

**STATISTICS PROJECT  
ON  
EXPLORATORY DATA ANALYSIS ON COVID'19 VACCINATION**

**SUBMITTED BY: -**

**Gulivindala Naveen Kumar (C22008)**

**Harendra Sai Nath Lella (C22009)**

**Kushagra Mohan Prasad (C22014)**



**Celebrate Your Worth**  
<http://www.praxis.ac.in>

**Objective: -**

Exploring and analyzing the COVID-19 data in terms of the number of cases and deaths and the vaccinations that are available in the world and in India and analyzing the correlation between the vaccination and the COVID-19 intensity.

**Abstract**

This project explores the data on the number of people getting vaccinated all around the world and in India. Specific to India, we will take a closer look at the vaccinations data from different perspectives like age, gender, place, vaccines types, etc and will try to draw some conclusions like, the trends that are happening in the vaccination drives in the maximum possible ways and the best possible strategies and suggestions that can be followed during this kind of pandemics.

The motivation behind the project is the COVID itself, as it is shaking the world with its new variants and the effects that it is causing to the world. The data that is available on the internet is so huge that we could make a lot of deductions from it and thereby make some strategies using them. Specific to the data on vaccinations, there is a lot that we can deduce in terms of human activity and also the economy generation all around the globe.

## **Introduction**

### **CORONAVIRUS (COVID-19) (What, Where, Why, When and How ?)**

**What** - Coronavirus (COVID-19) is a novel infectious disease caused by a new variant of virus called SARS-CoV-2 virus. They are called “corona” because of crown-like spikes on the surface of the virus. Severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS) and the common cold are the examples of the viruses that cause illness in humans.

SARS-CoV-2, the virus that causes COVID-19, enters your body through your mouth, nose or eyes (directly from the airborne droplets or from the transfer of the virus from your hands to your face). It then travels to the back of your nasal passages and mucous membrane in the back of your throat. It attaches to cells there, begins to multiply and moves into lung tissue. From there, the virus can spread to other body tissues.

#### **How does the virus spread?**

The two possible ways of getting affected by the virus are,

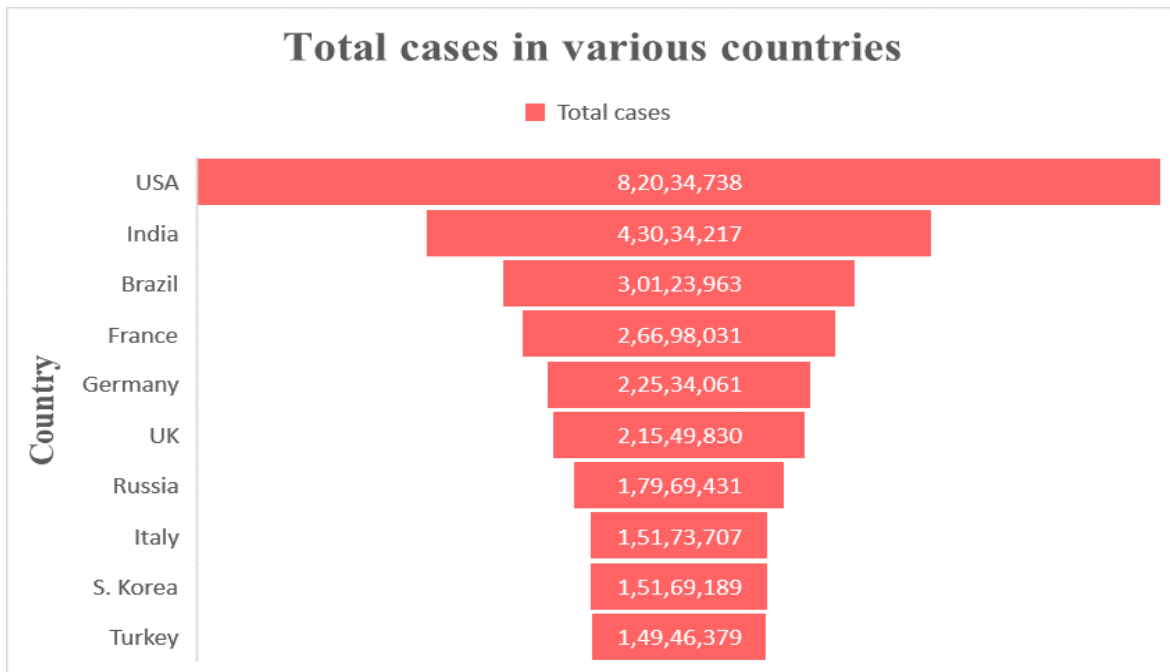
One, the virus travels in respiratory droplets released into the air when an infected person coughs, sneezes, talks, sings or breathes near you. You may be infected if you inhale those droplets.

Two, through the close contact (touching, shaking hands) with an infected person.

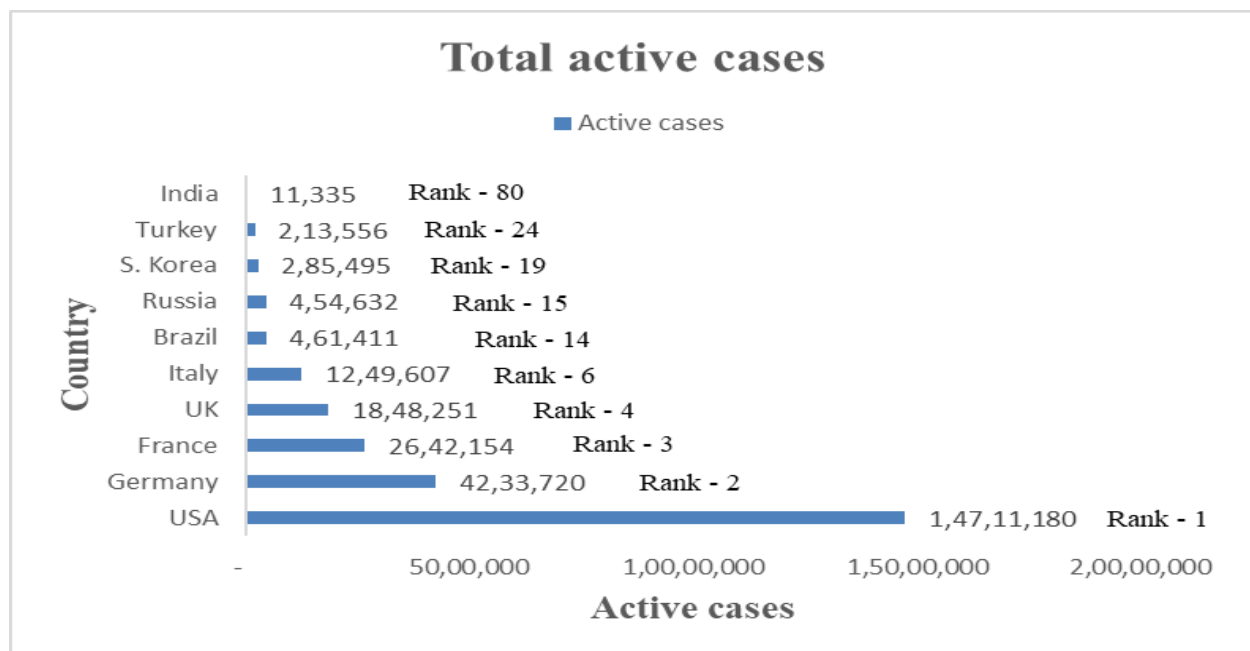
**Origin** - The COVID-19 outburst was seen in Wuhan, China in November 2019. But the first COVID case reported in India was on January 27th, 2020. It was the day when no one (including doctors) had any clue about the virus and the way to treat it, or the vaccines that can be given to stop the rise of the cases. But, today we are equipped with clinically proven vaccines along with medication for the treatment, where we could safeguard ourselves from the novel virus.

## COVID-19 - The Big Picture

Now, we will try to look at the covid cases all over the world with the help of charts and try to make some deductions from it.



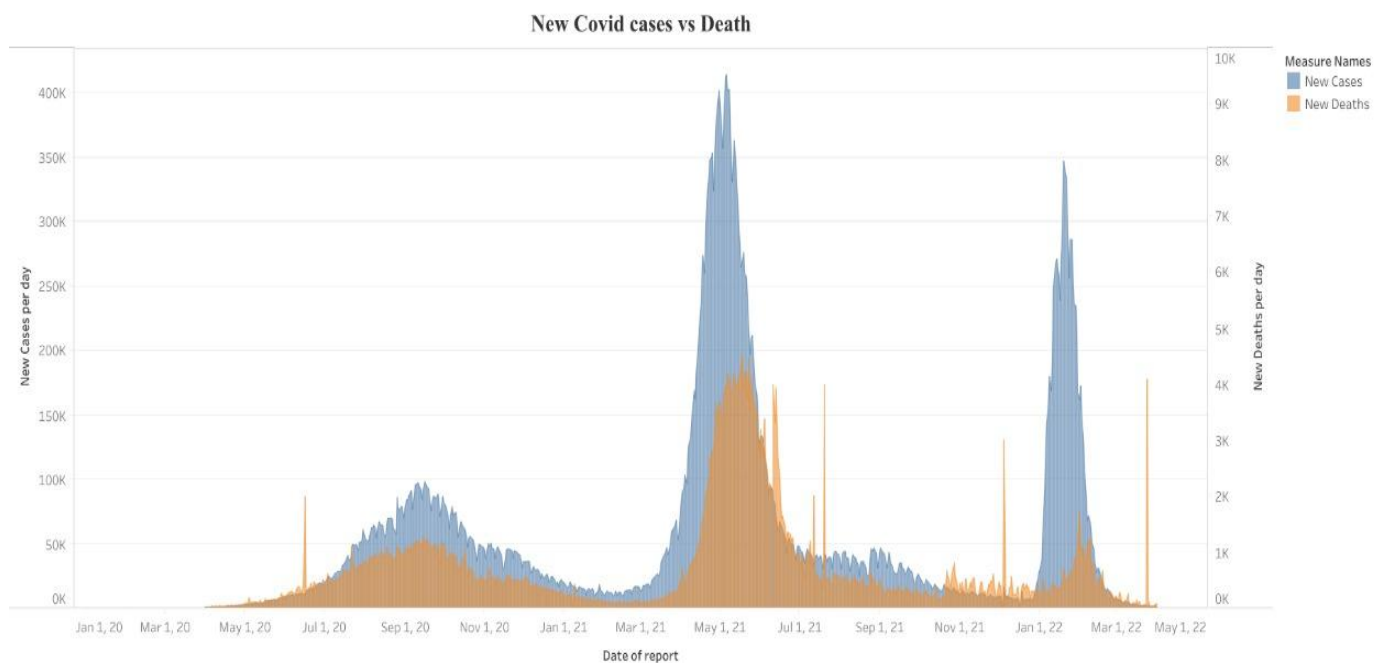
The above Funnel chart represents the top 10 countries with the highest number of cases. We see that India stands in the 2nd position in terms of the number of cases.



From the above chart we see that India ranks 80 in terms of the active cases. When we compare with the total cases, where India stands at 2nd position, the rank in the active cases has dropped down, which means that the vaccinations that are happening in India are making the path for letting the virus out of the country.

## COVID in India

The Novel virus has spread in all parts of the world within a very short period of time. Similarly, we have found the first case on Jan 27th, 2020 and we could observe the rise in the number of cases from June-October, 2020. There after the first wave, we have seen the rise in the second wave during Feb-July, 2021. Where we have seen a steep rise in the number. Coming to the third wave during Jan-Feb, 2022, we again got to see the maximum rise in the number of cases.



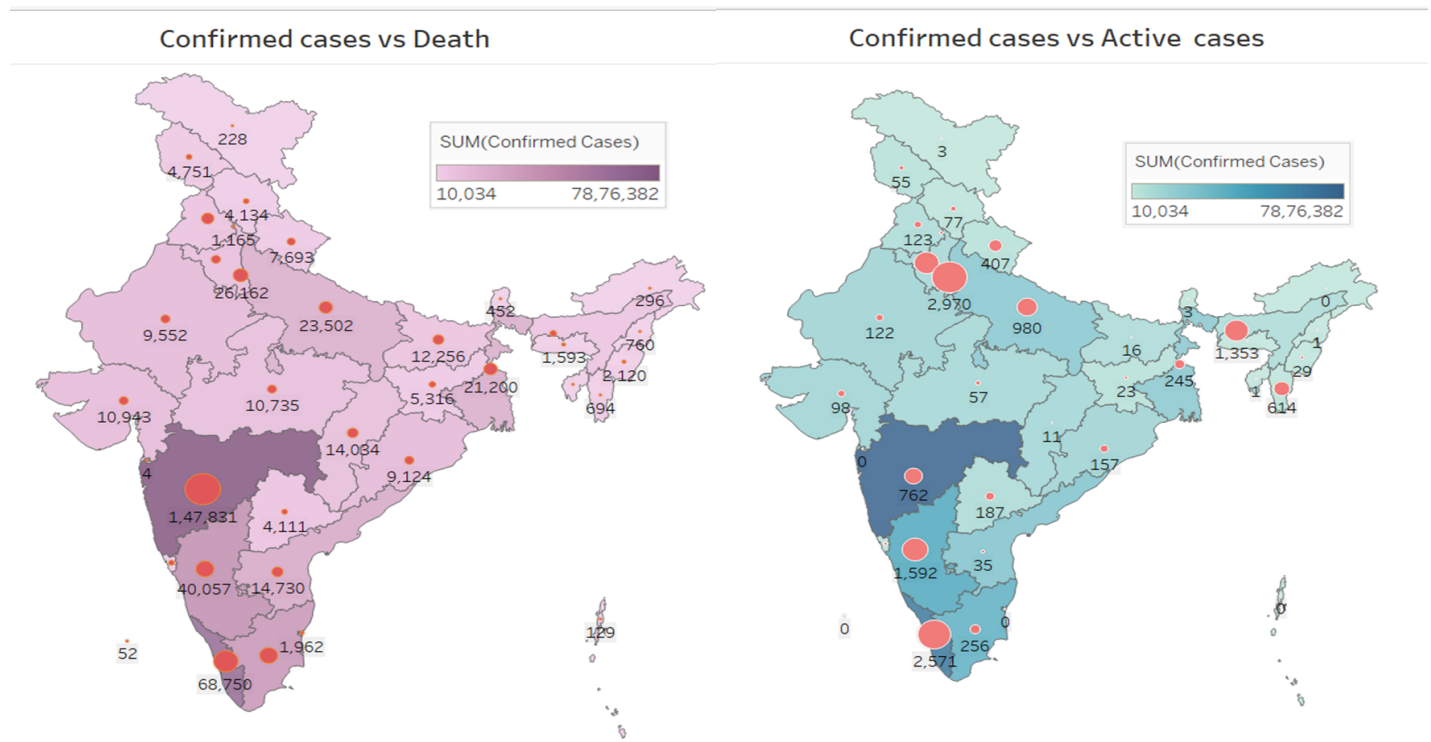
Speaking specifically about our country, from the graph on new cases and the deaths, we can observe that there is a steep decrease in the number of cases from July 2021. This was possible because of the precautionary measures taken by the government as well as the public in safeguarding themselves, by wearing masks and maintaining social distance. Later again in Jan 2022, we could see the spike up high because of the new variant and some negligence of the public in maintaining the precautionary

measures. But, from March 2022, we could again see the downfall in the no of cases. This was truly because of the vaccination drives that have happened all over the country very rigorously. Today, we are able to see a very low count in the number of cases because people are getting vaccinated even with the booster dose as well. This booster dose might take a further long step in removing the strains of the virus from our country.

The next thing we are supposed to talk about here is the deaths that are happening in India. From the above graph we could see that the no of deaths occurring have been reduced substantially from May-June 2021. This was possible just because of the availability of the medical treatment along with the vaccination drives that took place.

People who are vaccinated with at least one dose were also affected in the mid 2021. But, due to the vaccinations, they got the viral symptoms very mildly and have prevented them from causing death. Also, these vaccinations have helped them build immunity to fight against the virus if again got affected by.

Here, it is a clear indication that the vaccination drives are happening at a quicker pace and all the people are getting vaccinated and trying to let the virus out of our country. As we have seen the importance of the vaccines in eliminating the virus from humans and from the country on a whole, let us dig deeper into the vaccines and try to look at the data from different aspects.



Here, we could see the maps of India representing the confirmed cases vs Active Cases and confirmed cases vs Deaths. This map shows the filled map of the attributes. According to this, we can say that the maximum number of cases were reported in Maharashtra followed by Kerala and so on. Speaking about the active cases, we can observe that the maximum number of cases were being reported from Delhi, next to it is Kerala and so on. If we look at the number of deaths that took place during these series of pandemic waves, it is clear that maximum deaths happened in Maharashtra and contradictorily, North Eastern states and the Jammu and Kashmir have reported the lowest.

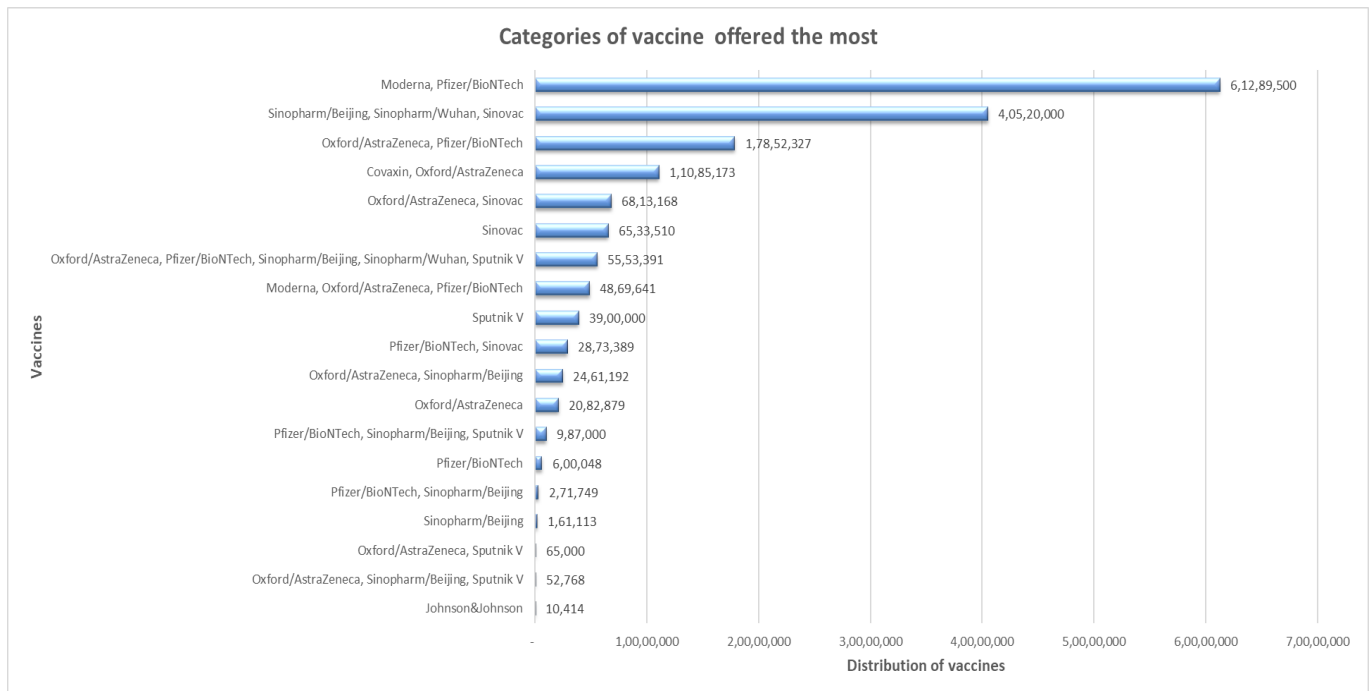
As we have seen the scenario of the COVID from the perspective of the number of cases and number of deaths, we will now try to look at the vaccines available and the vaccination drives that are happening and the positive outcomes from getting vaccinated.

## The COVID-19 Vaccines

Vaccines can save the lives of the humans from the respective viruses. Vaccines work by training and preparing the body's natural defense system - the immune system to recognize and fight off the

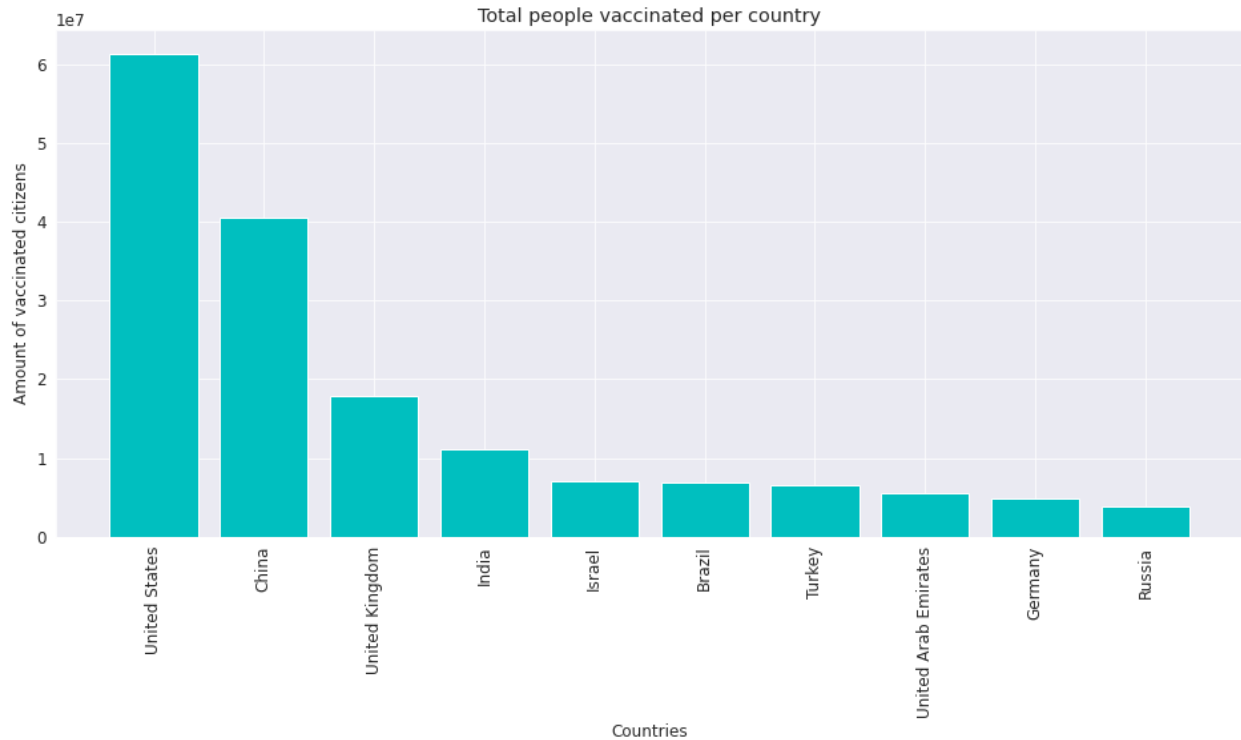
viruses and bacteria they target. Even if the body gets attacked by the virus, the body immediately destroys the virus and prevents the illness.

On a similar note, the vaccinations for fighting the COVID have been developed by various organizations of various countries. Currently, there are a number of vaccine types that are available all around the globe.

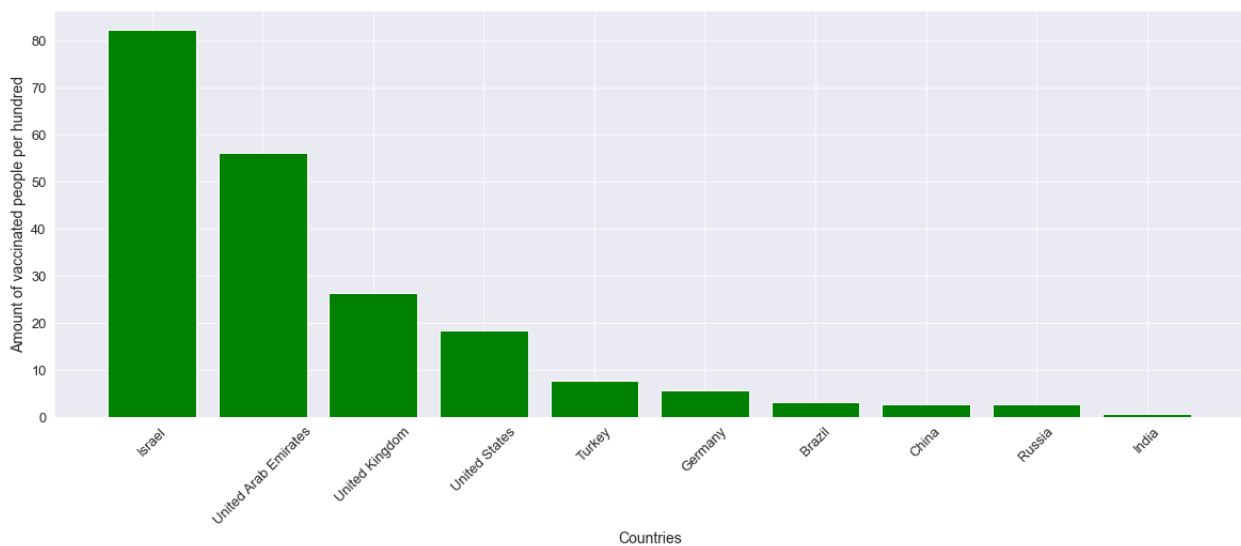


The above chart represents the different types of vaccines used all around the world. Pfizer tops the list of most-used COVID-19 vaccines in the world, followed by Moderna, and Oxford–AstraZeneca.





If we look at the countries that have been vaccinated, the USA stands first in the list, whereas, India stands at the 4th position.



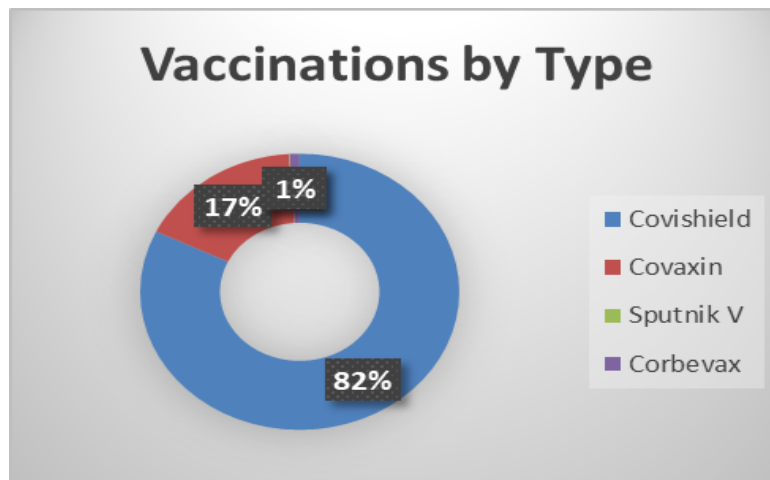
It is clear that Israel leads the world in terms of the number of doses per head of population, with more than 80 doses given for every 100 people.

## Vaccinations in India

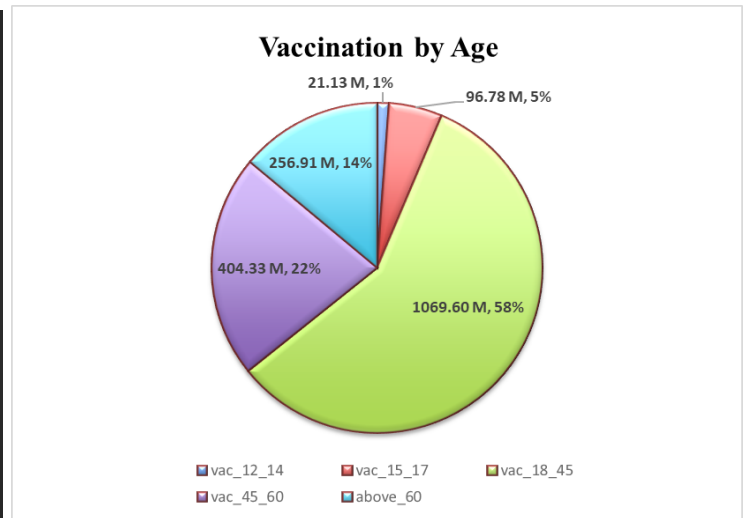
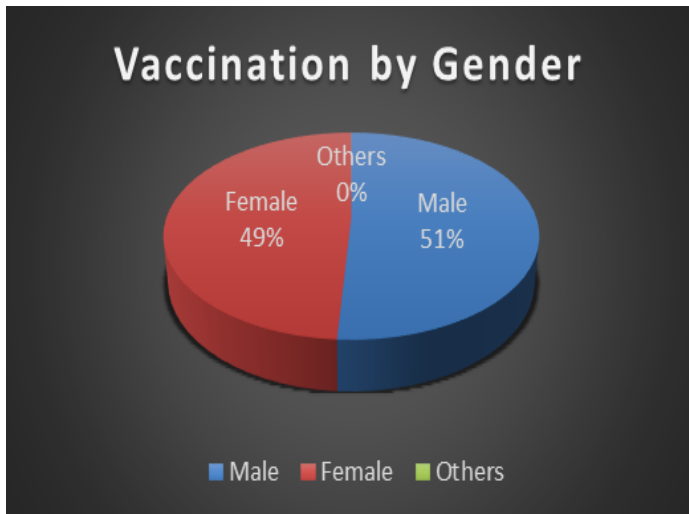
There are several COVID-19 vaccines validated for use by WHO. But, considering only India, there are four types of vaccines that are being used.

Vaccines by type			
Covishield	Covaxin	Sputnik V	Corbevax
1,51,37,49,429	30,81,86,515	12,21,722	1,85,82,502

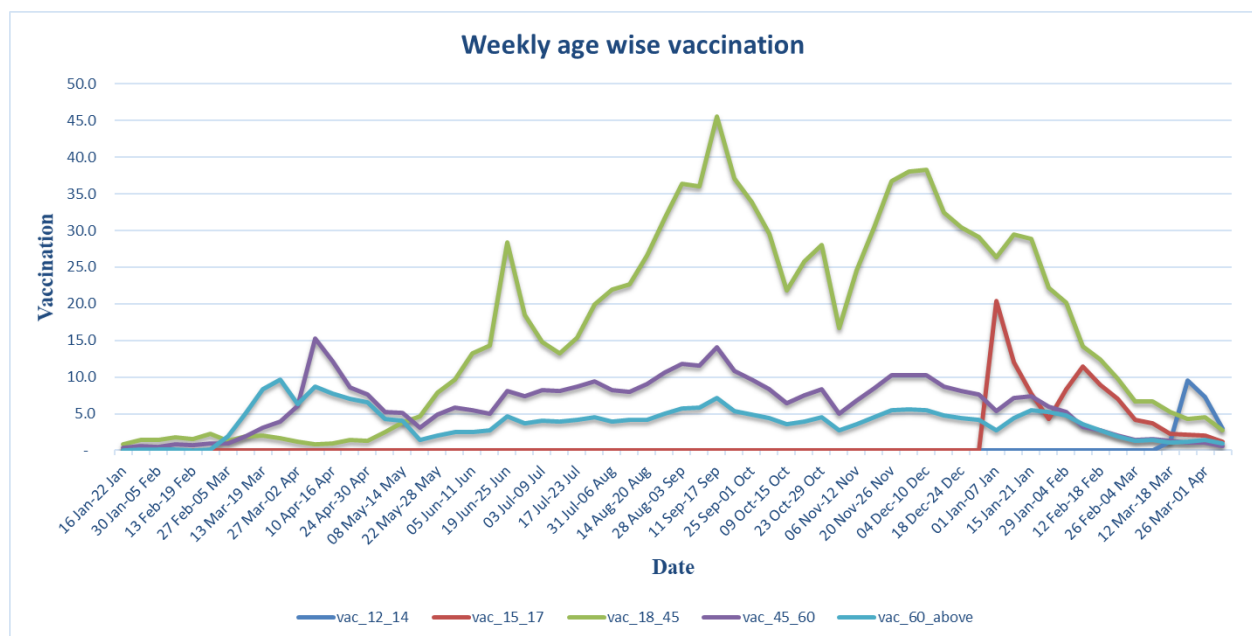
From the data it is evident that Covishield is mostly used for vaccines in India. Though the other vaccines such as Covaxin, Sputnik and Corbevax were also available, Covishield is mostly distributed all around the country. One such reason could be the acceptability of the vaccine overseas.



Now, let us dig deeper into the vaccinations. As we have seen the different types of vaccinations that are being used in our country, let us look at the vaccinations from the gender and the age perspectives.



The above two pie charts show the data of the people getting vaccinated with respect to their age and gender. If we closely look into the data, we can see that both the male and the female are vaccinated in proportion with each other. There is no gender bias in terms of people getting vaccinated, which is really a good sign that equality is being maintained all over the country.

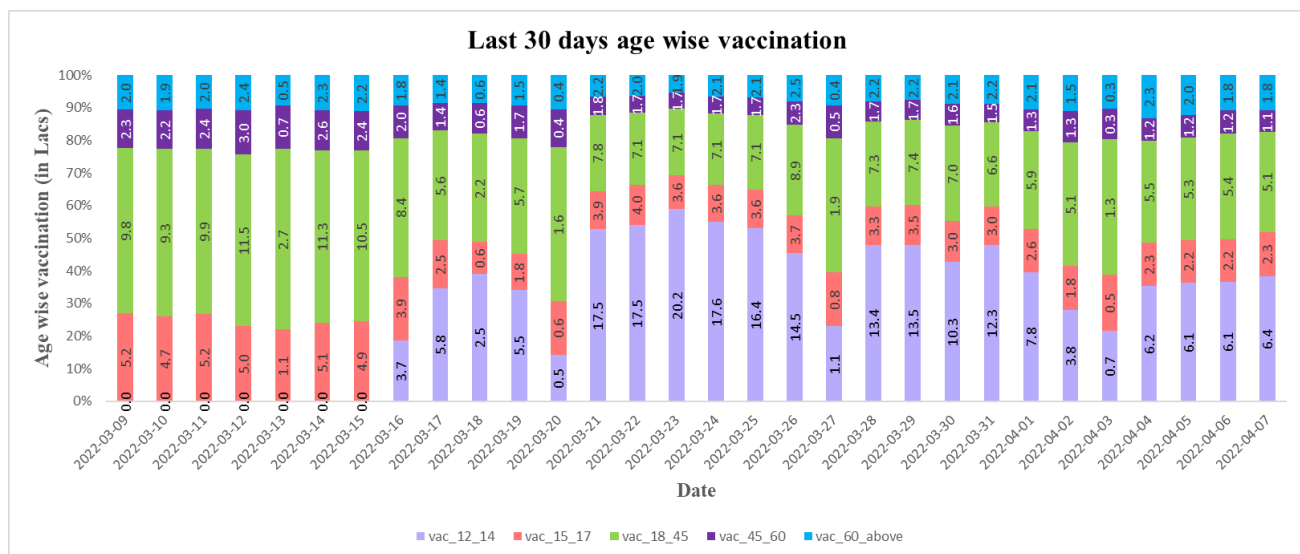


If we look at the above table, which tells us that the vaccination drives have been conducted in a systematic way, where each set of age groups is being targeted at every level. Like, the first spike tells that the people above 60 years old were vaccinated because they were very susceptible to the disease and also require the proper immunity levels to fight the virus if affected by. Next comes the age group

of 45-60 years, where these sets of people stand second in the requirement of the proper immunity levels.

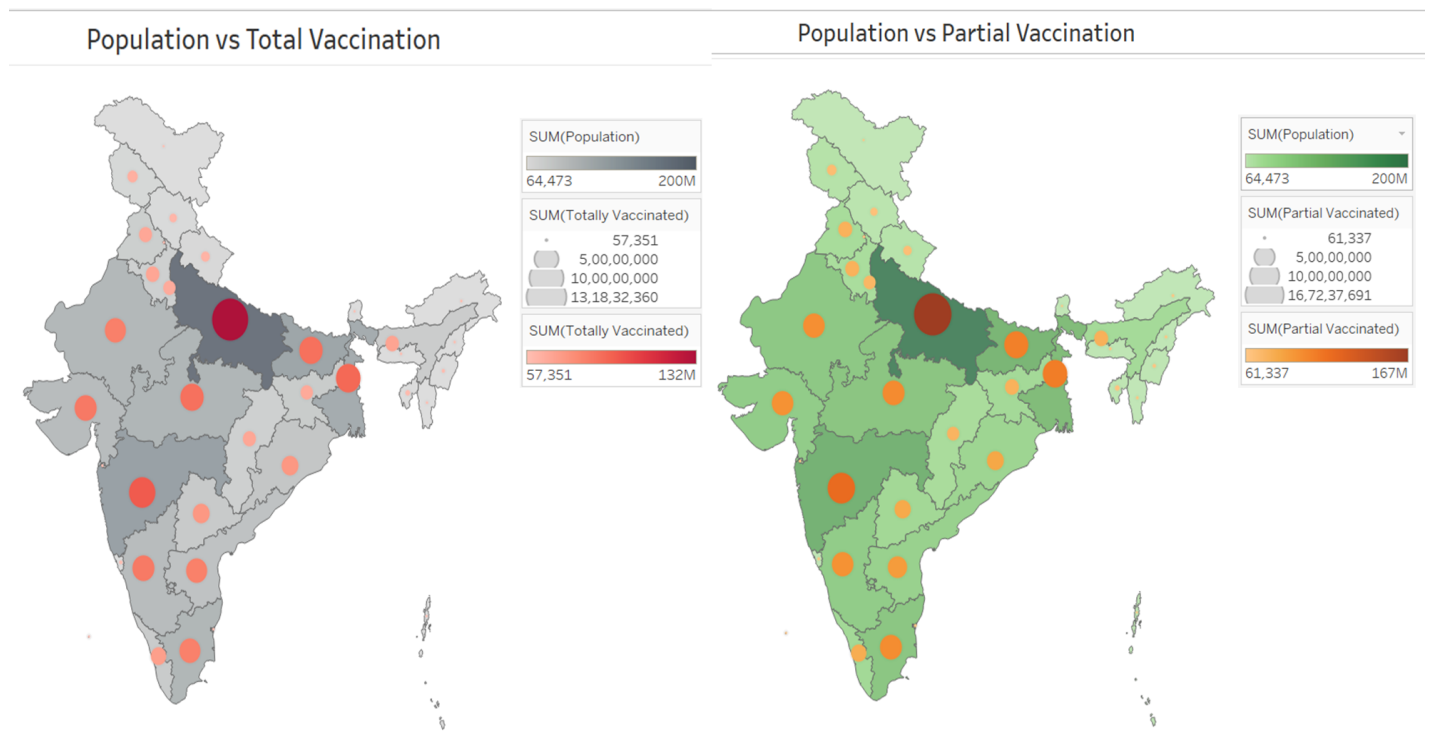
In the next level we can observe that the people in the age group of 18-45 are being vaccinated the most and also we can observe that this age group got their vaccinations right after the second wave. Which means that the young generation of the country have been warned by the negative repercussions of the COVID and so they all have been vaccinated. The other reason is that there is a huge population in this age band. So, huge vaccination drives have been done in this period.

Now, in the recent past, i.e, since March 2022, the children who come under the age group of 12-14 have been vaccinated. The reason could be that, the number of the cases have fallen down to a large extent and these people, who are basically the school goers that require a safety precaution have got their first dose of vaccination.



If we take a closer view at the above stacked bar chart, we see that the vaccination that happened in the last 30 days were more towards the people in the age group of 12-14 and 18-45. This stacked bar chart justifies our statement made above.

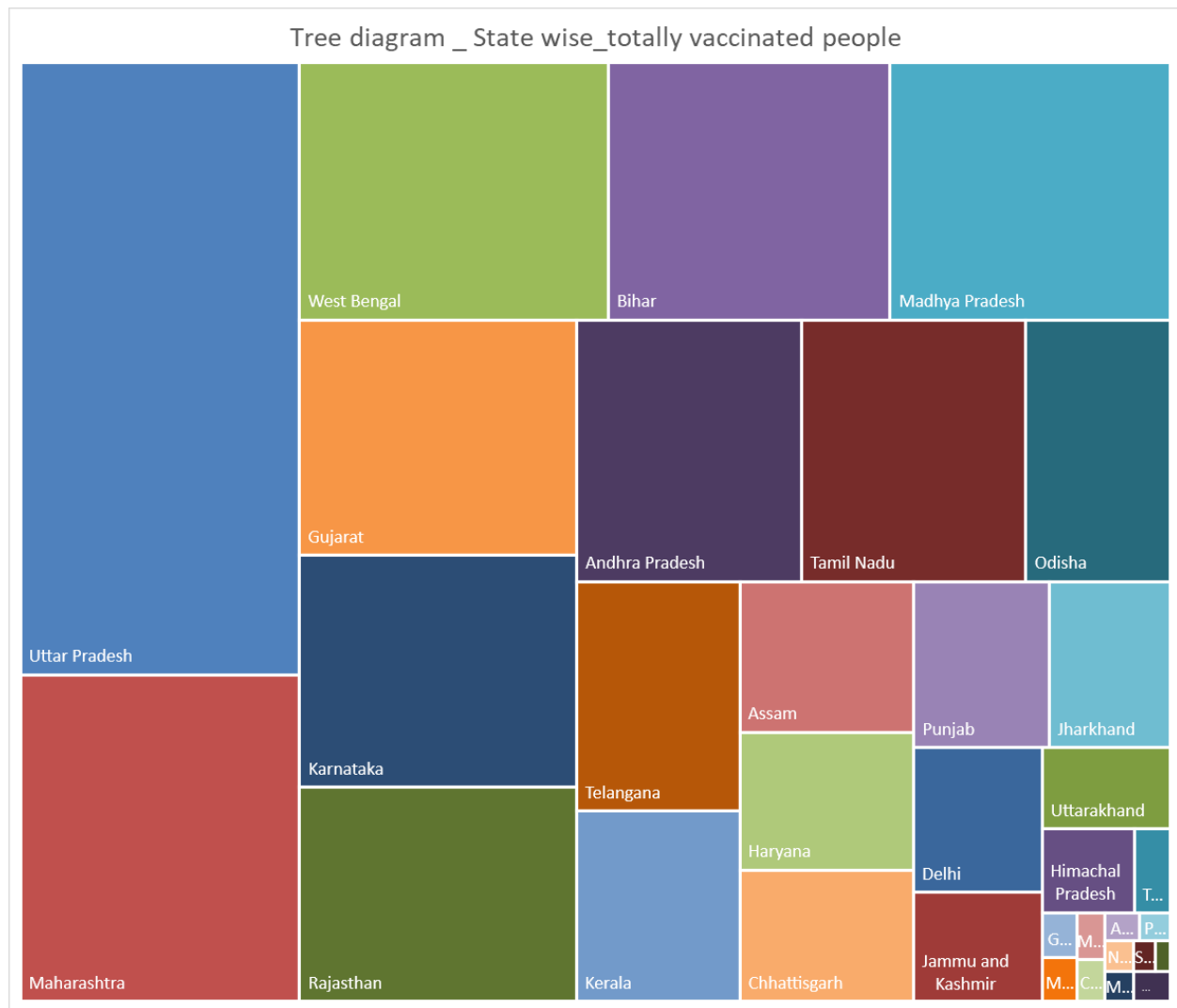
Till now we have gone through the vaccination data with respect to age, gender & type. Now, let us look into the data state wise.



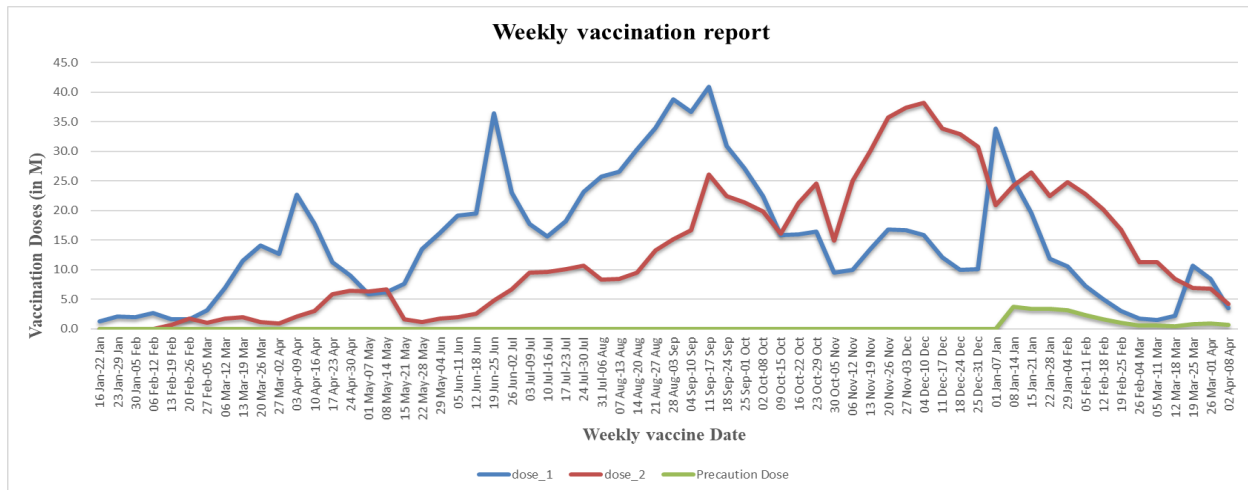
From the above filled maps, we can look at the data from the perspective of the vaccinations that have happened till date. The above map represents the data of the Population vs People who are partially vaccinated and also population vs people who are fully vaccinated.

From the above maps we can infer that, the most populous state is Uttar Pradesh and the percentage of vaccination, may be partially or fully vaccinated, were high followed by Maharashtra, Bihar, West Bengal, and so on.

It is good to see that the people are getting vaccinated in tune with the total population of the state. It shows that the vaccination awareness campaigns that took place had finally paid off. It is very strategic to see that the place with more population is more prone to get affected the most and the vaccination has happened in those places on par with the requirement.



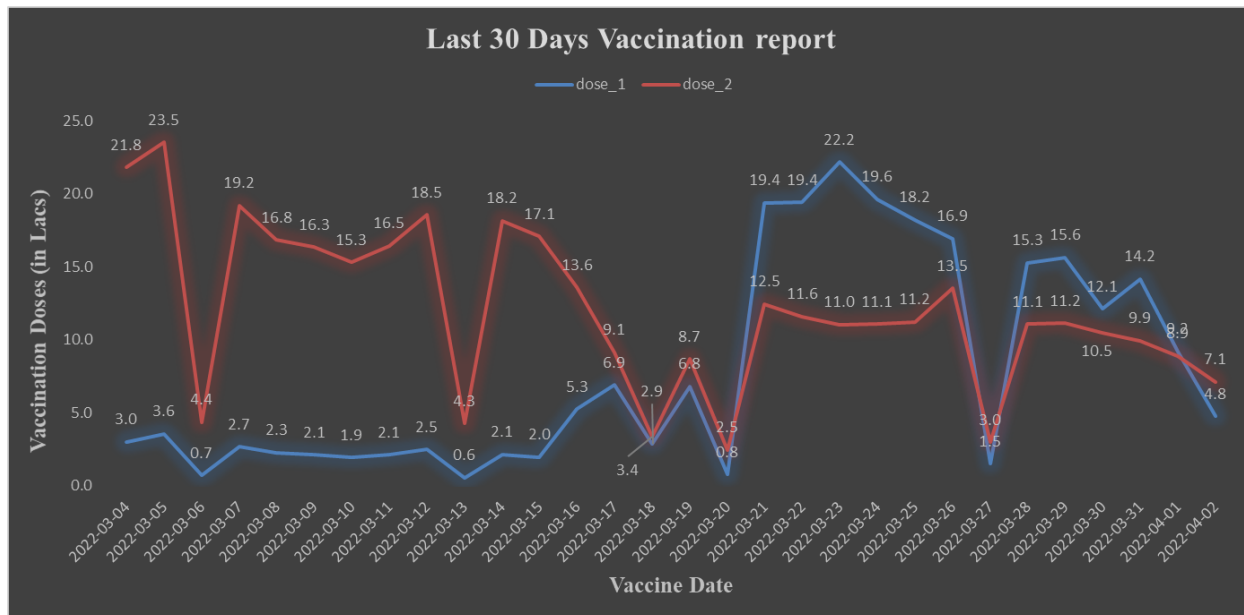
From the above Tree map, we can see that the majority of the vaccinations are happening in Uttar Pradesh followed by Maharashtra and West Bengal. We can infer that the population in the top three states is huge and that itself makes the difference, i.e, more the population, more prone to get the virus and so vaccination is required for sure.



The on-ground reality is that, though we are being vaccinated against the Coronavirus, we are being attacked by the virus again and again and the virus is changing its facets for every wave. This is because of the reason like, the economy was shut down during the COVID lockdown period. So, post the lockdown, the people got to their routine work and the cases which were supposed to get suppressed, surprisingly hit the strike again. So, looking at the series of events, the government enlightened the people about the importance of the vaccination and more specifically, the second dose of the vaccination. The above graph shows the spike in dose 2 after the second wave.

The booster dose also plays a major role in getting immunized over the coronavirus. The main purpose of the booster dose is to keep the human body immuned with the required antibodies, so that, if there is any attack after getting the two vaccinations being over, this booster dose would act as a real booster for the previously taken vaccine.

Post the third wave, the majority of the population are completely vaccinated, if not completely, at least by single dose and this is where the booster dose came into the picture. In general, any vaccine works for a specific time period, post that the efficiency of the vaccine decreases. Keeping this in view, the WHO and the governments have suggested getting the booster dose after 9 months of taking the second dose. In this way, we could protect ourselves from the virus in the future.



Atlast, here comes the report on vaccinations that took place in the last 30 days. Here, we could see a major hike in the second dose, where we have mentioned the reason above. From the past 15 days, it could be viewed that there is a rise in the first dose as well. The reason behind it is the vaccinations for the children below 18 years of age.

The above graph clearly shows that there is huge difference in the vaccinations, which can be said as an exponential increase right from its starting time. The vaccination is one of the major reasons behind reduction in the number of the covid cases being reported from the past 6 months. Strictly considering the data of the last 30 days, it could be said that there is a steep decrease in the registration of new cases and exponential increase in the number of vaccination registrations.



## **Conclusion**

With the help of insights generated from the report, we could conclude that Exploratory Data Analytics played a major role in deducing the insights from the available data. With the help of EDA, we were able to look at the data from different perspectives and use them in building a more robust system for vaccination in India.

In this case study, we have generated graphs on COVID cases along with the deaths that have happened since the outbreak of the virus. Later, we have generated graphs on vaccinations used in all around the globe and also on the countries and their vaccination progress. Later, we focused completely on the data of our country, where we have seen from the perspectives of age, gender, type and also percentage of vaccinations in each and every state.

If we dig deeper into the insights that we have arrived at, it is evident that the vaccinations, whatever the company it is, have played a key role in getting people immunized and making them antibodies ready to fight against the virus if affected.

India stood 2nd in terms of the total cases reported and later stood at the 80th position with respect to the active cases. Which means that the vaccinations that are happening in India are making a path for letting the virus out of the country. We could also notice that there was a substantial decrease in the number of cases that are being reported on a daily basis and this was quite evident from the data and the representations made above, saying that vaccinations are very key in decreasing the spread of the novel virus.

## **Suggestions or Recommendations to control the spread of the virus**

The best possible way to control the spread of the virus is through taking precautions and maintaining few guidelines.

The first thing would be maintaining social distance combined with wearing face masks.

Second comes the vaccinations, the major player in this pandemic. Now, the caution dose should also be mandated for each and every individual for their own safety.

## References

1. <https://dashboard.cowin.gov.in/>
2. <https://ourworldindata.org/covid-vaccinations?country=~IND>
3. <https://ourworldindata.org/covid-vaccinations>
4. <https://www.analyticsvidhya.com/blog/2021/05/covid-19-vaccination-data-analysis-visualization/>
5. <https://portal.who.int/report/eios-covid19-counts/#display=Countries and Territories&nrow=2&ncol=3&arr=row&pg=1&labels=&sort=cur case who;desc&filter=&sidebar=4&fv=>
6. <https://worldhealthorg.shinyapps.io/covid/>