



## **Microdata User Guide**

# **LONGITUDINAL SURVEY OF IMMIGRANTS TO CANADA**

## **Wave 1**



## Table of Contents

|            |  |           |
|------------|--|-----------|
| <b>1.0</b> | <b>Introduction .....</b>                                    | <b>1</b>  |
| <b>2.0</b> | <b>Background .....</b>                                      | <b>3</b>  |
| <b>3.0</b> | <b>Objectives .....</b>                                      | <b>5</b>  |
| <b>4.0</b> | <b>Concepts and Definitions .....</b>                        | <b>7</b>  |
| <b>5.0</b> | <b>Survey Methodology .....</b>                              | <b>11</b> |
| 5.1        | Target Population .....                                      | 11        |
| 5.2        | Survey Frame .....   | 11        |
| 5.3        | Survey Design .....  | 11        |
| 5.3.1      | Longitudinal Sample .....                                    | 12        |
| 5.3.2      | Stratification .....   | 12        |
| 5.4        | Sample Selection and Sample Size .....                       | 12        |
| <b>6.0</b> | <b>Data Collection .....</b>                                 | <b>15</b> |
| 6.1        | Computer-assisted Interviewing .....                         | 15        |
| 6.2        | Collection .....   | 16        |
| <b>7.0</b> | <b>Data Processing .....</b>                                 | <b>19</b> |
| 7.1        | Initial Application Editing .....                            | 19        |
| 7.2        | Minimum Completion Requirements .....                        | 20        |
| 7.3        | Coding .....   | 21        |
| 7.3.1      | Coding of Open-ended Questions .....                         | 21        |
| 7.3.2      | Coding of Census Type Variables .....                        | 21        |
| 7.3.3      | Coding of “Other – Specify” Answers .....                    | 22        |
| 7.4        | Head Office Editing .....                                    | 22        |
| 7.5        | Data Model Structure .....                                   | 23        |
| 7.5.1      | File Structure .....   | 24        |
| 7.6        | Constructing Custom Datasets .....                           | 26        |
| 7.7        | Consistency Edit .....                                       | 27        |
| 7.8        | Naming Convention .....                                      | 28        |
| 7.9        | Derived Variables .....                                      | 30        |
| <b>8.0</b> | <b>Non-response .....</b>                                    | <b>31</b> |
| 8.1        | Definitions of Responding Unit .....                         | 31        |
| <b>9.0</b> | <b>Imputation .....</b>                                      | <b>35</b> |
| 9.1        | Mass Imputation .....  | 35        |
| 9.2        | Imputation for Income and Other Quantitative Variables ..... | 36        |

|                     |   |           |
|---------------------|---|-----------|
| <b>10.0</b>         | <b>Treatment of Total Non-response and Weighting .....</b>                            | <b>39</b> |
| 10.1                | Representativity of the Weights .....   | 39        |
| 10.2                | Overview of the Weight Adjustments .....  | 39        |
| 10.3                | Longitudinal Weighting for Responding Immigrants .....                                | 41        |
| 10.3.1              | Design Weight .....   | 42        |
| 10.3.2              | Non-response and Unresolved Adjustment Weights .....                                  | 43        |
| 10.3.3              | Post-stratification.....  | 44        |
| 10.3.4              | Adjustment Classes: Homogeneous Groups .....  | 46        |
| <b>11.0</b>         | <b>Data Quality and Coverage .....</b>  | <b>47</b> |
| 11.1                | Sampling Errors .....   | 47        |
| 11.2                | Non-sampling Errors .....   | 48        |
| 11.3                | Non-response and Unresolved Cases .....   | 48        |
| 11.4                | Coverage.....   | 49        |
| <b>12.0</b>         | <b>Guidelines for Tabulation, Analysis and Release.....</b>                           | <b>51</b> |
| 12.1                | Rounding Guidelines.....  | 51        |
| 12.2                | Sample Weighting Guidelines for Tabulation.....                                       | 52        |
| 12.3                | Definitions of Types of Estimates: Categorical and Quantitative.....                  | 52        |
| 12.3.1              | Tabulation of Categorical Estimates .....   | 53        |
| 12.3.2              | Tabulation of Quantitative Estimates .....  | 53        |
| 12.4                | Guidelines for Statistical Analysis.....  | 54        |
| 12.5                | Coefficient of Variation Release Guidelines .....                                     | 55        |
| <b>13.0</b>         | <b>Variance Calculation.....</b>  | <b>57</b> |
| 13.1                | Importance of the Variance.....   | 57        |
| 13.2                | SAS and STATA Macros to Calculate the Variance Using the Bootstrap Weights.....       | 57        |
| 13.3                | Excel Based Coefficient of Variation Extraction Module.....                           | 58        |
| 13.3.1              | Statistics Canada Quality Standards .....   | 58        |
| 13.4                | How to Derive the Coefficient of Variation for Categorical Estimates .....            | 60        |
| 13.5                | How to Use the Coefficient of Variation to Obtain Confidence Limits .....             | 60        |
| 13.6                | Hypothesis Testing (t-test) .....   | 62        |
| 13.7                | Coefficients of Variation for Quantitative Estimates .....                            | 62        |
| 13.8                | Approximate Quality Release Cut-offs .....  | 62        |
| <b>14.0</b>         | <b>Record Layout with Univariate Frequencies .....</b>                                | <b>65</b> |
| <b>Appendix I</b>   | <b>Industry Codes – North American Industry Classification System (NAICS) 1997 ..</b> | <b>67</b> |
| <b>Appendix II</b>  | <b>Occupation Codes – Standard Occupational Classification (SOC) 1991.....</b>        | <b>73</b> |
| <b>Appendix III</b> | <b>Country Codes – 2001 Census .....</b>  | <b>83</b> |
| <b>Appendix IV</b>  | <b>Country of Citizenship Codes – 2001 Census .....</b>                               | <b>87</b> |
| <b>Appendix V</b>   | <b>Intended Occupation Codes – National Occupational Classification (NOC) 2001 ..</b> | <b>91</b> |

## **1.0 Introduction**

The Longitudinal Survey of Immigrants to Canada (LSIC), conducted jointly by Statistics Canada and Citizenship and Immigration Canada under the Policy Research Initiative, is a comprehensive survey designed to study the process by which new immigrants adapt to Canadian society.

The first wave of collection for the LSIC was conducted between April 2001 and March 2002 by Statistics Canada. This manual has been produced to facilitate the manipulation of the microdata file of the survey results and to document data quality and other analytical issues regarding the LSIC.

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

### Statistics Canada

Client Services  
Special Surveys Division, Statistics Canada  
2500 - R, Main Building  
Tunney's Pasture  
Ottawa, Ontario  
K1A 0T6

Telephone: (613) 951-3321 or call toll free: 1 800 461-9050  
Fax: (613) 951-4527  
Email: [ssd@statcan.ca](mailto:ssd@statcan.ca)



## **2.0 Background**

The Longitudinal Survey of Immigrants to Canada is a comprehensive survey designed to study the process by which new immigrants adapt to or integrate into Canadian society, including the timing of stages in the integration process, as well as the factors which influence integration.

As part of adapting to life in Canada, many immigrants face challenges such as finding suitable accommodation, learning or becoming more fluent in one or both of Canada's official languages, participating in the labour market or accessing education and training opportunities.

The results of this survey will provide indicators of how immigrants are meeting these and other challenges, and what resources are most helpful to their settlement in Canada. The survey also examines how the socio-economic characteristics of immigrants influence the process by which they integrate into Canadian society.

The topics covered by the survey include language proficiency, housing, education, foreign credentials recognition, employment, health, values and attitudes, the development and use of social networks, income, and impressions about life in Canada. The questions address respondents' situation before coming to Canada and their current situation since their arrival. The questionnaire also covers the respondent's impressions and perceptions of his or her current settlement situation.

With the exception of the module on income - in which the person most knowledgeable about the subject is asked to respond - no interview may be conducted by proxy. Some modules also contain questions on members of the household, such as questions on employment, income or demographic characteristics, and on children, such as education questions. However, the unit of analysis for the survey is the selected immigrant, referred to as the longitudinal respondent (LR).





### **3.0 Objectives**

There exists a growing need for information on recent immigrants to Canada. While full integration may take several generations to achieve, the Longitudinal Survey of Immigrants to Canada is designed to examine the process during the critical first four years of settlement, a time when newcomers establish economic, social and cultural ties to Canadian society. To this end, the objectives of the survey are two-fold:

- to study how new immigrants adjust to life in Canada over time; and,
- to provide information on the factors that can facilitate or hinder this adjustment.



## 4.0 Concepts and Definitions

There are many variables and concepts that are critical to the analysis of Longitudinal Survey of Immigrants to Canada (LSIC) data. The following is an explanation of the key concepts in the LSIC. Derived variables are those that are not asked directly to the respondents but are calculated using information they have provided.

The unit of analysis for the LSIC is the longitudinal respondent. See Section 5.3 for information on the survey design.

**Census family:** Refers to a married couple (with or without children of either or both spouses), a couple living common-law (with or without children of either or both partners) or a lone parent of any marital status, with at least one child living in the same dwelling. A couple living common-law may be of opposite or same sex. “Children” in a census family include grandchildren living with their grandparent(s) but with no parents present. A census family is also referred to as an “immediate family” in the survey.

**Citizenship:** The status of being a citizen, either native-born or naturalized, sharing equally in the rights, privileges and responsibilities belonging to each individual.

**Common-law partner:** The person who, though not legally married to the respondent, is living with the respondent as his/her spouse. This partner may be of the same or opposite sex.

**Credentials:** Includes any formal education higher than a high school diploma, such as professional or technical qualifications and any other degrees, diplomas or certificates received from outside Canada.

**Fully Accepted:** The employer/institution recognizes a certificate, diploma, etc. as being legitimate within determined standards.

**Partially Accepted:** The employer/institution partially recognizes a certificate, diploma, etc. as being legitimate within determined standards.

**Not Accepted:** Certificate, diploma, etc. is not recognized as being legitimate within determined standards.

**Discrimination:** The unfavourable treatment of individuals on the basis of their personal characteristics, which may include race or skin colour, ethnicity or culture, language or accent, religion etc.

**Economic family:** Refers to a group of two or more persons who live in the same dwelling and are related to each other by blood, marriage, common-law or adoption.

**Ethnic or Cultural Group:** A group of individuals having a distinct culture in common. The term “ethnic or cultural group” implies that values, norms, behaviour and language, *not necessarily physical appearance*, are the important distinguishing characteristics.

**FOSS:** The acronym stands for “Field Operations Support System” and is an administrative database maintained by Citizenship and Immigration Canada. The FOSS was used as the sample frame of the survey.

**Full-time Employment:** Persons who usually work 30 hours or more per week at their main or only job.

**Host Program:** This program matches newcomers with a volunteer who is familiar with Canadian ways. Someone who can teach newcomers about available services, make contacts, help with employment, housing, etc. This program is intended to facilitate the integration process of newcomers.

**Immigrant Categories:**

**Economic Class:** Immigrants selected for their skills or other assets that will contribute to the Canadian economy (includes skilled workers, investors, entrepreneurs, and self-employed persons).

**Family class:** Immigrants sponsored by close relatives or family members already living in Canada.

**Independent immigrants:** Immigrants who qualify for certain types of jobs or have other important assets to bring to Canada. They apply on their own or have more distant relatives living in Canada.

**Refugees:** Persons seeking protection in Canada.

**Immigrant Settlement and Adaptation Program (ISAP):** A program in which funds are provided to deliver direct and essential services to newcomers. These services include reception and orientation, translation and interpretation, referral to community resources, para-professional counselling, general information and employment-related services.

**Immigration Consultant:** A professional who gives advice or services related to immigration issues.

**Immigrant or Refugee Serving Agency:** An organized body catering to the needs of immigrants or refugees.

**Immigrating Unit:** Refers to a group of people who applied to come to Canada under the same visa form and, for the purpose of the survey, who arrived either with the longitudinal respondent or three months before or after the longitudinal respondent.

**Immigration Officer:** A Canadian official who processes the authorization of immigrants upon arrival in Canada.

**Integration:** The process through which newcomers participate in and shape Canadian community.

**Joiner:** A person who was not a member of the longitudinal respondent's (LR) immigrating unit, but who was living in the same household at the time of the interview. This includes people who were already living in Canada when the LR arrived.

**Labour Force Status:**

**Employed:** Were employed or self-employed at the time at the time of the interview.

**Unemployed:** Persons who have not worked since they came to Canada but have looked for work at some point between their arrival and the Wave 1 interview. It also includes persons who had a job between their arrival and the Wave 1 interview, but who were not working at the time of the interview.

**Not in the Labour Force:** Persons who have neither worked nor looked for work since their arrival.

**Longitudinal respondent (LR):** The longitudinal respondent is the person selected to answer the LSIC questions at each of the three waves.

**Mover:** A person who was a member of the longitudinal respondent's immigrating unit, but who was not living in the same household at the time of the interview.

**Part-time Employment:** Part-time employment consists of persons who usually work less than 30 hours per week at their main or only job.

**PMK:** Person Most Knowledgeable about a specific subject. In the LSIC, the only questions asked of the PMK were questions on family income within the Income Module. If the PMK is not available, the questions are asked to the LR.

**Population Group:** Refers to the population group to which the respondent belongs. It includes visible minorities (see definition below) as well as Aboriginal peoples, Caucasian in race or white in colour.

**Reference period:** Period of time between the landing date and the date of the interview.

**Sponsor:** Canadian Citizens, or permanent residents aged 19 or over, living in Canada that commit to provide the sponsored immigrant with basic assistance in the form of accommodation, clothing, food and settlement assistance for a specific period of time.

**Visible Minority:** Refers to "persons, other than Aboriginal peoples, who are non-Caucasian in race or non-white in colour".



## 5.0 Survey Methodology

The Longitudinal Survey of Immigrants to Canada (LSIC) was designed to collect longitudinal data on immigrants in order to better understand the process by which new immigrants adapt to Canadian society. This survey will provide information on factors which facilitate or impede their adaptation and the ways that they contribute to Canadian society and the Canadian economy.

To produce reliable estimates, a representative sample of approximately 20,300 new immigrants to Canada was selected. This chapter describes the sample selection method and the sample size.

### 5.1 Target Population

The target population for the survey consists of immigrants who meet all of the following criteria:

- arrived in Canada between October 1, 2000 and September 30, 2001;
- were age 15 or older at the time of landing;
- landed from abroad, must have applied through a Canadian Mission Abroad.

Individuals who applied and landed from *within* Canada are excluded from the survey. These people may have been in Canada for a considerable length of time before officially "landing" and would therefore likely demonstrate quite different integration characteristics to those recently arrived in Canada. Refugees claiming asylum from within Canada are also excluded from the scope of the survey.

The target population accounts for approximately 164,200 of the 250,000 persons admitted to Canada during this period. Coverage of the survey included all Census Metropolitan Areas and non-remote Census Agglomerations.

### 5.2 Survey Frame

The target population is represented by the survey frame from which the sample is selected. The sampling frame for the LSIC is an administrative database of all landed immigrants to Canada that comes from Citizenship and Immigration Canada. The database, known as the FOSS (Field Operation Support System), includes various characteristics of each immigrant that can be used for survey design purposes, such as: name, age, sex, mother tongue, country of origin, knowledge of English and/or French, class of immigrant, date of landing, and intended province of destination in Canada.

Detailed information from the FOSS on each immigrant landing during the survey reference period, i.e., October 2000 to September 2001, is provided to Statistics Canada two months after the reference month. This allows for the sampling frame to be built month after month by simply adding new monthly landings.

### 5.3 Survey Design

The survey was designed based on probability sample theory. The sample is created using a two-stage stratified sampling method. The first stage involves the selection of Immigrating Units (IU) using a probability proportional to size (PPS) method. The second stage involves the selection of one IU member within each selected IU. The selected member of the IU is called the longitudinal respondent (LR) and will be contacted to participate in the survey. Only

the LR will be followed throughout the survey and no interviews will be conducted with other members of the IU or the LR's household.

### **5.3.1 Longitudinal Sample**

The survey involves a longitudinal design with immigrants being interviewed at three different times: at six months, two years, and four years after landing in Canada. The sample design has been developed using a "funnel-shaped" approach, therefore only immigrants that respond to the Wave 1 interview will be traced for the Wave 2 interview and only those that respond to the Wave 2 interview will be traced for the Wave 3 interview.

The funnel-shape approach was chosen because of the nature of the survey and its analytical objectives. The survey collects information on perceptions, values and attitudes at specific points in time, in order to assess the immigrant's integration during their initial years in Canada. If data were collected only once (i.e., during the fourth year in Canada), significant recall and response errors could be encountered. Furthermore, to facilitate a complete study of the immigrant's adaptation, the full range of longitudinal data must be obtained from each longitudinal respondent.

### **5.3.2 Stratification**

The first stratification variable used is the month of landing in Canada; there are 12 cohorts of immigrants, i.e. one for each reference month. Within each month, two other stratification variables are used: the intended province of destination as stated by the immigrant and the class of immigrant.

Provinces are grouped into five categories: Québec, Ontario, Alberta, British Columbia and the remaining provinces (the territories are excluded).

There are six categories for classes of immigrant: family class, economic-skilled, economic-business, government-sponsored refugees, other refugees and other immigrants. Since a stratum is created by the intersection of the above categories, there are 30 strata for each cohort of immigrants for a total of 360 strata.

## **5.4 Sample Selection and Sample Size**

The sample can be divided into two components - the core and the additional samples. The core sample represents the target population, while the additional samples target specific sub-populations. These specific sub-populations were determined by analysing the expected sample allocation at Wave 3 and also by various requirements of federal and provincial government departments. The following subgroups have been over-sampled:

- 1) government sponsored refugees;
- 2) refugees other than government sponsored;
- 3) contractor and investor immigrants (economic-business);
- 4) family immigrants in British Columbia;
- 5) overall immigrants in Alberta; and
- 6) economic immigrants in Québec (economic-skilled and economic-business).

The stratification allows for control over the sample sizes for each of the additional samples' subgroups.



Tables 5.1, 5.2 and 5.3 provide a breakdown of the population based on the sampling frame and of the sample allocation for the core and additional samples expected at Wave 3.

For the core sample, it was determined that 5,000 completed interviews at Wave 3 would produce reliable estimates<sup>1</sup> at the national level, the provincial level where the in-flow of immigrants is the most significant (Québec, Ontario and British Columbia) and for certain classes of immigrants (family and economic classes). Also it would be possible to obtain reliable estimates for other combinations of variables as long as a minimum number requirement is met. After taking into account the requirements for the additional samples outlined above, the minimum number of completed interviews at Wave 3 is expected to be 5,755 immigrants.

The determination of the sample size for Wave 1 is based several sample attrition hypotheses applied to the Wave 3 minimum sample size requirement. Examining results from various longitudinal studies of the Canadian population, a combined response rate (resolved cases and respondent) of 75% was estimated for Waves 2 and 3 - i.e. 75% of Wave 1 respondents would respond in Wave 2 and 75% of Wave 2 respondents in Wave 3. In addition, various sources were used to estimate a combined return rate, i.e. after tracing and classification as in-source or out-of-scope. Results from the pilot study and a coverage study on language<sup>2</sup> were used as a source of information. Finally, Statistics Canada's Reverse Record Check Study (RRC)<sup>3</sup> was used to estimate the expected tracing rates or rates of resolved cases.

The sample is selected over a 12-month period. A sample allocation proportional to the number of immigrants in each month of landing as well as between strata within a month minimizes the total sampling variance. However, for operational reasons, such as maintaining a constant number of interviews in each month of collection, an equal allocation has been performed between the months of landing, even though immigration shows a seasonal pattern. Table 5.4 presents the final sample size at Wave 1.

**Table 5.1 Total Number of Immigrants, 15 Years and Over, by Province and Class of Immigrant, October 2000 to September 2001**

| Province         | Family        | Economic-skilled | Economic-business | Government-refugee | Other Refugee | Other        | Total          |
|------------------|---------------|------------------|-------------------|--------------------|---------------|--------------|----------------|
| Québec           | 4,680         | 12,694           | 2,977             | 1,238              | 887           | 78           | 22,554         |
| Ontario          | 26,579        | 64,346           | 3,591             | 2,054              | 2,123         | 216          | 98,909         |
| Alberta          | 3,250         | 5,651            | 444               | 623                | 307           | 125          | 10,400         |
| British Columbia | 8,532         | 15,048           | 2,489             | 679                | 317           | 235          | 27,300         |
| Other provinces  | 1,199         | 2,074            | 494               | 948                | 427           | 707          | 5,849          |
| <b>Canada</b>    | <b>44,240</b> | <b>99,813</b>    | <b>9,995</b>      | <b>5,542</b>       | <b>4,061</b>  | <b>1,361</b> | <b>165,012</b> |

1 By reliable estimates we mean being able to estimate a minimal proportion of 10% with a coefficient of variation of 16.5%. A cell size of 450 responding units is necessary to meet this requirement.

2 Given operational constraints, namely the requirement and associated costs to translate the questionnaire in several languages, a study has been performed to identify the population coverage according to languages. It has been determined that the translation could be performed in 13 languages other than English or French, and that it would allow a national coverage of around 93% of landed immigrants.

3 The 1996 RRC study was undertaken following the 1996 Census to estimate Census under-coverage. This study makes use of an immigrant frame that covers immigrants who landed in Canada between the 1991 and 1996 censuses.

**Table 5.2 Expected Allocation of Respondents in Wave 3 - Core Sample**

| Province         | Family       | Economic-skilled | Economic-business | Government-refugee | Other Refugee | Other     | Total        |
|------------------|--------------|------------------|-------------------|--------------------|---------------|-----------|--------------|
| Québec           | 151          | 312              | 94                | 46                 | 25            | 5         | 633          |
| Ontario          | 810          | 1,870            | 125               | 46                 | 72            | 12        | 2,935        |
| Alberta          | 104          | 156              | 21                | 13                 | 6             | 4         | 304          |
| British Columbia | 287          | 505              | 108               | 12                 | 10            | 10        | 932          |
| Other provinces  | 41           | 74               | 19                | 25                 | 12            | 25        | 196          |
| <b>Canada</b>    | <b>1,393</b> | <b>2,917</b>     | <b>367</b>        | <b>142</b>         | <b>125</b>    | <b>56</b> | <b>5,000</b> |

**Table 5.3 Expected Allocation of Respondents in Wave 3 - Core and Additional Samples**

| Province         | Family       | Economic-skilled | Economic-business | Government-refugee | Other Refugee | Other     | Total        |
|------------------|--------------|------------------|-------------------|--------------------|---------------|-----------|--------------|
| Québec           | 151          | 346              | 125               | 146                | 28            | 5         | 801          |
| Ontario          | 810          | 1,870            | 153               | 146                | 79            | 12        | 3,070        |
| Alberta          | 154          | 231              | 36                | 47                 | 9             | 6         | 483          |
| British Columbia | 450          | 505              | 132               | 38                 | 11            | 10        | 1,146        |
| Other provinces  | 41           | 74               | 23                | 79                 | 13            | 25        | 255          |
| <b>Canada</b>    | <b>1,606</b> | <b>3,026</b>     | <b>469</b>        | <b>456</b>         | <b>140</b>    | <b>58</b> | <b>5,755</b> |

**Table 5.4 Final Sample Allocation at Wave 1**

| Province         | Family       | Economic-skilled | Economic-business | Government-refugee | Other Refugee | Other      | Total         |
|------------------|--------------|------------------|-------------------|--------------------|---------------|------------|---------------|
| Québec           | 463          | 1,230            | 437               | 377                | 111           | 12         | 2,630         |
| Ontario          | 2,653        | 6,920            | 599               | 630                | 269           | 23         | 11,094        |
| Alberta          | 531          | 928              | 93                | 234                | 59            | 22         | 1,867         |
| British Columbia | 1,560        | 1,634            | 423               | 210                | 40            | 26         | 3,893         |
| Other provinces  | 121          | 225              | 81                | 293                | 46            | 72         | 838           |
| <b>Canada</b>    | <b>5,328</b> | <b>10,937</b>    | <b>1,633</b>      | <b>1,744</b>       | <b>525</b>    | <b>155</b> | <b>20,322</b> |

## 6.0 Data Collection

### 6.1 Computer-assisted Interviewing

Data collection for the Longitudinal Survey of Immigrants to Canada (LSIC) relied heavily on computer-assisted interviewing (CAI) technology. The use of CAI technology allows for high quality collection of complex population-specific content sections. For example, the system facilitates the collection of the relationships of all household members to each other (i.e., the relationship grid). This wealth of information will enable a detailed analysis of family structures, an important concept for analysis. This type of collection would be very difficult to implement in a paper and pencil environment.

The CAI system has two main parts:

#### 1) Case Management

The Case Management system controls the case assignment and data transmission for the survey. For this survey, a case refers to an individual selected for the LSIC sample. The Case Management system also automatically records management information for each contact (or attempted contact) with respondents and provides reports for the management of the collection process.

The Case Management system routes the questionnaire applications and sample file from headquarters to the regional offices and from the regional offices to the interviewer laptops. The returning data takes the reverse route. To assure confidentiality, all data is encrypted before transmission. The data are unencrypted only once they are on a separate secure computer with no external access.

#### 2) Survey-specific Components

##### Locating Respondents

The Wave 1 LSIC target population consists of immigrants who have been in Canada for only six months. During the first few months after landing, new immigrants are a very mobile population. Because of this mobility, and the short period of time living in Canada prior to the interview, the resources to trace respondents are limited.

To help locate respondents a contact questionnaire was designed to request the immigrant's address in Canada (if known) as well as the address of a contact person in Canada. The form also contained a consent statement asking the respondent to grant Statistics Canada permission to access information held by other federal/provincial organizations, such as a provincial health department, for tracing purposes only. The form was enclosed in the materials provided to immigrants when they receive their landing visa from a Canadian Mission Abroad.

Access to additional tracing information was only granted with consent from the potential respondent. This consent allowed Statistics Canada to obtain access to tracing-related information from health card records of all provincial health departments, with the exception of Nova Scotia. This source of information was considered to be the most current address information for the respondents.

### **Longitudinal Respondent Contact**

The first contact was established with these households using the address and telephone number provided on the sample file by Head Office. The interviewer confirmed that the respondent lived at that address. Once it was established that the interviewer was speaking to the correct person further steps were taken to ensure it was the proper respondent. Verification of respondent was done in two ways: matching of birth date and landing date.

Once the interviewer verified they had the correct respondent, the interviewer confirmed or updated the contact information (mailing and residence address, telephone number), as well as the list of household members. An appointment was then made to continue the interview in person.

If the interviewer was unable to locate the respondent the case was transferred to the trace folder, which was then transferred to a designated tracing team in the regional offices, for further follow up.

### **Tracing Respondents**

Within the regional offices, designated tracing teams followed up with further tracing sources to try and locate the respondent. Electronic phone books were the only effective public source used for tracing. The following sources of information were used for tracing the selected respondents of the first wave:

- administrative files from Citizenship and Immigration Canada;
- survey contact questionnaires;
- addresses from provincial health cards (where an agreement with the province was reached and consent was given by the respondent); and
- electronic phone books (Québec, Ontario and British Columbia).

### **Person Most Knowledgeable**

The LSIC is a non-proxy interview, with the exception of one section of questions – family income questions within the Income Module. The person most knowledgeable (PMK) about the family's income was asked to respond to these questions only.

## **6.2 Collection**

### **Collection Period**

The survey uses a longitudinal design, meaning the same selected respondent will be interviewed at three different points in time. The first of the three interviews is conducted six months after the respondent arrives in Canada; since it is desirable to assess their integration as soon as possible after they arrive. The second interview takes place two years after their arrival, and the final interview is conducted four years after their arrival.

To adequately represent the different immigration patterns in Canada over a one-year period, the sample is made up of 12 cohorts, consisting of 12 independent monthly samples selected over a period of 12 consecutive months. Collection is done separately for each month of landing. For example, immigrants who arrived in October 2000 will be interviewed in April 2001, October 2002 and October 2004, while those who arrived in September 2001 will be interviewed in March 2002, November 2003 and September 2005. Each monthly sample can remain in the field for up to three months.

| <b>Landing date: October 2000 to September 2001</b> |                         |                       |
|---|-------------------------|-----------------------|
| <b>Wave</b>   | <b>Collection Start</b> | <b>Collection End</b> |
| 1   | April 2001              | March 2002            |
| 2   | December 2002           | November 2003         |
| 3   | October 2004            | September 2005        |

**Wave 1 Collection**

Collection for the first wave of the survey occurred between April 2001 and May 2002. The majority of interviews, 68%, were conducted face-to-face, while the remaining 32% were conducted over the telephone for various reasons (location of interview, specific language requirements, etc.).

Interviews were conducted in one of the 15 languages most frequently spoken by the target population: English, French, Chinese (Mandarin, Cantonese), Punjabi, Farsi/Dari (one language), Arabic, Spanish, Russian, Serbo-Croatian, Urdu, Korean, Tamil, Tagalog, and Gujarati. The 15 languages selected cover approximately 93% of the immigrant population in Canada.

**Interview Length for Household Collection**

On average the Wave 1 interview lasted approximately 90 minutes. Fifteen minutes were devoted to the Entry and Exit components and the remaining 75 minutes to the survey.



## 7.0 Data Processing

The main output of the Longitudinal Survey of Immigrants to Canada (LSIC) is a "clean" master data file. This chapter presents a brief summary of some of the processing steps involved in producing this file.

### 7.1 Initial Application Editing

#### Computer Generated Edits

As discussed earlier, all of the information for the sampled individuals was collected in a face-to-face, or telephone interview when a face-to-face was not possible, using a computer-assisted personal interviewing (CAPI) application. As such, it was possible to build various edits and checks into the questionnaire in order to ensure that a high quality of the information was collected. Below are specific examples of the types of edits used in the LSIC computer-assisted interviewing (CAI) application:

#### Flow Pattern Edits

All flow patterns were automatically built into the CAI system. For example, for questions pertaining to a spouse/partner or child, the CAI system would automatically refer to the relationship information of all household members collected in the Entry Module to determine whether the longitudinal respondent (LR) had a spouse/partner or child living with them. If a spouse/partner or child was present, the CAI system continued with the specific questions related to them. If not, the CAI system automatically skipped these questions.

#### General Consistency Edits

Some consistency edits were included as part of the CAI system, and interviewers were able to "slide back" to previous questions to correct for inconsistencies. Instructions were displayed to interviewers for handling or correcting problems such as incomplete or incorrect data. For example, in the Language Module, if the respondent indicated that English was the language he/she most often spoke at home, the respondent could then not answer that they do not speak English to a following question. If this happened, an edit screen popped up and the interviewer had to change one of the answers.

#### Range Edits in Numeric Fields

Range edits were also built into the CAI system for questions asking for numeric values. If numbers entered were outside the range, the system generated a pop-up window which stated the error and instructed the interviewer to make corrections to the appropriate question. For example, in the collection of the Employment Details sub-module, the number of hours worked per week was set to a maximum of 168 hours (the number of hours in a week). If the respondent indicated that he/she worked more than 168 hours a week, the range edit was triggered.

## **7.2 Minimum Completion Requirements**

One of the first steps in the LSIC processing was to define the requirements for a responding immigrant.

### **No Information Collected**

In some cases, no LSIC information was collected for a sampled individual. This happened when an interviewer was unable to trace a selected immigrant or was not able to make contact for the entire collection period. In other cases, the individual refused to participate in the survey, was away for the duration of the collection period or language barriers (an individual who did not speak one of the 15 survey languages) prevented an interview from taking place.

For cases where no information at all was collected for an immigrant, the individual was dropped from the LSIC file and the sampling weights for responding immigrants were inflated to account for these "dropped" immigrants.

### **Partial Information**

In other cases, it was possible to carry out some of the interview, but a complete interview was not obtained for a variety of reasons. Some respondents were willing to give only a certain amount of time to the completion of the survey. In some cases an interviewer completed a portion of the survey with the respondent and made an appointment to continue at another time but was unable to re-contact the respondent.

### **Criteria for Partial Response**

It was necessary to come up with criteria for deciding what to do with these "partial" interviews. It was decided to consider as partial response any immigrant who answered the entire Entry and Background Modules (the first two modules). This was the information required in order to establish imputation strategies to complete the remainder of the questionnaire. The partial responding immigrants were then maintained in the responding sample.

### **Missing Components and Mass Imputation**

For the partial responding individuals, all variables from the missing components were set to not stated or imputed, with the exception of two modules - "Values and Attitudes" and "Perceptions of Settlement". The questions in these two modules asked about the LR's opinions and perceptions, which vary too much to establish a solid mass imputation strategy.

In total, mass imputation to complete partial responses was performed on 5% of all responding records. For more information on imputation, see Chapter 9.0.

### **Total Responding Records**

In total, 12,040 longitudinal respondent records were determined to be complete enough to be kept in the final file.

These immigrants had resided in a total of 6,411 places prior to their current place of residence (collected in the Where Lived sub-module). They had taken a total of 6,375 courses or training sessions. They reported 13,149 credentials of various kinds. They had a total of 7,554 jobs or businesses since landing in Canada. Moreover, there were 289 people who arrived in Canada with the LR's immigrating unit, but were not living with them at the time of the interview (movers).



## **7.3 Coding**

In Wave 1, coding was an important task of processing. Three different levels of coding were done: open-ended questions, census type of questions, and text recorded in the “Other - Specify” fields. Given the number of new categories that were added to questions during the coding step, coding was done before the pre-edit step, in order to minimize adjustments to the pre-edit and flow edits stages.

### **7.3.1 Coding of Open-ended Questions**

A few data items on the LSIC questionnaire were recorded by interviewers in an open-ended format. For example, in the Employment Module, a LR who had worked since they arrived in Canada was asked a series of open-ended questions about each job they have held:

- What kind of business, industry or service is/was it?
- What kind of work do/did you do in this job?
- In this job what are/were your most important duties?

In the Perceptions of Settlement Module, the last two questions were:

- What is the single most useful thing that was done to help you settle in Canada?
- What is the single most useful thing that could have been done to help you settle in Canada?

#### **How they are recorded**

The interviewer recorded, in words, the answer provided by the respondent to these questions. At Head Office, these written descriptions were converted into codes (e.g., industry or occupation) to make the data comparable. These codes only have one valid description in English and in French to ensure for consistency and comparability.

#### **How they are coded**

The open-ended questions were coded using various standard classifications. Occupation questions were coded using the 1991 Standard Occupational Classification codes (SOC) and the industry questions were coded using the 1997 North American Industry Classification System (NAICS).

Variables asking about the major field of study in the Education Module were coded using a code set called “Major Field of Study (MFS)”. Survey-specific code sets were developed in order to code questions such as the two examples from the Perceptions of Settlement Module.

### **7.3.2 Coding of Census Type Variables**

A few of the LSIC questions were also asked in the 2001 Census. These include questions on country of birth, country of citizenship, language, religion, ethnic group and visible minority.

#### **How they are recorded**

For most of these questions, a pick-list was included in the questionnaire. In many cases, the “Other - Specify” category was chosen by interviewers and a text entry was recorded.

### **How they are coded**

At Head Office, each of these questions were coded using the corresponding Census code set in order to match the 2001 Census data dictionary. The groupings resulting from the coding were then perfectly comparable with Census data released recently.

### **7.3.3 Coding of “Other – Specify” Answers**

In the LSIC Wave 1 questionnaire, several questions included an “Other - Specify” category, which allowed the interviewers to enter a text entry for an answer they could not find in the pick-list. These categories were looked at and in many cases new categories were added to the original questions. A category was added to the question if it accounted for 5% of all answers. The new categories were also added to the Wave 2 questionnaire to ensure consistency between the two waves.

### **How they are coded**

Variable-specific code sets were created for each variable containing an “Other - Specify” response category, and codes were assigned from that list. Categories entered in the “Other - Specify” fields might also have been coded up to an existing category when the response element was from the same concept.

## **7.4 Head Office Editing**

### **Pre-edits**

Before proceeding with the pre-edits, databases were created for the main section of the questionnaire, for the information collected on the LR’s household as well as for each of the roster files.

The pre-edit step is the first of various stages of editing. The purpose of the pre-edit is to carry out basic formatting and preliminary editing.

As a first pre-edit step, “Mark all that apply” questions were de-strung and values converted to Yes (1) or No (2) responses. Non-response values from the CAI system were also recoded to standard non-response codes for refusals, don't know and not stated.

### **Converting non-response codes to standard codes**

#### **Don't know**

During a CAI interview, the respondent may not know the answer to a particular item. The CAI system has a specific function key to describe this situation.

In the LSIC files, the code used to indicate that the respondent did not know the answer to an item is "7". For a variable that is two digits long the code is "97", for a three-digit variable "997", etc.

#### **Refusals**

The respondent may choose to refuse to provide an answer for a particular item. The CAI system has a specific function key that the interviewer presses to indicate a refusal. This information is recorded for the specific item refused and transmitted back to Head Office.

In the LSIC files, an item which was refused is indicated by a code "8". For a variable that is two digits long the code is "98", for a three-digit variable "998", etc.

### **Not stated**

In some cases, as part of Head Office processing, the answer to an item has been set to “not stated”. The not stated code indicates that the question was not asked of the respondent. These codes were assigned for three main reasons:

- 1) As part of the CAI interview, the interviewer was permitted to enter a “refusal” or “don't know” code, as described above. When this happened the CAI system was often programmed to skip out of this particular section of the questionnaire. In the case of refusal, it was assumed that the line of questioning was sensitive and it was likely that the respondent would not answer any more questions on this particular topic area. In the case of a “don't know” it was assumed that the respondent was not well enough informed to answer further questions and it was not known if the subsequent questions were applicable. As part of the LSIC processing system, it was decided that all of these subsequent questions should be assigned a “not stated” code.
- 2) In some cases, sections or entire modules of the questionnaire were not started or they were started but ended prematurely. For example, there may have been some kind of interruption, or the respondent decided that he/she wished to terminate the interview. If there was enough information collected to consider the module as responded, the questions that were not answered would be coded to “valid skip”. If an entire module was not answered, mass imputation was performed - with the exception of the Values and Attitudes Module and the Perceptions of Settlement Module, where questions not answered remained as “not stated”.
- 3) The third situation in which “not stated” codes were used was as a result of consistency edits. When the relationship between groups of variables was checked for consistency, if there was an error, often one or more of the variables were set to “not stated”.

In the case of derived variables, if one or more of the input variables contained a “not stated”, then the derived variable was also set to “not stated”.

An item which was coded as “not stated” is indicated by a code "9". For a variable that is two digits long the code is "99", for a three-digit variable "999", etc.

### **Flow edits and assignment of valid skip codes**

As the last step of the pre-edits, the flow patterns for each of the files were processed and standard codes for “valid skips” were assigned (6, 96, and 996).

For example, for all questions where the LR did not have a spouse or common-law partner residing in the household, all "spouse" variables have been set to “valid skip”.

## **7.5 Data Model Structure**

After pre-edit and flow edits were completed, the data were split up into smaller databases, called entities, following a data model approach. This approach allows for better processing efficiency and provides a good storage structure for longitudinal data, minimizing the impact of changes made to the questionnaires between waves.

All questions measuring the same general concepts (education, housing, etc.) were regrouped into the same database. Most of the time, the questionnaire was already divided by theme and therefore, the contents of the processing files were very similar to the questionnaire modules. The following table presents the LSIC entities with a description of their collected or derived content.

**Table 7.1 List of the Data Model Entities and Their Contents**

| Acronym | Entity Name (concept)         | Unique Key Variables | Collected or Derived From:  |
|---------|-------------------------------|----------------------|---|
| CI      | Citizenship                   | HHLDID               | Background Module questions BG_Q06 to BG_Q09B, BG_Q16 and BG_Q17 and Values and Attitudes Module questions VAS_Q01 to VAS_Q04A  |
| CR      | List of education credentials | HHLDID, EDCID        | Education Credentials - sub-module of the Education Module  |
| ED      | Education                     | HHLDID               | Education Module  |
| EM      | Employment                    | HHLDID               | Employment Module   |
| GO      | Groups and organizations      | HHLDID               | Group Organizations - sub-module of the Social Network Module   |
| HH      | Household                     | HHLDID               | Entry Module (includes aggregated derived variables from the relationships questions)   |
| HL      | Health                        | HHLDID               | Health Module   |
| HS      | Housing                       | HHLDID               | Housing Module and Background Module questions BG_Q14 and BG_Q15  |
| IN      | Income                        | HHLDID               | Income Module   |
| JB      | List of jobs                  | HHLDID, JOBID        | Employment Details and Employment Roster - sub-modules of the Employment Module   |
| LR      | Longitudinal respondent       | HHLDID               | Entry Module and Background Module questions BG_Q01 to BG_Q05 and BG_Q18 to BG_Q20, in addition to some variables from a Citizenship and Immigration Canada administrative database |
| LS      | Language skills               | HHLDID               | Language Skills Module, excluding language test questions LS_Q11E to LS_Q16E and LS_Q11F to LS_Q16F   |
| PS      | Perceptions of settlement     | HHLDID               | Perceptions of Settlement Module  |
| SI      | Social interactions           | HHLDID               | Social Network Module   |
| ST      | List of studies               | HHLDID, STUDYID      | Education Details and Education Roster - sub-modules of the Education Module  |
| VA      | Values and attitudes          | HHLDID,              | Values and Attitudes Module, excluding questions VAS_Q01 to VAS_Q04A  |
| WL      | List of places lived          | HHLDID, WLID         | Where Lived - sub-module of the Housing Module  |

### 7.5.1 File Structure

The unit of analysis for the majority of entities is the longitudinal respondent (LR). Therefore, these entities have one record per longitudinal respondent. However, the units of analysis are unique to each roster file: HH is the respondent's household, CR is credentials, JB is jobs, ST is training or courses and WL is places where the LR lived. While the minimum number of records for the longitudinal respondent on the roster files is zero (except for HH where the LR is the unit of analysis), the maximum varies for each (CR = 12, JB = 7, ST = 7, WL = 5).

Note that when producing estimates, the final weights are only to be used for the LR's records. Estimates cannot be produced for the places lived records, the employment records or the education records. The estimates must only be associated to the LR. For more information on weighting see Chapter 10.0.

The LSIC files are available in two different structures:

- 1) Text files (ASCII format) - Data from each entity is included in one large text file, with the exception of the roster information. Each roster file (CR, JB, ST, WL) and household file (HH) has its own separate text file. SAS and SPSS syntax cards are available for the formatting of these files (names of these files end by SASE and SPSSE for English syntax cards and SASF and SPSSF for French syntax cards).

**Table 7.2 Text File Structures**

| <b>Filenames</b>        | <b>File Description</b>   |
|-------------------------|---|
| LSIC_W1_MAIN_Master.txt | This file includes the following entities: LR, CI, SI, GO, HS, HL, LS, ED, EM, VA, IN, PS           |
| LSIC_W1_HH_Master.txt   | This file includes information collected on the respondent's household.                             |
| LSIC_W1_CR_Master.txt   | This file includes variables collected in the Education Credentials sub-module.                     |
| LSIC_W1_JB_Master.txt   | This file includes variables collected in the Employment Roster and Employment Details sub-modules. |
| LSIC_W1_ST_Master.txt   | This file includes variables collected in the Education Roster and Education Details sub-modules.   |
| LSIC_W1_WL_Master.txt   | This file includes variables collected in the Where Lived sub-module.                               |

- 2) Entities – Each entity (described in Table 7.1) is contained in its own SAS file. These files, or specific variables from these files, can be merged using the Integrated Extraction Program (IXP), which is explained in Section 7.6 or other analytical software such as SAS, SPSS, STATA, etc.

All LSIC files include a unique identifier key referred to as the Household identifier (variable name HHLDID) that only pertains to the longitudinal respondent. All LSIC files can be merged using this key variable. Other identifiers are included on all roster files to make each record unique since there might have been zero to many events collected for one HHLDID. The list of the key variables for all files is shown in Table 7.1. For example, a unique STUDYID is associated with each course or training event reported in the ST entity.

## 7.6 Constructing Custom Datasets

Although data can be extracted and merged in many ways, the LSIC team has developed an easy-to-use data extraction and merging application. The application named IXP (Integrated Extraction Program) enables users to create custom files by pointing and clicking. The following should be considered when working with the IXP:

- 1) The IXP requires SAS input files and has the capacity to output files in ASCII, SAS and SPSS formats. The IXP allows the creation of SAS or SPSS syntax cards that will enable users to obtain fully formatted result files.
- 2) For each table selected, users must select all the variables they want to extract by pointing and clicking. An option enables the selection of all variables in the selected table at once.
- 3) If variables are selected from a roster file, the IXP will automatically show an option box to flatten the file. Flattening will create a single row of data for each LR by transposing the variables across. To illustrate this, a respondent may have reported between 0 and 12 credentials. In the example below, we have two respondents; the first reported one credential and the second, three. The user is interested in the three variables at the right side of the table. The user wants to match this file with some LR characteristics from the LR file.

Variables to select from the Credentials (CR) file:

| HHLIDID   | EDCID | cr1q002 | cr1q003 | cr1q004 |
|-----------|-------|---------|---------|---------|
| hhldid001 | 1     | 2       | 6       | 6       |
| hhldid002 | 1     | 1       | 2       | 1       |
| hhldid002 | 2     | 1       | 1       | 1       |
| hhldid002 | 3     | 1       | 2       | 1       |

For the credentials roster, the maximum number of possible entries per LR is 12. The IXP will automatically create a table with 12 columns for each selected variable from the roster (in the example above 3 x 12). This new table will then be put beside the LR variables to create the merged file.

When creating the 12 columns for the selected variables, the names are changed to distinguish between them. This is done by adding a letter to the end of the variable name. The letter will correspond to the roster id. For example, all variables related to EDCID = 1 will get an “a” at the end of the variable names, EDCID = 2 will be get a “b”, etc. If a respondent only reported two credentials, the remaining ten sets of variables will automatically be filled with missing values (represented by a period (.) in SAS). For the purpose of the following example of an output file, the maximum number of credentials is assumed to be 4 rather than 12. The EDCID is included to show the relationship between the roster id and the new variable name.

| HHLID     | lr1d005 | lr1q008 | lr1q009 | lr1g039 | EDCIDa | cr1q002a |
|-----------|---------|---------|---------|---------|--------|----------|
| hhldid001 | 73      | 2       | 1       | 7       | 1      | 2        |
| hhldid002 | 36      | 1       | 4       | 7       | 1      | 1        |

(Continued)

| cr1q003a | cr1q004a | EDCIDb | cr1q002b | cr1q003b | cr1q004b | EDCIDc |
|----------|----------|--------|----------|----------|----------|--------|
| 6        | 6        | .      | .        | .        | .        | .      |
| 2        | 1        | 2      | 1        | 1        | 1        | 3      |

(Continued)

| cr1q002c | cr1q003c | cr1q004c | EDCIDd | cr1q002d | cr1q003d | cr1q004d |
|----------|----------|----------|--------|----------|----------|----------|
| .        | .        | .        | .      | .        | .        | .        |
| 1        | 2        | 1        | .      | .        | .        | .        |

If you choose not to flatten the file the output would appear as shown below with the LR variables repeated for each record in the roster.

| HHLID     | lr1d005 | lr1q008 | lr1q009 | lr1g039 | EDCID | cr1q002 | cr1q003 | cr1q004 |
|-----------|---------|---------|---------|---------|-------|---------|---------|---------|
| hhldid001 | 73      | 2       | 1       | 7       | 1     | 2       | 6       | 6       |
| hhldid002 | 36      | 1       | 4       | 7       | 1     | 1       | 2       | 1       |
| hhldid002 | 36      | 1       | 4       | 7       | 2     | 1       | 1       | 1       |
| hhldid002 | 36      | 1       | 4       | 7       | 3     | 1       | 2       | 1       |

For more information on how to use the IXP, see the Integrated Extraction Program User's Guide.

## 7.7 Consistency Edit

### The Goal of Consistency Editing

Once the data model was built and data split into entities, consistency editing was carried out to verify the relationship between two or more variables.

### Relationship edits

For various reasons, relationship data collected in the Entry Module at times contained errors. The relationship edit step ensures a clean file and consistency in the relationships among members of the same household and between households.

For example some respondents whose spouses had children reported their relationship to them as "unrelated". In fact, according to the Census definitions, these people should have been step-parents, which is not a well-known concept for some recent immigrants to Canada. Similarly some foster parents reported being unrelated to a foster child, when they should have reported being foster parents.

### Examples of consistency edits

In the Language Skills Module, for the question "Other languages that LR can speak or read?" (ls1g080 to ls1g086) some respondents reported languages they had already mentioned in previous questions of the module. They should have excluded these. Those languages were then blanked out from the file. If these were the only languages reported for this variable, the

question “Are there any other languages, other than those you have already mentioned, that you can speak or read at least to some degree?” (ls1q079) was changed to “No”.

### **Outcome of the consistency edits**

When a consistency edit problem was found, there were two possible actions: either to change the response to what should have been the correct answer or change the response to “not stated”.

## **7.8 Naming Convention**

The LSIC microdata file documentation system has employed certain standards to label variable names and values. The intent is to make data interpretation more straight-forward for the user. All variable names are, at most, eight characters long (most are 7 long) so that these names can easily be used with analytical software packages such as SAS or SPSS.

### **Format for Variable Names**

- The **first two** characters are the acronym of the entity to which the item belongs. See Table 7.1 for descriptions.
- The **third** digit of the variable name refers to the LSIC wave:
  - “1” indicates the first wave,
  - “2” will indicate the second wave and
  - “3” will indicate the third wave.

Since this is the first release, all variables will have a “1” in the third position.

- The **fourth** character provides information on the type of variable. There are six different types of variables.
  - c** Coded variable: A variable coded with standard exhaustive code sets (SOC91 - Standard Occupational Classification system, NAICS – North American Industry Classification System, and the Census Country Code set).
  - d** Derived variable: A variable calculated usually from two or more collected or coded variables (e.g., household size, labour force status, etc.).
  - g** Grouped variable: Collected, coded or derived variables collapsed into groups (e.g., age groups, world region, etc.).
  - i** Imputation flag: Indicates that values in a variable for a respondent were imputed (field imputation), or that an entire entity was imputed (massive imputation). Field imputation flag variables directly follow the questions imputed and have an “i” at the fourth position instead of a “q”. For example, the imputation flag variable for in1q003 would be named in1i004.
  - q** Collected variable: A variable that refers to question which was directly asked to the respondent.
  - z** Variables obtained from a linkage with administrative records of Citizenship and Immigration Canada.



- The **fifth, sixth and seventh** characters are a sequential number (from 001 to nnn) assigned to the question inside a file. The order of the questions will closely match the order of the Wave 1 questionnaire, but changes in the Wave 2 questionnaire will have a serious impact on this order starting in Wave 2. **Users should be very cautious when using the order and the name of questions. They should always refer to the survey’s documentation.**

Whenever possible grouped variable names correspond to the variable that it is grouping (whether it is a derived variable or a regular question). Derived variables usually follow the question variables and have their own sequential number from 001 to nnn.

- The LSIC Wave 1 variables are comprised of seven characters. **The eighth and last character** (a letter) will be reserved to indicate important changes to a variable from one wave to another that could affect the comparability of the two variables. For example, it will be used to mark an addition or deletion to the answer categories associated with a question. Also, some revisions were made to the content of the questionnaire between waves. If the revision resulted in a change to the meaning or the values of a question, the variable will be treated as new and will have an eighth character. *This last character does not apply to Wave 1 questions.*

**Table 7.3 Examples of Variable Names**

| Variable<br>ci1q002 | Refers to:  |
|---------------------|---|
| ci                  | Question from the Citizenship entity  |
| 1                   | Wave 1 variable   |
| q                   | An item asked directly on the questionnaire   |
| 002                 | The second question on the Citizenship file: “Why did you choose to come to Canada: To join family or close friends (e.g., join spouse)?” |

| Variable<br>em1d009 | Refers to:   |
|---------------------|--|
| em                  | Question from the Employment entity  |
| 1                   | Wave 1 variable  |
| d                   | Derived variable   |
| 009                 | The ninth derived variable on the Employment file: “Number of days of all jobless spells.” |

## **7.9 Derived Variables**

### **Combining Items**

A number of variables have been derived by combining questions on the questionnaire in order to facilitate data analysis. For example, in the Employment Module, one of the questions is on labour force status since landing in Canada (em1d003). This variable uses two questions: “Since you came to Canada, have you worked at a job or business?” (em1q047) and if not, “Since you came to Canada, have you looked for work?” (em1q048).

### **Where to find the Derived Variables on the Files**

With the exception of the Longitudinal Respondent’s entity, which is mostly comprised of derived variables, the derived variables are usually placed after the questions in each entity to which they belong. The aggregated type of derived variables (for example, number of credentials reported, number of places lived in before current place) can be found on the main file (for example, number of credentials will be on the Education entity).

### **Derived Variable Name**

All derived variables on the LSIC data files have a "d" as the fourth character of the variable name. For example, the name of the variable for the “Total hours per week currently in class or training” is ed1d008.

## 8.0 Non-response

A survey's response rates are a measure of the effectiveness of the population being sampled, the collection process and are also a good indicator of the quality of the estimates produced. Perhaps more so than other surveys, the Longitudinal Survey of Immigrants to Canada (LSIC) is faced with a certain level of non-response. This chapter will provide a summary that distinguishes between three types of non-response: total, partial and unit non-response.

### **Total non-response:**

No information was collected for the sampled unit. For total non-response, some weighting adjustment methods were used to compensate. This topic is discussed in more detail in Chapter 10.0.

### **Partial non-response:**

At least some socio-demographic variables from the Background Module are collected, whether or not the other modules are reported.

### **Item non-response:**

The absence of information is limited only to some pre-determined variables, but all other variables within the modules are collected.

Partial and item non-response are corrected by various techniques of imputation. The two latter topics are discussed in the next section.

## 8.1 Definitions of Responding Unit

The following definitions outline the content of the tables below.

The **out-of-scope population** is the population of immigrants who were listed on the sampling frame but after some verification steps did not meet the criteria of the target population. Other examples of the out-of-scope population are immigrants who arrived under student visas, were deceased, resided in one of the territories or were institutionalized.

Note that the out-of-scope population also includes the **out-of-interest population** which is defined as immigrants in the target population who are not included in the population of interest because they are no longer residing in Canada. This population of immigrants meets all criteria but are not included in the final weights for the population of interest. The **population of interest** is the population for which the variables of interest are being estimated in relation to the survey objectives. The LSIC survey objectives pertain to the integration of immigrants in Canada, hence, immigrants no longer residing in Canada are not included in the population of interest.

A **responding immigrant** is the immigrating unit member selected as longitudinal respondent (LR), for whom there is usable minimal information on the questionnaire. After verification that the appropriate selected immigrant had been contacted and responded to the survey, some mandatory fields were identified in order to determine the extent of usable information and subsequent imputation. After this verification, 12,040 usable records were identified as responding units. More details are available in Chapter 5.0.

**Unresolved or untraced** refers to cases identified during collection where there was no contact at all with the selected immigrant. No information was collected as to their whereabouts.

**Non-respondents** refers to cases identified during collection where the selected immigrant was somehow located and confirmed to be in Canada, but for a given reason could not respond to the interview. The interviewers' notes were also used to complete the coding. Even though, both unresolved and non-respondent cases result in unusable records, the main difference between the two is that in cases of non-response the respondent was confirmed to be residing in Canada.

**Table 8.1 Results of Wave 1 Collection by Reference Month and Year – Unweighted**

| Month and Year | Respondents   | Non-respondents | Out-of-interest population | Unresolved   | Total         |
|----------------|---------------|-----------------|----------------------------|--------------|---------------|
| October 2000   | 992           | 233             | 36                         | 533          | 1,794         |
| November 2000  | 1,041         | 206             | 29                         | 497          | 1,773         |
| December 2000  | 987           | 186             | 34                         | 437          | 1,644         |
| January 2001   | 972           | 178             | 28                         | 500          | 1,678         |
| February 2001  | 1,056         | 176             | 36                         | 431          | 1,699         |
| March 2001     | 989           | 167             | 35                         | 496          | 1,687         |
| April 2001     | 982           | 134             | 36                         | 530          | 1,682         |
| May 2001       | 1,048         | 130             | 39                         | 450          | 1,667         |
| June 2001      | 1,039         | 158             | 29                         | 456          | 1,682         |
| July 2001      | 1,012         | 155             | 35                         | 479          | 1,681         |
| August 2001    | 995           | 186             | 48                         | 452          | 1,681         |
| September 2001 | 927           | 211             | 26                         | 490          | 1,654         |
| <b>Total</b>   | <b>12,040</b> | <b>2,120</b>    | <b>411</b>                 | <b>5,751</b> | <b>20,322</b> |

**Table 8.2 Results of Wave 1 Collection by Class of Immigrant – Unweighted**

| Class of Immigrant | Respondents   | Non-respondents | Out-of-interest population | Unresolved   | Total         |
|--------------------|---------------|-----------------|----------------------------|--------------|---------------|
| Economic           | 6,973         | 1,117           | 294                        | 4,186        | 12,570        |
| Family             | 3,365         | 799             | 98                         | 1,066        | 5,328         |
| Refugees           | 1,590         | 190             | 18                         | 471          | 2,269         |
| Other              | 112           | 14              | 1                          | 28           | 155           |
| <b>Total</b>       | <b>12,040</b> | <b>2,120</b>    | <b>411</b>                 | <b>5,751</b> | <b>20,322</b> |

**Table 8.3 Results of Wave 1 Collection by Age Groups – Unweighted**

| Age Groups   | Respondents   | Non-respondents | Out-of-interest population | Unresolved   | Total         |
|--------------|---------------|-----------------|----------------------------|--------------|---------------|
| 15-24        | 2,325         | 410             | 73                         | 949          | 3,757         |
| 25-34        | 4,597         | 676             | 143                        | 2,566        | 7,982         |
| 35-44        | 3,020         | 470             | 107                        | 1,497        | 5,094         |
| 45-64        | 1,755         | 429             | 70                         | 656          | 2,910         |
| 65 +         | 343           | 135             | 18                         | 83           | 579           |
| <b>Total</b> | <b>12,040</b> | <b>2,120</b>    | <b>411</b>                 | <b>5,751</b> | <b>20,322</b> |

**Table 8.4 Results of Wave 1 Collection by Sex – Unweighted**

| Sex          | Respondents   | Non-respondents | Out-of-interest population | Unresolved   | Total         |
|--------------|---------------|-----------------|----------------------------|--------------|---------------|
| Male         | 6,039         | 1,028           | 215                        | 2,920        | 10,202        |
| Female       | 6,001         | 1,092           | 196                        | 2,831        | 10,120        |
| <b>Total</b> | <b>12,040</b> | <b>2,120</b>    | <b>411</b>                 | <b>5,751</b> | <b>20,322</b> |

**Table 8.5 Results of Wave 1 Collection by Intended Province of Destination – Unweighted**

| Province                  | Respondents   | Non-respondents | Out-of-interest population | Unresolved   | Total         |
|---------------------------|---------------|-----------------|----------------------------|--------------|---------------|
| Newfoundland and Labrador | 27            | 6               | 0                          | 14           | 47            |
| Prince Edward Island      | 9             | 3               | 0                          | 10           | 22            |
| Nova Scotia               | 63            | 24              | 4                          | 63           | 154           |
| New Brunswick             | 51            | 9               | 2                          | 17           | 79            |
| Québec                    | 1,708         | 208             | 34                         | 666          | 2,616         |
| Ontario                   | 6,215         | 1,121           | 227                        | 3,528        | 11,091        |
| Manitoba                  | 254           | 34              | 10                         | 67           | 365           |
| Saskatchewan              | 111           | 15              | 4                          | 45           | 175           |
| Alberta                   | 1,303         | 208             | 39                         | 317          | 1,867         |
| British Columbia          | 2,299         | 492             | 91                         | 1,024        | 3,906         |
| <b>Canada</b>             | <b>12,040</b> | <b>2,120</b>    | <b>411</b>                 | <b>5,751</b> | <b>20,322</b> |

**Table 8.6 Results of Wave 1 Collection by Place of Birth – Unweighted**

| <b>Place of Birth</b> | <b>Respondents</b> | <b>Non-respondents</b> | <b>Out-of-interest population</b> | <b>Unresolved</b> | <b>Total</b>  |
|-----------------------|--------------------|------------------------|-----------------------------------|-------------------|---------------|
| Africa                | 1,199              | 171                    | 34                                | 527               | 1,931         |
| America               | 873                | 109                    | 21                                | 370               | 1,373         |
| Asia                  | 7,695              | 1,442                  | 291                               | 4,234             | 13,662        |
| Europe                | 2,189              | 377                    | 64                                | 600               | 3,230         |
| Oceania               | 84                 | 21                     | 1                                 | 20                | 126           |
| <b>Total</b>          | <b>12,040</b>      | <b>2,120</b>           | <b>411</b>                        | <b>5,751</b>      | <b>20,322</b> |

## 9.0 Imputation

For partial and item non-response, imputation techniques were used. Imputation is basically the process during which a plausible value is produced to replace a missing or inconsistent value. The goal is to construct values that will lead to approximately unbiased estimators. There are many well-known techniques available to impute values for a given record or variable. When carried out properly, imputation improves data quality by reducing non-response bias. This was done to ensure that a complete data set of variables or records was produced and to minimize the “not stated” fields in the microdata file.

For item non-response, deterministic imputation was performed. Deterministic imputation is the process by which another source of data is used for a similar concept and from the exact same respondent. For example, date of birth is not reported on the questionnaire by Mr. X. However, there is an administrative file which contains the date of birth of Mr. X. The use of the date of birth found on the administrative file is considered deterministic imputation. For the Longitudinal Survey of Immigrants to Canada (LSIC), if a respondent did not report information for certain pre-determined variables, the information was imputed from the Field Operations Support System (FOSS). The FOSS values were used as if they were reported information.

Two imputation techniques were also performed specifically for the Income Module: nearest-neighbour donor imputation for some fields and median imputation for certain identified outliers. The next two sections include, respectively, a description of nearest-neighbour donor imputation and the techniques used for imputation in the Income Module.

### 9.1 Mass Imputation

For partial non-response, mass imputation for the non-reported modules was carried out using the nearest-neighbour donor technique. The donor imputation method generally will not alter the distribution of the data, which is a drawback of many other imputation techniques.

The nearest-neighbour donor technique is a widely used technique for treating non-response. It aims at replacing missing information for a respondent with values provided from another respondent who is “similar” to him/her. It works in the following manner: based on a statistical distance calculated on selected socio-demographic information, a donor (responding unit) determined to be the closest to the partial respondent (the recipient) is identified and the values of the donor are used to replace the missing values for the recipient.

For the LSIC, this was conducted module by module. The first step was to identify which modules had to be imputed. Key fields were identified and verified in order to determine if the respondent reported valid values. If all key fields were reported appropriately then the modules were defined as complete. If at least one field was invalid, the module was defined as incomplete and was imputed. For a unit which had more than one module missing, the same donor record was used. Note that only complete and edited records were in the donor pool that could be used as potential donors. To keep consistency within variables, the complete set of variables for a given module of the donor was imputed into the recipient record. At the end of this process, all records had fully completed modules. A file of flags indicating whether a variable or an entire module was imputed was created.

Table 9.1 presents the different patterns of module non-response for all partially responding records. In the table, a “1” denotes that the module is complete, i.e. all key fields have valid values, while a “2” indicates that the module is incomplete.

**Table 9.1 Distribution of Module Completion**

| BG | SN | LS | HS | ED | EM | HL | IN | Number of Cases | Percent |
|----|----|----|----|----|----|----|----|-----------------|---------|
| 1  | 1  | 1  | 1  | 1  | 1  | 1  | 2  | 370             | 3.06%   |
| 1  | 1  | 1  | 1  | 1  | 1  | 2  | 1  | 2               | 0.02%   |
| 1  | 1  | 1  | 1  | 1  | 1  | 2  | 2  | 16              | 0.13%   |
| 1  | 1  | 1  | 1  | 1  | 2  | 1  | 2  | 2               | 0.02%   |
| 1  | 1  | 1  | 1  | 1  | 2  | 2  | 2  | 11              | 0.09%   |
| 1  | 1  | 1  | 1  | 2  | 1  | 1  | 1  | 6               | 0.05%   |
| 1  | 1  | 1  | 1  | 2  | 2  | 2  | 2  | 13              | 0.11%   |
| 1  | 1  | 1  | 2  | 1  | 1  | 1  | 1  | 9               | 0.07%   |
| 1  | 1  | 1  | 2  | 1  | 1  | 1  | 2  | 8               | 0.07%   |
| 1  | 1  | 1  | 2  | 1  | 2  | 2  | 2  | 1               | 0.01%   |
| 1  | 1  | 1  | 2  | 2  | 2  | 2  | 2  | 6               | 0.05%   |
| 1  | 1  | 2  | 1  | 1  | 1  | 1  | 1  | 12              | 0.10%   |
| 1  | 1  | 2  | 2  | 2  | 2  | 2  | 2  | 6               | 0.05%   |
| 1  | 2  | 1  | 1  | 1  | 1  | 1  | 1  | 8               | 0.07%   |
| 1  | 2  | 1  | 1  | 1  | 1  | 1  | 2  | 7               | 0.06%   |
| 1  | 2  | 1  | 2  | 1  | 1  | 2  | 2  | 2               | 0.02%   |
| 1  | 2  | 1  | 2  | 2  | 2  | 2  | 2  | 2               | 0.02%   |
| 1  | 2  | 2  | 2  | 2  | 2  | 2  | 2  | 8               | 0.07%   |

BG - Background; SN - Social Network; LS – Language Skills; HS - Housing; ED - Education; EM - Employment; HL - Health; IN - Income.

Table 9.1 shows that the Income Module was the least reported module with 3% non-response. For the Income Module, a different processing approach was used. This approach is described in the next section.

## 9.2 Imputation for Income and Other Quantitative Variables

For quantitative variables such as wages, income and total earnings, editing as well as imputation was performed. The first step in the imputation process was to transform all values to the same yearly basis. As respondents could report values for different periods, all values had to be processed on the same basis. Values were transformed back to their original state after the imputation was completed. In some cases, respondent answers, especially government-sources income, were changed to preset maximum or minimum allowable values for that variable.

Outlier detection was also performed and records identified as outliers were sent for further manual verification and review. Missing data were imputed using nearest-neighbour donor imputation as well, but instead of performing mass imputation of all variables for the entire module, imputation was carried out variable by variable. Rules for identifying the respondent most similar to the non-respondent varied depending on the variable being imputed. Once the nearest neighbour imputation was done, within-record editing was performed again to ensure consistency of the data.



A list of variables for which imputation was carried out can be found in Table 9.2. The table shows an overall imputation rate for all jobs combined for each of the variables. Note that although imputation generally improves overall data quality, the artificial data created are used in estimation and can lead to underestimation of the sampling errors. This would only be a concern for variables with high imputation rates.

**Table 9.2 Imputation Rates for Income and Earnings**

| Variable Description                                       | Variable Name      | Number of Non-skips | Number of Imputed Values | Imputation Rate |
|--|--------------------|---------------------|--------------------------|-----------------|
| Income from all jobs                                       | in1i004            | 7,899               | 1,210                    | 15.32%          |
| Income from self-employment                                | in1i006            | 493                 | 258                      | 52.33%          |
| Pension from a Canadian business or company                | in1i028            | 41                  | 14                       | 34.15%          |
| Private sponsor  | in1i031            | 86                  | 4                        | 4.65%           |
| Investments  | in1i034            | 312                 | 68                       | 21.79%          |
| Other sources  | in1i037            | 516                 | 34                       | 6.59%           |
| Social Assistance  | in1i009            | 1,530               | 36                       | 2.35%           |
| Employment Insurance                                       | in1i012            | 325                 | 45                       | 13.85%          |
| Child tax benefits or credits                              | in1i015            | 3,781               | 221                      | 5.85%           |
| Canadian or Quebec Pension                                 | in1i018            | 176                 | 28                       | 15.91%          |
| Other government sources                                   | in1i021<br>in1i024 | 842                 | 52                       | 6.18%           |
| Longitudinal respondent's personal income from all sources | in1i007            | 12,040              | 355                      | 2.95%           |



## 10.0 Treatment of Total Non-response and Weighting

The Longitudinal Survey of Immigrants to Canada (LSIC) is a probability survey. As is the case with any probability survey, the sample is selected to represent a reference population - the immigrant population - at a specific date within the context of the survey as accurately as possible. Each unit in the sample must therefore represent a certain number of units in the population. If all selected units were traced, contacted and completed, and if the frame used was perfect, (covering exactly the population of interest), then the design weight assigned to each unit would represent accurately and exactly the number of immigrants in the target population. In this situation, using this weight would yield unbiased estimates. However, this is not the case when surveys are faced with non-response, unresolved/untraceable units and imperfect frames. Weight adjustments are traditionally used to compensate for these different issues.

### 10.1 Representativity of the Weights

For most surveys, the sum of the final weights represents the estimated target population counts which usually equate to the population of interest. However, in the case of the LSIC, because of the mobility of the population and the survey objectives (Chapter 3.0), the population of interest is actually a portion of the target population. Recall that the survey frame covers the target population - immigrants who meet all of the following criteria:

- arrived in Canada between October 1, 2000 and September 30, 2001;
- were age 15 or older at the time of landing;
- landed from abroad, must have applied through a Canadian Mission Abroad.

However, some of these immigrants resided in Canada for only a short period of time before returning to their original country or migrating to another country. These immigrants do not have similar adaptation characteristics as the ones who are permanently residing in Canada. It is biased to include in the same weight adjustment the immigrants who moved out of Canada and those who still reside in Canada. Therefore, the target population includes two basic sub-groups as defined in Section 8.1. The **population of interest** consists of immigrants who live in Canada for more than six months of the year. The final weight yields unbiased estimates of the population of interest. The **out-of-interest population** consists of immigrants who no longer live in Canada, i.e., who have left since landing in Canada.

### 10.2 Overview of the Weight Adjustments

During collection, there were four possible classifications for a selected immigrant; respondent, non-respondent, not in the population of interest, and unresolved. The first three categories resulted in an initial contact with the immigrant or with someone who was able to confirm their status. These cases are defined as resolved cases as the immigrant had a known status. The last collection outcome is the unresolved cases. For these, no contact was established and they remained unresolved. No information on whether they were still in Canada was available. The weight adjustments reflect these outcomes.

The sample can first be split between the resolved and the unresolved cases:

$$\text{Sample } S = S_U + S_R$$

where  $S_U$  = sampled units unresolved  
 $S_R$  = sampled units resolved

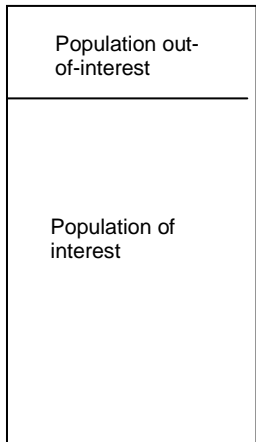
$$\text{Furthermore, in the resolved portion } S_R = S_{RR} + S_{RN} + S_{RO}$$

where  $S_{RR}$  = sampled units resolved that are respondents  
 $S_{RN}$  = sampled units resolved that are non-respondents  
 $S_{RO}$  = sampled units resolved that are not in the population of interest, referred to as OOI, out-of-interest.

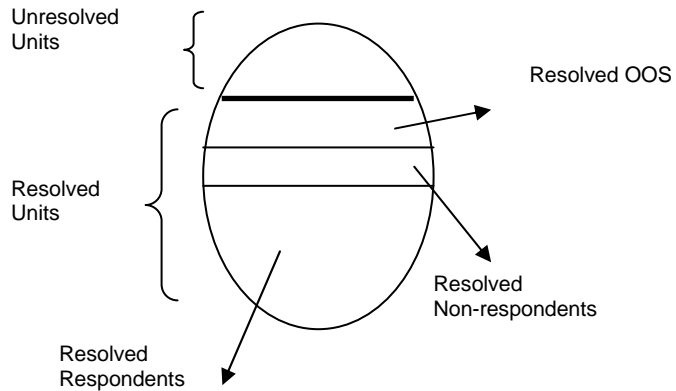
Individuals who are out-of-scope are represented by OOS.

The following diagram presents an overview of these concepts as they relate to weighting.

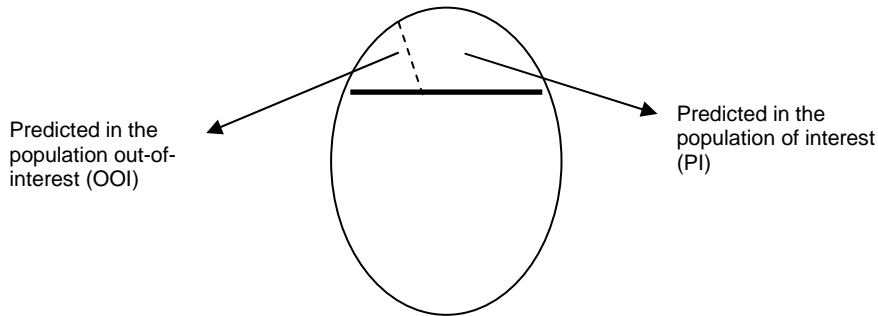
**Frame for Selection**



**Sample and Collection Outcome**



Conceptually, for the set of units that remained unresolved ( $S_U$ ), it is fair to assume it is composed of units in the population of interest (PI) and the population out-of-interest (OOI). However, at that point of the process, there was no information available. Consequently, the first step of the weighting process was to predict for the unresolved units whether they would have been in the population of interest or not. Through models, using the information available on the frame and from the resolved units, the status of the unresolved units was predicted as: PI or OOI as shown in the following diagram.



After this first step, we have a status (predicted or confirmed) for each selected unit indicating if they were part of the population of interest or not in the population of interest. Note that in the resolved units, the population of interest is composed of respondents and non-respondents. Thus the following notation, which will be used in subsequent sections:

For the unresolved units ( $S_U$ ):

$$j \in S_U = \text{unresolved sampled units } j \text{ where } S_U = \hat{S}_{U\_II} + \hat{S}_{U\_OOI}$$

$$j \in \hat{S}_{U\_II} = \text{unresolved sampled units } j \text{ predicted as PI}$$

$$j \in \hat{S}_{U\_OOI} = \text{unresolved sampled units } j \text{ predicted as OOI}$$

For the resolved units ( $S_R$ ):

$$i \in S_R = \text{resolved sampled units } i \text{ where } S_R = S_{RR} + S_{RN} + S_{RO}$$

$$i \in S_{RN} = \text{resolved non-respondents units } i$$

$$i \in S_{RR} = \text{resolved respondents units } i$$

$$i \in S_{RO} = \text{resolved OOI units } i$$

### 10.3 Longitudinal Weighting for Responding Immigrants

The LSIC weighting strategy is based on a series of cascading adjustments. The final longitudinal weight is obtained by applying various adjustments to the basic initial design weight. There are four weights involved in the weighting process which will compose the final weight; the design weight, the non-response adjustment weight, the resolved adjustment weight and finally the post-stratification weight. Table 10.1 shows the relationship between the different categories of outcomes related to the adjustment.

**Table 10.1 Process of Classifying the Respondents Outcome Status**

| Selection        | Tracing        | Status                       | Response             |                   |
|------------------|----------------|------------------------------|----------------------|-------------------|
| Sampled units    | Resolved units | PI:<br>In scope units        | Responding units     |                   |
|                  |                |                              | Non-responding units | Refusal           |
|                  |                |                              |                      | Language problems |
|                  |                |                              |                      | LR absent         |
|                  |                | OOI (Left Canada, dead, etc) |                      |                   |
| Unresolved units |                |                              |                      |                   |

Note that on the microdata file, only the responding resolved units, ( $i \in S_{RR}$ ), have a final weight as they are the only units which have fully completed records. As for the out-of-interest population, ( $i \in S_{RO}$ ) they also have a final weight, but are not available on the microdata file as they do not have full records. Only tabulations of this sub-population using the final weights are available.

The subsequent sections describe the design weights (Section 10.3.1), the two weight adjustments, i.e. non-response and unresolved (Section 10.3.2) and finally post-stratification is explained in Section 10.3.3.

### 10.3.1 Design Weight

At the time of selection, an initial design weight was assigned to the selected person. It is simply the inverse of the probability of selection of the selected immigrants. The probability of selection is a function of the selection method. Section 5.3 presents the details of the two-stage sample approach chosen for the LSIC. Thus, the design weight of each selected person is equal to the inverse of the probability of selection of the immigrating unit (IU) in which the person is selected, multiplied by the number of eligible persons in this IU. The probability of selection of the group itself is equal to the product of the size measure of the IU and the number of IUs selected in the stratum, divided by the total size of the IU in the stratum.

Probability of selection of one immigrant = probability of selection of an immigrating unit \* the probability of selection of one immigrant among the group

Or algebraically, the probability of selection of immigrating unit  $g$  in stratum  $h$  is:

$$\pi_{gh} = \frac{S_{gh} m_h}{\sum_{g=1}^{M_h} S_{gh}^*}$$

and the probability of selection of immigrant  $i$  in stratum  $h$  is:

$$\pi_{igh} = \frac{1}{S_{gh}}$$

$h$  = stratum  $h = 1, \dots, H$

$g$  = immigrating unit  $g = 1, \dots, M$

$M_h$  = number of immigrating units belonging to stratum  $h$

$m_h$  = number of selected immigrating units belonging to stratum  $h$

$S_{gh}$  = number of immigrants (size) in immigrating unit  $g$ , in stratum  $h$

$S_{gh}^*$  = number of immigrants (size) adjusted according to a maximum size, in immigrating unit  $g$ , in stratum  $h$

Obtaining design weight  $W_{igh}$  for each of the immigrants  $i$  selected in immigrating unit  $g$  in stratum  $h$  :

$$w_{igh} = \pi_{igh}^{-1} \pi_{gh}^{-1} = \frac{S_{gh} \sum_{g=1}^{M_h} S_{gh}^*}{m_h S_{gh}}$$

Because systematic selection was used, a maximum  $S_{gh}^*$  was introduced in strata with a high sampling fraction. This was to ensure the selection of only one immigrant per group as mathematically speaking, a large group could end up with more than one immigrant selected. Hence, the truncation by some size measure in those strata. Note that this only applies when systematic selection is used.

### 10.3.2 Non-response and Unresolved Adjustment Weights

For the resolved responding units ( $i \in S_{RR}$ ), the weight adjustment has the following formulation [before the post-stratification adjustment]:

Intermediate weight = design weight \* non-response adjustment \* unresolved adjustment

or

= design weight \*  $\frac{\text{weighted sum of resolved unit}}{\text{weighted sum of respondent}}$  \*  $\frac{\text{weighted sum of resolved and predicted resolved}}{\text{weighted sum of resolved}}$

or algebraically

$$\sum_{i \in S_{RR}} w_{int\_PI} = \sum_{i \in S_{RR}} w_D * \left[ \frac{\sum_{G_1} \sum_{i \in S_{RR}} w_D + \sum_{G_1} \sum_{i \in S_{RN}} w_D}{\sum_{G_1} \sum_{i \in S_{RR}} w_D} \right] * \left[ \frac{\sum_{G_2} \sum_{j \in \hat{S}_{U\_PI}} w_D + \sum_{G_2} \sum_{i \in S_{R\_PI}} w_1}{\sum_{G_2} \sum_{i \in S_{R\_PI}} w_1} \right]$$

$$\text{where } w_1 = w_D * \left[ \frac{\sum_{G_1} \sum_{i \in S_{RR}} w_D + \sum_{G_1} \sum_{i \in S_{RN}} w_D}{\sum_{G_1} \sum_{i \in S_{RR}} w_D} \right]$$

where  $G_1$  = non-response adjustment class  
 $G_2$  = unresolved adjustment class  
 $W_{D_{int\_PI}}$  = intermediary weight of the population of interest PI  
 $W_D$  = design weight (see Section 10.3 for more detail on the design weight calculations)

For the resolved out-of-interest population ( $i \in S_{RO}$ ), there is only one adjustment, i.e., one adjustment to compensate for the predicted out-of-interest ( $j \in \hat{S}_{U\_OOI}$ ) in the unresolved one.

$$\sum_{i \in S_{RO}} W_{int\_OOI} = \sum_{i \in S_{RO}} W_D * \left[ \frac{\sum_{G_2, j \in \hat{S}_{U\_OOI}} W_D + \sum_{G_2, i \in S_{RO}} W_D}{\sum_{G_2, i \in S_{RO}} W_D} \right]$$

### 10.3.3 Post-stratification

Post-stratification is one of the calibration estimation techniques widely used in social surveys. It allows benchmarking on new updated population counts. It has the same objective as the stratification; however, the groups are defined more appropriately with the estimation domains in mind, i.e. stratification a posteriori. The post-stratification file still represents the target population. The file was created with the same definitions and criteria as the survey frame, but with more up-to-date files. For example, it included new units, excluded deaths and/or updated missing or improperly specified variables that were on the survey frame. The post-stratification variables used were: age group, sex, place of birth (collapsed by world area) and class of immigrant.

Tables 10.2 through 10.5 provide the detailed categories.

**Table 10.2 Age Group**

|             |
|-------------|
| 15 – 24     |
| 25 – 34     |
| 35 – 44     |
| 45 and over |

**Table 10.3 Sex**

|        |
|--------|
| Male   |
| Female |



**Table 10.4 Place of Birth**

| Region                            | World Area (WA) |
|-----------------------------------|-----------------|
| Central Africa                    | 1 - Africa      |
| Eastern Africa                    |                 |
| Northern Africa                   |                 |
| Southern Africa                   |                 |
| Western Africa                    |                 |
| Central America                   | 2 - America     |
| Northern America                  |                 |
| Southern America                  |                 |
| Caribbean and Bermuda             |                 |
| Eastern Asia                      | 3 - Asia        |
| Southeast Asia                    |                 |
| Southern Asia                     |                 |
| West Central Asia and Middle East |                 |
| Eastern Europe                    | 4 - Europe      |
| Northern Europe                   |                 |
| Southern Europe                   |                 |
| Western Europe                    |                 |
| Oceania                           | 5 - Oceania     |

**Table 10.5 Class of Immigrant**

|   |
|---|
| Family  |
| Economic skilled worker                             |
| Economic business independent and other independent |
| Refugees government sponsored                       |
| Refugees other                                      |

The variables are cross-tabulated except in the following situations:

- For Oceania, there is only one other cross-tabulation: Family versus all other immigration classes collapsed together. There is neither sex nor age grouping for the post-stratification.
- For Government Sponsored Refugees the age groups 35 to 44 years and 45 years and over are collapsed.
- For Other Refugees, there is neither sex nor age grouping for the post-stratification.
- For Family class of immigrants from Africa, age 35 to 44 years, sex was collapsed.
- For Economic Business Independent and Other Independent from America, there is no age grouping for the post-stratification.

The adjustment has the following form:

$$\text{Final weight} = \text{intermediate weight} * \frac{\text{new counts at the post - strata}}{\text{estimated population counts using the intermediary weights}}$$

or algebraically for  $i \in S_{RR}$ ,

$$\sum_{i \in S_{RR}} W_f = \sum_{i \in S_{RR}} w_{\text{int\_II}} * \frac{N''_{pst}}{\sum_{pst} \sum_{i \in S_{RR}} w_{\text{int\_II}} + \sum_{pst} \sum_{i \in S_{RO}} w_{\text{int\_OOI}}}$$

### 10.3.4 Adjustment Classes: Homogeneous Groups

The weight adjustment classes, as well as the post-stratification groups, are constructed under the same assumption. They must be homogeneous groups related to the correction being made (the non-response adjustment classes are constructed based on the homogeneity of responses within a class, meaning that they have the same probability of response). The unresolved adjustment classes were constructed based on homogeneity or a similar propensity of being resolved and being in scope.

For the LSIC, the non-response and the unresolved adjustment classes were derived based on a logistic regression predicting respectively, the response probability and the resolution probability. For the latter model, the explanatory variables for predicting the population of interest status were included by default in the model.

The predictors or explanatory variables for the model predicting **responses** were; **class of immigrant, age group, level of education, knowledge of official language, and mother tongue.**

The explanatory variables for the model predicting the propensity of being **resolved** were **quality of tracing source, reference month, and number of years of school.** In this model, the predictor of being **in the population of interest, level of education and age** were included by default. The classes were constructed using similar probabilities obtained from each respective model. The number of classes for each adjustment was defined based on a convergence algorithm ensuring unbiased estimates.

## 11.0 Data Quality and Coverage

This chapter provides the user with information about the various factors affecting the quality of the survey data. There are two main types of errors: sampling errors and non-sampling errors. A sampling error is the difference between an estimate derived from a sample and the one that would have been obtained from a census that used the same procedures to collect data from every person in the population. All other types of errors such as frame coverage, response, processing and non-response are non-sampling errors. Many of these errors are difficult to identify and quantify. These are discussed in Section 11.2.

### 11.1 Sampling Errors

The estimates derived from this survey are based on a sample of immigrants and not from a complete enumeration (census) under similar conditions. This difference is the sampling error of the estimates. Statistics Canada's data quality guidelines have provided users with an indication of the magnitude of the sampling error. It is highly recommended that users analyzing data or producing estimates from the Longitudinal Survey of Immigrants to Canada (LSIC) data file do so as well.

The basis for measuring sampling error is the standard error of the estimates derived from survey results. However, because of the large variety of estimates that can be produced from a survey, the standard error of an estimate is usually expressed relative to the estimate to which it pertains. This measure, known as the coefficient of variation (CV) of an estimate, is obtained by expressing the standard error of the estimate as a percentage of the estimate. The smaller the CV, the smaller the sampling variability, meaning smaller CVs are more desirable. The CV depends on the size of the sample on which the estimate is based, the population size and on the distribution of the sample, i.e. the sampling fraction of the units of the domains being estimated. The following diagram presents the characteristics of some coefficients of variation and the Statistics Canada guidelines for release.

#### Characteristics

|   |
|---|
| 0.0% - 1.0% Excellent<br>1.0% - 5.0% Very Good<br>5.0% - 10.0% Good<br>10.0% - 16.5% Moderate |
|---|

|               |
|---------------|
| 16.6% - 33.3% |
|---------------|

|         |
|---------|
| 33.4% + |
|---------|

#### Guidelines for Release

|                                   |
|-----------------------------------|
| Reliable enough for most purposes |
|-----------------------------------|

|                   |
|-------------------|
| Use with caution! |
|-------------------|

|                     |
|---------------------|
| Data not acceptable |
|---------------------|

## 11.2 Non-sampling Errors

There are many sources of non-sampling errors that are not related to sampling, but may occur at almost any phase of a survey operation. Interviewers may misunderstand survey instructions, respondents may make a mistake in answering the questions, responses may be recorded in the questionnaire incorrectly or errors may be made in the processing or tabulating of the data. For the LSIC, quality assurance measures were implemented at each phase of the data collection and processing cycles to monitor the quality of the data. These measures included precise interviewer training with respect to the survey procedures and questionnaire, observation of interviews to detect questionnaire design problems or misinterpretation of instructions, monitoring of final coding, and coding and edit quality checks to verify the processing logic. Chapter 7.0 outlines data processing procedures. Other kinds of non-sampling error are more easily quantifiable, especially non-response and the population frame coverage, the topics of the next two sections.

## 11.3 Non-response and Unresolved Cases

Non-response and unresolved cases, if not appropriately corrected, are the types of error that can lead to bias in the survey estimates. For the LSIC, these two types of response categories reduced significantly the number of usable records. Biased estimates can occur when unusable units have significantly different characteristics from the usable ones. In the evaluation stage, studies were completed to understand the non-response mechanism. Results showed that non-response units and unresolved units displayed different patterns and different rates were obtained for different characteristics of immigrants.

Table 11.1 provides the sample distribution by response category for each class of immigrant. It shows that for the refugees and economic classes, there were higher rates of responding units (once resolved) than for the family class, even though they had lower resolved rates. For the family class, about 79.9% of the sample was resolved while 20.1% remained unresolved. In the economic classes, 66.5% of cases were resolved while 80.0% were resolved for the refugee classes. The rates of responding units once resolved for the economic, family and refugee classes were 83.2%, 78.9% and 88.4% respectively. Family class had a higher rate of non-response. One possible explanation for the higher non-response for this class may be due to language barriers which occur for many family class immigrants coming to Canada to join family members. This hypothesis was studied and this variable was one of the explanatory variables in the response prediction model described in Section 10.3.4.

**Table 11.1 Results of Wave 1 Data Collection by Class of Immigrant**

| Results                 | Economic | Family | Refugees | Other | Total |
|-------------------------|----------|--------|----------|-------|-------|
|                         | %        |        |          |       |       |
| <b>Resolved</b>         | 66.5     | 79.9   | 80.0     | 82.0  | 71.8  |
| In-scope responding     | 83.2     | 78.9   | 88.4     | 88.2  | 82.7  |
| In-scope non-responding | 13.3     | 18.7   | 10.6     | 11.0  | 14.5  |
| Out-of-scope            | 3.5      | 2.3    | 1.0      | 0.8   | 2.8   |
| <b>Unresolved</b>       | 33.5     | 20.1   | 20.0     | 18.0  | 28.3  |

After numerous studies of the different rates and characteristics, it was fair to assume non-random response and resolved patterns. Both responding and non-responding units as well as resolved and unresolved units showed different patterns. Every non-random

pattern must be corrected with the use of appropriate weight adjustment classes, taking into account the characteristics that lead to these different patterns. For example, if sex is an explanatory variable in the response prediction model, (i.e. different response rates for male and female), then sex must be used in the correction.

For these reasons, the adjustment weights were calculated in distinct steps for the responding units and for the resolved units as described in Section 10.3. Response and resolution models were used to construct the proper adjustment weights to correct for the fact that there were different response rates and different resolved rates. It also stresses the importance of using the final weights in any tabulation or analysis using the LSIC data. Any estimation done without the use of weights will produce biased results.

#### **11.4 Coverage**

Coverage is an indication of how a survey frame covers the target population or in the case of the LSIC, the population of interest. There could be over-coverage if the survey frame contains units that should not have been included, such as death, duplicates, or incorrect date of birth captured on the file. There could also be under-coverage, if the survey frame missed some units that should have been included. For the LSIC, there was a slight over-coverage which was corrected using a post-stratification technique on a more up to date file (see Section 10.3.3).



## 12.0 Guidelines for Tabulation, Analysis and Release

This chapter of the documentation outlines the guidelines to be adhered to by users tabulating, analyzing, publishing or otherwise releasing any data derived from the survey microdata files. With the aid of these guidelines, users of microdata should be able to produce the same figures as those produced by Statistics Canada and, at the same time, will be able to develop currently unpublished figures in a manner consistent with these established guidelines.

### 12.1 Rounding Guidelines

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

- a) Estimates in the main body of a statistical table are to be rounded to the nearest hundred units using the normal rounding technique. In normal rounding, if the first or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is raised by one. For example, in normal rounding to the nearest 100, if the last two digits are between 00 and 49, they are changed to 00 and the preceding digit (the hundreds digit) is left unchanged. If the last digits are between 50 and 99 they are changed to 00 and the preceding digit is incremented by 1.
- b) Marginal sub-totals and totals in statistical tables are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units using normal rounding.
- c) Averages, proportions, rates and percentages are to be computed from unrounded components (i.e. numerators and/or denominators) and then are to be rounded themselves to one decimal using normal rounding. In normal rounding to a single digit, if the final or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is increased by 1.
- d) Sums and differences of aggregates (or ratios) are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units (or the nearest one decimal) using normal rounding.
- e) In instances where, due to technical or other limitations, a rounding technique other than normal rounding is used resulting in estimates to be published or otherwise released which differ from corresponding estimates published by Statistics Canada, users are urged to note the reason for such differences in the publication or release document(s).
- f) Under no circumstances are unrounded estimates to be published or otherwise released by users. Unrounded estimates imply greater precision than actually exists.

## 12.2 Sample Weighting Guidelines for Tabulation

The sample design used for the LSIC was self-weighting. When producing simple estimates, including the production of ordinary statistical tables, users must apply the final weight. If final weights are not used, the estimates derived from the microdata files cannot be considered to be representative of the survey population, and will not correspond to those produced by Statistics Canada. The weight assigned to each immigrant reflects the number of immigrants represented by a particular respondent.

For any analysis dealing with correlation analysis or any other statistics where a significance measure is required, it is recommended that an adjusted weight be used. This weight is obtained by multiplying the final weight by the sample size and dividing this total by the total estimated population. This produces a mean weight of 1 and a sum of weights equal to the sample size.

The benefit of this adjusted weight is that an overestimation of the significance level (which is very sensitive to sample size) is avoided while maintaining the same distributions as those obtained when using the demographic weight. The disadvantage is that the numerator is not weighted up to the target population.

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

## 12.3 Definitions of Types of Estimates: Categorical and Quantitative

The LSIC file has been set up so that the longitudinal respondent is the unit of analysis. The weight that can be found on each record (WT1L) is an “immigrant” (the longitudinal respondent) weight. Estimates of the respondent’s children or spouse cannot be generated from the LSIC microdata file.

### Categorical Estimates

Categorical estimates are estimates of the number, or percentage of the surveyed population possessing certain characteristics or falling into some defined category. The number or the proportion of immigrants who plan to purchase a house or an apartment in the next few years are examples of such estimates. An estimate of the number of persons possessing a certain characteristic may also be referred to as an estimate of an aggregate.

#### Examples of Categorical Questions:

Q: Do you or your family have plans to purchase a house or an apartment in the next few years?

R: Yes / No / Not sure

Q: How many rooms are there where you live (Include kitchen, bedrooms, finished rooms in the attic or basement, etc.)? Do not count bathrooms, halls, vestibules and rooms used solely for business purposes.

R: One / Two / Three / Four / Five or more



### Quantitative Estimates

Quantitative estimates are estimates of totals or of means, medians and other measures of central tendency of quantities based upon some or all of the members of the surveyed population. They also specifically involve estimates of the form  $\hat{X}/\hat{Y}$  where  $\hat{X}$  is an estimate of surveyed population quantity total and  $\hat{Y}$  is an estimate of the number of persons in the surveyed population contributing to that total quantity.

An example of a quantitative estimate is the average monthly amount paid in rent/housing costs. The numerator is an estimate of the total amount paid each month for the immigrants who live in dwelling units and the denominator is the number of immigrants who live in dwelling units.

#### Examples of Quantitative Questions

Q: How much do you or your family pay each month towards rent/housing costs? (Include heat, water, electricity, parking, etc., but exclude telephone.)

R: |\_|\_|\_|\_| \$/month

Q: In this job, what is/was your wage or salary before taxes or other deductions?

R: |\_|\_|\_|\_|\_| \$

### 12.3.1 *Tabulation of Categorical Estimates*

Estimates of the number of immigrants with a certain characteristic can be obtained from the microdata file by summing the final weights of all records possessing the characteristic(s) of interest. These estimates may be cross-sectional or longitudinal. Proportions and ratios of the form  $\hat{X}/\hat{Y}$  are obtained by:

- a) summing the final weights of records having the characteristic of interest for the numerator ( $\hat{X}$ ),
- b) summing the final weights of records having the characteristic of interest for the denominator ( $\hat{Y}$ ), then
- c) divide estimate a) by estimate b) ( $\hat{X}/\hat{Y}$ ).

### 12.3.2 *Tabulation of Quantitative Estimates*

Estimates of quantities can be obtained from the microdata file by multiplying the value of the variable of interest by the final weight for each record, then summing this quantity over all records of interest. For example, to obtain an estimate of the total amount paid monthly in housing costs, multiply the monthly amount of the immigrant's housing costs by the final weight for the record, then sum this value over all records for immigrants who live in dwelling units.

To obtain a weighted average of the form  $\hat{X}/\hat{Y}$ , the numerator ( $\hat{X}$ ) is calculated as for a quantitative estimate and the denominator ( $\hat{Y}$ ) is calculated as for a categorical estimate. For example, to estimate the average monthly amount paid for housing by immigrants living in dwelling units,

- a) estimate the total monthly amount paid in housing costs ( $\hat{X}$ ) as described above,
- b) estimate the number of immigrants who live in dwelling units ( $\hat{Y}$ ) by summing the final weights of all records for this category, then
- c) divide estimate a) by estimate b) ( $\hat{X} / \hat{Y}$ ).

## 12.4 Guidelines for Statistical Analysis

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

While many analysis procedures found in statistical packages allow weights to be used, the meaning or definition of the weight in these procedures differs from that which is appropriate in a sample survey framework, with the result that while in many cases the estimates produced by the packages are correct, the variance estimates that are calculated are poor. Approximate variances for simple estimates such as totals, proportions and ratios (for qualitative variables and for common domains) can be derived using the LSIC Coefficients of Variation Extraction Module, which is provided as a companion tool..

For other analysis techniques (for example, linear regression, logistic regression and analysis of variance), a method exists which can make the variances calculated by the standard packages more meaningful, by incorporating the unequal probabilities of selection. The method rescales the weights so that there is an average weight of 1.

For example, suppose that analysis of all male respondents is required. The steps to rescale the weights are as follows:

- 1) select all immigrants from the file who reported Ir1q008 = male;
- 2) calculate the AVERAGE weight for these records by summing the original person weights (WT1L) from the microdata file for these records and then dividing by the number of respondents who reported Ir1q008 = male;
- 3) for each of these respondents, calculate a RESCALED weight equal to the original person weight divided by the AVERAGE weight;
- 4) perform the analysis for these respondents using the RESCALED weight.

However, because the stratification and clustering of the sample's design are still not taken into account, the variance estimates calculated in this way are likely to be under-estimates.

The calculation of more precise variance estimates requires detailed knowledge of the design of the survey. Such detail cannot be given in this microdata file because of confidentiality. Variances that take the complete sample design into account can be calculated for many statistics by Statistics Canada on a cost-recovery basis.

## **12.5 Coefficient of Variation Release Guidelines**

Before releasing and/or publishing any estimate from the LSIC, users should first determine the quality level of the estimate. The quality levels are *acceptable*, *marginal* and *unacceptable*. As discussed in Chapter 11.0, sampling and non-sampling errors both influence data quality. For the purposes of this document, however, estimate quality is based solely on the sampling error illustrated by the coefficient of variation, as shown in the table below.

First, the number of immigrants who contribute to the calculation of the estimate should be determined. If this number is less than 10, the weighted estimate should be considered to be of unacceptable quality.

For weighted estimates based on sample sizes of 10 immigrants or more, users should determine the coefficient of variation of the estimate and follow the guidelines below. These quality level guidelines should be applied to weighted rounded estimates.

All estimates can be considered releasable. However, those of marginal or unacceptable quality level must be accompanied by a warning to caution subsequent users.

**Quality Level Guidelines**

| <b>Quality Level of Estimate</b> | <b>Guidelines</b>   |
|----------------------------------|---|
| 1) Acceptable                    | <p>Estimates have:<br/>a sample size of 10 or more, and<br/>low coefficients of variation in the range of 0.0% - 16.5%</p> <p>No warning is required.</p>   |
| 2) Marginal                      | <p>Estimates have:<br/>a sample size of 10 or more, and<br/>high coefficients of variation in the range of 16.6% - 33.3%.</p> <p>Estimates should be flagged with the letter M (or some similar identifier). They should be accompanied by a warning to caution subsequent users about the high levels of error, associated with the estimates.</p>   |
| 3) Unacceptable                  | <p>Estimates have:<br/>a sample size of less than 10, or<br/>very high coefficients of variation in excess of 33.3%.</p> <p>Statistics Canada recommends not to release estimates of unacceptable quality. However, if the user chooses to do so then estimates should be flagged with the letter U (or some similar identifier) and the following warning should accompany the estimates:</p> <p>"Please be warned that these estimates [flagged with the letter U] do not meet Statistics Canada's quality standards. Conclusions based on these data will be unreliable, and most likely invalid."</p> |

## 13.0 Variance Calculation

The Longitudinal Survey of Immigrants to Canada (LSIC) is a probabilistic survey, i.e. a sample has been selected to represent the target population. A given variability is inherent in any random selection. This variability is known as the sampling error, as described in Section 11.1. In addition, adjustments have been made to take into account non-responding and unresolved units which are part of the variability of the estimates. This chapter explains why it is important to calculate the variance and presents different tools to do so.

### 13.1 Importance of the Variance

The variance of an estimate is a good indicator of the quality of the indicator. A high variance estimate is considered unreliable. In order to quantify large variance, a relative measure of the variability is used, namely the coefficient of variation (CV). The coefficient of variation is defined as the ratio of the square root of the variance over the estimate. The square root of the variance is also known as a standard deviation. The coefficient of variation, as opposed to the variance, allows the analyst to compare estimates of different magnitudes along the same scale. As a result, it is possible to assess the quality of any estimate with the CV.

Most importantly variance or the CV is required for statistical tests such as hypothesis tests, which determine if two estimates are statistically different. Consequently, variance or CV calculation is mandatory.

#### Method to Obtain the Variance of an Estimate

It is almost impossible to derive an exact formula to calculate the variance for the LSIC due to the complex sample design, weight adjustments and post-stratification. A very good way to approximate the true variance is to use the replicate method, namely the bootstrap method. This method is based on a replicate technique and is known to correctly approximate the true value of the variance. A file containing 1,000 bootstrap weights is available. Variance calculation using 1,000 bootstrap weights involves calculating the estimates with each of these 1,000 weights and then, calculating the variance of these 1,000 estimates.

Two user-friendly tools, both using the bootstrap weights, have been developed to help users calculate the variance and the CVs for their estimates. These tools are:

- **Macros to calculate the variance**, using bootstrap weights programmed for SAS and STATA users.
- **An Excel based CV extraction module (CVEM)** for totals and proportions, which produces approximate CVs for a large number of domains.

The use of one or more of these tools depends on the type of analysis and the level of precision required.

### 13.2 SAS and STATA Macros to Calculate the Variance Using the Bootstrap Weights

SAS and STATA macros have been developed to calculate the variance using the bootstrap weights. Of the different methods available, this method produces the best approximation of the true variance. With this method, it is possible to calculate the variance of any estimate, for any domain. The variance calculated using this method

takes into account the sample design and the specifics of the variable of interest. Finally, as opposed to the other methods, the user is not restricted to pre-defined domains.

This method has many advantages but requires more work from the researcher. Variance calculation using these macros is more time consuming than the other method presented (i.e. CVEM). The user must first become familiar with the macros before using them. However, these macros have been developed in such a way that they are easy to use.

Despite the time required to run these macros, it is strongly recommended to use this method to calculate the variance of any estimates to be published. This method provides a more precise and accurate measure of the true variance.

### **13.3 Excel Based Coefficient of Variation Extraction Module**

The second tool available for users to obtain approximate coefficients of variation is the Excel based CV extraction module (CVEM). This application, developed with Excel macros and accessed through a user-friendly interface, allows user to extract the desired information in two ways. One is by describing the domain of interest with the nine available variables, and the other is by specifying the size of the domain. The information displayed consists of the proportion estimate, the number of respondents in the specified domain, the estimated population in that domain, basic statistics and the coefficient of variation for the selected proportion. Here, a domain is defined as being the cross-tabulation of the variables listed in the table in Section 13.3.1.

Over 32,500 domains are covered by the set of spreadsheets, giving an approximate CV for eight different proportions in each of the domains, for a total of over 260,000 CV's. Simulations were run to calculate variances, coefficients of variation and confidence intervals at the 95% level for different proportions, i.e. 1%, 5%, 10%, 15%, 20%, 30%, 40% and 50%. These proportions were based on population distribution. For a given repetition, the observed proportion in the random sample can be different from that of the targeted proportion. Therefore the mean of 100 repetitions was used to account for that variability.

#### **13.3.1 Statistics Canada Quality Standards**

Users should note that for disclosure issues, when using a dichotomous variable, both the sample size and the CV should be publishable simultaneously. Users should always ensure the quality of the estimates, especially for smaller proportions obtained from small domains. To help users identify high CVs, color coding has been used in the Excel application when displaying a CV. Using the markers described below, the colors used are red for CV's in excess of 33.3% and yellow for the ones in the range of 16.6% to 33.3%. More details are provided in the CVEM User's Guide. Below is a list of the variables available in the CVEM.

| Field                          | Description  |
|--------------------------------|--|
| Class of immigrant             |  |
| Age group                      |  |
| Geographical residence         |  |
| Place of birth                 |  |
| Gender                         |  |
| Marital status                 |  |
| Labour force status            |  |
| Highest level of education     |  |
| Knowledge of official language |  |
| Target proportion              | The theoretical proportion used to simulate a variable. Can take the values 1%, 5%, 10%, 15%, 20%, 30%, 40% or 50% |
| Yhat                           | The mean of 100 calculated proportions. This figure should be close to the target proportion.                      |
| N                              | The average sample size of the specified domain from 100 repetitions.  |
| Bs_var                         | The mean of 100 variances for the specified domain.  |
| Bs_sd                          | The mean of 100 standard errors for the specified domain.  |
| Cil95                          | The mean of 100 at the 95% confidence interval lower boundary.   |
| Ciu95                          | The mean of 100 at the 95% confidence interval upper boundary.   |

As a reference, the following quality standards should be used:

- 1) An estimate is said to be **acceptable** if it has a sample size of 10 or more and low coefficient of variation in the range of 0.0% to 16.5%.
- 2) An estimate is said to be **marginal** if it has a sample size of 10 or more and high coefficient of variation in the range of 16.6% to 33.3%. This estimate should be accompanied by a warning to caution subsequent users about the high level of error, associated with the estimate.
- 3) An estimate is said to be **unacceptable** if it has a sample size of less than 10 or very high coefficient of variation in excess of 33.3%. Statistics Canada recommends not to release estimates of unacceptable quality (see Section 12.5).

For more information see the publication *Statistics Canada Quality Guidelines*, Catalogue no. 12-539-XIE.

### **13.4 How to Derive the Coefficient of Variation for Categorical Estimates**

#### **Rule 1: Estimates of Number of Immigrants Possessing a Characteristic (Aggregates)**

The coefficient of variation depends only on the size of the estimate itself. It is safe to say that an estimate's CV is close (though slightly greater) than the proportion it represents. Hence, to get an approximation of an estimate's CV, users could use the CVEM by specifying the domain's size and deriving the appropriate proportion. For example, suppose we have an estimate  $Y_{hat} = 30,000$  individuals possessing a certain characteristic. If we are to compare them to the  $100,000$  people in the domain of interest, then the CV for  $Y_{hat}$  should be close to the proportion i.e.  $30,000 / 100,000 = 30.0\%$ . To have an exact CV, the programs that use the bootstrap weights should be used. Bootstrap programs are available for SAS and STATA users.

#### **Rule 2: Estimates of Proportions or Percentages of Immigrants Possessing a Characteristic**

The CV's calculated in the CVEM are for proportions. Hence, they can be used directly as they are given on the spreadsheet.

#### **Rule 3: Estimates of Differences Between Aggregates, Percentages and ratios**

To obtain the CV for a difference, the Bootstrap programs are best suited as there is no easy way to derive it from each of the individual CV's. The programs offer the possibility to derive CV's for differences of totals and ratios.

#### **Rule 4: Estimates of Ratios**

If the denominator of a ratio is considered as a "domain size", one can use the CVEM just as it is used in Rule 2. Otherwise, the Bootstrap programs can be used by defining properly the numerator and the denominator.

### **13.5 How to Use the Coefficient of Variation to Obtain Confidence Limits**

Although coefficients of variation are widely used, a more intuitively meaningful measure of sampling error is the confidence interval of an estimate. A confidence interval constitutes a statement on the level of confidence that the true value for the population lies within a specified range of values. For example, a 95% confidence interval can be described as follows:

If sampling of the population is repeated indefinitely, with each sample leading to a new confidence interval for an estimate, then in 95% of the samples, the interval will cover the true population value.

Using the standard error of an estimate, confidence intervals for estimates may be obtained under the assumption that under repeated sampling of the population, the various estimates obtained for a population characteristic are normally distributed about the true population value. Under this assumption, the chances are about 68 out of 100 that the difference between a sample estimate and the true population value would be less than one standard error, about 95 out of 100 that the difference would be less than two standard errors, and about



99 out of 100 that the differences would be less than three standard errors. These different degrees of confidence are referred to as the confidence levels.

Confidence intervals for an estimate,  $\hat{X}$ , are generally expressed as two numbers, one below the estimate and one above the estimate, as  $(\hat{X} - k, \hat{X} + k)$  where  $k$  is determined depending upon the level of confidence desired and the sampling error of the estimate.

The 95% confidence intervals for an estimate are available directly in the CV spreadsheet. If the user wants to determine other confidence intervals, the following formula will convert to a confidence interval ( $CI_{\hat{x}}$ ):

$$CI_{\hat{x}} = (\hat{X} - t\hat{X}\alpha_{\hat{x}}, \hat{X} + t\hat{X}\alpha_{\hat{x}})$$

where  $\alpha_{\hat{x}}$  is the determined coefficient of variation for  $\hat{X}$  and

- $t = 1$  if a 68% confidence interval is desired;
- $t = 1.6$  if a 90% confidence interval is desired;
- $t = 2.6$  if a 99% confidence interval is desired.

**Warning Note on Confidence Intervals**

Release guidelines which apply to the estimate also apply to the confidence interval. For example, if the estimate is “marginal”, then the confidence interval is marginal and should be accompanied by a warning note to caution subsequent users about high levels of error, associated with the estimate.

**Example of Using the Coefficient of Variation to Obtain Confidence Limits**

A 90% confidence interval for the estimated proportion of women having a university degree would be calculated as follows:

$$\hat{X} = 47.4\% \text{ (or expressed as a proportion 0.474)}$$

$$t = 1.6$$

$$\alpha_{\hat{x}} = 1.21\% \text{ (0.0121 expressed as a proportion) is the coefficient of variation of this estimate as derived using the bootstrap weights.}$$

$$CI_{\hat{x}} = \{0.474 - (1.6) (0.474) (0.0121), 0.474 + (1.6) (0.474) (0.0121)\}$$

$$CI_{\hat{x}} = \{0.474 - 0.009, 0.474 + 0.009\}$$

$$CI_{\hat{x}} = \{0.465, 0.483\}$$

Hence, with a 90% level of confidence, it can be said that between 46.5% and 48.3% of women have a university degree.

### 13.6 Hypothesis Testing (t-test)

Standard errors may also be used to perform hypothesis testing, a procedure for distinguishing between population parameters using sample estimates. The sample estimates can be numbers, averages, percentages, ratios, etc. Tests may be performed at various levels of significance, where a level of significance is the probability of concluding that the characteristics are different when, in fact, they are identical.

Let  $\hat{X}_1$  and  $\hat{X}_2$  be sample estimates for two characteristics of interest. The standard error for the difference  $\hat{X}_1 - \hat{X}_2$  can be obtained through the programs that use the bootstrap weights. Let the standard error on the difference be  $\sigma_{\hat{d}}$ .

$$\text{If } t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_{\hat{d}}}$$

is between -2 and 2, then no conclusion about the difference between the characteristics is justified at the 5% level of significance. If however, this ratio is smaller than -2 or larger than +2, the observed difference is significant at the 0.05 level. That is to say that the difference between the estimates is significant.

### 13.7 Coefficients of Variations for Quantitative Estimates

For quantitative estimates, special tables would have to be produced to determine their sampling error. Since most of the variables for the LSIC are primarily categorical in nature, this has not been done.

As a general rule, however, the coefficient of variation of a quantitative total will be larger than the coefficient of variation of the corresponding category estimate (i.e., the estimate of the number of persons contributing to the quantitative estimate). If the corresponding category estimate is not releasable, the quantitative estimate will not be either. For example, the coefficient of variation of the total number of hours of class for women attending university courses would be greater than the coefficient of variation of the corresponding proportion of women attending university courses. Hence if the coefficient of variation of the proportion is not releasable, then the coefficient of variation of the corresponding quantitative estimate will also not be releasable.

#### Pseudo Replication

Coefficients of variation of such estimates can be derived as required for a specific estimate using a technique known as pseudo replication. This involves dividing the records on the microdata files into subgroups (or replicates) and determining the variation in the estimate from replicate to replicate. Users wishing to derive coefficients of variation for quantitative estimates may contact Statistics Canada for advice on the allocation of records to appropriate replicates and the formulae to be used in these calculations.

### 13.8 Approximate Quality Release Cut-offs

The tables below provide the approximate release cut-offs for two selected domains. These population estimates provide a rough indication of acceptable, marginal and unacceptable domain sizes. They are meant to be used as approximate guidelines only. Users are still responsible to calculate precise CVs before releasing results. The use of the CVEM is strongly recommended for better precision.

### Approximate Release Cut-offs by Class of Immigrant

| <b>Class of Immigrants</b> | <b>Acceptable<br/>CV 0.0% - 16.5%</b> | <b>Marginal<br/>CV 16.6% - 33.3%</b> | <b>Unacceptable<br/>CV &gt; 33.3%</b> |
|----------------------------|---------------------------------------|--------------------------------------|---------------------------------------|
| Family                     | 600 & over                            | 150 to < 600                         | under 150                             |
| Economic                   | 450 & over                            | 125 to < 450                         | under 125                             |
| Refugees                   | 210 & over                            | 60 to < 210                          | under 60                              |
| <b>Total</b>               | <b>400 &amp; over</b>                 | <b>110 to &lt; 400</b>               | <b>under 110</b>                      |

### Approximate Release Cut-offs by Geographical Regions

| <b>Province</b>  | <b>Acceptable<br/>CV 0.0% - 16.5%</b> | <b>Marginal<br/>CV 16.6% - 33.3%</b> | <b>Unacceptable<br/>CV &gt; 33.3%</b> |
|------------------|---------------------------------------|--------------------------------------|---------------------------------------|
| Québec           | 450 & over                            | 120 to < 450                         | under 120                             |
| Ontario          | 475 & over                            | 120 to < 475                         | under 120                             |
| Alberta          | 310 & over                            | 70 to < 310                          | under 70                              |
| British Columbia | 400 & over                            | 140 to < 400                         | under 140                             |
| Other            | 300 & over                            | 150 to < 300                         | under 150                             |
| <b>Canada</b>    | <b>400 &amp; over</b>                 | <b>110 to &lt; 400</b>               | <b>under 110</b>                      |



## **14.0 Record Layout with Univariate Frequencies**

**\*Available in the Research Data Centres only.**



## Appendix I

### Industry Codes

#### North American Industry Classification System (NAICS) 1997

#### Industry groups:

|      |  |      |   |
|------|--|------|---|
| 1111 | Oilseed and Grain Farming                                | 2213 | Water, Sewage and Other Systems                                 |
| 1112 | Vegetable and Melon Farming                              | 2311 | Land Subdivision and Land Development                           |
| 1113 | Fruit and Tree Nut Farming                               | 2312 | Building Construction   |
| 1114 | Greenhouse, Nursery and Floriculture Production          | 2313 | Engineering Construction  |
| 1119 | Other Crop Farming                                       | 2314 | Construction Management   |
| 1121 | Cattle Ranching and Farming                              | 2321 | Site Preparation Work   |
| 1122 | Hog and Pig Farming                                      | 2322 | Building Structure Work   |
| 1123 | Poultry and Egg Production                               | 2323 | Building Exterior Finishing Work                                |
| 1124 | Sheep and Goat Farming                                   | 2324 | Building Interior Finishing Work                                |
| 1125 | Animal Aquaculture                                       | 2325 | Building Equipment Installation                                 |
| 1129 | Other Animal Production                                  | 2329 | Other Special Trade Contracting                                 |
| 1131 | Timber Tract Operations                                  | 3111 | Animal Food Manufacturing                                       |
| 1132 | Forest Nurseries and Gathering of Forest Products        | 3112 | Grain and Oilseed Milling                                       |
| 1133 | Logging  | 3113 | Sugar and Confectionery Product Manufacturing                   |
| 1141 | Fishing  | 3114 | Fruit and Vegetable Preserving and Specialty Food Manufacturing |
| 1142 | Hunting and Trapping                                     | 3115 | Dairy Product Manufacturing                                     |
| 1151 | Support Activities for Crop Production                   | 3116 | Meat Product Manufacturing                                      |
| 1152 | Support Activities for Animal Production                 | 3117 | Seafood Product Preparation and Packaging                       |
| 1153 | Support Activities for Forestry                          | 3118 | Bakeries and Tortilla Manufacturing                             |
| 2111 | Oil and Gas Extraction                                   | 3119 | Other Food Manufacturing  |
| 2121 | Coal Mining  | 3121 | Beverage Manufacturing  |
| 2122 | Metal Ore Mining   | 3122 | Tobacco Manufacturing   |
| 2123 | Non-Metallic Mineral Mining and Quarrying                | 3131 | Fibre, Yarn and Thread Mills                                    |
| 2131 | Support Activities for Mining and Oil and Gas Extraction | 3132 | Fabric Mills  |
| 2211 | Electric Power Generation, Transmission and Distribution | 3133 | Textile and Fabric Finishing and Fabric Coating                 |
| 2212 | Natural Gas Distribution                                 | 3141 | Textile Furnishings Mills                                       |
|      |  | 3149 | Other Textile Product Mills                                     |

---

|      |  |      |   |
|------|--|------|---|
| 3151 | Clothing Knitting Mills  | 3311 | Iron and Steel Mills and Ferro-Alloy Manufacturing  |
| 3152 | Cut and Sew Clothing Manufacturing   | 3312 | Steel Product Manufacturing from Purchased Steel  |
| 3159 | Clothing Accessories and Other Clothing Manufacturing                                    | 3313 | Alumina and Aluminum Production and Processing  |
| 3161 | Leather and Hide Tanning and Finishing   | 3314 | Non-Ferrous Metal (except Aluminum) Production and Processing                               |
| 3162 | Footwear Manufacturing   | 3315 | Foundries   |
| 3169 | Other Leather and Allied Product Manufacturing   | 3321 | Forging and Stamping  |
| 3211 | Sawmills and Wood Preservation   | 3322 | Cutlery and Hand Tool Manufacturing   |
| 3212 | Veneer, Plywood and Engineered Wood Product Manufacturing                                | 3323 | Architectural and Structural Metals Manufacturing   |
| 3219 | Other Wood Product Manufacturing   | 3324 | Boiler, Tank and Shipping Container Manufacturing   |
| 3221 | Pulp, Paper and Paperboard Mills   | 3325 | Hardware Manufacturing  |
| 3222 | Converted Paper Product Manufacturing  | 3326 | Spring and Wire Product Manufacturing   |
| 3231 | Printing and Related Support Activities  | 3327 | Machine Shops, Turned Product, and Screw, Nut and Bolt Manufacturing                        |
| 3241 | Petroleum and Coal Products Manufacturing  | 3328 | Coating, Engraving, Heat Treating and Allied Activities                                     |
| 3251 | Basic Chemical Manufacturing   | 3329 | Other Fabricated Metal Product Manufacturing  |
| 3252 | Resin, Synthetic Rubber, and Artificial and Synthetic Fibres and Filaments Manufacturing | 3331 | Agricultural, Construction and Mining Machinery Manufacturing                               |
| 3253 | Pesticide, Fertilizer and Other Agricultural Chemical Manufacturing                      | 3332 | Industrial Machinery Manufacturing  |
| 3254 | Pharmaceutical and Medicine Manufacturing  | 3333 | Commercial and Service Industry Machinery Manufacturing                                     |
| 3255 | Paint, Coating and Adhesive Manufacturing  | 3334 | Ventilation, Heating, Air-Conditioning and Commercial Refrigeration Equipment Manufacturing |
| 3256 | Soap, Cleaning Compound and Toilet Preparation Manufacturing                             | 3335 | Metalworking Machinery Manufacturing  |
| 3259 | Other Chemical Product Manufacturing   | 3336 | Engine, Turbine and Power Transmission Equipment Manufacturing                              |
| 3261 | Plastic Product Manufacturing  | 3339 | Other General-Purpose Machinery Manufacturing   |
| 3262 | Rubber Product Manufacturing   | 3341 | Computer and Peripheral Equipment Manufacturing   |
| 3271 | Clay Product and Refractory Manufacturing  | 3342 | Communications Equipment Manufacturing  |
| 3272 | Glass and Glass Product Manufacturing  | 3343 | Audio and Video Equipment Manufacturing   |
| 3273 | Cement and Concrete Product Manufacturing  |      |   |
| 3274 | Lime and Gypsum Product Manufacturing  |      |   |
| 3279 | Other Non-Metallic Mineral Product Manufacturing   |      |   |

---



---

|      |  |      |   |
|------|--|------|---|
| 3344 | Semiconductor and Other Electronic Component Manufacturing                   | 4143 | Home Furnishings Wholesaler-Distributors  |
| 3345 | Navigational, Measuring, Medical and Control Instruments Manufacturing       | 4144 | Personal Goods Wholesaler-Distributors  |
| 3346 | Manufacturing and Reproducing Magnetic and Optical Media                     | 4145 | Pharmaceuticals, Toiletries, Cosmetics and Sundries Wholesaler-Distributors                           |
| 3351 | Electric Lighting Equipment Manufacturing                                    | 4151 | Motor Vehicle Wholesaler-Distributors   |
| 3352 | Household Appliance Manufacturing  | 4152 | New Motor Vehicle Parts and Accessories Wholesaler-Distributors                                       |
| 3353 | Electrical Equipment Manufacturing   | 4153 | Used Motor Vehicle Parts and Accessories Wholesaler-Distributors                                      |
| 3359 | Other Electrical Equipment and Component Manufacturing                       | 4161 | Electrical, Plumbing, Heating and Air-Conditioning Equipment and Supplies Wholesaler-Distributors     |
| 3361 | Motor Vehicle Manufacturing  | 4162 | Metal Service Centres   |
| 3362 | Motor Vehicle Body and Trailer Manufacturing                                 | 4163 | Lumber, Millwork, Hardware and Other Building Supplies Wholesaler-Distributors                        |
| 3363 | Motor Vehicle Parts Manufacturing  | 4171 | Farm, Lawn and Garden Machinery and Equipment Wholesaler-Distributors                                 |
| 3364 | Aerospace Product and Parts Manufacturing                                    | 4172 | Construction, Forestry, Mining, and Industrial Machinery, Equip. and Supplies Wholesaler-Distributors |
| 3365 | Railroad Rolling Stock Manufacturing   | 4173 | Computer and Communications Equipment and Supplies Wholesaler-Distributors                            |
| 3366 | Ship and Boat Building   | 4179 | Other Machinery, Equipment and Supplies Wholesaler-Distributors                                       |
| 3369 | Other Transportation Equipment Manufacturing                                 | 4181 | Recyclable Material Wholesaler-Distributors   |
| 3371 | Household and Institutional Furniture and Kitchen Cabinet Manufacturing      | 4182 | Paper, Paper Product and Disposable Plastic Product Wholesaler-Distributors                           |
| 3372 | Office Furniture (including Fixtures) Manufacturing                          | 4183 | Agricultural Supplies Wholesaler-Distributors   |
| 3379 | Other Furniture-Related Product Manufacturing                                | 4184 | Chemical (except Agricultural) and Allied Product Wholesaler-Distributors                             |
| 3391 | Medical Equipment and Supplies Manufacturing                                 | 4189 | Other Miscellaneous Wholesaler-Distributors   |
| 3399 | Other Miscellaneous Manufacturing  | 4191 | Wholesale Agents and Brokers  |
| 4111 | Farm Product Wholesaler-Distributors   | 4411 | Automobile Dealers  |
| 4121 | Petroleum Product Wholesaler-Distributors                                    | 4412 | Other Motor Vehicle Dealers   |
| 4131 | Food Wholesaler-Distributors   | 4413 | Automotive Parts, Accessories and Tire Stores   |
| 4132 | Beverage Wholesaler-Distributors   | 4421 | Furniture Stores  |
| 4133 | Cigarette and Tobacco Product Wholesaler-Distributors                        | 4422 | Home Furnishings Stores   |
| 4141 | Textile, Clothing and Footwear Wholesaler-Distributors                       |      |   |
| 4142 | Home Entertainment Equipment and Household Appliance Wholesaler-Distributors |      |   |

---

---

|      |  |      |   |
|------|--|------|---|
| 4431 | Electronics and Appliance Stores                       | 4854 | School and Employee Bus Transportation              |
| 4441 | Building Material and Supplies Dealers                 | 4855 | Charter Bus Industry                                |
| 4442 | Lawn and Garden Equipment and Supplies Stores          | 4859 | Other Transit and Ground Passenger Transportation   |
| 4451 | Grocery Stores   | 4861 | Pipeline Transportation of Crude Oil                |
| 4452 | Specialty Food Stores                                  | 4862 | Pipeline Transportation of Natural Gas              |
| 4453 | Beer, Wine and Liquor Stores                           | 4869 | Other Pipeline Transportation                       |
| 4461 | Health and Personal Care Stores                        | 4871 | Scenic and Sightseeing Transportation, Land         |
| 4471 | Gasoline Stations                                      | 4872 | Scenic and Sightseeing Transportation, Water        |
| 4481 | Clothing Stores  | 4879 | Scenic and Sightseeing Transportation, Other        |
| 4482 | Shoe Stores  | 4881 | Support Activities for Air Transportation           |
| 4483 | Jewellery, Luggage and Leather Goods Stores            | 4882 | Support Activities for Rail Transportation          |
| 4511 | Sporting Goods, Hobby and Musical Instrument Stores    | 4883 | Support Activities for Water Transportation         |
| 4512 | Book, Periodical and Music Stores                      | 4884 | Support Activities for Road Transportation          |
| 4521 | Department Stores                                      | 4885 | Freight Transportation Arrangement                  |
| 4529 | Other General Merchandise Stores                       | 4889 | Other Support Activities for Transportation         |
| 4531 | Florists   | 4911 | Postal Service                                      |
| 4532 | Office Supplies, Stationery and Gift Stores            | 4921 | Couriers  |
| 4533 | Used Merchandise Stores                                | 4922 | Local Messengers and Local Delivery                 |
| 4539 | Other Miscellaneous Store Retailers                    | 4931 | Warehousing and Storage                             |
| 4541 | Electronic Shopping and Mail-Order Houses              | 5111 | Newspaper, Periodical, Book and Database Publishers |
| 4542 | Vending Machine Operators                              | 5112 | Software Publishers                                 |
| 4543 | Direct Selling Establishments                          | 5121 | Motion Picture and Video Industries                 |
| 4811 | Scheduled Air Transportation                           | 5122 | Sound Recording Industries                          |
| 4812 | Non-Scheduled Air Transportation                       | 5131 | Radio and Television Broadcasting                   |
| 4821 | Rail Transportation                                    | 5132 | Pay TV, Specialty TV and Program Distribution       |
| 4831 | Deep Sea, Coastal and Great Lakes Water Transportation | 5133 | Telecommunications                                  |
| 4832 | Inland Water Transportation                            | 5141 | Information Services                                |
| 4841 | General Freight Trucking                               | 5142 | Data Processing Services                            |
| 4842 | Specialized Freight Trucking                           | 5211 | Monetary Authorities - Central Bank                 |
| 4851 | Urban Transit Systems                                  |      |   |
| 4852 | Interurban and Rural Bus Transportation                |      |   |
| 4853 | Taxi and Limousine Service                             |      |   |

---

---

|      |   |      |  |
|------|---|------|--|
| 5221 | Depository Credit Intermediation                                      | 5611 | Office Administrative Services   |
| 5222 | Non-Depository Credit Intermediation                                  | 5612 | Facilities Support Services  |
| 5223 | Activities Related to Credit Intermediation                           | 5613 | Employment Services  |
| 5231 | Securities and Commodity Contracts Intermediation and Brokerage       | 5614 | Business Support Services  |
| 5232 | Securities and Commodity Exchanges                                    | 5615 | Travel Arrangement and Reservation Services                                      |
| 5239 | Other Financial Investment Activities                                 | 5616 | Investigation and Security Services  |
| 5241 | Insurance Carriers  | 5617 | Services to Buildings and Dwellings  |
| 5242 | Agencies, Brokerages and Other Insurance Related Activities           | 5619 | Other Support Services   |
| 5261 | Pension Funds   | 5621 | Waste Collection   |
| 5269 | Other Funds and Financial Vehicles                                    | 5622 | Waste Treatment and Disposal   |
| 5311 | Lessors of Real Estate  | 5629 | Remediation and Other Waste Management Services                                  |
| 5312 | Offices of Real Estate Agents and Brokers                             | 6111 | Elementary and Secondary Schools   |
| 5313 | Activities Related to Real Estate                                     | 6112 | Community Colleges and C.E.G.E.P.s   |
| 5321 | Automotive Equipment Rental and Leasing                               | 6113 | Universities   |
| 5322 | Consumer Goods Rental   | 6114 | Business Schools and Computer and Management Training                            |
| 5323 | General Rental Centres  | 6115 | Technical and Trade Schools  |
| 5324 | Commercial and Industrial Machinery and Equipment Rental and Leasing  | 6116 | Other Schools and Instruction  |
| 5331 | Lessors of Non-Financial Intangible Assets (Except Copyrighted Works) | 6117 | Educational Support Services   |
| 5411 | Legal Services  | 6211 | Offices of Physicians  |
| 5412 | Accounting, Tax Preparation, Bookkeeping and Payroll Services         | 6212 | Offices of Dentists  |
| 5413 | Architectural, Engineering and Related Services                       | 6213 | Offices of Other Health Practitioners  |
| 5414 | Specialized Design Services   | 6214 | Out-Patient Care Centres   |
| 5415 | Computer Systems Design and Related Services                          | 6215 | Medical and Diagnostic Laboratories  |
| 5416 | Management, Scientific and Technical Consulting Services              | 6216 | Home Health Care Services  |
| 5417 | Scientific Research and Development Services                          | 6219 | Other Ambulatory Health Care Services  |
| 5418 | Advertising and Related Services                                      | 6221 | General Medical and Surgical Hospitals   |
| 5419 | Other Professional, Scientific and Technical Services                 | 6222 | Psychiatric and Substance Abuse Hospitals  |
| 5511 | Management of Companies and Enterprises                               | 6223 | Specialty (except Psychiatric and Substance Abuse) Hospitals                     |
|      |   | 6231 | Nursing Care Facilities  |
|      |   | 6232 | Residential Developmental Handicap, Mental Health and Substance Abuse Facilities |
|      |   | 6233 | Community Care Facilities for the Elderly  |
|      |   | 6239 | Other Residential Care Facilities  |

---

---

|      |   |      |  |
|------|---|------|--|
| 6241 | Individual and Family Services  | 8132 | Grant-Making and Giving Services                                     |
| 6242 | Community Food and Housing, and<br>Emergency and Other Relief Services  | 8133 | Social Advocacy Organizations  |
| 6243 | Vocational Rehabilitation Services  | 8134 | Civic and Social Organizations                                       |
| 6244 | Child Day-Care Services   | 8139 | Business, Professional, Labour and<br>Other Membership Organizations |
| 7111 | Performing Arts Companies   | 8141 | Private Households   |
| 7112 | Spectator Sports  | 9111 | Defence Services   |
| 7113 | Promoters (Presenters) of Performing<br>Arts, Sports and Similar Events   | 9112 | Federal Protective Services  |
| 7114 | Agents and Managers for Artists,<br>Athletes, Entertainers and Other Public<br>Figures                            | 9113 | Federal Labour, Employment and<br>Immigration Services               |
| 7115 | Independent Artists, Writers and<br>Performers  | 9114 | Foreign Affairs and International<br>Assistance                      |
| 7121 | Heritage Institutions   | 9119 | Other Federal Government Public<br>Administration                    |
| 7131 | Amusement Parks and Arcades   | 9121 | Provincial Protective Services                                       |
| 7132 | Gambling Industries   | 9122 | Provincial Labour and Employment<br>Services                         |
| 7139 | Other Amusement and Recreation<br>Industries  | 9129 | Other Provincial and Territorial Public<br>Administration            |
| 7211 | Traveller Accommodation   | 9131 | Municipal Protective Services  |
| 7212 | RV (Recreational Vehicle) Parks and<br>Recreational Camps   | 9139 | Other Local, Municipal and Regional<br>Public Administration         |
| 7213 | Rooming and Boarding Houses   | 9141 | Aboriginal Public Administration                                     |
| 7221 | Full-Service Restaurants  | 9191 | International and Other Extra-Territorial<br>Public Administration   |
| 7222 | Limited-Service Eating Places   |      |  |
| 7223 | Special Food Services   |      |  |
| 7224 | Drinking Places (Alcoholic Beverages)   |      |  |
| 8111 | Automotive Repair and Maintenance   |      |  |
| 8112 | Electronic and Precision Equipment<br>Repair and Maintenance  |      |  |
| 8113 | Commercial and Industrial Machinery<br>and Equipment (except Automotive and<br>Electronic) Repair and Maintenance |      |  |
| 8114 | Personal and Household Goods Repair<br>and Maintenance  |      |  |
| 8121 | Personal Care Services  |      |  |
| 8122 | Funeral Services  |      |  |
| 8123 | Dry Cleaning and Laundry Services   |      |  |
| 8129 | Other Personal Services   |      |  |
| 8131 | Religious Organizations   |      |  |

---

## Appendix II

### Occupation Codes

#### Standard Occupational Classification (SOC) 1991

**Unit groups:**

|      |   |      |   |
|------|---|------|---|
| A011 | Legislators   | A303 | Other Business Services Managers  |
| A012 | Senior Government Managers and Officials  | A311 | Telecommunication Carriers Managers   |
| A013 | Senior Managers - Financial, Communications Carriers and Other Business Services                | A312 | Postal and Courier Services Managers  |
| A014 | Senior Managers - Health, Education, Social and Community Services and Membership Organizations | A321 | Managers in Health Care   |
| A015 | Senior Managers - Trade, Broadcasting and Other Services, n.e.c.                                | A322 | Administrators in Post-Secondary Education and Vocational Training                      |
| A016 | Senior Managers - Goods Production, Utilities, Transportation and Construction                  | A323 | School Principals and Administrators of Elementary and Secondary Education              |
| A111 | Financial Managers  | A324 | Managers in Social, Community and Correctional Services                                 |
| A112 | Human Resources Managers  | A331 | Government Managers in Health and Social Policy Development and Program Administration  |
| A113 | Purchasing Managers   | A332 | Government Managers in Economic Analysis, Policy Development and Program Administration |
| A114 | Other Administrative Services Managers  | A333 | Government Managers in Education Policy Development and Program Administration          |
| A121 | Engineering, Science and Architecture Managers  | A334 | Other Managers in Public Administration   |
| A122 | Information Systems and Data Processing Managers  | A341 | Library, Archive, Museum and Art Gallery Managers                                       |
| A131 | Sales, Marketing and Advertising Managers   | A342 | Managers in Publishing, Motion Pictures, Broadcasting and Performing Arts               |
| A141 | Facility Operation and Maintenance Managers   | A343 | Recreation and Sport Program and Service Directors                                      |
| A211 | Retail Trade Managers   | A351 | Commissioned Police Officers  |
| A221 | Restaurant and Food Service Managers  | A352 | Fire Chiefs and Senior Fire-fighting Officers   |
| A222 | Accommodation Service Managers  | A353 | Commissioned Officers, Armed Forces   |
| A301 | Insurance, Real Estate and Financial Brokerage Managers   | A361 | Other Services Managers   |
| A302 | Banking, Credit and Other Investment Managers   | A371 | Construction Managers   |
|      |   | A372 | Residential Home Builders and Renovators  |

---

|      |   |      |   |
|------|---|------|---|
| A373 | Transportation Managers                                       | B413 | Supervisors, Library, Correspondence and Related Information Clerks |
| A381 | Primary Production Managers (except Agriculture)              | B414 | Supervisors, Mail and Message Distribution Occupations              |
| A391 | Manufacturing Managers  | B415 | Supervisors, Recording, Distributing and Scheduling Occupations     |
| A392 | Utilities Managers  | B511 | General Office Clerks   |
| B011 | Financial Auditors and Accountants                            | B512 | Typists and Word Processing Operators                               |
| B012 | Financial and Investment Analysts                             | B513 | Records and File Clerks   |
| B013 | Securities Agents, Investment Dealers and Traders             | B514 | Receptionists and Switchboard Operators                             |
| B014 | Other Financial Officers                                      | B521 | Computer Operators  |
| B021 | Specialists in Human Resources                                | B522 | Data Entry Clerks   |
| B022 | Professional Occupations in Business Services to Management   | B523 | Typesetters and Related Occupations                                 |
| B111 | Bookkeepers   | B524 | Telephone Operators   |
| B112 | Loan Officers   | B531 | Accounting and Related Clerks                                       |
| B113 | Insurance Adjusters and Claims Examiners                      | B532 | Payroll Clerks  |
| B114 | Insurance Underwriters  | B533 | Tellers, Financial Services   |
| B115 | Assessors, Valuers and Appraisers                             | B534 | Banking, Insurance and Other Financial Clerks                       |
| B116 | Customs, Ship and Other Brokers                               | B535 | Collectors  |
| B211 | Secretaries (except Legal and Medical)                        | B541 | Administrative Clerks   |
| B212 | Legal Secretaries   | B542 | Personnel Clerks  |
| B213 | Medical Secretaries   | B543 | Court Clerks  |
| B214 | Court Recorders and Medical Transcriptionists                 | B551 | Library Clerks  |
| B311 | Administrative Officers                                       | B552 | Correspondence, Publication and Related Clerks                      |
| B312 | Executive Assistants  | B553 | Customer Service, Information and Related Clerks                    |
| B313 | Personnel and Recruitment Officers                            | B554 | Survey Interviewers and Statistical Clerks                          |
| B314 | Property Administrators                                       | B561 | Mail, Postal and Related Clerks                                     |
| B315 | Purchasing Agents and Officers                                | B562 | Letter Carriers   |
| B316 | Conference and Event Planners                                 | B563 | Couriers and Messengers   |
| B317 | Court Officers and Justices of the Peace                      | B571 | Shippers and Receivers  |
| B318 | Immigration, Unemployment Insurance and Revenue Officers      | B572 | Storekeepers and Parts Clerks                                       |
| B411 | Supervisors, General Office and Administrative Support Clerks | B573 | Production Clerks   |
| B412 | Supervisors, Finance and Insurance Clerks                     | B574 | Purchasing and Inventory Clerks                                     |
|      |   | B575 | Dispatchers and Radio Operators                                     |

---

---

|      |   |      |  |
|------|---|------|--|
| B576 | Transportation Route and Crew Schedulers                  | C121 | Biological Technologists and Technicians   |
| C011 | Physicists and Astronomers                                | C122 | Agricultural and Fish Products Inspectors  |
| C012 | Chemists  | C123 | Forestry Technologists and Technicians   |
| C013 | Geologists, Geochemists and Geophysicists                 | C124 | Conservation and Fishery Officers  |
| C014 | Meteorologists  | C125 | Landscape and Horticultural Technicians and Specialists                            |
| C015 | Other Professional Occupations in Physical Sciences       | C131 | Civil Engineering Technologists and Technicians and Construction Estimators        |
| C021 | Biologists and Related Scientists                         | C132 | Mechanical Engineering Technologists and Technicians                               |
| C022 | Forestry Professionals                                    | C133 | Industrial Engineering and Manufacturing Technologists and Technicians             |
| C023 | Agricultural Representatives, Consultants and Specialists | C141 | Electrical and Electronics Engineering Technologists and Technicians               |
| C031 | Civil Engineers   | C142 | Electronic Service Technicians (Household and Business Equipment)                  |
| C032 | Mechanical Engineers                                      | C143 | Industrial Instrument Technicians and Mechanics                                    |
| C033 | Electrical and Electronics Engineers                      | C144 | Aircraft Instrument, Electrical and Avionics Mechanics, Technicians and Inspectors |
| C034 | Chemical Engineers  | C151 | Architectural Technologists and Technicians  |
| C041 | Industrial and Manufacturing Engineers                    | C152 | Industrial Designers   |
| C042 | Metallurgical and Materials Engineers                     | C153 | Drafting Technologists and Technicians   |
| C043 | Mining Engineers  | C154 | Survey Technologists and Technicians   |
| C044 | Geological Engineers                                      | C155 | Mapping and Related Technologists and Technicians                                  |
| C045 | Petroleum Engineers                                       | C161 | Nondestructive Testers and Inspectors  |
| C046 | Aerospace Engineers                                       | C162 | Engineering Inspectors and Regulatory Officers                                     |
| C047 | Computer Engineers  | C163 | Inspectors in Public and Environmental Health and Occupational Health and Safety   |
| C048 | Other Professional Engineers, n.e.c.                      | C164 | Construction Inspectors  |
| C051 | Architects  | C171 | Air Pilots, Flight Engineers and Flying Instructors                                |
| C052 | Landscape Architects                                      | C172 | Air Traffic Control Occupations  |
| C053 | Urban and Land Use Planners                               | C173 | Deck Officers, Water Transport   |
| C054 | Land Surveyors  |      |  |
| C061 | Mathematicians, Statisticians and Actuaries               |      |  |
| C062 | Computer Systems Analysts                                 |      |  |
| C063 | Computer Programmers                                      |      |  |
| C111 | Applied Chemical Technologists and Technicians            |      |  |
| C112 | Geological and Mineral Technologists and Technicians      |      |  |
| C113 | Meteorological Technicians                                |      |  |

---

---

|      |  |      |  |
|------|--|------|--|
| C174 | Engineer Officers, Water Transport                                 | D232 | Midwives and Practitioners of Natural Healing                                    |
| C175 | Railway and Marine Traffic Controllers                             | D233 | Registered Nursing Assistants  |
| D011 | Specialist Physicians  | D234 | Ambulance Attendants and Other Paramedical Occupations                           |
| D012 | General Practitioners and Family Physicians                        | D235 | Other Technical Occupations in Therapy and Assessment                            |
| D013 | Dentists   | D311 | Dental Assistants  |
| D014 | Veterinarians  | D312 | Nurse Aides and Orderlies  |
| D021 | Optometrists   | D313 | Other Aides and Assistants in Support of Health Services                         |
| D022 | Chiropractors  | E011 | Judges   |
| D023 | Other Professional Occupations in Health Diagnosing and Treating   | E012 | Lawyers and Quebec Notaries  |
| D031 | Pharmacists  | E021 | Psychologists  |
| D032 | Dietitians and Nutritionists                                       | E022 | Social Workers   |
| D041 | Audiologists and Speech-Language Pathologists                      | E023 | Family, Marriage and Other Related Counsellors                                   |
| D042 | Physiotherapists   | E024 | Ministers of Religion  |
| D043 | Occupational Therapists  | E025 | Probation and Parole Officers and Related Occupations                            |
| D044 | Other Professional Occupations in Therapy and Assessment           | E031 | Natural and Applied Science Policy Researchers, Consultants and Program Officers |
| D111 | Head Nurses and Supervisors  | E032 | Economists and Economic Policy Researchers and Analysts                          |
| D112 | Registered Nurses  | E033 | Economic Development Officers and Marketing Researchers and Consultants          |
| D211 | Medical Laboratory Technologists and Pathologists' Assistants      | E034 | Health and Social Policy Researchers, Consultants and Program Officers           |
| D212 | Medical Laboratory Technicians                                     | E035 | Education Policy Researchers, Consultants and Program Officers                   |
| D213 | Animal Health Technologists  | E036 | Recreation and Sports Program Supervisors and Consultants                        |
| D214 | Respiratory Therapists and Clinical Perfusionists                  | E037 | Program Officers Unique to Government  |
| D215 | Medical Radiation Technologists                                    | E038 | Other Professional Occupations in Social Science                                 |
| D216 | Medical Sonographers   | E111 | University Professors  |
| D217 | Cardiology Technologists   | E112 | Post-Secondary Teaching and Research Assistants                                  |
| D218 | Electroencephalographic and Other Diagnostic Technologists, n.e.c. | E121 | College and Other Vocational Instructors   |
| D219 | Other Medical Technologists and Technicians (except Dental Health) | E131 | Secondary School Teachers  |
| D221 | Denturists   |      |  |
| D222 | Dental Hygienists and Dental Therapists                            |      |  |
| D223 | Dental Technicians and Laboratory Bench Workers                    |      |  |
| D231 | Opticians  |      |  |

---



---

|      |  |      |  |
|------|--|------|--|
| E132 | Elementary School and Kindergarten Teachers  | F127 | Support and Assisting Occupations in Motion Pictures, Broadcasting and the Performing Arts |
| E133 | School and Guidance Counsellors  | F131 | Announcers and Other Broadcasters  |
| E211 | Paralegal and Related Occupations  | F132 | Other Performers   |
| E212 | Community and Social Service Workers   | F141 | Graphic Designers and Illustrating Artists   |
| E213 | Employment Counsellors   | F142 | Interior Designers   |
| E214 | Instructors and Teachers of Disabled Persons   | F143 | Theatre, Fashion, Exhibit and Other Creative Designers                                     |
| E215 | Other Instructors  | F144 | Artisans and Craftspersons   |
| E216 | Other Religious Occupations  | F145 | Patternmakers - Textile, Leather and Fur Products  |
| F011 | Librarians   | F151 | Athletes   |
| F012 | Conservators and Curators  | F152 | Coaches  |
| F013 | Archivists   | F153 | Sports Officials and Referees  |
| F021 | Writers  | F154 | Program Leaders and Instructors in Recreation and Sport                                    |
| F022 | Editors  | G011 | Retail Trade Supervisors   |
| F023 | Journalists  | G012 | Food Service Supervisors   |
| F024 | Professional Occupations in Public Relations and Communications                      | G013 | Executive Housekeepers   |
| F025 | Translators, Terminologists and Interpreters   | G014 | Dry Cleaning and Laundry Supervisors   |
| F031 | Producers, Directors, Choreographers and Related Occupations                         | G015 | Cleaning Supervisors   |
| F032 | Conductors, Composers and Arrangers  | G016 | Other Service Supervisors  |
| F033 | Musicians and Singers  | G111 | Sales Representatives, Wholesale Trade (Non-Technical)                                     |
| F034 | Dancers  | G121 | Technical Sales Specialists, Wholesale Trade   |
| F035 | Actors   | G131 | Insurance Agents and Brokers   |
| F036 | Painters, Sculptors and Other Visual Artists   | G132 | Real Estate Agents and Salespersons  |
| F111 | Library and Archive Technicians and Assistants                                       | G133 | Retail and Wholesale Buyers  |
| F112 | Technical Occupations Related to Museums and Galleries                               | G134 | Grain Elevator Operators   |
| F121 | Photographers  | G211 | Retail Salespersons and Sales Clerks   |
| F122 | Film and Video Camera Operators  | G311 | Cashiers   |
| F123 | Graphic Arts Technicians   | G411 | Chefs  |
| F124 | Broadcast Technicians  | G412 | Cooks  |
| F125 | Audio and Video Recording Technicians  | G511 | Maîtres d'hôtel and Hosts/Hostesses  |
| F126 | Other Technical Occupations in Motion Pictures, Broadcasting and the Performing Arts | G512 | Bartenders   |
|      |  | G513 | Food and Beverage Servers  |
|      |  | G611 | Police Officers (except Commissioned)  |

---

---

|      |  |      |   |
|------|--|------|---|
| G612 | Fire-fighters  | G941 | Butchers and Meat Cutters, Retail and Wholesale   |
| G621 | Sheriffs and Bailiffs  | G942 | Bakers  |
| G622 | Correctional Service Officers  | G951 | Elemental Medical and Hospital Assistants   |
| G623 | By-law Enforcement and Other Regulatory Officers, n.e.c.             | G961 | Food Service Counter Attendants and Food Preparers  |
| G624 | Other Ranks, Armed Forces  | G962 | Kitchen and Food Service Helpers  |
| G625 | Other Protective Service Occupations                                 | G971 | Service Station Attendants  |
| G631 | Security Guards and Related Occupations                              | G972 | Grocery Clerks and Shelf Stockers   |
| G711 | Travel Counsellors   | G973 | Other Elemental Sales Occupations   |
| G712 | Pursers and Flight Attendants  | G981 | Dry Cleaning and Laundry Occupations  |
| G713 | Airline Sales and Service Agents                                     | G982 | Ironing, Pressing and Finishing Occupations   |
| G714 | Ticket and Cargo Agents and Related Clerks (except Airline)          | G983 | Other Elemental Service Occupations   |
| G715 | Hotel Front Desk Clerks  | H011 | Supervisors, Machinists and Related Occupations   |
| G721 | Tour and Travel Guides   | H012 | Contractors and Supervisors, Electrical Trades and Telecommunications Occupations           |
| G722 | Outdoor Sport and Recreational Guides                                | H013 | Contractors and Supervisors, Pipefitting Trades   |
| G731 | Attendants in Amusement, Recreation and Sport                        | H014 | Contractors and Supervisors, Metal Forming, Shaping and Erecting Trades                     |
| G732 | Other Attendants in Accommodation and Travel (except Airline Travel) | H015 | Contractors and Supervisors, Carpentry Trades   |
| G811 | Visiting Homemakers, Housekeepers and Related Occupations            | H016 | Contractors and Supervisors, Mechanic Trades  |
| G812 | Elementary and Secondary School Teacher Assistants                   | H017 | Contractors and Supervisors, Heavy Construction Equipment Crews                             |
| G813 | Early Childhood Educators and Assistants                             | H018 | Supervisors, Printing and Related Occupations   |
| G814 | Babysitters, Nannies and Parents' Helpers                            | H019 | Contractors and Supervisors, Other Construction Trades, Installers, Repairers and Servicers |
| G911 | Hairstylists and Barbers   | H021 | Supervisors, Railway Transport Operations   |
| G912 | Funeral Directors and Embalmers                                      | H022 | Supervisors, Motor Transport and Other Ground Transit Operators                             |
| G921 | Image, Social and Other Personal Consultants                         | H111 | Plumbers  |
| G922 | Estheticians, Electrologists and Related Occupations                 | H112 | Steamfitters, Pipefitters and Sprinkler System Installers                                   |
| G923 | Pet Groomers and Animal Care Workers                                 | H113 | Gas Fitters   |
| G924 | Other Personal Service Occupations                                   |      |   |
| G931 | Light Duty Cleaners  |      |   |
| G932 | Specialized Cleaners   |      |   |
| G933 | Janitors, Caretakers and Building Superintendents                    |      |   |

---

---

|      |  |      |   |
|------|--|------|---|
| H121 | Carpenters   | H413 | Refrigeration and Air Conditioning Mechanics                  |
| H122 | Cabinetmakers  | H414 | Railway Carmen/women  |
| H131 | Bricklayers  | H415 | Aircraft Mechanics and Aircraft Inspectors                    |
| H132 | Cement Finishers   | H416 | Machine Fitters   |
| H133 | Tilesetters  | H417 | Textile Machinery Mechanics and Repairers                     |
| H134 | Plasterers, Drywall Installers and Finishers, and Lathers          | H418 | Elevator Constructors and Mechanics                           |
| H141 | Roofers and Shinglers  | H421 | Motor Vehicle Mechanics, Technicians and Mechanical Repairers |
| H142 | Glaziers   | H422 | Motor Vehicle Body Repairers                                  |
| H143 | Insulators   | H431 | Oil and Solid Fuel Heating Mechanics                          |
| H144 | Painters and Decorators  | H432 | Electric Appliance Servicers and Repairers                    |
| H145 | Floor Covering Installers  | H433 | Electrical Mechanics  |
| H211 | Electricians (except Industrial and Power System)                  | H434 | Motorcycle and Other Related Mechanics                        |
| H212 | Industrial Electricians  | H435 | Other Small Engine and Equipment Mechanics                    |
| H213 | Power System Electricians  | H511 | Upholsterers  |
| H214 | Electrical Power Line and Cable Workers                            | H512 | Tailors, Dressmakers, Furriers and Milliners                  |
| H215 | Telecommunications Line and Cable Workers                          | H513 | Shoe Repairers and Shoemakers                                 |
| H216 | Telecommunications Installation and Repair Workers                 | H514 | Jewellers, Watch Repairers and Related Occupations            |
| H217 | Cable Television Service and Maintenance Technicians               | H521 | Printing Press Operators                                      |
| H221 | Stationary Engineers and Auxiliary Equipment Operators             | H522 | Commercial Divers   |
| H222 | Power Systems and Power Station Operators                          | H523 | Other Trades and Related Occupations                          |
| H311 | Machinists and Machining and Tooling Inspectors                    | H531 | Residential and Commercial Installers and Servicers           |
| H312 | Tool and Die Makers  | H532 | Waterworks and Gas Maintenance Workers                        |
| H321 | Sheet Metal Workers  | H533 | Automotive Mechanical Installers and Servicers                |
| H322 | Boilermakers   | H534 | Pest Controllers and Fumigators                               |
| H323 | Structural Metal and Platework Fabricators and Fitters             | H535 | Other Repairers and Servicers                                 |
| H324 | Ironworkers  | H611 | Heavy Equipment Operators (except Crane)                      |
| H325 | Blacksmiths and Die Setters  | H612 | Public Works Maintenance Equipment Operators                  |
| H411 | Construction Millwrights and Industrial Mechanics (except Textile) | H621 | Crane Operators   |
| H412 | Heavy-Duty Equipment Mechanics                                     |      |   |

---

---

|      |  |      |   |
|------|--|------|---|
| H622 | Drillers and Blasters - Surface Mining, Quarrying and Construction | I022 | Nursery and Greenhouse Workers                                    |
| H623 | Water Well Drillers  | I111 | Supervisors, Logging and Forestry                                 |
| H711 | Truck Drivers  | I121 | Supervisors, Mining and Quarrying                                 |
| H712 | Bus Drivers and Subway and Other Transit Operators                 | I122 | Supervisors, Oil and Gas Drilling and Service                     |
| H713 | Taxi and Limousine Drivers and Chauffeurs                          | I131 | Underground Production and Development Miners                     |
| H714 | Delivery Drivers   | I132 | Oil and Gas Well Drillers, Servicers, Testers and Related Workers |
| H721 | Railway and Yard Locomotive Engineers                              | I141 | Underground Mine Service and Support Workers                      |
| H722 | Railway Conductors and Brakemen/women                              | I142 | Oil and Gas Well Drilling Workers and Services Operators          |
| H731 | Railway Yard Workers   | I151 | Logging Machinery Operators                                       |
| H732 | Railway Track Maintenance Workers                                  | I161 | Chain-saw and Skidder Operators                                   |
| H733 | Deck Crew, Water Transport   | I162 | Silviculture and Forestry Workers                                 |
| H734 | Engine Room Crew, Water Transport                                  | I171 | Fishing Masters and Officers                                      |
| H735 | Lock and Cable Ferry Operators and Related Occupations             | I172 | Fishing Vessel Skippers and Fishermen/women                       |
| H736 | Boat Operators   | I181 | Fishing Vessel Deckhands  |
| H737 | Air Transport Ramp Attendants                                      | I182 | Trappers and Hunters  |
| H811 | Longshore Workers  | I211 | Harvesting Labourers  |
| H812 | Material Handlers  | I212 | Landscaping and Grounds Maintenance Labourers                     |
| H821 | Construction Trades Helpers and Labourers                          | I213 | Aquaculture and Marine Harvest Labourers                          |
| H822 | Other Trades Helpers and Labourers                                 | I214 | Mine Labourers  |
| H831 | Public Works and Maintenance Labourers                             | I215 | Oil and Gas Drilling, Servicing and Related Labourers             |
| H832 | Railway and Motor Transport Labourers                              | I216 | Logging and Forestry Labourers                                    |
| I011 | Farmers and Farm Managers  | J011 | Supervisors, Mineral and Metal Processing                         |
| I012 | Agricultural and Related Service Contractors and Managers          | J012 | Supervisors, Petroleum, Gas and Chemical Processing and Utilities |
| I013 | Farm Supervisors and Specialized Livestock Workers                 | J013 | Supervisors, Food, Beverage and Tobacco Processing                |
| I014 | Nursery and Greenhouse Operators and Managers                      | J014 | Supervisors, Plastic and Rubber Products Manufacturing            |
| I015 | Landscaping and Grounds Maintenance Contractors and Managers       | J015 | Supervisors, Forest Products Processing                           |
| I016 | Supervisors, Landscape and Horticulture                            | J016 | Supervisors, Textile Processing                                   |
| I017 | Aquaculture Operators and Managers                                 | J021 | Supervisors, Motor Vehicle Assembling                             |
| I021 | General Farm Workers   | J022 | Supervisors, Electronics Manufacturing                            |

---

---

|      |   |      |  |
|------|---|------|--|
| J023 | Supervisors, Electrical Products Manufacturing                      | J152 | Weavers, Knitters and Other Fabric-Making Occupations                                  |
| J024 | Supervisors, Furniture and Fixtures Manufacturing                   | J153 | Textile Dyeing and Finishing Machine Operators   |
| J025 | Supervisors, Fabric, Fur and Leather Products Manufacturing         | J154 | Textile Inspectors, Graders and Samplers   |
| J026 | Supervisors, Other Mechanical and Metal Products Manufacturing      | J161 | Sewing Machine Operators   |
| J027 | Supervisors, Other Products Manufacturing and Assembly              | J162 | Fabric, Fur and Leather Cutters  |
| J111 | Central Control and Process Operators, Mineral and Metal Processing | J163 | Hide and Pelt Processing Workers   |
| J112 | Petroleum, Gas and Chemical Process Operators                       | J164 | Inspectors and Testers, Fabric, Fur and Leather Products Manufacturing                 |
| J113 | Pulping Control Operators   | J171 | Process Control and Machine Operators, Food and Beverage Processing                    |
| J114 | Papermaking and Coating Control Operators                           | J172 | Industrial Butchers and Meat Cutters, Poultry Preparers and Related Workers            |
| J121 | Machine Operators, Mineral and Metal Processing                     | J173 | Fish Plant Workers   |
| J122 | Foundry Workers   | J174 | Tobacco Processing Machine Operators   |
| J123 | Glass Forming and Finishing Machine Operators and Glass Cutters     | J175 | Testers and Graders, Food and Beverage Processing                                      |
| J124 | Concrete, Clay and Stone Forming Operators                          | J181 | Printing Machine Operators   |
| J125 | Inspectors and Testers, Mineral and Metal Processing                | J182 | Camera, Platemaking and Other Pre-Press Occupations                                    |
| J131 | Chemical Plant Machine Operators                                    | J183 | Binding and Finishing Machine Operators  |
| J132 | Plastics Processing Machine Operators                               | J184 | Photographic and Film Processors   |
| J133 | Rubber Processing Machine Operators and Related Workers             | J191 | Machining Tool Operators   |
| J134 | Water and Waste Plant Operators                                     | J192 | Forging Machine Operators  |
| J141 | Sawmill Machine Operators   | J193 | Woodworking Machine Operators  |
| J142 | Pulp Mill Machine Operators   | J194 | Metalworking Machine Operators   |
| J143 | Papermaking and Finishing Machine Operators                         | J195 | Welders and Soldering Machine Operators  |
| J144 | Other Wood Processing Machine Operators                             | J196 | Other Metal Products Machine Operators   |
| J145 | Paper Converting Machine Operators                                  | J197 | Other Products Machine Operators   |
| J146 | Lumber Graders and Other Wood Processing Inspectors and Graders     | J211 | Aircraft Assemblers and Aircraft Assembly Inspectors                                   |
| J151 | Textile Fibre and Yarn Preparation Machine Operators                | J212 | Motor Vehicle Assemblers, Inspectors and Testers                                       |
|      |   | J213 | Electronics Assemblers, Fabricators, Inspectors and Testers                            |
|      |   | J214 | Assemblers and Inspectors, Electrical Appliance, Apparatus and Equipment Manufacturing |

---

|      |   |
|------|---|
| J215 | Assemblers, Fabricators and Inspectors,<br>Industrial Electrical Motors and<br>Transformers |
| J216 | Mechanical Assemblers and Inspectors  |
| J217 | Machine Operators and Inspectors,<br>Electrical Apparatus Manufacturing                     |
| J221 | Boat Assemblers and Inspectors  |
| J222 | Furniture and Fixture Assemblers and<br>Inspectors  |
| J223 | Other Wood Products Assemblers and<br>Inspectors  |
| J224 | Furniture Finishers and Refinishers   |
| J225 | Plastic Products Assemblers, Finishers and<br>Inspectors                                    |
| J226 | Painters and Coaters, Manufacturing   |
| J227 | Plating, Metal Spraying and Related<br>Operators  |
| J228 | Other Assemblers and Inspectors   |
| J311 | Labourers in Mineral and Metal Processing   |
| J312 | Labourers in Metal Fabrication  |
| J313 | Labourers in Chemical Products<br>Processing and Utilities                                  |
| J314 | Labourers in Wood, Pulp and Paper<br>Processing   |
| J315 | Labourers in Rubber and Plastic Products<br>Manufacturing                                   |
| J316 | Labourers in Textile Processing   |
| J317 | Labourers in Food, Beverage and Tobacco<br>Processing                                       |
| J318 | Labourers in Fish Processing  |
| J319 | Other Labourers in Processing,<br>Manufacturing and Utilities                               |

### Appendix III Country Codes – 2001 Census

(Minor differences exist between this list and the one used by the 2001 Census.)

|     |                           |     |                                  |
|-----|---------------------------|-----|----------------------------------|
| 014 | Canada                    | 319 | Saint Kitts and Nevis            |
| 101 | Greenland                 | 320 | Saint Lucia                      |
| 102 | Saint-Pierre and Miquelon | 321 | Saint Vincent and the Grenadines |
| 103 | United States             | 322 | Trinidad and Tobago              |
| 105 | North America, N.O.S      | 323 | Turks and Caicos Islands         |
| 201 | Belize                    | 324 | British Virgin Islands           |
| 202 | Costa Rica                | 325 | U.S. Virgin Islands              |
| 203 | El Salvador               | 326 | West Indies, N.O.S.              |
| 204 | Guatemala                 | 329 | Caribbean and Bermuda, N.O.S.    |
| 205 | Honduras                  | 401 | Argentina                        |
| 206 | Mexico                    | 402 | Bolivia                          |
| 207 | Nicaragua                 | 403 | Brazil                           |
| 208 | Panama                    | 404 | Chile                            |
| 211 | Central America, N.O.S.   | 405 | Colombia                         |
| 301 | Anguilla                  | 406 | Ecuador                          |
| 302 | Antigua and Barbuda       | 407 | Falkland Islands (Malvinas)      |
| 303 | Aruba                     | 408 | French Guiana                    |
| 304 | Bahamas                   | 409 | Guyana                           |
| 305 | Barbados                  | 410 | Paraguay                         |
| 306 | Bermuda                   | 411 | Peru                             |
| 307 | Cayman Islands            | 412 | Suriname                         |
| 308 | Cuba                      | 413 | Uruguay                          |
| 309 | Dominica                  | 414 | Venezuela                        |
| 310 | Dominican Republic        | 420 | South America, N.O.S.            |
| 311 | Grenada                   | 501 | Austria                          |
| 312 | Guadeloupe                | 502 | Belgium                          |
| 313 | Haiti                     | 503 | France                           |
| 314 | Jamaica                   | 505 | Germany                          |
| 315 | Martinique                | 506 | Liechtenstein                    |
| 316 | Montserrat                | 507 | Luxembourg                       |
| 317 | Netherlands Antilles      | 508 | Monaco                           |
| 318 | Puerto Rico               | 509 | Netherlands                      |

---

|     |                                    |     |  |
|-----|------------------------------------|-----|--|
| 511 | Switzerland                        | 568 | Macedonia, Former Yugoslav Republic of Macedonia |
| 512 | Western Europe, N.O.S.             | 569 | Malta  |
| 521 | Bulgaria                           | 570 | Montenegro                                       |
| 522 | Czechoslovakia, N.I.E.             | 571 | Portugal   |
| 523 | Czech Republic                     | 572 | San Marino                                       |
| 524 | Estonia                            | 573 | Serbia   |
| 525 | Hungary                            | 574 | Slovenia   |
| 526 | Latvia                             | 575 | Spain  |
| 527 | Lithuania                          | 576 | Vatican City State                               |
| 528 | Poland                             | 577 | Yugoslavia, N.O.S.                               |
| 529 | Romania                            | 578 | Yugoslavia                                       |
| 530 | Slovakia                           | 579 | Macedonia (Region)                               |
| 531 | U.S.S.R., N.I.E.                   | 580 | Southern Europe, N.O.S.                          |
| 532 | Commonwealth of Independent States | 582 | Kosovo   |
| 533 | Belarus                            | 583 | Vojvodina  |
| 534 | Republic of Moldova                | 584 | Austria-Hungary                                  |
| 535 | Russia Federation                  | 587 | Europe, N.O.S.                                   |
| 536 | Ukraine                            | 601 | Benin  |
| 539 | Eastern Europe, N.O.S.             | 602 | Burkina Faso                                     |
| 541 | Republic of Ireland (EIRE)         | 603 | Cape Verde                                       |
| 542 | Ireland, N.I.E                     | 604 | Côte d'Ivoire                                    |
| 543 | United Kingdom                     | 605 | Gambia   |
| 546 | Denmark                            | 606 | Ghana  |
| 547 | Finland                            | 607 | Guinea   |
| 548 | Iceland                            | 608 | Guinea-Bissau                                    |
| 549 | Norway                             | 609 | Liberia  |
| 550 | Sweden                             | 610 | Mali   |
| 551 | Scandinavia, N.O.S                 | 611 | Mauritania                                       |
| 556 | Northern Europe, N.O.S.            | 612 | Niger  |
| 561 | Albania                            | 613 | Nigeria  |
| 562 | Andorra                            | 614 | St Helena and Ascension                          |
| 563 | Bosnia and Herzegovina             | 615 | Senegal  |
| 564 | Croatia                            | 616 | Sierra Leone                                     |
| 565 | Gibraltar                          | 617 | Togo   |
| 566 | Greece                             | 618 | Western Africa, N.O.S.                           |
| 567 | Italy                              |     |  |

---



---

|     |                             |     |  |
|-----|-----------------------------|-----|--|
| 621 | Burundi                     | 669 | Democratic Republic of the Congo (Zaire) |
| 622 | Comoros                     | 670 | Central Africa                           |
| 623 | Djibouti                    | 681 | Botswana                                 |
| 624 | Eritrea                     | 682 | Lesotho                                  |
| 625 | Ethiopia                    | 683 | Namibia                                  |
| 626 | Kenya                       | 684 | Republic of South Africa                 |
| 627 | Madagascar                  | 685 | Swaziland                                |
| 628 | Malawi                      | 686 | Southern Africa                          |
| 629 | Mauritius                   | 696 | Africa, N.O.S.                           |
| 630 | Mayotte                     | 701 | Afghanistan                              |
| 631 | Mozambique                  | 702 | Cyprus                                   |
| 632 | Reunion                     | 703 | Iran                                     |
| 633 | Rwanda                      | 704 | Turkey                                   |
| 634 | Seychelles                  | 705 | Bahrain                                  |
| 635 | Somalia                     | 706 | Iraq                                     |
| 636 | United Republic of Tanzania | 707 | Israel                                   |
| 637 | Uganda                      | 708 | Jordan                                   |
| 638 | Zambia                      | 709 | Kuwait                                   |
| 639 | Zimbabwe                    | 710 | Lebanon                                  |
| 640 | Eastern Africa, N.O.S.      | 711 | Oman                                     |
| 651 | Algeria                     | 712 | Palestine                                |
| 652 | Egypt                       | 713 | Qatar                                    |
| 653 | Libya                       | 714 | Saudi Arabia                             |
| 654 | Morocco                     | 715 | Syria                                    |
| 655 | Sudan                       | 716 | United Arab Emirates                     |
| 656 | Tunisia                     | 717 | West Bank/Gaza Strip                     |
| 657 | Western Sahara              | 718 | Yemen                                    |
| 658 | Northern Africa, N.O.S.     | 719 | Middle East, N.O.S.                      |
| 661 | Angola                      | 720 | Armenia                                  |
| 662 | Cameroon                    | 721 | Azerbaijan                               |
| 663 | Central African Republic    | 722 | Georgia                                  |
| 664 | Chad                        | 723 | Kazakhstan                               |
| 665 | Republic of the Congo       | 724 | Kyrgyzstan                               |
| 666 | Equatorial Guinea           | 725 | Tajikistan                               |
| 667 | Gabon                       | 726 | Turkmenistan                             |
| 668 | Sao Tome and Principe       | 727 | Uzbekistan                               |

---

|     |                            |     |                                |
|-----|----------------------------|-----|--------------------------------|
| 728 | West Central Asia, N.O.S.  | 803 | Cook Islands                   |
| 731 | China                      | 804 | Fiji                           |
| 732 | People's Republic of China | 805 | French Polynesia               |
| 733 | Hong Kong                  | 806 | Guam                           |
| 734 | Japan                      | 807 | Kiribati                       |
| 735 | North Korea                | 808 | Marshall Islands               |
| 736 | South Korea                | 809 | Federated States of Micronesia |
| 737 | Korea, N.O.S.              | 810 | Nauru                          |
| 738 | Macau                      | 811 | New Caledonia                  |
| 739 | Mongolia                   | 812 | New Zealand                    |
| 740 | Taiwan                     | 813 | Palau                          |
| 741 | Eastern Asia, N.O.S.       | 814 | Papua New Guinea               |
| 751 | Brunei Darussalam          | 815 | Pitcairn Island                |
| 752 | Cambodia                   | 816 | Solomon Islands                |
| 753 | Indonesia                  | 817 | Tonga                          |
| 754 | Laos                       | 818 | Tuvalu                         |
| 755 | Malaysia                   | 819 | U.S. Pacific Trust Territories |
| 756 | Myanmar                    | 820 | Vanuatu                        |
| 757 | Philippines                | 821 | Wallis and Futuna              |
| 758 | Singapore                  | 822 | Samoa                          |
| 759 | Thailand                   | 829 | Oceania, N.O.S.                |
| 760 | Vietnam                    | 994 | Other                          |
| 761 | South East Asia, N.O.S.    |     |                                |
| 771 | Bangladesh                 |     |                                |
| 772 | Bhutan                     |     |                                |
| 773 | India                      |     |                                |
| 774 | Maldives                   |     |                                |
| 775 | Nepal                      |     |                                |
| 776 | Pakistan                   |     |                                |
| 777 | Sri Lanka                  |     |                                |
| 778 | Southern Asia, N.O.S.      |     |                                |
| 781 | East Timor                 |     |                                |
| 782 | Kurdistan                  |     |                                |
| 786 | Asia, N.O.S.               |     |                                |
| 801 | American Samoa             |     |                                |
| 802 | Australia                  |     |                                |

**Appendix IV**  
**Country of Citizenship Codes - 2001 Census**  
(Minor differences exist between this list and the one used by the 2001 Census.)

|    |                                  |    |                                   |
|----|----------------------------------|----|-----------------------------------|
| 3  | United States                    | 35 | Uruguay                           |
| 4  | Belize                           | 36 | Venezuela                         |
| 5  | Costa Rica                       | 37 | Austria                           |
| 6  | El Salvador                      | 38 | Belgium                           |
| 7  | Guatemala                        | 39 | France                            |
| 8  | Honduras                         | 40 | Germany                           |
| 9  | Mexico                           | 41 | Liechtenstein                     |
| 10 | Nicaragua                        | 42 | Luxembourg                        |
| 11 | Panama                           | 43 | Monaco                            |
| 12 | Antigua and Barbuda              | 44 | Netherlands                       |
| 13 | Bahamas                          | 45 | Switzerland                       |
| 14 | Barbados                         | 46 | Bulgaria                          |
| 15 | Cuba                             | 47 | Czechoslovakia, N.I.E.            |
| 16 | Dominica                         | 48 | Czech Republic                    |
| 17 | Dominican Republic               | 49 | Estonia                           |
| 18 | Grenada                          | 50 | Hungary                           |
| 19 | Haiti                            | 51 | Latvia                            |
| 20 | Jamaica                          | 52 | Lithuania                         |
| 21 | Saint Kitts and Nevis            | 53 | Poland                            |
| 22 | Saint Lucia                      | 54 | Romania                           |
| 23 | Saint Vincent and the Grenadines | 55 | Slovakia                          |
| 24 | Trinidad and Tobago              | 56 | U.S.S.R., N.I.E.                  |
| 25 | Argentina                        | 57 | Belarus                           |
| 26 | Bolivia                          | 58 | Republic of Moldova               |
| 27 | Brazil                           | 59 | Russian Federation                |
| 28 | Chile                            | 60 | Ukraine                           |
| 29 | Colombia                         | 61 | Republic of Ireland (EIRE)        |
| 30 | Ecuador                          | 62 | United Kingdom                    |
| 31 | Guyana                           | 63 | British Dependent Territories     |
| 32 | Paraguay                         | 64 | United Kingdom - British Citizens |
| 33 | Peru                             | 65 | Denmark                           |
| 34 | Suriname                         | 66 | Finland                           |

---

|     |   |     |   |
|-----|---|-----|---|
| 67  | Iceland   | 103 | Eritrea                                     |
| 68  | Norway  | 104 | Ethiopia                                    |
| 69  | Sweden  | 105 | Kenya                                       |
| 70  | Albania   | 106 | Madagascar                                  |
| 71  | Andorra   | 107 | Malawi                                      |
| 72  | Bosnia and Herzegovina                              | 108 | Mauritius                                   |
| 73  | Croatia   | 109 | Mozambique                                  |
| 74  | Greece  | 110 | Rwanda                                      |
| 75  | Italy   | 111 | Seychelles                                  |
| 76  | Macedonia, Former Yugoslav<br>Republic of Macedonia | 112 | Somalia                                     |
| 77  | Malta   | 113 | United Republic of Tanzania                 |
| 78  | Portugal  | 114 | Uganda                                      |
| 79  | San Marino  | 115 | Zambia                                      |
| 80  | Slovenia  | 116 | Zimbabwe                                    |
| 81  | Spain   | 117 | Algeria                                     |
| 82  | Vatican City State                                  | 118 | Egypt                                       |
| 83  | Yugoslavia  | 119 | Libya                                       |
| 84  | Benin   | 120 | Morocco                                     |
| 85  | Burkina Faso  | 121 | Sudan                                       |
| 86  | Cape Verde  | 122 | Tunisia                                     |
| 87  | Côte d'Ivoire                                       | 123 | Western Sahara                              |
| 88  | Gambia  | 124 | Angola                                      |
| 89  | Ghana   | 125 | Cameroon                                    |
| 90  | Guinea  | 126 | Central African Republic                    |
| 91  | Guinea-Bissau                                       | 127 | Chad  |
| 92  | Liberia   | 128 | Republic of the Congo                       |
| 93  | Mali  | 129 | Equatorial Guinea                           |
| 94  | Mauritania  | 130 | Gabon                                       |
| 95  | Niger   | 131 | Sao Tome and Principe                       |
| 96  | Nigeria   | 132 | Democratic Republic of the Congo<br>(Zaire) |
| 97  | Senegal   | 133 | Botswana                                    |
| 98  | Sierra Leone  | 134 | Lesotho                                     |
| 99  | Togo  | 135 | Namibia                                     |
| 100 | Burundi   | 136 | Republic of South Africa                    |
| 101 | Comoros   | 137 | Swaziland                                   |
| 102 | Djibouti  | 138 | Afghanistan                                 |

---

---

|     |                                |     |                                |
|-----|--------------------------------|-----|--------------------------------|
| 139 | Cyprus                         | 176 | Singapore                      |
| 140 | Iran                           | 177 | Thailand                       |
| 141 | Turkey                         | 178 | Vietnam                        |
| 142 | Bahrain                        | 179 | Bangladesh                     |
| 143 | Iraq                           | 180 | Bhutan                         |
| 144 | Israel                         | 181 | India                          |
| 145 | Jordan                         | 182 | Maldives                       |
| 146 | Kuwait                         | 183 | Nepal                          |
| 147 | Lebanon                        | 184 | Pakistan                       |
| 148 | Oman                           | 185 | Sri Lanka                      |
| 149 | Palestine/West Bank/Gaza Strip | 186 | East Timor                     |
| 150 | Qatar                          | 187 | Australia                      |
| 151 | Saudi Arabia                   | 188 | Fiji                           |
| 152 | Syria                          | 189 | Kiribati                       |
| 153 | United Arab Emirates           | 190 | Marshall Islands               |
| 154 | Yemen                          | 191 | Federated States of Micronesia |
| 155 | Armenia                        | 192 | Nauru                          |
| 156 | Azerbaijan                     | 193 | New Zealand                    |
| 157 | Georgia                        | 194 | Palau                          |
| 158 | Kazakhstan                     | 195 | Papua New Guinea               |
| 159 | Kyrgyzstan                     | 196 | Solomon Islands                |
| 160 | Tajikistan                     | 197 | Tonga                          |
| 161 | Turkmenistan                   | 198 | Tuvalu                         |
| 162 | Uzbekistan                     | 199 | Vanuatu                        |
| 163 | China                          | 200 | Samoa                          |
| 164 | Japan                          | 201 | French responses               |
| 165 | North Korea                    | 202 | Stateless                      |
| 166 | South Korea                    | 300 | Canada                         |
| 167 | Mongolia                       | 996 | Valid skip                     |
| 168 | Taiwan                         | 997 | Don't know                     |
| 169 | Brunei Darussalam              | 998 | Refused                        |
| 170 | Cambodia                       | 999 | Not stated                     |
| 171 | Indonesia                      |     |                                |
| 172 | Laos                           |     |                                |
| 173 | Malaysia                       |     |                                |
| 174 | Myanmar                        |     |                                |
| 175 | Philippines                    |     |                                |

---



**Appendix V**  
**Intended occupation codes based on the**  
**National Occupational Classification (NOC) 2001**  
**(Only codes present in Wave 1 of the Longitudinal Survey of Immigrants to Canada are listed.)**

|      |   |      |   |
|------|---|------|---|
| 0013 | Senior Managers - Financial, Communications and Other Business Services                         | 0621 | Retail Trade Managers   |
|      |   | 0631 | Restaurant and Food Service Managers                          |
| 0014 | Senior Managers - Health, Education, Social and Community Services and Membership Organizations | 0632 | Accommodation Service Managers                                |
|      |   | 0651 | Other Services Managers                                       |
| 0015 | Senior Managers - Trade, Broadcasting and Other Services, n.e.c.                                | 0711 | Construction Managers   |
|      |   | 0713 | Transportation Managers                                       |
| 0016 | Senior Managers - Goods Production, Utilities, Transportation and Construction                  | 0721 | Facility Operation and Maintenance Managers                   |
| 0111 | Financial Managers  |      |   |
| 0112 | Human Resources Managers  | 0811 | Primary Production Managers (Except Agriculture)              |
| 0113 | Purchasing Managers   | 0911 | Manufacturing Managers  |
| 0114 | Other Administrative Services Managers  | 1111 | Financial Auditors and Accountants                            |
| 0121 | Insurance, Real Estate and Financial Brokerage Managers   | 1112 | Financial and Investment Analysts                             |
| 0122 | Banking, Credit and Other Investment Managers   | 1113 | Securities Agents, Investment Dealers and Brokers             |
| 0123 | Other Business Services Managers  | 1114 | Other Financial Officers                                      |
| 0131 | Telecommunication Carriers Managers   | 1121 | Specialists in Human Resources                                |
| 0211 | Engineering Managers  | 1122 | Professional Occupations in Business Services to Management   |
| 0213 | Computer and Information Systems Managers   | 1211 | Supervisors, General Office and Administrative Support Clerks |
| 0311 | Managers in Health Care   | 1221 | Administrative Officers                                       |
| 0611 | Sales, Marketing and Advertising Managers   | 1222 | Executive Assistants  |

---

|      |  |      |   |
|------|--|------|---|
| 1223 | Personnel and Recruitment Officers                     | 1442 | Personnel Clerks  |
| 1225 | Purchasing Agents and Officers                         | 1453 | Customer Service, Information and Related Clerks          |
| 1226 | Conference and Event Planners                          | 1461 | Mail, Postal and Related Clerks                           |
| 1228 | Immigration, Employment Insurance and Revenue Officers | 1471 | Shippers and Receivers                                    |
| 1231 | Bookkeepers  | 1472 | Storekeepers and Parts Clerks                             |
| 1232 | Loan Officers  | 1473 | Production Clerks   |
| 1233 | Insurance Adjusters and Claims Examiners               | 1474 | Purchasing and Inventory Clerks                           |
| 1234 | Insurance Underwriters                                 | 1475 | Dispatchers and Radio Operators                           |
| 1235 | Assessors, Valuators and Appraisers                    | 2111 | Physicists and Astronomers                                |
| 1241 | Secretaries (Except Legal and Medical)                 | 2112 | Chemists  |
| 1242 | Legal Secretaries                                      | 2113 | Geologists, Geochemists and Geophysicists                 |
| 1243 | Medical Secretaries                                    | 2114 | Meteorologists  |
| 1411 | General Office Clerks                                  | 2115 | Other Professional Occupations in Physical Sciences       |
| 1414 | Receptionists and Switchboard Operators                | 2121 | Biologists and Related Scientists                         |
| 1421 | Computer Operators                                     | 2122 | Forestry professionals                                    |
| 1422 | Data Entry Clerks                                      | 2123 | Agricultural Representatives, Consultants and Specialists |
| 1423 | Desktop Publishing Operators and Related Occupations   | 2131 | Civil Engineers   |
| 1431 | Accounting and Related Clerks                          | 2132 | Mechanical Engineers                                      |
| 1432 | Payroll Clerks   | 2133 | Electrical and Electronics Engineers                      |
| 1433 | Customer Service Representatives - Financial Services  | 2134 | Chemical Engineers  |
| 1434 | Banking, Insurance and Other Financial Clerks          | 2141 | Industrial and Manufacturing Engineers                    |
| 1435 | Collectors   | 2142 | Metallurgical and Materials Engineers                     |
| 1441 | Administrative Clerks                                  | 2143 | Mining Engineers  |

---



---

|      |  |      |  |
|------|--|------|--|
| 2144 | Geological Engineers   | 2241 | Electrical and Electronics Engineering Technologists and Technicians               |
| 2145 | Petroleum Engineers  | 2242 | Electronic Service Technicians (Household and Business Equipment)                  |
| 2146 | Aerospace Engineers  | 2243 | Industrial Instrument Technicians and Mechanics                                    |
| 2147 | Computer Engineers (Except Software Engineers)                         | 2244 | Aircraft Instrument, Electrical and Avionics Mechanics, Technicians and Inspectors |
| 2148 | Other Professional Engineers, n.e.c.                                   | 2251 | Architectural Technologists and Technicians  |
| 2151 | Architects   | 2252 | Industrial Designers   |
| 2153 | Urban and Land Use Planners  | 2253 | Drafting Technologists and Technicians   |
| 2154 | Land Surveyors   | 2255 | Mapping and Related Technologists and Technicians                                  |
| 2161 | Mathematicians, Statisticians and Actuaries                            | 2262 | Engineering Inspectors and Regulatory Officers                                     |
| 2162 | Computer Systems Analysts  | 2263 | Inspectors in Public and Environmental Health and Occupational Health and Safety   |
| 2163 | Computer Programmers   | 2264 | Construction Inspectors  |
| 2211 | Chemical Technologists and Technicians                                 | 2271 | Air Pilots, Flight Engineers and Flying Instructors                                |
| 2212 | Geological and Mineral Technologists and Technicians                   | 2272 | Air Traffic Control and Related Occupations  |
| 2213 | Meteorological Technicians   | 3111 | Specialist Physicians  |
| 2221 | Biological Technologists and Technicians                               | 3112 | General Practitioners and Family Physicians  |
| 2222 | Agricultural and Fish Products Inspectors                              | 3113 | Dentists   |
| 2223 | Forestry Technologists and Technicians                                 | 3114 | Veterinarians  |
| 2225 | Landscape and Horticulture Technicians and Specialists                 | 3122 | Chiropractors  |
| 2231 | Civil Engineering Technologists and Technicians                        | 3131 | Pharmacists  |
| 2232 | Mechanical Engineering Technologists and Technicians                   | 3132 | Dietitians and Nutritionists   |
| 2233 | Industrial Engineering and Manufacturing Technologists and Technicians |      |  |
| 2234 | Construction Estimators  |      |  |

---

---

|      |   |      |  |
|------|---|------|--|
| 3141 | Audiologists and Speech-Language Pathologists                                     | 3235 | Other Technical Occupations in Therapy and Assessment                            |
| 3142 | Physiotherapists  | 3411 | Dental Assistants  |
| 3143 | Occupational Therapists   | 3413 | Nurse Aides, Orderlies and Patient Service Associates                            |
| 3144 | Other Professional Occupations in Therapy and Assessment                          | 3414 | Other Assisting Occupations in Support of Health Services                        |
| 3152 | Registered Nurses   | 4112 | Lawyers and Quebec Notaries  |
| 3211 | Medical Laboratory Technologists and Pathologists' Assistants                     | 4121 | University Professors  |
| 3212 | Medical Laboratory Technicians  | 4122 | Post-Secondary Teaching and Research Assistants                                  |
| 3213 | Veterinary and Animal Health Technologists and Technicians                        | 4131 | College and Other Vocational Instructors   |
| 3214 | Respiratory Therapists, Clinical Perfusionists and Cardio-Pulmonary technologists | 4141 | Secondary School Teachers  |
| 3215 | Medical Radiation Technologists   | 4142 | Elementary School and Kindergarten Teachers                                      |
| 3216 | Medical Sonographers  | 4143 | Educational Counsellors  |
| 3218 | Electroencephalographic and Other Diagnostic Technologists, n.e.c.                | 4151 | Psychologists  |
| 3219 | Other Medical Technologists and Technicians (except Dental Health)                | 4152 | Social Workers   |
| 3221 | Denturists  | 4153 | Family, Marriage and Other Related Counsellors                                   |
| 3222 | Dental Hygienists and Dental Therapists   | 4154 | Ministers of Religion  |
| 3223 | Dental Technologists, Technicians and Laboratory Bench Workers                    | 4161 | Natural and Applied Science Policy Researchers, Consultants and Program Officers |
| 3231 | Opticians   | 4162 | Economists and Economic Policy Researchers and Analysts                          |
| 3232 | Midwives and Practitioners of Natural Healing                                     | 4163 | Business Development Officers and Marketing Researchers and Consultants          |
| 3233 | Licensed Practical Nurses   | 4164 | Social Policy Researchers Consultants and Program Officers                       |
| 3234 | Ambulance Attendants and Other Paramedical Occupations                            | 4166 | Education Policy Researchers, Consultants and Program Officers                   |

---

---

|      |   |      |  |
|------|---|------|--|
| 4169 | Other Professional Occupations in Social Science, n.e.c.        | 5226 | Other Technical and Co-ordinating Occupations in Motion Pictures, Broadcasting and the Performing Arts |
| 4211 | Paralegal and Related Occupations                               | 5231 | Announcers and Other Broadcasters  |
| 4212 | Community and Social Service Workers                            | 5232 | Other Performers   |
| 4214 | Early Childhood Educators and Assistants                        | 5241 | Graphic Designers and Illustrators   |
| 4215 | Instructors and Teachers of Persons with Disabilities           | 5242 | Interior Designers   |
| 4217 | Other Religious Occupations                                     | 5243 | Theatre, Fashion, Exhibit and Other Creative Designers   |
| 5111 | Librarians  | 5244 | Artisans and Craftspersons   |
| 5112 | Conservators and Curators                                       | 5245 | Patternmakers, Textile, Leather and Fur Products   |
| 5121 | Authors and Writers   | 5251 | Athletes   |
| 5122 | Editors   | 5252 | Coaches  |
| 5123 | Journalists   | 5254 | Program Leaders and Instructors in Recreation and Sport  |
| 5124 | Professional Occupations in Public Relations and Communications | 6211 | Retail Trade Supervisors   |
| 5125 | Translators, Terminologists and Interpreters                    | 6212 | Food Service Supervisors   |
| 5131 | Producers, Directors, Choreographers and Related Occupations    | 6221 | Technical Sales Specialists, Wholesale Trade   |
| 5133 | Musicians and Singers   | 6231 | Insurance Agents and Brokers   |
| 5134 | Dancers   | 6232 | Real Estate Agents and Salespersons  |
| 5135 | Actors and Comedians  | 6233 | Retail and Wholesale Buyers  |
| 5136 | Painters, Sculptors and Other Visual Artists                    | 6241 | Chefs  |
| 5211 | Library and Archive Technicians and Assistants                  | 6242 | Cooks  |
| 5221 | Photographers   | 6251 | Butchers and Meat Cutters, Retail and Wholesale  |
| 5222 | Film and Video Camera Operators                                 | 6252 | Bakers   |
| 5225 | Audio and Video Recording Technicians                           |      |  |

---

---

|      |   |      |   |
|------|---|------|---|
| 6261 | Police Officers (Except Commissioned)   | 6622 | Grocery Clerks and Store Shelf Stockers   |
| 6262 | Firefighters  | 6623 | Other Elemental Sales Occupations   |
| 6271 | Hairstylists and Barbers  | 6631 | Elemental Medical and Hospital Assistants   |
| 6411 | Sales Representatives, Wholesale Trade (Non-Technical)                              | 6641 | Food Counter Attendants, Kitchen Helpers and Related  |
| 6421 | Retail Salespersons and Sales Clerks  | 6642 | Kitchen and Food Service Helpers  |
| 6431 | Travel Counsellors  | 6651 | Security Guards and Related Occupations   |
| 6432 | Pursers and Flight Attendants   | 6661 | Light Duty Cleaners   |
| 6433 | Airline Sales and Service Agents  | 6663 | Janitors, Caretakers and Building Superintendents   |
| 6434 | Ticket Agents and Cargo Service Representatives and Related Clerks (Except Airline) | 7212 | Contractors and Supervisors, Electrical Trades and Telecommunications Occupations           |
| 6441 | Tour and Travel Guides  | 7216 | Contractors and Supervisors, Mechanic Trades  |
| 6442 | Outdoor Sport and Recreational Guides   | 7217 | Contractors and Supervisors, Heavy Construction Equipment Crews                             |
| 6443 | Casino Occupations  | 7219 | Contractors and Supervisors, Other Construction Trades, Installers, Repairers and Servicers |
| 6451 | Maîtres d'hôtel and Hosts/Hostesses   | 7221 | Supervisors, Railway Transport Operations   |
| 6452 | Bartenders  | 7231 | Machinists and Machining and Tooling Inspectors   |
| 6453 | Food and Beverage Servers   | 7232 | Tool and Die Makers   |
| 6465 | Other Protective Service Occupations  | 7241 | Electricians (Except Industrial and Power System)   |
| 6471 | Visiting Homemakers, Housekeepers and Related Occupations                           | 7242 | Industrial Electricians   |
| 6472 | Elementary and Secondary School Teacher Assistants                                  | 7244 | Electrical Power Line and Cable Workers   |
| 6473 | Early Childhood Educator Assistants   | 7245 | Telecommunications Line and Cable Workers   |
| 6474 | Babysitters, Nannies and Parent's Helpers   | 7246 | Telecommunications Installation and Repair Workers  |
| 6482 | Estheticians, Electrologists and Related Occupations                                |      |   |
| 6621 | Service Station Attendants  |      |   |

---

---

|      |  |      |   |
|------|--|------|---|
| 7251 | Plumbers   | 7344 | Jewellers, Watch Repairers and Related Occupations                |
| 7253 | Gas Fitters  | 7351 | Stationary Engineers and Auxiliary Equipment Operators            |
| 7261 | Sheet Metal Workers  | 7352 | Power Systems and Power Station Operators                         |
| 7265 | Welders and Related Machine Operators                                    | 7371 | Crane Operators   |
| 7271 | Carpenters   | 7381 | Printing Press Operators  |
| 7272 | Cabinetmakers  | 7383 | Other Trades and Related Occupations                              |
| 7281 | Bricklayers  | 7411 | Truck Drivers   |
| 7282 | Concrete Finishers   | 7412 | Bus Drivers, and Subway and Other Transit Operators               |
| 7283 | Tilesetters  | 7413 | Taxi and Limousine Drivers and Chauffeurs                         |
| 7294 | Painters and Decorators  | 7414 | Delivery Drivers and Courier Service Drivers                      |
| 7311 | Construction Millwrights and Industrial Mechanics (Except Textile)       | 7421 | Heavy Equipment Operators (Except Crane)                          |
| 7312 | Heavy-Duty Equipment Mechanics   | 7445 | Other Repairers and Servicers                                     |
| 7313 | Refrigeration and Air Conditioning Mechanics                             | 7452 | Material Handlers   |
| 7315 | Aircraft Mechanics and Aircraft Inspectors                               | 7611 | Construction Trades Helpers and Labourers                         |
| 7316 | Machine Fitters  | 7612 | Other Trades Helpers and Labourers                                |
| 7317 | Textile Machinery Mechanics and Repairers                                | 7622 | Railway and Motor Transport Labourers                             |
| 7321 | Automotive Service Technicians, Truck Mechanics and Mechanical Repairers | 8222 | Supervisors, Oil and Gas Drilling and Service                     |
| 7322 | Motor Vehicle Body Repairers   | 8232 | Oil and Gas Well Drillers, Servicers, Testers and Related Workers |
| 7332 | Electric Appliance Servicers and Repairers                               | 8251 | Farmers and Farm Managers   |
| 7333 | Electrical Mechanics   | 8253 | Farm Supervisors and Specialized Livestock Workers                |
| 7342 | Tailors, Dressmakers, Furriers and Milliners                             |      |   |
| 7343 | Shoe Repairers and Shoemakers  |      |   |

---

---

|      |  |      |                                   |  |
|------|--|------|-----------------------------------|--|
| 8254 | Nursery and Greenhouse Operators and Managers  |      |                                   | <b><u>List of Citizenship and Immigration Canada Synthetic Codes</u></b> |
| 8262 | Fishing Vessel Skippers and Fishermen/women  | 9910 | Investors                         |  |
| 8422 | Silviculture and Forestry Workers  | 9920 | Students                          |  |
| 8431 | General Farm Workers   | 9930 | New workers                       |  |
| 8614 | Mine Labourers   | 9940 | Retirees                          |  |
| 9216 | Supervisors, Textile Processing  | 9950 | Open employment authorization     |  |
| 9414 | Concrete, Clay and Stone Forming Operators   | 9960 | Fiancées                          |  |
| 9415 | Inspectors and Testers, Mineral and Metal Processing                                   | 9970 | Homemakers                        |  |
| 9432 | Pulp Mill Machine Operators  | 9980 | Other non-workers                 |  |
|      |  | 9990 | Software designer - pilot program |  |
| 9441 | Textile Fibre and Yarn Preparation Machine Operators                                   |      |                                   | <b><u>Longitudinal Survey of Immigrants to Canada Reserved Code</u></b>  |
| 9451 | Sewing Machine Operators   | 9999 | Not stated                        |  |
| 9462 | Industrial Butchers and Meat Cutters, Poultry Preparers and Related Workers            |      |                                   |  |
| 9465 | Testers and Graders, Food and Beverage Processing                                      |      |                                   |  |
| 9483 | Electronics Assemblers, Fabricators, Inspectors and Testers                            |      |                                   |  |
| 9484 | Assemblers and Inspectors, Electrical Appliance, Apparatus and Equipment Manufacturing |      |                                   |  |
| 9511 | Machining Tool Operators   |      |                                   |  |
| 9619 | Other Labourers in Processing, Manufacturing and Utilities                             |      |                                   |  |

---