

Vartalaap: What Drives #AirQuality Discussions: Politics, Pollution or Pseudo-science?

Rishiraj Adhikary, Zeel B Patel, Tanmay Srivastava, Nipun Batra, Mayank Singh, Udit Bhatia and Sarath Guttikunda#



Indian Institute of Technology (IIT)
Gandhinagar, Gujarat, India



#UrbanEmission.Info
New Delhi, India

In 2019 alone, air pollution was responsible for 1.67M deaths in India (17.8% of total deaths).*

*Health and economic impact of air pollution in the states of India: the Global Burden of Disease Study 2019: Pandey et al.

In 2019 alone, air pollution was responsible for 1.67M deaths in India (17.8% of total deaths)

Ambient fine particulate (PM_{2.5}) is the most significant risk factor for premature death, shortening life expectancy at birth by 1.5 to 1.9 years.*

*Ambient PM_{2.5} reduces global and regional life expectancy. Environmental Science & Technology Letters: Apte et al.

Adverse Health Effects

*Short-term exposures (up to 24-hours duration) have been associated with,

1. Premature mortality
2. Increased hospital admissions for heart or lung causes,
3. Acute and chronic bronchitis,
4. Asthma attacks,
5. Emergency room visits,
6. Respiratory symptoms, and
7. Restricted activity days

*Global Burden of Disease Project.

Which place is more polluted ?



A



B

While **B** is more polluted. Pollution level in **A** is 6 times above the WHO limit!



A



B

While **B** is more polluted. Pollution level in **A** is 6 times above the WHO limit!

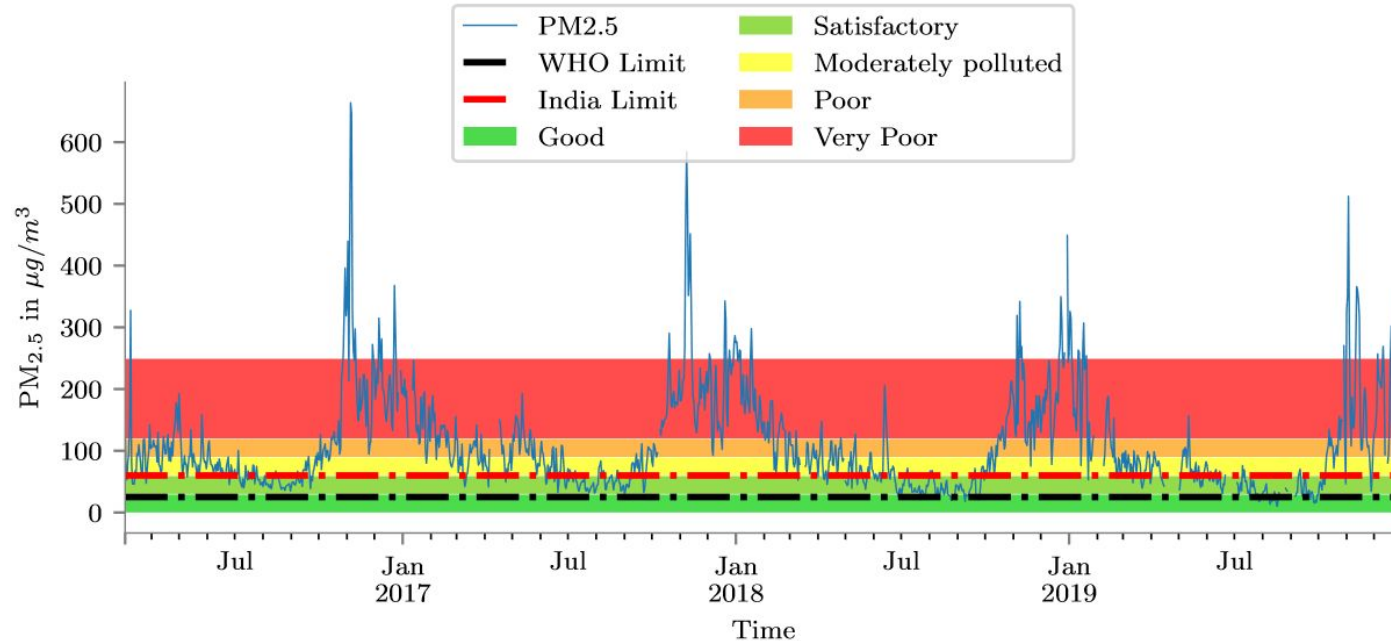


According to WHO, daily mean $PM_{2.5}$ should not be more than $15 \mu\text{g}/\text{m}^3$.

Air pollution is not always visible, leading to incorrect perception.



Air quality in Delhi (India) is a year long problem



Research Questions



What is the public perception towards air pollution mitigation strategies?

Research Questions



What is the public perception towards air pollution mitigation strategies?



Does air pollution cause Twitter discussion?

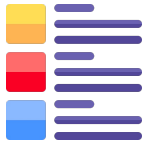
Research Questions



What is the public sentiment towards mitigation strategies?



Does air pollution cause twitter discussion?



What are the air pollution topics discussed on social media?

Related Work

1. Questionnaire based surveys for perception studies are not scalable and do not account for temporal reaction changes.

Related Work

1. Questionnaire based surveys for perception studies are not scalable and do not account for temporal reaction changes.
2. Social media has been used for sentiment analysis on climate change data and air quality. But, air quality studies are on smaller dataset and shorter period.

Approach Intuition

Our approach is to leverage the advances in NLP and statistics to use social media for gauging public perception.

Dataset

We use Twitter to curate a dataset of 1.2 Million tweets which talk about Delhi air pollution between January 2016 to March 2020.

Dataset

We use Twitter to curate a dataset of 1.2 Million tweets which talk about Delhi air pollution between January 2016 to March 2020.

We used query based and hashtag based snowball sampling approach to retrieve the tweets.

Dataset

We use Twitter to curate a dataset of 1.2 Million tweets which talk about Delhi air pollution between January 2016 to March 2020.

We used query based and hashtag based snowball sampling approach to retrieve the tweets.

```
#DelhiAgainstPollution, delhi pollution,  
#DelhiPollution, #SmogTower, #DelhiAir,  
#DelhiAirEmergency, #DelhiChokes, #DelhiAirQuality,  
#DelhiSmog, #OddEven, #DelhiFog
```


Dataset

Twitter Data	Time Period	Jan'16 to Apr'20
	Total Tweets	1.2 Million
	Total Queries	34
Air Pollution Data (PM_{2.5})	Time Period	Jan'16 to Dec'19
	Location	13
	Granularity Level	1 Hour

Word Cloud of Twitter Dataset



Research Question 1



What is the public perception towards air pollution mitigation strategies?

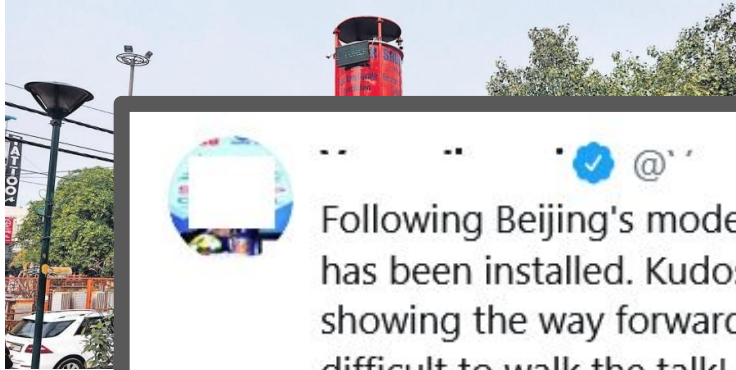
Mitigation Strategies

1. Smog Tower Installation



Mitigation Strategies

1. Smog Tower Installation



@

Jan 3

Following Beijing's model of curbing pollution, the first **smog tower** in Delhi has been installed. Kudos to for this initiative and showing the way forward. Its so easy to keep shifting blame but rather difficult to walk the talk!

Mitigation Strategies

2. Vehicle Rationing Policy



Mitigation Strategies

2. Vehicle Rationing Policy

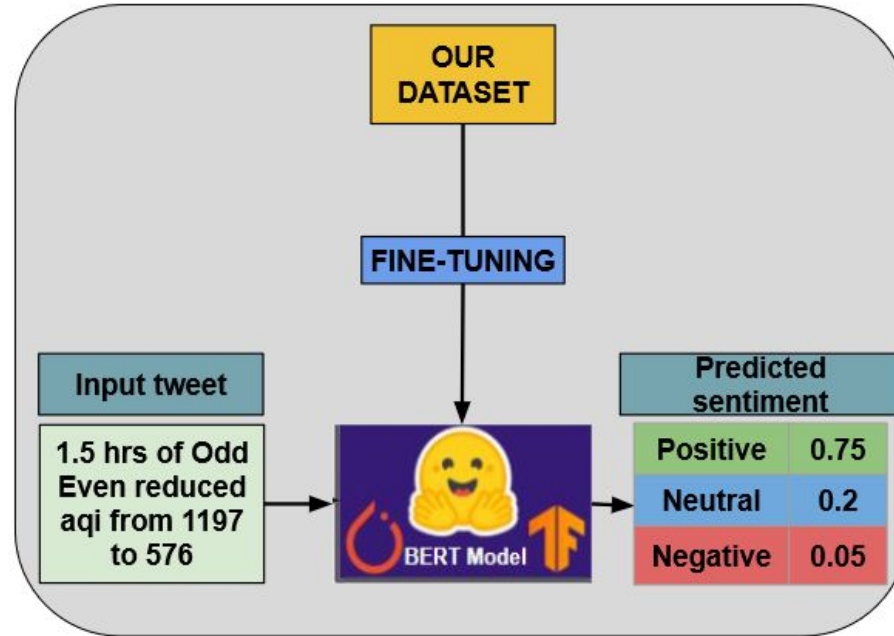


Approach (Sentiments towards mitigation strategies)

In order to understand the sentiment associated with mitigation strategies, we use a language model called BERT.

- The vanilla BERT is trained on Wikipedia data.
- Social media data is different from wikipedia data.
- It is important to contextualise the embeddings to the domain, i.e. social media

Approach (Sentiments towards mitigation strategies)



Evaluation

Total Tweets	“Odd-Even”	“Smog Tower”
	516	80,343
Tweets Annotated (Manual)	516	1100

Evaluation

Total Tweets	“Odd-Even”	“Smog Tower”
	516	80,343
Tweets Annotated (Manual)	516	1100

- Annotated dataset was split into 5 stratified folds. We used nested cross-validation for tuning hyperparameters via grid search.

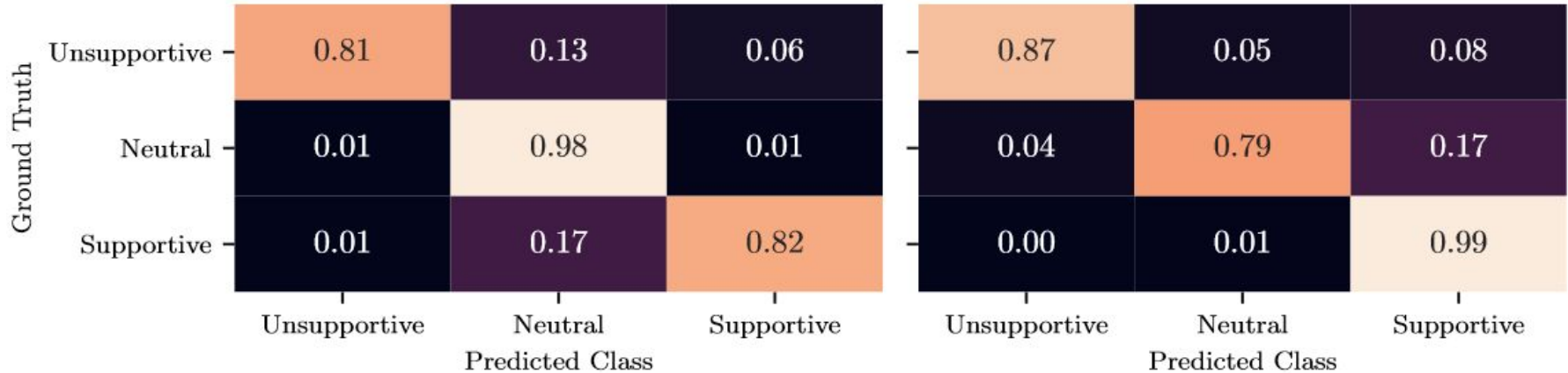
Evaluation

Total Tweets	“Odd-Even”	“Smog Tower”
	516	80,343
Tweets Annotated (Manual)	516	1100

- Annotated dataset was split into 5 stratified folds. We used nested cross-validation for tuning hyperparameters via grid search.
- We select the best performing set of hyperparameters with the highest macro averaged F1 score obtained.

Result

Confusion matrices for untested mitigation strategies:



A. Odd-Even

B. Smog Tower

Challenge

Some tweets can contain multiple sentiments.

Tweet	Ground Truth Sentiment	Predicted Sentiment
#hatersgonnahate but our #smogtower doing its job. fabricated news about rains and wind bringing relief. Tomorrow, when weather reverses, even then smog tower will stand. #smogtowersogood that we are breathing. in your face, haters #myrighttobreathe #Saracsm	Unsupportive	Supportive

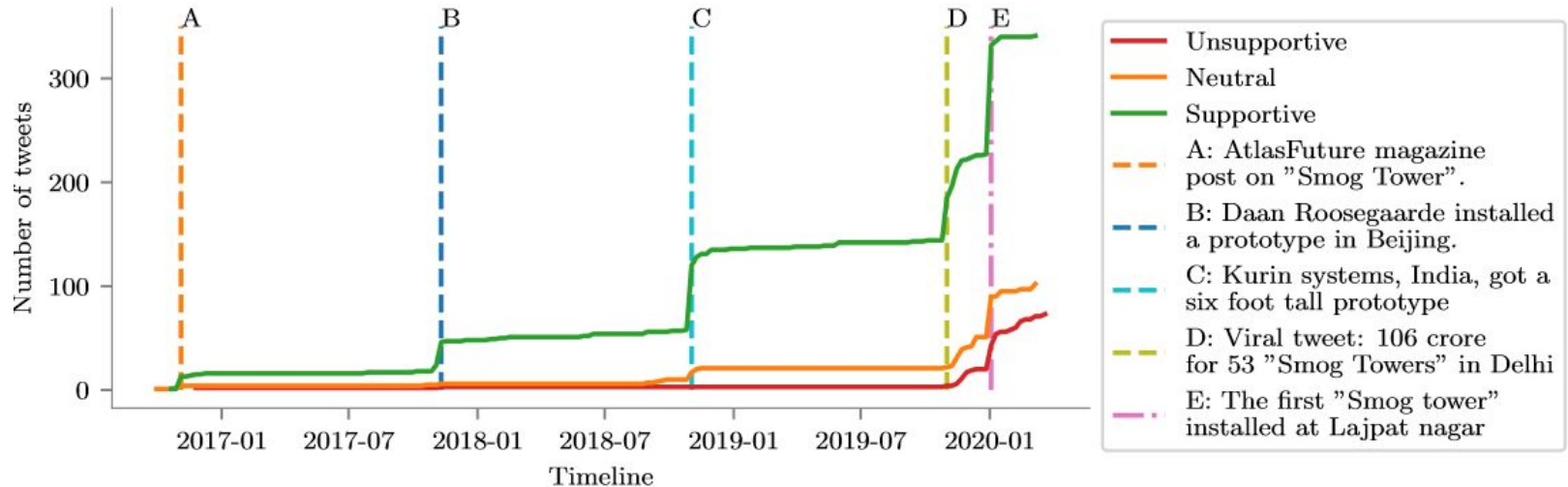
Challenge

Some tweets can contain multiple sentiments.

Tweet	Ground Truth Sentiment	Predicted Sentiment
#hatersgonnahate but our #smogtower doing its job. fabricated news about rains and wind bringing relief. Tomorrow, when weather reverses, even then smog tower will stand. #smogtowersogood that we are breathing. in your face, haters #myrighttobreathe #Saracsm	Unsupportive	Supportive
Wow clean Air in Delhi... Reason.. #OddEven no.. Thank you Nature for cleaning Delhi Air so that we can breath now.. @ArvindKejriwal sir please stop this odd even business and do something to stop stubble burning in adjoining states.. People will appreciate your efforts.. TY <p>pic.twitter.com/ILtw1QQYio</p>	Unsupportive	Supportive

Analysis

Cumulative sum of supportive, neutral and unsupportive tweets for “Smog Tower” over time.



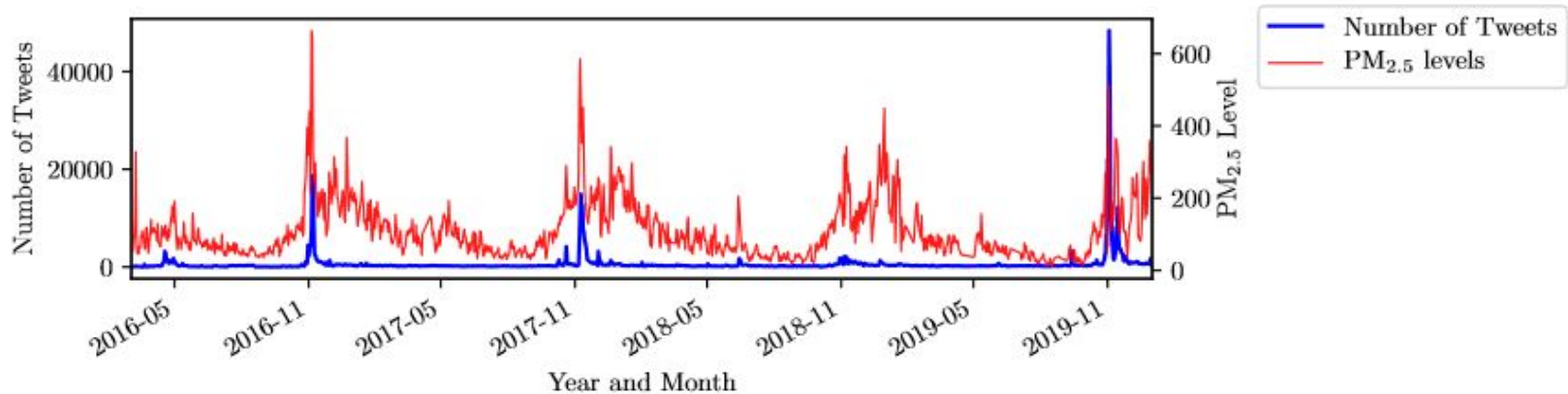
Research Question 2



Does air pollution cause Twitter discussion?

Approach

Our dataset revealed that the frequency of Tweets is high only during the winters when air pollution is high

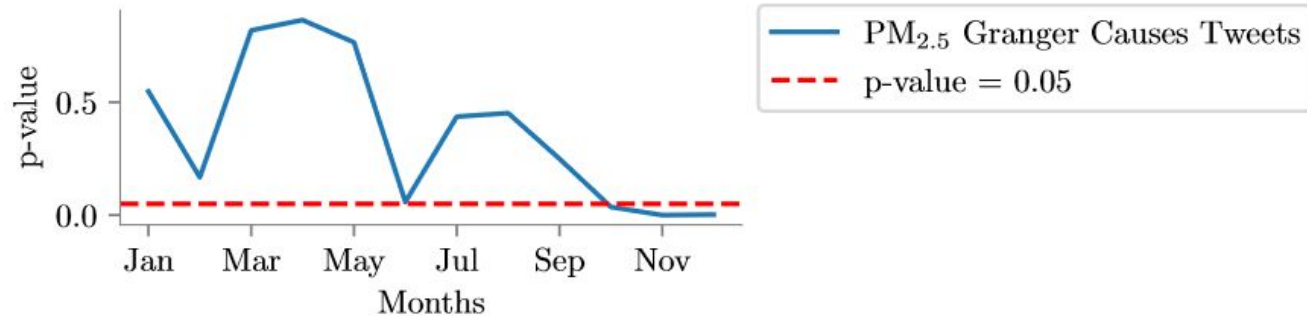


Granger-Causality

We studied whether air pollution ‘Granger-causes’ Twitter discussion in terms of the raw number of tweets.

Granger-Causality

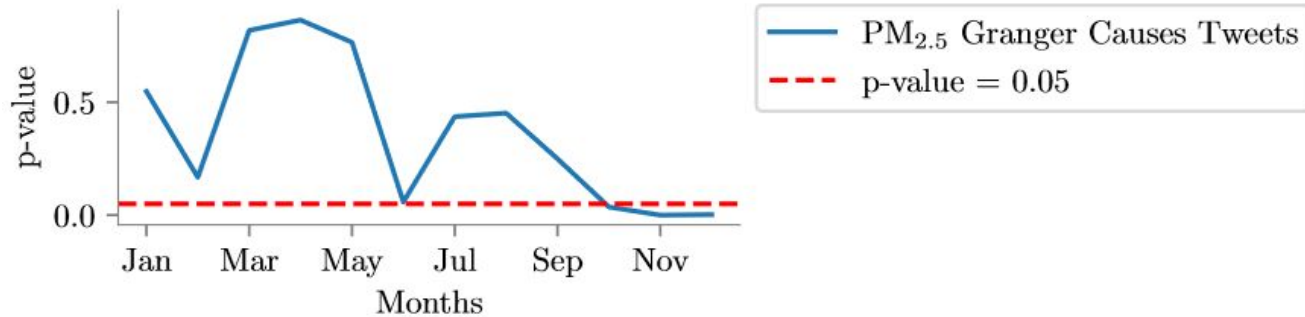
We studied whether air pollution ‘Granger-causes’ Twitter discussion in terms of the raw number of tweets.



We found out that, for most months of the year (January to September), PM_{2.5} does not ‘Granger-cause’ Twitter discussion on air pollution.

Granger-Causality

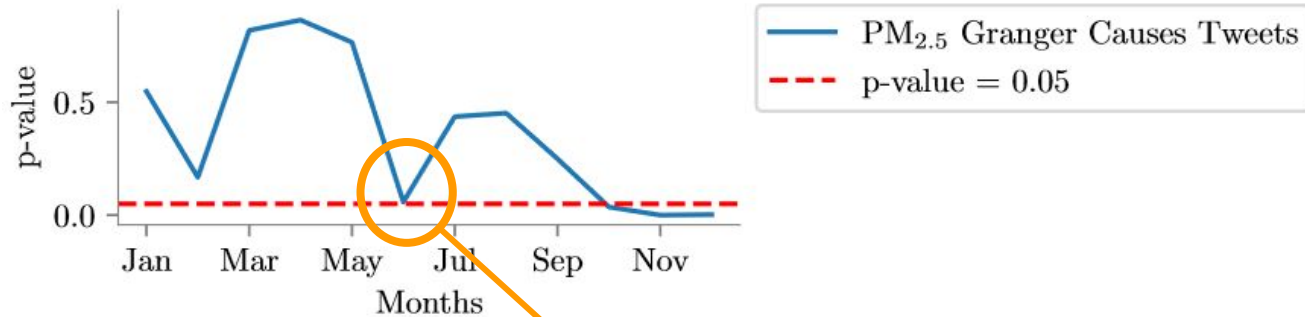
We studied whether air pollution ‘Granger-causes’ Twitter discussion in terms of the raw number of tweets.



During October, November and December ‘Granger-causality’ holds and it is at this time that air pollution becomes visible as smog. Consequently, the public reacts to it via social media

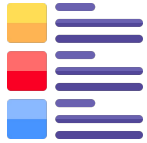
Granger-Causality

We studied whether air pollution ‘Granger-causes’ Twitter discussion in terms of the raw number of tweets.



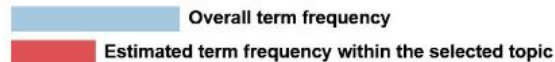
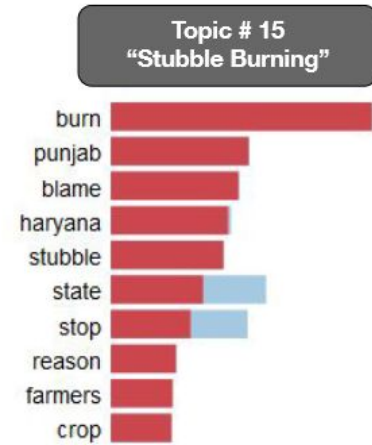
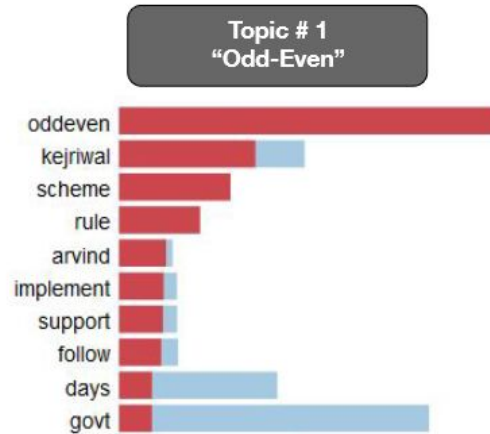
World Environment Day

Research Question 3

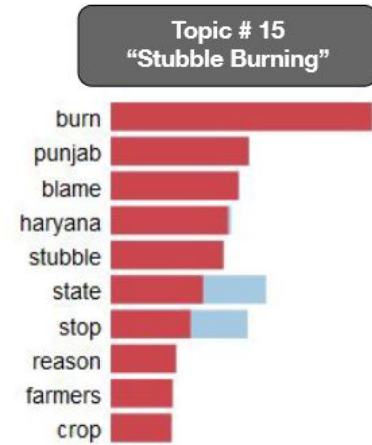
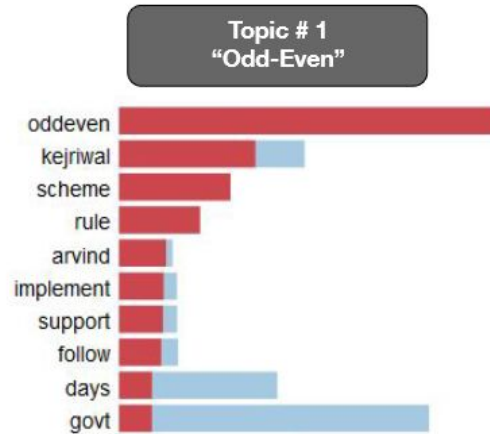


What are the air pollution topics discussed?

Topic Modeling

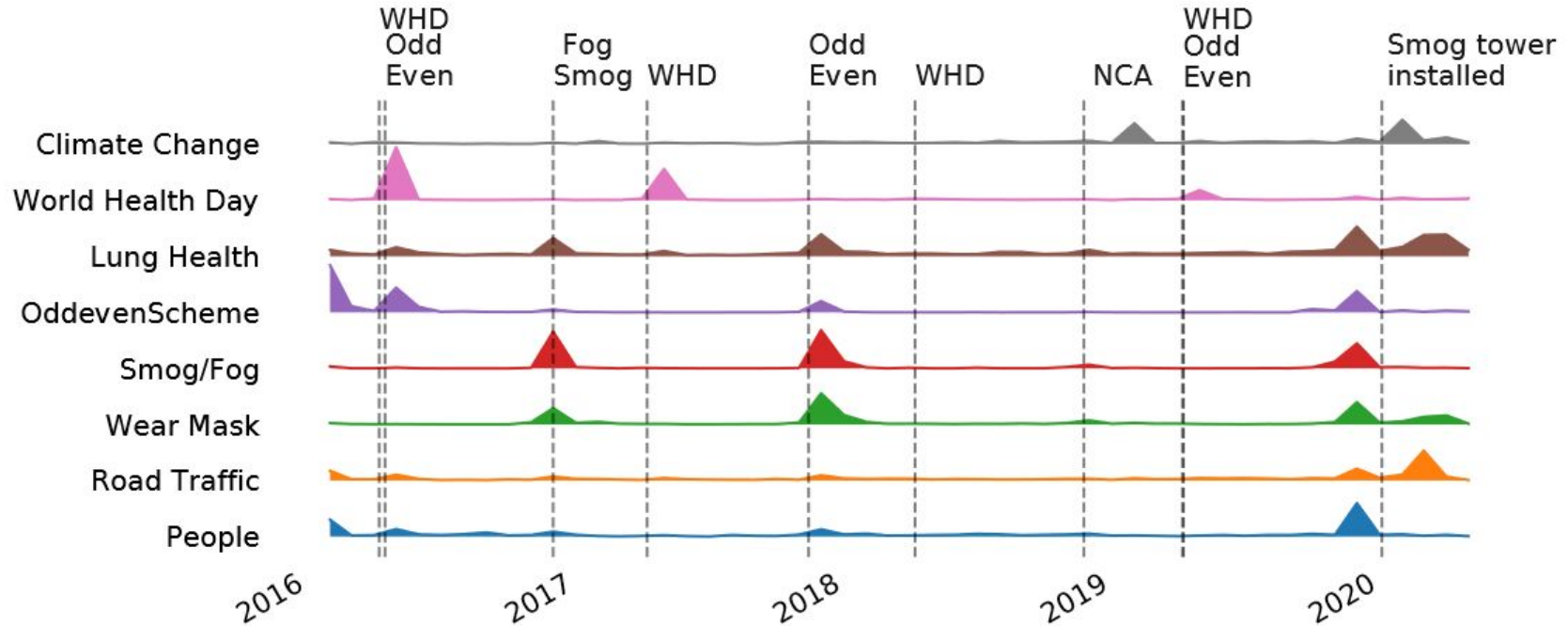


Topic Modeling



Overall term frequency
 Estimated term frequency within the selected topic

Topic Modeling



Summary

- We saw that public perception of untested mitigation strategies is positive which is unfortunate.
- While air pollution is a year long problem, twitter discussions are not.
- Discussions on most of the topics are short lived.