**RUSTAMJI INSTITUTE OF TECHNOLOGY**

### (Affiliated to Rajiv Gandhi Proudyogiki Vishwavidyalaya,Bhopal)

BSF ACADEMY,TEKANPUR-475005

**A**

**Minor Project Report On**

**CORONALYSIS**

**Submitted By:**

Chirag Sharma 0902IT181011

Mansi Jain 0902IT181026

Minesh Sarawogi 0902IT181027

Vanshika Sikarwar 0902IT181051

#### **Under the guidance of**

Prof. Aradhana Saxena

# **Department of Information Technology**

**Rustamji Institute of Technology Tekanpur, Gwalior (MP)**

**Session 2020-2021**

# **RUSTAM JI INSTITUTE OF TECHNOLOGY**



**Department of Information Technology**

## **Session 2020-2021**

CERTIFICATE

This is to certify that minor project work entitled **“Coronalysis”**

### is a bona fide work carried out by

**Chirag Sharma 0902IT181011**

**Mansi Jain 0902IT181026**

**Minesh Sarawogi 0902IT181027**

**Vanshika Sikarwar 0902IT181051**

**Head Of Department**

Prof. Anand Jha

Assistant Professor

Department of IT

R.J.I.T. TEKANPUR

**Project Guide**

Prof. Aradhana Saxena

Designation

Department of IT

R.J.I.T. TEKANPUR

# 

# 

# **Abstract**

The project aims to predict the future conditions of India by visualizing available datasets in the form of graphs, bar graphs and pie charts which can aid experts and law makers in defining the future course of actions to fight against coronavirus using Pandas for data manipulation, Numpy for numerical computation and Matplotlib for visualization with Seaborn for better and interactive visualization.

1. **Table of Content**

| **S. No.** | **Topic** | **Page No.** |
| --- | --- | --- |
| **01** | **Abstract** | **03** |
| **02** | **Table of Contents** | **04** |
| **03** | **Introduction** | **05** |
| **04** | **Design/ Implementation** | **06** |
| **05** | **Testing/Result and Analysis** | **20** |
| **06** | **Conclusion and Future Plans** | **24** |

1. **Introduction**

COVID-19 is the greatest global humanitarian challenge the world is facing since World War II. The Corona virus has spread widely and the number of cases are rising daily not only in India but in the whole world.

Government of India (GoI) has been taking many precautions and scientists are working for the vaccine of this disease but even after so many efforts the cases are only rising. Even if they started declining after the first wave due to community immunity, new virus strains started emerging and brought havoc in the form new waves. Experts can only tell much about when this pandemic is going to end and when we’ll finally be able to go outside our homes without fear.

So, we as the future engineers and responsible citizens of our country, have created a project to help the government and experts in devising ways effectively in tackling the prevailing pandemic. We used the current technology of ML and Data Science in our project Coronalysis to predict the number of future cases of the COVID 19 in India, particularly the Maharashtra state by using the data sets which are currently available with the help of Random Forest algorithm**.**

We have done the visualization of data in the form of graphs, pie charts and bar graphs which can aid experts and lawmakers in defining the future course of actions to fight against coronavirus using Pandas for data manipulation, Numpy for numerical computation and Matplotlib for visualization with Seaborn for better and interactive visualization.

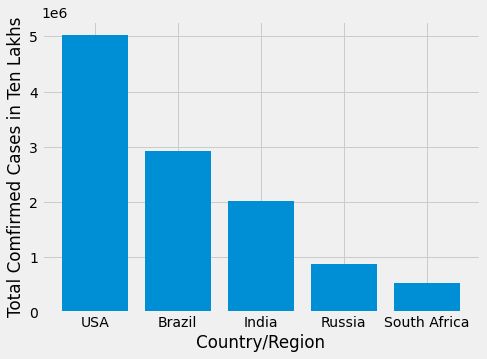
1. **Design/Implementation**

**Part 1: Exploring World wide data**

Visualizing: Worldwide COVID-19 cases

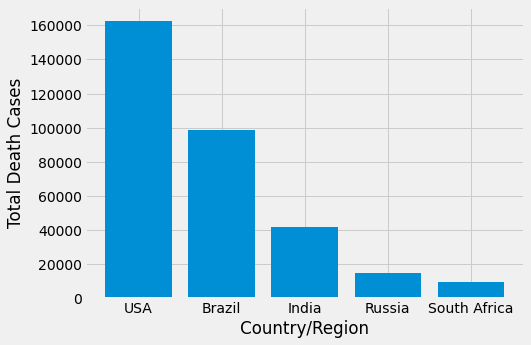
The trend of Active,Death,Total Confirmed and Recovered Cases in early phase of CORONA:

Total Confirmed-



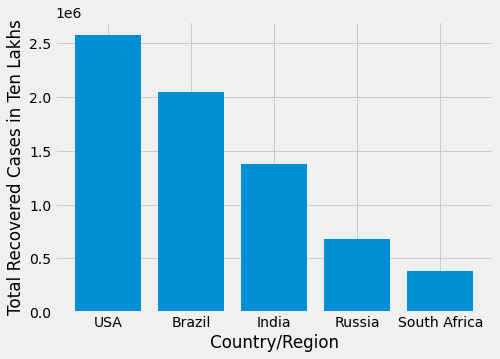
(1)

Total Death Cases



(2)

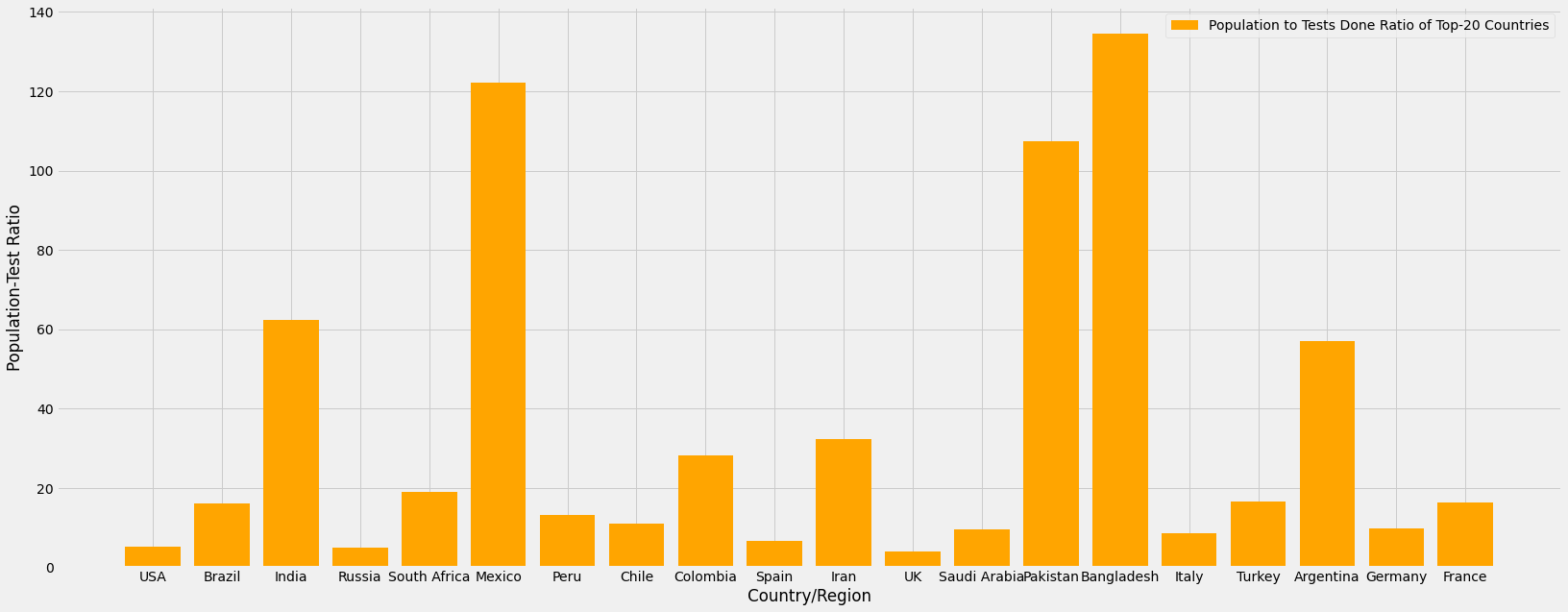
Total Recovered



(3)

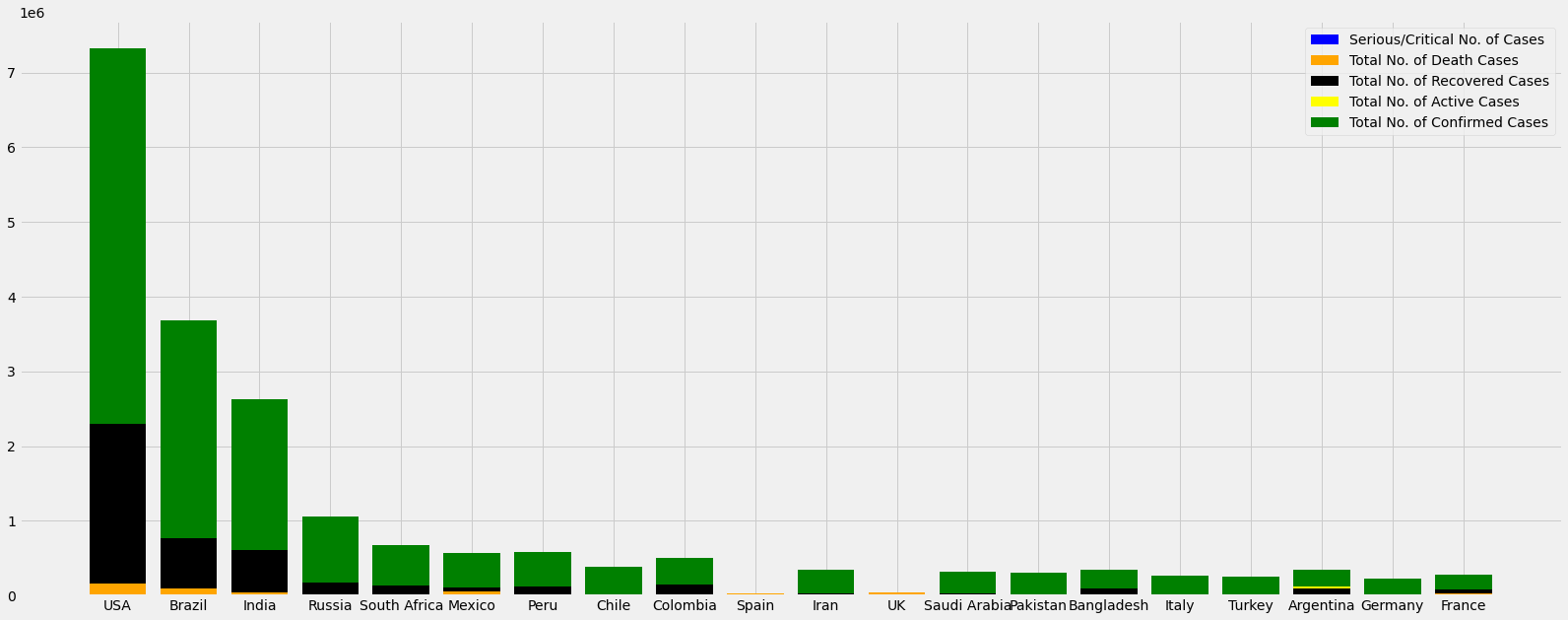
* Through Upper Analysis, We found that USA has maximum TotalCases,Active Cases,Recovered Cases and Death Cases
* Top 5 total Cases Countries are USA,Brazil,India,Russia and SouthAfrica

***BarPlot Representation of Population to Tests Done Ratio of Top 20 Countries***



(4)

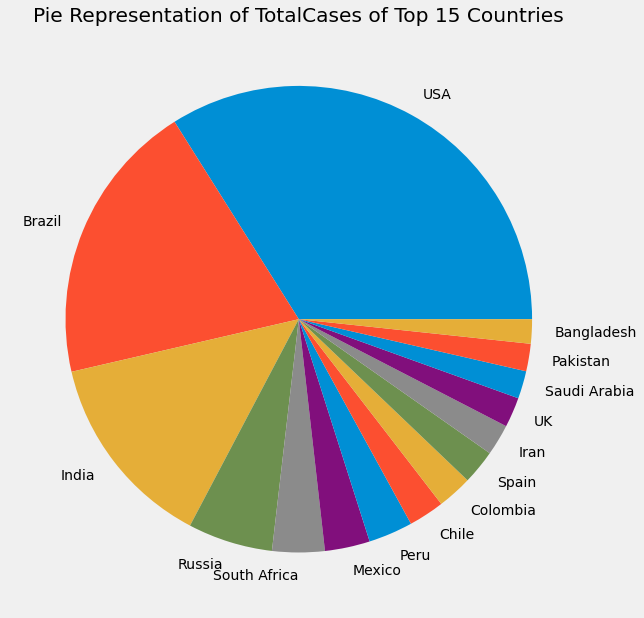
***BarPlot Representation of CoronaViruses Cases w.r.t Time***



(5)

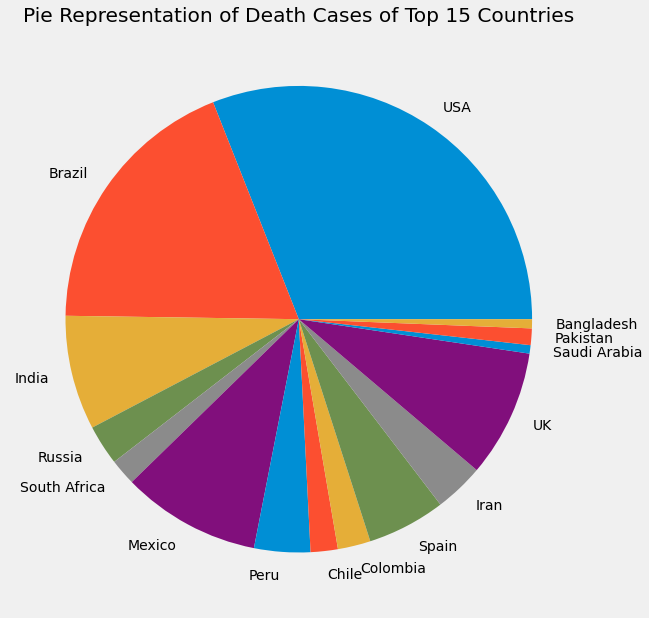
**Pie Chart Representation of stats of worst affected countries**

***Pie Representation of TotalCases of Top 15 Countries***



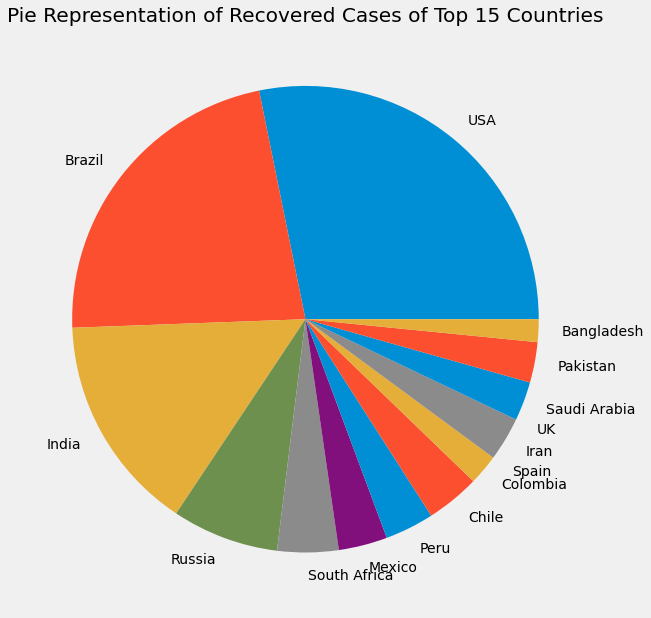
(6)

***Pie Representation of Death Cases of Top 15 Countries***



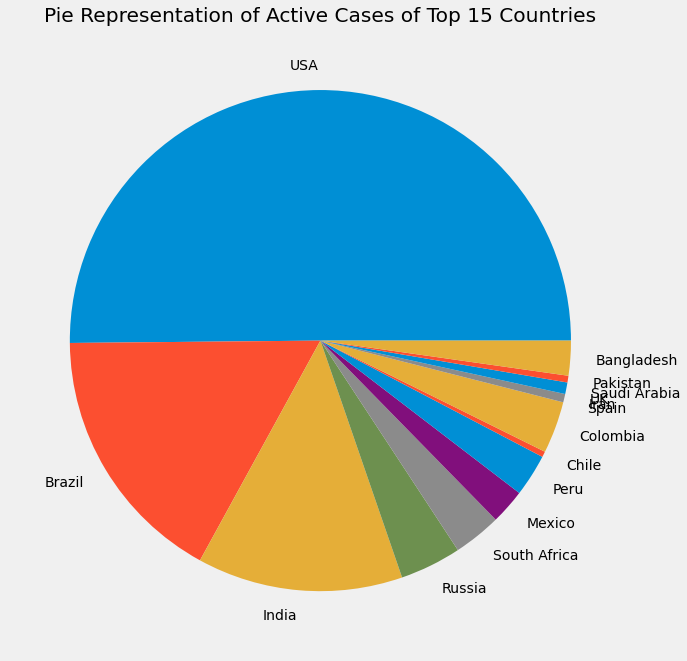
(7)

***Pie Representation of Recovered Cases of Top 15 Countries***



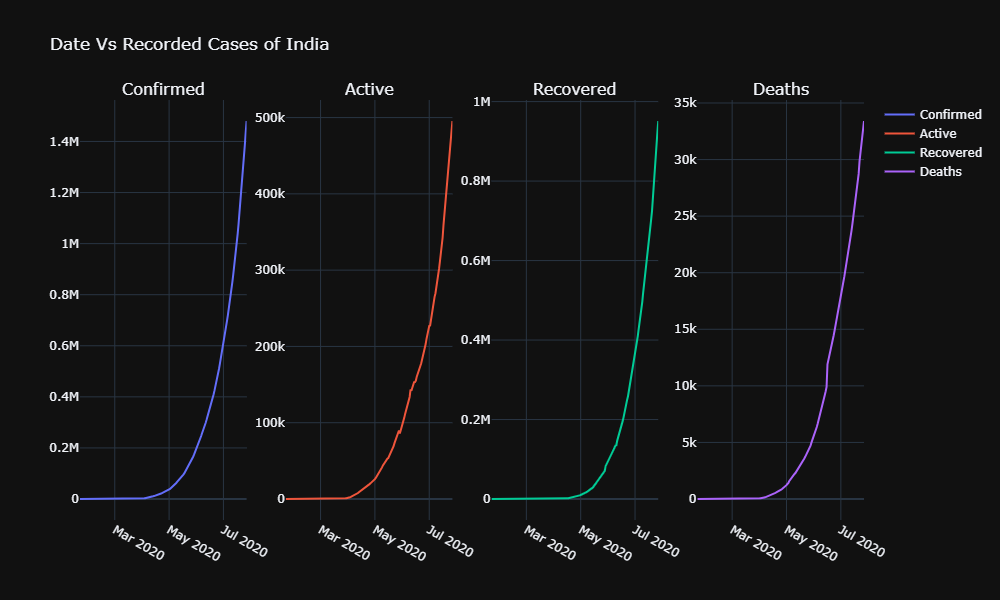
(8)

***Pie Representation of Active Cases of Top 15 Countries***



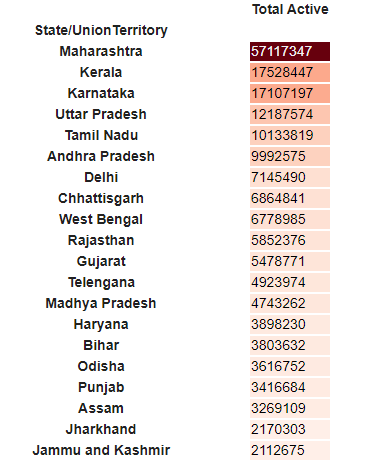
(9)

***Visualize Confirmed, Active, Recovered , Deaths Cases(entire statistics ) of India:***



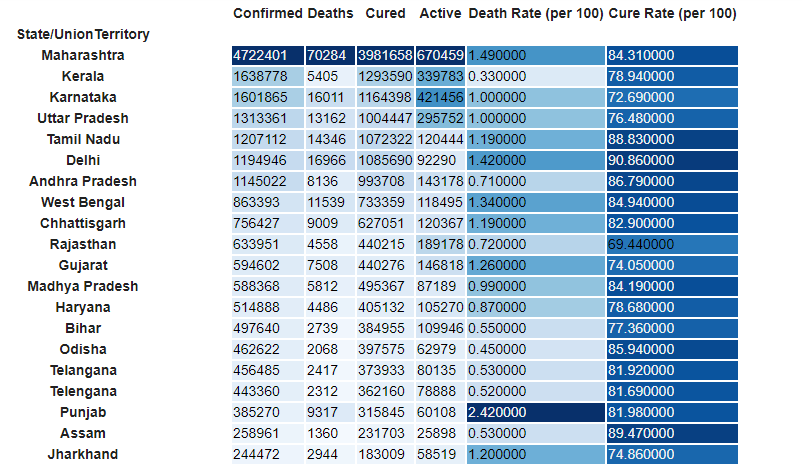
(10)

***Top-20 States Of India which have the highest Number of Total Active Cases in the First Wave of Covid19.***



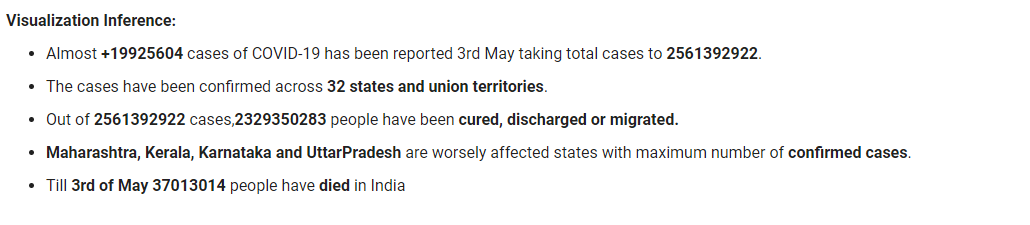
(11)

***Top-20 States of India which are worsley affected, With the confirmed, death, cured and active cases, as well as death and cure rate per 100 are also mentioned.***



(12)

***This is the inference (Till 3rd May),which we have accomplished after visualizing the state wise Covid19 Data of India.***



# **Part 2: Analysing the past/present condition in India**

How did it start in India?

The first COVID-19 case was reported on 30th January 2020 when a student arrived in Kerala from Wuhan. Just in next 2 days, Kerala reported 2 more cases.

For almost a month, no new cases were reported in India, however, on 2nd March 2020, five new cases of coronavirus were reported in Kerala again and since then the cases have been rising affecting all the states,

Here is a brief timeline of the cases in India.

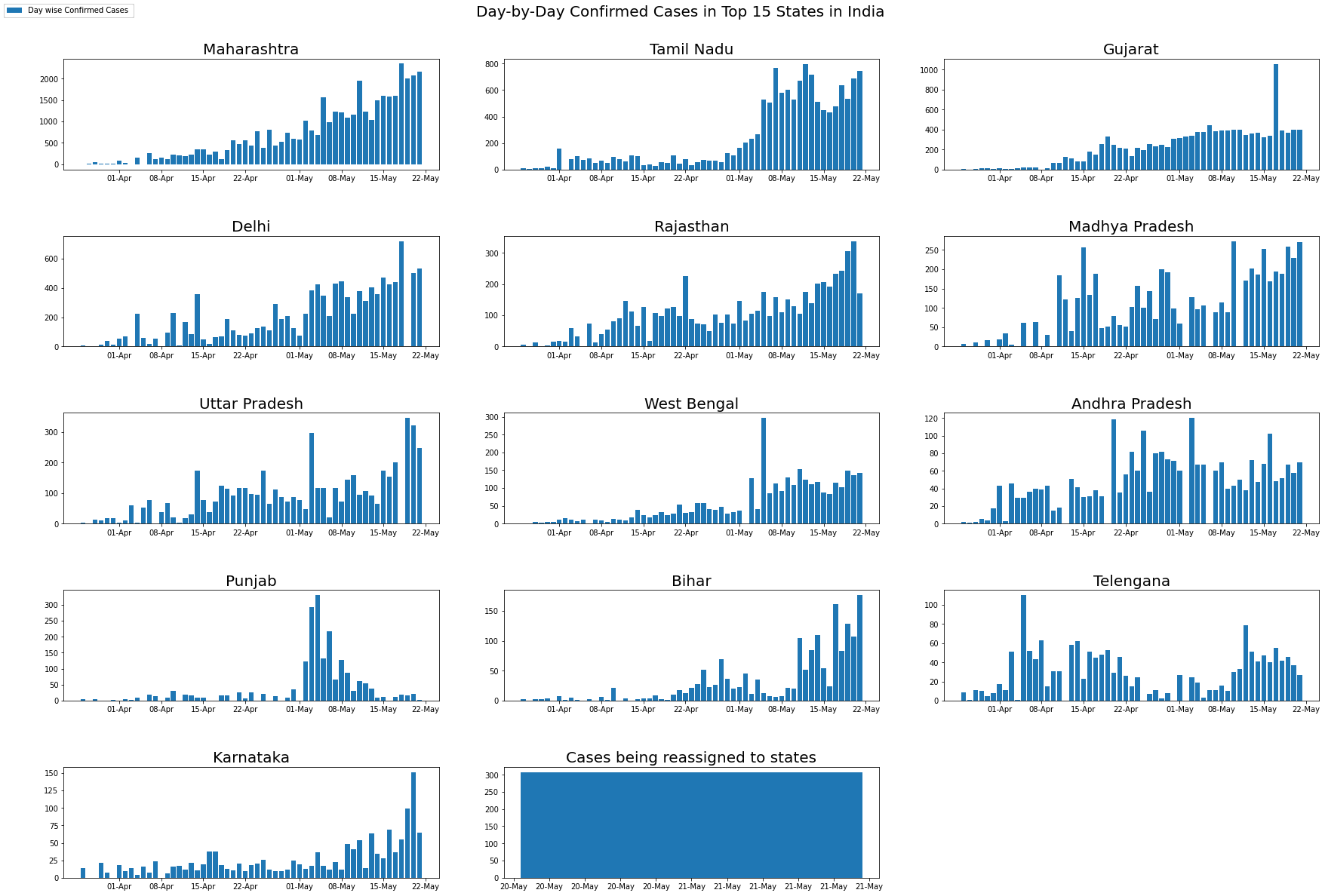


(13)

How is AI-ML useful in fighting the COVID-19 pandemic?

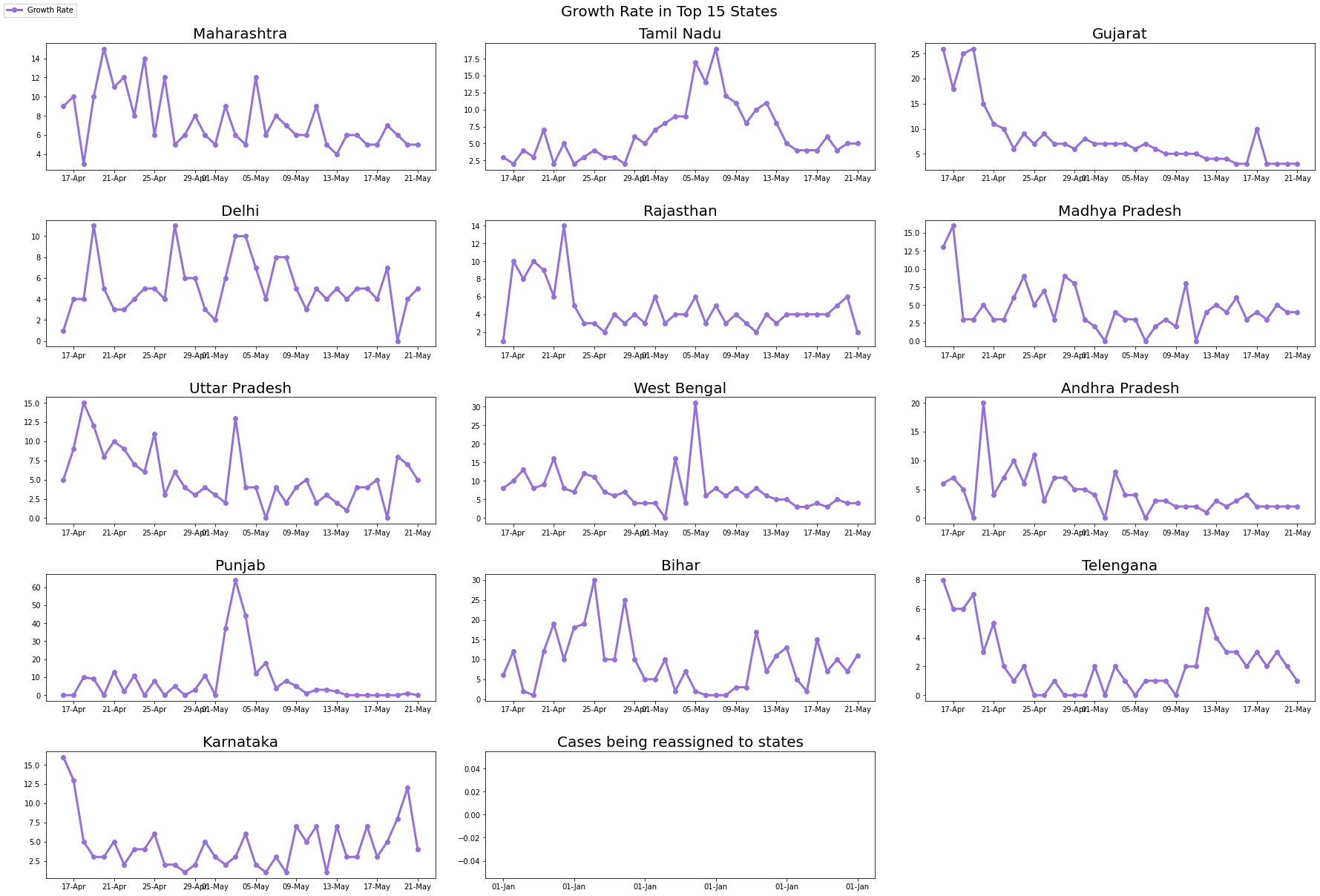
* Medical resource optimization
* Ensuring demand planning stability
* Contact tracing
* Situational awareness and critical response analysis

Day-by-Day Confirmed Cases in Top 15 States in India:



(14)

Growth Rate in top 15 States in India :



(15)

### Exploring Urban and Rural Healthcare Facility :

(16)

Data Source:

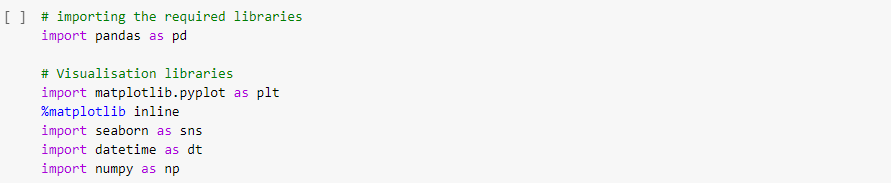
* <https://www.mohfw.gov.in/>
* <https://drive.google.com/drive/folders/1QNgK2juTicvO4Q9y_2EDRKwW8J4OEldB?usp=sharing>
* <https://drive.google.com/file/d/1QXtDtJWSMXajEAxtpo2GsmvCWFjQmOti/view?usp=sharing>
* <https://drive.google.com/file/d/1Jm1QMfr8pYv07InnXwl5MS6Ul0jMsUnp/view?usp=sharing>
* <https://colab.research.google.com/drive/1bmTi1ZOp1_bHigd-qbkUgYXUBkEj1_Wk#scrollTo=8sugsh4VxhQq>
* <https://colab.research.google.com/drive/1kNBq4Btkb1fUUZS22g5EPot8cKPUPz9Q#scrollTo=akmgm27xte1y>
* https://colab.research.google.com/drive/1bHzSrQRH7OopyCszd75vdUDrQYVkdVVW

Note: 1. Prediction may vary, because of the non availability of real-time Data Sets.

2. Here we can make our prediction for all the states in India Only.

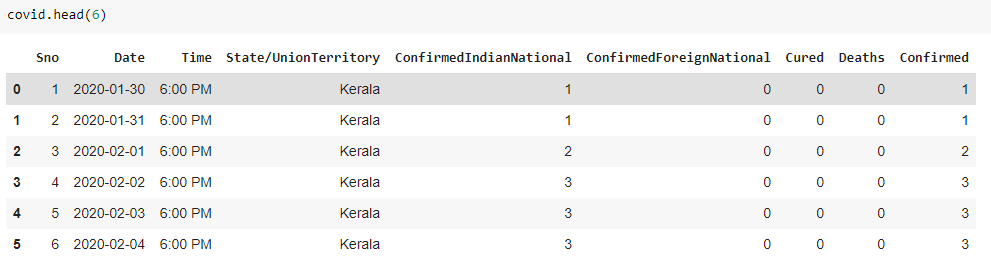
**Let's Start with the predictions :**

Here we have imported some useful python libraries to apply numerous computational operations on data and visualize that to extract some meaningful insights from that data.



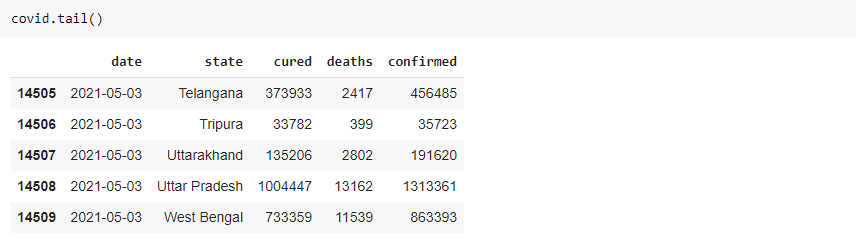
(17)

Here, we want to show the first few rows of covid data



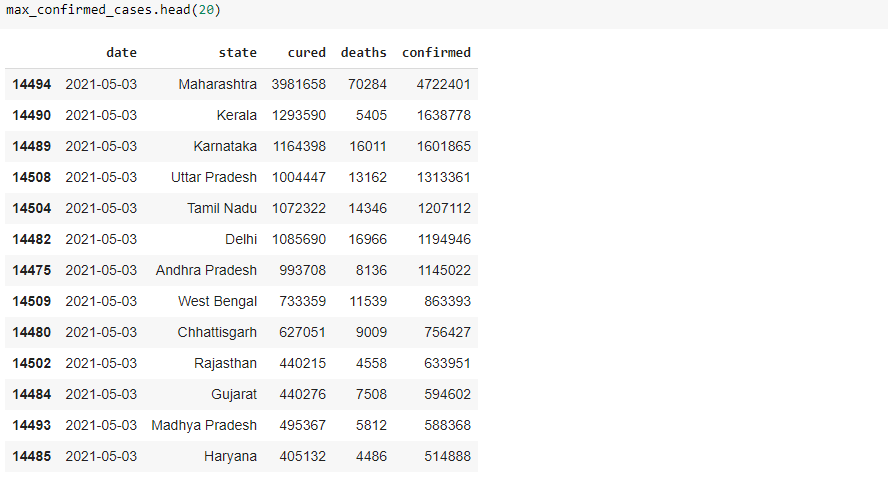
(18)

It will return some last rows to make it clear that the data is upto which date



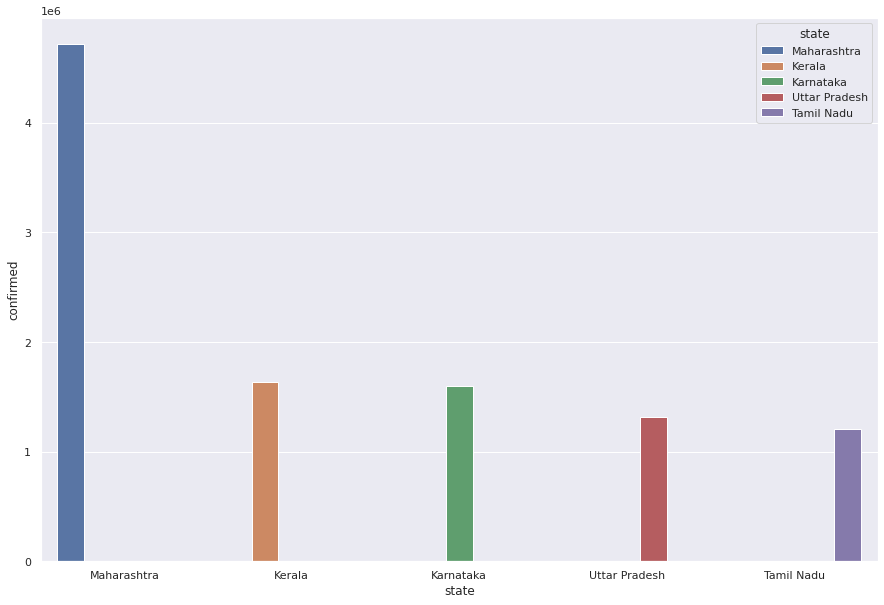
(19)

Here, we give the names of Top-20 states of India which have Maximum Confirmed Cases according to our datasets.



(20)

This is the bar which is used to show confirmed cases corresponding to Top-5 different states of India which have maximum confirmed cases.

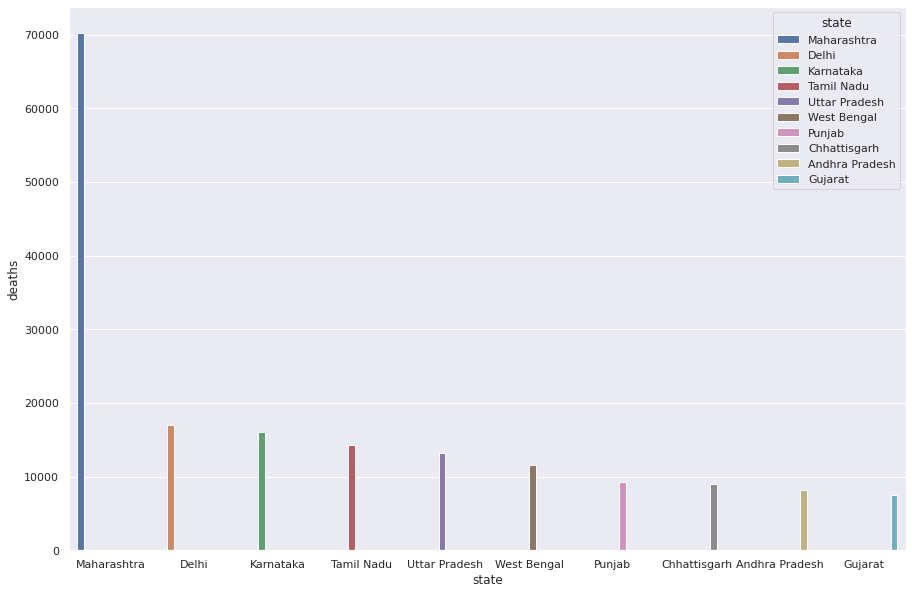


(21)

Here, we give the names of Top-10 states of India which have Maximum Death Cases reported according to our datasets



(22)

This is the bar plot which shows Top-10 different states of India which have maximum death cases.

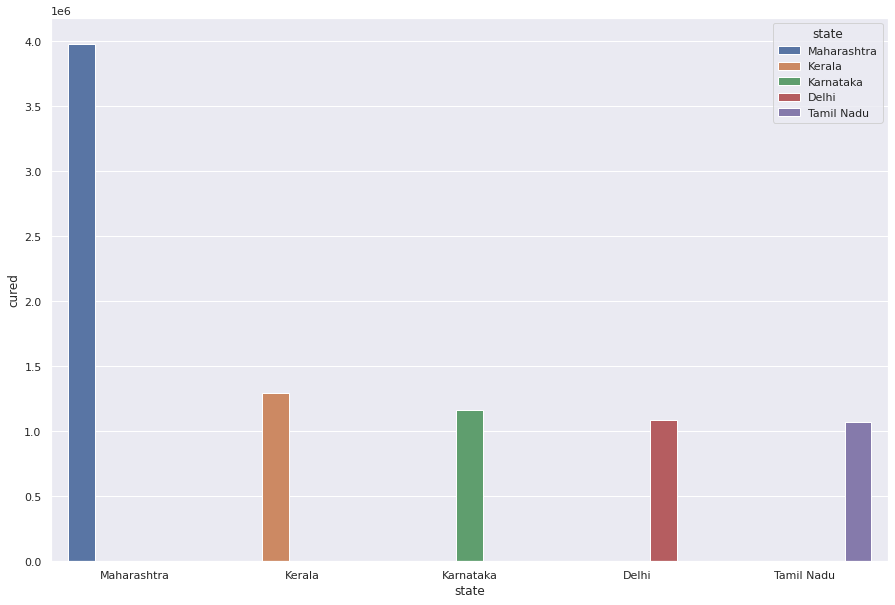
(23)

These are the some few states which have highest Covid-19 Recovery Rate in India



(24)

The following bar plot shows Top-5 States which have the highest Recovery Rate with the number of Cured Patients till now.

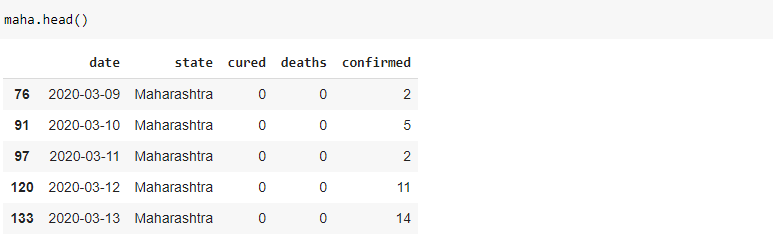


(25)

1. **Testing/Result and Analysis**

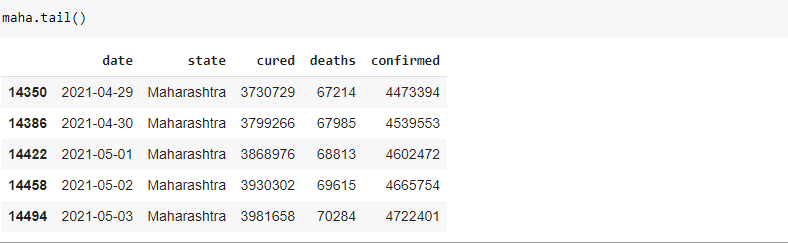
Let’s Predict on a single state => Maharashtra :

This statement will give us the first few records of Maharashtra which have different columns like date,state,cured cases,death cases as well as confirmed cases.



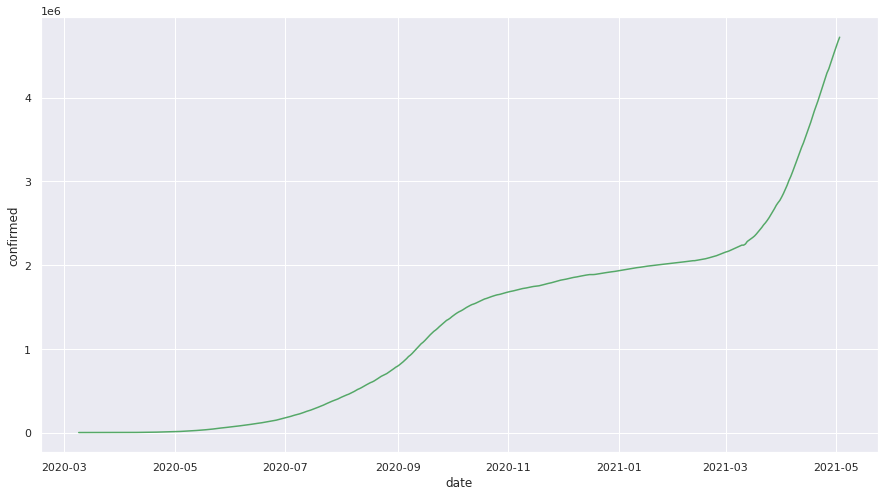
(26)

Tail() function returns the last few records of maha dataframe.



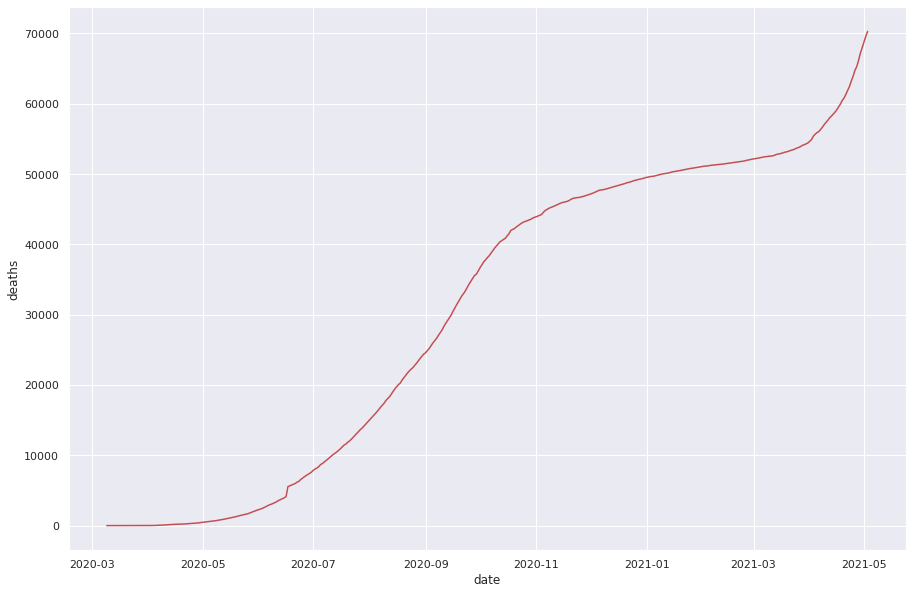
(27)

This is the line plot showing the confirmed cases of Maharashtra corresponding to the date on which confirmed cases have been reported.



(28)

This is the line plot showing the death cases of Maharashtra corresponding to the date on which death cases have been reported.



(29)

**Part 3: Predicting the future condition in India**

In these following lines of code, we have used a random forest regressor algorithm which is used for regression as well as classification problems.A Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees.

Here we will predict the number of confirmed cases on the basis of the date specified.Here x represents the independent variable which has to be used to calculate the dependent variable and y represents the dependent variable which has to be predicted.In model validation, we randomly split the complete data into training and test sets. Then Perform the model training on the training set and use the test set for validation purpose, ideally split the data into 70:30 or 80:20.Here we split the data into 70:30 for model training and validation.



(30)

*fitting* is equal to *training*. Then, after it is trained, the model can be used to make predictions, usually with a .predict() method call.Here we have used fit function to fit the model to the training set of the data.



(31)

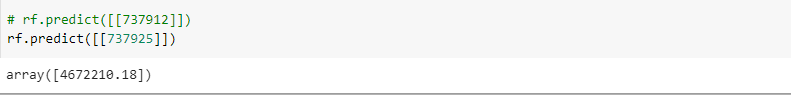
Use Timestamp.toordinal() function to return the Gregorian ordinal for the given Timestamp object.



(32)

We predicted the confirmed cases of Maharashtra after changing the date into Gregorian ordinal.Because we can’t predict the confirmed cases directly without changing the timestamp

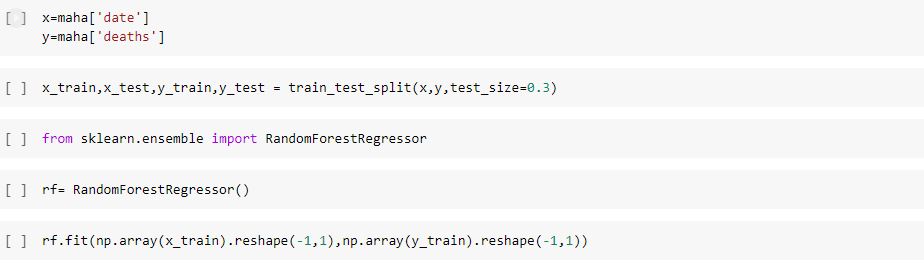
into it.



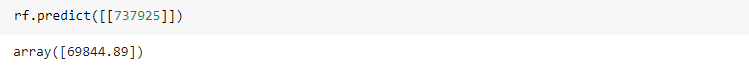
(33)

***We have come into this conclusion that On 21 May 2021 , Maharashtra must have more than 46 lakhs Confirmed Cases.***

Here we will predict the number of death cases on the basis of the date specified.Here x represents the independent variable which has to be used to calculate the dependent variable and y represents the dependent variable which has to be predicted.We have used random forest regressor algorithm for model preparation.



(34)

We predicted the death cases of Maharashtra after changing the date into Gregorian ordinal.

(35)

***We have predicted that Maharashtra Death Cases must rise to around 70 thousand on 15 May 2021.***

1. **Conclusion and Future Plan**

We have analyzed the datasets of coronavirus cases in India. We analyzed the day-by-day confirmed cases and growth rate in top states in India. We also analyzed rural and urban healthcare facilities to better analyze the prevailing conditions in India.

We visualized the Confirmed cases, Active cases and Deaths in India. After the visualization of datasets in the form of graphs, pie charts and bar graphs, we found that the first COVID-19 case was reported on 30th January 2020 when a student arrived in Kerala from Wuhan. Just in next 2 days, Kerala reported 2 more cases. For almost a month, no new cases were reported in India, however, on 2nd March 2020, 5 new cases of coronavirus were reported in Kerala again and since then the cases have been rising affecting all the states.

After the analysis and visualization of data, we predicted the future conditions of India, specifically for the Maharashtra state. Thus, we have predicted the following -

i) On 21 May 2021, Maharashtra must have more than 46 lakhs confirmed cases.

ii) Maharashtra death cases must rise to around 70 thousand on 15 May 2021.

After doing all these predictions, we have thought for future plans that we can make this prediction model more accurate and precise. So we have planned that we will improvise our data sets further, analyzing and retrieving the data including data-sets regarding hospital beds availability and requirement also, keeping the count of vaccinated and unvaccinated will help us to detect further status. Hence, more data will make us predict better.

----------------------------------------------------xxx---------------------------------------------------