

Chapter 1 - SA3 BROWSER

The last release of the Australian Bureau of Statistics Census (2016), in conjunction with the establishment of the Primary Health Networks (PHN) (DOH, Primary Health Networks, 2015), has created the need to group the PHN geographies into peer groups with similar determinants of health needs.

However, the 31 PHN areas exhibit large variation in areal extent, population composition and population size. Therefore, government agencies reported the variation using smaller local areas (SA3).

As a result, AIHW entered in a collaborative agreement with CMCRC to explore possible ways of identifying comparable SA3s. Consequently, I developed the **SA3 Browser Tool (SA3B)**. **SA3B** is the first tool able to compare any two geographic areas and divide a set of geographies into groups with similar characteristics.

More precisely, each SA3 has been grouped into one of seven peer groups based on socioeconomic status and remoteness. The tool also includes different scenarios with larger and smaller number of peer groups.

This allows:

- PHN's SA3s to be compared to other PHN's SA3s within the same metropolitan, regional or rural socioeconomic peer group.
- SA3 to be compared with the average for their peer group.

This enables fairer comparisons of individual PHN catchments and provides a summary of the variation across Australia diverse metropolitan, regional and rural population by presenting aggregate results for each peer group.

In the next sections I give a brief description of the main functionalities available in the SA3B demo version.

1.1 - Software Architecture

The primary data source for the tool is the Census but additional data will be included to provide for more robust analysis.

The tool is composed of 4 modules, shown in Figure 18 below:

1. **SA3 Compare:** This module provides a dashboard that allows users to compare any two SA3s and explore the IRSD decile distribution by looking at the summary statistics indices and related curves introduced in the previous chapters of this dissertation (e.g. CI, BCFA, LI, DI, HI). It also provides the SA3-PHN concordance and SA3s' peer group for the solution of 5 and 7 clusters.
2. **SA3 Cluster Atlas:** In this module the user can explore the map of the SA3 peer grouping solution for 7 clusters of the IRSD variable across and within the 31 PHNs.
3. **SA3 Cluster Analysis:** This module contains the cluster analysis of the IRSD SA3 peer groups for 5, 7, 10 and 15 clusters. It also contains an ATLAS of the SA3 classification in terms of the socioeconomic decile category (LI), HI and population size. The dashboard presents this information according to the state of the SA3.
4. **SA3 Distance DB:** Lastly, this module contains all the pairwise distances (similarity measures) for each SA3. The Data Base allows the users to find the most similar or different geographic unit for each SA3.

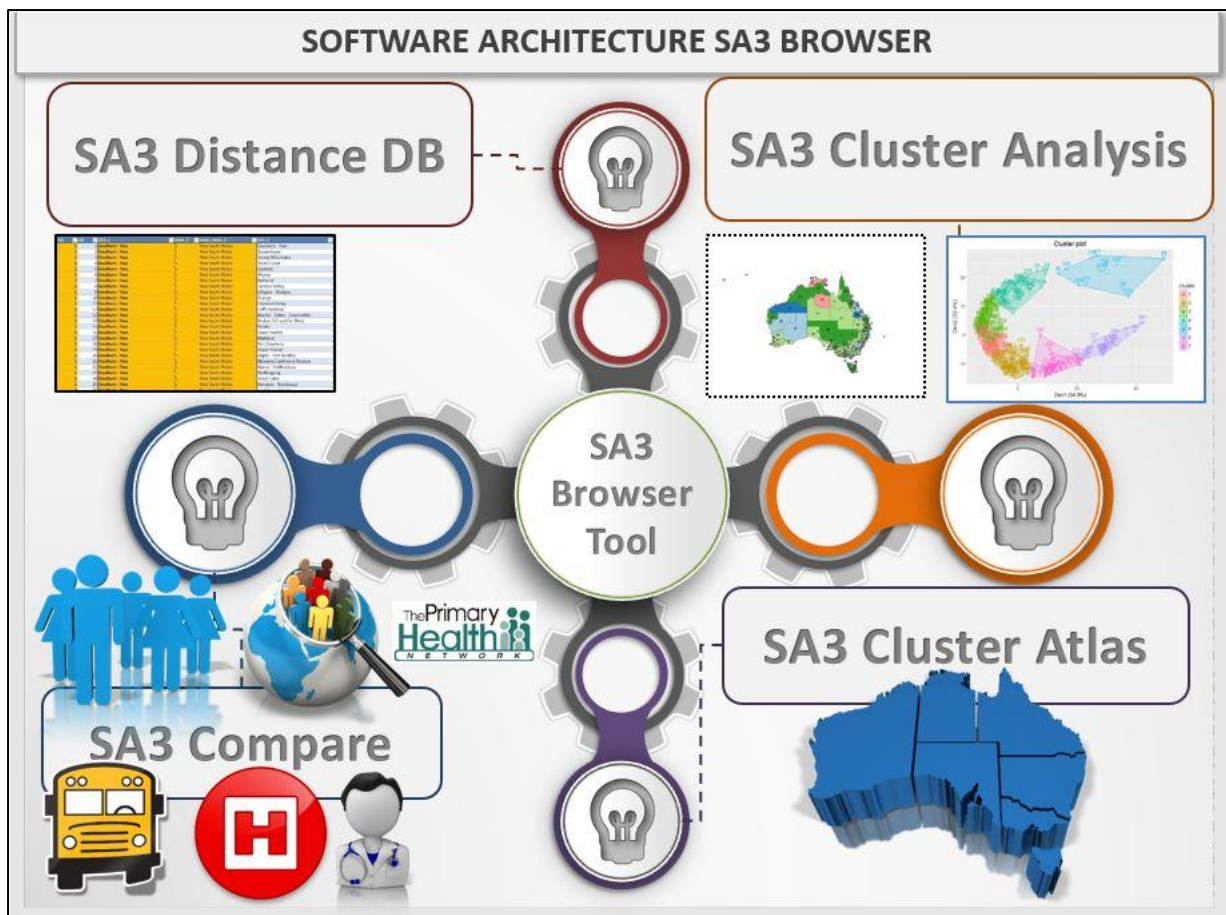


Figure 1 SA3 Browser Software Architecture

1.1.1 - SA3 Compare

When SA3 compare program starts (screen shot in Figure 19), a series of tabs are visible across the top of the screen. The interface is designed to guide the user in the socioeconomic analysis of the selected geographies. Let's just focus on the main functionalities of this module.

For instance, the two ribbon buttons allow the user to select the two SA3s to be compared. In this example, I selected the SA3s of Ku-ring-gai and Auburn. The state and SA3-PHN concordance grey windows show the SA3's state and PHN. In this case, the population living in Ku-ring-gai is located in the PHN of Northern Sydney, while 99.08 % of the population in Auburn is located in the PHN of Western Sydney and the remaining in South Western Sydney.

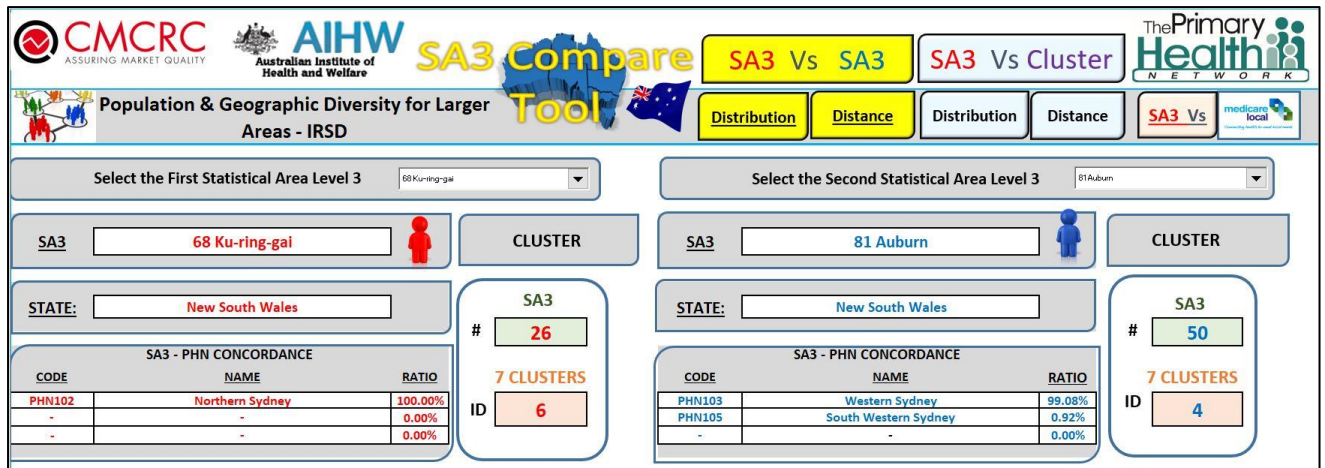


Figure 2 SA3 Compare Main Menu

Then, the numbers below the CLUSTERER grey area are the number of SA3s in the Peer group (green box) and the Peer group identification number (rose box), respectively. For example, there are 26 SA3s in the Peer Group number 6 and Ku-ring-gai is one of these SA3s. By clicking the distribution yellow button on the top of the Main Menu, the Compare module shows several information about the IRSD decile distributions, such as the IRSD score (PWAVGS, see Appendix), the socioeconomic disadvantage rank position, the decile and percentile. These information are indicated in the IRSD SCORE area of the screen shot illustrated in Figure 20. This screen shot shows also the Concentration, Divergence, Location and Homogeneity indices of the two SA3s.

In this example, we see that Ku-ring-gai is one of the least disadvantaged SA3 (rank 331) among the 333 units analysed. This information is further confirmed by looking at the LI (10) and HI score (91.78%) shown on the right hand side of Figure 20. On the other hand, the population of Auburn is less homogeneous (57.94 %) and more disadvantaged (LI=2) than Ku-ring-gai.

To have an idea of what does the distribution of Ku-ring-gai or Auburn look like, the user can explore the Estimated Residential population (ERP) of the IRSD decile distribution by clicking on the IRSD DECILE ERP button on the left hand side of Figure 20.

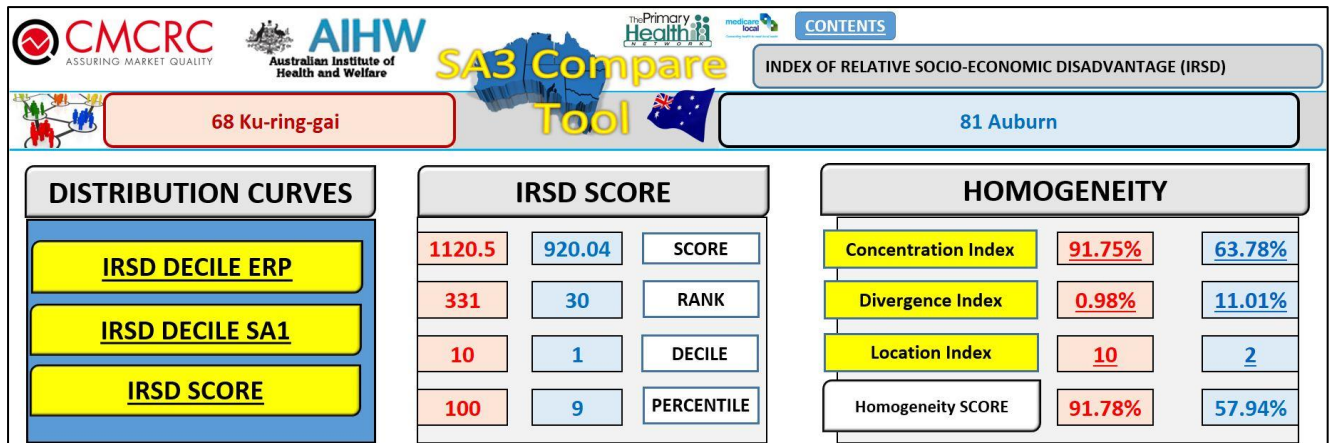


Figure 3 SA3 Compare Distribution Menu

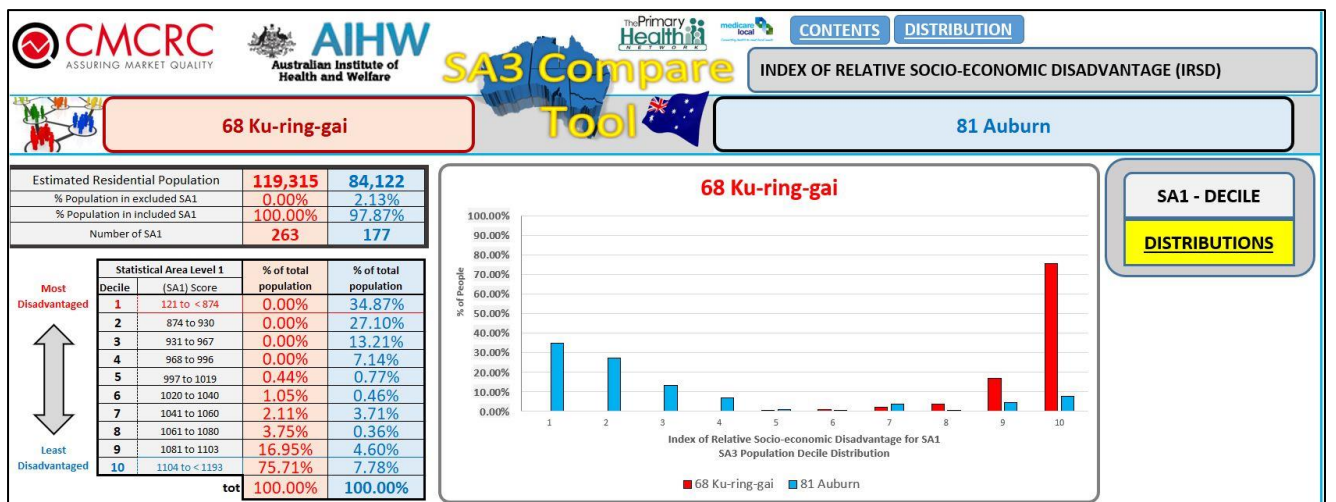


Figure 4 SA3 Compare IRSD DECILE ERP

The screen shot in Figure 21 shows the ERP IRSD decile distribution of the two SA3s. It also shows the population size and the number of SA1 within the SA3s. For instance, the SA3 of Auburn has a residential population of 84,122 people and 177 SA1s. We also notice that 2.31 % of the SA1s within Auburn have been excluded from the analysis for confidentiality reasons.

Finally, to see the CI, DI and Bin Concentration function of the LI, the user can simply click on the numbers displayed in the Homogeneity grey area shown in Figure 20.

For instance, the screen shots in Figure 22 and 23 show the CI and DI curves of Auburn. Naturally, there are other information in these charts that I do not discuss in this document, but the objective of this presentation is to give an idea of the capabilities of this module.

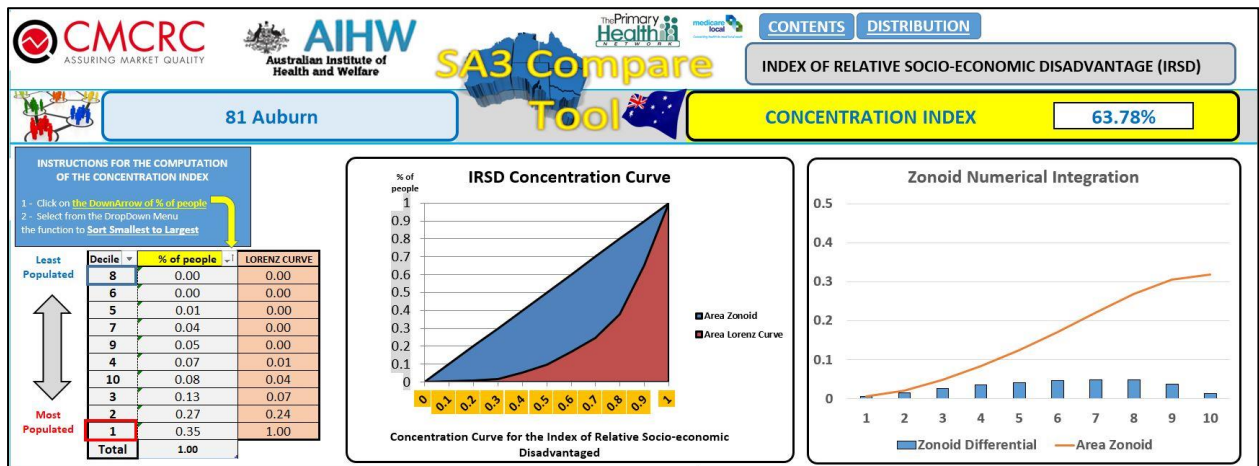


Figure 5 SA3 Compare IRSD Concentration Index Curve

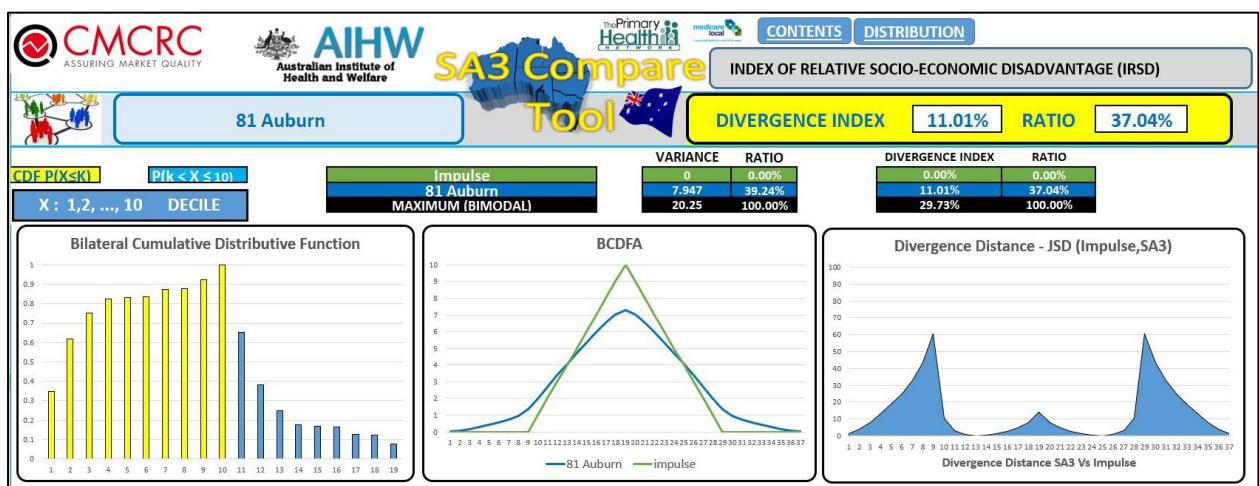


Figure 6 SA3 Compare IRSD Divergence Index Curve

Before I move on to the next module, it is worth showing two more functionalities of the SA3 compare module. The Primary Health Network (PHN) image and the distance tab in the main menu represented in Figure 20.

By clicking the PHN image on the top, the user can see the distribution of the selected SA3 and the corresponding PHN.

For example, the histograms in red and blue of Figure 24 show the IRSD decile distribution of Ku-ring-gai and the PHN of Northern Sydney, respectively.

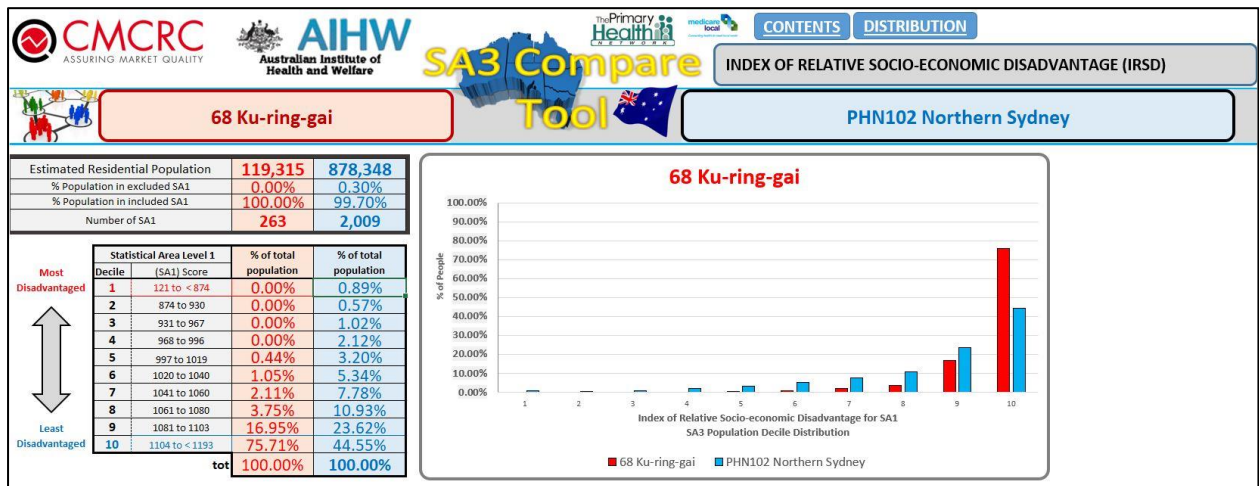


Figure 7 SA3 Compare PHN distribution

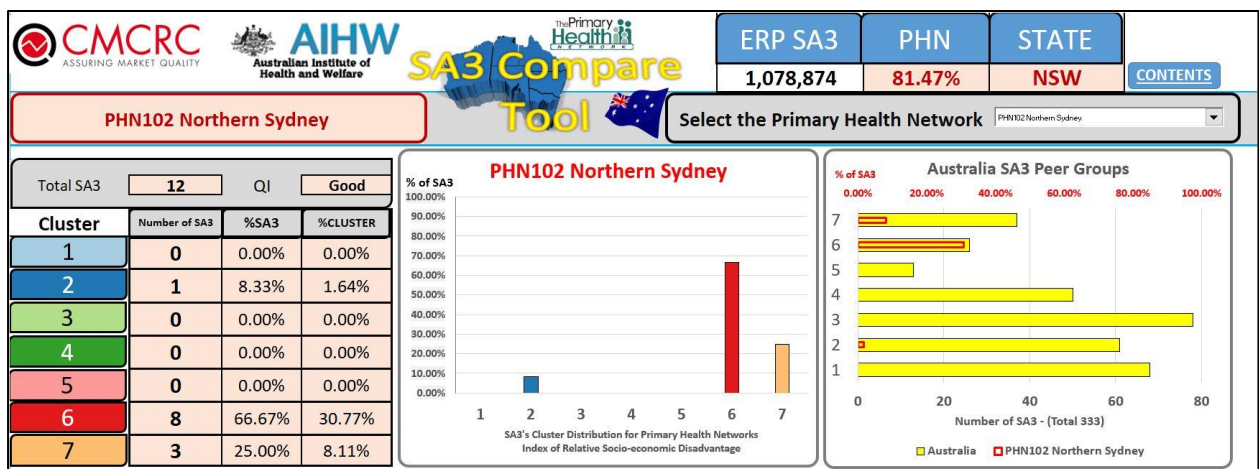


Figure 8 SA3 Compare PHN characteristics

Moreover, Figure 25 shows the number of SA3s within the PHN of Northern Sydney and how many of these SA3s are in the 7 peer groups. For example, this PHN has 12 SA3s and 8 of them are in the cluster number 5. This screen shot also show the total population of the SA3s within this PHN and the total correspondence. We see that the total population of the SA3s is 1,078,874

and 81.47 % live in the PHN of Northern Sydney. This information is further confirmed by the QI (Quality Index correspondence) shown on the left hand side of the screen shot (QI=good). Lastly, the distance tab in the main menu allows the users to compare the two SA3s and tells how different they are. For instance, the dissimilarity pie chart in figure 26 shows that the population living in Ku-ring-gai is not comparable to the population living in Auburn, since the dissimilarity between the two SA3s is 77.51%.

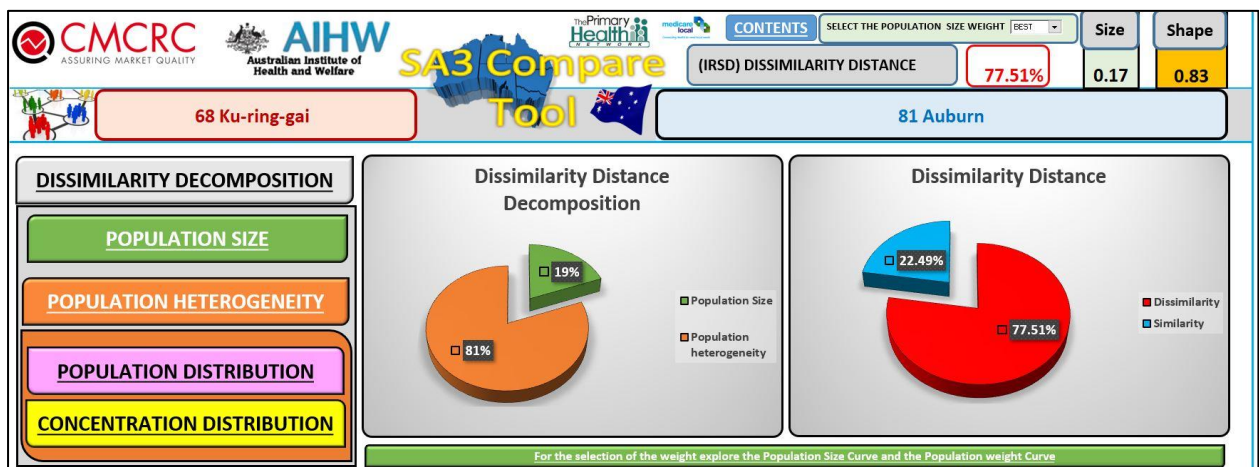


Figure 9 SA3 Compare Distance

1.1.2 - SA3 Cluster Atlas

This module shows the map of the PHNs within each state. It also shows the map of the SA3s' peer group within each PHN. For instance, Figure 27 shows the PHN within the state of New South Wales. Figure 28 shows the SA3s within the PHN of Northern Sydney and their corresponding peer group.

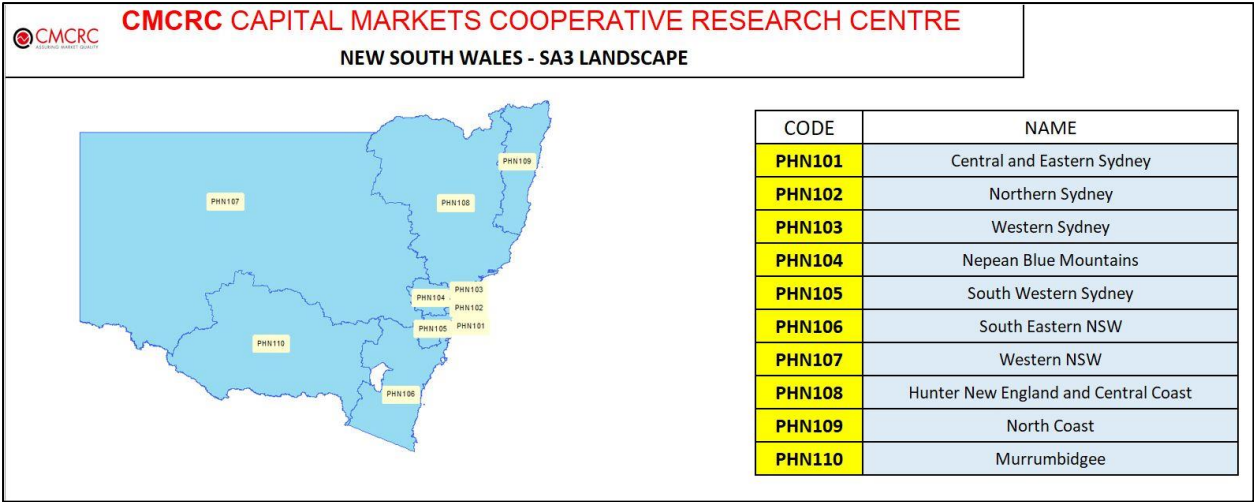


Figure 10 PHN Atlas mapping

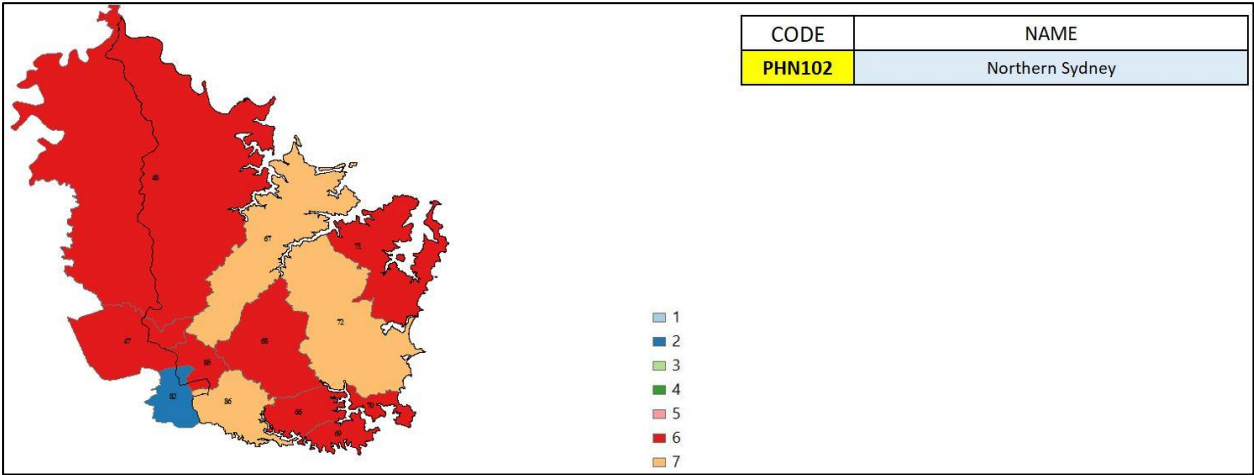


Figure 11 SA3 Atlas - PHN mapping

1.1.3 - SA3 Cluster Analysis

The approach taken in this module was to perform cluster analysis based on a metric that embeds the notion of similarity in terms of population size, population age distribution, distribution of remoteness using the ARIA+ classification (ABS, Remoteness classification, 2006) and distribution of socio-economic status (IRSD).

In order to evaluate the degree of dissimilarity between any two SA3s, I designed a metric composed of three terms:

1. **Population size:** This term considers the size of the histogram population curve. The goal is to capture the property that the SA3s in the same cluster should have roughly the same number of residents.
2. **Population Heterogeneity:** The second term considers the socio-economic characteristics of the distribution curve.
3. **Concentration Location:** Finally, another important feature is where the concentration of the population is located in the distribution.

The clustering algorithm used was a k-medoids algorithm called the Partitioning Around Medoids (PAM) algorithm. The number of clusters, k , was determined by maximizing the average silhouette width, a measure of how well the data has been clustered. I split SA3s into 5,7,10 and 15 clusters. As several clustering solutions have been provided, it is advised that the users consider each option and decide which, if any, sets of results conforms to their needs.

For instance, the snapshot in figure 29 shows the shape and the silhouette width of the clustering solutions for 5 and 7 peer groups.

CLUSTER ANALYSIS – PEER GROUPS

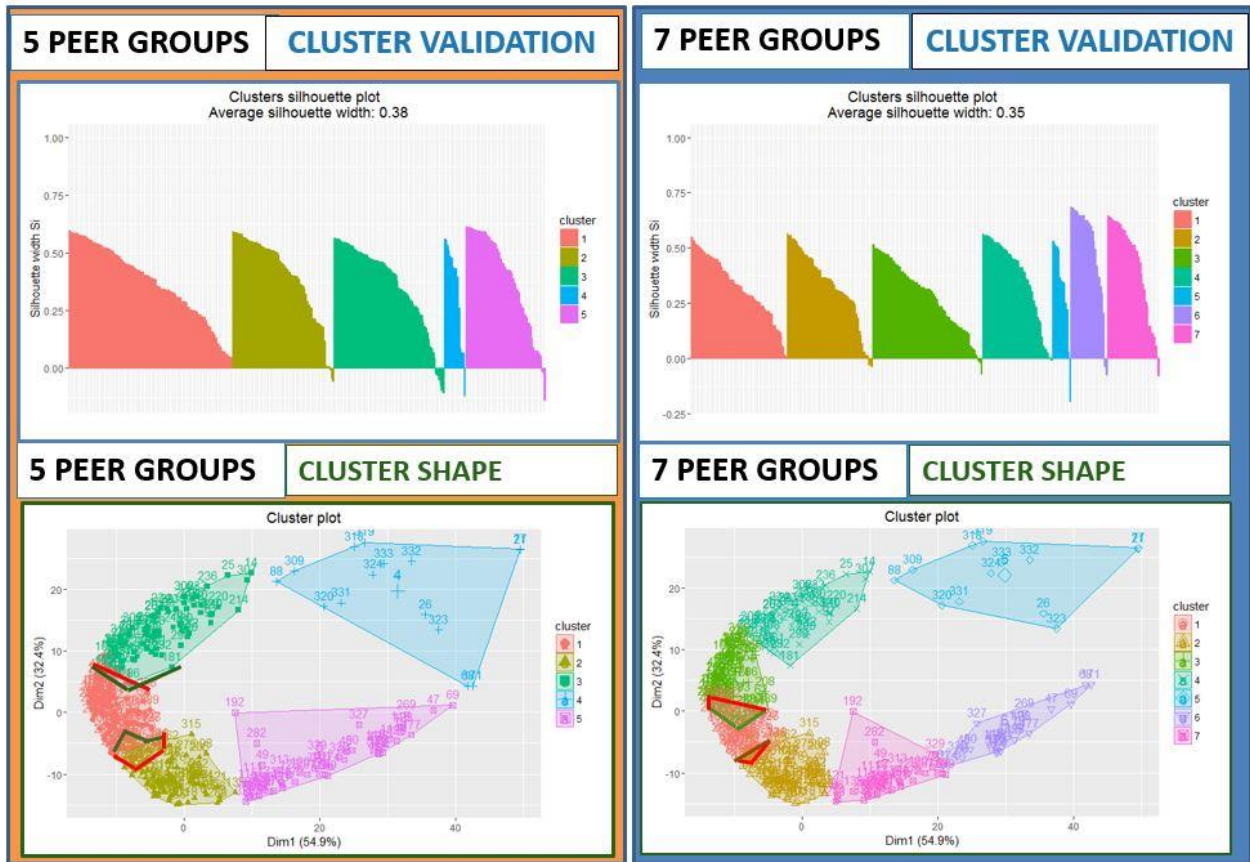


Figure 12 SA3 Cluster Analysis

1.1.4 - SA3 Distance DB

This module contains an easy to use spreadsheet for the analysis of the dissimilarity between any two SA3s. The spreadsheet contains 13 columns illustrated in the snapshot of Figure 30. The orange and blue fields represent the attributes of the first and second SA3, while the last column contains the dissimilarity value. A dissimilarity value of zero clearly indicates that the two SA3s are identical. On the other hand, a value of 1 means maximum dissimilarity.

An easy way to start using this table is the comparison of a selected SA3 to the other SA3s in the table. The Data Filter option in Excel allows the user to select the SA3 to compare and the criteria for sorting the records in the table.

In the example of Figure 30, I filtered the first column of the table by selecting the SA3 of Ku-ring-gai and sorted the dissimilarity measure (Distance) in ascending order. Once the filtering and sorting criteria are selected, the user can easily identify the most similar SA3 for the filtered SA3 (Ku-ring-gai) and compare the population size, HI and LI.

For instance, the second row in the table shows that the most similar SA3 is Baulkham Hills located in the state of New South Wales. The population of this SA3 is 144,084 (pop_2) and the HI is 0.859 or 86% (Hom_2). Lastly, the socio-economic classification or LI is 10 (X0*_2). Clearly, the user can apply other criteria to filter and sort the table.

SA3_1	state_1	state_name_1	SA3_2	state_2	state_name_2	pop_1	pop_2	Hom_1	Hom_2	x0*_1	x0*_2	Distance
Ku-ring-gai	1	New South Wales	Ku-ring-gai	1	New South Wales	119315	119315	0.933430224	0.933430224	10	10	0
Ku-ring-gai	1	New South Wales	Baulkham Hills	1	New South Wales	119315	144084	0.933430224	0.859756209	10	10	0.1023462
Ku-ring-gai	1	New South Wales	Kenmore - Brookfield - Moggill	3	Queensland	119315	46829	0.933430224	0.944764295	10	10	0.112479615
Ku-ring-gai	1	New South Wales	North Sydney - Mosman	1	New South Wales	119315	97707	0.933430224	0.889542937	10	10	0.113077769
Ku-ring-gai	1	New South Wales	Cottesloe - Claremont	5	Western Australia	119315	71477	0.933430224	0.823055857	10	10	0.147447092
Ku-ring-gai	1	New South Wales	Gungahlin	8	Australian Capital Territory	119315	57051	0.933430224	0.795042241	10	10	0.221547552
Ku-ring-gai	1	New South Wales	Brisbane Inner - West	3	Queensland	119315	58695	0.933430224	0.808388637	10	9	0.228355648
Ku-ring-gai	1	New South Wales	Boroondara	2	Victoria	119315	170553	0.933430224	0.779874353	10	9	0.22836227
Ku-ring-gai	1	New South Wales	Bayside	2	Victoria	119315	98368	0.933430224	0.741303129	10	10	0.229570683
Ku-ring-gai	1	New South Wales	Hills District	3	Queensland	119315	85712	0.933430224	0.750638283	10	9	0.249928809
Ku-ring-gai	1	New South Wales	Manly	1	New South Wales	119315	44232	0.933430224	0.790541499	10	9	0.260889561

Figure 1 SA3 Distance DB