Coronavirus Time Series Analysis

# Introduction

The current date of this analysis is March 17, 2020.

All code and data can be found on [Github](https://github.com/howardnewyork/corona).

# Summary of Results

This report analyzes the growth rate in Covid-19 confirmed cases and deaths by country. In particular, I have reviewed the “aligned” growth in confirmed cases and deaths. That is, for each country I aligned the data to start:

* For confirmed case analysis: The day when 200 cases were confirmed in each respective country
* For deaths analysis: The day when 10 deaths were confirmed in each respective country

My view is that in an exponential growth scenario (which we are in), it is the growth rate that is all important rather than the nominal number of confirmed cases or deaths. This is what the analysis focuses on. A country with low infected numbers but a high growth rate will quickly overtake a country with high infections but low growth rate. For example, compare Italian vs. Chinese deaths below.

The data has some good news. China and South Korea have managed to remarkably bring down the growth of new infections and deaths. Other countries do not have such good news and the confirmed cases and deaths are doubling every four or five days. Although Italy is experiencing a total disaster, it appears as if the severe exponential growth it was experiencing is slowing a little. The growth rates in Italy are still very high but trending in the right direction. Let’s see if this trend will continue.

For US and other European countries, the news is uniformally bad. These countries seem to be following more the Italian model rather than the South Korean one.

I note that the Chinese data is so good that it seems suspect. They had only 25 new infections and 13 deaths on March 19. Given their early problems and wide exposure, the new data seem too good to be true.

In summary, there is an inkling of good news in the Italian data and hopefully the trend continues. South Korea and China (if the data are correct) show that the virus can be beat, BUT at this stage radical control of population movement is likely necessary (per Italian and Chinese model).

# Methodology

I used a Gaussian Process model to establish the mean growth rates. Raw growth rates were determined as:

* Confirmed case growth rate:
* Confirmed death growth rate:

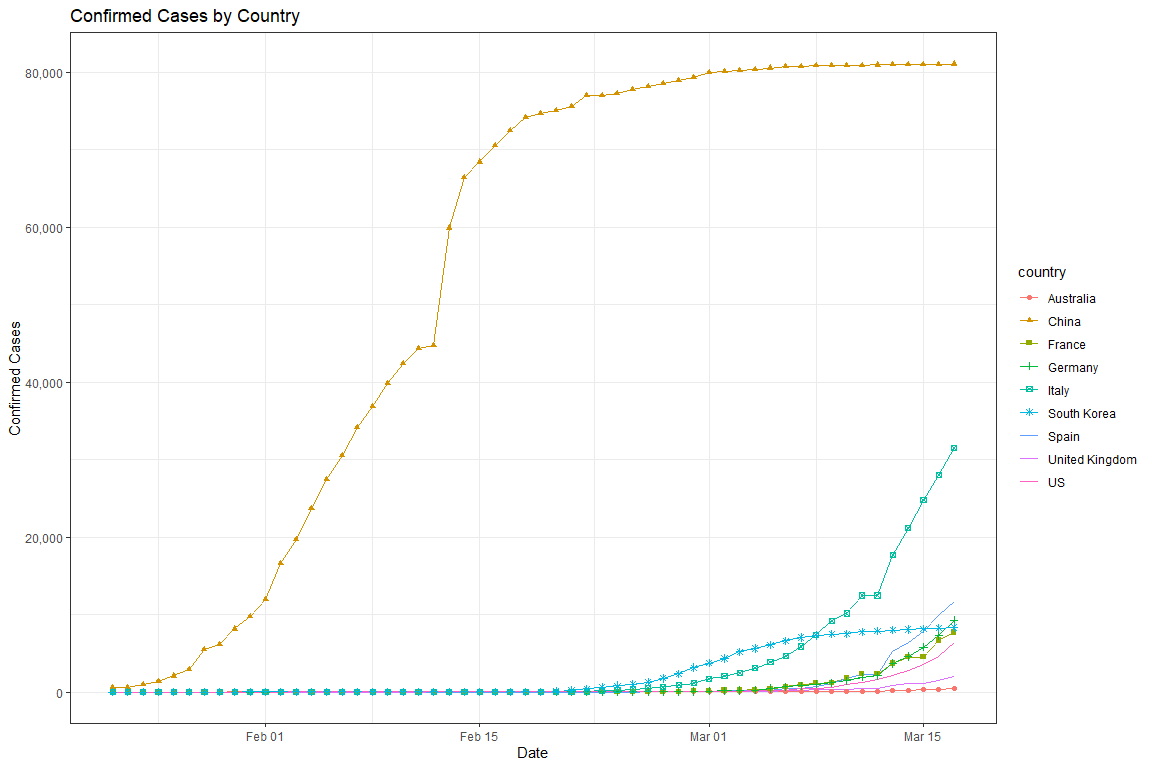
I then calculate a 95% credible around such mean estimate.

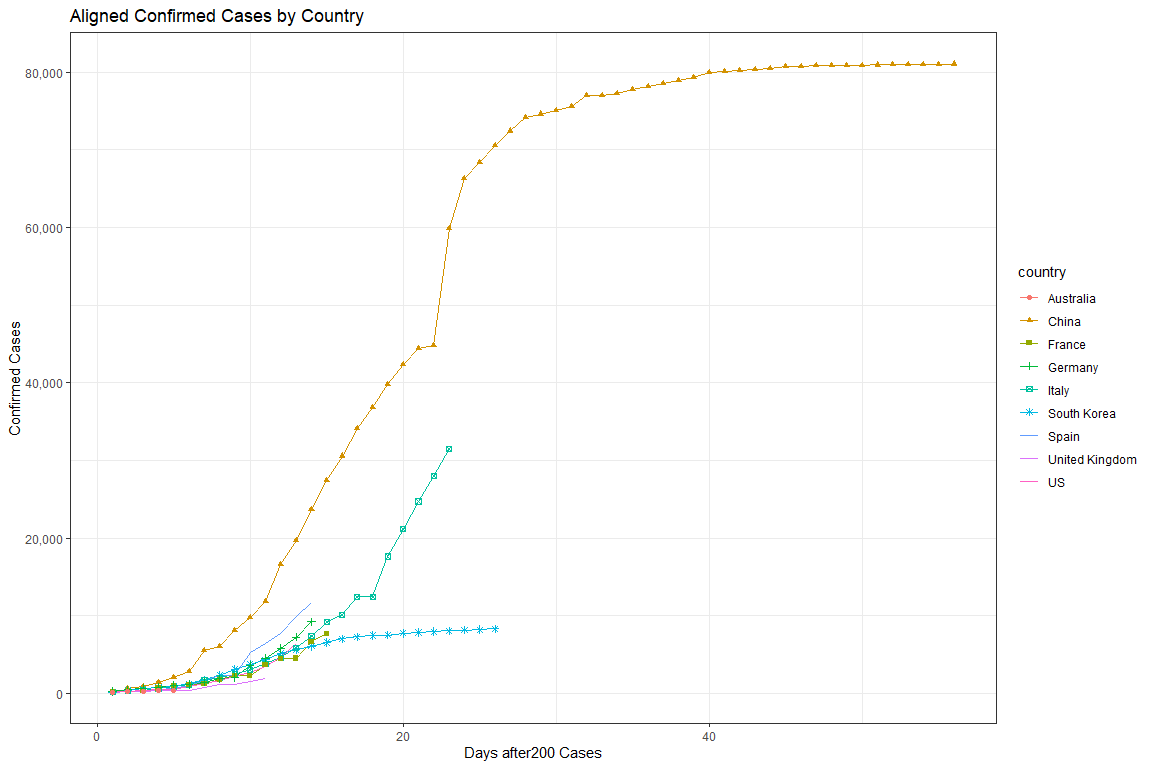
# Data Sources

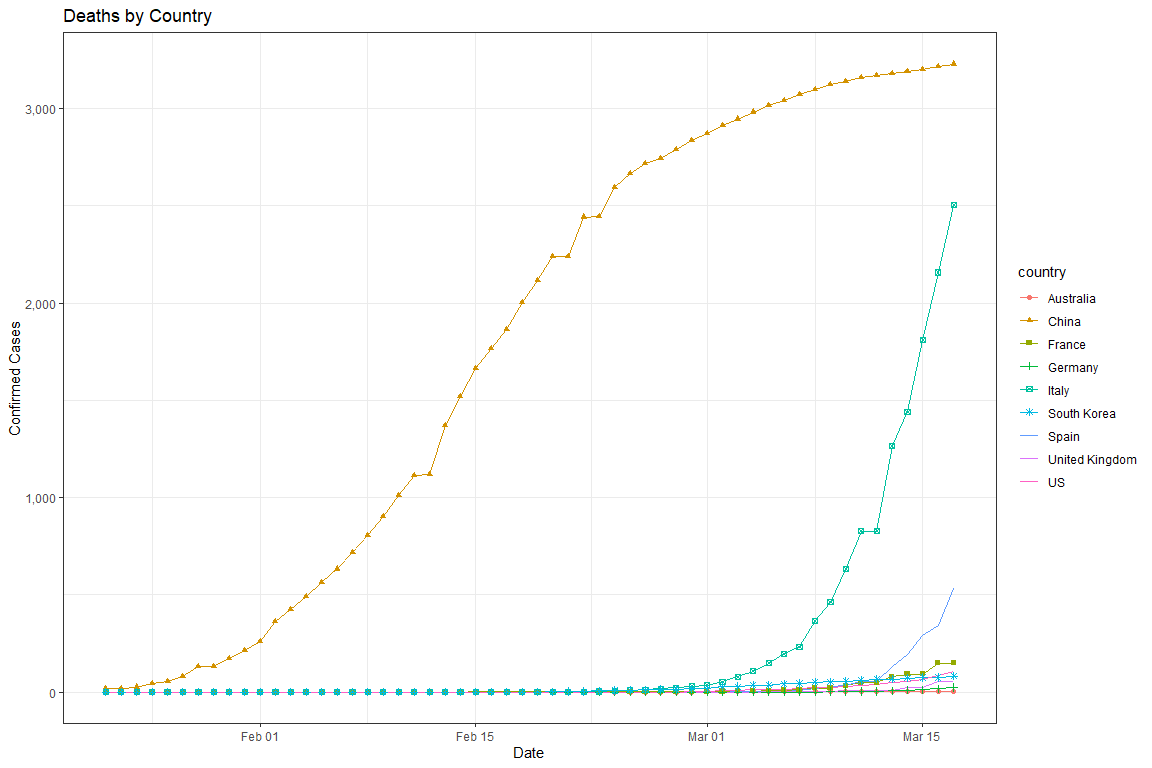
[Johns Hopkins COID-19 Data Repository](https://github.com/CSSEGISandData/COVID-19/compare?expand=1)

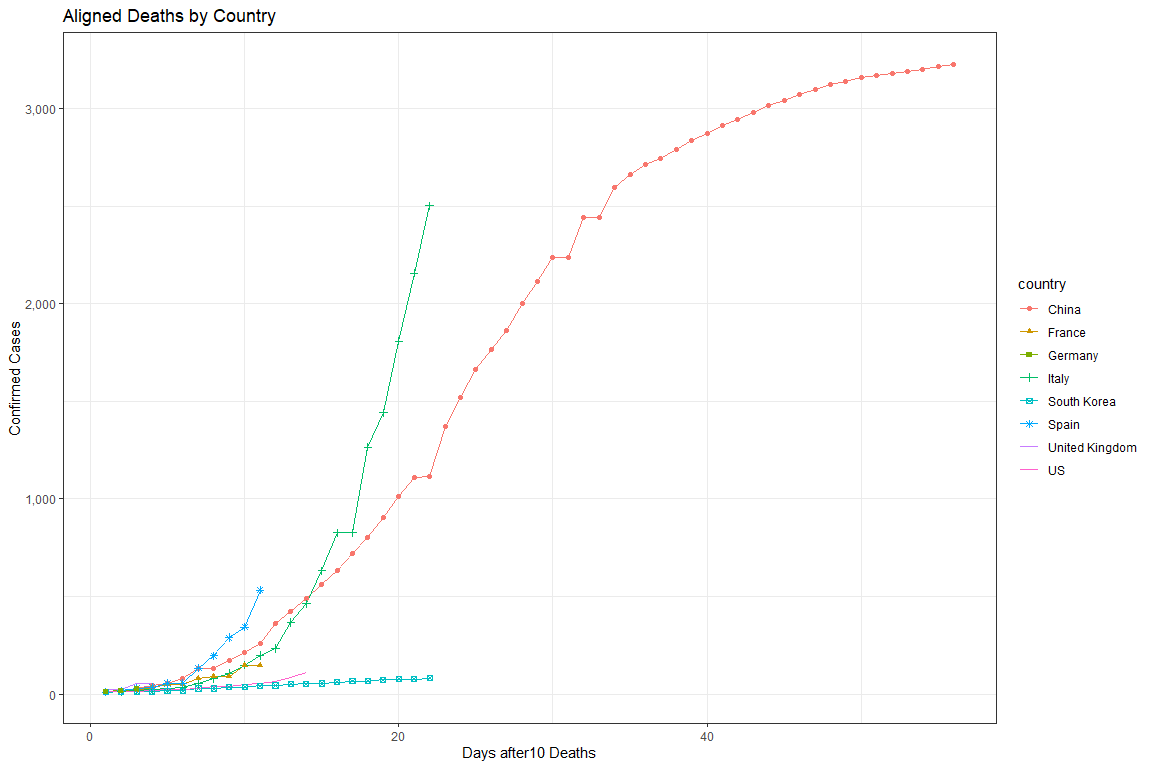
Note: the data are very messy and dependent on differing testing regimes rather than true infection experience.

# Data Exploration



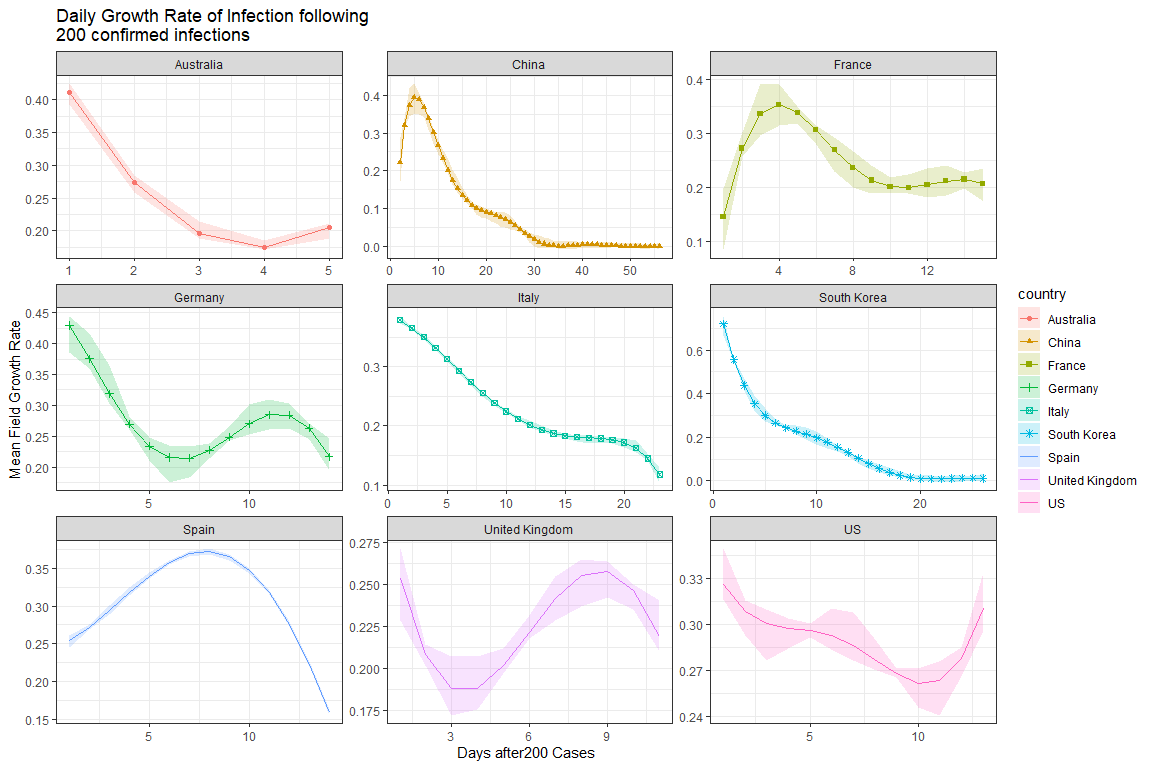




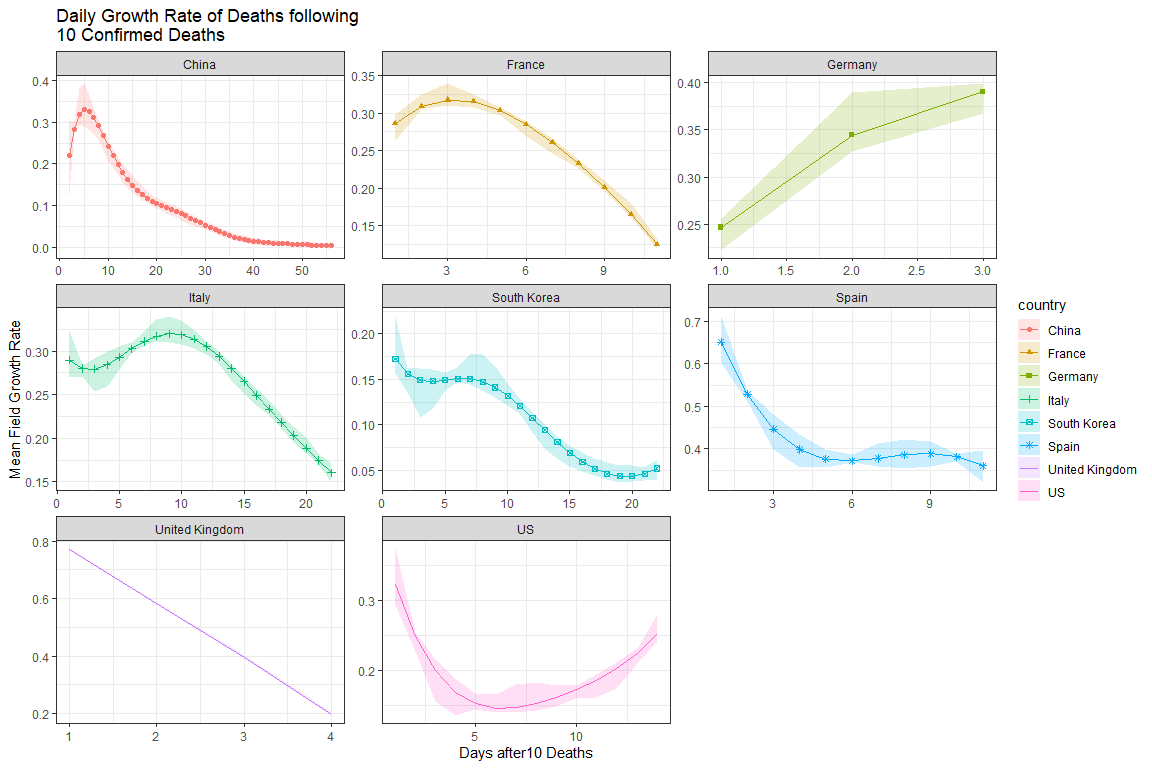


# Growth Rate Analysis

## Confirmed Cases



## Deaths



Note: UK death records are too limited to be reliable.