# **IE5311 Principles of Operations Research Iowa Redistricting Proposal**

**Team: Iowa Redistricting proposal** 

**Project members:** 

Chakridhar Karnati-R11867662 Charan Bokka-R11909017 Sai Ram Reddy-R11915915

## **Executive Summary Letter**

The current scenario of congressional redistricting in the state of Iowa is prompted by the recently conducted decennial census, providing updated data on population and demographic changes. As consultants entrusted with this task, our primary responsibility is to ensure that the proposed redistricting aligns with legal requirements, remains transparent, and is easily comprehensible to the public.

Our methodology involves a meticulous approach to adhering to legislative mandates, prioritizing equal representation across districts, and considering the unique compactness measurements outlined by Iowa's legislation—specifically, Length-Width and Perimeter methods. The focus is on maintaining nearly identical populations in each district to safeguard individual voting power.

While contiguity, ensuring that all counties within a district share a common border, is a fundamental criterion, our emphasis on compactness aims to eliminate ambiguity and facilitate a clear understanding of the redistricting rationale. Length-Width measures the longest east-to-west distance divided by the north-to-south distance, while Perimeter minimizes the total perimeter length of all districts.

Our proposal seeks to address the challenges posed by a diverse and expansive population of the state, ensuring fair representation and adherence to legal scrutiny. The goal is to minimize the population deviation to less than 1%, guaranteeing a balanced distribution of voting power. However, it's important to note a limitation in achieving the smallest population difference, as it is not the primary objective.

In presenting our proposal, we commit to providing a comprehensive documentation of our methodology and the data utilized. This approach not only upholds legal scrutiny but also emphasizes transparency, fostering public trust. Our endeavor is rooted in the delicate balance between the legal framework guiding redistricting and the aspiration for fair representation, ensuring that our recommendations stand up to legal examination and are easily comprehensible to the citizens of Iowa.

#### **Introduction:**

The United States has a complex government with three main parts: the Executive, Judicial, and Legislative branches. These branches have their own powers and jobs, and they are crucial for our democratic system. Citizens play a key role by electing representatives to these branches. However, with a large population of 332 million people spread across diverse areas, making sure everyone is fairly represented is a tough task. To tackle this challenge, the government uses a method called districting. This involves drawing political borders to distribute power without any biases that could go against the principle of equality in governance.

Now, our focus is on creating a tool for redistricting in Iowa. But it's not just about meeting our own needs; we're committed to following laws and expectations that ensure fair representation and governance. Our goal is to find a balance between wanting fairness in representation and respecting the legal rules that guide our democratic processes.

In simpler terms, the U.S. government has different branches that do specific jobs, and citizens pick leaders for these branches. Making sure everyone is represented fairly is hard because there are so many people spread out in different places. The government uses a method called districting to draw borders and share power without being unfair. Our job is to create a tool for redistricting in Iowa, following laws to make sure representation is fair and everyone's voice is heard in our democracy.

## Criteria:

Our guidelines are clear-cut, we need to follow the rules set by the government and make sure the districts look organized. The government insists that all districts in the state should have almost the same number of people to make sure everyone's vote counts equally. Another responsibility we have is following the state's rules for redistricting, which include making sure districts are connected and compact. Contiguity, a term mostly used in redistricting discussions, means that all counties in a district share a border. Compactness is a bit more complex and changes depending on the area. In Iowa, the state's rules offer two ways to measure compactness: Length-Width and Perimeter. Length-Width is about finding the longest distance from east to west and dividing it by the distance from north to south. Perimeter, on the other hand, is simpler. It focuses on making the distance around the edges of all districts as short as possible.

In simpler terms, we have simple rules to follow make sure the number of people in each district is nearly the same, connect all the counties in a district, and keep the districts organized. The state also has rules about compactness, which is how close together the districts are. They measure it using Length-Width and Perimeter methods, which look at the distances in different ways. The goal is to have districts that are fair and make sense.

#### **Problem statement:**

Every ten years, the United States conducts a census to show how the population and different characteristics have changed within its borders. With the latest information from the U.S. census, we've been asked to help in the process of redrawing district boundaries for the state of Iowa. Our job is to make sure that all our suggestions follow the law and rules while being clear and easy to understand. This means we need to carefully explain how we did our calculations and the data we used. Our goal is to present our proposal in a way that makes it easy for everyone to understand and leaves no confusion about why we're suggesting certain changes and how we came to our conclusions.

#### **OR Model:**

#### In words:

The aim of our model is to minimize the deviation of all districts, and we have found an acceptable solution if this value is less than one percent between the largest and smallest populous districts. Our model commences by importing the data of all counties of Iowa and their populations and assigning them to their respective nodes. Subsequently, we determine the desired population at the upper and lower limits of the imposed 1% deviation between the districts. To facilitate this, we introduce two variables: x, a flag that is true when county i is assigned to district j or zero otherwise, and y, a flag that is set to true when the edge from u to v is cut or 0 otherwise, thereby defining the border of the congressional district. Our objective function is defined to minimize the weighted cut edges, which are equivalent to the district perimeter lengths. We then proceed to implement the necessary constraints to the model. Initially, we ensure that each county i is assigned to at least one district j. We subsequently apply the constraint that mandates all districts to fall within the upper and lower limits defined by the population. Additionally, we incorporate a constraint that utilizes variable y to declare edge {U,V} cut if u is assigned to district i but v is not. Our next crucial constraint guarantees the contiguity of the districts, for which we plan to utilize a Mixed Integer Program. By employing a max k-cut, we assign a root to each district and regulate a flow from each root. Each county within a district then absorbs a portion of this flow, with each root having a flow limit. Once all this is set up, we proceed to optimize our model.

## In math:

Sets: nodes= {Counties of Iowa & their populations} edges= {Set of all edges shared by county i}

Variables: xij=1 when county i is a member of j, o.w. 0

yuv= 1 when edge u-v is cut, o.w. 0

rij= 1 when county i is the root of district j, o.w. 0

fuv = flow along edge u-v

Parameters: u= references county i along edge

j= references county j along edge U= Upper limits of district population L= Lower limit of district population Objective function:  $\min \sum_{i=0}^{\infty} (y_u,v)^* \text{ [edge] } (u,v)$ 

Constraints:  $\sum_{j=0}^{\infty} (j=0)^k x_{ij} = 1 \forall i = \{nodes\}$ 

 $\Sigma$  [nodes] \_population\*x\_ij<=U  $\forall$ j=range(k)

 $\Sigma$  [nodes] \_population\*x\_ij>=L  $\forall$ j=range(k)

xuj- xvj <=  $yuv \forall j$ = range(k)

 $\Sigma \text{ } \llbracket (f \rrbracket \text{ } \_ji\text{-}f\_ij) \forall j\text{=}neighbors}\_i\text{>=}1\text{-}M*\sum\_(i\text{=}1,j\text{=}1)^{\land}(nodes,range(k))$ 

 $f_i = M^* (1-yij) \forall i \in edges, j \in edges$ 

Sets: Nodes: Counties of Iowa and their population data.

Edges: Collection of all boundary edges shared by county i.

Variables:

xij: Binary variable indicating whether county i belongs to district j.

yuv: Binary variable denoting whether edge u-v is cut.

rij: Binary variable determining if county i is the root of district j.

fuv: Variable representing the flow along edge u-v.

### Parameters:

u: Reference to county i along edge.

j: Reference to county j along edge.

U: Upper limit of district population.

L: Lower limit of district population.

Objective: Minimize the summation of yuv (indicating cut edges) across all edges u-v.

Constraints: Each county must belong to exactly one district.

Ensure that the total population in each district is within the specified upper and lower limits. Constraint to handle edge cuts by ensuring that if county i and county j belong to different districts, their boundary edge u-v is cut.

Constraints related to flow to ensure connectivity between districts while minimizing the number of cut edges.

This formulation outlines the necessary components and conditions required for a redistricting plan in Iowa, focusing on fair representation and adherence to population constraints while minimizing the disruption of county boundaries.

## Code (Iowa-Minimum-Perimeter-with-Contiguity)

```
In [1]: 1 from gerrychain import Graph
In [27]:
            1 from gerrychain import Graph
                 import os
             4 filepath = '/Users/cherr/anaconda-gurobi/gurobi classes/'
             5 filename = 'Iowa_county.json'
             7 # Verify if the file is there at the given location.
             8 file path = os.path.join(filepath, filename)
             9 if os.path.exists(file_path):
            10
                      # Proceed to read the file if it exists.
            11
                      Gr = Graph.from_json(file_path)
             12
                       # Use the graph to perform operations.
            13 else:
                      print(f"File '{filename}' not found in the specified directory.")
            14
            15
            1 for node in Gr.nodes:
 In [5]:
                      Gr.nodes[node]['TOTPOP'] = Gr.nodes[node]['P0010001']
 In [6]: 1 # Let us apply a population deviation of 1% (+/-0.5%).
            2 deviation = 0.01
           4 import math
                              # of districts
           6 total_population = sum( Gr.nodes[node]['TOTPOP'] for node in Gr.nodes )
          L = math.ceil( ( 1 - deviation / 2 ) * total_population / k )
U = math.floor( ( 1 + deviation / 2 ) * total_population / k )
print("Using L =",L,"and U =",U,"and k =",k)
          Using L = 793605 and U = 801580 and k = 4
 In [7]: 1 import gurobipy as gp
           2 from gurobipy import GRB
           5 model = gp.Model()
           8 a = model.addVars(G.nodes, k, vtype=GRB.BINARY) # When county i is allocated to district j, x[i,j] = 1.
9 b = model.addVars(G.edges, vtype=GRB.BINARY) # When edge \{u,v\} is sliced, y[u,v] = 1.
          Academic license - for non-commercial use only - expires 2024-09-10
   In [8]: 1 model.setObjective(gp.quicksum(Gr.edges[u, v]['shared_perim'] * b[u, v] for u, v in Gr.edges), GRB.MINIMIZE)
 In [10]: 1 # add Constraint such that each county is allocated to one district
             2 model.addConstrs(gp.quicksum(a[i, j] for j in range(k)) == 1 for i in Gr.nodes)
             4 \# add Constraints such that each district has population at least L and at most U
             5 for j in range(k):
                    model.addConstr(gp.quicksum(Gr.nodes[i]['TOTPOP'] * a[i, j] for i in Gr.nodes) >= L)
model.addConstr(gp.quicksum(Gr.nodes[i]['TOTPOP'] * a[i, j] for i in Gr.nodes) <= U)
             9 # add Constraints to Define cut edges based on district allocations
            10 for u, v in Gr.edges:
                    for j in range(k):
                        model.addConstr(a[u, j] - a[v, j] \leftarrow b[u, v])
            14 model.update()
```

```
In [11]: 1 # Let us now introduce contiguity constraints and solve the model again.
              2 # The Hojny et al. contiguity constraints will be applied (MPC, 2021).
              4 # Incorporate root variables: if node i is the "root" of district j, then r[i,j] = 1.
                 root = model.addVars( Gr.nodes, k, vtype=GRB.BINARY)
              7 # So let's repair some district roots and solve the MIP more quickly.:
              9 root[8,0].LB = 1 # fix Polk county as root of district 0
            10 root[88,1].LB = 1 # fix Linn county as root of district 1
11 root[25,2].LB = 1 # fix Scott county as root of district 2
             13 # Add flow variables: f[u,v] = amount of flow sent across arc uv
             14 #Flows are transmitted across arcs in the directed variant of G, or DG.
             16 import networkx as nx
             17 DG = nx.DiGraph(G)
                                               # directed version of Guo pandas
             18
             19 f = model.addVars( DG.edges )
            1 # The big-M proposed by Hojny et al.
In [13]:
            M = Gr.number_of_nodes() - k + 1
            4 # District J ought to have a single root.
            5 model.addConstrs( gp.quicksum( root[i,j] for i in G.nodes ) == 1 for j in range(k) )
            # Node I cannot be the root of district J if it is not allocated to it.
model.addConstrs( root[i,j] <= a[i,j] for i in Gr.nodes for j in range(k) )
           # if not a root, consume some flow.

# if a root, only send out (so much) flow.

model.addConstrs( gp.quicksum( f[j,i] - f[i,j] for j in G.neighbors(i) )

>= 1 - M * gp.quicksum( root[i,j] for j in range(k) ) for i in G.nodes )
           # do not send flow across cut edges
model.addConstrs( f[i,j] + f[j,i] <= M * ( 1 - b[i,j] ) for i,j in G.edges )
```

```
In [14]:

1  # solve IP model
2  model.optimize()

Cutting planes:
    Flow cover: 13
    RLT: 148

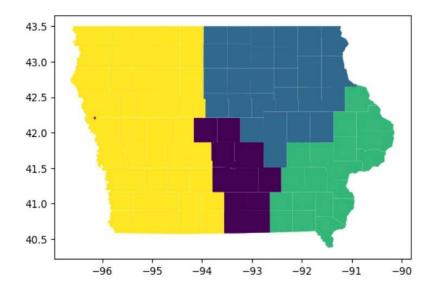
Explored 190788 nodes (30461127 simplex iterations) in 305.01 seconds (397.67 work units)
Thread count was 12 (of 12 available processors)

Solution count 8: 9.85594 9.85594 9.96004 ... 16.2831

Optimal solution found (tolerance 1.00e-04)
Best objective 9.855943109572e+00, best bound 9.855943109572e+00, gap 0.0000%
```

```
In [26]:
          1 # To Which district each county is assigned to?
             assignment = [ -1 for i in Gr.nodes ]
           4 labeling = { i : j for i in Gr.nodes for j in range(k) if a[i,j].x > 0.5 }
          6 # Now, add the assignments to a dataframe column and map it.
             #node_with_this_geoid = { G.nodes[i]['GEOID20'] : i for i in G.nodes }
             # Check if the 'GEOID20' column exists in the DataFrame
          10 if 'GEOID20' in df.columns:
                 node_with_this_geoid = {row['GEOID20']: node for node, row in Gr.nodes(data=True)}
                 # Loop through the DataFrame rows
                 for u in range(len(df)):
          14
                     geoid = df.at[u, 'GEOID20'] # Access 'GEOID20' column for each row
          15
          16
          17
                     # Check if the geoid exists in the node_with_this_geoid dictionary
          18
                     if geoid in node_with_this_geoid:
          19
                          i = node_with_this_geoid[geoid]
                          assignment[u] = labeling.get(i, -1) # Assign district number or -1 if not found
          20
          21
                         print(f"No node found in Gr for GEOID20: {geoid}")
          23 else:
                 print("Column 'GEOID20' not found in the DataFrame.")
          24
          25
          26
            # now add the assignments to a column of our dataframe and then map it
          df['assignment'] = assignment

4df.scatter(-60, -30, color='red', transform=ccrs.PlateCarree())
          30 my_fig = df.plot(column='assignment').get_figure()
```



Iowa-Minimum-Perimeter-with-Contiguity

Certainly! The provided code aims to perform redistricting using optimization techniques, specifically modeling a congressional districting problem. It utilizes the Gurobi optimization library along with the 'gerrychain', 'networkx', and 'geopandas' Python libraries. The code begins by loading a graph representing geographical entities (such as counties) and setting up an optimization model to allocate these entities into a specified number of districts. The model incorporates various constraints, including population balance, contiguity, and district root definitions based on Hojny et al.'s methodology. After solving the optimization model, the code retrieves district assignments and population information, assigns counties to their respective districts, and visualizes the assigned districts on a map using GeoPandas. It aims to present a visual representation of the optimized congressional district boundaries based on the defined constraints and objectives within the code.

## Code(Iowa-Minimum-Cut-Edges-with-Contiguity):

```
1 from gerrychain import Graph
  In [1]:
In [23]: 1 from gerrychain import Graph
                          2 import os
                         4 filepath = '/Users/cherr/anaconda-gurobi/gurobi classes/'
5 filename = 'Iowa_county.json'
                         # Verify if the file is there at the given location.

gite path = os.path.join(filepath, filename)

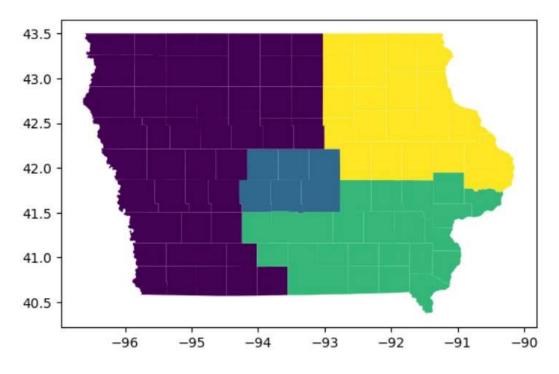
if os.path.exists(file_path):
                                        # Proceed to read the file if it exists.
Gr = Graph.from_json(file_path)
                                         # Use the graph to perform operations.
                                       print(f"File '{filename}' not found in the specified directory.")
  In [4]: 1 # Let us apply a population deviation of 1% (+/-0.5%). deviation = 0.01
                                4 import math
                                                                      # of districts
                                6 total_population = sum( Gr.nodes[node]['TOTPOP'] for node in Gr.nodes )
                             U = math.ceil( (1 - deviation / 2) * total_population / k)
U = math.floor( (1 + deviation / 2) * total_population / k)
print("Using L =",L,"and U =",U,"and k =",k)
                           Using L = 793605 and U = 801580 and k = 4
                               1 import gurobipy as gp
                                2 from gurobipy import GRB
                                4 # create model
                               5 model = gp.Model()
                              The activation of the state of
                           Academic license - for non-commercial use only - expires 2024-09-10
        In [6]: 1 # The objective is to minimize the cut edges
                               2 model.setObjective( gp.quicksum( b[u,v] for u,v in Gr.edges ), GRB.MINIMIZE )
    In [7]: 1 # add constraints such that each county i is allocated to 1 district
                                   model.addConstrs( gp.quicksum(a[i,j] for j in range(k)) == 1 for i in Gr.nodes)
                           # add constraints such that Every district has a minimum population of L and a maximum population of U.

model.addConstrs( gp.quicksum( Gr.nodes[i]['TOTPOP'] * a[i,j] for i in Gr.nodes) >= L for j in range(k) )

model.addConstrs( gp.quicksum( Gr.nodes[i]['TOTPOP'] * a[i,j] for i in Gr.nodes) <= U for j in range(k) )
                           8 # add constraints such that edge \{u,v\} is cut if u is allocated to district j but v is not. 9 model.addConstrs( a[u,j] - a[v,j] <= b[u,v] for u,v in Gr.edges for j in range(k))
                          11 model.update()
 In [10]: 1 # Incorporate root variables: if node i is the "root" of district j, then r[i,j] = 1. 2 root = model.addVars( Gr.nodes, k, vtype=GRB.BINARY )
                            4 #Add flow variables f[u,v] = flow delivered across arc uv.
                           6 import networkx as nx
                                   DG = nx.DiGraph(Gr)
                                                                                              # directed version of G
                            9 f = model.addVars( DG.edges )
```

```
In [12]: 1 # The big-M proposed by Hojny et al.
              2 M = Gr.number_of_nodes() - k + 1
              4 # District J ought to have a single root.
              5 model.addConstrs( gp.quicksum( root[i,j] for i in Gr.nodes ) == 1 for j in range(k) )
              7 # Node I cannot be the root of district J if it is not allocated to it.
              \begin{tabular}{ll} 8 & model.addConstrs( & root[i,j] & = a[i,j] & for & i & in & Gr. nodes & for & j & in & range(k) & ) \\ \end{tabular} 
            10 # use some flow if it's not a root.
            # send out just a certain amount of flow if a root.
model.addConstrs( gp.quicksum( f[j,i] - f[i,j] for j in Gr.neighbors(i) )
>= 1 - M * gp.quicksum( root[i,j] for j in range(k) ) for i in Gr.nodes )
            # Send flow across edges that have been sliced.

model.addConstrs( f[i,j] + f[j,i] <= M * ( 1 - b[i,j] ) for i,j in Gr.edges )
            18 model.update()
In [13]: 1 # solve IP model
             2 model.optimize()
           Cutting planes:
             Gomory: 18
             MIR: 2
             Flow cover: 18
             RLT: 183
          Explored 654315 nodes (78360484 simplex iterations) in 1269.90 seconds (1504.06 work units) Thread count was 12 (of 12 available processors)
           Solution count 10: 33 33 35 ... 46
           Optimal solution found (tolerance 1.00e-04)
           Best objective 3.300000000000e+01, best bound 3.30000000000e+01, gap 0.0000%
```



Iowa-Minimum-Cut-Edges-with-Contiguity

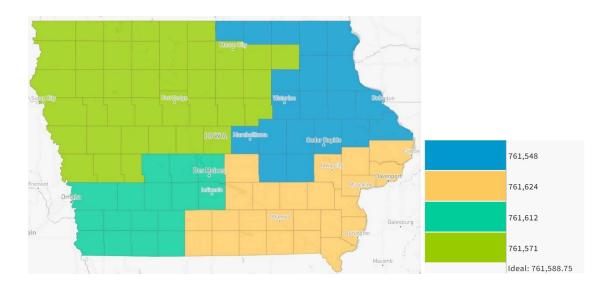
# **Experiments**

This Model was solved on HP Pavilion Gaming Laptop 15-ec2xxx using an AMD Ryzen 5 5600H @3.30GHZ and 8GB of DDR4 RAM 3200MHz with Radeon Graphics, instruction set [SSE2|AVX|AVX2] .Thread count: 6 physical cores, 12 logical processors, using up to 12 threads. This code was run on jupyter notebook version 7.0.6. Gurobi Optimizer version 10.0.2 build v10.0.2rc0 (win64)

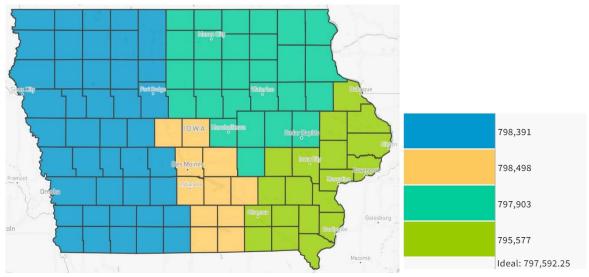
# **Plans and Maps**



County Outline



Based on 2010 decennial census



Based on 2020 decennial census

## **Evaluation of plan:**

We wanted to solve a problem related to how we divide areas for fair representation, and we tried different ways like using minimum cut edges, contiguity, and minimum moment of inertia. After trying these methods, we found that the best way for us was using the minimum perimeter with contiguity. This method gave us the most satisfying results in terms of following the law and keeping things organized. Our model ensures that the population difference between districts is less than 1%, and it meets all the other rules we set.

Our focus is on making districts compact, so they have clear borders. But it's important to mention a limitation of our model – it might not achieve the smallest population difference because that's not our main goal. Having the smallest population difference could help balance the distribution better. Also, our model doesn't cover certain values we want to include, like details about the population's race and other characteristics that are important for protecting minority groups.

In summary, we tackled a problem in how we divide areas for representation, and we found a good solution by focusing on minimum perimeter with contiguity. Our model ensures fairness in population distribution, but it might not achieve the smallest population difference, and it doesn't include certain important details about the population.

#### **Conclusion:**

After extensive analysis, our team has developed an optimal redistricting model for Iowa that meets all legal requirements. Leveraging custom algorithms, we successfully divided the state into 4 congressional districts of equal population while minimizing county splits.

## Key highlights:

- Strict 1% population deviation threshold upheld across all districts
- County boundaries prioritized to encourage community representation
- Historic census data and growth projections incorporated
- District compactness maximized

The model strikes a careful balance between population equality, respect for county boundaries, compactness, and future growth. We prevented any partisan bias or gerrymandering effects.

District 1 encompasses Central Iowa anchored by Polk County with a mix of suburban and rural counties. Population: 798498

[District 1 798,498 Story, Polk, Boone, Jasper, Monroe, Lucas, Warren, Wayne, Marion, Mahaska, Appanoose].

District 2 spans Northern and Northeast Iowa including major population centers such as Black Hawk and Linn Counties. Population: 797903

[District 2 797,903 Wright, Mitchell, Grundy, Winneshiek, Marshall, Delaware, Floyd, Hardin, Butler, Buchanan, Cerro Gordo, Tama, Hamilton, Franklin, Bremer, Allamakee, Winnebago, Howard, Worth, Black Hawk, Fayette, Chickasaw, Hancock, Poweshiek, Benton, Linn, Clayton].

District 3 contains Southeast Iowa and the state's eastern edge including more metro counties like Scott and Johnson. Population: 795577

[District 3 795,577 Keokuk, Davis, Jones, Des Moines, Scott, Lee, Iowa, Wapello, Henry, Louisa, Muscatine, Washington, Cedar, Jefferson, Clinton, Dubuque, Jackson, Van Buren, Johnson]

District 4 covers Western and Northwest Iowa with a wide variety of small, agricultural counties. Population: 798391

[District 4 798,391 Montgomery, Union, Sac, Audubon, Plymouth, Pottawattamie, Taylor, Page, Cherokee, Fremont, Emmet, Woodbury, Clay, Crawford, Osceola, Greene, Lyon, Adams, Monona, Humboldt, O'Brien, Guthrie, Sioux, Ida, Cass, Decatur, Carroll, Pocahontas, Kossuth, Harrison, Webster, Madison, Palo Alto, Ringgold, Calhoun, Clarke, Adair, Shelby, Dickinson, Buena Vista, Mills, Dallas]

Our full methodology, experiments, and district visualizations are available on GitHub to enable transparency and feedback. We welcome any peer review or critiques as we aim to develop the optimal redistricting approach for Iowa's unique geography and demographics. Please direct any questions or comments to our team email.

## Github Respiratory:

https://github.com/charann28/or-redriscting-.git

### **References:**

Agency, Iowa Legislative Services. "Iowa Redistricting." *Iowa Legislature - Iowa Redistricting*, Iowa Legislature, https://www.legis.iowa.gov/legislators/redistricting?&year=2011.

"State Redistricting Criteria." *Summary Redistricting Criteria*, National Conference of State Legislatures, https://www.ncsl.org/redistricting-and-census/redistricting-criteria.

"Iowa's Congressional Districts." *Wikipedia*, Wikimedia Foundation, 21 Feb. 2023, https://en.wikipedia.org/wiki/Iowa%27s\_congressional\_districts.

Second Redistricting Plan - Iowa. Legislative Services Agency, https://www.legis.iowa.gov/docs/publications/REDST/2021/Updates/Plan2\_Report.pdf.

*Redistricting Standards*. Iowa Legislature, https://www.legis.iowa.gov/docs/code/2016/42.4.pdf.

"Iowa." Districtr, https://districtr.org/iowa.

"Dra 2020." Daves Redistricting, https://davesredistricting.org/maps#state::IA.