# Analysing Australian Election Results

Carlos Yanez Santibanez

2023-06-12

## Contents



```
uofgdataanalyticsreport::cd_page_title()
```

DATA ANALYTICS
GLASGOW

University of Glasgow

School of Mathematics and Statistics

Data Analytics MSc (ODL)

# My Course

# Analysing Australian Election Results

Carlos Yanez Santibanez

‘r Sys.Date()‘

# 1 Acknowledgements

# 2 Introduction

# 3 Data

## 3.1 Data Sources

The first step in the process was to source demographic and electoral data, which has been provided from two sources:

- **The Australian Electoral Commission** (AEC) [Commission, 2023a] . The AEC contains detailed online records for every federal election held in the 21st century, through their Tally Room website [Commission, 2023b].
- The **Australian Bureau of Statistics** (ABS) [of Statistics, 2023a]. The ABS provides a wide number of national statistics and is responsible to conduct a national census of population and housing every 5 years. Comprehensive census data is provided in multiple formats, including csv files through Census Data Packs [of Statistics, 2023b], which are available for censuses from 2006 onwards.

Both organisations are the authoritative source for electoral and statistical data in Australia, and the data is provided openly. Although there are no quality issues, the way that data is provided presents other challenges, namely:

- In both cases, data are provided in large volumes and exhaustive granularity. If not done effectively, data extraction and aggregation can be time-consuming and resource intensive.
- Census data points are provided using the ABS own geographical standard - and only a small selection of census data is provided at the electoral division level. Conversion between ABS geographical structures and electoral divisions is not straightforward as there is no 1:1 correspondence. Both geographical systems change from election to election and census to census.
- Despite the best efforts of both organisations in keeping consistency, names of electorates, parties, and census attributes change over time - to compare similar statistics manual mapping is necessary.

To address these issues and ensure repeatability, three R packages have been written to undertake this task:

- **{auspol}** [**Yáñez Santibáñez, 2023b**], which extracts and presents electoral results.
- **{auscensus}** [**Yáñez Santibáñez, 2023a**], which allows to interact with Census Data Packs to extract different statistics across geographical units, and across censuses.

- **{aussiemaps}** [**Yáñez Santibáñez, 2023c**], which assists with aggregating census data into electoral divisions, by matching and apportioning different geographical structures.

The appendix contains a vignette for each package, explaining their respective *modus operandi*. At a higher level, the extraction pipeline for this project is represented by figure 1.

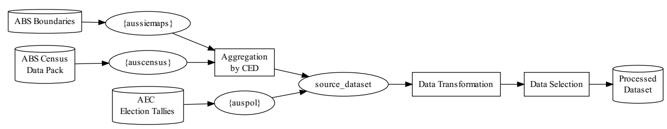



Figure 1: Flow of data from sources to dataset

In summary, the process followed consisted of the below steps:

1. Census data was extracted from the respective Census Data Pack using **{auscensus}**. Using the package workflow, key attributes were identified in each census, extracted from the respective files and given common names. Data were extracted for statistical areas and apportioned into Commonwealth Electoral Divisions by overlapping area, with the help of functions written into **{aussiemaps}**

2. Primary vote results for each division were extracted using the **auspol** package.

3. All the data was stored in a local database, from where was extracted and put together in a single dataset.

4. From there, the "raw" data was further processed and stored in a single "consolidated" dataset.

## 3.2 Data Selection

### 3.2.1 Census and Election Years

The first to address when extracting the data is to establish a correspondence between census and election data. Since election the census cycle (5 years) does not match the electoral cycle (determined by the incumbent government, with a 3-year term for the House of Representatives), there is a potential problem of the census data not being completely representative of the population on a given election day. Figure 2 presents the best matches between both events held in the 21st century.

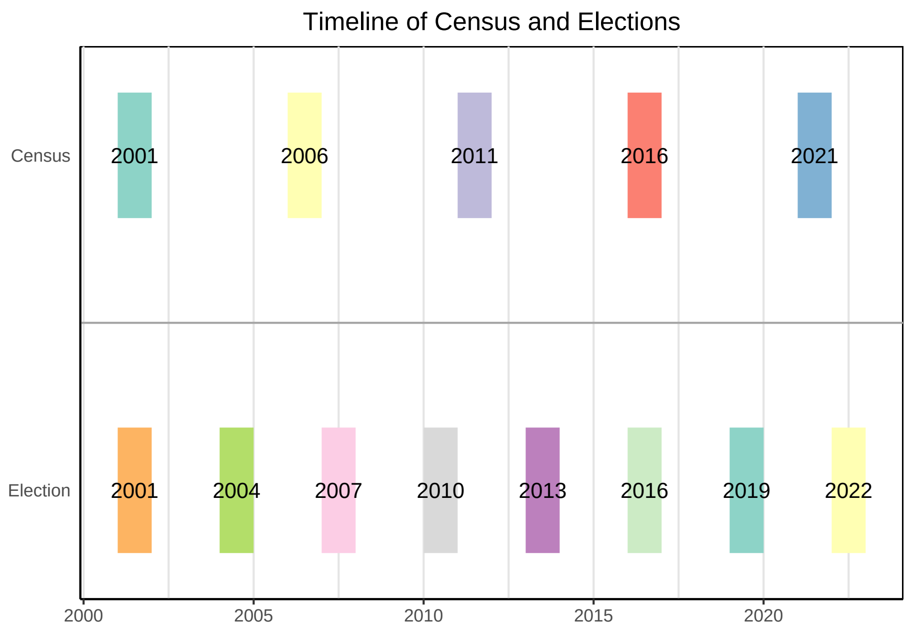



Figure 2: Census and Elections Timeline

Considering the census data available and selecting the elections closer to each census, four pairs of events were selected for data extraction. there are presented in table 1

Table 1: Selected Census-Election pairings

| Census | Election |
|---|---|
| 2006 | 2007 |
| 2010 | 2011 |

Table 1: Selected Census-Election pairings

| Census | Election |
|---|---|
| 2016 | 2016 |
| 2021 | 2022 |

Please note that this selection will remove half of the elections within the period, which may have an effect on model accuracy. However, since the objective is not to obtain an accurate prediction this has been accepted as a trade-off to avoid having to interpolate demographic attributes between censuses - which is also subject to inaccuracies given the rapid demographic changes experienced in Australia's main cities.

#### 3.2.2 Electoral Data

In the case of the electoral data. not much processing was required. The source data already contains records of primary voting for each electorate and only percentages have been calculated. In addition, the number of total votes per party at the national and state level have been calculated. A sample of the extracted data is presented in table 2.

Table 2: Sample extraction - Canberra 2022

| Year | Division | Abbreviation | Party | Votes | Percentage |
|---|---|---|---|---|---|
| 2022 | Canberra | ALP | ALP | 34,574 | 45.2% |
| 2022 | Canberra | GRN | GRN | 19,240 | 25.2% |
| 2022 | Canberra | COAL | Liberal (Coalition) | 16,264 | 21.3% |
| 2022 | Canberra | Other | Other Parties | 6,417 | 8.4% |

#### 3.2.3 Census Data

As mentioned in section 3.1, a major challenge with respect of census data is the large volume of data points collected. For instance, the data pack for the 2022 Census contains 62 different tables, ranging from 8 [1] to 1,590 [2] attributes.

To select which variables to extract, literature and journalistic sources were consulted ([Biddle and McAllister, 2022], [Parliament], [Jakubowicz and Ho, a])

[1] *02 -Selected Medians and Averages*
[2] *09 - Country of Birth of Person by Age by Sex*

to inform an initial set of covariates. In total (XYZ) variables were selected, which correspond to below to the following groups:

1. **Income** : Distribution of population in pre-set income brackets.
2. **Education Level**: Distribution of educational achievement (from incomplete secondary to vocational education and academic degrees).
3. **Age**: Distribution of the population in generational cohorts. Taking into account the selected elections, the four groups of interest are Baby Boomers (1946 to 1964), Generation X (1965 to 1980), Generation Y (1981 to 1996) and Generation Z (1997 to 2021).
4. **Relationship status**: Variables describing civil status (e.g. living alone, married, in a de facto relationship).
5. **Household type:** Descriptors of type of housing , (e.g. standalone house, semi-detached, flats).
6. **Household tenure:** Descriptors of house ownership, rental or other arrangement (e.g. public housing).
7. **Citizenship**: Percentage of the population that hold Australian citizenship. Although non-citizens are not entitled to vote, this variable can be taken as a proxy for relative integration of migrant communities into civic life.
8. **Religion:** Percentage of the population declaring to profess a religion. For this analysis, largest and high growth religious groups were selected (No religion/secular, Roman Catholic, Anglican-Presbyterian-Uniting, Christian Orthodox, Other Christianity, Islam, Hinduism, Buddhism).
9. **Language**: Languages spoken in the community. Similar to religion, a selection of relevant language have been included to reflect the historic and current migrant communities.

Apart from those, each electorate has been classified as **metropolitan** if it lies within the boundaries of Australian capital cities or **non-metropolitan** if not. Altogether, these variables try to reflect wealth and education (cited by [Biddle and McAllister, 2022] as key factors in deciding political persuasion), as well as stage in life and belonging to a particular migrant community (sometimes cited as an influential factor, for instance in [Jakubowicz and Ho, b]).

A sample of the resulting dataset is present in table 3.

Table 3: Dataset sample

| election_year | DivisionNm | ALP | COAL | GRN | Other | Australian_Citizens | Age_Baby_Boomers | Age_Gen_X |
|---|---|---|---|---|---|---|---|---|
| 2010 | Blaxland | 13.13 | -10.09 | -5.55 | 0.55 | 82.56 | 14.29 | 20.11 |
| 2010 | Casey | -2.40 | 4.52 | -0.45 | | 90.32 | 17.58 | 22.96 |
| 2007 | Oxley | 15.26 | -9.44 | -2.68 | -2.65 | 82.92 | 23.06 | 25.20 |
| 2007 | Rankin | 12.15 | -7.61 | -3.03 | -3.16 | 82.32 | 25.46 | 22.68 |
| 2022 | Ballarat | 11.90 | -8.09 | 2.41 | -5.50 | 90.27 | 22.56 | 20.18 |
| 2010 | Kooyong | -10.75 | 9.10 | 6.78 | | 83.82 | 16.47 | 22.63 |
| 2007 | Prospect | 14.54 | -10.51 | -3.27 | -1.50 | 86.43 | 25.83 | 22.14 |
| 2007 | Greenway | -5.14 | 8.44 | -2.03 | -0.31 | 86.67 | 24.30 | 25.51 |

### 3.3 Data Exploration

In total, the resulting dataset is made up of 4 response variables and 55 potential predictors, plus identificatory attributes like division mane and election year. As expected, an initial inspection shows that some of the covariates are loosely correlated with primary vote. Also expected, many of the covariates exhibit medium to high correlation levels amongst themselves, e.g. negative correlation between high and low level income groups, and certain age brackets with houselhold type and tenure.

As examples, figure 3 show a somewhat weak correlation between Coalition primary vote and percentage of baby boomer population. Figure 4 presents the correlation values for religion and language attributes that aside from expected pairings (e.g. Hinduism and South Asian languages or Italian speakers and percentage of declared catholics), there is a almost exclusive positive correlation between membership to Anglican, Presbyterian and Uniting churches and percentage of monolingual English speakers. The percentage of monolingual English speakers is also negative correlated to all other language groups.

Besides from this, it is worth noticing that :

- There is no apparent change in the relationship between a given covariate and the responses when broken down by state or capital city.

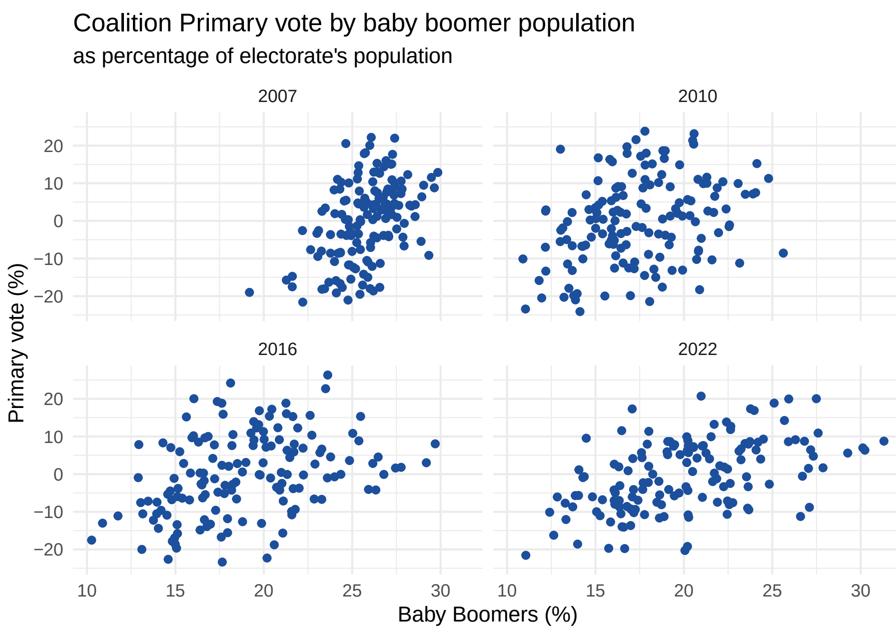



Figure 3: Correlation between Coalition vote and Baby boomer population

- There are also no obviously distinguishable differences when splitting results by each election.

### 3.3.1 Dimensionality reduction using Multiple Factor Analysis

Given the large number of variables and considering their correlation, it is worth exploring if a change of space could help to better identify variation, and whether the number of covariates can be reduced in a meaningful way. For this **multiple factor analysis** (MFA) [Escofier and Pagès, 2008] was used, given that:

- MFA allows to use variables that belong to groups.
- Allows to combine quantitative and qualitative variables.

The resulting scree plot and cumulative variance is presented in figure 5.

In terms of interpretability of the new dimensions, figure 6 present group biplots for the 8 most important dimensions. From there, it is possible show that there is not straightforward representation expect with Dimension 2 and Education variables.

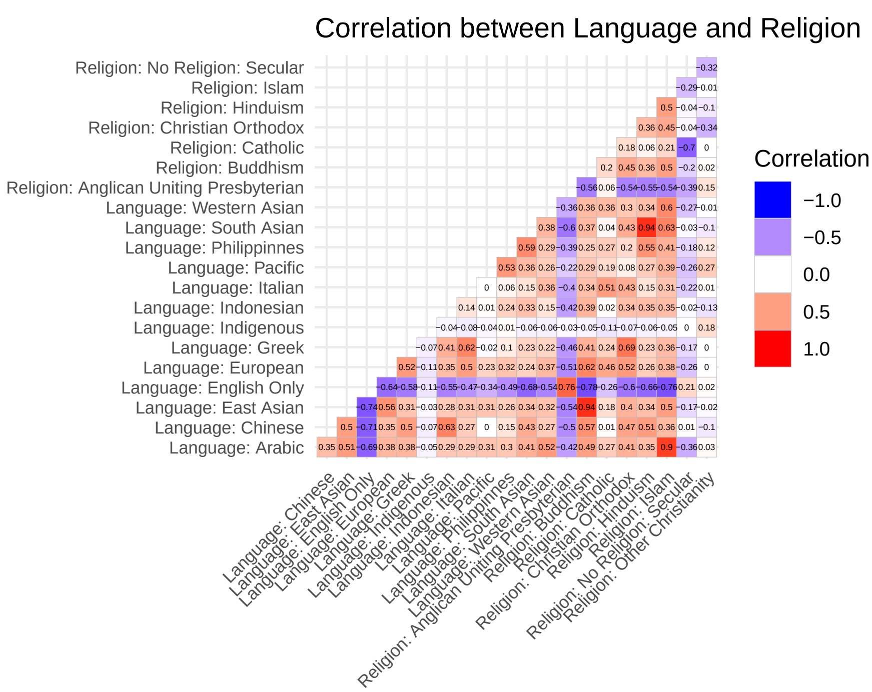



Figure 4: Correlation for selected covariates

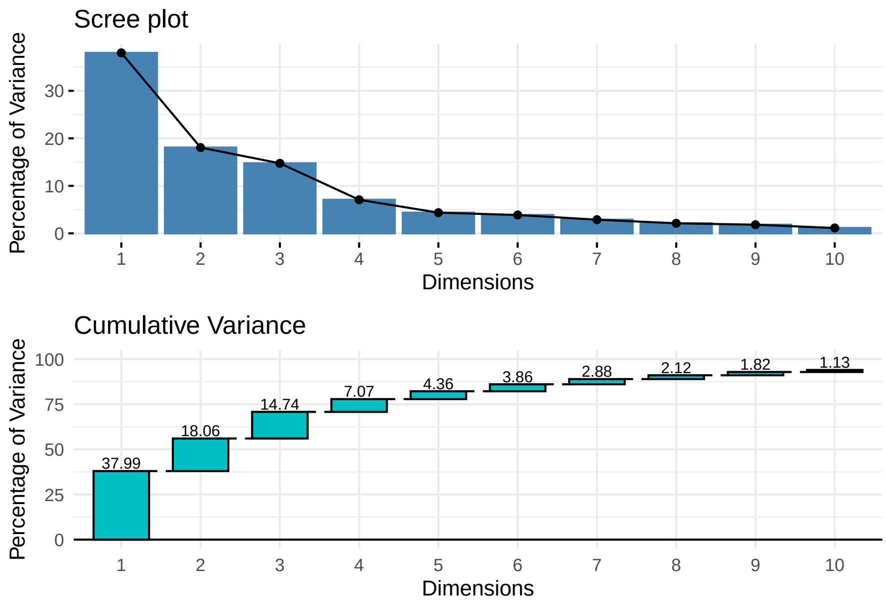



Figure 5: Scree plot and cumulative variance

#### 3.3.2 Electorate segments

A common trope of Australian politics is to map voters's political persuasion to whether they live in the inner cities, suburbia or regional areas. Therefore, it is relevant to explore if this can be substantiated by demographic attributes, or if there any other grouping of voters that may influence primary voting.

To explore this, a clustering algorithm has been applied using all electorates for election from 2006 up to 2016. After trial and error, the clustering procedure consists in:

- Transforming all demographic attributes to represent the difference of each data point and their corresponding national value (for the relevant year).
- Using the HDBSCAN [Campello et al., 2013], a density-based hierarchical clustering algorithm optimises the number of clusters and assignment

This results in 3 distinct clusters of electorates. When presented in a map, it is possible to obtain something similar to figure 7 for 2016.

This mapping seems to align with popular political knowledge, where:

- **cluster 0** seems to mostly contain electorates located in the inner cities, especially in Sydney and Melbourne. These areas tend to be more affluent, either "established" or "gentrified" suburbs. Notably, it also contains the

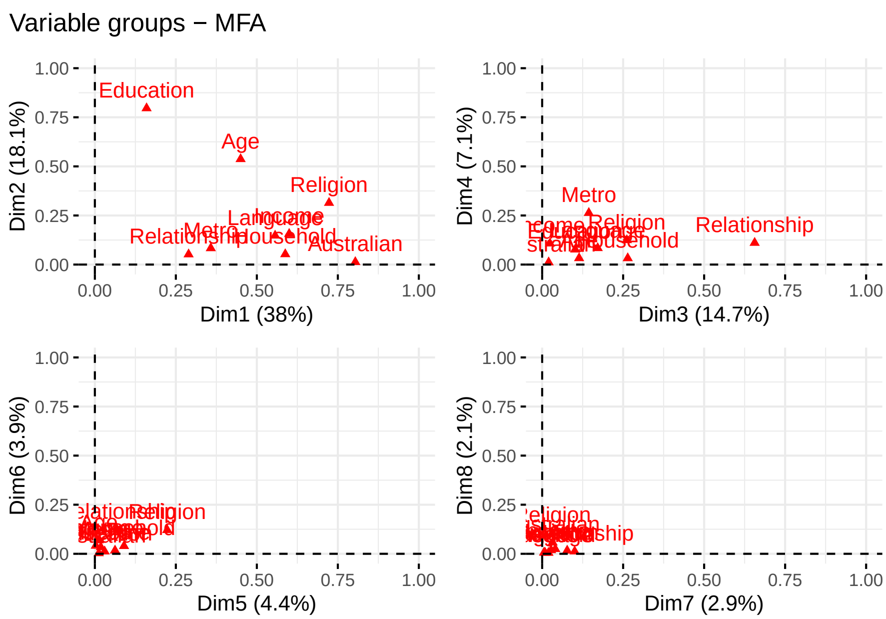



Figure 6: Group plots for first 8 dimensions

three northernmost, remote electorates.

- **cluster 1** comprises all regional areas outside state capitals (with the exception of Hobart in Tasmania).
- **cluster 2** largely represents "suburbia". It it also more prevalent in Brisbane and Perth compared when comparing capital cities.

Revisiting the demographic attributescan help to understand how this clusters differ from each other. A selection of those variables is presented in figure 8.

# 4 Method

Having obtained the data, it modelling time! Starting from a simple point, research in question can be expressed by equation (1).

$$\mathbf{Y} = f(\mathbf{X}) \tag{1}$$

where $\mathbf{Y}$ represents a vector with primary voting for an electorate, and $\mathbf{X}$ represents the vector of respective demographic attributes.

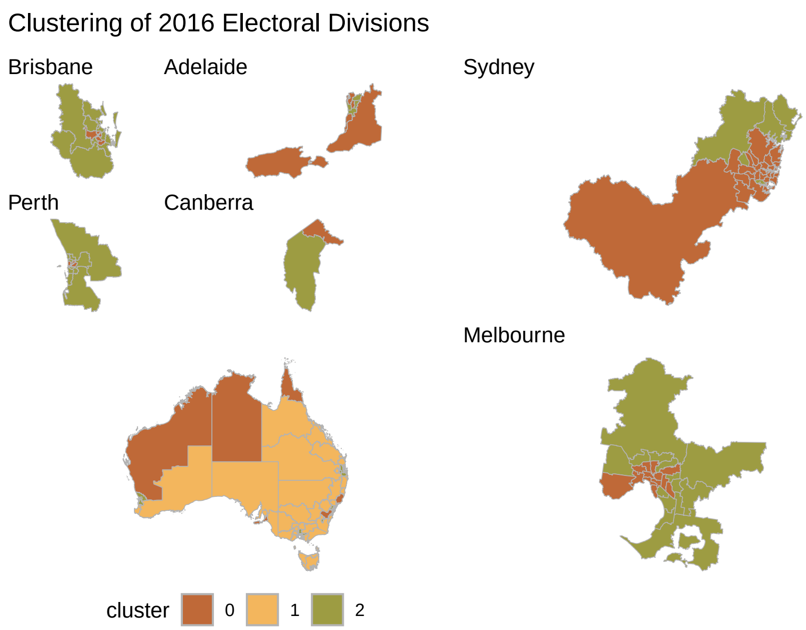



Figure 7: Clusters for 2016 Election

Considering the problem in question, this model does not consider other factors that may influence voting. These factors may difficult to quantify as they may related to a myriad of factors including the state of the economy, foreign affairs, perceptions about the governing party or any party in the election, or the mood of the times. However, it is possible to make the naive assumption that electoral polling aims to capture all those factors - thus it is possible to restate the objective : instead of calculating the primary vote itself, the aim is to determine how demographics influences the voting when compared to a the national values (as captured by polls). Furthermore, it is possible to redefine the problem once more, to aiming to determine how differences in demographic attributes, when compared against a national baseline makes primary voting in an electoral to differ from the national measure. This is states in (10), which allow accounts to error attributable to polling error and demographic drift between census and election.

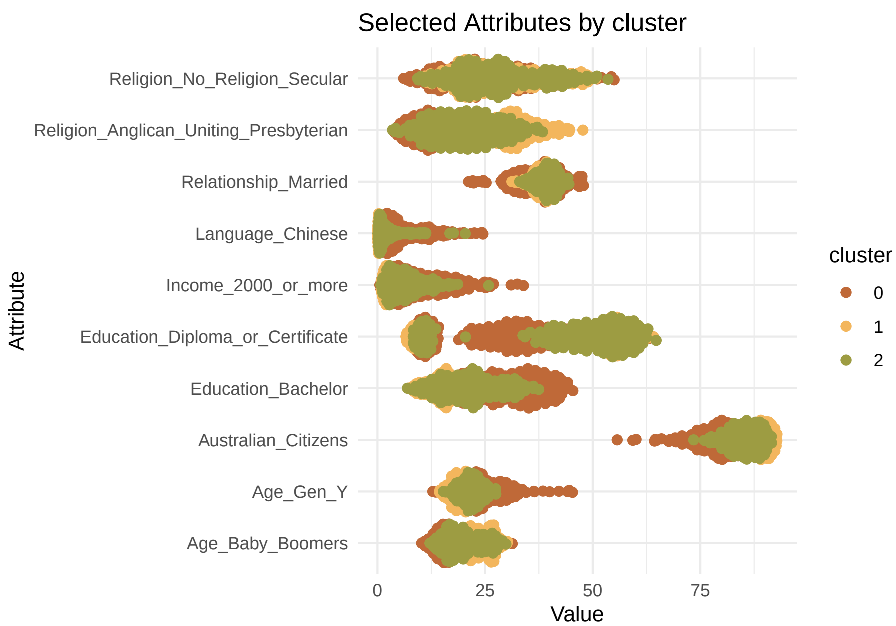



Figure 8: Selected attributes, coloured by cluster.

$$
\begin{aligned}
\mathbf{Y} &= f(\mathbf{X}) + \epsilon && (2)\\
& && (3)\\
\text{where} & && (4)\\
\mathbf{Y} &= \mathbf{Y} - \mathbf{Yp} && (5)\\
\mathbf{X} &= \mathbf{X} - \mathbf{Xn} && (6)\\
& && (7)\\
\text{and} & && (8)\\
\mathbf{Yp} &: \text{Primary voting polling results} && (9)\\
\mathbf{Xn} &: \text{Demographic values at national level} && (10)
\end{aligned}
$$

This second iteration does not take into account that in different electorates, different demographic attributes may have a different effect on primary vote. For instance, in more progressive areas a higher proportion of younger people may result in better results for left leaning parties, whereas in other areas younger people may be more conservative and therefore the youth-left leaning effect may be attenuated. Thus, a refinement on the model needs to acknowledge that not all electorates are made equal - this is presented in equation (16).

$$
\begin{aligned}
\mathbf{Y_i} &= f_i(\mathbf{X_i}) + \epsilon && (11)\\
\text{where} & && (12)\\
\mathbf{Y} &= \mathbf{Y} - \mathbf{Yp_i} && (13)\\
\mathbf{X} &= \mathbf{X} - \mathbf{Xn_i} && (14)\\
& && (15)\\
& && (16)
\end{aligned}
$$

Please note that in this $i$ represents a particular grouping of electorate and for each group predictors can be different, as different attributes may have different impact.

In term of choosing appropriate $f_i$, it would depend on the objective of the model. Considering the large number of predictors and requirements on interpretability and accuracy this could be a complex task. In this particular case, the focus is on understanding the factors that influence vote rather than produce accurate electoral predictions (which is attempted nevertheless), for which the use of **regularised regression** models is an appropriate choice. Thus the task involves using regularised regression to find the coefficients for the set of formulas represented by equation (17)

$$
\begin{pmatrix} y_{i1} \\ y_{i2} \\ \ldots \\ y_{in} \end{pmatrix} = \begin{pmatrix} \beta_{i11} & \beta_{i12} & \ldots & \beta_{i1m} \\ \beta_{i21} & \beta_{i22} & \ldots & \beta_{i2m} \\ \ldots & \ldots & \ldots & \ldots \\ \beta_{in1} & \beta_{in2} & \ldots & \beta_{inm} \end{pmatrix} \begin{pmatrix} x_{i1} \\ x_{i2} \\ \ldots \\ x_{im} \end{pmatrix} + \begin{pmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \ldots \\ \epsilon_{in} \end{pmatrix} \quad (17)
$$

A complication of this approach is that requires to separate the electorates in different segments and potentially have a separate method to place them if assignment is not provided. In this case the previously identified clusters seem a good choice a multiple step model is needed, comprised of:

- A classification model to identify belonging to cluster of similar electorates.
- A regularised regression model to predict difference from average vote.

# 5 Fitting and analysing a model

As mentioned in the previous section, this requires fitting both a classification and regularised regression model.

## 5.1 Cluster classification

Although in section XXX, clusters were obtained using HDBSCAN, which can be used to map new data points into the existing clusters, a different approach has been taken - to “reverse engineer” the clusters by training a classification model. The intent behind this is to leverage the trained model to identify the main contributors to the classification.

Different models where tried, starting with a basic tree partitioning. After trial and error, **random forest** was the chosen algorithm. The model training with:

- Census data from 2007 to 2016 (mirroring elections between 2006 to 2016) was used from training and testing.
- Values for demographic attributes where centred around the overall percentage for said attribute, for the respective cluster.
- clusters previously obtained with HDBSCAN were used as the response.
- Since year has been “discounted”, all values will considered one pool to divide, i.e. year will be ignored. An assumption has been made that the period in question is short enough to drastically affect the clustering model. If demographic values change - cluster assignment (for instance because of re-distribution), the effect is similar to being a different electorate.

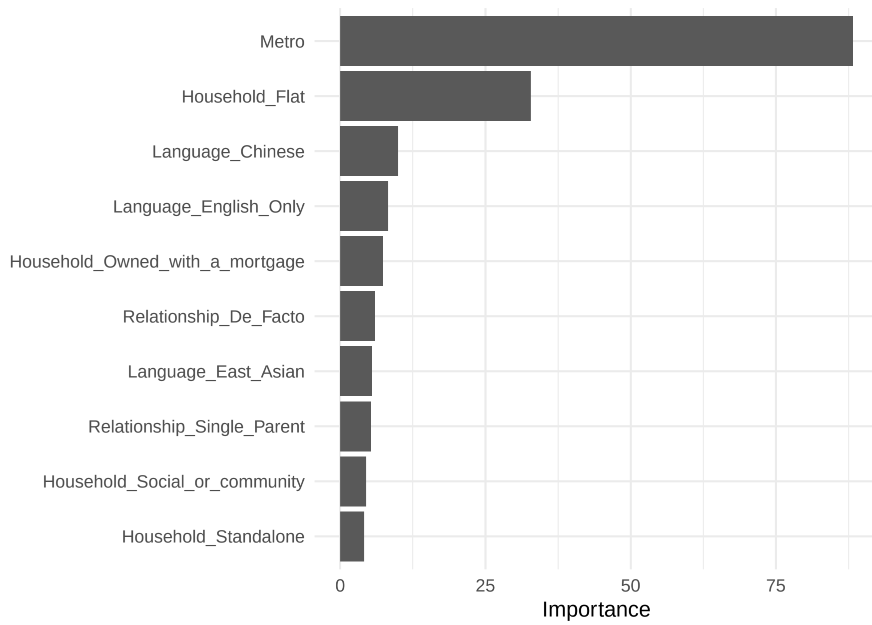
The initial fitting produces the results presented in table XXX

Table 4: First Model

| .metric | .estimate |
|---|---|
| accuracy | 0.8333333 |
| roc_auc | 0.9620648 |

Table 5: First Model

| cluster | FALSE | TRUE |
|---|---|---|
| 0 | 0.2258065 | 0.7741935 |
| 1 | | 1.0000000 |
| 2 | 0.2424242 | 0.7575758 |



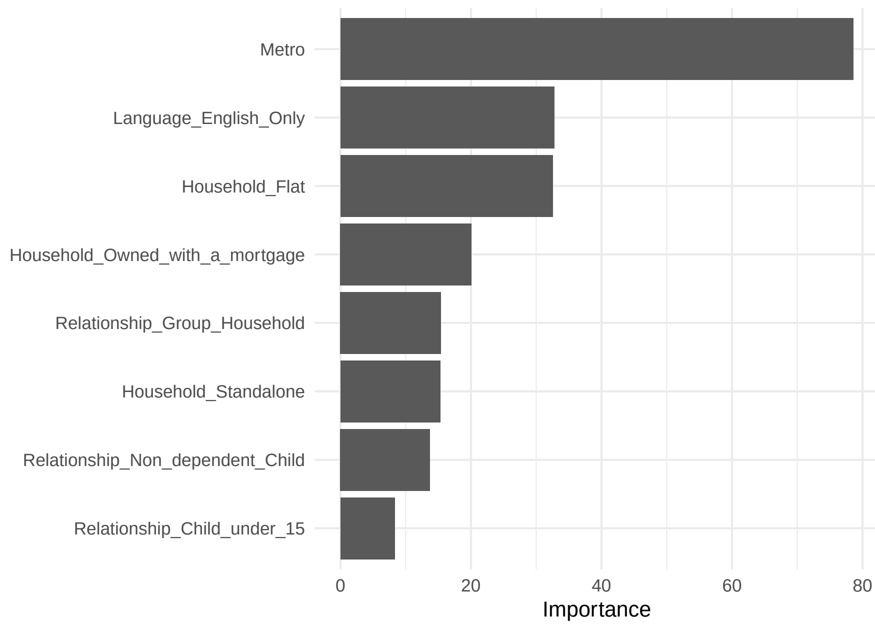

From the chart above, it is possible to see that only a handful of variables have a significant contribution to the cluster selection. Aiming for simplification , a random forest model with reduced variables was also trained , achieving similar results in accuracy and variable importance.

Table 6: First Model

| .metric | .estimate |
|---|---|
| accuracy | 0.8666667 |
| roc_auc | 0.9531230 |

Table 7: First Model

| cluster | FALSE | TRUE |
|---|---|---|
| 0 | 0.2258065 | 0.7741935 |
| 1 | | 1.0000000 |
| 2 | 0.2424242 | 0.7575758 |




Looking at variable importance , it is possible to appreciate that cluster placement can be driven:

Location in a large metropolitan area or the regions. Population density, (type of household) Life stage (relationship) Wealth (type of household ownership) Multicultural make up of the area - first and second generation migrants are more likely to be bilingual - thus the proportion of monolingual people is a proxy variable for this.

This picture fits with the media narrative about differences in the electorate (quote).

## 5.2 Regularised regression

Due to the large number of variables, the first step is to see if it is possible to identify which factor may influence. For this, a Lasso regression was conducted with the sole intent of variable selection. Then an elastic net was fitted, with the goal to optimise the root square mean error (RMSE). This process was done separately for each cluster. Although precision is not a key objective of this exercise, table 8 presents the best RMSE result per cluster, alongside the selected tuning parameters.

Table 8: Best Results for each cluster

| Cluster | □ | λ | RMSE | | | | |
|---|---|---|---|---|---|---|---|
| | | | Overall | GRN | ALP | COAL | Other |
| 0 | 0.5416 | 0.5191 | 5.9363 | 5.7787 | 5.9668 | 7.4913 | 3.9802 |
| 1 | 0.9976 | 1.5452 | 5.9408 | 2.8817 | 6.1343 | 6.8756 | 6.9258 |
| 2 | 0.8408 | 0.5437 | 4.8961 | 1.6438 | 5.4578 | 6.0699 | 5.1531 |

However, the main objective is to undestand the coefficients for each covariate, which are presented in figure 9.

It is worth noticing that some of the selected covariates may not relevant in all electorates, by account of their small absolute various or being relative uniform across the segment . For this reason, the covariates in figure 9 have been ordered by their respective variance - when assessing their overall effect / relevance this must also be taken into account.

When looking at each cluster, it is possible to summarise the different demographic effects as follows:

- In **cluster 0** (mostly inner metropolitan areas) political divides are drawn across wealth, religiosity (i.e. values) and generational lines.
  - In these areas, coalition vote is associated with higher percentages of followers of Anglican, Uniting and Presbyterian churches, people on higher income and Baby Boomers.
  - Labor vote is turn driven by followers of the Catholic Church (partially a reflection of the historic association between the Australian Catholic Church and the labour movement, and Irish and Italian migration) and Millennials. There is some association with less-advantaged population by social and community housing.
  - Green vote is also driven by Millennials, but unlike Labor there is positive association with higher income groups. Green votes are also related to irreligiosity o secular population groups.
- In **cluster 1** (regional areas, including midsize cities and rural areas), demographic variance is smaller. However when it happens, it follows a different pattern from the main cities.
  - In this area Coalition vote has also a positive association with religiosity - this is not disimilar to cluster 1, especially when considering that Anglicanism/Presbytiarism/Unitiarism are the largest religious groups in the area) . However, a key difference with the cities is that in case higher wealth groups have a negative association with Coalition vote.

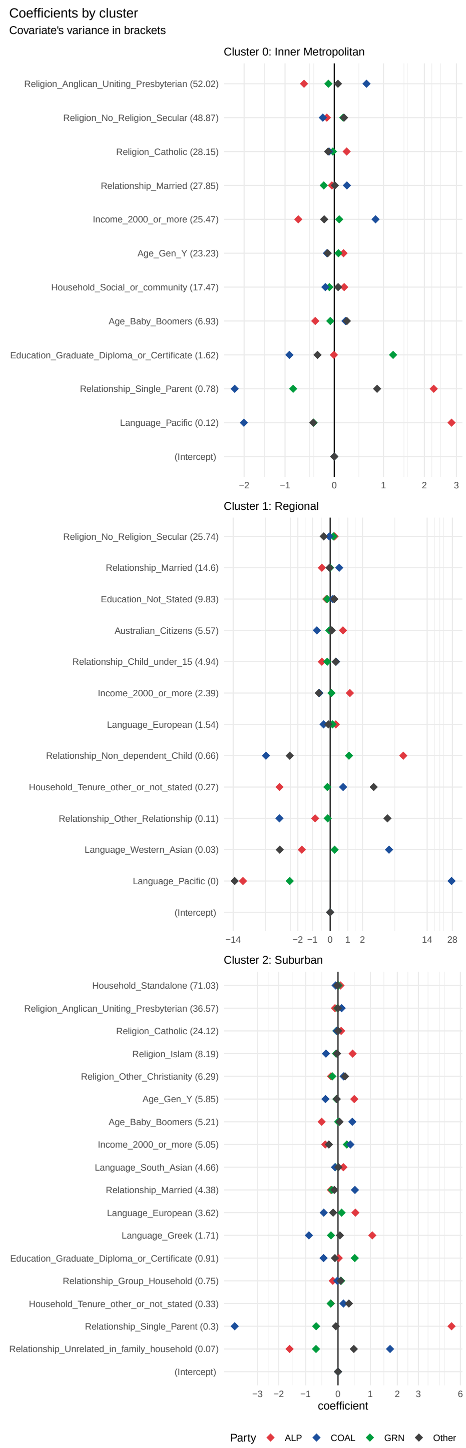



Figure 9: Resulting coefficients per cluster

- Labor vote in these areas is driven by a larger roportion of Australian citizens and higher income voters.
  - Overall, it seems there are no demographic factors influencing Green votes in these areas.
  - Interestingly, age does not rank as a variable of importance.

- As expected, **cluster 2** (metropolitian suburbia) , shares some traits with their inner city counterparts, showing the same associations along religious, age and wealth lines. However, there are larger number of predictors associated to the multicultural makeup of the electorates. Those covariates tend to have a positive effect on Labor vote and negative influence on Coalition and Green voting. This difference is interesting specially considering inner city areas are as multicultural as the suburbs.

# 6 Results

## 6.1 Forecasting the 2022 Federal Election

The previously fitted model can be used to attempt to retroactively forecast the 2022 Federal Election. Through this process it is possible to illustrate the model's strength and shortcomings at capturing how demographic factors succeed and fail to capture the change in voting patterns.

This exercise uses the results from the 2021 Census of Population and Housing. The base voting percentages are taken from the last Newspoll prior to the election [Benson, 2023] . Newspoll is usually considered a good predictor of the Australian election. The values are shown in table 9. Please note these values are national - but since there is no data at cluster level these are anyway.

Table 9: Newspoll Primary Vote Forecast, 20 May 2022

| Party | Forecast |
|---|---|
| COAL | 35% |
| ALP | 36% |
| GRN | 12% |
| Other | 17% |

The first step in the forecasting process is to map the electorates into the 3 clusters. The result is presented in figure 10.

After clustering, the regression models have been used to calculate a predicted outcome. Results have been transformed back to absolute values and then compared against actual results. This is presented in figure 11. The RMSE values are presened in table 10.

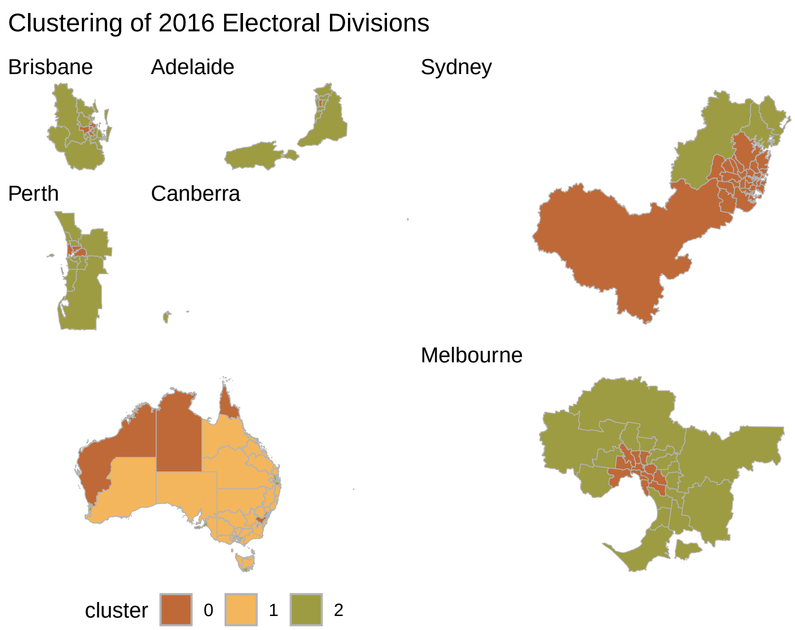



Figure 10: Cluster in 2022 Election

Table 10: RMSE Values for 2022 Forecast

| cluster | Overall | GRN | ALP | COAL | Other |
|---|---|---|---|---|---|
| 0 | 10.56 | 9.48 | 9.42 | 8.43 | 15.04 |
| 1 | 10.37 | 4.16 | 10.21 | 14.35 | 10.11 |
| 2 | 8.26 | 3.16 | 7.51 | 6.05 | 13.40 |

As expected, the results fail to adequately forecast election results, especially when it comes to Other parties and independents. However, it can be used as a tool to analyse the vote dynamics.

## 6.2 The Teal Wave

A particular phenomenon of the last election consisted in the so called “Teal Wave”, where centrist independents campaigned in traditional Coalition electorates. Most of these electorates are located in inner city, wealthy areas of Melbourne and Sydney, where voters have consistently voted Coalition since the Australian Federation. Right leaning voters in these areas are perceived as moderate, socially liberal (“little-l liberals”) who were dissatisfied with a

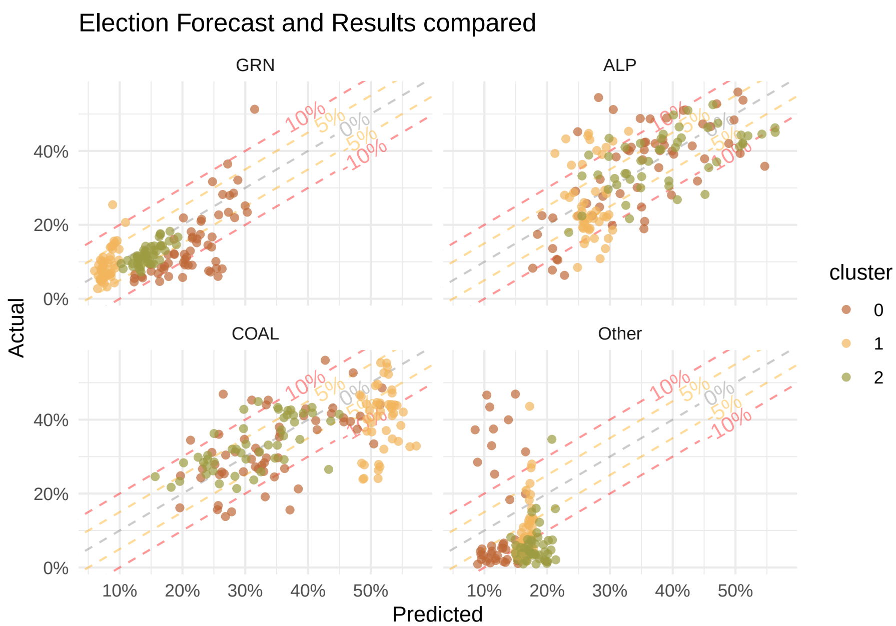



Figure 11: Comparison between prediction and election results

perceived conservative turn in Coalition politics. Teal candidates managed to unseat incumbent MPs - did they in effect capture the dissatisfied Coalition base? The results and predictions for 4 cases are presented in figure 12.

The answer is this case seems to indicate that dwindling Coalition vote is not entirely related to a new teal offering. Observing figure 13, demographic changes in figure Liberal (Coalition) voting decline aligns with increase in higher income population and secular views - trend that predates the emergence of the teal movement. Thus, it is valid to assume that these new candidates growth is largely attributable to the same growing group of people that voted Green in the past, plus Labor voters.

## 6.3 The Green Wave

Another feature of the past election was the increase in the number of Green Party MPs. In addition to the division of Melbourne, green candidates also won the seats of Griffith and Ryan in Brisbane. Again, do these victories have a demographic driver? Are there any differences between these electorates and contigous divisions, and between them and other electorates where the Green have been strong contender? 14 shows the prediction the latest and historic election results. figure 15 presents selected demographic attributes for those areas.

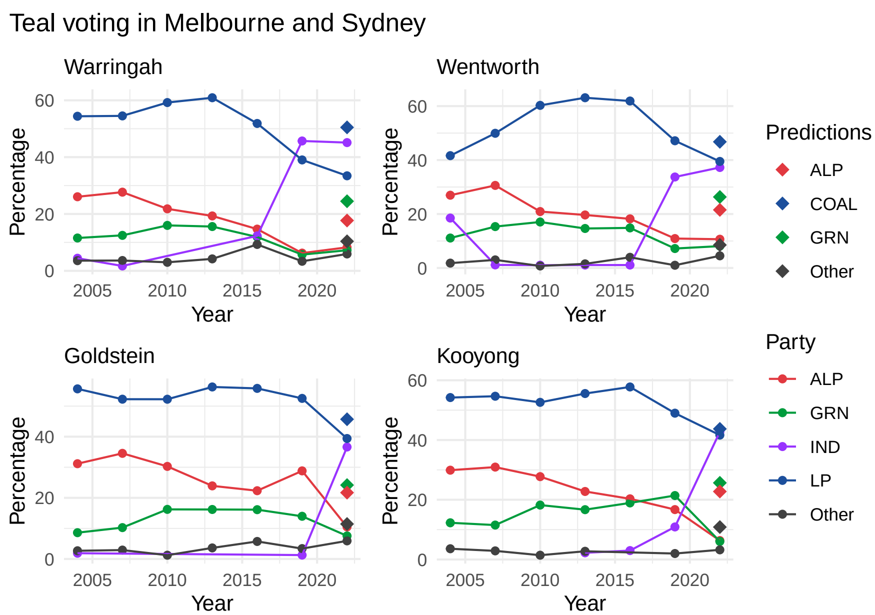



Figure 12: Example 1: Teal Wave

From the carts, three is clear correlation between demographic attributes and primary vote. This is clear for Giffith and Ryan, but also for Brisbane - where Greens didn't win the seat. In wills, this effect is more difficult to visuall inspect in Wills, however actual results lies very closed to the predictions.

## 6.4 The Changing face of suburbia

# References

Simon Benson. Newspoll: Labor lead over coalition narrows, 2023. URL https://www.theaustralian.com.au/nation/politics/newspoll-labor-lead-over-coalition-narrows/news-story/937dbfe8479e9380d93da4121f63c09d.

Nicholas Biddle and Ian McAllister. Explaining the 2022 australian federal election result. Technical report, 06 2022. URL https://apo.org.au/node/318286.

Ricardo J. G. B. Campello, Davoud Moulavi, and Joerg Sander. Density-based clustering based on hierarchical density estimates. Lecture Notes in Computer Science, pages 160–172, Berlin, Heidelberg, 2013. Springer. doi: 10.1007/978-3-642-37456-2_14.

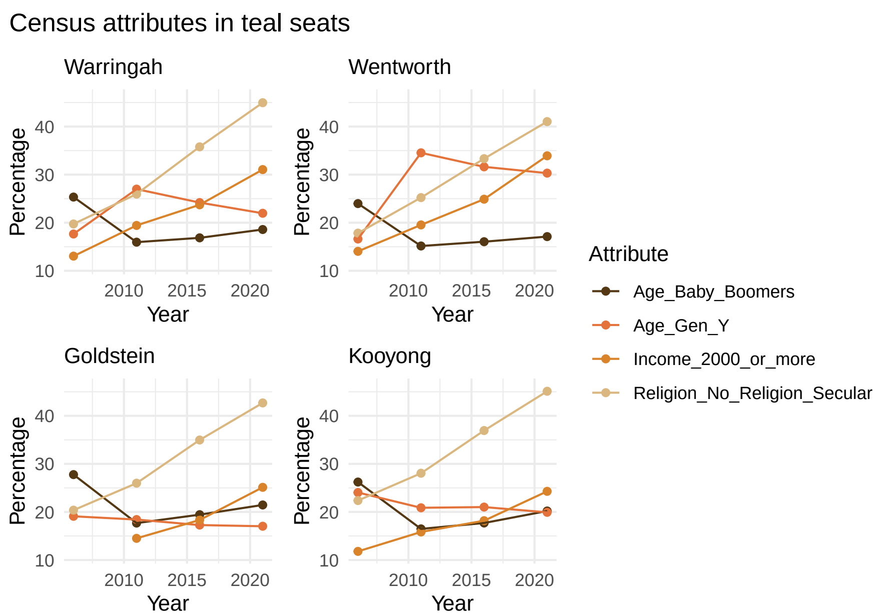



Figure 13: Selected demographics for teal seats

Australian Electoral Commission. Website, 2023a. URL https://www.aec.gov.au/.

Australian Electoral Commission. Tally room archive, 2023b. URL https://results.aec.gov.au/.

Brigitte Escofier and Jérome Pagès. *Analyses factorielles simples et multiples. Objectifs méthodes et interprétation*. Sciences Sup. Dunod, 2008. URL https://hal.science/hal-00382085.

Andrew Jakubowicz and Christina Ho. Was there an 'ethnic vote' in the 2019 election and did it make a difference?, a. URL http://theconversation.com/was-there-an-ethnic-vote-in-the-2019-election-and-did-it-make-a-difference-117911.

Andrew Jakubowicz and Christina Ho. Was there an 'ethnic vote' in the 2019 election and did it make a difference?, b. URL http://theconversation.com/was-there-an-ethnic-vote-in-the-2019-election-and-did-it-make-a-difference-117911.

Australian Bureau of Statistics. Website, 2023a. URL https://abs.gov.au/.

Australian Bureau of Statistics. Census data packa, 2023b. URL https://abs.gov.au/census/find-census-data/datapacks/.

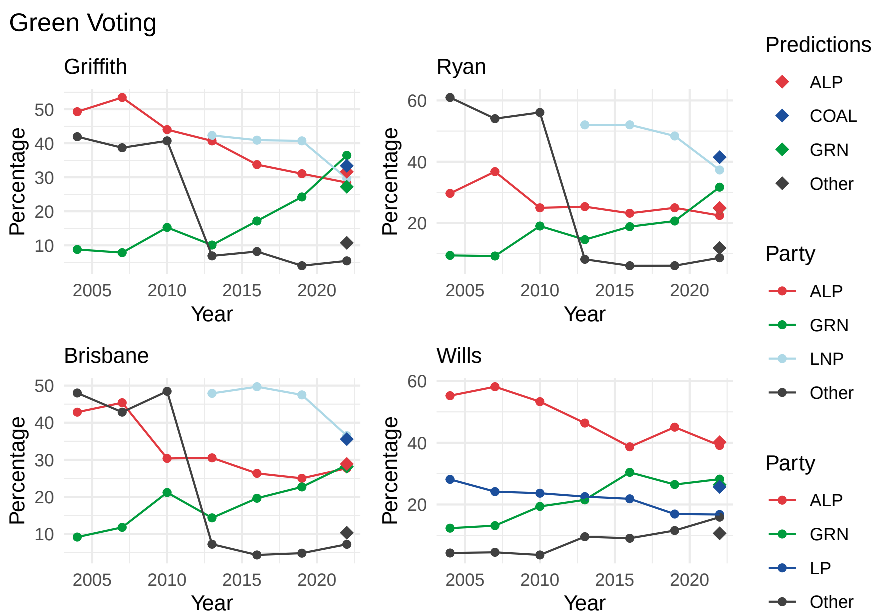



Figure 14: Green Voting

Commonwealth Parliament. Voting patterns by generation. URL https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/FlagPost/2022/April/Voting_patterns_by_generation. Archive Location: Australia Last Modified: 2022-04-29 Publisher: corporateName=Commonwealth Parliament; address=Parliament House, Canberra, ACT, 2600; contact=+61 2 6277 7111.

Carlos Yáñez Santibáñez. *auscensus: Access Australian Census Data (2006-2021)*, 2023a. URL https://carlosyanez.github.io/auscensus/. R package version 0.0.1.0008.

Carlos Yáñez Santibáñez. *auspol: Australian Federal Election Results (2004-2022)*, 2023b. URL https://carlosyanez.github.io/auspol/. R package version 0.0.1.0000.

Carlos Yáñez Santibáñez. *aussiemaps: Maps of Australia*, 2023c. URL https://carlosyanez.github.io/aussiemaps/. R package version 0.2.0.0013.

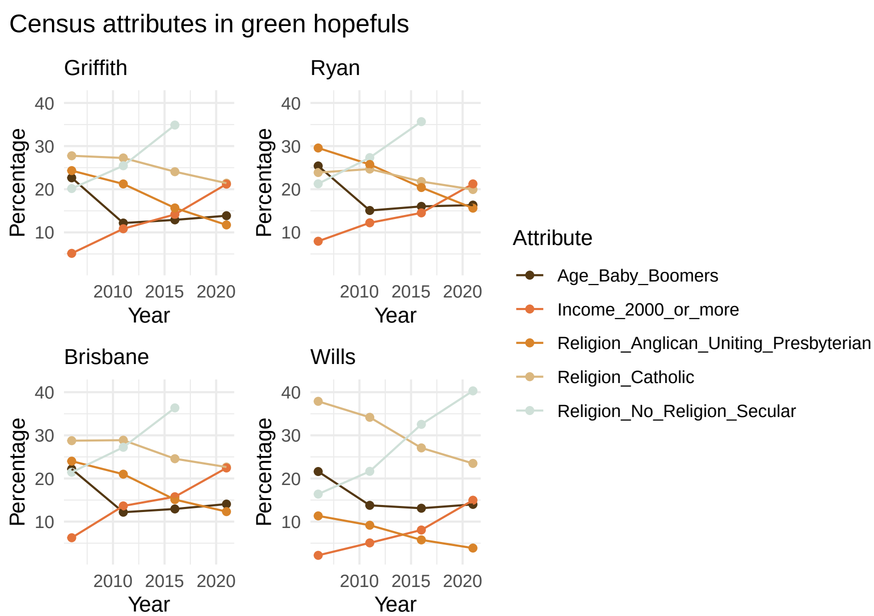



Figure 15: Demographics in Green seats

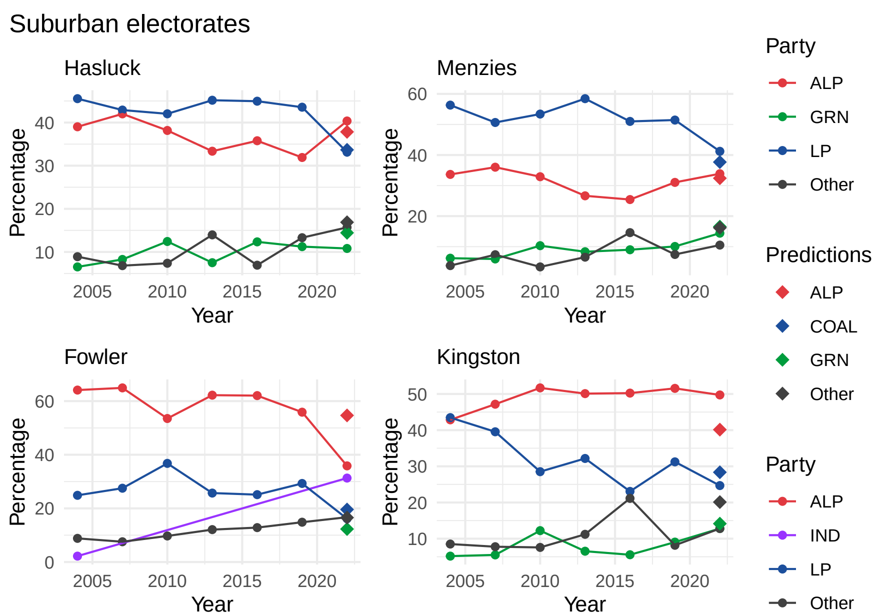



Figure 16: Green Voting

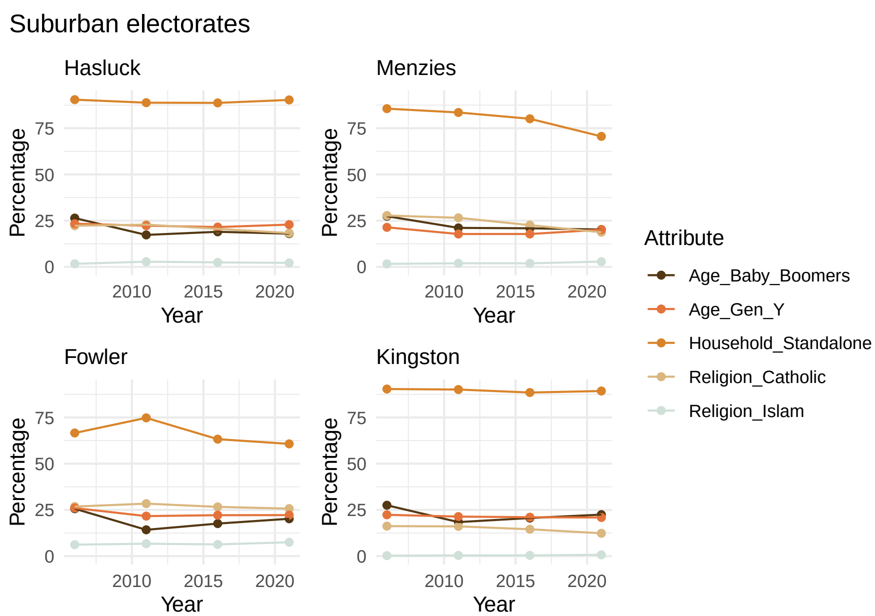



Figure 17: Demographics in Green seats