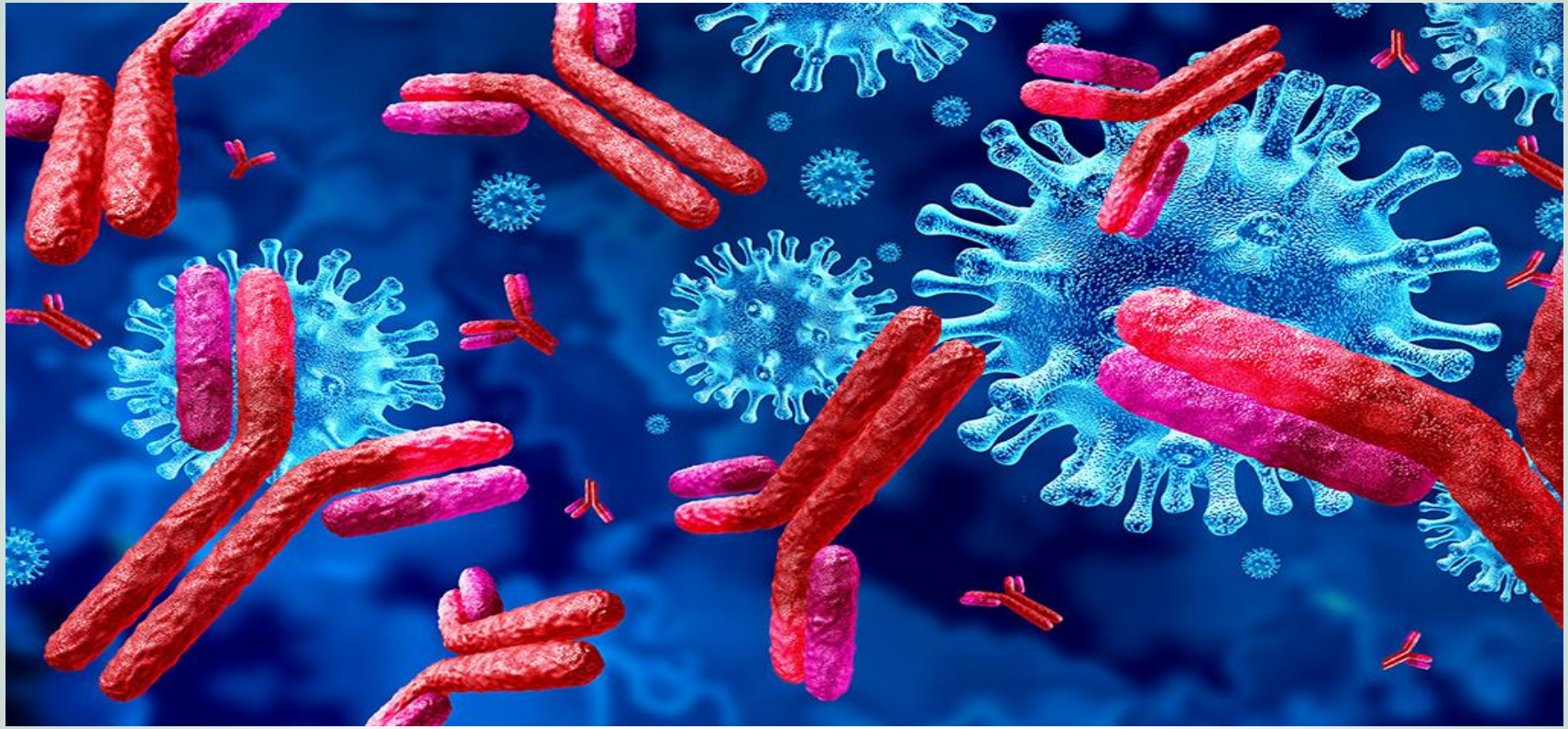
COVID-19 INDIA ANALYSIS



Prachi Gorwadkar

September 2021

Introduction

Coronavirus (COVID-19) in India

- The coronavirus outbreak has become one of the most notorious events of the decade. Coronavirus disease caused by severe acute respiratory syndrome. India reported the first confirmed case of the coronavirus infection on 30 January 2020 in the state of Kerala. The affected had a travel history from Wuhan, China. New confirmed cases are being reported in multiple cities such as New Delhi, Mumbai, Bengaluru, Hyderabad, and Patna.
- Despite a strong response at the outset of the pandemic, as of Sept 22, India has the world's fastest growing outbreak of COVID-19 in absolute numbers according to WHO, reporting more than 5·6 million infections. Restrictions began to be lifted in June, and this relaxation has continued in the face of a continuing dramatic increase in case numbers nationally.
- Even so, India is clearly facing a dangerous period. The country has responded well in many regards, especially for such a large and diverse nation.
- India instigated a national lockdown in March, which was praised by WHO.

CONTENTS

- ✓ Project Approach / Objectives
- ✓ What does the DATA tell us?
- ✓ Visualisations of the Data
- ✓ Conclusions / Notes

Objectives

Project Approach :-

Task Details Following are the inspirations to work in this dataset :

- Visualizing Covid-19 data in different ways.
- What are the Top-N States / Top-N districts affected by the disease?
- Out of different age groups, which of them were the most vulnerable to the disease?
- Which regions are more affected by the virus? / How many states are fully vaccinated?
- What is the number of positive cases rising in India?
- What are the Top-N States with the highest Fatality Ratio.

So, here I analyze the data to show the insights regarding coronavirus.

Data Analysis

Lets go for an EDA on the dataset.

❖ DATA SETS USED :-

1. covid_19_india.csv - This file gives the date-wise details of all states affected .
2. StatewiseTestingDetails.csv - This file gives the testing details of India & its states with dates.
3. covid_vaccine_statewise.csv - This file gives the vaccination details of India & its states with dates.
4. district_level_latest.csv - This file gives the details of all states affected & its related districts.

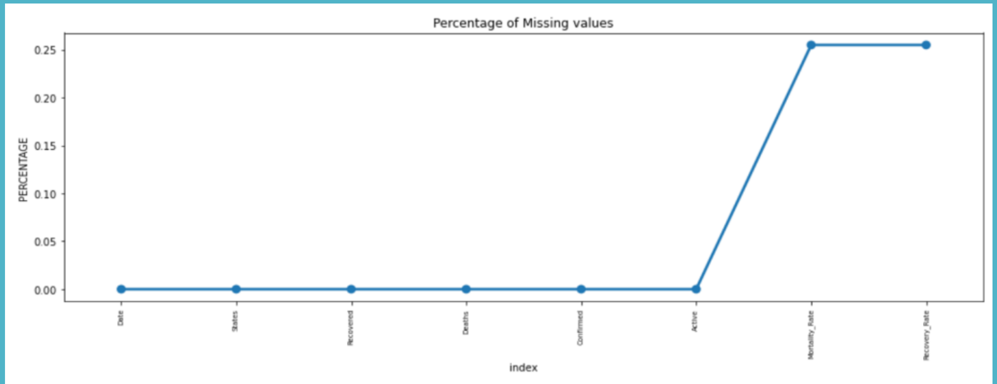
❖ Steps involved in the EDA :-

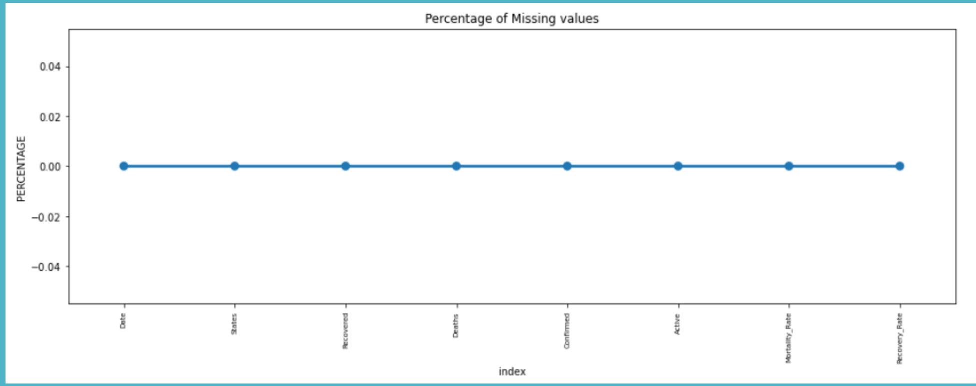
I. Data Sourcing
II. Data Cleaning
III. Data Processing
IV. Data Visualisations
V. Conclusions from the Visualisations

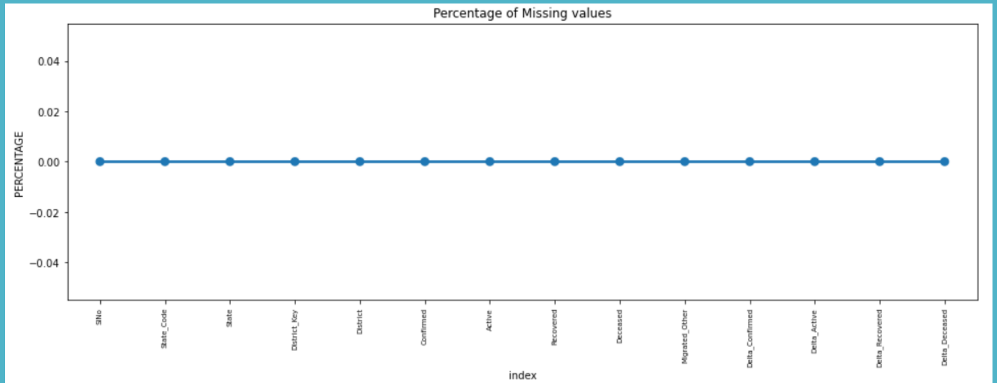
Missing DATA Treatment

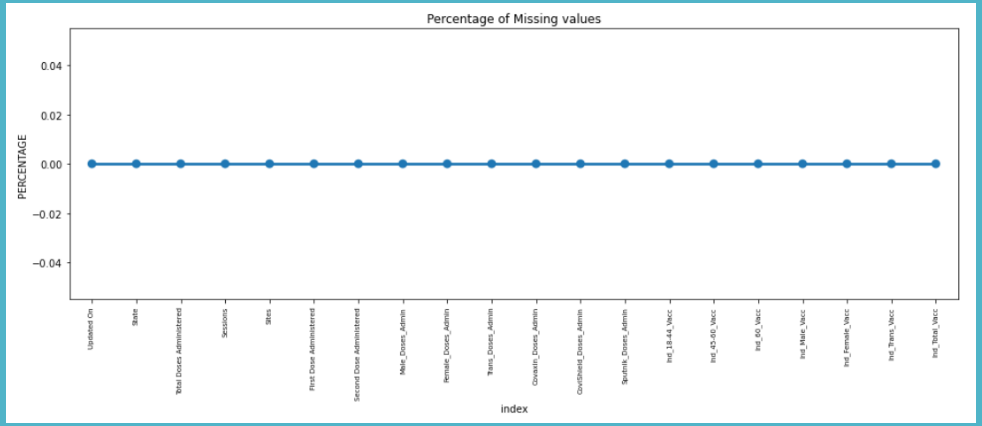
For 2 files, I need to work on Missing/NULL values.

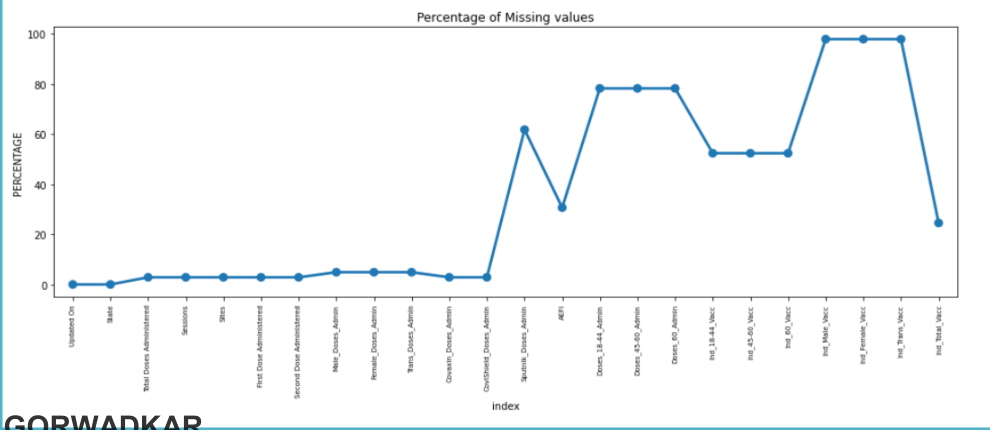
And for other 2 files, I already had the clean Data.

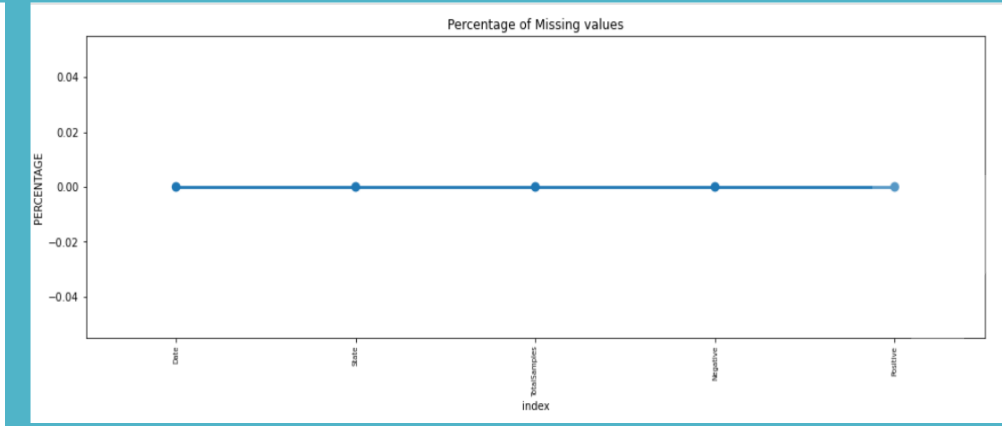










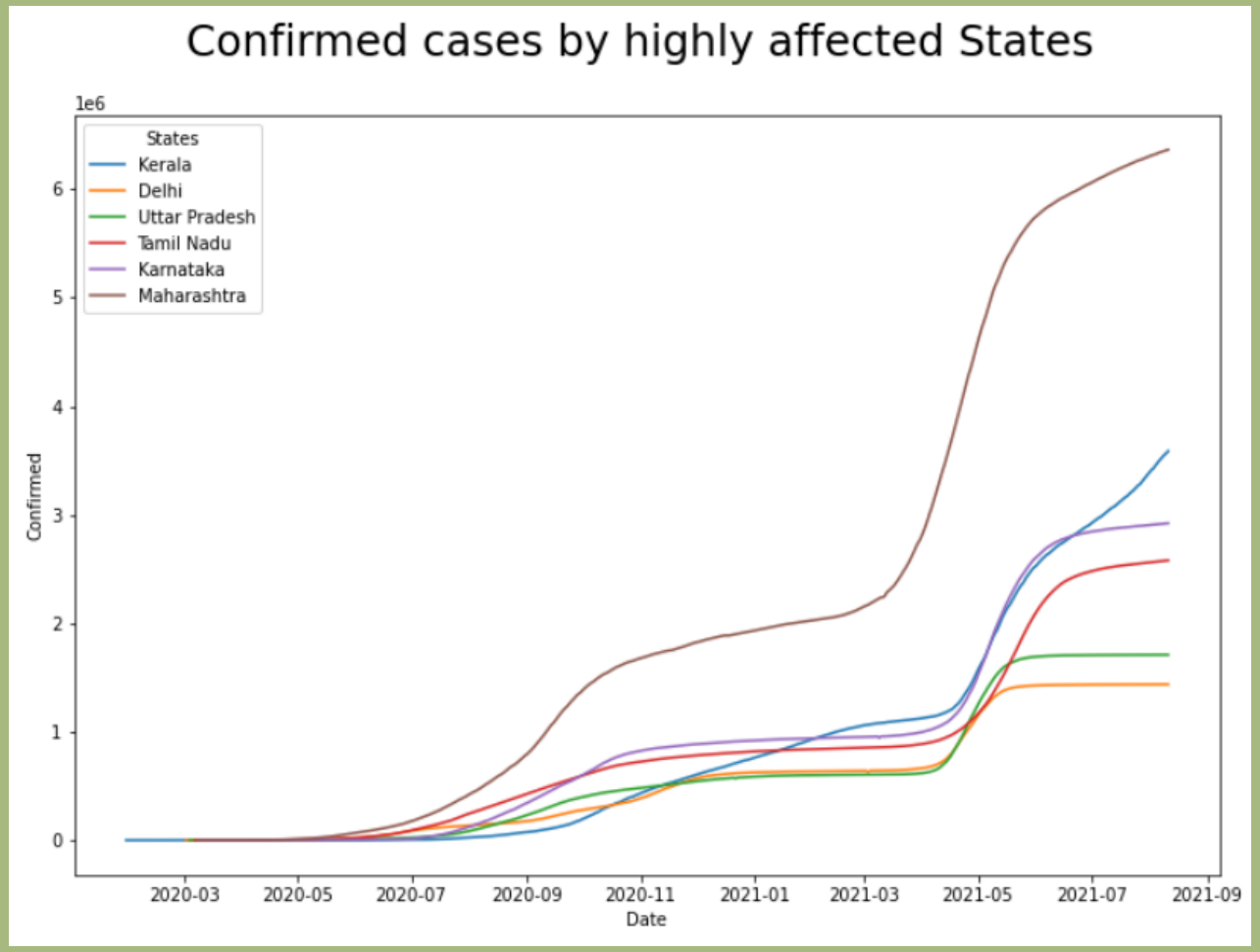


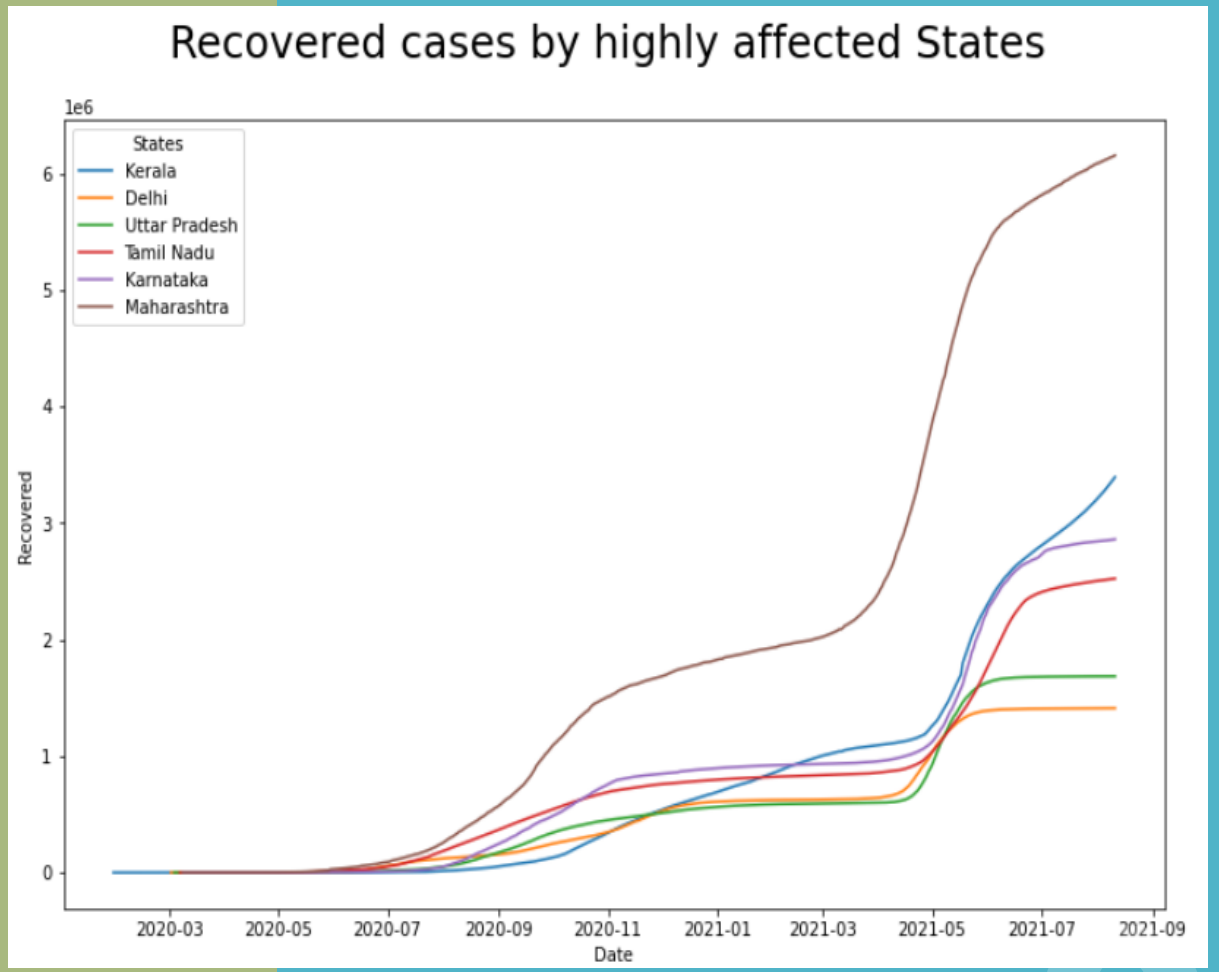
State-wise Data for Confirmed / Recovered / Death Cases and the related ratios

- Maharashtra tops this list in all aspects excluding the Mortality rate.
- Considering the number of cases Kerala, Goa, Manipur, Tripura, Arunachal Pradesh, Mizoram, Ladakh and Andaman and Nicobar Islands has done a great job when it comes to recovery rate.
- Similarly in terms of Mortality rate Bihar and Punjab seem to have higher numbers.
- Though being second in the country w.r.t. total confirmed cases, Kerala seem to have established the healthcare facility to a higher level which is evident from the lower mortality rate (0.89%).

States	Confirmed	Deaths	Mortality_Rate	Recovered	Recovery_Rate
Maharashtra	6363442	134201	7.510000	6159676	96.800000
Kerala	3586693	18004	0.890000	3396184	100.000000
Karnataka	2921049	36848	16.670000	2861499	99.280000
Tamil Nadu	2579130	34367	3.850000	2524400	98.070000
Andhra Pradesh	1985182	13564	3.040000	1952736	99.130000
Uttar Pradesh	1708812	22775	3.110000	1685492	98.640000
West Bengal	1534999	18252	11.110000	1506532	98.150000
Delhi	1436852	25068	14.290000	1411280	99.750000
Chhattisgarh	1003356	13544	1.350000	988189	98.490000
Odisha	988997	6565	2.380000	972710	99.260000
Rajasthan	953851	8954	2.860000	944700	99.040000
Gujarat	825085	10077	9.520000	814802	98.750000
Madhya Pradesh	791980	10514	7.450000	781330	99.230000
Haryana	770114	9652	2.040000	759790	98.660000
Bihar	725279	9646	50.000000	715352	99.890000
Telangana	650353	3831	4.430000	638410	98.900000
Punjab	599573	16322	50.000000	582791	98.640000
Assam	576149	5420	3.450000	559684	99.250000
Jharkhand	347440	5130	10.530000	342102	98.740000
Uttarakhand	342462	7368	2.150000	334650	97.790000
Jammu and Kashmir	322771	4392	7.690000	317081	98.290000
Himachal Pradesh	208616	3537	33.330000	202761	99.310000
Goa	172085	3164	1.840000	167978	100.000000
Puducherry	121766	1800	7.690000	119115	98.310000
Manipur	105424	1664	1.660000	96776	100.000000

LINE Graph representing Confirmed & Recovered Trolls in India – for Top5 State



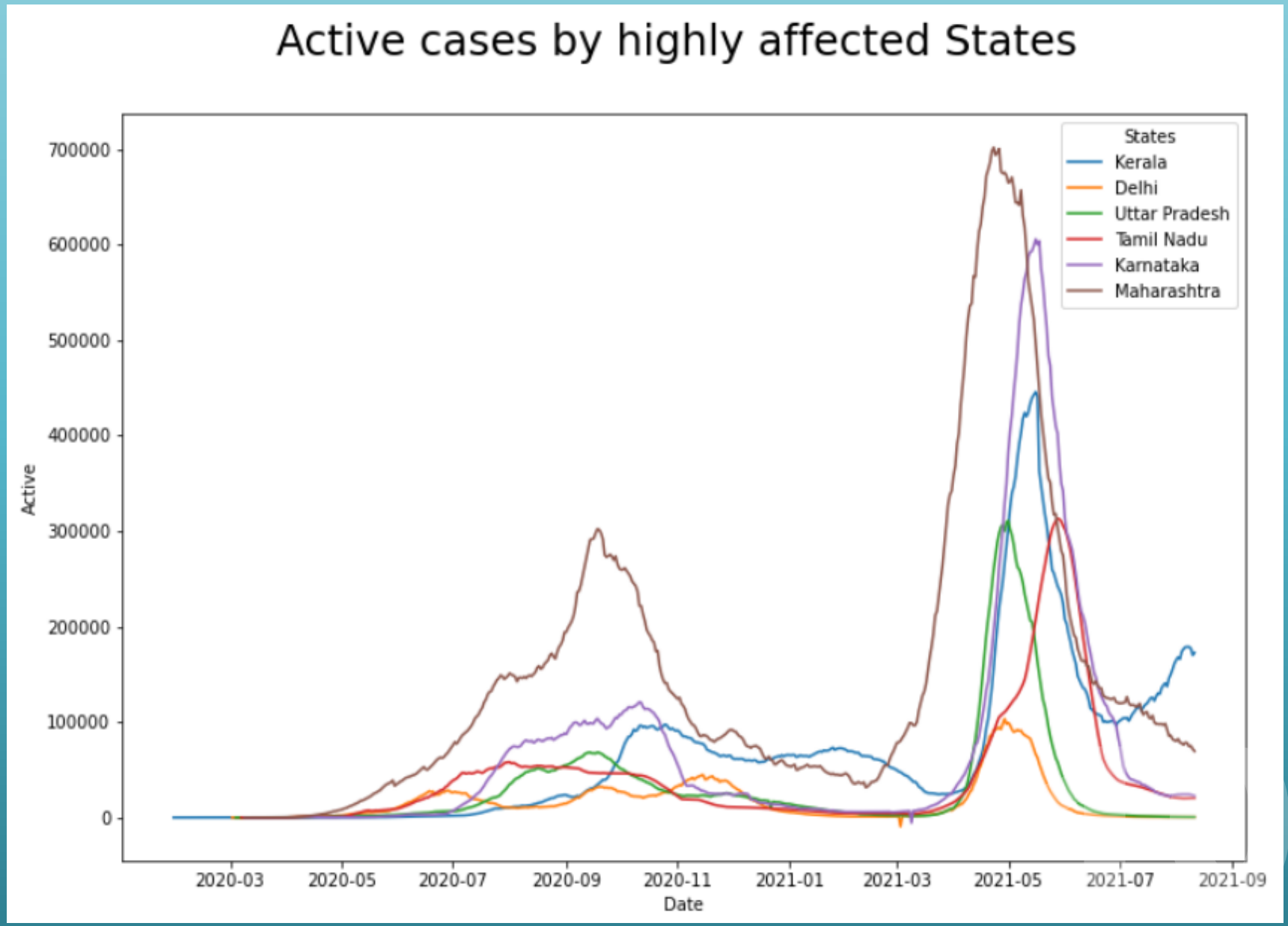


Among all highly affected states, Maharashtra has maximum number of confirmed, recovered and death cases.

Then Karnataka & Kerala are affected respectively.

LINE Graph representing Active Trolls in India – Top5 States

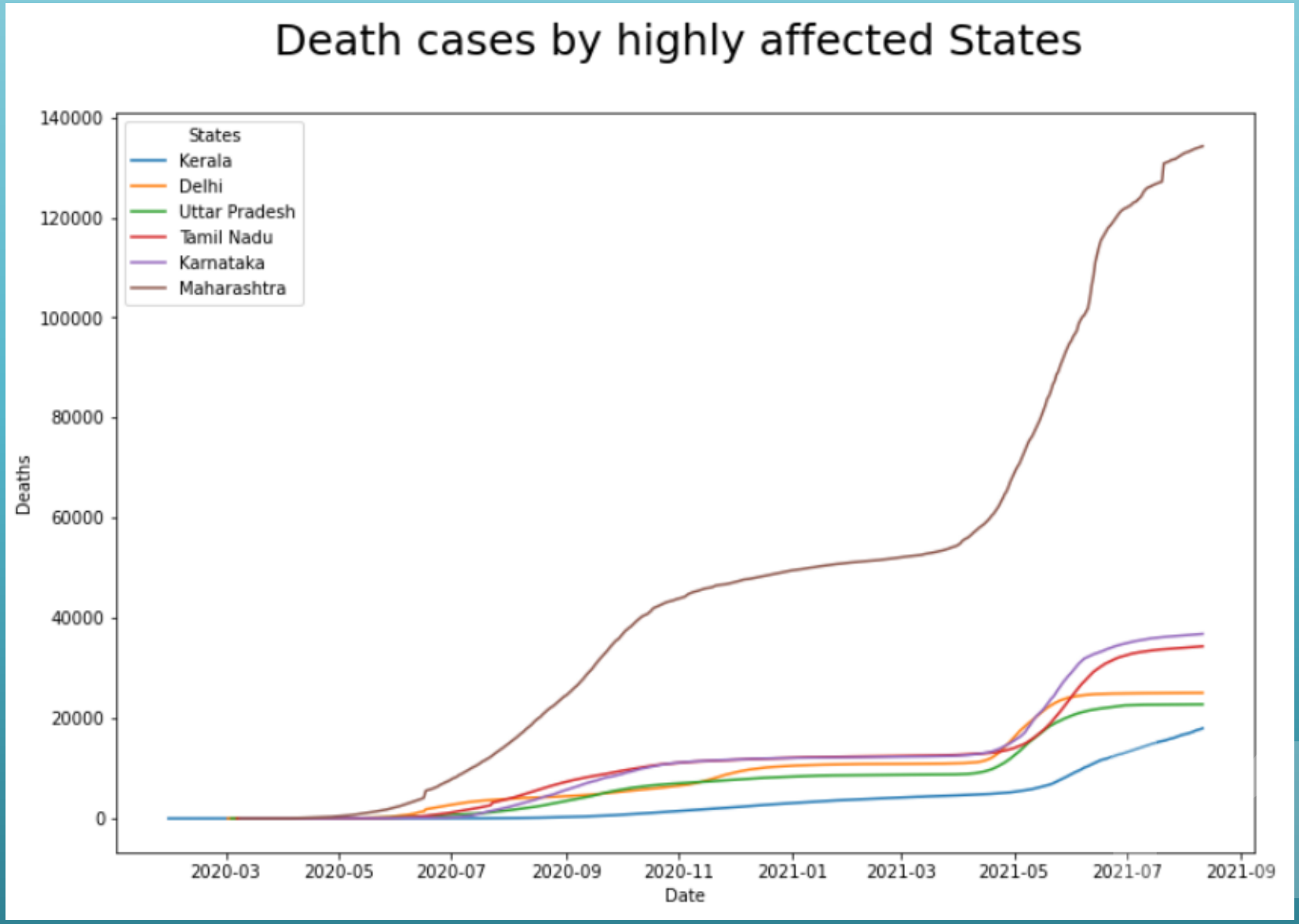
- ✓ Among all highly affected states, Maharashtra has maximum number of Active cases.
- ✓ Then Karnataka & Kerala are affected respectively.
- ✓ **Good news is that**, here we can see the there is a drop in actives cases after the vaccination process started from 1st May 21, i.e. from May21.
- ✓ There is a great drop in Delhi & Uttar Pradesh actives cases.



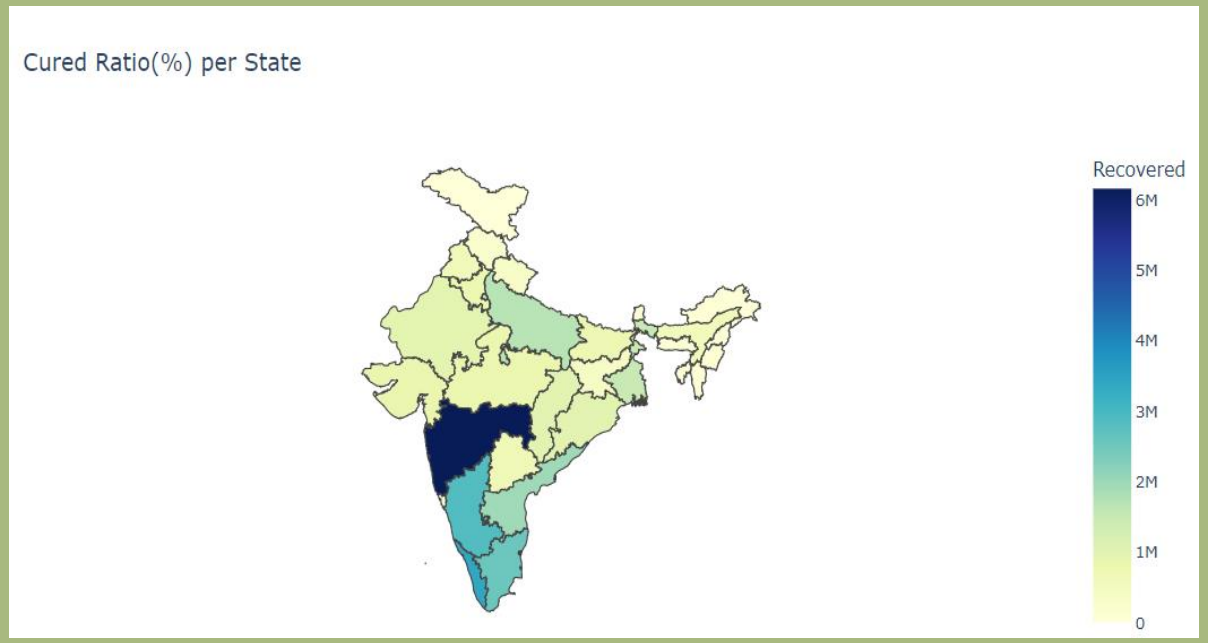
LINE Graph representing Death Trolls in India – Top5 States

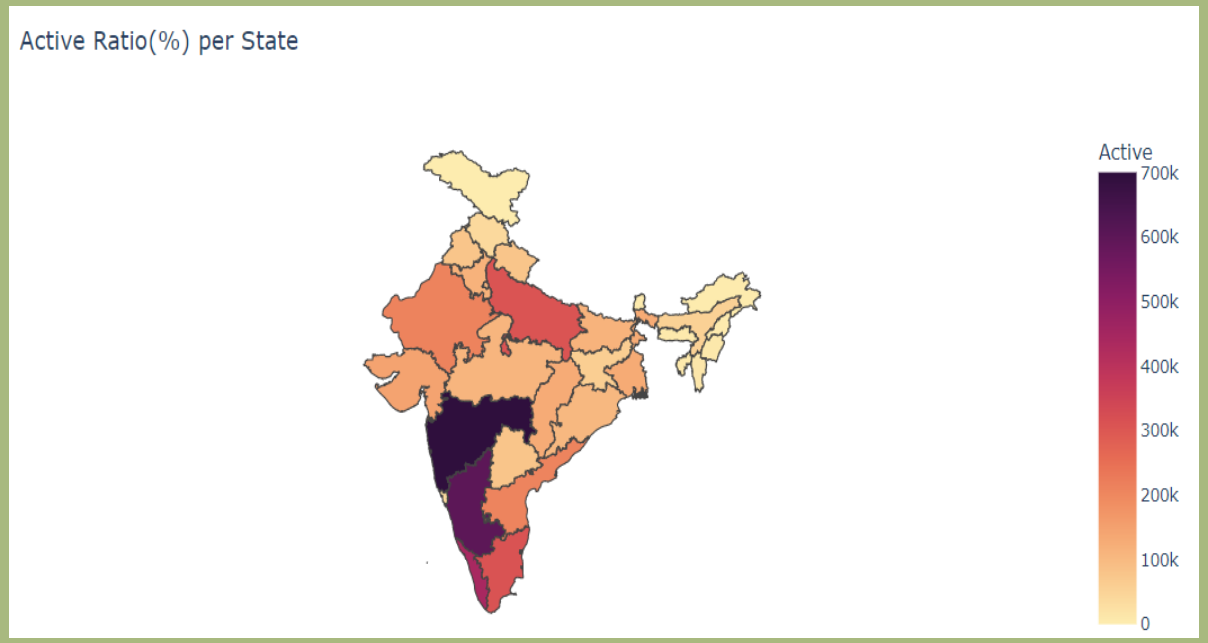
- Among all highly affected states, **Maharashtra** has maximum number of Death cases.

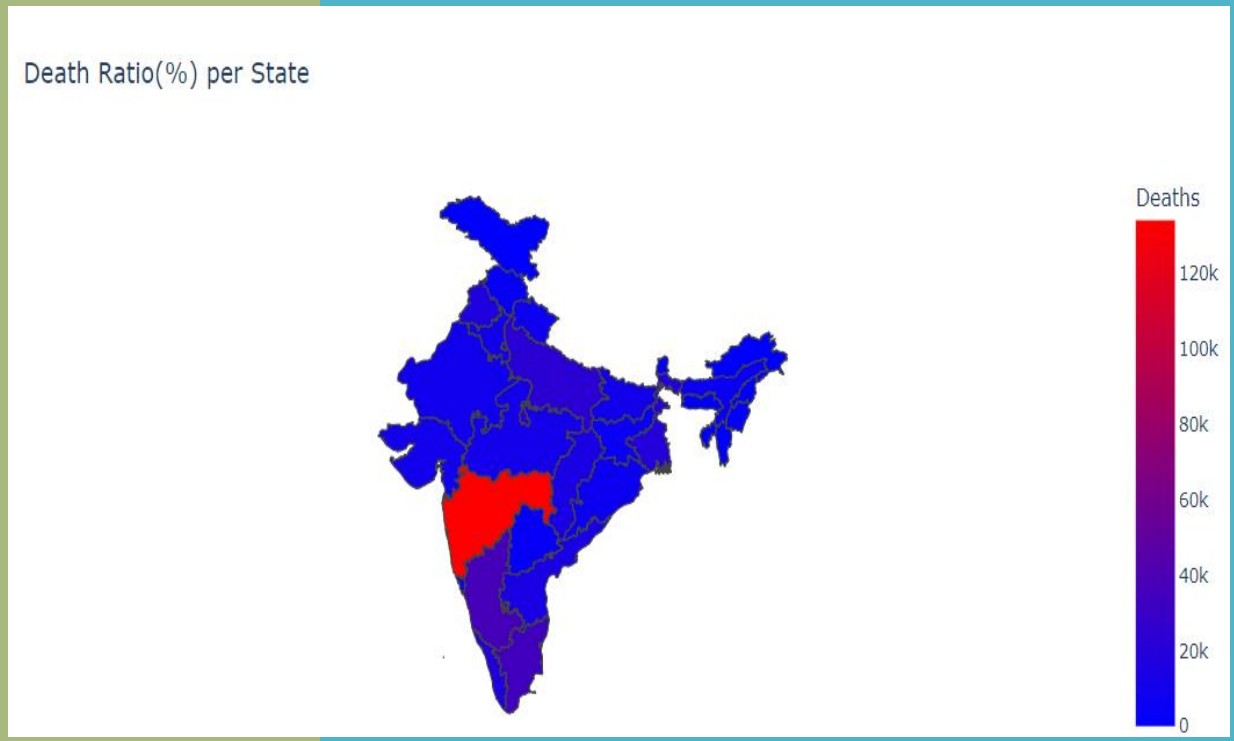
- **Good news is that**, here we can see that there is a **drop in death cases for Uttar Pradesh & Delhi** after May21 (in few days from the vaccination process started from 1st May 21 for 18+).



Cured/Active/Death Ratios representation of the data geographically.



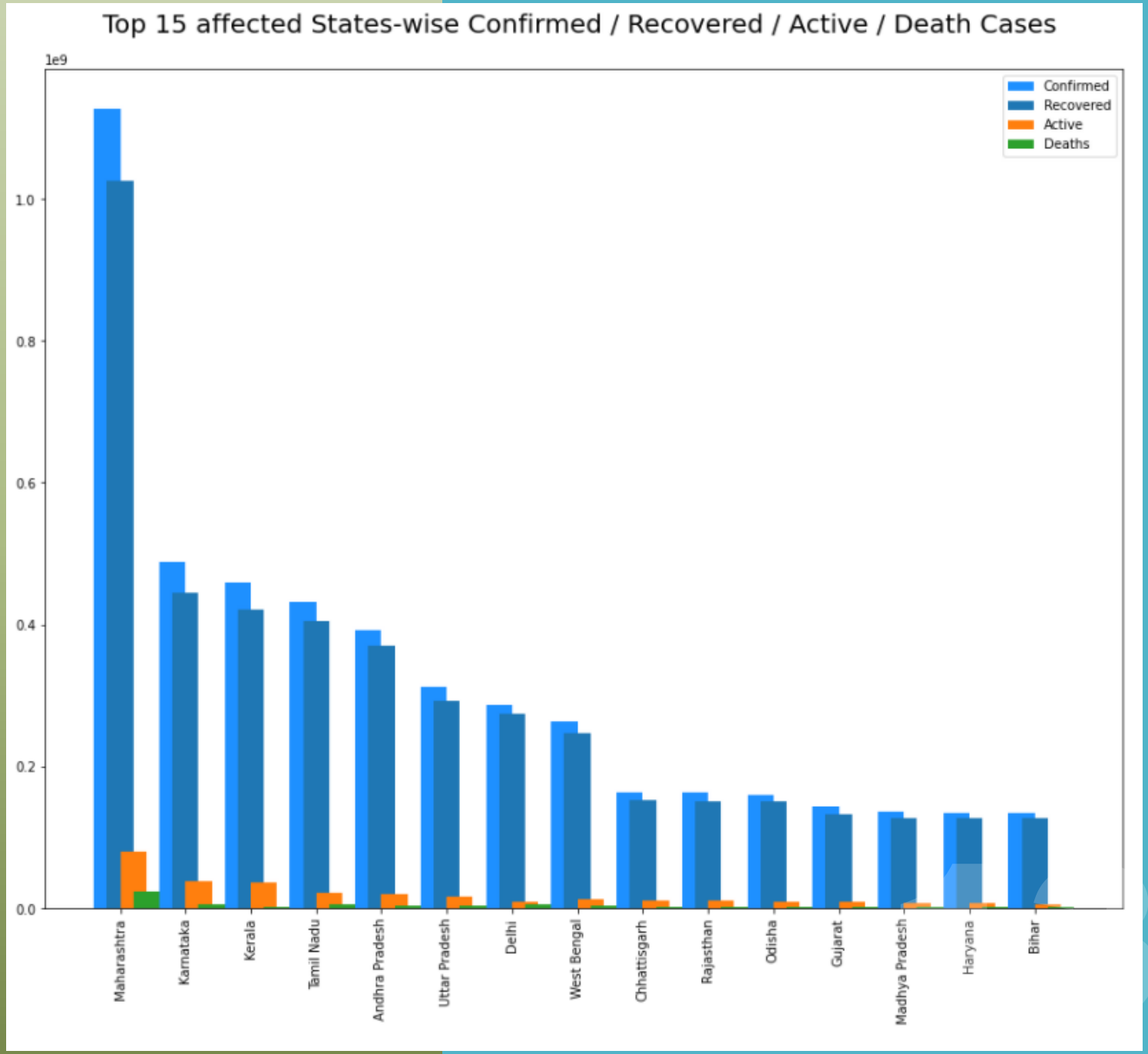




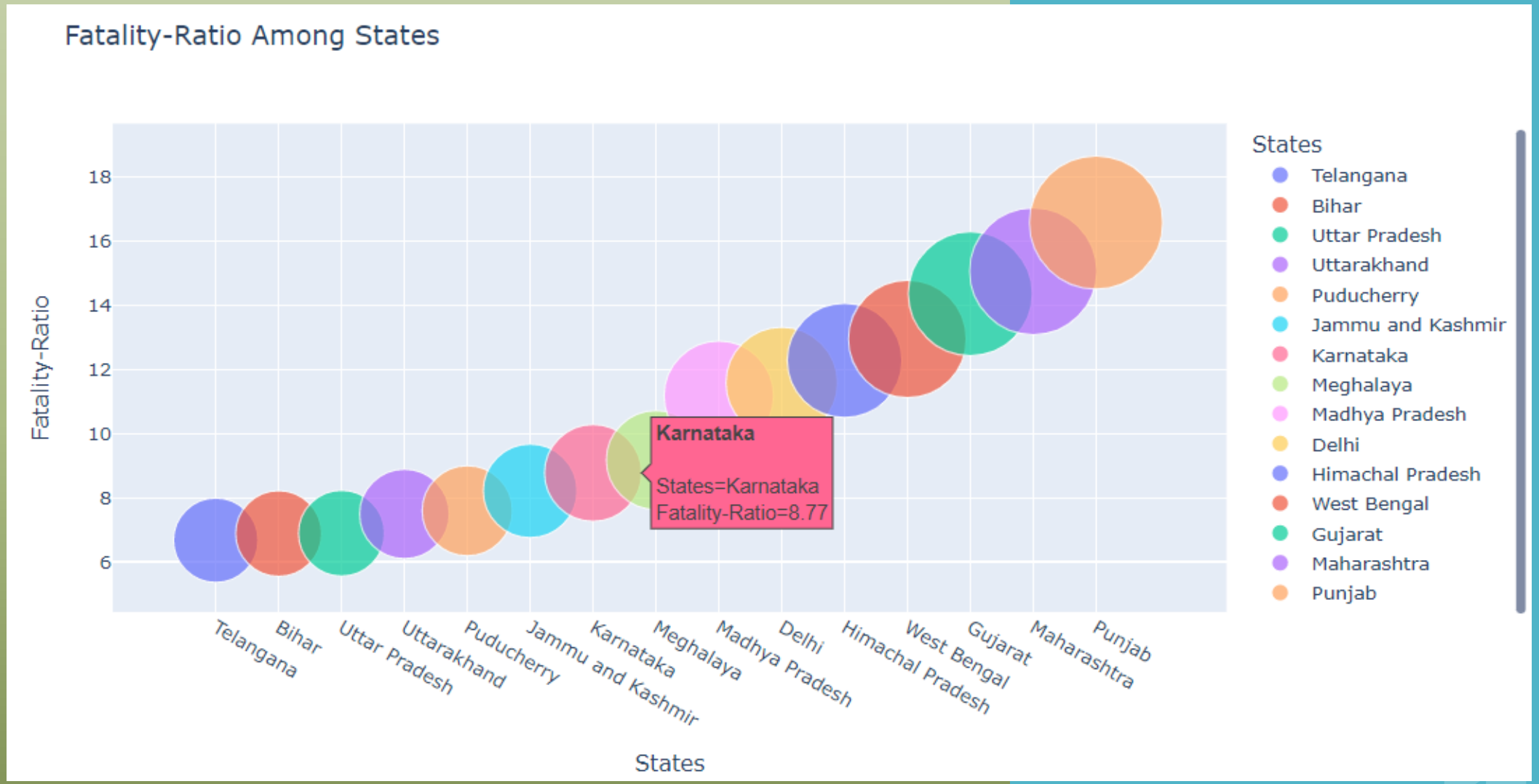
From these graphs, we can see that Maharashtra, Karnataka, Kerala are at high risks of having COVID 3rd wave.

States-wise Confirmed / Recovered / Active / Death Cases

Maharashtra is the state, which faced maximum number of Cases, maximum deaths, maximum number of cured cases, maximum number of Active cases, maximum number of cured cases during this period.



Severity of the disease State-wise



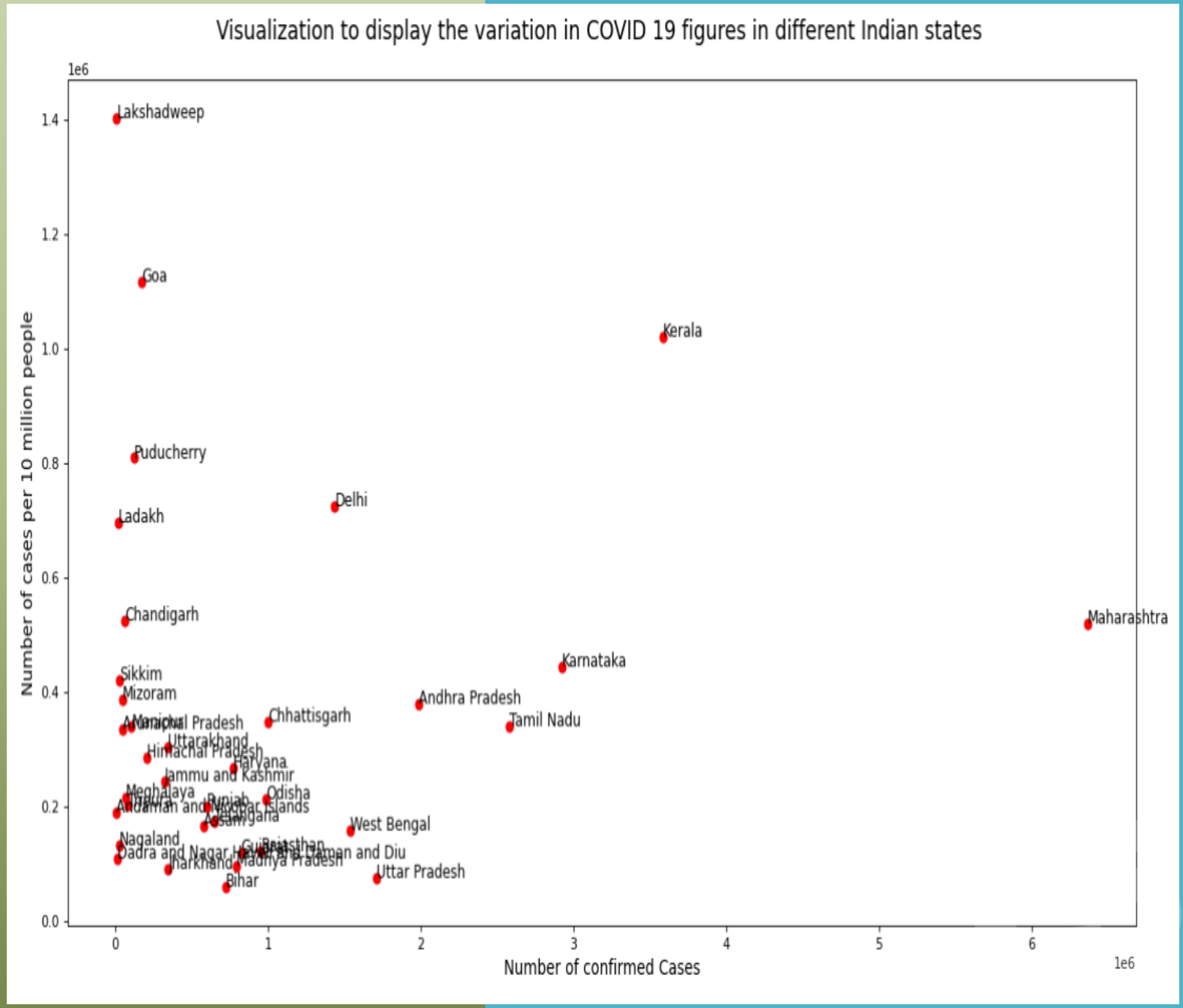
Punjab has the highest Fatality Ratio (severity of disease), and then Maharashtra & Gujarat.

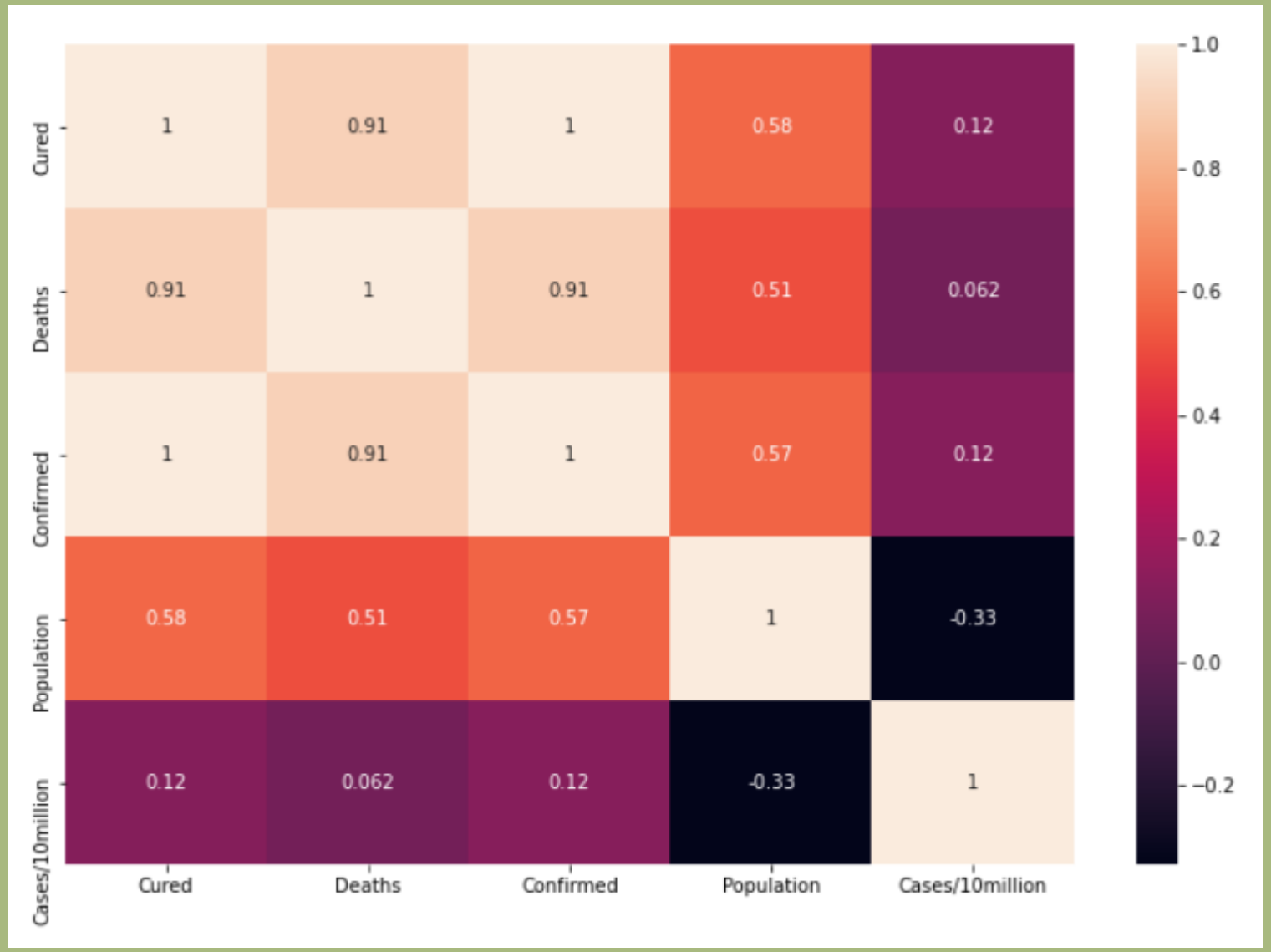
COVID 19 Cases/10million in different Indian states

- Even if we take the State population in consideration, Maharashtra, Kerala, Karnataka and Tamil Nadu are badly hit indeed (One observation here is, all these are near Coastal States).

- In addition to these states come up other states like Andhra Pradesh, Uttar Pradesh, West Bengal, Delhi and Chandigarh.

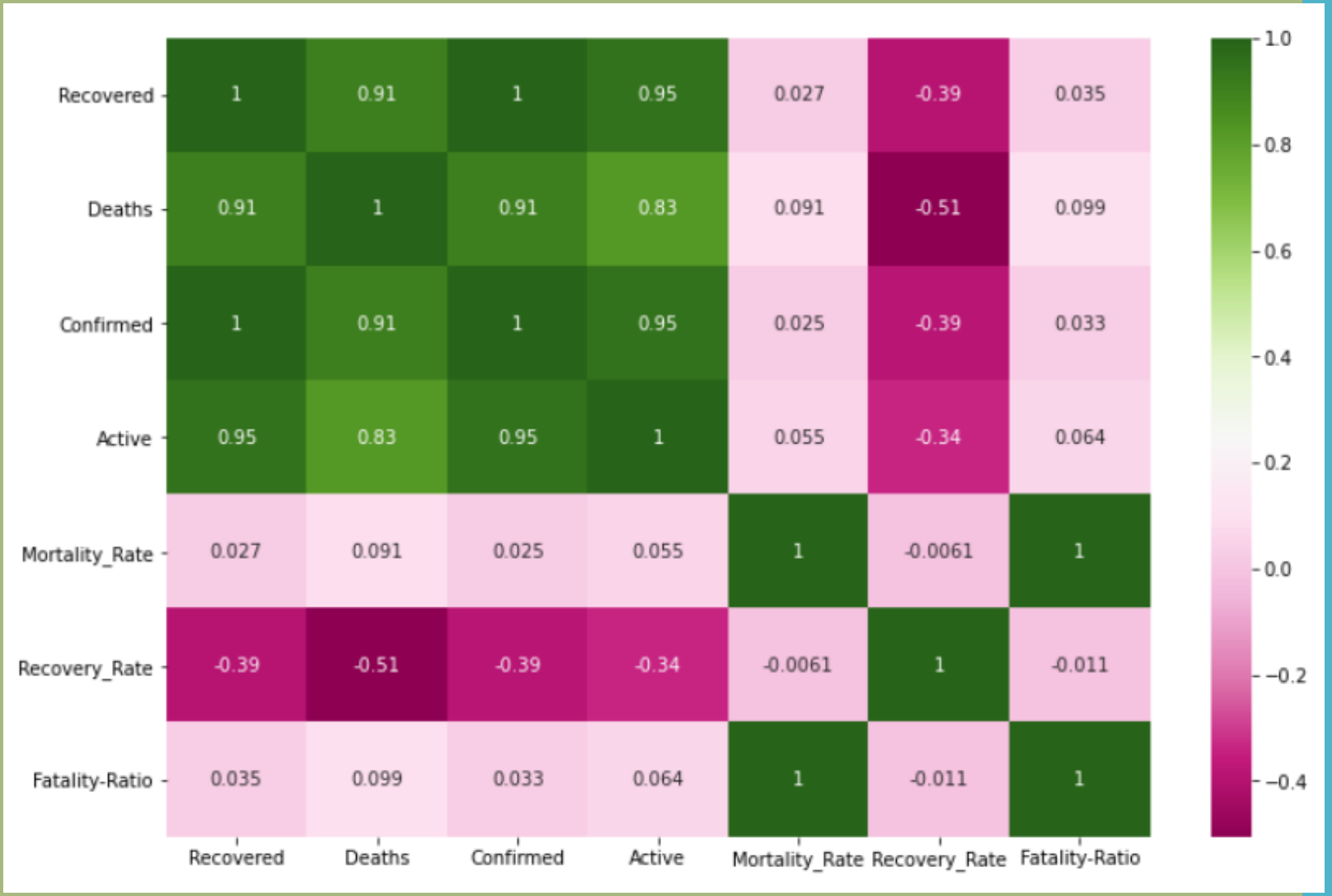
- However, **things do look good for** Lakshadweep, Chandigarh, Ladakh, Sikkim, Mizoram where many cases have recovered.





Features Correlation with Population

- We notice that some measures like Confirmed, Cured, Deaths and Cases/10 million are very much co-related.
- Population & Cases/10 million are negatively correlated.



Features Correlation with Ratios

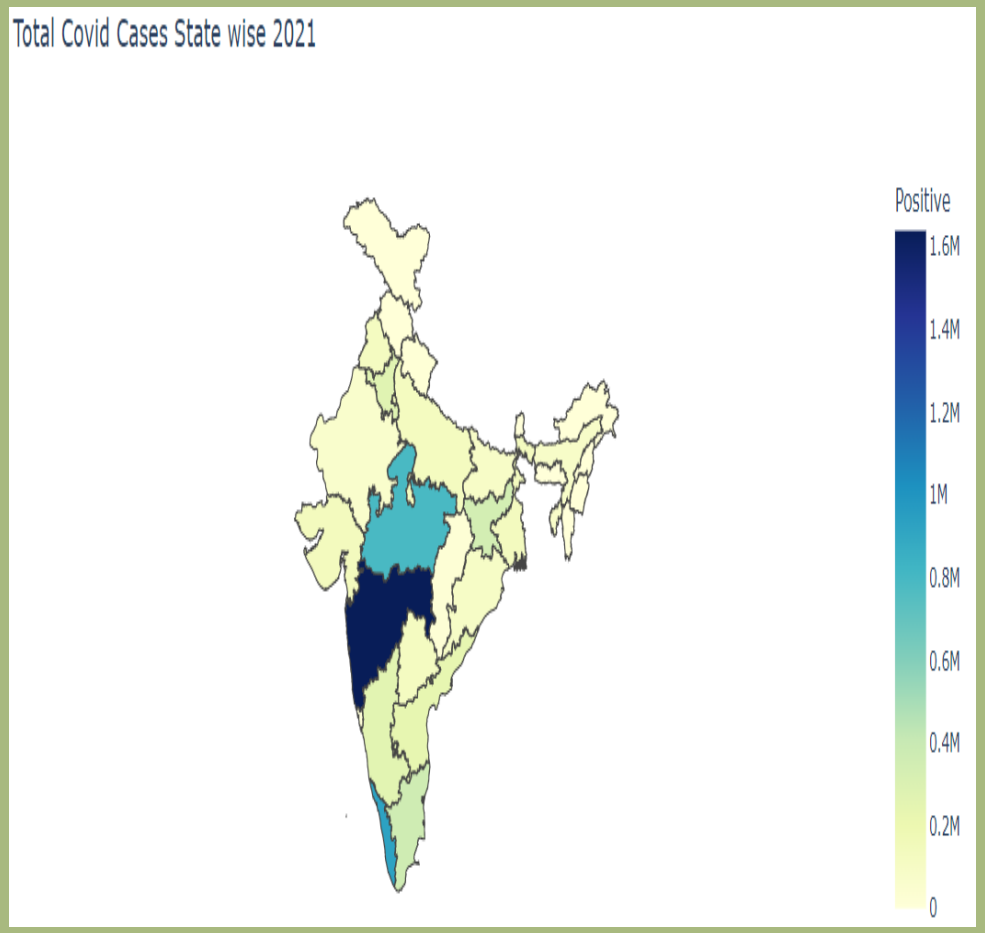
- Recovered & confirmed rates highly correlated.
- Recovery Rate & Deaths are -vely correlated, that means when recovery rate increases the mortality rate will be decreased.

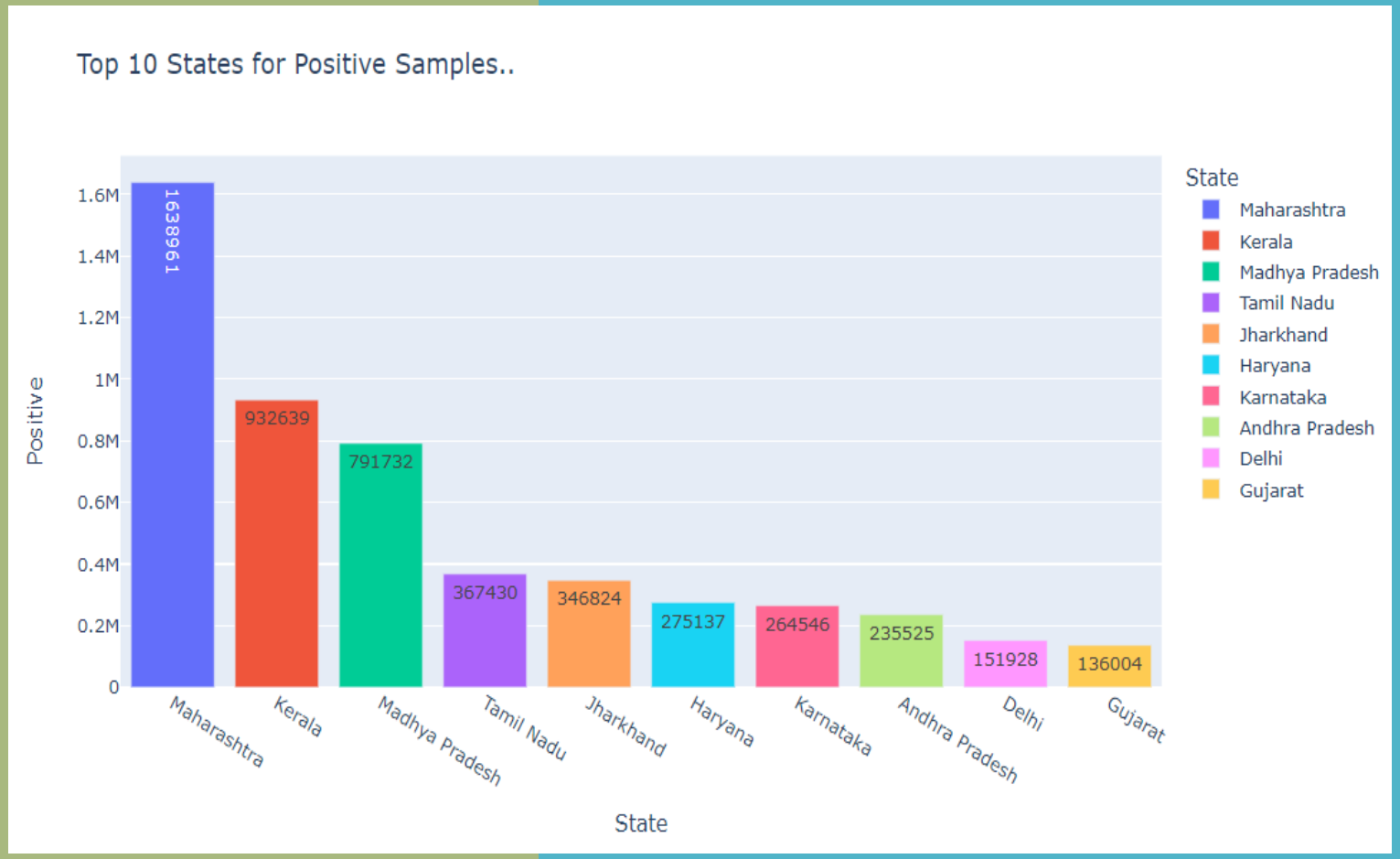
State	Positive	TotalSamples	Positive_Case_Rate
Uttar Pradesh	126722.000000	67897856.000000	0
Maharashtra	1638961.000000	49905065.000000	3
Karnataka	264546.000000	40104915.000000	0
Tamil Nadu	367430.000000	39002757.000000	0
Bihar	90553.000000	38820518.000000	0
Kerala	932639.000000	28745545.000000	3
Gujarat	136004.000000	26192626.000000	0
Andhra Pradesh	235525.000000	25311733.000000	0
Delhi	151928.000000	24333906.000000	0
Telangana	124963.000000	22991849.000000	0
Assam	87908.000000	19850867.000000	0
Odisha	97920.000000	16683764.000000	0
West Bengal	135596.000000	16162814.000000	0
Madhya Pradesh	791732.000000	15144644.000000	5
Rajasthan	67954.000000	13185136.000000	0
Punjab	124535.000000	12475529.000000	0
Jammu and Kashmir	31371.000000	12307566.000000	0
Jharkhand	346824.000000	12184347.000000	2
Chhattisgarh	19459.000000	11762041.000000	0
Haryana	275137.000000	11135555.000000	2
Uttarakhand	14083.000000	6526861.000000	0
Himachal Pradesh	3993.000000	2961627.000000	0
Tripura	80413.000000	1630572.000000	4
Puducherry	44037.000000	1557320.000000	2
Manipur	4765.000000	1136573.000000	0
Goa	12333.000000	1102474.000000	1
Arunachal Pradesh	2658.000000	986281.000000	0

Positive cases ratio State-wise

Uttar Pradesh and **Maharashtra** had ramped up their testing capabilities and tested the highest samples respectively.

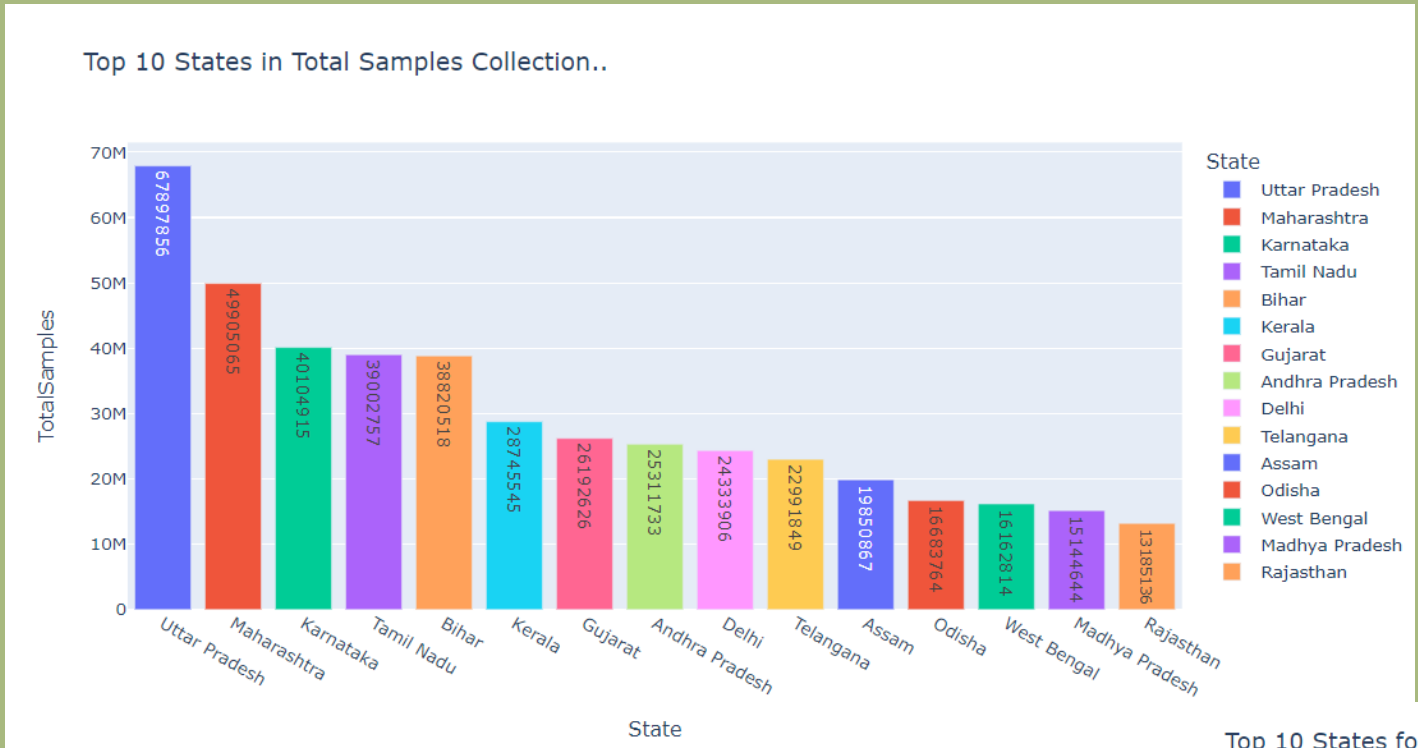
Top States in Positive Samples collection





We can see above that among all states, Maharashtra has highest Positive Samples collected.

Then comes Kerala & Madhya Pradesh.

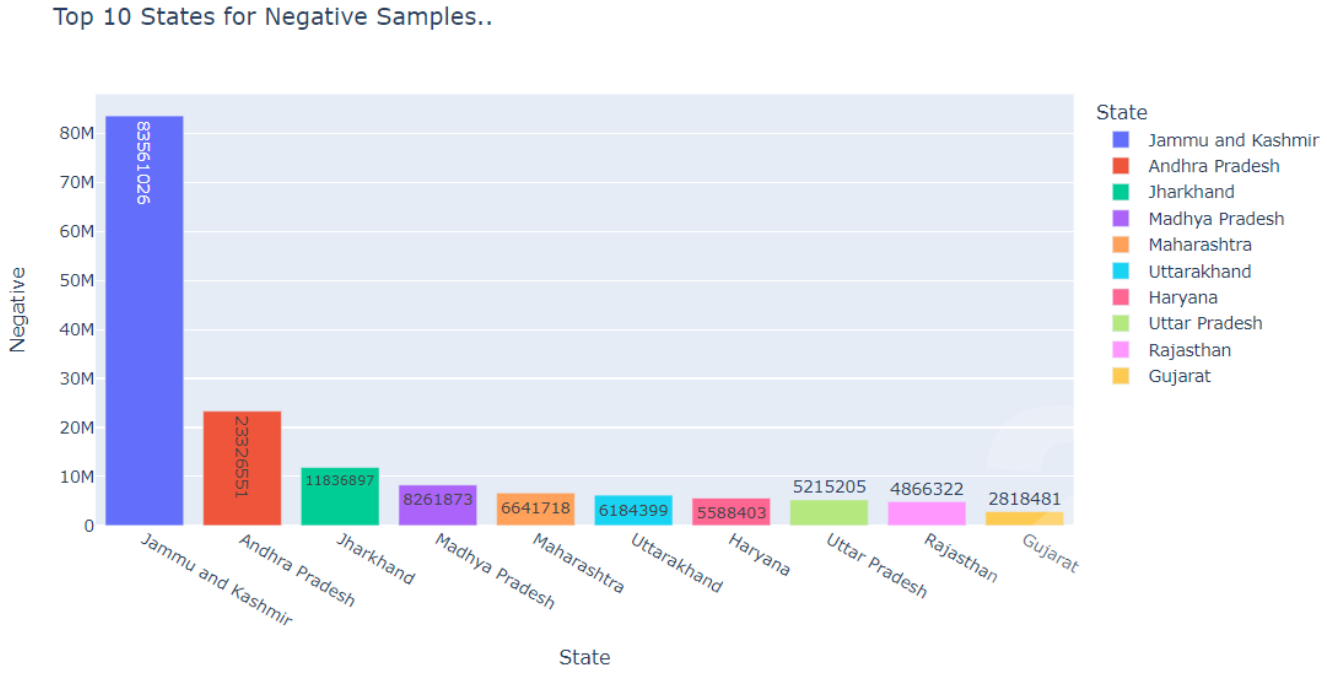


Top States in Total Samples collection

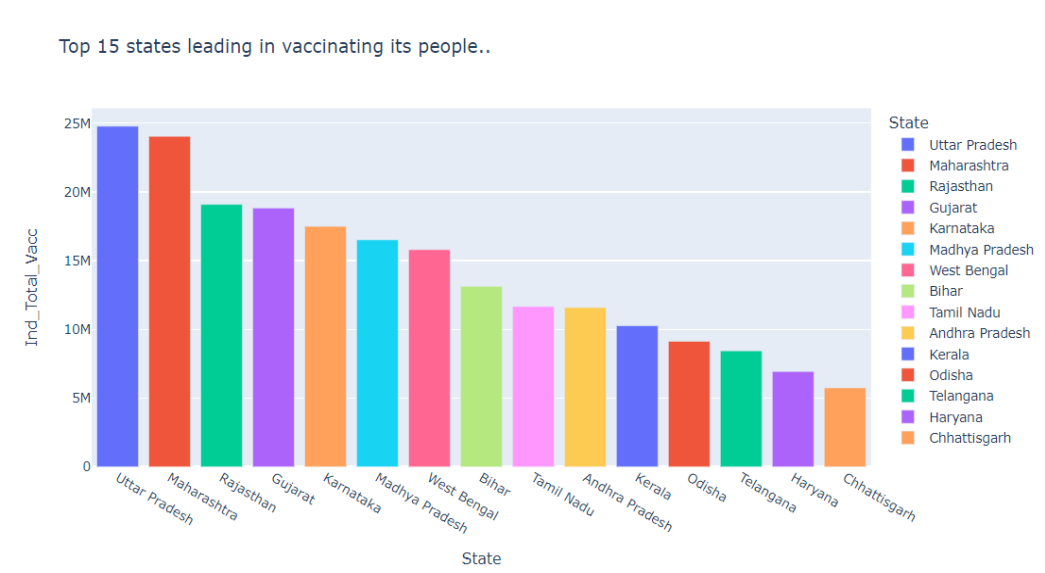
we can see that State UP collected highest Total Samples among others. That means UP is good at Testing.

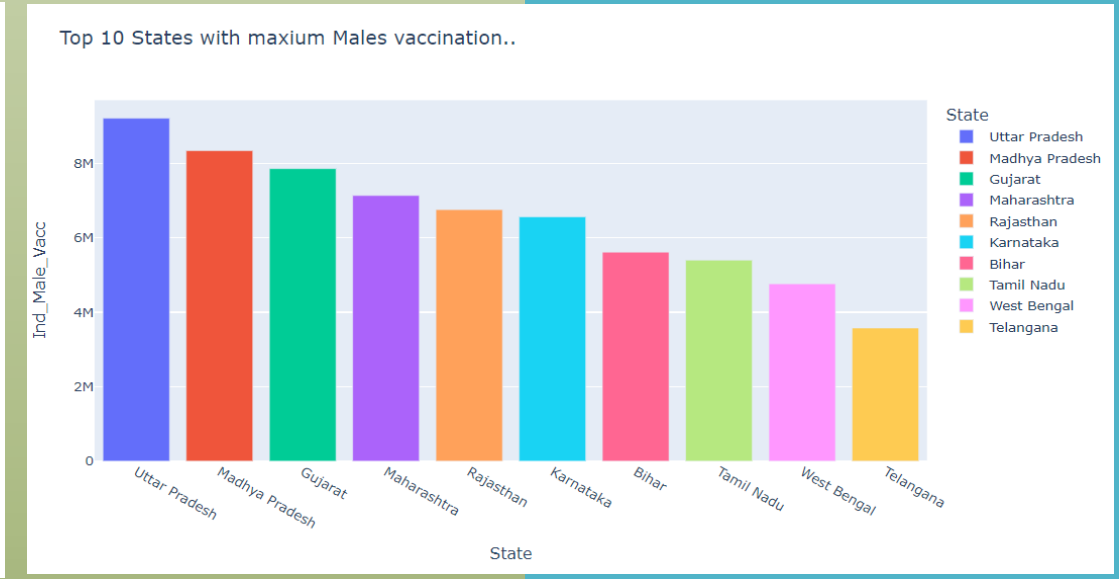
Top States in Negative Samples collection

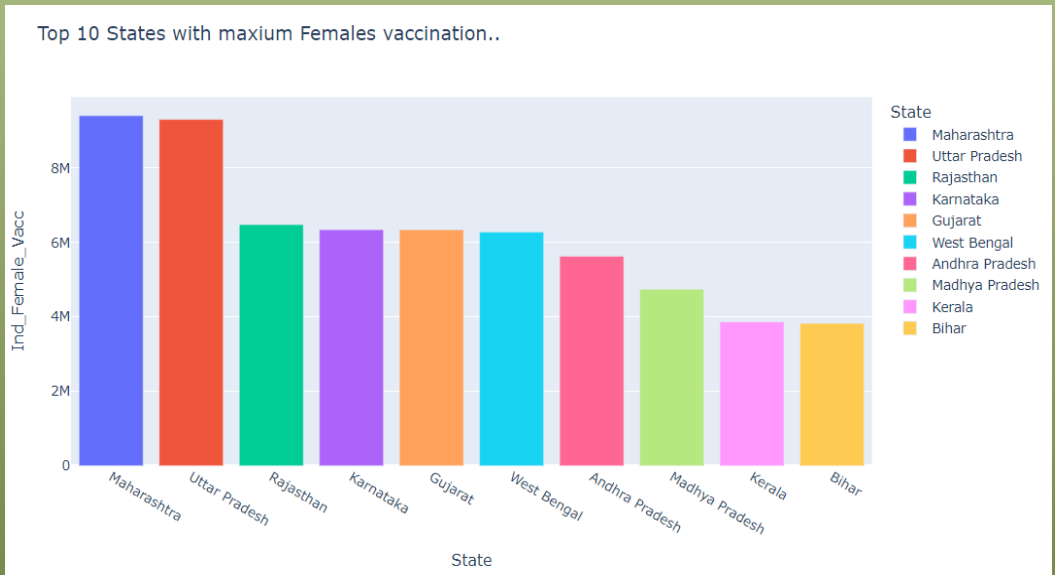
we can see that State Jammu and Kashmir is highest in Negative Samples collection among others.

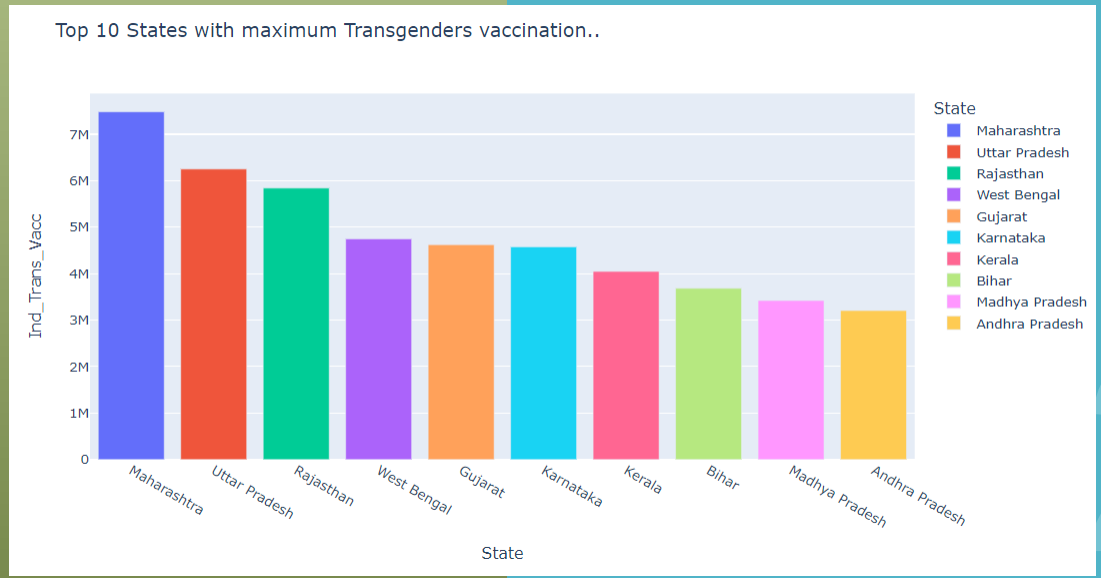


Vaccination Data – Total / Males / Females / Transgenders





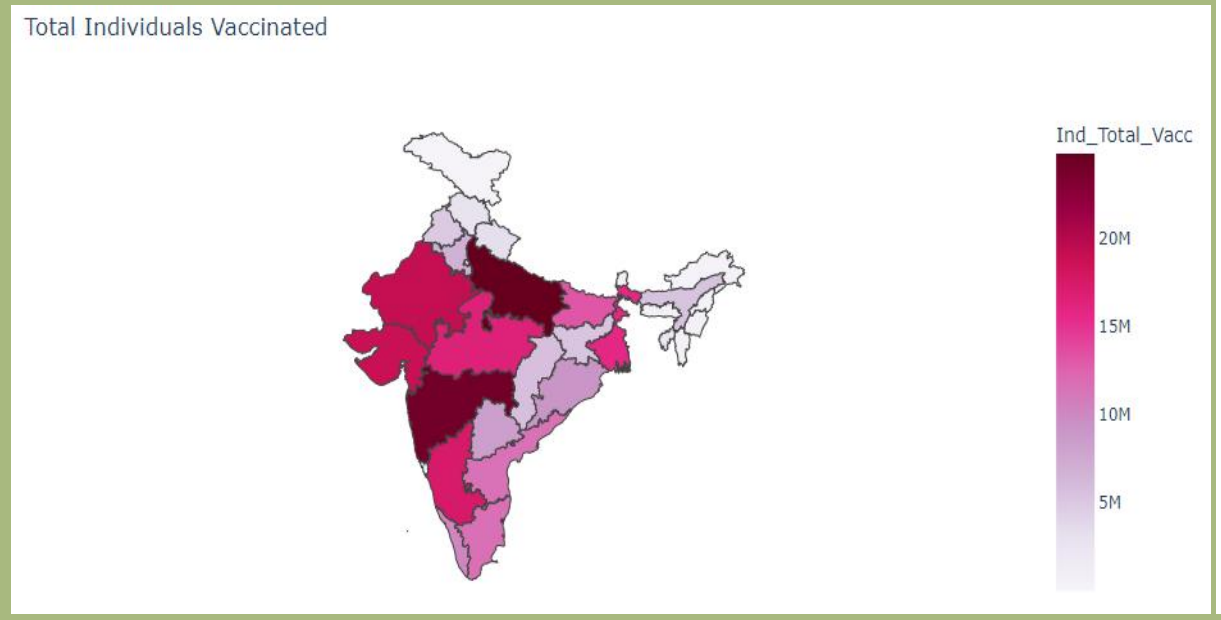


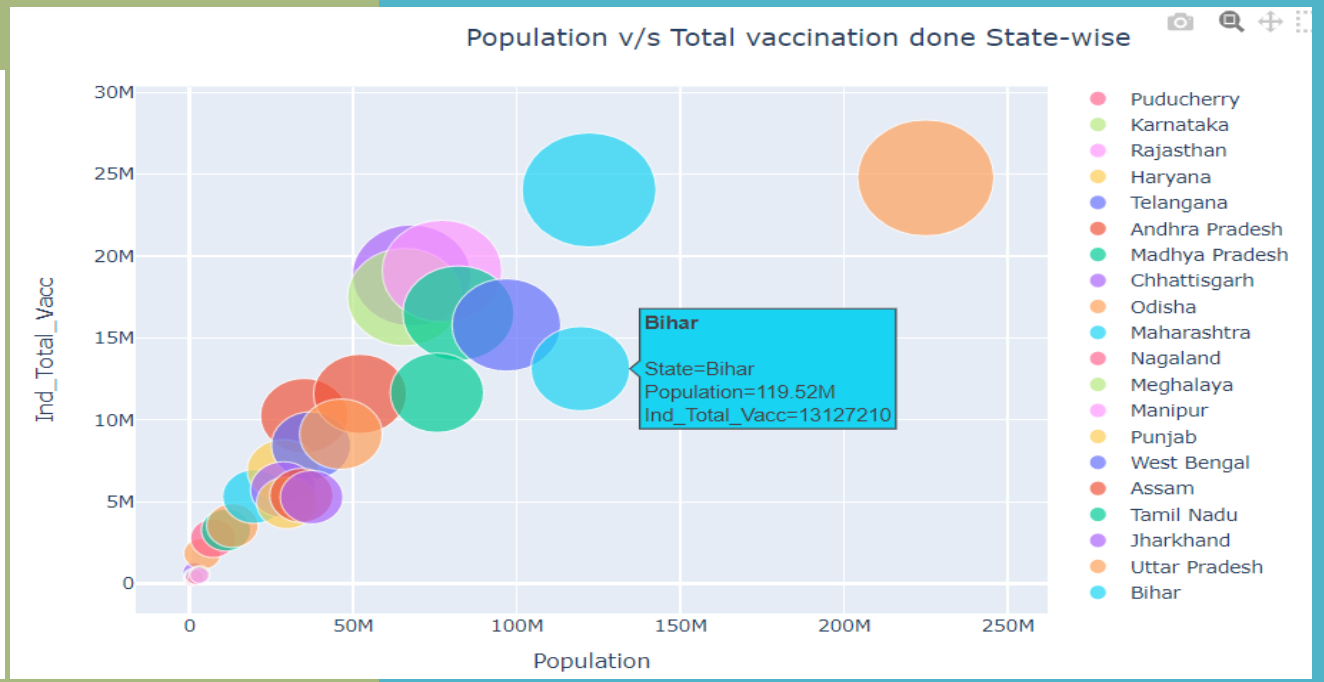


Graphs state that Maharashtra & Uttar Pradesh are the leading States among all, in having highest samples collection.

Then follows, Gujarat, Rajasthan, Madhya Pradesh & Karnataka.

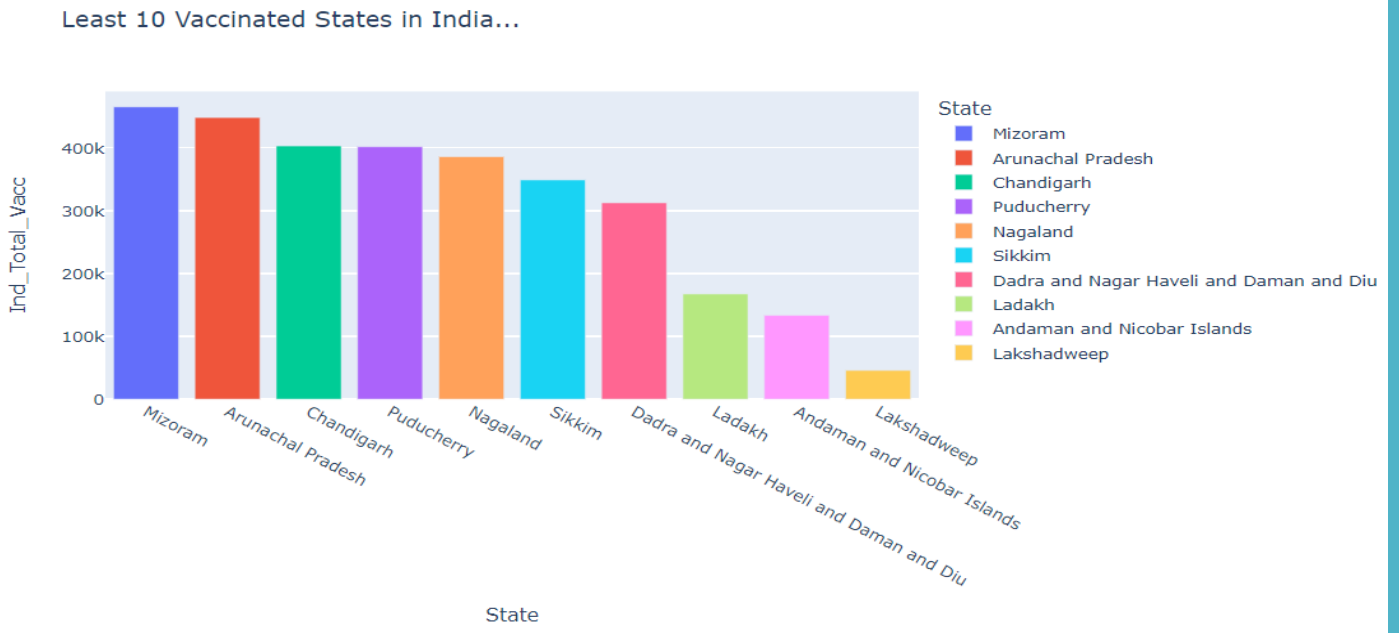
1. Display of highly vaccinated States on India Map.

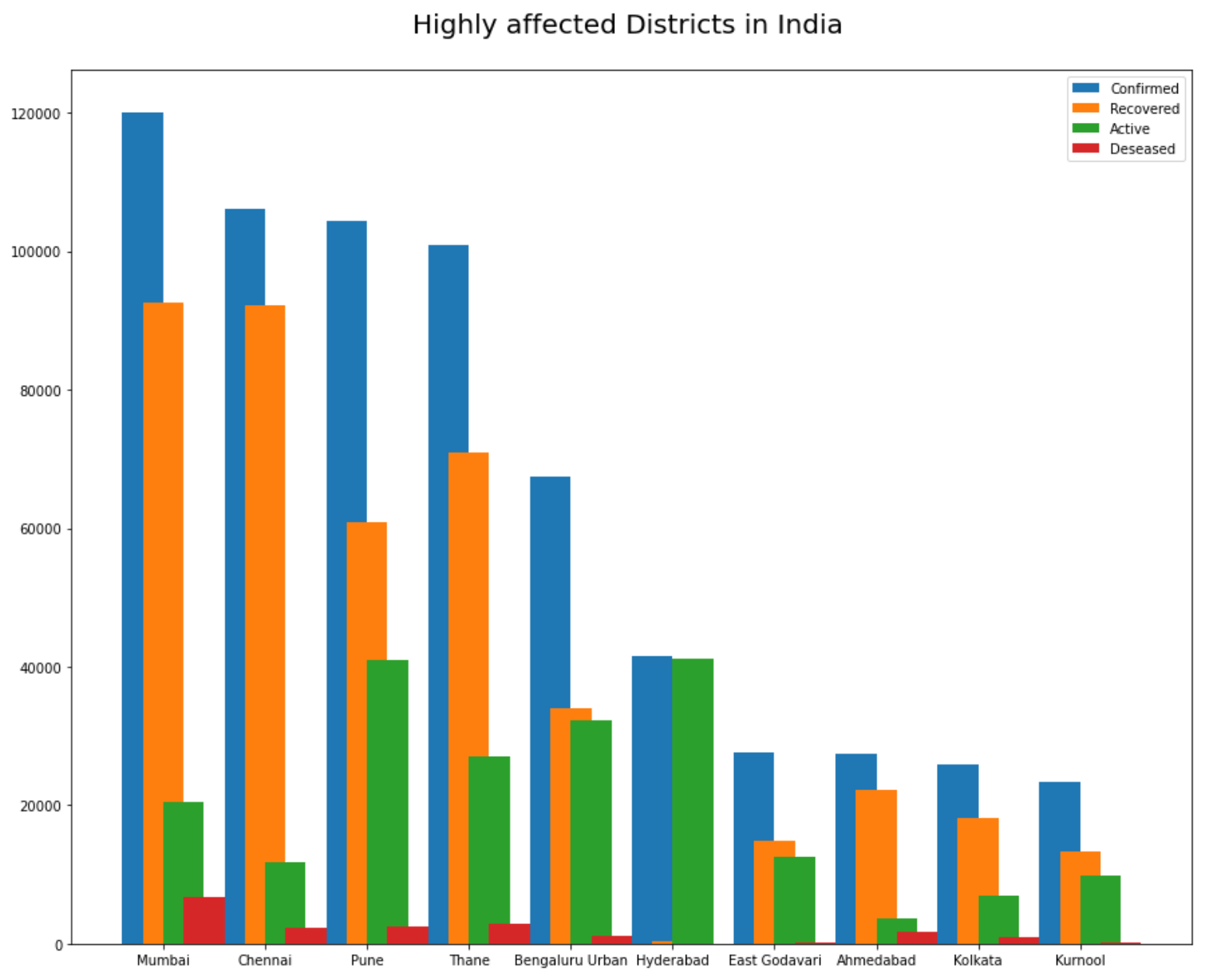




2. Least 10 vaccinated States in India.

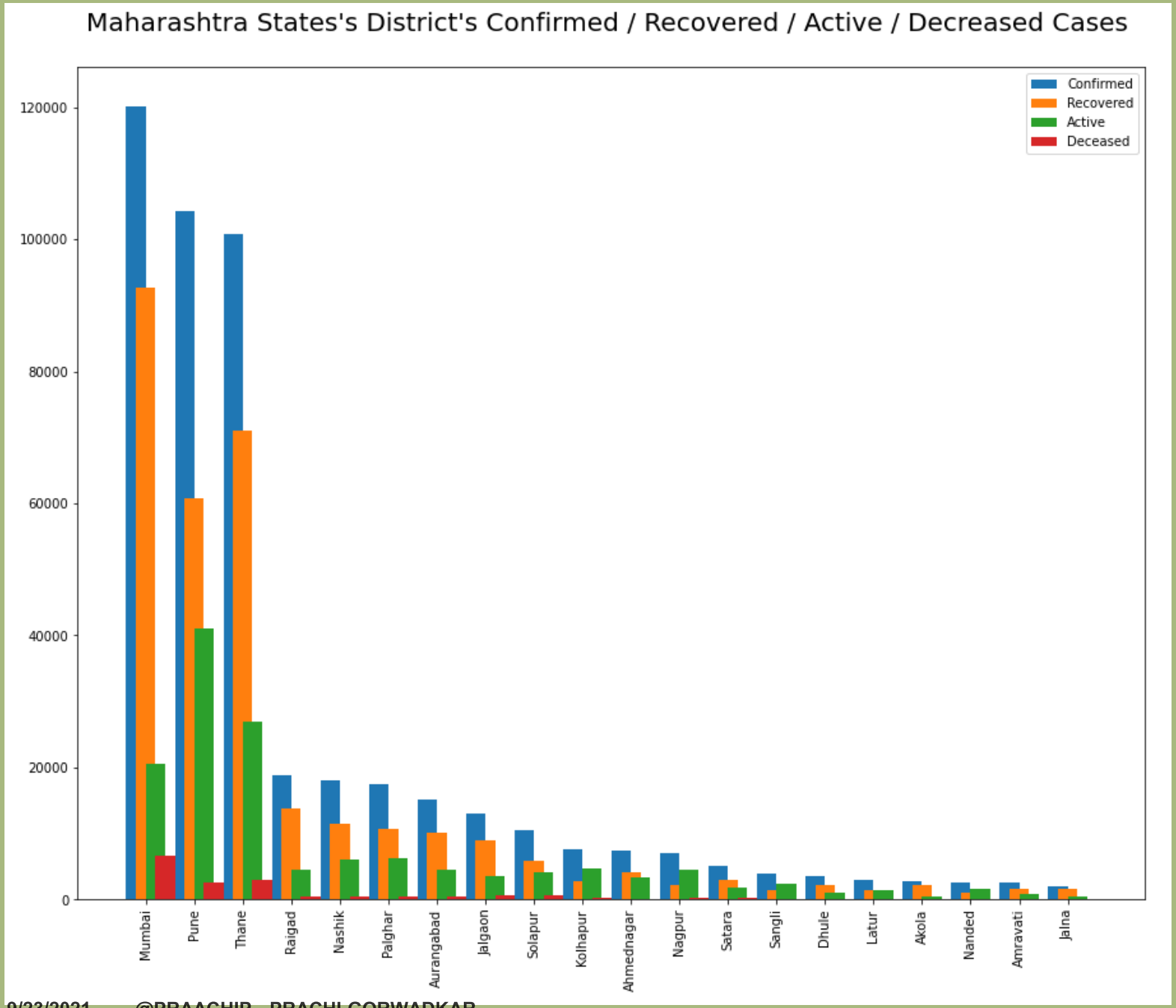
we can observe that, the states we need to concentrate on, as these are the least vaccinated states, as per population also. These include Mizoram, Arunachala Pradesh and Chandigarh from north east and some Union territories.





Highly affected Districts in India

- As we have seen Maharashtra is the highest affected State in India, its District Mumbai is highly affected among all Districts of Maharashtra, then Chennai, Pune & Thane.
- Mumbai has maximum number of Confirmed & Diseased Cases, whereas Pune & Hyderabad have maximum Active cases.
- Chennai, Mumbai, Thane & Pune have maximum number of Recovered cases respectively, among the top 10 highly affected Districts



Highly affected Districts of Maharashtra

- As we have seen Maharashtra is the highest affected State in India, so here is the comparison of its districts.
- **Districts like Mumbai, Pune & Thane are highly affected** among all Districts of Maharashtra.
- Mumbai has maximum number of Confirmed & Death Cases, whereas Pune& Hyderabad has maximum Active cases.
- Chennai, Mumbai & Thane have maximum number of Recovered cases respectively, among the top 10 highly affected Districts

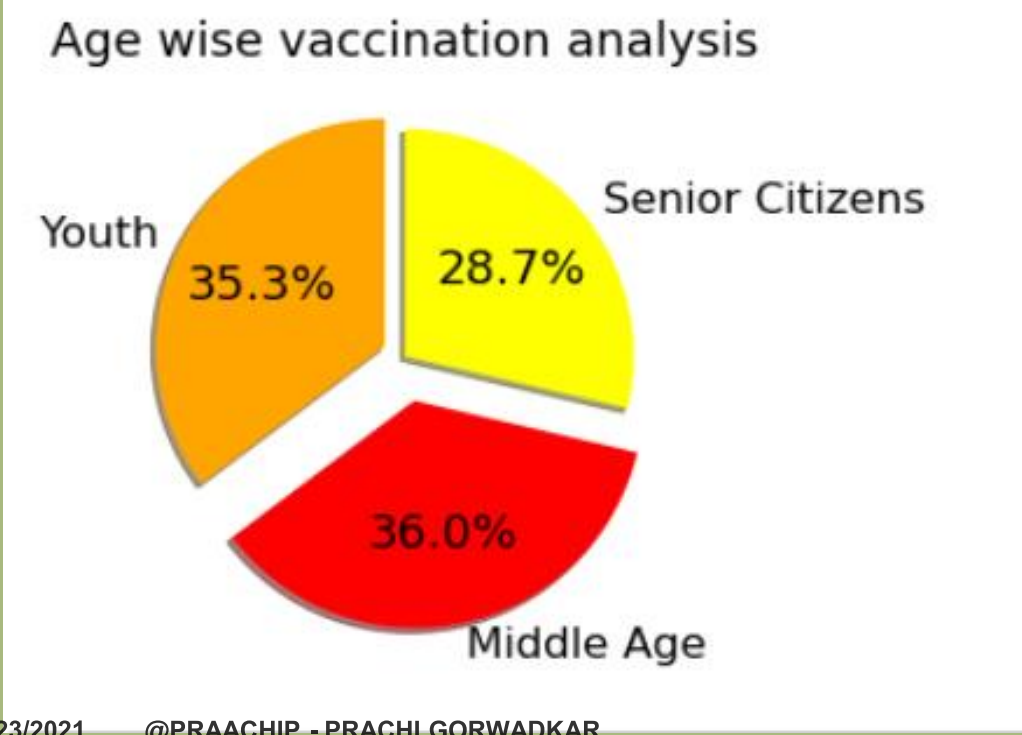
Vaccination Data Age-wise & Date-wise

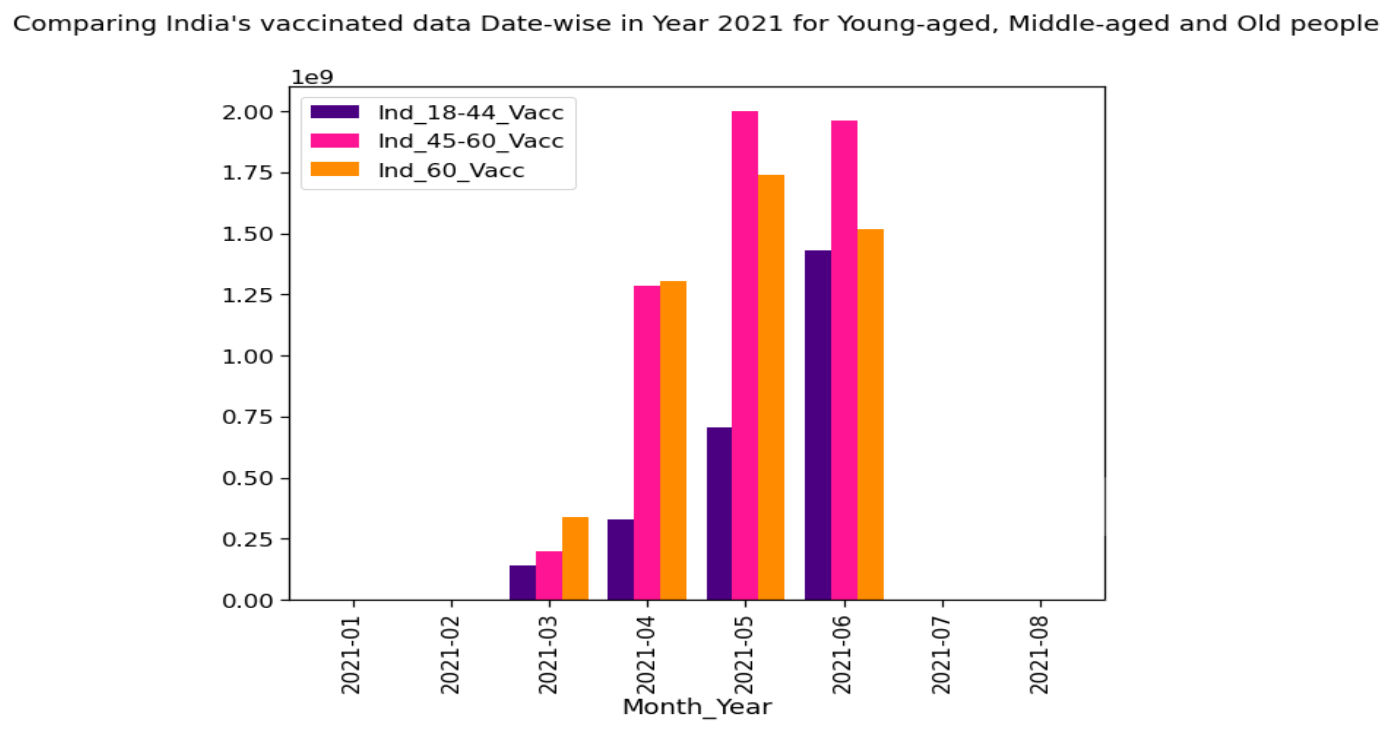
1. Middle Aged (45-60 years) people have high% (36%) of vaccination, followed by Youth (35.3%) and senior citizens (28.7%).

2. In March 2021, very less no of people are vaccinated, as 2nd phase of vaccination process was just started from 1st March 21 for 60+ & for aged 45 and above with specified co-morbid conditions , before it was only for the medical & related staff.

3. For 45+, it was started from 1st April 21 and for 18+, it was started from 1st May 21. **So, we can clearly observe the changes accordingly in the graph.**

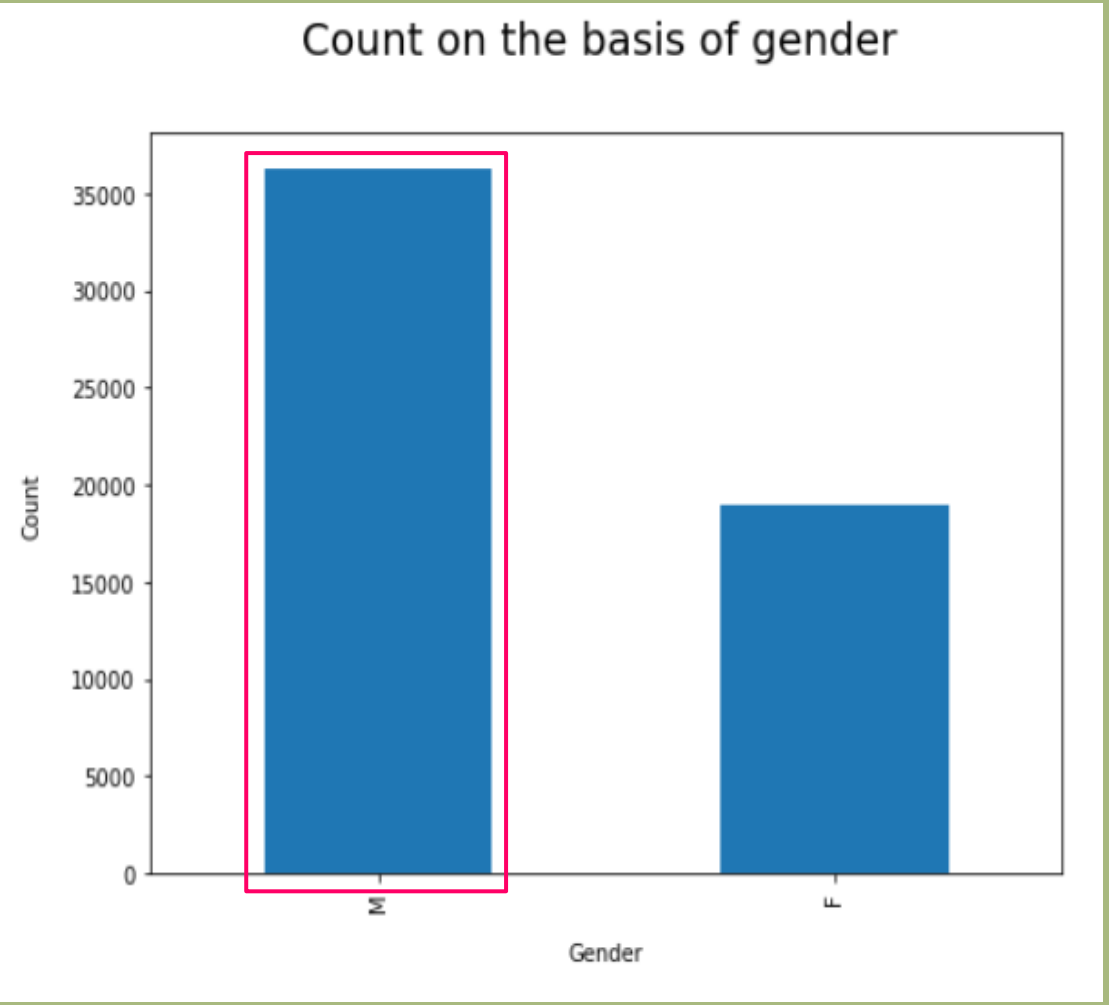
4. June 2021 has the maximum no of people vaccinated. People aged 45-60 are maximum in having vaccination.

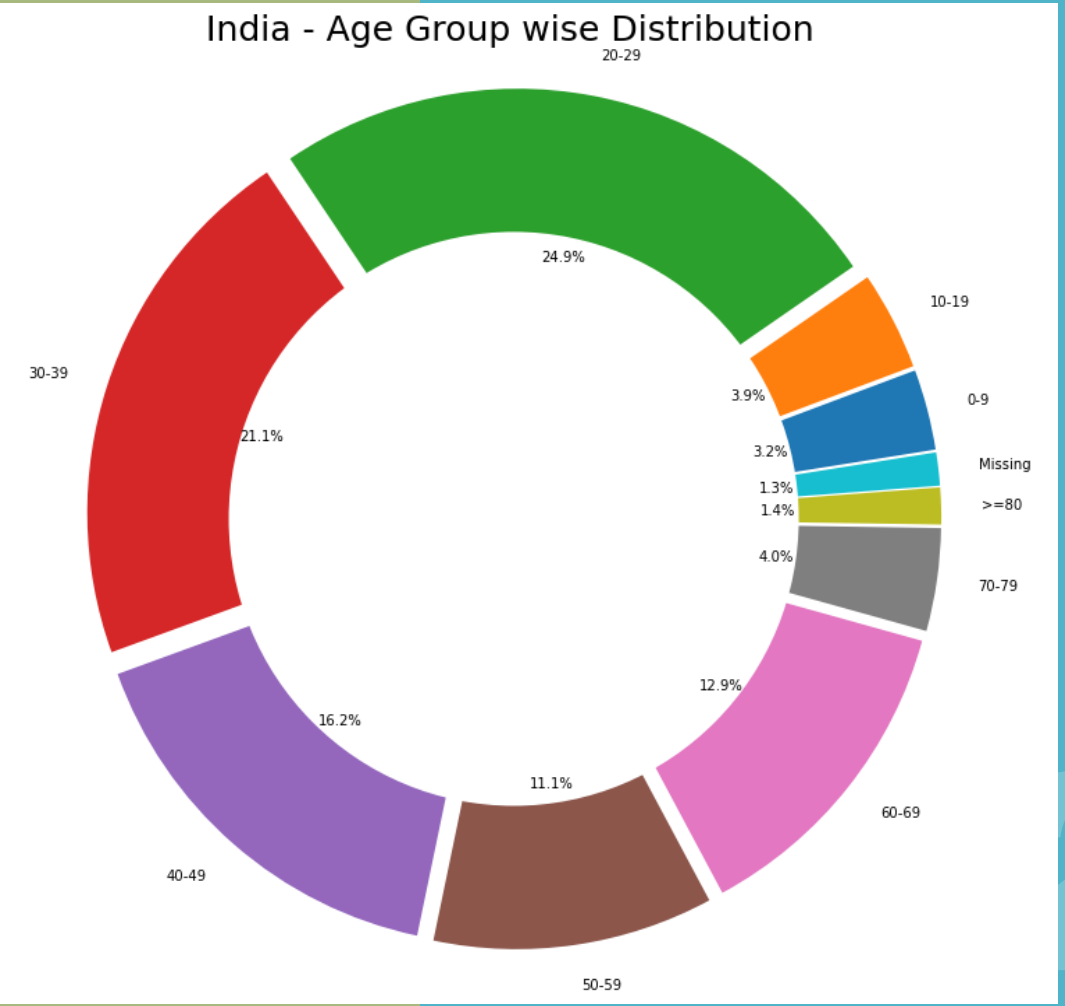




Gender based Vaccination count & age group-wise affected %

1. Males are highly affected than females.
2. Age group between 20-40 is highly affected by the pandemic.





CONCLUSIONS

Good News :-

1. We have a good news that almost all the states have the Discharge rate greater than 85% & Maharashtra has the highest % of Cured cases.

2. 1st phase → 1st Jan21 ; 2ns phase → 1st Mar21 → for 60+ & for aged 45 and above with specified co-morbid conditions

1st April 21 → All 45+ people; 1st May 21 → 18+

We can see the there is a drop in actives cases after the vaccination process started & there is a great drop in active cases for Delhi & Uttar Pradesh.

3. Things do look good for Lakshadweep, Chandigarh, Ladakh, Sikkim, Mizoram (compared to the population) where many cases have recovered.

4. Uttar Pradesh and Maharashtra had ramped up their testing capabilities and tested the highest samples respectively.

5. Being second in the country w.r.t. total confirmed cases, Kerala seem to have established the healthcare facility to a higher level which is evident from the lower mortality rate (0.89%).

Bad News :-

1. From the EDA we can say that most affected States by covid 2nd wave are Maharashtra, Kerela, Karnataka, Tamil Nadu and Andhra Pradesh out of which four are the south states so these states might have a high risk of getting covid 3rd wave and should be more concerned about taking the necessary steps. (One observation here is, all these States are near Coastal area).

END

Thank you