Catalogue No. 92-390-XIE

Journey to Work

2001 Census Technical Report

Statistics Canada Statistique Canada

Canada


ELECTRONIC PUBLICATIONS AVAILABLE AT
www.statcan.ca

2001 Census Technical Report

Journey to Work

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.

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 the Place of Work and Mode of Transportation variables. It has been prepared by the Place of Work Section, with the support of staff from Census Operations Division and 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. A *de jure* census counts people at their usual place of residence and not where they happen to be on Census Day.

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. For self-enumeration, a census representative drops 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 mid-November 2004.

1.2 Questionnaire and Instructions

Six types of questionnaires were used to collect 2001 Census data. Of these, the 2A questionnaire (short form) was distributed to four out of five households. The remainder of Canadians used either a 2B, a 2C or a 2D (long form) questionnaire. The 3A and 3B questionnaires were used in private dwellings to enumerate usual residents who wanted to be enumerated separately, and in collective dwellings.

Journey-to-work data were collected from the responses to two questions on the census long form: the Place of Work and Mode of Transportation questions. Information on commuting patterns is developed by combining the responses to these questions with Place of Residence information. The Place of Work question was first asked in the 1971 Census, and the Mode of Transportation question in the 1996 Census.

1.2.1 Place of Work

All non-institutional residents aged 15 years or older who had worked at some time since January 1, 2000 were asked to respond to the Place of Work question. The Place of Work question asked in the 2001 Census of Population is shown in Figure 1.

The following instructions appeared in the *2001 Census Guide:*

QUESTION 46 - Place of work

Mark **Worked at home** for persons aged 15 and over who worked at home-for example, farmers, private consultants, apartment building superintendents, etc.

For persons who worked part of the time at home **and** part of the time at an employer's address:

- mark **Worked at home** if most of the time was spent working at home (for example, three days out of five)
- mark **Worked at the address specified below** and print the employer's address if more time was spent working at an employer's address.

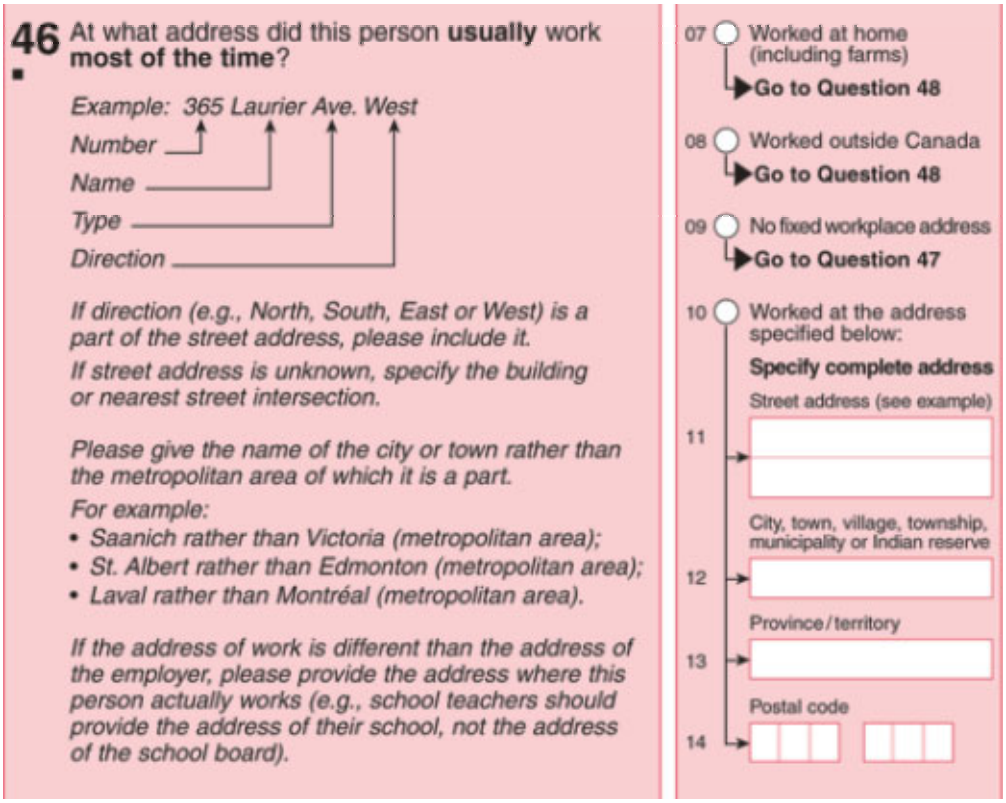
Mark **No fixed workplace address** for persons who go from home to various work locations--for example, building and landscape contractors, travelling salespersons, or independent truck drivers.

Mark **Worked at the address specified below** for persons who worked at an employer's address most of the time. Give a **complete address** including street number, name, type and, if applicable, direction. If the employer's address is unknown, print the name of the building or of the nearest street intersection. Report this person's regular place of work, even if he or she was temporarily on assignment, training or holiday last week.

For persons who **work at a different job site or location** every day:

- mark **Worked at the address specified below** and specify the address or name of the headquarters or depot **if they report there before starting work each day**.

Figure 1 - 2001 Place of Work Question



The Place of Work question has two objectives. The first is to identify the general workplace location of the respondent through the check-off categories and the second, to identify a precise workplace location using the specified address.

1.2.2 Mode of Transportation

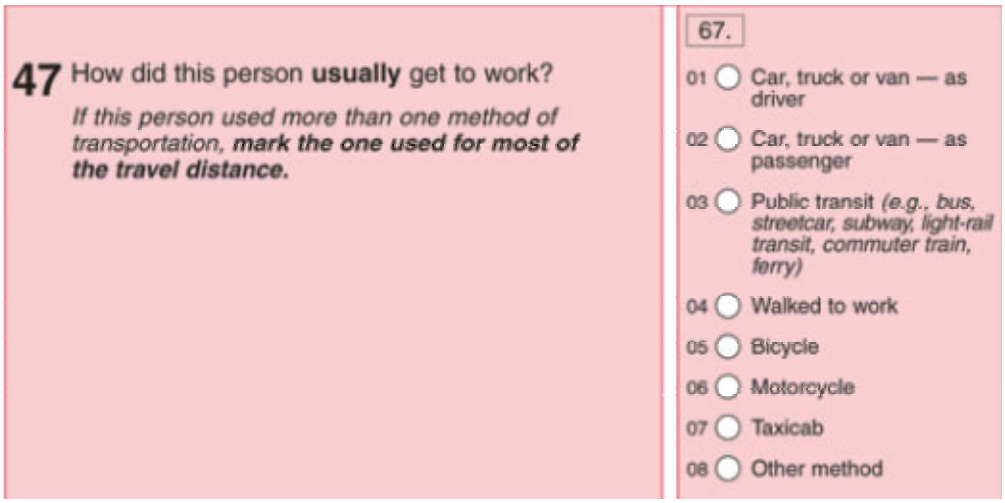
All persons responding to the Place of Work question who marked "No fixed workplace address" or "Worked at the address specified below" were asked to answer the Mode of Transportation question. The Mode of Transportation question asked in the 2001 Census of Population is shown in Figure 2.

The following instructions appeared in the *2001 Census Guide:*

QUESTION 47 - Transportation to work

Mark the type of transportation usually used to get to work. Mark only one circle indicating the type of transportation used for most of the distance travelled.

Figure 2 - 2001 Mode of Transportation Question



1.3 Variables

Place of Work Status

The Place of Work Status variable is created directly from the responses to the Place of Work question. The variable has four categories:

(i) Worked at home

Persons whose job is located in the same building as their place of residence, persons who live and work on the same farm, building superintendents and teleworkers who spend most of their work week working at home.

Note: Since the Census asks about a person's main job and where they work **most** of the time, not all persons who work at home are included in this category. Persons who have a job and also run a business at home part time are likely not counted in the work-at-home category. Also, teleworkers who work at home one to two days a week and in a central office the rest of the week may not be counted at home. Other surveys which ask respondents if they work **some** of their hours at home provide a much higher estimate of the work-at-home population than the census.[1]

(ii) Worked outside Canada

Persons who work at a location outside Canada. This can include diplomats, Armed Forces personnel and other persons enumerated abroad. This category also includes recent immigrants who may not currently be employed, but whose job of longest duration since January 1, 2000 was held outside Canada.

(iii) No fixed workplace address

Persons who do not go from home to the same workplace location at the beginning of each shift. Such persons include building and landscape contractors, travelling salespersons, independent truck drivers, etc.

[1] E. Akyeampong and R. Nadwodny, "Evolution of the Canadian Workplace: Work From Home," *Perspectives on Labour and Income.* 13, 4, (Winter 2001): 30-31.

(iv) Worked at the address specified below (Usual place of work)

Persons who are not included in the categories described above and who report to the same (usual) workplace location at the beginning of each shift are included here. Respondents were asked to provide the street address, city, town, village, township, municipality or Indian reserve, province/territory and postal code of their workplace. If the full street address was not known, the name of the building or nearest street intersection could be substituted.

Workplace Location (Place of Work Geography)

The second objective of the Place of Work question is to identify the specific location of a workplace by assigning it a precise geographic location. This is accomplished through the workplace location coding system. Workplace location data are available for most standard geographic areas either directly from the database or by aggregating other standard areas. Outside census metropolitan areas (CMA) and census agglomerations (CA), the data are coded to the census subdivision level. Within census metropolitan areas and census agglomerations, the coding is much more detailed. Consequently, for certain levels of geography, the workplace location is not available for the entire country. The following geographic areas are available only within census metropolitan areas and census agglomerations: census tracts, urban areas, dissemination areas, and blocks. Postal codes and designated places are not supported.

By using the workplace location and place of residence, origin-destination commuting flows can be produced for a variety of levels of geography.

Commuting Distance

Commuting distance is derived by calculating the straight-line distance in kilometres between a respondent's residential block and their workplace location representative point. Caution should be exercised when interpreting the distances for persons who work outside CMA or CA coverage since their workplace location is coded to the census subdivision level.

Mode of Transportation

The Mode of Transportation variable is derived from the check-off responses to the Mode of Transportation question. The question is asked of persons whose Place of Work status is either "Worked at the address specified below" (usual place of work) or "No fixed workplace address". It has eight categories: Car, truck or van--as driver; Car, truck or van--as passenger; Public transit (e.g., bus, streetcar, subway, light-rail transit, commuter train, ferry); Walked to work; Bicycle; Motorcycle; Taxicab; Other method.

Persons who use more than one mode of transportation were asked to identify the single mode they used for most of the travel distance. As a result, the responses to this question provide data on the primary mode of transportation to work. The question does not measure multiple modes of transportation, nor does it measure the seasonal variation in mode of transportation or trips made for purposes other than commuting from home to work.

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 (over 4 billion keystrokes), Regional Processing has been contracting this work out to the Canada Customs and Revenue Agency (CCRA), formerly called Revenue Canada, since 1981. 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 used by 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 a 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, to produce industry statistics under the North American Free Trade Agreement (NAFTA). This meant a change for industry coding--in 1996, industry was coded to the 1980 Standard Industrial Classification (SIC). In order to allow longitudinal comparisons, the 2001 Industry question was also coded to the 1980 SIC, using more automated means during Automated Coding (see Section 2.1.4).

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 the 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).

The 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 on 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 for their work.

As the EA boxes arrived at the 2001 processing site, they were registered and 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 and 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 archival purposes. 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 the 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 the 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, and consequently this process had an impact on the census counts.

As the census data arrived on cartridges from CCRA, they were loaded onto Statistics Canada's computers, ready for the Interactive Verification activities. A series of automated "structural" edits were performed, mainly to verify the information on the front cover of the questionnaire filled out by the census representative. 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 the 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 can be identified as belonging to a particular block.

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

The 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) against 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 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 them 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 **N**earest-neighbour **I**mputation **M**ethod (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 (**CAN**adian **C**ensus **E**dit and **I**mputation **S**ystem) to include Edit and Imputation for such variables as industry, 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 **I**nstructions from **D**irectly **E**ntered **R**equirements) was used to process the remaining census variables such as mother tongue, dwelling, income, etc. This tool translates 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 (DCS) 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, 111,628 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 143,681 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 Journey to Work - Processing

2.2.1 Workplace Location Coding

Place of Work responses are among the most complex census responses to code given the extremely large number of workplace locations. Successful coding of workplace location depends on a number of factors including the completeness of the response, the coverage of the reference files and the availability of additional reference materials.

The ideal workplace response includes the civic number, the street name, the street type and direction (if applicable), the official municipality name (i.e. city, town, village, township, reserve, etc.), the province/territory name and the postal code. However, since respondents do not always know the full address of their workplace, the responses are often less than ideal. Responses may be missing information, the information provided may be ambiguous, and it may contain abbreviations and locally known names. In some cases, the response consists of a building name or nearby street intersection instead of a street address.

The coding system has been designed to handle various types of responses through the use of a series of reference files. The reference files used in coding were derived from the June 2001 version of Statistics Canada's National Geographic Base.

All questionnaires providing a written response to the Industry or Place of Work questions were processed through the workplace location coding system. For 2001, this represented approximately

3.5 million responses over an eight-month period. Workplace locations were coded to a point location represented by a latitude-longitude coordinate.

When the workplace was located outside a census metropolitan area (CMA) or census agglomeration (CA), the point location represents a census subdivision (CSD). For CMAs and CAs, the point location usually represents a blockface or a block; however, it can also represent a dissemination area or a census tract (for areas that are census tracted). When the combination of response and reference materials was insufficient to code to these detailed levels, the record was either coded to a census subdivision, or was considered uncodeable.

The workplace location coding system consisted of two distinct components: the automated component and the interactive component.

Automated System

The automated coding system parsed the address responses into their component parts and then matched the component parts to several distinct reference files, in sequence, until the response was successfully coded. Responses which could not be coded automatically were sent to the interactive coding system.

The first step involved filtering out responses such as "Worked at home", "Worked outside Canada", "No fixed workplace address", and not applicable responses such as "didn't work", "full-time student", "retired", "on disability", etc.

Responses that were not filtered out were sent to the Place Name module. The province and city/town response were matched to a list of province/territory names and place names within the Place Name Reference File. When a match was found and the reference file record was located outside CMA- and CA-covered areas, then the response was coded at the census subdivision level. When a matching record was located within a CMA- or CA-covered area, a preliminary CSD code was assigned and the response was sent through subsequent coding system modules for more detailed coding. The preliminary code was used to control the matching process.

Additional coding modules were applied in the following order until a match was found:

- the Postal Code module matched the response to the Postal Code Reference File;
- the Business/Building module matched responses to a list of large employment locations, business names and addresses;
- the Street Address module matched the response to the Street Address File;
- the Street Intersection module matched the response to the Street Intersection Reference File.

Interactive System

Responses which were not coded by the automated system were coded in the interactive coding system. Responses were separated into 172 regional databases based on the preliminary geography assigned in automated coding. The design of the interactive system helped to compensate for less than ideal responses.

Coding staff working on a regional coding database could see all responses within the database. The responses were presented to the coders in list format, similar to that of a spreadsheet, one response per line. Coders could see both the parsed and unparsed versions of the responses, residence information, and Place of Work check-off responses for each record within the selected regional database.

Coding staff created groups of similar responses, searched the reference files and selected the best-matching record, and coded the entire group of records together. Groups of records varied in size from as few as one or two records to several hundred records.

Coding staff had access to the reference files used in the automated system as well as to additional files such as enumeration area and census tract files. The most valuable additional file was the electronic map of Canada, which allowed coders to search for addresses and postal codes, zoom in on specific streets, see civic address ranges, and code responses to a specific block. Coding staff also had access to additional reference materials such as city directories, business directories, telephone number lookup software, commercially produced maps and some maps provided by municipal planning departments.

Quality Control (QC) coding was applied to all cases where a coder had coded a number of similar responses at the same time as well as to a sample of all other coded responses. QC coders checked to ensure that all responses in the group represented the same workplace location and then recoded the group. The original and QC codes were compared and, if there was a difference, the correct code was determined by an adjudication coder. Responses which were coded by the automated system were not included in QC coding, but were subject to further analysis.

Coding Fixes

The coded responses were analyzed by subject-matter specialists to identify coding errors. When errors were identified, the responses were extracted to a coding fix system for recoding.

Geography Update

To account for any changes made to the geographic bases between the time the coding reference files were produced and the time the geographic base was finalized, the latitude and longitude coordinates assigned during coding were used to update the interim geography to the final geography. After this update, the coded data were loaded to the Edit and Imputation processing databases.

2.2.2 Edit and Imputation

Within the Edit and Imputation (E & I) process, the Place of Work universe was identified, inconsistent responses were resolved, non-responses were resolved through donor imputation, and a number of variables were derived. Throughout processing, a series of tabulations were used and carefully reviewed in monitoring the changes to the various variables and to confirm that the processing worked correctly.

The Journey to Work E & I process was divided into four separate modules. The first module finalized the Place of Work Status variable, the second and third modules, the Workplace Location variables, while the fourth module finalized the Mode of Transportation variable.

Place of Work Status

The Place of Work universe includes all non-institutional residents 15 years of age and over who had worked at some time since January 1, 2000. Place of Work Status was derived from the four check-off responses to the Place of Work question and also took into account whether the response was coded to a workplace location. Table 1 shows the check-off response patterns observed prior to Edit and Imputation processing for persons in the Place of Work universe and indicates whether or not responses were coded to a location during Workplace Location coding. The distribution of responses is very similar to that observed in the 1996 Census.

Table 1: Place of Work Response Patterns

Check-off response	Workplace coded	%
At home	No	7.3
At home	Yes	0.5
Outside Canada	No	0.6
No fixed workplace	No	8.6
No fixed workplace	Yes	0.3
Address specified	Yes	37.2
Address specified	No	0.5
At home and address specified	Yes	0.2
Other multiple responses	No	0.2
Other multiple responses	Yes	0.1
No checkoff	Yes	37.8
No checkoff	No	6.7

For persons who were part of the Place of Work universe, responses were quite clean overall. Nevertheless, only about half of the respondents who provided their address of work marked the "Worked at the address specified below" (usual place of work) checkoff. These cases were deterministically resolved as a "usual place of work."

When a single Place of Work check-off response was provided, the corresponding Place of Work Status value was assigned. A series of deterministic edits were performed to resolve inconsistent and missing responses. Where Place of Work coding assigned "Worked at home," "Worked outside Canada," or "No fixed workplace address," and the Place of Work status was not the same, it was changed to be consistent with the coded response. Where coding assigned a geographic location and both "Worked at home" and "Worked at the address specified below" were checked, the workplace location was compared to the residence geography. If the geographies were the same, then the status was set to "Worked at home," otherwise it was set to "Usual place of work."

Respondents who did not provide a Place of Work Status response were assigned "Worked at home" if they reported being an employee of the head of the household or if they were a farm operator and the dwelling was a farm dwelling. Persons enumerated outside Canada who worked at a job or were absent or on layoff from a job in the week before Census Day were assigned "Worked outside Canada."

When it was not possible to deterministically resolve the Place of Work status, a value was imputed from a donor record. The record to be imputed (failed record) was compared to a series of donor records in order to select a donor record having similar characteristics. The variables used in the matching process were Document Type (which identifies persons enumerated outside Canada), Class of Worker, Industry Code, Occupation Code, Age and Sex. Where multiple responses were provided, the imputed value must be the same as one of the provided responses. No stratification was used for the imputation of Place of Work Status.

Workplace Location

The second imputation module focused on assigning workplace locations at the CMA, CA and CSD levels of geography to persons with a usual place of work status. Records requiring imputation were either non-responses, or else the response was incomplete or too ambiguous to be geocoded to the CSD. The

variables used in the matching process were Industry Code, Occupation Code, Age and Sex. Stratification was not used in this imputation module.

The third module focused on assigning detailed workplace locations (i.e. census tract, dissemination areas, block and representative point) for persons with a usual Place of Work status and who worked within a CMA or CA. The records imputed in this module fall into two groups: records which had no workplace location and were imputed for the CMA, CA and CSD levels in the previous imputation module, and records which could only be coded to the CSD level because of the limited information provided. This module used the same matching variables as the previous module. For imputation purposes, the records were stratified (i.e. separated into mutually exclusive groups) based on the CSD of work. Therefore, the donor record was restricted to donors working in the same CSD as that of the failed record.

After imputation, the workplace locations were finalized by assigning the residence locations as the workplace locations for all persons who worked at home. The commuting distance was derived by calculating the straight line distance between the residence block latitude and longitude coordinates and the workplace location representative point latitude and longitude coordinates.

Mode of Transportation

The fourth module finalized the Mode of Transportation variable. Persons with a single check-off response to the Mode of Transportation question were assigned the corresponding Mode of Transportation value. When there was no response, or when more than one response was provided, the mode of transportation was imputed from a donor record. Where two responses were provided, the imputed value must be the same as one of the provided responses.

The Mode of Transportation universe includes all non-institutional residents 15 years of age and over, with a Place of Work status of "No fixed workplace address" or "usual place of work", who had worked at some time since January 1, 2000. Response patterns prior to Edit and Imputation processing are shown in Table 2. Respondents were asked to mark the single mode used for most of the travel distance. The distribution of responses is similar to that observed in the 1996 Census.

Table 2: Mode of Transportation Responses

Check-off response	%
Driver	67.0
Passenger	7.3
Public transit	9.5
Walked	7.3
Bicycle	1.2
Motorcycle	0.1
Taxicab	0.2
Other method	1.1
No checkoff	4.4
Total multiple response	2.0
- two checkoffs	1.8
- more than two checkoffs	0.3

The Place of Work question instructed persons who worked at home or outside Canada to skip the Mode of Transportation question. This instruction was often missed, possibly because the Place of Work and Mode of Transportation questions appeared on different pages. In these cases, the mode of transportation was deterministically set to "Not applicable."

For imputation purposes, the records were stratified based on workplace location, with one stratum being for persons with no fixed workplace, one for each tracted CMA or CA of work, one for each non-tracted CA of work, and, in the case of persons working outside CMAs or CAs, one for each province or territory.

The matching process considered the distance between the donor record's workplace and the failed record's workplace in order for both workplaces to be in the same area. Other matching variables used included Age, Sex, Commuting Distance, Full- or Part-time status, Industry, Occupation, and whether the respondent had a spouse (legal or common-law). The relative importance of the matching variables was adjusted according to which stratum was being processed. For example, since distances cannot be calculated for persons with no fixed workplace, distances were not used for this stratum. The importance of distance was less for strata outside CMAs and CAs since the workplace in these areas was coded to the CSD level.

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 Journey to Work

3.2.1 Non-response and Multiple Response

Place of Work

Persons who did not mark a check-off response to the Place of Work question and who could not be coded to a workplace location are considered to be non-respondents. Persons who marked two or more check-off responses are multiple respondents.

Table 3 shows the provincial and territorial non-response and multiple response rates for the Place of Work universe (i.e. non-institutional residents 15 years of age and over who worked at some time since January 1, 2000) as well as for the employed labour force. The employed labour force, a subset of the Place of Work universe, is used for all published Place of Work tabulations.

Table 3: Place of Work Non-response and Multiple Response Rates

Province/Territory	Place of Work universe		Employed labour force	
	Non-response	Multiple response (%)	Non-response (%)	Multiple response (%)
Newfoundland and Labrador	5.0	0.4	3.2	0.3
Prince Edward Island	5.5	0.4	3.8	0.5
Nova Scotia	4.9	0.4	3.6	0.5
New Brunswick	5.7	0.4	4.3	0.4
Quebec	5.8	0.4	4.6	0.4
Ontario	6.4	0.5	4.9	0.5
Manitoba	4.3	0.5	3.5	0.5
Saskatchewan	4.4	0.8	3.6	0.8
Alberta	5.5	0.6	4.5	0.6
British Columbia	6.7	0.5	5.3	0.5
Yukon	7.0	0.4	5.5	0.5
Northwest Territories	6.0	0.3	4.9	0.3
Nunavut	4.2	0.2	3.7	0.3
Canada	**6.0**	**0.5**	**4.7**	**0.5**

Higher non-response rates were observed in British Columbia, the Yukon and Ontario and lower rates, in Manitoba, Saskatchewan, Nunavut and the Atlantic provinces. The non-response rates for the employed labour force are about 1.3% lower than those for the Place of Work universe. Little variation is observed in the multiple response rates.

Table 4 shows the non-response and multiple response rates for Canada by sex and age. Higher non-response rates are observed for youth aged 15 to 24 and seniors 65 years and older. Overall, the non-response rate for men is higher than for women in the lower age groups, but starting at 55 years, the rate for women is higher. Multiple response rates are higher for persons in the older age groups.

Table 4: Place of Work Non-response and Multiple Response Rates by Sex and Age

Sex	Age	Place of Work universe		Employed labour force	
		Non-response (%)	Multiple response (%)	Non-response (%)	Multiple response (%)
Both sexes	**Total**	**6.0**	**0.5**	**4.7**	**0.5**
Both sexes	15-24	9.8	0.3	6.9	0.3
Both sexes	25-34	5.0	0.4	4.2	0.4
Both sexes	35-44	4.5	0.5	3.9	0.6
Both sexes	45-54	4.7	0.6	4.0	0.6
Both sexes	55-64	6.3	0.7	5.1	0.8
Both sexes	65 and over	12.9	1.1	10.2	1.2
Male	**Total**	**6.1**	**0.6**	**4.8**	**0.6**
Male	15-24	10.4	0.3	7.4	0.3
Male	25-34	5.2	0.4	4.4	0.5
Male	35-44	4.5	0.6	4.0	0.6
Male	45-54	4.6	0.7	4.0	0.7
Male	55-64	6.0	0.8	4.9	0.8
Male	65 and over	11.1	1.1	8.8	1.2
Female	**Total**	**5.9**	**0.5**	**4.5**	**0.5**
Female	15-24	9.1	0.2	6.4	0.2
Female	25-34	4.9	0.4	3.9	0.4
Female	35-44	4.6	0.5	3.8	0.5
Female	45-54	4.8	0.6	4.0	0.6
Female	55-64	6.6	0.7	5.3	0.7
Female	65 and over	16.4	1.0	13.0	1.1

Mode of Transportation

Persons who did not mark a check-off response to the Mode of Transportation question are non-respondents. Persons who marked two or more check-off responses are multiple respondents.

Table 5 shows the provincial and territorial non-response and multiple response rates for the Mode of Transportation universe (i.e. non-institutional residents 15 years of age and over who worked at some time since January 1, 2000 and who have a Place of Work status of "No fixed workplace address" or "Usual place of work") as well as for the employed labour force.

Table 5: Mode of Transportation Non-response and Multiple Response Rates

Province/Territory	Mode of Transportation universe		Employed labour force	
	Non-response (%)	Multiple response (%)	Non-response (%)	Multiple response (%)
Newfoundland and Labrador	3.4	1.9	2.1	1.8
Prince Edward Island	3.6	1.1	2.4	1.0
Nova Scotia	3.3	1.9	2.2	1.8
New Brunswick	3.5	1.6	2.3	1.5
Quebec	3.7	2.4	2.8	2.3
Ontario	4.1	2.0	2.8	2.0
Manitoba	2.8	2.1	2.1	2.0
Saskatchewan	3.0	1.9	2.2	1.9
Alberta	3.4	1.8	2.5	1.8
British Columbia	4.5	2.1	3.3	2.1
Yukon	6.0	1.8	5.0	1.8
Northwest Territories	5.8	1.9	4.7	2.0
Nunavut	4.4	1.1	3.9	1.2
Canada	**3.8**	**2.1**	**2.7**	**2.1**

The provincial and territorial variations are similar to those observed for Place of Work. The non-response rates for the employed labour force are about 1.1% lower than for the Mode of Transportation universe.

Table 6 shows the non-response and multiple response rates for Canada by sex and age. Higher non-response rates are observed for youth aged 15 to 24 and seniors 65 years and older. Overall, the non-response rate for women is higher than for men except in the 15-to-24 age group. Multiple response rates are higher for persons in the 15-to-24 age group.

Table 6: Mode of Transportation Non-response and Multiple Response Rates by Sex and Age

Sex	Age	Mode of Transportation universe		Employed labour force	
		Non-response (%)	Multiple response (%)	Non-response (%)	Multiple response (%)
Both sexes	**Total**	**3.8**	**2.1**	**2.7**	**2.1**
Both sexes	15-24	6.3	3.0	3.8	3.1
Both sexes	25-34	3.0	2.2	2.4	2.2
Both sexes	35-44	2.8	1.9	2.3	1.9
Both sexes	45-54	2.9	1.7	2.4	1.7
Both sexes	55-64	4.4	1.5	3.3	1.5
Both sexes	65 and over	11.7	1.6	8.6	1.7
Male	**Total**	**3.7**	**1.9**	**2.7**	**1.9**
Male	15-24	6.6	2.9	4.0	2.9
Male	25-34	2.9	2.1	2.4	2.0
Male	35-44	2.6	1.8	2.2	1.8
Male	45-54	2.7	1.5	2.3	1.5
Male	55-64	4.0	1.3	3.1	1.3
Male	65 and over	9.3	1.4	6.9	1.5
Female	**Total**	**4.0**	**2.3**	**2.8**	**2.3**
Female	15-24	6.0	3.1	3.6	3.2
Female	25-34	3.2	2.3	2.4	2.3
Female	35-44	3.0	2.1	2.3	2.1
Female	45-54	3.1	1.9	2.5	1.9
Female	55-64	5.0	1.9	3.7	1.9
Female	65 and over	16.6	1.9	12.6	2.1

3.2.2 Coding Results

Table 7 shows the geographic level the responses were coded to during the coding operation. Non-respondent records are excluded since they were never sent to coding. Responses coded to a CSD in a CMA or CA area and uncoded responses were resolved through imputation. The uncoded category can also include responses such as "going to school", "he retired in 1993", "on disability", etc.

Table 7: Geographic Level Coded for Employed Labour Force, Excluding "Worked at Home"

Geographic Level	Canada (%)	Tracted CMAs and CAs (%)	Non-tracted CAs (%)
Block-face	62.4	69.2	7.8
Block	25.5	22.5	49.8
Dissemination area	4.2	0.4	34.6
Census tract	0.1	0.1	0.0
CSD in a CMA/CA area	4.8	4.8	5.3
Not codeable	3.0	3.1	2.5

3.2.3 Imputation Rates

Table 8 shows the imputation rates for each of the imputation modules discussed in Section 2.2.2, for Canada and the provinces and territories, based on the employed labour force. Three imputation rates are shown for workplace location. For areas within CMAs and CAs, the first rate is for those cases that had no workplace location (i.e. all workplace location levels were imputed). The second rate is for those cases that were coded to a workplace location at the CSD level but lacked the necessary information to be coded to a more detailed level. The third imputation rate is for areas outside CMAs and CAs which only required coding to the CSD level. Taking Nova Scotia as an example, for areas outside CMAs and CAs, 3.4% of the workplace location areas were imputed. Within CMAs and CAs, 3.9% of the workplace location areas were imputed to the CSD level, and a further 6.3% were imputed below the CSD level. Therefore, in Nova Scotia, for data tabulated at a lower level (e.g. census tracts), 10.2% (3.9% + 6.3%) of the workplace locations were imputed.

Table 8: Employed Labour Force, Imputation Rates

Province/Territory	Place of Work Status (module 1) (%)	Workplace Location CSD and higher levels (areas inside CMAs/CAs) (module 2)(%)	Workplace Location Below CSD (areas inside CMAs/CAs) (module 3) (%)	Workplace Location (areas outside CMAs/CAs) (module 2) (%)	Mode of Transportation (module 4) (%)
Newfoundland and Labrador	3.3	3.4	4.6	3.3	4.2
Prince Edward Island	4.1	4.0	5.0	3.7	3.7
Nova Scotia	3.7	3.9	6.3	3.4	4.1
New Brunswick	4.4	4.1	4.1	5.4	4.1
Quebec	4.7	5.1	4.0	4.8	5.2
Ontario	5.5	5.5	5.1	5.2	5.4
Manitoba	3.6	3.5	3.1	3.4	4.5
Saskatchewan	3.5	3.4	3.5	3.7	4.2
Alberta	4.9	4.9	5.5	5.0	4.8
British Columbia	6.3	5.8	4.8	6.1	6.4
Yukon	8.6	4.8	4.4	8.6	9.3
Northwest Territories	4.9	4.3	7.5	6.1	6.7
Nunavut	3.8	--	--	4.1	5.1
Canada	**5.1**	**5.2**	**4.7**	**4.8**	**5.3**

The provincial/territorial pattern of imputation rates is similar to that observed for non-response rates. The imputation rates are slightly higher for the full Place of Work universe.

For module 3, imputation below the CSD level in CMAs and CAs, the imputation rates do not exhibit the same provincial patterns as the other four imputation rates, and the difference between the employed labour force and the full Place of Work universe is only 0.3%. Cases imputed in module 3 generally provided an address response, but not enough information for responses to be fully coded.

The proportion of cases imputed for each Place of Work status was as follows: 5.7% of all "Worked at home," 8.1% of all "No fixed workplace address,"6.1% of all "Worked outside Canada," and 6.3% of all "Usual place of work." The distributions of the unimputed responses and of the final Place of Work status were virtually identical.

Table 9 shows the proportion of cases imputed for each mode of transportation. For example, the mode of transportation was imputed for 9.6% of all persons who biked to work. The final Mode of Transportation distribution, while similar to the unimputed mode of transportation, shows a decrease in driving to work and an increase in passenger, public transit and walking. The higher imputation rate among youth aged 15 to 24 years and seniors is likely responsible for the differences observed between the unimputed and final data.

Table 9: Proportion of Cases Imputed by Mode of Transportation

Mode of transportation	Imputed (%)
Bike	9.6
Driver	5.3
Motorcycle	5.8
Other	8.6
Passenger	8.3
Taxi	8.5
Transit	8.8
Walked	10.4

3.3 Sources of Errors and Evaluation Studies

The census is collected through self-enumeration (respondents fill out the questionnaire themselves). As a result, the data are dependent on the respondent understanding the question and being able to answer with appropriate detail. There is generally no opportunity to probe a respondent for more information.

The number of workers counted by the census in a given geographic area may differ from the counts derived from other sources--business and establishment surveys, for example--since companies with more than one location often report all of their workers as working at one location (e.g. head office). In addition, the census only collects detailed information for one job. Persons having multiple jobs are only counted at their main job.

The census assumes that, in commuting flows, the commute to work originates from the home, but this may not always be the case. In some cases, respondents may be on a business trip and may have reported based on where they were working during the trip. Some persons maintain a residence close to work and commute to their home on weekends. Primary industries operating in remote areas often bring in workers from far away and schedule them to work two weeks at a time, followed by a two-week period

when they may return home. Students often work after school at a location near their school. As a result, the data can show unusual commutes and an unusual mode of transportation.

Journey-to-work data were evaluated throughout the processing operations. Distributions for check-off responses to the Place of Work and Mode of Transportation questions were reviewed shortly after data capture completion and found to be very similar to those from the 1996 Census. During workplace location coding, considerable efforts were taken to identify and correct systematic coding errors. Monitoring tabulations were produced during Edit and Imputation processing to ensure that everything was proceeding correctly. Afterwards, the data were subjected to further evaluation intended to assess their suitability for release.

The analyses and evaluations performed for each of the Journey to Work variables are described in the following sections.

3.3.1 Place of Work Status

Comparisons of 2001 data were carried out between geographic areas. Data from the 2001 Census were also compared to equivalent data from the 1996 Census.

The comparisons began at the Canada level and proceeded to increasingly detailed levels of geography. Data were evaluated for the provinces and territories, census metropolitan areas and census agglomerations, census divisions, and census subdivisions.

Table 10 shows a Place of Work Status distribution very similar to that of 1996. Slight declines are observed for "Worked at home" and for "Usual place of work," while increases are observed for the "Worked outside Canada" and "No fixed workplace address" categories.

Table 10: Place of Work Status Final Distribution, Canada

	At home (%)	No fixed workplace (%)	Outside Canada (%)	Usual place of work (%)
2001 distribution	8.0	8.7	0.5	82.9
1996 distribution	8.2	7.6	0.4	83.9
2001 vs 1996 change	-0.2	1.1	0.1	-1.0

Other Statistics Canada surveys show that the proportion of the workforce working at home increased from 16% to 17% in the 1995 to 2000 period.[2] This proportion is much higher than that observed in the census due to differences in the wording of the question. The surveys ask whether the respondents work any of their regular work hours at home, while the census asks where respondents usually work most of the time. Between 1996 and 2001, the number of farms in Canada declined almost 11%.[3] If the farming occupations are removed from the labour force, according to the census the proportion of the labour force working at home actually increases from 6.2% to 6.4% over the 1996 to 2001 period.

The decrease in "Worked at home" may result from two opposing trends: an increase in working at home (teleworking and home-based businesses) combined with a decrease in the number of workers in the agriculture sector, which historically accounted for about one-quarter of the census work-at-home count.

[2] E. Akyeampong and R. Nadwodny, "Evolution of the Canadian Workplace: Work From Home," *Perspectives on Labour and Income*. 2, 9 (September 2001): 31.
[3] Statistics Canada, "2001 Census of Agriculture--Canadian farm operations in the 21st century," *The Daily* (http://www.statcan.ca). May 15, 2002.

In 2001, more respondents provided a clean "Worked outside Canada" check-off response than in 1996. Changes in coding did not affect the number of persons coded as working outside Canada. The increase in the "Worked outside Canada" category is occurring in most areas of the country. However, a few CMAs are showing an increase in their share of the national total. The most notable case is Windsor, now home to 10.2% of all persons who reported working outside Canada, up from 7% in 1996.

In analyzing the increase in "No fixed workplace address" reporting, the employment distribution among occupations at the 4-digit-code level was compared for 1996 and 2001. It was found that this distribution was very similar for both census years. Respondents working in certain occupations were much more likely to report "No fixed workplace address" than others. Approximately 25% of the "No fixed workplace" count is for persons working in construction; a further 12% is for persons working in transportation, 7%, for occupations such as cleaners and janitors, and 3.5%, for couriers and messengers and survey interviewers.

The analyses conducted resulted in the conclusion that there are no major errors in the 2001 Census Place of Work Status variable distribution.

3.3.2 Workplace Location

Analysis of workplace locations is the most difficult and time-consuming portion of the evaluation process. During the coding operation, all the inter-provincial commuting flows were reviewed to ensure that assigned codes were consistent with responses. CSD-to-CSD commuting flows were tabulated separately for automatically coded and interactively coded records, and all commuting flows over 75 kilometres including at least five persons were then subjected to review. As a result of these reviews, a number of cases were identified where responses had been coded to the wrong CSD due to the existence of duplicate place names. These cases were corrected by recoding the responses through the coding fix system.

After processing, the workplace location data were assessed by comparing distributions--within the 2001 Census and between the 1996 and 2001 censuses--for a variety of geographic levels. Comparisons were done at the Canada, province/territory, census division, census subdivision, census metropolitan area and census agglomeration, and census tract levels. However, the focus was on the census subdivision and census tract levels.

Since geographic boundaries may change between censuses, the data were resolved to a common geography. For the CSD level, the 1996 data were reconstituted to the extent possible to the new 2001 CSD boundaries. For the CT level, the 2001 data were reconstituted to the 1996 CT boundaries.

A number of different indicators were tabulated to assess workplace location data:

Resident employed labour force counts (RELF) - the total number of persons in the employed labour force living in a given census subdivision.

Worker employed labour force counts (WELF) - the total number of persons in the employed labour force whose workplace location is in a given census subdivision.

Residence equals workplace (R=W) - the total number of persons in the CSD who live and work in the same CSD.

Commuting distance statistics such as mean, median, and percentiles (P75, P90, P95, P99).

In addition, three ratios were constructed for use in the assessment:

WELF/RELF - this ratio gives the number of persons working in a given area, divided by the number of employed labour force persons living in that area. This ratio is the most valuable ratio for the analysis of workplace locations. Within a self-contained labour market area, this ratio is approximately equal to 1. For smaller areas like CSDs, a rural area will typically have a low WELF/RELF ratio (less than 1), while cities and towns that are employment centres will have a higher value (greater than 1) since these areas attract workers from neighbouring areas. A WELF/RELF ratio that exceeds 2 is rare. These values are also very dependent on municipal structure. For example, a city which has a large land area incorporating rural areas will have a lower ratio than a city which incorporates only developed areas. Municipal amalgamations such as those in Halifax, Ottawa, Toronto and Hamilton have resulted in a noticeable reduction of this ratio for these cities.

R=W/RELF - this ratio gives a measure of the locally employed residents (i.e. the proportion of the resident employed labour force that works in their CSD of residence). It can range from 0 to 1. For a CSD that has a high WELF/RELF ratio, the R=W/RELF would also be high. It would be unusual for an area that attracts workers from other municipalities to not also supply a number of jobs to local residents.

R=W/WELF - this ratio gives a measure of the locally hired workers (i.e. the proportion of workers working in a given CSD who also live in that CSD). It can range from 0 to 1. It is very similar to the previous ratio but a different base is used for comparison.

For the census subdivision level, the indicators and ratios were produced for each CSD and the differences calculated between 2001 and 1996 data. Comparisons were made and CSDs having suspicious values were subjected to further investigation. This could include reviewing the CSD-to-CSD commuting flows as well as reviewing the responses and codes from the coding process. As a result of these investigations and reviews, the responses from 15 CSDs were recoded and corrected on the database.

A slightly different approach was employed for the census tract (CT) level. While census tracts are relatively stable in terms of population, they vary greatly in their job numbers. A great number of CTs have little or no employment (e.g. residential suburbs) while a fairly small number have very high employment.

In evaluating the census tracted areas, the number of persons working in each CT (WELF) was tabulated, each census tract's share of total CMA or CA employment was calculated, and each CT was rank ordered within the CMA or CA according to its employment. This procedure was done for both the 1996 and 2001 census data. Differences between 2001 and 1996 were calculated for WELF, for CMA/CA employment share, and for employment rank.

Where large differences were observed, the responses were retrieved from the coding file and reviewed. As a result of these investigations and reviews, responses from seven CTs were recoded and corrected on the database.

3.3.3 Commuting Distance

Commuting distances were tabulated and used in evaluating workplace locations, but were not directly assessed since commuting distance is calculated from the residence and workplace locations.

3.3.4 Mode of Transportation

Mode of Transportation data analysis was performed by assessing both the Mode of Transportation distribution across a number of geographic levels and the change in distribution from 1996 to 2001.

Table 11 compares the Mode of Transportation distribution for 2001 and 1996. The distribution is quite stable. The notable changes are increases in the Driver and Transit categories and decreases in the Passenger and Walking categories.

Table 11: Mode of Transportation Final Distribution, Canada

Mode of transportation	2001 (%)	1996 (%)
Driver	73.8	73.3
Passenger	6.9	7.4
Transit	10.5	10.1
Walking	6.6	7.0
Bike	1.2	1.1
Taxi	0.2	0.2
Motorcycle	0.1	0.1
Other	0.8	0.8

The changes observed are consistent with the trends of increasing population in suburban areas and growth of employment outside central areas of cities. While very little data are directly comparable to census mode-of-transportation data, increases have been observed in both the number of registered vehicles[4] and in the number of passengers on urban transit systems.[5]

Public transit shares are highest in the most urbanized provinces and in the larger CMAs. The Northwest Territories and Nunavut are quite different from the rest of the country since they have much higher shares for "Walking" and "Other" modes.

It must be noted that transit strikes were experienced in Calgary and Vancouver in 2001. In Calgary, the strike ended before the census was taken, but in Vancouver it lasted over three months and included the census period. According to the Translink website (responsible for public transit in the Greater Vancouver area), ferry service between Vancouver and North Vancouver and local bus services were shut down for the strike. However, the light-rail service (Skytrain) was operating during the strike.

[4] Statistics Canada, *Road Motor Vehicles - Registration* (1999 and 2000 annual surveys), CANSIM, Table 405-0004.
[5] Statistics Canada, *Passenger Bus and Urban Transit Statistics*, Catalogue no. 53C0001, and *Canadian Urban Transit Association* website, http://www.cutaactu.on.ca (accessed on various dates between 2002 and 2003).

4. Historical Comparability

4.1 Conceptual Changes

Journey to work data are generally comparable from 1971, 1981, 1991, 1996 and 2001. However, there have been some changes between censuses.

4.1.1 Changes to the Place of Work Question

The place of work question was included in almost every census since 1971 in almost the same format. Nonetheless, some changes were made over the years.

In 1991, a write-in box for postal code responses was added. Three changes were implemented for 1996. The "No fixed workplace address" response category was made explicit. The "County" write-in box was removed after investigation indicated that most persons did not respond and many of those who did respond confused "county" with "country" and wrote "Canada" as a response. Clearer instructions, with an example of a complete civic address response, were included to assist respondents in providing a complete workplace address.

The 2001 Place of Work question was changed slightly from 1996 to include the phrase "most of the time". The question reads "At what address did this person usually work most of the time?"

4.1.2 No Fixed Workplace Address

The "No fixed workplace address" category did not exist in 1971. Persons who had no usual place of work address, and who did not report to a headquarters or depot at the beginning of each shift, were instructed to write the address where they most often worked in the job described. Persons having no fixed workplace address were included with persons who worked at a usual place or in the "Not stated" category. In 1971, the "Not stated" category accounted for 9.2% of the respondents, many of whom worked in the construction or primary industries - industries which typically account for much of the "No fixed workplace" category.

In 1981 and 1991, persons who had no fixed workplace address, and who did not report to a headquarters or depot at the beginning of each shift, were instructed to write "No usual place of work" in the address write-in boxes. These instructions did not appear on the questionnaire, but were included in the Census Guide. Persons with "no usual place of work" were sometimes coded to a workplace location.

In 1996, a "No fixed workplace address" response category was explicitly added to the place of work question, thereby reducing response burden. Persons with "no fixed workplace" do not have a workplace location coded.

4.2 Processing Changes

4.2.1 Workplace Coding

In 2001, the workplace location of persons working in census metropolitan areas or census agglomerations was coded to a block-face, block or dissemination area representative point. The workplace location of persons working outside census metropolitan areas and census agglomerations was **coded** to a census subdivision representative point.

In 1996 the situation was similar to that of 2001. The workplace location of persons working in most urban areas was coded to a block-face, street intersection or enumeration area representative point. The

workplace location of persons working in rural areas and some urban areas was **coded** to census subdivision representative points.

Prior to 1996, workplace location data were first coded to the census subdivision and then were coded to the census tract in separate coding operations, therefore the census subdivision and census tract data may give different results.

In 1991 the census tract level was coded for 10% of residents of Ontario and part of west Quebec. In 1981 the census tract level was coded for workplaces within census tracts that were covered by street indexes. In 1971 the census tract level was coded for a 1/9 sample of residents of census divisions located within a 50 mile radius of tracted CMAs or CAs and who worked within a tracted CMA or CA.

4.2.2 Edit and Imputation (E & I)

A significant change occurred in the edit and imputation of 1981 data. In 1971, non-responses to the place of work question were reported as "Not stated". However, in 1981, the "Not stated" category was dropped and non-responses to the place of work question were changed to a specific response through imputation. Imputation was performed on both the Place of Work Status and Workplace Location (census subdivision level) variables. However, census tract data were not imputed. Since the 1991 Census, location data have been imputed for all missing workplace geographies.

4.2.3 Calculation of Commuting Distance

In 1971, commuting distance was calculated to the nearest half mile for workplace census subdivisions and for workplace census tracts. The distance was calculated between the residential enumeration area representative point and the workplace census subdivision and census tract representative points. Values of 251 miles or more were all stored as 251 miles. Commuting distance data are not available from the 1981 or 1991 censuses.

In 1996, the distance was calculated in kilometres, to the nearest 0.1km, between the residential enumeration area representative point and the workplace location representative point. In 2001, the distance was calculated in kilometres, to the nearest 0.1km, between the residential block representative point and the workplace location representative point. Values of 200.1 kilometres or more are all stored as 201.0 kilometres.

4.3 Changes in Geographic Framework

The comparability of workplace location data between censuses is affected by changes in the census subdivision, census metropolitan area, census agglomeration and census tract boundaries. Because of the large number of geographic areas and possible boundary changes between censuses, data users are encouraged to exercise caution when comparing workplace location data between censuses.

Municipal restructuring in many provinces has resulted in a number of changes at the census subdivision level. There were 910 dissolutions recorded between 1996 and 2001. In the same period, 519 incorporations were recorded, with the majority being newly amalgamated municipalities.

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.

DEFINITIONS - CENSUS VARIABLES

Commuting Distance

Refers to the distance, in kilometres, between the respondent's residence and his or her usual workplace location. The variable relates to non-institutional residents 15 years of age and over who worked at some time since January 1, 2000. The variable usually relates to the individual's job held in the week prior to enumeration. However, if the person did not work during that week but had worked at some time since January 1, 2000, the information relates to the job held longest during that period.

Employed (in Reference Week)

Persons who, during the week (Sunday to Saturday) prior to Census Day (May 15, 2001):
(a) did any work at all for pay or in self-employment or without pay in a family farm, business or professional practice;
(b) were absent from their job or business, with or without pay, for the **entire week** because of a vacation, an illness, a labour dispute at their place of work, or any other reasons.

Experienced Labour Force (in Reference Week)

Persons who, during the week (Sunday to Saturday) prior to Census Day (May 15, 2001), were employed or unemployed who worked for pay or in self-employment since January 1, 2000.

Experienced Labour Force = Employed + (Unemployed who last worked in 2000) + (Unemployed who last worked in 2001)

The **experienced labour force** can also be derived by excluding from the labour force those unemployed persons 15 years of age and over who have never worked or who had last worked prior to January 1, 2000 only.

Industry

Refers to the general nature of the business carried out in the establishment where the person worked. If the person did not have a job during the week (Sunday to Saturday) prior to enumeration (May 15, 2001), the data relate to the job of longest duration since January 1, 2000. Persons with two or more jobs were required to report the information for the job at which they worked the most hours.

Labour Force Activity (in Reference Week)

Refers to the labour market activity of the population 15 years of age and over in the week (Sunday to Saturday) prior to Census Day (May 15, 2001). Respondents were classified as either **employed**, or **unemployed**, or as **not in the labour force**. The **labour force** includes the **employed** and the **unemployed**.

Mode of Transportation

Refers to the mode of transportation to work of non-institutional residents 15 years of age and over who worked at some time since January 1, 2000. Persons who indicate in the place of work question that they either had no fixed workplace address, or specified a usual workplace address, are asked to identify the mode of transportation they most frequently use to commute from home to work. The variable usually relates to the individual's job in the week prior to enumeration. However, if the person did not work during that week but had worked at some time since January 1, 2000, the information relates to the job held longest during that period.

Not in the Labour Force (in Reference Week)

Refers to persons who, in the week (Sunday to Saturday) prior to Census Day (May 15, 2001), were neither employed nor unemployed. It includes students, homemakers, retired workers, seasonal workers in an "off" season who were not looking for work, and persons who could not work because of a long-term illness or disability.

On Temporary Lay-off or Absent from Job or Business (in Reference Week)

Refers to whether persons were, during the week (Sunday to Saturday) prior to Census Day (May 15, 2001), (a) on temporary lay-off from a job to which they expected to return or (b) absent from their job or business for the entire week. Reasons for absence included a vacation, an illness, a strike or lock-out at the respondent's place of employment, or other reasons, including maternity leave, bad weather, fire, personal or family responsibilities, and attendance at training courses if wages or salaries were received from the employer. Data were collected for persons who did not work for pay or in self-employment in the week prior to enumeration.

This variable is used to derive the respondent's labour force activity status. Other than for specialized research in consultation with census labour market analysts, the use of this variable on its own is not recommended.

Place of Work Status

Refers to the place of work of non-institutional residents 15 years of age and over who worked at some time since January 1, 2000. The variable usually relates to the individual's job held in the week prior to enumeration. However, if the person did not work during that week but had worked at some time since January 1, 2000, the information relates to the job held longest during that period.

Remarks: Respondent-completed responses

Worked at home - Persons whose job is located in the same building as their place of residence, persons who live and work on the same farm, building superintendents and teleworkers who spend most of their work week working at home.

Worked outside Canada - Persons who work at a location outside Canada. This can include diplomats, Armed Forces personnel and other persons enumerated abroad. This category also includes recent immigrants who may not currently be employed, but whose job of longest duration since January 1, 2000 was held outside Canada.

No fixed workplace address - Persons who do not go from home to the same workplace location at the beginning of each shift. Such persons include building and landscape contractors, travelling salespersons, independent truck drivers, etc.

Worked at the address specified below - Persons who are not included in the categories described above and who report to the same (usual) workplace location at the beginning of each shift are included here. Respondents are asked to provide the street address, city, town, village, township, municipality or Indian reserve, province/territory and postal code of their workplace. If the full street address was not known, the name of the building or nearest street intersection could be substituted.

Teleworkers who spend less than one-half of their workweek working at their home office are asked to report the full address of their employer. Persons whose workplace location varied, but who reported regularly to an employer's address at the beginning of each shift, are asked to report the full address of the employer.

The chart below indicates which journey-to-work data are available for each place of work status.

Place of Work Status	Workplace Location	Commuting Distance	Mode of Transportation
Worked at home	Available	Not applicable	Not applicable
Worked outside Canada	Not applicable	Not applicable	Not applicable
No fixed workplace address	Not applicable	Not applicable	Available
Usual place of work	Available	Available	Available

Unemployed (in Reference Week)

Persons who, during the week (Sunday to Saturday) prior to Census Day (May 15, 2001), were **without paid work or without self-employment work** and were **available for work** and either:

(a) had actively looked for paid work in the past four weeks; or
(b) were on temporary lay-off and expected to return to their job; or
(c) had definite arrangements to start a new job in four weeks or less.

When Last Worked for Pay or in Self-employment

Refers to the year or period in which persons last worked for pay or in self-employment, even if only for a few days.

Workplace Location

Refers to the geographic location of the workplace of non-institutional residents 15 years of age and over who worked at some time since January 1, 2000. The variable usually relates to the individual's job held in the week prior to enumeration. However, if the person did not work during that week, but had worked at some time since January 1, 2000, the information relates to the job held longest during that period.

DEFINITIONS - CENSUS GEOGRAPHY

Block

A block is an area bounded on all sides by roads and/or boundaries of standard geographic areas. Blocks cover all the territory of Canada. The block is the smallest geographic area for which population and dwelling counts are disseminated.

Block-face

A block-face is one side of a street between two consecutive features intersecting that street. The features can be other streets, boundaries of standard geographic areas, or limits of map tiles.

Block-faces are used for generating block-face representative points, which in turn are used for geocoding and census data extraction when the street and address information is available.

Census Division (CD)

Census division (CD) is the general term for provincially legislated areas (such as county, *municipalité régionale de comté* and regional district) or their equivalents. Census divisions are intermediate geographic areas between the province level and the municipality (census subdivision).

Census divisions have been established in provincial law to facilitate regional planning, as well as the provision of services that can be more effectively delivered on a scale larger than a municipality. In Newfoundland and Labrador, Manitoba, Saskatchewan, Alberta, Yukon Territory, Northwest Territories and Nunavut, provincial/territorial law does not provide for these administrative geographic areas. Therefore, Statistics Canada, in cooperation with these provinces and territories, has created equivalent areas called census divisions for the purpose of disseminating statistical data. In the Yukon Territory, the census division is equivalent to the entire territory.

Census Metropolitan Area (CMA) and Census Agglomeration (CA)

A census metropolitan area (CMA) or a census agglomeration (CA) is formed by one or more adjacent municipalities centred on a large urban area (known as the **urban core**). The census population count of the urban core is at least 10,000 to form a census agglomeration and at least 100,000 to form a census metropolitan area. To be included in the CMA or CA, other adjacent municipalities must have a high degree of integration with the central urban area, as measured by commuting flows derived from census place of work data.

If the population of the urban core of a CA declines below 10,000, the CA is retired. However, once an area becomes a CMA, it is retained as a CMA even if the population of its urban core declines below 100,000. The urban areas in the CMA or CA that are not contiguous to the **urban core** are called the **urban fringe**. Rural areas in the CMA or CA are called the **rural fringe**.

When a CA has an urban core of at least 50,000 based on census counts, it is subdivided into **census tracts**. Census tracts are maintained for the CA even if the population of the urban core subsequently falls below 50,000. All CMAs are subdivided into census tracts.

Census Subdivision (CSD)

Census subdivision (CSD) is the general term for municipalities (as determined by provincial legislation) or areas treated as municipal equivalents for statistical purposes (for example, Indian reserves, Indian settlements and unorganized territories).

Census Tract (CT)

Census tracts (CTs) are small, relatively stable geographic areas that usually have a population of 2,500 to 8,000. They are located in census metropolitan areas and in census agglomerations with an urban core population of 50,000 or more in the previous census.

A committee of local specialists (for example, planners, health and social workers and educators) initially delineates CTs in conjunction with Statistics Canada. Once a census metropolitan area (CMA) or census

agglomeration (CA) has been subdivided into census tracts, the census tracts are maintained even if the urban core population subsequently declines below 50,000.

Dissemination Area (DA)

The dissemination area (DA) is a small, relatively stable geographic unit composed of one or more blocks. It is the smallest standard geographic area for which all census data are disseminated. DAs cover all the territory of Canada.

Geocoding

Geocoding is the process of assigning geographic identifiers (codes) to map features and data records. The resulting **geocodes** permit data to be linked geographically.

Households and postal codes are linked to block-face representative points when the street and address information is available; otherwise, they are linked to block representative points.

Province or Territory

Province and territory refer to the major political units of Canada. From a statistical point of view, province and territory are basic areas for which data are tabulated. Canada is divided into ten provinces and three territories.

Representative Point

A representative point is a single point that represents a linear or areal feature. The point is centrally located along the linear feature or centrally within the areal feature.

Representative points are generated for block-faces, blocks, enumeration areas, dissemination areas, census subdivisions and designated places.

Standard Geographical Classification (SGC)

The Standard Geographical Classification (SGC) is Statistics Canada's official classification for three types of geographic areas: **provinces and territories, census divisions (CDs)** and **census subdivisions (CSDs)**. The SGC provides unique numeric identification (codes) for these hierarchically related geographic areas.

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.

5. Journey-to-Work Data

- Journey-to-work data were released on February 11, 2003 under the topics "Place of Work" and "Commuting to Work."

Appendix C. Tables

Bibliography

Singbeil, B.J. 2002. *Certification Report: Journey to Work Variables*. 2001 Census of Population. Unpublished document. Ottawa. Statistics Canada.

Statistics Canada. 2002. *2001 Census Dictionary*. Catalogue no. 92-378-XPE or 92-378-XIE. Ottawa.

Statistics Canada. 2001. *E & I Narrative: Mode of Transportation Process*. 2001 Census of Population. Unpublished document. Ottawa.

Statistics Canada. 2001. *E & I Narrative: Place of Work Process*. 2001 Census of Population. Unpublished document. Ottawa.

Statistics Canada. 2000. *Place of Work Batch Automated Coding System Specifications*. 2001 Census of Population. Unpublished document. Ottawa.

Statistics Canada. 2000. *Place of Work Interactive On-line Coding System Specification*s. 2001 Census of Population. Unpublished document. Ottawa.

Statistics Canada. 1994. *Place of Work*. 1991 Census Technical Reports. Reference Products Series. Catalogue no. 92-343-XPE. Ottawa.