

Predicting the Midterm Election

Problem statement

Midterm elections, like the one in November, often swing the power of the political parties and can provide a marker for how a president is performing. Historically, the president's party loses seats: on average 27 seats in the House of Representatives ($t = 5.09, p < 0.0001$) and 3.8 seats in the Senate ($t = 3.21, p = 0.0044$) ([data](#)). This significant change could be due voter fatigue, a reaction by the public to the President's actions, or simply a form of 'regression to the mean'. Essentially, the political direction of the country often flips or flops based on the result of the midterms.

The goal of this project is to predict the outcomes of the midterm elections: predicting the results of all federal House of Representative elections in November (Senate seats are a bonus). Any piece of information can be used to perform this prediction, but the suggested feature set should include past elections, polling data, and demographic information of congressional districts (see below).

For an example prediction model for 2018, see [fivethirtyeight](#).

Data Resources

1. Election Results

There are a few sources of recent election results: [2016 Election on Wiki](#), the [FEC's 2016 Report](#), and [Archived Election Results since 1982](#).

2. Polling Data

Recent (and a little historical) polling data can be found at [Real Clear Politics](#).

3. Congressional District Data

Demographic data on Congressional districts can be found on [Census.gov](#), and geospatial data are found on [Data.gov](#).

High-level project goals

1. Obtain publically available data from various public sources (some scraping will definitely be required).
2. Build a predictive model for the midterm election in November (not using the actual 2018 election results).
3. Use the 2018 results to determine why and when the predictive model was right or wrong.
4. Determine which of the polls got it right or wrong, and why.

References

1. Campbell, J., et al. (2015). *Evaluations of the 2014 Midterm Election Forecasts*. PS: Political Science & Politics, 48(2), 295-300.
2. Cohen, Jeffrey E., et al. (1991). *The Impact of Presidential Campaigning on Midterm U.S. Senate Elections*. The American Political Science Review, 85(1), 165-178.
3. Erikson, R.S. (1988). *The Puzzle of Midterm Loss*. The Journal of Politics, 50(4), 1011-1029.
4. [What went wrong with the 2016 election](#)