# Political districting to minimize county splits

**Maral Shahmizad and Austin Buchanan**

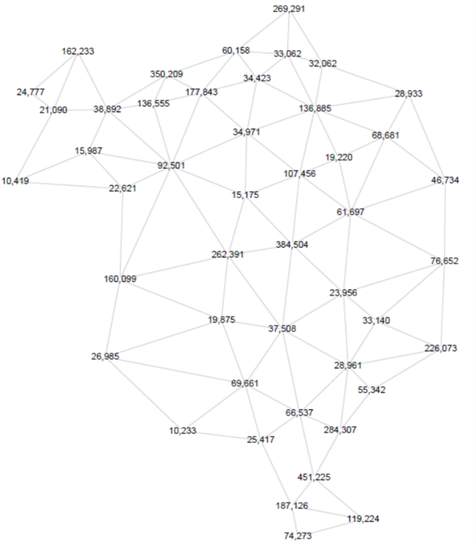
School of Industrial Engineering & Management
Oklahoma State University

March 2023

# My Coauthor

Maral Shahmizad

Can you find 7 contiguous districts for South Carolina such that each has population between 657,463 and 664,070? Provide a solution or an argument for infeasibility.



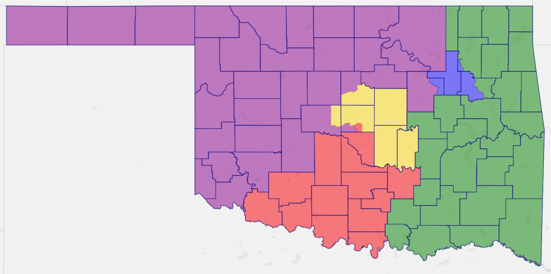
# What is Political Districting?

**Traditional criteria:**

- Population balance
- Voting Rights Act
- Contiguity
- Compactness
- Political subdivisions
- ...

**Emerging criteria:**

- Partisan fairness or proportionality
- Competitiveness
- ...



Oklahoma congressional districts (2022)

# Districting is Hard

# Optimization is a Work in Progress


Photo: Melanie Acevedo

George Box said:

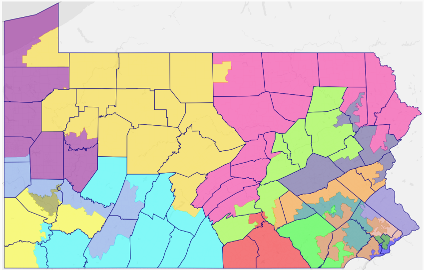
*All models are wrong, but some are useful.*

Optimizers should know:

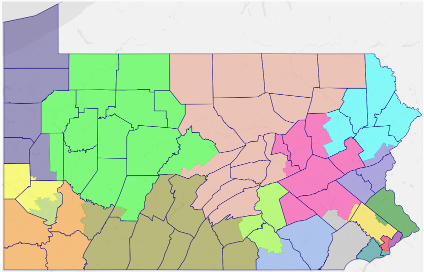
*All optimization models are wrong, but some are useful.*

# Republican Gerrymandering in Pennsylvania

*League of Women Voters of Pennsylvania v. the Commonwealth of Pennsylvania*



Overturned districts (2013-2018)
28 split counties, 37 county splits



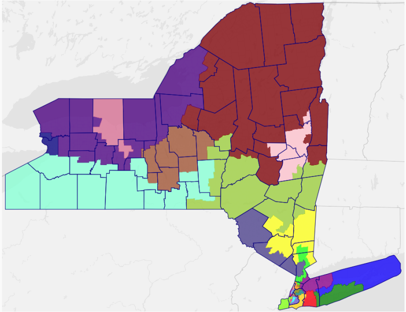
Court-mandated districts (2018)
13 split counties, 17 county splits

## Quote from Pennsylvania Supreme Court:

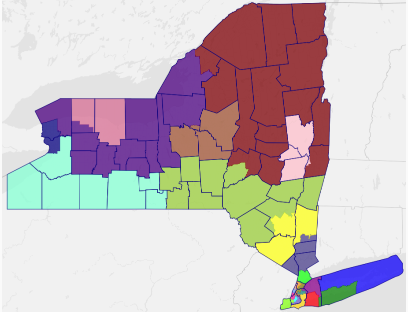
*[The new plan shall] not divide any county, city, incorporated town, borough, township, or ward, except where necessary to ensure equality of population.*

# Democratic Gerrymandering in New York

*Harkenrider v. Hochul*



Overturned districts (2022)
34 split counties, 56 county splits

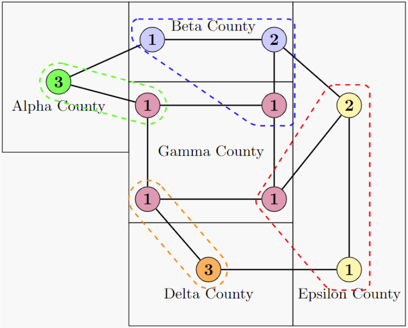


Court-mandated districts (2022)
16 split counties, 26 county splits

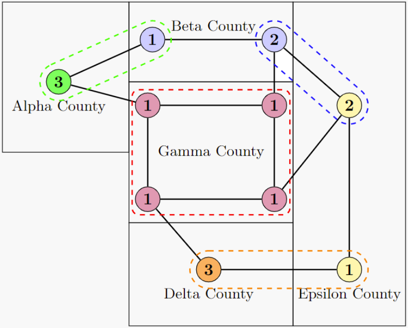
Quote from Special Master Jonathan Cervas:

*While I was quite successful in limiting the number of counties and cities that were split, some splits are simply inevitable. . . I can assure you that if yours was split it was not because of any kind of animus but was essentially due to the* ***mathematical necessity of splitting some units****.* (Bold added)

What is the **minimum number of county splits** possible (in a contiguous and population-balanced districting plan)?




1 split county, 3 county splits




2 split counties, 2 county splits

# Answers from the Literature

## John Nagle (2022) says:

The minimum number of county splits equals the number of districts minus one.

## Dave's Redistricting App (2023) says:

The minimum number of split counties is at most the number of districts minus one.

## McCartan and Imai (2020) say:

Our algorithm generates plans with (number of districts minus one) county splits.

## Autry, Carter, Herschlag, Hunter, and Mattingly (2021) say:

Our algorithm generates plans with (number of districts minus one) split counties.

## Carter, Hunter, Teague, Herschlag, and Mattingly (2020) say:

The minimum number of county splits equals the number of districts minus the maximum number of county clusters.

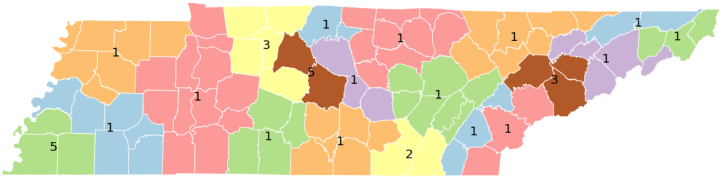
# County Clustering

Suppose we want $k$ districts, each with a population between $L$ and $U$.

**Definition (Carter et al., 2020)**

Let $G_C$ be the county-level graph. A county clustering is a partition $(C_1, C_2, \ldots, C_q)$ of the counties along with associated cluster sizes $(k_1, k_2, \ldots, k_q)$ such that:

1. the cluster sizes $(k_1, k_2, \ldots, k_q)$ are positive integers that sum to $k$;
2. each cluster $C_j$ is contiguous, i.e., induces a connected subgraph of $G_C$;
3. each cluster $C_j$ satisfies population balance, i.e., $Lk_j \leq p(C_j) \leq Uk_j$.

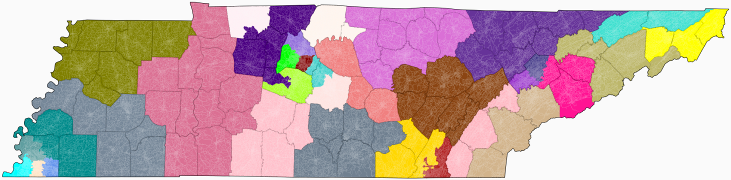



A **maximum** county clustering for the Tennessee State Senate
($k = 33$ districts, $q = 20$ county clusters)

# County Clustering

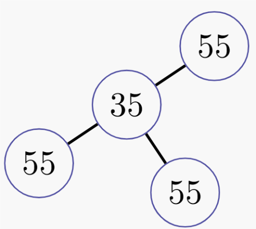
## Theorem (Carter et al., 2020)

*The minimum number of county splits equals the number of districts minus the maximum number of county clusters, "except in rare circumstances which affect the optimal districting".*



A districting plan for the Tennessee State Senate
($k = 33$ districts, $q = 20$ county clusters, $k - q = 13$ county splits)

*Enter the claw...*





Divide this county-level graph into $k = 2$ districts
with populations between $L = 95$ and $U = 105$.

The maximum number of county clusters is $q = 1$.
Carter's theorem suggests $k - q = 2 - 1 = 1$ split.

*But, we actually need 2 split counties and 2 county splits!*

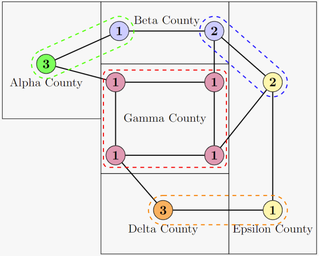
## Definition (Split Duality)

A districting instance exhibits **weak split duality** if the minimum number of county splits is at least the number of districts minus the maximum number of county clusters. It exhibits **strong split duality** if equality also holds.

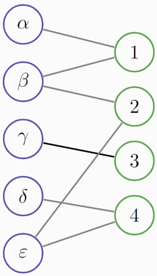
## Proposition

*Weak split duality always holds. Strong split duality does not always hold.*

Proof. Take a districting plan with a minimum number $s$ of splits. Construct the county-district incidence graph. It has $n = |C| + k$ vertices and $m = |C| + s$ edges, so its number of connected components is at least $n - m = (|C| + k) - (|C| + s) = k - s$. Construct a cluster from each.

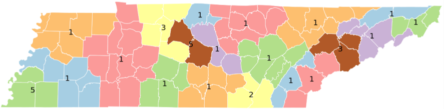



$\rightarrow$

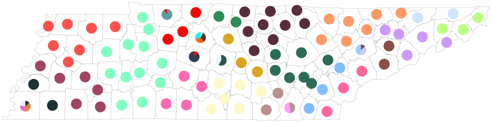


Our overall approach for min-split districting has three steps:

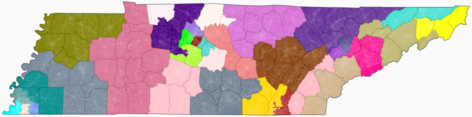
1. **Cluster**. Find a maximum number of county clusters $(C_1, C_2, \ldots, C_q)$ with associated cluster sizes $(k_1, k_2, \ldots, k_q)$.



2. **Sketch**. For each cluster $C_j$, sketch a plan using $k_j - 1$ county splits.



3. **Detail**. For each cluster $C_j$, find a plan that abides by the sketch's *support*.



If **Cluster** is successful, then $k - q$ is a lower bound. If **Sketch** and **Detail** are successful, then $k - q$ is an upper bound. If all, then **optimal!**

# MIP for Cluster

- $x_{ij} = \begin{cases} 1 & \text{if county } i \text{ is assigned to the cluster rooted at county } j \\ 0 & \text{otherwise} \end{cases}$
- $y_j$ = size of the cluster rooted at county $j$

$$
\begin{aligned}
\max \quad & \sum_{j \in C} x_{jj} & \\
\text{s.t.} \quad & \sum_{j \in C} x_{ij} = 1 & \forall i \in C \\
& \sum_{j \in C} y_j = k & \\
& C_j = \{i \in C \mid x_{ij} = 1\} \text{ is connected} & \forall j \in C \\
& L y_j \leq \sum_{i \in C} p_i x_{ij} \leq U y_j & \forall j \in C \\
& x_{ij} \leq x_{jj} & \forall i, j \in C \\
& x_{ij} \in \{0, 1\} & \forall i, j \in C \\
& y_j \in \mathbb{Z}_+ & \forall j \in C.
\end{aligned}
$$

# MIP for Cluster

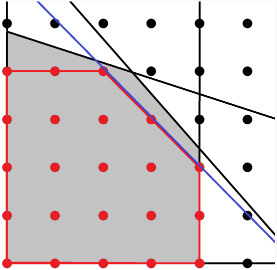
Key implementation details for max cluster:

- MIP is slow out-of-the-box; needs computational tricks!
- Symmetry handling (asymmetric representatives)
- MIP-based construction heuristic: carving à la McCartan and Imai (2020)
- MIP-based local search: *t*-opt recombination à la DeFord et al. (2021)
- Valid inequalities to strengthen the linear programming relaxation (!)

**Theorem (Rounding Inequalities)**

*Let t be a positive integer, and let j be a county. The following inequality is valid.*

$$\sum_{i \in C} \left\lfloor \frac{tp_i}{U+1} \right\rfloor x_{ij} \leq ty_j - x_{jj}.$$



# MIP for Sketch

- $x_{ij} = \begin{cases} 1 & \text{if some of county } i \text{ is assigned to district number } j \in [k] \\ 0 & \text{otherwise} \end{cases}$
- $z_{ij}$ = the proportion of county $i$ that is assigned to district number $j \in [k]$
- $s_i$ = the number of times that county $i$ is split

Constraints for splitting and population balance:

$$\sum_{i \in C} s_i = k - 1$$

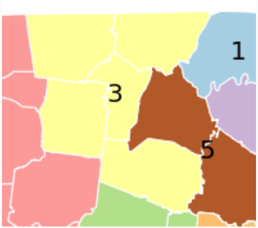
$$\sum_{j=1}^{k} x_{ij} = s_i + 1 \qquad \forall i \in C$$

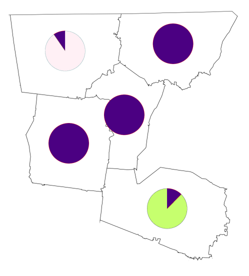
$$\sum_{j=1}^{k} z_{ij} = 1 \qquad \forall i \in C$$

$$L \le \sum_{i \in C} p_i z_{ij} \le U \qquad \forall j \in [k]$$

$$0 \le z_{ij} \le x_{ij} \qquad \forall i \in C,\ \forall j \in [k]$$

$$x_{ij} \in \{0, 1\} \qquad \forall i \in C,\ \forall j \in [k].$$



# MIP for Sketch

- $y_{ej} = \begin{cases} 1 & \text{if county edge } e \text{ is preserved in district number } j \in [k] \\ 0 & \text{otherwise} \end{cases}$

Edge consistency constraints:

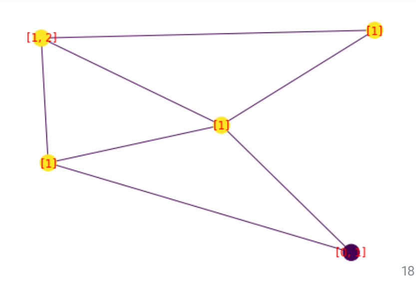
$$\begin{aligned} y_{ej} &\leq x_{ij} && \forall i \in e \in E(C),\ \forall j \in [k] \\ y_{ej} &\geq \sum_{i \in e} x_{ij} - 1 && \forall e \in E(C),\ \forall j \in [k] \\ \sum_{j=1}^{k} y_{ej} &\leq 1 && \forall e \in E(C) \\ y_{ej} &\in \{0, 1\} && \forall e \in E(C),\ \forall j \in [k]. \end{aligned}$$

Edge district has (# edges) $\geq$ (# nodes) - 1:

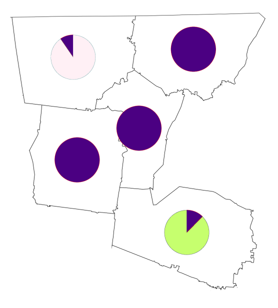
$$\sum_{e \in E(C)} y_{ej} \geq \sum_{i \in C} x_{ij} - 1 \qquad \forall j \in [k].$$

Objective: maximize preserved edges:

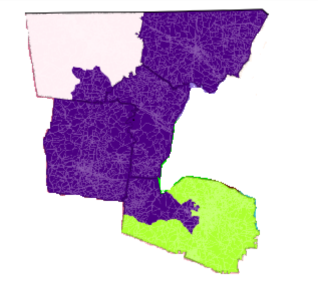
$$\max \sum_{e \in E(C)} \sum_{j=1}^{k} y_{ej}.$$



- Task: convert cluster's sketch into detailed plan



$\rightarrow$



- Challenge: large tract- and block-level instances $G = (V, E)$
- Approach: capacitated *k*-means algorithm, similar to Hess et al. (1965), Bradley et al. (2000), Validi et al. (2022)

# MIP for Detail

$$x_{ij} = \begin{cases} 1 & \text{if tract/block } i \text{ is assigned to district number } j \in [k] \\ 0 & \text{otherwise} \end{cases}$$

Solve this MIP, with better and better district means $(m_1, m_2, \ldots, m_k)$, with additional contiguity and sketch support constraints.

$$\min \sum_{i \in V} \sum_{j=1}^{k} (1 + p_i) \operatorname{dist}(i, m_j)^2 x_{ij}$$

$$\sum_{j=1}^{k} x_{ij} = 1 \qquad \forall i \in V$$

$$L \leq \sum_{i \in V} p_i x_{ij} \leq U \qquad \forall j \in [k]$$

$$x_{ij} \in \{0, 1\} \qquad \forall i \in V,\ \forall j \in [k].$$

Setup:

- Maral's Desktop PC has Intel Xeon Processor E52630 v4 (10 cores, 2.2GHz, 3.1GHz Turbo) and 32 GB RAM. MIP solver is Gurobi v10.0.
- P.L. 94-171 data from Census; initial processing by Redistricting Data Hub; then by Daryl DeFord; # enacted splits calculated from 2022 plans
- $\pm 0.5\%$ deviation for congressional instances; $\pm 5\%$ deviation for legislative

Questions:

- Does strong split duality hold in practice?
- How does the minimum number of splits compare to enacted plans?
- How strong is the "obvious lower bound"?

**Proposition**

*County $c$ must be divided across $\lceil p_c/U \rceil$ or more districts, so at least this many splits:*

$$\text{(obvious lower bound)} \quad \sum_{c \in C} \left( \lceil p_c/U \rceil - 1 \right).$$

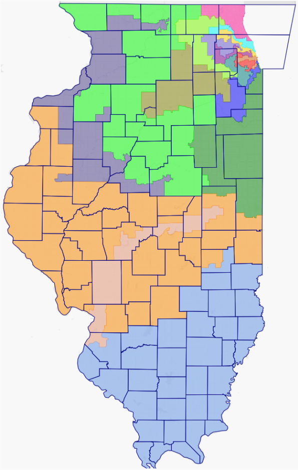
# Congressional Results (1/2)

| state | $\lvert C \rvert$ | $k$ | $L$ | $U$ | obvious LB | max clusters | split LB | min splits | enacted splits |
|---|---|---|---|---|---|---|---|---|---|
| AL | 67 | 7 | 714,166 | 721,342 | 0 | 7 | 0 | 0 | 6 |
| AR | 75 | 4 | 749,117 | 756,645 | 0 | 4 | 0 | 0 | 3 |
| AZ | 15 | 9 | 790,639 | 798,584 | 6 | 2 | 7 | 7 | 15 |
| CA | 58 | 52 | 756,549 | 764,152 | 39 | 11 | 41 | 41 | 72 |
| CO | 64 | 8 | 718,106 | 725,322 | 1 | 6 | 2 | 2 | 20 |
| CT | 8 | 5 | 717,583 | 724,794 | 3 | 1 | 4 | 4 | 10 |
| FL | 67 | 28 | 765,375 | 773,067 | 10 | 9 | 19 | 19 | 31 |
| GA | 159 | 14 | 761,311 | 768,961 | 2 | 12 | 2 | 2 | 21 |
| IA | 99 | 4 | 793,605 | 801,580 | 0 | 4 | 0 | 0 | 0 |
| ID | 44 | 2 | 914,956 | 924,150 | 0 | 2 | 0 | 0 | 1 |
| IL | 102 | 17 | 749,909 | 757,445 | 7 | 8 | 9 | 9 | 53 |
| IN | 92 | 9 | 750,178 | 757,717 | 1 | 8 | 1 | 1 | 8 |
| KS | 105 | 4 | 730,798 | 738,142 | 0 | 4 | 0 | 0 | 4 |
| KY | 120 | 6 | 747,218 | 754,727 | 1 | 5 | 1 | 1 | 6 |
| LA | 64 | 6 | 772,412 | 780,174 | 0 | 6 | 0 | 0 | 15 |
| MA | 14 | 9 | 777,197 | 785,007 | 5 | 2 | 7 | 7 | 22 |
| MD | 24 | 8 | 768,293 | 776,013 | 3 | 4 | 4 | 4 | 9 |
| ME | 16 | 2 | 677,774 | 684,585 | 0 | 2 | 0 | 0 | 1 |
| MI | 83 | 13 | 771,304 | 779,055 | 4 | 9 | 4 | 4 | 21 |
| MN | 87 | 8 | 709,746 | 716,878 | 1 | 6 | 2 | 2 | 12 |
| MO | 115 | 8 | 765,518 | 773,210 | 1 | 7 | 1 | 1 | 10 |
| MS | 82 | 4 | 736,619 | 744,021 | 0 | 4 | 0 | 0 | 4 |

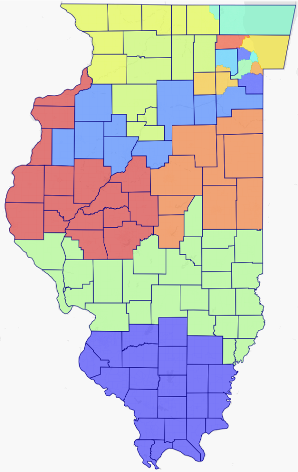
# Congressional Results (2/2)

| state | $\lvert C \rvert$ | $k$ | $L$ | $U$ | obvious LB | max clusters | split LB | min splits | enacted splits |
|---|---|---|---|---|---|---|---|---|---|
| MT | 56 | 2 | 539,402 | 544,823 | 0 | 2 | 0 | 0 | 1 |
| NC | 100 | 14 | 741,943 | 749,398 | 2 | 11 | 3 | 3 | 13 |
| NE | 93 | 3 | 650,566 | 657,103 | 0 | 3 | 0 | 0 | 2 |
| NH | 10 | 2 | 685,321 | 692,208 | 0 | 1 | 1 | 1 | 5 |
| NJ | 21 | 12 | 770,213 | 777,953 | 3 | 3 | 9 | 9 | 20 |
| NM | 33 | 3 | 702,312 | 709,369 | 0 | 3 | 0 | 0 | 10 |
| NV | 17 | 4 | 772,273 | 780,034 | 2 | 2 | 2 | 2 | 5 |
| NY | 62 | 26 | 773,087 | 780,855 | 13 | 8 | 18 | 18 | 26 |
| OH | 88 | 15 | 782,697 | 790,563 | 3 | 11 | 4 | 4 | 14 |
| OK | 77 | 5 | 787,912 | 795,829 | 1 | 4 | 1 | 1 | 7 |
| OR | 36 | 6 | 702,679 | 709,740 | 1 | 5 | 1 | 1 | 16 |
| PA | 67 | 17 | 761,041 | 768,689 | 4 | 10 | 7 | 7 | 17 |
| RI | 5 | 2 | 545,947 | 551,432 | 1 | 1 | 1 | 1 | 1 |
| SC | 46 | 7 | 727,548 | 734,859 | 0 | 6 | 1 | 1 | 10 |
| TN | 95 | 9 | 764,032 | 771,710 | 1 | 7 | 2 | 2 | 11 |
| TX | 254 | 38 | 763,153 | 770,821 | 19 | 19 | 19 | 19 | 59 |
| UT | 29 | 4 | 813,815 | 821,993 | 1 | 3 | 1 | 1 | 7 |
| VA | 133 | 11 | 780,749 | 788,595 | 1 | 9 | 2 | 2 | 11 |
| WA | 39 | 10 | 766,676 | 774,380 | 4 | 6 | 4 | 4 | 11 |
| WI | 72 | 8 | 733,032 | 740,398 | 1 | 7 | 1 | 1 | 13 |
| WV | 55 | 2 | 892,374 | 901,342 | 0 | 2 | 0 | 0 | 0 |

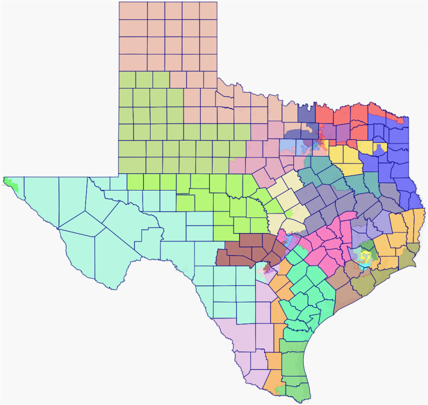
# Examples



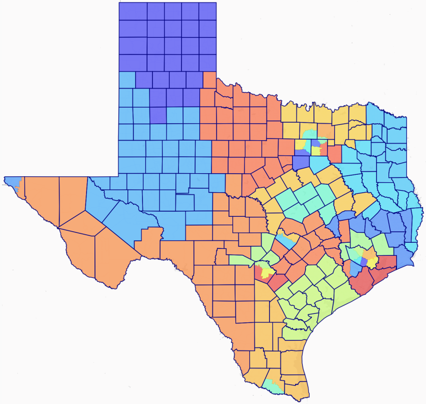
IL enacted (53 county splits)



IL min-split (9 county splits)



TX enacted (59 county splits)



TX min-split (19 county splits)

# State Senate Results (1/2)

| state | $\lvert C \rvert$ | $k$ | $L$ | $U$ | obvious LB | max clusters | split LB | min splits | enacted splits |
|---|---|---|---|---|---|---|---|---|---|
| AK | 30 | 20 | 34,837 | 38,503 | 12 | 7 | 13 | 13 | 19 |
| AL | 67 | 35 | 136,374 | 150,728 | 13 | 19 | 16 | 16 | 35 |
| AR | 75 | 35 | 81,742 | 90,345 | 14 | 21 | 14 | 14 | 51 |
| AZ | 15 | 30 | 226,465 | 250,302 | 22 | 6 | 24 | 24 | 44 |
| CA | 58 | 40 | 939,033 | 1,037,878 | 23 | 14 | 26 | 26 | 56 |
| CO | 64 | 35 | 156,716 | 173,211 | 22 | 13 | 22 | 22 | 42 |
| CT | 8 | 36 | 95,157 | 105,173 | 31 | 4 | 32 | 32 | 49 |
| DE | 3 | 21 | 44,784 | 49,497 | 18 | 3 | 18 | 18 | 20 |
| FL | 67 | 40 | 511,532 | 565,377 | 18 | 16 | 24 | 24 | 32 |
| GA | 159 | 56 | 181,720 | 200,848 | 23 | 31 | 25 | 25 | 60 |
| IA | 99 | 50 | 60,618 | 66,997 | 20 | 29 | 21 | 21 | 46 |
| ID | 44 | 35 | 49,919 | 55,173 | 19 | 14 | 21 | 21 | 25 |
| IL | 102 | 59 | 206,304 | 228,019 | 39 | 20 | 39 | 39 | 135 |
| IN | 92 | 50 | 128,926 | 142,496 | 20 | 28 | 22 | 22 | 48 |
| KS | 105 | 40 | 69,775 | 77,119 | 19 | 21 | 19 | 19 | 36 |
| KY | 120 | 38 | 112,646 | 124,503 | 11 | 26 | 12 | 12 | 21 |
| LA | 64 | 39 | 113,459 | 125,401 | 20 | 18 | 21 | 21 | 77 |
| MA | 14 | 40 | 166,961 | 184,535 | 31 | 6 | 34 | 34 | 59 |
| MD | 24 | 47 | 124,859 | 138,001 | 35 | 10 | 37 | 37 | 45 |
| ME | 16 | 35 | 36,979 | 40,870 | 24 | 8 | 27 | 27 | 40 |
| MI | 83 | 38 | 251,934 | 278,452 | 19 | 18 | 20 | 20 | 64 |
| MN | 87 | 67 | 80,913 | 89,430 | 38 | 26 | 41 | 41 | 100 |
| MO | 115 | 34 | 171,976 | 190,078 | 14 | 20 | 14 | 14 | 16 |
| MS | 82 | 52 | 54,101 | 59,795 | 19 | 29 | 23 | 23 | 64 |

# State Senate Results (2/2)

| state | $\lvert C \rvert$ | $k$ | $L$ | $U$ | obvious LB | max clusters | split LB | min splits | enacted splits |
|---|---|---|---|---|---|---|---|---|---|
| MT | 56 | 50 | 20,601 | 22,768 | 30 | 19 | 31 | 31 | 56 |
| NC | 100 | 50 | 198,349 | 219,227 | 20 | 28 | 22 | 22 | 24 |
| ND | 53 | 47 | 15,748 | 17,405 | 28 | 18 | 29 | 29 | 49 |
| NE | 93 | 49 | 38,030 | 42,032 | 26 | 21 | 28 | 28 | 37 |
| NH | 10 | 24 | 54,528 | 60,266 | 19 | 4 | 20 | 20 | 40 |
| NJ | 21 | 40 | 220,614 | 243,836 | 28 | 10 | 30 | 30 | 56 |
| NM | 33 | 42 | 47,897 | 52,938 | 28 | 13 | 29 | 29 | 64 |
| NV | 17 | 21 | 140,447 | 155,230 | 17 | 3 | 18 | 18 | 21 |
| NY | 62 | 63 | 304,623 | 336,687 | 42 | 20 | 43 | 43 | 66 |
| OH | 88 | 33 | 339,682 | 375,436 | 12 | 20 | 13 | 13 | 20 |
| OK | 77 | 48 | 78,363 | 86,610 | 22 | 25 | 23 | 23 | 59 |
| OR | 36 | 30 | 134,180 | 148,303 | 17 | 11 | 19 | 19 | 47 |
| PA | 67 | 50 | 247,052 | 273,056 | 26 | 23 | 27 | 27 | 47 |
| RI | 5 | 38 | 27,435 | 30,322 | 33 | 3 | 35 | 35 | 41 |
| SC | 46 | 46 | 105,707 | 116,833 | 26 | 16 | 30 | 30 | 68 |
| SD | 66 | 35 | 24,067 | 26,600 | 17 | 16 | 19 | 19 | 29 |
| TN | 95 | 33 | 198,949 | 219,890 | 13 | 20 | 13 | 13 | 15 |
| TX | 254 | 31 | 893,169 | 987,186 | 12 | 19 | 12 | 12 | 41 |
| UT | 29 | 29 | 107,174 | 118,455 | 22 | 7 | 22 | 22 | 41 |
| VA | 133 | 40 | 204,996 | 226,574 | 16 | 24 | 16 | 16 | 34 |
| VT | 14 | 30 | 20,365 | 22,507 | 23 | 6 | 24 | 24 | 18 |
| WA | 39 | 49 | 149,389 | 165,113 | 34 | 13 | 36 | 36 | 59 |
| WI | 72 | 33 | 169,668 | 187,527 | 12 | 20 | 13 | 13 | 73 |
| WV | 55 | 17 | 100,238 | 110,788 | 2 | 13 | 4 | 4 | 13 |
| WY | 23 | 30 | 18,267 | 20,189 | 17 | 10 | 20 | 20 | 25 |

# State House Results (1/2)

| state | $\lvert C \rvert$ | $k$ | $L$ | $U$ | obvious LB | max clusters | split LB | min splits | enacted splits |
|---|---|---|---|---|---|---|---|---|---|
| AK | 30 | 40 | 17,419 | 19,251 | 28 | 10 | 30 | 30 | 39 |
| AL | 67 | 105 | 45,458 | 50,242 | 71 | 29 | 76 | 76 | 115 |
| AR | 75 | 100 | 28,610 | 31,621 | 61 | 33 | 67 | 67 | 128 |
| AZ | 15 | 30 | 226,465 | 250,302 | 22 | 6 | 24 | 24 | 44 |
| CA | 58 | 80 | 469,517 | 518,939 | 57 | 20 | 60 | 60 | 95 |
| CO | 64 | 65 | 84,386 | 93,267 | 46 | 18 | 47 | 47 | 73 |
| CT | 8 | 151 | 22,687 | 25,074 | 139 | 8 | 143 | 143 | 162 |
| DE | 3 | 41 | 22,938 | 25,352 | 38 | 2 | 39 | 39 | 40 |
| FL | 67 | 120 | 170,511 | 188,459 | 89 | 26 | 94 | 94 | 112 |
| GA | 159 | 180 | 56,536 | 62,486 | 118 | 57 | 123 | 123 | 209 |
| IA | 99 | 100 | 30,309 | 33,498 | 54 | 38 | 62 | 62 | 92 |
| ID | 44 | 35 | 49,919 | 55,173 | 19 | 14 | 21 | 21 | 25 |
| IL | 102 | 118 | 103,152 | 114,009 | 85 | 31 | 87 | 87 | 220 |
| IN | 92 | 100 | 64,463 | 71,248 | 55 | 39 | 61 | 61 | 129 |
| KS | 105 | 125 | 22,328 | 24,678 | 87 | 34 | 91 | 91 | 127 |
| KY | 120 | 100 | 42,806 | 47,311 | 47 | 45 | 55 | 55 | 80 |
| LA | 64 | 105 | 42,142 | 46,577 | 72 | 29 | 76 | 76 | 116 |
| MA | 14 | 160 | 41,741 | 46,133 | 145 | 10 | 150 | 150 | 182 |
| MD | 24 | 47 | 124,859 | 138,001 | 35 | 10 | 37 | 37 | 67 |
| ME | 16 | 151 | 8,572 | 9,473 | 137 | 11 | 140 | 140 | 166 |
| MI | 83 | 110 | 87,032 | 96,192 | 73 | 32 | 78 | 78 | 154 |
| MN | 87 | 134 | 40,457 | 44,715 | 92 | 34 | 100 | 100 | 176 |
| MO | 115 | 163 | 35,873 | 39,648 | 110 | 47 | 116 | 116 | 137 |
| MS | 82 | 122 | 23,060 | 25,486 | 79 | 36 | 86 | 86 | 181 |

# State House Results (2/2)

| state | $\|C\|$ | $k$ | $L$ | $U$ | obvious LB | max clusters | split LB | min splits | enacted splits |
|---|---|---|---|---|---|---|---|---|---|
| MT | 56 | 100 | 10,301 | 11,384 | 74 | 22 | 78 | 78 | 99 |
| NC | 100 | 120 | 82,646 | 91,344 | 71 | 40 | 80 | 80 | 80 |
| ND | 53 | 47 | 15,748 | 17,405 | 28 | 18 | 29 | 29 | 53 |
| NH | 10 | 400 | 3,272 | 3,616 | 375 | 10 | 390 | 390 | 154 |
| NJ | 21 | 40 | 220,614 | 243,836 | 28 | 10 | 30 | 30 | 56 |
| NM | 33 | 70 | 28,738 | 31,762 | 52 | 15 | 55 | 55 | 86 |
| NV | 17 | 42 | 70,224 | 77,615 | 35 | 4 | 38 | 38 | 43 |
| NY | 62 | 150 | 127,942 | 141,408 | 115 | 26 | 124 | 124 | 179 |
| OH | 88 | 99 | 113,228 | 125,145 | 57 | 35 | 64 | 64 | 77 |
| OK | 77 | 101 | 37,242 | 41,161 | 67 | 30 | 71 | 71 | 134 |
| OR | 36 | 60 | 67,090 | 74,151 | 44 | 13 | 47 | 47 | 79 |
| PA | 67 | 203 | 60,851 | 67,255 | 158 | 39 | 164 | 164 | 186 |
| RI | 5 | 75 | 13,901 | 15,363 | 70 | 4 | 71 | 71 | 75 |
| SC | 46 | 124 | 39,214 | 43,341 | 94 | 24 | 100 | 100 | 145 |
| SD | 66 | 35 | 24,067 | 26,600 | 17 | 16 | 19 | 19 | 31 |
| TN | 95 | 99 | 66,317 | 73,296 | 55 | 36 | 63 | 63 | 74 |
| TX | 254 | 150 | 184,589 | 204,018 | 99 | 50 | 100 | 100 | 101 |
| UT | 29 | 75 | 41,441 | 45,802 | 59 | 12 | 63 | 63 | 72 |
| VA | 133 | 100 | 81,999 | 90,629 | 55 | 38 | 62 | 62 | 98 |
| VT | 14 | 150 | 4,073 | 4,501 | 137 | 12 | 138 | 138 | 118 |
| WA | 39 | 49 | 149,389 | 165,113 | 34 | 13 | 36 | 36 | 59 |
| WI | 72 | 99 | 56,556 | 62,509 | 63 | 30 | 69 | 69 | 159 |
| WV | 55 | 100 | 17,041 | 18,834 | 70 | 24 | 76 | 76 | 89 |
| WY | 23 | 60 | 9,134 | 10,094 | 44 | 11 | 49 | 49 | 56 |

# What About 1-Person Deviation?

## John Nagle (2022) writes:

Forcing districts to satisfy a 1-person deviation makes it “highly probable that the minimum number of county splits is uniquely given as the number of districts minus one.”

## Autry, Carter, Herschlag, Hunter, Mattingly (2021) write:

“It is reasonable to assume that there is no subset of counties that perfectly can accommodate a subset of the congressional districts. . . [which] may be used to demonstrate that $k-1$ splits is optimal.”

| state | $\lvert C\rvert$ | CD | SS | SH | state | $\lvert C\rvert$ | CD | SS | SH | state | $\lvert C\rvert$ | CD | SS | SH |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AK | 30 | ■ | ✗ | ✗ | MA | 14 | ✗ | ✗ | ✗ | OK | 77 | ✓ | ✓ | ✓ |
| AL | 67 | ✓ | ✓ | ✓ | MD | 24 | ✗ | ✗ | ✗ | OR | 36 | ✓ | ✓ | ✓ |
| AR | 75 | ✓ | ✓ | ✓ | ME | 16 | ✗ | ✗ | ✓ | PA | 67 | ✓ | ✓ | ✓ |
| AZ | 15 | ✗ | ✗ | ✗ | MI | 83 | ✓ | ✓ | ✓ | RI | 5 | ✗ | ✗ | ✗ |
| CA | 58 | ✓ | ✓ | ✓ | MN | 87 | ✓ | ✓ | ✓ | SC | 46 | ✓ | ✓ | ✓ |
| CO | 64 | ✓ | ✓ | ✓ | MO | 115 | ✓ | ✓ | ✓ | SD | 66 | ■ | ✓ | ✓ |
| CT | 8 | ✗ | ✗ | ✓ | MS | 82 | ✓ | ✓ | ✓ | TN | 95 | ✓ | ✓ | ✓ |
| DE | 3 | ■ | ✗ | ✗ | MT | 56 | ✓ | ✓ | ✓ | TX | 254 | ✓ | ✓ | ✓ |
| FL | 67 | ✓ | ✓ | ✓ | NC | 100 | ✓ | ✓ | ✓ | UT | 29 | ✗ | ✓ | ✓ |
| GA | 159 | ✓ | ✓ | ✓ | ND | 53 | ■ | ✓ | ✓ | VA | 133 | ✓ | ✓ | ✓ |
| IA | 99 | ✓ | ✓ | ✓ | NE | 93 | ✓ | ✓ | ■ | VT | 14 | ■ | ✗ | ✓ |
| ID | 44 | ✓ | ✓ | ✓ | NH | 10 | ✗ | ✗ | ✓ | WA | 39 | ✓ | ✓ | ✓ |
| IL | 102 | ✓ | ✓ | ✓ | NJ | 21 | ✗ | ✗ | ✗ | WI | 72 | ✓ | ✓ | ✓ |
| IN | 92 | ✓ | ✓ | ✓ | NM | 33 | ✓ | ✓ | ✓ | WV | 55 | ✓ | ✓ | ✓ |
| KS | 105 | ✓ | ✓ | ✓ | NV | 17 | ✗ | ✗ | ✗ | WY | 23 | ■ | ✓ | ✓ |
| KY | 120 | ✓ | ✓ | ✓ | NY | 62 | ✓ | ✓ | ✓ | | | | | |
| LA | 64 | ✓ | ✓ | ✓ | OH | 88 | ✓ | ✓ | ✓ | | | | | |

## Summary

- Many state constitutions say to preserve counties (cities, towns, ...).
- Prior to our work, the minimum number of splits was not known.
- We propose first exact approach: Cluster-Sketch-Detail.
- **Carter et al. (2020) are right!** Strong split duality *does* hold in practice.
- Many states' districting plans divide counties much more than necessary.

## Disclaimer

- We do not claim that our computer maps should be enacted in practice.
- They do not consider the Voting Rights Act or any laws that vary by state.
- Create your own maps using the county clusterings as starting points!

## Future work

- Determine the VRA-constrained minimum number of splits.
- Determine the tradeoffs between splits and population deviation.

# References I


[1] Eric A Autry, Daniel Carter, Gregory J Herschlag, Zach Hunter, and Jonathan C Mattingly.
**Metropolized multiscale forest recombination for redistricting.**
*Multiscale Modeling & Simulation*, 19(4):1885–1914, 2021.

[2] Paul S Bradley, Kristin P Bennett, and Ayhan Demiriz.
**Constrained *k*-means clustering.**
Technical Report MSR-TR-2000-65, Microsoft Research, Redmond, WA, May 2000.

[3] Daniel Carter, Zach Hunter, Dan Teague, Gregory Herschlag, and Jonathan Mattingly.
**Optimal legislative county clustering in North Carolina.**
*Statistics and Public Policy*, 7(1):19–29, 2020.

[4] Daryl DeFord, Moon Duchin, and Justin Solomon.
**Recombination: A family of Markov chains for redistricting.**
*Harvard Data Science Review*, 3(1), 2021.

[5] DRA.
**Dave's redistricting.**
`https://davesredistricting.org/`, 2023.
Accessed: 2023-02-17.

# References II


[6] Moon Duchin and Olivia Walch, editors.
***Political Geometry: Rethinking Redistricting in the US with Math, Law, and Everything In Between.***
Birkhauser, 2022.

[7] SW Hess, JB Weaver, HJ Siegfeldt, JN Whelan, and PA Zitlau.
**Nonpartisan political redistricting by computer.**
*Operations Research*, 13(6):998–1006, 1965.

[8] Cory McCartan and Kosuke Imai.
**Sequential Monte Carlo for sampling balanced and compact redistricting plans.**
*arXiv preprint arXiv:2008.06131*, 2020.

[9] John Nagle.
**Euler's formula determines the minimum number of splits in maps of election districts.**
*Available at SSRN 4115039*, 2022.

[10] Hamidreza Validi, Austin Buchanan, and Eugene Lykhovyd.
**Imposing contiguity constraints in political districting models.**
*Operations Research*, 70(2):867–892, 2022.