

2013 Locality to 2011 Statistical Area Level 2 (SA2) Coding Index

Background

The Locality Coding Index was developed to replace the 'Localities Index' within the National Localities Index (NLI) which was discontinued in 2007 (the 2007 edition was the final release of the NLI).

Since 2008 the Locality Coding Index was built from localities in the Geocoded National Address File (G-NAF) plus PSMA Australia localities along with a list of Localities/Aliases maintained by the ABS. G-NAF was then used as a weighting factor to determine which Statistical Local Area (SLA) boundary would be associated with a given Locality/Postcode/State combination.

In 2011 the ABS introduced a new statistical geography, the Australian Statistical Geography Standard (ASGS). In the ASGS the Statistical Area Level 2 (SA2) is the equivalent region to the old Australian Standard Geographical Classification (ASGC) SLAs, therefore from 2011 the Locality Coding Index was built to link localities to the new SA2s.

This 'Locality to SA2 Coding Index' utilises a new and improved methodology to allocate the 2013 localities to the 2011 SA2s. This new process is discussed in this documentation.

Description

The 'Locality to SA2 Coding Index' is a text based list of Localities with ASGS SA2 boundaries associated with them. The objective of this coding index is to aid the linking of data geocoded to Suburb/Locality to the 2011 edition of the ASGS. In cases where users have unit record data that does not contain whole addresses, but does contain suburb/locality and Australia Post postcode information, an association with the SA2 level of the ASGS can be obtained. In this instance the postcode is important as Locality names are not necessarily unique, even within a State/Territory – the postcode aids in identifying the correct Locality.

The index functions by searching for a Locality/Postcode/State combination to see which SA2 is associated with that locality. In cases where a Locality/Postcode/State combination covers more than one SA2 the boundary with the greater population is chosen as the associated SA2. Care should be taken when applying coding indexes, as the transformed areas may not be wholly within their associated ASGS region. Coding indexes do not apportion data – they assign whole areas to the most appropriate region within the new geography. As a general rule, small areas can be assigned more accurately than large areas.

During the development of the ASGS, SA2 boundaries were designed to reflect officially gazetted Suburb/Locality boundaries as close as possible. If collection units contain the suburb/locality information, they generally can be accurately coded to the SA2.

Format

The 'Locality to SA2 Coding Index' is provided as a delimited text file. Commas are used as the delimiter and double quotes are used to identify text.

Format of the 'Locality to SA2 Coding Index':

```
"LOCALITY_ID","LOCALITY_NAME","LOCALITY_TYPE","POSTCODE","STATE","SA2_MAINCODE","SA2_NAME",RATIO,PERCENT,
"ACT107","ACTON","GAZETTED LOCALITY","2601","ACT","801051049","Acton",0.9999982,99.9998158,
"ACT107","ANU","ALIAS LOCALITY","2601","ACT","801051049","Acton",0.9999982,99.9998158,
"ACT107","AUSTRALIAN NATIONAL UNIVERSITY","ALIAS LOCALITY","2601","ACT","801051049","Acton",0.9999982,99.9998158,
"ACT107","SPINNAKER ISLAND","ALIAS LOCALITY","2601","ACT","801051049","Acton",0.9999982,99.9998158,
"ACT107","SPRINGBANK ISLAND","ALIAS LOCALITY","2601","ACT","801051049","Acton",0.9999982,99.9998158,
"ABS195","ACTON PARK","EXTRA LOCALITY","2601","ACT","801021028","Molonglo",1,100,
```

Source Data

The source of localities and Aliases for the 'Locality to SA2 Coding Index' are detailed below:

Localities are sourced from:

- PSMA Australia Gazetted Localities
- ABS Extra Localities

Aliases are sourced from:

- PSMA Australia Aliases
- ABS Aliases

The attributes for the 'Locality to SA2 Coding Index' are as follows:

Data Item	Source	Description
LOCALITY_ID	August 2013 PSMA edition of Gazetted Localities and common aliases. ABS Extra Localities and ABS Aliases file.	ID associated with the locality, also indicates the different datasets that the locality is sourced from.
LOCALITY_NAME	August 2013 PSMA edition of Gazetted Localities and common aliases. ABS Extra Localities and ABS Aliases file.	The Name of the gazetted locality, alias or extra locality.
LOCALITY_TYPE	August 2013 PSMA edition of Gazetted Localities and common aliases. ABS Extra Localities and ABS Aliases file.	Informs the client whether the locality is: Gazetted Locality, Alias Locality or Extra Locality.
POSTCODE	May 2013 Pitney Bowes / MapInfo edition of the Australia Post Postcodes.	The Australia Post Postcode associated with each gazetted locality, alias or extra locality.
STATE	2011 ASGS	State/Territory Abbreviation
SA2_MAINCODE	2011 ASGS	Maincode for the Statistical Area Level 2 (SA2) from the Main Structure of the ASGS.

SA2_NAME	2011 ASGS	Name of the Statistical Area Level 2 (SA2) from the Main Structure of the ASGS.
RATIO	Created during the association process in FME.	Describes the Ratio of the 'FROM' region that is being donated to the 'TO' region. The Ratio is a figure between 0 and 1. This field gives the client an indication as to how good the association between the localities and the SA2 is.
PERCENT	Created during the association process in FME.	Describes the Percentage of the 'FROM' region that is being donated to the 'TO' region. The Percentage is the Ratio multiplied by 100. This field gives the client an indication as to how good the association between the localities and the SA2 is.

Further information on the main structure of the ASGS and the SA2 digital boundaries can be found in the following product:

[Australian Statistical Geography Standard \(ASGS\): Volume 1 – Main Structure and Greater Capital City Statistical Areas, July 2011 \(cat no. 1270.0.55.001\)](#)

The Process

The 'Locality to SA2 Coding Index' is built using an FME script process. The FME script is based on the FME script that is used to build the population weighted Mesh Block (MB) grid based correspondences. This method is essentially a series of grid points that represent the underlying geographical distribution of the weighting unit, which in this case is the MB total population. Each grid point is assigned a value based on this weighting. The next step in this process is to determine the proportion that the locality, as the FROM unit, is donating to the respective SA2 TO units. The proportion is calculated by dividing the population found in each of the TO regions by the total population of the FROM region. Further information on the generation of correspondences can be found in the following publication:

[Australian Statistical Geography Standard \(ASGS\): Correspondences, July 2011 \(cat no. 1270.0.55.006\)](#)

However, unlike the population weighted grid based correspondences, instead of outputting all ratios for statistical boundaries covering a given Locality/State/Postcode combination, only the locality with the largest ratio of population for the given statistical boundary is output. This means where a locality is covered by two or more SA2s the locality has been allocated to the SA2 where most of its population are located according to the underlying grid. This ensures a specific gazetted locality/alias/extra locality entry is represented only once in the file and is allocated wholly to a specific SA2.

The ratio and percentage have remained in the file to allow clients to evaluate how close a match the locality is to the SA2. It provides the client with an indication as to how good the association between the specific locality and the SA2 is. The higher the ratio/percentage the better the match between locality and SA2, and the more confident you can be in the quality of the association.

Working with the 'Locality to SA2 Coding Index'

The 'Locality to SA2 Coding Index' is a comma delimited ASCII text file (.txt) which has been designed to be easily loaded into a wide variety of computer systems and database software. All the fields (columns) within this file are delimited with a comma and all the data items are surrounded by double quotes. The double quotes allow software packages (such as Microsoft Excel) to identify the data items as text. This is important when dealing with Postcodes that can be associated with leading zeros.

The ABS recommends that all ASGS codes should be stored as text. This prevents the application of numerical operators to the classification codes.

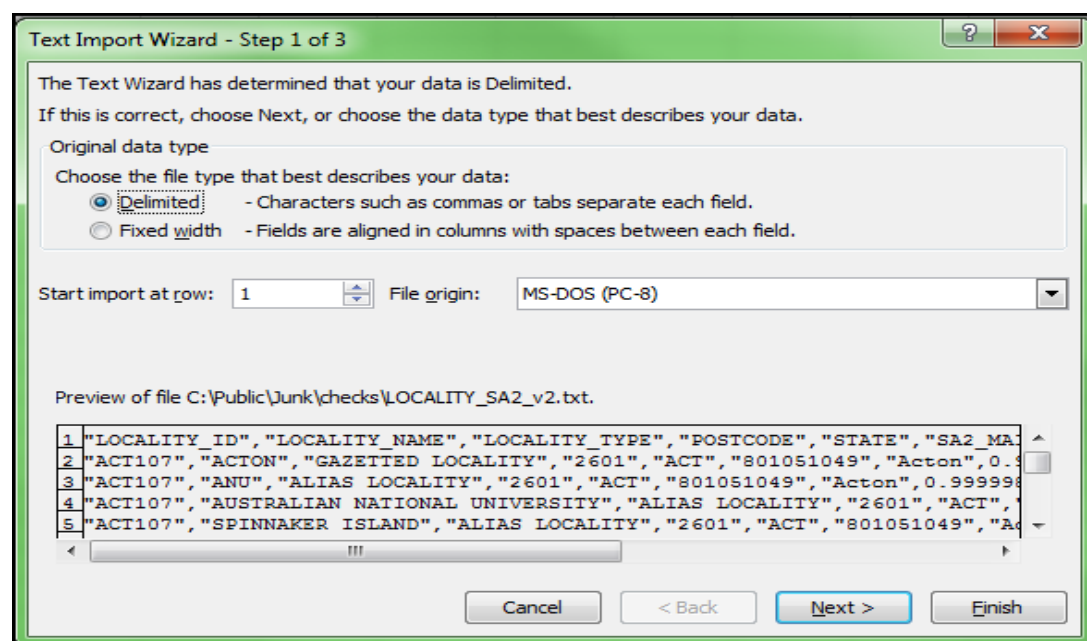
Loading the 'Locality to SA2 Coding Index' with Microsoft Excel 2010/2011

The 'Locality to SA2 Coding Index' can be read into Microsoft Excel, but to force Microsoft Excel 2010/2011 to format fields as text rather than numeric the imported file needs to be processed by the 'Text Import Wizard'.

Step 1

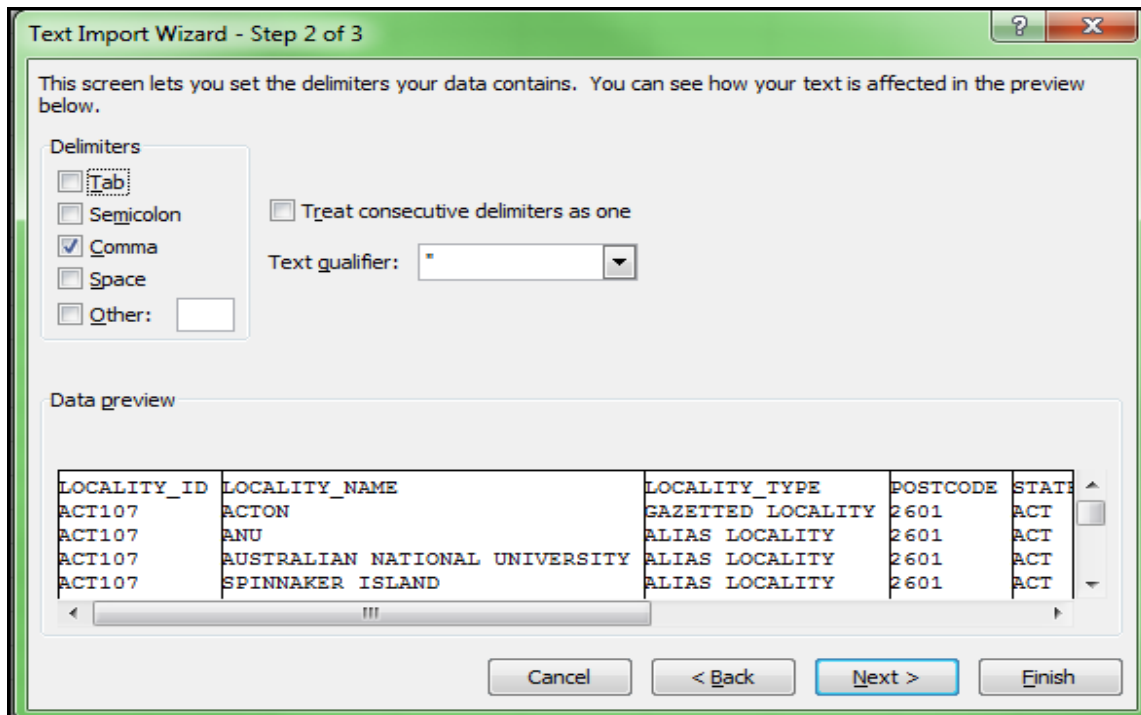
Open 'Text Import Wizard' by using the File | Open option. In the associated Open dialogue box choose the Text Files (*.prn, *.txt, *.csv) option.

This will start the first step of the 'Text Import Wizard'. Select the 'Delimited' option and click the 'Next' button.



Step 2

Select the 'Comma' option from the list of delimiters and make sure the double quote is selected as the 'Text qualifier'. Click the 'Next' button.



This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters

☐ Tab
☐ Semicolon
☒ Comma
☐ Space
☐ Other:

☐ Treat consecutive delimiters as one

Text qualifier: "

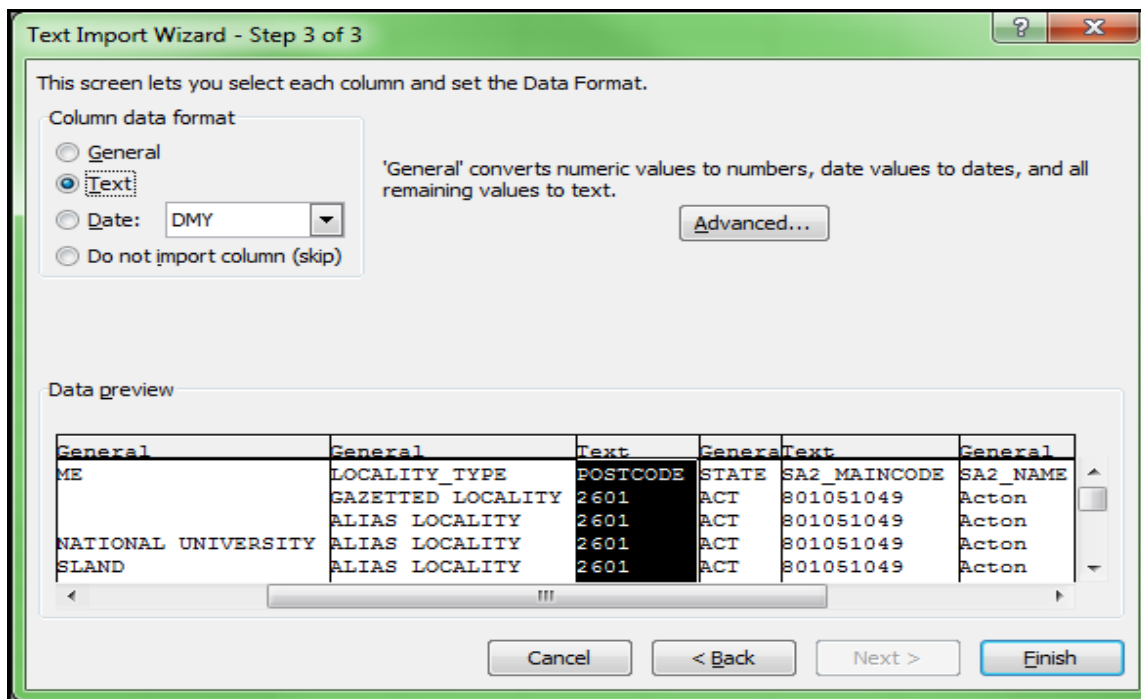
Data preview

LOCALITY_ID	LOCALITY_NAME	LOCALITY_TYPE	POSTCODE	STATE
ACT107	ACTON	GAZETTED LOCALITY	2601	ACT
ACT107	ANU	ALIAS LOCALITY	2601	ACT
ACT107	AUSTRALIAN NATIONAL UNIVERSITY	ALIAS LOCALITY	2601	ACT
ACT107	SPINNAKER ISLAND	ALIAS LOCALITY	2601	ACT

Buttons: Cancel, < Back, Next >, Finish

Step 3

Select the column holding the POSTCODE in the 'Data preview' frame and change the 'Column data format' from 'General' to 'Text'. Repeat this for the SA2_MAINCODE column. Click the 'Finish' button.



This screen lets you select each column and set the Data Format.

Column data format

☐ General
☒ Text
☐ Date: DMY
☐ Do not import column (skip)

'General' converts numeric values to numbers, date values to dates, and all remaining values to text.

Advanced...

Data preview

General	General	Text	General	Text	General
ME	LOCALITY_TYPE	POSTCODE	STATE	SA2_MAINCODE	SA2_NAME
	GAZETTED LOCALITY	2601	ACT	801051049	Acton
	ALIAS LOCALITY	2601	ACT	801051049	Acton
NATIONAL UNIVERSITY	ALIAS LOCALITY	2601	ACT	801051049	Acton
SLAND	ALIAS LOCALITY	2601	ACT	801051049	Acton

Buttons: Cancel, < Back, Next >, Finish

Once loaded the POSTCODE and SA2_MAINCODE should be stored as text. This is indicated by the green triangle in the top right hand corner of the storage cell.

	A	B	C	D	E	F	G	H	I
1	LOCALITY_ID	LOCALITY_NAME	LOCALITY_TYPE	POSTCODE	STATE	SA2_MAINCODE	SA2_NAME	RATIO	PERCENT
2	ACT107	ACTON	GAZETTED LOCALITY	2601	ACT	801051049	Acton	0.999998	99.99982
3	ACT107	ANU	ALIAS LOCALITY	2601	ACT	801051049	Acton	0.999998	99.99982
4	ACT107	AUSTRALIAN NATIONAL UNIVERSITY	ALIAS LOCALITY	2601	ACT	801051049	Acton	0.999998	99.99982
5	ACT107	SPINNAKER ISLAND	ALIAS LOCALITY	2601	ACT	801051049	Acton	0.999998	99.99982
6	ACT107	SPRINGBANK ISLAND	ALIAS LOCALITY	2601	ACT	801051049	Acton	0.999998	99.99982
7	ABS195	ACTON PARK	EXTRA LOCALITY	2601	ACT	801021028	Molonglo	1	100
8	ACT104	AINSLIE	GAZETTED LOCALITY	2602	ACT	801051050	Ainslie	1	100

Save the formatted Excel file to your computer.