

# **Politech Mathematical Group**

## Political Fatness Report

Darren Kong  
110770716

Hugo Mainguy  
111747982

Jeffrey Zhong  
112299792

Feb 1, 2021

## 1 Abstract

## 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 (Cook) such that there is less need to appeal to the other side. 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 as of today. Furthermore, state constitutions tend to be even further behind, because of the amount of work 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 for an object D by:

$$4 * \pi * A(D)) / (P(D)^2$$

Where  $A(D)$  is the area of the object  $D$  and  $P(D)$  its perimeter. This gives us a ratio between 0 and 1, and is 0 precisely when the area is 0 -  $D$  is a line - and 1 when  $D$  is a circle. In particular, this favors 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.

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. To calculate this measure, we use the following formula:

$$\frac{pop(DISTRICT)}{pop(BOUNDING CIRCLE)}$$

Where  $pop(x)$  is the population in  $x$ . 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. This 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 comes when the bounding circle encompasses many populous areas, large areas of land, or a combination of both.

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 maps are from 2016).

## 4 Results

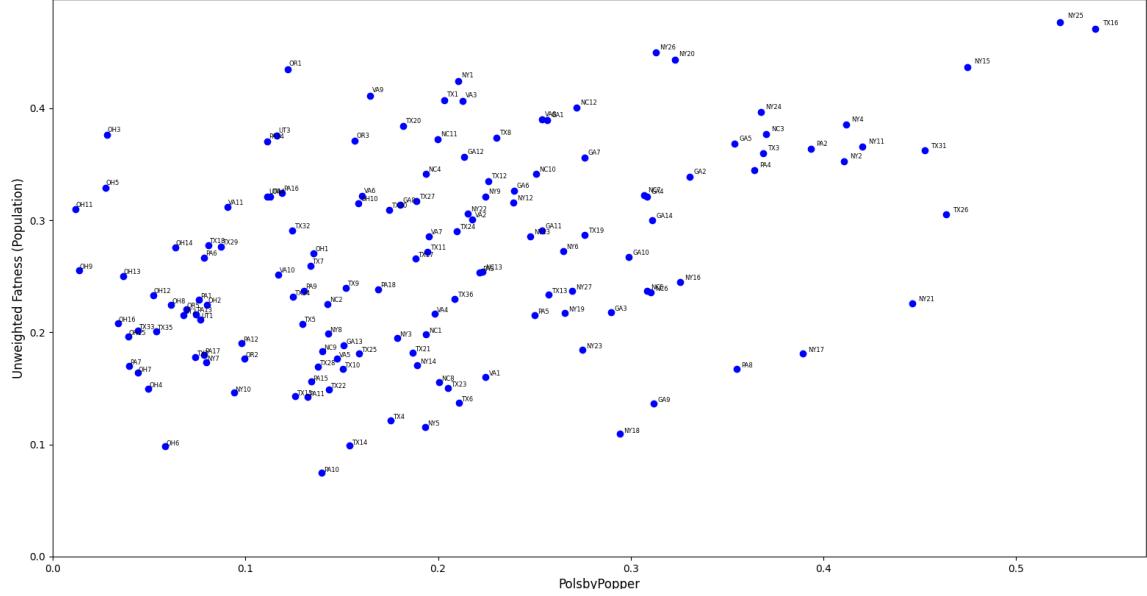


Figure 1: Plot of Population Fatness v. PolsbyPopper

## 5 Research

From the results, we can see compare each district's 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 extremities in each of the district scores to see in what cases these two scores may differ wildly.

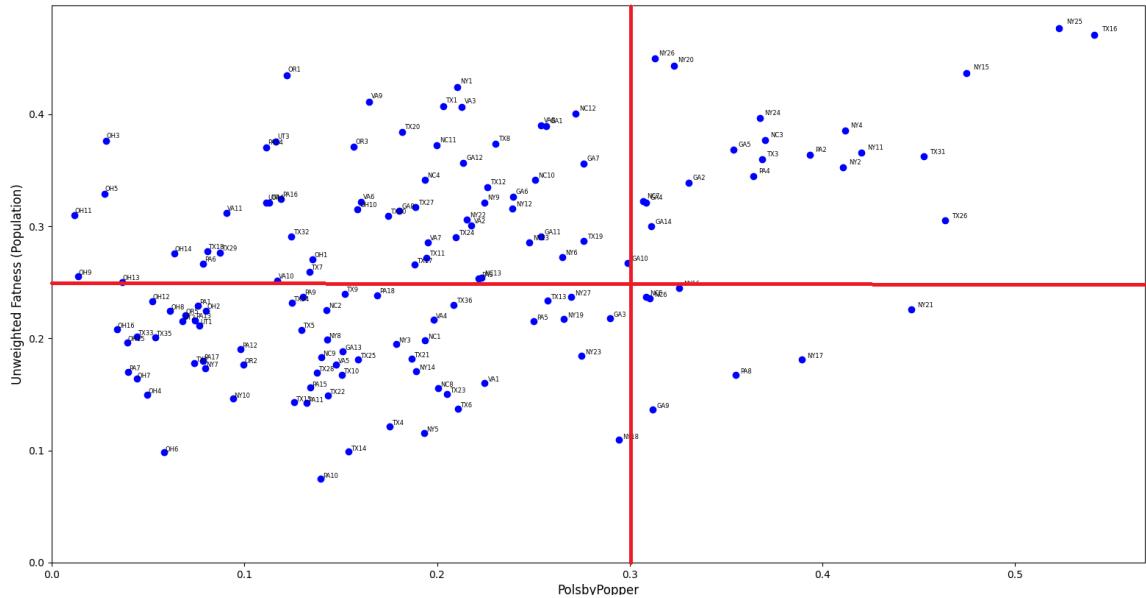


Figure 2: Plot of Population Fatness v. PolsbyPopper Edited

We will begin our look on the first quadrant, specifically NY25, NY15, and TX16. The district scores for both Fatness and Polsby Popper are both high for these districts. It is rather easy to tell why.

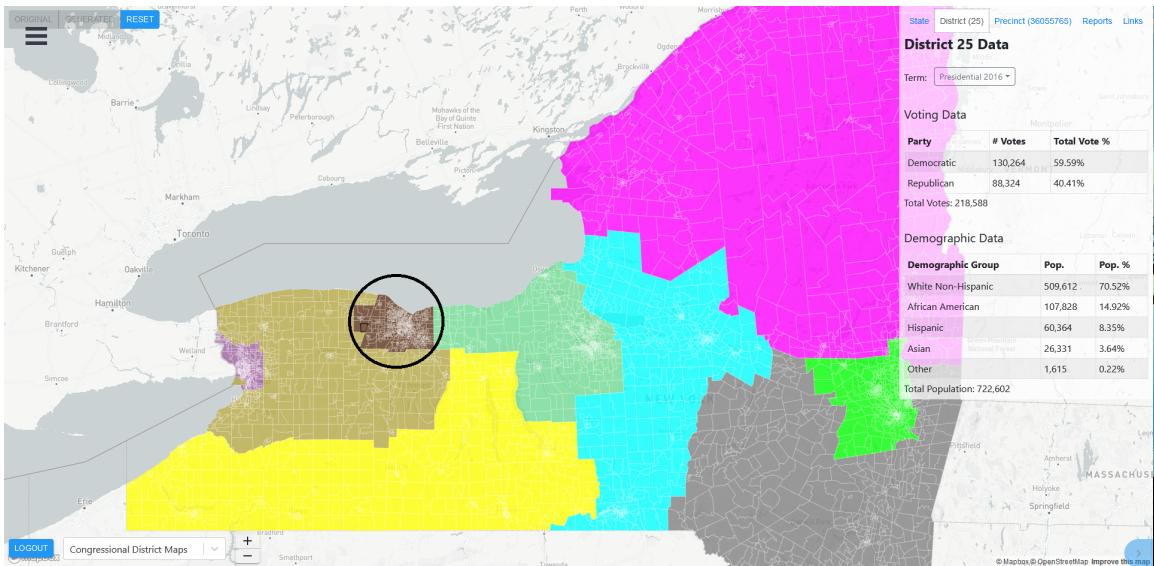


Figure 3: NY-25 Bounding Circle

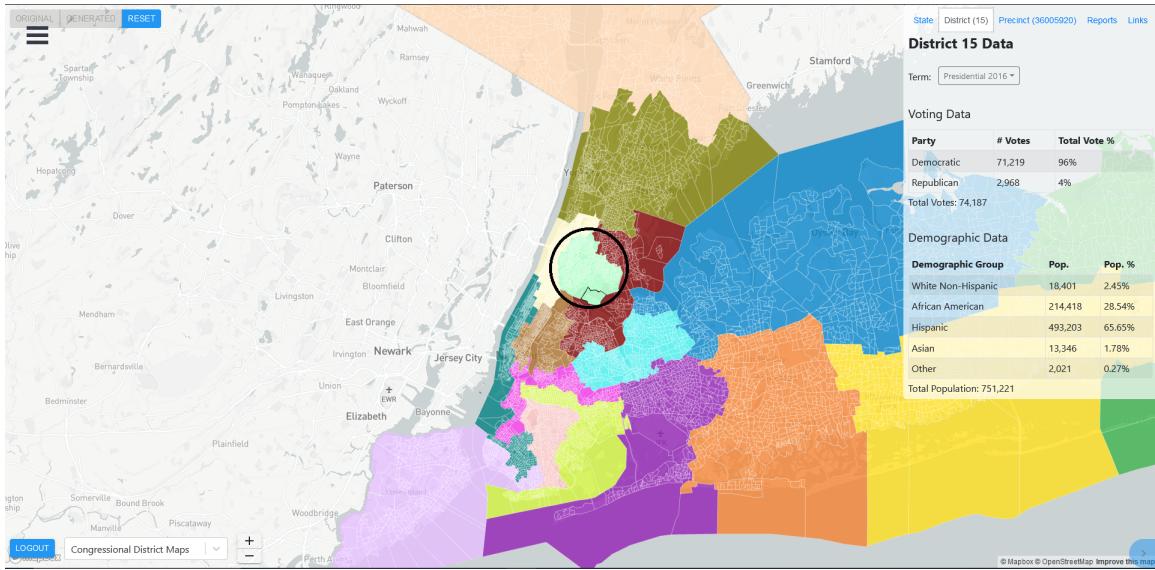


Figure 4: NY-15 Bounding Circle

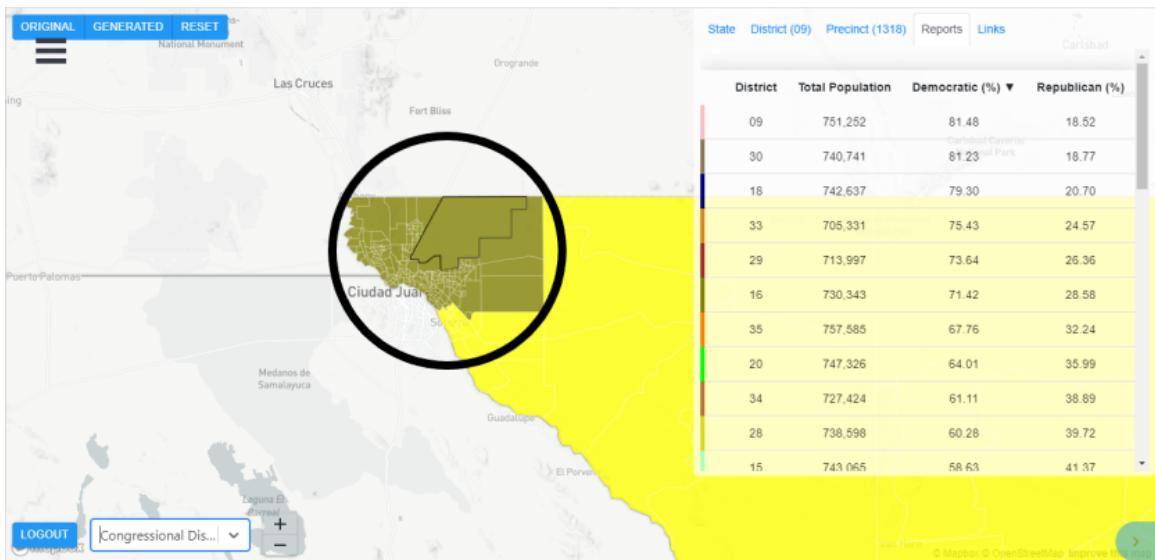


Figure 5: TX-16 Bounding Circle

These districts are all centered around a major population hub. NY25 is centered on Rochester, NY15 is centered on the Bronx, and TX16 is centered on El Paso. These districts geometrically all look compact and we believe that no one would consider them abnormal in any sense. This explains the high Polsby Popper score and also partially explains the high Fatness score.

Specifically for the example of TX16 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.

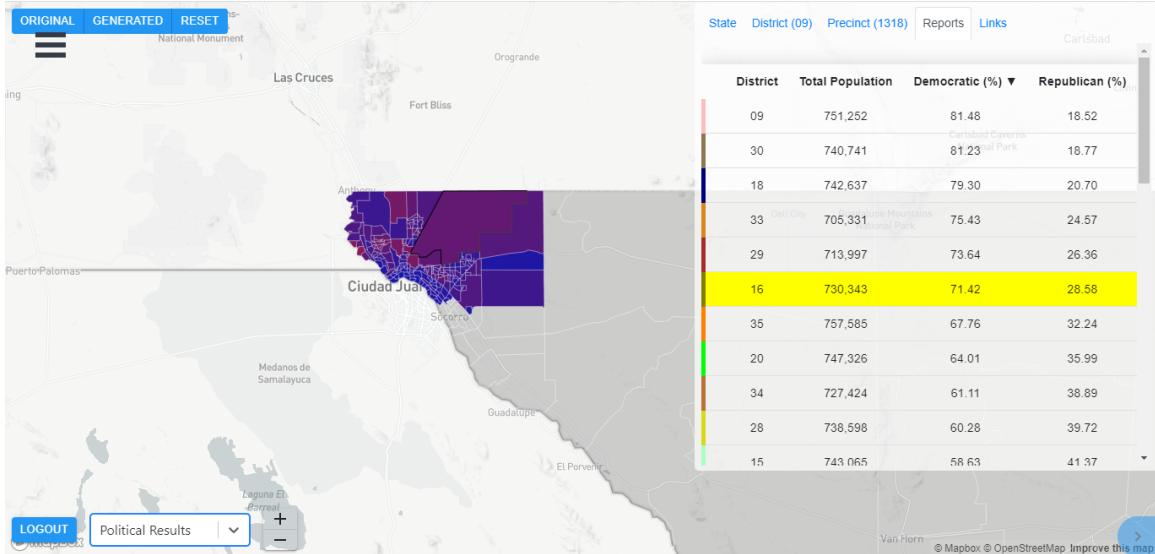


Figure 6: TX-16

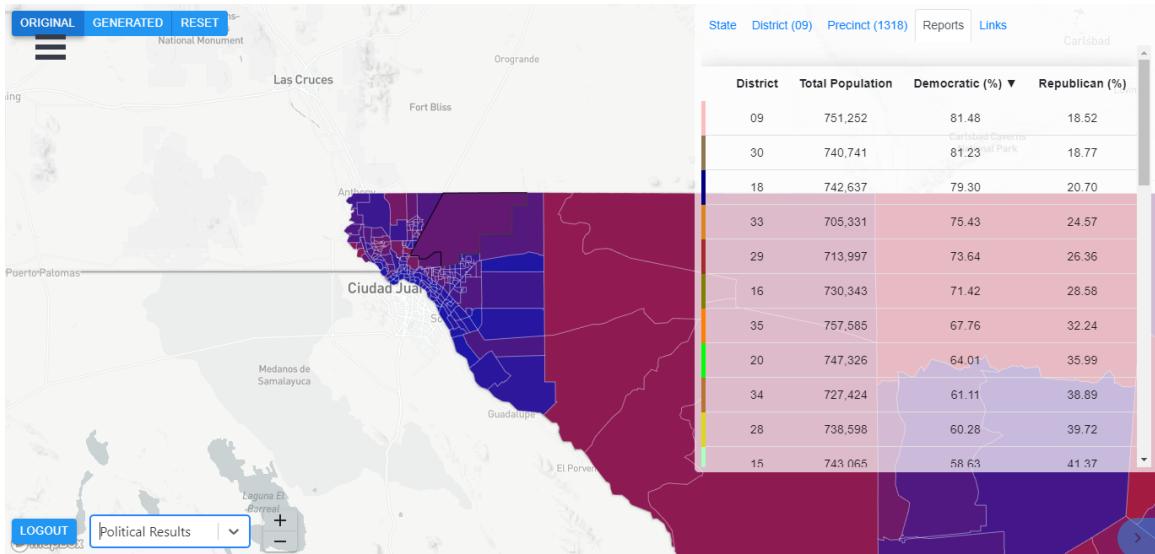


Figure 7: TX-16 Political Results

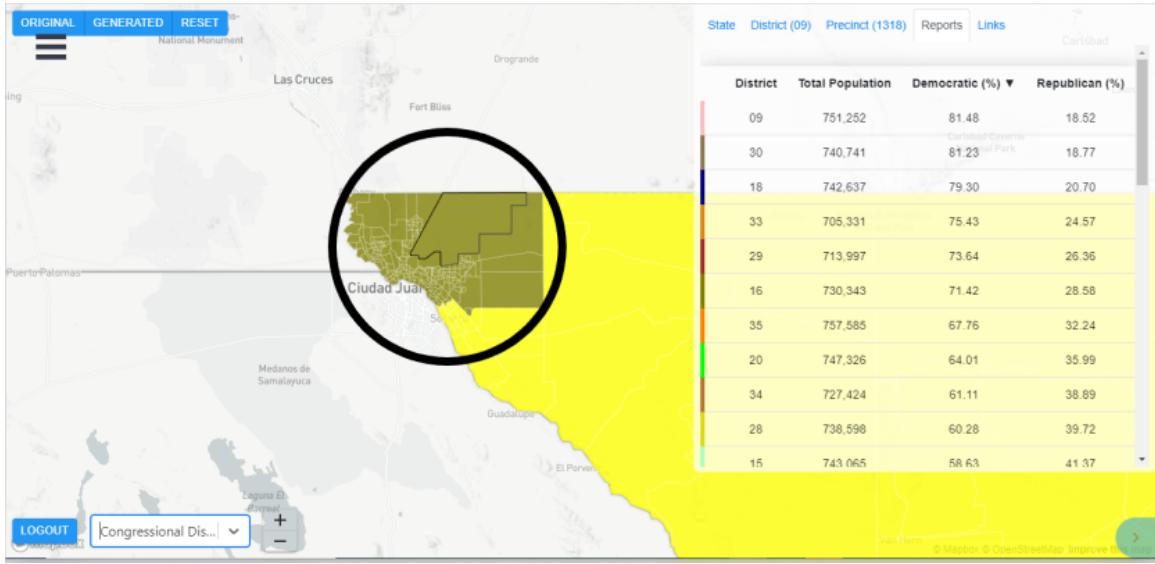


Figure 8: TX-16 Bounding Circle

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

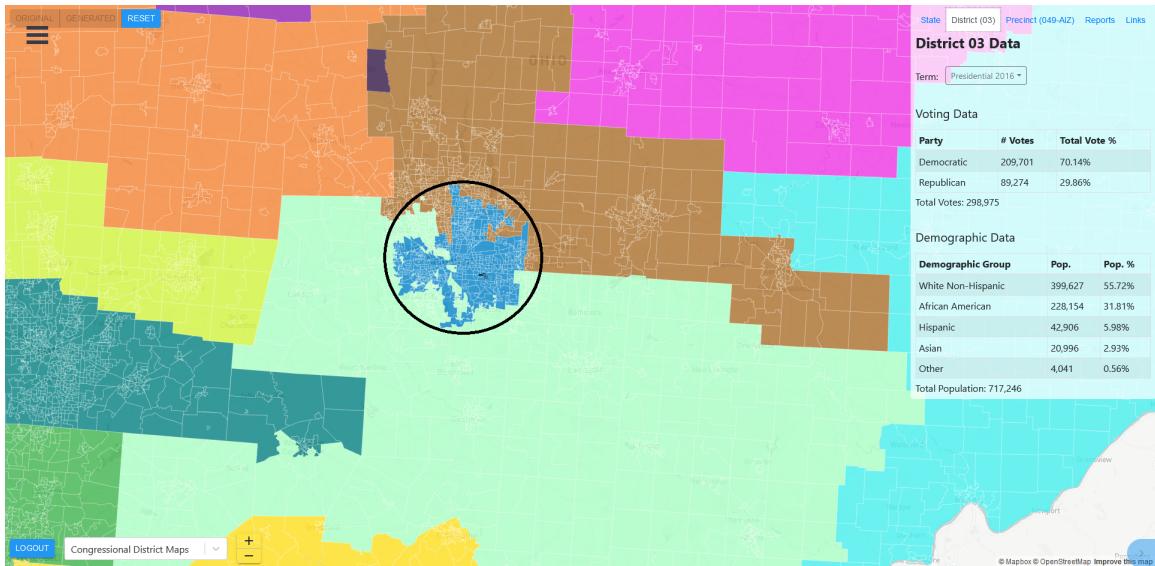


Figure 9: OH-03 Bounding Circle

We will focus on Ohio's 3rd congressional district. It was named in a lawsuit specifically for gerrymandering. This is where our 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 OH03 has relative to its area. The fact of the matter is OH03 is incredibly fractured on the city of Columbus. The fact that OH03 is centered around the city of Columbus is also the reason why it scores highly in the Fatness measure.

Furthermore, if we take a look at Ohio's 11th congressional district, it has a similar Fatness score yet a lower Polsby Popper score. This again is easy to explain. The population of the district contains both the cities of Cleveland and Akron. There are no other major cities matching the population size near them.

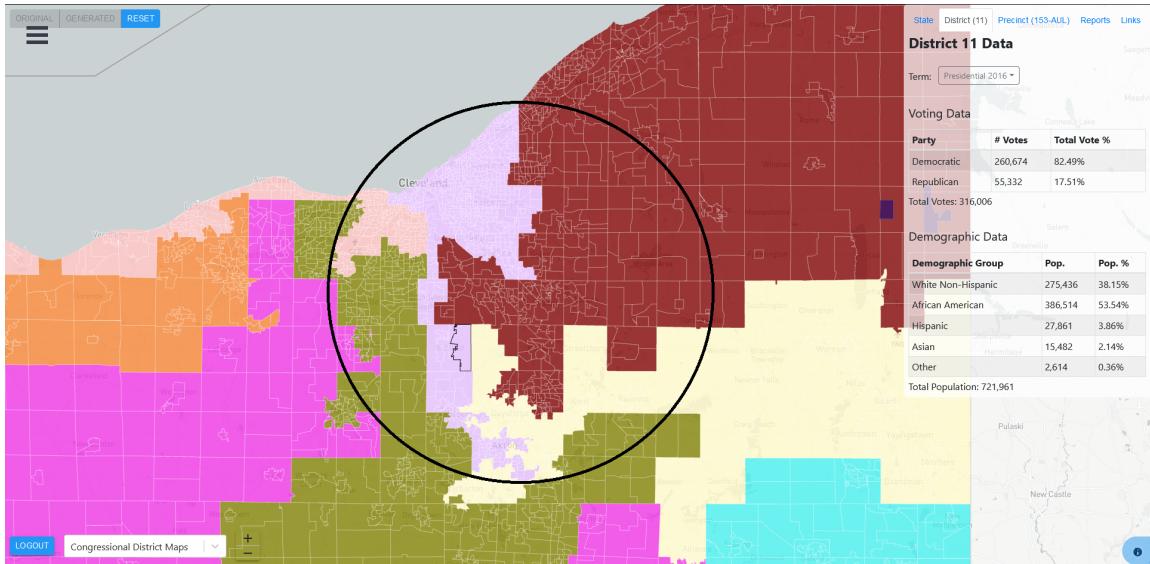


Figure 10: OH-11 Bounding Circle

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.

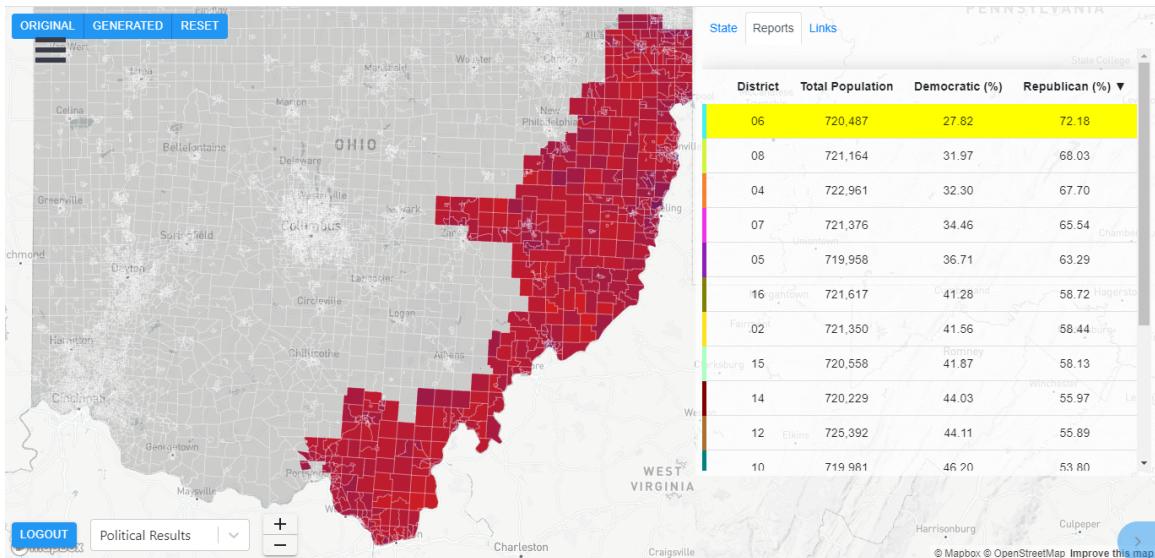


Figure 11: Ohio District 06

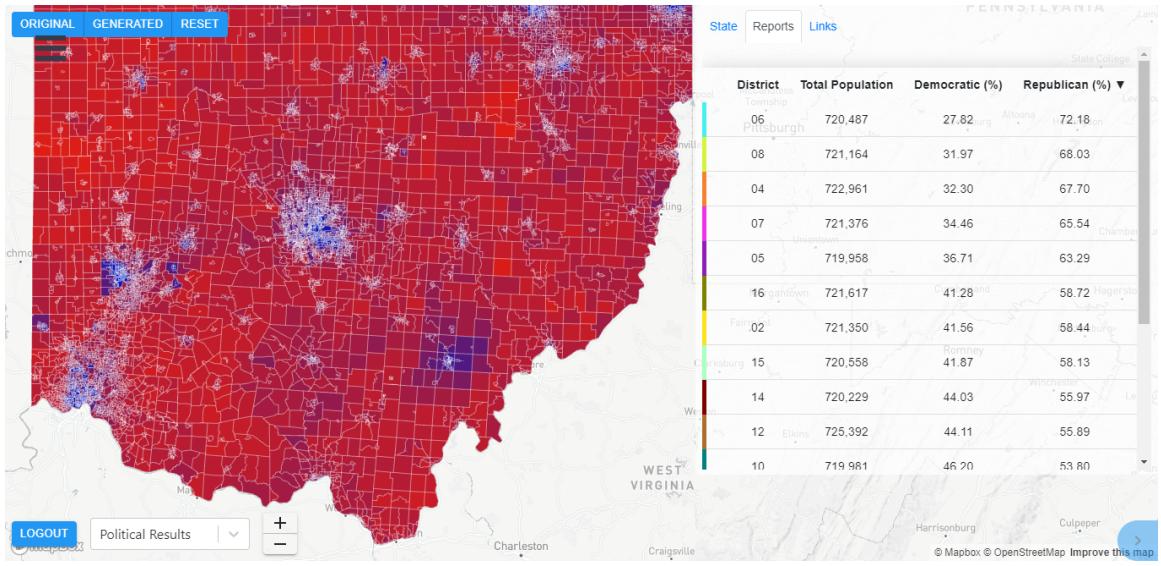


Figure 12: OH-06 Political Results

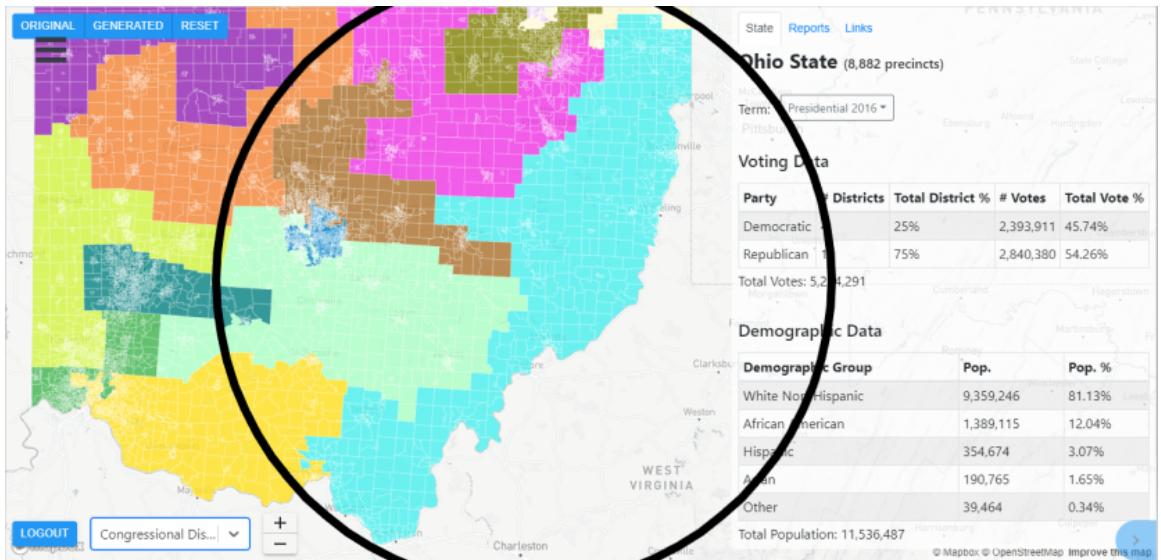


Figure 13: OH-06 Bounding Circle

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 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.

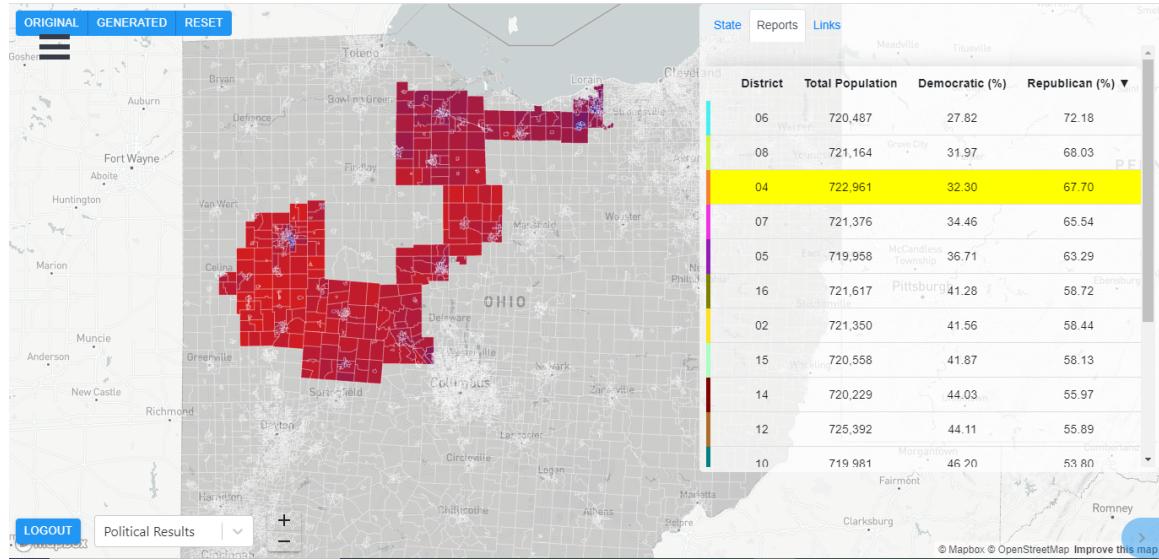


Figure 14: Ohio District 04

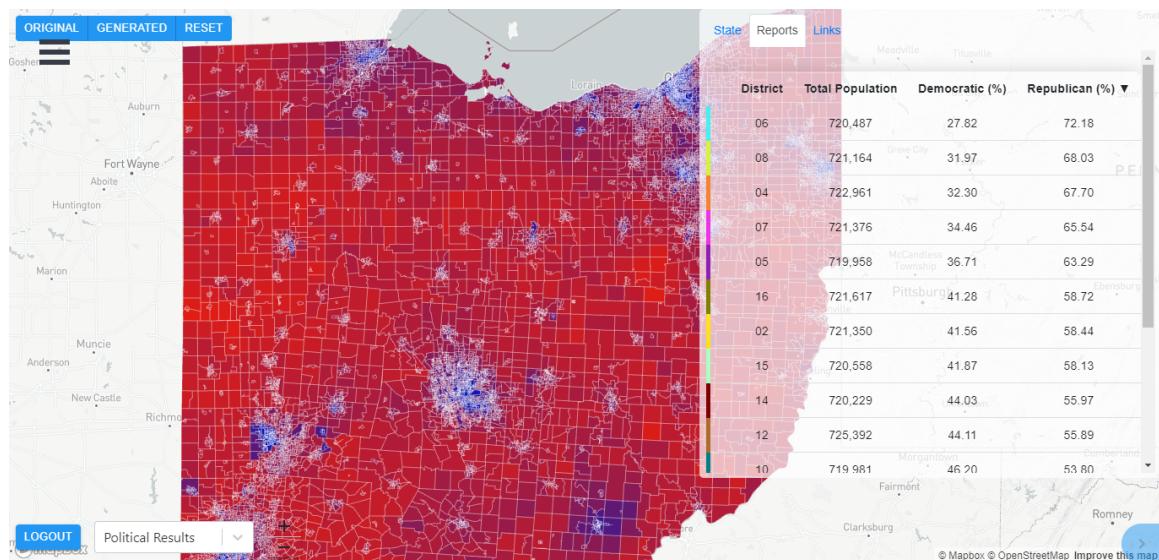


Figure 15: OH-04 Political Results

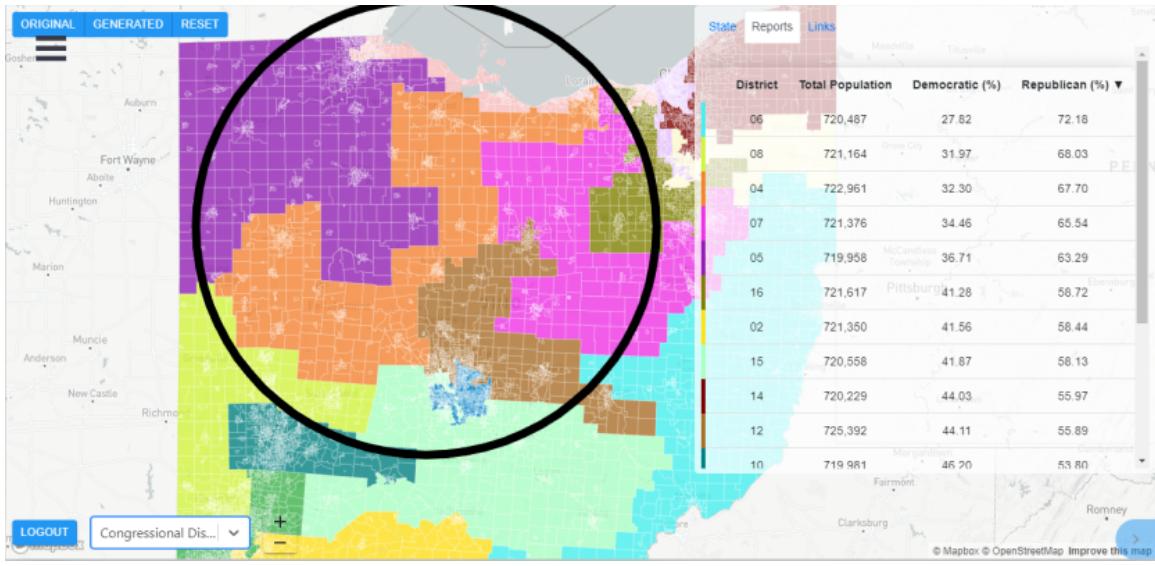


Figure 16: OH-04 Bounding Circle

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>

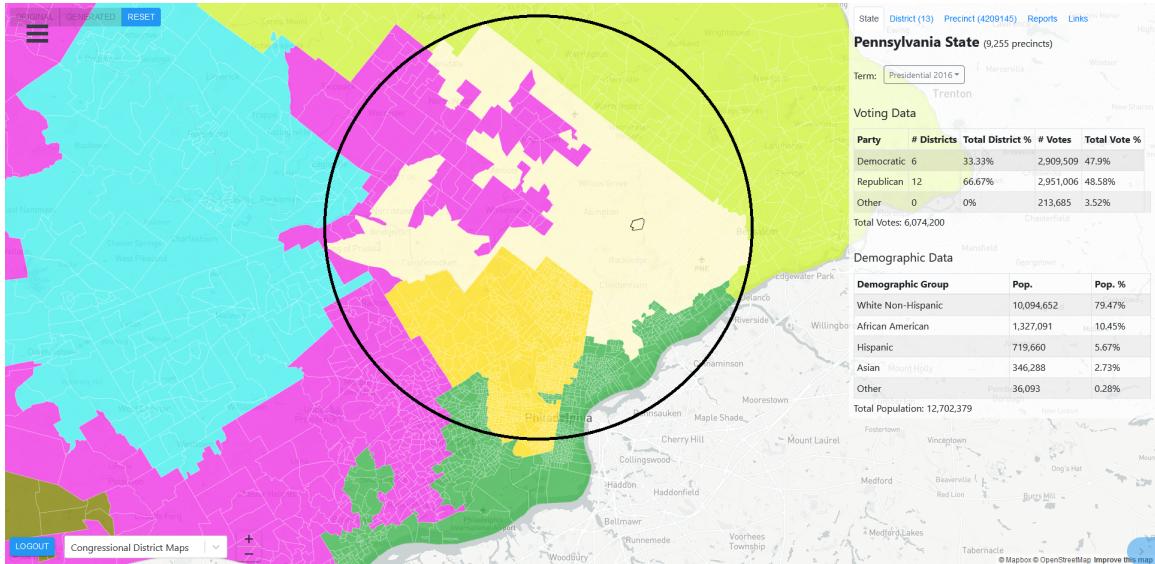


Figure 17: PA-13 Bounding Circle

We will discuss two districts in the fourth quadrant, PA-13 and PA-09. PA-13 lies in the southeastern part of Pennsylvania. It scores a comparatively high Polsby-Popper score, but this is due to the “box-like” nature of the intrusions that it has.. It suffers from its size, and being around slightly more densely populated districts, which lower its population fatness score while maintaining a decent

Polsby-Popper score. Furthermore, the intrusion, specifically from PA-07 into the area to the north and south of Fort Washington causes the low population fatness score.

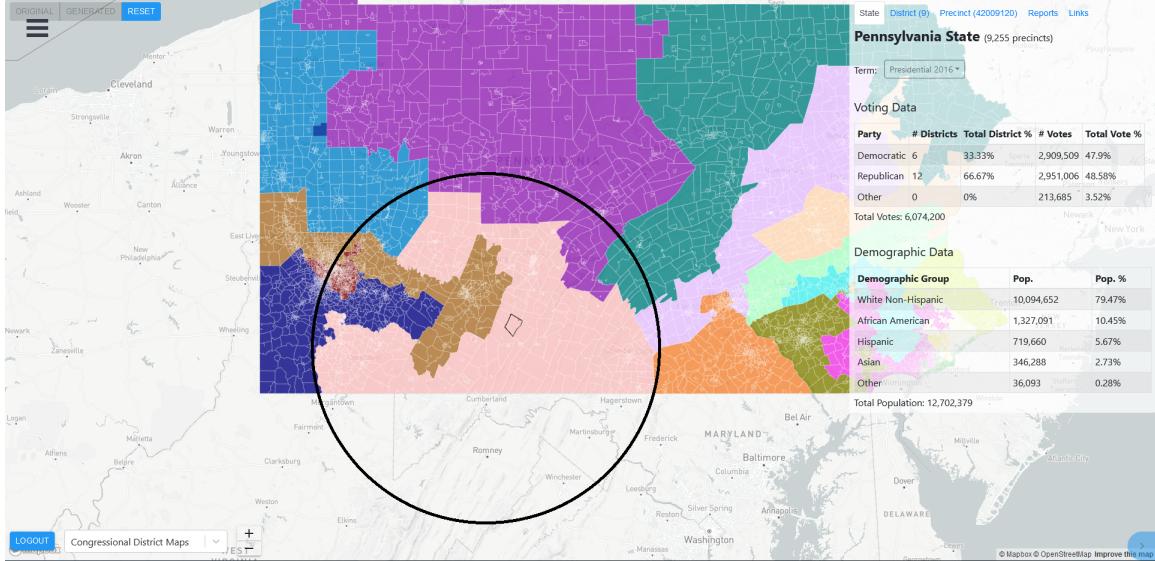


Figure 18: PA-09 Bounding Circle

PA-09 suffers from the same issues. It has a high PolsbyPopper score due to its low perimeter compared to its area yet it scores roughly the same as PA-13 on the population fatness measure. It is being intruded upon by PA-12 and PA-18, giving it a sort of Pac-Man shape.

This map has been redrawn already during the decade, so is less gerrymandered than it used to be, and in the middle of rural Pennsylvania, losing some compactness mostly due to following county lines, this is not the main culprit.

We now move our focus to the second quadrant, which is more populated than the fourth. One of its representants is CA-26, which achieves a low Polsby-Popper but high compactness score. The low Polsby-Popper comes from the fact that the borders are rather jagged, especially with CA-25, but the area around it is rather sparsely populated, as can be seen by the size of the districts. Furthermore, its bounding circle is not much bigger than the circle itself. As all of this area is Democrat, there is no significant evidence of gerrymandering, and this mostly follows county lines. Since it seems that there is always something to say about Ohio, we will now go back to Cleveland. 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.

Before concluding our work on the two measures, we will go to a somewhat miscellaneous example in California. CA-08 scores the worst fatness score across all ten states studied, and the Polsby-Popper score is mediocre as well. It is true that this is perhaps not the best district, however, its score is

harmed by what is going on behind the district lines. While it does have a rather jagged western boundary and a rather long rectilinear shape, the fatness score is deserving of the deeper analysis. It is no coincidence that the worst fatness measure is in the most populous state, and specifically, one with both very densely and sparsely populated areas. The diameter of the bounding circle of this large district is almost the northeast border of the district - this means that besides northern California and the Bay Area, virtually all of California is within that bounding circle. The fatness score is close to 0.025, which would indicate that about a fortieth of the bounding circle population is within the district: this means there are about forty districts inside of this circle (note that since we do not consider populations outside of the state, this would be impossible in any other state)!

## 6 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 smoothed, a less straightforward and undebatable process than it may appear.

The population 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 state. 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. Additionally, this measure tends to favor districts on the edge of the state, since on at least one side, there is no population to consider. 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), CA-26 (not too densely populated but in the middle of less densely populated areas), OH-03 (with a good proportion of Columbus, with few of the suburbs concerned), or PA-02 (in the middle of Philadelphia). 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.