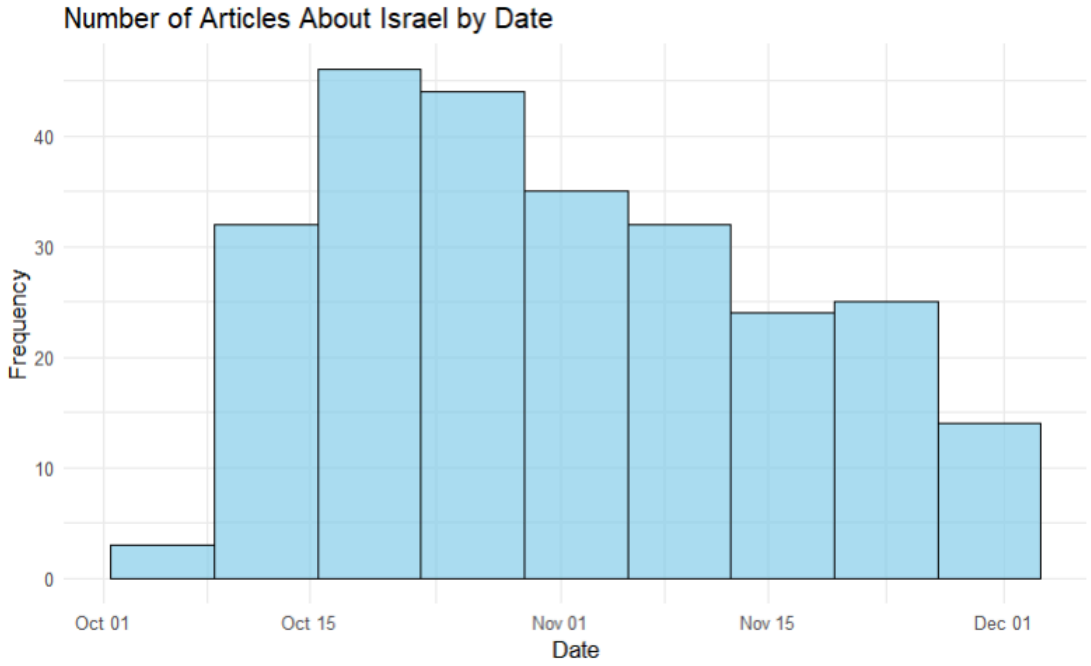
While the U.S. public has a reputation for being less informed on global issues than other countries, occasionally an international conflict occurs that consumes mainstream media and public discourse in the United States. Two prominent and recent examples are the Hamas attacks on Israel on October 7th, and the Russian invasion of Ukraine in February of last year. Media coverage on these events shapes the way the public views the conflict, and the U.S.'s role (if any) in what is happening in a specific region. In this project I decided to hone in specifically on the characterization of different political leaders in these conflicts, Putin, Netanyahu, and Biden. Putin and Netanyahu were chosen as the leaders of particular interest, as they are leaders spearheading their own individual military campaigns with varying degrees of popularity, and Biden was chosen as a pseudo-control, as he is mentioned a lot in both conflicts. To study their characterization, I performed a sentiment analysis on sentences that contained their names in them and compared them to sentences that didn't contain the name of any leader in them, to see if there was a noticeable difference. The results showed that sentences that contain a leader's name are more likely to be positive than sentences that don't contain the leader's name, and that Biden has a particularly strong influence on whether or not a sentence has a positive sentiment or not.

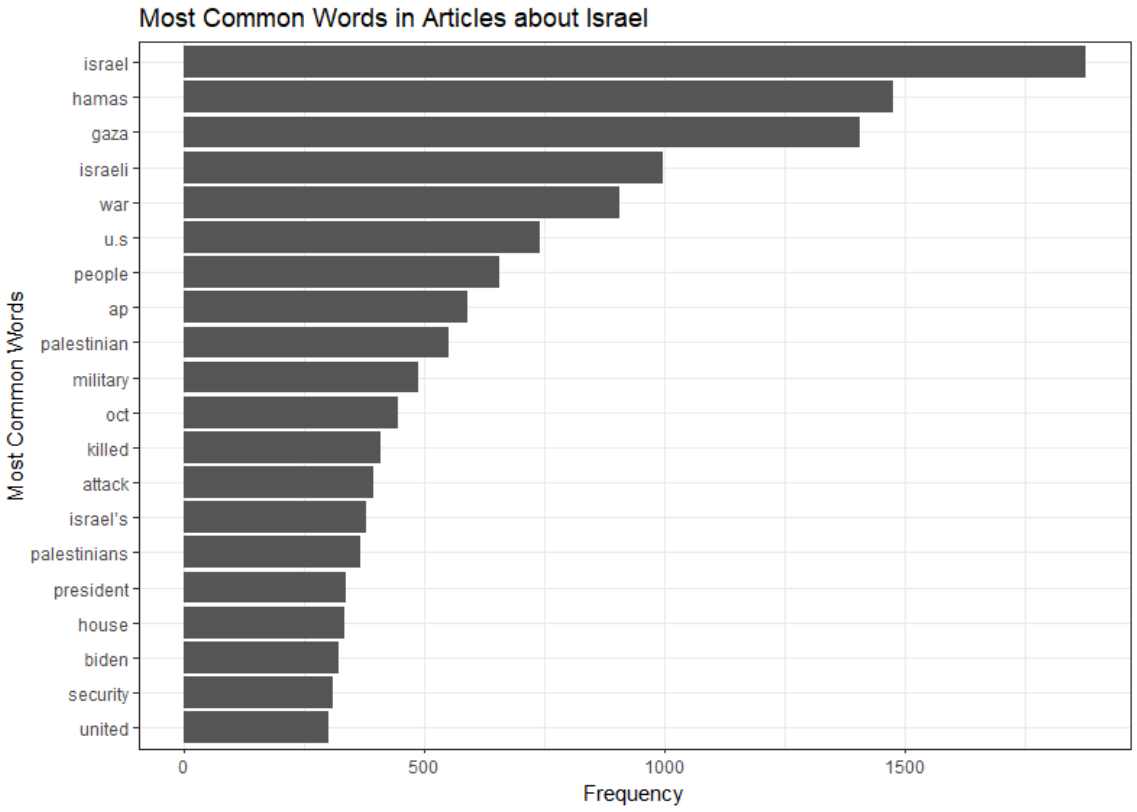
To study this, I decided to use articles from the Associated Press, as most major news outlets in the U.S. get their stories from them, and it seemed like a good way to account for partisan bias in major outlets like CNN or Fox. Using the GDELT database, which compiles global news coverage, I webscraped AP for all articles about Ukraine or Israel in a 2 month timeframe from the outset of the conflict. For Ukraine there were over 2000 articles in the period, but for Israel there were only around 200. Despite both conflicts being top of mind in the U.S. it is clear how the Russia Ukraine conflict would get more coverage on a global scale, as the implications for global energy prices, gas in Europe, grain exports, and global food supply are

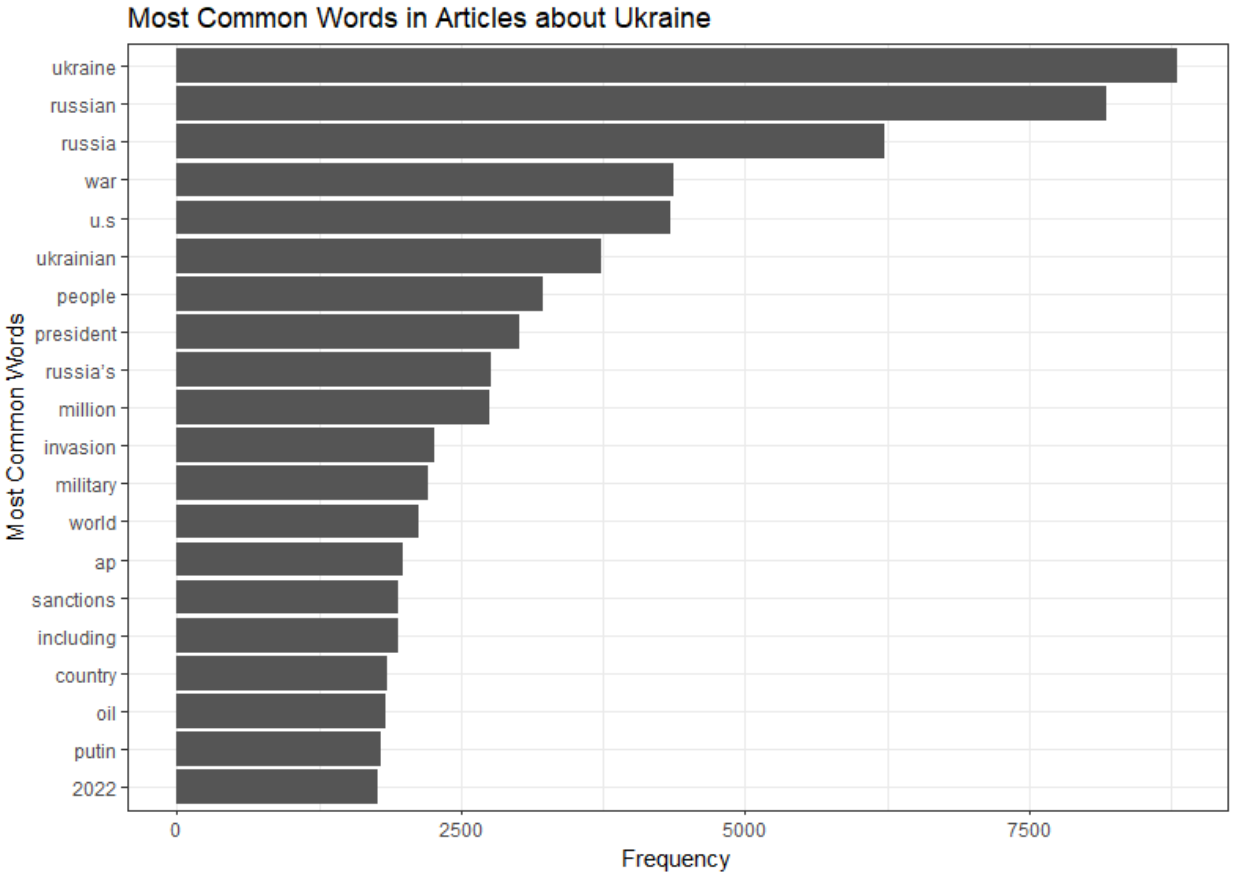
much broader. The conflict in Gaza doesn't have as wide an impact as the conflict in Ukraine. To visualize this we can look at the histograms detailed below.

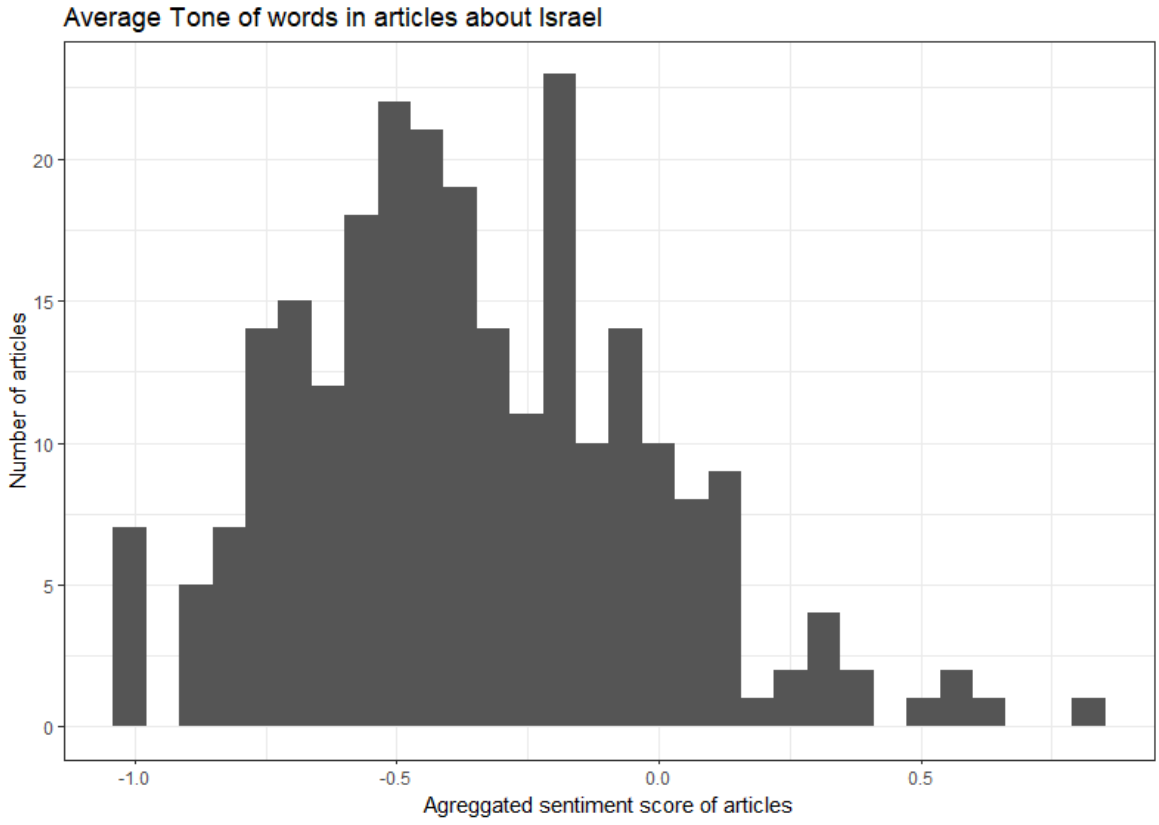


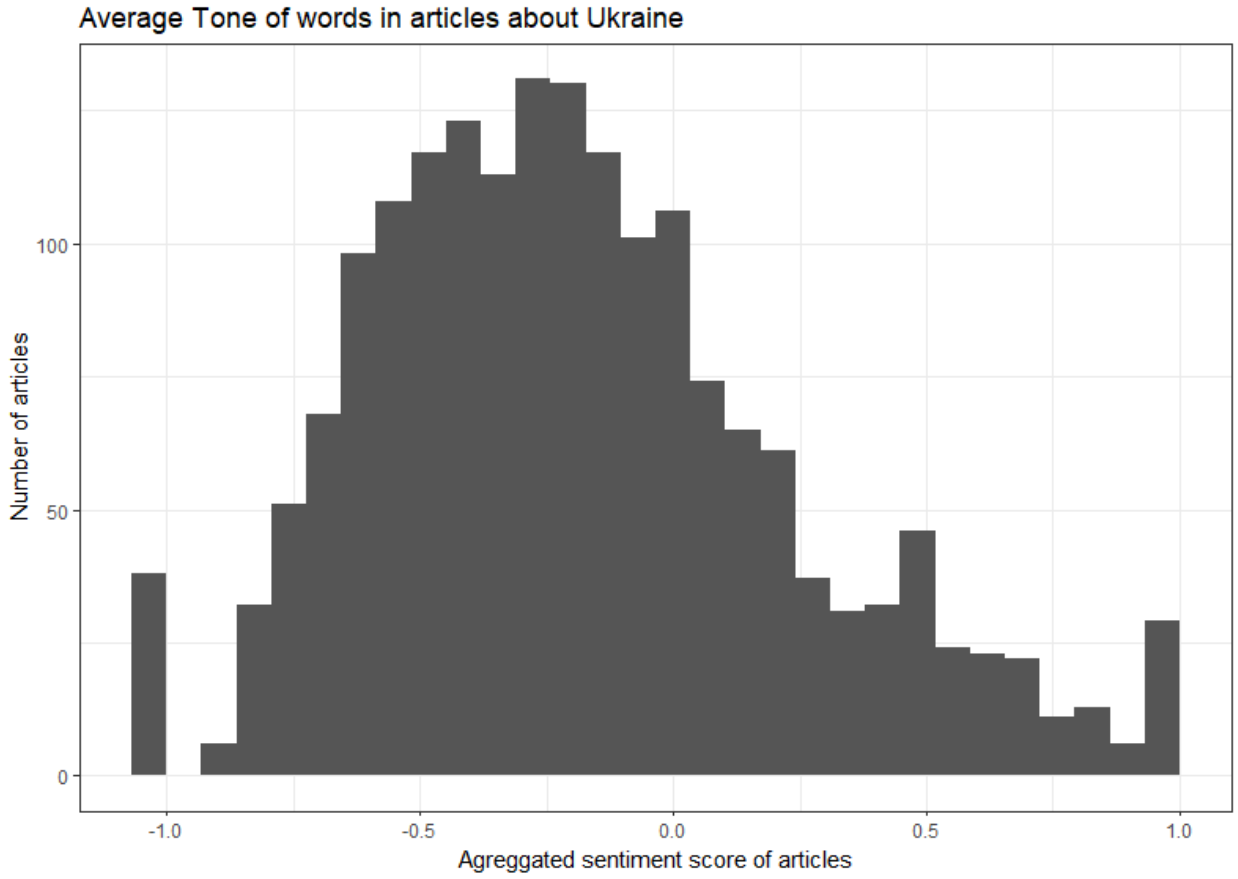
Both conflicts have similar distributions, with the most articles the week that the conflict started, and as the fighting rages on, articles slowly decrease in frequency. It is important to note that the Israel histogram is measured on a scale the 1/10th the size of the Ukraine histogram, as they got about 1/10th of the coverage that Ukraine did.

Now onto the text analysis portion. Text was analyzed in a few ways, the first being just a general aggregated word count of each article for each topic. The most common words are shown below for both Ukraine and Israel based on their total number of appearances. After this, each word in the article is given a sentiment score based on the Bing sentiment dictionary. Common filler words, or stop words are filtered out of this analysis, and all other words are given a positive or negative score based on how frequently that word is used in a negative or positive connotation. Though the model can occasionally mislabel the sentiment of individual words, on a large aggregated level it tends to be correct. The figures below show the results of text analysis on the article level, detailing both word counts and average sentiments.




Most Common Words in Articles about Ukraine
ukraine
russian
russia
war
u.s
ukrainian
people
president
russia's
million
invasion
military
world
ap
sanctions
including
country
oil
putin
2022
Most Common Words
0
2500
5000
7500
Frequency

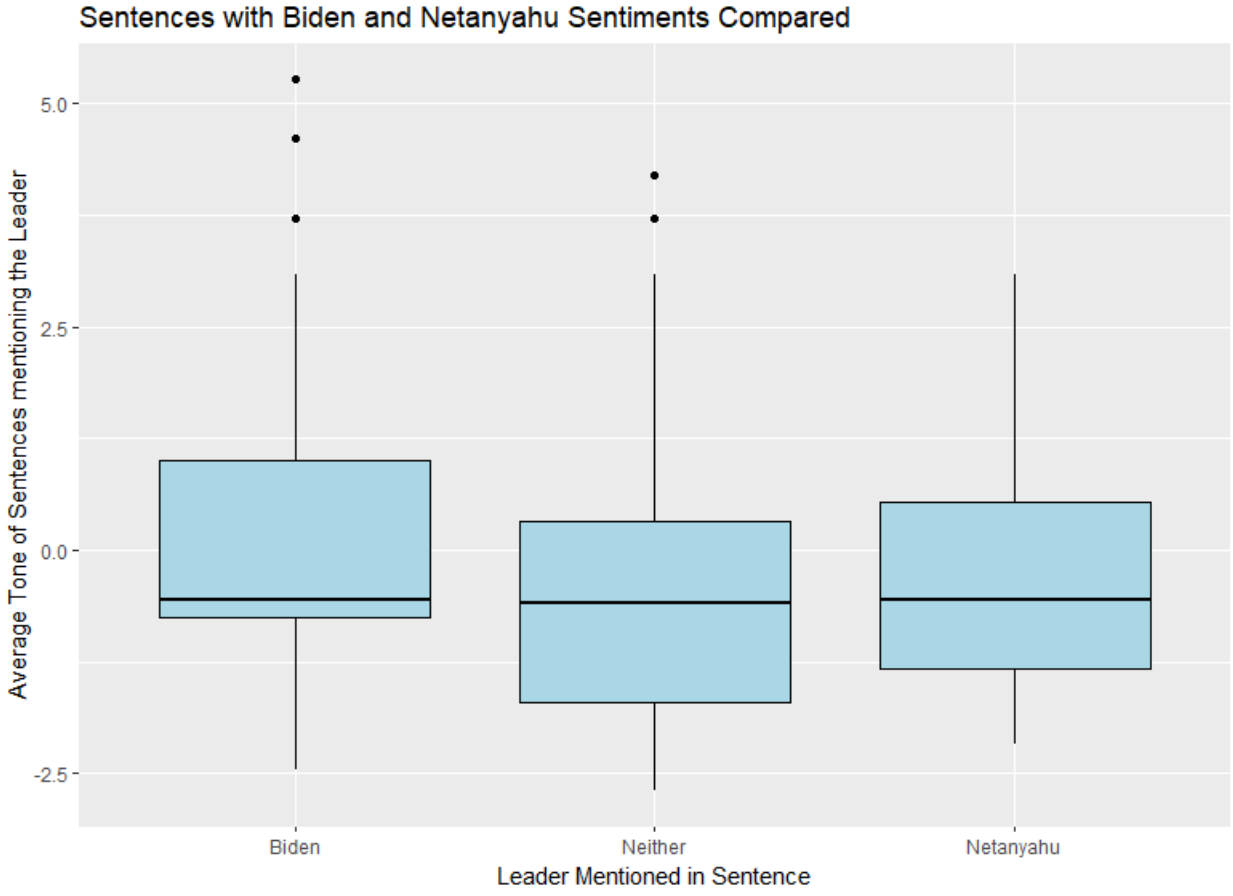

Average Tone of words in articles about Israel
Number of articles
0
5
10
15
20
-1.0
-0.5
0.0
0.5
Agreggated sentiment score of articles



Articles about both Ukraine and Israel had low average sentiment scores, which is generally unsurprising given the topic matter. The most common words in each category are words associated with the military, conflict, international politics, and specific groups of people involved in the conflict like Russians, or Hamas. Being that there is no real “control” group to base what the sentiment of an article should be, conducting any analysis at this level is quite difficult. Yes, all the body text of each article has an average sentiment score, but with no comparison these values don’t have all that much real world meaning. To change this, I decided to do sentiment analysis on an individual sentence level, comparing sentences that contain a leader’s name, or that don’t. In doing this I can compare sentences within the same article to determine if there is a statistically significant difference between sentences that mention a leader's name and sentences that don’t. This method accounts for the fact that all the articles

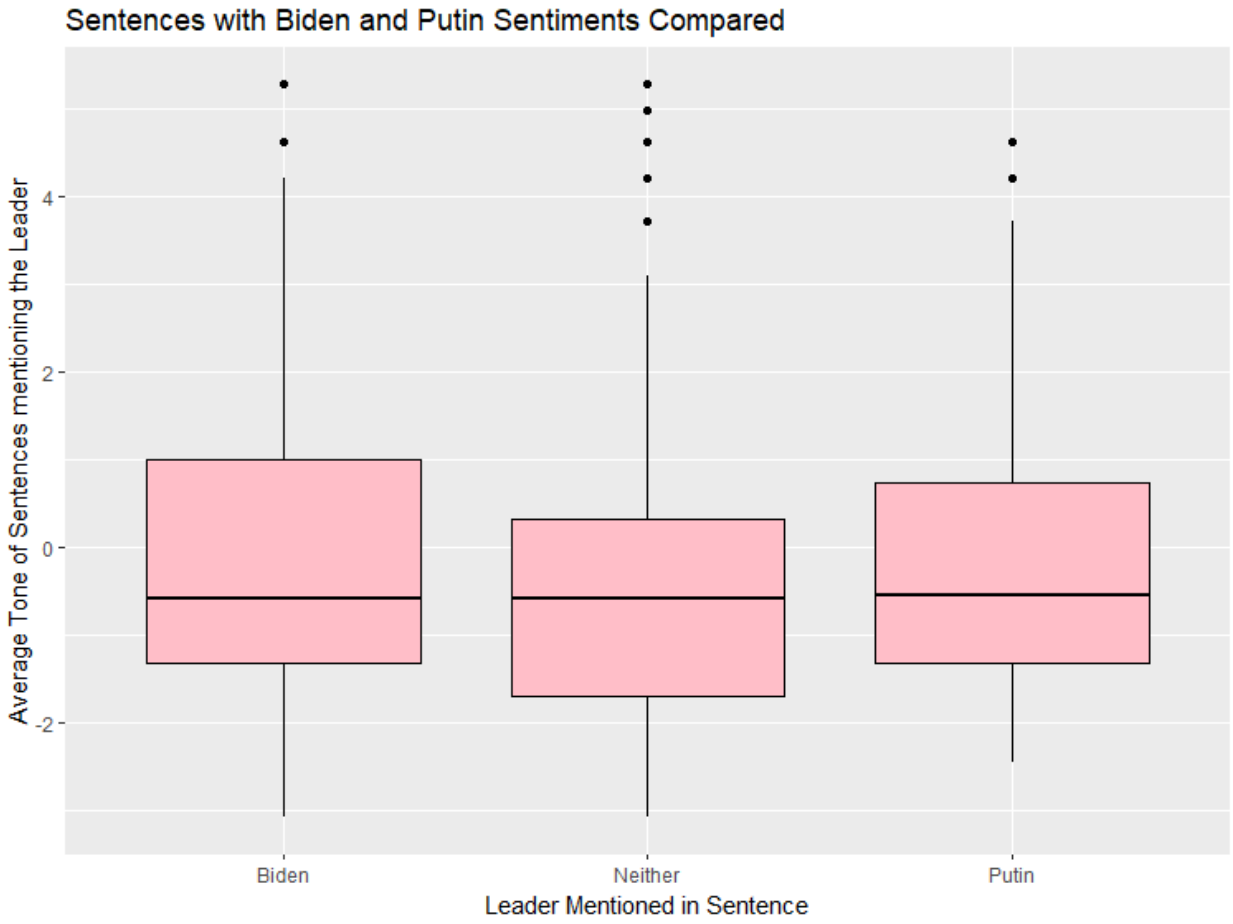
will have a low sentiment, due to their topic area. The results of sentence analysis are modeled in the regression and box plots below.

Israel Regression	Beta Coefficients	Std. Error	P value
Sentence With Biden (Intercept)	-.012	.07	.862
Sentence With Neither	-.48	.07	>.001
Sentence with Netanyahu	-.18	.12	.173



Ukraine Regression	Beta Coefficients	Std. Error	P value
Sentence With Biden	-.19	.03	>.001

(Intercept)			
Sentence With Neither	-.31	.03	>.001
Sentence with Putin	-.02	.04	.62



The coefficient differences in sentiment score are small enough that it is easier to see differences between them in the table rather than the box plots of the distribution of sentence's average sentiments. Starting off with the Israel regression, neither values for Biden nor Netanyahu leaders have statistical significance, which is likely a result of the small sample size. However, there are enough aggregate sentences with and without leader's names to show a significant negative difference in the sentences that don't contain leaders' names.

In the Ukraine analysis, there is still insignificance with the Putin coefficient, but the Biden and “Neither Leader” coefficients are significant and show that sentences where no leaders are mentioned are generally more negative than those that do.

On looking into the box plots visualization of the data, the plots for both Ukraine and Israel show that the upper half of Biden’s distributions tend to be higher than the others. This is likely a result of two factors, Biden not being directly involved in the conflict meaning higher sentiment words are used in sentences with him in them, and quotes from Biden that would contain statements of support for either Israel or Ukraine. If Biden were to “strongly stand behind Israel/Ukraine” that would be considered a positive sentiment sentence due to the adjectives and adverbs used to describe how the U.S. will support the regime involved in the conflict. Another note is that Netanyahu’s coefficient is quite close to statistical significance in such a small sample size, which could denote for future work that there may be a relationship to examine between the mentioning of his name and the sentiment of the sentence surrounding it.

Though it is disappointing to not see the statistical significance that we had hoped for, there are still conclusions to draw from this project for future work. We did find a significant difference in news coverage between the conflicts, higher than average sentence sentiment in sentences that contain Biden’s name, and a possible lead on whether or not the mention of Netanyahu would bring down the average sentiment of sentence (it would seem so but we can’t be sure without a larger sample size).