

# Impact of Abortion Rights on US Midterm Elections 2022

Ami Chauhan

State University of New York at  
Binghamton University  
Binghamton, New York, USA  
achauha4@binghamton.edu

Kirtika Jawerilal

State University of New York at  
Binghamton University  
Binghamton, New York, USA  
kjaweri1@binghamton.edu

Prarthna Mohanraj

State University of New York at  
Binghamton University  
Binghamton, New York, USA  
pmohanr1@binghamton.edu

Suryavardhan Thummala

State University of New York at  
Binghamton University  
Binghamton, New York, USA  
sthumma1@binghamton.edu

Ravi Teja Tadiboina

State University of New York at  
Binghamton University  
Binghamton, New York, USA  
rtadibo1@binghamton.edu

## ABSTRACT

Abortion has always been a controversy around the world. In 1973, the U.S. Supreme Court determined the verdict of the case of Roe vs. Wade. This decision proved to be a landmark one for the nation. The U.S. Supreme Court ruled that the decision whether to continue or terminate a pregnancy depends on the individual, not the government. However, in June 2022, the Supreme Court reversed the decision of Roe vs. Wade and declared that there is no longer a federal constitutional right to an abortion. In the light of these recent events, it seems that this issue will have a high impact on the midterm elections of the United States.

In this project, we intend to gain an overview of the general public and their thoughts on this particular issue and how it will affect the midterm elections being held in November of this year. Social media is the best option to gather data for this. We will collect data from various forms of social media such as Twitter, Reddit, and YouTube to analyze the impact of abortion rights on the midterm elections being held.

## KEYWORDS

Data sets	Abortion
Twitter	Elections
Reddit	Midterms
YouTube	US

\*Article Title Footnote needs to be captured as Title Note

\*Author Footnote to be captured as Author Note

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

WOODSTOCK'18, June 2018, El Paso, Texas USA

© 2018 Copyright held by the owner/author(s). 978-1-4503-0000-0/18/06...\$15.00

## ACM Reference format:

Ami Chauhan, Kirtika Jawerilal, Prarthna Mohanraj, Suryavardhan Thummala and Ravi Teja Tadiboina. 2022. IMPACT OF ABORTION RIGHTS ON US MIDTERM ELECTIONS 2022. In *Proceedings of ACM Woodstock conference (WOODSTOCK'18)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/1234567890>

## 1 Introduction

Data is being collected from three different sources: Twitter, Reddit, and YouTube.

We use Twitter as the primary source of data. Twitter is one of the most popular forms of social media for political discussions. Teenagers to politicians use this form of social media to express their thoughts. Hence, a huge amount of data is available. Twitter has a number of APIs which can be used to retrieve and analyze Twitter data. We use Twitter's Sampled Stream API which gives us 1% of all tweets in real time. Twitter API consists of a set of programmatic endpoints which can be used to understand or build a conversation on Twitter. With the use of a persistent HTTP GET connection, the streaming endpoint delivers tweet objects.

Reddit was chosen as our second source of data. There are many active political discussion communities present on Reddit. We use Reddit's Search API to collect data. Data will be collected that is related to our keywords (i.e., 2022Elections, politics, abortion, Roe v Wade).

Lastly, data was collected from YouTube using the Google Client API. Data is in the form of comments that are posted for specific videos we have chosen that are related to how abortion is influencing the midterm elections.

Following the implementation of code for data collection, we carried out a test run for 2 days for each source of data. The data collected during this period of time helped us understand its quantitative and qualitative measure. It also gave us an idea of how we will achieve our goal.

## 2 Implementation

### 2.1 Twitter

Twitter's Sampled Stream API is used to collect data. Roughly 1% of random sample of publicly available tweets are delivered in real time using this API. Using this data, we can identify and track trends and monitor general events/sentiments according to Twitter's API documentation. Per session, one client can be connected. This client can disconnect and reconnect up to 50 times every 15 minutes. The code for this is implemented using python and the data is stored in MongoDB.

### 2.2 Reddit

Reddit's Search API is used to collect data. All the subreddits that are related to our query are returned by the Search API. Comments are fetched from each subreddit and stored in MongoDB. The code for this is implemented using python and is run every 5 minutes.

### 2.3 YouTube

Google Client API is used to collect comments from YouTube. Comments will be collected for specific videos that we have given. A build() function is used to gain access. Comments are fetched from YouTube and stored in MongoDB. The code for this is implemented using python and is executed every 5 minutes.

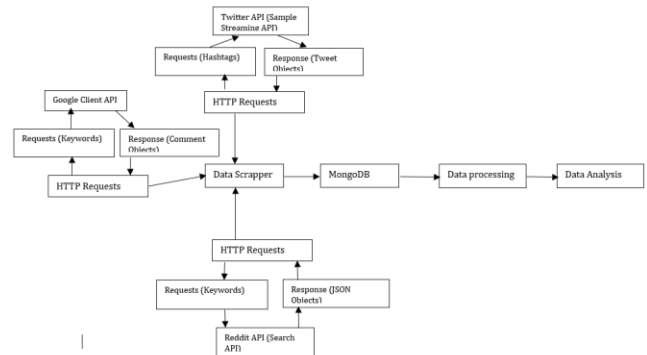


Figure 1: Data collection flow chart

## 3 Modifications

Data from Twitter is being collected using the Sampled Stream API. This API gives out a huge amount of data and the projected data estimation mentioned in our proposal is not valid.

For Reddit, we are still using the Search API as proposed. The comment objects are fetched by querying the URLs built using the subreddits. A few of the subreddits we are getting comments from are abortion, GOP, us elections 2022.

Lastly, for YouTube we had initially planned to get the comments for just one particular video related to our topic. But we have decided that it would be more beneficial and useful if we gather comments from multiple videos. There will be more different thoughts and ideas that can be derived from these comments now.

## 4 Preliminary Exploration of Data

### 4.1 Twitter

We collected data using Twitter's Sampled Stream API. Tweets that are found to be relevant to our topic are the only ones that are stored in our database. Approximately, we collected around a hundred tweets per hour for our specific hashtags. We filtered out the ones that are irrelevant since they will be of no help to our project. Based off our preliminary exploration of tweets, we determined that many tweets were pro-choice. For example, one user had tweeted "This is not just an assault on women's rights, but women's health. It will, as it has been proven historically, force women into unsafe abortion practices."

## 4.2 Reddit

For Reddit, we collected around ten thousand comments per hour using the Search API. We used subreddits that are related to the midterm elections and abortion as our query. The Search API returns everything that matches the query in our code. We only store the comments in MongoDB.

## 4.3 YouTube

Google Client API is used to get comments from YouTube. We selected six to seven videos that were about our topic. An example of a comment we stored is "If the economy is the #1 issue, then abortion comes with it... The top reason for abortion is financial hardship.... An unwanted pregnancy makes people suffer even more than they already..... The voters will not look kindly upon compulsory hardship.....". For all of those videos combined, we extracted around 500 comments. We cannot expect a drastic increase in the number of comments for these videos however. Thus, YouTube will provide us with less data than the other two sources.

## 5 Projected Data

We believe that the data that will be collected in the future will not exceed the storage limit on the virtual machines given.

## 6 Challenges/Hardships

### 6.1 Twitter

People are constantly tweeting on many different topics (tweets on what they are currently doing to tweets about world issues). We are storing tweets that are only related to our project. However, the number of tweets that are related to abortion and the US midterms is not as much as we expected. We hope to get more tweets each day as the election gets closer.

### 6.2 Reddit

The Reddit API documentation is not very specific or well documented. We had to learn how to get comments from reddits/subreddits without using any libraries like PRAW.

### 6.3 YouTube

We are gathering comments from multiple YouTube videos that are related to our project. Some of these videos are 5

months old whereas some are 3 days old. One challenge we discovered with YouTube is that we will not get many new comments on these videos each day.

## ACKNOWLEDGMENTS

We would like to thank our Professor, Jeremy Blackburn, for his assistance with this project. We are very grateful for the feedback, support, and learning opportunities he has provided.