



Catalogue no. 71M0001GIE

Labour Force Survey Microdata User Guide

2005



Statistics
Canada

Statistique
Canada

Canada

How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: Client Services, Labour Statistics Division, Statistics Canada, Ottawa, Ontario, K1A OT6, (telephone: (613) 951-4090, toll free number: 1 (866) 873-8788, fax: (613) 951-2869 or by e-mail address: labour@statcan.ca).

For information on the wide range of data available from Statistics Canada, you can contact us by calling one of our toll-free numbers. You can also contact us by e-mail or by visiting our Web site.

National inquiries line	1 800 263-1136
National telecommunications device for the hearing impaired	1 800 363-7629
Depository Services Program inquiries	1 800 700-1033
Fax line for Depository Services Program	1 800 889-9734
E-mail inquiries	infostats@statcan.ca
Web site	www.statcan.ca

Ordering and subscription information

This product, Catalogue no. 71M0001GIE, is available on Internet free. Users can obtain single issues at www.statcan.ca.

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner and in the official language of their choice. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll free at 1 800 263-1136.



Statistics Canada

Labour Force Survey Microdata User Guide

2005

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2005

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission from Licence Services, Marketing Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6.

May 2005

Catalogue no. 71M0001GIE

Frequency: Annual

Ottawa

La version française de cette publication est disponible sur demande (n° 71M0001GIF au catalogue).

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

ELECTRONIC PUBLICATIONS AVAILABLE AT
www.statcan.ca



Table of contents

1.0	Introduction	6
2.0	Background and objectives of the Labour Force Survey	7
2.1	Determining labour force status	8
3.0	Labour Force Survey concepts and definitions	10
4.0	Survey methodology	21
4.1	Population coverage	21
4.2	Sample design	21
4.2.1	Primary sampling unit (PSU) delineation	21
4.2.2	Primary stratification	21
4.2.3	Types of areas	22
4.2.4	Secondary stratification	22
4.2.5	Cluster dwelling and selection	23
4.2.6	Person selection	23
4.3	Sample size	24
5.0	Data collection	25
5.1	Interviewing for the Labour Force Survey	25
5.2	Supervision and quality control	25
5.3	Non-response to the Labour Force Survey	25
6.0	Data processing	27
6.1	Data capture	27
6.2	Editing	27
6.3	Imputation	27
6.4	Creation of derived variables	28
6.5	Weighting	28
6.6	Suppression of confidential information	28
7.0	Data quality	29
7.1	Response rates	29
7.2	Survey errors	29
8.0	Guidelines for tabulation, analysis and release	31
8.1	Sample weighting guidelines for tabulation	32
8.2	Weighting procedures for the Labour Force Survey	32
8.3	Rounding guidelines	32
8.4	Guidelines for statistical analysis	33
8.5	Approximate sampling variability	34
8.6	Release criteria for the Labour Force Survey	39
8.7	Composite estimation	39
9.0	The Labour Force Survey questionnaire	41
9.1	Guidelines for using the LFS questionnaire	40
10.0	Record layout	41
10.1	Public use	41
10.2	Confidential use	41

1.0 Introduction

This manual has been produced to facilitate the manipulation of the microdata file of the Labour Force Survey (LFS) results. The public use microdata file contains non-aggregated data for a wide variety of variables collected from the LFS.

This file contains both personal characteristics for all individuals in the household and detailed labour force characteristics for household members 15 years of age and over. See Section 3 of this publication for a list of the variables in this product. These variables are available by province and for the three largest census metropolitan areas (Montreal, Toronto, and Vancouver). This is a monthly file, and is available going back to 1976. This product is for users who prefer to do their own analysis by focusing on specific subgroups in the population or by cross-classifying variables that are not in our catalogued products.

Any questions about the data set or its use should be directed to:

Statistics Canada

Labour Statistics Division

Client Services

Telephone: (613) 951-4090 or call toll-free 1 866 873-8788

Fax: (613) 951-2869

E-mail: labour@statcan.ca

2.0 Background and objectives of the Labour Force Survey

The Canadian Labour Force Survey was developed following the Second World War to satisfy a need for reliable and timely data on the labour market. Information was urgently required on the massive labour market changes involved in the transition from a war-time to a peace-time economy. The survey was designed to provide estimates of employment by industry and occupation at the regional as well as the national level.

A quarterly survey initially, the LFS became a monthly survey in 1952. In 1960, the Interdepartmental Committee on Unemployment Statistics recommended that the LFS be designated the source of the official measure of unemployment in Canada. This endorsement was followed by a demand for a broader range of labour market statistics, in particular more detailed regional data. The information generated by the survey has expanded considerably over the years with a major redesign of the survey content in 1976 and again in 1997, and provides a rich and detailed picture of the Canadian labour market.

The LFS is the only source of monthly estimates of total employment including the self-employed, full and part-time employment, and unemployment. It publishes monthly standard labour market indicators such as the unemployment rate, the employment rate and the participation rate. The LFS is a major source of information on the personal characteristics of the working-age population, including age, sex, educational attainment, and family characteristics.

Employment estimates include detailed breakdowns by demographic characteristics, industry and occupation, job tenure, and usual and actual hours worked. The survey incorporates questions permitting analyses of many topical issues, such as involuntary part-time employment, multiple job-holding, and absence from work. Since January 1997, it also provides monthly information on the wages and union status of employees, as well as the number of employees at their workplace and the temporary or permanent nature of their job.

Unemployment estimates are produced by demographic group, duration of unemployment, and activity before looking for work. Information on industry and occupation, and reason for leaving last job is also available for persons currently unemployed or not in the labour market with recent labour market involvement.

In addition to providing national and provincial estimates, the LFS also releases estimates of labour force status for sub-provincial areas such as Economic Regions and Census Metropolitan Areas.

2.1 Determining labour force status

The concepts of employment and unemployment are derived from the theory of the supply of labour as a factor of production. The production referred to is in turn defined as those goods and services included in the System of National Accounts. For this reason, unpaid housework and volunteer work are not counted as work by the survey, although these activities need not differ from paid work, either in purpose or in the nature of the tasks completed.

While the logical and precise unit of measurement of total labour supply is person-hours, the conceptual terms of reference for the survey require that individual members of the population be classified as employed, unemployed, or not in the labour force. Accordingly, persons who are

supplying services in the reference period, regardless of the quantity supplied, are classified as employed while those who provide evidence that they are offering their labour services to the market (again regardless of quantity) are classified as unemployed. The remainder of the population, those neither currently supplying nor offering their labour services, are referred to as persons not in the labour force.

The concepts and definitions of employment and unemployment adopted by the survey are based on those endorsed by the International Labour Organisation (ILO).

Employment: Employed persons are those who, during the reference week:

- a) did any work at all at a job or business, that is, paid work in the context of an employer-employee relationship, or self-employment. It also includes unpaid family work, which is defined as unpaid work contributing directly to the operation of a farm, business or professional practice owned and operated by a related member of the same household; or
- b) had a job but were not at work due to factors such as own illness or disability, personal or family responsibilities, vacation, labour dispute or other reasons (excluding persons on layoff, between casual jobs, and those with a job to start at a future date).

Unemployment: Given the concept of unemployment as the unutilized supply of labour, the operational definition of unemployment is based primarily on the activity of job search and the availability to take a job. In addition to being conceptually appropriate, job search activities can, in a household survey, be objectively and consistently measured over time. The definition of unemployment is therefore the following:

Unemployed persons are those who, during reference week:

- a) were on temporary layoff during the reference week with an expectation of recall and were available for work, or
- b) were without work, had actively looked for work in the past four weeks, and were available for work, or
- c) had a new job to start within four weeks from reference week, and were available for work.

Persons are regarded as available if they reported that they could have worked in the reference week if a suitable job had been offered (or recalled if on temporary layoff); or if the reason they could not take a job was of a temporary nature such as: because of own illness or disability, personal or family responsibilities, because they already have a job to start in the near future, or because of vacation (prior to 1997, those on vacation were not considered available). Full-time students currently attending school and looking for full-time work are not considered to be available for work during the reference week. They are assumed to be looking for a summer or co-op job or permanent job to start sometime in the future, and are therefore not part of the current labour supply.

Note that in the above definition there are two groups for which job search is not required: persons on temporary layoff and persons with a job to start at a definite date in the future. Persons on layoff are included among the unemployed on the grounds that their willingness to supply labour services is apparent in their expectation of returning to work. A similar argument is applied for persons who will be starting at a new job in four weeks or less.

Finally, for the purposes of measuring job search as part of the identification of the unemployed, the LFS uses a four-week search period although the reference period for identifying the employed is that of one week. The justification for the difference is that delays inherent in job search (for example, periods spent awaiting the results of earlier job applications) require that the active element of looking for work be measured over a period greater than one week if a comprehensive measure of job search is to be obtained.

Not in the labour force: Persons not in the labour force are those who, during the reference week, were unwilling or unable to offer or supply labour services under conditions existing in their labour markets, that is, they were neither employed nor unemployed.

3.0 Labour Force Survey concepts and definitions

Provided below is the LFS dictionary with definitions of terms and variables associated with the survey. Where appropriate, changes to definitions through time are documented.

Absence from work: See *Hours (Hours lost)*.

Activity prior to unemployment: Main activity before looking for work. Distinguishes between those who were **working** (that is, job leavers, job losers and temporary layoffs) and those who were not in the labour force but were **keeping house, going to school**, or involved in some **other** type of activity.

Age: Age is collected for every household member in the survey, and the information on labour market activity is collected for all persons aged 15 and over. Prior to 1966, information on labour market activity was collected for persons aged 14 and over. Beginning January 1997, date of birth is collected to ensure inclusion of respondents who turn 15 during their six month rotation in the survey.

Availability: Persons are regarded as available if they reported that they could have worked in the reference week if a suitable job had been offered (or recalled if on temporary layoff); or if they could not take a job because of their own illness or disability, personal or family responsibilities, because they already have a job to start in the near future, or because of vacation (prior to 1997, those on vacation were not considered available). Full-time students currently attending school and looking for full-time work are not considered to be available for work during the reference week. They are assumed to be looking for a summer or co-op job or permanent job to start sometime in the future.

Average hours worked: See *Hours*.

Average weekly/hourly wages: See *Wages*.

Census metropolitan area (CMA) and census agglomeration (CA): Are large urban areas (known as urban cores) together with adjacent urban and rural areas (known as urban and rural fringes) that have a high degree of social and economic integration with the urban cores. A CMA has an urban core population of at least 100,000 and a CA has an urban core population between 10,000 and 100,000 based on the previous census.

The following areas distinguish between central and peripheral urban and rural areas **within** CMAs and CAs:

Urban core: is a large urban area within a CMA or a CA that must have a population of at least 100,000 in the case of a CMA, or between 10,000 and 99,999 in the case of a CA based on the previous census and have a population density of at least 400 per square kilometre.

Urban fringe is the urban area within a CMA or CA that is not contiguous to the urban core. It has a minimum population concentration of 1,000 and a population density of at least 400 per square kilometre, based on the previous census population counts.

Rural fringe is all territory within a CMA or CA not classified as urban core or urban fringe.

Class of worker: See *Status in employment*.

Discouraged searcher: (also called Discouraged worker) Since 1997, discouraged searchers are defined as those persons who reported wanting to work at a job or business during reference week and were available but who did not look for work because they believed no suitable work was available. Prior to January 1997, the definition of discouraged searcher was limited to those who looked for work within the previous 6 months but not during the last 4 weeks although they were available, and did not look because they believed no suitable work was available. The change in concept and question wording results in a complete break in the series.

Duration of joblessness: Number of months or years elapsed since persons who are not currently employed last worked, provided that they worked at some time in the past.

Duration of unemployment: Number of continuous weeks during which a person has been on temporary layoff or without work and looking for work. Respondents are required to look for work at least once every four weeks, they are not required to undertake job search activities each week in order to be counted as unemployed. The LFS measures the duration of incomplete spells of unemployment, since the information is collected only from those currently unemployed. A spell of unemployment is interrupted or completed by any period of work or withdrawal from the labour force.

Dwelling: Any set of living quarters that is structurally separate and has a private entrance outside the building or from a common hall or stairway inside the building.

Earnings: See *Wages*.

Economic region: LFS economic regions (ERs) have been established at each decennial sample redesign in consultation with the provinces. The regions generally correspond to regions used by the province for administrative and statistical purposes. The LFS ERs coincide with the official Subprovincial Regions (SPRs) defined by Standards Division in consultation with the provinces, for use in dissemination of subprovincial data by Statistics Canada.

Educational attainment: Highest level of schooling completed. Questions relating to educational attainment were changed in 1990, to better capture the relationship between educational attainment and labour market outcomes.

From 1976 to 1989: data on primary and secondary education reflected the number of years of primary and secondary education completed. In the case of those whose highest level was grades 11 through 13, no attempt was made to determine if the respondent had actually graduated. However, post-secondary education was limited to the education which normally requires high school graduation. In addition, information on type of post-secondary was limited to three categories: 1) some post-secondary; 2) post-secondary certificate or diploma; 3) university degree.

Beginning January 1990: data on primary and secondary education reflects the highest grade completed. This provides a more consistent measure for those who accelerate or fail a grade than did years of school. A question on high school graduation has also been added since it is generally believed that persons who have never completed their secondary education have greater difficulty competing in the labour market. With the new questions, any education that could be counted towards a degree, certificate or diploma from an educational institution is taken as post-secondary education. The change allows more persons into the post-secondary education category. For example, trades programs offered through apprenticeship, vocational schools or private trade schools do not always require high school graduation. Such education is considered as post-secondary while only primary or secondary would have been recognised prior to 1990. Finally, more information is collected on the type of post-secondary education: 1) some post-secondary; 2) trades certificate or diploma from a vocational or apprenticeship training; 3) Non-university certificate or diploma from a community college, CEGEP,

school of nursing, etc.; 4) University certificate below bachelor's degree; 5) Bachelors degree; and 6) University degree or certificate above bachelor's degree.

Employee: See *Status in employment*.

Employment: Employed persons are those who, during the reference week did any work for pay or profit, or had a job and were absent from work. See *Section 2.1: Determining labour force status for more detail*.

Employment rate: (employment/population ratio) Number of employed persons expressed as a percentage of the population 15 years of age and over. The employment rate for a particular group (age, sex, marital status, province, etc.) is the number employed in that group expressed as a percentage of the population for that group.

Establishment size: See *Workplace size*.

Extra hours worked: See *Hours*.

Family: The LFS identifies families according to the criteria for "Economic families": a group of two or more persons who live in the same dwelling and who are related by blood, marriage (including common-law) or adoption. A person living alone or who is related to no one else in the dwelling where he or she lives is classified as an unattached individual.

Firm size: Beginning January 1998, the number of employees at all locations of the employer is collected from employees. Responses are recorded according to the following size groups: less than 20, 20 to 99, 100 to 500, more than 500.

Flows into unemployment: Characterises the unemployed in terms of their activity immediately prior to looking for work. See *Job leavers, Job losers, Re-entrants and New entrants*.

Full-time employment: See *Type of work*.

Future starts: Persons who did not have a job during the survey reference week and did not search for work within the previous four weeks, but were available to work and had a job to start within the next four weeks. These persons are classified as unemployed, despite the lack of job search within the previous four weeks, since it is apparent that they are part of the current supply of labour. In contrast, those with jobs to start at a later time than the next four weeks are designated as long-term future starts and are classified as not in the labour force since they are not part of current labour supply.

Goods-producing industries (or goods sector, or goods industries): Includes agriculture; forestry, fishing, mining, and oil and gas extraction; utilities (electric power, gas and water); construction; and manufacturing.

Government sector: See *Public/private sector*.

Head of family: See *Reference person*.

Hours

Variable hours: Beginning January 1997, information is collected to determine if the number of hours worked varies from week to week. In these cases, usual hours worked are calculated as the average of the hours worked in the last 4 weeks.

Actual hours worked: Number of hours actually worked by the respondent during the reference week, including paid and unpaid hours.

Usual hours worked: Prior to January 1997, usual hours were the number of hours usually worked by the respondent in a typical week, regardless of whether they were paid. Beginning January 1997, usual hours for employees refers to their normal paid or contract hours, not counting any overtime. However, the definition of usual hours remains unchanged for the self-employed and unpaid family workers.

Average hours worked: Average number of hours worked per week, usual or actual, is calculated by dividing total hours by the total number of employed persons.

Also available is the average number of actual hours worked per week calculated by excluding persons who were not at work during the reference week.

Hours lost (absence from work): A distinction is made between those who lose hours from work because they missed part of the work week or the full work week. Reasons for the absence are collected for both situations.

Part-week absence: Collected for employees only. Reasons for absence include: own illness or disability, personal or family responsibilities, maternity leave (females only), vacation, weather, labour dispute, job started or ended during reference week, holiday, working short time, other reasons.

Full-week absence: Collected for all employed persons. Reasons for absence include: own illness or disability, personal or family responsibilities, maternity leave (females only), vacation, labour dispute, work schedule, self-employed (no work available), seasonal business (self-employed), other reasons. The number of full weeks absent from work are recorded. In addition, employees and self-employed with an incorporated business are asked if they received wages or salary for any time off in reference week.

Extra hours worked: The number of hours worked during the reference week in excess of the usual hours reported in all jobs combined. It includes all extra hours, whether the work was done at a premium or regular wage rate, or without pay. Since January 1997, extra hours are collected from employees only, in the form of 2 questions: number of paid overtime hours worked in reference week, and number of extra hours worked without pay.

Paid overtime: includes any hours worked during the reference week over and above standard or scheduled paid hours, for overtime pay or compensation (including time off in lieu).

Extra hours without pay (unpaid overtime): refers to time spent directly on work or work-related activities over and above scheduled paid hours. These must be extra hours worked for which the respondent received no additional compensation.

Household: Any person or group of persons living in a dwelling. A household may consist of any combination of: one person living alone, one or more families, a group of people who are not related but who share the same dwelling.

Industry: The general nature of the business carried out by the employer for whom the respondent works (main job only). If a person did not have a job during the survey reference week, the information is collected for the last job held, providing the person worked within the previous twelve months.

Involuntary part-time: See *Reason for working part-time*.

Job leavers: Persons currently not employed, who last worked within the previous year and left that job voluntarily. That is, the employer did not initiate the termination. Detailed reasons collected are: own illness, personal or family responsibilities, going to school, no specific reason, changed residence, dissatisfied with job, retired. Since 1997, further detail is available, including business sold or closed down (self-employed only), pregnancy.

Job losers: Persons currently not employed, who last worked within the previous year and left that job involuntarily (employer initiated because of business conditions, downsizing etc.). Prior to 1997, this category was broken down into those on temporary layoff and those laid off on a permanent basis. Since January 1997, more detail for reason for permanent layoff is available: end of seasonal job; end of temporary, term or contract job; casual job, no work; company moved; company went out of business; laid off due to business conditions with no expectation of recall; dismissal by employer; other reasons.

Job permanency: Beginning January 1997, information is collected to allow the classification of paid jobs as either permanent or temporary. This classification is based on the intentions of the employer, and characteristics of the job, rather than the intentions of the employee. If a job that was formerly considered permanent is ending in the near future because of downsizing or closure, it is still regarded as permanent.

Permanent: A permanent job is one that is expected to last as long as the employee wants it, given that business conditions permit. That is, there is no pre-determined termination date.

Temporary: A temporary job has a predetermined end date, or will end as soon as a specified project is completed. Information is collected to allow the sub-classification of temporary jobs into four groups: seasonal; temporary, term or contract, including work done through a temporary help agency; casual job; and other temporary work.

Job search: See *Methods of job search*.

Job security: See *Job permanency*.

Job tenure: The number of consecutive months or years a person has worked for the current (or, if employed within the previous twelve months, the most recent) employer. The employee may have worked in one or more occupations or in one or more locations, or have experienced periods of temporary layoff with recall and still be considered to have continuous tenure if the employer has not changed. But if a person has worked for the same employer over different periods of time, job tenure measures the most recent period of uninterrupted work.

Labour force: Civilian non-institutional population 15 years of age and over who, during the survey reference week, were employed or unemployed. Prior to 1966, persons aged 14 and over were covered by the survey.

Labour force by industry of occupation: See *Unemployment by industry or occupation*.

Labour force status: Designates the status of the respondent vis-à-vis the labour market: a member of the non-institutional population 15 years of age and over is either **employed**, **unemployed**, or **not in the labour force**. See *Section 2.1: Determining labour force status for more detail*.

Main job: When a respondent holds more than one job or business, the job or business involving the greatest number of usual hours worked is considered to be the main job. The full or part-time status and industry and occupation information available from the survey refer to the main job, as does information for employees on wages, union status, job permanency, and workplace size.

Marital status: Refers to the marital status reported by the respondent. No differentiation is made between married and common-law relationships, both are classified as married in the survey. The classification of single is reserved for those who have never married, otherwise, respondents are classified as either widowed or separated/divorced.

Methods of job search: Identifies the various methods of job search activities undertaken by unemployed persons in the previous four weeks. If more than one method is used, each one is recorded. Search methods include: checked with public employment agency, private employment agency, union, employers directly, friends or relatives, placed or answered ads, looked at job ads, other methods.

Multiple jobholders: Persons who, during the reference week, were employed in two or more jobs simultaneously. This group is sometimes referred to as “Moonlighters”.

New entrants: Persons entering the labour force in search of their first job (unemployed).

Not in the labour force: See *Section 2.1: Determining labour force status for more detail.*

Occupation: Refers to the kind of work persons were doing during the reference week, as determined by the kind of work reported and the description of the most important duties. For those not currently employed, information on occupation is collected for the most recent job held within the previous year.

Other job: (See also *Main job*) Information collected on the second job of multiple job holders and the old job of those who changed jobs during reference week is limited to: usual hours, actual hours worked, and status in employment.

Paid worker: See *Status in employment.*

Participation rate: Total labour force expressed as a percentage of the population aged 15 years and over. The participation rate for a particular group (for example, women aged 25 years and over) is the labour force in that group expressed as a percentage of the population for that group.

Part-time employment: See *Type of work and Reason for working part-time.*

Permanent job: See *Job permanency.*

Personal or family responsibilities: Beginning January 1997, more detail is collected on the personal or family reasons for the following data items: reason for absence from work, reason for leaving last job, reason for working part-time, and reason for not looking for work. The greater detail includes a) caring for own children; b) caring for elder relative, and c) other personal or family reasons. Pregnancy is also included in the response list for the question on reason for leaving last job, and maternity is included in the response list for the question on reason for absence from work.

Population: The target population covered by the survey corresponds to all persons aged 15 years and over residing in the provinces of Canada, with the exception of the following: persons living on Indian reserves, full-time members of the regular Armed Forces, and persons living in institutions (for example, inmates of penal institutions and patients in hospitals or nursing homes who have resided in the institution for more than six months).

Public/private sector employment: The public sector includes employees in public administration at the federal, provincial and municipal levels, as well as in Crown corporations, liquor control boards and other government institutions such as schools (including universities), hospitals and public libraries. The private sector comprises all other employees and self-employed owners of businesses (including unpaid family workers in those businesses), and self-employed persons without businesses. The definition was changed in January 1999 in order to

harmonize LFS data for the public and private sectors to the System of National Accounts standard. Prior to January 1999, “ownership” rules were used as the basis for classification of health care institutions and universities to the public sector by the LFS. Since January 1999, “funding” rules are used. As a result, all employees in universities and hospitals are now classified in the public sector. All historical data were revised to reflect this new definition. Thus, there is no break in public and private sector series.

Reason for leaving last job: Asked of all persons classified as unemployed or not in the labour force who last worked within the previous year. *See Job Losers and Job Leavers for detailed reasons.*

Reason for not looking for work: Beginning January 1997, asked of those who were not employed and did not search for work, but said they wanted work during reference week. Prior to 1997, asked of persons who had looked for work in the previous six months but not during the past four weeks. *See also Discouraged searchers.*

Reason for time lost/absence from work: *See Hours lost.*

Reason for working part-time: (*See also Type of work*) Prior to the introduction of the revised questionnaire in January 1997, the question on reason for working part-time was asked of all persons whose total usual work hours at all jobs or businesses were below 30 per week. Reasons included: own illness, personal or family responsibilities, going to school, could only find part-time work, did not want full-time work, other, and full-time work less than 30 hours. This last category of respondents were redefined as full-time workers and not counted in any part-time estimates. The involuntary part-time rate was calculated by dividing those who reported they could only find part-time work by the total employed part-time.

Beginning January 1997, all respondents who usually worked less than 30 hours per week at their main or only job are asked if they want to work more or less than 30 hours at a (single) job or business. Depending on the response, the main reason for working part-time is collected. For those who respond that they want to work less than 30 hours, the main reason for not wanting to work 30 or more hours per week is collected. Responses include: own illness, personal or family responsibilities, going to school, personal preference, other.

For those who respond that they want to work 30 or more hours per week, the main reason for working less than 30 hours is collected. Responses include: own illness, personal or family responsibilities, going to school, business conditions, could not find work with 30 or more hours, other. Those whose response is “business conditions” or “could not find work with 30 or more hours” are then asked if they looked for work with 30 or more hours during the past four weeks. Those who searched for full-time work are considered to be involuntary part-time workers. The involuntary part-time rate is calculated by dividing this group by the total number of persons working part-time at their main or only job. The change in concepts and definitions introduced in January 1997 results in a complete break in the involuntary part-time series.

Re-entrants: Persons currently unemployed who had worked in the past and were out of the labour force for some time following separation from their last job.

Reference person: At the time of interview the respondent designates a reference person for the family. The reference person is normally an adult with responsibility for the care or support of the family. The relationship of each family member to that reference person is recorded. *See also Relationship to family reference person.*

Reference week: The entire calendar week (from Sunday to Saturday) covered by the Labour Force Survey each month. It is usually the week containing the 15th day of the month. The interviews are conducted during the following week, called the Survey Week, and the labour force status determined is that of the reference week.

Relationship to family reference person: Relationship of each family member to the person who has been identified as the reference person (i.e. someone with responsibility for the care or support of the family). Relationships include: self, spouse, son or daughter, grandchild, son or daughter-in-law, foster child, parent, parent-in-law, brother or sister, other relative.

Retirement age: The Labour Force Survey asks people who are not working, and who have left their last job within the year prior to being surveyed, why they left this job. One of the response categories is “retired.” The average or median retirement age is calculated from this variable. For a complete description of who is represented and how the age is calculated, please refer to the article “Measuring retirement age” in *Perspectives on Labour and Income*, catalogue number 75-001-XWE, Summer 1997 issue.

Returning students: Since a majority of students are not attending school during the summer, supplementary questions are asked from May to August to identify those who are on summer break so that their labour market situation can be monitored. Youths (aged 15 to 24) are given the status of “returning student” if they reported that they were attending school full-time in the previous March and intend to return to school full-time in the fall. Information is also available for those who were full-time students in the previous March but do not intend to return to school full-time or are unsure of their intentions.

Rural and small town areas: Areas outside the commuting zone of Census Metropolitan Areas (CMAs) and Census Agglomerations (CAs). This includes:

Rural areas, which are sparsely populated lands lying outside small towns, villages and other populated places, with less than 1,000 population according to the previous census.

Small towns, which are urban areas with a population of 1,000 to 9,999 and with a population density of 400 inhabitants per square kilometre, based on the previous census.

School attendance: Establishes whether or not a respondent is attending an educational establishment. For those who are students, information is collected on the type of school, and whether enrolment is full or part-time, as designated by the educational establishment.

Seasonal adjustment: Fluctuations in economic time series are caused by seasonal, cyclical and irregular movements. A seasonally adjusted series is one from which seasonal movements have been eliminated. Seasonal movements are defined as those which are caused by regular annual events such as climate, holidays, vacation periods and cycles related to crops, production and retail sales associated with Christmas and Easter. It should be noted that the seasonally adjusted series contain irregular as well as longer-term cyclical fluctuations.

The seasonal adjustment program is a complicated computer program which differentiates between these seasonal, cyclical and irregular movements in a series over a number of years and, on the basis of past movements, estimates appropriate seasonal factors for current data. On an annual basis, the historic series of seasonally adjusted data are revised in light of the most recent information on changes in seasonality.

Self-employed: See *Status in employment*.

Seniority: See *Job tenure*.

Service-producing industries (or service sector or service industries): Includes trade; transportation and warehousing; finance, insurance, real estate and leasing; professional, scientific and technical services; management, administrative and other support; educational services; health care and social assistance; information, culture and recreation; accommodation and food services; other services; and public administration.

Status in employment: There are two broad categories of workers: Those who work for others and those who work for themselves. These two groups can be subdivided into a total of seven classes of workers as described below. (*See also Public and private employment*).

Employees:

Employees (private): Those who work as employees of a private firm or business.

Employees (government or public): Those who work for a local, provincial or federal government, for a government service or agency, a crown corporation, or a government funded public establishment such as a school (including universities) or a hospital. The definition was changed in January 1999 in order to harmonize LFS data for the public and private sectors to the System of National Accounts standard. Prior to January 1999, “ownership” rules were used as the basis for classification of health care institutions and universities to the public sector by the LFS. Since January 1999, “funding” rules are used. As a result, all employees in universities and hospitals are now classified in the public sector. All historical data were revised to reflect this new definition. Thus, there is no break in public and private sector series.

Self-employed:

Working owners of incorporated businesses: Working owners of an incorporated business, farm or professional practice. This group is further subdivided as follows:

With paid help
Without paid help

Working owners of unincorporated businesses and other self-employed: Working owners of a business, farm or professional practice that is not incorporated and self-employed persons who do not have a business (for example, baby-sitters, newspaper carriers). This group is further subdivided as follows:

With paid help
Without paid help

Unpaid family workers: Persons who work without pay on a farm or in a business or professional practice owned and operated by another family member living in the same dwelling.

Note: The definition of a **paid worker** may vary depending on the nature of the analysis. Those concerned with estimating the number of workers associated with total labour income usually include both employees and the self-employed with an incorporated business in estimates of paid workers. In contrast, most labour market analysts include only employees in paid worker estimates, while incorporated owners are grouped with the rest of the self-employed.

Student: *See School attendance and returning students.*

Temporary layoff: Persons on temporary layoff are employees who did not work during the reference week because they had been temporarily released by their employer due to business conditions (not enough work, drop in orders or sales, retooling etc.). They must have a definite date to return to work, or an indication from their employer that they will be recalled in the future, and they must be available for work during the reference week. Persons on temporary layoff are not required to undertake any job search in order to be counted as unemployed.

Prior to January 1997 the wording and structure of the questionnaire was such that it was likely that a number of persons on temporary layoff were not identified as such, and were classified as “not in the labour force” rather than “unemployed”. The 1997 redesign addressed this problem, resulting in a higher number of identified persons on temporary layoff. These changes result in a break in the temporary layoff series. Since those on temporary layoff account for a small proportion of the unemployed (less than 10%) the impact of these changes on the overall unemployment rate is negligible.

Temporary work: See *Job permanency*.

Tenure: See *Job tenure*.

Type of work: Full-time or part-time work schedule.

Full-time employment consists of persons who usually work 30 hours or more per week at their main or only job.

Part-time employment consists of persons who usually work less than 30 hours per week at their main or only job. This information is available for those currently employed or who last worked within the previous year. Note: prior to 1996, full-time and part-time had been defined according to usual hours at all jobs, and those who considered their work schedule of less than 30 hours per week to be full-time work were classified as full-time workers. In January 1996, when the definition was revised, all historical data and records were adjusted to reflect this new definition. Thus, there is no break in part-time and full-time data series.

Type of work sought: Identifies whether a job searcher is looking for full-time or part-time work. Unemployed persons on temporary layoff are classified as looking for full or part-time on the basis of their usual hours at their former job. This information is not available for non-searchers who are classified as unemployed because they have a job to start in the next four weeks (future-starts).

Unattached individuals: Persons who live alone or who are not related to anyone else in the household. They are excluded from the family member counts.

Unemployment: Unemployed persons are those who, during reference week, were available for work and were either on temporary layoff, had looked for work in the past four weeks or had a job to start within the next four weeks. See *Section 2.1: Determining labour force status for more detail*.

Unemployment by industry/occupation: The LFS produces information on the number of unemployed, the unemployment rate and the labour force by industry and occupation. The basis for these categories is industry or occupation of last job for those currently unemployed who have held a job in the previous year. It is important to note that no information is collected on industry or occupation of job search. Thus, these data should be interpreted with caution. For example, a recent graduate of law school looking for work as a lawyer in a law firm, may have last held a job as a waiter in a restaurant. For this person, unemployment is attributed to the personal service industry and the services occupation.

Unemployment rate: Number of unemployed persons expressed as a percentage of the labour force. The unemployment rate for a particular group (age, sex, marital status, etc.) is the number unemployed in that group expressed as a percentage of the labour force for that group.

Union status: Beginning January 1997, employees are classified as to their union status: a) union member; b) not a member but covered by a union contract or collective agreement; or c) non-unionised.

Unpaid family workers: *See Status in employment.*

Usual hours worked: *See Hours.*

Wages: Since January 1997, information is collected on the usual wages or salary of employees at their main job. Respondents are asked to report their wage/salary before taxes and other deductions, and include tips, commissions and bonuses. Weekly and hourly wages/salary are calculated in conjunction with usual paid work hours per week. Average hourly wages, average weekly wages, and wage distributions can then be cross-tabulated by other characteristics such as age, sex, education, occupation, and union status. Those who are paid on an hourly basis are also identified.

Work: Includes any work for pay or profit, that is, paid work in the context of an employer-employee relationship, or self-employment. It also includes work performed by those working in family business without pay (unpaid family workers).

Workplace size: Beginning January 1997, the number of employees at the location of employment (i.e. building or compound) is collected from employees. Responses are recorded according to the following size groups: less than 20, 20 to 99, 100 to 500, more than 500. The concept of location of employment approximates the concept of establishment used by many Statistics Canada business surveys.

4.0 Survey methodology

4.1 Population coverage

The LFS is a monthly household survey of a sample of individuals who are representative of the civilian, non-institutionalized population 15 years of age or older in Canada's ten provinces. Specifically excluded from the survey's coverage are residents of the Yukon, Northwest Territories and Nunavut, persons living on Indian Reserves, full-time members of the Canadian Armed Forces and inmates of institutions. These groups together represent an exclusion of approximately 2% of the population aged 15 or over.

4.2 Sample design

The LFS has undergone a redesign, culminating in the introduction of the new design at the end of 2004. Every five years, Labour Force Survey (LFS) estimates undergo extensive revisions. At the beginning of 2005, revisions have included four major changes. All estimates have been adjusted to reflect 2001 Census population counts; industry estimates have been classified from the 1997 to the 2002 North American Industry Classification System (NAICS); the occupation estimates will be classified from the 1991 Standard Occupation Classification to the National Occupational Classification for Statistics 2001 (NOC-S). Lastly, geography boundaries are based on the 2001 Standard Geographical Classification (SGC), which affect boundaries of census metropolitan areas only.

The LFS sample is based upon a stratified, multi-stage design employing probability sampling at all stages of the design. At the first stage, geographical areas, named primary sampling units (PSU), are selected. A list of dwellings is created for each selected PSU. At the second stage, dwellings are selected from these lists.

The LFS follows a rotating panel sample design, in which households remain in sample for six consecutive months. The total sample consists of six representative sub-samples or panels, and each month a panel is replaced after completing its six month stay in the survey. Outgoing households are replaced by households in the same or a similar area. This results in a five-sixths month-to-month sample overlap, which makes the design efficient for estimating month-to-month changes. The rotation after six months prevents undue respondent burden for households that are selected for the survey. Because of the rotation group feature, it is possible to readily conduct supplementary surveys using the LFS sample but employing less than the full size sample.

The design principles are the same for each province.

4.2.1 Primary sampling unit (PSU) delineation

For the first time in 2004, PSUs were delineated for the entire country. These were created by grouping blocks until each PSU covered the targeted number of households. The methodology used to create the PSUs ensured that they are made of contiguous pieces and are compact.

4.2.2 Primary stratification

In order to increase the efficiency of the sample design, PSUs are grouped into strata. The strata must respect broad geographical areas to improve the estimate's quality for these areas.

Provinces are divided into Economic Regions (ER) and Employment Insurance Economic Regions (EIER). ERs are geographic areas of more or less homogeneous economic structure formed on the basis of federal-provincial agreements. They are relatively stable over time. EIERs are geographic areas used by Human Resources and Skills Development Canada, and are roughly the same size and number as ERs, but they do not share the same definitions. Labour force estimates are produced for both sets of regions.

The intersections of the two sets of regions form the first level of stratification for the LFS. These ER/EIER intersections are treated as primary strata and further stratification is carried out within them (see Section 4.2.4). Note that a third set of regions, Census Metropolitan Areas (CMA), is also respected by stratification in the current LFS design. Most of these CMAs are also EIERs.

[Click here](#) to view the Labour Force Survey Reference Maps by Census metropolitan areas (CMAs) or Economic regions (ERs)

4.2.3 Types of areas

The primary strata (ER/EIER intersections) are further disaggregated into five types of areas: high vacancy, aboriginal, rural, urban and remote areas. The previous design included an apartment frame. This frame no longer exists in the current design. It was dismissed since it was expensive to maintain and did not provide large gains in efficiency. The 1994 design also used a three stage rural design. This design is no longer used since it did not deal properly with specific collection issues associated with these low population density areas. Many of the areas covered by the three stage rural design in 1994 are now part of the high vacancy type.

The high vacancy type was used for the first time in 2004. It groups PSUs that have a high vacancy rates. It was introduced to deal properly with the special collection needs of these areas.

PSUs having a large proportion of aboriginal people in the four western provinces are set aside in aboriginal strata. This innovation was introduced to address the growing demand for aboriginal estimates. Aboriginal strata were not defined in the remaining six provinces since their aboriginal population is much smaller.

Urban and rural areas are loosely based on the Census definitions of urban and rural, with some exceptions to allow for the formation of strata in some areas. Urban areas include the largest CMAs down to the smallest villages categorized by the 2001 Census as urban (1,000 people or more), while rural areas are made up of areas not designated as urban or remote.

Approximately 1% of the LFS population is found in remote areas which are less accessible to LFS interviewers than other areas. For administrative purposes, this portion of the population is sampled separately through the remote area frame. Some populations, in areas having very low population density, are excluded from the sampling frame.

4.2.4 Secondary stratification

CMAs are further stratified into regular urban strata, high income strata and immigrant strata. The immigrant strata are introduced in the design to provide more flexibility to sample this sub-population. As in the previous design, the high income strata are used to improve coverage of the high income sub-population.

Large urban areas outside of CMAs also fall in the regular urban stratum type. Where it is possible and/or necessary, these regular strata are further stratified to improve the design's efficiency. Most urban areas fall into the regular urban strata, which, in fact, cover the majority of Canada's population.

Small urban areas surrounded by remote areas are set aside in a special type labelled isolated urban towns. For these, strata are formed geographically. It is important to note that an additional sampling stage is used for these areas in order to reduce collection costs.

When possible and/or necessary, PSUs in rural areas are stratified optimally to improve design efficiency. In each province, remote settlements are sampled proportional to the number of dwellings in the settlement, with no further stratification taking place. Dwellings are selected using systematic sampling in each of the places sampled.

4.2.5 Cluster and dwelling selection

For all stratum types, except the isolated urban town type, the selection of a sample of PSUs (always six or a multiple of six) from each stratum represents the first stage of sampling. For each selected PSU, a list of dwellings is needed to select the sample of dwellings.

For the first time in 2004, the Address Register (AR) was used to reduce listing costs and reduce frame under-coverage. The Address Register is a frame of residential civic numbered addresses. Addresses are linked to both STC and Canada Post Corporation geographies and were verified during the 1991, 1996 and 2001 Censuses. Since 2001 the base has been supplemented with addresses from administrative files, the Telephone Billing Files and the GST Housing Rebate File. Although the file is not national it currently covers approximately 88% of the dwelling stock. Different strategies were developed to use the AR. If the AR is of very good quality for a PSU, the sample of dwellings is selected directly from a list of addresses extracted from the AR. When the AR is of good quality, the list of addresses extracted from the AR, along with a PSU map, is sent to a field interviewer for validation. The sample is then selected from the validated list. Finally, when AR quality is inadequate, only a map is sent to a field interviewer. The interviewer visits the PSU and lists all private dwellings. Sampled dwellings are selected from this list.

For isolated urban town strata, two towns are first selected from each stratum. Within selected towns, we select a sample of smaller geographical areas or Secondary Sampling Unit (SSU). Maps of selected SSUs are sent to field interviewers. The interviewers list all private dwellings in the SSUs and samples of dwellings are selected from these lists.

The number of dwellings selected in a PSU (or SSU) depends on its stratum type. For example, in the urban area frame, sample yields are either six or eight dwellings, depending on the size of the city. In rural strata, each PSU yields 10 dwellings. In all PSUs and SSUs, dwellings are sampled systematically. This represents the final stage of sampling.

4.2.6 Person selection

Demographic information is obtained for all persons in a household for whom the selected dwelling is the usual place of residence. All civilian household members 15 years of age or older are automatically selected and labour force information is obtained for each one of them. Respondent burden is minimized for the elderly (age 70 and over)

by carrying forward their responses for the initial interview to the subsequent five months in the survey.

4.3 Sample size

The sample size of eligible persons in the LFS is determined so as to meet the statistical precision requirements for various labour force characteristics at the provincial and sub-provincial level, to meet the requirement of federal, provincial and municipal governments as well as a host of other data users.

The monthly LFS sample consists of approximately 50,000 dwellings. After excluding dwellings found to be vacant, dwellings demolished or converted to non-residential uses, dwellings containing only ineligible persons, dwellings under construction, and seasonal dwellings, about 53,400 dwellings remain which are occupied by one or more eligible persons. From these dwellings, LFS information is obtained for approximately 102,000 civilians aged 15 or over.

The following table shows the targeted number of households in the LFS sample. The LFS sample size fluctuates from month to month and can evolve over time since partners sometimes finance a sample increase.

The figures provided in the table include 1,250 and 1,100 additional sampled households in Alberta and British Columbia, respectively. These increases are financed by the corresponding provincial governments to improve the quality of aboriginal estimates.

Province	Sample size
Newfoundland and Labrador	1,965
Prince Edward Island	1,400
Nova Scotia	3,005
New Brunswick	2,858
Quebec	9,956
Ontario	15,605
Manitoba	3,783
Saskatchewan	3,926
Alberta	5,368
British Columbia	5,783
Canada	53,648

5.0 Data collection

Data collection for the Labour Force Survey (LFS) is carried out each month during the week following the LFS reference week. The reference week is normally the week containing the 15th day of the month.

5.1 Interviewing for the Labour Force Survey

Statistics Canada interviewers are employees hired and trained to carry out the LFS and other household surveys. Each month they contact the sampled dwellings to obtain the required labour force information.

LFS interviews are conducted by telephone by interviewers working out of a regional office CATI (Computer Assisted Telephone Interviews) site or by personal visit from a field interviewer. Since 2004, dwellings new to the sample in urban areas are contacted by telephone if the telephone number is available from administrative files, otherwise the dwelling is contacted by a field interviewer. The interviewer first obtains socio-demographic information for each household member and then obtains labour force information for all members aged 15 and over who are not members of the regular armed forces. The majority of subsequent interviews are conducted by telephone. In subsequent monthly interviews the interviewer confirms the socio-demographic information collected in the first month and collects the labour force information for the current month.

In each dwelling, information about all household members is usually obtained from one knowledgeable household member. Such 'proxy' reporting, which accounts for approximately 65% of the information collected, is used to avoid the high cost and extended time requirements that would be involved in repeat visits or calls necessary to obtain information directly from each respondent.

If, during the course of the six months that a dwelling normally remains in the sample, an entire household moves out and is replaced by a new household, information is obtained about the new household for the remainder of the six-month period.

5.2 Supervision and quality control

All LFS interviewers are under the supervision of a staff of senior interviewers who are responsible for ensuring that interviewers are familiar with the concepts and procedures of the LFS and its many supplementary surveys, and also for periodically monitoring their interviewers. The senior interviewers are, in turn, under the supervision of the LFS program managers.

5.3 Non-response to the Labour Force Survey

Non-response to the LFS tends to average about 7% of eligible households. Interviewers are instructed to make all reasonable attempts to obtain LFS interviews with members of eligible households. For individuals who at first refuse to participate in the LFS, a letter is sent from the Regional Office to the dwelling address stressing the importance of the survey and the household's co-operation. This is followed by a second call (or visit) from the interviewer. For cases in which the timing of the interviewer's call (or visit) is inconvenient, an appointment is arranged to call back at a more convenient time. For cases in which there is no one home, numerous call backs are made. Under no circumstances are sampled dwellings replaced by other dwellings for reasons of non-response.

Each month, after all attempts to obtain interviews have been made, a small number of non-responding households remain. For households non-responding to the LFS, a weight adjustment is applied to account for non-responding households.

6.0 Data processing

This chapter presents a brief summary of the processing steps involved in producing this file.

6.1 Data capture

Since 1994, responses to survey questions are captured directly by the interviewer at the time of the interview using a computerized questionnaire on a lap-top or desktop computer. The computerized questionnaire reduces processing time and costs associated with data entry, transcription errors, and data transmission. The response data are encrypted to ensure confidentiality and sent via modem to the appropriate Statistics Canada Regional Office. From there they are transmitted over a secure line to Ottawa for further processing. Prior to the introduction of computer assisted interviewing (CAI), information was recorded by the interviewer on a paper questionnaire, which was then sent for data capture in the Regional Office before transmission to Ottawa.

6.2 Editing

Some editing is done directly at the time of interview. Where the information entered is out of range (too large or small) of expected values, or inconsistent with previous entries, the interviewer is prompted, through message screens on the computer, to modify the information. However, interviewers have the option of bypassing the edits, and of skipping questions if the respondent does not know the answer or refuses to answer. Therefore, the response data are subjected to further edit and imputation processes once they arrive in head office.

The editing and imputation phases of processing involve the identification of logically inconsistent or missing information items, and the modification of such conditions. Since the true value of each entry on the questionnaire is not known, the identification of errors can be done only through recognition of obvious inconsistencies (for example, a 15 year-old respondent who is recorded as having last worked in 1940). If a value is suspicious but reasonable, the erroneous value will find its way into the monthly statistics. For that reason emphasis must be placed on quality controls and interviewer training to ensure that errors are both minimal in number and non-systematic in nature.

Where errors or omissions are detected, the erroneous or missing items are replaced by the imputation of logically consistent values. Such changes are made automatically by the edit and imputation system or through intervention of experts. These changes are based on pre-specified criteria, and may involve the internal logic of the questionnaire, reference to earlier month's information (if available), or the use of similar records to impute one or more values. In all cases, editing changes are recorded and this information is used to assess various aspects of survey performance. These records of errors are also used to advise interviewers of mistakes made in the past in order to avoid repetition of these mistakes in the future.

6.3 Imputation

We can distinguish between three types of non-response to the LFS. Complete household non-response is when no response has been received for any member of a household. These households are dropped and accounted for in the weighting process. Partial household non-response is when at least one, but not all, members have responded to the survey. Finally, item non-response implies that some questionnaire answers have been received, but there are either missing responses to certain questions, or the interview was terminated part way through, and

responses are missing for the remainder of the questionnaire. Both partial household non-response and item non-response are resolved using imputation.

The LFS uses a hot deck imputation system, where responses from other complete questionnaires are used to fill in the missing data on incomplete records. To have the best possible match, the system finds a matching person based on socio-demographic and labour force information. This method assures that the selected donor record matches, as closely as possible, the characteristics of the incomplete record.

6.4 Creation of derived variables

A number of data items (variables) on the microdata file are derived by combining items on the questionnaire according to classification rules. For example, labour force status is derived from specific combinations of responses to a number of survey questions regarding work activity, status in employment, job search, availability, etc.

[Click here](#) to view the List of derived variables

6.5 Weighting

The principle behind estimation in a probability sample such as the LFS is that each person in the sample “represents”, besides himself or herself, several other persons not in the sample. For example, in a simple random 2% sample of the population, each person in the sample represents 50 persons in the population.

The weighting phase is a step which calculates, for each record, what this number is. This weight appears on the microdata file, and **must** be used to derive meaningful estimates from the survey. For example (* if the number of individuals enrolled full-time at a university during the reference week) is to be estimated, it is done by selecting the records referring to those (* households/individuals *) in the sample with that characteristic and summing the weights entered on those records.

Details of the method used to calculate these weights are presented in Section 8.2.

6.6 Suppression of confidential information

It should be noted that the “Public Use” microdata files described above differ in a number of important respects from the survey “master” files held by Statistics Canada. These differences are the result of actions taken to protect the anonymity of individual survey respondents. Users requiring access to information excluded from the microdata files may purchase custom tabulations. Estimates generated will be released to the user, subject to meeting the guidelines for analysis and release outlined in sections 8.4 and 8.6 of this document.

The Public Use Microdata File (PUMF) contains data for Canada, the provinces and the 3 largest census metropolitan areas – Toronto, Montreal and Vancouver. The survey master data file includes explicit geographic identifiers for province, economic region and census metropolitan area. It is also possible to obtain, where sample sizes permit, estimates by other areas. Please contact our Client Services unit by calling toll-free 1 866 873-8788, by fax: (613) 951-2869 or by e-mail: labour@statcan.ca

7.0 Data quality

LFS estimates, like those produced from any other sample survey, contain sampling errors and nonsampling errors. Accordingly, if estimates from this survey are to be interpreted correctly, knowledge of their quality is required. To ensure and monitor the quality of its data, the LFS has an extensive data quality program. A whole range of quality indicators are produced on a regular basis and carefully analysed. Among them, the response rate is one of the key indicators.

7.1 Response rates

The following table summarizes the response rates to the Labour Force Survey (LFS), by province.

Province	Sample size as of January 2005	LFS response rate*
		%
Newfoundland and Labrador	1,965	93
Prince Edward Island	1,400	93
Nova Scotia	3,005	93
New Brunswick	2,858	93
Quebec	9,956	92
Ontario	15,605	92
Manitoba	3,783	94
Saskatchewan	3,926	94
Alberta	5,368	93
British Columbia	5,783	93
Canada	53,648	93

Note: The LFS sample size is defined as the number of selected households.

*The LFS response rate is the number of LFS responding households as a percentage of the number of LFS selected households.

7.2 Survey errors

The estimates derived from this survey are based on a sample of households. Somewhat different estimates might have been obtained if a complete census had been taken using the same questionnaire, interviewers, supervisors, processing methods, etc. as those actually used in the survey. The difference between the estimates obtained from the sample and those resulting from a complete count taken under similar conditions, is called the sampling error of the estimate.

Errors which are not related to sampling may occur at almost every phase of a survey operation. Interviewers may misunderstand instructions, respondents may make errors in answering questions, the answers may be incorrectly entered on the questionnaire and errors may be introduced in the processing and tabulation of the data. These are all examples of non-sampling errors.

Over a large number of observations, randomly occurring errors will have little effect on estimates derived from the survey. However, errors occurring systematically will contribute to biases in the survey estimates. Considerable time and effort is made to reduce non-sampling errors in the survey. Quality assurance measures are implemented at each step of the data collection and processing cycle to monitor the quality of the data. These measures include the use of highly skilled interviewers, extensive training of interviewers with respect to the survey procedures and questionnaire, procedures to ensure that data capture errors are minimized and coding and edit quality checks to verify the processing logic.

8.0 Guidelines for tabulation, analysis and release

This chapter of the documentation outlines the guidelines to be adhered to by users tabulating, analyzing, publishing or otherwise releasing any data derived from survey microdata files. Note that the Labour Force Survey publishes both seasonally adjusted and unadjusted estimates. By following these guidelines, users should be able to produce the same figures as the unadjusted labour force estimates produced by Statistics Canada and, at the same time, will be able to develop currently unpublished figures in a manner consistent with these established guidelines.

8.1 Sample weighting guidelines for tabulation

If proper weights are not used, the estimates derived from the microdata files cannot be considered to be representative of the survey population, and will not correspond to those produced by Statistics Canada.

Users should also note that some software packages may not allow the generation of estimates that exactly match those available from Statistics Canada, because of their treatment of the weight field.

8.2 Weighting procedures for the Labour Force Survey

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. Each is described below.

Basic weight

In a probability sample, the sample design itself determines weights which must be used to produce unbiased estimates of population. Each record must be weighted by the inverse of the probability of selecting the person to whom the record refers. In the example of a 2% simple random sample, this probability would be 0.02 for each person and the records must be weighted by $1 / 0.02 = 50$. Due to the complex LFS design, dwellings in different regions will have different basic weights. Because all eligible individuals in a dwelling are interviewed (directly or by proxy), this probability is essentially the same as the probability with which the dwelling is selected.

Cluster sub-weight

The cluster delineation is such that the number of dwellings in the sample increases very slightly with moderate growth in the housing stock. Substantial growth can be tolerated in an isolated cluster before the additional sample represents a field collection problem. However, if growth takes place in more than one cluster in an interviewer assignment, the cumulative effect of all increases may create a workload problem. In clusters where substantial growth has taken place, sub-sampling is used as a means of keeping interviewer assignments manageable. The cluster sub-weight represents the inverse of this sub-sampling ratio in clusters where sub-sampling has occurred.

Stabilization weight

Sample stabilization is also used to address problems with sample size growth. Cluster sub-sampling addressed isolated growth in relatively small areas whereas sample stabilization accommodates the slow sample growth over time that is the result of a fixed sampling rate along with a general increase in the size of the population. Sample stabilization is the random dropping of dwellings from the sample in order to maintain the sample size at its desired level. The basic weight is adjusted by the ratio of the sample size, based on the fixed sampling rate, to the desired

sample size. This adjustment factor is known as the stabilization weight. The adjustment is done within stabilization areas defined as dwellings belonging to the same employment insurance economic region and the same rotation group.

Non-response

For certain types of non-response (i.e. household temporarily absent, refusal), data from a previous month's interview with the household if any, is brought forward and used as the current month's data for the household.

In other cases, non-response is compensated for by proportionally increasing the weights of responding households. The weight of each responding record is increased by the ratio of the number of households that should have been interviewed, divided by the number that were actually interviewed. This adjustment is done separately for non-response areas, which are defined by employment insurance economic region, type of area, and rotation group. It is based on the assumption that the households that have been interviewed represent the characteristics of those that should have been interviewed within a non-response area.

Labour Force Survey sub-weight

The product of the previously described weighting factors is called the LFS sub-weight. All members of the same sampled dwelling have the same sub-weight.

Sub-provincial and province-age-sex adjustments

The sub-weight can be used to derive a valid estimate of any characteristic for which information is collected by the LFS. However, these estimates will be based on a frame that contains some information that may be several years out of date and therefore not representative of the current population. Through the use of more up-to-date auxiliary information about the target population, the sample weights are adjusted to improve both the precision of the estimates and the sample's representation of the current population.

Independent estimates are available monthly for various age and sex groups by province. These are population projections based on the most recent census data, records of births and deaths, and estimates of migration. In the final step, this auxiliary information is used to transform the sub-weight into the final weight. This is done using a calibration method. This method ensures that the final weights it produces sum to the census projections for the auxiliary variables, namely totals for various age-sex groups, economic regions, census metropolitan areas, rotation groups, household and economic family size. Weights are also adjusted so that estimates of the previous month's industry and labour status estimates derived from the present month's sample, sum up to the corresponding estimates from the previous month's sample. This is called composite estimation. The entire adjustment is applied using the generalized regression technique.

This final weight is normally not used in the weighting for a supplement to the LFS. Instead, it is the sub-weight which is used, as explained in the following paragraphs.

8.3 Rounding guidelines

In order that estimates for publication or other release purposes derived from these microdata files correspond to those produced by Statistics Canada, users are urged to adhere to the following guidelines regarding the rounding of such estimates:

- a) Estimates in the main body of a statistical table are to be rounded to the nearest hundred units 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, if the last two digits are between 00 and 49, they are

changed to 00 and the preceding digit (the hundreds digit) is left unchanged. If the last digits are between 50 and 99 they are changed to 00 and the preceding digit is incremented by 1.

- b) Marginal sub-totals and totals in statistical tables are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units using normal rounding.
- c) For annual and three-month moving averages, unrounded estimates are added together and divided by the number of months in the average (either 12 or 3). Proportions, rates and percentages are to be computed from rounded components (i.e. rounded numerators and/or denominators) and then are to be rounded themselves to one decimal using normal rounding.
- d) Sums and differences of aggregates (or ratio) are to be derived from their corresponding rounded components to the nearest one decimal place.
- e) In instances where, due to technical or other limitations, a rounding technique other than normal rounding is used resulting in estimates to be published or otherwise released which differ from corresponding estimates published by Statistics Canada, users are urged to note the reason for such differences in the publication or release document(s).
- f) Under no circumstances are unrounded estimates to be published or otherwise released by users. Unrounded estimates imply greater precision than actually exists.

Aside: It will be difficult to produce identical seasonally adjusted estimates. These estimates depend on the seasonal adjustment methodology, the models used, the length of the series being adjusted and the raking of independently derived series. Raking refers to a procedure that ensures consistency between series. For example, employment for province level industry groups may not add up to the province level employment because they are seasonally adjusted separately. Raking refers to ratio adjusting the industry groups to ensure consistency with the province level total.

8.4 Guidelines for statistical analysis

The LFS is based upon a complex sample design, with stratification, multiple stages of selection, and unequal probabilities of selection of respondents. Using data from such complex surveys presents problems to analysts because the survey design and the selection probabilities affect the estimation and variance calculation procedures that should be used. In order for survey estimates and analyses to be free from bias, the survey weights must be used.

While many analysis procedures found in statistical packages allow weights to be used, the meaning or definition of the weight in these procedures may differ from that which is appropriate in a sample survey framework, with the result that while in many cases the estimates produced by the packages are correct, the variances that are calculated are poor. Approximate variances for simple estimates such as totals, proportions and ratios (for qualitative variables) can be derived using the accompanying Approximate Sampling Variability Tables.

8.5 Approximate sampling variability

Sampling variability

The Labour Force Survey collects information from a sample of households. Somewhat different figures might have been obtained if a complete census had been taken using the same questionnaires, interviewers, supervisors, processing methods, etc. as those actually used in the Labour Force Survey.

The sampling error, or standard error, is a measure that quantifies how different the sample estimate might be from the Census value. It is based on the idea of selecting several samples, although in a survey only one sample is drawn and information is collected on units in that sample. Using the same sampling plan, if a large number of samples were to be drawn from the same population, then about 68% of the samples would produce a sample estimate that is within one standard error of the census value and in about 95% of the samples it will be within two standard errors of the census value.

Sampling variability can also be expressed relative to the estimate itself. The standard error as a percentage of the estimate is called the coefficient of variation (CV) or the relative standard error. Probability statements can also be made about CVs; for example, if the CV is 7% then in 68% of the samples the census value will lie within 7% or one CV and in 95% of the samples the census value will lie within 14% or two times the CV of the estimate.

Indicators of sampling variability

Small CV's are desirable because they indicate that the sampling variability is small relative to the estimate. The CV depends on the size of the estimates, the sample size the estimate is based on, the distribution of the sample, and the use of post-censal population estimates in the estimation procedure. The size of the estimates is important because the CV is the sampling error expressed as a percentage of the estimate. The smaller the estimate the larger the CV (all other things being equal). For example, when the unemployment rate is high the CV may be small. If the unemployment rate falls due to improved economic conditions then the corresponding CV will become larger. Typically, of similar estimates, the one with largest sample size will yield the smaller CV. This is because the sampling error is smaller.

Also, estimates referring to characteristics that are more clustered will have a higher CV. For example, persons employed in forestry, fishing, mining, oil and gas in Canada are more clustered geographically than employed women aged 55 to 64 years in Ontario. The latter will have a smaller sampling variability although the estimates are of approximately the same size.

Finally estimates referring to age and sex are usually more reliable than other similar estimates because the LFS sample is calibrated to post-censal population estimates of various age and sex groupings. Continuing the previous example, persons employed part-time in Alberta will have a larger sampling variability than employed men aged 35 to 44 years in British Columbia although the estimates are of similar size.

Variability of monthly estimates

To look up an approximate measure of the CV of an estimate of a monthly total, please consult Table A, which gives the size of the estimate as a function of the geography and the CV. The rows give the geographic area of the estimate while the columns indicate the resulting level of accuracy in terms of the CV, given the size of the estimate. To determine the CV for an estimate of size X in an area A, look across the row for area A, find the estimate that is less than or equal to X. Then the title of the column will give the approximate CV. For example, to determine the sampling error for an estimate of 40.2 thousand unemployed in Newfoundland and Labrador in September 2004, we find the closest but smaller estimate of 27.1 thousand giving a CV of 5%.

Therefore, the estimate of 40,200 unemployed in Newfoundland and Labrador has a CV of roughly 5%.

Table A: CVs for estimates* of monthly totals for Canada and the provinces

Geographic Area	Coefficient of variation								
	1.0%	2.5%	5.0%	7.5%	10.0%	15.0%	20.0%	25.0%	30.0%
Canada	1,112.2	324.7	142.3	82.6	46.4	27.0	18.0	13.0	9.9
Newfoundland and Labrador	241.7	66.2	27.1	15.2	8.6	4.8	3.2	2.2	1.7
Prince Edward Island	59.7	18.0	7.9	4.6	2.7	1.6	1.1	0.8	0.6
Nova Scotia	278.4	79.8	33.7	19.3	11.1	6.4	4.2	3.1	2.3
New Brunswick	242.7	67.0	27.0	15.2	8.9	5.0	3.3	2.4	1.8
Quebec	1,173.9	327.3	135.9	77.0	43.8	24.8	16.3	11.7	8.9
Ontario	1,060.6	304.0	131.5	75.8	42.2	24.3	16.1	11.6	8.8
Manitoba	231.5	68.3	30.1	17.6	9.9	5.8	3.9	2.8	2.1
Saskatchewan	193.6	56.6	24.8	14.4	8.1	4.7	3.1	2.3	1.7
Alberta	530.2	158.6	72.2	42.6	23.5	13.8	9.3	6.8	5.2
British Columbia	761.8	222.8	97.7	56.8	31.9	18.6	12.4	9.0	6.8

* Estimates are in thousands and rounded to the nearest hundred.

Table A is supplied as a rough guide to the sampling variability. The sampling variability is modeled so that, given an estimate, approximately 75% of the CVs will be less than or equal to the CVs derived from the table. There will, however, be 25% of the CVs that will be somewhat higher than the ones given by the table.

The CV values given in Table A are derived from models based on 2001, 2002, 2003 and most of 2004 LFS sample data. It is important to bear in mind that these values are approximations. For more accurate measures of variability, please contact Client Services at 1 866 873-8788 or e-mail us at labour@statcan.ca.

Variability of annual estimates

To look up an approximate measure of the CV of an estimate of an annual average, please consult Table B, which gives the size of the estimate as a function of the geography and the CV. The rows give the geographic level of the estimate while the columns indicate the resulting level of accuracy in terms of the CV, given the size of the estimate. To determine the CV for an estimate of size X in an area A, look across the row for area A, find the estimate that is less than or equal to X. Then the title of the column will give the approximate CV. For example, to determine the sampling error for an annual average estimate of 39.8 thousand unemployed in Newfoundland and Labrador, we find the closest but smaller estimate of 23.5 thousand giving a CV of 2.5%. Therefore, the estimate of 39,800 unemployed in Newfoundland and Labrador has a CV of roughly 2.5%.

Table B: CVs for estimates* of annual averages for Canada and the provinces

Geographic Area	Coefficient of variation								
	1.0%	2.5%	5.0%	7.5%	10.0%	15.0%	20.0%	25.0%	30.0%
Canada	424.4	123.8	56.6	33.2	17.5	10.3	6.9	5.0	3.8
Newfoundland and Labrador	82.6	23.5	10.6	6.1	3.2	1.9	1.2	0.9	0.7
Prince Edward Island	20.1	6.2	3.0	1.8	1.0	0.6	0.4	0.3	0.2
Nova Scotia	99.5	29.1	13.4	7.9	4.2	2.4	1.6	1.2	0.9
New Brunswick	75.2	21.6	9.8	5.7	3.0	1.8	1.2	0.8	0.6
Quebec	307.3	92.3	44.8	27.0	13.7	8.2	5.6	4.1	3.1
Ontario	335.4	102.1	50.0	30.3	15.5	9.4	6.4	4.7	3.6
Manitoba	82.0	24.2	11.7	7.0	3.5	2.1	1.4	1.0	0.8
Saskatchewan	70.1	20.0	9.4	5.5	2.7	1.6	1.1	0.8	0.6
Alberta	205.1	60.5	29.1	17.4	8.7	5.2	3.5	2.6	1.9
British Columbia	246.6	74.7	36.4	22.0	11.2	6.8	4.6	3.4	2.6

* Estimates are in thousands and rounded to the nearest hundred.

Table B is supplied as a rough guide to the sampling variability. The sampling variability is modeled so that, given an estimate, approximately 75% of the CVs will be less than or equal to the CVs derived from the table. There will, however, be 25% of the CVs that will be somewhat higher than the ones given by the table.

The CV values given in Table B are derived from a model based on 2000, 2001, 2002, 2003 and most of 2004 LFS sample data. It is important to bear in mind that these values are approximations. For more accurate measures of variability, please contact Client Services at 1 866 873-8788 or e-mail us at labour@statcan.ca.

Variability of rates

Estimates that are rates and percentages are subject to sampling variability that is related to the variability of the numerator and the denominator of the ratio.

Unemployment rate

The unemployment rate is the ratio of X, the total number of unemployed in a group, to Y, which is the total number of participants in the labour force in the same group. Here the group may be a province or CMA and/or it may be an age-sex group. For example, in September 2004, there were approximately 40,200 unemployed persons in Newfoundland and Labrador and 253,700 participants in the labour force, giving an unemployment rate of 15.8%.

To determine the CV for the unemployment rate, the following formula can be used:

$$CV(X/Y) = \sqrt{[CV(X)]^2 - [CV(Y)]^2}$$

where, CV(X) is the CV for the total number of unemployed in a specific geographic or demographic subgroup and CV(Y) is the CV for the total number of participants in the labour force in the same subgroup. Continuing the example for Newfoundland and Labrador, the CV for the unemployment rate of 15.8% would be $\sqrt{5^2 - (1)^2} = 4.9\%$, where from Table A the CVs of monthly estimates of 40,200 and 253,700 in Newfoundland and Labrador are 5% and 1%, respectively.

Participation rate and employment rate

The participation rate represents the number of persons in the labour force expressed as a percentage of the total population size. The employment rate is the total number of employed divided by the total population size. For both the above rates, the numerator and the denominator represent the same geographic and demographic group.

For Canada, the provinces, CMAs and age-sex groups the LFS population estimates are not subject to sampling variability because they are calibrated to independent sources. Therefore, in the case of the participation rate and the employment rate of these geographic and demographic groups, the CV is equal to that of the contributing numerator.

Subgroups of Canada, the provinces and age-sex groups are called domains; for example, persons employed in agriculture in Manitoba are a domain. To determine the CV of rates in the case of domains, the variability of both the numerator and the denominator have to be taken into account because the denominator is no longer a controlled total and is subject to sampling variability. Therefore, for participation rates and employment rates of domains, the CV can be determined similar to the unemployment rate. The totals in the numerator and denominator for the relevant rate should reflect the same domain or subgroup.

Variability of estimate of change

The difference of estimates from two time periods gives an estimate of change that is also subject to sampling variability. An estimate of year-to-year or month-to-month change is based on two samples which may have some households in common. Hence, the CV of change depends on the CV of the estimates for both periods and the sample overlap, ρ , between the periods. The following formula can be used to approximate the CV of the estimate of change:

$$CV(Y_2 - Y_1) = \sqrt{1 - \rho} \frac{\sqrt{Y_1^2 CV(Y_1)^2 + Y_2^2 CV(Y_2)^2}}{(Y_2 - Y_1)}$$

where, Y_1 and Y_2 are the estimates for the two periods; and, the value of ρ is 0.5 for change between consecutive months and ρ is zero for changes over all other time periods. When comparing the annual averages of two years, the CV of the annual estimates (Table B) should be used. For month-to-month change, seasonally adjusted estimates should be used in conjunction with the CVs of the monthly estimates from Table A. Note that the above formula gives approximate estimates of the sampling variability associated to an estimate of change.

How to interpret the CV value

Household surveys within Statistics Canada generally use the following guidelines and reliability categories in interpreting CV values for data accuracy and in the dissemination of statistical information. These guidelines are provided for reference purposes only.

Category 1 - If the CV is $\leq 16.5\%$ - no release restrictions: data are of sufficient accuracy that no special warnings to users or other restrictions are required.

Category 2 - If the CV is $> 16.5\%$ and $\leq 33.3\%$ - release with caveats: data are potentially useful for some purposes but should be accompanied by a warning to users regarding their accuracy.

Category 3 - If the CV $> 33.3\%$ - not for release: data contain a level of error that makes them so potentially misleading that they should not be released in most circumstances. If users insist on inclusion of Category 3 data in a non-standard product, even after being advised of their accuracy, the data should be accompanied by a disclaimer. The user should acknowledge the

warnings given and undertake not to disseminate, present or report the data, directly or indirectly, without this disclaimer.

8.6 Release criteria for the Labour Force Survey

It is not possible to calculate exact measures of variability for each estimate released by the LFS. Since the size of the estimate is highly correlated with the degree of variability, the LFS suppresses estimates below the following levels:

	Minimum size for release (000s)
Canada	1.5
Newfoundland and Labrador	0.5
Prince Edward Island	0.2
Nova Scotia	0.5
New Brunswick	0.5
Quebec	1.5
Ontario	1.5
Manitoba	0.5
Saskatchewan	0.5
Alberta	1.5
British Columbia	1.5

8.7 Composite estimation

Composite estimation is a term that refers to any point estimator that is constructed by combining two or more other estimators. For surveys with a design like the LFS, it traditionally is used to refer to a weighted average of the present estimator and a “change based” estimator. The change based estimate updates last months estimate with an estimate of change derived from the current and previous month’s sample. A method of composite estimation has been developed that yields a single set of survey weights making the method operationally feasible for the LFS.

The method of composite estimation uses the rotation pattern of the LFS to improve estimates. Households selected for the LFS remain in the sample for six consecutive months. In any given month, about one sixth of households are interviewed for the first time, one sixth for the second time, and so on. Each group of households is referred to as a rotation group or a panel. In any two consecutive months, five sixths of the households form a common sample. Part of the apparent change in month to month estimates may arise because an old rotation group is replaced by one that has quite different characteristics. This contributes noise (or more precisely sampling error) to estimates of change and makes short-term trends difficult to identify.

For a set of predetermined variables, composite estimation exploits the rotation design by putting more emphasis on the change that comes from the common sample. By improving the change estimate while maintaining the quality of the level estimate, data reliability is enhanced. An important step in composite estimation is to select the predetermined, or key composite variables. As these are the variables which will exhibit the largest reduction in sampling error, and there is a limit to how many can be used, much thought was given at this stage. In terms of a strategy, two rationales were employed. The first was to select a set of variables that are deemed highly important and to which many other published series are summed. To meet this criteria, at the province level, total employed, unemployed, and not in the labour force for adult men and women and for young men and women (and implicitly the province level totals), were chosen. The second rationale was to choose variables that were of high interest but which were more prone to sampling error. This led to the selection of province level industry groupings and class of worker. All remaining data series are referred to as non-composite data series.

The main benefit for the key composite variables is a reduction in sampling error. There are modest gains for level estimates and substantial gains for estimates of change. This reduced error improves LFS time series because seasonal patterns are more easily detected. Also, the reduced volatility in data series means that series that are published on a three month moving average basis (province level industry series), can now be published on a monthly basis. Little or no impact on non-composite data series has been observed.

9.0 The Labour Force Survey questionnaire

The Labour Force Survey questionnaire is Statistics Canada's main vehicle for collecting monthly data about the working or job-seeking activities of the Canadian population 15 years of age or over.

The questionnaire is completed for every household member 15 years of age and over, excluding full-time members of the regular Armed Forces.

In subsequent interviews the Labour Force Survey will not be asked of persons 70 years of age and over, unless they have been a non-respondent for at least two consecutive months.

[Click here](#) to view the Labour Force Survey Questionnaire

9.1 Guidelines for using the LFS questionnaire

[Click here](#) to view the Guidelines for using the LFS questionnaire

10.0 Record layout

10.1 Public use

[Click here](#) to view the Public use record layout

[Click here](#) to view the List of derived variables

10.2 Confidential use

[Click here](#) to view the Confidential use record layout