From Pre-Election to Post-Election: Measuring Sentiment Shifts in New York Times Headlines

Dana Popky and Ted Van Green

2025-12-10

Contents

Introduction

Our github is located here (https://github.com/danapopky/Popky_VanGreen)

Media outlets play an important role in U.S. elections. Decisions about which stories to cover and how to report on the candidates can influence voter perception and behavior. But in the 2024 election, media outlets came under fire from critics alleging that their reporting was biased, inaccurate, or favored one political side. With most Americans tuning into election news, a partisan tilt could have an outsize impact on turnout, vote choice, and even trust in democracy. In this paper, we ask to what extent the data supports allegations of media bias by analyzing the sentiment and emotions of news headlines before and after the 2024 election.

In this analysis, we focus on the New York Times. In the lead up to the 2024 election, about four in 10 Americans say they used the New York Times as a source of political and election news. But, Democrats were much more likely than Republicans to say they use and trust the NYT coverage. To evaluate a potential bias in their reporting, we scrape articles from the 30 days before and after the election and identify how the sentiment and emotions of NYT headlines shifted after Donald Trump's election victory.
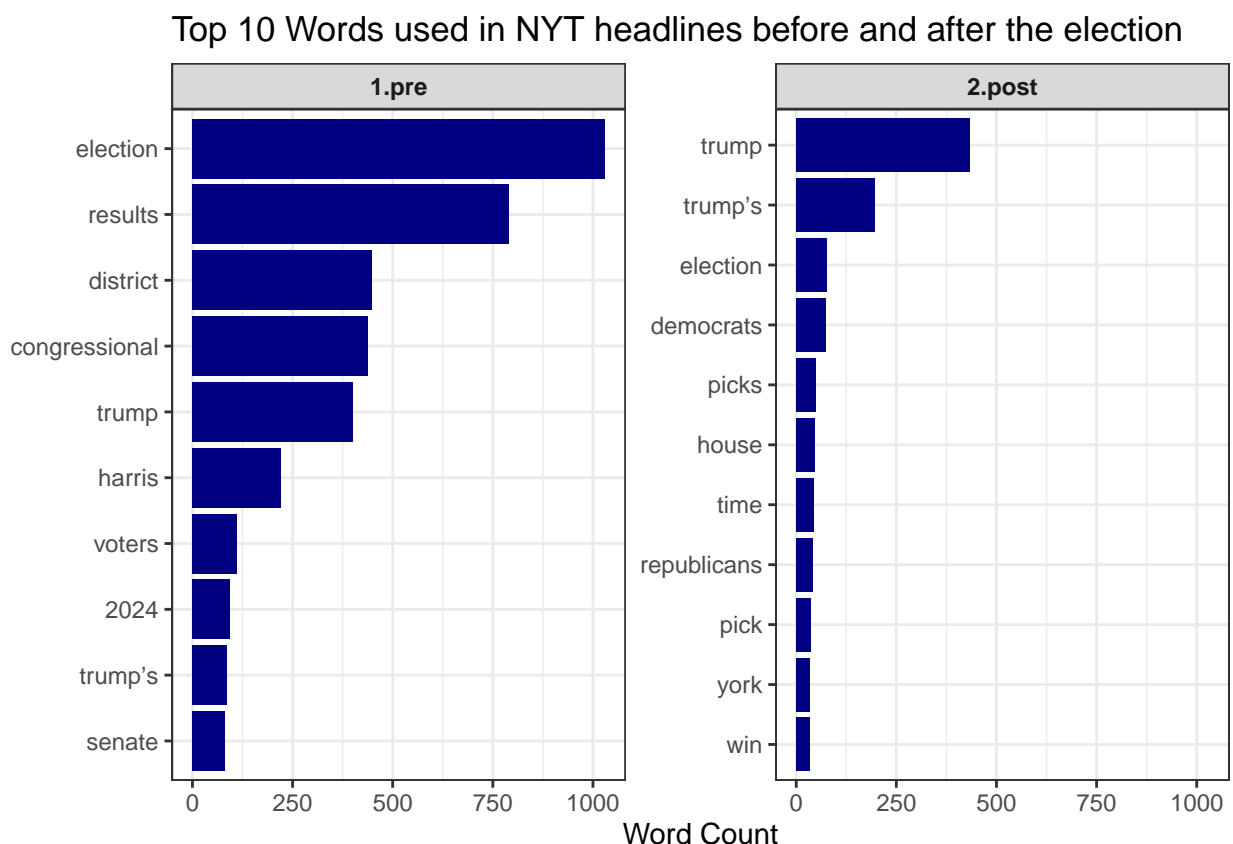
Methodology

In this project, we first leverage the New York Times' Archive API to scrape articles posted in the 30 days before and after the 2024 presidential election (Oct. 7 - Dec. 6, 2024). This corpus of articles - 8,873 in total - includes everything the New York Times printed or posted in this time frame. To avoid capturing the sentiment of headlines about unrelated topics like pop culture or cooking, we identified a set of keywords about the election and only kept articles whose headlines contained at least one of these key words. At the end of this data cleaning, we were left with a total of 3,173 articles.

To identify the sentiment and emotion of each headline, we use the distilBERT and the English DistilRoBERTa-base Hugging Face models. These pre-trained transformer models are trained on large text
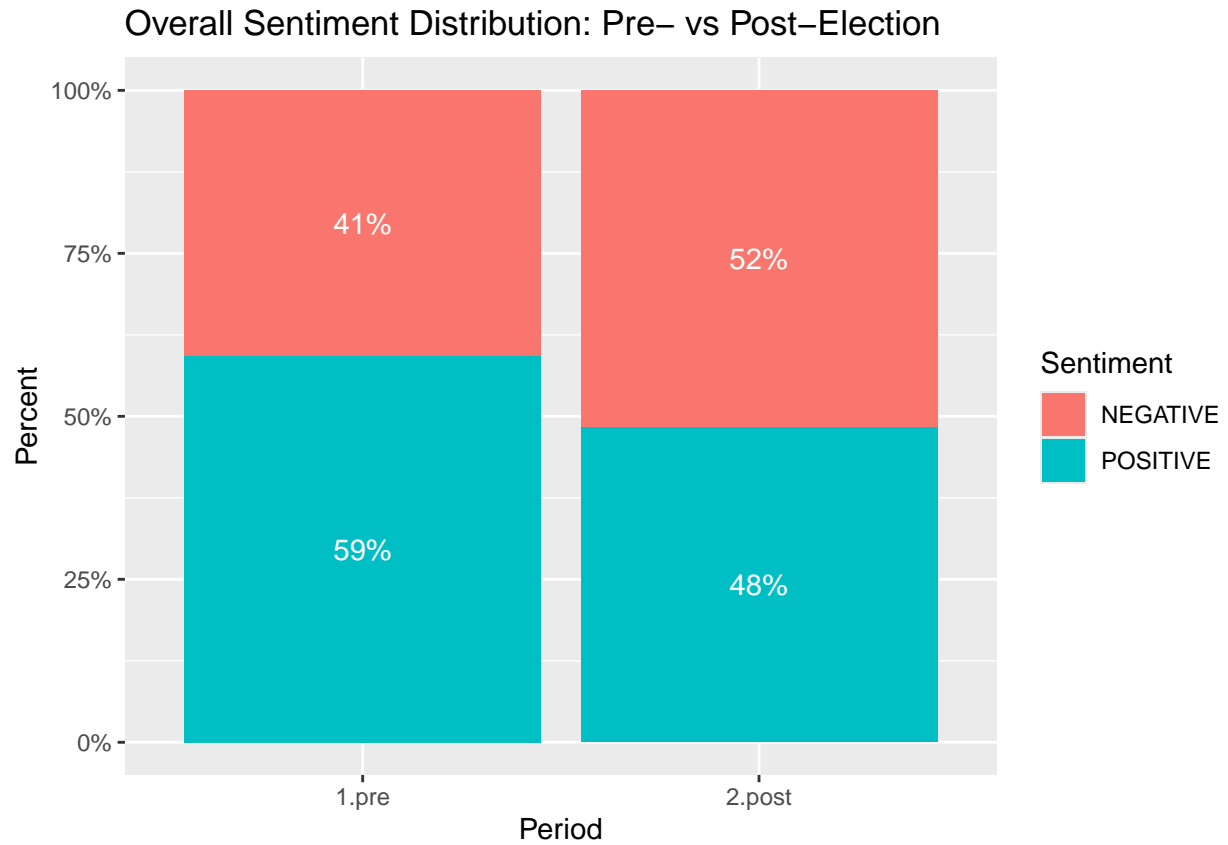
datasets and can understand the meaning of words in context, meaning they can pick up on things like word order and sarcasm. The distilBERT model returns a positive or negative classification of each headline, along with an associated confidence score. The distilRoBERTa-base model returns an emotion classification (anger, disgust, fear, joy, neutral, sadness, and surprise) and an associated confidence score.
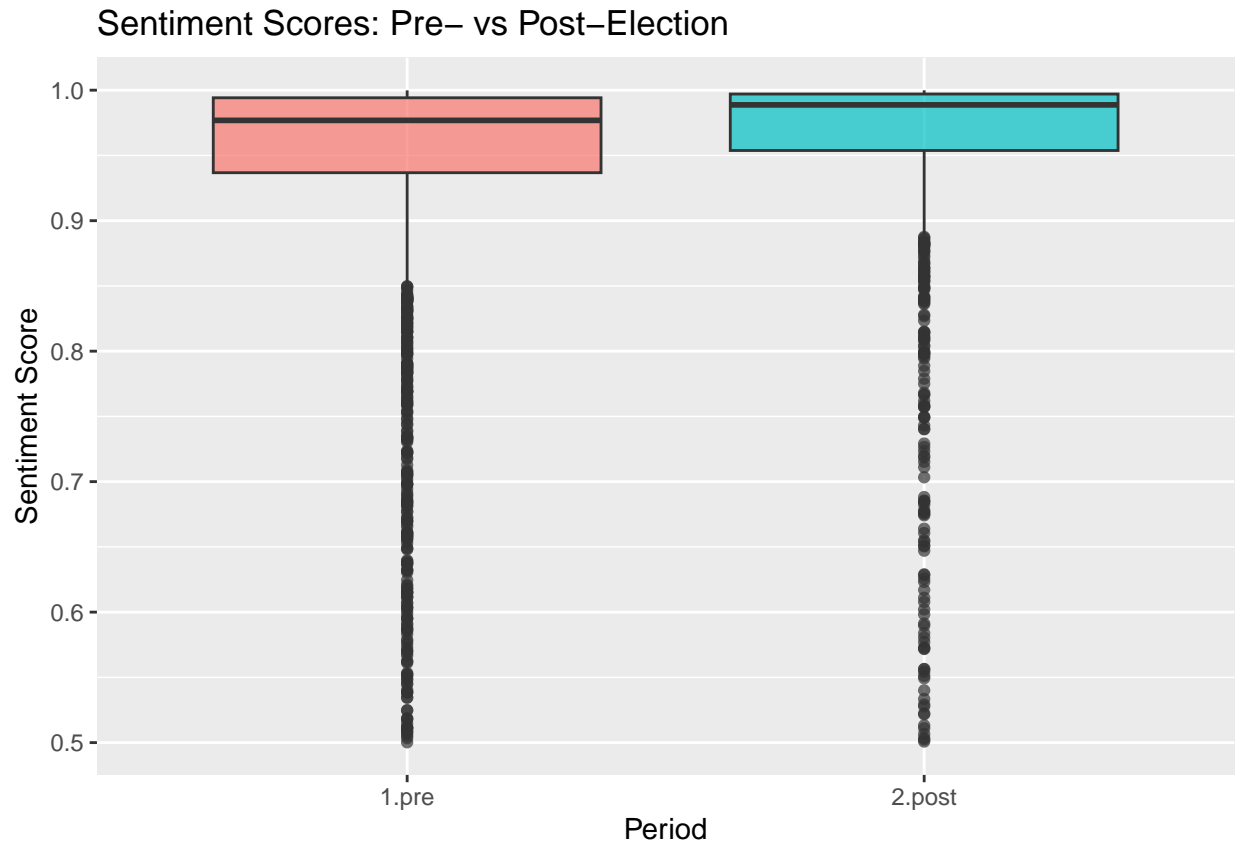
Results

The first thing we looked at was the top words used in headlines both before and after the election. Unsurprisingly, the top word used in the period before the election was the word "election," with this word appearing in more than 1,000 headlines out of around 2,100 headlines from before the election. The other top terms before the election included terms like district, congressional, as well as references to both Trump and Harris. After the election, we see Trump become the top two most used words in headlines, with most other words not appearing very frequently in headlines. For reference, Trump appeared in 630 out of 1,044 headlines post-election, with no other term appearing more than 100 times.

Top 10 Words used in NYT headlines before and after the election

In the 30 days before the election, a 59% majority of New York Times articles had a positive sentiment, with 41% being coded as having a negative sentiment. After the election, this dynamic flipped: Articles leaned slightly more negative (52% vs. 48%). One of our key research questions was whether or not the sentiment changed after the election and based on this topline analysis it sure seems like it did. We also briefly looked at if the OpEd section may have driven some of this negativity. Before the election, 57% of the headlines in that section were categorized as negative, after the election that figure was 59%. So while the OpEd headlines did get slightly more negative, it would not be accurate to say this drove the change.

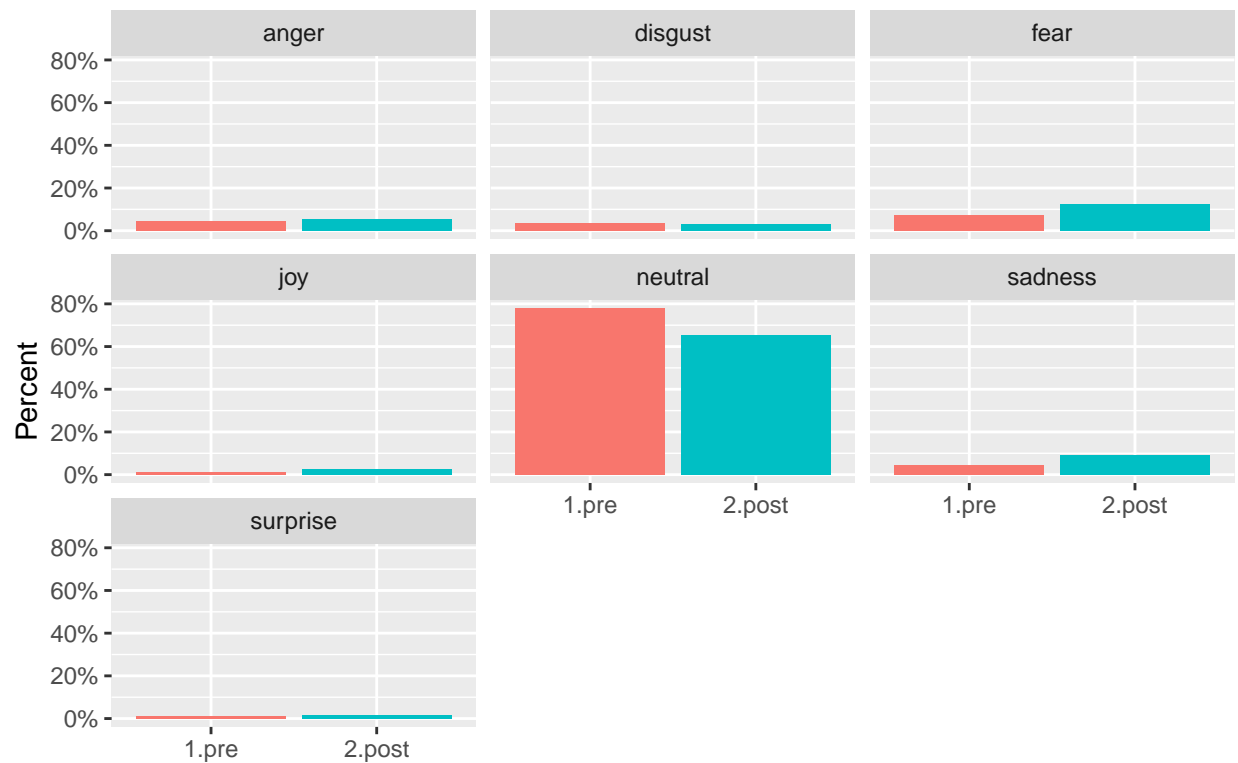## Overall Sentiment Distribution: Pre– vs Post–Election

We also checked to see how confident the model was in the sentiment scores that it produced and as you can see by the boxplot below, the mean sentiment score was over 0.95 indicating that the model was usually pretty confident it had categorized a headline correctly.
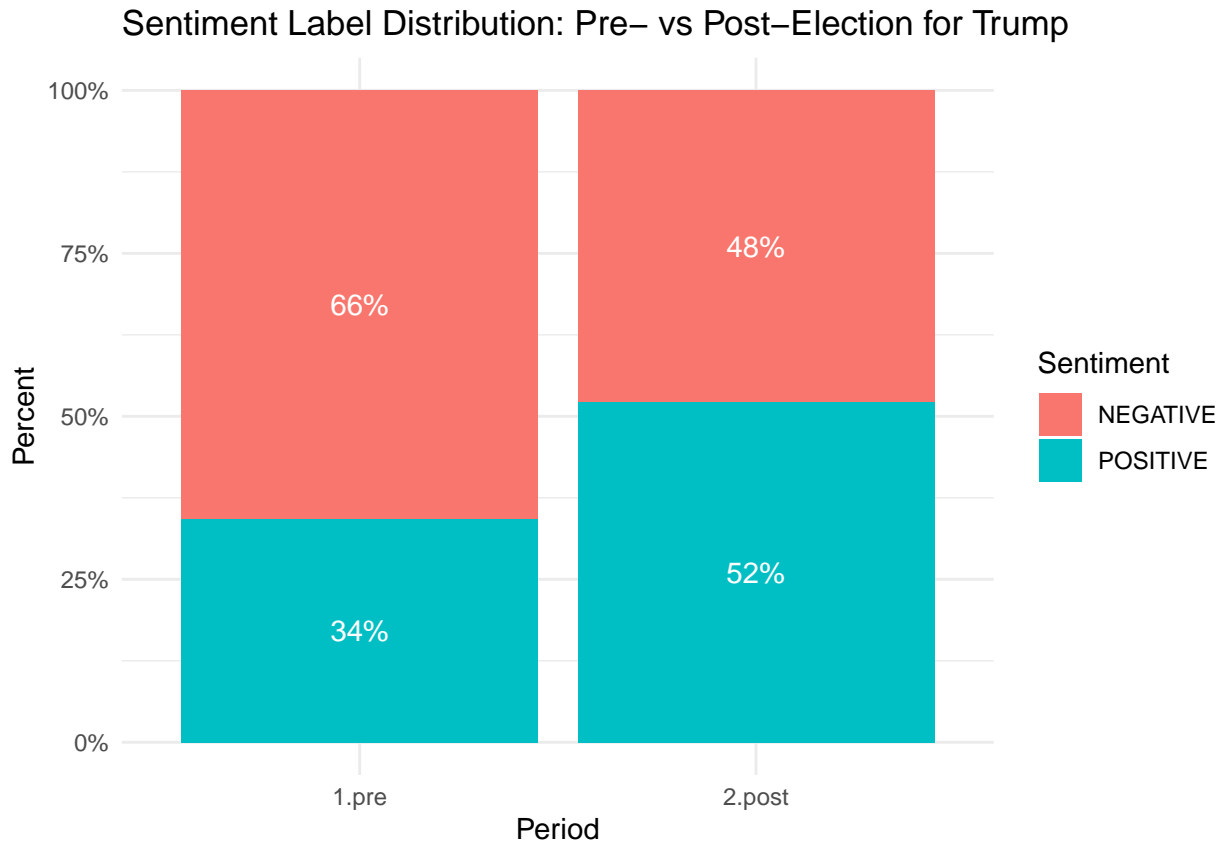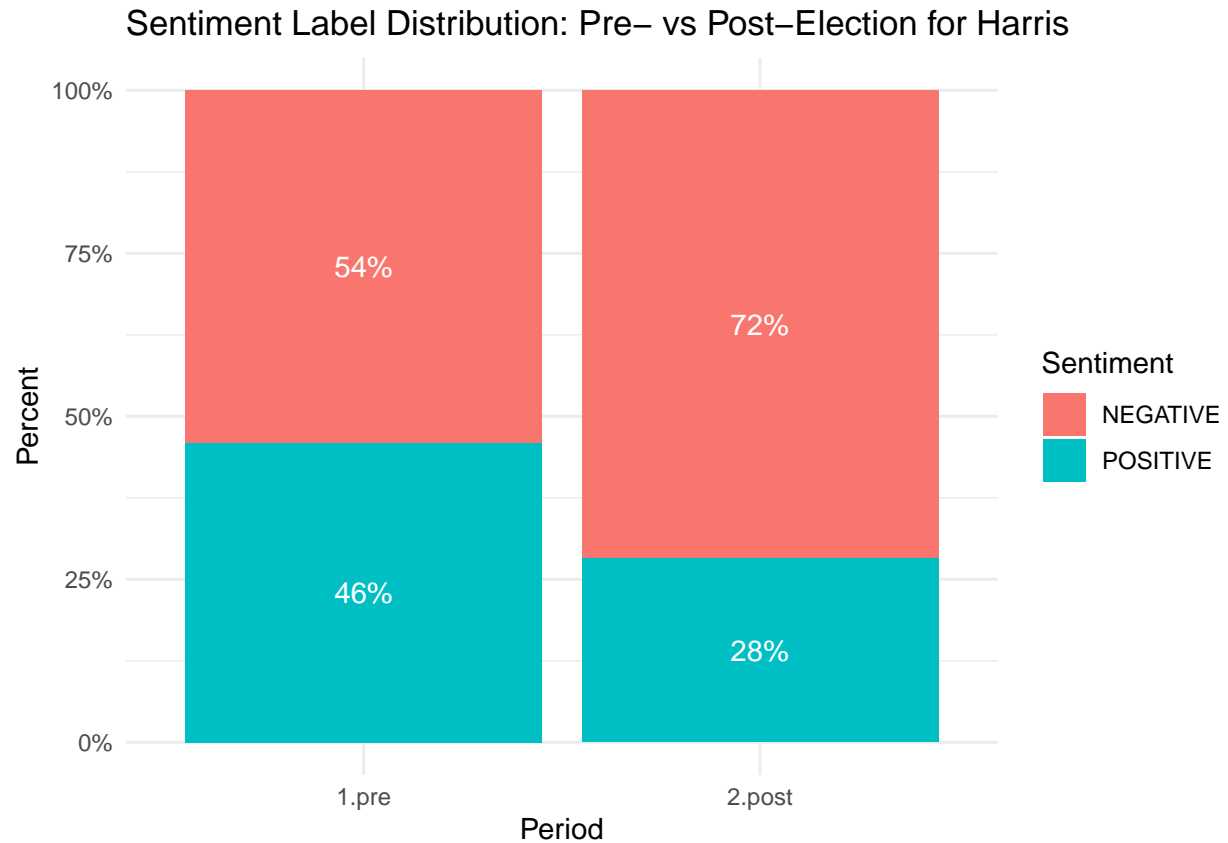
Sentiment Scores: Pre– vs Post–Election

The next step that we took was to look at the distribution of emotions that headlines had before and after the election. In both periods, the neutral category dominated, with 78% of headlines being coded as neutral before the election, and a somewhat smaller share (66%) being coded this way after the election. In the post election period, headlines were more likely to be categorized as fearful (7% before the election vs. 13% afterwards) and sadness (4% vs. 9%). The rest of the emotions did not change by more than one percent. This is another example of how the sentiment and emotion of headlines changed after the election compared with the pre-election period.

Emotion Category Distribution: Pre– vs Post–Election

The final piece that we looked at was to narrow in on headlines that contained the words "Trump" or "Harris." The dataset of over 3,000 headlines was subset and we ran similar analysis as we did above to see if headlines became more negative around the two candidates after the election or not. In the case of Donald Trump, about two-thirds of headlines before the election were negative. However, after the election, headlines were slightly more positive than negative (52% vs. 48%). For Kamala Harris, the opposite pattern emerged. Before the election, a small majority of headlines were negative, but that number grew to 72% in the period after the election.

Sentiment Label Distribution: Pre– vs Post–Election for Trump

Sentiment Label Distribution: Pre- vs Post-Election for Harris

Conclusion

Our core research question we were interested in asking was whether or not the sentiment of New York Times headlines changed meaningfully after the election. Based on the charts above I think we can conclude that they did. Overall, headlines became more negative. And more specifically, they became a little more fearful and sad. However, we saw headlines surrounding Kamala Harris get much more negative after the election, while those dealing with Trump became more positive. While we can't answer why this is the case, one could imagine winning an election (or losing in Harris' case) would bring about a change in the tone of coverage. It would be interesting for future research to explore the reason for this change beyond what we were able to do in this report.