1216.0.55.004

Australian Bureau of Statistics

Information Paper

# Converting Data to the Australian Statistical Geography Standard

## Australia

## 2012

# Information Paper

# Converting Data to the Australian Statistical Geography Standard

## Australia

## 2012

**Brian Pink**
**Australian Statistician**

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30AM (CANBERRA TIME) TUES 28 FEB 2012

ABS Catalogue No. 1216.0.55.004



Produced by the Australian Bureau of Statistics

## INQUIRIES

- For further information about these and related statistics, contact the National Information and Referral Service on 1300 135 070.

# CONTENTS

# PREFACE



The Australian Bureau of Statistics (ABS) has recently developed a new geography to replace the Australian Standard Geographical Classification (ASGC) that has been used for the collection and dissemination of geographically classified statistics since its introduction in 1984. The Australian Statistical Geography Standard (ASGS) was released in December 2010, and the first statistics based on the ASGS were released in August 2011.

The 2011 Census of Population and Housing will be the first Census to feature data released on the ASGS, and Census data will become available later in 2012.

The ABS has in place a number of strategies to assist clients in making the transition from the ASGC to the ASGS. For instance, data from the 2011 Census will be released in time series for 2001, 2006 and 2011 on both Statistical Local Areas (ASGC) and Statistical Areas Level 2 (ASGS) to enable users to maintain continuity. Other collections will release parallel datasets for a short time to allow statistical benchmarking.

A summary of the ASGS implementation schedule for other ABS publications is available on the web site <http://www.abs.gov.au/geography>.

Strategies for converting existing data onto the ASGS are detailed in this paper. Allocation tables, coding indexes and geographic correspondence tables will all be made available free of charge through the ABS website, whilst address coding of unit record data to the ASGS is also possible if address information is held with the data.

Specific information relating to the current versions of the ASGC and ASGS respectively can be found in the following ABS publications: *Australian Standard Geographical Classification (ASGC) July 2011* (cat. no. 1216.0), *Australian Statistical Geography Standard (ASGS): Volume 1 - Main Structure and Greater Capital City Statistical Areas, July 2011* (cat. no. 1270.0.55.001), *Australian Statistical Geography Standard (ASGS) Volume 2 - Indigenous Structure, July 2011* (cat. no. 1270.0.55.002) and *Australian Statistical Geography Standard (ASGS) Volume 3 - Non ABS Structures, July 2011* (cat. no. 1270.0.55.003). These publications are available from the ABS website.

Any enquiries regarding the ASGS can be made by emailing: <geography@abs.gov.au>.

Brian Pink
Australian Statistician

# ABBREVIATIONS

| | |
|---|---|
| **ABS** | Australian Bureau of Statistics |
| **ASGC** | Australian Standard Geographical Classification |
| **ASGS** | Australian Statistical Geography Standard |
| **CD** | collection district |
| **GCCSA** | Greater Capital City Statistical Area |
| **LGA** | local government area |
| **MB** | Mesh Block |
| **SA1** | Statistical Area Level 1 |
| **SA2** | Statistical Area Level 2 |
| **SA3** | Statistical Area Level 3 |
| **SA4** | Statistical Area Level 4 |
| **SD** | statistical division |
| **SLA** | statistical local area |
| **Vic.** | Victoria |

# INTRODUCTION



The Australian Standard Geographical Classification (ASGC) has provided the ABS statistical geography framework since its inception in 1984. A review of the ASGC in 1996 revealed some shortcomings with the application and ongoing maintenance of the ASGC, and these were reinforced in a further review in 2007. The main issue identified was with the provision of data over time as the areas within the ASGC were not stable, with annual updates largely driven by changes in Local Government Area (LGA) boundaries to which the ASGC had to conform. This requirement to conform to LGAs compromised the ABS' ability to maintain accurate time series datasets, and required them to be recast on to the latest boundaries, which was an ongoing process. Other issues identified included inconsistent area sizes and populations, and variable statistical quality in very small areas. In addition, in 2004 the ABS had commenced the design of a new very small area-level geography - the Mesh Block.

The outcome of the 2007 review resulted in the creation of a completely new statistical framework, the Australia Statistical Geography Standard (ASGS). The design of the ASGS addressed many of the issues previously identified with the ASGC. Primary among these was stability of design and a concerted effort to standardise the design of statistical regions at all levels, optimising each level of the new geography for consistent and comparable quality statistical outputs. Additionally, the ASGS incorporated the Mesh Block as its basic building block. The ASGS is designed to be stable over 5 years, with updates to the ASGS promulgated in the lead up to each Census of Population and Housing (the Census).

Units within the ASGS are internally consistent in population size, and based on functional and geographic criteria. Each level of the main structure of the ASGS is optimised for the release of key data collections - the Statistical Area Level 1 (SA1) is optimised for the release of Census data, the Statistical Area Level 2 (SA2) for the release of demographic Estimated Resident Population statistics, and the Statistical Area Level 4 (SA4) optimised for the release of Labour Force Survey data.

The ABS released the first data on the ASGS in August 2011. The 2011 Census, conducted on 9 August 2011, was the first Census to be collected on the ASGS, and Census data will become available in June 2012. As the ASGS will be the basis for the release of all geographically based statistics from the ABS, there is an opportunity for other organisations to modify their own statistical compilation methods to align with the ASGS. The advantage of doing so includes access to a wider variety of geographically comparable statistics that will inevitably lead to improved information for decision making and policy formulation.

The change to a new geographic standard has created a need for users to move historical and existing time series datasets onto the new geography to enable the ongoing comparison of data. There are a number of strategies that can be applied to aid this process, and the aim of this information paper is to look at the options that are available to users when converting data from other geographical classifications, including the ASGC, to the ASGS. This paper will examine options such as address coding and correspondences and will discuss the concepts behind these methods of converting data, as well as discussing some of the issues that the use of these methods can present to users. These options are not only available for ABS data collections, they can also be applied to data held by other agencies and organisations.

The options available to users to convert unit record data will depend largely on the information contained within those records. If these records contain full address information, or latitude and longitude, then it will be possible to geocode them to the finer levels of the ASGS such as Mesh Block or SA1. However, if only partial address information is available then coding indexes can be used to convert data, but this will only be possible at the higher levels of the ASGS such as SA2 and above.

If the collection data contains records that already include Mesh Block or SA1 codes, then allocation tables can be used to aggregate data to higher levels of the ASGS. It needs to be noted though, that an allocation table cannot be used to convert data from higher levels of the ASGS to the finer levels.

If no address information is retained with the unit records then geocoding and indexes will not be options for the conversion of data and correspondences will need to be utilised. There are many issues associated with using correspondences, most notably that attempting to convert data from large scale geographies to the finer levels of the ASGS will not work well and the data returned will not be reliable. The type of data being converted also needs to be considered as the weighting unit used can impact on how well data will be converted. All of these issues are discussed in detail.

## ADDRESS CODING

Address coding, a form of geocoding, is a process where a unit record is assigned to a latitude/longitude point on the ground, or "coded" to a small area, such as a Mesh Block, by using the address. Once this has been established, the unit record can then be assigned to a larger geographic area. The data can then be added to other unit records that are assigned to the same area, thus providing aggregate values for the new geographic area.

If latitudes and longitudes are available, then it is possible to code these units to any level of the ASGS. This is the most accurate method for the transition of existing data. This removes the need to apply a mathematical transformation, as data coded to a point can be aggregated accurately to any area, regardless of its geographic base. The overall effect is an accurate recoding of the original data.

If a complete street address is available - that is street number and name, Suburb or Locality, State or Territory (S/T) and postcode - it is also possible to code the units to any level of the ASGS using address coding software. However, if some unit records do not contain complete addresses, such as missing street numbers, misspelt street names or incorrect Suburb or Locality names, then achieving an accurate match to a point on the ground may not be possible. As a result accurate coding to the smaller levels of the ASGS, such as Mesh Block or SA1, may not be achieved. It should be noted though, that incomplete or inaccurate addresses may still code accurately to the higher levels of the ASGS such as SA2 and above. Other addresses that can prove problematic include lot numbers and Rural Mail Box and Post Office Box addresses. In these cases manual data repair may be required to achieve the desired results.

Rural addressing is not consistent across Australia, though this is in the process of being standardised, and as a result a collection unit address may not achieve a match that will enable coding to the smaller levels of the ASGS.

When considering converting data to the ASGS, address coding and geocoding are considered to be the most accurate method of converting data and these options should be explored prior to using any of the other options discussed in this paper.

At present the ABS does not provide an address geocoding service, as we do not own the intellectual property of the Geocoded National Address File (G-NAF) file which is owned by PSMA Australia. The ABS does however have an agreement with PSMA Australia to include our Mesh Block codes in the G-NAF file. This allows G-NAF users to relate addresses to the ABS geography.

There are several commercial organisations that offer geocoding services, based on G-NAF. More details on these companies can be found on the PSMA Australia website: <http://www.psma.com.au/wheretobuy/fullaccessresellers.html>.

## CODING INDEXES

Coding indexes enable data collected with a partial address to be linked to a geographic area. For ABS purposes, coding indexes are tables that list a geographic area against its most appropriate match in the ASGS. For example, a coding index will indicate a Locality and the SA2 it is most appropriately associated with based on the distribution of population within the locality. Table 1 shows an example of a Locality to SA2 coding index.

**1** EXAMPLE OF LOCALITY TO SA2 CODING INDEX

| *LOCALITY NAME* | *STATE* | *POSTCODE* | *SA2 MAINCODE 2011* | *SA2 NAME 2011* |
|---|---|---|---|---|
| IRYMPLE | VIC | 3498 | 215021395 | Irymple |
| NICHOLS POINT | VIC | 3501 | 215021395 | Irymple |
| NICHOLS PT | VIC | 3501 | 215021395 | Irymple |
| EAST MILDURA | VIC | 3500 | 215021397 | Mildura |
| MILDURA | VIC | 3500 | 215021397 | Mildura |
| MILDURA EAST | VIC | 3500 | 215021397 | Mildura |
| MILDURA SOUTH | VIC | 3500 | 215021397 | Mildura |
| MILDURA WEST | VIC | 3500 | 215021397 | Mildura |
| SOUTH MILDURA | VIC | 3500 | 215021397 | Mildura |
| WEST MILDURA | VIC | 3500 | 215021397 | Mildura |
| HATTAH | VIC | 3501 | 215021398 | Mildura Region |
| KOORLONG | VIC | 3501 | 215021398 | Mildura Region |
| LINDSAY POINT | VIC | 3496 | 215021398 | Mildura Region |
| LINDSAY PT | VIC | 3496 | 215021398 | Mildura Region |
| MERINEE | VIC | 3496 | 215021398 | Mildura Region |
| MERINGUR | VIC | 3496 | 215021398 | Mildura Region |
| MERRINEE | VIC | 3496 | 215021398 | Mildura Region |
| NEDS CORNER | VIC | 3496 | 215021398 | Mildura Region |
| WARGAN | VIC | 3496 | 215021398 | Mildura Region |
| WERRIMULL | VIC | 3496 | 215021398 | Mildura Region |
| KARADOC | VIC | 3496 | 215021399 | Red Cliffs |
| RED CLIFFS | VIC | 3496 | 215021399 | Red Cliffs |
| SUNNYCLIFFS | VIC | 3496 | 215021399 | Red Cliffs |

The coding index in Table 1 lists a series of Locality names, along with State and Australia Post postcode. For each Locality it then lists the associated SA2 Code and SA2 Name. 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 accurate match to the SA2 level of the ASGS can be obtained. In this instance the Australia Post postcode is important as Locality names are not necessarily unique, even within a State - the postcode aids in identifying the correct Locality.

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.

CODING INDEXES *continued*

During the development of the ASGS, SA2 boundaries were designed to reflect officially gazetted Suburb and Locality boundaries. As a result, if collection units contain this level of information, they can be accurately coded to the SA2 level of the ASGS. The ABS will supply this index free of charge, and it is available upon request by emailing <geography@abs.gov.au>.

Similarly, coding indexes have been created that enable users to match postcode data to the SA4 and Greater Capital City Statistical Area (GCCSA) boundaries, as well as 2006 Remoteness Areas and 2011 Statistical Divisions (SD).

There are issues that users need to be aware of when using postcode data as the sole source of deriving a match to the ASGS however. Firstly, there is no definitive set of postcode boundaries available, as Australia Post does not maintain a definitive set of boundaries as a matter of course. Various organisations have created their own boundaries, but these can differ quite significantly from one dataset to another. Another issue is that postcodes do not cover the whole of Australia. Postcodes are also, in general, larger than Suburbs and as a result cannot be coded to the more detailed levels of the ASGS. As a result, these coding indexes will only be released at the SA4 and GCCSA levels as the quality of these coding indexes at this level are deemed to be acceptable.

## ALLOCATION TABLES

Allocation tables, also known as hierarchy tables, describe a hierarchy of regions where the smaller regions fit precisely within the larger regions. The ASGS represents a structural hierarchy where areas at the smaller levels aggregate directly to those above them. An example of this is the relationship between SA1s and SA2s where an SA2 contains one or more whole SA1s. Allocation tables are usually supplied as text files.

Allocation tables only describe the relationship within a single structure and cannot be used to convert data from unrelated structures or data over time. It should also be noted that allocation tables cannot be used to disaggregate data from the higher levels of the ASGS.

The allocation tables for the ASGS were released with the *Australian Statistical Geography Standard (ASGS): Volume 1 - Main Structure and Greater Capital City Statistical Areas* (cat. no. 1270.0.55.001) and are available from the ABS website <http://www.abs.gov.au/geography>. Figure 1 below details the ASGS hierarchy.

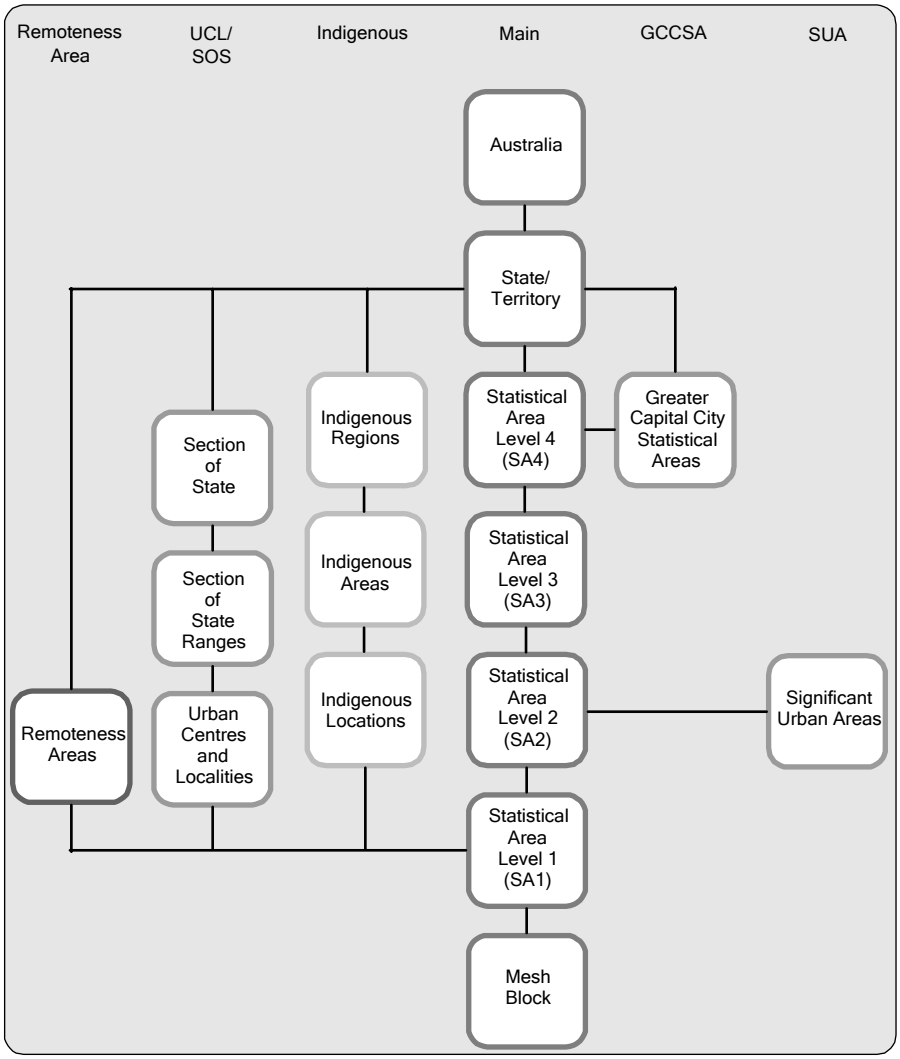


*Figure 1*: ASGS ABS STRUCTURES

ALLOCATION TABLES
*continued*

Table 2 illustrates an example of an allocation table, and details the ASGS main structure and how the smaller units of the ASGS are components of the larger ASGS units.

**2** EXAMPLE OF SA1 ALLOCATION TABLE

| *SA1 MAINCODE* | *SA2 MAINCODE* | *SA2 NAME* | *SA3 CODE* | *SA3 NAME* | *SA4 CODE* | *SA4 NAME* | *GCCSA CODE* | *GCCSA NAME* |
|---|---|---|---|---|---|---|---|---|
| 12602150130 | 126021501 | North Ryde - East Ryde | 12602 | Ryde - Hunters Hill | 126 | Sydney - Ryde | 1GSYD | Greater Sydney |
| 12602150131 | 126021501 | North Ryde - East Ryde | 12602 | Ryde - Hunters Hill | 126 | Sydney - Ryde | 1GSYD | Greater Sydney |
| 12602150132 | 126021501 | North Ryde - East Ryde | 12602 | Ryde - Hunters Hill | 126 | Sydney - Ryde | 1GSYD | Greater Sydney |
| 12602150201 | 126021502 | Ryde - Putney | 12602 | Ryde - Hunters Hill | 126 | Sydney - Ryde | 1GSYD | Greater Sydney |
| 12602150202 | 126021502 | Ryde - Putney | 12602 | Ryde - Hunters Hill | 126 | Sydney - Ryde | 1GSYD | Greater Sydney |
| 12602150203 | 126021502 | Ryde - Putney | 12602 | Ryde - Hunters Hill | 126 | Sydney - Ryde | 1GSYD | Greater Sydney |

## GEOGRAPHIC CORRESPONDENCES

Where other options for converting data are not available, a correspondence can be generated, enabling users to convert data to the ASGS. Correspondence tables detail a mathematical transformation that can be utilised to convert data from one geographic area onto another, unrelated, area. There are many issues that arise with the use of correspondences however, and users need to be aware of these issues and exercise caution when using them to convert data.

Correspondence tables specify the proportion of data for an area that should be donated to another area, effectively converting the geographic base of the data. The area that is donating is known as the "From" unit, and data is allocated to a "To" unit. Examples of the results from a generated correspondence are shown below.

**3** EXAMPLE OF SLA TO SA2 CORRESPONDENCE FILE

| *FROM UNIT: SLA CODE* | *FROM UNIT: SLA NAME* | *TO UNIT: SA2 CODE* | *TO UNIT: SA2 NAME* | *RATIO OF SLA WITHIN SA2* | *PERCENT OF SLA WITHIN SA2* |
|---|---|---|---|---|---|
| 105051100 | Botany Bay (C) | 117011320 | Banksmeadow | .0009346 | .09346 |
| 105051100 | Botany Bay (C) | 117011321 | Botany | .2219453 | 22.19453 |
| 105051100 | Botany Bay (C) | 117011322 | Mascot - Eastlakes | .491931 | 49.1931 |
| 105051100 | Botany Bay (C) | 117011323 | Pagewood - Hillsdale - Daceyville | .2851891 | 28.51891 |
| 105054800 | Leichhardt (A) | 120021387 | Balmain | .3055177 | 30.55177 |
| 105054800 | Leichhardt (A) | 120021388 | Leichhardt - Annandale | .4496677 | 44.96677 |
| 105054800 | Leichhardt (A) | 120021389 | Lilyfield - Rozelle | .2448146 | 24.48146 |

As can be seen from Table 3, the first two columns contain Statistical Local Area (SLA) codes and SLA names. In this correspondence SLAs represent the "From" units, or the units that data is being converted from. The SA2s represent the "To" units, or the units that data is being converted to. Ratio details the proportion of data that the "From" unit is donating to the respective "To" units.

In this example the SLA of Botany Bay is divided across four separate SA2s. 0.09% of the data for the SLA of Botany Bay is allocated to the SA2 of Banksmeadow, and 22.19% of the data for the SLA of Botany Bay is allocated to the SA2 of Botany. 49.19% is allocated to Mascot - Eastlakes and 28.52% to Pagewood - Hillsdale - Daceyville. The correspondence is then applied to a data value for each donating area, and the aggregate data for each "To" area calculated.

*Weighting of Correspondences*

When correspondences are created the ABS uses a weighting unit to assist with the allocation of "From" data to their respective "To" units. The unit that is used to weight the correspondence impacts on how the "From" units are apportioned, which in turn can have a major impact on the data values and hence the quality of the converted, or corresponded, data.

The unit used to weight a correspondence can vary, depending on the intended use and the nature of the data being converted. For example, correspondences can be weighted by area, by Mesh Block dwelling or population counts, or by particular population characteristics. The weighting unit is most effective when it is smaller than the geographic units being converted. Research and testing has also shown that the relationship between the weighting unit and the data being converted is critical. For instance, if agricultural or environmental data is being converted then using area as the weighting unit is quite effective as it assumes an even distribution of data, and these types of data tend to be uniform. However, using area as the weighting unit when converting other types of data, such as population-based data, can lead to poor results.

An example to consider is a case where there is a requirement to convert data from SLA to SA2 and one SLA encompasses two SA2s of equal size. One of the SA2s contains nursing homes whereas the other is made up of an industrial estate which contained no population on Census night. If an area weighted correspondence was used the population of the SLA would be evenly distributed between the SA2s. Given that only one of the SA2s contains population, distributing the population using this method will lead to incorrect and misleading results. However if a Census population weighted correspondence was used, the SA2 that contains the nursing homes would be allocated the entire population, and the SA2 containing the industrial estate would receive a zero allocation of population. This would result in an accurately distributed population and would reflect the true characteristics of the two SA2s.

A further example shows the comparison of actual and converted Deaths data using different SLA to SA2 correspondences. Figure 2 represents data converted by an area weighted correspondence whereas the data in Figure 3 has been weighted using Collection District (CD) counts of persons aged 65 and over from the 2006 Census.

Each point represents the actual data value plotted against its converted value for a given geographic area in the correspondence table. For example, in Figure 2 the highest point plotted for converted deaths is 292 whereas the actual deaths figure for the SA2 in question is 7. This indicates that the data transformation for that area was not accurate. A perfect correspondence would be represented by a plot showing a straight line rising at an angle representing a 1 to 1 ratio.

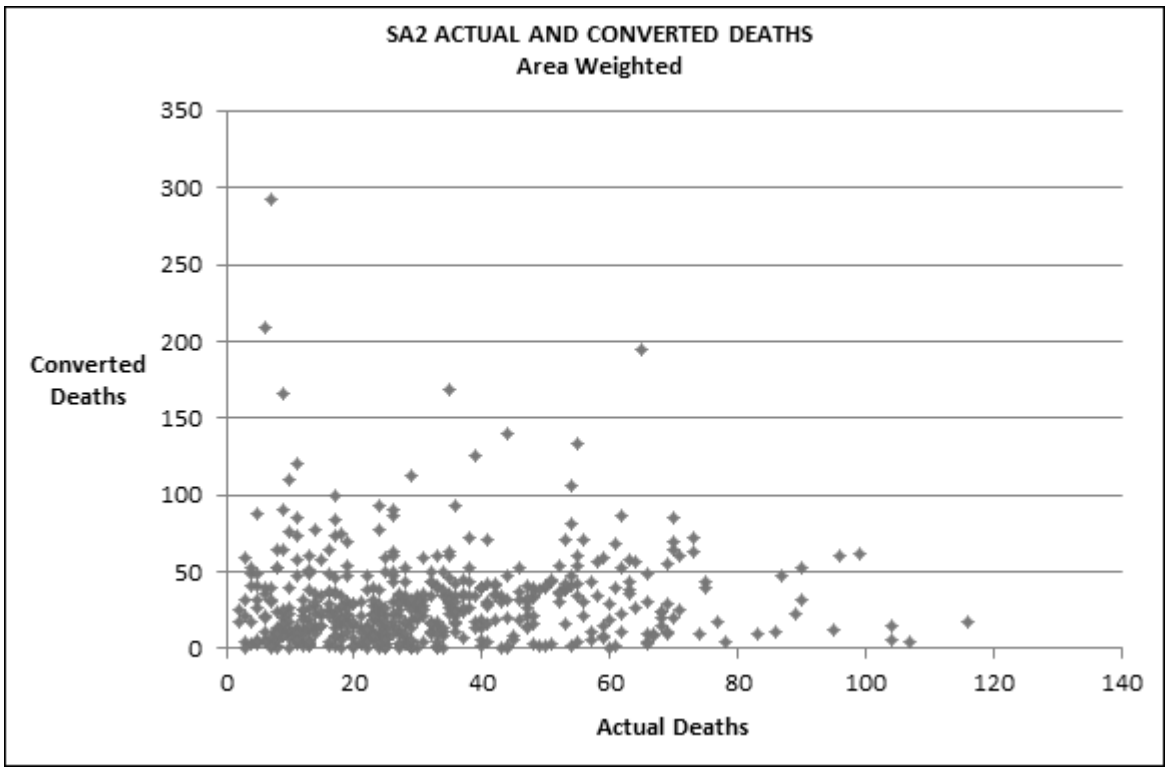


*Figure 2*: CORRESPONDED DEATHS DATA - Area weighted.

*Weighting of Correspondences continued*

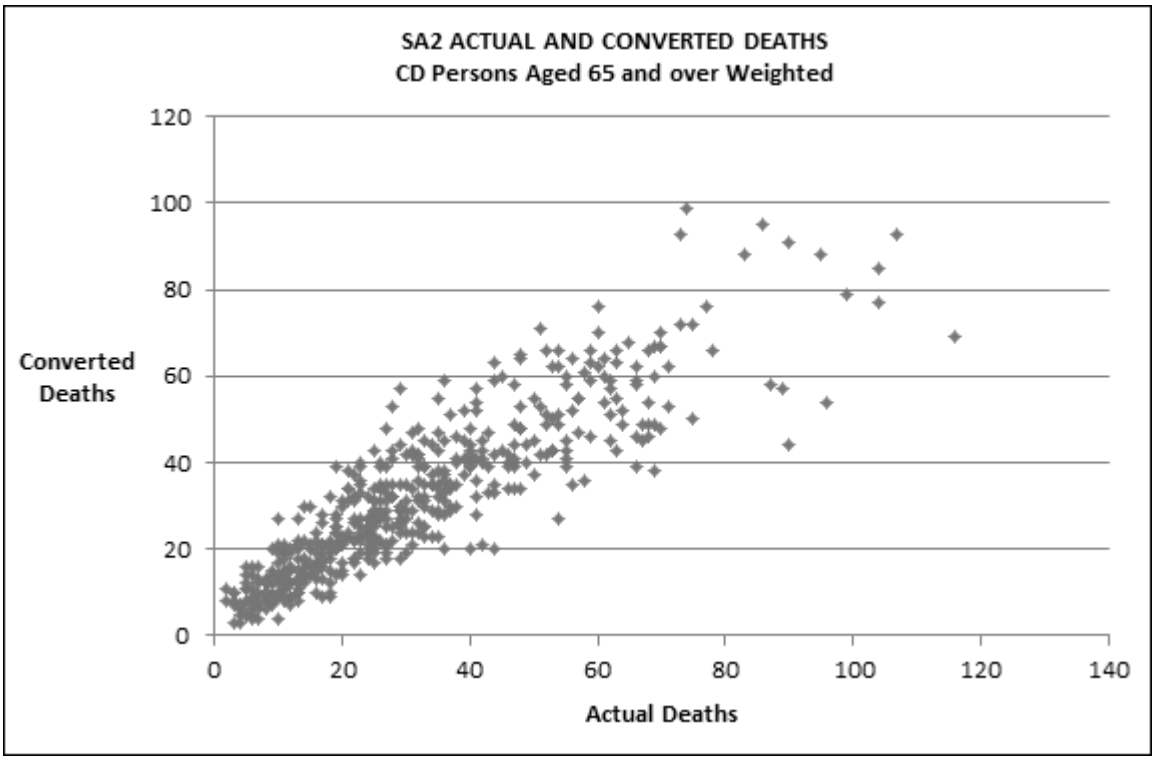


*Figure 3*: CORRESPONDED DEATHS DATA - Weighted by Persons aged 65 and over.

*Weighting of Correspondences continued*

It can be clearly seen in this instance that the area based correspondence has produced a markedly inferior result when compared to the population weighted data.

As can be seen in Figure 2, there are many points where the actual and converted counts are not similar at all, so the conclusion can be made that using an area weighted correspondence to convert this data is not suitable. However, in Figure 3, where the weighting unit is CD counts of persons aged 65 years and over, the actual and converted counts are more similar and show a better, though not perfect, result. The conclusion here is that this correspondence is converting data to a higher degree of accuracy. This is not surprising, given that the data is population-based, and is being converted using a population weighted correspondence.

Figures 2 and 3 showed examples of data being converted to differing degrees of accuracy. To contrast this is an example where data is being converted to a high level of accuracy. Figure 4 shows the results of geocoded address points being converted from SLA to Statistical Area Level 3 (SA3) where estimated dwellings at the Mesh Block level have been used as the weighting unit.

Figure 4 highlights two issues. Firstly, that converting smaller geographic units to larger units will generally result in more accurate data conversion than will be returned when converting areas of similar size, or when converting larger areas to smaller areas. In this example converting SLA level data to represent larger SA3 level data produces excellent results as SLAs are a smaller geographic unit than SA3s. The second issue demonstrated in Figure 4 is the importance of the weighting unit when converting data. The weighting unit used with this correspondence was estimated dwellings at the Mesh Block level. As dwelling estimates relate closely to the G-NAF, which is the data source, this correspondence was ideal for converting G-NAF counts from SLA to SA3.

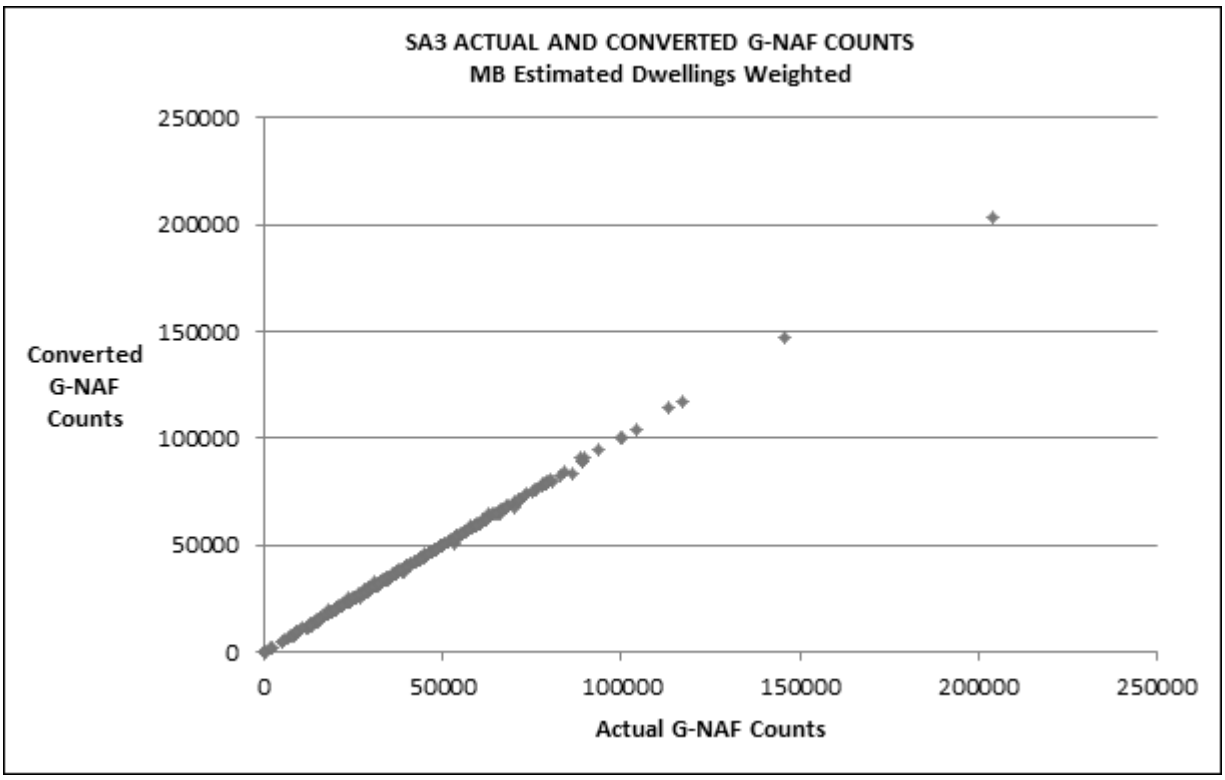


*Figure 4*: CORRESPONDED G-NAF COUNTS - SLA to SA3.

*Weighting of Correspondences continued*

Another issue that needs to be considered is that the same correspondence will convert different types of data to differing degrees of accuracy. This is highlighted in Figures 5 and 6.

Figure 5 details the results of converting G-NAF counts from SLA to SA2 where estimated dwellings at Mesh Block level is used as the weighting unit, whereas Figure 6 shows the results of converting deaths data from SLA to SA2 using the same weighting unit. In this instance the same correspondence is being used to convert different types of data, and as can be seen by the results, the G-NAF counts have been converted to a higher degree of accuracy than the Deaths data. The reason in this case is that the weighting unit being used, Mesh Block estimated dwellings, relates more closely to G-NAF data than Deaths data. This highlights the fact that the relationship between the data to be converted and the weighting unit used in the correspondence is critical to the accuracy of the output data.

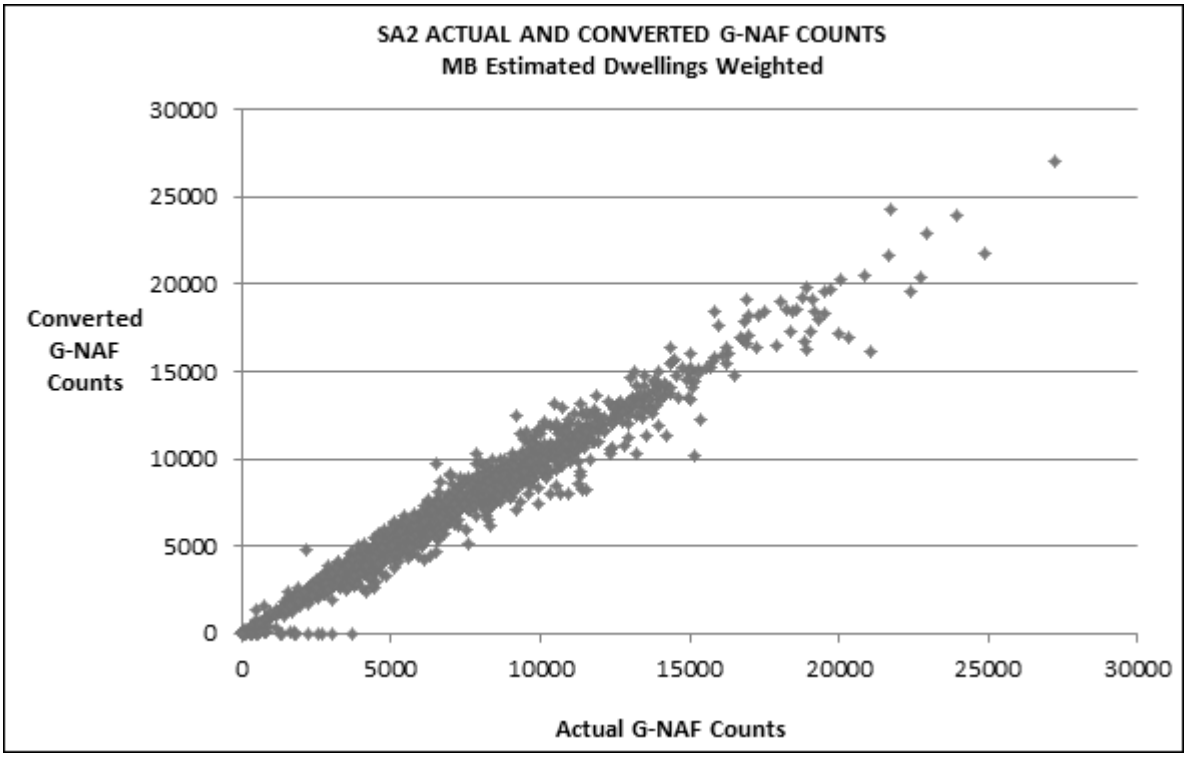


*Figure 5*: CORRESPONDED G-NAF DATA - Weighted by estimated dwelling counts.

*Weighting of Correspondences continued*

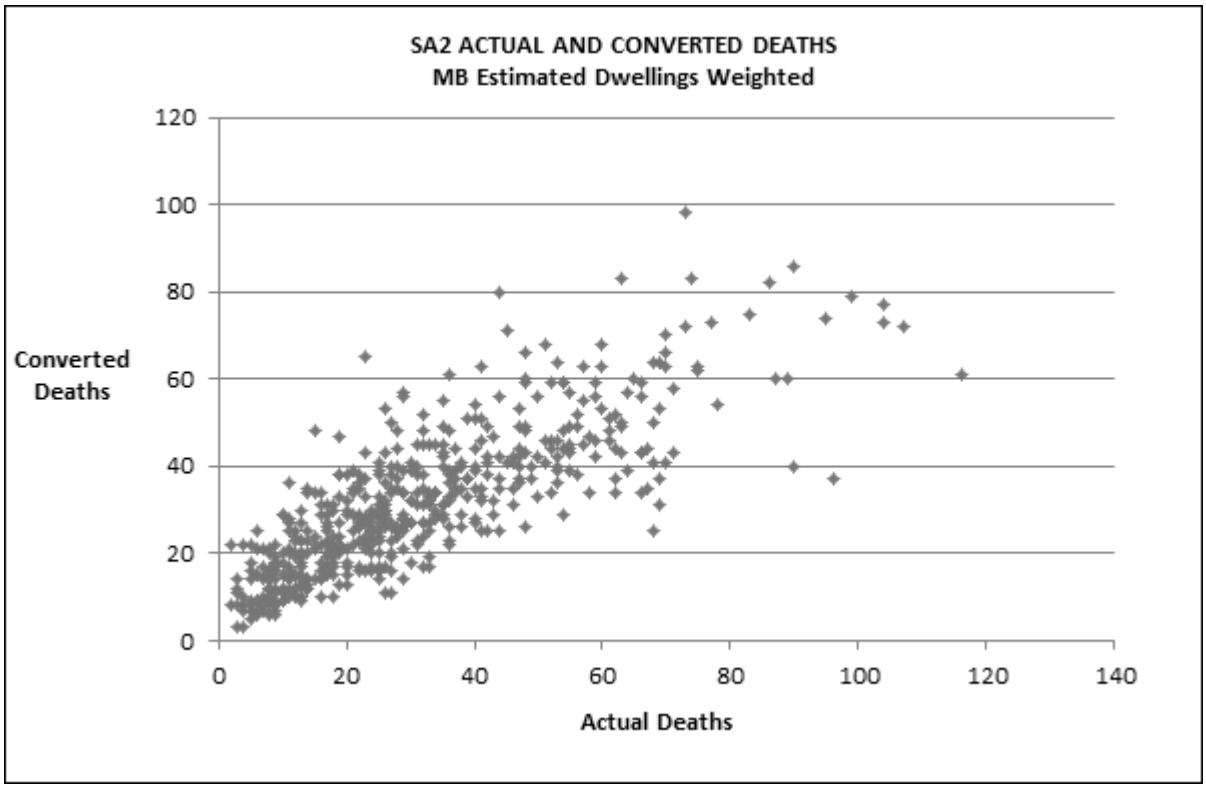


*Figure 6*: CORRESPONDED DEATHS DATA - Weighted by estimated dwelling counts.

## QUALITY INDICATOR

The ABS has determined that a quality indicator will be developed to advise users about the reliability of a correspondence. The quality indicator will be incorporated in geographic correspondences produced from June 2012 onwards. There will be two indicators reported for each correspondence. There will be an overall indicator of how well the ABS expects a correspondence to convert data, and an individual indicator will also be provided for each "To" record. This will allow users to see where the data is likely to be converted accurately, and conversely where the converted data is most likely to cause issues. However, the indicator should be used as a guide only and caution should always be used when examining the results of data that is converted using correspondences.

The quality indicator is calculated by creating a ratio of the area that is being donated, as a proportion of the whole "From" area. The area that is being donated is then proportioned against the whole of the "To" area to create another ratio. These two ratios are then multiplied together. This process is repeated for each "From" area and the resulting ratios are then aggregated to form the overall figure. From these ratios a textual description is then applied to each "To" area, and an overall description for the entire geographical correspondence is derived from averaging the ratios of every "To" area.

The quality indicator provides a considered measure of the quality of the correspondence in relation to the weighting unit. However, it cannot take into account the nature of the data that is to be converted. Data that is similarly structured to the weighting unit (e.g. population-based data when the weighting units are based on population distribution) will convert accurately, but that will not follow for data that has no relationship to the weighting unit. The textual descriptions that will be applied are as follows:

Good - The ABS expects that this correspondence or record will convert data overall to a high degree of accuracy and users can expect the converted data will reflect the actual characteristics of the geographic areas involved.

Acceptable - The ABS expects that this correspondence or record will convert data to a reasonable degree of accuracy, though caution needs to be applied as the quality of the converted data will vary and may differ in parts from the actual characteristics of the geographic areas involved.

Poor - The ABS expects that there is a high likelihood the correspondence or record will not convert data accurately and that the converted data should be used with caution and may not reflect the actual characteristics of many of the geographic areas involved.

There will be some cases where the ABS deems that a correspondence is not of an acceptable overall quality. In these cases they will be deemed as not being fit for purpose and will not be released. The ABS will release a series of correspondences from mid 2012 and these will be available, either for download or by email, free of charge from the ABS website. A list of correspondences that will be available can be found on the '*Correspondences*' chapter of the ABS Statistical Geography web site <http://www.abs.gov.au/geography>. The ABS will also produce custom correspondences, based on user requests, to convert data to the ASGS.

## CONCLUSIONS

The options available for converting data to the ASGS depends largely on the information stored with a user's unit record data. If a set of full addresses are available, or latitude and longitude coordinates are stored with the unit records, then geocoding the data to the new geography should be utilised as the most accurate method of converting the data.

If only partial address information is available, then coding indexes may be used to convert data to the ASGS. It needs to be noted that this will only be possible to the higher levels of the ASGS.

Allocation tables can be used if unit record data is contained at the finer levels of the ASGS. However, these tables can only be used to aggregate data to higher levels of the ASGS. They cannot be used to disaggregate data from higher levels of the ASGS.

If these options cannot be utilised, then a correspondence may be used to convert data to the ASGS. However, there are a number of issues that need to be investigated when using a correspondence. Correspondences work best when converting data from small regions onto large regions, and where the conversion is weighted by related criteria. Caution should always be taken when analysing the results of data that have been converted using correspondences, and the potential limitations of the data taken into account. A quality indicator will be incorporated in the future to ABS correspondences to assist users in determining whether they are fit for purpose.

Further information on converting data to the ASGS can be found by visiting the ABS Statistical Geography web site <http://www.abs.gov.au/geography>. Questions relating to the conversion of data to the ASGS, or in regards to the ASGS itself, can be emailed to <geography@abs.gov.au>.

## FOR MORE INFORMATION . . .

*INTERNET* **www.abs.gov.au** the ABS website is the best place for data from our publications and information about the ABS.

### INFORMATION AND REFERRAL SERVICE

Our consultants can help you access the full range of information published by the ABS that is available free of charge from our website. Information tailored to your needs can also be requested as a 'user pays' service. Specialists are on hand to help you with analytical or methodological advice.

*PHONE* 1300 135 070

*EMAIL* client.services@abs.gov.au

*FAX* 1300 135 211

*POST* Client Services, ABS, GPO Box 796, Sydney NSW 2001

## FREE ACCESS TO STATISTICS

All statistics on the ABS website can be downloaded free of charge.

*WEB ADDRESS* www.abs.gov.au