

# Flipr Hackathon 6.0

## Machine Learning

---

# Covid-19 in India

The COVID-19 pandemic in India is part of the worldwide pandemic of coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first case of COVID-19 in India, which originated from China, was reported on 30 January 2020. India currently has the largest number of confirmed cases in Asia, and has the second-highest number of confirmed cases in the world after the United States, with the number of total confirmed cases breaching the 100,000 mark on 19 May, and 1,000,000 confirmed cases on 17 July 2020. On 29 August 2020, India recorded the global highest spike in COVID-19 cases on a day with 78,761 cases surpassing the previous global highest daily spike of 77,368 cases which was recorded in the US on 17 July 2020.

India's case fatality rate is among the lowest in the world at 2.41% as of 23 July and is steadily declining. By mid-May 2020, six cities accounted for around half of all reported cases in the country – Mumbai, Delhi, Ahmedabad, Chennai, Pune and Kolkata. As of 24 May 2020, Lakshadweep is the only region which has not reported a case. On 10 June, India's recoveries exceeded active cases for the first time.

On 22 March, India observed a 14-hour voluntary public curfew at the instance of Prime Minister Narendra Modi. It was followed by mandatory lockdowns in COVID-19 hotspots and all major cities. Further, on 24 March, the prime minister ordered a nationwide lockdown for 21 days, affecting the entire 1.3 billion-person population of India. On 14 April, India extended the nationwide lockdown till 3 May which was followed by two-week extensions starting 3 and 17 May with substantial relaxations. From 1 June, the government started "unlocking" the country (barring "containment zones") in three unlock phases.

---

---

# Problem Statement

The spread of COVID-19 in the whole world has put the humanity at risk. The resources of some of the largest economies are stressed out due to the large infectivity and transmissibility of this disease. Due to the growing magnitude of number of cases and its subsequent stress on the administration and health professionals, some prediction methods would be required to predict the number of cases in future.

When will the coronavirus pandemic come to an end? The question is on everyone's mind, and while astrologers and politicians have answers, few scientists want to be drawn into hazarding a prediction.

The challenge of the Covid-19 prediction is the most crucial component for countries and global health institutions. A successful and accurate prediction to the future covid cases ultimately results in better management of the pandemic.

## **Part -01:**

The objective of the first part of the problem statement is to predict the Covid Cases of a City on 1<sup>st</sup> September 2020. The output file 01 should contain only City and the respective Covid Cases for the test data.

## **Part -02:**

The Foreign Visitors of a city is a time-dependent parameter, for which you have to come up with a Time-series prediction model. Using the Foreign Visitors predicted by the model, you need to calculate the Covid Cases on 1<sup>st</sup> Oct 2020 for every City in the test data. . The output file 02 should contain only City and the respective Covid Cases on 1<sup>st</sup> October.

---

---

There are 3 files provided:

**1. Variable\_Description.xlsx:**

This file contains description of all the variables available in the dataset

**2. Training\_data.xlsx:**

This is the training dataset on which model has to be trained, which contains parameters of a city on 1<sup>st</sup> September 2020

**3. Test\_data.xlsx:**

This is the test data on which accuracy of the model will be computed. It also contains Time Series data of Foreign Visitors to be used for Part – 02

## **Competition Rules**

- There should only be **one submission per participant/team**
  - Privately sharing of code is not permitted. In case of plagiarism, the participant shall be disqualified
  - Those attempting both the parts should send 2 separate .csv/.xlsx file, containing **City** and **Covid Cases** on 1st September and 1st October respectively
  - The **solution\_sheet** should also be attached along with the results
  - Share all your files in this Google form link: <https://forms.gle/5CHsgn8cbpwEDsSG7>
-