CS504 Data Mechanics
Final Project Report

”Determining how Progressive or Conservative a specific locality, in Massachusetts, is using answers from Ballot Questions”

Efstathios Karatsiolis | Chengjun Wu | Nikhilesh Murugavel | Nathan Mokhtarzadeh

May 3rd, 2019

Introduction

Mass Alliance is a coalition for political advocacy organization that work together to build a politically aware Massachusetts. As part of our collaboration we were tasked to gather data from Ballot Questions and by using specific metrics provided by Mass Alliance output meaningful results that showcase how progressive or conservative a specific locality is.

Ballot questions are a very important aspect of the elections procedure. However, these questions can infer more information rather than a simple Yes or No. By combining years of results and data we managed to get insight on the voting trends in Massachusetts, by district and precinct.

Goal

The major goals of the project was:

1. To allow Mass Alliance to collect and organize Massachusetts districts' voter data based on ballot questions over 2002 - 2018.
2. Assign scores to the districts on a progressive-conservative scale based on the data.
3. Statistical inferences on the aggregated data to identify a district as progressive or conservative.
4. Present creative visualizations to display our findings on Massachusetts voter districts.

Datasets Used

- Ballot questions from 2002-2018 from electionstats.state.ma.us
- A searchable database of Massachusetts historical election information (PD43) from electionstats.state.ma.us

Algorithms and Analysis

We have 24 ballot questions and can basically identify the conservative vs progressive attitude of respondents through their answer. Hence we was trying to use constraint satisfaction to see the conservative vs progressive attitude of regions across Massachusetts according to the statistics of response to the questions we have.

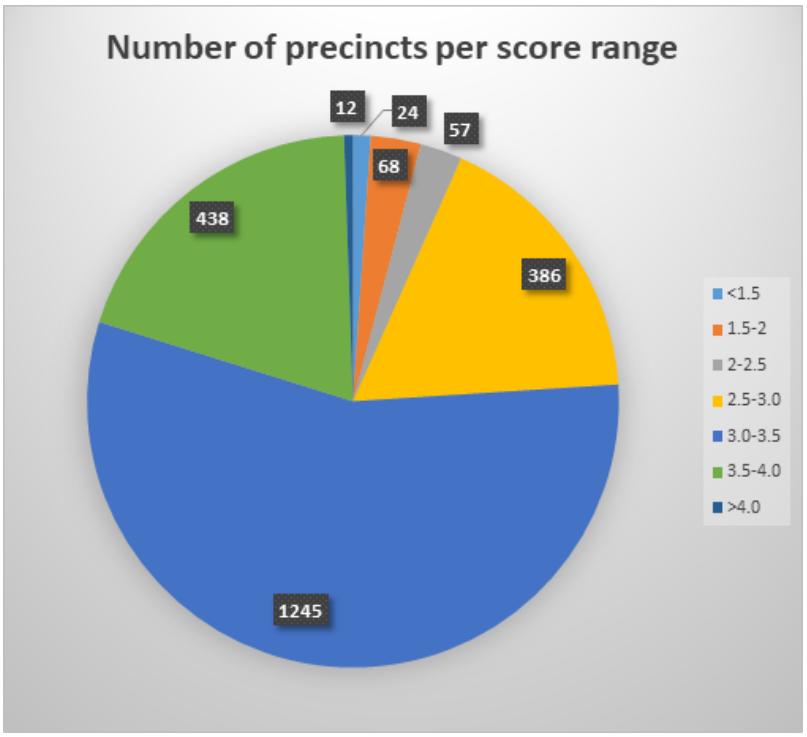
Generally, Mass Alliance only take advantage of the data by focusing the majority of response of a region that is only by looking at the dominants of respondents with “Yes” and “No”. However we think a better way to invest deep into it is to compare the ratio of those responses. At some point, respondents answered with “Yes”, considered as the conservative, may still have less impact than the minority of progressive people because they are not likely to take actions to show their political pursue when compare to progressive. Therefore, we come up with three constraint problems to show how many localities are likely to be dominated by the conservative.

1. Find out localities that have more people respond answers with “Yes” than “No”
2. Find out localities that have more people respond answers with “Yes” than 1.5 times of “No”
3. Find out localities that have more people respond answers with “Yes” than 1.5 times of “No” and also the overall respondents is more than people that give up the vote.

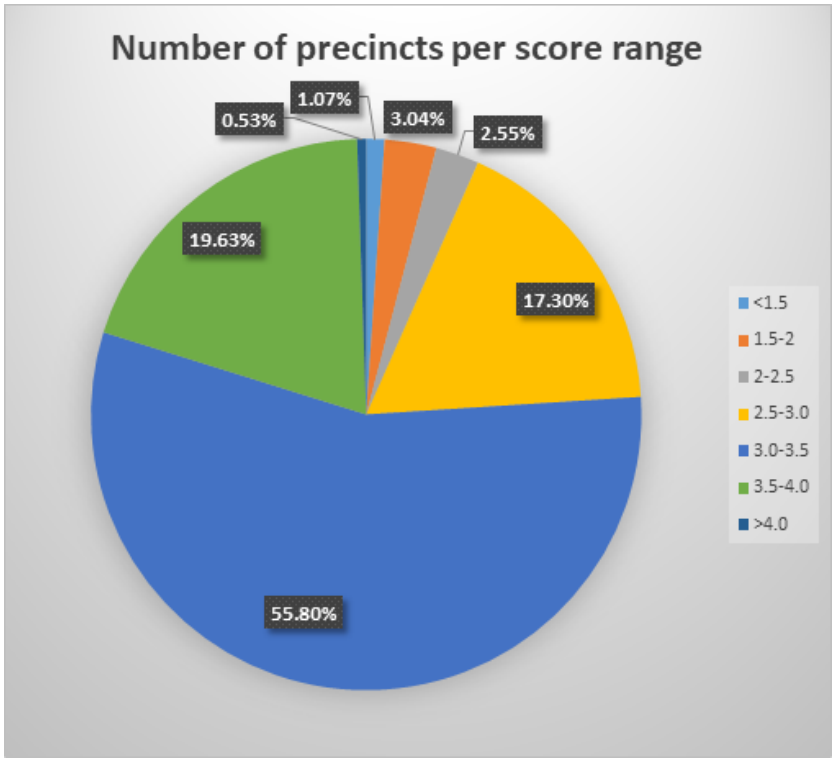
From the result, we found only just a few localities, 24 over 386 to be specific, satisfy the third constraint that means less than 0.1 of overall localities across Massachusetts are likely to show distinct conservative attitude as a whole.

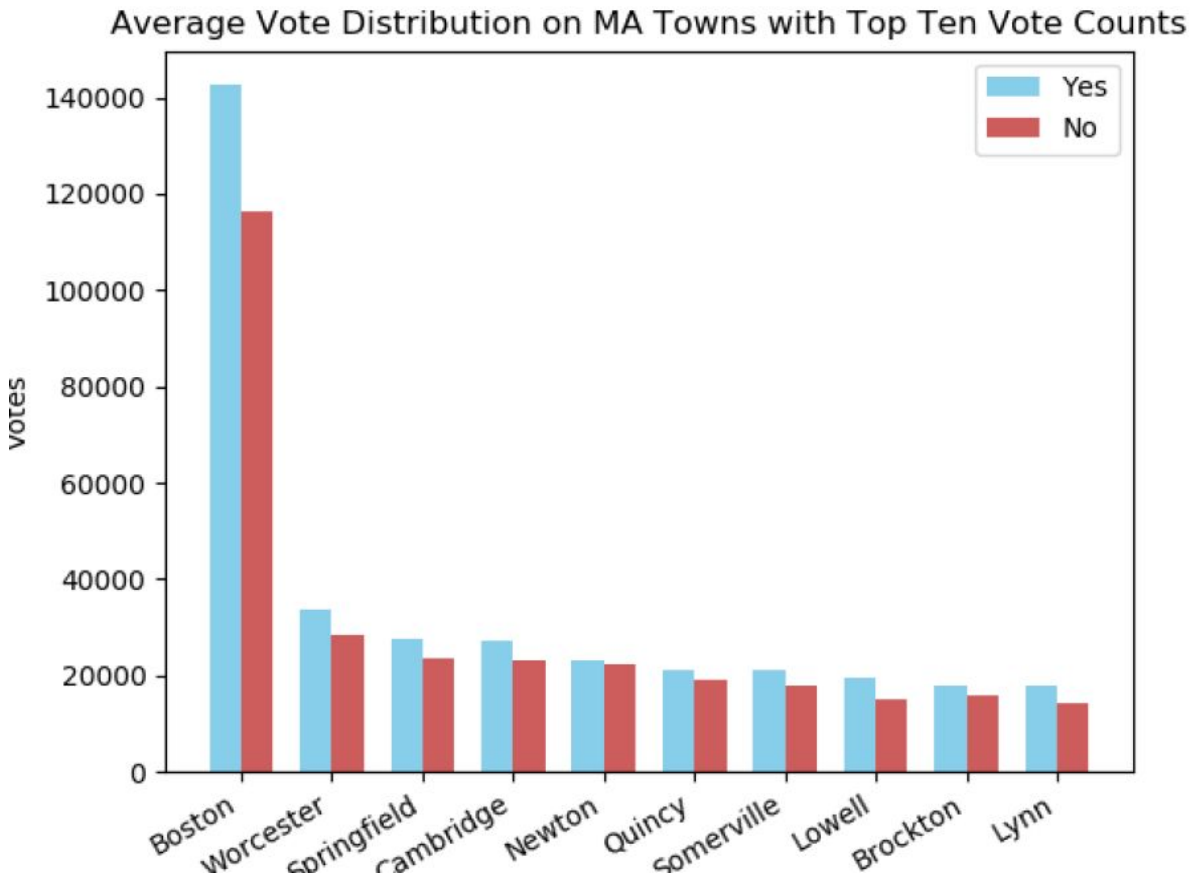
Results

Results for precincts



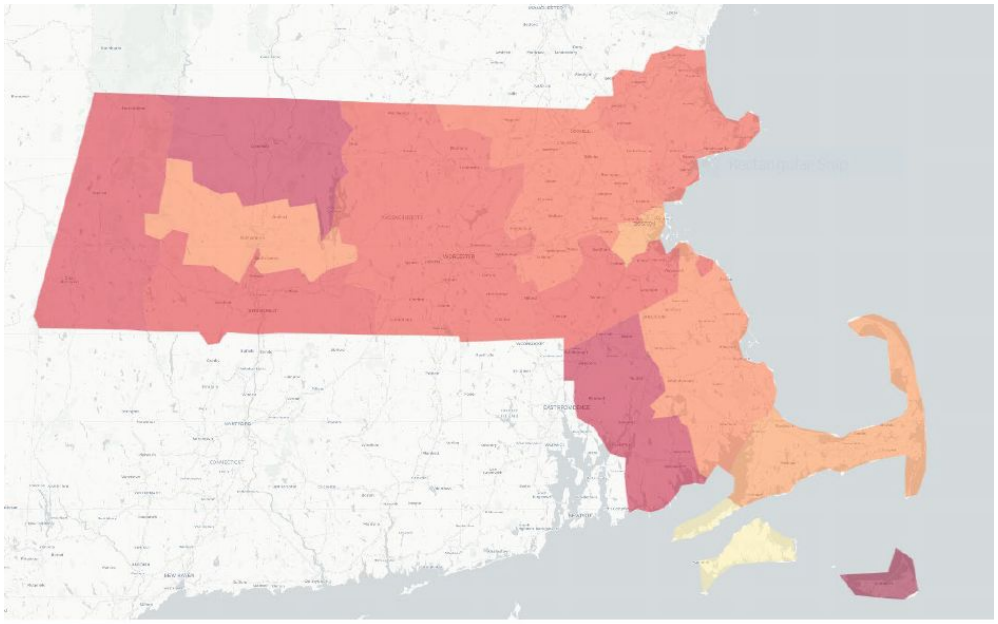
Results show a generally neutral stance from most of the precincts throughout the years.
About 75% of the precincts fall between a score of 2.50-3.0.
However in total 75% of the precincts show a relatively more progressive stance in contrast with 25% of them that have a more conservative one.





Voter data on yes vs. no questions on propositions. Data was aggregated based on different towns. Diagram does not take blanks into account.
This shows the tendencies of the cities with the large population in the state.

Heatmap data for counties in Massachusetts



Geographic heat map. The progressive scale aggregated by county. As color goes deep, the progressive scale increase.

Limitations

The dataset we were given was limited, we had to scrape data from online records and sometimes develop our own datasets from online information to operate on.
Another limitation is the scoring metric on what towns are considered to be “conservative” vs. “progressive”, since we’re scoring on a metric of 1 to 5 inclusive so the difference between what is considered to be conservative and progressive could be swayed by a small number of votes as most of the towns seem to be in the middle and easily influenced.

Conclusion

We provide patterns on the score ranges by mapping the different districts voter data. We also calculate the progressive and average scores for each precinct in Massachusetts based on the metrics used by Mass Alliance. An interesting pattern we found in our computations was that the scores average to a neutral stance meaning the voters are not always consistent in their choices. The scores averaged over the ballot questions through the years show that there are no extreme peaks in the scores,

implying a constant change in the mindset of the voters. Any choice the voters made over the years depends on the current circumstances and the content of the ballot questions.

Future Work

Our findings can be used to compare the results of this election to the results of the next with the same code (and some small modifications), the same charts and visualizations can be generated as well. In such a divisive election there will most likely be some very interesting conclusions to draw upon comparison between the primary and general results.