**Module 5 Assignment — Technique Practice  
Project: Martin Luther King's "I Have A Dream Speech”**

**Prepared by: Sarvesh Thorve**

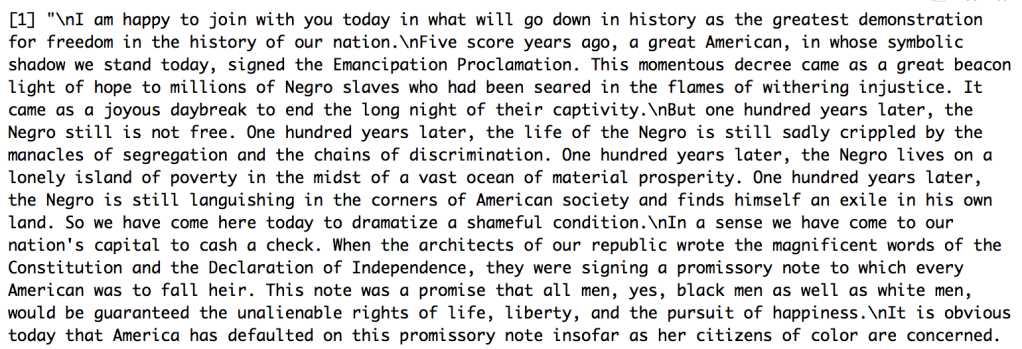
**Email: throve.s@northeastern.edu**

# INTRODUCTION

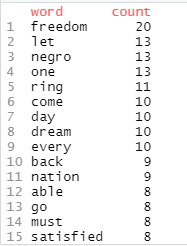
The major purpose of this research is to look at the words used in Martin Luther King's "I Have A Dream Speech" speech in order to better comprehend the emotions that go along with it. We'll do a text sentiment analysis of his speech to see what sentiments are associated with the words he used and what the overall sentiment of his speech is based on the text, whether it's a bad or good message.

# Data cleaning

Using the 'rvest' package, which lets you to read the HTML code of any URL into R, I scrape the spoken text from a webpage. I was able to import the text after locating the HTML node for the spoken text, which in this case was blockquote.

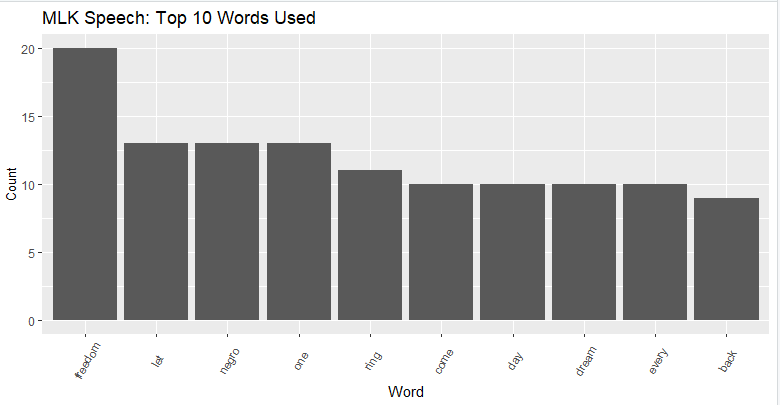


Turning the speech text into a Corpus, which is a collection of texts, is required before performing text sentiment analysis. Following that, we categorize the text's words and generate a data frame with each word and its frequency. The text is then cleaned by deleting punctuation and stop words, such as "a," "the," and "and," which add no context or sense to the text.



# Text Analysis

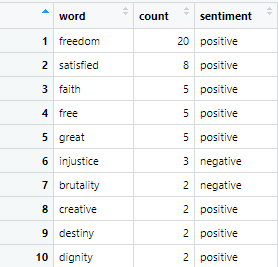
We constructed a bar graph using the new data frame to show the top 10 words used in MLK's speech. The most frequently used word in the speech is "freedom," which appears 20 times in the graph. Other key terms are "negro," "dream," and "country," all of which are important to the speech's meaning. These words will aid us in later comparisons with the text sentiment analysis version.



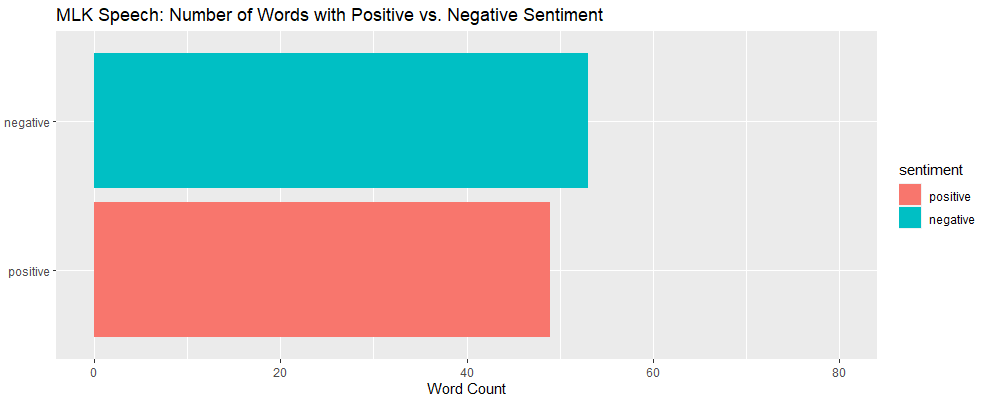
To perform text sentiment analysis, we now use the tidytext package, which offers a data collection named sentiments. The Bing and NRC dictionaries are used with this data set. The Bing lexicon assigns terms to positive and negative emotions, but the NRC dictionary assigns words to positive, negative, trust, sadness, joy, fear, and other emotions.

Now we're attempting to connect these dictionaries' thoughts to the words used in the "I Have a Dream" speech. This entails connecting the data frame containing the speech's words and each of the tidytext data frames containing the lexicon of words and their feelings. The terms that are common between the data frames will be used to connect them.

We acquire a new data frame with columns "word," "count," and "sentiment" after joining the voice words data frame with both sentiment dictionaries. As a result, we examine the text for positive and negative sentiments and compare the number of terms with positive sentiment to those with negative mood.

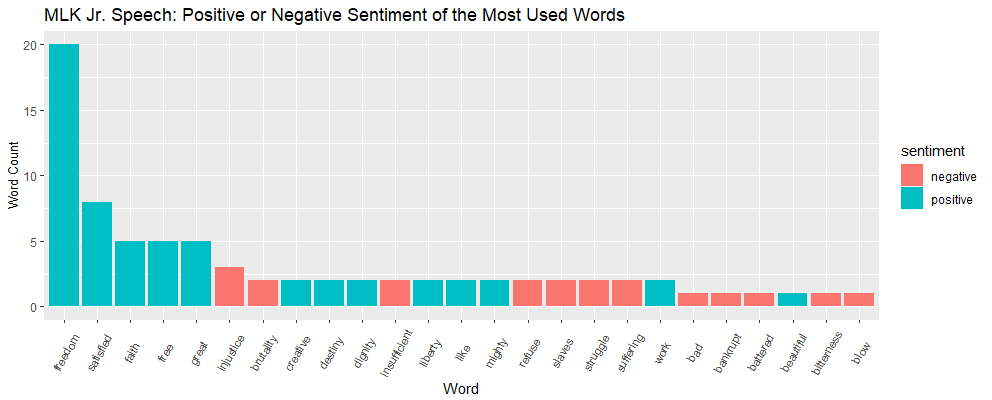


We make a bar graph to show the difference in the quantity of words with positive and negative feelings.



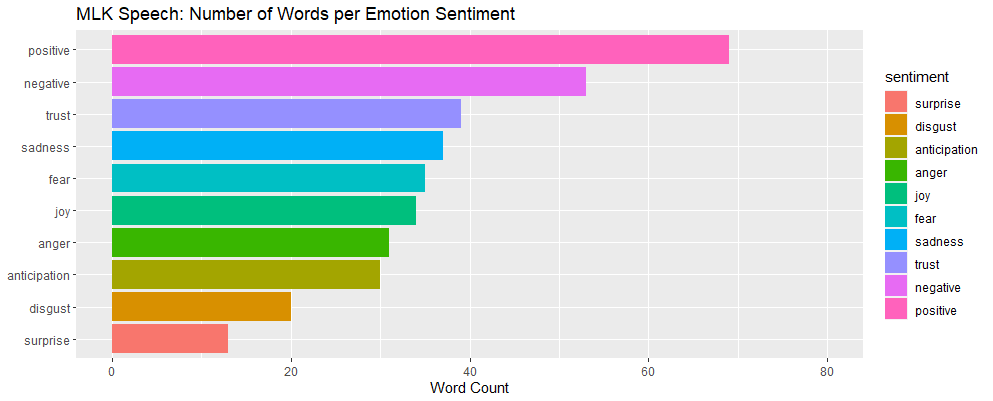
We can see from the graph that negative sentiment is related with more words in the speech than positive sentiment. The difference between the two, is fairly tiny.

Then we figure out which words conveyed different emotions. We examine it using the data frame and visualization. We plotted the top 25 words used for this. While the word count for each sentiment is shown in the bar graph above, we must keep in mind that words appear many times. This is why we'd like to see how people feel about the most commonly used words.

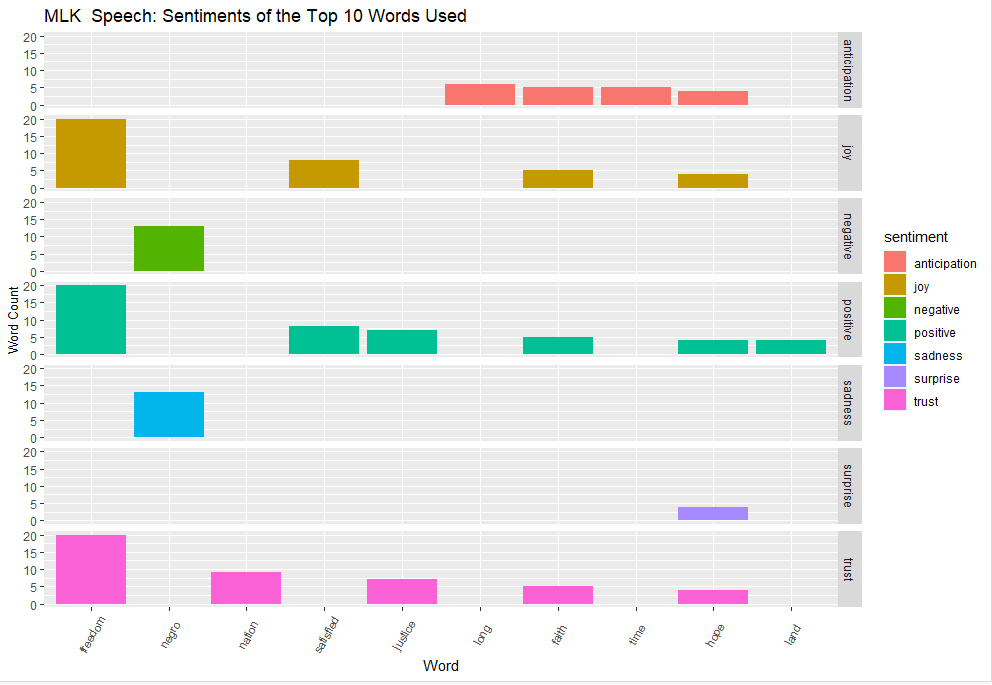


The top five words used all have positive connotations: "freedom," "satisfied," "faith," "free," and "wonderful," as can be seen. This contradicts the previous bar graph, which indicated that there is more negative emotion than positive sentiment. Words with a favorable connotation emerge more frequently in the speech. A wonderful example of why it's critical to examine data from several angles.

Next, we used the NRC dictionary to look at additional particular emotions after covering good and negative sentiments in the speech. Beginning at the top, we can see that there are more positive words than negative terms.



The emotions with the most words, after positive and negative, are trust, sadness, joy, and fear. It's vital to realize that not only do words exist multiple times, but they can also have multiple meanings. As a result, we created a bar graph of the top 10 words used, faceted by sentiment, to see the many emotions associated with each phrase, as well as how the feelings are reflected in the top 10 words used:



The emotions associated with each phrase may now be seen in the above plot. Positive-related sentiments appear 4-5 times in the top 10 terms used, whilst negative-related sentiments appear only twice. We receive a sense of a more positive message being given in the speech, similar to the top terms utilized from the Bing dictionary.

# Conclusion

According to the study described above, this is how these words are represented in a modern dictionary in terms of attitudes. Martin Luther King campaigned for equal treatment of people of different races. While his speech portrays a great image of trust, joy, and expectation for a change to come based on what we've seen in these graphs, the most crucial words still don't carry the meaning that they should. The campaign for racial equality will continue until attitudes toward "negro" alter. This is the message that Martin Luther King hoped to communicate in his iconic "I Have a Dream" speech.

# References

* I Have A Dream Speech (TEXT). (2012, January` 17). Retrieved from 'https://www.huffpost.com/entry/i-have-a-dream-speech-text\_n\_809993'
* Wongvorachan, T. (2020, August 14). Text mining and word cloud fundamentals. Retrieved from <https://rstudio-pubs-static.s3.amazonaws.com/667853_71fce9a2f903499487b89d0d7a883a6c.html>