COMP20008 Project Phase 2A

Urban Planning, Housing and Implications on Cultural Diversity

Domain

Community, Urban Planning & Real Estate

Question

This project seeks to provide an optimistic solution to the following questions:

- *Does housing policies and prices affect language and cultural diversity in Victoria?*
- *Could proper urban planning and artificial manipulation of real estate pricings be used to improve spatial segregation if present in Victoria?*

The Local Cities, State and Federal governments should all likely be interested in the results of these questions regarding future planning for the socio-cultural environment of the country. The government departments of social services, immigration and infrastructure development as well as finance would most likely want answers to these long-term questions. The information could be used for urban planning and setting housing prices and policies to promote cultural and linguistic diversity in Victoria for the sake of national identity and racial harmony.

Highlights of Information

Social polarisation in residential areas by cultural and linguistic means poses a long-term threat to the social harmony and national identity of countries. By looking at property prices, population density and languages data sets, the severity of this issue present in Victoria can be determined. Correlations could be found while looking at population density and housing prices, while referencing which language group is dominant in the area. Properties which are priced inefficiently for a certain culture/language group for locals/ immigrants alike could have their purchase policies or pricings altered by the governments as to have a slightly more uniform blend of people of all cultures and languages.

Data

The following datasets were used:

- Victorian Property Sales Report – Median House by Suburb 2005-2015
  Victorian State Government Open Data. Contains Data of the median property prices of each suburb from 2005-2015 with some calculated rate of change and growth rate.
  URI: https://www.data.vic.gov.au/data/dataset/victorian-property-sales-report-median-house-by-suburb-2004-2014
- ABS Census Data Pack – T01 Selected Person Characteristics
  Australian Bureau of Statistics Data. Contains Data of the population estimate for each area categorised by age, citizenship status, sex etc.
  URI: http://www.abs.gov.au/websitedbs/censushome.nsf/home/datapacks
- ABS Census Data Pack – T10 Language Spoken at Home by Sex
  Australian Bureau of Statistics Data. Contains Data of all known spoken languages in Australia grouped by area categorised by each census year, sex and language itself.
  URI: http://www.abs.gov.au/websitedbs/censushome.nsf/home/datapacks

Data Processing Benefits

As the issue addressed is about urban planning, mapping out the data of spoken languages and population density would be the best way to visualise these data sets. 6 language groups were extracted from the languages data set, the top 5 languages spoken in Victoria other than English and the number of people speaking English only per region. By comparing the densely-populated areas map and the individual maps for each language, we can locate the areas of concentrated language groups that would need to have some policies implemented depending on the density of the speakers. The ratio/percentage of the language spoken and of the population estimate for that area would be found for any area with a high concentration of one language. This would allow easier visualisation rather than look at all the languages of Australia and every region.

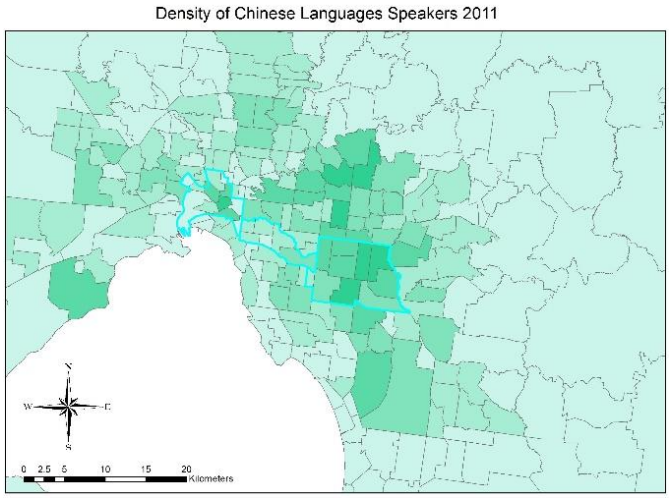
As the data for population estimates and languages spoken are census data taken only during the census years, the data for property prices must be selected out for the years 2006 and 2011. The median housing price for an area deemed to be too densely populated of people of a certain language group would be examined to determine if any actions are required to change it.

Initial Investigation

Languages Dataset:
Only the Data for Total Persons speaking the language were looked at instead of each sex. 6 languages were picked, one for English-only speakers, the others being the top five spoken languages other than English. They are Arabic, Chinese, Greek, Italian and Vietnamese. Work was done to combine language groups like Arabic type languages and Chinese type languages into individual ones for a more presentable result and map. Only data from the census years of 2006 and 2011 were taken into consideration as there are no data for the Property Prices of 2001 suburbs. A new csv file was created from data of all persons of the 6 languages for 2006 and 2011 after checking that there were no null cells without data.

The data set (csv file) was then mapped out onto maps for each year's density for each language using ArcMap. This helps projects a rough visualisation of the density of a spoken language in every area but not yet the population density. The following map is an example of just one in many.

Density of Chinese Speakers in regions in Victoria 2011



Population Dataset:
The csv file came with population estimates grouped by categories such as age, sex, citizenship status, aboriginal status, birthplace etc. Only the columns of 2006 and 2011 for total persons count for each area were extracted along with the region IDs and placed into a new csv file. Some cleaning was done on region IDs with zero population as it would be better ignoring them rather than dividing by zero in later calculations, not impacting the mapping and visualisation of the population density of Victoria. ArcMap was used to map out the population density of each area.

Property Sales Dataset:
After checking that there were no data for each suburb that was completely null, the excel file was first converted into a csv file for processing. A python function using pandas and numpy libraries was written to replace missing values for some regions for certain years. Since a growth rate per annum for each region is already given in the last column of the data, it was incorporated into the calculations done by the function. The following formula was used:

$$P_1 = P_0 \times (1 + r)$$

Where P1 is the current cell with no data and P0 is the previous cell from the previous year with data. In the case where no data for 2005 exists, the formula is reversed. After the csv file was successfully run through the function and each null value is successfully replaced, the columns of 2006 and 2011 were extracted with the suburb names into a new clean csv file for later.

Correlation and Confidence:
A function was written using the pandas and numpy library to calculate the density of each spoken language given the estimated population of the area. The function skips areas that do not appear on the data frame of the population estimate csv. It then matches each value of each language in that region with the total population estimate of the region using a loop and inner function. The correlation coefficient is then calculated with the simple formula:

$$c_{region} = \frac{Speakers_L}{Estimated\ Population_{region}}$$

The coefficients were observed to be high mostly in areas which already has a high population of certain individual language speakers and thus it could be concluded that these suburbs would yield results if observed in depth.

Feasibility of Project

By observing the trends of specific suburbs that have the highest coefficients and considering the details unique to those areas with property pricing taken into consideration, it can be derived if the region needs urban planning policies which promote better cultural/language diversity. Detailed analysis on these suburbs, while considering the concentrated language could yield different conclusions. Suburbs with high coefficient for languages from historically newer immigrants like Chinese, Arabs and Vietnamese might need to have housing policies implemented and housing prices increased. Suburbs with high coefficient for English and languages from historically older immigrants like Italians and Greeks might need to have housing prices artificially decreased. Newer immigrants need to be promoted into moving into areas with lower housing prices. Further research and analysis of the data would prove if these areas are also areas which historically older immigrants reside mostly. This would highlight an unexpected segregation problem in Victoria.