Phase I: Establish a system to analyze all available public-news related documents for the: 1) 2016 presidential election, 2) 2018 mid-term election & 3) 2020 presidential election— Republican/Democrat, for/against, positive/negative, by US State.

Search for identifiable features/patterns (i.e. phrases, descriptors, words, policies, etc.) that are: 1) tied to *contemporary polling information by state (may be difficult to get)* & 2) used to support the successful campaign of Donald Trump or the unsuccessful campaign of Hilary Clinton in 2016 or Senate/House seats in 2018.

Phase II: Apply the Phase I learnings to the 2020 campaign to-date to determine if patterns are consistent for one or both sides. Run analysis & capture results of a logistic regression that attempts to predict the 2020 presidential election. Training data to be taken as 2016 & 2018 information, test data to be taken as 2020 information. Features used in the logistic regression to be the output of Phase I.