

Microdata User Guide
Survey of Staffing - Candidates
2009 - Cycle 1



Statistics
Canada

Statistique
Canada

Canada

Table of Contents

1.0	Introduction	5
2.0	Background	7
3.0	Objectives	9
4.0	Concepts and Definitions	11
5.0	Survey Methodology	13
5.1	Population Coverage	13
5.2	Participating Departments and Agencies	14
5.3	Sample Design	15
5.4	Sample Size	15
6.0	Data Collection	17
6.1	Questionnaire Design	17
6.2	Data Collection	17
7.0	Data Processing	19
7.1	Data Capture	19
7.2	Verification and editing	19
7.3	Coding of Open-ended Questions	20
7.4	Imputation	20
7.5	Creation of Derived Variables	20
7.6	Weighting	20
7.7	Suppression of Confidential Information	21
8.0	Data Quality	23
8.1	Response Rates – Departments and Agencies	23
8.2	Survey Errors	24
8.2.1	The Frame	25
8.2.2	Data Collection	25
8.2.3	Data Processing	25
8.2.4	Non-response	25
8.2.5	Measurement of Sampling Error	26
9.0	Guidelines for Tabulation, Analysis and Release	27
9.1	Rounding Guidelines	27
9.2	Sample Weighting Guidelines for Tabulation	27
9.3	Definitions of Types of Estimates: Categorical and Quantitative	28
9.3.1	Categorical Estimates	28
9.3.2	Quantitative Estimates	28
9.3.3	Results from Scale-type Questions: Percentage of Favourable Response	29
9.3.4	Tabulation of Categorical Estimates	29
9.3.5	Tabulation of Quantitative Estimates	29
9.3.6	Percentage of Favourable Response: Evaluation Guidelines	30
9.4	Guidelines for Statistical Analysis	31
9.5	Coefficient of Variation Release Guidelines	31
9.6	Release Cut-off's for the Survey of Staffing - Candidates	32
9.6.1	Precision Guidelines	32
9.6.2	Confidentiality Guidelines	34

9.7	Other Types of Analysis	35
10.0	Approximate Sampling Variability Tables	37
10.1	How to Use the Coefficient of Variation Tables for Categorical Estimates.....	38
10.1.1	Examples of Using the Coefficient of Variation Tables for the Categorical Estimates	40
10.2	How to Use the Coefficient of Variation Tables to Obtain Confidence Limits.....	47
10.2.1	Examples of Using the Coefficient of Variation Tables to Obtain Confidence Limits.....	48
10.3	How to Use the Coefficient of Variation Tables to Do a T-test	48
10.3.1	Examples of Using the Coefficient of Variation Tables to Do a T-test.....	49
10.4	Coefficients of Variation for Quantitative Estimates.....	49
10.5	Coefficient of Variation Tables	50
11.0	Weighting	51
11.1	Initial Design Weight	51
11.2	Non-response Adjustment	51
11.3	Adjustment for Non-sharing Units.....	52
11.4	Post-stratification Adjustment and Final Weight	52
12.0	Questionnaire	53
13.0	Record Layout with Univariate Frequencies.....	55

1.0 Introduction

The Survey of Staffing – Candidates (SOS) was conducted by Statistics Canada from January 7th, 2009 to February 20th, 2009 on behalf of the Public Service Commission (PSC). This manual has been produced to facilitate the manipulation of the microdata files of the survey results.

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

Statistics Canada

Client Services
Special Surveys Division
Telephone: 613-951-3321 or call toll-free 1 800 461-9050
Fax: 613-951-4527
E-mail: ssd@statcan.gc.ca

Public Service Commission

Milan Jayasinghe
Manager, Survey Division
Audit, Evaluation & Studies Branch
Public Service Commission of Canada
L'esplanade Laurier, West Tower
Ottawa, Ontario, K1A 0M7
Tel: 613-992-9329
Fax: 613-947-7739
Milan.Jayasinghe@psc-cfp.gc.ca

2.0 Background

The Public Service Commission (PSC) conducted the Survey of Appointments (SOA) from 2000 to 2007 as a means of monitoring staffing in the public service. When the new *Public Service Employment Act (PSEA)* came into force on December 31st, 2005, it became apparent that the PSC needed to expand the scope of its survey activity to non-appointees to meet its obligations under the legislation and to provide a richer set of data for analysis of the staffing system under the new *PSEA*.

As a follow-up to the SOA, the PSC asked Statistics Canada to conduct the first cycle of the Survey of Staffing – Candidates to a sample of public service employees working in federal departments and agencies with at least 350 employees or more, which fall under the *PSEA*. In the spring of 2007, English and French focus groups that included employees from different departments, at various groups and levels were held across the country. In November and December of 2007, a pilot version of this survey was conducted by Statistics Canada in some departments.

The content of the questionnaire and the methods of collection for Cycle 1 of the survey were designed and implemented using the information gathered in the focus groups and the pilot survey,

The survey collects data on the appointment process, staffing strategies, the area of competition and the experiences of public servants who have participated in a staffing activity. Since the Public Service Commission is also mandated to oversee the political impartiality of the public service, the survey collected information on political activities.

3.0 Objectives

The major objective of the survey is to collect data on staffing experiences of federal public service employees who participated in specific staffing activity types during the reference period and, on political activities for all public service employees.

The information gathered by the survey will contribute to a government-wide perspective on the staffing process and will help to identify areas where improvement is needed to the staffing system. It will also be useful in providing information to determine whether any changes to the *Public Service Employment Act (PSEA)* and/or related policies are needed when the legislation is revisited in 2010. The questions on political activities will provide an indication of the participation by public servants in political activities, the information sources that they typically consulted regarding political activities, and their knowledge of their rights and responsibilities.

4.0 Concepts and Definitions

This chapter outlines concepts and definitions of interest to the users.

The population for the survey included all public service employees working in federal departments and agencies with at least 350 employees or more, which fall under the *Public Service Employment Act (PSEA)*.

Staffing activity is any action or activity intended to result in one or more appointments within, or into the federal public service. This could involve a change in group and/or level (e.g. CR-03 to CR-05), and/or a change in work status (e.g. term to indeterminate).

In-scope staffing activity is any staffing activity which **excludes** staffing activities for a deployment, a casual, consulting or acting positions, group and/or level changes resulting from a simultaneous reclassification for a number of employees at once, incumbent-based promotions, promotions arising from the completion of a specific training or development program; and automatic conversions of term positions to indeterminate positions. Staffing activities from which respondents voluntarily withdraw before their conclusion were excluded.

Some definitions were included on the questionnaire to ensure that all respondents had the same understanding of the terms.

These include:

Acting Position

A position created by the temporary assignment of an employee to the duties of a higher level, with an adjustment to pay.

Casual

A person hired into the public service on a short term temporary basis. In most instances, a casual employee has not worked in a single department/agency more than 90 days during a calendar year.

Concluded for you

Means one of several possible outcomes, such as receiving an offer of appointment, being placed in a pool or qualified candidates for future referral or consideration, or being removed from further consideration.

Consulting Position

A position held by a professional who provides advice or services in a particular area of expertise. This person is self-employed or works for a consulting firm.

Deployment

The transfer of an employee within an organization, without promotion, from one position to another in the same occupational group. Where authorized by the Public Service Employment Regulations (PSER), employees may be transferred to another occupational group.

Development Program

An employee development program implemented via a series of rotations or placements. Employees in these programs will usually get a promotion within a set amount of time (e.g. ES-1 to ES-2 after a year).

Hiring Manager

A person who chairs a board that makes a selection among candidates, or a person who provides the written rationale for the case of a non-advertised (non-competitive) appointment.

Incumbent-based promotion

A promotion based on a track record with identifiable accomplishments such as promotion for scientific researcher.

Political activities

A carrying on of any activity in support of, within or in opposition to a political party; carrying on any activity in support of or in opposition to a candidate; or seeking nomination as or being a candidate in an election before or during the election period.

Pool (of assessed or partially assessed individuals)

A group of candidates for future consideration who have been assessed on at least one of the merit criteria common to a number of positions. Individuals in the pool may subsequently undergo additional assessment on the remaining merit criteria as positions become available to be filled.

Reclassification

A change in either the occupational group or level of the position (or both) as a result of a classification decision. This can happen on an individual basis, or for multiple employees belonging to the same group and level.

Specific Training Program

In some instances, employees enter a training program, as part of their employment, with the understanding that they will be promoted to a higher position when the training is successfully completed.

Work Unit

A group of people who have the same objective or who work on the same project and come into regular contact, or meet regularly, with each other.

5.0 Survey Methodology

The Survey of Staffing - Candidates (SOS) was administered from January 7th, 2009 to February 20th, 2009 to a sample of public service employees who worked in federal departments and agencies which fall under the *Public Service Employment Act (PSEA)* and that had at least 350 employees or more on the last day of the reference period (i.e. September 30, 2008). A similar survey was being sent out by the Public Service Commission to a sample of managers involved in staffing processes during the same time period to explore their views and practices with staffing. Each person in the sample was contacted by e-mail and asked to complete an electronic questionnaire available on the Statistics Canada website. People who could not be contacted by e-mail or those who did not have access to the Statistics Canada website (or required an internet browser) were asked to complete a paper questionnaire.

5.1 Population Coverage

The SOS targeted public service employees who worked in federal departments and agencies with at least 350 employees or more which fell under the *Public Service Employment Act* as of September 30th, 2008 with the following exceptions:

- non-civilians;
- governor-in-council appointments; and
- minister's exempt staff.

The questions targeted three distinct populations. The first group was made up of all employees who were asked questions on political activities. The second included those employees who, during the reference period, had been appointed to a new position. The third group were those employees who participated in a staffing activity (candidates), but were not appointed to a position. This group included internal employees only.

The targeted employees for the candidate portion of the survey (i.e. groups 2 and 3 described in the previous paragraph) were defined as all candidates who participated in a staffing process whether it resulted in an appointment or not. The types of appointments that were of interest for the survey were the appointments to the public service, the promotions as well as a fraction of the lateral movements (for example, lateral movements through staffing processes but not in deployments). If the respondent indicated that the type of staffing activity was for

- a casual, consulting or acting position,
- a deployment,
- was part of the simultaneous reclassification of many employees,
- a promotion arising from the completion of a specific training or development program, or
- an incumbent-based promotion,

they were not asked the details about that staffing activity and went directly to the section on political activities.

For all other types of staffing activities the respondent was asked to continue with the questionnaire providing details of that staffing activity and their position at the time of the process.

If the applicant withdrew prior to the in-scope staffing activity's completion, the respondent went directly to the section on political activities.

5.2 Participating Departments and Agencies

Department/Agency Name	In-scope Active Population
National Defence	25,627
Service Canada and Human Resources and Social Development Canada	23,545
Correctional Service Canada	16,277
Canada Border Services Agency	14,172
Public Works and Government Services Canada	12,672
Fisheries and Oceans Canada	11,005
Health Canada	9,759
Agriculture and Agri-Food Canada	7,134
Royal Canadian Mounted Police	6,833
Environment Canada	6,310
Statistics Canada	5,731
Industry Canada	5,339
Transport Canada	5,114
Indian and Northern Affairs Canada and Indian Residential Schools Resolution Canada	4,657
Foreign Affairs and International Trade Canada	4,672
Natural Resources Canada	4,442
Department of Justice Canada	4,459
Veterans Affairs Canada	4,153
Citizenship and Immigration Canada	4,076
Passport Canada	2,756
Canadian Heritage	2,367
Public Health Agency of Canada	2,262
Canadian International Development Agency	1,884
Library and Archives Canada	1,142
Treasury Board of Canada Secretariat	1,155
Department of Finance Canada	1,077
Public Safety Canada	1,048
Public Service Commission of Canada	1,029
Immigration and Refugee Board of Canada	874
Canada School of Public Service	911
Privy Council Office	824
Public Prosecution Service of Canada	796

Department/Agency Name	In-scope Active Population
Atlantic Canada Opportunities Agency	746
Canadian Grain Commission	661
Canadian Space Agency	591
Canada Public Service Agency	598
Courts Administration Service	587
Office of the Superintendent of Financial Institutions Canada	479
Canadian Radio-television and Telecommunications Commission	441
Office of the Chief Electoral Officer	533
Economic Development Agency of Canada for the Regions of Quebec	410
Western Economic Diversification Canada	417
National Parole Board	373
National Energy Board	322
Total In-scope active population	200,260

5.3 Sample Design

The sampling frame was made up of all in-scope employees that were on the Public Works and Government Services Canada's Incumbent file. Since the contact information (e-mail and postal address) was not available on the Incumbent file, it had to be collected by Statistics Canada from the departments through Article 13 of the Statistics Act.

The sampling unit was the employee. In each department, a systematic sample of employees was selected from the sampling frame.

5.4 Sample Size

The required target population depends on the following factors:

- the targeted accuracy for the estimations (targeted coefficient of variation (CV)),
- the response rate,
- the share rate (proportion of respondents who agree to share their data with the Public Service Commission (PSC)),
- the minimum proportion to examine and
- the hit rate (proportion of all in-scope employees who are part of the targeted group – the three targeted groups were defined in Section 5.1).

A conservative approach was used for the calculation of the sample size. The expected hit rate for the smaller group (those who were appointed to a new position during the reference period) was used. An estimate of this hit rate was obtained from the PSC historical data (2005 to 2008).

The following parameters were used to calculate the sample size for the SOS:

- target CV – 16.5%
- combined response rate and share rate – 40.0%
- estimated minimum proportion – 13.5%
- hit rates based on the PSC historical data by department

By using these parameters, the required sample size was 77,214 employees.

6.0 Data Collection

6.1 Questionnaire Design

In the spring of 2007, English and French focus groups that included employees from different departments at various groups and levels were held across the country. In November and December of 2007, a pilot version of this survey was conducted in a few departments. The results of the pilot survey were used to improve numerous aspects of the survey.

The electronic format of the questionnaire was designed to follow standard practices and wording, when applicable, in an Internet-based environment. This included the automatic control of question wording and flows that depended upon answers to earlier questions and the use of on-line edits to check for logical inconsistencies and capture errors, such as out-of-range values. The electronic application for data collection was subjected to extensive testing.

Initially the main topic of the survey was staffing activities, for a targeted group of respondents. Later in the development stage, questions on political activities were added that everyone who received the questionnaire was required to answer.

6.2 Data Collection

Responding to this survey was voluntary. Data were collected directly from survey respondents.

In December 2008, as part of the communication plan, two official letters announcing the initiative were sent by the Public Service Commission to the participating departments; one communiqué to the Deputy Ministers and another communiqué to the Heads of Human Resources.

From January 7th to 9th, 2009, each person in the sample was contacted by e-mail and invited to complete an electronic questionnaire available on the Statistics Canada website. Those who could not be contacted by e-mail or who did not have access to the Statistics Canada website (or required an internet browser) were invited to complete a paper questionnaire.

About 90% of the sample was sent out electronically, 1% as surface mail packages containing the introductory letter, the questionnaire and the postage-paid return envelope and the remaining 9% of the sampled respondents didn't have enough mailing information to be contacted either by e-mail or by surface mail.

For the electronic sample, respondents received an e-mail containing the invitation with a link to the Statistics Canada Electronic Portal. The link had an embedded access code that provided access to the Survey of Staffing – Candidates questionnaire.

Paper questionnaires were sent out by regular mail. Once completed, the questionnaire was returned directly to Statistics Canada in a postage-paid return envelope. Statistics Canada accepted completed questionnaires until February 27th, 2009.

During collection five reminder e-mails were sent to participants in the electronic collection who had not already submitted their electronic questionnaire. Only one reminder was sent to participants in the paper collection. Individuals who refused to participate did not receive the reminder notifications.

Participants of the survey received support during the collection period through the Statistics Canada Help Desk (1-800 and e-mail). The participants of the electronic collection could request a paper questionnaire.

7.0 Data Processing

The main outputs of the Survey of Staffing – Candidates (SOS) are a “clean” Master and Share file. The Master File consists of data processed from the electronic and paper modes of the questionnaire. The Share File contains a subset of the records from the Master File. Respondents who refused to share their information with the sponsor of the survey, the Public Service Commission were removed from the share file. This section presents a brief summary of the processing steps involved in producing these files.

7.1 Data Capture

The data capture of the paper questionnaires received was done between February 4th and February 27th, 2009. Data was captured using imaging and automated data entry technology. A small proportion of questionnaires, those that could not be read by the optical scanners, were captured using heads-down keying by experienced operators. Questionable zones method with standard quality control measures were used to verify the error rate of the capture operations.

For the electronic questionnaire, responses to survey questions were entered directly by the respondents. The electronic questionnaire reduces processing time and costs associated with data entry, transcription errors and data transmission. The responses were secure through industry standard encryption protocols, firewalls and encryption layers.

Some editing was done directly at the time the electronic questionnaire was completed. Where the information was outside the range (too large or small) of expected values, or inconsistent with the previous entries, the respondent was prompted, through message screens, to verify the information. However, the respondents had the option of bypassing the edits, and of skipping questions if they did not know the answer or refused to answer. Therefore, the data were subjected to further edit processes after they were submitted. When the electronic data was received it was converted to readable text files.

7.2 Verification and editing

Electronic text files containing the daily transmissions of submitted cases coming from the Statistics Canada website collection and those coming from the paper data capture process were combined to create the “raw” survey files. Before further processing was done, verification was performed to identify and eliminate potential duplicate records and to identify non-response and out-of-scope records.

To be considered a response record in the SOS, respondents must have completed at least some of the items required to derive the type of staffing activity covered by the survey or answered part of the first two political activity questions. If these response criteria were not met, the record was considered as a non-response.

Editing consisted in modifying the data at the employee variable level. The first step in editing was to determine which items from the survey collection needed to be kept on the survey master file. Subsequently, invalid characters were deleted and the data items were formatted appropriately.

The first type of error treated involved a lack of information in questions that should have been answered. For this type of error, a non-response or “not-stated” code was assigned to the item.

The second type of error treated was errors in questionnaire flow, where questions that did not apply to the respondent (and should therefore not have been answered) were found to contain answers. In this case a computer edit automatically eliminated superfluous data by following the flow of the questionnaire implied by answers to previous, and in some cases, subsequent questions.

As well, data inconsistencies were corrected. Some verification was done to check if the respondent's age group was compatible with the number of years they worked in the Public Service, and by the date they started working in their general work unit before the staffing activity concluded. Occupational levels within specific occupation groups were also verified as to their validity, either by the position the respondent applied for, or by the position the respondent held before the staffing activity concluded.

7.3 Coding of Open-ended Questions

There were no open-ended questions in the SOS.

7.4 Imputation

There was no imputation in the SOS. Item and partial non-response were coded as "Not stated" during the editing process.

7.5 Creation of Derived Variables

A number of variables included on the Master file have been derived by combining variables on the questionnaire in order to facilitate data analysis. The following is a list of the derived variables for the SOS.

MOVEMENT	Was this position that the employee applied for or was appointed to, a lateral movement, a promotion or other?
CHNGREG	Did the employee apply to or get appointed to, a position in the same region where they were already working/living?
EXTNAOS	Was this a NAOS (National Area of Selection) process?
UNDERREP	Was the employee a member of any of three employment equity groups?
FAIRASSM	Taken together, how fairly were all the factors being considered assessed?

7.6 Weighting

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

The weighting phase is a step which calculates, for each record, what this number is. This weight appears on the microdata file, and must be used to derive meaningful estimates from the survey. For example if the number of employees who worked in Quebec and participated in a staffing activity is to be estimated, it is done by selecting the records referring to those employees in the sample with these characteristics and summing the weights entered on those records.

Details of the method used to calculate these weights are presented in Chapter 11.0.

See Chapter 9.0 for the guidelines for tabulation, analysis and release.

7.7 *Suppression of Confidential Information*

It should be noted that the “Share” file differs from the survey “Master” file held by Statistics Canada. The Share File contains a subset of the records from the Master File. Respondents who refused to share their information with the sponsor of the survey, the Public Service Commission were removed from the share file. The overall share rate for Cycle 1 of the SOS was 89.8%.

Users requiring access to information excluded from the “Master” microdata file may purchase custom tabulations. Estimates generated will be released to the user, subject to meeting the guidelines for analysis and release outlined in Chapter 9.0 of this document.

8.0 Data Quality

8.1 Response Rates – Departments and Agencies

The following table summarizes the response rates for the Survey of Staffing – Candidates (SOS) Cycle 1.

Response Rates by Department/Agency – Unweighted

Department/Agency Name	Sample Size	Responding Employees		Response Rate (%)	
		Master	Share	Master	Share
National Defence	2,554	980	878	38.4	34.4
Service Canada and Human Resources and Social Development Canada	3,045	1,524	1,375	50.0	45.2
Correctional Service Canada	3,856	1,426	1,270	37.0	32.9
Canada Border Services Agency	2,389	1,011	886	42.3	37.1
Public Works and Government Services Canada	3,443	1,470	1,325	42.7	38.5
Fisheries and Oceans Canada	3,275	967	880	29.5	26.9
Health Canada	2,415	1,093	1,001	45.3	41.4
Agriculture and Agri-Food Canada	2,188	895	805	40.9	36.8
Royal Canadian Mounted Police	2,405	976	881	40.6	36.6
Environment Canada	3,356	1,532	1,395	45.6	41.6
Statistics Canada	2,212	1,505	1,370	68.0	61.9
Industry Canada	3,033	1,491	1,287	49.2	42.4
Transport Canada	3,627	1,800	1,615	49.6	44.5
Indian and Northern Affairs Canada and Indian Residential Schools Resolution Canada	2,601	1,080	966	41.5	37.1
Foreign Affairs and International Trade Canada	2,512	1,129	1,024	44.9	40.8
Natural Resources Canada	3,315	1,512	1,362	45.6	41.1
Department of Justice Canada	2,752	1,211	1,065	44.0	38.7
Veterans Affairs Canada	2,414	1,324	1,203	54.8	49.8
Citizenship and Immigration Canada	2,664	1,281	1,135	48.1	42.6
Passport Canada	1,631	758	651	46.5	39.9
Canadian Heritage	2,367	1,008	916	42.6	38.7
Public Health Agency of Canada	2,262	910	824	40.2	36.4
Canadian International Development Agency	1,884	661	589	35.1	31.3
Library and Archives Canada	1,142	530	472	46.4	41.3
Treasury Board of Canada Secretariat	1,155	562	524	48.7	45.4
Department of Finance Canada	1,077	449	397	41.7	36.9
Public Safety Canada	1,048	438	411	41.8	39.2
Public Service Commission of Canada	1,029	534	495	51.9	48.1
Immigration and Refugee Board of Canada	874	423	366	48.4	41.9

Department/Agency Name	Sample Size	Responding Employees		Response Rate (%)	
		Master	Share	Master	Share
Canada School of Public Service	911	333	307	36.6	33.7
Privy Council Office	824	230	203	27.9	24.6
Public Prosecution Service of Canada	796	340	293	42.7	36.8
Atlantic Canada Opportunities Agency	746	403	370	54.0	49.6
Canadian Grain Commission	661	334	286	50.5	43.3
Canadian Space Agency	591	362	320	61.3	54.1
Canada Public Service Agency	598	224	204	37.5	34.1
Courts Administration Service	587	230	206	39.2	35.1
Office of the Superintendent of Financial Institutions Canada	479	218	198	45.5	41.3
Canadian Radio-television and Telecommunications Commission	441	187	160	42.4	36.3
Office of the Chief Electoral Officer	533	241	212	45.2	39.8
Economic Development Agency of Canada for the Regions of Quebec	410	203	185	49.5	45.1
Western Economic Diversification Canada	417	227	211	54.4	50.6
National Parole Board	373	159	147	42.6	39.4
National Energy Board	322	163	153	50.6	47.5
Survey of Staffing - Candidates Response Rate	77,214	34,334	30,823	44.5	39.9

8.2 Survey Errors

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

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

Over a large number of observations, randomly occurring errors will have little effect on estimates derived from the survey. However, errors occurring systematically will contribute to biases in the survey estimates. Considerable time and effort were taken to reduce non-sampling errors in the survey. Quality assurance measures were implemented at each step of the questionnaire development, data collection and processing cycle to monitor the quality of the data. These measures include focus group testing to detect problems of questionnaire design or misunderstanding of instructions, the use of highly tested computerized questionnaire applications, procedures to ensure that data capture errors were minimized, and edit quality checks to verify the processing logic.

8.2.1 The Frame

The sampling frame was made up of all in-scope employees from the Public Works and Government Services Canada's Incumbent file. Since the contact information for each employee (e-mail and postal address) was not available from the Incumbent file, it had to be collected by Statistics Canada from the departments through Article 13 of the Statistics Act. The files received from the departments were linked to the Public Works and Government Services Canada's Incumbent file to create the survey frame.

The record linkage was done using the variables "Last name", "First name initial", and "Department" that are available on both the SOS employee files and the Incumbent File. In the case of duplicates, the record linkage was done using the variables "Initial", "Province of work", "Sex", "Position occupational group and level", "Employee occupational group and level" and "Department Start Date" when provided by the departments. The average link rate was 92%, which resulted in only 8% of the sampled employees with no contact information.

8.2.2 Data Collection

A description of the objectives of the survey was provided to the respondents, as well as a glossary of terms. A set of questions and answers was also provided on the Information for Survey Participants, on the Statistics Canada Internet site.

The Statistics Canada Help Desk (1-800 and e-mail) provided support for participants who had questions during collection or needed technical assistance.

The survey was conducted from January 7th to February 20th, 2009.

8.2.3 Data Processing

Data processing of the SOS was done in a number of steps including verification, editing, estimation, confidentiality, etc. At each step a picture of the output files is taken and an easy verification can be made comparing files at the current and previous step. This greatly improved the data processing stage.

8.2.4 Non-response

A major source of non-sampling errors in surveys is the effect of non-response on the survey results. The extent of non-response varies from partial non-response (failure to answer just one or some questions) to total non-response. Total non-response occurred because employee contact information from the department was not obtained, the contact information was incorrect, the respondent had problems accessing the electronic questionnaire, or the respondent refused to participate in the survey. Total non-response was handled by adjusting the weight of employees who responded to the survey to compensate for those who did not respond.

In most cases, item non-response to the survey occurred when the respondent did not understand or misinterpreted a question, refused to answer a question, or could not recall the requested information. For item non-response a "Not stated" code was assigned to the item.

During the electronic collection, partial non-response occurred when the respondent saved the questionnaire but did not submit it for various reasons. In the case of the SOS,

only 0.5% of respondents saved the questionnaire but did not submit it. The missed questions were treated as multiple item non-response and coded to “Not stated”.

8.2.5 Measurement of Sampling Error

Since it is an unavoidable fact that estimates from a sample survey are subject to sampling error, sound statistical practice calls for researchers to provide users with some indication of the magnitude of this sampling error. This section of the documentation outlines the measures of sampling error which Statistics Canada commonly uses and which it urges users producing estimates from this microdata file to use also.

The basis for measuring the potential size of sampling errors 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 resulting measure, known as the coefficient of variation (CV) of an estimate, is obtained by dividing the standard error of the estimate by the estimate itself and is expressed as a percentage of the estimate.

For example, suppose that, based upon the survey results, one estimates that 52% of federal public servants participated in a staffing activity between October 1, 2007 and September 30, 2008 and this estimate is found to have a standard error of 0.005. Then the coefficient of variation of the estimate is calculated as:

$$\left(\frac{0.005}{0.52} \right) \times 100 \% = 0.96 \%$$

There is more information on the calculation of coefficients of variation in Chapter 10.0.

9.0 Guidelines for Tabulation, Analysis and Release

This chapter of the documentation outlines the guidelines to be adhered to by users tabulating, analysing, 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.

9.1 Rounding Guidelines

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

- a) Estimates in the main body of a statistical table are to be rounded to the nearest hundred units using the normal rounding technique. In normal rounding, if the first or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is raised by one. For example, in normal rounding to the nearest 100, if the last two digits are between 00 and 49, they are changed to 00 and the preceding digit (the hundreds digit) is left unchanged. If the last digits are between 50 and 99 they are changed to 00 and the preceding digit is incremented by 1.
- b) Marginal sub-totals and totals in statistical tables are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units using normal rounding.
- c) 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.

9.2 Sample Weighting Guidelines for Tabulation

The sample design used for the Survey of Staffing – Candidates (SOS) Cycle 1 was not self-weighting. When producing simple estimates including the production of ordinary statistical tables, users must apply the proper survey weight.

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

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

9.3 Definitions of Types of Estimates: Categorical and Quantitative

Before discussing how the SOS data can be tabulated and analyzed, it is useful to describe the two main types of point estimates of population characteristics which can be generated from the microdata file for the SOS.

9.3.1 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 of employees who participated in a staffing activity for a deployment or the proportion of employees who worked in the National Capital Region when the staffing activity concluded for them, are examples of such estimates. An estimate of the number of employees possessing a certain characteristic may also be referred to as an estimate of an aggregate.

Examples of Categorical Questions:

Q: Was it a staffing activity for a deployment?

R: Yes / No

Q: Where were you working before this staffing activity concluded for you?

R: National Capital Region / Atlantic / Quebec / Ontario / Prairies, Nunavut, Northwest Territories / British Columbia, Yukon / Outside Canada

9.3.2 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 number of consecutive months employees were in an acting position when the staffing activity concluded for them. The numerator is an estimate of the total number of consecutive months employees were in an acting position when the staffing activity concluded for them, and its denominator is the estimate of the number of employees who were in an acting position when the staffing activity concluded for them.

Examples of Quantitative Questions:

Q: How many consecutive months were you in that acting position when this staffing activity concluded for you?

R: |_|_|_| months

Q: In months, what was the total elapsed time for this staffing activity?

R: |_|_| months

9.3.3 Results from Scale-type Questions: Percentage of Favourable Response

The Survey of Staffing contains scale-type questions where the respondents are asked to rate their agreement or disagreement. The total number of responses is composed of "favourable" and "unfavourable" responses. Reporting the results in terms of the percentage of favourable responses is a standard practice that is widely used for scale-type surveys. This is because evaluating the results is easier when all of the favourable ratings on a question are combined into a single rating. In addition, the results from question to question are consistent.

The percentage of favourable responses is obtained by:

- a) summing the weights of records having a favourable response to obtain the numerator (\hat{X}),
- b) summing the weights of all records having a response (do not include the "Not stated") to obtain the denominator (\hat{Y}),
- c) dividing the numerator (\hat{X}) by the denominator (\hat{Y}),
- d) multiply by 100, then
- e) round to units.

For scale questions with more than three points on the scale, the favourable groups "To a great extent" and "To a moderate extent" may be grouped to obtain the percentage of favourable responses.

For example, for Question D_Q10A, "To what extent do you think your knowledge was considered important by the hiring manager in this staffing activity, as it pertained to you?" the responses "To a great extent" and "To a moderate extent" should be grouped to obtain the percentage of favourable responses.

9.3.4 Tabulation of Categorical Estimates

Estimates of the number of people 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. 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) dividing estimate a) by estimate b) (\hat{X} / \hat{Y}).

9.3.5 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 (\hat{X}) of the total number of consecutive months employees were in an acting position when the staffing activity concluded for them, multiply the value reported in variable C_Q08 (x_i) (number of

months) by the final weight (w_i) for the record, then sum this value over all records with C_Q07 = 1 (employees who were in an acting position)

$$\hat{X} = \sum_{Q_C7=1} x_i w_i$$

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 number of consecutive months employees were in an acting position when the staffing activity concluded for them.

- a) estimate the total number of consecutive months (\hat{X}) as described above,
- b) estimate the number of employees who were in an acting position when the staffing activity concluded for them (\hat{Y}) by summing the final weights of all records with C_Q07 = 1, then
- c) divide estimate a) by estimate b) (\hat{X} / \hat{Y}).

9.3.6 **Percentage of Favourable Response: Evaluation Guidelines**

Before releasing and/or publishing any estimate from the SOS users should first determine the data quality of the estimate. Data quality is affected by non-sampling errors as discussed in Chapter 8.0. Users should be sure to read this chapter to be more fully aware of the quality characteristics of these data.

The following table, extracted from William Davidson's (1979) *How to Develop and Conduct Successful Employee Attitude Surveys*, may be used as a guide to evaluate the percentage of favourable responses.

<u>Favourable Response</u>	<u>Evaluation</u>
90% or more	Highly meaningful favourable response
75% - 89%	Quite meaningful favourable response
65% - 74%	Suggestive of favourable response
35% - 64%	Requires further study
25% - 34%	Suggestive of unfavourable response
11% - 24%	Quite meaningful unfavourable response
10% or less	Highly meaningful unfavourable response

Davidson explained that the above table is based on the fact that favourable responses in the range of 35% to 64% do not show either favourable or unfavourable responses. It is clear that a 50% favourable response on an item indicates no trend whatsoever, as equal numbers of employees reacted both favourably and unfavourably. Questions that receive favourable responses in the 35% to 64% range should be further explored through, perhaps, follow-up discussions. Favourable response reactions below 34% indicate problem areas and may warrant immediate attention.

In addition, the number of respondents who contribute to the calculation of the percentage of favourable response should be determined. When comparing percentages, users should be cautious if the percentages are of different total quantities.

9.4 Guidelines for Statistical Analysis

The SOS is based upon a sample design with stratification. In order for survey estimates and analyses to be free from bias, the survey weights must be used.

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

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 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 respondents from the file who reported G_Q07 = male;
- 2) calculate the AVERAGE weight for these records by summing the original employee weights from the microdata file for these records and then dividing by the number of employees who reported G_Q07 = male;
- 3) for each of these respondents, calculate a RESCALED weight equal to the original employee weight divided by the AVERAGE weight;
- 4) perform the analysis for these employees using the RESCALED weight.

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.

9.5 Coefficient of Variation Release Guidelines

Before releasing and/or publishing any estimates from the SOS, users should first determine the quality level of the estimate. The quality levels are *acceptable*, *marginal* and *unacceptable*. Data quality is affected by both sampling and non-sampling errors as discussed in Chapter 8.0. However for this purpose, the quality level of an estimate will be determined only on the basis of sampling error as reflected by the coefficient of variation as shown in the table below. Nonetheless users should be sure to read Chapter 8.0 to be more fully aware of the quality characteristics of these data.

First, the number of respondents who contribute to the calculation of the estimate should be determined. If this number is less than 30, the weighted estimate should be considered to be of unacceptable quality.

For weighted estimates based on sample sizes of 30 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 rounded weighted 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

Quality Level of Estimate	Guidelines
1) Acceptable	<p>Estimates have a sample size of 30 or more, and low coefficients of variation in the range of 0.0% to 16.5%.</p> <p>No warning is required.</p>
2) Marginal	<p>Estimates have a sample size of 30 or more, and high coefficients of variation in the range of 16.6% to 33.3%.</p> <p>Estimates should be flagged with the letter E (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 a sample size of less than 30, or 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 F (or some similar identifier) and the following warning should accompany the estimates:</p> <p>"Please be warned that these estimates [flagged with the letter F] do not meet Statistics Canada's quality standards. Conclusions based on these data will be unreliable, and most likely invalid."</p>

9.6 Release Cut-off's for the Survey of Staffing - Candidates

9.6.1 Precision Guidelines

The following table provides an indication of the precision of population estimates as it shows the release cut-offs associated with each of the three quality levels presented in the previous section. These cut-offs are derived from the coefficient of variation (CV) tables discussed in Chapter 10.

For example, the table shows that the quality of a weighted estimate of 1,000 employees possessing a given characteristic at National Defence is marginal.

Note that these cut-offs apply to estimates of population totals only. To estimate ratios, users should not use the numerator value (nor the denominator) in order to find the

corresponding quality level. Rule 4 in Section 10.1 and Example 4 in Section 10.1.1 explains the correct procedure to be used for ratios.

Department/Agency Name	Acceptable CV 0.0% to 16.5%	Marginal CV 16.6% to 33.3%	Unacceptable CV > 33.3%
National Defence	1,250 & over	320 to < 1,250	under 320
Service Canada and Human Resources and Social Development Canada	580 & over	150 to < 580	under 150
Correctional Service Canada	500 & over	120 to < 500	under 120
Canada Border Services Agency	550 & over	140 to < 550	under 140
Public Works and Government Services Canada	320 & over	80 to < 320	under 80
Fisheries and Oceans Canada	630 & over	160 to < 630	under 160
Health Canada	330 & over	80 to < 330	under 80
Agriculture and Agri-Food Canada	390 & over	100 to < 390	under 100
Royal Canadian Mounted Police	340 & over	90 to < 340	under 90
Environment Canada	130 & over	30 to < 130	under 30
Statistics Canada	110 & over	30 to < 110	under 30
Industry Canada	110 & over	30 to < 110	under 30
Transport Canada	80 & over	20 to < 80	under 20
Indian and Northern Affairs Canada and Indian Residential Schools Resolution Canada	150 & over	40 to < 150	under 40
Foreign Affairs and International Trade Canada	150 & over	40 to < 150	under 40
Natural Resources Canada	80 & over	20 to < 80	under 20
Department of Justice Canada	120 & over	30 to < 120	under 30
Veterans Affairs Canada	100 & over	30 to < 100	under 30
Citizenship and Immigration Canada	110 & over	30 to < 110	under 30
Passport Canada	120 & over	30 to < 120	under 30
Canadian Heritage	60 & over	20 to < 60	under 20
Public Health Agency of Canada	70 & over	20 to < 70	under 20
Canadian International Development Agency	80 & over	20 to < 80	under 20
Library and Archives Canada	50 & over	10 to < 50	under 10
Treasury Board of Canada Secretariat	50 & over	10 to < 50	under 10
Department of Finance Canada	60 & over	20 to < 60	under 20
Public Safety Canada	70 & over	20 to < 70	under 20
Public Service Commission of Canada	40 & over	10 to < 40	under 10
Immigration and Refugee Board of Canada	60 & over	10 to < 60	under 10
Canada School of Public Service	90 & over	30 to < 90	under 30
Privy Council Office	120 & over	30 to < 120	under 30
Public Prosecution Service of Canada	60 & over	20 to < 60	under 20
Atlantic Canada Opportunities Agency	40 & over	10 to < 40	under 10
Canadian Grain Commission	50 & over	10 to < 50	under 10

Department/Agency Name	Acceptable CV 0.0% to 16.5%	Marginal CV 16.6% to 33.3%	Unacceptable CV > 33.3%
Canadian Space Agency	30 & over	10 to < 30	under 10
Canada Public Service Agency	80 & over	20 to < 80	under 20
Courts Administration Service	60 & over	20 to < 60	under 20
Office of the Superintendent of Financial Institutions Canada	50 & over	10 to < 50	under 10
Canadian Radio-television and Telecommunications Commission	70 & over	20 to < 70	under 20
Office of the Chief Electoral Officer	80 & over	20 to < 80	under 20
Economic Development Agency of Canada for the Regions of Quebec	50 & over	10 to < 50	under 10
Western Economic Diversification Canada	40 & over	10 to < 40	under 10
National Parole Board	70 & over	20 to < 70	under 20
National Energy Board	40 & over	10 to < 40	under 10
Survey of Staffing – Target Population	470 & over	110 to < 470	under 110

9.6.2 Confidentiality Guidelines

In order to protect the confidentiality of the SOS respondents, the following guidelines provide the rules to use before releasing the employee estimates.

The number of respondents who contribute to the calculation of the demographic estimate should be determined. If this number is less than 10, the estimate should not be released.

Examples of number of respondents by demographic cell:

Member of a visible minority group (G_Q10 = Yes)
Staffing activity stopped before an appointment decision was made
(E_Q05 = Yes)

- a) Member of a visible minority group; 1050 respondents
- b) The staffing activity stopped before an appointment decision was made; 165 respondents
- c) Member of a visible minority group * The staffing activity stopped before an appointment decision was made; 8 respondents

The examples show that the estimate of employees who were “members of a visible minority group” and the estimate of the number of employees who had a “staffing activity stopped before an appointment decision was made” could be released. However the estimate produced when the two variables are cross-tabulated should not be released to protect the confidentiality of the respondents.

9.7 Other Types of Analysis

The opportunities for other types of statistical analysis (e.g., hypothesis testing, ANOVA, factor analysis) are numerous, particularly if a specialist is involved. It is beyond the scope of this paper to describe all the various possibilities. In order for results to be free from bias, the weights must be used.

The sequence in which survey findings are analysed usually follows some predetermined pattern. Typically general level results are produced first, followed by analysis at finer levels. For example, it may be useful to compare results across different occupational groups of employees. Further insight into the results can be gained by examining different tenure groups, by gender, by age group, etc.

10.0 Approximate Sampling Variability Tables

In order to supply coefficients of variation (CV) which would be applicable to a wide variety of categorical estimates produced from this microdata file and which could be readily accessed by the user, a set of Approximate Sampling Variability Tables has been produced. These CV tables allow the user to obtain an approximate coefficient of variation based on the size of the estimate calculated from the survey data.

The coefficients of variation are derived using the variance formula for simple random sampling and incorporating a factor which reflects the multi-stage, clustered nature of the sample design. This factor, known as the design effect, was determined by first calculating design effects for a wide range of characteristics and then choosing from among these a conservative value (usually the 75th percentile) to be used in the CV tables which would then apply to the entire set of characteristics.

The table below shows the conservative value of the design effects as well as sample sizes (number of respondents on the share file) and population counts by department which were used to produce the Approximate Sampling Variability Tables for the Survey of Staffing (SOS).

Department/Agency Name	Design Effect	Sample Size	Population
National Defence	1.27	878	25,627
Service Canada and Human Resources and Social Development Canada	1.01	1,375	23,545
Correctional Service Canada	1.18	1,270	16,277
Canada Border Services Agency	1.04	886	14,172
Public Works and Government Services Canada	1.06	1,325	12,672
Fisheries and Oceans Canada	1.59	880	11,005
Health Canada	1.06	1,001	9,759
Agriculture and Agri-Food Canada	1.44	805	7,134
Royal Canadian Mounted Police	1.42	881	6,833
Environment Canada	1.03	1,395	6,310
Statistics Canada	1.00	1,370	5,731
Industry Canada	1.01	1,287	5,339
Transport Canada	1.03	1,615	5,114
Indian and Northern Affairs Canada and Indian Residential Schools Resolution Canada	1.08	966	4,657
Foreign Affairs and International Trade Canada	1.20	1,024	4,672
Natural Resources Canada	1.01	1,362	4,442
Department of Justice Canada	1.06	1,065	4,459
Veterans Affairs Canada	1.16	1,203	4,153
Citizenship and Immigration Canada	1.20	1,135	4,076
Passport Canada	1.07	651	2,756
Canadian Heritage	1.08	916	2,367
Public Health Agency of Canada	1.16	824	2,262
Canadian International Development Agency	1.06	589	1,884
Library and Archives Canada	1.09	472	1,142

Department/Agency Name	Design Effect	Sample Size	Population
Treasury Board of Canada Secretariat	1.07	524	1,155
Department of Finance Canada	1.08	397	1,077
Public Safety Canada	1.34	411	1,048
Public Service Commission of Canada	1.12	495	1,029
Immigration and Refugee Board of Canada	1.18	366	874
Canada School of Public Service	1.45	307	911
Privy Council Office	1.28	203	824
Public Prosecution Service of Canada	1.06	293	796
Atlantic Canada Opportunities Agency	1.16	370	746
Canadian Grain Commission	1.12	286	661
Canadian Space Agency	1.02	320	591
Canada Public Service Agency	1.30	204	598
Courts Administration Service	1.05	206	587
Office of the Superintendent of Financial Institutions Canada	0.99	198	479
Canadian Radio-television and Telecommunications Commission	1.24	160	441
Office of the Chief Electoral Officer	1.64	212	533
Economic Development Agency of Canada for the Regions of Quebec	1.17	185	410
Western Economic Diversification Canada	1.26	211	417
National Parole Board	1.40	147	373
National Energy Board	1.10	153	322
Survey of Staffing – Target Population	2.32	30,823	200,260

All coefficients of variation in the Approximate Sampling Variability Tables are approximate and, therefore, unofficial. Estimates of actual variance for specific variables may be obtained from Statistics Canada on a cost-recovery basis. Since the approximate CV is conservative, the use of actual variance estimates may cause the estimate to be switched from one quality level to another. For instance a *marginal* estimate could become *acceptable* based on the exact CV calculation.

Remember: If the number of observations on which an estimate is based is less than 30, the weighted estimate is most likely unacceptable and Statistics Canada recommends not to release such an estimate, regardless of the value of the coefficient of variation.

10.1 How to Use the Coefficient of Variation Tables for Categorical Estimates

The following rules should enable the user to determine the approximate coefficients of variation from the Approximate Sampling Variability Tables for estimates of the number, proportion or percentage of the surveyed population possessing a certain characteristic and for ratios and differences between such estimates.

Rule 1: Estimates of Numbers of Employees Possessing a Characteristic (Aggregates)

The coefficient of variation depends only on the size of the estimate itself. On the Approximate Sampling Variability Table for the appropriate department, locate the estimated number in the left-most column of the table (headed “Numerator of Percentage”) and follow the asterisks (if any) across to the first figure encountered. This figure is the approximate coefficient of variation.

Rule 2: Estimates of Proportions or Percentages of Employees Possessing a Characteristic

The coefficient of variation of an estimated proportion or percentage depends on both the size of the proportion or percentage and the size of the total upon which the proportion or percentage is based. Estimated proportions or percentages are relatively more reliable than the corresponding estimates of the numerator of the proportion or percentage, when the proportion or percentage is based upon a subset of the total population. For example, if the subset under consideration is all employees who participated as a candidate in a staffing activity during the period from October 1st, 2007 to September 30th, 2008 (A_Q01), then the estimated proportion of candidates for whom the staffing activity concluded during that period (A_Q02) is more reliable than the estimated number of candidates for whom the staffing activity concluded during that period. (Note that in the tables the coefficients of variation decline in value reading from left to right).

When the proportion or percentage is based upon the total population, the CV of the proportion or percentage is the same as the CV of the numerator of the proportion or percentage. In this case, Rule 1 can be used.

When the proportion or percentage is based upon a subset of the total population (as in the example above) reference should be made to the proportion or percentage (across the top of the table) and to the numerator of the proportion or percentage (down the left side of the table). The intersection of the appropriate row and column gives the coefficient of variation.

Rule 3: Estimates of Differences Between Aggregates or Percentages

The standard error of a difference between two estimates is approximately equal to the square root of the sum of squares of each standard error considered separately. That is, the standard error of a difference $(\hat{d} = \hat{X}_1 - \hat{X}_2)$ is:

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1\alpha_1)^2 + (\hat{X}_2\alpha_2)^2}$$

where \hat{X}_1 is estimate 1, \hat{X}_2 is estimate 2, and α_1 and α_2 are the coefficients of variation of \hat{X}_1 and \hat{X}_2 respectively. The coefficient of variation of \hat{d} is given by $\sigma_{\hat{d}} / \hat{d}$. This formula is accurate for the difference between separate and uncorrelated characteristics, but is only approximate otherwise.

Rule 4: Estimates of Ratios

In the case where the numerator is a subset of the denominator, the ratio should be converted to a percentage and Rule 2 applied. This would apply, for example, to the case where the denominator is the number of employees who participated as a candidate in a staffing activity during the reference period and the numerator is the number of candidates for whom the staffing activity concluded during that period.

In the case where the numerator is not a subset of the denominator, as for example, the ratio of

the number of employees who participated as a candidate in a staffing activity during the reference period to the number of employees who did not participate as a candidate in a staffing activity during that period, the standard error of the ratio of the estimates is approximately equal to the square root of the sum of squares of each coefficient of variation considered separately multiplied by \hat{R} . That is, the standard error of a ratio ($\hat{R} = \hat{X}_1 / \hat{X}_2$) is:

$$\sigma_{\hat{R}} = \hat{R} \sqrt{\alpha_1^2 + \alpha_2^2}$$

where α_1 and α_2 are the coefficients of variation of \hat{X}_1 and \hat{X}_2 respectively. The coefficient of variation of \hat{R} is given by $\sigma_{\hat{R}} / \hat{R}$. The formula will tend to overstate the error, if \hat{X}_1 and \hat{X}_2 are positively correlated and understate the error if \hat{X}_1 and \hat{X}_2 are negatively correlated.

Rule 5: Estimates of Differences of Ratios

In this case, Rules 3 and 4 are combined. The CVs for the two ratios are first determined using Rule 4, and then the CV of their difference is found using Rule 3.

10.1.1 Examples of Using the Coefficient of Variation Tables for the Categorical Estimates

The following examples based on the Survey of Staffing – Candidates (SOS) Cycle 1 are included to assist users in applying the foregoing rules.

Example 1: Estimates of Numbers of Employees Possessing a Characteristic (Aggregates)

Suppose that a user estimates that 42,393 candidates (D_Q01 = 1) underwent a formal assessment as part of their participation in a staffing activity. How does the user determine the coefficient of variation of this estimate?

- 1) Refer to the coefficient of variation table for the Target Population (All departments which fell under the Public Service Employment Act and had at least 350 employees or more).
- 2) The estimated aggregate 42,393 does not appear in the left-hand column (the “Numerator of Percentage” column), so it is necessary to use the figure closest to it, namely 40,000.
- 3) The coefficient of variation for an estimated aggregate is found by referring to the first non-asterisk entry on that row, namely, 1.6%.
- 4) So the approximate coefficient of variation of the estimate is 1.6%. The estimate that 42,393 (to be rounded according to the rounding guidelines in Section 9.1) candidates underwent a formal assessment as part of their participation in a staffing activity is publishable with no qualifications.

SURVEY OF STAFFING – CANDIDATES, 2009 CYCLE 1 - SHARE FILE

Approximate Sampling Variability Tables

Target population (All departments which fell under the Public Service Employment Act and had at least 350 employees or more)

NUMERATOR OF PERCENTAGE	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
10	112.9	112.4	111.8	110.1	107.1	104.1	101.0	97.8	94.5	91.0	87.5	79.9	61.9	35.7
20	79.8	79.5	79.1	77.8	75.8	73.6	71.4	69.2	66.8	64.4	61.9	56.5	43.7	25.3
30	65.2	64.9	64.5	63.6	61.9	60.1	58.3	56.5	54.6	52.6	50.5	46.1	35.7	20.6
40	56.4	56.2	55.9	55.0	53.6	52.1	50.5	48.9	47.2	45.5	43.7	39.9	30.9	17.9
50	50.5	50.3	50.0	49.2	47.9	46.6	45.2	43.7	42.3	40.7	39.1	35.7	27.7	16.0
60	46.1	45.9	45.6	44.9	43.7	42.5	41.2	39.9	38.6	37.2	35.7	32.6	25.3	14.6
70	42.7	42.5	42.3	41.6	40.5	39.4	38.2	37.0	35.7	34.4	33.1	30.2	23.4	13.5
80	39.9	39.7	39.5	38.9	37.9	36.8	35.7	34.6	33.4	32.2	30.9	28.2	21.9	12.6
90	37.6	37.5	37.3	36.7	35.7	34.7	33.7	32.6	31.5	30.3	29.2	26.6	20.6	11.9
100	35.7	35.5	35.4	34.8	33.9	32.9	31.9	30.9	29.9	28.8	27.7	25.3	19.6	11.3
.														
.														
.														
10,000	*****	*****	*****	3.5	3.4	3.3	3.2	3.1	3.0	2.9	2.8	2.5	2.0	1.1
15,000	*****	*****	*****	*****	2.8	2.7	2.6	2.5	2.4	2.4	2.3	2.1	1.6	0.9
20,000	*****	*****	*****	*****	2.4	2.3	2.3	2.2	2.1	2.0	2.0	1.8	1.4	0.8
30,000	*****	*****	*****	*****	*****	1.9	1.8	1.8	1.7	1.7	1.6	1.5	1.1	0.7
40,000	*****	*****	*****	*****	*****	*****	1.6	1.5	1.5	1.4	1.4	1.3	1.0	0.6
50,000	*****	*****	*****	*****	*****	*****	*****	1.4	1.3	1.3	1.2	1.1	0.9	0.5
60,000	*****	*****	*****	*****	*****	*****	*****	*****	1.2	1.2	1.1	1.0	0.8	0.5
70,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	1.1	1.0	1.0	0.7	0.4
80,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	1.0	0.9	0.7	0.4
90,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.8	0.7	0.4
100,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.8	0.6	0.4
125,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.6	0.3
150,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.3

NOTE: For correct usage of these tables, please refer to the microdata documentation.

Example 2: Estimates of Proportions or Percentages of Employees Possessing a Characteristic

Suppose that the user estimates that $23,656 / 42,393 = 55.8\%$ of the candidates who underwent a formal assessment as part of their participation in a staffing activity received an offer of appointment ($D_Q01 = 1$ and $E_Q06 = 1$). How does the user determine the coefficient of variation of this estimate?

- 1) Refer to the coefficient of variation table for the Target Population (All departments which fell under the Public Service Employment Act and had at least 350 employees or more).
- 2) Because the estimate is a percentage which is based on a subset of the total population (i.e., candidates who underwent a formal assessment as part of their participation in a staffing activity), it is necessary to use both the percentage itself

(55.8%) and the numerator portion of the percentage (23,656) in determining the coefficient of variation.

- 3) The numerator (23,656) does not appear in the left-hand column (the “Numerator of Percentage” column) so it is necessary to use the figure closest to it, namely 20,000. Similarly, the percentage estimate (55.8%) does not appear as any of the column headings, so it is necessary to use the percentage closest to it, namely 50.0%.
- 4) The figure at the intersection of the row and column used, namely 1.8%, is the coefficient of variation to be used.
- 5) So the approximate coefficient of variation of the estimate is 1.8%. The estimate that 55.8% of the candidates who underwent a formal assessment as part of their participation in a staffing activity received an offer of appointment can be published with no qualifications.

SURVEY OF STAFFING – CANDIDATES, 2009 CYCLE 1 - SHARE FILE

Approximate Sampling Variability Tables

Target population (All departments which fell under the Public Service Employment Act and had at least 350 employees or more)

NUMERATOR OF PERCENTAGE	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
10	112.9	112.4	111.8	110.1	107.1	104.1	101.0	97.8	94.5	91.0	87.5	79.9	61.9	35.7
20	79.8	79.5	79.1	77.8	75.8	73.6	71.4	69.2	66.8	64.4	61.9	56.5	43.7	25.3
30	65.2	64.9	64.5	63.6	61.9	60.1	58.3	56.5	54.6	52.6	50.5	46.1	35.7	20.6
40	56.4	56.2	55.9	55.0	53.6	52.1	50.5	48.9	47.2	45.5	43.7	39.9	30.9	17.9
50	50.5	50.3	50.0	49.2	47.9	46.6	45.2	43.7	42.3	40.7	39.1	35.7	27.7	16.0
60	46.1	45.9	45.6	44.9	43.7	42.5	41.2	39.9	38.6	37.2	35.7	32.6	25.3	14.6
70	42.7	42.5	42.3	41.6	40.5	39.4	38.2	37.0	35.7	34.4	33.1	30.2	23.4	13.5
.														
.														
.														
10,000	*****	*****	*****	3.5	3.4	3.3	3.2	3.1	3.0	2.9	2.8	2.5	2.0	1.1
15,000	*****	*****	*****	*****	2.8	2.7	2.6	2.5	2.4	2.4	2.3	2.1	1.6	0.9
20,000	*****	*****	*****	*****	2.4	2.3	2.3	2.2	2.1	2.0	2.0	1.8	1.4	0.8
30,000	*****	*****	*****	*****	*****	1.9	1.8	1.8	1.7	1.7	1.6	1.5	1.1	0.7
40,000	*****	*****	*****	*****	*****	*****	1.6	1.5	1.5	1.4	1.4	1.3	1.0	0.6
50,000	*****	*****	*****	*****	*****	*****	*****	1.4	1.3	1.3	1.2	1.1	0.9	0.5
60,000	*****	*****	*****	*****	*****	*****	*****	*****	1.2	1.2	1.1	1.0	0.8	0.5
70,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	1.1	1.0	1.0	0.7	0.4
80,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	1.0	0.9	0.7	0.4
90,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.8	0.7	0.4
100,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.8	0.6	0.4
125,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.6	0.3
150,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.3

NOTE: For correct usage of these tables, please refer to the microdata documentation.

Example 3: Estimates of Differences Between Aggregates or Percentages

Suppose that a user estimates that $23,656 / 42,393 = 55.8\%$ of the candidates who underwent a formal assessment as part of their participation in a staffing activity received an offer of appointment ($D_Q01 = 1$ and $E_Q06 = 1$). The user also estimates that $2,383 / 6,924 = 34.4\%$ of the candidates who did not undergo a formal assessment also received an offer of appointment ($D_Q01 = 2$ and $E_Q06 = 1$). How does the user determine the coefficient of variation of the difference between these two estimates?

- 1) Using the Target Population coefficient of variation table in the same manner as described in Example 2 gives the CV of the estimated percentage for candidates who underwent a formal assessment and received an offer of appointment as 1.8%. Using the same table in a similar manner gives a CV of the estimated percentage for candidates who did not undergo a formal assessment and also received an offer of appointment as 5.8%.

- 2) Using Rule 3, the standard error of a difference ($\hat{d} = \hat{X}_1 - \hat{X}_2$) is:

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

where \hat{X}_1 is estimate 1 (percentage of candidates who underwent a formal assessment and received an offer of appointment), \hat{X}_2 is estimate 2 (percentage of candidates who did not undergo a formal assessment and also received an offer of appointment), and α_1 and α_2 are the coefficients of variation of \hat{X}_1 and \hat{X}_2 respectively.

That is, the standard error of the difference $\hat{d} = 0.558 - 0.344 = 0.214$ (21.4%) is:

$$\begin{aligned} \sigma_{\hat{d}} &= \sqrt{[(0.558)(0.018)]^2 + [(0.344)(0.058)]^2} \\ &= \sqrt{(0.000101) + (0.000399)} \\ &= 0.022 \end{aligned}$$

- 3) The coefficient of variation of \hat{d} is given by $\sigma_{\hat{d}} / \hat{d} = 0.022 / 0.214 = 0.103$.
- 4) So the approximate coefficient of variation of the difference between the estimates is 10.3%. The difference between the estimates is considered acceptable and this estimate can be released with no qualifications.

SURVEY OF STAFFING – CANDIDATES, 2009 CYCLE 1 - SHARE FILE

Approximate Sampling Variability Tables

Target population (All departments which fell under the Public Service Employment Act and had at least 350 employees or more)

NUMERATOR OF PERCENTAGE	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
10	112.9	112.4	111.8	110.1	107.1	104.1	101.0	97.8	94.5	91.0	87.5	79.9	61.9	35.7
20	79.8	79.5	79.1	77.8	75.8	73.6	71.4	69.2	66.8	64.4	61.9	56.5	43.7	25.3
30	65.2	64.9	64.5	63.6	61.9	60.1	58.3	56.5	54.6	52.6	50.5	46.1	35.7	20.6
40	56.4	56.2	55.9	55.0	53.6	52.1	50.5	48.9	47.2	45.5	43.7	39.9	30.9	17.9
50	50.5	50.3	50.0	49.2	47.9	46.6	45.2	43.7	42.3	40.7	39.1	35.7	27.7	16.0
.														
.														
.														
1,250	*****	10.1	10.0	9.8	9.6	9.3	9.0	8.7	8.5	8.1	7.8	7.1	5.5	3.2
1,500	*****	9.2	9.1	9.0	8.7	8.5	8.2	8.0	7.7	7.4	7.1	6.5	5.1	2.9
2,000	*****	7.9	7.9	7.8	7.6	7.4	7.1	6.9	6.7	6.4	6.2	5.6	4.4	2.5
2,500	*****	*****	7.1	7.0	6.8	6.6	6.4	6.2	6.0	5.8	5.5	5.1	3.9	2.3
3,000	*****	*****	6.5	6.4	6.2	6.0	5.8	5.6	5.5	5.3	5.1	4.6	3.6	2.1
3,500	*****	*****	6.0	5.9	5.7	5.6	5.4	5.2	5.1	4.9	4.7	4.3	3.3	1.9
4,000	*****	*****	5.6	5.5	5.4	5.2	5.1	4.9	4.7	4.6	4.4	4.0	3.1	1.8
4,500	*****	*****	*****	5.2	5.1	4.9	4.8	4.6	4.5	4.3	4.1	3.8	2.9	1.7
5,000	*****	*****	*****	4.9	4.8	4.7	4.5	4.4	4.2	4.1	3.9	3.6	2.8	1.6
7,500	*****	*****	*****	4.0	3.9	3.8	3.7	3.6	3.5	3.3	3.2	2.9	2.3	1.3
10,000	*****	*****	*****	3.5	3.4	3.3	3.2	3.1	3.0	2.9	2.8	2.5	2.0	1.1
15,000	*****	*****	*****	*****	2.8	2.7	2.6	2.5	2.4	2.4	2.3	2.1	1.6	0.9
20,000	*****	*****	*****	*****	2.4	2.3	2.3	2.2	2.1	2.0	2.0	1.8	1.4	0.8
30,000	*****	*****	*****	*****	*****	1.9	1.8	1.8	1.7	1.7	1.6	1.5	1.1	0.7
40,000	*****	*****	*****	*****	*****	*****	1.6	1.5	1.5	1.4	1.4	1.3	1.0	0.6
50,000	*****	*****	*****	*****	*****	*****	*****	1.4	1.3	1.3	1.2	1.1	0.9	0.5
60,000	*****	*****	*****	*****	*****	*****	*****	*****	1.2	1.2	1.1	1.0	0.8	0.5
70,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	1.1	1.0	1.0	0.7	0.4
80,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	1.0	0.9	0.7	0.4
90,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.8	0.7	0.4
100,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.8	0.6	0.4
125,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.6	0.3
150,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.3

NOTE: For correct usage of these tables, please refer to the microdata documentation.

Example 4: Estimates of Ratios

Suppose that a user estimates that 23,656 candidates who underwent a formal assessment as part of their participation in a staffing activity received an offer of appointment (D_Q01 = 1 and E_Q06 = 1). The user also estimates that 2,383 candidates who did not undergo a formal assessment also received an offer of appointment (D_Q01 = 2 and E_Q06 = 1). The user is interested in comparing the estimate of candidates who underwent a formal assessment versus those who did not undergo a formal assessment in the form of a ratio. How does the user determine the coefficient of variation of this estimate?

- 1) First of all, this estimate is a ratio estimate, where the numerator of the estimate (\hat{X}_1) is the number of candidates who underwent a formal assessment. The denominator of the estimate (\hat{X}_2) is the number of candidates who did not undergo a formal assessment.
- 2) Refer to the approximate coefficient of variation table for the Target Population (All departments which fell under the Public Service Employment Act and had at least 350 employees or more).
- 3) The numerator of this ratio estimate is 23,656. The figure closest to it is 20,000. The coefficient of variation for this estimate is found by referring to the first non-asterisk entry on that row, namely, 2.4%.
- 4) The denominator of this ratio estimate is 2,383. The figure closest to it is 2,500. The coefficient of variation for this estimate is found by referring to the first non-asterisk entry on that row, namely, 7.1%.

SURVEY OF STAFFING - CANDIDATES, 2009 CYCLE 1 - SHARE FILE														
Approximate Sampling Variability Tables														
Target population (All departments which fell under the Public Service Employment Act and had at least 350 employees or more)														
NUMERATOR OF PERCENTAGE	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
10	112.9	112.4	111.8	110.1	107.1	104.1	101.0	97.8	94.5	91.0	87.5	79.9	61.9	35.7
20	79.8	79.5	79.1	77.8	75.8	73.6	71.4	69.2	66.8	64.4	61.9	56.5	43.7	25.3
30	65.2	64.9	64.5	63.6	61.9	60.1	58.3	56.5	54.6	52.6	50.5	46.1	35.7	20.6
.														
.														
.														
1,500	*****	9.2	9.1	9.0	8.7	8.5	8.2	8.0	7.7	7.4	7.1	6.5	5.1	2.9
2,000	*****	7.9	7.9	7.8	7.6	7.4	7.1	6.9	6.7	6.4	6.2	5.6	4.4	2.5
2,500	*****	*****	7.1	7.0	6.8	6.6	6.4	6.2	6.0	5.8	5.5	5.1	3.9	2.3
3,000	*****	*****	6.5	6.4	6.2	6.0	5.8	5.6	5.5	5.3	5.1	4.6	3.6	2.1
3,500	*****	*****	6.0	5.9	5.7	5.6	5.4	5.2	5.1	4.9	4.7	4.3	3.3	1.9
4,000	*****	*****	5.6	5.5	5.4	5.2	5.1	4.9	4.7	4.6	4.4	4.0	3.1	1.8
4,500	*****	*****	*****	5.2	5.1	4.9	4.8	4.6	4.5	4.3	4.1	3.8	2.9	1.7
5,000	*****	*****	*****	4.9	4.8	4.7	4.5	4.4	4.2	4.1	3.9	3.6	2.8	1.6
7,500	*****	*****	*****	4.0	3.9	3.8	3.7	3.6	3.5	3.3	3.2	2.9	2.3	1.3
10,000	*****	*****	*****	3.5	3.4	3.3	3.2	3.1	3.0	2.9	2.8	2.5	2.0	1.1
15,000	*****	*****	*****	*****	2.8	2.7	2.6	2.5	2.4	2.4	2.3	2.1	1.6	0.9
20,000	*****	*****	*****	*****	2.4	2.3	2.3	2.2	2.1	2.0	2.0	1.8	1.4	0.8
30,000	*****	*****	*****	*****	*****	1.9	1.8	1.8	1.7	1.7	1.6	1.5	1.1	0.7
.														
.														
.														
125,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.6	0.3
150,000	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	0.3

NOTE: For correct usage of these tables, please refer to the microdata documentation.

- 5) So the approximate coefficient of variation of the ratio estimate is given by Rule 4, which is:

$$\alpha_{\hat{R}} = \sqrt{\alpha_1^2 + \alpha_2^2}$$

where α_1 and α_2 are the coefficients of variation of \hat{X}_1 and \hat{X}_2 respectively. That is:

$$\begin{aligned} \alpha_{\hat{R}} &= \sqrt{(0.024)^2 + (0.071)^2} \\ &= \sqrt{0.000576 + 0.005041} \\ &= 0.075 \end{aligned}$$

- 6) The obtained ratio of candidates who underwent a formal assessment as part of their participation in a staffing activity and received an offer of appointment versus candidates who did not undergo a formal assessment and also received an offer of appointment is $23,656 / 2,383 = 9.9$ (to be rounded according to the rounding guidelines in Section 9.1). The coefficient of variation of this estimate is 7.5% which makes the estimate releasable with no qualifications.

Example 5: Estimates of Differences of Ratios

Suppose that the user estimates that the ratio of candidates who underwent a formal assessment as part of their participation in a staffing activity and received an offer of appointment to candidates who did not undergo a formal assessment and received an offer of appointment is $3,517 / 277 = 12.7$ for National Defence, while it is $976 / 94 = 10.4$ for Fisheries and Oceans Canada. The user is interested in comparing the two ratios to see if there is a statistical difference between them. How does the user determine the coefficient of variation of the difference?

- 1) First calculate the approximate coefficient of variation for the National Defence ratio (\hat{R}_1) and the approximate coefficient of variation for the Fisheries and Oceans Canada ratio (\hat{R}_2) as in Example 4. The approximate CV for the National Defence ratio is 35.4%. The approximate CV for the Fisheries and Oceans Canada ratio is 46.7%.
- 2) Using Rule 3, the standard error of a difference ($\hat{d} = \hat{R}_1 - \hat{R}_2$) is:

$$\sigma_{\hat{d}} = \sqrt{(\hat{R}_1\alpha_1)^2 + (\hat{R}_2\alpha_2)^2}$$

where α_1 and α_2 are the coefficients of variation of \hat{R}_1 and \hat{R}_2 respectively.

That is, the standard error of the difference $\hat{d} = 12.7 - 10.4 = 2.3$ is:

$$\begin{aligned}\sigma_{\hat{d}} &= \sqrt{[(12.7)(0.354)]^2 + [(10.4)(0.467)]^2} \\ &= \sqrt{(20.2122) + (23.5885)} \\ &= 6.618\end{aligned}$$

- 3) The coefficient of variation of \hat{d} is given by $\sigma_{\hat{d}} / \hat{d} = 6.618 / (2.3) = 2.877$.
- 4) So the approximate coefficient of variation of the difference between the estimated ratios is 287.7%. The difference between the estimated ratios is considered unacceptable and Statistics Canada recommends this estimate not be released. However, should the user choose to do so, the estimate should be flagged with the letter F (or some similar identifier) and be accompanied by a warning to caution subsequent users about the high levels of error associated with the estimate.

10.2 How to Use the Coefficient of Variation Tables 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, 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 difference 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.

Confidence intervals for an estimate can be calculated directly from the Approximate Sampling Variability Tables by first determining from the appropriate table the coefficient of variation of the estimate \hat{X} , and then using the following formula to 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 of \hat{X} , and

- $t = 1$ if a 68% confidence interval is desired;
- $t = 1.6$ if a 90% confidence interval is desired;
- $t = 2$ if a 95% confidence interval is desired;
- $t = 2.6$ if a 99% confidence interval is desired.

Note: Release guidelines which apply to the estimate also apply to the confidence interval. For example, if the estimate is not releasable, then the confidence interval is not releasable either.

10.2.1 Examples of Using the Coefficient of Variation Tables to Obtain Confidence Limits

A 95% confidence interval for the estimated proportion of candidates who underwent a formal assessment as part of their participation in a staffing activity and received an offer of appointment (from Example 2, Section 10.1.1) would be calculated as follows:

$$\begin{aligned} \hat{X} &= 55.8\% \text{ (or expressed as a proportion 0.558)} \\ t &= 2 \\ \alpha_{\hat{x}} &= 1.8\% \text{ (0.018 expressed as a proportion) is the coefficient of variation} \\ &\text{of this estimate as determined from the tables} \\ CI_{\hat{x}} &= \{0.558 - (2)(0.558)(0.018), 0.558 + (2)(0.558)(0.018)\} \\ CI_{\hat{x}} &= \{0.558 - 0.020, 0.558 + 0.020\} \\ CI_{\hat{x}} &= \{0.538, 0.578\} \end{aligned}$$

With 95% confidence it can be said that between 53.8% and 57.8% of candidates who underwent a formal assessment as part of their participation in a staffing activity received an offer of appointment.

10.3 How to Use the Coefficient of Variation Tables to Do a 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. Let the standard error on the difference $\hat{X}_1 - \hat{X}_2$ be σ_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.

10.3.1 Examples of Using the Coefficient of Variation Tables to Do a T-test

Let us suppose that the user wishes to test, at 5% level of significance, the hypothesis that there is no difference between the proportion of candidates who underwent a formal assessment as part of their participation in a staffing activity and received an offer of appointment (55.8%) and the proportion of candidates who did not undergo a formal assessment and also received an offer of appointment (34.4%). In Example 3, Section 10.1.1, the standard error of the difference between these two estimates was found to be 0.022. Hence,

$$t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_{\hat{d}}} = \frac{0.558 - 0.344}{0.022} = \frac{0.214}{0.022} = 9.73$$

Since $t = 9.73$ is greater than 2, it must be concluded that there is a significant difference between the two estimates at the 0.05 level of significance.

10.4 Coefficients of Variation 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 SOS 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, to obtain an estimate of the total number of consecutive months candidates were in an acting position when the staffing activity concluded for them, multiply the value reported in variable C_Q08 (number of months) by the final weight for the record, then sum this value over all records with C_Q07 = 1 (candidates in an acting position).

For example, the coefficient of variation of the total number of consecutive months candidates were in an acting position when the staffing activity concluded for them would be greater than the coefficient of variation of the corresponding number of candidates who were in an acting position. Hence, if the coefficient of variation of the number of candidates who were in an acting position is unacceptable (making this estimate not releasable), then the coefficient of variation of the corresponding quantitative estimate of the total number of months candidates were in an acting position will also be unacceptable (making the quantitative estimate not releasable).

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 formula to be used in these calculations.

10.5 Coefficient of Variation Tables

Refer to SOS2009C1_CVTabE.pdf for the coefficient of variation tables.

11.0 Weighting

The Survey of Staffing (SOS) is a probability survey. As is the case with any probability survey the sample is selected to represent a reference 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.

SOS weighting strategy overview

Weighting steps
1. Initial design weight
2. Non-response adjustment
3. Non-sharing adjustment
4. Post-stratification

11.1 Initial Design Weight

At the time of selection, an initial design weight was assigned to each person, as the inverse of its probability of selection. Since the SOS design is stratified with simple random sampling within strata, the probability of selection of the employee i in stratum h is:

$$\pi_{ih}^{initial} = \frac{n_h}{N_h}$$

where, n_h and N_h denote respectively the sample and population size of stratum h . The initial design weight ($w1_i$) is then given by:

$$w1_{ih} = \frac{N_h}{n_h}$$

11.2 Non-response Adjustment

It was observed that non-response did not occur randomly or uniformly within the population since different response rates were obtained for different sub-populations. The use of an appropriate technique is required to correct non-response bias that may be introduced. The chosen technique for the Survey of Staffing was based on response homogeneous groups (RHG). RHGs were developed with the premise of identifying sample units with similar response probabilities. In other words, it is assumed that persons pertaining to a given RHG are equally likely to respond to the survey in a similar fashion. Many factors, among them gender and age are traditionally known to be factors associated with different non-response patterns. Analyses were completed and the RHGs were identified. The implementation, i.e. the calculation of the weight adjustment, was carried out using Statistics Canada's StatMx software. This approach also ensures the use of the proper variance formula.

For employee i in RHG j the response probability is calculated as:

$$\pi_{ij}^{response} = \frac{\text{number of responding units}}{\text{number of sample units}}$$

and the non-response adjustment factor is given by the inverse of the response probability. To obtain the weight for person i after the non-response adjustment, we multiply $w1_i$ by the non-response adjustment factor:

$$w2_i = w1_i \times (\pi_{ij}^{response})^{-1}.$$

11.3 Adjustment for Non-sharing Units

People who refused to share their data cannot be included in the share file. An additional adjustment has to be made to the weights of the employees who agreed to share their data in order to compensate for the employees who refused to share. A process similar to the non-response adjustment procedure was used in order to estimate the probability of sharing and to calculate the non-sharing adjustment, in each Sharing Homogeneous Group.

For employee i in Sharing Homogeneous Group j the sharing probability is calculated as:

$$\pi_{ij}^{share} = \frac{\text{number of responding employees who agreed to share their data}}{\text{number of responding employees}}$$

and the non-sharing adjustment factor is given by the inverse of the sharing probability. To obtain the weight for person i after the non-sharing adjustment, we multiply $w2_i$ by the non-sharing adjustment factor:

$$w3_i = w2_i \times (\pi_{ij}^{share})^{-1}.$$

11.4 Post-stratification Adjustment and Final Weight

Post-stratification is one of the calibration estimation techniques widely used in social surveys. It allows benchmarking on new updated population counts. Note that the post-stratification file still represents the target population. The post-stratification adjustment is calculated at the post-stratum level (department level) using the following formula:

$$\frac{\text{total number of persons in a given post – stratum}}{\text{estimated total number of persons in a given post – stratum}}$$

The final weight consists of cascading the design weight. The non-response adjustment, the non-sharing adjustment and the post-stratification adjustment are used to calculate the final weight.

$$W_{final} = \text{design weight} * \text{non – response adjustment} * \text{non – sharing adjustment} * \text{post – stratification adjustment}$$

12.0 Questionnaire

The Survey of Staffing – Candidates questionnaire (SOS) was used in January and February 2009 to collect information for the survey. The file SOS2009C1_QuestE.pdf contains the English questionnaire.

13.0 Record Layout with Univariate Frequencies

See SOS2009C1_CdBk.pdf for the record layout with univariate counts.