#EndSARS Twitter Analysis

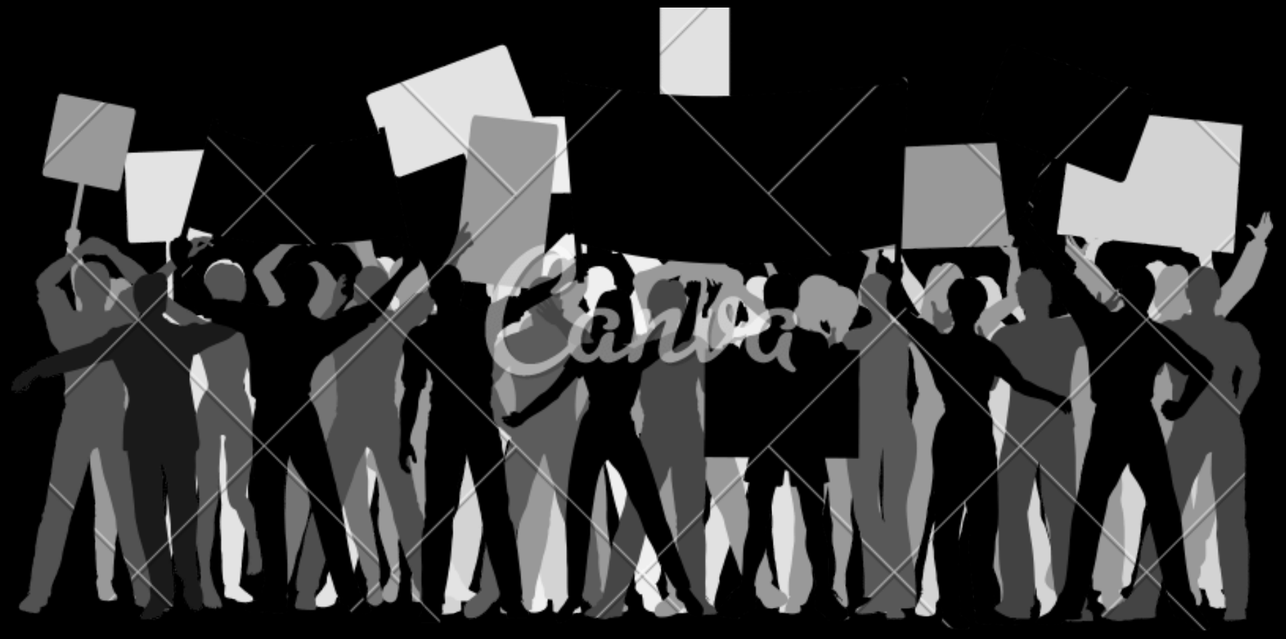
Sentiment and Network Analysis of the #EndSARS protest movement that occured in October 2020 in Nigeria.

Problem Definition

We believe the #EndSARS movement is one of the most powerful movements in Africa over the last two decades. In this analysis, we seek to quantify the impact and provide data-informed answers to some questions.

Methodology

- Data gathering
- Data cleaning
- Data analysis
- Sentiment analysis
- Network analysis



Data Gathering

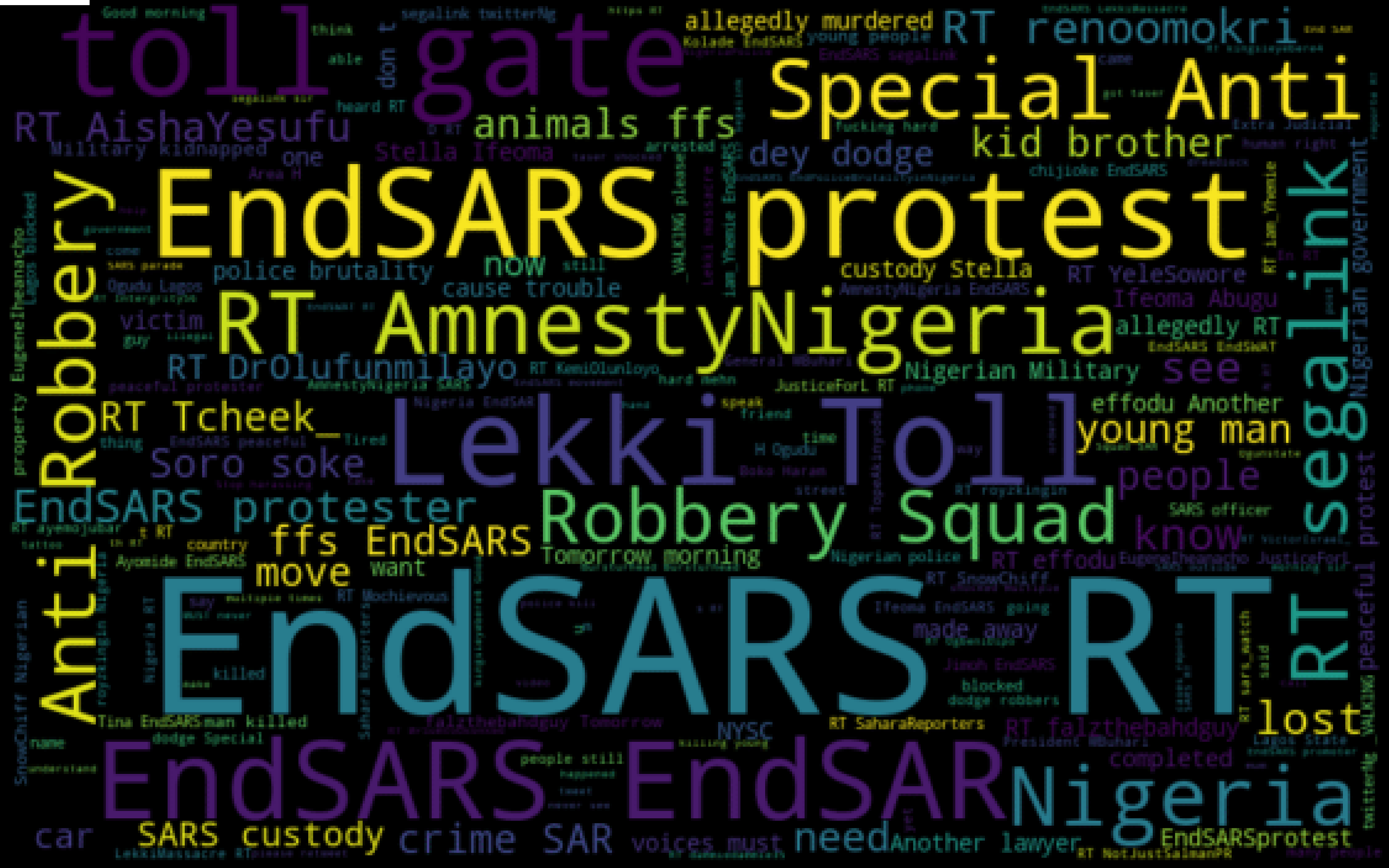
- Data was collected using the official Twitter API.
- Keywords relevant to the movement were passed in as query parameters.
- The time period of data collected was Jan. 2020 to Feb. 2021

Data Analysis

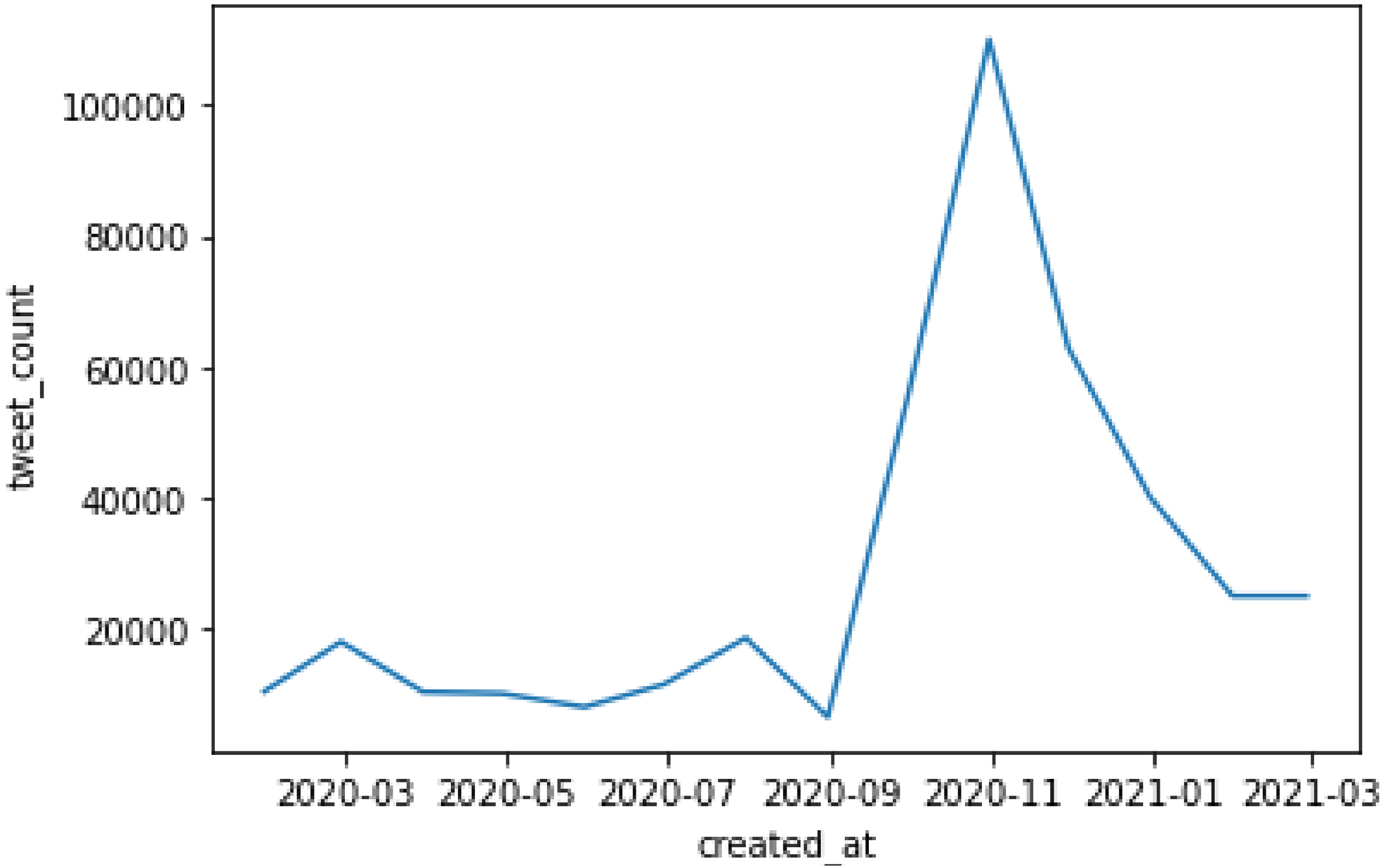
Popular Tweets

TWEETS WITH:	TEXT	VALUE	LOCATION	USER
Highest like count	#EndSARS NOW! □	30,684	US	@JheneAiko
Highest retweet count	#EndSARS	211,671	US	@Rihanna
Highest reply count	'I just lost my kid brother to these animals ffs! □□□ #EndSARS'	1,349	Nigeria	@Tcheek_

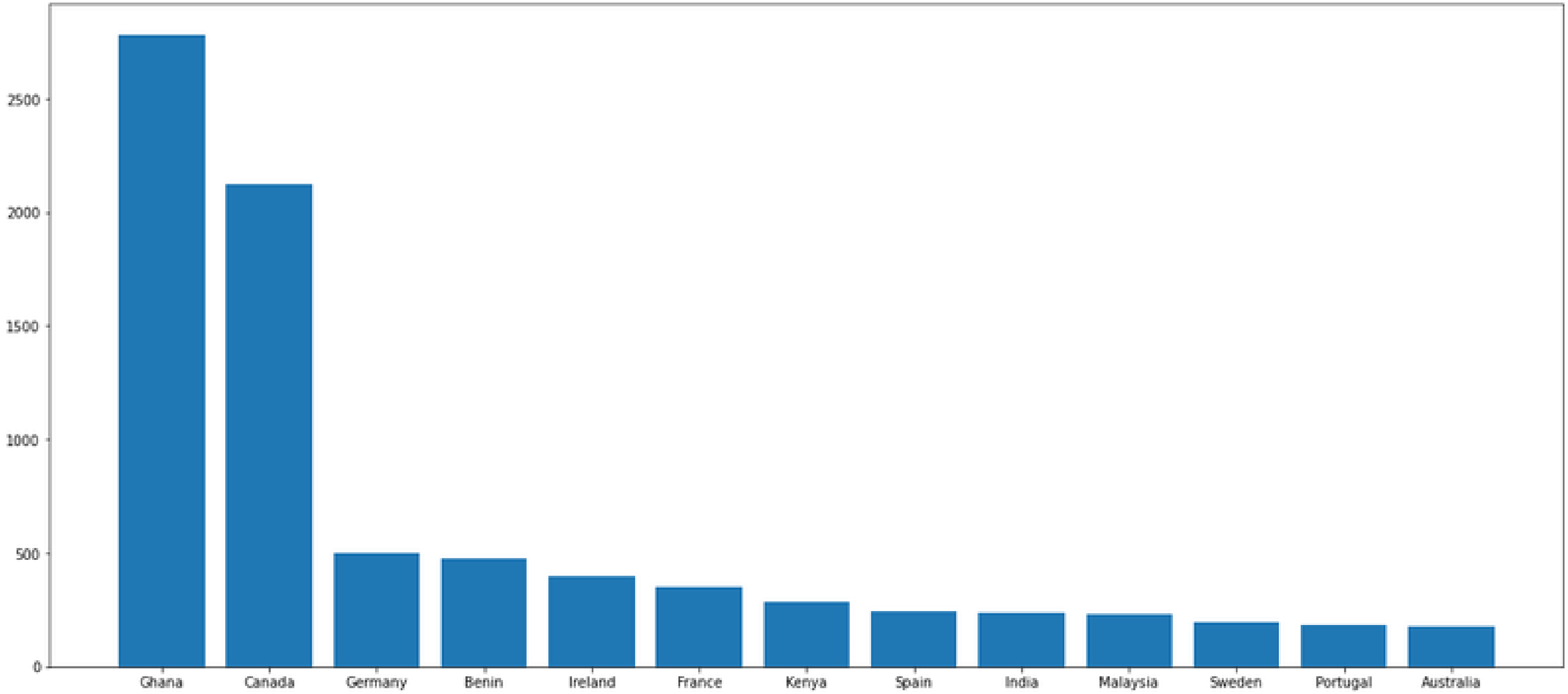
WordCloud


toll gate
Special Anti
RT renoomokri
allegedly murdered
RT AishaYesufu
animals ffs
kid brother
dey dodge
EndSARS protest
segalink
Robbery
RT AmnestyNigeria
police brutality
cause trouble
custody Stella
RT YeleSowore
Ifeoma Abugu
RT DrOlufunmilayo
Nigerian Military
see
Lekki Toll
RT Tcheek_
Soro soke
young man
effodu Another
people
EndSARS protester
Robbery Squad
know
ffs EndSARS
Tomorrow morning
move
Anti
EndSARS RT
peaceful protest
Nigerian government
lost
EndSARS EndSAR
Nigeria
SARS custody
crime SAR
voices must
need
Another lawyer
EndSARSprotest
completed
RT falzthebahdguy
car

Tweet count by month



Top countries with the most interactions



Sentiment Analysis

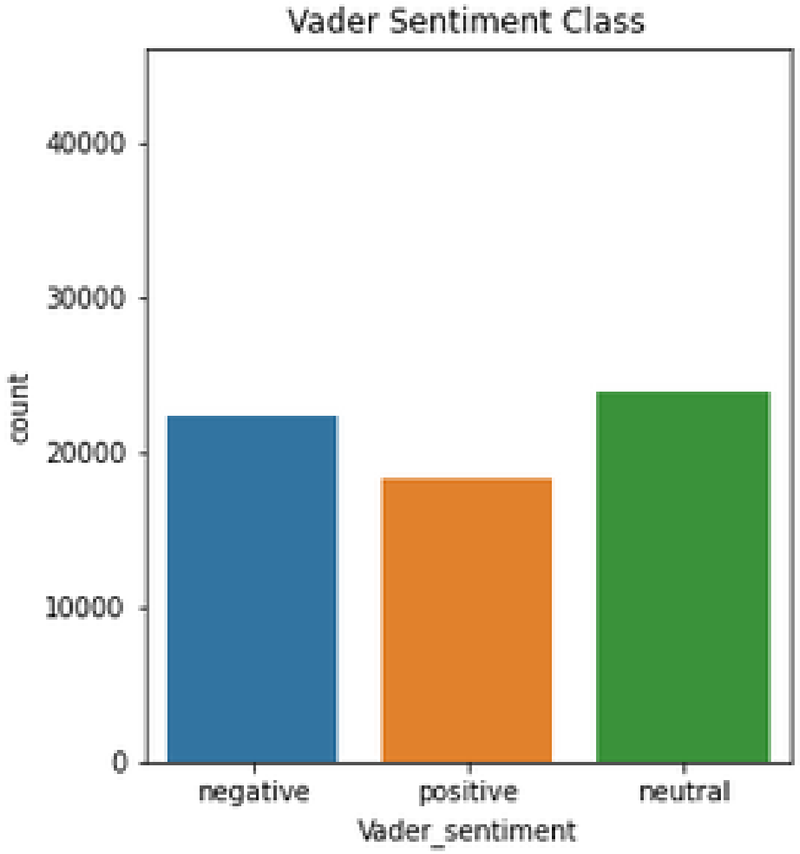
Approach:

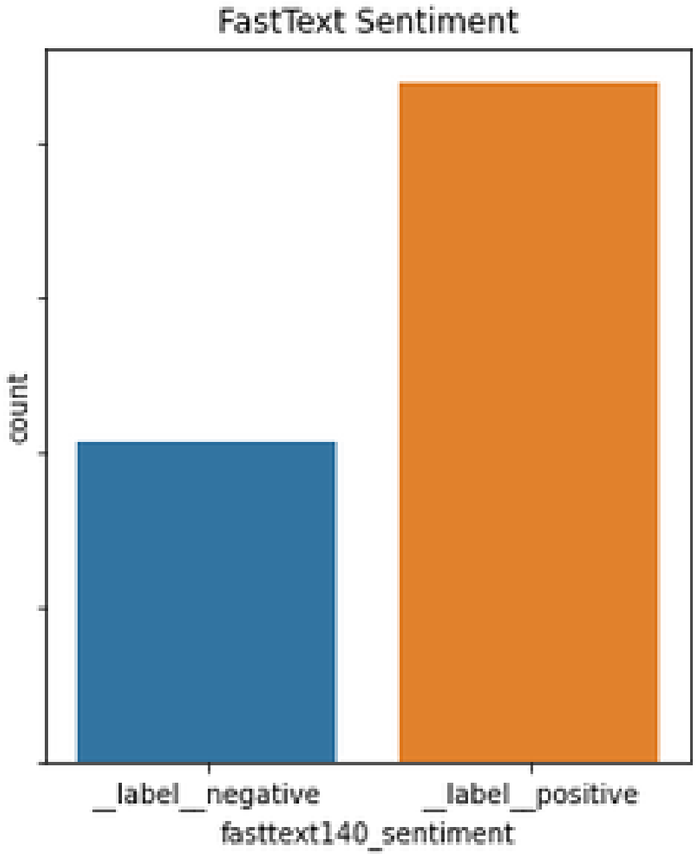
- Three models used
- Vader lexicon: uses a lexicon based approach
- Textblob: uses a rule based approach
- FastText: Supervised learning. Model was trained on tweets that do not necessarily represent the context of our data. (Precision and recall on validation set was about 0.8)

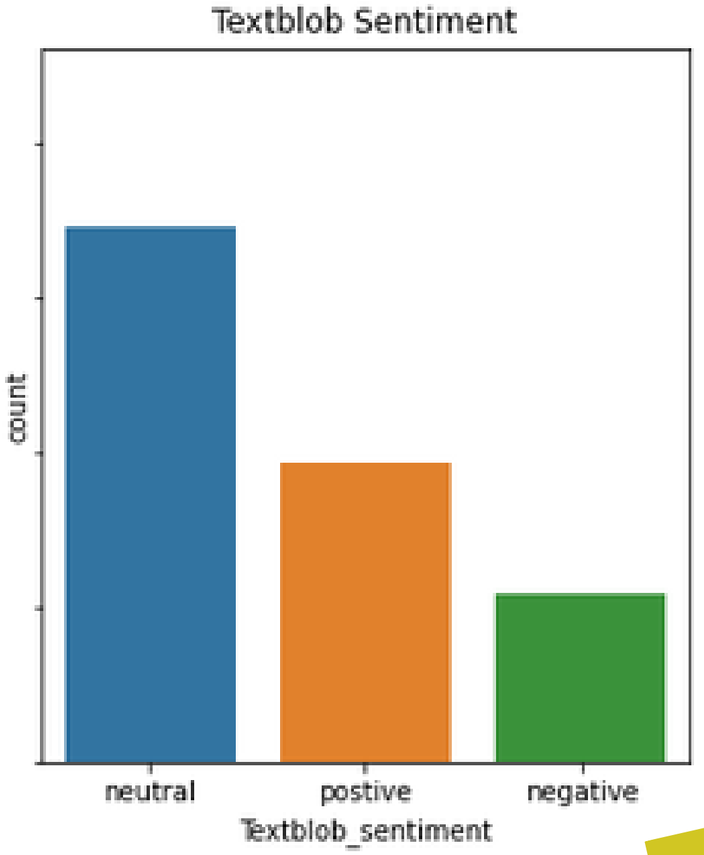
Challenge: Difficulty is getting labelled dataset with similar context to train on.

Sentiment Analysis

Model comparison







Sentiment Analysis

Outcome:

- Sentiments varied across the three models.
- FastText seemed to generalize on the dataset best.
- Positive sentiments mostly constitute tweets in favor of the movement, jokes, adverts.
- Negative sentiments were mostly captured by tweets expressing anger towards SARS and the government. Some adverts and jokes were also classified as negative sentiment.
- Many false negatives and false positives were captured by all three models.

Sentiment Analysis

Sample tweets:

positive sentiments:

```
array(['you nitwitts hide behind keypads under the guise of twitter influencers the highest you ve ever done is trend',
       'if you clean up they will harass you for being fresh if you look rugged they will harass you for being unkempt there s not
hing you can do to avoid sars brutality and harassment in a country with doltish sars operatives endsars',
       'the long wait is over dontshoot video and audio dropping this friday endsars please dontshoot',
       'dstvstepup is changing levels just pay now and get upgraded to the next level oscars2020 endsars',
       'let forget maryambooth iphone 11 madoh endsars who are you suppoing now like for retweet for me segalink',
       'forget about endsars maryambooth madoh and iphone11 agbalumo is overated like if you agree',
       'why did she fall like that wetin you notice endsars arewatwitter madoh askmeek maryambooth ki',
       'madoh maryambooth endsars',
       'no other theme is qualified enough to describe this endsars campaign madoh',
       'being a statement by the lnc mnn joint secretariat in reaction to the setting up of a 57 member constitution review com'],
      dtype=object)
```

negative sentiments:

```
array(['iphone 8plus mint condition 64gb 130k endsars maryambooth madoh sega iphone11',
       'forget endsars sega and maryambooth this is how lagosians would be entering bikes in the areas it wasn t banned',
       'we ve been having series of protest here in nigeria but the govt turns deaf ears to them all why don t you ladies tr',
       'the internet is just filled with bitter people endsars',
       'this is getting serious endsars now', 'sars again endsars',
       'i stand against brutality in nigeria endimpunity endpolicebrutality endsars ht',
       'had my first experience with sars officers they looked like hardcore criminals endsars',
       'finally segafraud unmasked the real sega that people are yet to know i don t have a problem with his height but no dou',
       'us air force vs nigeria air force i think the army man mistakenly joined army instead of police force check is appli'],
      dtype=object)
```

Network Analysis

- Network Analysis is a set of integrated techniques for depicting relationship among actors and to analyze relationship between these structures.
- It shows the relationship between nodes (actors) using edges.
- We created a network using mentioned users from the tweets collected.
- Eigenvector centrality is a measure of influence of node in a network and it was used to measure the top rated users in the campaign.
- The top rated users are listed below.

Top Influencers using Network Analysis

Segun Awosanya (@segalink)

- Eigen Vector Centrality - 0.41

Nigeria Police Force (@NGPolice)

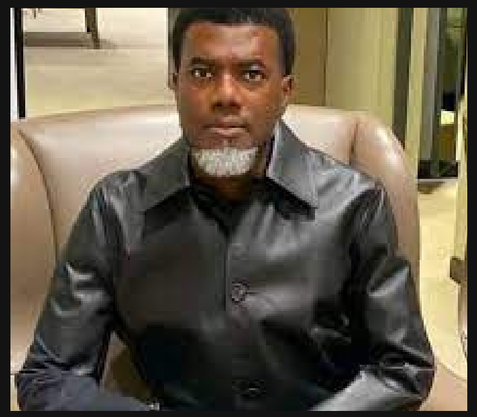
- Eigen Vector Centrality - 0.377

Aisha Yesufu (@AishaYesufu)

- Eigen Vector Centrality - 0.08

Amnesty International Nigeria (@AmnestyNigeria)

- Eigen Vector Centrality - 0.16



Reno Omokri(@renoomokri)

- Eigen Vector Centrality - 0.082

Omoyele Sowore (@YeleSowore)

- Eigen Vector Centrality - 0.081

Conclusion

Social Media has become a tool for people to make their voices heard and this is evident in the #endsars protest.

People are embracing this new way of demanding for change all over the world. This data is accessible and can be analysed compared to past protest

Analysis such as this is important in discovering important information on people's opinion.

Next Steps

- Improve on the existing model with hyper-parameter tuning.
- Calculate sentiment with respect to top accounts.
- Calculate change in sentiment with respect to time.
- Deploy an interactive dashboard where visitors can get insights about the movement.
- Turn model into an API.