

# SARS-CoV-2

---

## An Analysis of Spread of Infection

Divyansh Chahar



 <https://www.linkedin.com/in/divyanshchahar/>

 <https://github.com/divyanshchahar>

Friday 12<sup>th</sup> June, 2020

# Nomenclature

## Greek Characters

$\bar{I}$	Average Infected
$A_n$	Moving Average
$G_t$	Growth Rate at time $t$
$I_t$	Individual Infected at time $t$
$N$	Total Population
$R$	Reproduction Number
$R_t$	Individual Removed at time $t$
$S_t$	Individual Susceptible at time $t$
$t$	Time in days
$x$	Parameter Under Consideration

## Roman Characters

$\beta$	Disease Transmission Rate Constant
$\gamma$	Recovery Rate

# 1 Introduction

SARS-CoV-2 or Severe Acute Respiratory Syndrome related Coronavir-2, is a virus that causes Covid-19, which is commonly known as Coronavir. The virus originated in Wuhan, capital of Hubei Province of People's Republic of China, Covid-19 was classified as *Pandemic* by World Health Organization on *11-March-2020*. At the time of writing this report Coronavirus has infected 7360239 and killed 416201 individuals across 188 countries.

## What is a Pandemic ?

Pandemic or Pandemic Influenza is described as follows. As per [1], pandemic is the worldwide spread of a new disease. An influenza pandemic occurs when a new Influenza virus emerge and spreads around world and most people don't have immunity against it. However this could lead to misconceptions, because every season the cases of seasonal flu rises in temperate southern and northern hemispheres, in a globally connected world, it also crosses international boundaries, but is not classified as Pandemic. As mentioned in [2], during the 2009-2010 H1N1 influenza pandemic spread, influenza spread early in the temperate southern hemisphere but out of season in the northern hemisphere. This out of season spread was the reason that H1N1 was classified as influenza pandemic.

Thus based on [1] and [2] we can say that an influenza pandemic is an unexpected worldwide transmission of a new disease to which the population has no immunity.

## Modelling an infectious disease - SIR model

Variety of natural and physical phenomenon can be simulated by using numerical models as described in [3]. In epidemiology one such model used to simulate the spread of infectious disease is the *S.I.R.* model.

The *SIR* model was proposed by Kermack and McKendrick [4] in 1927. This model classifies the population to be evaluated into three categories:

- Susceptible  $\rightarrow S$
- Infected  $\rightarrow I$
- Removed  $\rightarrow R$

Susceptible population comprises of individuals who have not contracted the infection yet but can be infected. Infected population represents individuals who are infected by the infection and can transmit it to other susceptible individuals. Removed population represents individuals who have either recovered (hence they cannot contract and transmit the virus again) from the infection or have died from the infection.

Thus we can say that,

$$N_t = S_t + I_t + R_t$$

The most important parameter that comes from this model is basic reproduction number, denoted by  $R$ .

$$R = \frac{\beta S_0}{\gamma}$$

If  $R > 1$ , then it is an indication that every primary case is producing more than one secondary case, this situation indicates that infection will spread if checks are not put in place. If  $R < 1$  this indicates that every primary case is resulting in lesser number of secondary cases, thus the infection will eventually die, even if no corrective measures are introduced.

Even though  $R_0$  can provide insights into whether the infection is spreading or not, we don't have the required data to calculate this parameter. Thus we will use some other parameters to evaluate the situation. These parameters will be discussed in more detail in the next section.

## 2 Parameters used for analysis

This report will only focus on analyzing the spread of infection. The data used in this report is a time series data taken from John Hopkins University's Coronavirus Database (last updated 10-June-2020)

### Average Cases

This parameter helps us normalize the total number of cases with the duration of infection (Number of days to reach  $I$  cases)

$$\bar{I} = \frac{\sum I}{t}$$

Since the concerned data is a time series data, hence summation in this case represents the last column of the database.

### Moving Average

The concept of moving average is used in slightly different manner here.

$$A_n = \frac{\sum_{i=n}^0 x}{t}$$

### Growth Rate

The number of confirmed cases could be confusing when analyzing the situation as the total number of cases will only have an upward trend. Thus to overcome this difficulty we will use the concept of growth rate. Growth Rate can be defined as the ratio of new cases on the present day to the ratio of new cases the previous day.

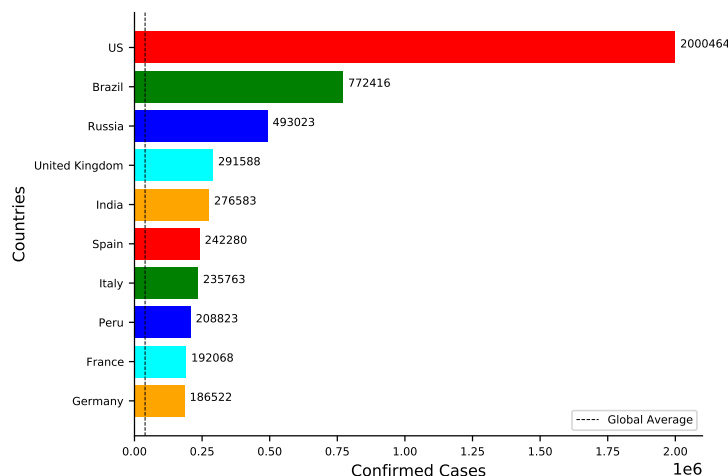
It can be mathematically expressed as

$$G_t = \frac{I_t - I_{t-1}}{I_{t-1} - I_{t-2}}$$

If the number of new cases on given day are less than the number of cases on the previous  $G_t < 1$ . This would mean each primary case of infection is infecting less than one individual. However if  $G_t > 1$  it would mean that each primary case is producing more than 1 secondary case. This is not a desired situation.

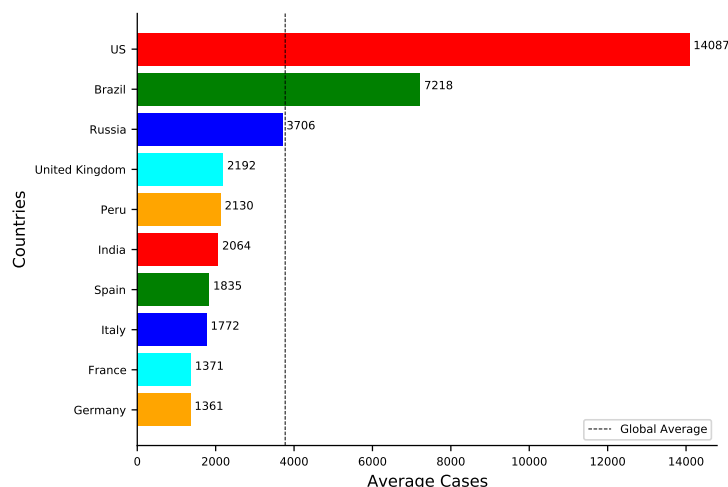
### 3 Comparative Analysis

The database used in this report consists of 188 countries but we will limit the scope of our analysis to the 10 countries with maximum number of confirmed cases.



**Figure 3-1: Number of Confirmed Cases**

As can be seen in figure 3-1 the number of confirmed cases in these ten countries exceed the global average. Even among the worst hit countries the number of cases in USA is exceptionally high. Number of confirmed cases is USA, which is at first place is 10.752 times higher than Germany which is at the tenth place. There also exists a huge difference between USA and Brazil, which is at second place. USA has 2.589 times more cases as that of Brazil. This goes on to reflect that the spread of Coronavirus in USA has a completely different magnitude than the rest of the world.

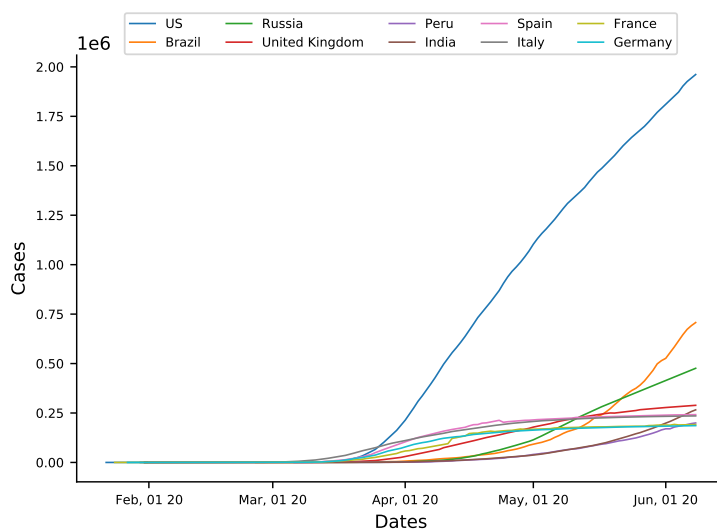


**Figure 3-2: Number of Confirmed Cases**

As mentioned in the previous section, that number of confirmed cases could be an elusive parameter. A country having higher number of confirmed cases might come across as having a more dire situation than a country which has lower number of confirmed cases. However, since infectious diseases follows an exponential curve hence the number of confirmed cases climb very quickly to higher numbers, hence normalizing

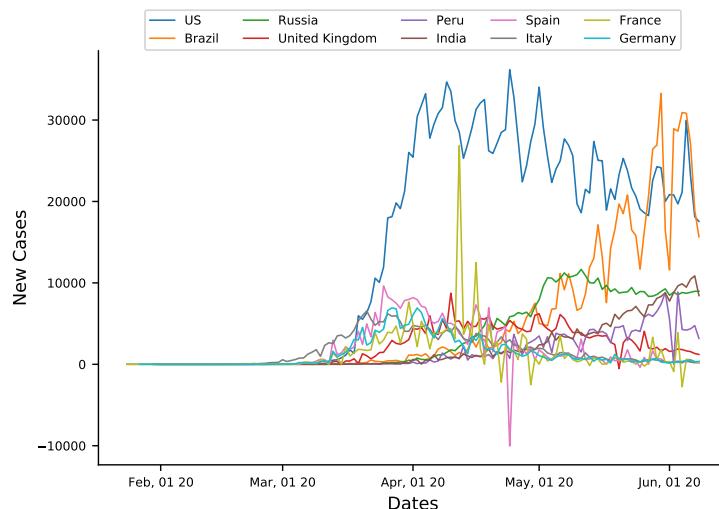
the total number of cases with the duration can give an indirect insight into the growth rate of confirmed cases.

As can be observed from figure 3-2, USA has the highest average number of cases among the worst hit countries. USA and Brazil are the only countries where average number of confirmed cases is higher than the global average. If we compare figure 3-1 and figure 3-2, we can observe even though India has higher number of confirmed cases, average cases in India are still lower than Peru, this indicates that Peru has a more aggressive growth rate than India. This analogy can be applied to other countries as well. In Europe the first major outbreak was in Italy, which has seventh highest number of confirmed cases, ranks eight in terms of average cases. France and Germany are the only countries which has the ninth and tenth rank in terms of confirmed cases and also have the same standing in terms of average number of cases.



**Figure 3-3: Number of Confirmed Cases over time**

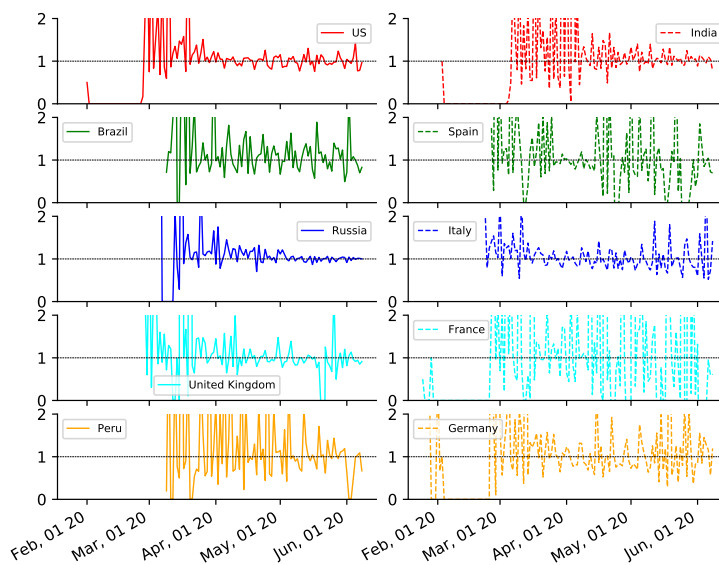
In order to paint a more clear picture of the spread of infection, rise in the number of confirmed cases needs to be studied. From figure 3-3 we can observe that USA and Italy had same number of cases between March and April, but the spread of Coronavirus is much more aggressive in USA than in Italy. Brazil and Russia had the same number of cases between mid April and May, But Russia had a more aggressive growth rate at this point of time hence as a result it developed more cases than Brazil. However Brazil ended up having more cases than Russia as between May and June, the spread of infection became more aggressive in Brazil than in Russia. A sudden spike in the number of cases can be observed for France between April and May. A similar pattern can be observed for Spain, where a small sharp decline can be observed for Spain, between April and May, this could be attributed to error in reporting or testing. India had the same amount of cases as Peru between May and June, but near the end of June India overtook France, Germany, Italy and Spain.



**Figure 3-4: Number of New Cases over time**

As it has been mentioned before that total number of cases do not give a complete perspective of the situation. Hence, proceeding further, number of new cases will be studied to better understand the rate of spread of Coronavirus. As can be seen from figure 3-4, occurrence of new cases in USA spiked after mid-March, a steep rise in new cases can be seen from mid of March till starting of April. This trend discontinued after April 1 and the pattern became more irregular. After April 1, several crests and troughs can be observed, even though the pattern seems to be irregular, but on a closer observation it can be seen that after March 15, the crests peak out to lower values and troughs fall to lower values, indicating a possible decline in the number of new cases. This could be an indicator that the measures taken to stop the spread of the Coronavirus are working.

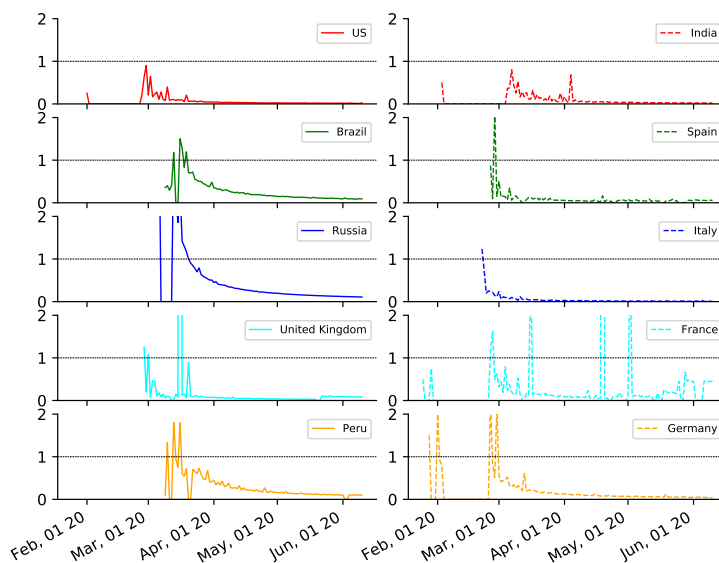
Except for Spain, UK, Italy and Germany, no other country reported less than 10000 cases in a single day. France has reported abrupt increase in the number of new cases which can be observed in the form of huge spikes. All the countries except USA, Brazil, India and Peru show a similar pattern in which the occurrence of new cases increases before decreasing. This is indicative of the fact that the measures taken to curb the spread of Coronavirus are working. We can also observe a peculiar feature in figure 3-3, Spain seems to have negative number of new cases this is not physically possible. As mentioned above, this might be due to erroneous reporting or erroneous testing.



**Figure 3-5: Growth Rate over time**

As was mentioned in the previous section of this report, growth rate could be the best indicator of the situation. The plot originally generated had huge spikes which made it difficult to focus on the important range (0 to 1) of the growth rate. Hence the maximum values of Growth rate was capped to 2 and data for each country was plotted separately.

As we can observe from figure 3-5, none of the country has managed to lower the Growth Rate below the critical value of 1. Oscillations close to the critical value can be observed for USA and Russia. Even though USA had the maximum number of confirmed cases, the growth rate observed after April 1 has dramatically decreased and oscillates very close to 1. Similar behavior can be seen with Russia after May 1. It can be observed that after May 1, variation in Growth Rate around the critical value is pronounced in USA than in Russia. Both India has also managed to keep the Growth Rate close to critical value after May 1.



**Figure 3-6: Moving Average of Growth Rate over time**



As can be seen from figure 3-6, moving averages have eliminated irregularities when applied to growth rate. USA and India are the only countries which has managed to keep the moving average below the critical value of 1. Although moving average eventually falls below 1 for all the countries, it can be observed that all the countries except India and USA have spikes in the moving average of Growth Rate which cross the critical value of 1.

## 4 Conclusion

Based on the observations made in previous section, following conclusions can be drawn:

- Coronavirus is highly contagious and the number of new cases can increase extremely rapidly as seen in the case of USA.
- Number of occurrence of new cases is a very strong indicator of evaluating the situation within a country
- Tracking the number of new cases could give an insight of the situation, but growth rate could be a better indicator that provides more conclusive analysis
- A drop in the number of new cases could be observed for some countries, this indicates that preventative measures such as lock-downs are actually effective in preventing the spread of Coronavirus

# Bibliography

- [1] “What is a pandemic?.” [https://www.who.int/csr/disease/swineflu/frequently\\_asked\\_questions/pandemic/en/](https://www.who.int/csr/disease/swineflu/frequently_asked_questions/pandemic/en/), Feb 2010. Accessed: 20 May 2020.
- [2] H. Kelly, “The classical definition of a pandemic is not elusive,” *Bulletin of the World Health Organization*, vol. 89, pp. 540–541, 2011. Accessed: 20 May 2020.
- [3] A. Scharnhorst, K. Börner, and P. van den Besselaar, *Models of science dynamics: Encounters between complexity theory and information sciences*. Springer Science & Business Media, 2012.
- [4] W. O. Kermack and A. G. McKendrick, “A contribution to the mathematical theory of epidemics,” *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, vol. 115, no. 772, pp. 700–721, 1927.