing information for the survey, respondents were encouraged to consult financial records or other family members as often as required.

5.3.3 Non-response errors

Non-response errors result from a failure to collect complete information on all units in the selected sample. There are two kinds of nonresponse: total nonresponse and item nonresponse.

Non-response produces errors in the survey estimates in two ways. First, non-respondents often have different characteristics from respondents, which can result in biased survey estimates. Secondly, non-response reduces the effective size of the sample since fewer units than expected responded to the survey. As a result, the sampling variance increases and the precision of the estimates decreases.

Total non-response can arise for a variety of reasons. For example, the interviewer is unable to contact the household, or no one in the household is able or willing to participate in the survey. Non-response adjustment of the survey weights for responding family units is performed in the SFS in order to reduce the non-response bias. For the 2019 SFS, the overall response rate was 59.4%.

Item non-response occurs when the information is available for certain questions only, in particular, because the respondent answers only to a portion of the questionnaire. It occurs for a number of reasons. A respondent may be unwilling or unable to answer the questions, the interviewer may fail to ask a question, or the wrong flow may have been followed through the questionnaire. Imputation of missing items compensates for this partial non-response.

Non-response error cannot be measured, but in general it is significant when non-respondents differ significantly from respondents with respect to particular characteristics that are important determinants of survey results.

5.3.4 Processing errors

Processing error is the error associated with activities conducted once survey responses have been received. It includes all data handling activities after collection and prior to estimation. Processing errors may occur in any of the data processing stages, for example, during data entry, coding, editing, imputation, weighting, and tabulation. Like all other errors, they can be random in nature, inflating the variance of the survey's estimates, or systematic, introducing bias. It is difficult to obtain direct measures of processing errors and their impact on data quality, especially since they are mixed in with other types of errors (non-response, measurement and coverage). The use of a generalized processing system reduces the processing errors that could occur. To minimize errors, diagnostic tests are carried out periodically to ensure that expected results have been obtained.

5.3.5 Estimation errors

Statistics Canada and other data-collecting agencies devote much effort to designing and monitoring surveys in order to make them as error-free as possible. If an inappropriate estimation method is used, then bias can still be introduced, regardless of how errorless the survey had been before estimation.

5.3.6 Analysis error

Analysis errors include any errors that occur when using the wrong analytical tools or when the preliminary results are used instead of the final ones. Errors that occur during the publication of these data results are also considered analysis errors.

5.4 Treatment of large values

For any sample, estimates can be affected disproportionately by the presence of extreme values from the population. In an asset and debt survey, a few extreme values are expected in the sample, as valid extreme values do exist in the population. Values outside defined bounds were identified and reviewed in relation to other information reported for that respondent. If the value was judged to be the result of a reporting or processing error, it was adjusted. Otherwise, it was retained and the weight of the family unit was reduced to ensure it has no undue influence on the survey estimates.

5.5 Impact of sampling and non-sampling errors on SFS estimates

Due to the combined effect of these errors, the quality of net worth data is judged to be lower than the quality of income data. This is largely because records of the current value of assets and the outstanding amount of debt are not as readily available as records of income. For example, respondents with numerous bank accounts and investments may receive several different statements, each with different reference period. Compiling this information can be difficult. Most income information, on the other hand, is obtained directly from income tax returns for the year in question.

5.6 Comparability of data and related sources

Of the variables that do have sources, comparison is often difficult because of differences in defining concepts, grouping of items, and how these items are valued.

Direct comparisons with outside sources, such as the National Balance Sheet Accounts (NBSA) of the System of National Accounts (SNA), do yield certain differences. Comparing both of these sources is difficult due to definitional, coverage and treatment differences.

Based on rough comparisons between the NBSA and the SFS, the following general conclusions can be drawn:

(a) The SFS appears to underestimate some net worth components, particularly financial assets and consumer debt.
(b) The quality of estimates of real assets (e.g., owner-occupied homes, vehicles) is much better than that of financial assets.

In theory – given similar valuation procedures and groupings – SNA data should be the same as that collected by an asset and debt survey. The SNA collects individual wealth data from institutional sources such as banks and insurance companies, net of corporations and governments. One major problem has been the SNA categorization of individuals and unincorporated business. Because the individual data and the unincorporated business cannot be separated out, these estimates will always be higher than the survey estimates alone.

The Census and other surveys are important sources for ensuring that the SFS sample is representative of the Canadian population. Despite conceptual differences with the SNA estimates, ensuring a representative sample is extremely important to the validity of the data. SFS estimates for pension variables such as membership and contributions were found to be very close to data produced by Statistics Canada's Pension Plans in Canada Survey.

5.7 Response rates

The overall response rate for the 2019 Survey of Financial Security was 59.4%. Table 5-2 gives a breakdown of response rates by province for the area frame sample and the DUF frame sample.

Table 5-2 Response rates for family units by frame type and province

Province	Urban sample (DUF frame)		Rural sample (LFS area frame)		Samples from both frames	
	Number of responding family units	Response rate (%)	Number of responding family units	Response rate (%)	Number of responding family units	Response rate (%)
Newfoundland & Labrador	164	71.6	216	78.0	380	75.1
Prince Edward Island	194	68.1	142	71.0	336	69.3
Nova Scotia	263	61.4	238	64.3	501	62.9
New Brunswick	256	63.5	149	73.8	405	67.1
Quebec	1,544	64.8	459	66.2	2,003	65.1
Ontario	2,025	53.6	673	60.6	2,698	55.3
Manitoba	277	48.2	253	63.6	530	54.5
Saskatchewan	271	58.3	379	61.6	650	60.2
Alberta	1,225	64.4	401	54.1	1,626	61.5
British Columbia	991	49.7	302	59.8	1,293	51.9
Total – All Provinces	**7,210**	**57.9**	**3,212**	**62.9**	**10,422**	**59.4**

6.0 Record layout, data dictionary and univariate distributions

Two additional information files are provided to assist users of the SFS PUMF microdata file: a record layout and a data dictionary with frequencies.

6.1 Columns of the record layout

Variable name: Public use microdata file (PUMF) variable name: This is the variable name assigned for the microdata file.

Type: Indicates whether the variable is numeric (in the sense that it can logically be used in mathematical operations) or character.

Number of categories: Shows the number of categories in the value set for the variable in question. The number applies only to "character" variables. Numeric variables have ranges, which are specified in the data dictionary.

Length: Indicates the number of spaces. For numeric variables, this includes the decimal point if there are decimal places and the number of decimal places, if any. For example, a variable which can have values of zero (00.0) to 99.9 would have a format expressed as 4.1. A variable which can have values of zero (00) to
99 would have a format express as 2.0.

Sequence number: Indicates the order that variables appear on the microdata file.

Start position: This shows the location of the variable on the public use microdata file.

Long variable name: A standardized name, with a maximum of 26 characters, which can be used to quickly identify variables, to label tables, and so on. Although still rather cryptic, it is considerably more revealing than the variable name. However, this longer name obviously excludes a lot of important information contained in the variable description shown in the data dictionary. In short, analysts are warned against making assumptions about the variable definition based on the long variable name.

6.2 Data dictionary with frequencies

The data dictionary presents the complete information about each survey variable on the file. For each variable are shown: the variable name, the description or definition, code lists with descriptions or alternatively the range of values that the variable can take on, the variable type, its length (or format), and the population to which the variable pertains, i.e. for whom it is applicable.

These distributions are provided to allow users of the public use microdata file to verify totals that they produce. These distributions relate to the public use files and not to the internal database; the distributions will be similar but not identical. To compare the public use file to the internal database, please see Appendices A, B, and C at the end of this user guide.

For character variables, the weighted and unweighted frequencies for each code, including reserved codes, are produced. For numeric variables, the range of values are included. The weighted and unweighted frequencies for both this range and the reserved codes are also provided.

6.3 Reserved codes

It is important to account for reserved codes in any analysis, particularly with numeric variables. If your calculation of means or aggregates seems too high, check to ensure that you have excluded reserved codes from the calculation. With only a few exceptions, the reserved codes are the highest four values permitted according to the length of

the variable. A brief explanation of reserved codes is provided below.

6, 96, 9…6, etc.: Valid skip - Not applicable

7, 97, 9…7, etc.: Don't know
(the respondent did not have an answer, or the value was rejected during processing without being replaced)

8, 98, 9…8, etc.: Refusal
(the respondent refused to answer the particular question in the questionnaire)

9, 99, 9...9, etc.: Not stated

7.0 Guidelines for analysis and presentation

7.1 Applying weights

The estimation of population characteristics from a survey is based on the premise that each sampled unit represents, in addition to itself, a certain number of non-sampled units in the population. The microdata on the public use file are unweighted. It is the responsibility of data users to apply the appropriate weights in any results they wish to produce. If proper weights are not used, the estimates derived from the microdata cannot be considered to be representative of the survey population, and will not correspond to those that would be produced by Statistics Canada. On the SFS PUMF, the weight variable is named PWEIGHT.

7.2 Rounding guidelines

Once it has been determined whether the results obtained are reliable, the level of rounding indicates the level of precision that the data can actually support. The following guidelines for rounding should be used:

- Estimates of population sub-groups should be rounded to the nearest hundred units.
- Rates and percentages should be rounded to one decimal point.

Note that all calculations are to be derived from their unrounded components, and then rounded using the normal rounding technique.

In normal rounding, if the first or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is raised by one. For example, in normal rounding to the nearest 100, the estimate 49,448 would be rounded down to 49,400 and an estimate of 49,252 would be rounded up to 49,300. The figure 1.78% would be rounded to 1.8%.

7.3 Missing values and reserved codes

There are a few types of missing values on the public use file. If the coverage of a variable does not extend to a certain population sub-group, then there are no valid values for that sub-group and the values that do appear are in the form of 6, 96, 9.6 and so on, depending on the format of the variable, which indicates that the variable is not applicable (marked at *'Valid Skip'* in the dictionary). The coverage of each variable on the file is referred to in the data dictionary as the "population".

For some variables for specific records, no valid value is available, although the variable is applicable. Possibly the respondent did not provide the information, or it failed an edit in processing and the value was not imputed, or it was suppressed by methodology to keep the data confidential. Such missing values appear with a reserved code such as 9, 99, 9.9, and so on, depending on the format of the variable. For certain variables, the number of missing values has been reduced through imputation. Missing values for the income variables have been entirely imputed, but other variables may have missing values.

The approach for dealing with missing values of this last kind depends on the type of analysis being carried out and the extent of missing data. Although the end solution may be to exclude the records with missing values from the analysis, a review should first be carried out to assess the impact of missing values on the overall representativeness of the data. Is it possible that a bias results from excluding records with missing values? For example, are the (other) characteristics of the people with missing values different from those of the observed part of the sample? It may be necessary to take into account the possible impact in some way. In all cases, analysts should note exclusions of records with missing values in their published results.

8.0 Guidelines for release

Microdata users should apply the rules for assessing data quality, below, to all estimates they produce, and retain only those that satisfy the release criteria. Estimates that do not satisfy the release criteria are not reliable.

8.1 Introduction

The guidelines for release and publication make use of the concept of sampling variability to determine whether estimates obtained from the microdata are reliable. Sampling variability is the error in the estimates caused by the fact that we survey a sample rather than the entire population. The concept of standard error and the related concepts of coefficient of variation and confidence interval provide an indication of the magnitude of the sampling variability.

The standard error and coefficient of variation do not measure any systematic biases in the survey data that might affect the estimate. Rather, they are based on the assumption that the sampling errors follow a normal probability distribution.

Subject to this assumption, it is possible to estimate the extent to which different samples that have the same design and the same number of observations would give different results. This indicates the margin of error that is likely to be included in the estimates derived from our single sample.

For a detailed description of the measures of sampling variability, see A. Satin and W. Shastry, Survey Sampling: A Non-Mathematical Guide, Statistics Canada, Catalogue no. 12-602-XPE.

8.2 Minimum sizes of estimates for release (Guideline for release)

In general, the smaller the sample, the greater the sampling variability. Therefore, estimates of small population subgroups are less reliable than estimates of large population subgroups. The minimum allowable sizes of estimates, also called the release cut-offs, are a quick rule for determining whether an estimate can be released, before applying the more rigorous test that uses the coefficient of variation. The release cut-offs are calculated specifically for the Survey of Financial Security, based on the sample size and the sample design.

The cut-off for the unweighted count must be satisfied:

- Unweighted count: The number of observations must be at least 30. If the unweighted count is less than 30, then the weighted estimate should not be released regardless of the value of its coefficient of variation.

For many estimates, coefficient of variation can be used to determine the quality of an estimate. However, for ratios and for variables such as net worth that can take on negative values, CV is a less useful measure of quality, and confidence intervals are more appropriate. For such estimates, whether or not to publish an estimate should be decided based on how its confidence interval reflects its fitness for use. Users should be sure to read Chapter 5 to be more fully aware of the quality characteristics of these data.

For estimates of assets and debts based on sample sizes of 30 or more, users can use the coefficient of variation of the estimate to determine the quality of the estimate. Estimates with CVs exceeding 16.6% should be used with caution. In these cases, estimates with a coefficient of variation between 16.6% and 33.3% can be considered to be of marginal quality and should be accompanied by a warning to use the estimates with caution. Estimates with a coefficient of variation over 33.3% can be considered to be of unacceptable quality. These quality level guidelines should be applied to rounded weighted estimates. Users are also reminded that since the variance estimates tend to be smaller for the generalized bootstraps available with the PUMF this means that the CVs will also be smaller (see section 5.2). This difference is greatest for estimates from smaller sample sizes, less than 100 observations. Therefore users should be cautious when reporting CVs that are close to the cut-off values for marginal (16.6%) or unacceptable (33.3%) quality. Appendix B shows a comparison of the CVs produced from the production files with the bootstrap weights and the PUMF file with the PUMF bootstraps. The difference in the confidence intervals and the CV is the most pronounced for Bonds, which is calculated from the fewest number of observations. The CV

from the PUMF indicates an acceptable quality, although the quality indicator for this variable should be marginal.

All estimates, other than estimates that cannot be released under the Disclosure Control Rules noted above, can be released. However, those of marginal or unacceptable quality level must be accompanied by a warning to caution subsequent users.

It is also recommended that 95% confidence intervals should be included for every released and/or published estimate. The following note may accompany estimates given with confidence intervals: "The 95% confidence interval illustrates the degree of variability associated with an estimate. Wide confidence intervals indicate high variability, thus, these estimates should be interpreted with due caution."

8.3 Statistical Packages for Variance Estimation

In order to determine the quality of the estimate a measure of its variability must be calculated. This can be the variance, the standard error, the coefficient of variation or a confidence interval. The SFS uses a multi-stage survey design and calibration, which means that there is no simple formula that can be used to calculate variance estimates. Therefore, an approximate method is needed - the generalized bootstrap method (see Section 5.2).

There are several different commercially available software packages that can be used to estimate variance using bootstrap weights. Examples include SAS, Stata, SUDAAN, WesVar, and BOOTVAR[3]. While the bootstrap method is often not directly supported, Balanced Repeated Replication (BRR) is supported, and its functionality can be used to estimate the variance for surveys using the bootstrap method. The generalized bootstrap and BRR differ in the way in which the replicates are built, but the variance estimation process is the same for both methods, a fact that is seldom mentioned in software documentation (Phillips, 2004). There are also methods implemented in some of the above software packages for using bootstrap weights to calculate confidence intervals for estimates.

A detailed description (along with examples) of how to implement the bootstrap method with the selected software is found in Statistics Canada's Research Data Centre's Information and Technical Bulletin – "Weighted estimation and bootstrap variance estimation for analyzing survey data: How to implement in selected software" (Gagné, 2014). http://www.statcan.gc.ca/pub/12-002-x/2014001/article/11901-eng.htm

8.4 Hypothesis tests provided by standard statistical software packages

Microdata users should be aware that the results of hypothesis tests (such as the p values accompanying t statistics or Pearson statistics) that are provided automatically by standard statistical software packages are incorrect for data provided by surveys with a complex survey design, such as the Survey of Financial Security. Such packages calculate these test results under the assumption of simple random sampling. That is, they do not take into account the special sample design features of SFS such as stratification, clustering, and unequal selection probabilities. While many of the standard packages can account for the unequal selection probabilities in the production of estimates by allowing the use of weights, these packages do not properly take the sample design into account when producing variance estimates that form part of most test statistics.

The statistical packages described in section 8.3 can be used to make hypothesis tests and to run regression and other statistical analyses using survey specific procedures that allow users to include the bootstrap weights for an appropriate calculation of the variance. Since the variance estimates calculated from the PUMF bootstrap weights tend to be smaller than they should be, hypothesis tests comparing groups may indicate significant differences that do not exist in the published data. It is therefore recommend that users increase the level of significance at which groups are considered to be different. For example, rather than using standard value of 0.05, it is recommend to use 0.01. This will help to ensure the validity of conclusions made from hypothesis tests. This is equivalent to changing from a 95% to a 99% confidence interval.

[3] BOOTVAR is a bootstrap variance estimation application developed by Statistics Canada. It was originally developed when alternatives were not available. Although it may still be used by some users, it is currently not maintained by Statistics Canada.

9.0 Appendix A – SFS production totals VS PUMF totals

The table below provides a comparison between Statistics Canada's internal SFS database and the public use microdata file.

Table 9-1 Comparison of SFS production totals to SFS PUMF, Canada, 2019

	Production Estimate				PUMF Estimate				% Difference 100*((PUMF/Production) - 1)			
	Total family units	Sum	Mean	Median	Total family units	Sum	Mean	Median	Total family units	Sum	Mean	Median
	number	$ millions	$	$	Number	$ millions	$	$	%	%	%	%
Assets (PWATOTPT)	**15,914,000**	**13,615,026**	**855,500**	**477,000**	**15,914,000**	**13,713,427**	**861,700**	**485,900**	**0.00**	**0.72**	**0.72**	**1.87**
RRSPs/LIRAs (PWARRSPL)	7,860,000	1,055,993	134,400	44,900	7,916,000	1,060,167	133,900	45,600	0.71	0.40	-0.37	1.56
RRIF (PWARRIF)	1,659,000	315,020	189,900	80,000	1,684,000	319,941	190,000	81,300	1.51	1.56	0.05	1.63
Retirement funds, other (PWAOTPEN)	1,056,000	61,176	57,900	11,800	1,059,000	62,195	58,700	11,900	0.28	1.67	1.38	0.85
Employer Pension Plans (PWARPPT)	8,017,000	2,513,121	313,500	164,900	8,074,000	2,508,611	310,700	160,000	0.71	-0.18	-0.89	-2.97
Deposits in financial institutions (PWASTDEP)	15,141,000	384,039	25,400	5,000	15,175,000	388,645	25,600	5,000	0.22	1.20	0.79	0.00
Mutual funds and other (PWASTMUI)	1,707,000	424,558	248,800	48,800	1,714,000	433,865	253,200	49,400	0.41	2.19	1.77	1.23
Stocks (PWASTSTK)	1,465,000	307,144	209,700	25,000	1,490,000	312,348	209,600	24,800	1.71	1.69	-0.05	-0.80
Bonds (PWASTBND)	363,000	27,891	76,900	4,600	367,000	28,516	77,600	5,100	1.10	2.24	0.91	10.87
Tax free Savings Accounts (PWATFS)	7,282,000	248,859	34,200	14,900	7,325,000	250,755	34,200	15,100	0.59	0.76	0.00	1.34
Other financial assets, non-pension (PWASTOIN)	3,265,000	152,225	46,600	12,000	3,249,000	151,652	46,700	12,100	-0.49	-0.38	0.21	0.83
Principal residence (PWAPRVAL)	9,850,000	4,992,343	506,800	399,000	9,931,000	5,023,104	505,800	384,100	0.82	0.62	-0.20	-3.73
Other real estate (PWASTRST)	2,789,000	1,272,081	456,100	199,700	2,787,000	1,278,586	458,700	201,600	-0.07	0.51	0.57	0.95
Vehicles (PWASTVHE)	12,801,000	337,048	26,300	17,000	12,856,000	337,261	26,200	16,700	0.43	0.06	-0.38	-1.76
Other non-financial assets (PWASTONF)	15,914,000	400,053	25,100	9,300	15,914,000	404,990	25,400	9,600	0.00	1.23	1.20	3.23
Equity in business (PWBUSEQ)	2,779,000	1,123,477	404,300	9,800	2,815,000	1,152,792	409,500	9,300	1.30	2.61	1.29	-5.10
Debts (PWDTOTAL)	**11,113,000**	**1,866,265**	**167,900**	**78,800**	**11,118,000**	**1,878,289**	**168,900**	**80,300**	**0.04**	**0.64**	**0.60**	**1.90**
Mortgage on principal residence (PWDPRMOR)	5,500,000	1,213,438	220,600	179,900	5,537,000	1,222,834	220,900	178,800	0.67	0.77	0.14	-0.61
Mortgage on other real estate (PWDSTOMR)	1,055,000	293,983	278,700	177,700	1,060,000	296,432	279,500	178,100	0.47	0.83	0.29	0.23
Line of credit (PWDSTLOC)	3,068,000	140,662	45,800	15,900	3,077,000	139,990	45,500	16,100	0.29	-0.48	-0.66	1.26
Credit card and instalment debt (PWDSTCRD)	5,930,000	40,190	6,800	3,000	5,940,000	40,360	6,800	3,100	0.17	0.42	0.00	3.33
Student loans (PWDSLOAN)	1,990,000	38,826	19,500	11,800	1,997,000	39,036	19,500	11,500	0.35	0.54	0.00	-2.54
Vehicle loans (PWDSTVHN)	4,800,000	112,170	23,400	18,000	4,786,000	111,730	23,300	18,000	-0.29	-0.39	-0.43	0.00
Other debt (PWDSTODB)	1,362,000	26,996	19,800	4,900	1,378,000	27,908	20,200	5,000	1.17	3.38	2.02	2.04
Net worth (PWNETWPT)	**15,900,000**	**11,748,761**	**738,900**	**330,500**	**15,913,000**	**11,835,138**	**743,800**	**338,400**	**0.08**	**0.74**	**0.66**	**2.39**

Notes:

Total family units estimates have been rounded to the nearest 1,000.

Mean and median estimates have been rounded to the nearest 100.

10.0 Appendix B – SFS production Quality Indicators VS PUMF Quality Indicators

The table below provides a comparison between Statistics Canada's internal SFS database and the public use microdata file for the 95% confidence intervals and CVs.

Table 10-1 Comparison of SFS production quality indicators to SFS PUMF quality indicators, Canada, 2019
Confidence intervals for total assets and debts held by family units, at the national level (2019 constant dollars x 1,000,000)

	Production Estimate and Quality Indicators				PUMF Estimate and Quality Indicators			
	Sum	95% Confidence intervals		Coefficients of variation (CVs)	Sum	95% Confidence intervals		Coefficients of variation (CVs)
		Lower bound	Upper bound			Lower bound	Upper bound	
	$ millions	$ millions	$ millions	%	$ millions	$ millions	$ millions	%
Assets (PWATOTPT)	**13,615,026**	**13,216,053**	**14,013,999**	**1.5**	**13,713,427**	**13,346,660**	**14,080,194**	**1.4**
RRSPs/LIRAs (PWARRSPL)	1,055,993	999,599	1,112,387	2.7	1,060,167	1,011,349	1,108,985	2.3
RRIF (PWARRIF)	315,020	247,021	383,019	11.0	319,941	275,927	363,956	7.0
Retirement funds, other (PWAOTPEN)	61,176	43,905	78,446	14.4	62,195	47,192	77,197	12.3
Employer Pension Plans (PWARPPT)	2,513,121	2,418,816	2,607,426	1.9	2,508,611	2,420,475	2,596,748	1.8
Deposits in financial institutions (PWASTDEP)	384,039	360,191	407,887	3.2	388,645	363,381	413,909	3.3
Mutual funds and other (PWASTMUI)	424,558	342,300	506,816	9.9	433,865	362,916	504,815	8.3
Stocks (PWASTSTK)	307,144	245,527	368,760	10.2	312,348	258,914	365,781	8.7
Bonds (PWASTBND)	27,891	9,962	45,820	32.8	28,516	17,588	39,443	19.5
Tax free Savings Accounts (PWATFS)	248,859	237,788	259,929	2.3	250,755	239,643	261,867	2.3
Other financial assets, non-pension (PWASTOIN)	152,225	119,036	185,413	11.1	151,652	120,168	183,135	10.6
Principal residence (PWAPRVAL)	4,992,343	4,872,438	5,112,248	1.2	5,023,104	4,898,241	5,147,968	1.3
Other real estate (PWASTRST)	1,272,081	1,119,185	1,424,977	6.1	1,278,586	1,137,679	1,419,492	5.6
Vehicles (PWASTVHE)	337,048	327,632	346,465	1.4	337,261	328,933	345,588	1.3
Other non-financial assets (PWASTONF)	400,053	377,259	422,846	2.9	404,990	382,992	426,988	2.8
Equity in business (PWBUSEQ)	1,123,477	955,386	1,291,568	7.6	1,152,792	999,101	1,306,482	6.8
Debts (PWDTOTAL)	**1,866,265**	**1,795,385**	**1,937,145**	**1.9**	**1,878,289**	**1,810,566**	**1,946,012**	**1.8**
Mortgage on principal residence (PWDPRMOR)	1,213,438	1,162,173	1,264,704	2.2	1,222,834	1,175,338	1,270,330	2.0
Mortgage on other real estate (PWDSTOMR)	293,983	257,454	330,512	6.3	296,432	261,269	331,595	6.0
Line of credit (PWDSTLOC)	140,662	125,777	155,548	5.4	139,990	127,007	152,972	4.7
Credit card and instalment debt (PWDSTCRD)	40,190	37,225	43,155	3.8	40,360	37,681	43,039	3.4
Student loans (PWDSLOAN)	38,826	34,516	43,136	5.7	39,036	35,038	43,034	5.2
Vehicle loans (PWDSTVHN)	112,170	106,721	117,619	2.5	111,730	106,618	116,842	2.3
Other debt (PWDSTODB)	26,996	19,391	34,600	14.4	27,908	21,105	34,710	12.4
Net worth (PWNETWPT)	**11,748,761**	**11,371,749**	**12,125,773**	**1.6**	**11,835,138**	**11,483,811**	**12,186,465**	**1.5**

Notes:

The SurveyMeans procedure in SAS has been used to calculate these 95% Confidence Intervals and CVs.

Therefore some discrepancies might exist between these counts and the ones in the CANSIM tables.

11.0 Appendix C – SFS totals

The following tables were generated from the SFS production database. Users may use these figures to compare their estimates from the microdata file with Statistics Canada's totals.

Table 11-1 Assets, debts, net worth showing millions of dollars and number for family units, Canada and regions, 2019

	Canada	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.
	Sum (2019 constant dollars x 1,000,000)										
Assets (WATOTPT)	13,615,026	129,314	39,859	268,060	155,870	2,397,970	5,882,379	396,729	419,534	1,576,465	2,348,845
Debts (WDTOTAL)	1,866,265	19,803	4,676	30,706	23,087	305,697	776,263	51,650	52,340	268,696	333,347
Net worth (WNETWPT)	11,748,761	109,511	35,183	237,355	132,783	2,092,273	5,106,115	345,079	367,194	1,307,769	2,015,499

	Canada	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.
	Number of family units										
Assets (WATOTPT)	15,914,000	235,000	70,000	443,000	350,000	3,849,000	6,002,000	529,000	466,000	1,725,000	2,245,000
Debts (WDTOTAL)	11,113,000	168,000	50,000	331,000	261,000	2,695,000	4,095,000	368,000	339,000	1,293,000	1,515,000
Net worth (WNETWPT)	15,914,000	235,000	70,000	443,000	350,000	3,849,000	6,002,000	529,000	466,000	1,725,000	2,245,000

Total family units estimates have been rounded to the nearest 1,000.

Table 11-2 Assets, debts, net worth showing median and average amounts for family units, Canada and regions, 2019

	Canada	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.
	Median amount for family units holding asset and debt										
Assets (WATOTPT)	477,000	326,300	291,000	335,800	265,000	344,600	608,900	445,000	492,500	520,200	617,000
Debts (WDTOTAL)	79,000	46,200	49,000	39,000	52,000	40,500	95,200	82,000	107,000	151,900	110,000
Net worth (WNETWPT)	329,900	247,300	211,400	257,900	185,000	237,800	434,500	295,700	330,500	317,300	423,700

	Canada	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.
	Average amount for family units holding asset and debt										
Assets (WATOTPT)	855,500	551,100	572,100	604,600	444,700	623,000	980,100	749,600	901,000	913,900	1,046,100
Debts (WDTOTAL)	167,900	118,200	93,800	92,800	88,400	113,400	189,600	140,400	154,600	207,800	220,100
Net worth (WNETWPT)	738,200	466,700	505,000	535,300	378,800	543,500	850,800	652,000	788,600	758,200	897,600

Median and average estimates have been rounded to the nearest 100.