# NEW ZEALAND HOUSEHOLD TRAVEL SURVEY

# 2018-2020

# Methodology Report

**Version 1**
**April 2022**

**Disclaimer**

All reasonable endeavours are made to ensure the accuracy of the information in this report. However, the information is provided without warranties of any kind including accuracy, completeness, timeliness or fitness for any particular purpose.

The Ministry of Transport excludes liability for any loss, damage or expense, direct or indirect, and however caused, whether through negligence or otherwise, resulting from any person or organisation's use of, or reliance on, the information provided in this report. Under the terms of the Creative Commons Attribution 4.0 International (BY) licence, this document, and the information contained within it, can be copied, distributed, adapted and otherwise used provided that –

> ▶ the Ministry of Transport is attributed as the source of the material
>
> ▶ the material is not misrepresented or distorted through selective use of the material
>
> ▶ images contained in the material are not copied

The terms of the Ministry's Copyright and disclaimer apply.

**Additional information**

For more information about the background to the survey see the Ministry of Transport website at www.transport.govt.nz/travelsurvey/

Enquires relating to the household travel survey may be directed to the Ministry of Transport, PO Box 3175, Wellington, or by email on info@transport.govt.nz or travelsurvey@transport.govt.nz .

**Citation**

Ministry of Transport. (2022). New Zealand Household Travel Survey 2018-2020 Methodology Report. Wellington: Ministry of Transport.

Published in April 2022 by the Ministry of Transport, PO Box 3175, Wellington 6140, New Zealand

ISBN 978-1-99-002810-6

This document is available at www.transport.govt.nz

# Contents

# Tables

# Figures

# Introduction

## Overview

The New Zealand Household Travel Survey (NZHTS) collects information about day-to-day travel in New Zealand – such as, how, where and when we travel. The results provide a picture of the travel patterns and choices of all types of people. The survey is used by a range of stakeholders for developing transport policy including road safety, public transport, walking and cycling.

Historically this has been done through a national travel survey, measuring the travel of New Zealanders by collecting 2 days of travel data for everyone in randomly selected households (1989/90, 1997/98 and continuously between 2003 and 2014).

In 2015, the survey changed methodology from a traditional 2-day personal interviewer-based survey, to a 7-day online survey with the option of using a GPS logger (NZHTS GPS7). In 2018, in response to respondent feedback, data collection was reduced back to 2 days, but kept the option of a GPS logger (NZHTS GPS2).

This document describes the methodology used to deliver the 2018-2020 iteration of the survey. This includes the sample selection, the questionnaires that were administered, data collection and data cleaning processes, and the calculation of weights included in the final datasets. A final section provides some notes on differences between the current survey and historical survey and other issues encountered, such as interruptions caused by COVID-19 lockdowns.

## Background

The Ministry provides advice to Government on issues across the whole of the transport system, and the regulatory framework that supports it. It is focused on developing a transport system that maximises economic and social benefits for New Zealand and minimises harm. To do this, we must have a good base knowledge of the transport sector. We also need to understand the future drivers for transport and their implications for Government policy and investment decisions.

The purpose of the Travel Survey is to understand who is travelling in New Zealand as well as when, why, where and how they are travelling. This provides an evidence base to inform central government, local government, academia and elsewhere about travel and the impacts of transport policies.

Examples of areas of interest include:
- exposure rates for relative safety risks
- changes in kilometres travelled, by vehicle, amongst different groups of New Zealanders and for different travel purposes
- cycling and walking patterns
- measuring changing travel patterns nationally and across different geographic areas and demographic groups

There have been numerous national travel surveys over the years. Details of the national surveys are shown in Table 1.

Table 1: Historical national travel surveys

| Survey | Time period | Sample | |
| --- | --- | --- | --- |
| | | Households | People |
| 1989/90 | July 1989 - June 1990 | 3,102 | 7,913 |
| | | | |
| 1997/98 | July 1997 - June 1998 | 5,367 | 14,251 |
| | | | |
| 2003/2004 | July 2003 - June 2004 | 1,382 | 3,365 |
| 2004/2005 | July 2004 - June 2005 | 1,405 | 3,431 |
| 2005/2006 | July 2005 - June 2006 | 1,491 | 3,723 |
| 2006/2007 | July 2006 - June 2007 | 1,439 | 3,539 |
| 2007/2008 | July 2007 - June 2008 | 1,378 | 3,473 |
| 2008/2009 | July 2008 - June 2009 | 3,149 | 7,940 |
| 2009/2010 | July 2009 - June 2010 | 3,430 | 8,840 |
| 2010/2011 | July 2010 - June 2011 | 3,457 | 8,794 |
| 2011/2012 | July 2011 - June 2012 | 3,321 | 8,585 |
| 2012/2013 | July 2012 - June 2013 | 3,273 | 8,300 |
| 2013/2014 | July 2013 - June 2014 | 3,194 | 7,966 |
| | | | |
| 2015/2016 | October 2015 – October 2016 | 1,531 | 3,230 |
| 2016/2017 | November 2016 – June 2017 | 1,355 | 2,473 |
| 2017/2018 | July 2017 – June 2018 | 1,258 | 2,564 |
| | | | |
| 2018/2019[1] | July 2018 – July 2019 | 2,553 | 5,432 |
| 2019/2020[2] | July 2019 – March 2020 | 2,194 | 4,797 |

From 1989-2014 the historical surveys were conducted using face-to-face interviews, based on 2 days of travel. Further information from these surveys is available from http://www.transport.govt.nz/travelsurvey/

In 2014, a pilot was conducted, testing GPS devices, online diaries and smartphones as an alternative method of data collection. At that time, the pilot demonstrated some difficulties with collecting information using smartphone apps but there were potential options for improving data collection as technology advances.

A 2014 review identified core household travel data requirements, after which the Ministry invited companies to participate in a competitive dialogue process as a precursor to delivering a new survey.

Ministry requirements for the new survey included:
- using new technology to survey people more efficiently
- the option of a panel survey approach that would allow the Ministry to expand its evidence base through additional surveying

---

[1] 2018/2019 sample included expansions in Auckland and Tauranga
[2] 2019/2020 sample included expansions in Wellington and Waikato but was also cut short due to COVID-19 restrictions

- a new operating model where the supplier undertook most of the data cleaning and analysis work
- an ability to expand the survey through additional samples purchased by local government.

The chosen supplier was also offered the chance to innovate throughout the life of the contract, allowing for the investigation of new technologies and potential cost savings.

From an initially open pool, a select few suppliers participated in a formal competitive dialogue process which included:
- discussing in detail the needs of the Ministry of Transport and refining the requirements,
- sharing the refined requirements with all invited suppliers, and
- identifying and developing potential solutions with each invited supplier.

As a result, CBG Public Sector Surveying (CBG) was awarded the contract.

The resulting NZHTS GPS7 which was delivered between 2015 and 2018 differed from earlier surveys in some key respects:
- The survey collected 7 consecutive days of travel from all household members
- GPS devices were issued to respondents if they agreed to use them
- Respondents had the option to self-code their travel using a website showing GPS tracks as a prompt
- A fresh sample of meshblocks (MBs) was drawn each year, based on primary sampling units (PSUs) supplied by Statistics NZ[3]

Some questions from previous surveys were removed (e.g. alcohol use) and sections on mobility difficulty and air travel were added.

Initial experiences from the first 2 years in the field found a lower fully completed household rate (40 percent) and a correspondingly higher partial completion rate (where at least 1 person in the household fully completed the 7 days, but not everyone, 22 percent) compared to the historical 2 day survey (67 percent fully completing, 2 percent partially completing).

The online access required excessive manual support, with over 70 percent of people requiring phone assistance to complete the survey at some point during the survey. This ranged from resetting forgotten passwords, to complete phone assistance/interviewing being required to fill in the travel diary. More detailed follow-up monitoring in the 3rd survey year found that only 41 percent of trips were self-completed online, with the rest done over the phone with an interviewer. Usage of the GPS units themselves was high, with 88 percent of those eligible (aged 12 years and over) using a GPS logger on at least one of their travel days. On any given survey day, about 77 percent of trips legs recorded have trip data from a GPS unit.

During the 3rd survey year, CBG piloted variations to the survey methodology and as a result, the decision was made to revert back to a 2-day travel diary, with face-to-face interviews for the 4th survey year (July 2018 –June 2019). Certain core features of the survey were retained. Sampling continued using an address-based sample frame, with face-to-face recruitment. All participants aged 12 years and over continued to be given the option of carrying a GPS unit which could act as a memory jogger. Personal face-to-face interviews

---

[3] More information about meshblocks is available from
http://archive.stats.govt.nz/methods/classifications-and-standards/classification-related-stats-standards/meshblock/definition.aspx

were arranged to complete the two-day travel diary where possible, with paper forms to supplement, to aid full completion by all in the household. Interviews were completed using a computer aided interface, where the available GPS tracks could be displayed to respondents to complete the travel diary for the two days. This has been confirmed to be the survey methodology going forward.

## Survey content

NZHTS collects information about households, the occupants of those houses and about trips made by those people.

The broad categories of information collected are:

**Household**
- Household type
- Number and type of vehicles
- Availability of bicycles in the household

**Occupants**
- Age / gender / ethnicity / income / household relationship matrix
- Addresses of work / school and other points of interest
- Whether a person travelled on a given day
- Frequency of use over the past year for:
    o cycling
    o public transport
    o domestic air travel
- Characteristics of  most recent domestic air travel, including travel to airport
- Distance (0 / 100 / 100+ km) travelled last year via different travel modes
- Recreational use of water transport
- Details of driver's license(s)
- Disabilities
- Employment status
- Occupation
- Education


**Trip data for 2 continuous assigned travel days**
- Start and stop location and time of trips, waypoints
- Trip time
- Trip purpose
- Travel mode
- Vehicle details
- Vehicle occupants
- Trip by Uber-like services

Copies of all questionnaires are available at:
http://www.transport.govt.nz/travelsurvey/

These documents contain the text of questions, possible response options and explanatory notes on questions and routing logic.

# Sampling

## Target and survey population

A target population is the population the survey aims to represent. A survey population is the population actually covered in the survey.

To be able to build a picture of national travel, the population targeted for this survey was the New Zealand 'usually resident' population of all ages.

The survey population consists of those people in New Zealand households.

A small proportion of the target population is excluded from the survey population, including:
- Most types of non-private dwellings (prisons, hospitals, hospices, dementia care units and hospital-level care in aged-care facilities)
- Households located on islands other than the North Island, South Island and Waiheke Island
- People living in aged-care facilities (rest homes) and those temporarily living away from their household in student accommodation (university hostels and boarding schools).

## Sample frame

The following section describes the sample frame for the NZHTS GPS2. For detail of the NZHTS GPS7, please refer to the New Zealand Household Travel Survey 2015-2017 Methodology Report. The NZHTS GPS2 was based on responses from randomly selected households from primary sampling units (PSUs). Statistics NZ chose PSUs to avoid clashes with other large population surveys, such as the NZ Health Survey, the NZ Crime and Victims Survey and the Household Labour Survey.

### Stage 1a – PSUs selected by Statistics NZ

The first stage in building the sample was a selection of 300 PSUs drawn probabilistically proportional to size (PPS) by Statistics NZ.

Statistics NZ provided a list of 300 PSUs along with the selection probabilities. PSUs were selected using a coordinated sampling model designed to make a PSU that has been chosen for other surveys less likely to be selected in a given time period.

### Stage 1b - Select a meshblock within each PSU

A Stats NZ PSU usually comprises one to five meshblocks. To help control travel costs, for the 2015-2020 survey years, one meshblock was selected within each PSU. A subsequent review was conducted in 2020, which revealed the cost benefit of sampling one meshblock from within a PSU to be marginal. As a result, for the 2020-2021 sample, within-PSU meshblock sampling was not undertaken.

The PSU selection was also probabilistic proportional to size, based on the 2013 Census resident count for that meshblock. To be eligible for selection, each meshblock/PSU had to have at least 9 houses.

## Stage 2

Within each meshblock, a systematic sample of 10 houses were drawn[4].  The NZ Post Postal Address File (PAF) was used as the basis for this sample, filtered to residential addresses only.

Houses were listed in house number order. A start house was randomly selected and a "skip" was calculated based on the number of houses in the Postal Address File. This ensured that if every $n^{th}$ house was chosen from this list, the sample would contain 10 houses. This systematic sampling provided a representative sample of houses from the meshblock.

Any house that has received a postal delivery is included in the Postal Address Frame, as well as any addresses notified to NZ Post from Councils. Therefore, in most meshblocks, the Postal Address Frame is a reasonably complete enumeration of households. However, if a house has not received a postal delivery, is erroneously classified as a business address, or does not have a household postal address it will not be included in the sample. An additional enumeration check is required to identify missing houses and to select them for the sample with the same probability as the other houses selected in the meshblock. This enumeration can be undertaken, and the selection applied, in-field when the meshblock is being surveyed for the first time. A meshblock with a low number of houses in the Postal Address Frame will be fully enumerated. This typically occurs in isolated rural meshblocks.

Advantages of using the hybrid "PAF/In-field Enumeration" approach are:
  (1) In most meshblocks, enumeration starts with a complete list that is likely better than one obtained by de novo in-field enumeration.
  (2) The sample can be drawn in advance and letters delivered to specific houses before the first house call by the interviewer. The household can then be provided with the name of the interviewer and their contact details in advance.
  (3) Surveying and in-field enumeration can happen at the same time. This reduces travel costs and reduces transcription error rates.

This process yielded approximately 3,000 houses[5]. The Postal Address File includes some addresses that might not have actual dwellings (typically from when Council's supplied addresses for houses not yet built, or from inaccuracies in address-type coding). For example, churches often appear as a residential address. The "Not a Dwelling or Empty Section" (NDE) rate was about 2.5% (based on having surveyed over 100,000 houses using this methodology). All such addresses were audited. Additionally, there were a proportion of addresses in the sample which were not occupied. The sum of the NDE and not occupied houses was referred to as the "sample loss". These addresses were considered to be ineligible for the survey and are excluded from the response rate calculation.

## Stage 3

All eligible occupants in every recruited house were invited to participate in the survey. The NZHTS GPS2 core sample was designed to deliver 1,670 fully completed households, in which all occupants completed 2 consecutive/continuous days data collection and answered a short questionnaire.

---

[4] Only 9 houses per meshblock were drawn in the 2018-19 year. This was increased to 10 from the 2019-20 to ensure that the target yield of fully-completed households was achieved.
[5] 2700 in 2018-19.

Table 2: Estimated yield per year

| Item | Value | Notes |
|---|---|---|
| PSU / MB | 300 | PSU supplied by Statistics NZ, meshblocks selected by CBG |
| Houses per MB | 10 | Systematic sample from Postal Address File supplied by NZ Post |
| Selected houses | 3,000 | |
| Sample loss | 12% | Based on previous large surveys using same methodology |
| Occupied dwellings | 2,640 | |
| Expected combined Response Rate | 65% | Calculated as 80% doorstep recruitment rate x 90% occupant recruitment rate x 90% occupant completion rate |
| Target fully completed households | 1,670 | |

## Further stratification by proximity to transport corridor

With respect to public transport access for major cities with local rail, a further level of stratification was applied to the sample to maintain time series. This ensured that annual NZHTS GPS7 and GPS2 samples would have approximately the same number of respondents living close to rail transport.

In the first year (2015/16), meshblocks in Auckland and Wellington were selected randomly and then flagged if they were in a railway corridor. This was based on the meshblock centroid distance to the closest railway station. The definition of "rail corridor" (3km or less for Wellington and 2km or less for Auckland) was based on the distribution of historical household travel survey distances travelled prior to a train trip. The proportion of meshblocks in a rail corridor in the subsequent annual samples was then fixed to the 2015/16 proportion by regarding Auckland and Wellington as having 2 substrata.

## Local authority expansions

Expansions are set up such that the expansion data is incorporated into the national dataset for reporting and analysis. Further information about the expansion samples is included in the Achieved Sample section.

4 local authorities have purchased additional expansions in their local areas:
- Auckland (2018/19)
- Tauranga (2018/19)
- Wellington (2019/20)
- Waikato (2019/20)

Unfortunately, the Wellington and Waikato expansions were curtailed by COVID-19.

## Survey time periods

Due to some initial issues, 2015/16 and 2016/17 surveying did not exactly align with the intended July through to June survey periods. Survey timing is approximately as follows:

Table 3: Survey time periods

| Survey year | Time period covered | Comments |
|---|---|---|
| 2015/16 | Oct 2015 – Oct 2016 | Surveying paused 16rd Dec – 8th Jan |
| 2016/17 | Nov 2016 – Jun 2017 | Surveying paused 23rd Dec – 2nd Jan |
| 2017/18 | Jul 2017 – Jun 2018 | Surveying paused 17th Dec – 4th Jan |
| 2018/19 | Jul 2018 – Jun 2019 | Surveying paused 23rd Dec – 3rd Jan |
| 2019/20 | Jul 2019 – Mar 2020 | Surveying paused 24th Dec – 5th Jan and was halted early in 2020, in response to COVID-19 restrictions. |

## Achieved sample

The response rate for 2018-2020 is described in the following table, in comparison with data from the 3 years of NZHTS GPS7, and the last three years of the previous survey.

Comparing historical response rates (2011-2014), to those achieved in NZHTS GPS7, we found a lower fully completed household rate (where every member of the household fully completed the 7 days of the travel diary) and a correspondingly higher partial completion rate (where at least 1 person in the household fully completed the 7 days of the travel diary).

Historically, 67 percent of eligible households fully completed while 2 percent of eligible household only partially completed.

In NZHTS GPS7, 39 percent of households fully completed the survey (43 percent 2015/16, 36 percent 2016/17, 38 percent in 2017/18) while 21 percent only partially completed (18 percent 2015/16, 25 percent 2016/17, 21 percent in 2017/18). This may have been a result of the longer travel survey period and increased respondent burden.

Once a household had been recruited by the interviewer, further sample loss may have occurred due to already recruited household occupants changing their minds or other household members no longer being willing or able to take part.

Of those eligible in recruited households, 61 percent completed the diary for 7 days of travel (66 percent 2015/16, 56 percent 2016/17, 61 percent in 2017/18); 33 percent changed their minds before the travel days started or did not agree to take part (26 percent 2015/16, 38 percent 2016/17, 35 percent in 2017/18); and the remainder did not complete the full course of the travel days. The greatest attrition rate occurred in the first few days.

Overall, 61 percent of the eligible households had one or more people in the household complete the new survey, compared to 69 percent of eligible households historically.

When comparing response rates of NZHTS GPS7 and NZHTS GPS2, we found a higher fully completed household rate (where every member of the household fully completed the 2 days of the travel diary) and a correspondingly lower partial completion rate (where at least 1 person in the household fully completed the 2 days of the travel diary).

In NZHTS GPS2, 64 percent of households fully completed the survey, while 5 percent only partially completed. The improvement in the fully complete rate is believed to be the result of lower respondent burden and the introduction of a second personal visit to the household by the interviewer.

Table 4: Household response rates.

| Household response status | 2 day diary, face to face interview | 7 day online travel diary, option of GPS tracker NZHTS GPS7 | | | 2 day diary, face to face interview, option of GPS tracker NZHTS GPS2 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2011-2014 (annual average) Years 10-12 | 2015/16 Year 14 | 2016/17 Year 15 | 2017/18 Year 16 | 2018/19 (total) Year 17 | Auckland expansion Year 17 | Tauranga expansion Year 17 | 2019/20 (total) Year 18 | Wellington expansion Year 18 | Waikato expansion Year 18 |
| **Eligible Households** | **4,746** | **3,530** | **3,758** | **3,296** | **3,991** | **645** | **915** | **3,516** | **739** | **592** |
| No contact with household | 222 | 407 | 414 | 406 | 351 | 80 | 74 | 255 | 76 | 65 |
| | 5% | 12% | 11% | 12% | 9% | 12% | 8% | 7% | 10% | 11% |
| Refused initial interview | 884 | 522 | 563 | 546 | 591 | 99 | 146 | 496 | 123 | 74 |
| | 19% | 15% | 15% | 17% | 15% | 15% | 16% | 14% | 17% | 13% |
| Language or health issues | 103 | 62 | 59 | 61 | 48 | 0 | 9 | 59 | 14 | 10 |
| | 2% | 2% | 2% | 2% | 1% | 0% | 1% | 2% | 2% | 2% |
| Partial recruitment – no-one fully responded | 274 | 377 | 418 | 339 | 259 | 15 | 101 | 195 | 38 | 25 |
| | 6% | 11% | 11% | 10% | 6% | 2% | 11% | 6% | 5% | 4% |
| Partial recruitment – 1 or more people fully responded | 106 | 631 | 952 | 686 | 189 | 28 | 39 | 317 | 76 | 82 |
| | 2% | 18% | 25% | 21% | 5% | 4% | 4% | 9% | 10% | 14% |
| **Full recruitment and response by all eligible people** | **3,157** | **1,531** | **1,352** | **1,258** | **2,553** | **423** | **546** | **2,194** | **412** | **336** |
| | **67%** | **43%** | **36%** | **38%** | **64%** | **66%** | **60%** | **62%** | **56%** | **57%** |
| Full response by at least one person | 3,263 | 2,162 | 2,304 | 1,944 | 2,742 | 451 | 585 | 2,511 | 488 | 418 |
| | 69% | 61% | 61% | 59% | 69% | 70% | 64% | 71% | 66% | 71% |

# Data collection for 2-day survey

Households were sent an introductory letter and information pamphlet before interviewers visited in-person to the house. Once contact was made, the interviewer made an attempt to recruit all eligible occupants who were present in the household at the time of their visit. The household members who had agreed to participate were asked to recruit other eligible household members (3rd party recruitment). Once a household was recruited, all occupant details were recorded, and the household was assigned 2 consecutive travel days. The start day for each recruited household was staggered to ensure an even distribution of start days within each geographical region.

All household members were left with a booklet containing instruction on how to take part, a two-day paper travel diary and a short exit questionnaire. Household members aged 12 and over were also given the option of using a GPS logger to record their travel. Where occupants were agreeable, the interviewer had the option of completing the exit questions for some or all of the occupants on the first visit. Alternatively, occupants could complete these in the booklet before the interviewer's return visit. Before the interviewer left, a date and time was agreed upon for the interviewer to return. Where possible, a time was agreed when all household occupants would be available.

Upon the interviewer's return visit, the interviewer reviewed the paper travel diary with each occupant that was present. If the occupant had also used a GPS logger, the GPS trace was displayed to the respondent as a prompt to ensure that any short or incidental trips were also recorded in the diary. Where no GPS data was available, interviewers were trained to probe for these types of trips.

Once all trips had been identified for an occupant, they were recorded in an electronic diary. This was completed from scratch for those with no GPS data. For those that had used the GPS logger, the trips were pre-populated in the electronic diary. The interviewer could then check and update the entries to ensure they aligned with the occupant's account of where they had been.

For the purposes of the survey, a trip is:
   a) any movement on a public street, footpath, railway line, etc. of **more than 100 metres.**
   b) any movement on a public street of less than 100 metres **if there is a change of purpose or a street is crossed.**
   c) any trips in a public aeroplane, train, ferry, water taxi.
   d) any trips on a private aeroplane, balloon, boat, that **do not finish where they started.**
   e) any trip not on public roads that does not finish where it started. This includes trips in a vehicle, on a bicycle or on foot.
   f) In outdoor malls, any time the shopper must walk along an outdoor footpath, across a street, car park or pedestrianised street (i.e. one that has been closed to traffic), then the above criteria apply.

A new stop is recorded if:
a) the respondent travels more than 100 metres on a public street or footpath, or
b) they cross a road used by any traffic, or
c) the purpose changes or
d) the mode changes.

In the event that an occupant was not available on the interviewer's return visit. The occupant could elect to have another member of the household provide their travel data to the interviewer on their behalf. Where this occurred, a flag was recorded in the dataset to indicate that the data had been provided by proxy. This approach was avoided wherever possible, with interviewers often arranging a third visit to collect the data directly from the occupant(s) concerned.

Once the interviewer had completed the digital diaries and exit questions for all participating occupants, all data was sent to a specialist coding team who were responsible for entering the trips into a dedicated website. The website contained the GPS traces where available and was programmed with a range of edit checks to alert the coder to any unusual trips (e.g. very fast/slow trips). If the coder was unsure about any of the information provided, they were able to contact the household for verification. In reality this occurred very infrequently.

Text message reminders were used for a variety of functions. For example, to remind participants to use and charge their device. Reminders were also sent if no motion was recorded for more than a day (checking to see if the person did not travel that day); or if the battery of the GPS unit was low/flat (battery levels were remotely monitored). Notification/reminder levels were tailored to suit respondent needs, to minimise intrusion.

## Data processing

GPS fixes (time stamped longitude and latitude) were constantly uploaded from the GPS device to a central secure server. A stop detection algorithm partitioned the GPS sequence of points into proposed trips. The GPS points along a trip were supplied to a Google Maps API which returned a trip speed and distance and rendered the trip on a map.

Coding staff were provided with a visual render of the GPS trips, along with the respondent's account of the trips they had made that day. The coder had the ability to delete incorrectly detected stops for example, being stopped outside a shop in a traffic jam. They could also add any missing trips, for example, when the respondent forgot to take their GPS logger or the GPS logger battery ran flat. The speed of trips for entered data were also calculated in real time using the Google Maps API.

The trip data entry screen included internal consistency checks and mode-specific travel speed checks. These took the form of a pop-up that asked the coding staff to confirm the data that had just been entered, allowing them to correct it.

The coding staff could choose to ignore a data check, so they could continue to code trips from the electronic diaries. The software records that there was an error with this trip and flag this for later investigation.

The most common data errors were speed errors that resulted from the manual entry of address data when a GPS tracker was not being used. If the address was not correctly geocoded, for example, entering the correct street name but in the wrong town, the trip

speed would exceed the test threshold. The subsequent trip would also usually have an excessive speed.

The household could be recontacted by telephone and coding checked and corrected if applicable.

Each quarter, the trips that failed any data checks were checked by an expert coder.

A final series of internal consistency checks and speed checks were also applied to the data each quarter.

# Weighting

Weights were applied to establish national-level results.

**Household dataset HH**

In NZHTS GPS7 and NZHTS GPS2, a random sample of 300 meshblocks were drawn from across NZ. There was no stratification by rurality.

Meshblocks were selected randomly (PPS) and within each meshblock, a household was selected randomly (with systematic sampling). The household sampling weight (hhsamwgt) was derived from these two components. The sum of household weights equals the number of houses in the sampling frame.

Not all houses provided a response. Within each meshblock, an adjustment was made to the household sampling weight so that the sum the weights of all responding houses was the same as the sum of all sampled houses. This is the household non-response weight (hhnrwgt).

The household non-response weights were then scaled too, so they summed to the estimated number of houses in New Zealand (hhweight).

The weights were then further stratified, so the sum of weights at the regional council level was equal to the estimated number of houses in the regional council (hhweight_poststrat). Nelson, Marlborough and Tasman were considered as a single unit.

**Person dataset PE**

The person weight (peweight) was constructed by restricting the dataset to all people that fully responded, ie completed 2 days of the travel diary, then post stratifying by Statistics NZ estimated of the number of people in each age / gender cell for each regional council.

The sum of peweight was the estimated number of people in each age / gender cell for each region and nationally. In Gisborne and West Coast some aggregation of age groups 0-15 and 15-40 was required.

As a further means of analysis, an additional person weight was generated for reporting on a national level, based only on the sample of fully responding people in households, where

every eligible member of the household completed the survey (peweight_fullhh). This weight was used for the December 2018 data release reports on the 2015-2018 data.[6]


**Trip dataset TR**

A weight (tripwgt_ann_mill) was then calculated for each trip so that the sum of the weights was the number of million trips per year.  An equivalent weight (tripwgt_ann_mill_fullhh) was calculated based on the person weight of those in fully responding households only. For 2015-2018, there was no adjustment for weekends ("day of week") as all people contributing to this dataset completed 7 days of travel diary.

With the shift to  a 2 day travel diary from 2018 onwards, day of week became more of an issue.

Originally, the trip days were planned to be evenly spread throughout the week except for the expansion Auckland and Tauranga samples.  In these households, the first day of trip reporting would exclude Saturday. Instead, Saturday travel patterns would be obtained from people whose 2-day diary period started on a Friday.  Similarly, Sunday travel patterns would be obtained from those whose first day of interviewing was a Sunday.  The trip data would be reweighted appropriately.

In practice, some interviews commenced on a Saturday in both areas, and the interviewer start day also varied by day in other regions.   Generally, more interviews commenced on a Sunday and less on a Saturday.  While fewer interviewers commenced on a Saturday throughout the country more interviews commenced on a Friday or a Sunday somewhat mitigating that effect.

For this reason, the data was reweighted by calendar day of travel.  Each person had two days of travel and received a different trip weight for each travel day.  The population counts were benchmarked to population values for each day by broad age group, broad region and Māori/non-Māori.

---

[6] https://www.transport.govt.nz/statistics-and-insights/household-travel/key-facts/

## Table 5: Weighting summary

| Level of sampling | Time period | | Numerator | Denominator | Level applied | NZHTS 2015-16 variable | Notes |
|---|---|---|---|---|---|---|---|
| **Meshblock (MB)** | | Sampling | 1 | Probabilty of selection | meshblock | **MB_weight** | sums to number of meshblocks in NZ in frame |
| **Household (hh)** | | Sampling | Number of households enumerated in meshblock | Number of households sampled | meshblock | **hhsamwgt** | sum weights for selected houses comes to total houses in NZ |
| | | Non-response | Summed weight of all eligible households | Summed weight of all responding households | hh non-response groups | **hhnrwgt** | use meshblock level response rate to scale sampling weight up for responding houses |
| | | National households | National estimate of household numbers | Summed weight of all responding households | national | **hhweight** | sum hhweight = national estimates of household numbers |
| | | Post-stratification by region | Predicted number of houses in region | Summed weight of all responding households | region | **hhweight_poststrat** | rescales the above so weights sum to number of houses in region and NZ |
| **Person (pe)** | | Post-stratification by region/age/sex | Total estimated of resident population in cell | Summed weight of all fully responding people | age/sex groups within region | **peweight** | in each region / age / sex group weights sum to Statistics NZ estimates |
| | | Post-stratification by region/age/sex | Total estimated of resident population in cell | Summed weight of all fully responding people in fully responding houses | age/sex groups within region | **peweight_fullhh** | in each region / age / sex group weights sum to Statistics NZ estimates |
| **Trip (tr)** | 2015-2018 | Post-stratification | 365/7 * nzpop/people *1/1000000 | | trip | **trwgt_ann_mill** | simple multiple of peweight to give million trips per annum |
| | 2015-2018 | Post-stratification | 365/7 * nzpop/people *1/1000000 | | trip | **trwgt_ann_mill_fullhh** | simple multiple of peweight_fullhh to give million trips per annum |
| | 2018/19 | Post-stratification | | | trip | **trwgt_ann_mill** | Adjusted by peweight and to account for day of week to give million trips per annum |
| | 2018/19 | Post-stratification | | | trip | **trwgt_ann_mill_fullhh** | Adjusted by peweight_fullhh and to account for day of week to give million trips per annum |

# Notes and issues

**Note on continuity with historical data**

Due to the large respondent burden in the 7-day self-complete diary some modification in methodology will be required to account for non-response bias and differential response to different survey methods, to provide comparability with historical datasets. This means that as of publication date (November 2019), time-series comparison of the 2015-2018 data to the historical data is not possible.

**Reception to new survey**

People that completed the new form of the travel survey (NZHTS GPS7 and GPS2) regarded the survey experience positively, with 80% saying they would do the travel survey again. Over 75% of people agreed to carry a GPS tracker device, and of these people, 96% subsequently used the device on any given day.

The proportion of households that agreed to take part in NZHTS GPS7 was similar to earlier surveys (around 70%) but the proportion of houses in which all occupants completed 7 days data collection and coding was considerably lower than the proportion of houses that completed 2 days data collection in the earlier survey. The result is that NZHTS GPS7 had a lower rate of "fully completed" houses. This trend has been reversed in the NZHTS GPS2, with x percent of households fully completing the survey.

A soft form of non-response occurs when respondents claim they did not travel on a given day, rather than formally withdrawing. A certain amount of non-travel was to have been expected, but as can be seen below, rates of non-travel increase as a function of travel day in NZHTS GPS7.

In the earlier survey (years 2011-2014), there was an increase between days 1 and 2, and this was also apparent in the 7-day survey. There was an increase from about 1 in 4 or 5 people not travelling at the start of the travel days, to nearly 1 in 3 people not travelling by the end of the travel days. As the start of the travel days was assigned to different days of the week for each household, this should not be a function of the day of the week travelled. With the survey reverting back to a two-day collection period, the rate of non-travel reported on days one and two is more in line with the historical survey.

Figure 1: Percentage of people who did not travel on given survey day, by survey year.
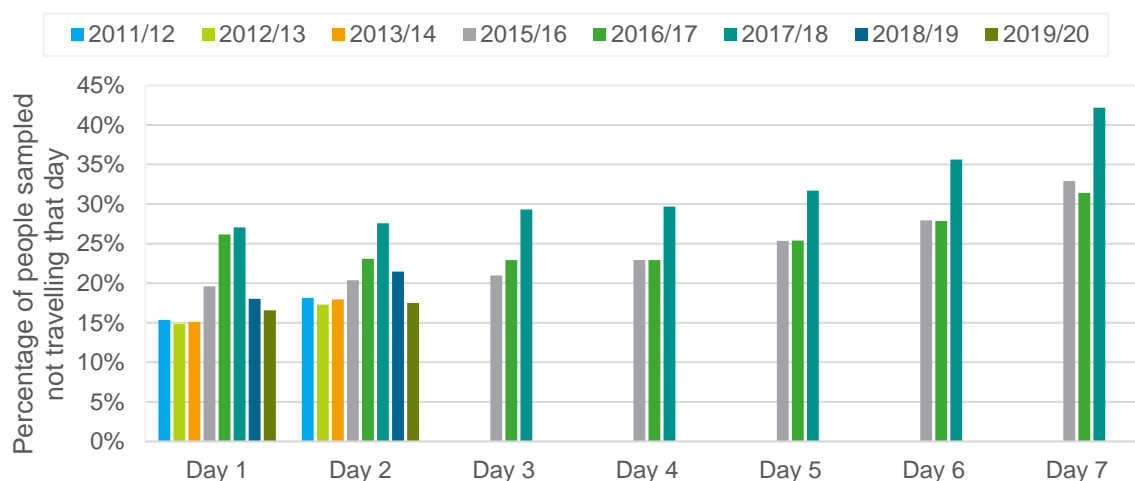
## Allowing for reducing trip counts over 7 days

In NZHTS GPS7, fewer trips were recorded by respondents towards the end of the 7-day collection period, so it is likely that trip numbers were slightly underestimated. Some controlling for dropout could be undertaken by scaling up estimates of trip numbers by a factor that is a function of day of survey and day of week.

## COVID-19 interruption and effect

Face-to-face surveying for the 2019/20 survey year was halted early, on 20 March 2020, in response to the Government's actions in response to COVID-19[7]. Households which had already been recruited to take part, but not yet been interviewed over their travel were followed up over the phone. Face-to-face recruiting and surveying was not able to commence again until August 2020 (the 2020/21 survey year), where upon procedures and protocols for established for interviewing at COVID-19 Alert Levels 2 and below.

Unfortunately, this meant a reduced sample of households for the 2019/20 survey year. Due to issues with interviewer availability, this reduction was not evenly spread through the regions, and certain regions were more heavily impacted. Southland was not sampled at all and Northland and Otago were heavily impacted.

In order to be able to report nationally, a synthetic Southland sample was imputed based on historically sampled Southland households. These have been specifically flagged in the datasets.

Weights have been set up to give an annual average travel estimate based on pre COVID-19 lockdown sampling. This means absolute annual average amounts of travel will be overestimated.

## Regional estimates.

Regional estimates rely on Stats NZ estimates of the numbers of households. More accurate household counts could have been used if they had been available.

---

[7] For more information about the COVID-19 pandemic and New Zealand response, please see https://en.wikipedia.org/wiki/COVID-19_pandemic_in_New_Zealand (accessed 10 May 2021)

**Size of regions**

Some aggregations have been undertaken: Nelson/Marlborough/Tasman have always been considered as a single group, and to perform age x gender x region post stratification, some age groups were merged in Gisborne and West Coast.

**Adjusting for person non-response at household level.**

In previous surveys all household members were expected to take part for two days. While there was some within-household non-response in earlier surveys, in the NZHTS GPS7 survey about 30% of occupants declined in houses where at least one person had agreed to take part. Of those occupants that started the survey, some did not complete all 7 days data collection.

All estimates were post stratified to age and gender distributions but there is an argument for weighting for household occupant non-response, on the assumption that a non-responding occupant is more like a responding occupant from the same house than another person of their age and gender.

**Sample composition by age and gender**

The following table and graph compare the sample composition in the historical survey (2011-2014) with NZHTS GPS7 (2015-2018), NZHTS GPS2 (2018-20) and with the current NZ population based on the 2018 Census. The 2015-2018 sample under-represents younger people (under 30 years old) and over-represents people 55 and older. The 2018-20 sample shows a similar skew, but is trending closer to the census distribution for those age groups. Both survey periods under-sampled people 20-24 years olds.

## Table 6: Composition of sample

| Age (years) | 2011-2014 sample | | 2015-2018 sample (NZHTS GPS7) | | 2018-2020 sample (NZHTS GPS2) | | NZ population (2018 Census) | |
|---|---|---|---|---|---|---|---|---|
| | Female | Male | Female | Male | Female | Male | Female | Male |
| 0-4 | 3.5% | 3.9% | 1.9% | 2.0% | 2.3% | 2.6% | 3.1% | 3.2% |
| 5-9 | 3.4% | 3.7% | 2.1% | 2.2% | 2.5% | 2.9% | 3.3% | 3.5% |
| 10-14 | 3.6% | 3.8% | 2.2% | 2.3% | 2.7% | 3.0% | 3.2% | 3.3% |
| 15-19 | 3.1% | 3.4% | 2.3% | 2.1% | 2.7% | 2.5% | 3.1% | 3.3% |
| 20-24 | 2.7% | 2.5% | 2.3% | 2.0% | 2.1% | 2.1% | 3.3% | 3.5% |
| 25-29 | 2.9% | 2.3% | 2.8% | 2.5% | 3.1% | 2.3% | 3.6% | 3.7% |
| 30-34 | 3.0% | 2.8% | 3.1% | 2.6% | 3.6% | 3.0% | 3.4% | 3.3% |
| 35-39 | 3.7% | 2.9% | 3.1% | 3.0% | 3.5% | 3.3% | 3.2% | 3.1% |
| 40-44 | 3.9% | 3.4% | 3.4% | 3.0% | 3.5% | 3.3% | 3.2% | 3.0% |
| 45-49 | 3.9% | 3.6% | 3.7% | 3.6% | 4.0% | 3.5% | 3.5% | 3.3% |
| 50-54 | 3.8% | 3.2% | 4.0% | 3.8% | 3.7% | 3.4% | 3.4% | 3.2% |
| 55-59 | 3.1% | 2.9% | 4.2% | 3.8% | 3.8% | 3.2% | 3.3% | 3.1% |
| 60-64 | 3.1% | 2.8% | 3.9% | 3.7% | 3.7% | 3.4% | 2.8% | 2.7% |
| 65-69 | 2.5% | 2.4% | 4.6% | 3.9% | 3.4% | 3.3% | 2.5% | 2.4% |
| 70-74 | 2.0% | 1.7% | 3.5% | 2.9% | 2.7% | 2.6% | 2.0% | 1.9% |
| 75-79 | 1.5% | 1.4% | 2.7% | 2.3% | 1.9% | 1.9% | 1.5% | 1.3% |
| 80-84 | 1.0% | 0.9% | 1.4% | 1.4% | 1.3% | 1.2% | 1.0% | 0.8% |
| 85+ | 0.8% | 0.6% | 1.0% | 1.0% | 1.0% | 0.8% | 1.1% | 0.7% |
| Total | 51.6% | 48.4% | 52.1% | 47.9% | 51.5% | 48.5% | 50.6% | 49.4% |

## Figure 2: Composition of sample compared to population