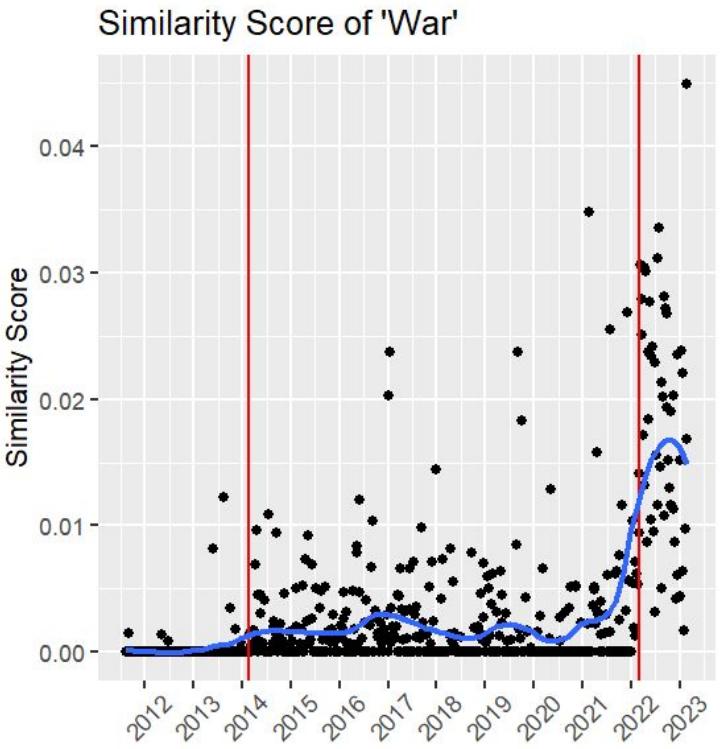
Analysis of Visualizations

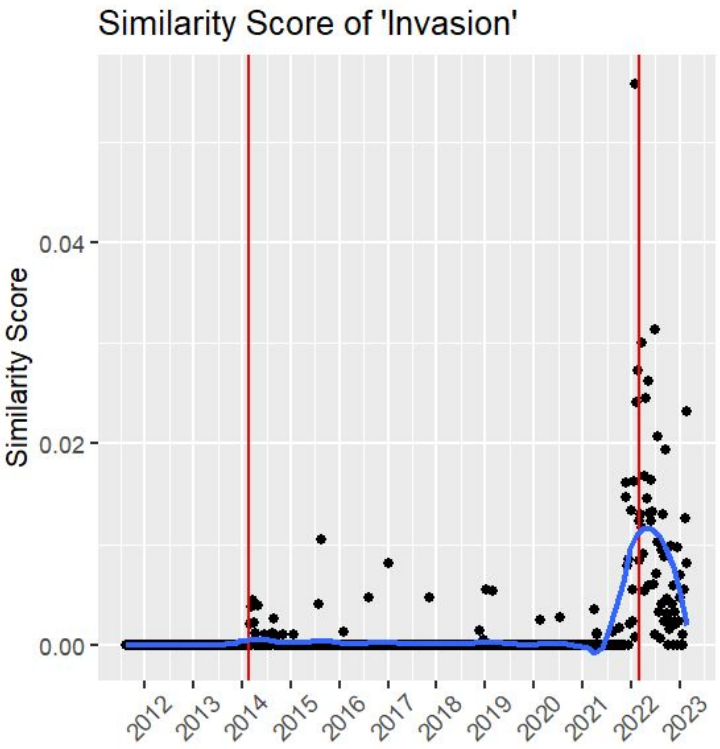
Nya Feinstein

Visualization 1: Similarity Score of “War”



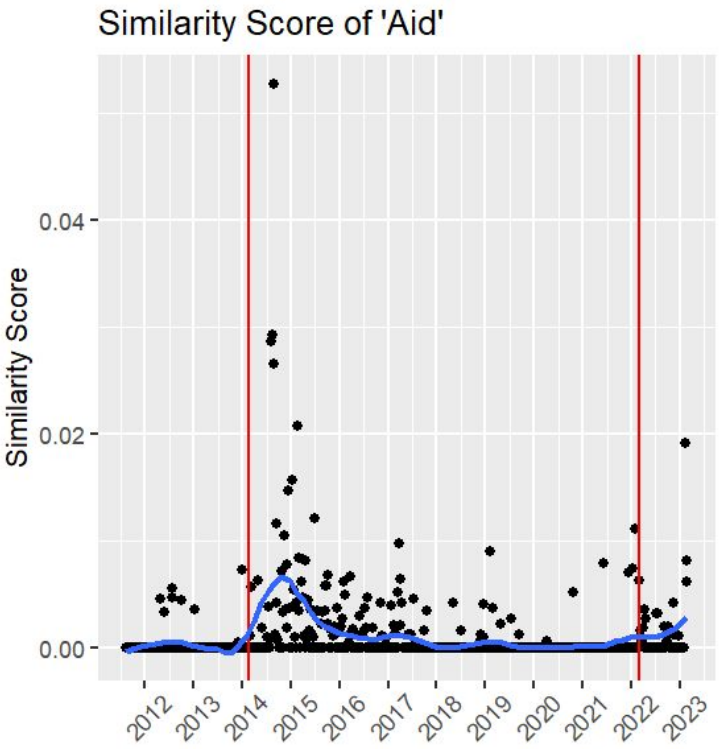
From this visualization, we see that there was a significant increase in similarity before the 2022 invasion of Ukraine. By contrast, there was only a slight increase in score before the 2014 annexation, but there was an overall increase and some consistency between the two nodes of conflict. With this information, it can be inferred that the 2022 invasion was seen broadly as having the classification of war and was not unexpected.

Visualization 2: Similarity Score of “Invasion”



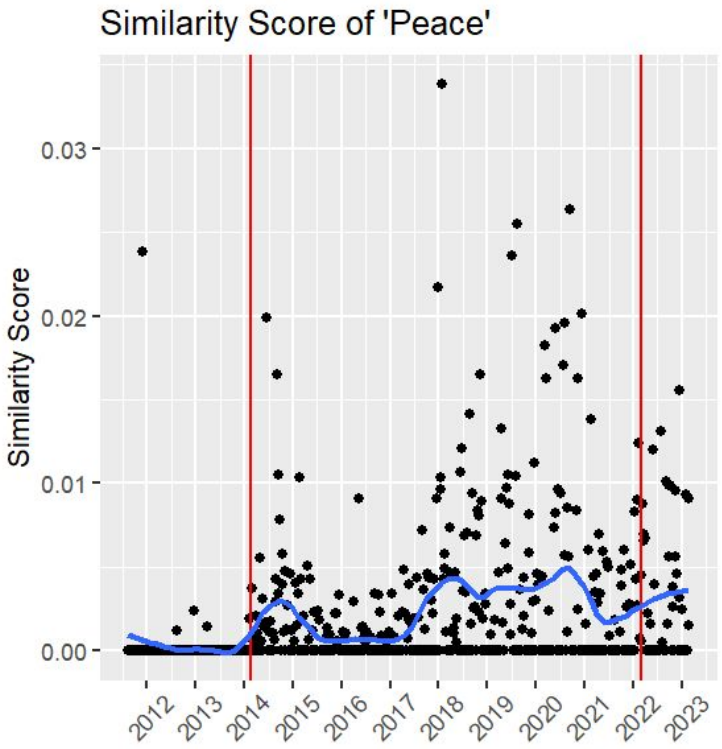
From this visualization, we see that the first similarity matches with the word “invasion” occurred on and after the annexation of Crimea. About two years before the Russian invasion of Ukraine, the scores began to increase. The maximum value is right on the eve of the invasion. From this, it appears that the annexation of Crimea (which included an invasion) was more of a surprise than the invasion of Ukraine in 2022, which began to have more mentions of invasion at least a year before the onset of the conflict.

Visualization 3: Similarity Score of “Aid”



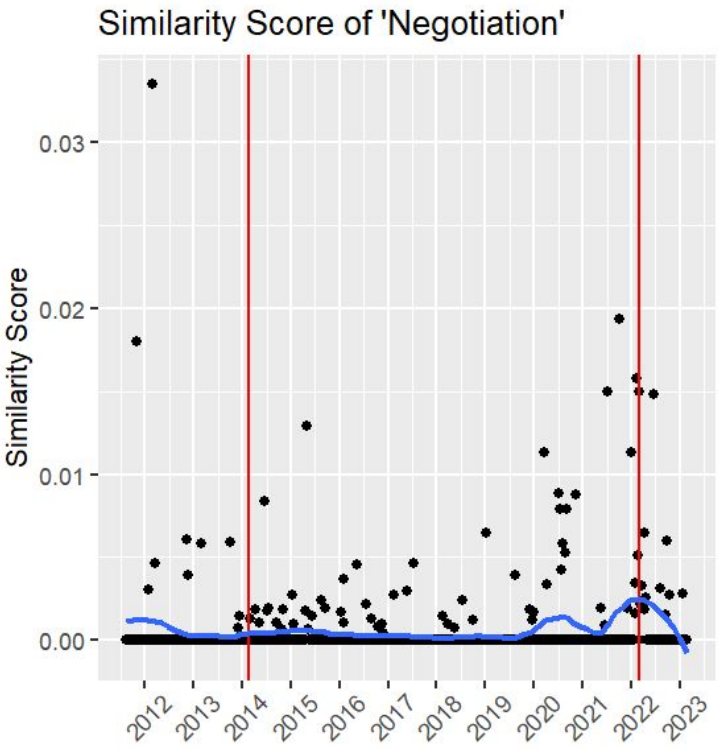
From this visualization, we see that the issue of aid was more of a conversation surrounding the annexation of Crimea and other events that could have been occurring during that time. The maximum value was in mid 2014, a few months after the annexation. There was an increase of articles with a similarity score to “aid” around mid-2021, about six months before the invasion of Ukraine, and it continues to increase with the local maximum occurring in early 2023. This makes sense — as more violence occurs, more aid is needed. Mentions of aid seem to usually come after the node of conflict.

Visualization 4: Similarity Score of “Peace”



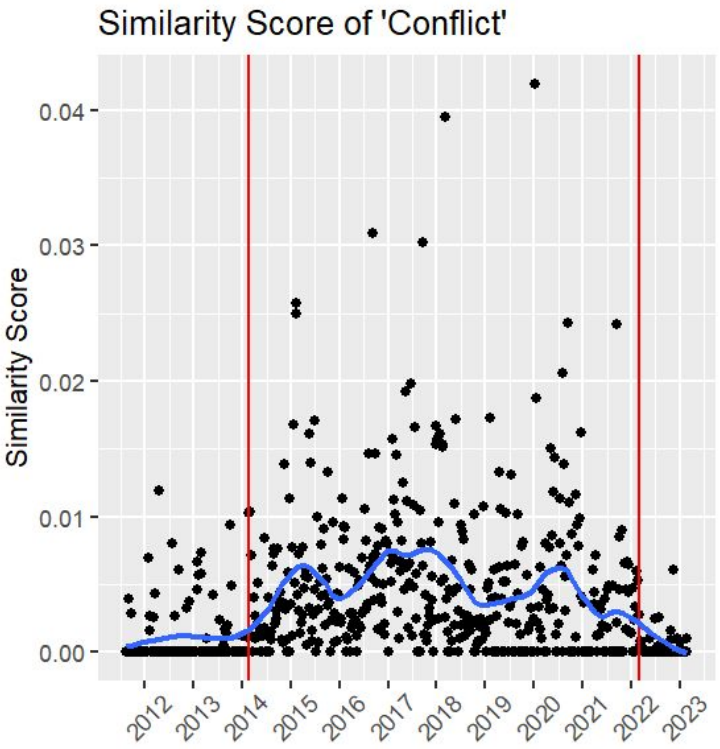
From this visualization, we see that the conversation surrounding peace was maximized between the two nodes of conflict. This is to be expected. I’m not sure whether this pattern is helpful as the term “peace” could be used both in times of peace (when peace is present) and in times of conflict (when peace is not). More investigation needs to be done on the sentiment surrounding this term and how it changes.

Visualization 5: Similarity Score of “Negotiation”



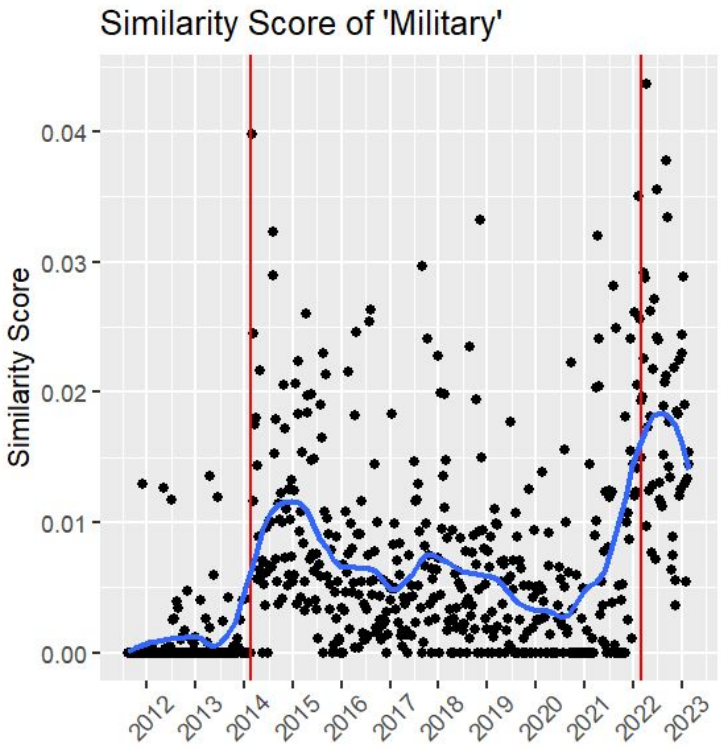
From this visualization, we see that there were not many points represented. From what we do have, the maximum datapoint occurs in 2012. The minimum is about halfway between the two nodes of conflict, which is expected as that is when no negotiations are needed. It should be noted that there was an increase of reporting surrounding this term before the 2022 invasion.

Visualization 6: Similarity Score of “Conflict”



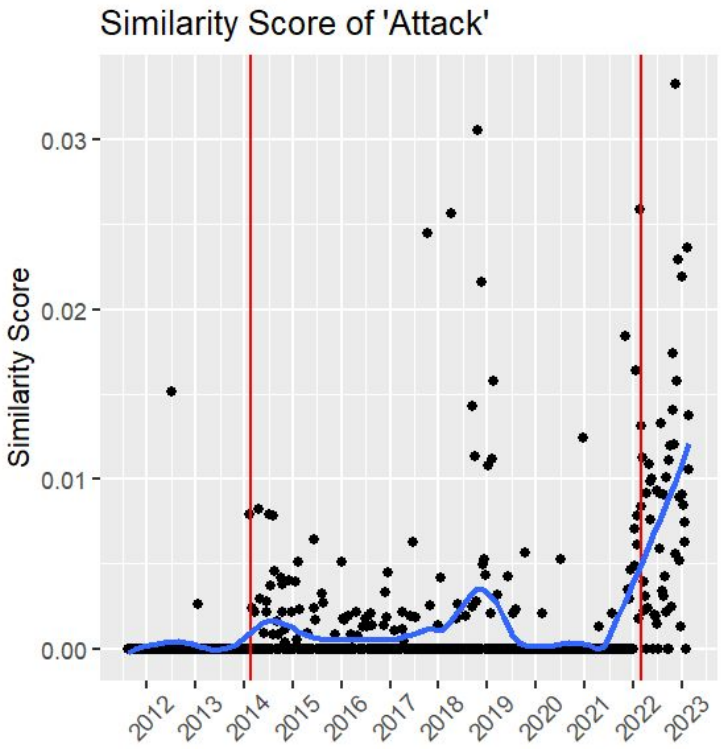
From this visualization, we see that the maximum fitted similarities with “conflict” actually peaked in around 2017 — around the halfway point between the two conflicts. It seems that this term is too broad to use for prediction as there is a clear trend, but it does not relate to the nodes of conflict. The maximum value was around 2020.

Visualization 7: Similarity Score of “Military”



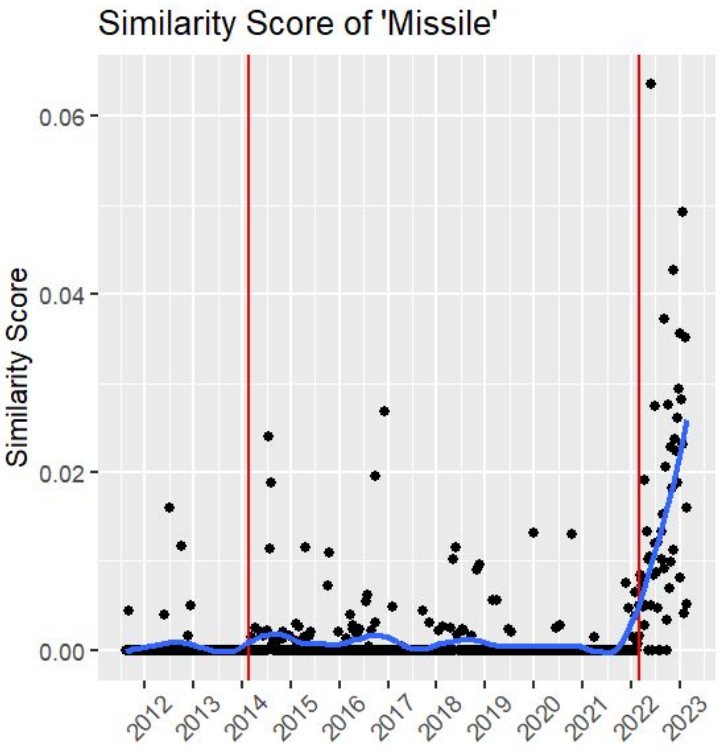
From this visualization, we see that there are clear trends around the nodes. About a year before the annexation of Crimea, there was an increase in similarity of articles with the word “military”. It continued to increase with a local maximum beyond on the date of the annexation. The fit gradually decreases, until a similar pattern appears about two years before the Russian invasion of Ukraine with a another increase (this time sharper) and a maximum value directly after the invasion.

Visualization 8: Similarity Score of “Attack”



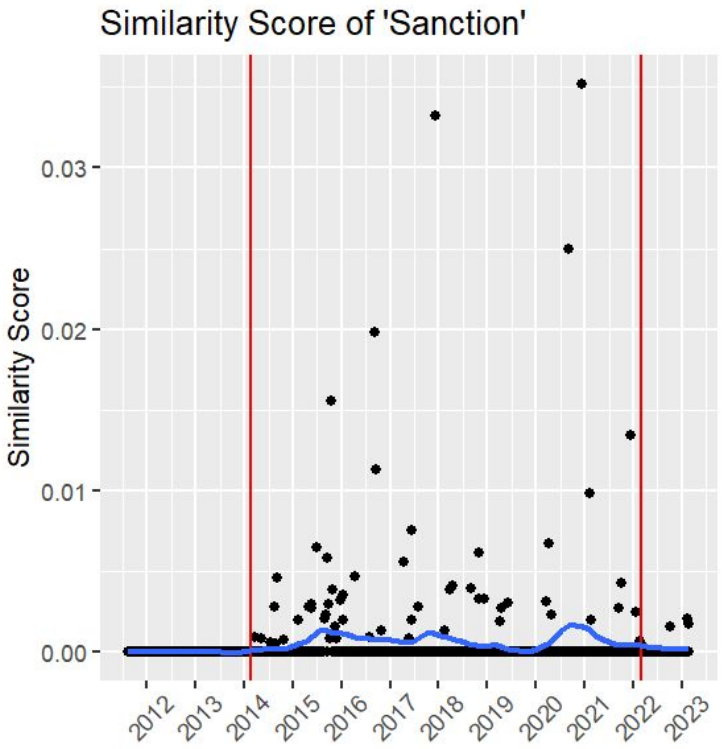
From this visualization, we see that there is not an incredibly clear trend surrounding the annexation of Crimea, though there were some medium values of similarity occurring around the conflict. It appears that there were some high similarity values around 2018 - 2019, but they appear to not be significant. There was a clear upward trend, however, beginning about two years before the Russian invasion of Ukraine. The maximum was around the beginning of 2023.

Visualization 9: Similarity Score of “Missile”



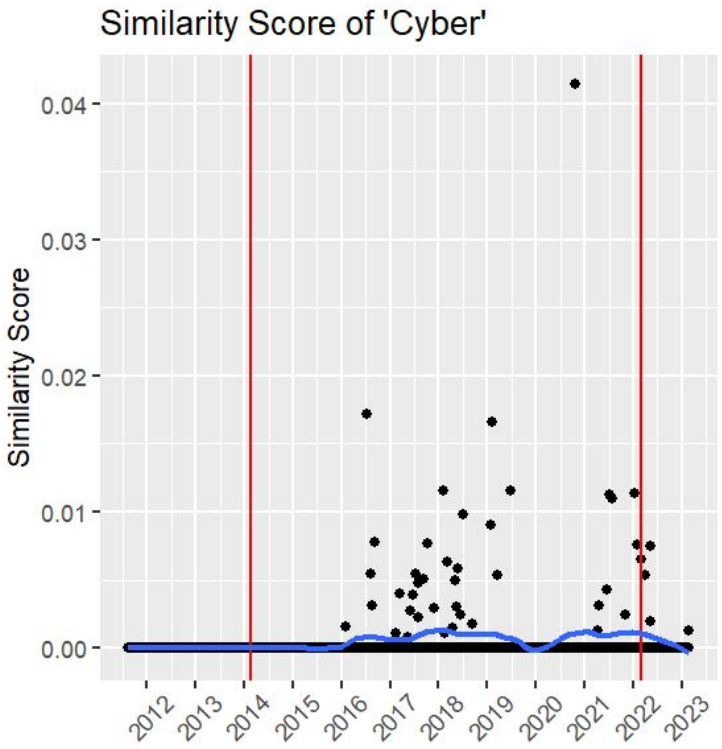
From this visualization, it appears that there was a most noticeable increase in reporting surrounding missiles on the eve of and during the 2022 invasion.

Visualization 10: Similarity Score of “Sanction”



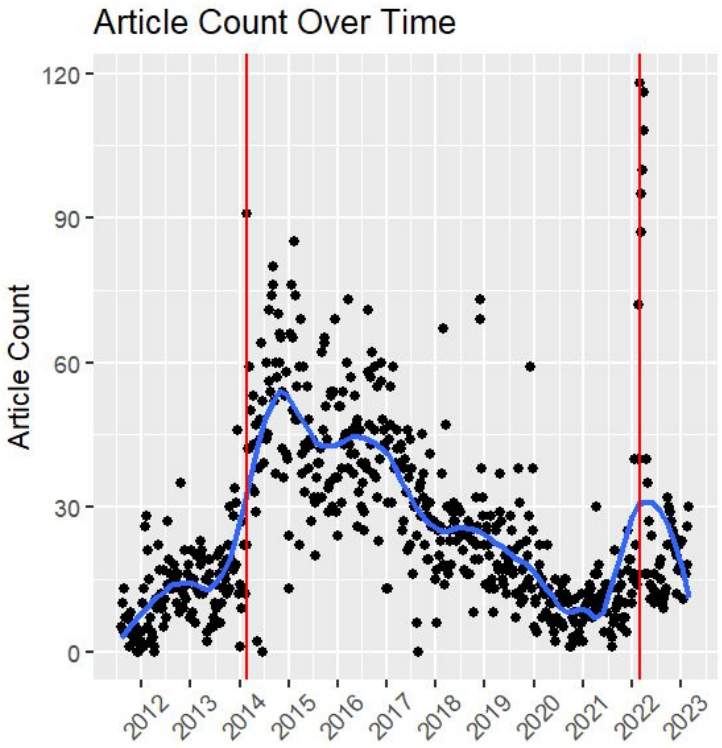
There is not a clear relationship between the nodes of conflict and the similarity of the word “sanction”. The maximum value of the similarity occurs in 2021. There was no similarity to this term before the annexation of Crimea

Visualization 11: Similarity Score of “Cyber”



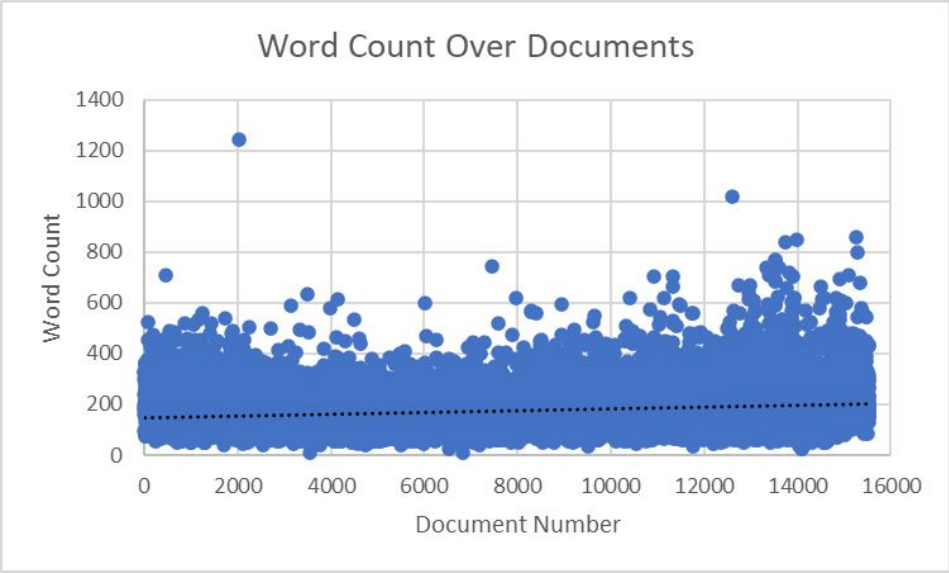
Unsurprisingly, the similarity of articles to the word “cyber” is a relatively new phenomenon. It seems that there are instances of similarity clustered around 2016 - 2018. I suspect that this has to do with the American election and Russia’s alleged meddling. There are some similarities around the invasion of Ukraine, but no gradual trend.

Visualization 12: Article Count



Keeping in mind that these articles were all queried from a Ukrainian news source with the keywords Russia, Putin, and Moscow, we see from this visualization that Ukrainian news was talking about Russian related topics in an increasing manner before the annexation of Crimea. This peaked about two years after the annexation and then gradually decreased, with an increase again about a year before the Russian invasion of Ukraine. The maximum articles published on this topic occurred on the exact time of the invasion.

Visualization 13: Word Count



One assumption made during the analysis was that the word count in all documents did not strongly vary. This visualization proves that this analysis was appropriate as the trend of word count is approximately zero.