
International Travel Survey
U.S. AND OVERSEAS VISITORS TO CANADA
MICRODATA USER GUIDE

2016



How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to:

Statistics Canada
150 Tunney's Pasture Driveway
Ottawa, Ontario K1A 0T6
Telephone: 613-951-8116 or call toll-free 1-800-263-1136
Fax: 613-951-0581 or 1-877-287-4369
Email :infostats@statcan.gc.ca

Accessing and ordering information

The International Travel Survey (ITS) produces two types of microdata files - master files, and public use microdata files (PUMF).

Master files

The master files contain all variables and all records from the survey collected during a collection period. These files are accessible at Statistics Canada for internal use and in Statistics Canada's Research Data Centres (RDC), and are also subject to custom tabulation requests.

Research Data Centre

The RDC Program enables researchers to use the survey data in the master files in a secure environment in several universities across Canada. Researchers must submit research proposals that, once approved, give them access to the RDC. For more information, please consult the following web page:

<http://www.statcan.gc.ca/eng/rdc/index>

Custom tabulations

Another way to access the master files, is to offer all users the option of having staff in Client Services of the ITS section prepare custom tabulations. This service is offered on a cost-recovery basis. It allows users who do not possess knowledge of tabulation software products to get custom results. The results are screened for confidentiality and reliability concerns before release. For more information, please contact Statistics Canada.

Public use microdata files (PUMF)

The public use microdata files are developed from the master files using a technique that balances the need to ensure respondent confidentiality with the need to produce the most useful data possible. The PUMF must meet stringent security and confidentiality standards required by the *Statistics Act* before they are released for public access. To ensure that these standards have been achieved, each PUMF goes through a formal review and approval process by an executive committee of Statistics Canada. Variables most likely to lead to identification of an individual are deleted from the data file or are collapsed to broader categories.

Note that there may be differences between the estimates published by Statistics Canada in its statistical products and the estimates reproduced by users with the public use microdata files, as Statistics Canada uses master files to produce its estimates. To obtain a copy of the PUMF contact Statistics Canada.

The Data Liberation Initiative

The Data Liberation Initiative (DLI) Program enables students and researchers to use the public use microdata files in several universities across Canada. For more information, please consult the following web page:

<http://www.statcan.gc.ca/eng/dli/dli>

Table of Contents

1.0 BACKGROUND	5
2.0 DATA COLLECTION METHODS	5
2.1 <i>Electronic questionnaire</i>	6
2.2 <i>Air Exit Survey of Visitors to Canada</i>	6
3.0 CONCEPTS AND DEFINITIONS	7
3.1 <i>Main Definitions</i>	7
4.0 SURVEY PROCESSING	13
4.1 <i>Editing</i>	13
4.2 <i>Imputation</i>	14
4.3 <i>Disclosure control</i>	15
5.0 WEIGHTING	16
5.1 <i>Port Factor Groups (PFGs)</i>	16
5.2 <i>Weighting procedures</i>	17
6.0 SURVEY ESTIMATION	18
7.0 SAMPLING VARIABILITY	19
7.1 <i>Coefficient of Variation for quality evaluation</i>	19
7.2 <i>Bootstrap method of variance estimation</i>	20
8.0 Public Use Microdata File (PUMF) Content	20

1.0 BACKGROUND

The International Travel Survey (ITS) is an ongoing survey conducted by Statistics Canada since 1920 to meet the requirements of the Balance of Payments of the Canadian System of National Accounts (BOP). Through the years, questions directed at collecting detailed characteristics of travellers for market research and industry planning were gradually incorporated in the survey. Today, the ITS provides a full range of statistics on the volume of international travellers and detailed characteristics of their trips such as expenditures, activities, places visited and length of stay.

The ITS covers both Canadian residents returning from trips outside Canada and international visitors to Canada.

In addition to fulfilling BOP requirements, the ITS is also being used by the Tourism Satellite Account (TSA), Canada Border Services Agency (CBSA), Destination Canada, provincial tourism agencies, the United States Department of Commerce and a number of private sector industries. The ITS is also used for reporting to international organizations such as the World Tourism Organization (WTO), the Organization for Economic Co-operation and Development (OECD) and the Pacific-Asia Tourism Association (PATA).

This manual has been produced to facilitate the use of the PUMF of the survey results.

2.0 DATA COLLECTION METHODS

The ITS consists of two main components: the frontier counts and the questionnaires. Both these components depend greatly on the co-operation of CBSA in the collection of the number of border crossings and the distribution of the invitation cards for the electronic questionnaires.

The frontier counts are used to provide estimates on the volume of trips. Information is collected at CBSA ports of entry into Canada on the travel party size and mode of entry using either custom declaration cards (E-311) completed by the travellers, the E62 and E63 tally sheets, or through the Integrated Primary Inspection Lane (IPIL) completed by CBSA officers. When processing the E-311 cards, a sampling of cards are used to estimate the total population of travellers, which then comprise the frontier counts.

The questionnaires are used to secure quarterly information on the characteristics of international travellers and trips. These details include purpose of trip, size of travelling party, places visited, activities participated in during the trip, length of trip and trip spending. These components of the survey are used to update the Canadian Balance of International Payments. In addition, the federal and provincial governments, the

ourism industry, businesses and the general public use quarterly and annual estimates of international trip and traveller characteristics.

The questionnaires are obtained via two methods: the electronic questionnaire surveys submitted on-line by travellers who received an invitation card and the Air Exit Survey of Visitors to Canada completed on paper questionnaire by Statistics Canada interviewers at select Canadian airports.

2.1 Electronic questionnaire

This is a sample survey where invitation cards are distributed only to part of the international traveller population. Canada Border Services officials hand out a questionnaire or an invitation card to travel parties on entry or re-entry to Canada. The cards are distributed to the following:

- Residents of United States or overseas countries visiting Canada;
- Canadian residents returning from trips outside Canada;

The recipients of the invitation card are asked to fill out the electronic questionnaire on-line after the trip is complete, using the web address provided on the card.

A stint distribution system has been developed to survey international travellers based on the previous year's traffic. A stint consists of a selected period of several days during which invitation cards are distributed to all eligible travellers as defined above. Each port of entry involved in this scheme receives, for each of its stints, a specific quantity of cards and a date to start the distribution. On the start date, the officers hand out the cards on a continuous basis to the appropriate travelling population until all cards have been distributed. The number of invitation cards distributed in 2016 was about 400,000 for all Canadian and overseas travellers.

2.2 Air Exit Survey of Visitors to Canada

Additional questionnaires are also obtained from US and overseas travellers returning directly to their country of origin by commercial air (the Air Exit Survey (AES)).

Countries from which we attract the most visitors are targeted by the AES. These countries include the United Kingdom, France, Germany and Australia as well as a number of emerging markets, such as China, Japan, India and South Korea. Statistics Canada interviewers conduct personal interviews with overseas travellers while they await their return flights to these targeted overseas countries. These interviews are conducted each month at the international airports in five cities (Halifax, Montreal, Toronto, Calgary and Vancouver) during a collection period lasting 5 to 7 days. In 2016, about 7,500 interviews were completed. The interviewing team consists of interviewers of different language skills, enabling interviews to be

conducted in the native language of the travellers where possible. The paper version of the questionnaire is available in 10 languages.

United States air commercial travelers complete the AES at the international airports in Toronto, Vancouver, Montreal, Ottawa, Calgary, Edmonton and Halifax. In 2016, about 4,000 interviews were conducted.

3.0 CONCEPTS AND DEFINITIONS

3.1 Main Definitions

3.1.1 Type of trips

Trip

A trip begins when an overseas traveller is cleared through a CBSA point of entry and enters Canada. The frontier counts collect data on the number of trips made and the ITS questionnaires are distributed to collect information about travellers from a given time period. A traveller may be included in the frontier counts more than once in the same time period if they enter Canada more than once through a CBSA point of entry during this period.

Trip duration

A trip starts when a person crosses the border to enter Canada and ends when he/she exits the country. The trip duration is measured by the number of nights a non-resident traveller spends in Canada. Trip durations are also divided into two mutually exclusive groups: same-day trips and overnight trips.

Same-day trip

Same-day travel encompasses travellers who enter and leave Canada in less than 24 hours. This category includes persons resident in another country and commuting daily to and from work in Canada.

Overnight trip

Overnight travel encompasses overseas travellers who stay in Canada for at least 24 hours before exiting the country. This classification closely approximates the term «tourist» although it includes temporary and seasonal workers, students, and in transit travellers.

Purpose of trip

The main reason for the trip to a country is generally summarized as follows:

- **Business.** Includes attending a meeting, convention, conference, trade show or seminar, and other work.
- **Visiting friends or relatives.**
- **Pleasure.** Includes a holiday or vacation, visiting a second home, cottage or condo, and attending events and attractions.
- **Other purposes.** Includes personal, in transit, shopping, educational study and other.

3.1.2 Type of traveller

International traveller

The term international traveller applies to all persons arriving in Canada who are cleared through CBSA points of entry. The method of data collection from the frontier counts serves to count each entry made into Canada. These frontier counts represent both the number of entries and re-entries into the country, therefore including the counts for travellers who have crossed the border more than once during the period of reference. The term international traveller is divided into three mutually exclusive groups: non-resident traveller, resident traveller and other traveller. Only non resident travellers are of interest to this document. For documentation concerning resident travellers, refer to the ITS Microdata User Guide for Canadian Resident Trips Abroad.

Non-resident Traveller

A non-resident traveller is an international traveller resident of a country other than Canada who is cleared through CBSA points of entry on a visit for a period of less than 8 months. Canadian citizens residing in other countries for more than 12 months coming to Canada on leave are included as a non-resident traveller.

Non-resident travellers are further distinguished by whether they are resident to the United States (United States traveller) or are resident to a country other than the United States (Overseas traveller).

Resident Traveller

A resident traveller is an international traveller residing in Canada who has travelled outside of Canada for a period of less than 12 months for the same reasons as given above and who is returning to Canada through CBSA. Foreign citizens who are residing in Canada, travelling abroad on leave or for other reasons are considered as a resident traveller.

Other Traveller

Other travellers include the following types of international travellers:

- **Immigrants.** Citizens of other countries entering Canada to take up permanent residence for more than 12 months.
- **Former residents.** Canadian citizens returning to Canada to re-establish permanent residence after residing outside Canada for more than 12 months.
- **Military personnel, diplomats and dependents.** Canadian citizens and dependents who have been stationed abroad for more than 12 months and who are returning from postings. Foreign citizens and dependents entering Canada on military or diplomatic postings.
- **Crews.** Persons engaged in the operation of a transport: air, sea, truck, train, etc.

3.1.3 Types of entry into Canada

Non-resident travellers are grouped by the type of entry used to enter Canada. There are three categories: direct to Canada, via the United States by land, and via the United States by other methods. Two examples follow: a non-resident traveller visiting the United States who enters Canada during his trip would be counted as entering via the United States; a non-resident entering direct to Canada who visits the United States during his trip would be counted direct on his first entry and via the United States on his second entry.

Direct to Canada

Non-residents entering directly to Canada from a country other than the United States are considered direct to Canada, regardless of the mode of entry to Canada. The number of countries visited on the trip, or while on board of the mode of transport to enter Canada, does not affect this category as long as the United States is not among the countries visited.

Via the United States - by land

Non-resident travellers entering through or by way of the United States are considered via the United States. The length of stay in the United States has no bearing on the classification; a person could be in transit through the United States or could have stayed several days. Travel by land includes automobile, bus and train as well as cycle (bicycle, motorcycle) and pedestrian. Also included is traffic entering on ferries either by car or foot.

Via the United States - by other methods

Non-resident travellers entering through or by way of the United States are considered via the United States. The length of stay in the United States has no bearing on the classification. This category includes non-resident travellers entering Canada by a mode of transportation other than land. These

other modes of transportation include persons travelling by commercial or private mode of transportation. These modes are either by air or sea.

Mode of entry

Frontier counts are collected according to the mode of transportation used to enter Canada. The possible modes collected are: commercial plane, private plane, commercial boat (including cruise ship and ferry), private boat, automobile, bus, and train. The questionnaire includes several additional modes. The modes included in the questionnaire are: plane (commercial and private), private boat, cruise ship, ferry, rented automobile, private automobile, bus, and train. Not all modes of transportation are possible for all CBSA points of entry. For example, an airport will only count entries by plane (commercial or private). See Appendix A for more information on the Canadian ports and their valid modes of entry. The mode of entry is used for weighting responses.

3.1.4 Travel party

A travel party refers to persons from the same country who are travelling together (whether they are from the same household or not) and who are willing to collectively report their trip spending and activities.

Party composition

The travel party composition refers to the number of adults and/or children who are part of the same travel party.

Origin or Residence

The origin or residence is the place of last permanent residence. This is normally summarized in the ITS by census region for United States residents and by country for other residents.

Person-trip

A person-trip begins each time a non-resident traveller enters Canada as recorded by Canada Border Services Agency (CBSA) officials. The person-trip concludes each time the non-resident traveller exits Canada. A person-trip differs from a trip in that a person-trip weighted by the travel party size, while a trip is taken by the whole travel party.

Person-trip-night

A person-trip-night is defined for each night a non-resident traveller spends within Canada during a person-trip.

3.1.5 Visits during a trip

Visit

A non-resident traveller may stay in several locations before leaving Canada. Each stay at a Canadian location within a given trip represents a place visited. Place visited are collected at the census division level, and up to 10 places visited can be recorded in the ITS questionnaire. Estimates of visits are derived solely from the questionnaire component. Note that for publishing purposes, a trip with multiple visits to the same geographical level, these will be counted as a single visit.

Visit duration

When reporting visit location, the respondent will declare the number of nights spent in each location visited. For publishing purposes, trip durations are divided into two mutually exclusive groups: same-day visits and overnight visits.

Same-day visit

A same-day visit is when the respondent have not declared any night for a location visited. On the course of a same-day trip, a traveller may visit several locations, however, only one location is captured on a same-day trip. During an overnight trip, every location visited where the respondent did not spend a night is recorded as a same-day visit.

Overnight visit

An overnight visit is when the respondent have declared one or more nights to the location visited.. Overnight visit can only be completed while on an overnight trip.

Person-visit

A person-visit is either a same-day or overnight visit made by a person taking a trip. A person-visit differs from a visit in that a person-visit is weighted by the travel party size, while a visit is taken by the whole travel party. The total number of person-visits is the total number of visits made by each person, in each travel party, on each trip taken.

Person-visit-night

A person-visit-night is considered to be one night spent in the visited location by a person taking a trip. For example, a person who spends three nights in the visited location has a count of three person-visit-nights in that location, as do three people who spend one night in the visited.

3.1.6 Trip components

Mode of transportation

The two modes of transportation used to *enter* and leave Canada is captured by the questionnaire as well as any commercial transportation used while in Canada. This information may be used when trip expenditures require imputation.

Accommodation

The accommodation used in each location visited is collected in the questionnaire. Coded responses include: Hotel, Motel, Home of friends or relatives, Camp or trailer park, Cottage or cabin, and Other.

Activities

Travellers report all activities in which they have taken part during the trip according to provided categories. All categories are contained within two broader groupings: participating in sports or outdoor activities, or other activities. No information is collected on activities during specific visits. For this reason, an activity cannot be associated with a precise location. For example, a person may visit Vancouver and Whistler during the same trip and declare they have done some downhill skiing. When compiling data, the «downhill skiing» activity will be attributed to both Vancouver and Whistler since activity information is only collected at the trip level. However, this activity may have occurred only in Vancouver or only in Whistler.

Activities are not collected from American travellers on a same-day trip to Canada by automobile.

Expenditures

Trip expenditures made by all members of the travelling party are collected in the questionnaire. Expenditures may have been paid for by individuals, by government or by a private sector business. The expenditures are broken down into the following categories: accommodation, transportation in Canada (include gasoline expenditures, rented car, intercity plane, bus and train fares, boat tours, local bus, taxi), food and beverage, recreation and entertainment, clothes and gifts, and other. Expenditures are reported in Canadian dollars or are converted to Canadian dollars.

For the non-resident travellers, the expenditures do not represent where money was spent on a specific visit, but rather how much money was spent, and in which category of expenditures for the trip as a whole.

Fares

For non-resident travellers, fares are the amounts spent by all members of the travelling party on commercial transportation (plane, bus, train or boat) to enter and exit Canada. Fares are reported in Canadian dollars or are converted to Canadian dollars.

Census Coding

For non-resident travellers, the Canadian locations visited are linked to the census information for that area.

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

CMAs are defined as the main labour market area of an urbanized core having 100,000 or more population. Smaller urban areas, with an urban core with populations of at least 10,000 are defined as CAs. For more details, refer to the Standard Geographical Classification SGC 2011.

4.0 SURVEY PROCESSING

Processing stages occur after obtaining survey responses and before estimation. Survey processing serves to make the data suitable for tabulation and producing estimates through data editing, coding, and imputation. All reported data on mail-back, electronic and AES questionnaires are captured, coded and verified.

4.1 Editing

Data is not always reported in a survey such that it can be used directly in further analysis. Editing reformats data such that it can be used in tabulation, analysis, and estimation.

4.1.1 Validity, and consistency

Some errors occur in the data capture process such as an incorrectly recorded field, or missing data on the questionnaire. These errors must be located and corrected. Edit rules are designed to ensure the data is consistent with known characteristics, and that questionnaire entries are valid responses given the question.

Some examples of edits made during the data capture process follow:

- Some ports of entry can only permit certain modes of entry. If an automobile is recorded in a response as a traveller's mode of entry into an airport, then the data entry is corrected. The instance is investigated and either the port is changed, or the mode of entry is changed accordingly.

-
- If all transportation costs are included in a package cost, but are also included in the entries for fare to enter Canada and fare to leave Canada, then the entries are corrected to remove the duplication of reported fares.

4.1.2 Derived variables creation

Expenses reported on package tours must be broken down into spending components, reported data for trips that involved multiple entries into Canada must be separated, and derived variables must be calculated.

The ITS asks respondents for information on any packages or inclusive trips purchased. This purchase is then not reported in the expenditure section on spending while in Canada. Consequently, packages are separated according to the components covered within the collective package cost, and added to the relevant expenditure categories.

From the ITS definition of a trip, any exit from Canada marks the end of a trip. Respondents might report several trips under this definition within their completed questionnaire. These trips must be separated to allow for consistent analysis.

Finally, all variables that are required in further analysis are created and calculated.

4.2 Imputation

A certain amount of data imputation is performed in the ITS whenever fields are missing within a questionnaire, or when an insufficient number of questionnaires are received to meet the ITS objectives.

Certain information must be completed by a respondent in order to impute any missing information. Similarly, estimates require certain information to be completed by the respondent in order to be reliable. Therefore, a questionnaire must have the following information to be kept: port of entry, mode of entry, entry route, date of entry, date of exit, reason for trip, travel party size, place visited, and whether a package was purchased along with the corresponding nights included in the package. These completed fields enable a questionnaire to be linked to the Port Factor Groups (PFGs) to impute missing values. For example, missing transportation fares and/or total trip spending are imputed using regression models based on the travel party size, trip duration, purpose of trip, ages of travel party members, accommodation information, port of entry.

Total imputation (i.e. imputation of entire questionnaires) is performed for any Port Factor Group (PFG) for which we have received an insufficient number of questionnaires for a given quarter. Estimates are generated by PFG, and thus a minimum number of questionnaires must be used to provide quality estimates for the PFG. In instances where the minimum number is not met by received questionnaires from the current quarter, a sufficient number of questionnaires, selected amongst questionnaires from the same

quarter of the five previous years belonging to the PFG are added to the current year's sample of that PFG for the reference quarter. For 2016, rates of total imputation are as follows: Canadian residents 97.6%, United States residents 83.1% and overseas residents 12.4%.

Total imputation is also carried out for several PFGs¹ or strata that are outside the scope of the questionnaire distribution. These PFGs refer to modes of entry that are not targeted by the questionnaire distribution (train, private plane or boat, motorcycle, bicycle, foot, etc.) or to ports of entry that do not distribute questionnaires. For these out-of scope PFGs, the characteristics of travellers are estimated using total imputation. These imputed questionnaires are duplicates of questionnaires that were obtained in 1990 for the same quarter and same PFGs. When total imputation has been used in this manner, the quality of the subsequent estimates may be compromised due to limitations in available questionnaires that can be brought forward to the current quarter's sample. For example, if there are very few questionnaires available in past quarters, the current quarter's estimates may be of poor quality.

4.3 Disclosure control

All survey recipients are assured that their identities will remain confidential and that no identifying information will be disclosed in the PUMF. This is a requirement by law that all Statistics Canada data adheres to. Consequently, any information that may divulge a respondent's identity without prior knowledge or written consent from the respondent remains confidential and cannot be released or published. Some data are therefore suppressed or removed entirely from the PUMF to prevent any direct or residual identification of respondents from the released data.

Subsequently, any variable that can relate a response to an individual has been removed. This includes variables with options selected by a small number of respondents, such as having only a few respondents identify as having participated in a given activity, or providing identifiable information such as visits made during a trip, which can be traced back to a respondent. Programs are run to ensure the confidentiality of all released data. Any compromising variables or combinations of variables are removed from the microdata file.

¹ In the International Travel Survey, the target population (international travellers) are partitioned into Port Factor Groups, based on selected traveller characteristics, such as country of residence, mode of entry and duration of stay.

5.0 WEIGHTING

5.1 Port Factor Groups (PFGs)

Records from United States residents are assigned weights based on post-stratification into particular PFGs. These PFGs are combinations of durations of trip, modes of entry to Canada and ports of entry that partition the United States travelling population. Due to low volumes of traffic, many ports are combined into one PFG; however, a PFG is always defined within a province. Records are weighted so that the weighted estimate of the number of travellers derived from the sample matches the frontier count for each PFG.

Factor Groups (PFGRN) for United States residents are based solely on the mode of entry used to enter Canada.

For more details, please refer to the « United States Port Factor Groups » section in Appendix A.

<u>Mode</u>	<u>No. Of Factor Groups</u>	<u>Factor Group Codes</u>
Automobile, same day	23	001 - 032
Automobile, overnight	29	101 - 190
Plane commercial	14	201 - 217
Plane private, same day	11	240 - 250
Plane private, overnight	11	251 - 261
Bus, same day	9	300 - 308
Bus, overnight	9	310 - 318
Train	8	400 - 407
Boat commercial	9	500 - 509
Boat private, same day	11	530 - 540
Boat private, overnight	11	541 - 551
Other methods	9	600 - 608
Total	154	

As with the United States flow, overseas travellers are stratified into PFGs. Each traveller is coded by country of residence (or, in some instances, combinations of countries) to obtain a country group code (SGR). The SGR is then cross-classified with the type of entry to obtain the PFG.

Factor Groups (PFGRN) for all non-resident travellers who are Overseas Residents are based on:

- country grouping; and
- type of entry into Canada (Direct, Via the United States by land, or Via the United States by other)

For more details, please refer to the « Overseas Factor Groups » section in Appendix E.

<u>Mode</u>	<u>No. Of Factor Groups</u>
Directly from overseas	27
Via the United States by land	27
Via the United States by other	27
Total	81

5.2 Weighting procedures

The initial ITS weighting procedure is based on Person Factor Weight (WEIGHT), where:

$$\text{WEIGHT} = \frac{\text{Volume counts (frontier counts) of persons in the factor group}}{\text{Number of sampled persons in the factor group}}$$

All information collected on the overseas and United States resident questionnaires refer specifically to the travelling party. The person weight is used to calculate all estimates.

5.2.1 United States residents

Once the initial weights are assigned based on the above weight equation, a second stage involving two calibrations are applied to the weights of certain records from United States residents. These calibrations are derived from the distributions seen in the frontier counts for some of the large factor groups.

The first calibration is performed on the initial weights of the records of the United States air travellers who came through the international airports of Vancouver, Toronto or Montreal. This calibration is based on information obtained from an independent administrative source regarding the purpose and duration of trip as obtained from the custom declaration cards (E-311) completed by these travelers. The calibration is therefore intended to better reflect the distribution of these two characteristics shown by the E311 sample.

The second calibration is performed on the initial weights of the records of the United States car travellers who entered through one of the 32 largest land ports equipped with the Integrated Primary Inspection Line (IPIL) system. This second calibration is referred to as the state bias adjustment and is based on the vehicle state registration information obtained from the IPIL system. The adjustment is therefore designed to better reflect the distribution of travellers from the states that provide the largest volume of travelers as characterized by the IPIL sample in each of these ports.

5.2.2. Overseas residents

All overseas records are weighted to ensure that the derived weighted estimates of the number of travellers correspond to the frontier counts by SGR.

Once the initial weights are assigned based on PERSFAC, a second stage called the bias adjustment is applied to the initial weights of the questionnaires completed by residents from the United Kingdom, Germany, France, Japan, China, India, Australia and Brazil. The adjustment is based on information regarding the purpose and duration of trip as obtained from the E-311 cards completed by these travelers. This adjustment is therefore designed to better reflect the distribution of these two characteristics as shown by the E-311 cards.

A third stage called the regional weighting adjustment is then performed on the weights of all overseas records, which uses the optimal model statistical technique. This procedure ensures that the derived estimates of the number of travellers from the weighted records closely match the frontier counts at the region of entry level (by type of entry), while maintaining an exact match at the SGR level. In addition, for some pre-identified SGR/region of entry/type of entry combinations, the weights assigned to the records are adjusted so as to obtain an exact match to the corresponding frontier counts. Referred to as Direct Weighting Groups (DWGs), these SGR/region of entry/type of entry combinations represent portions of the overseas travel market with high volumes of travellers and for which we obtain a sufficient number of questionnaires. Also, the DWGs have been selected so that there are sufficient questionnaires left over in the non-selected combinations to ensure convergence of the raking-ratio technique.

6.0 SURVEY ESTIMATION

The strata of interest are divided by port (or for smaller ports, by a group of ports), mode of entry, and quarter. For estimation purposes, the responses obtained through the questionnaires must be treated as a simple random sample from the total traffic in each stratum. The data are subject to some degree of «distribution bias» due to the fact that the distribution of respondents may be different from the distribution of the population. For example, many travellers with a particular spending habit may respond to the survey, but if that spending habit isn't very common in the population, then there's distribution bias. The data are also subject to some «non-response bias» due to the fact that the individuals who choose not to reply may share a characteristic. This may make the responses to some degree unrepresentative of the travelling population.

Weighting techniques used in the estimation process attempt to reduce the effect of biases. Responses are disaggregated by known characteristics into homogeneous groups. Using the calculated weights, estimates can be obtained for a variety of trip and traveller characteristics. These estimates take the form of totals, averages and percentages, and can be obtained if a sufficient record sample exists for the desired grouping of traveller characteristics and geography.

7.0 SAMPLING VARIABILITY

The ITS is based on data gathered from only a part (sample) of the population. It is important to mention that results are only estimations of «true» values, where the «true» values for the travelling population could only be obtained through a census. Since differences always exist between sample and census data, statistical measures have been developed which indicate the expected size of the differences. Such measures allow analysts to assess the reliability of statistics produced from sample surveys.

Various factors contribute to the quality of an estimate such as distribution bias or non-response bias. The size of the sample producing an estimate is important to produce quality estimates; if the number of respondents with the characteristics of interest is too small, particularly relative to the census volume, then the estimate will be less reliable.

7.1 *Coefficient of Variation for quality evaluation*

Before releasing and/or making inferences and/or publishing any estimates from the International Travel Survey, 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. 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 below.

First, the number of records selected for the calculation of the estimate should be determined by rule thumb, 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 should be accompanied by a warning to caution subsequent users.

Quality level guidelines

Category 1: Acceptable

Coefficients of variation between 0.0% and 16.5%.

If an estimate's CV falls within this range, it is considered to be of acceptable quality. No warnings or restrictions will be associated with the estimate in the data.

Category 2: Marginal

Coefficients of variation between 16.6% and 33.3%.

If an estimate's CV falls within this range, it is considered to be of marginal quality. The estimate is released as it may be useful for some purposes, but is displayed with a warning regarding the accuracy of the estimate.

Category 3: Unacceptable

Coefficients of variation larger than 33.3%.

If an estimate's CV is this high, it is considered to be of unacceptable quality. The estimate should not be released as it does not meet the quality standards of Statistics Canada. Any conclusions made using an estimate of unacceptable quality is unreliable and likely invalid.

However, if a user insist upon the inclusion of estimates considered to be of unacceptable quality, the estimates are displayed with a disclaimer. The user is also required to always include the disclaimer if he disseminate, report, or present the data which use these estimates either directly or indirectly. Any estimate calculated from fewer than 30 records is also considered to be of unacceptable quality regardless of its corresponding CV.

7.2 Bootstrap method of variance estimation

Because of the complex nature of the ITS design and processing, simple variance calculations using standard methods tend to under-estimate the variance, making the estimates appear more precise than they are. For the ITS, Statistics Canada uses the bootstrap method of variance estimation. Under this method, the sampling units are sub-sampled and re-weighted many times over (the ITS currently uses 500 bootstrap sub-samples). From these sub-samples an estimated variance is obtained to derive the CV.

8.0 Public Use Microdata File (PUMF) Content

The ITS PUMF contains information on the travel habits of United States and overseas residents travelling to Canada. This information is collected primarily from the electronic questionnaires and the Air Exit Survey. Here are the different characteristics available on the United States and Overseas PUMF:

Trip characteristics:

- Quarter of entry
- Province of entry
- Mode of transportation used to enter and exit Canada and while in Canada
- Purpose of trip
- Activities
- International transportation costs (fares paid to all carriers)
- Total trip spending (less fares)
- Spending by category
- Number of nights spent in Canada
- Number of nights by province
- Number of person-visits by province
- Place of residence (US region and country)
- Travelling party size
- Port factor group or SGR (county grouping) code
- Weight of records

Demographic characteristics:

- Composition of travelling party by age group and sex

Places visited:

- Place visited (province, Census Metropolitan Area) (max. of 10)
- Length of stay in place visited (number of nights)
- Accommodations used in place visited

A complete listing of the variables on the PUMF, with their frequencies, can be found in the data dictionary.

Appendix A: United States Port Factor Groups

See file User_Guide_E_MICRODATA2016-Visitors-Appendix.pdf

Appendix B: Canadian Custom Ports Lists

See file User_Guide_E_MICRODATA2016-Visitors-Appendix.pdf

Appendix C: U.S. States and Regions List

See file User_Guide_E_MICRODATA2016-Visitors-Appendix.pdf

Appendix D: Overseas Factor Groups

See file User_Guide_E_MICRODATA2016-Visitors-Appendix.pdf

Appendix E: Foreign Countries List

See file User_Guide_E_MICRODATA2016-Visitors-Appendix.pdf

Appendix F: Variance estimation for public use microdata files

Approximate sampling variability tables

In order to supply coefficients of variation (CV) which would be applicable to a wide variety of estimates produced from this microdata file and which could be readily accessed by the user, a set of Approximate Sampling Variability Tables have 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. It should be used for continuous estimates only, such as expenditures, person nights or person trips.

How to Use the Coefficient of Variation Tables for Continuous Estimates

The approximate coefficients of variation from the Approximate Sampling Variability Tables for estimates of the number of person trips, person nights and expenditures (i.e., continuous variables) can be obtained by using the appropriate table based on the province of residence and the quarter of reference, or for an annual estimate, the province of residence and annual table.

Estimates of Continuous Variables

In the case of the estimate of a continuous variable, the coefficient of variation depends only on the size of the estimate itself. On the Approximate Sampling Variability Table for the appropriate geographic area and time period, first locate the column that corresponds to the variable of interest (i.e., person trips, person nights or expenditures). The intersection of this column with the corresponding magnitude of the estimate itself should give the CV of interest. Note that there are two separate set of columns with different scales for the estimates. The last column on the right should be used for expenditures estimates only, while the set of columns on the left should be used for estimates of the remaining continuous variables.

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 $\alpha_1 \alpha_2$ 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.

Examples of Using the Coefficient of Variation Tables for Continuous Estimates

A series of examples from the fictitious data for the 2009 “Pangea International Travel Survey” are presented below to familiarize the user with the Continuous Estimates tables. Please note that the data for these examples are different than the results obtained from the current survey.

Example 1: Estimates of Expenditures

Suppose that a user estimates that non-resident travellers visiting Pangea, spent \$786 million on international travel in May 2009. How does the user determine the coefficient of variation of this estimate?

- 1) Refer to the Person coefficient of variation table for Pangea – May 2009 below. Note that there are separate tables for each quarter. In addition, there are annual tables available for the year as a whole.
- 2) Since this is an Expenditure estimate, the estimate value can be found in the last column on the right side of the table.
- 3) The value of the estimate itself (\$786 million) is not on the list, so it should be rounded to \$800 million.
- 4) By looking on the “Expenditures” column, the approximate coefficient of variation of the estimate is 10.3%. The finding that the non-resident travellers to Pangea spent \$786 million in international traveling in the reference period is publishable with no qualifications.

Example 2: Estimates of Person Trips

Suppose that a user estimates that non-resident travellers visiting Pangea, aged between the ages of 18 and 24 took 63,226 same-day person trips in May 2009. How does he/she determine the coefficient of variation of this estimate?

- 1) Using the same table as Example 1, now the value of the estimate must be found in the first columns to the left since it refers to person trips. This would be the case for estimates of person nights as well.
- 2) Similarly to the previous example, the value of the estimate itself (63,226) is not on the list, so it should be rounded to 60,000.
- 3) By looking on the “Person Trips” column, the approximate coefficient of variation of the estimate is 36.6%. This estimate is considered unacceptable and Statistics Canada recommends this estimate not to be released.

Approximate Sampling Variability Tables Pangea Domestic Travel Survey - May 2009 Pangea					
Estimate ('000)	Coefficient of variation			Expenditures	Estimate ('000,000)
	Household Trips	Person Trips	Person Nights		
5	70.2	80.2	142.3	43.7	5
10	57.4	64.4	109.8	35.8	10
20	47	51.8	84.7	29.4	20
30	41.7	45.6	72.7	26.2	30
40	38.4	41.6	65.3	24.1	40
50	36	38.8	60.1	22.6	50
60	34.1	36.6	56.1	21.5	60
70	32.6	34.9	52.9	20.6	70
80	31.4	33.4	50.4	19.8	80
90	30.3	32.2	48.2	19.1	90
100	29.4	31.2	46.3	18.6	100
200	24.1	25	35.7	15.2	200
300	21.4	22	30.7	13.6	300
400	19.7	20.1	27.5	12.5	400
500	18.4	18.7	25.3	11.7	500
600	17.5	17.7	23.7	11.1	600
700	16.7	16.9	22.3	10.7	700
800	16.1	16.2	21.2	10.3	800
900	15.5	15.6	20.3	9.9	900
1,000	15.1	15.1	19.5	9.6	1,000
2,000	12.3	12.1	15.1	7.9	2,000
3,000	11	10.6	12.9	7	3,000
4,000	10.1	9.7	11.6		4,000
5,000	9.4	9.1	10.7		5,000
6,000	9	8.6	10		6,000
7,000	8.6	8.1	9.4		7,000
8,000	8.2	7.8	9		8,000
9,000	8	7.5	8.6		9,000
10,000	7.7	7.3	8.2		10,000
15,000	6.9	6.4	7.1		15,000
20,000		5.8			20,000
25 ,000					25 ,000
NOTE: FOR CORRECT USAGE OF THESE TABLES PLEASE REFER TO THE MICRODATA DOCUMENTATION					