2012–13 Home purchase assistance data collection

Jurisdiction data manual

June 2013

Australian Institute of Health and Welfare

Canberra

For further information or queries please contact:

Peta Nicholson

Housing Unit

Australian Institute of Health and Welfare

GPO Box 570

Canberra ACT 2601

Phone: (02) 6249 5116

Email: housing@aihw.gov.au

Contents

1 Introduction

This data manual has been developed by the Australian Institute of Health and Welfare (AIHW) to assist jurisdictions with the collection and reporting of data for home purchase assistance. This manual provides the standard concepts, definitions and procedures to ensure data are comparable across jurisdictions. There are several changes relative to last year's collection. These changes are outlined in section 3 and include the introduction of a data validator for jurisdictional use, additional clarification of data item definitions, and a data quality information form. We encourage jurisdictions to contact the AIHW (email housing@aihw.gov.au) to address any queries you might have regarding these changes, or any other aspect of the collection.

The manual describes:

- the scope of the data collection
- specifications for quantitative and qualitative data
- tools used in compiling data
- steps for collecting and supplying data.

2   Process overview

The processes for the 2012–13 Home Purchase Assistance (HPA) data collection are illustrated in Table 2.1. There are three main steps in the process:

*   Step 1—The AIHW sends data collection documentation to jurisdictions.
*   Step 2—Jurisdictions compile data (both quantitative and qualitative), conduct quality assurance checks and send data to AIHW.
*   Step 3—The AIHW quality checks data and prepares final results for approval and reporting.

Key dates

Table 2.1: Key dates

Date	Task	Further information
19 Jun 2013	AIHW circulates data collection documentation to jurisdictions	Section 4
2 Aug 2013	FINAL date for jurisdictions to send both quantitative and qualitative data to AIHW	Section 4
30 Aug 2013	FINAL date for jurisdictions to send data quality information to AIHW	Section 4
Aug-Sept 2013	Ongoing consultation with jurisdictions to finalise data and associated footnotes	
6 Sept 2013	FINAL date for jurisdictions to submit revised data and provide information for footnote content	
13 Sept 2013	AIHW to send data and footnotes to jurisdictions for final sign-off	
20 Sept 2013	Jurisdictional data custodians to sign-off on data and footnotes	
27 Sept 2013	AIHW to send final data to Productivity Commission for national reporting	

Processing and quality assurance

Jurisdictions are required to utilise the AIHW HPA data validator prior to submitting data. The AIHW will undertake processing and further quality assurance checks of the data supplied by jurisdictions. The AIHW will liaise with jurisdictions to address any data issues identified in the finalisation of data, footnotes and data quality information.

National reporting

Once all data are finalised and approved by the jurisdictional Housing and Homelessness Information Management Group (HHIMG) representative, they are used to report to the Council of Australian Governments (COAG) against the National Affordable Housing Agreement (NAHA) Output (e) 'Number of people receiving home purchase assistance'. The measure is 'Number of households receiving home purchase assistance from State Housing Authorities', defined as the "Count of all households receiving home purchase assistance from State Housing Authorities and is presented as a number".

3   Scope and coverage

Home purchase assistance for the purpose of this collection relates to the provision of financial assistance to households to improve their access to home ownership and **<u>includes</u>**:

- direct lending (including government loans, shared equity loans and bridging loans)
- deposit assistance
- interest rate assistance
- mortgage relief
- other assistance grants.

<u>The collection</u> **excludes**:

- non-financial assistance, e.g. home purchase advisory and counselling services
- home renovation/maintenance services
- relocation or start up assistance
- sale to tenant programs
- the provision of housing or any share of it
- any expense incurred in providing assistance to a households that is not the value of financial assistance directly received by the household
- any aspect of a shared equity loan that is not direct lending, deposit assistance, interest rate assistance or mortgage relief
- any assistance not provided expressly for the purpose of purchasing a home.

Time period reference

Data to be provided for the 2012-13 collection is for the collection reference period 1 July 2012 to 30 June 2013, and should include records of assistance to households that:

- commenced receiving assistance in the 2012-13 financial year

- commenced receiving an ongoing form of assistance in a previous financial year <u>and</u> continued to receive this assistance in the 2012-13 financial year (in this instance both these criteria must be fulfilled for the record to be included in the 2012-13 collection), and those households that

- received a repayable form of assistance in a previous financial year for which repayable monies remained outstanding at the commencement of the 2012-13 collection year, i.e., where monies had not been completely repaid by COB 30 June 2012.

Reporting frequency

This data collection occurs annually.

Data definitions

Data definitions are provided in section 4; Data provision. The main counting unit for the outputs is households and a household is defined as:

'A group of two or more related or unrelated people who usually reside in the same dwelling, and who make common provision for food or other essentials for living; or a single person living in a dwelling who makes provision for his or her own food and other essentials for living, without combining with any other person'.

The 2012–13 data collection

As per last year all quantitative data should be provided in a single comma delimited (csv) file. A csv file has been sent out with this data manual to aid jurisdictions in compiling their quantitative data. Also consistent with last year's collection is that qualitative data should be supplied in the template provided and that all data must be supplied to the AIHW using the AIHW Secure Messaging (ASM) service (see Section 7).

There are a number of changes to this year's collection tools. These are:

- Introduction of a data validator (see section 5)

- Introduction of a data quality information form (see section 4)

- Additional clarification of the collection reference period (see section 3)

- Additional clarification of in-scope assistance (see section 3)

- Additional qualifying information on data items (see section 4)

- Explicit definition of payment type categories (see section 4)

- Inclusion of a summary box highlighting key features of quantitative data provision (see section 4).

Each of these changes is aimed at addressing issues that arose during last year's collection. We hope you find the changes helpful and user friendly.

There are also changes to the list of summary data items (see section 4). These are:

- Disaggregation of households assisted for year ending 30 June 2013 into households that received transfers of assistance in the 2012-13 collection year and those that had monies outstanding on repayable assistance provided in a previous financial year as at 1/7/2012.

- Change of labels for HP3-HP7. These outputs previously listed the total number of households assisted by regional breakdown. These outputs are now listed as reporting the total number of instances of assistance provided by regional breakdown. This change does not reflect a change in the underlying methodology used to calculate these outputs, but rather the new label better reflects the methodology used to calculate these figures.

4   Data provision

Jurisdictions' data submission to the AIHW must comprise the following:

- a csv file containing quantitative data
- a data validator output file pertaining to the submitted csv file
- an Excel file providing qualitative data on assistance programs
- a completed data quality information form.

The first three of these must be submitted by the 2nd of August 2013. The completed data quality information form is to be submitted by the 30th of August 2013.

Quantitative data

All data items should be provided in a single comma delimited (csv) file. A csv file has been sent out with this data manual to aid jurisdictions in compiling their quantitative data. This csv file includes headings for each data item within the data file, and data must be entered in line with the headings.

A separate record should be provided for each type of assistance provided to a household. For each type of assistance where assistance was ongoing, provide only a single record for that assistance in the 2012-13 collection, regardless of the financial year in which assistance commenced.

Guidance on scope, that is, on the household assistance records that should be included in the data set in terms of both types of assistance and timing of assistance, is provided in section 3.

The following table lists the data items to be reported by jurisdictions for the 2012–13 home purchase assistance data collection, and the order in which the data items should be provided. The data items are grouped into household data items, property data items and assistance data items. Tables 4.2, 4.3 and 4.4 display the applicable values and formats for each of these data item groupings and use the following notation.

- *Numeric n* specifies a string of up to 'n' digits.
- *Alphanumeric n* specifies a string of any combination of digits and characters up to a length of 'n'.
- *Alphabetic n* specifies a string of up to 'n' characters.
- *DD/MM/YYYY* specifies any date formatted as shown, e.g. 05/08/2006.

Table 4.1: Data items for 2012–13 HPA data collection

Household data items	Household identifier
	Main applicant identifier
	Date of birth
	Indigenous status
	Gross weekly income
Property data items	Suburb/town/locality name
	State
	Postcode
Assistance data items	Type of assistance received
	Payment type
	Date assistance received
	Amount of assistance received

Household data items

Table 4.2: household data items formats and values

Data item	Description	Format	AIHW values
ID	Unique household identifier	Alphanumeric 15	
APPLICANTID	Unique main tenant identifier	Alphanumeric 15	
DOB	Date of birth of main applicant	DD/MM/YYYY	
INDIG	Indigenous status of household	Numeric 1	1. Yes
			2. No
			9. Not stated/inadequately described
INC_GH	Gross weekly income of household	Numeric 8	

Data qualifications:

- If household identifiers are not assigned as part of general management processes, please assign a unique number to each household for the purposes of this collection.

- **The main applicant identifier should be the same person identifier that is used by the social housing authority across different housing programs. This field may be used to link home purchase assistance records with corresponding public rental housing, state owned and managed Indigenous housing, and/or private rent assistance records. If no such identifier exists, this field should be left blank.**

- Date of birth should be for the person or principal person whose name appears first on the home purchase assistance application. Where this is not clear, it should be for the person who is responsible for mortgage repayments.

- For the purposes of this collection, a household is considered Indigenous if the household contains one or more persons who identifies as being of Aboriginal or Torres Strait Islander origin.

- Gross weekly income is the value of weekly income from all sources before any deductions such as income tax, superannuation, etc. for all household members. Gross income is regarded as all receipts that are received regularly and are of a recurrent nature. Certain receipts such as lump sum receipts, windfall gains and withdrawals from savings are not considered to conform to these criteria and are not included as income.

Property data items

Table 4.3: Property data items formats

Data item	Description	Format	AIHW values
SUBURB	Suburb/town/locality name	Alphanumeric 15	
STATE	State	Alphabetic 3	
POSTCODE	Postcode	Numeric 4	

Data qualifications:

- Suburb/town/locality name may be a town, city, suburb or commonly used location name such as a large agricultural property or Aboriginal community where the dwelling for which assistance was provided is located. Postcode is the numeric descriptor for a postal delivery area, aligned with locality, suburb or place for the address of the dwelling for which assistance was provided.
- State is the state in which the dwelling for which assistance was provided is located.

Assistance data items

Table 4.4: Assistance data items formats and values

Data item	Description	Format	AIHW values
HPA_TYPE	Type of assistance received	Numeric 2	1. Direct lending
			2. Deposit assistance
			3. Interest rate assistance
			4. Mortgage relief
			99. Other
HPA_PAYMENT	Payment type	Numeric 1	1. One-off repayable
			2. One-off non-repayable
			3. Ongoing repayable
			4. Ongoing non-repayable
HPA_DATE	Date assistance received	DD/MM/YYYY	
HPA_AMOUNT	Amount of assistance received	Numeric 8	

Data qualifications:

- For the data item HPA_PAYMENT, which describes the HPA payment type, jurisdictions must assign one of the following four categories to each record;
  - one-off repayable
  - one-off non-repayable

- ongoing repayable

- ongoing non-repayable

For the purposes of this collection, 'one-off' denotes assistance consisting of a single transfer to a household, while 'ongoing' denotes assistance that comprises multiple transfers to a household at different points in time. Thus, 'ongoing repayable' assistance is where repayable assistance consists of multiple repayable transfers to a household that occur at different points in time, whereas an instance of 'one-off repayable assistance' consists of a single repayable transfer to a household. <u>In this context, 'ongoing' does not refer to a loan for which monies remain outstanding, but refers to whether the initial assistance consisted of one or multiple transfers to the household.</u>

- For all households, report the originating date of the assistance (i.e. the date the household first received the assistance), including those households who commenced receiving assistance in a previous financial year (i.e. prior to 1 July 2012).

- For those households who commenced receiving assistance in a previous financial year (i.e. prior to 1 July 2012) and they continued to receive this assistance, include only the value of assistance provided for the year ending 30 June 2012 (i.e. disregard the value of assistance provided prior to 1 July 2011).

Include:

- The value of assistance provided to a household during the 2012–13 financial year for:
  - all households who commenced receiving assistance for the year ending 30 June 2013 regardless of the form of assistance (i.e. whether assistance is one-off, ongoing, repayable or non-repayable); and
  - all households who commenced receiving an ongoing form of assistance in a previous financial year (i.e. prior to 1 July 2012) and they continued to receive this assistance from 1 July 2012.

Exclude:

- Related administrative and operational costs associated with providing the home purchase assistance.

- The value of assistance provided to a household prior to 1 July 2012 (i.e. for households continuing to receive ongoing forms of assistance that commenced prior to 1 July 2012).

- The value of outstanding repayable monies where a repayable form of assistance was provided prior to 1 July 2012 and outstanding monies had not been repaid by COB of 30 June 2012 (i.e. report HPA_TYPE, HPA_PAYMENT and HPA_DATE but *do not* report HPA_AMOUNT).

- For those households that had monies outstanding on repayable assistance provided in a previous financial year assistance that received no new assistance in the 2012-13 financial year, leave HPA_AMOUNT blank.

Summary Box: Key points on quantitative data provision

- Provide a separate record for each type of assistance provided to a household.

- For each type of assistance where ongoing assistance was provided, only a single record should be provided for that assistance in the 2012-13 collection, regardless of which financial year assistance commenced.

- Include any household that received assistance in the 2012-13 financial year.

    o This includes households who commenced receiving assistance in the 2012-13 financial year.

        § In this case, the value of assistance, HPA_AMOUNT, should be the value of assistance paid to the household in the 2012-13 financial year.

    o It also includes those households that commenced receiving ongoing assistance in a previous year and continued to receive this assistance in the 2012-13 financial year.

        § In this case, the value of assistance, HPA_AMOUNT, should be the value of assistance paid to the household in the 2012-13 financial year. In this context, this means excluding the value of assistance provided in any previous financial year.

- Include any household that has repayable monies outstanding for assistance provided in a previous financial year. That is, where repayable monies have not been completely repaid by COB of the 30 of June 2012.

        § In this case, the value of assistance, HPA_AMOUNT, should be left blank.

The above summary box outlines three different household assistance contexts and the AIHW's data item requirement for the amount of assistance received for each of these three contexts. In each of these contexts only the value of assistance received during the 2012-13 financial year should be included. In all cases, the value of assistance provided in a previous financial year should be excluded. Thus, outstanding loan amounts or loan principles provided in previous financial years for which monies remain outstanding are not captured by the HPA data collection. Hence, while the collection captures some households that may not have received new financial assistance in the 2012-13 financial year (households that continue to benefit from home purchase assistance due to repayable assistance provided in a previous financial period for which monies remain outstanding), the collection only captures the dollar value of *new* financial assistance provided in the 2012-13 financial year.

Summary data items

The following table provides details of the summary output items that will be calculated and published by the AIHW based on the household, property and assistance data items reported by jurisdictions.

Table 4.5: List of output items

Code	Description
HP1	Total number of households assisted for year ending 30 June 2013[a]
HP1a	Total number of newly assisted households for year ending 30 June 2013
HP1b	Total number of households continuing to hold assistance loans provided prior to 30 June 2012
HP2	Total number of Indigenous households assisted for year ending 30 June 2013
HP3	Total number of instances of assistance provided to households in Major cities of Australia for year ending 30 June 2013
HP4	Total number of instances of assistance provided to households in Inner regional areas of Australia for year ending 30 June 2013
HP5	Total number of instances of assistance provided to households in Outer regional areas of Australia for year ending 30 June 2013
HP6	Total number of instances of assistance provided to households in Remote areas of Australia for year ending 30 June 2013
HP7	Total number of instances of assistance provided to households in Very remote areas of Australia for year ending 30 June 2013
HP8	Total number of new households receiving direct lending for year ending 30 June 2013
HP9	Total number of new households receiving deposit assistance for year ending 30 June 2013
HP10	Total number of households receiving interest rate assistance for year ending 30 June 2013
HP11	Total number of households receiving mortgage relief for year ending 30 June 2013
HP12	Total number of households receiving other assistance for year ending 30 June 2013
HP13	Total value of assistance to new households receiving direct lending for year ending 30 June 2013
HP14	Total value of assistance to new households receiving deposit assistance for year ending 30 June 2013
HP15	Total value of assistance to households receiving interest rate assistance for year ending 30 June 2013
HP16	Total value of assistance to households receiving mortgage relief for year ending 30 June 2013
HP17	Total value of assistance to households receiving other assistance for year ending 30 June 2013

(a)   Includes those households with outstanding monies to repay for assistance provided in a previous financial year where no assistance was provided for year ending 30 June 2012 (i.e. direct lending and deposit assistance).

Qualitative data

Jurisdictions are requested to provide details of all programs of home purchase assistance (financial only) available to households during the 2012–13 reporting period. An Excel spreadsheet is provided for this purpose and includes the following fields:

- name of program
- description of program
- eligibility criteria.

The final date for jurisdictions to send this qualitative data to the AIHW is the same as that for quantitative data, the 2nd of August 2013.

Data quality information

Collection materials circulated to jurisdictions included a data quality information form. The form provides space for data quality information to be provided against each data item as well as an open ended section where jurisdictions should provide any other data quality information. The AIHW will use this information to compile collection data quality statements and footnotes to published data.

For data quality information going into the data quality statements we're very interested in things that affect 'accuracy' and 'coherence'. These are the names of the two sections in the data quality statements that are generally populated by information received from jurisdictions.

The section on 'accuracy' should include information on:

- exactly what has been reported, i.e. any deviation from the data items specifications outlined in section 4

- any aspect of collection methodology that potentially affects what actually ends up being reported against a particular data item, or summary data item, such as the point in time where information is collected, i.e., is the household information current at the end of financial year or the date assistance is received?

- anything that might introduce disparity between the reported data items and the 'true' values

- any factors that might have affected data quality, and if it is known, the direction of any bias that may have been introduced

- any deviation from the collection scope as outlined in section 3.

The section on 'coherence' should include information on:

- changes in what has been reported over time, in terms of alignment to data item specifications outlined in section 4

- changes in collection methodology that might affect what actually ends up being reported against a particular data item, or summary data item

- any factors that may have resulted in a change in data quality, and if it is known, whether the change in data quality is thought to be an improvement or not

- changes in scope compared to previous collection years

- anything that affects the comparability of data across time and between jurisdictions.

For both of these sections of the data quality statements we're interested in information that affects the whole of the collection, individual data items or summary data items, or groups of data items or summary data items. We're keen for you to be as specific and exhaustive as possible. In light of this explanation, please complete the data quality information form providing as much data quality information as possible. The completed form is due by the 30th of August.

5   Processing performed by jurisdictions

Jurisdictions are required to undertake the data quality checks outlined in the following table prior to submission to the AIHW. Jurisdictions are also required to utilise the HPA data validator prior to submission to the AIHW. Utilisation of the AIHW data validator is not a substitute for the data quality checks outlined in the table below. The data validator will assist with ensuring that the data conforms to formatting requirements and checks some readily testable data item relationships.

Data quality checks

Data item & definition	Data quality check
Household identifier A unique identifier for a household. If household identifiers are not assigned as part of general management processes, please assign a unique number to each household for the purposes of this collection.	Must be completed for all household records and must be unique for each household.
Main applicant identifier A unique identifier for the person or principal person whose name appears first on the home purchase assistance application form. This identifier should be the same person identifier that is used by the social housing authority across different housing programs. This field may be used to link home purchase assistance records with corresponding public rental housing, state owned and managed Indigenous housing, and/or private rent assistance records. If a consistent social housing authority identifier is not available, please leave this field blank.	Must be completed for all household records where a consistent social housing authority person identifier that can be used for data linkage exists. Must be left blank for all household records where a consistent social housing authority person identifier *does not* exist.
Date of birth of main applicant The date of birth of the person or principal person whose name appears first on the home purchase assistance application form. Where this is not clear, it should be the person who is responsible for mortgage repayments.	Please record in the DD/MM/YYYY format and leave blank if unknown.
Indigenous status of household A household that contains one or more persons who identifies as being of Aboriginal or Torres Strait Islander origin.	Only valid codes are accepted (i.e. 1, 2 or 9). Must be completed for all household records.
Gross weekly household income The value of weekly income from all sources before any deductions such as income tax, superannuation, etc. for all household members. Gross income is regarded as all receipts that are received regularly and are of a recurrent nature. Certain receipts such as lump sum receipts, windfall gains and withdrawals from savings are not considered to conform to these criteria and are not included as income.	Please check records where weekly income is above $1,500 or below $150. Report to 2 decimal places and leave blank if unknown.
Suburb/town/locality name of property The suburb/town/locality name may be a town, city, suburb or commonly used location name such as a large agricultural property or Aboriginal community.	Must be completed for all household records.
Postcode of property The numeric descriptor for a postal delivery area, aligned with locality, suburb or place for the address of a dwelling.	Please enter a valid 4-digit Australian postcode. Must be completed for all household records.

Using the AIHW HPA data validator

This section provides guidance on using the new HPA data validator. Jurisdictions are required to run the data validator on their data prior to submitting it to the AIHW. The data validator runs edit checks that highlight where data does not conform to the allowable formats and values, and where some readily testable data item relationships are violated. Jurisdictions should run the data validator, amend data as necessary and repeat this process as necessary to ensure that all edit flags have been addressed. Jurisdictions are required to submit a single data validator output file that verifies that all records conform to the data manual and return no edit flags.

Using the data validator involves 6 steps:

1. Compiling data in the csv file provided
2. Preparing the csv file for the data validator
3. Opening the data validator
4. Loading the csv file
5. Interpreting the data validator output file
6. Addressing data errors

1. Compiling data in the csv file

Input the data into the csv file provided. For definitions of each data item and allowable values and formats, please refer to section 4 Data provision.

When saving the csv file one or both of the following dialogue boxes may appear. Click 'OK' and/or 'Yes' to save.
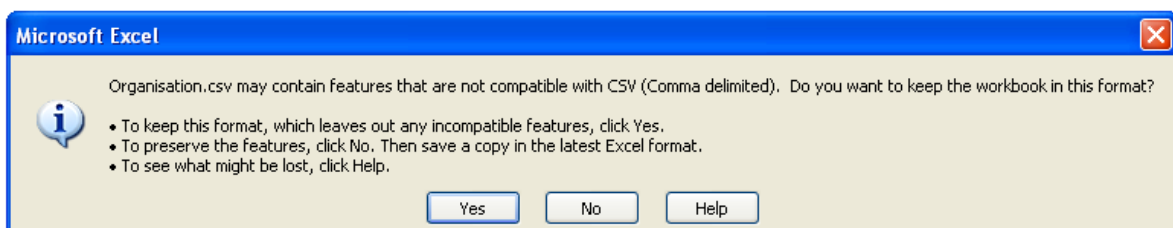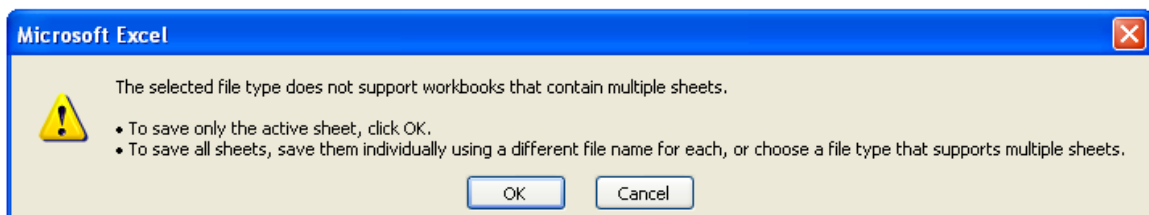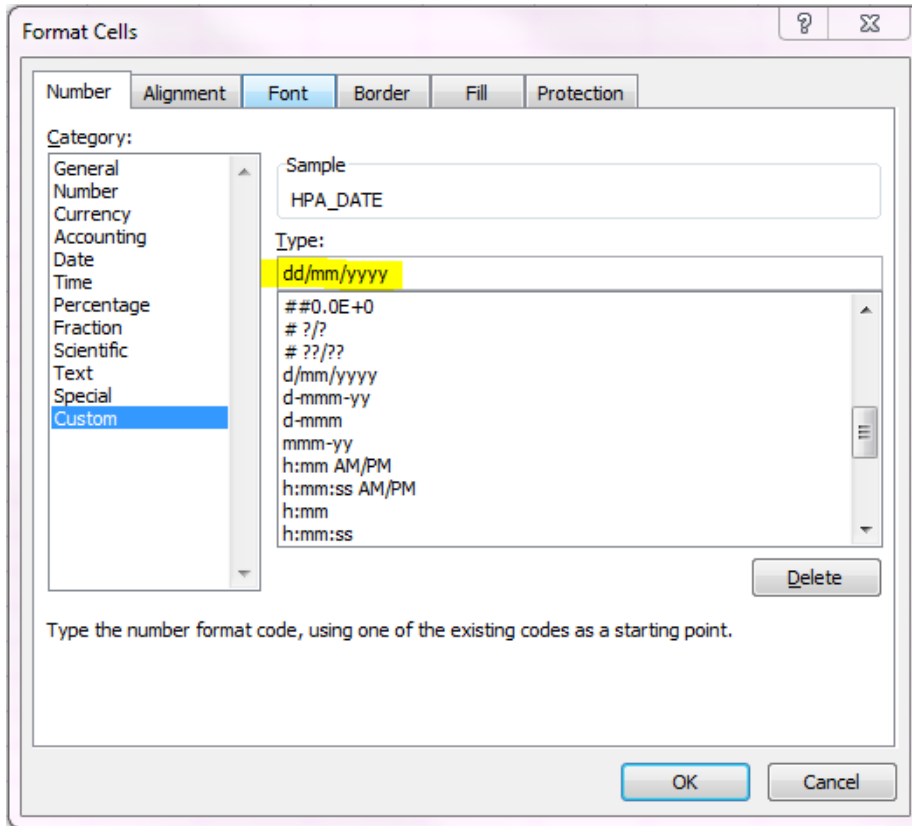

2. Preparing the csv file for the data validator

The data validator requires the csv file to be in a specific format. **Please ensure that you have performed the steps in the checklist below before using the data validator.**

Files not consistent with the checklist below may result in an error message *or false positives*. The most common error reported is the "subscript out of range (9)" error message. If you get this error, please try loading the same csv file again without making any amendments to the csv file. If the error message recurs, it indicates that the csv file is in a format that prevents the data validator from running, in which case you will need to work through the data validator preparatory formatting checklist again to ensure the data in your csv file is formatted correctly. Failure to address all points of the checklist will cause the data validator to continue to generate the error message "subscript out of range (9)". The csv file must contain at least three records and be in the correct format for the data validator to run.
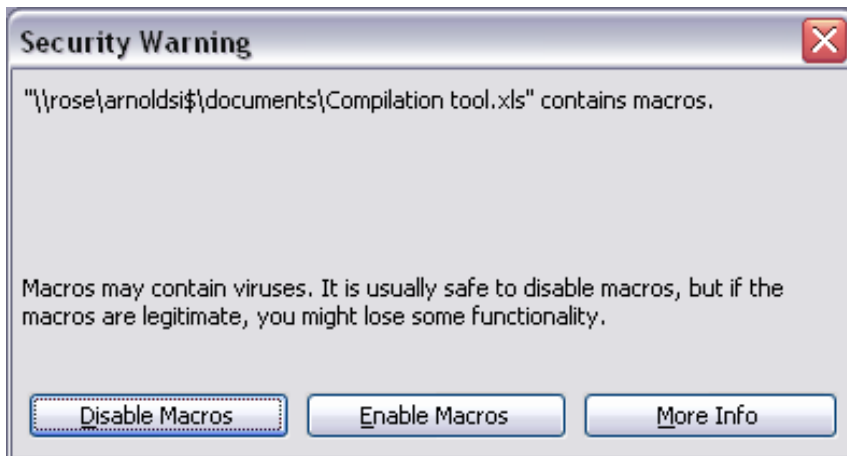
Data validator preparatory formatting checklist:

- Delete any blank rows.

- Remove all commas (,) from the data. This is most quickly done by using the 'Find and Replace' function in Excel. The short cut to the 'Find and Replace' function is Ctrl and f. Within the 'Replace' tab enter a comma (,) in the 'Find what' box and leave the 'Replace with' box empty and click 'Replace All'.

- Remove spaces from blank cells and any trailing spaces from populated cells. As a starting point to achieve this, highlight the columns that should not have any spaces in them and use the 'Find and Replace' function entering a space in the 'Find what' box and leave the 'Replace with' box empty and click 'Replace All'.

- Remove any carriage returns that may exist. These are perhaps most likely to occur in address fields.

- Make sure to remove any dashes from variables which are not identifiers. For example, if a dash has been used in the currency fields, replace with a decimal place if appropriate.

- Ensure all variables take on the appropriate format by highlighting a column and selecting Format / Cells. By default, cells will have a 'General' format before data is entered into the csv file. Once data has been entered:

  o Ensure all number fields are not formatted to separate 1000's by a comma (,)

  o Ensure all date variables (e.g. HPA_DATE and DOB) take on the "dd/mm/yyyy" format. This can be done by creating and applying a custom format as shown below:

Format Cells

| Number | Alignment | Font | Border | Fill | Protection |

Category:
General
Number
Currency
Accounting
Date
Time
Percentage
Fraction
Scientific
Text
Special
Custom

Sample
HPA_DATE

Type:
dd/mm/yyyy

##0.0E+0
?/?
??/??
d/mm/yyyy
d-mmm-yy
d-mmm
mmm-yy
h:mm AM/PM
h:mm:ss AM/PM
h:mm
h:mm:ss

Delete

Type the number format code, using one of the existing codes as a starting point.

OK   Cancel

3. Opening the data validator

When you open the Excel file, the following security warning will appear:

Security Warning

"\\rose\arnoldsi$\documents\Compilation tool.xls" contains macros.

Macros may contain viruses. It is usually safe to disable macros, but if the macros are legitimate, you might lose some functionality.

Disable Macros   Enable Macros   More Info

Click on 'Enable Macros' to continue and the file should open successfully.

4. Loading the csv file into the data validator

First:

- o Ensure that 'Check this box if the data file contains a header row' is ticked.
- o Enter the maximum number of errors to display for each item (AIHW recommends 100).

16

Then:

        o   Click on the purple 'HPA File' button
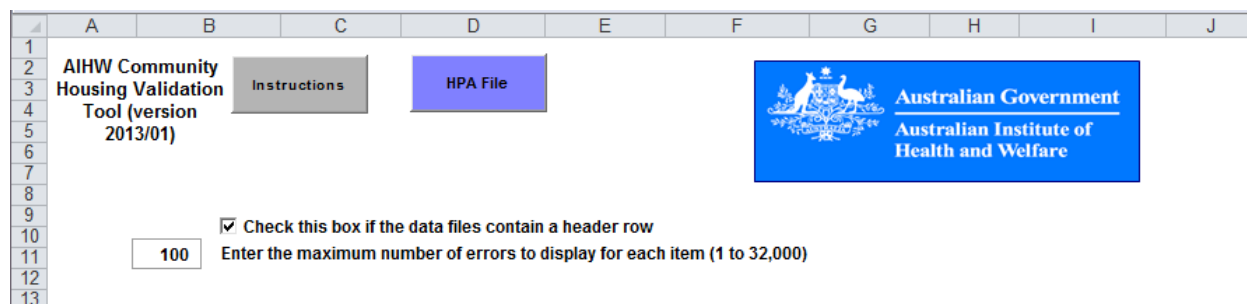
You will be asked to select a comma delimited file to load (i.e. a csv file) and click 'Open'.
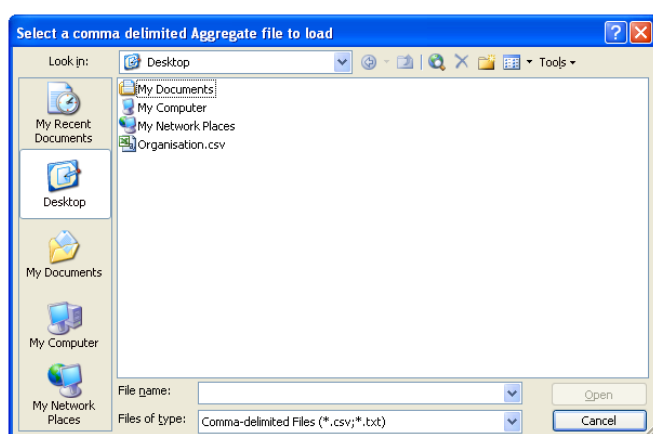
The next dialogue box asks you to enter an output filename and to select a location for the file to be saved. The html data validator output file will be saved in the location you nominate. It should also open automatically when the data validator has finished running. If this does not occur, the file can be opened manually.

5. Interpreting the data validator output file

Data validator output file layout

The date and time of data validation is given at the top of the file, along with how many records the data validator loaded from the csv file. Any failed edit checks will follow.

The different edits are separated by the 'Pre-Dispatch Edits' headings. Each 'Pre-Dispatch Edit' has;

- an error code and title (e.g. H02 – Missing household identifier),
- the reason the listed records failed the edit check (e.g. ID is missing),
- a message indicating your required action (e.g. this field must be populated),
- a table that contains the details of the records that have failed the edit check, and
- the total number of records failing the edit check displayed beneath the table. If the number of failed records exceeds the number entered for 'the maximum number of errors displayed for each item' (nominated when loading the csv file,

the default is 100), the following message will be displayed "Problem count exceeds 100; No more records will be shown...".

Edits will not appear in the output if no records failed the edit check.

16/05/2013 11:13:30 AM

HPA File loaded: 1,067 records

PRE-DISPATCH EDITS

H02 - Missing household identifier:
ID is missing.

This field can not be missing.

Line Number	ID	APPLICANTID	DOB	INDIG	INC_GH	SUBURB	POSTCODE	HPA_TYPE	HPA_PAYMENT	HPA_DATE	HPA_AMOUNT
6			21/05/1967	9		SUNSHINE	3021	4	3	30/06/1988	

Total: 1

PRE-DISPATCH EDITS

H04 - Missing Indigenous status of household:
Indigenous status of household is missing.

This field can not be missing.

Line Number	ID	APPLICANTID	DOB	INDIG	INC_GH	SUBURB	POSTCODE	HPA_TYPE	HPA_PAYMENT	HPA_DATE	HPA_AMOUNT
22	4				65000	GEELONG	321	1	1	1/09/1967	

Total: 1

Identifying which records failed an edit check

Each row of the Pre-Dispatch Edit table represents a record which has failed the edit check. The first column, 'Line Number', advises which record in the csv file has failed the edit check. The Line Number plus 1 equates to the excel row number in the csv file. For example, if the Line Number is 3 you will find the corresponding record in row 4 of the csv file that you uploaded.

The data item that has failed the edit check is coloured red. For relational edits checks, those that involve more than one data item, all the data items relevant to the edit check will be coloured red.

The Pre-Dispatch Edit table for duplicate records has a different format. As illustrated below, it shows the data items common to both records (ID and HPA_TYPE) coloured red, followed by 'Line No. 1' and 'Line No. 2' (the line the on which the record first occurs and the line on which a record with the same ID and HPA_TYPE occurs for the second time) coloured green.

H10 - Duplicate records:
Records have the same ID and HPA_TYPE.

Review records and delete or amend records as necessary to remove any true duplicates.

18

ID	HPA_TYPE	Line No. 1	Line No. 2
795	4	7	8
26	1	44	45

Total: 4

In the example above, the first row of the Pre-Dispatch Edit table shows that two records share a common household identifier of 795 and common type of assistance of 4, mortgage relief. The table also indicates that the records in the csv file that correspond to the record's first instance and repetition are in row 8 and row 9 of the csv file.

6. Addressing data errors

Using the information provided in the data validator output, locate the records that have failed an edit check in the csv file (Remember: Line number + 1 = row number) and make any necessary changes to the csv file. Guidance on appropriate amendments to the data is provided in the data validator output file. After the necessary changes have been made to the csv file, run the data validator again and resolve any other errors identified by the data validator. Continue this process of re-running the data validator and resolving edit returns until all data errors have been addressed.

Your quantitative data submission should include your csv file and the most recent data validator html output file that verifies that there are no records flagged by the data validator edit checks.

The AIHW will consult with jurisdictions about any data quality issues identified through internal AIHW processing. Internal processing involves more extensive quality assurance checks and may identify data issues not found by the preliminary checks conducted by the data validator. Where jurisdiction submit revised data in light of data issues identified by AIHW data processing, jurisdictions should run the data validator on their revised data and amend records as necessary prior to submission of data to the AIHW.

Send your HPA data submission to AIHW via the AIHW secure messaging (ASM) service. Refer to Section 7 of the data manual or the ASM Quick Guide for information on using the ASM.

The data validator runs the following edit checks.

Table 5.1: Data validator edit checks

Edit check	Error description
H02	Missing household ID
H04	Missing Indigenous status of household
H10	Duplicate records
H11	Invalid format for DOB
H12	Invalid format for gross weekly income of household
H13	Invalid Indigenous status of household (ie not 1, 2 or 9)
P04	Invalid postcode
A01	Invalid type of assistance received (i.e. not 1, 2, 3, 4 or 99)
A02	Invalid payment type (i.e. not 1, 2, 3 or 4)
A03	Missing type of assistance received
A04	Missing payment type
A05	Missing date assistance received
A07	Missing amount of assistance received and dated assistance received is not before 1 July 2012
A08	The date assistance received is after 30 June 2013
A09	Invalid format for date assistance received
A10	Invalid format for the amount of assistance received
A11	Amount of assistance reported for assistance before 1 July 2012 where the assistance payment is not on-going
A12	Type of assistance reported as 'Direct lending assistance' and payment type is reported as non-repayable

Note that in the edit codes above

"H" denotes a household data item edit check

"P" denotes a property data item edit check

"A" denotes an assistance data item edit check.

6   AIHW importing and edits process

The following table outlines the edits which will be undertaken by the AIHW. The results from the edits applied will be provided to each jurisdiction in a Processing Workbook. The workbook will include a summary of the edit results and samples of the records that failed each edit. The AIHW will liaise with jurisdictions to improve data quality by seeking to resolve edit flags raised during AIHW internal processing. Where the process involves jurisdictions submitting revised csv files, these files should be run through the data validator using the process outlined in section 5.

Edit ID	Description
H01	Invalid Indigenous status of household (i.e. not 1, 2 or 9)
H02	Missing household ID
H03	Missing date of birth of main applicant
H04	Missing Indigenous status of household
H05	Missing gross weekly income of household
H06	The age of the main applicant is greater than 100 years
H07	The age of the main applicant is less than 16 years
H08	The gross weekly income of the household is high (i.e. > $1,500)
H09	The gross weekly income of the household is low (i.e. < $150)
H10	Duplicate record
P01	Invalid postcode
P02	Missing suburb/town/locality name
P03	Missing postcode
A01	Invalid type of assistance received (i.e. not 1, 2, 3, 4 or 99)
A02	Invalid payment type (i.e. not 1, 2, 3 or 4)
A03	Missing type of assistance received
A04	Missing payment type
A05	Missing date assistance received
A06	Amount of assistance reported for assistance before 1 July 2012 and amount of assistance is not recorded as ongoing
A07	Missing amount of assistance received and date assistance received is not before 1 July 2012
A08	The date assistance received is after 30 June 2013

7   How to supply data to the AIHW

Sending files securely with the AIHW Secure Messaging (ASM) service

For the 2012–13 home purchase assistance data collection, the AIHW Secure Messaging (ASM) service will be used to send and receive IN-CONFIDENCE emails.

What is the AIHW Secure Messaging (ASM)

ASM is a service and procedure for sending and receiving emails and attached data and other files on the Internet more securely and reliably. It should be used for all data submissions and whenever IN-CONFIDENCE materials need to be sent to the Institute.

Registering to use ASM as an external user

To register as an external user, you will need an AIHW user to send you a message through ASM.

Once this is done, an email (shown below) will appear in your current email inbox:
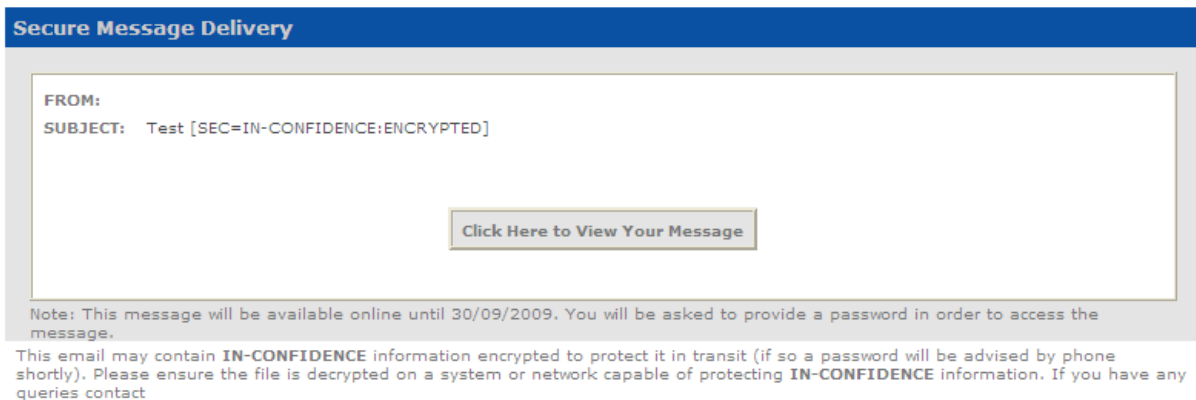
You will then need to:
1.   Click "**Click Here to View Your Message**"
2.   Fill in the blanks on the next web page that appears then click "**Continue**".

3. On successfully filling out the form, you will be logged on to ASM and your message will be displayed.

4. You can continue to use this logon until the account expires.

Logging on

1. Enter the following URL **https://envoy.aihw.gov.au** in your Internet browser.

2. Enter your email address and password and click the "**Sign In**" button when the ASM welcome\logon screen appears. (You must enter the full email address and the password used originally when registering.) Your ASM mailbox will then open and be ready for use.

Note: As a security precaution you will be required to re-logon after 5 minutes of inactivity. Waiting for a file to attach, is counted as activity.

Viewing Messages

You do not need to check your ASM mailbox regularly. When a message arrives in this mailbox you will be notified by an email in your normal email inbox Logon on to your ASM mailbox.
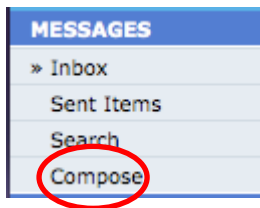
1. To see your messages click "**Inbox**" in the navigation pane on the left.
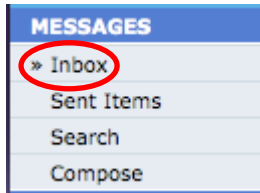
Sending Messages

As an external user, you will only be able to send emails to "**aihw.gov.au**" email addresses.

1.   Logon on to your ASM mailbox.
2.   To create a message, click "Compose" in the navigation pane on the left.

3. Enter the Recipient's email address, Subject and Message body in the spaces provided.
4. Add any Attachments.
    a. Click Browse in the Attachments section.
    b. Navigate to the file to be attached, select **–> "OK"**
    c. Click "**Upload**"

5.   Repeat until all files have been attached.
6.   Click "**Send**".
7.   A confirmation message will be displayed when the message has been sent.
8.   Once the recipient collects the message an email notification will be sent to your ASM mailbox and your Outlook inbox.

Replying to Messages

1.   Logon on to your ASM mailbox.
2.   Click "**Inbox**" in the navigation pane on the left.
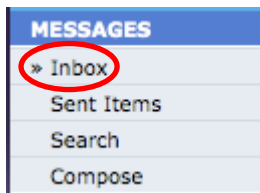
3.   Click the message you are replying to.
4.   Once the message has opened, click "**Reply**" from the toolbar in the message header.


5.   Recipients will be re-entered and the message body will be attached within the email.
6.   Compose your reply and upload any attachments, as explained in the "Sending Messages from Your ASM Web Site Mailbox" section.
7.   Click "**Send"**.

8.   Once the recipient collects the message an email notification will be sent to your ASM inbox and your Outlook inbox.

Deleting Messages

1.   Logon on to your ASM mailbox.
2.   Click "**Inbox**" in the navigation pane on the left.
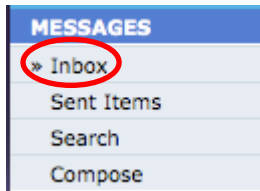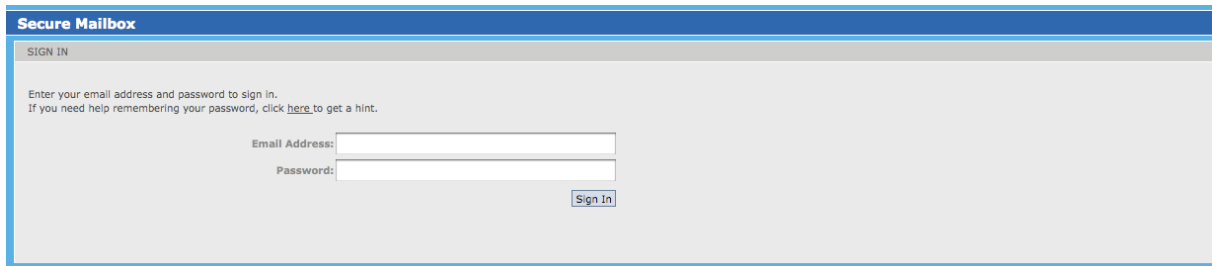

3.   Click the message you wish to delete.
4.   Once the message has opened, click "**Delete**" from the toolbar in the message header.

Important information about ASM

Session Expired Message

If you are logged into the ASM web site, but do not use it for 5 minutes you are automatically logged off and returned to the logon screen. This is a security precaution. If you still need to use the system you need to logon again.

Message No Longer Available

If a message hasn't been opened before it has expired then the following advisory message is sent to the recipient. When a message expires it is removed from the system. If the message is still required, it will need to be resent.

Further information on using ASM can be found in the ASM Quick Guide.

The AIHW would like to thank you for your time and diligence in preparing a data submission that accords with this data manual. We sincerely appreciate your efforts in providing the public and the Commonwealth with accurate and relevant information on home purchase assistance. This information informs policy makers and facilitates identification of best practice.