Covid-19

Spread of virus in india.

Step 1:

In the first step of this project I downloaded a csv file from OurWorldinData Website and Uploaded that data into MySQL workbench.

StLater I tried to open a file using Spreadsheet software but eventually I realized that it will take a long time to process all the data in the spreadsheet. Hence I decided to Use MySQL workbench.

Step 2:

I applied Few basic queries in SQL and did a data cleaning process which included the following.

- Created new columns to access the information quickly for visualization.
- Removed the rows containing data about countries other than India.
- Added zero in place of Nan(Null) values.
- Saved data in Covid_deaths_cleaned_India.csv

Step 3:

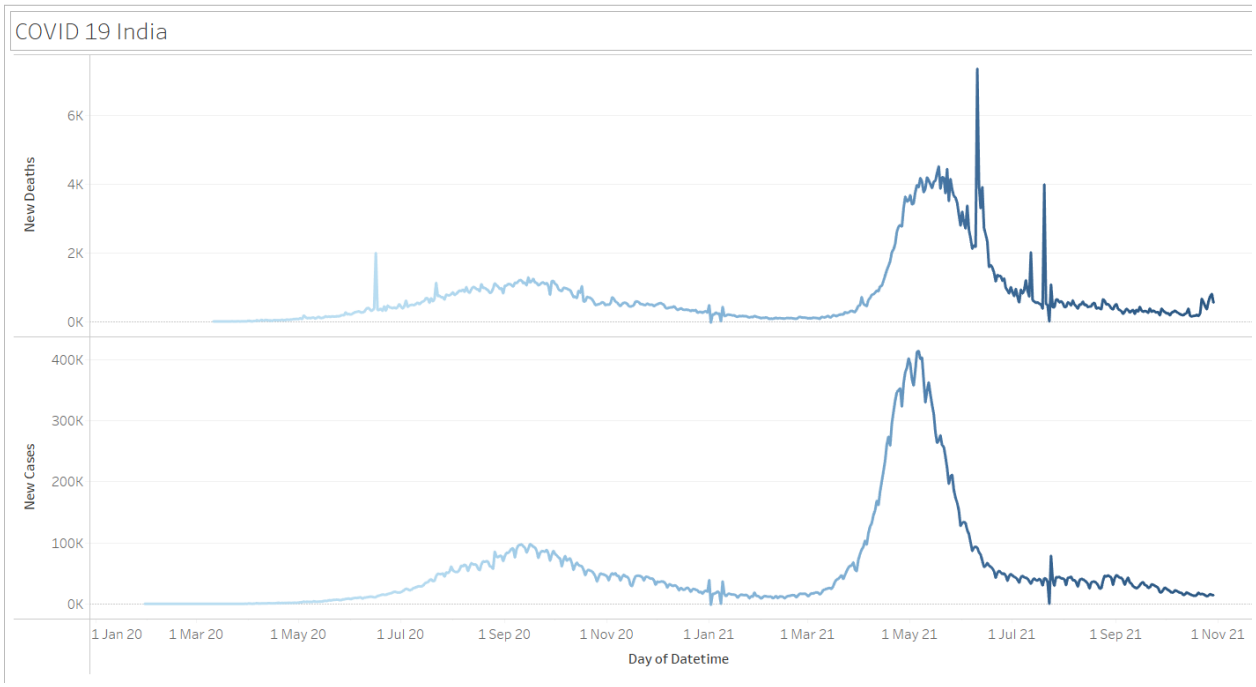
List of queries;

```
1 •  SELECT
2        *
3    FROM
4        covid.coviddeathsindia;
```

```sql
###max cases on a day
SELECT
    MAX(new_cases)
FROM
    covid.coviddeathsindia;

 ###max new cases per million
 SELECT
     MAX(new_cases_per_million), MAX(new_deaths_per_million)
 FROM
     covid.coviddeathsindia;


### total case vs total deaths
SELECT
    location,
    datetime,
    total_cases,
    total_deaths,
    (total_deaths / total_cases) * 100 AS DeathPercentage
FROM
    coviddeathsindia
ORDER BY datetime;


###what % got covid
SELECT
    location,
    datetime,
    total_cases,
    population,
    (total_cases / population) * 100 AS PercentgotCOVID
FROM
    coviddeathsindia
ORDER BY datetime;

###on what date most cases are detected? and on what day most deaths were recorded?

SELECT
    datetime, new_cases, new_deaths
FROM
    coviddeathsindia
WHERE
    new_cases = 414188 OR new_deaths = 7374
```

Step 4:

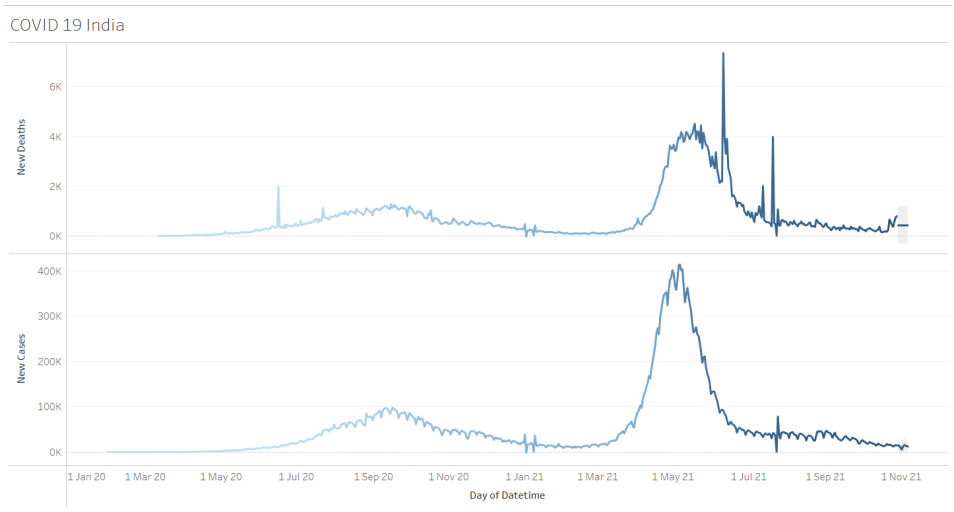
During this step I uploaded the data into Tableau Desktop.
I plotted a graph of Deaths in India and cases in India.



Now data I used was downloaded earlier so I had data until Nov 2021.
So why not use Predict features on tableau desktop to predict the spread of cases and deaths in India.

Step 5:

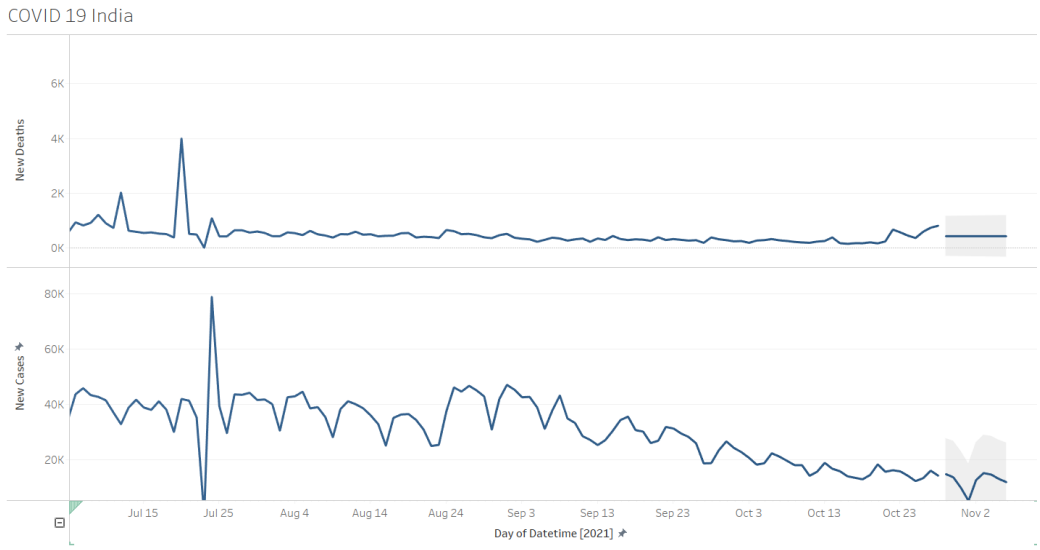
I Used the Forecast method in tableau. Let's see the result.



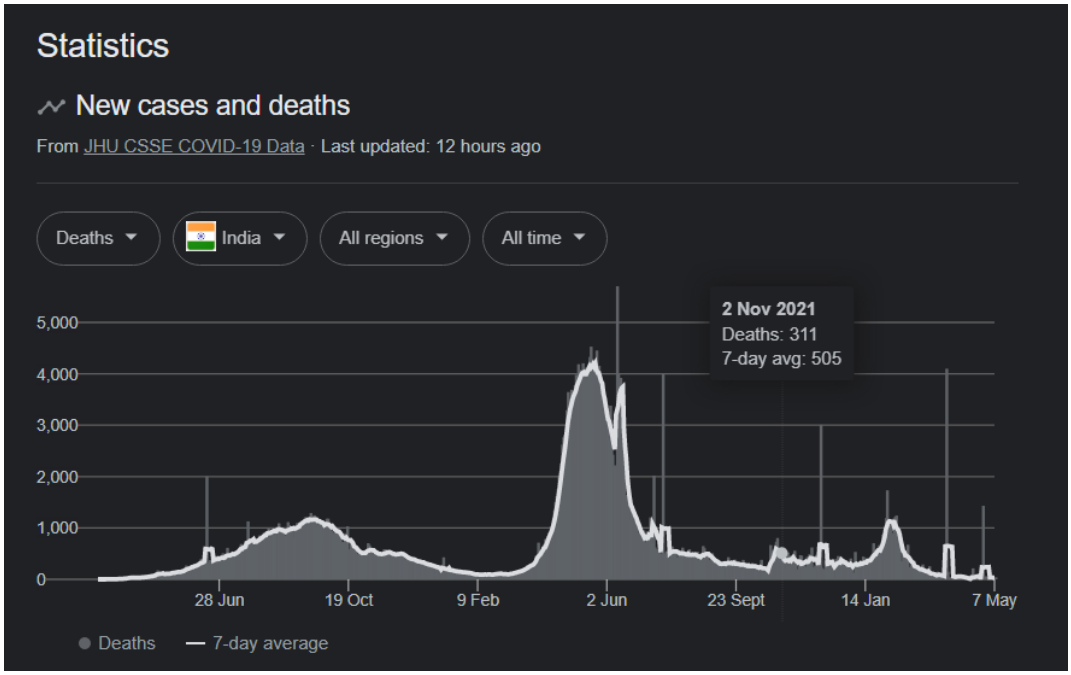
Those tiny Gray boxes Are the result.

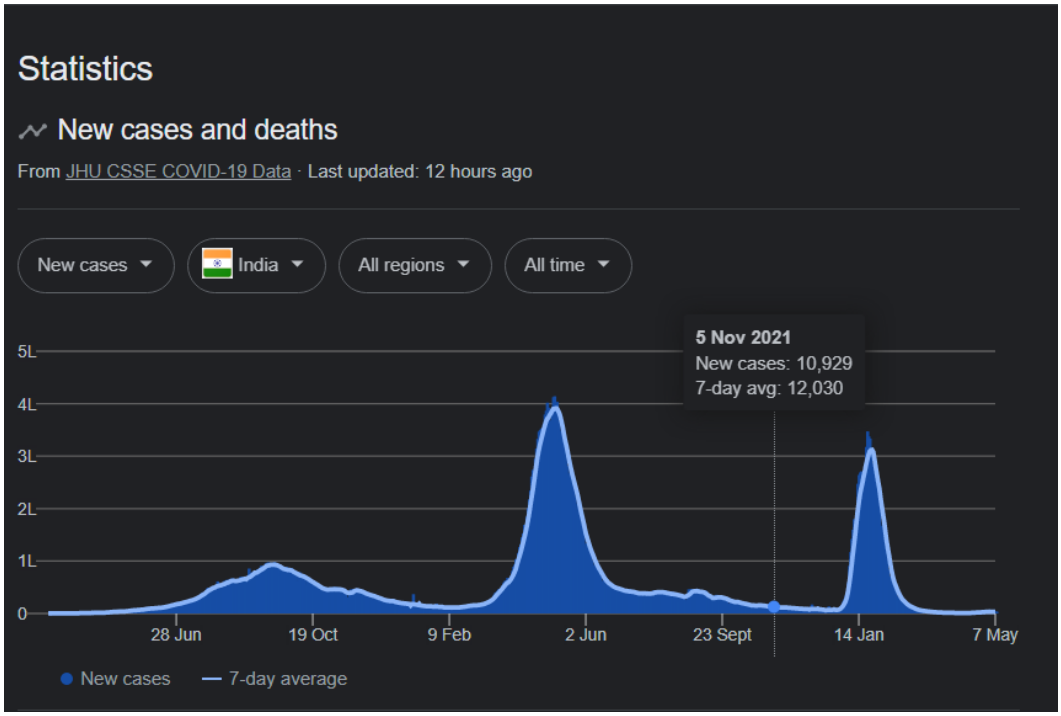
It really looks Unclear but let's interpret the result.
Line after 1 NOV 2021 is our prediction of cases in India.

Here's a closer look.



Now let's check with google to Test the results.





Our prediction graphs look very similar but let's check The exact info.

Info	Date	Actual	Predicted	Difference in %
Death	2 Nov 2021	311	413	24.7%
Cases	5 Nov 2021	10929	13206	17.2%

Conclusion:

Forecast predicted the trend correctly but there was an error in the results.