

## Mini Project

In this project we are going to work with COVID19 dataset, published by **John Hopkins University**, which consist of the data related to cumulative number of confirmed cases, per day, in each Country. Also we have another dataset consist of various life factors, scored by the people living in each country around the globe. **Your task is to merge these two datasets to see if there is any relationship between the spread of the virus in a country and how happy people are, living in that country.**

The Mini-Project on COVID19 Data Analysis using Python is divided into following Section:

### Section 1: COVID19 dataset & Libraries

*Hint:*

- Import all the relevant libraries and “Covid19\_Confirmed\_dataset.csv” dataset as a Data Frame.
- Pre-process the Data Frame and drop ‘Lat’ & ‘Long’ columns. Grouped the data with ‘Country/Region’ using sum as an aggregate function.

**Task:**

**Visualize the pre-processed data for India, China & US countries**

### Section 2: Finding a good Measure

*Hint:*

Find the first derivative for the data (let’s call it an ‘Infection Rate’), calculate the maximum infection rate for India, China & US. You may plot the Infection Rate for some countries also for a better understanding of data

**Task:**

**Calculate the maximum ‘Infection Rate’ for each country and store it in a new column named ‘max\_infection\_rate’**

**Create a New Data Frame name ‘Corona Data’ with ‘Country/Region’ as an index and ‘max\_infection\_rate’ as a column**

### Section 3: World happiness report dataset

*Hint:*

- Import the *Worldwide\_happiness\_report.csv* file as happiness\_report.

- Pre-process the DataFrame and drop 'Overall rank', 'Score', 'Generosity', 'Perceptions of corruption' columns. Set 'Country or region' column as Index

**Task:**

**Create a DataFrame named '*data*' by merging 'happiness\_report' with 'Corona Data' and find correlation among all variables**

**Section4: Visualize our results using Seaborn.**

*Hint:*

- Plotting GDP vs maximum Infection rate
- Plotting Social support vs maximum Infection rate
- Plotting Healthy life expectancy vs maximum Infection rate
- Plotting Freedom to make life choices vs maximum Infection rate

**Task:**

**Based on the plot above, comment on the Indicators having strong relationship with COVID19 Infection?**

**&&&&&**