Tege, J

IBM Data Science Capstone Assignment

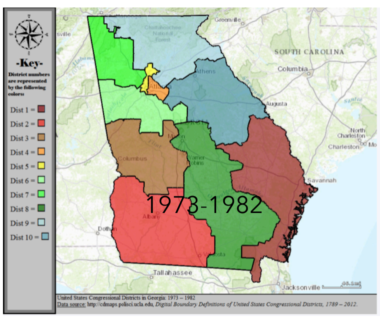
Red & Blue Gerrymandering:

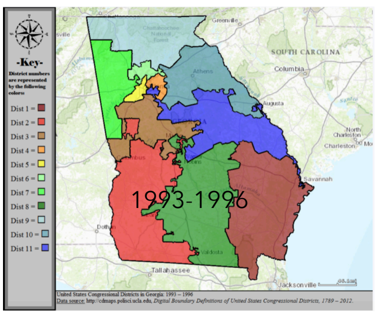
Responding to (alleged) unfair Georgia redistricting

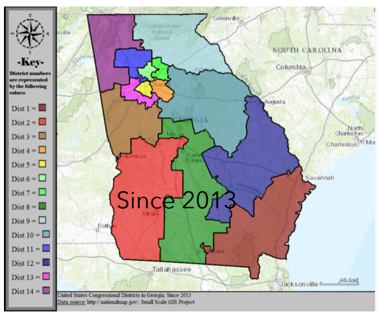
Introduction:

Amongst the issues affecting election campaigning, Gerrymandering is one of the most vexing and arcane. Given the contentious nature of this topic in Georgia (USA), the number of years that have passed since the last material changes (7 years, as interpreted from wikipedia Georgia district charts), and the close results of recent elections, it is important that political campaign management firms (as well as voters) understand alternatives to counter effects of this issue in order to better target candidate marketing strategies during the next Georgia election season.

Gerrymandering is a practice intended to establish an unfair political advantage for a particular party or group by manipulating US congressional election district boundaries. "Districts" define geographical boundaries, with each district within a state being geographically contiguous and having about the same number of state voters. Gerrymandered redistricting would result in engineered clusters of voters who are generally inclined to support a particular political party. Currently, the Republican party are benefitting most from these arrangements. The redistricting process is also managed by committee, currently chaired by the Republican party. Another redistricting is expected to happen following the 2020 US census. The redistricting process in Georgia last occurred in 2015 and had come under fire due to suspected gerrymandering to suppress minority votes (source "[https://democraticredistricting.com/state/georgia/"](https://democraticredistricting.com/state/georgia/)). A quick look at a sample of redistricting results in Georgia may explain why people are curious:







(source: wikipedia - https://en.wikipedia.org/wiki/Georgia%27s_congressional_districts)

The redistricting process in Georgia appears to be well-established and unlikely to change in terms of its (presumed) partisan nature and results. Therefore, candidates who wish to "make a difference" and their campaign advisors may find that shifting focus to include other demographic groupings that expose demographic-specific voter concerns may create a more relevant, stronger and winning election platform.

The analysis in this report explores the demographics for the districts as-drawn and then the demographics of counties across the state *ignoring district boundaries* in order to uncover potential avenues for combatting gerrymandering effects. **The target audience for this analysis are political campaign management firms who do business in Georgia.** Such firms typically control campaign investment strategies, including advertisements, focus locations, and delivery channels for candidate communication. These strategies would typically be aimed at achieving the highest voter turnout for the candidate, **therefore, demographics analysis is vital to the success and reputation of the campaign management firm** (as well as the candidate).

Caveats: this is exploratory analysis (in the loosest sense); results and observations are limited to datasets available to students and in the context of the IBM Data Science Professional course schedule. Moreover, I am not a lobbyist or expert at campaign management and this report does not express a political preference. At a time where complete and accurate information ("truth") is under stress, it is important that readers bear these points in mind re: this notebook and report.

Also be aware, this analysis required quite a few disparate files, including US government tables, which were awkward to parse; the accompanying Jupyter notebook is somewhat longer than the notebook referenced in course submission sample.

Data Requirements:

Certain characteristics of the 14 "red" (Republican) and "blue" (Democratic) districts will be reviewed:

- voter population size
- ethnicity (using 2015 survey estimates related by BallotPedia)
- education levels
- age ranges
- poverty levels in each district
- local amenities (venue categories) for each district

('Gender' has been excluded as there would be much more to say than course time permits.)

The features were sourced from:

- https://ballotpedia.org/Redistricting_in_Georgia - provides congressional districts by number, current representative by full name, party affiliation, election victory margins, and district ethnic demographic estimates (BallotPedia referenced the United States Census Bureau, "American Fact Finder: 2015 1-year estimates," for ethnicity data). *Note - the election results were ignored (too old) however the ethnicity estimates were used as I could not find any usable district-level ethnicity data in the time available.*
- https://en.wikipedia.org/wiki/2018_United_States_House_of_Representatives_elections_in_Georgia - provides congressional district 2018 election results for Georgia.
- https://www2.census.gov/programs-surveys/demo/tables/voting/table02c.xlsx - "Characteristics (Educational Attainment) of the Citizen Voting-Age Population for Congressional Districts: 2018" - provides education attainment per district.
- https://www2.census.gov/programs-surveys/demo/tables/voting/table02a.xlsx - "Characteristics (Age) of the Citizen Voting-Age Population for Congressional Districts: 2018" - provides voter age per district.
- https://www2.census.gov/programs-surveys/demo/tables/voting/table02b.xlsx - "Characteristics (Sex and Poverty) of the Citizen Voting-Age Population for Congressional Districts: 2018" - provides number of people living in poverty for each district.
- "./Documents/Coursera IBM Data Sci Spec/9. Applied Data Science Capstone/Coursera_Capstone/georgia_districts_addr" - manually created from the .gov websites of each congressional representative (as per BallotPedia Georgia Districts html table) for use with FourSquare venue api (using the first office address where representatives may have had multiple addresses).
- https://developer.foursquare.com/docs/build-with-foursquare/categories - FourSquare used to select venue information for each district. Such data will further inform the philosophy and voting pattern potential for a given district.
- https://www.census.gov/geographies/mapping-files/2018/geo/kml-cartographic-boundary-files.html ("cb_2018_13_sldu_500k") for map of Georgia USA with county boundaries, converted to geojson using the open-source tool kml2geojson.

Below data sources are county-level detail that will be used in the experimental reclustering of counties based on demographics and not congressional district *(important note - counties are NOT subsets of districts):*

- https://apps.bea.gov/iTable/iTable.cfm?reqid=70&step=1&isuri=1&acrdn=8#reqid=70&step=1&isuri=1&acrdn=8 "CAINC30_GA_1969_2018.csv" - manually extracted from CAINC30 zip file downloaded from the "US Dept of Commerce - BEA Regional Economics Data") - provides population and income by county ('LineCodes' 100 and 110).

- http://sos.ga.gov/admin/files/Active_Voters_by_Race_Gender_and_Age_as_of_May_1_2020.xlsx - from Georgia Voter Registration Statistics goverment site - active voter numbers by race and gender.
- "/Users/jtege/Documents/Coursera IBM Data Sci Spec/9. Applied Data Science Capstone/Coursera_Capstone/Georgia_Congressional_Distr.png" - for map overlay of Georgia state 2013, courtesy of National Atlas.gov. (*Note: congressional districts have not changed dramatically since 2013 however this map should be used as an indicative reference only).*

Methodology:

In order to suggest meaningful focus areas for campaign advisors, several steps were followed:

1. **Data sourcing for Districts** - Useful data sources on Georgia congressional districts will be identified with information that is as current as possible (see "Data" section above). The data will then be loaded into a common 'merged' data frame which can be easily charted and evaluated (with '.head()' checks along the way to check data preprocessing). Note that data will mostly be sourced from public and government websites with the exception of:
    - district center-point addresses, which are more difficult to source and required a manually-derived table
    - amenities information for districts of interest; this will be provided via FourSquare apis.
2. **Data Visualisation and Analysis for Districts** - Data visualisations will be created that illustrate demographic breakdowns for the districts. Highlights and anomalies will be noted, with a view to narrowing focus to a few districts with the most interesting characteristics. Further elaboration data will be charted for the districts to understand types of amenities available in each.
3. **Data Sourcing, Visualisation and Analysis for Counties** - Repeat the Data sourcing and visualisation steps above for Counties (smaller geographical representations) in Georgia to identify clusters of counties that might further inform better ways of targetting voters than reliance on (allegedly) gerrymandered district boundaries.

The following describes these steps in more detail.

1. Data Sourcing for Districts

A dataframe of district demographic data for the 14 congressional districts in the state of Georgia was initially seeded from “Electoral Margins of Victory” data in the Ballotpedia site (see Data section). BeautifulSoup was used. Given the number of tables embedded in the Ballotpedia page, this was an interactive process. Also, an amount of cleanup was required for usability (including column name and datatype changes). As it was not possible to not locate a district-level report on ethnicity that could be easily parsed, voting data by ethnicity was pulled from the Ballotpedia page, following the same cleaning process as before, and merged these into a consolidated table. This data was acquired by Ballotpedia from the United States Census Bureau, "American Fact Finder: 2015 1-year estimates," for ethnicity data; most other data pulled in this analysis is from 2018.

Education, age and poverty information by district was easier to locate on US government websites and load into the consolidated dataframe, although an amount of preprocessing was required due to various header and data type differences.

In preparation for later mapping and venue contrasting across districts, district address information was required; unfortunately it was not possible to find usable source for this and so a small .txt file of district address information was created with data manually sourced from the “offices” listed on each individual congressional district representatives’ website. Where multiple offices existed, the first address in the list was used. The data was then run through geocoder to apply location information for each district (for later use).

Lastly for the district information gather, the respective political party affiliation for the district and the party colours was added: “indianred” for Republicans, “cornflowerblue” for Democrats.

After the data had been sourced, it was time to chart observations to evaluate districts that may be of most interest for further analysis.

2. Data visualisation and Analysis

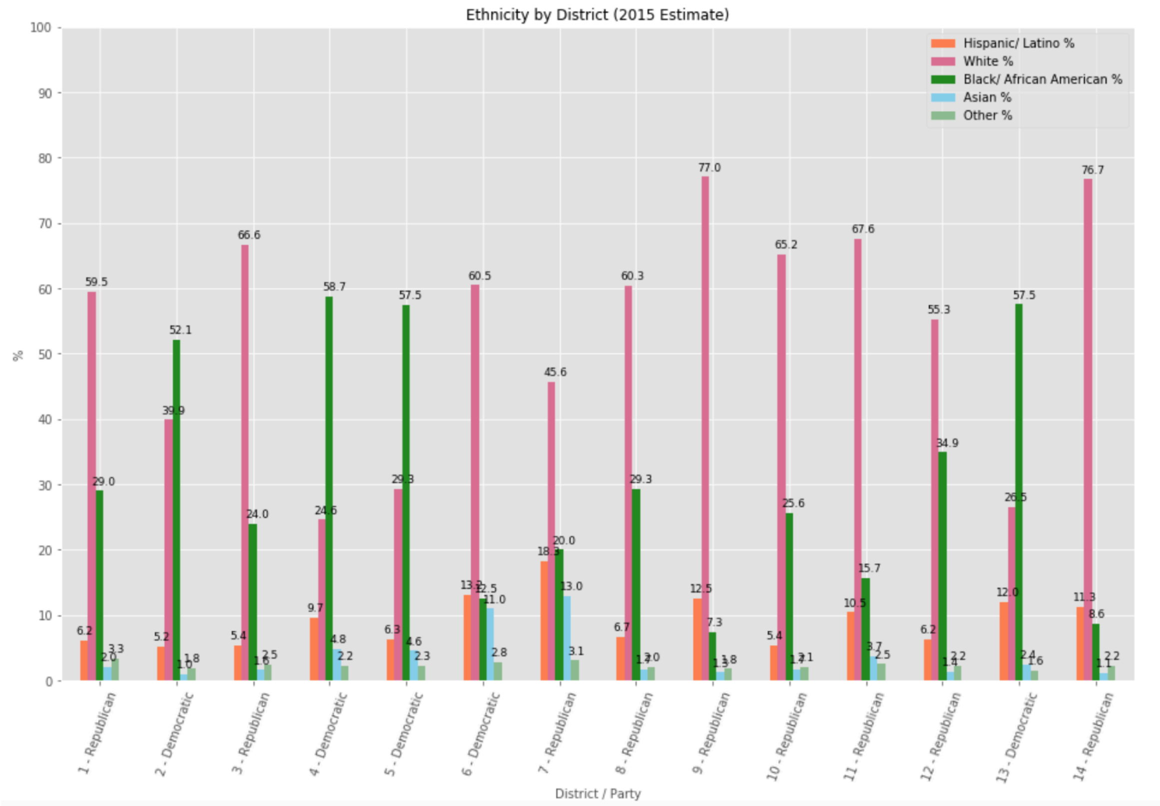
Given the amount of categorical information to review across 14 districts, bar charts seemed the most appropriate visualisation tool, organised by demographic type (ethnicity, education, age, income/poverty). Totals at the tops of each bar were added as long as doing so did not clutter or make the bar chart unreadable.

Additionally, due to the number of districts to assess, a “top 3s” spotlight dataframe was created reflecting the top 3 districts for each demographic category. *The top3 table has not been included in the main body of this report but is available in the Appendix and in the accompanying Jupyter notebook*.

Ethnicity, age and education information for all 14 districts was plotted in 3 separate bar graphs. Poverty data was also plotted juxtaposed with voter population in each district. Lastly, venue preferences in the districts were plotted.

Ethnicity

For ethnicity, an additional amount of preprocessing was necessary; smaller columns of data (“American Indian and Alaska Native”, “Native Hawaiian and other Pacific Islander”, and “Multiple races”) were consolidated into the “Other” column category for bar chart reading ease (not because those numbers were unimportant).

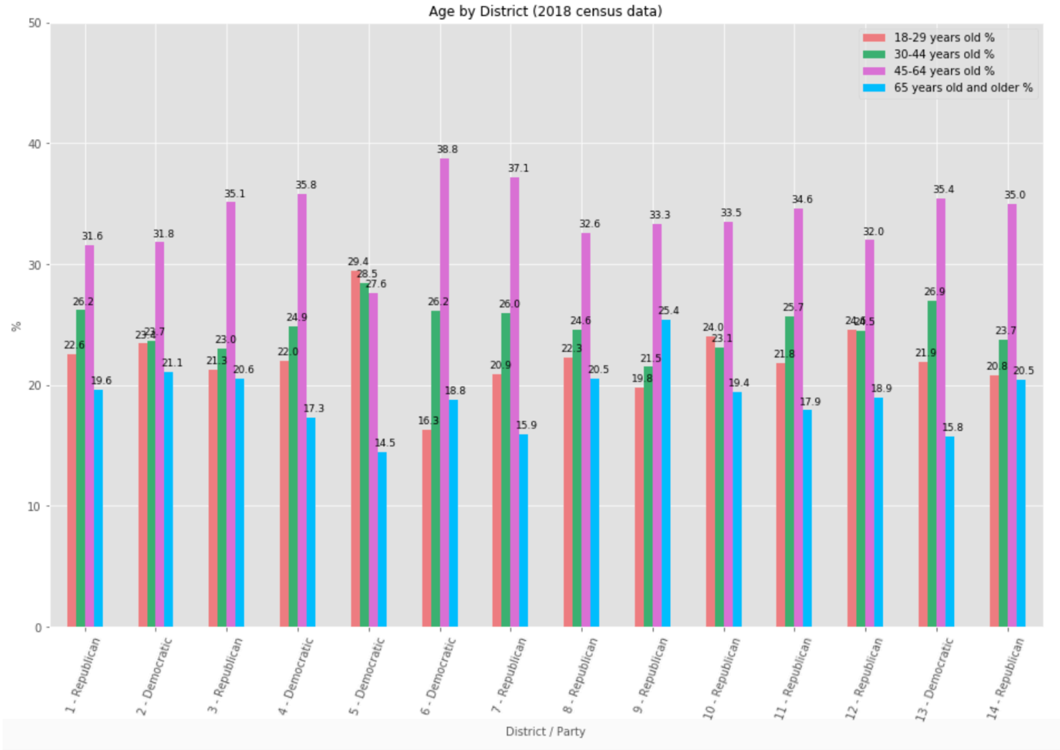


Data shows that, broadly speaking, the top 3 districts with a higher number of white constituents relative to the rest of the district population (Districts 9, 14 and 11 with 77.0%, 76.7%, and 67.6% white constituent voters, respectively) appear to have elected Republican candidates.

The top 3 districts with a higher number of black constituents (Districts 4, 13 and 5 with 58.7%, 57.5% and 57.5%, respectively) appear to have elected Democratic candidates.

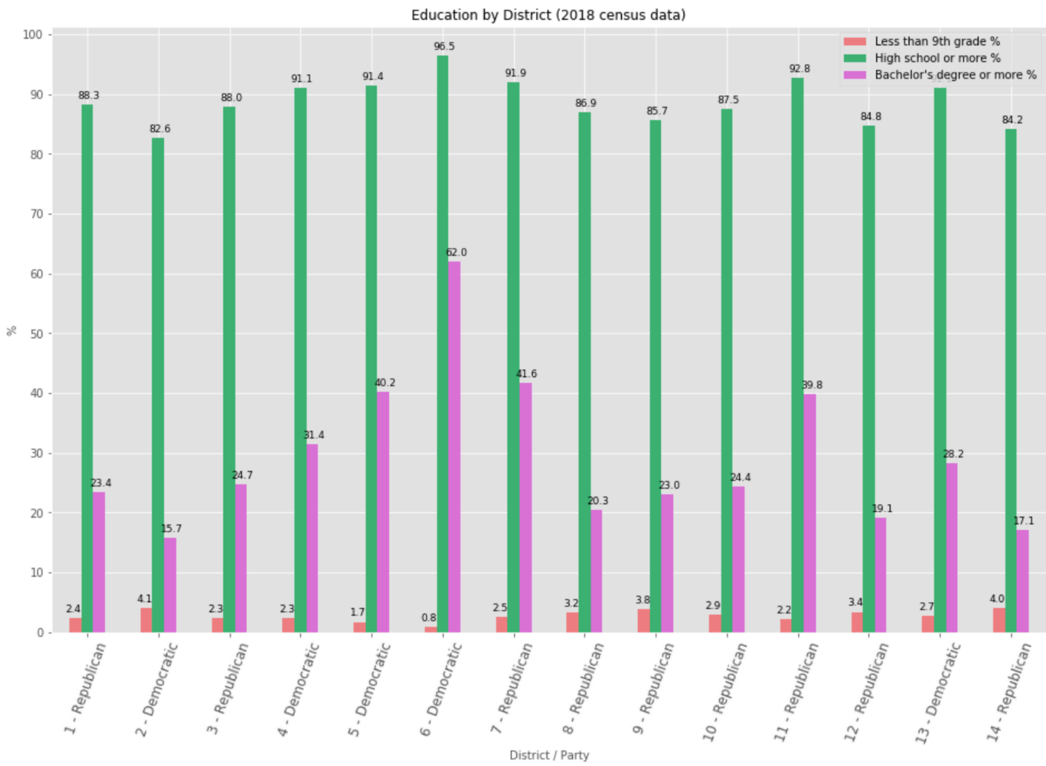
The exception is District 6, where the bar graph shows the proportion of white constituents to be higher relative to the rest of the district population but a Democratic candidate was elected.

Age



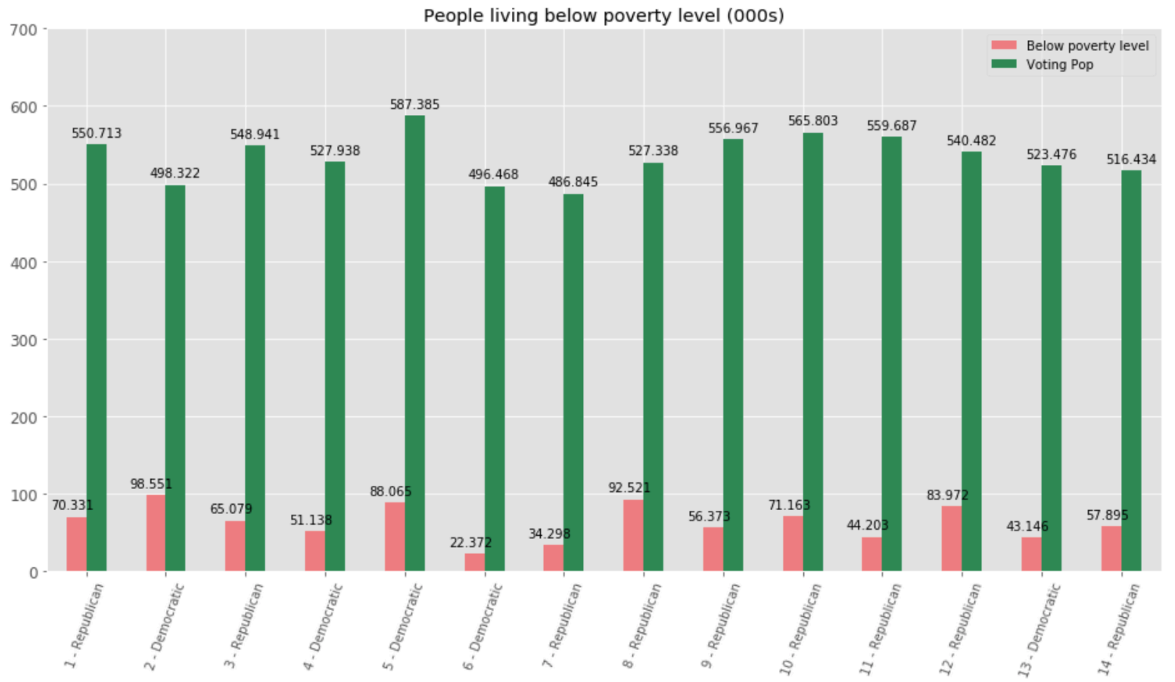
Age data shows District 5 to have the largest voting participation from 18-44 year old constituents whilst District 6 appears to have the greatest participation from 45-64 year old constituents. District 9 has the greatest participation from constituents 65 years and older.

Education



Education data shows that District 6 has the highest number of college graduates (308.1k), followed by District 5 (236.3k) and District 11 (222.6k). District 9 has the highest number of people who did not enter highschool, followed by District 14 and District 2 (20.9k, 20.7k, and 20.2k, respectively).

<u>Poverty</u>



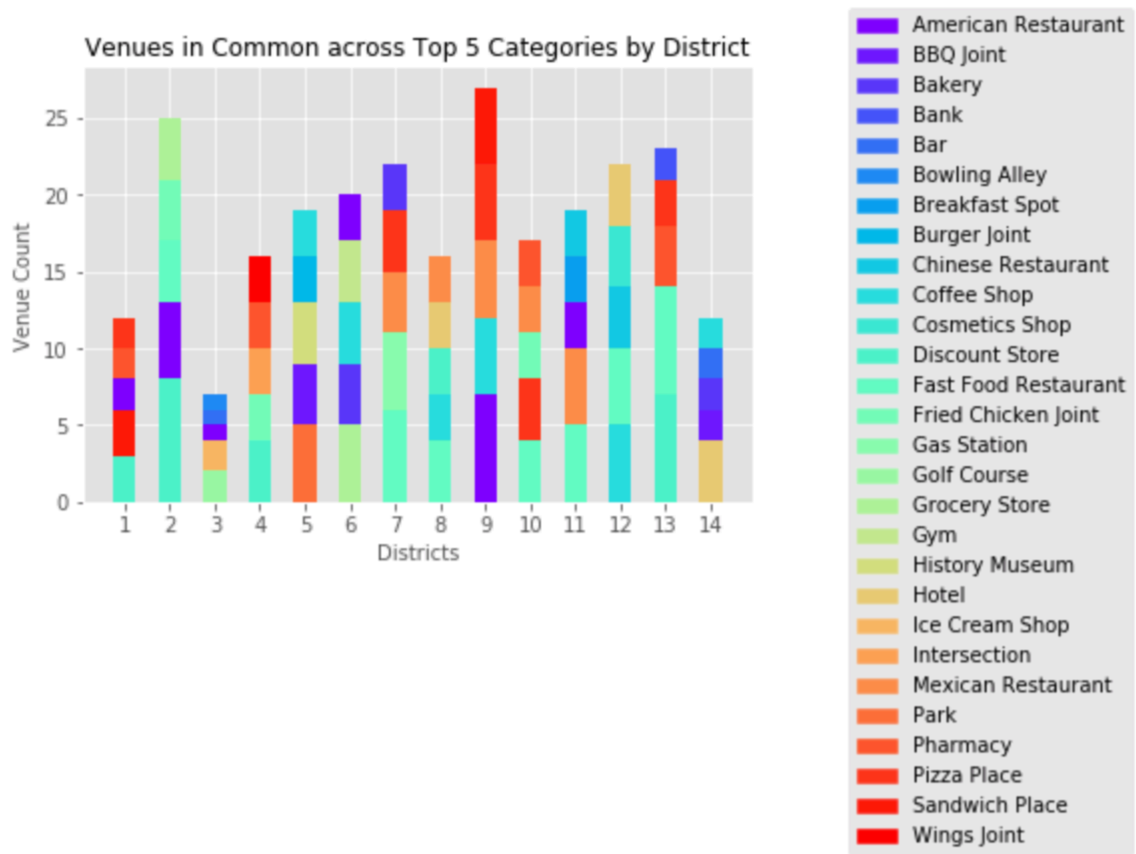
Poverty and population data were plotted next to each other - whilst it is unclear that the poverty figures only include voters, it is useful nonetheless to understand the voting size relative to poverty size in each district.

Poverty data shows highest levels to be in Districts 2 (98.6k people) followed by Districts 8 (92.5k people), and District 5 (88.1k people), with District 12 in close proximity (at ~84k people). Note District 2 and District 5 are Democratic whilst District 8 and District 12 are Republican, suggesting that poverty level is not necessarily a party vote differentiator.

Note that District 6 (Democratic) appears to have the lowest level of poverty in Georgia, at only 22.4k people.

<u>Venues of interest</u>

As prerequisite for the course Capstone, amenities data the FourSquare service apis were incorporated into this analysis to pull venue data for all Georgia congressional districts to identify potentially defining characteristics in districts. The locations used were office addresses manually retrieved previously. Up to 1,000 venues located within 3,200 meters (approximately 2 miles) from each district office were pulled for cross-district comparison. The analysis was limited to 5 preferred venue types (this seemed a reasonable number to review). The FourSquare results were sorted and grouped by district, and venue categories per district were plotted in a stacked bar chart (below). Because it was not possible to predict the number of categories that would be returned, colours were set from a range created using a technique applied in an earlier lab (noted in the accompanying Jupyter notebook). Additionally, it was necessary to shift the chart legend to outside of the normal plot borders using 'bbox_' parameters.



The data suggests that there are certain venue preferences in common across a number of districts, for example, Fast Food restaurants, American restaurants and Sandwich places (unsurprisingly). The limited selection size and venue categorisation issues suggests this chart is of limited value to campaign strategy discussion. If there was more time to work through the FourSquare developer forums to get cleaner venue types, the exercise may have proved more useful.

Districts summary

Based on data visualised and analysis thus far, interesting districts for suggested review are:

- District 6 - unusual ethnicity demographic relative to political party elected, second-highest populations of Hispanic, Asia and Other ethnicity, highest educational attainment, highest number of voters aged between 45-64 years old, and lowest poverty level. It is also the only district to have elected a different party in 2018, albeit by the second narrowest margin (3,264 votes - see “2018 Election Results” dataframe in Appendix).
- District 2 - highest poverty level, lowest college attainment, second-highest percentage of older voters (65 and older).
- District 9 - highest population of white voters, lowest educational attainment, greatest percentage of older voters (65 and older) and largest election victory margin (166,749 votes - see “2018 Election Results” dataframe in Appendix, noting unopposed elections in District 5 and District 8 were ignored).

Note that District 1, District 3, District 10 and District 12 appear in the ethnicity, education and age top3 list the least; interesting but not interesting enough to pursue in this analysis.

Having consolidated the districts analysis, it was time to review alternatives to voter groupings. In the next section, county data was gathered and assessed to arrive at alternative means of grouping voters than the (allegedly) gerrymandered district boundaries.

3. Data sourcing, visualisation and analysis (including K-means) for Counties

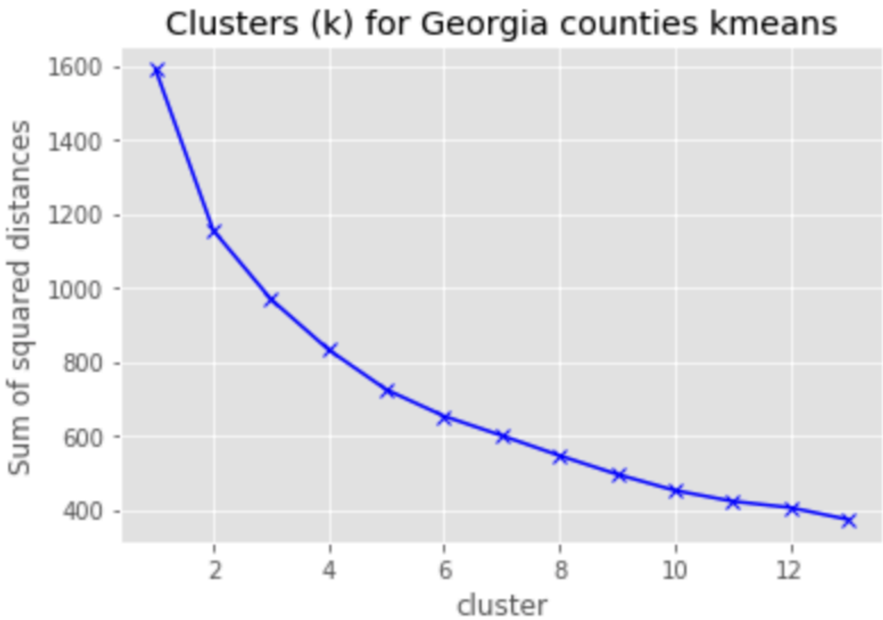
As with district data, a dataframe of demographic data was created for the 159 Georgia counties represented in US Census Bureau data (see Data section). This data was primarily excel tables, all of which required preprocessing.

<u>Demographics data sourcing</u>
Whilst it was possible to find voter population sizes, age ranges and ethnicity, it was not possible to find poverty numbers by county (as had been found by district) in the time allotted. Instead, per capita income was used as a proxy metric to poverty. It was also not possible to find education information by county as had been found by district; as no suitable proxy was found, this data was excluded. Moreover, ethnicity data for “Asians” is incorporated into the group “Asian or Pacific Islander” for

counties ("Asians" as a demographic was split out in the district-level detail); this is unlikely to impact the modelling significantly given the low values.

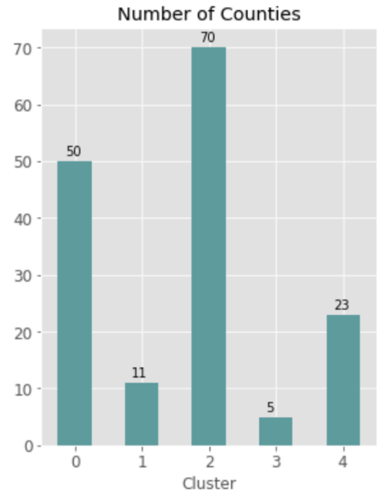
Cluster analysis

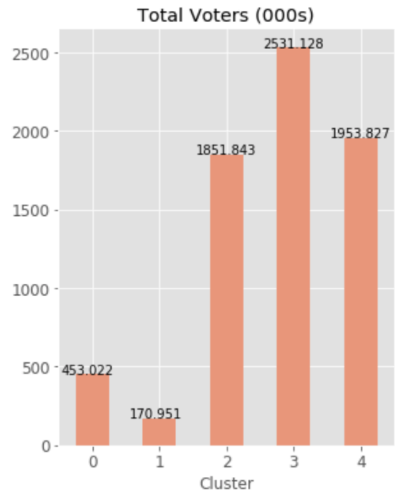
After consolidating the demographic data, the clustering technique K-means was used to create experimental groupings of counties that could provide an alternative cohort to 'districts'. K-means was well-suited to this use case based on the need to segment counties in an unsupervised, easy to run and interpret manner. However, using Kmeans required a starting 'k' number of clusters. In an attempt to identify the ideal 'k', the algorithm was executed a number of times (14, although this was a somewhat arbitrary selection) and results plotted each k against the sum of squared distances for the respective county datapoints. The plot was then inspected for the most dramatic change between k values ("the elbow"):

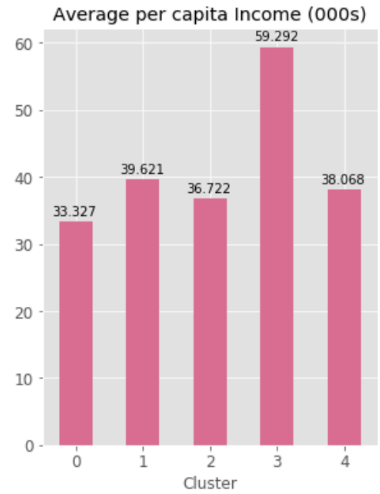


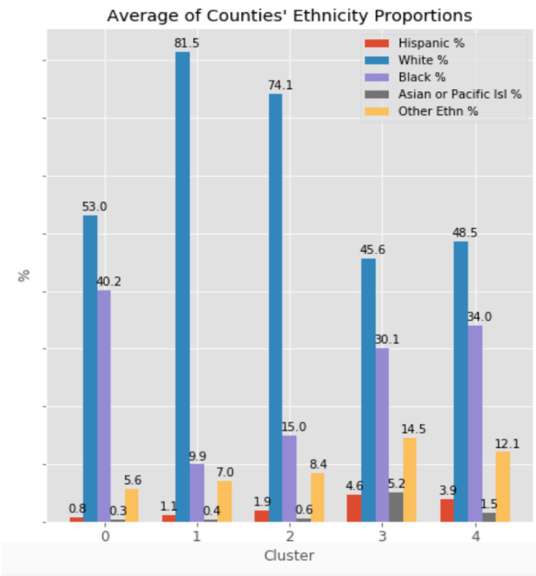
Unfortunately, the graph yielded an unsatisfactory result, in that the most dramatic change in difference appeared to be at 2, a very limited cluster number offering little campaigning flexibility, and next most dramatic seemed to be 10 and 12, which were too many clusters to consider for this cursory analysis. The next nearest change (as far as visual inspection could relate) was deemed to be 5. The K-means was rerun with k = 5, resulting in 5 distinctive cohorts.

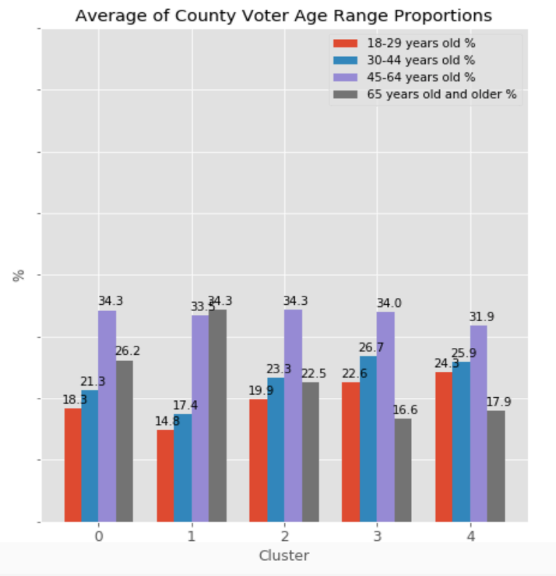
Each county-cluster assignment was recorded into the merged data frame. Demographics-oriented bar chart groups (matplotlib 'subplots') were then used to identify and contrast distinguishing features of the clusters:











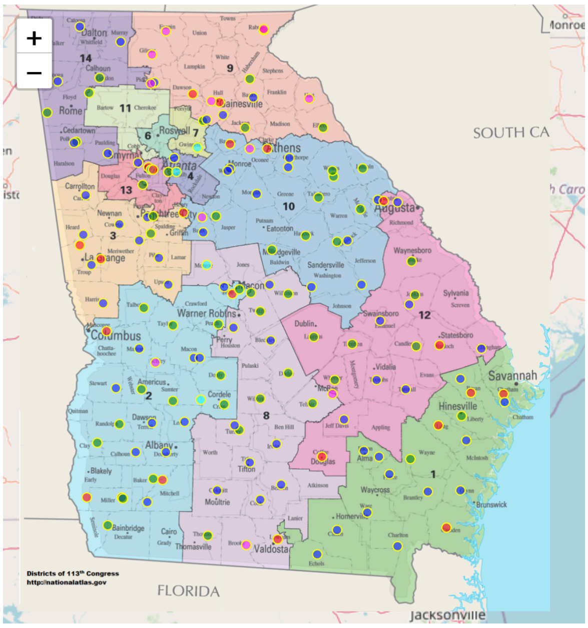
Based on these visualisations:

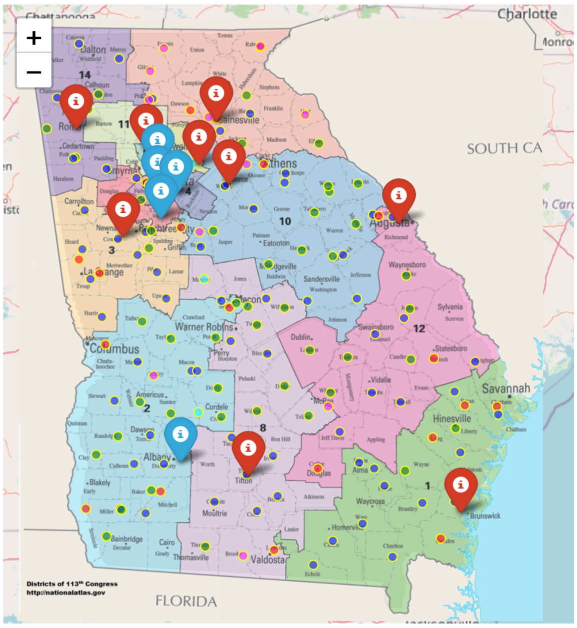
- **Cluster 0** - has the highest proportion of ‘black’ voters across clusters (40.2%), lowest combined proportion of ‘hispanic’, ‘asian’ and ‘other’ voter constituents (6.7%), second-lowest number of voters (~453k) across the second-highest number of counties (50, suggesting broad geographic dispersion 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 ethinicities (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.

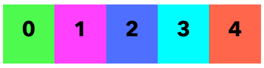
Note that, as these visualisations were easy to read, a special ‘top 3’ table for evaluating counties’ distinguishing features (as used for districts) was not necessary.

Clusters were then depicted geographically (to achieve this, each county was ‘geo-located’ using the ‘geocoder’ api). District office addresses were also applied to the map for more context (including political party colour). Because it was not possible to source useful kml or geojson data in order to contrast clusters against Georgia districts, a .png file illustrating Georgia district borders was used to overlay onto a choropleth map of Georgia. The colour-coded depictions of districts in the picture provided useful backdrop to the clusters. It was, however, quite tricky to align the picture and map state borders perfectly so a ‘good-enough’ alignment was used:





Cluster marker colours



Note that the district colours set in the .png overlay of the Georgia map appear to be random.

As district boundaries are not aligned to county boundaries in many cases (due to the interesting redistricting process described in the Introduction), it is difficult to cleanly depict how counties of particular clusters align to existing districts. These maps, however, give a reasonable approximation and suggests potential campaign alignment strategy that avoids normal 'district-centric' targetting.

Although counties of the same clusters are geographically dispersed, there are dominant clusters in certain districts. District-cluster *dominance is* defined (for this analysis) as any sole cluster or a cluster with at least a 2 county majority based on map inspection (as county borders do not cleanly align to district borders). For example, District 1 (Brunswick) appears to have a much greater prevalence of Cluster 2 ("blue") counties, whereas District 10 (Monroe) appears to have a mostly even distribution between Cluster 2 and Cluster 0 ("green") counties. See list below (manually created):

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*

Additional observations were that, for District 6, District 2, and District 9 (identified earlier as districts of interest):

- District 6 - (part of Atlanta) appears to have no clusters shown. This strangeness is likely to be a consequence of the (allegedly) gerry-mandered re-

districting process not aligning to county boundaries coupled with the positioning defaults from the geocoder api (there are also issues in the .png district map overlay alignment however these are inconsequential). There is a Cluster 2 ("blue") county nearest the district, although that county appears to also straddle District 4 (Decatur).
- District 2 - (Albany) appears to have more Cluster 0 ("green") counties, although by a narrow margin over Cluster 2 ("blue") counties.
- District 9 - (Gainesville) appears to have a greater number of Cluster 1 ("magenta") counties.

Finally, given the redistricting requirement of equal approximate population size, it is unsurprising to see the noticeable cluster of district offices in the Atlanta region (a densely populated area of approximately 469,000 people in 2015, the year of the last redistricting. (source: wikipedia)).

Results/Discussion:

It would seem that demographic 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. District-level issue campaigning (eg, for state-wide elections) should incorporate a broader set of issues than those usually associated with the particular political party in office in a given district. Whilst that should always be the ‘right’ thing to do, there are cost implications (analysis of which is out of scope for this report).

There may be exceptions to the idea of lack of issue alignment. For example, District 6 (Democratic) seems comprised of a single “blue” county (note that the county-agnostic redistricting process makes this an unsafe observation). If there is no alignment issue, the fact that only one county is shown makes campaign targeting by ‘Cluster 2’ issues probably similar to campaigning by ‘District 6’ issues; thus no new information and no requirement to change from a district-focussed to county-cluster-focussed strategy. This may be unsatisfactory considering the low 2018 victory margin in District 6 (see Appendix).

In terms of prioritising issues based on cluster (and not district), note 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. However, Cluster 3 in particular 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.

Conclusions

Although this analysis did not yield earth-shattering revelations, there are some useful conclusions drawn and helpful suggestions that can be made.

Elections are about appealing to the most voters, each of whom may have issues that one could assert broadly align to demographics. Based on this assertion and analysis results indicating the potential misalignments between district and voter issues, it would appear that “it’s all to play for in 2020” for campaigners of any party,

As a suggestion to the state-wide campaign advisors, it would make sense to prioritise the issues for Cluster 3 (“cyan”) counties (in the context of the geography of those counties) based on voting population size and less-skewed demographics, with strategies adjusted to spend less time and money on income/poverty-level-focussed conversation (as troubling as that idea may be to this author).

It would also make sense for campaign advisors to prioritise the issues for Cluster 2 (“blue”) as it appears to be dominant in the most districts, ie, cluster-specific campaigning could also be used for district-level campaigning. This could prove especially useful for elections in District 6 (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.

In the author’s opinion, these suggestions hold despite the various data sourcing and prepping issues, namely:

- No district-level report on ethnicity that could be easily parsed - data used was related from the United States Census Bureau, "American Fact Finder: 2015 1-year estimates” as related by Ballotpedia website.
- No district address information - information manually sourced from the first office listed on each individual congressional district representatives’ website.
- County per capita income used as a proxy metric for county poverty.
- No county education information.
- The use of 5 clusters instead of 2.
- No gender information included at all.

Note this is cursory, high-level analysis of subject matter that is very complex and and highly nuanced. This report barely scratches the surface.

Appendix

1. Top 3 districts by demographic type (see Jupyter notebook dataframe)

District - Party	Hispanic/ Latino %	White %	Black/ African American %	Asian %	Other %	Bachelor's degree or more	Less than 9th grade	18-29 years old %	30-44 years old %	45-64 years old %	65 years old and older %
1 - Republican	-	-	-	-	-	-	-	-	26.200000	-	-
2 - Democratic	-	-	-	-	-	-	20.248000	-	-	-	21.060000
3 - Republican	-	-	-	-	-	-	-	-	-	-	20.560000
4 - Democratic	-	-	58.660000	4.840000	2.240000	-	-	-	-	35.790000	-
5 - Democratic	-	-	57.510000	-	-	236.287000	-	29.450000	28.470000	-	-
6 - Democratic	13.180000	-	-	10.990000	2.830000	308.067000	-	-	-	38.780000	-
7 - Republican	18.300000	-	-	12.960000	3.100000	-	-	-	-	37.150000	-
9 - Republican	12.540000	77.040000	-	-	-	-	20.915000	-	-	-	25.390000
10 - Republican	-	-	-	-	-	-	-	24.020000	-	-	-
11 - Republican	-	67.620000	-	-	-	222.558000	-	-	-	-	-
12 - Republican	-	-	-	-	-	-	-	24.590000	-	-	-
13 - Democratic	-	-	57.530000	-	-	-	-	-	26.940000	-	-
14 - Republican	-	76.670000	-	-	-	-	20.663000	-	-	-	-

2. 2018 Congressional Election Results (see Jupyter notebook dataframe)

District - Party	Election summary	Republican Votes	Democratic Votes	Rep-Dem Margin
1 - Republican	Republican Hold	144741	105942	38799
2 - Democratic	Democratic Hold	92472	136699	-44227
3 - Republican	Republican Hold	191996	101010	90986
4 - Democratic	Democratic Hold	61092	227717	-166625
5 - Democratic	Democratic Hold	0	275406	-275406
6 - Democratic	Democratic Gain	156875	160139	-3264
7 - Republican	Republican Hold	140430	140011	419
8 - Republican	Republican Hold	198152	0	198152
9 - Republican	Republican Hold	224661	57912	166749
10 - Republican	Republican Hold	190396	112339	78057
11 - Republican	Republican Hold	191887	118653	73234
12 - Republican	Republican Hold	148986	101503	47483
13 - Democratic	Democratic Hold	69760	223157	-153397
14 - Republican	Republican Hold	175743	53981	121762