COVID-19 cases on world map using python

1. **Introduction**

SARS Corona virus COVID-19 is a global pandemic now. There has been an immense interest in analyzing the spread of this virus among researchers and data scientists. The data is updated on daily basis via Nemours data collection resources and is available to download for free.

It is a proven fact that analyzing the data with predictive modeling and machine learning algorithms helps reduce the spread, develop mitigation plans, prepare for future scenario and make effective decisions among multiple stakeholders. In current study, I will use jupyter notebook 6.0.3 to visualize COVID-19 data on world map. This is one of the best way to visualize the spread of the virus.

1. **Description of data set**

Data was downloaded from [Our World in Data](https://ourworldindata.org/). This website provides datasets in csv, xlsx and json formats. I used xlsx format but all three formats can easily be imported to the python data frame for further analysis.

1. **Data Science Methodology**

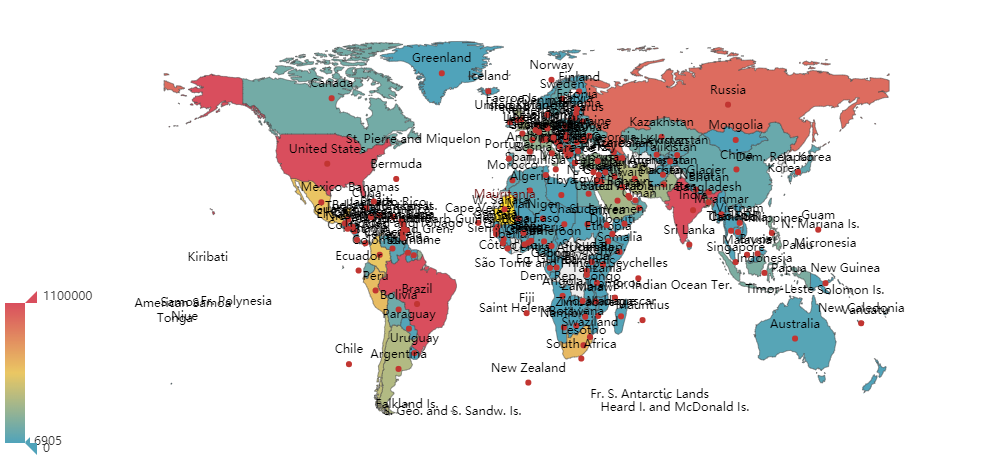
Python Pyecharts library was used to create world map. The data in tabular format was uploaded on pandas data frame. The size of the table was 210 rows and 40 columns. Data set was sorted in ascending dates format and a data was chosen for the report was prepared (Aug 25, 2020). Since I was only interested in total number of cases so rest of the data was ignored.

1. **Results**

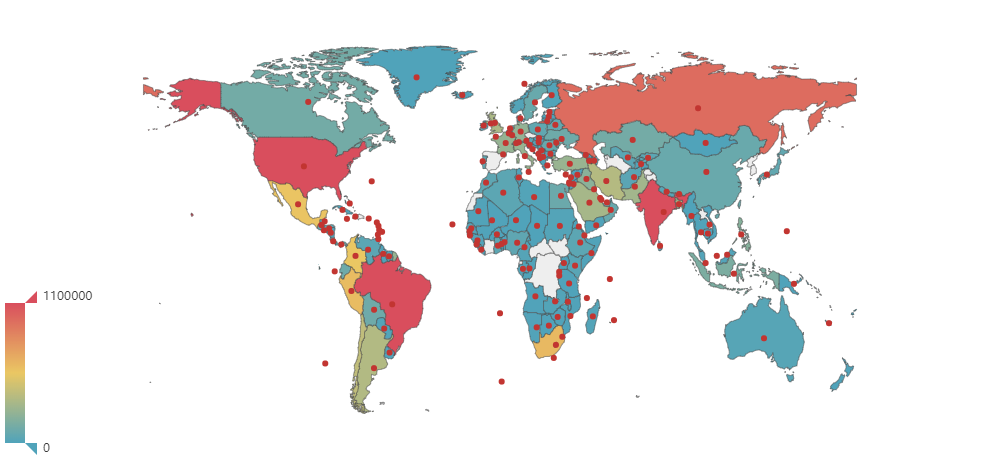
The results of data when plotted on the world’s map are shown as Fig.1, 2, 3 and 4. Different visuals of the same purpose were created for a better visualization.

The map is interactive and one can simple see the cases by hovering over to a specific location on the map or clicking on the dots shown in Fig.2.

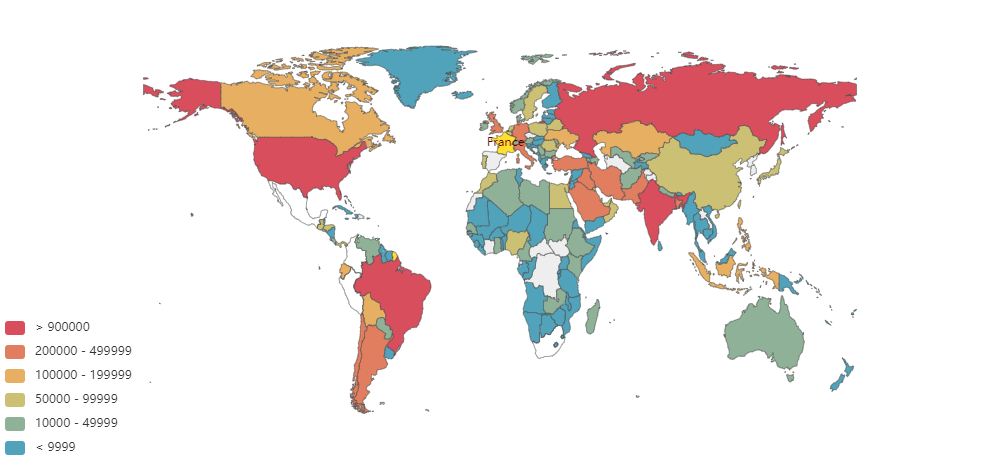
Fig.1 was messy as 212 countries were shown on the map. I got rid of the countries name and used a dot instead as shown in Fig. 2.



**Fig.1: Number of COVID-19 cases in 212 countries as of Aug 25, 2020.**



**Fig.2: Interactive map where one can find the total number of cases by a single click on the particular dot associated with each country.**



**Fig.3: Color coded countries list based on number of COVID-19 cases.**



**Fig. 4: Added title and subtitle on the map for more lucrative visualization of the cases.**

1. **Discussions**

Python provided a very simple and robust way to create COVID 19 spread visualization on world’s map. The cases in dataset were reported country wise and similar script can be modified to create maps for regional data. Other significant data like people effected with pre-existing conditions, total number of deaths, age group affected, change in number of cases after imposing lockdowns, mandatory masks etc. would have been more interesting to analyze but seems beyond the scope of the project.

Another important aspect could be to split the data in test and train sets and apply Machine learning algorithms to make predictions. This is a difficult problem as the number of assumptions could be very large resulting into a high degree of uncertainty in the results.

1. **Conclusions**

Data visualization of COVID-19 was successfully demonstrated using mapping the number of cases on a world’s map. Different ways of displaying the results were explored in order to construct a dashboard for end user for easy to work with interactive map. The data can be further explored using machine learning algorithms to predict the spread of the pandemic while keeping restrictions in place or without employing restrictions.