**REPORT**

**ON CORONAVIRUS CASES**

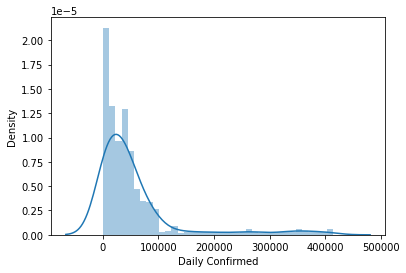
**IN INDIA**

**By**

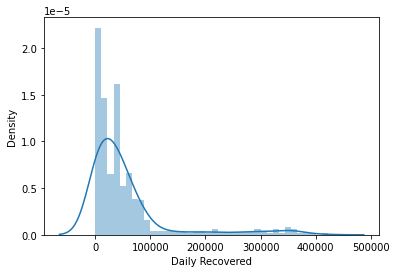
**Ayush Pandey**

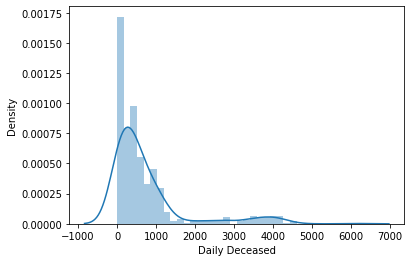
The first symptomatic infected individuals of coronavirus (Covid-19) was confirmed in December 2020 in the city of Wuhan, China. In India, the first reported case of Covid-19 was confirmed on 30th January 2020.

I am having a csv file of corona cases in India from 30-January-2020 to 28-September-2021 and from this csv file I see that during the month of February, the number of cases reported was two and remained constant during the entire month and from the start of March 2020 till 5th June 2020 there were total of 2,36,189 confirmed corona cases in India.During this period the daily cases discovered in India were less than 10,000 (In march it was less than 400) and after that the daily confirmed cases increased from July 2020 and this period(July to November) is also known as the “First Wave In India” where the daily corona cases were more than 50,000 and the maximum cases were found in the month of September which was close to 1,00,000 and after that we see that there was a drop in the corona cases per day, when I researched for the cause of the drop of the corona cases I found that the cases droped as the government imposed lockdown in the country as a result there was no much physical contact of people as a result the cases droped and during the start of the year 2021, the cases reported per day were less than 15,000 but again the daily cases rose from February end where India suffered the “Second Wave Of Covid” due to which the cases rose again and we see that the increase of cases was much large than the first wave and daily more than 1,00,000 cases were confirmed, the maximum daily cases came in April which was 4,14,280.



From the Daily Confirmed column,my mean is 55452.01974, my median is 34736.5, my skewness is 2.803722989 and from the graph which I have plotted I see that the Daily Confirmed cases is Right Skewed also known as positively skewed as most data falls to the right, or positive side,of the graph's peak(The density of the daily confirmed cases is forming a peak for the range(0,10000)).





Similarly the Daily Recovered cases and the Daily Deceased cases is also having a Right Skewed (Skewness for Daily Recovered is 2.704460986, the mean is 54240.85362, the median is 34567.5 and for the Daily Deceased the skewness is 2.396052631, the mean is 735.5115132 and the median is 429) as daily there were people who got deceased and recovered from the coronavirus and during the first wave and the second wave when the cases rose the amount of people who recovered from the virus also increased And the maximum of Daily Recovered cases was seen in the month of May 2021 where 4,22,391 people recovered from the coronavirus.

The maximum Daily Deceased cases was seen in the month of May where 6,139 people deceased from the coronavirus.

And same goes for the Total Confirmed,Total Recovered and Total

Deceased columns as they are the total of the cases confirmed,

recovered and deceased people in India and they are also having a positive skewness.

The Total Confirmed cases found in India from 30-January-2020 till 28-Septembr-2021 is 3,37,14,828 and the Total Recovered cases is 3,29,78,439. This shows the recovery rate in India is 97.81%, and the Total Deceased cases in India is 4,47,191. This shows that the deceased rate in India is 1.32%.

(The mean which I obtained for the Total Confirmed cases is 11620820.28, the median is 9412481, the skewness is 0.768645198,

The mean for Total Recovered is 10887584.27, the median is 8823570 and the skewness is 0.831826244, the mean for Total Deceased is 155795.3076, the median is 136364 and the skewness is 0.789147895)

In this report, I have provided a simple statistical analysis of the novel Coronavirus (COVID-19) outbreak in India. And I would like to conclude that all the columns present are Right Skewed and for infering useful information from the data (for making predictions in future) skewness should be removed so that each and every column can have a normal distribution which can be achieved by taking the square root or a cube root(ways of removing the skewness) of the data's present in the columns.