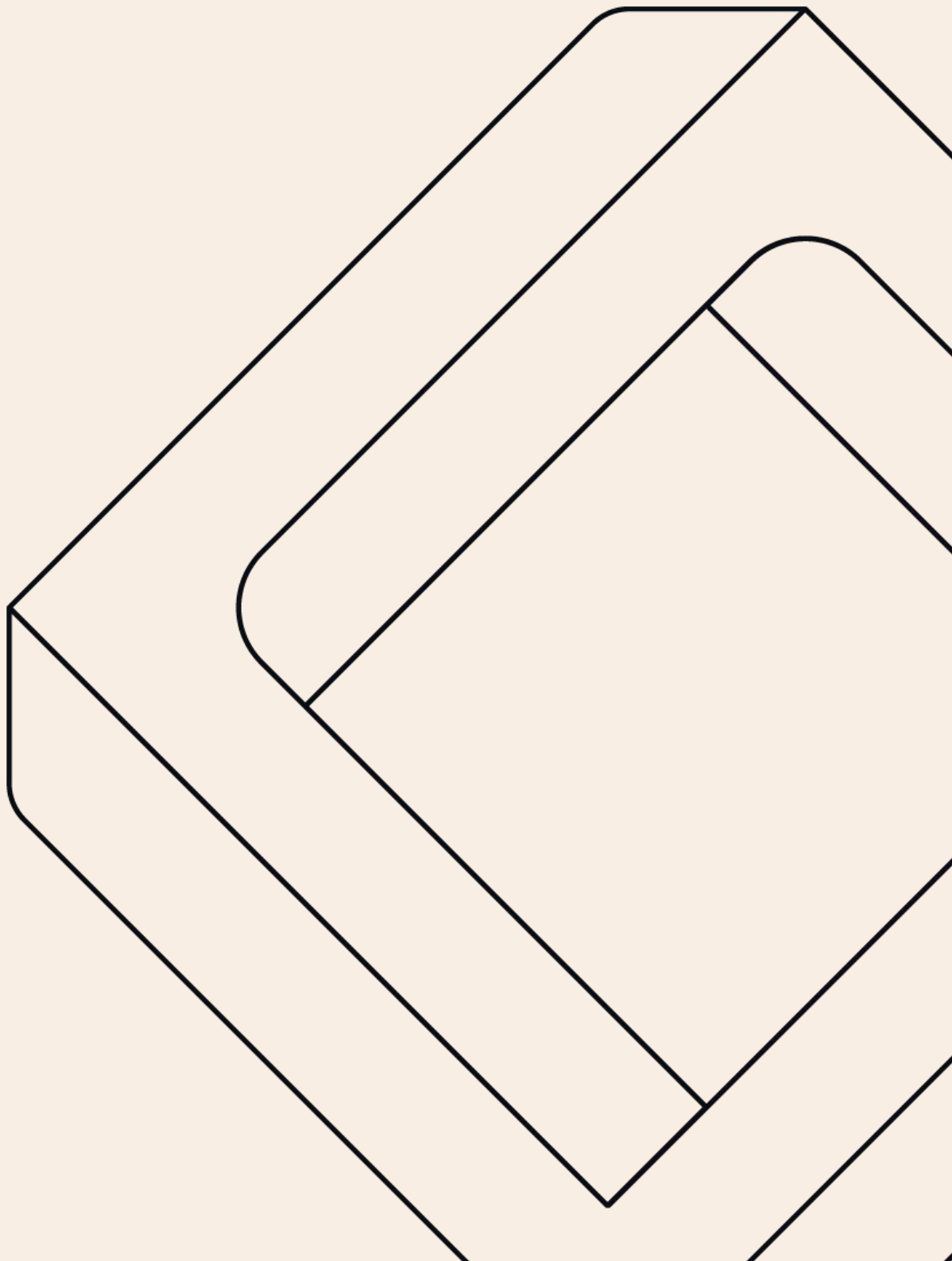


IVS Technical Description

Version: 5.0

Release date:



Contents

1. Background	4
1.1 Purpose	4
1.2 History	4
1.3 Responsibility	5
2. Objectives	6
2.1 Survey objectives	6
2.2 Frequency	6
2.3 Main outputs	6
3. Population	7
3.1 Target population	7
3.2 Survey population	7
3.3 Specific coverage differences between the Population and Survey Population	8
3.4 Sample frame	9
4. Sample design	10
4.1 Accuracy requirements	10
4.2 Sampling structure	10
4.3 Allocation of sample	11
4.3.1 Superstrata allocation	11
4.3.2 Substrata allocation	11
4.3.3 Selection of PSU	12
4.4 Respondent Sample Collection	13
5. Data collection	14
5.1 Recruitment	14
5.2 IVS data collection	15
6. Editing	16
6.1 Variables requiring editing	16
6.2 Data cleaning edits	16
6.2.1 Date of arrival	16
6.2.4 Expenditure	17
6.3 Derived or constructed variables	17
6.3.1 Exchange rates	18
6.3.2 Travel party	18
6.3.3 Removal of airfares	18
6.3.4 Multiple countries	19

6.3.5 Travel style	20
6.3.6 Length of stay (LOS)	20
7. Weighting	21
7.1 Stage 1 weighting	21
7.2 Stage 2 weighting	22
7.3 Stage 3 weighting	24
7.4 Rim weighting	27
7.5 Stage 4 weighting	28
7.5 Final weight	30
8. Outliers	31
8.1 Outlier detection	31
8.2 Outlier treatment	31
9 Response rate calculation	33
9.1 Screener response rate	33
9.2 Survey response rate	33
10 Estimation	34
11 Sample errors	36
11.1 Variables requiring sample errors	36
11.2 Sample error models	36
12. Appendix	38
12.1 Abbreviations	38

1. Background

1.1 Purpose

The International Visitor Survey (IVS) is the primary source of data on tourism-related expenditure by overseas visitors to New Zealand. This data is a Tier 1 official Statistic¹ and contributes to the production of several official statistics such as balance of payments, international trade in services, gross domestic product, national accounts, and the tourism satellite account. Tier 1 statistics are the most important official statistics – they are essential to understanding how well New Zealand is performing and are used to inform critical decision making.

IVS data is also used by several government agencies to inform policy advice, enhance understanding of international visitors, and monitor progress of strategy.

A key benefit of the IVS is that it provides a long-term, consistent time series to evaluate.

1.2 History

The IVS started in 1984 with the International Survey of Visitor Expenditure. Prior to 1984, international visitor expenditure information was collected from the Bank of New Zealand travel receipts. From 1984 onwards, the IVS was conducted at varying intervals, becoming a bi-annual survey through the late 1980s and early 1990s.

In January 1995, the IVS became a continuous ongoing survey, with the questionnaire changing in January 1997 to accommodate more expenditure details as well as activity and attraction information. Throughout this time various market research companies have been contracted to conduct the IVS. NFO New Zealand Limited conducted the IVS from 1997 through to 2002 in conjunction with the then client, Tourism New Zealand. In July 2002, in line with the New Zealand Tourism Strategy 2010, responsibility for the IVS was transferred to the Ministry of Tourism. The Nielsen Company took over the contract at the beginning of 2003. In 2008 this contract was transferred to Covec who became responsible for the data collection, with the Ministry of Tourism and then subsequently the Ministry of Economic Development managing the data processing and final analysis of the survey.

In 2011 the IVS was reviewed as part of a review of all the tourism data collected by the Ministry of Economic Development. This review concluded that the IVS needed to be redesigned to reflect changes in the industry and to improve its reliability. The major changes being:

- A move from a full face-to-face interview at the point of departure to a two-stage process – in-person recruitment, followed by an email with a link to an online survey.

¹ [Legislation, policies, and guidelines | Stats NZ](#)

- An increase in the sample size target from 5,200 to 8,900.

The redesigned IVS commenced in January 2013 with TNS New Zealand (subsequently renamed Kantar New Zealand) being responsible for collection and processing of the data and the Ministry of Business, Innovation and Employment being responsible for the analysis and output of the data.

A further review of the IVS was conducted in 2018. This review led to improvements in the processes associated with data collection and greater documentation and systemisation of those processes. The review also led to a change in the sampling approach, a move from time-based sampling to flight-based sampling. This change was made to improve the ability to recruit visitors from target countries.

Since 2018, the IVS questionnaire has gone through two major overhauls, the first in 2019 and the second in 2022. The purpose of the overhauls was to: (a) reduce the length of the questionnaire, and (b) improve the clarity and understandability of the questions. The IVS was paused in March 2020 due to COVID-19 and re-started in July 2022.

1.3 Responsibility

The Ministry of Business, Innovation and Employment is responsible for commissioning the IVS. Verian (Verian Group New Zealand Limited) are responsible for the collection, processing, and delivery of data obtained through the IVS survey – having taken over from Kantar after the merger of Kantar and Colmar Brunton, and subsequent decoupling into Kantar (providing research to commercial entities) and Verian (providing research-based evidence and advisory services to the public sector). MBIE is responsible for the analysis and reporting of the data and results. Annual and quarterly results are provided on MBIE's Tourism and Evidence Insights Centre: [Tourism Evidence and Insights Centre \(mbie.govt.nz\)](https://www.mbie.govt.nz/tourism-evidence-insights-centre).

Detailed information about the survey methodology, quality and outputs are provided publicly on MBIE's website [International Visitor Survey \(IVS\)](https://www.mbie.govt.nz/international-visitor-survey).

The survey is carried out with the assistance of Auckland International Airport, Wellington International Airport, Christchurch International Airport, Queenstown Airport, and Statistics New Zealand.

2. Objectives

2.1 Survey objectives

The purpose of the IVS is to provide accurate, quarterly national information on the characteristics, behaviour and expenditure of international visitors to New Zealand (excluding visitors whose primary purpose of visit was to attend a recognised educational institute as foreign fee-paying students). The specific objectives are:

- To measure the amount of expenditure of international visitors;
- To provide data for determining the travel credits component in the Balance of Payments, and tourism expenditure in the Tourism Satellite Account;
- To measure the amount of expenditure by country;
- To determine the activities international visitors participate in, the transport and accommodation types used and the places stayed overnight.

2.2 Frequency

The IVS is a continuous online survey with recruitment of international visitors conducted on a daily basis with the exclusion of public holidays. Results are reported on a quarterly basis.

2.3 Main outputs

A range of data is collected from international visitors, including their country of origin, age, gender, expenditure, accommodation usage, length of stay, transportation, activities and attractions and regions visited.

The main outputs for the IVS include:

- Estimates of annual and quarterly total expenditure of international visitors to New Zealand.
- Estimates of average expenditure per visitor to New Zealand.
- Estimates of average expenditure per night of international visitors to New Zealand.

3. Population

3.1 Target population

The population of interest for estimating the total spend of short-term visitors to New Zealand are people who've been on a tourism visit to New Zealand. A tourism visit being defined as: "the term 'tourism visit' refers to a stay in a place during a tourism trip", and a tourism trip being defined as: "A tourism trip is one that takes a traveller to a main destination outside his/her usual environment, for less than a year, for any main purpose (business, leisure, or other personal) other than to be employed by a resident entity in the place that they visited. A traveller who takes a tourism trip is called a visitor²".

Given the target population definition above, the population excludes:

- Individuals who reside in New Zealand (New Zealand was their last place of residence for twelve months or more).
- Individuals who have been in New Zealand for longer than 365 days.
- Those in transit who have not formally entered New Zealand (they have not filled out an arrival card and handed it into Customs).
- Individuals who visit New Zealand to serve in the armed forces, for diplomatic or consular business, or aircrew on duty or between flights.
- Individuals whose next country of permanent residence is New Zealand.

The target population also excludes:

- Those aged less than 15 years because parental consent is required to interview those under 15 and this would not be practical.
- Those whose primary purpose of visit was to attend a recognised educational institute and will be foreign fee-paying students. This is because the spending by this group will be estimated separately as the spending patterns from this segment create problems as they are so different from other international visitors.

3.2 Survey population

The survey population differs in a number of ways from the population due to legal and administrative reasons. The survey population for IVS is the population as defined in the target population and are departing from Auckland, Wellington, Christchurch, and

² The definition of 'tourism visit' and 'tourism trip' are taken from the United Nations and World Tourism Organisation Recommendations on Tourism Statistics, 2008.

Queenstown international airports. Specific areas of under-coverage and over-coverage are outlined in section 3.3 below.

3.3 Specific coverage differences between the Population and Survey Population

Individuals in the target population not covered by the survey population are:

- Individuals who have formally entered New Zealand but have not stayed in New Zealand for at least one night during their trip.

This is because international visitors who arrive in New Zealand and do not stay at least one night are generally considered to be in transit on their way to another country and are therefore deemed not to have “visited” New Zealand.

- Individuals whose port of departure from New Zealand is not Auckland, Wellington, Christchurch or Queenstown international airport. This includes departures from any seaport.
- Individuals who depart New Zealand on private aircraft.

Individuals departing on private aircraft cannot be covered because of airport security arrangements.

- Individuals who do not understand one of the survey languages (English, Japanese, Simplified Chinese, Korean, French, German, and Spanish) sufficiently to be able to answer the questions.

This is because the survey is only available in the languages listed. Other languages will be considered if there is sufficient demand.

- Individuals who do not have at least one of the following: an email address, a phone capable of receiving an SMS, a device with a camera that is connected to the internet (to be able to scan a QR code).

Individuals not in the target population but included in the survey population are:

- Those whose main purpose of visit to New Zealand is to work in New Zealand.

This is due to the difficulty in distinguishing between visiting New Zealand for business purposes (included in the target population) and visiting New Zealand for the purpose of working for an employer in New Zealand (excluded from the target population). Excluding individuals whose purpose of visit is to work in New Zealand would potentially risk also excluding those who visit for business purposes.

Additionally, individuals who have been employed in New Zealand may have multiple reasons for visiting New Zealand, particularly those on working holidays.

3.4 Sample frame

The sample frame is international visitors to New Zealand who have been through immigration and security and are waiting to board international flights departing from Auckland, Wellington, Christchurch and Queenstown international airports.

4. Sample design

4.1 Accuracy requirements

The sample is designed to achieve a 5% relative sample error (at the 95% confidence level) for total expenditure and <10% for the top six countries (Australia, United Kingdom, United States, China, Japan, and Germany).

4.2 Sampling structure

1. **Superstrata:** Auckland; Christchurch; Wellington; Queenstown.
2. **Substrata:** For Wellington and Queenstown there is no further breakdown. For Auckland and Christchurch a set of New Zealand airports to overseas airports. Substrata defined on:
 - i) Direct New Zealand to country of residence flights (Australia, China, United States, Japan, Korea and Canada). This group is sometimes further refined to known transiting airports and direct only.
 - ii) New Zealand to airports that are major transiting hubs (United Kingdom, Germany)
 - iii) Remainder.
3. **Primary sampling unit (PSU)** The PSUs are flights out of Auckland, Christchurch, Wellington and Queenstown.

The substrata are:

Strata	Destination
1	Auckland to Sydney*, Melbourne*
2	Auckland to other Australia
3	Auckland to Shanghai*, Hong Kong*
4	Auckland to other China
5	Auckland to Singapore*, Doha*, Dubai* and Kuala Lumpur*
6	Auckland to Canada
7	Auckland to Japan
8	Auckland to South Korea
9	Auckland to USA
10	Auckland to other airports
11	Christchurch to Sydney*, Melbourne*
12	Christchurch to other Oceania

13	Christchurch to Singapore*, Dubai*
14	Christchurch to China
15	Wellington to all airports
16	Queenstown to all airports

*Transit airports

4.3 Allocation of sample

4.3.1 Superstrata allocation

1400 flights are to be sampled every quarter to provide consistent interviewer workload. The allocation of flights for Wellington and Queenstown is fixed (70:40) based on interviewer numbers. The allocation of flights for Auckland and Christchurch are based on available flights in each strata and the best allocation to achieve country quotas.

4.3.2 Substrata allocation

Substrata Kish Allocation on the current sample size (8,875).

$$n_c = n \times \frac{\sqrt{N_c^2 + N^2}}{\sum_c \sqrt{N_c^2 + N^2}}$$

where N_c is the number of departures in country $c \in C = \{Australia, United Kingdom, \dots\}$, and N is the average size.

For the top 8 countries excluding Australia, double that sample size.

The table below shows an example country distribution using the above Kish allocation and a sample size of 8875. Note that the country data is 2011 data. This is currently being used but will be reviewed when overseas visitor travel patterns are redeemed to have stabilised after COVID-19.

Country	Annual target response size
Australia	1,513
United Kingdom	932
United States of America	868
China	816
Japan	752
Germany	748
Korea	742

Canada	740
Other	1,765
Total	8,875

4.3.3 Selection of PSU

In the month before the quarter the number of scheduled flights for each strata is obtained. While the actual number of flights may vary from the initial schedule used in the selection of the PSUs, experience shows they are generally minor. This gives the maximum number of flights from each strata that can be selected. The table illustrates the flights available for Quarter 2 (April to June), 2024.

Strata	Destination	Available flights
1	Auckland to Sydney, Melbourne	1631
2	Auckland to other Australia	1071
3	Auckland to Shanghai*, Hong Kong*	267
4	Auckland to other China	91
5	Auckland to Singapore*, Doha*, Dubai* and Kuala Lumpur*	560
6	Auckland to Canada	84
7	Auckland to Japan	89
8	Auckland to South Korea	39
9	Auckland to USA	467
10	Auckland to other airports	963
11	Christchurch to Sydney*, Melbourne*	631
12	Christchurch to other Oceania	287
13	Christchurch to Singapore*, Dubai*	91
14	Christchurch to China	0
15	Wellington to all airports	605
16	Queenstown to all airports	660
	Total	7,536

As described above, the flight allocation for Wellington and Queenstown is fixed at 70 and 40 flights, respectively. The fixed allocation leaves 1290 flights to be selected across the remaining eligible flights.

Statistics New Zealand provide data on the number of travellers for the target countries that have departed from New Zealand to all the destination airports. The destination airport numbers are aggregated to strata and estimates of the number of travellers out of the six recruited per flight are made.

4.4 Respondent Sample Collection

In order to achieve the final desired sample size – the actual number of respondents recruited will need to take into account the dropout rate due to the number of stages in the collection of the survey – the first being invalid contact details (email or mobile number), the second being non-response, and the third invalid response.

The assumed response rate is 30% for each country. In practice response rates vary by country and as such the practice of assuming a blanket response rate across country may be reviewed.

Country	Annual final sample size (from 4.3.2)	Response rate	Annual initial recruitment targets
Australia	1,513	30%	5,043
United Kingdom	932	30%	3,107
United States of America	868	30%	2,893
China,	816	30%	2,720
Japan	752	30%	2,507
Germany	748	30%	2,493
Korea	742	30%	2,473
Canada	740	30%	2,467
Other	1,765	30%	5,883
TOTAL	8,875		29,586

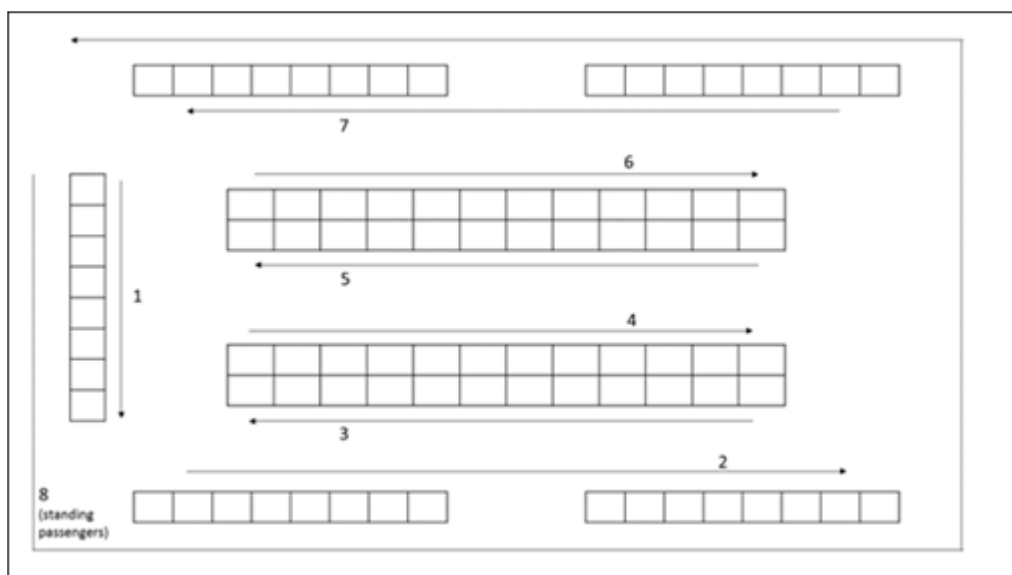
Although international visitor arrivals are highly seasonal the sample size is allocated evenly over the year.

5. Data collection

The IVS uses a two-part survey design. The first part involves screening departing international passengers for eligibility and recruiting them to participate in the second part. The second part is an online survey which the visitor completes in their own time.

5.1 Recruitment

Interviewers arrive at the departure lounge of selected flights (PSUs) approximately 75 minutes before the scheduled departure time of the flight. On arrival, the interviewer selects a start point and a direction and sequence for approaching every second passenger. An example (taken from the interviewer fieldwork manual) of the start, direction, and sequence an interviewer may take is illustrated below. In the illustration seated passengers are approached first, before moving to standing passengers.



Passengers are taken through a short screening questionnaire (see the Appendix) to determine their eligibility to participate in the survey and to collect their contact information (email or mobile phone number) so a link to the online questionnaire can be sent to them. Passengers can also opt to take a QR code to link them to the online survey.

As part of the recruitment process, passengers are offered a flyer which outlines the purpose of the survey, where to find more information about the survey, and details about the prize draw (those completing the online survey go into a draw to win a gift voucher valued at NZ\$500). The flyer is offered in English, Simplified Chinese, Japanese, Korean, French, German, and Spanish.

5.2 IVS data collection

Following recruitment, those international visitors who provided an email address or phone number are sent an email or SMS, with a unique link to the online survey. The emails and SMSs are sent automatically in batches every five minutes, so that the longest a visitor has to wait to start the survey is five minutes. Those who opt to take a QR code can begin the online survey immediately.

Visitors who have not completed the questionnaire after seven days are sent a reminder email or SMS, and those that still haven't completed after 14 days are sent a second reminder. The visitors who opted to take a QR code initially, were also asked for their email address as a method of sending them a reminder. Those that provided an email address for reminders, are sent reminders on the same schedule as those who opted to receive the initial link to the survey by email or SMS.

6. Editing

6.1 Variables requiring editing

All variables in the IVS are subject to micro editing to ensure validity and consistency in responses.

Expenditure and weighting variables are subject to macro editing, as these variables are affected by the expenditure outlier treatment process.

6.2 Data cleaning edits

The following checks are performed on the data through algorithms designed to identify potentially incorrect data that is then investigated manually. All edits and respondent deletions are recorded in the 'edits' component of the quarterly report completed at the end of each quarter.

6.2.1 Date of arrival

The date of arrival is checked for any likely errors by comparing the calculated length of stay (the number of days between the date of arrival (IVS_ITI2) and the date of departure (day the visitor was recruited)) and the stated length of stay (IVS_ITI8) and, if available, the sum of the number of nights they stayed (based on the answers to the map question, IVS_ITI5). Checks are done for the following:

- Date of arrival is after the departure date. If this occurs that the date of arrival is calculated from the stated length of stay (IVS_ITI8).
- Date of arrival of is the same as the departure date. If this occurs and the stated length of stay (IVS_ITI8) is one day, then the visitor is removed from the data as they are deemed to be 'in transit'. If date of arrival is the same as the departure date and the stated length of stay is greater than one day, then the date of arrival is calculated from the stated length of stay (IVS_ITI8).
- Date of arrival is more than 365 days before the date of departure. If this occurs and the stated length of stay (IVS_ITI8) is also more than 365 days, then the respondent is removed from the data as they are not a short-term visitor. If the stated length of stay is less than 365 days, then the date of arrival is calculated from the stated length of stay (IVS_ITI8).
- The calculated length of stay is a month (+/- a day), or multiples of a month, different than the stated length of stay. If this occurs, the respondent is judged to have entered the wrong month when completing the date of arrival question and the date of arrival is calculated from IVS_ITI8.

6.2.4 Expenditure

After expenditure has been processed/calculated – i.e., converted into New Zealand dollars, divided by the number of people the respondent was answering the spend questions for, the removal of the flight component and days spent in other countries from package amounts (see 6.3) – expenditure is checked for the following:

- Is the total value above \$0? If the total is \$0 then the respondent is deleted.
- Has the respondent entered the same low value in each category (e.g., \$1 in each category)? If so, the respondent is deleted.
- Does the respondent have an extremely low spend and if so, is this possible? For example, if they are in New Zealand for a reason other than visiting family and friends and their spend is less than \$10 per day.

Extremely high or low expenditure is also checked for two occasional respondent data entry errors:

- The number of people the respondent said they would answer the expenditure questions about and the number they actually answered about. This error is identified by dividing the airfare by the travel party size. If the airfare is less than 25% of the average airfare for the destination, then the respondent is deemed to have made an error and travel party size is reduced to one.
- The currency the respondent said they would answer in and the currency they actually used when entering amounts into the expenditure table. This is identified by comparing the airfare (once converted into New Zealand dollars) to the average airfare to the country of residence. If the airfare is deemed to be highly unlikely (e.g., NZ\$250,000), then the airfare expenditure and other expenditure is reviewed in New Zealand dollars versus the currency of the country of residence (or selected other currency). If the airfare (and other expenditure is more consistent) with the currency reviewed against than the currency stated in the survey by the respondent, then the currency selected is changed. Note that this check only identifies errors in currency selection with currencies very different in value to the New Zealand dollar (i.e., South Korean Won, Japanese Yen, and Indian Rupee).

6.3 Derived or constructed variables

Respondents are given the option to report their expenditure:

- In a foreign currency
- Include one or more members of their travel party (including children)
- Include travel in other countries before or after their travel to New Zealand.

The expenditure data is adjusted to represent expenditure in New Zealand dollars, per person for travel in New Zealand. The adjusted expenditure data is used for imputation, weighting, outlier treatment and estimation.

6.3.1 Exchange rates

For respondents that report their expenditure in a foreign currency it is adjusted to New Zealand dollars based on the average exchange rate over the quarter.

$$\text{Adjusted Expenditure (Currency)} = \text{Original Expenditure} / \text{Exchange Rate}$$

The exchange rates are calculated from x-rates.com and exchangerates.org.uk, with the monthly data averaged.

6.3.2 Travel party

Where the respondent has reported expenditure including other members of their travelling party the expenditure is adjusted to represent expenditure per person.

$$\text{Adjusted Expenditure (Travel Party)} = \text{Adjusted Expenditure (Currency)} / \text{Party Size}$$

Note that where children have been included in the expenditure then they represent ½ of an adult.

$$\text{Party Size} = \text{Number of People} - \frac{1}{2} \text{Number of Children}$$

6.3.3 Removal of airfares

For those travellers on package and tours who state that they don't know the flight cost component of their package, the flight cost component of their package is calculated and then removed from their expenditure using the following method.

Airfares are sourced from Cirium, an online source of airlines booking and air traffic data. Extreme values (extreme flight costs) are excluded from the analysis by removing the top and bottom 5% of airfares. The choice of a 5% boundary balances the need to exclude outliers with the need to preserve sample size.

The airfares taken from Cirium are based on 'discount coach' (which includes 'discount economy') from the point of origin (first airport on a passenger's itinerary) to the destination airport (Auckland) and return. The average fare using this data is calculated by dividing the total revenue by the total number of passengers on this route.

Average return airfares are calculated for the following countries: Argentina, Australia, Belgium, Brazil, Canada, China, Cook Islands, Denmark, Fiji, France, French Polynesia, Germany, Hong Kong, India, Indonesia, Ireland, Israel, Italy, Japan, Malaysia, Netherlands, Philippines, Samoa, Saudi Arabia, Singapore, South Africa, South Korea, Spain, Sweden, Switzerland, Taiwan, Thailand, United Arab Emirates, United Kingdom, United States.

When an airfare data is needed as input to the IVS for a country not in the list, the airfare data of the nearest country listed is used.

Multi-country package

In the case of a multi-country package the airfare data is calculated as the sum of average fares from:

- an origin country to Sydney international airport,
- Sydney international airport to Auckland international airport,
- Auckland international airport to the origin country.

The result is an average fare for a round-trip itinerary. The multi-country packages are calculated for the following countries: United Kingdom, China, United States. Where an airfare is needed for a country not on the list, the nearest country in the region is used – i.e., for countries in the Americas, the United States fare are used; for countries in Europe, the United Kingdom fare is used; for countries in Asia, the China fare is used; for all other countries, the United Kingdom fare is used.

The airfare must be removed prior to allocating the proportion of costs between the countries as per section 6.3.4 below. The average airfares are supplied by MBIE at the end of the quarter.

Checks are done to ensure that the airfare removed is not greater than 75% of the package cost (industry estimates that for a 7-10 day trip the airfares will be 40 – 60% of the cost). Where the airfare is greater than 75% of the cost, the airfare component is recalculated to be 75% of the package cost.

6.3.4 Multiple countries

Respondents that have packages or tours that span more than one country must have some of the tour/package cost for the other countries removed as well as the airfare costs. The airfares are first removed using the method outlined in section 6.3.4 then the adjustment made for the countries visited as given below.

$$\text{Adjusted Expenditure (Multi Country)} = \text{Multiplier} * \text{Adjusted Expenditure (Travel Party)}$$

The multiplier is based solely on the proportion of the nights spent in New Zealand.

$$\text{Multiplier} = \text{Nights in NZ} / (\text{Nights in NZ} + \text{Nights spent in other counties as part of the Tour/Package})$$

Note if New Zealand is the only country visited on the trip then the multiplier is equal to one. The multiplier is only applied to prepaid tour/package expenditure.

6.3.5 Travel style

Each respondent is assigned a travel style classification, defined as:

- Tour – travelled around New Zealand with others (aside from those visited New Zealand with) in a multi-day tour organised by a company.
- Package – did not travel as part of a tour and part (or all) of the visit to NZ was a package deal that included international travel to New Zealand as well as something else such as accommodation or transport.
- Independent Traveller – not recorded as Tour or Package.

6.3.6 Length of stay (LOS)

Length of stay is calculated by:

- a) Calculating the number of days between when they said they arrived (question IVS_ITI2) and when they left New Zealand (date they were recruited at the airport).
- b) If the visitor states they don't know when they arrived, then the number of days they said they were in New Zealand is used (IVS_ITI8).

If a length of stay cannot be calculated from using either IVS_ITI2 or IVS_ITI8 then the visitor is removed from the dataset and is recorded in the 'edits report' as not providing a length of stay.

An additional length of stay calculation is done for visitors arriving into New Zealand on a cruise ship, this is used for weighting purposes only (to align with how Stats NZ record the length of stay of cruise ship passengers). The 'length of stay for weighting for cruise ship arrivals' is calculated by calculating the numbers of days between when they said they finally disembarked from the cruise ship (IVS_ITI2a) and when they left New Zealand (date they were recruited at the airport). If the visitor states that they don't know when they finally disembarked the cruise ship, then the response to IVS_ITI8 is used as the 'length of stay for weighting for cruise ship arrivals'.

The process for editing the date of arrival is outlined in 6.2.1.

7. Weighting

The IVS sample is weighted to account for:

- Unequal probabilities of selection
- Known discrepancies between the sample and the population.

A four-stage weighting process is used. The weighting is undertaken quarterly after the data has been edited and cleaned.

- Stage 1 weighting benchmarks to the number of international visitor departures (by air) by age and departure airport.
- Stage 2 weighting benchmarks to the number of international visitor departures (by air) by country of residence, gender and age.
- Stage 3 weighting benchmarks to the number of international visitor departures (by air) by country of residence, purpose of visit and length of stay.
- Stage 4 weighting benchmarks to the marginal totals for six key variables.

7.1 Stage 1 weighting

The stage 1 weighting process adjusts for known discrepancies between the sample and the population, specifically, disproportional age distribution within strata compared to the population as no age quotas are set.

The stage 1 weight for individual i in weighting cell g is:

$$\text{Stage1 Weight}_g = \frac{N_g}{n_g}$$

Where N_g = Number of individuals in the population in weighting cell g and n_g = Number of individuals in the sample in weighting cell g .

Population

The population for stage 1 is all short-term international visitors aged 15 years and over departing from New Zealand by air from Auckland, Christchurch, Wellington and Queenstown Airports during the reference period excluding those whose primary purpose of visit was to attend a recognised educational institute and will be foreign fee-paying students. This information is obtained from Statistics NZ.

Weighting Cells

All individuals in the sample are assigned to a weighting cell. The weighting cells for stage 1 are based on age and airport, a two-way table: Airport x Age.

The airports are:

- 1) Auckland
- 2) Christchurch
- 3) Wellington
- 4) Queenstown

Age cells are defined as:

- 1) 15 to 24 years of age
- 2) 25 to 34 years of age
- 3) 35 to 44 years of age
- 4) 45 to 54 years of age
- 5) 55 years and above

Where the age is missing the individual is assigned to the age group that has the highest count for that country and gender (for the purposes of the weighting). Note that for the purpose of weighting, those respondents who recorded their gender as 'another gender' are randomly assigned as either male or female.

Cell Collapsing

Cell collapsing should not be necessary for Stage 1 as there should be a minimum cell size of 6 sample units.

7.2 Stage 2 weighting

The stage 2 weighting process adjusts for known discrepancies between the sample distribution of country of residence compared to the population and also disproportional sex and age groups within country of residence compared to the population.

The stage 2 weight for individual i in weighting cell k is:

$$\text{Stage2 Weight}_k = \frac{N_k}{n_k}$$

Where N_k = Number of individuals in the population in weighting cell k and n_k = Number of individuals in the sample in weighting cell k .

Population

The population for stage 2 is all short-term international visitors aged 15 years and over departing from New Zealand by air during the reference period excluding those whose primary purpose of visit was to attend a recognised educational institute and will be foreign fee-paying students.

Weighting Cells

All individuals in the sample are assigned to a weighting cell. The weighting cells for stage 2 are based on the complete three-way table of country of residence, age and gender.

Gender cells are defined as:

- 1) male
- 2) female

Age cells are defined as:

- 1) 15 to 24 years of age
- 2) 25 to 34 years of age
- 3) 35 to 44 years of age
- 4) 45 to 54 years of age
- 5) 55 to 64 years of age
- 6) 65 years and above

Country of residence cells are defined as:

- 1) Australia
- 2) United Kingdom
- 3) United States of America
- 4) China, Peoples Republic of
- 5) Japan
- 6) Germany
- 7) South Korea, Republic of
- 8) Canada
- 9) Rest of Europe
- 10) Rest of Asia
- 11) Rest of Oceania
- 12) Africa and the Middle East

13) Rest of the Americas

Individuals in the sample cannot have missing country of residence as this information is required to complete the interview. Those recording their gender as 'another gender' are recoded randomly into male or female (as described in 7.1)

Cell Collapsing

Where necessary *adjacent* age range cells are collapsed to achieve a minimum cell size of 6 sample units.

Cell collapsing for Stage 2 is carried out for the age category and twice - once from the bottom up and once from the top down. The solution with the lowest variance is then selected. For example, if a set is collapsed from bottom up and gives a result of 8:18:6 and top down the result is 12:14:6, then the latter solution is chosen as this has the lowest variance. Where each solution gives the same variance, the top down solution is used.

The variance is calculated as follows:

$$\sigma^2 = \frac{\sum(x - u)^2}{N}$$

Cells are never collapsed by gender or country. Cell collapsing only happens *once* within the Stage 2 weighting process i.e. it is not iterative.

7.3 Stage 3 weighting

The stage 3 weighting process adjusts for disproportional purpose of visit and length of stay bands samples within country of residence compared to the population.

The stage 3 weight for individual *i* in weighting cell *c* is:

$$\text{Stage3 Weight}_c = \frac{N_c}{n_c}$$

Where N_c = Number of individuals in the population in weighting cell *c* and n_c = Number of individuals in the sample in weighting cell *c*.

Population

The population for stage 3 is all short-term international visitors aged 15 years and over departing from New Zealand by air during the reference period excluding those whose primary purpose of visit was to attend a recognised educational institute and will be foreign fee-paying students.

Weighting Cells

All individuals in the sample are assigned to a weighting cell. The weighting cells for stage 3 are based on country of residence, POV and LOS.

The cells are defined as:

Country/Region	POV (purpose of visit)	LOS (length of stay)
Australia	Business	1-3 days
Australia	Business	4 - 364 days
Australia	Holiday	1-6 days
Australia	Holiday	7-9 days
Australia	Holiday	10-15 days
Australia	Holiday	16-364 days
Australia	VFR	1-6 days
Australia	VFR	7-9 days
Australia	VFR	10-15 days
Australia	VFR	16-364 days
Australia	Education	1-364 days
Australia	Conference, Other	1-364 days
Canada	Business	1-364 days
Canada	Holiday	1-364 days
Canada	VFR	1-364 days
Canada	Conference, Other	1-364 days
Canada	Education	1-364 days
China	Business	1-364 days
China	Holiday	1-4 days
China	Holiday	5-364 days
China	VFR	1-364 days
China	Conference, Other	1-364 days
China	Education	1-364 days
Germany	Business	1-364 days
Germany	Holiday	1-364 days
Germany	VFR	1-364 days
Germany	Conference, Other	1-364 days
Germany	Education	1-364 days
Japan	Business	1-364 days

Japan	Holiday	1-6 days
Japan	Holiday	7-364 days
Japan	VFR	1-364 days
Japan	Conference, Other	1-364 days
Japan	Education	1-364 days
South Korea	Business	1-364 days
South Korea	Holiday	1-6 days
South Korea	Holiday	7-364 days
South Korea	VFR	1-364 days
South Korea	Conference, Other	1-364 days
South Korea	Education	1-364 days
United Kingdom	Business	1-364 days
United Kingdom	Holiday	1-12 days
United Kingdom	Holiday	13-364 days
United Kingdom	VFR	1-15 days
United Kingdom	VFR	16-364 days
United Kingdom	Conference, Other	1-364 days
United Kingdom	Education	1-364 days
United States	Business	1-364 days
United States	Holiday	1-6 days
United States	Holiday	7-15 days
United States	Holiday	16-364 days
United States	VFR	1-364 days
United States	Conference, Other	1-364 days
United States	Education	1-364 days
Rest of Europe	Business	1-364 days
Rest of Europe	Holiday	1-15 days
Rest of Europe	Holiday	16-364 days
Rest of Europe	VFR	1-364 days
Rest of Europe	Conference, Other	1-364 days
Rest of Europe	Education	1-364 days
Rest of Asia	Business	1-364 days
Rest of Asia	Holiday	1-12 days
Rest of Asia	Holiday	13-364 days
Rest of Asia	VFR	1-364 days

Rest of Asia	Conference, Other	1-364 days
Rest of Asia	Education	1-364 days
Africa and the Middle East	Business	1-364 days
Africa and the Middle East	Holiday	1-364 days
Africa and the Middle East	VFR	1-364 days
Africa and the Middle East	Conference, Other	1-364 days
Africa and the Middle East	Education	1-364 days
Rest of the Americas	Business	1-364 days
Rest of the Americas	Holiday	1-364 days
Rest of the Americas	VFR	1-364 days
Rest of the Americas	Conference, Other	1-364 days
Rest of the Americas	Education	1-364 days
Rest of the Oceania	Business	1-364 days
Rest of the Oceania	Holiday	1-364 days
Rest of the Oceania	VFR	1-364 days
Rest of the Oceania	Conference, Other	1-364 days
Rest of the Oceania	Education	1-364 days

Note that POV population cells are defined using the POV from the arrival card. If the purpose of visit is missing then this is assigned to the other category. Also note that population LOS is derived by comparing arrival date on arrival cards with departure records.

Survey sample LOS cannot be missing as those who do not provide sufficient information for a LOS to be calculated are deleted from the sample.

Cell collapsing

Cell definitions are adjusted to achieve a minimum of 6 sample units per cell. Note the COPR x POV x LOS bands are pre-defined. However, given there are sample variations each quarter some adjustments may need to be made. Note there are only a minor number of adjustments each quarter. For example, if there is a pre-defined LOS band for respondents who have spent 1-15 nights in NZ and 6 cases are required but there are only 4 in the sample, then an adjustment such as collapsing the nights bands from 1-15 and 16-366 nights to 1-366 nights may be made.

7.4 Rim weighting

The weights in the above three stages are combined using a process called rim weighting or iterative proportional fitting. This works by the weights of each stage as a pre-weight for the

next stage. The process of iterations is repeated 100 times to ensure the weights converge. Note there is very little change in the data after 20 iterations. Therefore, a process incorporating 100 iterations guarantees that for each weighting round the weights will be stable. A series of statistical outputs are made throughout the iterative process to ensure this is the case.

7.5 Stage 4 weighting

The stage four weighting process aims to adjust the rim weights to preserve the marginal totals for following six key variables:

- 1) Country of permanent residence
- 2) Purpose of Visit
- 3) Age
- 4) Departure Airport
- 5) Length of Stay
- 6) Gender

To do this a second iterative proportional fitting process (also known as rim weighting or raking) is used to adjust cell weights so that the marginal totals of the survey sample match to the known population margins from the list of tables below. The process being repeated until the weights converge.

The marginal tables to be used in the stage four weighting are defined below:

Country

- 1) Australia
- 2) China, People's Republic of
- 3) Germany
- 4) United States of America
- 5) Japan
- 6) Korea, Republic of
- 7) Canada
- 8) UK
- 9) Rest of Europe
- 10) Rest of Asia
- 11) Rest of Oceania
- 12) Africa and Middle East
- 13) Rest of Americas

Purpose of Visit

- 1) Business
- 2) Holiday/Vacation
- 3) Other
- 4) Visit Friends/Relatives
- 5) Education

Age

- 1) 15 to 24 years of age
- 2) 25 to 34 years of age
- 3) 35 to 44 years of age
- 4) 45 to 54 years of age
- 5) 55 years of age and over

Departure Airport

- 1) Auckland
- 2) Wellington
- 3) Christchurch
- 4) Queenstown

Length of Stay

- 1) 1
- 2) 2
- 3) 3
- 4) 4
- 5) 5
- 6) 6
- 7) 7
- 8) 8
- 9) 9
- 10) 10
- 11) 11
- 12) 12
- 13) 13

- 14) 14
- 15) 15 to 28
- 16) 29 to 62
- 17) 63 to 93
- 18) 93 to 183
- 19) 184 to 364

Gender

- 1) Male
- 2) Female

Population

Population for stage five is all short-term international visitors aged 15 years and over departing from New Zealand by air during the reference period excluding those whose primary purpose of visit was to attend a recognised educational institute and will be foreign fee-paying students.

7.5 Final weight

The final weight, stage 4 weight, is used to calculate all survey estimates. A population weight is also produced from the stage 4 weight, this projects the survey sample to the population survey in terms of actual numbers of people.

8. Outliers

The Statistics Canada BANFF method is used for outlier detection and treatment.

8.1 Outlier detection

In the BANFF method an observation is identified as an outlier if falls outside the lower bound (median – (median – 1st quartile) x influence factor) or the upper bound (median + (median – 3rd quartile) x influence factor). The influence factor for the IVS has been set at 25 based on the extensive testing by Statistics Research Associates Limited.

The BANFF method is used to detect outliers in the following categories:

1. unweighted total spend
2. unweighted average spend per day
3. unweighted total spend by country
4. unweighted business total spend
5. unweighted business average spend per day
6. weighted total spend
7. weighted average spend per day
8. weighted total spend by country
9. weighted business total spend
10. weighted business average spend per day.

The rationale for detecting outliers in both unweighted and weighted spend categories is that the unweighted set identifies population uniqueness and the weighted set identifies influence. Only outliers that are flagged as being both unique and influential are treated as outliers.

8.2 Outlier treatment

Once cases have been identified as requiring outlier treatment, a factor is calculated in order to bring the weighted total expenditure figure down to the 99th percentile value for weighted expenditure within that category.

Cases are treated in order of size, with the largest outliers treated first. This is repeated through the expenditure groups above. If an identified outlier has already been treated in a previous category, the *treated* case is still considered for treatment in the next categories but with expenditure adjusted to the 99th percentile.

Outlier detection is only conducted at the total expenditure level. If a total expenditure outlier is detected and treated, then the individual expenditure categories have the same scaling factor applied as total expenditure.

9 Response rate calculation

9.1 Screener response rate

The screener response rate is a measure of the potential initial non-response bias caused by travellers who refused to be spoken to, refused to provide their email address, mobile phone number, and accept a QR code. This is calculated as the number of people who supply an email address, mobile phone number, or accept a QR code as a percentage of those approached that are estimated as being eligible.

9.2 Survey response rate

The survey response rate is a measure of the non-response bias caused by eligible travellers who declined to complete the survey once they were recruited. This is calculated as the number of people who complete the survey up to the end of the IVS section (the demographic section) as a percentage of those that supplied an email address, mobile phone number, or accepted a QR code.

10 Estimation

The key estimates calculated from the IVS include quarterly estimates of:

- total expenditure
- total expenditure by purpose of visit
- total expenditure by country of permanent residence

- total nights
- total nights by purpose of visit
- total nights by country of permanent residence

- mean and median expenditure per visitor
- mean and median expenditure per visitor by purpose of visit
- mean and median expenditure per visitor by country of permanent residence

All estimates are attributed to the quarter in which the respondent left the country.

Overall Population Estimates - Expenditure and Nights

Overall population estimates are calculated by summing up the weighted values for all responding individuals in the sample.

$$\hat{X} = \sum_{i \in R} w_i x_i$$

where:

\hat{X} = population estimate of variable x e.g. total expenditure

w_i = final weight for individual i

x_i = adjusted value (after all derivations outlined in section 6 and 9) of x for individual i

R = responding sample

This formula is to be used when calculating estimates of total expenditure and total nights spent in New Zealand.

Subpopulation Estimates – Expenditure and Nights

Estimates of expenditure or nights by either purpose of visit or country of permanent residence are calculated by summing up all weighted values for all responding individuals in the sample for the purpose of visit or country of permanent residence concerned.

$$\hat{X}_D = \sum_{i \in D} w_i z_i$$

where:

$$z_i = \begin{cases} x_i & i \in D \\ 0 & \text{otherwise} \end{cases}$$

D = subpopulation of interest where $D \subseteq R$ e.g. business travellers, Australian visitors.

Mean and median expenditure estimates

Estimates of median expenditure per visitor are calculated by finding the median value for the expenditure for the appropriate population using a statistical package, typically R.

Estimates of mean expenditure per visitor are calculated by dividing the sum of expenditure for the population of interest by the corresponding count in the population of interest.

11 Sample errors

Sample estimates are subject to two kinds of errors. These include:

- Non sampling error, which occurs as a result of anomalies in recording, processing or reporting of data
- Sampling error, which occurs as a result of data being obtained from a sample rather than a census

Non-sampling error is difficult to quantify and there are no standard measures of non-sampling error in the IVS. Standard errors are used to quantify the sampling error in the IVS.

11.1 Variables requiring sample errors

The variables requiring standard errors in the IVS are as follows:

- 1) Total expenditure and expenditure by top 8 markets
- 2) Number of visitors
- 3) Number of nights

11.2 Sample error models

For the IVS, standard errors of the number of visitors, number of nights and total expenditure have been modelled as a function of the estimate itself. This simplifies the calculation of standard errors, as it enables approximate standard errors to be calculated for a range of estimates without needing individual data records or sampling information.

Samples errors are reported by the Relative Margin of Error (RME) at the 95% confidence level. The calculations are performed using statistical process called bootstrapping with 500 replications.

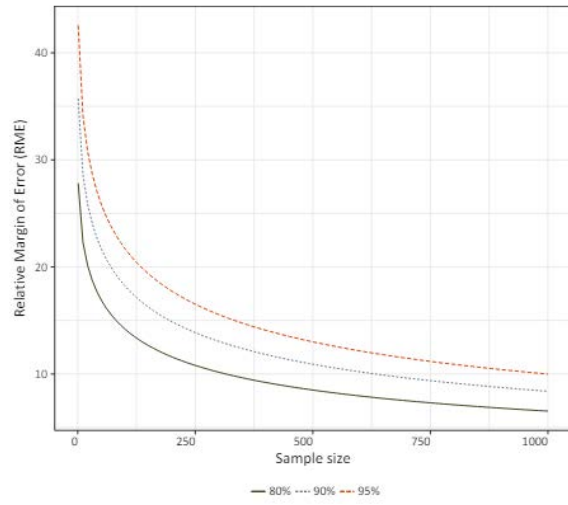
Relative Margin of Error are calculated using the following model:

$$RME_{1-\alpha} = \frac{z_{1-\alpha/2}}{1.96} \times \frac{1}{a + b\sqrt{n}}$$

Where a=0.021, b=0.0025 for RME of sample size at confidence level 100(1-a) % and z value is 1.96 for 95%, 1.64 for 90% and 1.28 for 80%, respectively.

The following table and figure show the RME when using 95, 90 and 80% confidence intervals,

Sample Size	Confidence Level		
	95%	90%	80%
20	31.1	26.0	20.3
50	25.9	21.6	16.9
100	21.7	18.2	14.2
200	17.7	14.8	11.6
500	13.0	10.9	8.5
1000	10.0	8.4	6.5
2000	7.5	6.3	4.9
5000	5.1	4.2	3.3
10000	3.7	3.1	2.4



12. Appendix

12.1 Abbreviations

COPR - Country of Permanent Residence

DEFF - Design Effect

DK – Don't Know

IVS – International Visitor Survey

LOS - Length of Stay

LP - Linear Program

POV - Purpose of Visit

SNZ – Statistics New Zealand

TMT - Ministry of Tourism

UK – United Kingdom

UNWTO –World Tourism Organisation

USA – United States of America



Powering decisions
that shape the world.

