



# redistrictR

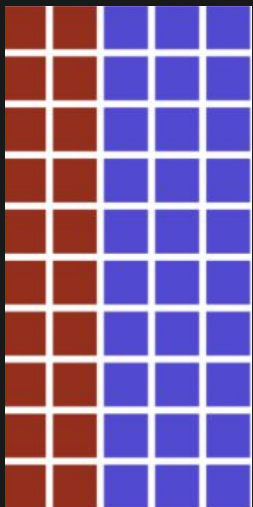
Joe Izenman | Nicholas Chen | Nikki Lee

# Gerrymander (v).

Manipulate districts with a bias during the mandatory redistricting every 10 years.

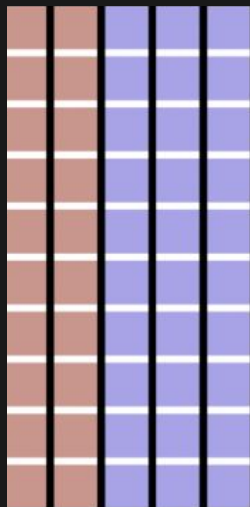
*Gerrymandering manipulates individuals' voting impact in elections.*

50 people



**60% Blue**  
40% Red

Perfect  
representation



**3 Blue**  
2 Red

Cracking



2 Blue  
3 Red

Packing



2 Blue  
3 Red

50 people



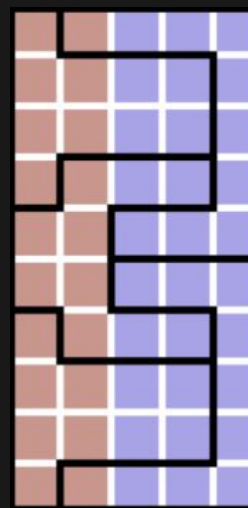
60% Blue  
40% Red

Perfect  
representation



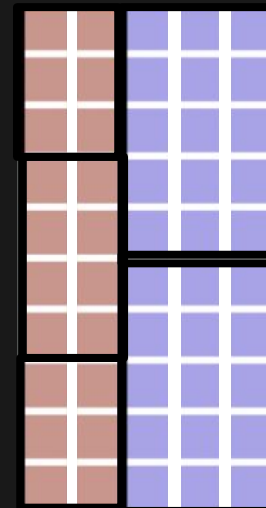
3 Blue  
2 Red

Cracking



2 Blue  
**3 Red**

Packing



2 Blue  
**3 Red**

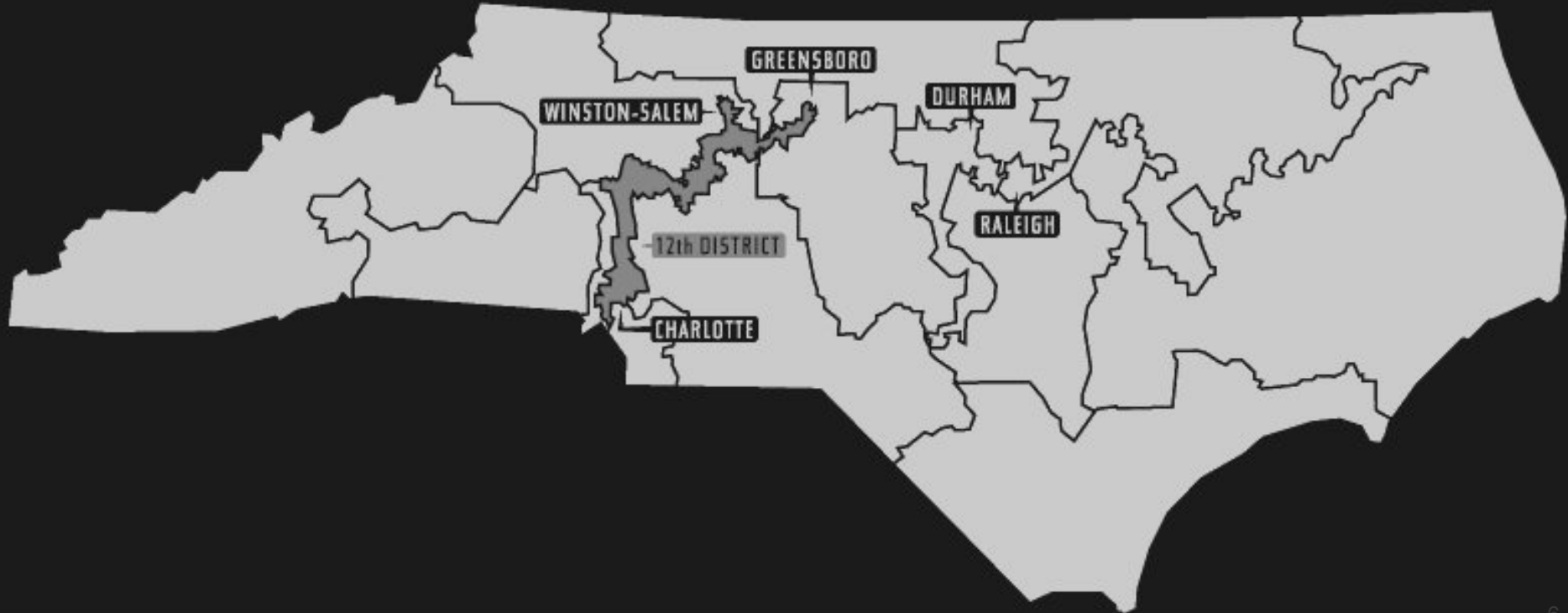
# 1812:

Governor Elbridge Gerry  
of Massachusetts

District of South Essex  
drawn like a salamander



# 2012: North Carolina



What is fair?

# Measures of Fairness & Scoring

Compactness

Vote Efficiency

Cluster Proximity



# The Building Blocks: Data Sources

- San Diego County
- Population and demographic data from US Census<sup>1</sup>
- Granular election data maintained by UC Berkeley<sup>2</sup>
- Join data from multiple sources
- Spatial tools for existing geographic and district relationships

[1] <https://cran.r-project.org/web/packages/acs/index.html>; <https://cran.r-project.org/web/packages/tigris/index.html>

[2] [http://statewidedatabase.org/d10/index\\_election.html](http://statewidedatabase.org/d10/index_election.html)

Demo

The word "Demo" is displayed in a sans-serif font. The letter 'D' is red, 'e' is white, 'm' is blue, and the final 'o' is white. A white hand cursor icon, consisting of a hand with the index finger pointing, is positioned over the 'o', with three concentric circles radiating from the tip of the finger to indicate a click or interaction.

# Optimization

3x10<sup>12</sup>50

Possible 5 district partitions of 1792 units

298,583,994,708,050,126,677,221,916,155,185,441,590,758,721,776,780,623,139,790,291,913,572,849,  
093,536,227,475,083,815,522,685,461,319,018,472,642,028,084,584,311,308,365,023,699,277,123,235,  
691,328,281,277,319,199,328,931,579,199,326,261,954,020,157,688,469,877,382,747,072,311,053,860,  
040,957,799,405,410,557,897,077,969,462,907,073,158,675,800,075,573,773,422,092,225,593,257,988,  
158,159,903,444,193,557,940,975,559,183,454,832,155,166,238,825,334,244,442,013,188,317,292,384,  
569,740,061,576,207,691,665,359,186,309,995,881,081,479,692,365,323,043,757,836,003,543,106,016,  
919,140,467,324,260,361,410,197,363,036,871,013,181,457,843,511,281,442,815,551,91,493,603,740,  
037,480,527,007,804,475,251,113,970,856,540,609,113,898,912,863,565,430,393,574,972,419,504,705,  
983,758,534,838,243,807,137,501,839,024,727,190,702,83,654,964,776,447,956,594,853,677,270,581,  
708,418,092,705,670,583,189,573,748,596,455,189,090,202,435,342,060,373,694,765,479,195,216,613,  
163,675,026,882,954,319,253,618,633,479,301,494,970,021,762,314,869,277,925,546,225,614,224,494,  
954,248,287,064,340,201,318,598,756,309,355,852,982,824,448,976,612,697,672,033,132,706,502,893,  
576,774,303,583,015,191,191,191,191,191,191,191,191,191,191,191,191,191,191,191,191,191,191,  
351,724,203,746,198,749,542,304,556,623,286,775,508,069,883,204,835,297,974,584,287,156,808,622,  
885,408,435,538,835,792,134,039,171,090,947,262,844,068,917,034,498,210,368,334,850,305,845,010,  
743,213,656,653,607,428,948,159,490,850,086,017,236,725,941,600,574,431,136,682,781,193,568,066,  
709,231,038,374,484,283,005,944,192,555,570,557,358,287,144,197,862,294,674,169,421,784,267,844,  
872,267,417,430,775,808,301,933,150

**Massive** solution space

**Highly Irregular** loss function

**Large plateaus**, with little feedback from small moves.

No **converging algorithm**, or method to evaluate overall quality

# The Evolutionary Algorithm

# Evolutionary Algorithm Purpose

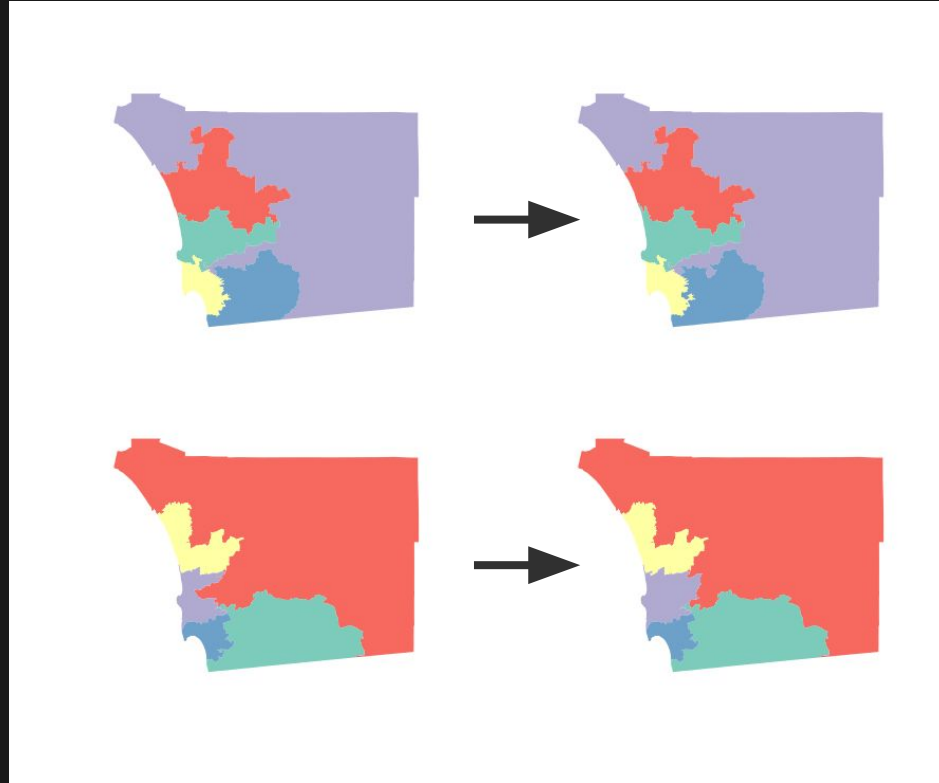
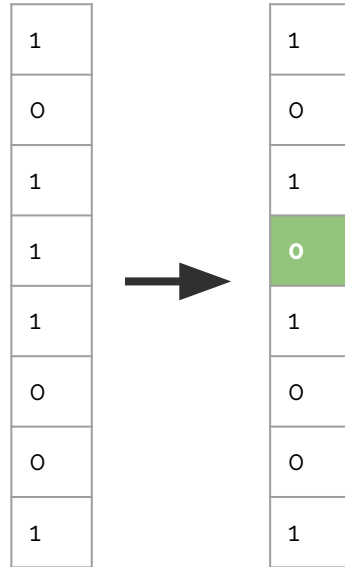
Without an elegant mathematical solution, the **optimization** problem becomes a **search** problem.

**Small moves** help find local improvement over a current solution.

**Large jumps** get away from local optima to drastically different options.

# Evolutionary Algorithm Operators

## Mutation

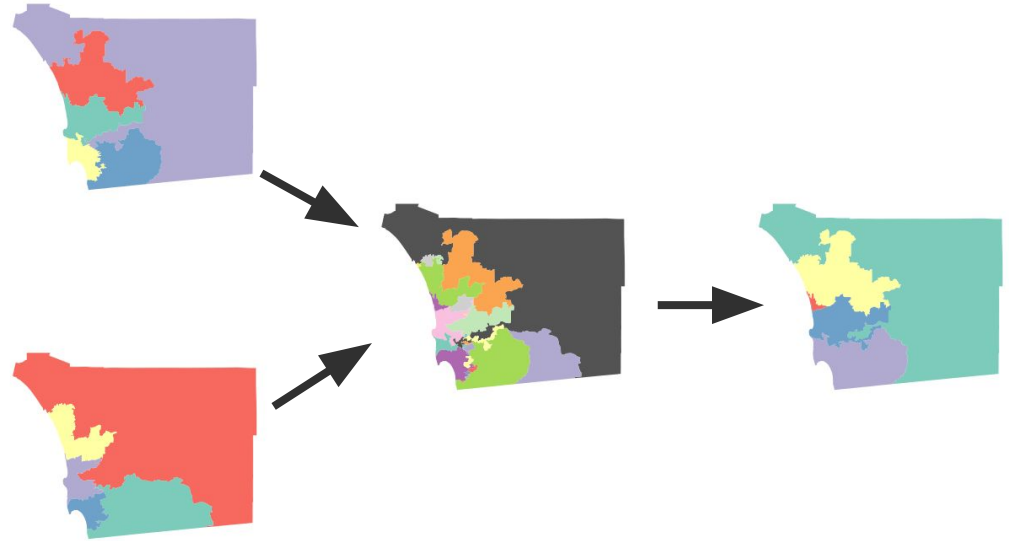




# Evolutionary Algorithm Operators

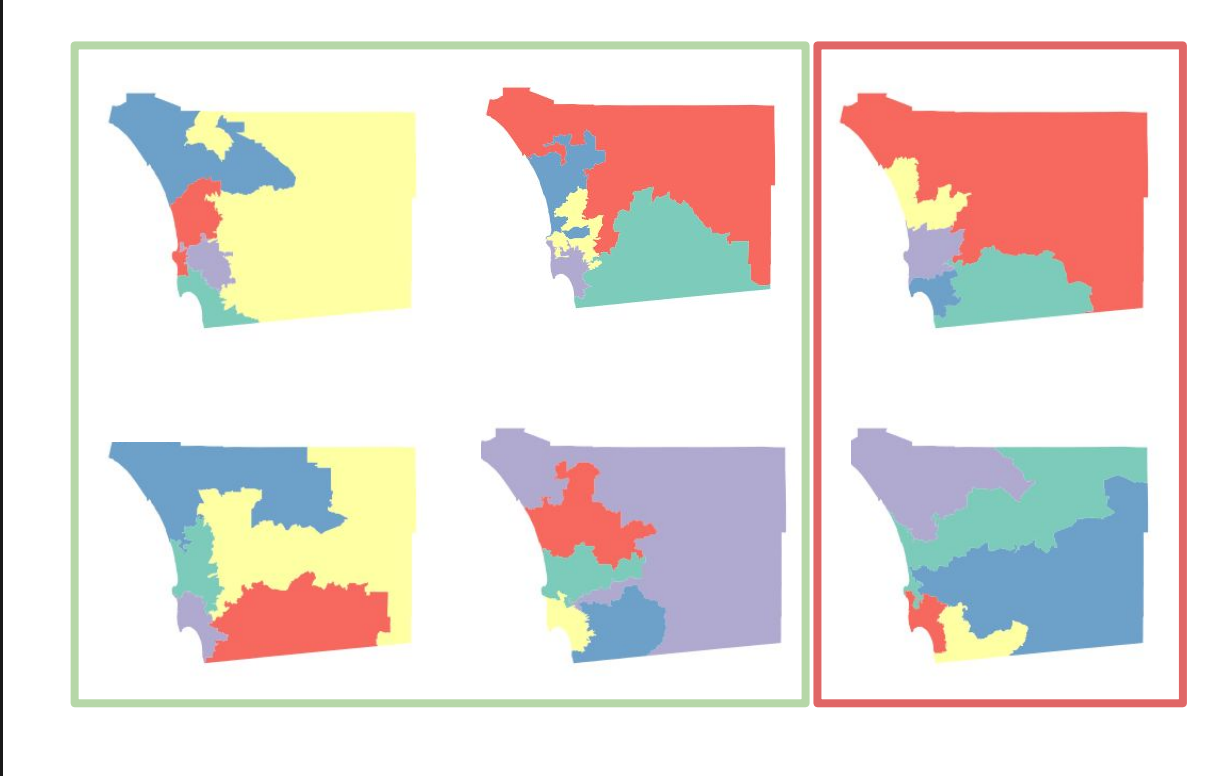
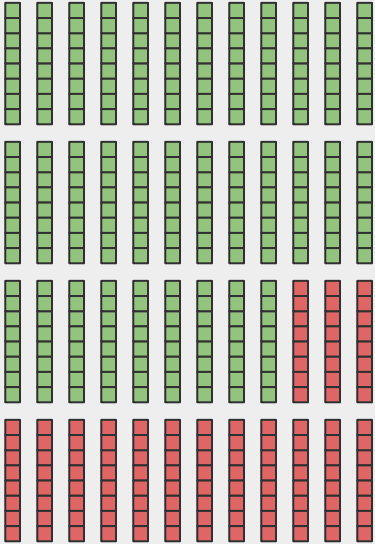
## Crossover

1	1	1	1
0	1	0	1
1	0	0	1
1	1	1	1
1	0	0	1
0	0	0	0
0	1	0	1
1	1	1	1



# Evolutionary Algorithm Operators

## Selection



# Evolutionary Algorithm Requirements

**R or Python** framework, for team familiarity.

**Custom operators** for unique geospatial constraints.

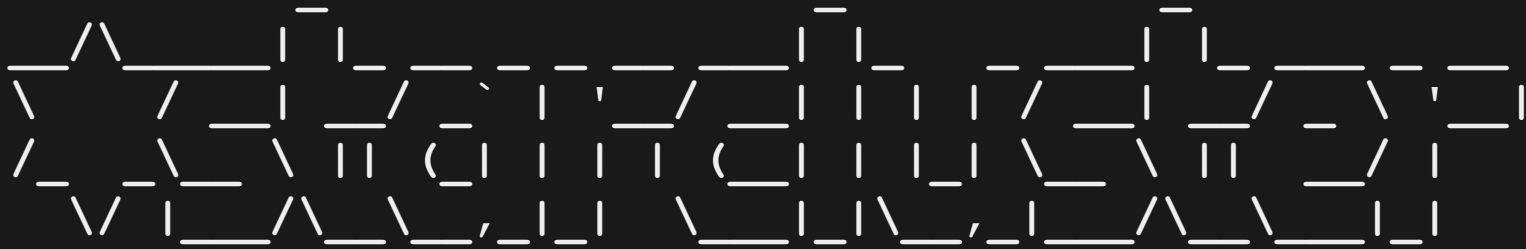
**Multi-objective fitness**, to optimize on several parameters.

Established methods for **parallel processing**.



DISTRIBUTED  
EVOLUTIONARY  
ALGORITHMS IN  
PYTHON

# Scaling Out



## StarCluster

# StarCluster Cost Comparison

Cores	Memory	Comparable EC2 Instance	\$ Per Week		
			Starcluster Built of t2.small Instances	Dedicated EC2 Instance	Starcluster Cost Savings per Week
2	4 GB	t2.medium	\$7.73	\$7.80	\$0.07
4	16 GB	t2.xlarge	\$15.46	\$31.18	\$15.72
8	32 GB	t2.2xlarge	\$30.91	\$62.36	\$31.45
48	192 GB	m5.12xlarge	\$185.47	\$387.07	\$201.60
96	384 GB	m5.24xlarge	\$370.94	\$774.14	\$403.20

# StarCluster Cost Comparison

Cores	Memory	Comparable EC2 Instance	\$ Per Week		
			Starcluster Built of t2.small Instances	Dedicated EC2 Instance	Starcluster Cost Savings per Week
2	4 GB	t2.medium	\$7.73	\$7.80	\$0.07
4	16 GB	t2.xlarge	\$15.46	\$31.18	\$15.72
8	32 GB	t2.2xlarge	\$30.91	\$62.36	\$31.45
48	192 GB	m5.12xlarge	\$185.47	\$387.07	\$201.60
96	384 GB	m5.24xlarge	\$370.94	\$774.14	\$403.20

# Parallelization Comparison

Technology	Cores	Time per Generation
Non Parallel Macbook Pro	1	600 seconds
Parallelized Macbook Pro	8	60 seconds
Starcluster	5	150 seconds
Starcluster	19	35 seconds

**RedistrictR** is (kind of) live at  
**[redistrictr.bigdatalittledata.net/app](https://redistrictr.bigdatalittledata.net/app)**