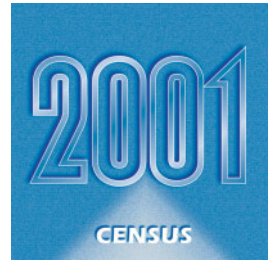




Catalogue No. 92-384-XIE

Mobility and Migration

2001 Census Technical Report



ELECTRONIC PUBLICATIONS AVAILABLE AT
www.statcan.ca



2001 Census Technical Report

Mobility and Migration

Table of Contents

	Page
INTRODUCTION.....	3
1. DATA COLLECTION AND COVERAGE	4
1.1 General.....	4
1.1.1 Collection Methods	4
1.1.2 Special Coverage Studies	4
1.2 Questionnaire and Instructions	4
2. DATA PROCESSING	7
2.1 General.....	7
2.1.1 Regional Processing.....	7
2.1.2 Imaging	8
2.1.3 Interactive Verification	8
2.1.4 Automated Coding	9
2.1.5 Edit and Imputation.....	10
2.1.5.1 General.....	10
2.1.5.2 Dwelling Classification Study (DCS)	11
2.1.5.3 Weighting.....	11
2.2 Mobility and Migration – Pre-processing.....	11
2.2.1 Coding of Mobility	11
2.2.2 Manual Coding.....	13
2.2.3 Places With the Same Name.....	13
2.2.4 Quality Control	13
2.3 Mobility and Migration – Processing	14
2.3.1 Stratification	14
2.3.2 Edit and Imputation.....	14
3. DATA QUALITY MEASUREMENT	16
3.1 General.....	16
3.2 Mobility 5 Years and 1 Year	16
3.2.1 Data Evaluation.....	16
3.2.2 Effect of Edit and Imputation.....	18
3.2.3 Evaluation and Comparison With Other Sources of Data	20
3.2.4 Comparison With Other Sources of Data	21
4. HISTORICAL COMPARABILITY	29
4.1 Comparability of Mobility Data With Those of Previous Censuses	29
4.2 Modifications Regarding Processing.....	31
5. CONCLUSION	32

Table of Contents – concluded

Page

APPENDICES

Appendix A. Glossary of Terms 33

Appendix B. 2001 Census Products and Services 36

BIBLIOGRAPHY **37**

Introduction

The 2001 Census required the participation of the entire population of Canada, i.e. some 30 million people distributed over a territory of 9 million square kilometres. An endeavour of this magnitude represented a tremendous challenge. Although there are high quality standards governing the collection and processing of the data, and in spite of efforts aimed at reducing non-response, for example through the use of communications, it is not possible to eliminate all errors. While this term does not necessarily imply any mistake as such, some element of error is bound to result in view of decisions to control census costs.

Statistics Canada is committed to explaining the methods and concepts used to collect and process its data and to providing users with information on the quality of the data produced, as well as other data characteristics which might limit their usefulness or interpretation. This report is aimed at informing users on the complexity of the data and on any difficulties that could affect their use. It explains the theoretical framework and the definitions used to gather the data, and describes unusual circumstances that could affect data quality. Moreover, the report touches upon data capture, edit and imputation, and deals with the historical comparability of the data.

The **2001 Census Technical Reports Series** includes 16 reports covering the variables of the 2001 Census of Population, as well as *Coverage* and *Sampling and Weighting*.

This report deals with mobility and migration. It has been prepared by the Demography Division, with the support of staff from the Census Operations Division and the Social Survey Methods Division.

Users will find additional information on census concepts, variables and geography in the *2001 Census Dictionary* (Catalogue No. 92-378-XIE), and an overview of the complete census process in the *2001 Census Handbook* (Catalogue No. 92-379-XIE).

1. Data Collection and Coverage

This stage of the census process ensures that each of the 11.8 million households in Canada is enumerated. The census enumerates the entire Canadian population, which consists of Canadian citizens (by birth and by naturalization), landed immigrants, and non-permanent residents, together with family members who live with them. Non-permanent residents are persons living in Canada who have a Minister's permit, a student or employment authorization, or who are claiming refugee status, and family members living with them.

The census also counts Canadian citizens and landed immigrants who are temporarily outside the country on Census Day. This includes federal and provincial government employees working outside Canada, Canadian embassy staff posted to other countries, members of the Canadian Armed Forces stationed abroad, and all Canadian crew members of merchant vessels. Because people outside the country are enumerated, the Census of Canada is considered a modified *de jure* census.

1.1 General

1.1.1. Collection Methods

To ensure the best possible coverage, the country is divided into small geographic areas called enumeration areas (EAs). Each census representative is responsible for at least one EA. The optimal number of households in an EA ranges from 175 in rural areas to 600 in urban areas. In the 2001 Census, there were 42,851 enumeration areas in Canada, and 38,000 people were engaged in collecting the data.

In 2001, approximately 98% of households were self-enumerated. Self-enumeration requires that a census representative drop off a questionnaire at each household during the two weeks before Census Day. An adult or responsible member of the household is asked to complete the questionnaire for all members of the household, and then mails the questionnaire in a pre-addressed envelope.

Approximately 2% of households were enumerated in the 2001 Census using the canvasser enumeration method. In this case, a census representative visits the household and completes a questionnaire for the household by interview. This method is normally used in remote and northern areas of the country, and on most Indian reserves. The canvasser enumeration method is also used in certain urban areas where it is considered highly possible that respondents would be unlikely to return a questionnaire.

1.1.2 Special Coverage Studies

Since 100% coverage is virtually impossible with such a large survey, a number of checks are performed on the collection of data. These studies measure the extent of coverage errors that occur when dwellings or individuals are missed, incorrectly included or double-counted. These checks are the Vacancy Check, the Reverse Record Check and the Overcoverage Study. These studies are discussed in the 2001 Census Technical Report on *Coverage* (Catalogue No. 92-394-XIE), planned for release in December 2004.

1.2 Questionnaire and Instructions

Mobility – Place of Residence 5 Years Ago

Refers to the relationship between a person's usual place of residence on Census Day and his or her usual place of residence five years earlier. A person is classified as a **non-mover** if no difference exists.

Otherwise, a person is classified as a **mover** and this categorization is called Mobility Status (5 Years Ago). Within the movers category, a further distinction is made between **non-migrants** and **migrants**; this difference is called **migration status**.

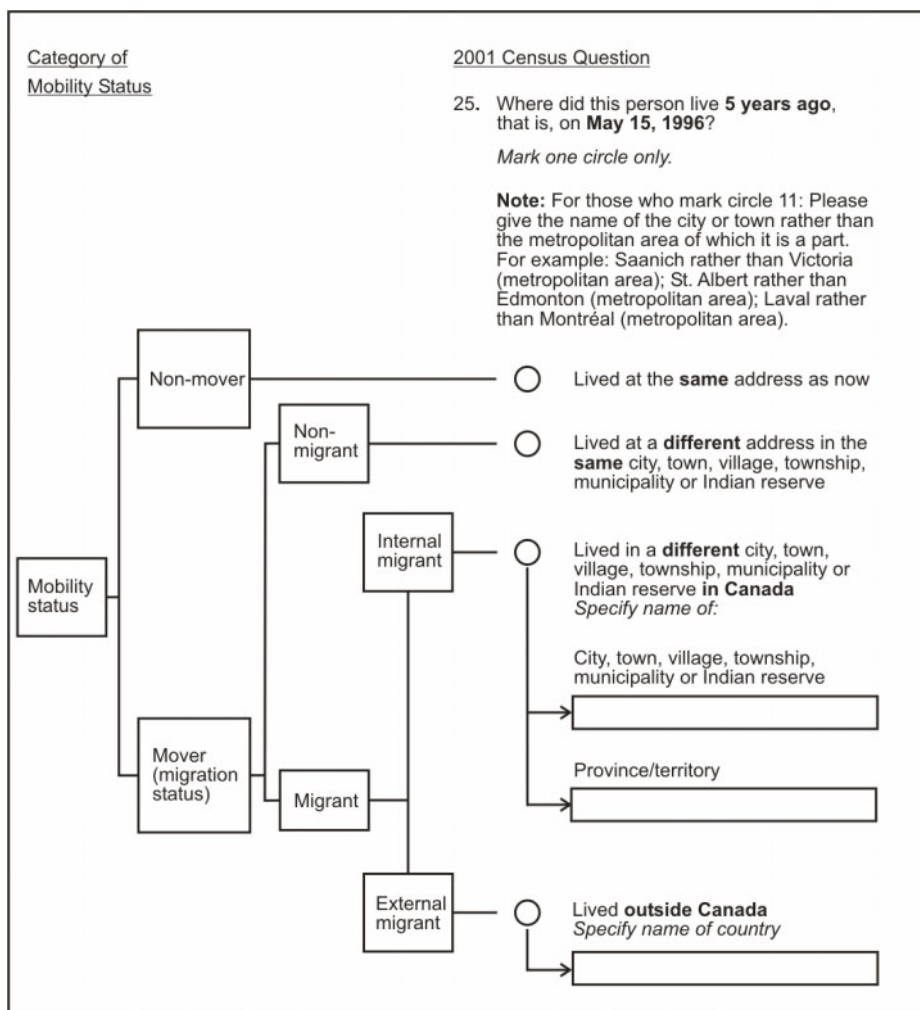
Non-movers are persons who, on Census Day, were living at the same address as the one at which they resided five years earlier.

Movers are persons who, on Census Day, were living at a different address than the one at which they resided five years earlier.

Non-migrants are movers who, on Census Day, were living at a different address, **but** in the same census subdivision (CSD) as the one they lived in five years earlier.

Migrants are movers who, on Census Day, were residing in a different CSD five years earlier (**internal migrants**) or who were living outside Canada five years earlier (**external migrants**).

Figure 1. Relationship Between the Category of Mobility Status and the 2001 Census Question on Place of Residence 5 Years Ago



Mobility – Place of Residence 1 Year Ago

Refers to the relationship between a person's usual place of residence on Census Day and his or her usual place of residence one year earlier. A person is classified as a **non-mover** if no difference exists. Otherwise, a person is classified as a **mover** and this categorization is called Mobility Status (1 Year Ago). Within the category of movers, a further distinction is made between **non-migrants** and **migrants**; this difference is called **migration status**.

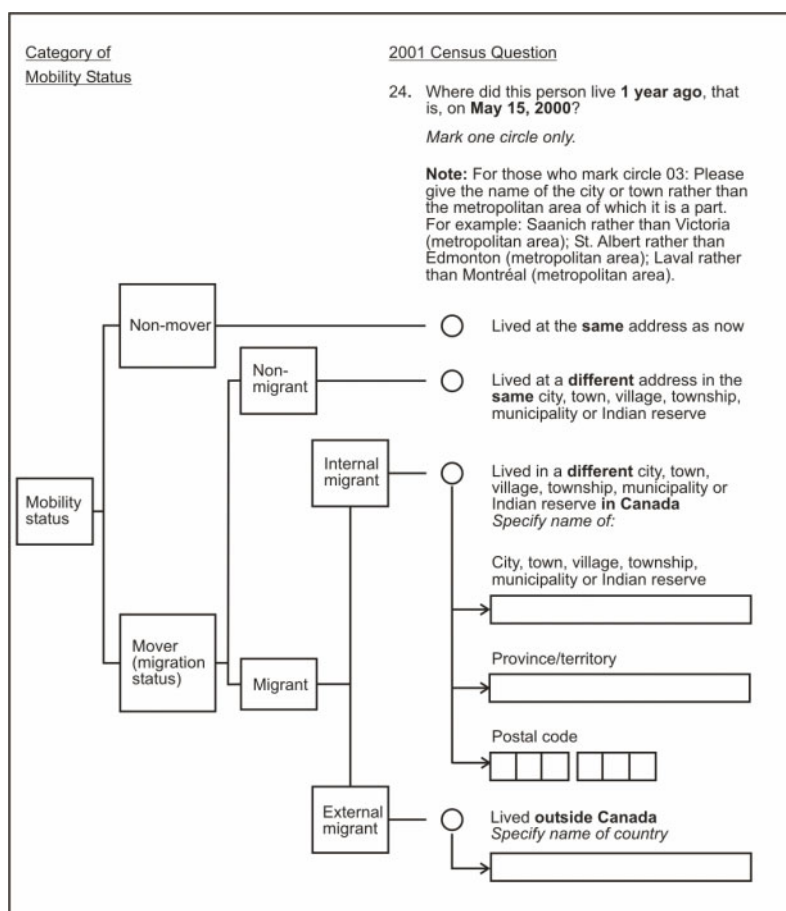
Non-movers are persons who, on Census Day, were living at the same address as the one at which they resided one year earlier.

Movers are persons who, on Census Day, were living at a different address than the one at which they resided one year earlier.

Non-migrants are movers who, on Census Day, were living at a different address, **but** in the same census subdivision (CSD) as the one they lived in one year earlier.

Migrants are movers who, on Census Day, were residing in a different CSD one year earlier (**internal migrants**) or who were living outside Canada one year earlier (**external migrants**).

Figure 2. Relationship Between the Category of Mobility Status and the 2001 Census Question on Place of Residence 1 Year Ago



2. Data Processing

This part of the census process involved the processing of all the completed questionnaires, from the data capture of the information through to the creation of an accurate and complete retrieval database. The final database was transferred to the Data Quality Measurement Project to determine the overall quality of the data, and to the Dissemination Project for the production and marketing of the 2001 Census products and services. A new objective for 2001 was to create an image retrieval system giving access to the images (pictures) of all the census questionnaires and visitation records, so that subsequent processes requiring access to original census forms would not have to handle the thousands of boxes and paper documents, as in previous censuses.

2.1 General

2.1.1 Regional Processing

Regional Processing was responsible for the manual coding of the industry and occupation responses and the data capture of the questionnaire information into a machine-readable format for subsequent processing systems. Given the enormous volume of census questionnaires and information to be captured (representing over 4 billion keystrokes), Regional Processing has been contracting this work out since 1981 to the Canada Customs and Revenue Agency (CCRA), formerly called Revenue Canada. By using the trained staff and infrastructure already in place at CCRA, the census realized cost savings by partnering with another government agency. For the 2001 Census, approximately 2,800 CCRA employees were sworn to secrecy under the *Statistics Act* to perform the census work, under the same rules and regulations as those which apply to the employees of Statistics Canada.

When the collection activities for a specific enumeration area (EA) were completed, the questionnaires, along with their maps and visitation records, were shipped in EA boxes from the field collection units to one of eight designated CCRA tax centres across the country.

The first step was to prepare the completed questionnaires for data capture. This traditionally included the manual assignment of codes to written answers that were provided by the respondents. For 2001, most of the written responses were converted to codes using automated systems (see Section 2.1.4). The only written responses that had to be manually coded for the 2001 Census were the questions on industry and occupation contained in the long-form questionnaires. Research into the automation of the coding of these questions has begun, and it is expected that an automated system will be operational for the 2006 Census.

The industry responses were coded at CCRA according to the North American Industry Classification System (NAICS), which was introduced as a standard within Statistics Canada a few years ago. NAICS is designed to provide a common framework for Canada, the United States and Mexico, which will enable the production of industry statistics under the North American Free Trade Agreement (NAFTA). This meant a change for industry coding - in 1996, industry was coded using the 1980 Standard Industrial Classification (SIC). In order to allow longitudinal comparisons, the 2001 industry question was also coded using the 1980 SIC during the Automated Coding phase (see Section 2.1.4). This phase was carried out with more automated means than in previous censuses.

Once the questionnaires were received and registered at one of the CCRA tax centres, and the industry and occupation codes assigned, the next step was to sort, label and batch the questionnaires in preparation for data capture. The labels affixed to each questionnaire contained a unique sequence number that was used to control the movement of the questionnaire throughout the CCRA operations. For the first time, the label also included a bar code to facilitate the scanning of the questionnaire in the imaging operation (see Section 2.1.2).

Data capture was then performed by traditional manual keying at mainly mainframe terminals. Verification of the accuracy of the data capture operation was done by selecting a sample of questionnaires that were already key-entered and capturing the information from the questionnaires in this sample a second time. Quality control statistics were produced by comparing the two sets of captured information.

As the data were keyed, they were transmitted in real time over dedicated communication lines to the CCRA computer in Ottawa. Within 24 hours, the data were then transferred to tape cartridges and transported by bonded carrier to Statistics Canada, where they were loaded into the mainframe computer. Questionnaires were reassembled into their EA boxes for shipment to Statistics Canada's 2001 processing site in Ottawa.

2.1.2 Imaging

In previous censuses, the remaining processing steps that required access to the questionnaires and visitation records used the paper documents. For 2001, the need to handle the paper was eliminated by imaging (scanning) all the questionnaires and visitation records as soon as they arrived at the 2001 processing site from the CCRA tax centres. Subsequent operations then had access to the questionnaires and visitation record images, using an image retrieval system, rather than using the paper documents.

As the EA boxes arrived at the 2001 processing site, they were registered. Then, the documents were prepared for imaging. Since the questionnaires and visitation records were in booklet format, they had to be cut into separate sheets in order to be run through the scanners. Following the cutting, since the 2A questionnaire was actually two booklets glued together (one English and the other French), the unused portion had to be separated from the completed portion. Extra material that was included with the questionnaires was removed (e.g., paper clips, notes). The questionnaires were then batched by EA for imaging.

The 13 million documents were imaged using 15 high-volume scanners running five days a week, two shifts per day. The geographic identifier that was required to identify each document image was automatically assigned using the bar code on the label affixed during the data-capture operations at CCRA (see Section 2.1.1). Quality control was performed to ensure that each document contained the right number of pages, and that the number of questionnaires by form type was correct for each EA. A problem-resolution operation resolved any problems that arose. The images were then written to optical platters for subsequent access and archiving. As the questionnaires were scanned, their images were also kept in magnetic storage for immediate access by the Interactive Verification activities (see Section 2.1.3).

The images on the optical platters are being kept in a secure location and are only accessible to authorized Statistics Canada employees from within the secure location.

2.1.3 Interactive Verification

The main objective of Interactive Verification was to identify and correct errors in the data, for which proper resolution required reference to the images of the questionnaires and/or visitation records. A detailed set of edits was applied to the captured data to identify possible errors, such as households with missing or duplicate persons, incorrect enumeration of foreign or temporary residents, questionnaires assigned to the wrong household, or misclassification of households as occupied or unoccupied. A thorough review of the information on all relevant census forms was conducted to determine the appropriate corrective action for each edit failure. In some cases, this required adding and/or deleting persons or dwellings; consequently, this process had an impact on the census counts.

As the census data arrived on cartridges from CCRA, they were loaded into Statistics Canada's computers, ready for the Interactive Verification activities. A series of automated "structural" edits were performed, mainly to verify the information filled out by the Census Representative on the front cover of the questionnaire. These edits included, among other things, matching questionnaire and household types, cross-checking the number of questionnaires and people enumerated, and verifying that the geographic identifiers were unique. Some edits were also performed on the income information, so that anomalies could be extracted and examined by income subject-matter experts.

All edits were done by EA. Errors were flagged, and then corrected by referring to the images of the questionnaires and visitation record for that EA. The corrections were made to the electronic data using an interactive PC-based system. Some of the corrections were also noted on the questionnaire images, using a process commonly called "annotation".

Once the EA edits were completed, automated and manual processes were used to verify the block number that the Census Representative had copied from the EA map onto the questionnaire and visitation record.

A National Block Program has been implemented for the first time in 2001. A "block" is basically the smallest area bounded by streets or roads, lakes and rivers. In urban centres, "blocks" are generally recognizable city blocks. In rural areas, "blocks" are much larger areas, but are still bounded by identifiable features, with no significant feature splitting an area. These blocks are added together to create the EAs for data collection purposes, and the dissemination areas (DAs) for the dissemination of census products and services.

During the field collection operations, as census representatives delivered a questionnaire to each dwelling within their EA, they wrote the person's name (if possible) and the address in their visitation records (VRs). At the same time, they copied the VR line number from the VR onto the questionnaire, to uniquely identify the questionnaire for that dwelling. As well, they identified the block number for the dwelling from their EA map and copied the number into the VR and onto the questionnaire. These block numbers were data-captured, so that all the dwellings in Canada could be identified as belonging to a particular block.

As a final step in the Interactive Verification process, the data were reformatted and forwarded for the final processing steps, namely Automated Coding and Edit and Imputation.

Interactive Verification also performed some special processing to ensure that Canadians living outside Canada on Census Day (people aboard coast guard and Canadian Armed Forces vessels, Canadian-registered merchant vessels, and diplomatic and military personnel) were enumerated.

2.1.4 Automated Coding

Automated coding matched the write-in responses that were "data-captured" from the long-form questionnaires during Regional Processing (see Section 2.1.1) to entries in an automated reference file/classification structure containing a series of words or phrases and corresponding numerical codes. Although a large percentage of write-in responses can be coded in a purely automated manner, a series of responses always remains unmatched. Specially trained coders and subject-matter experts reviewed all unmatched responses and, with the assistance of PC-based interactive coding systems, assigned the appropriate numerical code after examining responses to other questions and from other members of the household. Automated coding was applied to write-in responses for the following questions on the long form (2B):

- relationship to Person 1;
- home language;
- non-official languages;

- first language learned in childhood (mother tongue);
- language of work (new in 2001);
- place of birth;
- place of birth of parents (new in 2001);
- citizenship;
- ethnic origin (ancestry);
- population group;
- Indian Band/First Nation;
- place of residence 1 year ago;
- place of residence 5 years ago;
- major field of study;
- religion (last asked in 1991);
- place of work;
- industry according to the 1980 SIC (first time for automated coding in 2001).

As the responses for a particular variable were coded, the data for that variable were sent to the Edit and Imputation phase.

2.1.5 Edit and Imputation

2.1.5.1 General

The data collected in any survey or census contain omissions or inconsistencies. These errors can be the result of respondents answering the questions incorrectly or incompletely, or they can be due to errors generated during processing. For example, a respondent may be reluctant to answer a question, may fail to remember the right answer or may misunderstand the question. Census staff may code responses incorrectly or may make other mistakes during processing.

Prior to Edit and Imputation, the questionnaires underwent some basic manual edits during collection. Field staff reviewed the questionnaires for missing responses or unacceptable multiple responses. Such problems were resolved by contacting the respondents and obtaining the required information. Following collection, Interactive Verification (see Section 2.1.3) performed some basic structural edits, where the images of the questionnaires and visitation records were referenced as necessary.

The final clean-up of the data was done in Edit and Imputation and was, for the most part, fully automated. It applied a series of detailed edit rules that identified any missing or inconsistent responses. These missing or inconsistent responses were corrected most of the time by changing the values of a few variables as possible through imputation. Imputation invoked "deterministic" and/or "minimum-change "hot deck"" methods. For deterministic imputation, errors were corrected by inferring the appropriate response value from responses to other questions. For minimum-change "hot deck" imputation, a record with a number of characteristics in common with the record in error was selected. Data from this "donor" record were borrowed and used to change the minimum number of variables necessary to resolve all the edit failures.

Two different automated systems were used to carry out this processing.

The **Nearest-neighbour Imputation Method (NIM)**, developed for the 1996 Census to perform Edit and Imputation for basic demographic characteristics such as age, sex, marital status, common-law status and relationship to Person 1, was expanded for 2001 and implemented in a system called **CANCEIS (CANadian Census Edit and Imputation System)** to include Edit and Imputation for such variables as labour, place of work, mode of transportation and mobility. As in 1996, CANCEIS continued to allow more extensive and exact edits to be applied to the response data, while preserving responses through minimum-change "hot deck" imputation.

SPIDER (**S**ystem for **P**rocessing Instructions from **D**irectly **E**ntered **R**equirements) was used to process the remaining census variables, such as mother tongue, dwelling and income. This tool translated subject-matter requirements, identified through decision logic tables, into computer-executable modules. SPIDER performed both deterministic and "hot deck" imputation.

2.1.5.2 Dwelling Classification Study (DCS)

The Dwelling Classification Study takes a sample of dwellings declared either unoccupied or absent during the collection process. Later, the DCS returns to these dwellings to determine if, on Census Day, they were occupied, unoccupied or should not have been listed because they did not meet the definition of a census dwelling. If a dwelling was occupied, one of two separate adjustments is made to the census database. If the dwelling was listed as vacant in the census, then a technique, called "random additions", was applied to add households and persons to the census database. In the 2001 Census, 111,626 households and 222,720 persons were added to the database to account for the estimated number of persons living in vacant dwellings. The second adjustment was concerned with absent households. These were adjusted by creating a new household size for all such dwellings on the census database. A total of 143,684 households with 317,587 persons were added to the census database through this adjustment.

2.1.5.3 Weighting

Data on age, sex, marital status, common-law status, mother tongue and relationship to Person 1 were collected from all Canadians. However, the bulk of the information gathered in the census came from the 20% sampling of the population. Weighting, applied to the respondent data after Edit and Imputation, was used to adjust the census sample to represent the whole population.

The weighting method produced fully representative estimates from the sample data. For the 2001 Census, weighting employed a methodology known as calibration (or regression) estimation. Calibration estimation started with initial weights of approximately 5 and then adjusted them by the smallest possible amount needed to ensure closer agreement between the sample estimates (e.g., number of males, number of people aged 15 to 19) and the actual population counts for age, sex, marital status, common-law status and household size.

Once invalid and non-response data were corrected, they were transferred to the final national retrieval databases for subsequent data quality studies and dissemination.

2.2 Mobility and Migration – Pre-processing

2.2.1 Coding of Mobility

The mobility questions contain four fields requiring automated coding. The four variables to be coded are as follows:

1. place of residence 1 year ago within Canada
2. place of residence 1 year ago outside Canada
3. place of residence 5 years ago within Canada
4. place of residence 5 years ago outside Canada

The coding of within-Canada mobility variables involves converting place names into seven-digit Standard Geographical Classification (SGC)¹ codes. The SGC is the official classification used at Statistics Canada for three types of geographic areas: provinces and territories, census divisions (CDs) and census subdivisions (CSDs). It provides unique numeric codes for those areas, which are hierarchically related.

The coding of outside-Canada mobility variables involves converting place names into three-digit codes. In contrast to the SGC, the digital codes used to represent countries are not part of an official structure. Each code stands for a country. The country codes used were identical to those used to code the "Place of Birth" variable.

Write-in responses are batch-coded or manually coded. Batch coding involves finding an exact match in detailed reference files. Unmatched responses are sent to manual coding, where coders assign a code to each response.

In all, more than 1.3 million write-in responses were coded in this operation. As shown in Table 1 below, the match rate was 66.7% for responses on place of residence 1 year ago in Canada and 68.7% for place of residence 5 years ago. Hence, manual coding was required for one third of all within-Canada write-ins.

Table 1. Automated Coding of Migration Variables, 2001 Census

Variables	System-coded Responses	Manually Coded Responses	Total	System Match Rate	System Error Rate (%)
5 years ago (within Canada)	589,581	268,066	857,647	68.7%	
5 years ago (outside Canada)	158,170	13,235	171,405	92.3%	
5 years ago - Total	747,751	281,301	1,029,052	72.7%	
1 year ago (within Canada)	204,771	102,462	307,233	66.7%	
1 year ago (outside Canada)	49,380	4,604	53,984	91.5%	
1 year ago - Total	254,151	107,066	361,217	70.4%	
Total, mobility variables	1,001,902	388,367	1,390,269	72.1%	
Total, sociocultural variables	30,590,059	1,768,630	32,358,689	94.5%	
Total, autcoded variables	32,811,182	3,175,112	35,986,294	91.2%	
Total for all variables (within Canada)					1.07
Total for all variables (outside Canada)					1.13

Source: Automated coding project progress report of November 21, 2001.

For the coding of outside-Canada responses, the match rate was 91.5% for "1 year ago" and 92.3% for 5 years ago". In other words, less than 10% of responses were referred to manual coding.

1 2001 Census Dictionary (Catalogue No. 92-378-XPE), page 198, or the Internet version of the same publication (Catalogue No. 92-378-XIE): see the "Geographic Classifications: Standard Geographical Classification (SGT)" section in the chapter entitled "Geography".

The difference between the within-Canada and outside-Canada match rates can be attributed to a number of factors. Because of the wide variety of official and unofficial place names in Canada and the volume of responses, the match rate for within-Canada variables is always lower than the match rate for outside-Canada variables.

The system's combined error rate was 1.07% for within-Canada variables and 1.13% for outside-Canada variables.

2.2.2 Manual Coding

Various types of responses are not matched in batch coding. There are two reasons for failing to find a match:

- the reference file is incomplete;
- there are spelling errors in the response.

The reference file may be considered incomplete in cases where the respondent entered an unofficial but commonly used name, such as "*la petite bourgogne*" (Montréal), "Mechanicsville" (Ottawa), or "Wrightville" (Hull). The file may also be incomplete in cases where write-in responses are not captured as was expected when the reference file was developed. For example, if the keying rules require the operator to abbreviate the name of the province and the rule is not followed, responses with an unabbreviated province name will not be matched because the reference file was based on the keying rules.

Another reason for non-matching is spelling errors. Because automated coding requires an exact match between the write-in and the reference file, the slightest difference will result in non-matching. The same thing happens when abbreviations are used. Many respondents use abbreviations to identify the place where they were living one year or five years ago. Among the most common abbreviations are HFX for Halifax, KIT for Kitchener and PTBO for Peterborough.

2.2.3 Places With the Same Name

Because of the wide range of place names (official and unofficial), a number of localities have the same name. To address this problem, temporary codes or pseudo-codes are created for such cases. Each record with a temporary code is reassigned to a unique SGC code for the "edit and imputation" step. Such codes are based on the place name's frequency and the population of each unit in the previous census.

The 2001 reference file contained nearly 1,000 pseudo-codes to code same-name localities. Among the place names that occur most frequently are Langley, B.C.; Cornwall, Ont.; Magog, Qc; North Vancouver, B.C.; and Cambridge, Ont.

2.2.4 Quality Control

Coding quality is checked by a quality control module, which measures the error rates of the system and the manual coders. This process involved recoding response samples taken from batches of previously coded phrases. Each sample was recoded by a different coder from the one who originally coded the batch. Any discrepancies between the two codes assigned were reviewed by an expert coder. The error was attributed to the coder who was wrong. This helped to assess the performance of coders throughout the production process. The error rates of manual coders were under 2% in 2001, comparable to the rates observed in 1996.

2.3 Mobility and Migration – Processing

2.3.1 Stratification

The purpose of stratification is to divide the population into strata or subgroups with the same type of migration behaviour. The processing units are divided into distinct groups, on the basis of characteristics that differ from one family to another. Those characteristics are referred to as the stratification criteria, and each group of families formed in this way constitutes a stratum. The stratification criteria are defined by the restrictions imposed on the choice of donors, by the nature of the edit rules used, and by the size of the processing units.

The hierarchy presented below is followed by all modules that use CANCEIS.

Six strata were created:

- a person living in a collective dwelling (non-institutional), or a non-census family person in a private household, or a parent with one or more children under age 15 in a private household;
- a couple consisting of two adults living together with no children in a private household, or two parents having only children under age 15 in a private household;
- one parent living alone with just one child aged 15 and over in a private household, with or without children under age 15;
- two parents living together with just one child aged 15 and over in a private household, with or without children under age 15;
- one parent living with at least two children aged 15 and over in a private household, with or without children under age 15;
- two parents living with at least two children aged 15 and over in a private household, with or without children under age 15.

2.3.2 Edit and Imputation

Two systems were used for edit and imputation. The System for Processing Instructions from Directly Entered Requirements (SPIDER) was used for the pre-derivation and post-derivation modules. The Canadian Census Edit and Imputation System (CANCEIS) was used for the first time for hot-deck imputation of mark-in and write-in responses.

In the 2001 Census, processing was divided into 11 modules: three pre-processing modules, four imputation modules, two intermediate processing modules and two post-processing modules. The new feature of the processing system is that mark-in responses and write-in responses are processed separately in hot-deck imputation.

The first two pre-processing modules create the derived variables for mark-in responses. The pseudo-codes are processed with a subroutine that assigns a unique SGC code using proportions based on the population of each unit in the 1996 Census. A number of invalid response patterns are processed by deterministic imputation. Only cases in which the outcome of the imputation is clear are imputed – for example, a case where a respondent marks no boxes and writes in a place of residence 1 year or 5 years ago that is different from his place of residence on Census Day. In such cases, the status “was living in another city” will be assigned to the respondent, and the write-in place name will be used as the place of residence 1 year or 5 years ago. The third pre-derivation module performs stratification.

Next comes the first CANCEIS module, which performs edit and hot-deck imputation of mark-in responses. A mobility status is assigned to each respondent processed. Both variables (1 year and 5 years) are imputed simultaneously, and the module is submitted independently for each stratum. The processing units are families, which means that imputation is carried out for all persons with a missing response in the same family at the same time. The following variables are used for matching in the search for donors: mother tongue, age, sex, marital status, common-law status, census subdivision of residence on Census Day, and 1-year and 5-year migration status.

Once the mark-in responses have been resolved, two intermediate processing modules are submitted. The first one ensures that there are no inconsistencies between the mark-in responses assigned in the previous step and the write-in responses in the mobility questions. For example, records whose final mobility status is “non-mover” must not have valid codes for city or country of residence. The second intermediate processing module creates family movement structure variables. Two types of variables are created. The first compares a person’s responses for place of residence 1 year and 5 years ago. The second compares one person’s responses with the responses of another person in the same family for the same period. Each individual is assigned a value for each of the five family movement variables:

1. 1-year mobility between persons based on comparison with parent #1;
2. 1-year mobility between persons based on comparison with parent #2;
3. 5-year mobility between persons based on comparison with parent #1;
4. 5-year mobility between persons based on comparison with parent #2;
5. person mobility based on comparison between 1-year and 5-year mobility status.

For example, the first family movement variable tells us whether the person was living with parent #1 five years ago or whether it was impossible to determine.

The second CANCEIS module performs edit and imputation on the family movement structure indicators. In cases where it was impossible to determine whether one person was living with another at a particular time, the indicators have to be imputed. For example, if one parent is classified as “migrant” but did not indicate his/her place of residence five years ago, it will be impossible to determine whether the other family members were living with him/her five years ago.

Only indicators with an undetermined value are imputed, since the purpose of this module is to complete the process of assigning codes to the indicators which began in MBD5. On exiting this module, all indicators will have a valid value, which will be used to impute write-ins (census subdivision of residence 1 year and 5 years ago, and country of residence 1 year and 5 years ago).

The third CANCEIS module simultaneously imputes 1-year and 5-year within Canada mobility write-ins for persons in families with mobility status “was living in another census subdivision”. An SGC code is assigned to records that have “migrant” status but still do not have a census subdivision of residence 1 year or 5 years ago. The processing units are families, which means that imputation is carried out at the same time for all persons in the same family for whom a correction is required.

The fourth CANCEIS module simultaneously imputes 1-year and 5-year country-of-residence write-ins for persons in families that have “external migrant” status, i.e. that were living outside Canada. Once again, the processing units are families.

The last two modules (post-processing) carry out deterministic imputation for three types of persons:

1. Children under age 15: They were assigned the response of one parent because they did not answer the question in the census.
2. Children aged 15 and over: Some of them were not processed in hot-deck imputation because the family was too large. They were assigned the response of one parent.
3. Persons under age 15 who were not members of a census family: They were assigned the response of the household’s Person 1.

3. Data Quality Measurement

3.1 General

Throughout the census-taking process, every effort was made to ensure high-quality results. Rigorous quality standards were set for data collection and processing, and the Public Communications Program assisted in minimizing non-response. A Data Quality Measurement Program was established to provide users with information on the quality and limitations of census data.

Although considerable effort is made throughout the entire process to ensure high standards of data quality, the resulting data are subject to a certain degree of inaccuracy. To assess the usefulness of census data for their purposes and to understand the risk involved in drawing conclusions or making decisions on the basis of these data, users should be aware of their inaccuracies and appreciate their origin and composition.

Within the **2001 Census Technical Reports Series**, users will find detailed 2001 Census information on *Coverage* and *Sampling and Weighting*. These two reports are scheduled to be released in November and December 2004 respectively.

3.2 Mobility 5 Years and 1 Year

3.2.1 Data Evaluation

This section of the report covers data quality indicators such as non-response and invalid response rates. The effect of edit and imputation and the certification of data will also be discussed.

Non-response and invalid responses:

Non-response means that there were no marks in any of the boxes and no entries in any of the write-in areas. The term "invalid response" refers to:

- (a) a response in which more than one box was marked; or
- (b) an inconsistent response, in which the marked box conflicts with the write-in response; or
- (c) a census subdivision or country that cannot be coded.

For the question on place of residence 5 years ago, the non-response rate was higher in the 2001 Census than in the 1996 Census. It jumped from 0.7% in 1996 to 1.6% in 2001. In contrast, the invalid response rate was almost identical in the two censuses: 1.4% in 1996 and 1.3% in 2001. As shown in Table 2, the pattern was the same in all four processing regions.

Table 2. Rates of Non-response and of Invalid Responses for the Question on Place of Residence 5 Years Ago, Based on Unweighted Data, for Canada and the Regions, 1996 and 2001 Censuses

Question	East		Quebec		Ontario		West		CANADA	
	1996	2001	1996	2001	1996	2001	1996	2001	1996	2001
Non-response and invalid responses										
25 = Mobility 5 years ago	1.7	2.3	2.5	3.0	2.1	3.2	1.7	2.8	2.0	2.9
25A = --- City and province	1.3	0.8	1.5	1.1	1.1	0.9	1.1	0.8	1.2	0.9
25B = --- Country	1.1	1.2	1.2	0.8	1.2	0.9	0.9	0.7	1.1	0.8
Non-response										
25 = Mobility 5 years ago	0.6	1.1	0.7	1.5	0.6	1.7	0.7	1.8	0.7	1.6
25A = --- City and province	1.3	0.8	1.5	1.0	1.1	0.8	1.1	0.8	1.2	0.9
25B = --- Country	0.8	1.2	1.2	0.8	1.1	0.9	0.8	0.7	1.0	0.8
Invalid responses										
25 = Mobility 5 years ago	1.1	1.1	1.8	1.5	1.5	1.4	1.0	1.1	1.4	1.3
25A = --- City and province	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.1
25B = --- Country	0.2	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.0

Source: *Clustered Non-response Study*, Social Survey Methods Division, June 2002 (Dave O'Grady).

For the question on place of residence 1 year ago, the non-response rate rose from 0.5% in 1996 to 1.3% in 2001. The invalid response rate, on the other hand, declined somewhat, from 1.2% in 1996 to 0.9% in 2001. As shown in Table 3, the pattern was the same in all four regions of Canada.

Table 3. Rates of Non-response and of Invalid Responses for the Question on Place of Residence 1 Year Ago, Based on Unweighted Data, for Canada and the Regions, 1996 and 2001 Censuses

Question	East		Quebec		Ontario		West		CANADA	
	1996	2001	1996	2001	1996	2001	1996	2001	1996	2001
Non-response and invalid responses										
24 = Mobility 1 year ago	1.3	1.7	2.2	2.2	1.6	2.4	1.3	2.3	1.6	2.3
24A = ---City and province	1.2	1.1	1.1	0.8	1.1	0.9	1.0	0.9	1.1	0.9
24B = ---Country	1.8	1.0	1.3	0.8	1.1	0.9	1.2	0.5	1.2	0.8
Non-response										
24 = Mobility 1 year ago	0.3	0.8	0.5	1.2	0.4	1.4	0.5	1.5	0.5	1.3
24A = ---City and province	1.2	1.1	1.1	0.8	1.1	0.8	0.9	0.8	1.0	0.8
24B = ---Country	1.7	0.8	1.2	0.7	0.9	0.8	1.0	0.5	1.0	0.7
Invalid responses										
24 = Mobility 1 year ago	0.9	0.9	1.7	1.0	1.2	0.9	0.8	0.8	1.2	0.9
24A = ---City and province	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
24B = ---Country	0.1	0.2	0.2	0.0	0.2	0.1	0.2	0.1	0.2	0.1

Source: *Clustered Non-response Study*, Social Survey Methods Division, June 2002 (Dave O'Grady).

These non-response and invalid response rates are not affected by data processing. They are computed before the data are processed by the edit and imputation system.

3.2.2 Effect of Edit and Imputation

Non-response and invalid response cases were processed in one of two ways: deterministic imputation or CANCEIS.

Table 4 provides a comparison of the proportions of imputed records by province and imputation method for the 5-year mobility question in the 1996 and 2001 Censuses. Overall, the imputation rate for the population aged 5 and over was higher in the 2001 Census. In the 1996 Census, 12.47% of records were imputed, compared with 20.23% in the 2001 Census. The sharp increase in imputation rates between 1996 and 2001 was due to the way in which CANCEIS operates. In 1996, records were imputed only once, whereas CANCEIS imputes erroneous records in several steps. It imputes the mark-in responses, the family movement indicators and the write-in responses separately. Consequently, a record may be imputed more than once.

Table 4. Proportion of Records Undergoing Imputation by Method, for Canada, Provinces and Territories, 1996 and 2001 Censuses

Province/territory	2001 Census		
	Deterministic Imputation	CANCEIS	Total
CANADA	3.90	16.32	20.23
Newfoundland and Labrador	2.32	14.11	16.43
Prince Edward Island	3.43	15.28	18.71
Nova Scotia	3.05	14.80	17.85
New Brunswick	3.00	14.43	17.43
Quebec	3.86	15.30	19.16
Ontario	4.36	16.08	20.44
Manitoba	2.39	18.31	20.69
Saskatchewan	2.45	19.37	21.83
Alberta	3.99	17.58	21.56
British Columbia	4.33	16.26	20.58
Yukon	3.13	22.19	25.32
Northwest Territories	4.02	24.96	28.98
Nunavut	0.59	30.7	31.29

Source: 2001 Census of Canada, unpublished tables.

Province/territory	1996 Census			Total
	Deterministic Imputation	Family-based Imputation	Probabilistic ("HotDeck") Imputation	
CANADA	3.79	7.34	1.33	12.47
Newfoundland and Labrador	2.11	8.12	0.80	11.03
Prince Edward Island	3.09	7.09	1.35	11.53
Nova Scotia	3.21	6.89	1.18	11.28
New Brunswick	2.60	6.64	1.29	10.52
Quebec	3.82	6.79	1.31	11.93
Ontario	4.10	6.47	1.27	11.84
Manitoba	2.25	10.28	1.12	13.66
Saskatchewan	2.16	11.35	1.16	14.68
Alberta	3.27	8.61	1.29	13.17
British Columbia	5.09	6.86	1.71	13.66
Yukon	5.11	8.98	4.48	18.56
Northwest Territories	2.61	16.84	2.18	21.64
Nunavut	-	-	-	-

Source: 1996 Census of Canada, unpublished tables.

Imputation has a greater effect on the under-15 population, as the mobility questions are to be answered only by persons aged 15 and over. Since the results are published for the 5-and-over population, the responses for persons aged 5 to 14 must be imputed. In the 2001 Census, 70% of imputations were on the under-15 population.

3.2.3 Evaluation and Comparison With Other Sources of Data

Prior to release, the mobility data were evaluated for quality. Two types of evaluations were performed. First, the data were evaluated from a trend's perspective. They were compared with the data from previous censuses. Second, they were compared with data from other sources.

Overall, migration has been declining since the 1991 Census. The proportion of movers in the population has been falling since 1991. According to Table 5B, it dropped from 23.6% in 1991 to 20.3% in 1996 and 19.5% in 2001. The proportion of non-movers has been rising since 1991. It increased from 53.3% in 1991 to 56.7% in 1996 and 58.1% in 2001 (see Table 5A).

Table 5A. Proportion of Movers 5 Years and Over, from the Question on Mobility Status (Place of Residence) 5 Years Ago, for Canada, Provinces and Territories, 1981 to 2001 Censuses

Province/territory	Movers				
	1981	1986	1991	1996	2001
Newfoundland and Labrador	32.0	29.1	28.4	26.6	27.0
Prince Edward Island	35.5	32.4	34.5	32.0	32.2
Nova Scotia	39.4	37.1	38.0	36.0	36.0
New Brunswick	39.0	34.0	34.4	32.6	32.8
Quebec	44.8	41.0	43.6	39.9	38.9
Ontario	46.9	44.5	48.0	43.1	42.8
Manitoba	45.2	42.9	43.3	39.9	38.8
Saskatchewan	44.4	41.9	39.0	38.6	38.4
Alberta	60.0	51.3	52.8	49.7	49.2
British Columbia	56.6	49.3	56.2	54.5	46.3
Yukon	70.1	63.6	64.5	58.3	48.6
Northwest Territories	65.1	62.1	65.5	64.1	54.2
Nunavut	-	-	-	-	55.7
CANADA	47.6	43.7	46.7	43.3	41.9

Source: Censuses of Canada, 1981 to 2001, unpublished tables.

Table 5B. Proportion of Migrants 5 Years and Over, from the Question on Mobility Status (Place of Residence) 5 Years Ago, for Canada, Provinces and Territories, 1981 to 2001 Censuses

Province/territory	Migrants				
	1981	1986	1991	1996	2001
Newfoundland and Labrador	14.0	12.3	13.8	12.1	12.3
Prince Edward Island	18.4	16.8	19.4	15.6	15.2
Nova Scotia	17.7	16.4	17.9	14.8	13.9
New Brunswick	17.3	14.4	16.3	14.3	14.8
Quebec	19.4	17.4	21.7	18.5	19.1
Ontario	21.8	20.2	25.3	20.4	19.6
Manitoba	17.6	15.5	15.3	13.8	14.1
Saskatchewan	22.1	19.2	17.7	17.5	17.3
Alberta	34.2	23.2	23.7	21.2	22.8
British Columbia	31.2	23.9	31.0	29.2	23.0
Yukon	40.9	33.2	35.2	28.9	21.5
Northwest Territories	33.5	30.2	30.1	25.7	24.9
Nunavut	-	-	-	-	19.4
CANADA	22.8	19.5	23.6	20.3	19.5

Source: Censuses of Canada, 1981 to 2001, unpublished tables.

There was some variation at the provincial level. Only two provinces bucked the national trend. Alberta posted a significant increase in its proportion of migrants, from 21.2% in 1996 to 22.8% in 2001. That was the largest provincial increase.

British Columbia experienced a substantial decline in migration. The proportion of migrants dropped from 29.2% in 1996 to 23.0% in 2001. The decrease was due to a rise in the proportion of non-movers, from 45.5% in 1996 to 53.7% in 2001.

The percentage of people living outside Canada five years ago has remained steady since the 1991 Census, at both national and provincial levels.

3.2.4 Comparison With Other Sources of Data

The interprovincial migration data were compared with data from two other sources: interprovincial migration estimates based on tax files from the Canada Customs and Revenue Agency, and estimates based on data from the Canada Child Tax Benefit (CCTB) program.

Table 6 shows the number of in-migrants and out-migrants and the net migration for each data source for the 1996 and 2001 Censuses for the 5-year mobility question. According to 2001 Census data, there were 905,670 interprovincial migrants in Canada for the 1996-2001 period. This is higher than the figure for the 1991-1996 period (820,275) but lower than the figure for 1986-1991 (977,075).

Table 6. Interprovincial Migration Estimates According to Different Sources, for Provinces and Territories, 1991-1996 and 1996-2001

Province/territory	CCTB(*)			Tax Files			Census			Difference Between Net Migration Balances		
	In-migrants	Out-migrants	Net Migration (1)	In-migrants	Out-migrants	Net Migration (2)	In-migrants	Out-migrants	Net Migration (3)	(1-2)	(3-1)	(3-2)
1991-1996												
Newfoundland and Labrador	51,039	77,217	-26,178	36,331	60,980	-24,649	16,227	39,465	-23,238	-1,529	2,940	1,411
Prince Edward Island	16,150	14,456	1,694	13,554	11,247	2,307	8,949	7,484	1,465	-613	-229	-842
Nova Scotia	99,323	106,273	-6,950	81,771	88,042	-6,271	47,453	53,905	-6,452	-679	498	-181
New Brunswick	73,683	78,160	-4,477	57,307	60,818	-3,511	34,058	36,025	-1,967	-966	2,510	1,544
Quebec	144,161	214,064	-69,903	119,190	169,432	-50,242	68,897	106,337	-37,440	-19,661	32,463	12,802
Ontario	429,727	449,198	-19,471	335,655	375,458	-39,803	194,022	241,031	-47,009	20,332	-27,538	-7,206
Manitoba	102,511	125,620	-23,109	77,907	101,591	-23,684	43,215	62,592	-19,377	575	3,732	4,307
Saskatchewan	110,810	137,127	-26,317	85,845	111,078	-25,233	47,521	67,301	-19,780	-1,084	6,537	5,453
Alberta	327,048	328,711	-1,663	271,225	265,577	5,648	162,641	159,059	3,582	-7,311	5,245	-2,066
British Columbia	434,288	257,765	176,523	369,726	202,715	167,011	252,625	102,675	149,950	9,512	-26,573	-17,061
Yukon	12,958	10,849	2,109	10,288	10,357	-69	5,954	5,290	664	2,178	-1,445	733
Northwest Territories	19,145	21,403	-2,258	16,032	17,536	-1,504	8,711	9,109	-398	-754	1,860	1,106
Total	1,820,843	1,820,843	0	1,474,831	1,474,831	0	890,273	890,273	0	0	0	0

Province/territory	CCTB(*)			Tax Files			Census			Difference Between Net Migration Balances		
	In-migrants	Out-migrants	Net Migration (1)	In-migrants	Out-migrants	Net Migration (2)	In-migrants	Out-migrants	Net Migration (3)	(1-2)	(3-1)	(3-2)
1996-2001												
Newfoundland and Labrador	53,936	86,331	-32,395	38 248	70,323	-32,075	16,065	47,100	-31,035	-320	1,360	1,040
Prince Edward Island	15,335	15,026	309	13 072	12,890	182	7,900	7,760	140	127	-169	-42
Nova Scotia	91,735	96,818	-5,083	79 296	85,659	-6,363	53,000	54,290	-1,290	1,280	3,793	5,073
New Brunswick	68,460	73,915	-5,455	54 429	62,841	-8,412	32,630	41,060	-8,430	2,957	-2,975	-18
Quebec	127,975	205,861	-77,886	102 786	171,833	-69,047	62,435	119,745	-57,310	-8,839	20,576	11,737
Ontario	427,537	371,038	56,499	371 434	302,528	68,906	242,495	190,610	51,885	-12,407	-4,614	-17,021
Manitoba	90,582	105,462	-14,880	69 472	90,513	-21,041	42,600	61,185	-18,585	6,161	-3,705	2,456
Saskatchewan	104,789	129,053	-24,264	78 209	103,633	-25,424	42,605	67,530	-24,925	1,160	-661	499
Alberta	413,834	273,728	140,106	362 422	224,729	137,693	242,230	122,805	119,425	2,413	-20,681	-18,268
British Columbia	291,705	322,593	-30,888	241 059	278,588	-37,529	151,715	175,335	-23,620	6,641	7,268	13,909
Yukon	8,816	12,637	-3,821	7 133	10,221	-3,088	3,735	6,490	-2,755	-733	1,066	333
Northwest Territories	14,861	17,473	-2,612	11 487	14,865	-3,378	5,740	8,910	-3,170	766	-558	208
Nunavut	3,606	3,236	370	5 011	5,435	-424	2,520	2,845	-325	794	-695	99
Total	1,713,171	1,713,171	0	1 434 058	1,434,058	0	905,670	905,665	5	0	5	5

* Canada Child Tax Benefit (CCTB) program.

The differences between census data on net migration and interprovincial migration estimates based on tax data are smaller than the differences between census data and the estimates based on CCTB data. The estimates based on tax records measure migration on an annual basis, while the estimates based on CCTB data do so on a monthly basis. The CCTB data are always higher because they capture cases in which a person moves more than once in the same year.

The data for the 1-year mobility question show that the number of interprovincial migrants has been declining since the 1991 Census. It fell from 319,205 in 1991 to 293,345 in 1996 and 284,645 in 2001. As in the case of 5-year mobility, the differences between the various data sources indicate that the estimates based on tax data are closer to the census figures. According to the data in Table 7, the differences between the estimates based on tax records and the census figures are smaller than the differences between the estimates based on CCTB data and the census data.

Table 7. Interprovincial Migration Estimates According to Different Sources, for Provinces and Territories, 1990-1991, 1995-1996 and 2000-2001

Province/territory	CCTB(*)			Tax Files			Census			Difference Between Net Migration Balances		
	In-migrants	Out-migrants	Net Migration (1)	In-migrants	Out-migrants	Net Migration (2)	In-migrants	Out-migrants	Net Migration (3)	(1-2)	(3-1)	(3-2)
1990-91												
Newfoundland and Labrador	10,666	14,099	-3,433	10,278	11,208	-930	10,429	10,595	-166	-2,503	3,267	764
Prince Edward Island	3,327	4,241	-914	2,809	3,134	-325	2,527	2,922	-395	-589	519	-70
Nova Scotia	22,190	22,797	-607	18,319	18,481	-162	18,290	18,128	162	-445	769	324
New Brunswick	16,101	16,731	-630	13,744	12,821	923	12,950	12,764	186	-1,553	816	-737
Quebec	31,386	43,727	-12,341	25,220	36,545	-11,325	27,579	35,376	-7,797	-1,016	4,544	3,528
Ontario	91,749	101,601	-9,852	72,923	84,550	-11,627	67,943	89,813	-21,870	1,775	-12,018	-10,243
Manitoba	20,305	28,476	-8,171	16,441	23,981	-7,540	16,722	21,320	-4,598	-631	3,573	2,942
Saskatchewan	21,270	33,584	-12,314	16,317	28,493	-12,176	17,318	27,225	-9,907	-138	2,407	2,269
Alberta	75,828	68,046	7,782	62,505	53,522	8,983	63,014	55,536	7,478	-1,201	-304	-1,505
British Columbia	88,330	47,790	40,540	72,417	38,364	34,053	77,042	39,666	37,376	6,487	-3,164	3,323
Yukon	2,574	2,206	368	2,145	1,902	243	1,971	1,878	93	125	-275	-150
Northwest Territories	4,214	4,642	-428	3,449	3,566	-117	3,419	3,981	-562	-311	-134	-445
Total	387,940	387,940	0	316,567	316,567	0	319,204	319,204	0	0	0	0

Province/territory	CCTB(*)			Tax Files			Census			Difference Between Net Migration Balances		
	In-migrants	Out-migrants	Net Migration (1)	In-migrants	Out-migrants	Net Migration (2)	In-migrants	Out-migrants	Net Migration (3)	(1-2)	(3-1)	(3-2)
1995-96												
Newfoundland and Labrador	10,665	18,238	-7,573	7,005	14,441	-7,436	7,288	13,296	-6,008	-137	1,565	1,428
Prince Edward Island	3,482	2,602	880	2,882	2,244	638	3,438	2,300	1,138	242	258	500
Nova Scotia	20,301	21,194	-893	16,263	17,508	-1,245	17,157	17,173	-16	352	877	1,229
New Brunswick	14,607	14,835	-228	11,770	12,139	-369	12,711	12,404	307	141	535	676
Quebec	28,090	41,307	-13,217	22,556	35,182	-12,626	25,408	33,844	-8,436	-591	4,781	4,190
Ontario	84,390	89,969	-5,579	69,059	71,881	-2,822	67,080	73,782	-6,702	-2,757	-1,123	-3,880
Manitoba	20,811	22,757	-1,946	15,075	18,641	-3,566	15,617	19,093	-3,476	1,620	-1,530	90
Saskatchewan	23,476	24,219	-743	17,411	19,572	-2,161	18,255	18,834	-579	1,418	164	1,582
Alberta	66,931	61,192	5,739	57,037	49,381	7,656	54,691	50,480	4,211	-1,917	-1,528	-3,445
British Columbia	81,141	57,681	23,460	66,959	44,934	22,025	66,208	46,921	19,287	1,435	-4,173	-2,738
Yukon	2,874	2,073	801	2,203	1,639	564	2,361	1,559	802	237	1	238
Northwest Territories	3,850	4,551	-701	3,065	3,723	-658	3,130	3,658	-528	-43	173	130
Total	360,618	360,618	0	291,285	291,285	0	293,344	293,344	0	0	0	0

Province/territory	CCTB(*)			Tax Files			Census			Difference Between Net Migration Balances		
	In-migrants	Out-migrants	Net Migration (1)	In-migrants	Out-migrants	Net Migration (2)	In-migrants	Out-migrants	Net Migration (3)	(1-2)	(3-1)	(3-2)
2000-2001												
Newfoundland and Labrador	10,611	14,152	-3,541	7,499	11,992	-4,493	8,355	11,080	-2,725	952	816	1,768
Prince Edward Island	2,887	2,816	71	2,567	2,402	165	3,015	2,705	310	-94	239	145
Nova Scotia	18,332	19,156	-824	15,313	17,390	-2,077	16,300	17,135	-835	1,253	-11	1,242
New Brunswick	13,716	13,797	-81	10,539	12,069	-1,530	11,415	12,390	-975	1,449	-894	555
Quebec	24,834	36,616	-11,782	21,341	30,783	-9,442	23,750	32,370	-8,620	-2,340	3,162	822
Ontario	85,761	67,884	17,877	74,516	55,893	18,623	75,295	59,640	15,655	-746	-2,222	-2,968
Manitoba	17,562	20,656	-3,094	12,623	16,946	-4,323	13,715	18,175	-4,460	1,229	-1,366	-137
Saskatchewan	18,497	28,950	-10,453	12,985	21,395	-8,410	13,265	22,740	-9,475	-2,043	978	-1,065
Alberta	79,371	53,623	25,748	64,129	43,672	20,457	67,965	48,485	19,480	5,291	-6,268	-977
British Columbia	50,019	62,708	-12,689	43,338	51,624	-8,286	47,030	54,200	-7,170	-4,403	5,519	1,116
Yukon	1,465	2,311	-846	1,153	1,725	-572	1,110	1,960	-850	-274	-4	-278
Northwest Territories	2,569	3,175	-606	2,177	2,337	-160	2,245	2,805	-560	-446	46	-400
Nunavut	1,495	1,275	220	1,040	992	48	1,170	960	210	172	-10	162
Total	327,119	327,119	0	269,220	269,220	0	284,630	284,645	-15	0	-15	-15

* Canada Child Tax Benefit (CCTB) program.

When we compare the 1-year mobility data from the 1996 and 2001 Censuses with the tax data for the same period, we find that the difference in the number of interprovincial migrants in Canada was only about 2,000 in the 1996 Census and about 15,000 in the 2001 Census. When we compare net migration by province, however, we find that the difference between the two estimation methods was smaller in the 2001 Census than in the 1996 Census.

4. Historical Comparability

4.1 Comparability of Mobility Data With Those of Previous Censuses

The following is a brief summary of the historical comparability of census mobility data, from the “Place of residence five years ago” question. More detailed information (including references to the 1941 and 1946 Censuses) is available in two user guides and in one technical report: *A User’s Guide to the 1976 Census Data on Mobility Status*, uncatalogued working paper, May 1980; *User’s Guide to 1986 Census Data on Mobility*, November 1990; and *Mobility and Migration, 1991 Census Technical Report*, (Catalogue No. 92-326), all of which are available through Statistics Canada.

A. Conceptual Changes

The mobility status question on place of residence five years ago has not differed significantly from the five-year questions of previous censuses. Therefore, the mobility data are generally comparable from 1961 on. The question has been based on a five-year reference interval and the census subdivision (CSD) has been used as the migration-defining unit. While the five-year census mobility data are generally comparable from 1961 to 2001, there are some conceptual differences users should be aware of.

- Since 1991, the term “address” has replaced the term “dwelling”, which had been used in all previous censuses since 1961. The current term “address” is used in the context of address of usual residence, not mailing address.
- From 1976 on, the primary classification of the population was made on the basis of mobility status (movers, non-movers) while, in some of the earlier censuses, the primary classification was based on migration status (migrants, non-migrants).
- There are also changes in related factors, such as question content, which users should be aware of when analysing mobility data.

Factors Affecting Conceptual Comparability

A number of factors affect historical data comparability of mobility in relation to the conceptual framework. Some of the areas in which changes have occurred are: coverage and universe, question content and structure, and geographic framework.

1. Changes in Coverage and Universe

From 1961 on, the universe for mobility status has included the population 5 years of age and over, with exclusions, which have varied from census to census.

- (a) In 1961, mobility status was reported for the population 5 years of age and over who are members of private households, excluding residents in collective dwellings, temporary residents, overseas military and government personnel and their families, and persons located after the regular census through postal check or re-enumeration. In 1971 and 1976, the universes of population 5 years of age and over excluded Canadian residents stationed abroad in the Armed Forces or in diplomatic services.
- (b) From 1981 on, the mobility universe comprises the population 5 years of age and over residing in Canada, excluding institutional residents and Canadian military and government personnel and their families posted abroad, in households outside Canada. This is in contrast to the 1971 and 1976 data that did include institutional residents.

2. Changes in Question Content and Structure

- (a) From 1961 to 1986, the previous country of residence was not collected for respondents indicating a place of residence outside Canada five years earlier. From 1991 on, respondents who indicated that they had lived outside Canada five years ago were asked to provide the name of the country.
- (b) From 1971 on, internal migrants were asked to specify only the name of their census subdivision of residence 5 years ago, whereas in previous censuses migrants were also asked whether or not their earlier residence was a farm.
- (c) A question on the number of intermunicipal moves was asked only in 1971.
- (d) Since 1986, emphasis was placed on ensuring that Indian reserves were accurately reported in mobility categories. From 1986 on, the answer categories refer to “city, town, village, township, other municipality or Indian reserve” compared to “city, town, village, borough or municipality” in 1981 and “city, town, village, municipality” in 1971 and 1976.
- (e) Instructions in the question referring to write-ins of place names were the same between 1971 and 1976, but they were expanded in 1981 to include examples. The 1981 instruction was repeated in 1986. In 1991, the instruction was revised with new wording and examples. In 1996, the wording and examples in the instruction were again slightly modified. For 2001, they were again slightly modified.
- (f) In 1991, revisions were made to both the structure and wording of the “Place of residence five years ago” question. In addition to the rewording of instructions and the replacement of the term “dwelling” with “address”, a filter question was introduced to serve as a screen for movers and non-movers. As well, answer categories were reworded and shortened. With these revisions, the basic content is still the same as the five-year questions of previous censuses; in general, historical comparability is retained.
- (g) In 1996, although the filter question was eliminated, the concept remained the same.
- (h) In 1996, the write-in box for county was eliminated. However, at the time of automated coding, the code for county was provided.

3. Changes in Geographic Framework

- (a) The comparability of the mobility data over the censuses has been affected by both conceptual changes in geography (such as the definitions of rural area [RA], urban area [UA], farm, non-farm, census metropolitan area [CMA] and census agglomeration [CA]) and the changes in census subdivision (CSD), census division (CD), census metropolitan area (CMA) and census agglomeration (CA) boundaries. Because the number of census geographic areas (CSDs, CMAs, etc.) and their boundaries change from census to census, the user must exercise caution when using mobility data over two or more censuses. For example, in 1986 there were 6,009 CSDs, 114 CAs and 25 CMAs compared to 5,710 CSDs, 88 CAs and 24 CMAs in 1981. The number of CSDs in 2001 is much smaller, reduced to 5,593, because of the amalgamation of cities. The changing number and boundaries of CSDs from one census to another will, to some extent, affect the comparability of the measure of “migrants” across censuses (since the volume of migrants is partly a function of the number and size of CSDs). Details of changes affecting the historical comparability of census geography from 1961 to 2001, as well as definitions and descriptions of available maps, are covered in a variety of census products.

- (b) Because of changes in geographic areas between censuses, places of residence five years ago must reflect boundaries of the census in question in order to obtain geographic consistency between current and previous place of residence. For example, when tabulating 1996 data on usual place of residence five years ago by current place of residence, all areas reflect 1996 boundaries, even when referred to as places of residence in 1991.

B. Collection and Processing Changes

The changes over censuses associated with each of the collection and processing stages have not significantly affected the comparability of mobility and migration data. However, there are some changes in processing that the user should be aware of when analysing mobility data.

- In 1991, autocoding (computerized coding) was introduced for converting write-ins of place names in the mobility question to the Standard Geographical Classification (SGC) codes. In previous censuses, write-ins were coded manually. Some of the manual coding procedures used in 1986 for resolving duplicate place names (e.g. Kingston township vs Kingston city, both in Ontario, but only “Kingston” reported) were automated during Edit and Imputation (E & I) as an extension of the autocoding system. The increased accuracy obtained with autocoding is expected to improve the quality of mobility data on out-migrants from CSDs, compared with previous censuses. The evaluation of the autocoding on the data quality has not yet been done.
- A significant change in E & I from earlier censuses occurred in 1981. Prior to 1981, non-response (partial/total) to the question on previous place of residence was reported as “Not stated”. However, for 1981, this “Not stated” category was dropped. Non-response to the question on previous place of residence was changed to a specific response via a combination of deterministic, family and “hot deck” imputation assignments. This imputation was achieved using the SPIDER program, which was introduced in 1981. For the 2001 Census, the Imputation portion of the E & I processing is done using the CANadian Census Edit and Imputation System (CANCEIS). The Edit portion is still being done using the SPIDER program.

C. Place of Residence 1 Year Ago

The question on the place of residence 1 year ago was asked in the 1991 Census for the first time. At this time, the migration-defining boundary was the province or the territory. From the 1996 Census on, the migration-defining boundary was changed to the census subdivision as in the case of the question on the place of residence 5 years ago.

4.2 Modifications Regarding Processing

The low non-response rate suggests that the introduction of a new imputation system (CANCEIS) had no effect on the comparability of data for the two mobility questions.

Similarly, the various waves of municipal mergers had no impact on comparability, since the geographic structure used was the one in effect on Census Day. The SGC codes for the old cities were converted to the codes for the new, post-merger cities.

5. Conclusion

This technical report has presented a variety of information about the 2001 Census mobility and migration data. The information provides users with an overview of data collection, processing and evaluation.

Appendix A. Glossary of Terms

The definitions of census terms, variables and concepts are presented here as they appear in the *2001 Census Dictionary* (Catalogue No. 92-378-XIE). Users should refer to the *2001 Census Dictionary* for full definitions and additional remarks related to any concepts, such as information on direct and derived variables and their respective universe.

Mobility Variables Available for Retrieval, 2001 Census

A1. Data from the Question on the Place of Residence 5 Years Ago

1. **Mob5:** Mobility – Place of Residence 5 Years Ago

Refers to the relationship between a person's usual place of residence on Census Day and his or her usual place of residence five years ago. On the basis of this relationship, the population is classified as **non-movers** and **movers** (mobility status). Within the category **movers**, a further distinction is made between **non-migrants** and **migrants** (migration status).

2. **PR5:** Province or Territory of Residence 5 Years Ago

Refers to the person's usual province or territory of residence on May 15, 1996, five years prior to Census Day.

3. **PR:** Province or Territory of Current Residence

Refers to the province or territory of Canada where the enumerated person lived on Census Day (May 15, 2001)

4. **PCD5:** Census Division of Residence 5 Years Ago

Refers to the person's usual census division (CD) of residence on May 15, 1996, five years prior to Census Day.

5. **PCD:** Census Division of Current Residence

Refers to the person's usual census division (DR) where the person's current residence was located on May 15, 2001.

6. **PCSD5:** Census Subdivision of Residence 5 Years Ago

Refers to the person's usual municipality or census subdivision (CSD) of residence on May 15, 1996, five years prior to Census Day.

7. **PCSD:** Census Subdivision of Current Residence

Refers to the person's municipality or census subdivision (CSD) where the person's current residence was located on May 15, 2001.

8. **CMA5:** Census Metropolitan Area or Census Agglomeration of Residence 5 Years Ago

Refers to the census metropolitan area (CMA), census agglomeration (CA) or non-CMA/CA where the person usually resided on May 15, 1996, five years prior to Census Day.

9. **CMA:** Census Metropolitan Area or Census Agglomeration of Current Residence

Refers to the census metropolitan area (CMA), census agglomeration (CA) or non-CMA/CA where the person's current residence was located on May 15, 2001.

10. **POP5:** Population Size of Census Subdivision of Residence 5 Years Ago

Refers to the current population of the municipality or census subdivision (CSD) where the person usually resided on May 15, 1996, five years prior to Census Day.

11. **POP:** Population Size of Current Census Subdivision of Residence

Refers to the 2001 population of the municipality or census subdivision (CSD) where the person usually resided on Census Day.

12. **CO5:** Country of Residence 5 Years Ago

Refers to the person's usual country of residence on May 15, 1996, five years prior to Census Day.

13. **RUUB5:** Rural/Urban Classification of Place of Residence 5 Years Ago

Refers to the rural or urban classification of the municipality or census subdivision (CSD) where the person usually resided on May 15, 1996, five years prior to Census Day.

14. **CSDTYPE5:** Census Subdivision Type of Residence 5 Years Ago

Refers to the census subdivision (CSD) type classification of the CSD (Indian reserve, village, town, township, city or municipality) where the person usually resided on May 15, 1996, five years prior to Census Day.

A2. Data from the Question on the Place of Residence 1 Year Ago

15. **Mob1:** Mobility – Place of Residence 1 Year Ago

Refers to the relationship between a person's usual place of residence on Census Day and his or her usual place of residence one year ago. On the basis of this relationship, the population is classified as **non-movers** and **movers** (mobility status). Within the category **movers**, a further distinction is made between **non-migrants** and **migrants** (migration status).

16. **PR1:** Province or Territory of Residence 1 Year Ago

Refers to the person's usual province or territory of residence on May 15, 2000, one year prior to Census Day.

17. **PCD1:** Census Division of Residence 1 Year Ago

Refers to the person's usual census division (CD) of residence on May 15, 2000, one year prior to Census Day.

18. **PCSD1:** Census Subdivision of Residence 1 Year Ago

Refers to the person's usual municipality or census subdivision (CSD) of residence on May 15, 2000, one year prior to Census Day.

19. **CMA1:** Census Metropolitan Area or Census Agglomeration of Residence 1 Year Ago

Refers to the census metropolitan area (CMA), census agglomeration (CA) or non-CMA/CA where the person usually resided on May 15, 2000, one year prior to Census Day.

20. **POP1:** Population Size of Census Subdivision of Residence 1 Year Ago

Refers to the current population of the municipality or census subdivision (CSD) where the person usually resided on May 15, 2000, one year prior to Census Day.

21. **CO1:** Country of Residence 1 Year Ago

Refers to the person's usual country of residence on May 15, 2000, one year prior to Census Day.

22. **RUUB1:** Rural/Urban Classification of Place of Residence 1 Year Ago

Refers to the rural or urban classification of the municipality or census subdivision (CSD) where the person usually resided on May 15, 2000, one year prior to Census Day.

23. **CSDTYPE1:** Census Subdivision Type of Residence 1 Year Ago

Refers to the census subdivision (CSD) type classification of the CSD (Indian reserve, village, town, township, city or municipality) where the person usually resided on May 15, 2000, one year prior to Census Day.

Appendix B. 2001 Census Products and Services

The census is a reliable source for describing the characteristics of Canada's people and dwellings. The range of products and services derived from census information is designed to produce statistics that will be useful, understandable and accessible to all users. Sources, such as the *2001 Census Catalogue*, the Statistics Canada Web site (<http://www.statcan.ca>) and, specifically, the On-Line Catalogue, contain detailed information about the full range of 2001 Census products and services.

There are several new product and service features for the 2001 Census:

1. Media

- The Internet is the preferred medium for disseminating standard data products and reference products.
- More census data are available to the public free of charge via the Internet.

2. Content

- Data tables for the 2001 Census are released by topics, that is, groups of variables on related subjects.
- Wherever possible, the language and vocabulary used in 2001 Census products available on the Internet is simplified to make the information accessible to more people.
- Users are offered various methods of searching and navigating through census standard products (including reference products) on the Internet.

3. Geography

- Geographic units such as dissemination areas, urban areas, designated places and metropolitan influenced zones were added to the standard products line. Some new units, such as dissemination areas, replace others.

4. Variables

- Information on the following new subjects was collected in the 2001 Census: birthplace of parents, other languages spoken at home and language of work. The 2001 questionnaire also included the question on religion, which is asked in every decennial census. The family structure variable was broadened to include same-sex couples.

Bibliography

Janes, D. *CANCEIS User Guide*. Social Survey Methods Division. Statistics Canada, unpublished document, June 2002.

Lachance, M. *Processus de Contrôle et d'Imputation pour les variables de la Mobilité*. Social Survey Methods Division. Statistics Canada, unpublished document, August 2001.

Shin, Y. E. *2001 Census. Retrieval Data Base Certification, Mobility and Migration Variables, Place of Residence 1 and 5 Years Ago*. Statistics Canada, unpublished document, December 2002.

Statistics Canada. *2001 Census Dictionary*. Ottawa: Industry Canada, 2002. 2001 Census of Canada. Catalogue No. 92-378-XPE or Catalogue No. 92-378-XIE.

Statistics Canada. *1996 Census Technical Reports: Mobility and Migration*. Ottawa: Industry Canada, 1997. 1996 Census of Canada. Catalogue No. 92-326E.