

Final Project 729B

Isabelle Gibson

This data is available on my public page on github, available for people to use.

My final project aims to analyze media framing and sentiment in coverage of BLM protest. This project does the following:

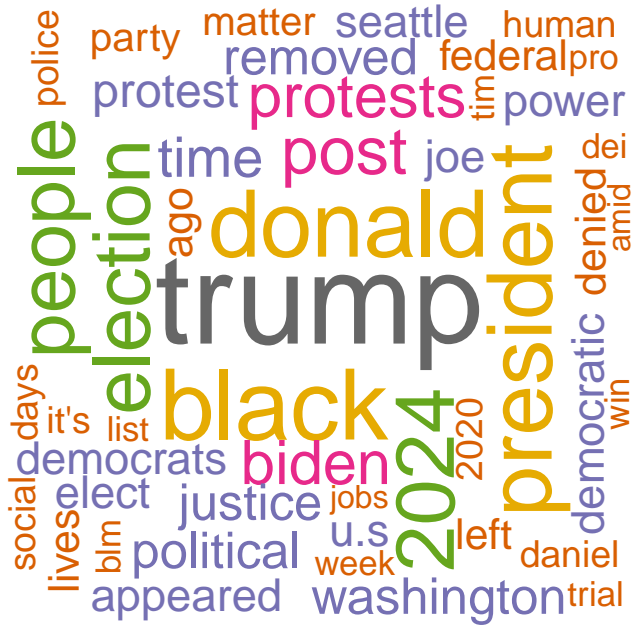
1. Identify common words/themes across a variety of articles gathered from NewsAPI.
2. Visualize the words/themes with word clouds and bar plots
3. Conduct a sentiment analysis to provide insights into the tone of the articles overall, negative or positive.

With my account with NewsAPI I gather multiple articles related to “Black Lives Matter protests”. To prepare the text data for analysis, I combine the title, description, and content fields from the articles into a single column called text for analysis.

- I then tokenize the combined text column into individual words using the `untest_tokens` function from tidy text and create `blm_words`.
- I remove all the stop words and irrelevant words like `http`, `amp`, etc.

Here I begin to build a word count to calculate the frequencies across all of the articles. Then I generate a word cloud that visualizes the most frequent words.

- I also group the word frequencies by the news source as well. I looped through each news source to generate individual word clouds, visualizing how different outlets emphasize different terms and words.



[1] "Word Cloud for Naples Daily News"



[1] "Word Cloud for BBC News"



A word cloud for Scientific American. The words are arranged in a vertical stack. The word 'plinth' is at the top in yellow. Below it is 'aid' in pink. Then 'forty' in pink. Then 'laque' in grey. Then 'colston' in pink. Then 'band' in pink. Then 'ng' in pink. The word 'plinth' is the largest, followed by 'laque'.

[1] "Word Cloud for Scientific American"



A word cloud for The Atlantic. The words are arranged in a vertical stack. The word 'generations' is at the top in pink. Below it is 'beatrice' in pink. Then 'project' in grey. Then 'carolyn' in pink. The word 'generations' is the largest, followed by 'project'.

[1] "Word Cloud for The Atlantic"



A word cloud for The Intercept. The words are arranged in a vertical stack. The word 'lazar' is at the top in grey. Below it is 'americ' in grey. Then 'alike' in grey. Then 'ebritie' in grey. The word 'lazar' is the largest, followed by 'americ'.

[1] "Word Cloud for The Intercept"

more to party
cost 250 blm reve
appeared
pro post an a fe

[1] "Word Cloud for Skidmore.edu"

emsp. rise alarming
identity olitica
ritarianism
function onalist

[1] "Word Cloud for Vox"

instituti hardcheap
academic
scientist postdo
busy faculty
competing

[1] "Word Cloud for NBC News"



[1] "Word Cloud for The Root"



[1] "Word Cloud for RT"



[1] "Word Cloud for Hyperallergic"



[1] "Word Cloud for IndieWire"

worth

[1] "Word Cloud for Washington Free Beacon"

black
prominent
run activist jury
daniel

[1] "Word Cloud for The Cut"

accessories
brothers
res aurora allies
luxury

[1] "Word Cloud for The Daily Caller"



[1] "Word Cloud for TheSocietypages.org"



[1] "Word Cloud for Infosecurity Magazine"



[1] "Word Cloud for USA Today"



[1] "Word Cloud for [Removed]"

move

[1] "Word Cloud for Thenation.com"

novembe
powerb
appear
bail
life
choosing
bosa
field
fear
post
amid
policy
orkers
bosses
2024

[1] "Word Cloud for Adweek"

ilgree
buc
wing
boots
brand
irget

[1] "Word Cloud for Truthout"



[1] "Word Cloud for Naturalnews.com"



[1] "Word Cloud for Yahoo Entertainment"



[1] "Word Cloud for Legalinsurrection.com"



A word cloud for menshealth.com. The words are arranged in a vertical stack. 'black' and 'friday' are the largest words in the center. To the right, 'boston' is written vertically. Above 'black', 'israel' is written in yellow, and 'ant' is written in yellow. Above 'friday', 'cities' is written in pink and 'disrupt' is written in pink.

[1] "Word Cloud for menshealth.com"



A word cloud for Cosmopolitan.com. The words are arranged in a vertical stack. 'aldis' and 'edge' are the largest words. 'aldis' is on the left and 'edge' is on the right. They are both written in a dark grey font.

[1] "Word Cloud for Cosmopolitan.com"



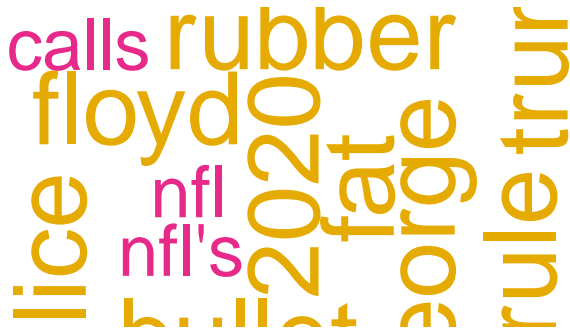
A word cloud for CounterPunch. The words are arranged in a vertical stack. 'election' is the largest word in the center. To the right, 'ovement' is written vertically. Above 'election', 'post' is written in pink. Below 'election', 'lue' is written vertically in pink.

[1] "Word Cloud for CounterPunch"



A word cloud for CounterPunch. The words are arranged in a vertical stack. 'protest' is the largest word in the center. Below it, 'seattle' is written. To the left, 'preside' is written vertically. Above 'protest', '2024' is written. To the right, 'tru' is written vertically. Below 'seattle', 'streets' is written. To the right, 'rarely' is written vertically. Below 'rarely', 'activists' is written. To the right, 'elect' is written vertically. Below 'streets', 'bet' is written. To the right, 'left' is written. Below 'bet', 'printed' is written. To the right, 'c' is written. Above 'protest', 'social' is written. To the right, 'c' is written.

[1] "Word Cloud for Newsweek"



A word cloud for Newsweek featuring words in yellow and pink. The words are arranged in a dense, overlapping manner. The most prominent words are 'calls', 'rubber', 'floyd', '2020', 'fat', 'forge', 'rule', 'trur', 'ice', 'nfl', 'nfl's', and 'lice'.

[1] "Word Cloud for Colorado Public Radio"



A word cloud for Colorado Public Radio featuring words in yellow and pink. The words are arranged in a dense, overlapping manner. The most prominent words are 'history', 'colorado', and 'art'.

[1] "Word Cloud for Breitbart News"



A word cloud for Breitbart News featuring words in yellow and pink. The words are arranged in a dense, overlapping manner. The most prominent words are 'time', 'mocr', 'position', and 'pollak'.

[1] "Word Cloud for Democracy Now!"

message
nehisi
coates
okwar

[1] "Word Cloud for The Week Magazine"

walm
dei
initiatives

[1] "Word Cloud for Lithub.com"

black
ago
active

[1] "Word Cloud for The Conversation Africa"

A word cloud featuring names of political figures and locations. The words are arranged in a vertical, overlapping manner. The most prominent words are 'tulsi', 'joe', 'dark', 'web', 'rfk', 'delon', and 'e'.

[1] "Word Cloud for Grist"

A word cloud with a single, large, vertical word: 'york'.

[1] "Word Cloud for Daily Signal"

A word cloud featuring names and terms. The words are arranged in a vertical, overlapping manner. The most prominent words are 'robert', 'kennedy', 'establishment', 'department', 'elected', 'can't', 'serves', 'it's', 'and', 'the', 'old', 'and', 'the', 'old', 'and', 'the', 'old'.

[1] "Word Cloud for Teen Vogue"

A word cloud with a single, large, vertical word: 'left'.

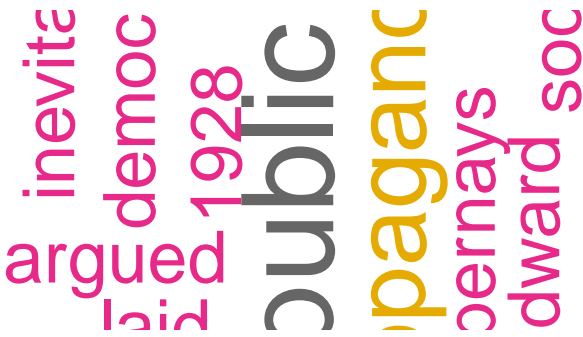
[1] "Word Cloud for New York Post"



[1] "Word Cloud for Celebitchy.com"



[1] "Word Cloud for Thecuriousbrain.com"



[1] "Word Cloud for Vulture"

keep
attack

[1] "Word Cloud for Juancole.com"

civic
attack
american
burn

[1] "Word Cloud for The Federalist"

political
claims
dirt
doj
joe
biden
son
casper
onents

[1] "Word Cloud for In These Times"

trumps
despair

[1] "Word Cloud for Advocate.com"



[1] "Word Cloud for Wnd.com"



[1] "Word Cloud for Boston Herald"



[1] "Word Cloud for Worksinprogress.co"

century
future

[1] "Word Cloud for Dianeravitch.net"

approve
educati
trump
campaign
artmer
u.s.
202
det

[1] "Word Cloud for Crossway.org"

justice
denied
love
delayed
heard

[1] "Word Cloud for Whyevolutionistrue.com"

christmas
celebrated
coynezaa
birthday
exte
day
.

[1] "Word Cloud for The New York Review of Books"

carolin club anuja
1827 slezkine
vuri sq action

[1] "Word Cloud for Slate Magazine"

implement
donald ate

[1] "Word Cloud for Freerepublic.com"

viral movement
sk comments
announcer
accidentally win
49ers ryan
2024 posted
tim car cartoon

[1] "Word Cloud for Chinafile.com"

china

[1] "Word Cloud for Substack.com"

feels
2016
critical
2024

[1] "Word Cloud for Americanthinker.com"

worse
deeper
denial
friends
left
deep
world

[1] "Word Cloud for Fortune"

o.j.
npson

[1] "Word Cloud for NDTV News"

attac
9186
2016
ip
ep

[1] "Word Cloud for HuffPost"

image
disco
1978
classic
ump

[1] "Word Cloud for Voxeurop.eu"



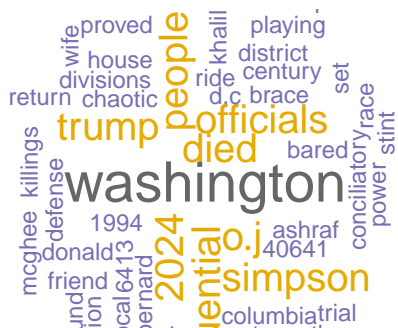
[1] "Word Cloud for Atlas Obscura"



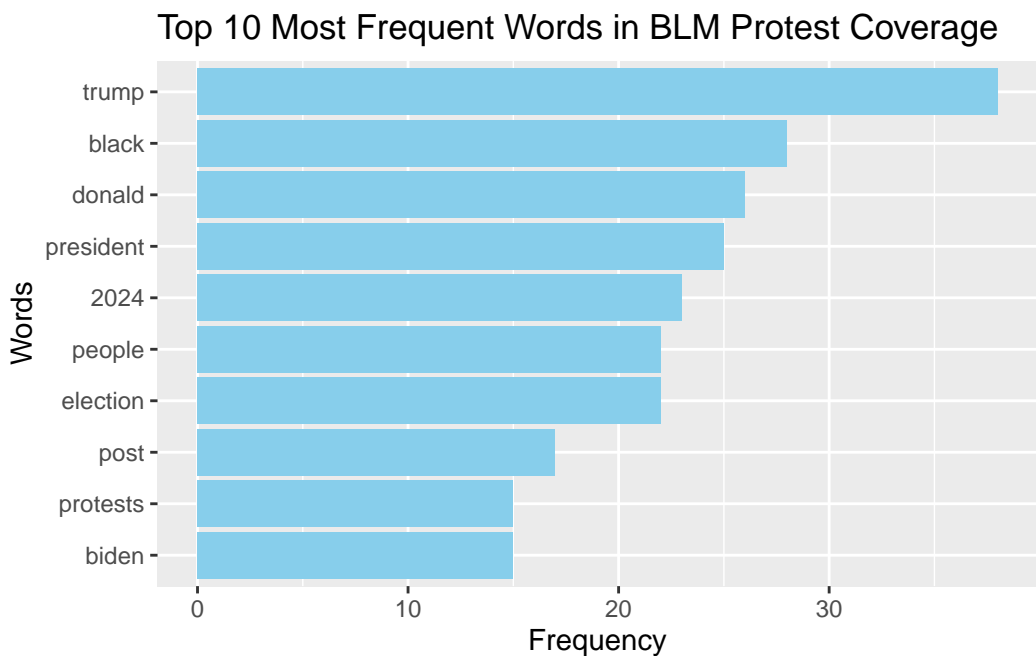
[1] "Word Cloud for USA Today"



[1] "Word Cloud for BBC News"

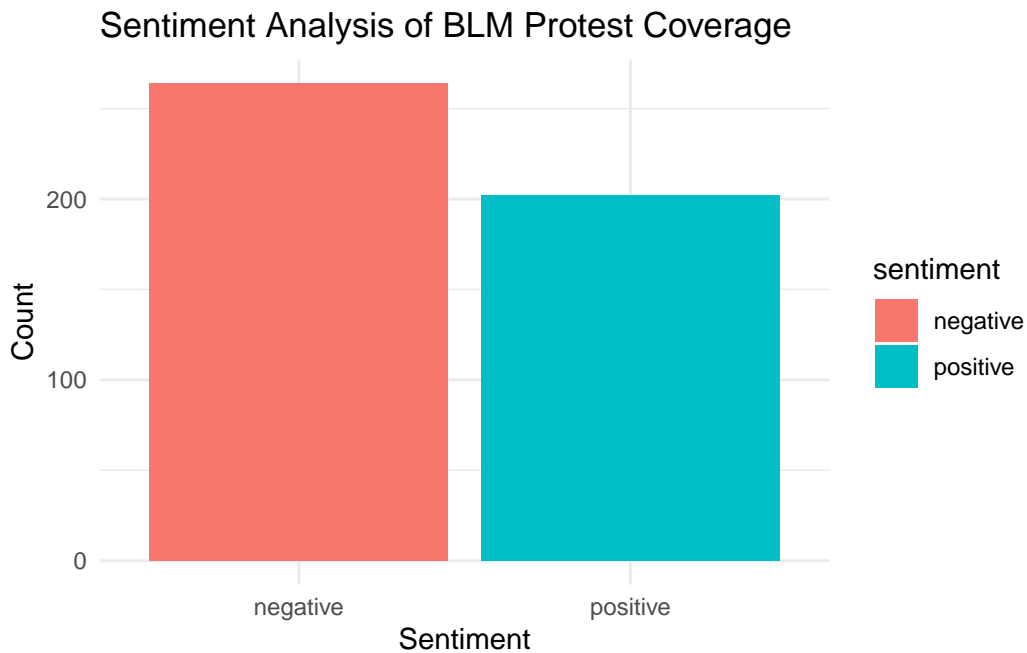


Here is where I created a Bar Plot to visualize the top words frequencies



Lastly, I conduct a sentiment analysis. Using Bing Lexicon, I match words from the tokenized blm_words with the bing sentiment lexicon, which classifies words as positive or negative. It counts the number of positive and negative words in the dataset.

- The bar chart shows the distribution of positive and negative sentiment words across the entire corpus.



Sentiment Analysis of BLM Coverage

- The bar chart in the first image displays the distribution of positive and negative sentiment words.

Key findings:

- The sentiment analysis reveals that negative words dominate over positive words.
- Negative sentiment count: ~250
- Positive sentiment count: ~200
- Implications:

The higher prevalence of negative sentiment aligns with research showing that media often emphasizes confrontational, disruptive, or violent aspects of protests.

- This framing can undermine public support for the BLM movement by fostering negative associations, even when the protests' goals focus on justice and equality.
- However, the substantial number of positive words suggests that some media outlets do highlight themes like justice, equality, and human rights.

Media Framing and Political Context: The findings suggest that BLM protests are framed not only as social justice events but also as politically charged phenomena, often tied to electoral cycles and political leaders.

- Sentiment and Public Perception: The prevalence of negative sentiment highlights the media's role in shaping public attitudes toward protests. Negative frames may contribute to polarization or delegitimization of the movement, while positive frames can foster empathy and support.
- Themes of Polarization: The focus on political figures (Trump, Biden) and elections suggests that BLM coverage may reflect and contribute to partisan divides in how protests are perceived.