

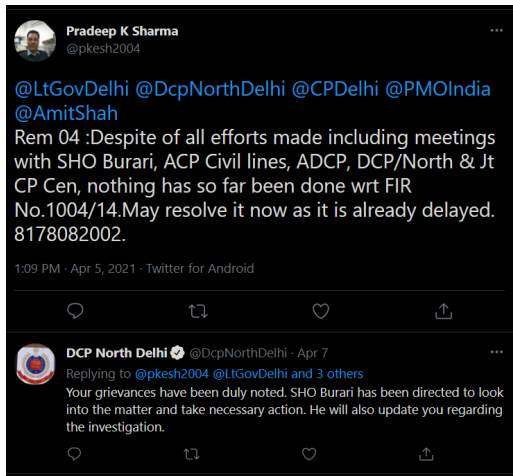
Name - Himanshu Yadav  
Roll No.- 2018286

## Section-1

### Q1-a)

PII Type	# of PII detected
mails	0
aadhar	0
phone	6
pincode	19

```
phone ==> {'9811373344', '0816601007', '9196275118', '8527379140', '8178082002', '9899727721'}  
Pincodes => {'110006', '234993', '210000', '219943', '116409', '273791', '178082', '196275', '110053', '166010', '110007', '113733', '229627', '106371', '136167', '230440', '110084', '110501', '113714'}
```




## Q1-b)

Total tweets contain images, videos or other media = 97

17:32

49%



**शांति सेवा न्याय**  
INFORMATION REPORT  
In respect of Article / Document Lost in Delhi  
SO No: 430 / 2014 Delhi Police  
LR No: 16021072021  
Date: 21/02/2021

**1. Complainant's Details**  
P.S. Crime Branch, Delhi  
Rajdeep Kumar Das, S/O W/O Ravi Das, Ward No. 1, Kirti Road, U.S. Nagar, Uttarakhand 20148  
Rajdeep@gmail.com

**2. Occurrence Details**  
(a) Date and Time of Report Date: 21/02/2021 Time: 16:00 hrs  
(b) Date and Time of Loss, if known Date: 2021-02-21 Time: 03:30 hrs  
(c) Place of Loss Sadar Bazar

**3. Lost Article's**

Sr.	Property Type	Property Description
1.	PAN CARD	AQHP0642R
2.	DRIVING LICENSE	UK0620130060980
3.	CREDIT CARD	37424255307006
4.	DEBIT CARD	ICICI BANK
5.	REGISTRATION CERTIFICATE OF VEHICLE	UK06AP9333
6.	WALLET	5000 CASH
7.	VOTER ID CARD	UTTARAKHAND
8.	POLLUTION CERTIFICATE	UK06AP9333
9.	MISC DOCUMENT	HEALTH INSURANCE CARD
10.	TPA CARD	HEPC

**4. Any Other Details**  
I have lost all the above mentioned original documents@sadar bazar delhi

Station House Officer:  
Crime Branch, Delhi

Note :  
(i) This is a digitally signed document.  
(ii) For verification visit "Lost Articles Report" module on www.dhpolice.ctic.in.  
(iii) Authority issuing duplicate document / article may obtain proof of identity.  
Disclaimer :  
(i) This application is for lodging report of Article / Document lost in Delhi.  
(ii) Report lodged under this application is not a subject matter for enquiry / investigation.  
(iii) In case loss is due to theft or any other crime, contact nearest Police Station.  
(iv) False report to Police is a punishable offence.

Arvind Singh

15 Oct •

वरिष्ठ संवाददाता:- \*फेडरेशन आफ आल इंडिया व्यापार मंडल' दिल्ली ईकाई\* के प्रदेश महामंत्री (संगठन) व सदर बाज़ार विकास मंच के अध्यक्ष श्री देवेन्द्र जैन (देवू) बताया कि सदर बाज़ार, दिल्ली 110006 में 30 हजार से अधिक छोटे-बड़े दुकानदार अपना कारोबार वर्षों से चलाते आ रहे हैं, जो दिल्ली सरकार व केन्द्र सरकार का सबसे बड़ा राजस्व का स्रोत भी है व लाखों दिल्लीवासियों की आजीविका भी इन्हीं मार्किटों से जुड़ी है ॥

दिपावली जैसे बड़े त्यौहारों के आगमन से बाजारों में भीड़ बढ़ने के साथ-साथ यहाँ बेतरतीब बैटरी रिक्शाओं व आटो-रिक्शाओं की अवैध पार्किंग/अवैध पटरी अतिक्रमण के साथ-साथ चोरी-झपटमारी व जेबतराशी की घटनाओं में \*अप्रत्याशित रूप से वृद्धि\* हो गई है ॥


सदर बाज़ार में अवैध पटरियों को माननीय कार्ट द्वारा भी अमान्य व अवैध करार दिया जा चुका है, कोरोना महामारी में मास्क का पालन करने हेतु दुकानदारों के तो चालान काटे जा रहे हैं परंतु अवैध पटरियों व ग्राहक बेलगाम घूम रहे हैं ॥

प्रशासन से अपील है कृपया यहाँ \*पुलिस व्यवस्था और अधिक दुरुस्त\* कर स्थानीय कारोबारी एसोसिएशन के साथ समन्वय कर रोजाना स्थिति का आकलन कर समस्याओं का निवारण किया जाए ॥

\*देवेन्द्र जैन\*

\*प्रदेश महामंत्री (संगठन)\*

\*फैम:- दिल्ली ईकाई\*



# Recents

+91 90984 74676  
India

4:48 PM ⓘ

+91 78730 95348 (2)  
India

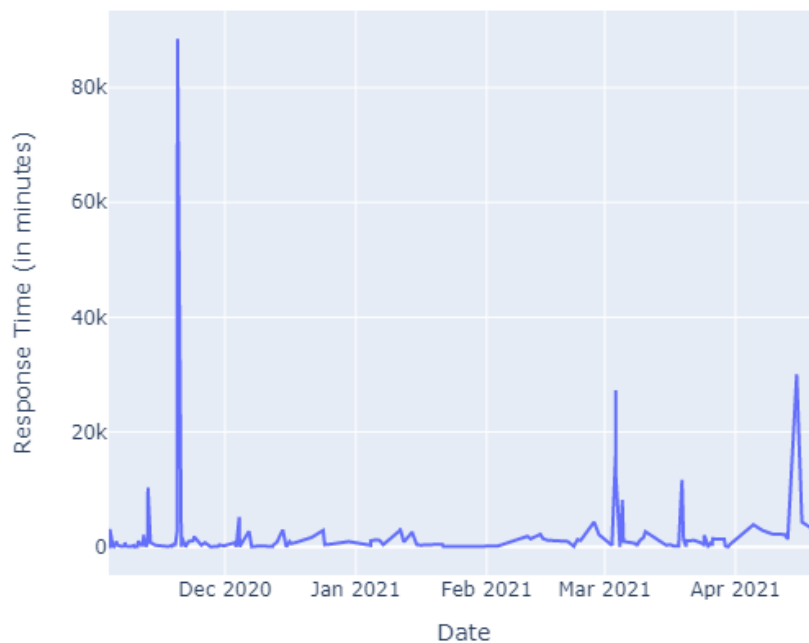
4:48 PM ⓘ

## Q2-a)

Min	Max	Mean	Std Dev
4 min	88485 min	2104 min	7690 min

- As we can see the std dev, we can say that this handle is most likely controlled by humans. And if this handle is handled by bot then mean should have been more less
- Minimum and maximum response time have big difference this shows human controlation, and if it was bot then difference b/w max and min would have been less.

responses times(Time series)



- In time series we can see that there is no recurring pattern in entire time line. By this we can say that delhi police is less active in social media
- From dec to mar delhi police was quite active and given fast responses.
- During nov there was spike that was maximum response time
- During saturday replies were slow
- During normal days(mon-fri) response times were minimum
- During second wave response time were fast in week days

### Q3-

	Tweet	Code
0	@rishirajshanke	other
1	@DcpNorthDelh	accountable
2	बुरारी के अंदर वेस्ट	report
3	@DcpNorthDelh	accountable
	@official_dda @	
4	Illegal encroachr	report
	@DcpNorthDelh	
5	Injustice me http	request
	@DcpNorthDelh	
6	5334 jawahar na	report
	@DcpNorthDelh	
7	No one wear Ma	accountable
8	अनाज वितरण प्रण	report
9	@DcpNorthDelh	report
10	@DcpNorthDelh	report
11	@DcpNorthDelh	accountable
12	@DcpNorthDelh	report
	Illegal encroachr	
	@PMOIndia	

**Coded file attached = annotated\_data.csv**

## **Distribution of every codes -**

Report = 13

Accountable = 6

Request = 10

Appreciation = 0

Other = 1

**The above categories of tweets can be modeled to provide actionable information and how can they act as bridge communication b/w police and user ->**

### **Report -**

- As we can see that most number of coded were in report category. Many user reports crime on twitter before even FIR. if police can look and make there model as who can tell them about the code of the coming tweet , then it will be big help. And it will lowersize the rate of crime drastically. Maybe police should need to use a hashtag(#Report\_handlename) , these hashtags can help to categorize the data fast and will be really helpful to fight crime.
- Example - @DcpNorthDelhi dear sir car side mirror being stolen of my mercedes car outside of my residence no police barricading is there at our locality during night no patrolling please look into this 5334 jawahar nagar delhi 110007
- I above example we can train our model to detect key words like(stolen, barricading) and then we can make it to detect address which will help police to take action fast.

### **Request -**

- Second most occurring code were 'request'. Many user tend to request and trying to reminder their main problem through tweet. Please and help can be used as keyword to detect request type tweets, many time request tweet are also kind of report tweet, if we can predict them right then we can use them to fight against crime.
- Example- @DcpNorthDelhi @DelhiPolice @CPDelhi Sir Please again request help me Injustice me <https://t.co/XRNgnXQ32a>

- In above example we can see that someone trying to request police to look into previous matter, we can train our model to look these kind of tweets and extract the initial tweet of this thread and take action fast.

### **Accountability:**

- This was 3rd most occurring code in the 30 tweet. Accountability tweets can help police to realize their mistakes and anything missing.
- Example -@DcpNorthDelhi Sir From last three months janta bazar is being set up regularly in sec15 block B & C without any permission . I have already complained about it several times but still the problem persists. No one wear Mask and social distancing in bazar .
- As we can see in the above tweet, prev complain has not been handled by police. We can train our model to look for previous or main thread of this tweet and trying to extract the useful information (address, category of case) and use them to fight against crime.

There were no appreciation post but we can train our model to detect keywords("thanks" etc) to find about them.

Many time it happens that user only report the problem on twitter, police can try to contact them in formal way and listen their problem and tell them each process to take which will lead to justice.

Social media is a sea where two person can be bridge for many other, if police can respond fast then the can lower the crime rate.

---

## Section-2

### Topic - Corona

#### Hashtag - #CoronaSecondWave

I have collected 3000 tweets which contains #CoronaSecondWave

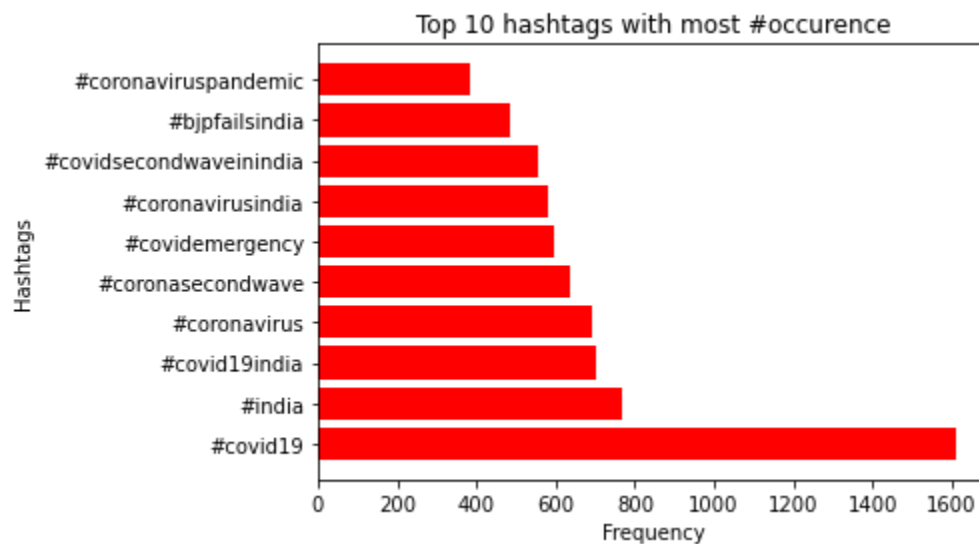
#### Top -10 Hashtags -

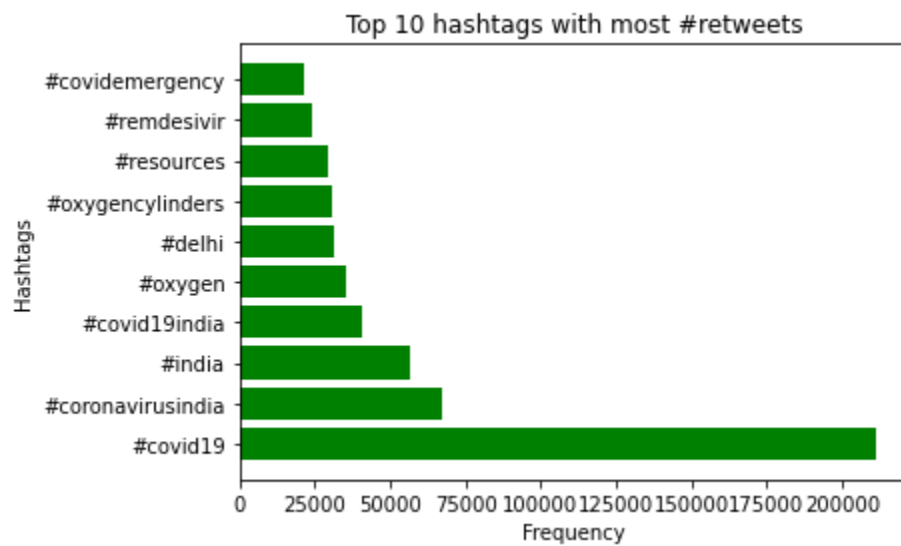
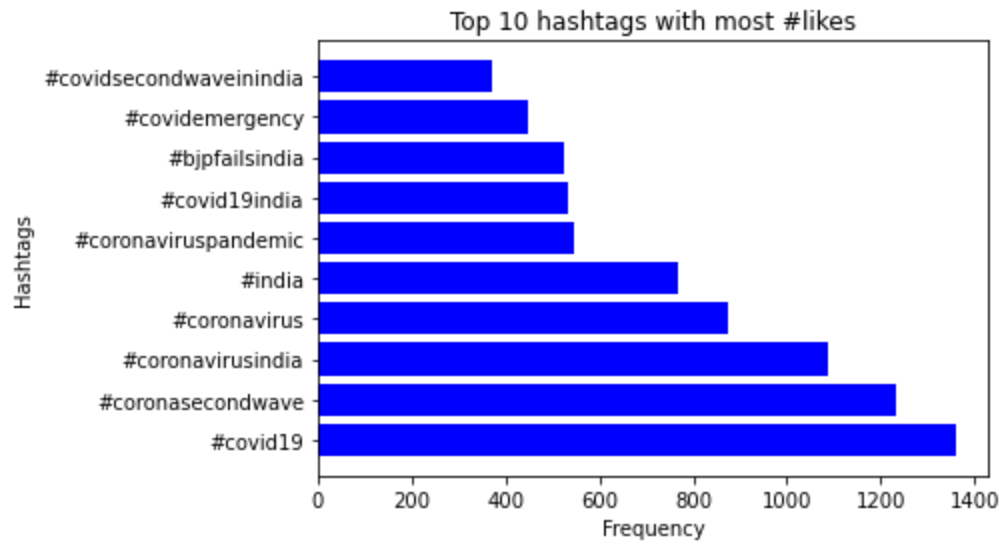
```
['#covid19india',  
'#covid19',  
'#covidemergency',  
'#coronavirus',  
'#coronavirusindia',  
'#covidemergency2021',  
'#covidsecondwaveinindia',  
'#coronaviruspandemic',  
'#india',  
'#bjpfailsindia']
```

and 2000 each for Top-10 Tweets.

Total dataset = 23000 tweets

After removing duplicates = 19207 tweets







## CTM Table for Top-10 hashtags with most # occurrence

	R/10	F	U	Coefficient
#covid19	8.38	7.19	1.27	16.84
#india	8.08	9.2	1.12	18.4
#covid19india	8.24	5.2	1.36	14.8
#coronavirus	7.65	4.75	1.33	13.73
#coronasecondwave	8.16	10.22	1.12	19.5
#covidemergency	8.39	5.15	1.35	14.89
#coronavirusindia	7.93	4.87	1.22	14.02
#covidsecondwaveinindia	8.26	5.39	1.32	14.97
#bjpfailsindia	6.53	4.43	1.29	12.25
#coronaviruspandemic	7.84	5.22	1.18	14.24

## CTM Table for Top-10 hashtags with most # like

	R/10	F	U	Coefficient
#covid19	8.38	7.19	1.27	16.84
#coronasecondwave	8.16	10.22	1.12	19.5
#coronavirusindia	7.93	4.87	1.22	14.02
#coronavirus	7.65	4.75	1.33	13.73
#india	8.08	9.2	1.12	18.4
#coronaviruspandemic	7.84	5.22	1.18	14.24
#covid19india	8.24	5.2	1.36	14.8
#bjpfailsindia	6.53	4.43	1.29	12.25
#covidemergency	8.39	5.15	1.35	14.89
#covidsecondwaveinindia	8.26	5.39	1.32	14.97

## CTM Table for Top-10 hashtags with most # Retweets

	R/10	F	U	Coefficient
#covid19	8.38	7.19	1.27	16.84
#coronavirusindia	7.93	4.87	1.22	14.02
#india	8.08	9.2	1.12	18.4
#covid19india	8.24	5.2	1.36	14.8
#oxygen	8.7	4.24	1.13	14.07
#delhi	8.79	3.93	1.28	14.0
#oxygencylinders	8.76	2.85	1.13	12.74
#resources	8.97	3.35	1.02	13.34
#remdesivir	8.81	2.98	1.05	12.84
#covidemergency	8.39	5.15	1.35	14.89

- As we can see in tables CTM is varied in range (12 - 20)
- U(avg posts/user) is in normal range, which indicates there is no major inorganic data
- F(Percentage of traffic from top 50) is little bit out of bound in some case that shows that top-50 user has more impact in case of some of hashtags(ex; #coronasecondwave)
- Based on current ctm we can say that it is not sufficient to say about inorganic or that there is not much of inorganic content present. Users do retweeted more but still that is not enough.
- No hashtags are unrelated with the topic
  - There is one (#bjpfailsindia) that is a political hashtag but since bjp in government it is connected with the topic.