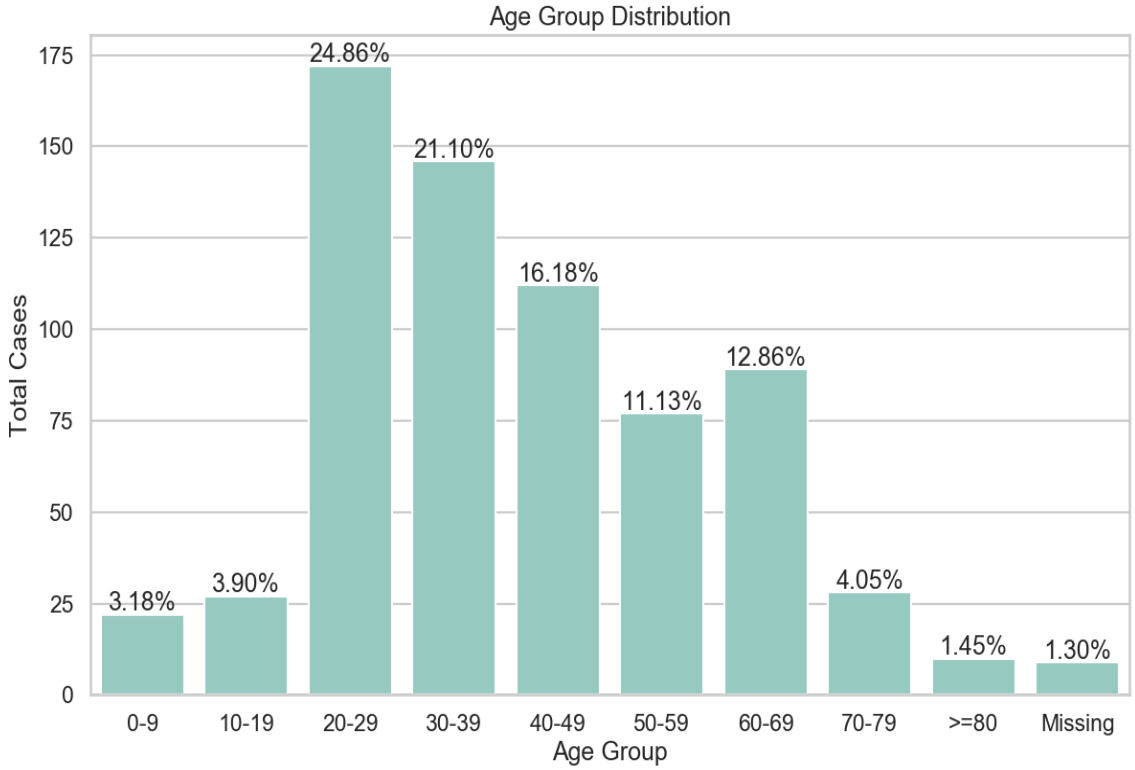
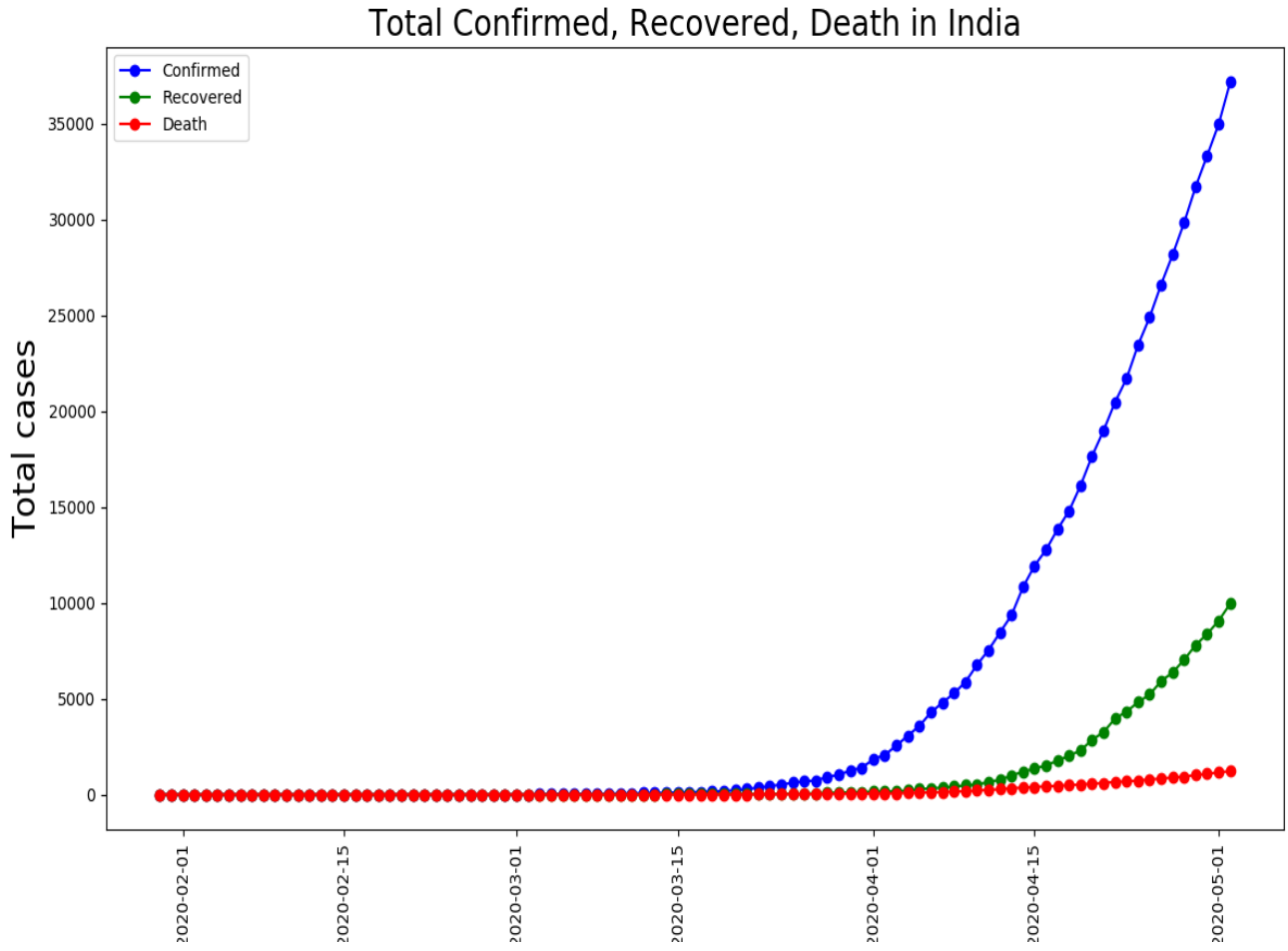
## Q1. AGE GROUP ANALYSIS



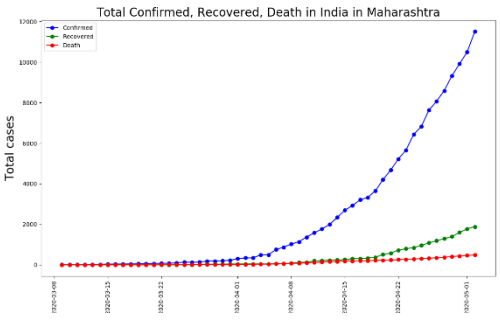

At a national level, the above chart best represents the breakup of patients across age brackets. From, the data given it is clear that the age group that has been most affected from 20-49. Ages 50-69 too seem to be affected to a major extent but the bulk of the infected are 20-49 which constitute for 62.1% of the infected. The ages from 50-69 constitute for 23.9% of the total cases and are the other major infected age group.

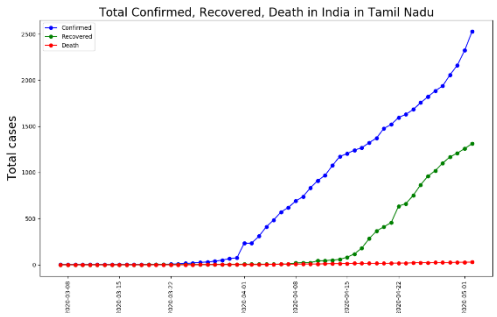
---

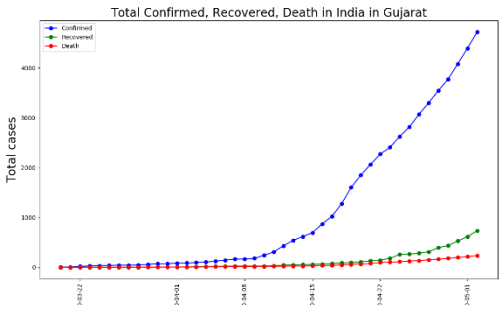
**Q2.**

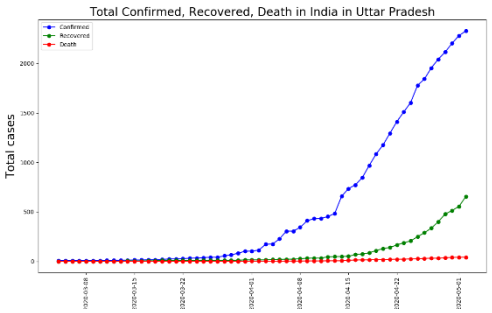


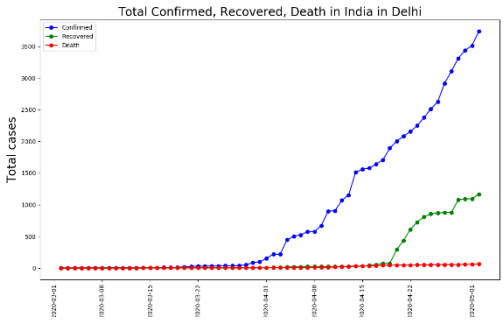
The above plot is the cases confirmed, recovered and deceased per day in India. Here we have assumed that the cases cured correspond to those who have recovered from the disease. The points on the graph correspond to a single day and this is plotted continuously for the entire timeframe of the dataset.

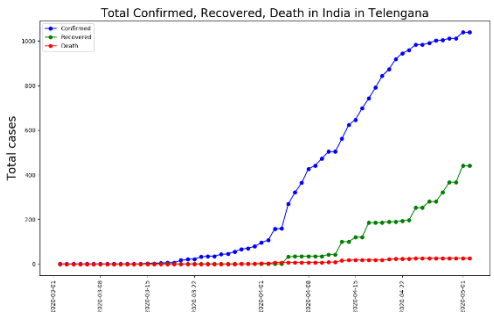


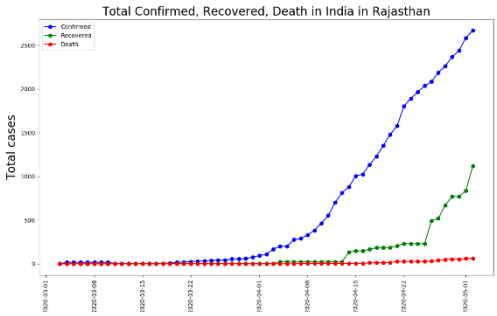



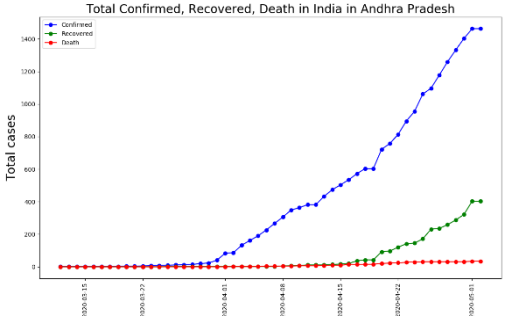



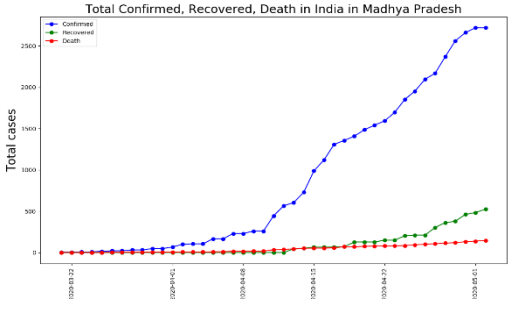



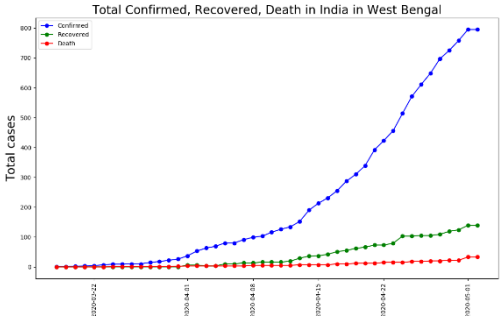



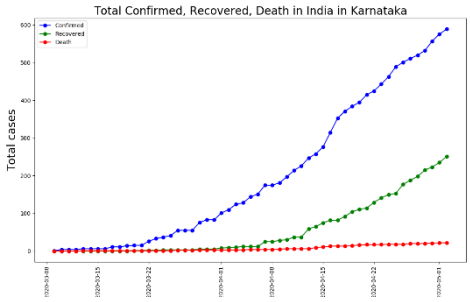



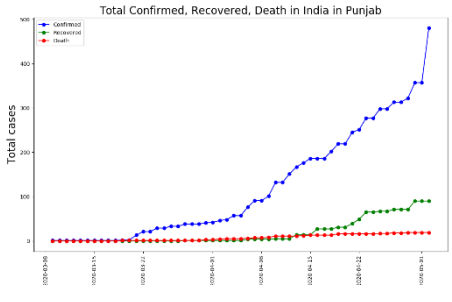



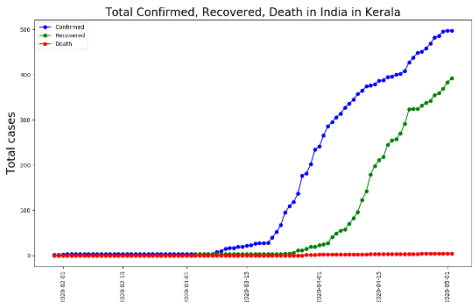



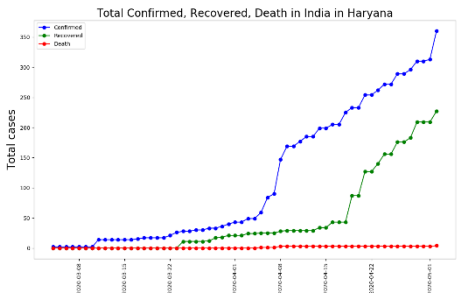



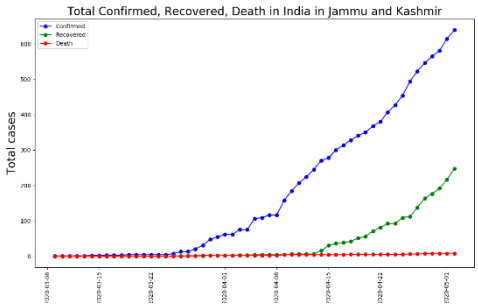



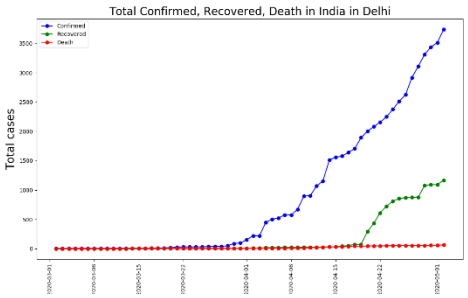

Total Confirmed, Recovered, Death in India in Karnataka
Confirmed
Recovered
Death
Total cases

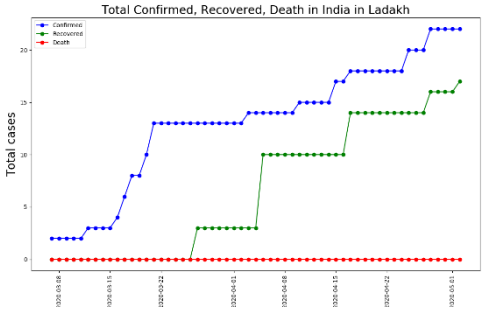

Total Confirmed, Recovered, Death in India in Punjab
Confirmed
Recovered
Death
Total cases

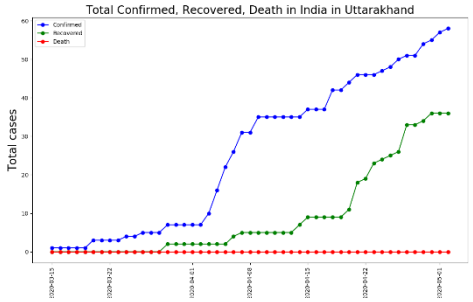

Total Confirmed, Recovered, Death in India in Kerala
Confirmed
Recovered
Death
Total cases

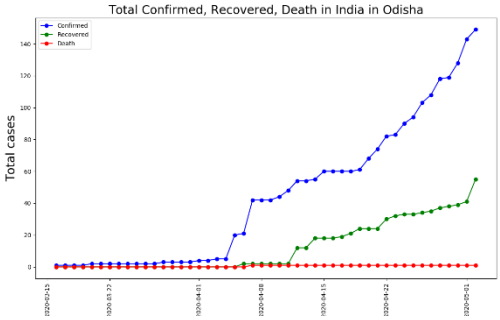

Total Confirmed, Recovered, Death in India in Haryana
Confirmed
Recovered
Death
Total cases

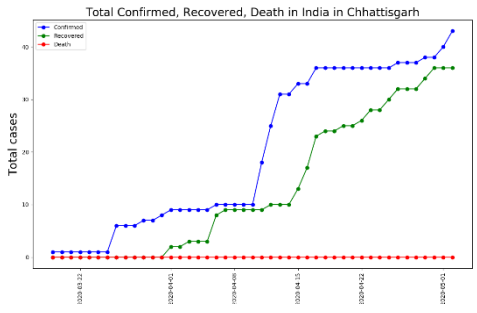

Total Confirmed, Recovered, Death in India in Jammu and Kashmir
Confirmed
Recovered
Death
Total cases

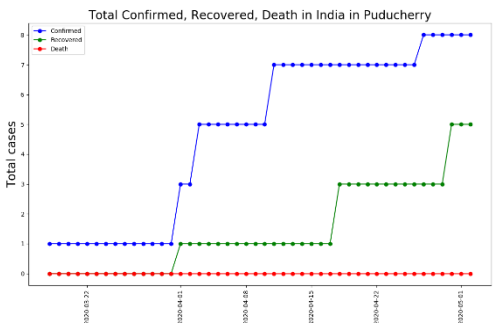

Total Confirmed, Recovered, Death in India in Delhi
Confirmed
Recovered
Death
Total cases

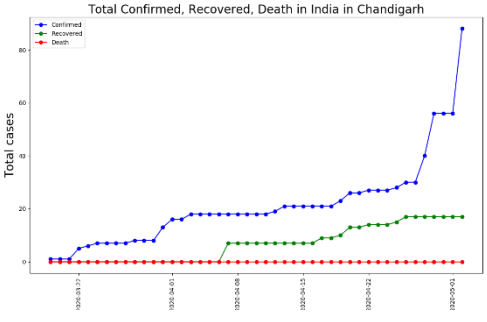

Total Confirmed, Recovered, Death in India in Ladakh
Confirmed
Recovered
Death
Total cases

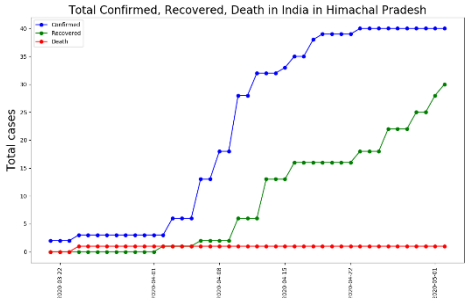

Total Confirmed, Recovered, Death in India in Uttarakhand
Confirmed
Recovered
Death
Total cases

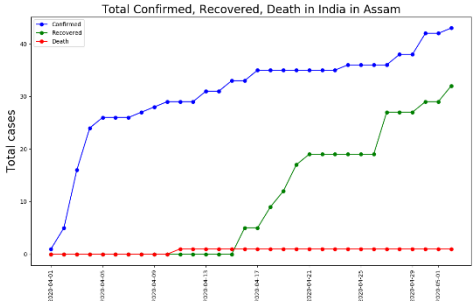

Total Confirmed, Recovered, Death in India in Odisha
Confirmed
Recovered
Death
Total cases

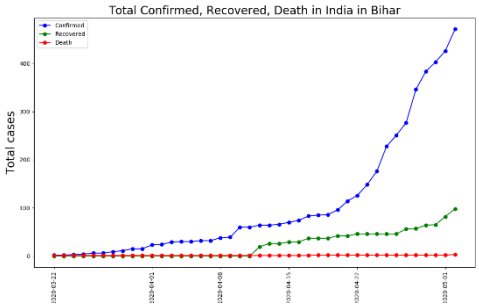

Total Confirmed, Recovered, Death in India in Chhattisgarh
Confirmed
Recovered
Death
Total cases

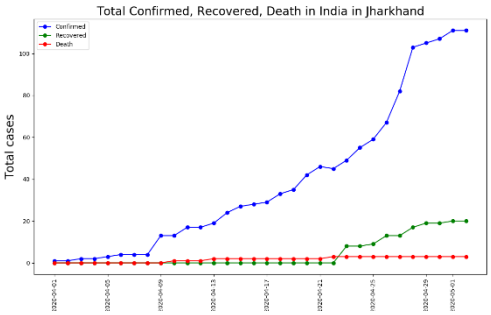

Total Confirmed, Recovered, Death in India in Puducherry
Confirmed
Recovered
Death
Total cases

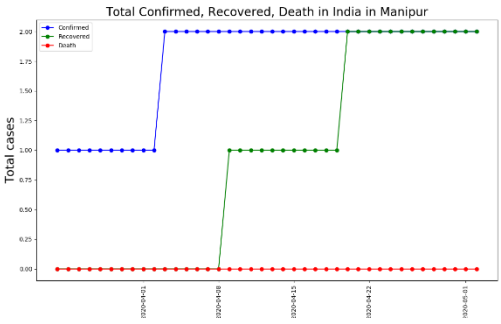

Total Confirmed, Recovered, Death in India in Chandigarh
Confirmed
Recovered
Death
Total cases

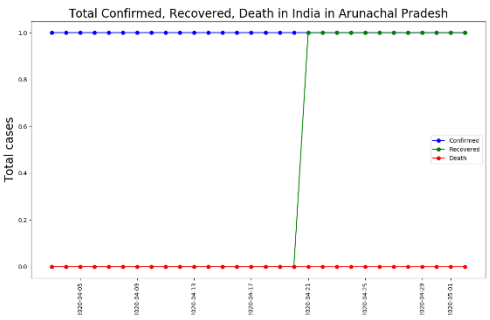

Total Confirmed, Recovered, Death in India in Himachal Pradesh
Confirmed
Recovered
Death
Total cases

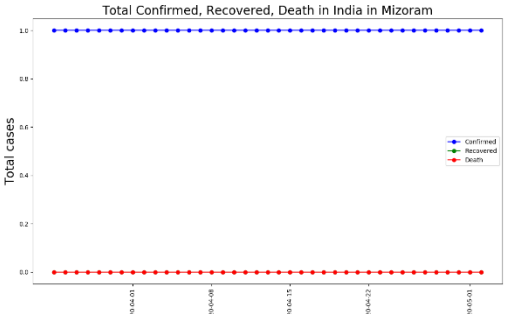

Total Confirmed, Recovered, Death in India in Assam
Confirmed
Recovered
Death
Total cases

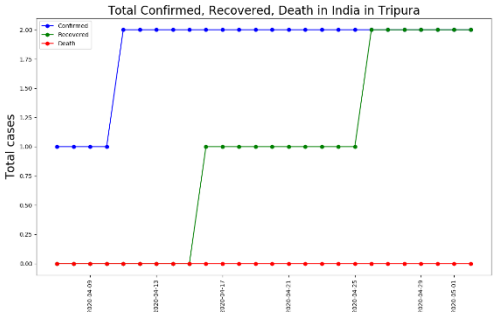

Total Confirmed, Recovered, Death in India in Bihar
Confirmed
Recovered
Death
Total cases

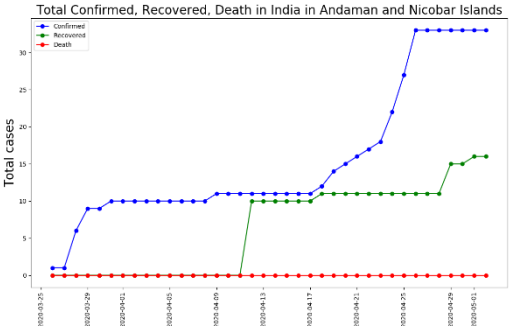

Total Confirmed, Recovered, Death in India in Jharkhand
Confirmed
Recovered
Death
Total cases

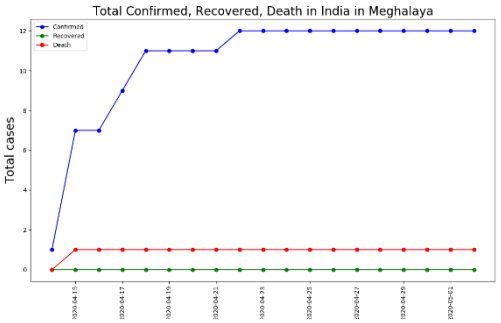

Total Confirmed, Recovered, Death in India in Manipur
Confirmed
Recovered
Death
Total cases

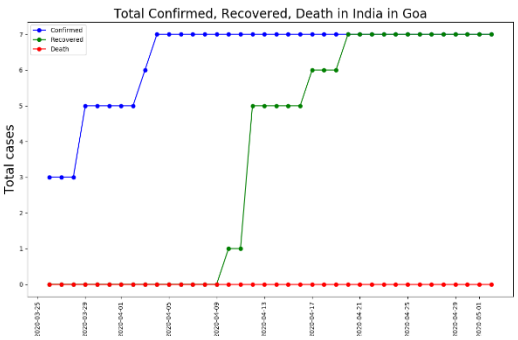

Total Confirmed, Recovered, Death in India in Arunachal Pradesh
Confirmed
Recovered
Death
Total cases

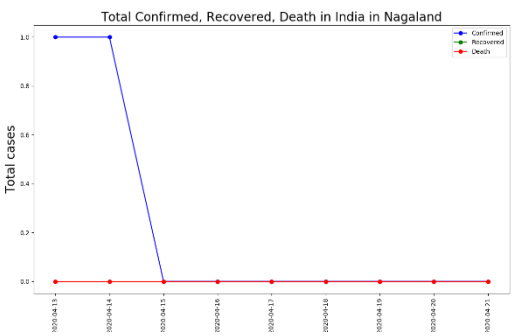

Total Confirmed, Recovered, Death in India in Mizoram
Confirmed
Recovered
Death
Total cases


Total Confirmed, Recovered, Death in India in Tripura
Confirmed
Recovered
Death
Total cases


Total Confirmed, Recovered, Death in India in Andaman and Nicobar Islands
Confirmed
Recovered
Death
Total cases


Total Confirmed, Recovered, Death in India in Meghalaya
Confirmed
Recovered
Death
Total cases

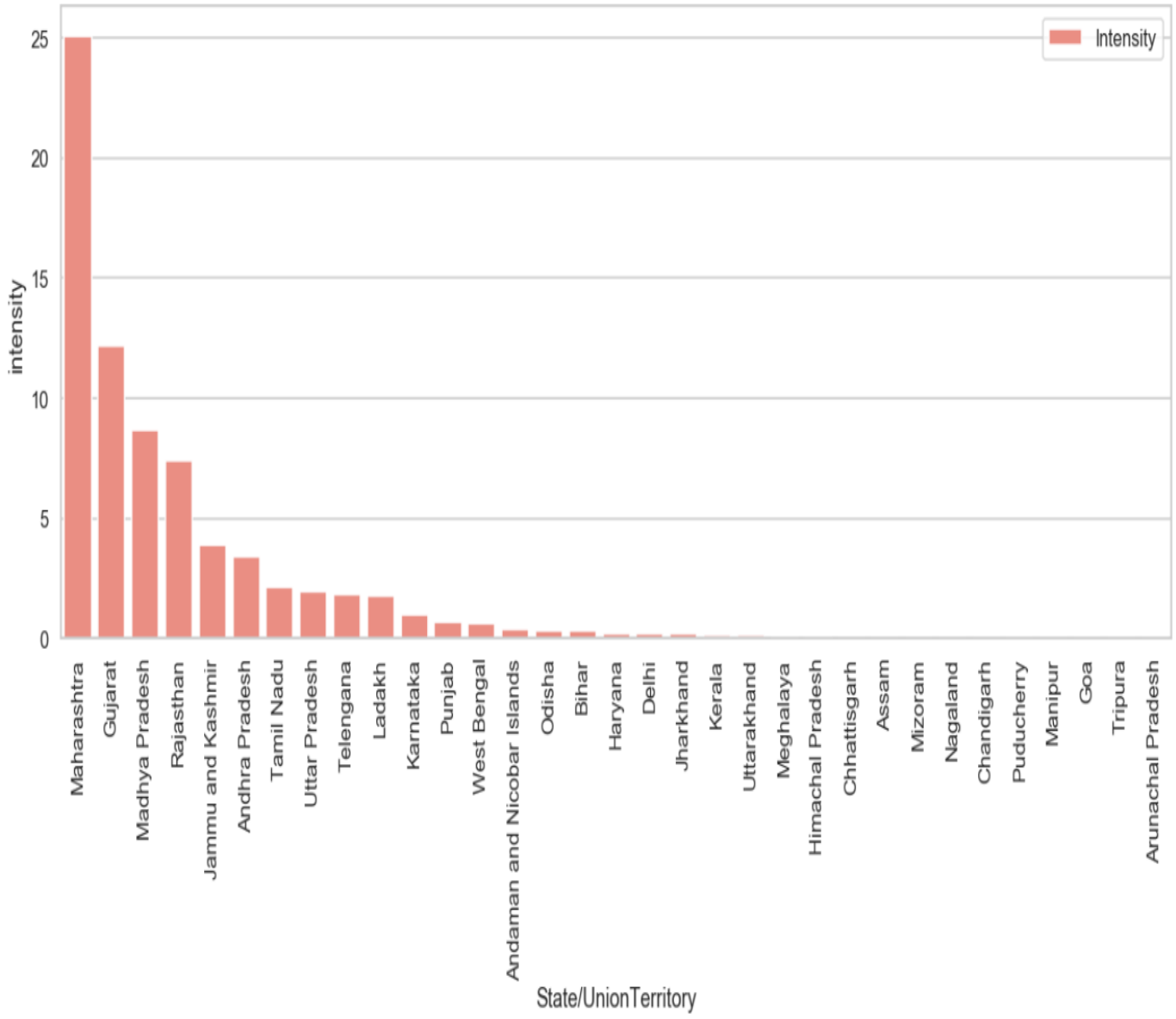

Total Confirmed, Recovered, Death in India in Goa
Confirmed
Recovered
Death
Total cases


Total Confirmed, Recovered, Death in India in Nagaland
Confirmed
Recovered
Death
Total cases

The above plots provide the same statistic of confirmed, recovered and deceased cases but in a state-wise sense and each state has a separate plot.

---

## Q3.




**FORMULA:**

Positive cases/active cases = Confirmed cases – Deaths – Cured

Intensity (in the state) = no. of active cases (in the state)/ population density (of state)

*(The units of population density were taken in people/km^2 but intensity used only as comparison parameter predominantly)*

**ASSUMPTIONS:**

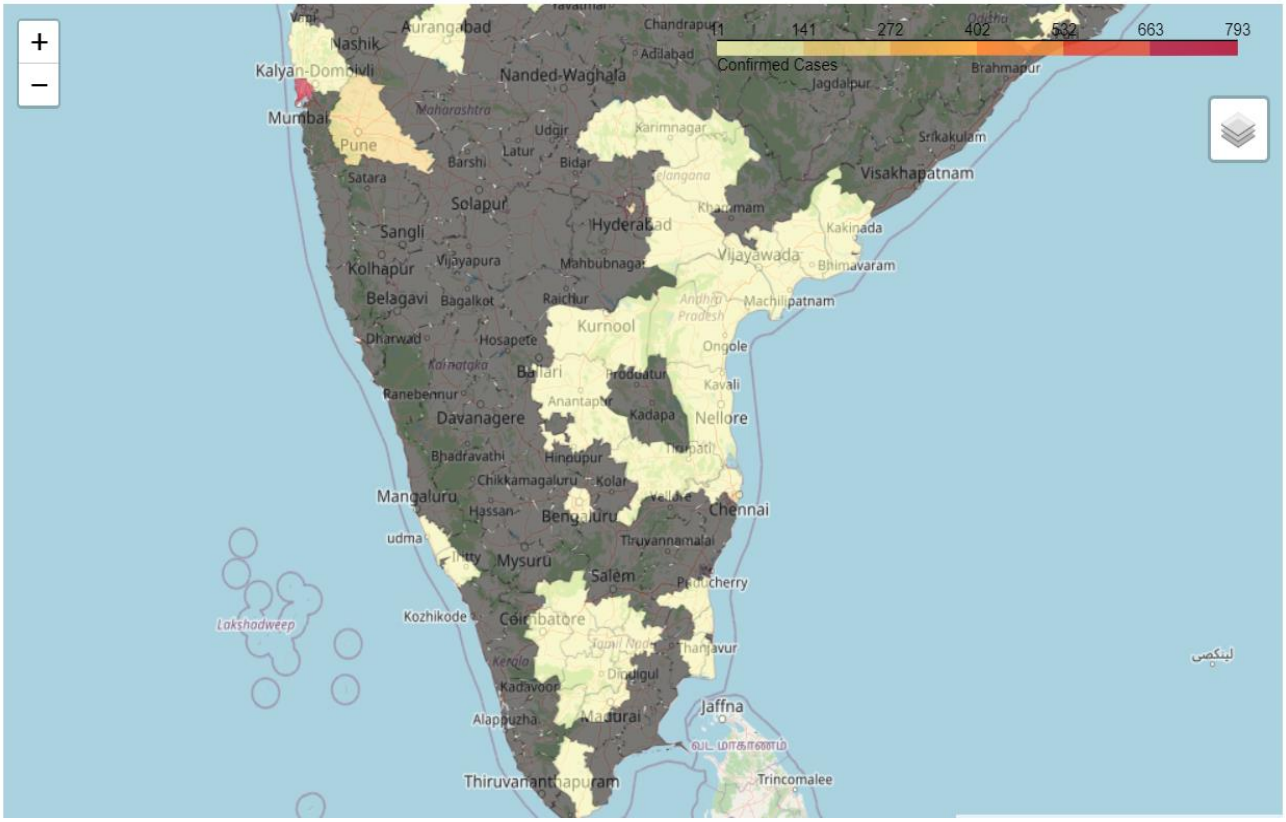
1. Confirmed cases include all cases that have tested positive including the deceased and cured.

Clearly the state of Maharashtra is the state with the most intensity.

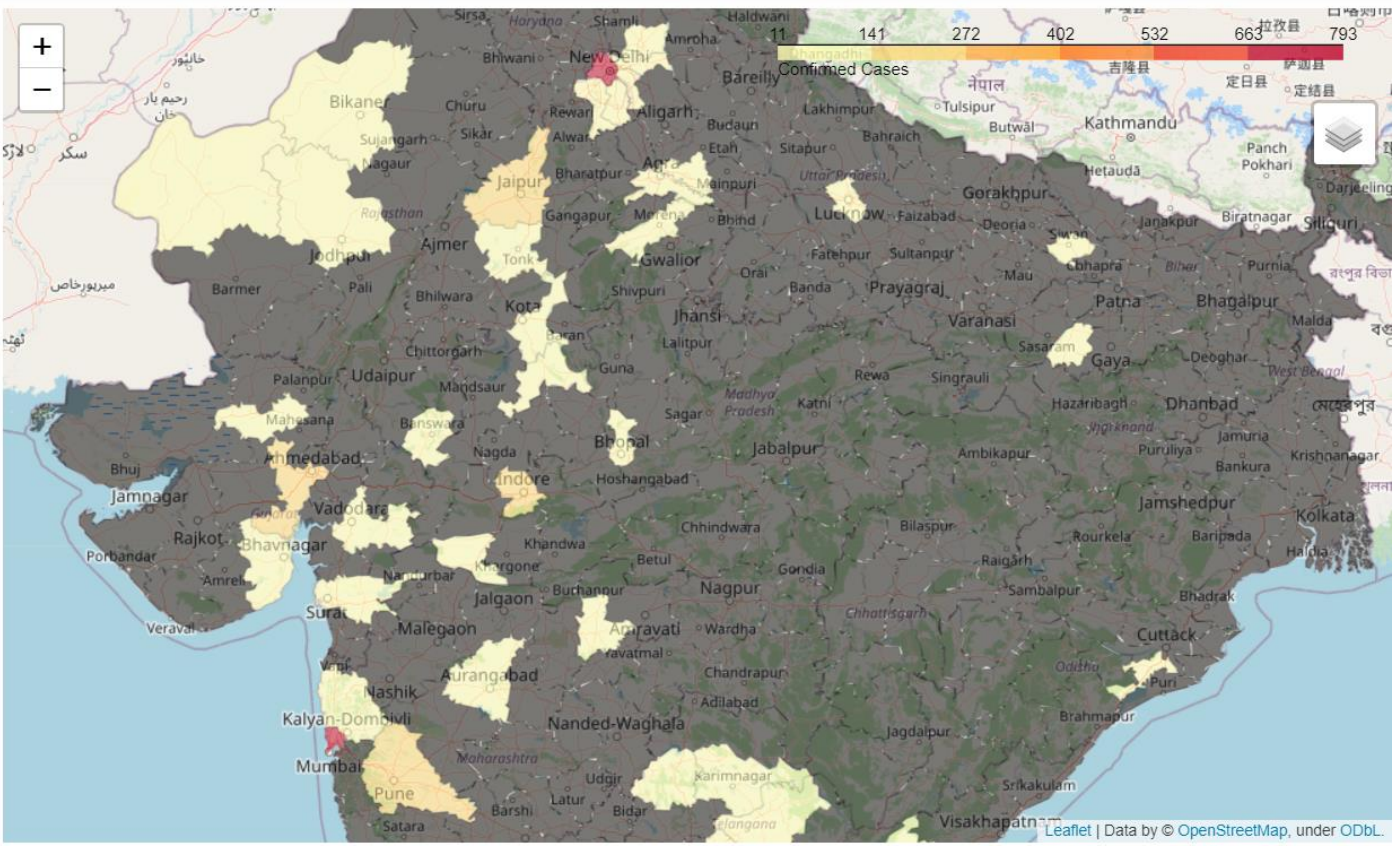
---

**Q4**.

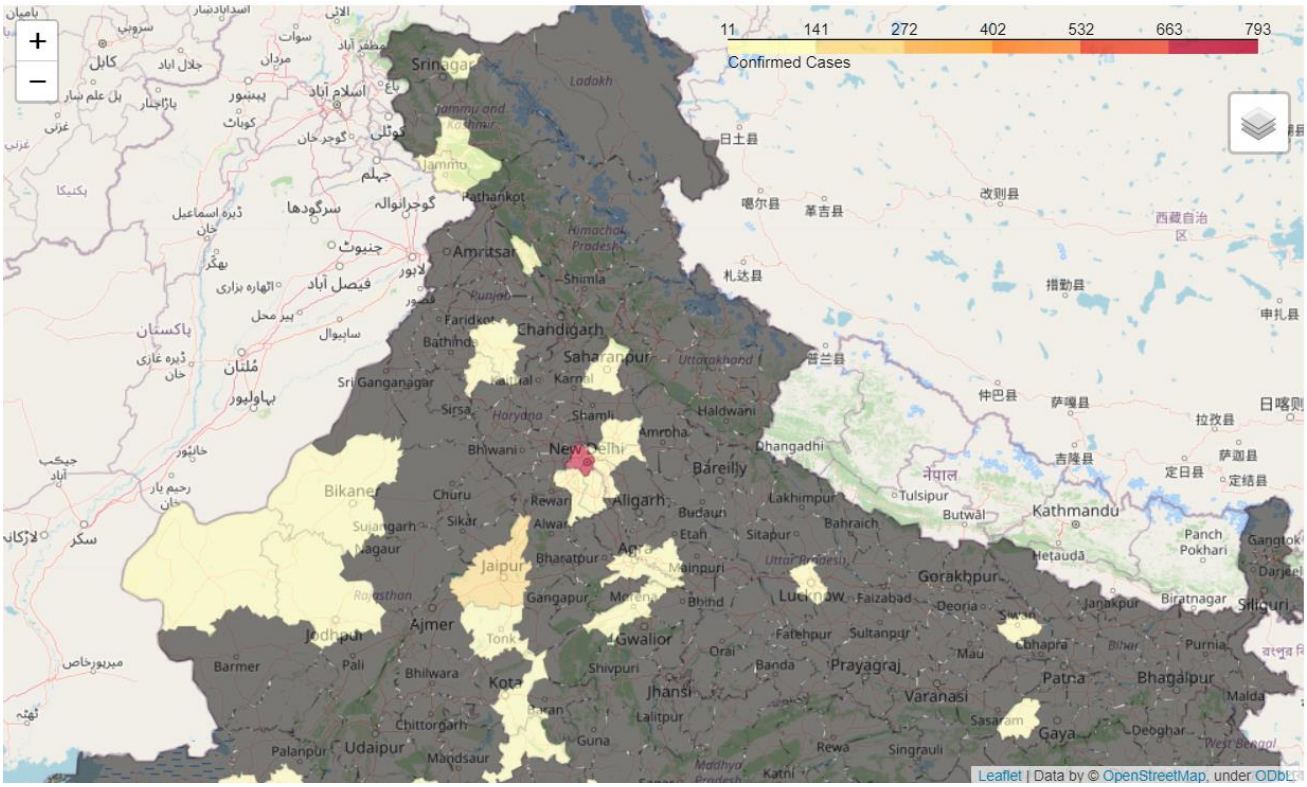
Southern Region: (Plot of Hotspots)



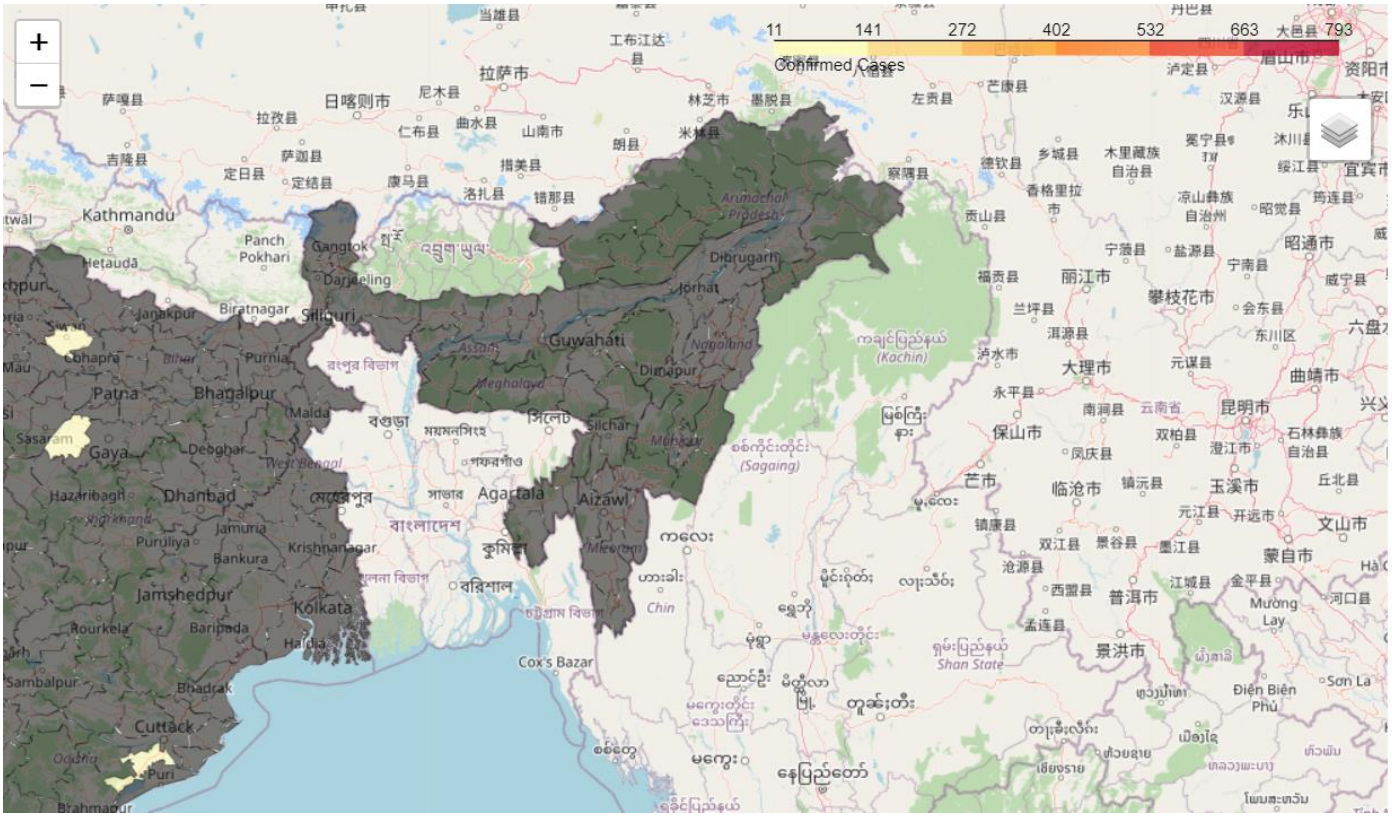

Central & Western Region: (Plot of Hotspots)

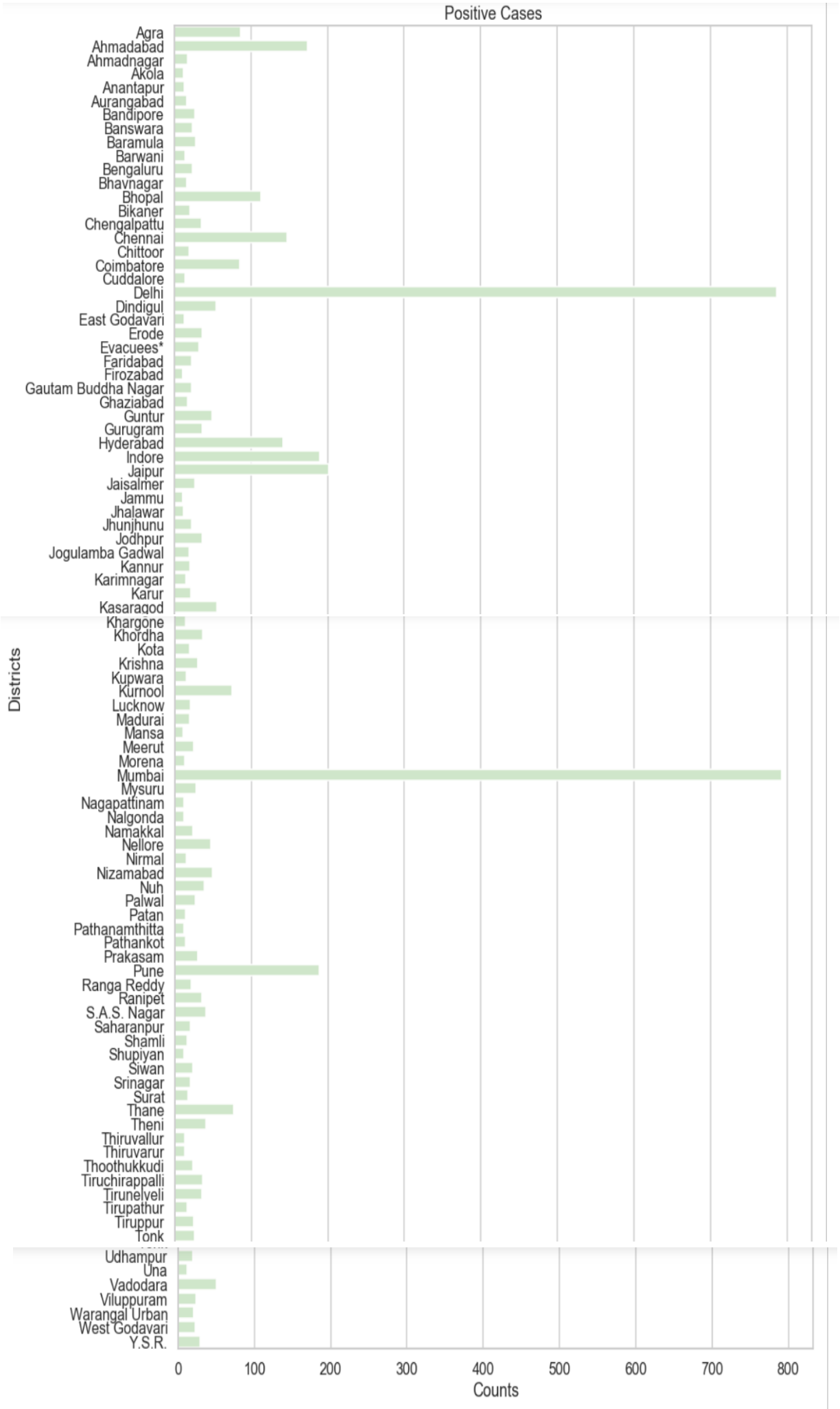


Nothern Region: (Plot of Hotspots)




Eastern Region: (plot of hotspots)




Positive Cases
Districts
Agra
Ahmadabad
Ahmadnagar
Akola
Anantapur
Aurangabad
Bandipore
Banswara
Baramula
Barwani
Bengaluru
Bhavnagar
Bhopal
Bikaner
Chengalpattu
Chennai
Chittoor
Coimbatore
Cuddalore
Delhi
Dindigul
East Godavari
Erode
Evacuees*
Faridabad
Firozabad
Gautam Buddha Nagar
Ghaziabad
Guntur
Gurugram
Hyderabad
Indore
Jaipur
Jaisalmer
Jammu
Jhalawar
Jhunjhunu
Jodhpur
Jogulamba Gadwal
Kannur
Karimnagar
Karur
Kasaragod
Khargone
Khordha
Kota
Krishna
Kupwara
Kurnool
Lucknow
Madurai
Mansa
Meerut
Morena
Mumbai
Mysuru
Nagapattinam
Nalgonda
Namakkal
Nellore
Nirmal
Nizamabad
Nuh
Palwal
Patan
Pathanamthitta
Pathankot
Prakasam
Pune
Ranga Reddy
Ranipet
S.A.S. Nagar
Saharanpur
Shamli
Shupiyan
Siwan
Srinagar
Surat
Thane
Theni
Thiruvallur
Thiruvarur
Thoothukkudi
Tiruchirappalli
Tirunelveli
Tirupathur
Tiruppur
Tonk
Udhampur
Una
Vadodara
Viluppuram
Warangal Urban
West Godavari
Y.S.R.
0
100
200
300
400
500
600
700
800
Counts

This is the list of all districts that are hotspots and the corresponding no. of people tested positive in these regions as of 10/04/2020.

HOTSPOT:

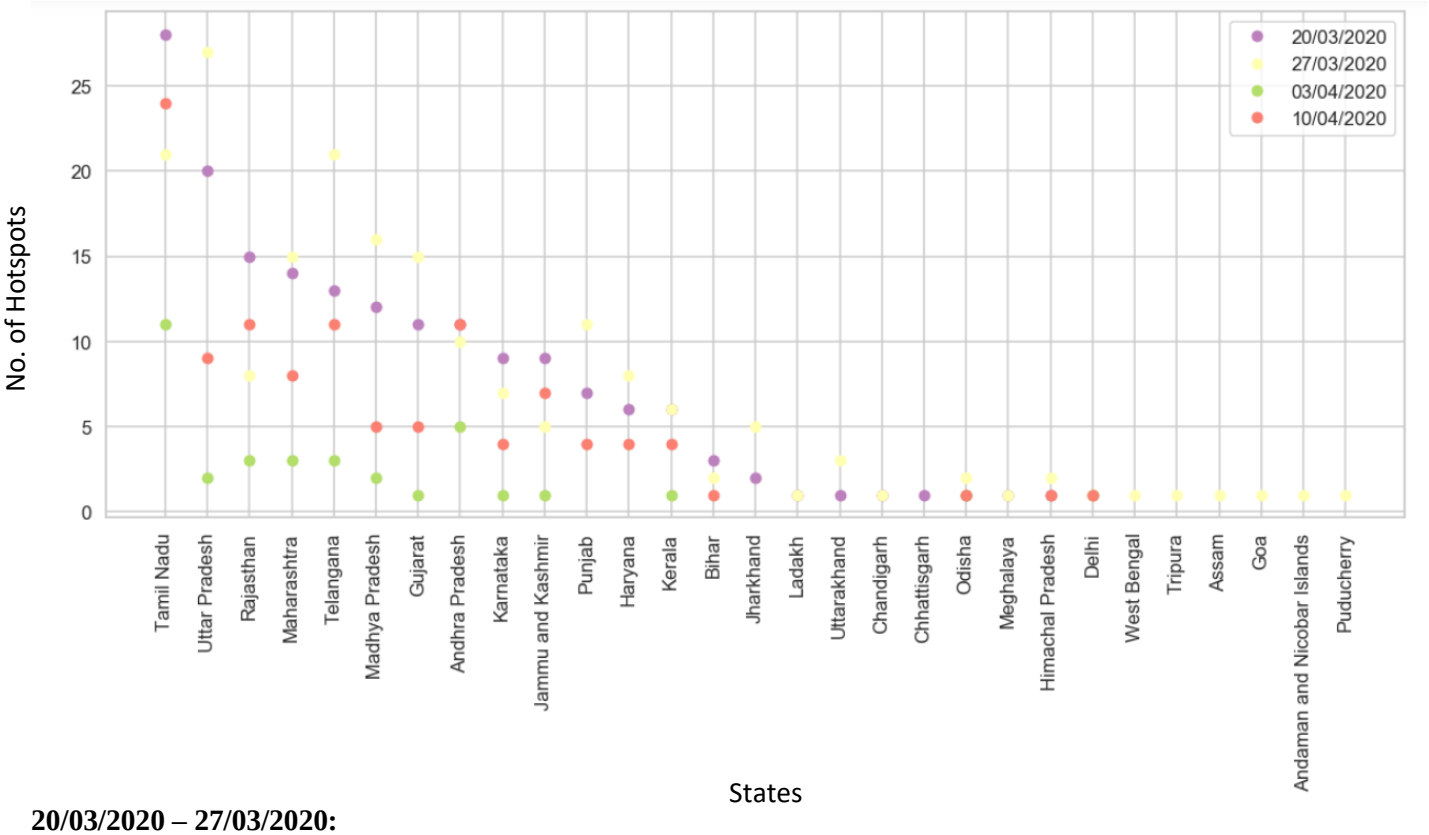
A hotspot is defined as that region in a city where the number of active cases is greater than 10.

ASSUMPTIONS:

1. Since the data only provides status of Hospitalized, Recovered or Deceased we assumed that all those whose status has remained Hospitalized till the date 10/04/2020 are considered as Active cases.

---

**Q5.** This plot looks at the number of hotspots in each state over a 3-week period from 20/03/2020 to 10/04/2020




**20/03/2020 – 27/03/2020:**

**States with maximal Change in Number of Hotspots (Increase):**

**Telangana:** (Increase in 8 Hotspots)

20/03/2020 – 13 Hotspots | 27/03/2020 – 21 Hotspots

**Uttar Pradesh:** (Increase in 7 Hotspots)

20/03/2020 – 20 Hotspots | 27/03/2020 – 27 Hotspots

**Madhya Pradesh:** (Increase in 4 Hotspots)

20/03/2020 – 12 Hotspots | 27/03/2020 – 16 Hotspots

**Gujarat:** (Increase in 4 Hotspots)

20/03/2020 – 11 Hotspots | 27/03/2020 – 15 Hotspots

**States with maximal Change in Number of Hotspots (Decrease):**

**Rajasthan:** (Decrease in 7 Hotspots)

20/03/2020 – 15 Hotspots | 27/03/2020 – 8 Hotspots

**Tamil Nadu:** (Decrease in 7 Hotspots)

20/03/2020 – 28 Hotspots | 27/03/2020 – 21 Hotspots

**Jammu & Kashmir:** (Decrease in 4 Hotspots)

20/03/2020 – 9 Hotspots | 27/03/2020 – 5 Hotspots

**27/03/2020 – 03/04/2020:**

**States with maximal Change in Number of Hotspots (Increase):**

No State has had an increase in the number of Hotspots in this period.

**States with maximal Change in Number of Hotspots (Decrease):**

**Uttar Pradesh:** (Decrease in 25 Hotspots)

27/03/2020 – 27 Hotspots | 03/04/2020 – 2 Hotspots

**Telangana:** (Decrease in 18 Hotspots)

27/03/2020 – 21 Hotspots | 03/04/2020 – 3 Hotspots

**Madhya Pradesh:** (Decrease in 14 Hotspots)

27/03/2020 – 16 Hotspots | 03/04/2020 – 2 Hotspots

**Gujarat:** (Decrease in 14 Hotspots)

27/03/2020 – 15 Hotspots | 03/04/2020 – 1 Hotspots

**Maharashtra:** (Decrease in 12 Hotspots)

27/03/2020 – 15 Hotspots | 03/04/2020 – 3 Hotspots

**03/04/2020 – 10/04/2020**

**States with maximal Change in Number of Hotspots (Increase):**

**Tamil Nadu:** (Increase in 10 Hotspots)

03/04/2020 – 11 Hotspots | 27/03/2020 – 21 Hotspots

**Rajasthan:** (Increase in 8 Hotspots)

03/04/2020 – 3 Hotspots | 27/03/2020 – 11 Hotspots

**Telangana:** (Increase in 8 Hotspots)

03/04/2020 – 3 Hotspots | 27/03/2020 – 11 Hotspots

**States with maximal Change in Number of Hotspots (Decrease):**

No State has had a decrease in the number of Hotspots in the period.

(Q6) The goal here is to classify the cases from the top 5 states into three categories: (1) with international travel history (primary), (2) personal contact with primary case (secondary) and (3) cases that do not fall into primary or secondary (tertiary).

**Datafile** – IndividualDetails.csv

*Step 1:* Read the csv file.

*Step 2:* Use the column 'detected_state' to plot the number of cases w.r.t. each state

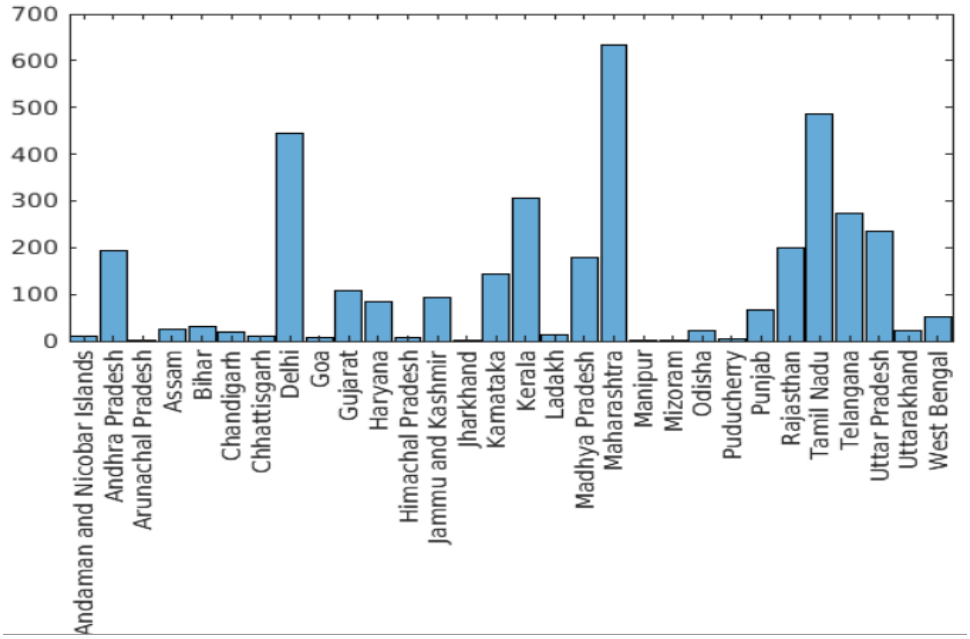


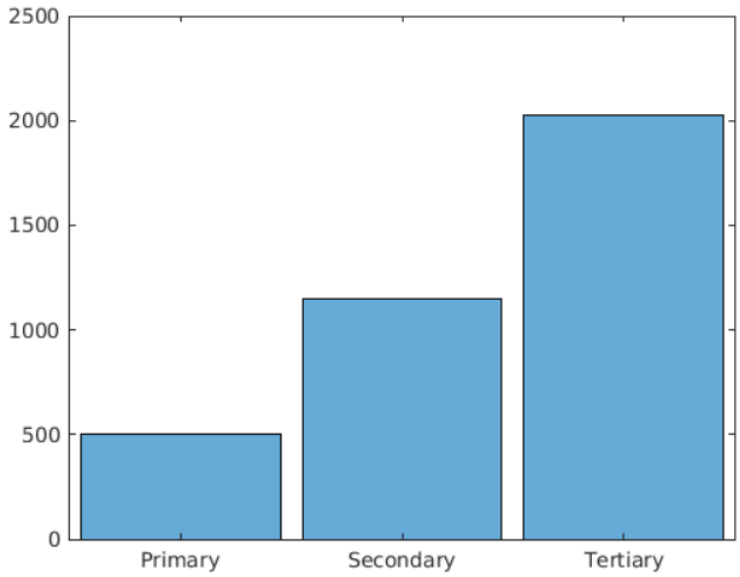

Figure 1: Number of cases vs. States

| | |
|---|---|
| 1 | Maharashtra |
| 2 | Tamil Nadu |
| 3 | Delhi |
| 4 | Kerala |
| 5 | Telangana |

Figure 2: Top 5 States

*Step 3:* Filter the top 5 states with highest number of cases from Step 2.

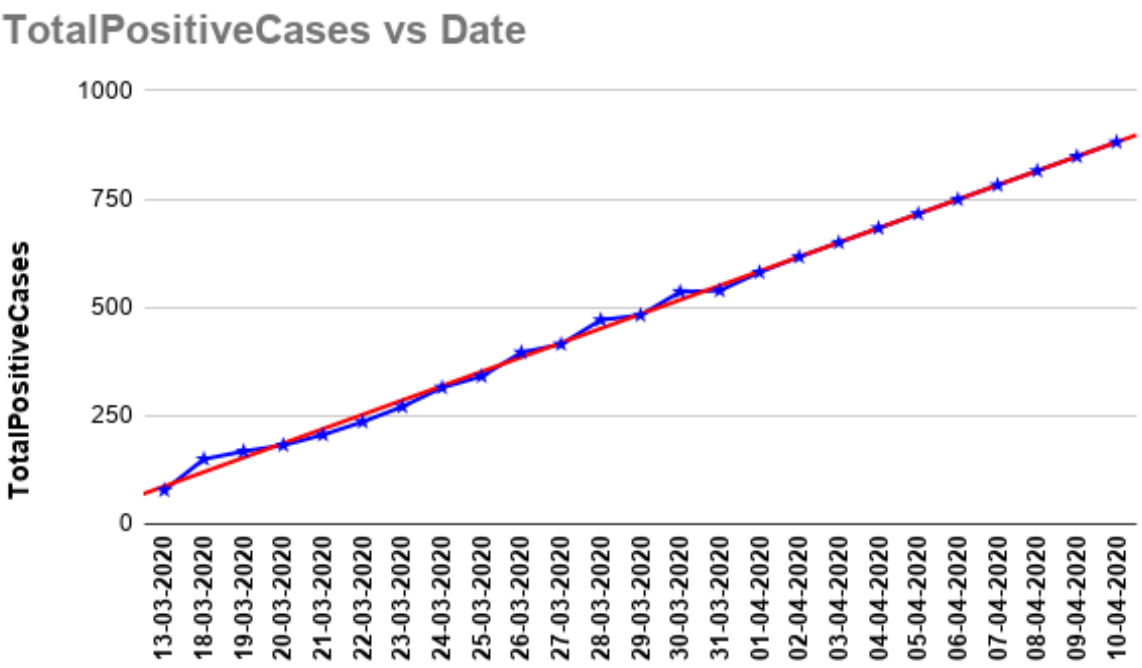
*Step 4:* For these 5 states, use the 'notes' column to identify primary, secondary or tertiary. Plot the count as a percentage plot.

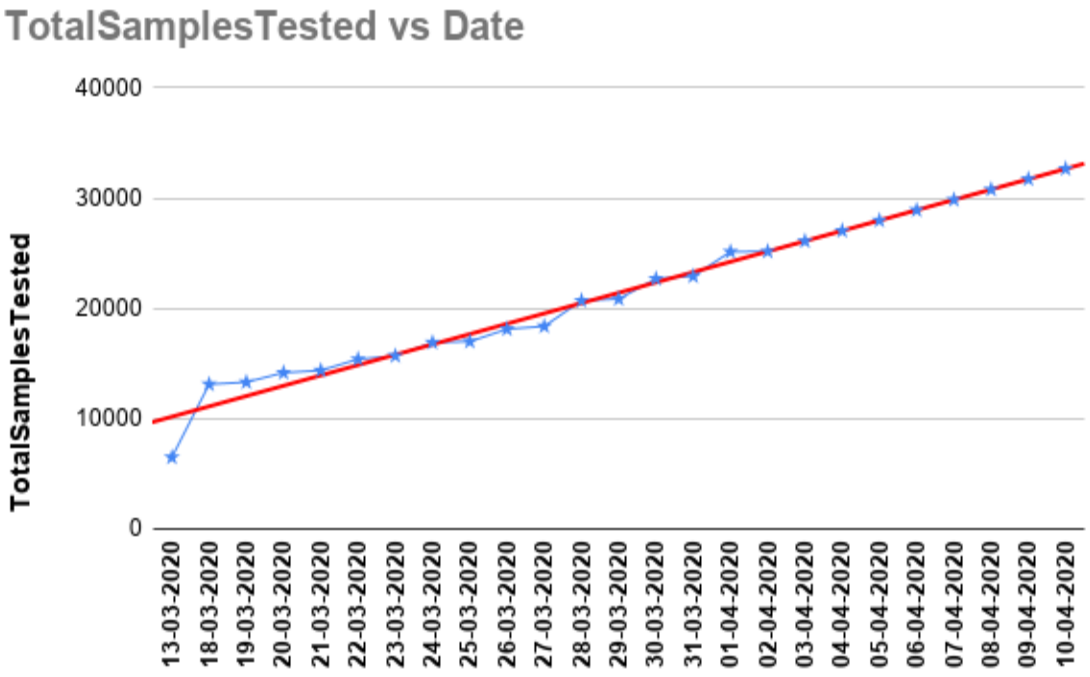


(Q7) The goal here is to find the number of additional labs required from the current existing labs with an increase rate of 10% cases per day from 11.04 to 10.04 (Assume 100 tests per lab per day)

**Datafile**: ICMRTestingDetails.csv

*Step 1:* The datafile contains data only up to 01.04. There is no available data from 02.04 to 10.04 for the number of samples tested nor for the number of positive cases. Therefore, we use the data available from 13.03 to 01.04 and extrapolate it to 10.04 (We assume that the number of positive cases and the number of samples tested will follow a similar trend during the period from 02.04 to 10.04)



*Step 2:* From 11.04 to 20.04, the number of cases increase 10% per day (given in question). We can therefore compute the number of cases using the formula, #cases on 11.04 = 1.1 * #cases on 10.04 etc.

*Step 3:* From 11.04 to 20.04, we need to predict the number of samples tested. For this we use the data from 13.03 to 10.04. It can be observed that both the number of positive cases and the number of samples tested increase by roughly 3% each day from 13.03 to 10.04. Since it is given that the number of cases increase by 10% each day from 11.04 to 20.04, we assume that the number of samples tested also increase similarly. We now obtain the data on the number of samples tested.

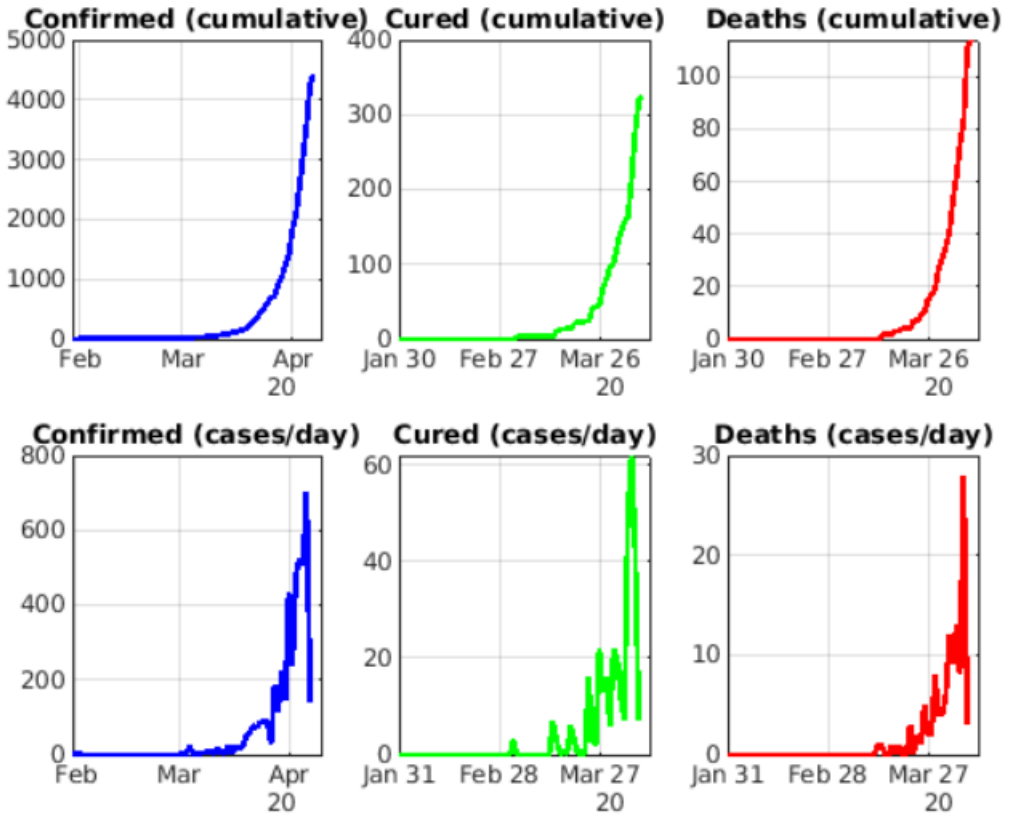
*Step 4:* Using the data obtained from Step 3, we compute the number of additional labs needed. The entire data is saved in Q7.csv. The number of additional labs needed each day are also listed here:

| | | | |
|---|---|---|---|
| 11-04-2020 | 33 | 16-04-2020 | 53 |
| 12-04-2020 | 36 | 17-04-2020 | 58 |
| 13-04-2020 | 40 | 18-04-2020 | 64 |
| 14-04-2020 | 43 | 19-04-2020 | 70 |
| 15-04-2020 | 48 | 20-04-2020 | 77 |

(Q8) Goal – Plot the number of cases from 01.03 to 10.04

Datafile: covid_19_india.csv

We use the data of the confirmed, cured and death cases and plot them against the dates (both cumulative and per each day)



On the notion of ‘flattening of the curve’, based on the data available and the plots till 10 April, the curve seems to be on the rise and the number of cases appear to only increase with time (at least for the near future (from 10 April)).

(Q9) Based on the data plots from Q8 above, it can be seen that although the curve for the number of cases seems to increase and not flatten anytime soon, it could have been much worse (the increase could have been much higher), if not for social distancing. These plots are during the time where social distancing is implemented. And it can be seen from the graphs that the number of new cases being recorded per day is decreasing (See the curve in blue after April 2020). This only goes to prove the point that social distancing is the only option we have currently to avoid the spread.