Politech Mathematical Group

Political Fatness Report

Darren Kong	Hugo Mainguy	Jeffrey Zhong
110770716	111747982	112299792

March 8, 2021

1 Abstract

The issue of gerrymandering has long been an issue in the United States, as many criteria are considered in the process of creating fair districting plans, and these criteria differ from state to state. Some cases in the past have used the Polsby-Popper test as one way to calculate compactness as a measure of gerrymandering, and the test generally favors geometrically compact shapes. However, Polsby-Popper does not account for the distribution of the population within the shape, and as such, may unnecessarily punish non-compact districtings. The paper therefore investigates an alternate compactness measure that factors in the distribution of the population alongside the geometry of a districting: population fatness. Using the ARS, we calculated the population fatness alongside the Polsby-Popper score for the districts of several U.S. states. Graphing the results, we examined the extrema in each quadrant to determine the attributes that most affect each measure. It may be of interest to weigh the Polsby-Popper test score and the population fatness score together for a better measure of compactness.

2 Introduction

The advent of technology, especially in more recent times, has made every aspect of life "more" something. For instance, we are more connected with the Internet and instant communication across the globe than we were with letters, or printed newspapers. It may seem that this intensification of long-distance communication should make us all closer to one another. Yet, it seems that technology has also made mankind more divided, with the decentralization of news outlets or American politics. With the advent of technology, the decennial process of redistricting states with more than one congressional district has become a partisan affair. Fewer districts are competitive now than even twenty years ago such that there is less need to appeal to the other side (Wasserman and Flinn 2017). Artificial intelligence applied to redistricting has contributed to a polarization of American politics; however, if used with good intentions, it could become the very solution to a problem that it fueled in the first place.

It is easy to call out gerrymandering, yet, at the same time, difficult to give precise measures for it. In Vieth v. Jubelirer, ruled in 2004, the Supreme Court argued that a districting was not gerrymandered for there was no measure to quantify inequality or lack of fairness. As courts – especially the Supreme Court – often rule by precedent, this has been used as an excuse to bring down some cases that made it to the highest court of the land. This has emboldened strategists to make increasingly "unfair" maps. The surge in gerrymandered districting plans is the combination of three factors that all came together at the same time: large political successes for one party, the Republicans, just before drawing districts, the sudden improvement of technology, and the increasing polarization in voting patterns.

This comes together in what could be considered a paradox: while there are ever more accurate methods to gerrymander at will, there is no mathematical standard accepted by the courts to identify the practice as of today. Furthermore, state constitutions tend to lag even further behind, because of the amount of the slow legislative process and support required to change them. As explained by Levitt, only twenty-three states require their congressional districts to be contiguous. While in practice, this is virtually always the case, it creates more potential for gerrymandering in over half of the states. (Contiguity refers to the fact that any point of the district must be accessible from any other point of the district without leaving it, except for example if part of it is an island and there is a method of transportation such as a bridge or a regular ferry between both sides.) Furthermore, only eighteen states have any mention of compactness for their congressional districts. Within those eighteen states, a minority have more requirements. For instance, Iowa requests that districts not be oddly shaped (the state is almost a rectangle that can be split equally in four rectangles), California

asks for districts not to bypass nearby large population areas for more distant populated areas, Arizona requires at least some of the districts to be competitive if feasible, and Rhode Island to represent the state fairly (although it currently has two congressional districts and will most likely lose one after the 2020 Census reapportionment.) Overall, there is a lot of leeway in what can be done by the redistricting committee.

3 Methodology

To counter this, we suggest setting a standard with a definition of compactness. Compactness is something easily perceivable by humans, yet, there is no measure that satisfies human perception all the time - and we sometimes disagree among ourselves, as shown by Kaufman et al. However, a generally sensible choice is using the Polsby-Popper measure, which is given by the formula:

$$PP(D) = \frac{4\pi A(D)}{P(D)^2}$$

Where A(D) is the area of the object D and P(D) its perimeter. This gives us a ratio between 0 and 1. It is 0 precisely when the area is 0 (D is a line or a point) and 1 when D is a circle. In particular, this favors compact round objects, and defavors those with longer or jagged perimeters. Both of these are important, and the measure has been used in real life, for instance with Arizona's redistricting in 2000. It is noteworthy that several states use different measures, particularly in the West, where more attention is paid to fairness. Other measures are used in different states: for instance, in Colorado, the districts should minimize the total perimeter. States like California and Michigan instead focus on dispersion rather than contorted boundaries, as Monorief points out.

It becomes evidently clear that the Polsby-Popper measure only takes into consideration the geometric compactness or fatness of the redistricting shape. A more suitable measure for compactness would also take into consideration the population requirement of the districtings. All congressional districtings have to be equal in population at the time of the redistricting. For example, a congressional district may seem to be shaped oddly but that is only due to the legal requirement for equal population size.

Simultaneously, it became clearer that while a non compact district is almost always subject to gerrymandering, the converse is not always true. One example of this is when districts bypass population centers nearby to instead include populations that are further away. Here we will propose an alternative to the Polsby-Popper that will be referred to as the Population Fatness measure, or Fatness for brevity. To calculate this measure, we used the following formula:

$$\frac{POP(\mathrm{D})}{POP(\mathrm{BC(D)})}$$

Where D represents the district D, BC(D) is the minimum bounding circle that contains D, and POP(D) is the population in D. This measure utilizes precincts as the base level of a population cluster. Since the bounding circle cuts through some precincts, in those cases, the ratio of the area of the precinct in the circle is used to estimate the population of the part of the precinct inside of the bounding circle, yielding good approximations. Since precincts form districts, we do not need to apply any ratio other than 1 for the population in the districts. Furthermore, this measure only takes into account the population that is inside the state, as the population outside of the state but inside the bounding circle should not influence the districts. Once again, a perfect score of 1 comes when the district is a perfect circle, and a low score close to 0 comes when the bounding circle

encompasses many populous precincts, large areas of land, or a combination of both. For the rest of this report, we will refer to this measure as the Circle Fatness measure.

Since these two measures are different, it is interesting to compare them to see when they are similar and when they differ, explain why that is the case, and conclude whether those districts are gerrymandered or not. Using the current congressional districts, here are the data points found (Pennsylvania and North Carolina maps are from 2016).

4 Results

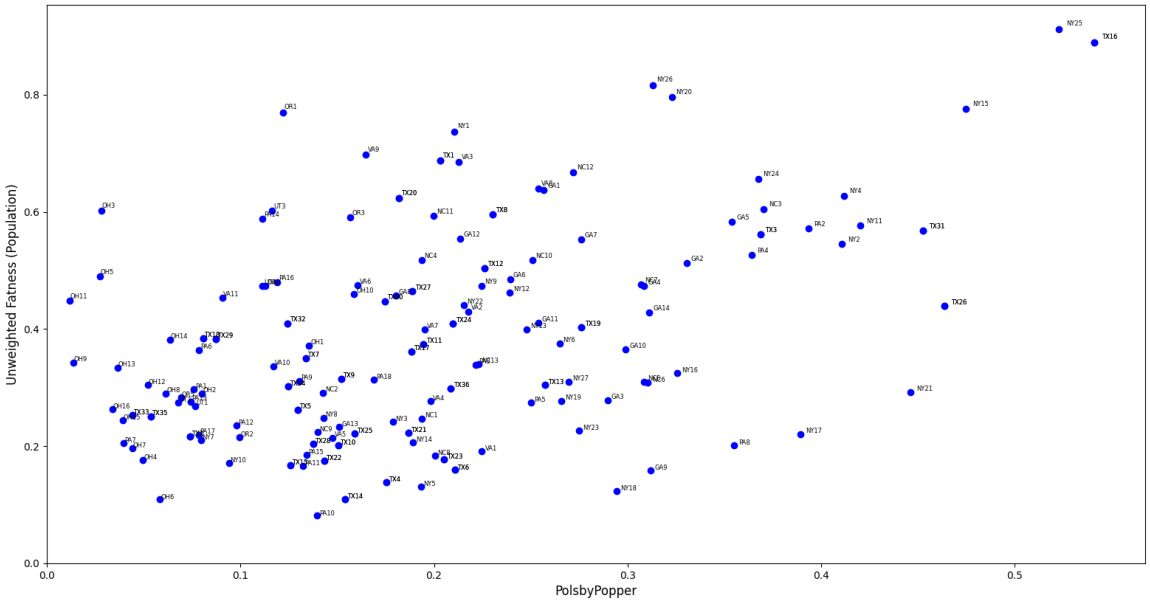

Figure 1: Plot of Population Fatness v. PolsbyPopper

5 Research

From the results, we can see compare each district's Circle Fatness score with its Polsby-Popper score. To help easily identify specific districts, we will split the graph into four quadrants. The quadrants are to be referred from 1 to 4 starting from the top right in a counterclockwise fashion. We do this to identify extremeties in each of the district scores to see in what cases these two scores may differ wildly.

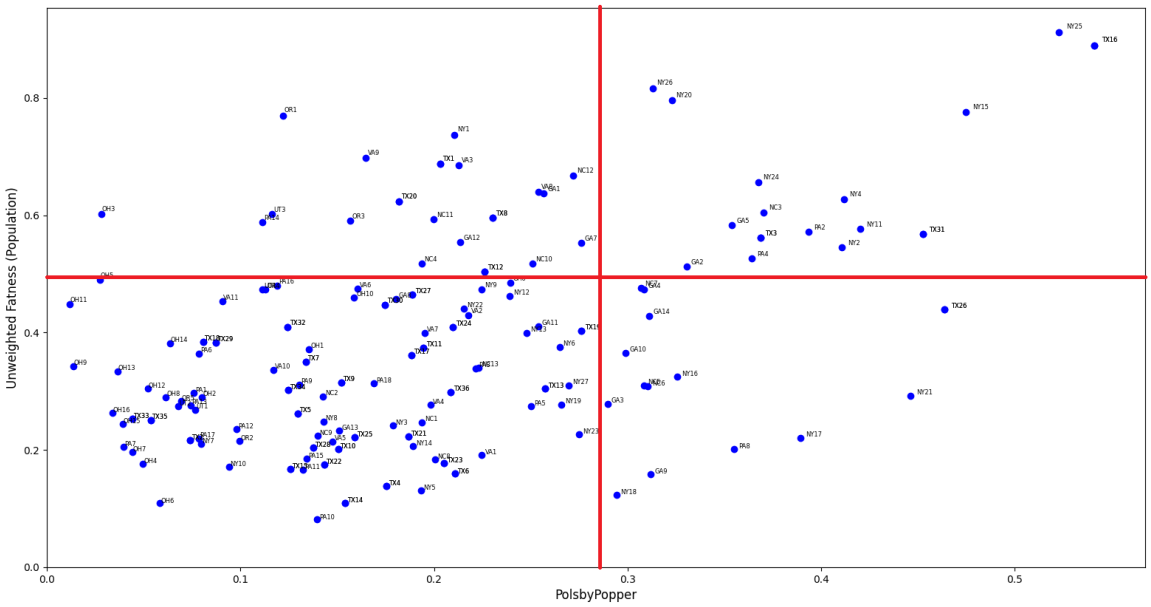

Figure 2: Plot of Population Fatness v. PolsbyPopper Edited

We will begin our look on the first quadrant, specifically NY-25, NY-15, and TX-16. The district scores for both Circle Fatness and Polsby-Popper are both high for these districts. It is rather easy to tell why.

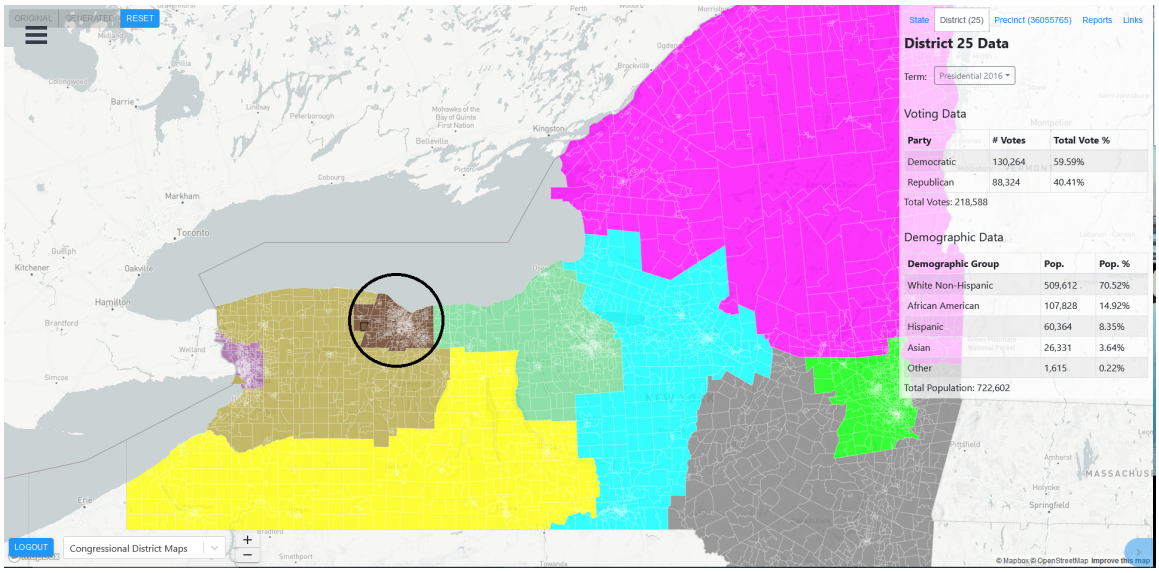

Figure 3: NY-25 Bounding Circle

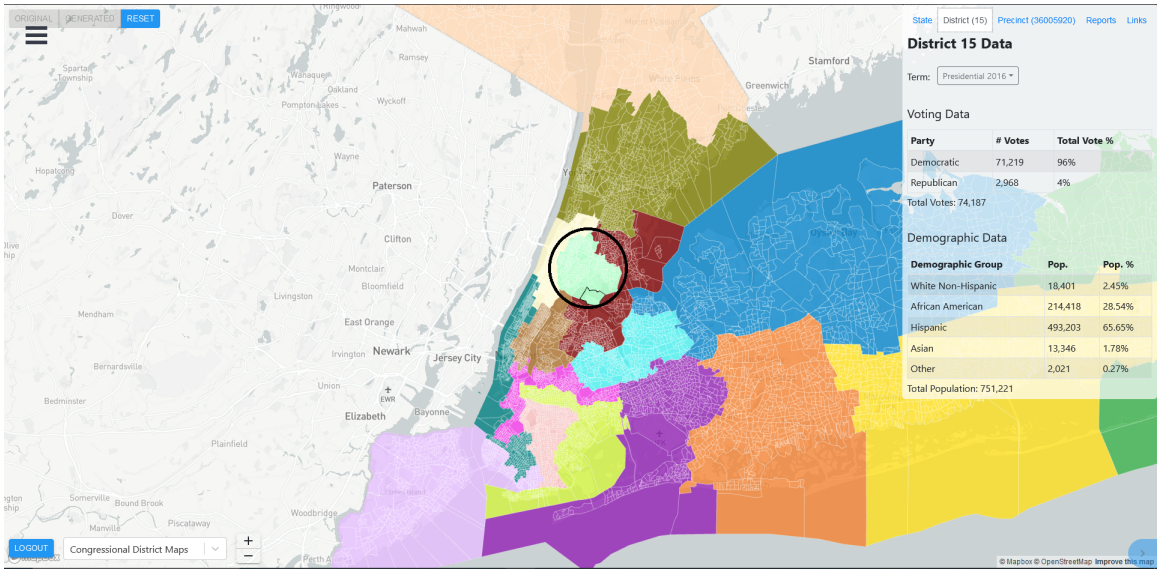

Figure 4: NY-15 Bounding Circle

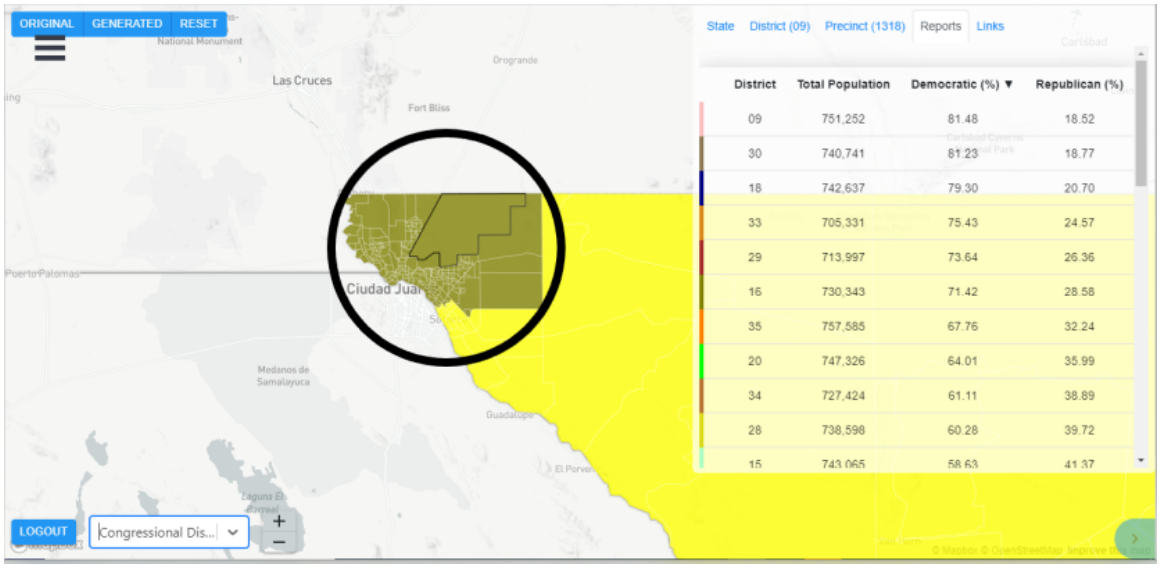

Figure 5: TX-16 Bounding Circle

These districts are all centered around a major population hub. NY-25 is centered on Rochester, NY-15 is centered on the Bronx, and TX-16 is centered on El Paso. These districts geometrically all look compact and we believe that no one would considered them abnormal in any sense. This explains the high Polsby-Popper score and also partially explains the high Circle Fatness score.

Specifically for the example of TX-16 which takes advantage of being small and densely populated with sparsely populated districts and state/country borders around, along with straight, perpendicular borders for most of its contour. It is a very Democratic area surrounded by moderately Republican areas, but it forms a coherent district encompassing one city and its direct surround-

ings, and contains all but the southernmost part of El Paso County, making it a very reasonable district. It is a similar story for the other districts.

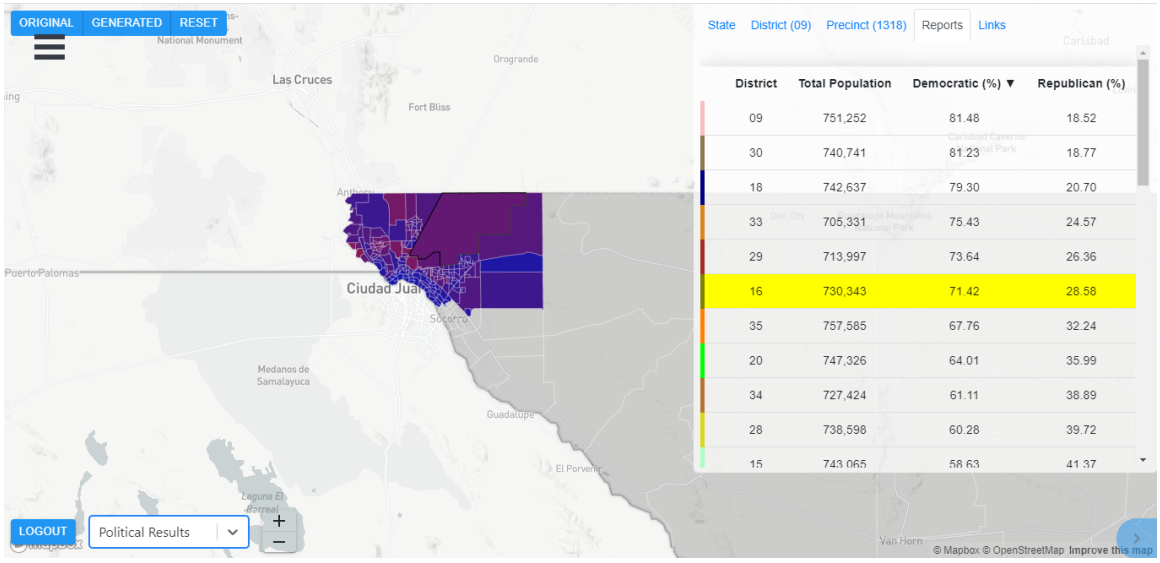

Figure 6: TX-16

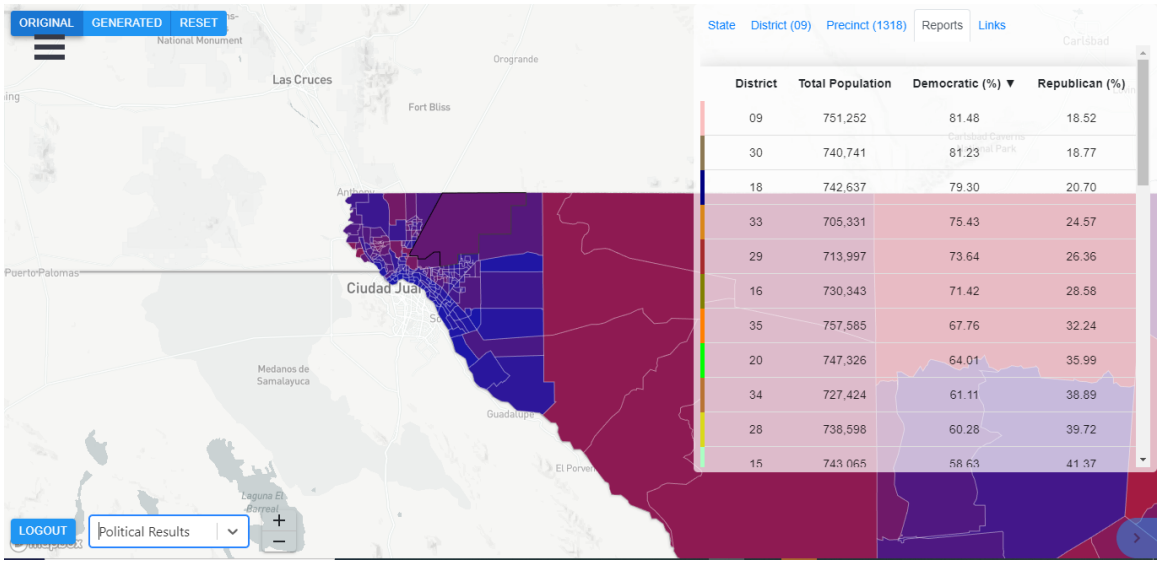

Figure 7: TX-16 Political Results

We now move onto the extremeties in the second quadrant. Specifically, many districts in Ohio form a cluster with low a Polsby-Popper score and yet a relatively high Circle Fatness score.

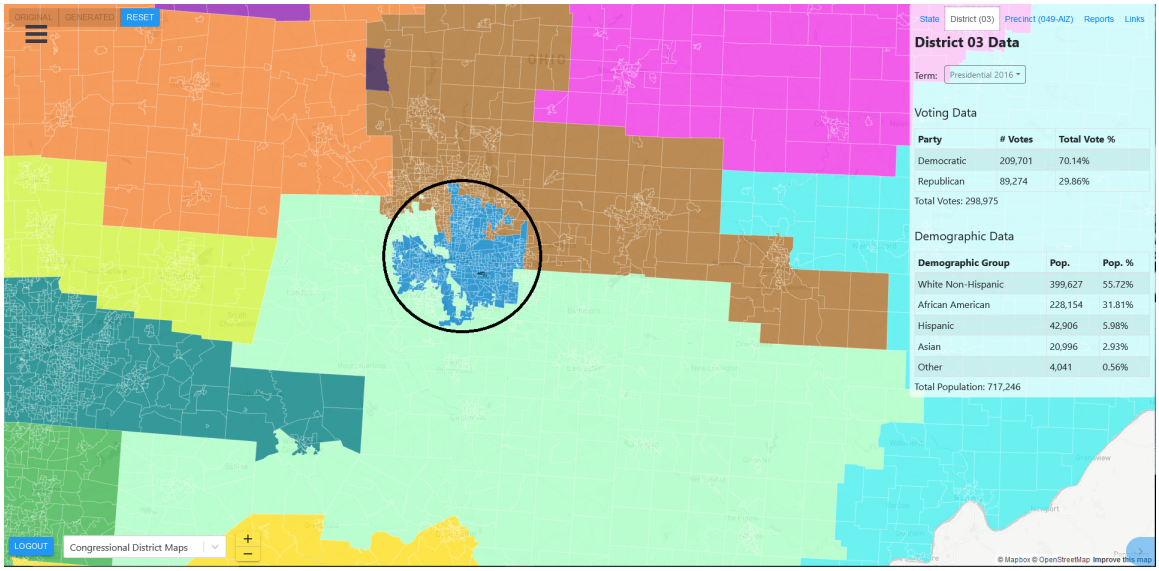

Figure 8: OH-03 Bounding Circle

We will focus on Ohio's 3rd congressional district. It was named in a lawsuit specifically for gerrymandering. This is where our Circle Fatness measure seems to fail as it gives a high score to the district while Polsby-Popper gives it a low score. Polsby-Popper's low score can be explained by the large perimeter that OH-03 has relative to its area. The fact of the matter is OH-03 is incredibly fractured on the city of Columbus. The fact that OH-03 is centered around the city of Columbus is also the reason why it scores highly in the Circle Fatness measure.

OH-11, the one with most of Cleveland, is heavily Democratic and connects two urban areas. The surrounding districts are barely Republican, and neighborhoods were carved carefully to ensure that this district would encompass as many Democrats as possible. Since Ohio's map was drawn by a Republican state legislature, it is quite safe to assume that this is an attempt at gerrymandering. OH-03, in the heart of Columbus, suffers from the same issue: it is surrounded by two moderately Republican districts, OH-12 and OH-15, meaning that there is only one Democratic district in Columbus when there could be two with a different drawing. While a potential excuse could be that this mostly follows the Columbus metropolitan area, it does not represent the will of the voters in the greater Columbus area quite fairly.

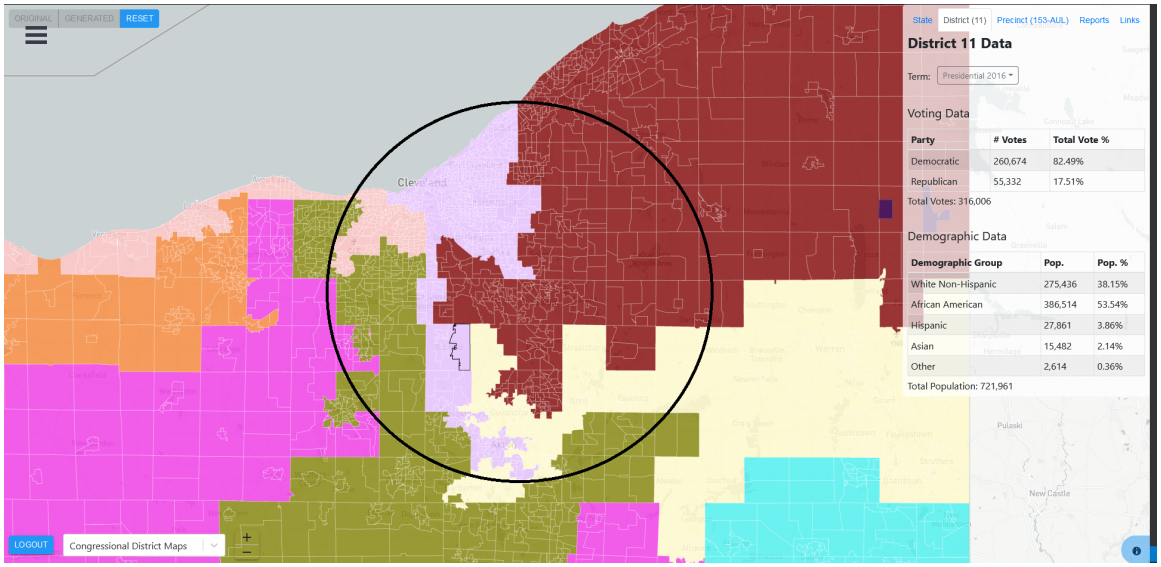

Figure 9: OH-11 Bounding Circle

We can now skip ahead to the third quadrant. These are the districts often pointed out by opponents of gerrymandering as being "bad", failing most standard measures for compactness.

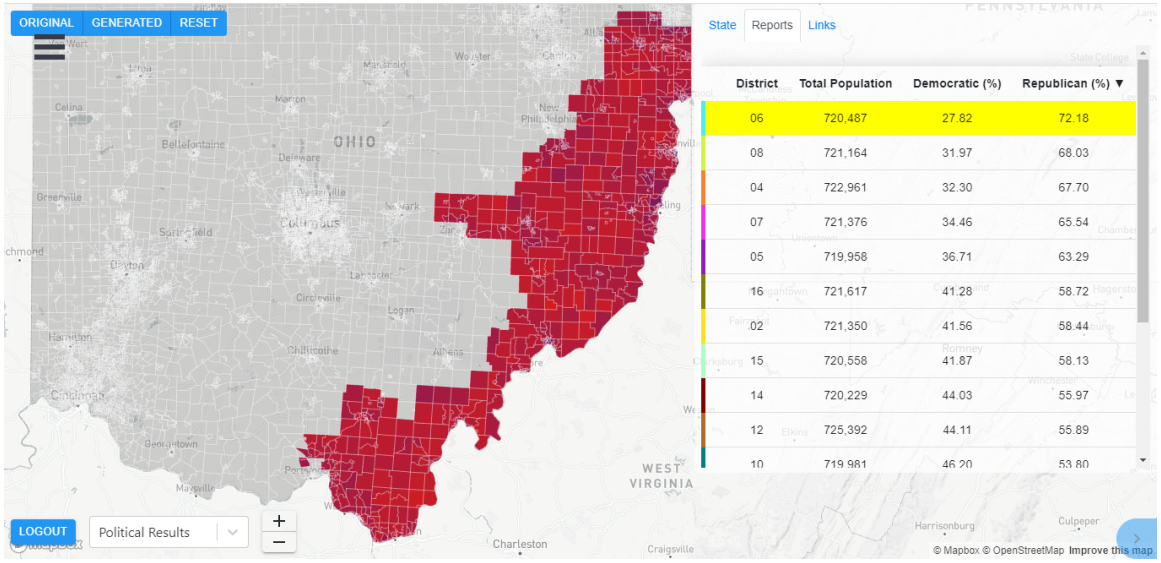

Figure 10: Ohio District 06

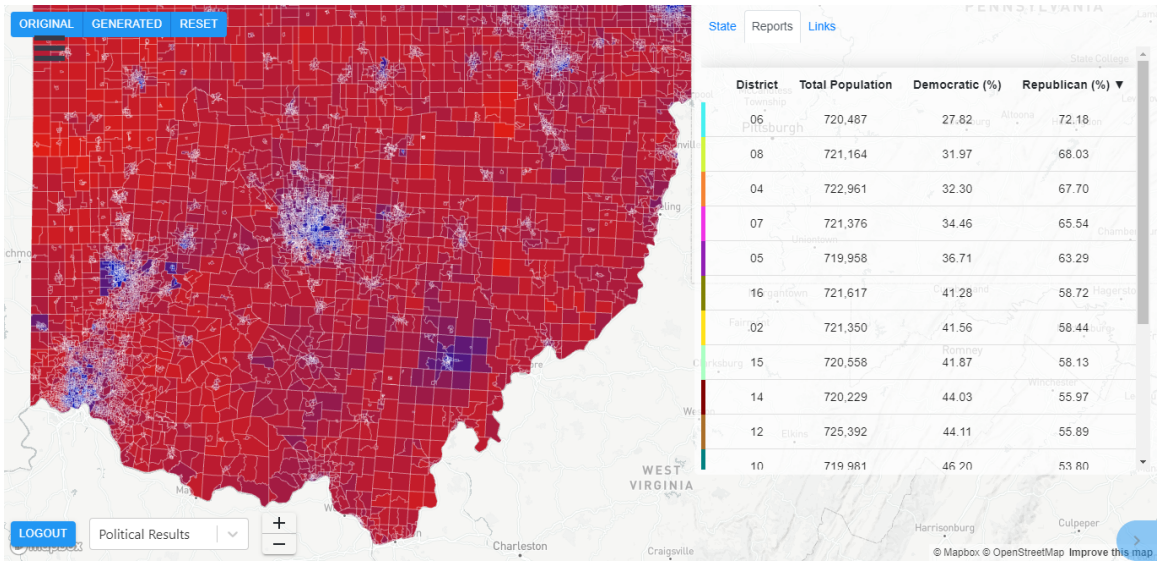

Figure 11: OH-06 Political Results

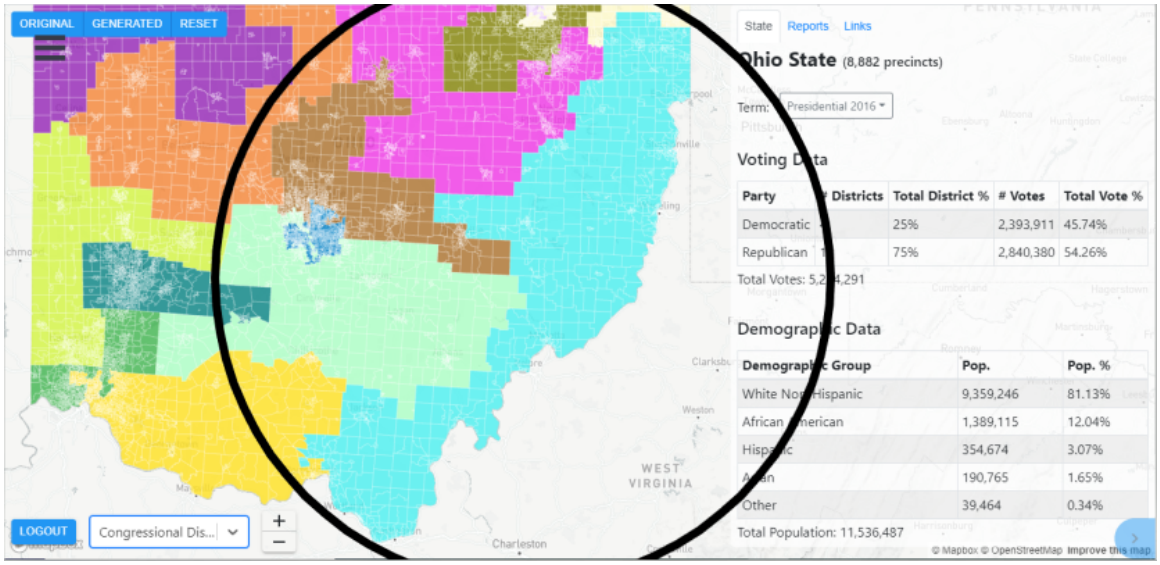

Figure 12: OH-06 Bounding Circle

OH-06 is a good example of this: it covers all of Southeast Ohio, but is unnecessarily stretched out, with several indents both ways from neighboring districts. For example, Athens or Youngstown, places that lean Democratic, are just barely out of the district. Once again in Ohio, OH-04 is also quite serpentine in shape and obviously avoids multiple areas, such as the indent made by OH-05, or the coast that is monopolized by OH-09 (it is worth mentioning that OH-09 is the only Democratic-leaning district of Northwest Ohio). As neighboring districts lean Republican like OH-04, we can see that small portions of urban areas such as Lorain or Strongsville were incorporated, but are not enough to switch the vote. What is currently Rep. Jim Jordan's district clearly could instead have taken the large swaths of Republican land that create a hole in the middle of the district. It's too

bad this district is not in the other state that begins with an O, as this "duck" district would have been a good explanation for that state's university's mascot.

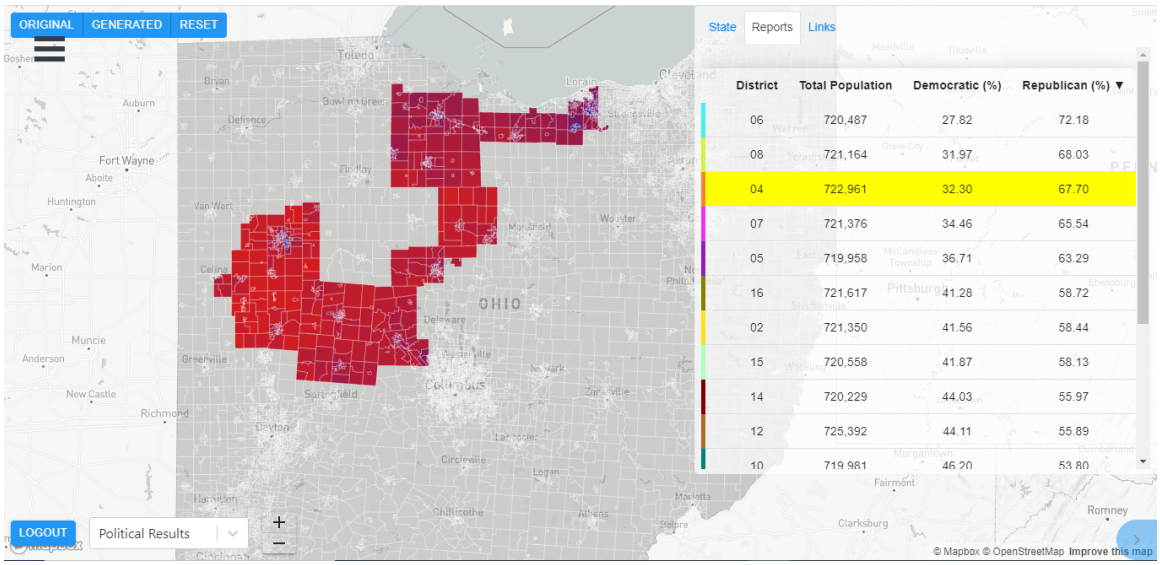

Figure 13: Ohio District 04

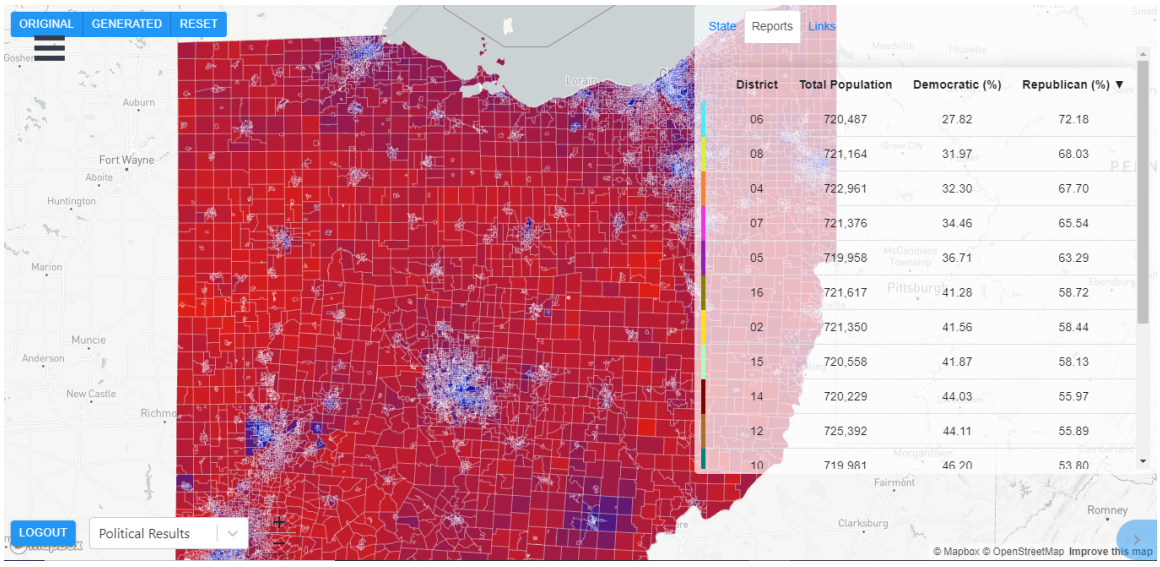

Figure 14: OH-04 Political Results

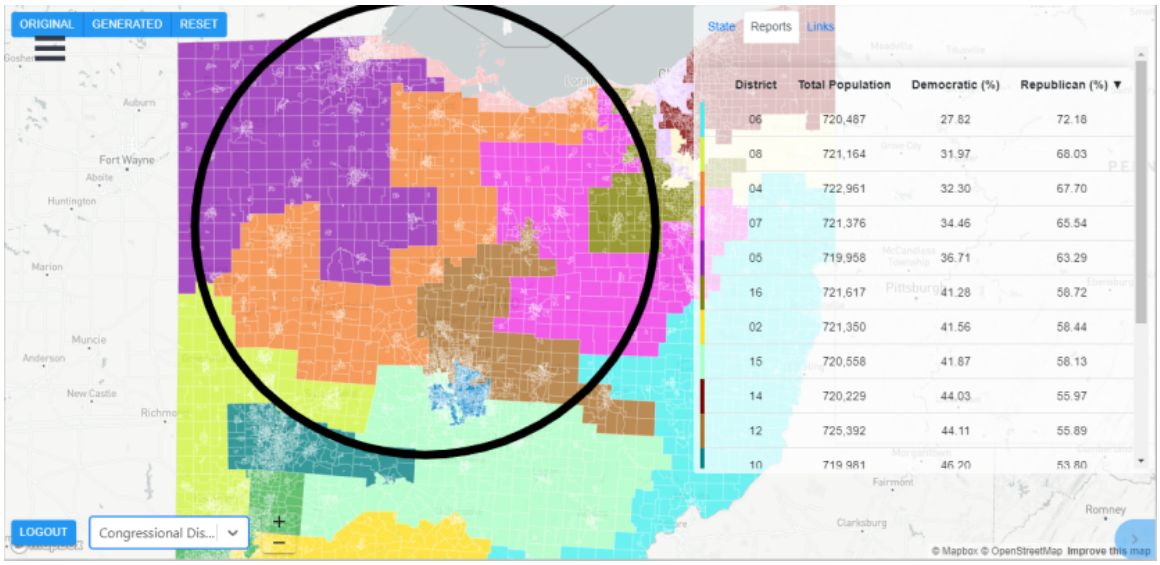

Figure 15: OH-04 Bounding Circle

As an aside: OH-04 has a very large prison and thousands of African American inmates that cannot vote. This also clearly does not respect county lines in the East (but look at OH-11, and Summit County, with four districts and no Congress representative...) https://www.wksu.org/government-politics/2019-11-15/how-did-ohios-most-liberal-city-end-up-with-its-most-conservative-congressman

We will finally take a look at the fourth and last quadrant of the plot. It is sparcely populated and represents districts that score high in Polsby-Popper but low in Circle Fatness. We will look at NY-17, NY-18, NY-21, PA-08, and GA-09.

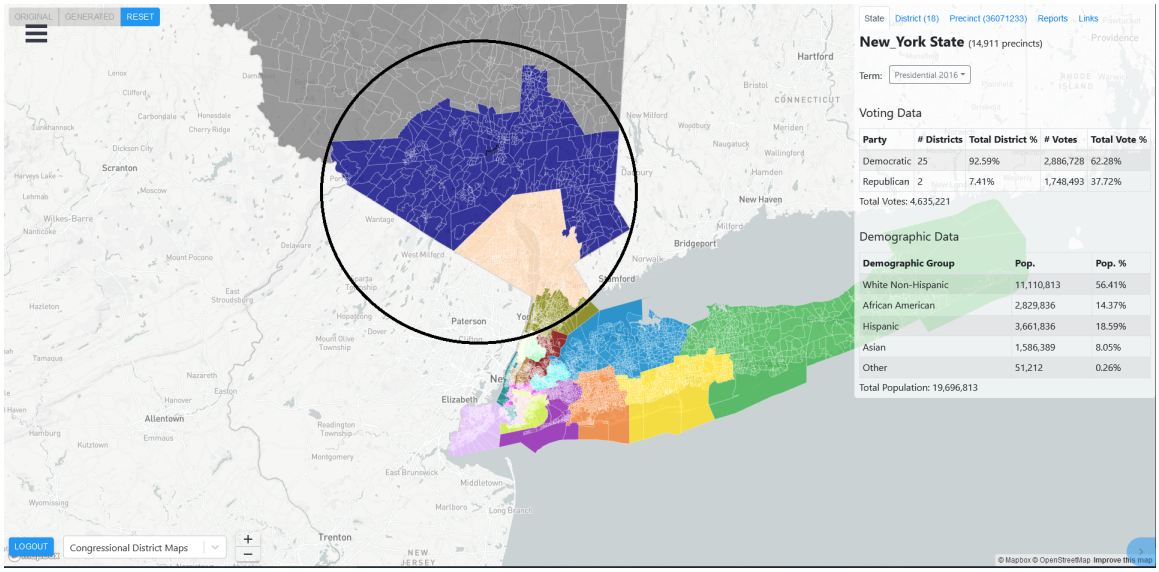

Figure 16: NY-18 Bounding Circle

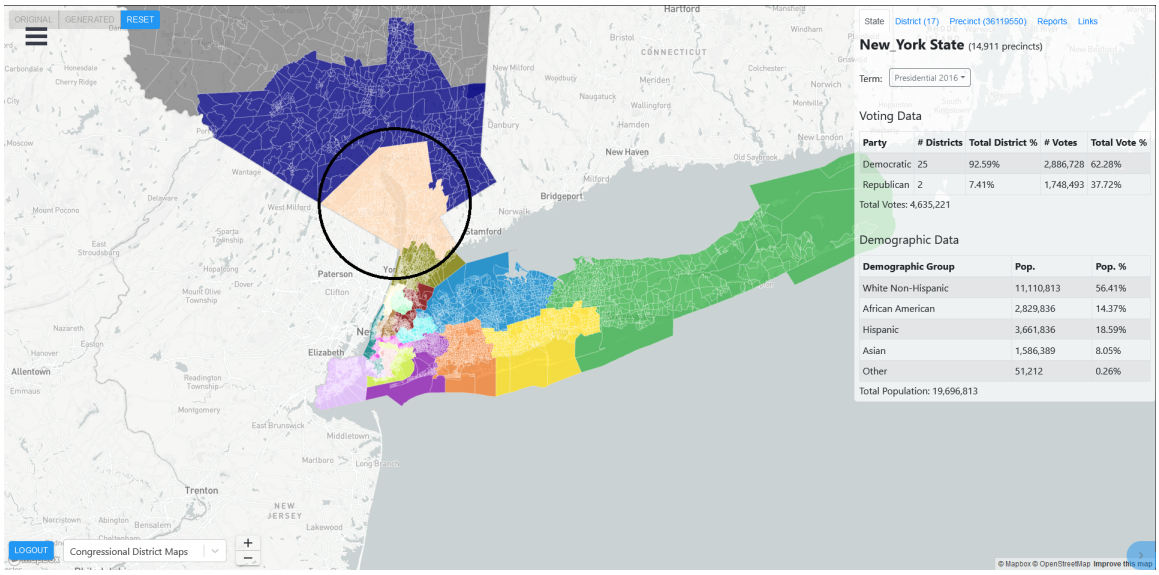

Figure 17: NY-17 Bounding Circle

These two districts from New York have low fatness scores yet high Polsby-Popper scores due to the unique shape of the state. These two districts are "funnel" shaped towards one of the most populous city in the entire country. Therefore, parts of the bounding circle enroach upon the city and its outer suburbs. This causes the fatness score to be low since it "misses" the more populous precincts of the state.

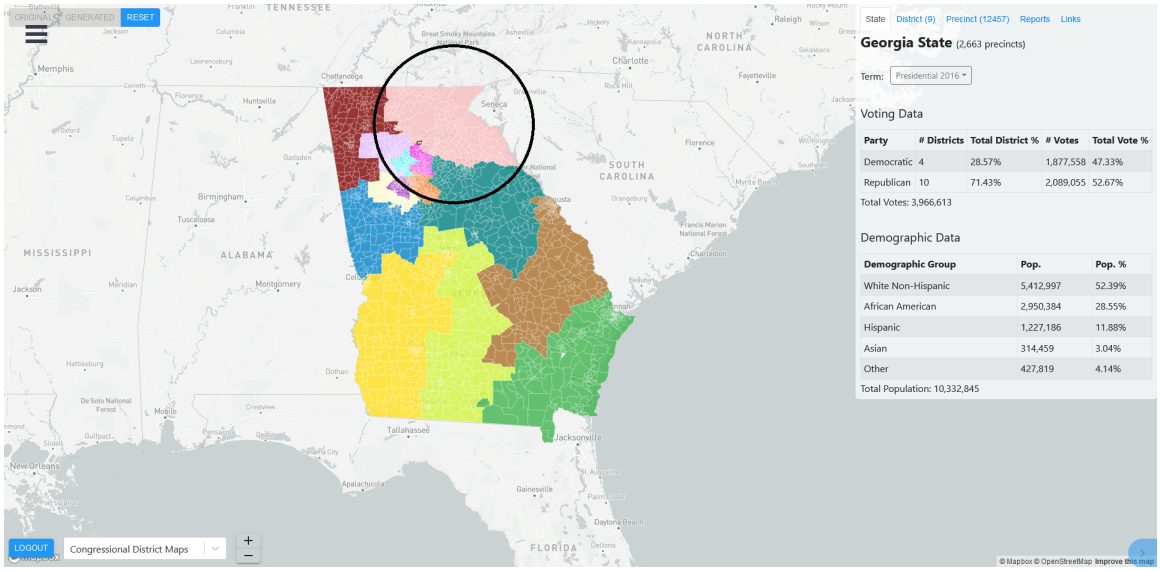

Figure 18: GA-09 Bounding Circle

The same situation can be applied to Georgia's 9th congressional district. The "long" shape of GA-09 going from northwest to southeast causes its bounding circle to contain the city of Atlanta. It is again a major population source. One thing to not is that although Atlanta predominately

leans towards the Democratic Party, GA-09 consistently leans Republican. It is one of the most pro-republican districts in the state of Georgia.

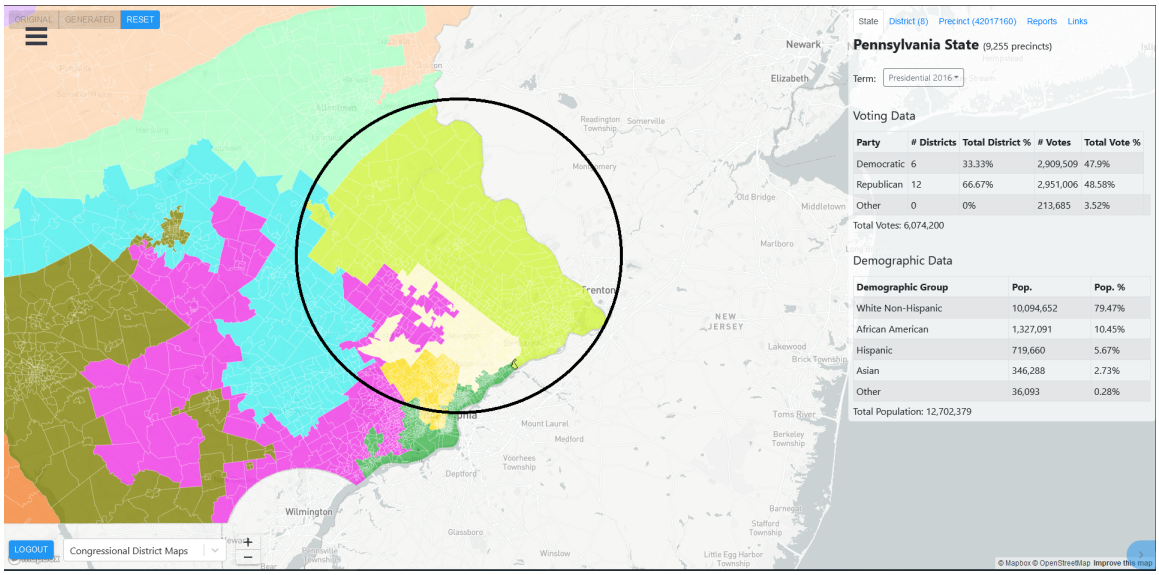

Figure 19: PA-08 Bounding Circle

Finally, Pennsylvania's 8th congressional district tells us the same story as the past three districts. As one can see, PA-08 sits right outside of Philidelphia, but its bounding circle does not. This again causes the high population of Philidelphia to be just outside of the district's borders but within the bounding circle, thus causing the low fatness score. It is a recurring theme in this quadrant of the graph that these districts seem to be just outside of major population zones and is part of the state's borders.

6 Improvements

At this point it is easy to suggest improvements to our original fatness measure. Instead of calculating population of the minimum bounding circle for the district, one can argue that a more geometric compact measure would calculate the population of the convex hull of the district as the denominator. The changed formula would be the following with $CH(D)$ obviously being the convex hull of a district:

$$\frac{POP(\mathrm{D})}{POP(\mathrm{CH(D)})}$$

We will refer to this measure as the Convex Hull Fatness. The idea is the Circle Fatness is too susceptible to the shape of the state, most notably the funnel shape discussed with NY-17 and NY-18. Below is our results from that thought.

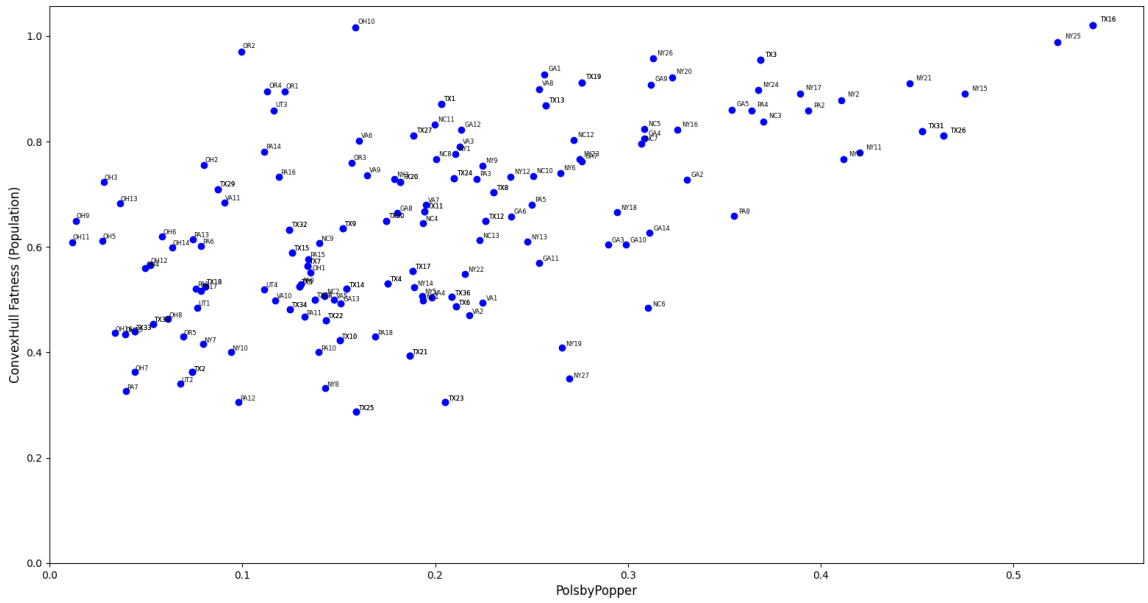

Figure 20: ConvexHull Population Fatness vs PolsbyPopper

The most notable change is that the Convex Hull Fatness score is much overall higher than the Circle Fatness score. Furthermore, it is to be noted that if you split the plot into four quadrants again, the fourth quadrant is now almost entirely empty. This proves that the Convex Hull Fatness measure is less susceptible to the funnel shape districts that were present in New York. Furthermore, looking at GA-09 geometrically, it is obviously a "compact" corner district that was simply too close to the city of Alanta.

However, there are many more new additions to the second quadrant. We will focus on OR-02 and OH-10 as these two districts had the highest increase in fatness score.

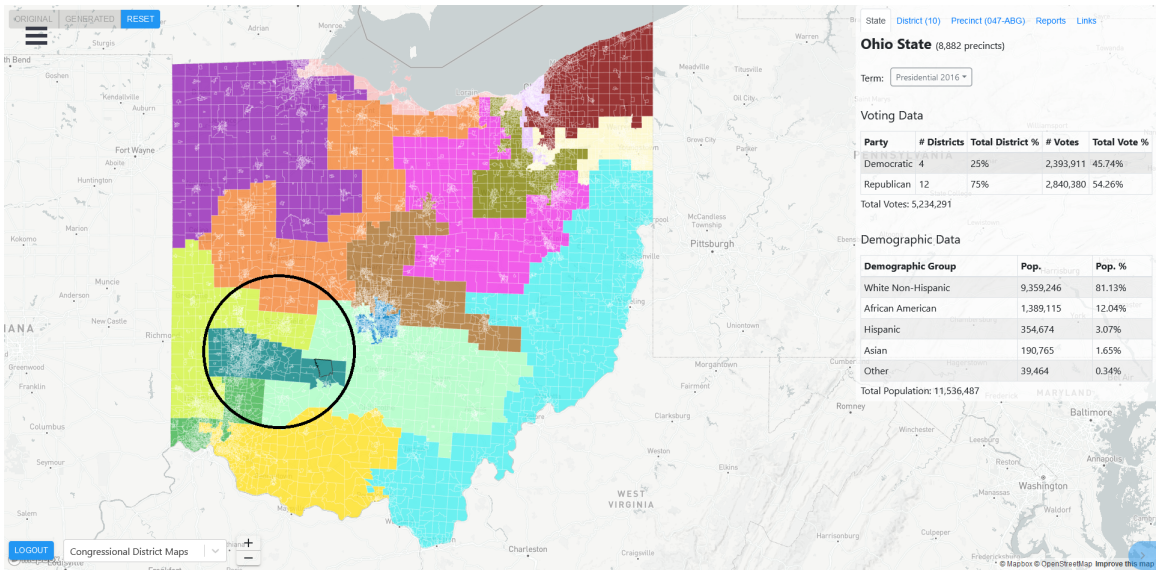

Figure 21: OH-10 Bounding Circle

For OH-10, the Polsby-Popper measure rated this district rather low due to its long rectangle shape. Obviously, with the convex hull fatness, the score is higher as the only major difference is the removal the staircase shaped northern border of the district. Looking at the district geometrically, by itself, it would be hard to justify malicious intent.

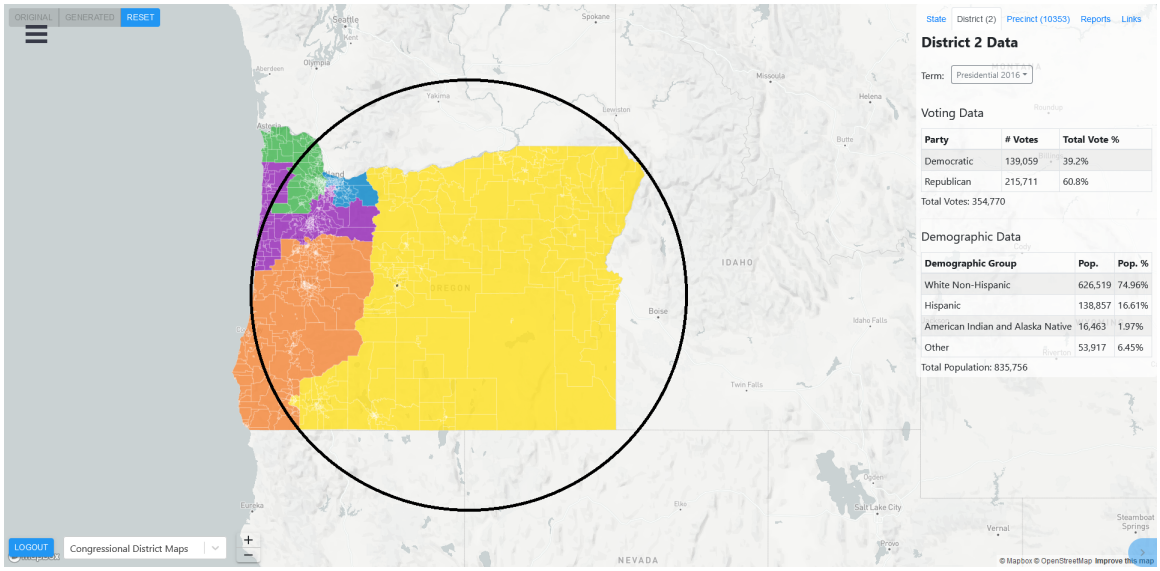

Figure 22: OR-02 Bounding Circle

A similar case can be made for OR-02. Its Circle Fatness score and Polsby-Popper score were both low, landing it in the third quadrant in the original plot. However, now it is situated comfortable in the second quadrant where the Polsby-Popper score is low and the Convex Hull Fatness score is high. We believe that this is a more accurate measure of the compactness of this district. The minimum

bounding circle nearly covers the entire state, including the most populous non coastal sections of the state. There are some theories to explain the low Polsby-Popper score. If one looks closely at the border of the district, specifically the western and northern half, it is made up of jagged edges. The jagged edges constitute the district following a border represented by rivers to the north and northeast. The western portion is simply following the precincts that have the jagged edges. The many jagged edges increase the perimeter portion of the Polsby-Popper measure, therby resulting in a low score for this district. In essence, the convex hull measure smooths out or fills in the jagged edges of a district, thereby resulting in a higher score.

The Convex Hull Fatness is not without its own faults. For example, OH-06 has scored much higher than in the Circle Fatness measure. An explaination of why it was considered a bad district was elaborated on previously. However, due to the new measure of fatness, it follows the same reasoning as OH-10. The difference here is that OH-10 does not seem to avoid any major population centers. However, an arguement can be made that OH-06 avoids the Democratic leaning city of Athens. It is not the most egregious district in terms of compactness but it seems to be elongated with intent. A better example of this comes with TX-15 and TX-34 which both have similar reasonings as OH-06.

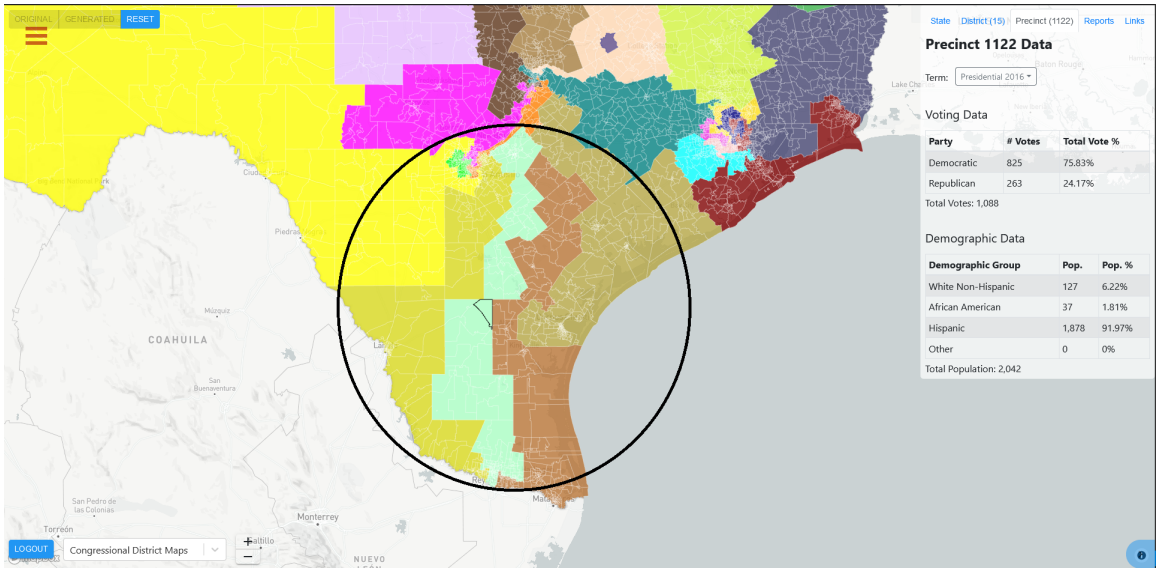

Figure 23: TX-15 Bounding Circle

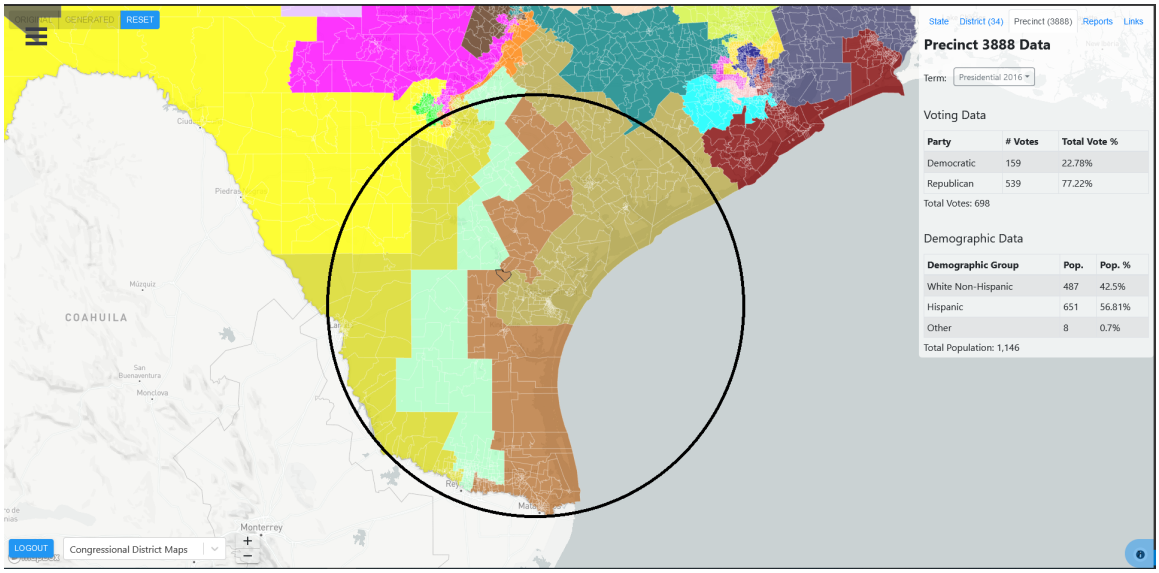

Figure 24: TX-34 Bounding Circle

A better districting would have been to cornerize at least one of these districts, but it seems that Texas has an over-abundance of these elongated districts. In addition, not only are these districts elongated, they seem to have bottlenecks in them for the sake of lengthening the district. Specifically Precinct 3888 and Preinct 1122 which are both highlighted in the figures. The reasoning behind the lengthening is unknown but it does seem to be done with intent.

7 Future Considerations

With all of these case studies, there are quite a few things we can say about these measures and their strengths and weaknesses compared to human perception - which is supposed to be the ultimate judge.

Polsby-Popper is good at catching districts that have unnecessarily long boundaries by penalizing them. As a result, districts with long straight lines tend to be favored. Something that is both positive and negative is that it does not depend on what is going on outside of the district, which means that we cannot see if population groups are avoided, but on the other hand, districts are not graded based on their size as a dilation of district by any strictly positive constant still yields the same score. However, sometimes, the border looks straight from a distance but is not on closer inspection: this can give unnecessarily low scores, unless the border is smoothened, a less straightforward and undebatable process than it may appear.

The Circle Fatness measure devised, on the other hand, will "punish" long districts, and offers a solution to the issue of Polsby-Popper not taking into account what is happening right outside of the district. However, it has inherent downsides. For instance, as long as a bounding circle is small enough, the score will remain high: that is the case for OH-03 where several areas are removed from the bounding circle deeply down towards the middle, but this has a relatively limited effect on the fatness score. The opposite is true as well, the Circle Fatness is too susceptible when the district is elongated in one direction, or just covers a large area due to the population requirements.

The Convex Hull Fatness improves on the Circle Fatness' faults but it leaves things to be desired. It increases the low scores given to the elongated districts like OH-10. But there are reservations

on whether this should apply to all elongated districts. Specifically, districts with a single precinct serving to create a bottleneck for the sake of elongating a district. Examples include TX-15 and TX-34.

As a result, we suggest that there could instead be a computation of the arithmetic or geometric mean of the fatness scores, and that good plans should score above a certain threshold. The arithmetic mean (adding all scores) would favor good outliers in the state, and the geometric mean (multiplying all some) would strongly hinder bad outliers, so both systems have their (different) merits. More importantly, population fatness favors smaller districts, and those surrounding less populated areas. We saw this with examples such as TX-16 (El Paso is in the middle of the desert), OH-03 (with a good proportion of Columbus, with few of the suburbs concerned). However, larger districts tend to have more districts in their bounding circle, which automatically drop the population fatness score. In fact, this might create a political bias, since in general, Democratic congressional districts are more urban and smaller than their Republican counterparts. If we minimize the number of "nonfat" districts, we might crack Republican districts, but if we maximize the number of "fat" districts, we would pack Democrats. There are still many more questions to be answered, but as in other instances, they might show political bias.

References

Wasserman, D., & Flinn, A. (2017). *Introducing the 2017 Cook Political Report Partisan Voter Index* (tech. rep.). The Cook Political Report.