Assignment 3: Integer Programming Example
Algorithmic Redistricting Fall 2024

Emily Ekdahl

MSDS 460: Decision Analytics

Dr. Thomas W. Miller

2024-11-03

Introduction

Redistricting is a hard data problem that has real implications for politics and community life. In the redistricting assignment, I am applying integer programming to the problem of fairly representing individual voters in the U.S. government. In the 2020 presidential election, Ohio voted 54.08% republican, 45.92% democratic according to the website https://districtr.org/plan. From both lived experience growing up in Ohio as well as the data, I know that those voters are not equally distributed among the counties.

Method

For this assignment, I gathered the population of each county in Ohio. Then, I separately got the information about the neighboring counties for each county. I cleaned and processed that data and joined it together.

The goal of the optimization problem is to assign each county in Ohio to one district and one district only and to have exactly 16 districts, corresponding to the number of representatives in Ohio. We are optimizing for roughly equal populations within districts, as well as compactness or the idea that we want districts to be close to one another in terms of physical proximity.

At first, I tried using only the adjacency data but that ultimately yielded a nonsensical result. Then, I switched over to calculating a distance score using the centroids and used a cost function that minimized that distance. This approach seems to make more sense but takes a prohibitively long time to run. I could see in the logs that the algorithm was attempting branch and bound and other methods to try and get to a solution faster, but ultimately it didn't resolve in a timely fashion.

Results

In the end, I wasn't able to get to a sensible district allocation with either method. You can see my map optimizing for only the population saved to the repo. The counties aren't grouped up but there were 16 districts.

Further research

If I had more time, I'd probably look into a different approach that runs faster than integer programming or some infra that would run the iterations faster. I also should have limited the amount of counties that I performed the centroid calculation on; I suspect that is why my project is running so long.

References

Districtr. "Draw Your Community's Districts." Accessed November 3, 2024. https://districtr.org/plan.

Ohio Secretary of State. "Ohio County Roster." Accessed November 3, 2024. https://ohioroster.ohiosos.gov/county_list.aspx.

U.S. Census Bureau. "County Adjacency File." Accessed November 3, 2024. https://www2.census.gov/geo/docs/reference/county_adjacency.txt.