



What comes to mind?



Abortion as a Public Issue: **A Sentiment Analysis** of Social Media versus Popular Media

Team Zebra: Bhavya, Wenhao, Xiaowei & Xiaochen

Objective + Rationale

The objective of our project is to **study the difference in views about abortion, on social media versus in popular media between January 2022 and January 2023.**

Social science research suggests that views about abortion can differ on social media and in popular media, with **social media tending to amplify extreme views** and **popular media presenting a more nuanced view** of the issue. It's important to be aware of these differences and to seek out a range of perspectives when engaging in discussions about volatile public issues, such as abortion.

Our project aims to help understand and deconstruct these differences through data visualizations, observations from data, and data modeling.

Social Science Perspective + Recent Developments

The question of whether abortion should be legal or not has led to the creation of two main schools of thought: **pro-choice** and **pro-life**.

However, just recently, the Supreme Court of the United States (SCOTUS) officially **overturned the Roe vs. Wade case on June 24, 2022**.

The Roe vs. Wade case first appeared in court in 1973 and was a **landmark decision** that stated that the Constitution of the United States granted **citizens the right to an abortion**.

However, with the recent overturning, **abortion rights are now decided by each individual state (regarding whether it is a right or not)**.

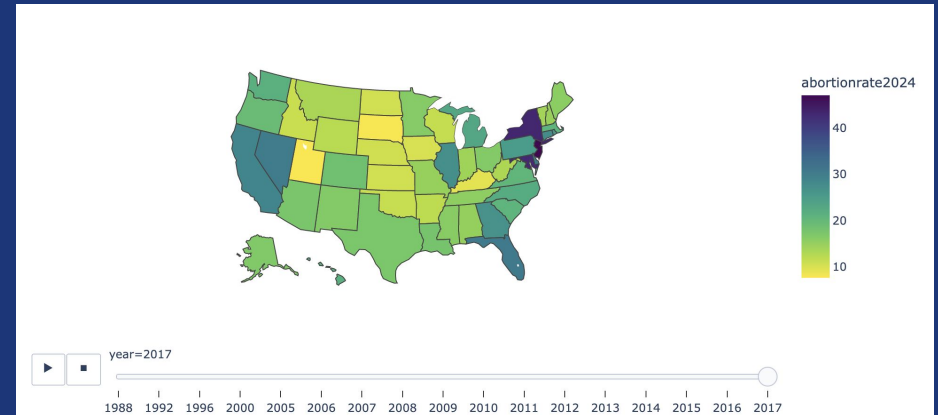
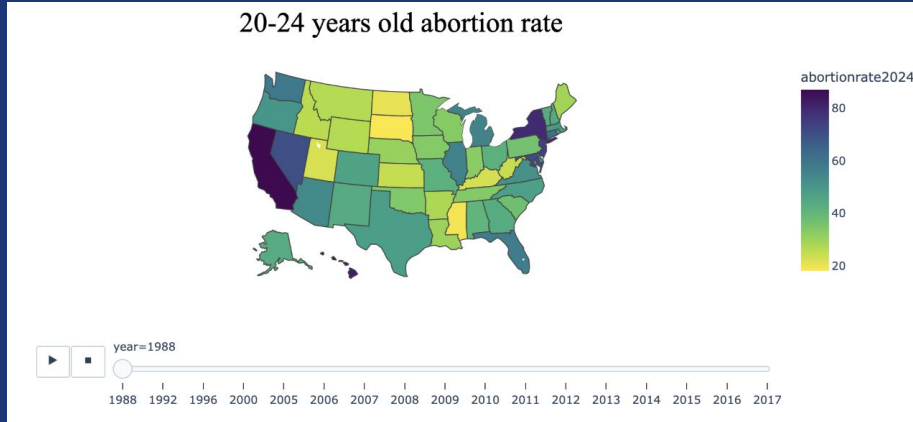


Literature Review

background, questions of interest to social scientists, recent developments, key concepts

- **Guttmacher Institute (2020)**
29 states have policies that are hostile to abortion rights.
- **Pew Research Centre (2019)**
61% of Democrats believe that abortion should be legal in all or most cases, while only 27% of Republicans hold this view.
- **Guttmacher Institute (2017)**
Abortion rate in the United States declined by 7% between 2014 and 2017.

Preface: Historical Changes in Abortion Rates



Research Questions...

How are people reacting to the topic of abortion online, particularly on Twitter?

What are the views being expressed through popular media, such as newspapers like the *New York Times*?

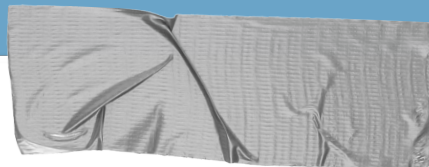
Is there a difference between these observations after the June 2022 ruling?

Based on social science research in the past, we **hypothesize** that social media tends to amplify extreme views on abortion. On the other hand, popular media, such as newspapers like *The New York Times*, tend to present a more nuanced view of abortion.

We expected a similar result on carrying out a **sentiment analysis** on both sources (more details on this follow). While media outlets may have their own editorial biases, they often present a range of perspectives on the issue of abortion and the various political and social factors that contribute to it.

Data! Where we went looking for answers...

1. Pregnancies, Births and Abortions in the United States: National and State Trends (~1000 data points, publicly available dataset by Guttmacher Institute, 1988-2017)
2. Tweets related to abortion views (~8500 Tweets, makes use of Python `snsrape` library, keyword: 'abortion', 4 major cities)
3. *New York Times* articles related to abortion (~500 articles, scraped using *New York Times* official API, and Python web-scraping libraries such as `requests` and `BeautifulSoup`, same keyword)



Data Collection

Here's some learnings from a few of our experiences...



Logistics

Methods and methodologies:
web-scraping, APIs, existing datasets



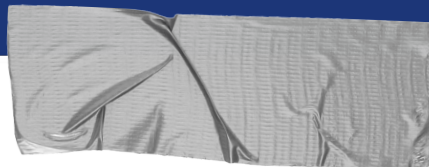
Maintaining internal validity

Ensuring methods are statistically
significant, allow room for errors



Maintaining external validity

Ensuring methods are clear,
synchronous, and replicable



Data Cleaning

Here's some learnings from a few of our experiences...

- **Data structures and maintaining them**
- **Leveraging functions and functionalities** in Python
- **Importance of pre-processing**
- **Cleaning and formatting real-world data**
- **Ensuring uniformity across sources**



Workflow

Details of our project workflow

- **Historical state-wise analysis as preface**
- **Sentiment analysis of Tweets and NYT articles**
- **Time-based and geography-based synchronous analysis of data**
- **Exploratory data analysis**



Detailed Workflow: Sentiment Analysis using RoBERTa

- We are using a pre-trained machine learning model for our analysis, called the **RoBERTa-Base Model**. This *transformer-based model* was trained on roughly 124 million Tweets from January 2018 to December 2021 and is specifically fine tuned for sentiment analysis involving media sources (like social media and newspapers) which makes it ideal for our project.
- This model comes with its own *classification labels and tokens*, which classify any piece of short text into Positive, Negative and Neutral sentiments. As seen previously, it is a very **nuanced and sophisticated** classification model which is being extensively used for sentiment analysis by researchers.

Let's begin to talk more about our analysis + findings by looking at some sample data and descriptive statistics...



Here's what some of the classification of Tweets looked like...

Positive: “Absolutely! It's also a men's issue. I am thankful that my partner sat next to me and held my hand during my #abortion, and I recognize the experience impacted both of us. I'm so grateful for @user @user #NY for being there for us. #abortionishealthcare”

Neutral: “Let me be clear: pro life excludes NO ONE. However if the woman's life is in danger, our first answer is to do what we can to save both lives and provide best care. Abortion should NEVER be an answer in such cases because it takes away a life.”

Negative: “Save the Babies. Also please tell young girls and women who are not sure if they should have their abortion and who do not know the horrific Truth about Abortion please Let them know the Truth.”



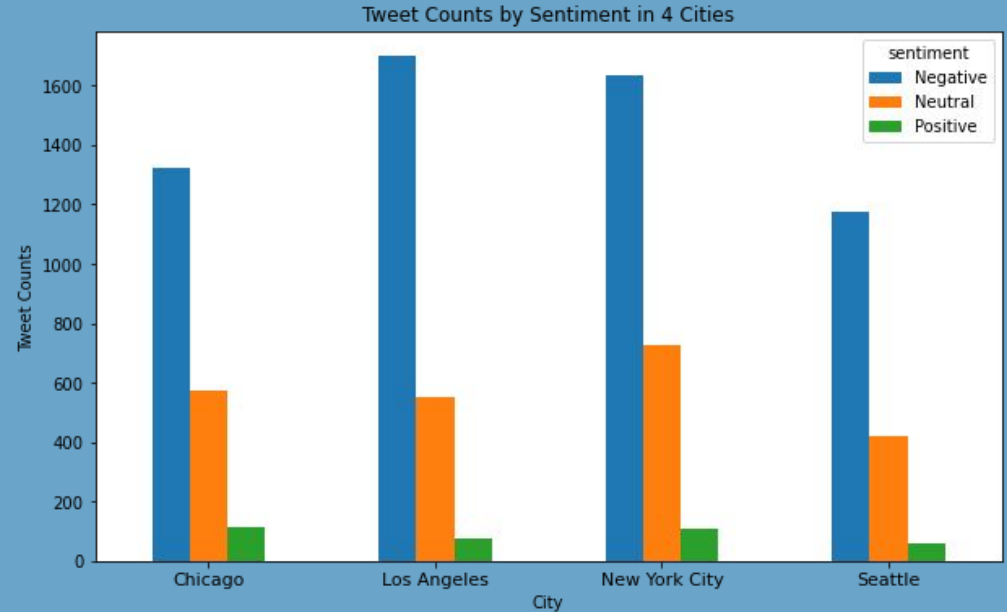
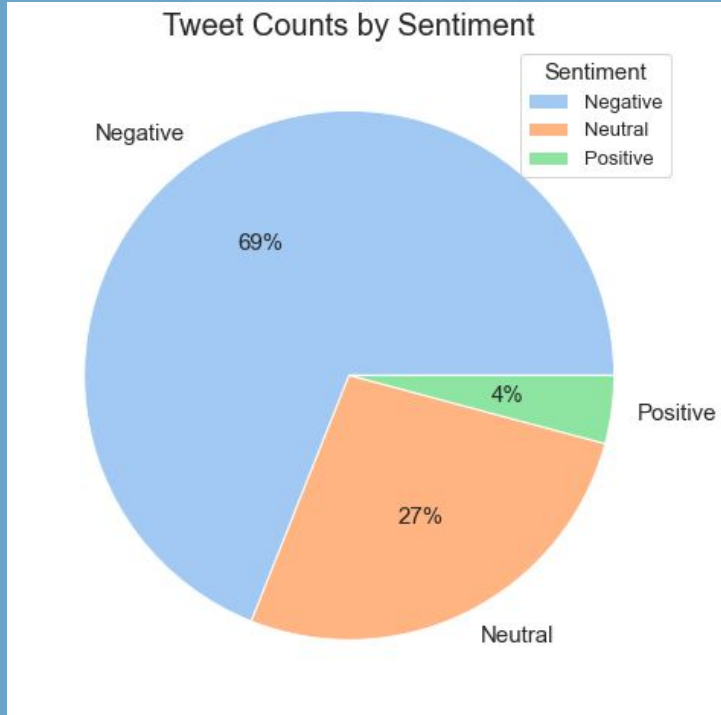
And here's what some of the classification of NYT article titles looked like...

Positive: “Weekend of Abortion Protests Brings Out Supporters”

Neutral: “The Women Who Had Abortions Before Roe v. Wade”

Negative: “Thousands Protest End of Constitutional Right to Abortion”

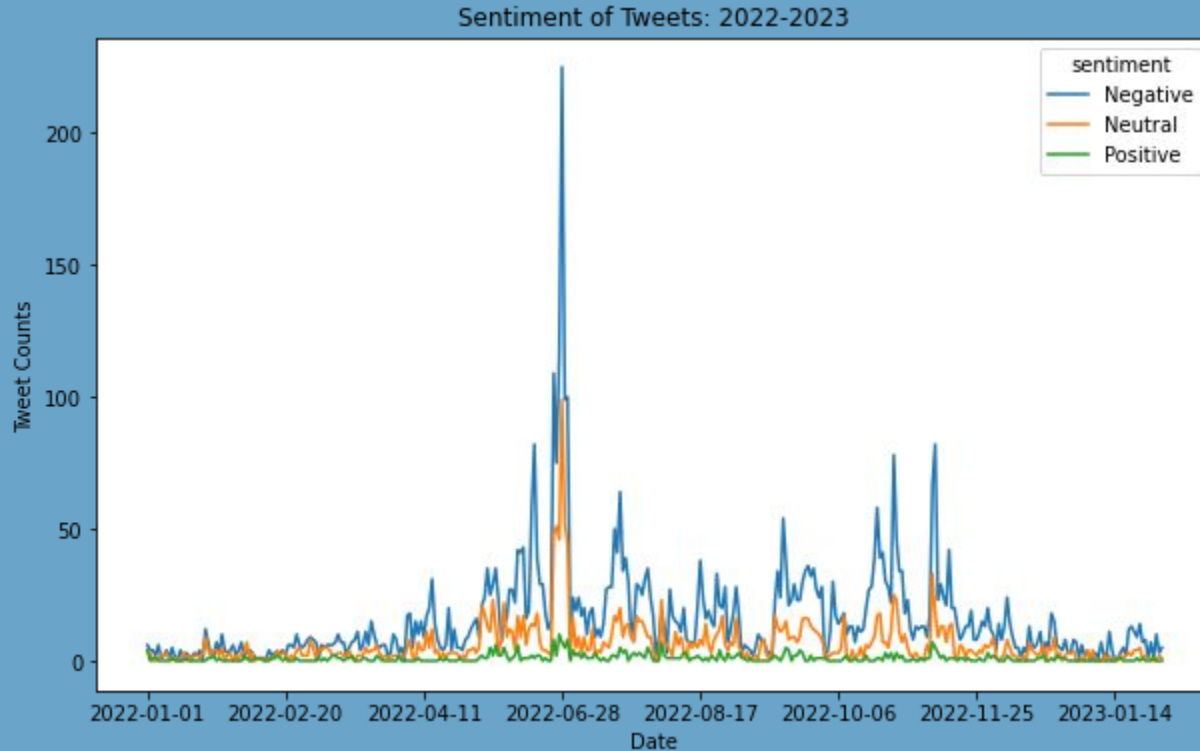
Tweet Counts by Sentiment and by City



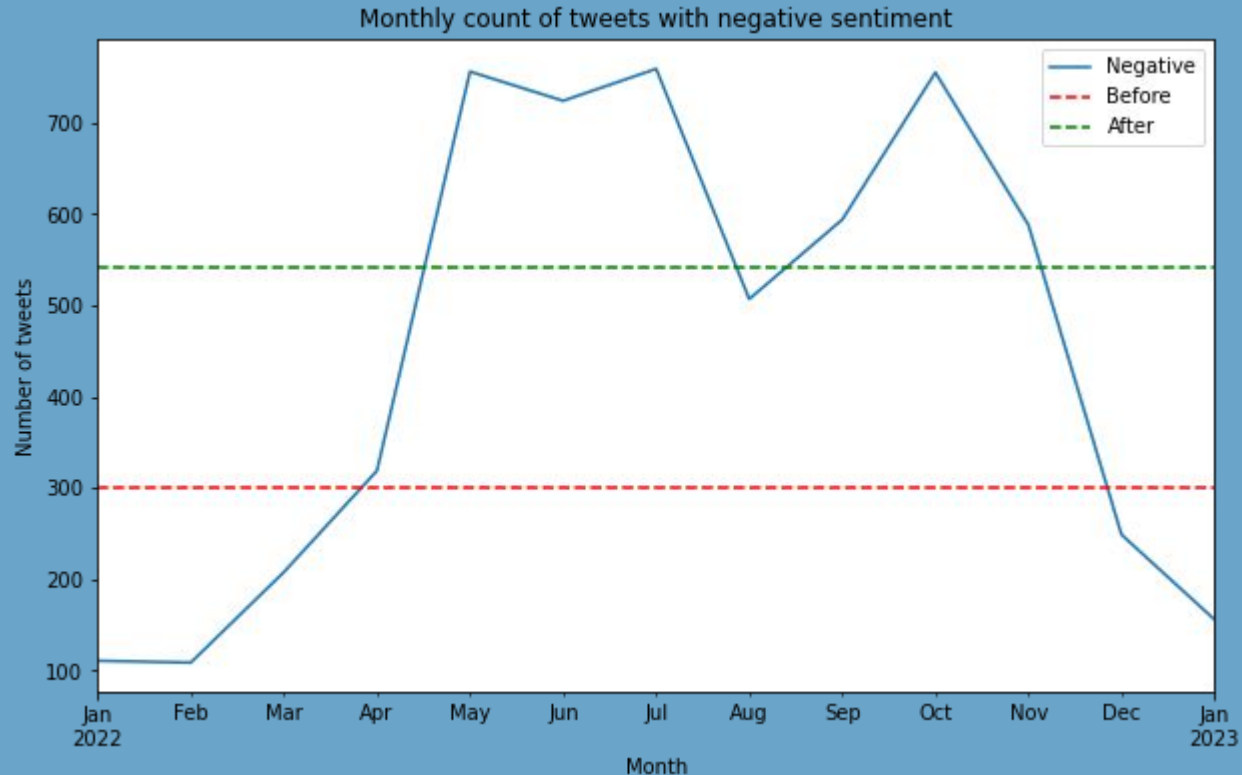
A snapshot...

Unnamed: 0	Date	Tweet	latitude	longitude	city	tweet_processed	sentiment
0	0	2022-01-31 Nazis, banned books, suppressed voting rights,...	40.7128	-74.006	New York City	Nazis, banned books, suppressed voting rights,...	Negative
1	1	2022-01-31 @SenatorLankford In case you haven't noticed a...	40.7128	-74.006	New York City	@user In case you haven't noticed abortion is ...	Neutral
2	2	2022-01-31 @SenatorLankford So you support pushing aborti...	40.7128	-74.006	New York City	@user So you support pushing abortion undergro...	Negative
3	3	2022-01-31 @RayRiosy @Gdad1 @Itwlauren @AngelMHart417 @Ji...	40.7128	-74.006	New York City	@user @user @user @user @user @user @user @use...	Neutral
4	4	2022-01-31 @alicee_pll @miaana_14 People who support the ...	40.7128	-74.006	New York City	@user @user People who support the pro abortio...	Negative

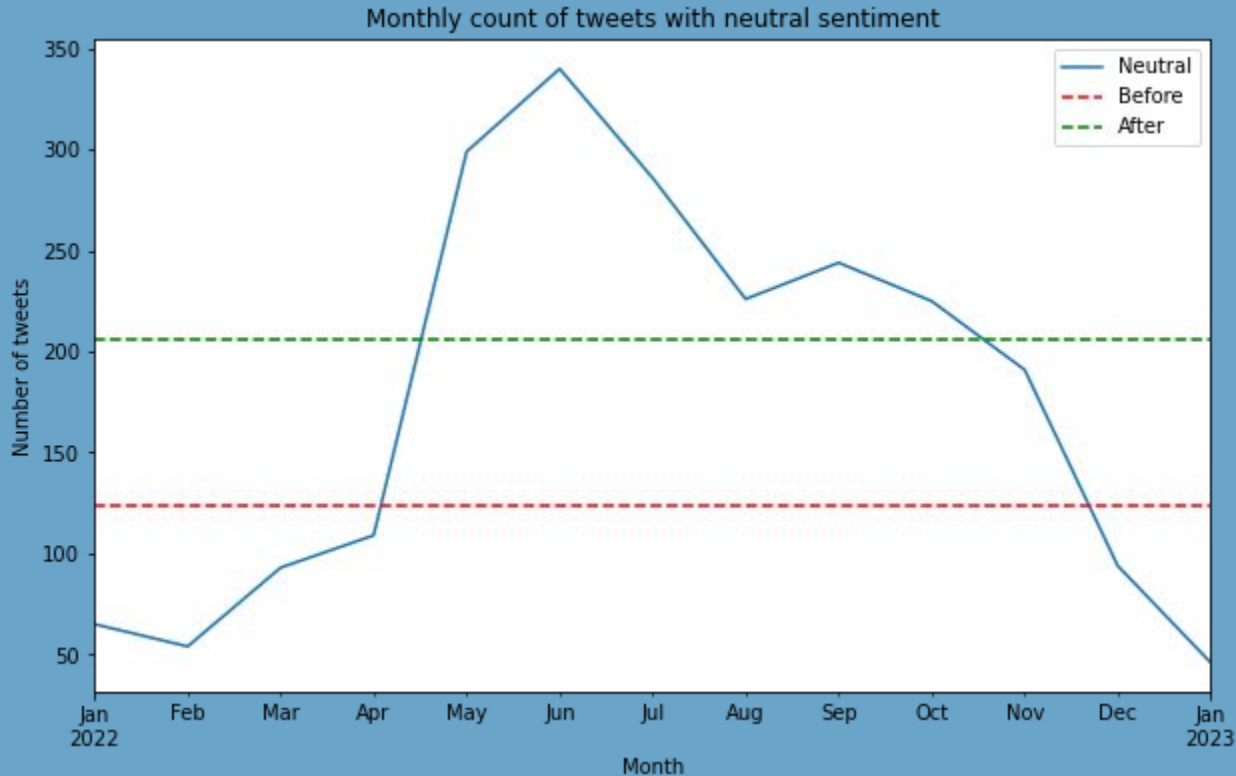
Progression of Sentiments in Tweets Over Time



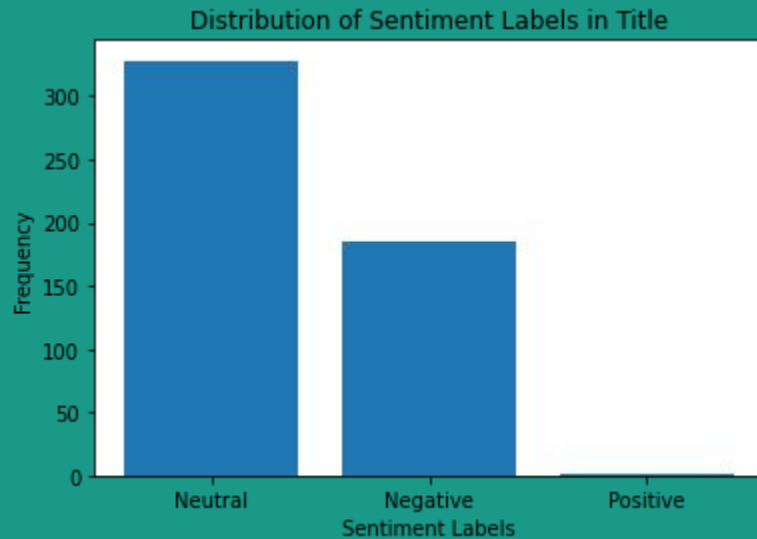
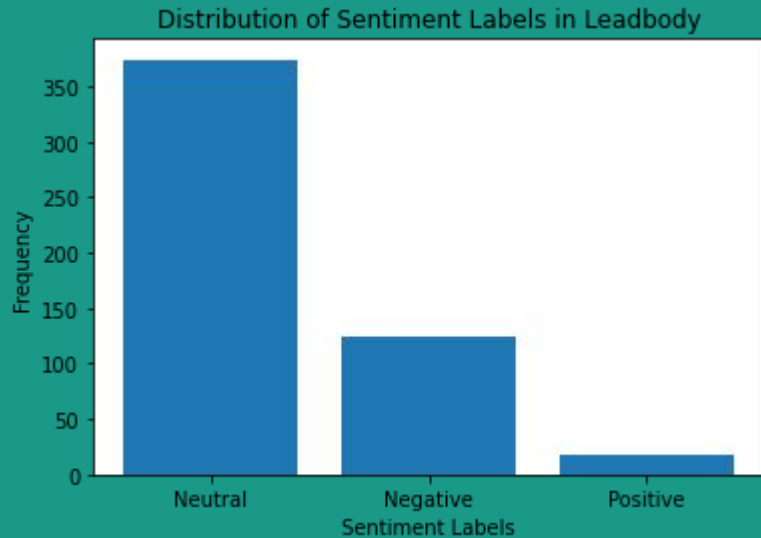
Progression of Monthly Count of Negative Tweets



Progression of Monthly Count of Neutral Tweets



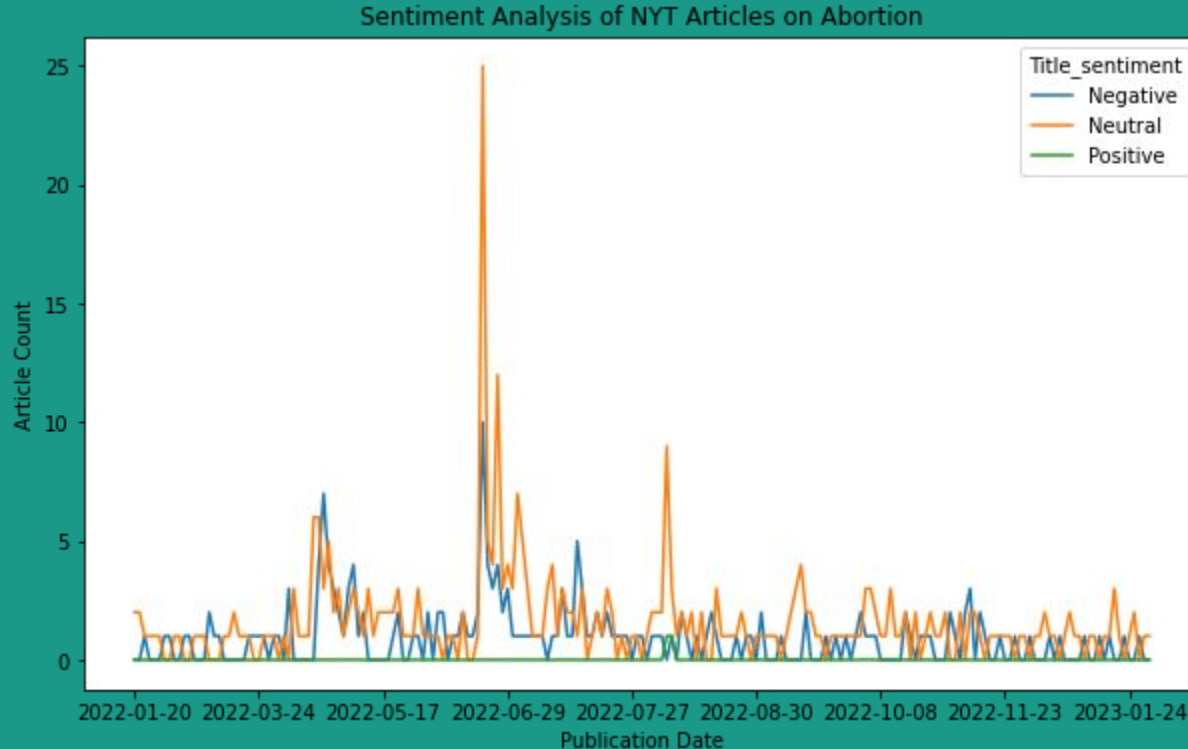
Distribution of Sentiments in NYT Articles



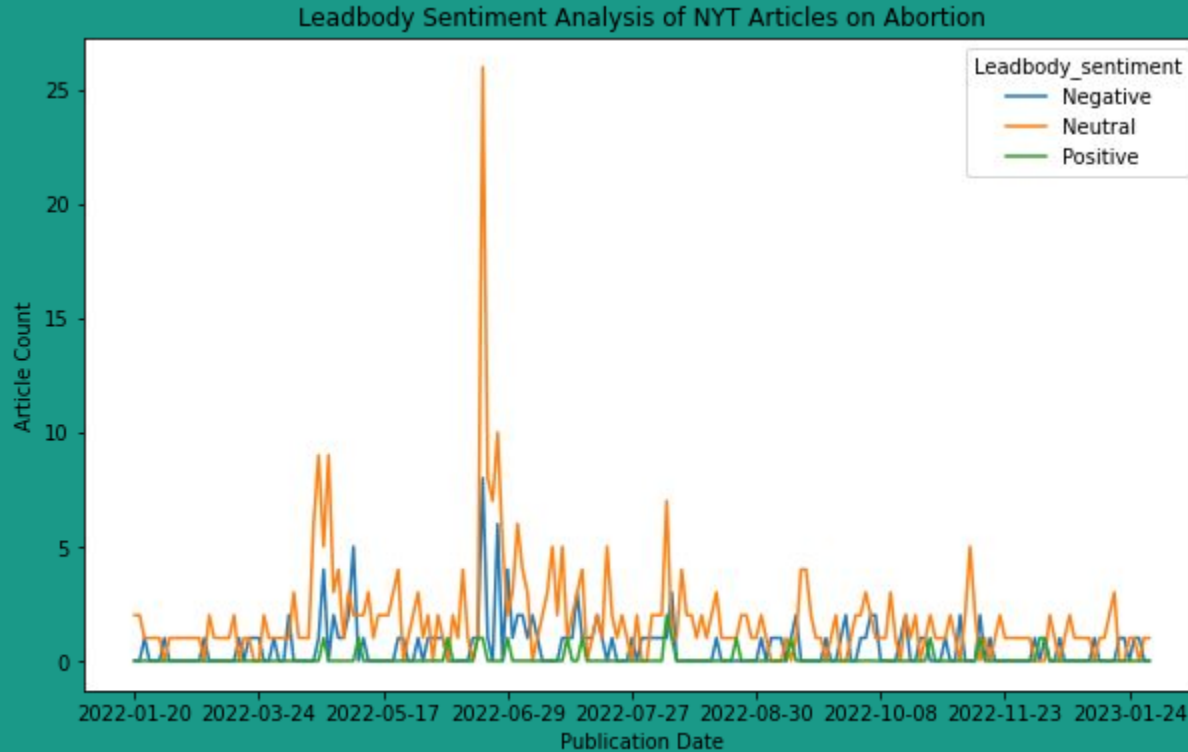
A snapshot...

	Time	Title	Link	Leadbody	Title_sentiment
0	2022-01-20	At Sundance, Two Films Look at Abortion and th...	https://www.nytimes.com/2022/01/20/movies/sund...	To hear more audio stories from publications l...	Neutral
1	2022-01-20	Supreme Court Lets Texas Abortion Law Stay in ...	https://www.nytimes.com/2022/01/20/us/politics...	WASHINGTON — The Supreme Court on Thursday rej...	Neutral
2	2022-01-21	The Women Who Had Abortions Before Roe v. Wade	https://www.nytimes.com/interactive/2022/01/21...	We must understand what happened before Roe v....	Neutral
3	2022-01-21	Anti-Abortion Marchers Gather With an Eye on t...	https://www.nytimes.com/2022/01/21/us/march-fo...	Anti-abortion protesters descended on Washingt...	Neutral
4	2022-01-22	On Abortion Law, the U.S. Is Unusual. Without ...	https://www.nytimes.com/2022/01/22/upshot/abor...	Chief Justice John G. Roberts Jr. said last mo...	Neutral
...
510	2023-01-25	New Lawsuit Challenges State Bans on Abortion ...	https://www.nytimes.com/2023/01/25/health/abor...	A company that makes an abortion pill filed a ...	Neutral
511	2023-01-25	2023's Biggest, Most Unusual Race Centers on A...	https://www.nytimes.com/2023/01/25/us/politics...	In 10 weeks, Wisconsin will hold an election t...	Neutral

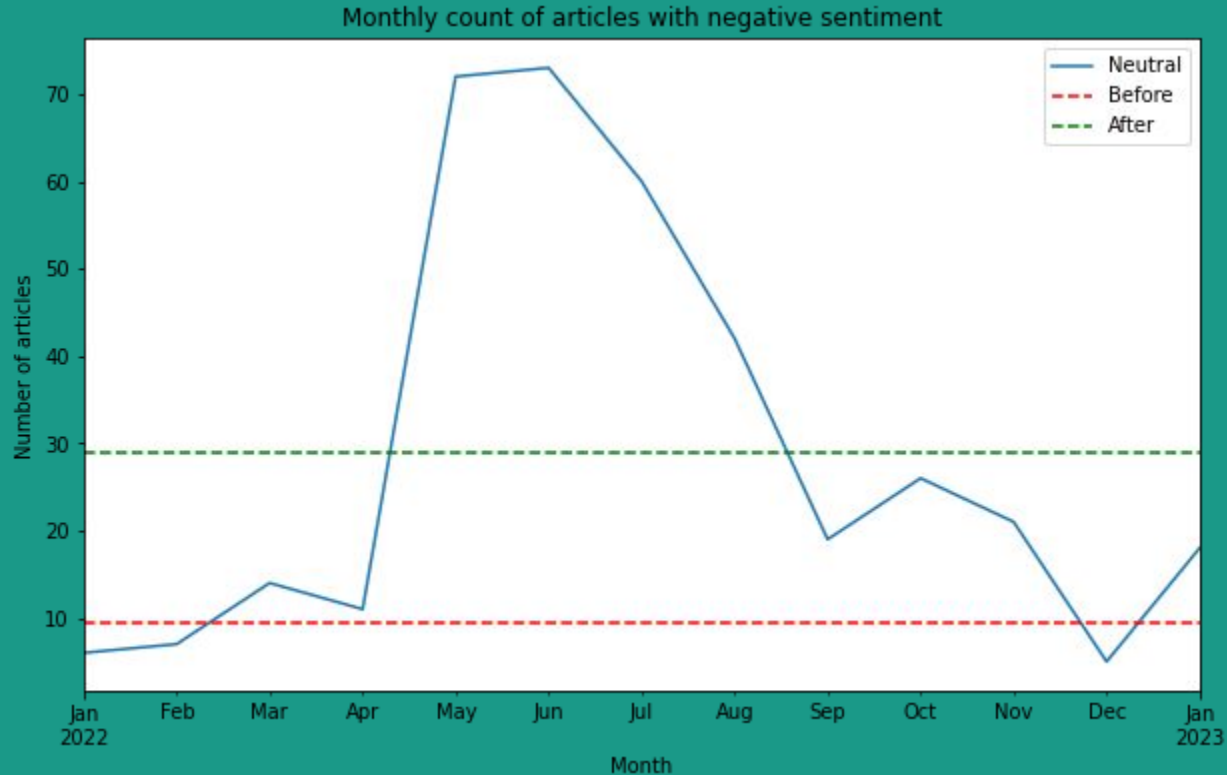
Progression of Sentiments in NYT Articles Over Time



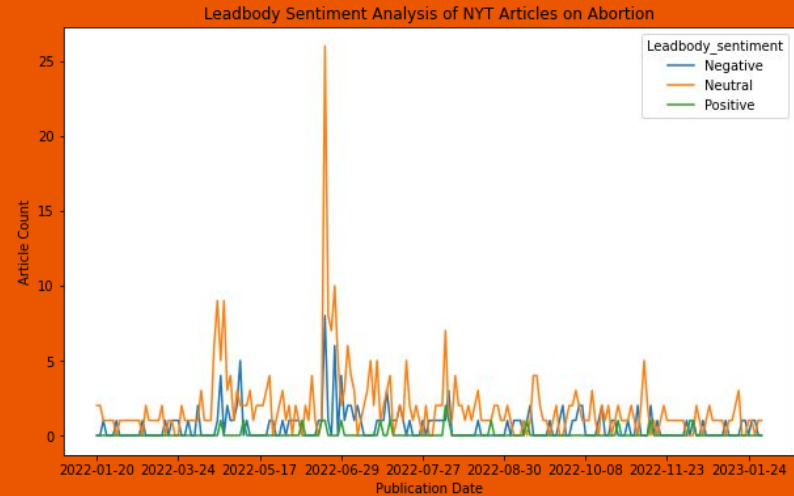
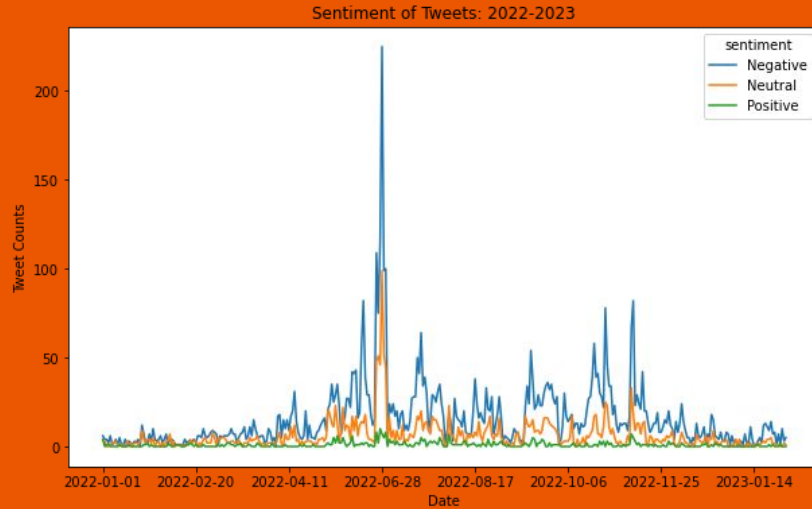
Progression of Sentiments in NYT Articles Over Time



Progression of Monthly Count of Negative Articles

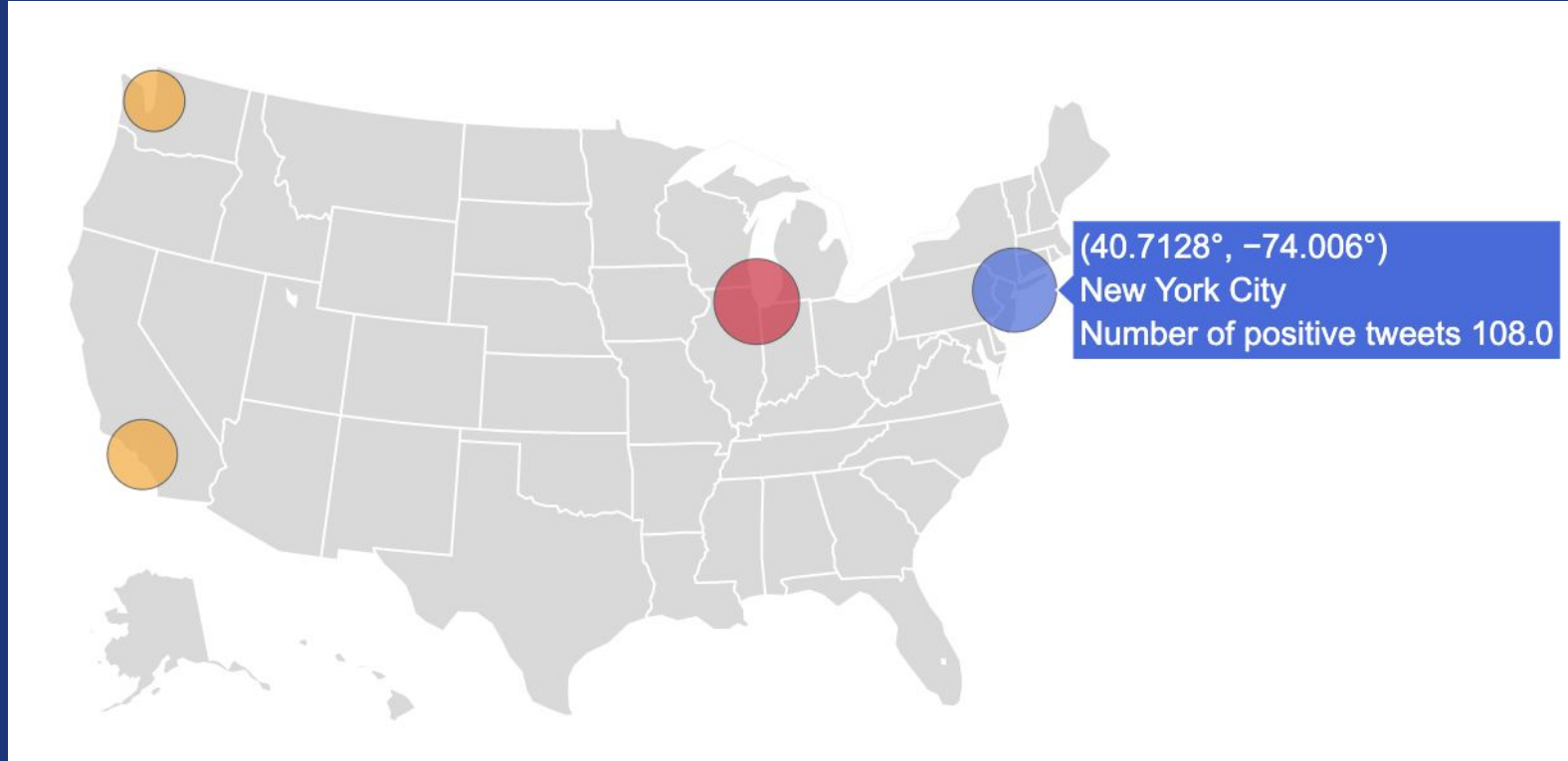


At the time of the overruling in June 2022...

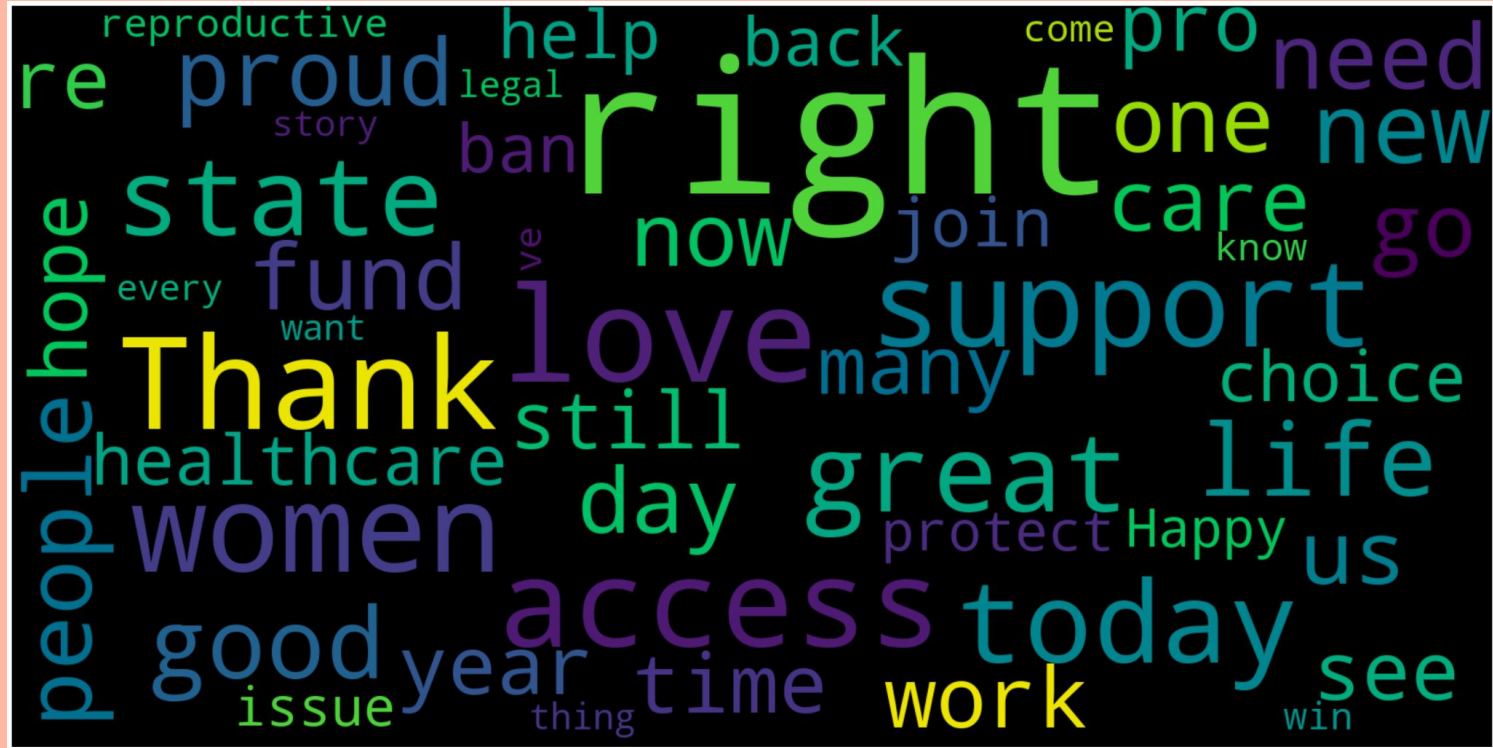


a negative peak for Tweet sentiments and a neutral peak for NYT data can be observed – which ratifies our conclusions regarding the trends of the two data so far

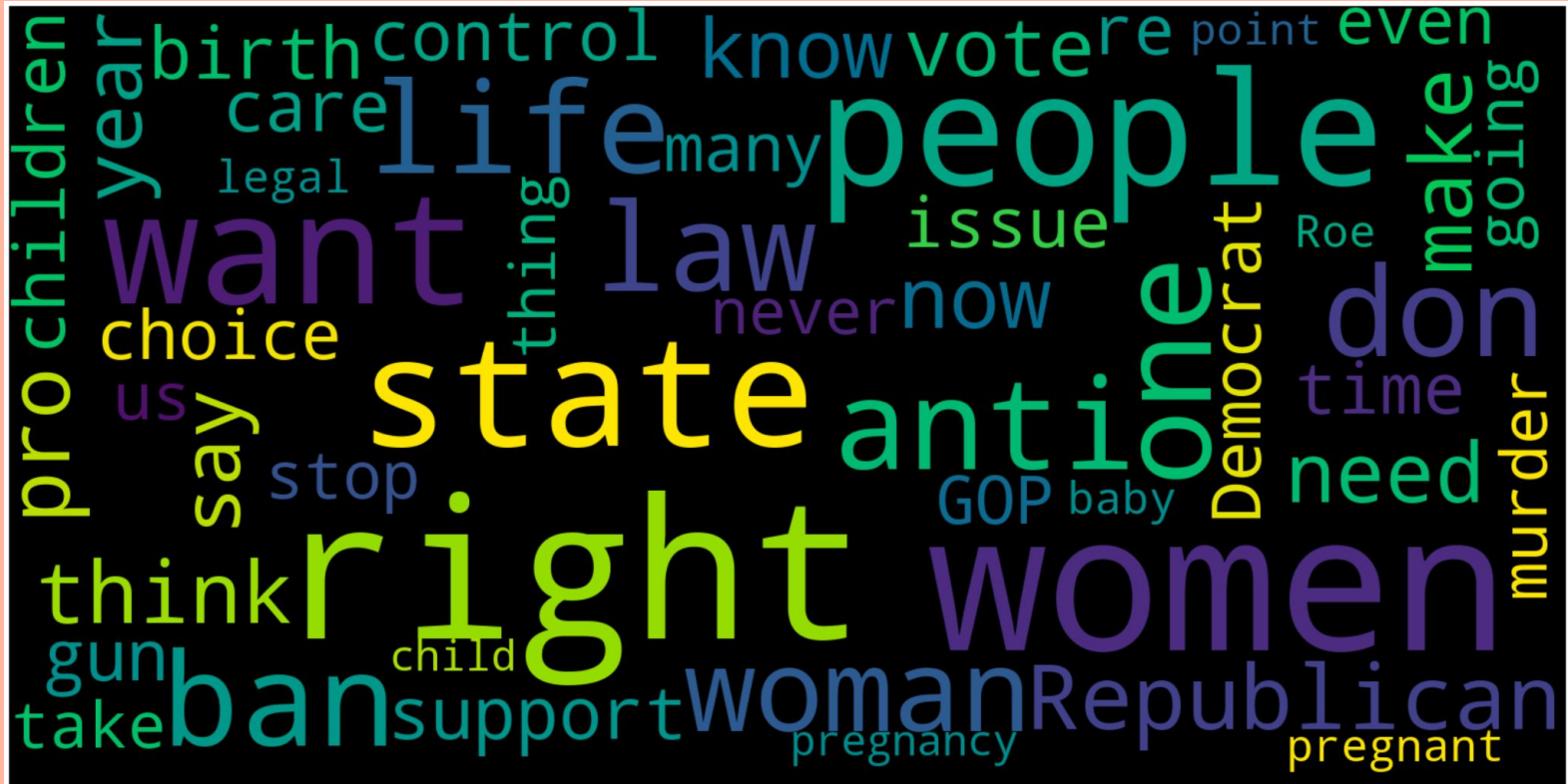
Bubble Map of Positive Tweets by City



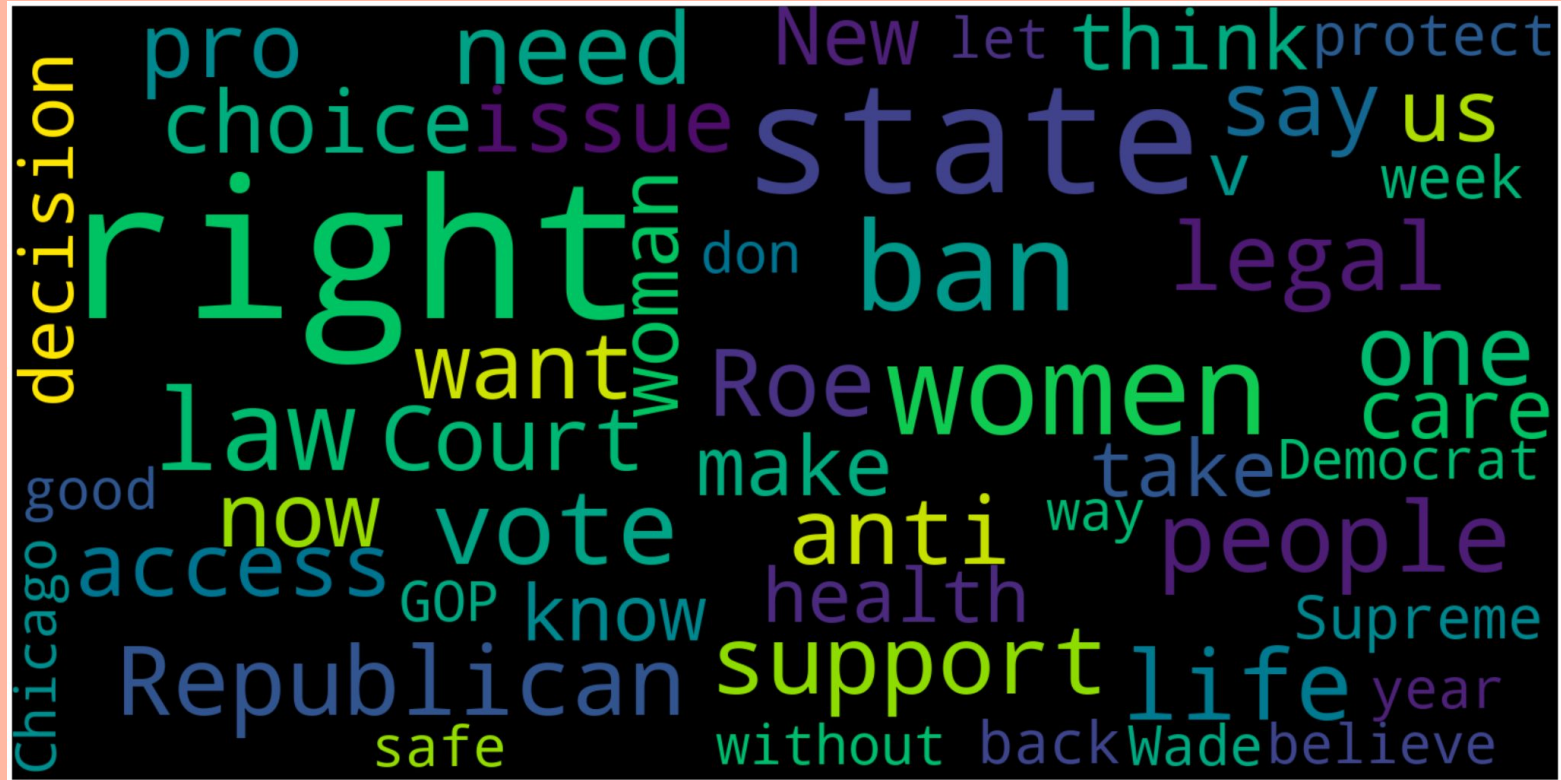
Here's what words in **Positive** Tweets look like...



...what words in **Negative** Tweets look like...



...and what words in **Neutral** Tweets look like.





Key Insights from the Analysis

- **Results of the sentiment analysis showed that:**
 - Overall, Tweets show a greater extent of extreme sentiments – particularly negative sentiment, across all four major cities we studied
 - On the other hand, the *New York Times* articles (both headlines and lead bodies) generally showcase ‘milder’ sentiments, such that a more neutral sentiment can be observed for them, across the time-period we studied (important to note here that positive sentiment is not observed in both cases)
- **Effect of the June 2022 ruling**
 - Our analysis showcase that the effect of the 2022 ruling was evident in the way the word ‘abortion’ appeared in both, Tweets and NYT articles: while there was an increase in volume of both, Tweets as well as articles, the net effect was convergent to the effect seen in general: negative sentiments for Tweets, and neutral for articles
 - There can be observed a peak in negative sentiment Tweets at the time of the overruling, and a synchronous peak of neutral sentiment headlines in NYT around this same time period - which ratifies our conclusions so far
- **Geographic incidence of views on social media**
 - We also observe that there is a difference in the way sentiments are expressed, on a geographic basis. We studied 4 large US cities. The number of positive sentiment Tweets were distributed as: New York > Chicago > Seattle > Los Angeles



Future Scope of Work

- Causal inference from geographic incidence of views is an aspect of our research question which can be explored further in the future
- Expanding the sentiment analysis to longer time periods, regions, social media platforms, and popular media platforms
- Juxtaposing the official statistics of births, pregnancies, and abortions with the corresponding sentiment analysis of social and popular media to further study the relationship regarding how real-world decisions of people are reflected through media channels, if at all



Key Insights from the Project

- Using social media data – scraping using various channels, pre-processing, and examining multiple use cases to ensure that the **data quality is maintained** + no biases creep into the data collection process
- Using data from web APIs and scraping – cleaning it, storing it properly for efficient usage, and **extracting best use cases** from it
- Looking for loopholes and possible shortcoming at each **step of the data management process, course-correcting along the way**
- Experimenting with different visualisation techniques, **storytelling using data**, gathering insights

Other important experiences and learnings:

- Coding as an effective team
- Iterating over problems
- Finding solutions together
- Sharing common learnings and experiences within the team

Thanks for your patience!

Please feel free to share your comments + suggestions...
