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

Task:

Visualize the data for India, China & US countries

Hint:

- Import all the relevant libraries and "Covid19_Confirmed_dataset.csv" dataset as a Data Frame.
- Drop 'Lat' & 'Long' columns. Grouped the data with 'Country/Region 'using sum as an aggregate function.

Section 2: Finding a good Measure

Hint:

Find the first derivative for the data using diff() command (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.
- 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

Section 4: Visualization using Folium Map

Task:

Add the Latitude & Longitude information of countries in 'data' and Visualize it using Folium world map

Hint: You may use Latitude & Longitude information available at COVID19 data set.

Section5: Visualization of 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?

&&&&&