

# **SoFIE-Health Data and Processing Systems**

## **SoFIE-Health Report 1**

*Version 1.0*

**Authors:**

**Ken Richardson  
Kristie Carter  
Michael Hayward**

**April 2008**

**ISBN 978-0-473-13474-7**

## ***Acknowledgements***

SoFIE-Health is primarily funded by the Health Research Council of New Zealand as part of the Health Inequalities Research Programme. Establishment funding was also received from the University of Otago, Accident Compensation Corporation of New Zealand (ACC), and the Alcohol Liquor Advisory Council (ALAC). Comments on this report were received from Tony Blakely and Statistics New Zealand.

## ***Statistics New Zealand Security Statement***

Access to the data used in this study was provided by Statistics New Zealand in a secure environment designed to give effect to the confidentiality provisions of the Statistics Act, 1975. The results in this study and any errors contained therein are those of the authors, not Statistics New Zealand.

## ***Copyright***

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopy, recording, or otherwise, without the permission of the authors.

## ***Reference***

Richardson, K., Carter, K., Hayward, M. (2008). SoFIE-Health Data and Processing Systems. SoFIE-Health Report 1. Wellington: University of Otago, Wellington.

Published by the Department of Public Health  
University of Otago, Wellington  
PO Box 7343  
Wellington South 6242  
New Zealand  
ISBN 978-0-473-13474-7

## Executive Summary

Statistics New Zealand (Statistics NZ) is running an 8-year longitudinal study of income, labour force participation, educational training, asset accumulation and family structure among 22,000 randomly selected adults called the Survey of Family, Income and Employment (SoFIE). SoFIE-Health is an extension to include 20 minutes of questions in years 3 (2004-05), 5 and 7. SoFIE-Health data has also been linked to hospitalisation, cancer registration data – and in the future will update these linkages to include mortality data.

After confidentialisation, SoFIE-Health wave 1-3 microdata were made available to researchers from the University of Otago, Wellington (UOW) as more than 70 SAS data files in the Statistics NZ Data Laboratory. The SoFIE-Health project team has designed more compact SAS data structures and associated processing code to provide SoFIE-Health researchers easier access to the complex and evolving Statistics NZ datasets.

This report has two general objectives. The first is to document the data structures and processing systems for the SoFIE-Health project. The second is to improve access to and the quality of SoFIE data over time.

Specific objectives include:

- To provide comprehensive system documentation to staff from the University of Otago, Wellington who manage SoFIE-Health data and processing systems.
- To collect into one document information about the SoFIE dataset and processing system deemed useful for SoFIE-Health research activities.
- To assist Statistics New Zealand maximise access to the SoFIE dataset.
- To assist Statistics New Zealand maximise the quality of the SoFIE dataset by providing a quality audit of selected SoFIE-Health variables.

This is the first version of this report. In subsequent versions, emphasis will shift to the first and second bullet points above as SoFIE-Health datasets evolve and become more stable. Note that the SoFIE-Health Technical report is, like SoFIE-Health itself, ‘work-in-progress’. More information and updates to information already included will be added to later versions.

# Table of Contents

Acknowledgements.....	- 1 -
Statistics New Zealand Security Statement.....	- 1 -
Copyright .....	- 1 -
Reference.....	- 1 -
<b>EXECUTIVE SUMMARY .....</b>	<b>- 2 -</b>
<b>TABLE OF CONTENTS .....</b>	<b>- 3 -</b>
<b>TABLE OF FIGURES .....</b>	<b>- 6 -</b>
<b>1 INTRODUCTION TO THE DOCUMENTATION .....</b>	<b>1</b>
1.1 Definitions.....	1
1.2 Getting more information.....	1
<b>2 OVERVIEW OF SOFIE.....</b>	<b>2</b>
2.1 History/development.....	2
2.1.1 SoFIE.....	2
2.1.2 SoFIE-Health.....	3
2.2 The SoFIE sample and following rules.....	4
2.2.1 Target population .....	4
2.2.2 Sampling frame .....	4
2.2.3 Sample selection.....	5
2.2.4 Sample Size .....	5
2.3 Data collection .....	5
2.3.1 Overview .....	5
2.3.2 Contact procedures .....	6
2.3.3 Interview period .....	6
2.3.4 Reference period.....	6
2.3.5 Scope and Coverage Rules .....	7
2.4 Survey instruments and Questionnaires .....	7
2.5 Confidentiality requirements .....	8
<b>3 SOFIE DATA AND SOFTWARE.....</b>	<b>9</b>
3.1 General Information.....	9

<b>3.2</b>	<b>Outline of Data Organisation and Processing Software .....</b>	<b>9</b>
3.2.1	Overview .....	9
3.2.2	Data organisation.....	10
3.2.3	Data processing and DataLab directory structure.....	14
3.2.3.1	SoFIE-Health DataLab directory structure .....	14
3.2.3.2	Code maintenance.....	14
3.2.3.3	Compiling and storing SAS macros.....	16
3.2.3.4	SAS data formats .....	16
3.2.3.5	User access to SoFIE-Health data and software.....	16
3.2.3.6	SoFIE-Health data processing.....	17
3.2.3.6.1	merge_setup.sas .....	17
3.2.3.6.2	merge_datasets.sas .....	17
3.2.3.6.3	check_merge_datasets.sas.....	18
3.2.3.6.4	run_dropvars.sas .....	18
3.2.3.6.5	create_user_dataset.sas.....	18
3.2.3.6.6	eqinc.sas .....	19
3.2.3.6.7	score_sf36.sas .....	19
3.2.3.6.8	drv_chron.sas .....	20
3.2.3.6.9	drv_nzidep.sas.....	20
3.2.3.6.10	drv_smoke_alc.sas.....	21
3.2.3.6.11	score_k10_PSS.sas .....	22
3.2.3.6.12	wealth.sas .....	22
3.2.3.6.13	maxqualrank_wsm.sas.....	23
3.2.3.6.14	merge_drv_vars.sas .....	23
3.2.3.6.15	create_research_dataset.sas .....	23
3.2.3.6.16	incwlth_pctiles.sas.....	23
3.2.3.6.17	rank_incwlth.sas .....	23
3.2.4	Variables.....	24
3.2.4.1	Naming conventions .....	24
3.2.4.2	Variable and data dictionary' .....	24
3.2.5	Imputation and variance estimation of missing data by Statistics New Zealand.....	24
3.2.5.1	Imputation for item non-response.....	24
3.2.5.2	Longitudinal imputation for wave non-response .....	25
3.2.5.3	Imputation for complete family/household wave non-response <sup>24</sup> .....	25
3.2.5.4	Types of data available for use in imputation .....	25
3.2.5.5	Imputation methodology.....	26
3.2.5.6	Implementation issues.....	26
3.2.5.7	Variance estimation (production of longitudinal standard errors) .....	27
3.2.6	Weight estimation by Statistics New Zealand .....	27
3.2.6.1	Longitudinal Weighting.....	28
3.2.6.2	Cross-sectional Weighting .....	28
3.2.7	Missing values and attrition.....	29
3.2.8	Future work .....	30
<b>4</b>	<b>MACROS AND OTHER USEFUL PROGRAMS .....</b>	<b>30</b>
<b>4.1</b>	<b>begin_sofie.sas .....</b>	<b>30</b>
<b>4.2</b>	<b>%sofsess.sas .....</b>	<b>30</b>
<b>4.3</b>	<b>%soffmts.sas .....</b>	<b>31</b>
<b>4.4</b>	<b>%fmt_sof_vars.sas .....</b>	<b>31</b>

<b>5</b>	<b>TESTING SOFIE WAVE 1-3 DATA.....</b>	<b>32</b>
5.1	Discrete variables .....	32
5.2	Continuous variables .....	33
5.2.1	PersIncAnn(DrvFamily).....	33
<b>6</b>	<b>ADDITIONAL TASKS PLANNED FOR FUTURE RELEASES OF THE SOFIE-HEALTH TECHNICAL REPORT .....</b>	<b>33</b>
<b>7</b>	<b>APPENDICES .....</b>	<b>35</b>
7.1	SoFIE-Health variable SAS formats.....	35
7.2	Statistics NZ SoFIE Health Questionnaire flowchart, pages 1-5 .....	54
7.3	SoFIE data dictionary waves 1-2 .....	59
7.4	SoFIE-Health Data Laboratory directory listing .....	95

## Table of Figures

Figure 1: Domains of SoFIE-Health.....	3
Figure 2: SoFIE reference periods. ....	6
Figure 3: Confidentialised view of the final long research dataset.....	12
Figure 4: Confidentialised view of the final wide research dataset. ....	13
Figure 5 Screen view of the UOW directory structure in the Data Laboratory..	15

# 1 Introduction to the documentation

This report contains the following information about the SoFIE-Health project:

- Definitions
- Background information
- Data organisation and software
- Data testing
- Additional work proposed for future releases of the Technical Report
- Appendices
  - SAS formats for SoFIE-Health variables
  - Example questionnaire flowchart
  - SoFIE data dictionary
  - SoFIE-Health Data Laboratory directory listing

## 1.1 Definitions

*OSM - Original Sample Member.* At the first wave of the SoFIE, every eligible responding resident of the selected households became an OSM. These are the people who are interviewed over successive waves of the survey, regardless of whether or not they still reside in their original dwelling.

*PQ - Personal Questionnaire.* The individual questionnaire filled out by survey respondents.

*HQ - Household Questionnaire.* The questionnaire filled out by a nominated individual in the household.

*HED - Household Enumeration Date.* The date when all of the eligible members of the household are identified and recorded in the household questionnaire.

*SAPD - Start of Annual Period Date* is the reference date for the start of the annual period e.g. if a respondent's Wave 1 HED interview is on 01 December 2002, the SAPD is 01 December 2001 (12 months before the HED).

*EAPD - End of Annual Period Date* is the reference date for the end of the annual period for each respondent e.g. if a respondent's Wave 1 HED interview is on 01 December 2002, the EAPD is 30 November 2002.

*PSU - Primary Sampling Unit.* The geographical units or groups of units considered for inclusion in the first stage of sample selection.

*CAI - Computer-Assisted Interviewing.* An interviewing technique that uses a computer to assist with data collection (usually face-to-face or self-interview).

*Cohabitant.* Someone living with an OSM, who is asked the PQ.

*Point-in-time data.* Respondent data collected at a specific date (usually the interview date).

*Spell data.* Respondent data relating to a period of time with a defined start and end date.

## 1.2 Getting more information

Further information about the SoFIE study and links to other longitudinal studies can be found at <http://www.stats.govt.nz/datasets/social-themes/Survey-of-family-income-and-employment.htm>.

Information about the SoFIE-Health project is available at <http://www.wnmeds.ac.nz/academic/dph/research/HIRP/SoFIE/index.html> or by emailing the Principal Investigator (Kristie Carter) at [kristie.carter@otago.ac.nz](mailto:kristie.carter@otago.ac.nz).

## **2 Overview of SoFIE**

### **2.1 History/development**

#### **2.1.1 SoFIE**

Statistics New Zealand was granted funding from the Foundation for Research, Science and Technology in 1997 to conduct a feasibility study for a longitudinal survey of income, employment and family dynamics. Following on from the feasibility study, the SoFIE study was developed and first went into the field in October 2002. SoFIE is the largest longitudinal survey ever run in New Zealand.

SoFIE is a single fixed panel longitudinal survey with an expected maximum duration of eight years. It will collect information once a year from the same individuals on income levels, sources and changes; and on the major influences on income such as employment and education experiences, household and family status and changes, demographic factors and health status. Every two years (waves 2, 4, and 6) it will also collect information on assets and liabilities to monitor net worth and savings. Every other year (waves 3, 5, and 7) it will collect detailed information on health.

The overall objective of SoFIE is to provide information about changes over time in the economic well-being of individuals and their families, and about factors influencing those changes.

The longitudinal objectives of the SoFIE are to:

- identify the pattern of income level experience over time for individuals and their families
- measure significant shifts in income levels for individuals and their families and explore the relationship of these shifts to labour market activity, receipt of government income support, and family status
- determine patterns over time of labour market activity, participation in education and training, and receipt of government income support
- identify transitions between spells of labour market activity, education and training participation and receipt of government income support, and examine the factors that influence transitions from one status to another.
- determine patterns of saving for retirement and relate these to income dynamics and life cycle stages
- determine patterns of change in family status.

The cross-sectional objectives of the SoFIE, are of lesser priority and concern the measurement and estimation of distributions and relationships at points in time for:

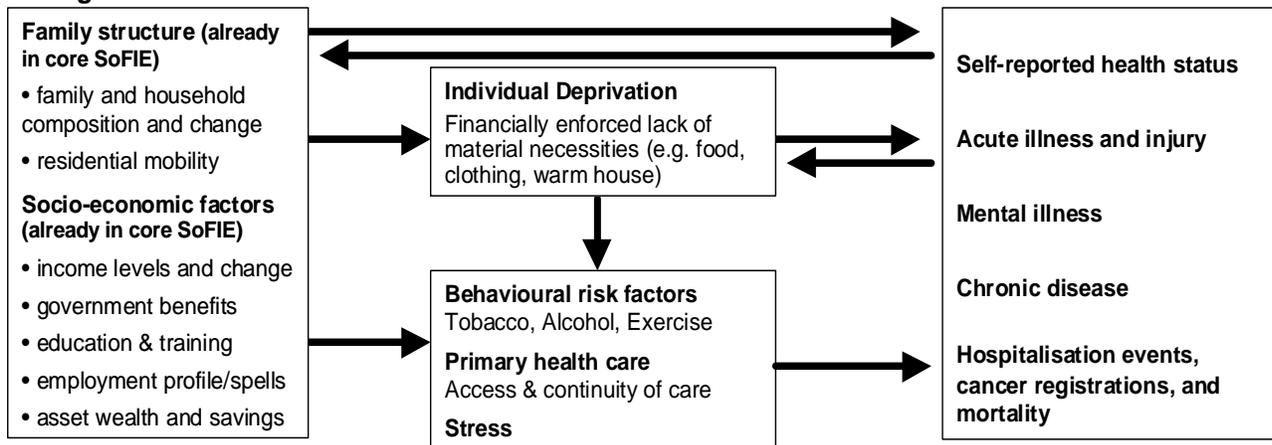
- individual, family and household income

- family and household composition
- labour market activity
- assets and liabilities

### 2.1.2 SoFIE-Health

In 2003 researchers from the Department of Public Health, UOW, applied for funding from the Health Research Council of New Zealand (HRC) to add a health module to the core SoFIE questionnaires to be asked every alternative year. The SoFIE-Health add-on consists of 20 minutes of questionnaire time in waves 3 (2004-05), 5 (2006-07) and 7 (2008-09), in the following health-related domains: health status (SF36 & Kessler scale), perceived stress, chronic conditions (heart disease, diabetes, and injury-related disability), tobacco smoking, alcohol consumption, health care utilisation, and access and continuity of primary health care, and an individual deprivation score.

**Figure 1: Domains of SoFIE-Health**



In Figure 1 the information on family structure and socio-economic factors is placed on the left of the diagram, and constitutes the ‘upstream determinants’ of health already collected by the core of SoFIE. The mid-stream determinants (data collected by SoFIE-Health) are individual deprivation, behavioural risk factors, primary care and stress. The right-hand box includes measures of self-reported health status and disease outcomes. The arrows identify likely causal relationships. Note that these domains cover in some way all four possible intervention points to reduce social inequalities in health identified by the Ministry of Health strategy for reducing inequalities in health. Further discussion of these issues can be found in Carter *et al.*, 2008<sup>1</sup>.

The three major HRC goals of SoFIE-Health for 2004-2008 are:

**Goal 1: Determining the impact of labour market factors, asset wealth, income and family dynamics on health**

Objective 1: Do labour market factors such as hours of work and job separations influence subsequent health status?

Objective 2: Does change in income predict change in self-rated health?

<sup>1</sup> Carter *et al.*, 2008. SoFIE-Health Baseline Report. University of Otago, Wellington.

Objective 3: How do labour market factors, family structure, asset wealth and income trajectories relate to health risk factors and health status?

**Goal 2: Determining the impact of health status on labour market factors, income trajectories, asset wealth and family dynamics**

Objective 4: Does pre-existing health status predict labour market, income and family mobility?

Objective 5: Do sex, ethnicity, and socio-economic position buffer the impact of pre-existing health status on subsequent labour market, income and family mobility?

Objective 6: Does pre-existing health status predict job separations for respondent-reported reasons other than “health status”?

**Goal 3: Determining the contribution of access, continuity and co-ordination of primary health care to health status and to social inequalities in health**

Objective 7: Are access to, continuity and coordination of primary care associated with better health?

Objective 8: Do access, continuity and coordination of primary health care explain some of the associations of ethnicity, family structure, and socio-economic factors with health?

## **2.2 The SoFIE sample and following rules<sup>2</sup>**

### **2.2.1 Target population**

The target longitudinal population for SoFIE is:

*the usually resident population of New Zealand living in permanent, private dwellings on the main islands in the North and South Islands, including Waiheke Island as at the first wave of the panel.*

The survey excludes overseas visitors resident in New Zealand for less than 12 months and who intend to stay in New Zealand for less than 12 months; non-New Zealand diplomats and diplomatic staff and their dependants; members of non-New Zealand armed forces stationed in New Zealand and their dependants; and people living in institutions or in other non-private dwelling establishments such as boarding houses, hotels, motels and hostels, as well as people living on offshore islands (excluding Waiheke Island). The cross-sectional population for a particular wave, wave  $t$ , can therefore be defined as: the usually resident New Zealand population living in private dwellings, as at wave  $t$  of the panel.

### **2.2.2 Sampling frame**

SoFIE used the standard Statistics New Zealand sampling frame used for other Statistics New Zealand household surveys. This frame partitions the North Island, South Island and Waiheke Island into approximately 19,000 small geographical areas known as Primary Sampling Units (PSUs). On average, a PSU contains around 70 dwellings and about 150-200 people, but can range in size from 30 to 260 dwellings. The only areas not included

---

<sup>2</sup> This section is summarised primarily from information on the SoFIE website <http://www.stats.govt.nz/datasets/social-themes/Survey-of-family-income-and-employment.htm> and other Statistics NZ SoFIE documentation

in this sampling frame are offshore islands (except Waiheke Island) and very remote areas of New Zealand.

### **2.2.3 Sample selection**

The selection of the random sample for SoFIE was a three-stage process. Firstly, PSUs were assigned to groups (strata) according to region, urban/rural, high/low Māori population density and other socio-economic variables derived from the most recent census. Then systematic sampling was used to select a sample of PSUs independently from each stratum.

The next stage of sampling involved taking a systematic random sample of permanent private dwellings within the PSUs selected. All of the eligible residents of each selected household who agreed to participate were then included in the sample. These people were designated as OSMs and will be followed and interviewed over time. The sample comprised 1,500 primary sampling units (PSUs), with an average 7.7 full and partially responding households obtained per PSU.

### **2.2.4 Sample Size**

For wave 1 of SoFIE a total of 15,000 randomly-selected households were approached, of which approximately 11,500 agreed to be interviewed (response rate of 77%), with data collected from over 22,000 individuals aged 15 and over. All responding individuals who were interviewed in Wave 1 were asked to become Original Sample Members (OSMs). Children under the age of 15 at Wave 1 are interviewed as OSMs from the first wave after they turn 15. All OSMs are re-interviewed in subsequent years, regardless of changes in their place of residence. Other members of an OSM's household who are not OSMs ('non-OSMs' or cohabitants) are interviewed from Wave 2 onwards while they remain living with an OSM. Non-OSMs are asked a reduced set of questions and are not followed up if they leave the OSM's household.

## **2.3 Data collection**

The following information is summarised from the Statistics NZ SoFIE Field Manual<sup>3</sup>.

### **2.3.1 Overview**

Respondents are interviewed eight times, once every year for eight years. The first time respondents were interviewed was known as Wave 1, and the second time as Wave 2, and so on. Each SoFIE Wave begins on the 1<sup>st</sup> of October each year and will continue for 12 months

On average, the Household questionnaire takes 10 minutes and the Personal questionnaire takes 40 minutes (per person).

If a person selected at Wave 1 refused to answer or was a non-contact at Wave 1, they are interviewed in future waves only if they are still living with an OSM i.e., they are treated as a cohabitant rather than an OSM. This means they will **not** be followed if they move.

---

<sup>3</sup> Survey of Family, Income and Employment Field Manual, October 2004. Statistics New Zealand

### 2.3.2 Contact procedures

One of the biggest problems in a longitudinal survey is attrition, which is the loss of respondents over time. Some attrition is unavoidable, such as people leaving the country or dying. During the course of the survey, respondents shifting house may also cause attrition. The respondent management system is designed to minimise attrition when people shift house in New Zealand.

The detailed procedures for making initial contact with respondents and completing the Household questionnaire, including guidelines for number of contact phone calls, household visits, etc, are outlined in the SoFIE Field Manual.

A household is eligible for the survey if it has been established that there is at least one OSM living there. The Household questionnaire is completed by an adult OSM at the household and **not** by a cohabitant or a child.

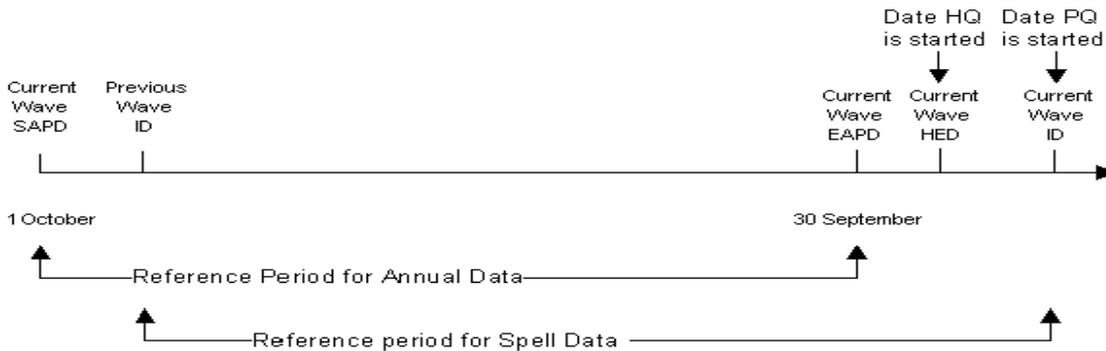
### 2.3.3 Interview period

At each wave, SoFIE interviewers attempt to find the people that were interviewed in Wave 1 (OSMs). During the three month interview period, the first (allocation) month is used to contact and interview OSMs, with an additional 2 months allowed to complete personal interviews in the household if this is not possible in the allocation month. This allowance is made because some people are away from home for extended periods of time or are difficult to get hold of.

### 2.3.4 Reference period

The SoFIE questionnaire collects point in time data (e.g., current health status) and spell data (e.g., the period of time spent in a job).

For spell data, different reference periods are used depending on the topic. For example, Labour Market spell data is collected from the previous interview date up until the current interview date. Income data for the self-employed is collected over a 12 month year with a pre-defined start of annual period date (SAPD) and end of annual period date (EAPD).



**Figure 2:** SoFIE reference periods.

Each household’s annual period is determined by their allocation month. The electronic questionnaire automatically sets these dates. Figure 2 illustrates the reference periods and dates used, including interview date (ID) and household enumeration date (HED).

### 2.3.5 Scope and Coverage Rules

Because this is a longitudinal survey, SoFIE is interested in all OSMs regardless of where they live after wave 1. Theoretically, all OSMs are 'in scope' for each wave. Only OSMs who live in either a private or non-private dwelling are interviewed. OSMs who live in an institution are not interviewed. See Appendix 2 of the SoFIE Field Manual for more information on non-private dwellings and Institutions.

Some people who live with an OSM after wave 1 are also interviewed:

- If an OSM lives in a *private* dwelling, all the people they live with (as defined by the person who completes the Household questionnaire) are eligible to be selected and interviewed.
- If an OSM lives in a *non-private* dwelling (e.g., a student in a hostel, an elderly person in a rest home), it is only the family members that they live with (partner/parent/children) who are eligible to be selected and interviewed.

Cohabitants who are “out of scope” are:

- overseas visitors who intend to stay in New Zealand for less than 12 months,
- foreign diplomats and diplomatic staff, and their families,
- members of non-New Zealand armed forces.

## 2.4 Survey instruments and Questionnaires

SoFIE is conducted using computer-assisted interviewing (CAI). In the field, interviewers use laptop computers to administer an electronic questionnaire (EQ), face-to-face, in respondents' homes. The EQ is organised into modules (a module is a series of questions on a particular topic).

There are two sets of questionnaires to SoFIE. For every household selected, a Household Questionnaire is firstly administered to a single adult OSM, which includes questions on the characteristics of the entire household, and then a Personal Questionnaire which was administered to every person aged 15 and over. In general, each adult answers their own Personal Questionnaire. However, in some cases a proxy is allowed to fill in the Personal Questionnaire for someone else, for reasons such as disability or language difficulties. Children (aged 14 or younger at the household enumeration date) do not fill out a Personal Questionnaire. Instead, the parent/caregiver nominated to answer for the child(ren) in the Household Questionnaire is asked a child module for each child as part of their Personal Questionnaire. Once a child turns 15, they answer the full PQ.

Copies of the questionnaires can be found on the SoFIE website: [www.stats.govt.nz/sofie](http://www.stats.govt.nz/sofie). A subset of the health questionnaire flowchart is shown in Section 7.2 of the Appendices (page 54).

The HQ is answered by one OSM in each household. It consists of two modules, the first of which is used to determine family type and household composition at the time of the interview. The modules are:

1. Household
2. Standard of living

The PQ is administered to all OSMs and cohabitants aged 15 and over, and consists of eight core modules:

1. Demographics
2. Child (if any)
3. Labour force involvement history
4. Education
5. Family
6. Current labour force involvement
7. Income
8. Contact

Depending on the module, the SoFIE questionnaire collects both point-in-time data and spell data. Point-in-time data relates to a single date, usually the interview date (e.g. the respondent's educational qualifications as at the interview date). Spell data relates to a period of time with a defined start and end date reported by the respondent (e.g. the period of time a respondent lived with a family member, or the length of time a person worked for a particular employer). For example, the start of a new household spell occurs when a person enters or leaves a household. The analysis in each release includes incomplete spells, i.e. spells that were ongoing at the time of the interview.

The first wave (or interview cycle) of SoFIE was conducted from 1 October 2002 to 30 September 2003 and sought responses to these modules: Household questionnaire; Non-contact short form; Demographics; Child; Labour market history; Education; Family; Labour market; Income; Contact; Standard series modules.

The second wave (or interview cycle) of SoFIE was conducted from 1 October 2003 to 30 September 2004 and sought responses to these modules: Household questionnaire; Non-contact short form; Demographics; Child; Labour market history; Education; Family; Labour market; Income; Contact; Standard series, Asset and liability modules. These modules are repeated in Waves 4 (1 October 2005 to 30 September 2006), 6 (1 October 2007 to 30 September 2008), and 8 (1 October 2009 to 30 September 2010).

The third wave (or interview cycle) of SoFIE was conducted from 1 October 2004 to 30 September 2005 and sought responses to these modules: Household questionnaire; Non-contact short form; Demographics; Child; Labour market history; Education; Family; Labour market; Income; Contact; Standard series; Health modules. These modules are repeated in Waves 5 (1 October 2006 to 30 September 2007) and 7 (1 October 2008 to 30 September 2009).

## **2.5 Confidentiality requirements**

Access to the unit record data was via Statistics NZ's Data Laboratory under protocols designed to protect the confidentiality of information. As noted in the SoFIE-Health Privacy Impact Assessment (PIA) Jan-Feb 2007, the protocols ensure:

- Accessed data are anonymised and confidentialised.
- Access is on-site and supervised by Statistics NZ employees.

- All output is subject to independent scrutiny by Statistics NZ staff to ensure the material is in the form of statistical aggregates that comply with the confidentiality requirements.

The current agreement with Statistics NZ for exporting summary SoFIE data from the DataLab includes a requirement that all stratum counts are randomly rounded to 5 units. Rounded counts less than 10 are replaced by 10.

Further details are available in the Data Laboratory Output Guide, Statistics NZ, 19 March 2008.

## 3 SoFIE Data and Software

### 3.1 General Information

Unit record SoFIE data is only available for processing and analysis in the Statistics NZ Data Laboratory. Statistics NZ additionally require that: (i) each new wave of SoFIE data requires a separate contract to be negotiated between e.g., Statistics NZ and the UOW; (ii) SoFIE data summaries must go through a confidentiality checking procedure before it can be taken outside the Data Laboratory (see Section 2.5).

### 3.2 Outline of Data Organisation and Processing Software

#### 3.2.1 Overview

Processing and analysis of SoFIE-Health data involves a number of people with varying levels of familiarity with data and software tools. As of 2007-08, there is a need to maximise ease of access to SoFIE-Health data for researchers (“users”). However, maintaining the stability and quality of user datasets require that access to key processing software and associated datasets be restricted to a subset of experienced UOW SoFIE-Health “administrators”.

As discussed in more detail later, this has been achieved by:

1. Developing a user/administrator directory structure to facilitate user access and administrator control of standard data processing tasks and software.
2. Simplifying the SoFIE (W1-3) dataset. (i) Non-spell Statistics NZ SAS files have been transposed by the cross-sectional unique identifier (*PersonCS*), merged into ‘themed’ files, and variables selected to generate the first SoFIE-Health research dataset. Each record in this research dataset contains at most one record for each cross-sectional identifier for a given wave<sup>4</sup>. This dataset has more records but fewer variables per record than the second dataset, and will be referred to as the ‘**long**’ **SoFIE-Health dataset** below. (ii) A second user dataset is created by further variable selection and transposition by the unique longitudinal identifier (*PersonLT*). It will be referred to as the ‘**wide**’ **research dataset** in what follows. Users have (read-only) access to both the long and wide research datasets.

---

<sup>4</sup> Though there could be up to three records for each unique longitudinal identifier i.e., one per wave

3. Developing (SAS) macros and programs to perform standard tasks. These programs are available to administrators and users to assist with processing and research activities.

We anticipate that Statistics NZ will generate stable SAS datasets with minimal need for code re-engineering, thus maximising productivity and minimising the potential for errors. Our design objective is that routine processing tasks (e.g., addition of each new wave of data) and standard analyses be driven via a metadata database i.e., a database that contains sufficient information about SoFIE-Health files and variables to enable automatic generation of user datasets without significant change to the software that generates the user datasets.

### 3.2.2 Data organisation

The general approach is as follows: SoFIE datasets have been grouped into a relatively small number of ‘themed’ files which retain all the information in the original Statistics NZ files. For non-spell<sup>5</sup> datasets, this approach has reduced 36 initial Statistics NZ files to 8. The no-loss-of-information rule simplifies the script design task and enables other (non-HIRP) users to more easily access SoFIE data.

The fidelity of the merging process requires a unique record key. This is achieved by an initial transposition of Statistics NZ files (by the cross-sectional identifier *PersonCS*) so that the files to be merged have one record per *PersonCS* in any given wave.

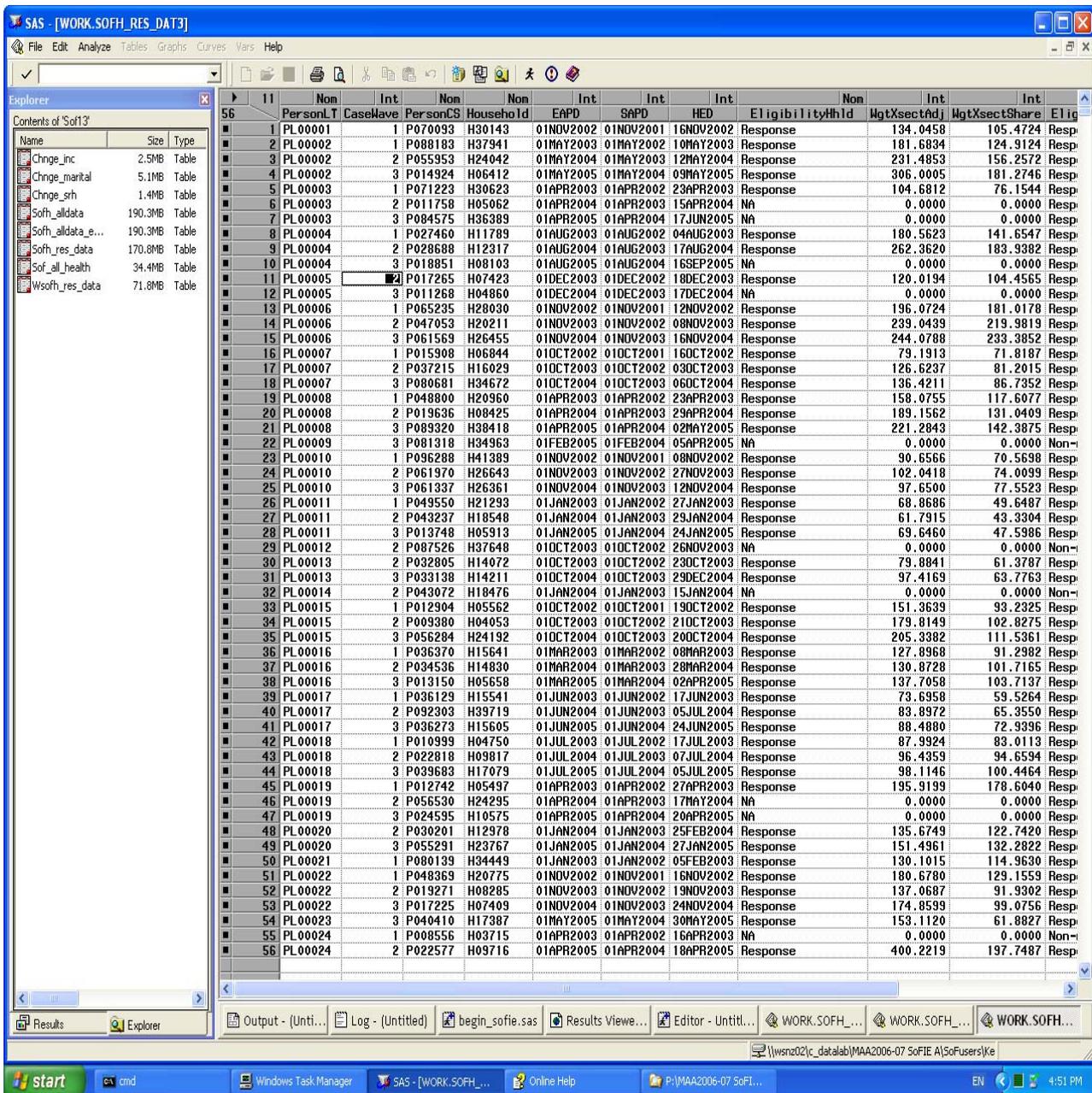
Top-level and intermediate-level ‘themed’ files (in bold) are organised as follows<sup>6</sup>:

- **Id\_data**: (Caseverperson + Casever + Case + Estcasever by Household) + Estcaverperson + Pqperson + Drvdemog + Hqreln + Hqethnicity by *PersonCS*.
- **Demogcomb**: (Id\_data + Demoethnicity + Demog + Drvdemog by *PersonCS*)
- **Dwellcomb**: Dwell + Drvhhldincome + Drvhhldcompn + Drvdwell + Dwellappliance + Dwellpayment by *Household*)
- **Educationcomb**: (Id\_data + Education + Drveducation + Qualpostschool + Studyymnth by *PersonCS*)
- **Familycomb**: (Id\_data + Drvfamil + Drvfamilynucleus by Household) + Family + Familyreln by *PersonCS*)
- **Incomecomb**: (Id\_data + Waveinchdr + Waveinc + Drvincome + Drvgovtincome + Incsubtypeincl by *PersonCS*)
- **Assliabcomb**: (Id\_data + Assliab + Assliabdetail + Assliabmisc + Drvassliab by *PersonCS*)
- **Healthcomb**: (Id\_data + Healthcaredepriv + Healthchronicdisease + Healthdrinksmoking + Healthgenphys + Healthmental + Healthprovtype+Healthrelnoccupation by *PersonCS*)

<sup>5</sup> Spell data contains information about events related to sub-annual time periods e.g., periods of activity in the labour market

<sup>6</sup> Original Statistics NZ files are given in (brackets), along with the merge *variable* – see Appendix 7.3 for more information about Statistics NZ files variables.

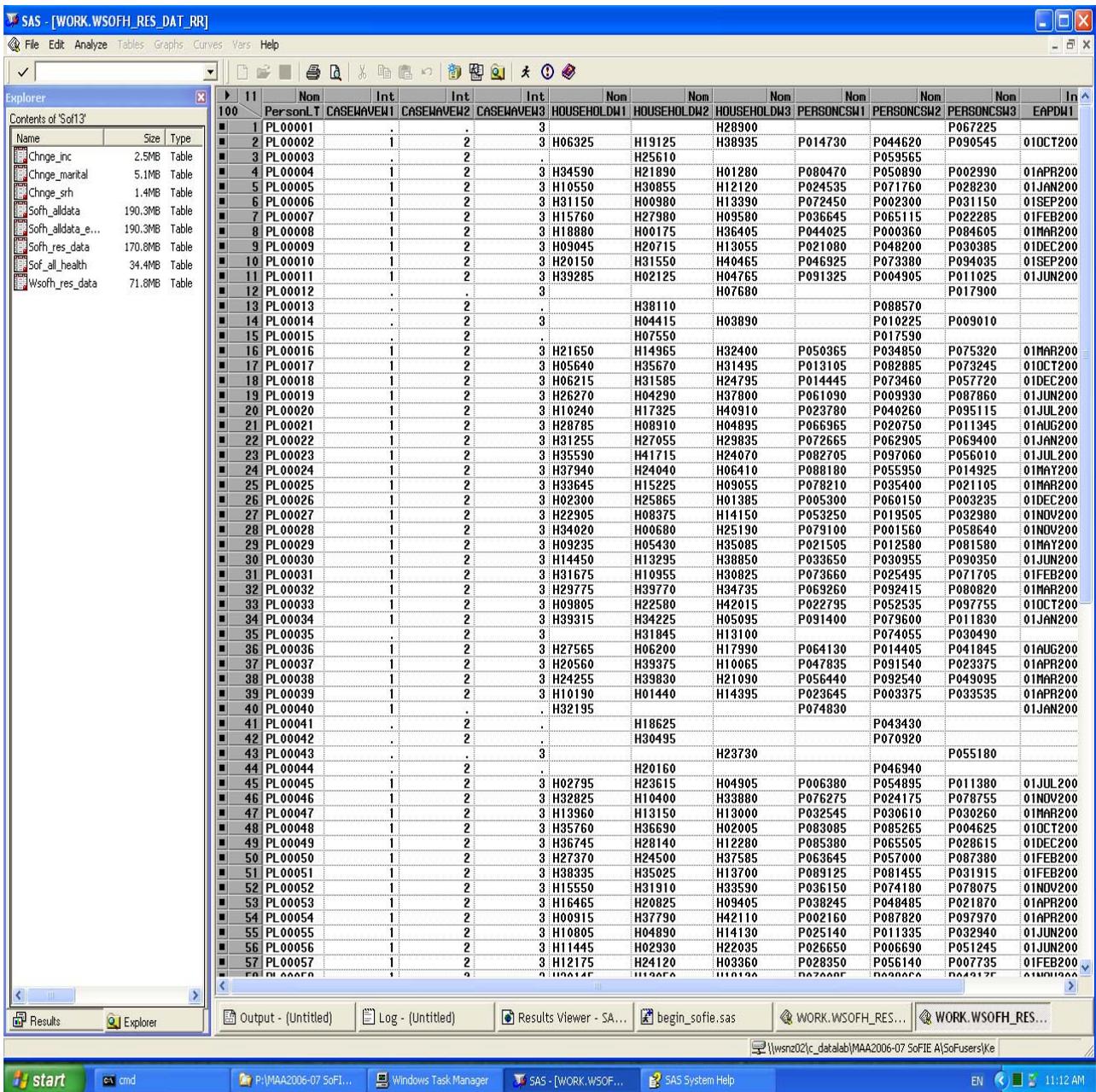
These intermediate-level themed datasets are merged into a single long wave 1-3 dataset by selecting the subset of variables that are potentially informative for SoFIE-Health research over the next 1-2 years (see Figure 3). Additional derived health, equivalised household income, education, wealth, etc; variables are added to the long research dataset at this stage. Each record in the long dataset contains all retained variables relevant to a respondent in a given wave since it has been created by merging on the cross-sectional identifier *PersonCS*. There is at most one such record each wave, though a given respondent can be associated with up to three different records since he/she will have a different *PersonCS* assigned in each wave.



**Figure 3:** Confidentialised view of the first few variables and records in the final long research dataset. Note there can be up to 3 records with constant *PersonLT* values, each with unique *PersonCS* values.

A final ‘wide’ dataset is created from the long dataset, by transposing on the longitudinal identifier (*PersonLT*) and selecting only those variables known to address immediate SoFIE-Health Goal 1-3 objectives. Additional derived variables are added to the ‘wide’ dataset at present<sup>7</sup> - e.g., percentiles for income and wealth. Each record in the ‘wide’ research dataset now contains all retained variables associated with a respondent over all waves (see Figure 4).

<sup>7</sup> However, the intention is to restrict addition of derived variables to a single stage (e.g., generation of the ‘long’ research dataset), so this additional step may be temporary.



**Figure 4:** Confidentialised view of the first few variables and records in the final wide research dataset. Missing numeric values are indicated with '.', and missing character values with blanks. In this dataset, each record has a unique value for the *PersonLT* variable, and wave-specific variables have the characters 'Wj' appended ( $j=1,2,3$ ).

SAS librefs<sup>8</sup> pointing to DataLab SoFIE server directories (e.g., containing the research datasets, and checking, and user work areas) are created automatically for each user via a simple script (*begin\_sofie.sas*). Further information about the processing scripts/macros, and the variables themselves are given in the following sections.

<sup>8</sup> A name that is temporarily associated with a SAS data library, often a disk directory. For example, the libref *sof13* created by *begin\_sofie.sas* references the directory *P:MAA2006-07 SoFIE A\SASdata\sof13*.

### 3.2.3 Data processing and DataLab directory structure

SAS programs for the above tasks, as well as a ‘library’ of community<sup>9</sup> SAS macros are available via a readonly SAS catalogue of compiled macros, or as scripts in subdirectories of the “SoFIE Admin” directory (see Figure 5).

#### 3.2.3.1 SoFIE-Health DataLab directory structure

Figure 5 is a screen view of the directory structure adopted to accommodate the requirements of a multi-user SoFIE-Health research effort, with varying levels of user-familiarity with data and software tools, and code maintenance tasks. A more detailed listing of directories and files is given in Section 7.4

The **Checked** and **Checking** directories satisfy the requirement that Statistics NZ staff must ensure the data meets confidentiality requirements. Each user is assigned a named subdirectory of the **Checking** directory in which they can store data to be checked (e.g. **Checking\Ken**) and a libref is created by the initialisation macro *%sofsess* (in *begin\_sofie.sas*). Statistics NZ staff move data files into subdirectories of the **Checked** directory once they have determined that the data meet confidentiality requirements.

The directory **SASdata\sof13** is the repository for the research datasets (long and wide) created by scripts described above. **SASdata\sof13A** stores the themed and derived variable SAS datasets. Librefs are created to these directories for users and administrators by the *%sofsess* macro, as appropriate.

The **SASprogs\SoFIE Admin** directory houses all the programs required to derive the themed and research datasets, as well as programs to derive variables required by research under the SoFIE-Health project. There is one directory for each task, with subdirectories for sub-tasks as required (e.g., **derive\_vars\CPI\_eqinc**, **derive\_vars\health**).

The grouping of software processing (and later analysis) tasks in this way, coupled with the identification of administration and maintenance responsibilities and design rules based on general solutions to processing and analytical tasks described above, greatly assists the maintenance of robust, stable SoFIE-Health community code. These principles are also important for macro development, which are stored in the **sofie macros** directory.

The **Documentation** directory is used as a repository for key Statistics NZ and UOW SoFIE documentation, and the **SoFusers** directory provides users with disk space which they can organise as they wish to support their own research.

#### 3.2.3.2 Code maintenance

Maintenance<sup>10</sup> of community SAS macros for SoFIE-Health is the responsibility of designated members (administrators<sup>11</sup>) of the SoFIE-Health team. The process for

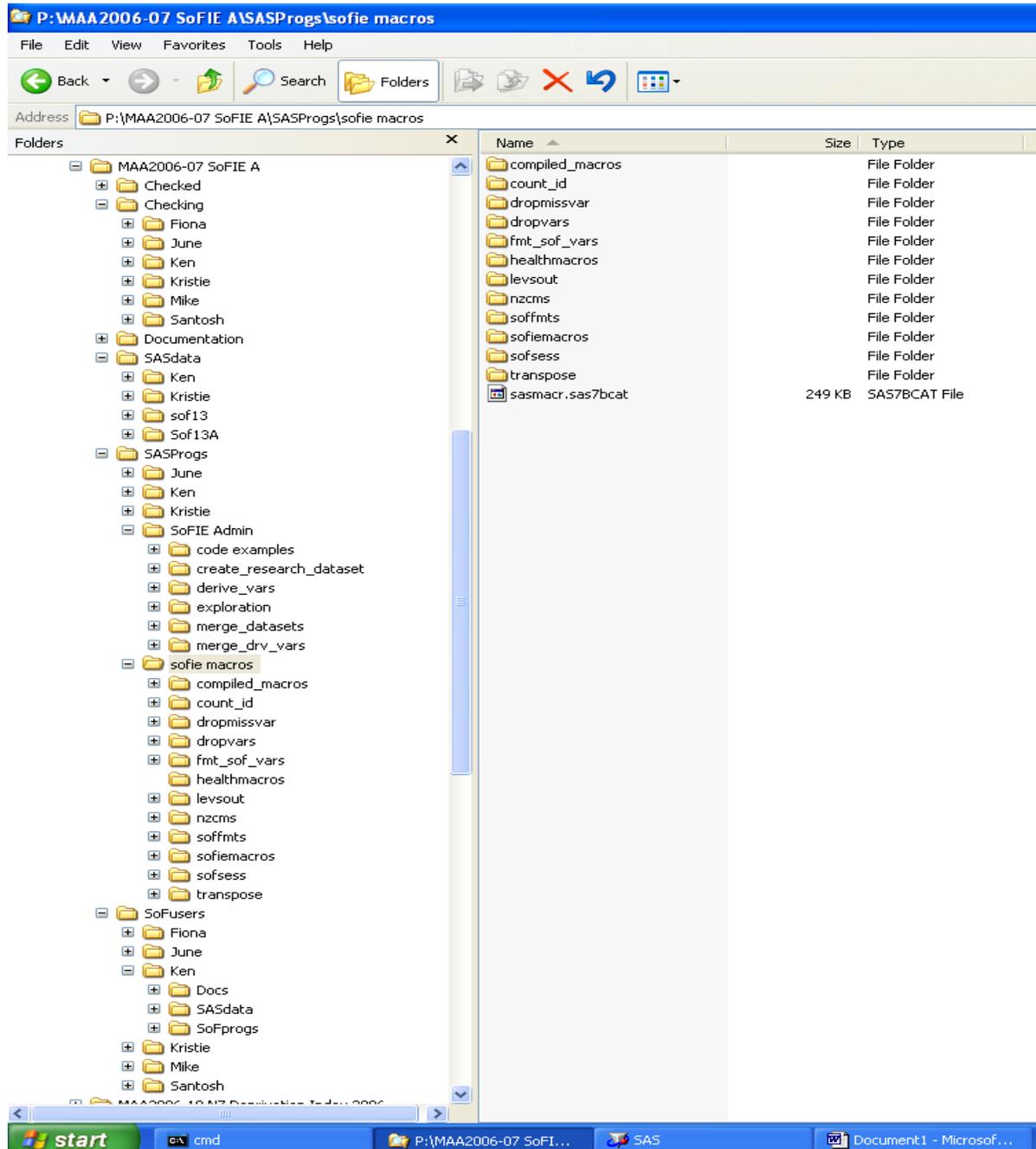
---

<sup>9</sup> Community SAS macros are contributed by SoFIE\_Health researchers that perform analysis or data management tasks required by all WSMHS SoFIE-Health researchers (users + administrators).

<sup>10</sup> Including the addition of new macros or other software, increasing the capability of existing software, fixing reported bugs, etc.

<sup>11</sup> Currently Ken Richardson and Kristie Carter

correcting errors and improving existing macros is as follows: Copy the existing code to a user directory, correct the error, and test. Comment out the STORE option in the %macro statement to avoid trying to overwrite the current compiled version which will generate an error. Once you have convinced an administrator that the macro runs as intended, and is properly documented, they will compile/store the macro into the macro database.



**Figure 5** Screen view of the Wellington School of Medicine directory structure in the Statistics NZ Data Laboratory. The top level directory (P: drive) is currently mapped to `\\wsnz02c_data\lab`.

If you are making new code available to the SoFIE-Health community, a similar process applies: Demonstrate to an administrator that the code runs as intended and is properly

documented. The administrator will then create the directory to store the macro code and compile/store the macro into the macro database.

### 3.2.3.3 Compiling and storing SAS macros

SAS compiles and stores macros that are ‘run’ from the SAS environment in the WORK.sasmacr catalog, and are therefore not permanent between SAS sessions. A permanent catalog can be created using the MSTORE (to allow compilation) and SASMSTORE (to define a libref that will ‘contain’ the compiled macro) options. For administrators, this is all taken care of by the %include statement in *begin\_sofie.sas* that references *compile\_macros.sas*. The end result is that macros compiled with the /STORE option are stored in a catalog called SOFMACR.sasmacr (SOFMACR being the libref referencing the catalog file on disk). These are automatically available to the user provided that libref has been defined and SAS ‘knows’ about it via the SASMSTORE option.

For users, access to compiled macros is provided by the %include statement in the user version of *begin\_sofie.sas* that references *use\_compiled\_macros.sas*. This forces readonly access<sup>12</sup> and reduces the potential for accidentally overwriting compiled macros by users<sup>13</sup>. Note that concurrent access of the catalog by users and administrators alike requires inclusion of *use\_compiled\_macros.sas* rather than *compile\_macros.sas*.

### 3.2.3.4 SAS data formats

A permanent format library of SAS data formats is maintained to assist the development of a coherent approach across SoFIE-Health research users. Maintenance of the library is the responsibility of the administrators. A listing and description of the formats available in the library is given in Appendix 7.1

The format library is created by the %soffmts macro, which uses the LIBRARY=LIBREF.CATALOG option in a PROC FORMAT statement (where CATALOG can be any valid SAS catalog name, but = FORMAT here). If libref is defined (in a LIBNAME statement) as LIBRARY, SAS automatically searches LIBRARY.FORMAT for user-defined formats. These details are all taken care of by the macro %sofsess (see *begin\_sofie.sas*), which guarantees automatic user access to formats defined in *library.format*.

### 3.2.3.5 User access to SoFIE-Health data and software

SoFIE\_Health data and project software is made available to users by running the two lines of code in the SAS program *begin\_sofie.sas*.

The recommended procedure is as follows: (i) Copy **P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\code examples\begin\_sofie.sas** to your user directory (ii) edit the copied file so that the *mycheck* and *mywork* arguments of the macro %sofsess point to your check and work directories respectively. (iii) run the %include statement that compiles the *use\_compiled\_macros.sas* program (iv) run the %sofsess macro.

---

<sup>12</sup> A warning of the catalog readonly status appears in the SAS log window in a green typeface.

<sup>13</sup> Further reduced by storing the catalog in a directory to which users do not have write access.

The *%include* statement in *begin\_sofie.sas* makes compiled macros available to the user (see above), and the *%sofsess* macro defines the following *librefs*:

- SOF13 which references the directory where the long and wide research datasets are stored, and to which SoFIE-Health users have read-only access
- FEXPORT which references the directory where Statistics NZ staff can locate files to check for confidentiality purposes prior to export to the UOW
- MYWORK which references a user work directory
- LIBRARY which references the permanent format library (thereby making these available to the user).

### 3.2.3.6 SoFIE-Health data processing

The following sections provide an overview of the SAS programs used to generate user datasets. Listings of the programs used currently are given in Appendix 7.2. They are described in this section in the order in which they are used.

#### 3.2.3.6.1 *merge\_setup.sas*

This program extracts information from Statistics NZ read-only files that will be required during the merging process that creates the themed datasets. Recall that each file that is merged to create a ‘themed’ dataset must be transposed (by *PersonCS*) if any variable in that file has a within-wave multiplicity<sup>14</sup> greater than 1. Furthermore, variables in a Statistics NZ file can have different multiplicities. The *%toward* macro (see later), used in *merge\_datasets.sas* to transpose the files prior to merging, requires information about the maximum multiplicity of the dataset to be transposed. This information is determined by the *merge\_setup* program using the *%count\_id\_vars* macro.

Currently, results from *merge\_setup.sas* are pasted by hand back into the program as comments and used later in *merge\_datasets.sas*. In future the *merge\_setup.sas* script will be turned into a macro, with the results loaded into macro variables that can be made available to *merge\_datasets.sas*.

#### 3.2.3.6.2 *merge\_datasets.sas*

Most of the work required for generating the ‘themed’ datasets is done in this script. The intended<sup>15</sup> design ‘rules’ are:

- No ‘loss’ of information (variables)
- Rename variables having the same name in different datasets e.g., ethnicity, sex by prepending the name of the dataset to the variable
- Use keep (rather than drop) statements, since keep variables are likely to be defined via metadata databases.

---

<sup>14</sup> The number of different values a variable can have for any given individual and any given wave in one of the read-only Statistics NZ datasets. If more than one response per respondent is possible, the *PersonCS* identifier variable for any given wave will be duplicated across several records in the file containing that variable. Many variables have a within-wave multiplicity of 1, but some variables, such as ethnicity, can have within-wave multiplicity > 1.

<sup>15</sup> Though not consistently (as at April 2008)

The main tasks include:

- Realisation of themes developed for intermediate datasets
- Inclusion of ID variables in each output dataset
- Conversion of character ‘numeric’ data (e.g., the string ‘10’) to numeric (the integer 10) using the *%char2num* macro (see below). This step is useful for viewing the files with other SAS software, and is consistent with NZCMS<sup>16</sup> practice.
- Transposition of datasets (by PersonCS) if required using the *%towitz* macro (see below).
- Merging the transposed datasets by *PersonCS* into themed datasets.

Variables that have been transposed for the long user dataset will typically have names like *var1-var10*, where *var* is the name of the original variable, and 10 is the maximum multiplicity for that dataset. Subsequently, in the wide user dataset (assuming 3 waves), these variable names will become *var1W1*, *var1W2*, *var1W3*, ..., *var10W1*, *var10W2*, *var10W3*.

A refinement of *%towitz* (not used for the current datasets, but important for technical reasons<sup>17</sup>) will prefix the transposition number (1-10 in the example above) with a user-defined character, namely ‘X’ at the cross-sectional level, and ‘W’ at the longitudinal level. Using this refinement the above example will in future generate variable names *varX1-varX10* in the long user dataset, and *varX1W1*, *varX1W2*, *varX1W3*, ..., *varX10W3* in the wide user dataset in a future re-processing of these data.

#### ***3.2.3.6.3 check\_merge\_datasets.sas***

Generates variable names and levels in the themed datasets (using the *%levs\_out* macro) for checking against the original datasets.

#### ***3.2.3.6.4 run\_dropvars.sas***

Some transposed variables end up completely empty if they have multiplicity less than the maximum file multiplicity. This program uses the *%dropmissvar* macro (see below) to remove such variables, and creates a dataset label.

#### ***3.2.3.6.5 create\_user\_dataset.sas***

Creates a precursor (*sof13a.Sofh\_alldata*) of the ‘long’ research dataset by merging the demographic, education, family, income, assets/liabilities, labour market, household, and health themed datasets. Selected variables, defined in **P:\MAA2006-07 SoFIE A\SoFusers\Ken\Docs\Data Dictionary required vars v1.4.xls**, are retained in the output dataset.

---

<sup>16</sup> The New Zealand Census-Mortality Study (NZCMS: see e.g., Blakely 2002)

<sup>17</sup> Some SoFIE-Health programs e.g., *fnt\_sof\_vars.sas* use character pattern matching of variable names to identify variables. Occasionally the method fails (e.g., for *DemogEthnicity* and *DemogEthnicityNR*) unless variable names are uniquely distinguished by the cross-sectional and/or longitudinal identifier characters .

### 3.2.3.6.6 *eqinc.sas*

Computes CPI-adjusted household equivalised income. Household income, derived by Statistics NZ, adjusted for inflation and equivalised for economies of scale, is used as a measure of income for each individual in the the same household. The assumption here is that a individual's standard of living is determined by the resources available to the household in which the individual lives rather than by their personal income.

The equivalisation procedure due to Jensen 1978<sup>18</sup>, and Jensen 1988<sup>19</sup> is used to adjust incomes to be more comparable across households with different sizes and age structures. Jensen gives two results depending on whether the ages of children are distinguished or not. Both are calculated by this program (*HhldIncAnn\_Jen1* and *HhldIncAnn\_Jen* respectively). Additionally, a household income equivalised using the Luxembourg Income Scale (Atkinson et al 1995<sup>20</sup>) is computed (*HhldIncAnn\_Lux*).

### 3.2.3.6.7 *score\_sf36.sas*

The SF-36 is comprised of 36 questions that fall into eight health domains: general health perceptions (GH – *health, sickquicker, ashealthy, expectworsen, healthexcellent*), physical functioning (PF – *limitvigorousacty, limitmoderateacty, limitliftcarry, limitmultiset, limitoneset* [see *Notes*], *limitbendkneel, limitwalk100m, limitwalk500m, limitwalk1km* [see *Notes*], *limitbathedress* [see notes]), role limitations due to physical functioning (RP – *physicallesstime, physicallessdone, physicallimited, physicaldifficulty*), bodily pain (BP – *painseverity, paininterferedeg*), general mental health (MH - 5), role limitations due to emotional problems (RE – *mentallesstime, mentallessdaone, mentallesscareful*), vitality (VT – *feltlively, feltenergetic, feltwornout, felttired*), and social functioning (SF – *probinterferedeg, probinterfererefreq*). The remaining item (HT – *healthnow*), relating to change in health compared to 12 months ago, is not scored as a separate dimension.

#### *Notes:*

In SoFIE the maximum walking (*limitwalk1km*) and steps (*limitmultiset*) is asked first. If a respondent is very limited on the maximum then the following walking and steps questions are not asked. Therefore, if a respondent was very limited this score was assigned to the subsequent questions.

The *limitbathedress* variable is derived from combining two questions (*limitbatheself, limitdressself*), if the respondent is not limited in either bathing or dressing then *limitbathedress* = 1 (not limited), otherwise if the respondent is limited in both bathing and dressing then *limitbathedress* = 3 (very limited), otherwise *limitbathedress* = 2 (slightly limited).

The variables above are transformed into SF36 variable names in a datastep and then checked for consistency using the *%sf36chk(sf36, PersonCS)* macro.

---

<sup>18</sup> Jensen, 1978. Minimum income levels and income equivalence scales. Unpublished report for the Department of Social Welfare.

<sup>19</sup> Jensen, 1988. Income equivalences and the estimation of family expenditure on children. Unpublished report for the Department of Social Welfare.

<sup>20</sup> Atkinson, A., Rainwater, L., and Seeding, T., 1995. Income distribution in OECD countries: Evidence for the Luxembourg Income Study. Paris: OECD.

The *%score\_SF36* (see Appendix 7.2) macro is used to score the SF-36 Health questionnaire using the Ware *et al.*, 2000<sup>21</sup> algorithm. Version 2 has new scoring for a number of questions, and uses the factor score coefficients from the NZ Health Survey 1996/7 for the calculation of Physical Component Score (PCS) and Mental Component Score (MCS) summary scores.

Each is scored from 0 (worst score) to 100 (best score). A score of 100 in the PF, RP, BP, SF, and RE domains and scores of 50 in the remaining three domains, GH, VT, and MH, indicate an absence of problems in those areas. The general health, Vitality and Mental Health scales cover both positive and negative health states with the maximum score on these scales indicates not just the absence of disability, but also the presence of a positive health state. For example, a score of 100 in physical functioning indicates an ability to perform all activities without limitations due to health; whereas a score of 50 in mental health indicates an ability to function without personal or emotional problems. A score of 100 in the mental health domain indicates that the respondent feels peaceful and happy and is calm all the time.

#### **3.2.3.6.8 *drv\_chron.sas***

The original data is presented as multiplicity 8 (8 records per *PersonCS*) for each chronic disease and age of diagnosis. This needs to be transposed by *PersonCS* to get 1 record per *PersonCS* and 8 columns of chronic diagnosis and 8 columns of age of diagnosis.

These columns are then renamed as *Asthma*, High Blood Pressure (*hi\_bp*), High Cholesterol (*hi\_chol*), Heart Disease (*heart\_dis*), Diabetes, Stroke, Migraines, Chronic Depression/Manic Depression/ Schizophrenia (*depress\_schiz*). Age of diagnosis of each disease are renamed as *asthma\_age*, *hibp\_age*, *hichol\_age*, *heartdis\_age*, *diabetes\_age*, *stroke\_age*, *migraine\_age*. NOTE: *depressschiz\_age* is empty for all respondents and is not included in the output dataset.

The number of chronic diseases were summed and coded into a co-morbidities index (*chronic\_comorbid*): 0, 1-2, >2 co-morbidities.

#### **3.2.3.6.9 *drv\_nzidep.sas***

There are eight questions used for the five-point individual-level index of socioeconomic deprivation (*NZiDep*): Unemployment (see *Notes*), benefits (see *Notes*), cheaperfood, saveheating, wornoutshoes, foodgrants, communityorg, oftennofruitveg (see *Notes*).

##### *Notes:*

Unemployment was defined using data on labour market activity, if a respondent was between 15 and 65 and was either unemployed or spent 4 or more weeks unemployed but actively looking for work.

Benefits was defined using information on government incomes earned during the year (*Drvgovtincom.GovtIncomeType* = 17 – last WINZ benefit payments during the reference period). Included subtype benefits were: youth, sickness, domestic purposes, invalids, widows, transitional retirement, orphans or unsupported child, special benefit, and accommodation supplement.

---

<sup>21</sup> Ware, J.E., Snow, K.K., Kosinski, M.A., Gandek, M.S., 2000. SF-36 Health Survey Manual and Interpretation Guide. Lincoln: Quality Metric Incorporated

*Oftennofruitveg* was calculated using information from *nofruitveg* and *nofruitvegoften*.

*Creating the NZiDep index*

- (i) Add the 'yes' responses (count any missing data as 'no').
- (ii) If a respondent had more than half (>4) of the questions missing then the NZiDep score was coded as missing.
- (iii) The deprivation score was coded into the following five ordinal categories (relatively few people have the largest number of deprivation characteristics).

<b>NziDep code</b>	<b>Description</b>
1	no deprivation
2	one deprivation characteristic
3	two deprivation characteristics
4	three or four deprivation characteristics
5	five or more deprivation characteristics

**3.2.3.6.10** *drv\_smoke\_alc.sas*

Smoking status (*smoke\_status*), current smoker, ex-smoker, never smoker, was calculated using information from *smokesreg* and *smokeeverreg*.

Binge drinking was calculated as if a respondent had > 0 occasions (*occmaydrinks*) in the last 4 weeks where they drank 8 (male) or 6 (female) drinks containing alcohol.

*Alc\_drink* measures frequency of drinking and is derived from *haddrink* and *daysdrink*.

<b>Alc_drink</b>	<b>Description</b>
0	Never drinks
1	Drinks monthly or less
2	Drinks 2-4 times a month
3	Drinks 2-3 times a week
4	Drinks 4 or more times a week.

*Alc\_num* measures the number of drinks on a typical drinking day and is derived from *drinksperday*.

<b>Alc_num</b>	<b>Description</b>
0	1-2 drinks per typical day
1	3-4 drinks per typical day
2	5-6 drinks per typical day
3	7-9 drinks per typical day
4	10 drinks per typical day

*Alc\_binge* measures the frequency of binge drinking and is derived from *occmnydrinks*.

<b>Alc_binge</b>	<b>Description</b>
0	Never binge drinks
1	binge drank monthly
2	binge drank 2 times per month
3	binge drank weekly
4	daily or almost daily binge drinking

Note there are a number of formats that have been created to make comparisons with the Alcohol Liquor Advisory Council of New Zealand’s AUDIT<sup>22</sup> questions.

### **3.2.3.6.11 score\_k10\_PSS.sas**

Create Kessler and Perceived Stress Score from SoFIE data (non-missing wave 3 only)

The Kessler 10 score is calculated as the sum of the following variables: *FeltDepress*, *FeltDepressExtreme*, *FeltEffortful*, *FeltFidgety*, *FeltFidgetyExtreme*, *FeltHopeless*, *FeltNervous*, *FeltNervousExtreme*, *FeltTiredNoRsn*, *FeltWorthless*.

A number of questions were ordered pair questions (*FeltDepress*, *FeltDepressExtreme*, *FeltFidgety*, *FeltFidgetyExtreme*, *FeltNervous*, *FeltNervousExtreme*). If a respondent answered “never” to the first question then the second was skipped. In that event, this code copies the first “never” score into the response for the second question.

The results of the 10 questions are summed into a K10 score. If there was 1 missing question *then* the individual mean was imputed for this question. If more than 1 question in the K10 was missing then the K10\_score was set to missing.

The scores were grouped into four levels according to the criteria developed by Andrews and Slade (2001)<sup>23</sup>:

<b>K10_group</b>	<b>K10 score</b>
Low	10-15
Moderate	16-21
High	22-29
Very high	30+

The perceived stress score consists of four variables: *FeltUnableControl*, *FeltProblemExtreme*, *FeltGoingWell*, *FeltConfident*. Each item is scored using a five-point Likert format, 4=never, 3=almost never, 2=sometimes, 1=fairly often and 0=very often for items stating positive experiences or response.

### **3.2.3.6.12 wealth.sas**

Personal wealth (*PersWealth*) was derived from *wealthcode* = “31”, couple wealth (*CoupWealth*) was derived from *wealthcode* = “32”. A measure of equivalised overall

<sup>22</sup> The WHO Alcohol Use Disorders Identification Test (e.g., see [http://whqlibdoc.who.int/hq/2001/WHO\\_MSD\\_MSB\\_01.6a.pdf](http://whqlibdoc.who.int/hq/2001/WHO_MSD_MSB_01.6a.pdf))

<sup>23</sup> Andrews, G., & Slade, T. (2001). Interpreting scores on the Kessler Psychological Distress Scale (K10). *Australian and New Zealand Journal of Public Health*, 25(6), 494-497

wealth (*Wealth*) was calculated by taking couple wealth divided by 2 (if couple wealth is not missing), otherwise personal wealth was used. An indicator variable of negative wealth (*wealth\_neg*) was calculated to identify respondents with wealth < 0.

#### **3.2.3.6.13** *maxqualrank\_wsm.sas*

The Statistics NZ derived variable *MaxQualRank* does not extrapolate maximum qualifications from one year to the next, leaving many missing values. This program does so by filling missing cells with valid responses (non-missing) from the previous year. Note that for the moment *maxqualrank\_WSM.sas* operates on the wide research dataset.

#### **3.2.3.6.14** *merge\_drv\_vars.sas*

Key variables derived by *drv\_chron.sas*, *score\_k10\_PSS.sas*, *drv\_nzidep.sas*, *score\_sf36.sas*, *drv\_smoke\_alc.sas*, *wealth.sas*, *eqinc.sas*, are merged by *PersonCS* with *sof13a.Sofh\_alldata*, the long research dataset generated by *create\_user\_dataset.sas*.

#### **3.2.3.6.15** *create\_research\_dataset.sas*

Generates the wide and long research datasets from the file created by *merge\_drv\_vars.sas*. Key tasks performed by *create\_research\_datatsets.sas* are (i) keep a final set of ‘research’ variables as defined by the file *Data Dictionary required vars v1.2KC sorted.xls* (ii) balance dataset by *CaseWave* to get wave labelling right (iii) transpose by *PersonLT* to form the wide research dataset (*SOF13.WSofH\_res\_data*) (iv) format research dataset variables.

#### **3.2.3.6.16** *incwlth\_pctiles.sas*

Derives cutpoints for income/wealth percentiles from income and wealth data, using the SAS **univariate** procedure. Mean (over waves 1-3) equivalised household income data is derived from *sof13.WSofH\_res\_data*. Other income or wealth (wave 2) data is taken from the same dataset.

Formats and informats based on the cutpoints derived by *incwlth\_pctiles.sas* are created by *%soffmts.sas*. Percentile variables are added to *sof13.WSofH\_res\_data* using the program *rank\_incwealth.sas*.

#### **3.2.3.6.17** *rank\_incwlth.sas*

Creates new percentiles variables in the wide dataset *sof13.WSofH\_res\_data* by applying informats based on the cutpoints derived by *incwlth\_pctiles.sas*. The program uses the macro **%num2pc** which requires as arguments the input and output datasets, a list of old income/wealth variables, and a list of informats to be applied to those variables.

The output dataset contains additional income/wealth variables which append a pre-defined **string** to the original variable name e.g., quintiles have the string ‘r5’ appended.

Note that the generation of income percentiles is done using the wide research dataset.

## 3.2.4 Variables

### 3.2.4.1 Naming conventions

Variable names developed by Statistics NZ are usually descriptive of purpose (see Appendix 7.3).

As noted in section 3.2.3, variables in the long user dataset that have multiplicities greater than 1 in the original Statistics NZ datasets will have names like *var1*, *var2*, ..., where *var* is the name of the original variable, and *1*, *2*, ... is the transposition number (an integer between 1 and the maximum multiplicity for that dataset). Subsequently, in the wide research dataset, *var1* becomes *var1W1*, *var1W2*, *var1W3*, ..., and so forth: In this case *W<sub>j</sub>*, *j*=1,2,3, ... are wave numbers.

In future waves, the above example will be generalised slightly, to become *varX1*, *varX2*, ..., in the long research dataset, whereas *varX1* becomes *varX1W1*, *varX1W2*, ..., in the wide research dataset.

### 3.2.4.2 Variable and data dictionary'

Appendix 7.3 reproduces the Statistics NZ dictionary of SoFIE datasets and variables, together with brief descriptions.

## 3.2.5 Imputation and variance estimation of missing data by Statistics New Zealand

Imputation will be used in some circumstances to supply missing data items for interviewed individuals. Imputation reduces the potential bias caused by missing data, which allows better analysis of the data. This is particularly important for a longitudinal survey, where attrition has a cumulative effect over the survey life. Because of this, imputation will increase in both importance and quantum as more waves of SoFIE become available. This section of the report is likely to be expanded in future versions, and will include details about how imputed information is updated in the database.

Some respondents were unable to provide complete information. In these cases, missing values were imputed for all key fields. The key fields for SoFIE, as defined by Statistics NZ, were age, ethnicity, income and pay details. Where possible, information was imputed deterministically, using other information reported by the respondent to provide a likely estimate for the missing value. When deterministic imputation was not possible, a 'hot deck' imputation method was used. This method involves identifying another respondent with similar characteristics to become the 'donor' and provide the imputed value. In future waves, values from previous waves will be used to impute missing values from the current wave, or information successfully collected in Wave 2 will be used to over-write information which had to be imputed in Wave 1.

### 3.2.5.1 Imputation for item non-response

Item non-response occurs when a respondent does not provide data for a selected item (or items) in the survey. Since the main objective of SoFIE is to provide a unit record longitudinal data file for analysis by users, ideally all item non-response by OSMs should be imputed. However, initially only the more important variables required for key outputs will be imputed. Those variables that contribute to household and family derived

variables will be imputed for cohabitants in order to derive the family and household variables for the OSMs. Details of these variables will be included in future releases of this report.

### *3.2.5.2 Longitudinal imputation for wave non-response*

Individual wave non-response occurs when an individual does not respond to a given wave of the survey in one of three patterns: attrition, re-entry, and late entry.

- Attrition<sup>24</sup>. The only circumstance when imputation under attrition will be considered is when individual non-response occurs for an OSM in the last wave only.
- Re-entry. The only circumstance when imputation under re-entry will occur is when one wave of non-response occurs and this missing wave is bounded on both sides by a response. The difficulty of developing accurate imputation methods makes it very unlikely that imputation would be extended in future to cases where two consecutive waves are missing<sup>25</sup>. In most cases respondents will be asked to recall data over the missing wave.
- Late entry. Imputation for late entry will not occur because no one who is a non-respondent in wave one will be considered to be an OSM. If they begin responding at a later date, they will be treated as a cohabitant.

### *3.2.5.3 Imputation for complete family/household wave non-response<sup>24</sup>*

Complete household/family (unit) wave non-response occurs when all eligible individuals (OSMs and cohabitants) in a household/family are non-respondents in a given wave of the survey. It is anticipated that if this situation occurs, attempts will be made to find and obtain an interview from all OSMs in the family or household in the following wave. If this is successful, and retrospective information about the family and household composition in the missing wave is available, then the more important variables for the OSMs will be imputed. In addition, the variables used to derive family and household variables for cohabitants will be imputed. Both imputations will use the same methods as for individual non-response.

### *3.2.5.4 Types of data available for use in imputation*

In addition to survey data collected from individuals in the interview, there are two other types of data that have the potential to improve the quality of the imputed data: retrospective data and proxy data.

Retrospective data is obtained when a respondent is asked questions about the previous wave in which they were a non-respondent. Collecting retrospective data can reduce the effects of non-response, but there are quality issues to be aware of. Overseas experiments have shown that fewer events are recorded retrospectively for the missing wave than for data collected for the current wave. The data may also be less accurate than that collected for the current wave.

---

<sup>24</sup> Not implemented in the current (wave 3) imputation specification according to Statistics NZ.

<sup>25</sup> This situation will not occur since no attempt is made to follow OSMs if they miss more than one wave.

Proxy data is obtained when other members of a family are asked to provide data about a non-contact in the same family. Variables that will be collected by proxy are:

- age
- sex
- ethnicity
- employed/not employed at the end of the annual period
- whether employed full-time or part-time at the end of the annual period.

Again, the quality of proxy data will not be as good as for data obtained from the individual themselves. However, if a direct interview is not possible, proxy or retrospective data may be useful in assisting with imputation.

#### *3.2.5.5 Imputation methodology*

Longitudinal surveys can provide rich sources of auxiliary data for use in imputation. If an OSM doesn't respond in one wave, there may be information about them from the previous or subsequent wave that will improve the quality of the imputed data and help to reduce the likelihood of imputing spurious changes. It is planned to use this information, along with methodologies such as regression imputation<sup>26</sup> or hot-deck imputation within classes, to impute the change in income, labour market information, assets and liabilities since the previous (known) wave.

A research study carried out as part of the Longitudinal Survey of Income, Employment and Family Dynamics feasibility study indicated that income from the previous year was by far the most significant variable in predicting income for the current year. Labour market information from the current year also improved the imputation model. This underlined the importance of collecting proxy or retrospective data when possible. Removal of outliers from the donor pool also improved the results of the imputation considerably.

#### *3.2.5.6 Implementation issues*

The following issues were considered by Statistics NZ when developing the imputation methodology for SoFIE.

- All imputed values must be identifiable so analysts have the option of removing imputed records if desired.
- Data files will be made available after each wave. Users may find that an imputed value in one data file is different in a subsequent data file if the imputed value changes as a result of information gained in the following wave. Documented and identifiable audit trails for revisions will be created so users can identify when and why an imputation was changed as a result of new data.

---

<sup>26</sup> Not yet implemented as at wave 3 according to Statistics NZ

- Linking the cross-sectional records of cohabitants that are in the survey in two or more consecutive waves is very likely to improve the imputation for cohabitants<sup>27</sup>.

### 3.2.5.7 Variance estimation (production of longitudinal standard errors)

After calibration and adjustments for non-response in a complex survey, the exact variance of an estimator is difficult to obtain. Due to this complexity, approximately unbiased methods of variance estimation are often used in practice.

The choice of method for variance estimation involves a complex trade-off of factors such as accuracy, cost, ease of implementation and flexibility. Re-sampling methods are becoming more popular as they are relatively simple to carry out, particularly with the availability of powerful computers.

The jack-knife is recommended as the method of variance estimation for SoFIE for the following reasons:

- It will produce accurate and consistent variance estimates for the key outputs.
- Statistics New Zealand already has considerable experience in using jack-knife methodology.
- It is easy to apply.

While the jack-knife will be appropriate for Statistics New Zealand's key outputs from SoFIE, analysts wanting to estimate variances for other variables will have to use alternative methods such as the bootstrap.

### 3.2.6 Weight estimation by Statistics New Zealand

A weight is an indication of how many units in the survey population are being represented by the sample unit. The weights are used to produce population estimates from the sample results. A basic survey weight is attached to each sample record. The size of weights depends on the sample design. The weights reflect the probability of selection of:

- the person within their household
- the household within its PSU
- the PSU within its stratum.

The weights also include:

- an adjustment for non-response
- an adjustment factor to ensure population counts for age and sex correspond to benchmark population estimates.

---

<sup>27</sup> Whereas the first two bullet points have been implemented as at wave 3 (using imputation indicators in versioned files), Statistics NZ note that the last bullet has not since cohabitants and OSMs are treated identically for imputation purposes.

Estimation is a method used to calculate population values from sample results. Estimation in SoFIE is done as follows: a basic survey weight is attached to each record to indicate the probability of that unit being included in the sample. Two types of adjustment are then applied to the basic survey weights to improve the reliability of the survey estimates. The basic weights are first inflated to adjust for non-response, and then further adjusted to ensure that estimates on relevant population characteristics match known population totals. The population totals used for SoFIE are derived from population estimates produced by Statistics New Zealand's Demography Division for counts for different age-sex groups.

### *3.2.6.1 Longitudinal Weighting*

From a longitudinal perspective, the population of interest is the usually resident population of NZ living in private dwellings as at wave one. Longitudinal estimates describe what happens to this population over the life of the survey. The sample will be weighted up to represent the longitudinal population of interest, which is the wave one population. These weights will then be used in producing population estimates and modelling the longitudinal data.

The longitudinal weighting methodology aims to:

- compensate for unequal selection probabilities
- compensate for (unit) non-response
- make use of available auxiliary information about the population to improve estimates.

The following steps will be taken to assign weights to responding Original Sample Members:

- calculation of selection weights
- adjustment for non-responding households at wave one
- adjustment for non-responding people within responding households at wave one
- adjustment for panel attrition since wave one
- adjustment to population benchmarks using post-stratification.

The population benchmarks will be demographic estimates of the wave one population. The exact benchmarks are still to be decided<sup>28</sup>, however they are likely to include sex and five-year age benchmarks, and possibly Māori and non-Māori and/or region benchmarks also.

### *3.2.6.2 Cross-sectional Weighting*

There are three key differences between cross-sectional and longitudinal outputs that influence the cross-sectional weighting procedure:

- Cohabitants will be included in the cross-sectional data files.

---

<sup>28</sup> Currently (wave 3), Statistics NZ use sex by 5 year age groups as benchmarks, but this choice is likely to be reviewed in later waves

- The cross-sectional outputs relate to the population as at wave  $t$ .
- Cross-sectional estimates of families and households will be produced.

It is intended that responding OSMs from partially responding households will be kept in the longitudinal outputs, primarily to reduce sample loss due to attrition. When it comes to producing the cross-sectional, outputs, though, it is only the individuals from the current wave's fully responding households that will be included.

The following steps will be taken to assign cross-sectional weights to households, families and individuals:

- calculate selection weights
- adjustment for non-responding households at wave one
- adjustment for non-responding people within responding households at wave one
- adjustment for panel attrition since wave one
- the weight share
- adjustment to population benchmarks using an integrated person-household weighting method.

### 3.2.7 Missing values and attrition

Attrition is the loss of respondents from the survey sample. The quality of longitudinal data through time is heavily affected by attrition as it occurs at each wave of the survey, resulting in fewer and fewer people who have complete records over the life of the survey. Minimising attrition is very important for a longitudinal survey.

Attrition from a longitudinal sample can be caused by:

- Death
- Emigration
- Movement into an institution
- Failure to trace respondents who have changed addresses between waves
- Failure to contact respondents
- Refusal to be interviewed.

The initial SoFIE sample comprised approximately 11,500 responding private households and 22,000 adults sampled within them, on a statistically representative basis from rural and urban areas throughout New Zealand. Information is collected from each member (including children) of a sampled household that falls within the scope of the survey and meets survey coverage rules. For wave one, approximately 77 percent of eligible households responded. In the second year of the survey, 87 percent of all respondents from wave one responded again, 91% of adults. In wave two there were just over 20,000 (91%) responding OSM adults. In wave three there were just over 18,200 responding OSM adults (90% of Wave 2 responders, 83% of Wave 1 responders)<sup>29</sup>.

---

<sup>29</sup> Statistics NZ note that response rates depend on the current output version. For example, response rates can change because all previous waves are reprocessed each time process the current wave is processed. Similarly, respondents who miss a wave may later be treated as responding for the period if they provide

### 3.2.8 Future work

It is worth noting that the SoFIE-Health team have expended considerable effort in making programs as generalisable and well-documented as possible to prepare for automation of the process. Robust processing systems also require careful attention to maintenance issues – such as the directory structure in which the ‘official’ version of the code is stored, who has responsibility for maintenance, and processes for updating and adding to the official code repository. Users are free, of course, to create their own software for specific applications.

## 4 Macros and other useful programs

There are a growing number of macros and other programs which have been written to support SoFIE-Health data processing (described above), to make access to research datasets easier for users, etc. This section describes some of these programs. Later versions of this report will provide more complete coverage of SoFIE-Health user software.

### 4.1 *begin\_sofie.sas*

A setup script for users to run at the beginning of new SAS sessions. Most of the work is done by the *%sofsess* macro.

The recommended process is:

- Copy *begin\_sofie.sas* to a subdirectory of SoFusers, and turn off read-only access status
- modify the *%sofsess* macro ‘mycheck’ and ‘mywork’ arguments to point to user *checking/working* subdirectories, and run the first two lines of the script (i.e., the *%include* and *%sofsess* statements).
- Run the program and check for any error messages

### 4.2 *%sofsess.sas*

The *%sofsess* macro enables access to the catalog of compiled macros and the permanent formats dataset, and defines several librefs to enable access to waves 1 to 3 datasets.

These include:

- FEXPORT. References the directory in which users can store files to be checked by Statistics NZ for confidentiality purposes (currently `\\wsnz02\c_datalab\MAA2006-07 SoFIE A\Checking` or a subdirectory thereof)
- SOF13\_R. References the directory where Statistics NZ read-only datasets are stored.

---

the missing information next wave. Changes may also occur due to further editing and changes to the methodology used.

- SOF13. References the directory where permanent SoFIE-Health wide and long research datasets are stored (currently `\\wsnz02\c_datalab\MAA2006-07 SoFIE A\SASdata\sof13`).
- MYWORK. References the directory (and subdirectories thereof) where users can store their own datasets (currently `\\wsnz02\c_datalab\MAA2006-07 SoFIE A\SoFusers`).

`%sofsess` also sets up the default for printed SAS output. Administrators can invoke the macros that define formats for SoFIE-Health variables (`%soffmts`), and applies them to the long and wide research datasets (`%fmt_sof_vars`).

### 4.3 `%soffmts.sas`

Creates formats and informats that are used by the macro `%fmt_sof_vars` to format variables in the long and wide research datasets. Formats created by `%soffmts` are stored in a permanent format library referenced by the libref LIBRARY (which is itself created by the macro `%sofsess` and points to the directory `\\wsnz02\c_datalab\MAA2006-07 SoFIE A\SASProgs\sofie macros\soffmts`). By default, SAS always searches the library that is referenced by the LIBRARY libref for a FORMATS catalog. Both `%soffmts` and `%fmt_sof_vars` can be run by specifying the ACTIONS argument to `%sofsess`.

The suggested format/informat naming convention is:

```
<f(format)><format descriptor><number non-missing levels><N(from numeric)>
<f(format)><format descriptor><number non-missing levels><C(from char)>
<i(informat)><format descriptor><number non-missing levels><N(from numeric)>
<i(informat)><format descriptor><number non-missing levels><C(from char)>
```

where "format descriptor" is a 3-5 character string defining the format.

The guidelines for constructing format descriptors are:

- If the format can be applied generically to more than 1 variable then provide a generic descriptor of the format e.g., Marit for *MaritalStatus* and *SocMarital*, yn12 for labelling (1|2) as (y|n)
- If the format is specific to a variable, then provide a descriptor for the variable. Note: SoFIE vars are often constructed from several words so sometimes the capitalised first letters of all or some words are used e.g., CSD for *CigSmokePerDay* for , MReg for *MajorReg*.

### 4.4 `%fmt_sof_vars.sas`

Formats variables in the dataset defined by argument 'indata' using formats created by `%soffmts.sas` (stored in LIBRARY.formats, since SAS automatically searches that library).

Currently the program `create_research_dataset.sas` runs this macro on the long `Sofh_res_data` (transposed by `PersonCS`) and on the wide research dataset `Wsofh_res_data` (transposed by `PersonCS` and `PersonLT`).

## 5 Testing SoFIE wave 1-3 data

The SoFIE-Health project is one of the first to use a substantial subset of SoFIE data. The data are complex, and a considerable effort is currently required of researchers to process the files made available by Statistics NZ into a useable structure and format. This section presents results from exploratory analyses of variables in the wide research dataset designed to highlight apparent errors and inconsistencies in both Statistics NZ and UOW derived variables. The population used in the error investigation process is that subset of the SoFIE population (OSMs and non-OSMs) who responded in at least one wave. There are 34205 such people. Wave 1 was chosen as baseline for this analysis because it has been around the longest, and some initial errors have been corrected by Statistics NZ.

Important non-health variables in the final research dataset were studied in turn using cross-wave statistics, cross-tabulations (for discrete variables) and histograms and scatter-plots (for continuous variables). It is assumed that the characteristics captured by these statistics, tables or plots are reasonably constant across waves, and transitions between levels of discrete variables are ‘permitted’<sup>30</sup>, or at least reasonable given expected characteristics of the SoFIE sample.

Given the confidentiality requirement to replace cell counts less than a pre-defined cut-off with a fixed constant in data exported from the Statistics NZ DataLab, it is not possible to tell whether cells have zero or ‘small’ non-zero counts. The problem can be particularly severe within variables that have many levels (e.g., *HhldCompn2*, *HhldIncSrce*, *LongitFamSrce*) or ‘rare’ levels (e.g., *don’t know*, *refused*). This report therefore does not deal with potential quality problems that might be detected if this information was available. Unit record data is, of course, available in the DataLab, so future updates to this report may deal with the small cell replacement problem if required.

Future releases of this report will also summarise results from on-going testing of the SoFIE data.

### 5.1 Discrete variables

Attrition rates (see Section 3.2.7) suggest that missing responses will tend to increase across waves (see Carter *et al.*, 2008<sup>1</sup>). Wave 3 missing values for most variables tabulated above increase significantly relative to waves 1 and 2: Typically, an increase in missing values between waves 1 and 2 is around 14%, but about 70% between waves 1 and 3. Some variables have different missing value behaviours: For example, the increase in missing values for *BornInNZ* between waves 1 and 2(3) is 0.6% and 14.7% respectively. It is not clear why this occurs, though attrition certainly contributes. It is likely that variability in missing responses across waves has a significant contribution from non-OSM respondents who have not been excluded from the analysis. The choice of wave 1 as a base year for comparison with later waves may thus be problematic for a test of missing values, since everyone is an OSM in wave 1. Furthermore, it could be difficult to come to any definitive conclusion about the behaviour of missing values until more waves are available for comparison. This issue requires further investigation, e.g., by restricting the analyses reported here to wave 1 OSMs though this will focus attention on

---

<sup>30</sup> For example, a transition from an adult in wave 1 to a child in wave 2 would be an error (not permitted).

attrition and longitudinal dynamics. For the moment we emphasise that the population used for the tests reported here were respondents (OSMs and non-OSMS) with *EligibilityPerson = D* (i.e., provided valid responses) *in at least one of waves 1-3*.

With the exception of *MaxQualRank*, *LmActy*, *MaritalStatus*, *PaidWorkIn*, and *SchoolAttend*, the temporal behaviour of non-missing values for the discrete variables surveyed here appears unremarkable, although the shortness of the time series renders this conclusion preliminary.

The following issues have been identified for the exceptions:

- As noted above, *MaxQualRank* has many missing values because qualifications are not entered unless there has been an improvement in rank over the previous wave. A new variable, *MaxQualRank\_WSM*, derived by the SoFIE-Health project, corrects this problem.
- *LmActy* and *Health* have some non-missing values for children but are asked only of adults.
- *MaritalStatus* has a number of unlikely transitions (e.g., from *divorced*, *widowed*, *separated*, or *married* to *never*).
- *PaidWorkIn* and *SchoolAttend* appear to have significant problems.

## 5.2 Continuous variables

The only continuous variables that are considered in this report are personal and household income, largely because they are important for immediate SoFIE-Health research objectives.

### 5.2.1 PersIncAnn(DrvFamily)

This Statistics NZ derived variable appears to have a consistent distribution at waves 1 and 2. However the distribution of wave 3 income data is noticeably different, with a heavy low-income tail relative to waves 1 and 2. According to Statistics NZ, this problem has arisen because some wave 3 benefit income spell data have not been correctly accounted for in the income calculation, a problem that is currently being corrected.

The above problem flows into other derived income variables e.g., *HhldIncAnn*.

## 6 Additional Tasks Planned for Future Releases of the SoFIE-Health Technical Report

Previous sections of this report have noted where information is incomplete or in progress, and for various reasons could not be included in version 1. It is also planned to include sections on the following topics:

- Comparison with Statistics NZ public summaries of SoFIE data.
- Benchmarking against other datasets.
- SoFIE quality issues/limitations and reliability/error e.g.,
  - + Limitations, confidentiality issues, attrition.

- + Known quality issues (including problems with derived variables, duplicated variable names in different Statistics NZ datasets, changes in variable names by wave, documentation, ...).
  - + Causes.
  - + Impacts.
  - + Resolution plans.
- 
- Supporting Information.
  - Technical references.

## 7 APPENDICES

### 7.1 SoFIE-Health variable SAS formats

As noted in Section 4.3, formats for SoFIE-Health variables are created by the macro `%soffmts`. The formats created by that macro are listed below.

```
*KC: age1x;
value fAgey8N
  0 -<15      =      ' 0-14'
  15-<25     =      '15-24'
  25-<35     =      '25-34'
  35-<45     =      '35-44'
  45-<55     =      '45-54'
  55-<65     =      '55-64'
  65-<75     =      '65-74'
  75-high    =      '75+'
  other      =      'NA';
```

```
*KC: age2x;
value fAgey6N
  0 -<15      =      ' 0-14'
  15-<25     =      '15-24'
  25-<45     =      '25-44'
  45-<65     =      '45-64'
  65-<75     =      '65-74'
  75-high    =      '75+'
  other      =      'NA';
```

```
invalue iAgey8N
  0-<15      =      5
  15-<25     =      15
  25-<35     =      25
  35-<45     =      35
  45-<55     =      45
  55-<65     =      55
  65-<75     =      65
  75-high    =      75
  other = .;
```

```
invalue iAgey6N
  0-<15      =      5
  15-<25     =      15
  25-<45     =      25
  45-<65     =      45
  65-<75     =      65
  75-high    =      75
  other = .;
```

```

/*
Geographical
Sof13 vars: MajorReg
*/
*KC: regionx;
value fMReg6N      1      =      'Auckland'
                   2      =      'Waikato'
                   3      =      'Wellington'
                   4      =      'Rest of North Island'
                   5      =      'Canterbury'
                   6      =      'Rest of South Island'
                   other =      'NA';

/*
MARITAL STATUS
Sof13 vars: MaritalStatus SocMarital
*/
value fMarit6N 11      =      'Never Married'
                   12     =      'Divorced'
                   13     =      'Widowed'
                   14     =      'Separated'
                   15     =      'Married'
                   other =      'NA:DK:REF';

*KC: marital;
value fMarit4N 11      =      'Never Married'
                   12,13,14 =      'Divorced Widowed Separated'
                   15      =      'Married'
                   other   =      'NA:DK:REF';

/*
Sex
Sof13 vars: Sex Demog_Sex
*/
value fSex5N      11      =      'Male'
                   12      =      'Female'
                   88      =      'DK'
                   99      =      'REF'
                   other   =      'NA';

*KC: sexx;
value fSex3N      11      =      'Male'
                   12      =      'Female'
                   other   =      'NA:DK:REF';

```

```

/*
Labour market
Sof13 vars: LmActy
*/
*KC: lmactyx;
Value fLmA5N      1      =      'Working'
                  2      =      'Not employed, looking for work'
                  3      =      'Not employed, not looking for work'
                  4      =      'Overseas'
                  5      =      'Under 15'
                  other = 'NA';

/*
Education/Quals
Sof13 vars: MaxQualRank
*/
*KC: qual1x;
value fQual3N    10,11,12,13      =      'School Qual'
                  14,15,16,17,18,19,20 =      'Post school qual'
                  77              =      'No Qualification'
                  other           =      'NA:DK:REF';

*KC: qual1x;
value fQual5N    10,11,12,13      =      'School Qual'
                  14,15,16,17,18 =      'Post school vocational qual'
                  19,20          =      'Degree or Higher'
                  77              =      'No Qualification'
                  other           =      'NA:DK:REF';

/*
ETHNICITY
Sof13 vars: EthGrpPrior Hqethnicity_ethnicity DemogEthnicity
*/
*KC: priethx;
value fPeth5N    1      =      'NZ/European'
                  2      =      'Maori'
                  3      =      'Pacific'
                  4      =      'Asian'
                  5      =      'Other'
                  other = 'NA:DK:REF';

/*

```

```

COUNTRY OF BIRTH
Sof13 vars BornInNz
*/
*KC: bornnzx;
value $fBINz3C      'n' = 'No'
                   'u' = 'Unknown'
                   'y' = 'Yes'
                   other = 'NA';

/*GENERAL*/
value $fyn4C        'y' = 'Yes'
                   'n' = 'No'
                   '88' = 'DK'
                   '99' = 'REF'
                   other = 'NA';

value $fyn2C        'y' = 'Yes'
                   'n' = 'No'
                   other = 'NA:DK:REF';

value $fyn124C      "1" = 'Yes'
                   "2" = 'No'
                   "88" = 'DK'
                   "99" = 'REF'
                   other = 'NA';

value $fyn122C      "1" = 'Yes'
                   "2" = 'No'
                   other = 'NA:DK:REF';

value fyn124N        1 = 'Yes'
                   2 = 'No'
                   88 = 'DK'
                   99 = 'REF'
                   other = 'NA';

*KC: yn12x;
value fyn122N        1 = 'Yes'
                   2 = 'No'
                   other = 'NA:DK:REF';

*KC: yn01x, yn10x;
value fyn104N        1 = 'Yes'
                   0 = 'No'
                   88 = 'DK'
                   99 = 'REF'
                   other = 'NA';

value fyn102N        1 = 'Yes'
                   0 = 'No'
                   other = 'NA:DK:REF';

```

```

/*
NEW ZEALAND DEPRIVATION
NZIDep:          Sof Vars NZIDEPNUM
NZIDep:          (Derived) nzidep_group
*/
*KC: nzdep;
value fNzD5N     1,2 = 'NZDepQ1(least)'
                 3,4 = 'NZDepQ2'
                 5,6 = 'NZDepQ3'
                 7,8 = 'NZDepQ4'
                 9,10 = 'NZDepQ5(most)'
                 other = 'NA';

*KC: fnzidep;
value fNziD5N    1 = '0 dep vars'
                 2 = '1 dep vars'
                 3 = '2 dep vars'
                 4 = '3-4 dep vars'
                 5 = '5+ dep vars'
                 other = 'NA';

invalue iNziD5N  1,2 = 1
                 3,4 = 2
                 5,6 = 3
                 7,8 = 4
                 9,10 = 5
                 other = .;

/*
HEALTH
*/
/*Sof13 vars: Health*/
value fSRH7N     11 = 'Excellent'
                 12 = 'Very Good'
                 13 = 'Good'
                 14 = 'Fair'
                 15 = 'Poor'
                 88 = 'DK'
                 99 = 'REF'
                 other = 'NA';

*KC: $SRHx;
value fSRH5N     11 = 'Excellent'
                 12 = 'Very Good'
                 13 = 'Good'
                 14 = 'Fair'
                 15 = 'Poor'
                 other = 'NA:DK:REF';

/*Sof13 vars chronic_comorbid*/
*KC: $comorb;
value fCCoM3N    0 = "No comorbid"
                 1 = "1-2 comorbid"
                 2 = ">2 comorbid"
                 other = "NA";

```

```

/*Sof13 vars InjuryType*/
value $fInjT5N    11  =    "Traffic Injury"
                  12  =    "Sports Injury"
                  13  =    "Other Injury"
                  88  =    "DK"
                  99  =    "REF"
                  other =    "NA";

*KC: $injtypx;
value fInjT3N    11  =    "Traffic Injury"
                  12  =    "Sports Injury"
                  13  =    "Other Injury"
                  other =    "NA:DK:REF";

/*Sof13 vars InjuryWhere*/
*KC: $injwhex;
value fInjW3N    11  =    "Injury at home"
                  12  =    "Injury at work"
                  13  =    "Injury at another place"
                  other =    "NA:DK:REF";

value fInjW5N    11  =    "Injury at home"
                  12  =    "Injury at work"
                  13  =    "Injury at another place"
                  88  =    "DK"
                  99  =    "REF"
                  other =    'NA';

/*Sof13 vars Healthrelntype_ReInType*/
*KC: $relatx;
value fHRT8N     11  =    'Mother'
                  12  =    'Father'
                  13  =    'Stepmother'
                  14  =    'Stepfather'
                  15  =    'Grandparents'
                  16  =    'Siblings'
                  17  =    'Other relatives'
                  18  =    'Other unrelated'
                  other =    'NA';

```

```

/*Sof13 vars: FeltCheerless FeltDownhearted FeltEnergetic FeltHappy
FeltLively FeltPeaceful
FeltTired FeltVeryNervous FeltWornOut MentalLessCareful MentalLessDone
MentalLessTime
PhysicalDifficulty PhysicalLessdone PhysicalLessTime PhysicalLimited
probInterfereFreq*/

```

```

*KC: $timex;
value ftime5N    11 = 'All of the time'
                 12 = 'Most of the time'
                 13 = 'Some of the time'
                 14 = 'A little of the time'
                 15 = 'None of the time'
                 other = 'NA:DK:REF';

```

```

value ftime7N    11 = 'All of the time'
                 12 = 'Most of the time'
                 13 = 'Some of the time'
                 14 = 'A little of the time'
                 15 = 'None of the time'
                 88 = 'DK'
                 99 = 'REF'
                 other = 'NA';

```

```

/*Sof13 vars HealthNow*/

```

```

*KC: $ratel2x;
value fHNow5N    11 = 'Much better than 12 months ago'
                 12 = 'A bit better than 12 months ago'
                 13 = 'About the same as 12 months ago'
                 14 = 'A bit worse than 12 months ago'
                 15 = 'Much worse than 12 months ago'
                 other = 'NA:DK:REF';

```

```

value fHNow7N    11 = 'Much better than 12 months ago'
                 12 = 'A bit better than 12 months ago'
                 13 = 'About the same as 12 months ago'
                 14 = 'A bit worse than 12 months ago'
                 15 = 'Much worse than 12 months ago'
                 88 = 'DK'
                 99 = 'REF'
                 other = 'NA';

```

```

/*Sof13 vars AsHealthy ExpectWorsen HealthExcellent SickQuicker*/
*KC: $truthx;
value ftrue5N      11    =    'Definitely true'
                  12    =    'Mostly true'
                  13    =    'Neither true nor false'
                  14    =    'Mostly false'
                  15    =    'Definitely false'
                  other =    'NA:DK:REF';

value ftrue7N      11    =    'Definitely true'
                  12    =    'Mostly true'
                  13    =    'Neither true nor false'
                  14    =    'Mostly false'
                  15    =    'Definitely false'
                  88    =    'DK'
                  99    =    'REF'
                  other =    'NA';

/*Sof13 vars AsHealthy ExpectWorsen HealthExcellent SickQuicker*/
*KC: $painx;
value fpSev6N      11    =    'Very severe'
                  12    =    'Severe'
                  13    =    'Moderate'
                  14    =    'Mild'
                  15    =    'Very mild'
                  16    =    'No bodily pain'
                  other =    'NA:DK:REF';

value fpSev8N      11    =    'Very severe'
                  12    =    'Severe'
                  13    =    'Moderate'
                  14    =    'Mild'
                  15    =    'Very mild'
                  16    =    'No bodily pain'
                  88    =    'DK'
                  99    =    'REF'
                  other =    'NA';

```

```

/*
Sof13 vars: LimitBatheSelf LimitBendKneel LimitDressSelf LimitLiftCarry
LimitModerateActivity LimitStepsMultiset LimitStepsOneset
LimitVigorousActy
LimitWalk100m LimitWalk1km LimitWalk500m
*/
*KC: $limitx;
value flimt3N      11 = 'Yes limited a lot'
                  12 = 'Yes limited a little'
                  13 = 'No not limited'
                  other = 'NA:DK:REF';

value flimt5N      11 = 'Yes limited a lot'
                  12 = 'Yes limited a little'
                  13 = 'No not limited'
                  88 = 'DK'
                  99 = 'REF'
                  other = 'NA';

/*Sof13 vars: PainInterfereDeg ProbInterfereDeg*/
*KC: $degreex;
value fdeg5N      11 = 'Extremely'
                  12 = 'Quite a lot'
                  13 = 'Moderatley'
                  14 = 'A little bit'
                  15 = 'Not at all'
                  other = 'NA:DK:REF';

value fdeg7N      11 = 'Extremely'
                  12 = 'Quite a lot'
                  13 = 'Moderatley'
                  14 = 'A little bit'
                  15 = 'Not at all'
                  88 = 'DK'
                  99 = 'REF'
                  other = 'NA';

```

```

/*
Sof13 vars: HlthProvAfterHrs HlthprovCheckUp HlthProvCultSens
HlthProvImportant
HlthProvKnowBest HlthProvKnowWell HlthProvNewProb HlthProvPhAdvice
HlthProvSameDay
HlthProvSamepers
*/
*KC: $certx;
value fHPrv4N      11 = 'Definitely'
                  12 = 'Probably'
                  13 = 'Probably not'
                  14 = 'Definitely not'
                  other = 'NA:DK:REF';

value fHPrv6N      11 = 'Definitely'
                  12 = 'Probably'
                  13 = 'Probably not'
                  14 = 'Definitely not'
                  88 = 'DK'
                  99 = 'REF'
                  other = 'NA';

/*Sof13 vars: provtype*/
*KC: $pcptypx;
value fPTy10N      11 = 'GP'
                  12 = 'Maori Provider'
                  13 = 'Pacific Provider'
                  14 = 'Health Centre'
                  15 = 'Nursing Service'
                  16 = 'A&E'
                  17 = 'A&M'
                  18 = 'Union Health Clinic'
                  19 = 'Other'
                  20 = 'None';

/*Sof13 vars: smoke_status, smokerstatus_snz(Drvhealth)*/
value fSmoke4N     1 = 'Current smoker'
                  2 = 'Ex-smoker'
                  3 = 'Never smoker'
                  4 = '???'
                  other = 'NA';

*KC: $smokex;
value fSmoke3N     1 = 'Current smoker'
                  2 = 'Ex-smoker'
                  3 = 'Never smoker'
                  other = 'NA';

```

```
/*Sof13 vars: CigSmokePerDay*/
```

```
value fCSD12N    11 = '<5/day'  
                 12 = '5-9/day'  
                 13 = '10-14/day'  
                 14 = '15-19/day'  
                 15 = '20-24/day'  
                 16 = '25-29/day'  
                 17 = '30-34/day'  
                 18 = '35-39/day'  
                 19 = '40-44/day'  
                 20 = '>45/day'  
                 88 = 'DK'  
                 99 = 'REF'  
                 other = 'NA';
```

```
*KC: $cigsx;
```

```
value fCSD10N    11 = '<5/day'  
                 12 = '5-9/day'  
                 13 = '10-14/day'  
                 14 = '15-19/day'  
                 15 = '20-24/day'  
                 16 = '25-29/day'  
                 17 = '30-34/day'  
                 18 = '35-39/day'  
                 19 = '40-44/day'  
                 20 = '>45/day'  
                 other = 'NA:DK:REF';
```

```
/*Sof13 vars: DrinkLastType DrinkMostType DaysDrink DrinksPerDay  
OccManyDrinks alc_drink alc_num*/
```

```
value fDT11N     11 = 'Low alcohol beer'  
                 12 = 'Normal beer'  
                 13 = 'Extra strength beer'  
                 14 = 'Home brew'  
                 15 = 'Wine'  
                 16 = 'RTDs'  
                 17 = 'Low alcohol spirits'  
                 18 = 'Full strength spirits'  
                 19 = 'Other'  
                 88 = 'DK'  
                 99 = 'REF'  
                 other = 'NA';
```

```
*KC: $alcx;
```

```
value fDT9N 11   = 'Low alcohol beer'  
                 12 = 'Normal beer'  
                 13 = 'Extra strength beer'  
                 14 = 'Home brew'  
                 15 = 'Wine'  
                 16 = 'RTDs'  
                 17 = 'Low alcohol spirits'  
                 18 = 'Full strength spirits'  
                 19 = 'Other'  
                 other = 'NA:DK:REF';
```

```

*KC: fdrnknm;
value fDD4N 1-4          = '<1 spell / wk'
                        5-8          = '1-2 spells / wk'
                        9-12         = '2-3 spells / wk'
                        13-high      = '3+ spells / wk'
                        other        = 'NA';

*KC: fdrnkqn;
value fDpD4N            1-2          = '1-2 / day'
                        3-4          = '3-4 / day'
                        5-6          = '5-6 / day'
                        7-8          = '7-8 / day'
                        9-high       = '9+ / day'
                        other        = 'NA';

*KC: fbingenm;
value fOMD4N            0            = '0 /4wks'
                        1-2          = '1-2 / 4wks'
                        3-4          = '3-4 / 4wks'
                        5-6          = '5-6 / 4wks'
                        7-high       = '7+ / 4wks'
                        other        = 'NA';

*KC: fauddrk;
value fAAD5N            0            = 'Never'
                        1            = '< monthly'
                        2            = '2-4 times / month'
                        3            = '2-3 times / week'
                        4            = '4 or more times / week'
                        other        = 'NA';

*KC: faudnum - no vars with this format?;
value fANum5N            0            = '1-2 / occasion'
                        1            = '3-4 / occasion'
                        2            = '5-6 / occasion'
                        3            = '7-9 / occasion'
                        4            = '10+ / occasion'
                        other        = 'NA';

*KC: faudbin - alc_num;
value fAAN5N            0            = 'Never binge'
                        1            = 'Binge monthly'
                        2            = 'Binge 2 times / month'
                        3            = 'Binge weekly'
                        4            = 'Binge daily or almost daily '
                        other        = 'NA';

```

```
/******  
Additional formats for Non-Health variables (MH 11/10/07)  
******/
```

```
* Marital Status;  
* Used in MaritalStatus;  
value fMari7N    11    = 'Never Married'  
                  12    = 'Divorced'  
                  13    = 'Widowed'  
                  14    = 'Separated'  
                  15    = 'Married'  
                  88    = 'DK'  
                  99    = 'REF'  
other            = 'NA';
```

```
* Ethnicity;  
* Used in EthnicityOccrrence;  
value fEth17N    11    = 'European'  
                  12    = 'Other European'  
                  13    = 'NZ Maori'  
                  14    = 'Samoan'  
                  15    = 'Cook Islands'  
                  16    = 'Tongan'  
                  17    = 'Niuean'  
                  18    = 'Tokelauan'  
                  19    = 'Fijian'  
                  20    = 'Other Pacific'  
                  21    = 'Asian'  
                  22    = 'Chinese'  
                  23    = 'Indian'  
                  24    = 'Other Asian'  
                  25    = 'Other'  
                  88    = 'DK'  
                  99    = 'REF'  
other            = 'NA';
```

```
* Type of Dwelling Tenure;  
* Used in DwellTenure;  
value fDwel6N    1      = 'Owned with mortgage'  
                  2      = 'Owned without mortgage'  
                  3      = 'Occupied rent-free'  
                  4      = 'Rented or leased'  
                  5      = 'Dwelling tenure unspecified'  
other            = 'NA';
```

```
* Region;  
* Used in MajorReg;  
value fRegi6N    1      = 'Auckland'  
                  2      = 'Waikato'  
                  3      = 'Wellington'  
                  4      = 'Rest of North Island'  
                  5      = 'Canterbury'  
                  6      = 'Rest of South Island'  
other            = 'NA';
```

```

* Highest Education Qualification;
* Used in MaxQualRank;
value fHEQ13N    10    = 'Other School Qualification'
                  11    = 'School Certificate/Fifth Form Qualification'
                  12    = 'Sixth Form Qualification'
                  13    = 'Higher School Qualification'
                  14    = 'Post-school qualification not identifiable'
                  15    = 'Basic Vocational Qualification'
                  16    = 'Skilled Vocational Qualification'
                  17    = 'Intermediate Vocational Qualification'
                  18    = 'Advanced Vocational Qualification'
                  19    = 'Bachelor degree'
                  20    = 'Higher degree'
                  77    = 'No qualification'
                  78    = 'REF:DK'
other           = 'NA';

* Social Marital Status;
* Used in SocMarital;
value fSMar3N    1      = 'Partnered'
                  2      = 'Non-partnered'
                  9      = 'Not elsewhere included'
other           = 'NA';

* Household composition;
* Used in HhldCompn2;
value fHHC11N    11      = 'Couple only'
                  12      = 'Couple with others'
                  13      = 'Couple with child(ren)'
                  14      = 'Couple with child(ren) and others'
                  15      = 'One parent with child(ren)'
                  16      = 'One parent with child(ren) and others'
                  20      = 'Two families'
                  30      = 'Three or more families'
                  40      = 'Other multiperson household'
                  50      = 'One person household'
                  90      = 'Household not elsewhere classified'
other           = 'NA';

* Income source;
* Used in HhldIncSrce, LongitFamSrce;
value fInS11N    10      = 'Employment earnings'
                  11      = 'Self-employment'
                  12      = 'NZ Superannuation'
                  13      = 'Other government income support'
                  14      = 'Private superannuation & pension schemes'
                  15      = 'Interest & other investments'
                  16      = 'All other sources'
                  17      = 'No single source of income'
                  18      = 'Self-employment loss'
                  19      = 'Investment losses'
                  20      = 'Self-employment & investment losses'
other           = 'NA';

```

```

* Household eligibility status;
* Used in EligibilityHhld;
value $FHES6C      'A' = 'Ineligible pre-contact household'
                   'B' = 'Ineligible post-contact household'
                   'C' = 'Non-responding household'
                   'D' = 'Full responding household'
                   'E' = 'Unknown eligible household'
                   'F' = 'Partial responding household'
                   other = 'NA';

* Person eligiblity status;
* Used in EligibilityPerson;
value $FPES4C      'B' = 'Out of scope'
                   'C' = 'Non-response'
                   'D' = 'Response'
                   'E' = 'Unknown eligiblity'
                   other = 'NA';

* Original sample member;
* Used in Osm;
value fOSMe2N      0      = 'Not an original sample member'
                   1      = 'Original Sample Member'
                   other = 'NA';

/*****
Formats and informats for income percentiles
*****/
value fPI5N low    -<    7602.18106 = "q1: low -< 7602"
                   7602.18106 -<    14310.61616 = "q2: 7602-< 14311"
                   14310.61616 -<    25652.12451 = "q3: 14311 -< 25652"
                   25652.12451 -<    41486.67558 = "q4: 25652 -< 41487"
                   41486.67558 -      high      = "q5: 41487 -high";
value fHJ5N low    -<    21077.94532 = "q1: low -< 21078"
                   21077.94532 -<    34009.74742 = "q2: 21078 -< 34010"
                   34009.74742 -<    49379.34645 = "q3: 34010 -< 49379"
                   49379.34645 -<    72280.40562 = "q4: 49379 -< 72280"
                   72280.40562 -      high      = "q5: 72280 -< high";
value fHJA5N low   -<    21499.70120 = "q1: low   -< 21500"
                   21499.70120 -<    34266.81327 = "q2: 21500 -< 34267"
                   34266.81327 -<    49666.51538 = "q3: 34267 -< 49667"
                   49666.51538 -<    71975.34656 = "q4: 49667 -< 71975"
                   71975.34656 -      high      = "q5: 71975 -high";
value fHLx5N low   -<    14749.31594 = "q1: low   -< 14749"
                   14749.31594 -<    24121.76431 = "q2: 14749 -< 24122"
                   24121.76431 -<    35156.65710 = "q3: 24122 -< 35157"
                   35156.65710 -<    51545.79975 = "q4: 35157 -< 51546"
                   51545.79975 -      high      = "q5: 51546 -high";
value fHI5N low    -<    24645.69062 = "q1: low -< 24646"
                   24645.69062 -<    43480.95504 = "q2: 24646 -< 43481"
                   43480.95504 -<    63189.43142 = "q3: 43481 -< 63189"
                   63189.43142 -<    92165.40190 = "q4: 63189 -< 92165"
                   92165.40190 -      high      = "q5: 92165 - high";

```

```

value fW5N  low      -< 4898.708978 = "q1: low -< 4899"
            4898.708978 -< 34976.99443 = "q2: 4899-< 34977"
            34976.99443 -< 92841.25536 = "q3: 34977 -< 92841"
            92841.25536 -< 195924.1142 = "q4: 92841 -< 195924"
            195924.1142 - high      = "q5: 195924 - high";

value fPI3N low      -< 11671.47347 = "t1: low -< 11671"
            11671.47347 -< 30165.87566 = "t2: 11671-< 30166"
            30165.87566 - high      = "t3: 30166-high";
value fHJ3N low      -< 29195.12645 = "t1: low -< 29195"
            29195.12645 -< 55536.84294 = "t2: 29195 -< 55536"
            55536.84294 - high      = "t3: 55536-< high";
value fHJA3N low     -< 29618.78166 = "t1: low -< 29619"
            29618.78166 -< 55634.89763 = "t2: 29619 -< 55635"
            55634.89763 - high      = "t3: 55635 -high";
value fHLx3N low     -< 20660.38576 = "t1: low -< 20660"
            20660.38576 -< 39703.18173 = "t2: 20660 -< 39703"
            39703.18173 - high      = "t3: 39703 -high";
value fHI3N low      -< 37671.63864 = "t1: low -< 37672"
            37671.63864 -< 70756.23363 = "t2: 37672 -< 70756"
            70756.23363 - high      = "t3: 70756 - high";
value fW3N  low      -< 21435.79218 = "t1: low -< 21436"
            21435.79218 -< 119556.9952 = "t2: 21436 -< 119557"
            119556.9952 - high      = "t3: 119557 - high";

value fPI10N low     -< 2441.573731 = "d1: low -< 2442"
            2441.573731 -< 7602.181060 = "d2: 2442-< 7602"
            7602.181060 -< 10613.09350 = "d3: 7602-< 10613"
            10613.09350 -< 14310.61616 = "d4: 10613 -< 14311"
            14310.61616 -< 19480.48283 = "d5: 14311 -< 19480"
            19480.48283 -< 25652.12451 = "d6: 19480 -< 25652"
            25652.12451 -< 32624.43561 = "d7: 25652 -< 32624"
            32624.43561 -< 41486.67558 = "d8: 32624 -< 41487"
            41486.67558 -< 56935.14264 = "d9: 41487 -< 56935"
            56935.14264 - high      = "d10: 56935 -high";
value fHJ10N low     -< 15569.39865 = "d1: low -< 15569"
            15569.39865 -< 21077.94532 = "d2: 15569 -< 21078"
            21077.94532 -< 27082.79884 = "d3: 21078 -< 27083"
            27082.79884 -< 34009.74742 = "d4: 27083 -< 34010"
            34009.74742 -< 41093.31399 = "d5: 34010 -< 41093"
            41093.31399 -< 49379.34645 = "d6: 41093 -< 49379"
            49379.34645 -< 58947.29746 = "d7: 49379 -< 58947"
            58947.29746 -< 72280.40562 = "d8: 58947 -< 72280"
            72280.40562 -< 96028.22157 = "d9: 72280 -< 96028"
            96028.22157 - high      = "d10: 96028 -high";

```

```

value fHJA10N low      -<    15665.19348 = "d1: low -< 15665"
      15665.19348 -<    21499.70120 = "d2: 15665 -< 21500"
      21499.70120 -<    27338.22759 = "d3: 21500 -< 27338"
      27338.22759 -<    34266.81327 = "d4: 27338 -< 34267"
      34266.81327 -<    41333.56141 = "d5: 34267 -< 41334"
      41333.56141 -<    49666.51538 = "d6: 41334 -< 49667"
      49666.51538 -<    59000.88364 = "d7: 49667 -< 59001"
      59000.88364 -<    71975.34656 = "d8: 59001 -< 71975"
      71975.34656 -<    95957.77675 = "d9: 71975 -< 95958"
      95957.77675 -      high      = "d10: 95958 -high";
value fHLx10N low     -<    10710.27645 = "d1: low -< 10710"
      10710.27645 -<   14749.31594 = "d2: 10710 -< 14749"
      14749.31594 -<   19186.83284 = "d3: 14749 -< 19187"
      19186.83284 -<   24121.76431 = "d4: 19187 -< 24122"
      24121.76431 -<   29183.90790 = "d5: 24122 -< 29184"
      29183.90790 -<   35156.65710 = "d6: 29184 -< 35157"
      35156.65710 -<   41991.73433 = "d7: 35157 -< 41992"
      41991.73433 -<   51545.79975 = "d8: 41992 -< 51546"
      51545.79975 -<   68504.85926 = "d9: 51546 -< 68505"
      68504.85926 -      high      = "d10: 68505 -high";
value fHI10N low      -<   16526.94997 = "d1: low -< 16527"
      16526.94997 -<   24645.69062 = "d2: 16527 -< 24646"
      24645.69062 -<   34601.92764 = "d3: 24646 -< 34602"
      34601.92764 -<   43480.95504 = "d4: 34602 -< 43481"
      43480.95504 -<   52872.11705 = "d5: 43481 -< 52872"
      52872.11705 -<   63189.43142 = "d6: 52872 -< 63189"
      63189.43142 -<   74902.50802 = "d7: 63189 -< 74903"
      74902.50802 -<   92165.40190 = "d8: 74903 -< 92165"
      92165.40190 -<  122560.5306 = "d9: 92165 -< 122561"
      122560.5306 -      high      = "d10: 122561 -high";
value fW10N low       -<    477.5784741 = "d1: low -< 478"
      477.5784741 -<   4898.708978 = "d2: 478 -< 4899"
      4898.708978 -<  15691.29086 = "d3: 4899 -< 15691"
      15691.29086 -<  34976.99443 = "d4: 15691 -< 34977"
      34976.99443 -<  62395.62764 = "d5: 34977 -< 62396"
      62395.62764 -<  92841.25536 = "d6: 62396 -< 92841"
      92841.25536 -< 134208.1476 = "d7: 92841 -< 134208"
      134208.1476 -< 195924.1142 = "d8: 134208 -< 195924"
      195924.1142 -< 313940.1408 = "d9: 195924 -< 313940"
      313940.1408 -      high      = "d10: 313940 -high";

```

\*Also create the corresponding informats for deriving incom/wealth ranks;

```

invalue iPI5N low     -<    7602.18106 = 1
      7602.18106 -<   14310.61616 = 2
      14310.61616 -<   25652.12451 = 3
      25652.12451 -<   41486.67558 = 4
      41486.67558 -      high      = 5;
invalue iHJ5N low     -<   21077.94532 = 1
      21077.94532 -<   34009.74742 = 2
      34009.74742 -<   49379.34645 = 3
      49379.34645 -<   72280.40562 = 4
      72280.40562 -      high      = 5;

```

```

invalue iHJA5N    low      -<    21499.70120 = 1
                  21499.70120 -<    34266.81327 = 2
                  34266.81327 -<    49666.51538 = 3
                  49666.51538 -<    71975.34656 = 4
                  71975.34656 -    high      = 5;
invalue iHLx5N    low      -<    14749.31594 = 1
                  14749.31594 -<    24121.76431 = 2
                  24121.76431 -<    35156.65710 = 3
                  35156.65710 -<    51545.79975 = 4
                  51545.79975 -    high      = 5;
invalue iHI5N     low      -<    24645.69062 = 1
                  24645.69062 -<    43480.95504 = 2
                  43480.95504 -<    63189.43142 = 3
                  63189.43142 -<    92165.40190 = 4
                  92165.40190 -    high      = 5;
invalue iW5N      low      -<    4898.708978 = 1
                  4898.708978 -<    34976.99443 = 2
                  34976.99443 -<    92841.25536 = 3
                  92841.25536 -<    195924.1142 = 4
                  195924.1142 -    high      = 5;

invalue iPI3N     low      -<    11671.47347 = 1
                  11671.47347 -<    30165.87566 = 2
                  30165.87566 -    high      = 3;
invalue iHJ3N     low      -<    29195.12645 = 1
                  29195.12645 -<    55536.84294 = 2
                  55536.84294 -    high      = 3;
invalue iHJA3N    low      -<    29618.78166 = 1
                  29618.78166 -<    55634.89763 = 2
                  55634.89763 -    high      = 3;
invalue iHLx3N    low      -<    20660.38576 = 1
                  20660.38576 -<    39703.18173 = 2
                  39703.18173 -    high      = 3;
invalue iHI3N     low      -<    37671.63864 = 1
                  37671.63864 -<    70756.23363 = 2
                  70756.23363 -    high      = 3;
invalue iW3N      low      -<    21435.79218 = 1
                  21435.79218 -<    119556.9952 = 2
                  119556.9952 -    high      = 3;

invalue iPI10N    low      -<    2441.573731 = 1
                  2441.573731 -<    7602.181060 = 2
                  7602.181060 -<    10613.09350 = 3
                  10613.09350 -<    14310.61616 = 4
                  14310.61616 -<    19480.48283 = 5
                  19480.48283 -<    25652.12451 = 6
                  25652.12451 -<    32624.43561 = 7
                  32624.43561 -<    41486.67558 = 8
                  41486.67558 -<    56935.14264 = 9
                  56935.14264 -    high      = 10;

```

```

invalue iHJ10N    low      -<    15569.39865 = 1
                  15569.39865 -<    21077.94532 = 2
                  21077.94532 -<    27082.79884 = 3
                  27082.79884 -<    34009.74742 = 4
                  34009.74742 -<    41093.31399 = 5
                  41093.31399 -<    49379.34645 = 6
                  49379.34645 -<    58947.29746 = 7
                  58947.29746 -<    72280.40562 = 8
                  72280.40562 -<    96028.22157 = 9
                  96028.22157 -     high      = 10;

invalue iHJA10N   low      -<    15665.19348 = 1
                  15665.19348 -<    21499.70120 = 2
                  21499.70120 -<    27338.22759 = 3
                  27338.22759 -<    34266.81327 = 4
                  34266.81327 -<    41333.56141 = 5
                  41333.56141 -<    49666.51538 = 6
                  49666.51538 -<    59000.88364 = 7
                  59000.88364 -<    71975.34656 = 8
                  71975.34656 -<    95957.77675 = 9
                  95957.77675 -     high      = 10;

invalue iHLx10N   low      -<    10710.27645 = 1
                  10710.27645 -<    14749.31594 = 2
                  14749.31594 -<    19186.83284 = 3
                  19186.83284 -<    24121.76431 = 4
                  24121.76431 -<    29183.90790 = 5
                  29183.90790 -<    35156.65710 = 6
                  35156.65710 -<    41991.73433 = 7
                  41991.73433 -<    51545.79975 = 8
                  51545.79975 -<    68504.85926 = 9
                  68504.85926 -     high      = 10;

invalue iHI10N    low      -<    16526.94997 = 1
                  16526.94997 -<    24645.69062 = 2
                  24645.69062 -<    34601.92764 = 3
                  34601.92764 -<    43480.95504 = 4
                  43480.95504 -<    52872.11705 = 5
                  52872.11705 -<    63189.43142 = 6
                  63189.43142 -<    74902.50802 = 7
                  74902.50802 -<    92165.40190 = 8
                  92165.40190 -<    122560.5306 = 9
                  122560.5306 -     high      = 10;

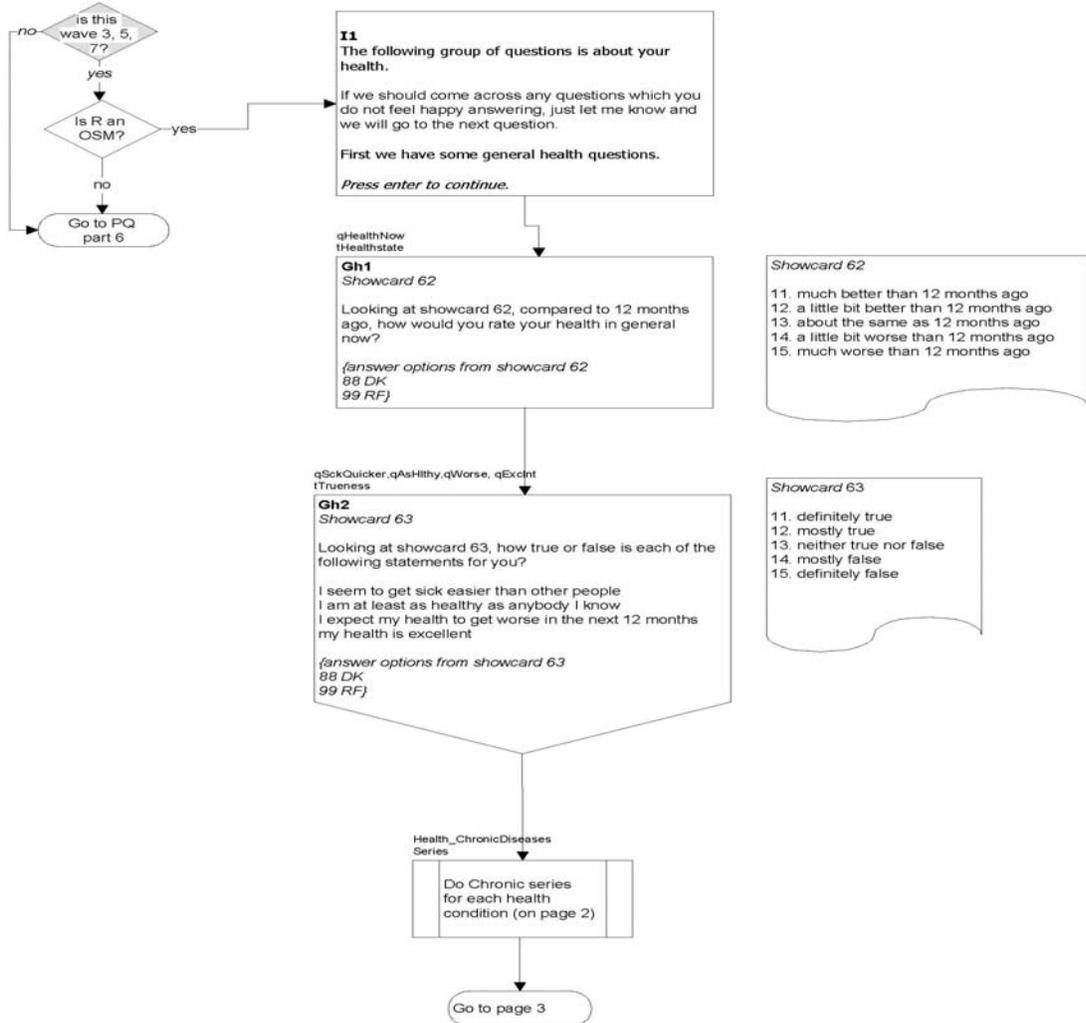
invalue iW10N     low      -<    477.5784741 = 1
                  477.5784741 -<    4898.708978 = 2
                  4898.708978 -<    15691.29086 = 3
                  15691.29086 -<    34976.99443 = 4
                  34976.99443 -<    62395.62764 = 5
                  62395.62764 -<    92841.25536 = 6
                  92841.25536 -<    134208.1476 = 7
                  134208.1476 -<    195924.1142 = 8
                  195924.1142 -<    313940.1408 = 9
                  313940.1408 -     high      = 10;

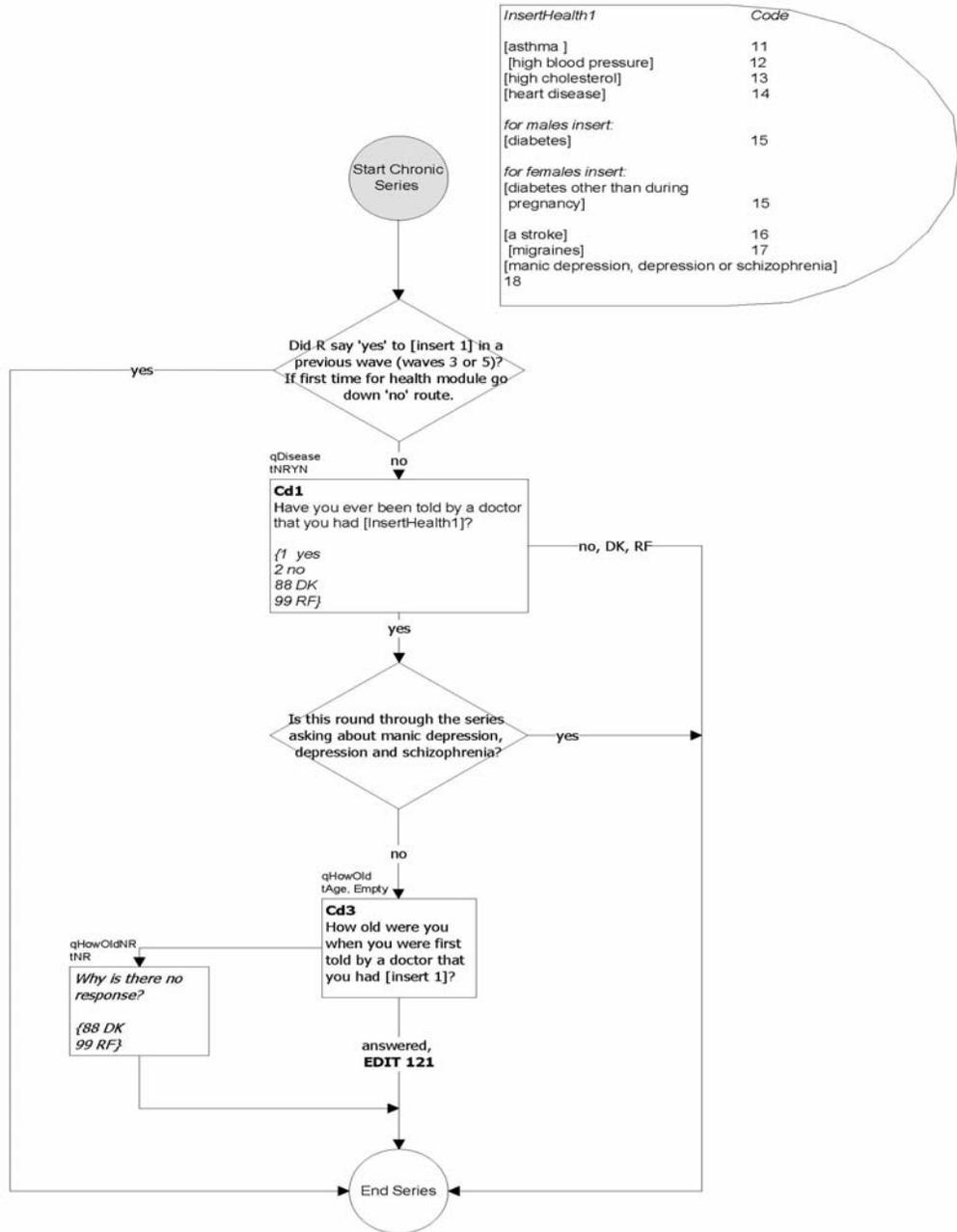
```

## 7.2 Statistics NZ SoFIE Health Questionnaire flowchart, pages 1-5

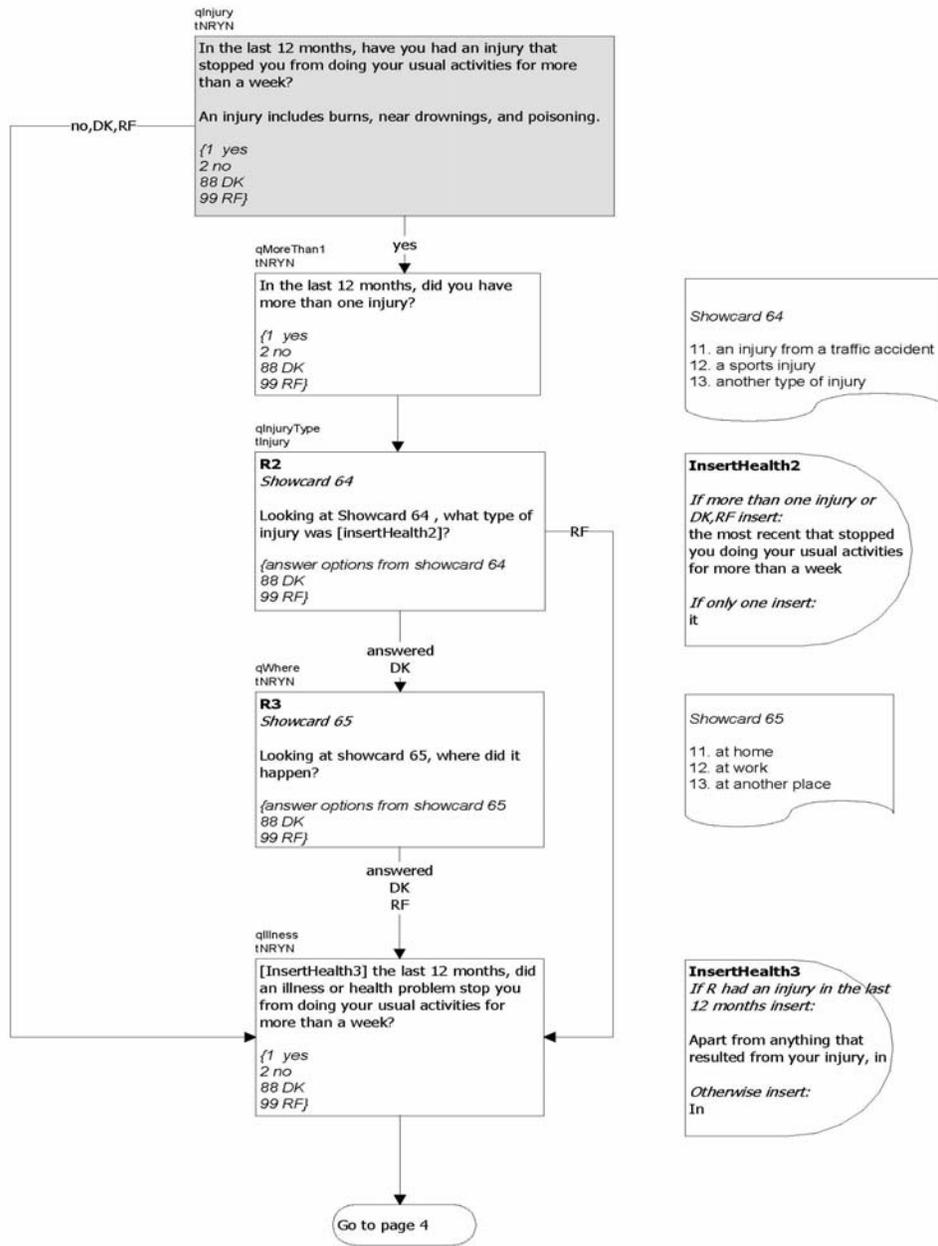
G:\SNZApps\Dev\LSID\Mainv1\Flowcharts\W3 consolidated flowcharts\PQ part 5 - Health.flo  
 Wednesday, 9 February 2005  
 3:24 pm

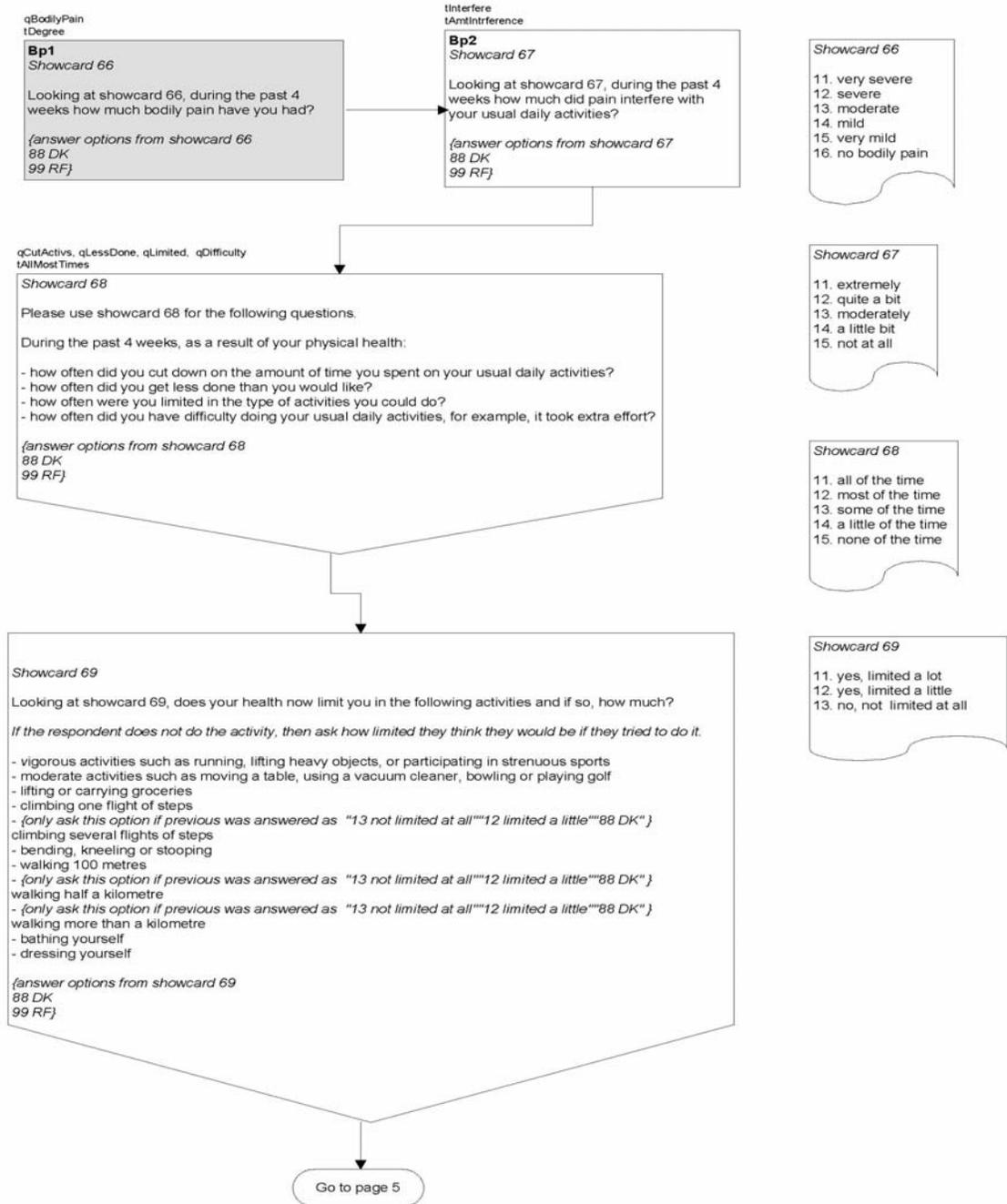
Introduction  
 SF36 General Health  
 Chronic Diseases





InsertHealth1	Code
[asthma ]	11
[high blood pressure]	12
[high cholesterol]	13
[heart disease]	14
for males insert: [diabetes]	15
for females insert: [diabetes other than during pregnancy]	15
[a stroke]	16
[migraines]	17
[manic depression, depression or schizophrenia]	18





qCutActvs,qGTLessDn,qLssCareff  
tAllMostTimes

**Showcard 70**

Please use showcard 70 for the following questions.

During the past 4 weeks, as a result of any emotional problems such as feeling depressed or anxious:

- how often did you cut down on the amount of time you spent on your usual daily activities?
- how often did you get less done than you would like?
- how often did you do your usual daily activities less carefully than usual?

{answer options from showcard 70  
88 DK  
99 RF}

**Showcard 70**

- 11. all of the time
- 12. most of the time
- 13. some of the time
- 14. a little of the time
- 15. none of the time

qTmSocActvs  
tAllMostTimes

**Sf2**  
**Showcard 70**

Looking at showcard 70, during the past 4 weeks, how much of the time have any physical or emotional problems interfered with your usual social activities with family, friends, neighbours or groups?

{answer options from showcard 70  
88 DK  
99 RF}

answered

qXtntsActvs  
tAmtInfrference

**Sf1**  
**Showcard 71**

Looking at showcard 71, during the past 4 weeks, to what extent have any physical or emotional problems interfered with your usual social activities?

{answer options from showcard 71  
88 DK  
99 RF}

**Showcard 71**

- 11. extremely
- 12. quite a bit
- 13. moderately
- 14. a little bit
- 15. not at all

Go to page 6

### 7.3 SoFIE data dictionary waves 1-2<sup>31</sup>

Dataset	Var_Name	Label
AssLiab		An asset or liability that a person has or may have (ie has responded that they have that asset or liability or have given a non-response).
AssLiab	AssLiabCnt	A count of the number of items with this sub-type that the person has.
AssLiab	AssLiabCntNr	Non-response to count of the number of items with this sub-type that the person has.
AssLiab	AssLiabSubType	A code to uniquely identify the sub-type of asset or liability within the type to which it belongs. The sub-type breaks the type down to a more specific level. A sub-type code of 10 is used for when the type does not break down into sub-types.
AssLiab	AssLiabType	A code to uniquely identify the type of asset or liability. The type classifies the asset or liability at a broad level.
AssLiab	EmplContrIncl	Whether the value of employer contributions were included in contributions to a superannuation scheme.
AssLiab	HasAssLiab	Whether the person has this asset or liability.
AssLiab	HowGrouped	How the items with this sub-type have been grouped for the purpose of recording their values, ie whether given separately for each item or for all of them combined.
AssLiab	OneOrMore	Whether the person has just one or more than one of this type/sub-type of asset or liability.
AssLiab	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
AssLiab	PropIncl	Whether the value of any residential property was included in the value(s) given for the asset.
AssLiab	SoleOwnerAll	Whether the person is the sole owner of all of the items with this sub-type that they own or partly own.
AssLiabDetail		Details of ownership and value for one or more asset or liability items of a particular sub-type. These details can be for an individual item, a grouping of items, or for all of that sub-type combined.
AssLiabDetail	AssLiabGrpType	A code to uniquely identify the group-type of asset or liability within the type to which it belongs. The group-type breaks the type down to a more specific level. A group-type code of 10 is used for when the type does not break down into group-types.

<sup>31</sup> There are differences between datasets and variables provided by Statistics NZ for waves 1-2, and 1-3. At the time of publication of this report, a data dictionary for waves 1-3 was not available.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
AssLiabDetail	AssLiabType	A code to uniquely identify the type of asset or liability. The type classifies the asset or liability at a broad level.
AssLiabDetail	CurrConvRate	The currency conversion rate to use to convert the income in foreign currency to NZ dollars. This is given as the amount in the foreign currency that equals a NZ dollar.
AssLiabDetail	EstMeth	How the estimate of the value was made
AssLiabDetail	ItemGrouping	A number to distinguish multiple groupings of items of the same sub-type for the same person. This has a value of zero when the grouping is for all items combined, otherwise the groupings are numbered sequentially starting from 1.
AssLiabDetail	NzDollars	Whether the amount quoted is in New Zealand dollars.
AssLiabDetail	OneOrAll	Indicates whether this grouping is of all items of this sub-type combined or for just one of those items. (If the person has only one item of this sub-type, the grouping is considered as being ONE rather than ALL.)
AssLiabDetail	OtherOwners	The number of other persons who jointly own the item with this person. (Only applies if the grouping consists of just one single item.)
AssLiabDetail	OtherOwnersNr	Non-response to the number of other persons who jointly own the item with this person.
AssLiabDetail	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
AssLiabDetail	PropValnYr	The year of the most recent valuation of any of the items in this grouping. (Only if the item is a residential property in New Zealand.)
AssLiabDetail	PropValnYrNr	Non-response to the year of the most recent valuation of any of the items in this grouping.
AssLiabDetail	SoleOwnerOne	Whether the person is the sole owner of this asset or liability item. (Only applies if the grouping consists of just one single item.)
AssLiabDetail	StdVal	The value of the grouping of items converted to a standard form. (If the value was given in foreign currency, and a currency conversion rate is included, the standard value has the conversion rate applied.)
AssLiabDetail	Val	The total value for this grouping of items.
AssLiabDetail	ValBand	The value band in the value band set that indicates the range of values that the estimated value lies in.
AssLiabDetail	ValBandSet	Indicates which set of value bands are used for when the value is given as a range.
AssLiabDetail	ValImp	Indicates whether and how the Value of the Asset or Liability is imputed.
AssLiabDetail	ValNr	Non-response to the total value for this grouping of items.
AssLiabMisc		Miscellaneous asset and liability details for a person.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
AssLiabMisc	FamTrstBeneficy	Whether the person is named as one of the people to benefit from a family trust.
AssLiabMisc	GiftedTrst	Whether the person has gifted money or other assets to a family trust. (Only applies if the person had not gifted any assets in their previous asset wave.)
AssLiabMisc	MaoriAssOwner	Whether the person is one of the owners or shareholders of Maori assets that belong to an iwi, hapu or whanau.
AssLiabMisc	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
AssLiabMisc	TrstOwned	Whether any of the assets gifted to any trusts during this wave are owned outright by the trusts.
AssLiabMisc	WillTrstBeneficy	Whether the person is named as one of the people to benefit from a trust set up by a will.
Case		A Case for which data was collected for the survey for one wave.(Cases do not have continuity across waves. If the same household is re-enumerated in a later wave, it is considered as being a different case.)
Case	Household	A randomised number assigned to uniquely identify the household
Case	HouseholdWave	The number of the wave in the wave sequence for the survey.
CaseVer		A version of a case that holds the complete set of information about the case as it is known for a particular output version.
CaseVer	Household	A randomised number assigned to uniquely identify the household
CaseVer	EAPD	End date of annual reference period (at 2400hrs on that date).
CaseVer	HED	Household enumeration date (at 0000hrs on that date).
CaseVer	HqiRespStatus	Household interview response status code. Derived from either HQI Response Status description or CAI Response Status description.
CaseVer	LastHhld1Chg	Last date of one-person household composition change.
CaseVer	LastHhld1ChgNr	Non-response to last date on one-person household composition change.
CaseVer	LastHhld2Chg	Last date of two-person household composition change.
CaseVer	LastHhld2ChgNr	Non-response to last date on twoe-person household composition change.
CaseVer	LastHhld3pChg	Last date of 3+ person household composition change.
CaseVer	LastHhld3pChgNr	Non-response to last date of 3+ person household composition change.
CaseVer	LatestHhldChg	Date of latest household composition change. (used from Wave 2 on.)

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
CaseVer	LatestHhldChgNr	Non-response to date of latest household composition change (from wave 2 on).
CaseVer	NewPeople	Whether there are any other people currently living with the answering person who started living with them since the HED of the previous wave.
CaseVer	NpWhenMovedIn	Date when the answering person started living at this address.
CaseVer	NpWhenMovedInNr	Non-response to date when the answering person started living at this address.
CaseVer	OutOfScopeMembers	Indicates whether someone in the household is out of scope.
CaseVer	PhoneYn	Indicates whether the household has a telephone.
CaseVer	PrivateYn	Check on whether this is a private dwelling.
CaseVer	RelatedMembersYn	Indicates whether any household members are related.
CaseVer	SAPD	Start date of the annual reference period (at 0000hrs on that day).
CaseVerPerson		One of the persons included in the Case Version. If a Case has multiple versions, each Version has its own set of Case Version Persons, though obviously multiple Case Version Persons are the same actual person.
CaseVerPerson	Age15plus	Whether person is aged 15 or more.
CaseVerPerson	AgeGiven	Given age of person.
CaseVerPerson	AgeGivenNr	Non-response to given age of person.
CaseVerPerson	Household	A randomised number assigned to uniquely identify the household
CaseVerPerson	Joined	Date when the reference person started living with the answering person.
CaseVerPerson	JoinedDateNr	Non-response to date when the reference person started living with the answering person.
CaseVerPerson	OutOfScopeMember	Whether the household member is out of scope.
CaseVerPerson	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
CaseVerPerson	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
CaseVerPerson	Sex	The sex of the person.
Demog		The demographic details of the person as at the interview date.
Demog	BornIn	The country of birth of the person.
Demog	Health	Self-rating of general health of the person.
Demog	MaritalStatus	Most recent legal marital status of the person.
Demog	PaidWorkIn	Whether the person is in paid work.
Demog	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
Demog	Sex	The sex of the person.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
Demog	SexImp	Indicates whether and how the Sex of the Person was imputed.
Demog	Wave1Eligibility	Whether the person lived in NZ at any time between the 1st of October 2002 and the 30th of September 2003.
Demog	WorkFullPart	Whether the person is working full-time or part-time.
DemogEthnicity		One of the ethnic groups that the person belongs to.
DemogEthnicity	DemogEthnicity	Code to identify the ethnic group.
DemogEthnicity	DemogEthnicityImp	Indicates whether and how the code to identify the ethnic group was imputed.
DemogEthnicity	EthnicityOccurrence	An occurrence number assigned to distinguish multiple ethnicity occurrences for the same Person. (This is used instead of the Ethnicity Code itself, as the latter can change, eg if a non-response code is replaced with an imputed value.)
DemogEthnicity	ImpGen	Indicates whether and how this row in the table was generated by the imputation process. This is usually as a result of imputing for a person who had a proxy response, and where therefore no information of this type was collected during the survey.
DemogEthnicity	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvAssLiab		Derived asset and liability variable for a person.
DrvAssLiab	AssLiabAmt	The amount for the asset/liability derived variable.
DrvAssLiab	AssLiabCode	A code to indicate which asset or liability derived variable is recorded.
DrvAssLiab	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvDemog		Derived Demography Variables about a person.
DrvDemog	Adult	Indicates whether the person is an adult (ie is aged 15 years or more) at Household Enumeration Date.
DrvDemog	AgeAtEapd	Age of the person at EAPD.
DrvDemog	AgeAtHed	Age of the person at HED.
DrvDemog	AgeAtWave1Sapd	Age of the person at SAPD of first wave.
DrvDemog	BornInNz	Indicates whether the person was born in NZ.
DrvDemog	EthGrpPrior	The prioritised ethnic group that the person belongs to.
DrvDemog	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvDemog	YrsInNz	The number of years since the person first arrived to live in NZ as a permanent or long-term resident (only if that person was not born in NZ).

Dataset	Var_Name	Label
DrvDwell		Derived Variables about the dwelling that a household lives in.
DrvDwell	Household	A randomised number assigned to uniquely identify the household
DrvDwell	DeprivIdx	The decile geographical ranking assigned to the geographical location of the dwelling of the household.
DrvDwell	DwellTenure	The type of tenure-of-dwelling status of the household.
DrvDwell	MajorReg	The major geographical region where the dwelling of the household is located.
DrvEarnGovtSpell		Derived Variables for an Earnings + Govt Transfers Spell (ie where the spells are based on changes of income for a person from Earnings plus Government Transfers).
DrvEarnGovtSpell	EarnGovtSpell	A unique number to distinguish an Earnings + Govt Transfers Spell from all other spells of this type for the same person.
DrvEarnGovtSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvEarnGovtSpell	SpellEnd	End date of the spell (at 2400 hrs).
DrvEarnGovtSpell	SpellStart	Start date of the spell (at 0000hrs).
DrvEarnGovtSpell	WklyEarnGovt	The usual gross weekly income of the person from earnings and/or government transfers during the Spell.
DrvEarnSpell		Derived Variables for an Earnings Spell (ie where Spells are formed based on changes to income of a person from earnings only).
DrvEarnSpell	EarnSpell	A unique number to distinguish an Earnings Spell from all other Earnings Spells for the same person.
DrvEarnSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvEarnSpell	SpellEnd	End date of the spell (at 2400 hrs).
DrvEarnSpell	SpellStart	Start date of the spell (at 0000hrs).
DrvEarnSpell	WklyEarn	The usual gross weekly income of a person from earnings during the Spell.
DrvEducation		Derived Educational Variables about a person.
DrvEducation	MaxQualRank	The highest educational qualification of the person at interview date.
DrvEducation	MnthStudyFull	The number of months during the annual reference period that the person was studying full-time towards a formally recognised qualification.
DrvEducation	MnthStudy	The number of months during the annual reference period that the person was studying towards a formally recognised qualification.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
DrvEducation	MnthStudyPart	The number of months during the annual reference period that the person was studying part-time towards a formally recognised qualification.
DrvEducation	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvFamAgesSpell		A Family Spell for a person by Age Categories of Children. Age Categories of the Children in the Family Nucleus of the person are constant for the duration of each spell, but change between spells if there is more than one spell in the same wave.
DrvFamAgesSpell	FamilyAge2	The Family Type of the person (Level 2) by Age Categories of Children in the family to which the person belonged.
DrvFamAgesSpell	FamilySpell	A unique number to distinguish a Family Spell from all other Family Spells for the same person.
DrvFamAgesSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvFamAgesSpell	SpellEnd	End date of the spell (at 2400 hrs).
DrvFamAgesSpell	SpellStart	Start date of the spell (at 0000hrs).
DrvFamAtomicSpell		An atomic family spell of the person. These spells are based on changes to the type of family that the person belonged to over time, including factors such as whether with a partner, children, one or two parents, siblings and ages of children.
DrvFamAtomicSpell	Child18plus	Indicates whether there were any children 18 years old or over in the Family Nucleus of the person.
DrvFamAtomicSpell	ChildUnder18	Indicates whether there were any children under 18 years of age in the Family Nucleus of the person.
DrvFamAtomicSpell	ChildUnknownAge	Indicates whether there were children of unknown age in the Family Nucleus of the person.
DrvFamAtomicSpell	FamilySpell	A unique number to distinguish a Family Spell from all other Family Spells for the same person.
DrvFamAtomicSpell	MinChildAge	The age of the youngest child in the family nucleus of the person (where a child is aged less than 18 years).
DrvFamAtomicSpell	NoFamNucleus	Indicates whether a person was categorised as being not in a family nucleus.
DrvFamAtomicSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvFamAtomicSpell	SpellEnd	End date of the spell (at 2400 hrs).
DrvFamAtomicSpell	SpellStart	Start date of the spell (at 0000hrs).
DrvFamAtomicSpell	With1parent	Indicates whether a person was categorised as being with 1 parent.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
DrvFamAtomicSpell	With2parents	Indicates whether a person was categorised as being with 2 parents.
DrvFamAtomicSpell	WithChild	Indicates whether a person was categorised as being with a child.
DrvFamAtomicSpell	WithPartnerInd	Indicates whether a person was categorised as being with a partner.
DrvFamAtomicSpell	WithSibling	Indicates whether a person was categorised as being with a sibling.
DrvFamily		Derived Variables about the marital status of a person at HED and how they relate to their Family Nucleus as at HED.
DrvFamily	Household	A randomised number assigned to uniquely identify the household
DrvFamily	ChildRoleInd	Indicates whether the person was in a child-role.
DrvFamily	DependStatus	Indicates the dependency status of the person (if flagged as being in a child role).
DrvFamily	FamilyEcon	The Economic Family Type of the Family of the person.
DrvFamily	FamilyNucleus	A unique number to distinguish a family nucleus from other family nuclei for the same household.
DrvFamily	FamilyRole	The role of the person in their family nucleus.
DrvFamily	ParentRole	Indicates whether the person was in a parent-role.
DrvFamily	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvFamily	SocMarital	The Social Marital Status of the person.
DrvFamily	WithPartnerInd	Indicates whether a person was categorised as being with a partner.
DrvFamilyNucleus		Derived Variables about a Family Nucleus of a Household as at Household Enumeration Date.
DrvFamilyNucleus	Household	A randomised number assigned to uniquely identify the household
DrvFamilyNucleus	ChildCntDepend	Categorisation of the number of dependent children in the family nucleus.
DrvFamilyNucleus	ChildDependent	The number of person in the family nucleus flagged as being in a child role and as dependent.
DrvFamilyNucleus	ChildIndependent	The number of person in a family nucleus flagged as being in a child role and as independent.
DrvFamilyNucleus	ChildRole	The number of persons in the family nucleus flagged as being in a child-role.
DrvFamilyNucleus	DependUnknown	The number of person in a family nucleus flagged as being in a child role and as dependency unknown.
DrvFamilyNucleus	FamilyNucleus	A unique number to distinguish a family nucleus from other family nuclei for the same household.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
DrvFamilyNucleus	FamilyStd2	The Standard Family Type (Level 2) for the person.
DrvFamilyNucleus	MinChildAge	The age of the youngest child in the family nucleus of the person (where a child is aged less than 18 years).
DrvFamilyNucleus	NoRole	The number of persons in the family nucleus not flagged as being in a child-role or in a parent role.
DrvFamilyNucleus	WithPartner	The number of persons in the family nucleus flagged as being with a partner.
DrvGovtIncome		Derived Variables about Government Transfers to a Person for a particular type/sub-type at Wave Level (ie as totals over the entire wave).
DrvGovtIncome	GovtAnn	The total annual income received by the person for this type/sub-type of regular government transfer.
DrvGovtIncome	GovtIncomeSubType	The type of government income broken down to a more specific sub-type.
DrvGovtIncome	GovtIncomeType	The type of government income at a broad level.
DrvGovtIncome	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvGovtIncomeSpell		Longitudinal Spell for a Type/SubType of Government Income.
DrvGovtIncomeSpell	GovtIncomeSpell	Spell number identifying this longitudinal government income spell out of all spells for the same type/sub-type of government income for this person.
DrvGovtIncomeSpell	GovtIncomeSubType	The type of government income broken down to a more specific sub-type.
DrvGovtIncomeSpell	GovtIncomeType	The type of government income at a broad level.
DrvGovtIncomeSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvGovtIncomeSpell	SpellEnd	End date of the spell (at 2400 hrs).
DrvGovtIncomeSpell	SpellStart	Start date of the spell (at 0000hrs).
DrvGovtIncomeSpell	WksGovtSpell	The number of weeks of a spell of government income support.
DrvHhldCompn		Derived Variables about Household Composition at Wave Level.
DrvHhldCompn	Household	A randomised number assigned to uniquely identify the household
DrvHhldCompn	HhldCompn2	Standard household composition (Level 2) for the household.
DrvHhldCompn	HhldCompnChg	The date of the most recent household composition change for the household.
DrvHhldIncome		Derived Variables for Income at Household Level.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
DrvHhldIncome	Household	A randomised number assigned to uniquely identify the household
DrvHhldIncome	HhldIncAnn	The Household Annual Income (ie the total income received from all sources by the household over the annual reference period).
DrvHhldIncome	HhldIncSrce	The main source of Household Annual Income of a household.
DrvIncome		Derived Variables about the Income of a person at wave level.
DrvIncome	EmplEarnAnn	The total gross earnings received by the person as a paid employee over the annual reference period.
DrvIncome	GovtIncSpells	The number of spells of government income support that the person had over the annual reference period.
DrvIncome	GovtXferAnn	The total gross and non-taxable government transfer income from all government transfers received by the person over the annual reference period
DrvIncome	LongitFamAnn	The Longitudinal Annual Family Income of a Person (ie the total income received from all sources by the family nucleus of that person as at the end of the annual period).
DrvIncome	LongitFamSrce	The main source of Longitudinal Annual Family Income of the Person.
DrvIncome	PersIncAnn	The total income received by the person from all sources over the annual reference period.
DrvIncome	PersIncSrce	The main source of the personal annual income of the person over the annual reference period.
DrvIncome	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvIncome	SubFamAnn	The Longitudinal Sub-Annual Family Income for the person.
DrvIncome	WksGovtAnn	The number of weeks that the person received government income support over the annual reference period.
DrvIncome	XsectFamAnn	The Cross-Sectional Annual Family Income of the Person (ie the total income received from all sources by the family nucleus of that person as at the Household Enumeration Date).
DrvIncome	XsectFamSrce	The main source of Cross-Sectional Annual Family income of the person.
DrvIncomeSpell		Derived Variables for an Income Spell (ie where spells are based on changes to income of a person from earnings + Govt Transfers + Regular Income).
DrvIncomeSpell	IncomeSpell	A unique number to distinguish an Income Spell from all other spells of this type for the same person.
DrvIncomeSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
DrvIncomeSpell	SpellEnd	End date of the spell (at 2400 hrs).
DrvIncomeSpell	SpellStart	Start date of the spell (at 0000hrs).
DrvIncomeSpell	WklyFamInc	Total average weekly family income of the person.
DrvIncomeSpell	WklyPersInc	The usual gross weekly income of the person from earnings and/or government transfers.
DrvLmActySpell		Derived variables for a Spell of Labour Market Activity.
DrvLmActySpell	LmActy	The type of Labour Market Activity of the person.
DrvLmActySpell	LmActySpell	A unique number to distinguish a Labour Market Activity Spell from other such spells for the same person.
DrvLmActySpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvLmActySpell	SpellEnd	End date of the spell (at 2400 hrs).
DrvLmActySpell	SpellStart	Start date of the spell (at 0000hrs).
DrvLmWave		Derived Variables about Labour Market aspects of a person at wave level.
DrvLmWave	HrsWorkTotal	Total hours worked by the person in all jobs over the annual reference period.
DrvLmWave	LmActy	The type of Labour Market Activity of the person.
DrvLmWave	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvLmWave	SpellsEmpl	Number of spells that the person had with an employer over the annual reference period.
DrvLmWave	SpellsLooking	Number of spells not employed but looking for work over the annual reference period.
DrvLmWave	SpellsOutOf	Number of spells not employed and not looking for work over the annual reference period.
DrvLmWave	WklyHrsWork	Total number of hours worked per week by the person over all jobs.
DrvLmWave	WksAveEmpl	Average number of weeks of spells that the person had with an employer over the annual reference period.
DrvLmWave	WksAveLooking	Average number of weeks of spells not employed but looking for work over the annual reference period.
DrvLmWave	WksAveOutOf	Average number of weeks of spells not employed and not looking for work over the annual reference period.
DrvLmWave	WksPaidTotal	The total number of weeks that the person worked in paid employment in Labour Market Spells over the annual reference period.
DrvLmWave	WksTotalEmpl	Total number of weeks that the person had with an employer over the annual reference period.
DrvLmWave	WksTotalLooking	Number of weeks not employed but looking for work over the annual reference period.
DrvLmWave	WksTotalOutOf	Number of weeks not employed and not looking for work over the annual reference period.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
DrvLmWave	YrsPaidEmpl	The total number of years that the person had spent in paid employment by the time of their first interview.
DrvSocMaritalSpell		Derived Variables about Social Marital Status at Spell level.
DrvSocMaritalSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvSocMaritalSpell	SocMarital	The Social Marital Status of the person.
DrvSocMaritalSpell	SocMaritalSpell	A unique number to distinguish this Social Marital Status Spell from other Social Marital Status Spells for the same person.
DrvSocMaritalSpell	SpellEnd	End date of the spell (at 2400 hrs).
DrvSocMaritalSpell	SpellStart	Start date of the spell (at 0000hrs).
DrvWorkSpell		Derived Variables for a Work Spell (ie a Labour Market Time Line Spell where the person did work).
DrvWorkSpell	EmplStatus	Status in employment over a Work Spell.
DrvWorkSpell	FullTimeWork	Indicates whether the person worked full-time during a Work Spell.
DrvWorkSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
DrvWorkSpell	TimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
DrvWorkSpell	WksPaid	The total number of weeks worked in paid employment.
DrvWorkSpell	WorkUsualHrs	Indicates whether the person worked usual weekly hours in a Work Spell.
Dwell		Details of the Dwelling of the Household.
Dwell	Bedrooms	The number of bedrooms in the dwelling.
Dwell	BedroomsNr	Non-response to the number of bedrooms in the dwelling.
Dwell	BodyCorpInclLand	Indicates whether land rates were included in the last body corporate payment.
Dwell	BodyCorpInclRates	Indicates whether council rates were included in the last body corporate payment.
Dwell	BodyCorpInclWater	Indicates whether water rates were included in the last body corporate payment.
Dwell	BodyCorpPay	Indicates whether any person in the household makes body corporate payments for the dwelling.
Dwell	Household	A randomised number assigned to uniquely identify the household
Dwell	DwellingOwnership	Indicates whether the dwelling is owned by any person in the household.
Dwell	FlexiLoan	Indicates whether the mortgage is revolving credit/flexiloan.
Dwell	MgePay	Indicates whether any person in the household makes mortgage payments for the dwelling.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
Dwell	MultiDwelling	Indicates whether the dwelling is part of a multi-dwelling block.
Dwell	RatesInclWater	Indicates whether council rates include water rates.
Dwell	RatesPay	Indicates whether any person in the household pays council rates for the dwelling.
Dwell	RentInclFood	Whether food is included in the rent payments for a non-private dwelling.
Dwell	RentPay	Indicates whether any person in the household pays rent for the dwelling.
Dwell	Vehicles	The number of vehicles available to members of the household.
Dwell	VehiclesNr	Non-response to the number of vehicles available to members of the household.
DwellAppliance		Indicates whether the dwelling has an appliance in working order of a particular type.
DwellAppliance	ApplianceType	A code to indicate the type of appliance.
DwellAppliance	Household	A randomised number assigned to uniquely identify the household
DwellPayment		Details of a payment made with respect to the dwelling of the household.
DwellPayment	Household	A randomised number assigned to uniquely identify the household
DwellPayment	DwellPay	The amount of the last payment of this type.
DwellPayment	DwellPayNr	Non-response to last payment of this type.
DwellPayment	DwellPayPeriod	The period covered by the last payment of this type.
DwellPayment	DwellPayType	A code to indicate the type of dwelling payment for which details of the last payment are being given.
DwellPayment	DwellPayWeeks	Number of weeks covered by the last payment of this type when the Period was specified as OTHER.
DwellPayment	DwellPayWeeksNr	Non-response to number of weeks covered by the last payment of this type (when the Period was specified as OTHER).
DwellPayment	StdDwellPayWeeks	The number of weeks that the dwelling payment applies to, as standardised from the period code that was collected from the survey or the actual number of weeks if the response was given as OTHER.
Education		Educational details of the person as at the interview date.
Education	GainedAnyQuals	Whether the person gained any other qualifications that took more than the equivalent of 3 months of full-time study since the last interview.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
Education	GainedSchoolQuals	Whether any school qualifications were gained since the last interview. This is the response recorded when this was directly asked, as opposed to when asked whether any other qualifications gained were school qualifications.
Education	IsSchoolQual	Whether any of the OTHER qualifications gained since the last interview were school qualifications.
Education	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
Education	QualOth	Whether the person has any post-school qualifications.
Education	QualStudyFor	Whether the reference person has studied for a post-school qualification.
Education	SchoolAttend	Whether the person still attends secondary school, if the person is under 20 years old.
Education	SchoolLeaveMnth	The month in which the person left school, if the person is under 20 years old.
Education	SchoolLeaveYr	The year in which the person left school, if the person is under 20 years old.
Education	SchoolLeaveYrNr	Non-response to the year in which the person left school.
EstCaseVer		Cross sectional weights and eligiblity status of the household
EstCaseVer	Household	A randomised number assigned to uniquely identify the household
EstCaseVer	EligibilityHhld	Household eligibility status
EstCaseVer	xwtfinalwgt0	Cross sectional weight (final full sample weight). Use to produce estimates
EstCaseVer	xwtfinalwgt1-xwtfinalwgt100	Replicate cross sectional weights (use for calculating standard errors)
EstCaseVerPerson		Weights and eligiblity status for each person
EstCaseVerPerson	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
EstCaseVerPerson	EligibilityPerson	Person eligibility status
EstCaseVerPerson	Osm	Indicates whether the person is a an original sample member
EstCaseVerPerson	wtfinalwgt0	Final weight. Use to produce estimates
EstCaseVerPerson	wtfinalwgt1-wtfinalwgt100	Replicate weights (use for calculating standard errors)
Family		Family details of the person that apply either as at the interview date or over the reference period.
Family	AnyChildren	Indicates whether the person had any children as at the interview date.
Family	AnySiblings	Indicates whether the person had any siblings as at the interview date.
Family	LivedWithChild	Indicates whether the person had lived with any of their children during the reference period.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
Family	LivedWithParent	Indicates whether the person had lived with any of their parents during the reference period.
Family	LivedWithPartner	Indicates whether the person had lived with any partner during the reference period.
Family	LivedWithSibling	Indicates whether the person had lived with any sibling during the reference period.
Family	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
FamilyReln		Details of a relation that was living with the person at any time over the reference period.
FamilyReln	ChildWholeTime	Indicates whether any of the children of the relation (child) lived with the relation for all of the reference period.
FamilyReln	HadChild	Indicates whether the relation (child) of the person had children of their own while living with the person over the reference period.
FamilyReln	HadPartner	Indicates whether the relation (child) of the person had a partner while living with the person over the reference period.
FamilyReln	LivingWithStart	Date that the person started living with this relation (partner), if living with the partner for the whole reference period.
FamilyReln	LivingWithStartNr	Non-response to date that the relation (partner) started living with the person.
FamilyReln	PartnerWholeTime	Indicates whether the partner of the relation (child) lived with the child for all of the reference period.
FamilyReln	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
FamilyReln	PrevRelnOccrnce	Occurrence Number of this relationship in the most recent previous wave.
FamilyReln	RelnOccrnce	An occurrence number to distinguish relations of the same type for the same reference person.
FamilyReln	RelnPerson	The person number of the relation as shown in the EQ or 77 for those not present on the EQ.
FamilyReln	RelationID	The randomised id number (PersonCS) of the relation. If this is missing then the relation was not a respondent.
FamilyReln	RelnType	A code to denote the type of relation of the person.
FamilyReln	StillWithReln	Whether the relative is still living with the person.
FamilyReln	WholeTime	Indicates whether the person lived with the relation for the whole time since the start of the reference period.
HqEthnicity		An Ethnic Group to which the Household Member belongs. (A person may identify as belonging to more than one ethnic group.)
HqEthnicity	Ethnicity	Code to identify the ethnic group.
HqEthnicity	PersonCS	A randomised number assigned to uniquely identify a person within each wave.

Dataset	Var_Name	Label
HqReln		Relationship between 2 members of a household for a particular wave.
HqReln	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
HqReln	RelationID	The randomised id number (PersonCS) of the relation. If this is missing then the relation was not a respondent.
HqReln	Reln	Relationship of the Related Person to the Primary Person. (For example, if the relationship is CHILD, the Related Person is a child of the Primary Person.
IncomeSubTypeIncl		A wave income sub-type that was included in a wave income record for a type where multiple sub-types can be included in the same income amount. Each sub-type that was included has a record at this level, including the one that was originally selected.
IncomeSubTypeIncl	IncomeSubTypeIncl	The code for the wave income subtype that was included as part of the income being quoted.
IncomeSubTypeIncl	IncomeTypeIncl	The type of income that this included sub-type is a sub-type of. It should be the same as the wave income type of the wave income record that this included sub-type is for.
IncomeSubTypeIncl	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
IncomeSubTypeIncl	WaveIncomeSeq	A sequence number to distinguish multiple wave incomes for the same person. (This is used as a single key column in place of the different combinations of type, sub-type and occurrence number that would otherwise be used as key columns.)
LmHistory		The Labour Market History of the Person.
LmHistory	AgeStartWork	Age of the person when first started paid work.
LmHistory	AgeStartWorkNr	Non-response to age of the person when first started paid work.
LmHistory	NotWorking	Indicates whether the person has had a time of 12 months or more out of paid work.
LmHistory	PaidWorkEver	Indicates whether the person has ever been in paid work.
LmHistory	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LmHistory	YrsOutOfWork	The total number of years the person has been out of paid work since first starting.
LmHistory	YrsOutOfWorkNr	Non-response to the total number of years that the person has been out of paid work since first starting.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LmInfoItem		An item of information about a labour market spell that indicates a reason or method that was used to do something during that spell. These are items of information that can generally be assigned as codes unless the option selected was OTHER.
LmInfoItem	InfoItem	A code to indicate the information applicable for this information item type for this spell (eg one of the methods that was used to look for work). There can be multiple items of information for each possible type.
LmInfoItem	InfoItemOthNr	Non-response to the description of the item of information when OTHER was selected.
LmInfoItem	InfoItemType	A code used to indicate the type of information, eg reason for leaving the spell, method used to look for work.
LmInfoItem	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LmInfoItem	TimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
LmSpellInfo		Labour Market information about a labour market spell.
LmSpellInfo	EmployingOth	Whether the person was employing others during the spell, if this was a spell of self employment.
LmSpellInfo	HrsPerDayImp	Indicates whether and how the usual number of hours worked per day is imputed.
LmSpellInfo	HrsPerDay	Usual number of hours worked per day.
LmSpellInfo	HrsPerDayNr	Non-response to usual number of hours worked per day.
LmSpellInfo	HrsPerFortnight	The number of irregular hours worked in a fortnight at the end of the spell.
LmSpellInfo	HrsPerFortnightNr	Non-response to number of irregular hours worked in a fortnight at the end of the spell.
LmSpellInfo	HrsPerFortniteImp	Indicates whether and how the number of irregular hours worked in a fortnight at the end of the spell is imputed.
LmSpellInfo	HrsPerSpellImp	Indicates whether and how the total number of hours worked over the duration of the spell is imputed.
LmSpellInfo	HrsPerSpell	Total number of hours worked over the duration of the spell.
LmSpellInfo	HrsPerSpellNr	Non-response to total number of hours worked over the duration of the spell.
LmSpellInfo	HrsPerWeekImp	Indicates whether and how the usual number of hours worked per week is imputed.
LmSpellInfo	HrsPerWeek	Usual number of hours worked per week.
LmSpellInfo	HrsPerWeekNr	Non-response to usual number of hours worked per week.
LmSpellInfo	ImpGen	Indicates whether and how the row was generated as a result of imputation

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LmSpellInfo	Industry	The industry code as per the CARS classification and version that is currently the standard for this database.
LmSpellInfo	LookForPaidWork	Indicates whether the person looked for paid work during the spell, if this is a spell of unpaid work.
LmSpellInfo	Occupation	The occupation code as per the CARS classification and version that is currently the standard for this database.
LmSpellInfo	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LmSpellInfo	SameSituation	Indicates whether the person is in the same situation for this labour market spell as in the same spell for the preceding wave. (Derived from whether any of employer address and business activity and occupation, tasks and duties have changed.)
LmSpellInfo	StdHrsPerWeek	The number of hours worked per week, as standardised from the different ways in which hours worked during the spell could have been specified.
LmSpellInfo	TimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
LmTimeLineSpell		A labour market time-line spell.
LmTimeLineSpell	ImpGen	Indicates whether and how the row was generated as a result of imputation
LmTimeLineSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LmTimeLineSpell	SpellContd	Whether this is a continuation of a spell from the last wave.
LmTimeLineSpell	SpellCoverage	Indicates whether the spell lasted from the start of the reference period to the interview date.
LmTimeLineSpell	SpellEnd	End date of the spell (at 2400 hrs).
LmTimeLineSpell	SpellStart	Start date of the spell (at 0000hrs).
LmTimeLineSpell	SpellType	The type of labour market activity of the person during the spell.
LmTimeLineSpell	SpellTypeImp	Indicates whether and how the type of labour market activity of the person during the spell is imputed.
LmTimeLineSpell	TimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
LongEarnSpell		A longitudinal Earnings spell.
LongEarnSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LongEarnSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongEarnSpell	SpellEnd	End date of the spell (at 2400 hrs).
LongEarnSpell	SpellStart	Start date of the spell (at 0000hrs).

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LongEarnSpell	WklyEarn	The usual gross weekly income of a person from earnings during the Spell.
LongEarnSpell	WklyEarnImp	Indicates whether the usual gross weekly income from earnings during the Spell is derived from imputed data, ie at least one of the variables that was used in the derivation logic was imputed.
LongFamAgesSpell		A longitudinal Family By Ages spell.
LongFamAgesSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LongFamAgesSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongFamAgesSpell	SpellEnd	End date of the spell (at 2400 hrs).
LongFamAgesSpell	SpellStart	Start date of the spell (at 0000hrs).
LongFamilyReIn		A longitudinal relationship of the reference person.
LongFamilyReIn	LongReInOccrnce	An occurrence number to distinguish multiple longitudinal relationships of the same type for the same person.
LongFamilyReIn	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongFamilyReIn	ReInType	A code to denote the type of relation of the person.
LongGovtEarnSpell		A longitudinal Govt Transfers + Earnings spell.
LongGovtEarnSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LongGovtEarnSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongGovtEarnSpell	SpellEnd	End date of the spell (at 2400 hrs).
LongGovtEarnSpell	SpellStart	Start date of the spell (at 0000hrs).
LongGovtEarnSpell	WklyEarnGovt	The usual gross weekly income of the person from earnings and/or government transfers during the Spell.
LongGovtEarnSpell	WklyEarnGovtImp	Indicates whether the usual gross weekly income of the person from earnings and/or government transfers during the Spell is derived from imputed data, ie at least one of the variables that was used in the derivation logic was imputed.
LongGovtIncomeSpell		A longitudinal Government Income spell.
LongGovtIncomeSpell	GovtIncomeSubType	The type of government income broken down to a more specific sub-type.
LongGovtIncomeSpell	GovtIncomeType	The type of government income at a broad level.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LongGovtIncomeSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LongGovtIncomeSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongGovtIncomeSpell	SpellEnd	End date of the spell (at 2400 hrs).
LongGovtIncomeSpell	SpellStart	Start date of the spell (at 0000hrs).
LongGovtIncomeSpell	WksGovtSpell	The number of weeks of a spell of government income support.
LongIncomeSpell		A longitudinal Income spell.
LongIncomeSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LongIncomeSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongIncomeSpell	SpellEnd	End date of the spell (at 2400 hrs).
LongIncomeSpell	SpellStart	Start date of the spell (at 0000hrs).
LongIncomeSpell	WklyFamInc	Total average weekly family income of the person.
LongIncomeSpell	WklyFamIncImp	Indicates whether the total average weekly family income of the person is derived from imputed data, ie at least one of the variables that was used in the derivation logic was imputed.
LongIncomeSpell	WklyPersInc	The usual gross weekly income of the person from earnings and/or government transfers.
LongIncomeSpell	WklyPersIncImp	Indicates whether the usual gross weekly income of the person from earnings and/or government transfers is derived from imputed data, ie at least one of the variables that was used in the derivation logic was imputed.
LongLmActySpell		A longitudinal Labour Market Time Activity Spell.
LongLmActySpell	LongLmActySpell	A spell number to distinguish multiple longitudinal labour market activity spells for the same person.
LongLmActySpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongLmActySpell	SpellEnd	End date of the spell (at 2400 hrs).
LongLmActySpell	SpellStart	Start date of the spell (at 0000hrs).
LongLmTISpell		A longitudinal Labour Market Time Line Spell.
LongLmTISpell	EmplStatus	Status in employment over a Work Spell.
LongLmTISpell	FullTimeWork	Indicates whether the person worked full-time during a Work Spell.
LongLmTISpell	IndustryLevel1	The industry at major group (1 digit) level according to the classification version applicable to the maximum wave number of the output version.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LongLmTlSpell	LongTlSpell	A spell number to distinguish multiple longitudinal time line spells for the same person.
LongLmTlSpell	OccupationLevel1	The occupation at major group (1 digit) level according to the classification version applicable to the maximum wave number of the output version.
LongLmTlSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongLmTlSpell	SpellEnd	End date of the spell (at 2400 hrs).
LongLmTlSpell	SpellStart	Start date of the spell (at 0000hrs).
LongLmTlSpell	SpellType	The type of labour market activity of the person during the spell.
LongLmTlSpell	StdAllowPerWeek	The amount of allowances per week over the spell or sub-spell, as standardised from the different ways in which allowances during the spell or sub-spell could have been specified.
LongLmTlSpell	StdHrsPerWeek	The number of hours worked per week, as standardised from the different ways in which hours worked during the spell could have been specified.
LongLmTlSpell	StdPayPerWeek	The amount of pay per week over the spell or sub-spell, as standardised from the different ways in which pay during the spell or sub-spell could have been specified.
LongLmTlSpell	StdTotalExtraPay	The total of any extra payments for the spell.
LongLmTlSpell	WksPaid	The total number of weeks worked in paid employment.
LongReInSpell		A longitudinal relationship spell.
LongReInSpell	LongReInOccrnce	An occurrence number to distinguish multiple longitudinal relationships of the same type for the same person.
LongReInSpell	LongReInSpell	A spell number to distinguish multiple spells for the same longitudinal relationship.
LongReInSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongReInSpell	ReInSpellType	Indicates whether this spell is 1) where the relation lived with the person 2) where the relation (child) lived with the person and with a child of their own 3) where the relation (child) lived with the person and with a partner of their own
LongReInSpell	ReInType	A code to denote the type of relation of the person.
LongReInSpell	SpellEnd	End date of the spell (at 2400 hrs).
LongReInSpell	SpellStart	Start date of the spell (at 0000hrs).
LongSocMaritalSpell		A longitudinal Social Marital Status spell.
LongSocMaritalSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LongSocMaritalSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongSocMaritalSpell	SpellEnd	End date of the spell (at 2400 hrs).
LongSocMaritalSpell	SpellStart	Start date of the spell (at 0000hrs).
LongWaveIncomeSpell		A longitudinal spell of a type/sub-type of income where the amount is recorded at wave level.
LongWaveIncomeSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LongWaveIncomeSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LongWaveIncomeSpell	SpellEnd	End date of the spell (at 2400 hrs).
LongWaveIncomeSpell	SpellStart	Start date of the spell (at 0000hrs).
LongWaveIncomeSpell	WaveIncomeSubType	A further breakdown of income type that applies to some types of wave income (eg WINZ Benefit). If the Wave Income Type allows an amount of income to be quoted that can cover multiple sub-types, this sub-type is the one that was originally selected.
LongWaveIncomeSpell	WaveIncomeType	The basic type of wave income (eg ACC, Student Allowance, IRD Lump Sum, Family Support, WINZ Benefit, Youth Benefit, NZ Super, Veteran Pension Fund, Private Superannuation Fund, Self Employment).
LwxcEarnSpell		A reference to the Cross-sectional Earnings Spell that is the Contributor to a Longitudinal Earnings Spell for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)
LwxcEarnSpell	EarnSpell	A unique number to distinguish an Earnings Spell from all other Earnings Spells for the same person.
LwxcEarnSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LwxcEarnSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LwxcEarnSpell	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
LwxcEarnSpell	Wave	The number of the wave in the wave sequence for the survey.
LwxcFamAgesSpell		A reference to the Cross-sectional Family By Ages Spell that is the Contributor to a Longitudinal Family By Ages Spell for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LwxcFamAgesSpell	FamilySpell	A unique number to distinguish a Family Spell from all other Family Spells for the same person.
LwxcFamAgesSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LwxcFamAgesSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LwxcFamAgesSpell	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
LwxcFamAgesSpell	Wave	The number of the wave in the wave sequence for the survey.
LwxcFamilyReIn		A reference to the Cross-sectional Relationship that is the Contributor to a Longitudinal Relationship for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)
LwxcFamilyReIn	LongReInOccrrnce	An occurrence number to distinguish multiple longitudinal relationships of the same type for the same person.
LwxcFamilyReIn	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LwxcFamilyReIn	LwxcReInOccrrnce	An occurrence number to distinguish relations of the same type for the same reference person.
LwxcFamilyReIn	LwxcReInType	A code to denote the type of relation of the person.
LwxcFamilyReIn	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
LwxcFamilyReIn	ReInType	A code to denote the type of relation of the person.
LwxcFamilyReIn	Wave	The number of the wave in the wave sequence for the survey.
LwxcGovtEarnSpell		A reference to the Cross-sectional Govt Transfers + Earnings Spell that is the Contributor to a Longitudinal Govt Transfers + Earnings Spell for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)
LwxcGovtEarnSpell	EarnGovtSpell	A unique number to distinguish an Earnings + Govt Transfers Spell from all other spells of this type for the same person.
LwxcGovtEarnSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LwxcGovtEarnSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LwxcGovtEarnSpell	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LwxcGovtEarnSpell	Wave	The number of the wave in the wave sequence for the survey.
LwxcGovtIncomeSpell		A reference to the Cross-sectional Government Income Spell that is the Contributor to a Longitudinal Government Income Spell for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)
LwxcGovtIncomeSpell	GovtIncomeSpell	Spell number identifying this longitudinal government income spell out of all spells for the same type/sub-type of government income for this person.
LwxcGovtIncomeSpell	GovtIncomeSubType	The type of government income broken down to a more specific sub-type.
LwxcGovtIncomeSpell	GovtIncomeType	The type of government income at a broad level.
LwxcGovtIncomeSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LwxcGovtIncomeSpell	LwxcGvIncSubType	The type of government income broken down to a more specific sub-type.
LwxcGovtIncomeSpell	LwxcGvIncType	The type of government income at a broad level.
LwxcGovtIncomeSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LwxcGovtIncomeSpell	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
LwxcGovtIncomeSpell	Wave	The number of the wave in the wave sequence for the survey.
LwxcIncomeSpell		A reference to the Cross-sectional Income Spell that is the Contributor to a Longitudinal Income Spell for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)
LwxcIncomeSpell	IncomeSpell	A unique number to distinguish an Income Spell from all other spells of this type for the same person.
LwxcIncomeSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LwxcIncomeSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LwxcIncomeSpell	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
LwxcIncomeSpell	Wave	The number of the wave in the wave sequence for the survey.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LwxcLmActySpell		A reference to the Cross-sectional Labour market Activity Spell that is the Contributor to a Longitudinal Labour market Activity Spell for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)
LwxcLmActySpell	LongLmActySpell	A spell number to distinguish multiple longitudinal labour market activity spells for the same person.
LwxcLmActySpell	LwxcLmActySpell	A unique number to distinguish a Labour Market Activity Spell from other such spells for the same person.
LwxcLmActySpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LwxcLmActySpell	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
LwxcLmActySpell	Wave	The number of the wave in the wave sequence for the survey.
LwxcLmTlSpell		A reference to the Cross-sectional Labour Market Time Line Spell that is the Contributor to a Longitudinal Labour Market Time Line Spell for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)
LwxcLmTlSpell	LongTlSpell	A spell number to distinguish multiple longitudinal time line spells for the same person.
LwxcLmTlSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LwxcLmTlSpell	LwxcTimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
LwxcLmTlSpell	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
LwxcLmTlSpell	Wave	The number of the wave in the wave sequence for the survey.
LwxcReInSpell		A reference to the Cross-sectional Relationship Spell that is the Contributor to a Longitudinal Relationship Spell for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)
LwxcReInSpell	LongReInOccrnce	An occurrence number to distinguish multiple longitudinal relationships of the same type for the same person.
LwxcReInSpell	LongReInSpell	A spell number to distinguish multiple spells for the same longitudinal relationship.
LwxcReInSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LwxcReInSpell	LwxcReInOccrnce	An occurrence number to distinguish relations of the same type for the same reference person.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LwxcReInSpell	LwxcReInSpell	A spell number to distinguish multiple spells for the same person and relation of the same relation spell type.
LwxcReInSpell	LwxcReInSpellType	Indicates whether this spell is 1) where the relation lived with the person 2) where the relation (child) lived with the person and with a child of their own 3) where the relation (child) lived with the person and with a partner of their own
LwxcReInSpell	LwxcReInType	A code to denote the type of relation of the person.
LwxcReInSpell	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
LwxcReInSpell	ReInSpellType	Indicates whether this spell is 1) where the relation lived with the person 2) where the relation (child) lived with the person and with a child of their own 3) where the relation (child) lived with the person and with a partner of their own
LwxcReInSpell	ReInType	A code to denote the type of relation of the person.
LwxcReInSpell	Wave	The number of the wave in the wave sequence for the survey.
LwxcSocMaritalSpell		A reference to the Cross-sectional Social Marital Status Spell that is the Contributor to a Longitudinal Social Marital Status Spell for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)
LwxcSocMaritalSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LwxcSocMaritalSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
LwxcSocMaritalSpell	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
LwxcSocMaritalSpell	SocMaritalSpell	A unique number to distinguish this Social Marital Status Spell from other Social Marital Status Spells for the same person.
LwxcSocMaritalSpell	Wave	The number of the wave in the wave sequence for the survey.
LwxcWaveIncomeSpell		A reference to the Cross-sectional Wave Income Spell that is the Contributor to a Longitudinal Wave Income Spell for a specific Wave. (LWXC = Longitudinal Wave Cross-Sectional Contributor.)
LwxcWaveIncomeSpell	LongSpell	A spell number to distinguish multiple longitudinal spells of the same type for the same person.
LwxcWaveIncomeSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
LwxcWaveIncomeSpell	PersonLT	A randomised number assigned to uniquely identify a person over each wave. For each wave the person has the same identifier
LwxcWaveIncomeSpell	WaveIncomeSeq	A sequence number to distinguish multiple wave incomes for the same person. (This is used as a single key column in place of the different combinations of type, sub-type and occurrence number that would otherwise be used as key columns.)
LwxcWaveIncomeSpell	WaveIncomeSpell	Spell number identifying this wave income spell out of all spells for the same wave income for this person.
LwxcWaveIncomeSpell	WaveIncomeSubType	A further breakdown of income type that applies to some types of wave income (eg WINZ Benefit). If the Wave Income Type allows an amount of income to be quoted that can cover multiple sub-types, this sub-type is the one that was originally selected.
LwxcWaveIncomeSpell	WaveIncomeType	The basic type of wave income (eg ACC, Student Allowance, IRD Lump Sum, Family Support, WINZ Benefit, Youth Benefit, NZ Super, Veteran Pension Fund, Private Superannuation Fund, Self Employment).
LwxcWaveIncomeSpell	Wave	The number of the wave in the wave sequence for the survey.
PayChg		Details of a pay change during a (paid work) spell.
PayChg	PayChgOccrnce	An occurrence number to distinguish multiple pay changes during the same spell.
PayChg	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
PayChg	TimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
PayChg	WhenChg	The date of the change (at 0000hrs).
PayChg	WhenChgNr	Non-response to the date of the change.
PayChgRsn		A reason for a pay change during a (paid work) spell.
PayChgRsn	PayChgOccrnce	An occurrence number to distinguish multiple pay changes during the same spell.
PayChgRsn	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
PayChgRsn	TimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
PayChgRsn	WhyChg	Code indicating which of a set of standard reasons applies.
PayChgRsn	WhyChgOthNr	Non-response to other reason for change.
PayRate		Details of income for one of the rates at which the person was paid at this part of the spell.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
PayRate	HrlyPayRate	The hourly pay rate.
PayRate	HrlyPayRateImp	Indicates whether and how the hourly pay rate is imputed.
PayRate	HrlyPayRateNr	Non-response to hourly pay rate.
PayRate	HrsAtRateImp	Indicates whether and how the number of hours worked at this hourly rate is imputed.
PayRate	HrsAtRate	Number of hours worked at this hourly rate.
PayRate	HrsAtRateNr	Non-response to number of hours worked at this rate.
PayRate	ImpGen	Indicates whether and how this row in the table was generated by the imputation process. This is usually as a result of imputing for a person who had a proxy response, and where therefore no information of this type was collected during the survey.
PayRate	PayRate	A number to distinguish multiple pay rates for the same sub-spell.
PayRate	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
PayRate	SubSpellType	Indicates whether the sub-spell is the start or the end of the spell.
PayRate	TimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
PqPerson		Personal Questionnaire administration details of a person.
PqPerson	EIPD	End date of the interview period (at time 2400hrs).
PqPerson	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
PqPerson	PqiRespStatus	Personal interview response status code. Derived from either PQI Response Status description or CAI Response Status description.
PqPerson	SIPD	Start date of the interview period (at time 0000hrs).
QualPostSchool		Details of a post-school qualification held by the person.
QualPostSchool	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
QualPostSchool	Qual	The qualification code as per the CARS classification and version that is currently the standard for this database.
QualPostSchool	QualGainedYr	The year in which the post-school qualification was gained.
QualPostSchool	QualGainedYrNr	Non-response to the year in which the qualification was gained.
QualPostSchool	QualOccrnce	Occurrence number to distinguish this qualification out of all post-school qualifications held by the person.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
QualPostSchool	QualOrig	The qualification code as originally coded using the Classification and Version for the wave that the code applies to.
ReInSpell		A spell during which the person had the relation living with them and during which the relation (if a child) may also have had other relations of theirs (children or a partner) also living with the person.
ReInSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
ReInSpell	ReInOccrnce	An occurrence number to distinguish relations of the same type for the same reference person.
ReInSpell	ReInSpell	A spell number to distinguish multiple spells for the same person and relation of the same relation spell type.
ReInSpell	ReInSpellType	Indicates whether this spell is 1) where the relation lived with the person 2) where the relation (child) lived with the person and with a child of their own 3) where the relation (child) lived with the person and with a partner of their own
ReInSpell	ReInType	A code to denote the type of relation of the person.
ReInSpell	SpellEnd	End date of the spell (at 2400 hrs).
ReInSpell	SpellStart	Start date of the spell (at 0000hrs).
SchoolQual		Details of a secondary school qualification of the person.
SchoolQual	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
SchoolQual	SchoolQual	Code for the secondary school qualification held by the person.
SpellExtraPay		Details of an extra payment from a (paid work) spell.
SpellExtraPay	ExtraPay	Gross amount of the extra payment.
SpellExtraPay	ExtraPayImp	Indicates whether and how the gross amount of the extra payment is imputed.
SpellExtraPay	ExtraPayNr	Non-response to the gross amount of the extra payment.
SpellExtraPay	ExtraPayOccrnce	An occurrence number used to distinguish multiple extra payments for the same spell.
SpellExtraPay	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
SpellExtraPay	TimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
SpellIncome		Details of income for a working spell.
SpellIncome	AvgMnthlyEarn	Average gross monthly earnings over the paid work spell.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
SpellIncome	AvgMnthlyEarnImp	Indicates whether the average gross monthly earnings over the paid work spell is imputed.
SpellIncome	AvgMnthlyEarnNr	Non-response to average gross monthly earnings over the paid work spell.
SpellIncome	GrossEarn	Gross earnings from either a casual work spell or from a paid work spell lasting more than 1 week.
SpellIncome	GrossEarnApprox	Gross earnings from the (casual work) spell approximated to the nearest \$1000.
SpellIncome	GrossEarnApproxNr	Non-response to gross earnings from the (casual work) spell approximated to the nearest \$1000.
SpellIncome	GrossEarnImp	Indicates whether and how the gross earnings from either a casual work spell or from a paid work spell lasting more than 1 week is imputed.
SpellIncome	GrossEarnNr	Non-response to gross earnings over the spell.
SpellIncome	LastGrossPay	Last gross pay amount in the (paid work) spell.
SpellIncome	LastGrossPayImp	Indicates whether and how the last gross pay amount in the (paid work) spell is imputed.
SpellIncome	LastGrossPayNr	Non-response to last gross pay amount in the (paid work) spell.
SpellIncome	LastMnthlyEarn	Last gross monthly earnings in the (paid work) spell.
SpellIncome	LastMnthlyEarnImp	Indicates whether and how the last gross monthly earnings in the (paid work) spell is imputed.
SpellIncome	LastMnthlyEarnNr	Non-response to last gross monthly earnings in the (paid work) spell.
SpellIncome	LastPayDaysImp	Indicates whether and how the number of days covered by the last gross pay in the (paid work) spell has been imputed.
SpellIncome	LastPayDays	The number of days covered by the last gross pay in the (paid work) spell.
SpellIncome	LastPayHrsImp	Indicates whether and how the number of hours covered by the last gross pay in the (paid work) spell is imputed.
SpellIncome	LastPayHrs	The number of hours covered by the last gross pay in the (paid work) spell.
SpellIncome	LastPayPeriod	The period covered by the last gross pay in the (paid work) spell.
SpellIncome	LastPayPeriodImp	Indicates whether and how the period covered by the last gross pay in the (paid work) spell is imputed.
SpellIncome	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
SpellIncome	RcvdExtraPay	Indicates whether extra payments were received over the (paid work) spell.
SpellIncome	RcvdRedundancy	Indicates whether redundancy pay was received during the (paid work) spell.
SpellIncome	Redundancy	The gross amount of redundancy payments in the (paid work) spell.
SpellIncome	RedundancyNr	Non-response to the gross amount of redundancy payments in the (paid work) spell.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
SpellIncome	SamePay	Indicates whether the pay was the same at the start and end of the (paid work) spell.
SpellIncome	ShortSpellEarn	Gross earnings for a short (paid work) spell lasting a week or less.
SpellIncome	ShortSpellEarnImp	Indicates whether and how the gross earnings for a short (paid work) spell lasting a week or less is imputed.
SpellIncome	ShortSpellEarnNr	Non-response to gross earnings for a short (paid work) spell lasting a week or less.
SpellIncome	StdAllowPerWeek	The amount of allowances per week over the spell or sub-spell, as standardised from the different ways in which allowances during the spell or sub-spell could have been specified.
SpellIncome	StdPayPerWeek	The amount of pay per week over the spell or sub-spell, as standardised from the different ways in which pay during the spell or sub-spell could have been specified.
SpellIncome	StdTotalExtraPay	The total of any extra payments for the spell.
SpellIncome	TimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
StudyYrMnth		Details of a year/month in which the person spent 5 or more days studying.
StudyYrMnth	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
StudyYrMnth	StudyFullTime	Whether the reference person was engaged in full-time study for this month.
StudyYrMnth	StudyYrMnth	The year/month in which the study occurred.
SubSpellIncome		Details of income for a a (paid work) sub-spell, ie either the start or the end of the spell.
SubSpellIncome	Allowances	Gross amount of allowances.
SubSpellIncome	AllowancesHrlyPay	Indicates whether the person received allowances additional to their hourly pay rate.
SubSpellIncome	AllowancesImp	Indicates whether and how the gross amount of allowances is imputed.
SubSpellIncome	AllowancesInclPay	Indicates whether allowances were included in pay amount given.
SubSpellIncome	AllowancesNr	Non-response to gross amount of allowances.
SubSpellIncome	AllowancesPartPay	Indicates whether allowances were received as part of pay.
SubSpellIncome	AllowancesPeriodIm	Indicates whether and how the period covered by allowances is imputed.
SubSpellIncome	AllowancesPeriod	The period covered by allowances.
SubSpellIncome	AllowancesWeeksImp	Indicates whether and how the number of weeks covered by allowances is imputed.
SubSpellIncome	AllowancesWeeks	The number of weeks covered by allowances when OTHER specified for the period.
SubSpellIncome	AllowancesWeeksNr	Non-response to number of weeks covered by the allowances.
SubSpellIncome	AnnSalary	Gross annual salary.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
SubSpellIncome	AnnSalaryImp	Indicates whether and how the gross annual salary is imputed.
SubSpellIncome	AnnSalaryNr	Non-response to gross annual salary.
SubSpellIncome	PayDetails	Whether the respondent has given income as a usual pay, annual salary or as 1 or more hourly rate(s).
SubSpellIncome	PayDetailsImp	Indicates whether and how the type of pay details has been imputed.
SubSpellIncome	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
SubSpellIncome	StdAllowPerWeek	The amount of allowances per week over the spell or sub-spell, as standardised from the different ways in which allowances during the spell or sub-spell could have been specified.
SubSpellIncome	StdPayPerWeek	The amount of pay per week over the spell or sub-spell, as standardised from the different ways in which pay during the spell or sub-spell could have been specified.
SubSpellIncome	SubSpellType	Indicates whether the sub-spell is the start or the end of the spell.
SubSpellIncome	TimeLineSpell	A spell number used to distinguish the spell out of all labour market time line spells for this person.
SubSpellIncome	UsualPay	Gross usual pay amount.
SubSpellIncome	UsualPayImp	Indicates whether and how the gross usual pay amount is imputed.
SubSpellIncome	UsualPayNr	Non-response to gross usual pay amount.
SubSpellIncome	UsualPayPeriod	Period covered by usual pay.
SubSpellIncome	UsualPayPeriodImp	Indicates whether and how the period covered by usual pay is imputed.
SubSpellIncome	UsualPayWeeksImp	Indicates whether and how the number of weeks covered by usual pay is imputed.
SubSpellIncome	UsualPayWeeks	Number of weeks covered by usual pay if OTHER was specified as the period.
SubSpellIncome	UsualPayWeeksNr	Non-response to number of weeks covered by usual pay.
Wave		One of the Waves of the Longitudinal Survey. (This is a reference table that is referred to by UV tables, but which has content that is independent of Output Version.)
Wave	PopnEstAsAt	As-at date that population estimates apply to for this wave.
Wave	Wave	The number of the wave in the wave sequence for the survey.
WaveIncome		Details of an income of the person that are collected at wave level (ie covers the whole reference or interview period, or is for the last payment of its kind for the reference period). Each income is for a specific type, and (for some types) sub-type.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
WaveIncome	Contd	Whether this type of wave income is still being received.
WaveIncome	Curr	The currency code as per the CARS classification and version that is currently that standard for this database.
WaveIncome	CurrConvRate	The currency conversion rate to use to convert the income in foreign currency to NZ dollars. This is given as the amount in the foreign currency that equals a NZ dollar.
WaveIncome	CurrOrig	The currency code as originally coded using the classification and version for the wave that the code belongs to.
WaveIncome	Income	Amount of income. If there is tax on this income, this is the amount before tax.
WaveIncome	IncomeApprox	The amount of the income expressed to the nearest \$x where x is an amount that depends on the type of income. (If the income is taxable, this is the amount of income before tax.)
WaveIncome	IncomeApproxNr	Non-response to approximate amount of income.
WaveIncome	IncomeBand	A code within the income band set that indicates which band was selected. The income (before tax if taxable) is between the minimum and maximum specified for that band.
WaveIncome	IncomeBandSet	The set of income bands that have been used to specify a band that the income (before tax if taxable) belongs to. (Only used when the income has been specified as within a band.)
WaveIncome	IncomeImp	Indicates whether and how the Amount of Income is imputed.
WaveIncome	IncomeNr	Non-response to amount of income.
WaveIncome	IncomePeriod	A code to indicate the period of time covered by the income.
WaveIncome	IncomePeriodImp	Indicates whether and how the code to indicate the period of time covered by the income is imputed.
WaveIncome	IncomeWeeksImp	Indicates whether and how the number of weeks covered by the income is imputed.
WaveIncome	IncomeWeeks	The number of weeks covered by the income (if OTHER selected from the list of standard periods).
WaveIncome	IncomeWeeksNr	Non-response to number of weeks covered by the income.
WaveIncome	MultiSubTypes	Indicates whether this (last) income payment covers multiple sub-types. (This can apply both for Family Support and WINZ benefits.)
WaveIncome	NetIncome	The net amount of the income after tax deducted. (Only used if the income is taxable.)
WaveIncome	NetIncomeApprox	Amount of net income after tax is deducted, expressed to the nearest \$x where x is a number that depends on the type of income. (Only used if this type of income is taxable.)
WaveIncome	NetIncomeApproxNr	Non-response to approximate net amount of income.

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
WaveIncome	NetIncomeNr	Non-response to net income amount.
WaveIncome	NzDollars	Whether the amount of income was in New Zealand dollars.
WaveIncome	OtherTime	Whether the person received this type of wave income at any other time during the reference period.
WaveIncome	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
WaveIncome	ProfitLoss	Indicates whether the amount of income quoted is a profit or loss. (Only used if the income is from self-employment.)
WaveIncome	ProfitLossImp	Indicates whether and how the indicator for whether the amount of income quoted is a profit or loss is imputed.
WaveIncome	Rcvd	Whether this type of wave income was received at any time during the reference period. Only if it was recorded for the last wave in error.
WaveIncome	StdIncome	The amount of income for the type (and possibly sub-type) of income, as standardised from the different ways in which the income could have been specified.
WaveIncome	StdIncomeWeeks	The number of weeks covered by the income quoted, as standardised from the period code that was collected from the survey or the actual number of weeks if the response was given as OTHER.
WaveIncome	Stop	When did the person stop receiving this type of wave income? Only if they received it last wave but then stopped receiving it.
WaveIncome	StopNr	Non-response to when the person stopped receiving this type of wave income.
WaveIncome	WaveIncomeOccrrnce	An occurrence number to distinguish multiple incomes that have the same type and/or sub-type. This is used to track back to the as-received information that this is based on. Only used if the income type (or sub-type) allows multiple occurrences.
WaveIncome	WaveIncomeSeq	A sequence number to distinguish multiple wave incomes for the same person. (This is used as a single key column in place of the different combinations of type, sub-type and occurrence number that would otherwise be used as key columns.)
WaveIncome	WaveIncomeSubType	A further breakdown of income type that applies to some types of wave income (eg WINZ Benefit). If the Wave Income Type allows an amount of income to be quoted that can cover multiple sub-types, this sub-type is the one that was originally selected.
WaveIncome	WaveIncomeType	The basic type of wave income (eg ACC, Student Allowance, IRD Lump Sum, Family Support, WINZ Benefit, Youth Benefit, NZ Super, Veteran Pension Fund, Private Superannuation Fund, Self Employment).

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
WaveIncome	WholeTime	Whether this type of wave income was being received for the whole of the reference period.
WaveIncomeHdr		A header under which information is collected about an income of the person at wave level, ie where the information applies to the entire wave (all of reference or interview period) or is for the last payment from the wave).
WaveIncomeHdr	ContdNzSuper	Whether the respondent is still receiving NZ Superannuation.
WaveIncomeHdr	ContdPension	Whether the respondent is still receiving a veteran pension.
WaveIncomeHdr	NzSuperWholeTime	Whether the respondent got NZ Superannuation for the whole of the reference period, if they got it at the start of the period and still do.
WaveIncomeHdr	PensionWholeTime	Whether the respondent got a veteran pension for the whole of the reference period, if they got it at the start of the period and still do.
WaveIncomeHdr	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
WaveIncomeHdr	PrivateSuper	The number of private superannuation schemes that the person received payments from over the reference period.
WaveIncomeHdr	PrivateSuperNr	Non-response to the number of private superannuation schemes that the person received payments from.
WaveIncomeHdr	RcvdAcc	Indicates whether the person received any ACC payments during the reference period.
WaveIncomeHdr	RcvdIrdLumpSum	Indicates whether the person received a lump sum IRD payment during the reference period.
WaveIncomeHdr	RcvdNzSuper	Indicates whether the person received payments from the NZ Superannuation fund over the reference period.
WaveIncomeHdr	RcvdPension	Indicates whether the person received payments from a veterans pension fund over the reference period.
WaveIncomeHdr	RcvdPrivateSuper	Indicagtes whether the person received payments from one or more private superannuation funds over the reference period.
WaveIncomeHdr	RcvdStudentAllow	Indicates whether the person received any Student Allowance payments during the reference period.
WaveIncomeHdr	RcvdYouthBenefit	Indicates whether the person received the youth benefit during the reference period.
WaveIncomeHdr	SelfEmpEarlierRec	Indicates whether self-employment income records are available for an earlier tax year. (Only used if the person was self-employed.)
WaveIncomeHdr	SelfEmpTaxYr	The year for which self-employment tax records are available. (Only used if the person was self-employed.)

<b>Dataset</b>	<b>Var_Name</b>	<b>Label</b>
WaveIncomeHdr	SelfPwvEarlierRec	Indicates whether self-employment income records for a missed previous wave are available for an earlier tax year. (Only used if the person was self-employed in the missed wave.)
WaveIncomeHdr	StopNzSuper	When the respondent stopped getting NZ Superannuation, if they got it last wave but no longer get it.
WaveIncomeHdr	StopNzSuperNr	Non-response to when stopped receiving NZ Superannuation.
WaveIncomeSpell		A spell during which the person received an income for which the amounts of income have been collected at wave level
WaveIncomeSpell	PersonCS	A randomised number assigned to uniquely identify a person within each wave.
WaveIncomeSpell	SpellEnd	End date of the spell (at 2400 hrs).
WaveIncomeSpell	SpellStart	Start date of the spell (at 0000hrs).
WaveIncomeSpell	WaveIncomeSeq	A sequence number to distinguish multiple wave incomes for the same person. (This is used as a single key column in place of the different combinations of type, sub-type and occurrence number that would otherwise be used as key columns.)
WaveIncomeSpell	WaveIncomeSpell	Spell number identifying this wave income spell out of all spells for the same wave income for this person.

## 7.4 SoFIE-Health Data Laboratory directory listing

Key files in the SoFIE-Health user/administrator directory structure are listed here. This provides a more detailed view of view provided in Figure 5.

Volume in drive P is C\_Datalab  
Volume Serial Number is 14FB-6F10

Directory of P:\MAA2006-07 SoFIE A\SASProgs

17/07/2007	09:51 AM	<DIR>	June
18/10/2007	01:48 PM	<DIR>	Ken
14/01/2008	02:06 PM	<DIR>	Kristie
06/11/2007	09:11 AM	<DIR>	SoFIE Admin
04/10/2007	12:02 PM	<DIR>	sofie macros

Directory of P:\MAA2006-07 SoFIE A\SASProgs\June

0 File(s) 0 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\Ken

31/07/2007	12:23 PM	<DIR>	autocall_macros
31/07/2007	10:45 AM	<DIR>	compiled_macros_setup
07/11/2007	04:02 PM	<DIR>	SASprog_backup
		0 File(s)	0 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\Ken\autocall\_macros

18/10/2007	12:30 PM		1,484 autocall_macros.sas
18/10/2007	01:48 PM	<DIR>	autocall_macros_code
18/10/2007	12:23 PM		1,971 autocall_macros_setup.sas
		2 File(s)	3,455 bytes

Directory of P:\MAA2006-07 SoFIE  
A\SASProgs\Ken\autocall\_macros\autocall\_macros\_code

04/09/2007	11:40 AM		2,103 catvar_levels.sas
31/08/2007	02:33 PM		5,998 char2num.sas
12/07/2007	03:48 PM		3,823 check_levs.sas
31/07/2007	11:24 AM		806 compile_macros.sas
12/07/2007	02:11 PM		1,879 count_ID.sas
12/07/2007	02:10 PM		2,984 count_ID_vars.sas
04/10/2007	09:31 AM		2,574 dropmissvar.sas
31/07/2007	04:13 PM		3,525 dropvars.sas
17/07/2007	11:33 AM		1,322 filepres.sas
18/10/2007	10:50 AM		16,205 fmt_sof_vars.sas
04/09/2007	05:07 PM		3,820 levsout.sas
17/07/2007	11:08 AM		3,490 recontab.sas
17/07/2007	11:34 AM		1,965 rndround.sas
05/09/2007	12:36 PM		11,884 sf36calc.sas
06/09/2007	04:34 PM		3,944 sf36chk.sas
19/09/2007	01:05 PM		15,167 soffmts.sas
18/10/2007	01:48 PM		4,846 sofsess.sas
27/09/2007	01:39 PM		8,850 towide.sas

31/07/2007	11:26 AM		533 use_compiled_macros.sas
27/09/2007	03:10 PM		2,143 var_ntl.sas
		20 File(s)	97,861 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\Ken\compiled\_macros\_setup

28/08/2007	09:47 AM		1,949 compile_macros_setup.sas
28/08/2007	09:47 AM		1,708 use_compiled_macros_setup.sas
		2 File(s)	3,657 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\Kristie

16/01/2008	03:25 PM		29,072 attrit_tabs.sas
14/08/2007	05:49 PM		5,317 begin_formats.sas
18/10/2007	01:04 PM		5,256 begin_formats_numeric.sas
04/12/2007	04:58 PM		6,024 early_sep.sas
14/01/2008	04:17 PM		11,793 food_security.sas
03/08/2007	10:55 AM		7,753 health_check.sas
03/08/2007	05:11 PM		6,325 health_check_weighted.sas
17/09/2007	12:49 PM		32,499 health_sort.sas
28/11/2007	04:11 PM		4,718 health_tabs.sas
15/08/2007	05:42 PM		6,842 income_checks.sas
20/11/2007	04:45 PM		12,086 logistic regression macros 2.sas
17/12/2007	03:08 PM		15,157 logistic regression macros 3.sas
20/11/2007	12:00 PM		12,436 logistic regression macros.sas
17/09/2007	12:49 PM		1,947 new_formats_KC.sas
20/11/2007	08:54 AM	<DIR>	Old
18/10/2007	04:30 PM		11,775 random_var_checking.sas
15/08/2007	05:29 PM		14,078 respondent_tabs.sas
28/06/2007	05:15 PM		15,744 sf36score_macro.sas
03/08/2007	03:22 PM		4,562 SF36_SoFIE_W3.sas
15/08/2007	04:49 PM		4,784 SRH_compare.sas
02/08/2007	05:06 PM		289 wealth_code.txt
16/01/2008	03:14 PM		24,411 wealth_health.sas
		21 File(s)	232,868 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin

02/10/2007	11:43 AM	<DIR>	code_examples
07/11/2007	04:00 PM	<DIR>	create_research_dataset
06/11/2007	09:12 AM	<DIR>	derive_vars
17/07/2007	03:18 PM	<DIR>	exploration
25/09/2007	02:11 PM	<DIR>	merge_datasets
06/11/2007	01:06 PM	<DIR>	merge_drv_vars
		0 File(s)	0 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\code\_examples

31/07/2007	11:22 AM		2,537 begin_sofie.sas
02/10/2007	11:43 AM	<DIR>	old
		1 File(s)	2,537 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\create\_research\_dataset

03/12/2007	11:18 AM		17,858 create_research_dataset.sas
------------	----------	--	------------------------------------

01/11/2007 05:07 PM <DIR> old  
1 File(s) 17,858 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\derive\_vars

24/10/2007 02:19 PM <DIR> change\_vars  
02/10/2007 09:08 AM <DIR> CPI\_eqinc  
18/10/2007 01:58 PM <DIR> health  
07/11/2007 03:09 PM <DIR> incwlth\_pctiles  
06/11/2007 09:13 AM <DIR> quals  
19/09/2007 02:54 PM <DIR> wealth  
0 File(s) 0 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\derive\_vars\change\_vars

07/11/2007 03:38 PM 3,755 chnge\_inc.sas  
07/11/2007 03:42 PM 5,020 chnge\_marital.sas  
07/11/2007 03:46 PM 3,248 chnge\_SRH.sas  
24/10/2007 02:19 PM <DIR> old  
3 File(s) 12,023 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\derive\_vars\CPI\_eqinc

09/10/2007 02:13 PM 6,879 CPI\_eqinc.sas  
19/09/2007 03:48 PM <DIR> old  
08/08/2007 05:47 PM 40,448 SNZ cpimar07 KC adapted KR.xls  
09/08/2007 10:53 AM 16,896 SNZ cpimar07 KC adapted  
KRv1.1.xls  
3 File(s) 64,223 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\derive\_vars\health

18/10/2007 02:02 PM 3,847 drv\_chron.sas  
16/10/2007 10:01 AM 6,092 drv\_nzidep.sas  
25/10/2007 04:27 PM 4,350 drv\_smoke\_alc.sas  
16/10/2007 09:59 AM <DIR> old  
25/10/2007 04:18 PM 8,533 score\_k10\_PSS.sas  
09/10/2007 12:35 PM 7,942 score\_SF36.sas  
28/08/2007 11:54 AM 11,800 sf36calc.sas  
16/10/2007 10:39 AM 14,217 test\_drv\_nzidep.sas  
7 File(s) 56,781 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\derive\_vars\incwlth\_pctiles

18/10/2007 03:41 PM 16,246 income\_checks\_KC.sas  
03/12/2007 12:34 PM 17,911 incwlth\_pctiles.sas  
07/11/2007 03:09 PM <DIR> old  
03/12/2007 12:20 PM 5,448 rank\_incwlth.sas  
3 File(s) 39,605 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\derive\_vars\quals

29/11/2007 04:46 PM <DIR> MaxQualRank\_WSM  
0 File(s) 0 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE  
Admin\derive\_vars\quals\MaxQualRank\_WSM

27/11/2007 01:56 PM 1,355 drv\_maxqualrank\_WSM.sas  
29/11/2007 04:00 PM 2,400 maxqualrank\_WSM.sas  
27/11/2007 12:05 PM <DIR> old  
2 File(s) 3,755 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE  
Admin\derive\_vars\wealth

19/09/2007 02:54 PM <DIR> old  
19/09/2007 03:25 PM 4,490 test\_wealth.sas  
03/12/2007 11:20 AM 2,463 wealth.sas  
2 File(s) 6,953 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\exploration

19/06/2007 02:44 PM <DIR> checking\_key\_vars  
28/06/2007 09:35 AM <DIR> sofvarinvestigate  
01/08/2007 02:48 PM <DIR> var\_levels  
22/05/2007 05:42 PM <DIR> wealth\_health  
0 File(s) 0 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE  
Admin\exploration\checking\_key\_vars

17/07/2007 02:49 PM 2,722 checking\_key\_vars.sas  
19/06/2007 02:44 PM <DIR> old  
1 File(s) 2,722 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE  
Admin\exploration\sofvarinvestigate

28/06/2007 09:34 AM <DIR> old  
17/07/2007 03:10 PM 19,939 sofvarinvestigateKRv1.sas  
1 File(s) 19,939 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE  
Admin\exploration\var\_levels

19/06/2007 02:45 PM <DIR> old  
12/07/2007 03:51 PM 12,312 var\_levels.sas  
01/08/2007 03:23 PM 2,501 var\_levels\_4SJ.txt  
17/07/2007 03:11 PM 3,386 var\_levels\_health.sas  
17/07/2007 06:10 PM 2,367 var\_levels\_report.sas  
4 File(s) 20,566 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE  
Admin\exploration\wealth\_health

22/05/2007 05:42 PM <DIR> old  
17/07/2007 03:14 PM 10,142 wealth\_health.sas  
1 File(s) 10,142 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\merge\_datasets

31/08/2007	04:52 PM		10,813	check_merge_datasets.sas
12/07/2007	09:29 AM	<DIR>		creating_id
11/09/2007	04:12 PM		7,361	dropvars.sas
25/09/2007	11:49 AM		42,070	merge_datasets.sas
05/09/2007	10:58 AM		56,039	merge_setup.sas
06/11/2007	05:37 PM		7,716	merge_themed_datasets.sas
25/09/2007	02:11 PM	<DIR>		old
		5 File(s)		123,999 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\merge\_datasets\creating\_id

12/07/2007	09:31 AM		1,450	creating_id.hlp
12/07/2007	09:27 AM	<DIR>		old
		1 File(s)		1,450 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\SoFIE Admin\merge\_drv\_vars

07/11/2007	02:52 PM		3,868	merge_drv_vars.sas
06/11/2007	01:06 PM	<DIR>		old
		1 File(s)		3,868 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros

31/07/2007	10:18 AM	<DIR>		compiled_macros
10/07/2007	12:31 PM	<DIR>		count_id
04/10/2007	09:23 AM	<DIR>		dropmissvar
31/07/2007	03:55 PM	<DIR>		dropvars
07/11/2007	12:10 PM	<DIR>		fmt_sof_vars
29/08/2007	12:52 PM	<DIR>		healthmacros
27/09/2007	03:11 PM	<DIR>		levsout
17/07/2007	11:45 AM	<DIR>		nzcms
03/12/2007	12:51 PM		254,976	sasmacr.sas7bcats
07/11/2007	05:13 PM	<DIR>		soffmts
12/07/2007	02:02 PM	<DIR>		sofiemacros
27/11/2007	09:39 AM	<DIR>		sofsess
10/07/2007	11:00 AM	<DIR>		transpose
		1 File(s)		254,976 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\compiled\_macros

31/07/2007	11:24 AM		806	compile_macros.sas
26/07/2007	10:41 AM	<DIR>		old
31/07/2007	11:26 AM		533	use_compiled_macros.sas
		2 File(s)		1,339 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\count\_id

12/07/2007	02:11 PM		1,879	count_ID.sas
12/07/2007	02:10 PM		2,984	count_ID_vars.sas
		2 File(s)		4,863 bytes

```

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\dropmissvar
04/10/2007  09:31 AM                2,574 dropmissvar.sas
04/10/2007  09:23 AM    <DIR>                old
                1 File(s)                2,574 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\dropvars
31/07/2007  04:13 PM                3,525 dropvars.sas
31/07/2007  03:55 PM    <DIR>                old
                1 File(s)                3,525 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\fmt_sof_vars
03/12/2007  11:14 AM            19,245 fmt_sof_vars.sas
07/11/2007  12:10 PM    <DIR>                old
                1 File(s)            19,245 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\healthmacros
05/09/2007  12:36 PM            11,884 sf36calc.sas
06/09/2007  04:34 PM            3,944 sf36chk.sas
                2 File(s)            15,828 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\levsout
04/09/2007  11:40 AM            2,103 catvar_levels.sas
12/07/2007  03:48 PM            3,823 check_levs.sas
04/09/2007  05:07 PM            3,820 levout.sas
                3 File(s)            9,746 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\nzcms
17/07/2007  11:06 AM    <DIR>                filepres
05/07/2007  09:50 AM    <DIR>                nzcmsmacros
17/07/2007  10:41 AM    <DIR>                recontab
17/07/2007  11:06 AM    <DIR>                rndround
                0 File(s)                0 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie
macros\nzcms\filepres
17/07/2007  11:33 AM            1,322 filepres.sas
                1 File(s)            1,322 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie
macros\nzcms\nzcmsmacros
05/07/2007  09:38 AM        459,031 nzcmsMacros.sas
05/07/2007  09:35 AM        458,193 nzcmsMacros_kr.sas
05/07/2007  09:49 AM    <DIR>                old
                2 File(s)        917,224 bytes

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie
macros\nzcms\recontab
17/07/2007  10:41 AM    <DIR>                .

```

```
17/07/2007 10:41 AM <DIR> ..
17/07/2007 11:08 AM 3,490 recontab.sas
1 File(s) 3,490 bytes
```

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\nzcms\rndround

```
17/07/2007 11:34 AM 1,965 rndround.sas
1 File(s) 1,965 bytes
```

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\soffmts

```
31/08/2007 02:33 PM 5,998 char2num.sas
03/12/2007 11:05 AM 148,480 formats.sas7bcats
07/11/2007 05:13 PM <DIR> old
03/12/2007 10:51 AM 32,721 soffmts.sas
19/09/2007 01:17 PM 164 test_soffmts.sas
4 File(s) 187,363 bytes
```

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\sofsess

```
27/11/2007 09:39 AM <DIR> old
27/11/2007 09:39 AM 5,107 sofsess.sas
1 File(s) 5,107 bytes
```

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\transpose

```
27/09/2007 12:46 PM <DIR> tolong
27/09/2007 03:11 PM <DIR> towide
10/07/2007 11:00 AM <DIR> tpose
0 File(s) 0 bytes
```

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\transpose\tolong

```
27/09/2007 12:46 PM <DIR> old
10/07/2007 10:16 AM 2,503 tolong.hlp
27/09/2007 01:11 PM 11,171 tolong.sas
2 File(s) 13,674 bytes
```

Directory of P:\MAA2006-07 SoFIE A\SASProgs\sofie macros\transpose\towide

```
25/09/2007 10:41 AM <DIR> old
10/07/2007 10:17 AM 3,119 towide.hlp
27/09/2007 01:39 PM 8,850 towide.sas
27/09/2007 03:10 PM 2,143 var_ntl.sas
3 File(s) 14,112 bytes
```