**Clustering countries based on Corona cases and analyse trending venues**

**Introduction**

The COVID-19 pandemic, also known as the coronavirus pandemic, is an ongoing pandemic of coronavirus disease 2019, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The outbreak was first identified in Wuhan, China, in December 2019. The World Health Organization declared the outbreak a Public Health Emergency of International Concern on 30 January, and a pandemic on 11 March.

The pandemic has impacted almost the entire world. There has been lockdowns implemented by different countries to contain the spread of the disease. It has caused global social and economic disruption, including the largest global recession since the Great Depression. It has led to the postponement or cancellation of sporting, religious, political, and cultural events, widespread supply shortages exacerbated by panic buying, and decreased emissions of pollutants and greenhouse gases. Schools, universities, and colleges have been closed either on a nationwide or local basis in 177 countries, affecting approximately 98.6 percent of the world's student population.

As of 31 May 2020, more than 6.07 million cases of COVID-19 have been reported in more than 188 countries and territories, resulting in more than 369,000 deaths; more than 2.57 million people have recovered.

The extend of spread of the pandemic has been different in different countries. For some countries, the spread has been exponential while some countries attributed a steady growth and there were others who did not see much spike in the cases.

This project aims to

1. Cluster different countries based on their daily corona cases

2. What are the trending venues in the capitals of most affected corona countries

**Data**

For this project, I will be using the data from below sources

1. WHO covid website

https://covid19.who.int/info

The site has date wise spread of corona cases for each country. The data will need to be cleansed and required features to be selected for doing the analysis. Then need to standardise the data and find out the optimum number of clusters using the elbow method. Apply the KMeans clustering method and label countries with similar features.

2. Fousquare API

I will also be using the Fousquare API to find the trending venues in some of the most affected corona country capitals.