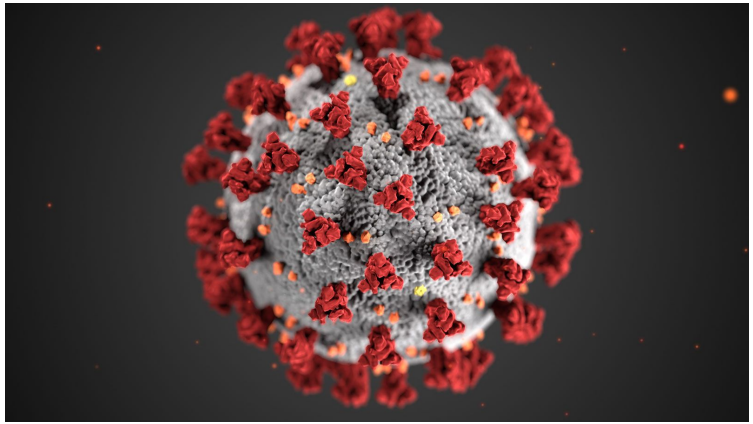
Live Covid Data Visualization

By : Raghav Mundra

Index :

Covid 19

Coronavirus disease 2019 (COVID-19) is a contagious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first known case was identified in Wuhan, China, in December 2019.The disease has since spread worldwide, leading to an ongoing pandemic.



Covid 19 : India

- The COVID-19 pandemic in India is a part of the worldwide pandemic of coronavirus .As of 27 September 2021, according to official figures, India has the second-highest number of confirmed cases in the world (after the United States of America) with 33,678,786 reported cases of COVID-19 infection and the third-highest number of COVID-19 deaths (after the United States and Brazil).
- The first cases of COVID-19 in India were reported on 30 January 2020 in three towns of Kerala, among three Indian medical students who had returned from Wuhan, the epicenter of the pandemic.
- Lockdowns were announced in Kerala on 23 March, and in the rest of the country on 25 March 2020.

Covid 19 : India

- On 10 June, India's recoveries exceeded active cases for the first time.Infection rates started to drop in September, along with the number of new and active cases.
- Daily cases peaked mid-September with over 90,000 cases reported per-day, dropping to below 15,000 in January 2021.
- A second wave beginning in March 2021 was much more devastating than the first, with shortages of vaccines, hospital beds, oxygen cylinders and other medical supplies in parts of the country.By late April, India led the world in new and active cases. On 30 April 2021, it became the first country to report over 400,000 new cases in a 24-hour period.

Covid 19 : India

India began its vaccination programme on 16 January 2021 with AstraZeneca vaccine (Covishield) and the indigenous Covaxin.Later, Sputnik V and the Moderna vaccine was approved for emergency use too.

As of 2nd December 2021, the country had administered over 1.1 billion vaccine doses.

It is speculated that by the end of first quarter of 2022,Covid 19 may reach an endemic stage in India rather than completely disappear i.e. India may be in some stage of endemicity where the country learns to live with the virus.

Project Overview

- The aim of this project was to scrape live covid data from trustable sources,generate visualizations and make a live dashboard.For this project ,State-wise data about covid 19 was used and insights were made regarding covid 19
- Tools used :
  - Languages : Python
  - Libraries : Pandas , bs4 , requests , json , folium and geopy
  - Softwares/IDEs : Jupyter notebook and Tableau

Stages of development

1. Data Scraping : This is the first step in the project which was done with help python and its libraries . In this step live data is scraped from the internet and converted to csv file
2. Data Cleaning : In this step the data is cleaned and feature engineering is done on data so that is ready for visualization
3. Data Visualization : This was the last step in which visualizations were generated and insights were made .

Data scraping

Webscrapping data

```
#Fetching covid vaccination data
path1='https://www.mygov.in/sites/default/files/covid/vaccine/vaccine_counts_today.json'
page1 = requests.get(path1)
soup1 = BeautifulSoup(page1.content, 'html.parser')
site_json=json.loads(soup1.text)
site_json['vacc_st_data'][0]
vaccine_data=pd.DataFrame(columns=['state','dose1','dose2','total_doses'])
c=0
for i in site_json['vacc_st_data']:
  vaccine_data.loc[c]=[i['covid_state_name'],i['dose1'],i['dose2'],i['total_doses']]
  c+=1
#Fetching covid cases data
path2='https://www.mygov.in/sites/default/files/covid/covid_state_counts_ver1.json'
page2 = requests.get(path2)
soup2 = BeautifulSoup(page2.content, 'html.parser')
site_json2=json.loads(soup2.text)
active_cases=site_json2['Active']
state_names=site_json2['Name of State / UT']
total_deaths=site_json2['Death']
cured_cases=site_json2['Cured/Discharged/Migrated']
total_cases=site_json2['Total Confirmed cases']
covid_data=pd.DataFrame(columns =['state', 'total_cases','active_cases','discharged_cases','total_deaths'])
covid_data.loc[0]=[state_names['0'],total_cases['0'],active_cases['0'],cured_cases['0'],total_deaths['0']]
for i in range(0,36):
  covid_data.loc[i]=[state_names[str(i)],total_cases[str(i)],active_cases[str(i)],cured_cases[str(i)],total_deaths[str(i)]]
```

Data Cleaning

: `covid_data`

:

	state	total_cases	active_cases	discharged_cases	total_deaths
0	Andaman and Nicobar	7683	5	7549	129
1	Andhra Pradesh	2073093	2149	2056501	14443
2	Arunachal Pradesh	55279	35	54964	280
3	Assam	617039	2612	608319	6108
4	Bihar	726225	32	716530	9663
5	Chandigarh	65471	66	64585	820
6	Chhattisgarh	1006833	316	992924	13593

`vaccine_data`

	state	dose1	dose2	total_doses
0	Andaman and Nicobar	296964	258561	555525
1	Andhra Pradesh	34991943	25032907	60024850
2	Arunachal Pradesh	795139	612865	1408004
3	Assam	21209271	12160743	33370014
4	Bihar	54637126	27792850	82429976
5	Chandigarh	942626	635712	1578338
6	Chhattisgarh	16510938	9528852	26039790

```
covid_data[['dose1','dose2','total_doses']] = vaccine_data[['dose1','dose2','total_doses']]
covid_data.head()
```

	state	total_cases	active_cases	discharged_cases	total_deaths	dose1	dose2	total_doses
0	Andaman and Nicobar	7683	5	7549	129	296964	258561	555525
1	Andhra Pradesh	2073093	2149	2056501	14443	34991943	25032907	60024850
2	Arunachal Pradesh	55279	35	54964	280	795139	612865	1408004
3	Assam	617039	2612	608319	6108	21209271	12160743	33370014
4	Bihar	726225	32	716530	9663	54637126	27792850	82429976

Data Cleaning

```
geolocator = Nominatim(timeout=10, user_agent = "myGeolocator")
covid_data['gcode'] = covid_data.state.apply(geolocator.geocode)
covid_data['lat'] = [g.latitude for g in covid_data.gcode]
covid_data['long'] = [g.longitude for g in covid_data.gcode]
covid_data.head()
```

	state	total_cases	active_cases	discharged_cases	total_deaths	dose1	dose2	total_doses	gcode	lat	long
0	Andaman and Nicobar	7683	5	7549	129	296964	258561	555525	(Andaman, India, (10.0001051, 93.0000194))	10.000105	93.000019
1	Andhra Pradesh	2073093	2149	2056501	14443	34991943	25032907	60024850	(Andhra Pradesh, India, (15.9240905, 80.1863809))	15.924091	80.186381
2	Arunachal Pradesh	55279	35	54964	280	795139	612865	1408004	(Arunachal Pradesh, India, (28.0937702, 94.592...	28.093770	94.592133

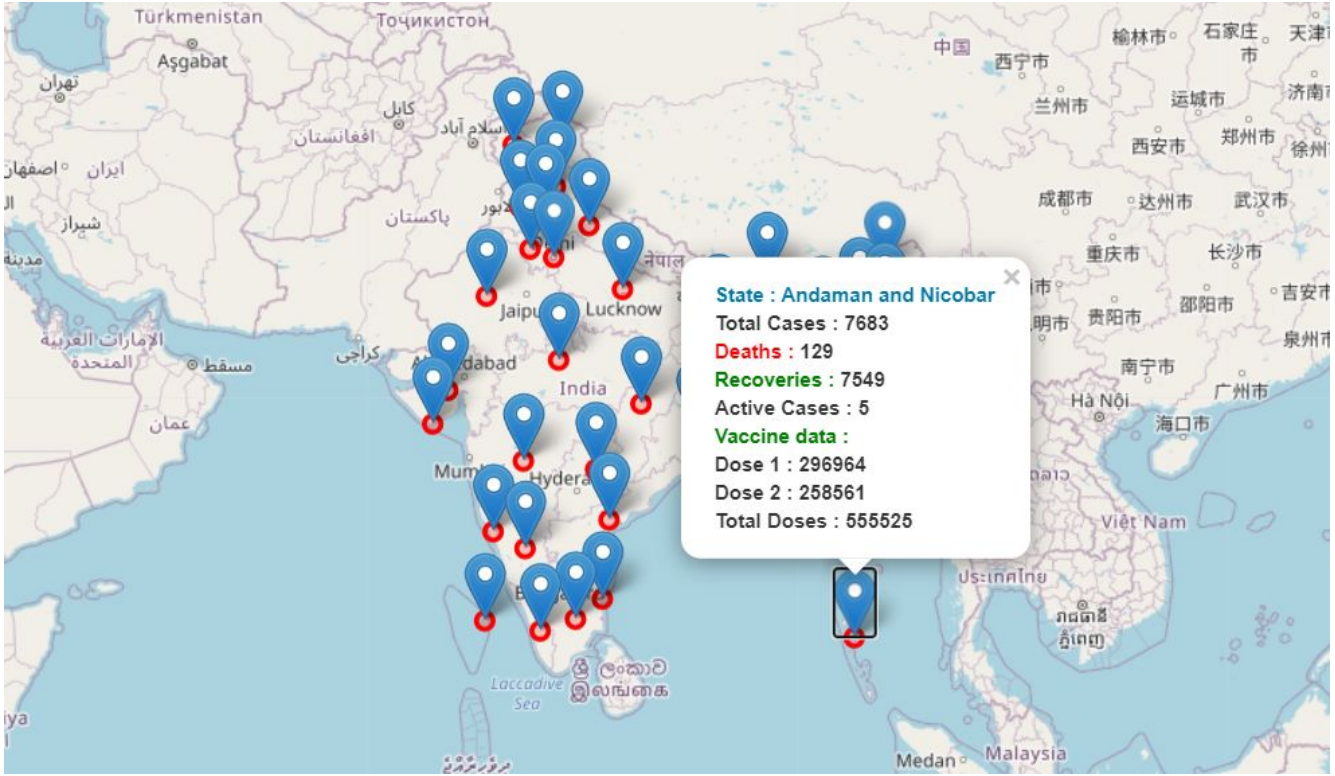
Data Visualization

```
for state,lat,long,total_cases,Death,Recov,Active,dose1,dose2,totaldoses in
zip(list(covid_data['state']),list(covid_data['lat']),list(covid_data['long']),
list(covid_data['total_cases']),list(covid_data['total_deaths']),list(covid_data['discharged_cases']
), list(covid_data['active_cases']),list(covid_data['dose1']),list(covid_data['dose2']),
list(covid_data['total_doses'])):

    #for creating circle marker

    folium.CircleMarker(location = [lat,long], radius = 5,color='red',fill =
True,fill_color="red").add_to(india)

    #for creating marker

    folium.Marker(location = [lat,long],
```

Data Visualization

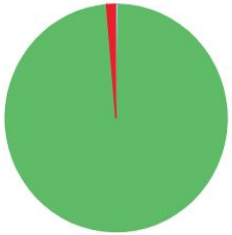
```
# adding information that need to be displayed on popup

                popup=folium.Popup(('<a href="https://www.youtube.com/watch?v=aJmaw3QKMvk"
target="_blank"><strong><b>State  : '+state+'</strong></a> <br>' +

                  '<strong><b>Total Cases : '+total_cases+'</strong><br>' +

                  '<strong><font color= red>Deaths : </font>'+Death+'</strong><br>' +

                  '<strong><font color=green>Recoveries : </font>'+Recov+'</strong><br>' +

                  '<strong><b>Active Cases : '+Active+'</strong><br>'+

                  '<strong><b><font color=green>Vaccine data : </font></strong> <br>' +

                  '<strong><b>Dose 1 : '+dose1+'</strong><br>' +

                  '<strong><b>Dose 2  : '+dose2+'</strong><br>' +

                  '<strong><b>Total Doses : '+totaldoses+'</strong><br>'),max_width=200)).add_to(india)
```

Data Visualization



Covid Data Visualization

Active Cases	Total Cases	Total Deaths
122,714	34,510,413	465,662

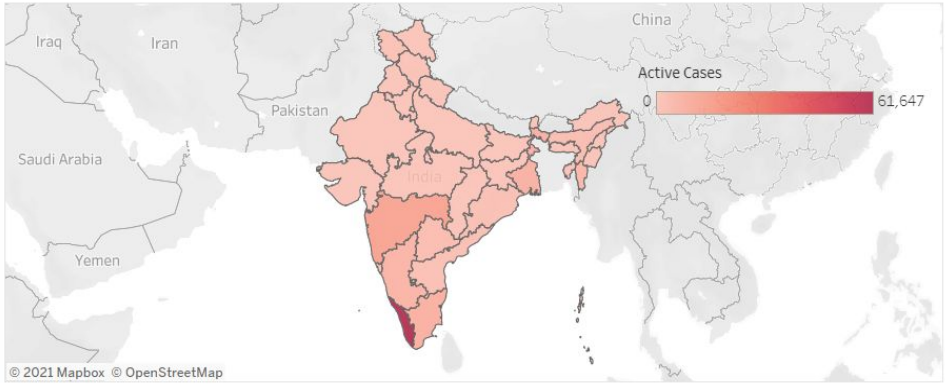


Death Rate :

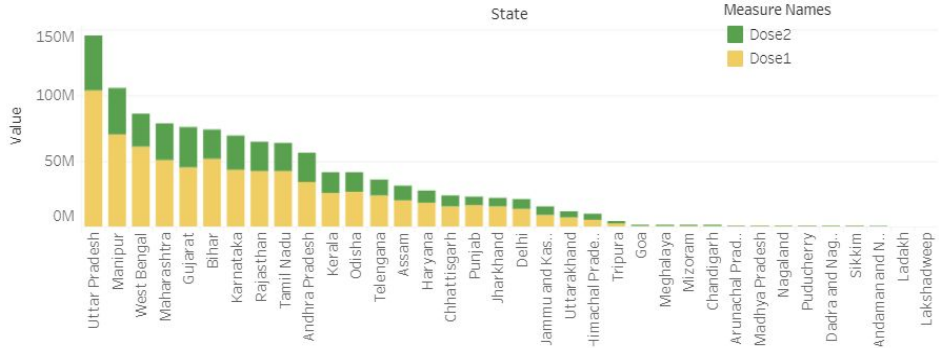
State	
Punjab	0.02750
Nagaland	0.02169
Uttarakhand	0.02152
Maharashtra	0.02123
Goa	0.01890
Delhi	0.01742
Meghalaya	0.01742
Himachal Pradesh	0.01692
Andaman and Nicobar	0.01681
Manipur	0.01570
Jharkhand	0.01472
Puducherry	0.01453
Chhattisgarh	0.01350

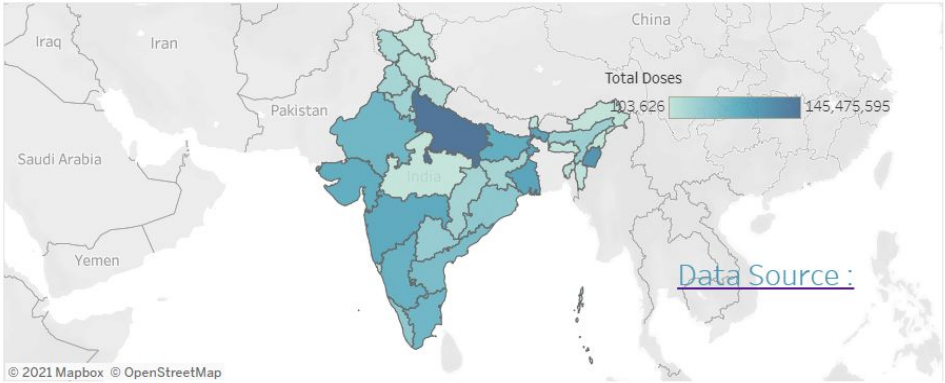
Death rate

0.00037 0.02750



Vaccination data





Insights

- Kerala is having the highest number of active cases despite its small size.
- Highest number of covid cases were found in Maharashtra .
- Out of 1.1 billion vaccine doses administered across India , Almost 30% of them were from states : Uttar Pradesh , Manipur and West Bengal .
- The vaccination drive in Manipur has been more successful than other states as they have got maximum percentage of their population vaccinated.
- The death rate i.e deaths/covid patient has been higher in smaller states like Punjab , Nagaland and Uttarakhand

Links

Jupyter notebook link : Jupyter Notebook Viewer (nbviewer.org)

Tableau link : Covid Data | Tableau Public

Data source : #IndiaFightsCorona COVID-19 in India, Vaccination, Dashboard , Corona Virus Tracker | mygov.in

Thank you