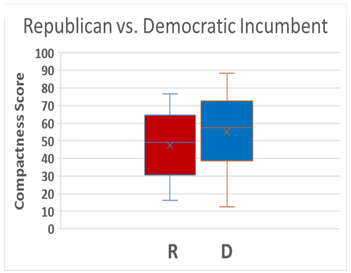
CHARACTERIZING MASSACHUSETTS LEGISLATIVE DISTRICT COMPACTNESS

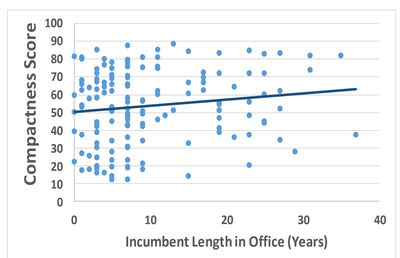
NATHAN FOSTER, ADVANCED GIS, MAY 2018

Introduction

Gerrymandering is the drawing of electoral districts to benefit one group over another, typically based on race, political party, or incumbency. One measure often used to identify gerrymandering is the "compactness" of a district, which refers to how irregularly shaped it is. If this sounds vague, that's because it is. There are a number of simple mathematical techniques for measuring compactness – one of the most popular, for example, is dividing the area of a district by its perimeter – but they all have major flaws, failing to identify what humans would clearly see as unnatural shapes, and incorrectly flagging shapes that humans would see as relatively compact.

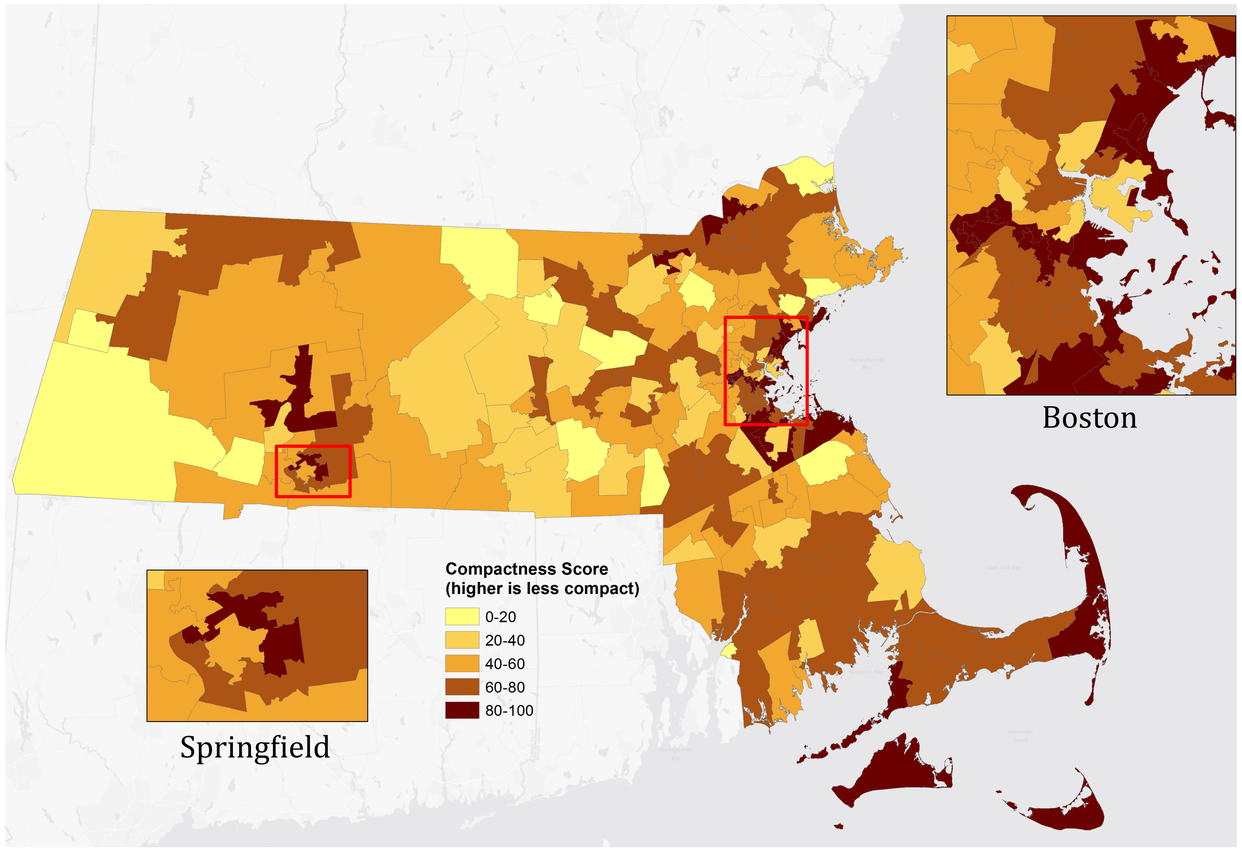
In a recently released working paper titled "How to Measure Legislative District Compactness If You Only Know It When You See It" (Kaufman, King, and Komisarchik), hundreds of individuals, including judges, politicians, redistricting consultants, undergraduate students, and Amazon Mechanical Turk workers, were asked to rank various legislative districts by compactness. The authors then trained a machine learning algorithm on that dataset, producing a measure meant to replicate *how a human would rank* the compactness of any given district. This measure raises some interesting problems – for example, a long thin district oriented horizontally will be judged less compact than the exact same district oriented vertically – but it also has the potential to be much more useful than existing measures. This analysis looks at the paper's compactness score as applied to Massachusetts' 160 State House of Representatives districts.

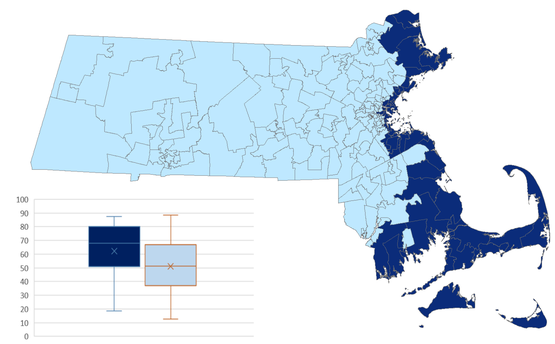




Politics and Compactness

Two political factors were examined: the length in office, in years, of the incumbent of a district, and the political party of the incumbent. There was a positive correlation between the non-compactness of a district and the number of years the incumbent has been in office; however, with a p value of 0.085, this correlation was not quite statistically significant. As for party affiliation, Republicans tend to represent more compact districts than Democrats. The average district with a Republican officeholder had a compactness score of 47, with a standard deviation of 18; the average district with a Democratic officeholder had a compactness score of 55, with a standard deviation of 21. Because Democrats tend to represent districts that are both more urban and more racially diverse, this result is not independent of the demographic analyses that looked at race and population density.



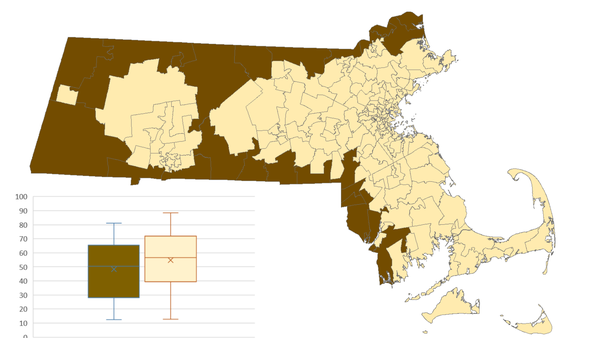


The Coastline Effect

The contours of a coastline can make it difficult or impossible to draw compact districts. The average compactness score of coastal districts was 62, with a standard deviation of 20, while non-coastal districts had an average compactness of 51, also with a standard deviation of 20. While the high standard deviations show that there is significant variation within both groups, coastal districts do tend to be significantly less compact.

Data Sources

Compactness scores from the Gary King Lab at Harvard. Legislative district shapefiles and race, per capita income, and population by district from the National Historical Geographic Information System (NHGIS). Ma State House of Representatives incumbent information (party and length in office) from Ballotopedia and Wikipedia. Basemap from ESRI.



Borders

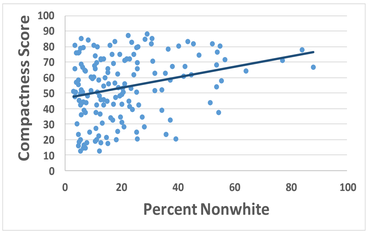
Land borders with neighboring states also constrain district geography. However, perhaps because all of Massachusetts' borders are straight lines, border districts actually tend to be slightly *more* compact than other districts. Border districts have an average compactness score of 48 with a standard deviation of 21, while non-border districts have an average compactness of 55 with a standard deviation of 20.

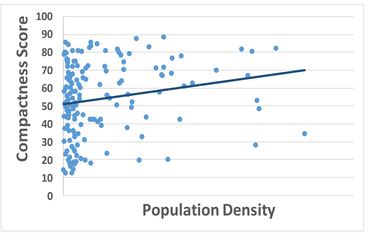
Methods

Maps were made in ArcMap. Correlation analysis was conducted using the scatterplot function in Geoda. Global Moran's I for compactness scores was calculated in Geoda using Queen's move contiguity. All non-geographic data visualizations, as well as the average and standard deviations for the three two-group analyses, were created in Excel.

Demographics and Compactness

Of the demographic factors analyzed, race was the most strongly correlated with compactness. Specifically, districts with a higher percentage of nonwhite residents tended to be significantly less compact than districts with fewer nonwhite residents. These results were highly statistically significant, with a p-value of 0.001. In addition, there was a substantial and statistically significant (p = 0.017) correlation between the population density of a district and non-compactness. Because Massachusetts cities tend to be much more racially diverse than the less dense areas of the state, these two results are not independent of one another. The relationship between the per capita income of a district and its compactness was also analyzed, but was not statistically significant.





Conclusion

Two types of extremely non-compact districts stand out: those that are constrained by geography, and those that are not. A number of coastal districts near Cape Cod and Boston Harbor are extremely non-compact due to geography; more broadly, Massachusetts coastal districts tend to be less compact than interior districts. If the intent of compactness scores is to identify possible gerrymandering, this reflects a limitation of the technique.

On the other hand, many non-compact districts have no visible geographic constraints. These include a number of districts in and immediately north and south of Boston; three districts surrounding Springfield; one large, strangely shaped district just north of Springfield; and several districts near the New Hampshire border. Demographic analysis sheds some light on the characteristics of these districts: they tend to be more racially diverse and more urban (and more often represented by a Democrat) than compact districts. Unfortunately, these factors all correlate with each other as well, making it difficult to determine causality.

The lack of a statistically significant relationship between length of incumbency and the compactness of a district does not mean that political factors are irrelevant. For future analyses with better access to data, it may be valuable to look at the partisan composition of a district in the year of redistricting, the length in office of the incumbent as of the year the district was created, rather than seven years later, and whether the incumbent held a leadership position when the district was created.

These results should be interpreted with caution. The compactness scores being used are still experimental, non-compactness is not always evidence of gerrymandering, and it is perfectly possible to gerrymander while keeping individual districts relatively compact. Nonetheless, when a district is extremely non-compact in the absence of geographic constraints, it warrants further examination.