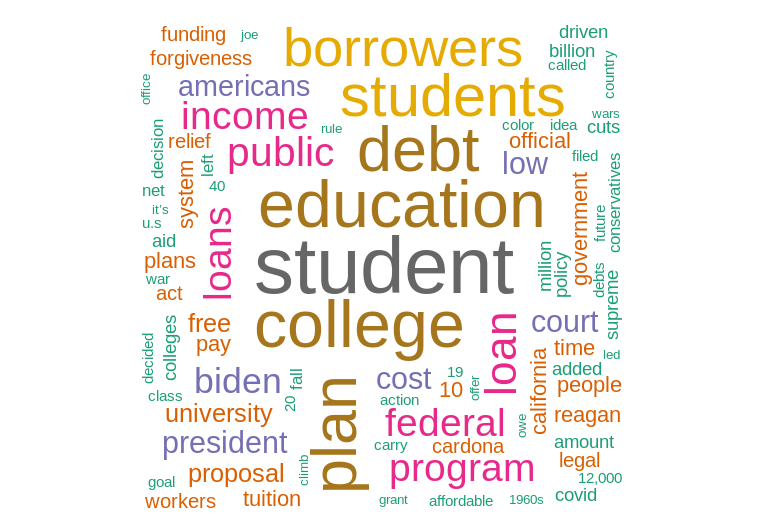
Unit 7 Project, MM325, Laurence Burden

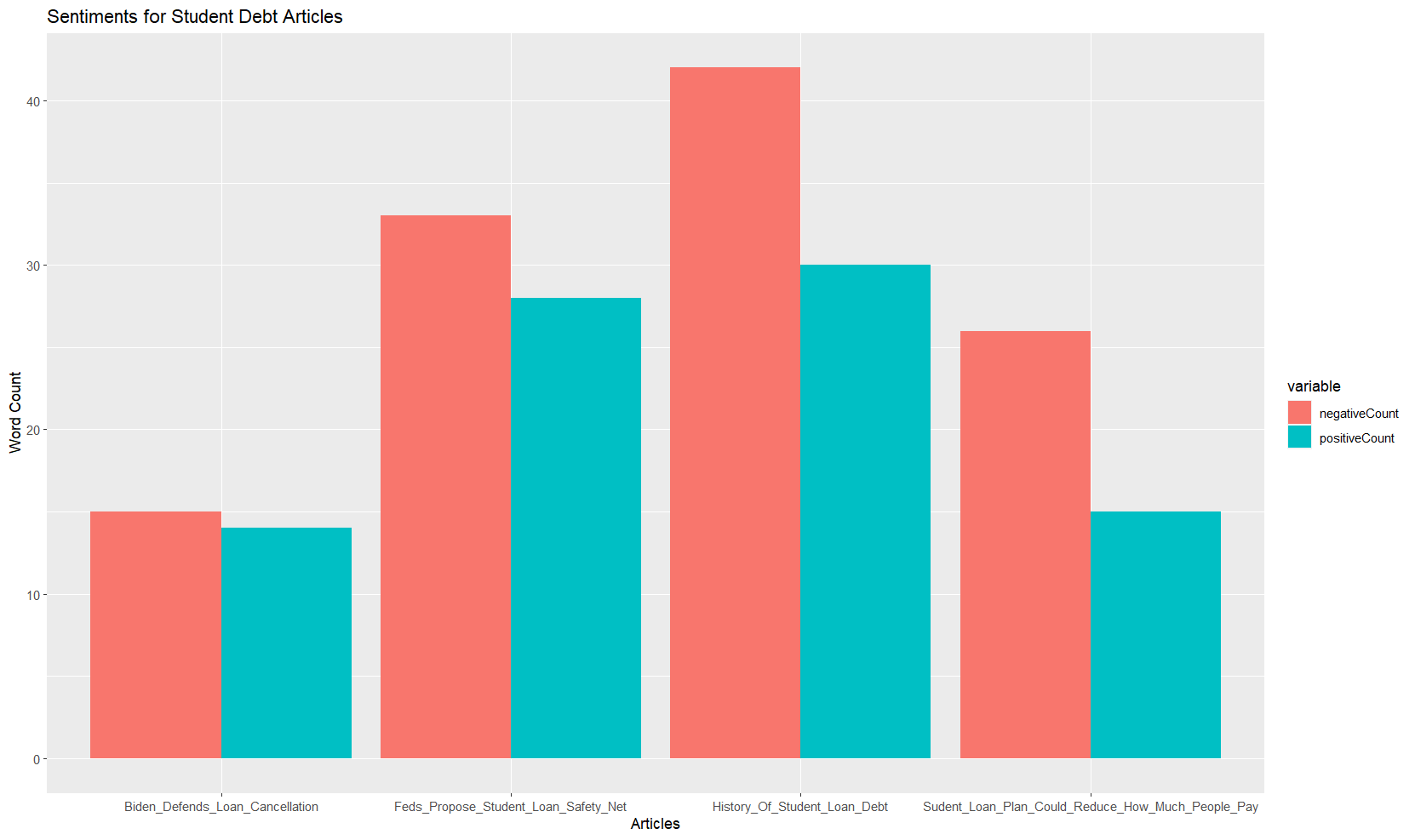
The text analysis for this project was completed using a variety of R packages such as tidytext, reshape2, and wordcloud to reformat the text file into binned groups of words that were then used to form the graphical charts. The text was first modified to a delimitated file using a colon (:) as the delimiter between article titles and article text. Then the text was stripped of “stop words” such as “the”, “of”, “to”, etc. Tidytext was used to make this process easier.

The word cloud was created by selecting a seed value to keep the output consistent, creating a count of each word, and then using the wordcloud function with specified parameters to create the output. The sentiment bar graph used the bing lexicon to separate each article into positive and negative words. Two counts were then created for each of the four articles, one for the positive words, and one for the negative words. The data was then “melted” to associate the counts with each article title. Finally, the data was plotted using ggplot2 and its geometry bar function.

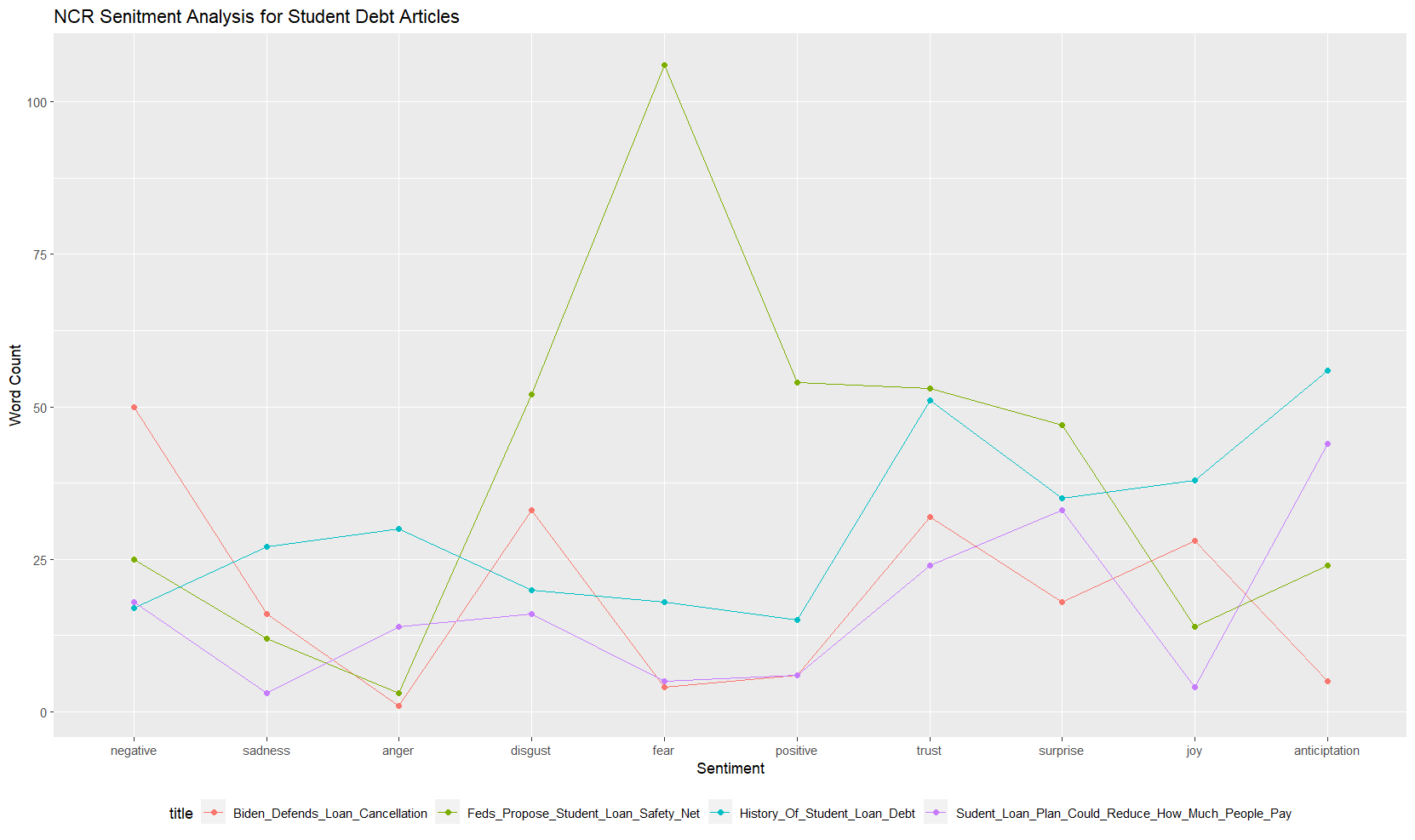
The sentiment line graph was created with the NRC lexicon. It places words into bins of sentiments such as “trust”, “fear”, etc. Each article is then looped through and then an inner loop runs a sentiment analysis for each sentiment bin. The final vector of word counts is then separated into the proper indexes for use in creating a data frame associating the sentiment word count with the proper article title. This section would need a bit of work to be “production ready”, but works for a quick script. The data is once again melted and then used with ggplot2 and its geometry line function.



A word cloud showing the most used words within the four chosen articles. The larger the word, the more times it appears in the selected text.



A bar graph showing the difference between positive and negative words within each of the four articles.



A line graph showing the word count of each sentiment bin as described in the NRC lexicon. Each line represents one article.

References

Sentiment Analysis: Automatically Detecting Valence, Emotions, and Other Affectual States rom ext. Saif M. Mohammad, arXiv:2005.11882, Jan 2021.

Haslett, C. (2023, January 10). *Biden's student loan plan could reduce how much people pay and for how long*. ABC News. Retrieved February 21, 2023, from https://abcnews.go.com/US/bidens-student-loan-plan-reduce-people-pay-long/story?id=96327219

Binkley, Collin (2023, January 10). *Feds propose 'student loan safety net' alongside forgiveness*. ABC News. Retrieved February 21, 2023, from https://abcnews.go.com/Business/wireStory/feds-propose-student-loan-safety-net-alongside-forgiveness-96331710

Baker, P. (2023, January 5). *Biden administration defends student loan cancellation at Supreme Court*. The New York Times. Retrieved February 21, 2023, from https://www.nytimes.com/2023/01/04/us/politics/biden-student-loans-supreme-court.html

Love, D. A. (2022, December 29). *Perspective | America's student loan crisis stems from a war on education as a public good*. The Washington Post. Retrieved February 21, 2023, from https://www.washingtonpost.com/made-by-history/2022/12/29/history-student-loan-debt/