

The bulk of our project was “Implementation,” but we decided to also give an empirical analysis to display/test the usefulness of our implementation. We have 4 “mini-experiments” which use our code to answer interesting questions.

First, we want to see if an article deemed as very optimistic or pessimistic can be manually confirmed to be optimistic and pessimistic, which will give us an idea that our vector space model is working.

1. Question: Do the most optimistic/pessimistic articles pass the eye test?

Hypothesis: To a human, the most optimistic article returned by the code is clearly more optimistic than the most pessimistic article returned by the code, and likewise the most pessimistic article returned by the code is clearly more pessimistic.

Experiment:

The following is one test we used to answer this question.

Filter: All regions, publisher = BBC News, anytime, all titles

From the 211 BBC News articles analyzed:

*Most positive article: <https://www.bbc.com/news/world-52470472>
with positivity 0.01811835866522568*

*Most negative article: <https://www.bbc.com/news/world-middle-east-52493608>
with negativity 0.024530989777740944*

*Biggest difference: <https://www.bbc.com/news/world-middle-east-52493608>
with more negativity by 0.01717235846359454*

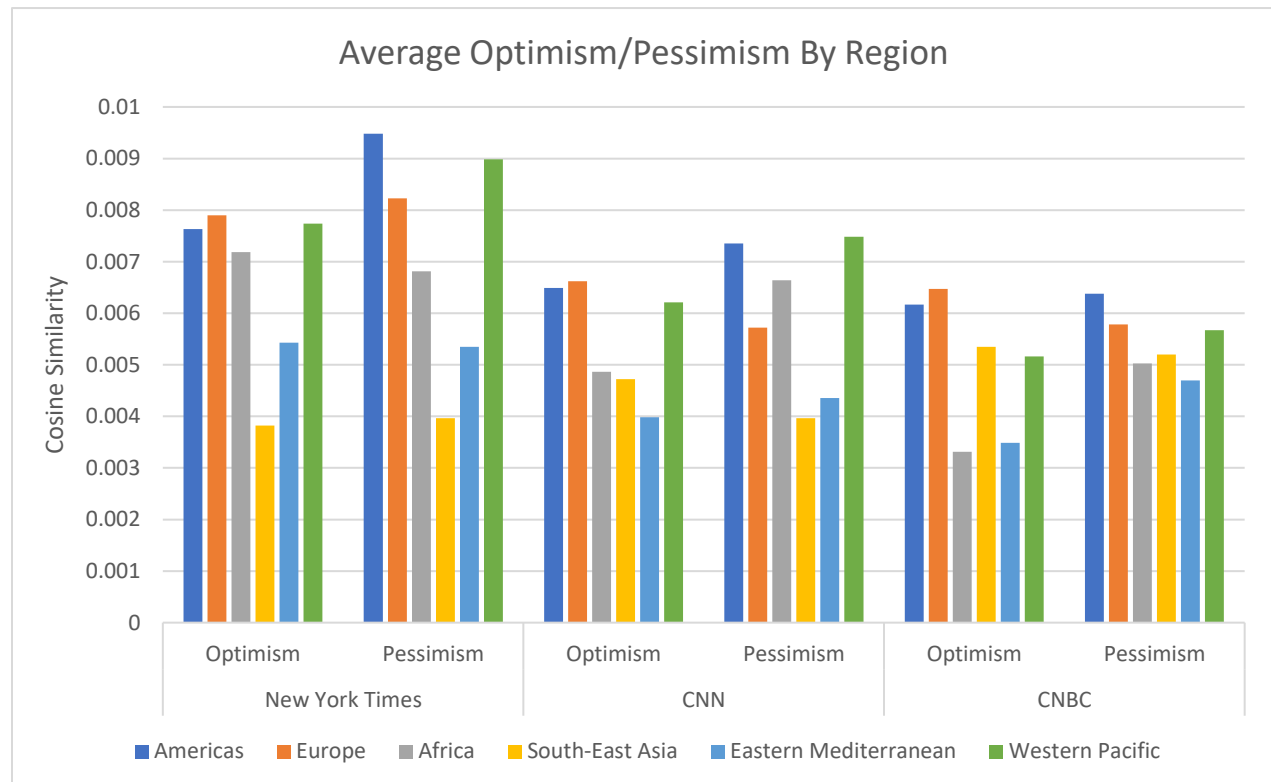
By manually reading these articles, it can clearly be seen that the most positive article, which talks about ideas and inventions that can reform back to normal life, is more optimistic than the most negative article, which discusses the tension, anger, and collapsing economies in the Middle East during the coronavirus. It also makes sense that this article had the biggest difference in positivity and negativity.

Conclusions: Although our code is definitely not going to be a completely accurate ranking of how optimistic and pessimistic articles are, in general highly optimistic and highly pessimistic articles will have higher optimism and pessimism ratings. In other words, the most optimistic and pessimistic articles usually pass the eye test.

2. Question: How does optimism and pessimism differ across regions?

Hypothesis: Due to the continued high coronavirus rate in the U.S., articles pertaining to the America region will have greater pessimism and lower optimism than for other regions.

Experiment: Filter: all regions, repeated for the top 3 worldwide publishers (New York Times, CNN, CNBC), anytime, all titles



It can be seen that for all the publications, America is among the highest two regions in pessimism, supporting our hypothesis that articles about America may be more pessimistic than articles about other regions. However, against our hypothesis, optimism for America is also high compared to other regions.

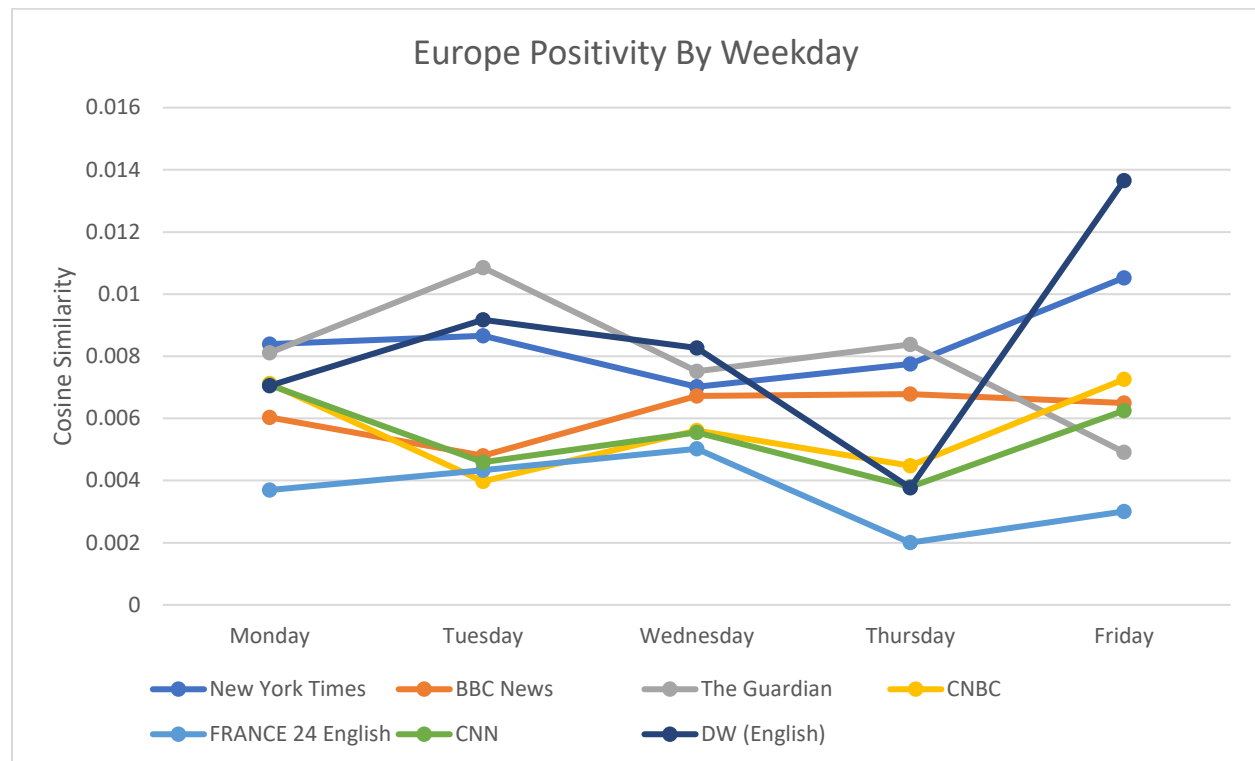
Conclusions: Although our hypothesis was not entirely correct, our results imply that regions with greater coronavirus impact, such as America, may have more articles with greater amounts of **both** optimism and pessimism, as authors may write from strong perspectives on both sides of the situation. In other words, our results could imply that a dire situation would lead to not just articles showing the “bad,” but also articles that are uplifting and discuss how things can be improved.

3. Question: How does optimism and pessimism differ across weekdays?

Hypothesis: Articles written on Fridays will generally have higher optimism than other weekdays.

Experiment: Filter: Region = Europe, repeated for top 7 European publishers, anytime, all titles.

Note that because Google News does not have articles from Saturday or Sunday for certain publications, only workdays are included. Europe was chosen just for fun!



It can be seen that, supporting our hypothesis, Fridays generally have slightly higher optimism than the rest of the weekdays. At the very least, the optimism increases between Thursday and Friday for almost all of the publications, jumping substantially for the publication DW.

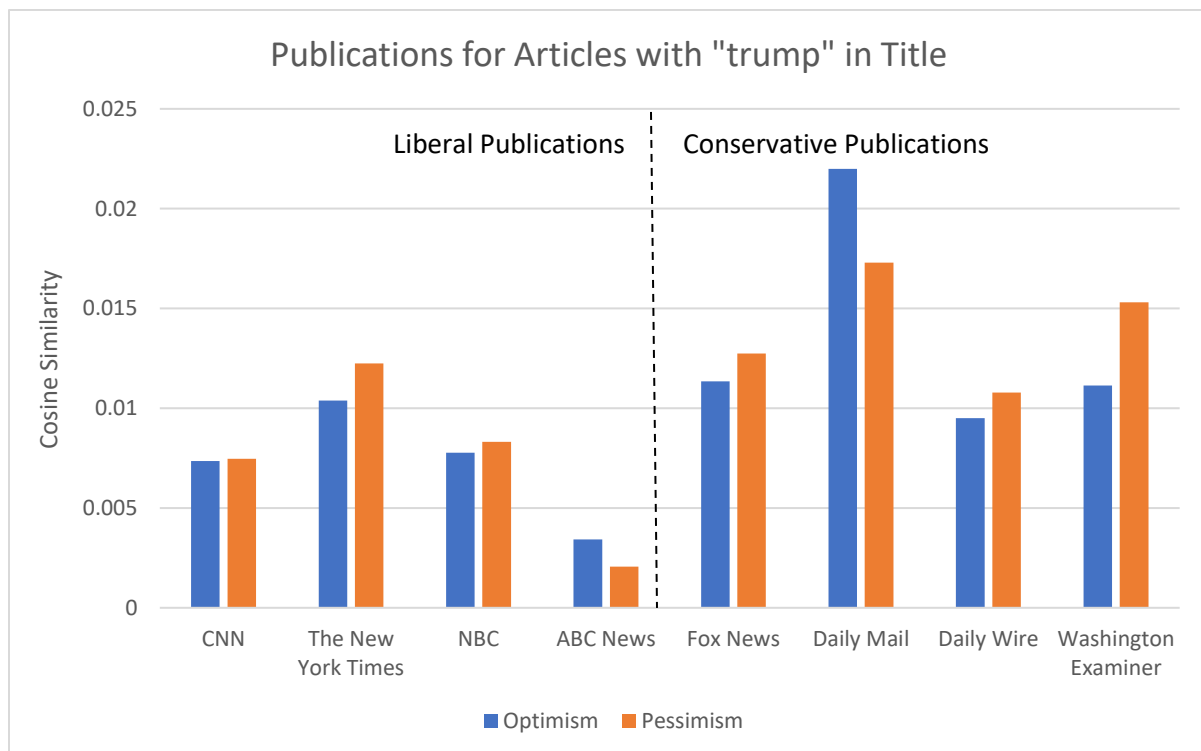
Conclusions: Although Fridays may not always have articles with higher optimism compared to the rest of the week, it appears that in the majority of publications the optimism for articles on Friday is higher than the optimism for articles on Thursday, and is usually among the highest weekdays in optimism.

4. Question: How do different publications view President Trump?

Hypothesis: For articles that contain “Trump” in the title, conservative publishers such as Fox News will have higher optimism than liberal publishers.

Experiment: Filter: Americas, all publishers, anytime, containing “trump” in title
External sites that ranked publications by their political leanings were used to find four liberal publications and four conservative publications.

Result:



The average optimism for liberal publications was 0.00723, while the average optimism for conservative publications was higher, at 0.0135. This result supported my hypothesis, but interestingly the average pessimism for conservative publications was also higher, at 0.014 compared to 0.0075 for liberal publications.

Conclusions: The results support our hypothesis: More conservative articles had greater positivity than the liberal articles. However, we did not predict or expect that the conservative publications would also have much higher pessimism. This implies that perhaps conservative publications use **more passionate and strong vocabulary** as a whole compared to the liberal publications, increasing both their optimism and pessimism ratings.