First there are 2 types of under-reporting that we need to take for consideration:

+ Death tolls under reporting – The number of fatality case in some countries.

+ Testing under reporting

Approaches: Influenza or Pneumonia

* Estimate the novelty of corona outbreak based on the comparison between the number of Severe Acute Respiratory Infection in 2020 and that of the period from 2016-2019 (Brazilian)
* It means that they assume people contracted with normal respiratory conditions as covid infected potential. They are detecting the anomolies in the novelty dataset to identify the possibility of under-reporting coronavirus data.
* We can apply this method for not only SARI but also for Pneumonia & Influenza. By having more data from other diseases, it helps to increase the confidence of our model
* According to WHO, from week 3 to week 36, there have been 194 cases suspected to be contracted with influenza, however, all patients are proven with negative testing results.
* Extrapolate expected deaths and hospitalizations with respect to observations in countries that passed the exponential growth curve
  + This is an unattainable approach since facts indicate that most death cases are correlated with pre-existed medical syndromes. Therefore, it will be even more difficult to cluster the ones died with covid infection
* Some use integer-valued autogressive hidden Markov model to identify the hidden process of recording covid cases in the world, then they applied for Viterbi estimation to predict the exact infection numbers in a specific location . For the moment I am trying to understand the math behind this algorithm to make it applicable for our research purpose. I am requesting the original research in 2014 applying this methodology and will look further into that.
* We have the spread rate of some nations, taking that as a model to compare with past viral pandemic to see our confidence
* Another popular methodology I found is that some applied susceptible-exposed-infected-recovered model with assuming the incubation period of 6 days to describe the dynamics of disease spread
  + However, their model did not take into account of covid importation which was our main target for this scientific research.
  + We assume that the time-rate of change of  **S(t)**,  the *number* of susceptibles,1 depends on the number already susceptible, the number of individuals already infected, and the amount of contact between susceptibles and infecteds. In particular, suppose that each infected individual has a fixed number  **b**  of contacts per day that are sufficient to spread the disease. Not all these contacts are with susceptible individuals. If we assume a homogeneous mixing of the population, the *fraction* of these contacts that are with susceptibles is  **s(t)**.  Thus, on average, each infected individual generates  **b s(t)**  new infected individuals per day. [With a large susceptible population and a relatively small infected population, we can ignore tricky counting situations such as a single susceptible encountering more than one infected in a given day.]

Limitations:

* Most of official medical datasets from government’s health webpage are outdated (latest data I can find was in 2017 in New Zealand)
* Data in Taiwan is integrated with Chinese intelligence records, CCP hardly provide the authenticated data for the public audience

ConfirmedGlobal = read.csv(url("https://data.humdata.org/hxlproxy/api/data-preview.csv?url=https%3A%2F%2Fraw.githubusercontent.com%2FCSSEGISandData%2FCOVID-19%2Fmaster%2Fcsse\_covid\_19\_data%2Fcsse\_covid\_19\_time\_series%2Ftime\_series\_covid19\_confirmed\_global.csv&filename=time\_series\_covid19\_confirmed\_global.csv"));

DiedGlobal = read.csv(url("https://data.humdata.org/hxlproxy/api/data-preview.csv?url=https%3A%2F%2Fraw.githubusercontent.com%2FCSSEGISandData%2FCOVID-19%2Fmaster%2Fcsse\_covid\_19\_data%2Fcsse\_covid\_19\_time\_series%2Ftime\_series\_covid19\_deaths\_global.csv&filename=time\_series\_covid19\_deaths\_global.csv"));

RecoveredGlobal = read.csv(url("https://data.humdata.org/hxlproxy/api/data-preview.csv?url=https%3A%2F%2Fraw.githubusercontent.com%2FCSSEGISandData%2FCOVID-19%2Fmaster%2Fcsse\_covid\_19\_data%2Fcsse\_covid\_19\_time\_series%2Ftime\_series\_covid19\_recovered\_global.csv&filename=time\_series\_covid19\_recovered\_global.csv"));