



# Red & Blue Gerrymandering

Alternatives to (allegedly) unfair districts in Georgia, USA

# Agenda

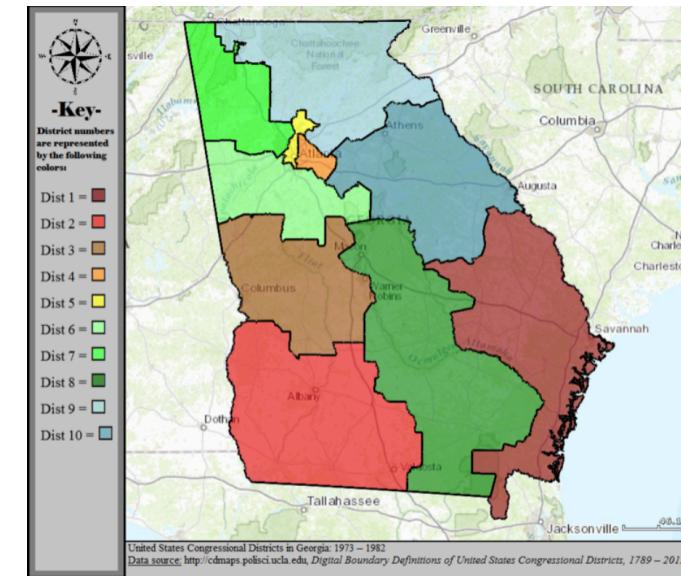
- The Problem
- Context - Gerrymandering in Georgia
- District focussed demographics
- An alternative: County-cluster focussed demographics
- Results
- Conclusion / Next Steps

# The Problem

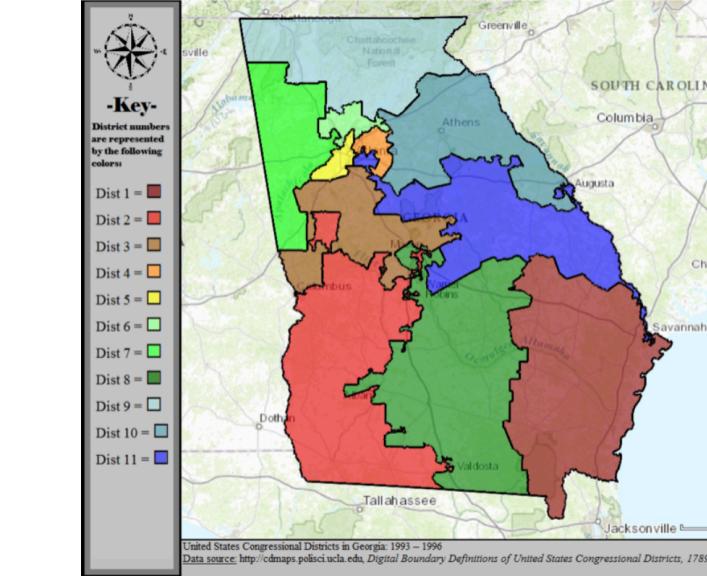
- Amongst the issues affecting US election campaigning, ‘gerrymandering’, the process of drawing district boundaries to achieve favourable political outcomes, is one of the most vexing and arcane.
- The topic is particularly contentious in the US state of Georgia, where court cases have been filed challenging the fairness of Georgia’s redistricting policies.
- A number of years have passed since the last material redistricting (7 years, as interpreted from wikipedia Georgia district charts); voter issues may have changed in that time and issues may no longer be in common within a district.
- Moreover, the next election is approaching (November 2020) and will occur before any further redistricting (which would likely also be gerrymandered) can be conducted.
- *For campaign advisors to be effective in 2020, they will need to reconcile their candidate’s platforms and messaging with current voter issues; such issues should include at minimum demographics-based voter concerns.*

# Context - Gerrymandering in Georgia

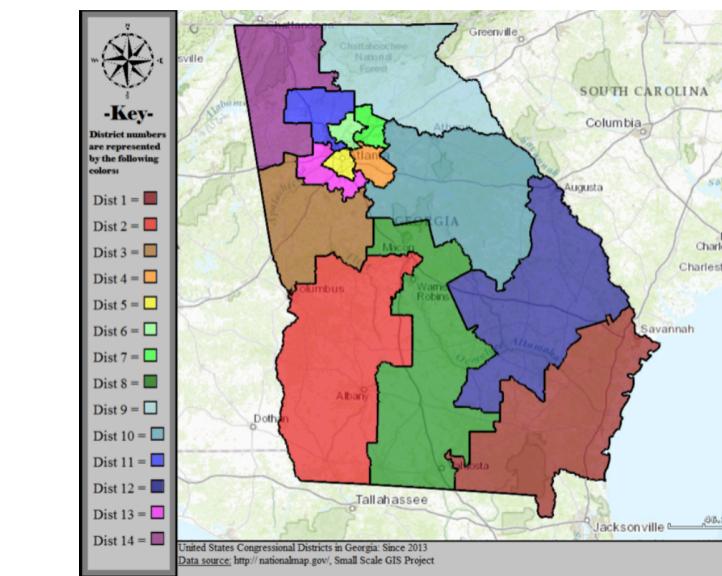
- Gerrymandering establishes unfair political advantage by manipulating US congressional election district boundaries to isolate voters with particular voting patterns.
  - “Districts” define geographical boundaries, with each district within a state being geographically contiguous and having about the same number of state voters.
- Georgia has come under fire due to suspected gerrymandering to suppress minority votes. Odd ‘contiguous’ shapes following redistricting fuel the debate:



1973-1982



1993-1996

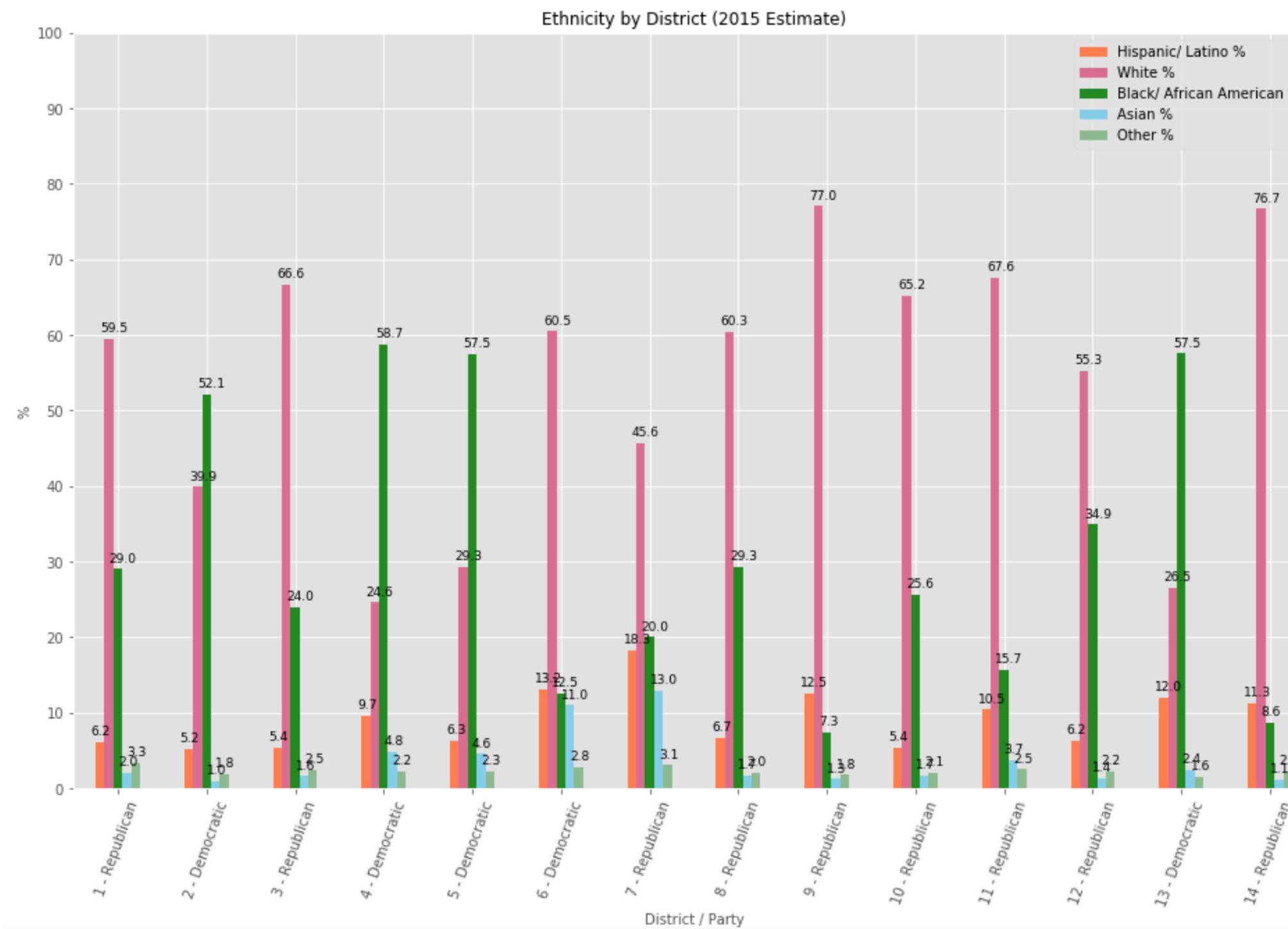


Since 2013

- Given current beneficiaries of gerrymandering and the party currently controlling the process are Republican, the practice is unlikely to change anytime soon.
- Gerrymandering has (allegedly) been conducted in Georgia for many years, although boundaries have not materially changed in the last 7 years. Voters captured within these districts may no longer have issues in common. It is therefore unclear that the current boundaries will yield the same electoral results in 2020.

# District-focussed demographics

- Highlights from 2015 & 2018 US Census Bureau demographic data on Georgia districts

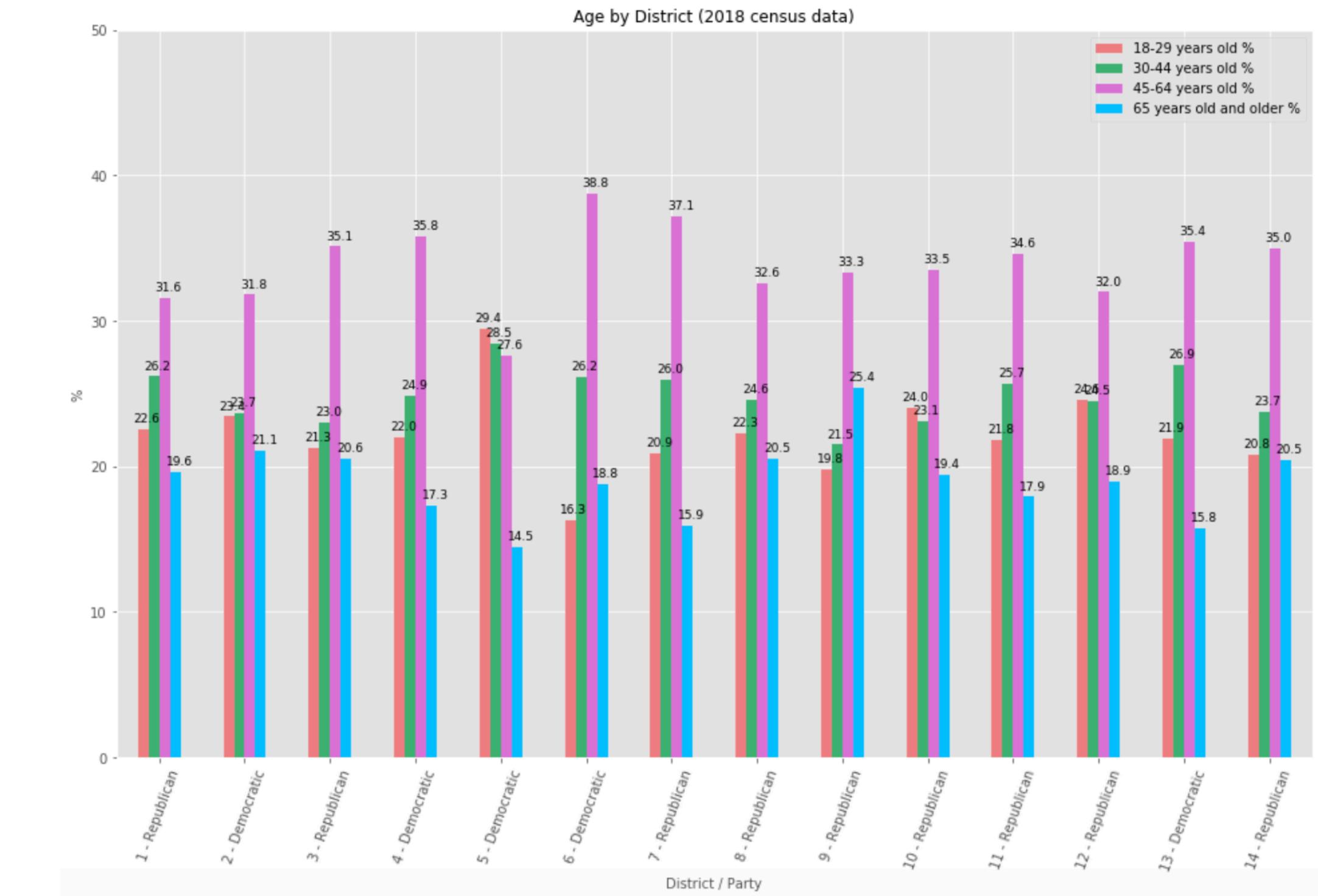


## Ethnicity

- Districts with larger numbers of white voters elected Republican candidates while districts with larger number of black voters elected Democratic candidates.
- The notable exception is District 6, which has a larger proportion of white voters and elected a Democratic candidate.

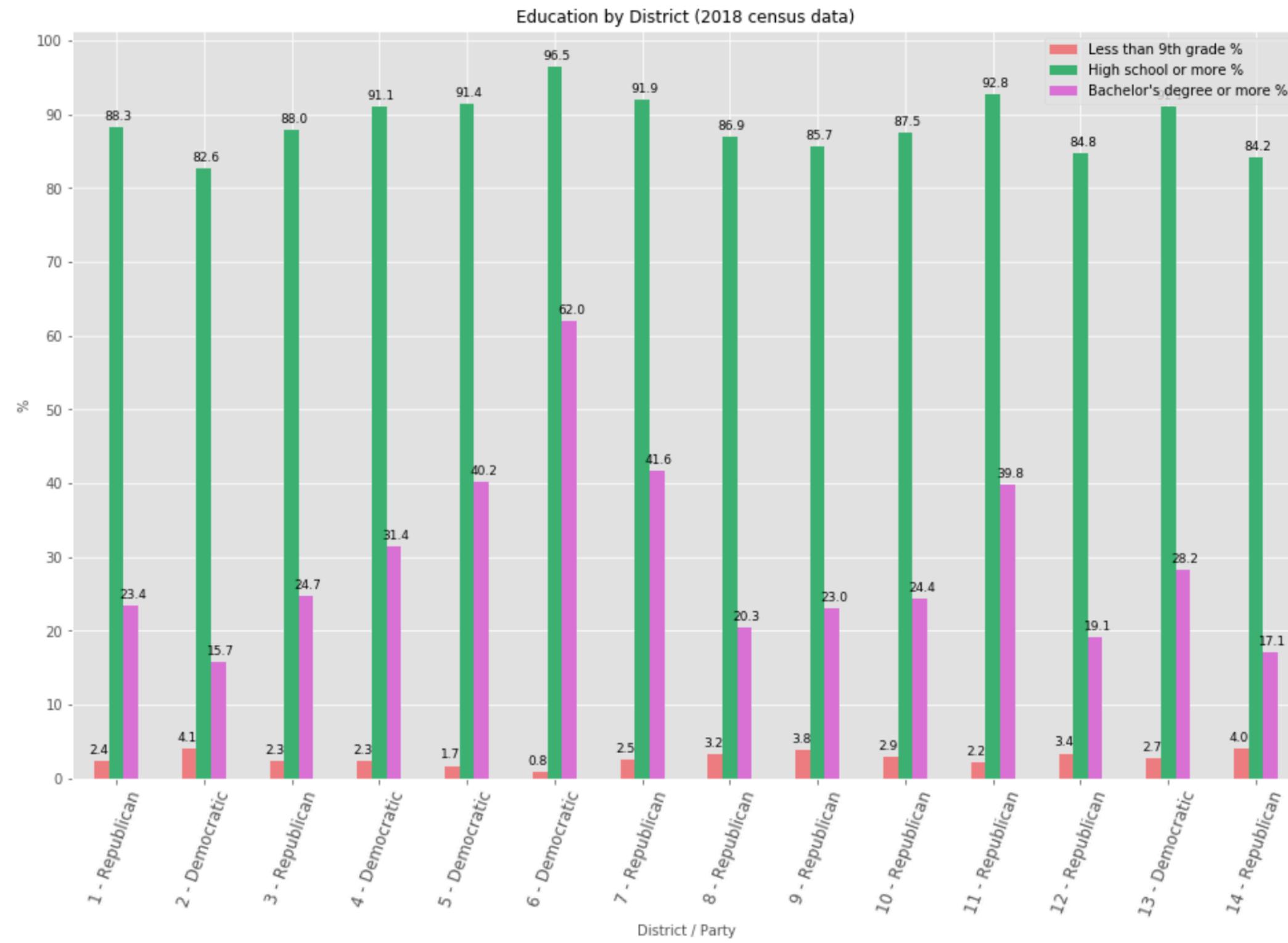
## Age

- District 5 has the youngest voting population
- District 6 has the largest proportion of voting age 45-64 y.o.
- District 9 has the oldest voting population



# District-focussed demographics (cont'd)

- Highlights from 2018 US Census Bureau demographic data on Georgia districts

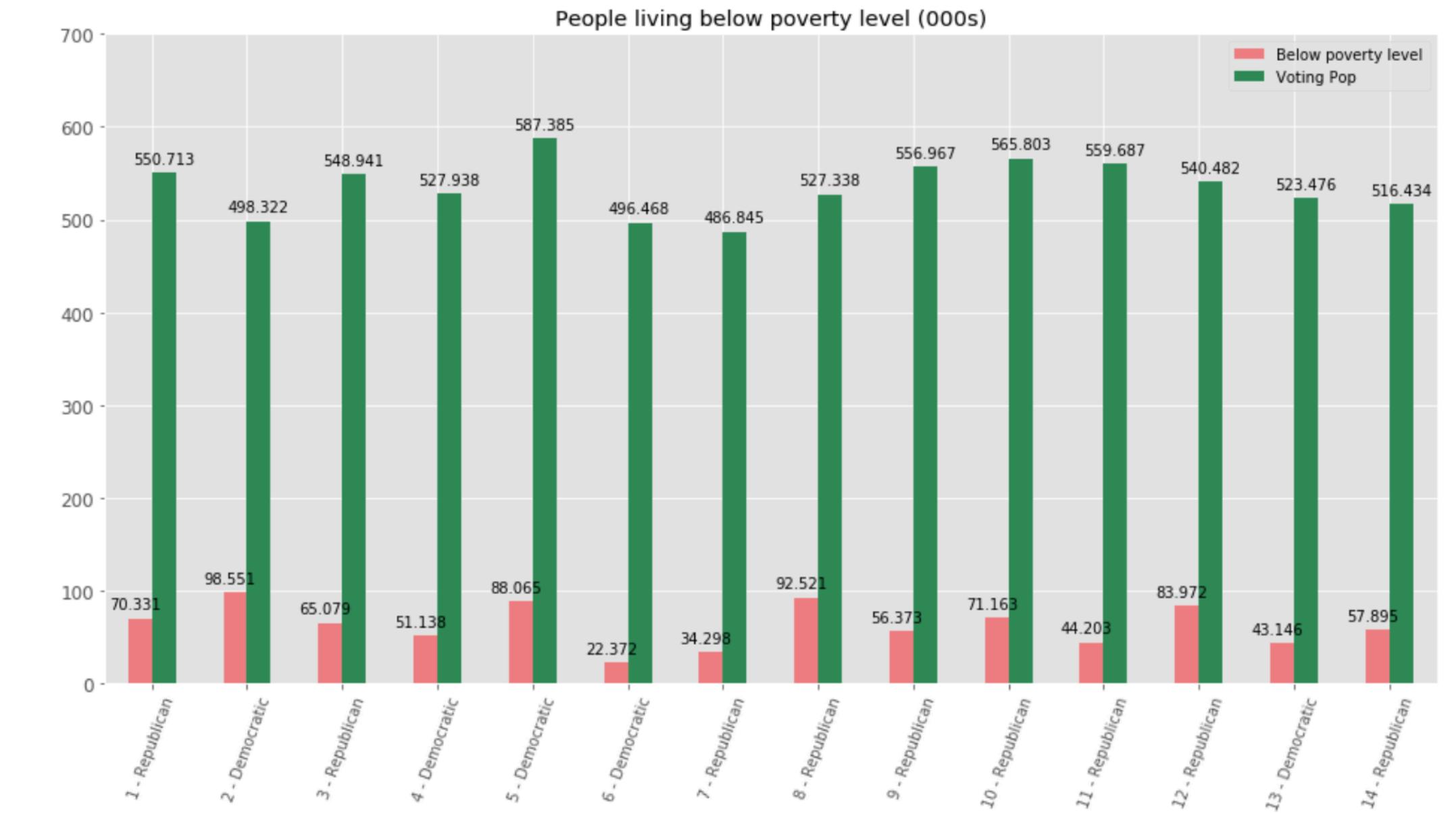


## Education

- District 6 has highest number of college graduates
- District 9 has highest number of people who did not finish 9th grade

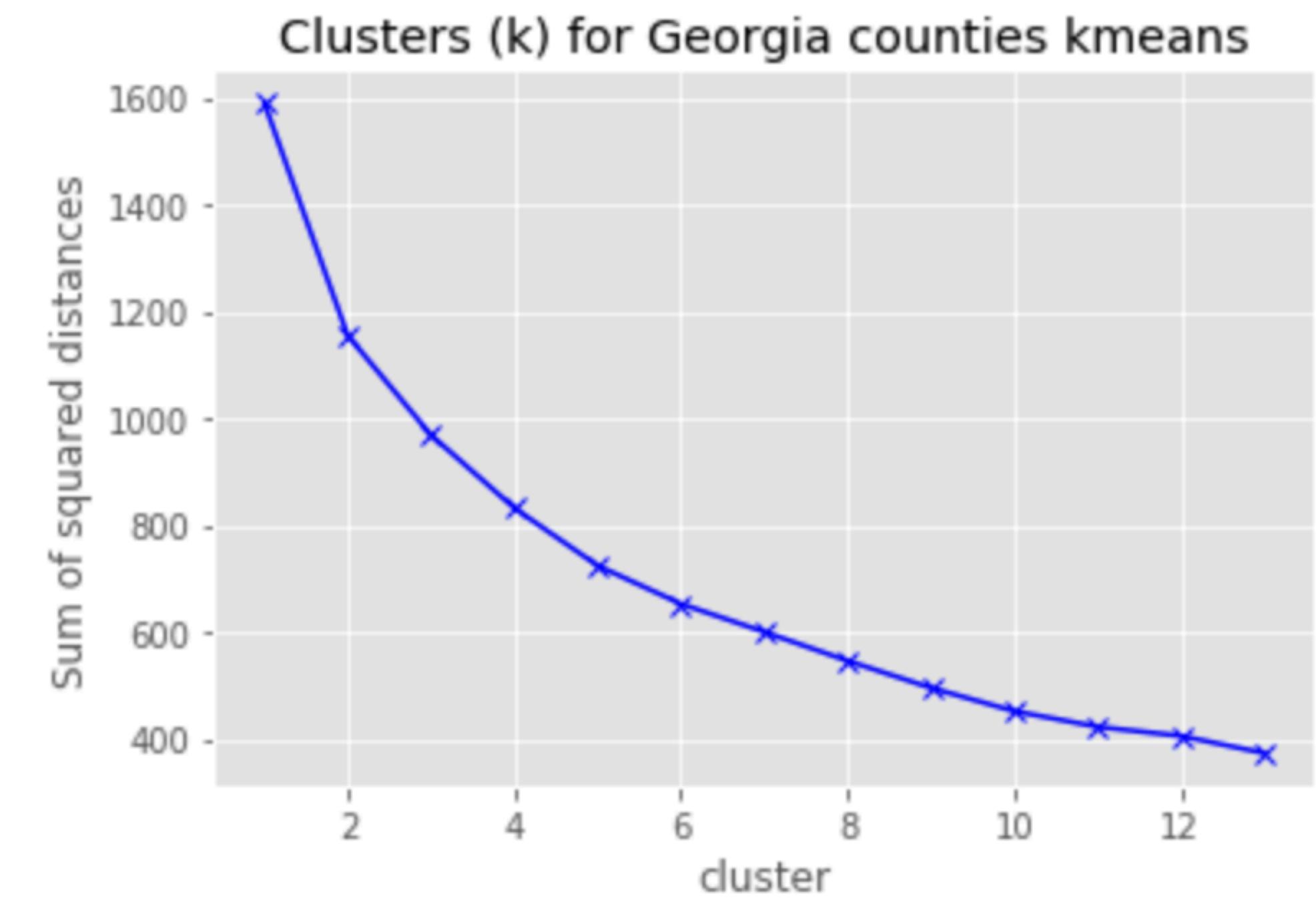
## Poverty

- District 2 has the highest number of people living below the poverty line.
- District 6 has the lowest number

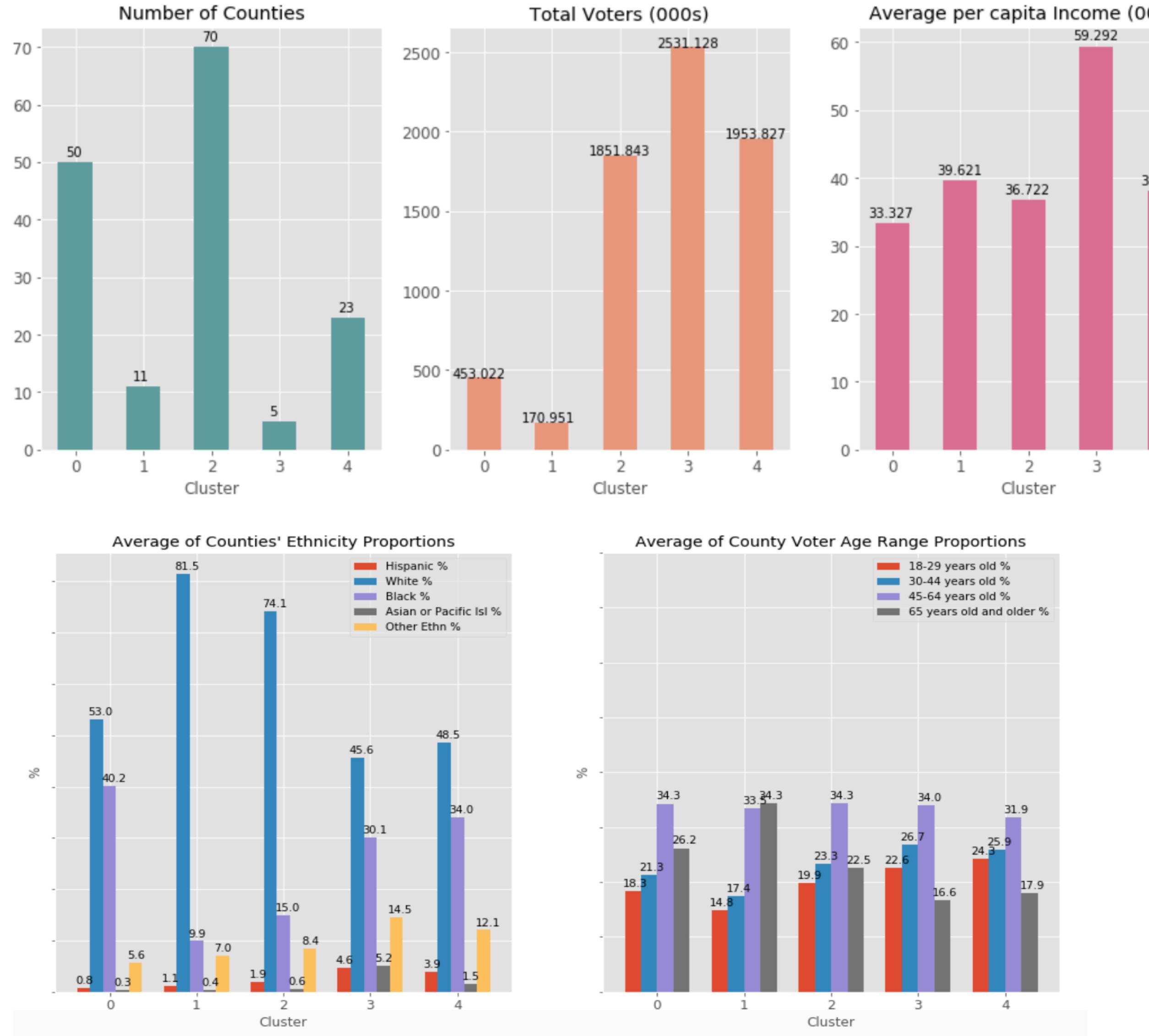


# Alternative: County-focussed Demographics

- To achieve a workable alternative clustering of voters, the K-means machine learning algorithm was used
  - K-means is well-suited to unsupervised segmenting of populations
  - K-means is less computationally-intensive and so is easy to rerun; this allows for more features experimentation
- K-means requires a starting number of clusters ‘k’. The “elbow method” was used to determine the best cluster size however, due to the small resulting size ( $k=2$ ),  $k=5$  was used to allow greater degree of campaigning flexibility.

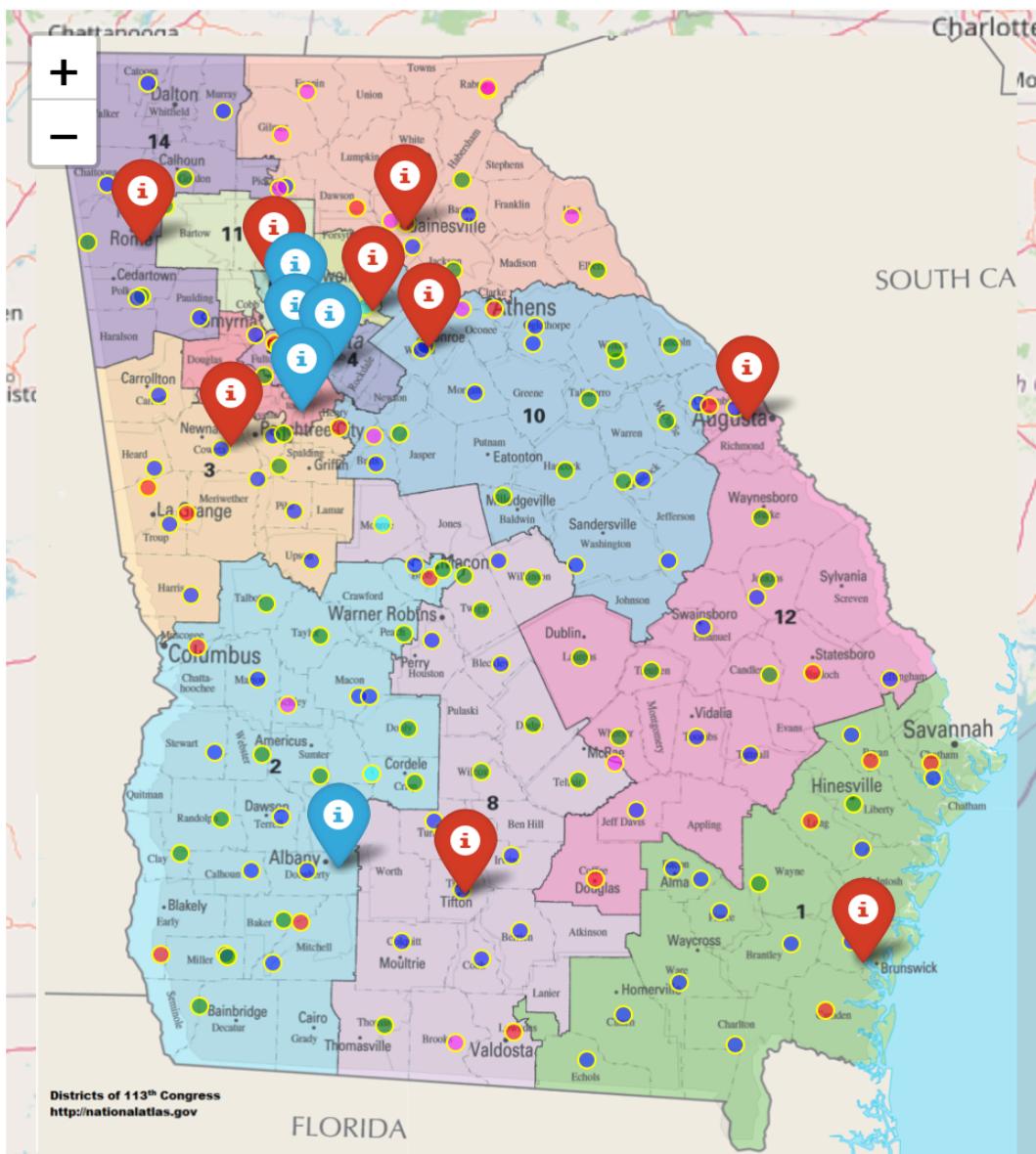
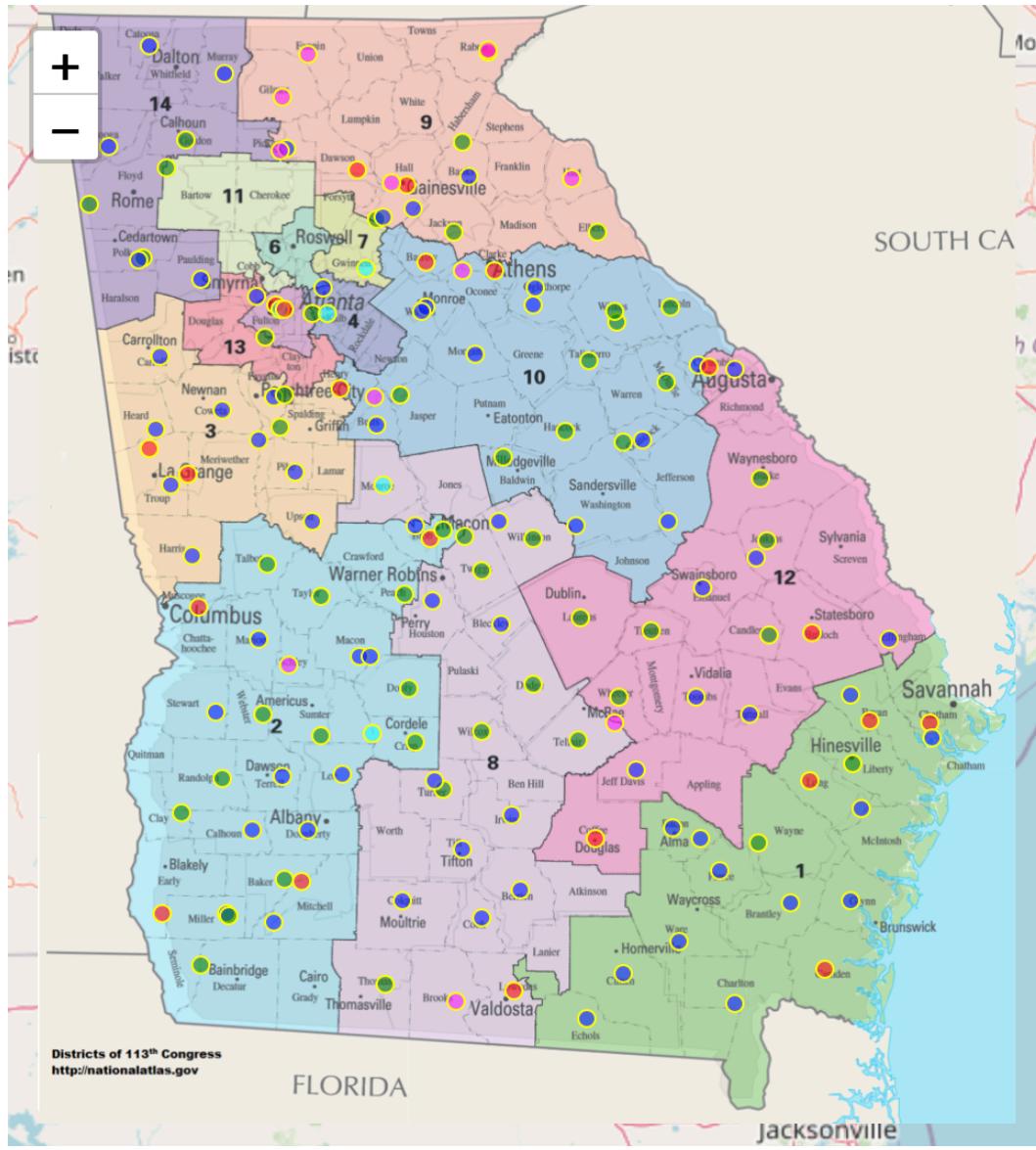


# Alternative: County-focussed Demographics



- **Cluster 0** - has the highest proportion of 'black' voters across clusters (40.2%), lowest combined proportion of 'hispanic', 'asian' and 'other' voter constituents (6.7%), second-lowest number of voters (~453k) across the second-highest number of counties (50, suggesting broad geographic dispersion of these voters). It also has the (marginally) lowest per capita income across clusters (~33.3k).
- **Cluster 1** - has the greatest proportion of 'white' voters relative to other ethnicities (81.5%), marginally oldest voter population (65 years old and older, 34.3%), lowest number of voters (~171k) across the second-lowest number of counties (11).
- **Cluster 2** - has the second greatest proportion of 'white' voters relative to other ethnicities (74.1%), largest number of counties (70) with third largest total number of voters (~1,851.8k).
- **Cluster 3** - has the greatest proportions of 'other', 'asian or pacific islander', and 'hispanic' ethnicities relative to other clusters (24.3%), marginally lowest proportion of 'white' voters relative to other clusters (45.6%) and older voters (65 years old and older, 16.6%). It also has the largest number of voters (~2,531.1k) across lowest number of counties (5) and the highest income by far (59.3k) across clusters.
- **Cluster 4** - has the second greatest proportion of black voters (marginally, 34%), youngest voting population considering 18-29 and 30-44 age ranges (50.2%), with a more distributed age range in general. It has second highest number of voters (~1,953.8k) across 23 counties.

# Alternative: County-focussed Demographics cont'd



District	Party	Dominant Cluster
District 1	Republican	Cluster 2
District 2	Democratic	Cluster 0
District 3	Republican	Cluster 2
District 4	Democratic	Not discernible
District 5	Democratic	Not discernible
District 6	Democratic	Cluster 2*
District 7	Republican	Cluster 3
District 8	Republican	Not discernible
District 9	Republican	Cluster 1
District 10	Republican	Not discernible
District 11	Republican	Cluster 0*
District 12	Republican	Not discernible
District 13	Democratic	Not discernible
District 14	Republican	Not discernible

\* clusters with possible map misalignment

Cluster marker colours

0	1	2	3	4
---	---	---	---	---

- These maps of Georgia depict locations of counties coloured by assigned cluster (see legend below). The bottom map shows District office addresses and colour indicating political party (red = Republican, blue = Democratic).
- County clusters are geographically dispersed and not explicitly aligned to districts, however there are dominant clusters in certain districts (described in middle table).
- District 6 (Democratic) seems comprised of a single Cluster 2 (“blue”) county; campaign targeting by ‘Cluster 2’ issues is probably similar to campaigning by ‘District 6’ issues.
- District 11 also appears to be an anomaly as no county is captured within the district boundary. This is likely due to a combination of map misalignment and geocoder api defaults

# Results

- Demographics-based county clusters, as described by the K-means process, do not cleanly align to district boundaries. This might suggest that, indeed, (allegedly) gerrymandered districts no longer reflect common demographics-based issues.
  - There may be exceptions for example where a district seems comprised of a single county (note that the county-agnostic redistricting process makes this an unsafe observation).
- In terms of prioritising issues based on cluster (and not district), District 2 and District 5 (both Democratic) and District 8 and District 12 (both Republican) had the highest levels of poverty which suggests that poverty level is not necessarily a party vote differentiator.
  - Whilst this is a limited sampling, it suggests that the clusters where income is a significant differentiator such as Cluster 3 (“cyan”) might not be as interesting to prioritise for campaigning where the political message relates to income. *Cluster 3 has the highest population of voters (compressed into the smallest number of counties), therefore, further modelling and analysis (out of scope for this report) would be required before making this cluster a lower priority.*
- Ethnicity may have been a determining factor in previous district-level election results (based on district demographic data).
  - The notable exception is District 6

# Conclusion / Next Steps

- District-level issue campaigning (eg, for state-wide elections) should incorporate a broader set of issues than usually associated with the particular political party in a given district. The cost implications are out of scope for this report.
- State-wide campaign advisors should prioritise issues for Cluster 3 (“cyan”) counties (in the context of the geography of those counties) given its voting population size and less-skewed demographics.
  - Spend less time and money on income/poverty-level- focussed conversation (as troubling an idea as that is)
- These advisors should also prioritise Cluster 2 (“blue”) county issues due to dominance in the most districts, ie, cluster-specific campaigning could also be used for district-level campaigning.
  - This leverage will be useful for District 6 elections (assuming the map misalignment impact is negligible), given its low Democratic victory margin in 2018 and possible Cluster 2 dominance.
- Lastly, Cluster 1 (“magenta”) should not be prioritised for strategic campaigning based on the low voting population and other demographics that are broadly in range with other clusters except for the rather large white ethnicity skew.

## Next Steps

This is cursory, high-level analysis of subject matter that is very complex and highly nuanced; this presentation barely scratches the surface. Campaign managers should incorporate more demographic information, such as gender, in further reclustering exercises. ***Warning: do not underestimate the time it will take to source and prep government data.***