Project Proposal: COVID'19 Data Analysis

Shantanu Tyagi (201801015)[*] and Arkaprabha Banerjee (201801408)[†]
Dhirubhai Ambani Institute of Information & Communication Technology,
Gandhinagar, Gujarat 382007, India
CS-306 Data Analysis and Visualization

Ever since the onset of the COVID-19 pandemic, researchers have been trying to model this epidemic in terms of its spread, intervention measures and various other social factors which are often time dependent. As a part of this project, we aim to analyze the trend of this pandemic primarily in terms of the infected populations for various region and try to explain it in accordance with various social factors.

We shall collect the case data ever since the inception of the pandemic. This shall enable us to analyze the case related data with respect to each region over the time span. Post that, we shall try to explain the trend with respect to pre-existing epidemic models (SIR model, SEIR model) and the effect of various interventions such as lockdown and vaccination which can be done using differential equations.

To compare the theoretical models, we shall try to fit a curve to our existing data and compare it to a generic graph of these models. On account of the extremely high complexity for getting the parameters of the exact theoretical curve , we shall engage in a qualitative analysis to explain the trend. To analyze interventions, we shall consider various government measures and the response of the generic public. The response shall be measured by the Google Searches on this topic which implicitly correlates to the awareness among the generic human population. We shall correlate this behavior to the change in the infected population.

We shall obtain our data from the following sources:
1. COVID 19 Case data: Coronavirus Source Data - Our World in Data
2. Government Measures: Blavatnik School of Government
3. Search behavior — Google pytrend API