

Background Research

What is gerrymandering?

- The manipulation of congressional districts and precincts to regroup voters and give more power to a certain political party

What are current ways to identify gerrymandering?

The Efficiency Gap:

- Uses number of votes that don't count towards electing representative
- Unreliable, can give false positives and negatives

Supercomputing:

- Generates millions of un-gerrymandered maps and sees if current districts are included
- Effective, but costly

Congressional districts correspond to representatives in Congress. They must have roughly equal population.



Precincts are smaller divisions of districts representing people who vote at one place.



HACKING THE ELECTION

Measuring and Solving Gerrymandering in Today's Political System

Arin Khare, James Lian, Kai Vernooy

Data Analysis

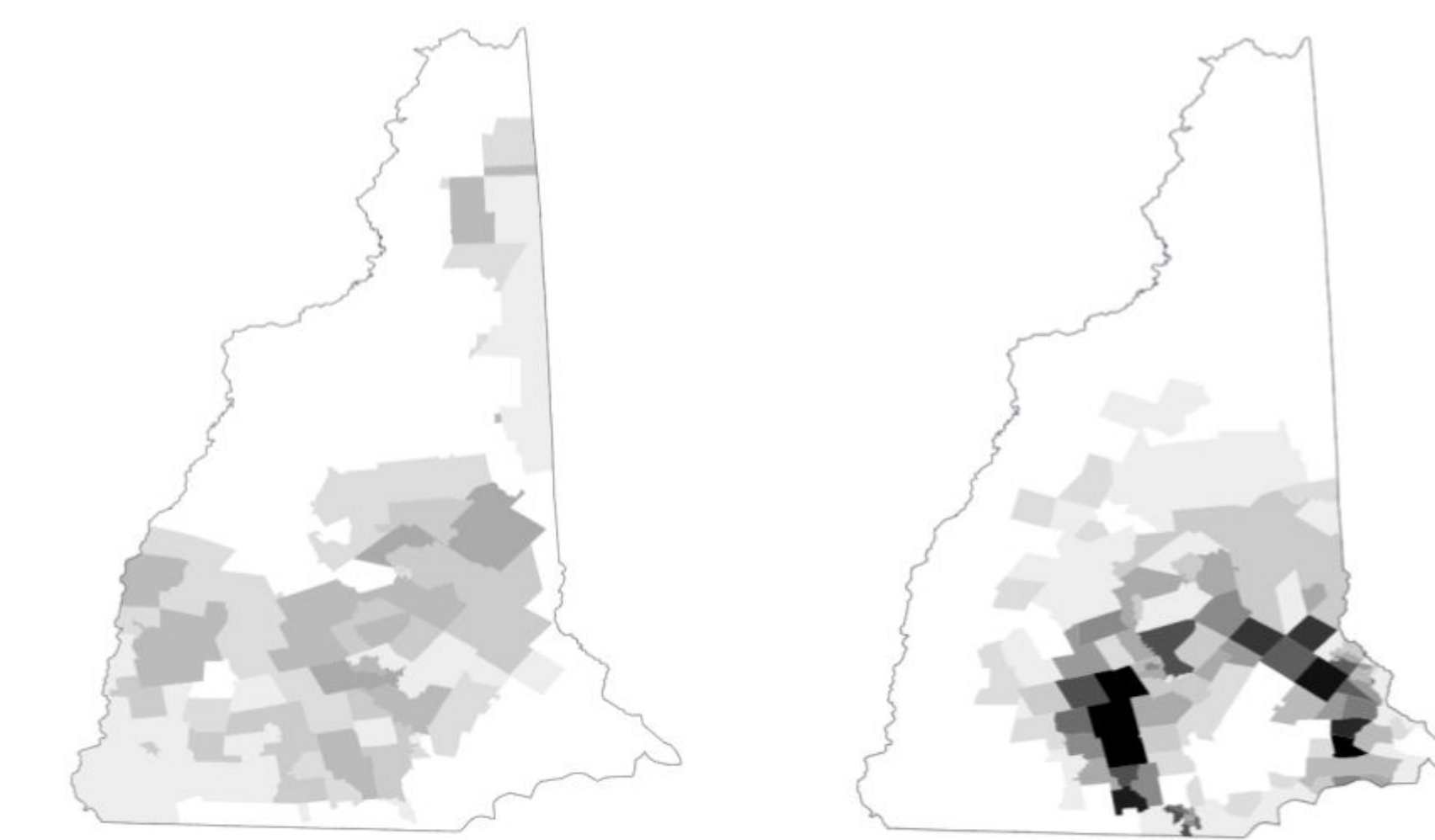


Fig 4 (Above) These two heatmaps show the number of times a precinct changed which community it belonged to, for base communities (left) and redistricting (right). While precinct exchanges for base communities were more spread out, redistricting was concentrated along a border, possibly reflecting more stringent constraints or a luckier initial configuration.

Hypotheses & Goals

Goals

- Create an algorithm that, when given election and geographical data, can determine how gerrymandered a state's congressional districts are.
- Create an algorithm that, when given election and geographical data, can fairly redistrict any state.

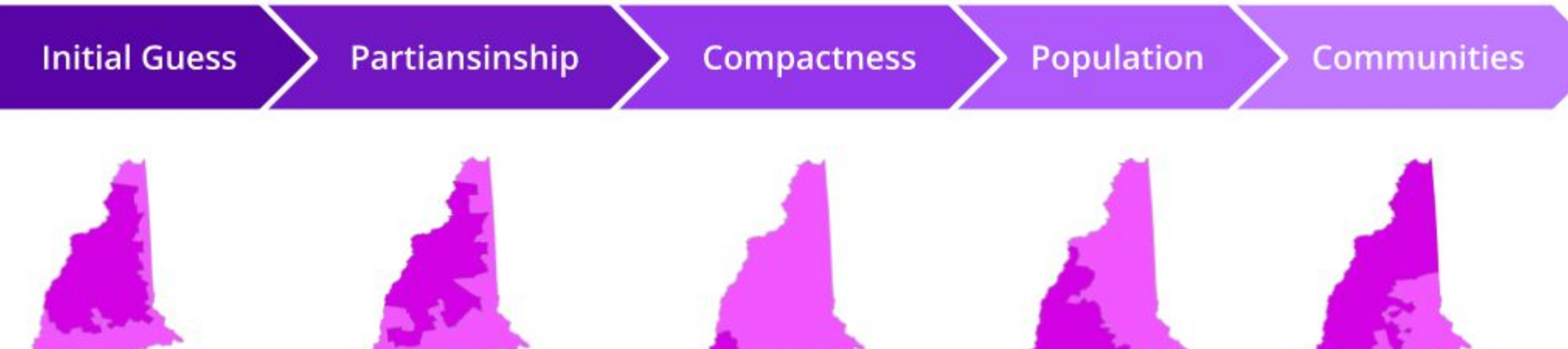
Hypotheses

- States such as North Carolina and Alabama with a history of gerrymandering will be found to still have noticeable traces of gerrymandering.
- Our redistricted map will vary greatly from the current districts in famously gerrymandered states, but states that are less gerrymandered will have more similar districts to ours.

Identifying Political Communities

- To detect gerrymandering, we must first identify political communities (in order to find how they are split up and combined)
- We define political communities to have minimal partisanship standard deviation, meaning the constituency thinks (and votes) alike

Our algorithm uses the **iterative method** to find the optimal solution to our constraints of population, compactness, and political diversity. This means it starts with a random guess and optimizes for each constraint until a configuration is found that works for them all.



Political Communities

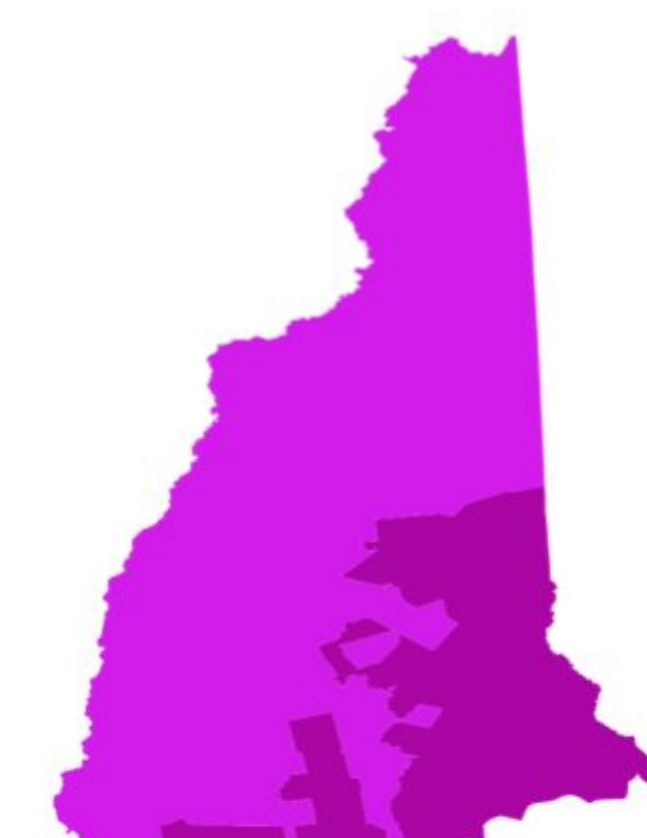
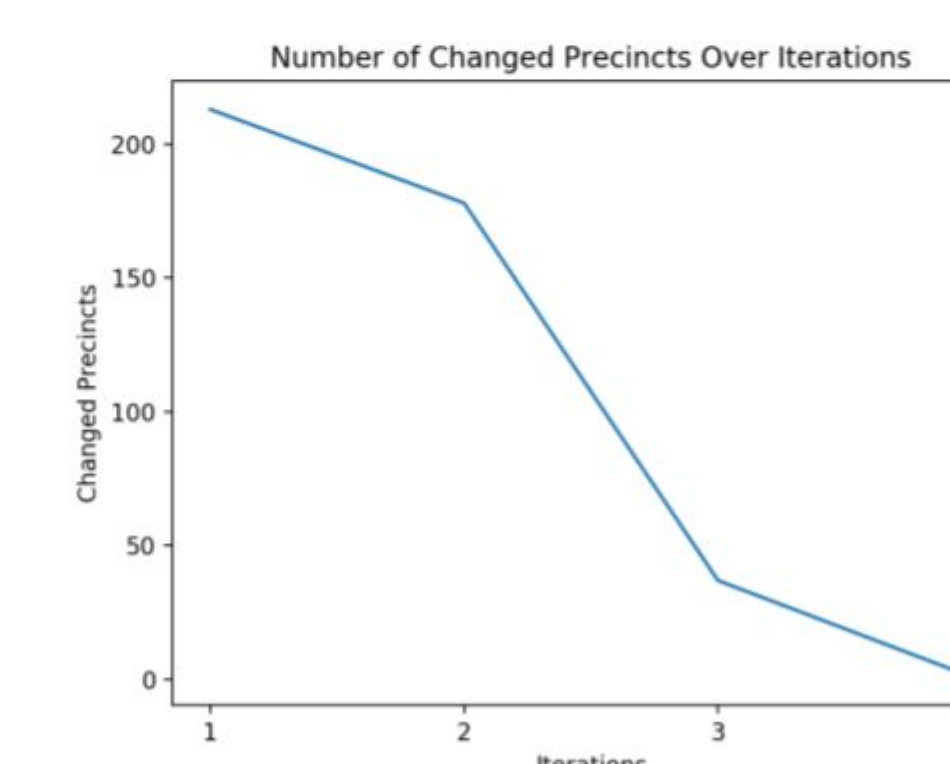
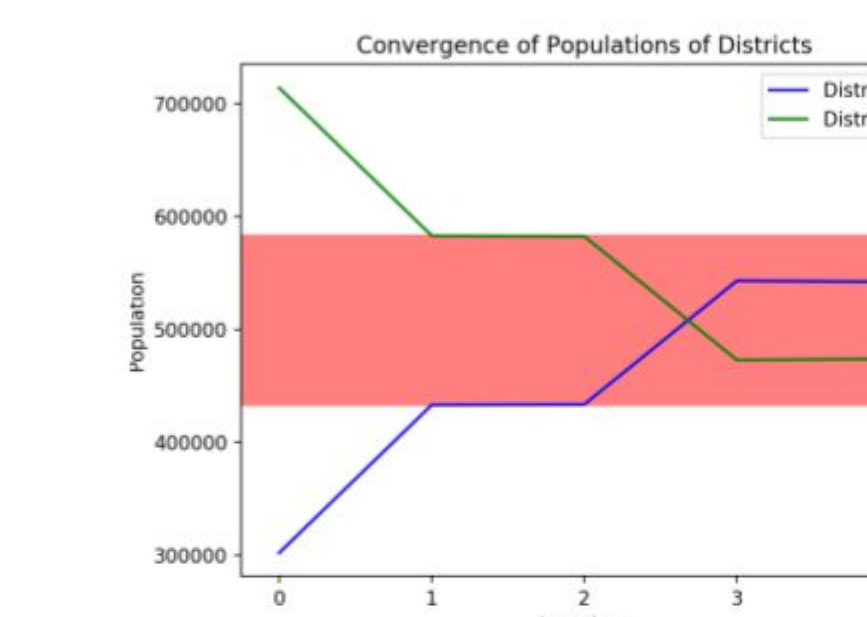
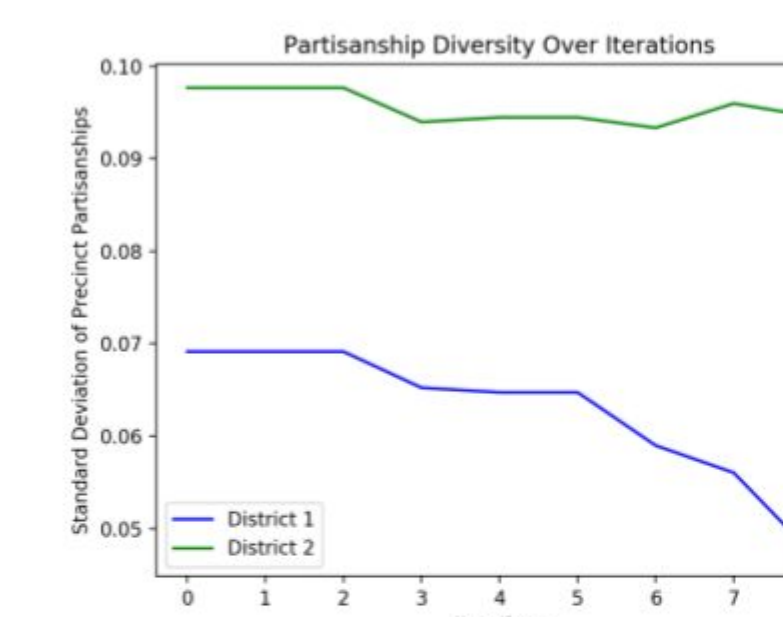


Fig 1.0 (Above) political communities we generated for New Hampshire.

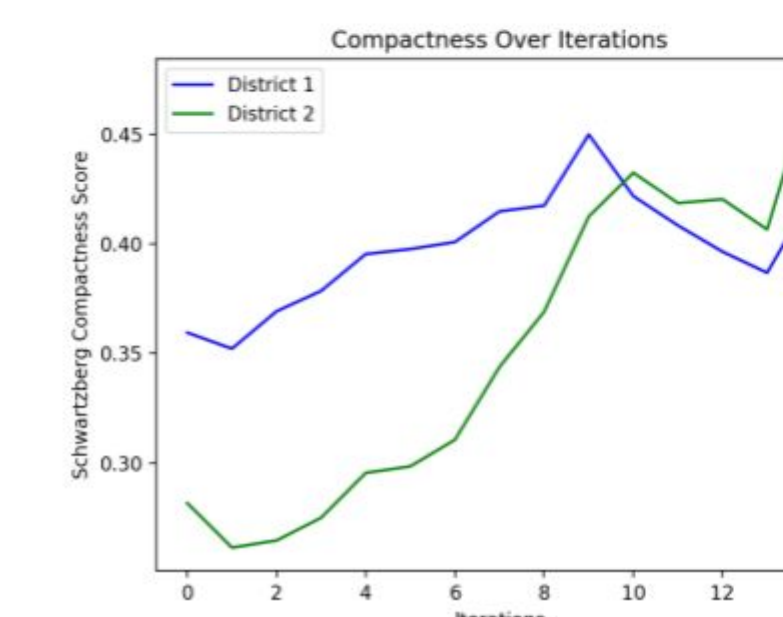
Fig 1.1 (Below) Measuring the number of precincts that moved communities each iteration is a way of measuring how the iterative method converged. Fewer precincts moving between refinement processes implies that the solution satisfies all of the constraints faster and is closer to the global optimum.



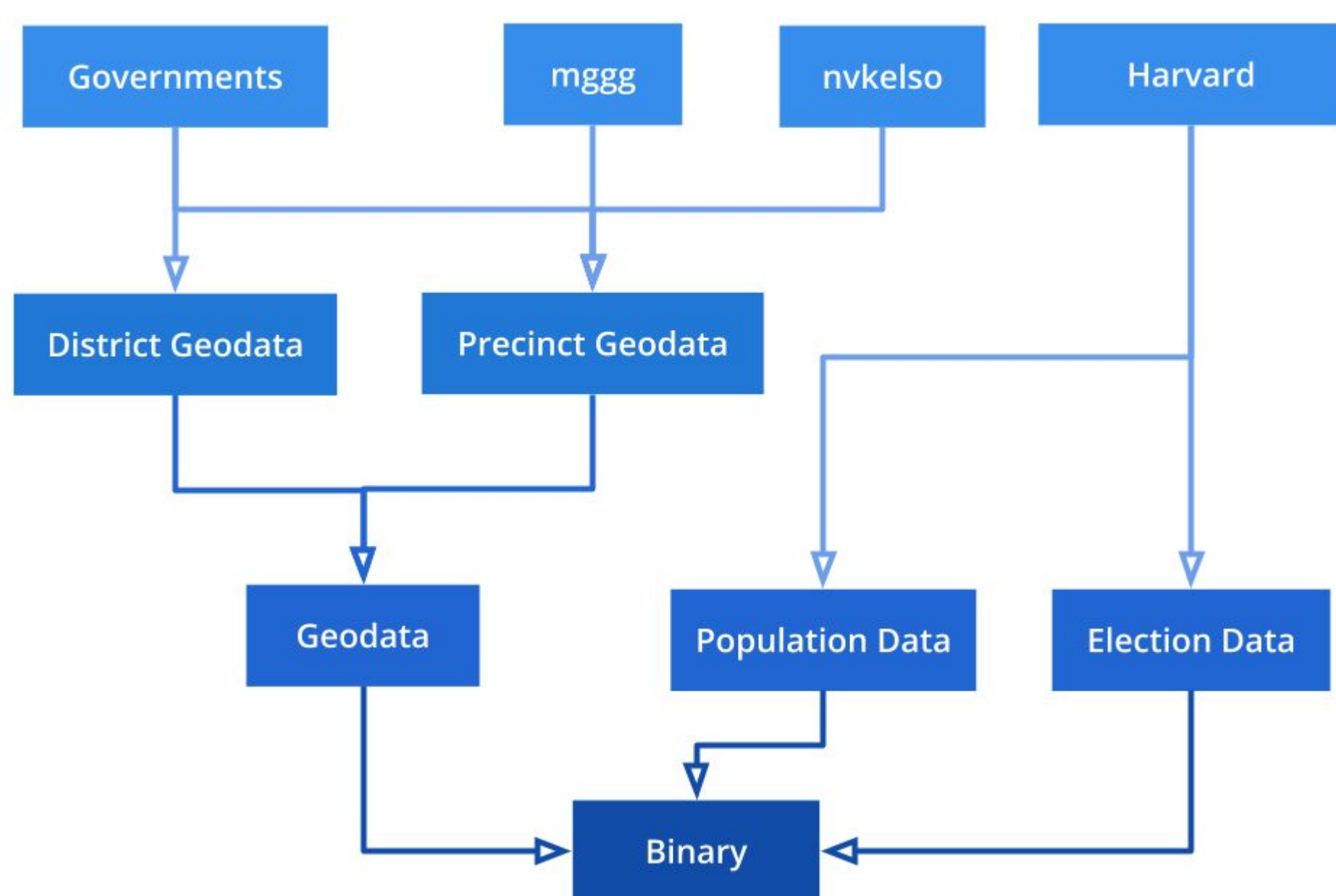
Data Analysis



An analysis of the refinement processes. As can be seen, as precincts are modified, the scores for partisanship, compactness and population either decrease, increase, or converge (respectively) to a range as expected, thus showing how our algorithm improves the various constraints simultaneously over iterations.



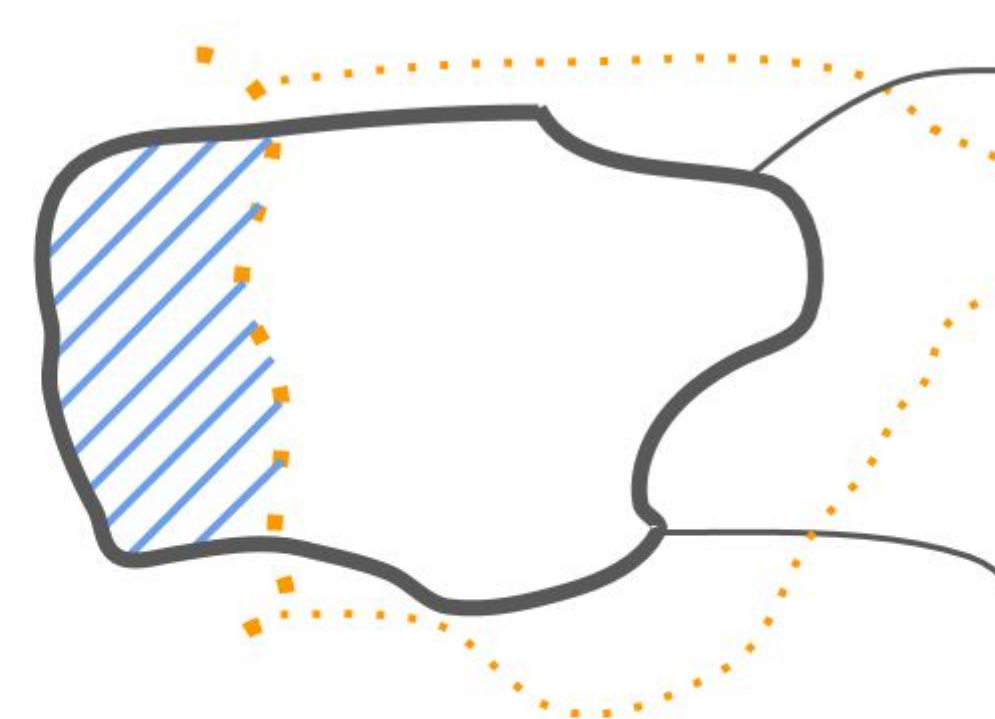
Data Processing



Quantifying Gerrymandering

Once we have identified political communities, we can see how current districts split up and divide them. We do this with the following process:

- Find the community that contains more of the district than any other community.
- Determine the percent area of the district that is outside of this community.



Using the actual districts of the state, we can calculate a gerrymandering score for the districts we generate, and we can do the same for our generated districts, thereby quantifying the decrease in gerrymandering and comparing our two maps fairly with each other.

Fig 2 (Above) An example that demonstrates how our quantification algorithm arrives at a score. Bold black is the district in question, and the orange dotted lines represent the borders of base communities. Shaded in blue is the part of the area of the district not in the community of largest area.

Redistricting Method

- Since we define gerrymandering breaking up and combining these political communities, our redistricted maps must try to follow these community lines
- We used the algorithm we created to generate communities to generate non-gerrymandered political districts. However, districts are limited by strict population and compactness criteria, so we gave less importance to political standard deviation

Communities

Partisanship: 9.5% stdev
Compactness: 0.35 ratio
Population: 15% +/- ideal

Districts

Partisanship: 9.8% stdev
Compactness: 0.45 ratio
Population: 0.5% +/- ideal

Conclusion

- New Hampshire has significant gerrymandering.
- Our redistricting algorithm has found the global optimum having found the same solution to our constraints starting from different initial guesses.
- Our new districts are significantly less gerrymandered than the current districts of New Hampshire.
- Our algorithm runs with limited computational resources and avoids the reliability problem found in the efficiency gap.

Open Source

This was one of the first major efforts in gathering precinct level election data for the entire United States. We contributed to the open-source political movement by compiling this data into an easily accessible GitHub repository, found here: [hacking-the-election/data](https://github.com/hacking-the-election/data)



Having written implementations for our algorithms in both C++ and Python, countless developers and researchers who would like to help the fight against gerrymandering will have access to our code.

Redistricting Results

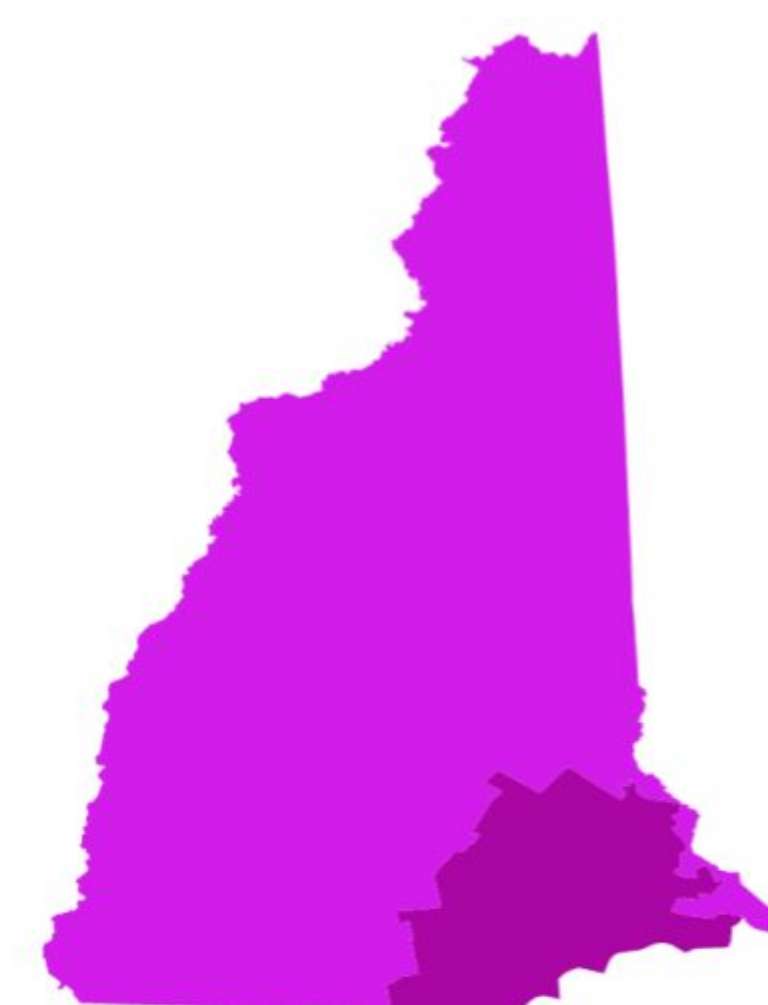


Fig 3.1 (Above) A map of our generated New Hampshire districts.

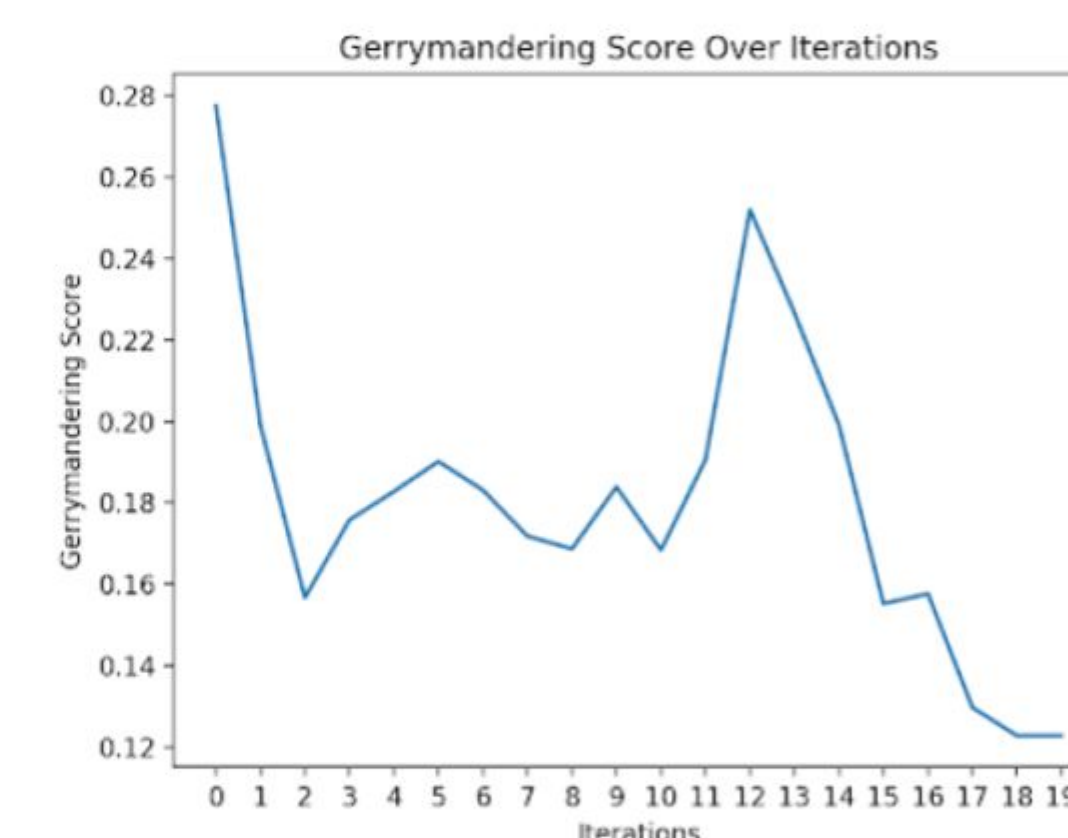
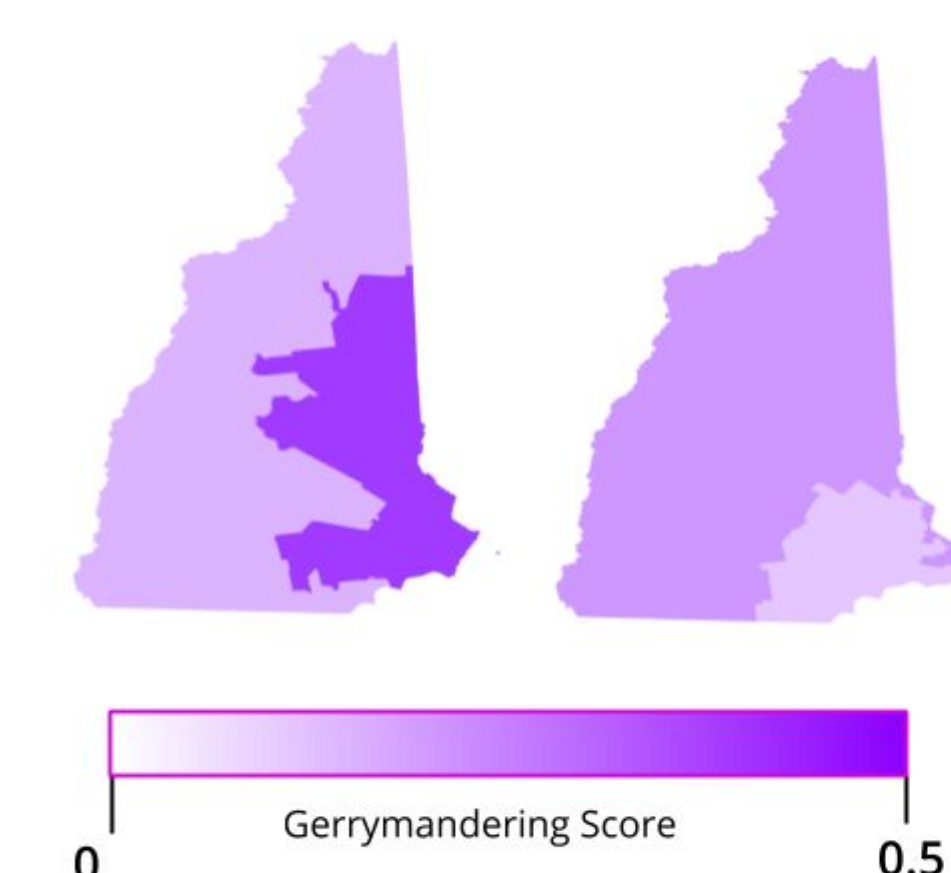
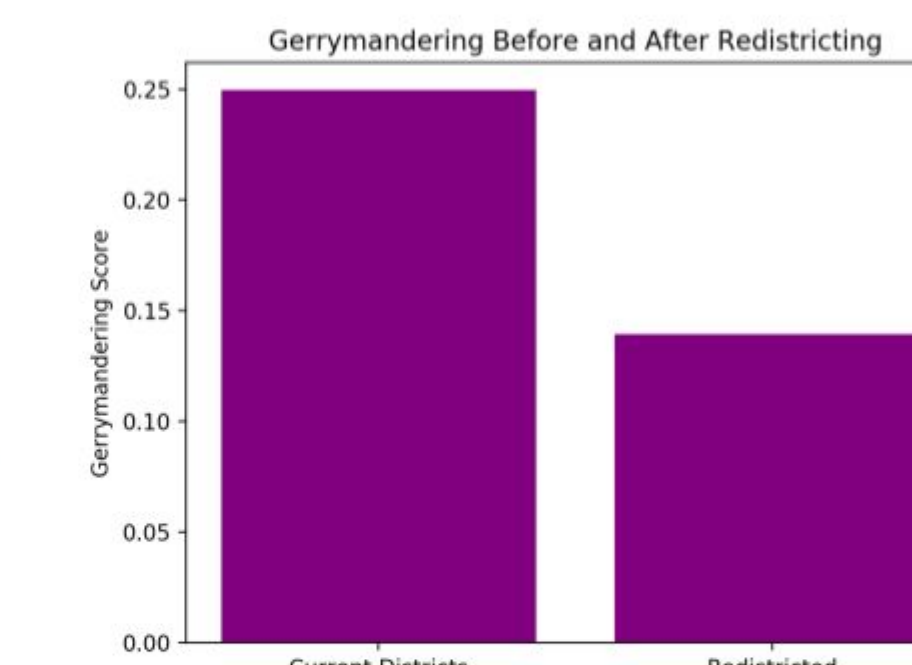


Fig 3.2 (Above) Over 19 iterations, our algorithm reduces the gerrymandering score of New Hampshire. The final score of our districts was nearly half that of the current congressional districts

Redistricting Quantification



The districts drawn by our algorithm (left) got significantly lower gerrymandering scores than the current Congressional Districts of New Hampshire (far left), reducing scores by around 50% (right).



Future Steps



States we have data for in red, states we ran in grey, and states we don't have data for in black.

In the future, we would like to collect data for all 50 states and run our algorithms on all of them. This will allow us to determine the extent of political gerrymandering throughout the country, and attempt to fix the problem.

- Gerrymandering on the basis of race is illegal, and we plan to conduct an analysis to see if it is still present in any states.
- We hope to continue working on our code, so that it is faster, more robust, and more friendly for contributors and users.
- Create an intuitive and functional website for redistricting and documentation

More can be found on our website, hacking-the-election.github.io