8 Statistical analysis

Introduction

Two sets of analyses have been undertaken to illustrate the extent of association between areas with low socioeconomic status and poor health. Correlation coefficients have been produced to indicate interdependence between the measures of socioeconomic status, health status and use of health services. Cluster analysis has been undertaken to indicate the extent to which areas display significantly similar characteristics from among the chosen measures of socioeconomic status, health status and use of health services.

Inequalities in health have traditionally been indicated by an approximation to social class, frequently based on a categorisation of occupations. The other major indicators traditionally used have included income, education, ethnicity and employment status (which allows for the inclusion of unemployed people and those not in the labour force). The measures of socioeconomic status included in this analysis include income, education, occupation, labour force status and Aboriginality.

Correlation analysis

Description

Correlation is the degree to which one variable is statistically associated with another. The correlation coefficient is a measure of the strength of this association. When high values for one variable are matched by high values for the other (or when low values are matched by low values), then they are positively correlated. Where the interdependence is inverse (ie. high values for one are matched by low values for the other), the two variables are negatively correlated.

Methods

The Pearson product-moment correlation (r) has been used in this analysis to indicate the degree of correlation between pairs of variables. Pearson correlation coefficients range from +1 (complete positive correlation) through 0 (complete lack of correlation) to –1 (complete negative correlation). As a general rule, correlations of plus or minus 0.5 or above are considered to be of meaningful statistical significance. Correlations of plus or minus 0.71 or above are of substantial statistical significance, because this higher value represents at least 50 per cent shared variation (r^2 greater than or equal to 0.5).

Correlation coefficients were calculated by comparing the value (expressed as a percentage, or as a standardised ratio) for each variable in each SLA with the value of each of the other variables. Correlation coefficients are generally referred to as being, for example, 'a correlation of low income families with the *paired* variable of hospital admissions of females'. However, to promote ease of reading where many correlation coefficients are quoted in the text, the word 'paired' has been omitted. For similar reasons the symbol used to indicate a correlation coefficient (r) has been omitted.

Two measures of socioeconomic status included in the analysis in this section have not been mapped. They are families receiving an income of $52,000 or more per annum and people in occupations classified as 'managers and administrators' and 'professionals'. These two measures were included as they indicate high socioeconomic status, in contrast to most other measures, which were chosen because they indicate low socioeconomic status.

The results of the correlation analysis, which was undertaken separately for **Melbourne** and the rest of the State, are shown in the following tables: coefficients of from 0.5 to 0.7 and from 0.71 to 1 (both positive and negative) are highlighted in the tables, and are referred to in the individual map commentaries, as appropriate. The analysis was not undertaken for **Geelong**, as there were too few SLAs for the analysis to be valid.

The different years for which the data is available, and changes in boundaries between those periods, have meant that there are three correlation matrices for **Melbourne** and two for the rest of Victoria. The first matrix (**Table 8.1**) comprises data mapped on 1996 boundaries; that is, the 1996 Census data in Chapter 3, the income support data in Chapter 4, the general medical practitioner and immunisation data from Chapter 6 and the population per GP data from Chapter 7. The second matrix (**Table 8.2**) comprises data mapped on a common set of the boundaries in existence over the period from 1992 to 1995 (boundaries in the 1991 to 1994 editions of the ASGC); that is, the 1996 Census data in Chapter 3 (re-cast, at the Collection District level, to approximate the 1994 SLA boundaries and then aggregated to equate with the 1991 boundaries) and the data from Chapter 5, other than the variables for fair/poor health and the Physical Component Summary. These two variable are shown in the third matrix for **Melbourne** (**Table 8.3**), with other data mapped on 1994 boundaries, namely the re-calculated Census data and the hospital admission data from Chapter 6.

There is one less matrix for non-metropolitan SLAs, as the deaths data was re-coded to 1994 boundaries (see note in Chapter 5), removing the necessity to have a matrix for the 1991 to 1994 boundaries. The first matrix (**Table 8.4**) comprises data mapped on 1996 boundaries; that is, the 1996 Census data in Chapter 3, the income support data in Chapter 4, the general medical practitioner and immunisation data from Chapter 6 and the population per GP data from Chapter 7. The second matrix (**Table 8.5**) comprises data mapped on 1994 boundaries; this includes the 1996 Census data in Chapter 3 (re-cast, at the Collection District level, to approximate the 1994 SLA boundaries), the data in Chapter 5 and the hospital admission data from Chapter 6.

When discussing the results of the correlation analysis in the text, mention is often made of 'the indicators of socioeconomic disadvantage'. This reference is to variables such as those for single parent families, the unemployed, the Indigenous population and housing authority rented dwellings. References to 'high socioeconomic status' are about the variables for high income families, female labour force participation and managers and administrators and professionals.

The associations discussed in the text are, in general, limited to associations between the variable under discussion and the indicators of socioeconomic status from Chapter 3. This approach is largely a response to the limited space available for comment. The extent of any association with the other variables analysed can be ascertained from an examination of the correlation matrices (**Table 8.1** and **8.2**).

Results

Melbourne

There were correlations of significance at the SLA level between the measures of socioeconomic disadvantage and a number of the health status variables. The strongest of these were with the variables for people reporting their health as fair or poor (as opposed to those reporting their health as being excellent, very good, or good); the PCS (the Physical Component Summary, a measure of physical health); years of potential life lost; and premature death from, in particular, circulatory system diseases. Similarly, strong associations were also evident in the correlation analysis with the health service use variables of GP services to males and females; and of admissions for neurotic, personality and other mental disorders and ischaemic heart disease, and admissions to a public hospital.

Non-metropolitan areas

SLAs in the non-metropolitan areas range in size from an estimated 9 square kilometres in Queenscliffe to 21,756 in Mildura [Part B]. They also range from sparsely populated rural and remote areas to large country towns. Despite these wide variations, the correlation analysis has been undertaken.

It is clear from an examination of the correlation tables that there are fewer correlations of significance at the SLA level in the non-metropolitan areas of Victoria than was the case in **Melbourne**. This is, in part, a result of the number of areas with relatively small numbers of cases (population, deaths, hospital admissions, etc.) which reduces the strength of the analysis.

However a number of variables are highly correlated with each other; these are the variables for low income families, unemployed people, single parent families, dwellings rented from the State housing authority and dwellings without a motor vehicle.

Various sub-sets of these are correlated with measures of health status and use of health services. The strongest correlations with the measures of socioeconomic disadvantage were with the variables for people reporting their health as fair or poor, and the PCS. There was a consistent, although weaker, pattern in the correlations between socioeconomic disadvantage and the variables for premature deaths of males and females, from respiratory and circulatory system diseases and years of potential life lost.

Table 8.1: Correlation matrix for SLAs in Melbourne, 1996 boundaries

Refer to file: ch8 correlation matrices

Table 8.2: Correlation matrix for SLAs in Melbourne, 1991 boundaries

Refer to file: ch8 correlation matrices

Table 8.3: Correlation matrix for SLAs in Melbourne, 1994 boundaries

Refer to file: ch8 correlation matrices

Table 8.3: Correlation matrix for SLAs in Melbourne, 1994 boundaries ...*cont*

Refer to file: ch8 correlation matrices

Table 8.4: Correlation matrix for SLAs in non-metropolitan areas of Victoria, 1996 boundaries

Refer to file: ch8 correlation matrices

This page intentionally left blank

Table 8.5: Correlation matrix for SLAs in non-metropolitan areas of Victoria , 1994 boundaries

Refer to file: ch8 correlation matrices

Table 8.5: Correlation matrix for SLAs in non-metropolitan areas of Victoria , 1994 boundaries ...*cont*

Refer to file: ch8 correlation matrices

Cluster analysis

Description

The intention of the cluster analysis is to produce summary measures of socioeconomic status, health status and health service use at the SLA level. It is useful to have this information, as the SLA is an important administrative and planning unit. However, the production of clusters at this level is problematic, as SLAs are often large, heterogeneous areas, and their average values sometimes disguise a wide range of sub-area variation in the values of the population characteristics under analysis.

It should also be noted that cluster analysis is an exploratory technique and, as with all such techniques, the real test of a solution is whether it makes any sense. Decisions as to the variables to be used, or the number of clusters in a solution, all impact on the final result.

The results of the cluster analysis, therefore, represent indicative groupings of areas with broadly similar characteristics among the variables analysed in each set. They will be a useful tool for some purposes: on other occasions, however, the individual variables on which they are based may also be relevant.

Methods

Cluster analysis (using the squared Euclidean measure) was undertaken by the Ward's method. This (hierarchic) clustering method seeks to partition a set of objects (eg. postcodes or, in this case, SLAs) into a set of non-overlapping groups so as to maximise some external criterion of 'goodness of clustering', typically the extent to which the within-cluster inter-object similarities are maximised and the between-cluster similarities minimised.

In cluster analysis, 10 records (ie. SLAs) per variable is considered desirable, with an absolute minimum of five. Had all the datasets been used in the analysis there would have been many fewer than this. A variety of techniques was used to attempt to overcome this problem, including applying a factor analysis or undertaking an experimental fit of the full data set, and using the results to reduce the number of variables included in the final analysis.

Table 8.6 lists the variables used in the analysis. The analysis was undertaken separately for **Melbourne** and the rest of the State. The datasets used in the cluster analysis (based on boundaries in existence from 1991 to 1997) were aggregated to a common set of boundaries (1994). Where the areas differ from the 1994 boundaries, the variations are noted in the text.

Table 8.6: Variables used in cluster analysis

Socioeconomic status	**Utilisation of health services**
% single parent families	***Hospital admissions*** (Standardised Admission Ratio)
% low income families	to public acute hospitals
% unskilled or semi-skilled workers	to private acute & private psychiatric hospitals
% unemployed	to public acute & private hospitals, admissions
% female labour force participation	total
People who left school at age 15 or earlier,	of males
or who did not attend school (Standardised Ratio)	of females
% Aboriginal & Torres Strait Islander people	for infectious diseases
% Housing authority rented dwellings	for all cancers
% Dwellings with no vehicles	for lung cancer
Health status	for breast cancer for women aged 40 years or more
Self-reported health status	for psychoses
Physical Component Summary score [SF-36]	for neuroses
Disability and handicap status (Standardised Ratio)	for circulatory system diseases
with a disability	for ischaemic heart diseases
with a handicap	for respiratory system diseases
Deaths (Standardised Death Ratio)	for respiratory system diseases in 0 to 4 year old children
Infant deaths	for bronchitis, emphysema & asthma
of males aged 15-64 years, from all causes	from accidents, poisonings and violence
of females aged 15-64 years, from all causes	for all surgical procedures
of persons aged 15-64 years	for all surgical procedures as same day admission
from cancer	for a tonsillectomy and/or adenoidectomy
from circulatory system diseases	for a myringotomy in children aged 0-9 years
from respiratory system diseases	for a Caesarean section in women aged 15-44 years
from accidents, poisonings & violence	for an hysterectomy in women aged 30 years and over
of persons aged 15-24 years	for an hip replacement
from accidents, poisonings & violence	for a lens insertion in people aged 50 years or more
Years of potential life lost as a result of deaths at ages 15-64 years	for an endoscopy
Total Fertility Rate	***General medical practitioner services*** (Standardised Ratio)
	for males
	for females
	Children fully immunised at 12 months

Results

Socioeconomic clusters in Melbourne

Variables considered for inclusion were those listed in **Table 8.6** under the heading *Socioeconomic status*. The ABS Index of Relative Socio-Economic Disadvantage (IRSD) was also used in the analysis, as an independent check on the solution.

Although a number of other variables were available for analysis, previous experience (Glover, 1996) has shown that the inclusion of variables regarding non-English speaking background is not beneficial to this analysis. The congregation of persons of the same ethnic group does not necessarily indicate a pocket of disadvantage. Although on average we may expect these variables to also show higher levels in disadvantaged areas, their inclusion in the cluster analyses does not assist in the search for viable and sensible solutions.

The variables relating to people born in predominantly non-English speaking countries (and their proficiency in English) were

Table 8.7: Composition of SLA clusters in Melbourne

SLA	Socioeconomic status	Health status	Health service utilisation	Social health[1]
Altona (C)	Low	Medium	Low	Low
Berwick (C)	Low	Good	Low	Medium
Box Hill (C)	High	Good	Low	Medium
Brighton (C)	High	Good	Medium	High
Broadmeadows (C)	Low	Medium	High	Low
Brunswick (C)	Low	Medium	High	Low
Bulla (S)	High	Good	Low	Medium
Camberwell (C)	High	Good	Medium	High
Caulfield (C)	High	Good	Medium	Medium
Chelsea (C)	Low	Medium	Low	Medium
Coburg (C)	Low	Medium	High	Low
Collingwood (C)	Medium	Poor	Low	Very low
Cranbourne (S)	Low	Good	High	Medium
Croydon (C)	High	Good	High	Medium
Dandenong (C)	Low	Medium	High	Low
Diamond Valley (S)	High	Good	Low	Medium
Doncaster and Templestowe (C)	High	Good	Low	High
Eltham (S)	High	Good	Low	High
Essendon (C)	Medium	Poor	Low	Very low
Fitzroy (C)	Medium	Poor	Low	Very low
Flinders (S)	Low	Medium	Medium	Medium
Footscray (C)	Low	Poor	Low	Low
Frankston (C)	Low	Medium	Low	Medium
Hastings (S)	Low	Medium	Medium	Medium
Hawthorn (C)	High	Good	Low	High
Healesville (S)	Low	Good	Low	Medium
Heidelberg (C)	Medium	Good	Low	Medium
Keilor (C)	Low	Medium	High	Low
Kew (C)	High	Good	Low	High
Knox (C)	High	Good	Low	Medium
Lillydale (S)	High	Good	Low	Medium
Malvern (C)	High	Good	Medium	High
Melbourne (C)	Medium	Poor	High	Very low
Melton (S)	Low	Good	Low	Medium
Moorabbin (C)	High	Good	Low	Medium
Mordialloc (C)	High	Good	Medium	Medium
Mornington (S)	High	Good	Medium	Medium
Northcote (C)	Low	Medium	High	Low
Nunawading (C)	High	Good	Low	High
Oakleigh (C)	Low	Medium	Low	Low
Pakenham (S)	High	Good	Low	Medium
Port Melbourne (C)	Medium	Medium	High	Very low
Prahran (C)	Medium	Good	Medium	Very low
Preston (C)	Low	Medium	High	Low
Richmond (C)	Medium	Poor	Low	Very low

Table 8.7: Composition of SLA clusters in Melbourne ... *cont*

SLA	Socioeconomic status	Health status	Health service utilisation	Social health[1]
Ringwood (C)	High	Good	Low	Medium
Sandringham (C)	High	Good	Medium	High
Sherbrooke (S)	High	Good	Medium	Medium
South Melbourne (C)	Medium	Poor	Medium	Very low
Springvale (C)	Low	Medium	Low	Low
St Kilda (C)	Medium	Poor	Low	Very low
Sunshine (C)	Low	Medium	High	Low
Upper Yarra (S) (Part A)	Low	Medium	Low	Medium
Waverley (C)	High	Good	Low	High
Werribee (C)	Low	Good	High	Medium
Whittlesea (C)	Low	Medium	Low	Low
Williamstown (C)	Medium	Medium	High	Medium

[1] **Social health 'clusters were produced by a joint analysis of the socioeconomic status and health status variable**

accordingly dropped from the analysis, leaving nine variables for inclusion. Problems of scale can affect the analysis as more common data items will dominate the solution. To avoid these problems, the variables were standardised and the resultant z scores were entered into the analysis.

There are 57 SLAs in **Melbourne** (the SLAs of Melbourne-Inner and -Remainder were analysed as one). These 57 records are theoretically sufficient to carry out a cluster analysis with nine input variables. A cluster analysis was performed on the available data, and the solution examined before attempting more complicated techniques to find a solution. This analysis provided a relatively clean three cluster solution which did not discriminate well between the Medium and Low socioeconomic clusters, although it lined up quite well against the IRSD.

The 57 records also provided enough information for an exploratory factor analysis, since this analysis has the same data requirements as the previous model. A factor analysis was attempted using principal components extraction and varimax rotation, and a reasonably sensible two factor solution was produced by this analysis.

A second factor analysis was run using maximum likelihood extraction and oblimin rotation, which resulted in the same two factor solution as above.

The main factor drivers of the two factor solution (ie. excluding low income families and Aboriginal people) were entered into a cluster analysis. This analysis indicated a three cluster solution, which was similar to the above but slightly worse quality. Again some of the lower SLAs for the IRSD grouped into the Medium socioeconomic status cluster.

The factor drivers of the first factor of the factor analysis solution (early school leavers, unemployed people, female labour force participation, unskilled and semi-skilled workers, low income families and Aboriginal people) were entered into a cluster analysis. This produced a two cluster solution, which was relatively uninformative.

None of the alternative analyses produced a superior solution to the original analysis including all input variables, the original solution was therefore accepted (see **Table 8.7** and **Map 8.1**).

The ABS Index of Relative Socio-Economic Disadvantage (IRSD) was also available for the specified SLAs, but was withheld from the analysis and used as an independent check on the solution. It was found that of the 23 SLAs with the lowest IRSD scores in **Melbourne**, 15 (65.2 per cent) were classified to the Low socioeconomic status group in this analysis; and that 20 of the 23 (87.0 per cent) with the highest scores for the IRSD were classified to the High socioeconomic status group.

After completion of the analysis for the **Melbourne**, the SLAs of **Geelong** were allocated to the clusters generated in **Melbourne** using the quick cluster command in SPSS. This procedure allocates the SLAs based on the minimum euclidean distance from each cluster centre. It therefore does not interfere with the formation of clusters in the capital city statistical division, but can be said to be on the same basis.

This resulted in Newtown being grouped into the High socioeconomic status cluster, and the SLAs of Bellarine Inner and Corio Inner being grouped into the Low socioeconomic status cluster (**Table 8.8** and **Map 8.1**).

The IRSD was again used as an independent check on the solution. It was found that, of the bottom two SLAs for **Geelong** as classified by the IRSD, one was classified to the Low socioeconomic status group in this analysis. Further, the top SLA under the IRSD, was classified to the High health status group.

Health status clusters in Melbourne

The data variables available for this analysis were the variables of premature death, disability and handicap status, the Total Fertility Rate and the two synthetically predicted estimates from the 1995 National Health Survey (the Physical Component Summary score and the measure of fair/poor health).

Thus there were 57 records to analyse 15 variables. Clearly this was not quite enough data. However, a cluster analysis of all the above variables was tried to see if it gave a sensible solution despite the lack of data. This produced a clean three cluster solution of high quality, which was accepted without further investigation (**Table 8.7** and **Map 8.2**).

Note that the Poor Status group did have higher status than the Good Status group for three variables (Total Fertility Rate, disability and deaths of people aged from 15 to 24 years from the combined causes of accidents, poisonings and violence). These results are understandable, in that females in socioeconomically disadvantaged areas have higher Total Fertility Rates; and that disability rates are higher in both socioeconomically disadvantaged areas and areas with high proportions of boarding houses and sheltered and other forms of specialist accommodation.
A check with the IRSD found that, of the bottom eight SLAs for **Melbourne** (as classified by the IRSD), four (50.0 per cent) were classified to the Poor health status group in this analysis. Further, of the top 30 SLAs under the IRSD, 27 (90.0 per cent) were classified to the Good health status group.

After completion of the analysis for **Melbourne** the SLAs in the major urban centre of **Geelong** were allocated to the clusters generated in **Melbourne** as discussed above under *Socioeconomic clusters in Melbourne*.

This resulted in two of the SLAs in **Geelong** (Bellarine Inner and Newton) being grouped into the Good health status cluster, and the SLAs of Geelong, Geelong West and Corio Inner being grouped into the Medium health status cluster (**Table 8.8** and **Map 8.2**).

The IRSD was again used as an independent check on the solution. It was found that of the bottom 3 SLAs for **Geelong**, as classified by the IRSD, all were classified to the Medium health status group in this analysis. Further, of the top two SLAs under the IRSD, both were classified to the Good health status group.

Health service utilisation clusters in Melbourne

Problems of scale can affect the analysis as more common data items will dominate the solution. To avoid these problems the variables were standardised and the resultant z scores were entered into the analysis. Thus there were 26 variables to analyse 57 records. Clearly this was not enough data. Alternative strategies were tried in an attempt to produce a useful solution.

A cluster analysis of all the above variables was tried to see if it gave a sensible solution despite the lack of data. This produced a four cluster solution of reasonable quality, although there was a suspicion that some of the SLA groupings to clusters were counter intuitive and may be explained by the lack of data to support the analysis.

An exploratory factor analysis was run on the data using Maximum Likelihood extraction and oblique (oblimin) rotation. The analysis produced a five factor solution. A further exploratory factor analysis was run on the data using Principal Component extraction and orthogonal (varimax) rotation. The analysis produced a five factor solution very similar to the previous solution. It should be noted that there was not enough data to sustain a factor analysis either.

The drivers of the first factor (admissions to a private hospital, admissions for a surgical procedure, same day admissions for a surgical procedure, and admissions for a myringotomy, endoscopy, lens insertion and hip replacement) were entered into a cluster analysis. The solution contained two clusters, which was not considered informative.

The drivers of the oblique factor solution were selected for entry into a cluster analysis. The first two drivers of each factor (admissions to public and private hospitals, admissions of males, admissions for cancer, neurotic, personality and other mental disorders, circulatory system diseases, ischaemic heart disease, respiratory system diseases, respiratory system diseases of children aged 0 to 4 years, and same day admissions for a surgical procedure) were entered into the analysis. This analysis again produced a three cluster solution of poor quality.

The drivers of the orthogonal factor solution were selected for entry into a cluster analysis. The first two drivers of each factor (admissions to a private hospital, admissions of males, admissions for cancer, breast cancer, neurotic, personality and other mental disorders, circulatory system diseases, ischaemic heart disease, respiratory system diseases of children aged 0 to 4 years, bronchitis, emphysema and asthma and same day admissions for a surgical procedure) were entered into the analysis.

This analysis produced a three cluster solution of acceptable quality. The solution does not discriminate particularly well between Medium and High health service use clusters, but it is the best solution found.

The analysis used ten variables to analyse 57 records, so the solution is supported by the data (see **Table 8.7** and **Map 8.3**).

A check with the IRSD showed that, of the bottom 14 SLAs for **Melbourne** as classified by the IRSD, eight (57.1 per cent) was classified to the High health service use group in this analysis. Further, of the top 31 SLAs under the IRSD, 18 (58.1 per cent) were classified to the Low health service use group.

After completion of the analysis for **Melbourne**, the SLAs in **Geelong** were allocated to the clusters generated in **Melbourne** as discussed above under *Socioeconomic clusters in Melbourne*.

This resulted in the **Geelong** SLA of Corio Inner being grouped into the High health service use cluster, with the SLAs of Geelong, Geelong West and Bellarine Inner being grouped into the Low health service use cluster (**Table 8.8** and **Map 8.3**).

A check with the IRSD showed that, of the bottom three SLAs for **Geelong** as classified by the IRSD, two (66.7 per cent) were classified to the High health service use group in this analysis. However, the top SLA under the IRSD was Newtown, which was not classified to the Low health service use cluster. Greater Geelong had the second lowest IRSD score, but was classified to the Low health service use cluster.

Social health clusters in Melbourne

The cluster analysis technique has also been applied to a combination of the socioeconomic status and health status data sets. The results of the cluster analysis for the combination of these data sets may be useful as a summary indicator of the 'social health' status of the population of each grouping of SLAs.

Data considered for inclusion were the variables in the final models for SLAs in **Melbourne**, used to examine socioeconomic status and health status.

There were 57 SLAs in **Melbourne** for this analysis (the same number as was available for the health status analysis). Clearly this was not enough data. A cluster analysis of all the above variables was tried to see if it gave a sensible solution despite the lack of data. This produced a clean four cluster solution of good quality, which was accepted without further investigation. The SLAs in each cluster are listed in **Table 8.7** and shown in **Map 8.4**. Note that the Low social health status group had a higher ranking than the High social health status group for disability and the Total Fertility Rate.

It was also found that, of the bottom nine SLAs for **Melbourne** as classified by the IRSD, four (44.4 per cent) were classified to the Very low social health status group in this analysis. Further, of the top 10 SLAs under the IRSD, 8 (80.0 per cent) were classified to the High social health status group.

After completion of the analysis for **Melbourne**, the SLAs in **Geelong** were allocated to the clusters generated in **Melbourne** as discussed above under *Socioeconomic clusters in Melbourne*.

This analysis produced four groupings, with the SLA of Newtown classified to the High social health status cluster and the SLA of Geelong West being classified to the Very low social health status cluster (**Table 8.8** and **Map 8.4**).

The IRSD was also available for the specified SLAs, and was used as an independent check on the solution. It was found that the bottom three SLAs for **Geelong** as classified by the IRSD were classified to the Medium social health status group in this analysis. Further, the top two SLAs under the IRSD were both classified to the High social health status group.

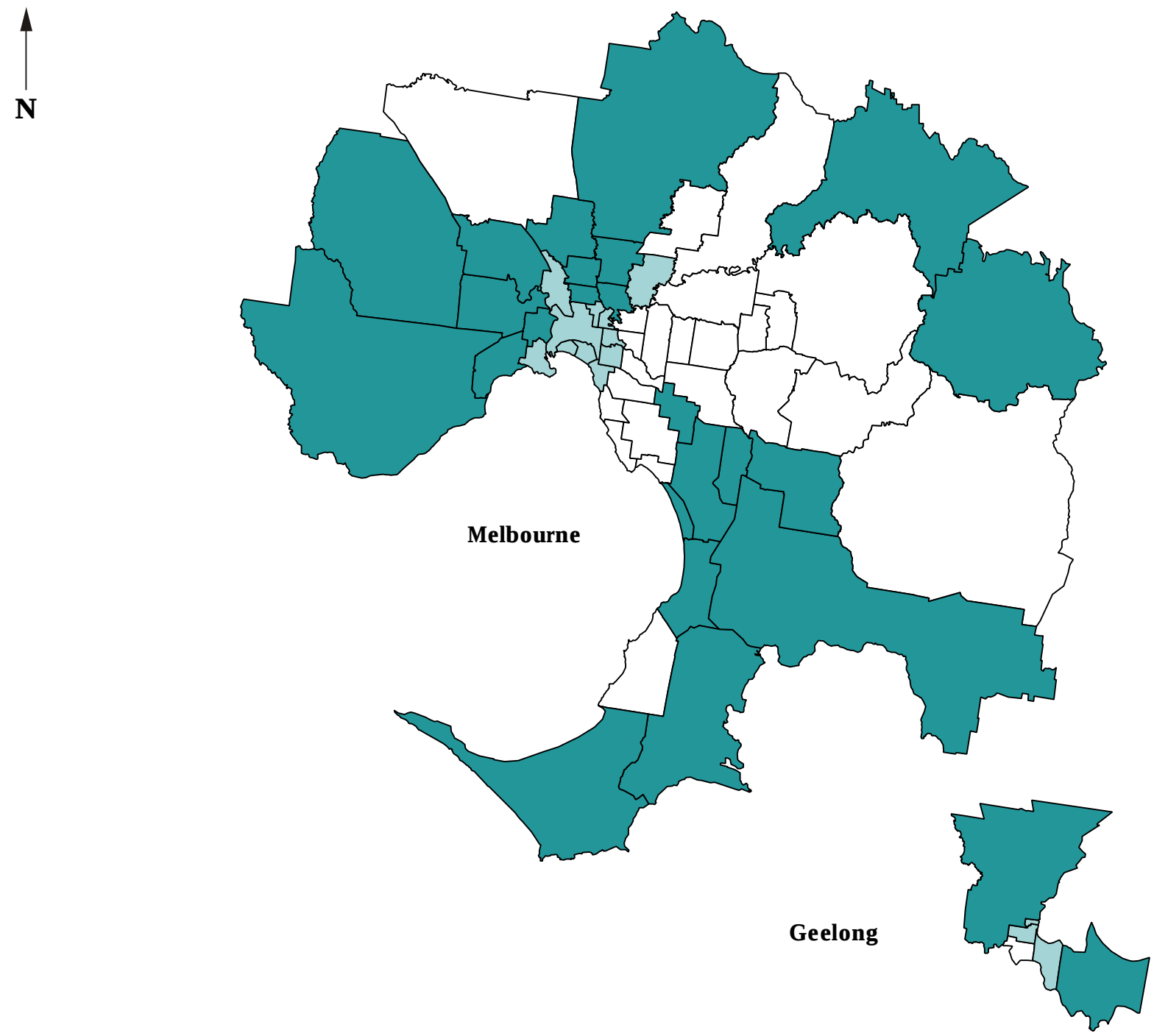
Table 8.8: Composition of SLA clusters in Geelong

SLA	Socioeconomic status	Health status	Health service utilisation	Social health status[1]
Bellarine - Inner	Low	Good	Low	Medium
Corio - Inner	Low	Medium	High	Low
Geelong	Medium	Medium	Low	Medium
Geelong West	Medium	Medium	Low	Very Low
Newtown	High	Good	Medium	High

[1] Social health 'status clusters were produced by a joint analysis of the socioeconomic status and health status variables

Map 8.1
Socioeconomic status clusters based on Statistical Local Areas, Melbourne and Geelong, 1996

clusters of SLAs with generally similar health status characteristics



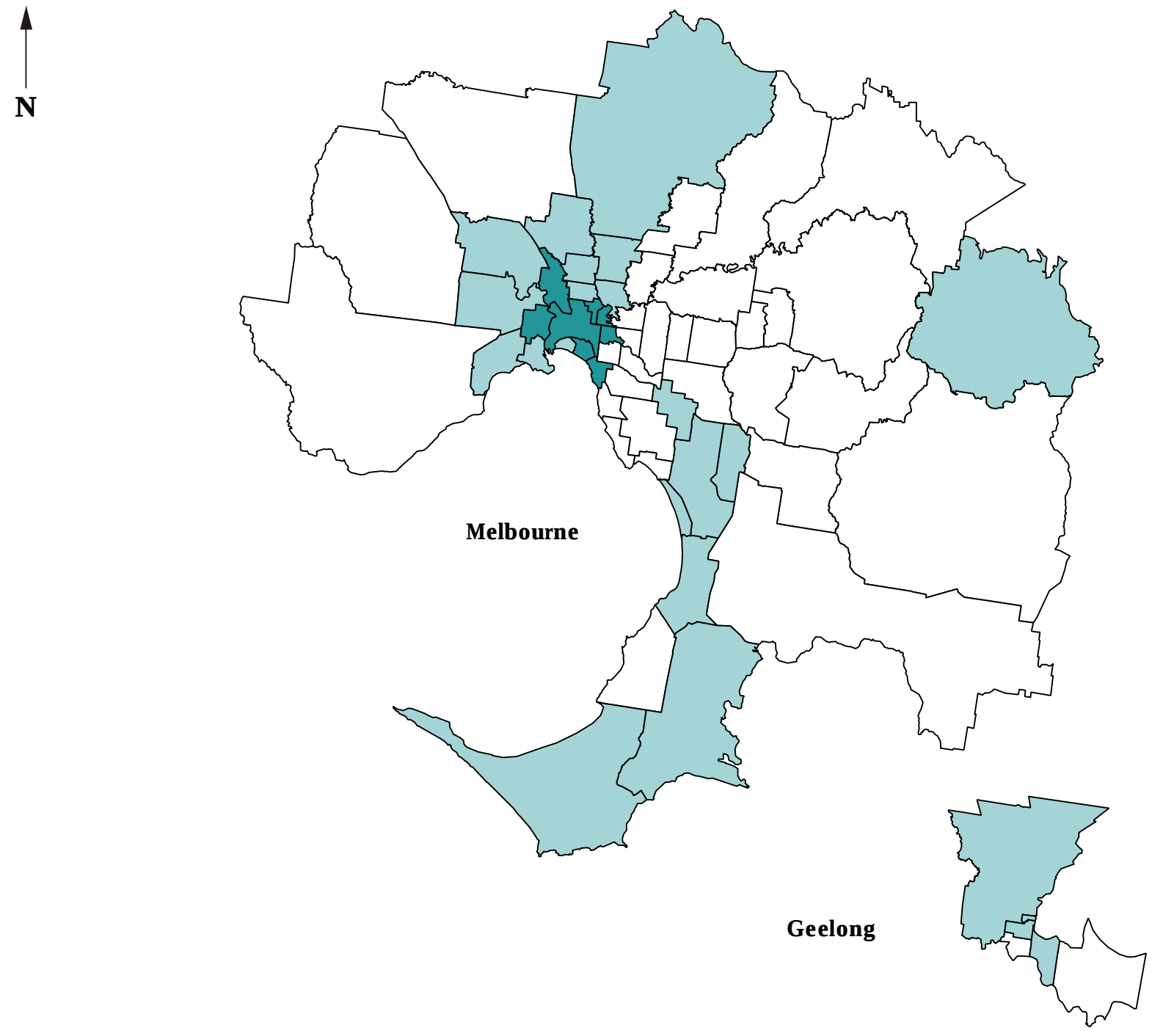
Source: Calculated on data from ABS 1996 Census

Details of map boundaries are in Appendix 1.2

National Social Health Atlas Project, 1999

Map 8.2
Health status clusters based on Statistical Local Areas, Melbourne and Geelong, 1996

clusters of SLAs with generally similar health status characteristics

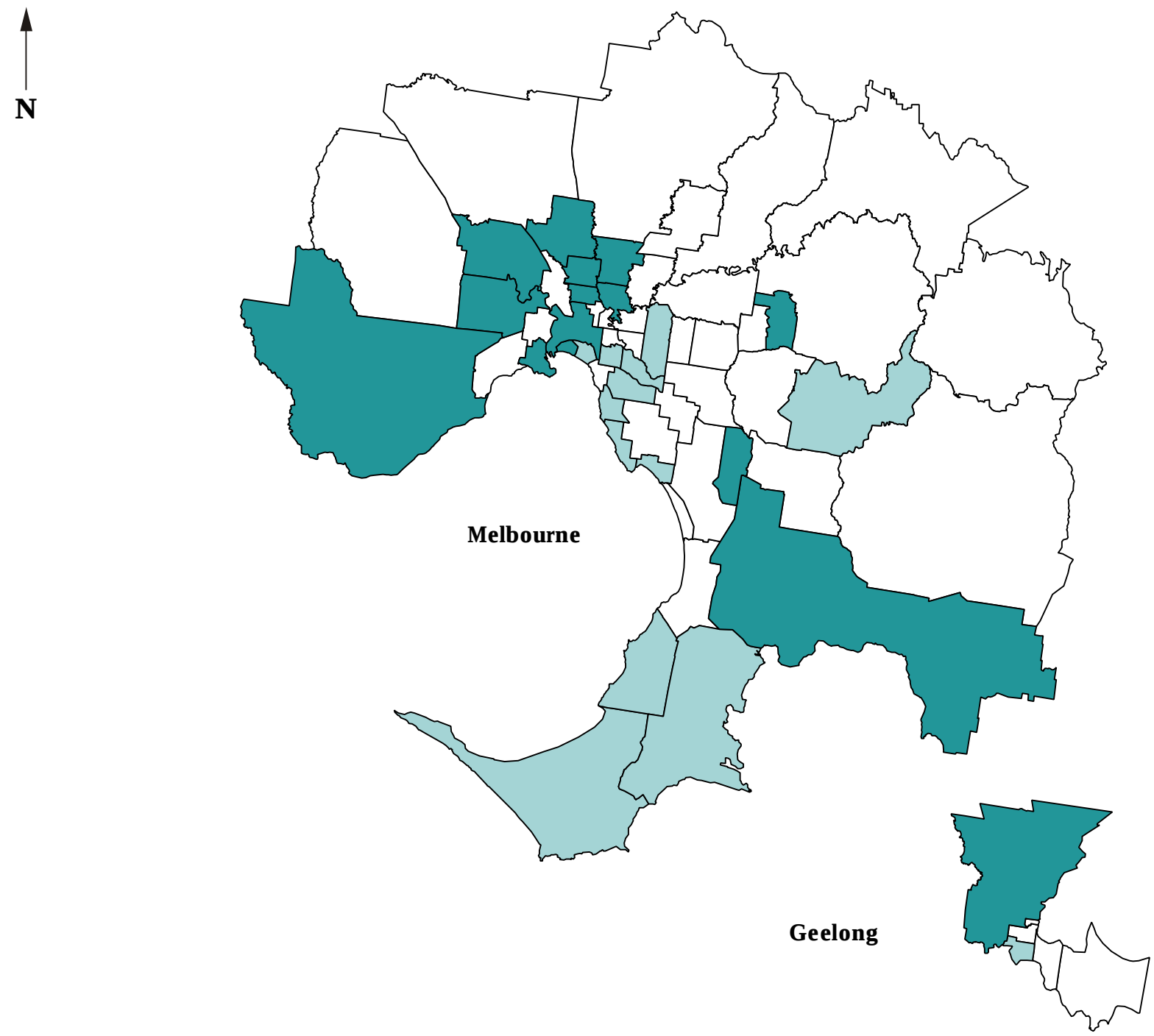


Source: Calculated on data from ABS 1996 Census

Details of map boundaries are in Appendix 1.2

Map 8.3
Health service utilisation clusters based on Statistical Local Areas, Melbourne and Geelong, 1996

clusters of SLAs with generally similar health service utilisation characteristics



Source: Calculated on data from ABS 1996 Census

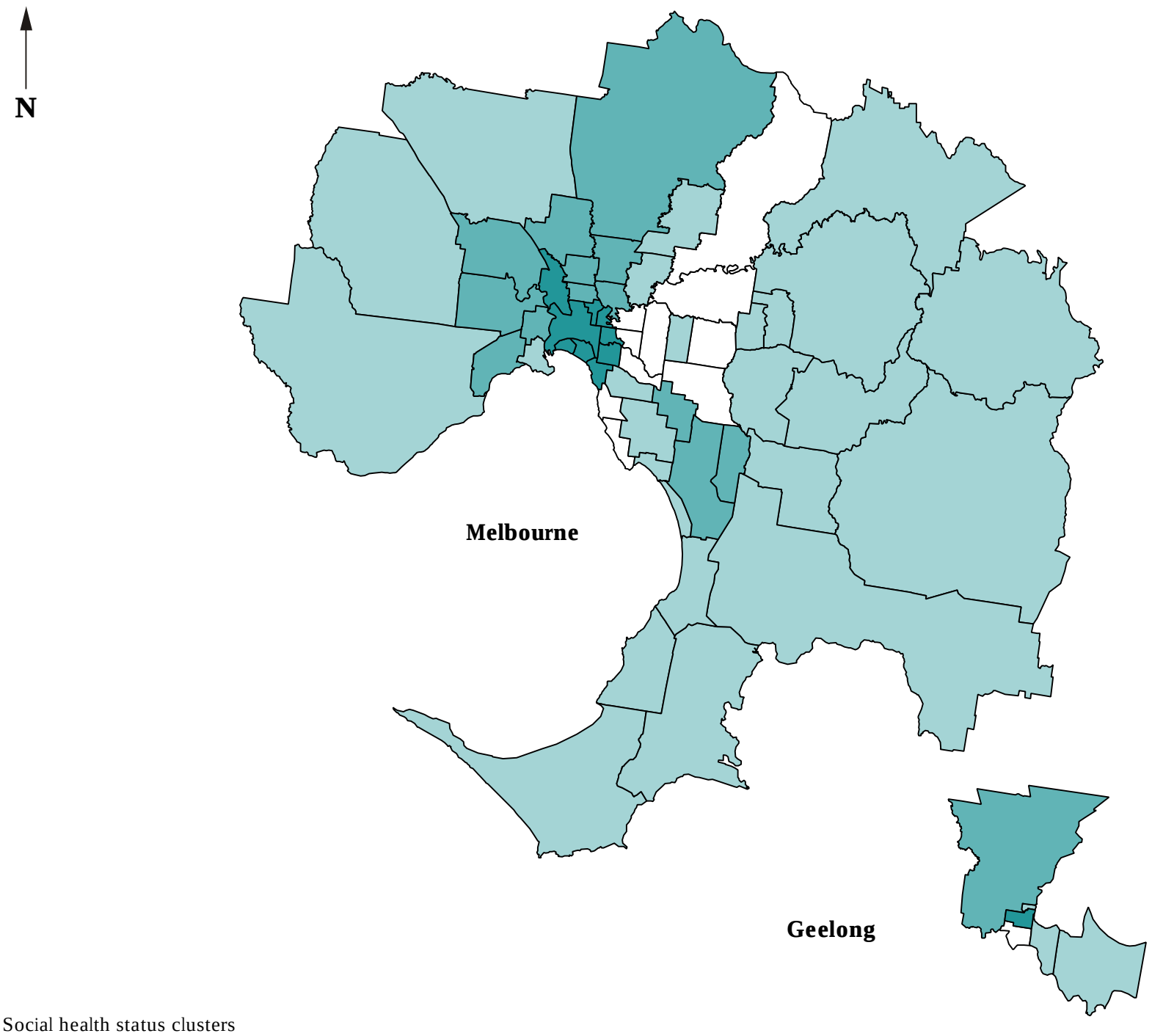
Details of map boundaries are in Appendix 1.2

National Social Health Atlas Project, 1999

Map 8.4

Social health status clusters based on Statistical Local Areas, Melbourne and Geelong, 1996

clusters of SLAs with generally similar social health status characteristics



Source: Calculated on data from ABS 1996 Census

Details of map boundaries are in Appendix 1.2

National Social Health Atlas Project, 1999

Socioeconomic clusters of non-metropolitan SLAs

The production of clusters at the SLA level in the non-metropolitan areas is even more problematic (than for **Melbourne**), with SLAs varying enormously in size and composition. For example, large urban centre SLAs such as Ballarat, Shepparton and Wodonga (with populations of 75,458, 39,629 and 29,190 respectively) stand in contrast to rural SLAs such as Yarra Ranges [Part B] (751 people) and Toowong [Part A] (2,287). Mildura [Part A], the SLA with the largest land area, occupies 9.6 per cent of Victoria's land mass yet has a population of only 40,646 (0.9 per cent of the State population). Indigenous Australians, generally the most disadvantaged population group, are unevenly distributed throughout these SLAs, from 9.2 per cent of the total population in Robinvale, 4.3 per cent in Orbost and 3.4 per cent in Swan Hill, to less than 0.5 per cent in some 43 non-metropolitan SLAs (37.1 per cent of all non-metropolitan SLAs). Despite these variations, the results of the cluster analysis are understandable.

There was data for 153 SLAs across Victoria. These 153 records are ample to carry out a cluster analysis with seven input variables. A cluster analysis was performed on the available data, and the solution examined. The three cluster solution produced was found to be of good quality and was accepted without further investigation. The SLAs in each cluster are listed in **Table 8.9** and **Map 8.5**.

Of the 11 lowest SLAs for the IRSD, only 2 (18.2 per cent) were classified to the Low socioeconomic status group and, of the top 82 SLAs for the IRSD, 72 (87.8 per cent) were classified to the High socioeconomic status group.

Health status clusters of non-metropolitan SLAs

The variables for infant deaths; deaths of 15 to 64 year olds from lung cancer and diseases of the respiratory system; and deaths of 15 to 24 year olds from the external causes of accidents, poisonings and violence were excluded from the analysis because five per cent or more of SLAs had no cases. Thus there were 11 variables to analyse 153 records.

A cluster analysis of all the above variables was tried to see if it gave a sensible solution. This produced a good quality three cluster solution which was accepted without further investigation. The SLAs in each cluster are listed in **Table 8.9** and shown in **Map 8.6**. Note that the Poor health status group did have higher status than the Good health status group for disability.

The ABS Index of Relative Socio-Economic Disadvantage (IRSD) was again used as an independent check on the solution. It was found that, of the bottom 46 SLAs for non-metropolitan SLAs in Victoria as classified by the IRSD, 34 (73.9 per cent) were classified to the Poor health status group in this analysis. Further, of the top 38 SLAs under the IRSD, 25 (65.8 per cent) were classified to the Good health status group.

Health service utilisation clusters of non-metropolitan SLAs

The variables for admissions for lung cancer, breast cancer, psychosis, tonsillectomy and/or adenoidectomy and hip replacement were excluded from the analysis because over five per cent of areas had no cases. Thus there were 20 variables to analyse 153 records. This was ample data to carry out a cluster analysis.

A cluster analysis of all the above variables was tried to see if it gave a sensible solution. This produced an extremely clean four cluster solution which was accepted without further investigation (see **Table 8.9** and **Map 8.7**).

There was moderate agreement with the IRSD: of the lowest 18 SLAs for the IRSD, none were classified to the Very high health service use cluster; and of the highest 20, four (20.0 per cent) were classified to the Low health service use cluster.

Social health clusters of non-metropolitan area SLAs

Data considered for inclusion were the variables in the final models for SLAs in the non-metropolitan areas of Victoria used to examine socioeconomic status and health status. The variables excluded from the health status model because of missing data were excluded from this model also. Thus there were 18 variables to analyse 153 records (SLAs). This is ample data on which to undertake a cluster analysis.

A cluster analysis of all the above variables was tried to see if it gave a sensible solution. This produced an very clean three cluster solution of very high quality, which was accepted without further investigation. The SLAs in each cluster are listed in **Table 8.9** and shown in **Map 8.8**. Note that the Low social health status group had a higher ranking than the High social health status group for disability.

Of the 37 lowest SLAs for the IRSD, 29 (78.4 per cent) were classified to the Low social health status cluster; and of the top 73 SLAs for the SEIFA index, 62 (84.9 per cent) were classified to the High social health status cluster.

Table 8.9: Composition of SLA clusters in non-metropolitan areas of Victoria

SLA	Socioeconomic status	Health status	Health service utilisation	Social health[1]
Alberton (S)	Medium	Poor	Medium	Low
Alexandra (S)	High	Poor	High	Medium
Arapiles (S)	High	Good	Low	High
Ararat (C)	Medium	Poor	High	Low
Ararat (S)	High	Good	High	High
Avoca (S)	Medium	Poor	Medium	Low
Avon (S)	High	Good	High	High
Bacchus Marsh (S)	High	Medium	High	Medium
Bairnsdale (C)	Low	Poor	Low	Low
Bairnsdale (S) (Part A)	Medium	Medium	Low	Medium
Bairnsdale (S) (Part B)	High	Good	Very high	High
Ballarat (C)	Medium	Poor	Medium	Medium
Ballarat (C) - North	High	Medium	Low	High
Bass (S)	Medium	Medium	High	Medium
Beechworth (S)	High	Medium	High	High
Belfast (S)	High	Medium	Very high	High
Benalla (C)	Medium	Poor	Very high	Low
Benalla (S)	High	Good	Medium	High
Bet Bet (S)	Medium	Poor	High	Low
Birchip (S)	High	Medium	Medium	High
Bright (S)	High	Medium	High	High
Broadford (S)	Medium	Medium	High	Medium
Buln Buln (S)	Medium	Medium	High	Medium
Camperdown (T)	Medium	Poor	High	Low
Castlemaine (C)	Medium	Poor	High	Low
Charlton (S)	High	Medium	Medium	High
Chiltern (C)	High	Medium	Medium	High
Cobram (S)	Medium	Medium	Very high	Medium
Cohuna (S)	High	Good	High	High
Colac (C)	Medium	Poor	High	Low
Colac(S) Surf Coast (S:Part B)	High	Good	Medium	High
Creswick (S)	Medium	Poor	Medium	Low
Daylesford and Glenlyon (S)	Medium	Poor	Medium	Low
Deakin (S)	High	Good	Very high	High
Dimboola (S)	Medium	Medium	High	Medium
Donald (S)	High	Medium	High	High
Dundas (S)	High	Good	High	High
Dunmunkle (S)	High	Medium	High	High
East Loddon (S)	High	Medium	Medium	High
Echuca (C)	Low	Poor	High	Low
Euroa (S)	Medium	Poor	Medium	Low
Gisborne (S)	High	Good	Medium	High
Glenelg (S)	Medium	Medium	High	High
Gordon (S)	High	Medium	High	High
Goulburn (S)	Medium	Poor	Medium	Medium
Gr. Bendigo (C)	Medium	Poor	Medium	Medium
Gr. Bendigo (C) - Huntly Bal	Medium	Poor	Medium	Low
Gr. Bendigo (C) - Marong Bal	Medium	Medium	Medium	Medium
Gr. Bendigo (C) - Strathfieldsaye Bal	High	Good	Medium	High
Gr. Geelong (C: Part C)	High	Good	Low	High
Greater Geelong [Balance]	High	Medium	Medium	High
Hamilton (C)	Medium	Medium	High	Medium
Hampden (S)	High	Medium	Very high	High
Heytesbury (S)	High	Good	Very high	High
Heywood(S)	High	Good	Medium	High
Horsham (C)	Medium	Medium	High	Medium

Table 8.9: Composition of SLA clusters in non-metropolitan areas of Victoria ... *cont*

SLA	Socioeconomic status	Health status	Health service utilisation	Social health[1]
Huntly - Inner	High	Medium	Low	High
Kaniva (S)	High	Medium	High	High
Kara Kara (S)	High	Good	Low	High
Karkarooc (S)	High	Medium	Very high	High
Kerang (B)	Low	Medium	Medium	Low
Kerang (S)	High	Good	High	High
Kilmore (S)	High	Good	High	High
Korong (S)	Medium	Poor	High	Low
Korumburra (S)	Medium	Medium	High	Medium
Kowree (S)	High	Medium	High	High
Kyabram (T)	Medium	Poor	Very high	Low
Kyneton (S)	High	Poor	Medium	Low
Lexton (S)	Medium	Poor	Low	Medium
Lowan (S)	High	Medium	High	High
Maffra (S)	Medium	Good	High	Medium
Maldon (S)	Medium	Medium	Medium	Medium
Mansfield (S)	High	Medium	Very high	High
Maryborough (C)	Medium	Poor	High	Low
McIvor (S)	Medium	Poor	High	Low
Metcalfe (S)	Medium	Medium	High	Medium
Mildura (C)	Low	Poor	High	Low
Mildura (S) (Part A)	Medium	Medium	High	Medium
Mildura (S) (Part B)	High	Medium	Low	High
Minhamite (S)	High	Medium	Medium	High
Mirboo (S)	High	Medium	High	Medium
Moe (C)	Medium	Poor	High	Low
Moorabool (S) - East	Medium	Medium	Medium	Medium
Mortlake (S)	High	Good	Medium	High
Morwell (C) (Part A)	Medium	Poor	High	Low
Morwell (C) (Part B)	Medium	Good	High	Medium
Mount Rouse (S)	High	Medium	Very high	High
Myrtleford (S)	High	Medium	Very high	Medium
Narracan (S) (Part A)	Medium	Medium	High	Medium
Narracan (S) (Part B)	High	Medium	Medium	High
Nathalia (S)	Medium	Poor	Very high	Medium
Newham and Woodend (S)	High	Good	High	High
Newstead (S)	Medium	Poor	Medium	Medium
Numurkah (S)	High	Medium	High	High
Omeo (S)	High	Good	Low	High
Orbost (S)	Low	Medium	Low	Low
Otway (S)	High	Medium	Medium	High
Oxley (S)	High	Good	Medium	High
Phillip Island (S)	Medium	Medium	Medium	Medium
Port Fairy (B)	Medium	Poor	High	Low
Portland (C)	Medium	Poor	Very high	Low
Pyalong (S)	High	Good	Low	High
Queenscliffe (B)	High	Poor	Medium	High
Ripon (S)	Medium	Medium	Very high	Medium
Rochester (S)	High	Good	High	High
Rodney (S) (Part A)	Low	Poor	Medium	Low
Rodney (S) (Part B)	High	Medium	High	Medium
Romsey (S)	High	Good	Medium	High
Rosedale (S)	Medium	Medium	High	Medium
Rutherglen (S)	High	Medium	High	High
Sale (C)	Medium	Medium	High	Medium
Seymour (S)	Medium	Medium	High	Medium

Table 8.9: Composition of SLA clusters in non-metropolitan areas of Victoria ... *cont*

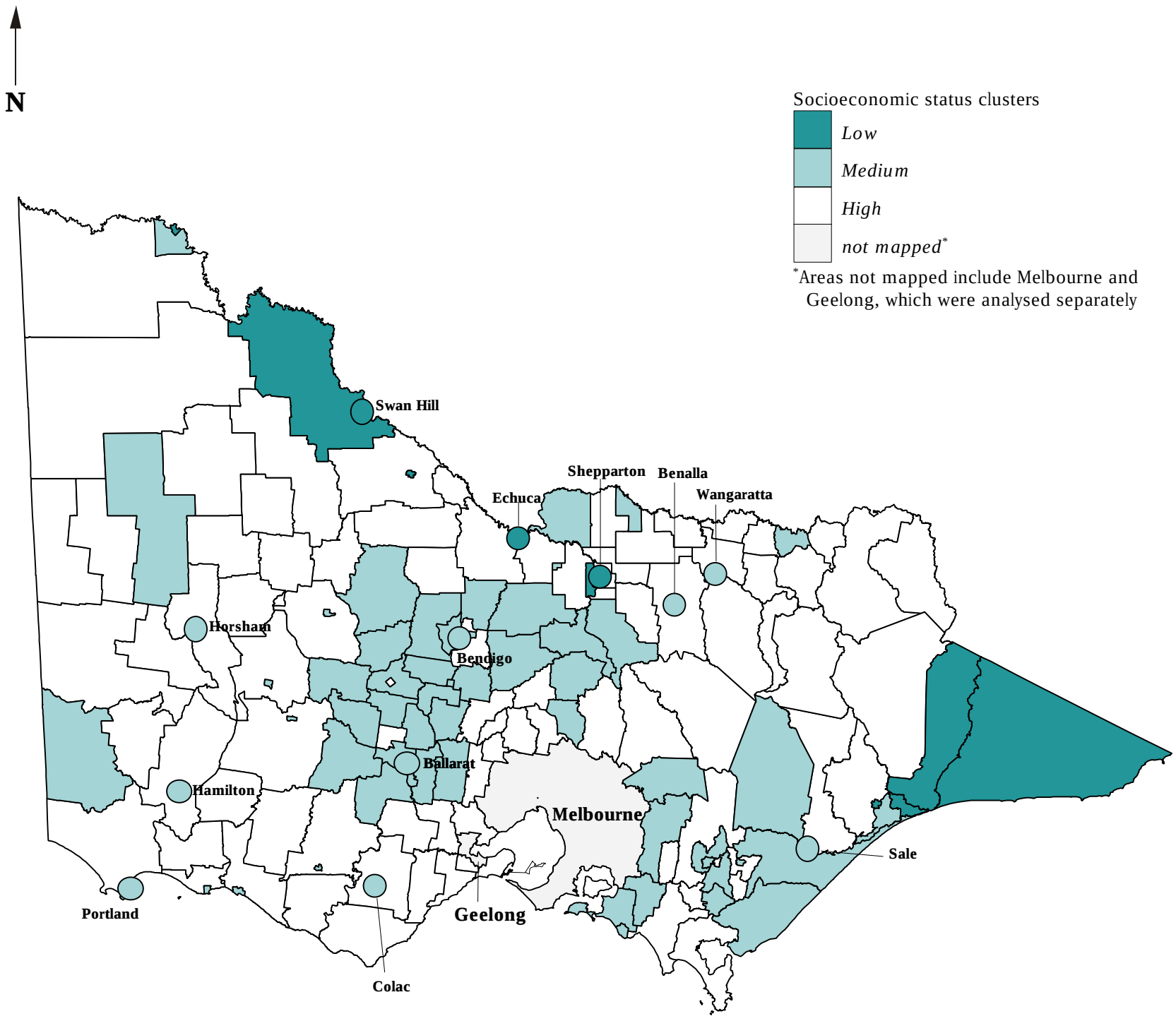
SLA	Socioeconomic status	Health status	Health service utilisation	Social health[1]
Shepparton (C)	Low	Poor	High	Low
Shepparton (S) (Part A)	High	Good	Medium	High
Shepparton (S) (Part B)	High	Medium	High	High
South Gippsland (S)	High	Good	High	Medium
Southern Rural (S) - Central	High	Medium	Medium	Medium
Southern Rural (S) - East	High	Good	Medium	High
St Arnaud (T)	Medium	Poor	High	Low
Stawell (C)	Medium	Poor	High	Low
Stawell (S)	High	Good	High	High
Swan Hill (C)	Low	Poor	High	Low
Swan Hill (S)	Low	Poor	High	Low
Talbot and Clunes (S)	Medium	Poor	High	Low
Tallangatta (S) (Part A)	High	Medium	High	High
Tallangatta (S) (Part B)	High	Good	Low	High
Tambo (S) (Part A)	Low	Poor	Low	Low
Tambo (S) (Part B)	Low	Poor	Low	Low
Traralgon (C)	Medium	Medium	High	Medium
Traralgon (S) (Part A)	High	Medium	High	High
Traralgon (S) (Part B)	High	Good	Low	High
Tullaroop (S)	Medium	Medium	Low	Medium
Tungamah (S)	High	Good	Low	High
Upper Murray (S)	High	Medium	Very high	High
Upper Yarra (S) (Part B)	Medium	Poor	Low	Low
Violet Town (S)	High	Good	High	Medium
Walpeup (S)	High	Good	Medium	High
Wangaratta (C)	Medium	Poor	Very high	Medium
Wangaratta (S)	High	Good	Medium	High
Wannon (S)	High	Medium	High	High
Waranga (S)	Medium	Poor	Medium	Medium
Warracknabeal (S)	High	Medium	High	Medium
Warragul (S)	High	Medium	High	High
Warrnambool (C)	Medium	Poor	Very high	Medium
Warrnambool (S)	High	Medium	High	High
Wimmera (S)	High	Good	Low	High
Wodonga (Rural City)	Medium	Medium	High	Medium
Wonthaggi (B)	Medium	Poor	High	Low
Woorayl (S)	High	Medium	High	High
Wycheproof (S)	High	Medium	High	High
Yackandandah (S)	High	Good	Medium	High
Yarrawonga (S)	High	Medium	High	High
Yea (S)	High	Medium	Medium	High

[1] Social health 'clusters were produced by a joint analysis of the socioeconomic status and health status variable

Map 8.5

Socioeconomic status clusters based on Statistical Local Areas, Victoria, 1994

clusters of SLAs with generally similar socioeconomic status characteristics

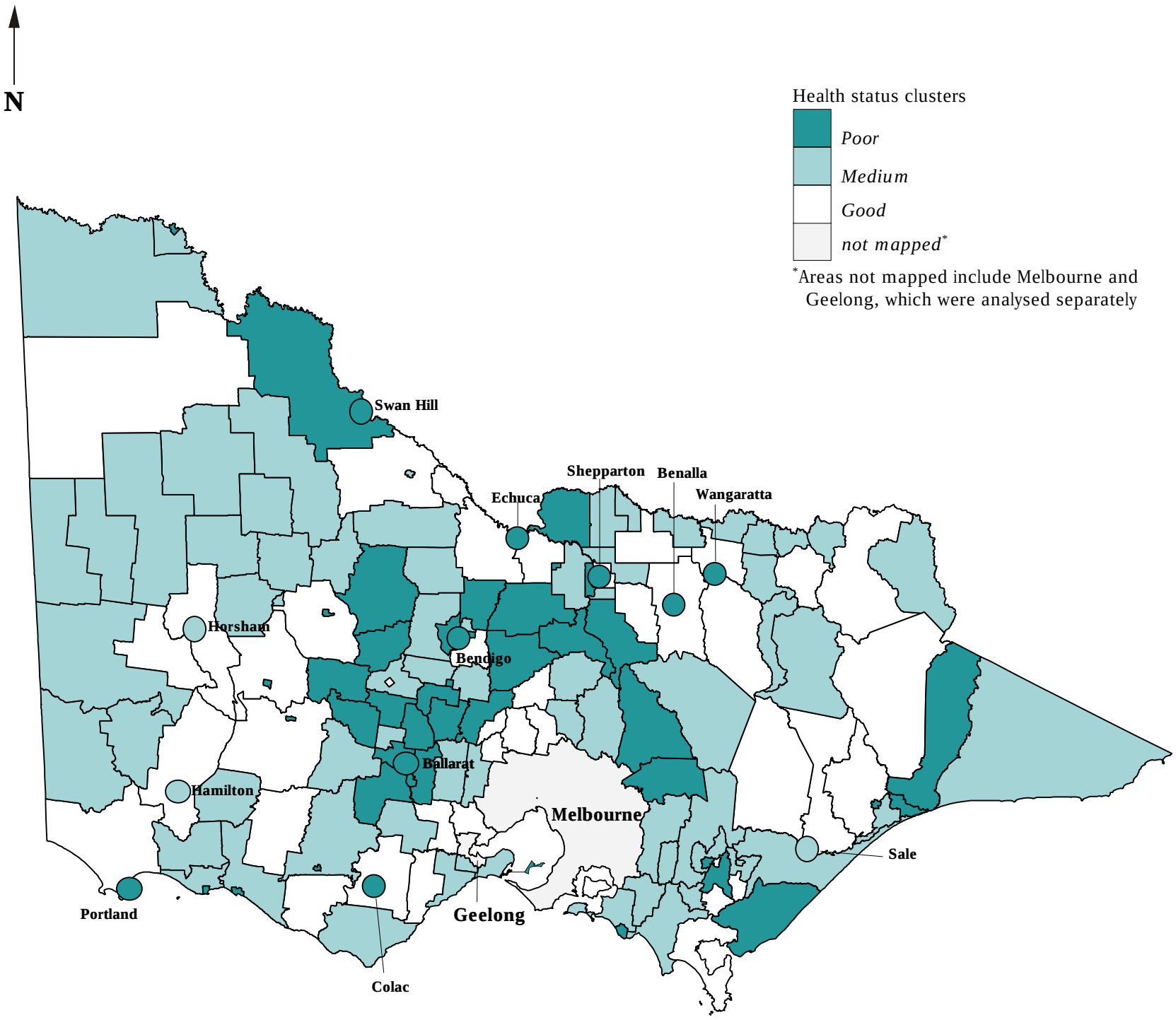


Source: Compiled from project sources

Details of map boundaries are in Appendix 1.2

Map 8.6
Health status clusters based on Statistical Local Areas, Victoria, 1994

clusters of SLAs with generally similar health status characteristics

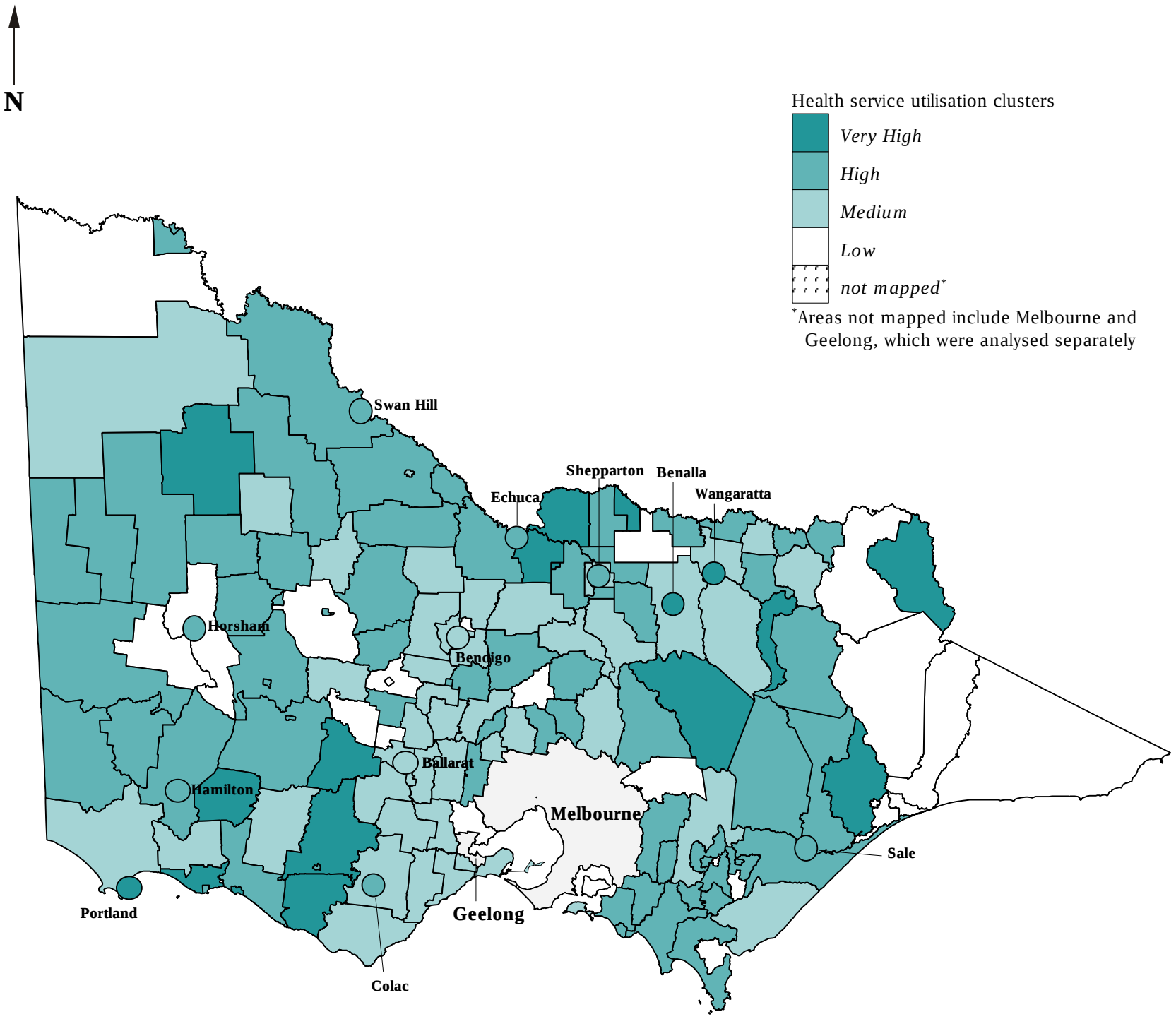


Source: Compiled from project sources

Details of map boundaries are in Appendix 1.2

Map 8.7
Health service utilisation clusters based on Statistical Local Areas, Victoria, 1994

clusters of SLAs with generally similar health service utilisation characteristics

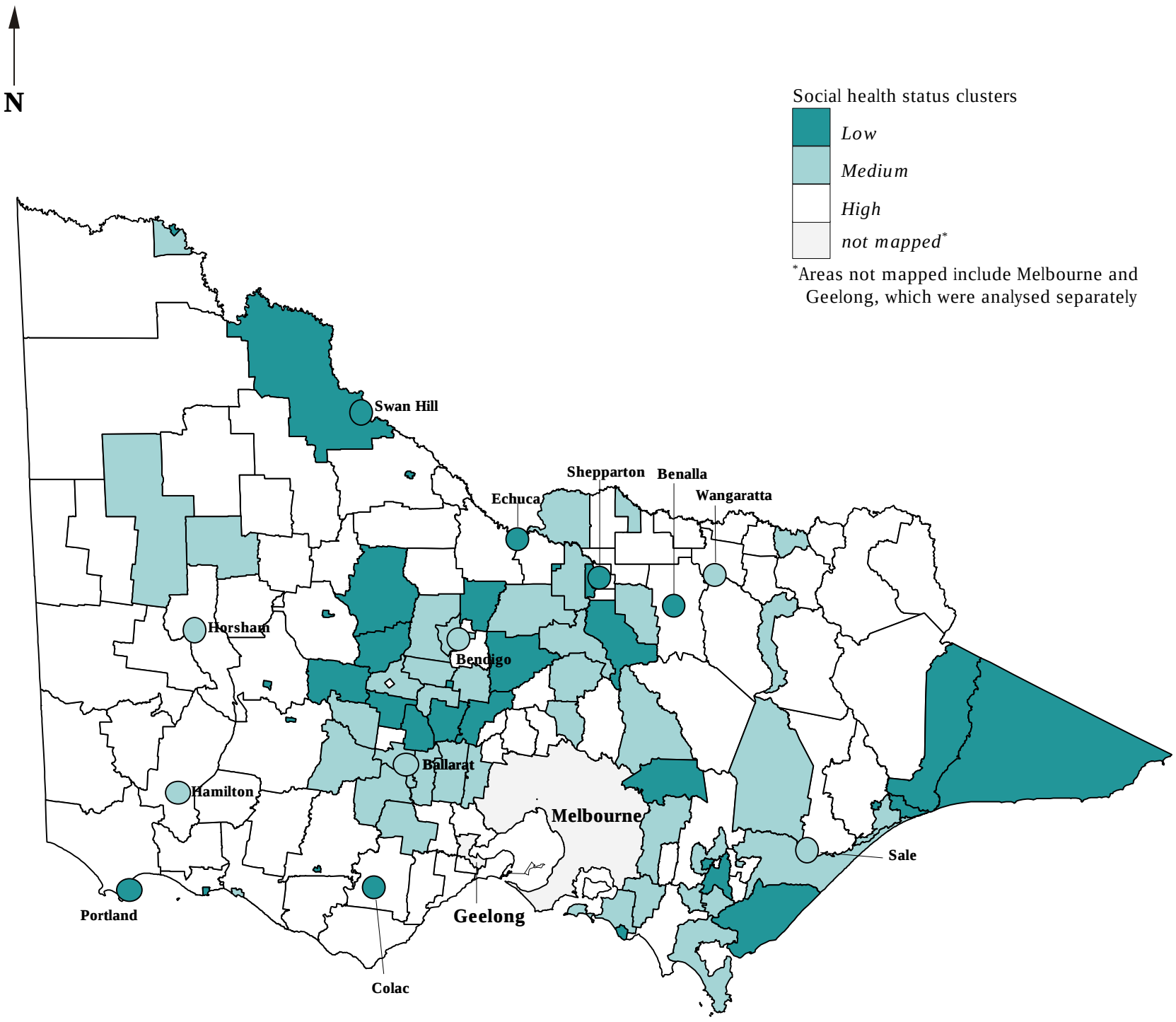


Source: Compiled from project sources

Details of map boundaries are in Appendix 1.2

Map 8.8
Social health status clusters based on Statistical Local Areas, Victoria, 1994

clusters of SLAs with generally similar social health status characteristics



Source: Compiled from project sources

Details of map boundaries are in Appendix 1.2

Socioeconomic clusters of towns

A cluster analysis was undertaken for the 55 towns (urban centres) across Australia that had populations of 7,500 or more at the 1996 Census and were identifiable in the non-Census datasets (see Appendix 1.2 for further details of the selection of these towns). These 55 records are sufficient to carry out a cluster analysis with the nine input variables.

As the analysis was somewhat complicated, only the main results are discussed below. The full description is in Appendix 1.6.

A cluster analysis was performed on the available data, and the solution examined before attempting more complicated techniques to find a solution. This analysis provided a three cluster solution of fair to average quality. It did not discriminate particularly well between clusters, and the High socioeconomic cluster did not perform particularly well against the IRSD.

The 55 records also provided enough information for an exploratory factor analysis, since this analysis has the same data requirements as the previous model.

Although several analyses were tried, the best solution was a four cluster solution (based on low income families, unemployed people, early school leavers, unskilled and semi-skilled workers, Indigenous people and single parent families). This solution is reproduced in **Table 8.10**.

The ABS Index of Relative Socio-Economic Disadvantage (IRSD) was available for the specified towns, but was withheld from the analysis and used as an independent check on the solution. It was found that, of the bottom 17 towns as classified by the IRSD, 16 (94.1 per cent) were classified to the Low socioeconomic group in this analysis. Further, of the top 20 towns under the IRSD, 15 (75.0 per cent) were classified to the High socioeconomic group.

Health status clusters of towns

There were 15 variables to analyse 55 records. This was not quite enough data. A number of alternative strategies were tried in an attempt to produce a satisfactory solution, with the outcome being a three cluster solution of good quality. The clusters were better spread than in other solutions, and it performed better against the IRSD than other solutions (**Table 8.10**).

The IRSD was again used as an independent check on the solution. It was found that, of the bottom 12 towns as classified by the IRSD, five (41.7 per cent) were classified to the Poor health status group in this analysis. Further, of the top 22 towns under the IRSD, 14 (63.6 per cent) were classified to the Good health status group.

Health service utilisation clusters of towns

There were 30 variables to analyse 55 records. This was not enough data. A number of alternative strategies were tried in an attempt to produce a satisfactory solution, with the outcome being a three cluster solution of good quality. The clusters were better spread than in other solutions, and it performed better against the IRSD than other solutions (**Table 8.10**).

A check with the IRSD showed that, of the bottom ten towns as classified by the IRSD, three (30.0 per cent) were classified to the High health service use group in this analysis. Further, of the top 26 towns under the IRSD, 13 (50.0 per cent) were classified to the Low health service use group.

Social health clusters of towns

The cluster analysis technique has also been applied to a combination of the socioeconomic status and health status data sets. Data considered for inclusion were the variables in the final models for towns used to examine socioeconomic status and health status.

There were 24 variables to analyse 55 records. This was clearly not enough data. A cluster analysis of all the above variables was tried to see if it gave a reasonable solution despite the lack of data. This produced a three cluster solution of fair to average quality. The solution did not perform at all well against the IRSD for the Low status group, and lacked definition between the Medium and Low status groups.

Alternative strategies were tried in an attempt to produce a better solution, with the outcome a three cluster solution of reasonable quality, with Charters Towers (C) not grouped. The clusters were better spread than in other solutions, and the solution performed better against the IRSD than other solutions (**Table 8.10**).

Of the 17 lowest towns for the IRSD, nine (52.9 per cent) were classified to the Low social health status cluster; and of the top 14 towns for the IRSD, seven (50.0 per cent) were classified to the High social health status cluster.

Table 8.10: Composition of town clusters in Australia

SLA	Socioeconomic status	Health status	Health service utilisation	Social health status[1]
Albany (T)	Very low	Medium	Low	Medium
Albury (C)	High	Medium	Low	Low
Alice Springs (T)	Low	Medium	Medium	Low
Armidale (C)	High	Good	High	High
Ballarat (C)	High	Good	Low	Medium
Bathurst (C)	High	Good	Low	High
Benalla	High	Medium	High	Medium
Bendigo (C)	High	Good	Low	Medium
Broken Hill (C)	Very low	Poor	Low	Medium
Broome (S)	Low	Medium	Medium	Medium
Bunbury (C)	Medium	Good	Medium	High
Burnie (C)	Very low	Poor	Low	Low
Cairns (C)	High	Good	Low	High
Casino (A)	Very low	Medium	Medium	Low
Charters Towers (C)	Medium	Poor	Medium	Not grouped
Colac	Medium	Poor	Low	Low
Dalby (T)	Medium	Medium	Low	High
Deniliquin (A)	High	Poor	Medium	Medium
Devonport (C)	Very low	Medium	Low	Low
Dubbo (C)	High	Good	Medium	Medium
Echuca	High	Medium	Low	Medium
Geraldton (C)	Very low	Medium	Low	Medium
Gladstone (C)	Medium	Good	Low	High
Goulburn (C)	Medium	Medium	Medium	Low
Grafton (C)	Very low	Medium	Medium	Medium
Hamilton	High	Good	Low	Medium
Hervey Bay (C)	Very low	Medium	Low	Low
Horsham (RC)	High	Good	Low	Medium
Inverell (A)	Very low	Medium	High	Medium
Kalgoorlie/Boulder (C)	Medium	Poor	Medium	High
Katherine (T)	Low	Poor	Medium	Low
Launceston (C)	High	Good	Low	Medium
Mandurah (C)	Very low	Medium	Low	Low
Maryborough (C)	Very low	Medium	Low	Medium
Mount Gambier (C)	Medium	Good	High	High
Mount Isa (C)	Medium	Medium	Medium	High
Murray Bridge (RC)	Very low	Medium	Low	Low
Noosa	High	Good	Low	Medium
Orange (C)	High	Good	Medium	Low
Port Augusta (C)	Very low	Poor	Medium	Low
Port Hedland (T)	Medium	Medium	Medium	High
Port Lincoln (C)	Very low	Poor	High	Low
Port Pirie (C)	Very low	Poor	High	Medium
Portland	Very low	Poor	High	Medium
Queanbeyan (C)	High	Good	High	High
Rockhampton (C)	Medium	Good	Low	High
Sale	High	Good	Low	Medium
Shepparton (C)	Medium	Good	Medium	Low
Swan Hill (RC)	High	Good	Low	Medium
Tamworth (C)	High	Medium	Medium	Medium
Toowoomba (C)	Medium	Good	Low	High
Wagga Wagga (C)	High	Good	Medium	High
Wangaratta (RC)	Medium	Good	Medium	Low
Warwick (S)	Medium	Poor	High	Medium
Whyalla (C)	Very low	Medium	High	Low

[1] **Social health 'status clusters were produced by a joint analysis of the socioeconomic status and health status variables**

This page intentionally left blank