Capstone Project Proposal - 2

Congressional Survival

Problem

Based on their voting history, what type of congress member is most likely to survive in what political environment? Election results are highly dependent on voter enthusiasm, partisan lean, and the national political environment. However candidates in swing districts must do a cost benefit analysis of meeting the partisan preferences of their district v. reducing the enthusiasm of their base.

Client

The theoretical client in this instance would be the DNC, who could distribute findings to individual members of congress who would use them in addition to polling to inform their choices and positions.

Approach

Unsupervised learning could be used to cluster congressional members based on their voting history, and partisain district lean. The similarity of these groups survival of these groups would then be checked via the bill descriptions and election results, It should also be possible to compare which votes were on average associated with the greatest electoral advantage over district partisan lean. Additionally, an informational interview with a DNC data scientist has indicated demonstrating competence with API calls should be a priority.

Data

FiveThirtyEight:

Which congress people voted on what bills, in 2017 and 2018, and Position Trump held on these bills The partisan lean of each district:

Pro-pubica:

Historic comprehensive voting records available via API