

Using Machine Learning to Counteract Gerrymandering

Alan Shen, University of New Mexico, Computer Science (B.S. Expected: May 2021), alshen@unm.edu

Mentor: Dr. Soraya Abad-Mota, Department of Computer Science, University of New Mexico

Introduction

The goal of my research is to determine if using machine learning is a viable option for generating fair maps during the process of drawing congressional districts.

Gerrymandering describes the process of drawing districts that skew elections to favor one party over the other.

Methods

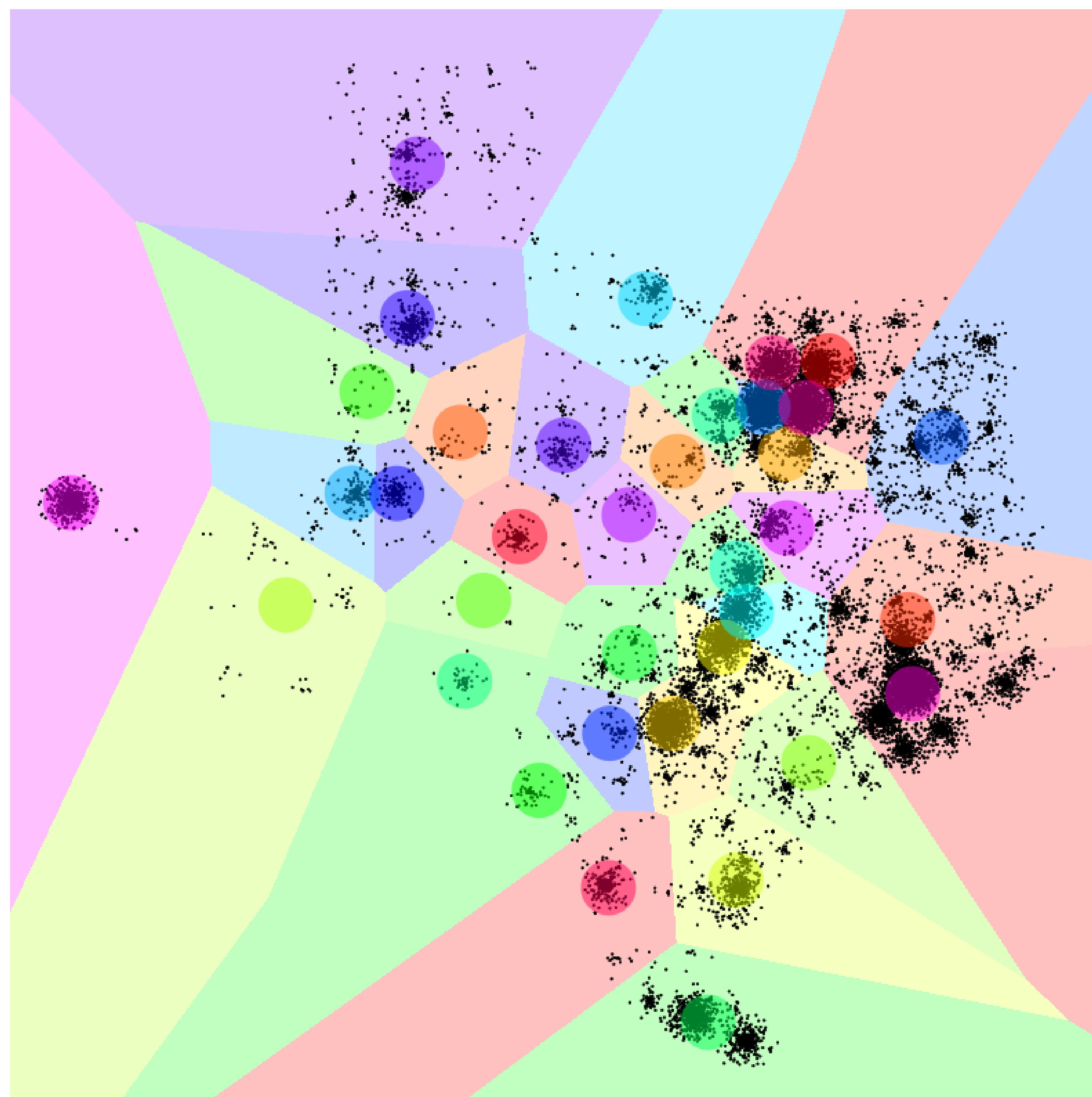
I used the k-means clustering algorithm to generate “fair” maps.

Every centroid in my model represents one congressional district, with the data points representing the population.

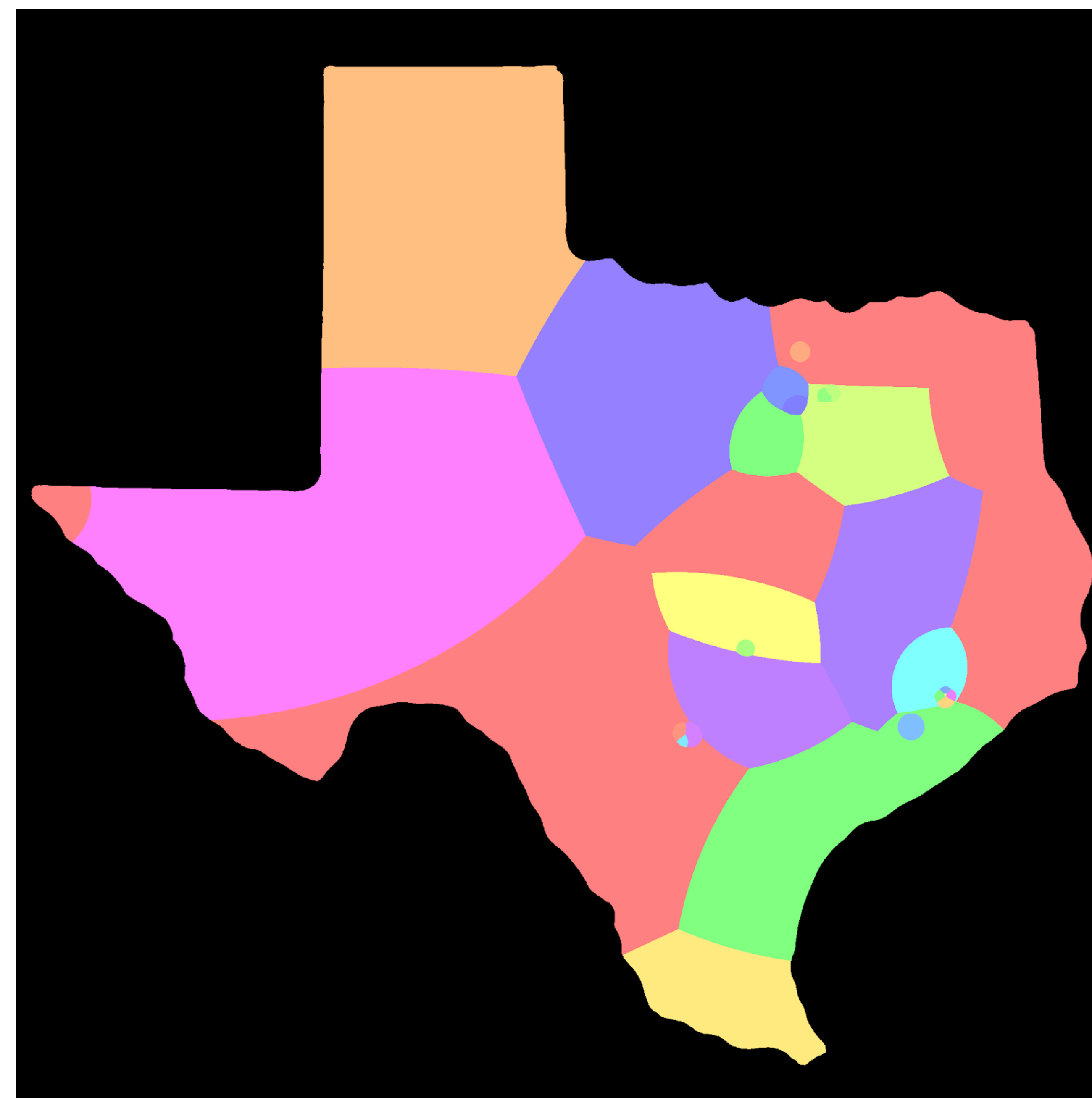
I added weights to the centroids to ensure that all districts had roughly equal populations.

Results

I generated congressional district maps for the state of Texas because of its large population and its currently heavily gerrymandered districts. The generated map (with equal populations of districts) is not too dissimilar from the current map but is much more representative of the population.

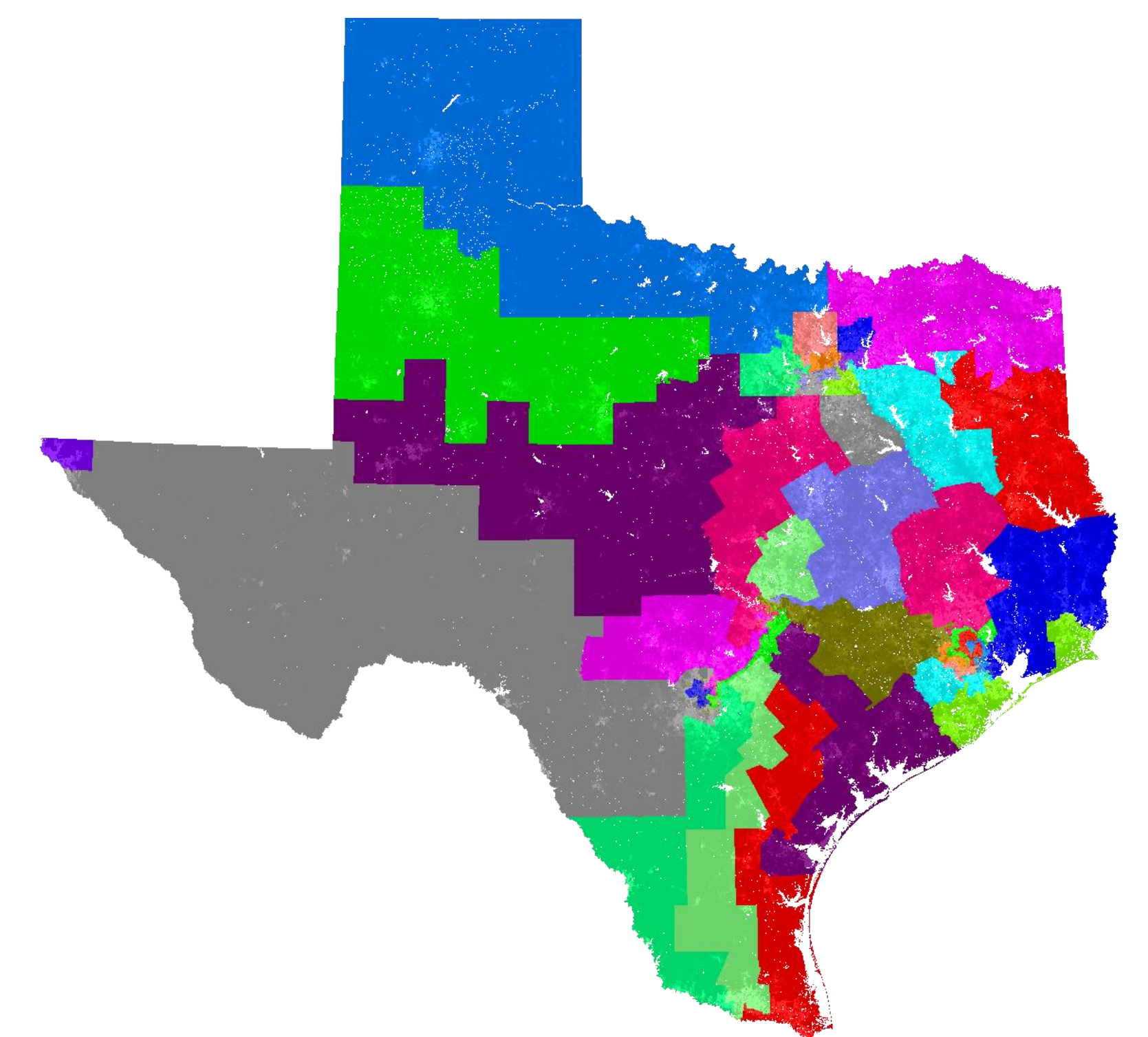


A map generated without population-correcting weights on the centroids. Populations of each district vary too much to satisfy federal district drawing laws.



A map generated with population-correcting weights on the centroids. The centroids all have roughly equal populations.

Current Congressional District Map of Texas



Current congressional district map has far fewer “compact” districts, which is typically a sign of gerrymandering.

Conclusions

Using k-means clustering to generate maps is an effective method.

Weights do need to be added to the centroids to ensure that federal laws are met, most notably the equal-populations rule.

More weights could be added to ensure that various other district drawing rules are met, such as grouping people of common interest together.

Acknowledgements

NSSP/CTSP programs

Upon using 2016 election voting data to simulate elections, the results are 20 republican districts to 16 democratic districts (55.55% republican). This is much more representative of the population, which voted 50.4 % Republican in the 2018 general election. The current ratio is 26 republic districts to 10 democratic districts (72.22% republican).

