**DSC 530 Final Project**

**Covid India Vaccination Analysis**

**The outcome of EDA:**

As per the current situation, the second wave of covid spreading in the Asia region especially in India is rapid, and it's spreading at a massive rate. So, I would like to analyze how the vaccination drive is going on there using the data available in the world meter. After the analysis, I found that the data provided is correct and there is no outlier in the data, and I found that the reason for rapid spreading is only due to fewer people got vaccinated in the entire Indian population.

**Missing During Analysis:**

Due to the very high dimensionality of the data, there are a lot of missing values and to do the analysis correctly, I have been forced to delete some of those rows. So, some of the data of that particular date are not included. However, I don’t feel that’s a mismatch for the analysis because the positive case count has been increased constantly, and skipping some days in the middle isn’t a problem in this analysis.

**Variables could have helped in the Analysis:**

Originally, I intended to attempt to define the difference between the number of vaccinated people based on the different types of vaccine. But due to the deletions of certain rows due to missing data, we couldn’t achieve that analysis correctly, so which made to skip that analysis.

**Possible Incorrect assumptions:**

Since all the data are from the world-meter and this data is very sensitive due to the pandemic, I couldn’t find any incorrect assumptions made in the analysis.

**Challenges:**

I knew that this project going to take enormous energy and amount of time, because of my curiosity to know the facts of vaccination going on in India, I learned a lot and after doing this analysis I came to know that there are many more I need to learn.

Especially, I think, I need to focus more on how to conduct the test and how to perform regression analysis.