Gerrymandering and Graph Theory

Put Together and Presented by Dylan Phelan

Roadmap

1

Understanding Gerrymandering

Brief history of gerrymandering
How to gerrymander a district
How to quantify gerrymandering

2

Gerrymandering as a Graph Problem

Redistricting as a graph problem
Markov-Chain-Monte-Carlo simulation
Approaches to generation of new districts

3

Implementation: Gerrymandering-MCMC

Implementation of MCMC
Live demo of code
Investigating results
Future work

Understanding Gerrymandering and Related Terminology

- Redistricting
- Gerrymandering
- Precinct
- District
- Packing
- Cracking
- Efficiency Gap

Redistricting: Ensuring for Representative Democracy

- Instituted after the creation of a Congress, the **House of Representatives** was designed to **improve proportional representation** of States in the US Legislature.
- 435 Representatives total, divided proportionally* among the states, based the on population of the most recent census
  - There's a process called apportionment that determines the number of representatives each state gets; beyond the scope of the current discussion
- **Redistricting is the redrawing of district boundaries to account for changes in population** while maintaining, in general:
  - Compactness
  - Contiguity
  - Equal population
  - Preservation of existing political communities
  - Partisan fairness
  - Racial fairness (Thanks to the Voting Rights Act of 1965)

Problem:
Sitting Representatives are in charge of redistricting

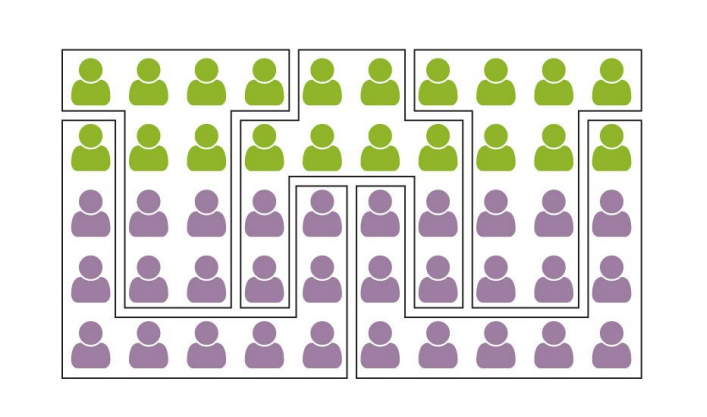
Gerrymandering: Among Boston's Many Great Exports

- 1812: Elbridge Gerry (Gov of Mass) passes a bill creating this partisan redistricting
- Boston news outlets balked at the gnarled and tendril-like nature of the new district, likening it to a salamander
- Gerry + Salamander = Gerrymander
- In general, Gerrymandering is the practice of redistricting to benefit one's political party
    - Racial Gerrymandering - strictly illegal, easier to prove because of censuses
    - Partisan Gerrymandering - illegal, but much harder to make a court-case for.

Elkanah Tisdale (1771-1835) - Originally published in the Boston Centinel, 1812.

Precinct: Fundamental Geographic Unit of a State

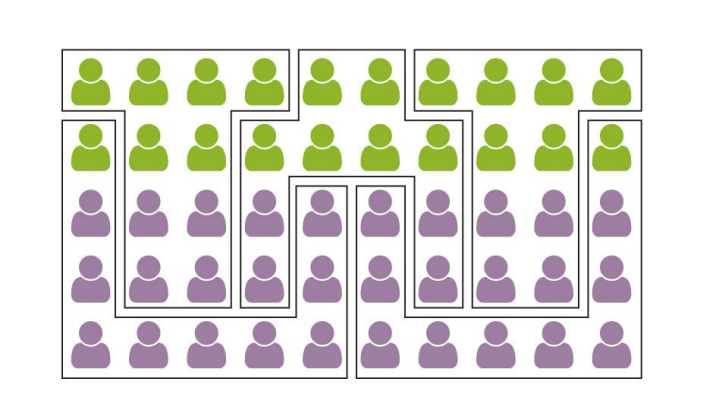
- Precinct: The smallest unit of space up for shuffling in a redistricting plan
  - Sometimes these are census tracts
  - Sometimes these can be as large as counties
  - Sufficiently abstract for our purposes
- Geographical entities with population numbers, demographic information, and other factors to consider in drawing up district boundaries



Precincts in this picture correspond to the individual people.
How Does the Efficiency Gap Measure Partisan Gerrymandering - Cameron, Darla
https://www.washingtonpost.com/graphics/2017/politics/courts-law/gerrymander/

District: Group of Precincts that Elect one Representative

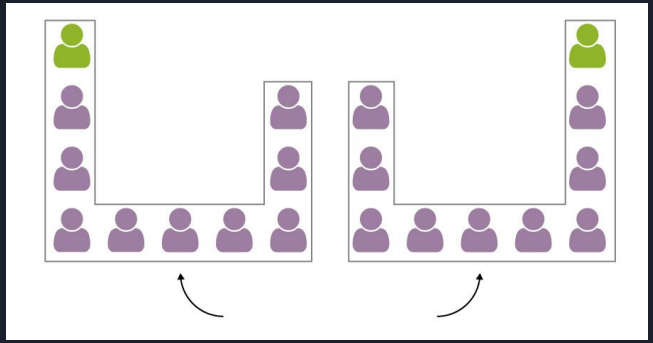
- One district corresponds to one representative.
- Districts are geographical entities encompassing multiple precincts.
- Districts need to meet the criteria discussed earlier:
  - Somewhat compact
  - Somewhat contiguous*
    - Islands? Disconnected parts of states? Corner cases abound
  - Somewhat equal in population, equal representation in race



Districts are defined with black borders:
How Does the Efficiency Gap Measure Partisan Gerrymandering - Cameron, Darla, https://www.washingtonpost.com/graphics/2017/politics/courts-law/gerrymander/

Packing: Pack Many Votes into Individual Districts

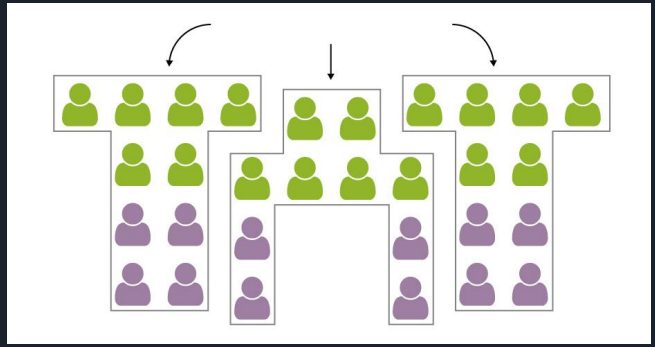
- In order to reduce the impact a group of voters can have, you **put all like-minded voters in a single district**
- Ensures they win a particular district
- Reduces the number of districts those voters can impact
- Increases the number of wasted votes due to surplus
    - Once >50% of the district votes one way, any additional votes are a surplus beyond what was needed to win



An example of packed districts:
How Does the Efficiency Gap Measure Partisan Gerrymandering - Cameron, Darla, https://www.washingtonpost.com/graphics/2017/politics/courts-law/gerrymander/

Cracking: Crack a Group Across Multiple Districts

- In order to reduce the impact a group of voters can have, you **split like-minded voters across multiple district**
- Spreads out voters for one part across enough districts to dilute their impact below require thresholds
- Increases the number of wasted votes due to losing districts
    - Any votes cast for Party A in a district that elects Party B are considered wasted



An example of cracked districts:
How Does the Efficiency Gap Measure Partisan Gerrymandering - Cameron, Darla, https://www.washingtonpost.com/graphics/2017/politics/courts-law/gerrymander/

Efficiency Gap: One Measure of Gerrymandering

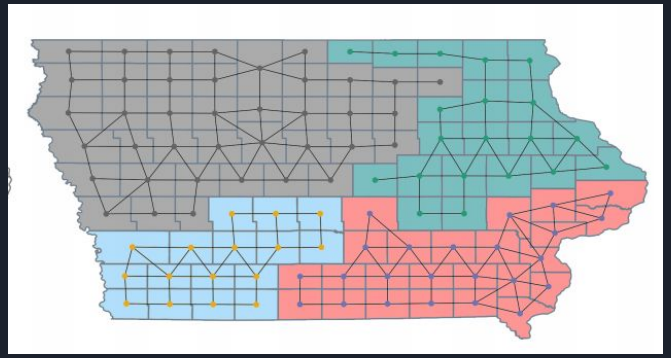
- Metric proposed in 2014 by Stephanopoulos and McGhee to quantify gerrymandering in terms of wasted votes
- “The efficiency gap, then, is simply the difference between the parties’ respective wasted votes, divided by the total number of votes cast in the election”
- Basic Formula:
  - (Party_A_Wasted_Votes - Party_B_Wasted_Votes) / Total Votes
- Important Note: This is only one measure in a world of many
  - Efficiency Gap has it’s benefits and drawbacks, detailed in the literature[1]
  - Other measures exist (Polsby-Popper score, variants on that)
  - We’ll focus on this measure for implementations-sake, but future work could expand to include other measures

[1] For more detailed comments on the pros/cons of Efficiency Gap, see https://www.theatlantic.com/science/archive/2018/01/efficiency-gap-gerrymandering/551492/ and section three of https://arxiv.org/pdf/1801.02064.pdf

Gerrymandering as a Graph Problem

Representing State Maps as Graphs: Looking at the Dual of our Planar Graph

- Consider a state graph, broken up by precinct.
- By looking at this dual graph, we create a graphical representation of the same geometry.
- Precincts:
  - Individual faces in the original graph;
  - A vertex in the
- Districts:
  - Collections of bordering faces in the original graph;
  - Connected subgraphs induced by a subset of vertices in the dual graph
- Redistricting:
  - Mutually exclusive but collectively exhaustive sets of faces that meet some set of minimum criteria (contiguity; equal pop; etc)
  - A mutually exclusive…. set of induced subgraphs that meet our minimum criteria



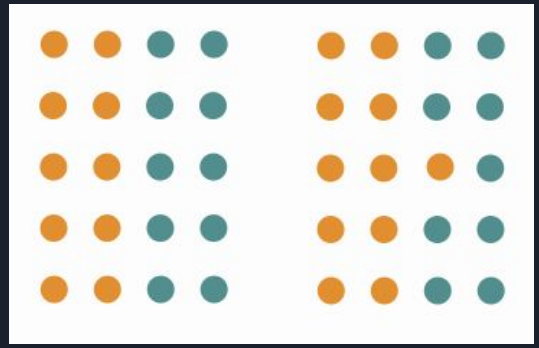
Recombination: A family of Markov chains for redistricting; Deford, Duchin and Solomon
https://arxiv.org/pdf/1911.05725.pdf

Markov Chain Monte Carlo Simulation: Helps us Measure Outliership

- Markov Chain Monte Carlo methods in short:
  - Some distributions are quite complicated, and sampling from them can be complicated
    - E.g. the distribution of all possible
  - Markov Chain Monte Carlo methods look at simulating that complicated distribution by sampling multiple random walks on an associated graph (our markov chain)
  - The random walk in our case - a walk over possible partitions of districts
- The idea: By talking a random walk on that space of possible district plans, we can generate an ensemble of redistricting plans whose distribution resembles the population of all possible plans.
- Key detail: **By comparing a proposed plan to the ensemble of generated plans, we can determine how much of an outlier a proposed plan is.**

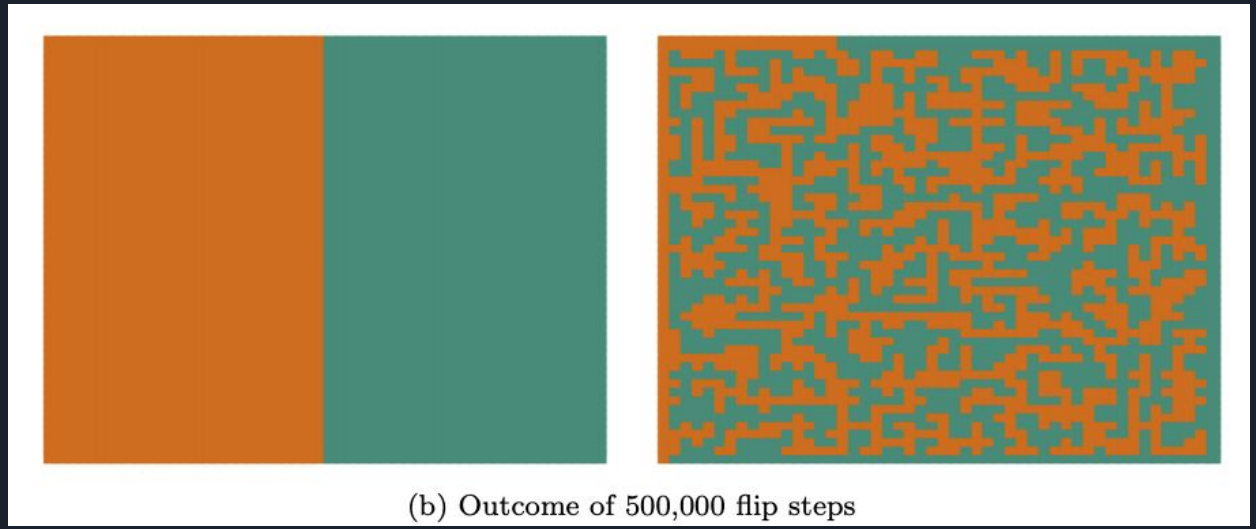
The Flip Method: One Approach to Generating New Districts

- Open question: How we generate new districting plans from old ones
- Solution: The Flip Method
  - Take an existing valid redistricting plan
  - Find two adjacent districts
  - **Pick one node on the boundary of adjacent districts, and flip it's district**
  - If this new plan is invalid, try again;
  - Else, add to our sample of possible plans
- Pro: Straightforward implementation
- Con: Can get stuck in peculiar local valleys of exploration



Recombination: A family of Markov chains for redistricting; Deford, Duchin and Solomon
https://arxiv.org/pdf/1911.05725.pdf

Shortcomings of the Flip Method

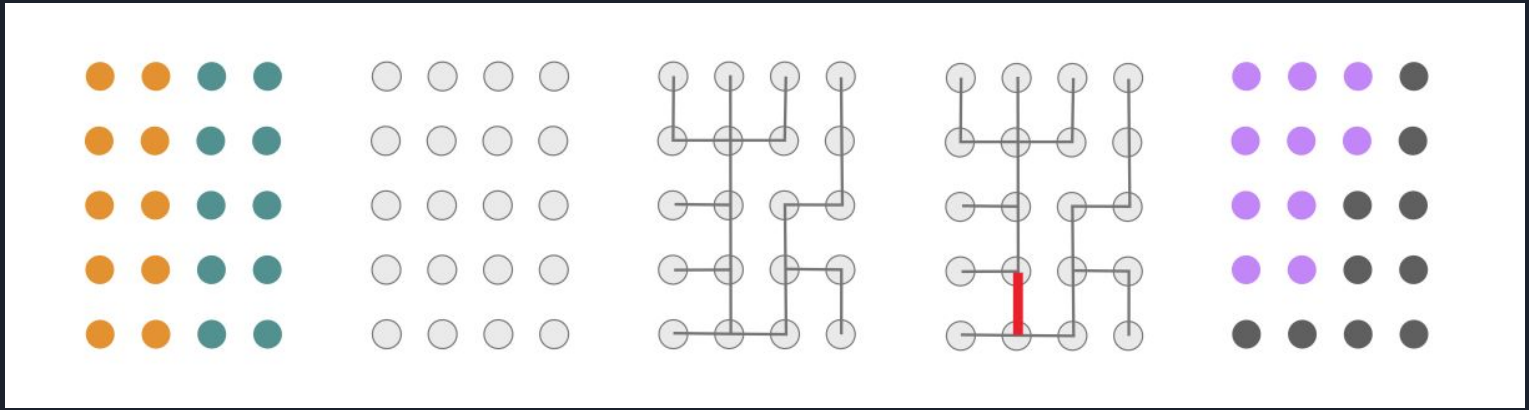


(b) Outcome of 500,000 flip steps

Recombination: A family of Markov chains for redistricting;
Deford, Duchin and Solomon
https://arxiv.org/pdf/1911.05725.pdf

New Proposal:
The Recombination Method

- New proposal (Nov 2019) by Deford, Duchin and Solomon - Recombination
- Instead of flipping individual precincts, why don't you combine and entire adjacent districts?



Recombination: A family of Markov chains for redistricting;
Deford, Duchin and Solomon
https://arxiv.org/pdf/1911.05725.pdf

Implementation: Gerrymandering-MCMC

Emulating the Work of Metric Geometry and Gerrymandering Group

- Based on a tool created by the MGGG (Moon Duchin and others) - GerryChain
- Create a python command line interface that can:
  - Take in a path to a “graph” file corresponding to some state’s dual graph;
    - That graph contains not only vertices and edges, but also relevant metadata about a precinct like it’s population, it’s voting history (e.g. Democrat or Republican), and it’s current district
  - Generate an ensemble of possible redistricting plans according to some required criteria
    - Contiguity is a must
    - Rough equality among populations - equality within some error margin
  - Use that ensemble to plot various metrics
    - Efficiency Gap
    - Number of districts democrat vs republican
  - Highlight where our original proposed graph lies relative to the ensemble’s distribution

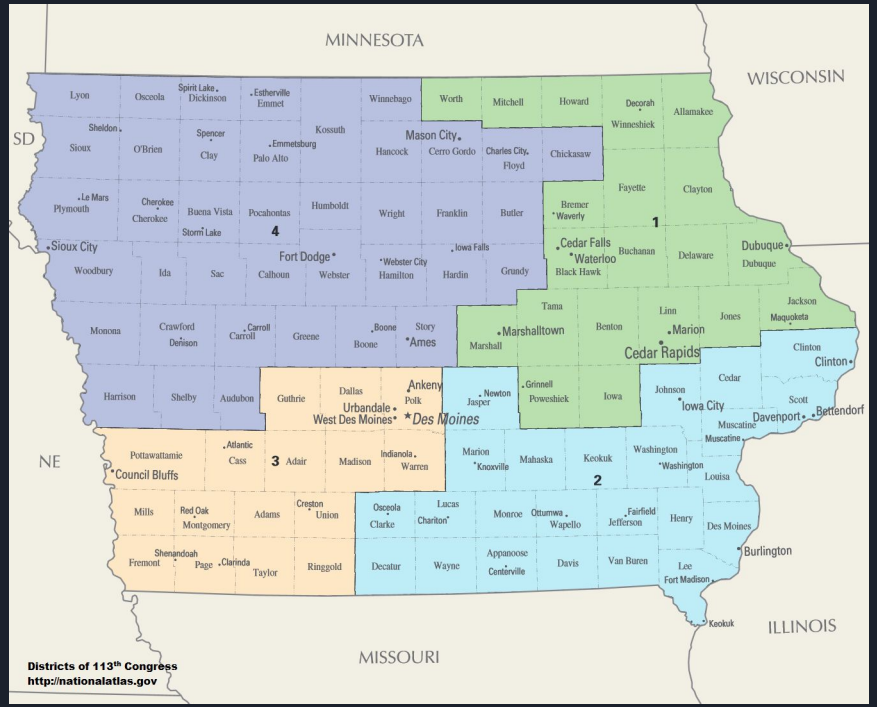
Basic Architecture and Command Line Interface

```
python cli.py [-h] [-g GRAPH_FILE] [-c COOLING_PERIOD] [-r ROUNDS] [-v]

Use MCMC Simulation to generate districting plans and plot relevant key
statistics to illustrate the possibility that a source plan was gerrymandered

optional arguments:
  -h, --help            show this help message and exit
  -g GRAPH_FILE, --graph_file GRAPH_FILE
                        A path to a potential districting plan specified in
                        this projects proprietary json schema; defaults to
                        ./src/data/iowa.json
  -c COOLING_PERIOD, --cooling_period COOLING_PERIOD
                        The number of plans you'd like to generate _before_
                        counting them towards your ensemble; defaults to 50
  -r ROUNDS, --rounds ROUNDS
                        The number of plans you'd like to generate and include
                        in your ensemble; defaults to 200
  -v, --verbose         Include this flag if you'd like real-time output to
                        the console
```

Gerrymandering-MCMC Part 1: Modeling Iowa




```
{
    "1": {
        "adjacent_nodes": [
            "2",
            "12"
        ],
        "population": 50,
        "voting_history": "D",
        "district": "A"
    },
    "2": {
        "adjacent_nodes": [
            "3",
            "13"
        ],
        "population": 50,
        "voting_history": "D",
        "district": "A"
    },
```

Gerrymandering-MCMC Part 2: Generating Alternatives Using Recom

Algorithm 6: `ReCom` (Spanning tree bipartitioning)

Input: Dual graph $G = (V, E)$, the current partition P, population tolerance ε
Output: The next partition Q

Select: $(u, v) \in \partial P$ uniformly
Set $W_1 = P(u)$ and $W_2 = P(v)$
Form the induced subgraph H of G on the nodes of $W_1 \cup W_2$.
Initialize: Cuttable = `False`
while *Cuttable = False* **do**
  Sample a spanning tree T of H
  Let EdgeList = []
  for *edge in T* **do**
    Let $T_1, T_2 = T \setminus edge$
    if $|T_1| - |T_2| < \varepsilon |T|$ **then**
      Add edge to EdgeList
      Cuttable = True
    end
  end
end
Select cut uniformly from EdgeList
Let $R = T \setminus cut$

$$\text{Define } Q(v) = \begin{cases} R(v) & v \in H \\ P(v) & \text{otherwise} \end{cases}$$

Return: Q

```
Randomizing the seed plan 0
Randomizing the seed plan 25
Finding recomb ...  0
Finding recomb ...  20
Finding recomb ...  40
Finding recomb ...  60
Finding recomb ...  80
Finding recomb ...  100
Finding recomb ...  120
Finding recomb ...  140
Finding recomb ...  160
Finding recomb ...  180
```

Gerrymandering-MCMC Part 3: Investigating Results

Problem: Looking to validate my algorithm; laborious to manually create redistricting proposals for Iowa that are/aren't gerrymandered

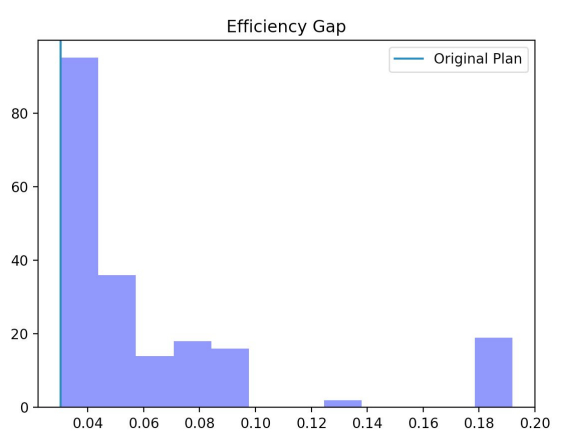
Solution: Leave the districts the same, but test data where party-allegiance is modified instead of districts

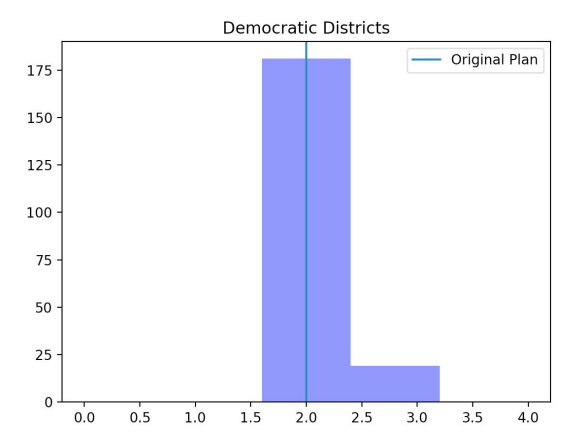
Consider two models of Iowa:

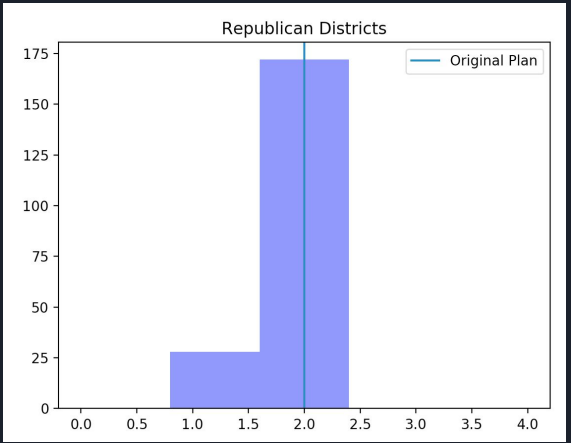
- One where party-allegiance is randomly assigned to each precinct;
- One where party-allegiance is moved around such that the particular redistricting plan packs and cracks across party lines

Gerrymandering-MCMC Part 3: Investigating Results

1. Uniformly distribute and randomly-assign party allegiance across precincts

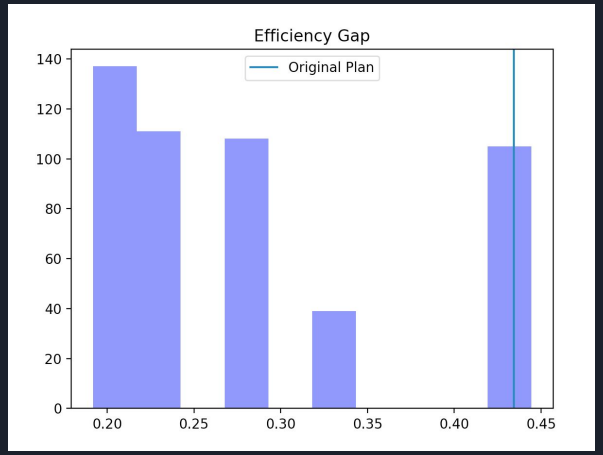


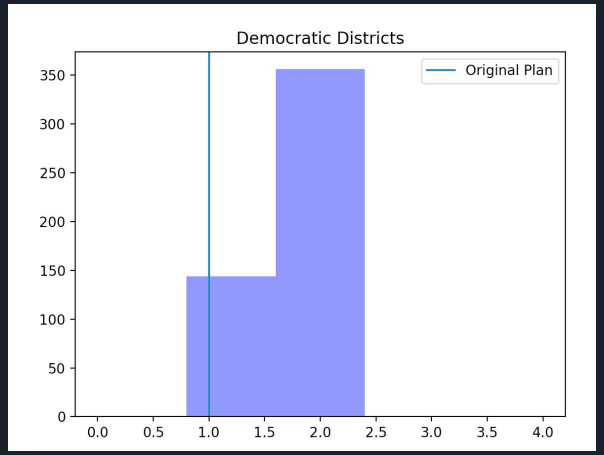


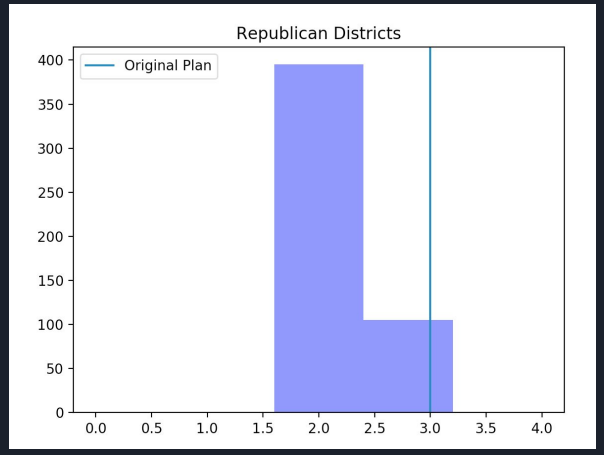


Gerrymandering-MCMC Part 3: Investigating Results

2. Fixed the parties of particular precincts s.t. Districts are very







Demo
Time Permitting

Next Steps and Future Work

- Compare and contrast alternative approaches to generating new plans
  - Novel work being done modifying the flip step; at alternative generation mechanisms
- Introduce additional requirements on what constitutes a valid district
  - Race and other demographics;
- Introduce more complex models of voting history
  - Use more sophisticated input data to deal with the complex problem of predicting voter turnout
- Support more established data formats
  - .shp files are the status quo w/r/t work done by others
- General improvements to performance
- Get involved in local initiatives to support for more democratic legislation
  - Organizations like Voter Choice Mass trying to enact alternatives to first-past-the-post voting (e.g. ranked choice)
  - Another shout out to the Metric Geometry and Gerrymandering Group, and their efforts to inform quantitatively-driven rulings on what partisan gerrymandering looks like.

Bibliography:

Videos:

- CGP Grey; Gerrymandering Explained: https://www.youtube.com/watch?v=Mky11UJb9AY&t=1s
- CGP Grey; The Problems with First Past the Post Voting Explained: https://www.youtube.com/watch?v=s7tWHJfhiyo
- Moon Duchin; Political Geometry: https://www.youtube.com/watch?v=VddLOevo7QY
- Moon Duchin; Redistricting and Representation: https://www.youtube.com/watch?v=PrUqbCnEpus
- Wesley Pegden; Mathematics of Gerrymandering: https://www.youtube.com/watch?v=NAS4AsPi1q4

Bibliography:

Additional Reading:

- Benjamin Fifield, Michael Higgins, Kosuke Imai and Alexander Tarr; A New Automated Redistricting Simulator Using Markov Chain Monte Carlo: https://imai.fas.harvard.edu/research/files/redist.pdf
- Darla Cameron; How Does the Efficiency Gap Measure Partisan Gerrymandering: https://www.washingtonpost.com/graphics/2017/politics/courts-law/gerrymander/
- Daryl DeFord, Moon Duchin, and Justin Solomon; Recombination: A family of Markov chains for redistricting: https://arxiv.org/pdf/1911.05725.pdf
- Moon Duchin; Gerrymandering Metrics: How to Measure, What's the Baseline: https://arxiv.org/pdf/1801.02064.pdf
- Sam Kean; The Flaw in America's 'Holy Grail' Against Gerrymandering: https://www.theatlantic.com/science/archive/2018/01/efficiency-gap-gerrymandering/551492/
- Nicholas Stephanopoulos and Eric McGhee; Partisan Gerrymandering and the Efficiency Gap: https://ssrn.com/abstract=2457468