Reach of the Pandemic

ANALYSIS ON CORONA VIRUS

R.H.S.BHARADWAJA

PROBLEM STATEMENT

- 2020 has been a nightmare to the entire due to an epidemic, Corona. There are hundreds of causalities daily due this virus and the number of positive cases are increasing as we speak.
- The main agenda is to find which are all countries that are handling this epidemic efficiently and visualizing all the details of each country on a map for better understanding of what is happening around the world regarding the virus.
- There will an analysis on country level for this project.

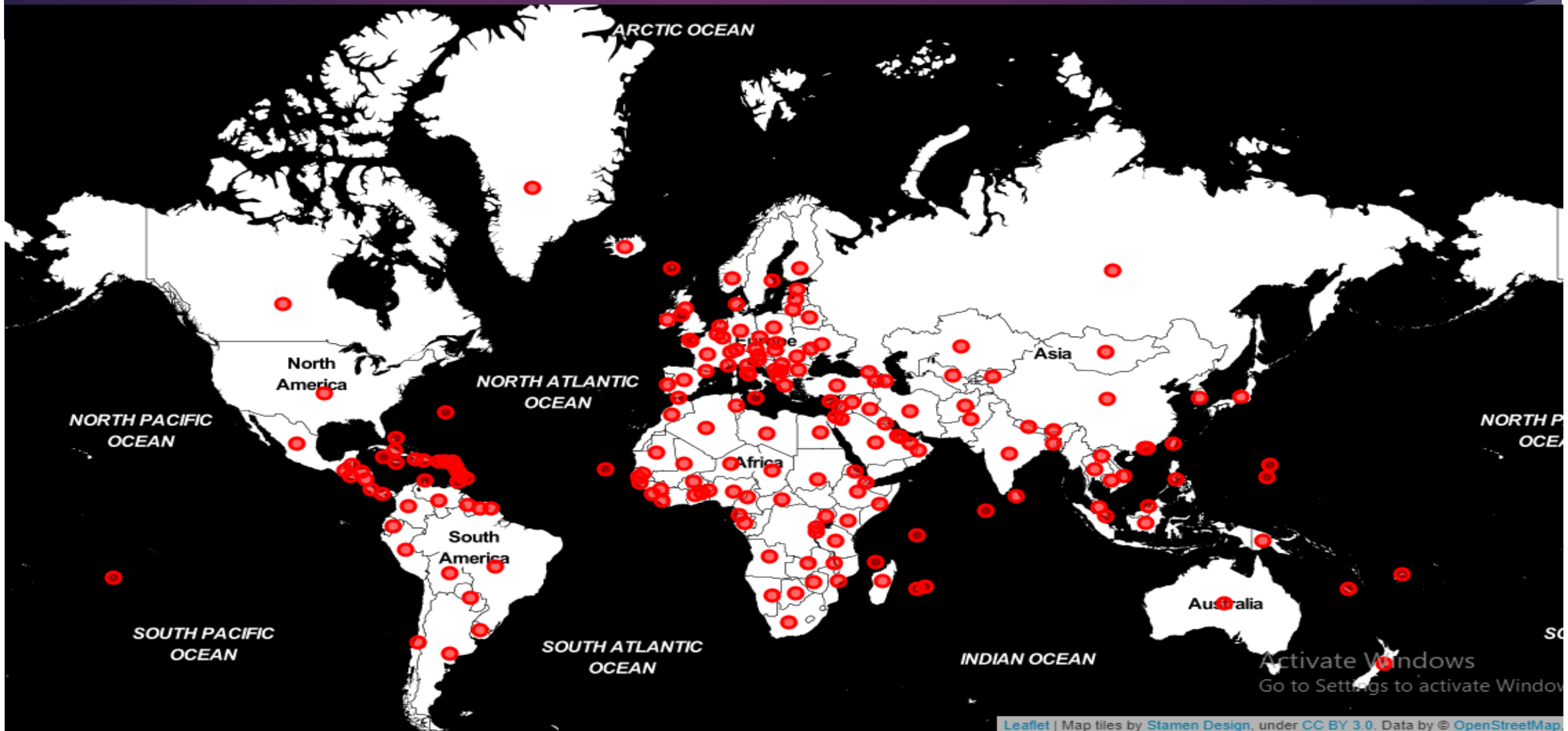
DATA ACQUISITION

- We actually in search of 2 sources. One, information of each country's statistics on corona and locations details of each country which is helpful for visualization purpose.
- Information regarding Confirmed cases, Recovered etc., are available in Wikipedia for us to use.
- Location of countries are available as a csv file provided by google under the name "Countries.csv" which contains latitudes and longitudes every country.

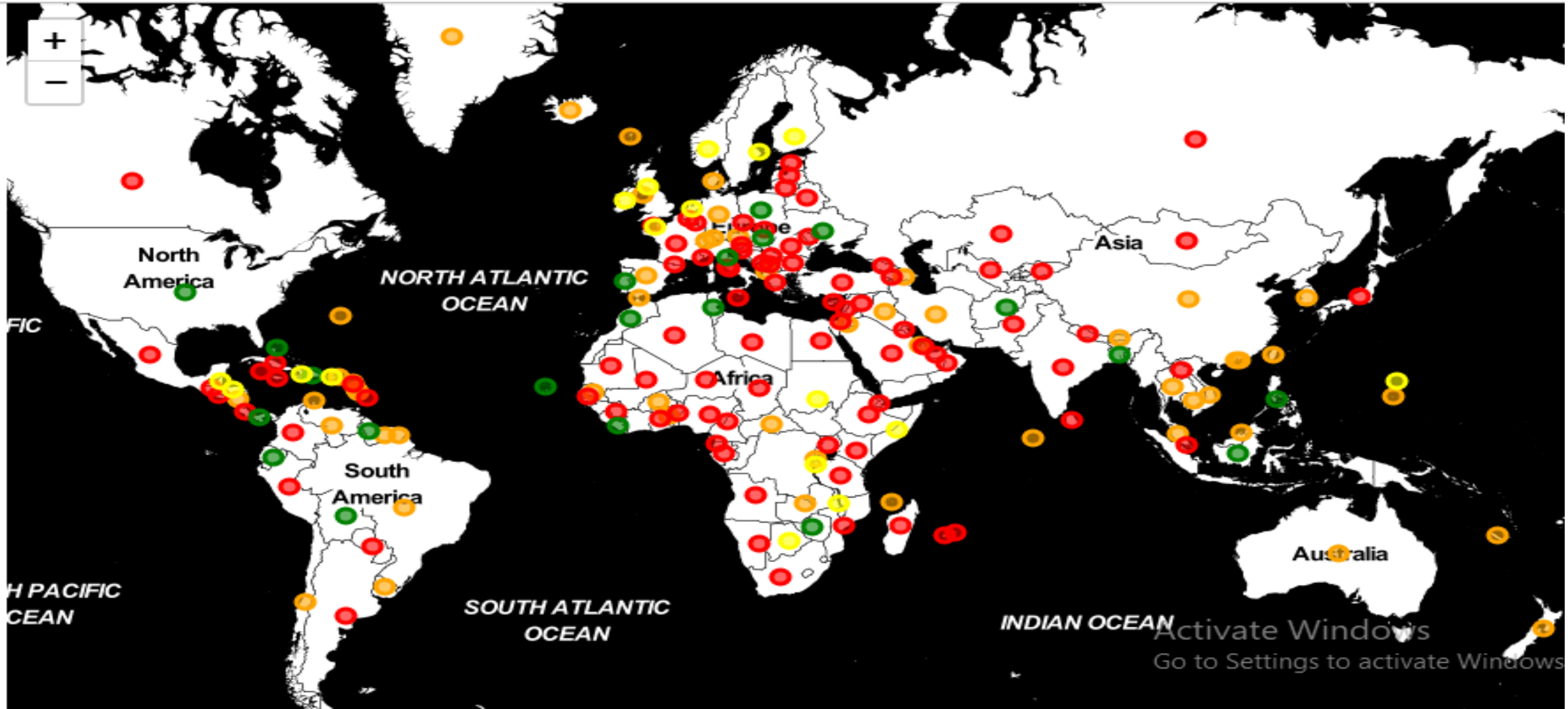
DATA SCIENCE METHODOLOGY

- It was extracted using BeautifulSoup and created a Data frame for analysis purpose.
- Data Cleaning was done to remove all Nan (Not a number) and characters like – or _ are replaced with 0.
- Joining of cleaned data frame is done with countries data frame that contained location details.
- Addition new columns like “Recovery Rate” , “Pending percentage”, “Live Cases” is done for better understanding purpose.
- Recovery Rate = (Recovered Count *100) / (Recovered count + Death count).
- Pending percentage = (Live cases * 100) / Confirmed cases.
- Used K means clustering algorithm to cluster countries into 4.

Before Clustering



After Clustering



CONCLUSION

- As we can see the countries with less pending percentage and high Recovery Rate are clustered into one zone and they are very few countries this cluster, namely Greenland, South Korea etc.,
- Observing Map over a period of time (Like a daily or weekly report) can help us with more insights like, which countries are moving between clusters and according we will get to know whether severity is increased or decreased.
- Many more observations can be made if we continue his project with each state's data.