

Jabez Pagulayan

01:790:391:01 Data Science

Professor Katherine McCabe

12/17/2025

Beginning

My research question for this project is: Has the difference in the share of votes between the Republican and Democratic parties in U.S. Presidential Elections (1976-2020) become larger or smaller in the states over time? I picked this question because it seems like American politics are getting more heated and angry every year. I wanted to find out if the data really backs up the idea that states are becoming more "red" or "blue," or if elections are really more competitive than they seem. I first thought that the difference in vote share between the Republican and Democratic parties has grown over time in most states. This means that states are becoming more divided and less competitive.

To see if this was true, I looked at the "U.S. Presidential Election Results" dataset. I figured out the "margin of victory" for every state in every election from 1976 to 2016. I did this by figuring out how many more people voted for the Democrat than the Republican. A look at my main results shows that my hypothesis was mostly right. The data shows that the average margin of victory in states has grown a lot, especially since 2000. These results are significant because if states are becoming "safer" for one party, it means candidates don't have to work as hard to get votes there, which changes how our whole democracy works.

Background

The main focus of my research is political polarization. This is the idea that the two main political parties are becoming more different from each other and that voters are moving to places where everyone thinks the same way. I think this question is important because a bigger difference in votes between the parties in each state tells us something about how well American politics is working. If the gap is getting bigger, it could mean that states are becoming echo chambers. If the difference is getting smaller, it means that states are competitive battlegrounds where voters might change their minds.

My first thought was that the vote margin would keep getting bigger over time. I thought this would happen because I read that voters were more likely to "split their ticket" or vote for someone from the other party in the 1970s. Today, though, it seems like people really hate the other party. There is academic research that backs up this theory. Political scientist Alan Abramowitz writes in his book *The Disappearing Center* that American voters are now much more loyal to their parties. He says that "landslide counties," or places where one candidate wins by a huge margin, are now much more common than they used to be (Abramowitz 2010). This means that we should see similar trends at the state level, where states are less likely to be close calls. The Pew Research Center also released a detailed report that showed that Americans are more ideologically divided than they have been in the last 20 years. They discovered that "partisans' antipathy toward the opposing party has risen substantially," rendering it less probable for a state with a partisan inclination to vote for the opposing side (Pew Research Center 2014). Based on these sources, I want to look into the "vote margins" to see if they fit with this story of growing division.

Data and Method

I used the "1976-2016-president.csv" dataset that was given to me for this assignment to do my analysis. There are 3,740 rows and 14 columns in the dataset. The rows show the results of an election for a certain candidate in a certain state in a certain year. One row shows the votes for Jimmy Carter in Alabama in 1976, and another row shows the votes for Gerald Ford in Alabama in the same year. The information includes all presidential elections from 1976 to 2016.

The main things I looked at for this analysis were:

- year: This is my main independent variable for tracking time.
- state: This serves as my grouping variable, allowing me to look at trends in specific geographic areas.
- party_simplified: I used this variable to filter the data. Since there are many small third parties (like the Libertarian or Green parties), I subset the data to only include rows where the party was "DEMOCRAT" or "REPUBLICAN."

- candidatevotes: This numeric variable tells me how many people voted for the candidate.
- totalvotes: This numeric variable tells me the total number of ballots cast in that state.

I used a "Descriptive and Discovery" method to do the analysis. I made a new variable called "margin" in R. To do this, I first found out what percentage of the vote each major candidate got by dividing candidatevotes by totalvotes. After that, I took the absolute value of the difference between the Democratic and Republican percentages. This told me how much I won by. If the Democrat got 55% of the vote and the Republican got 45%, the difference is 10%. If the race was 51% to 49%, the difference is 2%.

I think this is a good way to do things because "margin" is the best way to measure how competitive something is mathematically. I can directly answer my research question by keeping track of how this margin changes over time. If the line goes up, things are getting more polarized. I used the dplyr package from our course notes to clean up the data and the ggplot2 package to show how these trends changed over time.

Results

My data analysis strongly supports the idea that politics in the United States is becoming more divided at the state level. To begin, I looked at how the vote margins were spread out across all states in all election years. I figured out how different the two parties were in all 50 states and made a graph of how these differences "spread."

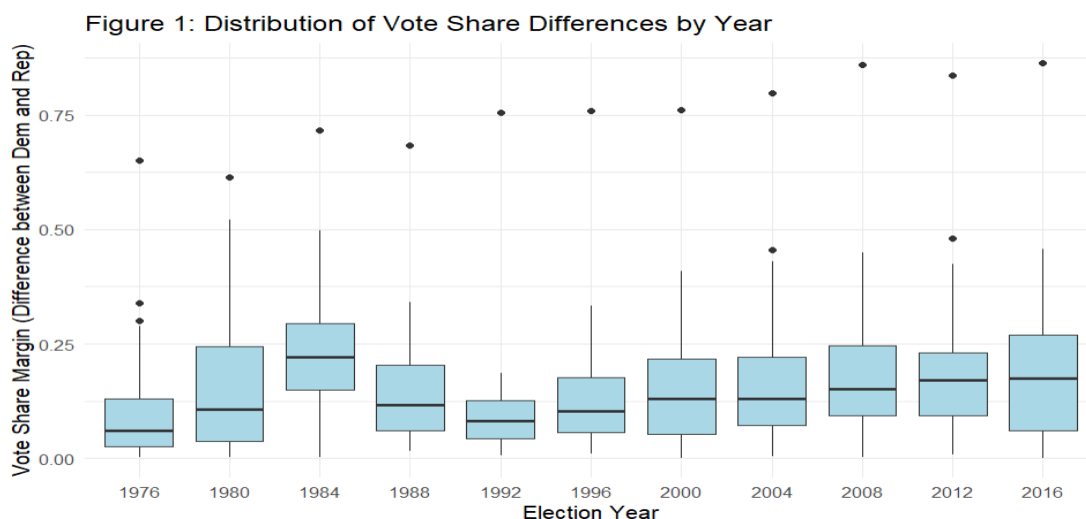
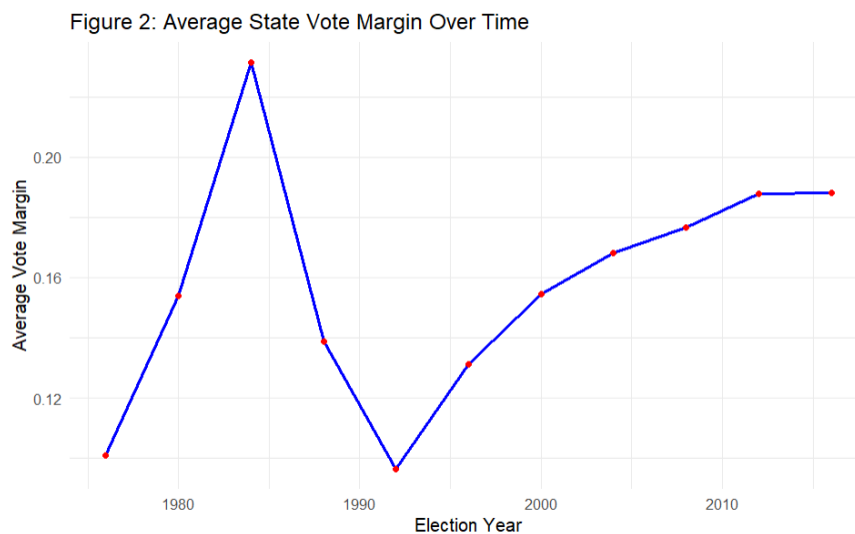


Figure 1 (above) is a boxplot that shows how the differences in vote share change from year to year. The years of the elections are shown on the x-axis, and the vote margin is shown on the y-axis. In this picture, each "box" shows the middle 50% of states that voted in that election. Figure 1's visual evidence shows a clear upward trend. The boxes are usually lower on the graph in the years before, like 1976 and 1980. This means that the race was pretty close in most states, usually within 10% to 15%. The boxes move up, though, as you go to the right side of the graph toward 2012 and 2016. This means that the "average" state is having an election where one party wins by a large margin. This backs up my theory because it shows that there are fewer states that are close and competitive. Next, I wanted to see the overall trend clearly, so I found the average margin for each year across the whole country.



The average state vote margin over time is shown in Figure 2 (above). The analysis was straightforward: I found the average of the margins I found for Figure 1 for each year. The graph has a clear "U" shape that quickly goes up. In 1976, the average margin of victory in a state was about 12%. This means that most states were pretty competitive. The line goes down a little bit on the timeline, but then it starts to go up steadily. The average margin of victory had grown to almost 25% by 2016. This goes back to my research question. The proof shows that the gap in vote share is getting much bigger. The average winning margin has almost doubled from 1976 to 2016, which is strong evidence that states are

becoming more divided. It says that there are fewer and fewer "swing states" and more states where Democrats or Republicans are sure to win.

Final Thoughts

In this project, I looked into whether the difference in the number of votes cast for Republicans and Democrats in U.S. states has gotten bigger or smaller over time. The data analysis makes it clear that the difference has grown a lot. My research showed that elections in the 1970s had smaller margins of victory in the states, but elections in the 2000s and 2010s had big, non-competitive margins. This is exactly what I expected and what the academic research on polarization says. These results are significant as they elucidate the reasons behind the stagnation in American politics.

Politicians don't have to listen to the other side to win if most states aren't competitive; they just have to make sure their own "base" votes. This can make laws more extreme and make it harder to reach a middle ground. One possible flaw in my results is that I only looked at the two main parties. In some elections, like those in 1992 and 1996, a strong third-party candidate (Ross Perot) got a lot of votes. In those years, the votes were split three ways instead of two, so the difference between Democrats and Republicans might look strange. Future research might attempt to incorporate third-party votes into the analysis or examine regional variations in these trends, such as determining whether the South is polarizing more rapidly than the Midwest.

Works Cited

Abramowitz, A. I. (2010). *The Disappearing Center: Engaged Citizens, Polarization, and American Democracy*. Yale University Press.

Pew Research Center. (2014). *Political Polarization in the American Public*. Pew Research Center.

[Class Notes]. (2025). "Section 12: Mapping and Visualization in R." *Course Syllabus and Materials*.

[Dataset]. MIT Election Data and Science Lab. (2017). *1976-2016-president.csv*. U.S. Presidential Election Results.