ASSIGNMENT 3: INTEGER PROGRAMMING–ALGORITHMIC REDISTRICTING FOR NEW JERSEY

Michael Mistarz, Alex McCorriston, Halton Pi
MSDS 460: Decision Analytics
Northwestern University
Dr. Thomas Miller
May 5, 2024

1.0 Data Sources

For our model, we relied on two primary sources of data: county population information and county adjacency information. County population data for our model was sourced from the World Population Review, utilizing data from 2024 (World Population Review n.d.). Additionally, county adjacency data was obtained from the United States Census Bureau (U.S. Census Bureau 2023). We had no concerns about the reliability of these sources, as they are well-established and reputable.

2.0 Model Specification

To solve the redistricting problem, we developed and tested two distinct models, as outlined below:

2.1 Primary Model

The standard form for our primary integer programming model is based on a redistricting model developed for Oregon (Havit 2021). The standard form for the model is as follows:

Index Sets:

- i: County index (21 counties)
- j: District index (12 districts)
- k: Additional county index for adjacency matrix

Data:

- P_i: Population in county i
- L: :Lower population limit for each district
- U: Upper population limit for each district
- A_{ik}: Adjacency matrix indicating whether county i borders county k (binary)
- M: large arbitrary number

Decision variables:

- Y_{ij}: Binary variable representing whether county i is assigned to district j
- X_{ij}: Integer variable representing the population allocated from county i to district j

Objective function:

- Minimize $\Sigma_i \Sigma_j Y_{ij}$ (minimize the total number of county-to-district assignments)

Subject to:

- Constraint 1: $\Sigma_j X_{ij} = \mathrm{P}_i \ \forall \ i$ (allocate 100% of county populations to districts)
- Constraint 2: $X_{ij} \leq M \cdot Y_{ij} \ \forall \ i$ (assignment required for population allocation)
- Constraint 3: $\Sigma_j Y_{ij} \leq 1 \ \forall \ i$ (each county must be assigned to one district)
- Constraint 4: $L \leq \Sigma_i x_{ij} \leq U \ \forall \ j$ (each district must be within population limits)
- Constraint 5: $Y_{ij} \leq \ \Sigma A_{ik} \cdot Y_{kj} \ \forall \ i, j, k$ (adjacency constraint)
- $X_{ij}\, Y_{ij} \geq 0 \ \forall \ i, j$ (non-negativity constraint)

Our model employs an objective function focused on minimizing the total number of county-to-district assignments, aiming to create the fewest number of county assignments to each district. To accommodate set covering, constraint 3 ensures that each county is assigned one district. To foster geographically compact districts, the model uses adjacency information to form constraint 5, encouraging neighboring counties to be grouped together. Although not directly incorporating district compactness measures or distance metrics, these constraints indirectly promote compactness by favoring contiguous county assignments. The principles of one-person-one-vote are implemented through constraints 1 and 4, ensuring that each county's population is fully allocated to districts and that districts are approximately equally populated within the specified lower and upper bounds.

2.2. Alternate Model

The alternate model integrates concepts from two political redistricting articles. The first article, by Garfinkel & Nemhauser (1970), introduces an objective function focused on minimizing the maximum deviation from the average district population. This average is calculated by dividing the total population across all counties by the number of districts. Additionally, there is a constraint ensuring that district populations remain within a specified percentage of this average. The second article, authored by Becker & Soloman (2020), introduces an additional decision variable, denoted as x', to represent whether the boundary between two counties is cut. This model also incorporates constraints to ensure the adjacency of counties assigned to the same district. By combining the objectives and constraints from both articles, we formed our alternate model.

3.0 Programming

We implemented our models using Python's PuLP package. The program code and output files are available in the GitHub repository. We used the GLPK solver for our first attempt at the primary model. This code can be found in "main.py" but we were unable to find a solution so we also ran the primary model without the adjacency constraints, which resulted in a solution. This code and output can be found in "secondary.py" and "secondary_solution.txt" Additionally, code and output for the alternate model can be found in “alternate_attempt.py” and “alternate_attempt.txt”. The GitHub repository also contains image files depicting the mapped distribution of counties to districts based on the model solutions. The primary and secondary models were tested using GLPK, CBC and PuLP’s built-in solvers and the alternate model was only tested using PuLP's built-in solver.

4.0 Solution

On our initial attempt to solve the primary model using the GLPK solver, the program ran for over 38 hours without finding a solution. Unfortunately, despite the extensive runtime, the solver did not converge. An illustration of this lengthy simulation is provided in Figure 1 of the appendix. Subsequently, we turned to the CBC solver, hoping for better results. However, the output from this attempt yielded infeasible solutions. Despite iteratively adjusting various constraints, such as considering the possibility of allowing counties to be split into multiple districts and widening the lower and upper population bounds, the solutions remained infeasible. Figure 2 in the appendix displays one of the instances of these unsuccessful attempts. Ultimately, we decided to remove the adjacency constraint and rerun the model using PuLP's solver. This decision led to an optimal solution with an objective function value of 21, indicating successful assignment of each county to a single district. The district population lower and upper limits were defined as 600,000 and 1,000,000 respectively. The resulting district populations range from 611,442 to 961,932. However, without enforcing the adjacency constraints, our districts lack practical feasibility as they are not in close geographic proximity to each other. Consequently, this plan cannot be deemed ready for submission to the state, governor, or courts. Further refinement is necessary to find an optimal solution while imposing the adjacency constraints.

Similarly, running the alternate model using PuLP's solver resulted in an optimal solution. However, the adjacency constraints were not properly integrated into the model, so this resulting redistricting map is again impractical because the districts are not geographically compact. Our testing of both models highlights the necessity of refining the model to ensure the adjacency constraints are accurately incorporated.

5.0 Maps and Discussion

Color-coded maps illustrating the algorithmic redistricting for both the primary and alternate models can be found in Figures 3 and 4 of the appendix, respectively. Additionally, Figure 5 in the appendix displays the map depicting the actual redistricting implemented by New Jersey (County of Union 2023). A comparison of our algorithmic redistricting maps with the actual map reveals that the actual map uses geographic splitting of almost every county to create geographically separated districts. This allows for a more compact district layout. We did not do this because the assignment specified for us to assign every county to exactly one district. To create a geographically compact map like the state's current redistricting plan, we would need county population density data to help us split the counties into multiple districts. The state's current redistricting plan is the most optimal for this reason and is what we would recommend.

Regarding the fairness and equity of the redistricting maps and their enforcement of equal representation in voting, our primary model solution displays a wide range of district populations from 611,442 to 961,932, indicating a disparity in voting power among citizens across districts. Comparatively, the state's redistricting model as of 2020 showed district populations ranging from 680,382 to 741,466, presenting a more balanced distribution conducive to equal voting power among citizens (The Official Web Site for the State of New Jersey n.d.). None of the maps achieve absolute fairness and equity yet. Ideally, achieving completely balanced populations across all districts is necessary to uphold the principle of "one person, one vote."

Appendix

```
Time used: 136898.3 secs.  Memory used: 1134.2 Mb.
+15599760: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988272; 122549)
+15600202: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988312; 122552)
+15600680: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988345; 122555)
+15601205: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988374; 122561)
+15601612: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988413; 122565)
+15602219: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988453; 122569)
+15602596: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988493; 122573)
+15603138: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988525; 122579)
+15603594: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988558; 122584)
+15604105: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988598; 122588)
+15604653: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988640; 122592)
Time used: 136958.6 secs.  Memory used: 1134.7 Mb.
+15605106: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988683; 122595)
+15605665: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988712; 122601)
+15606121: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988754; 122604)
+15606676: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988793; 122608)
+15607222: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988835; 122612)
+15607651: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988872; 122616)
+15608050: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988909; 122620)
+15608448: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988947; 122624)
+15608965: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (988977; 122630)
+15609353: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989018; 122633)
+15609998: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989059; 122636)
Time used: 137018.6 secs.  Memory used: 1135.1 Mb.
+15610454: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989101; 122639)
+15610813: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989145; 122642)
+15611221: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989184; 122646)
+15611681: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989206; 122654)
+15612146: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989234; 122659)
+15612458: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989279; 122662)
+15612754: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989308; 122667)
+15613310: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989349; 122670)
+15613955: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989389; 122674)
+15614469: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989439; 122676)
+15614961: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989478; 122680)
+15615432: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989506; 122686)
Time used: 137078.7 secs.  Memory used: 1135.5 Mb.
+15615901: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989552; 122689)
+15616376: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989590; 122693)
+15616872: mip =   2.700000000e+01 >=   2.100000000e+01  22.2% (989631; 122697)
Traceback (most recent call last):
  File "C:\Users\micha\OneDrive - Northwestern University\MSDS460\Module 6\assignme
e 392, in <module>
```

Figure 1. Initial 38 hour run of primary model (main.py) using GLPK solver with no solution found.

```
The model status is:  Infeasible
The objective value is:  0.9999999976999998
County 0 assigned to district 0:  27924.0
County 0 assigned to district 6:  93815.0
County 0 assigned to district 7:  90303.889
County 0 assigned to district 11:  63001.111
County 1 assigned to district 4:  454501.83
County 1 assigned to district 8:  485461.5
County 1 assigned to district 9:  21968.667
County 2 assigned to district 1:  245069.22
County 2 assigned to district 7:  227163.78
County 3 assigned to district 6:  150745.0
County 3 assigned to district 7:  315996.89
County 3 assigned to district 11:  63001.111
County 4 assigned to district 6:  93815.0
County 5 assigned to district 7:  90303.889
County 5 assigned to district 11:  63001.111
County 6 assigned to district 3:  398008.17
County 6 assigned to district 4:  454501.83
County 7 assigned to district 6:  310079.0
County 8 assigned to district 3:  223101.5
County 8 assigned to district 6:  1.0050652e-10
County 8 assigned to district 8:  485461.5
County 9 assigned to district 1:  16207.778
County 9 assigned to district 2:  69101.0
County 9 assigned to district 7:  24344.222
County 9 assigned to district 9:  20908.0
County 10 assigned to district 1:  16207.778
County 10 assigned to district 9:  366355.22
County 11 assigned to district 0:  233353.0
County 11 assigned to district 2:  29840.0
County 11 assigned to district 3:  174906.67
County 11 assigned to district 4:  1.6960475e-10
County 11 assigned to district 5:  261277.0
County 11 assigned to district 9:  1.2249232e-10
County 11 assigned to district 11:  166775.33
County 12 assigned to district 2:  167126.5
County 12 assigned to district 6:  37384.5
County 12 assigned to district 9:  345447.22
County 12 assigned to district 10:  91411.778
County 13 assigned to district 7:  123577.28
County 13 assigned to district 10:  394049.72
County 14 assigned to district 0:  27924.0
County 14 assigned to district 1:  245069.22
County 14 assigned to district 2:  167126.5
County 14 assigned to district 6:  131199.5
County 14 assigned to district 10:  91411.778
County 15 assigned to district 9:  119106.28
County 15 assigned to district 10:  394049.72
County 16 assigned to district 6:  65519.0
County 17 assigned to district 2:  89360.0
County 17 assigned to district 5:  261277.0
County 18 assigned to district 4:  30959.667
County 18 assigned to district 7:  18591.722
County 18 assigned to district 9:  97137.611
County 19 assigned to district 0:  233353.0
County 19 assigned to district 3:  174906.67
County 19 assigned to district 11:  166775.33
County 20 assigned to district 4:  30959.667
County 20 assigned to district 7:  80641.333
```

Figure 2. Example infeasible solution for primary model after relaxing constraints and using CBC solver.

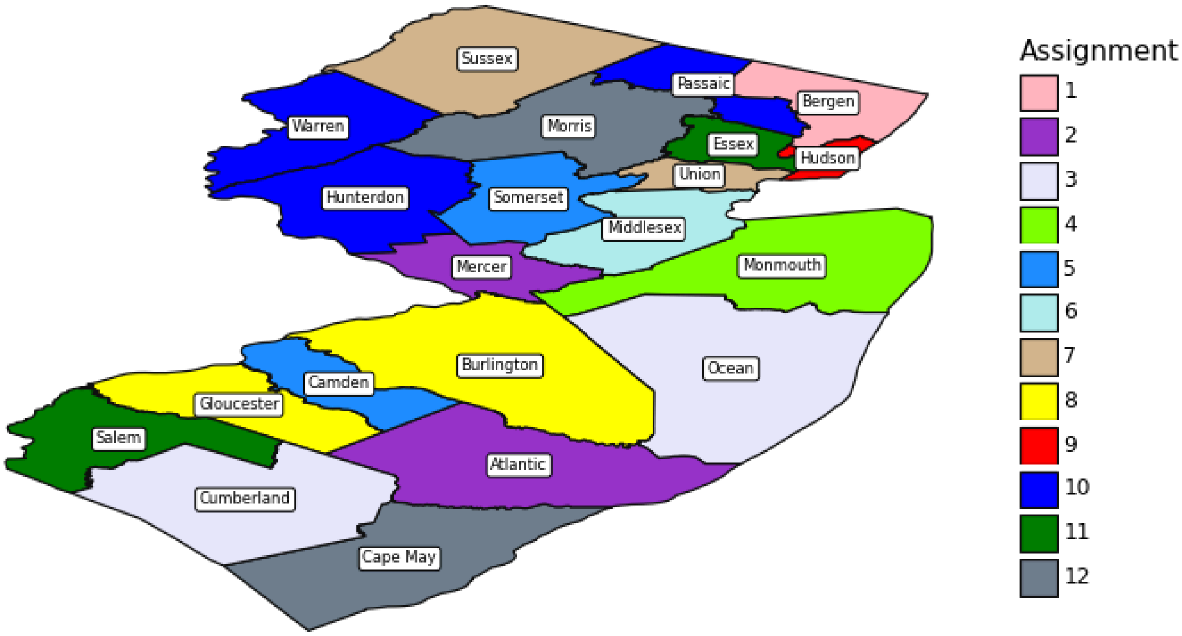

Figure 3. Color-coded map for primary model redistricting.

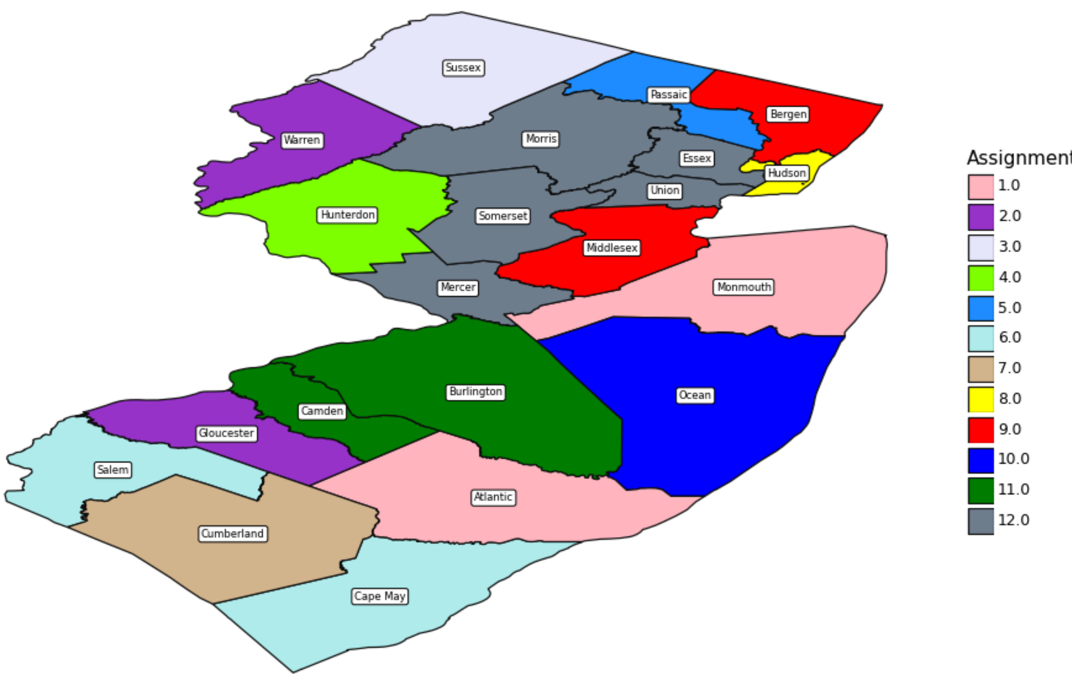

Figure 4. Color-coded map for alternate attempted model redistricting.

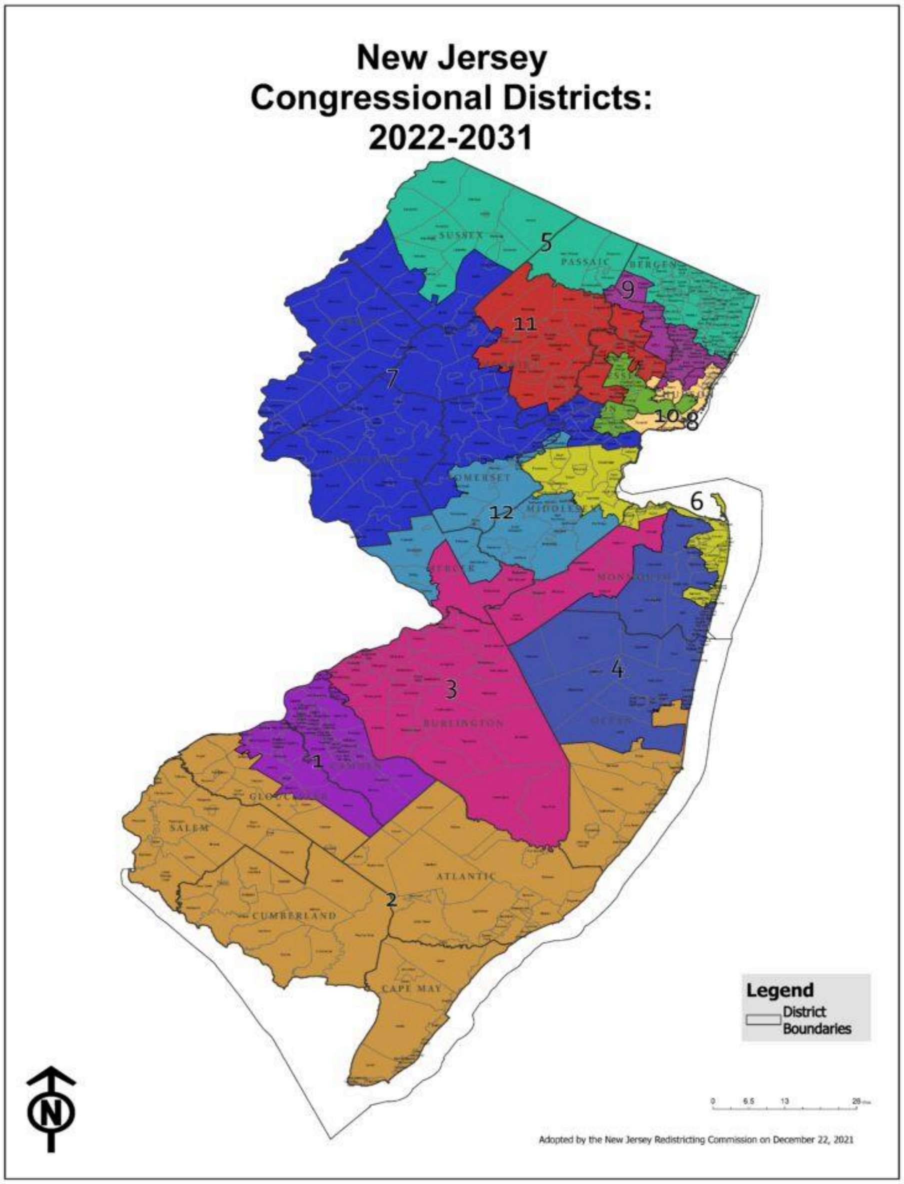

Figure 5. Color-coded map for actual redistricting implemented by New Jersey.

References

Becker, Amariah, and Justin Solomon. 2020. "Redistricting Algorithms." *Springer International Publishing, Cham*: 303–340. https://doi.org/10.1007/978-3-319-69161-9_16.

County of Union. 2023. "Changes to Union County Congressional, Municipal Ward & District Maps." https://ucnj.org/press-releases/public-info/2022/04/04/changes-to-union-county-congressional-municipal-ward-district-maps/.

Garfinkel, R. S., and G. L. Nemhauser. 1970. "Optimal Political Districting by Implicit Enumeration Techniques." *Management Science* 16, no. 8: B495–508. http://www.jstor.org/stable/2628656.

Horvat, Sabi. 2021. "How to Draw Congressional Districts in Python With Linear Programming." Medium. https://towardsdatascience.com/how-to-draw-congressional-districts-in-python-with-linear-programming-b1e33c80bc52.

The Official Web Site for the State of New Jersey. n.d. "P1. POPULATION BY RACE AND HISPANIC OR LATINO ORIGIN". "https://nj.gov/state/assets/2020-census-data/unadjusted-data/congressional-and-legislative-districts/congressional-and-legislative-districts.xlsx.

US Census Bureau. 2023. "County Adjacency File." https://www.census.gov/geographies/reference-files/2023/geo/county-adjacency.html.

World Population Review. n.d. "Population of Counties in New Jersey 2024." Accessed on May 5, 2024. https://worldpopulationreview.com/states/new-jersey/counties.