

# Horse Race Journalism – News Headline Sentiment And Presidential Approval Ratings

Noah Julian Angara  
Introduction to Web Mining for Social Scientists

June 6, 2021

## 1 Introduction

Arguably the single most important number to a US president is their approval rating. As an indicator for political capital and public support, much of the president’s agenda hinges on this number. Hence, it can make or break a presidency, in particular concerning legislative success (Canes-Wrone & De Marchi, 2002). The importance of metrics such as approval ratings or polls has also influenced the way the media covers presidential issues. This phenomenon has been coined “horse race journalism”, as it focuses the coverage on data and polls rather than policy positions (Broh, 1980, p. 515). This potential disconnect between the media and what citizens care about was highlighted by former US president Trump. Against this background, the goal of this project is to explore the association between news headlines and presidential approval ratings, focusing on Joe Biden’s term from January 23, 2021 <sup>1</sup> until June 2, 2021. More precisely, by computing a sentiment score of news headlines this project analyzes whether it can be used to predict presidential approval ratings.

Previous research has outlined the main drivers of presidential approval ratings. Several studies identify the importance of economic factors in explaining approval ratings. For example, MacKuen, Erikson, & Stimson (1992) find that future economic conditions, measured by changes in inflation and unemployment are important determinants of presidential approval. This is confirmed by Jung & Oh (2020), who find that GDP growth and unemployment rate are significant predictors of approval. However, consensus is not clear as Berlemann & Enkelmann (2014) show in their literature review that the effect depends highly on the sampling period. Non-economic effects are explored in Brace & Hinckley (1991), who find that approval can be decomposed into a time, circumstance and administration-specific component. This project contributes to the literature by exploring a new channel which might influence approval ratings, namely the one via news headline sentiment. Intuitively, it is reasonable to conjecture that the public’s view of the president and by extension the approval ratings are influenced by the headlines they consume.

To investigate this issue, this paper is structured as follows. Section 2 outlines the data and web scraping strategy in R to obtain news headlines from *The New York Times* (NYT) and *The Guardian*. Then, section 3 explains how Granger causality is used to test the association between news headline sentiment and approval ratings, before discussing the results and providing concluding remarks. This paper shows that news headline sentiment contains no predictive power over approval ratings, but that the reverse is true; presidential approval ratings can to a certain extent predict news headline sentiment.

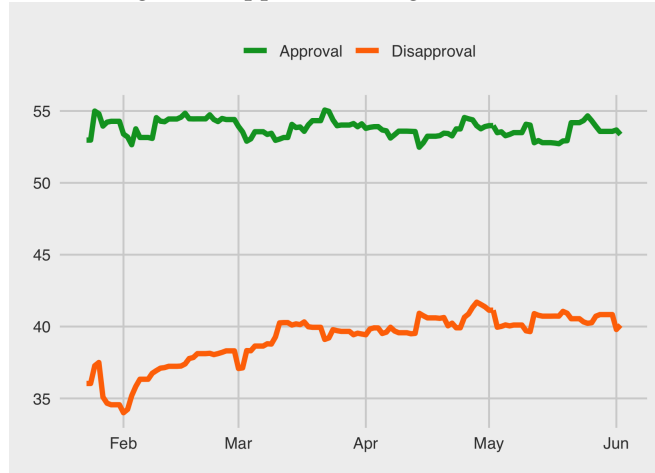
---

<sup>1</sup>This is when the first approval ratings were published

## 2 Data

The data for this project consists of two parts: data on Joe Biden’s approval rating pulled directly from the *FiveThirtyEight* website and news headlines scraped from NYT and The Guardian. NYT and The Guardian were the news outlets with the most straightforward ways of extracting news headlines, compared to pricier alternatives such as LexisNexis or Factiva. Figure 1 shows the evolution of Joe Biden’s approval and disapproval ratings since the start of his term in January. As is typical with US presidents, the disapproval rating is low at the beginning of the term and slowly increases as the honeymoon period wears off and the president faces more scrutiny. Table 1 provides summary statistics on the approval and disapproval ratings. Interestingly, there is more variation in the disapproval rating compared to the approval rating. This may suggest that Biden’s supporters are less swayed by his performance, while critics are prone to change their opinion more often.

Figure 1: Approval Ratings of Joe Biden



**Notes:** This figure shows the approval and disapproval ratings of US president Joe Biden. The data covers the period from 23rd January 2021 until the 2nd of June 2021. Source: FiveThirtyEight

Table 1: Descriptive Statistics of Approval Ratings

	N	Min	Max	Mean	Median	Var	SD	Skew	Exc. Kurtosis
Approval	131	52.480	55.065	53.776	53.761	0.343	0.586	−0.051	−0.852
Disapproval	131	34.004	41.696	39.115	39.704	3.182	1.784	−1.079	0.393

**Notes:** This table shows descriptive statistics for the approval and disapproval ratings (in %) of US president Joe Biden. The data covers the period from the 23rd January 2021 until the 2nd of June 2021. Kurtosis is indicated in excess to the one of the normal distribution (3). Hence, negative values are possible even though it is an even order moment.

Headlines from NYT were scraped using the Selenium framework within R. This allows the developer to automate data extraction tasks and control the client-server interaction, which is particularly important when the website is dynamic. In this project, Selenium was used to automate the process of loading all the articles in the search tool of NYT. A script was written which automatically clicks `SHOW MORE` until all the articles between January 23, 2021 and June 2, 2021 with the key word `Joe Biden` were loaded. Then, the headlines and dates were extracted from the web page’s source code. The script was executed in two parts, first extracting the headlines from March 22 until June 2 and then from January 23 until March 22. This resulted from the fact that the `SHOW MORE` button would not work after March 22 when querying the full

time window. This may be due to network issues on either the client or server side. Another aspect worth highlighting is the importance of letting the web page load before executing the rest of the script. Often the script would keep running but the virtual session would not show more results loading, resulting in needing to run the script again. Lastly, while the script handled the data extraction automatically, it was less efficient compared to more direct methods such as `rvest/httr`, taking  $\approx 27$  minutes.

Table 2: Number of News Articles With the Key Words *Joe Biden* by Source

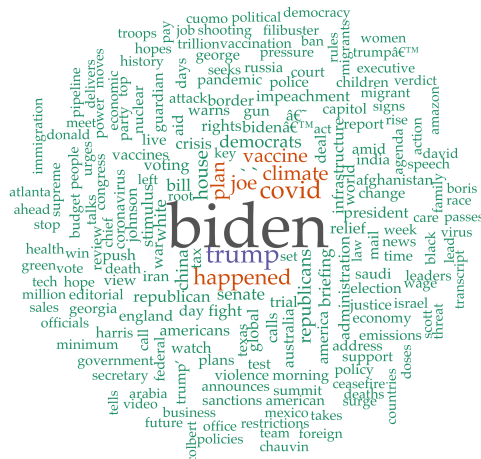
	N
The New York Times	1'921
The Guardian	2'199

**Notes:** This table shows the number of articles returned by querying the terms *Joe Biden* in NYT and The Guardian. The data covers the period from the 23rd of January 2021 until the 2nd of June 2021.

The headlines from The Guardian were obtained using its Open Platform API, querying all articles containing the search term `Joe Biden` from January 23, 2021 until June 2, 2021. The maximum number of results is set to 200 articles per query. Hence, the script had to download each page individually within a loop to circumvent this and get all the articles and pages in the time window. The output of The Guardian’s API is in the form of JavaScript Object Notation (JSON), a widely-used standard for formatting and transporting data. However, for the purposes of this project they had to be converted into a more usable format within R.

Headlines from NYT and The Guardian were then combined into one large sample. Table 2 shows how many articles were obtained – in total 4’120. Figure 2 shows the result of the data collection procedure in a word cloud with the most common words found in the news headlines. This is a good plausibility check to see whether the search engine of NYT and The Guardian’s API returned sensible headlines. Unsurprisingly, the main words are “Biden”, “Joe” and words which have been relevant over the past months such as “covid” and “vaccine”.

Figure 2: Word Cloud of News Headlines



**Notes:** This figure shows a word cloud with the most used words in headlines containing the search term *Joe Biden*. The data covers the period from 23rd January 2021 until the 2nd of June 2021.

### 3 Empirical Strategy

The aim is to test whether the sentiment score constructed from news headlines has any meaningful association with presidential approval ratings. The headlines are assigned a sentiment value via the “Valence Aware Dictionary and sEntiment Reasoner” (Vader) package in R. Vader uses a dictionary to map a sentiment score onto each word in a string (i.e., how positive or negative a word is) – in our case news headlines. The individual sentiment scores are then aggregated into a compound score for the whole string which is standardized to the  $[-1, 1]$  interval. To have the same sampling frequency as the approval ratings, the mean of all articles on a specific day is taken. This results in a daily time series of sentiment values, plotted in figure 3.

To test whether news headline sentiment has any relation with approval ratings, this project considers the simple test proposed by Granger (1969), as has also been done in MacKuen et al. (1992). While causality is perhaps a misnomer, Granger causality tries to make statements about the predictive power of one time series over another. More precisely, there are two models, one containing only an autoregressive component (equation 1) and another model with the lags of the explanatory variable, given by equation 2.

$$y_t = a_0 + a_1 y_{t-1} + \dots + a_k y_{t-k} + \epsilon_t \quad (1)$$

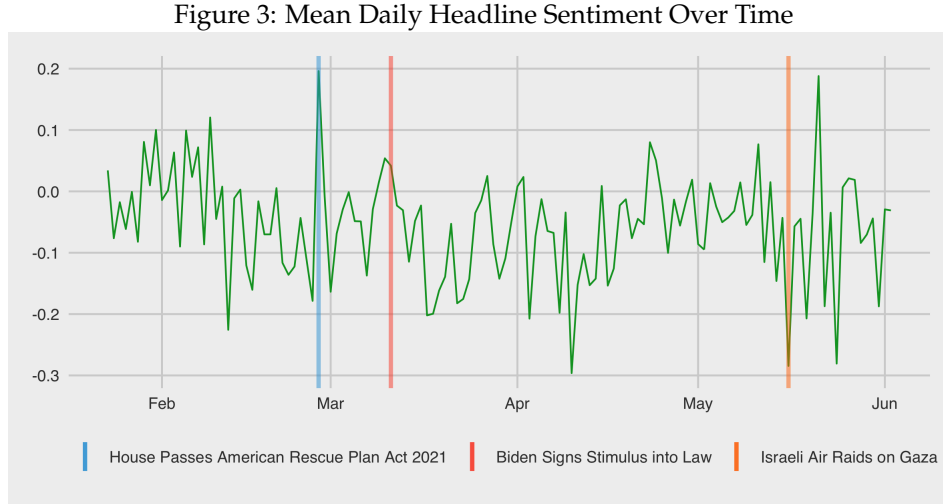
$$y_t = a_0 + a_1 y_{t-1} + \dots + a_k y_{t-k} + b_1 x_{t-1} + \dots + b_p x_{t-p} + \epsilon_t \quad (2)$$

In the above equations,  $\mathbf{E}[\epsilon_t | \mathcal{F}_{t-1}] = 0$ , because the error term is assumed to be white noise and hence not predictable given the information set at time  $t - 1$ . Using the restricted and unrestricted model, a simple F-test is performed to test whether one time series Granger causes the other. Under the null hypothesis,  $b_1 = \dots = b_p = 0$ , meaning that the lagged values of  $x$  provide no explanatory power about  $y$ . In our application, we are interested if the lagged values of the sentiment score have any explanatory power for presidential approval ratings (or vice-versa).

To avoid spurious regressions, the dependent variable is operationalized as the first difference of the approval rating, as both the correlograms and the Ljung-Box test indicated non-stationarities in the time series. As there is a bias-power trade off in selecting the lags in equation 2, a common approach is to base this decision on information criterion such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Following this procedure, the regression in equation 2 is run for different lags, with the chosen model being the one with the lowest AIC and/or BIC. Then, the chosen model is used to perform the Granger causality test.

## 4 Results

Figure 3 shows the result of the sentiment score constructed with Vader. The time series shows no obvious trend, but certain important events are evident. For example, the headline sentiment is highly positive when the US House of Representatives passed the Covid Stimulus Bill on February 26, 2021 or when Joe Biden signed it into law. Conversely, it is highly negative during international crises such as the Israeli air raids on Gaza in May 2021.



**Notes:** This figure shows the time series of computed sentiment scores for news headlines on Joe Biden. The data covers the period from 23rd January 2021 until the 2nd of June 2021.

Table 3 shows the results from the Granger test, using lags of order 2 based on the AIC and BIC. There is no empirical support for the hypothesis that news headline sentiment can help explain future values of presidential approval ratings. Hence, it does not appear that news headlines influence how citizens perceive Joe Biden's performance, at least based on articles collected from NYT and The Guardian. This is partially in line with the literature in that citizens may not care about the noise generated by headlines, but rather focus on bread-and-butter issues such as economic conditions. However, it still seems counter-intuitive that public perception is not shaped by the consumption of news, especially in the current polarized media environment.

Table 3: Granger Causality From Approval Ratings to Sentiment

	Residual Df	Df	F-Stat.	Pr(>F)
Model 1: Only Lags of Approval Rating	123			
Model 2: With Lags of Sentiment	125	-2	0.117	0.890

**Notes:** This table shows the output from running a Granger causality test from approval ratings to news headline sentiment, using lags of the order 2. The data covers the period from the 23rd January 2021 until the 2nd of June 2021. Statistical significance is given by \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

As an additional test, table 4 reports the results from running the Granger test in the opposite direction, namely if approval ratings can be used to predict news headline sentiment. The order of the lags was set to 1, as it is a different model and therefore the information criterion had to be recomputed. Interestingly, the test rejects the null hypothesis at the 10% level. Hence, there is tentative evidence that approval ratings can

be used to explain future news headline sentiment. Intuitively, this would mean that news headlines – to a certain extent and with a lag – follow public opinion concerning the president’s performance. This makes sense if news outlets focus on issues that are of interest to the public. If the public is currently debating an issue which negatively reflects on the president and their performance, the media will pick up on this and start covering this negative issue which will also be reflected in the sentiment of the headlines. Furthermore, news outlets want to maximize their readership. Hence, it makes sense from an editorial and business perspective to cover issues which are of public interest. Lastly, given the polarized media environment today, the fact that headline sentiment can be explained by public perception raises questions around objectivity, as news should not necessarily be influenced by public opinion but rather follow the facts.

Table 4: Granger Causality From Sentiment to Approval Rating

	Residual Df	Df	F-Stat.	Pr(>F)
Model 1: Only Lags of Sentiment	126			
Model 2: With Lags of Approval Rating	127	-1	3.428	0.066*

**Notes:** This table shows the output from running a Granger causality test from news headline sentiment to approval ratings, using lags of the order 1. The data covers the period from the 23rd January 2021 until the 2nd of June 2021. Statistical significance is given by \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

There are several limiting factors which may have influenced the above results. First, there could be confounding if there is a variable (e.g. economic conditions) which influences both news headlines sentiment and approval ratings. Future research should explore this angle to make these results more robust. Second, the data set only covers two news outlets, one being based in the United Kingdom. Therefore, the sentiment captured in the headlines is not comparable in its representativeness to the sample of citizens questioned for approval ratings. Third, the Vader lexicon consists of only roughly 7’500 words, which resulted in several articles being assigned a sentiment score of 0 as they contained words which are not in Vader’s lexicon. Hence, a more granular dictionary may provide more accurate estimates of news headline sentiment. Lastly, an inspection of the articles returned by The Guardian’s API showed several articles which were unrelated to Joe Biden. Therefore, future research needs to filter out these headlines to get a more accurate sentiment score. Nonetheless, this project provides first evidence of how news and public perception interact with respect to presidential approval ratings.

## 5 Conclusion

This paper sought to explore the relation between news headline sentiment and presidential approval ratings. By scraping news headlines from the NYT search portal with RSelenium and obtaining headlines via The Guardian’s API we were able to compute a sentiment score. Using the Granger causality test from approval ratings to news headline sentiment, the results show that news headline sentiment contains no predictive power for presidential approval ratings. Conversely, the Granger test for sentiment to approval ratings showed at the 10% level that approval ratings have some explanatory power for news headline sentiment. While tentative, this shows that news outlets cover presidential issues in a way that tracks public opinion of a president.

Future research should consider aggregating headlines from more news sources and using a richer lexicon to calculate sentiment. More importantly, an analysis should be conducted with a more extensive set of control variables (including economic conditions), as Granger (1980) outlines that the estimates can be spurious if relevant variables are omitted.

## References

- Berlemann, M., & Enkelmann, S. (2014). The economic determinants of us presidential approval: A survey. *European Journal of Political Economy*, 36, 41–54.
- Brace, P., & Hinckley, B. (1991). The structure of presidential approval: Constraints within and across presidencies. *The Journal of Politics*, 53(4), 993–1017.
- Broh, C. A. (1980). Horse-race journalism: Reporting the polls in the 1976 presidential election. *Public Opinion Quarterly*, 44(4), 514–529.
- Canes-Wrone, B., & De Marchi, S. (2002). Presidential approval and legislative success. *Journal of Politics*, 64(2), 491–509.
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, 424–438.
- Granger, C. W. (1980). Testing for causality: a personal viewpoint. *Journal of Economic Dynamics and control*, 2, 329–352.
- Jung, J. W., & Oh, J. (2020). Determinants of presidential approval ratings: Cross-country analyses with reference to latin america. *International Area Studies Review*, 23(3), 251–267.
- MacKuen, M. B., Erikson, R. S., & Stimson, J. A. (1992). Peasants or bankers? the american electorate and the us economy. *The American Political Science Review*, 597–611.

**Word Count:** 2'134