

Kaylan Wallace

POLI 381

Data Project: Description

Are elections becoming more competitive or less competitive?

Introduction

For the first part of the project, I will conceptualize and measure electoral competitiveness. To do this, I will use polling data sourced from 538, a website for opinion poll analysis, politics, economics, and sports in the US. This data set logs the daily polling averages for every Presidential candidate between 1968 and 2016. It shows each candidate's daily polling average in every state leading up to the election date.

In this project, electoral competitiveness will be conceptualized as a comparison of Presidential candidates' polling averages. A competitive election can therefore be conceived as one where the average rating of the top two Presidential candidates is relatively close, given that there are at least two candidates.

Here are examples of competitive and non-competitive elections under these parameters:

- Example 1, Competitive election: Candidate A has a rating of 32%, candidate B has a rating of 48%, and the remaining 20% is split among several other candidates.
- Example 2, Competitive election: Three candidates have an approximately equal rating.
- Example 3, Non-Competitive election: Candidate A has a rating of 90%, candidate B has 10%.
- Example 4, Non-Competitive election: There is only one candidate.

Measurement

Electoral competitiveness will be measured using the 538 data set. For each election:

1. Only data from April 9th onward will be considered, so that each election's data covers an approximately equivalent time frame. This is to ensure a fair comparison.
2. The states will be put into 9 groups, based on the Census Regions and Divisions of the United States (Bureau). This is to simplify analysis.
3. Each candidate's polling average from April 9th to election day per division will be calculated.
4. Each candidate's national polling average will be computed by taking a weighted average based on the number of electoral votes each division gets (National Archives).

5. A comparison of the standard deviation of candidates' national polling average will provide information about the election's competitiveness.

Scope and Coverage of the Measure

This measure limits the sample to US Presidential elections. While it may provide a meaningful comparison of the competitiveness of this type of election, it excludes elections in that use different democratic structures and processes outside of the US. It also excludes all elections within the US that are not Presidential. As such, the findings from using this measure may not be generalizable to other countries or electoral systems. Nonetheless, it still ought to provide insights into the competitiveness of the elections that are considered.

Statistical Analyses

After performing the above tidying and calculations, I was able to calculate the standard deviation of candidate's national polling averages for each presidential election. **Table 1** shows these results.

Election Year	Mean of Weighted Ratings	Standard Deviation of Weighted Ratings
1972	40.08	18.75
1976	34.14	1.79
1980	29.38	14.38
1984	44.99	12.69
1988	43.98	0.16
1992	31.65	7.31
1996	30.42	19.37
2000	29.65	22.4
2004	45.34	3.05
2008	45.31	0.27
2012	45.62	0.08
2016	30.52	19.99

Table 1: The number of candidates and the standard deviation of their polling averages per election

For 1972, while the high standard deviation in **Table 1** suggests a close race, looking at individual candidates' ratings (see Appendix A) shows Nixon with a substantially higher score (53.34) compared to McGovern (26.83). This indicates a dominant victory for Nixon despite some public disagreement. In this case, we can conclude the election was not very competitive.

2012 had a standard deviation of 0.08 between two candidates, the smallest standard deviation value in the entire table. We can conclude that the polling averages for these two candidates were extremely close in value. Looking at the actual ratings (see Appendix A), Barack Obama had 45.67% of popular support across America while Mitt Romney was rated 45.56%. Indeed, their ratings are very close. This election can be considered very competitive.

From **Table 1**, we also see that 2016 was the year with the largest standard deviation amongst candidate ratings. Appendix A shows a miniscule difference between Clinton's rating (42.61%) and Trump's (41.5%). The large standard deviation is due to a third candidate, Gary

Johnson, who only received a 7.45% rating. The results in **Table 1** may lead us to believe this was not a competitive election as there is a large deviation amongst candidates' ratings. The actual data reveals a different scenario though, where two candidates have relatively close scores and the third candidate having a much lower rating increases the standard deviation. This election still be considered competitive under the parameters set out in our conceptualization.

These analyses show the importance of considering the standard deviations in tandem with the actual data, as the figures in **Table 1** can lead to incorrect conclusions if considered in isolation.

Table 1 also indicates a general downward trend in average weighted ratings between 1972 and 2016. This could suggest a possible decline in public approval for presidential candidates across both parties. Spikes in years like 2004-2012 may be indicative of elections where the populace was more active in political matters. **Figure 1** shows us how the average weighted ratings for all candidates has fluctuate significantly between 1972 and 2016.

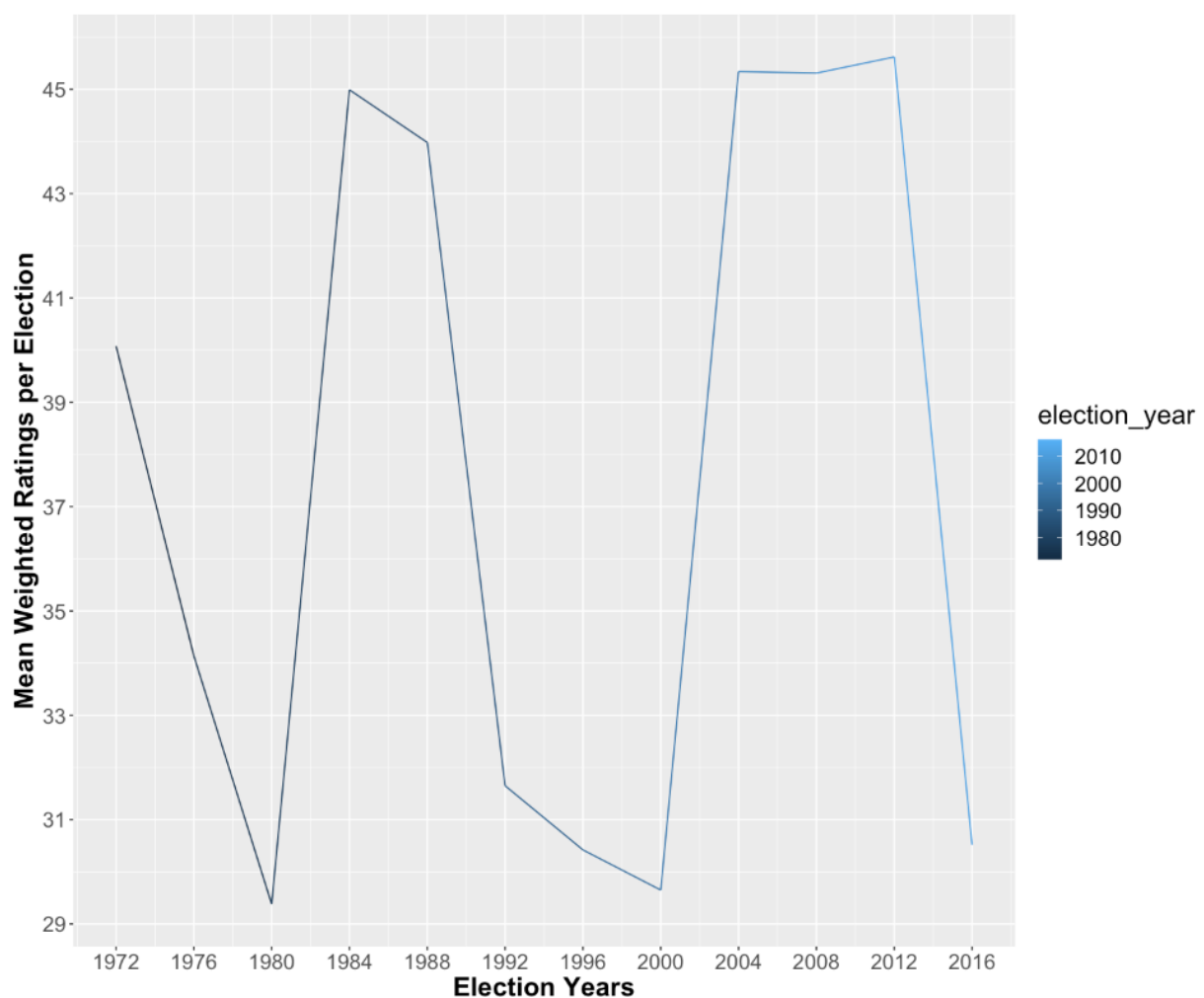


Figure 1: Graph showing trends in average weighted ratings for all candidates per election

References

- Bureau, U. C. (2021, October 8). *Geographic Levels*. Census.gov.
<https://www.census.gov/programs-surveys/economic-census/guidance-geographies/levels.html>.
- National Archives and Records Administration. (n.d.). *Distribution of electoral votes*. National Archives and Records Administration. <https://www.archives.gov/electoral-college/allocation>.

Appendix A

Table showing candidate's weighted polling score for every US Presidential election.

Candidate Name	Election Year	Weighted Polling Score	Number of Divisions
George S. McGovern	1972	26.83	8
Richard M. Nixon	1972	53.34	8
Gerald R. Ford	1976	32.88	7
Jimmy Carter	1976	35.41	7
Jimmy Carter	1980	35.7	9
John B. Anderson	1980	12.93	9
Ronald Reagan	1980	39.52	9
Ronald Reagan	1984	53.96	9
Walter F. Mondale	1984	36.02	9
George Bush	1988	44.09	9
Michael S. Dukakis	1988	43.87	9
Bill Clinton	1992	34.15	9
George Bush	1992	37.38	9
H. Ross Perot	1992	23.42	9
Bill Clinton	1996	45.56	9
Bob Dole	1996	37.1	9
H. Ross Perot	1996	8.59	9
Al Gore	2000	39.79	9
George W. Bush	2000	45.18	9
Ralph Nader	2000	3.97	9
George W. Bush	2004	47.5	9
John Kerry	2004	43.18	9
Barack Obama	2008	45.12	9
John McCain	2008	45.5	9
Barack Obama	2012	45.67	9
Mitt Romney	2012	45.56	9
Donald Trump	2016	41.5	9
Gary Johnson	2016	7.45	9
Hillary Rodham Clinton	2016	42.61	9

This table was the result of the R code that tidied and performed calculations on the original dataset.