

## Media in the midterms: What did the media cover in 2018, and what can we expect in 2022?

2022 is a crucial mid-term election year in the United States. Since around 2016, it has been observed time and again in various elections around the world that the news cycle and media consumption strongly influence election results. In this scenario, campaign strategists are interested in knowing: - What are the important issues in the media around election time? - When is a good time to start playing up these issues? - Which news outlets could give maximum exposure to your ideas?

In this project, we looked at a news dataset that scraped news articles by American sources in the previous midterm year, 2018. Then, we utilized standard techniques in Natural Language Processing to answer some of these questions. We were able to provide quantifiers that signify how long an important topic stayed in the spotlight, and which news publications are more susceptible to dominant topics.

After classifying unlabeled articles and using Nonnegative Matrix Factorization to identify the most relevant topics in the politics news in 2018, we observed that most topics are tied to specific events that occurred throughout 2018. We defined a quantifier called *topic lifecycle* which measures the popularity of a particular issue relative to its annual peak value on a month-to-month basis. Using this metric, we could see that topics like “Bernie Sanders” and “Polls” start getting traction in the news as early as June (given that the elections are held in November). It is noteworthy that people begin publishing polls around the same time. This momentum is carried through to election time. Thus, we infer that it is important to start campaigning and guiding the conversation towards favorable polls around June.

Finally, we also defined a *myopicity score* for each publication in the dataset which corresponds to the fraction of articles dedicated to the single most dominant story at any given time. Outlets like People magazine, and TMZ were found to be “myopic” as more than 75% of the politics news stories in them were focused on a single story. These myopic outlets might be ideal to start promoting/attacking candidates and guiding the conversation around polls.

Given our headstart in analyzing the news cycle in 2018, and our simple metrics to analyze news articles, one can gain key insight into the effect it has on elections. It would be interesting to extend our analysis with news datasets from earlier election years in order to see recurring patterns.