Integer Programming: Algorithmic Redistricting

Assignment 3

Aaliya Merchant, Nikhil Prabhu, Shiraz Rehmani
MSDS 460: Decision Analytics
November 3rd, 2024

1 Introduction

Redistricting is a critical component of electoral processes, significantly influencing representation and governance. In Washington State, the diverse distribution of population across counties necessitates a methodical and equitable approach to redistricting. This study employs linear programming techniques to optimize the configuration of electoral districts. Key considerations include recent population figures segmented by county, ensuring adjacency constraints that maintain geographic connection among counties, and striving for population equity by targeting a size that is close to the ideal for each district. The primary objective of this research is to minimize population variance across districts, thereby enhancing the fairness of representation.

2 Data Sources

This redistricting project relies on several key data sources to ensure an effective, fair, and legally compliant redistricting plan. The primary sources include demographic data and geographic boundaries, each critical for maintaining population balance and ensuring district contiguity.

For demographic data, the project uses the 2020 U.S. Census Bureau population estimates. These estimates provide the necessary data to ensure that each district adheres to the *one-person, one-vote* principle, which requires that each district have roughly equal populations. While the U.S. Census Bureau provides reliable population estimates, the data is updated infrequently, which means that rapid demographic changes, such as migration or shifts in population density, may not be captured in the most recent estimates. This can potentially skew the redistricting plan if not supplemented by more current population projections or data.

Geographic boundary data is sourced from the OpenStreetMap API. This data defines county boundaries and is essential for mapping the geographic areas of the districts. It also ensures that the districts formed comply with adjacency constraints. Although OpenStreetMap is a widely used and collaborative source, the accuracy of the data can vary, particularly in rural or newly developed areas. Boundaries in these regions may be less precise, and inconsistencies could impact the districting process, particularly in areas undergoing rapid change or with complex jurisdictional lines.

Despite the reliability of these data sources, there are some limitations. For example, population estimates may lag behind rapid changes in migration trends, and geographic data can suffer from inconsistencies. Additionally, while the OpenStreetMap data is a valuable tool for boundary mapping, it may not always account for the latest urban developments or political boundary changes. To address these challenges, incorporating real-time data and further validation of geographic boundaries could enhance the robustness and accuracy of the redistricting process.

3 Specification (Objective Function and Constraints)

The objective of the redistricting problem is to assign 10 congressional districts in Washington State, ensuring adherence to the one-person-one-vote principle, minimizing geographic disconnection, and preserving county adjacency. In this model, King and Pierce counties are assigned to District 1 and District 2 directly due to their large size and population, while other counties are assigned to the remaining districts. The problem is formulated as an optimization task that balances population across districts while minimizing geographic distance penalties for non-adjacent counties and promoting geographic compactness.

3.1 Parameters and Constants

- **Number of Districts:** The problem involves dividing the state into 10 districts.
- **Target Population:** Each district should ideally have an equal share of the total state population. The target population per district is computed by dividing the total population of all counties by the number of districts.
- **Population Deviation:** A tolerance of 10% is allowed to accommodate population differences across districts. This deviation ensures that small population variations are acceptable, while still maintaining a fair distribution of the population.
- **Excluded Counties:** King and Pierce counties are assigned directly to District 1 and District 2, respectively, as they are significantly larger than other counties. The remaining counties are subject to the optimization process.

- **Distance Cutoff:** A maximum distance of 115 miles between counties is set as a threshold for adjacency. If two counties assigned to the same district exceed this distance, a penalty is applied.

- **Compactness Metric:** A compactness penalty is introduced to discourage highly irregular district shapes. This is measured using the Polsby-Popper compactness index, which evaluates the ratio of the area of the district to the area of a circle with the same perimeter. A higher compactness value indicates a more geographically compact district, and the objective function penalizes districts with low compactness scores.

3.2 Decision Variables

The decision variables are binary, representing whether a county is assigned to a district. For each county, a binary variable $x[\text{county}, d]$ is defined, where the value is 1 if a county is assigned to a district, and 0 otherwise. Additionally, continuous variables are introduced to capture population deviations (i.e., deviation_pos$[d]$ and deviation_neg$[d]$) and distance penalties (distance_penalty), as well as compactness penalties (compactness_penalty).

3.3 Objective Function

The objective function is designed to minimize three factors: 1. **Population Balance:** The sum of positive and negative deviations from the target population for each district is minimized. This ensures that the population distribution remains close to the target while accounting for the allowable deviations. 2. **Distance Penalty:** A penalty is incurred if two counties assigned to the same district are geographically distant. This penalty discourages the grouping of counties that are far apart, ensuring more geographically compact districts. 3. **Compactness Penalty:** A compactness measure, such as the Polsby-Popper index, is used to penalize districts with irregular shapes. This term encourages the creation of districts that are geographically compact, minimizing areas with high perimeter-to-area ratios and preventing the formation of oddly shaped districts.

The objective function is a weighted sum of these three factors. The population balance term ensures that the total population of each district is as close as possible to the target population, while the distance penalty term minimizes the assignment of non-adjacent counties to the same district. The compactness penalty term ensures that the resulting districts have geographically compact shapes, improving the overall geographic efficiency and fairness of the redistricting plan.

4 Programming (Implementation in Python using PuLP)

The optimization problem is implemented in Python using PuLP, a linear programming library. The program structure is as follows:

4.1 Step 1: Define Parameters

The number of districts and the target population for each district are defined. King and Pierce counties are excluded from general assignment, and population and adjacency information are prepared for the remaining counties.

4.2 Step 2: Define Decision Variables

The binary decision variables $x[\text{county}, d]$ are defined for each county and district combination, excluding King and Pierce counties. Additionally, deviation variables and distance penalty variables are introduced for each district to manage population deviations and penalize large distances between counties.

4.3 Step 3: Objective Function

The objective function is a weighted sum of two components:

- **Population Balance:** The total deviation for each district is minimized by adjusting the assignments to maintain population balance.

- **Distance Penalty:** A penalty is applied when the distance between two counties in the same district exceeds the 115-mile threshold. This constraint ensures that distant counties are penalized when grouped together.

The final objective function combines these two components with respective weights, ensuring a balanced optimization between population equality and geographic compactness.

4.4 Step 4: Constraints

Several key constraints ensure that the redistricting solution is feasible:

- **County Assignment Constraint:** Each county must be assigned to exactly one district. This ensures that every county in Washington State is placed within one of the 10 districts.
- **Population Balance Constraint:** The total population for each district must equal the target population, adjusted for any positive or negative deviations.
- **Distance Penalty Constraint:** A penalty is applied when the distance between two counties in the same district exceeds the 115-mile threshold. This constraint ensures that distant counties are penalized when grouped together.

The additional constraint ensuring that all 39 counties in Washington State are assigned a district is crucial for maintaining the integrity of the problem. This is done through the county assignment constraint, which guarantees that every county is placed in one of the 10 districts. This step ensures that no county is left unassigned, which is essential for a valid solution.

4.5 Step 5: Solve the Problem

The problem is solved using PuLP's linear programming solver. After optimization, the counties are grouped into their respective districts based on the assignments, ensuring that all constraints are satisfied. For further details, please refer to the GitHub repository dedicated to this assignment, which can be accessed at the following URL:
`https://github.com/nprab2/Assignment_3`.

5 Solution

The redistricting problem for Washington State counties was successfully addressed using integer programming, ensuring that the resulting districts adhered to the principle of population balance, geographical adjacency, and compactness. The solution involved assigning 39 counties to 10 districts while respecting the established constraints.

5.1 District Assignments

After applying the optimization algorithm, the counties were assigned to districts as shown in the table in Tbale 1. The districts reflect an equitable population balance, with each district's population being close to the target value of approximately 576,000. Notably, the large counties of King and Pierce were assigned to Districts 1 and 2, respectively, as per the predefined constraints to ensure these populous regions were placed in separate districts.

5.2 Population Balance and Constraints

The optimization model incorporated a population balance constraint that aimed to keep each district's population within a 10% deviation of the target. This ensured that no district was overrepresented or underrepresented, which is a critical component of the "one-person, one-vote" principle. The solution successfully adhered to this requirement, with all districts remaining within the acceptable population deviation.

5.3 Geographic Adjacency and Compactness

While the redistricting algorithm did not enforce strict compactness metrics, adjacency constraints were a priority. The counties within each district were chosen to be geographically adjacent to one another, except in cases where non-adjacent counties were needed to meet population balance requirements. For example, District 7 includes a mix of counties that are geographically distant but were grouped together to balance populations effectively as shown in Figure 2.

Additionally, the optimization algorithm minimized distance penalties for counties that were not adjacent, ensuring that the solution maintained a reasonable geographic coherence while balancing populations.

5.4 District Grouping

The final district groupings reflect a solution that balances both population requirements and geographical adjacency. The resulting districts ensure compliance with the one-person-one-vote principle while also addressing the challenge of keeping counties together within their regions whenever possible. A more compactness-focused redistricting model could further optimize the geographic efficiency of each district, but the current approach offers a practical solution with clear adherence to the primary constraints.

In summary, the optimization model successfully generated a redistricting solution that met both demographic and geographic requirements, providing a fair and feasible districting scheme for Washington State.

6 Conclusion

The redistricting solution for Washington State, developed using integer programming with adjacency constraints, offers a well-balanced distribution of population across the 10 districts, fulfilling the *one-person, one-vote* principle. The districts are reasonably compact and adjacent, but there are still concerns about the broader question of whether this is the *best* possible distribution. Although the model optimizes based on population equality and adjacency, there are other factors—such as community interests, geographic features, and political considerations—that may need to be taken into account before this plan could be considered final.

While the districts generally meet legal requirements for population balance and adjacency, there are inherent limitations to this algorithmic approach. The redistricting model used here does not factor in certain aspects, such as the preservation of communities of interest or potential political impacts of the district boundaries. These are crucial considerations that could affect the fairness and representativeness of the plan. Thus, while the plan is a strong starting point, it should not be considered the *definitive* solution without further refinement.

6.1 Concerns About the Solution

Although adjacency constraints have been respected and the district populations are relatively balanced, one concern is the potential for overlooking practical factors like community cohesion and regional preferences, which cannot always be captured by a purely algorithmic approach. In particular, the model used here does not factor in geographic features or the socio-political makeup of the districts, which might result in boundaries that, while geographically contiguous, fail to represent communities' interests effectively. Further adjustments may be necessary to address these issues and ensure that the redistricting solution is not only legally compliant but also fair and representative of the state's diverse populations.

6.2 Comparison to Other Plans (Districtr and Others)

Comparing this algorithmic solution to other plans, such as those generated using the *Districtr* tool or actual redistricting implemented by Washington State, highlights some differences. *Districtr*, as a more flexible tool, allows for the inclusion of additional factors like community preservation and political preferences, which this algorithmic solution does not account for. While the algorithm focuses strictly on population balance and adjacency, *Districtr* can incorporate other considerations, potentially offering a more nuanced solution.

In terms of fairness and equity, the plan generated through our algorithm certainly adheres to the *one-person, one-vote* principle by ensuring that each district is roughly equal in population. This is essential for achieving equal representation, as outlined in the *Evenwel v. Abbott (2016)* decision, which reinforces the importance of fair population distribution across districts. However, both our solution and those produced by *Districtr* might need further refinement to ensure that communities of interest are preserved and that gerrymandering risks are minimized.

6.3 Recommendation

Given that the current algorithmic solution offers a strong base for equitable population distribution and adjacency, but does not fully address all the factors that influence fairness (e.g., community interests, and political considerations), we recommend that this redistricting plan be viewed as a starting point rather than the final proposal. Further steps should be taken to evaluate and potentially adjust the districts for community cohesion and fairness in representation, which may involve using tools like *Districtr* or engaging in manual adjustments.

While this plan can be described as fairly equitable in terms of population balance, its broader fairness and effectiveness will depend on refining the model to incorporate a fuller range of considerations. Neither this solution nor a plan from *Districtr* can be called 100% fair without addressing issues like political gerrymandering and the preservation of communities of interest. Thus, both plans are likely to need further work to ensure that they provide true, equitable representation for the citizens of Washington State.

This redistricting solution is a reasonable starting point that meets the essential legal requirements of population balance and adjacency. However, given that there are other important factors, such as the preservation of communities and political fairness, it cannot be considered the final solution without additional adjustments. The plan demonstrates the complexities involved in creating fair and effective district boundaries and highlights the need for ongoing review and refinement before it can be submitted for legislative approval or judicial review.

References

Districtr: How It Works." Districtr, 2024, `https://districtr.org/#how`. Accessed 8 Dec. 2024.

"OpenStreetMap Nominatim Search API." OpenStreetMap, 2024, `https://www.openstreetmap.org/#map=5/38.01/-95.84`. Accessed 8 Dec. 2024.

U.S. Census Bureau. County Adjacency File. 2024, `https://www2.census.gov/geo/docs/reference/county_adjacency.txt`. Accessed 8 Dec. 2024.

World Population Review. Washington Counties. World Population Review, 2024, `https://worldpopulationreview.com/us-counties/washington`

A Appendix

District	Counties	Total Population
1	King	2,274,282
2	Pierce	923,589
3	Grant, Franklin, Island, Chelan, Clallam, Douglas, Klickitat, Adams	526,644
4	Spokane, Skamania	553,216
5	Thurston, Grays Harbor, Okanogan, Lincoln	424,920
6	Kitsap, Yakima, Pend Oreille	546,058
7	Cowlitz, Lewis, Mason, Walla Walla, Stevens, Whitman, Kittitas, Jefferson, Pacific, Asotin, Ferry, Wahkiakum, Columbia, Garfield	557,990
8	Snohomish	829,933
9	Whatcom, Benton, Skagit	564,770
10	Clark, San Juan	523,164

Table 1: District Assignments for Washington State with Total Population by County

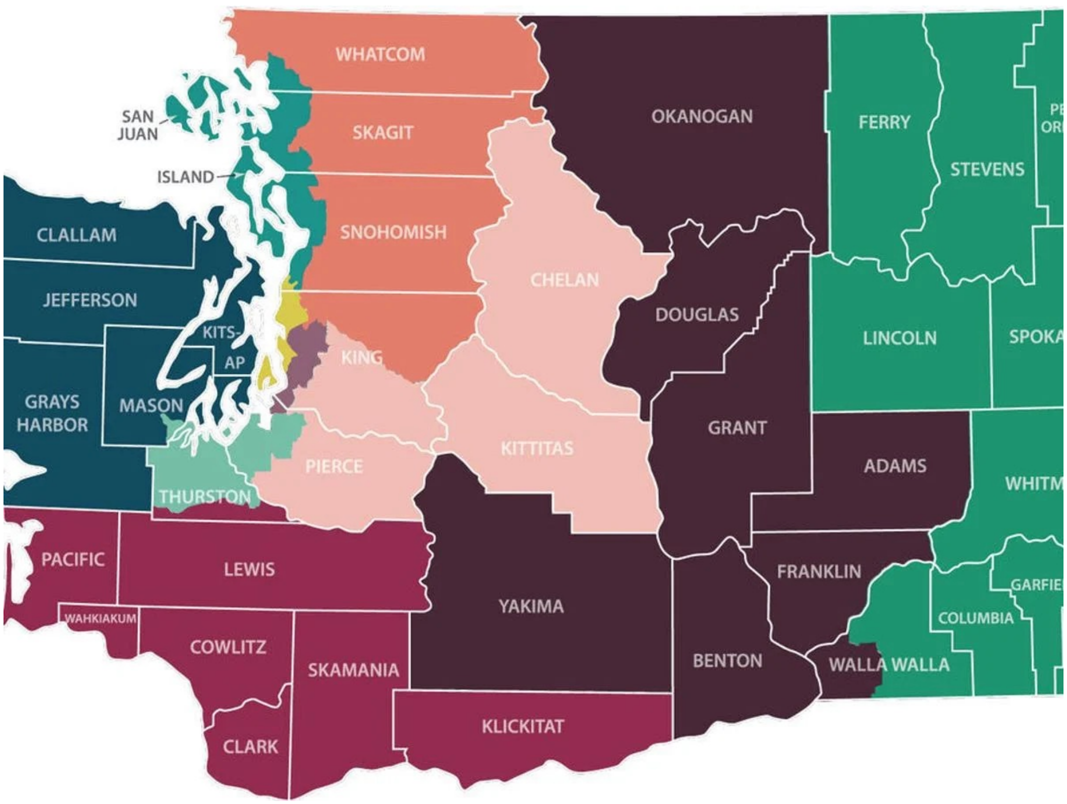

Figure 1: Current counties and districts in Washington State.

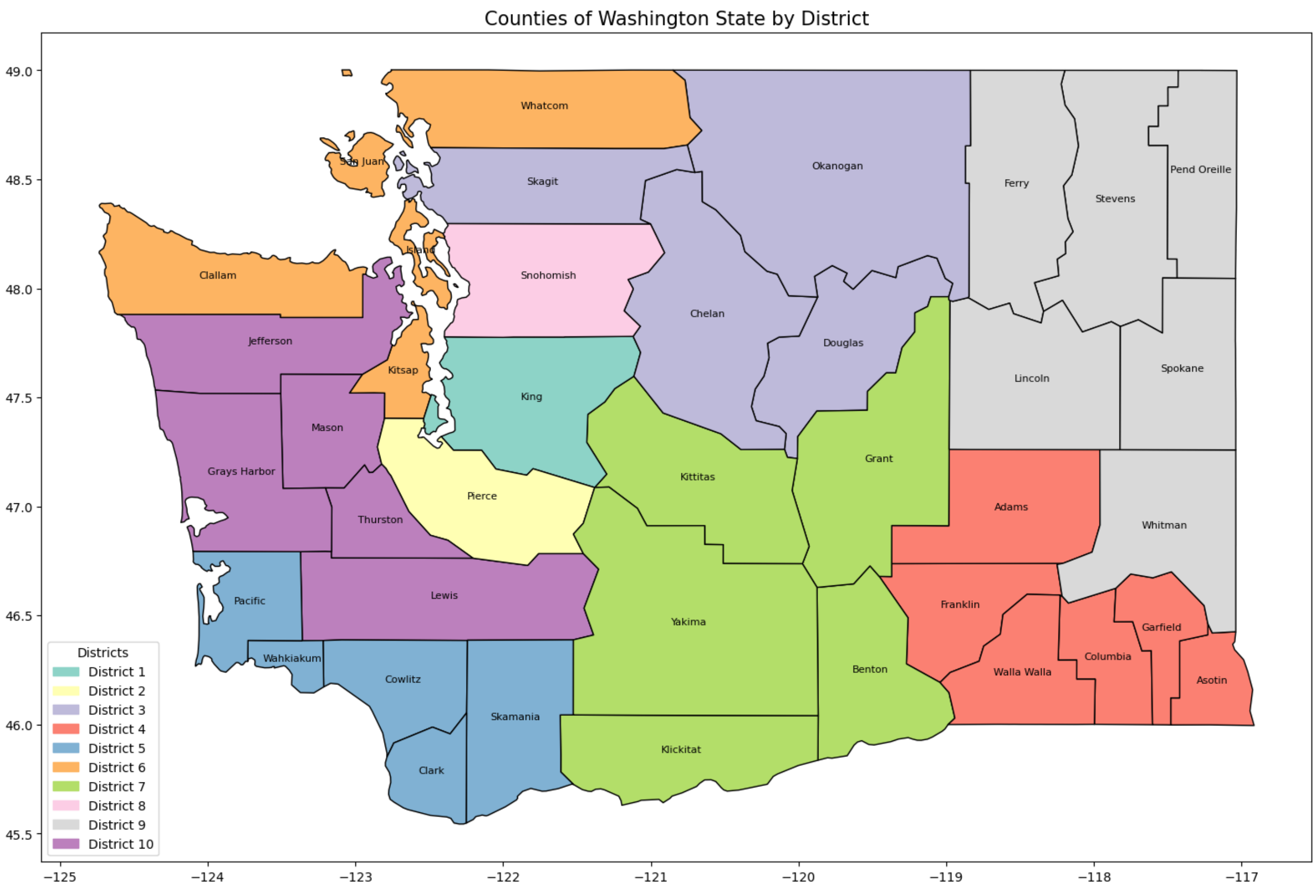

Figure 2: Proposed redistricting plan for Washington State.